

3612

Fortseh. Physik., 1968, Bd. 16,
№ 6, s. 325-355

ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

P1 - 3612



ЛАБОРАТОРИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ
И АВТОМАТИЗАЦИИ

А.Пазман, В.В.Федоров

ПЛАНИРОВАНИЕ ФИЗИЧЕСКИХ ЭКСПЕРИМЕНТОВ

(Обзор)

1967.

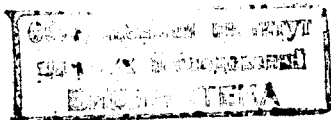
P1 - 3612

А.Пазман, В.В.Федоров*

ПЛАНИРОВАНИЕ ФИЗИЧЕСКИХ ЭКСПЕРИМЕНТОВ

(Обзор)

* Московский Государственный университет им. Ломоносова



дм
1/9814

При современном уровне развития теории и техники многие физические исследования (в частности, исследование свойств элементарных частиц) требуют постановки очень сложных, дорогостоящих и длительных экспериментов. Как показал имеющийся опыт, планирование подобных экспериментов с помощью методов математической статистики позволяет в большинстве случаев в несколько раз более эффективно использовать средства (время, денежные или материальные затраты) по сравнению с пассивным экспериментом. В связи с этим становится необходимым широкое применение методов математического планирования экспериментов.

В настоящем обзоре излагаются существующие методы планирования экспериментов в форме, доступной для широкого круга физиков-экспериментаторов. В обзоре рассматриваются примеры из области экспериментов по фазовому анализу, упругого рассеяния элементарных частиц и по измерению дифференциального эффективного сечения. Для простоты изложения они сильно упрощены и схематизированы. Примеры выбраны из ядерной физики не случайно, так как именно здесь затраты на эксперименты особенно большие.

По сравнению с книгой [1], которая содержит главу о планировании физических экспериментов, настоящая работа охватывает значительно более широкий круг вопросов.

О Г Л А В Л Е Н И Е:

	стр.
Список символов.....	5
§1. Метод максимума правдоподобия.....	7
§2. Критерии оптимальности регрессионного эксперимента.....	12
§3. Статическое планирование регрессионных экспериментов.....	18
§4. Последовательное и непрерывное планирование.....	28
§5. Планирование дискриминирующих экспериментов.....	34
§6. Информационный метод планирования экспериментов.....	38

Список символов

- $E\{ \}$ - оператор усреднения,
 $\mathfrak{D}\{ \}$ - оператор вычисления дисперсии случайной величины или ковариационной матрицы многомерной случайной величины,
 T - транспозиция матрицы,
 $^{-1}$ - обращение матрицы,
 θ_α - неизвестный параметр,
 $\vec{\theta} = (\theta_1, \dots, \theta_m)'$ - вектор всех неизвестных параметров,
 $x_l - \hat{l}$ - та-я контролируемая переменная,
 \vec{x} - вектор всех контролируемых переменных (допустимая точка измерения),
 $\vec{x}^{(k)}$ - вектор оптимальных значений, контролируемых переменных (k -тая точка измерения),
 $\eta(\vec{\theta}, \vec{x})$ - теоретическая зависимость измеряемой величины от параметров $\vec{\theta}$ и от контролируемых переменных \vec{x} ,
 y, y_l - результаты (непосредственного) измерения величины η
 t - время измерения,
 $t^{(k)}$ - оптимальное время измерения в точке измерения $\vec{x}^{(k)}$,
 T - общее количество времени (или затрат), отведенное на эксперимент,
 $\sigma^2(\vec{x}), \sigma_1^2(\vec{x})$ (или σ^2, σ_1^2) - дисперсия величин y, y_l
 $(\sigma_1^2 = \mathfrak{D}\{y_l\})$,
 w_1, w - веса измерений,
 $\lambda(\vec{x})$ - функция эффективности
 $(\lambda(\vec{x}) = 1/\sigma^2(\vec{x})t)$,
 $f_\alpha(\vec{x})$ - коэффициент линейного приближения функции $\eta(\vec{\theta}, \vec{x})$
 $(f_\alpha(\vec{x}) = \frac{\partial \eta(\vec{\theta}, \vec{x})}{\partial \theta_\alpha} \Big|_{\vec{\theta} = \hat{\vec{\theta}}^{(0)}})$,
 F - матрица коэффициентов ($F_{\alpha i} = f_\alpha(\vec{x}^{(i)})$),
 $L(\vec{\theta})$ - функция правдоподобия,
 $\hat{\theta}$ - оценка вектора $\vec{\theta}$,
 $\hat{\vec{\theta}}^{(0)}$ - предварительная оценка вектора $\vec{\theta}$,

D (или $D(0)$, или $D(T, \vec{x})$) - матрица ошибок вектора оценок $\vec{\theta}$

$$(D = \mathcal{D}\{\vec{\theta}\}),$$

M - информационная матрица

$$(M = D^{-1}),$$

$\vec{\omega}$ - оценка вектора $\vec{\omega}$,

$D(\vec{\omega})$ (или $D(0; \vec{\omega})$ или $D(T, \vec{x}; \vec{\omega})$) матрица ошибок, соответствующая параметрам $\vec{\omega}$

$$(D(\vec{\omega}) = \mathcal{D}\{\vec{\omega}\}),$$

$\hat{\eta}(\vec{x})$ - оценка функции $\eta(\vec{\theta}, \vec{x})$

$$(\hat{\eta}(\vec{x}) = \eta(\vec{\theta}, \vec{x})),$$

$\hat{\sigma}(\vec{x})$ - коридор ошибок оценки $\hat{\eta}(\vec{x})$

$$(\hat{\sigma}^2(\vec{x}) = \mathcal{D}\{\hat{\eta}(\vec{x})\}),$$

$s(\vec{x}, \vec{x}^*)$ - ковариационная (корреляционная) функция, соответствующая случайной функции $\hat{\eta}(\vec{x})$,

δ_{ij} - элемент единичной матрицы

$$(\delta_{ij} = 1, \text{ если } i=j, \delta_{ij} = 0, \text{ если } i \neq j).$$

§1. Метод максимума правдоподобия

В настоящем параграфе будут изложены некоторые вопросы обработки экспериментальных данных в размерах, необходимых для построения математического аппарата планирования экспериментов.

Пусть в точке \vec{x} измеряется некоторая физическая величина $y(\vec{x})$, причём закон распределения $p[y(\vec{x})/\vec{\theta}]$ предполагается известным и

$$E\{y(\vec{x})\} = \eta(\vec{\theta}, \vec{x}). \quad (1.1)$$

Здесь оператор E означает операцию усреднения по всем возможным выборкам y в точке \vec{x} ,

$$\vec{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_m \end{pmatrix} \quad - \text{неизвестные параметры,}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad - \text{контролируемые переменные}$$

(температура, давление, угол рассеяния и т.д.). В дальнейшем вместо термина "значение контролируемых переменных" нам будет иногда удобней пользоваться термином "положение измерительного прибора".

Целью эксперимента обычно является определение неизвестных параметров $\vec{\theta}$, точнее, нахождение их оценок $\hat{\vec{\theta}}$. Предположим, что измерения проведены в точках $\vec{x}^{(1)}, \dots, \vec{x}^{(n)}$ с результатами y_1, \dots, y_n . Составим функцию правдоподобия^{/1/}

$$L(\vec{\theta} / y_1, \dots, y_n) = p_1[y_1(\vec{x}^{(1)}) / \vec{\theta}] \dots p_n[y_n(\vec{x}^{(n)}) / \vec{\theta}]. \quad (1.2)$$

За оценки параметров $\vec{\theta}$ принимаются те значения $\hat{\vec{\theta}}$, при которых функция правдоподобия достигает абсолютного максимума

$$L(\hat{\vec{\theta}} / y_1, \dots, y_n) = \max_{\vec{\theta}} L(\vec{\theta} / y_1, \dots, y_n). \quad (1.3)$$

Для нас особый интерес будет представлять случай, когда

$$p_1(y_1 / \vec{\theta}) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2} \left[\frac{y_1 - \eta(\vec{\theta}, \vec{x}^{(1)})}{\sigma_1} \right]^2} \quad (1.4)$$

Можно показать^{/1/}, что при выполнении (1.4) оценки $\hat{\vec{\theta}}$, полученные из (1.3), будут состоятельными, т.е. они сходятся по вероятности к истинным параметрам при безграничном увеличении объема выборки $\{y_1, \dots, y_n\}$, асимптотически эффективными, т.е. определитель матрицы ошибок искомых параметров будет наименьшим из всех возможных, и асимптотически нормально распределенными:

$$p(\hat{\vec{\theta}} | \vec{\theta}) = \frac{e^{-\frac{1}{2}(\hat{\vec{\theta}} - \vec{\theta})' D^{-1}(\hat{\vec{\theta}} - \vec{\theta})}}{(2\pi)^{\frac{m}{2}} |D|^{1/2}}, \quad (1.5)$$

где штрих означает транспонирование.

Общее выражение для матрицы ошибок равно^{/1/}:

$$D_{\alpha\beta}^{-1} = -E \left\{ \frac{\partial^2 \log L}{\partial \theta_\alpha \partial \theta_\beta} \Big|_{\vec{\theta} = \hat{\vec{\theta}}} \right\} = E \left\{ \frac{\partial \log L}{\partial \theta_\alpha} \frac{\partial \log L}{\partial \theta_\beta} \Big|_{\vec{\theta} = \hat{\vec{\theta}}} \right\}. \quad (1.6)$$

$\alpha, \beta = 1, \dots, m$

Оценки $\hat{\vec{\theta}}$ и матрица ошибок D особенно просто отыскиваются, если измеряемые величины линейно зависят от искомых параметров:

$$\eta(\vec{\theta}, \vec{x}) = \sum_{\alpha=1}^m \theta_\alpha f_\alpha(\vec{x}) = \vec{f}(\vec{x}) \vec{\theta}, \quad (1.7)$$

где

$$\vec{f}(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x})) \text{ и } f_\alpha(\vec{x}) -$$

известные функции. Из (1.4), (1.7) и (1.3) нетрудно получить, что

$$\hat{\vec{\theta}} = M^{-1} \vec{Y}, \quad (1.8)$$

где

$$Y_\beta = \sum_{i=1}^n w_i f_\beta(\vec{x}^{(i)}) y_i, \quad w_i = \frac{1}{\sigma_1^2}, \quad (1.9)$$

$$M_{\alpha\beta} = \sum_{i=1}^n w_i f_\alpha(\vec{x}^{(i)}) f_\beta(\vec{x}^{(i)}). \quad (1.10)$$

Матрица M называется информационной матрицей Фишера и обратной матрице ошибок:

$$D = M^{-1}. \quad (1.11)$$

Если проделано два эксперимента с информационными матрицами M_I и M_{II} , то из (1.10) вытекает, что их совокупность характеризуется информационной матрицей.

$$M = M_I + M_{II} \quad (1.12)$$

(аддитивность информационной матрицы Фишера).

Зная матрицу ошибок оценок $\hat{\theta}$, можно найти матрицу ошибок произвольной линейной комбинации $\hat{A} = C' \hat{\theta} / 2$:

$$D\{\hat{A}\} = C' D C. \quad (1.13)$$

Воспользовавшись (1.13), можно найти корреляционную функцию случайной величины $\hat{\eta}(\vec{x}) = \vec{f}'(\vec{x}) \hat{\theta}$:

$$s(\vec{x}, \vec{x}^*) = \vec{f}'(\vec{x}) D \vec{f}(\vec{x}^*). \quad (1.14)$$

Величина

$$\hat{\sigma}(\vec{x}) = \sqrt{s(\vec{x}, \vec{x})} = \sqrt{\vec{f}'(\vec{x}) D \vec{f}(\vec{x})} \quad (1.15)$$

называется коридором ошибок кривой $\hat{\eta}(\vec{x})$.

Если объем выборки $\{y_1, \dots, y_n\}$ настолько велик, что в области, ограниченной эллипсоидом рассеяния $/3/$ для некоторой предварительной оценки $\hat{\theta}^{(0)}$ справедливо:

$$\eta(\vec{\theta}, \vec{x}) = \eta(\hat{\theta}^{(0)}, \vec{x}) + \sum_{\alpha=1}^m \frac{\partial \eta(\vec{\theta}, \vec{x})}{\partial \theta_{\alpha}} \Big|_{\vec{\theta}=\hat{\theta}^{(0)}} (\theta_{\alpha} - \hat{\theta}_{\alpha}^{(0)}), \quad (1.16)$$

то можно пользоваться формулами (1.9)–(1.15) как приближенными, положив

$$f_{\alpha}(\vec{x}) = \frac{\partial \eta(\vec{\theta}, \vec{x})}{\partial \theta_{\alpha}} \Big|_{\vec{\theta}=\hat{\theta}^{(0)}} \text{ и считая } \theta_{\alpha} - \hat{\theta}_{\alpha}^{(0)}$$

искомыми параметрами.

Пример 1.1. Фазовый анализ упругого рассеяния элементарных частиц

Целью многих экспериментов по рассеянию элементарных частиц является определение фазовых сдвигов и коэффициентов смешивания парциальных волновых функций, соответствующих разным значениям орбитального количества движения $/3,4/$. В таких экспериментах фазовые сдвиги и коэффициенты смешивания (далее – фазовые сдвиги) составляют набор неизвестных параметров $\theta_1, \dots, \theta_m$.

В каждом отдельном эксперименте измеряется некоторая величина y_1 , характеризующая исследуемое взаимодействие (например, эффективное сечение, поляризация, величины, определяющие поворот вектора поляризации при рассеянии и т.д.). Сама величина y_1 – результат исследования большого количества рассеянных (случайно) частиц. Поэтому можно определить и ее дисперсию $\sigma_1^2 = D\{y_1\}$. При достаточно большой статистике (числе исследованных треков) величину y_1 можно считать распределенной по нормальному закону (1.4). Метод максимума правдоподобия сводится в этом случае к методу наименьших квадратов, т.е. к отысканию $\vec{\theta} = \vec{\theta}^*$, в котором достигается

$$\min_{\vec{\theta}} \sum_{i=1}^n \left[\frac{y_i - \eta_1(\vec{\theta})}{\sigma_1} \right]^2. \quad (1.17)$$

Функции $\eta_1(\vec{\theta})$ известны из теории исследуемого рассеяния.

Пример 1.2. Измерение дифференциального эффективного сечения точечными детекторами (счётчиками)

Обсудим эксперимент по рассеянию элементарных частиц, в котором дифференциальное эффективное сечение $\eta(x)$ (x – угол рассеяния) непосредственно измеряется на выбранных (дискретных) углах $x^{(1)}, \dots, x^{(n)}$. Цель эксперимента – оценка дифференциального сечения в некотором интервале углов рассеяния. Предположим, что в рассматриваемом интервале углов возможно приближение дифференциального сечения $\eta(x)$ многочленом $m-1$ -ой степени:

$$\eta(\vec{\theta}, \vec{x}) = \sum_{\alpha=1}^m \theta_{\alpha} p_{\alpha}(x). \quad (1.18)$$

Оценку $\hat{\eta}(\mathbf{x})$ вычислим косвенно, через оценки параметров $\hat{\theta}_\alpha$. В (1.18) $p_\alpha(\mathbf{x})$ — удобно подобранные многочлены порядка $\alpha - 1$. Число параметров $\hat{\theta}_\alpha$ не может превышать числа точек измерения $\mathbf{x}^{(i)}$ ($n \geq m$) и обыкновенно $n = m$.

