

00384



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Facultad de Ciencias

"REGIONES DE PREDICTIBILIDAD EN SUPERFICIES DE RESPUESTA".

T E S I S
Que para obtener el grado de
DOCTORA EN CIENCIAS
(MATEMATICAS)
P r e s e n t a
BLANCA ROSA PEREZ SALVADOR

México D. F.

1991.

TESIS CON VALOR DE ORIGEN



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# REGRESOS DE PARETO Y MODELOS DE SUPERFICIES DE RESPUESTA

## Tesis doctoral

### Resumen

Ramona Rosa Pérez Salvador.

El tema de superficies de respuesta supone la existencia de una función continua en  $n$  factores estandarizados,  $Y = f(X_1, \dots, X_n)$ , que define el resultado de un proceso (que puede ser químico, industrial, etc.); el propósito del estudio es estimar el óptimo de  $Y$  suponiendo que la información es de la forma:

$$y = a + (x - \bar{x})^T \beta + \epsilon \quad \text{donde } \epsilon \sim N(0, \sigma^2) \text{ independientes.}$$

Box y Wilson (1951) fueron los primeros en formalizar el problema, a partir de ellos, son varios los autores que han estudiado el tema. Una referencia muy detallada de estos trabajos fue hecha por Myers, Khuri y Carter (1988).

El trabajo se conforma de 3 capítulos y dos apéndices.

En el primero se hace una recapitulación de los aspectos relevantes escritos sobre el tema, con la finalidad de ubicar al lector. Se enlistan una serie de algoritmos para estimar funciones de primer y segundo grado, así como los diferentes métodos conocidos para estimar el óptimo.

En el segundo capítulo se proponen 7 nuevos métodos para estimar el óptimo, basados en el método de ascenso por pendiente máxima, con la finalidad de mejorar las estimaciones del óptimo tanto en la precisión de los estimadores, como en el tamaño de las muestras requeridas.

En el tercer capítulo se somete la bondad de los 7 métodos propuestos, utilizando simulaciones de Monte Carlo, ANOVA y pruebas de una serie de coviadas. Se confrontan los 7 métodos propuestos y 2 ya conocidos usados como testigos. Los criterios de comparación son: la desviación media de  $\hat{X}_{opt}$  y la desviación media de  $\hat{Y}_{opt}$ . La

descripción mediante la "relativización" y el "ajuste" de la muestra. De esta confrontación destaca, entre todos los criterios, uno de los métodos propuestos el cual mejora la precisión de los estimadores.

En el cuarto capítulo se discute la construcción de las regiones de confianza para  $s_{11}$  y para  $\frac{s_{11}}{s_{22}}$ . Se propone una región de confianza para  $s_{11}$  de medida nominal. Se encuentran las condiciones para el diseño y para el método de búsqueda que otorgan un nivel de confianza de menor medida. Estas justifica parcialmente, de manera teórica, los resultados de la simulación.

El quinto capítulo es de conclusiones, y se proponen algunas acciones a seguir en trabajos futuros.

En el primer apéndice se presentan los diagramas de flujo de los métodos confrontados en la simulación.

En el segundo apéndice se presentan en diferentes tablas los resultados de la simulación.

#### Bibliografía señalada en este resumen.

- Box G. E. P., and Wilson K. R. (1951) "On the Experimental Attainment of Optimum Conditions", *Journal of the Royal Statistical Society, Ser. B*, 13, p. 1-45.
- Myers R. H., Khuri A. I. y Carter J. W. H. (1988) "Response Surface Methodology: 1966-1988", *Technometrics*, Vol 31, No. 2, p. 107-158.

Visto y firmado:



Dr. Federico Jorge O'Felly Fogno,  
Director de Tesis.

ABSTRACT

Blanca Rosa Perez-Sanchez

The subject of response surface analysis assumes there exist a continuous function of  $m$  standardized factors,  $y = f(x)$ , which defines the "yield" of a process. It can be chemical, industrial, and so on. The researcher is engaged in the exercise of finding conditions that give rise to an optimum of  $y$  when the information is from the model

$$y = \alpha_0 + \sum_{i=1}^m \alpha_i x_i \quad \text{where } \alpha_0, \alpha_1, \dots, \alpha_m \text{ are independent.}$$

Box and Wilson (1951) were the first who defined the conditions that give rise to an optimal response. After them, many researchers have studied the subject. An excellent relation about these works was written by Myers, Khuri and Carter (1988).

This work consist of five chapters and two appendices.

In Chapter 1 a review is done of the outstanding points on this subject. Here some well known designs for fitting first and second order models are mentioned, as well as various methods to estimate the optimum.

In chapter 2 seven new methods are proposed to estimate the optimum seeking the improvements on both the precision of the estimators and the sample size. These methods are a modification of steepest descent method.

In chapter 3 the relative use of the seven proposed methods using Monte Carlo simulations are obtained. They are also compared with two known methods ("controls"). The criteria of comparison are: mean deviation of  $\hat{\alpha}_i$ , mean deviation of  $\hat{\alpha}_i^2$ , mean deviation of  $\hat{\alpha}_{ij}$ .

standardized by  $\sigma$  and  $\mu$  and the standard deviation  $\sigma$  is assumed to be known. The study aims at improving the precision of estimates and the sample size.

In chapter 4 construction of confidence regions for  $S_{\mu}$  is done. Conditions are given for the adjoined confidence region. This justifies theoretically (although partially) the outcome of the simulation study.

In chapter 5 some actions for future works are proposed and some conclusions are given.

In appendix 1 flow charts of the compared methods are presented.

Appendix 2 the outcomes of the simulation study are presented.

#### Bibliography:

Box G. E. P. and Wilson K. B. (1951) "On the Experimental Attainment of Optimum Conditions", *Journal of the Royal Statistical Society, Ser. B*, 13, p. 1-45.

Myers R. H., Khuri A. I. and Carter, Jr. W. H. (1988) "Response Surface Methodology: 1956-1988", *Technometrics*, Vol 31, No. 2, p. 137-158.

Visto, Buenos



Dr. Federico Jorge O'Keilly Togno.  
Director de tesis.

## O. INTRODUCCIÓN.

En la industria, en la agricultura, en la administración y en muchas otras áreas, es frecuente encontrar situaciones en las que se requiere conocer la combinación de los niveles de un conjunto de factores que optimiza un resultado específico.

No siempre se necesita recurrir a la estadística para tomar una decisión; sin embargo, existen casos en los que por su misma complejidad, es imposible hacerlo sin un análisis estadístico, largo y minucioso, para obtener la combinación óptima de los factores. Por ejemplo: nadie, o casi nadie, pensaría en obtener una aleación para recubrir la cabeza del próximo cohete que será lanzado al espacio, sin antes haber realizado una serie de experimentos, tendientes a encontrar la combinación de materiales que proporcionan una mayor resistencia a las presiones a que estará expuesto.

A la serie de acciones (desde la elección del diseño experimental), para encontrar la combinación de factores donde se alcanza un máximo (o un mínimo, en su caso), se le conoce como Metodología de Superficies de Respuesta.

Por su misma naturaleza, la metodología de superficies de respuesta tiene múltiples aplicaciones de entre las cuales se presentan los siguientes ejemplos:

**EN LA PRODUCCIÓN DE CERÁMICA:** En esta rama industrial resulta de interés conocer la proporción de los ingredientes, la temperatura y el tiempo de cocción necesario con los que se obtienen piezas con mayor dureza, brillantez o suavidad al tacto.

EN LA INDUSTRIA FARMACÉUTICA. En este campo hay múltiples procesos en los que la metodología de superficies de respuesta tiene aplicación, por ejemplo: en la elaboración de un analgésico es necesario conocer que proporción de las diferentes sustancias produce un efecto calmante más efectivo, o bien, minimiza los efectos colaterales no deseados.

EN LA INDUSTRIA AUTOMOVILÍSTICA: Aquí resulta de interés conocer, por ejemplo, el tiempo de elaboración necesario y la proporción en que se deben mezclar los materiales para obtener neumáticos con mayor resistencia al desgaste.

EN LA FRUTICULTURA: En el tratamiento para la comercialización y conservación de la fruta es importante conocer el tiempo de corte, la cantidad de cera para recubrir la superficie y la temperatura a la que debe mantenerse una determinada fruta para que el grado de maduración se obtenga en un tiempo máximo.

Con estos ejemplos se ilustra la importancia que en la práctica tiene la Metodología de Superficies de Respuesta.

El tema de superficies de respuesta supone la existencia de una función continua,  $\eta : \mathbb{R}^m \rightarrow \mathbb{R}$ , que define el resultado de un proceso (que puede ser químico, industrial, etc.) en función de  $m$  factores estandarizados; el propósito del estudio, es obtener el óptimo de  $\eta$  cuando la información observable que se tiene de ella, está alterada por un término aleatorio de acuerdo al modelo:

$$y_i = \eta(x_i) + \varepsilon_i; \text{ donde } \varepsilon_i \sim N(0, \sigma^2) \text{ independientes.}$$

El primer trabajo publicado que tuvo como finalidad encontrar la solución práctica al problema de optimizar una función de  $m$  factores, cuyas observaciones están alteradas por el azar, fue el de Box y Wilson (1951). Ellos establecieron el método de ascenso por pendiente máxima para encontrar un estimador de la combinación de factores óptima. (Myers, Khuri y Carter, 1988). Posteriormente, otros autores abordaron el tema, entre ellos se encuentra un libro editado por Davies (1954) donde se exploran las propiedades de los diseños para estimar funciones de primer y segundo

## INDICE

	PAG.
0. INTRODUCCIÓN.....	1
1. METODOLOGÍA DE SUPERFICIES DE RESPUESTA.....	5
1.1 UNA REVISIÓN GENERAL.....	5
1.2 EL DISEÑO.....	8
1.3 MÉTODOS DE BÚSQUEDA DEL ÓPTIMO.....	14
2. VARIANTES QUE SE PROPONEN AL MÉTODO DE ASCENSO POR PENDIENTE MÁXIMA.....	23
2.1 PRIMERA VARIANTE.....	24
2.2 SEGUNDA VARIANTE.....	27
2.3 TERCERA VARIANTE.....	28
2.4 CUARTA, QUINTA, SEXTA Y SÉPTIMA VARIANTE.....	31
2.5 RESUMEN.....	32
3. LA SIMULACIÓN.....	33
3.1 GENERALIDADES.....	33
3.2 PROCEDIMIENTO.....	35
3.3 ANÁLISIS DE LOS RESULTADOS DE LA SIMULACIÓN.....	41
4. UNA REGIÓN DE CONFIANZA PARA $x_{op}$ Y PARA $\pi_{op}$ .....	43
4.1 INTERVALOS DE CONFIANZA EN CASO DE TENER UN ÚNICO FACTOR.....	44
4.2 REGIÓN DE CONFIANZA PARA $x_{op}$ Y PARA $\pi_{op}$ . CASO M FACTORES.....	75
5. CONCLUSIONES Y COMENTARIOS FINALES.....	103
BIBLIOGRAFÍA.....	107
APENDICE 1. DIAGRAMAS DE BLOQUE DE LOS MÉTODOS ANALIZADOS.....	109
APENDICE 2. ALGUNOS RESULTADOS DE LA SIMULACIÓN.....	117

grado. Box y Hunter (1954) establecen una región de confianza para la combinación de factores óptima basada en el teorema de Fieller. Box y Hunter (1957) analizan el comportamiento de la varianza del predictor de la función de respuesta,  $\hat{\eta}(x)$ , con relación a los diseños y definen el concepto de diseños rotables. Myers y Khuri (1979) proporcionan un nuevo método de ascenso por pendiente máxima. Carter, Chinchilli, Campbell y Wampler (1984); y Carter, Chinchilli, Myers y Campbell (1986) proporcionan una nueva región de confianza. Notz (1988) propuso los diseños D-eficientes, lo que originó el concepto de diseños D-óptimos, abordado por Atkinson y Donev (1989) entre otros. Una excelente relación, muy detallada, de los trabajos que se han hecho sobre superficies de respuesta fue hecha por Myers, Khuri y Carter (1988)

En este trabajo se abordan 3 de los aspectos relevantes de las superficies de respuesta: el método de estimación del óptimo, las regiones de confianza del óptimo, y el diseño. El trabajo se conforma de 5 capítulos y dos apéndices.

En el primero se hace una recapitulación de los aspectos relevantes escritos sobre el tema, con la finalidad de ubicar al lector. Se enlistan, sin ser exhaustivos, una serie de diseños para estimar funciones de primer y segundo grado, así como los diferentes métodos conocidos para estimar el óptimo.

En el segundo capítulo se proponen 7 nuevos métodos para estimar el óptimo, todos ellos basados en el método de ascenso por pendiente máxima. Se proponen y se justifican una serie de modificaciones y adiciones al método de ascenso por pendiente máxima, con la finalidad de mejorar las estimaciones del óptimo, tanto en la precisión de los estimadores, como en el tamaño de las muestras requeridas.

En el tercer capítulo se sondea la bondad de los 7 métodos propuestos, utilizando simulaciones de Monte Carlo. Así, a través de una serie de corridas, se confrontan los 7 métodos propuestos, junto con otros dos ya conocidos (usados como testigos).

Los criterios de confrontación utilizados son: la desviación media de  $\hat{x}_{op}$ ; la desviación media de  $\hat{\eta}_{op}$ ; la desviación media de  $\hat{\eta}_{op}$  relativa a  $\sigma$ ; y el tamaño de la muestra.

De esta confrontación destaca como el mejor, ante todos los

critérios, uno de los métodos propuestos. Este método mejora la precisión de los estimadores.

Los estimadores obtenidos con los otros métodos tienen una desviación media empírica hasta  $10^5$  más grande que la de los estimadores obtenidos con este método.

En el cuarto capítulo se discute la construcción de las regiones de confianza para  $x_{op}$ , y para  $\eta_{op}$ . Se propone la construcción de una región de confianza para  $x_{op}$  de medida mínima; se encuentran las condiciones para el diseño y para el método de búsqueda que originan una región de confianza de menor medida. Esto justifica parcialmente, de manera teórica, los resultados de la simulación.

El quinto capítulo es de conclusiones, y se proponen algunas acciones a seguir en trabajos futuros.

En el primer apéndice se presentan los diagramas de flujo de los métodos confrontados en la simulación.

En el segundo apéndice se presentan en diferentes tablas algunos de los resultados de la simulación.

Las aportaciones del trabajo son:

- i) Un nuevo método de búsqueda del óptimo derivado del método de ascenso por pendiente máxima el cual, de acuerdo a los resultados de una serie de simulaciones, proporciona estimadores del óptimo con menor varianza y menor sesgo.
- ii) Una región de confianza para la combinación de factores óptima. Esta región tiene mayor probabilidad de estar acotada que la región de confianza reportada en la literatura.
- iii) Una región de confianza para el resultado óptimo.
- iv) Un diseño con el cual se construyen, fácilmente, las regiones de confianza anteriormente mencionadas.

## 1. METODOLOGIA DE SUPERFICIES DE RESPUESTA.

### 1.1 UNA REVISIÓN GENERAL.

La metodología de superficies de respuesta tiene como fin determinar las condiciones que proporcionan un resultado óptimo en un proceso específico. Esta técnica, como ya se dijo, tiene múltiples aplicaciones, por ejemplo: en los procesos químicos donde se pretende maximizar el rendimiento, o la pureza de un producto; o bien, minimizar el costo de producción variando las condiciones iniciales del proceso específico (las condiciones son los niveles de los diferentes factores involucrados).

Cuando los factores involucrados son variables numéricas, como por ejemplo: la temperatura, el tiempo, la cantidad de una substancia, etc. es natural considerar que el resultado del proceso está en función de dichos factores.

Esto es:

$$\eta^* = \eta^*(x_1, x_2, \dots, x_m). \quad (1.1)$$

donde  $\eta^*$  es la respuesta real de interés, medida numéricamente.

Esta función se conoce como función de respuesta y su gráfica como superficie de respuesta.

Las observaciones de la función  $\eta^*$  son de la forma:

$$Y_1 = \eta^*(x_{11}, x_{12}, \dots, x_{1m}) + \varepsilon_1 \quad (1.2)$$

Con  $\varepsilon_1 \sim N(0, \sigma^2)$  independientes.

El conocimiento que se tiene de la respuesta  $\eta^*$  determina la manera de estimar el punto óptimo así como la precisión de la estimación.

Si se conoce completamente la forma de  $\eta^*$  es posible encontrar el o los puntos donde la superficie de respuesta tiene un óptimo; esto es un máximo o un mínimo. En esta situación no hay problema de tipo estadístico y no se requiere de la información de una muestra.

Cuando sólo se conoce la familia paramétrica de funciones a la que pertenece  $\eta^*$ , es posible estimar la función de respuesta específica con base en los datos experimentales  $\{Y_1, Y_2, \dots, Y_n\}$  y  $\{x_1, x_2, \dots, x_n\}$ , el diseño utilizado. La precisión de la estimación depende, como es de esperarse, de la magnitud de  $\sigma$ , la desviación estándar del error de las observaciones y del diseño. El óptimo de la respuesta  $\eta^*$  se estima con el óptimo de la respuesta estimada  $\hat{\eta}$ .

En la mayoría de los casos, sin embargo, no se conoce la forma paramétrica de  $\eta^*$ ; pero suponiendo que esta función es continua, se puede aproximar su comportamiento local mediante una función paramétrica, principalmente polinomios de primer y segundo grado (de primer grado si el punto crítico está lejos de los puntos donde se toma la muestra y de segundo grado si está cerca); posteriormente, se estiman los parámetros de la función elegida. Estas estimaciones sólo tienen validez en una región reducida alrededor de los puntos del diseño; dado que, como ya se dijo, la aproximación es local.

Cuando se estiman los parámetros que definen una función, se

---

‡  $x_{ij}$  es resultado de la transformación  $x_{ij} = (\xi_{ij} - \bar{\xi}_j)/c$  cuya finalidad es escalar y centrar a  $\xi_{ij}$ , el  $i$ -ésimo nivel observado del factor  $j$  ( $c$  es una constante y  $\bar{\xi}_j = \sum \xi_{ij}/n$ ; por lo tanto para cada  $j$ ,  $\sum x_{ij} = 0$ ). Esta transformación tiene como fin simplificar los cálculos numéricos.

puede considerar que existen tres funciones involucradas:

- i) La función  $\eta^*$  que describe la respuesta real del proceso.
- ii) La función  $\eta$ , seleccionada por el investigador para aproximar el comportamiento de  $\eta^*$  en el interior de una región que contiene los puntos del diseño. ( $\eta = \eta^*$  cuando se conoce la familia paramétrica a la que pertenece  $\eta^*$ ).
- iii) Y por último, la función  $\hat{\eta}$ , estimador de la función  $\eta$ . (o de  $\eta^*$  si es conocida su forma paramétrica).

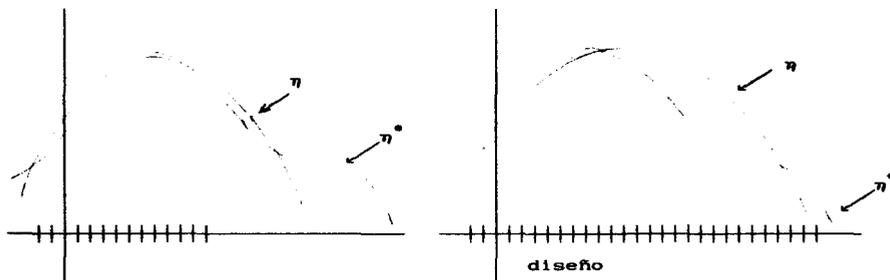


Figura 4.1. El cambio de diseño puede hacer que la función  $\eta$  sea diferente. (caso  $\eta \neq \eta^*$ )

En estas condiciones la función  $\hat{\eta}$ , como estimador de  $\eta^*$ , presenta dos fuentes de variación: una que proviene de estimar a la función  $\eta$ , y la otra que proviene de que tan bien  $\eta$  se aproxima a  $\eta^*$ . Por esta razón, si  $\eta$  se parece muy poco a  $\eta^*$ , también  $\hat{\eta}$  va a diferir de  $\eta^*$ .

Esto justifica la necesidad de tener un criterio para determinar, en función del diseño, la región donde la predictibilidad de  $\eta^*$  mediante el uso de  $\hat{\eta}$  es aceptable. ( $\hat{\eta}$  predice a  $\eta^*$  usando la función intermedia  $\eta$ ).

De principio se considera que  $\eta$  describe razonablemente bien a  $\eta^*$  en una región alrededor de los puntos del diseño.

## 1.2 EL DISEÑO.

La elección del diseño, i. e. la combinación de factores en donde se toma la muestra, va a responder a los diferentes intereses del investigador, como por ejemplo: la economía, (tanto por el número de observaciones, como por el número de cálculos implicados) y la precisión de la estimación (i. e.. que  $V(\eta)$  sea pequeña). (Notz, 1988)

Entonces, es conveniente contar con criterios para elegir aquellos diseños que sean "económicos" y "precisos". Con la finalidad de establecer dichos criterios se discuten las características del modelo lineal para ajustar un polinomio de grado  $r$  a los datos ( $r = 1, 2, \dots$ ); con  $m$  factores en estudio.

El modelo lineal general es:

$$Y = W\beta + \epsilon \quad (1.3)$$

donde

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}; \quad \beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix}; \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

y

$$W = \begin{bmatrix} 1 & x_{11} \dots x_{1m} & x_{11}^2 & x_{11}x_{12} & x_{11}x_{13} \dots & x_{11}^2 & x_{11}^3 & \dots & x_{11}^r \\ \vdots & \vdots \\ 1 & x_{n1} \dots x_{nm} & x_{n1}^2 & x_{n1}x_{n2} & x_{n1}x_{n3} \dots & x_{n1}^2 & x_{n1}^3 & \dots & x_{n1}^r \end{bmatrix}$$

la dimensión del vector  $\beta$  es  $k+1 = C_r^{m+r}$ , y  $\epsilon$  es el vector de errores.

El estimador de mínimos cuadrados de  $\beta$  satisface la ecuación matricial  $W^T W \beta = W^T Y$  conocida como "sistema de ecuaciones normales", y

de esa manera el estimador del polinomio evaluado en  $(x_1, x_2, \dots, x_m)$  es

$$\hat{\eta}(\mathbf{x}) = \mathbf{w}^T \hat{\beta};$$

donde  $\mathbf{w}^T = (1, x_1, x_2, \dots, x_m, x_1^2, x_1 x_2, \dots, x_m^2, x_1^3, \dots, x_m^r)$ .

El vector  $\hat{\beta}$ , que siempre existe, es único si y sólo si la matriz  $\mathbf{W}^T \mathbf{W}$  es invertible; ésto es equivalente a que el rango de  $\mathbf{W}$  sea igual a  $k+1$ , por lo que se requiere de al menos  $k+1$  puntos para estimar a  $\eta(\mathbf{x})$  de manera única.

De aquí se sigue que el diseño más económico es: por su tamaño, el que tiene  $k+1$  puntos con  $\mathbf{W}$  de rango completo; y, por los cálculos numéricos involucrados, el que proporciona una matriz  $\mathbf{W}^T \mathbf{W}$  diagonal o con mayor número de ceros lo que facilita su inversión.

Por otro lado, la magnitud de  $V(\hat{\eta}) = \mathbf{w}^T (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{w} \sigma^2$  determina la precisión de las estimaciones, y  $V(\hat{\eta})$ , por regla general, es más pequeña cuando se tienen más datos. Por lo que claramente se ve que los conceptos de economía y precisión son opuestos y debe buscarse un equilibrio entre estos dos conceptos al elegir el diseño.

### 1.2.1 DISEÑOS PARA ESTIMAR UN POLINOMIO DE PRIMER GRADO.

En esta parte, se enumeran únicamente aquellos diseños que tradicionalmente han sido utilizados para estimar el modelo lineal general (1.3) cuando  $r = 1$ ; o sea, cuando la superficie descrita por  $\hat{\eta}$  es un hiperplano.

El modelo que describe el comportamiento de las observaciones, en este caso, es:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} + \varepsilon_i \quad (1.4)$$

el cual tiene  $m + 1$  parámetros  $\beta_i$  (además de  $\sigma^2$ , la varianza de las observaciones). Por lo tanto para estimar los parámetros de este modelo es necesario un diseño con al menos  $m + 1$  puntos, distribuidos de forma tal que la matriz  $\mathbf{W}$  sea no singular. Si además, se quiere estimar a  $\sigma^2$  es necesario, cuando menos,  $m + 2$  puntos.

### 1.2.1.1 Diseño simplex.

Este es el diseño más pequeño que reporta la literatura; consta de  $m+1$  puntos.  $x_1 = (x_{11}, x_{12}, \dots, x_{1m})$ , equidistantes al origen y equidistantes entre sí, de norma igual a uno. El diseño es insuficiente para estimar a  $\sigma^2$ , por ello es necesario agregar, al menos, un punto más y por simetría se sugiere que este punto se localice en el centro del diseño.

Por ejemplo: para  $m = 2$ , el diseño simplex está formado por los tres vértices de un triángulo equilátero con centro en el origen. Para  $m = 3$ , el diseño está formado por los vértices de un tetraedro regular con centro también en el origen.

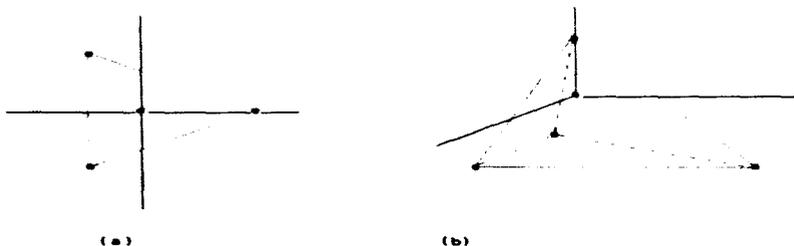


figura 1.1 (a) diseño simplex  $m=2$ . (b) diseño simplex  $m=3$ .

Con este diseño la matriz  $W^T W$  es diagonal; esto implica que los cálculos para obtener los estimadores de  $\beta$  y  $\eta$  son muy simples.

El diseño simplex es invariante a las rotaciones, ya que si se aplica una rotación a los puntos de un diseño simplex, el resultado es otro diseño simplex.

Por último, se puede ver fácilmente que para un diseño simplex con  $\alpha$  puntos en el centro,  $V(\hat{\eta})$  evaluada en el punto

$\mathbf{x} = (x_1, x_2, \dots, x_m)$  es igual a:

$$V(\hat{\eta}) = [1/(m+\alpha+1) + m \sum x_i^2 / (m+1)] \sigma^2. \quad (1.5)$$

### 1.2.1.2 Diseño factorial $2^m$ .

Este es el diseño más utilizado, posiblemente por lo fácil de construir, ya que cada factor presenta únicamente dos niveles, 1 y -1. La matriz  $\mathbf{W}^T \mathbf{W}$  es diagonal, y la varianza de la respuesta estimada, evaluada en el punto  $\mathbf{x}$  es:

$$V(\hat{\eta}) = \sigma^2 (1 + \sum x_i^2) / 2^m \quad (1.6)$$

La división entre  $2^m$  implica que  $V(\hat{\eta})$  es pequeña, esto es: los estimadores son precisos; sin embargo, al tener  $2^m$  puntos el diseño resulta caro para  $m$  "grande".

### 1.2.1.3 Diseño factorial fraccionado $2^{m-k}$ .

El diseño factorial fraccionado  $2^{m-k}$  tiene la estructura básica del factorial, pero con sólo  $2^{m-k}$  puntos, por lo que resulta más económico que el factorial; con este diseño la matriz  $\mathbf{W}^T \mathbf{W}$  también es diagonal, y  $V(\hat{\eta})$  es:

$$V(\hat{\eta}) = \sigma^2 (1 + \sum x_i^2) / 2^{m-k} \quad (1.7)$$

### 1.2.2 DISEÑOS PARA ESTIMAR POLINOMIOS DE SEGUNDO GRADO.

Después de revisar brevemente los diseños para ajustar un polinomio de primer grado, se enumeran los diseños utilizados para el ajuste de polinomios de segundo grado. Los polinomios de segundo grado tienen más parámetros, por lo que se necesita un diseños con más puntos para estimarlos. Es común usar, en este caso, los diseños utilizados para el ajuste de un polinomio de primer grado, agregándole convenientemente algunos puntos más para tener el número necesario de observaciones.

