П Р О М Ы Ш Л Е Н Н А Я   А К А Д Е М И Я

А.С. ЛУНИН

ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА В НАУКЕ И ТЕХНИКЕ.

МЕТОДЫ МНОГОФАКТОРНОЙ ОПТИМИЗАЦИИ I - ОГО И 2 - ГО ПОРЯДКА

новый взгляд на проблему планирования эксперимента
с учётом имеющихся противоречий в теории
и сложностей в практическом воплощении

МОСКВА - I992 ГОД

§ I. Введение.

Экспериментальный поиск экстремума некоторой функции неизвестного вида (функции отклика) осуществляют через оценку ее значений опытным путем в ряде точек факторного пространства. Экспериментальная оптимизация, как составная часть теории планирования эксперимента, пользуется значительным вниманием со времени появления работы Дж.Бокса и К.Уилсона /I; 2, с.4I0-42I/. Однако, до настоящего времени в планировании эксперимента вообще применяют односторонний, дисперсионный подход к выбору плана эксперимента. Этот подход нельзя признать обоснованным, поскольку не учитывается систематическое отклонение модели от реальной функции. Кроме того, методы экспериментальной оптимизации можно более упорядочить и разнообразить, используя их логическую взаимосвязь с соответствующими численными методами поиска экстремума.

Цель настоящего исследования - выявить и рассмотреть в этой связи наиболее рациональные основные статистические методы I-ого и 2-ого порядка для поиска локального экстремума функции многих переменных по ее экспериментальным значениям.

Исходной базой для их построения, естественно, являются численные методы Коши ("градиентный"), Ньютона и модификации последних /3/. В классической форме эти методы I-ого и 2-ого порядка требуют вычисления в каждой исходной точке движения соответственно градиента и вектора Ньютона. Если их расчет затруднен, то применяют аппроксимацию методом конечных разностей. Для этого нужно вычислять только значения функции в ряде точек, т.е. по определенному плану. Отклонение получаемой разностной оценки от истинной характеристики (систематическая ошибка) зависит как от производных высших порядков, так и от плана вычислений (расположения точек, где вычисляют значения функции).

В рассматриваемом случае вид функции неизвестен, а ее значения определяют экспериментально и поэтому неточно. Следовательно, разностная аппроксимация - здесь единственно возможный путь для оценки градиента или вектора Ньютона. Наряду с систематической ошибкой, имеет место и случайная ошибка, зависящая как от точности эксперимента, так и его плана (расположения точек, где находят опытные значения функции). Очевидно, что этот план тем лучше, чем меньшую общую ошибку он дает по отношению к числу опытов.

Еще одна особенность заключается в том, что методы должны применяться в шаговой (неитерационной) форме. Т.е. градиент или вектор Ньютона лишь задают направление движения, а местоположение промежуточного экстремума определяют также экспериментально по профилю пути. Это означает сведение исходной многомерной задачи к серии одномерных задач и обусловлено существенным вкладом случайной ошибки, резко сужающей область сходимости итерационной процедуры.

К настоящему времени при поиске локального экстремума функции отклика широко применяют метод Бокса-Уилсона, который использует статистическую оценку градиента, т.е. является статистическим методом I-ого порядка. Для оценки градиента в этом методе служит план полного факторного эксперимента (план n-КУБ) или его подходящая дробная реплика.

В данной работе для этих целей обосновано применение значительно более простого плана n -КРЕСТ. При этом принимали во внимание как случайную, так и систематическую ошибку модели. Причем пригодность данной модели к применению определяют по ее общей относительной погрешности. Разработаны также методы экспериментальной оптимизации 2-го порядка с привлечением метода регуляризации Тихонова /4/ для корректного решения уравнений с неточными коэффициентами.

Основные обозначения и определения:

x_i - фактор i (независимая переменная i), $i = 1, n$;

$t_i = \frac{x_i - x_{iR}}{\Delta x_i}$ - нормированный фактор i (путем переноса начала координат в точку R и выбора масштабов Δx_i);

$x = (x_1, \ldots, x_n)$, $t = (t_1, \ldots, t_n)$ - точки n -мерных координатных линейных пространств $\langle x \rangle$ и $\langle t \rangle$;

M - диагональная масштабная матрица с диагональными элементами $\Delta x_1, \ldots, \Delta x_n$; $x - x_R = Mt$;

$y = f(x) = y(t)$ - неизвестная однозначная скалярная функция (функция или поверхность отклика);

$$g = g(t) = \frac{dy}{dt}, \quad G = G(t) = \frac{d^2y}{dt^2} \text{ и } V = V(t) = \frac{d^3y}{dt^3}$$

- градиент, матрица Гессе и $n \times n \times n$ -матрица-производная для y ;

в точке R : $y = y_R$, $g = g_R$, $G = G_R$ и $V = V_R$ ($t_R = 0$);

аналогично $j = j(x) = \frac{dy}{dx}$, $J = J(x) = \frac{d^2y}{dx^2}$ и $W = W(x) = \frac{d^3y}{dx^3}$;

$y = y_R + g_R t + \frac{1}{2!} G_R t^2 + \frac{1}{3!} V_R t^3 + \ldots$ - символическое представление аналитической функции $y(t)$ степенным рядом;

аналогично для $y = f(x)$:

$$y = y_R + j_R(x - x_R) + \frac{1}{2!} J_R (x - x_R)^2 + \frac{1}{3!} W_R (x - x_R)^3 + \ldots$$

(формулы масштабных преобразований тензорных производных:

$g = Mj$, $G = MJM$, $V = M\overset{M}{W}M$ и т.д. и наоборот)

$P_R = -G_R^{-1} g_R$ – вектор Ньютона для функции $y(t)$ в точке R ;

$D = D(t)$ – диагональная матрица-производная для y с диагональными элементами: $\frac{\partial^2 y}{\partial t_1^2}$, ... , $\frac{\partial^2 y}{\partial t_m^2}$; в точке R :

$D = D_R$;

$h_R = -D_R^{-1} g_R$ – некоторый вектор для функции $y(t)$ в точке R ;

$\tau > 0$ – параметр регуляризации по Тихонову;

$P_R^{(\tau)}$ и $h_R^{(\tau)}$ – регуляризованные по Тихонову векторы P_R и h_R ;

$\hat{g}$, $\hat{G}$, $\hat{P}$, $\hat{D}$ и $\hat{h}$ – разностные оценки соответствующих характеристик;

$\tilde{y} = \tilde{y}(t)$ – статистическая функция отклика (случайное значение y в данной точке; оно может быть единственным опытным значением или выборочным среднеарифметическим);

$y_{(k)}$, $\tilde{y}_{(k)}$, $t_{i(k)}$ – значения y , $\tilde{y}$ и t_i в k –ой точке плана;

$\tilde{g}$, $\tilde{G}$, $\tilde{P}$, $\tilde{D}$ и $\tilde{h}$ – статистические оценки соответствующих характеристик;

$y = \Psi(\varepsilon) = y(\varepsilon \tau_R)$ – профиль пути при движении из точки R по поверхности $y(t)$ в соответствии с направлением вектора τ_R ;

$\tilde{y} = \tilde{\Psi}(\varepsilon) = \tilde{y}(\varepsilon \tau_R)$ – аналогичный статистический профиль пути;

назностные модели функции $y(t)$ или поверхности аппроксимации:

$$Z_1 = y_R + \hat{g}_R t \tag{1}$$

$$Z_2 = y_R + \hat{g}_R t + \frac{1}{2} \hat{G}_R t^2 \tag{2}$$

$$Z_3 = y_R + \hat{g}_R t + \frac{1}{2} \hat{D}_R t^2 \tag{3}$$

$$Z_4 = y_k + \hat{g}_k t + \frac{1}{2}(\hat{G}_k - \hat{D}_k) t^2 \qquad (4)$$

регресионные модели функции или поверхности регрессии:

$$\bar{Z}_1 = \tilde{b}_0 + \sum_{i=1}^{n} \tilde{b}_i t_i \qquad (5)$$

$$\bar{Z}_2 = \tilde{b}_0 + \sum_{i=1}^{n} \tilde{b}_i t_i + \sum_{i=1}^{n} \tilde{b}_i t_i^2 + \sum_{i=1}^{j-1} \sum_{j=1}^{n} b_{ij} t_i t_j \qquad (6)$$

$$\bar{Z}_3 = \tilde{b}_0 + \sum_{i=1}^{n} \tilde{b}_i t_i + \sum_{i=1}^{n} \tilde{b}_{ii} t_i^2 \qquad (7)$$

$$\bar{Z}_4 = \tilde{b}_0 + \sum_{i=1}^{n} \tilde{b}_i t_i + \sum_{i=1}^{j-1} \sum_{j=1}^{n} \tilde{b}_{ij} t_i t_j \qquad (8)$$

МНК – метод наименьших квадратов;

$E(\ldots)$, $\sigma^2(\ldots)$, $s^2(\ldots)$ – математическое ожидание, дисперсия и выборочная (несмещенная) дисперсия;

$s^2_{ad}(\tilde{y}, \bar{z})$ – выборочная (несмещенная) дисперсия адекватности $\tilde{y}$ и $\bar{Z}$ (причем $E(\bar{Z}) = Z$);

B – количество коэффициентов в модели регрессии;

N – количество точек в плане;

область плана n -шар минимального радиуса, содержащий все точки плана (для центральносимметричных планов);

n -КУБ – план полного факторного эксперимента (т.е. в вершинах центрального единичного n -куба, грани которого перпендикулярны осям координат; $N = 2^n$);

n -КРЕСТ – план, в котором точки расположены в вершинах центрального единичного n -креста, совпадающего с осями координат, и одна точка в центре, $N = 2n + 1$;

центральный композиционный план (n -ЦКП) – композиция из плана n -КРЕСТ и плана n -КУБ или дробной реплики последнего субпорядка m , $N = 2^m + 2n + 1$;

α – отношение плеча n -креста к ребру n -куба в n -ЦКП (в частности для ортогонального плана: $\alpha_\perp = \sqrt{\dfrac{-2^m + \sqrt{2^m \cdot N}}{2}}$

$\xi^{\bullet}$, ξ^{+} и $\xi^{\square}$ – количество опытов соответственно в центре плана, в вершинах n -креста и в вершинах n -куба; верхние индексы I, II и III означают, что соответствующие характеристики вычислены с использованием соответственно плана n -КУБ (или его дробной реплики) с моделью 4 или 8, плана n -КРЕСТ с моделью 3 или 7 и n -ЦКП с моделью 2 или 6.

С целью аппроксимации по МНК функции отклика регрессионными моделями 5 – 8, а также для оценки значимости коэффициентов регрессии принимают гипотезы:

I) функция $y(t)$ аналитична в рассматриваемой области (или, по крайней мере, дважды непрерывно дифференцируема);

2) $E(\tilde{y}_{(k)} - y_{(k)}) = 0$, $E(\tilde{y}_{(k)} - y_{(k)})^2 = \sigma^2(y) = const$,

$$E\left[(\tilde{y}_{(j)} - y_{(j)})\cdot(\tilde{y}_{(k)} - y_{(k)})\right] = 0 \qquad (j \neq k);$$

3) $\tilde{y}_{(k)}$ имеют нормальное распределение.

