

Определение значимости факторов и их взаимодействия в многофакторном эксперименте

Р. Алалами, С.С. Торбунов

После изучения объекта исследования и его физической сущности возникает ряд представлений о действии различных параметров (факторов) и необходимость получить экспериментальные данные о их совокупном влиянии на какой-либо показатель (критерий), характеризующий объект исследований. На языке математики этот критерий называется откликом. Функция, связывающая независимые факторы с критерием, называется функцией отклика, а ее геометрический образ – поверхностью отклика.

При составлении плана эксперимента прежде всего назначают (выбирают) независимые факторы, исходя из априорной (доопытной) информации или предварительного изучения объекта исследования.

Факторы бывают количественные и качественные. Количественные можно измерить и выразить в числах.

Далее при составлении плана эксперимента назначаются уровни варьирования факторами, или градации. При качественных факторах их можно условно пронумеровать (закодировать). Допустим, подача материала в машину самотексом – уровень 0, а принудительно – 1.

В многофакторных экспериментах, основанных на дисперсионном анализе, обычно берут два уровня факторов (верхний и нижний). Ведь в таких экспериментах важно проверить, значимо ли влияет тот или иной фактор и есть ли факторные взаимодействия. Варьирование переменными на двух уровнях позволяет значительно уменьшить объем экспериментальной и счетной работы.

В многофакторном эксперименте уровни одного фактора должны сочетаться с уровнями другого, образуя тем самым вариант испытаний. Интервал варьирования того или иного фактора должен быть таким, чтобы можно было реализовать любой вариант испытаний.

В теории планирования эксперимента условно принято обозначать нижнюю границу или нижний уровень фактора знаком -1 , а верхнюю $+1$.

В теории планирования эксперимента, если факторы устанавливаются на двух уровнях, принято обозначать – эксперимент типа 2^n , где n – число факторов. Если факторы устанавливаются на трех уровнях, то называют эксперимент типа 3^n и т.д.[14].

Комбинации условий эксперимента 2^2 можно выразить в виде таблицы, если обозначить нижний уровень фактора -1 , а верхний $+1$. Такая таблица называется матрицей планирования эксперимента (табл. 1).

Таблица 1 - Матрица планирования двухфакторного эксперимента

№ опыта	Факторы и их взаимодействия			Кодовое обозначение
	A	B	AB	
1	2	3	4	5
1	-	-	+	(1)
2	+	-	-	a
3	-	+	-	b
4	+	+	+	ab

Примечание. С целью упрощения записи символы 1 не указаны, а проставлены только их знаки.

В первом столбце рассматриваемой матрицы эксперимента записаны номера опытов (без повторностей), которые при реализации необходимо рандомизировать. Это варианты испытаний. Второй и третий столбцы представляют собственно планирование, образуя возможные комбинации знаков факторов (условий испытаний). Четвертый столбец образуется

перемножением знаков факторов А×В. Он показывает возможные взаимодействия факторов и их знаки в том или ином опыте.

Взаимодействие факторов в многофакторном эксперименте следует понимать так: изменение одного фактора сопровождается непропорциональными изменениями результатов эксперимента при изменении уровней другого.

Рассмотрим планирование экспериментов с тремя независимыми переменными (факторами) А,В,С, которые также варьируют на двух уровнях. Это будет эксперимент типа 2³.

Таблица 2 - Матрица планирования трехфакторного эксперимента

№ опыта	Факторы и их взаимодействия							Кодовое обозначение строк
	А	В	С	АВ	АС	ВС	АВС	
1	-	-	-	+	+	+	-	(1)
2	+	-	-	-	-	+	+	<i>a</i>
3	-	+	-	-	+	-	+	<i>b</i>
4	+	+	-	+	-	-	-	<i>ab</i>
5	-	-	+	+	-	-	+	<i>c</i>
6	+	-	+	-	+	-	-	<i>ac</i>
7	-	+	+	-	-	+	-	<i>bc</i>
8	+	+	+	+	+	+	+	<i>abc</i>

Однако при постановке полного многофакторного эксперимента даже только на двух уровнях уже требуется проводить большое число опытов, поэтому для сокращения затрат средств и времени можно не реализовать полный факторный эксперимент, а ограничиться лишь некоторой частью его. В этих случаях используют и часть матрицы, называемую дробной репликой.

