

Capítulo 4:

Variabilidad cuando Y depende de 1 factor

ruido: Estudio de $V(Y)$

4 Variabilidad cuando Y depende de 1 factor ruido: Estudio de V(Y)

En el capítulo 2 dedicado al estado del arte, presentábamos las dos maneras principales de analizar la variabilidad transmitida por los factores ruido a la respuesta:

- A) Estimando directamente la variabilidad utilizando métricas resumen. Para lograr este objetivo todas las condiciones experimentales en los factores de control se han de someter a las mismas condiciones de ruido y se aproxima $V(Y)$ por polinomios del tipo:

$$V(Y) = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \varepsilon \quad i, j = 1, \dots, k \quad \varepsilon \text{ -- iid -- } \rightarrow N(0, \sigma) \quad (4.1)$$

- B) Estimando primero el modelo que relaciona la respuesta Y con los factores de control y ruido,

$$Y = F(X's) + \sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{il} X_i Z_l + \varepsilon \quad (4.2)$$

$i = 1, \dots, k \quad l, m = 1, \dots, r \quad \varepsilon \text{ -- iid -- } \rightarrow N(0, \sigma)$

y deduciendo el comportamiento de $V(Y)$ a partir de este modelo por dos métodos:

- B.1) Método gráfico: En el que se estudian las interacciones entre factores de control y factores ruido para determinar las condiciones de los factores de control que minimizan el efecto de los factores ruido;
- B.2) Método analítico: A partir del modelo estimado para Y en la ecuación (4.2) se obtiene el modelo para $V(Y)$ donde la parte principal se debe a la varianza transmitida por los factores ruido $V_Z(Y)$,

$$V(Y) = V\left(\sum_{l=1}^r \theta_l Z_l + \sum_{l \neq m} \theta_{lm} Z_l Z_m + \sum_{i=1}^k \sum_{l=1}^r \delta_{il} X_i Z_l\right) + V(\varepsilon) = V_Z(Y) + V_\varepsilon(Y)$$

$i = 1, \dots, k \quad l, m = 1, \dots, r$

$$V_Z(Y) = \beta_0^* + \sum_{i=1}^k \beta_i^* X_i + \sum_{i \neq j} \beta_{ij}^* X_i X_j \quad i, j = 1, \dots, k \quad (4.3)$$

Estas diferencias de partida dan lugar en general a diferentes estrategias de experimentación: Puesto que con el primer método se persigue estimar el modelo para la variabilidad, se seleccionan diseños producto donde una parte importante del presupuesto se asigna a los factores ruido, que son los que inducen esta variabilidad (Taguchi and Wu (1980)). En cambio, cuando se trata de modelar el valor esperado de la respuesta, es decir, la ecuación (4.2), se tiende a seleccionar las condiciones experimentales de una manera menos redundante, de tal manera que permitan estimar los coeficientes de este modelo con un mínimo coste (Jones (1990), Shoemaker et al. (1991), Wu et al. (1999), Bisgaard (2000) y Wu y Hamada (2000) entre otros).

En general, un diseño seleccionados para estimar un modelo para la respuesta Y sólo permite llegar a $V(Y)$ por medio del modelo descrito en (4.3) mientras que si se selecciona el diseño a partir de una matriz producto, se puede aproximar $V(Y)$ por las dos vías descritas en (4.1) y en (4.3) y se espera llegar a las mismas conclusiones.

Sin embargo existen casos donde esto no ocurre y prueba de ellos son los dos ejemplos desarrollados ampliamente en el capítulo 3. Brevemente, las dos situaciones con las que nos encontrábamos eran:

- Ejemplo 1: Diseño fraccionado, $2^{7-4} \times 2^{3-1}$, de Resolución III. En este ejemplo al analizar directamente la variabilidad vía métricas resumen se seleccionan como significativos los factores C y E y en cambio al estimar la variabilidad apoyándose en el modelo para la respuesta se selecciona el factor F.
- Ejemplo 2: Diseño completo, $2^3 \times 2^2$. Cuando se utilizan métricas resumen surgen como significativos los factores B y C; en cambio al modelar directamente la respuesta el factor A es el único que aparece interaccionando de una manera claramente significativa con un factor ruido.

Objetivos

Motivados por los diferentes resultados que se puede llegar a obtener por los dos métodos de análisis, en este capítulo y en los tres siguiente pretendemos:

- Partiendo de un modelo teórico para la respuesta Y, estudiar la estructura de los modelos asociados a las superficies de respuesta más utilizadas en la práctica en este tipo de estudios: $\sigma_Z^2(Y)$, $\sigma_Z(Y)$ y $\log(\sigma_Z)$;
- Reflexionar sobre las consecuencias de analizar los problemas de robustez tal y como se hacen habitualmente a partir de diseños factoriales a 2 niveles y por medio de métricas resumen o métricas no-resumen.

- Establecer unas pautas metodológicas a introducir en la práctica experimental que nos permitan abordar el estudio de $V(Y)$ con éxito.

Hipótesis

A pesar de que el objetivo de los problemas de robustez es el estudio de la variabilidad transmitida por los factores ruido en las distintas condiciones de los factores de control, el experimentador está menos familiarizado con este tipo de respuestas que con la respuesta Y , y por ello en la práctica se seleccionan factores de control y ruido partiendo de su efecto lineal en esta última respuesta. Por esta razón, en todo nuestro trabajo asumiremos que se ha seleccionado una respuesta Y cuya relación con los factores de control, X_i , y factores ruido, Z_l , puede ser representada, en la región experimental, por un polinomio de primer orden con a lo sumo términos cruzados de segundo orden,

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i \neq j} \beta_{ij} X_i X_j + \sum_{l=1}^r \theta_l Z_l + \sum_{i=1}^k \sum_{l=1}^r \delta_{ij} X_i Z_l + \sum_{l \neq m}^r \theta_{lm} Z_l Z_m + \varepsilon \quad (4.4)$$

$$i, j = 1, \dots, k \quad l, m = 1, \dots, r \quad \varepsilon - iid -> N(0, \sigma)$$

y realizaremos un estudio de la respuesta $V(Y)$ asociada a este modelo.