Гипотезы 2 и 3 подвергаемы статистической проверке. Им эквивалентна гипотеза, что ($\tilde{y}$ – y) имеет центральное однородное нормальное распределение. Кроме того, предположим: а) учтены все влияющие факторы, б) значения $t_{i(k)}$ задают без ошибок, в) функция $y(t)$ имеет не более одного экстремума. В соответствии с гипотезой I функция отклика представима степенным рядом:

$$y = y(t) = \beta_0 + \sum_{i=1}^{n} \beta_i t_i + \sum_{i=1}^{j} \sum_{j=1}^{n} \beta_{ij} t_i t_j + \dots =$$

$$= \beta_0 + \sum_{k=1}^{\infty} \underbrace{\beta_{i_1 \dots i_m}}_{k \geq m} t_{i_1}^{q_1} \dots t_{i_m}^{q_m} \qquad (9)$$

Причем $\beta_{i_1 \dots i_m} = \dfrac{1}{q_1! \dots q_m!} \dfrac{\partial^k y}{\partial t_{i_1}^{q_1} \dots \partial t_{i_m}^{q_m}} \left(\sum_{p=1}^{m} q_p = k \right)$,

где $q_1, \dots, q_m$ – количества одинаковых индексов. Аналогично $\alpha_{i_1 \dots i_m}$ – коэффициенты разложения $y = f(x)$ в степенной ряд.

§ 2. Дискуссия.

Согласно методу Бокса-Уилсона движение к экстремуму функции отклика осуществляют сериями шагов вдоль или против статистического градиента $\tilde{g}_k = (\tilde{b}_1, \ldots, \tilde{b}_n)$. Коэффициенты регрессии $\tilde{b}_i$ находят по плану n-КУБ или его подходящей дробной реплике с центром в каждой исходной точке движения. К $\bar{z}$ предъявляют требование адекватности $\tilde{y}$ на множестве точек плана по критерию Фишера. В зависимости от выполнения этого требования последовательно применяют модели 5, 8 и 6. На последнем этапе в случае адекватности модели 6 и функции $\tilde{y}$ задача сводится к вычислению координат экстремума (при соблюдении достаточных условий).

В связи с этим методом покажем некорректность некоторых положений, используемых при экспериментальном поиске экстремума с применением регрессионных моделей.

Первый вопрос связан с требованием адекватности $\bar{z}$ и $\tilde{y}$ на этапе движения к экстремуму. В статистике под адекватностью понимают высокую вероятность тождественности в пределах случайной ошибки. С увеличением случайной ошибки вероятность адекватности модели и функции возрастает. Следовательно, эта характеристика сама по себе не может являться критерием применимости полученной модели для оценки направления движения.

Адекватность $\bar{z}$ и $\tilde{y}$ устанавливают по критерию Фишера $\tilde{F} = s^2_{ad}(\tilde{y}, \bar{z}) / s^2(\tilde{y})$. В отличие от табулированной величины (F_{kp}) статистика $\tilde{F}$ имеет в общем случае нецентральное распределение Фишера, т.е. нецентральное χ^2-распределение числителя:

$$E\left[s^2_{ad}(\tilde{y},\bar{z})\right] = \sigma^2(\tilde{y}) + \frac{1}{\varphi_1}\cdot\sum_{j=1}^{N}\Delta^2 z_{(j)} = \sigma^2(\tilde{y}) + \overline{\Delta}^2 z\,,$$

где $\Delta z_{(j)} = z_{(j)} - y_{(j)}$, φ_1 – число степенной свободы. Чем меньше $\sigma^2(\tilde{y})$, тем хуже адекватность данной модели: $\lim\limits_{s^2(\tilde{y})\to 0}\hat{F} = \infty$, так как $\lim\limits_{s^2(\tilde{y})\to 0} s^2_{ad}(\tilde{y},\bar{z}) = \overline{\Delta}^2$ (если $\overline{\Delta}^2 \neq 0$). Поэтому неадекватность ($\tilde{\hat{F}} > F_{кр}$) априори не обусловливает непригодность модели для удовлетворительной аппроксимации функции отклика. Например, разностная модификация численного метода Коши использует заведомо неадекватную модель I и тем не менее может с успехом работать при численной оптимизации.

Иными словами при традиционном планировании имеется следующий парадокс: адекватная (и поэтому пригодная к использованию) модель становится неадекватной (непригодной к использованию) при уменьшении ее ошибки за счет более точного эксперимента и наоборот – неадекватную модель можно сделать адекватной за счет менее точного эксперимента!? Избежать этого можно, если принять следующую концепцию:

I) разность между статистической моделью и функцией в каждой точке факторного пространства складывается из случайной и систематической ошибок, а именно: $\bar{z}_{(j)} - y_{(j)} = (\bar{z}_{(j)} - z_{(j)}) + (z_{(j)} - y_{(j)})$;

2) точность модели повышают за счет уменьшения явно большей из этих ошибок (что оценивается статистически);

3) пригодность полученной модели к использованию определяют по ее относительной погрешности (общей).

Чтобы практически реализовать эту концепцию, достаточно уметь оценивать относительную погрешность модели и относительный вклад систематической ошибки. Относительную погрешность модели определим с учетом числа степеней свободы как

$$\rho = \frac{s_{ad}(\tilde{y}, \bar{z})}{\tilde{y}_{max} - \tilde{y}_{min}}$$

где $\tilde{y}_{max}$ и $\tilde{y}_{min}$ - максимальное и минимальное значение $\tilde{y}$ на множестве точек плана. Нетрудно видеть, что при данном подходе точность планирования согласована нужным образом с величинами ошибок: ρ уменьшается при уменьшении как случайной, так и систематической ошибки и наоборот. Отсюда видна применимость подобного критерия и в соответствующих разностных численных методах.

В свою очередь, относительный вклад систематической ошибки можно оценить следующим образом. Пусть $\Delta z = z - y$. Тогда статистика $\frac{s^2_{ad}[\tilde{y}, (\bar{z} - \Delta z)]}{s^2(\tilde{y})}$ имеет центральное распределение Фишера. Поэтому справедливо вероятностное неравенство

$$P\left\{F^{H}_{кр} < \frac{s^2_{ad}[\tilde{y}, (\bar{z} - \Delta z)]}{s^2(\tilde{y})} < F^{B}_{кр}\right\} = I - 2\alpha$$

где P - заданная вероятность; $\alpha = \frac{I - P}{2}$;

$$F^{B}_{кр} = F_{\nu_1, \nu_2 (1-\alpha)} \ ; \ F^{H}_{кр} = F_{\nu_1, \nu_2, \alpha} = F^{-1}_{\nu_2, \nu_1 (1-\alpha)}$$

Учитывая реальное значение критерия Фишера, а именно

$$\tilde{F} = \frac{s^2_{ad}(\tilde{y}, \bar{z})}{s^2(\tilde{y})},$$

преобразуем исходное неравенство таким образом

$$F^{H}_{кр} < \left\{I - \frac{s^2_{ad}(\tilde{y}, \bar{z}) - s^2_{ad}[\tilde{y}, (\bar{z} - \Delta z)]}{s^2_{ad}(\tilde{y}, \bar{z})}\right\} \cdot \tilde{F} < F^{B}_{кр}$$

или

$$F^{H}_{кр} < (1 - \eta^2) \cdot \tilde{F} < F^{B}_{кр}$$

или $$\mathrm{I} - \frac{F_{кр}^{в}}{\tilde{F}} < \eta^2 < \mathrm{I} - \frac{F_{кр}^{н}}{\tilde{F}}$$

где $\eta^2 = \frac{s_{ad}^2(\tilde{y}, \bar{z}) - s_{ad}^2[\tilde{y}, (\bar{z} - \Delta z)]}{s_{ad}^2(\tilde{y}, \bar{z})}$ характеризует относительный вклад систематической ошибки. Так как эти преобразования тождественны, то в итоге

$$P\left\{\mathrm{I} - \frac{F_{кр}^{в}}{\tilde{F}} < \eta^2 < \mathrm{I} - \frac{F_{кр}^{н}}{\tilde{F}}\right\} = \mathrm{I} - 2\alpha \qquad (\mathrm{I0})$$

Наиболее вероятное значение $\eta^2_{(0,5)} = \mathrm{I} - \frac{F_{кр}^{(0,5)}}{\tilde{F}}$

В частности, если $\tilde{F} < F_{кр}^{в}$, то принимается, что ошибка носит случайный характер. Если же $F \gg F_{кр}^{в}$, то принимается, что ошибка в основном систематическая. В общем случае показатель η^2 даёт возможность принять достаточно обоснованное решение, как повышать точность аппроксимации, и оценивается по неравенству (I0). Снижение систематической ошибки достигается уменьшением интервалов варьирования факторов, а снижение случайной ошибки, наоборот их увеличением или увеличением количества опытов в точках плана.

Другой вопрос касается выбора наиболее рационального плана эксперимента. Ряд используемых при этом критериев (ротатабельность; А,Д,... - оптимальность и др.) совершенно не учитывают вклад систематической ошибки, тоже зависящей от плана эксперимента; не принимают во внимание удобство применения плана, включая ортогональность, и, что самое противоречивое, - они неинвариантны по отношению к выбираемым масштабам по осям, которые в сущности произвольны.

Рассмотрим точность статистической аппроксимации градиента.

Вектор $\tilde{g}_R = (\tilde{b}_1, \ldots, \tilde{b}_n)$ является в общем случае смещенной оценкой вектора $g_R = (\beta_1, \ldots, \beta_n)$. Средний квадрат ошибки:

$$\delta^2(\tilde{g}_R) = E|\tilde{g}_R - g_R|^2 = \sum_{i=1}^{n} E(\tilde{b}_i - \beta_i)^2 = \sum_{i=1}^{n} \sigma^2(\tilde{b}_i) + \Delta^2 \tilde{b}_i ,$$

где $\Delta \tilde{b}_i = b_i - \beta_i$. Характеристики $\sigma^2(\tilde{b}_i)$, $\Delta^2 \tilde{b}_i$, а следовательно и $\delta^2(g_R)$, зависят от плана эксперимента. Пусть для примера $n = 3$. Тогда для плана 3-КУБ:

$$\tilde{b}_i^{I} = \frac{\sum_{k=1}^{8} \tilde{y}_{(k)} \cdot t_{i(k)}}{8}, \quad \sigma^2(\tilde{b}_i^{I}) = \frac{\sigma^2(\tilde{y})}{8};$$

с учетом (9): $\Delta \tilde{b}_i^{I} = E(\tilde{b}_i^{I} - \beta_i) = \dfrac{\sum_{k=1}^{8} y_{(k)} \cdot t_{i(k)}}{8} - \beta_i =$

$= \beta_{i11} + \beta_{i22} + \beta_{i33} + \beta_{i111} + \beta_{i1122} + \beta_{i1133} + \beta_{i2222} + \beta_{i2233} + \beta_{i3333} + \ldots$ и $N^{I} = 8$;

для плана 3-КРЕСТ: $\tilde{b}_i^{II} = \dfrac{\tilde{y}_{(2i-1)} - \tilde{y}_{(2i)}}{2}$, $\sigma^2(\tilde{b}_i^{II}) = \dfrac{\sigma^2(\tilde{y})}{2}$;

с учетом (9): $\Delta \tilde{b}_i^{II} = E(\tilde{b}_i^{II} - \beta_i) = \dfrac{y_{(2i-1)} - y_{(2i)}}{2} - \beta_i =$

$= \beta_{iii} + \beta_{iiii} + \ldots$ и $N^{II} = 6$.

Отсюда видно, что $\sigma^2(\tilde{b}_i^{I}) < \sigma^2(\tilde{b}_i^{II})$, но $\Delta^2 \tilde{b}_i^{I}$ может, наоборот, значительно превосходить $\Delta^2 \tilde{b}_i^{II}$ и кроме того $N^{I} > N^{II}$. Поэтому априори нельзя сказать, какой из этих планов даст меньшую общую ошибку. Поскольку коэффициенты разложения функций $y = y(t)$ и $y = f(x)$ в степенной ряд связаны соотношениями: $\beta_{ij\ldots q} = \Delta x_i \cdot \Delta x_j \ldots \Delta x_q \cdot \alpha_{ij\ldots q}$, то с увеличением интервалов варьирования это расхождение в точности может заметно возрастать, а дисперсия $\tilde{b}_i$ не изменяется. Для дробных реплик плана n -КУБ коэффициенты регрессии определяются еще менее точно.