При реализации полного факторного эксперимента можно определить взаимодействия всех порядков, но в этом нет необходимости. При реализации же дробной реплики от полного факторного эксперимента теряется часть информации о влиянии факторов, а именно, в дробных репликах эффекты ряда взаимодействий (в зависимости от структуры дробной реплики) приравниваются либо какому-нибудь эффекту фактора, либо ошибке эксперимента. Или иначе – эффект фактора смешивается с взаимодействием высокого порядка. Если известно заранее, что взаимодействие, с которым смешан фактор, не значимо, то вычисленный эффект будет равен эффекту фактора.

Дробная реплика может быть построена с различной степенью дробности. Допустим, что требуется получить полуреплику от трехфакторного эксперимента: *a, b, ab, c, ac, bc, abc*. Из восьми условий эксперимента нужно получить четыре. Факторы обозначим через x_1, x_2, x_3 . Если известно, что взаимодействие x_1, x_2 не значимо, то можно с ним смешать фактор x_3 и можно записать $x_3 = x_1 x_2$. Выражение $x_3 = x_1 x_2$ называется генерирующим соотношением [7]. При помощи генерирующих соотношений строятся дробные реплики.

В трехфакторном эксперименте генерирующие соотношения можно принять в виде: $x_3 = -x_1 x_2, x_2 = x_1 x_3, x_1 = -x_1 x_3$ и т.д.

Задавись определяющим контрастом, можно найти генерирующие соотношения, которые покажут, какие эффекты в эксперименте будут смешаны.

Для трехфакторного эксперимента определяющие контрасты можно задать так:

$$I = x_1 x_2 x_3, I = -x_1 x_2 x_3.$$

Определяющий контраст I всегда при кодированных факторах (+1 или -1) будет равен единице. Для того, чтобы найти генерирующие соотношения, или, как еще говорят, определить совместимые оценки эффектов факторов, необходимо последовательно помножить независимые переменные на определяющий контраст, учитывая $x_1^2 = 1 = I$.

Поясним подробнее, почему так получается. Задавись определяющим контрастом

$I = x_1 x_2 x_3$, можно получить одну полуреплику от трехфакторного эксперимента. Допустим, для анализа выбрали полуреплику с нечетным сочетанием символов a, b, c, abc . Составим для этой полуреплики развернутую матрицу планирования эксперимента (табл.3).

Из табл. 3 видно, что эффект фактора x_1 равен

$$X_1 = +a - b - c + abc.$$

Эффект взаимодействия $x_2 x_3$ равен

$$X_2 X_3 = +a - b - c + abc.$$

Значит, при реализации этой полуреплики нельзя различить эффекты факторов x_1 и $x_2 x_3$, т.е. они будут совместными. Из табл.3 видны и другие совместные эффекты. В этом плане нельзя различить отдельно эффект фактора и парных взаимодействий при наличии эффектов взаимодействий.

Таблица 3 - Полуреплика факторного эксперимента 2^3

Кодовое обозначение строк (комбинаций условий)	Эффекты факторов						
	X_1	X_2	X_1X_2	X_3	X_1X_3	X_2X_3	$X_1X_2X_3$
a	+	-	-	-	-	+	+
b	-	+	-	-	+	-	+
c	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+

Поэтому, если экспериментатор не знает заранее, что взаимодействий нет, то реализация такого плана не имеет практической ценности.

Как показывает опыт экспериментальной работы, полуреплики от полного факторного эксперимента целесообразно получать лишь при числе факторов более трех.

Вышеизложенное о разрешающей способности дробных реплик и правилах их построения можно свести в общую схему. Для этого необходимо проделать следующие операции.

1. Задавшись общим количеством факторов n , необходимо выписать общее количество условий испытаний или опытов N по формуле $N=2^n$, если факторы варьируют на двух уровнях.

2. Исходя из конкретной обстановки, условий, бюджета времени и т.п., решается вопрос об объеме реализации опытов. Если эксперимент решено провести не в полном объеме, то назначается дробная реплика (1/2, 1/8 и т.д.).

3. Задавшись дробной репликой, вычисляют величину p в формуле $Q=2^{n-p}$, где Q – число опытов в дробной реплике.

Это необходимо для того, чтобы знать, сколько эффектов факторов смешать с эффектами взаимодействий.

4. Решается вопрос о том, какие факторы и с какими взаимодействиями можно смешать.