El modelo lineal general con  $r = 2$  es:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_m x_{im} + \beta_{11} x_{i1}^2 + \beta_{12} x_{i1} x_{i2} + \dots + \beta_{mm} x_{im}^2 + \epsilon_i \quad (1.8)$$

el cual tiene: un parámetro independiente,  $\beta_0$ ;  $m$  parámetros en los términos de primer grado  $\beta_i$  ( $i = 1, 2, \dots, m$ );  $m$  parámetros en los términos cuadráticos  $\beta_{ii}$  ( $i = 1, 2, \dots, m$ ); y por último,  $C_2^m$  parámetros en los términos cruzados,  $\beta_{ij}$  ( $1 \leq i < j \leq m$ ). En total hay

$$1 + m + m + C_2^m = C_0^m + C_1^m + C_1^m + C_2^m =$$

$$C_1^{m+1} + C_2^{m+1} = C_2^{m+2}$$

parámetros. Entonces para estimar un polinomio de segundo grado se necesitan al menos  $C_2^{m+2}$  puntos (uno más por lo menos si se desea estimar también a  $\sigma^2$ ) y, por supuesto, que  $W$  sea de rango completo.

#### 1.2.2.1 Diseño simplex compuesto.

Este diseño se compone de los  $m+1$  puntos de un diseño simplex,  $p_i$   $i = 1, 2, \dots, m, m+1$ ; más  $C_2^{m+1}$  puntos de la forma  $k(p_i + p_j)$ , donde  $i \neq j$  y  $k$  es una constante de escalamiento. Por construcción, este diseño tiene el mínimo número de puntos, necesarios para estimar al vector  $\beta$ ; sin embargo, la matriz  $W$  no es de rango completo. Para completar el rango de  $W$  es suficiente agregar al diseño un punto más en el centro.

diseño simplex compuesto.  $m = 2$ .

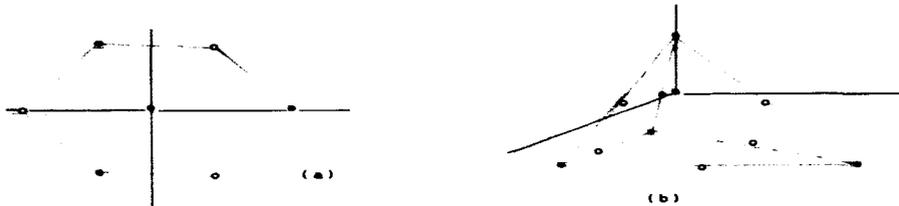


figura 1.2 Diseño simplex compuesto ( a) en dos factores, b) en tres factores) los puntos que se agregan al diseño simplex son los marcados con "o".

#### 1.2.2.2 Diseño compuesto.

Es el diseño más utilizado; está formado por los puntos de un diseño factorial  $2^m$ , o de un factorial fraccionado  $2^{m-k}$ , más los puntos de la forma  $\pm \alpha e_i$ , donde  $\{e_1, e_2, \dots, e_m\}$  es la base canónica de  $\mathbb{R}^m$  y  $\alpha$  es una constante de escalamiento. Este diseño tiene  $2^{m-k} + 2m$  puntos.

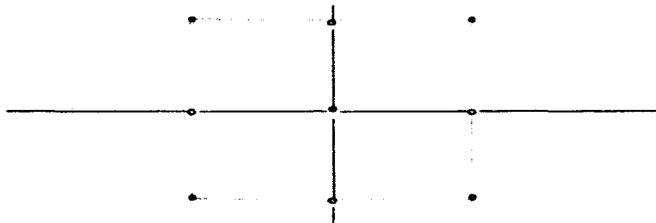


figura 1.3 a) Diseño compuesto en dos factores, con  $\alpha=1$ , los puntos que se agregan al diseño factorial están marcados con "o"

### 1.2.3 OTROS DISEÑOS MÁS GENERALES.

No es la finalidad de este trabajo hacer un listado exhaustivo de los diseños que se pueden encontrar en la literatura correspondiente al tema; sin embargo, es conveniente mencionar otros diseños frecuentemente mencionados, que se utilizan para estimar polinomios más generales.

#### 1.2.3.1 Diseño poligonal.

Los puntos de estos diseños son los vértices de poliedros regulares. Para  $r = 2$  los puntos de los diseños poligonales forman un polígono regular; los diseños simplex, factorial  $2^m$  y factorial fraccionado  $2^{m-k}$  caen en esta clasificación.

#### 1.2.3.2 Diseño rotable.

Los rotables son una clase particular de diseños y se utilizan para la estimación de un polinomio de cualquier grado. Con estos diseños  $V(\eta) = g(\#\#\#)$ , esto es  $V(\eta)$  sólo depende de la norma de  $\#$ . El diseño simplex, el factorial y el simplex compuesto con  $m = 2$  son rotables.

#### 1.2.3.3 Diseños D-óptimos.

Los D-óptimos son una clase de diseños con los que se maximiza el determinante de la matriz  $W^T W$ .

## 1.3 MÉTODOS DE BÚSQUEDA DEL ÓPTIMO.

Como ya se mencionó, en la mayoría de los casos se utilizan aproximaciones locales a la función de respuesta  $\eta^*$ . Por ello, para estimar el o los puntos donde  $\eta^*$  alcanza un óptimo relativo (o absoluto)

se necesita contar con un método que garantice un mínimo de confiabilidad en los resultados.

Por último, debe mencionarse que los métodos de búsqueda presentados solamente consideran el caso en que  $\eta^*$  presenta un máximo único y bien definido (o al menos, que se puede llegar a un máximo bien definido).

En esta parte, se revisarán los métodos más conocidos, los cuales son:

- El método aleatorio. (Cochran y Cox, 1960)
- El método del factor único. (Cochran y Cox, 1960)
- El método de un solo experimento. (Cochran y Cox, 1960)
- El método de ascenso por pendiente máxima. (Box y Wilson, 1951)

### 1.3.1 EL METODO ALEATORIO.

Este método consta de tres pasos

1°. Se eligen al azar  $n$  puntos  $\mathbf{x}_1^T = (x_{11}, x_{12}, \dots, x_{1m})$  cuyas coordenadas son variables aleatorias independientes con distribución uniforme:  $x_{1j} \sim U(a_j, b_j)$ ; así, estos puntos resultan ser variables aleatorias multivariadas, independientes, con distribución uniforme en el conjunto  $A = (a_1, b_1) \times (a_2, b_2) \times \dots \times (a_m, b_m)$  el cual se llama región de muestreo, o región de exploración.

2°. Se obtienen las observaciones  $Y_1 = \eta(\mathbf{x}_1) + \varepsilon_1$ , en cada uno de los puntos  $\mathbf{x}_1$ , antes seleccionados.

3°. Se escoge como  $\hat{\mathbf{x}}_{op}$  al punto  $\mathbf{x}_1$  que presente la mayor respuesta experimental.

Cuando  $\mathbf{x}_{op} \in A$ ,  $\hat{\mathbf{x}}_{op}$  resulta ser un estimador sesgado de  $\mathbf{x}_{op}$ , (esto es  $E(\hat{\mathbf{x}}_{op}) \neq \mathbf{x}_{op}$ ), el sesgo es mayor conforme  $\mathbf{x}_{op}$  se encuentre más lejos de la región de muestreo.

Cuando  $\mathbf{x}_{op} \in A$ , para cada número natural  $N$ , se puede

establecer una partición de la región de muestreo dada por los  $N^m$  hiper-rectángulos:

$$A(i_1, i_2, \dots, i_m) = (a_{1i_1}, a_{1(i_1+1)}) \times (a_{2i_2}, a_{2(i_2+1)}) \times \dots \times (a_{mi_i}, a_{m(i_m+1)}) \quad (1.9)$$

donde para cada  $k = 1, m$ ;  $0 \leq i_k \leq N-1$  y  $a_{ki_i} = a_k + i_k(b_k - a_k)/N$ .

y encontrar la probabilidad que alguno de los puntos  $\mathbf{x}_i$  esté en el mismo elemento de la partición que contiene a  $\mathbf{x}_{op}$ . (Como  $\mathbf{x}_{op} \in A$ , se sigue que  $\mathbf{x}_{op}$  pertenece a uno y sólo uno de los conjuntos de la partición)

Para calcular esta probabilidad recuerde que cada punto  $\mathbf{x}_i$  es un vector aleatorio distribuido uniformemente en  $A$ , y que cada conjunto  $A(i_1, i_2, \dots, i_m)$  tiene la misma medida de Lebesgue.

Así el número de observaciones que caen en la misma celdilla que  $\mathbf{x}_{op}$  es una variable aleatoria binomial con parámetros  $n$  y  $1/N^m$ , por lo que, la probabilidad de que al menos una observación esté en esta celdilla es igual a:

$$1 - (1 - 1/N^m)^n. \quad (1.10)$$

Un resultado semejante se obtiene cuando en lugar de usar un hiper-rectángulo que contiene a  $\mathbf{x}_{op}$ , se usa una hiper-esfera de probabilidad  $1/N^m$  con centro en  $\mathbf{x}_{op}$ .

Note que, independientemente del valor de esta probabilidad, un vector  $\mathbf{x}_i$  en la misma celdilla que  $\mathbf{x}_{op}$  no necesariamente presenta la mayor respuesta observada y, por lo tanto, no necesariamente es el punto elegido como estimador de  $\mathbf{x}_{op}$ .

### 1.3.2 EL MÉTODO DEL FACTOR ÚNICO.

Con este método se propone descomponer el proceso multifactorial, en  $m$  procesos sencillos de búsqueda individual del óptimo de cada coordenadas; convirtiéndose en un proceso de búsqueda

unidimensional.

El método consta de los siguientes pasos:

1°. Se ordenan los factores en orden de importancia, de acuerdo a su contribución en el resultado de la respuesta: así,  $x_1$  es el factor que más contribuye al resultado;  $x_2$  es el segundo en orden de contribución, etc. La ordenación se hace de acuerdo al criterio del investigador.

2°. Se elige un punto inicial,  $P_0 = (x_{10}, x_{20}, \dots, x_{m0})$ , ubicado donde, de acuerdo a la experiencia del investigador, es más probable que esté el óptimo.

3°. Se busca, sucesivamente, el óptimo de cada coordenada y se reemplaza en  $P_0$ .

Este último paso, se realiza de la siguiente manera:

i) Se obtiene al menos tres observaciones de la forma:

$$Y_i = \eta(x_{11}, x_{20}, \dots, x_{m0}) + \epsilon_i,$$

variando el nivel del primer factor y manteniendo constante el nivel de los otros factores.

ii) Con estas observaciones se estima la función cuadrática

$$\hat{y}_t = \hat{a}t^2 + \hat{b}t + \hat{c}, \quad (1.11)$$

la cual es un estimador de la tendencia de  $\eta$  en dirección de la primera coordenada; si  $\hat{a} < 0$ , la función en (1.11) alcanza el máximo en el punto  $x_{11} = -\hat{b}/2\hat{a}$ ; este valor sustituye la primer coordenada de  $P_0$ , resultando el nuevo punto  $(x_{11}, x_{20}, \dots, x_{m0})$ . (Cuando  $\hat{a} = 0$ , (1.11) es una recta y no presenta ningún máximo relativo; cuando  $\hat{a} > 0$ , (1.11) presenta un mínimo como único punto crítico).

iii) A partir de este punto, se toman al menos tres observaciones de la forma  $Y_i = \eta(x_{11}, x_{21}, \dots, x_{m0}) + \epsilon_i$ ; variando ahora, únicamente, el segundo factor; con esas observaciones se estima el máximo en dirección

de la segunda coordenada, y este valor sustituye a  $x_{20}$ . Así se continúa hasta tener el vector  $P_1 = (x_{11}, x_{21}, \dots, x_{n1})$ .

El proceso de búsqueda se detiene, cuando  $\|P_1 - P_{1+1}\|$  es menor a un valor preestablecido y  $\bar{x}_{op} = P_{1+1}$ . Si esto no ocurre, se inicia nuevamente el proceso de búsqueda con el punto inicial en  $P_{1+1}$ .

Este método no toma en cuenta las interacciones; es decir cuando no hay interacciones el método puede producir buenos resultados, de no ser así va a requerir de muchos pasos.

Note que con estos dos métodos no se determina la forma funcional de  $\eta^*$ . Si se desea estimarla es necesario hacerlo por otro camino.

### 1.3.3 EL MÉTODO DE UN SOLO EXPERIMENTO.

Con este método se considera que cerca del óptimo  $\eta^*$  se parece a un polinomio de segundo grado, por lo que, para buscar el punto óptimo de  $\eta^*$  es necesario ajustar a los datos un polinomio de este tipo.

El método consta de los dos pasos siguientes:

- i) Se elige un diseño para el ajuste de un polinomio de segundo grado y con las observaciones en estos puntos se estima dicho polinomio.
- ii) El óptimo estimado resulta ser el óptimo de la respuesta estimada.

Las estimaciones así obtenidas pueden ser muy malas si el diseño está alejado del óptimo real; porque  $V(\hat{\eta}(z))$  crece conforme  $z$  se encuentre más lejos del diseño.

### 1.3.4. EL MÉTODO DE ASCENSO POR PENDIENTE MÁXIMA.

Este método basa el proceso de búsqueda del óptimo en considerar que lejos de un punto crítico,  $\eta^*$  puede ser aproximada por una función

lineal; mientras que cerca del punto crítico  $\eta^*$  se puede aproximar mediante una función cuadrática.

El método consta de los siguientes pasos.

1°.- Se elige un diseño para el ajuste de un polinomio de primer grado (por ejemplo un diseño simplex o un factorial), con repeticiones en el centro; y se obtienen las observaciones en los puntos de este diseño.

2°.- Se estima el polinomio de primer grado.

3°.- Se hace la prueba de falta de ajuste, para determinar si el plano se ajusta bien a los datos, o si existe ya una evidente curvatura. Si hay evidencia de curvatura se pasa al punto 7. El error puro se estima con las observaciones en el centro del diseño.

4°.- Se efectúa la prueba de hipótesis  $H_0: \beta_i = 0$ , para toda  $i = 1, m$  contra  $H_A: \beta_i \neq 0$ , para alguna  $i = 1, m$ ; con el fin de determinar evidencia que los factores influyen en el resultado de la respuesta. Si no se rechaza  $H_0$  implica que se tienen evidencias que  $\eta$  se ajusta bien a un hiperplano horizontal y, por lo tanto, se está cerca de un punto crítico, en este caso se pasa al punto 7.

5°.- Se encuentra el vector de dirección de máxima pendiente en  $\hat{\eta}$ ,  $v = (v_1, v_2, \dots, v_m)^T$  esto se hace porque al considerar que  $\eta_{op}$  aún está lejos, se puede suponer que éste se localiza en la dirección de mayor pendiente de  $\eta^*$ , y la tendencia local de la respuesta  $\eta^*$  es estimada por el hiperplano  $\hat{\eta}$ .

El vector  $v$  se encuentra maximizando la función

$$\hat{\eta} = \hat{a}_0 + \hat{a}_1 x_1 + \dots + \hat{a}_m x_m \quad (1.11)$$

sujeta a la restricción  $\sum x_i^2 = 1$ .

6°.- Con un espaciamento constante determinado por  $d = 1/\max\{|v_i|\}$ , cantidad llamada longitud del "paso", y en dirección de  $v$  se toman, sucesivamente, nuevas observaciones,

$$Y_i = \eta(idv) + \varepsilon_i, \quad i = 1, \dots, m$$

hasta detectar un cambio que indique la existencia de un máximo de  $\eta^{\circ}$  en dirección de  $v$ ; esto es, si  $Y_i < Y_{i+1}$  se obtiene la siguiente observación  $Y_{i+2}$ ; si por el contrario,  $Y_i \geq Y_{i+1}$  se detiene el proceso y con centro en el punto  $idv$ , que fue donde se obtuvo la observación de máximo valor, se toma un nuevo diseño con el que se reinicia el proceso de búsqueda desde el paso 1.

Debe notarse que al reiniciar el proceso solamente se utilizan las observaciones del nuevo diseño que al centrarse y escalarse resulta ser igual al anterior, así se tiene la misma matriz  $W$  y no es necesario invertir nuevamente la matriz  $W^T W$ .

7°.- Si las pruebas indican la existencia de curvatura o que  $\eta^{\circ}$  es localmente horizontal, se completa el diseño para el ajuste de una función cuadrática.  $\hat{x}_{op}$  es el punto donde alcanza el máximo esta función cuadrática.

Si  $\hat{\eta}$  no tiene un máximo, sino un punto silla, es conveniente hacer un análisis de cordillera posterior al análisis de ascenso por pendiente máxima, dicho análisis no se incluye en este trabajo.

El método de ascenso por pendiente máxima falla cuando  $\eta^{\circ}$  tiene más de un máximo relativo y el diseño inicial no está localizado en el lugar adecuado, ya que puede detectarse un óptimo; pero, éste puede ser un óptimo local. Por experiencia se sabe que en la mayoría de los casos  $\eta^{\circ}$  presenta un máximo único (Cochran & Cox, 1980)

### 1.3.5 COMENTARIOS SOBRE LOS DIFERENTES MÉTODOS.

El método aleatorio.

Se recomienda usar este método como un proceso exploratorio, tanto para determinar el punto de inicio del método del factor único, como para encontrar el centro del diseño en los métodos de un solo experimento y de ascenso por pendiente máxima.

El método, por sí solo, proporciona para cada N, bajo ciertas hipótesis adicionales, una región de confianza para  $x_{op}$ ; pero, es necesario una muestra muy grande para que la confianza sea aceptable.

Por ejemplo: con cuatro factores y A particionada con diez subdivisiones por coordenada (i.e.  $N=10$  y  $m = 4$ ); la probabilidad de que al menos uno de los puntos en la muestra esté en la misma celdilla de  $x_{op}$  (siempre y cuando  $x_{op} \in A$ ) es:

$$1 - (1 - 1/N^m)^n = 1 - (1 - .1^4)^n = 1 - .99999^n.$$

Para tener una confianza de 95%, el tamaño de la muestra debe cumplir la relación:  $1 - .99999^n = .95$ ; y esto se da si:

$$n = \frac{\log(.05)}{\log(.99999)} = 299571,$$

y para tener una confianza de 90%,  $n = \frac{\log(.10)}{\log(.99999)} = 230256$ . Con un tamaño de muestra así, el método resulta necesariamente caro.

Cabe aclarar que la evaluación de esta probabilidad supone que el error de observación tiene varianza pequeña, de modo que el punto  $x_1$  con respuesta observada mayor realmente está próximo a  $x_{op}$ . Si esto no ocurre la confiabilidad de la región es menor y de hecho no se puede calcular.

El método de un solo experimento.

Si el diseño se localiza cerca de  $x_{op}$  las estimaciones obtenidas con este método coinciden con las estimaciones obtenidas con el método de ascenso por pendiente máxima; pero, si el diseño se encuentra lejos del óptimo, las estimaciones pueden diferir mucho, porque mientras  $\hat{x}_{op}$  esté más lejos del diseño, menos confiable es  $\hat{\eta}$  evaluada en  $\hat{x}_{op}$ . Si se aumenta el área que abarca el diseño (i. e. los puntos del diseño se separan más); se puede evitar, parcialmente esto; sin embargo, cuando el área del diseño es grande la superficie de respuesta en esta región puede ser muy complicada y quizás una función cuadrática no sea lo más adecuado para describirla, en estos casos se necesita un diseño con más puntos para ajustar un polinomio de mayor grado.

#### Método del factor unico.

Este método es muy sencillo porque descompone el proceso en sus componentes elementales. Se ha visto en la práctica, que si los efectos de cada factor en lo individual, son independientes de los otros factores, el método da buenos resultados. Esto significa que la búsqueda realizada sobre cada factor, manteniendo a los otros fijos, va conduciendo paulatinamente al máximo de manera independiente. Pero, si los factores interactúan fuertemente entre si, es necesario una muestra mayor, e inclusive, puede registrarse un óptimo sin que éste exista, principalmente si se está en una cordillera ascendente, como se ve en la figura 1.4 .



figura 1.4 Curvas de nivel de una cordillera transversal ascendente en dirección de  $(x, y) \rightarrow (\infty, \infty)$ . Se muestra la búsqueda de las coordenadas del óptimo usando el método del factor unico. Se busca el punto más alto manteniendo constante la segunda coordenada, este se localiza cerca del "lomo" de la cordillera. En dirección de los ejes, el máximo se encuentra cerca del primer punto localizado.

#### Método de ascenso por pendiente máxima.

Este es el método más utilizado, es un método secuencial que considera en cada etapa el comportamiento local de  $\eta^*$  alrededor de los puntos del diseño. Como ya se dijo, el método no funciona si  $\eta^*$  tiene más de un máximo relativo porque el que se detecta es el más evidente en la región del diseño y no necesariamente es el máximo global.

En el apéndice 1 se muestra el diagrama de flujo de este método.

## 2. VARIANTES QUE SE PROPONEN AL METODO DE ASCENSO POR PENDIENTE MAXIMA.

En el primer capitulo se revisaron, brevemente, los métodos de búsqueda del óptimo, así como los diseños más utilizados; y se dijo, que el más conocido es el de ascenso por pendiente máxima; en este capitulo, se presentan siete formas modificadas de éste, el cual de aquí en adelante se llamará "método original". El propósito de estas modificaciones es encontrar "mejores" formas de estimar al óptimo.

Myers y Khuri (1979), proponen una modificación al método de ascenso por pendiente máxima, la cual consiste en establecer una prueba de hipótesis para determinar si en dirección de la pendiente máxima  $v$ , ya se dió el cambio de ascendente a descendente, las propuestas de este trabajo van en otro sentido.

En las dos primeras variantes se modifica la forma de obtener la longitud del paso en dirección de la máxima pendiente de  $\hat{\eta}$ , y la manera de encontrar el centro del nuevo diseño. En la tercera se incluyen estas modificaciones y, además, se propone que los ajustes, tanto del polinomio de primer grado como el de segundo grado, se hagan con toda la información disponible y no sólo con los datos experimentales de la última etapa. Las cuatro últimas variantes son una modificación del método original y de sus tres primeras variantes, respectivamente; esta modificación consiste en condicionar la finalización de la búsqueda, al hecho que  $\hat{x}_{op}$  se encuentre "cerca" del diseño.

## 2.1 PRIMERA VARIANTE.

Uno de los aspectos que está sujeto a la subjetividad del investigador, es la distancia que existe entre dos puntos diferentes del diseño. Si estos se encuentran muy alejados, se corre el riesgo de no detectar cambios importantes en la función de respuesta  $\eta^*$ ; si por el contrario, los puntos se encuentran muy cerca, se corre el riesgo de que su tendencia quede completamente confundida con la variación aleatoria del proceso. Este hecho despierta la inquietud de explorar nuevos criterios para determinar la longitud del paso adecuada.

Recuerde que la longitud del "paso" es una constante que determina el espaciamento entre dos observaciones subsecuentes en dirección de la pendiente máxima de  $\eta$ . Luego que las pruebas realizadas dan evidencia que el óptimo aun se encuentra "lejos" y que, un plano se ajusta bien a los datos.

En esta parte se propone escoger una longitud del "paso" con la cual se tenga "poca" probabilidad de confundir la tendencia de  $\eta^*$  con el error de las observaciones.

### 2.1.1 OBTENCIÓN DE LA LONGITUD DEL PASO $d$ .

Suponga que la tendencia local de  $\eta^*$ , está satisfactoriamente estimada en una región circundante al diseño mediante un polinomio de primer grado; suponga también, que el vector  $v$  indica la dirección de máxima pendiente de  $\eta$ ; entonces "cerca" del diseño se cumple la relación:

$$\eta^*(idv) < \eta^*((i+1)dv).$$

Dado que las observaciones tienen un error aleatorio, que  $Y_1$  resulte ser mayor que  $Y_{1+1}$ , no necesariamente implica que la tendencia de  $\eta^*$  ya cambió, de ascendente a descendente; este hecho podría deberse únicamente a la aleatoriedad del proceso. Es por ésto, que tiene sentido pedir que la longitud del paso sea tal que permita detectar la tendencia de  $\eta^*$  con poca influencia del error de observación.

Se desea encontrar la magnitud del paso tal que si la tendencia de  $\eta^{\circ}$  en dirección de  $v$  es creciente, la probabilidad que  $Y_1$  resulte ser mayor o igual a  $Y_{1+1}$  no sea mayor a un valor prefijado.

Como:

$$P(Y_1 \geq Y_{1+1}) = P(\eta^{\circ}(1dv) + \epsilon_1 \geq \eta^{\circ}((i+1)dv) + \epsilon_{1+1}) = \quad (2.1)$$

$$P(\eta^{\circ}((i+1)dv) - \eta^{\circ}(1dv) \leq \epsilon_1 - \epsilon_{1+1}) \quad (2.2)$$

con  $\epsilon_1 \sim N(0, \sigma^2)$  independientes, se sigue que:  $\epsilon_1 - \epsilon_{1+1} \sim N(0, 2\sigma^2)$ ; y

$$P(Y_1 \geq Y_{1+1}) = \alpha \quad (2.3)$$

cuando

$$\eta^{\circ}((i+1)dv) - \eta^{\circ}(1dv) = z_{\alpha} \sigma \sqrt{2} \quad (2.4)$$

Por ejemplo: si la probabilidad de detectar equivocadamente un cambio de ascenso a descenso, es igual a .025, la constante  $z_{\alpha}$  en (2.4) es 1.96; si dicha probabilidad es igual a .0013, entonces  $z_{\alpha}$  es 3.

Dado que tanto  $\sigma$  como  $\eta^{\circ}$  son desconocidos, se puede utilizar el hecho que  $SCE/\sigma^2 = (n-2)\hat{\sigma}^2/\sigma^2 \sim \chi^2$  con  $n-2$  grados de libertad y que por lo tanto  $(\epsilon_1 - \epsilon_{1+1})/\sqrt{2}\hat{\sigma} = t$  con  $n-2$  grados de libertad. De ésto se sigue que (2.4) se sustituye por

$$\eta^{\circ}((i+1)dv) - \eta^{\circ}(1dv) = t_{\alpha} \hat{\sigma} \sqrt{2} \quad (2.5)$$

con esta ecuación se puede encontrar un estimador de la longitud del paso  $d$ , al sustituir  $\eta^{\circ}$  por  $\hat{\eta}$  y despejar a  $d$  de la ecuación resultante, con lo que se llega a:

$$\hat{d} = \frac{t_{\alpha} \hat{\sigma} \sqrt{2}}{\hat{\beta}_1 v_1 + \dots + \hat{\beta}_m v_m} \quad (2.6)$$

### 2.1.2 LOCALIZACIÓN DEL CENTRO DEL NUEVO DISEÑO, c.

Para estimar el punto donde  $\eta$  alcanza el óptimo en dirección de  $v$ , se propone utilizar un criterio semejante al que se usa para encontrar el máximo de cada factor en el método del factor único. De esta manera se utiliza toda la información disponible en dirección de  $v$ , y no únicamente la información de las dos últimas observaciones, como ocurre en el método original donde se considera que hubo un cambio en la tendencia de  $\eta$  si la última observación es menor o igual a la penúltima.

A diferencia del método del factor único donde se maximiza cada factor en el orden indicado por el investigador; aquí se busca el máximo en la dirección óptima estimada; la cual no depende, directamente, del criterio del investigador, y sí de la información contenida en los datos. Además, de esta manera  $v$  y el subespacio ortogonal a  $v$ , tienen poca interacción.