Более того, покажем, что план n -КУБ (или его дробная реплика субпорядка m) и план n -КРЕСТ, точки которых расположены на одной сфере, т.е. равноудалены от центра, при равном количестве опытов дают идентичные случайные ошибки в направлении градиента. Разумеется, что для сравнения точности этого направления в обоих случаях необходимо привести значения коэффициентов регрессии в одну систему координат, например, в $\langle X \rangle$. Пусть K_1 и K_2 - количество опытов в каждой точке I-ого и 2-ого плана. Тогда $2^m K_1$ и $2n K_2$ - общее количество используемых опытов для этих планов. Дисперсии коэффициентов регрессии $\tilde{b}_i$, преобразованных в $\langle X \rangle$:

$$\sigma^2(\tilde{a}_i^{I}) = \sigma^2(\tilde{b}_i^{I}/\Delta x_i) = \sigma^2(\tilde{y})/2^m K_1 \cdot \Delta^2 x_i ;$$

$$\sigma^2(\tilde{a}_i^{II}) = \sigma^2(\tilde{b}_i^{II}/\sqrt{n} \cdot \Delta x_i) = \sigma^2(\tilde{y})/2K_2 \cdot n\, \Delta^2 x_i .$$

Поэтому $\sigma^2(\tilde{a}_i^{I}) = \sigma^2(\tilde{a}_i^{II})$, если $2^m K_1 = 2n K_2$.

Учитывая вышеизложенное, а также произвольность выбора масштабов по осям, можно сделать вывод, что никакого априорного преимущества в отношении точности оценки направления градиента (по крайней мере, в условиях лабораторного эксперимента) план n -КУБ не имеет в сравнении с планом n -КРЕСТ. С другой стороны, план n -КРЕСТ имеет ряд преимуществ: I) варьировать только один из факторов, а остальные поддерживать на заданном нулевом уровне, значительно проще; 2) при неудачном выборе масштаба по оси (интервала варьирования фактора) и его замене новым для плана n -КУБ нужно переделать опыты во всех точках, а для плана n -КРЕСТ только в 2-х; 3) количество точек в плане n -КУБ или его дробных репликах растет значительно быстрее с увеличением числа факторов, чем в плане n -КРЕСТ; 4) при выявлении значимых коэффициентов регрессии $\tilde{b}_{ii}$, как показано далее, план n -КРЕСТ дает возможность применить более точную регрессионную модель 7, что значительно увеличивает скорость сходимости процедуры.

Нетрудно видеть, что некоторые недостатки плана n -КУБ относятся и к плану n -СИМПЛЕКС, хотя последний имеет минимально возможное число точек для оценки направления градиента. Кроме того, расположение симплекса в факторном пространстве несимметрично относительно всех координатных осей, что создает неопределенность в выборе плана.

Других простейших ортогональных планов I-ого порядка не имеется. Все это позволяет сделать обоснованный вывод, что для статистической оценки градиента, по крайней мере в лабораторных условиях, наиболее рациональным является план n -КРЕСТ.

При переходе от лабораторных условий к опытно- промышленным и промышленным область оптимума может значительно смещаться. Могут также появиться новые факторы (влияющие параметры). Основная задача промышленного эксперимента в рассматриваемом аспекте состоит в том, чтобы эволюционно корректировать факторы, влияющие на управляемый объект. Этому направлению в планировании эксперимента положила начало работа также Дж.Бокса /5; 2, с.438-446/. Согласно методу эволюционного планирования Бокса (ЭВОП) движение к экстремуму функции отклика или ее постепенное улучшение осуществляют путем использования направления статистического градиента, находимого по плану n -КУБ или его подходящей дробной реплике. Выбор этих планов в данном случае оправдывают тем, что возможные интервалы варьирования факторов незначительны и строго регламентированы (ограничены), а дисперсия воспроизводимости слишком высока.

С другой стороны, в промышленных условиях точное задание факторов, предусмотренное планом, часто практически неосуществимо из-за их колебаний. В таком случае статистический градиент функции отклика, на наш взгляд, более целесообразно находить по тому случайному плану, который удается реализовать, фиксируя одновременно значения функции отклика и влияющих параметров. Дополнительное требование для корректного расчета в данном случае - отсутствие заметного инерционного запаздывания в изменении функции отклика от изменения параметров.

Последний из обсуждаемых вопросов заключается в выборе масштабов по осям. Они оказывают большое влияние на скорость сходимости процедуры поиска экстремума. В отсутствии информации о вторых производных масштабы совершенно случайны. Рассмотрим начало оптимизации.

Градиент $g_0 = (\beta_1, \dots, \beta_n) \in \langle t \rangle$ в пространстве $\langle X \rangle$ представляется вектором $\tau_0 = M g_0 = (\Delta x_1 \beta_1, \dots, \Delta x_1 \beta_n) = M^2 j_0 = (\Delta^2 x_1 \alpha_1, \dots, \Delta^2 x_n \alpha_n)$, где α_i являются инвариантами масштабов. Подходящим выбором вектору τ_0 придается любое направление в пределах одного (градиентного) n-квадранта. Следовательно, в данном случае градиент определяется только n-квадрант, в котором функция при движении из начальной точки возрастает во всех направлениях:

$$dy = j_0 dx = |j_0| \cdot |dx| \cos \gamma_x = g_0 dt = |g_0| \cdot |dt| \cdot \cos \gamma_t ,$$

где γ_x и $\gamma_t < 90^0$ – углы между градиентом и направлением движения. В других n –квадрантах характер изменения функции зависит от этого угла. Можно также убедиться в отличии профилей пути по градиенту в $\langle t \rangle$ и $\langle x \rangle$ (если, конечно, масштабы не прямо пропорциональны):

$$\Psi_t(\varepsilon) = y_t(\varepsilon g_0) = y_0 + g_0(\varepsilon g_0) + \frac{1}{2} G_0(\varepsilon g_0)^2 + \frac{1}{6} V_0(\varepsilon g_0)^3 + \ldots = y_0 + J g_0^2 \varepsilon + \frac{1}{2} G_0 g_0^2 \varepsilon^2 + \frac{1}{6} V_0 g_0^3 \varepsilon^3 + \ldots,$$

$$\left(\frac{d\Psi_t(\varepsilon)}{d\varepsilon}\right)_0 = J g_0^2 = |g_0|^2;$$

$$\Psi_x(\varepsilon) = y_x(\varepsilon j_0) = y_0 + j_0(\varepsilon j_0) + \frac{1}{2} J_0(\varepsilon j_0)^2 + \frac{1}{6} W(\varepsilon j_0)^3 + \ldots = y_0 + M^{-2} g_0^2 \varepsilon + \frac{1}{2}(M^{-2} G_0 M^{-2}) g_0^2 \varepsilon^2 + \frac{1}{6}(M^{-2} \overset{M^{-2}}{V_0} M^{-2}) g_0^3 \varepsilon^3 +$$

$$+ \ldots, \quad \left(\frac{d\Psi_x(\varepsilon)}{d\varepsilon}\right)_0 = M^{-2} g_0^2 = |M^{-1} g_0|^2.$$

Итак, в начале оптимизации имеется неопределенность в направлении движения в пределах одного n -квадранта, которая обусловлена случайным выбором масштабов. Хотя это затем и отражается в сильной мере на скорости сходимости, но ее порядка не изменяет: при любом масштабе метод остается тем же. В случае неудачного выбора масштаба скорость сходимости (с заданной точностью) может быть сколь угодно медленной.

X X X

Дальнейшее уточнение направления движения к экстремуму, а, следовательно, и увеличение скорости сходимости, возможно лишь с применением более точных регрессионных моделей.

Движение к экстремуму с применением модели 6 (или 2) основано на численном методе Ньютона, использующем приближение:

$$y(t) \approx y_R + g_R t + \frac{I}{2} G_R t^2 = Q(t)$$

Экстремум $Q(t)$, если он существует, определяется вектором Ньютона $P_R = -G_R^{-1} g_R$, находимым из условия: $\frac{dQ}{dt} = G_R t + g_R = 0$. Когда G_R либо отрицательна (при поиске максимума), либо положительна (при поиске минимума), приближение $y \approx Q(t)$ с геометрической точки зрения представляет собой аппроксимацию выпуклой или вогнутой поверхности отклика n -поверхностью эллиптического параболоида (ЭП). Главные эллиптические оси ЭП параллельны собственным векторам матрицы Гессе, а параболическая ось параллельна оси y .

В свою очередь, для отрицательной G_R все $\left(\frac{\partial^2 y}{\partial t_i^2}\right)_R < 0$, для положительной G_R все $\left(\frac{\partial^2 y}{\partial t_i^2}\right)_R > 0$. Ввиду этого выпуклую или вогнутую (хотя бы в области плана) поверхность отклика можно аппроксимировать с промежуточной точностью между моделями 5 и 6 n -поверхностью ЭП, все главные оси которого параллельны осям координат. Аналитически это выражается приближением:

$$y(t) \approx y_R + g_R t + \frac{I}{2} D_R t^2 = q(t).$$

Экстремум $q(t)$, если он существует, определяется вектором $h_R = -D_R^{-1} g_R$. Учитывая вышесказанное с $\left(\frac{\partial^2 y}{\partial t_i^2}\right)_R$, можно видеть, что в рассматриваемых случаях вектор h_R расположен либо в градиентном, либо антиградиентном n -квадранте. Метод, основанный на движении к экстремуму по вектору h_R , аналитически тождественнен масштабной модификации метода Коши. Естественно, что скорость его сходимости занимает промежуточное положение по отношению к методам Коши и Ньютона.

Можно убедиться в инвариантности методов 2-ого порядка к выбираемым масштабам. Вектор Ньютона (или вектор h_R) преобразуется в соответствии с изменением масштабов:

$$P_R = -G_R^{-1} g_R = -(M J_R M)^{-1} M j_R = -M^{-1}(J_R^{-1} j_R).$$

Профили пути по векторам в $\langle t \rangle$ и $\langle X \rangle$ тождественны:

$$\psi(\varepsilon) = y_t(\varepsilon P_R) = y_R + g_R P_R \cdot \varepsilon + \frac{I}{2} G_R P^2 \cdot \varepsilon^2 + \frac{I}{6} V_R P_R^3 \cdot \varepsilon^3 + \ldots =$$

$$= y_R + j_R (M P_R) \cdot \varepsilon + \frac{I}{2} J_R (M P_R)^2 \varepsilon^2 + \frac{I}{6} W_R (M P_R)^3 \cdot \varepsilon^3 + \ldots =$$

$$= y_x(\varepsilon M P_R) ;$$

$$\frac{d\psi(\varepsilon)}{d\varepsilon} = g_R P_R = j_R (M P_R) .$$

Если $y(t) \equiv Q(t)$, то $\varepsilon_{extr} = I$.

Соответственно разностно-статистические модификации этих методов используют приближения:

$$y(t) \approx \tilde{y}_R + \tilde{g}_R t + \frac{I}{2} \tilde{G}_R t^2 = \tilde{Q}(t)$$

$$y(t) \approx \tilde{y}_R + \tilde{g}_R t + \frac{I}{2} \tilde{D}_R t^2 = \tilde{q}(t)$$

Возможно, что процедура приведет не в экстремальную, а седловую точку S. Тогда дальнейший выбор направления движения становится неопределенным в пределах подпространства, заданном точкой S и аномальными собственными векторами матрицы $\tilde{G}_S$ или $\tilde{D}_S$. Задача поиска следующего промежуточного экстремума становится многозначной. На наш взгляд, для ее упрощения можно изучить поведение функции при движении из точки S по каждому аномальному собственному вектору. В частности, если таковой один, то поиск следующего промежуточного экстремума является одномерной задачей. Если же функция $\tilde{Q}(t)$ или $\tilde{q}(t)$ имеет противоположный экстремум, то методы 2-ого порядка не пригодны.

