

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

С. П. МАНУКЯН, Э. П. МАНУКЯН, В. М. ПИСАРЕНКО

МЕТОД ПОСТРОЕНИЯ ОПТИМАЛЬНЫХ
 ПОСЛЕДОВАТЕЛЬНЫХ ПЛАНОВ РЕГРЕССИОННЫХ
 ЭКСПЕРИМЕНТОВ СО СВОЙСТВАМИ ϵ -ОПТИМАЛЬНОСТИ

Проблема извлечения наибольшего количества сведений об изучаемых процессах при ограниченных затратах на эксперименты является весьма актуальной. В этой связи оказывается необходимым применение методов, которые давали бы не только способ обработки экспериментальных данных, измеряемых при воздействии некоторых случайных помех, но и позволили бы оптимальным образом организовать проведение регрессионных экспериментов. Изложим следующую математическую постановку проблемы планирования оптимальных экспериментов, являющуюся достаточно общей.

Пусть на основе априорной информации о механизме исследуемого процесса экспериментатором выбран аналитический вид функции, описывающей связь между координатами X_1, X_2, \dots, X_n факторного пространства и исследуемой величиной Y ,

$$Y = F(X, Y, Q). \quad (1)$$

Параметры Q_1, Q_2, \dots, Q_n , входящие в (1), определяются на основе обработки экспериментальных данных и выбираются, исходя из принципа максимума правдоподобия.

Проблемой планирования экспериментов является определение совокупности тех точек факторного пространства, в которых при реализации экспериментов получаем статистический материал, позволяющий наилучшим образом оценить искомые значения параметров Q_1, Q_2, \dots, Q_n при минимальных затратах на эксперименты. В настоящее время при статистической обработке результатов экспериментов наибольшее распространение получил метод наименьших квадратов. Найденные по этому методу совместные эффективные оценки неизвестных параметров Q_1, Q_2, \dots, Q_n имеют дисперсионную матрицу $\|1\|$:

$$D(\bar{Q}) = M^{-1}(\bar{Q});$$

$$M(Q) = \sum_{i=1}^k \varepsilon_i^{-2} f(X^i) \cdot f^T(X^i);$$

$$f(X) = \nabla \cdot Y(X, Q); \quad (2)$$

$$\sigma^2 = \left\| \frac{\sigma}{\sigma Q_1} : \frac{\sigma}{\sigma Q_2} : \dots : \frac{\sigma}{\sigma Q_n} \right\|.$$

где σ_i^2 — дисперсия измерений исследуемой величины при i -ом эксперименте;

$M(\bar{Q})$ — информационная матрица Фишера.

Дисперсионная матрица $D(\bar{Q})$ геометрически интерпретируется как некоторый эллипсоид рассеяния точек \bar{Q} в пространстве неизвестных параметров, объем, площадь поверхности, сумма полуосей и подобные геометрические параметры которого характеризуют близость полученных оценок \bar{Q} к их истинным значениям [2]. Существует целый ряд методов по составлению оптимальных планов, минимизирующих параметры эллипсоида рассеяния на множестве планов в заданной области. Большинство разработанных к настоящему времени планов хорошо зарекомендовали себя в простых задачах, при небольшом числе неизвестных параметров Q_i и линейных моделях. Встречаемые же на практике сложные прикладные задачи отличаются сильно выраженной нелинейностью и наличием большого числа неизвестных параметров в предполагаемой математической модели исследуемого процесса. Как показывает практика, применение имеющихся методов по составлению планов экспериментов, минимизирующих обобщенные параметры эллипсоида рассеяния, в этих случаях не оправдывают себя. Эллипсоиды рассеяния при них получаются сильно вытянутыми, что увеличивает коридор ошибок искоемых параметров и затрудняет процедуру поиска этих параметров по методу наименьших квадратов. Если учесть, что полуоси эллипсоида рассеяния характеризуются собственными числами дисперсионной матрицы $[D(\bar{Q})]$, то это явление можно устранить, если регрессионные эксперименты проводятся по так называемым E -оптимальным планам, при обработке результатов которых получаются минимальные значения максимальных собственных чисел, соответствующих дисперсионной матрице. Но при построении подобных планов возникает необходимость многократного выполнения трудоемкой операции определения собственных чисел матрицы, что и является существенным препятствием при построении и применении E -оптимальных планов. Для разрешения вышеназванных проблем в настоящей статье предлагается метод построения планов последовательных экспериментов, позволяющий после обработки их результатов уменьшить вытянутость эллипсоида рассеяния, при этом опускается трудоемкая операция вычисления собственных чисел матрицы.

