ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

P10-5279

Г.А. Емельяненко

ВЫБОР ПРЕДСТАВЛЕНИЯ ФУНКЦИИ ПРАВДОПОДОБИЯ ПРИ ОБРАБОТКЕ КАМЕРНЫХ СНИМКОВ

ЛАБОРАТОРИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ
И АВТОМАТИЗАЦИИ

1970

P10-5279

Г.А. Емельяненко

ВЫБОР ПРЕДСТАВЛЕНИЯ ФУНКЦИИ ПРАВДОПОДОБИЯ ПРИ ОБРАБОТКЕ КАМЕРНЫХ СНИМКОВ

Направлено в ПТЭ

Постановка задачи

При обработке снимков с пузырьковых камер приходится оценивать по элементам единственной случайной выборки – множеству из n независимо измеренных координат точек на треке оптимальные значения кинематических параметров (и их ошибок) частицы, оставившей след в эффективной части рабочего объема камеры.

Будем рассматривать всевозможные наборы кинематических параметров $\{\theta_1^j, \theta_2^j, \theta_3^j, \ldots, \theta_m^j\}$ как множества упорядоченных систем из m вещественных чисел. Каждый из указанных наборов определяет точку в параметрическом пространстве Θ.

Нетрудно показать[x/], что Θ является метрическим пространством, изометричным обычному m-мерному Евклидову пространству R_m. Поэтому в дальнейшем будем говорить о вещественном параметрическом пространстве R_m.

[x/] Определим понятие "близости" (метрику) в пространстве Θ, $\rho(\Theta^i, \Theta^j) = \sqrt{(\Theta^i - \Theta^j, \Theta^i - \Theta^j)}$ (*) и введем понятие неограниченного "сближения" (т.е. предела последовательности в Θ) как $\rho(\Theta^i, \Theta^0)_{i \to \infty} 0$ (**). Как следует из (**), сходимость в Θ является сходимостью по θ_k^i координатам, которая полностью определяется введенной метрикой (*).

$\Theta^j = (\theta_1^j, \theta_2^j, \theta_3^j, \dots, \theta_m^j) \subseteq R_m$ назовем j –гипотезой о частице, где $j \subseteq J$ – счётное множество неотрицательных целых чисел. Пусть L – есть информация о движении частицы, а $Y = (y_0, y_1, y_2, \dots, y_{n-1})$ – случайный вектор независимых измерений y –ных координат точек на треке.

Теперь можем сформулировать описанную выше статистическую задачу:

Найти такой элемент $\Theta^k \subseteq R_m$, чтобы условная вероятность $P(\Theta/Y,L)$ принимала в R_m максимальное значение, т.е.

$$P(\Theta^k / Y,L) \overset{k\subseteq J}{=} \max_{j\subseteq J} P(\Theta^j / Y,L). \qquad (1.0)$$

Здесь воспользовались тем обстоятельством, что элементы R_m можно рассматривать как непересекающиеся (одновременно по всей совокупности параметров θ_ℓ) события.

Определим основной подход к решению поставленной задачи. Пусть $\Delta Y = Y - \langle Y \rangle$, где $\langle Y \rangle$ – вектор средних значений для Y. Отметим, что $\langle Y \rangle = \langle Y(\Theta,L) \rangle$. Предположим, что ΔY распределен по закону $F(\eta, \Sigma)$, где η – вектор средних значений для ΔY, а Σ – ковариационная матрица распределения.

Если известно событие Θ^k, т.е. задана точная гипотеза $(\theta_1^k, \theta_2^k, \dots, \theta_m^k) \subseteq R_m$ для данного события Y, то можно вычислить $\langle Y \rangle^k$ и, следовательно, ΔY^k. Тогда вероятность получить ΔY, "близкие" к ΔY^k, запишется:

$$P(\Delta Y^k < \Delta Y < \Delta Y^k + \delta) = f(\Delta Y / \eta, \Sigma) \cdot \Delta,$$

где $f(\Delta Y / \eta, \Sigma)$ – значение плотности распределения вероятностей

$\eta \Rightarrow$ g. δ. = 0·
как/ки маленькой
случ. величин,
[redacted]

многомерного случайного вектора ΔY^k, δ – вектор малых положительных смещений. Однако событие $\Theta^k \subseteq R_m$ является искомым в нашей задаче. Поэтому остается только выбирать различные элементы $\Theta^\ell \subseteq R_m$, $\ell \subseteq J$; получать для них ΔY^ℓ при заданном событии Y и, вычислив вероятности для ΔY^1, "близких" к ΔY^ℓ, в виде:

$$P[(\Delta Y^\ell < \Delta Y^1 < \Delta Y^\ell + \delta)/(\theta_1^\ell, \theta_2^\ell, ..., \theta_m^\ell)], \qquad (1.1)$$

сравнивать эти вероятности.

