ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

11-2001-114

На правах рукописи
УДК 519.612

ДУШАНОВ
Эрмухаммад Бердимурадович

АЛГОРИТМЫ И МАШИННО-НЕЗАВИСИМЫЙ ПАКЕТ ПРОГРАММ JINRLINPACK РЕШЕНИЯ ПЛОХО ОБУСЛОВЛЕННЫХ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

Специальность: 01.01.07 — вычислительная математика

Автореферат диссертации на соискание ученой степени
кандидата физико-математических наук

Дубна 2001

Работа выполнена в Лаборатории информационных технологий (ЛИТ) Объединенного института ядерных исследований.

Научные руководители:	доктор физико-математических наук профессор Емельяненко Геннадий Андреевич, кандидат физико-математических наук Рахмонов Турдимухаммад Тухтаматович
Официальные оппоненты:	доктор физико-математических наук профессор Тыртышников Евгений Евгеньевич, доктор физико-математических наук профессор Гердт Владимир Петрович
Ведущая организация:	Российский университет дружбы народов

Защита состоится "____" ___________ 2001 г. в ____ часов на заседании диссертационного совета Д720.001.04 в ОИЯИ по адресу: г.Дубна Московской области, ОИЯИ, ЛИТ.

С диссертацией можно ознакомиться в библиотеке ОИЯИ.

Афтореферат разослан "____" ___________ 2001 г.

Ученый секретарь
диссертационного совета З.М.Иванченко

1. Общая характеристика работы

Целью настоящей *диссертационной работы является:*

— разработка эффективного численного *непараметрического* метода решения плохо обусловленных систем линейных алгебраических уравнений вида

$$AZ \doteq F, \tag{1}$$

где $A = (a_{ij})$ — квадратная порядка m или прямоугольная размерности $m \times n$ вещественная матрица общего вида, $i = 1, ..., m, j = 1, ..., n, Z = (z_1, ..., z_n)^T$ — искомый и $F = (f_1, ..., f_m)^T$ — заданный n и m-мерные векторы соответственно;

— разработка эффективных универсальных алгоритмов и программ автоматического определения констант вещественной арифметики ЭВМ;

— создание на основе разработанных алгоритмов эффективного пакета `JINRLINPACK` стандартных программ на Фортране для ЭВМ любого типа для решения систем линейных алгебраических уравнений.

Актуальность работы. Поставленная задача решается достаточно просто, если матрица системы хорошо обусловлена и, следовательно, решение системы устойчиво относительно погрешностей. В настоящее время проблеме решения систем линейных алгебраических уравнений и исследованию свойств матриц системы, а также разработке и описанию пакетов программ уделено большое место в фундаментальных монографиях, обзорах, справочниках и многочисленных оригинальных публикациях. Что же касается алгоритмов автоматического получения констант вещественной арифметики ЭВМ, то этой проблеме уделено меньшее внимание в литературе.

С учетом сказанного интерес к поставленным в диссертационной работе задачам по-прежнему велик, что обусловлено также следующими обстоятельствами:

— *во-первых,* проблема решения систем линейных алгебраических уравнений (1) занимает одно из центральных мест в вычислительной алгебре. При численном решении этой задачи точными по сути методами имеется несколько источников появления погрешности решения. Один из них обусловлен округлением вещественных чисел в процессе вычисления на ЭВМ. Второй — неточностью задания самих исходных данных задачи (1). Третий — ошибками собственно метода, применяемого для решении задач (1). При этом к численному методу предъявляются жесткие требования к устойчивости получаемого им решения относительно указанных погрешностей.

Численное решение систем (1) с плохо обусловленными матрицами сводится, как известно, к проблеме устойчивого решения редуцированных систем вида

$$C_3 X = Y, \tag{2}$$
$$C_2 \hat{X} = \hat{Y}, \tag{3}$$

где C_3 и C_2 соответственно трехдиагональная и двухдиагональная матрицы

$$C_3 = \begin{bmatrix} q_1 & r_2 & & 0 \\ p_2 & q_2 & \ddots & \\ & \ddots & \ddots & r_m \\ 0 & & p_m & q_m \end{bmatrix}, \quad C_2 = \begin{bmatrix} q_1 & r_2 & & 0 \\ & q_2 & \ddots & \\ & & \ddots & r_m \\ 0 & & & q_m \end{bmatrix}, \tag{4}$$

а $X = (x_1, x_2, ..., x_m)^T$ и $\hat{X} = (\hat{x}_1, \hat{x}_2, ..., \hat{x}_m)^T$ — искомые, $Y = (y_1, y_2, ..., y_m)^T$ и $\hat{Y} = (\hat{y}_1, \hat{y}_2, ..., \hat{y}_m)^T$ — m-мерные векторы, $\{q_i\}_{i=1}^m$ — диагональные и $\{p_i, r_i\}_{i=2}^m$ — под(над)диагональные элементы матриц C_3 и C_2.