El estudio se realizará principalmente a partir de ejemplos y podremos observar que:

- La superficie de respuesta para $V(Y)$ tiene en general una estructura más compleja que la superficie Y , y por lo tanto para estimarla adecuadamente se necesitan diseños más complejos que los diseños 2^k que se utilizan habitualmente;
- Cuando se estima $V(Y)$ con los diseños 2^k es fácil interpretar erróneamente el comportamiento de esta superficie.

4.1 Estudio de $V(Y)$ para 1 factor ruido y 1 factor de control

Supongamos que el modelo para la respuesta Y puede aproximarse por un modelo como el que sigue,

$$Y = \mu + \theta Z + \delta XZ + \varepsilon \quad \text{con } \delta \neq 0 \quad (4.5)$$

donde: “ μ ” representa la parte de Y afectada sólo por factores de control (X y otros); “ $\theta Z + \delta XZ$ ” representa la influencia del factor ruido Z en la respuesta y “ ε ” representa a la variación aleatoria sobre el valor esperado de Y, variación que suponemos independiente del resto de la expresión.

Así, la expresión para la varianza de Y en función de sus dos fuentes de variación no controlables, Z y ε , queda:

$$V(Y) = V(Z)(\theta + \delta X)^2 + V(\varepsilon) = V_Z(Y) + V_\varepsilon(Y) \quad (4.6)$$

Caso $\theta \neq 0$

Forma teórica para $V_Z(Y)$

Salvo una constante de proporcionalidad, la expresión para la varianza transmitida únicamente por el factor ruido Z toma la forma de un polinomio de segundo grado en X que llamaremos “Q”:

$$V_Z(Y) \approx (\theta + \delta X)^2 = \theta^2 + (2\theta\delta)X + (\delta^2)X^2 \quad (4.7)$$

$$Q = \beta_0 + \beta_1 X + \beta_{11} X^2 \quad \text{con } \beta_0 = \theta^2; \quad \beta_1 = 2\theta\delta; \quad \beta_{11} = \delta^2$$

Podemos ver que la relación entre $V(Y)$ y el factor de control X es más compleja que la relación entre Y y el mismo factor, ya que ahora aparece el término cuadrático puro X^2 .

La superficie que describe la ecuación (4.7) es una parábola con un mínimo (ya que el coeficiente de X^2 es positivo), y por lo tanto existe una condición en X donde la variabilidad transmitida por el factor ruido Z a Y es mínima:

$$X_Q = -\frac{\theta}{\delta}$$

Si $|\theta| \leq |\delta|$ la condición robusta está dentro de la región experimental (-1, 1)

Aproximación a $V_Z(Y)$ a partir de los puntos de un diseño factorial 2^1

En la práctica, la aproximación a $V(Y)$ se realiza habitualmente a partir de diseños factoriales a dos niveles. En estos diseño se cumple que $X^2 = 1$ y por ello no permiten

estimar el coeficiente de X^2 en la expresión “Q”. Por consiguiente se aproxima $V_Z(Y)$ por un polinomio de primer orden que hemos denominado “L”:

$$\hat{V}_Z(Y) \approx b_0 + b_1 X = L \quad \text{con} \quad E(b_0) = \beta_0 + \beta_{11} = \theta^2 + \delta^2 \quad \text{y} \quad E(b_1) = \beta_1 = 2\theta\delta \quad (4.8)$$

“L” representa a una recta de pendiente no nula que corta a “Q” en los puntos $X = \pm 1$. El experimentador interpreta que siguiendo la dirección de pendiente negativa se conseguiría reducir “indefinidamente” la varianza. Si, como se hace habitualmente, el estudio se restringe a las dos condiciones donde se ha experimentado, $X = -1$ y $X = 1$, la condición de mínima varianza ha de cumplir,

$$X_L = -1 * \text{signo}(b_0 b_1) = -1 * (\text{signo} \theta \delta)$$

Notamos que X_L no tiene por qué coincidir con la solución robusta teórica X_Q .

Caso $\theta = 0$

Forma teórica para $V_Z(Y)$

Si en el modelo para Y el efecto principal del factor ruido es nulo ($\theta = 0$), pero aparece como significativa la interacción del factor ruido con el factor de control ($\delta \neq 0$), el modelo teórico “Q” para la varianza transmitida por el factor ruido Z queda de la siguiente forma:

$$Q = V_Z(Y) = (\delta^2) X^2 \quad (4.9)$$

Es decir, un polinomio puro de segundo grado que representa a una parábola con mínimo en $X = 0$. En este caso la solución teórica al problema es,

$$X_Q = 0$$

Aproximación a $V_Z(Y)$ a partir de los puntos de un diseño factorial 2^1

La aproximación “L” a la que se llega a partir de un diseño 2^1 representa a una recta constante para cualquier valor de X ,

$$L = V_Z(Y) = \delta^2 \quad (4.10)$$

En este caso el experimentador llega a la conclusión errónea de que la varianza transmitida a Y por el factor Z es independiente del nivel en que se encuentre X y que por lo tanto el problema de robustez no tiene solución. Sin embargo cualquier condición en X dentro del intervalo $(-1, 1)$ reduce $V_Z(Y)$ y, lo que es más peligroso, cualquier condición de X fuera de este intervalo la aumenta.

En general, la diferencia entre los valores estimados por “Q” y “L” es:

$$Q - L = \delta^2 (X^2 - 1)$$

y por lo tanto “L” sobrestima la variabilidad en puntos interiores de la zona experimental y subestima fuera de esta zona, estimando adecuadamente en los puntos del diseño.

Concluimos que la aproximación “L” discrimina bien entre las dos opciones, $X=1$ y $X=-1$, aquella de mínima varianza, X_L , pero no tiene por qué coincidir con la solución robusta teórica X_Q . Además “L” no identifica adecuadamente la naturaleza de la superficie por lo que interpreta erróneamente el problema.

4.1.1 Ejemplos

En la Figura 4.1 presentamos la representación gráfica de las superficies “Q” y “L” asociadas a $V_Z(Y)$ para distintos modelos de Y :

- | | | |
|----------------------|----------------------|---------------------|
| a) $Y = 4 Z + 1 XZ;$ | b) $Y = 2 Z + 1 XZ;$ | c) $Y = 1 Z + 1 XZ$ |
| d) $Y = 0 Z + 1 XZ;$ | e) $Y = 1 Z + 5 XZ;$ | f) $Y = 0 Z + 5 XZ$ |

El rango de variación de X ha sido seleccionado de una manera mucho más amplia que lo habitual en la experimentación ordinaria con el fin de dar cabida a la solución teórica al problema de robustez.