Отметим, что условные вероятности (1.1) можно считать также и вероятностями получить вектор Y при измерениях, поскольку Y выражается через ΔY^ℓ и $< Y >^\ell$. Следовательно, перебор различных гипотез $\Theta^\ell \subseteq R_m$ приводит к сравнению условных вероятностей для данного случайного вектора измерений Y.

Чтобы отдать предпочтение какой-либо из гипотез, воспользуемся теоремой Бейеса /1/, согласно которой апостериорная вероятность события Θ равна априорной вероятности, умноженной на вероятность экспериментальных данных как функцию Θ, т.е.

$$P(\Theta/Y,L) = P(\Theta/L)P(Y/\Theta,L) \qquad (1.2)$$

для любого события $\Theta \subseteq R_m$. Таким образом, решение задачи (1.0) свелось к поиску абсолютного экстремума функции апостериорной вероятности (1.2). Поскольку $P(\Theta/L)$ – априорные вероятности, как правило, не известны, то, поступая в согласии с постулатом Бейеса[1], полагаем все гипотезы $\Theta \subseteq R_m$ равновероятными в некоторой области, которая может быть сделана достаточно малой в окрестности неизвестной точкой гипотезы, т.е. полагаем $P(\Theta/L)=1$. Следовательно, выбор оптимальной гипотезы накладывает требование максимальности на функцию правдоподобия $P(Y/\Theta,L)$. Поскольку оценки параметров и их ошибок, полученные из условий максимизации функции $P(Y/\Theta,L)$, обладают эффективными свойствами[2], то решение задачи свелось к построению функции правдоподобия[x]. Идея построения такой функции была изложена выше.

Способ определения параметров частиц, развитый в[4,5,6,7,8], несмотря на свою простоту и привлекательность обладает существенным недостатком - он не позволяет эффективным образом учитывать ядерное рассеяние наряду с учётом множественных случайных факторов (многократное кулоновское рассеяние, измерительные ошибки) и "непрерывных" (ионизационные потери, неоднородность поля, уравнение движения). Попытка учесть этот недостаток, предпринятая в[9,10], приводит, как нетрудно показать[3], к множественной корреляции оценок вектора измерительных ошибок и вектора кулоновских углов рассеяния, что существенным образом затрудняет учёт ядерного рассеяния на следе частицы.

Все сказанное приводит к необходимости выбора такого представления функции правдоподобия, которое позволит, сохраняя сильные стороны обоих упомянутых методов, в значительной мере исключить их недостатки.

---

[x] Нетрудно догадаться, что состоятельность оценок, полученных по ММП, равносильна сходимости в R_m последовательности (**) в смысле метрики (*).

Выбор представления

Предположим, что на треке измерено n – точек, которым присвоены номера начиная с 0, и система координат выбирается таким образом, чтобы ось OX совпала с вектором скорости частицы в первой точке на треке (см. рис. 1).

Обозначим через ξ_i смещение в i точке по отношению к $i-1$ точке, а через δ_i – случайное угловое отклонение в i -точке по отношению к направлению касательной в $(i-1)$ точке[x/].

$$M(\xi_i, \xi_j) = M(\delta_i, \delta_j) = M(\xi_j, \delta_i) = 0, \qquad \text{если } i \neq j. \tag{1.3}$$

Пусть $\ell_i = x_i - x_{i-1}$. Тогда, воспользовавшись распределением Ферми /11/

$$W(\xi_k, \delta_k, \ell_k) = \frac{2\sqrt{3}}{\pi} \cdot \frac{1}{\theta_s^2 \ell_k^2} \exp\left[-\frac{4}{\theta_s^2}\left(\frac{\delta_k^2}{\ell_k} - \frac{3\xi_k\delta_k}{\ell_k^2} + \frac{3\xi_k^2}{\ell_k^3}\right)\right], \tag{1.4}$$

в каждом из слоев ℓ_k можем записать в силу условий независимости

---

x/ $M(\alpha) = M(\beta) = 0$ /11/, где α и β принимают случайные значения ξ или δ. Здесь и далее такая запись будет означать математическое ожидание случайных величин.