Esto sugiere que se puede obtener el estimador del óptimo con pocos "datos".

Específicamente el proceso para localizar el centro del nuevo diseño es el siguiente:

i) Se toman  $n_0$  observaciones en dirección de  $v$ ;

$$Y_i = \eta(i dv) + \epsilon_i; \quad i = 1, \dots, n_0.$$

ii) Estos datos, junto con las observaciones repetidas en el centro del diseño, se utilizan para estimar la función

$$\hat{Y}_t = \hat{\alpha}_0 + \hat{\alpha}_1 t + \hat{\alpha}_2 t^2 \quad (2.7)$$

Note que esta función describe una parábola en dirección de  $v$  y que la variable  $t$  determina el factor real que multiplica al vector  $dv$  (esto es,  $\hat{Y}_t$  es el valor estimado de la respuesta en el vector  $tdv$ )

$$\text{Si } \hat{\alpha}_2 < 0, \text{ el vector dado por: } c = -\frac{\hat{\alpha}_1}{2\hat{\alpha}_2} dv \quad (2.8)$$

es donde  $\hat{Y}_t$  alcanza el óptimo; y por consiguiente resulta ser el centro del nuevo diseño.

En este caso, el nuevo diseño está formado por los puntos de la forma  $c + x_1$ ; donde  $x_1$  es un punto del diseño original. Con este diseño, previamente centrado y escalado; se inicia, nuevamente, el proceso de búsqueda del óptimo.

Si  $\hat{\alpha}_2$  es mayor o igual a cero, (2.7) no presenta un valor máximo. Si esto ocurre, se recomienda estimar a  $c$  nuevamente; la propuesta es que en la nueva estimación de  $c$ , el espaciado entre dos observaciones subsecuentes en dirección de  $v$  se determine por la máxima distancia recorrida en el ajuste anterior, o sea que la longitud del paso sea igual a  $n_0 d$ . Si aún así  $\hat{\alpha}_2$  es mayor a 0, se puede estimar nuevamente el vector  $c$ , pero con una longitud de paso igual a  $n_0$  veces la anterior, o bien, puede hacerse un análisis de cordillera; porque es posible que, al menos localmente, se esté en una cordillera ascendente.

## 2.2 SEGUNDA VARIANTE.

La segunda variante sólo difiere de la primera en la manera de obtener la longitud del paso  $d$ . Para encontrar la longitud del paso, ahora se utilizan los resultados debidos a O'Reilly (1975). Este autor considera que:

si  $Y_1 \sim N(x_1^T \beta, \sigma^2)$  con  $x_1^T$  el  $i$ -ésimo renglón de una matriz  $X$  de rango completo, entonces se dice que una v. a.  $Y$  tiene una distribución estimable si existe una función medible  $h(Y_1, Y_2, \dots, Y_n)$  con la misma distribución de  $Y$ .

La extrapolación en un punto  $x_0$  es válida si  $Y(x_0) \sim N(x_0^T \beta, \sigma^2)$  tiene una distribución estimable.

O'Reilly muestra que la región donde es válido extrapolar con un modelo de regresión lineal está dada por:

$$S = \{x \in \mathbb{R}^m \mid V(\hat{\eta}) \leq \sigma^2\} = \{x \in \mathbb{R}^m \mid w^T (W^T W)^{-1} w \leq 1\}$$

con  $w^T = (1, x_1, x_2, \dots, x_m, x_1^2, x_1 x_2, \dots, x_m^r)$

Entonces se puede suponer que entre dos elementos de  $S$  no existen fluctuaciones importantes que pasen inadvertidas al modelo y de esta manera si  $d$  es solución de la ecuación:

$$(1 \ d v^T) (W^T W)^{-1} \begin{bmatrix} 1 \\ d v \end{bmatrix} = 1 \quad (2.9)$$

se minimiza el riesgo de confundir la tendencia de  $\eta^*$  con el error aleatorio manteniéndose dentro de la región de extrapolación.

Se considera que un punto  $x$  está cerca del diseño, si  $x \in S$ .

### 2.3 TERCERA VARIANTE.

La tercer variante incluye las mismas modificaciones que se hicieron a la segunda variante; pero, además utiliza toda la información disponible y no sólo la del último diseño. Esto permite tener estimadores de menor varianza.



fig.2.1 Datos de dos diseños factoriales,  $2^2$  con repeticiones en el centro; unidos por varios puntos alineados. Tal como quedan después de la primera fase del proceso de búsqueda del máximo.

Con toda la información, la varianza de  $\hat{\eta}$  disminuye; pero, no uniformemente en todas direcciones, ya que el conjunto de puntos de dos diferentes diseños presentan una distribución a lo largo de una recta.

Si los dos diseños de la figura 2.1 están muy alejados, la estimación de  $\hat{\eta}$  se hace con los puntos localizados casi en una recta, por lo que  $V(\hat{\eta})$  es "grande" en el subespacio ortogonal a  $v$ . Para evitar en parte que los puntos del diseño presenten un patrón semejante, proponemos que el nuevo diseño sea de la forma.

$$c + cx_1 \quad (2.10)$$

donde:  $x_1$  es un punto del diseño original;  $c$ , el nuevo centro y  $c$ , una constante

$$c = \|c\| / \|\bar{x}\|$$

donde  $\|\bar{x}\| = (\|x_1\| + \|x_2\| + \dots + \|x_n\|) / n_1$  con  $n_1$  igual al número de vectores del diseño diferentes de cero. De esta manera, el diseño no se encuentra localizado casi en una recta. (Si  $c$  se localiza "lejos" del origen,  $c$  es mayor a 1; Ver figura 2.2).

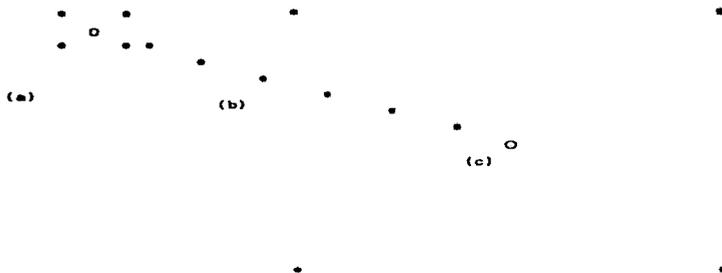


figura 2.2. (a) primer diseño, tamaño "normal". (b) puntos en dirección de la máxima pendiente. (c) segundo diseño, aumentado por un factor igual a  $c$ .

Para comparar la varianza de  $\hat{\eta}$ , en los casos que:

- i)  $\eta$  se estima sólo con los datos del último diseño sin utilizar la

constante  $c$ . (cuando  $c = 1$ , como en el método original)

ii)  $\eta$  se estima con los datos del último diseño, escalado con la constante  $c$ , ó cuando  $c > 1$ .

iii)  $\eta$  se estima con toda la información disponible.

se debe considerar que los puntos del último diseño son:  $z_1 = c + cx_1$  y que éstos se escalan y centran de la siguiente manera

$$(z_1 - c)/c = x_1,$$

entonces la matriz  $W$  asociada a este diseño es igual a la del primer diseño y

$$V(\hat{\eta}((z - c)/c)) = w^T(W^T W)^{-1} w \sigma^2 \quad (2.11)$$

es la varianza de  $\hat{\eta}$ .

Note que  $c$  es una constante de escalamiento; por tal razón el conjunto  $S = \{z \mid w^T(W^T W)^{-1} w \leq 1\}$  tiene mayor medida de Lebesgue conforme el valor de  $c$  aumenta; de esto se sigue que  $S$  con  $c = 1$  es "menor" que  $S$  con  $c > 1$ .

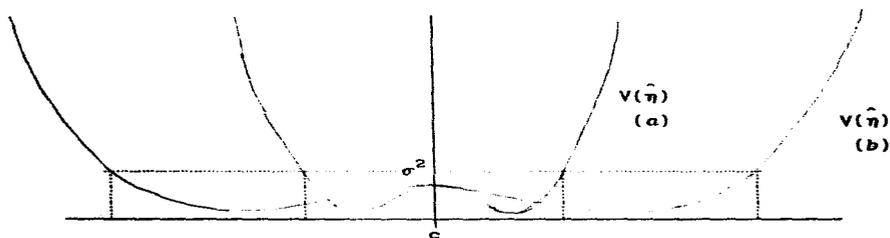


fig. 2.3 Grafica de  $V(\hat{\eta})$  en el caso unifactorial con un diseño simétrico con respecto a  $C$ ; (a) cuando  $c=1$ ; (b) cuando  $c>1$ . Se puede ver el conjunto  $S$  de puntos donde  $V(\hat{\eta}) < \sigma^2$ .

Al incorporar toda la información acumulada, la matriz asociada

al modelo es  $W^T W + V^T V$  ( $W$  contiene la información del último diseño y  $V$ , la de los anteriores).  $V(\hat{\eta})$  entonces resulta ser:

$$V(\hat{\eta}) = W^T (W^T W + V^T V)^{-1} W \sigma^2, \quad (2.12)$$

Y como

$$(W^T W + V^T V)^{-1} = (W^T W)^{-1} - (W^T W)^{-1} V^T (I - V(W^T W)^{-1} V^T)^{-1} V (W^T W)^{-1}$$

se sigue que

$$V(\hat{\eta}) = W^T (W^T W)^{-1} W \sigma^2 - W^T (W^T W)^{-1} V^T (I + V(W^T W)^{-1} V^T)^{-1} V (W^T W)^{-1} W \sigma^2$$

las matrices involucradas en esta expresión son positivas definidas por lo tanto (2.12) siempre es menor que (2.11). De aquí se sigue que  $V(\hat{\eta})$  es menor cuando se utiliza la información acumulada.

Cabe aclarar que si  $c$  es mayor a uno, el área que abarca cada nuevo diseño, es mayor. En un área mayor, como ya se dijo, puede ser que la función de respuesta sea compleja y no se ajuste bien a un polinomio de segundo grado; sin embargo, al encontrar el centro del nuevo diseño, se vió que entre 0 y  $c$ , un polinomio de segundo grado se ajusta bien a la tendencia de  $\eta$ . Como ésto no garantiza que  $\eta^*$  tenga un máximo local que se pueda encontrar con el ajuste de una función cuadrática, se recomienda hacer una prueba de falta de ajuste sobre la función cuadrática estimada para determinar que tan bueno es el ajuste.

#### 2.4 CUARTA, QUINTA, SEXTA Y SÉPTIMA VARIANTE.

Una nueva modificación es que para terminar la búsqueda del óptimo se pide que  $\hat{x}_{op}$  pertenezca a  $S$ , lo que equivale a pedir que  $\hat{x}_{op}$  se encuentre cerca del diseño.

El método original con esta modificación es la cuarta variante.  
 La primera variante con esta modificación es la quinta variante.  
 La segunda variante con esta modificación es la sexta variante.  
 La tercera variante con esta modificación es la séptima variante.

Con esto se garantiza que en  $\hat{x}_{op}$  la extrapolación es aceptable, y que  $\hat{\eta}_{op}$  tiene una varianza menor a  $\sigma^2$ .

## 2.5 RESUMEN.

Las modificaciones propuestas al método original se concretaron en los siguientes cinco puntos:

- i) La forma de determinar la distancia entre dos observaciones sucesivas en dirección de la pendiente máxima.
- ii) La manera de encontrar el punto donde  $\eta$  alcanza su máximo en dirección de  $v$ , que es el centro del nuevo diseño.
- iii) La incorporación de toda la información disponible para efectuar las estimaciones subsecuentes.
- iv) El aumento del área considerada por cada nuevo diseño con la incorporación del término  $c$ .
- v) La condición de que  $\hat{x}_{op} \in S$  para terminar el proceso de búsqueda.

Note que en cada nuevo proceso de búsqueda, el conjunto  $S$  es recalculado.

Para sondear las bondades de los diferentes métodos de búsqueda, se llevaron a cabo una serie de simulaciones cuyas particularidades y resultados se presentan en el siguiente capítulo.

En el apéndice 1, se presentan los diagramas de flujo de los métodos presentados en este capítulo.

### 3 LA SIMULACIÓN.

#### EL MÉTODO DE MONTE CARLO. UNA MANERA DE SONDEAR LA BONDAD DE LOS DIFERENTES MÉTODOS DE BÚSQUEDA.

##### 3.1 GENERALIDADES.

En el capítulo 1 se hizo una revisión de los métodos de búsqueda del óptimo más conocidos; en el capítulo 2 se mencionó que el método de uso más generalizado es el de ascenso por pendiente máxima: el cual, parece proporcionar "mejores" estimadores tanto de  $x_{op}$  como de  $\eta_{op}$ . También se dijo que con algunas modificaciones este método se puede mejorar y en esa línea se propusieron siete variantes. En este capítulo se presentan las características y los resultados de una serie de simulaciones cuya finalidad es sondear la bondad de los diferentes métodos propuestos.

Es razonable pensar que la bondad de los estimadores obtenidos con los diferentes métodos de búsqueda, no es igual en todos los casos. Para cada función de respuesta  $\eta$ , y para cada valor de  $\sigma^2$ , debe existir un método, y posiblemente también un diseño, con el que se obtengan mejores estimadores.

Estudiar analíticamente el sesgo y la varianza de los estimadores, resulta inaccesible por la misma naturaleza de sus distribuciones. Bajo esta circunstancia, se propone sondear la bondad de cada método utilizando el método de Monte Carlo (mediante simulación); y así, posteriormente, elegir al mejor, con base en una serie de criterios relacionados con la bondad del procedimiento.

El estudio comparativo no incluyó a los métodos aleatorio y de un solo experimento; ya que, como se dijo, su confiabilidad es pobre. Los

métodos incluidos en el análisis son: el del factor único, el de ascenso por pendiente máxima y las siete variantes de este último propuestas en este trabajo.

Para identificar fácilmente estos métodos se propone la siguiente nomenclatura.

METODO	IDENTIFICACIÓN
i) El del factor unico .....	M1
ii) El de ascenso por pendiente máxima.....	M2
iii) La primera variante a M2.....	M2a
iv) La segunda variante a M2.....	M2b
v) La tercera variante a M2.....	M2c
vi) M2 con restricción para terminar.....	M2T
vii) M2a con restricción para terminar.....	M2aT
viii) M2b con restricción para terminar.....	M2bT
ix) M2c con restricción para terminar.....	M2cT

La simulación se realizó con un programa que:

1°.- Realiza 500 veces una rutina consistente en:

i) La generación de una muestra aleatoria de una función de segundo grado del tipo:

$$Y = - (x_{op} - x)^T A^T A (x_{op} - x) + \epsilon_1 \quad (3.1)$$

$$\epsilon_1 \sim N(0, \sigma^2) \quad \text{Independientes}$$

y

ii) El cálculo de  $\hat{x}_{op}$  y de  $\hat{\eta}_{op}$ , usando cada uno de los diferentes métodos de búsqueda.

2°.- Obtiene la variación promedio empírica de los estimadores con base en los 500 resultados de la simulación.

El programa se elaboró para hacer simulaciones de superficies de respuesta en función de cualquier número de factores; sin embargo, para este trabajo sólo se probó con dos factores.

Se consideraron dos diferentes esquemas de diseño.

El primero consiste en el factorial  $2^2$  para la estimación del plano, y el compuesto con  $\alpha = 1$  para estimar la función cuadrática.

El segundo, consiste en el simplex para las estimaciones del plano y el simplex compuesto con  $\alpha = 1$  para las estimaciones del polinomio de segundo grado.

En ambos casos, se agregaron tres repeticiones en el centro para estimar el error puro.

Se aclara que el método del factor único no necesita de un diseño específico, por lo que su aplicabilidad es independiente del diseño utilizado.

### 3.2 PROCEDIMIENTO.

Para generar la muestra se consideró:

1°.- Que la función  $\eta^*$  pertenece a la familia de funciones cuadráticas cuya gráfica tiene un máximo único y bien definido, o sea aquellas de la forma:

$$\begin{aligned}\eta^*(x_1, x_2, \dots, x_m) &= - \sum (a_{10} + \sum a_{1j} x_j)^2 & (3.2) \\ &= -(\mathbf{x}_{op} - \mathbf{x})^T \mathbf{A}^T \mathbf{A} (\mathbf{x}_{op} - \mathbf{x})\end{aligned}$$

con

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix} \quad \text{una matriz real} \quad (3.3)$$

por lo que  $\eta^*$  tiene un máximo único y bien definido, si y sólo si,  $\mathbf{A}$  es invertible.

En particular, fueron 3 las matrices usadas.

$$a) A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}; \quad b) A = \begin{bmatrix} -.09 & .09 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad c) A = \begin{bmatrix} -1 & 1 \\ .09 & .09 \end{bmatrix}$$

En el caso a) las curvas de nivel de la función  $\eta^*$  son círculos concéntricos al punto óptimo. Esta función proporciona un ejemplo de estudio para determinar la naturaleza de los estimadores, cuando la rapidez de ascenso al punto máximo no depende de la dirección que se tome para llegar a él.

En el caso b) las curvas de nivel son elipses concéntricas al óptimo, con una marcada diferencia entre la magnitud de sus ejes, el ángulo formado entre el eje mayor de las elipses y el eje  $x_1$  es igual a  $135^\circ$ . Así entonces, la evidencia de curvatura se manifiesta de manera diferente desde las distintas direcciones por la que se ascienda al óptimo, además los factores se relacionan a través de la respuesta.

En el último caso, las curvas de nivel de la función de respuesta son elipses concéntricas al óptimo, similares al caso b), sólo que el ángulo de inclinación de su eje mayor es de  $45^\circ$  con respecto al eje  $x_1$ .

La elección de la matriz A, obedeció a razones puramente prácticas; se buscó tener un representante de la familia de funciones:

$$\eta(\mathbf{x}) = -(\mathbf{x} - \mathbf{x}_{op})^T c^2 A^T (\mathbf{x} - \mathbf{x}_{op})$$

con  $c=1$ , en 3 diferentes casos para analizar situaciones contrastantes.

Note que para valores "grandes" de  $c$  la tendencia ascendente de  $\eta$  es más pronunciada, por lo que parece lógico pensar, que en esos casos es más fácil estimar al óptimo.

Para cada matriz A se utilizaron tres diferentes posiciones del óptimo dadas por los puntos.

$$\mathbf{x}_{op} = (0, 25), \quad \mathbf{x}_{op} = (25, 25) \quad \text{y} \quad \mathbf{x}_{op} = (200, 200) \quad (3.4)$$

Considerando que el proceso de búsqueda siempre se inicia en el origen de  $\mathbb{R}^2$ , la localización del óptimo presupone una dirección para llegar a él y una distancia que debe recorrerse para alcanzarlo.

Finalmente, para cada combinación de matriz A y vector  $x_{op}$ , se utilizaron 10 diferentes valores para  $\sigma^2$ :  $\sigma^2 = .1, .2, .3, \dots, 1$ .

En resumen: se utilizaron 9 métodos de búsqueda con dos diferentes esquemas de diseño, con 9 diferentes funciones de respuesta, con 10 diferentes varianzas para el error; lo que da, finalmente, 180 casos estudiados por simulación para cada método; esto es 1620 casos-método estudiados. (con 500 repeticiones cada uno). Sin embargo, se reconoce que se está lejos de tener resultados generales.

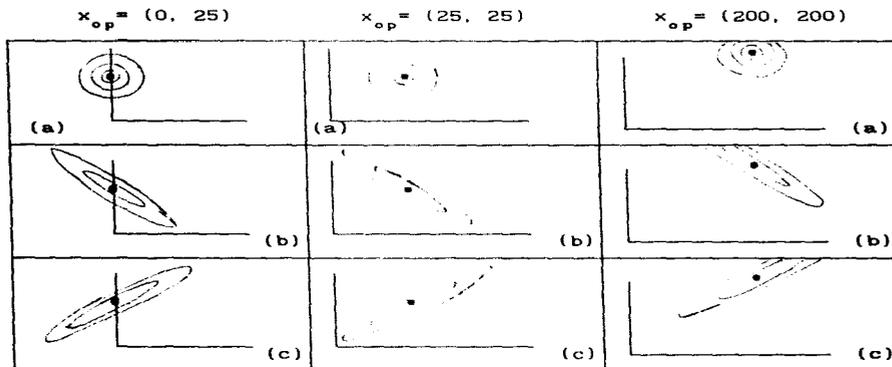


Fig. 3.1 Curvas de nivel de las funciones utilizadas en la simulación, la letra entre parentesis indica cada caso de matriz A utilizada.

Los criterios para elegir al "mejor" método son:

a) La distancia empírica promedio entre  $\hat{x}_{op}$  y  $x_{op}$ , la cual está dada por la expresión:

$$\sum \| \hat{x}_{op} - x_{op} \| / 500 \quad (3.5)$$

b) La distancia empírica promedio entre  $\hat{\eta}_{op}$  y  $\eta_{op}$ , dada por la expresión

$$\sum | \eta(\hat{x}_{op}) - \eta_{op} | / 500 \quad (3.6)$$

c) El tamaño de muestra promedio requerido para converger: la muestra incluyó todas las observaciones: desde las del primer diseño, hasta las del último, incluyendo las observaciones que se tomaron en dirección de v.

d) La variación empírica promedio de  $\hat{\eta}_{op}$  entre  $\sigma$ :

$$\sum | \eta(\hat{x}_{op}) - \eta_{op} | / 500\sigma$$

e) Y el tiempo promedio de proceso computacional.

En el apéndice 2 se presenta una selección de los resultados de la simulación.

**TABLA 3.1**  
 Fragmento de la tabla 1 del apéndice 2

$$Y = -160,000 + 800x_1 + 800x_2 - 2x_1^2 - 2x_2^2 + c$$

$$\bar{x}_{op} = (200, 200)$$

diseño factorial.

$$A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M2T	M2aT	M2bT	M2cT
.1	.00017	456.0	437.9	454.9	158.4	.0124	.00282	.00352	4.7E-8
	.00054	1441	1381	1438	500.9	.0392	.00892	.01113	1.5E-8
	.00834	13.4	13.14	13.26	12.46	.0333	.327	.0339	.00006
	30.07	11.0	11.0	11.0	11.0	35.7	44.4	44.6	20.0
.2	.00025	289.3	286.2	299.0	589.0	.0697	.00755	.00892	4.7E-8
	.00056	1446.	4901.	4497.	2945.	.0345	.037	.0446	2.3E-8
	.00980	18.5	18.6	18.42	18.8	.0449	.0507	.0512	.00009
	30.02	21.3	11.6	11.6	11.8	42.07	46.6	47.60	20.08
.3	.00043	1198.	1395.	4652.	1293.	.0852	.0073	.00814	5.3E-8
	.00145	4919.	4652.	3995.	4318.	.0285	.0244	.0271	1.3E-7
	.01304	16.14	23.13	22.04	20.59	.0535	.0510	.0522	.00001
	30.20	54.47	13.88	14.00	12.67	62.31	47.94	51.65	20.18
.4	.00051	12.38	2310.	1716.	1539.	.0115	.0112	.0119	4.8E-8
	.00127	32.08	5775.	4290.	3846.	.0288	.0280	.0292	9.5E-8
	.0140	19.94	29.09	25.45	19.86	.0646	.0595	.0637	.00014
	30.07	71.25	15.45	15.9	13.0	86.9	46.9	52.3	20.27
.5	.00081	72.7	1945.	2422	89.3	.0186	.0223	.0199	5.7E-8
	.00162	145	3890	4845	17.3	.0308	.0372	.0333	9.3E-8
	.01834	17.35	26.93	29.94	16.17	.0729	.0834	.0865	.00017
	30.27	85.54	17.86	17.65	14.46	95.56	47.11	55.15	20.47
.6	.00096	3.27	3393	3502	82.4	.0205	.0252	.0224	6.7E-8
	.00161	5.4	5655	5837	17.6	.0299	.0360	.0321	9.5E-8
	.01954	18.58	35.84	35.48	19.09	.0719	.0981	.0846	.00016
	30.42	100.3	20.12	21.32	14.47	110.2	49.5	57.3	20.46
.7	.00104	9.97	3446	3175	1432	.0247	.0248	.0259	7.8E-8
	.00142	14.25	4923	4536	19.20	.0311	.0309	.0324	9.5E-8
	.02034	18.4	35.08	34.55	22.68	.0721	.0959	.0946	.00018
	31.41	116.5	20.0	22.52	15.6	135.7	44.67	61.74	20.49
.8	.00130	.0130	4695	5155	.1254	.0151	.0178	.0167	7.6E-8
	.00163	.016	5869	6443	.1752	.0304	.0356	.0334	8.8E-8
	.0227	15.25	41.9	40.82	18.25	.0646	.0758	.0745	.00020
	30.08	133	25.5	26.36	16.4	92.37	46.17	54.40	20.49

(1)\*  
 (2)\*  
 (3)\*  
 (4)\*

$$(1) \cdot \sum | \eta(\hat{x}_{op}) - \eta_{op} | / 500$$

$$(3) \cdot \sum \| \hat{x}_{op} - x_{op} \| / 500$$

$$(2) \cdot \sum | \eta(\hat{x}_{op}) - \eta_{op} | / 5000\sigma$$

$$(4) \cdot \text{tamaño de la muestra}$$

TABLA 3.2.

Tiempo de proceso computacional

Cada renglón equivale al tiempo de cómputo requerido para obtener los resultados en conjunto para los 10 valores de  $\sigma^2$ , la comilla (') representa los minutos; las comillas representan los segundos (").

Si el programa requirió de mas de 10 minutos para terminar, se consideró que el proceso no converge

A y $\hat{x}_{op}$	M1	tiempo acumulado para		M2c	tiempo acumulado para		M2cT
		M2	M2a y M2b		M2T	M2aT y M2bT	
$\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$ (25, 25)	40"	1' 51"	1' 04"	4' 29"	1' 20"		
$\begin{bmatrix} -1 & 1 \\ .09 & .09 \end{bmatrix}$ (0, 25)	1' 01"	2' 42"	1' 08"	no hubo convergencia	2' 44"		
$\begin{bmatrix} -.09 & .09 \\ 1 & 1 \end{bmatrix}$ (25, 25)	1' 13"	2' 01"	1' 08"	no hubo convergencia	2' 44"		
$\begin{bmatrix} -.09 & .09 \\ 1 & 1 \end{bmatrix}$ (0, 25)	2' 00"	1' 51"	1' 04"	4' 29"	2' 44"		
$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (200, 200)	1' 31"	2' 42"	1' 08"	no hubo convergencia	2' 18"		
$\begin{bmatrix} 1 & 1 \\ .09 & .09 \end{bmatrix}$ (200, 200)	7' 20"	2' 02"	1' 08"	no hubo convergencia	2' 18"		
$\begin{bmatrix} -.09 & .09 \\ 1 & 1 \end{bmatrix}$ (200, 200)	7' 40"	4' 12"	1' 40"	8' 00"	2' 54"		

### 3.3 ANALISIS DE LOS RESULTADOS DE LA SIMULACIÓN.

De acuerdo a los resultados de la simulación se puede observar que sí:

- i) La matriz que define la función es  $A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$  ;
- ii) El punto óptimo está "cerca" del centro del diseño ;

y

- iii)  $\sigma^2$  es pequeña,

los estimadores de  $x_{op}$  y de  $\eta_{op}$  obtenidos con cualquiera de los métodos tienen menor varianza que para otras condiciones diferentes a i), ii) y a iii).

El punto i) sugiere que al utilizar una transformación  $T: R^n \rightarrow R^n$  que convierte elipses en círculos, se pueden escoger diseños que mejoren las estimaciones.

El punto ii) justifica, al menos parcialmente, la bondad de un método secuencial de búsqueda como el de ascenso por pendiente máxima, el cual en cada paso, al menos en teoría, utiliza un diseño más cercano a  $x_{op}$ .

