

11.- Diseño del experimento y optimización

11.1- Diseño de experimentos

11.1.1- Terminología

Anteriormente hemos tratado el problema de escoger el mejor, atendiendo a ciertos criterios, de entre un conjunto dado de sistemas. En ese tipo de problemas las alternativas vienen dadas. En esta sección, por el contrario, trataremos la situación en la cual el objetivo del estudio de simulación esta menos estructurado; se desea o bien estudiar cómo afectan al comportamiento del sistema algunos de sus parámetros o algunas de las hipótesis realizadas sobre su estructura, o bien estimar qué conjunto de especificaciones conducen a un comportamiento óptimo.

En la terminología del diseño de experimentos, los parámetros y las hipótesis acerca de su estructura, que componen un modelo, se llaman *factores* y la medida su comportamiento se llama *respuesta*.

Los factores pueden ser:

- *cuantitativos*: aquellos que naturalmente toman valores numéricos, por ejemplo, el número de empleados en un banco.
- *cualitativos*: normalmente representan hipótesis acerca de la estructura de del modelo que no tienen un significado numérico, por ejemplo, la disciplina de la cola (FIFO, LIFO, etc.).

En los experimentos de simulación, también pueden clasificarse los factores en:

- *factores controlables*: representan opciones modificables en el sistema real, por ejemplo, el número de empleados en un banco o la disciplina de la cola.
- *factores incontrolables*: aquellos no manipulables en el sistema real, por ejemplo, la frecuencia de llegada de los clientes al banco o el volumen medio de demandas de un producto.

En un modelo determinado, la elección de que parámetros o hipótesis acerca de su estructura son considerados fijos y cuales otros son los factores experimentales depende más del objetivo del estudio que de la naturaleza del modelo.

El *diseño de experimentos* proporciona un modo de decidir qué simular, para obtener la información deseada al mínimo coste, es decir, con la menor cantidad de simulación. Los experimentos cuidadosamente diseñados son mucho más eficientes que una secuencia de ejecuciones en las cuales se simula un número de sistemas alternativos "para ver lo que pasa". El diseño del experimento es particularmente útil cuando no se tiene demasiado conocimiento acerca de qué factores son importantes y afectan a la respuesta.

Si el modelo sólo tiene un factor, el diseño del experimento es relativamente simple: se ejecuta la simulación para varios valores, o *niveles*, del factor, estableciendo, posiblemente, un intervalo de confianza para la respuesta verdadera en cada uno de los niveles del factor. Para factores cualitativos, suele ser útil construir un gráfico de la respuesta en función del nivel del factor.

Supongamos que el modelo contiene k ($k > 1$) factores que influyen en su respuesta, y que se desea obtener una estimación inicial de cómo influye cada factor sobre la respuesta y de cómo los factores interactúan entre sí, es decir, en qué medida depende el efecto de un factor sobre la respuesta del nivel de otro. Un modo de medir el efecto de un factor en particular, puede ser fijar los niveles de los $k-1$ restantes a un conjunto de valores dados y realizar simulaciones para cada uno de los niveles de este factor. El proceso completo es entonces repetido para examinar el resto de los factores, cada uno a un tiempo. Esta estrategia es, sin embargo, bastante poco eficiente en términos del número de simulaciones necesario y no permite medir las interacciones entre los factores.

11.1.2- Diseño 2^k factorial

Una estrategia menos costosa y que permite medir interacciones entre factores es la estrategia de *diseño 2^k factorial*. Esta técnica de diseño de experimentos requiere que se seleccionen sólo dos niveles para cada factor (notándose normalmente "-" el nivel de menor valor numérico y "+" el de mayor) y consiste en la simulación de cada una de las 2^k posibles combinaciones de niveles de los factores. La forma del experimento puede representarse en forma de tabla, llamada *matriz de diseño*. Por ejemplo, para $k=3$:

Combinación	Factor 1	Factor 2	Factor 3	Respuesta
1	-	-	-	R_1
2	+	-	-	R_2
3	-	+	-	R_3
4	+	+	-	R_4
5	-	-	+	R_5
6	+	-	+	R_6
7	-	+	+	R_7
8	+	+	+	R_8

Las variables R_i (para $i=1,2,\dots,8$), son los valores de la respuesta cuando se ejecuta la simulación con la i -ésima combinación de los niveles.

