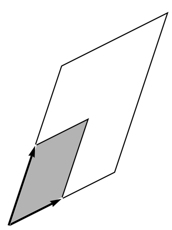
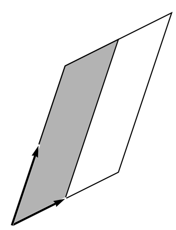
Answers to Exercises from

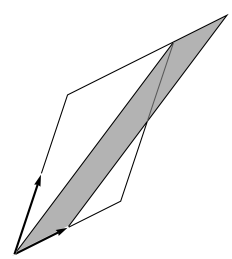
Linear Algebra



$$\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}$$



$$\begin{vmatrix} x\cdot 1 & 2 \\ x\cdot 3 & 1 \end{vmatrix}$$



$$\begin{vmatrix} 6 & 2 \\ 8 & 1 \end{vmatrix}$$

Jim Hefferon

Notation

$\mathbb{R}$	real numbers
$\mathbb{N}$	natural numbers: $\{0, 1, 2, \ldots\}$
$\mathbb{C}$	complex numbers
$\{\ldots \mid \ldots\}$	set of ... such that ...
$\langle \ldots \rangle$	sequence; like a set but order matters
V, W, U	vector spaces
$\vec{v}, \vec{w}$	vectors
$\vec{0}$, $\vec{0}_V$	zero vector, zero vector of V
B, D	bases
$\mathcal{E}_n = \langle \vec{e}_1, \ldots, \vec{e}_n \rangle$	standard basis for $\mathbb{R}^n$
$\vec{\beta}, \vec{\delta}$	basis vectors
$\operatorname{Rep}_B(\vec{v})$	matrix representing the vector
$\mathcal{P}_n$	set of n-th degree polynomials
$\mathcal{M}_{n\times m}$	set of $n\times m$ matrices
$[S]$	span of the set S
$M \oplus N$	direct sum of subspaces
$V \cong W$	isomorphic spaces
h, g	homomorphisms
H, G	matrices
t, s	transformations; maps from a space to itself
T, S	square matrices
$\operatorname{Rep}_{B,D}(h)$	matrix representing the map h
$h_{i,j}$	matrix entry from row i, column j
$\lvert T \rvert$	determinant of the matrix T
$\mathscr{R}(h), \mathscr{N}(h)$	rangespace and nullspace of the map h
$\mathscr{R}_\infty(h), \mathscr{N}_\infty(h)$	generalized rangespace and nullspace

Lower case Greek alphabet

name	symbol	name	symbol	name	symbol
alpha	α	iota	ι	rho	ρ
beta	β	kappa	κ	sigma	σ
gamma	γ	lambda	λ	tau	τ
delta	δ	mu	μ	upsilon	υ
epsilon	ϵ	nu	ν	phi	ϕ
zeta	ζ	xi	ξ	chi	χ
eta	η	omicron	o	psi	ψ
theta	θ	pi	π	omega	ω

Cover. This is Cramer's Rule applied to the system $x + 2y = 6$, $3x + y = 8$. The area of the first box is the determinant shown. The area of the second box is x times that, and equals the area of the final box. Hence, x is the final determinant divided by the first determinant.

These are answers to the exercises in Linear Algebra *by J. Hefferon. Corrections or comments are very welcome, email to jimjoshua.smcvt.edu*

An answer labeled here as, for instance, 1.II.3.4, matches the question numbered 4 from the first chapter, second section, and third subsection. The Topics are numbered separately.

Chapter 1. Linear Systems

. . . .

Answers for subsection 1.I.1

1.I.1.22 This system with more unknowns than equations

$$\begin{aligned} x+y+z&=0\\ x+y+z&=1 \end{aligned}$$

has no solution.

1.I.1.23 Yes. For example, the fact that the same reaction can be performed in two different flasks shows that twice any solution is another, different, solution (if a physical reaction occurs then there must be at least one nonzero solution).

1.I.1.25

(a) Yes, by inspection the given equation results from $-\rho_1+\rho_2$.

(b) No. The given equation is satisfied by the pair $(1,1)$. However, that pair does not satisfy the first equation in the system.

(c) Yes. To see if the given row is $c_1\rho_1+c_2\rho_2$, solve the system of equations relating the coefficients of x, y, z, and the constants:

$$\begin{aligned} 2c_1+6c_2&=6\\ c_1-3c_2&=-9\\ -c_1+c_2&=5\\ 4c_1+5c_2&=-2 \end{aligned}$$

and get $c_1=-3$ and $c_2=2$, so the given row is $-3\rho_1+2\rho_2$.

1.I.1.26 If $a\neq 0$ then the solution set of the first equation is $\{(x,y)\mid x=(c-by)/a\}$. Taking $y=0$ gives the solution $(c/a,0)$, and since the second equation is supposed to have the same solution set, substituting into it gives that $a(c/a)+d\cdot 0=e$, so $c=e$. Then taking $y=1$ in $x=(c-by)/a$ gives that $a((c-b)/a)+d\cdot 1=e$, which gives that $b=d$. Hence they are the same equation.

When $a=0$ the equations can be different and still have the same solution set: e.g., $0x+3y=6$ and $0x+6y=12$.

1.I.1.29 For the reduction operation of multiplying ρ_i by a nonzero real number k, we have that $(s_1,\ldots,s_n)$ satisfies this system

$$\begin{aligned} a_{1,1}x_1+a_{1,2}x_2+\cdots+a_{1,n}x_n&=d_1\\ &\vdots\\ ka_{i,1}x_1+ka_{i,2}x_2+\cdots+ka_{i,n}x_n&=kd_i\\ &\vdots\\ a_{m,1}x_1+a_{m,2}x_2+\cdots+a_{m,n}x_n&=d_m \end{aligned}$$

if and only if

$$\begin{aligned} a_{1,1}s_1+a_{1,2}s_2+\cdots+a_{1,n}s_n&=d_1\\ &\vdots\\ \text{and } ka_{i,1}s_1+ka_{i,2}s_2+\cdots+ka_{i,n}s_n&=kd_i\\ &\vdots\\ \text{and } a_{m,1}s_1+a_{m,2}s_2+\cdots+a_{m,n}s_n&=d_m \end{aligned}$$

by the definition of 'satisfies'. But, because $k \neq 0$, that's true if and only if

$$\begin{aligned} a_{1,1}s_1 + a_{1,2}s_2 + \cdots + a_{1,n}s_n &= d_1 \\ &\vdots \\ \text{and } a_{i,1}s_1 + a_{i,2}s_2 + \cdots + a_{i,n}s_n &= d_i \\ &\vdots \\ \text{and } a_{m,1}s_1 + a_{m,2}s_2 + \cdots + a_{m,n}s_n &= d_m \end{aligned}$$

(this is straightforward cancelling on both sides of the i-th equation), which says that $(s_1, \ldots, s_n)$ solves

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= d_1 \\ &\vdots \\ a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n &= d_i \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= d_m \end{aligned}$$

as required.

For the pivot operation $k\rho_i + \rho_j$, we have that $(s_1, \ldots, s_n)$ satisfies

$$\begin{aligned} a_{1,1}x_1 + \cdots + \qquad a_{1,n}x_n &= \qquad d_1 \\ &\vdots \\ a_{i,1}x_1 + \cdots + \qquad a_{i,n}x_n &= \qquad d_i \\ &\vdots \\ (ka_{i,1} + a_{j,1})x_1 + \cdots + (ka_{i,n} + a_{j,n})x_n &= kd_i + d_j \\ &\vdots \\ a_{m,1}x_1 + \cdots + \qquad a_{m,n}x_n &= d_m \end{aligned}$$

if and only if

$$\begin{aligned} a_{1,1}s_1 + \cdots + a_{1,n}s_n &= d_1 \\ &\vdots \\ \text{and } a_{i,1}s_1 + \cdots + a_{i,n}s_n &= d_i \\ &\vdots \\ \text{and } (ka_{i,1} + a_{j,1})s_1 + \cdots + (ka_{i,n} + a_{j,n})s_n &= kd_i + d_j \\ &\vdots \\ \text{and } a_{m,1}s_1 + a_{m,2}s_2 + \cdots + a_{m,n}s_n &= d_m \end{aligned}$$

again by the definition of 'satisfies'. Subtract k times the i-th equation from the j-th equation (remark: here is where $i \neq j$ is needed; if $i = j$ then the two d_i's above are not equal) to get that the previous compound statement holds if and only if

$$\begin{aligned} a_{1,1}s_1 + \cdots + a_{1,n}s_n &= d_1 \\ &\vdots \\ \text{and } a_{i,1}s_1 + \cdots + a_{i,n}s_n &= d_i \\ &\vdots \\ \text{and } (ka_{i,1} + a_{j,1})s_1 + \cdots + (ka_{i,n} + a_{j,n})s_n & \\ -(ka_{i,1}s_1 + \cdots + ka_{i,n}s_n) &= kd_i + d_j - kd_i \\ &\vdots \\ \text{and } a_{m,1}s_1 + \cdots + a_{m,n}s_n &= d_m \end{aligned}$$

which, after cancellation, says that $(s_1, \ldots, s_n)$ solves

$$\begin{aligned}
a_{1,1}x_1 + \cdots + a_{1,n}x_n &= d_1 \\
&\vdots \\
a_{i,1}x_1 + \cdots + a_{i,n}x_n &= d_i \\
&\vdots \\
a_{j,1}x_1 + \cdots + a_{j,n}x_n &= d_j \\
&\vdots \\
a_{m,1}x_1 + \cdots + a_{m,n}x_n &= d_m
\end{aligned}$$

as required.

1.I.1.30 Yes, this one-equation system:

$$0x + 0y = 0$$

is satisfied by every $(x, y) \in \mathbb{R}^2$.

1.I.1.32 Swapping rows is reversed by swapping back.

$$\begin{array}{c} a_{1,1}x_1 + \cdots + a_{1,n}x_n = d_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = d_m \end{array} \xrightarrow{\rho_i \leftrightarrow \rho_j} \xrightarrow{\rho_j \leftrightarrow \rho_i} \begin{array}{c} a_{1,1}x_1 + \cdots + a_{1,n}x_n = d_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = d_m \end{array}$$

Multiplying both sides of a row by $k \neq 0$ is reversed by dividing by k.

$$\begin{array}{c} a_{1,1}x_1 + \cdots + a_{1,n}x_n = d_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = d_m \end{array} \xrightarrow{k\rho_i} \xrightarrow{(1/k)\rho_i} \begin{array}{c} a_{1,1}x_1 + \cdots + a_{1,n}x_n = d_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = d_m \end{array}$$

Adding k times a row to another is reversed by adding $-k$ times that row.

$$\begin{array}{c} a_{1,1}x_1 + \cdots + a_{1,n}x_n = d_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = d_m \end{array} \xrightarrow{k\rho_i + \rho_j} \xrightarrow{-k\rho_i + \rho_j} \begin{array}{c} a_{1,1}x_1 + \cdots + a_{1,n}x_n = d_1 \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = d_m \end{array}$$

Remark: observe for the third case that if $i = j$ then the result doesn't hold:

$$3x + 2y = 7 \xrightarrow{2\rho_1 + \rho_1} 9x + 6y = 21 \xrightarrow{-2\rho_1 + \rho_1} -9x - 6y = -21$$

1.I.1.33 Let p, n, and d be the number of pennies, nickels, and dimes. For real variables, this system

$$\begin{array}{rl} p + n + d = 13 \\ p + 5n + 10d = 83 \end{array} \xrightarrow{-\rho_1 + \rho_2} \begin{array}{rl} p + n + d = 13 \\ 4n + 9d = 70 \end{array}$$

has infinitely many solutions. However, it has a limited number of solutions in which p, n, and d are non-negative integers. Running through $d = 0, \ldots, d = 8$ shows that $(p, n, d) = (3, 4, 6)$ is the only sensible solution.

1.I.1.34 Solving the system

$$\begin{aligned}
(1/3)(a + b + c) + d &= 29 \\
(1/3)(b + c + d) + a &= 23 \\
(1/3)(c + d + a) + b &= 21 \\
(1/3)(d + a + b) + c &= 17
\end{aligned}$$

we obtain $a = 12$, $b = 9$, $c = 3$, $d = 21$. Thus the second item, 21, is the correct answer.

1.I.1.36 Eight commissioners voted for B. To see this, we will use the given information to study how many voters chose each order of A, B, C.

The six orders of preference are ABC, ACB, BAC, BCA, CAB, CBA; assume they receive a, b, c, d, e, f votes respectively. We know that

$$\begin{aligned}
a + b + e &= 11 \\
d + e + f &= 12 \\
a + c + d &= 14
\end{aligned}$$

from the number preferring A over B, the number preferring C over A, and the number preferring B over C. Because 20 votes were cast, we also know that

$$\begin{aligned} c+d+f&=9\\ a+b+c&=8\\ b+e+f&=6 \end{aligned}$$

from the preferences for B over A, for A over C, and for C over B.

The solution is $a=6$, $b=1$, $c=1$, $d=7$, $e=4$, and $f=1$. The number of commissioners voting for B as their first choice is therefore $c+d=1+7=8$.

Comments. The answer to this question would have been the same had we known only that *at least* 14 commissioners preferred B over C.

The seemingly paradoxical nature of the commissioners's preferences (A is preferred to B, and B is preferred to C, and C is preferred to A), an example of "non-transitive dominance", is not uncommon when individual choices are pooled.

1.I.1.37 (This is how the solution appeared in the *Monthly*. We have not used the word "dependent" yet; it means here that Gauss' method shows that there is not a unique solution.) If $n\geq 3$ the system is dependent and the solution is not unique. Hence $n<3$. But the term "system" implies $n>1$. Hence $n=2$. If the equations are

$$\begin{aligned} ax+\ (a+d)y&=a+2d\\ (a+3d)x+(a+4d)y&=a+5d \end{aligned}$$

then $x=-1$, $y=2$.

. . . .

Answers for subsection 1.I.2

1.I.2.21 For each problem we get a system of linear equations by looking at the equations of components.

(a) Yes; take $k=-1/2$.
(b) No; the system with equations $5=5\cdot j$ and $4=-4\cdot j$ has no solution.
(c) Yes; take $r=2$.
(d) No. The second components give $k=0$. Then the third components give $j=1$. But the first components don't check.

1.I.2.22 This system has 1 equation. The leading variable is x_1, the other variables are free.

$$\{\begin{pmatrix}-1\\1\\\vdots\\0\end{pmatrix}x_2+\cdots+\begin{pmatrix}-1\\0\\\vdots\\1\end{pmatrix}x_n \;\big|\; x_1,\ldots,x_n\in\mathbb{R}\}$$

1.I.2.26
(a) $\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}$ **(b)** $\begin{pmatrix}2&1\\-3&1\end{pmatrix}$ **(c)** $\begin{pmatrix}5&10\\10&5\end{pmatrix}$ **(d)** $\begin{pmatrix}1&1&0\end{pmatrix}$

1.I.2.28 On plugging in the five pairs (x,y) we get a system with the five equations and six unknowns a, ..., f. Because there are more unknowns than equations, if no inconsistency exists among the equations then there are infinitely many solutions (at least one variable will end up free).

But no inconsistency can exist because $a=0,\ldots,f=0$ is a solution (we are only using this zero solution to show that the system is consistent — the prior paragraph shows that there are nonzero solutions).

1.I.2.29

(a) Here is one — the fourth equation is redundant but still OK.

$$\begin{aligned} x+y-\ \ z+\ \ w&=0\\ y-\ \ z\qquad\ &=0\\ 2z+2w&=0\\ z+\ \ w&=0 \end{aligned}$$

(b) Here is one.

$$\begin{aligned} x+y-z+w&=0\\ w&=0\\ w&=0\\ w&=0 \end{aligned}$$

(c) This is one.

$$\begin{aligned} x+y-z+w&=0\\ x+y-z+w&=0\\ x+y-z+w&=0\\ x+y-z+w&=0 \end{aligned}$$

1.I.2.30

(a) Formal solution of the system yields

$$x=\frac{a^3-1}{a^2-1}\qquad y=\frac{-a^2+a}{a^2-1}.$$

If $a+1\neq 0$ and $a-1\neq 0$, then the system has the single solution

$$x=\frac{a^2+a+1}{a+1}\qquad y=\frac{-a}{a+1}.$$

If $a=-1$, or if $a=+1$, then the formulas are meaningless; in the first instance we arrive at the system

$$\begin{cases}-x+y=1,\\ \ \ x-y=1,\end{cases}$$

which is a contradictory system. In the second instance we have

$$\begin{cases}x+y=1,\\ x+y=1,\end{cases}$$

which has an infinite number of solutions (for example, for x arbitrary, $y=1-x$).

(b) Solution of the system yields

$$x=\frac{a^4-1}{a^2-1}\qquad y=\frac{-a^3+a}{a^2-1}.$$

Here, is $a^2-1\neq 0$, the system has the single solution $x=a^2+1$, $y=-a$. For $a=-1$ and $a=1$, we obtain the systems

$$\begin{cases}-x+y=-1,\\ \ \ x-y=\ \ 1\end{cases}\qquad \begin{cases}x+y=1,\\ x+y=1,\end{cases}$$

both of which have an infinite number of solutions.

1.I.2.31 (This is how the answer appeared in *Math Magazine*.) Let u, v, x, y, z be the volumes in cm^3 of Al, Cu, Pb, Ag, and Au, respectively, contained in the sphere, which we assume to be not hollow. Since the loss of weight in water (specific gravity 1.00) is 1000 grams, the volume of the sphere is 1000 cm^3. Then the data, some of which is superfluous, though consistent, leads to only 2 independent equations, one relating volumes and the other, weights.

$$\begin{aligned} u+\ \ \ \ v+\ \ \ \ \ \ x+\ \ \ \ \ \ y+\ \ \ \ \ \ z&=1000\\ 2.7u+8.9v+11.3x+10.5y+19.3z&=7558 \end{aligned}$$

Clearly the sphere must contain some aluminum to bring its mean specific gravity below the specific gravities of all the other metals. There is no unique result to this part of the problem, for the amounts of three metals may be chosen arbitrarily, provided that the choices will not result in negative amounts of any metal.

If the ball contains only aluminum and gold, there are 294.5 cm^3 of gold and 705.5 cm^3 of aluminum. Another possibility is 124.7 cm^3 each of Cu, Au, Pb, and Ag and 501.2 cm^3 of Al.

. . . .

Answers for subsection 1.I.3

1.I.3.16 The answers from the prior subsection show the row operations.

(a) The solution set is

$$\{\begin{pmatrix}2/3\\-1/3\\0\end{pmatrix}+\begin{pmatrix}1/6\\2/3\\1\end{pmatrix}z \bigm| z\in\mathbb{R}\}.$$

A particular solution and the solution set for the associated homogeneous system are

$$\begin{pmatrix}2/3\\-1/3\\0\end{pmatrix}\quad\text{and}\quad\{\begin{pmatrix}1/6\\2/3\\1\end{pmatrix}z \bigm| z\in\mathbb{R}\}.$$

(b) The solution set is

$$\{\begin{pmatrix}1\\3\\0\\0\end{pmatrix}+\begin{pmatrix}1\\-2\\1\\0\end{pmatrix}z \bigm| z\in\mathbb{R}\}.$$

A particular solution and the solution set for the associated homogeneous system are

$$\begin{pmatrix}1\\3\\0\\0\end{pmatrix}\quad\text{and}\quad\{\begin{pmatrix}1\\-2\\1\\0\end{pmatrix}z \bigm| z\in\mathbb{R}\}.$$

(c) The solution set is

$$\{\begin{pmatrix}0\\0\\0\\0\end{pmatrix}+\begin{pmatrix}-1\\0\\1\\0\end{pmatrix}z+\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix}w \bigm| z,w\in\mathbb{R}\}.$$

A particular solution and the solution set for the associated homogeneous system are

$$\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\quad\text{and}\quad\{\begin{pmatrix}-1\\0\\1\\0\end{pmatrix}z+\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix}w \bigm| z,w\in\mathbb{R}\}.$$

(d) The solution set is

$$\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}+\begin{pmatrix}-5/7\\-8/7\\1\\0\\0\end{pmatrix}c+\begin{pmatrix}-3/7\\-2/7\\0\\1\\0\end{pmatrix}d+\begin{pmatrix}-1/7\\4/7\\0\\0\\1\end{pmatrix}e \bigm| c,d,e\in\mathbb{R}\}.$$

A particular solution and the solution set for the associated homogeneous system are

$$\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}\quad\text{and}\quad\{\begin{pmatrix}-5/7\\-8/7\\1\\0\\0\end{pmatrix}c+\begin{pmatrix}-3/7\\-2/7\\0\\1\\0\end{pmatrix}d+\begin{pmatrix}-1/7\\4/7\\0\\0\\1\end{pmatrix}e \bigm| c,d,e\in\mathbb{R}\}.$$

1.I.3.19 The first is nonsingular while the second is singular. Just do Gauss' method and see if the echelon form result has non-0 numbers in each entry on the diagonal.

1.I.3.22 Because the matrix of coefficients is nonsingular, Gauss' method ends with an echelon form where each variable leads an equation. Back substitution gives a unique solution.

(Another way to see the solution is unique is to note that with a nonsingular matrix of coefficients the associated homogeneous system has a unique solution, by definition. Since the general solution is the sum of a particular solution with each homogeneous solution, the general solution has (at most) one element.)

1.I.3.23 In this case the solution set is all of $\mathbb{R}^n$, and can be expressed in the required form

$$\{c_1\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}+c_2\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}+\cdots+c_n\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\bigm| c_1,\ldots,c_n\in\mathbb{R}\}.$$

1.I.3.25 First the proof.

Gauss' method will use only rationals (e.g., $-(m/n)\rho_i+\rho_j$). Thus the solution set can be expressed using only rational numbers as the components of each vector. Now the particular solution is all rational.

There are infinitely many (rational vector) solutions if and only if the associated homogeneous system has infinitely many (real vector) solutions. That's because setting any parameters to be rationals will produce an all-rational solution.

. . . .

Answers for subsection 1.II.1

1.II.1.5 The vector

$$\begin{pmatrix}2\\0\\3\end{pmatrix}$$

is not in the line. Because

$$\begin{pmatrix}2\\0\\3\end{pmatrix}-\begin{pmatrix}-1\\0\\-4\end{pmatrix}=\begin{pmatrix}3\\0\\7\end{pmatrix}$$

that plane can be described in this way.

$$\{\begin{pmatrix}-1\\0\\4\end{pmatrix}+m\begin{pmatrix}1\\1\\2\end{pmatrix}+n\begin{pmatrix}3\\0\\7\end{pmatrix}\bigm| m,n\in\mathbb{R}\}$$

1.II.1.8 The "if" half is straightforward. If $b_1-a_1=d_1-c_1$ and $b_2-a_2=d_2-c_2$ then

$$\sqrt{(b_1-a_1)^2+(b_2-a_2)^2}=\sqrt{(d_1-c_1)^2+(d_2-c_2)^2}$$

so they have the same lengths, and the slopes are just as easy:

$$\frac{b_2-a_2}{b_1-a_1}=\frac{d_2-c_2}{d_1-a_1}$$

(if the denominators are 0 they both have undefined slopes).

For "only if", assume that the two segments have the same length and slope (the case of undefined slopes is easy; we will do the case where both segments have a slope m). Also assume, without loss of generality, that $a_1<b_1$ and that $c_1<d_1$. The first segment is $\overline{(a_1,a_2)(b_1,b_2)}=\{(x,y)\bigm| y=mx+n_1,\ x\in[a_1..b_1]\}$ (for some intercept n_1) and the second segment is $\overline{(c_1,c_2)(d_1,d_2)}=\{(x,y)\bigm| y=mx+n_2,\ x\in[c_1..d_1]\}$ (for some n_2). Then the lengths of those segments are

$$\sqrt{(b_1-a_1)^2+((mb_1+n_1)-(ma_1+n_1))^2}=\sqrt{(1+m^2)(b_1-a_1)^2}$$

and, similarly, $\sqrt{(1+m^2)(d_1-c_1)^2}$. Therefore, $|b_1-a_1|=|d_1-c_1|$. Thus, as we assumed that $a_1<b_1$ and $c_1<d_1$, we have that $b_1-a_1=d_1-c_1$.

The other equality is similar.

1.II.1.9 We shall later define it to be a set with one element — an "origin".

1.II.1.11 Euclid no doubt is picturing a plane inside of $\mathbb{R}^3$. Observe, however, that both $\mathbb{R}^1$ and $\mathbb{R}^3$ also satisfy that definition.

. . . .

Answers for subsection 1.II.2

1.II.2.13 Solve $(k)(4)+(1)(3)=0$ to get $k=-3/4$.

1.II.2.14 The set

$$\{\begin{pmatrix}x\\y\\z\end{pmatrix} \bigm| 1x+3y-1z=0\}$$

can also be described with parameters in this way.

$$\{\begin{pmatrix}-3\\1\\0\end{pmatrix}y+\begin{pmatrix}1\\0\\1\end{pmatrix}z \bigm| y,z\in\mathbb{R}\}$$

1.II.2.16 Clearly $u_1u_1+\cdots+u_nu_n$ is zero if and only if each u_i is zero. So only $\vec{0}\in\mathbb{R}^n$ is perpendicular to itself.

1.II.2.18

(a) Verifying that $(k\vec{x})\cdot\vec{y}=k(\vec{x}\cdot\vec{y})=\vec{x}\cdot(k\vec{y})$ for $k\in\mathbb{R}$ and $\vec{x},\vec{y}\in\mathbb{R}^n$ is easy. Now, for $k\in\mathbb{R}$ and $\vec{v},\vec{w}\in\mathbb{R}^n$, if $\vec{u}=k\vec{v}$ then $\vec{u}\cdot\vec{v}=(k\vec{u})\cdot\vec{v}=k(\vec{v}\cdot\vec{v})$, which is k times a nonnegative real.

The $\vec{v}=k\vec{u}$ half is similar (actually, taking the k in this paragraph to be the reciprocal of the k above gives that we need only worry about the $k=0$ case).

(b) We first consider the $\vec{u}\cdot\vec{v}\geq 0$ case. From the Triangle Inequality we know that $\vec{u}\cdot\vec{v}=\|\vec{u}\,\|\,\|\vec{v}\,\|$ if and only if one vector is a nonnegative scalar multiple of the other. But that's all we need because the first part of this exercise shows that, in a context where the dot product of the two vectors is positive, the two statements 'one vector is a scalar multiple of the other' and 'one vector is a nonnegative scalar multiple of the other', are equivalent.

We finish by considering the $\vec{u}\cdot\vec{v}<0$ case. Because $0<|\vec{u}\cdot\vec{v}|=-(\vec{u}\cdot\vec{v})=(-\vec{u})\cdot\vec{v}$ and $\|\vec{u}\,\|\,\|\vec{v}\,\|=\|-\vec{u}\,\|\,\|\vec{v}\,\|$, we have that $0<(-\vec{u})\cdot\vec{v}=\|-\vec{u}\,\|\,\|\vec{v}\,\|$. Now the prior paragraph applies to give that one of the two vectors $-\vec{u}$ and $\vec{v}$ is a scalar multiple of the other. But that's equivalent to the assertion that one of the two vectors $\vec{u}$ and $\vec{v}$ is a scalar multiple of the other, as desired.

1.II.2.19 No. These give an example.

$$\vec{u}=\begin{pmatrix}1\\0\end{pmatrix}\qquad\vec{v}=\begin{pmatrix}1\\0\end{pmatrix}\qquad\vec{w}=\begin{pmatrix}1\\1\end{pmatrix}$$

1.II.2.22 Assume that $\vec{v}\in\mathbb{R}^n$ has components $v_1,\ldots,v_n$. If $\vec{v}\neq\vec{0}$ then we have this.

$$\begin{aligned}&\sqrt{\left(\frac{v_1}{\sqrt{{v_1}^2+\cdots+{v_n}^2}}\right)^2+\cdots+\left(\frac{v_n}{\sqrt{{v_1}^2+\cdots+{v_n}^2}}\right)^2}\\&\qquad=\sqrt{\left(\frac{{v_1}^2}{{v_1}^2+\cdots+{v_n}^2}\right)+\cdots+\left(\frac{{v_n}^2}{{v_1}^2+\cdots+{v_n}^2}\right)}\\&\qquad=1\end{aligned}$$

If $\vec{v}=\vec{0}$ then $\vec{v}/\|\vec{v}\,\|$ is not defined.

1.II.2.23 For the first question, assume that $\vec{v}\in\mathbb{R}^n$ and $r\geq 0$, take the root, and factor.

$$\|r\vec{v}\,\|=\sqrt{(rv_1)^2+\cdots+(rv_n)^2}=\sqrt{r^2({v_1}^2+\cdots+{v_n}^2}=r\|\vec{v}\,\|$$

For the second question, the result is r times as long, but it points in the opposite direction in that $r\vec{v}+(-r)\vec{v}=\vec{0}$.

1.II.2.25 Write

$$\vec{u}=\begin{pmatrix}u_1\\\vdots\\u_n\end{pmatrix}\qquad\vec{v}=\begin{pmatrix}v_1\\\vdots\\v_n\end{pmatrix}$$

and then this computation works.

$$\begin{aligned}\|\vec{u}+\vec{v}\,\|^2+\|\vec{u}-\vec{v}\,\|^2 &= (u_1+v_1)^2+\cdots+(u_n+v_n)^2\\ &\quad+(u_1-v_1)^2+\cdots+(u_n-v_n)^2\\ &= {u_1}^2+2u_1v_1+{v_1}^2+\cdots+{u_n}^2+2u_nv_n+{v_n}^2\\ &\quad+{u_1}^2-2u_1v_1+{v_1}^2+\cdots+{u_n}^2-2u_nv_n+{v_n}^2\\ &= 2({u_1}^2+\cdots+{u_n}^2)+2({v_1}^2+\cdots+{v_n}^2)\\ &= 2\|\vec{u}\,\|^2+2\|\vec{v}\,\|^2\end{aligned}$$

1.II.2.26 We will prove this demonstrating that the contrapositive statement holds: if $\vec{x}\neq\vec{0}$ then there is a $\vec{y}$ with $\vec{x}\cdot\vec{y}\neq 0$.

Assume that $\vec{x}\in\mathbb{R}^n$. If $\vec{x}\neq\vec{0}$ then it has a nonzero component, say the i-th one x_i. But the vector $\vec{y}\in\mathbb{R}^n$ that is all zeroes except for a one in component i gives $\vec{x}\cdot\vec{y}=x_i$. (A slicker proof just considers $\vec{x}\cdot\vec{x}$.)

1.II.2.27 Yes. We prove this by induction.

Assume that the vectors are in some $\mathbb{R}^k$. Clearly the statement applies to one vector. The Triangle Inequality is this statement applied to two vectors. For an inductive step assume the statement is true for n or fewer vectors. Then this

$$\|\vec{u}_1+\cdots+\vec{u}_n+\vec{u}_{n+1}\|\leq\|\vec{u}_1+\cdots+\vec{u}_n\|+\|\vec{u}_{n+1}\|$$

follows by the Triangle Inequality for two vectors. Now the inductive hypothesis, applied to the first summand on the right, gives that as less than or equal to $\|\vec{u}_1\|+\cdots+\|\vec{u}_n\|+\|\vec{u}_{n+1}\|$.

1.II.2.28 By definition

$$\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\,\|\,\|\vec{v}\,\|}=\cos\theta$$

where θ is the angle between the vectors. Thus the ratio is $|\cos\theta|$.

1.II.2.29 So that the statement 'vectors are orthogonal iff their dot product is zero' has no exceptions.

1.II.2.30 The angle between (a) and (b) is found (for $a,b\neq 0$) with

$$\arccos(\frac{ab}{\sqrt{a^2}\sqrt{b^2}}).$$

If a or b is zero then the angle is $\pi/2$ radians. Otherwise, if a and b are of opposite signs then the angle is π radians, else the angle is zero radians.

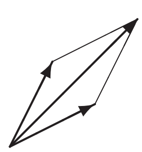
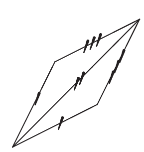
1.II.2.31 The angle between $\vec{u}$ and $\vec{v}$ is acute if $\vec{u}\cdot\vec{v}>0$, is right if $\vec{u}\cdot\vec{v}=0$, and is obtuse if $\vec{u}\cdot\vec{v}<0$. That's because, in the formula for the angle, the denominator is never negative.

1.II.2.33 Where $\vec{u},\vec{v}\in\mathbb{R}^n$, the vectors $\vec{u}+\vec{v}$ and $\vec{u}-\vec{v}$ are perpendicular if and only if $0=(\vec{u}+\vec{v})\cdot(\vec{u}-\vec{v})=\vec{u}\cdot\vec{u}-\vec{v}\cdot\vec{v}$, which shows that those two are perpendicular if and only if $\vec{u}\cdot\vec{u}=\vec{v}\cdot\vec{v}$. That holds if and only if $\|\vec{u}\,\|=\|\vec{v}\,\|$.

1.II.2.34 Suppose $\vec{u}\in\mathbb{R}^n$ is perpendicular to both $\vec{v}\in\mathbb{R}^n$ and $\vec{w}\in\mathbb{R}^n$. Then, for any $k,m\in\mathbb{R}$ we have this.

$$\vec{u}\cdot(k\vec{v}+m\vec{w})=k(\vec{u}\cdot\vec{v})+m(\vec{u}\cdot\vec{w})=k(0)+m(0)=0$$

1.II.2.35 We will show something more general: if $\|\vec{z}_1\|=\|\vec{z}_2\|$ for $\vec{z}_1,\vec{z}_2\in\mathbb{R}^n$, then $\vec{z}_1+\vec{z}_2$ bisects the angle between $\vec{z}_1$ and $\vec{z}_2$

 gives 

(we ignore the case where $\vec{z}_1$ and $\vec{z}_2$ are the zero vector).

The $\vec{z}_1+\vec{z}_2=\vec{0}$ case is easy. For the rest, by the definition of angle, we will be done if we show this.

$$\frac{\vec{z}_1\cdot(\vec{z}_1+\vec{z}_2)}{\|\vec{z}_1\|\,\|\vec{z}_1+\vec{z}_2\|}=\frac{\vec{z}_2\cdot(\vec{z}_1+\vec{z}_2)}{\|\vec{z}_2\|\,\|\vec{z}_1+\vec{z}_2\|}$$

But distributing inside each expression gives

$$\frac{\vec{z}_1\cdot\vec{z}_1+\vec{z}_1\cdot\vec{z}_2}{\|\vec{z}_1\|\,\|\vec{z}_1+\vec{z}_2\|}\qquad\frac{\vec{z}_2\cdot\vec{z}_1+\vec{z}_2\cdot\vec{z}_2}{\|\vec{z}_2\|\,\|\vec{z}_1+\vec{z}_2\|}$$

and $\vec{z}_1\cdot\vec{z}_1=\|\vec{z}_1\|=\|\vec{z}_2\|=\vec{z}_2\cdot\vec{z}_2$, so the two are equal.

1.II.2.36 We can show the two statements together. Let $\vec{u},\vec{v}\in\mathbb{R}^n$, write

$$\vec{u}=\begin{pmatrix}u_1\\ \vdots\\ u_n\end{pmatrix}\qquad\vec{v}=\begin{pmatrix}v_1\\ \vdots\\ v_n\end{pmatrix}$$

and calculate.

$$\cos\theta=\frac{ku_1v_1+\cdots+ku_nv_n}{\sqrt{(ku_1)^2+\cdots+(ku_n)^2}\sqrt{{b_1}^2+\cdots+{b_n}^2}}=\frac{k}{|k|}\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}=\pm\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}$$

1.II.2.39 *This is how the answer was given in the cited source.* The actual velocity $\vec{v}$ of the wind is the sum of the ship's velocity and the apparent velocity of the wind. Without loss of generality we may assume $\vec{a}$ and $\vec{b}$ to be unit vectors, and may write

$$\vec{v}=\vec{v}_1+s\vec{a}=\vec{v}_2+t\vec{b}$$

where s and t are undetermined scalars. Take the dot product first by $\vec{a}$ and then by $\vec{b}$ to obtain

$$\begin{aligned}s-t\vec{a}\cdot\vec{b}&=\vec{a}\cdot(\vec{v}_2-\vec{v}_1)\\ s\vec{a}\cdot\vec{b}-t&=\vec{b}\cdot(\vec{v}_2-\vec{v}_1)\end{aligned}$$

Multiply the second by $\vec{a}\cdot\vec{b}$, subtract the result from the first, and find

$$s=\frac{[\vec{a}-(\vec{a}\cdot\vec{b})\vec{b}]\cdot(\vec{v}_2-\vec{v}_1)}{1-(\vec{a}\cdot\vec{b})^2}.$$

Substituting in the original displayed equation, we get

$$\vec{v}=\vec{v}_1+\frac{[\vec{a}-(\vec{a}\cdot\vec{b})\vec{b}]\cdot(\vec{v}_2-\vec{v}_1)\vec{a}}{1-(\vec{a}\cdot\vec{b})^2}.$$

1.II.2.40 We use induction on n.

In the $n=1$ base case the identity reduces to

$$(a_1b_1)^2=({a_1}^2)({b_1}^2)-0$$

and clearly holds.

For the inductive step assume that the formula holds for the $0,\ldots,n$ cases. We will show that it then holds in the $n+1$ case. Start with the right-hand side

$$\begin{aligned}&\Bigl(\sum_{1\le j\le n+1}{a_j}^2\Bigr)\Bigl(\sum_{1\le j\le n+1}{b_j}^2\Bigr)-\sum_{1\le k<j\le n+1}\bigl(a_kb_j-a_jb_k\bigr)^2\\
&=\Bigl[\Bigl(\sum_{1\le j\le n}{a_j}^2\Bigr)+{a_{n+1}}^2\Bigr]\Bigl[\Bigl(\sum_{1\le j\le n}{b_j}^2\Bigr)+{b_{n+1}}^2\Bigr]-\Bigl[\sum_{1\le k<j\le n}\bigl(a_kb_j-a_jb_k\bigr)^2+\sum_{1\le k\le n}\bigl(a_kb_{n+1}-a_{n+1}b_k\bigr)^2\Bigr]\\
&=\Bigl(\sum_{1\le j\le n}{a_j}^2\Bigr)\Bigl(\sum_{1\le j\le n}{b_j}^2\Bigr)+\sum_{1\le j\le n}{b_j}^2{a_{n+1}}^2+\sum_{1\le j\le n}{a_j}^2{b_{n+1}}^2+{a_{n+1}}^2{b_{n+1}}^2\\
&\qquad-\Bigl[\sum_{1\le k<j\le n}\bigl(a_kb_j-a_jb_k\bigr)^2+\sum_{1\le k\le n}\bigl(a_kb_{n+1}-a_{n+1}b_k\bigr)^2\Bigr]\\
&=\Bigl(\sum_{1\le j\le n}{a_j}^2\Bigr)\Bigl(\sum_{1\le j\le n}{b_j}^2\Bigr)-\sum_{1\le k<j\le n}\bigl(a_kb_j-a_jb_k\bigr)^2+\sum_{1\le j\le n}{b_j}^2{a_{n+1}}^2+\sum_{1\le j\le n}{a_j}^2{b_{n+1}}^2+{a_{n+1}}^2{b_{n+1}}^2\\
&\qquad-\sum_{1\le k\le n}\bigl(a_kb_{n+1}-a_{n+1}b_k\bigr)^2\end{aligned}$$

and apply the inductive hypothesis

$$
\begin{aligned}
&= \big(\sum_{1\le j\le n} a_jb_j\big)^2 + \sum_{1\le j\le n} {b_j}^2{a_{n+1}}^2 + \sum_{1\le j\le n} {a_j}^2{b_{n+1}}^2 + {a_{n+1}}^2{b_{n+1}}^2 \\
&\qquad - \big[\sum_{1\le k\le n} {a_k}^2{b_{n+1}}^2 - 2\sum_{1\le k\le n} a_kb_{n+1}a_{n+1}b_k + \sum_{1\le k\le n} {a_{n+1}}^2{b_k}^2\big] \\
&= \big(\sum_{1\le j\le n} a_jb_j\big)^2 - 2\big(\sum_{1\le k\le n} a_kb_{n+1}a_{n+1}b_k\big) + {a_{n+1}}^2{b_{n+1}}^2 \\
&= \big[\big(\sum_{1\le j\le n} a_jb_j\big) + a_{n+1}b_{n+1}\big]^2
\end{aligned}
$$

to derive the left-hand side.

. . . .

Answers for subsection 1.III.1

1.III.1.10 Routine Gauss' method gives one:

$$
\xrightarrow[-(1/2)\rho_1+\rho_3]{-3\rho_1+\rho_2}
\begin{pmatrix} 2 & 1 & 1 & 3 \\ 0 & 1 & -2 & -7 \\ 0 & 9/2 & 1/2 & 7/2 \end{pmatrix}
\xrightarrow{-(9/2)\rho_2+\rho_3}
\begin{pmatrix} 2 & 1 & 1 & 3 \\ 0 & 1 & -2 & -7 \\ 0 & 0 & 19/2 & 35 \end{pmatrix}
$$

and any cosmetic change, like multiplying the bottom row by 2,

$$
\begin{pmatrix} 2 & 1 & 1 & 3 \\ 0 & 1 & -2 & -7 \\ 0 & 0 & 19 & 70 \end{pmatrix}
$$

gives another.

1.III.1.13

(a) The $\rho_i \leftrightarrow \rho_i$ operation does not change A.

(b) For instance,

$$
\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}
\xrightarrow{-\rho_1+\rho_1}
\begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix}
\xrightarrow{\rho_1+\rho_1}
\begin{pmatrix} 0 & 0 \\ 3 & 4 \end{pmatrix}
$$

leaves the matrix changed.

(c) If $i \neq j$ then

$$
\begin{pmatrix} \vdots & & \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & & \\ a_{j,1} & \cdots & a_{j,n} \\ \vdots & & \end{pmatrix}
\xrightarrow{k\rho_i+\rho_j}
\begin{pmatrix} \vdots & & \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & & \\ ka_{i,1}+a_{j,1} & \cdots & ka_{i,n}+a_{j,n} \\ \vdots & & \end{pmatrix}
$$

$$
\xrightarrow{-k\rho_i+\rho_j}
\begin{pmatrix} \vdots & & \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & & \\ -ka_{i,1}+ka_{i,1}+a_{j,1} & \cdots & -ka_{i,n}+ka_{i,n}+a_{j,n} \\ \vdots & & \end{pmatrix}
$$

does indeed give A back. (Of course, if $i = j$ then the third matrix would have entries of the form $-k(ka_{i,j}+a_{i,j})+ka_{i,j}+a_{i,j}$.)

. . . .

Answers for subsection 1.III.2

1.III.2.12 First, the only matrix row equivalent to the matrix of all 0's is itself (since row operations have no effect).

Second, the matrices that reduce to

$$\begin{pmatrix} 1 & a \\ 0 & 0 \end{pmatrix}$$

have the form

$$\begin{pmatrix} b & ba \\ c & ca \end{pmatrix}$$

(where $a, b, c \in \mathbb{R}$).

Next, the matrices that reduce to

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

have the form

$$\begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}$$

(where $a, b \in \mathbb{R}$).

Finally, the matrices that reduce to

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

are the nonsingular matrices. That's because a linear system for which this is the matrix of coefficients will have a unique solution, and that is the definition of nonsingular. (Another way to say the same thing is to say that they fall into none of the above classes.)

1.III.2.13

(a) They have the form

$$\begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix}$$

where $a, b \in \mathbb{R}$.

(b) They have this form (for $a, b \in \mathbb{R}$).

$$\begin{pmatrix} 1a & 2a \\ 1b & 2b \end{pmatrix}$$

(c) They have the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

(for $a, b, c, d \in \mathbb{R}$) where $ad - bc \neq 0$. (This is the formula that determines when a 2×2 matrix is nonsingular.)

1.III.2.14 Infinitely many.

1.III.2.15 No. Row operations do not change the size of a matrix.

1.III.2.16

(a) A row operation on a zero matrix has no effect. Thus each zero matrix is alone in its row equivalence class.

(b) No. Any nonzero entry can be rescaled.

1.III.2.20

(a) If there is a linear relationship where c_0 is not zero then we can subtract $c_0\vec{\beta}_0$ and divide both sides by c_0 to get $\vec{\beta}_0$ as a linear combination of the others. (Remark. If there are no others — if the relationship is, say, $\vec{0} = 3 \cdot \vec{0}$ — then the statement is still true because zero is by definition the sum of the empty set of vectors.)

If $\vec{\beta}_0$ is a combination of the others $\vec{\beta}_0 = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n$ then subtracting $\vec{\beta}_0$ from both sides gives a relationship where one of the coefficients is nonzero, specifically, the coefficient is -1.

(b) The first row is not a linear combination of the others for the reason given in the proof: in the equation of components from the column containing the leading entry of the first row, the only nonzero entry is the leading entry from the first row, so its coefficient must be zero. Thus, from the prior part of this question, the first row is in no linear relationship with the other rows. Hence, to see if the second row can be in a linear relationship with the other rows, we can leave the first row out of the equation. But now the argument just applied to the first row will apply to the second row. (Technically, we are arguing by induction here.)

1.III.2.22

(a) The inductive step is to show that if the statement holds on rows 1 through r then it also holds on row $r + 1$. That is, we assume that $\ell_1 = k_1$, and $\ell_2 = k_2$, ... , and $\ell_r = k_r$, and we will show that $\ell_{r+1} = k_{r+1}$ also holds (for r in $1\,..\,m - 1$).

(b) Lemma 2.3 gives the relationship $\beta_{r+1} = s_{r+1,1}\delta_1 + s_{r+2,2}\delta_2 + \cdots + s_{r+1,m}\delta_m$ between rows. Inside of those rows, consider the relationship between entries in column $\ell_1 = k_1$. Because $r + 1 > 1$, the row β_{r+1} has a zero in that entry (the matrix B is in echelon form), while the row δ_1 has a nonzero entry in column k_1 (it is, by definition of k_1, the leading entry in the first row of D). Thus, in that column, the above relationship among rows resolves to this equation among numbers: $0 = s_{r+1,1} \cdot d_{1,k_1}$, with $d_{1,k_1} \neq 0$. Therefore $s_{r+1,1} = 0$.

With $s_{r+1,1} = 0$, a similar argument shows that $s_{r+1,2} = 0$. With those two, another turn gives that $s_{r+1,3} = 0$. That is, inside of the larger induction argument used to prove the entire lemma is here an subargument by induction that shows $s_{r+1,j} = 0$ for all j in $1\,..\,r$. (We won't write out the details since it is just like the induction done in Exercise 21.)

(c) First, $\ell_{r+1} < k_{r+1}$ is impossible. In the columns of D to the left of column k_{r+1} the entries are are all zeroes as $d_{r+1,k_{r+1}}$ leads the row $k+1$) and so if $\ell_{k+1} < k_{k+1}$ then the equation of entries from column ℓ_{k+1} would be $b_{r+1,\ell_{r+1}} = s_{r+1,1} \cdot 0 + \cdots + s_{r+1,m} \cdot 0$, but $b_{r+1,\ell_{r+1}}$ isn't zero since it leads its row. A symmetric argument shows that $k_{r+1} < \ell_{r+1}$ also is impossible.

1.III.2.23 The zero rows could have nonzero coefficients, and so the statement would not be true.

1.III.2.25 If multiplication of a row by zero were allowed then Lemma 2.6 would not hold. That is, where

$$\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \xrightarrow{0\rho_2} \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$$

all the rows of the second matrix can be expressed as linear combinations of the rows of the first, but the converse does not hold. The second row of the first matrix is not a linear combination of the rows of the second matrix.

1.III.2.27 Define linear systems to be equivalent if their augmented matrices are row equivalent. The proof that equivalent systems have the same solution set is easy.

. . . .

Answers for Topic: Computer Algebra Systems

1

(a) The commands

```
> A:=array( [[40,15],
             [-50,25]] );
> u:=array([100,50]);
> linsolve(A,u);
```

yield the answer $[1,4]$.

(b) Here there is a free variable:

```
> A:=array( [[7,0,-7,0],
             [8,1,-5,2],
             [0,1,-3,0],
             [0,3,-6,-1]] );
> u:=array([0,0,0,0]);
> linsolve(A,u);
```

prompts the reply $[\_t_1, 3\_t_1, \_t_1, 3\_t_1]$.

2 These are easy to type in. For instance, the first

```
> A:=array( [[2,2],
             [1,-4]] );
> u:=array([5,0]);
> linsolve(A,u);
```

gives the expected answer of $[2,1/2]$. The others are entered similarly.

(a) The answer is $x = 2$ and $y = 1/2$.

(b) The answer is $x = 1/2$ and $y = 3/2$.

(c) This system has infinitely many solutions. In the first subsection, with z as a parameter, we got $x = (43-7z)/4$ and $y = (13-z)/4$. Maple responds with $[-12+7\_t_1, \_t_1, 13-4\_t_1]$, for some reason preferring y as a parameter.

(d) There is no solution to this system. When the array A and vector u are given to Maple and it is asked to `linsolve(A,u)`, it returns no result at all, that is, it responds with no solutions.

(e) The solutions is $(x,y,z) = (5,5,0)$.

(f) There are many solutions. Maple gives $[1, -1+\_t_1, 3-\_t_1, \_t_1]$.

3 As with the prior question, entering these is easy.

(a) This system has infinitely many solutions. In the second subsection we gave the solution set as

$$\{\begin{pmatrix}6\\0\end{pmatrix} + \begin{pmatrix}-2\\1\end{pmatrix} y \mid y \in \mathbb{R}\}$$

and Maple responds with $[6-2\_t_1, \_t_1]$.

(b) The solution set has only one member

$$\{\begin{pmatrix}0\\1\end{pmatrix}\}$$

and Maple has no trouble finding it $[0,1]$.

(c) This system's solution set is infinite

$$\{\begin{pmatrix}4\\-1\\0\end{pmatrix} + \begin{pmatrix}-1\\1\\1\end{pmatrix} x_3 \mid x_3 \in \mathbb{R}\}$$

and Maple gives $[\_t_1, -\_t_1+3, -\_t_1+4]$.

(d) There is a unique solution

$$\{\begin{pmatrix}1\\1\\1\end{pmatrix}\}$$

and Maple gives $[1,1,1]$.

(e) This system has infinitely many solutions; in the second subsection we described the solution set with two parameters

$$\{\begin{pmatrix}5/3\\2/3\\0\\0\end{pmatrix} + \begin{pmatrix}-1/3\\2/3\\1\\0\end{pmatrix} z + \begin{pmatrix}-2/3\\1/3\\0\\1\end{pmatrix} w \mid z,w \in \mathbb{R}\}$$

as does Maple $[3-2\_t_1+\_t_2, \_t_1, \_t_2, -2+3\_t_1-2\_t_2]$.

(f) The solution set is empty and Maple replies to the `linsolve(A,u)` command with no returned solutions.

4 In response to this prompting

```
> A:=array( [[a,c],
             [b,d]] );
> u:=array([p,q]);
> linsolve(A,u);
```

Maple thought for perhaps twenty seconds and gave this reply.

$$[-\frac{-d\,p+q\,c}{-b\,c+a\,d},\frac{-b\,p+a\,q}{-b\,c+a\,d}]$$

. . . .

Answers for Topic: Input-Output Analysis

1 These answers were given by Octave.
(a) $s = 33\,379$, $a = 43\,304$
(b) $s = 37\,284$, $a = 43\,589$
(c) $s = 37\,411$, $a = 43\,589$

2 Octave gives these answers.
(a) $s = 24\,244$, $a = 30\,309$
(b) $s = 24\,267$, $a = 30\,675$

. . . .

Answers for Topic: Accuracy of Computations

1 Sceintific notation is convienent to express the two-place restriction. We have $.25 \times 10^2 + .67 \times 10^0 = .25 \times 10^2$. The $2/3$ has no apparent effect.

2 The reduction

$$\xrightarrow{-3\rho_1+\rho_2}\begin{array}{rl} x+\ 2y= & 3 \\ -8= & -7.992 \end{array}$$

gives a solution of $(x, y) = (1.002, 0.999)$.

3
(a) The fully accurate solution is that $x = 10$ and $y = 0$.
(b) The four-digit conclusion is quite different.

$$\xrightarrow{-(.3454/.0003)\rho_1+\rho_2}\left(\begin{array}{cc|c} .0003 & 1.556 & 1.569 \\ 0 & 1789 & -1805 \end{array}\right) \Longrightarrow x = 10460,\ y = -1.009$$

4
(a) For the first one, first, $(2/3) - (1/3)$ is $.666\,666\,67 - .333\,333\,33 = .333\,333\,34$ and so $(2/3) + ((2/3) - (1/3)) = .666\,666\,67 + .333\,333\,34 = 1.000\,000\,0$. For the other one, first $((2/3) + (2/3)) = .666\,666\,67 + .666\,666\,67 = 1.333\,333\,3$ and so $((2/3) + (2/3)) - (1/3) = 1.333\,333\,3 - .333\,333\,33 = .999\,999\,97$.
(b) The first equation is $.333\,333\,33\cdot x+1.000\,000\,0\cdot y = 0$ while the second is $.666\,666\,67\cdot x+2.000\,000\,0\cdot y = 0$.

5

(a) This calculation

$$\xrightarrow[-(1/3)\rho_1+\rho_3]{-(2/3)\rho_1+\rho_2} \left(\begin{array}{ccc|c} 3 & 2 & 1 & 6 \\ 0 & -(4/3)+2\varepsilon & -(2/3)+2\varepsilon & -2+4\varepsilon \\ 0 & -(2/3)+2\varepsilon & -(1/3)-\varepsilon & -1+\varepsilon \end{array}\right)$$

$$\xrightarrow{-(1/2)\rho_2+\rho_3} \left(\begin{array}{ccc|c} 3 & 2 & 1 & 6 \\ 0 & -(4/3)+2\varepsilon & -(2/3)+2\varepsilon & -2+4\varepsilon \\ 0 & \varepsilon & -2\varepsilon & -\varepsilon \end{array}\right)$$

gives a third equation of $y-2z=-1$. Substituting into the second equation gives $((-10/3)+6\varepsilon)\cdot z = (-10/3)+6\varepsilon$ so $z=1$ and thus $y=1$. With those, the first equation says that $x=1$.

(b) The solution with two digits kept

$$\left(\begin{array}{ccc|c} .30\times10^{1} & .20\times10^{1} & .10\times10^{1} & .60\times10^{1} \\ .10\times10^{1} & .20\times10^{-3} & .20\times10^{-3} & .20\times10^{1} \\ .30\times10^{1} & .20\times10^{-3} & -.10\times10^{-3} & .10\times10^{1} \end{array}\right)$$

$$\xrightarrow[-(1/3)\rho_1+\rho_3]{-(2/3)\rho_1+\rho_2} \left(\begin{array}{ccc|c} .30\times10^{1} & .20\times10^{1} & .10\times10^{1} & .60\times10^{1} \\ 0 & -.13\times10^{1} & -.67\times10^{0} & -.20\times10^{1} \\ 0 & -.67\times10^{0} & -.33\times10^{0} & -.10\times10^{1} \end{array}\right)$$

$$\xrightarrow{-(.67/1.3)\rho_2+\rho_3} \left(\begin{array}{ccc|c} .30\times10^{1} & .20\times10^{1} & .10\times10^{1} & .60\times10^{1} \\ 0 & -.13\times10^{1} & -.67\times10^{0} & -.20\times10^{1} \\ 0 & 0 & .15\times10^{-2} & .31\times10^{-2} \end{array}\right)$$

comes out to be $z=2.1$, $y=2.6$, and $x=-.43$.

. . . .

Answers for Topic: Analyzing Networks

5

(a) $40/13$
(b) 8 ohms
(c) $R = 1/((1/R_1)+(1/R_2))$

Chapter 2. Vector Spaces

. . . .

Answers for subsection 2.I.1

2.I.1.17

(a) $0 + 0x + 0x^2 + 0x^3$

(b) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

(c) The constant function $f(x) = 0$

(d) The constant function $f(n) = 0$

2.I.1.21 The usual operations $(v_0+v_1i)+(w_0+w_1i) = (v_0+w_0)+(v_1+w_1)i$ and $r(v_0+v_1i) = (rv_0)+(rv_1)i$ suffice. The check is easy.

2.I.1.23 The natural operations are $(v_1x+v_2y+v_3z)+(w_1x+w_2y+w_3z) = (v_1+w_1)x+(v_2+w_2)y+(v_3+w_3)z$ and $r\cdot(v_1x+v_2y+v_3z) = (rv_1)x+(rv_2)y+(rv_3)z$. The check that this is a vector space is easy; use Example 1.3 as a guide.

2.I.1.24 The '+' operation is not commutative; producing two members of the set witnessing this assertion is easy.

2.I.1.25

(a) It is not a vector space.

$$(1+1)\cdot\begin{pmatrix}1\\0\\0\end{pmatrix} \neq \begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}1\\0\\0\end{pmatrix}$$

(b) It is not a vector space.

$$1\cdot\begin{pmatrix}1\\0\\0\end{pmatrix} \neq \begin{pmatrix}1\\0\\0\end{pmatrix}$$

2.I.1.29

(a) No: $1\cdot(0,1) + 1\cdot(0,1) \neq (1+1)\cdot(0,1)$.

(b) Same as the prior answer.

2.I.1.30 It is not a vector space since it is not closed under addition since $(x^2) + (1 + x - x^2)$ is not in the set.

2.I.1.31

(a) 6

(b) nm

(c) 3

(d) To see that the answer is 2, rewrite it as

$$\{\begin{pmatrix} a & 0 \\ b & -a-b \end{pmatrix} \mid a, b \in \mathbb{R}\}$$

so that there are two parameters.

2.I.1.34 Addition is commutative, so in any vector space, for any vector $\vec{v}$ we have that $\vec{v} = \vec{v} + \vec{0} = \vec{0} + \vec{v}$.

2.I.1.36 Each element of a vector space has one and only one additive inverse.

For, let V be a vector space and suppose that $\vec{v} \in V$. If $\vec{w}_1, \vec{w}_2 \in V$ are both additive inverses of $\vec{v}$ then consider $\vec{w}_1 + \vec{v} + \vec{w}_2$. On the one hand, we have that it equals $\vec{w}_1 + (\vec{v} + \vec{w}_2) = \vec{w}_1 + \vec{0} = \vec{w}_1$. On the other hand we have that it equals $(\vec{w}_1 + \vec{v}) + \vec{w}_2 = \vec{0} + \vec{w}_2 = \vec{w}_2$. Therefore, $\vec{w}_1 = \vec{w}_2$.