De los dos esquemas de diseño utilizados, el factorial y el compuesto proporcionan estimadores ligeramente más precisos que con el esquema formado por el simplex y el simplex compuesto; esto se debe, posiblemente, a que el primero tiene más puntos y abarca una mayor área.

De los diferentes métodos de búsqueda del óptimo, la séptima variante (M2cT) proporciona con una muestra pequeña, estimadores tanto de  $x_{op}$  como de  $\eta_{op}$ , con menos variación empírica promedio.

Si  $\alpha$  es la variación promedio empírica de los estimadores con

cualquier método, (excepto el M&T) la variación promedio de la séptima variante es, en todos los casos, menor a  $\alpha/100$ .

La séptima variante necesitó de relativamente pocos datos para obtener las estimaciones. El tamaño de la muestra fue casi invariante en todos los casos (de 16 a 26). El método del factor único también presentó estimadores con "buena" precisión; sin embargo, sólo con el óptimo cerca del origen y con la función de respuesta de curvas de nivel circulares la muestra fue aceptable, en los otros casos la muestra es muy grande.

Se pudo observar que la séptima variante ocupa más tiempo computacional que otros métodos de precisión pobre; pero esto se debe, seguramente, a que en cada etapa se invierte la matriz  $W^T W$  y porque también se ocupa más memoria; sin embargo, con el uso cada vez más generalizado de computadoras de enorme rapidez, este aspecto no constituye un problema.

La introducción del término  $c$  para agrandar al diseño tiene el siguiente efecto en la función de respuesta:

$$\eta(c\mathbf{x}) = (c\mathbf{x} - \mathbf{x}_{op})^T A^T A (c\mathbf{x} - \mathbf{x}_{op}) = (\mathbf{x} - \frac{1}{c} \mathbf{x}_{op})^T c^2 A^T A (\mathbf{x} - \frac{1}{c} \mathbf{x}_{op})$$

o sea que, es equivalente a tener el mismo diseño, con una función de respuesta de tendencia ascendente más acentuada (porque  $c > 1$ ) y con el punto óptimo más cercano al diseño. (el centro del diseño es el origen) Esto justifica, de alguna manera, los resultados de la simulación con respecto de la séptima variante la cual incluyó este término.

Se efectuaron dos simulaciones más con tres factores, los resultados obtenidos fueron semejantes a los de dos factores.

Las simulaciones se efectuaron limitando el tiempo de cómputo a 10 minutos de proceso. Si en ese tiempo no se llegó a encontrar el resultado se consideró que el método no convergía.

En resumen, los resultados de la simulación señalan que la séptima variante proporciona "buenos" estimadores tanto de  $\mathbf{x}_{op}$  como de  $\eta_{op}$ .

#### 4. UNA REGION DE CONFIANZA PARA $x_{op}$ Y PARA $\eta_{op}$ .

En el capítulo anterior se sondeó, mediante una serie de simulaciones, la "bondad" de los diferentes métodos de búsqueda del óptimo; los resultados indican que la séptima variante al método de ascenso por pendiente máxima es el mejor. En este capítulo se van a dar los elementos para construir las regiones de confianza, tanto para el vector  $x_{op}$ , como para el número  $\eta_{op}$ ; un análisis de estas regiones de confianza, permite observar cuales son las condiciones experimentales que proporcionan una región de confianza más pequeña.

Este tema fue primeramente tratado por Box y Hunter (1954 y 1957) quienes se basaron en el teorema de Fieller para encontrar una región de confianza del punto donde  $\eta$  alcanza el óptimo, posteriormente Carter, Chinchilli, Campbell y Wampler (1984) siguiendo una idea de Rao (1973) desarrollan una región de confianza para los valores de  $g(\beta)$  una función de los valores del vector de parámetros  $\beta$ . Escogiendo adecuadamente la función  $g$ , se puede encontrar los intervalos de confianza para cada una de las coordenadas del punto donde  $\eta$  alcanza el óptimo. Carter, Chinchilli, Myers y Campbell (1986) exploran la misma idea anterior para definir los intervalos de confianza aplicandolos a problemas específicos.

Es importante notar que de acuerdo al teorema de Fieller, (Fieller, 1944) las regiones de confianza de  $x_{op}$ , propuestas por Box y Hunter, tienen la característica que para una muestra dada, siempre existe un valor  $\alpha > 0$  tal que  $P(x_{op} \in R^\alpha) = \alpha$ .

Sin embargo, como se verá más adelante, se puede probar que para una significancia igual a este valor de  $\alpha$ , se aceptaría la hipótesis  $H: \eta = c$ , constante, y como una función constante no presenta máximos ni

mínimos locales, tiene sentido la suposición que  $x_{op} \in R^n$  con probabilidad positiva.

El desarrollo de las regiones de confianza para  $x_{op}$  que aquí se dan se basan en la idea original de Box y Hunter (1954), a estas regiones se les denota "regiones de confianza usuales".

Debe resaltarse que en este contexto "óptimo" significa "máximo".

#### 4.1 INTERVALO DE CONFIANZA EN EL CASO DE TENER UN ÚNICO FACTOR.

En esta parte se analiza la situación en la que  $\eta$  es función de un único factor. Éste es el caso trivial; pero, ejemplifica de manera muy simple la construcción de las regiones de confianza.

##### 4.1.1 INTERVALO DE CONFIANZA PARA $x_{op}$ .

En esta sección se discuten tres regiones de confianza para el término  $x_{op}$ .

La primera es la región reportada en la literatura, (Box & Hunter, 1954) por tal razón en este trabajo se le llama "la región de confianza usual".

La segunda es una generalización de la región usual, pero de menor longitud, por tal razón se denota como "la región de longitud mínima", y es una aportación de este trabajo.

La tercera es una región de confianza sesgada, ya que siempre contiene al origen; pero presenta las siguientes características que la hacen interesante.

- i) Se puede construir en condiciones en las que tanto la región usual como la región de longitud mínima no existen; y

ii) La región de confianza de longitud mínima se puede representar como la intersección de una región usual y de esta región, y esta forma de representación es fácilmente generalizable a más de dos factores.

Esta última región de confianza también es una aportación de este trabajo y se le llama "región auxiliar", ya que se utiliza para encontrar el intervalo de longitud mínima.

En este contexto, decir que un intervalo existe es equivalente a decir que el intervalo está acotado.

**Consideraciones generales:** Cuando la respuesta depende de un único factor, las observaciones se pueden aproximar, en un entorno del óptimo, con el modelo:

$$Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon \quad (4.1)$$

donde  $\beta_0$ ,  $\beta_1$  y  $\beta_2$  son parámetros desconocidos,  $\beta_2 < 0$  y  $\epsilon \sim N(0, \sigma^2)$ , son variables aleatorias independientes.

La forma matricial de (4.1) es:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} = W\beta + \epsilon; \quad (4.2)$$

con  $\epsilon \sim N(0, \sigma^2 I)$

El estimador de mínimos cuadrados de  $\beta$ , es el vector aleatorio:  
 $\hat{\beta} \sim N(\beta, V)$  con

$$V = \sigma^2 \begin{bmatrix} n & \sum x_1 & \sum x_1^2 \\ \sum x_1 & \sum x_1^2 & \sum x_1^3 \\ \sum x_1^2 & \sum x_1^3 & \sum x_1^4 \end{bmatrix}^{-1} \quad (4.3)$$

Si los puntos del diseño se escogen simétricos respecto del origen; entonces, tanto  $\sum x_1$ , como  $\sum x_1^3$  resultan ser cero, por lo que la matriz de varianza covarianza es igual a:

$$V = \sigma^2 \begin{bmatrix} n & 0 & \sum x_1^2 \\ 0 & \sum x_1^2 & 0 \\ \sum x_1^2 & 0 & \sum x_1^4 \end{bmatrix}^{-1} \quad (4.4)$$

y la distribución de los estimadores  $\hat{\beta}_0$ ,  $\hat{\beta}_1$  y  $\hat{\beta}_2$  es tal que:

$$\begin{aligned} \hat{\beta}_0 &\sim N(\beta_0, \sum x_1^4 \sigma^2 / (n \sum x_1^4 - (\sum x_1^2)^2)), \\ \hat{\beta}_1 &\sim N(\beta_1, \sigma^2 / \sum x_1^2) \\ \hat{\beta}_2 &\sim N(\beta_2, n \sigma^2 / (n \sum x_1^4 - (\sum x_1^2)^2)) \end{aligned} \quad (4.5)$$

y  $\hat{\beta}_1$  es independiente de  $\hat{\beta}_0$  y  $\hat{\beta}_2$ .

Un estimador insesgado de  $\sigma^2$  es  $\hat{\sigma}^2 = \text{SCE}/(n-3)$  y se tiene que  $\text{SCE}/\sigma^2 \sim \chi_{n-3}^2$ .

El estimador de la matriz de varianza covarianza se encuentra sustituyendo  $\sigma^2$  por su estimador; en otras palabras,  $\hat{V} = \hat{\sigma}^2 V / \sigma^2$ .

$\hat{\sigma}^2$  y  $\hat{\beta}$  son independientes, ya que  $\text{SCE} = y^T(I - W(W^T W)^{-1} W^T)y$ , y  $\hat{\beta} = (W^T W)^{-1} W^T y$  lo que implica que  $(W^T W)^{-1} W^T (I - W(W^T W)^{-1} W^T) = 0$ . (Searle, 1971).

De aquí en adelante se considera que el diseño elegido es simétrico con respecto al cero, por lo que los estimadores  $\hat{\beta}_0$ ,  $\hat{\beta}_1$  y  $\hat{\beta}_2$  presentan las propiedades en (4.5)

La estimación de la función de respuesta es entonces:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2 \quad (4.6)$$

y presenta un máximo en

$$\hat{x}_{op} = -\hat{\beta}_1 / 2\hat{\beta}_2 \quad (4.7)$$

si y sólo si  $\hat{\beta}_2$  es negativo; en cuyo caso, un estimador de la respuesta máxima es

$$\hat{\eta}_{op} = \hat{\eta}(\hat{x}_{op})$$

Para obtener el intervalo de confianza usual, se utiliza la variable aleatoria

$$\hat{\eta}'(x_{op}) = \hat{\ell}(x_{op}) = \hat{\beta}_1 + 2\hat{\beta}_2 x_{op} \quad (4.8)$$

normal con media cero y varianza

$$V(\hat{\ell}) = V(\hat{\beta}_1) + 4x_{op} \text{cov}(\hat{\beta}_1, \hat{\beta}_2) + 4x_{op}^2 V(\hat{\beta}_2) = V(\hat{\beta}_1) + 4x_{op}^2 V(\hat{\beta}_2)$$

por lo que

$$\frac{\hat{\ell}(x_{op}) / \sqrt{V(\hat{\ell})}}{\hat{\sigma}/\sigma} = \hat{\ell}(x_{op}) / \sqrt{V(\hat{\ell})}$$

es una variable aleatoria con distribución  $t$  de student con  $n-3$  grados de libertad.

**Definición.**

**Intervalo de confianza usual.** El conjunto

$$\{ x \mid -t_{\alpha/2} \leq \hat{\ell}(x) / \sqrt{V(\hat{\ell})} \leq t_{\alpha/2} \} \quad (4.9)$$

es el intervalo de confianza usual para  $x_{op}$  con un nivel de significancia igual a  $\alpha$ .

Note que  $\hat{x}_{op}$  está contenido en (4.9), por lo que el intervalo usual no es vacío.

**Observaciones:**

— Si  $t_1 < t_2$  entonces

$$\{ x \mid -t_1 \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_1 \} \subseteq \{ x \mid -t_2 \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_2 \}$$

— Dado que el cuadrado de una v.a.  $t$  es igual a una v.a.  $F$ , el intervalo usual es igual al conjunto:

$$\{ x \mid \hat{\ell}^2(x) / \hat{V}(\hat{\ell}) < F_{1, n-3} \} \quad (4.10)$$

donde  $F_{1, n-3} = t_{\alpha/2}^2$  es el valor tabulado de la distribución  $F$  y que por comodidad de aquí en adelante se escribirá sin subíndices.

**Teorema 4.1** (Box y Hunter, 1954)

El intervalo de confianza usual está acotado si y sólo si

$$\hat{\beta}_2^2 > F_{\alpha} \hat{V}(\hat{\beta}_2).$$

Note que esto equivale a decir que el intervalo de confianza usual está acotado si y sólo si en la prueba

$$H_0: \beta_2 = 0 \quad \text{contra} \quad H_1: \beta_2 \neq 0$$

se rechaza la hipótesis nula con un nivel de significancia igual a  $\alpha$ .

**Demostración.**

Si se sustituyen  $\hat{\ell}(x)$  y  $\hat{V}(\hat{\ell})$  en (4.10) y se desarrolla la expresión algebraica, se tiene que el intervalo de confianza usual es:

$$\{ x \mid \hat{\beta}_1^2 + 4\hat{\beta}_1\hat{\beta}_2x + 4\hat{\beta}_2^2x^2 < F_{1, n-3}(\hat{V}(\hat{\beta}_1) + 4x^2\hat{V}(\hat{\beta}_2)) \} =$$

$$\{ x \mid (\hat{\beta}_1^2 - F\hat{V}(\hat{\beta}_1) + 4\hat{\beta}_1\hat{\beta}_2x + 4(\hat{\beta}_2^2 - F\hat{V}(\hat{\beta}_2))x^2 < 0 \} \quad (4.11)$$

para que este conjunto esté acotado es necesario que la función cuadrática sea convexa, y esto se da si y sólo si  $\hat{\beta}_2^2 - FV(\hat{\beta}_2)$  es un número positivo, y esto es equivalente a  $\hat{\beta}_2^2 > FV(\hat{\beta}_2)$ , que es lo que se quería probar.



fig. 4.1 . Intervalo de confianza para  $x_1$ , (remarcado en negro) según la relación (4.11), en el caso que  $\hat{\beta}_2^2 > FV(\hat{\beta}_2)$ .

**Corolario.**

Si el intervalo de confianza usual está acotado, sus puntos extremos son las dos raíces de la ecuación de segundo grado

$$(\hat{\beta}_1^2 - FV(\hat{\beta}_1)) + 4\hat{\beta}_1\hat{\beta}_2x + 4(\hat{\beta}_2^2 - FV(\hat{\beta}_2))x^2 = 0$$

por lo que el intervalo usual es igual al conjunto

$$= \{ x \mid x_1 < x < x_2 \}$$

$$\text{donde } x_1 = \frac{-\hat{\beta}_1\hat{\beta}_2 - \sqrt{t_{\alpha/2}^2(V(\hat{\beta}_1)\hat{\beta}_2^2 + V(\hat{\beta}_2)\hat{\beta}_1^2) - t_{\alpha/2}^4 V(\hat{\beta}_2)V(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_{\alpha/2}^2 V(\hat{\beta}_2))}$$

$$y \quad x_2 = \frac{-\hat{\beta}_1 \hat{\beta}_2 + \sqrt{t_{\alpha/2}^2 (\hat{V}(\hat{\beta}_1) \hat{\beta}_2^2 + \hat{V}(\hat{\beta}_2) \hat{\beta}_1^2) - t_{\alpha/2}^4 \hat{V}(\hat{\beta}_2) \hat{V}(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_{\alpha/2}^2 \hat{V}(\hat{\beta}_2))}$$

**Definición.**

Intervalo de confianza de longitud mínima: (aportación de este trabajo) Es el intervalo de la forma

$$\{ x \mid -t_\epsilon \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon} \} \quad (4.12)$$

para  $0 \leq \epsilon < \alpha$ ; con longitud mínima.

Note que (4.9) es el caso particular de (4.12), cuando  $\epsilon = \alpha/2$ .

**Lema 4.1**

Si la función  $\hat{\eta}$  tiene un punto máximo ( $\hat{\beta}_2 < 0$ ) entonces la función  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  es decreciente en el conjunto

$$\{ x \mid \hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} \} \quad (4.13)$$

**Demostración.**

Observe que  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell}(x))}$  es una función continua en  $\mathbb{R}$ , y que:

$$i) \quad \lim_{x \rightarrow \infty} \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} = -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} > 0; \quad y$$

$$ii) \quad \lim_{x \rightarrow -\infty} \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} = \hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} < 0.$$

por lo tanto  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell}(x))}$  es una función acotada.

Además,  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell}(x))}$  presenta un único punto crítico en

$$x_0 = \hat{\beta}_2 \hat{V}(\hat{\beta}_1) / \hat{\beta}_1 \hat{V}(\hat{\beta}_2) \quad y$$

$$\hat{\ell}(x_0) / \sqrt{\hat{V}(\hat{\ell}(x_0))} = (\text{signo de } \hat{\beta}_1) \sqrt{\hat{\beta}_1^2 / \hat{V}(\hat{\beta}_1) + \hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2)}$$

Caso  $\hat{\beta}_1 > 0$ :

Como  $\hat{\beta}_1 > 0$  y  $\hat{\beta}_2 < 0$ ; se sigue que  $x_0$  es negativo, que  $x_0$  es un máximo del cociente y que el valor máximo es:

$$\hat{\ell}(x_0) / \sqrt{\hat{V}(\hat{\ell}(x_0))} > \sqrt{\hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2)}$$

entonces en el intervalo  $(-\infty, x_0)$  la función  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  es creciente

$$\text{y } 0 < -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} < \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} < \sqrt{\hat{\beta}_1^2 / \hat{V}(\hat{\beta}_1) + \hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2)}; \quad (4.14)$$

mientras que en  $(x_0, \infty)$  la función  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  es decreciente y

$$\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} < \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} < \sqrt{\hat{\beta}_1^2 / \hat{V}(\hat{\beta}_1) + \hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2)}$$

de (4.14) se sigue que

$$(-\infty, x_0) \cap \{x \mid \hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)}\} = \emptyset;$$

y por lo tanto  $\{x \mid \hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)}\} \subseteq (x_0, \infty)$ ;

ésto implica que la función  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  es decreciente en el conjunto (4.13); y es lo que se quería probar.

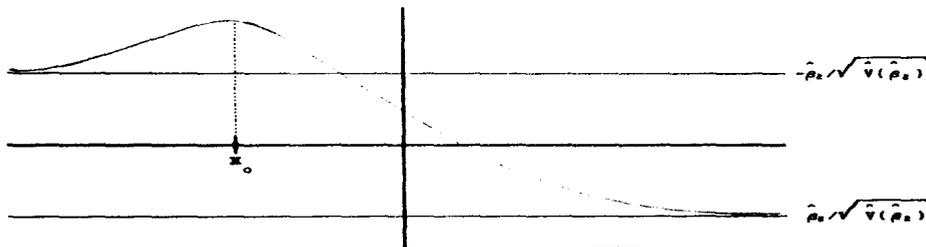


Fig. 4.2 Gráfica de la función  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  cuando  $\hat{\beta}_1 > 0$ .

Caso  $\hat{\beta}_1 < 0$ .

Como  $\hat{\beta}_1 < 0$  y  $\hat{\beta}_2 < 0$ ; se sigue que  $x_0$  es positivo, que  $x_0$  es un mínimo del cociente y que el valor mínimo es:

$$\hat{\ell}(x_0) / \sqrt{\hat{V}(\hat{\ell}(x_0))} < -\sqrt{\hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2)}$$

entonces en el intervalo  $(-\infty, x_0)$  la función  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  es decreciente y

$$-\sqrt{\hat{\beta}_1^2 / \hat{V}(\hat{\beta}_1) + \hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2)} < \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} < -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} \quad (4.15)$$

mientras que en  $(x_0, \infty)$  la función  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  es creciente y

$$-\sqrt{\hat{\beta}_1^2 / \hat{V}(\hat{\beta}_1) + \hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2)} < \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} < \hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} < 0$$

de (4.16) se sigue que

$$(x_0, \infty) \cap \{x \mid \hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)}\} = \emptyset;$$

y por lo tanto  $\{x \mid \hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq -\hat{\beta}_2 / \sqrt{\hat{V}(\hat{\beta}_2)}\} \subseteq (-\infty, x_0)$ , o sea que (4.13) está contenido en el conjunto donde  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  es decreciente; y es lo que se quería probar.

Caso  $\hat{\beta}_1 = 0$ .

Si  $\hat{\beta}_1 = 0$ ,  $\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}$  no tiene puntos crítico, y por i) y ii) en  $-\infty$  tiende a un número positivo, mientras que en  $\infty$  tiende a un número negativo, lo que implica que la función es decreciente en  $\mathbb{R}$ , por lo tanto es decreciente en (4.13). Queda demostrado.

**Corolario 1.**

Si  $\hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2) > \max\{t_c^2, t_{a-c}^2\}$  y  $\hat{\eta}$  tiene un máximo, entonces

$$\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \text{ es decreciente en (4.12).}$$

**Demostración.**

Como (4.12) está contenido en (4.13) se sigue que el corolario es verdadero.

**Corolario 2.**

Si  $\alpha > 0$  es tal que  $P(x_{op} \in R) \geq \alpha$ , entonces para la prueba:  
 $H_0: \beta_1 = \beta_2 = 0$  contra  $H_a$ : al menos una,  $\beta_1$  ó  $\beta_2$  son diferente de cero.  
 se acepta  $H_0$  con una significancia igual a  $\alpha$ .

**Demostración.**

Si  $P(x_{op} \in R) \geq \alpha$ , entonces

$$\{ x \mid -t_{\alpha/2} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha/2} \} = R$$

lo que implica que  $\hat{\ell}^2(x) / \hat{V}(\hat{\ell}) \leq t_{\alpha/2}^2$  para toda  $x \in R \rightarrow$

$$\max_x \{ \hat{\ell}^2(x) / \hat{V}(\hat{\ell}) \} = \hat{\beta}_1^2 / \hat{V}(\hat{\beta}_1) + \hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2) \leq t_{\alpha/2}^2 < 2F_\alpha$$

donde los grados de libertad de la distribución  $t$  son  $n-3$  y los de la distribución  $F$  son  $2$  y  $n-3$  ( como  $\beta_1$  y  $\beta_2$  son independientes entonces  $\hat{\beta}_1^2 / \hat{V}(\hat{\beta}_1) + \hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2) \sim \chi^2$  con  $2$  grados de libertad), por lo tanto se está en la región de aceptación de  $H_0$ , y queda demostrado el corolario.

**Teorema 4.2**

Si  $\hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2) > \max\{t_c^2, t_{\alpha-\epsilon}^2\}$  la longitud del intervalo (4.12) es igual a  $L = x_2 - x_1$ ; con

$$x_2 = \frac{-\hat{\beta}_1 \hat{\beta}_2 + \sqrt{t_c^2 (V(\hat{\beta}_1) \hat{\beta}_2^2 + V(\hat{\beta}_2) \hat{\beta}_1^2) - t_c^4 V(\hat{\beta}_2) V(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_c^2 V(\hat{\beta}_2))}$$

$$x_1 = \frac{-\hat{\beta}_1 \hat{\beta}_2 - \sqrt{t_{\alpha-\epsilon}^2 (V(\hat{\beta}_1) \hat{\beta}_2^2 + V(\hat{\beta}_2) \hat{\beta}_1^2) - t_{\alpha-\epsilon}^4 V(\hat{\beta}_2) V(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_{\alpha-\epsilon}^2 V(\hat{\beta}_2))}$$

**Demostración.**

Note que (4.12) es igual a la unión ajena de

$$\{ x \mid -t_\epsilon \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq 0 \} \cup \{ x \mid 0 \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon} \}$$

(4.17)

y como  $\hat{\ell}(\hat{x}_{op}) = 0$ , entonces  $\hat{x}_{op}$  es el extremo común de los dos intervalos en (4.17); por otro lado, el corolario 1 del lema 4.1 asegura que  $\hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})}$  es una función decreciente en (4.12); por lo que si

$$x_1 \in \{ x \mid 0 < \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon} \}$$

y

$$x_2 \in \{ x \mid -t_\epsilon \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} < 0 \}$$

entonces  $x_1 < \hat{x}_{op} < x_2$

Por esta razón y por el corolario del teorema 4.1 se sigue que

$$\{ x \mid 0 \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon} \} =$$

$$\{ x \mid \hat{\ell}^2(x) / \hat{V}(\hat{\ell}) \leq t_{\alpha-\epsilon}^2 \text{ y } x \leq \hat{x}_{op} \} =$$

$$\left\{ x \mid \frac{-\hat{\beta}_1 \hat{\beta}_2 - \sqrt{t_{\alpha-\epsilon}^2 (\hat{V}(\hat{\beta}_1) \hat{\beta}_2^2 + \hat{V}(\hat{\beta}_2) \hat{\beta}_1^2) - t_{\alpha-\epsilon}^4 \hat{V}(\hat{\beta}_2) \hat{V}(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_{\alpha-\epsilon}^2 \hat{V}(\hat{\beta}_2))} \leq x \leq \hat{x}_{op} \right\}$$

cuando  $\hat{\beta}_2^2 > t_{\alpha-\epsilon}^2 \hat{V}(\hat{\beta}_2)$ .

De igual manera

$$\{ x \mid -t_\epsilon \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq 0 \} =$$

$$\left\{ x \mid \hat{x}_{op} \leq x \leq \frac{-\hat{\beta}_1 \hat{\beta}_2 + \sqrt{t_\epsilon^2 (\hat{V}(\hat{\beta}_1) \hat{\beta}_2^2 + \hat{V}(\hat{\beta}_2) \hat{\beta}_1^2) - t_\epsilon^4 \hat{V}(\hat{\beta}_2) \hat{V}(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_{\alpha-\epsilon}^2 \hat{V}(\hat{\beta}_2))} \right\}$$

cuando  $\hat{\beta}_2^2 > t_\epsilon^2 \hat{V}(\hat{\beta}_2)$ .

Finalmente, se concluye que si  $\hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2) > \max\{t_c^2, t_{\alpha-c}^2\}$

$$\{x \mid -t_c \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq 0\} \cup \{x \mid 0 \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-c}\} =$$

$$\{x \mid -t_c \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-c}\} = (x_1, x_2)$$

con

$$x_2 = \frac{-\hat{\beta}_1 \hat{\beta}_2 + \sqrt{t_c^2 (V(\hat{\beta}_1) \hat{\beta}_2^2 + V(\hat{\beta}_2) \hat{\beta}_1^2) - t_c^4 V(\hat{\beta}_2) V(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_c^2 V(\hat{\beta}_2))}$$

$$x_1 = \frac{-\hat{\beta}_1 \hat{\beta}_2 - \sqrt{t_{\alpha-c}^2 (V(\hat{\beta}_1) \hat{\beta}_2^2 + V(\hat{\beta}_2) \hat{\beta}_1^2) - t_{\alpha-c}^4 V(\hat{\beta}_2) V(\hat{\beta}_1)}}{2(\hat{\beta}_2^2 - t_{\alpha-c}^2 V(\hat{\beta}_2))}$$

que es lo que se quería probar.

**Corolario.**

En la prueba de hipótesis:

$$H_0: \beta_2 = 0 \quad \text{contra} \quad H_a: \beta_2 \neq 0$$

se rechaza la hipótesis nula con un nivel de significancia igual a  $\min(\alpha-c, c)$ , si y sólo si (4.12) es un intervalo acotado. Esto significa que (4.12) puede no estar acotado aun cuando el intervalo usual lo esté.

**Demostración.**

Se rechaza  $H_0$  si y sólo si  $\hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2) > \max\{t_c^2, t_{\alpha-c}^2\}$  y esta es la condición para que (4.12) este acotada. Por otro lado se sabe que la condición para que el intervalo usual este acotado es  $\hat{\beta}_2^2 / \hat{V}(\hat{\beta}_2) > t_{\alpha/2}^2$  (teorema 4.1); y como  $\max\{t_c^2, t_{\alpha-c}^2\} \geq t_{\alpha/2}^2$  para  $0 \leq c \leq \alpha$ , sigue que (4.12) puede no existir aun cuando el intervalo usual esté acotado, y queda demostrado el corolario.