Если исследование проводится впервые, то делают случайный выбор и на этом основании строят генерирующие отношения. Для удобства построения дробной реплики эффект факторов лучше смешивать с эффектами взаимодействий, которые содержат нижние индексы их начальных цифр, например $x_2 = x_1 x_2 x_3$ или $x_3 = x_1 x_2 x_3 x_4 x_5$ и т.п.

5. Строится матрица планирования эксперимента дробной реплики на основании выбранных генерирующих соотношений.

6. При выбранных генерирующих соотношениях записываются обобщающие контрасты и затем определяется обобщающий определяющий контраст.

7. Обобщающий определяющий контраст умножается на тот или иной фактор (взаимодействие) для определения совместных эффектов факторов (взаимодействий). Как правило, взаимодействиями высоких порядков, начиная с тройных, пренебрегают.

Математическая модель двухфакторного эксперимента представляется в следующем виде

$$X_{ij} = \mu + A_i + B_j + AB_{ij} + \varepsilon_{ij},$$

где: μ – общий эффект всего эксперимента

A_i - эффект фактора A на i-ом уровне;

B_j - эффект фактора B на j-м уровне;

AB_{ij} - эффект парного взаимодействия факторов A и B на уровнях i и j;

ε_{ij} - ошибка эксперимента.

Математическая модель трехфакторного эксперимента представлена выражением

$$X_{ijk} = \mu + A_i + B_j + C_k + AB_{ij} + AC_{ik} + BC_{jk} + ABC_{ijk} + \varepsilon_{ijk}.$$

Здесь по сравнению с двухфакторным экспериментом добавляется эффект C_k третьего фактора, а также парные взаимодействия AC_{ik} , BC_{jk} и взаимодействие третьего порядка ABC_{ijk} . Подобные модели эксперимента (7 и 10) можно составить для любого числа факторов, однако при большом количестве их значительно увеличивается число опытов и общий объем последующих вычислений.

Если при исследовании динамического процесса задана функция цели (критерий оптимизации), то значимость факторов и алгоритм оптимизации согласно методу сечений соответствующей гиперповерхности (по числу факторов регрессионной модели) можно определить в виде следующей последовательности операций [4].

1. Задаем начальные значения параметров $x_i(0)$, $i=1, \dots, n$; величины исследующих шагов оптимизации δx_i , $i=1, \dots, n$; величины шагов оптимизации Δx_i , $i=1, \dots, n$; ограничения параметров

$$a_i \leq x_i \leq b_i, \quad i=1, \dots, n.$$

2. Вычисляем приращение параметров оптимизации

$$x'_i(N) = x_i(N) + \delta x_i, \quad i=1, \dots, n.$$

3. Вычисляем значение функции

$$Q(N) = Q [x_1(N), \dots, x_n(N)].$$

4. Определяем значения функции при варьировании параметров

$$Q_i(N) = Q [x'_i(N), x_j (j \neq i) (N)], \quad i=1, \dots, n.$$

5. Вычисляем приращения функции

$$\Delta Q_i(N) = Q(N) - Q_i(N), \quad i=1, \dots, n.$$

6. Вычисляем нормированные приращения функции

$$v_i(N) = \frac{\Delta Q_i(N)}{\Delta Q_i(1)} \leq \mu_i, \quad i=1, \dots, n.$$

7. Вычисляем знаки единичных векторов

$$e_i = \frac{Q(N) - Q_i(N)}{|Q(N) - Q_i(N)|}, \quad i=1, \dots, n.$$

8. Вычисляем значения параметров для перевода системы в новое состояние, приближающее ее к экстремуму,

$$x_i(N) = x_i(N-1) + \Delta x_i v_i(N) e_i, \quad i=1, \dots, n.$$

9. Проверяем ограничения:

если $x_i(N) < a_i$, то $x_i(N) = a_i$, $i=1, \dots, n$;

если $x_i(N) > b_i$, то $x_i(N) = b_i$, $i=1, \dots, n$.

10. Вычисляем приращение функции

$$\Delta Q(N) = Q(N) - Q(N+1)$$

(при $N=1$ не вычисляется).