Si comparamos los dos ejemplos a) y e) la superficie “L” selecciona el punto $X = -1$ como condición robusta y en ambos casos interpreta que reduciendo X más allá de este valor se reduce la varianza cuando esto no es así en el caso e).

Si comparamos los dos ejemplos a) y b) vemos que siguiendo la dirección de decrecimiento de varianza estimada por “L” se puede llegar a entrar, en el caso b) en una zona de aumento real de varianza. Esto se debe a que “L” no detecta la proximidad o lejanía de la solución robusta X_Q .

En los casos d) y f) “L” interpreta que el problema no tiene solución robusta y sin embargo no es así. Además, cualquier condición de X fuera del intervalo $(-1, 1)$ da lugar a una variabilidad mayor que la observada en los puntos del diseño.

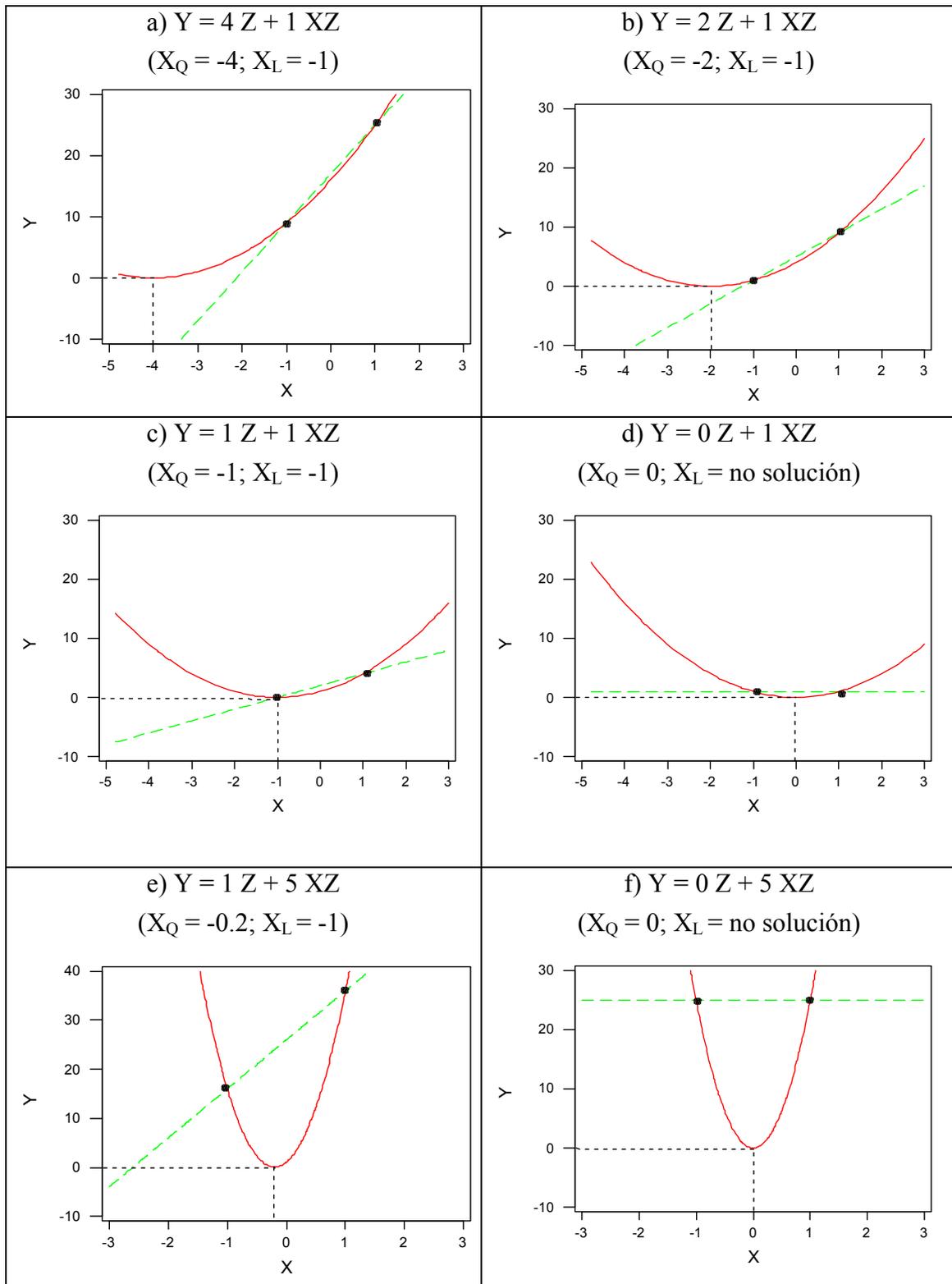


Figura 4.1. Representación teórica de $V_Z(Y)$ (en rojo) y aproximación por polinomios de primer grado (en verde) con $Y = \theta Z + \delta XZ$

4.2 Estudio de V(Y) para 1 factor ruido y 2 factores de control

Supongamos en este caso que el modelo que relaciona la característica de interés Y con los dos factores de control X₁ y X₂ y el factor ruido Z puede aproximarse por un polinomio de primer orden con términos cruzados de segundo orden,

$$Y = \mu + \theta Z + \delta_1 X_1 Z + \delta_2 X_2 Z + \varepsilon \quad \text{con } \delta_1 \neq 0 \text{ y } \delta_2 \neq 0 \quad (4.11)$$

Asumiendo que “ε” es independiente del resto de la expresión, podemos descomponer V(Y) en dos componentes independientes, una dependiente de Z y otra de “ε”,

$$V(Y) = V_Z(Y) + V_\varepsilon(Y)$$

Así, la varianza transmitida por el factor ruido Z a la característica Y puede ser expresada, salvo un factor de proporcionalidad, de la siguiente forma,

$$V_Z(Y) \approx (\theta + \delta_1 X_1 + \delta_2 X_2)^2$$

$$Q = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \beta_{11} X_1^2 + \beta_{22} X_2^2 \quad (4.12)$$

$$\text{con } \beta_0 = \theta^2; \beta_1 = 2\theta\delta_1; \beta_2 = 2\theta\delta_2; \beta_{12} = 2\delta_1\delta_2; \beta_{11} = \delta_1^2; \beta_{22} = \delta_2^2$$

A igual que en el apartado anterior, se puede observar que el modelo para V(Y) es más complejo en X₁ y X₂ que el modelo inicial para Y, ya que necesita términos de segundo orden: interacciones y cuadráticos puros.