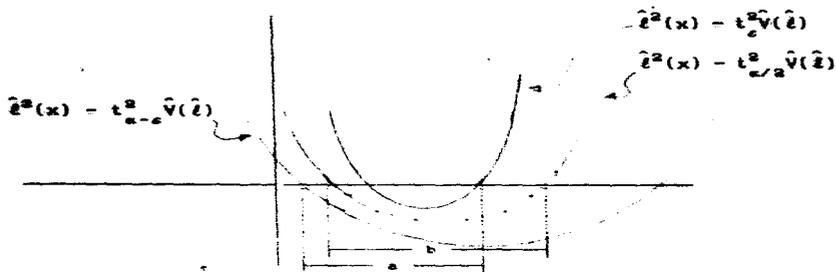


fig. 4.3 a) intervalo (4.12) con  $\hat{\rho}_2 > 0$ , y  $\hat{\rho}_2^2 > \max(t_{\alpha}^2, t_{\alpha-\epsilon}^2) \hat{V}(\hat{\rho}_2)$ ,  
 b) intervalo usual.

#### Teorema 4.3

Los números  $t_{\epsilon}$  y  $t_{\alpha-\epsilon}$  que minimizan la longitud  $L$  son solución de la ecuación

$$g(t) = \frac{(\hat{V}(\hat{\beta}_1) \hat{\beta}_2^4 + \hat{V}(\hat{\beta}_2) \hat{\beta}_1^2 \hat{\beta}_2^2) + t^2 (\hat{V}^2(\hat{\beta}_2) \hat{\beta}_1^2 - \hat{V}(\hat{\beta}_2) \hat{V}(\hat{\beta}_1) \hat{\beta}_2^2)}{2\lambda(\hat{\beta}_2^2 - t^2 \hat{V}(\hat{\beta}_2))^2 \sqrt{(\hat{V}(\hat{\beta}_1) \hat{\beta}_2^2 + \hat{V}(\hat{\beta}_2) \hat{\beta}_1^2) - t^2 \hat{V}(\hat{\beta}_2) \hat{V}(\hat{\beta}_1)}} +$$

$$\frac{2\hat{V}(\hat{\beta}_2) \hat{\beta}_1 \hat{\beta}_2 t}{2\lambda(\hat{\beta}_2^2 - t^2 \hat{V}(\hat{\beta}_2))^2} \quad (4.18)$$

donde  $g(t)$  es la función de densidad  $t$  de student con  $n-3$  grados de libertad y  $\lambda$  de tal manera que se cumpla la restricción  $G(t_{\alpha-\epsilon}) - G(-t_{\epsilon}) = 1 - \alpha$ .

#### Demostración.

Se va a minimizar la longitud  $L$  sujeta a la condición que  $G(t_{\alpha-\epsilon}) - G(-t_{\epsilon}) = 1 - \alpha$ , donde  $G$  es la función de distribución de la

variable aleatoria  $t$  con  $n-3$  grados de libertad.

Para ello se aplica la técnica de los multiplicadores de Lagrange; derivando y resolviendo las ecuaciones.

$$\frac{\partial}{\partial t_c} (L - \lambda(G(t_{\alpha-c}) - G(-t_c) - (1 - \alpha))) = 0.$$

$$\frac{\partial}{\partial t_{\alpha-c}} (L - \lambda(G(t_{\alpha-c}) - G(-t_c) - (1 - \alpha))) = 0.$$

$$\frac{\partial}{\partial \lambda} (L - \lambda(G(t_{\alpha-c}) - G(-t_c) - (1 - \alpha))) = 0.$$

Al derivar la primer ecuación teniendo en cuenta que  $t = -t_c < 0$ , resulta la ecuación (4.18) para  $t \leq 0$ ; al derivar la segunda ecuación considerando que  $t = t_{\alpha-c} > 0$ , resulta la ecuación (4.18) para  $t \geq 0$ ; al derivar la tercer ecuación se obtiene la restricción, esta restricción se manifiesta en el valor de  $\lambda$ . Queda demostrado el teorema.

**Corolario.**

Si el intervalo de longitud mínima existe entonces se cumplen las relaciones de orden:

$$t_c > t_{\alpha-c} \quad \text{si} \quad \hat{\beta}_1 < 0 \quad \text{ó} \quad t_c < t_{\alpha-c} \quad \text{si} \quad \hat{\beta}_1 > 0. \quad (4.19)$$

**Demostración.**

Observe que el segundo miembro de la ecuación (4.18) es la suma de dos funciones de  $t$ : la primer función es par y es positiva, al menos para  $t^2 < \hat{\beta}^2 / \hat{V}(\hat{\beta}_2)$ ; y la segunda es una función impar, y es positiva para los valores de  $t$  de signo contrario a  $\hat{\beta}_1$ . (como  $\hat{\beta}_2 < 0$ , si  $\hat{\beta}_1 t < 0$ ; entonces  $\hat{\beta}_1 \hat{\beta}_2 t > 0$ ).

Por lo tanto la función en el lado derecho de la ecuación (4.18) es mayor en  $-(\text{signo de } \hat{\beta}_1)t$  que en el número  $(\text{signo de } \hat{\beta}_1)t$ , con  $t > 0$ .

Además, esta misma función tiende a infinito cuando  $t^2$  tiende a  $\hat{\beta}^2 / \hat{V}(\hat{\beta}_2)$ .

Por otro lado, la función  $g(t)$ , como se sabe, es una función par, positiva y decrece conforme  $t$  se aleja del cero; por lo tanto, la intersección de las funciones en ambos lados de (4.18) más cerca de cero es en  $t$  de signo contrario a  $\hat{\beta}_1$ ; un valor más cerca de cero tiene menor valor

absoluto y por lo tanto las soluciones de (4.18) cumplen la relación de orden en (4.19). Y es lo que se quería demostrar.

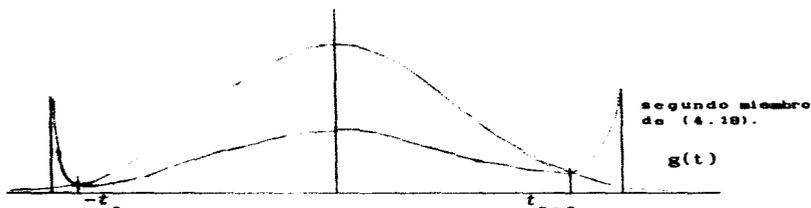


fig. 4.4 solución de  $\varepsilon$  que minimiza la longitud del intervalo de confianza cuando  $\beta < 0$ .

**Definición.**

El intervalo de confianza auxiliar es el conjunto

$$\{x \mid \hat{\beta}_1 x + 2\hat{\beta}_2 x^2 \geq -t_\alpha \sqrt{\hat{V}(\hat{\beta}_1)x^2 + 4x^4 \hat{V}(\hat{\beta}_2)}\} \quad (4.20)$$

donde  $x_{op} \hat{\ell}(x_{op}) = \hat{\beta}_1 x_{op} + 2\hat{\beta}_2 x_{op}^2 \sim N(0, \hat{V}(\hat{\beta}_1)x_{op}^2 + 4x_{op}^4 \hat{V}(\hat{\beta}_2))$

**Lema 4.2**

El intervalo de confianza auxiliar es igual al conjunto

$$\{x \mid \hat{\ell}(x) \geq -t_\alpha \sqrt{\hat{V}(\hat{\ell})} \text{ y } x > 0\} \cup \{x \mid \hat{\ell}(x) \leq t_\alpha \sqrt{\hat{V}(\hat{\ell})} \text{ y } x < 0\} \cup \{0\}$$

**Demstración.**

Si  $x \neq 0$

$$\hat{\beta}_1 x + 2\hat{\beta}_2 x^2 \geq -t_\alpha \sqrt{\hat{V}(\hat{\beta}_1)x^2 + 4x^4 \hat{V}(\hat{\beta}_2)} \text{ es equivalente a a}$$

$$x(\hat{\beta}_1 + 2\hat{\beta}_2 x) \geq -t_\alpha |x| \sqrt{\hat{V}(\hat{\beta}_1) + 4x^2 \hat{V}(\hat{\beta}_2)} \rightarrow$$

$$\frac{x}{|x|} (\hat{\beta}_1 + 2\hat{\beta}_2 x) \approx -t_\alpha \sqrt{\frac{\hat{V}(\hat{\beta}_1)}{V(\hat{\beta}_1)} + 4x^2 \frac{\hat{V}(\hat{\beta}_2)}{V(\hat{\beta}_2)}} = \begin{cases} \hat{\ell}(x) \approx -t_\alpha \sqrt{\hat{V}(\hat{\ell})} & \text{si } x > 0 \\ \hat{\ell}(x) \approx t_\alpha \sqrt{\hat{V}(\hat{\ell})} & \text{si } x < 0 \end{cases}$$

Si  $x = 0$ ,  $x$  pertenece al intervalo de confianza auxiliar.

Combinando estos dos resultados se llega a lo que afirma el lema.

#### Lema 4.3

El intervalo de confianza auxiliar está contenido en el conjunto

$$\{x \mid \hat{\beta}_1 x + 2\hat{\beta}_2 x^2 \geq -t_\alpha [2\sqrt{\hat{V}(\hat{\beta}_2)} x^2 + \hat{V}(\hat{\beta}_1) / 4\sqrt{\hat{V}(\hat{\beta}_2)}] \} \quad (4.21)$$

**Demostración.**

$$\begin{aligned} \hat{V}(\hat{\beta}_1)x^2 + 4x^4\hat{V}(\hat{\beta}_2) &= 4\hat{V}(\hat{\beta}_2)[\hat{V}(\hat{\beta}_1)x^2/4\hat{V}(\hat{\beta}_2) + x^4] \leq \\ 4\hat{V}(\hat{\beta}_2)[\hat{V}^2(\hat{\beta}_1)/64\hat{V}^2(\hat{\beta}_2) + \hat{V}(\hat{\beta}_1)x^2/4\hat{V}(\hat{\beta}_2) + x^4] &= 4\hat{V}(\hat{\beta}_2)[\hat{V}(\hat{\beta}_1)/8\hat{V}(\hat{\beta}_2) + x^2]^2 \end{aligned}$$

de esta manera

$$\begin{aligned} \sqrt{\hat{V}(\hat{\beta}_1)x^2 + 4x^4\hat{V}(\hat{\beta}_2)} &\leq \sqrt{4\hat{V}(\hat{\beta}_2)[\hat{V}(\hat{\beta}_1)/8\hat{V}(\hat{\beta}_2) + x^2]^2} = \\ &2\sqrt{\hat{V}(\hat{\beta}_2)} x^2 + \hat{V}(\hat{\beta}_1) / 4\sqrt{\hat{V}(\hat{\beta}_2)} \end{aligned}$$

por lo que si  $\hat{\ell}(x) > -t_\alpha \sqrt{\hat{V}(\hat{\beta}_1)x^2 + 4x^4\hat{V}(\hat{\beta}_2)}$  entonces necesariamente

$$\hat{\ell}(x) > -t_\alpha [2\sqrt{\hat{V}(\hat{\beta}_2)} x^2 + \hat{V}(\hat{\beta}_1) / 4\sqrt{\hat{V}(\hat{\beta}_2)}]$$

y esto implica que (4.20) esté contenido en (4.21).

Por lo tanto, (4.21) es un intervalo de confianza conservador para  $x_{op}$ .

#### Lema 4.4

El intervalo de confianza conservador, (4.21), está acotado si y sólo si se rechaza la hipótesis nula de la prueba:

$$H_0: \beta_2 = 0 \quad \text{contra} \quad H_1: \beta_2 \neq 0$$

con una significancia igual a  $2\alpha$ , que es equivalente a  $\hat{\beta}_2^2 > t_{\alpha}^2 \hat{V}(\hat{\beta}_2)$ .

**Demstración.**

El intervalo de confianza (4.21) es equivalente a

$$= \{x \mid 2(\hat{\beta}_2 + t_{\alpha} \sqrt{\hat{V}(\hat{\beta}_2)})x^2 + \hat{\beta}_1 x + t_{\alpha} \hat{V}(\hat{\beta}_1) / 4\sqrt{\hat{V}(\hat{\beta}_2)} \geq 0\}$$

este conjunto contiene al 0, por lo que no es vacío, y está acotado si y sólo si el polinomio de segundo grado es cóncavo, y esto se da si y sólo si  $\hat{\beta}_2 + t_{\alpha} \sqrt{\hat{V}(\hat{\beta}_2)}$  es negativo.

Se sabe que  $\hat{\beta}_2$  es negativo, y que  $t_{\alpha} \sqrt{\hat{V}(\hat{\beta}_2)}$  es positivo, por lo tanto para que  $\hat{\beta}_2 + t_{\alpha} \sqrt{\hat{V}(\hat{\beta}_2)}$  sea negativo es necesario que el valor absoluto de  $\hat{\beta}_2$  sea mayor que  $t_{\alpha} \sqrt{\hat{V}(\hat{\beta}_2)}$ .

Por lo tanto, para que (4.21) está acotado se se debe cumplir la desigualdad  $\hat{\beta}_2^2 > t_{\alpha}^2 \hat{V}(\hat{\beta}_2)$ ; que es equivalente a rechazar la hipótesis nula, y es lo que se quería probar.

**Corolario 1.**

Una condición suficiente para que (4.20) esté acotada es que  $\hat{\beta}_2^2 > t_{\alpha}^2 \hat{V}(\hat{\beta}_2)$ ;

**Demstración.**

Como el intervalo en (4.20) está contenido en el intervalo (4.21), Si  $\hat{\beta}_2^2 > t_{\alpha}^2 \hat{V}(\hat{\beta}_2)$  entonces (4.21) está acotado; y necesariamente (4.20) lo está. Queda demostrado el corolario.

**Corolario 2.**

La región de confianza auxiliar puede estar acotada aún cuando la región de confianza usual no lo esté.

**Demstración.**

Como  $t_{\alpha}^2 < t_{\alpha/2}^2 = F_{\alpha}$  entonces se puede cumplir la relación  $\hat{\beta}_2^2 >$

$t_{\alpha}^2 \hat{V}(\hat{\beta}_2)$  aún cuando no se cumpla la relación  $\hat{\beta}_2^2 > t_{\alpha/2}^2 \hat{V}(\hat{\beta}_2)$ , y queda demostrado el corolario.



fig. 4.8 Intervalo de confianza para  $x_{op}$  (resarcado en negro) según la relación (4.17), en el caso que  $\hat{\beta}_2^2 > t_{\alpha}^2 \hat{V}(\hat{\beta}_2)$ .

#### Teorema 4.4

Si el intervalo de longitud mínima está acotado y no contiene al cero, entonces este intervalo es igual a la intersección de un intervalo usual y un intervalo auxiliar.

#### Demostración.

Como el intervalo de longitud mínima está acotado y no contiene al cero, este intervalo está totalmente contenido en los reales positivos o bien en los reales negativos.

El intervalo de longitud mínima con  $100(1 - \alpha)\%$  de confianza es:

$$\{ x \mid -t_{\alpha} \sqrt{\hat{V}(\ell)} \leq \ell(x) \leq t_{\alpha-\epsilon_0} \sqrt{\hat{V}(\ell)} \} =$$

**Caso 1.** el intervalo de longitud mínima está contenido en  $\mathbb{R}^+$ .

El número  $\hat{x}_{op}$  pertenece al intervalo de longitud mínima, entonces  $\hat{x}_{op} = -\hat{\beta}_1 / 2\hat{\beta}_2$  es mayor a cero, y como  $\hat{\beta}_2 < 0$  entonces  $\hat{\beta}_1$  es mayor a cero; y por el corolario 1 del teorema 4.2 se sigue que  $t_{\alpha} < t_{\alpha-\epsilon_0}$ .

El intervalo usual con  $100(1 - \alpha + \epsilon_0)\%$  de confianza es

$$\{ x \mid -t_{\alpha-\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\ell)} \leq t_{\alpha-\epsilon_0} \} = \quad (4.22)$$

$$\{ x \mid -t_{\alpha-\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\ell)} \leq -t_{\epsilon_0} \} \cup \{ x \mid -t_{\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\ell)} \leq t_{\alpha-\epsilon_0} \}$$

Como  $\hat{\ell}(x)/\sqrt{\hat{V}(\ell)}$  es decreciente en (4.22) y el cero no está contenido en el intervalo de longitud mínima, entonces tampoco (4.22) contiene al cero, y por lo tanto (4.22) está totalmente contenido en  $R^+$ .

El intervalo auxiliar con  $100(1 - \epsilon_0)\%$  de confianza es:

$$\{ x \mid -t_{\epsilon_0} \leq x\hat{\ell}(x)/\sqrt{\hat{V}(x\ell)} \} = \quad (4.23)$$

$$\{ x \mid -t_{\epsilon_0}\sqrt{\hat{V}(\ell)} \leq \hat{\ell}(x) \text{ y } x > 0 \} \cup \{ x \mid t_{\epsilon_0}\sqrt{\hat{V}(\ell)} \geq \hat{\ell}(x) \text{ y } x < 0 \} \cup \{0\}$$

Así la intersección de los intervalos (4.22) y (4.23) es igual a

$$\{ x \mid -t_{\alpha-\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\ell)} \leq t_{\alpha-\epsilon_0} \} \cap \{ x \mid -t_{\epsilon_0}\sqrt{\hat{V}(\ell)} \leq \hat{\ell}(x) \text{ y } x > 0 \} =$$

$$\{ x \mid -t_{\epsilon_0}\sqrt{\hat{V}(\ell)} \leq \hat{\ell}(x) \leq t_{\alpha-\epsilon_0}\sqrt{\hat{V}(\ell)} \}$$

y esto es lo que se quería demostrar.

Caso 2. el intervalo de longitud mínima está contenido en  $R^-$ .

El número  $\hat{x}_{op}$  pertenece al intervalo de longitud mínima, entonces  $\hat{x}_{op} = -\hat{\beta}_1/2\hat{\beta}_2$  es menor a cero, y como  $\hat{\beta}_2 < 0$  entonces  $\hat{\beta}_1$  es menor a cero; y por el corolario 1 del teorema 4.2 se sigue que  $t_{\epsilon_0} > t_{\alpha-\epsilon_0}$ .

El intervalo usual con  $100(1 - \epsilon_0)\%$  de confianza es

$$\{ x \mid -t_{\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\ell)} \leq t_{\epsilon_0} \} \quad (4.24)$$

$$\{ x \mid -t_{\alpha-\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon_0} \} \cup \{ x \mid t_{\alpha-\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})} \leq t_{\epsilon_0} \}$$

Como  $\hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})}$  es decreciente en (4.24) y el cero no está contenido en el intervalo de longitud mínima, entonces tampoco (4.24) contiene al cero, y se sigue que (4.24) está totalmente contenido en los reales negativos.

El intervalo auxiliar con  $100(1 - (\alpha - \epsilon_0))\%$  de confianza es:

$$\{ x \mid -t_{\alpha-\epsilon_0} \leq x\hat{\ell}(x)/\sqrt{\hat{V}(x\hat{\ell})} \} = \quad (4.25)$$

$$\{x \mid -t_{\alpha-\epsilon_0} \sqrt{\hat{V}(\hat{\ell})} \leq \hat{\ell}(x) \text{ y } x > 0\} \cup \{x \mid t_{\alpha-\epsilon_0} \sqrt{\hat{V}(\hat{\ell})} \geq \hat{\ell}(x) \text{ y } x < 0\} \cup \{0\} =$$

Así la intersección de los intervalos (4.24) y (4.25),

$$\{x \mid -t_{\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})} \leq t_{\epsilon_0}\} \cap \{x \mid t_{\alpha-\epsilon_0} \sqrt{\hat{V}(\hat{\ell})} \geq \hat{\ell}(x) \text{ y } x < 0\} =$$

$$\{ x \mid -t_{\epsilon_0} \sqrt{\hat{V}(\hat{\ell})} \leq \hat{\ell}(x) \leq t_{\alpha-\epsilon_0} \sqrt{\hat{V}(\hat{\ell})} \}$$

y ésto es lo que se quería demostrar.

#### Teorema 4.5

Si el intervalo de longitud mínima está acotado y contiene al cero, entonces el intervalo auxiliar de igual confianza está contenido en él.

#### Demostración.

Como el intervalo de longitud mínima contiene al cero, entonces

$$\{ x \mid -t_{\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon_0} \}$$

es igual a la unión de los conjuntos no vacíos,

$$\{x \mid -t_{\epsilon_0} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})} \leq \hat{\ell}(0)/\sqrt{\hat{V}(\hat{\ell})}\} \cup \{x \mid \hat{\ell}(0)/\sqrt{\hat{V}(\hat{\ell})} \leq \hat{\ell}(x)/\sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon_0}\}$$

Por otro lado, el intervalo auxiliar con  $100(1 - \alpha)\%$  de confianza es:

$$\{x \mid -t_{\alpha} \leq x \hat{\ell}(x) / \sqrt{\hat{V}(x\hat{\ell})}\} =$$

$$\{x \mid -t_{\alpha} \sqrt{\hat{V}(\hat{\ell})} \leq \hat{\ell}(x) \text{ y } x > 0\} \cup \{x \mid t_{\alpha} \sqrt{\hat{V}(\hat{\ell})} \geq \hat{\ell}(x) \text{ y } x < 0\} \cup \{0\}$$

Como  $t_{\alpha} < \min\{t_{\alpha-\epsilon}, t_{\epsilon}\}$  entonces

$$\{x \mid -t_{\alpha} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}, \text{ y } x > 0\} \subseteq \{x \mid -t_{\alpha} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq \hat{\ell}(0) / \sqrt{\hat{V}(\hat{\ell})}\}$$

y

$$\{x \mid t_{\alpha} \geq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}, \text{ y } x < 0\} \cup \{0\} \subseteq \{x \mid \hat{\ell}(0) / \sqrt{\hat{V}(\hat{\ell})} \leq \hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})} \leq t_{\alpha-\epsilon}\}$$

de estas dos últimas relaciones se sigue la veracidad del teorema.

Los teoremas 4.4 y 4.5 muestran la importancia que tiene el intervalo de confianza auxiliar.

#### 4.1.2 INTERVALO DE CONFIANZA PARA $\eta_{op}$

En esta sección se discute la construcción de dos regiones de confianza para la respuesta máxima: la primera es una región conjunta para  $(x_{op}, \eta_{op})$ , la segunda es una región para el valor individual de  $\eta_{op}$ .

La región de confianza conjunta para  $(x_{op}, \eta_{op})$  es el conjunto

$$\{(x, y) \mid |\hat{\ell}(x) / \sqrt{\hat{V}(\hat{\ell})}| \leq t_{r(a)}, |\hat{k}(x) - y| / \sqrt{\hat{V}(\hat{k})} \leq \sqrt{\frac{t_{r(a)}^2 - \hat{\ell}^2(x) / \hat{V}(\hat{\ell})}{r(a)}}\} \quad (4.26)$$

donde

$$\hat{k}(x_{op}) = \hat{\beta}_0 + \hat{\beta}_1 x_{op} + \hat{\beta}_2 x_{op}^2 - \frac{1}{2} x_{op} \hat{\ell} = \hat{\beta}_0 + \frac{1}{2} \hat{\beta}_1 x_{op}$$

es normal con media  $\eta_{op}$  y varianza

$$V(\hat{\beta}_0) + \text{cov}(\hat{\beta}_0, \hat{\beta}_1) x_{op} + \frac{1}{4} V(\hat{\beta}_1) x_{op}^2 = V(\hat{\beta}_0) + \frac{1}{4} V(\hat{\beta}_1) x_{op}^2$$

y  $\hat{\ell}(x)$  es igual que antes.

El número  $\gamma(\alpha)$  es tal que la significancia de la región (4.26) es  $\alpha$ .

La manera más simple de calcular el valor de  $\gamma(\alpha)$  es cuando las variables  $\hat{\ell}(x)/\sqrt{V(\hat{\ell})}$  y  $(\hat{k}(x) - \eta(x))/\sqrt{V(\hat{k})}$  son independientes. Sin embargo, es difícil caracterizar los diseños que hacen posible esto.

Por suerte existe otra condición, mediante la cual se calcula el valor de  $t_{\gamma(\alpha)}$  y fácilmente se pueden caracterizar los diseños que la cumplen.

#### Teorema 4.6

Si  $X, Y \sim N(0, 1)$  y  $Z \sim \chi_m^2$ ;  $X, Y$  y  $Z$  independientes, entonces la función de densidad conjunta de las variables  $U = X/\sqrt{Z/m}$  y  $V = Y/\sqrt{Z/m}$  es

$$f_{UV}(u, v) = \frac{1}{2\pi(u^2/m + v^2/m + 1)^{(m+2)/2}} \quad (4.27)$$

#### Demostración.

Si  $X, Y$  y  $Z$  son independientes, su función de densidad conjunta es

$$f_{XYZ}(x, y, z) = \frac{z^{m/2-1} e^{-(x^2 + y^2 + z)/2}}{2\pi\Gamma(m/2)2^{m/2}} \quad \text{para } -\infty \leq x, y \leq \infty \text{ y } 0 \leq z \leq \infty.$$

para obtener la densidad conjunta de  $(u, v)$  se utiliza la transformación

$$U = X/\sqrt{Z/m}, \quad V = Y/\sqrt{Z/m} \quad \text{y} \quad W = Z$$

y su transformación inversa

$$X = UV\sqrt{W/m}, \quad Y = V\sqrt{W/m} \quad \text{y} \quad Z = W,$$

en la que el jacobiano es

$$|J| = W/m$$

entonces la f. d. de (U, V, W) es

$$f_{UVW}(u, v, w) = \frac{w^{(m+2)/2-1} e^{-(u^2 w/m + v^2 w/m + w)/2}}{2\pi \Gamma(m/2) 2^{m/2} m} =$$

la función de densidad conjunta de (U, V) es

$$\int_0^\infty f_{UVW}(u, v, w) dw = \frac{1}{2\pi \Gamma(m/2) 2^{m/2} m} \int_0^\infty w^{(m+2)/2-1} e^{-w(u^2/m + v^2/m + 1)/2} dw$$

esta integral se resuelve por el método de sustitución de variables de la forma

$$y = w(u^2/m + v^2/m + 1)/2; \quad dw = 2dy \cdot (u^2/m + v^2/m + 1)$$

$$= \frac{2^{(m+2)/2}}{2\pi \Gamma(m/2) 2^{m/2} m (u^2/m + v^2/m + 1)^{(m+2)/2}} \int_0^\infty y^{(m+2)/2-1} e^{-y} dy =$$

$$\frac{\Gamma((m+2)/2)}{\pi \Gamma(m/2) m (u^2/m + v^2/m + 1)^{(m+2)/2}}$$

y como  $\Gamma(a+1) = a\Gamma(a)$ , finalmente se tiene que

$$\frac{1}{2\pi(u^2/m + v^2/m + 1)^{(m+2)/2}}$$

y es lo que se quería probar.

**Teorema 4.7**

Si la función de densidad conjunta de

$$\hat{z}(x)/\sqrt{V(\hat{z})} \text{ y } (\hat{k}(x) - \eta(x))/\sqrt{V(\hat{k})}$$

es  $f(u, v) = \frac{1}{2\pi(u^2/k + v^2/k + 1)^{(k+2)/2}}$  para  $(u, v) \in \mathbb{R}^2$

entonces el valor de  $t_{\alpha}(x)$  en (4.26) es la solución positiva de la ecuación

$$t = \sqrt{[(1/\alpha)^{2/k} - 1]k} \quad (4.28)$$

**Demostración.**

La probabilidad de estar dentro de la región (4.26) es

$$\frac{1}{2\pi} \int_{-t}^t \int_{-\sqrt{t^2 - v^2}}^{\sqrt{t^2 - v^2}} \frac{dudv}{(u^2/k + v^2/k + 1)^{(k+2)/2}} =$$

y considerando la transformación  $u = r \cos\theta$  y  $v = r \sin\theta$ ; el jacobiano de la transformación es  $|r|$ , y la integral de arriba es igual a

$$\frac{1}{2\pi} \int_{-t}^t \int_0^{2\pi} \frac{|r| dr d\theta}{(r^2/k + 1)^{(k+2)/2}}$$

$$= \frac{1}{(r^2/k + 1)^{k/2}} \Big|_{r=0}^t = 1 - \frac{1}{(t^2/k + 1)^{k/2}} = 1 - \alpha \Rightarrow t = \sqrt{[(1/\alpha)^{2/k} - 1]k}$$

Queda demostrado el teorema.