В случае плохо обусловленной $\tilde{G}_R$ (или $\tilde{D}_R$), т.е. $\det \tilde{G}_R \approx 0$, рассматриваемые методы 2-ого порядка дают явно некорректные результаты: неточность направления движения значительно превышает неточность коэффициентов модели. Совершенно неопределенным становится решение задачи при $\det \tilde{G}_R = 0$. Чтобы избежать этой неопределенности и сделать методы в этом смысле универсальными, достаточно их модифицировать регуляризацией по Тихонову /4/. Корректным в смысле Тихонова решением уравнения $\left|\frac{d\tilde{Q}}{dt}\right| = \left|\tilde{g}_R + \tilde{G}_R t\right| = min$ является решение $\tilde{t}^{(\tau)}$, находимое либо путем численной минимизации функции Тихонова: $Th(t,\tau) = |\tilde{g}_R + \tilde{G}_R t|^2 + \tau |t|^2$; либо по итоговой формуле: $\tilde{t}^{(\tau)} = -(\tilde{G}_R^2 + \tau J)^{-1} G_R g_R$, где $\tau > 0$ (так как $\lim\limits_{\tau \to 0} \tilde{t}^{(\tau)}$ есть нормальное решение этого же уравнения и $\det(\tilde{G}_R^2 + \tau J) = 0$). Обращение заведомо определенной матрицы можно упростить. Значение параметра τ и неточность в определении коэффициентов модели должны быть согласованы так, чтобы стремиться к нулю одновременно. Причем коэффициенты матрицы определяют значительно менее точно, чем у градиента.

Диагональные элементы $\tilde{G}_k^2$ есть сумма $\sum_{j=1 (j\neq i)}^{n} \tilde{b}_{ij}^2 + 4\tilde{b}_{ii}^2$; диагональные элементы $\tilde{D}_k^2$ есть $4\tilde{b}_{ii}^2$. Поэтому целесообразно принять в первом случае: $\tau = 4s^2(\tilde{b}_{ii}) + (n - 1)\cdot s^2(\tilde{b}_{ij})$, а во втором случае: $\tau = 4s^2(\tilde{b}_{ii})$.

Процедуру экспериментальной оптимизации завершают, когда регрессионная модель 6 имеет удовлетворительную относительную погрешность. Тогда можно полагать, что точность оценки экстремума по модели 6 (с регуляцией по Тихонову) соответствует точности экспериментальной оценки функции отклика. Дополнительно можно потребовать, чтобы оцениваемый экстремум принадлежал области последнего плана, а оцениваемое по модели экстремальное значение функции соответствовало с имеющейся точностью экспериментальной оценке.

§ 3. Изложение методов.

На первом этапе оптимизации осуществляют статистическую проверку гипотез 2 и 3, одновременно оценивая $s^2(\tilde{y})$. Заключительная стадия – оценка экстремума по модели 6, имеющей допустимую погрешность ρ. Эти окаймляющие операции в изложении методов опускаются.

<u>Метод I-ого порядка</u>. Процедура основана на последовательном движении вдоль (или против) векторов $\tilde{g}_k$, находимых по модели 5; коэффициенты модели устанавливают по плану n – КРЕСТ (таблица I).

Согласно МНК и модели 5: $\tilde{b}_i = \dfrac{\tilde{y}_{(2i-1)} - \tilde{y}_{(2i)}}{2}$

$$s^2(\tilde{b}_i) = \frac{s^2(\tilde{y})}{2\,\xi^+} \qquad (II)$$

Таблица I

План n -КРЕСТ

Номер точки, k	t_1	t_2		t_{n-1}	t_n	$\tilde{y}_{(k)}$	$\bar{z}_{(k)}$
0	0	0	...	0	0	$\tilde{y}_0 = \tilde{b}_0$	$\bar{z}_0 = \tilde{b}_0$
I	+I	0	...	0	0	$\tilde{y}_1$	$\bar{z}_1$
2	-I	0	...	0	0	$\tilde{y}_2$	$\bar{z}_2$
...	...	...		...		...	...
$2n$ - I	0	0	...	0	+I	$\tilde{y}_{(2n-1)}$	$\bar{z}_{(2n-1)}$
$2n$	0	0	...	0	-I	$\tilde{y}_{(2n)}$	$z_{(2n)}$

Условие значимости коэффициентов регрессии:

$$|\tilde{b}_i| \geq T_{кр}|s(\tilde{b}_i)| = T_{кр}\frac{s(\tilde{y})}{\sqrt{2\,\xi^+}}\,, \quad |\tilde{b}_0| \geq T_{кр}\frac{s(\tilde{y})}{\sqrt{\xi^\cdot}},$$

где $T_{кр}$ - критическое значение критерия Стьюдента для принятого уровня вероятности и числа степеней свободы $s^2(\tilde{y})$. Незначимый коэффициент $\tilde{b}_i$, если он относится к влияющему фактору, можно сделать значимым, увеличив интервал варьирования Δx_i. Затем находят относительную погрешность модели ρ и делают вывод об удовлетворительности данной аппроксимации.

После получения удовлетворительной оценки градиента далее проводят одномерную экспериментальную оптимизацию функции

$\tilde{y} = \tilde{y}(\varepsilon \tilde{g}_0) = \tilde{\psi}(\varepsilon)$, причем $\tilde{\psi}(0) = \tilde{y}_{(0)}$, и т.д. Движение прекращают, когда все коэффициенты $\tilde{b}_i$ становятся незначимыми.

Квадратичный метод. Процедура основана на последовательном движении вдоль векторов $\tilde{h}_k^{(r)}$, находимых по модели 7; коэффициенты модели устанавливают по плану n -КРЕСТ.

Согласно МНК и модели 7 коэффициенты $\tilde{b}_0$ и $\tilde{b}_i$ вычисляют, как и раньше;

$$\tilde{b}_{ii} = \frac{\tilde{y}_{(2i-1)} + \tilde{y}_{(2i)} - 2\tilde{y}_{(0)}}{2}$$

$$s^2(\tilde{b}_{ii}) = \left(\frac{I}{\xi^{\cdot}} + \frac{I}{2\xi^{+}}\right) \cdot s^2(\tilde{y}) . \qquad (I2)$$

Фактор x_i исключают только тогда, когда как $\tilde{b}_i$, так и $\tilde{b}_{ii}$ незначимы, даже с увеличением Δx_i . Далее, как и выше, устанавливают удовлетворительность аппроксимации.

Следует иметь ввиду, что для неопределенной матрицы $\tilde{D}_k$ (и вообще $\tilde{G}_k$):

$$\begin{aligned} &\tilde{b}_{ii} \leq 0 \text{ (поиск максимума)} \\ &\tilde{b}_{ii} \geq 0 \text{ (поиск минимума)} \end{aligned} \qquad (I3)$$

Движение осуществляют по вектору $\tilde{h}_k^{(r)}$ с координатами:

$$h_{ik}^{(r)} = - \frac{0{,}5 \cdot \tilde{b}_i \cdot \tilde{b}_{ii}}{\tilde{b}_{ii}^2 + s^2(\tilde{b}_{ii})} .$$

Если все значимые коэффициенты $\tilde{b}_{ii}$ имеют противоположный знак, то остается использовать метод I-ого порядка. В случае знакопеременных коэффициентов $\tilde{b}_{ii}$ предварительно устанавливают знак произведения $\tilde{g}_k \cdot \tilde{h}_k^{(r)}$ (при поиске максимума) или -

$-\tilde{g}_k \cdot \tilde{h}_k^{(r)}$ (при поиске минимума). Если он положителен (т.е. угол между векторами меньше 90^0 и функция изменяется в нужную сторону), то можно использовать данный метод; в противоположном случае применяют метод I-ого порядка.

Метод 2-ого порядка. Процедура основана на последовательном движении вдоль векторов $\tilde{P}_k^{(r)}$, находимых по модели 6; коэффициенты модели устанавливают по n -ЦКП (таблица 2).

Таблица 2

Центральный композиционный план (в полной форме)

Номер точки, k	t_1	t_2		t_{n-1}	t_n	$\tilde{y}_{(k)}$	$\bar{z}_{(k)}$
0	0	0	...	0	0	$\tilde{y}_0 = \tilde{b}_0$	$\bar{z}_0 = \tilde{b}_0$
I	$+\alpha$	0	...	0	0	$\tilde{y}_1$	$\bar{z}_1$
2	$-\alpha$	0	...	0	0	$\tilde{y}_2$	$\bar{z}_2$
...	...	...	...	...	...	...	...
$2n$ - I	0	0	...	0	$+\alpha$	$\tilde{y}_{(2n-1)}$	$\bar{z}_{(2n-1)}$
$2n$	0	0	...	0	$-\alpha$	$\tilde{y}_{(2n)}$	$\bar{z}_{(2n)}$
$2n$ + I	+I	+I	...	+I	+I	$\tilde{y}_{(2n+1)}$	$\bar{z}_{(2n+1)}$
$2n$ + 2	+I	+I	...	+I	-I	$\tilde{y}_{(2n+2)}$	$\bar{z}_{(2n+2)}$
...	...	...	...	...	...	...	...
$2n$ +2^n -I	+I	-I		-I	-I	$\tilde{y}_{(2n+2^n-1)}$	$\bar{z}_{(2n+2^n-1)}$
$2n$ +2^n	-I	-I		-I	-I	$\tilde{y}_{(2n+2^n)}$	$\bar{z}_{(2n+2^n)}$

При использовании n -ЦКП с дробной репликой (от плана n -КУБ) последняя не должна смешивать коэффициенты b_{ij}. Это достигают использованием при ее построении генерирующих соотношений типа $t_j = \pm t_p \cdot t_q \ldots t_r$, где число сомножителей больше 3-х. Соответственно определяющие контрасты, т.е. произведения $\pm t_p \cdot t_q \cdot \ldots t_r \cdot t_j$, содержат более 4-х сомножителей.

Чтобы избежать ненужных опытов, планирование целесообразно осуществлять в два этапа. Сначала ставят I-ую часть эксперимента по плану n -КРЕСТ и ориентировочно находят коэффициенты $\tilde{b}_i$ и $\tilde{b}_{ii}$:

$$\tilde{b}_i^{II} = \frac{\tilde{y}_{(2i-1)} - \tilde{y}_{(2i)}}{2\alpha}, \quad s^2(\tilde{b}_i^{II}) = \frac{I}{2\alpha^2 \cdot \xi^+} \cdot s^2(\tilde{y}),$$

$$\tilde{b}_{ii}^{II} = \frac{\tilde{y}_{(2i-1)} + \tilde{y}_{(2i)} - 2\tilde{y}_{(0)}}{2},$$

$$s^2(\tilde{b}_{ii}^{II}) = \frac{\frac{2}{\xi^{\circ}} + \frac{I}{\xi^+}}{2\alpha^4} \qquad (I4)$$

Если условие I3 не выполняется, то дальнейшее движение осуществляют по вектору $\tilde{h}_R^{(r)}$ или $\pm\ \tilde{g}_R$. На этом же этапе можно, когда требуется, улучшить масштабы по осям. Если $\tilde{b}_i^{II} \neq 0$, а $\tilde{b}_{ii}^{II}$ незначимым, то увеличивают Δx_i , так как $\beta_{ii} = \alpha_{ii} \cdot \Delta^2 x_i$. Если $\tilde{b}_i^{II}$ и $\tilde{b}_{ii}^{II}$ оба незначимы даже с увеличением Δx_i , то фактор x_i исключают.