При нашем стремлении к уменьшению вытянутости эллипсоида рассеяния и коридора ошибок искоемых параметров, естественно было бы поставить перед каждым новым $(N+1)$ -ым экспериментом задачу минимизации выражения

$$|1 + \alpha\Phi(N+1)|L[D(N+1)] = \min_{X \in X} |1 + \alpha\Phi(N+1)|L[D(N+1)], \quad (3)$$

где α — постоянный коэффициент;

X — точка в факторном пространстве, где должен проводиться очередной $(N+1)$ -ый эксперимент;

L — функционал, ставящий в соответствие каждой матрице $D(N+1)$ некоторую скалярную величину, характеризующую точность определения оценок искомых параметров.

В частности он может быть $\det D(N+1)$, использование которого приводит к минимизации объема эллипсоида рассеяния, $\text{Sp } D(N+1)$, минимизирующий сумму полуосей эллипсоида, и т. д.

$\Phi(N+1) = \Phi(N, X)$ — положительная функция, находящаяся в пропорциональной зависимости от максимального разброса собственных чисел дисперсионной матрицы $D(N+1)$.

Присутствие в (3) функции $\Phi(N+1)$ позволит наряду с уменьшением обобщенных параметров эллипсоида рассеяния, как это делается в широко распространенных методах планирования, уменьшить также и его вытянутость. В качестве функции $\Phi(N+1)$, удовлетворяющей вышеуказанным требованиям, предлагается применять функцию вида:

$$\Phi(N, X) = \sum_{\substack{i=1 \\ j=i+1 \\ i+j}}^n \left(\frac{\nu_i - \nu_j}{\nu_i} \right)^2 = \sum_{\substack{i=1 \\ j=i+1 \\ i+j}}^n \left(\frac{\lambda_j - \lambda_i}{\lambda_j} \right)^2, \quad (4)$$

где ν_i — i -ое собственное число дисперсионной матрицы оценок искомых параметров, полученных в результате обработки данных $N+1$ экспериментов;

$\lambda_i = \frac{1}{\nu_i}$ — i -ое собственное число, соответствующее информационной матрице Фишера.

При непосредственном вычислении функции $\Phi(N, X)$ по формуле возникает необходимость выполнения трудоемких операций по определению собственных чисел информационной матрицы. Покажем, что выражение (4) можно вычислять и без определения самых собственных чисел λ_i .

Заменим функцию (4) некоторым ее приближением, которое также достаточно хорошо характеризует разброс собственных чисел дисперсионной матрицы, т. е.

$$\Phi(N, X) = \sum_{\substack{i=1 \\ j=i+1 \\ i+j}}^n \left(\frac{\lambda_i - \lambda_j}{\lambda_{cp}} \right)^2, \quad (5)$$

где

$$\lambda_{cp} = \frac{1}{n} \sum_{i=1}^n \lambda_i.$$

Преобразуем последнее выражение:

$$\Phi(N, X) = \frac{1}{2} \sum_{\substack{i=1 \\ j=1 \\ (i,j)}}^n (x_i - x_j)^2 = \frac{1}{2} \left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right].$$

Используя известные соотношения матричной алгебры [3], отражающие связи между собственными числами матрицы с ее элементами, получим:

$$\begin{aligned} \Phi(N, X) &= \frac{n^2}{\left(\sum_{i=1}^n m_{ii} \right)^2} \left[n \sum_{j=1}^n m_{jj}^2 - \left(\sum_{i=1}^n m_{ii} \right)^2 \right] = \\ &= \frac{n^2 \left\{ n \sum_{j=1}^n m_{jj}^2 - |\text{Sp} M(N+1)|^2 \right\}}{|\text{Sp} M(N+1)|^2} \end{aligned} \quad (6)$$

где m_{ij} — элементы информационной матрицы $M(N+1)$.