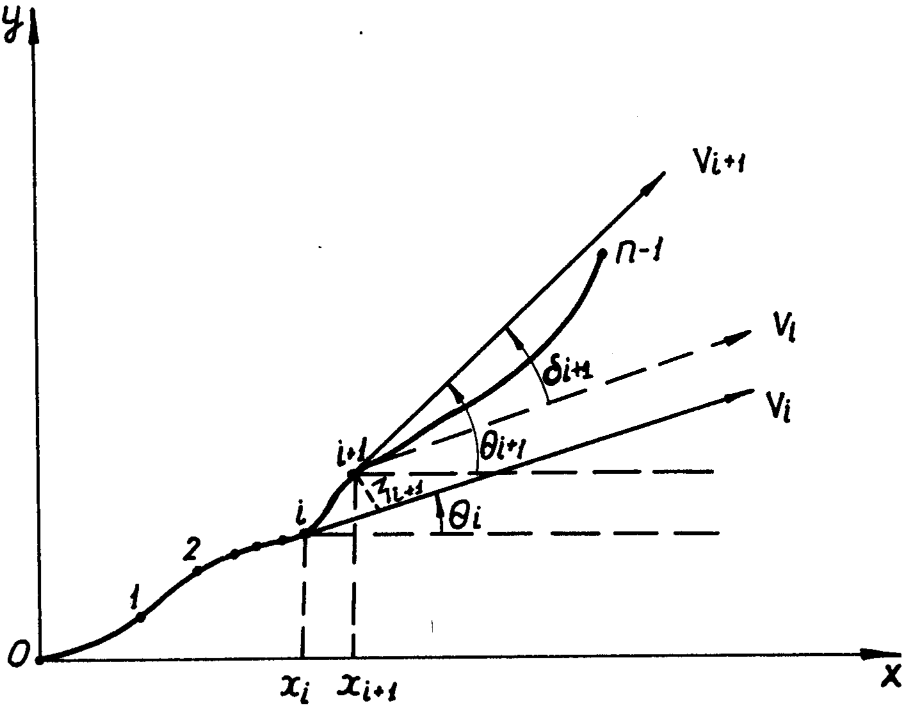

Рис. 1

(1,3) многомерную нормальную[x/] плотность распределения боковых смещений и угловых кулоновских отклонений одновременно для всей совокупности слоев

$$\Phi(\xi_1,\delta_1,\ell_1,\dots,\xi_{n-1},\delta_{n-1},\ell_{n-1}) = \frac{2^{(n-1)}\cdot 3^{(n-1)/2}\cdot\pi^{-(n-1)}\cdot\theta_s^{-2(n-1)}}{\prod_{i=1}^{n-1}\ell_i^2} \times \tag{1.5}$$
$$\times\exp\Big[-\frac{4}{\theta_s^2}\sum_{i=1}^{n-1}\Big(\frac{\delta_i^2}{\ell_i}-\frac{3\xi_i\delta_i}{\ell_i^2}+\frac{3\xi_i^2}{\ell_i^3}\Big)\Big]$$

Пусть θ_i - случайное угловое отклонение в i -ой точке по отношению к направлению касательной в первой точке на треке (оси ox).

Выразим смещение относительно ox вдоль оси OY через наборы независимых случайных величин $(\xi_1, \xi_2, \dots, \xi_{n-1})$ $(\delta_1, \delta_2, \dots, \delta_{n-1})$ с известными дисперсиями. Для этого воспользуемся справедливыми[xx/] в каждом слое ℓ_i при малых углах θ_i соотношениями [/5/]