El *efecto principal* del factor j , e_j , es el cambio promedio en la respuesta producido por el paso del factor j de su nivel - a su nivel +. Mide el cambio promedio en la respuesta debido al cambio en un factor individual. Se calcula como la diferencia entre la respuesta promedio cuando el factor j está en nivel + y la respuesta promedio cuando está en nivel -.

Para el diseño 2^3 factorial de la tabla:

$$\begin{aligned}\text{efecto principal del factor 1: } e_1 &= \frac{-R_1 + R_2 - R_3 + R_4 - R_5 + R_6 - R_7 + R_8}{4} \\ \text{efecto principal del factor 2: } e_2 &= \frac{-R_1 - R_2 + R_3 + R_4 - R_5 - R_6 + R_7 + R_8}{4} \\ \text{efecto principal del factor 3: } e_3 &= \frac{-R_1 - R_2 - R_3 - R_4 + R_5 + R_6 + R_7 + R_8}{4}\end{aligned}$$

Puede suceder que el efecto del factor j_1 dependa del nivel de otro factor, j_2 , en cuyo caso se dice que los factores j_1 y j_2 interactúan. El *efecto interactivo*, $e_{j_1 j_2}$, entre los factores j_1 y j_2 , se define como la semidiferencia entre el promedio del efecto del factor j_1 cuando el factor j_2 está en su nivel - y el promedio del efecto del factor j_1 cuando el factor j_2 está en su nivel +.

Para el diseño 2^3 factorial de la tabla:

$$\begin{aligned}e_{12} &= \frac{1}{2} \left[\frac{-R_3 + R_4 - R_7 + R_8}{2} - \frac{-R_1 + R_2 - R_5 + R_6}{2} \right] \\ e_{13} &= \frac{1}{2} \left[\frac{-R_5 + R_6 - R_7 + R_8}{2} - \frac{-R_1 + R_2 - R_3 + R_4}{2} \right] \\ e_{23} &= \frac{1}{2} \left[\frac{-R_5 - R_6 + R_7 + R_8}{2} - \frac{-R_1 - R_2 + R_3 + R_4}{2} \right]\end{aligned}$$

Los efectos interactivos son simétricos, es decir: $e_{j_i j_k} = e_{j_k j_i}$

Ejemplo. En el ejemplo de inventario, existen dos factores experimentales: s y S. Los niveles - y + que escogemos para los factores s y S vienen dados en la *tabla de codificación*.

	-	+
s	20	60
S	70	120

Tabla de codificación

Factor	s	S	Respuesta
1	-	-	118.280
2	+	-	141.060
3	-	+	136.807
4	+	+	152.789

Matriz de diseño

Los efectos principales son:

$$\begin{aligned}e_s &= \frac{-118.280 + 141.060 - 136.807 + 152.789}{2} = 19.381 \\ e_S &= \frac{-118.280 - 141.060 + 136.807 + 152.789}{2} = 15.128\end{aligned}$$

y el efecto interactivo entre s y S es:

$$e_s = \frac{118.280 - 141.060 - 136.807 + 152.789}{2} = -3.399$$

Así pues, el efecto promedio de aumentar s de 20 a 60 es aumentar el coste mensual en 19.381, y el de aumentar S de 70 a 120 es aumentar el coste mensual en un promedio de 15.128. Así pues, parece que si se desea el mínimo coste mensual posible son preferibles valores pequeños de s y S.

El efecto interactivo negativo es una indicación de que se consiguen costes menores escogiendo, o bien s y S a -, o bien s y S a +, en vez de escoger uno a - y el otro a +.

Como las respuestas R_i 's son variables aleatorias, los efectos son también aleatorios. Para determinar su fiabilidad deben estimarse sus varianzas. Un método sencillo es replicar independientemente el experimento n veces, obteniendo n valores independientes de e_s , e_S y e_{sS} . Escogiendo n=10, se obtienen:

$$\begin{aligned}\bar{e}_s &= 21.539 \\ \text{Var}(\bar{e}_s) &= 0.879\end{aligned}$$

un intervalo del 90% de confianza (basado en una distribución t con 9 grados de libertad) para el efecto principal esperado, E(s), es: 21.539 ± 1.719 .