2.I.1.37

(a) Every such set has the form $\{r\cdot\vec{v}+s\cdot\vec{w} \mid r,s\in\mathbb{R}\}$ where either or both of $\vec{v},\vec{w}$ may be $\vec{0}$. With the inherited operations, closure of addition $(r_1\vec{v}+s_1\vec{w})+(r_2\vec{v}+s_2\vec{w})=(r_1+r_2)\vec{v}+(s_1+s_2)\vec{w}$ and scalar multiplication $c(r\vec{v}+s\vec{w})=(cr)\vec{v}+(cs)\vec{w}$ are easy. The other conditions are also routine.

(b) No such set can be a vector space under the inherited operations because it does not have a zero element.

2.I.1.39 Yes. A theorem of first semester calculus says that a sum of differentiable functions is differentiable and that $(f+g)'=f'+g'$, and that a multiple of a differentiable function is differentiable and that $(r\cdot f)'=r\,f'$.

2.I.1.40 The check is routine. Note that '1' is $1+0i$ and the zero elements are these.

(a) $(0+0i)+(0+0i)x+(0+0i)x^2$

(b) $\begin{pmatrix}0+0i & 0+0i\\ 0+0i & 0+0i\end{pmatrix}$

2.I.1.41 Notably absent from the definition of a vector space is a distance measure.

2.I.1.43

(a) We outline the check of the conditions from Definition 1.1.

Item (1) has five conditions. First, additive closure holds because if $a_0+a_1+a_2=0$ and $b_0+b_1+b_2=0$ then

$$(a_0+a_1x+a_2x^2)+(b_0+b_1x+b_2x^2)=(a_0+b_0)+(a_1+b_1)x+(a_2+b_2)x^2$$

is in the set since $(a_0+b_0)+(a_1+b_1)+(a_2+b_2)=(a_0+a_1+a_2)+(b_0+b_1+b_2)$ is zero. The second through fifth conditions are easy.

Item (2) also has five conditions. First, closure under scalar multiplication holds because if $a_0+a_1+a_2=0$ then

$$r\cdot(a_0+a_1x+a_2x^2)=(ra_0)+(ra_1)x+(ra_2)x^2$$

is in the set as $ra_0+ra_1+ra_2=r(a_0+a_1+a_2)$ is zero. The second through fifth conditions here are also easy.

(b) This is similar to the prior answer.

(c) Call the vector space V. We have two implications: left to right, if S is a subspace then it is closed under linear combinations of pairs of vectors and, right to left, if a nonempty subset is closed under linear combinations of pairs of vectors then it is a subspace. The left to right implication is easy; we here sketch the other one by assuming S is nonempty and closed, and checking the conditions of Definition 1.1.

Item (1) has five conditions. First, to show closure under addition, if $\vec{s}_1,\vec{s}_2\in S$ then $\vec{s}_1+\vec{s}_2\in S$ as $\vec{s}_1+\vec{s}_2=1\cdot\vec{s}_1+1\cdot\vec{s}_2$. Second, for any $\vec{s}_1,\vec{s}_2\in S$, because addition is inherited from V, the sum $\vec{s}_1+\vec{s}_2$ in S equals the sum $\vec{s}_1+\vec{s}_2$ in V and that equals the sum $\vec{s}_2+\vec{s}_1$ in V and that in turn equals the sum $\vec{s}_2+\vec{s}_1$ in S. The argument for the third condition is similar to that for the second. For the fourth, suppose that $\vec{s}$ is in the nonempty set S and note that $0\cdot\vec{s}=\vec{0}\in S$; showing that the $\vec{0}$ of V acts under the inherited operations as the additive identity of S is easy. The fifth condition is satisfied because for any $\vec{s}\in S$ closure under linear combinations shows that the vector $0\cdot\vec{0}+(-1)\cdot\vec{s}$ is in S; showing that it is the additive inverse of $\vec{s}$ under the inherited operations is routine.

The proofs for item (2) are similar.

. . . .

Answers for subsection 2.I.2

2.I.2.23

(a) Yes; it is in that span since $1\cdot\cos^2 x+1\cdot\sin^2 x=f(x)$.

(b) No, since $r_1\cos^2 x+r_2\sin^2 x=3+x^2$ has no scalar solutions that work for all x. For instance, setting x to be 0 and π gives the two equations $r_1\cdot 1+r_2\cdot 0=3$ and $r_1\cdot 1+r_2\cdot 0=3+\pi^2$, which are not consistent with each other.

(c) No; consider what happens on setting x to be $\pi/2$ and $3\pi/2$.

(d) Yes, $\cos(2x) = 1 \cdot \cos^2(x) - 1 \cdot \sin^2(x)$.

2.I.2.27 Technically, no. Subspaces of $\mathbb{R}^3$ are sets of three-tall vectors, while $\mathbb{R}^2$ is a set of two-tall vectors. Clearly though, $\mathbb{R}^2$ is "just like" this subspace of $\mathbb{R}^3$.

$$\{\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \bigm| x, y \in \mathbb{R}\}$$

2.I.2.29 It can be improper. If $\vec{v} = \vec{0}$ then this is a trivial subspace. At the opposite extreme, if the vector space is $\mathbb{R}^1$ and $\vec{v} \neq \vec{0}$ then the subspace is all of $\mathbb{R}^1$.

2.I.2.30 No, such a set is not closed. For one thing, it does not contain the zero vector.

2.I.2.31 No. The only subspaces of $\mathbb{R}^1$ are the space itself and its trivial subspace. Any subspace S of $\mathbb{R}$ that contains a nonzero member $\vec{v}$ must contain the set of all of its scalar multiples $\{r \cdot \vec{v} \bigm| r \in \mathbb{R}\}$. But this set is all of $\mathbb{R}$.

2.I.2.32 Item (1) is checked in the text.

Item (2) has five conditions. First, for closure, if $c \in \mathbb{R}$ and $\vec{s} \in S$ then $c \cdot \vec{s} \in S$ as $c \cdot \vec{s} = c \cdot \vec{s} + 0 \cdot \vec{0}$. Second, because the operations in S are inherited from V, for $c, d \in \mathbb{R}$ and $\vec{s} \in S$, the scalar product $(c+d) \cdot \vec{s}$ in S equals the product $(c+d) \cdot \vec{s}$ in V, and that equals $c \cdot \vec{s} + d \cdot \vec{s}$ in V, which equals $c \cdot \vec{s} + d \cdot \vec{s}$ in S.

The check for the third, fourth, and fifth conditions are similar to the second conditions's check just given.

2.I.2.33 An exercise in the prior subsection shows that every vector space has only one zero vector (that is, there is only one vector that is the additive identity element of the space). But a trivial space has only one element and that element must be this (unique) zero vector.

2.I.2.35

(a) It is not a subspace because these are not the inherited operations. For one thing, in this space,

$$0 \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

while this does not, of course, hold in $\mathbb{R}^3$.

(b) We can combine the arguments showing closure under addition and scalar multiplication into one single argument showing closure under linear combinations of two vectors. If $r_1, r_2, x_1, x_2, y_1, y_2, z_1, z_2$ are in $\mathbb{R}$ then

$$r_1 \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + r_2 \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} r_1 x_1 - r_1 + 1 \\ r_1 y_1 \\ r_1 z_1 \end{pmatrix} + \begin{pmatrix} r_2 x_2 - r_2 + 1 \\ r_2 y_2 \\ r_2 z_2 \end{pmatrix} = \begin{pmatrix} r_1 x_1 - r_1 + r_2 x_2 - r_2 + 1 \\ r_1 y_1 + r_2 y_2 \\ r_1 z_1 + r_2 z_2 \end{pmatrix}$$

(note that the first component of the last vector does not say '$+\,2$' because addition of vectors in this space has the first components combine in this way: $(r_1 x_1 - r_1 + 1) + (r_2 x_2 - r_2 + 1) - 1$). Adding the three components of the last vector gives $r_1(x_1 - 1 + y_1 + z_1) + r_2(x_2 - 1 + y_2 + z_2) + 1 = r_1 \cdot 0 + r_2 \cdot 0 + 1 = 1$.

Most of the other checks of the conditions are easy (although the oddness of the operations keeps them from being routine). Commutativity of addition goes like this.

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 - 1 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix} = \begin{pmatrix} x_2 + x_1 - 1 \\ y_2 + y_1 \\ z_2 + z_1 \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

Associativity of addition has

$$(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}) + \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix} = \begin{pmatrix} (x_1 + x_2 - 1) + x_3 - 1 \\ (y_1 + y_2) + y_3 \\ (z_1 + z_2) + z_3 \end{pmatrix}$$

while

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + (\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ y_3 \\ z_3 \end{pmatrix}) = \begin{pmatrix} x_1 + (x_2 + x_3 - 1) - 1 \\ y_1 + (y_2 + y_3) \\ z_1 + (z_2 + z_3) \end{pmatrix}$$

and they are equal. The identity element with respect to this addition operation works this way

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} x+1-1 \\ y+0 \\ z+0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

and the additive inverse is similar.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} -x+2 \\ -y \\ -z \end{pmatrix} = \begin{pmatrix} x+(-x+2)-1 \\ y-y \\ z-z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

The conditions on scalar multiplication are also easy. For the first condition,

$$(r+s)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (r+s)x-(r+s)+1 \\ (r+s)y \\ (r+s)z \end{pmatrix}$$

while

$$r\begin{pmatrix} x \\ y \\ z \end{pmatrix} + s\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} rx-r+1 \\ ry \\ rz \end{pmatrix} + \begin{pmatrix} sx-s+1 \\ sy \\ sz \end{pmatrix} = \begin{pmatrix} (rx-r+1)+(sx-s+1)-1 \\ ry+sy \\ rz+sz \end{pmatrix}$$

and the two are equal. The second condition compares

$$r\cdot(\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}) = r\cdot\begin{pmatrix} x_1+x_2-1 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} = \begin{pmatrix} r(x_1+x_2-1)-r+1 \\ r(y_1+y_2) \\ r(z_1+z_2) \end{pmatrix}$$

with

$$r\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + r\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} rx_1-r+1 \\ ry_1 \\ rz_1 \end{pmatrix} + \begin{pmatrix} rx_2-r+1 \\ ry_2 \\ rz_2 \end{pmatrix} = \begin{pmatrix} (rx_1-r+1)+(rx_2-r+1)-1 \\ ry_1+ry_2 \\ rz_1+rz_2 \end{pmatrix}$$

and they are equal. For the third condition,

$$(rs)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} rsx-rs+1 \\ rsy \\ rsz \end{pmatrix}$$

while

$$r(s\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = r(\begin{pmatrix} sx-s+1 \\ sy \\ sz \end{pmatrix}) = \begin{pmatrix} r(sx-s+1)-r+1 \\ rsy \\ rsz \end{pmatrix}$$

and the two are equal. For scalar multiplication by 1 we have this.

$$1\cdot\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1x-1+1 \\ 1y \\ 1z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Thus all the conditions on a vector space are met by these two operations.

Remark. A way to understand this vector space is to think of it as the plane in $\mathbb{R}^3$

$$P = \{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \bigm| x+y+z=0\}$$

displaced away from the origin by 1 along the x-axis. Then addition becomes: to add two members of this space,

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$$

(such that $x_1+y_1+z_1=1$ and $x_2+y_2+z_2=1$) move them back by 1 to place them in P and add as usual,

$$\begin{pmatrix} x_1-1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2-1 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1+x_2-2 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} \qquad \text{(in } P\text{)}$$

and then move the result back out by 1 along the x-axis.

$$\begin{pmatrix} x_1+x_2-1 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix}.$$

Scalar multiplication is similar.

(c) For the subspace to be closed under the inherited scalar multiplication, where $\vec{v}$ is a member of that subspace,

$$0\cdot\vec{v}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

must also be a member.

The converse does not hold. Here is a subset of $\mathbb{R}^3$ that contains the origin

$$\{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\}$$

(this subset has only two elements) but is not a subspace.

2.I.2.36

(a) $(\vec{v}_1+\vec{v}_2+\vec{v}_3)-(\vec{v}_1+\vec{v}_2)=\vec{v}_3$

(b) $(\vec{v}_1+\vec{v}_2)-(\vec{v}_1)=\vec{v}_2$

(c) Surely, $\vec{v}_1$.

(d) Taking the one-long sum and subtracting gives $(\vec{v}_1)-\vec{v}_1=\vec{0}$.

2.I.2.37 Yes; any space is a subspace of itself, so each space contains the other.

2.I.2.38

(a) The union of the x-axis and the y-axis in $\mathbb{R}^2$ is one.

(b) The set of integers, as a subset of $\mathbb{R}^1$, is one.

(c) The subset $\{\vec{v}\}$ of $\mathbb{R}^2$ is one, where $\vec{v}$ is any nonzero vector.

2.I.2.39 Because vector space addition is commutative, a reordering of summands leaves a linear combination unchanged.

2.I.2.40 We always consider that span in the context of an enclosing space.

2.I.2.41 It is both 'if' and 'only if'.

For 'if', let S be a subset of a vector space V and assume $\vec{v}\in S$ satisfies $\vec{v}=c_1\vec{s}_1+\cdots+c_n\vec{s}_n$ where $c_1,\ldots,c_n$ are scalars and $\vec{s}_1,\ldots,\vec{s}_n\in S$. We must show that $[S\cup\{\vec{v}\}]=[S]$.

Containment one way, $[S]\subseteq[S\cup\{\vec{v}\}]$ is obvious. For the other direction, $[S\cup\{\vec{v}\}]\subseteq[S]$, note that if a vector is in the set on the left then it has the form $d_0\vec{v}+d_1\vec{t}_1+\cdots+d_m\vec{t}_m$ where the d's are scalars and the $\vec{t}$'s are in S. Rewrite that as $d_0(c_1\vec{s}_1+\cdots+c_n\vec{s}_n)+d_1\vec{t}_1+\cdots+d_m\vec{t}_m$ and note that the result is a member of the span of S.

The 'only if' is clearly true — adding $\vec{v}$ enlarges the span to include at least $\vec{v}$.

2.I.2.44 It is; apply Lemma 2.9. (You must consider the following. Suppose B is a subspace of a vector space V and suppose $A\subseteq B\subseteq V$ is a subspace. From which space does A inherit its operations? The answer is that it doesn't matter — A will inherit the same operations in either case.)

2.I.2.46 Call the subset S. By Lemma 2.9, we need to check that $[S]$ is closed under linear combinations. If $c_1\vec{s}_1+\cdots+c_n\vec{s}_n, c_{n+1}\vec{s}_{n+1}+\cdots+c_m\vec{s}_m\in[S]$ then for any $p,r\in\mathbb{R}$ we have

$$p\cdot(c_1\vec{s}_1+\cdots+c_n\vec{s}_n)+r\cdot(c_{n+1}\vec{s}_{n+1}+\cdots+c_m\vec{s}_m)=pc_1\vec{s}_1+\cdots+pc_n\vec{s}_n+rc_{n+1}\vec{s}_{n+1}+\cdots+rc_m\vec{s}_m$$

which is an element of $[S]$. (*Remark.* If the set S is empty, then that 'if ... then ... ' statement is vacuously true.)

2.I.2.47 For this to happen, one of the conditions giving the sensibleness of the addition and scalar multiplication operations must be violated. Consider $\mathbb{R}^2$ with these operations.

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}+\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}\qquad r\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The set $\mathbb{R}^2$ is closed under these operations. But it is not a vector space.

$$1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

. . . .

Answers for subsection 2.II.1

2.II.1.22 No, that equation is not a linear relationship. In fact this set is independent, as the system arising from taking x to be 0, $\pi/6$ and $\pi/4$ shows.

2.II.1.23 To emphasize that the equation $1 \cdot \vec{s} + (-1) \cdot \vec{s} = \vec{0}$ does not make the set dependent.

2.II.1.26

(a) A singleton set $\{\vec{v}\}$ is linearly independent if and only if $\vec{v} \neq \vec{0}$. For the 'if' direction, with $\vec{v} \neq \vec{0}$, we can apply Lemma 1.4 by considering the relationship $c \cdot \vec{v} = \vec{0}$ and noting that the only solution is the trivial one: $c = 0$. For the 'only if' direction, just recall that Example 1.11 shows that $\{\vec{0}\}$ is linearly dependent, and so if the set $\{\vec{v}\}$ is linearly independent then $\vec{v} \neq \vec{0}$.

(*Remark.* Another answer is to say that this is the special case of Lemma 1.15 where $S = \varnothing$.)

(b) A set with two elements is linearly independent if and only if neither member is a multiple of the other (note that if one is the zero vector then it is a multiple of the other, so this case is covered). This is an equivalent statement: a set is linearly dependent if and only if one element is a multiple of the other.

The proof is easy. A set $\{\vec{v}_1, \vec{v}_2\}$ is linearly dependent if and only if there is a relationship $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$ with either $c_1 \neq 0$ or $c_2 \neq 0$ (or both). That holds if and only if $\vec{v}_1 = (-c_2/c_1)\vec{v}_2$ or $\vec{v}_2 = (-c_1/c_2)\vec{v}_1$ (or both).

2.II.1.27 This set is linearly dependent set because it contains the zero vector.

2.II.1.28 The 'if' half is given by Lemma 1.13. The converse (the 'only if' statement) does not hold. An example is to consider the vector space $\mathbb{R}^2$ and these vectors.

$$\vec{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \vec{z} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

2.II.1.29

(a) The linear system arising from

$$c_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has the unique solution $c_1 = 0$ and $c_2 = 0$.

(b) The linear system arising from

$$c_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

has the unique solution $c_1 = 8/3$ and $c_2 = -1/3$.

(c) Suppose that S is linearly independent. Suppose that we have both $\vec{v} = c_1\vec{s}_1 + \cdots + c_n\vec{s}_n$ and $\vec{v} = d_1\vec{t}_1 + \cdots + d_m\vec{t}_m$ (where the vectors are members of S). Now,

$$c_1\vec{s}_1 + \cdots + c_n\vec{s}_n = \vec{v} = d_1\vec{t}_1 + \cdots + d_m\vec{t}_m$$

can be rewritten in this way.

$$c_1\vec{s}_1 + \cdots + c_n\vec{s}_n - d_1\vec{t}_1 - \cdots - d_m\vec{t}_m = \vec{0}$$

Possibly some of the $\vec{s}$'s equal some of the $\vec{t}$'s; we can combine the associated coefficients (i.e., if $\vec{s}_i = \vec{t}_j$ then $\cdots + c_i\vec{s}_i + \cdots - d_j\vec{t}_j - \cdots$ can be rewritten as $\cdots + (c_i - d_j)\vec{s}_i + \cdots$). That equation is a linear relationship among distinct (after the combining is done) members of the set S. We've assumed that S is linearly independent, so all of the coefficients are zero. If i is such that $\vec{s}_i$ does not equal any $\vec{t}_j$ then c_i is zero. If j is such that $\vec{t}_j$ does not equal any $\vec{s}_i$ then d_j is zero. In the final case, we have that $c_i - d_j = 0$ and so $c_i = d_j$.

Therefore, the original two sums are the same, except perhaps for some $0 \cdot \vec{s}_i$ or $0 \cdot \vec{t}_j$ terms that we can neglect.

(d) This set is not linearly independent:

$$S = \{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}\} \subset \mathbb{R}^2$$

and these two linear combinations give the same result

$$\begin{pmatrix}0\\0\end{pmatrix} = 2 \cdot \begin{pmatrix}1\\0\end{pmatrix} - 1 \cdot \begin{pmatrix}2\\0\end{pmatrix} = 4 \cdot \begin{pmatrix}1\\0\end{pmatrix} - 2 \cdot \begin{pmatrix}2\\0\end{pmatrix}$$

Thus, a linearly dependent set might have indistinct sums.

In fact, this stronger statement holds: if a set is linearly dependent then it must have the property that there are two distinct linear combinations that sum to the same vector. Briefly, where $c_1\vec{s}_1 + \cdots + c_n\vec{s}_n = \vec{0}$ then multiplying both sides of the relationship by two gives another relationship. If the first relationship is nontrivial then the second is also.

2.II.1.30 In this 'if and only if' statement, the 'if' half is clear — if the polynomial is the zero polynomial then the function that arises from the action of the polynomial must be the zero function $x \mapsto 0$. For 'only if' we write $p(x) = c_n x^n + \cdots + c_0$. Plugging in zero $p(0) = 0$ gives that $c_0 = 0$. Taking the derivative and plugging in zero $p'(0) = 0$ gives that $c_1 = 0$. Similarly we get that each c_i is zero, and p is the zero polynomial.

2.II.1.31 The work in this section suggests that an n-dimensional non-degenerate linear surface should be defined as the span of a linearly independent set of n vectors.

2.II.1.32

(a) For any $a_{1,1}, \ldots, a_{2,4}$,

$$c_1 \begin{pmatrix}a_{1,1}\\a_{2,1}\end{pmatrix} + c_2 \begin{pmatrix}a_{1,2}\\a_{2,2}\end{pmatrix} + c_3 \begin{pmatrix}a_{1,3}\\a_{2,3}\end{pmatrix} + c_4 \begin{pmatrix}a_{1,4}\\a_{2,4}\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

yields a linear system

$$\begin{aligned} a_{1,1}c_1 + a_{1,2}c_2 + a_{1,3}c_3 + a_{1,4}c_4 &= 0 \\ a_{2,1}c_1 + a_{2,2}c_2 + a_{2,3}c_3 + a_{2,4}c_4 &= 0 \end{aligned}$$

that has infinitely many solutions (Gauss' method leaves at least two variables free). Hence there are nontrivial linear relationships among the given members of $\mathbb{R}^2$.

(b) Any set five vectors is a superset of a set of four vectors, and so is linearly dependent.

With three vectors from $\mathbb{R}^2$, the argument from the prior item still applies, with the slight change that Gauss' method now only leaves at least one variable free (but that still gives infintely many solutions).

(c) The prior item shows that no three-element subset of $\mathbb{R}^2$ is independent. We know that there are two-element subsets of $\mathbb{R}^2$ that are independent — one is

$$\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}$$

and so the answer is two.

2.II.1.34 Yes. The two improper subsets, the entire set and the empty subset, serve as examples.

2.II.1.35 In $\mathbb{R}^4$ the biggest linearly independent set has four vectors. There are many examples of such sets, this is one.

$$\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\}$$

To see that no set with five or more vectors can be independent, set up

$$c_1 \begin{pmatrix}a_{1,1}\\a_{2,1}\\a_{3,1}\\a_{4,1}\end{pmatrix} + c_2 \begin{pmatrix}a_{1,2}\\a_{2,2}\\a_{3,2}\\a_{4,2}\end{pmatrix} + c_3 \begin{pmatrix}a_{1,3}\\a_{2,3}\\a_{3,3}\\a_{4,3}\end{pmatrix} + c_4 \begin{pmatrix}a_{1,4}\\a_{2,4}\\a_{3,4}\\a_{4,4}\end{pmatrix} + c_5 \begin{pmatrix}a_{1,5}\\a_{2,5}\\a_{3,5}\\a_{4,5}\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

and note that the resulting linear system

$$\begin{aligned} a_{1,1}c_1 + a_{1,2}c_2 + a_{1,3}c_3 + a_{1,4}c_4 + a_{1,5}c_5 &= 0 \\ a_{2,1}c_1 + a_{2,2}c_2 + a_{2,3}c_3 + a_{2,4}c_4 + a_{2,5}c_5 &= 0 \\ a_{3,1}c_1 + a_{3,2}c_2 + a_{3,3}c_3 + a_{3,4}c_4 + a_{3,5}c_5 &= 0 \\ a_{4,1}c_1 + a_{4,2}c_2 + a_{4,3}c_3 + a_{4,4}c_4 + a_{4,5}c_5 &= 0 \end{aligned}$$

has four equations and five unknowns, so Gauss' method must end with at least one c variable free, so there are infinitely many solutions, and so the above linear relationship among the four-tall vectors has more solutions than just the trivial solution.

The smallest linearly independent set is the empty set.

The biggest linearly dependent set is $\mathbb{R}^4$. The smallest is $\{\vec{0}\}$.

2.II.1.38

(a) Assuming first that $a \neq 0$,

$$x \begin{pmatrix} a \\ c \end{pmatrix} + y \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

gives

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned} \xrightarrow{-(c/a)\rho_1+\rho_2} \begin{aligned} ax + \qquad\qquad by &= 0 \\ (-(c/a)b + d)y &= 0 \end{aligned}$$

which has a solution if and only if $0 \neq -(c/a)b + d = (-cb + ad)/d$ (we've assumed in this case that $a \neq 0$, and so back substitution yields a unique solution).

The $a = 0$ case is also not hard — break it into the $c \neq 0$ and $c = 0$ subcases and note that in these cases $ad - bc = 0 \cdot d - bc$.

Comment. An earlier exercise showed that a two-vector set is linearly dependent if and only if either vector is a scalar multiple of the other. That can also be used to make the calculation.

(b) The equation

$$c_1 \begin{pmatrix} a \\ d \\ g \end{pmatrix} + c_2 \begin{pmatrix} b \\ e \\ h \end{pmatrix} + c_3 \begin{pmatrix} c \\ f \\ i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gives rise to a homogeneous linear system. We proceed by writing it in matrix form and applying Gauss' method.

We first reduce the matrix to upper-triangular. Assume that $a \neq 0$.

$$\xrightarrow{(1/a)\rho_1} \left(\begin{array}{ccc|c} 1 & b/a & c/a & 0 \\ d & e & f & 0 \\ g & h & i & 0 \end{array}\right) \xrightarrow[-g\rho_1+\rho_3]{-d\rho_1+\rho_2} \left(\begin{array}{ccc|c} 1 & b/a & c/a & 0 \\ 0 & (ae-bd)/a & (af-cd)/a & 0 \\ 0 & (ah-bg)/a & (ai-cg)/a & 0 \end{array}\right)$$

$$\xrightarrow{(a/(ae-bd))\rho_2} \left(\begin{array}{ccc|c} 1 & b/a & c/a & 0 \\ 0 & 1 & (af-cd)/(ae-bd) & 0 \\ 0 & (ah-bg)/a & (ai-cg)/a & 0 \end{array}\right)$$

(where we've assumed for the moment that $ae - bd \neq 0$ in order to do the row reduction step). Then, under the assumptions, we get this.

$$\xrightarrow{((ah-bg)/a)\rho_2+\rho_3} \left(\begin{array}{ccc|c} 1 & \frac{b}{a} & \frac{c}{a} & 0 \\ 0 & 1 & \frac{af-cd}{ae-bd} & 0 \\ 0 & 0 & \frac{aei+bgf+cdh-hfa-idb-gec}{ae-bd} & 0 \end{array}\right)$$

shows that the original system is nonsingular if and only if the $3,3$ entry is nonzero. This fraction is defined because of the $ae - bd \neq 0$ assumption, and it will equal zero if and only if its numerator equals zero.

We next worry about the assumptions. First, if $a \neq 0$ but $ae - bd = 0$ then we swap

$$\left(\begin{array}{ccc|c} 1 & b/a & c/a & 0 \\ 0 & 0 & (af-cd)/a & 0 \\ 0 & (ah-bg)/a & (ai-cg)/a & 0 \end{array}\right) \xrightarrow{\rho_2\leftrightarrow\rho_3} \left(\begin{array}{ccc|c} 1 & b/a & c/a & 0 \\ 0 & (ah-bg)/a & (ai-cg)/a & 0 \\ 0 & 0 & (af-cd)/a & 0 \end{array}\right)$$

and conclude that the system is nonsingular if and only if either $ah - bg = 0$ or $af - cd = 0$. That's the same as asking that their product be zero:

$$\begin{aligned} ahaf - ahcd - bgaf + bgcd &= 0 \\ ahaf - ahcd - bgaf + aegc &= 0 \\ a(haf - hcd - bgf + egc) &= 0 \end{aligned}$$

(in going from the first line to the second we've applied the case assumption that $ae - bd = 0$ by substituting ae for bd). Since we are assuming that $a \neq 0$, we have that $haf - hcd - bgf + egc = 0$. With $ae - bd = 0$ we can rewrite this to fit the form we need: in this $a \neq 0$ and $ae - bd = 0$ case, the given system is nonsingular when $haf - hcd - bgf + egc - i(ae - bd) = 0$, as required.

The remaining cases have the same character. Do the $a = 0$ but $d \neq 0$ case and the $a = 0$ and $d = 0$ but $g \neq 0$ case by first swapping rows and then going on as above. The $a = 0$, $d = 0$, and $g = 0$ case is easy — a set with a zero vector is linearly dependent, and the formula comes out to equal zero.

(c) It is linearly dependent if and only if either vector is a multiple of the other. That is, it is not independent iff

$$\begin{pmatrix} a \\ d \\ g \end{pmatrix} = r \cdot \begin{pmatrix} b \\ e \\ h \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} b \\ e \\ h \end{pmatrix} = s \cdot \begin{pmatrix} a \\ d \\ g \end{pmatrix}$$

(or both) for some scalars r and s. Eliminating r and s in order to restate this condition only in terms of the given letters a, b, d, e, g, h, we have that it is not independent — it is dependent — iff $ae - bd = ah - gb = dh - ge$.

(d) Dependence or independence is a function of the indices, so there is indeed a formula (although at first glance a person might think the formula involves cases: "if the first component of the first vector is zero then ... ", this guess turns out not to be correct).

2.II.1.40

(a) This check is routine.

(b) The summation is infinite (has infinitely many summands). The definition of linear combination involves only finite sums.

(c) No nontrivial finite sum of members of $\{g, f_0, f_1, \dots\}$ adds to the zero object: assume that

$$c_0 \cdot (1/(1-x)) + c_1 \cdot 1 + \cdots + c_n \cdot x^n = 0$$

(any finite sum uses a highest power, here n). Multiply both sides by $1-x$ to conclude that each coefficient is zero, because a polynomial describes the zero function only when it is the zero polynomial.

2.II.1.41 It is both 'if' and 'only if'.

Let T be a subset of the subspace S of the vector space V. The assertion that any linear relationship $c_1\vec{t}_1 + \cdots + c_n\vec{t}_n = \vec{0}$ among members of T must be the trivial relationship $c_1 = 0, \dots, c_n = 0$ is a statement that holds in S if and only if it holds in V, because the subspace S inherits its addition and scalar multiplication operations from V.

. . . .

Answers for subsection 2.III.1

2.III.1.18 A natural basis is $\langle 1, x, x^2 \rangle$. There are bases for $\mathcal{P}_2$ that do not contain any polynomials of degree one or degree zero. One is $\langle 1 + x + x^2, x + x^2, x^2 \rangle$. (Every basis has at least one polynomial of degree two, though.)

2.III.1.19 The reduction

$$\left(\begin{array}{cccc|c} 1 & -4 & 3 & -1 & 0 \\ 2 & -8 & 6 & -2 & 0 \end{array}\right) \xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cccc|c} 1 & -4 & 3 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

gives that the only condition is that $x_1 = 4x_2 - 3x_3 + x_4$. The solution set is

$$\{\begin{pmatrix} 4x_2 - 3x_3 + x_4 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \bigm| x_2, x_3, x_4 \in \mathbb{R}\} = \{x_2 \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \bigm| x_2, x_3, x_4 \in \mathbb{R}\}$$

and so the obvious candidate for the basis is this.

$$\langle \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle$$

We've shown that this spans the space, and showing it is also linearly independent is routine.

2.III.1.22 We will show that the second is a basis; the first is similar. We will show this straight from the definition of a basis, because this example appears before Theorem 2.III.1.12.

To see that it is linearly independent, we set up $c_1 \cdot (\cos\theta - \sin\theta) + c_2 \cdot (2\cos\theta + 3\sin\theta) = 0\cos\theta + 0\sin\theta$. Taking $\theta = 0$ and $\theta = \pi/2$ gives this system

$$\begin{array}{rl} c_1 \cdot 1 + c_2 \cdot 2 = 0 & \xrightarrow{\rho_1 + \rho_2} \quad c_1 + 2c_2 = 0 \\ c_1 \cdot (-1) + c_2 \cdot 3 = 0 & \qquad\qquad\quad +5c_2 = 0 \end{array}$$

which shows that $c_1 = 0$ and $c_2 = 0$.

The calculation for span is also easy; for any $x, y \in \mathbb{R}^4$, we have that $c_1 \cdot (\cos\theta - \sin\theta) + c_2 \cdot (2\cos\theta + 3\sin\theta) = x\cos\theta + y\sin\theta$ gives that $c_2 = x/5 + y/5$ and that $c_1 = 3x/5 - 2y/5$, and so the span is the entire space.

2.III.1.25 Yes. Linear independence and span are unchanged by reordering.

2.III.1.26 No linearly independent set contains a zero vector.

2.III.1.28 Each forms a linearly independent set if $\vec{v}$ is ommitted. To preserve linear independence, we must expand the span of each. That is, we must determine the span of each (leaving $\vec{v}$ out), and then pick a $\vec{v}$ lying outside of that span. Then to finish, we must check that the result spans the entire given space. Those checks are routine.

(a) Any vector that is not a multiple of the given one, that is, any vector that is not on the line $y = x$ will do here. One is $\vec{v} = \vec{e}_1$.

(b) By inspection, we notice that the vector $\vec{e}_3$ is not in the span of the set of the two given vectors. The check that the resulting set is a basis for $\mathbb{R}^3$ is routine.

(c) For any member of the span $\{c_1 \cdot (x) + c_2 \cdot (1 + x^2) \bigm| c_1, c_2 \in \mathbb{R}\}$, the coefficient of x^2 equals the constant term. So we expand the span if we add a quadratic without this property, say, $\vec{v} = 1 - x^2$. The check that the result is a basis for $\mathcal{P}_2$ is easy.

2.III.1.30 No; no linearly independent set contains the zero vector.

2.III.1.31 Here is a subset of $\mathbb{R}^2$ that is not a basis, and two different linear combinations of its elements that sum to the same vector.

$$\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}\} \qquad 2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 0 \cdot \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Subsets that are not bases can possibly have unique linear combinations. Linear combinations are unique if and only if the subset is linearly independent. That is established in the proof of the theorem.

2.III.1.34 We have (using these peculiar operations with care)

$$\{\begin{pmatrix} 1 - y - z \\ y \\ z \end{pmatrix} \bigm| y, z \in \mathbb{R}\} = \{\begin{pmatrix} -y + 1 \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} -z + 1 \\ 0 \\ z \end{pmatrix} \bigm| y, z \in \mathbb{R}\} = \{y \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \bigm| y, z \in \mathbb{R}\}$$

and so a natural candidate for a basis is this.

$$\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rangle$$

To check linear independence we set up

$$c_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

(the vector on the right is the zero object in this space). That yields the linear system

$$\begin{array}{rl} (-c_1+1)+(-c_2+1)-1 &= 1 \\ c_1 \qquad\qquad\qquad &= 0 \\ c_2 &= 0 \end{array}$$

with only the solution $c_1 = 0$ and $c_2 = 0$. Checking the span is similar.

. . . .

Answers for subsection 2.III.2

2.III.2.15 The solution set is

$$\{\begin{pmatrix} 4x_2 - 3x_3 + x_4 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mid x_2, x_3, x_4 \in \mathbb{R}\}$$

so a natural basis is this

$$\langle \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \rangle$$

(checking linear independence is easy). Thus the dimension is three.

2.III.2.17

(a) As in the prior exercise, the space $\mathcal{M}_{2\times 2}$ of matrices without restriction has this basis

$$\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \rangle$$

and so the dimension is four.

(b) For this space

$$\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a = b - 2c \text{ and } d \in \mathbb{R}\} = \{b \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + c \cdot \begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix} + d \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mid b, c, d \in \mathbb{R}\}$$

this is a natural basis.

$$\langle \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \rangle$$

The dimension is three.

(c) Gauss' method applied to the two-equation linear system gives that $c = 0$ and that $a = -b$. Thus, we have this description

$$\{\begin{pmatrix} -b & b \\ 0 & d \end{pmatrix} \mid b, d \in \mathbb{R}\} = \{b \cdot \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} + d \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mid b, d \in \mathbb{R}\}$$

and so this is a natural basis.

$$\langle \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \rangle$$

The dimension is two.

2.III.2.19 First recall that $\cos 2\theta = \cos^2\theta - \sin^2\theta$, and so deletion of $\cos 2\theta$ from this set leaves the span unchanged. What's left, the set $\{\cos^2\theta, \sin^2\theta, \sin 2\theta\}$, is linearly independent (consider the relationship $c_1\cos^2\theta + c_2\sin^2\theta + c_3\sin 2\theta = Z(\theta)$ where Z is the zero function, and then take $\theta = 0$, $\theta = \pi/4$, and $\theta = \pi/2$ to conclude that each c is zero). It is therefore a basis for its span. That shows that the span is a dimension three vector space.

2.III.2.20 Here is a basis

$$\langle (1+0i, 0+0i, \ldots, 0+0i),\ (0+1i, 0+0i, \ldots, 0+0i), (0+0i, 1+0i, \ldots, 0+0i), \ldots \rangle$$

and so the dimension is $2 \cdot 47 = 94$.

2.III.2.21 A basis is

$$\langle \begin{pmatrix} 1&0&0&0&0 \\ 0&0&0&0&0 \\ 0&0&0&0&0 \end{pmatrix}, \begin{pmatrix} 0&1&0&0&0 \\ 0&0&0&0&0 \\ 0&0&0&0&0 \end{pmatrix}, \ldots, \begin{pmatrix} 0&0&0&0&0 \\ 0&0&0&0&0 \\ 0&0&0&0&1 \end{pmatrix} \rangle$$

and thus the dimension is $3 \cdot 5 = 15$.

2.III.2.23

(a) The diagram for $\mathcal{P}_2$ has four levels. The top level has the only three-dimensional subspace, $\mathcal{P}_2$ itself. The next level contains the two-dimensional subspaces (*not* just the linear polynomials; any two-dimensional subspace, like those polynomials of the form $ax^2 + b$). Below that are the one-dimensional subspaces. Finally, of course, is the only zero-dimensional subspace, the trivial subspace.

(b) For $\mathcal{M}_{2\times2}$, the diagram has five levels, including subspaces of dimension four through zero.

2.III.2.25 We need only produce an infinite linearly independent set. One is $\langle f_1, f_2, \ldots \rangle$ where $f_i\colon \mathbb{R} \to \mathbb{R}$ is

$$f_i(x) = \begin{cases} 1 & \text{if } x = i \\ 0 & \text{otherwise} \end{cases}$$

the function that has value 1 only at $x = i$.

2.III.2.26 Considering a function to be a set, specifically, a set of ordered pairs $(x, f(x))$, then the only function with an empty domain is the empty set. Thus this is a trivial vector space, and has dimension zero.

2.III.2.27 Apply Corollary 2.8.

2.III.2.28 The first chapter defines a plane — a '2-flat' — to be a set of the form $\{\vec{p} + t_1\vec{v}_1 + t_2\vec{v}_2 \mid t_1, t_2 \in \mathbb{R}\}$ (also there is a discussion of why this is equivalent to the description often taken in Calculus as the set of points (x, y, z) subject to some linear condition $ax + by + cz = d$). When the plane passes throught the origin we can take the particular vector $\vec{p}$ to be $\vec{0}$. Thus, in the language we have developed in this chapter, a plane through the origin is the span of a set of two vectors.

Now for the statement. Asserting that the three are not coplanar is the same as asserting that no vector lies in the span of the other two — no vector is a linear combination of the other two. That's simply an assertion that the three-element set is linearly independent. By Corollary 2.12, that's equivalent to an assertion that the set is a basis for $\mathbb{R}^3$.

2.III.2.29 Let the space V be finite dimensional. Let S be a subspace of V.

(a) The empty set is a linearly independent subset of S. By Corollary 2.10, it can be expanded to a basis for the vector space S.

(b) Any basis for the subspace S is a linearly independent set in the superspace V. Hence it can be expanded to a basis for the superspace, which is finite dimensional. Therefore it has only finitely many members.

2.III.2.30 It ensures that we exhaust the $\vec{\beta}$'s. That is, it justifies the first sentence of the last paragraph.

2.III.2.32 First, note that a set is a basis for some space if and only if it is linearly independent, because in that case it is a basis for its own span.

(a) The answer to the question in the second paragraph is "yes" (implying "yes" answers for both questions in the first paragraph). If B_U is a basis for U then B_U is a linearly independent subset of W. Apply Corollary 2.10 to expand it to a basis for W. That is the desired B_W.

The answer to the question in the third paragraph is "no", which implies a "no" answer to the question of the fourth paragraph. Here is an example of a basis for a superspace with no sub-basis forming a basis for a subspace: in $W = \mathbb{R}^2$, consider the standard basis $\mathcal{E}_2$. No sub-basis of $\mathcal{E}_2$ forms a basis for the subspace U of $\mathbb{R}^2$ that is the line $y = x$.

(b) It is a basis (for its span) because the intersection of linearly independent sets is linearly independent (the intersection is a subset of each of the linearly independent sets).

It is not, however, a basis for the intersection of the spaces. For instance, these are bases for $\mathbb{R}^2$:

$$B_1 = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle \quad \text{and} \quad B_2 = \langle \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \rangle$$

and $\mathbb{R}^2 \cap \mathbb{R}^2 = \mathbb{R}^2$, but $B_1 \cap B_2$ is empty. All we can say is that the intersection of the bases is a basis for a subset of the intersection of the spaces.

(c) The union of bases need not be a basis: in $\mathbb{R}^2$

$$B_1 = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle \quad \text{and} \quad B_2 = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \rangle$$

have a union $B_1 \cup B_2$ that is not linearly independent. A necessary and sufficient condition for a union of two bases to be a basis

$$B_1 \cup B_2 \text{ is linearly independent} \quad \iff \quad [B_1 \cap B_2] = [B_1] \cap [B_2]$$

it is easy enough to prove (but perhaps hard to apply).

(d) The complement of a basis cannot be a basis because it contains the zero vector.

2.III.2.34 The possibilities for the dimension of V are 0, 1, $n-1$, and n.

To see this, first consider the case when all the coordinates of $\vec{v}$ are equal.

$$\vec{v} = \begin{pmatrix} z \\ z \\ \vdots \\ z \end{pmatrix}$$

Then $\sigma(\vec{v}) = \vec{v}$ for every permutation σ, so V is just the span of $\vec{v}$, which has dimension 0 or 1 according to whether $\vec{v}$ is $\vec{0}$ or not.

Now suppose not all the coordinates of $\vec{v}$ are equal; let x and y with $x \neq y$ be among the coordinates of $\vec{v}$. Then we can find permutations σ_1 and σ_2 such that

$$\sigma_1(\vec{v}) = \begin{pmatrix} x \\ y \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \quad \text{and} \quad \sigma_2(\vec{v}) = \begin{pmatrix} y \\ x \\ a_3 \\ vdots \\ a_n \end{pmatrix}$$

for some $a_3, \ldots, a_n \in \mathbb{R}$. Therefore,

$$\frac{1}{y-x}\big(\sigma_1(\vec{v}) - \sigma_2(\vec{v})\big) = \begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

is in V. That is, $\vec{e}_2 - \vec{e}_1 \in V$, where $\vec{e}_1$, $\vec{e}_2$, ..., $\vec{e}_n$ is the standard basis for $\mathbb{R}^n$. Similarly, $\vec{e}_3 - \vec{e}_2$, ..., $\vec{e}_n - \vec{e}_1$ are all in V. It is easy to see that the vectors $\vec{e}_2 - \vec{e}_1$, $\vec{e}_3 - \vec{e}_2$, ..., $\vec{e}_n - \vec{e}_1$ are linearly independent (that is, form a linearly independent set), so $\dim V \geq n - 1$.

Finally, we can write

$$\begin{aligned} \vec{v} &= x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n \\ &= (x_1 + x_2 + \cdots + x_n)\vec{e}_1 + x_2(\vec{e}_2 - \vec{e}_1) + \cdots + x_n(\vec{e}_n - \vec{e}_1) \end{aligned}$$

This shows that if $x_1 + x_2 + \cdots + x_n = 0$ then $\vec{v}$ is in the span of $\vec{e}_2 - \vec{e}_1$, ..., $\vec{e}_n - \vec{e}_1$ (that is, is in the span of the set of those vectors); similarly, each $\sigma(\vec{v})$ will be in this span, so V will equal this span and $\dim V = n - 1$. On the other hand, if $x_1 + x_2 + \cdots + x_n \neq 0$ then the above equation shows that $\vec{e}_1 \in V$ and thus $\vec{e}_1, \ldots, \vec{e}_n \in V$, so $V = \mathbb{R}^n$ and $\dim V = n$.

. . . .

Answers for subsection 2.III.3

2.III.3.16

(a) $\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & 6 \\ 4 & 7 \\ 3 & 8 \end{pmatrix}$ **(d)** $\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ **(e)** $\begin{pmatrix} -1 \\ -2 \end{pmatrix}$

2.III.3.22 Only the zero matrices have rank of zero. The only matrices of rank one have the form

$$\begin{pmatrix} k_1 \cdot \rho \\ \vdots \\ k_m \cdot \rho \end{pmatrix}$$

where ρ is some nonzero row vector, and not all of the k_i's are zero. (*Remark.* We can't simply say that all of the rows are multiples of the first because the first row might be the zero row. *Another Remark.* The above also applies with 'column' replacing 'row'.)

2.III.3.24 The column rank is two. One way to see this is by inspection — the column space consists of two-tall columns and so can have a dimension of at least two, and we can easily find two columns that together form a linearly independent set (the fourth and fifth columns, for instance). Another way to see this is to recall that the column rank equals the row rank, and to perform Gauss' method, which leaves two nonzero rows.

2.III.3.25 We apply Theorem 2.III.3.13. The number of columns of a matrix of coefficients A of a linear system equals the number n of unknowns. A linear system with at least one solution has at most one solution if and only if the space of solutions of the associated homogeneous system has dimension zero (recall: in the 'General = Particular + Homogeneous' equation $\vec{v} = \vec{p} + \vec{h}$, provided that such a $\vec{p}$ exists, the solution $\vec{v}$ is unique if and only if the vector $\vec{h}$ is unique, namely $\vec{h} = \vec{0}$). But that means, by the theorem, that $n = r$.

2.III.3.27 There is little danger of their being equal since the row space is a set of row vectors while the column space is a set of columns (unless the matrix is 1×1, in which case the two spaces must be equal).

Remark. Consider

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$

and note that the row space is the set of all multiples of $\begin{pmatrix} 1 & 3 \end{pmatrix}$ while the column space consists of multiples of

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

so we also cannot argue that the two spaces must be simply transposes of each other.

2.III.3.28 First, the vector space is the set of four-tuples of real numbers, under the natural operations. Although this is not the set of four-wide row vectors, the difference is slight — it is "the same" as that set. So we will treat the four-tuples like four-wide vectors.

With that, one way to see that $(1, 0, 1, 0)$ is not in the span of the first set is to note that this reduction

$$\begin{pmatrix} 1 & -1 & 2 & -3 \\ 1 & 1 & 2 & 0 \\ 3 & -1 & 6 & -6 \end{pmatrix} \xrightarrow[-3\rho_1+\rho_3]{-\rho_1+\rho_2} \xrightarrow{-\rho_2+\rho_3} \begin{pmatrix} 1 & -1 & 2 & -3 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and this one

$$\begin{pmatrix} 1 & -1 & 2 & -3 \\ 1 & 1 & 2 & 0 \\ 3 & -1 & 6 & -6 \\ 1 & 0 & 1 & 0 \end{pmatrix} \xrightarrow[\substack{-3\rho_1+\rho_3 \\ -\rho_1+\rho_4}]{-\rho_1+\rho_2} \xrightarrow[-(1/2)\rho_2+\rho_4]{-\rho_2+\rho_3} \xrightarrow{\rho_3 \leftrightarrow \rho_4} \begin{pmatrix} 1 & -1 & 2 & -3 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & -1 & 3/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

yield matrices differing in rank. This means that addition of $(1, 0, 1, 0)$ to the set of the first three four-tuples increases the rank, and hence the span, of that set. Therefore $(1, 0, 1, 0)$ is not already in the span.

2.III.3.30 This can be done as a straightforward calculation.

$$\begin{aligned}
(rA+sB)^{\text{trans}} &= \begin{pmatrix} ra_{1,1}+sb_{1,1} & \ldots & ra_{1,n}+sb_{1,n} \\ \vdots & & \vdots \\ ra_{m,1}+sb_{m,1} & \ldots & ra_{m,n}+sb_{m,n} \end{pmatrix}^{\text{trans}} \\
&= \begin{pmatrix} ra_{1,1}+sb_{1,1} & \ldots & ra_{m,1}+sb_{m,1} \\ \vdots & & \\ ra_{1,n}+sb_{1,n} & \ldots & ra_{m,n}+sb_{m,n} \end{pmatrix} \\
&= \begin{pmatrix} ra_{1,1} & \ldots & ra_{m,1} \\ \vdots & & \\ ra_{1,n} & \ldots & ra_{m,n} \end{pmatrix} + \begin{pmatrix} sb_{1,1} & \ldots & sb_{m,1} \\ \vdots & & \\ sb_{1,n} & \ldots & sb_{m,n} \end{pmatrix} \\
&= rA^{\text{trans}} + sB^{\text{trans}}
\end{aligned}$$

2.III.3.32 It cannot be bigger.

2.III.3.33 The number of rows in a maximal linearly independent set cannot exceed the number of rows. A better bound (the bound that is, in general, the best possible) is the minimum of m and n, because the row rank equals the column rank.

2.III.3.35 False. The first is a set of columns while the second is a set of rows.

This example, however,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad A^{\text{trans}} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

indicates that as soon as we have a formal meaning for "the same", we can apply it here:

$$\text{Columnspace}(A) = [\{\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix}\}]$$

while

$$\text{Rowspace}(A^{\text{trans}}) = [\{\begin{pmatrix} 1 & 4 \end{pmatrix}, \begin{pmatrix} 2 & 5 \end{pmatrix}, \begin{pmatrix} 3 & 6 \end{pmatrix}\}]$$

are "the same" as each other.

2.III.3.37 A linear system

$$c_1\vec{a}_1 + \cdots + c_n\vec{a}_n = \vec{d}$$

has a solution if and only if $\vec{d}$ is in the span of the set $\{\vec{a}_1, \ldots, \vec{a}_n\}$. That's true if and only if the column rank of the augmented matrix equals the column rank of the matrix of coefficients. Since rank equals the column rank, the system has a solution if and only if the rank of its augmented matrix equals the rank of its matrix of coefficients.

2.III.3.38

(a) Row rank equals column rank so each is at most the minimum of the number of rows and columns. Hence both can be full only if the number of rows equals the number of columns. (Of course, the converse does not hold: a square matrix need not have full row rank or full column rank.)

(b) If A has full row rank then, no matter what the right-hand side, Gauss' method on the augmented matrix ends with a leading one in each row and none of those leading ones in the furthest right column (the "augmenting" column). Back substitution then gives a solution.

On the other hand, if the linear system lacks a solution for some right-hand side it can only be because Gauss' method leaves some row so that it is all zeroes to the left of the "augmenting" bar and has a nonzero entry on the right. Thus, if A does not have a solution for some right-hand sides, then A does not have full row rank because some of its rows have been eliminated.

(c) The matrix A has full column rank if and only if its columns form a linearly independent set. That's equivalent to the existence of only the trivial linear relationship.

(d) The matrix A has full column rank if and only if the set of its columns is linearly independent set, and so forms a basis for its span. That's equivalent to the existence of a unique linear representation of all vectors in that span.

2.III.3.39 Instead of the row spaces being the same, the row space of B would be a subspace (possibly equal to) the row space of A.

. . . .

Answers for subsection 2.III.4

2.III.4.22 It is. Showing that these two are subspaces is routine. To see that the space is the direct sum of these two, just note that each member of $\mathcal{P}_2$ has the unique decomposition $m+nx+px^2 = (m+px^2)+(nx)$.

2.III.4.24 Each of these is $\mathbb{R}^3$.

(a) These are broken into lines for legibility.

$$\begin{aligned}&W_1+W_2+W_3,\ W_1+W_2+W_3+W_4,\ W_1+W_2+W_3+W_5,\ W_1+W_2+W_3+W_4+W_5,\\ &\quad W_1+W_2+W_4,\ W_1+W_2+W_4+W_5,\ W_1+W_2+W_5,\\ &W_1+W_3+W_4,\ W_1+W_3+W_5,\ W_1+W_3+W_4+W_5,\\ &W_1+W_4,\ W_1+W_4+W_5,\\ &W_1+W_5,\\ &W_2+W_3+W_4,\ W_2+W_3+W_4+W_5,\\ &W_2+W_4,\ W_2+W_4+W_5,\\ &W_3+W_4,\ W_3+W_4+W_5,\\ &W_4+W_5\end{aligned}$$

(b) $W_1 \oplus W_2 \oplus W_3$, $W_1 \oplus W_4$, $W_1 \oplus W_5$, $W_2 \oplus W_4$, $W_3 \oplus W_4$

2.III.4.26 It is W_2.

2.III.4.27 True by Lemma 4.8.

2.III.4.28 Two distinct direct sum decompositions of $\mathbb{R}^4$ are easy to find. Two such are $W_1 = [\{\vec{e}_1, \vec{e}_2\}]$ and $W_2 = [\{\vec{e}_3, \vec{e}_4\}]$, and also $U_1 = [\{\vec{e}_1\}]$ and $U_2 = [\{\vec{e}_2, \vec{e}_3, \vec{e}_4\}]$. (Many more are possible, for example $\mathbb{R}^4$ and its trivial subspace.)

In contrast, any partition of $\mathbb{R}^1$'s single-vector basis will give one basis with no elements and another with a single element. Thus any decomposition involves $\mathbb{R}^1$ and its trivial subspace.

2.III.4.29 Set inclusion one way is easy: $\{\vec{w}_1 + \cdots + \vec{w}_k \bigm| \vec{w}_i \in W_i\}$ is a subset of $[W_1 \cup \ldots \cup W_k]$ because each $\vec{w}_1 + \cdots + \vec{w}_k$ is a sum of vectors from the union.

For the other inclusion, to any linear combination of vectors from the union apply commutativity of vector addition to put vectors from W_1 first, followed by vectors from W_2, etc. Add the vectors from W_1 to get a $\vec{w}_1 \in W_1$, add the vectors from W_2 to get a $\vec{w}_2 \in W_2$, etc. The result has the desired form.

2.III.4.30 One example is to take the space to be $\mathbb{R}^3$, and to take the subspaces to be the xy-plane, the xz-plane, and the yz-plane.

2.III.4.32 It can contain a trivial subspace; this set of subspaces of $\mathbb{R}^3$ is independent: $\{\{\vec{0}\}, x\text{-axis}\}$. No nonzero vector from the trivial space $\{\vec{0}\}$ is a multiple of a vector from the x-axis, simply because the trivial space has no nonzero vectors to be candidates for such a multiple (and also no nonzero vector from the x-axis is a multiple of the zero vector from the trivial subspace).

2.III.4.35

(a) The intersection and sum are

$$\{\begin{pmatrix}0&0\\c&0\end{pmatrix} \bigm| c \in \mathbb{R}\} \qquad \{\begin{pmatrix}0&b\\c&d\end{pmatrix} \bigm| b,c,d \in \mathbb{R}\}$$

which have dimensions one and three.

(b) We write $B_{U\cap W}$ for the basis for $U\cap W$, we write B_U for the basis for U, we write B_W for the basis for W, and we write B_{U+W} for the basis under consideration.

To see that that B_{U+W} spans $U+W$, observe that any vector $c\vec{u}+d\vec{w}$ from $U+W$ can be written as a linear combination of the vectors in B_{U+W}, simply by expressing $\vec{u}$ in terms of B_U and expressing $\vec{w}$ in terms of B_W.

We finish by showing that B_{U+W} is linearly independent. Consider

$$c_1\vec{\mu}_1+\cdots+c_{j+1}\vec{\beta}_1+\cdots+c_{j+k+p}\vec{\omega}_p=\vec{0}$$

which can be rewritten in this way.

$$c_1\vec{\mu}_1+\cdots+c_j\vec{\mu}_j=-c_{j+1}\vec{\beta}_1-\cdots-c_{j+k+p}\vec{\omega}_p$$

Note that the left side sums to a vector in U while right side sums to a vector in W, and thus both sides sum to a member of $U\cap W$. Since the left side is a member of $U\cap W$, it is expressible in terms of the members of $B_{U\cap W}$, which gives the combination of $\vec{\mu}$'s from the left side above as equal to a combination of $\vec{\beta}$'s. But, the fact that the basis B_U is linearly independent shows that any such combination is trivial, and in particular, the coefficients c_1, ... , c_j from the left side above are all zero. Similarly, the coefficients of the $\vec{\omega}$'s are all zero. This leaves the above equation as a linear relationship among the $\vec{\beta}$'s, but $B_{U\cap W}$ is linearly independent, and therefore all of the coefficients of the $\vec{\beta}$'s are also zero.

(c) Just count the basis vectors in the prior item: $\dim(U+W)=j+k+p$, and $\dim(U)=j+k$, and $\dim(W)=k+p$, and $\dim(U\cap W)=k$.

(d) We know that $\dim(W_1+W_2)=\dim(W_1)+\dim(W_2)-\dim(W_1\cap W_2)$. Because $W_1\subseteq W_1+W_2$, we know that W_1+W_2 must have dimension greater than that of W_1, that is, must have dimension eight, nine, or ten. Substituting gives us three possibilities $8=8+8-\dim(W_1\cap W_2)$ or $9=8+8-\dim(W_1\cap W_2)$ or $10=8+8-\dim(W_1\cap W_2)$. Thus $\dim(W_1\cap W_2)$ must be either eight, seven, or six. (Giving examples to show that each of these three cases is possible is easy, for instance in $\mathbb{R}^{10}$.)

2.III.4.36 Expand each S_i to a basis B_i for W_i. The concatenation of those bases $B_1{}^\frown\cdots{}^\frown B_k$ is a basis for V and thus its members form a linearly independent set. But the union $S_1\cup\cdots\cup S_k$ is a subset of that linearly independent set, and thus is itself linearly independent.

2.III.4.37

(a) Two such are these.

$$\begin{pmatrix}1&2\\2&3\end{pmatrix}\qquad\begin{pmatrix}0&1\\-1&0\end{pmatrix}$$

For the antisymmetric one, entries on the diagonal must be zero.