Использование формулы (6) для вычисления функции $\Phi(N+1)$ уже не представляет особой трудности и не требует выполнения таких трудоемких операций, как определение собственных чисел, обращение матриц и т. д.

Уменьшению вытянутости эллипсоида рассеяния способствует также и обоснованный выбор скаляра $l|D(N+1)|$. Практика применения так называемых A -оптимальных планов, при которых минимизируется $\text{Sp} D(N+1)$, показывает, что при достаточной эффективности каждого эксперимента, проведенного по такому плану, эллипсоиды получаются менее вытянутыми. Основным препятствием применению таких планов на практике можно считать необходимость выполнения обращения матриц на каждом шагу поисковой процедуры оптимизации функционала (3); при поиске необходимо вычислять значение $\text{Sp} D(N+1) = \text{Sp} [M(N) + \sigma^{-2}(X) f(X) f'(X)]^{-1}$ при различных значениях X . Заменяя трудоемкую операцию обращения информационной матрицы более простыми, можно облегчить вычислительный процесс при поиске минимума функционала (3). Покажем, что такая замена при вычислении $\text{Sp} M^{-1}(N+1)$ возможна, если известно значение $\text{Sp} M^{-1}(N)$ и измеряемая величина Y , характеризующая исследуемый процесс, является одномерной. Для этого преобразуем выражение для $\text{Sp} D(N+1)$ к виду:

$$\begin{aligned} \text{Sp} D(N+1) &= \text{Sp} [M(N) + \sigma^{-2}(X) f(X) f'(X)]^{-1} = \\ &= \text{Sp} [I + M^{-1}(N) \sigma^{-2}(X) f(X) f'(X)]^{-1} M^{-1}(N). \end{aligned}$$

Примем $M^{-1}(N) = D(N)$; $D(N) \sigma^{-2}(X) f(X) = A$; $f'(X) = B$ и воспользуемся следующей леммой [1]:

$$[I_p + A \cdot B]^{-1} = I_p - A [I_q + BA]^{-1} \cdot B.$$

где I_p и I_q — единичные матрицы размера $p \times p$ и $q \times q$;
 A — матрица размера $p \times q$;
 B — матрица размера $q \times p$.

Тогда

$$\text{Sp} D(N+1) = \text{Sp} [I_m - D(N) \sigma^{-2}(X) f(X) \{ [1 + f^T(X) D(N) \sigma^{-2}(X) f(X)]^{-1} + f^T(X) \} D(N)].$$

Обозначив $[1 + f^T(X) D(N) \sigma^{-2}(X) f(X)]^{-1} = R$, получим:

$$\begin{aligned} \text{Sp} D(N+1) &= \text{Sp} [I_m - R D(N) \sigma^{-2}(X) f(X) f^T(X)] \cdot D(N) = \\ &= \text{Sp} D(N) - \text{Sp} R D(N) \sigma^{-2}(X) f(X) f^T(X) D(N) = \\ &= \text{Sp} D(N) - \text{Sp} R \sigma^{-2}(X) f^T(X) D(N) D(N) f(X). \end{aligned}$$

Так как в последнем выражении второе слагаемое представляет собой скалярную величину, то окончательно можно записать:

$$\text{Sp} D(N+1) = \text{Sp} D(N) - [1 + f^T(X) D(N) \sigma^{-2}(X) f(X)]^{-1} f^T(X) \sigma^{-2}(X) \times \\ \times D^2(N) f(X). \quad (7)$$

Очевидно, что при вычислении $\text{Sp} D(N+1)$ по формуле (7) не требуется выполнения операции обращения матриц. Вместо нее выполняются операции умножения матриц, что не представляет собой особых вычислительных трудностей.