Los coeficientes de los términos cuadráticos puros de la superficie “Q” son positivos, por lo tanto la superficie teórica es un paraboloide con una región de valores mínimos que depende de los valores θ, δ₁ y δ₂. Las coordenadas de los puntos situados en esta región de mínima varianza (X_{1Q}, X_{2Q}) cumplen la ecuación,

$$\delta_1 X_{1Q} + \delta_2 X_{2Q} + \theta = 0$$

La distancia de esta región al centro del diseño (0,0) aumenta conforme θ crece para valores fijos de δ₁ y δ₂,

$$d[(X_{1Q}, X_{2Q}), (0,0)] = \frac{\theta}{\sqrt{\delta_1^2 + \delta_2^2}}$$

Si se aproxima $V_Z(Y)$ a partir de un diseño factorial a dos niveles, por cumplirse $X_1^2 = X_2^2 = 1$ en todas las condiciones experimentales, la expresión de la varianza queda reducida a un polinomio de primer orden en X_1 y X_2 con términos cruzados, que denominamos “L”:

$$\hat{V}_Z(Y) \approx b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2 = L \quad (4.13)$$

con

$$E(b_0) = \beta_0 + \beta_{11} + \beta_{22} = (\theta^2 + \delta_1^2 + \delta_2^2)$$

$$E(b_1) = \beta_1 = 2\theta\delta_1 \quad (4.14)$$

$$E(b_2) = \beta_2 = 2\theta\delta_2$$

$$E(b_{12}) = \beta_{12} = 2\delta_1\delta_2$$

La ecuación en (4.13) para “L” contiene únicamente el término de segundo orden $X_1 X_2$ y representa a una “silla de montar”.

La diferencia entre los valores estimados con ambas superficies no depende del parámetro θ y puede obtenerse a partir de la expresión

$$Q - L = \delta_1^2 (X_1^2 - 1) + \delta_2^2 (X_2^2 - 1)$$

Por lo tanto, ambas superficies coinciden en los puntos del diseño factorial y son diferentes en el resto. Dentro de la región experimental, el error de estimación se hace máximo en el centro (0,0)

$$Q - L = \delta_1^2 + \delta_2^2$$

Si el punto (X_1, X_2) está fuera de la región experimental, “L” subestima $V_Z(Y)$; en cambio, la variabilidad en las condiciones internas a la zona experimental, condiciones que interesan más en la experimentación, es superestimada.

4.2.1 Ejemplos

A continuación mostramos gráficamente distintas situaciones de las superficies “Q” y “L” a partir de 5 ejemplos de modelos para Y donde se ha mantenido la estructura del modelo y se ha variado la magnitud de los coeficientes para mover la zona de mínima varianza.

Figura 4.2) $\theta = 4; \delta_1 = 1; \delta_2 = 1;$

Figura 4.3) $\theta = 2; \delta_1 = 1; \delta_2 = 1;$

Figura 4.4) $\theta = 0; \delta_1 = 1; \delta_2 = 1;$

Figura 4.5) $\theta = 2; \delta_1 = 5; \delta_2 = 1;$

Figura 4.6) $\theta = 0; \delta_1 = 5; \delta_2 = 1.$

$$\underline{Y = 4Z + 1X_1Z + 1X_2Z}$$

En este caso, se puede apreciar cierta curvatura en “Q” dentro de la región seleccionada, aunque alejada de lo que acostumbra a ser la zona de experimentación cuando se trabaja con diseños factoriales a dos niveles.

La curvatura la marca la zona de mínima varianza representada por las coordenadas (X_{1Q}, X_{2Q}) que cumplen,

$$X_{2Q} + X_{1Q} + 4 = 0$$

Al coincidir la superficie “L” con la superficie “Q” en las condiciones del cubo representadas por $X_1 = \pm 1$ y $X_2 = \pm 1$, “L” discrimina bien entre las cuatro condiciones del cubo la de mínima varianza, estimando adecuadamente su valor.

Dentro del cubo “L” identifica adecuadamente la dirección a seguir para minimizar la variabilidad aunque la estimación de esta no sea correcta. La zona de mínima varianza teórica queda alejada del cubo y en esta zona “L” continua su tendencia decreciente adoptando valores negativos, es decir, no es una buena aproximación.

Por lo tanto teniendo en cuenta sólo la forma de las superficies, “L” resulta una buena aproximación en cuanto a forma, dentro de la zona de experimentación aunque los valores obtenidos no son correctos (salvo en las condiciones del cubo). En consecuencia, se debería utilizar “L” como una superficie orientativa que marca una dirección a seguir, pero no para estimar $V_Z(Y)$ en condiciones que no sean las experimentales.