— *Во-вторых,* трехдиагональные матрицы играют и самостоятельную (выделенную) роль при решении многих задач вычислительной математики и математической физики. В самом деле, численное решение краевых задач математической физики сводится в общем случае к решению алгебраических задач с трехдиагональными, ленточными и блочно-трехдиагональными матрицами. Методам решения систем линейных алгебраических уравнений с трехдиагональными матрицами посвящена обширная литература. Эти методы в значительной степени связаны с построением экономичных алгоритмов для характерных типов систем, возникающих при решении задач математической физики. Основная задача при их разработке заключается или в минимизации числа арифметических действий с учетом специфики уравнений, или в обеспечении надежности, т.е. устойчивости к накоплению погрешностей округления при реализации на ЭВМ.

— *В-третьих,* практическое использование существующих пакетов программ и сравнительный анализ результатов расчетов на их основе показывают [1], что в случае плохой обусловленности (особенно патологически плохой обусловленности $\text{cond}(A) > 1/\varepsilon_1$, где ε_1 — относительная погрешность вычислений с вещественными числами данной ЭВМ) поставленная задача разработки эффективного пакета программ еще не решена окончательно. Не решена, по-прежнему, также проблема разработки универсального (инвариантного относительно различных типов ЭВМ) пакета программ.

Научная новизна.

1. Разработан эффективный метод критических компонент решения плохо обусловленных (особенно при $\text{cond}(A) > 1/\varepsilon_1$) систем линейных алгебраических уравнений.
2. Разработан на основе метода критических компонент алгоритм решения плохо обусловленных систем линейных алгебраических уравнений.
3. Разработан алгоритм автоматического определения констант машинной арифметики ЭВМ, обеспечивающий *универсальность* программ на различных типах ЭВМ.

Практическая значимость настоящей диссертационной работы определяется созданным [2] на основе новых разработанных алгоритмов машинно-независимым пакетом `JINRLINPACK`, включающим в себя программные модули `INIT_CONST`, `CHEXP`, `GTEXP`, `LinsysccmSolver`, `Lin2dsysccmSolver`, `Lin3dsysccmSolver`, `PseudsysccmSolver` на Фортране 77. Многочисленные эксперименты показали, что новые программы обладают в среднем лучшими основными показателями, чем подобные программы из наиболее известных пакетов:

`CERNLIB` — библиотека программ CERN;
`NAG` — пакет математических программ (Numerical Algorithms Group, Oxford);
`LIBJINR` — библиотека программ ОИЯИ;
`LINA` — пакет программ.

Новый машинно-независимый пакет программ `JINRLINPACK` тестирован и поставлен [3] в библиотеку `LIBJINR` базовых ЭВМ ЛИТ ОИЯИ. Ему присвоен индекс `F499` по классификации CERN. Фортранный текст всех программ пакета `JINRLINPACK` находится в файле `f499.f` на сервере Convex и доступен через WWW по адресу `http://www.jinr.ru/~tsap/Koi/jinrlib`.

Высокая эффективность программ пакета JINRLINPACK подтверждена [11] при численном решении плохо обусловленных систем линейных алгебраических уравнений в задаче рассеяния с вариационным функционалом Швингера.

Апробация работы. Основные результаты диссертации докладывались на научных семинарах: по методам вычислительной и прикладной математики ЛВТА (ЛИТ), по применению вычислительной техники в научных исследованиях ЛВТА (ЛИТ) ОИЯИ, а также на международных конференциях: "Численное моделирование и вычисления в физике" (СМСР–96, Дубна, 16–21 сентября 1996 г.), "Актуальные проблемы вычислительной физики" (МТСР-98, Дубна, 15–20 июня 1998 г.), "Актуальные проблемы вычислительной физики" (МТСР–2000, Дубна, 24–30 июля 2000 г.), III научная конференция молодых ученых и специалистов (Дубна, 15–19 февраля 1999 г.), на семинаре Института вычислительной математики РАН (Москва, 17 апреля 2001 г.).

Публикации. По теме диссертации опубликовано 10 работ [1–10]. Из них в диссертацию вошли лишь некоторые из результатов работ [1–10].

Структура и построение диссертации. Диссертация состоит из 3 глав, 33 рисунков и 30 таблиц.

В диссертации принята следующая трехзначная нумерация (а.б.в) формул, лемм, теорем, замечаний, следствий и представлений, где а — номер главы, б — номер параграфа, в — номер соответствующих формул, лемм, теорем, замечаний, следствий и представлений в этом параграфе. Например, теорема 3.2.1 есть первая теорема второго параграфа третьей главы диссертации.