Análogamente, los intervalos del 90% de confianza para E(e_s) y E(e_{sS}) son, respectivamente:

$$10.459 \pm 2.002 \quad \text{y} \quad -2.971 \pm 0.794$$

11.1.3- Diseños 2^{k-p} factoriales fraccionales

Para un modelo con k factores, el diseño de la sección anterior requiere al menos una simulación para cada una de las 2^k combinaciones posibles de factores. Si el modelo tiene, por ejemplo, k=11 parámetros, hay $2^{11} = 2048$ combinaciones. Si se desea estimar las varianzas de los efectos, realizando, por ejemplo, n=5 replicaciones en cada punto de diseño, el número total de simulaciones que hay que realizar es 10240. Si cada simulación tiene un tiempo de ejecución de 1 minuto, se necesitarán más de 7 días de computación ininterrumpida para realizar las 10240 simulaciones.

Los diseños factoriales fraccionales son muy útiles en las primeras fases de la experimentación, cuando están presentes muchos factores y se desean descartar los factores relativamente no importantes sin que para ello haya que realizar una cantidad excesiva de simulación.

Básicamente, un diseño factorial fraccional se construye seleccionando un subconjunto, de tamaño 2^{k-p} , de las 2^k combinaciones posibles de factores y ejecutando las simulaciones sólo para las combinaciones elegidas. La elección de las 2^{k-p} combinaciones depende de qué efectos principales e interacciones son de mayor interés.

Ejemplo. Consideremos una generalización del modelo del inventario en el cual se introducen dos nuevos factores, además de s y S :

- El primer factor es el intervalo de evaluación del inventario, m , que es el número de meses entre evaluaciones sucesivas del inventario, en las cuales se determina si debe realizarse un pedido al suministrador. En el modelo original $m=1$, pero desea estudiarse el efecto de fijar $m=3$.
- El segundo factor surge debido a que el suministrador ha introducido una opción "expres". Originalmente, si se ordenan Z unidades de producto, el coste es $3200+300Z$ y el tiempo que tardará en recibirse el producto es una variable aleatoria $U(0.5,1.0)$. Si se utiliza el servicio "expres", el producto se recibirá en la mitad de tiempo: el tiempo que transcurre desde que se ordena el pedido hasta que se recibe es una variable aleatoria $U(0.25,0.5)$; a cambio, el coste del servicio "expres" es $4800+400Z$. Así pues, se introduce el factor cualitativo de prioridad, P , del pedido, que puede ser "normal" o "expres".

En este modelo generalizado, hay $k=4$ parámetros:

	-	+
s	20	60
S	70	120
m	1	3
P	Normal	Exprés

Tabla de codificación

Un diseño completo 2^4 factorial requeriría 16 combinaciones posibles de factores. Sin embargo, si el objetivo es obtener estimas aceptables de los efectos principales, podemos considerar sólo la mitad de esta combinaciones, al coste de perder la capacidad de calcular las interacciones. Se escogen las siguientes 8 combinaciones de los factores:

Factor	s	S	m	P	Respuesta
1	-	-	-	-	118.686
2	+	-	-	+	196.119
3	-	+	-	+	166.674
4	+	+	-	-	149.542
5	-	-	+	+	148.045
6	+	-	+	-	135.754
7	-	+	+	-	153.202
8	+	+	+	+	181.885

Matriz de diseño

Una vez realizadas las simulaciones, se calculan los efectos del mismo modo que en los diseños 2^k factoriales; por ejemplo, el efecto principal de s se calcula aplicando los signos de la columna " s " a las correspondientes respuestas, realizando la suma y dividiendo el resultado por $2^{k-p-1} = 4$. Se obtiene:

$$e_s = 19.173$$

$$e_S = 13.175$$

$$e_m = -3.034$$

$$e_p = 33.885$$

Para estimar las varianzas de los efectos principales, se replica el experimento $n=10$ veces y se construye un intervalo del 90% de confianza para los valores esperados de los efectos principales:

para $E(e_s)$:	15.478 ± 1.831
para $E(e_S)$:	14.431 ± 2.017
para $E(e_m)$:	-1.217 ± 1.449
para $E(e_p)$:	34.915 ± 1.775

Los efectos principales de s y S son significativamente positivas. El efecto principal de m es muy pequeño y estadísticamente insignificante, lo cual indica que no hay que incluirlo como factor en futuras simulaciones, al menos, en el rango en que hemos especificado sus niveles. La opción "expres" aumenta considerablemente el coste de operación, con lo cual parece que no compensa.