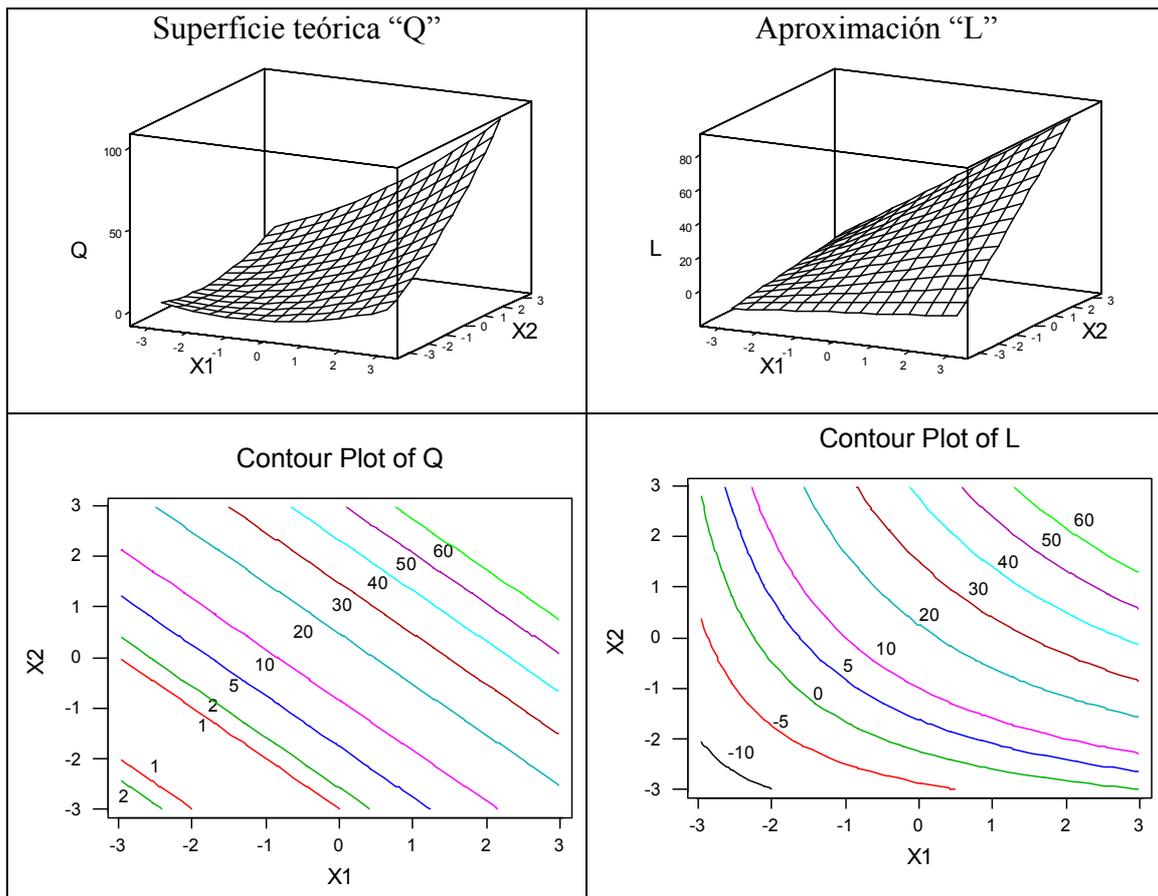


Figura 4.2. Representación de la superficie $V_Z(Y)$ con $Y = 4Z + 1X_1Z + 1X_2Z$
(Q: superficie teórica; L: aproximación por un polinomio sin términos cuadráticos puros)

$$\underline{Y = 2Z + 1X_1Z + 1X_2Z}$$

En este caso, la superficie teórica “Q” posee una zona de mínima varianza representada por las condiciones (X_{1Q}, X_{2Q}) que cumplen:

$$X_{2Q} + X_{1Q} + 2 = 0$$

Esta zona está más centrada en la zona de experimentación ya que incluye al vértice del cubo $(-1, -1)$; por esto la curvatura es más evidente que en el caso anterior.

La aproximación “L” es peor que en el caso anterior, necesitando la superficie “doblarse” para adaptarse a la curvatura y como consecuencia, en el modelo aparecerá el término de la interacción.

Si las direcciones de decrecimiento de “L” en que salen valores negativos fueran interpretadas como direcciones de minimización de la variabilidad y se experimentara en estas condiciones, nos encontraríamos, contrariamente a lo esperado, con un aumento de la variabilidad. Por lo tanto la forma de “L” no se aproxima a la de “Q” en estas direcciones que están muy próximas a la zona de mínima varianza teórica.

Como conclusión, la superficie “L” estima adecuadamente la variabilidad en las cuatro condiciones del cubo representadas por $X_1 = \pm 1$ y $X_2 = \pm 1$ pero no así en el resto de condiciones. Además, es peligroso tomar a “L” como una superficie orientativa para seleccionar la dirección de reducción de $V_Z(Y)$ ya que se logrará el efecto contrario.