Личный вклад автора. Автор диссертации, работая в коллективе соавторов, объединяющих сотрудников ЛВТА (ЛИТ) ОИЯИ, ИЯФ АН РУ (г. Ташкент), ВМК МГУ (г. Москва), университет "Дубна" внес большой вклад в *разработку алгоритмов* на основе метода критических компонент для решения плохо обусловленных систем линейных алгебраических уравнений, в разработку алгоритмов получения констант вещественной арифметики ЭВМ, а также в создание основных модулей нового машинно–независимого пакета JINRLINPACK.

2. Содержание диссертации

Во введении обсуждается содержание предмета исследований, обосновывается их актуальность и приводится краткий анализ состояния проблемы. Сформулирована также цель, научная новизна и практическая значимость работы. Указаны основные результаты, выносимые на защиту.

Первая глава, состоящая из трех параграфов, посвящена описанию разработанных в [1; 4 – 6] численных методов решения плохо обусловленных систем линейных алгебраических уравнений. Приводятся также результаты исследования структуры представлений трех(двух)диагональных и им обратных матриц. Для решения плохо обусловленных систем с трех(двух)диагональными матрицами применен метод оптимального анализа структуры и построения обратных матриц к трех(двух)диагональным. В основе приводимых в Главе 1 диссертации численных методов лежат результаты оптимального вычисления ведущих верхних (нижних) угловых миноров трехдиагональных матриц.

В параграфе 1.1 приведена теорема 1.1.1, доказанная в [1]. В ней установлено, что для внедиагональных элементов B_{ij} обратной матрицы $B = C_3^{-1}$, где C_3 — невырожденная трехдиагональная матрица вида (4), некоторые ведущие угловые миноры которой обращаются в нуль (т.е. $\Delta_1^i = 0$ и/или $\Delta_j^m = 0$, для любых i из $1 \le i < m$, и j из $1 < j \le m$), имеют место следующие представления:

$$B_{ij} = \begin{cases} \omega_i \prod\limits_{\xi=j+1}^{i} \beta_\xi, \text{ если } 1 \le j < i, 1 \le i \le m, \\ 0 \text{ для всех } i \text{ из } j < i \le m, \text{ если } \Lambda_j = 0, \\ 0 \text{ для всех } j \text{ из } 1 \le j < i, \text{ если } \mathrm{G}_i = 0, \\ \omega_i \prod\limits_{\xi=i+1}^{j} \hat{\beta}_\xi, \text{ если } 1 \le i < j \le m, \\ 0 \text{ для всех } i \text{ из } 1 \le i < j, \text{ если } \mathrm{G}_j = 0, \\ 0 \text{ для всех } j \text{ из } i < j \le m, \text{ если } \Lambda_i = 0. \end{cases} \tag{5}$$

$$B_{ij} = \begin{cases} \prod\limits_{\xi=j+1}^{i} \hat{c}_\xi \omega_j, \text{ если } 1 \le j < i \le m, \\ 0 \text{ для всех } j \text{ из } 1 \le j < i, \text{ если } G_i = 0, \\ 0 \text{ для всех } i \text{ из } j < i \le m, \text{ если } \Lambda_j = 0, \\ \prod\limits_{\xi=i+1}^{j} c_\xi \omega_j, \text{ если } 1 \le i < j \le m, \\ 0 \text{ для всех } j \text{ из } i < j \le m, \text{ если } \Lambda_i = 0, \\ 0 \text{ для всех } i \text{ из } 1 \le i < j, \text{ если } G_j = 0. \end{cases} \quad (6)$$

Диагональные же элементы B_{ii} матрицы B, а также величины ω_i в (5)–(6) представимы в виде:

$$\begin{cases} B_{11} = \begin{cases} G_0^{-1}, \text{ если } G_1 \ne 0, G_0 \ne 0 \text{ и при этом } \omega_1 = B_{11}, \\ 0, \text{ если } G_1 = 0 \quad \text{ и при этом } \omega_1 = (-r_2 p_2)^{-1}; \end{cases} \\ B_{ii} = \begin{cases} (\Lambda_{i+1} + G_{i-1} - q_i)^{-1}, \text{ если } \Lambda_i \ne 0 \ne G_i \text{ и при этом } \omega_i = B_{ii}, \\ 0, \text{ если } \Lambda_i = 0 \quad \text{ и при этом} \\ \quad B_{i-1i-1} = G_{i-1}\omega_i, B_{i+1i+1} = G_i^{-1}, \omega_i = (-p_i r_i)^{-1}, \\ 0, \text{ если } G_i = 0 \quad \text{ и при этом} \\ \quad B_{i-1i-1} = \Lambda_i^{-1}, B_{i+1i+1} = \Lambda_{i+1}\omega_i, \omega_i = (-r_{i+1}p_{i+1})^{-1}, \\ i = 2, 3, ..., m-1; \end{cases} \\ B_{mm} = \begin{cases} \Lambda_{m+1}^{-1}, \text{ если } \Lambda_m \ne 0, \Lambda_{m+1} \ne 0 \text{ и при этом } \omega_m = B_{mm}, \\ 0, \text{ если } \Lambda_m = 0 \quad \text{ и при этом } \omega_m = (-p_m r_m)^{-1}. \end{cases} \end{cases} \quad (7)$$