[x/] Нетрудно показать, что в каждом слое ℓ_k распределение Ферми является двумерным гауссовским распределением. Рассмотрим несобственные интегралы:

$$\int_{-\infty}^{\infty} W(\xi,\delta,\ell)\,d\xi = \frac{1}{\sqrt{\pi}}\cdot\frac{1}{\theta_s\ell^{1/2}}\exp\Big[-\frac{\delta^2}{\theta_s^2\ell}\Big] \qquad \int_{-\infty}^{\infty} W(\xi,\delta,\ell)\,d\delta = \frac{\sqrt{3}}{\sqrt{\pi}}\cdot\frac{1}{\theta_s\ell^{3/2}}\exp\Big[-\frac{3\xi^2}{\theta_s^2\ell^3}\Big].$$

Первый из них означает, что на любом слое ℓ_k распределение проекции угла рассеяния безотносительно к положению ξ имеет гауссовский вид с дисперсией $\sigma^2(\delta)=\frac{1}{2}\theta_s^2\ell$. Второй интеграл означает, что боковое смещение безотносительно к угловому отклонению тоже имеет гауссовский вид с дисперсией $\sigma^2(\xi)=\frac{\theta_s^2}{6}\ell_k^3$. Поскольку частные распределения с указанными параметрами нормальны и совместное распределение существует, то (1.4) является двумерным гауссовским распределением с коэффициентом корреляции $\rho=\sqrt{3}/2$ (см. т. 2.3.1[/12/]).

[xx/] Следует отметить, что здесь и далее будем пользоваться обозначениями, принятыми в [/12/]. Латинскими жирными буквами будем обозначать векторы.

$$\xi_i = y_i - y_{i-1} - \ell_i \theta_{i-1},$$

$$\delta_i = \theta_i - \theta_{i-1}. \qquad (1 \le i \le n-1) \qquad (1.6)$$

Из соотношений (1.6) нетрудно получить выражение для кулоновских смещений в общем случае

$$y_i = \sum_{j=1}^{i} \xi_j + \sum_{j=1}^{i-1} \sum_{k=1}^{j} \ell_{j+1} \delta_k. \qquad (1.7)$$

Будем рассматривать теперь множества независимых случайных величин как два случайных вектора $Z^{(1)}$ и $Z^{(2)}$

$$Z^{(1)} = [\xi_1, \xi_2, \ldots, \xi_{n-1}]^T, \; Z^{(2)} = [\delta_1, \delta_2, \ldots, \delta_{n-1}]^T,$$

где T - означает транспонирование. Если ввести вектор

$$Z = [Z^{(1)}, Z^{(2)}]^T, \qquad (1.8)$$

то, как нетрудно увидеть, распределение (1.5) является совместным нормальным распределением $2(n-1)$ случайных величин со средними значениями

$$M(Z^{(1)}) = M(Z^{(2)}) = 0$$

и ковариациями

$$M(Z^{(1)}Z^{T(1)}) = \Sigma_{11}\ ,\quad M(Z^{(1)}Z^{T(2)}) = \Sigma_{12}\ ,$$

$$M(Z^{(2)}Z^{T(1)}) = \Sigma_{21}\ ,\quad M(Z^{(2)}Z^{T(2)}) = \Sigma_{22}\ .$$

Выпишем явный вид ковариационной матрицы распределения (1.5)

$$\Sigma = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right) = \left(\begin{array}{ccc|ccc} \sigma^2_{\xi_1} & & 0 & \sigma_{\xi_1\delta_1}\cdot\rho & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \sigma^2_{\xi_{n-1}} & 0 & & \rho\sigma_{\xi_{n-1}\delta_{n-1}} \\ \hline \sigma_{\delta_1\xi_1}\cdot\rho & & 0 & \sigma^2_{\delta_1} & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \rho\sigma_{\delta_{n-1}\xi_{n-1}} & 0 & & \sigma^2_{\delta_{n-1}} \end{array}\right)$$

т.е. это трехдиагональная блочная матрица, элементы которой известны.