(b) A square symmetric matrix equals its transpose. A square antisymmetric matrix equals the negative of its transpose.

(c) Showing that the two sets are subspaces is easy. Suppose that $A\in\mathcal{M}_{n\times n}$. To express A as a sum of a symmetric and an antisymmetric matrix, we observe that

$$A=(1/2)(A+A^{\text{trans}})+(1/2)(A-A^{\text{trans}})$$

and note the first summand is symmetric while the second is antisymmetric. Thus $\mathcal{M}_{n\times n}$ is the sum of the two subspaces. To show that the sum is direct, assume a matrix A is both symmetric $A=A^{\text{trans}}$ and antisymmetric $A=-A^{\text{trans}}$. Then $A=-A$ and so all of A's entries are zeroes.

2.III.4.38 Assume that $\vec{v}\in(W_1\cap W_2)+(W_1\cap W_3)$. Then $\vec{v}=\vec{w}_2+\vec{w}_3$ where $\vec{w}_2\in W_1\cap W_2$ and $\vec{w}_3\in W_1\cap W_3$. Note that $\vec{w}_2,\vec{w}_3\in W_1$ and, as a subspace is closed under addition, $\vec{w}_2+\vec{w}_3\in W_1$. Thus $\vec{v}=\vec{w}_2+\vec{w}_3\in W_1\cap(W_2+W_3)$.

This example proves that the inclusion may be strict: in $\mathbb{R}^2$ take W_1 to be the x-axis, take W_2 to be the y-axis, and take W_3 to be the line $y=x$. Then $W_1\cap W_2$ and $W_1\cap W_3$ are trivial and so their sum is trivial. But W_2+W_3 is all of $\mathbb{R}^2$ so $W_1\cap(W_2+W_3)$ is the x-axis.

2.III.4.39 It happens when at least one of W_1, W_2 is trivial. But that is the only way it can happen.

To prove this, assume that both are non-trivial, select nonzero vectors $\vec{w}_1,\vec{w}_2$ from each, and consider $\vec{w}_1+\vec{w}_2$. This sum is not in W_1 because $\vec{w}_1+\vec{w}_2=\vec{v}\in W_1$ would imply that $\vec{w}_2=\vec{v}-\vec{w}_1$ is in W_1, which violates the assumption of the independence of the subspaces. Similarly, $\vec{w}_1+\vec{w}_2$ is not in W_2. Thus there is an element of V that is not in $W_1\cup W_2$.

2.III.4.42 No. The standard basis for $\mathbb{R}^2$ does not split into bases for the complementary subspaces the line $x = y$ and the line $x = -y$.

2.III.4.43

(a) Yes, $W_1 + W_2 = W_2 + W_1$ for all subspaces W_1, W_2 because each side is the span of $W_1 \cup W_2 = W_2 \cup W_1$.

(b) This one is similar to the prior one — each side of that equation is the span of $(W_1 \cup W_2) \cup W_3 = W_1 \cup (W_2 \cup W_3)$.

(c) Because this is an equality between sets, we can show that it holds by mutual inclusion. Clearly $W \subseteq W + W$. For $W + W \subseteq W$ just recall that every subset is closed under addition so any sum of the form $\vec{w}_1 + \vec{w}_2$ is in W.

(d) In each vector space, the identity element with respect to subspace addition is the trivial subspace.

(e) Neither of left or right cancelation needs to hold. For an example, in $\mathbb{R}^3$ take W_1 to be the xy-plane, take W_2 to be the x-axis, and take W_3 to be the y-axis.

2.III.4.44

(a) They are equal because for each, V is the direct sum if and only if each $\vec{v} \in V$ can be written in a unique way as a sum $\vec{v} = \vec{w}_1 + \vec{w}_2$ and $\vec{v} = \vec{w}_2 + \vec{w}_1$.

(b) They are equal because for each, V is the direct sum if and only if each $\vec{v} \in V$ can be written in a unique way as a sum of a vector from each $\vec{v} = (\vec{w}_1 + \vec{w}_2) + \vec{w}_3$ and $\vec{v} = \vec{w}_1 + (\vec{w}_2 + \vec{w}_3)$.

(c) Any vector in $\mathbb{R}^3$ can be decomposed uniquely into the sum of a vector from each axis.

(d) No. For an example, in $\mathbb{R}^2$ take W_1 to be the x-axis, take W_2 to be the y-axis, and take W_3 to be the line $y = x$.

(e) In any vector space the trivial subspace acts as the identity element with respect to direct sum.

(f) In any vector space, only the trivial subspace has a direct-sum inverse (namely, itself). One way to see this is that dimensions add, and so increase.

. . . .

Answers for Topic: Fields

1 These checks are all routine; most consist only of remarking that property is so familiar that it does not need to be proved.

2 For both of these structures, these checks are all routine. As with the prior question, most of the checks consist only of remarking that property is so familiar that it does not need to be proved.

3 There is no multiplicative inverse for 2, so the integers do not satisfy condition (5).

4 These checks can be done by listing all of the possibilities. For instance, to verify the commutativity of addition, that $a + b = b + a$, we can easily check it for all possible pairs a, b, because there are only four such pairs. Similarly, for associativity, there are only eight triples a, b, c, and so the check is not too long. (There are other ways to do the checks, in particular, a reader may recognize these operations as arithmetic 'mod 2'.)

5 These will do.

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

As in the prior item, the check that they satisfy the conditions can be done by listing all of the cases, although this way of checking is somewhat long (making use of commutativity is helpful in shortening the work).

. . . .

Answers for Topic: Crystals

1 Each fundamental unit is 3.34×10^{-10} cm, so there are about $0.1/(3.34 \times 10^{-10})$ such units. That gives 2.99×10^8, so there are something like $300,000,000$ (three hundred million) units.

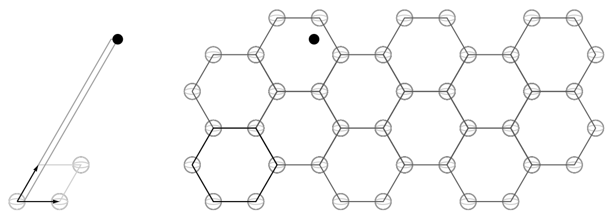
2

(a) We solve

$$c_1 \begin{pmatrix} 1.42 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1.23 \\ 0.71 \end{pmatrix} = \begin{pmatrix} 5.67 \\ 3.14 \end{pmatrix} \quad \Longrightarrow \quad \begin{aligned} 1.42c_1 + 1.23c_2 &= 5.67 \\ 0.71c_2 &= 3.14 \end{aligned}$$

to get $c_2 =\approx 4.42$ and $c_1 \approx 0.16$.

(b) Here is the point located in the lattice. In the picture on the left, superimposed on the unit cell are the two basis vectors $\vec{\beta}_1$ and $\vec{\beta}_2$, and a box showing the offset of $0.16\vec{\beta}_1 + 4.42\vec{\beta}_2$. The picture on the right shows where that appears inside of the crystal lattice, taking as the origin the lower left corner of the hexagon in the lower left.



So this point is in the next column of hexagons over, and either one hexagon up or two hexagons up, depending on how you count them.

(c) This second basis

$$\langle \begin{pmatrix} 1.42 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1.42 \end{pmatrix} \rangle$$

makes the computation easier

$$c_1 \begin{pmatrix} 1.42 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1.42 \end{pmatrix} = \begin{pmatrix} 5.67 \\ 3.14 \end{pmatrix} \quad \Longrightarrow \quad \begin{aligned} 1.42c_1 \qquad &= 5.67 \\ 1.42c_2 &= 3.14 \end{aligned}$$

(we get $c_2 \approx 2.21$ and $c_1 \approx 3.99$), but it doesn't seem to have to do much with the physical structure that we are studying.

3 In terms of the basis the locations of the corner atoms are $(0,0,0)$, $(1,0,0)$, ..., $(1,1,1)$. The locations of the face atoms are $(0.5,0.5,1)$, $(1,0.5,0.5)$, $(0.5,1,0.5)$, $(0,0.5,0.5)$, $(0.5,0,0.5)$, and $(0.5,0.5,0)$. The locations of the atoms a quarter of the way down from the top are $(0.75,0.75,0.75)$ and $(0.25,0.25,0.25)$. The locations of the atoms a quarter of the way up from the bottom are $(0.75,0.25,0.25)$ and $(0.25,0.75,0.25)$. Converting to Ångstroms is easy.

4

(a) $195.08/6.02 \times 10^{23} = 3.239 \times 10^{-22}$

(b) 4

(c) $4 \cdot 3.239 \times 10^{-22} = 1.296 \times 10^{-21}$

(d) $1.296 \times 10^{-21}/21.45 = 6.042 \times 10^{-23}$ cubic centimeters

(e) 3.924×10^{-8} centimeters.

(f) $\langle \begin{pmatrix} 3.924 \times 10^{-8} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3.924 \times 10^{-8} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 3.924 \times 10^{-8} \end{pmatrix} \rangle$

. . . .

Answers for Topic: Voting Paradoxes

1 This is one example that yields a non-rational preference order for a single voter.

	character	*experience*	*policies*
Democrat	most	middle	least
Republican	middle	least	most
Third	least	most	middle

The Democrat is preferred to the Republican for character and experience. The Republican is preferred to the Third for character and policies. And, the Third is preferred to the Democrat for experience and policies.

2 First, compare the $D > R > T$ decomposition that was done out in the Topic with the decomposition of the opposite $T > R > D$ voter.

$$\begin{pmatrix}-1\\1\\1\end{pmatrix}=\frac{1}{3}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}+\frac{2}{3}\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix}+\frac{2}{3}\cdot\begin{pmatrix}-1\\0\\1\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}1\\-1\\-1\end{pmatrix}=d_1\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}+d_2\cdot\begin{pmatrix}-1\\1\\0\end{pmatrix}+d_3\cdot\begin{pmatrix}-1\\0\\1\end{pmatrix}$$

Obviously, the second is the negative of the first, and so $d_1 = -1/3$, $d_2 = -2/3$, and $d_3 = -2/3$. This principle holds for any pair of opposite voters, and so we need only do the computation for a voter from the second row, and a voter from the third row. For a positive spin voter in the second row,

$$\begin{array}{rl} c_1 - c_2 - c_3 = & 1 \\ c_1 + c_2 \qquad = & 1 \\ c_1 \qquad + c_3 = & -1 \end{array} \;\xrightarrow[-\rho_1+\rho_3]{-\rho_1+\rho_2}\;\xrightarrow{(-1/2)\rho_2+\rho_3}\; \begin{array}{rl} c_1 - \quad c_2 - \qquad c_3 = & 1 \\ 2c_2 + \qquad c_3 = & 0 \\ (3/2)c_3 = & -2 \end{array}$$

gives $c_3 = -4/3$, $c_2 = 2/3$, and $c_1 = 1/3$. For a positive spin voter in the third row,

$$\begin{array}{rl} c_1 - c_2 - c_3 = & 1 \\ c_1 + c_2 \qquad = & -1 \\ c_1 \qquad + c_3 = & 1 \end{array} \;\xrightarrow[-\rho_1+\rho_3]{-\rho_1+\rho_2}\;\xrightarrow{(-1/2)\rho_2+\rho_3}\; \begin{array}{rl} c_1 - \quad c_2 - \qquad c_3 = & 1 \\ 2c_2 + \qquad c_3 = & -2 \\ (3/2)c_3 = & 1 \end{array}$$

gives $c_3 = 2/3$, $c_2 = -4/3$, and $c_1 = 1/3$.

3 The mock election corresponds to the table on page 150 in the way shown in the first table, and after cancellation the result is the second table.

positive spin	*negative spin*
$D > R > T$ 5 voters	$T > R > D$ 2 voters
$R > T > D$ 8 voters	$D > T > R$ 4 voters
$T > D > R$ 8 voters	$R > D > T$ 2 voters

positive spin	*negative spin*
$D > R > T$ 3 voters	$T > R > D$ –
$R > T > D$ 4 voters	$D > T > R$ –
$T > D > R$ 6 voters	$R > D > T$ –

All three come from the same side of the table (the left), as the result from this Topic says must happen. Tallying the election can now proceed, using the cancelled numbers

$$3\cdot\begin{tikzcd}[column sep=small] & D \arrow[dl, "-1\text{ voter}"'] & \\ T \arrow[rr, bend right, "1\text{ voter}"'] & & R \arrow[ul, "1\text{ voter}"'] \end{tikzcd} +4\cdot\begin{tikzcd}[column sep=small] & D \arrow[dl, "1\text{ voter}"'] & \\ T \arrow[rr, bend right, "1\text{ voter}"'] & & R \arrow[ul, "-1\text{ voter}"'] \end{tikzcd} +6\cdot\begin{tikzcd}[column sep=small] & D \arrow[dl, "1\text{ voter}"'] & \\ T \arrow[rr, bend right, "-1\text{ voter}"'] & & R \arrow[ul, "1\text{ voter}"'] \end{tikzcd} = \begin{tikzcd}[column sep=small] & D \arrow[dl, "7\text{ voters}"'] & \\ T \arrow[rr, bend right, "1\text{ voter}"'] & & R \arrow[ul, "5\text{ voters}"'] \end{tikzcd}$$

to get the same outcome.

4

(a) The two can be rewritten as $-c \leq a - b$ and $-c \leq b - a$. Either $a - b$ or $b - a$ is nonpositive and so $-c \leq -|a - b|$, as required.

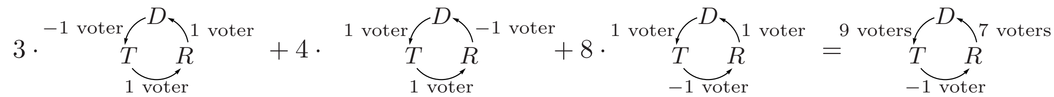
(b) This is immediate from the supposition that $0 \leq a + b - c$.

(c) A trivial example starts with the zero-voter election and adds any one voter. A more interesting example is to take the Political Science mock election and add two $T > D > R$ voters (they can be added one at a time, to satisfy the "addition of one more voter" criteria in the question). Observe that the additional voters have positive spin, which is the spin of the votes remaining after cancellation in the original mock election. This is the resulting table of voters, and next to it is the result of cancellation.

positive spin	*negative spin*
$D>R>T$ 5 voters	$T>R>D$ 2 voters
$R>T>D$ 8 voters	$D>T>R$ 4 voters
$T>D>R$ 10 voters	$R>D>T$ 2 voters

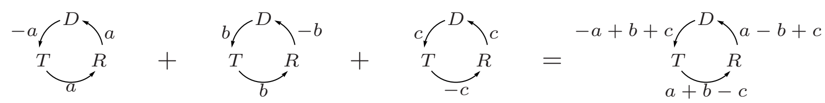
positive spin	*negative spin*
$D>R>T$ 3 voters	$T>R>D$ –
$R>T>D$ 4 voters	$D>T>R$ –
$T>D>R$ 8 voters	$R>D>T$ –

The election, using the cancelled numbers, is this.



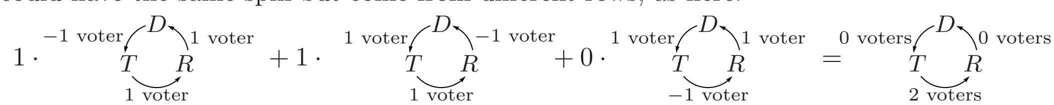
The majority cycle has indeed disappeared.

(d) One such condition is that, after cancellation, all three be nonnegative or all three be nonpositive, and: $|c|<|a+b|$ and $|b|<|a+c|$ and $|a|<|b+c|$. This follows from this diagram.

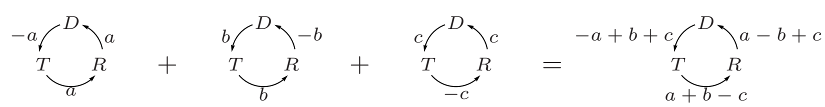


5

(a) A two-voter election can have a majority cycle in two ways. First, the two voters could be opposites, resulting after cancellation in the trivial election (with the majority cycle of all zeroes). Second, the two voters could have the same spin but come from different rows, as here.



(b) There are two cases. An even number of voters can split half and half into opposites, e.g., half the voters are $D>R>T$ and half are $T>R>D$. Then cancellation gives the trivial election. If the number of voters is greater than one and odd (of the form $2k+1$ with $k>0$) then using the cycle diagram from the proof,



we can take $a=k$ and $b=k$ and $c=1$. Because $k>0$, this is a majority cycle.

6 It is nonempty because it contains the zero vector. To see that it is closed under linear combinations of two of its members, suppose that $\vec{v}_1$ and $\vec{v}_2$ are in $U^\perp$ and consider $c_1\vec{v}_1+c_2\vec{v}_2$. For any $\vec{u}\in U$,

$$(c_1\vec{v}_1+c_2\vec{v}_2)\cdot\vec{u}=c_1(\vec{v}_1\cdot\vec{u})+c_2(\vec{v}_2\cdot\vec{u})=c_1\cdot 0+c_2\cdot 0=0$$

and so $c_1\vec{v}_1+c_2\vec{v}_2\in U^\perp$.

. . . .

Answers for Topic: Dimensional Analysis

1

(a) This relationship

$$(L^1M^0T^0)^{p_1}(L^1M^0T^0)^{p_2}(L^1M^0T^{-1})^{p_3}(L^0M^0T^0)^{p_4}(L^1M^0T^{-2})^{p_5}(L^0M^0T^1)^{p_6}=L^0M^0T^0$$

gives rise to this linear system

$$\begin{aligned} p_1+p_2+\ \ p_3\ +\ p_5\qquad &=0\\ 0&=0\\ -p_3\ -2p_5+p_6&=0 \end{aligned}$$

(note that there is no restriction on p_4). The natural paramatrization uses the free variables to give $p_3 = -2p_5 + p_6$ and $p_1 = -p_2 + p_5 - p_6$. The resulting description of the solution set

$$\{\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix} = p_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + p_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + p_5 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + p_6 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \bigm| p_2, p_4, p_5, p_6 \in \mathbb{R}\}$$

gives $\{y/x, \theta, xt/{v_0}^2, v_0t/x\}$ as a complete set of dimensionless products (recall that "complete" in this context does not mean that there are no other dimensionless products; it simply means that the set is a basis). This is, however, not the set of dimensionless products that the question asks for.

There are two ways to proceed. The first is to fiddle with the choice of parameters, hoping to hit on the right set. For that, we can do the prior paragraph in reverse. Converting the given dimensionless products gt/v_0, gx/v_0^2, gy/v_0^2, and θ into vectors gives this description (note the *?*'s where the parameters will go).

$$\{\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix} = \underline{\ ?\ } \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \underline{\ ?\ } \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \underline{\ ?\ } \begin{pmatrix} 0 \\ 1 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + p_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \bigm| p_2, p_4, p_5, p_6 \in \mathbb{R}\}$$

The p_4 is already in place. Examining the rows shows that we can also put in place p_6, p_1, and p_2.

The second way to proceed, following the hint, is to note that the given set is of size four in a four-dimensional vector space and so we need only show that it is linearly independent. That is easily done by inspection, by considering the sixth, first, second, and fourth components of the vectors.

(b) The first equation can be rewritten

$$\frac{gx}{{v_0}^2} = \frac{gt}{v_0}\cos\theta$$

so that Buckingham's function is $f_1(\Pi_1, \Pi_2, \Pi_3, \Pi_4) = \Pi_2 - \Pi_1 \cos(\Pi_4)$. The second equation can be rewritten

$$\frac{gy}{{v_0}^2} = \frac{gt}{v_0}\sin\theta - \frac{1}{2}\left(\frac{gt}{v_0}\right)^2$$

and Buckingham's function here is $f_2(\Pi_1, \Pi_2, \Pi_3, \Pi_4) = \Pi_3 - \Pi_1 \sin(\Pi_4) + (1/2){\Pi_1}^2$.

2 We consider

$$(L^0M^0T^{-1})^{p_1}(L^1M^{-1}T^2)^{p_2}(L^{-3}M^0T^0)^{p_3}(L^0M^1T^0)^{p_4} = (L^0M^0T^0)$$

which gives these relations among the powers.

$$\begin{array}{rcl} p_2 - 3p_3 &= 0 \\ -p_2 + p_4 &= 0 \\ -p_1 + 2p_2 &= 0 \end{array} \xrightarrow{\rho_1\leftrightarrow\rho_3} \xrightarrow{\rho_2+\rho_3} \begin{array}{rcl} -p_1 + 2p_2 &= 0 \\ -p_2 + p_4 &= 0 \\ -3p_3 + p_4 &= 0 \end{array}$$

This is the solution space (because we wish to express k as a function of the other quantities, p_2 is taken as the parameter).

$$\{\begin{pmatrix} 2 \\ 1 \\ 1/3 \\ 1 \end{pmatrix} p_2 \bigm| p_2 \in \mathbb{R}\}$$

Thus, $\Pi_1 = \nu^2 k N^{1/3} m$ is the dimensionless combination, and we have that k equals $\nu^{-2}N^{-1/3}m^{-1}$ times a constant (the function $\hat{f}$ is constant since it has no arguments).

3

(a) Setting

$$(L^2M^1T^{-2})^{p_1}(L^0M^0T^{-1})^{p_2}(L^3M^0T^0)^{p_3} = (L^0M^0T^0)$$

gives this

$$\begin{aligned} 2p_1 \quad\quad\; + 3p_3 &= 0 \\ p_1 \quad\quad\quad\quad\; &= 0 \\ -2p_1 - p_2 \quad\quad &= 0 \end{aligned}$$

which implies that $p_1 = p_2 = p_3 = 0$. That is, among quantities with these dimensional formulas, the only dimensionless product is the trivial one.

(b) Setting

$$(L^2M^1T^{-2})^{p_1}(L^0M^0T^{-1})^{p_2}(L^3M^0T^0)^{p_3}(L^{-3}M^1T^0)^{p_4} = (L^0M^0T^0)$$

gives this.

$$\begin{array}{rcl} 2p_1 \quad + 3p_3 - 3p_4 = 0 & & 2p_1 \quad + \quad 3p_3 - \quad 3p_4 = 0 \\ p_1 \quad\quad\quad + \; p_4 = 0 & \xrightarrow[\rho_1+\rho_3]{(-1/2)\rho_1+\rho_2} \xrightarrow{\rho_2\leftrightarrow\rho_3} & -p_2 + \quad 3p_3 - \quad 3p_4 = 0 \\ -2p_1 - p_2 \quad\quad\quad = 0 & & (-3/2)p_3 + (5/2)p_4 = 0 \end{array}$$

Taking p_1 as parameter to express the torque gives this description of the solution set.

$$\{\begin{pmatrix} 1 \\ -2 \\ -5/3 \\ -1 \end{pmatrix} p_1 \bigm| p_1 \in \mathbb{R}\}$$

Denoting the torque by τ, the rotation rate by r, the volume of air by V, and the density of air by d we have that $\Pi_1 = \tau r^{-2}V^{-5/3}d^{-1}$, and so the torque is $r^2V^{5/3}d$ times a constant.

4

(a) These are the dimensional formulas.

quantity	*dimensional formula*
speed of the wave v	$L^1M^0T^{-1}$
separation of the dominoes d	$L^1M^0T^0$
height of the dominoes h	$L^1M^0T^0$
acceleration due to gravity g	$L^1M^0T^{-2}$

(b) The relationship

$$(L^1M^0T^{-1})^{p_1}(L^1M^0T^0)^{p_2}(L^1M^0T^0)^{p_3}(L^1M^0T^{-2})^{p_4} = (L^0M^0T^0)$$

gives this linear system.

$$\begin{array}{rcl} p_1 + p_2 + p_3 + \; p_4 = 0 & & \\ 0 = 0 & \xrightarrow{\rho_1+\rho_4} & p_1 + p_2 + p_3 + p_4 = 0 \\ -p_1 \quad\quad\quad\quad - 2p_4 = 0 & & p_2 + p_3 - p_4 = 0 \end{array}$$

Taking p_3 and p_4 as parameters, the solution set is described in this way.

$$\{\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} p_3 + \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} p_4 \bigm| p_3, p_4 \in \mathbb{R}\}$$

That gives $\{\Pi_1 = h/d, \Pi_2 = dg/v^2\}$ as a complete set.

(c) Buckingham's Theorem says that $v^2 = dg \cdot \hat{f}(h/d)$, and so, since g is a constant, if h/d is fixed then v is proportional to $\sqrt{d}$.

5 Checking the conditions in the definition of a vector space is routine.

6

(a) The dimensional formula of the circumference is L, that is, $L^1M^0T^0$. The dimensional formula of the area is L^2.

(b) One is $C + A = 2\pi r + \pi r^2$.

(c) One example is this formula relating the the length of arc subtended by an angle to the radius and the angle measure in radians: $\ell - r\theta = 0$. Both terms in that formula have dimensional formula L^1. The relationship holds for some unit systems (inches and radians, for instance) but not for all unit systems (inches and degrees, for instance).

Chapter 3. Maps Between Spaces

. . . .

Answers for subsection 3.I.1

3.I.1.12 To verify it is one-to-one, assume that $f_1(c_1x+c_2y+c_3z)=f_1(d_1x+d_2y+d_3z)$. Then $c_1+c_2x+c_3x^2=d_1+d_2x+d_3x^2$ by the definition of f_1. Members of $\mathcal{P}_2$ are equal only when they have the same coefficients, so this implies that $c_1=d_1$ and $c_2=d_2$ and $c_3=d_3$. Therefore $f_1(c_1x+c_2y+c_3z)=f_1(d_1x+d_2y+d_3z)$ implies that $c_1x+c_2y+c_3z=d_1x+d_2y+d_3z$, and so f_1 is one-to-one.

To verify that it is onto, consider an arbitrary member of the codomain $a_1+a_2x+a_3x^2$ and observe that it is indeed the image of a member of the domain, namely, it is $f_1(a_1x+a_2y+a_3z)$. (For instance, $0+3x+6x^2=f_1(0x+3y+6z)$.)

The computation checking that f_1 preserves addition is this.

$$\begin{aligned}
f_1\,(\,(c_1x+c_2y+c_3z)+(d_1x+d_2y+d_3z)\,) &= f_1\,(\,(c_1+d_1)x+(c_2+d_2)y+(c_3+d_3)z\,)\\
&=(c_1+d_1)+(c_2+d_2)x+(c_3+d_3)x^2\\
&=(c_1+c_2x+c_3x^2)+(d_1+d_2x+d_3x^2)\\
&=f_1(c_1x+c_2y+c_3z)+f_1(d_1x+d_2y+d_3z)
\end{aligned}$$

The check that f_1 preserves scalar multiplication is this.

$$\begin{aligned}
f_1(\,r\cdot(c_1x+c_2y+c_3z)\,) &= f_1(\,(rc_1)x+(rc_2)y+(rc_3)z\,)\\
&=(rc_1)+(rc_2)x+(rc_3)x^2\\
&=r\cdot(c_1+c_2x+c_3x^2)\\
&=r\cdot f_1(c_1x+c_2y+c_3z)
\end{aligned}$$

3.I.1.14 It is one-to-one and onto, a correspondence, because it has an inverse (namely, $f^{-1}(x)=\sqrt[3]{x}$). However, it is not an isomorphism. For instance, $f(1)+f(1)\neq f(1+1)$.

3.I.1.16 Here are two.

$$a_0+a_1x+a_2x^2\mapsto\begin{pmatrix}a_1\\a_0\\a_2\end{pmatrix}\quad\text{and}\quad a_0+a_1x+a_2x^2\mapsto\begin{pmatrix}a_0+a_1\\a_1\\a_2\end{pmatrix}$$

Verification is straightforward (for the second, to show that it is onto, note that

$$\begin{pmatrix}s\\t\\u\end{pmatrix}$$

is the image of $(s-t)+tx+ux^2$).

3.I.1.18 Here are two:

$$\begin{pmatrix}r_1\\r_2\\\vdots\\r_{16}\end{pmatrix}\mapsto\begin{pmatrix}r_1&r_2&\ldots&\\&&\ldots&r_{16}\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}r_1\\r_2\\\vdots\\r_{16}\end{pmatrix}\mapsto\begin{pmatrix}r_1&\\r_2&\\\vdots&\vdots\\&r_{16}\end{pmatrix}$$

Verification that each is an isomorphism is easy.

3.I.1.20 If $n\geq 1$ then $\mathcal{P}_{n-1}\cong\mathbb{R}^n$. (If we take $\mathcal{P}_{-1}$ and $\mathbb{R}^0$ to be trivial vector spaces, then the relationship extends one dimension lower.) The natural isomorphism between them is this.

$$a_0+a_1x+\cdots+a_{n-1}x^{n-1}\mapsto\begin{pmatrix}a_0\\a_1\\\vdots\\a_{n-1}\end{pmatrix}$$

Checking that it is an isomorphism is straightforward.

3.I.1.21 This is the map, expanded.

$$\begin{aligned}
f(a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5) &= a_0+a_1(x-1)+a_2(x-1)^2+a_3(x-1)^3 \\
&\qquad +a_4(x-1)^4+a_5(x-1)^5 \\
&= a_0+a_1(x-1)+a_2(x^2-2x+1) \\
&\qquad +a_3(x^3-3x^2+3x-1)+a_4(x^4-4x^3+6x^2-4x+1) \\
&\qquad +a_5(x^5-5x^4+10x^3-10x^2+5x-1) \\
&= (a_0-a_1+a_2-a_3+a_4-a_5)+(a_1-2a_2+3a_3-4a_4+5a_5)x \\
&\qquad +(a_2-3a_3+6a_4-10a_5)x^2+(a_3-4a_4+10a_5)x^3 \\
&\qquad +(a_4-5a_5)x^4+a_5x^5
\end{aligned}$$

This map is a correspondence because it has an inverse, the map $p(x)\mapsto p(x+1)$.

To finish checking that it is an isomorphism, we apply item (2) of Lemma 1.9 and show that it preserves linear combinations of two polynomials. Briefly, the check goes like this.

$$\begin{aligned}
&f(c\cdot(a_0+a_1x+\cdots+a_5x^5)+d\cdot(b_0+b_1x+\cdots+b_5x^5)) \\
&\quad =\cdots=(ca_0-ca_1+ca_2-ca_3+ca_4-ca_5+db_0-db_1+db_2-db_3+db_4-db_5)+\cdots+(ca_5+db_5)x^5 \\
&\quad =\cdots=c\cdot f(a_0+a_1x+\cdots+a_5x^5)+d\cdot f(b_0+b_1x+\cdots+b_5x^5)
\end{aligned}$$

3.I.1.22 No vector space has the empty set underlying it. We can take $\vec{v}$ to be the zero vector.

3.I.1.23 Yes; where the two spaces are $\{\vec{a}\}$ and $\{\vec{b}\}$, the map sending $\vec{a}$ to $\vec{b}$ is clearly one-to-one and onto, and also preserves what little structure there is.

3.I.1.24 A linear combination of $n=0$ vectors adds to the zero vector and so Lemma 1.8 shows that the three statements are equivalent in this case.

3.I.1.25 Consider the basis $\langle 1\rangle$ for $\mathcal{P}_0$ and let $f(1)\in\mathbb{R}$ be k. For any $a\in\mathcal{P}_0$ we have that $f(a)=f(a\cdot 1)=af(1)=ak$ and so f's action is multiplication by k. Note that $k\neq 0$ or else the map is not one-to-one. (Incidentally, any such map $a\mapsto ka$ is an isomorphism, as is easy to check.)

3.I.1.27 One direction is easy: by definition, if f is one-to-one then for any $\vec{w}\in W$ at most one $\vec{v}\in V$ has $f(\vec{v}\,)=\vec{w}$, and so in particular, at most one member of V is mapped to $\vec{0}_W$. The proof of Lemma 1.8 does not use the fact that the map is a correspondence and therefore shows that any structure-preserving map f sends $\vec{0}_V$ to $\vec{0}_W$.

For the other direction, assume that the only member of V that is mapped to $\vec{0}_W$ is $\vec{0}_V$. To show that f is one-to-one assume that $f(\vec{v}_1)=f(\vec{v}_2)$. Then $f(\vec{v}_1)-f(\vec{v}_2)=\vec{0}_W$ and so $f(\vec{v}_1-\vec{v}_2)=\vec{0}_W$. Consequently $\vec{v}_1-\vec{v}_2=\vec{0}_V$, so $\vec{v}_1=\vec{v}_2$, and so f is one-to-one.

3.I.1.28 We will prove something stronger—not only is the existence of a dependence preserved by isomorphism, but each instance of a dependence is preserved, that is,

$$\begin{aligned}
&\vec{v}_i=c_1\vec{v}_1+\cdots+c_{i-1}\vec{v}_{i-1}+c_{i+1}\vec{v}_{i+1}+\cdots+c_k\vec{v}_k \\
&\qquad\iff f(\vec{v}_i)=c_1f(\vec{v}_1)+\cdots+c_{i-1}f(\vec{v}_{i-1})+c_{i+1}f(\vec{v}_{i+1})+\cdots+c_kf(\vec{v}_k).
\end{aligned}$$

The $\Longrightarrow$ direction of this statement holds by item (3) of Lemma 1.9. The $\Longleftarrow$ direction holds by regrouping

$$\begin{aligned}
f(\vec{v}_i)&=c_1f(\vec{v}_1)+\cdots+c_{i-1}f(\vec{v}_{i-1})+c_{i+1}f(\vec{v}_{i+1})+\cdots+c_kf(\vec{v}_k) \\
&=f(c_1\vec{v}_1+\cdots+c_{i-1}\vec{v}_{i-1}+c_{i+1}\vec{v}_{i+1}+\cdots+c_k\vec{v}_k)
\end{aligned}$$

and applying the fact that f is one-to-one, and so for the two vectors $\vec{v}_i$ and $c_1\vec{v}_1+\cdots+c_{i-1}\vec{v}_{i-1}+c_{i+1}f\vec{v}_{i+1}+\cdots+c_kf(\vec{v}_k$ to be mapped to the same image by f, they must be equal.

3.I.1.30 First, the map $p(x)\mapsto p(x+k)$ doesn't count because it is a version of $p(x)\mapsto p(x-k)$. Here is a correct answer (many others are also correct): $a_0+a_1x+a_2x^2\mapsto a_2+a_0x+a_1x^2$. Verification that this is an isomorphism is straightforward.

3.I.1.31

(a) For the 'only if' half, let $f\colon \mathbb{R}^1 \to \mathbb{R}^1$ to be an isomorphism. Consider the basis $\langle 1\rangle \subseteq \mathbb{R}^1$. Designate $f(1)$ by k. Then for any x we have that $f(x) = f(x\cdot 1) = x\cdot f(1) = xk$, and so f's action is multiplication by k. To finish this half, just note that $k \neq 0$ or else f would not be one-to-one.

For the 'if' half we only have to check that such a map is an isomorphism when $k \neq 0$. To check that it is one-to-one, assume that $f(x_1) = f(x_2)$ so that $kx_1 = kx_2$ and divide by the nonzero factor k to conclude that $x_1 = x_2$. To check that it is onto, note that any $y \in \mathbb{R}^1$ is the image of $x = y/k$ (again, $k \neq 0$). Finally, to check that such a map preserves combinations of two members of the domain, we have this.

$$f(c_1x_1 + c_2x_2) = k(c_1x_1 + c_2x_2) = c_1kx_1 + c_2kx_2 = c_1f(x_1) + c_2f(x_2)$$

(b) By the prior item, f's action is $x \mapsto (7/3)x$. Thus $f(-2) = -14/3$.

(c) For the 'only if' half, assume that $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is an automorphism. Consider the standard basis $\mathcal{E}_2$ for $\mathbb{R}^2$. Let

$$f(\vec{e}_1) = \begin{pmatrix} a \\ c \end{pmatrix} \quad\text{and}\quad f(\vec{e}_2) = \begin{pmatrix} b \\ d \end{pmatrix}.$$

Then the action of f on any vector is determined by by its action on the two basis vectors.

$$f(\begin{pmatrix} x \\ y \end{pmatrix}) = f(x\cdot\vec{e}_1 + y\cdot\vec{e}_2) = x\cdot f(\vec{e}_1) + y\cdot f(\vec{e}_2) = x\cdot\begin{pmatrix} a \\ c \end{pmatrix} + y\cdot\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$$

To finish this half, note that if $ad - bc = 0$, that is, if $f(\vec{e}_2)$ is a multiple of $f(\vec{e}_1)$, then f is not one-to-one.

For 'if' we must check that the map is an isomorphism, under the condition that $ad - bc \neq 0$. The structure-preservation check is easy; we will here show that f is a correspondence. For the argument that the map is one-to-one, assume this.

$$f(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}) = f(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}) \quad\text{and so}\quad \begin{pmatrix} ax_1+by_1 \\ cx_1+dy_1 \end{pmatrix} = \begin{pmatrix} ax_2+by_2 \\ cx_2+dy_2 \end{pmatrix}$$

Then, because $ad - bc \neq 0$, the resulting system

$$\begin{aligned} a(x_1-x_2) + b(y_1-y_2) &= 0 \\ c(x_1-x_2) + d(y_1-y_2) &= 0 \end{aligned}$$

has a unique solution, namely the trivial one $x_1 - x_2 = 0$ and $y_1 - y_2 = 0$ (this follows from the hint).

The argument that this map is onto is closely related—this system

$$\begin{aligned} ax_1 + by_1 &= x \\ cx_1 + dy_1 &= y \end{aligned}$$

has a solution for any x and y if and only if this set

$$\{\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}\}$$

spans $\mathbb{R}^2$, i.e., if and only if this set is a basis (because it is a two-element subset of $\mathbb{R}^2$), i.e., if and only if $ad - bc \neq 0$.

(d)

$$f(\begin{pmatrix} 0 \\ -1 \end{pmatrix}) = f(\begin{pmatrix} 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 4 \end{pmatrix}) = f(\begin{pmatrix} 1 \\ 3 \end{pmatrix}) - f(\begin{pmatrix} 1 \\ 4 \end{pmatrix}) = \begin{pmatrix} 2 \\ -1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

3.I.1.32 There are many answers; two are linear independence and subspaces.

To show that if a set $\{\vec{v}_1,\dots,\vec{v}_n\}$ is linearly independent then its image $\{f(\vec{v}_1),\dots,f(\vec{v}_n)\}$ is also linearly independent, consider a linear relationship among members of the image set.

$$0 = c_1f(\vec{v}_1) + \cdots + c_nf(\vec{v_n}) = f(c_1\vec{v}_1) + \cdots + f(c_n\vec{v_n}) = f(c_1\vec{v}_1 + \cdots + c_n\vec{v_n})$$

Because this map is an isomorphism, it is one-to-one. So f maps only one vector from the domain to the zero vector in the range, that is, $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$ equals the zero vector (in the domain, of course). But, if $\{\vec{v}_1,\dots,\vec{v}_n\}$ is linearly independent then all of the c's are zero, and so $\{f(\vec{v}_1),\dots,f(\vec{v}_n)\}$ is linearly independent also. (*Remark.* There is a small point about this argument that is worth mention. In a set, repeats collapse, that is, strictly speaking, this is a one-element set: $\{\vec{v},\vec{v}\}$, because the things listed as in it are the same thing. Observe, however, the use of the subscript n in the above argument. In moving from the domain set $\{\vec{v}_1,\dots,\vec{v}_n\}$ to the image set $\{f(\vec{v}_1),\dots,f(\vec{v}_n)\}$, there is no collapsing, because the image set does not have repeats, because the isomorphism f is one-to-one.)

To show that if $f\colon V \to W$ is an isomorphism and if U is a subspace of the domain V then the set of image vectors $f(U) = \{\vec{w} \in W \bigm| \vec{w} = f(\vec{u}) \text{ for some } \vec{u} \in U\}$ is a subspace of W, we need only show that it is closed under linear combinations of two of its members (it is nonempty because it contains the image of the zero vector). We have

$$c_1 \cdot f(\vec{u}_1) + c_2 \cdot f(\vec{u}_2) = f(c_1\vec{u}_1) + f(c_2\vec{u}_2) = f(c_1\vec{u}_1 + c_2\vec{u}_2)$$

and $c_1\vec{u}_1 + c_2\vec{u}_2$ is a member of U because of the closure of a subspace under combinations. Hence the combination of $f(\vec{u}_1)$ and $f(\vec{u}_2)$ is a member of $f(U)$.

3.I.1.33

(a) The association

$$\vec{p} = c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + c_3\vec{\beta}_3 \overset{\mathrm{Rep}_B(\cdot)}{\longmapsto} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

is a function if every member $\vec{p}$ of the domain is associated with at least one member of the codomain, and if every member $\vec{p}$ of the domain is associated with at most one member of the codomain. The first condition holds because the basis B spans the domain—every $\vec{p}$ can be written as at least one linear combination of $\vec{\beta}$'s. The second condition holds because the basis B is linearly independent—every member of the domain $\vec{p}$ can be written as at most one linear combination of the $\vec{\beta}$'s.

(b) For the one-to-one argument, if $\mathrm{Rep}_B(\vec{p}) = \mathrm{Rep}_B(\vec{q})$, that is, if $\mathrm{Rep}_B(p_1\vec{\beta}_1+p_2\vec{\beta}_2+p_3\vec{\beta}_3) = \mathrm{Rep}_B(q_1\vec{\beta}_1+q_2\vec{\beta}_2 + q_3\vec{\beta}_3)$ then

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$$

and so $p_1 = q_1$ and $p_2 = q_2$ and $p_3 = q_3$, which gives the conclusion that $\vec{p} = \vec{q}$. Therefore this map is one-to-one.

For onto, we can just note that

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

equals $\mathrm{Rep}_B(a\vec{\beta}_1 + b\vec{\beta}_2 + c\vec{\beta}_3)$, and so any member of the codomain $\mathbb{R}^3$ is the image of some member of the domain $\mathcal{P}_2$.

(c) This map respects addition and scalar multiplication because it respects combinations of two members of the domain (that is, we are using item (2) of Lemma 1.9): where $\vec{p} = p_1\vec{\beta}_1 + p_2\vec{\beta}_2 + p_3\vec{\beta}_3$ and $\vec{q} = q_1\vec{\beta}_1 + q_2\vec{\beta}_2 + q_3\vec{\beta}_3$, we have this.

$$\begin{aligned}
\mathrm{Rep}_B(c \cdot \vec{p} + d \cdot \vec{q}) &= \mathrm{Rep}_B(\,(cp_1 + dq_1)\vec{\beta}_1 + (cp_2 + dq_2)\vec{\beta}_2 + (cp_3 + dq_3)\vec{\beta}_3\,) \\
&= \begin{pmatrix} cp_1 + dq_1 \\ cp_2 + dq_2 \\ cp_3 + dq_3 \end{pmatrix} \\
&= c \cdot \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + d \cdot \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \\
&= \mathrm{Rep}_B(\vec{p}) + \mathrm{Rep}_B(\vec{q})
\end{aligned}$$

(d) Use any basis B for $\mathcal{P}_2$ whose first two members are $x + x^2$ and $1 - x$, say $B = \langle x + x^2, 1 - x, 1\rangle$.

3.I.1.34 See the next subsection.

3.I.1.35

(a) Most of the conditions in the definition of a vector space are routine. We here sketch the verification of part (1) of that definition.

For closure of $U \times W$, note that because U and W are closed, we have that $\vec{u}_1+\vec{u}_2 \in U$ and $\vec{w}_1+\vec{w}_2 \in W$ and so $(\vec{u}_1 + \vec{u}_2, \vec{w}_1 + \vec{w}_2) \in U \times W$. Commutativity of addition in $U \times W$ follows from commutativity of addition in U and W.

$$(\vec{u}_1, \vec{w}_1) + (\vec{u}_2, \vec{w}_2) = (\vec{u}_1 + \vec{u}_2, \vec{w}_1 + \vec{w}_2) = (\vec{u}_2 + \vec{u}_1, \vec{w}_2 + \vec{w}_1) = (\vec{u}_2, \vec{w}_2) + (\vec{u}_1, \vec{w}_1)$$

The check for associativity of addition is similar. The zero element is $(\vec{0}_U,\vec{0}_W)\in U\times W$ and the additive inverse of $(\vec{u},\vec{w})$ is $(-\vec{u},-\vec{w})$.

The checks for the second part of the definition of a vector space are also straightforward.

(b) This is a basis

$$\langle\,(1,\begin{pmatrix}0\\0\end{pmatrix}),(x,\begin{pmatrix}0\\0\end{pmatrix}),(x^2,\begin{pmatrix}0\\0\end{pmatrix}),(1,\begin{pmatrix}1\\0\end{pmatrix}),(1,\begin{pmatrix}0\\1\end{pmatrix})\,\rangle$$

because there is one and only one way to represent any member of $\mathcal{P}_2\times\mathbb{R}^2$ with respect to this set; here is an example.

$$(3+2x+x^2,\begin{pmatrix}5\\4\end{pmatrix})=3\cdot(1,\begin{pmatrix}0\\0\end{pmatrix})+2\cdot(x,\begin{pmatrix}0\\0\end{pmatrix})+(x^2,\begin{pmatrix}0\\0\end{pmatrix})+5\cdot(1,\begin{pmatrix}1\\0\end{pmatrix})+4\cdot(1,\begin{pmatrix}0\\1\end{pmatrix})$$

The dimension of this space is five.

(c) We have $\dim(U\times W)=\dim(U)+\dim(W)$ as this is a basis.

$$\langle(\vec{\mu}_1,\vec{0}_W),\dots,(\vec{\mu}_{\dim(U)},\vec{0}_W),(\vec{0}_U,\vec{\omega}_1),\dots,(\vec{0}_U,\vec{\omega}_{\dim(W)})\rangle$$

(d) We know that if $V=U\oplus W$ then each $\vec{v}\in V$ can be written as $\vec{v}=\vec{u}+\vec{w}$ in one and only one way. This is just what we need to prove that the given function an isomorphism.

First, to show that f is one-to-one we can show that if $f\left((\vec{u}_1,\vec{w}_1)\right)=\left((\vec{u}_2,\vec{w}_2)\right)$, that is, if $\vec{u}_1+\vec{w}_1=\vec{u}_2+\vec{w}_2$ then $\vec{u}_1=\vec{u}_2$ and $\vec{w}_1=\vec{w}_2$. But the statement 'each $\vec{v}$ is such a sum in only one way' is exactly what is needed to make this conclusion. Similarly, the argument that f is onto is completed by the statement that 'each $\vec{v}$ is such a sum in at least one way'.

This map also preserves linear combinations

$$\begin{aligned}f(\,c_1\cdot(\vec{u}_1,\vec{w}_1)+c_2\cdot(\vec{u}_2,\vec{w}_2)\,)&=f(\,(c_1\vec{u}_1+c_2\vec{u}_2,c_1\vec{w}_1+c_2\vec{w}_2)\,)\\&=c_1\vec{u}_1+c_2\vec{u}_2+c_1\vec{w}_1+c_2\vec{w}_2\\&=c_1\vec{u}_1+c_1\vec{w}_1+c_2\vec{u}_2+c_2\vec{w}_2\\&=c_1\cdot f(\,(\vec{u}_1,\vec{w}_1)\,)+c_2\cdot f(\,(\vec{u}_2,\vec{w}_2)\,)\end{aligned}$$

and so it is an isomorphism.

. . . .

Answers for subsection 3.I.2

3.I.2.14 There are many answers, one is the set of $\mathcal{P}_k$ (taking $\mathcal{P}_{-1}$ to be the trivial vector space).

3.I.2.15 False (except when $n=0$). For instance, if $f\colon V\to\mathbb{R}^n$ is an isomorphism then multiplying by any nonzero scalar, gives another, different, isomorphism. (Between trivial spaces the isomorphisms are unique; the only map possible is $\vec{0}_V\mapsto 0_W$.)

3.I.2.16 No. A proper subspace has a strictly lower dimension than it's superspace; if U is a proper subspace of V then any linearly independent subset of U must have fewer than $\dim(V)$ members or else that set would be a basis for V, and U wouldn't be proper.

3.I.2.19 We must show that if $\vec{a}=\vec{b}$ then $f(\vec{a})=f(\vec{b})$. So suppose that $a_1\vec{\beta}_1+\dots+a_n\vec{\beta}_n=b_1\vec{\beta}_1+\dots+b_n\vec{\beta}_n$. Each vector in a vector space (here, the domain space) has a unique representation as a linear combination of basis vectors, so we can conclude that $a_1=b_1,\ \dots,\ a_n=b_n$. Thus,

$$f(\vec{a})=\begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix}=\begin{pmatrix}b_1\\\vdots\\b_n\end{pmatrix}=f(\vec{b})$$

and so the function is well-defined.

3.I.2.20 Yes, because a zero-dimensional space is a trivial space.

3.I.2.21

(a) No, this collection has no spaces of odd dimension.

(b) Yes, because $\mathcal{P}_k\cong\mathbb{R}^{k+1}$.

(c) No, for instance, $\mathcal{M}_{2\times3} \cong \mathcal{M}_{3\times2}$.

3.I.2.22 One direction is easy: if the two are isomorphic via f then for any basis $B \subseteq V$, the set $D = f(B)$ is also a basis (this is shown in Lemma 2.3). The check that corresponding vectors have the same coordinates: $f(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n) = c_1 f(\vec{\beta}_1) + \cdots + c_n f(\vec{\beta}_n) = c_1\vec{\delta}_1 + \cdots + c_n\vec{\delta}_n$ is routine.

For the other half, assume that there are bases such that corresponding vectors have the same coordinates with respect to those bases. Because f is a correspondence, to show that it is an isomorphism, we need only show that it preserves structure. Because $\mathrm{Rep}_B(\vec{v}\,) = \mathrm{Rep}_D(f(\vec{v}\,))$, the map f preserves structure if and only if representations preserve addition: $\mathrm{Rep}_B(\vec{v}_1 + \vec{v}_2) = \mathrm{Rep}_B(\vec{v}_1) + \mathrm{Rep}_B(\vec{v}_2)$ and scalar multiplication: $\mathrm{Rep}_B(r \cdot \vec{v}\,) = r \cdot \mathrm{Rep}_B(\vec{v}\,)$ The addition calculation is this: $(c_1 + d_1)\vec{\beta}_1 + \cdots + (c_n + d_n)\vec{\beta}_n = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n + d_1\vec{\beta}_1 + \cdots + d_n\vec{\beta}_n$, and the scalar multiplication calculation is similar.

3.I.2.23

(a) Pulling the definition back from $\mathbb{R}^4$ to $\mathcal{P}_3$ gives that $a_0 + a_1x + a_2x^2 + a_3x^3$ is orthogonal to $b_0 + b_1x + b_2x^2 + b_3x^3$ if and only if $a_0b_0 + a_1b_1 + a_2b_2 + a_3b_3 = 0$.

(b) A natural definition is this.

$$D(\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}) = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{pmatrix}$$

3.I.2.25 Because $V_1 \cap V_2 = \{\vec{0}_V\}$ and f is one-to-one we have that $f(V_1) \cap f(V_2) = \{\vec{0}_U\}$. To finish, count the dimensions: $\dim(U) = \dim(V) = \dim(V_1) + \dim(V_2) = \dim(f(V_1)) + \dim(f(V_2))$, as required.

3.I.2.26 Rational numbers have many representations, e.g., $1/2 = 3/6$, and the numerators can vary among representations.

. . . .

Answers for subsection 3.II.1

3.II.1.19 Each of these projections is a homomorphism. Projection to the xz-plane and to the yz-plane are these maps.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ y \\ z \end{pmatrix}$$

Projection to the x-axis, to the y-axis, and to the z-axis are these maps.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$$

And projection to the origin is this map.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Verification that each is a homomorphism is straightforward. (The last one, of course, is the zero transformation on $\mathbb{R}^3$.)

3.II.1.20 The first is not onto; for instance, there is no polynomial that is sent the constant polynomial $p(x) = 1$. The second is not one-to-one; both of these members of the domain

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

are mapped to the same member of the codomain, $1 \in \mathbb{R}$.

3.II.1.21 Yes; in any space $\mathrm{id}(c \cdot \vec{v} + d \cdot \vec{w}) = c \cdot \vec{v} + d \cdot \vec{w} = c \cdot \mathrm{id}(\vec{v}) + d \cdot \mathrm{id}(\vec{w})$.

3.II.1.24

(a) Let $\vec{v} \in V$ be represented with respect to the basis as $\vec{v} = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n$. Then $h(\vec{v}) = h(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n) = c_1 h(\vec{\beta}_1) + \cdots + c_n h(\vec{\beta}_n) = c_1 \cdot \vec{0} + \cdots + c_n \cdot \vec{0} = \vec{0}$.

(b) This argument is similar to the prior one. Let $\vec{v} \in V$ be represented with respect to the basis as $\vec{v} = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n$. Then $h(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n) = c_1 h(\vec{\beta}_1) + \cdots + c_n h(\vec{\beta}_n) = c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n = \vec{v}$.

(c) As above, only $c_1 h(\vec{\beta}_1) + \cdots + c_n h(\vec{\beta}_n) = c_1 r\vec{\beta}_1 + \cdots + c_n r\vec{\beta}_n = r(c_1\vec{\beta}_1 + \cdots + c_n\vec{\beta}_n) = r\vec{v}$.

3.II.1.29 Let $h\colon \mathbb{R}^1 \to \mathbb{R}^1$ be linear. A linear map is determined by its action on a basis, so fix the basis $\langle 1 \rangle$ for $\mathbb{R}^1$. For any $r \in \mathbb{R}^1$ we have that $h(r) = h(r \cdot 1) = r \cdot h(1)$ and so h acts on any argument r by multiplying it by the constant $h(1)$. If $h(1)$ is not zero then the map is a correspondence—its inverse is division by $h(1)$—so any nontrivial transformation of $\mathbb{R}^1$ is an isomorphism.

This projection map is an example that shows that not every transformation of $\mathbb{R}^n$ acts via multiplication by a constant when $n > 1$, including when $n = 2$.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

3.II.1.30

(a) Where c and d are scalars, we have this.

$$\begin{aligned}
h(c \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + d \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}) &= h(\begin{pmatrix} cx_1 + dy_1 \\ \vdots \\ cx_n + dy_n \end{pmatrix}) \\
&= \begin{pmatrix} a_{1,1}(cx_1 + dy_1) + \cdots + a_{1,n}(cx_n + dy_n) \\ \vdots \\ a_{m,1}(cx_1 + dy_1) + \cdots + a_{m,n}(cx_n + dy_n) \end{pmatrix} \\
&= c \cdot \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{pmatrix} + d \cdot \begin{pmatrix} a_{1,1}y_1 + \cdots + a_{1,n}y_n \\ \vdots \\ a_{m,1}y_1 + \cdots + a_{m,n}y_n \end{pmatrix} \\
&= c \cdot h(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) + d \cdot h(\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix})
\end{aligned}$$

(b) Each power i of the derivative operator is linear because of these rules familiar from calculus.

$$\frac{d^i}{dx^i}(\, f(x) + g(x)\,) = \frac{d^i}{dx^i} f(x) + \frac{d^i}{dx^i} g(x) \quad \text{and} \quad \frac{d^i}{dx^i}\, r \cdot f(x) = r \cdot \frac{d^i}{dx^i} f(x)$$

Thus the given map is a linear transformation of $\mathcal{P}_n$ because any linear combination of linear maps is also a linear map.

3.II.1.31 (This argument has already appeared, as part of the proof that isomorphism is an equivalence.) Let $f\colon U \to V$ and $g\colon V \to W$ be linear. For any $\vec{u}_1, \vec{u}_2 \in U$ and scalars c_1, c_2 combinations are preserved.

$$\begin{aligned}
g \circ f(c_1\vec{u}_1 + c_2\vec{u}_2) = g(\, f(c_1\vec{u}_1 + c_2\vec{u}_2)\,) &= g(\, c_1 f(\vec{u}_1) + c_2 f(\vec{u}_2)\,) \\
&= c_1 \cdot g(f(\vec{u}_1)) + c_2 \cdot g(f(\vec{u}_2)) = c_1 \cdot g \circ f(\vec{u}_1) + c_2 \cdot g \circ f(\vec{u}_2)
\end{aligned}$$

3.II.1.33 Recall that the entry in row i and column j of the transpose of M is the entry $m_{j,i}$ from row j

and column i of M. Now, the check is routine.

$$
\begin{aligned}
[r\cdot\begin{pmatrix} & \vdots & \\ \cdots & a_{i,j} & \cdots \\ & \vdots & \end{pmatrix}+s\cdot\begin{pmatrix} & \vdots & \\ \cdots & b_{i,j} & \cdots \\ & \vdots & \end{pmatrix}]^{\text{trans}}
&=\begin{pmatrix} & \vdots & \\ \cdots & ra_{i,j}+sb_{i,j} & \cdots \\ & \vdots & \end{pmatrix}^{\text{trans}} \\
&=\begin{pmatrix} & \vdots & \\ \cdots & ra_{j,i}+sb_{j,i} & \cdots \\ & \vdots & \end{pmatrix} \\
&=r\cdot\begin{pmatrix} & \vdots & \\ \cdots & a_{j,i} & \cdots \\ & \vdots & \end{pmatrix}+s\cdot\begin{pmatrix} & \vdots & \\ \cdots & b_{j,i} & \cdots \\ & \vdots & \end{pmatrix} \\
&=r\cdot\begin{pmatrix} & \vdots & \\ \cdots & a_{j,i} & \cdots \\ & \vdots & \end{pmatrix}^{\text{trans}}+s\cdot\begin{pmatrix} & \vdots & \\ \cdots & b_{j,i} & \cdots \\ & \vdots & \end{pmatrix}^{\text{trans}}
\end{aligned}
$$

The domain is $\mathcal{M}_{m\times n}$ while the codomain is $\mathcal{M}_{n\times m}$.

3.II.1.34

(a) For any homomorphism $h\colon \mathbb{R}^n\to\mathbb{R}^m$ we have

$$h(\ell)=\{h(t\cdot\vec{u}+(1-t)\cdot\vec{v}) \bigm| t\in[0..1]\}=\{t\cdot h(\vec{u})+(1-t)\cdot h(\vec{v}) \bigm| t\in[0..1]\}$$

which is the line segment from $h(\vec{u})$ to $h(\vec{v})$.

(b) We must show that if a subset of the domain is convex then its image, as a subset of the range, is also convex. Suppose that $C\subseteq\mathbb{R}^n$ is convex and consider its image $h(C)$. To show $h(C)$ is convex we must show that for any two of its members, $\vec{d}_1$ and $\vec{d}_2$, the line segment connecting them

$$\ell=\{t\cdot\vec{d}_1+(1-t)\cdot\vec{d}_2 \bigm| t\in[0..1]\}$$

is a subset of $h(C)$.