В (5)–(7) элементы Λ_i и G_i определяются [1] как

$$\begin{cases} \Lambda_{i+1} = q_i - p_i \Lambda_i^{-1} r_i, \ \Lambda_2 = q_1, \ i = 2, ..., m, \quad \text{если } \Lambda_i \ne 0 \\ \text{для всех } 2 \le i \le m. \\ \text{Если } \Lambda_i = 0 \text{ для любого } i \text{ из } (2 \le i \le m), \text{ то} \\ \Lambda_{i+1} \text{— неопределено, но } \Lambda_{i+2} = q_{i+1}, \Lambda_i \Lambda_{i+1} = -p_i r_i; \end{cases} \quad (8)$$

$$\begin{cases} G_{i-1} = q_i - r_{i+1} G_i^{-1} p_{i+1}, \ G_{m-1} = q_m, \ i = m-1, ..., 1, \\ \text{если } G_i \ne 0 \text{ для всех } 1 \le i \le m-1. \\ \text{Если } G_i = 0 \text{ для любого } i \text{ из } (1 \le i \le m-1), \text{ то} \\ G_{i-1} \text{— неопределено, но } G_{i-2} = q_{i-1}, G_i G_{i-1} = -r_{i+1} p_{i+1}. \end{cases} \quad (9)$$

При этом в (5)–(6) структурные элементы $\beta, \hat{\beta}, c, \hat{c}$ и произведения $\prod \beta_\xi, \prod \hat{\beta}_\xi, \prod c_\xi, \prod \hat{c}_\xi$ имеют вид:

$$\beta_i = \begin{cases} -p_i \Lambda_i^{-1}, \text{ если } \Lambda_i \ne 0, \\ -p_i, \text{ если } \Lambda_i = 0, \text{ при} \\ \text{этом } \beta_{i+1} = -p_{i+1}\omega_i; \end{cases} \quad c_i = \begin{cases} -\Lambda_i^{-1} r_i, \text{ если } \Lambda_i \ne 0, \\ -r_i, \text{ если } \Lambda_i = 0, \text{ при} \\ \text{этом } c_{i+1} = -\omega_i r_{i+1}; \end{cases} \quad (10)$$

$$\hat{c}_{i+1} = \begin{cases} -G_i^{-1} p_{i+1}, \text{ если } G_i \ne 0, \\ -p_{i+1}, \text{ если } G_i = 0, \\ \text{при этом } \hat{c}_i = -\omega_i p_i; \end{cases} \quad \hat{\beta}_{i+1} = \begin{cases} -r_{i+1} G_i^{-1}, \text{ если } G_i \ne 0, \\ -r_{i+1}, \text{ если } G_i = 0, \\ \text{при этом } \hat{\beta}_i = -r_i \omega_i; \end{cases} \quad (11)$$

$$\prod_{\xi=j+1}^{i} \beta_\xi = \begin{cases} \beta_i \cdots \beta_{j+1}, \text{ если } j < i, \\ 1, \quad \text{ если } j \ge i; \end{cases} \quad \prod_{\xi=i+1}^{j} c_\xi = \begin{cases} c_{i+1} \cdots c_j, \text{ если } i < j, \\ 1, \quad \text{ если } i \ge j; \end{cases} \quad (12)$$

$$\prod_{\xi=i+1}^{j} \hat{\beta}_\xi = \begin{cases} \hat{\beta}_{i+1} \cdots \hat{\beta}_j, \text{ если } i < j, \\ 1, \quad \text{ если } i \ge j; \end{cases} \quad \prod_{\xi=j+1}^{i} \hat{c}_\xi = \begin{cases} \hat{c}_i \cdots \hat{c}_{j+1}, \text{ если } j < i, \\ 1, \quad \text{ если } j \ge i. \end{cases} \quad (13)$$

Представление элементов B_{ij} обратной матрицы $B = C_3^{-1}$, имеет место как для всей матрицы C_3 (4), так и для любых ее подматриц C_ρ^ν того же вида.