Найдем теперь совместное распределение вектора случайных смещений вдоль OY и вектора угловых отклонений. Рассмотрим для этого линейное преобразование вектора Z

$$\begin{aligned} y_i^{(1)} &= \sum_{j=1}^{i} \xi_j + \sum_{j=1}^{i-1}\sum_{k=1}^{j} \ell_{j+1}\,\delta_k \\ y_i^{(2)} &= \delta_i \end{aligned} \qquad (1 \le i \le n-1) \qquad (2.0)$$

или в векторной записи

$$\begin{pmatrix} Y^{(1)} \\ \\ Y^{(2)} \end{pmatrix} = \left(\begin{array}{c|c} \Sigma E_{\Delta} & L_{\Delta} \\ \hline 0 & E \end{array}\right) \begin{pmatrix} Z^{(1)} \\ \\ Z^{(2)} \end{pmatrix} \qquad (2.0)'$$

где

$$E_{\Delta} = \begin{pmatrix} 1 & & & & 0 \\ 1 & 1 & & & \\ \vdots & \vdots & \ddots & & \\ \vdots & \vdots & & 1 & \\ & & & & \ddots \\ 1 & 1 & \cdots & \cdots & 1 \end{pmatrix}, \quad L_{\Delta} = \begin{pmatrix} 0 & & & & 0 \\ \ell_2 & 0 & & & \\ \sum\limits_{k=2}^{3} \ell_k & \ell_3 & \ddots & & \\ \vdots & \vdots & \ddots & \vdots & \\ \sum\limits_{k=2}^{n-1} \ell_k & \sum\limits_{k=3}^{n-1} \ell_k & \cdots & \ell_{n-1} & 0 \end{pmatrix},$$

$$E = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix},$$

т.е. это матрицы размерностей $[n-1,\ n-1]$.

Используя явный вид блочной матрицы преобразования, нетрудно показать, что (2.0) является невырожденным линейным преобразованием с якобианом

$$D\left[\frac{y_1, y_2, \dots, y_{n-1}; \delta_1, \delta_2, \dots, \delta_{n-1}}{\xi_1, \xi_2, \dots, \xi_{n-1}; \delta_1, \delta_2, \dots, \delta_{n-1}}\right] = 1.$$

Поэтому можем применить теорему (Т 2.4.1)[12] и найти совместное нормальное распределение для вектора $\tilde{Y} = [Y^{(1)}, Y^{(2)}]$. Поскольку вектор Z распределен $N(0, \Sigma)$, то $\tilde{\tilde{Y}}$ распределен $N(0, C\Sigma C^T)$, здесь C – есть краткое обозначение невырожденной матрицы преобразования (2.0).

Чтобы найти ковариационную матрицу $C\Sigma C^T$, воспользуемся правилами умножения клеточных матриц[13]. В результате получим

$$C\Sigma C^T = \left(\begin{array}{c|c} E_\Delta \Sigma_{11} E_\Delta^T + E_\Delta \Sigma_{12} L_\Delta^T + L_\Delta \Sigma_{21} E_\Delta^T + L_\Delta \Sigma_{22} L_\Delta^T & E_\Delta \Sigma_{12} + L_\Delta \Sigma_{22} \\ \hline \Sigma_{21} E_\Delta^T + \Sigma_{22} L_\Delta^T & \Sigma_{22} \end{array}\right)$$

где матрицы-множители определены выше.

Проделав довольно громоздкую работу по вычислению отдельных клеток, получим в итоге следующий вид ковариационной матрицы для вектора $\tilde{\tilde{Y}}$

$${}_1\Sigma = C\Sigma C^T = \left(\begin{array}{c|c} {}_1\Sigma_{11} & {}_1\Sigma_{12} \\ \hline {}_1\Sigma_{21} & {}_1\Sigma_{22} \end{array}\right).$$

где

$$ {}_1\Sigma_{11} = M\,[\,Y^{(1)} Y^{T(1)}] = \frac{\Theta_s^2}{12}\, x_i^2\,(3x_j - x_i) \qquad (j \geq i) $$

$$ {}_1\Sigma_{22} = M\,[\,Y^{(2)} Y^{T(2)}] = \frac{\Theta_s^2}{2}\,(x_i - x_{i-1}) $$

$$ {}_1\Sigma_{12} = {}_1\Sigma_{21}^{T} = \begin{cases} 0, & j > i \\ \dfrac{\Theta_s^2}{4}\,[(x_j - x_{j-1})(2x_i - x_j - x_{j-1})], & i \geq j\,. \end{cases} \qquad (2.2) $$

В этих выражениях

$$ \Theta_s^2 = \text{const}\;\frac{m^2 + p^2}{p^4}, $$

где m и p – масса и импульс частиц[11]. Отметим, что ковариационная матрица вектора $\underset{\approx}{Y}$ обозначена ${}_1\Sigma$ в отличие от ковариационной матрицы Σ вектора Z.