**Teorema 4.8**

Si el diseño es simétrico, y cumple la relación

$$n\sum_1^4 - 5(\sum_1^2)^2 = 0.$$

entonces la función de densidad conjunta de las v.a.'s  $(\hat{k}(x) - \eta(x))/\sqrt{V(\hat{k})}$  y  $\hat{\ell}(x)/\sqrt{V(\hat{\ell})}$  es

$$f(u, v) = \frac{1}{2\pi(u^2/(n-3) + v^2/(n-3) + 1)^{(n-1)/2}} \quad \text{para } (u, v) \in \mathbb{R}^2$$

**Demostración.**

Por el teorema 4.6, si  $(\hat{k}(x) - \eta(x))/\sqrt{V(\hat{k})}$  y  $\hat{\ell}(x)/\sqrt{V(\hat{\ell})} \sim N(0, 1)$  y  $\hat{\sigma}^2/\sigma^2 \sim \chi_{n-3}^2$ , son independientes, entonces la función de densidad conjunta de  $(\hat{k}(x) - \eta(x))/\sqrt{V(\hat{k})}$  y de  $\hat{\ell}(x)/\sqrt{V(\hat{\ell})}$  es la que afirma el teorema.

Ya se sabe que  $\hat{\sigma}^2/\sigma^2$  es independiente de  $(\hat{k}(x) - \eta(x))/\sqrt{V(\hat{k})}$  y de  $\hat{\ell}(x)/\sqrt{V(\hat{\ell})}$ , por lo que basta probar que  $\hat{\ell}(x)$  y  $\hat{k}(x)$  son independientes para que las tres variables lo sean.

$\hat{\ell}(x)$  y  $\hat{k}(x)$  son independientes si y sólo si  $\text{cov}(\hat{\ell}(x), \hat{k}(x)) = 0$ . (es una propiedad de las v.a. normales)

Así, se necesita determinar cuando  $\text{cov}(\hat{\ell}(x_{op}), \hat{k}(x_{op}))$  es igual a cero.

$$\begin{aligned} \text{cov}(\hat{\ell}(x_{op}), \hat{k}(x_{op})) &= \\ \text{cov}(\hat{\beta}_0, \hat{\beta}_1) + x_{op}(2\text{cov}(\hat{\beta}_0, \hat{\beta}_2) + V(\hat{\beta}_1)/2) + x_{op}^2 \text{cov}(\hat{\beta}_1, \hat{\beta}_2) \end{aligned}$$

y como  $\text{cov}(\hat{\beta}_0, \hat{\beta}_1)$  y  $\text{cov}(\hat{\beta}_1, \hat{\beta}_2)$  son cero,

$$\text{cov}(\hat{\ell}(x_{op}), \hat{k}(x_{op})) = x_{op}(2\text{cov}(\hat{\beta}_0, \hat{\beta}_2) + V(\hat{\beta}_1)/2)$$

suponiendo que  $x_{op}$  es diferente de cero, la independencia de ambas variables se da si  $2\text{cov}(\hat{\beta}_0, \hat{\beta}_2) + V(\hat{\beta}_1)/2$  es igual a 0.

De la relación (4.4) se sigue que

$$\text{cov}(\hat{\beta}_0, \hat{\beta}_2) = -\sigma^2 \sum x_i^2 / (n \sum x_i^4 - (\sum x_i^2)^2) \quad \text{y} \quad V(\hat{\beta}_1) = \sigma^2 / \sum x_i^2$$

por lo que sustituyendo estos términos

$$2\text{cov}(\hat{\beta}_0, \hat{\beta}_2) + V(\hat{\beta}_1)/2 = -2\sigma^2 \sum x_1^2 / (n \sum x_1^4 - (\sum x_1^2)^2) + \sigma^2 / 2 \sum x_1^2 = 0.$$

de esta última ecuación se sigue que

$$2\text{cov}(\hat{\beta}_0, \hat{\beta}_2) + V(\hat{\beta}_1)/2 = 0 \quad \Leftrightarrow \quad n \sum x_1^4 - 5(\sum x_1^2)^2 = 0$$

así entonces, si el diseño es tal que  $n \sum x_1^4 - 5(\sum x_1^2)^2 = 0$ , las variables aleatorias  $\hat{\ell}(x_{op})$ ,  $\hat{k}(x_{op})$  son independientes, y por el teorema 4.6 se sigue que la densidad de conjunta de  $(\hat{k}(x) - \eta(x))/\sqrt{V(\hat{k})}$  y  $\hat{\ell}(x)/\sqrt{V(\hat{\ell})}$  es la dada en (4.27) y queda demostrado el teorema.

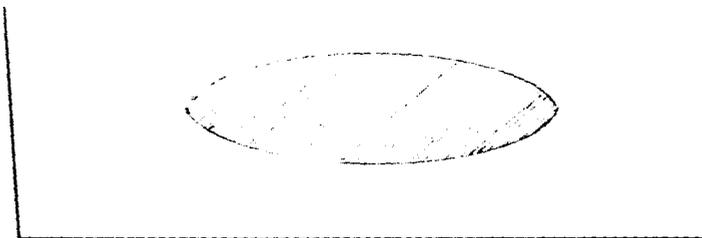


fig. 4.7 Región de confianza conjunta para  $x_{op}$  y para  $\eta_{op}$ .

De todo esto, se sigue que si el diseño cumple la condición

$$n \sum x_1^4 - 5(\sum x_1^2)^2 = 0$$

la región de confianza conjunta es

$$\{(x, y) \mid |\hat{\ell}(x)/\sqrt{V(\hat{\ell})}| \leq t_{\gamma(\alpha)}, \quad |\hat{k}(x) - y|/\sqrt{V(\hat{k})} \leq \sqrt{t_{\gamma(\alpha)}^2 - \ell^2(x)/V(\hat{\ell})}\}$$

$$\text{con } t_{\gamma(\alpha)} = \sqrt{((1/\alpha)^{2/(n-3)} - 1)(n-3)}$$

Por ejemplo el diseño formado por los puntos  $x_1 = 1$ ,  $x_2 = -1$ ,  $x_3 = \sqrt{3}$ ,  $x_4 = \sqrt{3}$  y 12 repeticiones en el centro, es simétrico con respecto al cero y cumple la condición del teorema 4.8.

Intervalo de confianza para  $\eta_{op}$ . Para encontrar el intervalo de confianza individual para  $\eta_{op}$  se va a construir una "banda" o cinturón de confianza para  $\eta(x)$  con  $x$  en  $\mathbb{R}$ .

**Lema 4.5 (Noble, 1969)**

Si  $A$  es una matriz de  $m \times m$  y  $B$  una matriz de  $m \times k$  de rango  $m$ , entonces  $A$  es positiva semidefinida, si y sólo si  $B^T A B$  es positiva semidefinida.

**Lema 4.6 (Scheffé)**

El número  $k$ , que cumple la relación

$$P(|\hat{\eta}(x) - \eta(x)| \leq (k \hat{V}(\hat{\eta}(x)))^{1/2} \text{ para todo } x \in \mathbb{R}) = 1 - \alpha$$

es igual a  $k = 3F_{3, n-3}$ .

**Demostración.**

Se sabe que  $\hat{\eta}(x) = w^T \hat{\beta} = w^T \beta + w^T (W^T W)^{-1} W^T \epsilon$ ; que  $\eta(x) = w^T \beta$ ; y que  $\hat{V}(\hat{\eta}(x)) = w^T (W^T W)^{-1} w$  por lo tanto  $\hat{\eta}(x) - \eta(x) = w^T (W^T W)^{-1} W^T \epsilon$ .

$$y$$

$$(\hat{\eta}(x) - \eta(x))^2 \leq k \hat{V}(\hat{\eta}(x)) \Leftrightarrow k w^T (W^T W)^{-1} w \epsilon^2 - w^T (W^T W)^{-1} W^T \epsilon \epsilon^T W^T (W^T W)^{-1} w \geq 0$$

donde  $w^T = (1, x, x^2)$ .

Una condición suficiente para que esta desigualdad se cumpla para toda  $x$  real, es que la matriz

$$k s^2 (W^T W)^{-1} - (W^T W)^{-1} W^T \epsilon \epsilon^T W^T (W^T W)^{-1}$$

sea positiva semidefinida, y por el lema 4.5, esto ocurre si y sólo si

$$k s^2 W (W^T W)^{-1} W^T - W (W^T W)^{-1} W^T \epsilon \epsilon^T W^T (W^T W)^{-1} W^T \quad (4.29)$$

sea positiva semidefinida.

Para que la matriz en (4.29) sea positiva semidefinida es necesario que todos sus valores propios sean mayores o iguales a cero.

Los valores propios de (4.29) son:

i)  $ks^2 - \epsilon^T W(W^T W)^{-1} W^T \epsilon$  con vector propio  $W(W^T W)^{-1} W^T \epsilon$ , porque

$$\begin{aligned} (ks^2 W(W^T W)^{-1} W^T - W(W^T W)^{-1} W^T \epsilon \epsilon^T W(W^T W)^{-1} W^T) W(W^T W)^{-1} W^T \epsilon = \\ (ks^2 - \epsilon^T W(W^T W)^{-1} W^T \epsilon) W(W^T W)^{-1} W^T \epsilon \end{aligned}$$

ii) 0 con vectores propios iguales a  $(I - W(W^T W)^{-1} W^T) y$  para todo  $y \in \mathbb{R}^n$ .

iii)  $ks^2$  con vectores propios  $W(W^T W)^{-1} W^T y$  con  $y$  ortogonal a  $W(W^T W)^{-1} W^T \epsilon$

Como  $\epsilon^T W(W^T W)^{-1} W^T \epsilon \geq 0$ , entonces una condición necesaria y suficiente para que i) e iii) sean mayores que o iguales a cero es que  $ks^2 - \epsilon^T W(W^T W)^{-1} W^T \epsilon \geq 0$ .

Por otro lado, se puede ver que la forma cuadrática  $\epsilon^T W(W^T W)^{-1} W^T \epsilon / \sigma^2$  se distribuye como una  $\chi^2_3$ , ya que la matriz  $(W(W^T W)^{-1} W^T / \sigma^2)(\sigma^2 I)$  es idempotente y el rango de  $W(W^T W)^{-1} W^T$  es igual a 3.

Además, se sabe que  $(n-3)s^2 / \sigma^2 = SCE / \sigma^2$  es una variable aleatoria  $\chi^2$  con  $n-3$  grados de libertad.

Por lo que finalmente se tiene que

$$\epsilon^T W(W^T W)^{-1} W^T \epsilon / 3s^2$$

tiene una distribución F con 3 y  $n-3$  grados de libertad.

De todo esto se sigue que  $ks^2 \geq \epsilon^T W(W^T W)^{-1} W^T \epsilon$  implica que

$$P(|\hat{\eta}(x) - \eta(x)| \leq (k\hat{V}(\hat{\eta}(x)))^{1/2} \text{ para todo } x \in \mathbb{R}) = 1 - \alpha$$

donde  $k = 3F_{3, n-3}$ .

Del lema 4.6 se obtiene una región para  $\eta(x)$  con un  $100(1 - \alpha)\%$  de confianza, dada por

$$(\hat{\eta}(x) - \sqrt{kV(\hat{\eta}(x))}, \hat{\eta}(x) + \sqrt{kV(\hat{\eta}(x))}) \quad (4.30)$$

En esta expresión  $\hat{\eta}(x)$  es una función cuadrática con un máximo y  $\sqrt{kV(\hat{\eta}(x))}$  es la raíz cuadrada de un polinomio de cuarto grado que es positivo para todo  $x$  real.

**Lema 4.7**

Si  $x_1, x_3$  son los puntos donde

$$\hat{\eta}(x) - \sqrt{kV(\hat{\eta}(x))} \quad \text{y} \quad \hat{\eta}(x) + \sqrt{kV(\hat{\eta}(x))}$$

alcanzan su valor máximo, respectivamente, entonces el intervalo

$$(\hat{\eta}(x_1) - \sqrt{kV(\hat{\eta}(x_1))}, \hat{\eta}(x_3) + \sqrt{kV(\hat{\eta}(x_3))}) \quad (4.31)$$

contiene a  $\eta_{op}$  con  $(1 - \alpha)100\%$  de confianza.

**Demstración.**

Por el lema 4.6, se sabe que

$$(\hat{\eta}(x) - \sqrt{kV(\hat{\eta}(x))}, \hat{\eta}(x) + \sqrt{kV(\hat{\eta}(x))})$$

es un intervalo de confianza de  $\eta(x)$  para todo  $x$  real.

Así con una confianza da  $(1 - \alpha)100\%$  de confianza

$$\hat{\eta}(x) - \sqrt{kV(\hat{\eta}(x))} \leq \eta(x) \leq \hat{\eta}(x) + \sqrt{kV(\hat{\eta}(x))} \quad \forall x \in \mathbb{R}$$

De esta relación se sigue que si  $x_1, x_2$  y  $x_3$ , son los puntos donde  $\hat{\eta}(x) - \sqrt{kV(\hat{\eta}(x))}$ ,  $\eta(x)$  y  $\hat{\eta}(x) + \sqrt{kV(\hat{\eta}(x))}$  alcanzan el valor máximo, respectivamente, entonces

$$\hat{\eta}(x_1) - \sqrt{k\hat{V}(\hat{\eta}(x_1))} \leq \eta(x_1) \leq \eta(x_2) \leq \hat{\eta}(x_2) + \sqrt{k\hat{V}(\hat{\eta}(x_2))} \leq \hat{\eta}(x_3) + \sqrt{k\hat{V}(\hat{\eta}(x_3))}$$

y esto implica que:

$$\max_x \{ \hat{\eta}(x) - \sqrt{k\hat{V}(\hat{\eta}(x))} \} \leq \eta_{op} \leq \max_x \{ \hat{\eta}(x) + \sqrt{k\hat{V}(\hat{\eta}(x))} \}$$

con al menos  $(1 - \alpha)100\%$  de confianza. Queda demostrado.

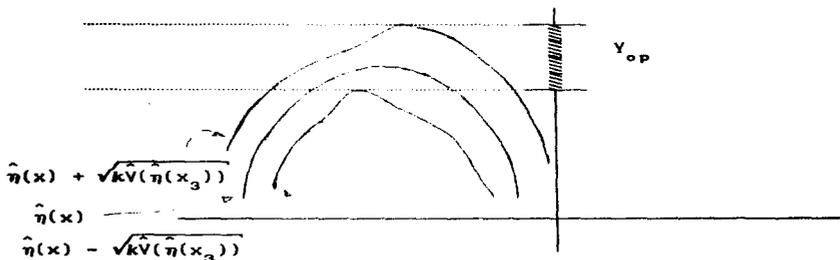


fig. 4.8 En esta grafica se muestra la banda de confianza para  $\eta(x)$ , así como el intervalo de confianza para  $\eta_{op}$ .

#### Teorema 4.8

Si el diseño es simétrico respecto al cero, una condición suficiente para que (4.31) esté acotado es que  $\hat{\beta}_2 + \sqrt{k\hat{V}(\hat{\beta}_2)} < 0$ .

#### Demostración.

Si el diseño es simétrico respecto al cero, entonces  $\hat{\beta}_1$  es independiente de  $\hat{\beta}_0$  y de  $\hat{\beta}_2$ , y

$$\hat{V}(\hat{\eta}(x)) = \hat{V}(\hat{\beta}_0) + (\hat{V}(\hat{\beta}_1) + \hat{Cov}(\hat{\beta}_0, \hat{\beta}_2))x^2 + \hat{V}(\hat{\beta}_2)x^4$$

completando cuadrados se sigue que

$$\hat{V}(\hat{\eta}(x)) = \hat{V}(\hat{\beta}_2)|x^2 + \hat{Cov}(\hat{\beta}_0, \hat{\beta}_2)/2\hat{V}(\hat{\beta}_2)|^2 + \hat{V}(\hat{\beta}_0) - \hat{Cov}^2(\hat{\beta}_0, \hat{\beta}_2)/4\hat{V}(\hat{\beta}_2)$$

y de aquí se sigue que

$$\sqrt{k\hat{V}(\hat{\eta}(x))} \leq \sqrt{k\hat{V}(\hat{\beta}_2)}x^2 + |\hat{\text{Cov}}(\hat{\beta}_0, \hat{\beta}_2)/2\sqrt{\hat{V}(\hat{\beta}_2)}| + |\hat{V}(\hat{\beta}_0) - \hat{\text{Cov}}^2(\hat{\beta}_0, \hat{\beta}_2)/4\hat{V}(\hat{\beta}_2)|$$

Si se considera que

$$c = |\hat{\text{Cov}}(\hat{\beta}_0, \hat{\beta}_2)/2\sqrt{\hat{V}(\hat{\beta}_2)}| + |\hat{V}(\hat{\beta}_0) - \hat{\text{Cov}}^2(\hat{\beta}_0, \hat{\beta}_2)/4\hat{V}(\hat{\beta}_2)|$$

entonces

$$\begin{aligned} \hat{\eta}(x) - \sqrt{k\hat{V}(\hat{\beta}_2)}x^2 - c &\leq \hat{\eta}(x) - \sqrt{k\hat{V}(\hat{\eta}(x))} \leq \hat{\eta}(x) + \sqrt{k\hat{V}(\hat{\eta}(x))} \\ &\leq \hat{\eta}(x) + \sqrt{k\hat{V}(\hat{\beta}_2)}x^2 + c. \end{aligned}$$

El coeficiente del término cuadrático en la función  $\hat{\eta}(x) - \sqrt{k\hat{V}(\hat{\beta}_2)}x^2 + c$  es negativo, lo que implica que tiene un máximo; la función  $\hat{\eta}(x) + \sqrt{k\hat{V}(\hat{\beta}_2)}x^2 + c$  tiene un máximo en  $\mathbb{R}$  si y sólo si,  $\hat{\beta}_2 + \sqrt{k\hat{V}(\hat{\beta}_2)}$ , el coeficiente del término cuadrático, es negativo.

Entonces  $\hat{\beta}_2 + \sqrt{k\hat{V}(\hat{\beta}_2)} < 0$ ,  $\hat{\eta}(x) + \sqrt{k\hat{V}(\hat{\beta}_2)}x^2 + c$  tiene un máximo. Queda demostrado el teorema.

**Corolario.**

Si  $\hat{\beta}_2 + \sqrt{k\hat{V}(\hat{\beta}_2)} < 0$ , entonces se rechaza la hipótesis nula

$$H_0: \beta_2 = 0.$$

con una significancia igual a  $\alpha$ .

**Demostración.**

El número  $k = 3F_{3, n-3} > F_{1, n-3} = t_{n-3}^2$ ; y la condición para rechazar  $H_0$  es que  $\hat{\beta}_2 < \sqrt{t_{n-3} \hat{V}(\hat{\beta}_2)}$ .

Entonces, si  $\hat{\beta}_2 < \sqrt{k\hat{V}(\hat{\beta}_2)}$  implica que  $\hat{\beta}_2 < \sqrt{t_{n-3} \hat{V}(\hat{\beta}_2)}$  y se rechaza la hipótesis nula.

## 4.2 REGIÓN DE CONFIANZA PARA $x_{op}$ Y PARA $\eta_{op}$

### CASO M FACTORES.

En esta sección se presenta la generalización del caso anterior, con desarrollos semejantes.

#### 4.2.1 REGIÓN DE CONFIANZA PARA $x_{op}$ .

En esta parte se presentan dos regiones de confianza para el vector  $x_{op}$ , la primera es la región usual y la segunda es la región auxiliar. Posteriormente, siguiendo un proceso semejante al caso de un factor, se define otra región con la intersección de una región usual y una auxiliar.

**Consideraciones generales:** Cuando la respuesta depende de  $m$  factores, las observaciones se pueden aproximar, en un entorno del óptimo, con el modelo:

$$Y = w^T \beta + \varepsilon = \beta_0 + b^T x + x^T B x + \varepsilon$$

$$\text{con } w^T = (1, x_1, \dots, x_m, x_1^2, x_1 x_2, \dots, x_1 x_m, x_2^2, \dots, x_m^2)$$
$$\text{y } \beta = (\beta_0, \beta_1, \dots, \beta_m, \beta_{11}, \beta_{12}, \dots, \beta_{1m}, \beta_{22}, \dots, \beta_{mm})$$

así:  $\beta_0$  es un número real,  $b$  un vector y  $B$  una matriz negativa definida, cuyos elementos son parte del vector desconocido  $\beta$ ; y  $\varepsilon \sim N(0, \sigma^2)$  son variables aleatorias independientes. De esta manera  $\eta$  presenta un máximo único en  $x_{op}$  si, y sólo si,  $B$  es una matriz negativa definida y

$$x_{op} = -\frac{1}{2} B^{-1} b.$$

El estimador de mínimos cuadrados de  $\beta$  es,  $\hat{\beta} \sim N(\beta, V)$ ; con

$$\hat{V} = \hat{\sigma}^2(W^T W)^{-1}$$

$$W = \begin{bmatrix} 1 & x_{11} & \dots & x_{1m} & x_{11}^2 & x_{11}x_{12} & \dots & x_{11}x_{1m} & x_{12}^2 & \dots & x_{1m}^2 \\ 1 & x_{21} & \dots & x_{2m} & x_{21}^2 & x_{21}x_{22} & \dots & x_{21}x_{2m} & x_{22}^2 & \dots & x_{2m}^2 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{nm} & x_{n1}^2 & x_{n1}x_{n2} & \dots & x_{n1}x_{nm} & x_{n2}^2 & \dots & x_{nm}^2 \end{bmatrix}$$

El estimador insesgado de  $\sigma^2$  es  $\hat{\sigma}^2 = \text{SCE}/(n-(m+1)(m+2)/2)$  y  $\hat{V} = \hat{\sigma}^2(W^T W)^{-1}$

Además se tiene que  $\text{SCE}/\sigma^2 \sim \chi^2_{(n-(m+1)(m+2)/2)}$ , es independiente de  $\hat{\beta}$ . (Searle, 1971), por lo que también es independiente de  $\hat{\beta}_0$ , de  $\hat{b}$  y de  $\hat{B}$ .

En el caso de un factor, se consideraron únicamente los diseños simétricos respecto al cero. Una condición equivalente a la simetría con respecto al cero, en el caso de  $m$  factores, es que el diseño cumpla las condiciones:

$$\sum_i x_{ij} = 0; \sum_i x_{ij}x_{ik} = 0; \sum_i x_{ij}^3 = 0; \sum_i x_{ij}^2x_{ik} = 0; \sum_i x_{ij}x_{ik}x_{il} = 0; \quad (4.32)$$

$$\sum_i x_{ij}^3x_{ik} = 0; \sum_i x_{ij}^2x_{ik}x_{il} = 0; \sum_i x_{ij}x_{ik}x_{il}x_{ip} = 0. \quad (4.33)$$

y

$$\sum_i x_{ij}^2 = \sum_i x_{ik}^2; \sum_i x_{ij}^4 = \sum_i x_{ik}^4; \sum_i x_{ij}^2x_{ik}^2 = \sum_i x_{il}^2x_{ip}^2 \quad (4.34)$$

para todo  $j = k; p = j; p = k$  y  $l = p$ .

Si el diseño cumple (4.32) resulta que el vector  $\hat{b}$  es independiente de  $\hat{\beta}_0$  y de la matriz  $\hat{B}$ ,

Si el diseño además cumple (4.33) resulta que las únicas covarianzas diferentes de cero son:  $\text{Cov}(\hat{\beta}_0, \hat{\beta}_{11})$  y  $\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$ .

Si el diseño cumple también (4.34), entonces  $\text{Cov}(\hat{\beta}_0, \hat{\beta}_{11})$ ,  $\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$ ,  $V(\hat{\beta}_1)$ ,  $V(\hat{\beta}_{11})$  y  $V(\hat{\beta}_{1j})$  no dependen de  $i$  ni de  $j$ .

Por ejemplo, el diseño compuesto cumple las tres condiciones: (4.32), (4.33) y (4.34).

Al sustituir las relaciones (4.32), (4.33) y (4.33) en  $W^T W$  e invertir esta matriz, se tiene

$$V(\hat{\beta}_0) = \frac{(m-1)\sum x_{1j}^2 x_{kj}^2 + \sum x_{1j}^4}{n(m-1)\sum x_{1j}^2 x_{kj}^2 - m(\sum x_{1j}^2)^2 + n\sum x_{1j}^4} \sigma^2 \quad (4.35)$$

$$\text{Cov}(\hat{\beta}_0, \hat{\beta}_{11}) = \frac{\sum x_{1j}^2}{n(m-1)\sum x_{1j}^2 x_{kj}^2 - m(\sum x_{1j}^2)^2 + n\sum x_{1j}^4} \sigma^2 \quad (4.36)$$

$$V(\hat{\beta}_1) = \frac{\sigma^2}{\sum x_{1j}^2} \quad (4.37)$$

$$V(\hat{\beta}_{1j}) = \frac{\sigma^2}{\sum x_{1j}^2 x_{kj}^2} \quad (4.38)$$

$$V(\hat{\beta}_{11}) = \frac{-n(m-2)\sum x_{1j}^2 x_{kj}^2 + (m-1)(\sum x_{1j}^2)^2 - n\sum x_{1j}^4}{(n(m-1)\sum x_{1j}^2 x_{kj}^2 - m(\sum x_{1j}^2)^2 + n\sum x_{1j}^4)(y-z)} \sigma^2 \quad (4.39)$$

$$\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj}) = \frac{n\sum x_{1j}^2 x_{kj}^2 - (\sum x_{1j}^2)^2}{(n(m-1)\sum x_{1j}^2 x_{kj}^2 - m(\sum x_{1j}^2)^2 + n\sum x_{1j}^4)(\sum x_{1j}^2 x_{kj}^2 - \sum x_{1j}^4)} \sigma^2 \quad (4.40)$$

las demás covarianzas son igual a cero.

Si el estimador de mínimos cuadrados de  $\eta$  es

$$\hat{\eta}(z) = \hat{\beta}_0 + \hat{b}^T z + z^T \hat{b} z$$

con  $\hat{b}$  negativa definida, el vector aleatorio  $\hat{\ell}(z) = \hat{b} + 2\hat{b}z$  (4.41) es normal con media cero y matriz de varianza-covarianza igual a

$$V = V(\hat{\ell}(z_{op})) = E(\hat{\ell}(z_{op})\hat{\ell}^T(z_{op}))$$

Por lo tanto,  $\hat{\ell}^T(\mathbf{x})V^{-1}\hat{\ell}(\mathbf{x}) \sim \chi^2$  con  $m$  grados de libertad; y, consecuentemente,  $\hat{\ell}^T(\mathbf{x})V^{-1}\hat{\ell}(\mathbf{x})/m = \hat{\ell}^T(\mathbf{x})V^{-1}\hat{\ell}(\mathbf{x})/m(\hat{\sigma}^2/\sigma^2) \sim F$  con  $m$  y  $n - (m+2)(m+1)/2$  grados de libertad.

#### Lema 4.8

Cada uno de los elementos de la matriz  $V(\hat{\ell}(\mathbf{x}))$  es una forma cuadrática en  $\mathbf{x}$ .