В параграфе 1.2 изложены результаты исследований, связанных с обоснованием эффективности предложенного в [1;4–6] прямого метода критических компонент для решения *плохо обусловленных* систем линейных алгебраических уравнений (1). Показано, что для систем типа (1) устойчивый к погрешностям $(h, \delta, \varepsilon_1, \varepsilon_0)$ метод критических компонент позволяет численно находить единственное нормальное минимальное по норме решение, обеспечивающее минимум нормы невязки

$$(Z^+ = A^+F): \ ||AZ^+ - F|| = \inf_{Z \in Z_A} ||AZ - F||, ||Z^+|| = \inf_{Z \in Z_A} ||Z||,$$

где Z_A — совокупность всех решений системы (1), а также единственную матрицу A^+, удовлетворяющую условиям

$$||A^+A - E|| = \inf_{\mathring{A}^{-1} \in \Omega_A} ||\mathring{A}^{-1}A - E||, ||A^+|| = \inf_{\mathring{A}^{-1} \in \Omega_A} ||\mathring{A}^{-1}||, A^+A = AA^+,$$

где E — единичная матрица и Ω_A — совокупность всех $\mathring{A}^{-1}$, "псевдообратных" к A. При этом даже в случае существенно плохой обусловленности задачи (1) Z^+ и A^+ устойчивы к малым изменениям входных данных (A, F).

Приводимая в этом параграфе теорема 1.2.1 обосновывает [1] получение численно методом критических компонент минимального по норме ($||X^+|| = \min$), единственного решения X системы $C_3X = Y$, удовлетворяющего условию минимума нормы невязки ($||\tilde{C}_3X^+ - \tilde{Y}|| = \min$), устойчивого к погрешностям $(\varepsilon_1, \varepsilon_0)$ вычислений и к малым изменениям (h, δ) входных данных (C_3, Y).

Ниже приводится *алгоритм* метода критических компонент, следующий из теоремы.

Начало вычислений:

$k = 1, i = m;$

$(1)^0$ $l_k = i;$

$(2)^0$ $\overset{[k]}{x}_i = \sum\limits_{\xi=1}^{l_k} \overset{[k]}{B}_{i\xi}\, y_\xi,\ \overset{[k]}{\varphi}_i = \begin{cases} 0, \text{ если } k = 1, \\ -\overset{[k]}{B}_{il_k}\, r_{l_k+1} x^+_{l_k+1}, \text{ если } k > 1; \end{cases}$

если $i = l_k$, то $(5)^0$, иначе $(3)^0$;

$(3)^0$ если $|\overset{[k]}{\varphi}_i| < 1/\varepsilon_1$, то $(4)^0$, иначе $k = k + 1$ и $(1)^0$;

$(4)^0$ $j = i + 1, \overset{[k]}{x}_{l_k+1} = 0;$

$$\Phi_j = \begin{cases} |y_j| - |p_j \overset{[k]}{x}_{j-1} + q_j \overset{[k]}{x}_j + r_{j+1} \overset{[k]}{x}_{j+1}|, \text{ при } |y_j| \le 1, \\ 1 - |p_j \overset{[k]}{x}_{j-1} + q_j \overset{[k]}{x}_j + r_{j+1} \overset{[k]}{x}_{j+1}|/|y_j|, \text{ при } |y_j| > 1; \end{cases}$$

если $|\Phi_j| \le 2\varepsilon_1$, то $(5)^0$, иначе $k = k + 1$ и $(1)^0$;

$(5)^0$ $x_i^+ = \overset{[k]}{x}_i + \overset{[k]}{\varphi}_i;$

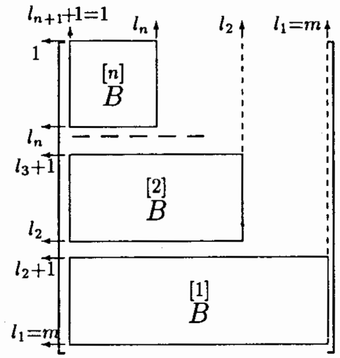
если $i = 1$, то конец вычислений, иначе $i = i - 1$ и $(2)^0$;

Конец вычислений.

Здесь:

$\overset{[k]}{B}_{ij}$ $(l_{k+1} \le i \le l_k,\ 1 \le j \le l_k$ и $k = 1, 2, ..., n)$ — элементы подматриц $\overset{[k]}{B}$ матрицы $\mathring{B} = \mathring{C}_3^{-1}$, обратной к хорошо обусловленной матрице $\mathring{C}_3$ вида:

$$\mathring{C}_3 = \left[\begin{array}{l}
\left[\begin{array}{cccc} q_1 & r_2 & & \\ p_2 & q_2 & r_3 & \\ & \cdots & & \\ & & p_{l_n} & q_{l_n} \end{array}\right] = \overset{[n]}{C}{}^{(l_{n+1}+1=1)}_{l_n} \quad 0 \\
\quad - - - - - - \\
\quad p_{l_3+1} \left[\begin{array}{cccc} \tilde{q}_{l_3+1} & r_{l_3+2} & & \\ p_{l_3+2} & q_{l_3+2} & r_{l_3+3} & \\ & \cdots\cdots & & \\ & & p_{l_2} & q_{l_2} \end{array}\right] = \overset{[2]}{C}{}^{l_3+1}_{l_2} \quad 0 \\
\qquad\qquad p_{l_2+1} \quad \overset{[1]}{C}{}^{l_2+1}_{(l_1=m)} = \left[\begin{array}{cccc} \tilde{q}_{l_2+1} & r_{l_2+2} & & \\ p_{l_2+2} & q_{l_2+2} & r_{l_2+3} & \\ & \cdots\cdots & & \\ & & p_m & q_m \end{array}\right]
\end{array}\right] \Rightarrow \qquad (14)$$



$$= \mathring{B}, \qquad (15)$$

где

$$\tilde{q}_{l_k+1} = q_{l_k+1} - p_{l_k+1} \overset{[k]}{B}_{l_k l_k} r_{l_k+1},\ k = 2, 3, .., n \qquad (16)$$

и $\overset{[k]}{B}_{l_k l_k}$ — последние диагональные элементы подматриц $\overset{[k]}{B}$, которые совпадают с последними диагональными элементами подматриц, обратных к хорошо обусловленным подматрицам $\overset{[k]}{C}{}^{l_{k+1}+1}_{l_k}$, выделяемых методом, n — число выделяемых подпространств. Элементы $\overset{[k]}{B}_{ij}$ вычисляются [1] по формулам (5) или (6).

В параграфе 1.3 рассматривается проблема итерационного уточнения решения, полученного методом критических компонент.

Вторая глава диссертации состоит из трех параграфов. В параграфе 2.1 изучаются форматы представления нормализованных чисел в памяти ЭВМ различного типа. В параграфе 2.2 приведен разработанный в [7] алгоритм вычисления единственных базисных констант (β, t, l, u и $\varepsilon_\infty, \varepsilon_0, \varepsilon_1$) вещественной арифметики ЭВМ, основанный на поразрядном исследовании формата представления вещественных чисел с плавающей точкой в памяти ЭВМ.

Пусть в формате, например, из 64 двоичных разрядов представлено нормализованное число $\gamma_1 = (1.)$. Копируем этот формат в другой формат τ такой же длины. Прежде чем описать дальнейшие шаги алгоритма отметим следующее. Если данная ЭВМ работает при основании $\beta = 2$, то известно, что содержимое старшего бита мантиссы может храниться и вне формата. Это обеспечивает увеличение длины мантиссы на один разряд. Поэтому при $\beta = 2$ в разрядах формата, отведенных для хранения мантиссы, в случае $\gamma_1 = (1.)$ будут все нули.

При $\beta > 2$ (и при $\beta = 2$, если единица мантиссы $\gamma_1 = (1.)$ хранится в формате) в разрядах, отведенных для хранения мантиссы числа $\gamma_1 = (1.)$ будут нули и лишь одна единица в старшем β-ичном разряде мантиссы. Эта единица является машинным представлением числа $1/\beta$. Для хранения любой значащей цифры от 0 до $\beta - 1$ (при основании β) будет отводиться n_β – двоичных разрядов формата. Следовательно мантисса числа $\gamma_1 = (1.)$ будет храниться в виде единицы в младшем из n_β – двоичных разрядов мантиссы формата τ.

Пусть мантисса m_1 и порядок e_1 числа $\gamma_1 = (1.)$ хранятся в прямом коде*) в формате τ в виде двух единиц в соответствующих разрядах. Все остальные разряды формата τ будут при этом нулевыми. Обозначим номера двоичных разрядов формата τ, в которых хранятся единицы, соответственно μ_i и μ_j (см. рис. 1).

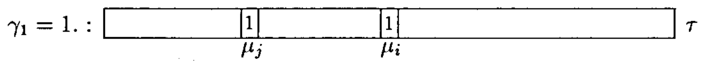


Рис. 1.

При этом дальнейшее описание алгоритма не зависит от того является ли $i > j$ или наоборот. Прибавим (1.)–единицу к числу γ_1, т.е. получим в формате τ число $\gamma_2 = (\gamma_1 + 1.)$. Тогда, если основание β было равно 2, то в формате τ, в котором хранится теперь число $\gamma_2 = 2.$, единица мантиссы останется в прежнем двоичном разряде μ_i, а единица порядка переместится в следующий старший разряд μ_{j+1}. При этом в разряде μ_j будет 0 (см. рис. 2).

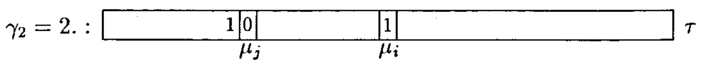


Рис. 2.