Fix any member $\hat{t}\cdot\vec{d}_1+(1-\hat{t})\cdot\vec{d}_2$ of that line segment. Because the endpoints of ℓ are in the image of C, there are members of C that map to them, say $h(\vec{c}_1)=\vec{d}_1$ and $h(\vec{c}_2)=\vec{d}_2$. Now, where $\hat{t}$ is the scalar that is fixed in the first sentence of this paragraph, observe that $h(\hat{t}\cdot\vec{c}_1+(1-\hat{t})\cdot\vec{c}_2)=\hat{t}\cdot h(\vec{c}_1)+(1-\hat{t})\cdot h(\vec{c}_2)=\hat{t}\cdot\vec{d}_1+(1-\hat{t})\cdot\vec{d}_2$ Thus, any member of ℓ is a member of $h(C)$, and so $h(C)$ is convex.

3.II.1.36 Suppose that $h\colon V\to W$ is a homomorphism and suppose that S is a subspace of V. Consider the map $\hat{h}\colon S\to W$ defined by $\hat{h}(\vec{s})=h(\vec{s})$. (The only difference between $\hat{h}$ and h is the difference in domain.) Then this new map is linear: $\hat{h}(c_1\cdot\vec{s}_1+c_2\cdot\vec{s}_2)=h(c_1\vec{s}_1+c_2\vec{s}_2)=c_1h(\vec{s}_1)+c_2h(\vec{s}_2)=c_1\cdot\hat{h}(\vec{s}_1)+c_2\cdot\hat{h}(\vec{s}_2)$.

3.II.1.37 This will appear as a lemma in the next subsection.

(a) The range is nonempty because V is nonempty. To finish we need to show that it is closed under combinations. A combination of range vectors has the form, where $\vec{v}_1,\ldots,\vec{v}_n\in V$,

$$c_1\cdot h(\vec{v}_1)+\cdots+c_n\cdot h(\vec{v}_n)=h(c_1\vec{v}_1)+\cdots+h(c_n\vec{v}_n)=h(c_1\cdot\vec{v}_1+\cdots+c_n\cdot\vec{v}_n),$$

which is itself in the range as $c_1\cdot\vec{v}_1+\cdots+c_n\cdot\vec{v}_n$ is a member of domain V. Therefore the range is a subspace.

(b) The nullspace is nonempty since it contains $\vec{0}_V$, as $\vec{0}_V$ maps to $\vec{0}_W$. It is closed under linear combinations because, where $\vec{v}_1,\ldots,\vec{v}_n\in V$ are elements of the inverse image set $\{\vec{v}\in V \bigm| h(\vec{v})=\vec{0}_W\}$, for $c_1,\ldots,c_n\in\mathbb{R}$

$$\vec{0}_W=c_1\cdot h(\vec{v}_1)+\cdots+c_n\cdot h(\vec{v}_n)=h(c_1\cdot\vec{v}_1+\cdots+c_n\cdot\vec{v}_n)$$

and so $c_1\cdot\vec{v}_1+\cdots+c_n\cdot\vec{v}_n$ is also in the inverse image of $\vec{0}_W$.

(c) This image of U nonempty because U is nonempty. For closure under combinations, where $\vec{u}_1,\ldots,\vec{u}_n\in U$,

$$c_1\cdot h(\vec{u}_1)+\cdots+c_n\cdot h(\vec{u}_n)=h(c_1\cdot\vec{u}_1)+\cdots+h(c_n\cdot\vec{u}_n)=h(c_1\cdot\vec{u}_1+\cdots+c_n\cdot\vec{u}_n)$$

which is itself in $h(U)$ as $c_1\cdot\vec{u}_1+\cdots+c_n\cdot\vec{u}_n$ is in U. Thus this set is a subspace.

(d) The natural generalization is that the inverse image of a subspace of is a subspace.

Suppose that X is a subspace of W. Note that $\vec{0}_W \in X$ so the set $\{\vec{v} \in V \mid h(\vec{v}) \in X\}$ is not empty. To show that this set is closed under combinations, let $\vec{v}_1, \ldots, \vec{v}_n$ be elements of V such that $h(\vec{v}_1) = \vec{x}_1$, ... , $h(\vec{v}_n) = \vec{x}_n$ and note that

$$h(c_1 \cdot \vec{v}_1 + \cdots + c_n \cdot \vec{v}_n) = c_1 \cdot h(\vec{v}_1) + \cdots + c_n \cdot h(\vec{v}_n) = c_1 \cdot \vec{x}_1 + \cdots + c_n \cdot \vec{x}_n$$

so a linear combination of elements of $h^{-1}(X)$ is also in $h^{-1}(X)$.

3.II.1.38 No; the set of isomorphisms does not contain the zero map (unless the space is trivial).

3.II.1.39 If $\langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ doesn't span the space then the map needn't be unique. For instance, if we try to define a map from $\mathbb{R}^2$ to itself by specifying only that $\vec{e}_1$ is sent to itself, then there is more than one homomorphism possible; both the identity map and the projection map onto the first component fit this condition.

If we drop the condition that $\langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ is linearly independent then we risk an inconsistent specification (i.e, there could be no such map). An example is if we consider $\langle \vec{e}_2, \vec{e}_1, 2\vec{e}_1 \rangle$, and try to define a map from $\mathbb{R}^2$ to itself that sends $\vec{e}_2$ to itself, and sends both $\vec{e}_1$ and $2\vec{e}_1$ to $\vec{e}_1$. No homomorphism can satisfy these three conditions.

3.II.1.40

(a) Briefly, the check of linearity is this.

$$F(r_1 \cdot \vec{v}_1 + r_2 \cdot \vec{v}_2) = \begin{pmatrix} f_1(r_1\vec{v}_1 + r_2\vec{v}_2) \\ f_2(r_1\vec{v}_1 + r_2\vec{v}_2) \end{pmatrix} = r_1 \begin{pmatrix} f_1(\vec{v}_1) \\ f_2(\vec{v}_1) \end{pmatrix} + r_2 \begin{pmatrix} f_1(\vec{v}_2) \\ f_2(\vec{v}_2) \end{pmatrix} = r_1 \cdot F(\vec{v}_1) + r_2 \cdot F(\vec{v}_2)$$

(b) Yes. Let $\pi_1 \colon \mathbb{R}^2 \to \mathbb{R}^1$ and $\pi_2 \colon \mathbb{R}^2 \to \mathbb{R}^1$ be the projections

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{\pi_1}{\longmapsto} x \quad \text{and} \quad \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{\pi_2}{\longmapsto} y$$

onto the two axes. Now, where $f_1(\vec{v}) = \pi_1(F(\vec{v}))$ and $f_2(\vec{v}) = \pi_2(F(\vec{v}))$ we have the desired component functions.

$$F(\vec{v}) = \begin{pmatrix} f_1(\vec{v}) \\ f_2(\vec{v}) \end{pmatrix}$$

They are linear because they are the composition of linear functions, and the fact that the compoistion of linear functions is linear was shown as part of the proof that isomorphism is an equivalence relation (alternatively, the check that they are linear is straightforward).

(c) In general, a map from a vector space V to an $\mathbb{R}^n$ is linear if and only if each of the component functions is linear. The verification is as in the prior item.

. . . .

Answers for subsection 3.II.2

3.II.2.25 The shadow of a scalar multiple is the scalar multiple of the shadow.

3.II.2.26

(a) Setting $a_0 + (a_0 + a_1)x + (a_2 + a_3)x^3 = 0 + 0x + 0x^2 + 0x^3$ gives $a_0 = 0$ and $a_0 + a_1 = 0$ and $a_2 + a_3 = 0$, so the nullspace is $\{-a_3x^2 + a_3x^3 \mid a_3 \in \mathbb{R}\}$.

(b) Setting $a_0 + (a_0 + a_1)x + (a_2 + a_3)x^3 = 2 + 0x + 0x^2 - x^3$ gives that $a_0 = 2$, and $a_1 = -2$, and $a_2 + a_3 = -1$. Taking a_3 as a parameter, and renaming it $a_3 = a$ gives this set description $\{2 - 2x + (-1 - a)x^2 + ax^3 \mid a \in \mathbb{R}\} = \{(2 - 2x - x^2) + a \cdot (-x^2 + x^3) \mid a \in \mathbb{R}\}$.

(c) This set is empty because the range of h includes only those polynomials with a $0x^2$ term.

3.II.2.29 For any vector space V, the nullspace

$$\{\vec{v} \in V \bigm| 2\vec{v} = \vec{0}\}$$

is trivial, while the rangespace

$$\{\vec{w} \in V \bigm| \vec{w} = 2\vec{v} \text{ for some } \vec{v} \in V\}$$

is all of V, because every vector $\vec{w}$ is twice some other vector, specifically, it is twice $(1/2)\vec{w}$. (Thus, this transformation is actually an automorphism.)

3.II.2.30 Because the rank plus the nullity equals the dimension of the domain (here, five), and the rank is at most three, the possible pairs are: $(3,2)$, $(2,3)$, $(1,4)$, and $(0,5)$. Coming up with linear maps that show that each pair is indeed possible is easy.

3.II.2.31 No (unless $\mathcal{P}_n$ is trivial), because the two polynomials $f_0(x) = 0$ and $f_1(x) = 1$ have the same derivative; a map must be one-to-one to have an inverse.

3.II.2.33

(a) One direction is obvious: if the homomorphism is onto then its range is the codomain and so its rank equals the dimension of its codomain. For the other direction assume that the map's rank equals the dimension of the codomain. Then the map's range is a subspace of the codomain, and has dimension equal to the dimension of the codomain. Therefore, the map's range must equal the codomain, and the map is onto. (The 'therefore' is because there is a linearly independent subset of the range that is of size equal to the dimension of the codomain, but any such linearly independent subset of the codomain must be a basis for the codomain, and so the range equals the codomain.)

(b) By Theorem 3.II.2.20, a homomorphism is one-to-one if and only if its nullity is zero. Because rank plus nullity equals the dimension of the domain, it follows that a homomorphism is one-to-one if and only if its rank equals the dimension of its domain. But this domain and codomain have the same dimension, so the map is one-to-one if and only if it is onto.

3.II.2.34 We are proving that $h\colon V \to W$ is nonsingular if and only if for every linearly independent subset S of V the subset $h(S) = \{h(\vec{s}) \bigm| \vec{s} \in S\}$ of W is linearly independent.

One half is easy—by Theorem 3.II.2.20, if h is singular then its nullspace is nontrivial (contains more than just the zero vector). So, where $\vec{v} \neq \vec{0}_V$ is in that nullspace, the singleton set $\{\vec{v}\}$ is independent while its image $\{h(\vec{v})\} = \{\vec{0}_W\}$ is not.

For the other half, assume that h is nonsingular and so by Theorem 3.II.2.20 has a trivial nullspace. Then for any $\vec{v}_1, \ldots, \vec{v}_n \in V$, the relation

$$\vec{0}_W = c_1 \cdot h(\vec{v}_1) + \cdots + c_n \cdot h(\vec{v}_n) = h(c_1 \cdot \vec{v}_1 + \cdots + c_n \cdot \vec{v}_n)$$

implies the relation $c_1 \cdot \vec{v}_1 + \cdots + c_n \cdot \vec{v}_n = \vec{0}_V$. Hence, if a subset of V is independent then so is its image in W.

Remark. The statement is that a linear map is nonsingular if and only if it preserves independence for *all* sets (that is, if a set is independent then its image is also independent). A singular map may well preserve some independent sets. An example is this singular map from $\mathbb{R}^3$ to $\mathbb{R}^2$.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x+y+z \\ 0 \end{pmatrix}$$

Linear independence is preserved for this set

$$\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\} \mapsto \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$$

and (in a somewhat more tricky example) also for this set

$$\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\} \mapsto \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$$

(recall that in a set, repeated elements do not appear twice). However, there are sets whose independence is not preserved under this map;

$$\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\2\\0\end{pmatrix}\}\mapsto\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}2\\0\end{pmatrix}\}$$

and so not all sets have independence preserved.

3.II.2.35 (We use the notation from Theorem 3.II.1.9.) Fix a basis $\langle\vec{\beta}_1,\ldots,\vec{\beta}_n\rangle$ for V and a basis $\langle\vec{w}_1,\ldots,\vec{w}_k\rangle$ for W. If the dimension k of W is less than or equal to the dimension n of V then the theorem gives a linear map from V to W determined in this way.

$$\vec{\beta}_1\mapsto\vec{w}_1,\ \ldots,\ \vec{\beta}_k\mapsto\vec{w}_k\quad\text{and}\quad\vec{\beta}_{k+1}\mapsto\vec{w}_k,\ \ldots,\ \vec{\beta}_n\mapsto\vec{w}_k$$

We need only to verify that this map is onto.

Any member of W can be written as a linear combination of basis elements $c_1\cdot\vec{w}_1+\cdots+c_k\cdot\vec{w}_k$. This vector is the image, under the map described above, of $c_1\cdot\vec{\beta}_1+\cdots+c_k\cdot\vec{\beta}_k+0\cdot\vec{\beta}_{k+1}\cdots+0\cdot\vec{\beta}_n$. Thus the map is onto.

3.II.2.36 By assumption, h is not the zero map and so a vector $\vec{v}\in V$ exists that is not in the nullspace. Note that $\langle h(\vec{v})\rangle$ is a basis for $\mathbb{R}$, because it is a size one linearly independent subset of $\mathbb{R}$. Consequently h is onto, as for any $r\in\mathbb{R}$ we have $r=c\cdot h(\vec{v})$ for some scalar c, and so $r=h(c\vec{v})$.

Thus the rank of h is one. Because the nullity is given as n, the dimension of the domain of h (the vector space V) is $n+1$. We can finish by showing $\{\vec{v},\vec{\beta}_1,\ldots,\vec{\beta}_n\}$ is linearly independent, as it is a size $n+1$ subset of a dimension $n+1$ space. Because $\{\vec{\beta}_1,\ldots,\vec{\beta}_n\}$ is linearly independent we need only show that $\vec{v}$ is not a linear combination of the other vectors. But $c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n=\vec{v}$ would give $-\vec{v}+c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n=\vec{0}$ and applying h to both sides would give a contradiction.

3.II.2.38 This is a simple calculation.

$$\begin{aligned}h([S])&=\{h(c_1\vec{s}_1+\cdots+c_n\vec{s}_n)\bigm| c_1,\ldots,c_n\in\mathbb{R}\text{ and }\vec{s}_1,\ldots,\vec{s}_n\in S\}\\&=\{c_1h(\vec{s}_1)+\cdots+c_nh(\vec{s}_n)\bigm| c_1,\ldots,c_n\in\mathbb{R}\text{ and }\vec{s}_1,\ldots,\vec{s}_n\in S\}\\&=[h(S)]\end{aligned}$$

3.II.2.40 Because the rank of t is one, the rangespace of t is a one-dimensional set. Taking $\langle h(\vec{v})\rangle$ as a basis (for some appropriate $\vec{v}$), we have that for every $\vec{w}\in V$, the image $h(\vec{w})\in V$ is a multiple of this basis vector—associated with each $\vec{w}$ there is a scalar $c_{\vec{w}}$ such that $t(\vec{w})=c_{\vec{w}}t(\vec{v})$. Apply t to both sides of that equation and take r to be $c_{t(\vec{v})}$

$$t\circ t(\vec{w})=t(c_{\vec{w}}\cdot t(\vec{v}))=c_{\vec{w}}\cdot t\circ t(\vec{v})=c_{\vec{w}}\cdot c_{t(\vec{v})}\cdot t(\vec{v})=c_{\vec{w}}\cdot r\cdot t(\vec{v})=r\cdot c_{\vec{w}}\cdot t(\vec{v})=r\cdot t(\vec{w})$$

to get the desired conclusion.

3.II.2.41 Fix a basis $\langle\vec{\beta}_1,\ldots,\vec{\beta}_n\rangle$ for V. We shall prove that this map

$$h\stackrel{\Phi}{\longmapsto}\begin{pmatrix}h(\vec{\beta}_1)\\\vdots\\h(\vec{\beta}_n)\end{pmatrix}$$

is an isomorphism from V^* to $\mathbb{R}^n$.

To see that Φ is one-to-one, assume that h_1 and h_2 are members of V^* such that $\Phi(h_1)=\Phi(h_2)$. Then

$$\begin{pmatrix}h_1(\vec{\beta}_1)\\\vdots\\h_1(\vec{\beta}_n)\end{pmatrix}=\begin{pmatrix}h_2(\vec{\beta}_1)\\\vdots\\h_2(\vec{\beta}_n)\end{pmatrix}$$

and consequently, $h_1(\vec{\beta}_1)=h_2(\vec{\beta}_1)$, etc. But a homomorphism is determined by its action on a basis, so $h_1=h_2$, and therefore Φ is one-to-one.

To see that Φ is onto, consider

$$\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}$$

for $x_1,\dots,x_n\in\mathbb{R}$. This function h from V to $\mathbb{R}$

$$c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n \stackrel{h}{\longmapsto} c_1x_1+\cdots+c_nx_n$$

is easily seen to be linear, and to be mapped by Φ to the given vector in $\mathbb{R}^n$, so Φ is onto.

The map Φ also preserves structure: where

$$\begin{aligned}c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n &\stackrel{h_1}{\longmapsto} c_1h_1(\vec{\beta}_1)+\cdots+c_nh_1(\vec{\beta}_n)\\ c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n &\stackrel{h_2}{\longmapsto} c_1h_2(\vec{\beta}_1)+\cdots+c_nh_2(\vec{\beta}_n)\end{aligned}$$

we have

$$\begin{aligned}(r_1h_1+r_2h_2)(c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n)&=c_1(r_1h_1(\vec{\beta}_1)+r_2h_2(\vec{\beta}_1))+\cdots+c_n(r_1h_1(\vec{\beta}_n)+r_2h_2(\vec{\beta}_n))\\ &=r_1(c_1h_1(\vec{\beta}_1)+\cdots+c_nh_1(\vec{\beta}_n))+r_2(c_1h_2(\vec{\beta}_1)+\cdots+c_nh_2(\vec{\beta}_n))\end{aligned}$$

so $\Phi(r_1h_1+r_2h_2)=r_1\Phi(h_1)+r_2\Phi(h_2)$.

3.II.2.42 Let $h\colon V\to W$ be linear and fix a basis $\langle\vec{\beta}_1,\dots,\vec{\beta}_n\rangle$ for V. Consider these n maps from V to W

$$h_1(\vec{v})=c_1\cdot h(\vec{\beta}_1),\quad h_2(\vec{v})=c_2\cdot h(\vec{\beta}_2),\quad \dots\quad ,h_n(\vec{v})=c_n\cdot h(\vec{\beta}_n)$$

for any $\vec{v}=c_1\vec{\beta}_1+\cdots+c_n\vec{\beta}_n$. Clearly h is the sum of the h_i's. We need only check that each h_i is linear: where $\vec{u}=d_1\vec{\beta}_1+\cdots+d_n\vec{\beta}_n$ we have $h_i(r\vec{v}+s\vec{u})=rc_i+sd_i=rh_i(\vec{v})+sh_i(\vec{u})$.

3.II.2.43 Either yes (trivially) or no (nearly trivially).

If V 'is homomorphic to' W is taken to mean there is a homomorphism from V into (but not necessarily onto) W, then every space is homomorphic to every other space as a zero map always exists.

If V 'is homomorphic to' W is taken to mean there is an onto homomorphism from V to W then the relation is not an equivalence. For instance, there is an onto homomorphism from $\mathbb{R}^3$ to $\mathbb{R}^2$ (projection is one) but no homomorphism from $\mathbb{R}^2$ onto $\mathbb{R}^3$ by Corollary 2.16, so the relation is not reflexive.*

3.II.2.44 That they form the chains is obvious. For the rest, we show here that $\mathscr{R}(t^{j+1})=\mathscr{R}(t^j)$ implies that $\mathscr{R}(t^{j+2})=\mathscr{R}(t^{j+1})$. Induction then applies.

Assume that $\mathscr{R}(t^{j+1})=\mathscr{R}(t^j)$. Then $t\colon\mathscr{R}(t^{j+1})\to\mathscr{R}(t^{j+2})$ is the same map, with the same domain, as $t\colon\mathscr{R}(t^j)\to\mathscr{R}(t^{j+1})$. Thus it has the same range: $\mathscr{R}(t^{j+2})=\mathscr{R}(t^{j+1})$.

. . . .

Answers for subsection 3.III.1

3.III.1.12
(a) $\begin{pmatrix}2\cdot4+1\cdot2\\3\cdot4-(1/2)\cdot2\end{pmatrix}=\begin{pmatrix}10\\11\end{pmatrix}$ **(b)** $\begin{pmatrix}4\\1\end{pmatrix}$ **(c)** Not defined.

3.III.1.18 Where the space is n-dimensional,

$$\mathrm{Rep}_{B,B}(\mathrm{id})=\begin{pmatrix}1&0\dots&0\\0&1\dots&0\\&\vdots&\\0&0\dots&1\end{pmatrix}_{B,B}$$

is the $n\times n$ identity matrix.

3.III.1.19 Taking this as the natural basis

$$B=\langle\vec{\beta}_1,\vec{\beta}_2,\vec{\beta}_3,\vec{\beta}_4\rangle=\langle\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\rangle$$

the transpose map acts in this way

$$\vec{\beta}_1\mapsto\vec{\beta}_1\quad\vec{\beta}_2\mapsto\vec{\beta}_3\quad\vec{\beta}_3\mapsto\vec{\beta}_2\quad\vec{\beta}_4\mapsto\vec{\beta}_4$$

*More information on equivalence relations is in the appendix.

so that representing the images with respect to the codomain's basis and adjoining those column vectors together gives this.

$$\mathrm{Rep}_{B,B}(\text{trans}) = \begin{pmatrix} 1&0&0&0 \\ 0&0&1&0 \\ 0&1&0&0 \\ 0&0&0&1 \end{pmatrix}_{B,B}$$

3.III.1.20

(a) With respect to the basis of the codomain, the images of the members of the basis of the domain are represented as

$$\mathrm{Rep}_B(\vec{\beta}_2) = \begin{pmatrix} 0\\1\\0\\0 \end{pmatrix} \quad \mathrm{Rep}_B(\vec{\beta}_3) = \begin{pmatrix} 0\\0\\1\\0 \end{pmatrix} \quad \mathrm{Rep}_B(\vec{\beta}_4) = \begin{pmatrix} 0\\0\\0\\1 \end{pmatrix} \quad \mathrm{Rep}_B(\vec{0}) = \begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}$$

and consequently, the matrix representing the transformation is this.

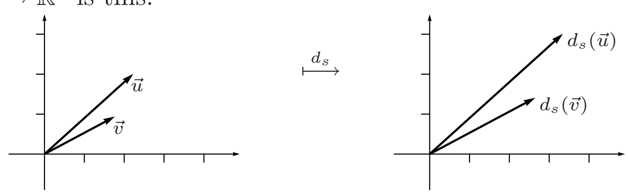
$$\begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \end{pmatrix}$$

(b) $\begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 0&0&0&0 \\ 0&0&1&0 \end{pmatrix}$

(c) $\begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&0 \end{pmatrix}$

3.III.1.21

(a) The picture of $d_s\colon \mathbb{R}^2 \to \mathbb{R}^2$ is this.



This map's effect on the vectors in the standard basis for the domain is

$$\begin{pmatrix} 1\\0 \end{pmatrix} \overset{d_s}{\longmapsto} \begin{pmatrix} s\\0 \end{pmatrix} \qquad \begin{pmatrix} 0\\1 \end{pmatrix} \overset{d_s}{\longmapsto} \begin{pmatrix} 0\\s \end{pmatrix}$$

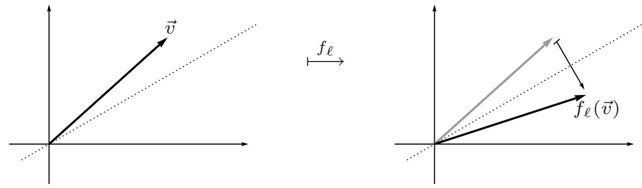
and those images are represented with respect to the codomain's basis (again, the standard basis) by themselves.

$$\mathrm{Rep}_{\mathcal{E}_2}(\begin{pmatrix} s\\0 \end{pmatrix}) = \begin{pmatrix} s\\0 \end{pmatrix} \qquad \mathrm{Rep}_{\mathcal{E}_2}(\begin{pmatrix} 0\\s \end{pmatrix}) = \begin{pmatrix} 0\\s \end{pmatrix}$$

Thus the representation of the dilation map is this.

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(d_s) = \begin{pmatrix} s&0 \\ 0&s \end{pmatrix}$$

(b) The picture of $f_\ell\colon \mathbb{R}^2 \to \mathbb{R}^2$ is this.



Some calculation (see Exercise I.29) shows that when the line has slope k

$$\begin{pmatrix}1\\0\end{pmatrix}\overset{f_\ell}{\longmapsto}\begin{pmatrix}(1-k^2)/(1+k^2)\\2k/(1+k^2)\end{pmatrix}\qquad\begin{pmatrix}0\\1\end{pmatrix}\overset{f_\ell}{\longmapsto}\begin{pmatrix}2k/(1+k^2)\\-(1-k^2)/(1+k^2)\end{pmatrix}$$

(the case of a line with undefined slope is separate but easy) and so the matrix representing reflection is this.

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_\ell)=\frac{1}{1+k^2}\cdot\begin{pmatrix}1-k^2&2k\\2k&-(1-k^2)\end{pmatrix}$$

3.III.1.23

(a) The images of the members of the domain's basis are

$$\vec{\beta}_1\mapsto h(\vec{\beta}_1)\quad\vec{\beta}_2\mapsto h(\vec{\beta}_2)\quad\ldots\quad\vec{\beta}_n\mapsto h(\vec{\beta}_n)$$

and those images are represented with respect to the codomain's basis in this way.

$$\mathrm{Rep}_{h(B)}(\,h(\vec{\beta}_1)\,)=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\quad\mathrm{Rep}_{h(B)}(\,h(\vec{\beta}_2)\,)=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}\quad\ldots\quad\mathrm{Rep}_{h(B)}(\,h(\vec{\beta}_n)\,)=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$$

Hence, the matrix is the identity.

$$\mathrm{Rep}_{B,h(B)}(h)=\begin{pmatrix}1&0&\ldots&0\\0&1&&0\\&&\ddots&\\0&0&&1\end{pmatrix}$$

(b) Using the matrix in the prior item, the representation is this.

$$\mathrm{Rep}_{h(B)}(\,h(\vec{v})\,)=\begin{pmatrix}c_1\\\vdots\\c_n\end{pmatrix}_{h(B)}$$

3.III.1.24 The product

$$\begin{pmatrix}h_{1,1}&\ldots&h_{1,i}&\ldots&h_{1,n}\\h_{2,1}&\ldots&h_{2,i}&\ldots&h_{2,n}\\&\vdots&&&\\h_{m,1}&\ldots&h_{m,i}&\ldots&h_{1,n}\end{pmatrix}\begin{pmatrix}0\\\vdots\\1\\\vdots\\0\end{pmatrix}=\begin{pmatrix}h_{1,i}\\h_{2,i}\\\vdots\\h_{m,i}\end{pmatrix}$$

gives the i-th column of the matrix.

3.III.1.26

(a) It is the set of vectors of the codomain represented with respect to the codomain's basis in this way.

$$\{\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\mid x,y\in\mathbb{R}\}=\{\begin{pmatrix}x\\0\end{pmatrix}\mid x,y\in\mathbb{R}\}$$

As the codomain's basis is $\mathcal{E}_2$, and so each vector is represented by itself, the range of this transformation is the x-axis.

(b) It is the set of vectors of the codomain represented in this way.

$$\{\begin{pmatrix}0&0\\3&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\mid x,y\in\mathbb{R}\}=\{\begin{pmatrix}0\\3x+2y\end{pmatrix}\mid x,y\in\mathbb{R}\}$$

With respect to $\mathcal{E}_2$ vectors represent themselves so this range is the y axis.

(c) The set of vectors represented with respect to $\mathcal{E}_2$ as

$$\{\begin{pmatrix}a&b\\2a&2b\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\mid x,y\in\mathbb{R}\}=\{\begin{pmatrix}ax+by\\2ax+2by\end{pmatrix}\mid x,y\in\mathbb{R}\}=\{(ax+by)\cdot\begin{pmatrix}1\\2\end{pmatrix}\mid x,y\in\mathbb{R}\}$$

is the line $y=2x$, provided either a or b is not zero, and is the set consisting of just the origin if both are zero.

3.III.1.28 We mimic Example 1.1, just replacing the numbers with letters.

Write B as $\langle\vec{\beta}_1,\dots,\vec{\beta}_n\rangle$ and D as $\langle\vec{\delta}_1,\dots,\vec{\delta}_m\rangle$. By definition of representation of a map with respect to bases, the assumption that

$$\mathrm{Rep}_{B,D}(h)=\begin{pmatrix}h_{1,1}&\dots&h_{1,n}\\ \vdots&&\vdots\\ h_{m,1}&\dots&h_{m,n}\end{pmatrix}$$

means that $h(\vec{\beta}_i)=h_{i,1}\vec{\delta}_1+\dots+h_{i,n}\vec{\delta}_n$. And, by the definition of the representation of a vector with respect to a basis, the assumption that

$$\mathrm{Rep}_B(\vec{v})=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}$$

means that $\vec{v}=c_1\vec{\beta}_1+\dots+c_n\vec{\beta}_n$. Substituting gives

$$\begin{aligned}h(\vec{v})&=h(c_1\cdot\vec{\beta}_1+\dots+c_n\cdot\vec{\beta}_n)\\ &=c_1\cdot h(\vec{\beta}_1)+\dots+c_n\cdot\vec{\beta}_n\\ &=c_1\cdot(h_{1,1}\vec{\delta}_1+\dots+h_{m,1}\vec{\delta}_m)+\dots+c_n\cdot(h_{1,n}\vec{\delta}_1+\dots+h_{m,n}\vec{\delta}_m)\\ &=(h_{1,1}c_1+\dots+h_{1,n}c_n)\cdot\vec{\delta}_1+\dots+(h_{m,1}c_1+\dots+h_{m,n}c_n)\cdot\vec{\delta}_m\end{aligned}$$

and so $h(\vec{v})$ is represented as required.

3.III.1.30

(a) Write B_U as $\langle\vec{\beta}_1,\dots,\vec{\beta}_k\rangle$ and then B_V as $\langle\vec{\beta}_1,\dots,\vec{\beta}_k,\vec{\beta}_{k+1},\dots,\vec{\beta}_n\rangle$. If

$$\mathrm{Rep}_{B_U}(\vec{v})=\begin{pmatrix}c_1\\ \vdots\\ c_k\end{pmatrix}\qquad\text{so that }\vec{v}=c_1\cdot\vec{\beta}_1+\dots+c_k\cdot\vec{\beta}_k$$

then,

$$\mathrm{Rep}_{B_V}(\vec{v})=\begin{pmatrix}c_1\\ \vdots\\ c_k\\ 0\\ \vdots\\ 0\end{pmatrix}$$

because $\vec{v}=c_1\cdot\vec{\beta}_1+\dots+c_k\cdot\vec{\beta}_k+0\cdot\vec{\beta}_{k+1}+\dots+0\cdot\vec{\beta}_n$.

(b) We must first decide what the question means. Compare $h\colon V\to W$ with its restriction to the subspace $h\restriction_U\colon U\to W$. The rangespace of the restriction is a subspace of W, so fix a basis $D_{h(U)}$ for this rangespace and extend it to a basis D_V for W. We want the relationship between these two.

$$\mathrm{Rep}_{B_V,D_V}(h)\quad\text{and}\quad\mathrm{Rep}_{B_U,D_{h(U)}}(h\restriction_U)$$

The answer falls right out of the prior item: if

$$\mathrm{Rep}_{B_U,D_{h(U)}}(h\restriction_U)=\begin{pmatrix}h_{1,1}&\dots&h_{1,k}\\ \vdots&&\vdots\\ h_{p,1}&\dots&h_{p,k}\end{pmatrix}$$

then the extension is represented in this way.

$$\mathrm{Rep}_{B_V,D_V}(h)=\begin{pmatrix}h_{1,1}&\dots&h_{1,k}&h_{1,k+1}&\dots&h_{1,n}\\ \vdots&&&&&\vdots\\ h_{p,1}&\dots&h_{p,k}&h_{p,k+1}&\dots&h_{p,n}\\ 0&\dots&0&h_{p+1,k+1}&\dots&h_{p+1,n}\\ \vdots&&&&&\vdots\\ 0&\dots&0&h_{m,k+1}&\dots&h_{m,n}\end{pmatrix}$$

(c) Take W_i to be the span of $\{h(\vec{\beta}_1),\ldots,h(\vec{\beta}_i)\}$.
(d) Apply the answer from the second item to the third item.
(e) No. For instance $\pi_x\colon\mathbb{R}^2\to\mathbb{R}^2$, projection onto the x axis, is represented by these two upper-triangular matrices

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(\pi_x)=\begin{pmatrix}1&0\\0&0\end{pmatrix}\quad\text{and}\quad\mathrm{Rep}_{C,\mathcal{E}_2}(\pi_x)=\begin{pmatrix}0&1\\0&0\end{pmatrix}$$

where $C=\langle\vec{e}_2,\vec{e}_1\rangle$.

. . . .

Answers for subsection 3.III.2

3.III.2.12 A general member of the domain, represented with respect to the domain's basis as

$$a\cos\theta+b\sin\theta=\begin{pmatrix}a\\a+b\end{pmatrix}_B$$

is mapped to

$$\begin{pmatrix}0\\a\end{pmatrix}_D\quad\text{representing}\quad 0\cdot(\cos\theta+\sin\theta)+a\cdot(\cos\theta)$$

and so the linear map represented by the matrix with respect to these bases

$$a\cos\theta+b\sin\theta\mapsto a\cos\theta$$

is projection onto the first component.

3.III.2.14 Let the matrix be G, and suppose that it rperesents $g\colon V\to W$ with respect to bases B and D. Because G has two columns, V is two-dimensional. Because G has two rows, W is two-dimensional. The action of g on a general member of the domain is this.

$$\begin{pmatrix}x\\y\end{pmatrix}_B\mapsto\begin{pmatrix}x+2y\\3x+6y\end{pmatrix}_D$$

(a) The only representation of the zero vector in the codomain is

$$\mathrm{Rep}_D(\vec{0})=\begin{pmatrix}0\\0\end{pmatrix}_D$$

and so the set of representations of members of the nullspace is this.

$$\{\begin{pmatrix}x\\y\end{pmatrix}_B\bigm| x+2y=0\text{ and }3x+6y=0\}=\{y\cdot\begin{pmatrix}-1/2\\1\end{pmatrix}_D\bigm| y\in\mathbb{R}\}$$

(b) The representation map $\mathrm{Rep}_D\colon W\to\mathbb{R}^2$ and its inverse are isomorphisms, and so preserve the dimension of subspaces. The subspace of $\mathbb{R}^2$ that is in the prior item is one-dimensional. Therefore, the image of that subspace under the inverse of the representation map—the nullspace of G, is also one-dimensional.
(c) The set of representations of members of the rangespace is this.

$$\{\begin{pmatrix}x+2y\\3x+6y\end{pmatrix}_D\bigm| x,y\in\mathbb{R}\}=\{k\cdot\begin{pmatrix}1\\3\end{pmatrix}_D\bigm| k\in\mathbb{R}\}$$

(d) Of course, Theorem 3.III.2.3 gives that the rank of the map equals the rank of the matrix, which is one. Alternatively, the same argument that was used above for the nullspace gives here that the dimension of the rangespace is one.
(e) One plus one equals two.

3.III.2.17 Yes. Consider

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

representing a map from $\mathbb{R}^2$ to $\mathbb{R}^2$. With respect to the standard bases $B_1 = \mathcal{E}_2, D_1 = \mathcal{E}_2$ this matrix represents the identity map. With respect to

$$B_2 = D_2 = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle$$

this matrix again represents the identity. In fact, as long as the starting and ending bases are equal—as long as $B_i = D_i$—then the map represented by H is the identity.

3.III.2.19 The first map

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{E}_2} \mapsto \begin{pmatrix} 3x \\ 2y \end{pmatrix}_{\mathcal{E}_2} = \begin{pmatrix} 3x \\ 2y \end{pmatrix}$$

stretches vectors by a factor of three in the x direction and by a factor of two in the y direction. The second map

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{E}_2} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}_{\mathcal{E}_2} = \begin{pmatrix} x \\ 0 \end{pmatrix}$$

projects vectors onto the x axis. The third

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{E}_2} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}_{\mathcal{E}_2} = \begin{pmatrix} y \\ x \end{pmatrix}$$

interchanges first and second components (that is, it is a reflection about the line $y = x$). The last

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{E}_2} \mapsto \begin{pmatrix} x+3y \\ y \end{pmatrix}_{\mathcal{E}_2} = \begin{pmatrix} x+3y \\ y \end{pmatrix}$$

stretches vectors parallel to the y axis, by an amount equal to three times their distance from that axis (this is a *skew*.)

3.III.2.20

(a) This is immediate from Theorem 3.III.2.3.

(b) Yes. This is immediate from the prior item.

To give a specific example, we can start with $\mathcal{E}_3$ as the basis for the domain, and then we require a basis D for the codomain $\mathbb{R}^3$. The matrix H gives the action of the map as this

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}_{\mathcal{E}_3} \mapsto \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}_D \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}_{\mathcal{E}_3} \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_D \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_{\mathcal{E}_3} \mapsto \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}_D$$

and there is no harm in finding a basis D so that

$$\mathrm{Rep}_D(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}_D \quad \text{and} \quad \mathrm{Rep}_D(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_D$$

that is, so that the map represented by H with respect to $\mathcal{E}_3, D$ is projection down onto the xy plane. The second condition gives that the third member of D is $\vec{e}_2$. The first condition gives that the first member of D plus twice the second equals $\vec{e}_1$, and so this basis will do.

$$D = \langle \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle$$

3.III.2.21

(a) Recall that the representation map $\mathrm{Rep}_B \colon V \to \mathbb{R}^n$ is linear (it is actually an isomorphism, but we do not need that it is one-to-one or onto here). Considering the column vector x to be a $n \times 1$ matrix gives that the map from $\mathbb{R}^n$ to $\mathbb{R}$ that takes a column vector to its dot product with $\vec{x}$ is linear (this is a matrix-vector product and so Theorem 3.III.2.1 applies). Thus the map under consideration $h_{\vec{x}}$ is linear because it is the composistion of two linear maps.

$$\vec{v} \mapsto \mathrm{Rep}_B(\vec{v}) \mapsto \vec{x} \cdot \mathrm{Rep}_B(\vec{v})$$

(b) Any linear map $g\colon V \to \mathbb{R}$ is represented by some matrix

$$\begin{pmatrix} g_1 & g_2 & \cdots & g_n \end{pmatrix}$$

(the matrix has n columns because V is n-dimensional and it has only one row because $\mathbb{R}$ is one-dimensional). Then taking $\vec{x}$ to be the column vector that is the transpose of this matrix

$$\vec{x} = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}$$

has the desired action.

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = g_1 v_1 + \cdots + g_n v_n$$

(c) No. If $\vec{x}$ has any nonzero entries then $h_{\vec{x}}$ cannot be the zero map (and if $\vec{x}$ is the zero vector then $h_{\vec{x}}$ can only be the zero map).

3.III.2.22 See the following section.

. . . .

Answers for subsection 3.IV.1

3.IV.1.8 Represent the domain vector $\vec{v} \in V$ and the maps $g, h\colon V \to W$ with respect to bases B, D in the usual way.

(a) The representation of $(g+h)\,(\vec{v}) = g(\vec{v}) + h(\vec{v})$

$$\begin{aligned}\big((g_{1,1}v_1 + \cdots + g_{1,n}v_n)\vec{\delta}_1 + \cdots + (g_{m,1}v_1 + \cdots + g_{m,n}v_n)\vec{\delta}_m\big)& \\ + \big((h_{1,1}v_1 + \cdots + h_{1,n}v_n)\vec{\delta}_1 + \cdots + (h_{m,1}v_1 + \cdots + h_{m,n}v_n)\vec{\delta}_m\big)&\end{aligned}$$

regroups

$$= ((g_{1,1} + h_{1,1})v_1 + \cdots + (g_{1,1} + h_{1,n})v_n) \cdot \vec{\delta}_1 + \cdots + ((g_{m,1} + h_{m,1})v_1 + \cdots + (g_{m,n} + h_{m,n})v_n) \cdot \vec{\delta}_m$$

to the entry-by-entry sum of the representation of $g(\vec{v})$ and the representation of $h(\vec{v})$.

(b) The representation of $(r \cdot h)\,(\vec{v}) = r \cdot \big(h(\vec{v})\big)$

$$\begin{aligned}r \cdot \big((h_{1,1}v_1 + h_{1,2}v_2 + \cdots + h_{1,n}v_n)\vec{\delta}_1 + \cdots + (h_{m,1}v_1 + h_{m,2}v_2 + \cdots + h_{m,n}v_n)\vec{\delta}_m\big)& \\ = (rh_{1,1}v_1 + \cdots + rh_{1,n}v_n) \cdot \vec{\delta}_1 + \cdots + (rh_{m,1}v_1 + \cdots + rh_{m,n}v_n) \cdot \vec{\delta}_m&\end{aligned}$$

is the entry-by-entry multiple of r and the representation of h.

3.IV.1.10 For any V, W with bases B, D, the (appropriately-sized) zero matrix represents this map.

$$\vec{\beta}_1 \mapsto 0 \cdot \vec{\delta}_1 + \cdots + 0 \cdot \vec{\delta}_m \quad \cdots \quad \vec{\beta}_n \mapsto 0 \cdot \vec{\delta}_1 + \cdots + 0 \cdot \vec{\delta}_m$$

This is the zero map.

There are no other matrices that represent only one map. For, suppose that H is not the zero matrix. Then it has a nonzero entry; assume that $h_{i,j} \neq 0$. With respect to bases B, D, it represents $h_1\colon V \to W$ sending

$$\vec{\beta}_j \mapsto h_{1,j}\vec{\delta}_1 + \cdots + h_{i,j}\vec{\delta}_i + \cdots + h_{m,j}\vec{\delta}_m$$

and with respcet to $B, 2 \cdot D$ it also represents $h_2\colon V \to W$ sending

$$\vec{\beta}_j \mapsto h_{1,j} \cdot (2\vec{\delta}_1) + \cdots + h_{i,j} \cdot (2\vec{\delta}_i) + \cdots + h_{m,j} \cdot (2\vec{\delta}_m)$$

(the notation $2 \cdot D$ means to double all of the members of D). These maps are easily seen to be unequal.

3.IV.1.13 That the trace of a sum is the sum of the traces holds because both $\text{trace}(H+G)$ and $\text{trace}(H) + \text{trace}(G)$ are the sum of $h_{1,1} + g_{1,1}$ with $h_{2,2} + g_{2,2}$, etc. For scalar multiplication we have $\text{trace}(r \cdot H) = r \cdot \text{trace}(H)$; the proof is easy. Thus the trace map is a homomorphism from $\mathcal{M}_{n \times n}$ to $\mathbb{R}$.

3.IV.1.14

(a) The i,j entry of $(G+H)^{\text{trans}}$ is $g_{j,i}+h_{j,i}$. That is also the i,j entry of $G^{\text{trans}}+H^{\text{trans}}$.
(b) The i,j entry of $(r\cdot H)^{\text{trans}}$ is $rh_{j,i}$, which is also the i,j entry of $r\cdot H^{\text{trans}}$.

. . . .

Answers for subsection 3.IV.2

3.IV.2.16
(a) Yes. **(b)** Yes. **(c)** No. **(d)** No.

3.IV.2.19 Technically, no. The dot product operation yields a scalar while the matrix product yields a 1×1 matrix. However, we usually will ignore the distinction.

3.IV.2.21 It is true for all one-dimensional spaces. Let f,g be transformations of a one-dimensional space. We must show that $g\circ f\,(\vec{v})=f\circ g\,(\vec{v})$ for all vectors. Fix a basis B for the space and then the transformations are represented by 1×1 matrices.

$$F=\text{Rep}_{B,B}(f)=\begin{pmatrix}f_{1,1}\end{pmatrix}\qquad G=\text{Rep}_{B,B}(g)=\begin{pmatrix}g_{1,1}\end{pmatrix}$$

Therefore, the compositions can be represented as GF and FG.

$$GF=\text{Rep}_{B,B}(g\circ f)=\begin{pmatrix}g_{1,1}f_{1,1}\end{pmatrix}\qquad FG=\text{Rep}_{B,B}(f\circ g)=\begin{pmatrix}f_{1,1}g_{1,1}\end{pmatrix}$$

These two matrices are equal and so the compositions have the same effect on each vector in the space.

3.IV.2.22 It would not represent linear map composition; Theorem 3.IV.2.6 would fail.

3.IV.2.25 We have not seen a map interpretation of the transpose operation, so we will verify these by considering the entries.
(a) The i,j entry of GH^{trans} is the j,i entry of GH, which is the dot product of the j-th row of G and the i-th column of H. The i,j entry of $H^{\text{trans}}G^{\text{trans}}$ is the dot product of the i-th row of H^{trans} and the j-th column of G^{trans}, which is the the dot product of the i-th column of H and the j-th row of G. Dot product is commutative and so these two are equal.
(b) By the prior item each equals its transpose, e.g., $(HH^{\text{trans}})^{\text{trans}}=H^{\text{trans}^{\text{trans}}}H^{\text{trans}}=HH^{\text{trans}}$.

3.IV.2.27 It doesn't matter (as long as the spaces have the appropriate dimensions).

For associativity, suppose that F is $m\times r$, that G is $r\times n$, and that H is $n\times k$. We can take any r dimensional space, any m dimensional space, any n dimensional space, and any k dimensional space—for instance, $\mathbb{R}^r$, $\mathbb{R}^m$, $\mathbb{R}^n$, and $\mathbb{R}^k$ will do. We can take any bases A, B, C, and D, for those spaces. Then, with respect to C,D the matrix H represents a linear map h, with respect to B,C the matrix G represents a g, and with respect to A,B the matrix F represents an f. We can use those maps in the proof.

The second half is done similarly, except that G and H are added and so we must take them to represent maps with the same domain and codomain.

3.IV.2.28
(a) The product of rank n matrices can have rank less than or equal to n but not greater than n.

To see that the rank can fall, consider the maps $\pi_x,\pi_y\colon\mathbb{R}^2\to\mathbb{R}^2$ projecting onto the axes. Each is rank one but their composition $\pi_x\circ\pi_y$, which is the zero map, is rank zero. That can be translated over to matrices representing those maps in this way.

$$\text{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(\pi_x)\cdot\text{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(\pi_y)=\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}$$

To prove that the product of rank n matrices cannot have rank greater than n, we can apply the map result that the image of a linearly dependent set is linearly dependent. That is, if $h\colon V\to W$ and $g\colon W\to X$ both have rank n then a set in the range $\mathscr{R}(g\circ h)$ of size larger than n is the image under g of a set in W of size larger than n and so is linearly dependent (since the rank of h is n). Now, the image of a linearly dependent set is dependent, so any set of size larger than n in the range is dependent. (By the way, observe that the rank of g was not mentioned. See the next part.)

(b) Fix spaces and bases and consider the associated linear maps f and g. Recall that the dimension of the image of a map (the map's rank) is less than or equal to the dimension of the domain, and consider the arrow diagram.

$$V \overset{f}{\longmapsto} \mathscr{R}(f) \overset{g}{\longmapsto} \mathscr{R}(g \circ f)$$

First, the image of $\mathscr{R}(f)$ must have dimension less than or equal to the dimension of $\mathscr{R}(f)$, by the prior sentence. On the other hand, $\mathscr{R}(f)$ is a subset of the domain of g, and thus its image has dimension less than or equal the dimension of the domain of g. Combining those two, the rank of a composition is less than or equal to the minimum of the two ranks.

The matrix fact follows immediately.

3.IV.2.29 The 'commutes with' relation is reflexive and symmetric. However, it is not transitive: for instance, with

$$G = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad J = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

G commutes with H and H commutes with J, but G does not commute with J.

3.IV.2.31 Note that $(S+T)(S-T) = S^2 - ST + TS - T^2$, so a reasonable try is to look at matrices that do not commute so that $-ST$ and TS don't cancel: with

$$S = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad T = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

we have the desired inequality.

$$(S+T)(S-T) = \begin{pmatrix} -56 & -56 \\ -88 & -88 \end{pmatrix} \qquad S^2 - T^2 = \begin{pmatrix} -60 & -68 \\ -76 & -84 \end{pmatrix}$$

3.IV.2.33 Here are four solutions.

$$T = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$$

3.IV.2.34

(a) The vector space $\mathcal{M}_{2\times 2}$ has dimension four. The set $\{T^4, \dots, T, I\}$ has five elements and thus is linearly dependent.

(b) Where T is $n\times n$, generalizing the argument from the prior item shows that there is such a polynomial of degree n^2 or less, since $\{T^{n^2}, \dots, T, I\}$ is a n^2+1-member subset of the n^2-dimensional space $\mathcal{M}_{n\times n}$.

(c) First compute the powers

$$T^2 = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} \qquad T^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad T^4 = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$$

(observe that rotating by $\pi/6$ three times results in a rotation by $\pi/2$, which is indeed what T^3 represents). Then set $c_4T^4 + c_3T^3 + c_2T^2 + c_1T + c_0I$ equal to the zero matrix

$$\begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix} c_4 + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} c_3 + \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} c_2 + \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix} c_1 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} c_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

to get this linear system.

$$\begin{array}{rcl}
-(1/2)c_4 \qquad\; + \;\; (1/2)c_2 + (\sqrt{3}/2)c_1 + c_0 &=& 0 \\
-(\sqrt{3}/2)c_4 - c_3 - (\sqrt{3}/2)c_2 - \;\; (1/2)c_1 \qquad &=& 0 \\
(\sqrt{3}/2)c_4 + c_3 + (\sqrt{3}/2)c_2 + \;\; (1/2)c_1 \qquad &=& 0 \\
-(1/2)c_4 \qquad\; + \;\; (1/2)c_2 + (\sqrt{3}/2)c_1 + c_0 &=& 0
\end{array}$$

Apply Gaussian reduction.

$$\xrightarrow[]{-\rho_1+\rho_4}\;\xrightarrow[]{\rho_2+\rho_3}\;
\begin{array}{rcl}
-(1/2)c_4 \qquad\; + \;\; (1/2)c_2 + (\sqrt{3}/2)c_1 + c_0 &=& 0 \\
-(\sqrt{3}/2)c_4 - c_3 - (\sqrt{3}/2)c_2 - \;\; (1/2)c_1 \qquad &=& 0 \\
0 &=& 0 \\
0 &=& 0
\end{array}$$

$$\xrightarrow[]{-\sqrt{3}\rho_1+\rho_2}\;
\begin{array}{rcl}
-(1/2)c_4 \qquad + (1/2)c_2 + (\sqrt{3}/2)c_1 + \quad c_0 &=& 0 \\
- c_3 - \quad \sqrt{3}c_2 - \qquad 2c_1 - \sqrt{3}c_0 &=& 0 \\
0 &=& 0 \\
0 &=& 0
\end{array}$$

Setting c_4, c_3, and c_2 to zero makes c_1 and c_0 also come out to be zero so no degree one or degree zero polynomial will do. Setting c_4 and c_3 to zero (and c_2 to one) gives a linear system

$$\begin{aligned} (1/2) + (\sqrt{3}/2)c_1 + \quad c_0 &= 0 \\ -\sqrt{3} - \qquad 2c_1 - \sqrt{3}c_0 &= 0 \end{aligned}$$

that can be solved with $c_1 = -\sqrt{3}$ and $c_0 = 1$. Conclusion: the polynomial $m(x) = x^2 - \sqrt{3}x + 1$ is minimal for the matrix T.

3.IV.2.35 The check is routine:

$$a_0 + a_1x + \cdots + a_nx^n \stackrel{s}{\longmapsto} a_0x + a_1x^2 + \cdots + a_nx^{n+1} \stackrel{d/dx}{\longmapsto} a_0 + 2a_1x + \cdots + (n+1)a_nx^n$$

while

$$a_0 + a_1x + \cdots + a_nx^n \stackrel{d/dx}{\longmapsto} a_1 + \cdots + na_nx^{n-1} \stackrel{s}{\longmapsto} a_1x + \cdots + a_nx^n$$

so that under the map $(d/dx \circ s) - (s \circ d/dx)$ we have $a_0 + a_1x + \cdots + a_nx^n \mapsto a_0 + a_1x + \cdots + a_nx^n$.

3.IV.2.36

(a) Tracing through the remark at the end of the subsection gives that the i, j entry of $(FG)H$ is this

$$\begin{aligned} \sum_{t=1}^{s}\Big(\sum_{k=1}^{r} f_{i,k}g_{k,t}\Big)h_{t,j} = \sum_{t=1}^{s}\sum_{k=1}^{r}(f_{i,k}g_{k,t})h_{t,j} &= \sum_{t=1}^{s}\sum_{k=1}^{r} f_{i,k}(g_{k,t}h_{t,j}) \\ &= \sum_{k=1}^{r}\sum_{t=1}^{s} f_{i,k}(g_{k,t}h_{t,j}) = \sum_{k=1}^{r} f_{i,k}\Big(\sum_{t=1}^{s} g_{k,t}h_{t,j}\Big) \end{aligned}$$

(the first equality comes from using the distributive law to multiply through the h's, the second equality is the associative law for real numbers, the third is the commutative law for reals, and the fourth equality follows on using the distributive law to factor the f's out), which is the i, j entry of $F(GH)$.

(b) The k-th component of $h(\vec{v})$ is

$$\sum_{j=1}^{n} h_{k,j}v_j$$

and so the i-th component of $g \circ h\,(\vec{v})$ is this

$$\sum_{k=1}^{r} g_{i,k}\Big(\sum_{j=1}^{n} h_{k,j}v_j\Big) = \sum_{k=1}^{r}\sum_{j=1}^{n} g_{i,k}h_{k,j}v_j = \sum_{k=1}^{r}\sum_{j=1}^{n}(g_{i,k}h_{k,j})v_j = \sum_{j=1}^{n}\sum_{k=1}^{r}(g_{i,k}h_{k,j})v_j = \sum_{j=1}^{n}\Big(\sum_{k=1}^{r} g_{i,k}h_{k,j}\Big)\,v_j$$

(the first equality holds by using the distributive law to multiply the g's through, the second equality represents the use of associativity of reals, the third follows by commutativity of reals, and the fourth comes from using the distributive law to factor the v's out).

. . . .

Answers for subsection 3.IV.3

3.IV.3.26 The product is the identity matrix (recall that $\cos^2\theta + \sin^2\theta = 1$). An explanation is that the given matrix represents, with respect to the standard bases, a rotation in $\mathbb{R}^2$ of θ radians while the transpose represents a rotation of $-\theta$ radians. The two cancel.

3.IV.3.28 No. In $\mathcal{P}_1$, with respect to the unequal bases $B = \langle 1, x\rangle$ and $D = \langle 1+x, 1-x\rangle$, the identity transformation is represented by by this matrix.

$$\mathrm{Rep}_{B,D}(\mathrm{id}) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}_{B,D}$$

3.IV.3.29 For any scalar r and square matrix H we have $(rI)H = r(IH) = rH = r(HI) = (Hr)I = H(rI)$.

There are no other such matrices; here is an argument for 2×2 matrices that is easily extended to $n\times n$. If a matrix commutes with all others then it commutes with this unit matrix.

$$\begin{pmatrix}0&a\\0&c\end{pmatrix}=\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}0&1\\0&0\end{pmatrix}=\begin{pmatrix}0&1\\0&0\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}c&d\\0&0\end{pmatrix}$$

From this we first conclude that the upper left entry a must equal its lower right entry d. We also conclude that the lower left entry c is zero. The argument for the upper right entry b is similar.

3.IV.3.30 It is false; these two don't commute.

$$\begin{pmatrix}1&0\\0&0\end{pmatrix}\qquad\begin{pmatrix}0&0\\1&0\end{pmatrix}$$

3.IV.3.32 The generalization is to go from the first and second rows to the i_1-th and i_2-th rows. Row i of GH is made up of the dot products of row i of G and the columns of H. Thus if rows i_1 and i_2 of G are equal then so are rows i_1 and i_2 of GH.

3.IV.3.33 If the product of two diagonal matrices is defined—if both are $n\times n$—then the product of the diagonals is the diagonal of the products: where G,H are equal-sized diagonal matrices, GH is all zeros except each that i,i entry is $g_{i,i}h_{i,i}$.

3.IV.3.34 One way to produce this matrix from the identity is to use the column operations of first multiplying the second column by three, and then adding the negative of the resulting second column to the first.

$$\begin{pmatrix}1&0\\0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0\\0&3\end{pmatrix}\longrightarrow\begin{pmatrix}1&0\\-3&3\end{pmatrix}$$

Column operations, in contrast with row operations) are written from left to right, so doing the above two operations is expressed with this matrix product.

$$\begin{pmatrix}1&0\\0&3\end{pmatrix}\begin{pmatrix}1&0\\-1&1\end{pmatrix}$$

Remark. Alternatively, we could get the required matrix with row operations. Starting with the identity, first adding the negative of the first row to the second, and then multiplying the second row by three will work. Because successive row operations are written as matrix products from right to left, doing these two row operations is expressed with: the same matrix product.

3.IV.3.36 Perhaps the easiest way is to show that each $n\times m$ matrix is a linear combination of unit matrices in one and only one way:

$$c_1\begin{pmatrix}1&0&\ldots&\\0&0&&\\\vdots&&&\end{pmatrix}+\cdots+c_{n,m}\begin{pmatrix}0&0&\ldots&\\\vdots&&&\\0&\ldots&&1\end{pmatrix}=\begin{pmatrix}a_{1,1}&a_{1,2}&\ldots&\\\vdots&&&\\a_{n,1}&\ldots&&a_{n,m}\end{pmatrix}$$

has the unique solution $c_1=a_{1,1}$, $c_2=a_{1,2}$, etc.