При достаточно большой "статистике" (большом количестве зарегистрированных частиц) вычисление оценок $\hat{\theta}_\alpha$ сводится к методу наименьших квадратов (формула (1.17)).

В описываемом эксперименте дисперсия σ_1^2 пропорциональна измеренному значению y_1 дифференциального сечения на угле $\mathbf{x}^{(1)}$ (см. /8/ г. 9, 81)

$$\sigma_1^2 \approx y_1 \quad (1.19)$$

(среднее и дисперсия совпадают, если случайная величина распределена по Пуассону).

Оценка дифференциального сечения при любом угле \mathbf{x} равна

$$\hat{\eta}(\mathbf{x}) = \sum_{\alpha=1}^m \hat{\theta}_\alpha p_\alpha(\mathbf{x}) \quad (1.20)$$

и дисперсия этой оценки (квадрат коридора ошибок) равен

$$\hat{\sigma}^2(\mathbf{x}) = \sum_{\alpha, \beta=1}^m p_\alpha(\mathbf{x}) D_{\alpha\beta} p_\beta(\mathbf{x}). \quad (1.21)$$

Здесь $D_{\alpha\beta}$ — элемент матрицы ошибок вычислений по формулам (1.10) и (1.11).

§2. Критерии оптимальности регрессионного эксперимента

Пусть в области $\vec{\mathbf{x}} \in X$ возможно измерение величины y , причём

$$E\{y\} = \eta(\vec{\theta}, \vec{\mathbf{x}}). \quad (2.1)$$

Производя измерения, экспериментатор может, вообще говоря, преследовать совершенно различные цели. Так, например, его может интересовать либо наиболее точное значение оценок $\vec{\theta}$, либо наиболее точная аппроксимация функции $\eta(\vec{\theta}, \vec{\mathbf{x}})$ в области, не принадлежащей X и т.д. Очевидно, что оптимальное размещение затрат будет существенно зависеть от требований, предъявляемых к эксперименту. Поэтому, прежде всего, необходимо сформулировать критерий оптимальности эксперимента. Формально это означает следующее. Для того чтобы спланировать оптимальный эксперимент, необходимо построить такой функционал, значения которого характеризовали бы количественно цель эксперимента. Обозначим этот функционал через

$$\psi(\vec{\mathbf{x}}^{(1)}, \dots, \vec{\mathbf{x}}^{(n)}, t^{(1)}, \dots, t^{(n)}), \quad (2.2)$$

где $\vec{\mathbf{x}}^{(1)}, \dots, \vec{\mathbf{x}}^{(n)}$ точки, в которых производятся измерения в течение времени $t^{(1)}, \dots, t^{(n)}$, соответственно. Если функционал ψ построен, то задача планирования заключается в отыскании максимума (или минимума) этого функционала по числу точек измерения n по значениям точек измерения (значениям контролируемых переменных) и по распределению времени, в течение которых производятся измерения:

$$\max_{\substack{\vec{\mathbf{x}}^{(1)}, \dots, \vec{\mathbf{x}}^{(n)} \\ t^{(1)}, \dots, t^{(n)}}} \psi(\vec{\mathbf{x}}^{(1)}, \dots, \vec{\mathbf{x}}^{(n)}, t^{(1)}, \dots, t^{(n)}) \quad (2.3)$$

при условии: $\sum_{i=1}^n t^{(i)} = T$,

где T — время, отведенное на эксперимент.

В дальнейшем вместо (2.3) мы будем писать

$$\max_{\epsilon} \psi(\epsilon), \quad (2.3a)$$

где эксперимент ϵ подразумевает проведение измерений в точках $\vec{\mathbf{x}}^{(1)}, \dots, \vec{\mathbf{x}}^{(n)}$ в течение времени $t^{(1)}, \dots, t^{(n)}$ соответственно, $\sum t^{(i)} = T$.

Замечаем, что абсолютно наилучшего эксперимента, который бы максимизировал все возможные ψ не существует. Поэтому в зависимости от целей экспериментатора необходимо применять подходящие критерии оптимальности.

Приведем наиболее важные и распространенные критерии оптимальности эксперимента.

1. D - оптимальность

Как правило, точность оценок $\vec{\theta}$ искомых параметров $\vec{\theta}$ характеризуют матрицей ошибок D этих оценок. Если цель эксперимента - определение параметров $\vec{\theta}$, то естественно обратиться к такой мере результатов эксперимента, которая разумным образом строится из элементов матрицы ошибок D .

Практика показывает, что наиболее удобно и естественно в этом случае взять в качестве функционала ψ определитель матрицы ошибок интересующих экспериментатора параметров $|D|$. Тогда планирование оптимального эксперимента будет заключаться в отыскании

$$\min_{\epsilon} |D|, \quad (2.4)$$

или в отыскании

$$\max_{\epsilon} |M|, \quad (2.5)$$

где M - информационная матрица Фишера (см. (1.10)).

Определителю $|D|$ можно дать геометрическую интерпретацию, а, именно, величина определителя $|D|$ пропорциональна объему эллипсоида рассеяния (см. главу 2.2 из [3]). Связь элементов матрицы ошибок с параметрами эллипсоида рассеяния представлена на рис. 1 (двумерный случай).

Если для экспериментатора особый интерес представляет k параметров $(\theta_1, \dots, \theta_k)$ из m неизвестных, то необходимо минимизировать определитель $|D(\vec{\omega})|$ матрицы ошибок, соответствующей параметрам $(\theta_1, \dots, \theta_k)' = \vec{\omega}$.

2. A -оптимальность

Некоторые авторы (см., например, [7]) при отыскании неизвестных параметров вместо определителя матрицы ошибок минимизируют сумму дисперсий параметров $\vec{\theta}$:

$$\min_{\epsilon} \text{Sp } D = \min_{\epsilon} \sum_{\alpha=1}^m D_{\alpha\alpha}. \quad (2.6)$$

Величина $\text{Sp } D$ пропорциональна квадрату длины диагонали прямоугольного параллелепипеда, описывающего эллипсоид рассеяния (см. рис. 1).

В случае необходимости наиболее точного определения k параметров из m планирование заключается в отыскании

$$\min_{\epsilon} \text{Sp } D(\vec{\omega}) = \min_{\epsilon} \sum_{\alpha=1}^k D_{\alpha\alpha}. \quad (2.6a)$$

3. d -оптимальность

Пусть экспериментатора интересует некоторая линейная функция искомых параметров

$$h(\vec{\theta}) = \sum_{\alpha=1}^m f_{\alpha} \theta_{\alpha}. \quad (2.7)$$

Эксперимент будет считаться оптимальным, если дисперсия оценки этой функции

$$D\{\hat{h}\} = \vec{f}' D \vec{f} \quad (2.8)$$

имеет наименьшее возможное значение. Здесь $\vec{f}' = (f_1, \dots, f_m)$. В случае нелинейной зависимости функции $h(\vec{\theta})$ от параметров в (2.7) следует положить

$$\vec{f}' = \left(\frac{\partial h}{\partial \theta_1}, \dots, \frac{\partial h}{\partial \theta_m} \right)_{\vec{\theta} = \vec{\theta}^{(0)}}. \quad (2.9)$$

Здесь $\vec{\theta}^{(0)}$ — предварительные оценки неизвестных параметров (сравни с (1.16)). Предполагается, что $\vec{\theta}^{(0)}$ близки к истинным значениям определяемых параметров.

4. Минимаксные критерии

Пусть $\eta(\vec{\theta}, \vec{x})$ — линейная функция неизвестных параметров:

$$\eta(\vec{\theta}, \vec{x}) = \vec{f}'(\vec{x}) \vec{\theta}.$$

Рассмотрим дисперсию функции $\eta(\vec{\theta}, \vec{x})$ в заданной точке \vec{x} . Согласно (2.8),

$$\mathcal{D}\{\eta(\vec{\theta}, \vec{x})\} = \vec{f}'(\vec{x}) \mathcal{D}\vec{f}(\vec{x}). \quad (2.10)$$

Если экспериментатора интересует точное определение "поверхности" $\eta(\vec{\theta}, \vec{x})$, то естественно потребовать, чтобы максимальная дисперсия $\mathcal{D}\{\eta(\vec{\theta}, \vec{x})\}$ в области $\vec{x} \in X'$ была минимальной:

$$\min_{\vec{\theta}} \max_{\vec{x} \in X'} \mathcal{D}\{\eta(\vec{\theta}, \vec{x})\}. \quad (2.11)$$

Отметим, что область X' может быть частью X (области, где измерения возможны), может включать в себя X полностью или частично и может не иметь с областью X общих точек (экстраполяция в область X'). В частном случае X' может вырождаться в точку.

5. I — оптимальность

При нелинейной зависимости $\eta(\vec{\theta}, \vec{x})$ от искомых параметров $\vec{\theta}$ и при не очень большом объеме выборки может оказаться, что описание степени точности наших знаний с помощью вторых моментов (матрицы ошибок D) либо невозможно, либо неполно. В этом случае удобно пользоваться энтропийной мерой ко-

личества информации, приобретенной в результате эксперимента

$$I = \int L(\vec{\theta}/\vec{y}) \log L(\vec{\theta}/\vec{y}) d\vec{\theta}, \quad (2.12)$$

где $L(\vec{\theta}/\vec{y})$ — нормированная функция правдоподобия.

Оптимальным будет тот эксперимент, который минимизирует средний по всем возможным выборкам прирост количества информации:

$$\min_{\epsilon} J = \min_{\epsilon} E\{I - I_0\}, \quad (2.13)$$

где I_0 — количество информации, имевшееся до начала планируемого эксперимента.

Если экспериментатора интересует k параметров из m , то в (2.12) необходимо вести интегрирование только по "нужным" параметрам $\vec{\omega}$.

Замечаем, что информационный критерий является некоторым обобщением описанных выше критериев регрессионных экспериментов и описанного в §5 критерия для дискриминирующих экспериментов (более подробно см. §6).

Все рассмотренные критерии оптимальности приведены в таблице 1. В этой таблице отражена связь между различными критериями.

Пример 2.1.

Критерии D — оптимальности и I — оптимальности употребляются при планировании экспериментов по упругому рассеянию (см. пример 1.1). Экспериментатор требует максимальной точности всех фазовых сдвигов до фазы, соответствующей некоторому максимальному значению момента количества движения. D — оптимальность требуется, если набор оценок фазовых сдвигов однозначен (одно решение в (1.17)) и I — оптимальность требуется в случае неоднозначности набора фазовых сдвигов (несколько решений в (1.17)).

Пример 2.2.

Минимаксный критерий подходит для планирования экспериментов из примера 1.2.

Условие (2.11) – это требование, чтобы в заданном интервале углов максимальное отклонение оценки дифференциального сечения $\hat{\eta}(\mathbf{x})$ от его истинного значения $\eta(\mathbf{x})$ было минимальным. Это требование эквивалентно требованию максимальной точности всех параметров θ_α в (1.18) (в смысле критерия D-оптимальности; см. таблицу 1).

Иногда можно потребовать, чтобы среднеквадратическое отклонение оценки $\hat{\eta}(\mathbf{x})$ от истинной кривой $\eta(\mathbf{x})$ было минимальным, т.е. чтобы достиглось

$$\min_{\theta} \int_a^b \hat{\sigma}^2(\mathbf{x}) d\mathbf{x}, \quad (2.14)$$

где $\langle a, b \rangle$ – интервал, на котором оцениваем дифференциальное сечение.

В работе [7] доказывалась эквивалентность (2.14) с критерием A-оптимальности в предположении, что $p = m$ (число параметров совпадает с числом точек измерения) и что многочлены $p_\alpha(\mathbf{x})$ ортогональны, т.е.

$$\int_a^b q(\mathbf{x}) p_\alpha(\mathbf{x}) p_\beta(\mathbf{x}) d\mathbf{x} = \delta_{\alpha\beta}, \quad (2.15)$$

где $q(\mathbf{x})$ – произвольная непрерывная весовая функция.

§3. Статическое планирование регрессионных экспериментов

Рассмотрим случай, когда измеряемая величина η линейно зависит от искомых параметров θ :

$$\eta(\theta, \mathbf{x}) = f'(\mathbf{x}) \theta, \quad (3.1)$$

где $f_\alpha(\mathbf{x})$ – известные функции. Предположим, что известна функция эффективности $\lambda(\mathbf{x})^{1/2}$, определенная выражением:

$$\lambda(\mathbf{x}) = \frac{1}{\sigma^2(\mathbf{x}) t}, \quad (3.2)$$

где $\sigma^2(\mathbf{x})$ – дисперсия измерения y , проведенного в точке \mathbf{x} , полученная за время t .

Статическое планирование – это планирование сразу всего эксперимента, т.е. от начала до его полного завершения. Очевидно, что для статического планирования необходимо, чтобы функция эффективности была неизменна на протяжении всего эксперимента или ее зависимость от времени была бы известна до опыта. В настоящем параграфе предполагается, что экспериментальные условия неизменны, и что начальная информация о параметрах θ отсутствует ($M=0$). Статическое планирование при изменяющихся экспериментальных условиях и при наличии начальной информации является сложной и громоздкой вычислительной задачей и не будет непосредственно рассматриваться нами. Некоторые рецепты по приближенному решению этих задач кратко излагаются в §4.

Как уже упоминалось в §2, оптимальный эксперимент определяется выбранным экспериментатором критерием оптимальности и соответствующим набором точек измерения $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}$ и временах измерения в этих точках $t^{(1)}, \dots, t^{(n)}$. При этом должно выполняться условие $\sum_{i=1}^n t^{(i)} = T$, где T – время всего эксперимента.

1. Одна контролируемая переменная

Ограничимся вначале предположением, что существует единственная контрольная переменная x , принимающая значение из некоторого интервала (a, b) . Общий случай, когда переменная x принимает значения из некоторого компактного множества, рассмотрен в [8].