11.2- Optimización. Metodología de la superficie de respuesta

Un programa de experimentación típico puede consistir en realizar diseños factoriales completos o fraccionarios para identificar los factores relevantes y, a continuación, examinar estos factores con más detalle para encontrar la combinación de niveles que conduce a que la respuesta se maximice o minimice.

Cuando todos los factores son cuantitativos, pueden usarse, para realizar esta optimización, el conjunto de técnicas conocidas como metodologías de la superficie de respuesta. La *superficie de respuesta* es el valor esperado de la respuesta en función de los factores.

El siguiente ejemplo ilustra la idea principal de estos métodos.

Ejemplo. Consideremos el modelo original de inventario con 2 factores: s y S . El método puede comenzar con diseño factorial 2^2 . Las combinaciones 1, 2, 3, 4 de los factores (s, S) (que son $(-, -)$, $(+, -)$, $(-, +)$ y $(+, +)$) pueden representarse como los vértices de un cuadrado en el plano (s, S). Las respuestas medias (de 10 replicaciones independientes) obtenidas en estos puntos son: 119.684, 144.194, 133.114 y 151.682 respectivamente.

Como una aproximación lineal de la superficie de respuesta cubierta por este diseño (el área comprendida dentro del cuadrado), podemos ajustar por mínimos cuadrados el modelo lineal

$$E[R(s, S)] = \beta_0 + \beta_1 s + \beta_2 S$$

a la respuesta en los 4 puntos de diseño. Realizando los cálculos, se obtienen los siguientes coeficientes del ajuste por mínimos cuadrados:

$$\hat{\beta}_0 = 95.757 \quad \hat{\beta}_1 = 0.538 \quad \hat{\beta}_2 = 0.209$$

Dado que en este caso interesa calcular la respuesta (coste) mínima, se examina el modelo lineal ajustado, $E[R(s, S)] = \beta_0 + \beta_1 s + \beta_2 S$, para determinar qué dirección en el espacio (s, S) es la de mayor decrecimiento de la respuesta.

El negativo del vector de las derivadas parciales de una función, particularizadas en un punto, apunta en la dirección de más rápido descenso de la función en ese punto. En el caso del modelo lineal, la dirección de más rápido descenso de la respuesta en el plano (s,S) , es paralela al vector $(-\hat{\beta}_1, -\hat{\beta}_2)$, que, en nuestro caso, es igual a $(-0.538, -0.209)$. Así pues, la búsqueda debe desplazarse, desde el punto central del cuadrado, por el plano (s,S) , en la dirección $(-0.538, -0.209)$.

La búsqueda debe desplazarse realizando simulaciones para (s,S) , avanzando en la dirección del vector, hasta que se observe que la respuesta comience a crecer. Entonces, se retrocede hasta la respuesta anterior menor y se usa como centro para un nuevo diseño 2^2 factorial, del cual obtenemos un nuevo ajuste lineal y una dirección de más rápido descenso.

Cuando el ajuste lineal sea aproximadamente plano, es decir, cuando $\hat{\beta}_1$ y $\hat{\beta}_2$ están ambos próximos a cero, es indicación de que estamos cerca de un mínimo (posiblemente, sólo local). En este punto, puede realizarse un diseño experimental más elaborado que el 2^2 factorial, que nos permita realizar el ajuste a un modelo cuadrático:

$$E[R(s, S)] = \beta_0 + \beta_1 s + \beta_2 S + \beta_{12} sS + \beta_{11} s^2 + \beta_{22} S^2$$

del cual calcularemos el punto mínimo.

El problema de la optimización en simulación es considerablemente difícil: la optimización determinista, de por sí, es un problema complejo, al cual hay que añadir, la dificultad de la naturaleza aleatoria de las salidas de la simulación, que hace imposible calcular exactamente la función objetivo (superficie respuesta) a optimizar. En ocasiones, los problemas de optimización requieren un número inalcanzable de simulaciones. Además, no existe garantía de que el extremo calculado sea un extremo absoluto.