Это обусловлено тем, что при любом основании β в n_β – двоичных разрядах будет представима единицами $(\beta - 1)$ — максимальная значащая цифра. Поэтому число β — основание системы счисления будет уже двузначным и оно представляется в формате единицей в разрядах мантиссы и двойкой в разрядах порядка. Следовательно, если произошел указанный сдвиг единицы в разрядах порядка, то основание $\beta = 2$ найдено. Если указанный сдвиг порядка не произошел, то $\beta > 2$. Но при этом единица мантиссы перемещается в следующий разряд μ_{i+1} (см. рис. 3).

*)Рассуждения, подобные приводимым ниже, имеют место и в случае любого другого кода представления нормализованных чисел. Отметим также, что при любом основании β *базисными* (значащими) цифрами являются $0, 1, 2, \ldots, \beta - 1$.

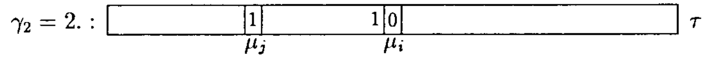


Рис. 3.

В разряде μ_i будет 0, а единица порядка останется в прежнем двоичном разряде μ_j. Далее прибавляем (1.)–единицу к числу $\gamma_2 = (2.)$, т.е. получим в формате τ число $\gamma_3 = (3.)$. Тогда, если основание β было равно 3, то в формате τ, в котором хранится теперь "двузначное" число $\gamma_3 = 3.$, единица мантиссы возвращается в разряд μ_i, разряд μ_{i+1} будет содержать 0, а единица порядка переместится в следующий старший разряд μ_{j+1} (см. рис. 4).

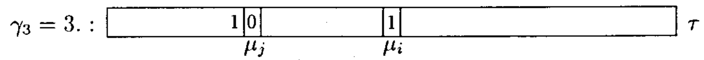


Рис. 4.

Если такой сдвиг порядка не произошел, то $\beta > 3$ и алгоритм поиска β продолжается. Итак, пусть $\beta > 3$. Тогда в соответствии с выше описанной процедурой заполнения двоичных n_β – разрядов работа алгоритма поиска β прекращается как только происходит перемещение единицы порядка из разряда μ_j в двоичный разряд μ_{j+1}. При этом в n_β – двоичных разрядах мантиссы должно было бы храниться (в двоичном коде) число β, которое, как отмечено уже выше, является двузначным при основании β. Поэтому, как и при основаниях $\beta = 2; 3$, в младший из n_β – двоичных разрядов переместится 1, а остальные $(n_\beta - 1)$ – старших из этих двоичных разрядов будут нулевыми. Таким образом, будет найдено число n_β – двоичных разрядов, в которых хранится в двоичном коде $(\beta - 1)$-максимальная из значащих цифр $0, 1, \ldots, \beta - 1$ при основании β в виде

$$\beta - 1 = (\mu_{i+n_\beta-1} \equiv 1) \cdot 2^{n_\beta - 1} + (\mu_{i+n_\beta-2}) \cdot 2^{n_\beta-2} + \cdots + (\mu_{i+1}) \cdot 2^1 + (\mu_i) \cdot 2^0,$$

где $(\mu_{i+n_\beta-1}), (\mu_{i+n_\beta-2}), (\mu_{i+1}), (\mu_i)$ — содержимое (1 или 0) указанных двоичных разрядов $\mu_{i+n_\beta-1}, \mu_{i+n_\beta-2}, \mu_{i+1}, \mu_i$. При этом $\gamma_{k-1} = \beta - 1$ и $\gamma_k = \beta$ хранятся в формате τ в виде (см. рис. 5).

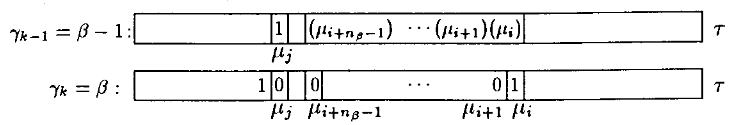


Рис. 5.

Итак, основание β и n_β-двоичных разрядов, в которых хранятся значащие цифры $0, 1, 2, ..., \beta - 1$, найдены.

В параграфе 2.3 разработаны алгоритмы для численного обращения матриц и решения систем линейных алгебраических уравнений с плохо обусловленными матрицами на основе метода критических компонент. Приведены блок-схемы программ.

В третьей главе приводится описание машинно-независимого пакета программ `JINRLINPACK` [2] на Фортране, в котором реализованы изложенные в Главе 2 алгоритмы и который поставлен [3] на базовых ЭВМ ОИЯИ. В пакет входят программы:

`INIT_CONST` — получения констант вещественной арифметики ЭВМ;
`Lin2dsysccmSolver` — решения систем линейных уравнений $C_2X = Y$;
`Lin3dsysccmSolver` — решения систем линейных уравнений $C_3X = Y$;
`LinsysccmSolver`, `PseudsysccmSolver` — решения систем линейных уравнений $AX = Y$.