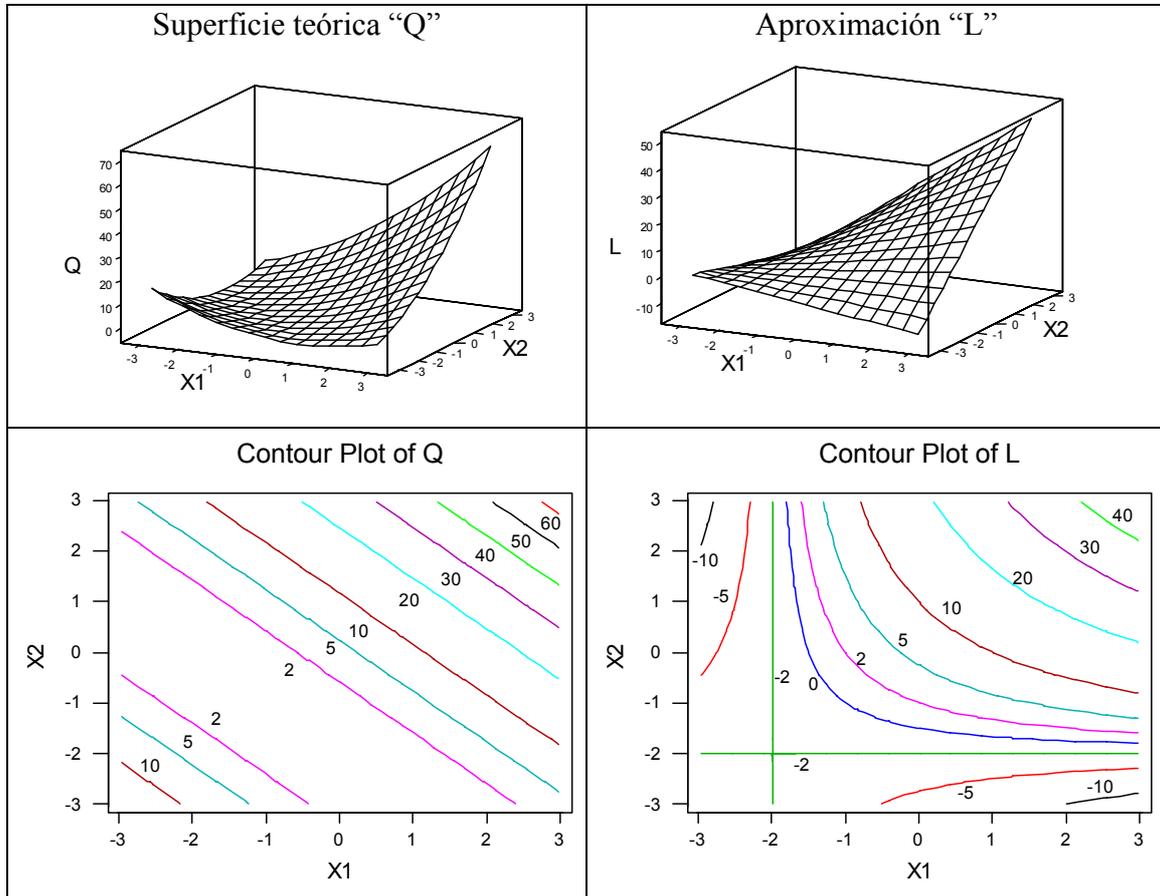


Figura 4.3. Representación de $V_Z(Y)$ con $Y = 2Z + 1X_1Z + 1X_2Z$
(Q: superficie teórica; L: aproximación por un polinomio sin términos cuadráticos puros)

$Y = 0Z + 1X_1Z + 1X_2Z$

En este caso en el que en el modelo para Y el efecto principal del factor ruido se anula, la superficie “Q” posee una zona de mínima varianza en las condiciones (X_{1Q}, X_{2Q}) situadas en una recta pasando por $(0, 0)$, $(-1, 1)$ y $(1, -1)$:

$$X_{2Q} + X_{1Q} = 0$$

La superficie “Q” muestra una curvatura significativa dentro de la región delimitada por $X_1 = \pm 1$ y $X_2 = \pm 1$. (Ver Figura 4.4).

La aproximación “L” adopta la forma de una “silla de montar” centrada en (0, 0). En este valor “L” superestima claramente la variabilidad.

La dirección de varianza mínima de $V_Z(Y)$ es la dirección peor estimada por la superficie “L”, ya que la toma como una dirección no acotada de reducción de la variabilidad, es decir, “L” invita erróneamente a alejarse del centro de la experimentación. Un indicio de la falta de ajuste del modelo a la superficie es la aparición de valores negativos de varianza dentro de la región de experimentación.

Por lo tanto “Q” y “L” representan a superficies que difieren notablemente en la zona de experimentación tanto en forma como en valores previstos.

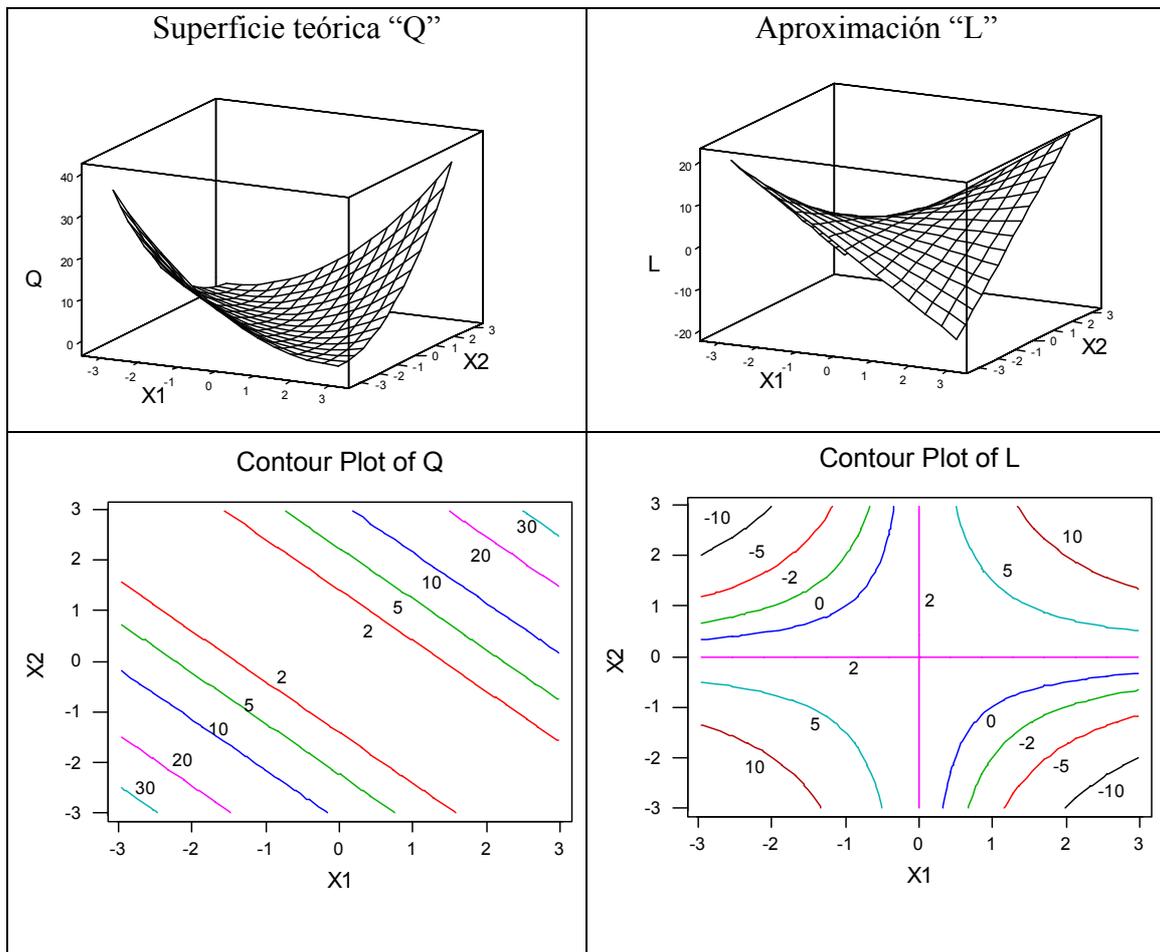


Figura 4.4. Representación de $V_Z(Y)$ con $Y = 0 Z + 1 X_1 Z + 1 X_2 Z$.