3.IV.3.37 Call that matrix F. We have

$$F^2=\begin{pmatrix}2&1\\1&1\end{pmatrix}\quad F^3=\begin{pmatrix}3&2\\2&1\end{pmatrix}\quad F^4=\begin{pmatrix}5&3\\3&2\end{pmatrix}$$

In general,

$$F^n=\begin{pmatrix}f_{n+1}&f_n\\f_n&f_{n-1}\end{pmatrix}$$

where f_i is the i-th Fibonacci number $f_i=f_{i-1}+f_{i-2}$ and $f_0=0$, $f_1=1$, which is verified by induction, based on this equation.

$$\begin{pmatrix}f_{i-1}&f_{i-2}\\f_{i-2}&f_{i-3}\end{pmatrix}\begin{pmatrix}1&1\\1&0\end{pmatrix}=\begin{pmatrix}f_i&f_{i-1}\\f_{i-1}&f_{i-2}\end{pmatrix}$$

3.IV.3.40 The sum along the i-th row of the product is this.

$$\begin{aligned}
p_{i,1}+\cdots+p_{i,n} &= (h_{i,1}g_{1,1}+h_{i,2}g_{2,1}+\cdots+h_{i,n}g_{n,1}) \\
&\quad + (h_{i,1}g_{1,2}+h_{i,2}g_{2,2}+\cdots+h_{i,n}g_{n,2}) \\
&\quad + \cdots + (h_{i,1}g_{1,n}+h_{i,2}g_{2,n}+\cdots+h_{i,n}g_{n,n}) \\
&= h_{i,1}(g_{1,1}+g_{1,2}+\cdots+g_{1,n}) \\
&\quad + h_{i,2}(g_{2,1}+g_{2,2}+\cdots+g_{2,n}) \\
&\quad + \cdots + h_{i,n}(g_{n,1}+g_{n,2}+\cdots+g_{n,n}) \\
&= h_{i,1}\cdot 1+\cdots+h_{i,n}\cdot 1 \\
&= 1
\end{aligned}$$

3.IV.3.42 The combination is to have all entries of the matrix be zero except for one (possibly) nonzero entry in each row and column. Such a matrix can be written as the product of a permutation matrix and a diagonal matrix, e.g.,

$$\begin{pmatrix}0&4&0\\2&0&0\\0&0&-5\end{pmatrix}=\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}4&0&0\\0&2&0\\0&0&-5\end{pmatrix}$$

and its action is thus to rescale the rows and permute them.

3.IV.3.43

(a) Each entry $p_{i,j}=g_{i,1}h_{1,j}+\cdots+g_{1,r}h_{r,1}$ takes r multiplications and there are $m\cdot n$ entries. Thus there are $m\cdot n\cdot r$ multiplications.

(b) Let H_1 be $5{\times}10$, let H_2 be $10{\times}20$, let H_3 be $20{\times}5$, let H_4 be $5{\times}1$. Then, using the formula from the prior part,

this association	*uses this many multiplications*
$((H_1H_2)H_3)H_4$	$1000+500+25=1525$
$(H_1(H_2H_3))H_4$	$1000+250+25=1275$
$(H_1H_2)(H_3H_4)$	$1000+100+100=1200$
$H_1(H_2(H_3H_4))$	$100+200+50=350$
$H_1((H_2H_3)H_4)$	$1000+50+50=1100$

shows which is cheapest.

(c) This is reported by Knuth as an improvement by S. Winograd of a formula due to V. Strassen: where $w=aA-(a-c-d)(A-C+D)$,

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}A&B\\C&D\end{pmatrix}=\begin{pmatrix}aA+bB & w+(c+d)(C-A)+(a+b-c-d)D\\ w+(a-c)(D-C)-d(A-B-C+D) & w+(a-c)(D-C)+(c+d)(C-A)\end{pmatrix}$$

takes seven multiplications and fifteen additions (save the intermediate results).

3.IV.3.44 *This is how the answer was given in the cited source.* No, it does not. Let A and B represent, with respect to the standard bases, these transformations of $\mathbb{R}^3$.

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\overset{a}{\longmapsto}\begin{pmatrix}x\\y\\0\end{pmatrix}\qquad\begin{pmatrix}x\\y\\z\end{pmatrix}\overset{a}{\longmapsto}\begin{pmatrix}0\\x\\y\end{pmatrix}$$

Observe that

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\overset{abab}{\longmapsto}\begin{pmatrix}0\\0\\0\end{pmatrix}\quad\text{but}\quad\begin{pmatrix}x\\y\\z\end{pmatrix}\overset{baba}{\longmapsto}\begin{pmatrix}0\\0\\x\end{pmatrix}.$$

3.IV.3.45 *This is how the answer was given in the cited source.*

(a) Obvious.

(b) If $A^{\text{trans}}A\vec{x}=\vec{0}$ then $\vec{y}\cdot\vec{y}=0$ where $\vec{y}=A\vec{x}$. Hence $\vec{y}=\vec{0}$ by (a).

The converse is obvious.

(c) By (b), $A\vec{x}_1,\ldots,A\vec{x}_n$ are linearly independent iff $A^{\text{trans}}A\vec{x}_1,\ldots,A^{\text{trans}}A\vec{v}_n$ are linearly independent.

(d) $\text{col rank}(A)=\text{col rank}(A^{\text{trans}}A)=\dim\{A^{\text{trans}}(A\vec{x})\bigm| \text{all }\vec{x}\}\leq\dim\{A^{\text{trans}}\vec{y}\bigm|\text{all }\vec{y}\}=\text{col rank}(A^{\text{trans}})$. Thus also $\text{col rank}(A^{\text{trans}})\leq\text{col rank}({A^{\text{trans}}}^{\text{trans}})$ and so $\text{col rank}(A)=\text{col rank}(A^{\text{trans}})=\text{row rank}(A)$.

3.IV.3.46 *This is how the answer was given in the cited source.* Let $\langle \vec{z}_1, \ldots, \vec{z}_k \rangle$ be a basis for $\mathscr{R}(A) \cap \mathscr{N}(A)$ (k might be 0). Let $\vec{x}_1, \ldots, \vec{x}_k \in V$ be such that $A\vec{x}_i = \vec{z}_i$. Note $\{A\vec{x}_1, \ldots, A\vec{x}_k\}$ is linearly independent, and extend to a basis for $\mathscr{R}(A)$: $A\vec{x}_1, \ldots, A\vec{x}_k, A\vec{x}_{k+1}, \ldots, A\vec{x}_{r_1}$ where $r_1 = \dim(\mathscr{R}(A))$.

Now take $\vec{x} \in V$. Write

$$A\vec{x} = a_1(A\vec{x}_1) + \cdots + a_{r_1}(A\vec{x}_{r_1})$$

and so

$$A^2\vec{x} = a_1(A^2\vec{x}_1) + \cdots + a_{r_1}(A^2\vec{x}_{r_1}).$$

But $A\vec{x}_1, \ldots, A\vec{x}_k \in \mathscr{N}(A)$, so $A^2\vec{x}_1 = \vec{0}, \ldots, A^2\vec{x}_k = \vec{0}$ and we now know

$$A^2\vec{x}_{k+1}, \ldots, A^2\vec{x}_{r_1}$$

spans $\mathscr{R}(A^2)$.

To see $\{A^2\vec{x}_{k+1}, \ldots, A^2\vec{x}_{r_1}\}$ is linearly independent, write

$$b_{k+1}A^2\vec{x}_{k+1} + \cdots + b_{r_1}A^2\vec{x}_{r_1} = \vec{0}$$

$$A[b_{k+1}A\vec{x}_{k+1} + \cdots + b_{r_1}A\vec{x}_{r_1}] = \vec{0}$$

and, since $b_{k+1}A\vec{x}_{k+1} + \cdots + b_{r_1}A\vec{x}_{r_1} \in \mathscr{N}(A)$ we get a contradiction unless it is $\vec{0}$ (clearly it is in $\mathscr{R}(A)$, but $A\vec{x}_1, \ldots, A\vec{x}_k$ is a basis for $\mathscr{R}(A) \cap \mathscr{N}(A)$).

Hence $\dim(\mathscr{R}(A^2)) = r_1 - k = \dim(\mathscr{R}(A)) - \dim(\mathscr{R}(A) \cap \mathscr{N}(A))$.

. . . .

Answers for subsection 3.IV.4

3.IV.4.13 Here is one way to proceed.

$$\xrightarrow{\rho_1 \leftrightarrow \rho_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{-\rho_1+\rho_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & -1 & 1 \end{array}\right)$$

$$\xrightarrow{(1/3)\rho_2+\rho_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & 0 & -4/3 & 1/3 & -1 & 1 \end{array}\right) \xrightarrow[-(3/4)\rho_3]{(1/3)\rho_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1 & -1/4 & 3/4 & -3/4 \end{array}\right)$$

$$\xrightarrow[-\rho_3+\rho_1]{(1/3)\rho_3+\rho_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/4 & 1/4 & 3/4 \\ 0 & 1 & 0 & 1/4 & 1/4 & -1/4 \\ 0 & 0 & 1 & -1/4 & 3/4 & -3/4 \end{array}\right)$$

3.IV.4.18

(a) The proof that the inverse is $r^{-1}H^{-1} = (1/r) \cdot H^{-1}$ (provided, of course, that the matrix is invertible) is easy.

(b) No. For one thing, the fact that $H + G$ has an inverse doesn't imply that H has an inverse or that G has an inverse. Neither of these matrices is invertible but their sum is.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Another point is that just because H and G each has an inverse doesn't mean $H + G$ has an inverse; here is an example.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Still a third point is that, even if the two matrices have inverses, and the sum has an inverse, doesn't imply that the equation holds:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}^{-1} \qquad \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1/3 & 0 \\ 0 & 1/3 \end{pmatrix}^{-1}$$

but

$$\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} 1/5 & 0 \\ 0 & 1/5 \end{pmatrix}^{-1}$$

and $(1/2)_+(1/3)$ does not equal $1/5$.

3.IV.4.20 Yes, the inverse of H^{-1} is H.

3.IV.4.21 One way to check that the first is true is with the angle sum formulas from trigonometry.

$$\begin{aligned}\begin{pmatrix}\cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2)\\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2)\end{pmatrix} &= \begin{pmatrix}\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2 & -\sin\theta_1\cos\theta_2-\cos\theta_1\sin\theta_2\\ \sin\theta_1\cos\theta_2+\cos\theta_1\sin\theta_2 & \cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2\end{pmatrix}\\ &= \begin{pmatrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{pmatrix}\begin{pmatrix}\cos\theta_2 & -\sin\theta_2\\ \sin\theta_2 & \cos\theta_2\end{pmatrix}\end{aligned}$$

Checking the second equation in this way is similar.

Of course, the equations can be not just checked but also understood by recalling that t_θ is the map that rotates vectors about the origin through an angle of θ radians.

3.IV.4.22 There are two cases. For the first case we assume that a is nonzero. Then

$$\xrightarrow{-(c/a)\rho_1+\rho_2}\left(\begin{array}{cc|cc}a & b & 1 & 0\\ 0 & -(bc/a)+d & -c/a & 1\end{array}\right)=\left(\begin{array}{cc|cc}a & b & 1 & 0\\ 0 & (ad-bc)/a & -c/a & 1\end{array}\right)$$

shows that the matrix is invertible (in this $a\neq 0$ case) if and only if $ad-bc\neq 0$. To find the inverse, we finish with the Jordan half of the reduction.

$$\xrightarrow[(a/ad-bc)\rho_2]{(1/a)\rho_1}\left(\begin{array}{cc|cc}1 & b/a & 1/a & 0\\ 0 & 1 & -c/(ad-bc) & a/(ad-bc)\end{array}\right)\xrightarrow{-(b/a)\rho_2+\rho_1}\left(\begin{array}{cc|cc}1 & 0 & d/(ad-bc) & -b/(ad-bc)\\ 0 & 1 & -c/(ad-bc) & a/(ad-bc)\end{array}\right)$$

The other case is the $a=0$ case. We swap to get c into the $1,1$ position.

$$\xrightarrow{\rho_1\leftrightarrow\rho_2}\left(\begin{array}{cc|cc}c & d & 0 & 1\\ 0 & b & 1 & 0\end{array}\right)$$

This matrix is nonsingular if and only if both b and c are nonzero (which, under the case assumption that $a=0$, holds if and only if $ad-bc\neq 0$). To find the inverse we do the Jordan half.

$$\xrightarrow[(1/b)\rho_2]{(1/c)\rho_1}\left(\begin{array}{cc|cc}1 & d/c & 0 & 1/c\\ 0 & 1 & 1/b & 0\end{array}\right)\xrightarrow{-(d/c)\rho_2+\rho_1}\left(\begin{array}{cc|cc}1 & 0 & -d/bc & 1/c\\ 0 & 1 & 1/b & 0\end{array}\right)$$

(Note that this is what is required, since $a=0$ gives that $ad-bc=-bc$).

3.IV.4.23 With H a $2{\times}3$ matrix, in looking for a matrix G such that the combination HG acts as the $2{\times}2$ identity we need G to be $3{\times}2$. Setting up the equation

$$\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}m & n\\ p & q\\ r & s\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

and solving the resulting linear system

$$\begin{array}{rl}m\qquad +r\qquad &=1\\ n\qquad +s&=0\\ p\qquad\qquad &=0\\ q\qquad\qquad &=1\end{array}$$

gives infinitely many solutions.

$$\{\begin{pmatrix}m\\ n\\ p\\ q\\ r\\ s\end{pmatrix}=\begin{pmatrix}1\\ 0\\ 0\\ 1\\ 0\\ 0\end{pmatrix}+r\cdot\begin{pmatrix}-1\\ 0\\ 0\\ 0\\ 1\\ 0\end{pmatrix}+s\cdot\begin{pmatrix}0\\ -1\\ 0\\ 0\\ 0\\ 1\end{pmatrix}\bigm| r,s\in\mathbb{R}\}$$

Thus H has infinitely many right inverses.

As for left inverses, the equation

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

gives rise to a linear system with nine equations and four unknowns.

$$\begin{array}{rcl} a & & =1 \\ \quad b & & =0 \\ a & & =0 \\ \quad\quad c & & =0 \\ \quad\quad\quad d & & =1 \\ \quad\quad c & & =0 \\ \quad\quad\quad\quad e & & =0 \\ \quad\quad\quad\quad\quad f & & =0 \\ \quad\quad\quad\quad e & & =1 \end{array}$$

This system is inconsistent (the first equation conflicts with the third, as do the seventh and ninth) and so there is no left inverse.

3.IV.4.24 With respect to the standard bases we have

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_3}(\eta)=\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$$

and setting up the equation to find the matrix inverse

$$\begin{pmatrix}a&b&c\\d&e&f\end{pmatrix}\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}=\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(\mathrm{id})$$

gives rise to a linear system.

$$\begin{array}{rcl} a & & =1 \\ \quad b & & =0 \\ \quad\quad d & & =0 \\ \quad\quad\quad e & & =1 \end{array}$$

There are infinitely many solutions in $a,\ldots,f$ to this system because two of these variables are entirely unrestricted

$$\{\begin{pmatrix}a\\b\\c\\d\\e\\f\end{pmatrix}=\begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix}+c\cdot\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix}+f\cdot\begin{pmatrix}0\\0\\0\\0\\0\\1\end{pmatrix}\;\big|\;c,f\in\mathbb{R}\}$$

and so there are infinitely many solutions to the matrix equation.

$$\{\begin{pmatrix}1&0&c\\0&1&f\end{pmatrix}\;\big|\;c,f\in\mathbb{R}\}$$

With the bases still fixed at $\mathcal{E}_2,\mathcal{E}_2$, for instance taking $c=2$ and $f=3$ gives a matrix representing this map.

$$\begin{pmatrix}x\\y\\z\end{pmatrix}\stackrel{f_{2,3}}{\longmapsto}\begin{pmatrix}x+2z\\y+3z\end{pmatrix}$$

The check that $f_{2,3}\circ\eta$ is the identity map on $\mathbb{R}^2$ is easy.

3.IV.4.25 By Lemma 4.3 it cannot have infinitely many left inverses, because a matrix with both left and right inverses has only one of each (and that one of each is one of both—the left and right inverse matrices are equal).

3.IV.4.27 Multiply both sides of the first equation by H.

3.IV.4.30 Assume that B is row equivalent to A and that A is invertible. Because they are row-equivalent, there is a sequence of row steps to reduce one to the other. That reduction can be done with matrices, for instance, A can be changed by row operations to B as $B=R_n\cdots R_1A$. This equation gives B as a product of invertible matrices and by Lemma 4.5 then, B is also invertible.

3.IV.4.31

(a) See the answer to Exercise 28.

(b) We will show that both conditions are equivalent to the condition that the two matrices be nonsingular.

As T and S are square and their product is defined, they are equal-sized, say $n\times n$. Consider the $TS=I$ half. By the prior item the rank of I is less than or equal to the minimum of the rank of T and the rank of S. But the rank of I is n, so the rank of T and the rank of S must each be n. Hence each is nonsingular.

The same argument shows that $ST=I$ implies that each is nonsingular.

3.IV.4.32 Inverses are unique, so we need only show that it works. The check appears above as Exercise 31.

3.IV.4.33

(a) See the answer for Exercise 25.

(b) See the answer for Exercise 25.

(c) Apply the first part to $I=AA^{-1}$ to get $I=I^{\text{trans}}=(AA^{-1})^{\text{trans}}=(A^{-1})^{\text{trans}}A^{\text{trans}}$.

(d) Apply the prior item with $A^{\text{trans}}=A$, as A is symmetric.

3.IV.4.35 No, there are at least four.

$$\begin{pmatrix}\pm 1&0\\0&\pm 1\end{pmatrix}$$

3.IV.4.36 It is not reflexive since, for instance,

$$H=\begin{pmatrix}1&0\\0&2\end{pmatrix}$$

is not a two-sided inverse of itself. The same example shows that it is not transitive. That matrix has this two-sided inverse

$$G=\begin{pmatrix}1&0\\0&1/2\end{pmatrix}$$

and while H is a two-sided inverse of G and G is a two-sided inverse of H, we know that H is not a two-sided inverse of H. However, the relation is symmetric: if G is a two-sided inverse of H then $GH=I=HG$ and therefore H is also a two-sided inverse of G.

3.IV.4.37 *This is how the answer was given in the cited source.* Let A be $m\times m$, non-singular, with the stated property. Let B be its inverse. Then for $n\leq m$,

$$1=\sum_{r=1}^{m}\delta_{nr}=\sum_{r=1}^{m}\sum_{s=1}^{m}b_{ns}a_{sr}=\sum_{s=1}^{m}\sum_{r=1}^{m}b_{ns}a_{sr}=k\sum_{s=1}^{m}b_{ns}$$

(A is singular if $k=0$).

. . . .

Answers for subsection 3.V.1

3.V.1.8 One way to go is to find $\text{Rep}_B(\vec{\delta}_1)$ and $\text{Rep}_B(\vec{\delta}_2)$, and then concatenate them into the columns of the desired change of basis matrix. Another way is to find the inverse of the matrices that answer Exercise 7.

(a) $\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$ **(b)** $\begin{pmatrix}1 & 1\\ 2 & 4\end{pmatrix}$ **(c)** $\begin{pmatrix}2 & -1/2\\ -1 & 1/2\end{pmatrix}$ **(d)** $\begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}$

3.V.1.11 This question has many different solutions. One way to proceed is to make up any basis B for any space, and then compute the appropriate D (necessarily for the same space, of course). Another, easier, way to proceed is to fix the codomain as $\mathbb{R}^3$ and the codomain basis as $\mathcal{E}_3$. This way (recall that the representation of any vector with respect to the standard basis is just the vector itself), we have this.

$$B = \langle \begin{pmatrix}3\\ 2\\ 0\end{pmatrix}, \begin{pmatrix}1\\ -1\\ 0\end{pmatrix}, \begin{pmatrix}4\\ 1\\ 4\end{pmatrix} \rangle \qquad D = \mathcal{E}_3$$

3.V.1.12 Checking that $B = \langle 2\sin(x) + \cos(x), 3\cos(x)\rangle$ is a basis is routine. Call the natural basis D. To compute the change of basis matrix $\mathrm{Rep}_{B,D}(\mathrm{id})$ we must find $\mathrm{Rep}_D(2\sin(x) + \cos(x))$ and $\mathrm{Rep}_D(3\cos(x))$, that is, we need x_1, y_1, x_2, y_2 such that these equations hold.

$$\begin{aligned} x_1 \cdot \sin(x) + y_1 \cdot \cos(x) &= 2\sin(x) + \cos(x)\\ x_2 \cdot \sin(x) + y_2 \cdot \cos(x) &= 3\cos(x) \end{aligned}$$

Obviously this is the answer.

$$\mathrm{Rep}_{B,D}(\mathrm{id}) = \begin{pmatrix}2 & 0\\ 1 & 3\end{pmatrix}$$

For the change of basis matrix in the other direction we could look for $\mathrm{Rep}_B(\sin(x))$ and $\mathrm{Rep}_B(\cos(x))$ by solving these.

$$\begin{aligned} w_1 \cdot (2\sin(x) + \cos(x)) + z_1 \cdot (3\cos(x)) &= \sin(x)\\ w_2 \cdot (2\sin(x) + \cos(x)) + z_2 \cdot (3\cos(x)) &= \cos(x) \end{aligned}$$

An easier method is to find the inverse of the matrix found above.

$$\mathrm{Rep}_{D,B}(\mathrm{id}) = \begin{pmatrix}2 & 0\\ 1 & 3\end{pmatrix}^{-1} = \frac{1}{6} \cdot \begin{pmatrix}3 & 0\\ -1 & 2\end{pmatrix} = \begin{pmatrix}1/2 & 0\\ -1/6 & 1/3\end{pmatrix}$$

3.V.1.13 We start by taking the inverse of the matrix, that is, by deciding what is the inverse to the map of interest.

$$\mathrm{Rep}_{D,\mathcal{E}_2}(\mathrm{id})\mathrm{Rep}_{D,\mathcal{E}_2}(\mathrm{id})^{-1} = \frac{1}{-\cos^2(2\theta) - \sin^2(2\theta)} \cdot \begin{pmatrix}-\cos(2\theta) & -\sin(2\theta)\\ -\sin(2\theta) & \cos(2\theta)\end{pmatrix} = \begin{pmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{pmatrix}$$

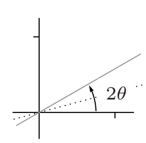
This is more tractable than the representation the other way because this matrix is the concatenation of these two column vectors

$$\mathrm{Rep}_{\mathcal{E}_2}(\vec{\delta}_1) = \begin{pmatrix}\cos(2\theta)\\ \sin(2\theta)\end{pmatrix} \qquad \mathrm{Rep}_{\mathcal{E}_2}(\vec{\delta}_2) = \begin{pmatrix}\sin(2\theta)\\ -\cos(2\theta)\end{pmatrix}$$

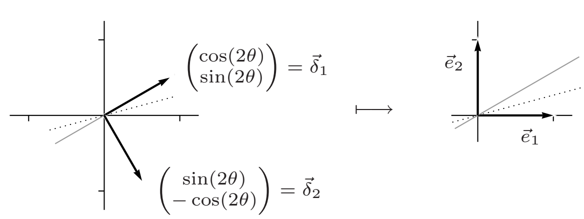
and representations with respect to $\mathcal{E}_2$ are transparent.

$$\vec{\delta}_1 = \begin{pmatrix}\cos(2\theta)\\ \sin(2\theta)\end{pmatrix} \qquad \vec{\delta}_2 = \begin{pmatrix}\sin(2\theta)\\ -\cos(2\theta)\end{pmatrix}$$

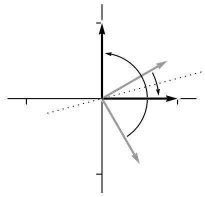
For illustration, taking θ as here,



this pictures the action of the map that transforms D to $\mathcal{E}_2$ (it is, again, the inverse of the map that is the answer to this question).



This action is easier to understand if we superimpost the domain and codomain planes.



Geometrically, the action of this map is to reflect vectors over the line through the origin at angle θ. Since reflections are self-inverse, the answer to the question is: it reflects about the line through the origin with angle of elevation θ. (Of course, it does this to any basis.)

3.V.1.15 Each is true if and only if the matrix is nonsingular.

3.V.1.16 What remains to be shown is that left multiplication by a reduction matrix represents a change from another basis to $B = \langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$.

Application of a row-multiplication matrix $M_i(k)$ translates a representation with respect to the basis $\langle \vec{\beta}_1, \ldots, k\vec{\beta}_i, \ldots, \vec{\beta}_n \rangle$ to one with respect to B, as here.

$$\vec{v} = c_1 \cdot \vec{\beta}_1 + \cdots + c_i \cdot (k\vec{\beta}_i) + \cdots + c_n \cdot \vec{\beta}_n \;\mapsto\; c_1 \cdot \vec{\beta}_1 + \cdots + (kc_i) \cdot \vec{\beta}_i + \cdots + c_n \cdot \vec{\beta}_n = \vec{v}$$

Applying a row-swap matrix $P_{i,j}$ translates a representation with respect to $\langle \vec{\beta}_1, \ldots, \vec{\beta}_j, \ldots, \vec{\beta}_i, \ldots, \vec{\beta}_n \rangle$ to one with respect to $\langle \vec{\beta}_1, \ldots, \vec{\beta}_i, \ldots, \vec{\beta}_j, \ldots, \vec{\beta}_n \rangle$. Finally, applying a row-combination matrix $C_{i,j}(k)$ changes a representation with respect to $\langle \vec{\beta}_1, \ldots, \vec{\beta}_i + k\vec{\beta}_j, \ldots, \vec{\beta}_j, \ldots, \vec{\beta}_n \rangle$ to one with respect to B.

$$\begin{aligned}\vec{v} = c_1 \cdot \vec{\beta}_1 + \cdots + c_i \cdot (\vec{\beta}_i + k\vec{\beta}_j) + \cdots + c_j\vec{\beta}_j + \cdots + c_n \cdot \vec{\beta}_n& \\ \mapsto\; c_1 \cdot \vec{\beta}_1 + \cdots + c_i \cdot \vec{\beta}_i + \cdots + (kc_i + c_j) \cdot \vec{\beta}_j + \cdots + c_n \cdot \vec{\beta}_n = \vec{v}&\end{aligned}$$

(As in the part of the proof in the body of this subsection, the various conditions on the row operations, e.g., that the scalar k is nonzero, assure that these are all bases.)

3.V.1.19 This is the topic of the next subsection.

. . . .

Answers for subsection 3.V.2

3.V.2.12 Recall the diagram and the formula.

$$\begin{array}{ccc} \mathbb{R}^2_{\text{w.r.t. } B} & \xrightarrow[T]{t} & \mathbb{R}^2_{\text{w.r.t. } D} \\ \text{id}\big\downarrow & & \text{id}\big\downarrow \\ \mathbb{R}^2_{\text{w.r.t. } \hat{B}} & \xrightarrow[\hat{T}]{t} & \mathbb{R}^2_{\text{w.r.t. } \hat{D}} \end{array} \qquad \hat{T} = \text{Rep}_{D,\hat{D}}(\text{id}) \cdot T \cdot \text{Rep}_{\hat{B},B}(\text{id})$$

(a) These two

$$\begin{pmatrix}1\\1\end{pmatrix} = 1 \cdot \begin{pmatrix}-1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}2\\1\end{pmatrix} \qquad \begin{pmatrix}1\\-1\end{pmatrix} = (-3) \cdot \begin{pmatrix}-1\\0\end{pmatrix} + (-1) \cdot \begin{pmatrix}2\\1\end{pmatrix}$$

show that

$$\text{Rep}_{D,\hat{D}}(\text{id}) = \begin{pmatrix}1 & -3\\1 & -1\end{pmatrix}$$

and similarly these two

$$\begin{pmatrix}0\\1\end{pmatrix} = 0 \cdot \begin{pmatrix}1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\end{pmatrix} \qquad \begin{pmatrix}1\\1\end{pmatrix} = 1 \cdot \begin{pmatrix}1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\end{pmatrix}$$

give the other nonsinguar matrix.

$$\text{Rep}_{\hat{B},B}(\text{id}) = \begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}$$

Then the answer is this.

$$\hat{T} = \begin{pmatrix} 1 & -3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -10 & -18 \\ -2 & -4 \end{pmatrix}$$

Although not strictly necessary, a check is reassuring. Arbitrarily fixing

$$\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

we have that

$$\mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}_B \qquad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}_{B,D} \begin{pmatrix} 3 \\ 2 \end{pmatrix}_B = \begin{pmatrix} 7 \\ 17 \end{pmatrix}_D$$

and so $t(\vec{v})$ is this.

$$7 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 17 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 24 \\ -10 \end{pmatrix}$$

Doing the calculation with respect to $\hat{B}, \hat{D}$ starts with

$$\mathrm{Rep}_{\hat{B}}(\vec{v}) = \begin{pmatrix} -1 \\ 3 \end{pmatrix}_{\hat{B}} \qquad \begin{pmatrix} -10 & -18 \\ -2 & -4 \end{pmatrix}_{\hat{B},\hat{D}} \begin{pmatrix} -1 \\ 3 \end{pmatrix}_{\hat{B}} = \begin{pmatrix} -44 \\ -10 \end{pmatrix}_{\hat{D}}$$

and then checks that this is the same result.

$$-44 \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} - 10 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 24 \\ -10 \end{pmatrix}$$

(b) These two

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{3} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

show that

$$\mathrm{Rep}_{D,\hat{D}}(\mathrm{id}) = \begin{pmatrix} 1/3 & -1 \\ 1/3 & 1 \end{pmatrix}$$

and these two

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

show this.

$$\mathrm{Rep}_{\hat{B},B}(\mathrm{id}) = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$$

With those, the conversion goes in this way.

$$\hat{T} = \begin{pmatrix} 1/3 & -1 \\ 1/3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} -28/3 & -8/3 \\ 38/3 & 10/3 \end{pmatrix}$$

As in the prior item, a check provides some confidence that this calculation was performed without mistakes. We can for instance, fix the vector

$$\vec{v} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

(this is selected for no reason, out of thin air). Now we have

$$\mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}_{B,D} \begin{pmatrix} -1 \\ 2 \end{pmatrix}_B = \begin{pmatrix} 3 \\ 5 \end{pmatrix}_D$$

and so $t(\vec{v})$ is this vector.

$$3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 5 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \end{pmatrix}$$

With respect to $\hat{B}, \hat{D}$ we first calculate

$$\mathrm{Rep}_{\hat{B}}(\vec{v}) = \begin{pmatrix} 1 \\ -2 \end{pmatrix} \qquad \begin{pmatrix} -28/3 & -8/3 \\ 38/3 & 10/3 \end{pmatrix}_{\hat{B},\hat{D}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}_{\hat{B}} = \begin{pmatrix} -4 \\ 6 \end{pmatrix}_{\hat{D}}$$

and, sure enough, that is the same result for $t(\vec{v})$.

$$-4 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 6 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ -2 \end{pmatrix}$$

3.V.2.17 Yes. Row rank equals column rank, so the rank of the transpose equals the rank of the matrix. Same-sized matrices with equal ranks are matrix equivalent.

3.V.2.18 Only a zero matrix has rank zero.

3.V.2.21 There are two matrix-equivalence classes of 1×1 matrices—those of rank zero and those of rank one. The 3×3 matrices fall into four matrix equivalence classes.

3.V.2.22 For $m\times n$ matrices there are classes for each possible rank: where k is the minimum of m and n there are classes for the matrices of rank 0, 1, ..., k. That's $k+1$ classes. (Of course, totaling over all sizes of matrices we get infinitely many classes.)

3.V.2.23 They are closed under nonzero scalar multiplication, since a nonzero scalar multiple of a matrix has the same rank as does the matrix. They are not closed under addition, for instance, $H+(-H)$ has rank zero.

3.V.2.24

(a) We have

$$\mathrm{Rep}_{B,\mathcal{E}_2}(\mathrm{id})=\begin{pmatrix}1&-1\\2&-1\end{pmatrix}\qquad \mathrm{Rep}_{\mathcal{E}_2,B}(\mathrm{id})=\mathrm{Rep}_{B,\mathcal{E}_2}(\mathrm{id})^{-1}=\begin{pmatrix}1&-1\\2&-1\end{pmatrix}^{-1}=\begin{pmatrix}-1&1\\-2&1\end{pmatrix}$$

and thus the answer is this.

$$\mathrm{Rep}_{B,B}(t)=\begin{pmatrix}1&-1\\2&-1\end{pmatrix}\begin{pmatrix}1&1\\3&-1\end{pmatrix}\begin{pmatrix}-1&1\\-2&1\end{pmatrix}=\begin{pmatrix}-2&0\\-5&2\end{pmatrix}$$

As a quick check, we can take a vector at random

$$\vec{v}=\begin{pmatrix}4\\5\end{pmatrix}$$

giving

$$\mathrm{Rep}_{\mathcal{E}_2}(\vec{v})=\begin{pmatrix}4\\5\end{pmatrix}\qquad \begin{pmatrix}1&1\\3&-1\end{pmatrix}\begin{pmatrix}4\\5\end{pmatrix}=\begin{pmatrix}9\\7\end{pmatrix}=t(\vec{v})$$

while the calculation with respect to B,B

$$\mathrm{Rep}_{B}(\vec{v})=\begin{pmatrix}1\\-3\end{pmatrix}\qquad \begin{pmatrix}-2&0\\-5&2\end{pmatrix}_{B,B}\begin{pmatrix}1\\-3\end{pmatrix}_{B}=\begin{pmatrix}-2\\-11\end{pmatrix}_{B}$$

yields the same result.

$$-2\cdot\begin{pmatrix}1\\2\end{pmatrix}-11\cdot\begin{pmatrix}-1\\-1\end{pmatrix}=\begin{pmatrix}9\\7\end{pmatrix}$$

(b) We have

$$\begin{array}{ccc}\mathbb{R}^2_{\text{w.r.t. }\mathcal{E}_2} & \xrightarrow[T]{t} & \mathbb{R}^2_{\text{w.r.t. }\mathcal{E}_2}\\ \mathrm{id}\big\downarrow & & \mathrm{id}\big\downarrow\\ \mathbb{R}^2_{\text{w.r.t. }B} & \xrightarrow[\hat{T}]{t} & \mathbb{R}^2_{\text{w.r.t. }B}\end{array}\qquad \mathrm{Rep}_{B,B}(t)=\mathrm{Rep}_{\mathcal{E}_2,B}(\mathrm{id})\cdot T\cdot\mathrm{Rep}_{B,\mathcal{E}_2}(\mathrm{id})$$

and, as in the first item of this question

$$\mathrm{Rep}_{B,\mathcal{E}_2}(\mathrm{id})=\left(\vec{\beta}_1\;\middle|\;\cdots\;\middle|\;\vec{\beta}_n\right)\qquad \mathrm{Rep}_{\mathcal{E}_2,B}(\mathrm{id})=\mathrm{Rep}_{B,\mathcal{E}_2}(\mathrm{id})^{-1}$$

so, writing Q for the matrix whose columns are the basis vectors, we have that $\mathrm{Rep}_{B,B}(t)=Q^{-1}TQ$.

3.V.2.25

(a) The adapted form of the arrow diagram is this.

$$\begin{array}{ccc}V_{\text{w.r.t. }B_1} & \xrightarrow[H]{h} & W_{\text{w.r.t. }D}\\ \mathrm{id}\big\downarrow Q & & \mathrm{id}\big\downarrow P\\ V_{\text{w.r.t. }B_2} & \xrightarrow[\hat{H}]{h} & W_{\text{w.r.t. }D}\end{array}$$

Since there is no need to change bases in W (or we can say that the change of basis matrix P is the identity), we have $\mathrm{Rep}_{B_2,D}(h)=\mathrm{Rep}_{B_1,D}(h)\cdot Q$ where $Q=\mathrm{Rep}_{B_2,B_1}(\mathrm{id})$.

(b) Here, this is the arrow diagram.

$$\begin{array}{ccc} V_{\text{w.r.t. } B} & \xrightarrow[H]{h} & W_{\text{w.r.t. } D_1} \\ \text{id}\big\downarrow Q & & \text{id}\big\downarrow P \\ V_{\text{w.r.t. } B} & \xrightarrow[\hat{H}]{h} & W_{\text{w.r.t. } D_2} \end{array}$$

We have that $\text{Rep}_{B,D_2}(h) = P \cdot \text{Rep}_{B,D_1}(h)$ where $P = \text{Rep}_{D_1,D_2}(\text{id})$.

3.V.2.26

(a) Here is the arrow diagram, and a version of that diagram for inverse functions.

$$\begin{array}{ccc} V_{\text{w.r.t. } B} & \xrightarrow[H]{h} & W_{\text{w.r.t. } D} \\ \text{id}\big\downarrow Q & & \text{id}\big\downarrow P \\ V_{\text{w.r.t. } \hat{B}} & \xrightarrow[\hat{H}]{h} & W_{\text{w.r.t. } \hat{D}} \end{array} \qquad \begin{array}{ccc} V_{\text{w.r.t. } B} & \xleftarrow[H^{-1}]{h^{-1}} & W_{\text{w.r.t. } D} \\ \text{id}\big\downarrow Q & & \text{id}\big\downarrow P \\ V_{\text{w.r.t. } \hat{B}} & \xleftarrow[\hat{H}^{-1}]{h^{-1}} & W_{\text{w.r.t. } \hat{D}} \end{array}$$

Yes, the inverses of the matrices represent the inverses of the maps. That is, we can move from the lower right to the lower left by moving up, then left, then down. In other words, where $\hat{H} = PHQ$ (and P, Q invertible) and $H, \hat{H}$ are invertible then $\hat{H}^{-1} = Q^{-1}H^{-1}P^{-1}$.

(b) Yes; this is the prior part repeated in different terms.

(c) No, we need another assumption: if H represents h with respect to the same starting as ending bases B, B, for some B then H^2 represents $h \circ h$. As a specific example, these two matrices are both rank one and so they are matrix equivalent

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

but the squares are not matrix equivalent—the square of the first has rank one while the square of the second has rank zero.

(d) No. These two are not matrix equivalent but have matrix equivalent squares.

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

. . . .

Answers for subsection 3.VI.1

3.VI.1.9 $$\frac{\begin{pmatrix} 1 \\ 2 \\ 1 \\ 3 \end{pmatrix} \bullet \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}}{\begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}} \cdot \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \frac{3}{4} \cdot \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3/4 \\ 3/4 \\ -3/4 \\ 3/4 \end{pmatrix}$$

3.VI.1.11 Suppose that $\vec{v}_1$ and $\vec{v}_2$ are nonzero and orthogonal. Consider the linear relationship $c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$. Take the dot product of both sides of the equation with $\vec{v}_1$ to get that

$$\vec{v}_1 \bullet (c_1\vec{v}_1 + c_2\vec{v}_2) = c_1 \cdot (\vec{v}_1 \bullet \vec{v}_1) + c_2 \cdot (\vec{v}_1 \bullet \vec{v}_2) = c_1 \cdot (\vec{v}_1 \bullet \vec{v}_1) + c_2 \cdot 0 = c_1 \cdot (\vec{v}_1 \bullet \vec{v}_1)$$

is equal to $\vec{v}_1 \bullet \vec{0} = \vec{0}$. With the assumption that $\vec{v}_1$ is nonzero, this gives that c_1 is zero. Showing that c_2 is zero is similar.

3.VI.1.12

(a) If the vector $\vec{v}$ is in the line then the orthogonal projection is $\vec{v}$. To verify this by calculation, note that since $\vec{v}$ is in the line we have that $\vec{v}=c_{\vec{v}}\cdot\vec{s}$ for some scalar $c_{\vec{v}}$.

$$\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}=\frac{c_{\vec{v}}\cdot\vec{s}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}=c_{\vec{v}}\cdot\frac{\vec{s}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}=c_{\vec{v}}\cdot 1\cdot\vec{s}=\vec{v}$$

(*Remark.* If we assume that $\vec{v}$ is nonzero then the above is simplified on taking $\vec{s}$ to be $\vec{v}$.)

(b) Write $c_{\vec{p}}\vec{s}$ for the projection $\operatorname{proj}_{[\vec{s}\,]}(\vec{v})$. Note that, by the assumption that $\vec{v}$ is not in the line, both $\vec{v}$ and $\vec{v}-c_{\vec{p}}\vec{s}$ are nonzero. Note also that if $c_{\vec{p}}$ is zero then we are actually considering the one-element set $\{\vec{v}\,\}$, and with $\vec{v}$ nonzero, this set is necessarily linearly independent. Therefore, we are left considering the case that $c_{\vec{p}}$ is nonzero.

Setting up a linear relationship

$$a_1(\vec{v})+a_2(\vec{v}-c_{\vec{p}}\vec{s})=\vec{0}$$

leads to the equation $(a_1+a_2)\cdot\vec{v}=a_2c_{\vec{p}}\cdot\vec{s}$. Because $\vec{v}$ isn't in the line, the scalars a_1+a_2 and $a_2c_{\vec{p}}$ must both be zero. The $c_{\vec{p}}=0$ case is handled above, so the remaining case is that $a_2=0$, and this gives that $a_1=0$ also. Hence the set is linearly independent.

3.VI.1.13 If $\vec{s}$ is the zero vector then the expression

$$\operatorname{proj}_{[\vec{s}\,]}(\vec{v})=\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}$$

contains a division by zero, and so is undefined. As for the right definition, for the projection to lie in the span of the zero vector, it must be defined to be $\vec{0}$.

3.VI.1.14 Any vector in $\mathbb{R}^n$ is the projection of some other into a line, provided that the dimension n is greater than one. (Clearly, any vector is the projection of itself into a line containing itself; the question is to produce some vector other than $\vec{v}$ that projects to $\vec{v}$.)

Suppose that $\vec{v}\in\mathbb{R}^n$ with $n>1$. If $\vec{v}\neq\vec{0}$ then we consider the line $\ell=\{c\vec{v}\mid c\in\mathbb{R}\}$ and if $\vec{v}=\vec{0}$ we take ℓ to be any (nondegenerate) line at all (actually, we needn't distinguish between these two cases—see the prior exercise). Let $v_1,\ldots,v_n$ be the components of $\vec{v}$; since $n>1$, there are at least two. If some v_i is zero then the vector $\vec{w}=\vec{e}_i$ is perpendicular to $\vec{v}$. If none of the components is zero then the vector $\vec{w}$ whose components are $v_2,-v_1,0,\ldots,0$ is perpendicular to $\vec{v}$. In either case, observe that $\vec{v}+\vec{w}$ does not equal $\vec{v}$, and that $\vec{v}$ is the projection of $\vec{v}+\vec{w}$ into ℓ.

$$\frac{(\vec{v}+\vec{w})\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\cdot\vec{v}=\Big(\frac{\vec{v}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}+\frac{\vec{w}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\Big)\cdot\vec{v}=\frac{\vec{v}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\cdot\vec{v}=\vec{v}$$

We can dispose of the remaining $n=0$ and $n=1$ cases. The dimension $n=0$ case is the trivial vector space, here there is only one vector and so it cannot be expressed as the projection of a different vector. In the dimension $n=1$ case there is only one (nondegenerate) line, and every vector is in it, hence every vector is the projection only of itself.

3.VI.1.16 Because the projection of $\vec{v}$ into the line spanned by $\vec{s}$ is

$$\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}$$

the distance squared from the point to the line is this (a vector dotted with itself $\vec{w}\cdot\vec{w}$ is written $\vec{w}^{\,2}$).

$$\begin{aligned}\|\vec{v}-\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}\,\|^2&=\vec{v}\cdot\vec{v}-\vec{v}\cdot\Big(\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}\Big)-\Big(\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}\,\Big)\cdot\vec{v}+\Big(\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\cdot\vec{s}\,\Big)^2\\&=\vec{v}\cdot\vec{v}-2\cdot\Big(\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\Big)\cdot\vec{v}\cdot\vec{s}+\Big(\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}\Big)\cdot\vec{s}\cdot\vec{s}\\&=\frac{(\vec{v}\cdot\vec{v}\,)\cdot(\vec{s}\cdot\vec{s}\,)-2\cdot(\vec{v}\cdot\vec{s}\,)^2+(\vec{v}\cdot\vec{s}\,)^2}{\vec{s}\cdot\vec{s}}\\&=\frac{(\vec{v}\cdot\vec{v}\,)(\vec{s}\cdot\vec{s}\,)-(\vec{v}\cdot\vec{s}\,)^2}{\vec{s}\cdot\vec{s}}\end{aligned}$$

3.VI.1.17 Because square root is a strictly increasing function, we can minimize $d(c)=(cs_1-v_1)^2+(cs_2-v_2)^2$ instead of the square root of d. The derivative is $dd/dc=2(cs_1-v_1)\cdot s_1+2(cs_2-v_2)\cdot s_2$. Setting it equal to zero $2(cs_1-v_1)\cdot s_1+2(cs_2-v_2)\cdot s_2=c\cdot(2s_1^2+2s_2^2)-(v_1s_1+v_2s_2)=0$ gives the only critical point.

$$c=\frac{v_1s_1+v_2s_2}{{s_1}^2+{s_2}^2}=\frac{\vec{v}\cdot\vec{s}}{\vec{s}\cdot\vec{s}}$$

Now the second derivative with respect to c

$$\frac{d^2\,d}{dc^2} = 2{s_1}^2 + 2{s_2}^2$$

is strictly positive (as long as neither s_1 nor s_2 is zero, in which case the question is trivial) and so the critical point is a minimum.

The generalization to $\mathbb{R}^n$ is straightforward. Consider $d_n(c) = (cs_1 - v_1)^2 + \cdots + (cs_n - v_n)^2$, take the derivative, etc.

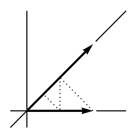
3.VI.1.21 The sequence need not settle down. With

$$\vec{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \vec{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

the projections are these.

$$\vec{v}_1 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1/4 \\ 1/4 \end{pmatrix}, \quad \ldots$$

This sequence doesn't repeat.



. . . .

Answers for subsection 3.VI.2

3.VI.2.9

(a)

$$\vec{\kappa}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\vec{\kappa}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \mbox{proj}_{[\vec{\kappa}_1]}(\begin{pmatrix} 2 \\ 1 \end{pmatrix}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{\begin{pmatrix} 2 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{3}{2} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}$$

(b)

$$\vec{\kappa}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\vec{\kappa}_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix} - \mbox{proj}_{[\vec{\kappa}_1]}(\begin{pmatrix} -1 \\ 3 \end{pmatrix}) = \begin{pmatrix} -1 \\ 3 \end{pmatrix} - \frac{\begin{pmatrix} -1 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}}{\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} - \frac{3}{1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

(c)

$$\vec{\kappa}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\vec{\kappa}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - \mbox{proj}_{[\vec{\kappa}_1]}(\begin{pmatrix} -1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - \frac{\begin{pmatrix} -1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}}{\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - \frac{0}{1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

The corresponding orthonormal bases for the three parts of this question are these.

$$\langle\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix},\begin{pmatrix}\sqrt{2}/2\\-\sqrt{2}/2\end{pmatrix}\rangle\qquad\langle\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}-1\\0\end{pmatrix}\rangle\qquad\langle\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}-1\\0\end{pmatrix}\rangle$$

3.VI.2.12 Reducing the linear system

$$\begin{array}{rl}x-y-z+w=0\\x\quad\;+z\quad\;=0\end{array}\xrightarrow{-\rho_1+\rho_2}\begin{array}{rl}x-y-\;\;z+w=0\\y+2z-w=0\end{array}$$

and paramatrizing gives this description of the subspace.

$$\{\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}\cdot z+\begin{pmatrix}0\\1\\0\\1\end{pmatrix}\cdot w\;\big|\;z,w\in\mathbb{R}\}$$

So we take the basis,

$$\langle\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\1\end{pmatrix}\rangle$$

go through the Gram-Schmidt process

$$\vec{\kappa}_1=\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}$$

$$\vec{\kappa}_2=\begin{pmatrix}0\\1\\0\\1\end{pmatrix}-\mbox{proj}_{[\vec{\kappa}_1]}(\begin{pmatrix}0\\1\\0\\1\end{pmatrix})=\begin{pmatrix}0\\1\\0\\1\end{pmatrix}-\frac{\begin{pmatrix}0\\1\\0\\1\end{pmatrix}\cdot\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}}{\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}\cdot\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}}\cdot\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}=\begin{pmatrix}0\\1\\0\\1\end{pmatrix}-\frac{-2}{6}\cdot\begin{pmatrix}-1\\-2\\1\\0\end{pmatrix}=\begin{pmatrix}-1/3\\1/3\\1/3\\1\end{pmatrix}$$

and finish by normalizing.

$$\langle\begin{pmatrix}-1/\sqrt{6}\\-2/\sqrt{6}\\1/\sqrt{6}\\0\end{pmatrix},\begin{pmatrix}-\sqrt{3}/6\\\sqrt{3}/6\\\sqrt{3}/6\\\sqrt{3}/2\end{pmatrix}\rangle$$

3.VI.2.13 A linearly independent subset of $\mathbb{R}^n$ is a basis for its own span. Apply Theorem 3.VI.2.7.

Remark. Here's why the phrase 'linearly independent' is in the question. Dropping the phrase would require us to worry about two things. The first thing to worry about is that when we do the Gram-Schmidt process on a linearly dependent set then we get some zero vectors. For instance, with

$$S=\{\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}3\\6\end{pmatrix}\}$$

we would get this.

$$\vec{\kappa}_1=\begin{pmatrix}1\\2\end{pmatrix}\qquad\vec{\kappa}_2=\begin{pmatrix}3\\6\end{pmatrix}-\mbox{proj}_{[\vec{\kappa}_1]}(\begin{pmatrix}3\\6\end{pmatrix})=\begin{pmatrix}0\\0\end{pmatrix}$$

This first thing is not so bad because the zero vector is by definition orthogonal to every other vector, so we could accept this situation as yielding an orthogonal set (although it of course can't be normalized), or we just could modify the Gram-Schmidt procedure to throw out any zero vectors. The second thing to worry about if we drop the phrase 'linearly independent' from the question is that the set might be infinite. Of course, any subspace of the finite-dimensional $\mathbb{R}^n$ must also be finite-dimensional so only finitely many of its members are linearly independent, but nonetheless, a "process" that examines the vectors in an infinite set one at a time would at least require some more elaboration in this question. A linearly independent subset of $\mathbb{R}^n$ is automatically finite—in fact, of size n or less—so the 'linearly independent' phrase obviates these concerns.

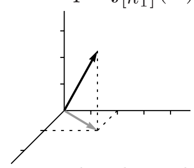
3.VI.2.15

(a) The argument is as in the $i = 3$ case of the proof of Theorem 3.VI.2.7. The dot product

$$\vec{\kappa}_i \cdot \Bigl(\vec{v} - \mathrm{proj}_{[\vec{\kappa}_1]}(\vec{v}\,) - \cdots - \mathrm{proj}_{[\vec{v}_k]}(\vec{v}\,)\Bigr)$$

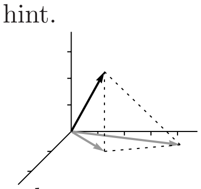
can be written as the sum of terms of the form $-\vec{\kappa}_i \cdot \mathrm{proj}_{[\vec{\kappa}_j]}(\vec{v}\,)$ with $j \neq i$, and the term $\vec{\kappa}_i \cdot (\vec{v} - \mathrm{proj}_{[\vec{\kappa}_i]}(\vec{v}\,))$. The first kind of term equals zero because the $\vec{\kappa}$'s are mutually orthogonal. The other term is zero because this projection is orthogonal (that is, the projection definition makes it zero: $\vec{\kappa}_i \cdot (\vec{v} - \mathrm{proj}_{[\vec{\kappa}_i]}(\vec{v}\,)) = \vec{\kappa}_i \cdot \vec{v} - \vec{\kappa}_i \cdot ((\vec{v} \cdot \vec{\kappa}_i)/(\vec{\kappa}_i \cdot \vec{\kappa}_i)) \cdot \vec{\kappa}_i$ equals, after all of the cancellation is done, zero).

(b) The vector $\vec{v}$ is shown in black and the vector $\mathrm{proj}_{[\vec{\kappa}_1]}(\vec{v}\,) + \mathrm{proj}_{[\vec{v}_2]}(\vec{v}\,) = 1 \cdot \vec{e}_1 + 2 \cdot \vec{e}_2$ is in gray.



The vector $\vec{v} - (\mathrm{proj}_{[\vec{\kappa}_1]}(\vec{v}\,) + \mathrm{proj}_{[\vec{v}_2]}(\vec{v}\,))$ lies on the dotted line connecting the black vector to the gray one, that is, it is orthogonal to the xy-plane.

(c) This diagram is gotten by following the hint.



The dashed triangle has a right angle where the gray vector $1 \cdot \vec{e}_1 + 2 \cdot \vec{e}_2$ meets the vertical dashed line $\vec{v} - (1 \cdot \vec{e}_1 + 2 \cdot \vec{e}_2)$; this is what was proved in the first item of this question. The Pythagorean theorem then gives that the hypoteneuse—the segment from $\vec{v}$ to any other vector—is longer than the vertical dashed line.

More formally, writing $\mathrm{proj}_{[\vec{\kappa}_1]}(\vec{v}\,) + \cdots + \mathrm{proj}_{[\vec{v}_k]}(\vec{v}\,)$ as $c_1 \cdot \vec{\kappa}_1 + \cdots + c_k \cdot \vec{\kappa}_k$, consider any other vector in the span $d_1 \cdot \vec{\kappa}_1 + \cdots + d_k \cdot \vec{\kappa}_k$. Note that

$$\begin{aligned}
&\vec{v} - (d_1 \cdot \vec{\kappa}_1 + \cdots + d_k \cdot \vec{\kappa}_k) \\
&\qquad = \bigl(\vec{v} - (c_1 \cdot \vec{\kappa}_1 + \cdots + c_k \cdot \vec{\kappa}_k)\bigr) + \bigl((c_1 \cdot \vec{\kappa}_1 + \cdots + c_k \cdot \vec{\kappa}_k) - (d_1 \cdot \vec{\kappa}_1 + \cdots + d_k \cdot \vec{\kappa}_k)\bigr)
\end{aligned}$$

and that

$$\bigl(\vec{v} - (c_1 \cdot \vec{\kappa}_1 + \cdots + c_k \cdot \vec{\kappa}_k)\bigr) \cdot \bigl((c_1 \cdot \vec{\kappa}_1 + \cdots + c_k \cdot \vec{\kappa}_k) - (d_1 \cdot \vec{\kappa}_1 + \cdots + d_k \cdot \vec{\kappa}_k)\bigr) = 0$$

(because the first item shows the $\vec{v} - (c_1 \cdot \vec{\kappa}_1 + \cdots + c_k \cdot \vec{\kappa}_k)$ is orthogonal to each $\vec{\kappa}$ and so it is orthogonal to this linear combination of the $\vec{\kappa}$'s). Now apply the Pythagorean Theorem (i.e., the Triangle Inequality).

3.VI.2.16 One way to proceed is to find a third vector so that the three together make a basis for $\mathbb{R}^3$, e.g.,

$$\vec{\beta}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

(the second vector is not dependent on the third because it has a nonzero second component, and the first is not dependent on the second and third because of its nonzero third component), and then apply the

Gram-Schmidt process.

$$\begin{aligned}
\vec{\kappa}_1 &= \begin{pmatrix}1\\5\\-1\end{pmatrix} \\
\vec{\kappa}_2 &= \begin{pmatrix}2\\2\\0\end{pmatrix} - \mbox{proj}_{[\vec{\kappa}_1]}(\begin{pmatrix}2\\2\\0\end{pmatrix}) = \begin{pmatrix}2\\2\\0\end{pmatrix} - \frac{\begin{pmatrix}2\\2\\0\end{pmatrix}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix}}{\begin{pmatrix}1\\5\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix}}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix} \\
&= \begin{pmatrix}2\\2\\0\end{pmatrix} - \frac{12}{27}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix} = \begin{pmatrix}14/9\\-2/9\\4/9\end{pmatrix} \\
\vec{\kappa}_3 &= \begin{pmatrix}1\\0\\0\end{pmatrix} - \mbox{proj}_{[\vec{\kappa}_1]}(\begin{pmatrix}1\\0\\0\end{pmatrix}) - \mbox{proj}_{[\vec{\kappa}_2]}(\begin{pmatrix}1\\0\\0\end{pmatrix}) \\
&= \begin{pmatrix}1\\0\\0\end{pmatrix} - \frac{\begin{pmatrix}1\\0\\0\end{pmatrix}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix}}{\begin{pmatrix}1\\5\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix}}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix} - \frac{\begin{pmatrix}1\\0\\0\end{pmatrix}\cdot\begin{pmatrix}14/9\\-2/9\\4/9\end{pmatrix}}{\begin{pmatrix}14/9\\-2/9\\4/9\end{pmatrix}\cdot\begin{pmatrix}14/9\\-2/9\\4/9\end{pmatrix}}\cdot\begin{pmatrix}14/9\\-2/9\\4/9\end{pmatrix} \\
&= \begin{pmatrix}1\\0\\0\end{pmatrix} - \frac{1}{27}\cdot\begin{pmatrix}1\\5\\-1\end{pmatrix} - \frac{7}{12}\cdot\begin{pmatrix}14/9\\-2/9\\4/9\end{pmatrix} = \begin{pmatrix}1/18\\-1/18\\-4/18\end{pmatrix}
\end{aligned}$$

The result $\vec{\kappa}_3$ is orthogonal to both $\vec{\kappa}_1$ and $\vec{\kappa}_2$. It is therefore orthogonal to every vector in the span of the set $\{\vec{\kappa}_1, \vec{\kappa}_2\}$, including the two vectors given in the question.

3.VI.2.18 First, $\|\vec{v}\,\|^2 = 4^2 + 3^2 + 2^2 + 1^2 = 50$.