В параграфах 3.1 и 3.2 приведены также некоторые результаты тестовых расчетов и примеры использования программ. В параграфе 3.3 приводятся результаты численных экспериментов в сравнении с результатами наиболее известных программ из пакетов:

`CERNLIB` — библиотека программ CERN;
`NAG` — пакет математических программ (Numerical Algorithms Group, Oxford);
`LIBJINR` — библиотека программ ОИЯИ;
`LINA` — пакет программ.

Приведен также краткий анализ полученных численных результатов, из которого следует, что метод критических компонент является лучшим по своим качественным показателям из наиболее известных методов решения плохо обусловленных систем линейных алгебраических уравнений.

В параграфе 3.4 приводятся тексты программ машинно – независимого пакета `JINRLINPACK` на Фортране–77.

3. Основные результаты и выводы

1. Получены новые типы мультипликативных и аддитивно – мультипликативных представлений для трехдиагональных и им обратных матриц общего вида, некоторые угловые миноры которых обращаются в нуль (или “машинный” нуль).

2. Разработан новый эффективный метод решения систем линейных алгебраических уравнений с трехдиагональными и двухдиагональными матрицами получивший название [1;4–6] метода критических компонент.

3. Изучены форматы представления нормализованных чисел в памяти ЭВМ различного типа и получен алгоритм вычисления единственных базисных констант $(\beta, t, l, u$ и $\varepsilon_\infty, \varepsilon_0, \varepsilon_1)$ вещественной арифметики ЭВМ, основанный на поразрядном исследовании форматов представления вещественных чисел с плавающей точкой в памяти ЭВМ. Разработаны алгоритмы для численного обращения матриц и решения систем линейных алгебраических уравнений с плохо обусловленными матрицами на основе метода критических компонент. Приведены блок-схемы программ, реализующих алгоритмы.

4. Приведено описание и тексты программ на Фортране машинно-независимого пакета `JINRLINPACK`, включающего в себя модули `INIT_CONST` (получение констант вещественной арифметики ЭВМ), `Lin2dsysccmSolver` (решение систем линейных уравнений $C_2X = Y$), `Lin3dsysccmSolver` (решение систем линейных уравнений $C_3X = Y$), `LinsysccmSolver`, `PseudsysccmSolver` (решение систем линейных уравнений $AX = Y$). В пакете реализованы алгоритмы Главы 1 и 2. Пакет поставлен [3] на базовых ЭВМ ОИЯИ. Приведены некоторые результаты тестовых расчетов, а также примеры использования программ.

5. Приведены результаты численных экспериментов по вычислению основных численных характеристик решений систем линейных алгебраических уравнений. Приведен также краткий анализ полученных численных результатов, из которого следует, что метод критических компонент является в среднем лучшим по своим качественным показателям из наиболее известных методов, реализованных в пакетах `NAG`, `CERNLIB`, `LIBJINR`, `LINA`.

4. Публикации по теме диссертации

[1] Emel'yanenko G.A., Emelianenko M.G., Rakhmonov T.T., Dushanov E.B., Konovalova G.Yu. JINR preprint, E11-98-302, Dubna, 1998.

[2] Емельяненко Г.А., Душанов Э.Б., Емельяненко М.Г., Рахмонов Т.Т., Сапожников А.П. Сообщение ОИЯИ, Р11-2000-287, Дубна, 2000.

[3] Новости ОИЯИ (JINR News). Информационный бюллетень ОИЯИ: 3, 1996, стр. 12; 1, 2001, стр. 7-8.

[4] Emel'yanenko G.A., Rakhmonov T.T., Dushanov E.B. JINR preprint, E11-96-105, Dubna, 1996.

[5] Emel'yanenko G.A., Rakhmonov T.T., Dushanov E.B. JINR preprint, E11-96-106, Dubna, 1996.

[6] Emel'yanenko G.A., Rakhmonov T.T., Dushanov E.B. JINR preprint, E11-96-107, Dubna, 1996.

[7] Душанов Э.Б., Емельяненко М.Г., Коновалова Г.Ю. Препринт ОИЯИ, Р11-2000-163, Дубна, 2000.

[8] Международная Конференция МТСР–98, Дубна, 15–20 июня 1998.

[9] Труды III научной Конференции молодых ученых и специалистов. Дубна, 15–19 февраля 1999 г.

[10] Труды Международной Конференции МТСР–2000, Дубна, 24–30 июля 2000 г.

. . . .

[11] Виницкий С.И., Пузынин И.В., Чулуунбаатар О. Препринт ОИЯИ, Р11-2001-61, Дубна, 2001.

Рукопись поступила в издательский отдел
30 мая 2001 года.