1. D-оптимальные планы

Минимум определителя матрицы ошибок соответствует максимуму определителя информационной матрицы M (см. (1.11))

$$M = F W F', \quad (3.3)$$

где

$$F_{\alpha i} = f_\alpha(\mathbf{x}^{(i)}) \quad i = 1, \dots, n, \alpha = 1, \dots, m$$

$$W_{ij} = \lambda(x^{(1)}) t^{(1)} \delta_{ij} \quad i, j = 1, \dots, n.$$

Из требования $|M| \neq 0$ следует, что нижняя граница числа точек, в которых необходимо производить измерения, должна быть не меньше числа неизвестных параметров:

$$n \geq m. \quad (3.4)$$

Так как информационная матрица симметрична, то она полностью задается своими $\frac{m(m+1)}{2}$ элементами. Основываясь на этом и используя методы линейной алгебры, можно показать (см. /9/), что для любого плана с числом точек измерения, большим $m(m+1)/2$ и известной матрицей M_0 найдется план с числом точек измерения

$$n \leq \frac{m(m+1)}{2} \quad (3.5)$$

и с той же информационной матрицей

$$M = M_0. \quad (3.6)$$

Таким образом, оптимальное число точек измерения должно быть заключено в границах

$$m \leq n \leq \frac{m(m+1)}{2}. \quad (3.7)$$

Если экспериментатора интересует k параметров из m , то аналогично /9/:

$$n \leq \frac{k}{2} (2m - k + 1). \quad (3.8)$$

Если число точек измерения совпадает с числом параметров ($n = m$), то из (3.3) вытекает, что оптимальный эксперимент должен удовлетворять условию

$$\max_{t^{(1)}, \dots, t^{(m)}=1} \prod_{i=1}^m t^{(i)} \quad \max_{x^{(1)}, \dots, x^{(m)}=1} \lambda(x^{(1)}) |F|^2 : \quad (3.9)$$

Максимизируя выражение $\prod_{i=1}^m t^{(i)}$ при условии $\sum_{i=1}^m t^{(i)} = T$, получим оптимальные времена измерения

$$t^{(i)} = \frac{T}{m} \quad (i = 1, \dots, m), \quad (3.10)$$

т.е. следует распределить время измерения равномерно по всем точкам измерения.

В настоящее время точное решение задачи планирования имеется только для случая, когда $\eta(\vec{\theta}, \vec{x})$ является алгебраическим полиномом, т.е.

$$f_{\alpha}(x) = x^{\alpha-1} \quad \alpha = 1, \dots, m. \quad (3.11)$$

В этом случае предположение о совпадении числа точек измерения с числом параметров $m = n$ оказывается верным для многих практически важных видов функций эффективности $\lambda(x)$.

Положения оптимальных точек измерения для различных случаев представлены в табл. 2, основанной на результатах работы /10/.

2. Планы, минимизирующие дисперсию линейной функции параметров (d-оптимальные планы)

Пусть $R(\vec{\theta}) = h' \vec{\theta}$ — функция, линейная комбинация параметров, интересующая экспериментатора. Удобно пользоваться обозначениями

$$h_1 = f_1(x_0), \quad h(\vec{\theta}) = \eta(\vec{\theta}, x_0). \quad (3.12)$$

Мы ввели абстрактную точку x_0 , которая становится реальной в задачах интерполяции ($x_0 \in (a, b)$) или экстраполяции ($x_0 < a$ или $x_0 > b$).

Из неравенства (3.8) для $k = 1$ вытекает, что число оптимальных точек измерения не больше числа неизвестных параметров.

Оптимальный эксперимент соответствует

$$\min_{t^{(1)}, \dots, t^{(m)}} \min_{x^{(1)}, \dots, x^{(m)}} \mathcal{D}\{h(\hat{\theta})\} \quad (3.13)$$

$$\sum t^{(i)} = T \quad x^{(i)} \in (a, b)$$

Для непосредственного вычисления (3.13) требуется минимизация $\mathcal{D}\{h(\hat{\theta})\}$ по $2m$ переменным.

Задачу определения оптимальных $t^{(1)}, \dots, t^{(m)}$ можно сравнительно просто решить при помощи интерполяционных функций Лагранжа (см. ниже).

Вычисление оптимальных точек измерения можно свести к задаче нахождения обобщенных чебышевских полиномов наилучшего приближения ^{/8/, /10/} или при малом количестве параметров пользоваться описанным ниже графическим методом ^{/11/}.

Система интерполяционных функций Лагранжа $l_i(x)$ ^{/12/}, соответствующая точкам $x^{(1)}, \dots, x^{(m)}$, определяется при помощи системы линейных уравнений:

$$\sum_{i=1}^m l_i(x) f_a(x^{(i)}) = f_a(x) \quad (3.14)$$

$$a = 1, \dots, m, \quad x \in (a, b).$$

Единственное решение системы (3.14) существует тогда и только тогда, когда

$$|F| \neq 0, \quad (3.15)$$

где

$$F_{ai} = f_a(x^{(i)}).$$

Предполагаем, что условие (3.15) выполнено для всех $x^{(1)}, \dots, x^{(m)}$ таких, что

$$a < x^{(1)} \leq \dots \leq x^{(m)} < b \quad (3.16)$$

(примером таких функций являются функции $f_a(x) = x^{a-1}$).

Умножая (3.14) на $\hat{\theta}_a$ и суммируя по a , получим

$$\sum_{i=1}^m l_i(x_0) \eta(x^{(i)}, \hat{\theta}) = h(\hat{\theta}). \quad (3.17)$$

Применим (3.17) и общую формулу (1.13) для вычисления дисперсии величины $h(\hat{\theta})$. Получим

$$\mathcal{D}\{h(\hat{\theta})\} = \sum_{i=1}^m l_i^2(x_0) \mathcal{D}\{\eta(x^{(i)}, \hat{\theta})\} = \sum_{i=1}^m \frac{l_i^2(x_0)}{\lambda(x^{(i)}) t^{(i)}}. \quad (3.18)$$

Минимизацией (3.18) по $t^{(i)}$ при условии $\sum t^{(i)} = T$ получим, что оптимальные времена измерения пропорциональны

$$t^{(i)} = \frac{|l_i(x_0)|}{\sqrt{\lambda(x^{(i)})}}. \quad (3.19)$$

При таком подборе времен измерения дисперсия $\mathcal{D}\{h(\hat{\theta})\}$ величины $h(\hat{\theta})$ равна

$$\sqrt{\mathcal{D}\{h(\hat{\theta})\}} = \frac{1}{\sqrt{T}} \sum_{i=1}^m \frac{|l_i(x_0)|}{\sqrt{\lambda(x^{(i)})}} \quad (3.20)$$

и зависит еще от выбора точек измерения $x^{(1)}, \dots, x^{(m)}$. Покажем метод вычисления этих точек в случае экстраполяции ($x_0 \leq a$ или $x_0 \geq b$).

Выражение (3.20) является значением величины

$$s(x) = \frac{1}{\sqrt{T}} \sum_{i=1}^m \frac{l_i(x)}{\sqrt{\lambda(x^{(i)})}} \text{sign}[l_i(x_0)], \quad (3.21)$$

вычисленным в точке $x = x_0$.

Используем соотношения

$$l_i(x^{(k)}) = \delta_{ik} = \begin{cases} 1 & \text{если } i = k \\ 0 & \text{если } i \neq k \end{cases}$$

вытекающие из (3.14), чтобы получить значения функций $s(x)$ в точках измерения:

$$|s(x^{(k)})| = \frac{1}{\sqrt{T \lambda(x^{(k)})}} \quad k = 1, \dots, m, \quad (3.22)$$

т.е. в точках измерения кривая $\sqrt{T} s(x)$ обязательно пересекается с кривыми $+\frac{1}{\sqrt{\lambda(x)}}$ и $-\frac{1}{\sqrt{\lambda(x)}}$ или касается их.

Очевидно, что дисперсия

$$D\{h(\theta)\} = \sigma^2(x_0)$$

будет минимальной, если и "разброс" кривой $s(x)$ на интервале $a \leq x \leq b$ будет минимальным (рис. 2), т.е. если кривую $\sqrt{T} s(x)$ удастся вписать в коридор $(-1/\sqrt{\lambda(x)}, 1/\sqrt{\lambda(x)})$ (полное доказательство см. /11/). Из (3.14) вытекает, что функции $f_a(x)$ являются линейными комбинациями функций $f_a(x)$. Поэтому (см. (3.21)) функцию $\sqrt{T} s(x)$ можно также представить в виде линейной комбинации функций $f_a(x)$

$$\sqrt{T} s(x) = \sum_{a=1}^m a_a f_a(x). \quad (3.23)$$

Таким образом, для определения оптимальных экспериментальных точек нужно подобрать a_1, \dots, a_m так, чтобы $\sum_a a_a f_a(x)$ касалась кривых $1/\sqrt{\lambda(x)}$ и $-1/\sqrt{\lambda(x)}$ в интервале $x \in (a, b)$. Абсциссы точек соприкосновения - это оптимальные точки измерения.

Когда кривую вида (3.23) нельзя вписать в указанный коридор, тогда оптимальный эксперимент является вырожденным, т.е. число точек измерения меньше числа параметров. В этом случае кривую (3.23) "построим" так, чтобы она немного выходила из указанного коридора. Абсциссы ее точек пересечения с границами коридора - это приближенно оптимальные точки измерения.

Если вид функций $f_a(x)$ достаточно прост и число параметров мало ($m \leq 6$), то можно описанное решение получить графически /11/ (см. рис. 2).

Пример 3.1 /11/

Известно, что дифференциальное эффективное сечение $\eta(x)$ (см. пример 1.2) нельзя измерять на угле $x_0 = 0$. Поэтому оцениваем $\eta(0)$ косвенным образом, через измерения на углах, отличных от нуля (экстраполяция).

Предположим, что на некотором интервале $0 \leq x \leq b$ дифференциальное сечение $\eta(x)$ не изменяется быстрее, чем многочлен второй степени:

$$\eta(x) = \eta(\theta, x) = \theta_1 + \theta_2 x + \theta_3 x^2, \quad (3.24)$$

где $\theta_1, \theta_2, \theta_3$ какие-то (пока неизвестные) коэффициенты. Нас интересует оценка $\hat{\eta}(0) = \hat{\theta}_1$. Требуем эксперимент, определяющий $\hat{\eta}(0)$ с минимальной дисперсией (d - оптимальность).

Функция эффективности $\lambda(x)$ в данной задаче пропорциональна неизвестному дифференциальному сечению (см. (1.19))

$$\lambda(x) \approx \eta(x).$$

Поэтому нужно приблизительно оценить $\eta(x)$ (из теории или на основе коротко-временных измерений).

На рисунке 2 (взятом из /11/) выше описанным графическим методом определены оптимальные точки измерения $x^{(1)}, x^{(2)}, x^{(3)}$

для случая $\lambda(x) = x$.

Времена измерения $t^{(i)}$ ($i = 1, 2, 3$) нужно вычислить по формуле (3.19). При этом пользуемся следующим выражением для интерполяционных полиномов Лежандра /12/:

$$l_i(x) = \prod_{\substack{k=1 \\ k \neq i}}^3 \frac{x - x^{(k)}}{x^{(i)} - x^{(k)}} \quad (3.25)$$

(аналогичное выражение верно всегда, когда $f_a(x) = x^a$), Получим

$$t^{(1)} : t^{(2)} : t^{(3)} = 0.84 : 0.12 : 0.04.$$

Пример 3.2.

Увеличение числа точек измерения и не оптимальное размещение времени измерения по этим точкам может привести к значительному ухудшению точности результатов эксперимента. На рис. 3 показана зависимость дисперсии величины

$$h(\vec{\theta}) = \theta_1 - \theta_2 + \theta_3$$

от числа точек измерения n в следующих экспериментах:

$$\eta(\vec{\theta}, x) = \theta_1 + \theta_2 x + \theta_3 x^2, \quad x \in \langle 0, 1 \rangle$$

$$\lambda(x) = \text{const}$$

с симметричными планами:

$$x^{(i)} = \frac{i-1}{n-1} \quad (i = 1, \dots, n)$$

$$t^{(i)} = \frac{T}{n}$$

Оптимальный план соответствует:

$$n = 3$$

$$x^{(1)} = 0, \quad x^{(2)} = 0.5, \quad x^{(3)} = 1,$$

$$t^{(1)}, t^{(2)}, t^{(3)} = 0.35 : 0.47 : 0.18.$$

II. Несколько контролируемых переменных

Планирование оптимальных экспериментов в многомерном случае (несколько контролируемых переменных) является технически значительно более трудной задачей, чем в одномерном. Однако кроме технических затруднений (например, резкое увеличение числа переменных, по которым необходимо производить поиск экстремума) возникают также и трудности принципиального характера. Так, например, для некоторых классов функций $\eta(\vec{\theta}, \vec{x})$ для заданного эксперимента

может оказаться бесконечно много D -оптимальных планов. Это практически исключает возможность применения в этих случаях приближенных методов расчёта (например, какого-либо градиентного метода поиска экстремума или метода последовательного планирования (см. § 4 данного обзора)), так как эти методы могут сходиться либо к планам с очень большим числом точек измерения (много большим числа неизвестных параметров), либо вообще не сойтись.

В настоящее время планы, оптимальные по минимаксному критерию (и тем самым и по критерию D -оптимальности; см. табл. 11), найдены лишь для простейших функций $\eta(\vec{\theta}, \vec{x})$, а именно, когда $\eta(\vec{\theta}, \vec{x})$ является линейной или квадратичной функцией контролируемых переменных

$$\eta(\vec{\theta}, \vec{x}) = \theta_1 + \theta_2 x_1 + \dots + \theta_{s+1} x_s \quad (3.26)$$

или

$$\eta(\vec{\theta}, \vec{x}) = \theta_1 + \sum_{j=1}^s \theta_{j+1} x_j + \sum_{j=1}^s \theta_{j+s+1} x_j^2 + \quad (3.27)$$

$$+ \sum_{j=1}^s \sum_{k=1}^s \theta_{2s+j+k-1} x_j x_k \quad k > j$$

Случай (3.26) достаточно тривиален и подробно изложен в книге /13/.

Рассмотрим случай (3.27) (см. /8, 19/) на примере. Будем предполагать, что $-1 \leq x_i \leq 1, i = 1, \dots, s$ (этого всегда можно достичь соответствующим изменением масштаба по осям x_i) и что $\lambda(\vec{x}) \equiv 1$ (это весьма строгое ограничение).

Как видно из (3.27), число искомых параметров равно

$$m = \frac{(s+1)(s+2)}{2}, \quad (3.28)$$

где s - число контролируемых переменных.

Воспользуемся следующим примером. Из всего множества планов мы будем рассматривать планы следующего вида: измерения с весами $w_1 = \alpha$ прово-

дятся в каждой из 2^s вершин s -мерного куба, с весами $w_i = \beta$ в каждой из $s 2^{s-1}$ точек, которые являются серединами ребер и с весами $w_i = \gamma$ в каждой из $s(s-1) 2^{s-3}$ центров двумерных граней.

Оптимальный минимаксный план должен удовлетворять условию //10/:

$$\max_{x_i \in \langle -1, 1 \rangle} \lambda(\vec{x}) \hat{\sigma}^2(\vec{x}) = m. \quad (3.29)$$

$i = 1, \dots, s$

Поэтому в нашем случае надо подобрать α, β, γ так, чтобы определитель информационной матрицы $|M|$ был максимальным при условии $\Sigma \alpha + \Sigma \beta + \Sigma \gamma = 1$ и для полученных весов α, β, γ проверить выполнение равенства (3.29).