(a) $c_1 = 4$ **(b)** $c_1 = 4$, $c_2 = 3$ **(c)** $c_1 = 4$, $c_2 = 3$, $c_3 = 2$, $c_4 = 1$

For the proof, we will do only the $k = 2$ case because the completely general case is messier but no more enlightening. We follow the hint (recall that for any vector $\vec{w}$ we have $\|\vec{w}\,\|^2 = \vec{w}\cdot\vec{w}$).

$$\begin{aligned}
0 &\leq \left(\vec{v} - (\frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\kappa}_1 + \frac{\vec{v}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot\vec{\kappa}_2)\right)\cdot\left(\vec{v} - (\frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\kappa}_1 + \frac{\vec{v}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot\vec{\kappa}_2)\right) \\
&= \vec{v}\cdot\vec{v} - 2\cdot\vec{v}\cdot\left(\frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\kappa}_1 + \frac{\vec{v}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot\vec{\kappa}_2\right) + \left(\frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\kappa}_1 + \frac{\vec{v}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot\vec{\kappa}_2\right)\cdot\left(\frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\kappa}_1 + \frac{\vec{v}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot\vec{\kappa}_2\right) \\
&= \vec{v}\cdot\vec{v} - 2\cdot\left(\frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot(\vec{v}\cdot\vec{\kappa}_1) + \frac{\vec{v}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot(\vec{v}\cdot\vec{\kappa}_2)\right) + \left((\frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1})^2\cdot(\vec{\kappa}_1\cdot\vec{\kappa}_1) + (\frac{\vec{v}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2})^2\cdot(\vec{\kappa}_2\cdot\vec{\kappa}_2)\right)
\end{aligned}$$

(The two mixed terms in the third part of the third line are zero because $\vec{\kappa}_1$ and $\vec{\kappa}_2$ are orthogonal.) The result now follows on gathering like terms and on recognizing that $\vec{\kappa}_1\cdot\vec{\kappa}_1 = 1$ and $\vec{\kappa}_2\cdot\vec{\kappa}_2 = 1$ because these vectors are given as members of an orthonormal set.

3.VI.2.19 It is true, except for the zero vector. Every vector in $\mathbb{R}^n$ except the zero vector is in a basis, and that basis can be orthogonalized.

3.VI.2.20 The 3×3 case gives the idea. The set

$$\{\begin{pmatrix}a\\d\\g\end{pmatrix}, \begin{pmatrix}b\\e\\h\end{pmatrix}, \begin{pmatrix}c\\f\\i\end{pmatrix}\}$$

is orthonormal if and only if these nine conditions all hold

$$\begin{array}{ccc}
\begin{pmatrix}a&d&g\end{pmatrix}\cdot\begin{pmatrix}a\\d\\g\end{pmatrix}=1 & \begin{pmatrix}a&d&g\end{pmatrix}\cdot\begin{pmatrix}b\\e\\h\end{pmatrix}=0 & \begin{pmatrix}a&d&g\end{pmatrix}\cdot\begin{pmatrix}c\\f\\i\end{pmatrix}=0 \\
\begin{pmatrix}b&e&h\end{pmatrix}\cdot\begin{pmatrix}a\\d\\g\end{pmatrix}=0 & \begin{pmatrix}b&e&h\end{pmatrix}\cdot\begin{pmatrix}b\\e\\h\end{pmatrix}=1 & \begin{pmatrix}b&e&h\end{pmatrix}\cdot\begin{pmatrix}c\\f\\i\end{pmatrix}=0 \\
\begin{pmatrix}c&f&i\end{pmatrix}\cdot\begin{pmatrix}a\\d\\g\end{pmatrix}=0 & \begin{pmatrix}c&f&i\end{pmatrix}\cdot\begin{pmatrix}b\\e\\h\end{pmatrix}=0 & \begin{pmatrix}c&f&i\end{pmatrix}\cdot\begin{pmatrix}c\\f\\i\end{pmatrix}=1
\end{array}$$

(the three conditions in the lower left are redundant but nonetheless correct). Those, in turn, hold if and only if

$$\begin{pmatrix}a&d&g\\b&e&h\\c&f&i\end{pmatrix}\begin{pmatrix}a&b&c\\d&e&f\\g&h&i\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

as required.

This is an example, the inverse of this matrix is its transpose.

$$\begin{pmatrix}1/\sqrt{2}&1/\sqrt{2}&0\\-1/\sqrt{2}&1/\sqrt{2}&0\\0&0&1\end{pmatrix}$$

3.VI.2.21 If the set is empty then the summation on the left side is the linear combination of the empty set of vectors, which by definition adds to the zero vector. In the second sentence, there is not such i, so the 'if … then … ' implication is vacuously true.

3.VI.2.22

(a) Part of the induction argument proving Theorem 3.VI.2.7 checks that $\vec{\kappa}_i$ is in the span of $\langle\vec{\beta}_1,\dots,\vec{\beta}_i\rangle$. (The $i=3$ case in the proof illustrates.) Thus, in the change of basis matrix $\operatorname{Rep}_{K,B}(\mathrm{id})$, the i-th column $\operatorname{Rep}_B(\vec{\kappa}_i)$ has components $i+1$ through k that are zero.

(b) One way to see this is to recall the computational procedure that we use to find the inverse. We write the matrix, write the identity matrix next to it, and then we do Gauss-Jordan reduction. If the matrix starts out upper triangular then the Gauss-Jordan reduction involves only the Jordan half and these steps, when performed on the identity, will result in an upper triangular inverse matrix.

3.VI.2.23 For the inductive step, we assume that for all j in $[1..i]$, these three conditions are true of each $\vec{\kappa}_j$: (i) each $\vec{\kappa}_j$ is nonzero, (ii) each $\vec{\kappa}_j$ is a linear combination of the vectors $\vec{\beta}_1,\dots,\vec{\beta}_j$, and (iii) each $\vec{\kappa}_j$ is orthogonal to all of the $\vec{\kappa}_m$'s prior to it (that is, with $m<j$). With those inductive hypotheses, consider $\vec{\kappa}_{i+1}$.

$$\begin{aligned}\vec{\kappa}_{i+1}&=\vec{\beta}_{i+1}-\operatorname{proj}_{[\vec{\kappa}_1]}(\beta_{i+1})-\operatorname{proj}_{[\vec{\kappa}_2]}(\beta_{i+1})-\dots-\operatorname{proj}_{[\vec{\kappa}_i]}(\beta_{i+1})\\&=\vec{\beta}_{i+1}-\frac{\beta_{i+1}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\kappa}_1-\frac{\beta_{i+1}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot\vec{\kappa}_2-\dots-\frac{\beta_{i+1}\cdot\vec{\kappa}_i}{\vec{\kappa}_i\cdot\vec{\kappa}_i}\cdot\vec{\kappa}_i\end{aligned}$$

By the inductive assumption (ii) we can expand each $\vec{\kappa}_j$ into a linear combination of $\vec{\beta}_1,\dots,\vec{\beta}_j$

$$\begin{aligned}&=\vec{\beta}_{i+1}-\frac{\vec{\beta}_{i+1}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\beta}_1\\&\quad-\frac{\vec{\beta}_{i+1}\cdot\vec{\kappa}_2}{\vec{\kappa}_2\cdot\vec{\kappa}_2}\cdot\Big(\text{linear combination of }\vec{\beta}_1,\vec{\beta}_2\Big)-\dots-\frac{\vec{\beta}_{i+1}\cdot\vec{\kappa}_i}{\vec{\kappa}_i\cdot\vec{\kappa}_i}\cdot\Big(\text{linear combination of }\vec{\beta}_1,\dots,\vec{\beta}_i\Big)\end{aligned}$$

The fractions are scalars so this is a linear combination of linear combinations of $\vec{\beta}_1,\dots,\vec{\beta}_{i+1}$. It is therefore just a linear combination of $\vec{\beta}_1,\dots,\vec{\beta}_{i+1}$. Now, (i) it cannot sum to the zero vector because the equation would then describe a nontrivial linear relationship among the $\vec{\beta}$'s that are given as members of a basis (the relationship is nontrivial because the coefficient of $\vec{\beta}_{i+1}$ is 1). Also, (ii) the equation gives $\vec{\kappa}_{i+1}$ as a combination of $\vec{\beta}_1,\dots,\vec{\beta}_{i+1}$. Finally, for (iii), consider $\vec{\kappa}_j\cdot\vec{\kappa}_{i+1}$; as in the $i=3$ case, the dot product of $\vec{\kappa}_j$ with $\vec{\kappa}_{i+1} = \vec{\beta}_{i+1} - \mbox{proj}_{[\vec{\kappa}_1]}(\vec{\beta}_{i+1}) - \cdots - \mbox{proj}_{[\vec{\kappa}_i]}(\vec{\beta}_{i+1})$ can be rewritten to give two kinds of terms, $\vec{\kappa}_j \cdot \left(\vec{\beta}_{i+1} - \mbox{proj}_{[\vec{\kappa}_j]}(\vec{\beta}_{i+1})\right)$ (which is zero because the projection is orthogonal) and $\vec{\kappa}_j \cdot \mbox{proj}_{[\vec{\kappa}_m]}(\vec{\beta}_{i+1})$ with $m \neq j$ and $m < i+1$ (which is zero because by the hypothesis (iii) the vectors $\vec{\kappa}_j$ and $\vec{\kappa}_m$ are orthogonal).

. . . .

Answers for subsection 3.VI.3

3.VI.3.12

(a) Paramatrizing the equation leads to this basis for P.

$$B_P = \langle \begin{pmatrix}1\\0\\3\end{pmatrix}, \begin{pmatrix}0\\1\\2\end{pmatrix} \rangle$$

(b) Because $\mathbb{R}^3$ is three-dimensional and P is two-dimensional, the complement $P^\perp$ must be a line. Anyway, the calculation as in Example 3.5

$$P^\perp = \{ \begin{pmatrix}x\\y\\z\end{pmatrix} \mid \begin{pmatrix}1&0&3\\0&1&2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \}$$

gives this basis for $P^\perp$.

$$B_{P^\perp} = \langle \begin{pmatrix}3\\2\\-1\end{pmatrix} \rangle$$

(c) $\begin{pmatrix}1\\1\\2\end{pmatrix} = (5/14)\cdot\begin{pmatrix}1\\0\\3\end{pmatrix} + (8/14)\cdot\begin{pmatrix}0\\1\\2\end{pmatrix} + (3/14)\cdot\begin{pmatrix}3\\2\\-1\end{pmatrix}$

(d) $\mbox{proj}_P(\begin{pmatrix}1\\1\\2\end{pmatrix}) = \begin{pmatrix}5/14\\8/14\\31/14\end{pmatrix}$

(e) The matrix of the projection

$$\begin{pmatrix}1&0\\0&1\\3&2\end{pmatrix}\left(\begin{pmatrix}1&0&3\\0&1&2\end{pmatrix}\begin{pmatrix}1&0\\0&1\\3&2\end{pmatrix}\right)^{-1}\begin{pmatrix}1&0&3\\0&1&2\end{pmatrix} = \begin{pmatrix}1&0\\0&1\\3&2\end{pmatrix}\begin{pmatrix}10&6\\6&5\end{pmatrix}^{-1}\begin{pmatrix}1&0&3\\0&1&2\end{pmatrix} = \frac{1}{14}\begin{pmatrix}5&-6&3\\-6&10&2\\3&2&13\end{pmatrix}$$

when applied to the vector, yields the expected result.

$$\frac{1}{14}\begin{pmatrix}5&-6&3\\-6&10&2\\3&2&13\end{pmatrix}\begin{pmatrix}1\\1\\2\end{pmatrix} = \begin{pmatrix}5/14\\8/14\\31/14\end{pmatrix}$$

3.VI.3.14 No, a decomposition of vectors $\vec{v} = \vec{m} + \vec{n}$ into $\vec{m} \in M$ and $\vec{n} \in N$ does not depend on the bases chosen for the subspaces—this was shown in the Direct Sum subsection.

3.VI.3.15 The orthogonal projection of a vector into a subspace is a member of that subspace. Since a trivial subspace has only one member, $\vec{0}$, the projection of any vector must equal $\vec{0}$.

3.VI.3.16 The projection into M along N of a $\vec{v} \in M$ is $\vec{v}$. Decomposing $\vec{v} = \vec{m} + \vec{n}$ gives $\vec{m} = \vec{v}$ and $\vec{n} = \vec{0}$, and dropping the N part but retaining the M part results in a projection of $\vec{m} = \vec{v}$.

3.VI.3.17 The proof of Lemma 3.7 shows that each vector $\vec{v} \in \mathbb{R}^n$ is the sum of its orthogonal projections onto the lines spanned by the basis vectors.

$$\vec{v} = \mbox{proj}_{[\vec{\kappa}_1]}(\vec{v}\,) + \cdots + \mbox{proj}_{[\vec{\kappa}_n]}(\vec{v}\,) = \frac{\vec{v}\cdot\vec{\kappa}_1}{\vec{\kappa}_1\cdot\vec{\kappa}_1}\cdot\vec{\kappa}_1 + \cdots + \frac{\vec{v}\cdot\vec{\kappa}_n}{\vec{\kappa}_n\cdot\vec{\kappa}_n}\cdot\vec{\kappa}_n$$

Since the basis is orthonormal, the bottom of each fraction has $\vec{\kappa}_i \cdot \vec{\kappa}_i = 1$.

3.VI.3.20 True; the only vector orthogonal to itself is the zero vector.

3.VI.3.21 This is immediate from the statement in Lemma 3.7 that the space is the direct sum of the two.

3.VI.3.25

(a) First note that if a vector $\vec{v}$ is already in the line then the orthogonal projection gives $\vec{v}$ itself. One way to verify this is to apply the formula for projection into the line spanned by a vector $\vec{s}$, namely $(\vec{v}\cdot\vec{s}/\vec{s}\cdot\vec{s})\cdot\vec{s}$. Taking the line as $\{k\cdot\vec{v} \bigm| k\in\mathbb{R}\}$ (the $\vec{v}=\vec{0}$ case is separate but easy) gives $(\vec{v}\cdot\vec{v}/\vec{v}\cdot\vec{v})\cdot\vec{v}$, which simplifies to $\vec{v}$, as required.

Now, that answers the question because after once projecting into the line, the result $\mathrm{proj}_{\ell}(\vec{v})$ is in that line. The prior paragraph says that projecting into the same line again will have no effect.

(b) The argument here is similar to the one in the prior item. With $V=M\oplus N$, the projection of $\vec{v}=\vec{m}+\vec{n}$ is $\mathrm{proj}_{M,N}(\vec{v})=\vec{m}$. Now repeating the projection will give $\mathrm{proj}_{M,N}(\vec{m})=\vec{m}$, as required, because the decomposition of a member of M into the sum of a member of M and a member of N is $\vec{m}=\vec{m}+\vec{0}$. Thus, projecting twice into M along N has the same effect as projecting once.

(c) As suggested by the prior items, the condition gives that t leaves vectors in the rangespace unchanged, and hints that we should take $\vec{\beta}_1,\ \ldots\ ,\vec{\beta}_r$ to be basis vectors for the range, that is, that we should take the range space of t for M (so that $\dim(M)=r$). As for the complement, we write N for the nullspace of t and we will show that $V=M\oplus N$.

To show this, we can show that their intersection is trivial $M\cap N=\{\vec{0}\}$ and that they sum to the entire space $M+N=V$. For the first, if a vector $\vec{m}$ is in the rangespace then there is a $\vec{v}\in V$ with $t(\vec{v})=\vec{m}$, and the condition on t gives that $t(\vec{m})=(t\circ t)\,(\vec{v})=t(\vec{v})=\vec{m}$, while if that same vector is also in the nullspace then $t(\vec{m})=\vec{0}$ and so the intersection of the rangespace and nullspace is trivial. For the second, to write an arbitrary $\vec{v}$ as the sum of a vector from the rangespace and a vector from the nullspace, the fact that the condition $t(\vec{v})=t(t(\vec{v}))$ can be rewritten as $t(\vec{v}-t(\vec{v}))=\vec{0}$ suggests taking $\vec{v}=t(\vec{v})+(\vec{v}-t(\vec{v}))$.

So we are finished on taking a basis $B=\langle\vec{\beta}_1,\ldots,\vec{\beta}_n\rangle$ for V where $\langle\vec{\beta}_1,\ldots,\vec{\beta}_r\rangle$ is a basis for the rangespace M and $\langle\vec{\beta}_{r+1},\ldots,\vec{\beta}_n\rangle$ is a basis for the nullspace N.

(d) Every projection (as defined in this exercise) is a projection into its rangespace and along its nullspace.

(e) This also follows immediately from the third item.

3.VI.3.26 For any matrix M we have that $(M^{-1})^{\mathrm{trans}}=(M^{\mathrm{trans}})^{-1}$, and for any two matrices M, N we have that $MN^{\mathrm{trans}}=N^{\mathrm{trans}}M^{\mathrm{trans}}$ (provided, of course, that the inverse and product are defined). Applying these two gives that the matrix equals its transpose.

$$\begin{aligned}\bigl(A(A^{\mathrm{trans}}A)^{-1}A^{\mathrm{trans}}\bigr)^{\mathrm{trans}}&=(A^{\mathrm{trans}^{\mathrm{trans}}})(\bigl((A^{\mathrm{trans}}A)^{-1}\bigr)^{\mathrm{trans}})(A^{\mathrm{trans}})\\&=(A^{\mathrm{trans}^{\mathrm{trans}}})(\bigl((A^{\mathrm{trans}}A)^{\mathrm{trans}}\bigr)^{-1})(A^{\mathrm{trans}})=A(A^{\mathrm{trans}}A^{\mathrm{trans}^{\mathrm{trans}}})^{-1}A^{\mathrm{trans}}=A(A^{\mathrm{trans}}A)^{-1}A^{\mathrm{trans}}\end{aligned}$$

. . . .

Answers for Topic: Line of Best Fit

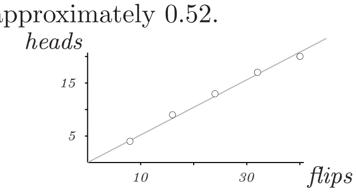
1 As with the first example discussed above, we are trying to find a best m to “solve” this system.

$$\begin{aligned}8m&=4\\16m&=9\\24m&=13\\32m&=17\\40m&=20\end{aligned}$$

Projecting into the linear subspace gives this

$$\frac{\begin{pmatrix}4\\9\\13\\17\\20\end{pmatrix}\cdot\begin{pmatrix}8\\16\\24\\32\\40\end{pmatrix}}{\begin{pmatrix}8\\16\\24\\32\\40\end{pmatrix}\cdot\begin{pmatrix}8\\16\\24\\32\\40\end{pmatrix}}\cdot\begin{pmatrix}8\\16\\24\\32\\40\end{pmatrix}=\frac{1832}{3520}\cdot\begin{pmatrix}8\\16\\24\\32\\40\end{pmatrix}$$

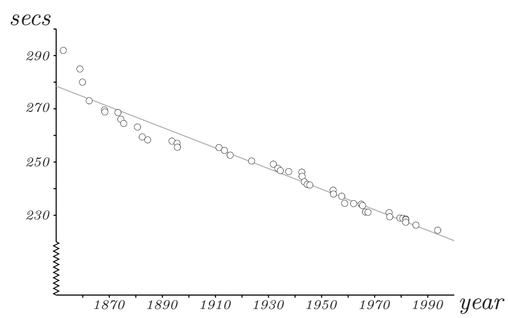
so the slope of the line of best fit is approximately 0.52.



2 With this input

$$A=\begin{pmatrix}1&1852.71\\1&1858.88\\\vdots&\vdots\\1&1985.54\\1&1993.71\end{pmatrix}\qquad b=\begin{pmatrix}292.0\\285.0\\\vdots\\226.32\\224.39\end{pmatrix}$$

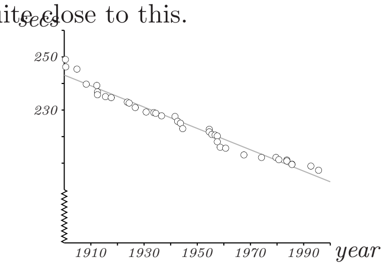
(the dates have been rounded to months, e.g., for a September record, the decimal .71 ≈ (8.5/12) was used), Maple responded with an intercept of $b = 994.8276974$ and a slope of $m = -0.3871993827$.



3 With this input (the years are zeroed at 1900)

$$A:=\begin{pmatrix}1&.38\\1&.54\\\vdots&\vdots\\1&92.71\\1&95.54\end{pmatrix}\qquad b=\begin{pmatrix}249.0\\246.2\\\vdots\\208.86\\207.37\end{pmatrix}$$

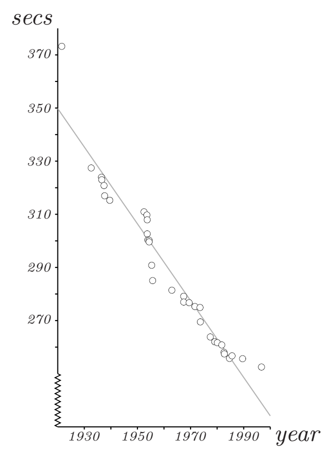
(the dates have been rounded to months, e.g., for a September record, the decimal .71 ≈ (8.5/12) was used), Maple gives an intercept of $b = 243.1590327$ and a slope of $m = -0.401647703$. The slope given in the body of this Topic for the men's mile is quite close to this.



4 With this input (the years are zeroed at 1900)

$$A = \begin{pmatrix} 1 & 21.46 \\ 1 & 32.63 \\ \vdots & \vdots \\ 1 & 89.54 \\ 1 & 96.63 \end{pmatrix} \qquad b = \begin{pmatrix} 373.2 \\ 327.5 \\ \vdots \\ 255.61 \\ 252.56 \end{pmatrix}$$

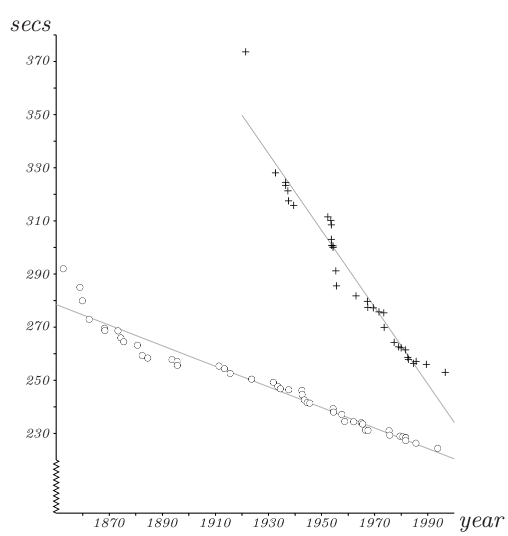
(the dates have been rounded to months, e.g., for a September record, the decimal $.71 \approx (8.5/12)$ was used), MAPLE gave an intercept of $b = 378.7114894$ and a slope of $m = -1.445753225$.



5 These are the equations of the lines for men's and women's mile (the vertical intercept term of the equation for the women's mile has been adjusted from the answer above, to zero it at the year 0, because that's how the men's mile equation was done).

$$y = 994.8276974 - 0.3871993827x$$
$$y = 3125.6426 - 1.445753225x$$

Obviously the lines cross. A computer program is the easiest way to do the arithmetic: MuPAD gives $x = 2012.949004$ and $y = 215.4150856$ (215 seconds is 3 minutes and 35 seconds). *Remark.* Of course all of this projection is highly dubious — for one thing, the equation for the women is influenced by the quite slow early times — but it is nonetheless fun.



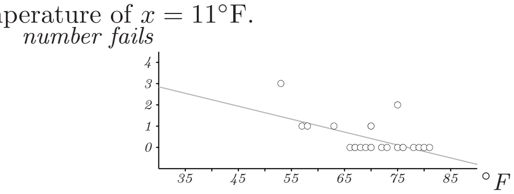
6

(a) A computer algebra system like MAPLE or MuPAD will give an intercept of $b = 4259/1398 \approx 3.239628$ and a slope of $m = -71/2796 \approx -0.025393419$ Plugging $x = 31$ into the equation yields a predicted number of O-ring failures of $y = 2.45$ (rounded to two places). Plugging in $y = 4$ and solving gives a temperature of $x = -29.94°$F.

(b) On the basis of this information

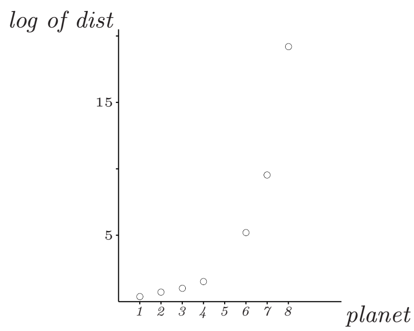
$$A = \begin{pmatrix} 1 & 53 \\ 1 & 75 \\ \vdots & \\ 1 & 80 \\ 1 & 81 \end{pmatrix} \qquad b = \begin{pmatrix} 3 \\ 2 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

MAPLE gives the intercept as $b = 187/40 = 4.675$ and the slope as $m = -73/1200 \approx -0.060833$. Here, plugging $x = 31$ into the equation predicts $y = 2.79$ O-ring failures (rounded to two places). Plugging in $y = 4$ failures gives a temperature of $x = 11°$F.

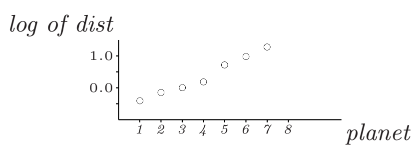


7

(a) The plot is nonlinear.

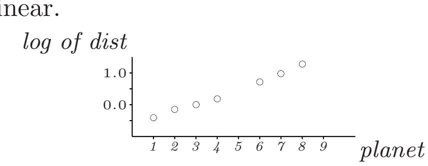


(b) Here is the plot.



There is perhaps a jog up between planet 4 and planet 5.

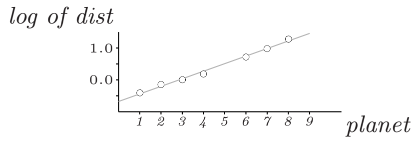
(c) This plot seems even more linear.



(d) With this input

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \end{pmatrix} \qquad b = \begin{pmatrix} -0.40893539 \\ -0.1426675 \\ 0 \\ 0.18184359 \\ 0.71600334 \\ 0.97954837 \\ 1.2833012 \end{pmatrix}$$

MuPAD gives that the intercept is $b = -0.6780677466$ and the slope is $m = 0.2372763818$.



(e) Plugging $x = 9$ into the equation $y = -0.6780677466 + 0.2372763818x$ from the prior item gives that the log of the distance is 1.4574197, so the expected distance is 28.669472. The actual distance is about 30.003.

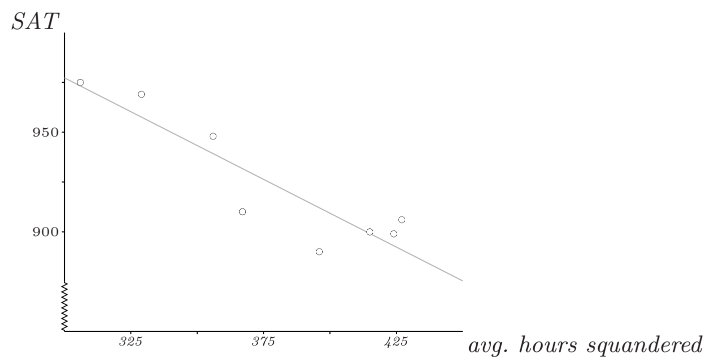
(f) Plugging $x = 10$ into the same equation gives that the log of the distance is 1.6946961, so the expected distance is 49.510362. The actual distance is about 39.503.

8

(a) With this input

$$A = \begin{pmatrix} 1 & 306 \\ 1 & 329 \\ 1 & 356 \\ 1 & 367 \\ 1 & 396 \\ 1 & 427 \\ 1 & 415 \\ 1 & 424 \end{pmatrix} \qquad b = \begin{pmatrix} 975 \\ 969 \\ 948 \\ 910 \\ 890 \\ 906 \\ 900 \\ 899 \end{pmatrix}$$

MAPLE gives that the intercept is $b = 34009779/28796 \approx 1181.0591$ and the slope is $m = -19561/28796 \approx -0.6793$.



. . . .

Answers for Topic: Geometry of Linear Maps

1

(a) To represent H, recall that rotation counterclockwise by θ radians is represented with respect to the standard basis in this way.

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(h) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

A clockwise angle is the negative of a counterclockwise one.

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(h) = \begin{pmatrix} \cos(-\pi/4) & -\sin(-\pi/4) \\ \sin(-\pi/4) & \cos(-\pi/4) \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$$

This Gauss-Jordan reduction

$$\xrightarrow{\rho_1+\rho_2} \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ 0 & \sqrt{2} \end{pmatrix} \xrightarrow[(1/\sqrt{2})\rho_2]{(2/\sqrt{2})\rho_1} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \xrightarrow{-\rho_2+\rho_1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

produces the identity matrix so there is no need for column-swapping operations to end with a partial-identity.

(b) The reduction is expressed in matrix multiplication as

$$\begin{pmatrix}1&-1\\0&1\end{pmatrix}\begin{pmatrix}2/\sqrt{2}&0\\0&1/\sqrt{2}\end{pmatrix}\begin{pmatrix}1&0\\1&1\end{pmatrix}H=I$$

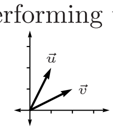
(note that composition of the Gaussian operations is performed from right to left).

(c) Taking inverses

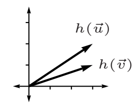
$$H=\underbrace{\begin{pmatrix}1&0\\-1&1\end{pmatrix}\begin{pmatrix}\sqrt{2}/2&0\\0&\sqrt{2}\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}}_{P}I$$

gives the desired factorization of H (here, the partial identity is I, and Q is trivial, that is, it is also an identity matrix).

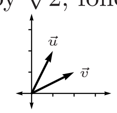
(d) Reading the composition from right to left (and ignoring the identity matrices as trivial) gives that H has the same effect as first performing this skew



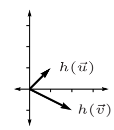
$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x+y\\y\end{pmatrix}\qquad\longrightarrow$$



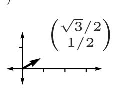
followed by a dilation that multiplies all first components by $\sqrt{2}/2$ (this is a "shrink" in that $\sqrt{2}/2\approx 0.707$) and all second components by $\sqrt{2}$, followed by another skew.



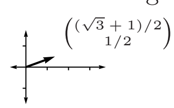
$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x\\-x+y\end{pmatrix}\qquad\longrightarrow$$



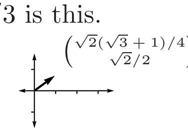
For instance, the effect of H on the unit vector whose angle with the x-axis is $\pi/3$ is this.



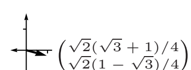
$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x+y\\y\end{pmatrix}\qquad\longrightarrow$$



$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}(\sqrt{2}/2)x\\\sqrt{2}y\end{pmatrix}\qquad\longrightarrow$$



$$\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x\\-x+y\end{pmatrix}\qquad\longrightarrow$$



Verifying that the resulting vector has unit length and forms an angle of $-\pi/6$ with the x-axis is routine.

2 We will first represent the map with a matrix H, perform the row operations and, if needed, column operations to reduce it to a partial-identity matrix. We will then translate that into a factorization $H=PBQ$. Subsitituting into the general matrix

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(r_\theta)\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}$$

gives this representation.

$$\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(r_{2\pi/3})\begin{pmatrix}-1/2&-\sqrt{3}/2\\\sqrt{3}/2&-1/2\end{pmatrix}$$

Gauss' method is routine.

$$\xrightarrow{\sqrt{3}\rho_1+\rho_2}\begin{pmatrix}-1/2&-\sqrt{3}/2\\0&-2\end{pmatrix}\xrightarrow[(-1/2)\rho_2]{-2\rho_1}\begin{pmatrix}1&\sqrt{3}\\0&1\end{pmatrix}\xrightarrow{-\sqrt{3}\rho_2+\rho_1}\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

That translates to a matrix equation in this way.

$$\begin{pmatrix}1&-\sqrt{3}\\0&1\end{pmatrix}\begin{pmatrix}-2&0\\0&-1/2\end{pmatrix}\begin{pmatrix}1&0\\\sqrt{3}&1\end{pmatrix}\begin{pmatrix}-1/2&-\sqrt{3}/2\\\sqrt{3}/2&-1/2\end{pmatrix}=I$$

Taking inverses to solve for H yields this factorization.

$$\begin{pmatrix}-1/2&-\sqrt{3}/2\\\sqrt{3}/2&-1/2\end{pmatrix}=\begin{pmatrix}1&0\\-\sqrt{3}&1\end{pmatrix}\begin{pmatrix}-1/2&0\\0&-2\end{pmatrix}\begin{pmatrix}1&\sqrt{3}\\0&1\end{pmatrix}I$$

3 This Gaussian reduction

$$\xrightarrow[-\rho_1+\rho_3]{-3\rho_1+\rho_2}\begin{pmatrix}1&2&1\\0&0&-3\\0&0&1\end{pmatrix}\xrightarrow{(1/3)\rho_2+\rho_3}\begin{pmatrix}1&2&1\\0&0&-3\\0&0&0\end{pmatrix}\xrightarrow{(-1/3)\rho_2}\begin{pmatrix}1&2&1\\0&0&1\\0&0&0\end{pmatrix}\xrightarrow{-\rho_2+\rho_1}\begin{pmatrix}1&2&0\\0&0&1\\0&0&0\end{pmatrix}$$

gives the reduced echelon form of the matrix. Now the two column operations of taking -2 times the first column and adding it to the second, and then of swapping columns two and three produce this partial identity.

$$B=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$$

All of that translates into matrix terms as: where

$$P=\begin{pmatrix}1&-1&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&-1/3&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&1/3&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\-1&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\-3&1&0\\0&0&1\end{pmatrix}$$

and

$$Q=\begin{pmatrix}1&-2&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$$

the given matrix factors as PBQ.

4 Represent it with respect to the standard bases $\mathcal{E}_1,\mathcal{E}_1$, then the only entry in the resulting 1×1 matrix is the scalar k.

5 We can show this by induction on the number of components in the vector. In the $n=1$ base case the only permutation is the trivial one, and the map

$$\begin{pmatrix}x_1\end{pmatrix}\mapsto\begin{pmatrix}x_1\end{pmatrix}$$

is indeed expressible as a composition of swaps—as zero swaps. For the inductive step we assume that the map induced by any permutation of fewer than n numbers can be expressed with swaps only, and we consider the map induced by a permutation p of n numbers.

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\mapsto\begin{pmatrix}x_{p(1)}\\x_{p(2)}\\\vdots\\x_{p(n)}\end{pmatrix}$$

Consider the number i such that $p(i)=n$. The map

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_i\\\vdots\\x_n\end{pmatrix}\stackrel{\hat{p}}{\longmapsto}\begin{pmatrix}x_{p(1)}\\x_{p(2)}\\\vdots\\x_{p(n)}\\\vdots\\x_n\end{pmatrix}$$

will, when followed by the swap of the i-th and n-th components, give the map p. Now, the inductive hypothesis gives that $\hat{p}$ is achievable as a composition of swaps.

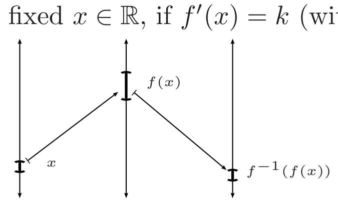
6

(a) A line is a subset of $\mathbb{R}^n$ of the form $\{\vec{v}=\vec{u}+t\cdot\vec{w}\bigm| t\in\mathbb{R}\}$. The image of a point on that line is $h(\vec{v})=h(\vec{u}+t\cdot\vec{w})=h(\vec{u})+t\cdot h(\vec{w})$, and the set of such vectors, as t ranges over the reals, is a line (albeit, degenerate if $h(\vec{w})=\vec{0}$).

(b) This is an obvious extension of the prior argument.

(c) If the point B is between the points A and C then the line from A to C has B in it. That is, there is a $t\in(0\,..\,1)$ such that $\vec{b}=\vec{a}+t\cdot(\vec{c}-\vec{a})$ (where B is the endpoint of $\vec{b}$, etc.). Now, as in the argument of the first item, linearity shows that $h(\vec{b})=h(\vec{a})+t\cdot h(\vec{c}-\vec{a})$.

7 The two are inverse. For instance, for a fixed $x\in\mathbb{R}$, if $f'(x)=k$ (with $k\neq 0$) then $(f^{-1})'(x)=1/k$.



. . . .

Answers for Topic: Markov Chains

1

(a) With this file `coin.m`

```
# Octave function for Markov coin game.  p is chance of going down.
function w = coin(p,v)
  q = 1-p;
  A=[1,p,0,0,0,0;
     0,0,p,0,0,0;
     0,q,0,p,0,0;
     0,0,q,0,p,0;
     0,0,0,q,0,0;
     0,0,0,0,q,1];
  w = A * v;
endfunction
```

this Octave session produced the output given here.

```
octave:1> v0=[0;0;0;1;0;0]
v0 =
  0
  0
  0
  1
  0
  0
octave:2> p=.5
p = 0.50000
octave:3> v1=coin(p,v0)
v1 =
  0.00000
  0.00000
  0.50000
  0.00000
  0.50000
  0.00000
octave:4> v2=coin(p,v1)
v2 =
  0.00000
  0.25000
  0.00000
  0.50000
  0.00000
  0.25000
      ⋮
octave:26> v24=coin(p,v23)
v24 =
  0.39600
  0.00276
  0.00000
  0.00447
  0.00000
  0.59676
```

(b) Using these formulas

$$p_{1,n+1} = 0.5 \cdot p_{2,n} \quad p_{2,n+1} = 0.5 \cdot p_{1,n} + 0.5 \cdot p_{3,n} \quad p_{3,n+1} = 0.5 \cdot p_{2,n} + 0.5 \cdot p_{4,n} \quad p_{5,n+1} = 0.5 \cdot p_{4,n}$$

and these initial conditions

$$\begin{pmatrix} p_{0,0} \\ p_{1,0} \\ p_{2,0} \\ p_{3,0} \\ p_{4,0} \\ p_{5,0} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

we will prove by induction that when n is odd then $p_{1,n} = p_{3,n} = 0$ and when n is even then $p_{2,n} = p_{4,n} = 0$. Note first that this is true in the $n = 0$ base case by the initial conditions. For the inductive step, suppose that it is true in the $n = 0$, $n = 1$, ..., $n = k$ cases and consider the $n = k + 1$ case. If $k + 1$ is odd then

$$p_{1,k+1} = 0.5 \cdot p_{2,k} = 0.5 \cdot 0 = 0 \quad p_{3,k+1} = 0.5 \cdot p_{2,k} + 0.5 \cdot p_{4,k} = 0.5 \cdot 0 + 0.5 \cdot 0 = 0$$

follows from the inductive hypothesis that $p_{2,k} = p_{4,k} = 0$ since k is even. The case where $k + 1$ is even is similar.

(c) We can use, say, $n = 100$. This Octave session

```
octave:1> B=[1,.5,0,0,0,0;
>           0,0,.5,0,0,0;
>           0,.5,0,.5,0,0;
>           0,0,.5,0,.5,0;
>           0,0,0,.5,0,0;
>           0,0,0,0,.5,1];
octave:2> B100=B**100
B100 =
  1.00000  0.80000  0.60000  0.40000  0.20000  0.00000
  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
  0.00000  0.20000  0.40000  0.60000  0.80000  1.00000
octave:3> B100*[0;1;0;0;0;0]
octave:4> B100*[0;1;0;0;0;0]
octave:5> B100*[0;0;0;1;0;0]
octave:6> B100*[0;1;0;0;0;0]
```

yields these outputs.

starting with:	\$1	\$2	\$3	\$4
$s_{0,100}$	0.80000	0.60000	0.40000	0.20000
$s_{1,100}$	0.00000	0.00000	0.00000	0.00000
$s_{2,100}$	0.00000	0.00000	0.00000	0.00000
$s_{3,100}$	0.00000	0.00000	0.00000	0.00000
$s_{4,100}$	0.00000	0.00000	0.00000	0.00000
$s_{5,100}$	0.20000	0.40000	0.60000	0.80000

2

(a) From these equations

$$\begin{aligned}
(1/6)s_{1,n} + \quad 0s_{2,n} + \quad 0s_{3,n} + \quad 0s_{4,n} + \quad 0s_{5,n} + \quad 0s_{6,n} &= s_{1,n+1} \\
(1/6)s_{1,n} + (2/6)s_{2,n} + \quad 0s_{3,n} + \quad 0s_{4,n} + \quad 0s_{5,n} + \quad 0s_{6,n} &= s_{2,n+1} \\
(1/6)s_{1,n} + (1/6)s_{2,n} + (3/6)s_{3,n} + \quad 0s_{4,n} + \quad 0s_{5,n} + \quad 0s_{6,n} &= s_{3,n+1} \\
(1/6)s_{1,n} + (1/6)s_{2,n} + (1/6)s_{3,n} + (4/6)s_{4,n} + \quad 0s_{5,n} + \quad 0s_{6,n} &= s_{4,n+1} \\
(1/6)s_{1,n} + (1/6)s_{2,n} + (1/6)s_{3,n} + (1/6)s_{4,n} + (5/6)s_{5,n} + \quad 0s_{6,n} &= s_{5,n+1} \\
(1/6)s_{1,n} + (1/6)s_{2,n} + (1/6)s_{3,n} + (1/6)s_{4,n} + (1/6)s_{5,n} + (6/6)s_{6,n} &= s_{6,n+1}
\end{aligned}$$

We get this transition matrix.

$$\begin{pmatrix} 1/6 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 2/6 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 3/6 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 4/6 & 0 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 5/6 & 0 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 6/6 \end{pmatrix}$$

(b) This is the Octave session, with outputs edited out and condensed into the table at the end.

```
octave:1>   F=[1/6,  0,   0,   0,   0,   0;
>      1/6,  2/6, 0,   0,   0,   0;
>      1/6,  1/6, 3/6, 0,   0,   0;
>      1/6,  1/6, 1/6, 4/6, 0,   0;
>      1/6,  1/6, 1/6, 1/6, 5/6, 0;
>      1/6,  1/6, 1/6, 1/6, 1/6, 6/6];
octave:2> v0=[1;0;0;0;0;0]
octave:3> v1=F*v0
octave:4> v2=F*v1
octave:5> v3=F*v2
octave:6> v4=F*v3
octave:7> v5=F*v4
```

These are the results.

	1	2	3	4	5
1	0.16667	0.027778	0.0046296	0.00077160	0.00012860
0	0.16667	0.083333	0.0324074	0.01157407	0.00398663
0	0.16667	0.138889	0.0879630	0.05015432	0.02713477
0	0.16667	0.194444	0.1712963	0.13503086	0.10043724
0	0.16667	0.250000	0.2824074	0.28472222	0.27019033
0	0.16667	0.305556	0.4212963	0.51774691	0.59812243

3

(a) It does seem reasonable that, while the firm's present location should strongly influence where it is next time (for instance, whether it stays), any locations in the prior stages should have little influence. That is, while a company may move or stay because of where it is, it is unlikely to move or stay because of where it was.

(b) This Octave session has been edited, with the outputs put together in a table at the end.

```
octave:1> M=[.787,0,0,.111,.102;
>           0,.966,.034,0,0;
>           0,.063,.937,0,0;
>           0,0,.074,.612,.314;
>           .021,.009,.005,.010,.954]
M =
  0.78700  0.00000  0.00000  0.11100  0.10200
  0.00000  0.96600  0.03400  0.00000  0.00000
  0.00000  0.06300  0.93700  0.00000  0.00000
  0.00000  0.00000  0.07400  0.61200  0.31400
  0.02100  0.00900  0.00500  0.01000  0.95400
octave:2> v0=[.025;.025;.025;.025;.900]
octave:3> v1=M*v0
octave:4> v2=M*v1
octave:5> v3=M*v2
octave:6> v4=M*v3
```

is summarized in this table.

$$\begin{array}{c|cccc}
\vec{p}_0 & \vec{p}_1 & \vec{p}_2 & \vec{p}_3 & \vec{p}_4 \\
\hline
\begin{pmatrix} 0.025000\\ 0.025000\\ 0.025000\\ 0.025000\\ 0.900000 \end{pmatrix} &
\begin{pmatrix} 0.114250\\ 0.025000\\ 0.025000\\ 0.299750\\ 0.859725 \end{pmatrix} &
\begin{pmatrix} 0.210879\\ 0.025000\\ 0.025000\\ 0.455251\\ 0.825924 \end{pmatrix} &
\begin{pmatrix} 0.300739\\ 0.025000\\ 0.025000\\ 0.539804\\ 0.797263 \end{pmatrix} &
\begin{pmatrix} 0.377920\\ 0.025000\\ 0.025000\\ 0.582550\\ 0.772652 \end{pmatrix}
\end{array}$$

(c) This is a continuation of the Octave session from the prior item.

```
octave:7> p0=[.0000;.6522;.3478;.0000;.0000]
octave:8> p1=M*p0
octave:9> p2=M*p1
octave:10> p3=M*p2
octave:11> p4=M*p3
```

This summarizes the output.

$$\begin{array}{c|cccc}
\vec{p}_0 & \vec{p}_1 & \vec{p}_2 & \vec{p}_3 & \vec{p}_4 \\
\hline
\begin{pmatrix} 0.00000\\ 0.65220\\ 0.34780\\ 0.00000\\ 0.00000 \end{pmatrix} &
\begin{pmatrix} 0.00000\\ 0.64185\\ 0.36698\\ 0.02574\\ 0.00761 \end{pmatrix} &
\begin{pmatrix} 0.0036329\\ 0.6325047\\ 0.3842942\\ 0.0452966\\ 0.0151277 \end{pmatrix} &
\begin{pmatrix} 0.0094301\\ 0.6240656\\ 0.3999315\\ 0.0609094\\ 0.0225751 \end{pmatrix} &
\begin{pmatrix} 0.016485\\ 0.616445\\ 0.414052\\ 0.073960\\ 0.029960 \end{pmatrix}
\end{array}$$

(d) This is more of the same Octave session.

```
octave:12> M50=M**50
M50 =
  0.03992  0.33666  0.20318  0.02198  0.37332
  0.00000  0.65162  0.34838  0.00000  0.00000
  0.00000  0.64553  0.35447  0.00000  0.00000
  0.03384  0.38235  0.22511  0.01864  0.31652
  0.04003  0.33316  0.20029  0.02204  0.37437
octave:13> p50=M50*p0
p50 =
  0.29024
  0.54615
  0.54430
  0.32766
  0.28695
octave:14> p51=M*p50
p51 =
  0.29406
  0.54609
  0.54442
  0.33091
  0.29076
```

This is close to a steady state.

4

(a) This is the relevant system of equations.

$$\begin{array}{rrrrrl}
(1-2p)\cdot s_{U,n} + & p\cdot t_{A,n} + & p\cdot t_{B,n} & & & = s_{U,n+1} \\
p\cdot s_{U,n} + & (1-2p)\cdot t_{A,n} & & & & = t_{A,n+1} \\
p\cdot s_{U,n} & & +(1-2p)\cdot t_{B,n} & & & = t_{B,n+1} \\
 & p\cdot t_{A,n} & & + s_{A,n} & & = s_{A,n+1} \\
 & & p\cdot t_{B,n} & & + s_{B,n} & = s_{B,n+1}
\end{array}$$

Thus we have this.

$$\begin{pmatrix} 1-2p & p & p & 0 & 0 \\ p & 1-2p & 0 & 0 & 0 \\ p & 0 & 1-2p & 0 & 0 \\ 0 & p & 0 & 1 & 0 \\ 0 & 0 & p & 0 & 1 \end{pmatrix} \begin{pmatrix} s_{U,n} \\ t_{A,n} \\ t_{B,n} \\ s_{A,n} \\ s_{B,n} \end{pmatrix} = \begin{pmatrix} s_{U,n+1} \\ t_{A,n+1} \\ t_{B,n+1} \\ s_{A,n+1} \\ s_{B,n+1} \end{pmatrix}$$

(b) This is the Octave code, with the output removed.

```
octave:1> T=[.5,.25,.25,0,0;
>             .25,.5,0,0,0;
>             .25,0,.5,0,0;
>             0,.25,0,1,0;
>             0,0,.25,0,1]
T =
  0.50000  0.25000  0.25000  0.00000  0.00000
  0.25000  0.50000  0.00000  0.00000  0.00000
  0.25000  0.00000  0.50000  0.00000  0.00000
  0.00000  0.25000  0.00000  1.00000  0.00000
  0.00000  0.00000  0.25000  0.00000  1.00000
octave:2> p0=[1;0;0;0;0]
octave:3> p1=T*p0
octave:4> p2=T*p1
octave:5> p3=T*p2
octave:6> p4=T*p3
octave:7> p5=T*p4
```

Here is the output. The probability of ending at s_A is about 0.23.

	$\vec{p}_0$	$\vec{p}_1$	$\vec{p}_2$	$\vec{p}_3$	$\vec{p}_4$	$\vec{p}_5$
s_U	1	0.50000	0.375000	0.31250	0.26562	0.22656
t_A	0	0.25000	0.250000	0.21875	0.18750	0.16016
t_B	0	0.25000	0.250000	0.21875	0.18750	0.16016
s_A	0	0.00000	0.062500	0.12500	0.17969	0.22656
s_B	0	0.00000	0.062500	0.12500	0.17969	0.22656

(c) With this file as `learn.m`

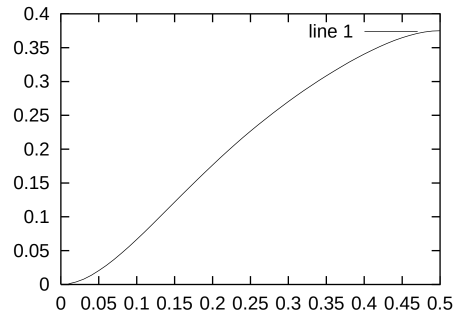
```
# Octave script file for learning model.
function w = learn(p)
   T = [1-2*p,p,    p,    0, 0;
        p,    1-2*p,0,    0, 0;
        p,    0,    1-2*p,0, 0;
        0,    p,    0,    1, 0;
        0,    0,    p,    0, 1];
  T5 = T**5;
  p5 = T5*[1;0;0;0;0];
  w = p5(4);
endfunction
```

issuing the command `octave:1> learn(.20)` yields `ans = 0.17664`.

(d) This Octave session

```
octave:1> x=(.01:.01:.50)';
octave:2> y=(.01:.01:.50)';
octave:3> for i=.01:.01:.50
>           y(100*i)=learn(i);
>         endfor
octave:4> z=[x, y];
octave:5> gplot z
```

yields this plot. There is no threshold value — no probability above which the curve rises sharply.



5

(a) From these equations

$$\begin{aligned} 0.90 \cdot p_{T,n} + 0.01 \cdot p_{C,n} &= p_{T,n+1} \\ 0.10 \cdot p_{T,n} + 0.99 \cdot p_{C,n} &= p_{C,n+1} \end{aligned}$$

we get this matrix.

$$\begin{pmatrix} 0.90 & 0.01 \\ 0.10 & 0.99 \end{pmatrix} \begin{pmatrix} p_{T,n} \\ p_{C,n} \end{pmatrix} = \begin{pmatrix} p_{T,n+1} \\ p_{C,n+1} \end{pmatrix}$$

(b) This is the result from Octave.

$n = 0$	1	2	3	4	5
0.30000	0.27700	0.25653	0.23831	0.22210	0.20767
0.70000	0.72300	0.74347	0.76169	0.77790	0.79233

6	7	8	9	10
0.19482	0.18339	0.17322	0.16417	0.15611
0.80518	0.81661	0.82678	0.83583	0.84389

(c) This is the $s_T = 0.2$ result.

$n = 0$	1	2	3	4	5
0.20000	0.18800	0.17732	0.16781	0.15936	0.15183
0.80000	0.81200	0.82268	0.83219	0.84064	0.84817

6	7	8	9	10
0.14513	0.13916	0.13385	0.12913	0.12493
0.85487	0.86084	0.86615	0.87087	0.87507

(d) Although the probability vectors start 0.1 apart, they end only 0.032 apart. So they are alike.

6 These are the $p = .55$ vectors,

	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
0-0	1	0	0	0	0	0	0	0
1-0	0	0.55000	0	0	0	0	0	0
0-1	0	0.45000	0	0	0	0	0	0
2-0	0	0	0.30250	0	0	0	0	0
1-1	0	0	0.49500	0	0	0	0	0
0-2	0	0	0.20250	0	0	0	0	0
3-0	0	0	0	0.16638	0	0	0	0
2-1	0	0	0	0.40837	0	0	0	0
1-2	0	0	0	0.33412	0	0	0	0
0-3	0	0	0	0.09112	0	0	0	0
4-0	0	0	0	0	0.09151	0.09151	0.09151	0.09151
3-1	0	0	0	0	0.29948	0	0	0
2-2	0	0	0	0	0.36754	0	0	0
1-3	0	0	0	0	0.20047	0	0	0
0-4	0	0	0	0	0.04101	0.04101	0.04101	0.04101
4-1	0	0	0	0	0	0.16471	0.16471	0.16471
3-2	0	0	0	0	0	0.33691	0	0
2-3	0	0	0	0	0	0.27565	0	0
1-4	0	0	0	0	0	0.09021	0.09021	0.09021
4-2	0	0	0	0	0	0	0.18530	0.18530
3-3	0	0	0	0	0	0	0.30322	0
2-4	0	0	0	0	0	0	0.12404	0.12404
4-3	0	0	0	0	0	0	0	0.16677
3-4	0	0	0	0	0	0	0	0.13645

and these are the $p = .60$ vectors.

	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
0-0	1	0	0	0	0	0	0	0
1-0	0	0.60000	0	0	0	0	0	0
0-1	0	0.40000	0	0	0	0	0	0
2-0	0	0	0.36000	0	0	0	0	0
1-1	0	0	0.48000	0	0	0	0	0
0-2	0	0	0.16000	0	0	0	0	0
3-0	0	0	0	0.21600	0	0	0	0
2-1	0	0	0	0.43200	0	0	0	0
1-2	0	0	0	0.28800	0	0	0	0
0-3	0	0	0	0.06400	0	0	0	0
4-0	0	0	0	0	0.12960	0.12960	0.12960	0.12960
3-1	0	0	0	0	0.34560	0	0	0
2-2	0	0	0	0	0.34560	0	0	0
1-3	0	0	0	0	0.15360	0	0	0
0-4	0	0	0	0	0.02560	0.02560	0.02560	0.02560
4-1	0	0	0	0	0	0.20736	0.20736	0.20736
3-2	0	0	0	0	0	0.34560	0	0
2-3	0	0	0	0	0	0.23040	0	0
1-4	0	0	0	0	0	0.06144	0.06144	0.06144
4-2	0	0	0	0	0	0	0.20736	0.20736
3-3	0	0	0	0	0	0	0.27648	0
2-4	0	0	0	0	0	0	0.09216	0.09216
4-3	0	0	0	0	0	0	0	0.16589
3-4	0	0	0	0	0	0	0	0.11059

(a) The script from the computer code section can be easily adapted.

```
# Octave script file to compute chance of World Series outcomes.
function w = markov(p,v)
  q = 1-p;
  A=[0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 0-0
     p,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 1-0
```

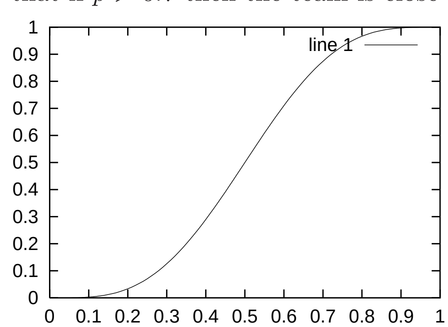
```
        q,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 0-1_
        0,p,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 2-0
        0,q,p,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 1-1
        0,0,q,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 0-2__
        0,0,0,p,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 3-0
        0,0,0,q,p,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 2-1
        0,0,0,0,q,p, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 1-2_
        0,0,0,0,0,q, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 0-3
        0,0,0,0,0,0, p,0,0,0,1,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 4-0
        0,0,0,0,0,0, q,p,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 3-1__
        0,0,0,0,0,0, 0,q,p,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 2-2
        0,0,0,0,0,0, 0,0,q,p,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0;  # 1-3
        0,0,0,0,0,0, 0,0,0,q,0,0, 0,0,1,0,0,0, 0,0,0,0,0,0;  # 0-4_
        0,0,0,0,0,0, 0,0,0,0,0,p, 0,0,0,1,0,0, 0,0,0,0,0,0;  # 4-1
        0,0,0,0,0,0, 0,0,0,0,0,q, p,0,0,0,0,0, 0,0,0,0,0,0;  # 3-2
        0,0,0,0,0,0, 0,0,0,0,0,0, q,p,0,0,0,0, 0,0,0,0,0,0;  # 2-3__
        0,0,0,0,0,0, 0,0,0,0,0,0, 0,q,0,0,0,0, 1,0,0,0,0,0;  # 1-4
        0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,p,0, 0,1,0,0,0,0;  # 4-2
        0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,q,p, 0,0,0,0,0,0;  # 3-3_
        0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,q, 0,0,0,1,0,0;  # 2-4
        0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,p,0,1,0;  # 4-3
        0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,0,0,0,0, 0,0,q,0,0,1]; # 3-4
  v7 = (A**7) * v;
  w = v7(11)+v7(16)+v7(20)+v7(23)
endfunction
```

Using this script, we get that when the American League has a $p = 0.55$ probability of winning each game then their probability of winning the first-to-win-four series is 0.60829. When their probability of winning any one game is $p = 0.6$ then their probability of winning the series is 0.71021.

(b) This Octave session

```
octave:1> v0=[1;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0];
octave:2> x=(.01:.01:.99)';
octave:3> y=(.01:.01:.99)';
octave:4> for i=.01:.01:.99
>          y(100*i)=markov(i,v0);
>         endfor
octave:5> z=[x, y];
octave:6> gplot z
```

yields this graph. By eye we judge that if $p > 0.7$ then the team is close to assurred of the series.



7

(a) They must satisfy this condition because the total probability of a state transition (including back to the same state) is 100%.

(b) See the answer to the third item.

(c) We will do the 2×2 case; bigger-sized cases are just notational problems. This product

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} = \begin{pmatrix} a_{1,1}b_{1,1} + a_{1,2}b_{2,1} & a_{1,1}b_{1,2} + a_{1,2}b_{2,2} \\ a_{2,1}b_{1,1} + a_{2,2}b_{2,1} & a_{2,1}b_{1,2} + a_{2,2}b_{2,2} \end{pmatrix}$$

has these two column sums

$$(a_{1,1}b_{1,1} + a_{1,2}b_{2,1}) + (a_{2,1}b_{1,1} + a_{2,2}b_{2,1}) = (a_{1,1} + a_{2,1}) \cdot b_{1,1} + (a_{1,2} + a_{2,2}) \cdot b_{2,1} = 1 \cdot b_{1,1} + 1 \cdot b_{2,1} = 1$$

and

$$(a_{1,1}b_{1,2} + a_{1,2}b_{2,2}) + (a_{2,1}b_{1,2} + a_{2,2}b_{2,2}) = (a_{1,1} + a_{2,1}) \cdot b_{1,2} + (a_{1,2} + a_{2,2}) \cdot b_{2,2} = 1 \cdot b_{1,2} + 1 \cdot b_{2,2} = 1$$

as required.