Далее приступают ко 2-ой части эксперимента по плану n -КУБ (или его дробной реплике). Согласно МНК и модели 6 для n -ЦКП коэффициенты $\tilde{b}_{ij}$ вычисляют только по опытам в вершинах n -куба:

$$\tilde{b}_{ij}^{III} = \tilde{b}_{ij}^{I} = \tilde{b}_{ij} = \frac{\sum\limits_{k=2n+1}^{2n+2^m} t_{i(k)} \cdot t_{j(k)} \cdot \tilde{y}_{(k)}}{2^m}$$

$$s^2(\tilde{b}_{ij}) = \frac{s^2(\tilde{y})}{2^m \cdot \xi^{\square}} \qquad (I5)$$

Остальные коэффициенты вычисляют также, исходя из МНК, по формулам /2, с.141-146/:

$$\tilde{b}_i = \sum_{k=1}^{2n+2^m} t_{i(k)} \cdot \tilde{y}_{(k)} \Big/ \sum_{k=1}^{2n+2^m} t_{i(k)}^2 , \tag{16}$$

где $\sum_{k=1}^{2n+2^m} t_{i(k)}^2 = 2^m + 2\alpha^2,$

$$s^2\left(\tilde{b}_i^{\overline{III}}\right) = \frac{2^m / \xi^{\square} + 2\alpha^2 / \xi^{+}}{\left(2^m + 2\alpha^2\right)^2} \cdot s^2(\tilde{y}) . \tag{17}$$

Коэффициенты b_{ii} оставляют ранее найденными по формуле 14, либо в случае ортогонального плана вычисляют по формуле:

$$\sum_{k=0}^{2n+2^m} t_{i(k)}^{*} \cdot \tilde{y}_{(k)} \Big/ \sum_{k=0}^{2n+2^m} \left(t_{i(k)}^{*}\right)^2 , \tag{18}$$

где $t_{i(k)}^{*} = t_{i(k)}^2 - \dfrac{\sum_{k=1}^{2n+2^m} t_{i(k)}^2}{1 + 2n + 2^m} = t_{i(k)}^2 - \dfrac{2^m + 2\alpha_{\perp}^2}{N^{\overline{III}}}$

- вспомогательные переменные, вводимые для ортогонализации матрицы планирования;

$$\sum_{k=0}^{2n+2^m} \left(t_{i(k)}^{*}\right)^2 = 2\alpha_{\perp}^4 = S \tag{19}$$

$$s^2\left(\tilde{b}_{ii}^{\overline{III}}\right) = \left(\frac{1}{S} - \frac{\xi' - \xi^{\square}}{\xi'} \cdot V_1 - \frac{\xi^{+} - \xi^{\square}}{\xi^{+}} \cdot V_2\right) \cdot \frac{s^2(\tilde{y})}{\xi^{\square}} , \tag{20}$$

где $V_1 = \dfrac{2^m}{S^2 \cdot N^{\overline{III}}}$, $\quad V_2 = \dfrac{N^{\overline{III}} - 2}{S \cdot N^{\overline{III}}} - V_1$, (значения параметров для ряда простейших планов даны в таблице 3 и 4).

Таблица 3

Основные параметры центральных композиционных планов с пробной репликой

n	5	6	7	8
m	4	5	6	6
$N = 1 + 2n + 2^m$	27	45	79	81
$B = \frac{(n+1)\cdot(n+2)}{2}$	21	28	36	45
генерирующие соотношения	$t_5 = \pm t_1 t_2 t_3 t_4$	$t_6 = \pm t_1 t_2 t_3 t_4 t_5$	$t_7 = \pm t_1 t_2 t_3 t_4 t_5 t_6$	$t_7 = \pm t_1 \cdot t_2 t_3 \cdot t_4$ $t_8 = \pm t_1 \cdot t_2 \cdot t_5 \cdot t_6$

Таблица 4

Вычислительные параметры простейших планов Бокса-Уилсона

n	2	3	4	5	5
m	2	3	4	5	4
$= 1 + 2n + 2^m$	9	15	25	43	27
$\alpha_\perp$	1	1,215	1,414	1,596	1,547
$S = 2\alpha_\perp^4$	2	4,356	8	12,98	11,45
g_1	0,5556	0,4340	0,3600	0,3100	0,2958
g_2	0,1975	0,2089	0,2048	0,1979	0,1820
V_1	0,1111	0,02811	0,01000	0,004419	0,004520
V_2	0,2778	0,1709	0,1050	0,06906	0,07635

И в данном случае перед движением по вектору $\tilde{P}_R^{(\tau)}$ предварительно находят знак произведения $\tilde{g}_R \cdot \tilde{P}_R^{(\tau)}$ (при поиске максимума) или $-\tilde{g}_R \cdot \tilde{P}_R^{(\tau)}$ (при поиске минимума). Если чункция изменяется не в нужную сторону, то применяют квадратичный метод.

Замечания.

I) Дисперсию адекватности оценивают по общей для всех методов формуле:

$$s^2(\tilde{y},\bar{Z}) = \frac{\sum_{\kappa=1}^{N-1} \xi_\kappa \cdot \left(\tilde{y}_{(\kappa)} - \bar{Z}_{(\kappa)}\right)^2}{\varphi_1},$$

где в вершинах n -куба $\xi_\kappa = \xi^{\square}$, в вершинах n -креста $\xi = \xi^+$; $\varphi_1 = N - B'$ - число степеней свободы (B' - количество значимых коэффициентов регрессии).

Дисперсию воспроизводимости можно оценить по общей формуле:

$$s^2(\tilde{y}) = \frac{\sum_{j=1}^{\xi^\circ}\left(\tilde{y}_{(0),j} - \tilde{y}_{(0)}\right)^2 + \sum_{\kappa=1}^{2n}\sum_{j=1}^{\xi^+}\left(\tilde{y}_{(\kappa),j} - \tilde{y}_{(\kappa)}\right)^2 + \sum_{\kappa=2n+1}^{2n+2^m}\sum_{j=1}^{\xi^{\square}}\left(\tilde{y}_{(\kappa),j} - \tilde{y}_{(\kappa)}\right)^2}{\varphi_2},$$

для которой число степеней свободы: $\varphi_2 = \xi^\circ + \xi^+ 2n + \xi^{\square} 2^m - N$.

2) Допустимая величина относительной погрешности модели устанавливается с учетом того, что: а) она не должна быть меньше относительной погрешности определения самой функции $|s(\tilde{y})| / (\tilde{y}_{max} - \tilde{y}_{min})$; б) при движении по векторам требование к точности модели меньше, чем при оценке стационарной точки по модели 6.

3) В последнем случае при удовлетворительной погрешности модели 6 оценивают стационарную точку, как $\tilde{t}_s^{(\tau)} = \tilde{P}_{(s-1)}^{(\tau)}$. Если она находится в области плана и достаточные условия экстремума выполнены, то остается сравнить в ней экспериментальное и теоретическое значение y .

4) При необходимости точность модели 6 повышают в зависимости от относительного вклада систематической ошибки либо изменением интервалов варьирования факторов (можно последовательно по одному), либо увеличением количества опытов в точках плана (можно в последовательности: центр, вершины n -креста, вершины n -куба), добиваясь на каком-нибудь этапе требуемой точности модели.

<u>Методы промышленной эволюционной оптимизации (ЭВОП)</u>. Движение к экстремуму или постепенное улучшение y осуществляют по векторам: $\pm\ \tilde{g}_R$, $\tilde{h}_R^{(\tau)}$ или $\tilde{P}_R^{(\tau)}$. Их координаты определяют посредством плана, который удается реализовать в окрестности R , и соответствующей модели 5, 7 или 6 (причем $N > B$). Два крайних случая: активный эксперимент, т.е. возможно весьма точное задание факторов по плану; пассивный эксперимент, т.е. возможно лишь одновременно измерять значения функции отклика и всех факторов, которые совершают спонтанные колебания около R . В общем же случае при реализации эксперимента следует стремиться по возможности к осуществлению вполне конкретного плана. Эта возможность и определит степень случайности плана.

Рассмотрим решение, общее для всех методов. Допустим, что инерционное запаздывание y по отношению к изменяющимся t_i пренебрежимо мало (равновесное планирование). Дополнительно обозначим:

$\bar{Z} = z(t) \cdot \tilde{b}$ - представление регрессионной модели линейной функцией $\tilde{b}$;

$\tilde{b} = (\ \tilde{b}_0,\ \tilde{b}_1\ ,\ldots) - K$ - вектор коэффициентов регрессии;

$\tilde{y} = (\ \tilde{y}_1\ ,\ldots,\ y_N\ )$ и $\bar{Z} = (\ \bar{Z}_1\ ,\ldots,\ \bar{Z}_N) - N$ -векторы;

R - $N \times K$ - матрица со строками $r(t_1), \ldots, r(t_N)$.

Согласно МНК: $|\tilde{y} - \bar{z}| = |\tilde{y} - R\tilde{b}| = min > 0$. Весьма маловероятно, что $rang R < K$. Поэтому $\tilde{b} = R^{-} \cdot \tilde{y} = (R' \cdot R)^{-1} \cdot R' \tilde{y}$. Если окажется, что $R' \cdot R$ плохо обусловлена, то лучше изменить план, например, добавлением еще нескольких точек.

Когда имеется предварительная оценка $s^2(\tilde{y})$, то далее проверяют значимость коэффициентов регрессии посредством диагональных элементов дисперсионной матрицы $C = (R' \cdot R)^{-1}$, а именно: $s^2(\tilde{b}_i) = c_{ii} \cdot s^2(\tilde{y})$, $i = 0, 1, \ldots$. Удовлетворительность аппроксимации оценивают аналогично по критерию $\rho =$ $= |s_{ad}(\tilde{y}, \bar{z}) / (\tilde{y}_{max} - \tilde{y}_{min})$, учитывая, что

$$s^2_{ad}(\tilde{y}, \bar{z}) = \frac{1}{\varphi_1} |\tilde{y} - \bar{z}|^2 .$$

Основной недостаток спонтанных планов состоит в том, что в силу корреляции между коэффициентами регрессии (согласно $K^2_{ij} = \frac{c^2_{ij}}{c_{ii} \cdot c_{jj}}$) последние вычисляют зависимо друг от друга, т.е. ошибки в их определении возрастают. Однако все равно имеется возможность оценить точность получаемой модели. Выбор же конкретного метода обусловлен степенью случайности плана и точностью измерений.

Замечания.

1) Представление $\bar{z} = r(t) \cdot \tilde{b}$ должно лишь аналитически быть тождественным модели, поскольку для приведения R к специальному виду иногда требуется тождественное преобразование модели.

2) Корреляция между коэффициентами регрессии отсутствует тогда и только тогда, когда матрица $R' \cdot R$ является диагональной, т.е. R принадлежит к некоторому специальному типу

матриц, включающему и ортогональные матрицы.

Методы поиска условного экстремума (при ограничении типа равенства).

Пусть $\langle t\rangle_\ell \ni \forall t : At = a$ - заданное линейное ограничение. Тогда в точке $t_R \in \langle t\rangle_\ell$ статистические векторы $\pm\,\tilde{g}_R$, $\tilde{h}_R^{(\tau)}$ и $\tilde{P}_R^{(\tau)}$ можно преобразовать в соответствующие условные векторы: $\tilde{g}_{u_R} = \vec{A}\,\tilde{g}_R$, $\tilde{P}_{u_R}^{(\tau)} = -\vec{A}\left(\tilde{G}_R^2 + \tau J\right)^{-1}\tilde{G}_R \vec{A}\,\tilde{g}_R$,

$\tilde{h}_{u_R}^{(\tau)} = -\vec{A}\left(\tilde{D}_R^2 + \tau J\right)\tilde{D}_R \vec{A}\,\tilde{g}_R$, которые определяют направление движения из точки R согласно условным методам, соответствующим вышерассмотренным. Матрица $\vec{A}$ в нормированном пространстве $\langle t\rangle$ осуществляет ортопроекцирование его элементов на линейное подпространство $\langle \ker A\rangle$. Она вычисляется либо по алгоритму, либо по формуле.