Получим следующие результаты для $s \leq 5$

s	α	β	γ
1	0,333	0,333	0,000
2	0,1458	0,8015	0,962
3	0,071975	0,01895	0,03280
4	0,03705	0,0038375	0,01185
5	0,01928	0,0003125	0,004475

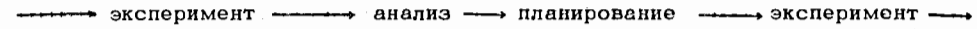
Для $s \geq 6$ нельзя получить оптимальный план с положительными α, β, γ .

84. Последовательное и непрерывное планирование

Если зависимость измеренной величины от искомых параметров (функция $\eta(\vec{\theta}, \vec{x})$) нелинейна или условия эксперимента (например, функция эффективности) изменяются в течение эксперимента, то изложенные выше методы планирования неприменимы. В этих случаях естественным и удобным является после-

довательное (или непрерывное) планирование. Такое планирование полезно и в случае, когда статическое планирование возможно, но трудно выполнимо с точки зрения математических расчётов. Тогда последовательное планирование является по сути дела некоторым приближенным методом расчёта статического планирования //14/.

Основная идея последовательного планирования определяется схемой:



Планирование определяет оптимальную точку измерения для каждого нового опыта, используя результаты предыдущих экспериментов. Время измерения T в новом эксперименте считается заданным. При $T \rightarrow 0$, стремящемся к нулю, последовательное планирование переходит в непрерывное планирование.

Если нелинейность функции $\eta(\vec{\theta}, \vec{x})$ существенна и набранная статистика из предыдущих экспериментов мала, то оценки параметров могут оказаться неоднозначными. Планирование для таких случаев обсуждается в §§ 5 и 6.

Если набор оценок параметров $\hat{\theta}^{(0)}$, полученных на основе предыдущих экспериментов, единственен, то можно воспользоваться линейным приближением (см. также § 1):

$$\eta(\vec{\theta}, \vec{x}) = \eta(\hat{\theta}^{(0)}, \vec{x}) + \vec{f}'(\vec{x}) [\vec{\theta} - \hat{\theta}^{(0)}], \quad (4.1)$$

где

$$f'_a(\vec{x}) = \frac{\partial \eta(\vec{\theta}, \vec{x})}{\partial \theta_a} \Big|_{\vec{\theta} = \hat{\theta}^{(0)}}. \quad (4.2)$$

Обозначим $D(0)$ матрицу ошибок, вычисленную на основе предыдущих экспериментов. После измерения в течение T в точке \vec{x} матрица ошибок примет значение $D(T, \vec{x})$, которое связано с $D(0)$ следующей формулой (см. приложение в //15/):

$$D(T, \vec{x}) = \left(I_m - \frac{\lambda(\vec{x}) T D(0) \vec{f}'(\vec{x}) \vec{f}'(\vec{x})}{1 + \lambda(\vec{x}) T \hat{\sigma}_0^2(\vec{x})} \right) D(0), \quad (4.3)$$

где $\hat{\sigma}_0^2(\vec{x})$ - квадрат коридора ошибок к моменту $T=0$ и I_m - единичная матрица $m \times m$. Данная формула является основой для всех расчетов при последовательном и непрерывном планировании.

Из (4.3) нетрудно получить, что

$$|D(T, \vec{x})| = \frac{|D(0)|}{1 + \lambda(\vec{x}) T \hat{\sigma}_0^2(\vec{x})} \quad (4.4)$$

для определителя матрицы ошибок и

$$\mathcal{D}\{\hat{h}; T\} = \hat{h}' D(T, \vec{x}) \hat{h} = \mathcal{D}\{\hat{h}; 0\} - \frac{[\hat{h}' D(0) \hat{f}(\vec{x})]^2}{[\lambda(\vec{x}) T]^{-1} + \hat{\sigma}_0^2(\vec{x})} \quad (4.5)$$

для дисперсии оценки величины $h(\theta) = \hat{h}' \vec{\theta}$.

Аналогичные формулы можно вывести для случая, когда основной целью эксперимента является определение k из m параметров: $(\theta_1, \dots, \theta_k)' = \vec{\omega}$ (см. /15/). Из этих формул и из (4.4) и (4.5) непосредственно следуют результаты, представленные в таблице 3.

Заметим, что в принципе можно спланировать эксперимент так, чтобы в течение данного времени T измерения производились в нескольких точках, число которых ранее задано /16/. Такое планирование требует довольно сложных вычислений.

Интересно сравнить экономию затрат (времени) при непрерывном (последовательном) планировании и при статическом планировании в задачах, допускающих статическое планирование.

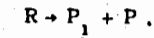
Если нет первоначальных пассивных экспериментов, то затраты на эксперимент при непрерывном и статическом планировании одни и те же. Если часть первоначальных экспериментов проводилась пассивно (без планирования), то непрерывное планирование сначала более экономично, чем статическое, но при большом общем времени измерения статическое планирование, учитывающее результаты непланированных экспериментов, немножко лучше непрерывного.

Для определения оптимальной точки измерения \vec{x}_0 необходимо найти максимум некоторой функции $\epsilon(\vec{x})$ (см. таблицу 3). Если существует не больше трех контролируемых переменных (\vec{x} - трехмерный вектор), то

\vec{x}_0 можно найти простым перебором. Полезно вычислить и значения $\epsilon(\vec{x})$ в окрестности \vec{x}_0 , чтобы оценить влияние малого смещения точки \vec{x}_0 на значение $\epsilon(\vec{x})$ (устойчивость решения). Если размерность \vec{x} велика, то необходимо прибегнуть к более эффективным методам поиска экстремума (см., например, /12/). Оценить устойчивость решения можно на основе вторых производных функций $\epsilon(\vec{x})$ в точке \vec{x}_0 (приближение $\epsilon(\vec{x})$ в окрестности \vec{x}_0 многомерным параболоидом).

Пример 4.1.

Смоделируем эксперимент, который покажет, насколько эффективно последовательное планирование. Пусть происходит химическая реакция типа



Скорость η этой реакции равна

$$\eta = \frac{\theta_1 \theta_3 x_1}{1 + \theta_1 x_1 + \theta_2 x_2} \quad (4.6)$$

где x_1 - парциальное давление исходного продукта R ;

x_2 - парциальное давление продукта P ;

$\theta_1, \theta_2, \theta_3$ - параметры, которые необходимо определить.

Для моделирования воспользуемся известными из литературы истинными значениями параметров:

$$\theta_1 = 2,9, \quad \theta_2 = 12,2, \quad \theta_3 = 0,69. \quad (4.7)$$

В дальнейшем мы будем считать, что измерения возможны в области

$$0 \leq x_1 \leq 3, \quad 0 \leq x_2 \leq 3. \quad (4.8)$$

Измерения η при любых x_1 и x_2 из указанной области будем моделировать следующим образом. Производится выборка из таблицы случайных чисел (нормальное распределение с параметрами $0, \sigma = 0,01$). Полученное число прибавляется к $\eta(\vec{\theta}, \vec{x})$, где $\vec{\theta}$ определяется равенствами (4.7).

Нетрудно видеть, что эффективность $\lambda(\vec{x})$ будет при этом постоянна в области (4.8) и вес измерения, проведенного в любой точке $\vec{x}^{(1)}$, равен

$$w_1 = \frac{N_1}{\sigma^2},$$

где N_1 - число отдельных измерений в точке $\vec{x}^{(1)}$. Первоначальные оценки параметров $\hat{\theta}_1^{(0)}, \hat{\theta}_2^{(0)}, \hat{\theta}_3^{(0)}$ были получены методом наименьших квадратов из первых четырех экспериментов, описанных в таблице 4.

Предполагая, что оценки $\hat{\theta}^{(0)}$ не сильно отличаются от истинных, воспользуемся развитым методом. Вычислим:

$$f_a(\vec{x}) = \frac{\partial}{\partial \theta_a} \left[\frac{\theta_3 \theta_1 x_1}{1 + \theta_1 x_1 + \theta_2 x_2} \right]_{\vec{\theta} = \hat{\theta}^{(0)}} \quad (4.9)$$

Согласно (1.12), после дополнительного (пятого) измерения информационная матрица Фишера будет иметь вид:

$$M = \frac{1}{\sigma^2} \begin{pmatrix} \sum_{i=1}^5 f_1^2(\vec{x}^{(1)}), & \sum_{i=1}^5 f_1(\vec{x}^{(1)})f_2(\vec{x}^{(1)}), & \sum_{i=1}^5 f_1(\vec{x}^{(1)})f_3(\vec{x}^{(1)}) \\ \sum_{i=1}^5 f_1(\vec{x}^{(1)})f_2(\vec{x}^{(1)}), & \sum_{i=1}^5 f_2^2(\vec{x}^{(1)}), & \sum_{i=1}^5 f_2(\vec{x}^{(1)})f_3(\vec{x}^{(1)}) \\ \sum_{i=1}^5 f_1(\vec{x}^{(1)})f_3(\vec{x}^{(1)}), & \sum_{i=1}^5 f_2(\vec{x}^{(1)})f_3(\vec{x}^{(1)}), & \sum_{i=1}^5 f_3^2(\vec{x}^{(1)}) \end{pmatrix} \quad (4.10)$$

При составлении (4.10) учтено, что на каждом числе проводится лишь одно измерение и что все измерения равнозначны, т.е. $N_1=1, \sigma_1=\sigma$ ($1=1,2,\dots,p$).

Подставим из табл. 4 в (4.10) значения $\vec{x}^{(1)}$ ($1=1,2,3,4$) и воспользовавшись (4.4), получим, что следующее измерение следует проводить в точке $\vec{x}^{(5)}$ с координатами

$$x_1^{(5)} = 0,1; \quad x_2^{(5)} = 0,0. \quad (4.11)$$

При этом поиск максимума определителя информационной матрицы производится методом сетки.

Эксперимент, "проведенный" в точке $\vec{x}^{(5)}$, дал результат $y_5 = 0,186$. Включая пятое измерение в метод наименьших квадратов, получим новые оценки параметров

$$\hat{\theta}_1^{(1)} = 3,11, \quad \hat{\theta}_2^{(1)} = 15,19, \quad \hat{\theta}_3^{(1)} = 0,73. \quad (4.12)$$

По формуле (4.9) подсчитаем производные при $\vec{\theta} = \hat{\theta}^{(0)}$, составим матрицу M , отыщем максимум ее определителя (максимум достигается при $x_1 = 3,0$ и $x_2 = 0,0$). Вслед за этим проводим измерение в точке $\vec{x}^{(6)}$, находим оценки $\hat{\theta}^{(2)}$ и т.д. Всего аналогичным образом было спланировано еще семь последующих измерений. Результаты планирования и опытов представлены в таблице 4.

Моделирование эксперимента было прекращено, как только разброс оценок параметров стал незначителен и значения оценок стали близки к теоретическим значениям параметров $\theta_1, \theta_2, \theta_3$. На практике подобные эксперименты обычно прекращаются, если определитель $|D|$ становится меньше некоторого, наперед заданного числа.

Обсудим результаты проведенного планирования. Нетрудно видеть, что значения $\vec{x}^{(1)}$ ($1=5,\dots,13$) сосредоточены вблизи трех точек А,В,С (см.рис.4) (для наглядности на этом же рисунке проведены линии постоянных значений $\eta(\vec{x}, \vec{\theta})$). Эти точки должны соответствовать плану, близкому к

D -оптимальному при нулевой начальной информации ($M=0$). При этом число измерений N_1, N_2, N_3 ($N_1 + N_2 + N_3 = N$), которое необходимо провести в каждой точке, должно быть пропорционально частоте появления этой точки в последовательном планировании. В данном случае число точек плана и число параметров равны, поэтому согласно (3.10), $N_1 = N_2 = N_3$.

Сравним полученный приближенно D -оптимальный план с симметричными планами, которые очень часто употребляются на практике. Эти планы и соответствующие определители матриц ошибок представлены в таблице 5.

В экспериментальной практике нередко возникают ситуации, когда на некоторых этапах экспериментальной работы появляются две или более гипотезы о зависимости неизвестных параметров от измеренных величин. Например, при фазовом анализе рассеяния функция правдоподобия, как правило, имеет несколько равноценных максимумов, каждому из которых соответствует свой набор оценок параметров. Возникает необходимость в планировании экспериментов, которые за заданное время позволили бы с наибольшей достоверностью выделить правильную гипотезу.

Предположим, что из теоретических исследований или из анализа предварительных экспериментов следует, что изучаемый процесс может быть описан одной из двух функций:

$$\eta_1(\vec{x}) = \eta_1(\vec{\theta}, \vec{x}) \quad (\text{гипотеза I}) \quad \text{и} \quad (5.1)$$

$$\eta_2(\vec{x}) = \eta_2(\vec{\phi}, \vec{x}) \quad (\text{гипотеза II}).$$

Через $\vec{\theta}$ и $\vec{\phi}$ обозначены неизвестные параметры. В частном случае неоднозначного анализа экспериментальных данных

$$\eta_1(\vec{x}) = \eta_1(\vec{\hat{\theta}}_1, \vec{x}) + \sum_{i=1}^n \frac{\partial \eta_1(\vec{\theta}, \vec{x})}{\partial \theta_i} \Big|_{\vec{\theta} = \vec{\hat{\theta}}_1} (\theta_i - \hat{\theta}_{1,i}),$$

где $\vec{\hat{\theta}}_1$ и $\vec{\hat{\theta}}_2$ соответствуют двум максимумам функции правдоподобия ($\vec{\hat{\theta}}_1 \leftrightarrow \vec{\theta}, \vec{\hat{\theta}}_2 \leftrightarrow \vec{\phi}$).

Пусть в результате измерений в точках $\vec{x}^{(1)}, \dots, \vec{x}^{(n)}$ для η получены значения y_1, \dots, y_n с дисперсиями $\sigma_1^2, \dots, \sigma_n^2$ и пусть набранной статистики недостаточно, чтобы из двух конкурирующих гипотез отбросить неверную с заданной достоверностью.

Это означает, что суммы

$$\chi_1^2(n) = \sum_{i=1}^n \frac{1}{\sigma_i^2} [y_i - \eta_1(\vec{\hat{\theta}}_1^{(n)}, \vec{x}^{(i)})]^2$$

и

$$\chi_2^2(n) = \sum_{i=1}^n \frac{1}{\sigma_i^2} [y_i - \eta_2(\vec{\hat{\phi}}_1^{(n)}, \vec{x}^{(i)})]^2 \quad (5.2)$$

несущественно отличаются друг от друга. Здесь $\vec{\hat{\theta}}^{(n)}$ и $\vec{\hat{\phi}}^{(n)}$ — оценка параметров $\vec{\theta}$ и $\vec{\phi}$ после n измерений. Задача экспериментатора заключается в том, чтобы отыскать такую точку \vec{x} , в результате измерения в которой разница

$$\chi_j^2(n+1) - \chi_k^2(n+1) \quad j, k = 1, 2, \quad j \neq k \quad (5.3)$$

(где k — индекс, соответствующий верной гипотезе), достигала бы максимальной величины.