. . . .

Answers for Topic: Orthonormal Matrices

1

(a) Yes.

(b) No, the columns do not have length one.

(c) Yes.

2 Some of these are nonlinear, because they involve a nontrivial translation.

(a) $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cdot \cos(\pi/6) - y \cdot \sin(\pi/6) \\ x \cdot \sin(\pi/6) + y \cdot \cos(\pi/6) \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \cdot (\sqrt{3}/2) - y \cdot (1/2) + 0 \\ x \cdot (1/2) + y \cdot \cos(\sqrt{3}/2) + 1 \end{pmatrix}$

(b) The line $y = 2x$ makes an angle of $\arctan(2/1)$ with the x-axis. Thus $\sin\theta = 2/\sqrt{5}$ and $\cos\theta = 1/\sqrt{5}$.

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cdot (1/\sqrt{5}) - y \cdot (2/\sqrt{5}) \\ x \cdot (2/\sqrt{5}) + y \cdot (1/\sqrt{5}) \end{pmatrix}$$

(c) $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cdot (1/\sqrt{5}) - y \cdot (-2/\sqrt{5}) \\ x \cdot (-2/\sqrt{5}) + y \cdot (1/\sqrt{5}) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x/\sqrt{5} + 2y/\sqrt{5} + 1 \\ -2x/\sqrt{5} + y/\sqrt{5} + 1 \end{pmatrix}$

3

(a) Let f be distance-preserving and consider f^{-1}. Any two points in the codomain can be written as $f(P_1)$ and $f(P_2)$. Because f is distance-preserving, the distance from $f(P_1)$ to $f(P_2)$ equals the distance from P_1 to P_2. But this is exactly what is required for f^{-1} to be distance-preserving.

(b) Any plane figure F is congruent to itself via the identity map $\mathrm{id}\colon \mathbb{R}^2 \to \mathbb{R}^2$, which is obviously distance-preserving. If F_1 is congruent to F_2 (via some f) then F_2 is congruent to F_1 via f^{-1}, which is distance-preserving by the prior item. Finally, if F_1 is congruent to F_2 (via some f) and F_2 is congruent to F_3 (via some g) then F_1 is congruent to F_3 via $g \circ f$, which is easily checked to be distance-preserving.

4 The first two components of each are $ax + cy + e$ and $bx + dy + f$.

5

(a) The Pythagorean Theorem gives that three points are colinear if and only if (for some ordering of them into P_1, P_2, and P_3), $\mathrm{dist}(P_1, P_2) + \mathrm{dist}(P_2, P_3) = \mathrm{dist}(P_1, P_3)$. Of course, where f is distance-preserving, this holds if and only if $\mathrm{dist}(f(P_1), f(P_2)) + \mathrm{dist}(f(P_2), f(P_3)) = \mathrm{dist}(f(P_1), f(P_3))$, which, again by Pythagoras, is true if and only if $f(P_1)$, $f(P_2)$, and $f(P_3)$ are colinear.

The argument for betweeness is similar (above, P_2 is between P_1 and P_3).

If the figure F is a triangle then it is the union of three line segments P_1P_2, P_2P_3, and P_1P_3. The prior two paragraphs together show that the property of being a line segment is invariant. So $f(F)$ is the union of three line segments, and so is a triangle.

A circle C centered at P and of radius r is the set of all points Q such that $\mathrm{dist}(P, Q) = r$. Applying the distance-preserving map f gives that the image $f(C)$ is the set of all $f(Q)$ subject to the condition that $\mathrm{dist}(P, Q) = r$. Since $\mathrm{dist}(P, Q) = \mathrm{dist}(f(P), f(Q))$, the set $f(C)$ is also a circle, with center $f(P)$ and radius r.

(b) Here are two that are easy to verify: (i) the property of being a right triangle, and (ii) the property of two lines being parallel.

(c) One that was mentioned in the section is the 'sense' of a figure. A triangle whose vertices read clockwise as P_1, P_2, P_3 may, under a distance-preserving map, be sent to a triangle read P_1, P_2, P_3 counterclockwise.

Chapter 4. Determinants

. . . .

Answers for subsection 4.I.1

4.I.1.2
(a) 6 **(b)** 21 **(c)** 27

4.I.1.5
(a) Nonsingular, the determinant is 3.
(b) Singular, the determinant is 0.
(c) Singular, the determinant is 0.

4.I.1.7 Using the formula for the determinant of a 3×3 matrix we expand the left side

$$1\cdot b\cdot c^2+1\cdot c\cdot a^2+1\cdot a\cdot b^2-b^2\cdot c\cdot 1-c^2\cdot a\cdot 1-a^2\cdot b\cdot 1$$

and by distributing we expand the right side.

$$(bc-ba-ac+a^2)\cdot(c-b)=c^2b-b^2c-bac+b^2a-ac^2+acb+a^2c-a^2b$$

Now we can just check that the two are equal. (*Remark.* This is the 3×3 case of *Vandermonde's determinant* which arises in applications).

4.I.1.9 We first reduce the matrix to echelon form. To begin, assume that $a\neq 0$ and that $ae-bd\neq 0$.

$$\xrightarrow{(1/a)\rho_1}\begin{pmatrix}1&b/a&c/a\\ d&e&f\\ g&h&i\end{pmatrix}\quad\xrightarrow[-g\rho_1+\rho_3]{-d\rho_1+\rho_2}\begin{pmatrix}1&b/a&c/a\\ 0&(ae-bd)/a&(af-cd)/a\\ 0&(ah-bg)/a&(ai-cg)/a\end{pmatrix}$$
$$\xrightarrow{(a/(ae-bd))\rho_2}\begin{pmatrix}1&b/a&c/a\\ 0&1&(af-cd)/(ae-bd)\\ 0&(ah-bg)/a&(ai-cg)/a\end{pmatrix}$$

This step finishes the calculation.

$$\xrightarrow{((ah-bg)/a)\rho_2+\rho_3}\begin{pmatrix}1&b/a&c/a\\ 0&1&(af-cd)/(ae-bd)\\ 0&0&(aei+bgf+cdh-hfa-idb-gec)/(ae-bd)\end{pmatrix}$$

Now assuming that $a\neq 0$ and $ae-bd\neq 0$, the original matrix is nonsingular if and only if the $3,3$ entry above is nonzero. That is, under the assumptions, the original matrix is nonsingular if and only if $aei+bgf+cdh-hfa-idb-gec\neq 0$, as required.

We finish by running down what happens if the assumptions that were taken for convienence in the prior paragraph do not hold. First, if $a\neq 0$ but $ae-bd=0$ then we can swap

$$\begin{pmatrix}1&b/a&c/a\\ 0&0&(af-cd)/a\\ 0&(ah-bg)/a&(ai-cg)/a\end{pmatrix}\xrightarrow{\rho_2\leftrightarrow\rho_3}\begin{pmatrix}1&b/a&c/a\\ 0&(ah-bg)/a&(ai-cg)/a\\ 0&0&(af-cd)/a\end{pmatrix}$$

and conclude that the matrix is nonsingular if and only if either $ah-bg=0$ or $af-cd=0$. The condition '$ah-bg=0$ or $af-cd=0$' is equivalent to the condition '$(ah-bg)(af-cd)=0$'. Multiplying out and using the case assumption that $ae-bd=0$ to substitute ae for bd gives this.

$$0=ahaf-ahcd-bgaf+bgcd=ahaf-ahcd-bgaf+aegc=a(haf-hcd-bgf+egc)$$

Since $a\neq 0$, we have that the matrix is nonsingular if and only if $haf-hcd-bgf+egc=0$. Therefore, in this $a\neq 0$ and $ae-bd=0$ case, the matrix is nonsingular when $haf-hcd-bgf+egc-i(ae-bd)=0$.

The remaining cases are routine. Do the $a=0$ but $d\neq 0$ case and the $a=0$ and $d=0$ but $g\neq 0$ case by first swapping rows and then going on as above. The $a=0$, $d=0$, and $g=0$ case is easy—that matrix is singular since the columns form a linearly dependent set, and the determinant comes out to be zero.

4.I.1.10 Figuring the determinant and doing some algebra gives this.

$$\begin{aligned} 0 &= y_1x + x_2y + x_1y_2 - y_2x - x_1y - x_2y_1 \\ (x_2 - x_1) \cdot y &= (y_2 - y_1) \cdot x + x_2y_1 - x_1y_2 \\ y &= \frac{y_2 - y_1}{x_2 - x_1} \cdot x + \frac{x_2y_1 - x_1y_2}{x_2 - x_1} \end{aligned}$$

Note that this is the equation of a line (in particular, in contains the familiar expression for the slope), and note that (x_1, y_1) and (x_2, y_2) satisfy it.

4.I.1.12 The determinant is $(x_2y_3 - x_3y_2)\vec{e}_1 + (x_3y_1 - x_1y_3)\vec{e}_2 + (x_1y_2 - x_2y_1)\vec{e}_3$. To check perpendicularity, we check that the dot product with the first vector is zero

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} = x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_1x_2y_3 + x_1x_3y_2 - x_2x_3y_1 = 0$$

and the dot product with the second vector is also zero.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \cdot \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} = x_2y_1y_3 - x_3y_1y_2 + x_3y_1y_2 - x_1y_2y_3 + x_1y_2y_3 - x_2y_1y_3 = 0$$

4.I.1.13

(a) Plug and chug: the determinant of the product is this

$$\begin{aligned} \det(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} w & x \\ y & z \end{pmatrix}) &= \det(\begin{pmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \end{pmatrix}) \\ &= acwx + adwz + bcxy + bdyz \\ &\quad -acwx - bcwz - adxy - bdyz \end{aligned}$$

while the product of the determinants is this.

$$\det(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) \cdot \det(\begin{pmatrix} w & x \\ y & z \end{pmatrix}) = (ad - bc) \cdot (wz - xy)$$

Verification that they are equal is easy.

(b) Use the prior item.

That similar matrices have the same determinant is immediate from the above two: $\det(PTP^{-1}) = \det(P) \cdot \det(T) \cdot \det(P^{-1})$.

4.I.1.15 The computation for 2×2 matrices, using the formula quoted in the preamble, is easy. It does also hold for 3×3 matrices; the computation is routine.

4.I.1.17 Bring out the c's one row at a time.

4.I.1.18 There are no real numbers θ that make the matrix singular because the determinant of the matrix $\cos^2\theta + \sin^2\theta$ is never 0, it equals 1 for all θ. Geometrically, with respect to the standard basis, this matrix represents a rotation of the plane through an angle of θ. Each such map is one-to-one — for one thing, it is invertible.

4.I.1.19 *This is how the answer was given in the cited source.* Let P be the sum of the three positive terms of the determinant and $-N$ the sum of the three negative terms. The maximum value of P is

$$9 \cdot 8 \cdot 7 + 6 \cdot 5 \cdot 4 + 3 \cdot 2 \cdot 1 = 630.$$

The minimum value of N consistent with P is

$$9 \cdot 6 \cdot 1 + 8 \cdot 5 \cdot 2 + 7 \cdot 4 \cdot 3 = 218.$$

Any change in P would result in lowering that sum by more than 4. Therefore 412 the maximum value for the determinant and one form for the determinant is

$$\begin{vmatrix} 9 & 4 & 2 \\ 3 & 8 & 6 \\ 5 & 1 & 7 \end{vmatrix}.$$

. . . .

Answers for subsection 4.I.2

4.I.2.8

(a) $\begin{vmatrix} 2 & -1 \\ -1 & -1 \end{vmatrix} = \begin{vmatrix} 2 & -1 \\ 0 & -3/2 \end{vmatrix} = -3$; **(b)** $\begin{vmatrix} 1 & 1 & 0 \\ 3 & 0 & 2 \\ 5 & 2 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & -3 & 2 \\ 0 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & -3 & 2 \\ 0 & 0 & 0 \end{vmatrix} = 0$

4.I.2.9 When is the determinant not zero?

$$\begin{vmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 0 \\ 1 & 0 & k & 0 \\ 0 & 0 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & k-1 & 1 \\ 0 & 0 & 1 & -1 \end{vmatrix}$$

Obviously, $k = 1$ gives nonsingularity and hence a nonzero determinant. If $k \neq 1$ then we get echelon form with a $(-1/k-1)\rho_3 + \rho_4$ pivot.

$$= \begin{vmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & k-1 & 1 \\ 0 & 0 & 0 & -1-(1/k-1) \end{vmatrix}$$

Multiplying down the diagonal gives $(k-1)(-1-(1/k-1)) = -(k-1)-1 = -k$. Thus the matrix has a nonzero determinant, and so the system has a unique solution, if and only if $k \neq 0$.

4.I.2.12 It is the trivial subspace.

4.I.2.14

(a) (1), $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$

(b) The determinant in the 1×1 case is 1. In every other case the second row is the negative of the first, and so matrix is singular and the determinant is zero.

4.I.2.15

(a) (2), $\begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$, $\begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{pmatrix}$

(b) The 1×1 and 2×2 cases yield these.

$$\begin{vmatrix} 2 \end{vmatrix} = 2 \qquad \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} = -1$$

And $n\times n$ matrices with $n \geq 3$ are singular, e.g.,

$$\begin{vmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{vmatrix} = 0$$

because twice the second row minus the first row equals the third row. Checking this is routine.

4.I.2.17 No, we cannot replace it. Remark 2.2 shows that the four conditions after the replacement would conflict — no function satisfies all four.

4.I.2.18 A upper-triangular matrix is in echelon form.

A lower-triangular matrix is either singular or nonsingular. If it is singular then it has a zero on its diagonal and so its determinant (namely, zero) is indeed the product down its diagonal. If it is nonsingular then it has no zeroes on its diagonal, and can be reduced by Gauss' method to echelon form without changing the diagonal.

4.I.2.19

(a) The properties in the definition of determinant show that $|M_i(k)| = k$, $|P_{i,j}| = -1$, and $|C_{i,j}(k)| = 1$.

(b) The three cases are easy to check by recalling the action of left multiplication by each type of matrix.

(c) If TS is invertible $(TS)M = I$ then the associative property of matrix multiplication $T(SM) = I$ shows that T is invertible. So if T is not invertible then neither is TS.

(d) If T is singular then apply the prior answer: $|TS| = 0$ and $|T| \cdot |S| = 0 \cdot |S| = 0$. If T is not singular then it can be written as a product of elementary matrices $|TS| = |E_r \cdots E_1 S| = |E_r| \cdots |E_1| \cdot |S| = |E_r \cdots E_1||S| = |T||S|$.

(e) $1 = |I| = |T \cdot T^{-1}| = |T||T^{-1}|$

4.I.2.20

(a) We must show that if

$$T \xrightarrow{k\rho_i + \rho_j} \hat{T}$$

then $d(T) = |TS|/|S| = |\hat{T}S|/|S| = d(\hat{T})$. We will be done if we show that pivoting first and then multiplying to get $\hat{T}S$ gives the same result as multiplying first to get TS and then pivoting (because the determinant $|TS|$ is unaffected by the pivot so we'll then have $|\hat{T}S| = |TS|$, and hence $d(\hat{T}) = d(T)$). That argument runs: after adding k times row i of TS to row j of TS, the j,p entry is $(kt_{i,1} + t_{j,1})s_{1,p} + \cdots + (kt_{i,r} + t_{j,r})s_{r,p}$, which is the j,p entry of $\hat{T}S$.

(b) We need only show that swapping $T \xrightarrow{\rho_i \leftrightarrow \rho_j} \hat{T}$ and then multiplying to get $\hat{T}S$ gives the same result as multiplying T by S and then swapping (because, as the determinant $|TS|$ changes sign on the row swap, we'll then have $|\hat{T}S| = -|TS|$, and so $d(\hat{T}) = -d(T)$). That argument runs just like the prior one.

(c) Not surprisingly by now, we need only show that multiplying a row by a nonzero scalar $T \xrightarrow{k\rho_i} \hat{T}$ and then computing $\hat{T}S$ gives the same result as first computing TS and then multiplying the row by k (as the determinant $|TS|$ is rescaled by k the multiplication, we'll have $|\hat{T}S| = k|TS|$, so $d(\hat{T}) = k\,d(T)$). The argument runs just as above.

(d) Clear.

(e) Because we've shown that $d(T)$ is a determinant and that determinant functions (if they exist) are unique, we have that so $|T| = d(T) = |TS|/|S|$.

4.I.2.21 We will first argue that a rank r matrix has a $r \times r$ submatrix with nonzero determinant. A rank r matrix has a linearly independent set of r rows. A matrix made from those rows will have row rank r and thus has column rank r. Conclusion: from those r rows can be extracted a linearly independent set of r columns, and so the original matrix has a $r \times r$ submatrix of rank r.

We finish by showing that if r is the largest such integer then the rank of the matrix is r. We need only show, by the maximality of r, that if a matrix has a $k \times k$ submatrix of nonzero determinant then the rank of the matrix is at least k. Consider such a $k \times k$ submatrix. Its rows are parts of the rows of the original matrix, clearly the set of whole rows is linearly independent. Thus the row rank of the original matrix is at least k, and the row rank of a matrix equals its rank.

4.I.2.23 *This is how the answer was given in the cited source.* The value $(1 - a^4)^3$ of the determinant is independent of the values B, C, D. Hence operation (e) does not change the value of the determinant but merely changes its appearance. Thus the element of likeness in (a), (b), (c), (d), and (e) is only that the appearance of the principle entity is changed. The same element appears in (f) changing the name-label of a rose, (g) writing a decimal integer in the scale of 12, (h) gilding the lily, (i) whitewashing a politician, and (j) granting an honorary degree.

. . . .

Answers for subsection 4.I.3

4.I.3.17 This is all of the permutations where $\phi(1) = 1$

$$\phi_1 = \langle 1,2,3,4 \rangle \quad \phi_2 = \langle 1,2,4,3 \rangle \quad \phi_3 = \langle 1,3,2,4 \rangle \quad \phi_4 = \langle 1,3,4,2 \rangle \quad \phi_5 = \langle 1,4,2,3 \rangle \quad \phi_6 = \langle 1,4,3,2 \rangle$$

the ones where $\phi(1) = 1$

$$\phi_7 = \langle 2,1,3,4 \rangle \quad \phi_8 = \langle 2,1,4,3 \rangle \quad \phi_9 = \langle 2,3,1,4 \rangle \quad \phi_{10} = \langle 2,3,4,1 \rangle \quad \phi_{11} = \langle 2,4,1,3 \rangle \quad \phi_{12} = \langle 2,4,3,1 \rangle$$

the ones where $\phi(1) = 3$

$\phi_{13} = \langle 3,1,2,4\rangle \quad \phi_{14} = \langle 3,1,4,2\rangle \quad \phi_{15} = \langle 3,2,1,4\rangle \quad \phi_{16} = \langle 3,2,4,1\rangle \quad \phi_{17} = \langle 3,4,1,2\rangle \quad \phi_{18} = \langle 3,4,2,1\rangle$

and the ones where $\phi(1) = 4$.

$\phi_{19} = \langle 4,1,2,3\rangle \quad \phi_{20} = \langle 4,1,3,2\rangle \quad \phi_{21} = \langle 4,2,1,3\rangle \quad \phi_{22} = \langle 4,2,3,1\rangle \quad \phi_{23} = \langle 4,3,1,2\rangle \quad \phi_{24} = \langle 4,3,2,1\rangle$

4.I.3.18 Each of these is easy to check.

(a)

permutation	ϕ_1	ϕ_2
inverse	ϕ_1	ϕ_2

(b)

permutation	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6
inverse	ϕ_1	ϕ_2	ϕ_3	ϕ_5	ϕ_4	ϕ_6

4.I.3.19 For the 'if' half, the first condition of Definition 3.2 follows from taking $k_1 = k_2 = 1$ and the second condition follows from taking $k_2 = 0$.

The 'only if' half also routine. From $f(\vec{\rho}_1,\ldots,k_1\vec{v}_1 + k_2\vec{v}_2,\ldots,\vec{\rho}_n)$ the first condition of Definition 3.2 gives $= f(\vec{\rho}_1,\ldots,k_1\vec{v}_1,\ldots,\vec{\rho}_n) + f(\vec{\rho}_1,\ldots,k_2\vec{v}_2,\ldots,\vec{\rho}_n)$ and the second condition, applied twice, gives the result.

4.I.3.20 To get a nonzero term in the permutation expansion we must use the $1,2$ entry and the $4,3$ entry. Having fixed on those two we must also use the $2,1$ entry and the the $3,4$ entry. The signum of $\langle 2,1,4,3\rangle$ is $+1$ because from

$$\begin{pmatrix} 0&1&0&0\\ 1&0&0&0\\ 0&0&0&1\\ 0&0&1&0 \end{pmatrix}$$

the two rwo swaps $\rho_1 \leftrightarrow \rho_2$ and $\rho_3 \leftrightarrow \rho_4$ will produce the identity matrix.

4.I.3.21 They would all double.

4.I.3.22 For the second statement, given a matrix, transpose it, swap rows, and transpose back. The result is swapped columns, and the determinant changes by a factor of -1. The third statement is similar: given a matrix, transpose it, apply multilinearity to what are now rows, and then transpose back the resulting matrices.

4.I.3.24 False.

$$\begin{vmatrix} 1&-1\\ 1&1 \end{vmatrix} = 2$$

4.I.3.25

(a) For the column index of the entry in the first row there are five choices. Then, for the column index of the entry in the second row there are four choices (the column index used in the first row cannot be used here). Continuing, we get $5\cdot 4\cdot 3\cdot 2\cdot 1 = 120$. (See also the next question.)

(b) Once we choose the second column in the first row, we can choose the other entries in $4\cdot 3\cdot 2\cdot 1 = 24$ ways.

4.I.3.26 $n\cdot(n-1)\cdots 2\cdot 1 = n!$

4.I.3.27 In $|A| = |A^{\text{trans}}| = |-A| = (-1)^n|A|$ the exponent n must be even.

4.I.3.30 Let T be $n\times n$, let J be $p\times p$, and let K be $q\times q$. Apply the permutation expansion formula

$$|T| = \sum_{\text{permutations } \phi} t_{1,\phi(1)}t_{2,\phi(2)}\ldots t_{n,\phi(n)}\,|P_\phi|$$

Because the upper right of T is all zeroes, if a ϕ has at least one of $p+1,\ldots,n$ among its first p column numbers $\phi(1),\ldots,\phi(p)$ then the term arising from ϕ is 0 (e.g., if $\phi(1) = n$ then $t_{1,\phi(1)}t_{2,\phi(2)}\ldots t_{n,\phi(n)}$ is 0). So the above formula reduces to a sum over all permutations with two halves: first $1,\ldots,p$ are rearranged, and after that comes a permutation of $p+1,\ldots,p+q$. To see this gives $|J|\cdot|K|$, distribute.

$$\Bigl[\sum_{\substack{\text{perms } \phi_1\\ \text{of } 1,\ldots,p}} t_{1,\phi_1(1)}\cdots t_{p,\phi_1(p)}\,|P_{\phi_1}|\Bigr]\cdot\Bigl[\sum_{\substack{\text{perms } \phi_2\\ \text{of } p+1,\ldots,p+q}} t_{p+1,\phi_2(p+1)}\cdots t_{p+q,\phi_2(p+q)}\,|P_{\phi_2}|\Bigr]$$

4.I.3.32 *This is how the answer was given in the cited source.* When two rows of a determinant are interchanged, the sign of the determinant is changed. When the rows of a three-by-three determinant are permuted, 3 positive and 3 negative determinants equal in absolute value are obtained. Hence the 9! determinants fall into 9!/6 groups, each of which sums to zero.

4.I.3.33 *This is how the answer was given in the cited source.* When the elements of any column are subtracted from the elements of each of the other two, the elements in two of the columns of the derived determinant are proportional, so the determinant vanishes. That is,

$$\begin{vmatrix}2&1&x-4\\4&2&x-3\\6&3&x-10\end{vmatrix}=\begin{vmatrix}1&x-3&-1\\2&x-1&-2\\3&x-7&-3\end{vmatrix}=\begin{vmatrix}x-2&-1&-2\\x+1&-2&-4\\x-4&-3&-6\end{vmatrix}=0.$$

4.I.3.34 *This is how the answer was given in the cited source.* Let

$$\begin{matrix}a&b&c\\d&e&f\\g&h&i\end{matrix}$$

have magic sum $N=S/3$. Then

$$\begin{aligned}N&=(a+e+i)+(d+e+f)+(g+e+c)\\&\quad-(a+d+g)-(c+f+i)=3e\end{aligned}$$

and $S=9e$. Hence, adding rows and columns,

$$D=\begin{vmatrix}a&b&c\\d&e&f\\g&h&i\end{vmatrix}=\begin{vmatrix}a&b&c\\d&e&f\\3e&3e&3e\end{vmatrix}=\begin{vmatrix}a&b&3e\\d&e&3e\\3e&3e&9e\end{vmatrix}=\begin{vmatrix}a&b&e\\d&e&e\\1&1&1\end{vmatrix}S.$$

4.I.3.35 *This is how the answer was given in the cited source.* Denote by D_n the determinant in question and by $a_{i,j}$ the element in the i-th row and j-th column. Then from the law of formation of the elements we have

$$a_{i,j}=a_{i,j-1}+a_{i-1,j},\qquad a_{1,j}=a_{i,1}=1.$$

Subtract each row of D_n from the row following it, beginning the process with the last pair of rows. After the $n-1$ subtractions the above equality shows that the element $a_{i,j}$ is replaced by the element $a_{i,j-1}$, and all the elements in the first column, except $a_{1,1}=1$, become zeroes. Now subtract each column from the one following it, beginning with the last pair. After this process the element $a_{i,j-1}$ is replaced by $a_{i-1,j-1}$, as shown in the above relation. The result of the two operations is to replace $a_{i,j}$ by $a_{i-1,j-1}$, and to reduce each element in the first row and in the first column to zero. Hence $D_n=D_{n+i}$ and consequently

$$D_n=D_{n-1}=D_{n-2}=\cdots=D_2=1.$$

. . . .

Answers for subsection 4.I.4

4.I.4.10 This is the permutation expansion of the determinant of a 2×2 matrix

$$\begin{vmatrix}a&b\\c&d\end{vmatrix}=ad\cdot\begin{vmatrix}1&0\\0&1\end{vmatrix}+bc\cdot\begin{vmatrix}0&1\\1&0\end{vmatrix}$$

and the permutation expansion of the determinant of its transpose.

$$\begin{vmatrix}a&c\\b&d\end{vmatrix}=ad\cdot\begin{vmatrix}1&0\\0&1\end{vmatrix}+cb\cdot\begin{vmatrix}0&1\\1&0\end{vmatrix}$$

As with the 3×3 expansions described in the subsection, the permutation matrices from corresponding terms are transposes (although this is disguised by the fact that each is self-transpose).

4.I.4.13 The pattern is this.

i	1	2	3	4	5	6	...
$\operatorname{sgn}(\phi_i)$	+1	−1	−1	+1	+1	−1	...

So to find the signum of $\phi_{n!}$, we subtract one $n!-1$ and look at the remainder on division by four. If the remainder is 1 or 2 then the signum is -1, otherwise it is $+1$. For $n>4$, the number $n!$ is divisible by four, so $n!-1$ leaves a remainder of -1 on division by four (more properly said, a remainder or 3), and so the signum is $+1$. The $n=1$ case has a signum of $+1$, the $n=2$ case has a signum of -1 and the $n=3$ case has a signum of -1.

4.I.4.14

(a) Permutations can be viewed as one-one and onto maps $\phi\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$. Any one-one and onto map has an inverse.

(b) If it always takes an odd number of swaps to get from P_ϕ to the identity, then it always takes an odd number of swaps to get from the identity to P_ϕ (any swap is reversible).

(c) This is the first question again.

4.I.4.15 If $\phi(i)=j$ then $\phi^{-1}(j)=i$. The result now follows on the observation that P_ϕ has a 1 in entry i,j if and only if $\phi(i)=j$, and $P_{\phi^{-1}}$ has a 1 in entry j,i if and only if $\phi^{-1}(j)=i$,

. . . .

Answers for subsection 4.II.1

4.II.1.8 For each, find the determinant and take the absolute value.

(a) 7 **(b)** 0 **(c)** 58

4.II.1.13 The starting area is 6 and the matrix changes sizes by -14. Thus the area of the image is 84.

4.II.1.14 By a factor of $21/2$.

4.II.1.15 For a box we take a sequence of vectors (as described in the remark, the order in which the vectors are taken matters), while for a span we take a set of vectors. Also, for a box subset of $\mathbb{R}^n$ there must be n vectors; of course for a span there can be any number of vectors. Finally, for a box the coefficients $t_1,\ldots,t_n$ are restricted to the interval $[0..1]$, while for a span the coefficients are free to range over all of $\mathbb{R}$.

4.II.1.18

(a) If it is defined then it is $(3^2)\cdot(2)\cdot(2^{-2})\cdot(3)$.

(b) $|6A^3+5A^2+2A|=|A|\cdot|6A^2+5A+2I|$.

4.II.1.24 Any permutation matrix has the property that the transpose of the matrix is its inverse.

For the implication, we know that $|A^{\text{trans}}|=|A|$. Then $1=|A\cdot A^{-1}|=|A\cdot A^{\text{trans}}|=|A|\cdot|A^{\text{trans}}|=|A|^2$.

The converse does not hold; here is an example.

$$\begin{pmatrix}3&1\\2&1\end{pmatrix}$$

4.II.1.25 Where the sides of the box are c times longer, the box has c^3 times as many cubic units of volume.

4.II.1.27

(a) The new basis is the old basis rotated by $\pi/4$.

(b) $\langle\begin{pmatrix}-1\\0\end{pmatrix},\begin{pmatrix}0\\-1\end{pmatrix}\rangle$, $\langle\begin{pmatrix}0\\-1\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\rangle$

(c) In each case the determinant is $+1$ (these bases are said to have *positive orientation*).

(d) Because only one sign can change at a time, the only other cycle possible is

$$\cdots\longrightarrow\begin{pmatrix}+\\+\end{pmatrix}\longrightarrow\begin{pmatrix}+\\-\end{pmatrix}\longrightarrow\begin{pmatrix}-\\-\end{pmatrix}\longrightarrow\begin{pmatrix}-\\+\end{pmatrix}\longrightarrow\cdots.$$

Here each associated determinant is -1 (such bases are said to have a *negative orientation*).

(e) There is one positively oriented basis $\langle(1)\rangle$ and one negatively oriented basis $\langle(-1)\rangle$.

(f) There are 48 bases (6 half-axis choices are possible for the first unit vector, 4 for the second, and 2 for the last). Half are positively oriented like the standard basis on the left below, and half are negatively oriented like the one on the right

In $\mathbb{R}^3$ positive orientation is sometimes called 'right hand orientation' because if a person's right hand is placed with the fingers curling from $\vec{e}_1$ to $\vec{e}_2$ then the thumb will point with $\vec{e}_3$.

4.II.1.28 We will compare $\det(\vec{s}_1,\dots,\vec{s}_n)$ with $\det(t(\vec{s}_1),\dots,t(\vec{s}_n))$ to show that the second differs from the first by a factor of $|T|$. We represent the $\vec{s}$'s with respect to the standard bases

$$\operatorname{Rep}_{\mathcal{E}_n}(\vec{s}_i)=\begin{pmatrix}s_{1,i}\\ s_{2,i}\\ \vdots\\ s_{n,i}\end{pmatrix}$$

and then we represent the map application with matrix-vector multiplication

$$\begin{aligned}\operatorname{Rep}_{\mathcal{E}_n}(\,t(\vec{s}_i)\,)&=\begin{pmatrix}t_{1,1}&t_{1,2}&\dots&t_{1,n}\\ t_{2,1}&t_{2,2}&\dots&t_{2,n}\\ &\vdots\\ t_{n,1}&t_{n,2}&\dots&t_{n,n}\end{pmatrix}\begin{pmatrix}s_{1,j}\\ s_{2,j}\\ \vdots\\ s_{n,j}\end{pmatrix}\\ &=s_{1,j}\begin{pmatrix}t_{1,1}\\ t_{2,1}\\ \vdots\\ t_{n,1}\end{pmatrix}+s_{2,j}\begin{pmatrix}t_{1,2}\\ t_{2,2}\\ \vdots\\ t_{n,2}\end{pmatrix}+\dots+s_{n,j}\begin{pmatrix}t_{1,n}\\ t_{2,n}\\ \vdots\\ t_{n,n}\end{pmatrix}\\ &=s_{1,j}\vec{t}_1+s_{2,j}\vec{t}_2+\dots+s_{n,j}\vec{t}_n\end{aligned}$$

where $\vec{t}_i$ is column i of T. Then $\det(t(\vec{s}_1),\,\dots,\,t(\vec{s}_n))$ equals $\det(s_{1,1}\vec{t}_1{+}s_{2,1}\vec{t}_2{+}\dots{+}s_{n,1}\vec{t}_n,\,\dots,\,s_{1,n}\vec{t}_1{+}s_{2,n}\vec{t}_2{+}\dots{+}s_{n,n}\vec{t}_n)$.

As in the derivation of the permutation expansion formula, we apply multilinearity, first splitting along the sum in the first argument $\det(s_{1,1}\vec{t}_1,\,\dots,\,s_{1,n}\vec{t}_1+s_{2,n}\vec{t}_2+\dots+s_{n,n}\vec{t}_n)+\dots+\det(s_{n,1}\vec{t}_n,\,\dots,\,s_{1,n}\vec{t}_1+s_{2,n}\vec{t}_2+\dots+s_{n,n}\vec{t}_n)$ and then splitting each of those n summands along the sums in the second arguments, etc. We end with, as in the derivation of the permutation expansion, n^n summand determinants, each of the form $\det(s_{i_1,1}\vec{t}_{i_1},s_{i_2,2}\vec{t}_{i_2},\,\dots,\,s_{i_n,n}\vec{t}_{i_n})$. Factor out each of the $s_{i,j}$'s $=s_{i_1,1}s_{i_2,2}\dots s_{i_n,n}\cdot\det(\vec{t}_{i_1},\vec{t}_{i_2},\,\dots,\,\vec{t}_{i_n})$.

As in the permutation expansion derivation, whenever two of the indices in $i_1,\,\dots,\,i_n$ are equal then the determinant has two equal arguments, and evaluates to 0. So we need only consider the cases where $i_1,\,\dots,\,i_n$ form a permutation of the numbers $1,\,\dots,\,n$. We thus have

$$\det(t(\vec{s}_1),\dots,t(\vec{s}_n))=\sum_{\text{permutations }\phi}s_{\phi(1),1}\dots s_{\phi(n),n}\det(\vec{t}_{\phi(1)},\dots,\vec{t}_{\phi(n)}).$$

Swap the columns in $\det(\vec{t}_{\phi(1)},\dots,\vec{t}_{\phi(n)})$ to get the matrix T back, which changes the sign by a factor of $\operatorname{sgn}\phi$, and then factor out the determinant of T.

$$=\sum_\phi s_{\phi(1),1}\dots s_{\phi(n),n}\det(\vec{t}_1,\dots,\vec{t}_n)\cdot\operatorname{sgn}\phi=\det(T)\sum_\phi s_{\phi(1),1}\dots s_{\phi(n),n}\cdot\operatorname{sgn}\phi.$$

As in the proof that the determinant of a matrix equals the determinant of its transpose, we commute the s's so they are listed by ascending row number instead of by ascending column number (and we substitute $\operatorname{sgn}(\phi^{-1})$ for $\operatorname{sgn}(\phi)$).

$$=\det(T)\sum_\phi s_{1,\phi^{-1}(1)}\dots s_{n,\phi^{-1}(n)}\cdot\operatorname{sgn}\phi^{-1}=\det(T)\det(\vec{s}_1,\vec{s}_2,\dots,\vec{s}_n)$$

. . . .

Answers for subsection 4.III.1

4.III.1.15 $\operatorname{adj}(T) = \begin{pmatrix} T_{1,1} & T_{2,1} & T_{3,1} \\ T_{1,2} & T_{2,2} & T_{3,2} \\ T_{1,3} & T_{2,3} & T_{3,3} \end{pmatrix} = \begin{pmatrix} +\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} & +\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \\ -\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} & +\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \\ +\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} & +\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}$

4.III.1.18 $\begin{pmatrix} T_{1,1} & T_{2,1} & T_{3,1} & T_{4,1} \\ T_{1,2} & T_{2,2} & T_{3,2} & T_{4,2} \\ T_{1,3} & T_{2,3} & T_{3,3} & T_{4,3} \\ T_{1,4} & T_{2,4} & T_{3,4} & T_{4,4} \end{pmatrix} = \begin{pmatrix} 4 & -3 & 2 & -1 \\ -3 & 6 & -4 & 2 \\ 2 & -4 & 6 & -3 \\ -1 & 2 & -3 & 4 \end{pmatrix}$

4.III.1.23 Consider this diagonal matrix.

$$D = \begin{pmatrix} d_1 & 0 & 0 & \dots & \\ 0 & d_2 & 0 & & \\ 0 & 0 & d_3 & & \\ & & & \ddots & \\ & & & & d_n \end{pmatrix}$$

If $i \neq j$ then the i, j minor is an $(n-1)\times(n-1)$ matrix with only $n-2$ nonzero entries, because both d_i and d_j are deleted. Thus, at least one row or column of the minor is all zeroes, and so the cofactor $D_{i,j}$ is zero. If $i = j$ then the minor is the diagonal matrix with entries $d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_n$. Its determinant is obviously $(-1)^{i+j} = (-1)^{2i} = 1$ times the product of those.

$$\operatorname{adj}(D) = \begin{pmatrix} d_2 \cdots d_n & 0 & & 0 \\ 0 & d_1 d_3 \cdots d_n & & 0 \\ & & \ddots & \\ & & & d_1 \cdots d_{n-1} \end{pmatrix}$$

By the way, Theorem 4.III.1.9 provides a slicker way to derive this conclusion.

4.III.1.25 It is false; here is an example.

$$T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \qquad \operatorname{adj}(T) = \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix} \qquad \operatorname{adj}(\operatorname{adj}(T)) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4.III.1.26

(a) An example

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

suggests the right answer.

$$\operatorname{adj}(M) = \begin{pmatrix} M_{1,1} & M_{2,1} & M_{3,1} \\ M_{1,2} & M_{2,2} & M_{3,2} \\ M_{1,3} & M_{2,3} & M_{3,3} \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} 4 & 5 \\ 0 & 6 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ 0 & 6 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} \\ -\begin{vmatrix} 0 & 5 \\ 0 & 6 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 0 & 6 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 0 & 5 \end{vmatrix} \\ \begin{vmatrix} 0 & 4 \\ 0 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 0 & 4 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 24 & -12 & -2 \\ 0 & 6 & -5 \\ 0 & 0 & 4 \end{pmatrix}$$

The result is indeed upper triangular.

A check of this is detailed but not hard. The entries in the upper triangle of the adjoint are $M_{a,b}$ where $a > b$. We need to verify that the cofactor $M_{a,b}$ is zero if $a > b$. With $a > b$, row a and column b of M,

$$\begin{pmatrix} m_{1,1} & \ldots & m_{1,b} & & \\ m_{2,1} & \ldots & m_{2,b} & & \\ \vdots & & \vdots & & \\ m_{a,1} & \ldots & m_{a,b} & \ldots & m_{a,n} \\ & & \vdots & & \\ & & m_{n,b} & & \end{pmatrix}$$

when deleted, leave an upper triangular minor, because entry i,j of the minor is either entry i,j of M (this happens if $a > i$ and $b > j$; in this case $i < j$ implies that the entry is zero) or it is entry $i,j+1$ of M (this happens if $i < a$ and $j > b$; in this case, $i < j$ implies that $i < j+1$, which implies that the entry is zero), or it is entry $i+1,j+1$ of M (this last case happens when $i > a$ and $j > b$; obviously here $i < j$ implies that $i+1 < j+1$ and so the entry is zero). Thus the determinant of the minor is the product down the diagonal. Observe that the $a-1,a$ entry of M is the $a-1,a-1$ entry of the minor (it doesn't get deleted because the relation $a > b$ is strict). But this entry is zero because M is upper triangular and $a-1 < a$. Therefore the cofactor is zero, and the adjoint is upper triangular. (The lower triangular case is similar.)

(b) This is immediate from the prior part, by Corollary 1.11.

4.III.1.27 We will show that each determinant can be expanded along row i. The argument for column j is similar.

Each term in the permutation expansion contains one and only one entry from each row. As in Example 1.1, factor out each row i entry to get $|T| = t_{i,1}\cdot\hat{T}_{i,1}+\cdots+t_{i,n}\cdot\hat{T}_{i,n}$, where each $\hat{T}_{i,j}$ is a sum of terms not containing any elements of row i. We will show that $\hat{T}_{i,j}$ is the i,j cofactor.

Consider the $i,j = n,n$ case first:

$$t_{n,n}\cdot\hat{T}_{n,n} = t_{n,n}\cdot\sum_{\phi} t_{1,\phi(1)}t_{2,\phi(2)}\ldots t_{n-1,\phi(n-1)}\operatorname{sgn}(\phi)$$

where the sum is over all n-permutations ϕ such that $\phi(n) = n$. To show that $\hat{T}_{i,j}$ is the minor $T_{i,j}$, we need only show that if ϕ is an n-permutation such that $\phi(n) = n$ and σ is an $n-1$-permutation with $\sigma(1) = \phi(1)$, …, $\sigma(n-1) = \phi(n-1)$ then $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\phi)$. But that's true because ϕ and σ have the same number of inversions.

Back to the general i,j case. Swap adjacent rows until the i-th is last and swap adjacent columns until the j-th is last. Observe that the determinant of the i,j-th minor is not affected by these adjacent swaps because inversions are preserved (since the minor has the i-th row and j-th column omitted). On the other hand, the sign of $|T|$ and $\hat{T}_{i,j}$ is changed $n-i$ plus $n-j$ times. Thus $\hat{T}_{i,j} = (-1)^{n-i+n-j}|T_{i,j}| = (-1)^{i+j}|T_{i,j}|$.

4.III.1.28 This is obvious for the 1×1 base case.

For the inductive case, assume that the determinant of a matrix equals the determinant of its transpose for all 1×1, …, $(n-1)\times(n-1)$ matrices. Expanding on row i gives $|T| = t_{i,1}T_{i,1}+\ldots+t_{i,n}T_{i,n}$ and expanding on column i gives $|T^{\text{trans}}| = t_{1,i}(T^{\text{trans}})_{1,i}+\cdots+t_{n,i}(T^{\text{trans}})_{n,i}$ Since $(-1)^{i+j} = (-1)^{j+i}$ the signs are the same in the two summations. Since the j,i minor of T^{trans} is the transpose of the i,j minor of T, the inductive hypothesis gives $|(T^{\text{trans}})_{i,j}| = |T_{i,j}|$.

4.III.1.29 *This is how the answer was given in the cited source.* Denoting the above determinant by D_n, it is seen that $D_2 = 1$, $D_3 = 2$. It remains to show that $D_n = D_{n-1}+D_{n-2}$, $n \geq 4$. In D_n subtract the $(n-3)$-th column from the $(n-1)$-th, the $(n-4)$-th from the $(n-2)$-th, …, the first from the third, obtaining

$$F_n = \begin{vmatrix} 1 & -1 & 0 & 0 & 0 & 0 & \ldots \\ 1 & 1 & -1 & 0 & 0 & 0 & \ldots \\ 0 & 1 & 1 & -1 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 1 & -1 & 0 & \ldots \\ . & . & . & . & . & . & \ldots \end{vmatrix}.$$

By expanding this determinant with reference to the first row, there results the desired relation.

. . . .

Answers for Topic: Cramer's Rule

1

(a) $x = 1$, $y = -3$ **(b)** $x = -2$, $y = -2$

2 $z = 1$

3 Determinants are unchanged by pivots, including column pivots, so $\det(B_i) = \det(\vec{a}_1, \ldots, x_1\vec{a}_1 + \cdots + x_i\vec{a}_i + \cdots + x_n\vec{a}_n, \ldots, \vec{a}_n)$ is equal to $\det(\vec{a}_1, \ldots, x_i\vec{a}_i, \ldots, \vec{a}_n)$ (use the operation of taking $-x_1$ times the first column and adding it to the i-th column, etc.). That is equal to $x_i \cdot \det(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) = x_i \cdot \det(A)$, as required.

4 Because the determinant of A is 1, Cramer's Rule applies, and shows that $x_i = |B_i|$. With B_i a matrix of integers, its determinant is an integer.

5 The solution of

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned}$$

is

$$x = \frac{ed - fb}{ad - bc} \qquad y = \frac{af - ec}{ad - bc}$$

provided of course that the denominators are not zero.

6 Of course, singular systems have $|A|$ equal to zero, but the infinitely many solutions case is characterized by the fact that all of the $|B_i|$ are zero as well.

. . . .

Answers for Topic: Speed of Calculating Determinants

1

(a) Under Octave, `rank(rand(5))` finds the rank of a 5×5 matrix whose entries are (uniformily distributed) in the interval $[0..1)$. This loop which runs the test 5000 times

```
octave:1> for i=1:5000
> if rank(rand(5))<5 printf("That's one."); endif
> endfor
```

produces (after a few seconds) returns the prompt, with no output.

The Octave script

```
function elapsed_time = detspeed (size)
  a=rand(size);
  tic();
  for i=1:10
     det(a);
  endfor
  elapsed_time=toc();
endfunction
```

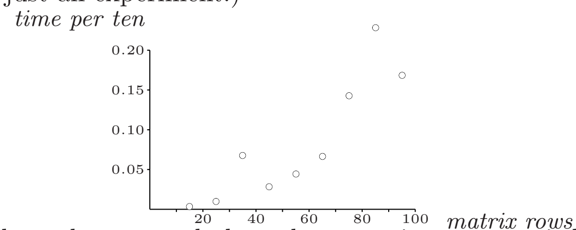
lead to this session.

```
octave:1> detspeed(5)
ans = 0.019505
octave:2> detspeed(15)
ans = 0.0054691
octave:3> detspeed(25)
ans = 0.0097431
octave:4> detspeed(35)
ans = 0.017398
```

(b) Here is the data (rounded a bit), and the graph.

matrix rows	15	25	35	45	55	65	75	85	95
time per ten	0.0034	0.0098	0.0675	0.0285	0.0443	0.0663	0.1428	0.2282	0.1686

(This data is from an average of twenty runs of the above script, because of the possibility that the randomly chosen matrix happens to take an unusually long or short time. Even so, the timing cannot be relied on too heavily; this is just an experiment.)



2 The number of operations depends on exactly how the operations are carried out.

(a) The determinant is -11. To row reduce takes a single pivot with two multiplications ($-5/2$ times 2 plus 5, and $-5/2$ times 1 plus -3) and the product down the diagonal takes one more multiplication. The permutation expansion takes two multiplications (2 times -3 and 5 times 1).

(b) The determinant is -39. Counting the operations is routine.

(c) The determinant is 4.

3 Because this question is open, any reasonable try is worthwhile. Here is a suggestion to get started: compare these under Octave: `tic(); det(rand(10)); toc()` versus `tic(); det(hilb(10)); toc()`, versus `tic(); det(eye(10)); toc()`, versus `tic(); det(zeroes(10)); toc()`.

4 This is a simple one.

```
DO 5 ROW=1, N
   PIVINV=1.0/A(ROW,ROW)
   DO 10 I=ROW+1, N
      DO 20 J=I, N
         A(I,J)=A(I,J)-PIVINV*A(ROW,J)
      20 CONTINUE
   10 CONTINUE
5 CONTINUE
```

5 Yes, because the J is in the innermost loop.

. . . .

Answers for Topic: Projective Geometry

1 From the dot product

$$0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} L_1 & L_2 & L_3 \end{pmatrix} = L_1$$

we get that the equation is $L_1 = 0$.

2

(a) This determinant

$$0 = \begin{vmatrix} 1 & 4 & x \\ 2 & 5 & y \\ 3 & 6 & z \end{vmatrix} = -3x + 6y - 3z$$

shows that the line is $L = \begin{pmatrix} -3 & 6 & -3 \end{pmatrix}$.

(b) $\begin{pmatrix} -3 \\ 6 \\ -3 \end{pmatrix}$

3 The line incident on

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

can be found from this determinant equation.

$$0 = \begin{vmatrix} u_1 & v_1 & x \\ u_2 & v_2 & y \\ u_3 & v_3 & z \end{vmatrix} = (u_2v_3 - u_3v_2) \cdot x + (u_3v_1 - u_1v_3) \cdot y + (u_1v_2 - u_2v_1) \cdot z$$

The equation for the point incident on two lines is the same.

4 If p_1, p_2, p_3, and q_1, q_2, q_3 are two triples of homogeneous coordinates for p then the two column vectors are in proportion, that is, lie on the same line through the origin. Similarly, the two row vectors are in proportion.

$$k \cdot \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \qquad m \cdot \begin{pmatrix} L_1 & L_2 & L_3 \end{pmatrix} = \begin{pmatrix} M_1 & M_2 & M_3 \end{pmatrix}$$

Then multiplying gives the answer $(km) \cdot (p_1L_1 + p_2L_2 + p_3L_3) = q_1M_1 + q_2M_2 + q_3M_3 = 0$.

5 The picture of the solar eclipse — unless the image plane is exactly perpendicular to the line from the sun through the pinhole — shows the circle of the sun projecting to an image that is an ellipse. (Another example is that in many pictures in this Topic, the circle that is the sphere's equator is drawn as an ellipse, that is, is seen by a viewer of the drawing as an ellipse.)

The solar eclipse picture also shows the converse. If we picture the projection as going from left to right through the pinhole then the ellipse I projects through P to a circle S.

6 A spot on the unit sphere

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

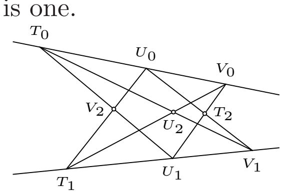
is non-equatorial if and only if $p_3 \neq 0$. In that case it corresponds to this point on the $z = 1$ plane

$$\begin{pmatrix} p_1/p_3 \\ p_2/p_3 \\ 1 \end{pmatrix}$$

since that is intersection of the line containing the vector and the plane.

7

(a) Other pictures are possible, but this is one.



The intersections $T_0U_1 \cap T_1U_0 = V_2$, $T_0V_1 \cap T_1V_0 = U_2$, and $U_0V_1 \cap U_1V_0 = T_2$ are labeled so that on each line is a T, a U, and a V.

(b) The lemma used in Desargue's Theorem gives a basis B with respect to which the points have these homogeneous coordinate vectors.

$$\mathrm{Rep}_B(\vec{t}_0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_B(\vec{t}_1) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_B(\vec{t}_2) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \mathrm{Rep}_B(\vec{v}_0) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(c) First, any U_0 on T_0V_0

$$\mathrm{Rep}_B(\vec{u}_0) = a \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a+b \\ b \\ b \end{pmatrix}$$

has homogeneous coordinate vectors of this form

$$\begin{pmatrix}u_0\\1\\1\end{pmatrix}$$

(u_0 is a parameter; it depends on where on the T_0V_0 line the point U_0 is, but any point on that line has a homogeneous coordinate vector of this form for some $u_0\in\mathbb{R}$). Similarly, U_2 is on T_1V_0

$$\mathrm{Rep}_B(\vec{u}_2)=c\begin{pmatrix}0\\1\\0\end{pmatrix}+d\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}d\\c+d\\d\end{pmatrix}$$

and so has this homogeneous coordinate vector.

$$\begin{pmatrix}1\\u_2\\1\end{pmatrix}$$

Also similarly, U_1 is incident on T_2V_0

$$\mathrm{Rep}_B(\vec{u}_1)=e\begin{pmatrix}0\\0\\1\end{pmatrix}+f\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}f\\f\\e+f\end{pmatrix}$$

and has this homogeneous coordinate vector.

$$\begin{pmatrix}1\\1\\u_1\end{pmatrix}$$

(d) Because V_1 is $T_0U_2\cap U_0T_2$ we have this.

$$g\begin{pmatrix}1\\0\\0\end{pmatrix}+h\begin{pmatrix}1\\u_2\\1\end{pmatrix}=i\begin{pmatrix}u_0\\1\\1\end{pmatrix}+j\begin{pmatrix}0\\0\\1\end{pmatrix}\qquad\Longrightarrow\qquad\begin{aligned}g+h&=iu_0\\hu_2&=i\\h&=i+j\end{aligned}$$

Substituting hu_2 for i in the first equation

$$\begin{pmatrix}hu_0u_2\\hu_2\\h\end{pmatrix}$$

shows that V_1 has this two-parameter homogeneous coordinate vector.

$$\begin{pmatrix}u_0u_2\\u_2\\1\end{pmatrix}$$

(e) Since V_2 is the intersection $T_0U_1\cap T_1U_0$

$$k\begin{pmatrix}1\\0\\0\end{pmatrix}+l\begin{pmatrix}1\\1\\u_1\end{pmatrix}=m\begin{pmatrix}0\\1\\0\end{pmatrix}+n\begin{pmatrix}u_0\\1\\1\end{pmatrix}\qquad\Longrightarrow\qquad\begin{aligned}k+l&=nu_0\\l&=m+n\\lu_1&=n\end{aligned}$$

and substituting lu_1 for n in the first equation

$$\begin{pmatrix}lu_0u_1\\l\\lu_1\end{pmatrix}$$

gives that V_2 has this two-parameter homogeneous coordinate vector.

$$\begin{pmatrix}u_0u_1\\1\\u_1\end{pmatrix}$$

(f) Because V_1 is on the T_1U_1 line its homogeneous coordinate vector has the form

$$p\begin{pmatrix}0\\1\\0\end{pmatrix}+q\begin{pmatrix}1\\1\\u_1\end{pmatrix}=\begin{pmatrix}q\\p+q\\qu_1\end{pmatrix} \qquad (*)$$

but a previous part of this question established that V_1's homogeneous coordinate vectors have the form

$$\begin{pmatrix}u_0u_2\\u_2\\1\end{pmatrix}$$

and so this a homogeneous coordinate vector for V_1.

$$\begin{pmatrix}u_0u_1u_2\\u_1u_2\\u_1\end{pmatrix} \qquad (**)$$

By $(*)$ and $(**)$, there is a relationship among the three parameters: $u_0u_1u_2=1$.

(g) The homogeneous coordinate vector of V_2 can be written in this way.

$$\begin{pmatrix}u_0u_1u_2\\u_2\\u_1u_2\end{pmatrix}=\begin{pmatrix}1\\u_2\\u_1u_2\end{pmatrix}$$

Now, the T_2U_2 line consists of the points whose homogeneous coordinates have this form.

$$r\begin{pmatrix}0\\0\\1\end{pmatrix}+s\begin{pmatrix}1\\u_2\\1\end{pmatrix}=\begin{pmatrix}s\\su_2\\r+s\end{pmatrix}$$

Taking $s=1$ and $r=u_1u_2-1$ shows that the homogeneous coordinate vectors of V_2 have this form.

Chapter 5. Similarity

. . . .

Answers for subsection 5.II.1

5.II.1.4 One way to proceed is left to right.

$$PSP^{-1}=\begin{pmatrix}4&2\\-3&2\end{pmatrix}\begin{pmatrix}1&3\\-2&-6\end{pmatrix}\begin{pmatrix}2/14&-2/14\\3/14&4/14\end{pmatrix}=\begin{pmatrix}0&0\\-7&-21\end{pmatrix}\begin{pmatrix}2/14&-2/14\\3/14&4/14\end{pmatrix}=\begin{pmatrix}0&0\\-11/2&-5\end{pmatrix}$$

5.II.1.6 Gauss' method shows that the first matrix represents maps of rank two while the second matrix represents maps of rank three.

5.II.1.7

(a) Because t is described with the members of B, finding the matrix representation is easy:

$$\mathrm{Rep}_B(t(x^2))=\begin{pmatrix}0\\1\\1\end{pmatrix}_B\qquad\mathrm{Rep}_B(t(x))=\begin{pmatrix}1\\0\\-1\end{pmatrix}_B\qquad\mathrm{Rep}_B(t(1))=\begin{pmatrix}0\\0\\3\end{pmatrix}_B$$

gives this.

$$\mathrm{Rep}_{B,B}(t)\begin{pmatrix}0&1&0\\1&0&0\\1&-1&3\end{pmatrix}$$

(b) We will find $t(1)$, $t(1+x)$, and $t(1+x+x^2$, to find how each is represented with respect to D. We are given that $t(1)=3$, and the other two are easy to see: $t(1+x)=x^2+2$ and $t(1+x+x^2)=x^2+x+3$. By eye, we get the representation of each vector

$$\mathrm{Rep}_D(t(1))=\begin{pmatrix}3\\0\\0\end{pmatrix}_D\qquad\mathrm{Rep}_D(t(1+x))=\begin{pmatrix}2\\-1\\1\end{pmatrix}_D\qquad\mathrm{Rep}_D(t(1+x+x^2))=\begin{pmatrix}2\\0\\1\end{pmatrix}_D$$

and thus the representation of the map.

$$\mathrm{Rep}_{D,D}(t)=\begin{pmatrix}3&2&2\\0&-1&0\\0&1&1\end{pmatrix}$$

(c) The diagram, adapted for this T and S,

$$\begin{array}{ccc}V_{\text{w.r.t. }D} & \xrightarrow[S]{t} & V_{\text{w.r.t. }D}\\ \mathrm{id}\big\downarrow P & & \mathrm{id}\big\downarrow P\\ V_{\text{w.r.t. }B} & \xrightarrow[T]{t} & V_{\text{w.r.t. }B}\end{array}$$

shows that $P=\mathrm{Rep}_{D,B}(\mathrm{id})$.

$$P=\begin{pmatrix}0&0&1\\0&1&1\\1&1&1\end{pmatrix}$$

5.II.1.9 The only representation of a zero map is a zero matrix, no matter what the pair of bases $\mathrm{Rep}_{B,D}(z)=Z$, and so in particular for any single basis B we have $\mathrm{Rep}_{B,B}(z)=Z$. The case of the identity is related, but slightly different: the only representation of the identity map, with respect to any B,B, is the identity $\mathrm{Rep}_{B,B}(\mathrm{id})=I$. (*Remark:* of course, we have seen examples where $B\neq D$ and $\mathrm{Rep}_{B,D}(\mathrm{id})\neq I$ — in fact, we have seen that any nonsingular matrix is a representation of the identity map with respect to some B,D.)