В точке $\tilde{t}_{u_S}$, куда приводит процедура, о характере условной стационарности судят по вырожденной условной матрице Гессе $\vec{A}\,G\,\vec{A}$.

В случае нелинейного ограничения $C(t) = 0$ метод I-ого порядка трансформируется в соответствующий условный метод путем использования при движении направления ортопроекции линии условного градиента на поверхность $C(t) = 0$. Уравнение линии условного статистического градиента, исходящей из точки R, в параметрической форме: $\tilde{t} = t_R + \varepsilon\cdot\tilde{g}_{u_R} = \tilde{t}_{(\varepsilon)}$. Искомую ортопроекцию находят, решая задачу на условный минимум функции $q(t) = J(t - \tilde{t})^2$ при ограничении $C(t) = 0$. ЕЕ решение в дискретной форме: $t = t_{(\varepsilon_k)}$, где $k = 1,2\ldots$ вплоть до промежуточного экспериментального экстремума и т.д.

§4. О точности коэффициентов регрессии.

Отклонение $\tilde{b}$ от β состоит из случайной и систематической ошибок: $\tilde{b}-\beta = (\tilde{b} - b) + (b - \beta)$. Первая из них характеризуется дисперсией $\sigma^2(\tilde{b})$, вторая - квадратом смещения $\Delta^2 \tilde{b}$. Рассмотрим последовательно точность приближений: $\beta_0 \approx \tilde{b}_0$, $\beta_i \approx \tilde{b}_i$, $\beta_{ij} \approx \tilde{b}_{ij}$ и $\beta_{ii} \approx \tilde{b}_{ii}$ - согласно первому, второму и третьему методу. Качество оценки $\tilde{b}$ характеризуется, исходя из дисперсии $\sigma^2(\tilde{b})$, а также порядка и количества коэффициентов β, содержащихся в смещении $\Delta \tilde{b}$.

Прежде всего в этих методах используется оценка $\tilde{b}_0 = \tilde{y}_{(0)}$. Для нее $\Delta \tilde{b}_0 = 0$, а $\sigma^2(\tilde{b}_0) = \frac{1}{\xi^{\cdot}} \sigma^2(\tilde{y})$. Согласно МНК и моделям 5 и 7 возможна также смещенная оценка в первых двух методах:

$$\tilde{b}_0 = \frac{\sum\limits_{k=1}^{2n} \tilde{y}_{(k)}}{2n}, \tag{21}$$

для которой $\sigma^2(\tilde{b}_0) = \frac{1}{2n \cdot \xi^{+}} \cdot \sigma^2(\tilde{y})$,

$$\Delta \tilde{b}_0 = \frac{1}{n} \sum_{k=1}^{\infty} \sum_{j=1}^{n} \beta_{\underbrace{jj\ldots jj}_{2k}} \tag{22}$$

т.е. $b_0 = \beta_0 + \frac{1}{n} (\beta_{11} + \ldots + \beta_{nn} + \beta_{1111} + \ldots)$. Это может проявляться в их существенном систематическом отклонении от β_0, несмотря на меньшую дисперсию. Промежуточными являются оценки типа:

$$b_0 = \frac{q \cdot \tilde{y}_{(0)} + \sum\limits_{k=1}^{2n} \tilde{y}_{(k)}}{q + 2n}$$

Аналогичный недостаток имеет оценка, используемая в методе Бокса-Уилсона:

$$\tilde{b}_0 = \tilde{b}_0^{I} = \frac{\sum\limits_{K=1}^{2^m} \tilde{y}_{(K)}}{2^m} ,$$

для которой $\sigma^2(\tilde{b}_0^{I}) = \frac{1}{2^m \cdot \xi^{\square}} \sigma^2(\tilde{y})$,

$$\Delta \tilde{b}_0 = \Delta \tilde{b}_0^{I} = \sum_{K=1}^{\infty} \sum_{l=1}^{m} \dots \sum_{p=1}^{q} \sum_{q=1}^{n} \beta_{\underbrace{ll\,mm \dots pp\,qq}_{2K}} \qquad (23)$$

$$(b_0^{I} = \beta_0 + \beta_{11} + \dots + \beta_{nn} + \beta_{1111} + \beta_{1122} + \dots)$$

Соответственно в третьем методе согласно МНК и модели 6 при $\alpha = \alpha_{\perp}$ справедлива также оценка /2, с.146/:

$$\tilde{b}_0^{III} = \frac{\sum\limits_{K=0}^{2n+2^m} \tilde{y}_{(K)} - (2^m + 2\alpha_{\perp}^2) \cdot \sum\limits_{i=1}^{n} \tilde{b}_{ii}^{III}}{N^{III}} \qquad (24)$$

которая смешивается с коэффициентом β , более чем 3-его порядка. Для нее

$$\sigma^2(\tilde{b}_0^{III}) = y_1 - \frac{\xi^{*} - \xi^{\square}}{\xi^{*}} y_1^2 - \frac{\xi^{+} - \xi^{\square}}{\xi^{+}} \cdot y_2 \quad \frac{\sigma^2(\tilde{y})}{\xi^{\square}} , \qquad (25)$$

где $y_1 = \frac{S + n \cdot 2^m}{S \cdot N^{III}}$, $y = 2n \left[\frac{2^m \cdot (n - \alpha_{\perp}^2)^2}{S \cdot N^{III}} \right]^2$,

(значения параметров для ряда простейших случаев даны в таблице 4).

$$\Delta \tilde{b}_0^{III} = \frac{2^m \cdot \Delta \tilde{b}_0^{I} + 2n \cdot \Delta \tilde{b}_0^{II} - (2^m + 2\alpha_{\perp}^2) \cdot \sum\limits_{i=1}^{n} \tilde{b}_{ii}^{III}}{N^{III}} \qquad (26)$$

где $\Delta \tilde{b}_0^{I}$ приведено выше,

$$\Delta \tilde{b}_0^{II} = \frac{1}{n} \sum_{K=1}^{\infty} \alpha^{2K} \sum_{j=1}^{n} \beta_{\underbrace{jj \dots jj}_{2K}} \qquad (27)$$

а $\Delta \tilde{b}_{ii}^{III}$ будут даны выше. Нетрудно представить, какой большой величины может достигать систематическое отклонение оценки $\tilde{b}_0^{III}$. С другой стороны, как видно из таблицы 5 (полученной с использованием формулы 25 и таблицы 4), в отношении дисперсии она не имеет существенного преимущества по сравнению с несмещенной оценкой $\tilde{b}_0 = \tilde{y}_{(0)}$.

В случае произвольного α справедлива оценка:

$$\tilde{b}_0^{III} = \frac{\sum\limits_{k=0}^{2n+2^m} \tilde{y}_{(k)} - \left(2^m + 2\alpha^2\right)\sum\limits_{i=1}^{n} \tilde{b}_{ii}^{II}}{N^{III}},$$

к которой относятся те же недостатки. Промежуточные оценки образуют аналогично.

Перейдем к оценке $\tilde{b}_i$. Для первого и второго метода $\sigma^2(\tilde{b}_i)$ выражается аналогично (II);

$$\Delta \tilde{b}_i = \sum_{k=1}^{\infty} \beta_{\underbrace{i \ldots ii}_{2k+1}} \tag{28}$$

Для третьего метода $\sigma^2(\tilde{b}_i^{III})$ выражается аналогично (I7);

$$\Delta \tilde{b}_i^{III} = \frac{2^m}{2^m + 2\alpha^2} \cdot \Delta \tilde{b}_i^{I} + \frac{2\alpha^2}{2^m + 2\alpha^2} \cdot \Delta \tilde{b}_i^{II}, \tag{29}$$

где $\Delta \tilde{b}_i^{I} = \sum\limits_{k=1}^{\infty} \sum\limits_{l=1}^{m} \ldots \sum\limits_{p=1}^{q} \sum\limits_{q=1}^{n} \beta_{\underbrace{illmm \ldots ppqq}_{2k+1}}$, (30)

$$\Delta \tilde{b}_i^{II} = \sum_{k=1}^{\infty} \alpha^{2k} \cdot \beta_{\underbrace{i \ldots ii}_{2k+1}} \tag{31}$$

Откуда видно, что $\Delta \tilde{b}_i^{III}$ занимает промежуточное положение между $\Delta \tilde{b}_i^{I}$ и $\Delta \tilde{b}_i^{II}$ и не сожержит коэффициентов β второго порядка. С другой стороны, в случае $\xi^{+} = \xi^{\square}$ дисперсия

Таблица 5

Соотношение дисперсий коэффициентов регрессии для простейших планов Бокса-Уилсона

n	2	3	4	5	$\xi^{\square}/\xi^{\bullet}$	$\xi^{\square}/\xi^{+}$
m	2	3	4	4		
$\dfrac{\sigma^2(\tilde{y}_{(0)})}{\sigma^2(\tilde{b}_0^{III})}$	I,8	2,3	2,8	3,4	I	I
	I,2	I,5	I,7	2,0	I/2	I
	I,7	2,I	2,6	3,I	I/2	I/2
	0,95	I,I	I,2	I,4	I/3	I
	I,3	I,6	I,9	2,3	I/3	I/2
$\dfrac{\sigma^2(\tilde{b}_{ii}^{II})}{\sigma^2(\tilde{b}_{ii}^{III})}$	3	3	3	3	I	I
	2,2	2,I	2,I	2,I	I/2	I
	2,5	2,6	2,8	2,8	I/2	I/2
	2,0	I,8	I,8	I,7	I/3	I
	2,0	2,I	2,2	2,2	I/3	I/2
$\sigma^2(\tilde{b}_i^{III}) \Big/ \dfrac{\sigma^2(\tilde{y})}{\xi^{\square}}$	0,I667	0,09I3	0,0500	0,048I	–	I
	0,I389	0,0790	0,0450	0,0426	–	I/2
$\sigma^2(\tilde{b}_{ii}^{III}) \Big/ \dfrac{\sigma^2(\tilde{y})}{\xi^{\square}}$	0,5000	0,2296	0,I250	0,0873	I	I
	0,4444	0,2I55	0,I200	0,085I	I/2	I
	0,3055	0,I30I	0,0675	0,0469	I/2	I/2
	0,4259	0,2I09	0,II83	0,0843	I/3	I
	0,2870	0,I254	0,0658	0,0462	I/3	I/2
$\sigma^2(\tilde{b}_{ij}^{III}) \Big/ \dfrac{\sigma^2(\tilde{y})}{\xi^{\square}}$	0,2500	0,I250	0,0625	0,0625	–	–

$\sigma^2(\tilde{b}_i^{III})$ меньше любой из дисперсий $\sigma^2(\tilde{b}_i^{I})$ и $\sigma^2(\tilde{b}_i^{II})$. Это оправдывает используемую оценку $\tilde{b}_i^{III}$, как предпочтительную из возможных оценок. Промежуточными являются оценки типа:

$$\tilde{b}_i = \frac{k_1 \cdot \tilde{b}_i^{I} + k_2 \cdot \tilde{b}_i^{II}}{k_1 + k_2} \quad .$$

Оценка $\tilde{b}_{ij}$ в третьем методе единственна и осуществляется только по опытам плана n –КУБ. Для нее $\sigma^2(\tilde{b}_{ij}^{III})$ выражается аналогично (I5):

$$\Delta \tilde{b}_{ij} = \Delta \tilde{b}_{ij}^{I} = \Delta \tilde{b}_{ij}^{III} = \sum_{k=1}^{\infty} \sum_{\ell=1}^{m} \ldots \sum_{p=1}^{q} \sum_{q=1}^{n} \beta_{\underbrace{ij\ell\ell mm \ldots ppqq}_{2k+2}} \tag{32}$$

(т.е. $\tilde{b}_{ij}^{III}$ смешиваются с коэффициентами β не ниже 4-ого порядка).