Рассмотрим сумму:

$$\chi_2^2(n+1) = \sum_{i=1}^{n+1} \frac{1}{\sigma_i^2} [y_i - \eta_2(\vec{\hat{\phi}}^{(n+1)}, \vec{x}^{(i)})]^2 \quad (5.4)$$

Значение $\chi_2^2(n+1)$ существенно зависит от результата $(n+1)$ -го измерения. Заранее величина y_{n+1} неизвестна, поэтому точнее значение разницы (5.3) определить до опыта невозможно. Однако мы можем вычислить ее среднее значение.

В предположении верности k -ой гипотезы случайная величина

$$u_k = \frac{y - \hat{\eta}_k(\vec{x})}{\sqrt{\hat{\sigma}^2(\vec{x}) + \sigma^2(\vec{x})}} \quad (5.5)$$

(где

$$\hat{\eta}_1(\vec{x}) = \eta_1(\vec{\hat{\theta}}^{(n)}, \vec{x}), \quad \hat{\eta}_2(\vec{x}) = \eta_2(\vec{\hat{\phi}}^{(n)}, \vec{x}), \quad \hat{\sigma}_k^2(\vec{x}) -$$

коридор ошибок для k -ой поверхности отклика после n измерений, $\sigma^2(\vec{x})$ — дисперсия дополнительного измерения y), распределена по нормальному закону с параметрами $E\{u_k\} = 0, D\{u_k\} = 1$. В (5.5) для краткости y величины y и y остальных величин индекс $n+1$ опущен.

Пусть верна первая гипотеза. Тогда, используя формулу (4.3), можно показать /15/, что

$$E_1 \{ \chi_2^2(n+1) - \chi_1^2(n+1) \} = \chi_2^2(n) - \chi_1^2(n) + \quad (5.6)$$

$$+ \frac{[\hat{\eta}_1(\vec{x}) - \hat{\eta}_2(\vec{x})]^2 + [\hat{\sigma}_2^2(\vec{x}) - \hat{\sigma}_1^2(\vec{x})]}{s_2^2(\vec{x})},$$

$$s_k^2(\vec{x}) = \hat{\sigma}_k^2(\vec{x}) + \sigma^2(\vec{x}).$$

где

Очевидно, что наиболее выгодным будет измерение в той точке \vec{x} , где среднее значение разницы суммы взвешенных квадратичных отклонений достигает наибольшего значения. Планирование дискриминирующего эксперимента сводится, таким образом, к отысканию

$$\max_{\vec{x}} \frac{[\hat{\eta}_1(\vec{x}) - \hat{\eta}_2(\vec{x})]^2 + \hat{\sigma}_2^2(\vec{x}) - \hat{\sigma}_1^2(\vec{x})}{s_2^2(\vec{x})}. \quad (5.7)$$

В предположении верности второй гипотезы измерение следует проводить в точке, соответствующей

$$\max_{\vec{x}} \frac{[\hat{\eta}_2(\vec{x}) - \hat{\eta}_1(\vec{x})]^2 + \hat{\sigma}_1^2(\vec{x}) - \hat{\sigma}_2^2(\vec{x})}{s_1^2(\vec{x})}. \quad (5.8)$$

Если точки, определенные из (5.7) и (5.8), не совпали, то измерение следует проводить в точке, соответствующей

$$\max_{\vec{x}} \{ W_1 E_1 \{ \chi_2^2(n+1) - \chi_1^2(n+1) \} + W_2 E_2 \{ \chi_1^2(n+1) - \chi_2^2(n+1) \} \}, \quad (5.9)$$

где веса W_1 и W_2 , вообще говоря, зависят от целей эксперимента. Если потери от принятия неверной гипотезы и отбрасывания верной можно считать равными, то веса определяются следующим образом:

$$W_1 \approx e^{-\frac{1}{2} \chi_1^2(n)} \quad W_2 \approx e^{-\frac{1}{2} \chi_2^2(n)}. \quad (5.10)$$

В некоторых случаях цель проводимых экспериментов не определена в таком виде, который позволил бы найти отношение потерь от неправильного принятия первой гипотезы к потерям от неправильного принятия второй гипотезы. В подобной ситуации измерение следует проводить в точке, соответствующей

$$\max_{\vec{x}} \min_{\substack{k=1,2 \\ j \neq k}} E_k \{ \chi_j^2(n+1) - \chi_k^2(n+1) \}. \quad (5.11)$$

Пример 5.1. Рассеяние поляризованных протонов (см. пр. 1.1 и 2.1)

Обсудим следующий эксперимент (рис. 5). Пучок поляризованных протонов рассеивается на неполяризованной нейтронной мишени T_1 . Вследствие этого рассеяния вектор поляризации \vec{P}_1 поворачивается. Этот поворот можно описать пятью параметрами Вольфенштейна $R(x)$, $R'(x)$, $A(x)$, $A'(x)$ и $D(x)$ (см. ^{14/}). Здесь x - угол рассеяния на мишени T_1 . Если подходящим образом выбрать угол между вектором \vec{n} (нормалью на плоскости первого рассеяния) и вектором поляризации первичного пучка \vec{P}_1 , то можно определить один из параметров Вольфенштейна, например, $R(x)$ при помощи второго рассеяния на T_2 .

Экспериментатору, подготавливающему эксперимент, обыкновенно известны оценки $\hat{\theta}^{(0)}$ фазовых сдвигов, уже полученные из других экспериментов, и он ограничен временем T , которое предоставлено ему для измерений на ускорителе. В ^{17/} описан случай, когда существуют два набора фазовых сдвигов $\hat{\theta}^{(1)}$ и $\hat{\theta}^{(2)}$ и целью некоторого эксперимента является дискриминирование неправильного набора. Вследствие неоднозначности оценок $\hat{\theta}$ возможны две оценки кривой $R(x)$ ($R(\hat{\theta}^{(1)}, x)$ и $R(\hat{\theta}^{(2)}, x)$; см. рис. 7) или $D(x)$ ($D(\hat{\theta}^{(1)}, x)$ и $D(\hat{\theta}^{(2)}, x)$; см. рис. 9).

Функция эффективности $\lambda(x)$ может быть вычислена на основе результатов предыдущих экспериментов (например, методом, описанным в ^{17/}; результат приведен на рис. 8).

Чтобы сравнить полученные результаты с результатами примера 6.1, будем планировать оптимальный эксперимент для случая, когда $T \rightarrow 0$ (непрерывное планирование дискриминирующего эксперимента). Планирование тогда состоит в вычислении функций:

$$\phi(x) = \lambda(x) \{ [\hat{\eta}_1(x) - \hat{\eta}_2(x)]^2 + \hat{\sigma}_2^2(x) - \hat{\sigma}_1^2(x) \} \quad (5.12)$$

(полученной из (5.7) для $T \rightarrow 0$)

и

$$\psi(x) = \lambda(x) \{ [\hat{\eta}_1(x) - \hat{\eta}_2(x)]^2 + \hat{\sigma}_1^2(x) - \hat{\sigma}_2^2(x) \} \quad (5.13)$$

(полученной из (5.8))

и определении максимума этих функций (здесь $\hat{\eta}_1(x) = R(\hat{\theta}^{(1)}, x)$ или $D(\hat{\theta}^{(1)}, x)$ и т.д.).

Результаты вычисления $\phi(x)$ и $\psi(x)$ приведены на рис. 8 (для $R(x)$) и на рис. 10 (для $D(x)$). Оказывается, что на обоих рисунках $\phi(x)$ и $\psi(x)$ почти совпадают в окрестности своих максимумов, т.е. применение формул (5.9) и (5.11) эквивалентно. Если сравнить эксперимент по измерению $R(x)$ с экспериментом по измерению $D(x)$, то измерение $D(x)$ более выгодно (главным образом, благодаря большому расстоянию между кривыми $D(\hat{\theta}^{(1)}, x)$ и $D(\hat{\theta}^{(2)}, x)$).

§6. Информационный метод планирования экспериментов

Если функция $\eta(\vec{\theta}, \vec{x})$ сильно нелинейна по параметрам $\vec{\theta}$ и набранная статистика мала, то возможно существование нескольких максимумов функции правдоподобия $L(\vec{\theta})$, слабо отличающихся по высоте. Принято считать положения этих максимумов наборами оценок параметров $\vec{\theta}$, полученными по методу максимума правдоподобия. Эксперименты, спланированные по методам, описанным в предыдущем параграфе, позволяют наиболее эффективно исключить неправильные наборы оценок. При этом экспериментатор не интересуется точностью оценок параметров (т.е. матрицей ошибок). При таком подходе приходится вначале делать различающие эксперименты (§5) и только после достоверного различения перейти на планирование уточняющих экспериментов (§4). Очевидно, наиболее перспективными критериями оптимальности являются те, которые позволяют отыскивать эксперименты, одновременно различающие и уточняющие наборы оценок.

Таким свойством обладает критерий, основанный на энтропийной мере информации^{/18/}. Количество информации, содержащееся в экспериментальных данных, равно

$$I_{\vec{\theta}} = \int L(\vec{\theta}) \ln L(\vec{\theta}) d\vec{\theta}, \quad (6.1)$$

где $L(\vec{\theta})$ — функция правдоподобия (см. §1). Функция $L(\vec{\theta})$ должна быть нормирована, т.е.

$$\int L(\vec{\theta}) d\vec{\theta} = 1, \quad (6.2)$$

а интегралы (6.1) и (6.2) взяты по всей области возможных значений параметров.

Аналогично, если экспериментатора интересует k из m параметров (см. §4), то соответствующее количество информации об этих полезных параметрах $\vec{\omega}$ равно:

$$I_{\vec{\omega}} = \int L(\vec{\omega}) \ln L(\vec{\omega}) d\vec{\omega}. \quad (6.3)$$

Здесь

$$L(\vec{\omega}) = \int L(\vec{\theta}) d\vec{\phi}, \quad (6.4)$$

где $\vec{\phi}$ — это мешающие параметры, $\vec{\phi} = (\theta_{k+1}, \dots, \theta_m)$.

Мера количества информации (6.1) и (6.3) по многим свойствам аналогична энтропии физической системы. Нетрудно показать^{/9/}, что в случае нормального распределения экспериментальных данных и существования единственного набора оценок параметров

$$I_{\vec{\theta}} = -\frac{1}{2} \ln |D| - \frac{1}{2} [m + \ln(2\pi)^m] \quad (6.5)$$

$$I_{\vec{\omega}} = -\frac{1}{2} \ln |D(\vec{\omega})| - \frac{1}{2} [k + \ln(2\pi)^k], \quad (6.6)$$

где D и $D(\vec{\omega})$ - матрицы ошибок, соответствующие параметрам $\vec{\theta}$ и $\vec{\omega}$. Из (6.5) и (6.6) вытекает, что в этом случае критерий D - оптимальности (82) и информационный критерий совпадают.

Планирование по информационному критерию соответствует выбору эксперимента, дающего максимальный прирост информации. Этот прирост зависит от результата планируемого эксперимента, не известного априори. Поэтому вычисляем средний прирост информации

$$\Delta J_{\vec{\theta}}(T, \vec{x}) = \int p(y) \{ \int L(\vec{\theta}/y) \ln L(\vec{\theta}/y) d\vec{\theta} - \int L(\vec{\theta}) \ln L(\vec{\theta}) d\vec{\theta} \} dy. \quad (6.7)$$

Здесь T - время, предназначенное для измерения, \vec{x} - точка измерения (значения контролируемых переменных) $L(\vec{\theta}/y)$ - функция правдоподобия после проведения планируемого эксперимента с результатом y ,

$$p(y) = \int p(y/\vec{\theta}) L(\vec{\theta}) d\vec{\theta} \quad (6.8)$$

и $p(y/\vec{\theta})$ - распределение вероятности результатов планируемого эксперимента. В большинстве физических экспериментов можно предполагать, что

$$p(y/\vec{\theta}) = \frac{\sqrt{\lambda(x) T}}{\sqrt{2\pi}} e^{-\frac{1}{2} [y - \eta(\vec{\theta}, \vec{x})]^2 \lambda(x) T}$$

(см. также §1).

Оптимальную точку измерения \vec{x}_0 следует подобрать так, чтобы

$$\Delta J_{\vec{\theta}}(T, \vec{x}_0) = \max_{\vec{x}} \Delta J_{\vec{\theta}}(T, \vec{x}) \quad (6.10)$$

(последовательное планирование)

или

$$\left. \frac{\partial J_{\vec{\theta}}(t, \vec{x}_0)}{\partial t} \right|_{t=0} = \max_{\vec{x}} \left. \frac{\partial J_{\vec{\theta}}(t, \vec{x})}{\partial t} \right|_{t=0} \quad (6.11)$$

(непрерывное планирование).

В^{19/} выведен следующий результат:

$$\begin{aligned} \left. \frac{\partial J_{\vec{\theta}}(t, \vec{x})}{\partial t} \right|_{t=0} &= \frac{\lambda(\vec{x})}{2} \int [\eta(\vec{\theta}, \vec{x}) - \int \eta(\vec{\theta}, \vec{x}) L(\vec{\theta}) d\vec{\theta}]^2 L(\vec{\theta}) d\vec{\theta} = \\ &= \frac{\lambda(\vec{x})}{2} \int [\eta - \int L_{\eta}(\eta) \eta d\eta]^2 L_{\eta}(\eta) d\eta. \end{aligned} \quad (6.12)$$

Здесь $L_{\eta}(\eta)$ - это функция правдоподобия, соответствующая измеряемой величине η :

$$L_{\eta}(\eta) = \int L(\vec{\theta}) \delta(\eta - \eta(\vec{\theta}, \vec{x})) d\vec{\theta}, \quad (6.13)$$

где $\delta(\cdot)$ - δ - функция Дирака.

Для расчёта (6.12) можно пользоваться следующим приближением:

$$L(\vec{\theta}) = \sum_{i=1}^N W_i p_i(\vec{\theta}), \quad (6.14)$$

где N - число наборов оценок $\vec{\theta}_i$,

$$W_i = L(\vec{\theta}_i^{(1)}) |D^{(1)}|^{1/2} \quad (6.15)$$

и $p_i(\vec{\theta})$ - мерное нормальное распределение со средним, равным $\vec{\theta}_i^{(1)}$ -тому набору оценок $\vec{\theta}_i^{(1)}$ и с ковариационной матрицей, равной матрице ошибок $D^{(1)}$. Подставляя (6.14) в (6.13), получим

$$L_{\eta}(\eta) = \sum_{i=1}^N W_i p_i(\eta), \quad (6.16)$$

где

$$p_i(\eta) = \frac{1}{\sqrt{2\pi} \hat{\sigma}_i(\vec{x})} e^{-\frac{1}{2} \frac{[\eta - \hat{\eta}_i(\vec{x})]^2}{\hat{\sigma}_i^2(\vec{x})}}$$

функция $\hat{\eta}_1(\vec{x})$ — это оценка "кривой" $\eta(\vec{x})$, соответствующая i -тому набору, и $\hat{\sigma}_1(\vec{x})$ — ширина соответствующего коридора ошибок (см. §1).