Обобщая полученные результаты, предложим следующий итерационный алгоритм метода последовательного планирования оптимальных экспериментов:

1. Проводится некоторый стартовый эксперимент из N_0 измерений;
2. Проводится обработка результатов измерений стартового эксперимента методом наименьших квадратов, в результате которой получаются однозначные оценки Q_i и приближенное значение их дисперсионной матрицы;
3. Определяются координаты точки факторного пространства, в которой необходимо проводить следующее $(N+1)$ -ое измерение как координаты вектора X , минимизирующего метрику (3)

$$\min_{X \in X} \{ [1 + \sigma^2 \Phi(N, X)] L | D(N, X) | \}.$$

На каждом шагу поиска минимума последнего функционала используются ранее полученные выражения (6) и (7);

4. Проводится $(N+1)$ -ый эксперимент при найденных значениях факторов X_1, X_2, \dots, X_n , после результатов которого уточняются ранее найденные значения оценок Q_i и соответствующей дисперсионной матрицы;

5. Проверяется условие

$$|Q_i(N+1) - Q_i(N)| < \varepsilon_i, \quad (i=1, 2, \dots, n) \quad (8)$$

где ε_i — заранее заданное малое число.

Если условие (8) не удовлетворяется, то выполняется переход к операциям, описанным в пункте 3, в противном случае итерационный процесс планирования и экспериментирования прекращается.

В качестве примера рассмотрим химическую реакцию типа



в которой реагент R — первичный спирт с длиной цепочки; P_1 — олефин; P_2 — вода. Теоретический анализ данной реакции показал, что она может быть описана следующей моделью:

$$Y(X, Q) = \frac{Q_1 Q_2 X_1}{1 + Q_1 X_1 + Q_2 X_2} \quad (8)$$

где Y — скорость реакции; X_1 и X_2 — парциальные давления продуктов R и R_1 соответственно; Q_i — константы реакции.

Требуется запланировать и реализовать эксперименты, результаты которых позволят как можно точнее определить Q_1 , Q_2 , Q_3 . Будем считать, что эксперименты можно реализовать в области

$$0 \leq X_1 \leq 3; \quad 0 \leq X_2 \leq 3.$$

Эксперименты и результаты измерения будем имитировать, используя таблицу случайных чисел. Предполагается, что Y нормально распределено около среднего $Y(X, Q_{ист})$ с дисперсией $\sigma^2 = 0.01$. Истинные значения параметров Q_i известны: $Q_{1,ист} = 2.9$; $Q_{2,ист} = 12.2$; $Q_{3,ист} = 0.69$. (Предполагается, конечно, что значения $Q_{ист}$ экспериментатору неизвестны.)

Приведем стартовый эксперимент в четырех точках факторного пространства $\{1; 1\}$; $\{2; 1\}$; $\{1; 2\}$; $\{2; 2\}$, в которых измерение Y дало результаты: $Y_1 = 0.126$; $Y_2 = 0.219$; $Y_3 = 0.076$; $Y_4 = 0.126$. Предварительные оценки Q_i , полученные на основе этих данных, равны $\bar{Q}_1 = 10.39$; $\bar{Q}_2 = 48.83$; $\bar{Q}_3 = 0.74$. Принимая, что оценки \bar{Q} (4) не слишком отличаются от истинных, вычисляем функцию $f(X) = \nabla \cdot Y(X, \bar{Q}(4))$ в четырех точках стартового эксперимента. Составляем информационную матрицу для пяти экспериментов:

$$M(5) = M(\bar{Q}, X) = \sum_{i=1}^4 f(\bar{Q}, X_i) \cdot f'(\bar{Q}, X_i) + f(\bar{Q}, X) f'(\bar{Q}, X).$$

где X — координаты точки пятого эксперимента в факторном пространстве.