11. Вычисляем относительные величины

$$\delta Q_i = \frac{Q(N) - Q_i(N)}{Q(N)} \geq \beta_i;$$

$$\delta x_i = \frac{x_i(N) - x_i(N+1)}{x_i(N)} \geq \varphi_i, i=1, \dots, n.$$

12. После выполнения операций 1-11 заканчивается первый шаг итерации N=1; далее, задавая N=2, выполняем операции 3-11.

13. Шаги итерации N=1, 2, 3, ... повторяем до тех пор, пока соблюдаются условия 6, 12. В момент нарушения условия 6 переходим к следующей операции.

14. Фиксируем значение $v_i(N)$ на постоянном уровне и далее приближаемся к экстремуму, повторяя шаги итерации с минимальным постоянным шагом. Момент перехода экстремума функции определяется по изменению знака величины $v_i(N) = v_i(N_0)$.

15. Определяем радиусы кривизны

$$p_i = \frac{[x_i(N_0-1) + x_i(N_0-2)] - [x_i(N_0-2) + x_i(N_0-3)]}{2[v_i(N_0-2) - v_i(N_0-1)]}, \quad i=1, \dots, n.$$

16. Определяем оптимальные значения параметров

$$x_i^* = \frac{x_i(N_0-1) + x_i(N_0-2)}{2} + p_i v_i(N_0-1).$$

Пример: Дана целевая функция

$$Q = [4,905 x_7 x_6^2 + x_4 x_6 + x_5 x_6 E] D$$

при ограничениях

$$E = \frac{tgx_1 + ctgx_1}{(ctgx_1 + ctgx_2)(tgx_1 + tgx_3)}; \quad D = \frac{ctgx_1 + ctgx_2}{ctg(x_1 + x_3) + ctg(x_2 + x_8)}.$$

В результате реализации алгоритма в виде программы и расчета получены значения составляющих оптимального вектора X:

Начальный вектор оптимизации может быть задан любым в пределах ограничений (табл. 4), при этом результат решения получается всегда одним и тем же (табл. 5).

Таблица 4 – Ограничения на параметры целевой функции

Параметры	x_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Размерность		град	град	град	кПа	кПа	м	т/м ³	град
Предел	a_i	30	30	22	0	40	0,12	1,6	16,5
Предел	b_i	80	40	40	100	60	0,25	2,2	30

Таблица 5 – Значения составляющих оптимального вектора

x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	x_6^*	x_7^*	x_8^*
55,75	30	22	0	40	0,2	1,6	16,5

Одной из трудных задач, которую приходится решать при оптимизации многопараметрических систем, является выбор минимального значения шага итерации $\Delta x_{i\min}$, находящегося близко к границе разрешающей способности ЭВМ. Это минимальное значение шага является ошибкой значений оптимизирующих параметров, под которой понимают наименьший сдвиг экспериментов, позволяющий обнаружить разницу значений целевой функции [4].

Ошибку оптимизирующего параметра оценивают относительной величиной [5]

$$E_i = \frac{x_i(N) - x_i(N_0-1)}{x_i(N)} 100, i=1, \dots, n.$$

Если к экстремуму функции подойти с минимальным шагом, то достаточная точность оптимизирующего параметра будет достигнута сразу в момент изменения знака функции (4.12). В противном случае для повышения точности оптимизации можно использовать методы дробления шагов оптимизации.

После получения оптимального решения в виде $Q(X^*)$ и совокупности X^* (табл. 6) решение задачи оптимизации считается выполненным. Однако, как уже отмечалось, такой информации оказывается недостаточно при оптимизации конструктивных параметров машины на этапе проектирования. Необходимо дополнительно к полученному решению осуществить оценку влияния каждого параметра на целевую функцию при помощи метода сечений n -мерных гиперповерхностей или по относительному изменению целевой функции при изменении параметра от оптимального значения x_i до границы a_i или b_i (табл. 6).

Таблица 6 – Оценка значимости влияния параметров на целевую функцию

Параметры x_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Относительное влияние, %	,36,3	29,3	7,4	100	46,2	119	2,6	82,5

Таким образом, разработанный метод оптимизации позволяет получить оптимальное решение и оценить влияние параметров на целевую функцию.

Литература

1. Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных условий. – М.: Наука, 1971.
2. Налимов В.В. Теория эксперимента. – М.: Физматгиз, 1971.
3. Математическое моделирование и параметрическая устойчивость динамических систем с детерминированными параметрами.– Томск: Изд-во ТПУ, 2002.