5.II.1.13 Let f_x and f_y be the reflection maps (sometimes called 'flip's). For any bases B and D, the matrices $\mathrm{Rep}_{B,B}(f_x)$ and $\mathrm{Rep}_{D,D}(f_y)$ are similar. First note that

$$S = \mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_x) = \begin{pmatrix}1&0\\0&-1\end{pmatrix} \qquad T = \mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_y) = \begin{pmatrix}-1&0\\0&1\end{pmatrix}$$

are similar because the second matrix is the representation of f_x with respect to the basis $A = \langle \vec{e}_2, \vec{e}_1\rangle$:

$$\begin{pmatrix}1&0\\0&-1\end{pmatrix} = P\begin{pmatrix}-1&0\\0&1\end{pmatrix}P^{-1}$$

where $P = \mathrm{Rep}_{A,\mathcal{E}_2}(\mathrm{id})$.

$$\begin{array}{ccc}
\mathbb{R}^2_{\text{w.r.t. } A} & \xrightarrow[T]{f_x} & V\mathbb{R}^2_{\text{w.r.t. } A} \\
\mathrm{id}\big\downarrow P & & \mathrm{id}\big\downarrow P \\
\mathbb{R}^2_{\text{w.r.t. } \mathcal{E}_2} & \xrightarrow[S]{f_x} & \mathbb{R}^2_{\text{w.r.t. } \mathcal{E}_2}
\end{array}$$

Now the conclusion follows from the transitivity part of Exercise 12.

To finish without relying on that exercise, write $\mathrm{Rep}_{B,B}(f_x) = QTQ^{-1} = Q\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_x)Q^{-1}$ and $\mathrm{Rep}_{D,D}(f_y) = RSR^{-1} = R\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_y)R^{-1}$. Using the equation in the first paragraph, the first of these two becomes $\mathrm{Rep}_{B,B}(f_x) = QP\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_y)P^{-1}Q^{-1}$ and rewriting the second of these two as $R^{-1}\cdot\mathrm{Rep}_{D,D}(f_y)\cdot R = \mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_y)$ and substituting gives the desired relationship

$$\begin{aligned}\mathrm{Rep}_{B,B}(f_x) &= QP\mathrm{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(f_y)P^{-1}Q^{-1}\\ &= QPR^{-1}\cdot\mathrm{Rep}_{D,D}(f_y)\cdot RP^{-1}Q^{-1} = (QPR^{-1})\cdot\mathrm{Rep}_{D,D}(f_y)\cdot(QPR^{-1})^{-1}\end{aligned}$$

Thus the matrices $\mathrm{Rep}_{B,B}(f_x)$ and $\mathrm{Rep}_{D,D}(f_y)$ are similar.

5.II.1.14 We must show that if two matrices are similar then they have the same determinant and the same rank. Both determinant and rank are properties of matrices that we have already shown to be preserved by matrix equivalence. They are therefore preserved by similarity (which is a special case of matrix equivalence: if two matrices are similar then they are matrix equivalent).

To prove the statement without quoting the results about matrix equivalence, note first that rank is a property of the map (it is the dimension of the rangespace) and since we've shown that the rank of a map is the rank of a representation, it must be the same for all representations. As for determinants, $|PSP^{-1}| = |P|\cdot|S|\cdot|P^{-1}| = |P|\cdot|S|\cdot|P|^{-1} = |S|$.

The converse of the statement does not hold; for instance, there are matrices with the same determinant that are not similar. To check this, consider a nonzero matrix with a determinant of zero. It is not similar to the zero matrix, the zero matrix is similar only to itself, but they have they same determinant. The argument for rank is much the same.

5.II.1.15 The matrix equivalence class containing all $n{\times}n$ rank zero matrices contains only a single matrix, the zero matrix. Therefore it has as a subset only one similarity class.

In contrast, the matrix equivalence class of $1{\times}1$ matrices of rank one consists of those $1{\times}1$ matrices (k) where $k \neq 0$. For any basis B, the representation of multiplication by the scalar k is $\mathrm{Rep}_{B,B}(t_k) = (k)$, so each such matrix is alone in its similarity class. So this is a case where a matrix equivalence class splits into infinitely many similarity classes.

5.II.1.16 Yes, these are similar

$$\begin{pmatrix}1&0\\0&3\end{pmatrix} \qquad \begin{pmatrix}3&0\\0&1\end{pmatrix}$$

since, where the first matrix is $\mathrm{Rep}_{B,B}(t)$ for $B = \langle\vec{\beta}_1,\vec{\beta}_2\rangle$, the second matrix is $\mathrm{Rep}_{D,D}(t)$ for $D = \langle\vec{\beta}_2,\vec{\beta}_1\rangle$.

5.II.1.19 There are two equivalence classes, (i) the class of rank zero matrices, of which there is one: $\mathscr{C}_1 = \{(0)\}$, and (2) the class of rank one matrices, of which there are infinitely many: $\mathscr{C}_2 = \{(k) \bigm| k \neq 0\}$.

Each $1{\times}1$ matrix is alone in its similarity class. That's because any transformation of a one-dimensional space is multiplication by a scalar $t_k\colon V \to V$ given by $\vec{v} \mapsto k\cdot\vec{v}$. Thus, for any basis $B = \langle\vec{\beta}\rangle$, the matrix representing a transformation t_k with respect to B, B is $(\mathrm{Rep}_B(t_k(\vec{\beta}))) = (k)$.

So, contained in the matrix equivalence class $\mathscr{C}_1$ is (obviously) the single similarity class consisting of the matrix (0). And, contained in the matrix equivalence class $\mathscr{C}_2$ are the infinitely many, one-member-each, similarity classes consisting of (k) for $k \neq 0$.

5.II.1.20 No. Here is an example that has two pairs, each of two similar matrices:

$$\begin{pmatrix}1&-1\\1&2\end{pmatrix}\begin{pmatrix}1&0\\0&3\end{pmatrix}\begin{pmatrix}2/3&1/3\\-1/3&1/3\end{pmatrix}=\begin{pmatrix}5/3&-2/3\\-4/3&7/3\end{pmatrix}$$

and

$$\begin{pmatrix}1&-2\\-1&1\end{pmatrix}\begin{pmatrix}-1&0\\0&-3\end{pmatrix}\begin{pmatrix}-1&-2\\-1&-1\end{pmatrix}=\begin{pmatrix}-5&-4\\2&1\end{pmatrix}$$

(this example is mostly arbitrary, but not entirely, because the the center matrices on the two left sides add to the zero matrix). Note that the sums of these similar matrices are not similar

$$\begin{pmatrix}1&0\\0&3\end{pmatrix}+\begin{pmatrix}-1&0\\0&-3\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}\qquad\begin{pmatrix}5/3&-2/3\\-4/3&7/3\end{pmatrix}+\begin{pmatrix}-5&-4\\2&1\end{pmatrix}\neq\begin{pmatrix}0&0\\0&0\end{pmatrix}$$

since the zero matrix is similar only to itself.

5.II.1.21 If $N = P(T-\lambda I)P^{-1}$ then $N = PTP^{-1} - P(\lambda I)P^{-1}$. The diagonal matrix λI commutes with anything, so $P(\lambda I)P^{-1} = PP^{-1}(\lambda I) = \lambda I$. Thus $N = PTP^{-1} - \lambda I$ and consequently $N + \lambda I = PTP^{-1}$. (So not only are they similar, in fact they are similar via the same P.)

. . . .

Answers for subsection 5.II.2

5.II.2.7

(a) Setting up

$$\begin{pmatrix}-2&1\\0&2\end{pmatrix}\begin{pmatrix}b_1\\b_2\end{pmatrix}=x\cdot\begin{pmatrix}b_1\\b_2\end{pmatrix}\qquad\Longrightarrow\qquad\begin{aligned}(-2-x)\cdot b_1+{}&&b_2&=0\\&&(2-x)\cdot b_2&=0\end{aligned}$$

gives the two possibilities that $b_2=0$ and $x=2$. Following the $b_2=0$ possibility leads to the first equation $(-2-x)b_1=0$ with the two cases that $b_1=0$ and that $x=-2$. Thus, under this first possibility, we find $x=-2$ and the associated vectors whose second component is zero, and whose first component is free.

$$\begin{pmatrix}-2&1\\0&2\end{pmatrix}\begin{pmatrix}b_1\\0\end{pmatrix}=-2\cdot\begin{pmatrix}b_1\\0\end{pmatrix}\qquad\vec{\beta}_1=\begin{pmatrix}1\\0\end{pmatrix}$$

Following the other possibility leads to a first equation of $-4b_1+b_2=0$ and so the vectors associated with this solution have a second component that is four times their first component.

$$\begin{pmatrix}-2&1\\0&2\end{pmatrix}\begin{pmatrix}b_1\\4b_1\end{pmatrix}=2\cdot\begin{pmatrix}b_1\\4b_1\end{pmatrix}\qquad\vec{\beta}_2=\begin{pmatrix}1\\4\end{pmatrix}$$

The diagonalization is this.

$$\begin{pmatrix}1&1\\0&4\end{pmatrix}^{-1}\begin{pmatrix}-2&1\\0&2\end{pmatrix}\begin{pmatrix}1&1\\0&4\end{pmatrix}^{-1}\begin{pmatrix}-2&0\\0&2\end{pmatrix}$$

(b) The calculations are like those in the prior part.

$$\begin{pmatrix}5&4\\0&1\end{pmatrix}\begin{pmatrix}b_1\\b_2\end{pmatrix}=x\cdot\begin{pmatrix}b_1\\b_2\end{pmatrix}\qquad\Longrightarrow\qquad\begin{aligned}(5-x)\cdot b_1+{}&&4\cdot b_2&=0\\&&(1-x)\cdot b_2&=0\end{aligned}$$

The bottom equation gives the two possibilities that $b_2=0$ and $x=1$. Following the $b_2=0$ possibility, and discarding the case where both b_2 and b_1 are zero, gives that $x=5$, associated with vectors whose second component is zero and whose first component is free.

$$\vec{\beta}_1=\begin{pmatrix}1\\0\end{pmatrix}$$

The $x = 1$ possibility gives a first equation of $4b_1 + 4b_2 = 0$ and so the associated vectors have a second component that is the negative of their first component.

$$\vec{\beta}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

We thus have this diagonalization.

$$\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 5 & 4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$$

5.II.2.9 These two are not similar

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

because each is alone in its similarity class.

For the second half, these

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \qquad \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

are similar via the matrix that changes bases from $\langle \vec{\beta}_1, \vec{\beta}_2 \rangle$ to $\langle \vec{\beta}_2, \vec{\beta}_1 \rangle$. (*Question.* Are two diagonal matrices similar if and only if their diagonal entries are permutations of each other's?)

5.II.2.10 Contrast these two.

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

The first is nonsingular, the second is singular.

5.II.2.12

(a) The check is easy.

$$\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 3 & 3 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

(b) It is a coincidence, in the sense that if $T = PSP^{-1}$ then T need not equal $P^{-1}SP$. Even in the case of a diagonal matrix D, the condition that $D = PTP^{-1}$ does not imply that D equals $P^{-1}TP$. The matrices from Example 2.2 show this.

$$\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 5 & -1 \end{pmatrix} \qquad \begin{pmatrix} 6 & 0 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -6 & 12 \\ -6 & 11 \end{pmatrix}$$

5.II.2.13 The columns of the matrix are chosen as the vectors associated with the x's. The exact choice, and the order of the choice was arbitrary. We could, for instance, get a different matrix by swapping the two columns.

5.II.2.14 Diagonalizing and then taking powers of the diagonal matrix shows that

$$\begin{pmatrix} -3 & 1 \\ -4 & 2 \end{pmatrix}^k = \frac{1}{3} \begin{pmatrix} -1 & 1 \\ -4 & 4 \end{pmatrix} + (\frac{-2}{3})^k \begin{pmatrix} 4 & -1 \\ 4 & -1 \end{pmatrix}.$$

5.II.2.16 Yes, ct is diagonalizable by the final theorem of this subsection.

No, $t + s$ need not be diagonalizable. Intuitively, the problem arises when the two maps diagonalize with respect to different bases (that is, when they are not *simultaneously diagonalizable*). Specifically, these two are diagonalizable but their sum is not:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

(the second is already diagonal; for the first, see Exercise 15). The sum is not diagonalizable because its square is the zero matrix.

The same intuition suggests that $t \circ s$ is not be diagonalizable. These two are diagonalizable but their product is not:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

(for the second, see Exercise 15).

5.II.2.18

(a) Using the formula for the inverse of a 2×2 matrix gives this.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\cdot\frac{1}{ad-bc}\cdot\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}=\frac{1}{ad-bc}\begin{pmatrix} ad+2bd-2ac-bc & -ab-2b^2+2a^2+ab \\ cd+2d^2-2c^2-cd & -bc-2bd+2ac+ad \end{pmatrix}$$

Now pick scalars $a,\ldots,d$ so that $ad-bc\neq 0$ and $2d^2-2c^2=0$ and $2a^2-2b^2=0$. For example, these will do.

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\cdot\frac{1}{-2}\cdot\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}=\frac{1}{-2}\begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix}$$

(b) As above,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ y & z \end{pmatrix}\cdot\frac{1}{ad-bc}\cdot\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}=\frac{1}{ad-bc}\begin{pmatrix} adx+bdy-acy-bcz & -abx-b^2y+a^2y+abz \\ cdx+d^2y-c^2y-cdz & -bcx-bdy+acy+adz \end{pmatrix}$$

we are looking for scalars $a,\ldots,d$ so that $ad-bc\neq 0$ and $-abx-b^2y+a^2y+abz=0$ and $cdx+d^2y-c^2y-cdz=0$, no matter what values x, y, and z have.

For starters, we assume that $y\neq 0$, else the given matrix is already diagonal. We shall use that assumption because if we (arbitrarily) let $a=1$ then we get

$$\begin{aligned} -bx-b^2y+y+bz&=0 \\ (-y)b^2+(z-x)b+y&=0 \end{aligned}$$

and the quadratic formula gives

$$b=\frac{-(z-x)\pm\sqrt{(z-x)^2-4(-y)(y)}}{-2y}\qquad y\neq 0$$

(note that if x, y, and z are real then these two b's are real as the discriminant is positive). By the same token, if we (arbitrarily) let $c=1$ then

$$\begin{aligned} dx+d^2y-y-dz&=0 \\ (y)d^2+(x-z)d-y&=0 \end{aligned}$$

and we get here

$$d=\frac{-(x-z)\pm\sqrt{(x-z)^2-4(y)(-y)}}{2y}\qquad y\neq 0$$

(as above, if $x,y,z\in\mathbb{R}$ then this discriminant is positive so a symmetric, real, 2×2 matrix is similar to a real diagonal matrix).

For a check we try $x=1$, $y=2$, $z=1$.

$$b=\frac{0\pm\sqrt{0+16}}{-4}=\mp 1\qquad d=\frac{0\pm\sqrt{0+16}}{4}=\pm 1$$

Note that not all four choices $(b,d)=(+1,+1),\ldots,(-1,-1)$ satisfy $ad-bc\neq 0$.

. . . .

Answers for subsection 5.II.3

5.II.3.20

(a) This

$$0=\begin{vmatrix} 10-x & -9 \\ 4 & -2-x \end{vmatrix}=(10-x)(-2-x)-(-36)$$

simplifies to the characteristic equation $x^2-8x+16=0$. Because the equation factors into $(x-4)^2$ there is only one eigenvalue $\lambda_1=4$.

(b) $0=(1-x)(3-x)-8=x^2-4x-5$; $\lambda_1=5$, $\lambda_2=-1$

(c) $x^2-21=0$; $\lambda_1=\sqrt{21}$, $\lambda_2=-\sqrt{21}$

(d) $x^2=0$; $\lambda_1=0$

(e) $x^2-2x+1=0$; $\lambda_1=1$

5.II.3.22 The characteristic equation

$$0=\begin{vmatrix}-2-x & -1\\ 5 & 2-x\end{vmatrix}=x^2+1$$

has the complex roots $\lambda_1=i$ and $\lambda_2=-i$. This system

$$\begin{array}{rr}(-2-x)\cdot b_1- & 1\cdot b_2=0\\ 5\cdot b_1 & (2-x)\cdot b_2=0\end{array}$$

For $\lambda_1=i$ Gauss' method gives this reduction.

$$\begin{array}{rr}(-2-i)\cdot b_1- & 1\cdot b_2=0\\ 5\cdot b_1-(2-i)\cdot b_2=0\end{array}\;\xrightarrow{(-5/(-2-i))\rho_1+\rho_2}\;\begin{array}{r}(-2-i)\cdot b_1-1\cdot b_2=0\\ 0=0\end{array}$$

(For the calculation in the lower right get a common denominator

$$\frac{5}{-2-i}-(2-i)=\frac{5}{-2-i}-\frac{-2-i}{-2-i}\cdot(2-i)=\frac{5-(-5)}{-2-i}$$

to see that it gives a $0=0$ equation.) These are the resulting eigenspace and eigenvector.

$$\{\begin{pmatrix}(1/(-2-i))b_2\\ b_2\end{pmatrix}\bigm| b_2\in\mathbb{C}\}\qquad\begin{pmatrix}1/(-2-i)\\ 1\end{pmatrix}$$

For $\lambda_2=-i$ the system

$$\begin{array}{rr}(-2+i)\cdot b_1- & 1\cdot b_2=0\\ 5\cdot b_1-(2+i)\cdot b_2=0\end{array}\;\xrightarrow{(-5/(-2+i))\rho_1+\rho_2}\;\begin{array}{r}(-2+i)\cdot b_1-1\cdot b_2=0\\ 0=0\end{array}$$

leads to this.

$$\{\begin{pmatrix}(1/(-2+i))b_2\\ b_2\end{pmatrix}\bigm| b_2\in\mathbb{C}\}\qquad\begin{pmatrix}1/(-2+i)\\ 1\end{pmatrix}$$

5.II.3.23 The characteristic equation is

$$0=\begin{vmatrix}1-x & 1 & 1\\ 0 & -x & 1\\ 0 & 0 & 1-x\end{vmatrix}=(1-x)^2(-x)$$

and so the eigenvalues are $\lambda_1=1$ (this is a repeated root of the equation) and $\lambda_2=0$. For the rest, consider this system.

$$\begin{array}{rrr}(1-x)\cdot b_1+ & b_2+ & b_3=0\\ & -x\cdot b_2+ & b_3=0\\ & & (1-x)\cdot b_3=0\end{array}$$

When $x=\lambda_1=1$ then the solution set is this eigenspace.

$$\{\begin{pmatrix}b_1\\ 0\\ 0\end{pmatrix}\bigm| b_1\in\mathbb{C}\}$$

When $x=\lambda_2=0$ then the solution set is this eigenspace.

$$\{\begin{pmatrix}-b_2\\ b_2\\ 0\end{pmatrix}\bigm| b_2\in\mathbb{C}\}$$

So these are eigenvectors associated with $\lambda_1=1$ and $\lambda_2=0$.

$$\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}\qquad\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix}$$

5.II.3.26 $\lambda=1,\begin{pmatrix}0&0\\0&1\end{pmatrix}$ and $\begin{pmatrix}2&3\\1&0\end{pmatrix}$, $\lambda=-2,\begin{pmatrix}-1&0\\1&0\end{pmatrix}$, $\lambda=-1,\begin{pmatrix}-2&1\\1&0\end{pmatrix}$

5.II.3.28 The determinant of the triangular matrix $T-xI$ is the product down the diagonal, and so it factors into the product of the terms $t_{i,i}-x$.

5.II.3.30 Any two representations of that transformation are similar, and similar matrices have the same characteristic polynomial.

5.II.3.33 The characteristic equation

$$0=\begin{vmatrix}a-x & b\\ c & d-x\end{vmatrix}=(a-x)(d-x)-bc$$

simplifies to $x^2+(-a-d)\cdot x+(ad-bc)$. Checking that the values $x=a+b$ and $x=a-c$ satisfy the equation (under the $a+b=c+d$ condition) is routine.

5.II.3.37

(a) Where the eigenvalue λ is associated with the eigenvector $\vec{x}$ then $A^k\vec{x}=A\cdots A\vec{x}=A^{k-1}\lambda\vec{x}=\lambda A^{k-1}\vec{x}=\cdots=\lambda^k\vec{x}$. (The full details can be put in by doing induction on k.)

(b) The eigenvector associated wih λ might not be an eigenvector associated with μ.

5.II.3.38 No. These are two same-sized, equal rank, matrices with different eigenvalues.

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\qquad\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$$

5.II.3.39 The characteristic polynomial has an odd power and so has at least one real root.

5.II.3.40 The characteristic polynomial x^3-5x^2+6x has distinct roots $\lambda_1=0$, $\lambda_2=-2$, and $\lambda_3=-3$. Thus the matrix can be diagonalized into this form.

$$\begin{pmatrix}0 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -3\end{pmatrix}$$

5.II.3.41 We must show that it is one-to-one and onto, and that it respects the operations of matrix addition and scalar multiplication.

To show that it is one-to-one, suppose that $t_P(T)=t_P(S)$, that is, suppose that $PTP^{-1}=PSP^{-1}$, and note that multiplying both sides on the left by P^{-1} and on the right by P gives that $T=S$. To show that it is onto, consider $S\in\mathcal{M}_{n\times n}$ and observe that $S=t_P(P^{-1}SP)$.

The map t_P preserves matrix addition since $t_P(T+S)=P(T+S)P^{-1}=(PT+PS)P^{-1}=PTP^{-1}+PSP^{-1}=t_P(T+S)$ follows from properties of matrix multiplication and addition that we have seen. Scalar multiplication is similar: $t_P(cT)=P(c\cdot T)P^{-1}=c\cdot(PTP^{-1})=c\cdot t_P(T)$.

5.II.3.42 *This is how the answer was given in the cited source.* If the argument of the characteristic function of A is set equal to c, adding the first $(n-1)$ rows (columns) to the nth row (column) yields a determinant whose nth row (column) is zero. Thus c is a characteristic root of A.

. . . .

Answers for subsection 5.III.1

5.III.1.8 For the zero transformation, no matter what the space, the chain of rangespaces is $V\supset\{\vec{0}\}=\{\vec{0}\}=\cdots$ and the chain of nullspaces is $\{\vec{0}\}\subset V=V=\cdots$. For the identity transformation the chains are $V=V=V=\cdots$ and $\{\vec{0}\}=\{\vec{0}\}=\cdots$.

5.III.1.9

(a) Iterating t_0 twice $a+bx+cx^2\mapsto b+cx^2\mapsto cx^2$ gives

$$a+bx+cx^2\overset{t_0^2}{\longmapsto}cx^2$$

and any higher power is the same map. Thus, while $\mathscr{R}(t_0)$ is the space of quadratic polynomials with no linear term $\{p+rx^2\bigm| p,r\in\mathbb{C}\}$, and $\mathscr{R}(t_0^2)$ is the space of purely-quadratic polynomials $\{rx^2\bigm| r\in\mathbb{C}\}$, this is where the chain stabilizes $\mathscr{R}_\infty(t_0)=\{rx^2\bigm| n\in\mathbb{C}\}$. As for nullspaces, $\mathscr{N}(t_0)$ is the space of purely-linear quadratic polynomials $\{qx\bigm| q\in\mathbb{C}\}$, and $\mathscr{N}(t_0^2)$ is the space of quadratic polynomials with no x^2 term $\{p+qx\bigm| p,q\in\mathbb{C}\}$, and this is the end $\mathscr{N}_\infty(t_0)=\mathscr{N}(t_0^2)$.

(b) The second power

$$\begin{pmatrix} a \\ b \end{pmatrix} \overset{t_1}{\longmapsto} \begin{pmatrix} 0 \\ a \end{pmatrix} \overset{t_1}{\longmapsto} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

is the zero map. Consequently, the chain of rangespaces

$$\mathbb{R}^2 \supset \{ \begin{pmatrix} 0 \\ p \end{pmatrix} \mid p \in \mathbb{C} \} \supset \{ \vec{0} \} = \cdots$$

and the chain of nullspaces

$$\{ \vec{0} \} \subset \{ \begin{pmatrix} q \\ 0 \end{pmatrix} \mid q \in \mathbb{C} \} \subset \mathbb{R}^2 = \cdots$$

each has length two. The generalized rangespace is the trivial subspace and the generalized nullspace is the entire space.

(c) Iterates of this map cycle around

$$a + bx + cx^2 \overset{t_2}{\longmapsto} b + cx + ax^2 \overset{t_2}{\longmapsto} c + ax + bx^2 \overset{t_2}{\longmapsto} a + bx + cx^2 \ \cdots$$

and the chains of rangespaces and nullspaces are trivial.

$$\mathcal{P}_2 = \mathcal{P}_2 = \cdots \qquad \{ \vec{0} \} = \{ \vec{0} \} = \cdots$$

Thus, obviously, generalized spaces are $\mathscr{R}_\infty(t_2) = \mathcal{P}_2$ and $\mathscr{N}_\infty(t_2) = \{ \vec{0} \}$.

(d) We have

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} a \\ a \\ b \end{pmatrix} \mapsto \begin{pmatrix} a \\ a \\ a \end{pmatrix} \mapsto \begin{pmatrix} a \\ a \\ a \end{pmatrix} \mapsto \cdots$$

and so the chain of rangespaces

$$\mathbb{R}^3 \supset \{ \begin{pmatrix} p \\ p \\ r \end{pmatrix} \mid p, r \in \mathbb{C} \} \supset \{ \begin{pmatrix} p \\ p \\ p \end{pmatrix} \mid p \in \mathbb{C} \} = \cdots$$

and the chain of nullspaces

$$\{ \vec{0} \} \subset \{ \begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix} \mid r \in \mathbb{C} \} \subset \{ \begin{pmatrix} 0 \\ q \\ r \end{pmatrix} \mid q, r \in \mathbb{C} \} = \cdots$$

each has length two. The generalized spaces are the final ones shown above in each chain.

5.III.1.10 Each maps $x \mapsto t(t(t(x)))$.

5.III.1.11 Recall that if W is a subspace of V then any basis B_W for W can be enlarged to make a basis B_V for V. From this the first sentence is immediate. The second sentence is also not hard: W is the span of B_W and if W is a proper subspace then V is not the span of B_W, and so B_V must have at least one vector more than does B_W.

5.III.1.12 It is both 'if' and 'only if'. We have seen earlier that a linear map is nonsingular if and only if it preserves dimension, that is, if the dimension of its range equals the dimension of its domain. With a transformation $t\colon V \to V$ that means that the map is nonsingular if and only if it is onto: $\mathscr{R}(t) = V$ (and thus $\mathscr{R}(t^2) = V$, etc).

5.III.1.13 The nullspaces form chains because because if $\vec{v} \in \mathscr{N}(t^j)$ then $t^j(\vec{v}) = \vec{0}$ and $t^{j+1}(\vec{v}) = t(\, t^j(\vec{v})\,) = t(\vec{0}) = \vec{0}$ and so $\vec{v} \in \mathscr{N}(t^{j+1})$.

Now, the "further" property for nullspaces follows from that fact that it holds for rangespaces, along with the prior exercise. Because the dimension of $\mathscr{R}(t^j)$ plus the dimension of $\mathscr{N}(t^j)$ equals the dimension n of the starting space V, when the dimensions of the rangespaces stop decreasing, so do the dimensions of the nullspaces. The prior exercise shows that from this point k on, the containments in the chain are not proper — the nullspaces are equal.

5.III.1.14 (Of course, many examples are correct, but here is one.) An example is the shift operator on triples of reals $(x, y, z) \mapsto (0, x, y)$. The nullspace is all triples that start with two zeros. The map stabilizes after three iterations.

5.III.1.15 The differentiation operator $d/dx\colon \mathcal{P}_1 \to \mathcal{P}_1$ has the same rangespace as nullspace. For an example of where they are disjoint — except for the zero vector — consider an identity map (or any nonsingular map).

. . . .

Answers for subsection 5.III.2

5.III.2.19 By Lemma 1.3 the nullity has grown as large as possible by the n-th iteration where n is the dimension of the domain. Thus, for the 2×2 matrices, we need only check whether the square is the zero matrix. For the 3×3 matrices, we need only check the cube.

(a) Yes, this matrix is nilpotent because its square is the zero matrix.

(b) No, the square is not the zero matrix.

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}^2 = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}$$

(c) Yes, the cube is the zero matrix. In fact, the square is zero.

(d) No, the third power is not the zero matrix.

$$\begin{pmatrix} 1 & 1 & 4 \\ 3 & 0 & -1 \\ 5 & 2 & 7 \end{pmatrix}^3 = \begin{pmatrix} 206 & 86 & 304 \\ 26 & 8 & 26 \\ 438 & 180 & 634 \end{pmatrix}$$

(e) Yes, the cube of this matrix is the zero matrix.

Another way to see that the second and fourth matrices are not nilpotent is to note that they are nonsingular.

5.III.2.23 A couple of examples

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ a & b \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ a & b & c \\ d & e & f \end{pmatrix}$$

suggest that left multiplication by a block of subdiagonal ones shifts the rows of a matrix downward. Distinct blocks

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a & b & c & d \\ 0 & 0 & 0 & 0 \\ i & j & k & l \end{pmatrix}$$

act to shift down distinct parts of the matrix.

Right multiplication does an analgous thing to columns. See Exercise 17.

5.III.2.24 Yes. Generalize the last sentence in Example 2.9. As to the index, that same last sentence shows that the index of the new matrix is less than or equal to the index of $\hat{N}$, and reversing the roles of the two matrices gives inequality in the other direction.

Another answer to this question is to show that a matrix is nilpotent if and only if any associated map is nilpotent, and with the same index. Then, because similar matrices represent the same map, the conclusion follows. This is Exercise 30 below.

5.III.2.26 No, by Lemma 1.3 for a map on a two-dimensional space, the nullity has grown as large as possible by the second iteration.

5.III.2.27 The index of nilpotency of a transformation can be zero only when the vector starting the string must be $\vec{0}$, that is, only when V is a trivial space.

5.III.2.29 We must check that $B\cup\hat{C}\cup\{\vec{v}_1,\dots,\vec{v}_j\}$ is linearly independent where B is a t-string basis for $\mathscr{R}(t)$, where $\hat{C}$ is a basis for $\mathscr{N}(t)$, and where $t(\vec{v}_1)=\vec{\beta}_1,\dots,t(\vec{v}_i=\vec{\beta}_i$. Write

$$\vec{0}=c_{1,-1}\vec{v}_1+c_{1,0}\vec{\beta}_1+c_{1,1}t(\vec{\beta}_1)+\cdots+c_{1,h_1}t^{h_1}(\vec{\vec{\beta}}_1)+c_{2,-1}\vec{v}_2+\cdots+c_{j,h_i}t^{h_i}(\vec{\beta}_i)$$

and apply t.

$$\vec{0}=c_{1,-1}\vec{\beta}_1+c_{1,0}t(\vec{\beta}_1)+\cdots+c_{1,h_1-1}t^{h_1}(\vec{\vec{\beta}}_1)+c_{1,h_1}\vec{0}+c_{2,-1}\vec{\beta}_2+\cdots+c_{i,h_i-1}t^{h_i}(\vec{\beta}_i)+c_{i,h_i}\vec{0}$$

Conclude that the coefficients $c_{1,-1},\dots,c_{1,h_i-1},c_{2,-1},\dots,c_{i,h_i-1}$ are all zero as $B\cup\hat{C}$ is a basis. Substitute back into the first displayed equation to conclude that the remaining coefficients are zero also.

5.III.2.30 For any basis B, a transformation n is nilpotent if and only if $N = \mathrm{Rep}_{B,B}(n)$ is a nilpotent matrix. This is because only the zero matrix represents the zero map and so n^j is the zero map if and only if N^j is the zero matrix.

5.III.2.31 It can be of any size greater than or equal to one. To have a transformation that is nilpotent of index four, whose cube has rangespace of dimension k, take a vector space, a basis for that space, and a transformation that acts on that basis in this way.

$$\begin{array}{ccccccccc}
\vec{\beta}_1 & \mapsto & \vec{\beta}_2 & \mapsto & \vec{\beta}_3 & \mapsto & \vec{\beta}_4 & \mapsto & \vec{0} \\
\vec{\beta}_5 & \mapsto & \vec{\beta}_6 & \mapsto & \vec{\beta}_7 & \mapsto & \vec{\beta}_8 & \mapsto & \vec{0} \\
 & & \vdots & & & & & & \\
\vec{\beta}_{4k-3} & \mapsto & \vec{\beta}_{4k-2} & \mapsto & \vec{\beta}_{4k-1} & \mapsto & \vec{\beta}_{4k} & \mapsto & \vec{0} \\
 & & \vdots & & & & & & \\
\end{array}$$

–possibly other, shorter, strings–

So the dimension of the rangespace of T^3 can be as large as desired. The smallest that it can be is one — there must be at least one string or else the map's index of nilpotency would not be four.

5.III.2.32 These two have only zero for eigenvalues

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

but are not similar (they have different canonical representatives, namely, themselves).

5.III.2.33 A simple reordering of the string basis will do. For instance, a map that is assoicated with this string basis

$$\vec{\beta}_1 \mapsto \vec{\beta}_2 \mapsto \vec{0}$$

is represented with respect to $B = \langle \vec{\beta}_1, \vec{\beta}_2 \rangle$ by this matrix

$$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

but is represented with respect to $B = \langle \vec{\beta}_2, \vec{\beta}_1 \rangle$ in this way.

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

5.III.2.35 For the matrices to be nilpotent they must be square. For them to commute they must be the same size. Thus their product and sum are defined.

Call the matrices A and B. To see that AB is nilpotent, multiply $(AB)^2 = ABAB = AABB = A^2B^2$, and $(AB)^3 = A^3B^3$, etc., and, as A is nilpotent, that product is eventually zero.

The sum is similar; use the Binomial Theorem.

5.III.2.36 Some experimentation gives the idea for the proof. Expansion of the second power

$$t_S^2(T) = S(ST - TS) - (ST - TS)S = S^2 - 2STS + TS^2$$

the third power

$$\begin{aligned} t_S^3(T) &= S(S^2 - 2STS + TS^2) - (S^2 - 2STS + TS^2)S \\ &= S^3T - 3S^2TS + 3STS^2 - TS^3 \end{aligned}$$

and the fourth power

$$\begin{aligned} t_S^4(T) &= S(S^3T - 3S^2TS + 3STS^2 - TS^3) - (S^3T - 3S^2TS + 3STS^2 - TS^3)S \\ &= S^4T - 4S^3TS + 6S^2TS^2 - 4STS^3 + TS^4 \end{aligned}$$

suggest that the expansions follow the Binomial Theorem. Verifying this by induction on the power of t_S is routine. This answers the question because, where the index of nilpotency of S is k, in the expansion of t_S^{2k}

$$t_S^{2k}(T) = \sum_{0 \leq i \leq 2k} (-1)^i \binom{2k}{i} S^i T S^{2k-i}$$

for any i at least one of the S^i and S^{2k-i} has a power higher than k, and so the term gives the zero matrix.

5.III.2.37 Use the geometric series: $I - N^{k+1} = (I-N)(N^k + N^{k-1} + \cdots + I)$. If N^{k+1} is the zero matrix then we have a right inverse for $I-N$. It is also a left inverse.

This statement is not 'only if' since

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

is invertible.

. . . .

Answers for subsection 5.IV.1

5.IV.1.15 Its characteristic polynomial has complex roots.

$$\begin{vmatrix} -x & 1 & 0 \\ 0 & -x & 1 \\ 1 & 0 & -x \end{vmatrix} = (1-x)\cdot(x-(-\frac{1}{2}+\frac{\sqrt{3}}{2}i))\cdot(x-(-\frac{1}{2}-\frac{\sqrt{3}}{2}i))$$

As the roots are distinct, the characteristic polynomial equals the minimal polynomial.

5.IV.1.18 The $n=3$ case provides a hint. A natural basis for $\mathcal{P}_3$ is $B = \langle 1, x, x^2, x^3\rangle$. The action of the transformation is

$$1 \mapsto 1 \quad x \mapsto x+1 \quad x^2 \mapsto x^2+2x+1 \quad x^3 \mapsto x^3+3x^2+3x+1$$

and so the representation $\mathrm{Rep}_{B,B}(t)$ is this upper triangular matrix.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Because it is triangular, the fact that the characteristic polynomial is $c(x) = (x-1)^4$ is clear. For the minimal polynomial, the candidates are $m_1(x) = (x-1)$,

$$T - 1I = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$m_2(x) = (x-1)^2$,

$$(T - 1I)^2 = \begin{pmatrix} 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$m_3(x) = (x-1)^3$,

$$(T - 1I)^3 = \begin{pmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and $m_4(x) = (x-1)^4$. Because m_1, m_2, and m_3 are not right, m_4 must be right, as is easily verified.

In the case of a general n, the representation is an upper triangular matrix with ones on the diagonal. Thus the characteristic polynomial is $c(x) = (x-1)^{n+1}$. One way to verify that the minimal polynomial equals the characteristic polynomial is argue something like this: say that an upper triangular matrix is 0-upper triangular if there are nonzero entries on the diagonal, that it is 1-upper triangular if the diagonal contains only zeroes and there are nonzero entries just above the diagonal, etc. As the above example illustrates, an induction argument will show that, where T has only nonnegative entries, T^j is j-upper triangular. That argument is left to the reader.

5.IV.1.19 The map twice is the same as the map once: $\pi \circ \pi = \pi$, that is, $\pi^2 = \pi$ and so the minimal polynomial is of degree at most two since $m(x) = x^2 - x$ will do. The fact that no linear polynomial will do follows from applying the maps on the left and right side of $c_1 \cdot \pi + c_0 \cdot \mathrm{id} = z$ (where z is the zero map) to these two vectors.

$$\begin{pmatrix}0\\0\\1\end{pmatrix} \qquad \begin{pmatrix}1\\0\\0\end{pmatrix}$$

Thus the minimal polynomial is m.

5.IV.1.20 This is one answer.

$$\begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix}$$

5.IV.1.21 The x must be a scalar, not a matrix.

5.IV.1.22 The characteristic polynomial of

$$T = \begin{pmatrix}a&b\\c&d\end{pmatrix}$$

is $(a-x)(d-x) - bc = x^2 - (a+d)x + (ad-bc)$. Substitute

$$\begin{aligned}\begin{pmatrix}a&b\\c&d\end{pmatrix}^2 &- (a+d)\begin{pmatrix}a&b\\c&d\end{pmatrix} + (ad-bc)\begin{pmatrix}1&0\\0&1\end{pmatrix}\\ &= \begin{pmatrix}a^2+bc&ab+bd\\ac+cd&bc+d^2\end{pmatrix} - \begin{pmatrix}a^2+ad&ab+bd\\ac+cd&ad+d^2\end{pmatrix} + \begin{pmatrix}ad-bc&0\\0&ad-bc\end{pmatrix}\end{aligned}$$

and just check each entry sum to see that the result is the zero matrix.

5.IV.1.25 A minimal polynomial must have leading coefficient 1, and so if the minimal polynomial of a map or matrix were to be a degree zero polynomial then it would be $m(x) = 1$. But the identity map or matrix equals the zero map or matrix only on a trivial vector space.

So in the nontrivial case the minimal polynomial must be of degree at least one. A zero map or matrix has minimal polynomial $m(x) = x$, and an identity map or matrix has minimal polynomial $m(x) = x - 1$.

5.IV.1.27 For a diagonal matrix

$$T = \begin{pmatrix}t_{1,1}&0&&\\0&t_{2,2}&&\\&&\ddots&\\&&&t_{n,n}\end{pmatrix}$$

the characteristic polynomial is $(t_{1,1} - x)(t_{2,2} - x)\cdots(t_{n,n} - x)$. Of course, some of those factors may be repeated, e.g., the matrix might have $t_{1,1} = t_{2,2}$. For instance, the characteristic polynomial of

$$D = \begin{pmatrix}3&0&0\\0&3&0\\0&0&1\end{pmatrix}$$

is $(3-x)^2(1-x) = -1 \cdot (x-3)^2(x-1)$.

To form the minimal polynomial, take the terms $x - t_{i,i}$, throw out repeats, and multiply them together. For instance, the minimal polynomial of D is $(x-3)(x-1)$. To check this, note first that Theorem 5.IV.1.8, the Cayley-Hamilton theorem, requires that each linear factor in the characteristic polynomial appears at least once in the minimal polynomial. One way to check the other direction — that in the case of a diagonal matrix, each linear factor need appear at most once — is to use a matrix argument. A diagonal matrix, multiplying from the left, rescales rows by the entry on the diagonal. But in a product $(T - t_{1,1}I)\cdots$, even without any repeat factors, every row is zero in at least one of the factors.

For instance, in the product

$$(D-3I)(D-1I) = (D-3I)(D-1I)I = \begin{pmatrix}0&0&0\\0&0&0\\0&0&-2\end{pmatrix}\begin{pmatrix}2&0&0\\0&2&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

because the first and second rows of the first matrix $D-3I$ are zero, the entire product will have a first row and second row that are zero. And because the third row of the middle matrix $D-1I$ is zero, the entire product has a third row of zero.

5.IV.1.29

(a) *This is a property of functions in general, not just of linear functions.* Suppose that f and g are one-to-one functions such that $f\circ g$ is defined. Let $f\circ g(x_1)=f\circ g(x_2)$, so that $f(g(x_1))=f(g(x_2))$. Because f is one-to-one this implies that $g(x_1)=g(x_2)$. Because g is also one-to-one, this in turn implies that $x_1=x_2$. Thus, in summary, $f\circ g(x_1)=f\circ g(x_2)$ implies that $x_1=x_2$ and so $f\circ g$ is one-to-one.

(b) If the linear map h is not one-to-one then there are unequal vectors $\vec{v}_1$, $\vec{v}_2$ that map to the same value $h(\vec{v}_1)=h(\vec{v}_2)$. Because h is linear, we have $\vec{0}=h(\vec{v}_1)-h(\vec{v}_2)=h(\vec{v}_1-\vec{v}_2)$ and so $\vec{v}_1-\vec{v}_2$ is a nonzero vector from the domain that is mapped by h to the zero vector of the codomain ($\vec{v}_1-\vec{v}_2$ does not equal the zero vector of the domain because $\vec{v}_1$ does not equal $\vec{v}_2$).

(c) The minimal polynomial $m(t)$ sends every vector in the domain to zero and so it is not one-to-one (except in a trivial space, which we ignore). By the first item of this question, since the composition $m(t)$ is not one-to-one, at least one of the components $t-\lambda_i$ is not one-to-one. By the second item, $t-\lambda_i$ has a nontrivial nullspace. Because $(t-\lambda_i)(\vec{v})=\vec{0}$ holds if and only if $t(\vec{v})=\lambda_i\cdot\vec{v}$, the prior sentence gives that λ_i is an eigenvalue (recall that the definition of eigenvalue requires that the relationship hold for at least one nonzero $\vec{v}$).

5.IV.1.30 This is false. The natural example of a non-diagonalizable transformation works here. Consider the transformation of $\mathbb{C}^2$ represented with respect to the standard basis by this matrix.

$$N=\begin{pmatrix}0&1\\0&0\end{pmatrix}$$

The characteristic polynomial is $c(x)=x^2$. Thus the minimal polynomial is either $m_1(x)=x$ or $m_2(x)=x^2$. The first is not right since $N-0\cdot I$ is not the zero matrix, thus in this example the minimal polynomial has degree equal to the dimension of the underlying space, and, as mentioned, we know this matrix is not diagonalizable because it is nilpotent.

5.IV.1.31 Let A and B be similar $A=PBP^{-1}$. From the facts that

$$A^n=(PBP^{-1})^n=(PBP^{-1})(PBP^{-1})\cdots(PBP^{-1})=PB(P^{-1}P)B(P^{-1}P)\cdots(P^{-1}P)BP^{-1}=PB^nP^{-1}$$

and $c\cdot A=c\cdot(PBP^{-1})=P(c\cdot B)P^{-1}$ follows the required fact that for any polynomial function f we have $f(A)=P\,f(B)\,P^{-1}$. For instance, if $f(x)=x^2+2x+3$ then

$$\begin{aligned}A^2+2A+3I&=(PBP^{-1})^2+2\cdot PBP^{-1}+3\cdot I\\&=(PBP^{-1})(PBP^{-1})+P(2B)P^{-1}+3\cdot PP^{-1}=P(B^2+2B+3I)P^{-1}\end{aligned}$$

shows that $f(A)$ is similar to $f(B)$.

(a) Taking f to be a linear polynomial we have that $A-xI$ is similar to $B-xI$. Similar matrices have equal determinants (since $|A|=|PBP^{-1}|=|P|\cdot|B|\cdot|P^{-1}|=1\cdot|B|\cdot 1=|B|$). Thus the characteristic polynomials are equal.

(b) As P and P^{-1} are invertible, $f(A)$ is the zero matrix when, and only when, $f(B)$ is the zero matrix.

(c) They cannot be similar since they don't have the same characteristic polynomial. The characteristic polynomial of the first one is x^2-4x-3 while the characteristic polynomial of the second is x^2-5x+5.

5.IV.1.32 Suppose that $m(x)=x^n+m_{n-1}x^{n-1}+\cdots+m_1x+m_0$ is minimal for T.

(a) For the 'if' argument, because $T^n+\cdots+m_1T+m_0I$ is the zero matrix we have that $I=(T^n+\cdots+m_1T)/(-m_0)=T\cdot(T^{n-1}+\cdots+m_1I)/(-m_0)$ and so the matrix $(-1/m_0)\cdot(T^{n-1}+\cdots+m_1I)$ is the inverse of T. For 'only if', suppose that $m_0=0$ (we put the $n=1$ case aside but it is easy) so that $T^n+\cdots+m_1T=(T^{n-1}+\cdots+m_1I)T$ is the zero matrix. Note that $T^{n-1}+\cdots+m_1I$ is not the zero matrix because the degree of the minimal polynomial is n. If T^{-1} exists then multiplying both $(T^{n-1}+\cdots+m_1I)T$ and the zero matrix from the right by T^{-1} gives a contradiction.

(b) If T is not invertible then the constant term in its minimal polynomial is zero. Thus,

$$T^n+\cdots+m_1T=(T^{n-1}+\cdots+m_1I)T=T(T^{n-1}+\cdots+m_1I)$$

is the zero matrix.

. . . .

Answers for subsection 5.IV.2

5.IV.2.17 We are required to check that

$$\begin{pmatrix}3&0\\1&3\end{pmatrix}=N+3I=PTP^{-1}=\begin{pmatrix}1/2&1/2\\-1/4&1/4\end{pmatrix}\begin{pmatrix}2&-1\\1&4\end{pmatrix}\begin{pmatrix}1&-2\\1&2\end{pmatrix}$$

That calculation is easy.

5.IV.2.18

(a) The characteristic polynomial is $c(x)=(x-3)^2$ and the minimal polynomial is the same.
(b) The characteristic polynomial is $c(x)=(x+1)^2$. The minimal polynomial is $m(x)=x+1$.
(c) The characteristic polynomial is $c(x)=(x+(1/2))(x-2)^2$ and the minimal polynomial is the same.
(d) The characteristic polynomial is $c(x)=(x-3)^3$ The minimal polynomial is the same.
(e) The characteristic polynomial is $c(x)=(x-3)^4$. The minimal polynomial is $m(x)=(x-3)^2$.
(f) The characteristic polynomial is $c(x)=(x+4)^2(x-4)^2$ and the minimal polynomial is the same.
(g) The characteristic polynomial is $c(x)=(x-2)^2(x-3)(x-5)$ and the minimal polynomial is $m(x)=(x-2)(x-3)(x-5)$.
(h) The characteristic polynomial is $c(x)=(x-2)^2(x-3)(x-5)$ and the minimal polynomial is the same.

5.IV.2.20 For each, because many choices of basis are possible, many other answers are possible. Of course, the calculation to check if an answer gives that PTP^{-1} is in Jordan form is the arbiter of what's correct.

(a) Here is the arrow diagram.

$$\begin{tikzcd}
\mathbb{C}^3_{\text{w.r.t. }\mathcal{E}_3} \arrow[r, "t", "T"'] \arrow[d, "\text{id}"', "P"] & \mathbb{C}^3_{\text{w.r.t. }\mathcal{E}_3} \arrow[d, "\text{id}"', "P"] \\
\mathbb{C}^3_{\text{w.r.t. }B} \arrow[r, "t", "J"'] & \mathbb{C}^3_{\text{w.r.t. }B}
\end{tikzcd}$$

The matrix to move from the lower left to the upper left is this.

$$P^{-1}=\big(\text{Rep}_{\mathcal{E}_3,B}(\text{id})\big)^{-1}=\text{Rep}_{B,\mathcal{E}_3}(\text{id})=\begin{pmatrix}1&-2&0\\1&0&1\\-2&0&0\end{pmatrix}$$

The matrix P to move from the upper right to the lower right is the inverse of P^{-1}.

(b) We want this matrix and its inverse.

$$P^{-1}=\begin{pmatrix}1&0&3\\0&1&4\\0&-2&0\end{pmatrix}$$

(c) The concatenation of these bases for the generalized null spaces will do for the basis for the entire space.

$$B_{-1}=\langle\begin{pmatrix}-1\\0\\0\\1\\0\end{pmatrix},\begin{pmatrix}-1\\0\\-1\\0\\1\end{pmatrix}\rangle\qquad B_3=\langle\begin{pmatrix}1\\1\\-1\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\0\\-2\\2\end{pmatrix},\begin{pmatrix}-1\\-1\\1\\2\\0\end{pmatrix}\rangle$$

The change of basis matrices are this one and its inverse.

$$P^{-1}=\begin{pmatrix}-1&-1&1&0&-1\\0&0&1&0&-1\\0&-1&-1&0&1\\1&0&0&-2&2\\0&1&0&2&0\end{pmatrix}$$

5.IV.2.23 The restriction of $t+2$ to $\mathscr{N}_\infty(t+2)$ can have only the action $\vec{\beta}_1 \mapsto \vec{0}$. The restriction of $t-1$ to $\mathscr{N}_\infty(t-1)$ could have any of these three actions on an associated string basis.

$$\begin{array}{lll} \vec{\beta}_2 \mapsto \vec{\beta}_3 \mapsto \vec{\beta}_4 \mapsto \vec{0} \qquad & \vec{\beta}_2 \mapsto \vec{\beta}_3 \mapsto \vec{0} \qquad & \vec{\beta}_2 \mapsto \vec{0} \\ & \vec{\beta}_4 \mapsto \vec{0} & \vec{\beta}_3 \mapsto \vec{0} \\ & & \vec{\beta}_4 \mapsto \vec{0} \end{array}$$

Taken together there are three possible Jordan forms, the one arising from the first action by $t-1$ (along with the only action from $t+2$), the one arising from the second action, and the one arising from the third action.

$$\begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

5.IV.2.25 There are two possible Jordan forms. The action of $t+1$ on a string basis for $\mathscr{N}_\infty(t+1)$ must be $\vec{\beta}_1 \mapsto \vec{0}$. There are two actions for $t-2$ on a string basis for $\mathscr{N}_\infty(t-2)$ that are possible with this characteristic polynomial and minimal polynomial.

$$\begin{array}{ll} \vec{\beta}_2 \mapsto \vec{\beta}_3 \mapsto \vec{0} \qquad & \vec{\beta}_2 \mapsto \vec{\beta}_3 \mapsto \vec{0} \\ \vec{\beta}_4 \mapsto \vec{\beta}_5 \mapsto \vec{0} & \vec{\beta}_4 \mapsto \vec{0} \\ & \vec{\beta}_5 \mapsto \vec{0} \end{array}$$

The resulting Jordan form matrics are these.

$$\begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

5.IV.2.29 Its characteristic polynomial is $c(x) = x^2 + 1$ which has complex roots $x^2 + 1 = (x+i)(x-i)$. Because the roots are distinct, the matrix is diagonalizable and its Jordan form is that diagonal matrix.

$$\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

To find an associated basis we compute the null spaces.

$$\mathscr{N}(t+i) = \{\begin{pmatrix} -iy \\ y \end{pmatrix} \bigm| y \in \mathbb{C}\} \qquad \mathscr{N}(t-i) = \{\begin{pmatrix} iy \\ y \end{pmatrix} \bigm| y \in \mathbb{C}\}$$

For instance,

$$T + i \cdot I = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}$$

and so we get a description of the null space of $t+i$ by solving this linear system.

$$\begin{array}{rl} ix - y = 0 \\ x + iy = 0 \end{array} \xrightarrow{i\rho_1 + \rho_2} \begin{array}{rl} ix - y = 0 \\ 0 = 0 \end{array}$$

(To change the relation $ix = y$ so that the leading variable x is expressed in terms of the free variable y, we can multiply both sides by $-i$.)

As a result, one such basis is this.

$$B = \langle \begin{pmatrix} -i \\ 1 \end{pmatrix}, \begin{pmatrix} i \\ 1 \end{pmatrix} \rangle$$

5.IV.2.30 We can count the possible classes by counting the possible canonical representatives, that is, the possible Jordan form matrices. The characteristic polynomial must be either $c_1(x) = (x+3)^2(x-4)$ or $c_2(x) = (x+3)(x-4)^2$. In the c_1 case there are two possible actions of $t+3$ on a string basis for $\mathscr{N}_\infty(t+3)$.

$$\begin{array}{ll} \vec{\beta}_1 \mapsto \vec{\beta}_2 \mapsto \vec{0} \qquad & \vec{\beta}_1 \mapsto \vec{0} \\ & \vec{\beta}_2 \mapsto \vec{0} \end{array}$$

There are two associated Jordan form matrices.

$$\begin{pmatrix} -3 & 0 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \qquad \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Similarly there are two Jordan form matrices that could arise out of c_2.

$$\begin{pmatrix} -3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 1 & 4 \end{pmatrix} \qquad \begin{pmatrix} -3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

So in total there are four possible Jordan forms.

5.IV.2.32 One example is the transformation of $\mathbb{C}$ that sends x to $-x$.

5.IV.2.33 Apply Lemma 2.7 twice; the subspace is $t-\lambda_1$ invariant if and only if it is t invariant, which in turn holds if and only if it is $t-\lambda_2$ invariant.

5.IV.2.34 False; these two 4×4 matrices each have $c(x)=(x-3)^4$ and $m(x)=(x-3)^2$.

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix} \qquad \begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

5.IV.2.35

(a) The characteristic polynomial is this.

$$\begin{vmatrix} a-x & b \\ c & d-x \end{vmatrix} = (a-x)(d-x)-bc = ad-(a+d)x+x^2-bc = x^2-(a+d)x+(ad-bc)$$

Note that the determinant appears as the constant term.

(b) Recall that the characteristic polynomial $|T-xI|$ is invariant under similarity. Use the permutation expansion formula to show that the trace is the negative of the coefficient of x^{n-1}.

(c) No, there are matrices T and S that are equivalent $S=PTQ$ (for some nonsingular P and Q) but that have different traces. An easy example is this.

$$PTQ = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Even easier examples using 1×1 matrices are possible.

(d) Put the matrix in Jordan form. By the first item, the trace is unchanged.

(e) The first part is easy; use the third item. The converse does not hold: this matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

has a trace of zero but is not nilpotent.

5.IV.2.36 Suppose that B_M is a basis for a subspace M of some vector space. Implication one way is clear; if M is t invariant then in particular, if $\vec{m}\in B_M$ then $t(\vec{m})\in M$. For the other implication, let $B_M=\langle\vec{\beta}_1,\dots,\vec{\beta}_q\rangle$ and note that $t(\vec{m})=t(m_1\vec{\beta}_1+\cdots+m_q\vec{\beta}_q)=m_1t(\vec{\beta}_1)+\cdots+m_qt(\vec{\beta}_q)$ is in M as any subspace is closed under linear combinations.

5.IV.2.38 One such ordering is the *dictionary ordering*. Order by the real component first, then by the coefficient of i. For instance, $3+2i<4+1i$ but $4+1i<4+2i$.

5.IV.2.39 The first half is easy—the derivative of any real polynomial is a real polynomial of lower degree. The answer to the second half is 'no'; any complement of $\mathcal{P}_j(\mathbb{R})$ must include a polynomial of degree $j+1$, and the derivative of that polynomial is in $\mathcal{P}_j(\mathbb{R})$.

5.IV.2.40 For the first half, show that each is a subspace and then observe that any polynomial can be uniquely written as the sum of even-powered and odd-powered terms (the zero polynomial is both). The answer to the second half is 'no': x^2 is even while $2x$ is odd.

5.IV.2.41 Yes. If $\mathrm{Rep}_{B,B}(t)$ has the given block form, take B_M to be the first j vectors of B, where J is the $j\times j$ upper left submatrix. Take B_N to be the remaining k vectors in B. Let M and N be the spans of B_M and B_N. Clearly M and N are complementary. To see M is invariant (N works the same way), represent any $\vec{m} \in M$ with respect to B, note the last k components are zeroes, and multiply by the given block matrix. The final k components of the result are zeroes, so that result is again in M.

5.IV.2.42 Put the matrix in Jordan form. By non-singularity, there are no zero eigenvalues on the diagonal. Ape this example:

$$\begin{pmatrix} 9 & 0 & 0 \\ 1 & 9 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 1/6 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}^2$$

to construct a square root. Show that it holds up under similarity: if $S^2 = T$ then $(PSP^{-1})(PSP^{-1}) = PTP^{-1}$.

. . . .

Answers for Topic: Computing Eigenvalues—the Method of Powers

1
- **(a)** The largest eigenvalue is 4.
- **(b)** The largest eigenvalue is 2.

3
- **(a)** The largest eigenvalue is 3.
- **(b)** The largest eigenvalue is -3.

5 In theory, this method would produce λ_2. In practice, however, rounding errors in the computation introduce components in the direction of $\vec{v}_1$, and so the method will still produce λ_1, although it may take somewhat longer than it would have taken with a more fortunate choice of initial vector.

6 Instead of using $\vec{v}_k = T\vec{v}_{k-1}$, use $T^{-1}\vec{v}_k = \vec{v}_{k-1}$.

. . . .

Answers for Topic: Stable Populations

. . . .

Answers for Topic: Linear Recurrences