Используя элементы полученной матрицы, составляем функционал $\Phi(4, X)$ по формуле (6). Подставим его в (3) и с помощью поисковой программы на ЦВМ находим точку в факторном пространстве, в которой функционал (3) получается минимальным.

Вычисления на ЦВМ дали следующие значения: $X_1 = 0.152$; $X_2 = 1.01$. При расчетах на ЦВМ в качестве $L(D(N+1))$ в формуле (3) использовался $\det D(N+1)$ [1]. Эксперимент, проведенный в этой точке, дал результат $Y = 0.352$. Используя его, получим:

$$\bar{Q}(5) = \| 3,05; 13,5; 0,8 \|.$$

Вышеописанным же образом планируем и реализуем следующие эксперименты. Результаты планирования и опытов представлены в табл. 1.

Таблица 1

Номер опыта	x_1	x_2	\bar{Q}_1	\bar{Q}_2	\bar{Q}_3
1	1,00	1,00	—	—	—
2	2,00	1,00	—	—	—
3	1,00	2,00	—	—	—
4	2,00	2,00	10,39	48,83	0,74
5	0,15	0,01	3,05	13,50	0,80
6	3,01	0,01	3,11	13,30	0,65
7	0,19	0,00	3,04	12,50	0,60
8	3,01	0,01	3,00	12,15	0,68
9	0,21	0,00	3,00	12,17	0,68
10	3,00	1,05	2,95	12,16	0,67

Сравнивая полученные результаты с результатами планирования и опытов, проведенных в работе [1] для выяснения механизма рассматриваемой здесь реакции, приходим к выводу, что предложенная в настоящей статье методика позволила при меньшем количестве экспериментов получить оценки некоторых параметров \bar{Q}_i с меньшим значением коридора ошибок. Так, например, в работе [1] составлялись D -оптимальные последовательные планы экспериментов, которые после 13 экспериментов позволяли получать оценки:

$$Q(13) = \| 3,57; 12,77; 0,63 \|.$$

Таким образом, предложенная методика позволяет получать планы экспериментов, близкие к E -оптимальным, при этом не требуется многократного выполнения таких трудоемких вычислительных операций, как вычисление детерминантов, собственных чисел, обращение матриц и т. д.

ГрНХ им. К. Маркса

Поступила 19.VI.1975

Ա. Ն. ՄԱՎՈՒԿՅԱՆ, Է. Ե. ՄԱՆՈՒԿՅԱՆ, Վ. Ն. ՊԻՍԱՐԵՆԿՈ

ՌԵԿԵՐԵՍԻՆՆ ԲՈՐՁԵՐԻ Է-ՕՊՏԻՄԱԼԱՅԻՆ ՀԱՏԿՈՒԹՅՈՒՆ ՈՒՆԵՑՈՂ
ՀՕՋՈՐԿՈՒՄԱՆ ՊԼԱՆԱՎՈՐՄԱՆ ՄԵԹՈԴ

Ա Վ Փ Ո Փ Ո Վ

Աշխատանքում առաջարկվում է փորձերի օպտիմալ պլանավորման մեթոդ, որի նպատակն է պարզել հետազոտվող երևույթի որակից մաթեմատիկական նկարագրի մեջ մտնող անհայտ հաստատունների արդյունքը Առաջարկ-

վող մեթոդը թույլ է տալիս փորձերի մինիմալ քանակի դեպքում ստանալ որոնվող հաստատունների արժեքների զննհատականը մեծ ճշտությամբ: Փորձերի պլանավորման առաջարկված մեթոդը իրականացնելու համար չի պահանջվում բարդ, աշխատատար հաշվողական գործողություններ:

Л И Т Е Р А Т У Р А

1. Федоров В. В. Теория оптимального эксперимента. Изд-во «Наука», 1971.
2. Sibson R. (1972) Discussion of paper by H. P. Wynn. I. R. Statist. Soc. B—34. 181—3.
3. Пароди М. Локализация характеристических чисел матриц и ее применение. Изд-во «Иностр. литература», 1960.