<u>Замечание</u>.

В случае дробной реплики плана n –КУБ индексы символически /2, с.I34/ умножают слева на алгебраическую сумму из I и правых частей определяющих контрастов.

Осталось охарактеризовать точность коэффициентов $\tilde{b}_{ii}$. Для второго метода $\sigma^2(\tilde{b}_{ii})$ выражается аналогично (I2):

$$\Delta \tilde{b}_{ii} = \sum_{k=1}^{\infty} \beta_{\underbrace{ii \ldots ii}_{2k+2}} \quad . \tag{33}$$

Для третьего метода $\sigma^2(\tilde{b}_{ii}^{III})$ выражается аналогично (20) при $\alpha = \alpha_\perp$:

$$\Delta \tilde{b}_{ii}^{III} = \Delta \tilde{b}_{ii}^{II} + \frac{2^m (I + 2n - 2\alpha_\perp^2)}{2\alpha_\perp^4 \cdot N^{III}} \Delta \tilde{b}_0^{I} - \frac{2n(2^m + 2\alpha_\perp^2)}{2\alpha_\perp^4 \cdot N^{III}} \cdot \Delta \tilde{b}_0^{II}, \tag{34}$$

где $$\Delta \tilde{b}_{ii}^{II} = \sum_{k=1}^{\infty} \alpha^{2k} \cdot \beta_{\underbrace{ii \ldots ii}_{2k+2}} \tag{35}$$

Из сравнения формул I4 и 20 видно, что определение b_{ii} по всем точкам плана при $\xi^{\bullet} = \xi^{+} = \xi^{\square}$ уменьшает их дисперсии в 3 раза. Хотя, конечно, $\Delta\widetilde{b}_{ii}^{\overline{\text{III}}}$ может изменяться в соответствии с (23), (27) и (35). Для ряда случаев соотношение $\dfrac{\sigma^2\left(\widetilde{b}_{ii}^{\overline{\text{II}}}\right)}{\sigma^2\left(\widetilde{b}_{ii}^{\overline{\text{III}}}\right)}$ представлено в таблице 5.

§6. Вывод формул.

Ниже изложены доказательства ряда формул данной работы. Вывод формул смещений коэффициентов регрессии для планов n -КРЕСТ и n -КУБ (№№ 22,23,27,28,30-33 и 35) аналогичен и исходит из подстановки разложения 9 в формулу соответствующего коэффициента регрессии с учетом структуры данного плана. Для примера укажем вывод формулы 22 (план n -КРЕСТ). Из (2I), (9) и таблицы I следует:

$$\Delta\widetilde{b}_0 = b_0 - \beta_0 = \frac{\sum\limits_{p=1}^{2n} y_{(p)}}{2n} - \beta_0 =$$

$$= \frac{\sum\limits_{p=1}^{2n}\left(\beta_0 + \sum\limits_{k=1}^{\infty} \underbrace{\beta_{i_1 \dots i_m}}_{k>m} \cdot t_{i_1(p)}^{q_1} \dots t_{i_m(p)}^{q_m}\right)}{2n} - \beta_0 =$$

$$= \frac{2n\beta_0 + \sum\limits_{j=1}^{n}\left(\beta_j - \beta_j + \beta_{jj} - \beta_{jj} + \beta_{jjj} - \beta_{jjj} + \dots\right)}{2n} - \beta_0 =$$

$$= \frac{I}{n}\sum\limits_{j=1}^{n}\left(\beta_{jj} + \beta_{jjjj} + \dots\right).$$

Для центральных композиционных планов коэффициенты регрессии можно выразить через таковые для составляющих простых планов: n -КРЕСТ (с плечом α) и n -КУБ (или его дробная реплика). В соответствии с этими формулами аналогично выражаются смещения коэффициентов регрессии. Чтобы выразить дисперсии этих коэффициентов, предварительно необходимо привести формулы к виду:

$$\frac{a_0 \cdot \tilde{y}_{(0)} + \sum_{p=1}^{2n} a_p \cdot \tilde{y}_{(p)} + \sum_{p=2n+1}^{2n+2^m} a_p \tilde{y}_{(p)}}{}$$, откуда следует

$$\sigma^2(b) = \frac{a_0^2/\xi^{\circ} + \sum_{p=1}^{2n} a_p^2/\xi^{+} + \sum_{p=2n+1}^{2n+2^m} a_p^2/\xi^{\square}}{C^2}\,\sigma^2(\tilde{y}) =$$

$$= \left[\frac{\sum_{p=0}^{2n+2^m} a_p^2}{C^2} - \frac{\xi^{\circ}-\xi^{\square}}{\xi^{\circ}} \cdot \frac{a_0^2}{C^2} - \frac{\xi^{+}-\xi^{\square}}{\xi^{+}} \cdot \frac{\sum_{p=1}^{2n} a_p^2}{C^2}\right] \cdot \frac{\sigma^2(\tilde{y})}{\xi^{\square}}$$

<u>Формула 26.</u> Из (24) следует

$$N^{\text{III}} \cdot \tilde{b}_0^{\text{III}} = \tilde{y}_{(0)} + 2n \frac{\sum_{k=1}^{2n} \tilde{y}_{(k)}}{2n} + 2^m \frac{\sum_{k=2n+1}^{2n+2^m} \tilde{y}_{(k)}}{2^m} - (2^m + 2\alpha_{\perp}^2)\sum_{i=1}^{n} \tilde{b}_{ii}^{\text{III}} =$$

$$= \tilde{y}_{(0)} + 2n\,\tilde{b}_0^{\text{II}} + 2^m \cdot \tilde{b}_0^{\text{I}} - (2^m + 2\alpha_{\perp}^2)\sum_{i=1}^{n} \tilde{b}_{ii}^{\text{III}} ,$$

откуда вытекат (26).

<u>Формула 25.</u> Обозначим: $N = N^{\text{III}} = \text{I} + 2n + 2^m$, $\tau = \dfrac{2^m + 2\alpha_{\perp}^2}{2\alpha_{\perp}^4}$,

$h = \dfrac{2^m + 2\alpha_{\perp}^2}{N^{\text{III}}}$, $\zeta = 2\alpha_{\perp}^4$. Параметр $\alpha_{\perp}$, как известно, удовлетворяет уравнению:

$$4\alpha_{\perp}^4 = 2^m(\text{I} + 2n - 4\alpha_{\perp}^2)$$

При преобразованиях используются легко проверяемые соотношения:

$$\frac{(2^m + 2\alpha_\perp^2)^2}{N} = 2^m \;, \text{ или } \tau h = \frac{2^m}{S}; \quad \tau\cdot(2^m\cdot n + S) = Nh(1 + n\tau h);$$

$$1 + n\tau h - \alpha_\perp^2\cdot\tau = \frac{2^m\cdot(n-\alpha_\perp^2)}{S}; \quad 1 + n\tau h - n\tau = -\frac{n-\alpha_\perp^2}{\alpha_\perp^2}.$$

С учетом (24) и (18)

$$N\tilde{b}_0^{\text{III}} = \sum_{k=0}^{2n+2^m}\tilde{y}_{(k)} - \tau\cdot\sum_{k=0}^{2n+2^m}\left[\sum_{i=1}^{n}\left(t_{i(k)}^2 - h\right)\right]\cdot\tilde{y}_{(k)} =$$

$$= \sum_{k=0}^{2n+2^m}\tilde{y}_{(k)} - \tau\left[-nh\tilde{y}_{(0)} + (\alpha_\perp^2 - nh)\cdot\sum_{k=1}^{2n}\tilde{y}_{(k)} + n(1-h)\cdot\sum_{k=2n+1}^{2n+2^m}\tilde{y}_{(k)}\right] =$$

$$= (1-n\tau h)\cdot\tilde{y}_{(0)} + (1+n\tau h - \tau\cdot\alpha_\perp^2)\cdot\sum_{k=1}^{2n}\tilde{y}_{(k)} + (1+n\tau h - n\tau)\sum_{k=2n+1}^{2n+2^m}\tilde{y}_{(k)} =$$

$$= \left(1+\frac{n\cdot 2^m}{S}\right)\cdot\tilde{y}_0 + \frac{2^m(n-\alpha_\perp^2)}{S}\cdot\sum_{k=1}^{2n}\tilde{y}_{(k)} - \frac{n-\alpha_\perp^2}{\alpha_\perp^2}\cdot\sum_{k=2n+1}^{2n+2^m}\tilde{y}_{(k)} \;;$$

$$\sigma^2(\tilde{b}_0^{\text{III}}) = \left\{\left(\frac{1+\frac{n\cdot 2^m}{S}}{N}\right)^2\cdot\frac{1}{\xi^{\cdot}} + \left[\frac{2^m\cdot(n-\alpha_\perp^2)}{SN}\cdot\frac{2n}{\xi^{+}}\right] + \left(\frac{n-\alpha_\perp^2}{\alpha_\perp^2\cdot N}\right)^2\frac{2^m}{\xi^{\square}}\right\}\sigma^2(\tilde{y})$$

$$\text{или } \sigma(\tilde{b}_0^{\text{III}}) = \left\{\frac{1}{N^2}\left[(1-n\tau h)^2 + 2n(1+n\tau h - \alpha_\perp^2\cdot\tau)^2 + 2^m(1+n\tau h - n\tau)^2\right] - \right.$$

$$\left. - \frac{\xi^{\cdot}-\xi^{\square}}{\xi^{\cdot}}\left(1+\frac{n\cdot 2^m}{S}\right)^2 - \frac{\xi^{+}-\xi^{\square}}{\xi^{+}}\cdot 2n\left[\frac{2^m(n-\alpha_\perp^2)}{S}\right]^2\right\}\cdot\frac{\sigma^2(\tilde{y})}{\xi^{\square}} \;;$$

выражение в квадратных скобках упрощается:

$$(I+n\tau h)^2+2h(I+n\tau h)^2-4n\tau\alpha_1^2(I+n\tau h)+2n\tau^2\alpha_1^4+$$
$$+2^m(I+n\tau h)^2-2n\tau\,2^m(I+n\tau h)+2^m n^2\tau^2=$$
$$=N(I+n\tau h)^2-N\,2n\tau h(I+n\tau h)+N\cdot n\tau h(I+n\tau h)=$$
$$=N(I+n\tau h)=N\left(I+\frac{n\,2^m}{S}\right),$$

откуда вытекает (25).

Формула 29. Учитывая, что

$$\tilde{b}_i^{I}=\frac{\sum\limits_{k=2n+1}^{2n+2^m} t_{i(k)}\cdot\tilde{y}_{(k)}}{2^m},\qquad \tilde{b}_i^{II}=\frac{\tilde{y}_{(2i-1)}-\tilde{y}_{(2i)}}{2\alpha}$$

получаем с учетом (I6):

$$\tilde{b}_i^{III}=\frac{2^m}{2^m+2\alpha^2}\,\tilde{b}_i^{I}+\frac{2\alpha^2}{2^m+2\alpha^2}\,\tilde{b}_i^{II}$$

откуда вытекает (29).