#### Demstración.

Por definición  $V(\hat{\ell}(\mathbf{x})) = E(\hat{\ell}(\mathbf{x}) - \ell(\mathbf{x}))^T(\hat{\ell}(\mathbf{x}) - \ell(\mathbf{x})) =$

$$E[(\hat{\mathbf{b}} - \mathbf{b}) + 2(\hat{\mathbf{B}} - \mathbf{B})\mathbf{x}]^T[(\hat{\mathbf{b}} - \mathbf{b}) + 2(\hat{\mathbf{B}} - \mathbf{B})\mathbf{x}] =$$

$$\begin{aligned} E(\hat{\mathbf{b}} - \mathbf{b})^T(\hat{\mathbf{b}} - \mathbf{b}) + 2E(\hat{\mathbf{b}} - \mathbf{b})^T(\hat{\mathbf{B}} - \mathbf{B})\mathbf{x} + 2E\mathbf{x}^T(\hat{\mathbf{B}} - \mathbf{B})(\hat{\mathbf{b}} - \mathbf{b}) + 4E\mathbf{x}^T(\hat{\mathbf{B}} - \mathbf{B})^2\mathbf{x} \\ = V(\hat{\mathbf{b}}) + 2\text{Cov}(\hat{\mathbf{b}}, \hat{\mathbf{B}}\mathbf{x}) + 2\text{Cov}(\hat{\mathbf{B}}\mathbf{x}, \hat{\mathbf{b}}) + 4V(\hat{\mathbf{B}}\mathbf{x}) \end{aligned}$$

Ahora considere la forma particionada de  $\hat{\mathbf{B}}$ :

$$\hat{\mathbf{B}} = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \dots & \hat{\beta}_{1m}/2 \\ \hat{\beta}_{12}/2 & \hat{\beta}_{22} & \dots & \hat{\beta}_{2m}/2 \\ \dots & \dots & \dots & \dots \\ \hat{\beta}_{1m}/2 & \hat{\beta}_{2m}/2 & \dots & \hat{\beta}_{mm} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{b}}_1^T \\ \vdots \\ \hat{\mathbf{b}}_m^T \end{bmatrix}$$

y el vector  $\hat{\mathbf{b}}$

$$\hat{\mathbf{b}}^T = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_m)$$

y como  $\hat{\mathbf{B}}\mathbf{x} = \begin{bmatrix} \hat{\mathbf{b}}_1^T\mathbf{x} \\ \vdots \\ \hat{\mathbf{b}}_m^T\mathbf{x} \end{bmatrix}$ , entonces el elemento  $ij$  de  $V(\hat{\ell}(\mathbf{x}))$  resulta ser:

$$v_{ij} = \text{Cov}(\hat{\beta}_i, \hat{\beta}_j) + 2\text{Cov}(\hat{\beta}_i, \hat{\mathbf{b}}_j^T)\mathbf{x} + 2\text{Cov}(\hat{\beta}_j, \hat{\mathbf{b}}_i^T)\mathbf{x} + 4\mathbf{x}^T\text{Cov}(\hat{\mathbf{b}}_i, \hat{\mathbf{b}}_j^T)\mathbf{x} \quad (4.42)$$

que, como se ve, es una forma cuadrática en  $\mathbf{x}$ . Queda demostrado el lema.

**Corolario**

La expresión

$$\hat{\ell}(\mathbf{x}_{op}) \hat{V}^{-1} \hat{\ell}(\mathbf{x}_{op}) = (\hat{\mathbf{b}} + 2\hat{\mathbf{B}}\mathbf{x}_{op})^T \hat{V}^{-1} (\hat{\mathbf{b}} + 2\hat{\mathbf{B}}\mathbf{x}_{op})$$

es una forma cuadrática donde, tanto la matriz como el vector, están en función de  $\mathbf{x}_{op}$ .

**Lema 4.8**

La matriz  $V(\hat{\ell}(\mathbf{x}))$  es diagonal, independientemente del vector  $\mathbf{x}$ , si y sólo si,

$$\text{Cov}(\hat{\beta}_i, \hat{\beta}_j) = 0; \text{ para } i \neq j \quad (4.43)$$

$$2\text{Cov}(\hat{\beta}_i, \hat{\beta}_{jj}) = -\text{Cov}(\hat{\beta}_j, \hat{\beta}_{ij}) \text{ y } \text{Cov}(\hat{\beta}_i, \hat{\beta}_{jk}) = 0, \text{ para } i \neq j \neq k \neq i. \quad (4.44)$$

$$\begin{aligned} \text{Cov}(\hat{\beta}_{ik}, \hat{\beta}_{jk}) &= \text{Cov}(\hat{\beta}_{ik}, \hat{\beta}_{ji}) = \text{Cov}(\hat{\beta}_{ii}, \hat{\beta}_{jj}) = \text{Cov}(\hat{\beta}_{ii}, \hat{\beta}_{ij}) = 0 \text{ y} \\ V(\hat{\beta}_{ij}) &= -4\text{Cov}(\hat{\beta}_{ii}, \hat{\beta}_{jj}); \text{ para } i, j, k \text{ y } l \text{ diferentes.} \end{aligned} \quad (4.45)$$

**Demostración.**

$V(\hat{\ell}(\mathbf{x}))$  es diagonal independientemente del vector  $\mathbf{x}$ , si  $v_{ij}$  en (4.42) es igual a cero para  $i \neq j$ ; y esto se da, si y sólo si

$$\text{Cov}(\hat{\beta}_i, \hat{\beta}_j) = 0; (\text{Cov}(\hat{\beta}_i, \hat{\mathbf{b}}_j^T) + \text{Cov}(\hat{\beta}_j, \hat{\mathbf{b}}_i^T))\mathbf{x} = 0 \text{ y } \mathbf{x}^T \text{Cov}(\hat{\mathbf{b}}_i, \hat{\mathbf{b}}_j^T)\mathbf{x} = 0.$$

De aquí inmediatamente se cumple (4.43).

Para ver que se cumple (4.44) observe que

$(\text{Cov}(\hat{\beta}_i, \hat{\mathbf{b}}_j^T) + \text{Cov}(\hat{\beta}_j, \hat{\mathbf{b}}_i^T))\mathbf{x} = 0$  para toda  $\mathbf{x}$  implica que

$$\text{Cov}(\hat{\beta}_i, \hat{\mathbf{b}}_j^T) + \text{Cov}(\hat{\beta}_j, \hat{\mathbf{b}}_i^T) = 0$$

$$y \ 2(\text{Cov}(\hat{\beta}_1, \hat{b}_j^T) + \text{Cov}(\hat{\beta}_j, \hat{b}_1^T)) = \begin{bmatrix} \text{cov}(\hat{\beta}_1, \hat{\beta}_{j1}) + \text{cov}(\hat{\beta}_j, \hat{\beta}_{11}) \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_{j2}) + \text{cov}(\hat{\beta}_j, \hat{\beta}_{12}) \\ \vdots \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_{j1}) + 2\text{cov}(\hat{\beta}_j, \hat{\beta}_{11}) \\ \vdots \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_{jm}) + \text{cov}(\hat{\beta}_j, \hat{\beta}_{1m}) \end{bmatrix}$$

de aquí se obtienen las relaciones  $\text{cov}(\hat{\beta}_1, \hat{\beta}_{jk}) + \text{cov}(\hat{\beta}_j, \hat{\beta}_{1k}) = 0$ ; y  $\text{cov}(\hat{\beta}_1, \hat{\beta}_{j1}) + 2\text{cov}(\hat{\beta}_j, \hat{\beta}_{11}) = 0$ ; con  $i, j$  y  $k$  diferentes.

En estas relaciones, los subíndices  $i, j$  y  $k$  se pueden permutar. Además se sabe que  $\hat{B}$  es simétrica, por lo que  $\hat{\beta}_{ij} = \hat{\beta}_{ji}$ ; así se obtiene el sistema de ecuaciones,

$$\begin{aligned} \text{cov}(\hat{\beta}_1, \hat{\beta}_{jk}) + \text{cov}(\hat{\beta}_j, \hat{\beta}_{1k}) &= 0 \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_{jk}) + \text{cov}(\hat{\beta}_k, \hat{\beta}_{1j}) &= 0 \\ \text{cov}(\hat{\beta}_k, \hat{\beta}_{1j}) + \text{cov}(\hat{\beta}_j, \hat{\beta}_{1k}) &= 0 \end{aligned}$$

La solución de este sistema de ecuaciones es el vector 0, por lo que se concluye que para  $i, j$  y  $k$  diferentes,  $\text{cov}(\hat{\beta}_i, \hat{\beta}_{jk}) = 0$ .

Con esto se prueba (4.44).

Para probar (4.45) observe que  $\mathbf{x}^T \text{Cov}(\hat{b}_1, \hat{b}_j^T) \mathbf{x} = 0$  implica que

$$\mathbf{x}^T (\text{Cov}(\hat{b}_1, \hat{b}_j^T) + \text{Cov}(\hat{b}_j, \hat{b}_1^T)) \mathbf{x} = 0;$$

y como los valores propios de una matriz simétrica son reales, entonces la única manera que ésto ocurra es que  $\text{Cov}(\hat{b}_1, \hat{b}_j^T) + \text{Cov}(\hat{b}_j, \hat{b}_1^T) = 0$

Ahora observe que

$$4\hat{\mathbf{b}}_1 \hat{\mathbf{b}}_j^T = \begin{bmatrix} \hat{\beta}_{11} \hat{\beta}_{j1} & \hat{\beta}_{11} \hat{\beta}_{j2} & \hat{\beta}_{11} \hat{\beta}_{j3} & \dots & 2\hat{\beta}_{11} \hat{\beta}_{jj} & \dots & \hat{\beta}_{11} \hat{\beta}_{j1} & \dots & \hat{\beta}_{11} \hat{\beta}_{jm} \\ \hat{\beta}_{12} \hat{\beta}_{j1} & \hat{\beta}_{12} \hat{\beta}_{j2} & \hat{\beta}_{12} \hat{\beta}_{j3} & \dots & 2\hat{\beta}_{12} \hat{\beta}_{jj} & \dots & \hat{\beta}_{12} \hat{\beta}_{j1} & \dots & \hat{\beta}_{12} \hat{\beta}_{jm} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 2\hat{\beta}_{11} \hat{\beta}_{j1} & 2\hat{\beta}_{11} \hat{\beta}_{j2} & 2\hat{\beta}_{11} \hat{\beta}_{j3} & \dots & 4\hat{\beta}_{11} \hat{\beta}_{jj} & \dots & 2\hat{\beta}_{11} \hat{\beta}_{j1} & \dots & 2\hat{\beta}_{11} \hat{\beta}_{jm} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \hat{\beta}_{1m} \hat{\beta}_{j1} & \hat{\beta}_{1m} \hat{\beta}_{j2} & \hat{\beta}_{1m} \hat{\beta}_{j3} & \dots & 2\hat{\beta}_{1m} \hat{\beta}_{jj} & \dots & \hat{\beta}_{1m} \hat{\beta}_{j1} & \dots & \hat{\beta}_{1m} \hat{\beta}_{jm} \end{bmatrix}$$

Así  $\text{Cov}(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_j^T) + \text{Cov}(\hat{\mathbf{b}}_j, \hat{\mathbf{b}}_1^T) = 0$  si y sólo si

- i)  $\text{Cov}(\hat{\beta}_{1k}, \hat{\beta}_{jk}) = 0$  y  $\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{j1}) = 0$ , con  $i, j$  y  $k$  diferentes.
- ii)  $\text{Cov}(\hat{\beta}_{1k}, \hat{\beta}_{j1}) + \text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jk}) = 0$ , con  $i, j, k$ , y  $l$  diferentes.
- iii)  $2\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jk}) + \text{Cov}(\hat{\beta}_{1k}, \hat{\beta}_{1j}) = 0$ , con  $i, j$  y  $k$  diferentes.
- iv) y  $4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj}) + V(\hat{\beta}_{1j}) = 0$ , con  $i = j$ .

en todas estas relaciones los índices se pueden permutar.

De la primera ecuación en i) y de iii) se sigue que

$$\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jk}) = 0.$$

Por otro lado, intercambiando los índices en ii) se tiene el sistema de ecuaciones

$$\text{Cov}(\hat{\beta}_{1k}, \hat{\beta}_{j1}) + \text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{11}) = 0.$$

$$\text{Cov}(\hat{\beta}_{1k}, \hat{\beta}_{j1}) + \text{Cov}(\hat{\beta}_{j1}, \hat{\beta}_{k1}) = 0.$$

$$\text{Cov}(\hat{\beta}_{j1}, \hat{\beta}_{k1}) + \text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{11}) = 0.$$

cuya solución es el vector 0, con lo que se prueba que  $\text{Cov}(\hat{\beta}_{1k}, \hat{\beta}_{j1}) = 0$ , para  $i, j, k$  y  $l$  diferentes. Queda probado (4.39), y el lema.

**Lema 4.10**

$V$  es múltiplo de la matriz identidad ( $v_{11} = v_{jj}$  para  $i \neq j$ ) si, además de ser una matriz diagonal, se cumple:

$$V(\hat{\beta}_i) = V(\hat{\beta}_j); \text{ para } i \neq j. \quad (4.46)$$

$$\text{Cov}(\hat{\beta}_j, \hat{\beta}_{1j}) = 2\text{Cov}(\hat{\beta}_1, \hat{\beta}_{11}) \text{ y } \text{Cov}(\hat{\beta}_1, \hat{\beta}_{1k}) = \text{Cov}(\hat{\beta}_j, \hat{\beta}_{jk});$$

$$\text{para } i, j \text{ y } k \text{ diferentes.} \quad (4.47)$$

$$V(\hat{\beta}_{1k}) = V(\hat{\beta}_{jk}) \text{ y } 4V(\hat{\beta}_{11}) = V(\hat{\beta}_{1j}); \text{ para } i \neq j. \quad (4.48)$$

**Demostración.**

Los elementos de la diagonal de  $V(\hat{\ell}(x))$  se encuentran con la relación (4.42) en el caso  $i = j$ :

$$v_{11} = \text{Cov}(\hat{\beta}_1, \hat{\beta}_1) + 4\text{Cov}(\hat{\beta}_1, \hat{b}_1^T)x + 4x^T\text{Cov}(\hat{b}_1, \hat{b}_1^T)x$$

por lo tanto  $v_{11} = v_{jj}$  si y sólo si  $\text{Cov}(\hat{\beta}_1, \hat{\beta}_1) = \text{Cov}(\hat{\beta}_j, \hat{\beta}_j)$ ;

$$\text{Cov}(\hat{\beta}_1, \hat{b}_1^T) = \text{Cov}(\hat{\beta}_j, \hat{b}_j^T) \text{ y } \text{Cov}(\hat{b}_1, \hat{b}_1^T) = \text{Cov}(\hat{b}_j, \hat{b}_j^T)$$

Y como  $\text{Cov}(\hat{\beta}_1, \hat{\beta}_1) = V(\hat{\beta}_1)$ , queda probado (4.46)

Para probar (4.47) observe que

$$\text{Cov}(\hat{\beta}_1, \hat{b}_1^T) - \text{Cov}(\hat{\beta}_j, \hat{b}_j^T) = \begin{bmatrix} \text{cov}(\hat{\beta}_1, \hat{\beta}_{11}) - \text{cov}(\hat{\beta}_j, \hat{\beta}_{j1}) \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_{12}) - \text{cov}(\hat{\beta}_j, \hat{\beta}_{j2}) \\ \vdots \\ 2\text{cov}(\hat{\beta}_1, \hat{\beta}_{11}) - \text{cov}(\hat{\beta}_j, \hat{\beta}_{j1}) \\ \vdots \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_{1m}) - \text{cov}(\hat{\beta}_j, \hat{\beta}_{jm}) \end{bmatrix} = 0$$

De aquí se sigue que para  $i, j$  y  $k$  diferentes  
 $\text{Cov}(\hat{\beta}_i, \hat{\beta}_{ik}) = \text{Cov}(\hat{\beta}_j, \hat{\beta}_{jk})$  y  $\text{Cov}(\hat{\beta}_j, \hat{\beta}_{ij}) = 2\text{Cov}(\hat{\beta}_i, \hat{\beta}_{ii})$

que es lo que se quería probar.

Para probar (4.48) observe que

$$4\hat{\mathbf{b}}_i \hat{\mathbf{b}}_i^T = \begin{bmatrix} \hat{\beta}_{i1}\hat{\beta}_{i1} & \hat{\beta}_{i1}\hat{\beta}_{i2} & \hat{\beta}_{i1}\hat{\beta}_{i3} & \dots & 2\hat{\beta}_{i1}\hat{\beta}_{ii} & \dots & \hat{\beta}_{i1}\hat{\beta}_{i1} & \dots & \hat{\beta}_{i1}\hat{\beta}_{im} \\ \hat{\beta}_{i2}\hat{\beta}_{i1} & \hat{\beta}_{i2}\hat{\beta}_{i2} & \hat{\beta}_{i2}\hat{\beta}_{i3} & \dots & 2\hat{\beta}_{i2}\hat{\beta}_{ii} & \dots & \hat{\beta}_{i2}\hat{\beta}_{i1} & \dots & \hat{\beta}_{i2}\hat{\beta}_{im} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 2\hat{\beta}_{i1}\hat{\beta}_{i1} & 2\hat{\beta}_{i1}\hat{\beta}_{i2} & 2\hat{\beta}_{i1}\hat{\beta}_{i3} & \dots & 4\hat{\beta}_{i1}\hat{\beta}_{ii} & \dots & 2\hat{\beta}_{i1}\hat{\beta}_{i1} & \dots & 2\hat{\beta}_{i1}\hat{\beta}_{im} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \hat{\beta}_{im}\hat{\beta}_{i1} & \hat{\beta}_{im}\hat{\beta}_{i2} & \hat{\beta}_{im}\hat{\beta}_{i3} & \dots & 2\hat{\beta}_{im}\hat{\beta}_{ii} & \dots & \hat{\beta}_{im}\hat{\beta}_{i1} & \dots & \hat{\beta}_{im}\hat{\beta}_{im} \end{bmatrix}$$

Entonces  $\text{Cov}(\hat{\mathbf{b}}_i, \hat{\mathbf{b}}_i^T) = \text{Cov}(\hat{\mathbf{b}}_j, \hat{\mathbf{b}}_j^T)$  si y sólo si

$$V(\hat{\beta}_{ik}) = V(\hat{\beta}_{jk}); \text{Cov}(\hat{\beta}_{ik}, \hat{\beta}_{ii}) = \text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{jj});$$

$$2\text{Cov}(\hat{\beta}_{ii}, \hat{\beta}_{ik}) = \text{Cov}(\hat{\beta}_{ij}, \hat{\beta}_{jk}) \text{ y } 4V(\hat{\beta}_{ii}) = V(\hat{\beta}_{ij})$$

por el lema 4.9, se sabe que  $\text{Cov}(\hat{\beta}_{ik}, \hat{\beta}_{ii}) = \text{Cov}(\hat{\beta}_{jk}, \hat{\beta}_{jj}) = 0$ , y también  $2\text{Cov}(\hat{\beta}_{ii}, \hat{\beta}_{ik}) = \text{Cov}(\hat{\beta}_{ij}, \hat{\beta}_{jk}) = 0$ , y queda demostrado (4.48) y el lema.

#### Teorema 4.10

No existe un diseño para el cual  $V(\hat{\ell}(\mathbf{x}))$  sea un múltiplo de la matriz identidad.

#### Demostración.

Para que  $V(\hat{\ell}(\mathbf{x}))$  sea múltiplo de la identidad es necesario que se cumpla simultáneamente (4.45) y (4.48); y esto ocurre si

$$V(\hat{\beta}_{ij}) = -4\text{Cov}(\hat{\beta}_{ii}, \hat{\beta}_{jj}) \text{ y } 4V(\hat{\beta}_{ii}) = V(\hat{\beta}_{ij})$$

lo que implica que  $V(\hat{\beta}_{ii}) = -\text{Cov}(\hat{\beta}_{ii}, \hat{\beta}_{jj})$

por lo tanto, la matriz de varianza covarianza de  $(\hat{\beta}_{11}, \hat{\beta}_{22}, \dots, \hat{\beta}_{mm})$  es de la forma:

$$\begin{bmatrix} c & -c & -c & -c & \dots & -c \\ -c & c & -c & -c & \dots & -c \\ -c & -c & c & -c & \dots & -c \\ \vdots & & & & & \\ -c & -c & -c & -c & \dots & c \end{bmatrix} \quad (4.49)$$

donde  $c = V(\hat{\beta}_{11}) > 0$ .

Se sabe que toda matriz de varianza covarianza es positiva definida, por lo tanto se espera que (4.49) sea positiva definida; pero,  $-c(m-1)$  es un valor propio de (4.49), con vector propio  $(1, 1, \dots, 1)^T$ , y como  $-c(m-1) < 0$  para  $m \geq 2$ , se tiene una contradicción ya que un valor propio de una matriz positiva definida no puede ser negativo. El supuesto que (4.45) y (4.48) se cumplen simultaneamente es falso, y el teorema queda probado.

**Corolario.**

Si  $V(\hat{\beta}_i)$ , y  $\text{Cov}(\hat{\beta}_i, \hat{\beta}_j)$  no dependen de  $i$  ni de  $j$ , entonces

$$V(\hat{\beta}_{11}) > -(m-1)\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj}) \quad \text{y} \quad V(\hat{\beta}_{11}) > \text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$$

**Demostración.**

Si  $V(\hat{\beta}_{11})$  y  $\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$  son constantes para  $i$  y  $j$ , entonces la matriz de varianza covarianza de  $(\hat{\beta}_{11}, \hat{\beta}_{22}, \dots, \hat{\beta}_{mm})$  es de la forma:

$$\begin{bmatrix} c & d & d & d & \dots & d \\ d & c & d & d & \dots & d \\ d & d & c & d & \dots & d \\ \vdots & & & & & \\ d & d & d & d & \dots & c \end{bmatrix}$$

donde  $c = V(\hat{\beta}_{11})$  y  $d = \text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$ .

los valores propios de esta matriz son:  $c + (m-1)d$  y  $c - d$ , y ambos deben ser positivos, por lo que se sigue que  $c > -(m-1)d$  y  $c > d$ . Queda demostrado el corolario.

**Lema 4.11**

Si el diseño cumple las propiedades de simetría: (4.32), (4.33) y (4.34); entonces  $V(\hat{\beta}_{1j}) = -4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$  si y sólo si  $V(\hat{\ell}(x))$  es diagonal.

**Demostración.**

Si se cumplen las propiedades (4.32), (4.33) y (4.34) entonces las covarianzas en (4.43), (4.44) y en parte de (4.45) son cero, por lo tanto se cumplen todas estas relaciones, excepto

$$V(\hat{\beta}_{1j}) = -4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj});$$

esta relación se cumple si y sólo si  $V(\hat{\ell}(x))$  es diagonal. Queda demostrado el lema.

**Lema 4.12**

Si el diseño cumple las propiedades de simetría (4.32), (4.33) y (4.34) entonces  $V(\hat{\ell}(x))$  es diagonal si y sólo si los puntos del diseño cumplen la relación

$$\begin{aligned} m(\sum x_{1j}^2)^2 \sum x_{1j}^4 - (m+4) \sum x_{1j}^2 x_{kj}^2 (\sum x_{1j}^2)^2 - n(\sum x_{1j}^4)^2 - (m-2) \sum x_{1j}^2 x_{kj}^2 \sum x_{1j}^4 + \\ (m+3)n(\sum x_{1j}^2 x_{kj}^2)^2 = 0 \end{aligned} \quad (4.99)$$

**Demostración.**

El lema 4.10 afirma que si las hipótesis del lema se cumplen entonces  $V(\hat{\beta}_{1j}) = -4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$  si y sólo si  $V(\hat{\ell}(x))$  es diagonal.

Si en la ecuación  $V(\hat{\beta}_{1j}) = -4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$  se sustituye los términos en (4.38) y en (4.40) y se desarrolla los términos en la ecuación resultante se llega a lo que afirma el teorema.

**Lema 4.13**

Si el diseño cumple las propiedades de simetría: (4.32), (4.33) y (4.34); y además  $V(\hat{\beta}_{1j}) = -4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$  entonces

$$v_{11} = V(\hat{\beta}_{11}) + 4e_2 x^T x \quad \text{donde } e_2 = V(\hat{\beta}_{11})$$

**Demostración.**

Por el lema 4.10, se sabe que al cumplirse las hipótesis del lema entonces la matriz  $V(\hat{\xi}(x))$  es diagonal, y por (4.42) se sigue que  $v_{11} = V(\hat{\beta}_1) + 4x^T \text{Cov}(\hat{b}_1, \hat{b}_1^T)x$

$$\text{con } x^T \text{Cov}(\hat{b}_1, \hat{b}_1^T)x = x^T \begin{pmatrix} V(\hat{\beta}_{11})/4 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & V(\hat{\beta}_{12})/4 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & V(\hat{\beta}_{11}) & 0 & \dots & 0 \\ \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & V(\hat{\beta}_{1m})/4 \end{pmatrix} x;$$

los valores de  $V(\hat{\beta}_{1k})$  y de  $V(\hat{\beta}_{11})$  con  $l = k$ , no dependen ni de  $i$ , ni de  $k$ ; y por el corolario del teorema 4.10  $V(\hat{\beta}_{11}) > -(m-1)\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{22}) = (m-1)V(\hat{\beta}_{12})/4$ , entonces  $x^T \text{Cov}(\hat{b}_1, \hat{b}_1^T)x \leq \epsilon_2 x^T x$  donde  $\epsilon_2 = V(\hat{\beta}_{11})$ , y por consecuencia  $v_{11} \leq V(\hat{\beta}_1) + 4\epsilon_2 x^T x$ . Queda demostrado el lema.

**Definición.**

La región de confianza usual para  $\mu_{op}$  es:

$$\{ x \mid \hat{\xi}^T(x) \hat{V}^{-1} \hat{\xi}(x) \leq F_{\alpha} \} \quad (4.51)$$

con  $(1-\alpha)100\%$  de confianza.

El teorema 4.10, afirma que no es posible tener una expresión analítica para la frontera de (4.51) considerando que  $\hat{V}$  es diagonal, lo más que se puede tener es una frontera conservadora para la misma, ésto se da en el siguiente lema.

**Lema 4.10**

Si el diseño cumple las propiedades de simetría: (4.32), (4.33) y (4.34); y además  $V(\hat{\beta}_{1j}) = -4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$  entonces

$$\{ x \mid \hat{\xi}^T(x) \hat{V}^{-1} \hat{\xi}(x) \leq F_{\alpha} \} \subseteq \{ x \mid \hat{\xi}^T(x) \hat{\xi}(x) \leq F_{\alpha} (V(\hat{\beta}_1) + 4\epsilon_2 x^T x) \}$$

**Demostración.**

Si el diseño cumple las propiedades de simetría dadas y  $V(\hat{\beta}_{1j}) = -4\text{Cov}(\hat{\beta}_{11}, \hat{\beta}_{jj})$ , entonces la matriz  $\hat{V} = V\epsilon^2/\sigma^2$  es diagonal y

$$\hat{\ell}^T(\mathbf{x})\hat{V}^{-1}\hat{\ell}(\mathbf{x}) = \sum \hat{\ell}_1^2(\mathbf{x})/\hat{v}_{11} \geq \sum \hat{\ell}_1^2(\mathbf{x})/(\hat{V}(\hat{\beta}_1) + 4\psi_2\mathbf{x}^T\mathbf{x}) =$$

$$\hat{\ell}^T(\mathbf{x})\hat{\ell}(\mathbf{x})/(\hat{V}(\hat{\beta}_1) + 4\psi_2\mathbf{x}^T\mathbf{x})$$

así, si  $\mathbf{x} \in \{ \mathbf{x} \mid \hat{\ell}^T(\mathbf{x})\hat{V}^{-1}\hat{\ell}(\mathbf{x}) \leq F_{\alpha, m} \}$  necesariamente

$$\mathbf{x} \in \{ \mathbf{x} \mid \hat{\ell}^T(\mathbf{x})\hat{\ell}(\mathbf{x})/(\hat{V}(\hat{\beta}_1) + 4\psi_2\mathbf{x}^T\mathbf{x}) \leq F_{\alpha, m} \} \rightarrow$$

$$\mathbf{x} \in \{ \mathbf{x} \mid \hat{\ell}^T(\mathbf{x})\hat{\ell}(\mathbf{x}) \leq F_{\alpha, m}(\hat{V}(\hat{\beta}_1) + 4\psi_2\mathbf{x}^T\mathbf{x}) \} \quad (4.52)$$

y queda demostrado el lema.