(Q: superficie teórica; L: aproximación por un polinomio sin términos cuadráticos puros).

$$\underline{Y = 2 Z + 5 X_1 Z + 1 X_2 Z}$$

Para este caso particular en el que en el modelo para Y el efecto principal del factor ruido es de magnitud relativamente “moderada”, dominando el efecto de la interacción del factor ruido con uno de los factores de control, puede verse cómo el modelo teórico

“Q” adopta la forma de una “teja inclinada” con una zona de mínima varianza determinada por las condiciones X_{1Q} y X_{2Q} que cumplen:

$$X_{2Q} + 5 X_{1Q} + 2 = 0$$

Esta recta pasa por el interior de la zona de experimentación representada por $X_1 = \pm 1$ y $X_2 = \pm 1$.

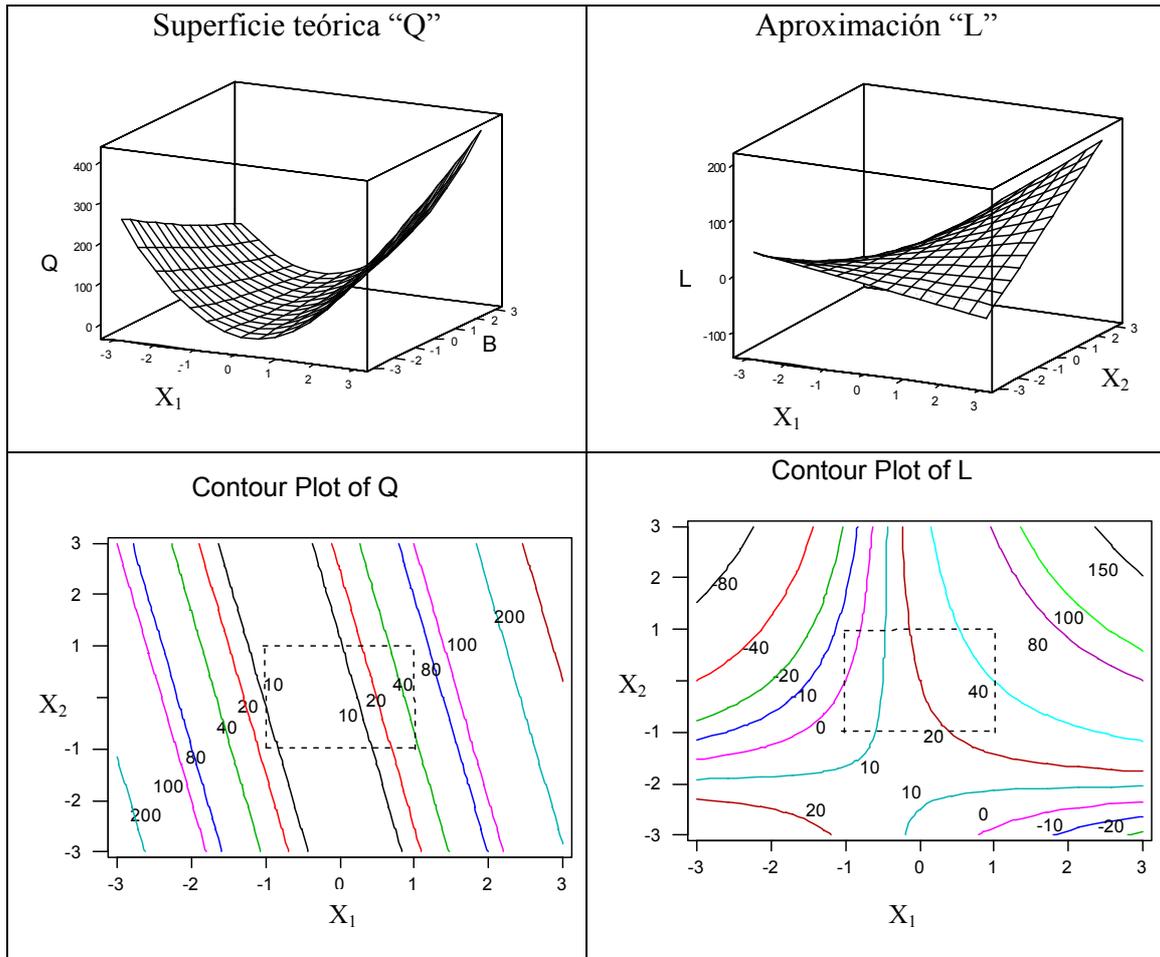


Figura 4.5. Representación de $V_Z(Y)$ con $Y = 2 Z + 5 X_1 Z + 1 X_2 Z$

(Q: superficie teórica; L: aproximación por un polinomio sin términos cuadráticos puros).

La aproximación “L” es muy diferente a la superficie teórica “Q” tanto en forma, es una “silla de montar”, como en valores observados. En la practica se interpretaría erróneamente la zona de mínima varianza constante tomándola como una dirección a seguir, alejándose del centro del diseño, de cara a minimizar la variabilidad. Hay que notar que el hecho de que aparezcan valores negativos en las estimaciones de variabilidad puede alertar al experimentador de que la aproximación no es adecuada.

Por lo tanto podemos decir que la interpretación que se hace de $V_Z(Y)$ a partir de “L” es incorrecta incluso dentro de la zona de experimentación.

$$Y = 0Z + 5X_1Z + 1X_2Z$$

Si en el modelo para Y el efecto principal del factor ruido se anula, como en un ejemplo anteriormente visto, y se aumenta la magnitud de al menos una de las interacciones del factor ruido con los factores de control, quedan más patentes las diferencias entre “Q” y “L” tal y como se puede ver en la Figura 4.6.