Подставляя (6.15) в (6.12), получим, что измерение следует провести в точке, дающей максимум выражения

$$\lambda(\vec{x}) \left\{ \sum_{i=1}^N W_i \hat{\sigma}_i^2(\vec{x}) + \sum_{i=1}^N W_i \left[\hat{\eta}_i(\vec{x}) - \sum_{k=1}^N W_k \hat{\eta}_k(\vec{x}) \right]^2 \right\}. \quad (6.17)$$

Если $N=1$, то планирование по формуле (6.17) совпадает с непрерывным планированием, описанным в §4.

Аналогичный, хотя более сложный результат можно получить для случая k "полезных" параметров, если использовать меру информации (6.3) (см. /19/).

Пример 6.1.

При помощи описанного метода спланируем эксперимент по R -рассеянию из примера 5.1.

Для вычисления формулы (6.17) нужна функция эффективности $\lambda(\vec{x})$ (см. рис. 6), кривые $\hat{\eta}_1(x)$, $\hat{\eta}_2(x)$, $\hat{\sigma}_1(x)$, $\hat{\sigma}_2(x)$ (см. рис. 7 для $\hat{\eta}(x) = \hat{R}(x)$ и рис. 9 для $\hat{\eta}(x) = \hat{D}(x)$) и веса W_1, W_2 (в результате вычисления (6.15) получаем $W_1:W_2 \approx 1:9$).

Вычисленная из (6.17) "скорость прироста количества информации"

$$\xi(x) = \frac{\partial J(t, x)}{\partial t} \Big|_{t=0} \quad (6.18)$$

приведена на рис. 11 для случая измерения $R(x)$ и на рис. 12 — для случая измерения $D(x)$. Результаты вычислены с точностью до несущественного постоянного коэффициента.

Из сравнения кривой на рис. 12 с кривыми на рис. 10 вытекает, что в случае измерения величины $D(x)$ оптимальный дискриминирующий эксперимент совпадает с экспериментом, оптимальным по информационному критерию. (В обоих случаях оптимальный угол рассеяния равен $x_0 = 60^\circ$). Это совпадение является следствием того, что во всех трех формулах (5.12), (5.13) и (6.17) доминирующим является член

$$\lambda(x) \left[D(\hat{\theta}^{(1)}, x) - D(\hat{\theta}^{(2)}, x) \right]^2.$$

Если измеряемой величиной является $R(x)$, то оптимальный дискриминирующий эксперимент должен быть проведен на угле $x = 25^\circ$ (рис. 8), а максимум информации из эксперимента получается на угле $x = 53^\circ$. Из рис. 8 ясно, что прирост информации на угле 53° получается только за счёт понижения ошибок фазовых сдвигов, а дискриминирование неправильного набора фазовых сдвигов будет незначительно. Поэтому более разумно провести измерение на каком-то промежуточном угле, например, $x = 28^\circ$, где прирост информации и дискриминация близки к оптимальным.

Оценим увеличение времени измерения t , необходимого для получения из эксперимента заданного (небольшого) прироста информации ΔJ . Из (6.18) вытекает, что

$$t = t(x, \Delta J) = \frac{1}{\xi(x)} \Delta J \quad (6.19)$$

и на основе результатов, приведенных на рис. 11, получим

$$\frac{t(x = 28^\circ, \Delta J) - t(x = 53^\circ, \Delta J)}{t(x = 53^\circ, \Delta J)} = 30\%.$$

Аналогичным образом, используя кривые на рис. 8, получим для дискриминирующего эксперимента

$$\frac{t(x = 28^\circ) - t(x = 25^\circ)}{t(x = 25^\circ)} = 7,5\%.$$

Л и т е р а т у р а

1. M.G.Kendall, A.Stuart. The Advanced Theory of Statistics, Volume 2. Charles Griffen and Comp. London 1961.
2. Н.П.Клепиков, С.Н.Соколов. Анализ и планирование экспериментов методом максимума правдоподобия, "Наука", Москва 1964.
3. Г.Крамер. Математические методы статистики. ГИИЛ, Москва 1948.
4. Г.Файснер. Поляризация нуклонов при рассеянии, ИИЛ, Москва 1960.

5. Л.Шифф. Квантовая механика. ИИЛ Москва 1957.
6. Л.Яноши. Теория и практика обработки результатов измерений "Мир" Москва 1965.
7. J.E.Monahan, A.Langsdorf, Jr. Annals of Phys., 34, 238 (1965).
8. J.Kiefer, J.Wolfowitz. Ann. Math. Stat. 30, 272 (1959).
9. M.Stone. Ann. Math. Stst., 30, 55 (1959).
10. S.Karlin, W.J.Studden. Ann. Math. Stat., 37, 783 (1966).
11. С.Н.Соколов, Н.П. Клепиков. Теория вероятностей и ее применения, 8, 238 (1963).
12. И.С.Березин, Н.П.Жидков, Методы вычислений. "Наука" Москва 1966.
13. В.В.Налимов, Н.А.Чернова. Статистические методы планирования экспериментальных экспериментов. "Наука" Москва 1965.
14. С.Н.Соколов. Теория вероятностей и ее применения, 8, 95 (1963); 8, 318 (1963).
15. А.Пазман, В.В.Федоров. ЯФ, 6, 853 (1967). Preprint E-3066. Dubna 1966.
16. A.Pazman, Preprint JINR E5-3047. Dubna 1966.
17. Ф.Легар, В.В.Федоров, З.Яноут. ЯФ, 5 887 (1967).
18. D.Lindley. Ann. Math. Stat. 27, 986 (1956).
19. V.V.Fedorov, A.Pa'zman. Preprint E5-3247 Dubna 1967.
20. J.Kiefer. Proc. 4-th Berkeley Symp. on Math. Stat. and Prob., vol.1, 381 (1961).
21. J.Kiefer. J.Royal Stat. Soc., B, 21, 272 (1959).
22. J.Kiefer. Ann. Math. Stat., 32, 298 (1961).

Рукопись поступила в издательский отдел
30 ноября 1967 года.

Таблица I
 условия совпадения оптимальных планов для различных критериев
 оптимальности (в предположении, что $\eta(\theta, x) = \sum_{\alpha=1}^m \theta_{\alpha} t_{\alpha}(x)$)

Критерии	D	A	A	d	$\min_{\epsilon} \max_{x \in X} \delta^2(x)$	I
D	/	Оптимальные планы имеют диагональную матрицу D		m = 1	всегда	γ_1 распределены по нормальному закону или $T \rightarrow \infty$
A		/	m = 1	условие A \leftrightarrow D	условие I \leftrightarrow D И условие A \leftrightarrow D	условие I \leftrightarrow D И условие A \leftrightarrow D
d			/	m = 1	m = 1	условие I \leftrightarrow D И m = 1
$\min_{\epsilon} \max_{x \in X} \delta^2(x)$				/	/	условие I \leftrightarrow D

Таблица 2
Оптимальные статические планы

Интеграл допустимых точек измерения	вид $\lambda(x)$	Оптимальные точки измерения $x^{(1)}, \dots, x^{(n)}$ это нули полинома n -той степени $\phi(x)$, где:
$\langle -1, 1 \rangle$	$\lambda(x) = 1$	$\phi(x) = (1-x^2) \frac{dP_{n-1}(x)}{dx}$ где $P_n(x)$ полином Лежандра
$\langle -1, 1 \rangle$	$\lambda(x) = \frac{(1-x)^{\alpha+1} (1+x)^{\beta+1}}{\alpha! \beta!}$ где $\alpha > -1, \beta > -1$	$\phi(x) = P_n^{(\alpha, \beta)}(x)$ где $P_n^{(\alpha, \beta)}(x)$ - полином Якоби
$\langle 0, \infty \rangle$	$\lambda(x) \approx e^{-x}$	$\phi(x) = x L_{n-1}^{(1)}(x)$ где $L_n^{(\alpha)}(x)$ - полином Лагера с параметром α
$\langle 0, \infty \rangle$	$\lambda(x) = x^{\alpha+1} e^{-x}$ где $\alpha > -1$	$\phi(x) = L_n^{(\alpha)}(x)$
$\langle -\infty, \infty \rangle$	$\lambda(x) \approx e^{-x^2}$	$\phi(x) = H_n(x)$ где $H_n(x)$ - полином Эрмита

Таблица 3

Критерий оптимальности	Непрерывное планирование	Оптимальная точка измерения соответствует
	Последовательное планирование	
$\min \mathcal{F} \{ \hat{h} \}$	$\max_x \lambda(x) [h' D(0) f(x)]^2$	$\max_x \lambda(x) [h' D(0) f(x)]^2$
$\min D $	$\max_x \lambda(x) \tau \hat{\sigma}_0^2(x)$	$\max_x \lambda(x) \hat{\sigma}_0^2(x)$
$\min D(\omega) $	$\max_x \frac{\sum_{\alpha, \beta=0}^k [D(\omega; 0)]_{\alpha\beta} C_{\alpha\beta}(x)}{[\lambda(x) \tau]^{-1} + \hat{\sigma}_0^2(x)}$ $C(x) = D(0) f(x) f(x) D(0)$	$\max_x \lambda(x) \sum_{\alpha, \beta=0}^k [D(0; \omega)]_{\alpha\beta} C_{\alpha\beta}(x)$

Таблица 4

Номер опыта	x_1	x_2	y	θ_1	θ_2	θ_3
I	1,0	1,0	1,26			
2	2,0	1,0	0,219			
3	1,0	2,0	0,076			
4	2,0	2,0	0,126	10,39	48,83	0,74
5	0,1	0,0	0,186	3,11	15,19	0,79
6	3,0	0,0	0,606	3,96	15,32	0,66
7	0,2	0,0	0,268	3,61	14,00	0,66
8	3,0	0,0	0,614	3,56	13,96	0,67
9	0,3	0,0	0,318	3,32	13,04	0,67
10	3,0	0,8	0,298	3,33	13,48	0,67
II	3,0	0,0	0,509	3,74	13,71	0,63
12	0,2	0,0	0,247	3,58	13,15	0,63
13	3,0	0,8	0,319	3,57	12,77	0,63

Таблица 5

№	Характеристики планов	отношение $ D / D _{opt}$
1	$x^{(1)} = (1; 1), x^{(2)} = (1; 2), x^{(3)} = (2; 2), x^{(4)} = (2; 1)$	$5 \cdot 10^7$
	$N_1 = 3, N_2 = 3, N_3 = 3, N_4 = 3$	
2	$x^{(1)} = (0; 0), x^{(2)} = (0; 3), x^{(3)} = (3; 3), x^{(4)} = (3; 0)$	10^9
	$N_1 = N_2 = N_3 = N_4 = 3$	
3	Первые 12 точек из табл. 4.2	1,5
	$N_i = 1 \quad (i = 1, \dots, 12)$	
4	$x^{(1)} = (0,2; 0), x^{(2)} = (3; 0), x^{(3)} = (3; 0,8)$	1
	$N_1 = N_2 = N_3 = 4$	

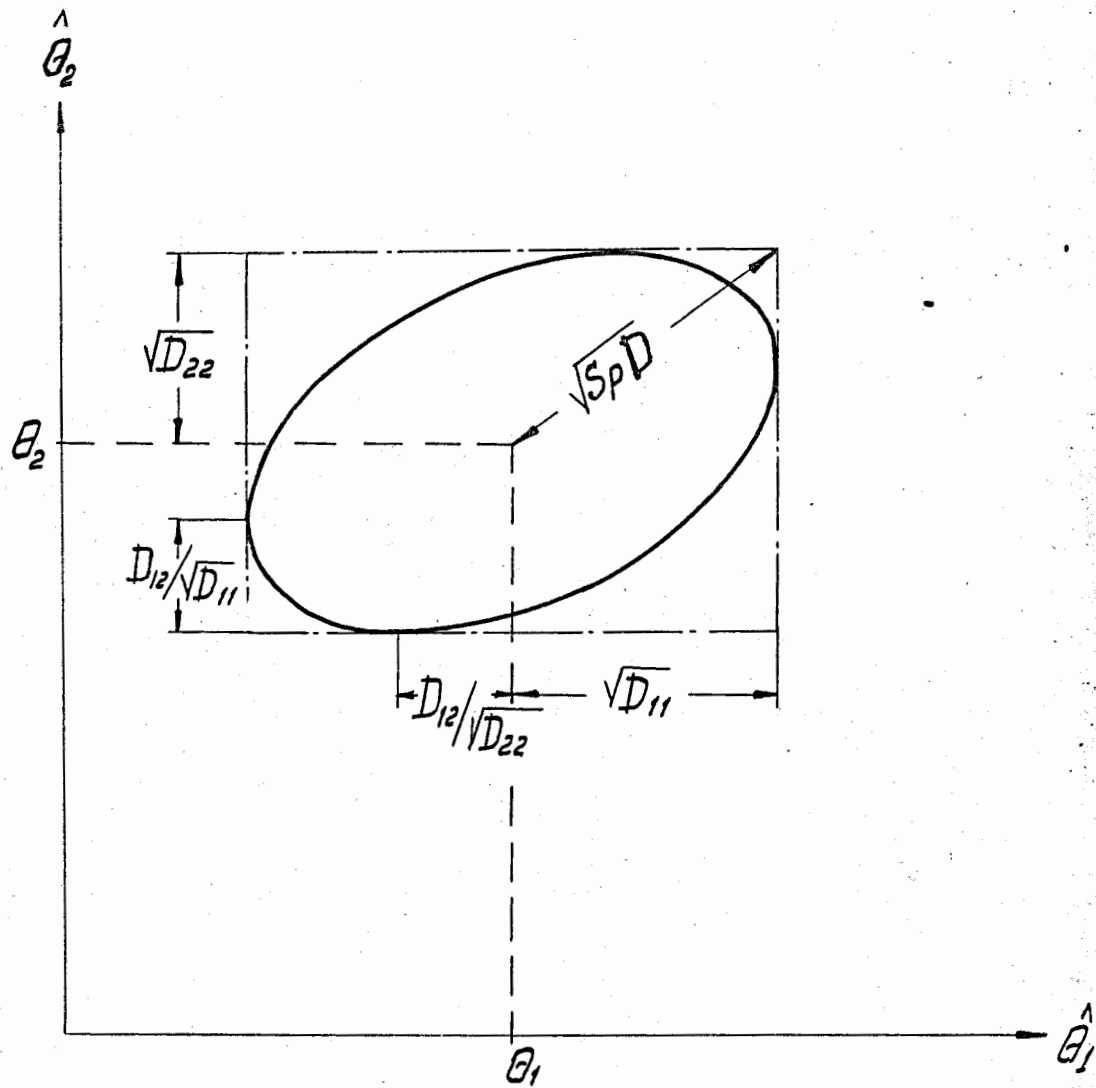


Рис. 1.