Формулы I9,34. Учитывая, что $\tilde{b}_0^{I}=\dfrac{\sum\limits_{k=2n+1}^{2n+2^m}\tilde{y}_{(k)}}{2^m}$,

$$\tilde{b}_0^{II}=\frac{\sum\limits_{k=1}^{2n}\tilde{y}_{(k)}}{2n}\quad\text{и}\quad \tilde{b}_{ii}=\frac{\tilde{y}_{(2i-1)}+\tilde{y}_{(2i)}-2\tilde{y}_{(0)}}{2\alpha^2}$$

получаем с учетом (I8):

$$\tilde{b}_{ii}^{III}=\frac{2\alpha^2-h}{S}\cdot\tilde{y}_{(0)}+\frac{2^m(I-h)}{S}\tilde{b}_0^{I}-\frac{2nh}{S}\tilde{b}_0^{II}+\frac{2\alpha^4}{S}\cdot\tilde{b}_{ii}^{II}\qquad(36)$$

где $h=\dfrac{2\alpha^2+2^m}{N^{III}}$, $S=\sum\limits_{k=0}^{2n+2^m}\left(t^*_{i(k)}\right)^2$.

Принимая во внимание (23), (27) и (35), нетрудно убедиться в том, что

$$\frac{2\alpha^2-h}{S}+\frac{2^m(I-h)}{S}-\frac{2nh}{S}=0\quad,$$

т.е. $b_{ii}^{\overline{\text{III}}}$ независимы от $\beta_\cdot$ (при любом α):

$$\frac{2^m(I-h)}{S}-\frac{2nh}{S}\;\frac{\alpha^2}{n}+\frac{2\alpha^4}{S}=I\quad,$$

т.е. $b_{ii}^{\overline{\text{III}}}=\beta_{ii}+\ldots$ (при любом α);

$$\frac{2^m(I-h)}{S}-\frac{2nh}{S}\;\frac{\alpha^2}{n}=0\quad,$$

т.е. $b_{ii}^{\overline{\text{III}}}$ независимо от β_{jj} ($i\neq j$), но только при $\alpha=\alpha_\perp$. Из последних двух уравнений следует (I9), т.е. $S=2\alpha^4$. С учетом (I9) после подстановки h и S получаем (34).

<u>Формула 20.</u> Предварительно преобразуем (36):

$$\tilde{b}_{ii}^{\overline{\text{III}}}=\frac{2\alpha_\perp^2-h}{S}\tilde{y}_{(0)}+\frac{2^m(I-h)}{S}\,\frac{\sum\limits_{k=2n+1}^{2n+2^m}\tilde{y}_{(k)}}{2^m}-\frac{2nh}{S}\,\frac{\sum\limits_{k=1}^{2n}\tilde{y}_{(k)}}{2n}+$$

$$+\frac{\tilde{y}_{(2i-1)}+\tilde{y}_{(2i)}-2\tilde{y}_{(0)}}{2\alpha_\perp^2}=\left(\frac{2\alpha_\perp^2-h}{S}-\frac{I}{\alpha_\perp^2}\right)\tilde{y}_{(0)}-\frac{h}{S}\sum_{\substack{k=1\\(k\neq 2i-1\;\;k\neq 2i)}}^{2n}\tilde{y}_{(k)}+$$

$$+\left(\frac{I}{2\alpha_\perp^2}-\frac{h}{S}\right)\cdot\left(\tilde{y}_{(2i-1)}+\tilde{y}_{(2i)}\right)+\frac{I-h}{S}\sum_{k=2n+1}^{2n+2^m}\tilde{y}_{(k)}=-\frac{h}{S}\tilde{y}_{(0)}-$$

$$-\frac{h}{S}\sum_{\substack{k=1\\ \left(\substack{k\neq 2i-1\\ k\neq 2i}\right)}}^{2n}\tilde{y}_{(k)}+\frac{\alpha_\perp^2-h}{S}\left(\tilde{y}_{(2i-1)}+\tilde{y}_{(2i)}\right)+\frac{I-h}{S}\sum_{k=2n+1}^{2n+2^m}\tilde{y}_{(k)};$$

$$s^2\left(\tilde{b}_{ii}^{\text{III}}\right) = \left[\frac{h^2}{\xi^{\cdot}} + \frac{2(n-1)h^2 + 2(\alpha_1^2 - h)^2}{\xi^{+}} + \frac{2^m(1-h)^2}{\xi^{\square}}\right] \cdot \frac{s^2(\tilde{y})}{S^2} =$$

$$= \left[\frac{2^m}{\xi^{\cdot}} + \frac{S\,(N^{\text{III}}-2) - 2^m}{\xi^{+}} + \frac{2^m(1+2n-2\alpha_1^2)}{\xi^{\square}}\right] \cdot \frac{s^2(\tilde{y})}{N^{\text{III}}\,S^2};$$

или
$$s^2\left(\tilde{b}_{ii}^{\text{III}}\right) = \frac{1}{S^2}\left\{\left[h^2 + 2(n-1)h^2 + 2\,(\alpha_1^2 - h)^2 + 2^m(1-h^2)\right] - \right.$$

$$\left. - \frac{\xi^{\cdot} - \xi^{\square}}{\xi^{\cdot}}\,\frac{2^m}{N^{\text{III}}} - \frac{\xi^{+} - \xi^{\square}}{\xi^{+}} \cdot \frac{S\,(N^{\text{III}}-2) - 2^m}{N^{\text{III}}}\right\} \frac{s^2(\tilde{y})}{\xi^{\square}};$$

выражение в квадратных скобках упрощается:

$$h^2 + 2nh^2 - 2h^2 + S - 4\alpha_1^2 h + 2h^2 + 2^m + 2^m h^2 - 2h\cdot 2^m =$$

$$= h^2(1 + 2n + 2^m) - 2h\,(2^m + 2\alpha_1^2) + 2^m + S =$$

$$= h^2 N^{\text{III}} - 2\,h^2 N^{\text{III}} + 2^m + S = S$$

так как $h^2 N^{\text{III}} = 2^m$; откуда вытекает (20).

§7. Заключение.

Строго говоря, рассмотренные статистические методы основаны на разностных численных методах оптимизации нулевого порядка, ибо непосредственно экспериментально оценивают только значения функции. Однако для лучшего использования аналогии с соответствующими численными методами Коши, Ньютона и их модификациями более целесообразно определить их как статистические методы многомерной экспериментальной оптимизации

I-ого и 2-ого порядка, т.е. по порядку используемой модели.

Различие между разностными численными и статистическими методами, по существу, заключается в характере случайной ошибки. Для первых она определяется в идеале только ошибкой округления, для вторых – ошибкой опытов. В случае достаточной точности опытов близость этих ошибок приводит к нарушению используемых статистических закономерностей. Применение основанных на них критериев становится некорректным. Поэтому ошибка округления при использовании этих критериев должна быть значительно меньше ошибки опыта.

Принимая во внимание вышеуказанное соответствие статистических методов численным, нетрудно построить ряд их модификаций, используя известные приемы. Например, модификация метода 2-ого порядка может быть основана на движении по вектору $p_{0k} = -G_0 g_k$ и т.д. Возможно последовательное применение методов: первый – второй – третий или второй – третий, например в зависимости от достигаемой точности аппроксимации. В то же время представляет интерес применение квадратичного метода в численных, в т.ч. разностных, математических процедурах оптимизации.

Анализ общих ошибок в определении коэффициентов регрессии позволяет заметить следующее. Если для данного плана возможно несколько оценок $\tilde{b}$, то в ряду этих оценок $\sigma^2(\tilde{b})$ и $\Delta^2\tilde{b}$ ведут себя до некоторой степени противоположным образом: с уменьшением $\sigma^2(\tilde{b})$ количество коэффициентов β в смещении $\Delta\tilde{b}$ возрастает. Однако МНК выбирает оценку лишь с наименьшей $\sigma^2(\tilde{b})$. В качестве примера можно привести оценки $\tilde{b}_{ii}^{II}$ и $\tilde{b}_{ii}^{III}$ для ортогонального n -ЦКП. Согласно (36):

$\tilde{b}_{ii}^{\overline{III}} = \tilde{b}_{ii}^{\overline{II}} + \delta$, где

$$\delta = \frac{2^m(I-h)}{S} \cdot {}_{\Delta}\tilde{b}_0^{\overline{I}} - \frac{2nh}{S} \cdot {}_{\Delta}\tilde{b}_0^{\overline{II}}$$

– дополнительное смещение коэффициента $\tilde{b}_{ii}$, т.е. ${}_{\Delta}\tilde{b}_{ii}^{\overline{III}} = {}_{\Delta}\tilde{b}_{ii}^{\overline{II}} + \delta$, сопровождаемое дополнительным уменьшением его дисперсии. Выбор одной из этих оценок является проблематичным, ибо априорного преимущества друг перед другом они не имеют. Хотя можно, конечно, сопоставить впоследствии точность обеих моделей 2-ого порядка (по критерию ρ). С другой стороны, оценка $\tilde{b}_{ii}^{\overline{II}}$ не смешивается с коэффициентами вплоть до 3-его порядка при любом α , а оценка $\tilde{b}_{ii}^{\overline{III}}$ – только при $\alpha = \alpha_{\perp}$.

Следовательно, выбор α в n -ЦКП должен выполняться из других соображений (с использованием оценки $\tilde{b}_{ii}^{\overline{II}}$). В отсутствии ограничений (по величине) на интервалы варьирования факторов целесообразно положить $\alpha = \sqrt{n}$ (шаровой n -ЦКП), т.е. чтобы полнее использовать область плана. Нетрудно видеть, что для этих значений α коэффициенты $\tilde{b}_i^{\overline{III}}$ и $\tilde{b}_{ii}^{\overline{III}}$ имеют гораздо меньшие дисперсии, чем $\tilde{b}_i^{\overline{III}}$ и $\tilde{b}_{ii}^{\overline{III}}$ при $\alpha = \alpha_{\perp}$(в ортогональном n -ЦКП). Если же интервалы варьирования факторов ограничены по величине (что бывает в промышленности, например), то целесообразно положить $\alpha = I$ (кубический n -ЦКП), т.е. чтобы полнее использовать допустимую область варьирования. (В родственном численном методе для простоты также целесообразно положить $\alpha = I$).

Аналогичная альтернатива имеет место при выборе интервалов варьирования и плана эксперимента вообще. Например, увеличение интервалов варьирования факторов приводит к снижению случайной ошибки модели, но при этом возрастает систематическая ошибка; или другая альтернатива: планы n -КРЕСТ и n -КУБ в методах I-ого порядка и т.п.

В данной работе рассматривали только основные методы экспериментальной оптимизации I-ого и 2-ого порядка. Следует однако заметить, что для экспериментальной оптимизации с большой погрешностью модели (например в промышленных условиях) представляет интерес и метод Зейделя, заключающийся в покоординатном изменении влияющих параметров. Этот метод нулевого порядка наименее чувствителен к ошибкам измерений и, пожалуй, наиболее прост. Кроме того он более других методов позволяет принимать комплексное решение на каждом этапе оптимизации с учетом совокупности оптимизируемых функций. А именно, после изменения факторов достаточно ответить на вопрос: лучше или нет? Хотя метод Зейделя имеет значительно меньшую скорость сходимости, но и степень риска эквивалентно снижается. Есть случаи, когда последнее очень важно.

Вообще теория промышленной экспериментальной оптимизации требует дальнейшего развития с учетом разнообразия практических условий, а также широкой практической реализации. Она должна стать важнейшей составной частью математической экономики.

Литература.

I. Box GEP, Wilson K.B., J. Roy. Stat. Soc., Ser B, 1951, v. 13.

2. Хартман К. и др., Планирование эксперимента, М., "Мир", 1977.

3. Карманов В.Г., Математическое программирование, М., "Наука", 1980, с.183-190.

4. Тихонов А.Н., Арсенин В.Я., Методы решения некорректных задач, М., "Наука", 1979, с.110-127.

5. Box GEP, Appl. Stat. 1957, v. 6.