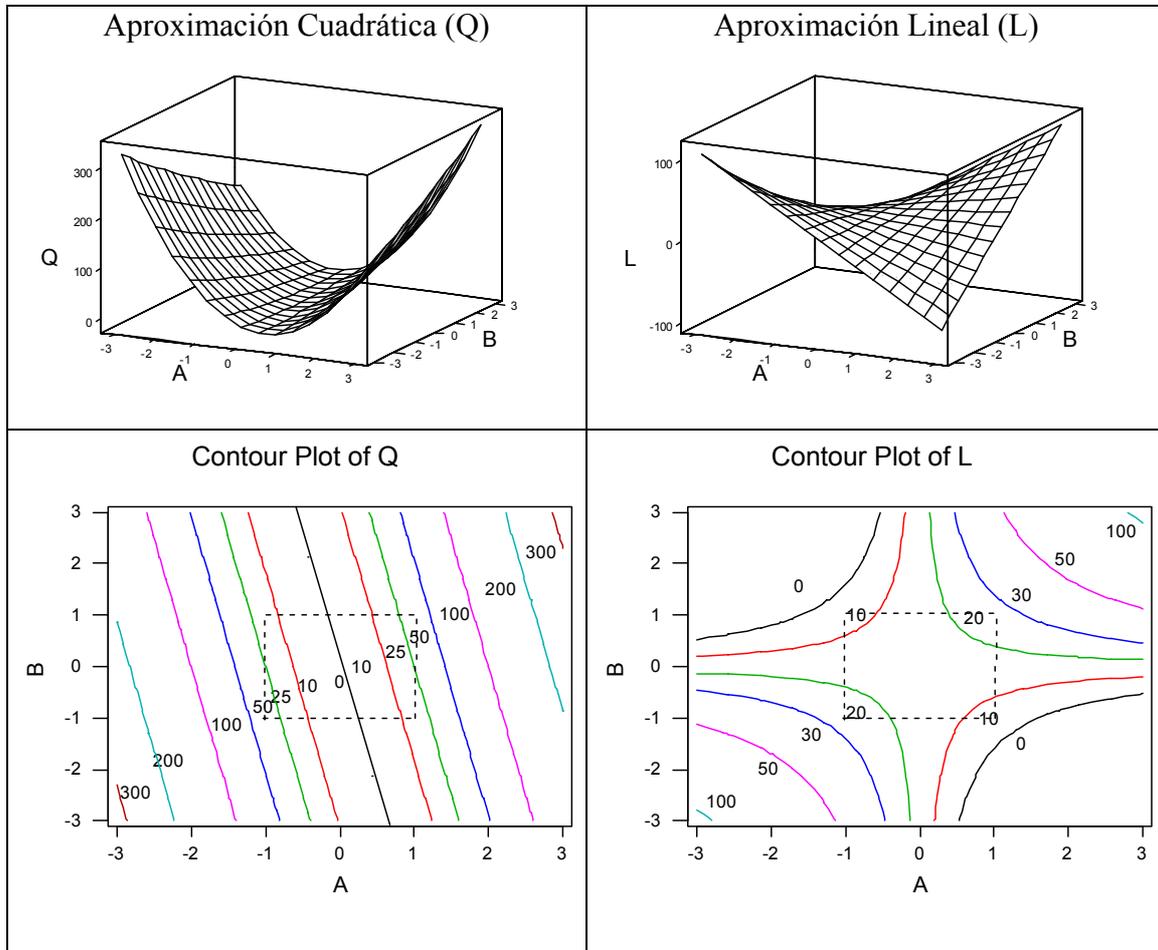


Figura 4.6. Representación de $V_Z(Y)$ con $Y = 0Z + 5X_1Z + 1X_2Z$.

(Q: superficie teórica; L: aproximación por un polinomio sin términos cuadráticos puros)

Comparando estos gráficos con los obtenidos en la Figura 4.4 se puede observar que en este caso la superficie teórica “Q” es similar a la obtenida en la Figura 4.4 pero girada respecto al centro y con un crecimiento más rápido en la dirección de X_1 . La zona de mínima varianza teórica también ha girado respecto al centro y “L” sigue sin identificarla como tal, tratándola una vez más como una dirección a tomar para alejarse del centro si se quiere reducir la variabilidad.

Esta vez “L” no toma valores negativos cerca de la zona determinada por $X_1 = \pm 1$ y $X_2 = \pm 1$ por lo que no tenemos una “pista” que nos ayude a detectar esta falta de ajuste.

4.3 Conclusiones

Partiendo de un modelo lineal para la respuesta Y dependiendo de un único factor ruido Z que interacciona con factores de control, se ha realizado un estudio de la superficie $V_Z(Y)$ asociada a la variabilidad transmitida por este factor ruido a la respuesta en la región determinada por los factores de control con los que interacciona. Se han comparado la superficie teórica denominada “Q” con la aproximación “L” que se obtiene por métricas resumen a partir de los resultados de un diseño factorial a dos niveles. Sin pérdida de generalidad, el estudio se ha llevado a cabo para $k = 2$ factores de control obteniéndose las siguientes conclusiones.

La superficie teórica “Q” es cuadrática mientras que el modelo “L” no posee términos cuadráticos puros.

El error de estimación que se comete al aproximar la superficie teórica “Q” por la superficie “L” no depende del coeficiente θ asociado al efecto del factor ruido en el modelo de Y ,

$$Q - L = \delta_1^2 (X_1^2 - 1) + \delta_2^2 (X_2^2 - 1)$$

- Cuanto mayor sea la magnitud de las interacciones entre los factores de control y el factor ruido, mayor será el error de estimación;
- La superficie “L” siempre coincide con la teórica en las condiciones experimentales y por lo tanto estima adecuadamente la variabilidad en esos puntos. Sin embargo, “L” superestima la variabilidad en la región interior a la experimentación (la que interesa normalmente al experimentador) e subestima en el resto.

La distancia de la zona de mínima varianza al centro del diseño si que depende de θ , del efecto del factor ruido en la respuesta

$$d((X_{1Q}, X_{2Q}), (0,0)) = \frac{\theta}{\sqrt{\delta_1^2 + \delta_2^2}}$$

Cuando la zona de mínima varianza pasa por la región experimental, la interpretación que se tiene de la forma de $V_Z(Y)$ a partir de “L” no es adecuada. Una consecuencia de esta errónea interpretación es que puede inducir al experimentador a alejarse de la región de interés.

El experimentador puede sospechar de la falta de ajuste del modelo “L” a $V_Z(Y)$ si la superficie “L” toma la forma de una silla de montar, b_{ij} significativo, o por la presencia de valores negativos en “L”.