Итак, получено совместное нормальное распределение для вектора случайных смещений точек трека от прямой линии и вектора случайных угловых отклонений из-за многократного кулоновского рассеяния

$$ n(\tilde{\tilde{y}}/0, {}_1\Sigma) = (2\pi)^{-(n-1)}\,|{}_1\Sigma|^{-1/2} \exp\left[-\frac{1}{2}\,\tilde{\tilde{Y}}^{T}\,{}_1\Sigma^{-1}\,\tilde{\tilde{Y}}\right], \qquad (2.3) $$

где элементы ковариационной матрицы $_1\Sigma$ определяются выражениями (2.2).

Рассматривая (1.4) как плотность условных вероятностей[x/] отклонения частицы на угол δ_j и смещения на величину ξ_j за промежуток времени $t_j - t_{j-1} \geq 0$, приходим к выводу[/5/], что процесс многократного кулоновского рассеяния является марковским[/14/] случайным процессом, для которого совместная плотность распределения вероятностей ординат определяется выражением (1.5).

Однако, как следует из (1.3), корреляционная функция такого процесса $K(t_i, t_j)$ отлична от нуля только при $t_i = t_j$, т.е. представление (1.5) функции правдоподобия "лишает" реальный кулоновский процесс известной "плавности".

Поэтому переход (2.0) к новым ординатам, приводящий к более сложной (2.3) модели реального процесса, чем (1.5), позволяет не только "вернуть плавность", т.е. рассмотреть процесс в единой системе координат, но и исследовать явления, возникающие вследствие наличия функциональной зависимости между разноименными ординатами процесса в различные моменты времени.

Учитывая вышесказанное, далее при построении полной функции правдоподобия $P(Y/\Theta, L)$ будем рассматривать вместо кулоновской матрицы $_1\tilde{\Sigma}_{11}$, полученной в[/5/], матрицу полного кулоновского процесса рассеяния $_1\Sigma$ и нормальную плотность $n(\tilde{\tilde{y}}/0, {}_1\Sigma)$, записанную в виде (2.3).

[x/] Такое предположение справедливо, если начало системы координат при движении частицы в слое ℓ_k совпадает с положением ее в момент t_{k-1}.

В заключение автор благодарит Н.Н. Говоруна и И.Н. Силина за постоянное внимание к работе и полезные обсуждения, а также В.И. Мороза, Г.А. Ососкова, В.П. Жигунова за ценные советы.

Литература

1. М. Дж. Кендал. Теория распределений т. 1, Наука, Москва, 1966.
2. У. Гренадер. Случайные процессы и статистические выводы ИИЛ. Москва, 1961.
3. В.И. Мороз. Автореферат диссертации. Дубна, 1968 г.
4. И.М. Граменицкий, Л.А. Тихонова, П.В. Шляпников. Препринт ОИЯИ, Р-2146, Дубна, 1961.
5. Б.А. Манюков, П.В. Шляпников. Препринт ОИЯИ, Р10-4256, Дубна, 1969.
6. Г.А. Ососков, П.В. Шляпников. Препринт ОИЯИ, Р10-4261, Дубна, 1969.
7. Б.А. Манюков, П.В. Шляпников. Препринт ОИЯИ, Р10-4574, Дубна, 1969.
8. Л.Н. Гердюков, Б.А. Манюков, П.В. Шляпников. Препринт ОИЯИ, Р10-4579, Дубна, 1969.
9. Г.А. Емельяненко, К.П. Ломов, Г.И. Макаренко, В.И. Мороз, И.С. Саитов, А.П. Стельмах. Препринт ОИЯИ, Р-2829, Дубна, 1966.
10. Г.А. Емельяненко. Препринт ОИЯИ, Р11-3977, Дубна, 1968.
11. Б. Росии. Частицы больших энергий. ГИТЛ, Москва, 1955.
12. Т. Андерсон. Введение в многомерный статистический анализ, ГИФМЛ, Москва, 1961.
13. Г.Е. Шилов. Математический анализ. Конечномерные линейные пространства. Изд. "Наука", 1969 г., Москва.
14. А.А. Свешников. Прикладные методы теории случайных функций. Изд. "Наука", 1968 г., Москва.

Рукопись поступила в издательский отдел
24 июля 1970 года.