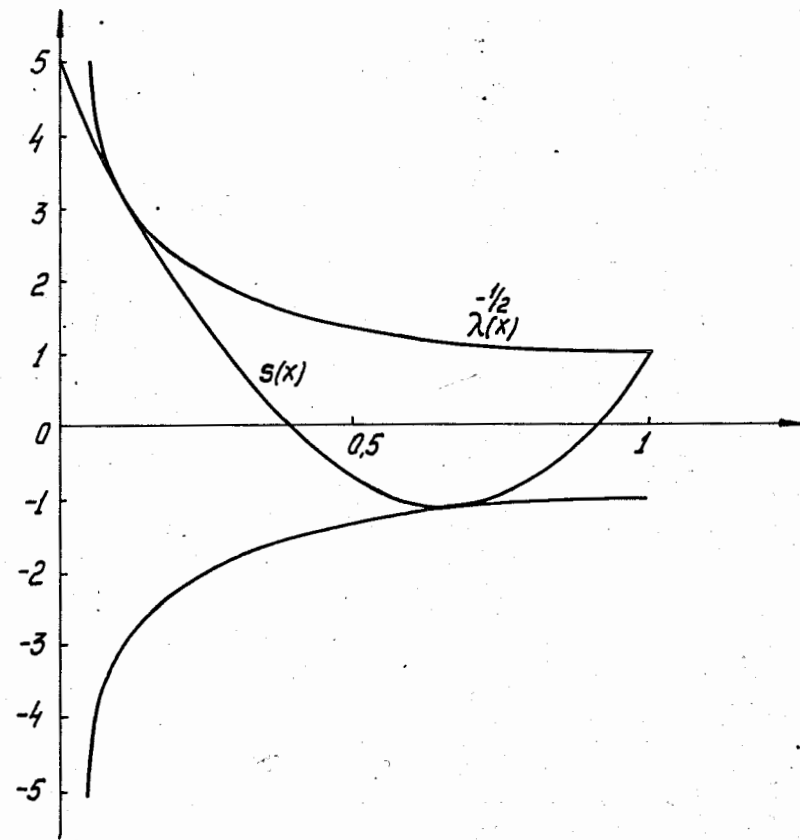


Рис.2.

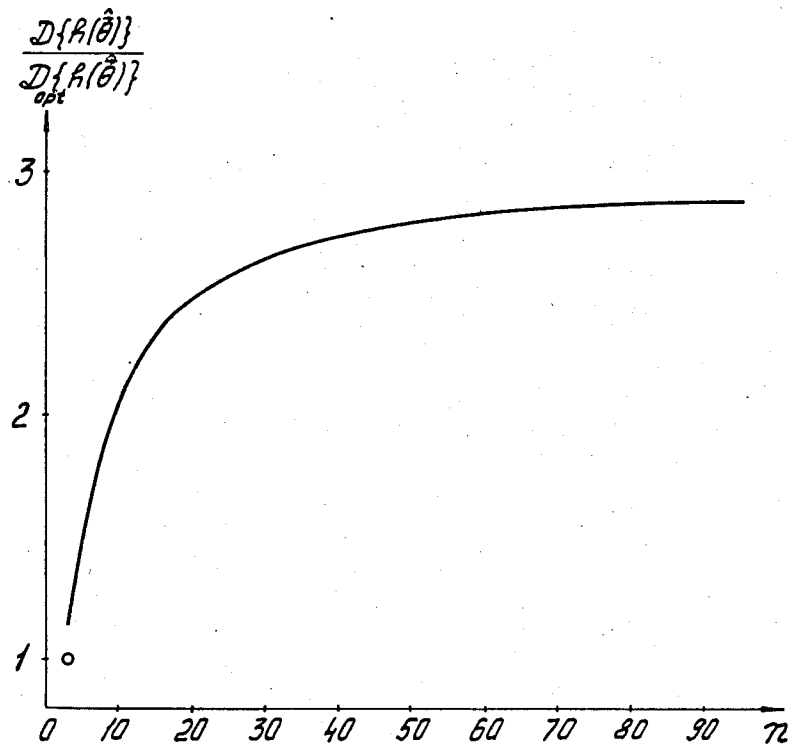


Рис. 3.

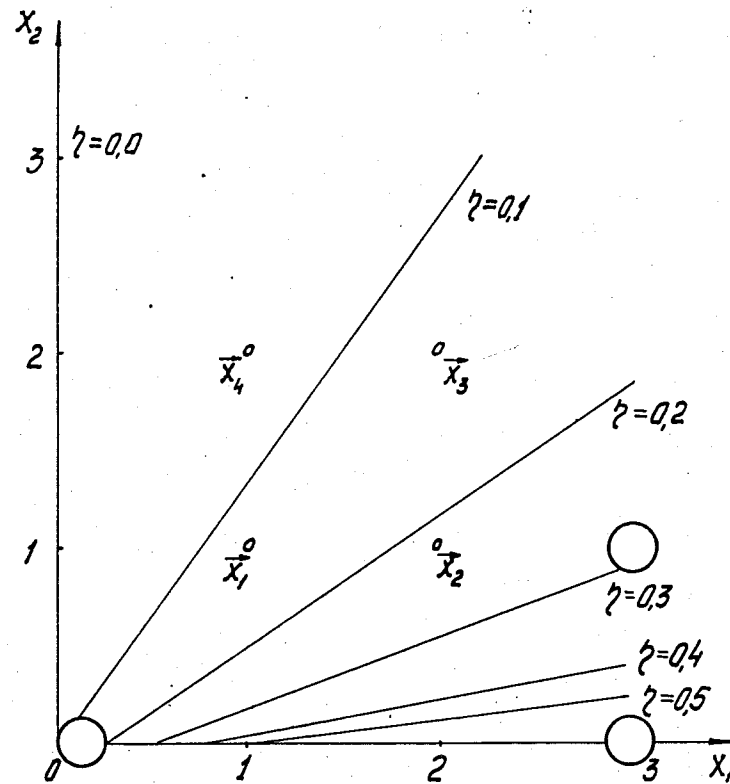


Рис. 4.

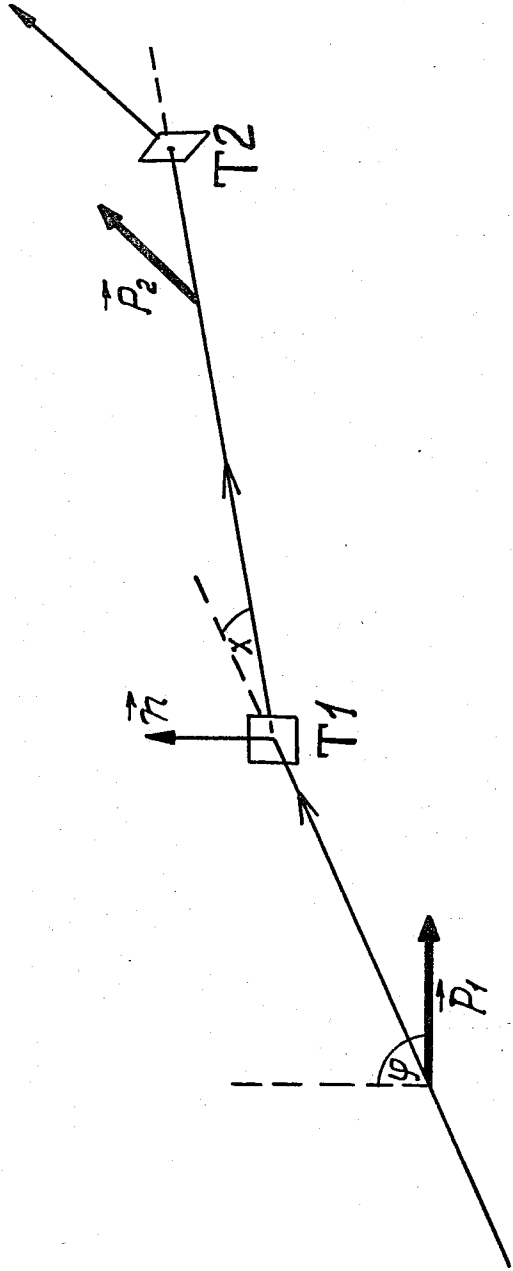


FIG. 5.

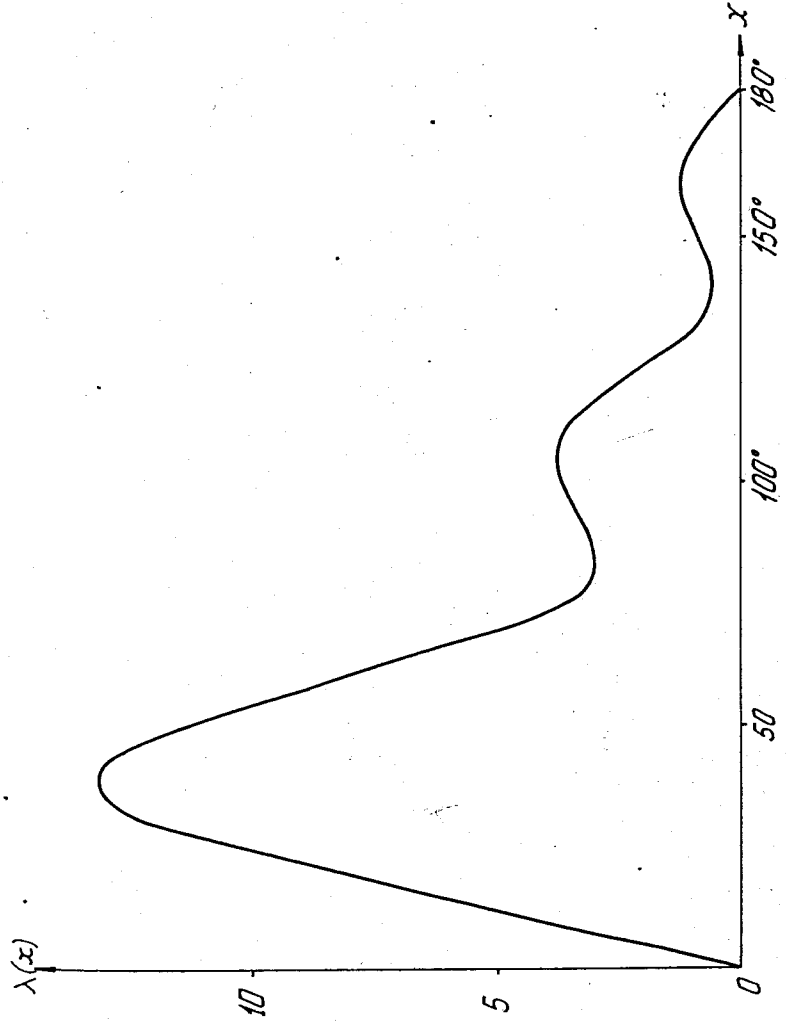


FIG. 6

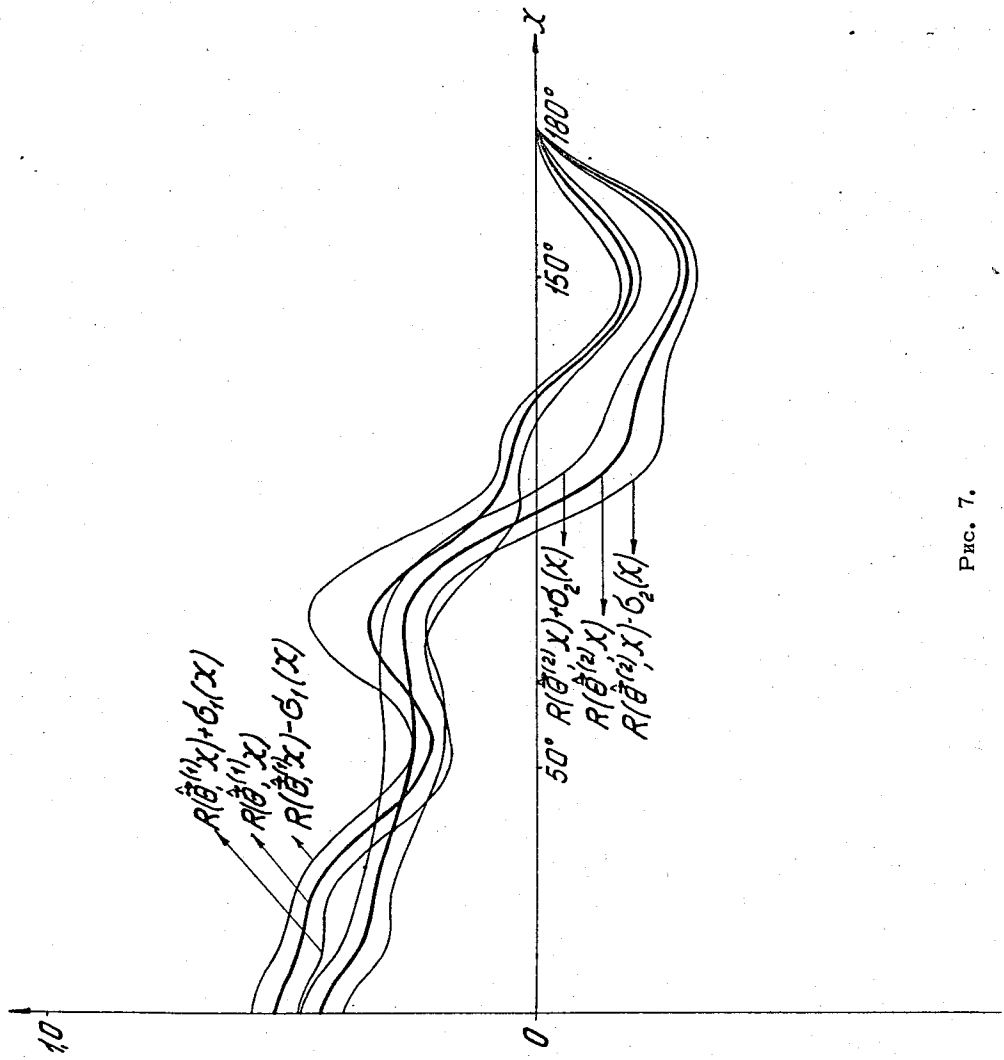


Рис. 7.

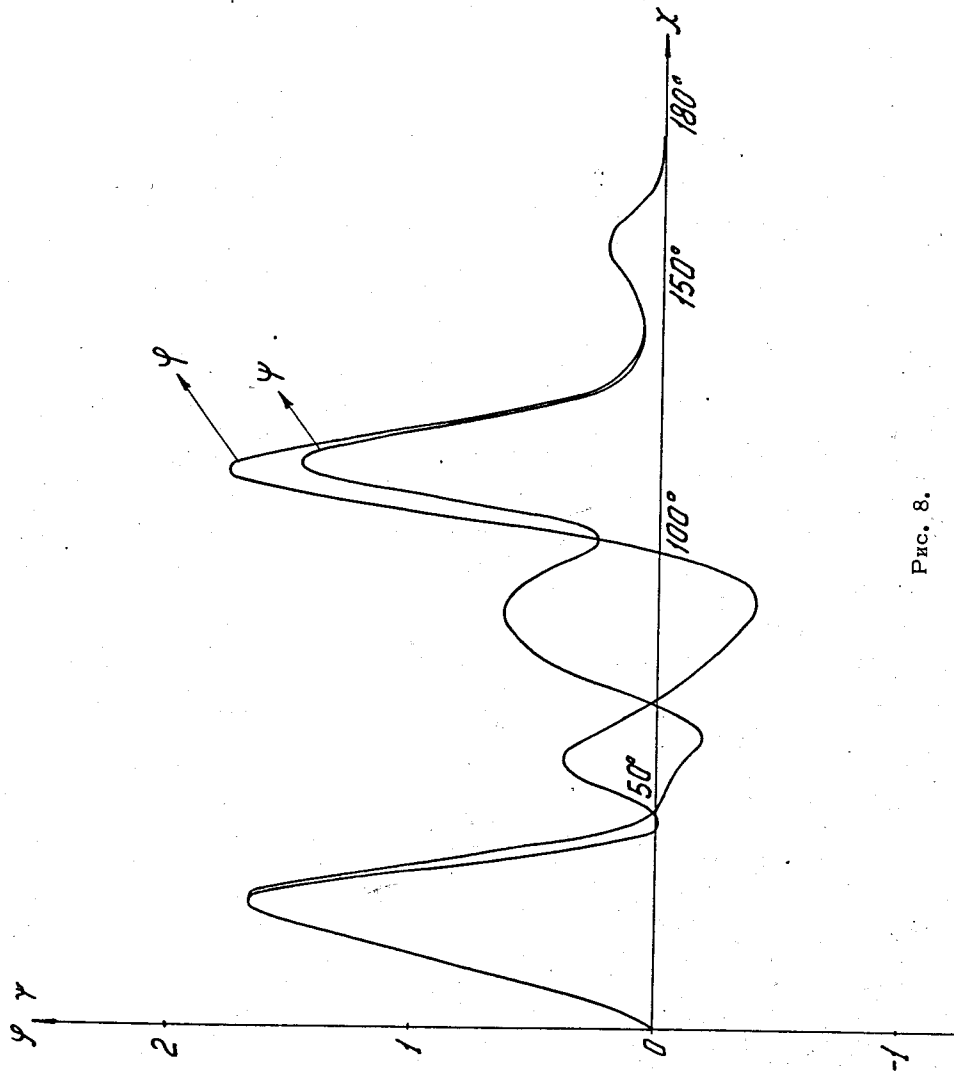


Рис. 8.

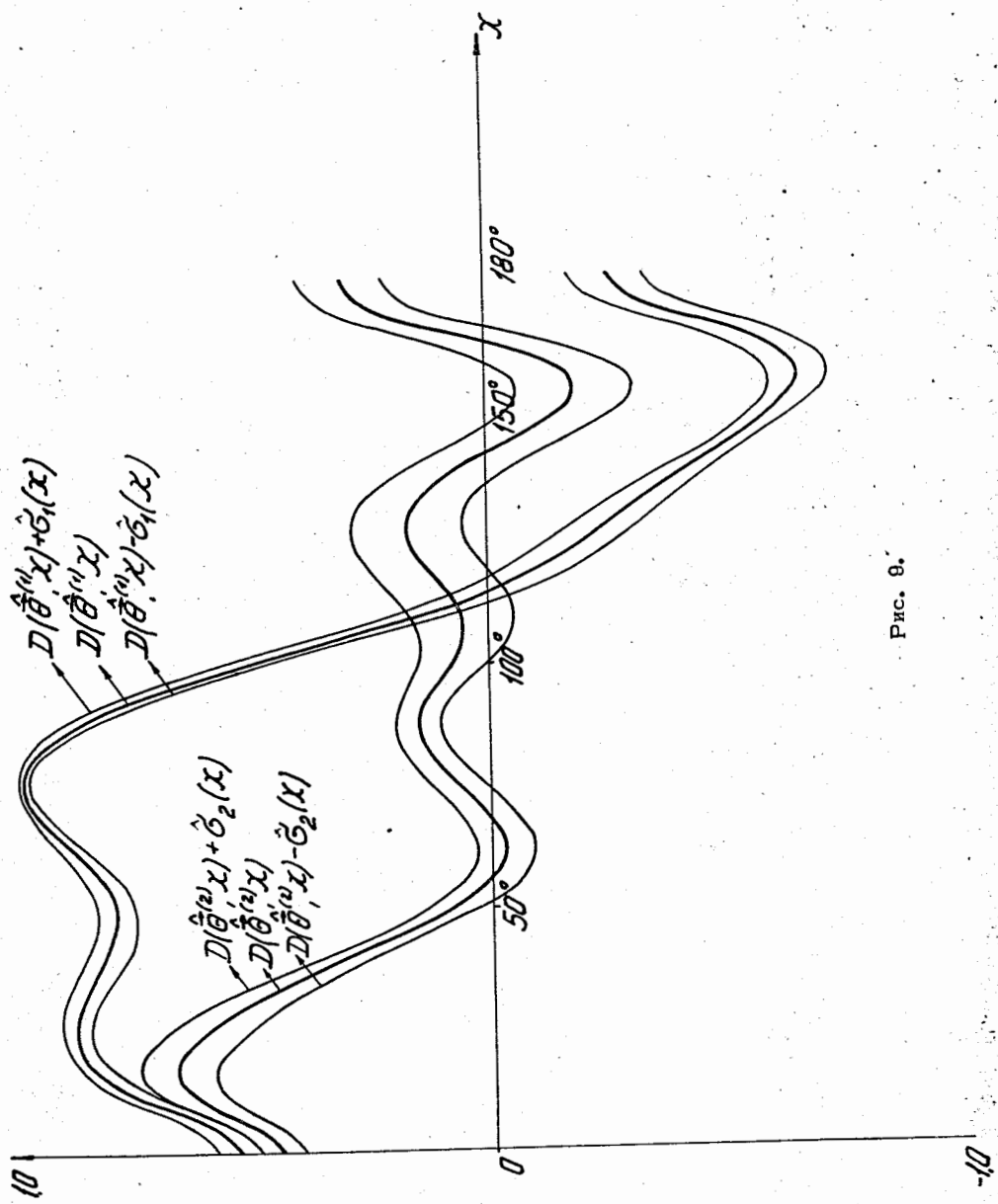


Рис. 9.

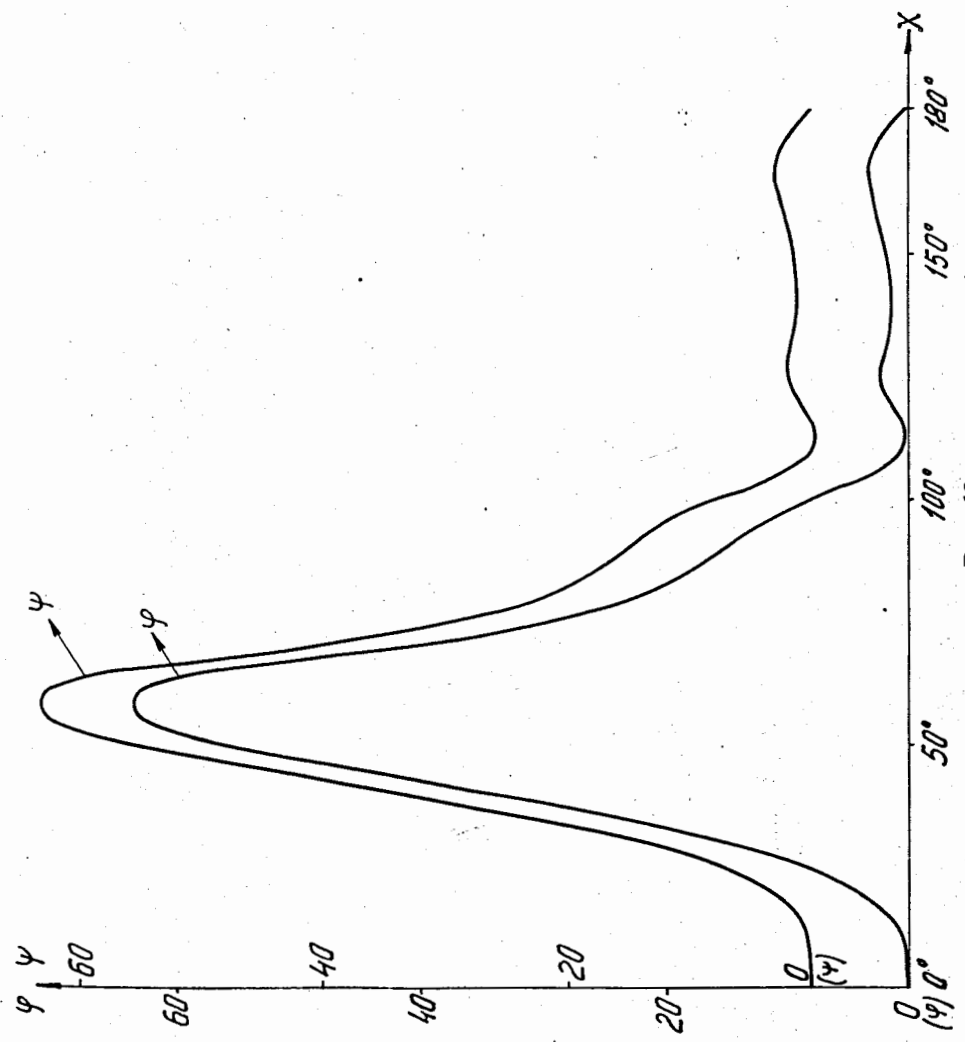


Рис. 10.

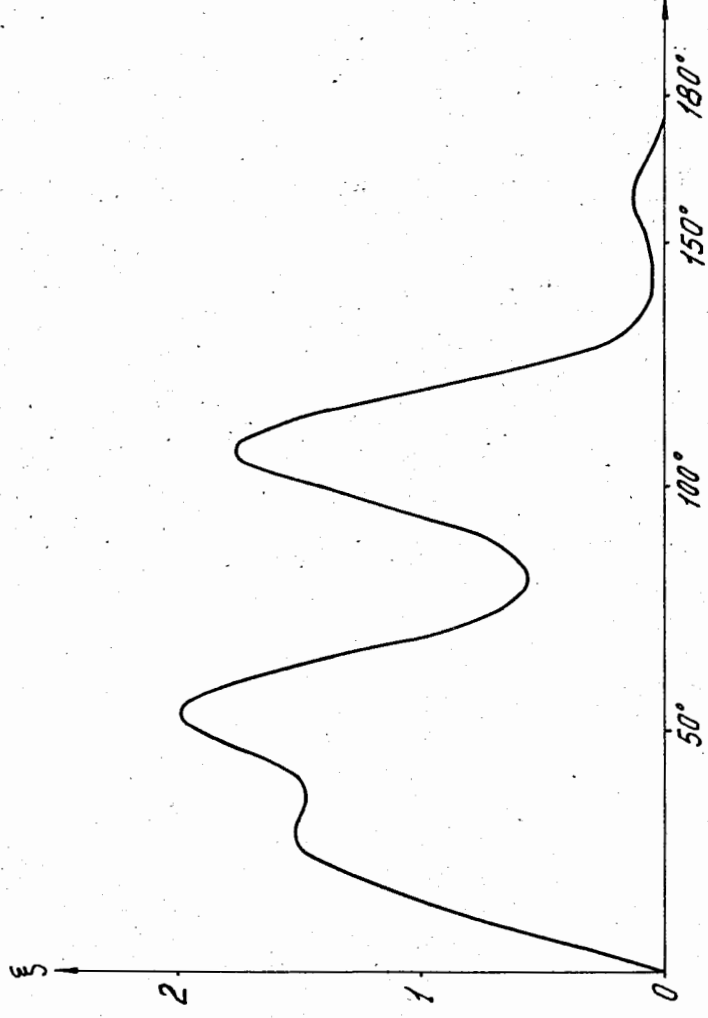


Рис. 11.

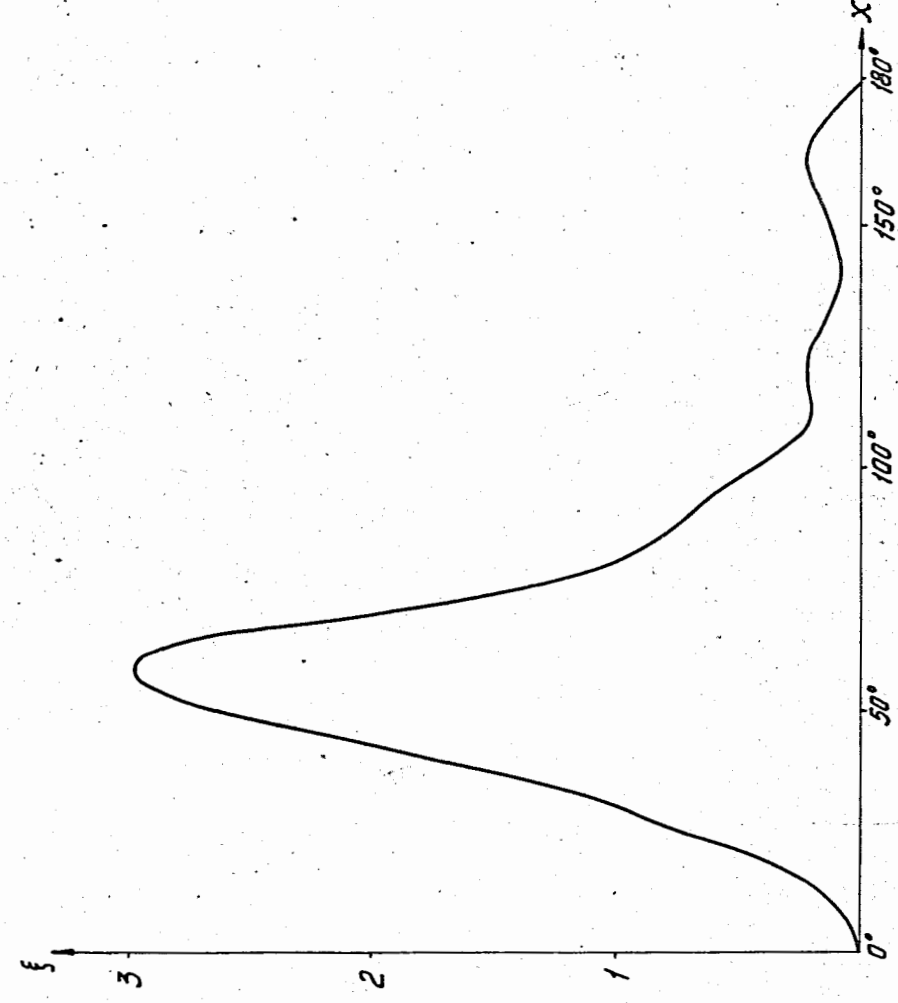


Рис. 12.