**Lema 4.15** (Noble, 1969)

El conjunto  $\{ \mathbf{x} \mid (\mathbf{x} + \mathbf{a})^T \mathbf{A} (\mathbf{x} + \mathbf{a}) \leq k \}$  está acotado, si y sólo si  $\mathbf{A}$  es positiva definida.

**Corolario.**

El conjunto  $\{ \mathbf{x} \mid (\mathbf{x} + \mathbf{a})^T \mathbf{A} (\mathbf{x} + \mathbf{a}) \geq k \}$  está acotado, si y sólo si  $\mathbf{A}$  es negativa definida.

**Demostración.**

Para probar el corolario, basta ver que si  $\mathbf{A}$  es negativa definida, entonces  $-\mathbf{A}$  es positiva definida.

**Lema 4.16** (Noble, 1969)

Si  $\lambda$  es un valor propio de  $\mathbf{A}$  y  $\mathbf{x}$  es su respectivo vector propio, entonces  $\lambda^k + c$  es un valor propio de  $\mathbf{A}^k + c\mathbf{I}$  con  $\mathbf{x}$  como vector propio.

**Lema 4.17**

Si el diseño cumple las propiedades de simetría: (4.32), (4.33) y (4.34); y  $\mathbf{V}$  es diagonal, una condición suficiente para que (4.51) esté acotada es que  $\lambda_1^2 > mF\psi_2$  para todo  $\lambda_1$  valor propio de  $\hat{\mathbf{B}}$ .

**Demostración.**

Como el conjunto  $\{ \mathbf{x} \mid \hat{\ell}^T(\mathbf{x})\hat{\ell}(\mathbf{x}) \leq mF(\hat{\psi}_1 + 4\psi_2\mathbf{x}^T\mathbf{x}) \}$  contiene a (4.51), cada vez que el primero esté acotado, el segundo necesariamente también lo estará. Ahora observe que:

$$\{ \mathbf{x} \mid \hat{\ell}^T(\mathbf{x})\hat{\ell}(\mathbf{x}) \leq mF(\hat{\sigma}_1 + 4\hat{\sigma}_2\mathbf{x}^T\mathbf{x}) \} =$$

$$\{ \mathbf{x} \mid \hat{\mathbf{b}}^T\hat{\mathbf{b}} + 4\hat{\mathbf{b}}^T\hat{\mathbf{B}}\mathbf{x} + 4\mathbf{x}^T(\hat{\mathbf{B}}^2 - mF\hat{\sigma}_2\mathbf{I})\mathbf{x} \leq mF\hat{\sigma}_1 \} =$$

$$\{ \mathbf{x} \mid (\mathbf{x} + \frac{1}{2}(\hat{\mathbf{B}}^2 - mF\hat{\sigma}_2\mathbf{I})^{-1}\hat{\mathbf{b}}\mathbf{b})^T(\hat{\mathbf{B}}^2 - mF\hat{\sigma}_2\mathbf{I})(\mathbf{x} + \frac{1}{2}(\hat{\mathbf{B}}^2 - mF\hat{\sigma}_2\mathbf{I})^{-1}\hat{\mathbf{b}}\mathbf{b}) \leq$$

$$mF\hat{\sigma}_1/4 + mF\hat{\sigma}_2\hat{\mathbf{b}}^T(\hat{\mathbf{B}}^2 - mF\hat{\sigma}_2\mathbf{I})^{-1}\hat{\mathbf{b}}/4 \}$$

Y por los lemas 4.15 y 4.16, se sigue que esta región está acotada si, y sólo si,  $(\hat{\mathbf{B}}^2 - mF\hat{\sigma}_2\mathbf{I})$  es una matriz positiva definida, que es equivalente a decir que para cada  $\lambda_1$ , valor propio de  $\hat{\mathbf{B}}$ ,  $\lambda_1^2 > mF\hat{\sigma}_2$ .

#### Lema 4.18

Si el diseño cumple las propiedades de simetría: (4.32), (4.33) y (4.34);  $V$  es diagonal y la región de confianza conservadora (4.52) está acotada, entonces se rechaza la hipótesis nula en la prueba:

$H_0: B^2$  es positiva semidefinida. contra  $H_a: B^2$  es positiva definida.

#### Demostración.

Bajo el supuesto que la hipótesis nula es cierta debe existir al menos un vector  $\mathbf{x} \in \mathbb{R}^m$  de norma 1, tal que  $B\mathbf{x} = 0$ . Ahora observe que la  $i$ -ésima coordenada de  $B\mathbf{x}$  es igual a  $\hat{\mathbf{b}}_i^T\mathbf{x}$ , y como  $V$  es diagonal, las v. a.s  $\hat{\mathbf{b}}_i^T\mathbf{x} \sim N(0, \mathbf{x}^T\text{Cov}(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_1)\mathbf{x})$  son independientes; así la región de rechazo de  $H_0$  es:

$$\sum (\hat{\mathbf{b}}_i^T\mathbf{x})^2 / \mathbf{x}^T\hat{\text{Cov}}(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_1)\mathbf{x} > mF$$

por el lema 4.17 se sabe que la región conservadora (4.52) está acotada, si y sólo si  $B^2 - mF\hat{\sigma}_2\mathbf{I}$  es positiva definida, lo que implica que para  $\mathbf{x} \in \mathbb{R}^m$  de norma 1:

$$\mathbf{x}^T(B^2 - mF\hat{\sigma}_2\mathbf{I})\mathbf{x} = \sum (\hat{\mathbf{b}}_i^T\mathbf{x})^2 - mF\hat{\sigma}_2 > 0 \Rightarrow \sum (\hat{\mathbf{b}}_i^T\mathbf{x})^2 / \hat{\sigma}_2 > mF$$

y por el lema 4.13, se sigue que  $\mathbf{x}^T\text{Cov}(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_1)\mathbf{x} < \hat{\sigma}_2\mathbf{x}^T\mathbf{x} = \hat{\sigma}_2$  lo cual implica que

$$\sum (\hat{\mathbf{b}}_i^T\mathbf{x})^2 / \mathbf{x}^T\hat{\text{Cov}}(\hat{\mathbf{b}}_1, \hat{\mathbf{b}}_1)\mathbf{x} > \sum (\hat{\mathbf{b}}_i^T\mathbf{x})^2 / \hat{\sigma}_2 > mF$$

de lo cual se concluye que cada vez que (4.52) está acotada, se rechaza la hipótesis nula del lema.

**Definición.**

El conjunto

$$\{ \mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) / \sqrt{V(\mathbf{x}^T \hat{\ell}(\mathbf{x}))} \geq -t_\alpha \} \quad (4.53)$$

con  $\mathbf{x}_{op}^T \hat{\ell}(\mathbf{x}_{op}) = \mathbf{x}_{op}^T \mathbf{b} + 2\mathbf{x}_{op}^T \hat{\mathbf{B}}\mathbf{x}_{op} \sim N(0, V(\mathbf{x}^T \hat{\ell}(\mathbf{x})))$ ; es la región de confianza auxiliar.

**Lema 4.19**

Si se cumplen las condiciones de simetría (4.32), (4.33) y (4.34) y la matriz  $V$  es diagonal entonces la región de confianza auxiliar (4.53) está contenida en la región

$$\{ \mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) / \sqrt{V(\hat{\beta}_1) \mathbf{x}^T \mathbf{x} + 4\sigma_2^2 (\mathbf{x}^T \mathbf{x})^2} \geq -t_\alpha \} \quad (4.54)$$

que a su vez, está contenido en la región

$$\{ \mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) \geq -2t_\alpha \sqrt{\sigma_2^2} ( \hat{V}(\hat{\beta}_1) / 8\sigma_2^2 + \mathbf{x}^T \mathbf{x} ) \}. \quad (4.55)$$

**Demostración.**

Observe que  $V(\mathbf{x}^T \hat{\ell}(\mathbf{x})) = \mathbf{x}^T V(\hat{\ell}(\mathbf{x})) \mathbf{x} = \mathbf{x}^T (V(\hat{\mathbf{b}}) + 4V(\hat{\mathbf{B}}\mathbf{x})) \mathbf{x}$ , y por el lema 4.12

$$\mathbf{x}^T (V(\hat{\mathbf{b}}) + 4V(\hat{\mathbf{B}}\mathbf{x})) \mathbf{x} \leq \mathbf{x}^T (V(\hat{\beta}_1) + 4\sigma_2^2 \mathbf{x}^T \mathbf{x}) \mathbf{x} = V(\hat{\beta}_1) \mathbf{x}^T \mathbf{x} + 4\sigma_2^2 (\mathbf{x}^T \mathbf{x})^2$$

Por otro lado,

$$V(\hat{\beta}_1) \mathbf{x}^T \mathbf{x} + 4\sigma_2^2 (\mathbf{x}^T \mathbf{x})^2 =$$

$$4\sigma_2^2 [ (\mathbf{x}^T \mathbf{x})^2 + V(\hat{\beta}_1) \mathbf{x}^T \mathbf{x} / 4\sigma_2^2 + V^2(\hat{\beta}_1) / 64\sigma_2^2 ] - V^2(\hat{\beta}_1) / 16\sigma_2^2 \leq$$

$$4\sigma_2^2 [ (\mathbf{x}^T \mathbf{x})^2 + V(\hat{\beta}_1) \mathbf{x}^T \mathbf{x} / 4\sigma_2^2 + V^2(\hat{\beta}_1) / 64\sigma_2^2 ] = 4\sigma_2^2 [ (\mathbf{x}^T \mathbf{x}) + V(\hat{\beta}_1) / 8\sigma_2^2 ]^2$$

y ésto implica que

$$\sqrt{V(\mathbf{x}^T \hat{\ell}(\mathbf{x}))} \leq \sqrt{\hat{V}(\hat{\beta}_1) \mathbf{x}^T \mathbf{x} + 4\sigma_2^2 (\mathbf{x}^T \mathbf{x})^2} \leq 2\sqrt{\sigma_2} (\mathbf{x} \mathbf{x})^T + V(\hat{\beta}_1) / 8\sqrt{\sigma_2}$$

Así finalmente se concluye que si

$$\mathbf{x} \in \{ \mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) \geq -t_\alpha \sqrt{V(\mathbf{x}^T \hat{\ell}(\mathbf{x}))} \}$$

+

$$\mathbf{x} \in \{ \mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) / \sqrt{\hat{V}(\hat{\beta}_1) \mathbf{x}^T \mathbf{x} + 4\sigma_2^2 (\mathbf{x}^T \mathbf{x})^2} \geq -t_\alpha \}$$

+

$$\mathbf{x} \in \{ \mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) \geq -2t_\alpha \sqrt{\sigma_2} ( \hat{V}(\hat{\beta}_1) / 8\sigma_2 + \mathbf{x}^T \mathbf{x} ) \}$$

y queda demostrado el lema.

#### Lema 4.20

Si se cumplen las condiciones de simetría (4.32), (4.33) y (4.34) y la matriz  $V$  es diagonal entonces una condición suficiente para que la región (4.53) esté acotada es que  $\lambda_i + t_\alpha \sqrt{\sigma_2} < 0$ , para toda  $i$ .

**Demstración.**

Una condición suficiente para que (4.53) esté acotada es que (4.55) lo esté, así

$$\begin{aligned} & \{ \mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) \geq -2t_\alpha \sqrt{\sigma_2} ( \hat{V}(\hat{\beta}_1) / 8\sigma_2 + \mathbf{x}^T \mathbf{x} ) \} \\ & \{ \mathbf{x} \mid \hat{\mathbf{b}}^T \mathbf{x} + 2\mathbf{x}^T (\hat{\mathbf{B}} + t_\alpha \sqrt{\sigma_2} \mathbf{I}) \mathbf{x} \geq -t_\alpha V(\hat{\beta}_1) / 4\sqrt{\sigma_2} \} = \\ & \{ \mathbf{x} \mid (\mathbf{x} + \frac{1}{4}(\hat{\mathbf{B}} + t_\alpha \sqrt{\sigma_2} \mathbf{I})^{-1} \hat{\mathbf{b}})^T (\hat{\mathbf{B}} + t_\alpha \sqrt{\sigma_2} \mathbf{I}) (\mathbf{x} + \frac{1}{4}(\hat{\mathbf{B}} + t_\alpha \sqrt{\sigma_2} \mathbf{I})^{-1} \hat{\mathbf{b}}) \geq \\ & \quad -tV(\hat{\beta}_1) / 8\sqrt{\sigma_2} + \hat{\mathbf{b}}^T (\hat{\mathbf{B}} + t_\alpha \sqrt{\sigma_2} \mathbf{I})^{-1} \hat{\mathbf{b}} / 16 \} \end{aligned}$$

Y esta región está acotada si, y sólo si  $\hat{\mathbf{B}} + t_\alpha \sqrt{\sigma_2} \mathbf{I}$  es negativa definida, y esto se da si, y sólo si  $\lambda_i + t_\alpha \sqrt{\sigma_2} < 0$ , donde  $\lambda_i$  son los valores propios de  $\hat{\mathbf{B}}$ .

Y queda demostrado el lema.

**Corolario.**

Si la región (4.52) está acotada, entonces la región (4.55) también lo está.

**Demostración.**

Como  $t_{\alpha}^2 < mF_{\alpha}$ , entonces si la región (4.52) existe eso quiere decir que  $\hat{\lambda}_1^2 > mF_{\alpha} \hat{\sigma}_2^2$  y esto implica que  $\hat{\lambda}_1^2 > t_{\alpha}^2 \hat{\sigma}_2^2$  que equivale a decir que (4.55) está acotada. Queda demostrado el corolario.

Observe que la frontera de las regiones (4.52) y (4.55) dependen del vector  $\hat{b}$  (y por lo mismo del vector  $\hat{x}_{op} = -\frac{1}{2}\hat{B}^{-1}\hat{b}$ , y de las varianzas del vector  $\hat{b}$  y de la matriz  $\hat{B}$ , y que estas regiones de confianza serán más pequeñas mientras  $\hat{b}$  esté más cerca del cero ó las varianzas de los coeficientes de  $\hat{\eta}$  sean más pequeñas.

**Lema 4.21**

Si se cumplen las condiciones de simetría (4.32), (4.33) y (4.34) y la matriz  $V$  es diagonal, una condición suficiente para rechazar la hipótesis nula en la prueba:

$H_0: B$  es positiva semidefinida.    contra     $H_a: B$  es negativa definida.

es que la región (4.53) este acotada.

**Demostración.**

Para tener evidencia que  $B$  es negativa definida es necesario que para toda  $x \in R^n$  de norma 1 se cumpla que el valor  $x^T B x = \sum x_i b_i^T x$  sea "pequeño". Así, considerando que  $V$  es diagonal, la región de rechazo de  $H_0$  es:

$$\sum x_i b_i^T x / \sqrt{\sum x_i^2 x^T \text{Cov}(b_i, b_i) x} < -t$$

Por otro lado, el lema 4.20 afirma que si la región (4.53) está acotada, si y sólo si  $\hat{B} + t\sqrt{\hat{\sigma}_2^2} I$  es negativa definida, de esto y del lema 4.13 se sigue que

$$\sum x_i b_i^T x / \sqrt{\sum x_i^2 x^T \text{Cov}(b_i, b_i) x} > \sum x_i b_i^T x / \sqrt{\sigma_2^2} > -t$$

de aquí se concluye que si (4.53) está acotada, entonces necesariamente se rechaza la hipótesis nula del lema.

#### Lema 4.22

Las regiones de confianza (4.52) y (4.55) son dos hiper-elipsoides cuyos ejes direccionales son paralelos.

#### Demostración.

La región de confianza (4.52) está formada por los puntos interiores de un hiper-elipsoide con los ejes direccionales iguales a los vectores propios de  $\hat{B} + t_a \sqrt{\sigma_2^2}$ , de magnitud proporcional a sus valores propios correspondientes.

Por otro lado, la región (4.55) está formada por los puntos interiores de un hiper-elipsoide con los ejes direccionales iguales a los vectores propios de  $\hat{B}^2 - F_a \sigma_2^2$ , de magnitud proporcional a sus valores propios correspondientes.

Pero  $\hat{B} + t_a \sqrt{\sigma_2^2}$  y  $\hat{B}^2 - F_a \sigma_2^2$  tienen los mismos vectores propios, por lo que se sigue que los ejes direccionales de las dos regiones son paralelos.

#### Lema 4.23

Si las regiones de confianza en (4.52) y en (4.55) existen, entonces el centro de la frontera de (4.52) está más lejos del origen que el centro de la frontera de (4.55)

#### Demostración.

De (4.52) y (4.55) se puede ver que el centro de las fronteras de ambas regiones de confianza son:  $\frac{1}{2}(\hat{B}^2 - mF_a \sigma_2^2 I)^{-1} \hat{B} b$  y  $\frac{1}{4}(\hat{B} + t \sqrt{\sigma_2^2} I)^{-1} \hat{B} b$ , respectivamente.

Por otro lado, se sabe que  $(\hat{B}^2 - mF_a \sigma_2^2 I)^{-1}$ ,  $\hat{B}$  y  $(\hat{B} + t \sqrt{\sigma_2^2} I)^{-1}$  tienen los mismos valores propios, por lo que existe una matriz ortogonal  $R$  que diagonaliza a las tres simultáneamente y  $R^T R = I = R R^T$ .

Ahora considere el centro del primer conjunto:

$$\frac{1}{2}R(\hat{B}^2 - mF\hat{\sigma}_2 I)^{-1}\hat{B}\hat{b} = \frac{1}{2}R(\hat{B}^2 - mF\hat{\sigma}_2 I)^{-1}R^T R R^T R \hat{b} = \frac{1}{2}\text{diag}[\lambda_1(\hat{\lambda}_1^2 - mF\hat{\sigma}_2 I)^{-1}]\hat{b}$$

con  $\hat{b} = R\hat{b}$

los elementos de este vector son  $\lambda_1 b_1^*(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)/2$  y el cuadrado de su norma es:

$$\frac{1}{4}\|(\hat{B}^2 - mF\hat{\sigma}_2 I)^{-1}\hat{B}\hat{b}\|^2 = \sum \lambda_1^2 b_1^{*2} / 4(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)^2 \quad (4.56)$$

de manera semejante, para el centro del segundo conjunto

$$\frac{1}{4}R(\hat{B} + t\sqrt{\hat{\sigma}_2} I)^{-1}\hat{b} = \frac{1}{4}R(\hat{B} + t\sqrt{\hat{\sigma}_2} I)^{-1}R^T R \hat{b} = \frac{1}{4}\text{diag}(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})^{-1}\hat{b}$$

y su norma al cuadrado es:

$$\frac{1}{16}\|(\hat{B} + t\sqrt{\hat{\sigma}_2} I)^{-1}\hat{b}\|^2 = \sum b_1^{*2} / 16(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})^2 \quad (4.57)$$

Si ambas regiones de confianza existen, entonces  $\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2} < 0$ ;  $\hat{\lambda}_1^2 - mF\hat{\sigma}_2 > 0$  (además  $\lambda_1 < 0$ );

Si se resta la relación (4.57) a la relación (4.56), se obtiene

$$\frac{1}{4}\|(\hat{B}^2 - mF\hat{\sigma}_2 I)^{-1}\hat{B}\hat{b}\|^2 - \frac{1}{16}\|(\hat{B} + t\sqrt{\hat{\sigma}_2} I)^{-1}\hat{b}\|^2$$

$$\sum [ \lambda_1^2 b_1^{*2} / 4(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)^2 - b_1^{*2} / 16(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})^2 ] =$$

$$\sum \frac{b_1^{*2}}{4} \left[ \frac{\lambda_1^2}{(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)^2} - \frac{1}{4(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})^2} \right] = \quad (4.58)$$

así, cada sumando de esta relación es:

$$\frac{b_1^{-2}}{4} \left[ \frac{\lambda_1}{(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)} - \frac{1}{2(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})} \right] \left[ \frac{\lambda_1}{(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)} + \frac{1}{2(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})} \right]$$

En esta expresión, hay 3 factores: el primero es positivo, porque es el cuadrado de un número real; el tercero es negativo, porque es la suma de dos términos negativos; y el segundo es una suma de dos términos de signo contrario. Así, para saber el signo de la expresión se debe determinar el signo del segundo término.

$$\frac{\lambda_1}{(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)} - \frac{1}{2(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})} = \frac{(\lambda_1 + t\sqrt{\hat{\sigma}_2})^2 + (mF - t^2)\hat{\sigma}_2}{2(\hat{\lambda}_1^2 - mF\hat{\sigma}_2)(\hat{\lambda}_1 + t\sqrt{\hat{\sigma}_2})}$$

El numerador de esta expresión es positivo, mientras que el denominador es negativo, así el término es negativo, y por lo tanto (4.58) es positivo, lo que implica que

$$\frac{1}{4} \# (\hat{B}^2 - mF\hat{\sigma}_2 I)^{-1} \hat{B} \hat{b} \#^2 \geq \frac{1}{16} \# (\hat{B} + t\sqrt{\hat{\sigma}_2} I)^{-1} \hat{b} \#^2$$

que es lo que se quería probar.

#### Lema 4.24

Si las regiones de confianza en (4.52) y (4.55) existen entonces al aplicar la rotación R:

- i) los ejes direccionales de sus fronteras son paralelos a los ejes coordenados,
- ii) y el centro de los dos hiper-elipsoides se encuentran en el mismo hipercuadrante.

#### Demostración.

Si R es la matriz que diagonaliza a la matriz simétrica  $A_{m \times m}$ , se tiene que para  $y = Rx$ ,

$$\mathbf{x}^T \mathbf{A}^{-1} \mathbf{x} = \mathbf{y}^T \text{diag}(1/\lambda_1) \mathbf{y} = k$$

y esta última relación es la ecuación de un hiper-elipsoide de ejes paralelos a los ejes coordenados, si y sólo si  $\lambda_1, \lambda_2, \dots, \lambda_m$  y  $k$  sean del mismo signo.

Este resultado aplicado directamente a las fronteras de las regiones de confianza (4.52) y (4.53), implican la primera parte del lema.

Por otro lado, la rotación aplicada a los dos centros de las hiper-elipsoides

$$\frac{1}{2} \mathbf{a}(\hat{\mathbf{B}}^2 - mF\hat{u}_2 \mathbf{I})^{-1} \hat{\mathbf{b}} = \frac{1}{2} \text{diag}[\lambda_1(\hat{\lambda}_1^2 - mF\hat{u}_2 \mathbf{I})^{-1}] \hat{\mathbf{b}}.$$

$$\frac{1}{4} \mathbf{a}(\hat{\mathbf{B}} + t\sqrt{\hat{u}_2} \mathbf{I})^{-1} \hat{\mathbf{b}} = \frac{1}{4} \text{diag}(\hat{\lambda}_1 + t\sqrt{\hat{u}_2})^{-1} \hat{\mathbf{b}}.$$

directamente se observa que las coordenadas correspondientes en ambos vectores son del mismo signo, por lo que se sigue que ambos están en el mismo hiper-cuadrante. Queda demostrado el lema.

Los lemas 4.20, 4.21 y 4.23 implican que la región de confianza usual se localiza más lejos del cero que la región de confianza auxiliar, exactamente como ocurre en el caso de un factor.

Con un factor, se encontró que la región de confianza de  $\mathbf{x}_{op}$  de longitud mínima, se podía escribir como la intersección de una región auxiliar y de una región usual.

Esta región se encuentra más cerca del cero que la usual, y más alejada del cero que la auxiliar.

En este momento se tiene la tentación de definir una región semejante para el caso de  $m$  factores, pero no es fácil determinar el valor de  $\epsilon_0$  que minimice su "medida". Por esta razón se propone que  $\epsilon_0 = \alpha/2$ .

$$\{\mathbf{x} \mid \hat{\ell}(\mathbf{x})^T \hat{\mathbf{V}}^{-1} \hat{\ell}(\mathbf{x}) \leq F_{\alpha/2, m}\} \cap \{\mathbf{x} \mid \mathbf{x}^T \hat{\ell}(\mathbf{x}) / \sqrt{\hat{\mathbf{V}}(\mathbf{x}^T \hat{\ell}(\mathbf{x}))} \geq -t_{\alpha/2}\} \quad (4.59)$$

Esta región está contenida en la intersección de (4.52) y (4.55), ambas con más de  $(1 - \alpha/2)100\%$  de confianza.

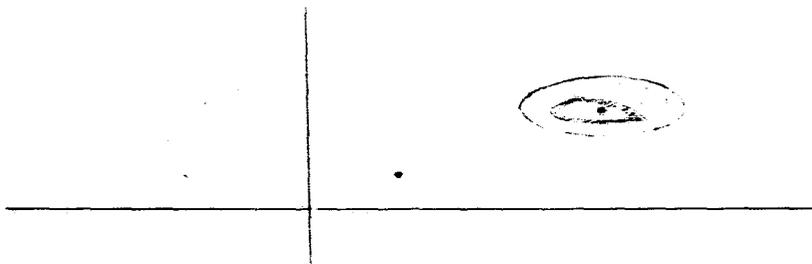


Fig. 4.8 La región de confianza (4.58); y la intersección de (4.52) y (4.55).

#### 4.2.2 LA REGIÓN DE CONFIANZA PARA $\eta_{op}$ .

En esta parte se discute la construcción de una región de confianza conjunta para  $(x_{op}, \eta_{op})$ , así como, la construcción de un intervalo de confianza para  $\eta_{op}$ .

##### Leem 4.25

Dados los conjuntos:

A =

$$\{(x, y) \mid |x^T \hat{\ell}(x) / \sqrt{V(x^T \hat{\ell})}| \leq t, \quad |\hat{k}(x) - y| / \sqrt{V(\hat{k})} \leq \sqrt{t^2 - (x^T \hat{\ell}(x))^2 / V(x^T \hat{\ell})}\}$$

(4.60)

y  $E_1 =$

$$\{(x, y) \mid |e_1^T \hat{\ell}(x) / \sqrt{V(e_1^T \hat{\ell})}| \leq t_1, \quad |\hat{k}(x) - y| / \sqrt{V(\hat{k})} \leq \sqrt{t_1^2 - (e_1^T \hat{\ell}(x))^2 / V(e_1^T \hat{\ell})}\}$$

(4.61)

para  $i = 1, 2, \dots, m$ ; donde  $e_i$  es cada uno de los vectores de la base canónica; y  $\hat{k}(x) = \hat{\beta}_0 + \hat{b}^T x/2 \sim N(\eta(x), V(\hat{k}(x)))$ ;

entonces  $P(A) = 1 - \alpha/2$  y  $P(E_1) = 1 - \alpha/2m$ , implica

$$P(A \cap E_1 \cap E_2 \cap \dots \cap E_m) \geq 1 - \alpha.$$

**Demostración.**

Se sabe que  $1 \geq P(A \cup E_1) = P(A) + P(E_1) - P(A \cap E_1)$  y esto implica que  $1 \geq 1 - \alpha/2 + 1 - \alpha/2m - P(A \cap E_1) \Rightarrow P(A \cap E_1) \geq 1 - \alpha/2 - \alpha/2m$ .

Por un proceso semejante se tiene que  $P(A \cap E_1 \cap E_2) \geq 1 - \alpha/2 - \alpha/m$ , y en general  $P(A \cap E_1 \cap E_2 \cap \dots \cap E_i) \geq 1 - \alpha/2 - i\alpha/2m$ ; El lema se cumple cuando  $i = m$ .

**Teorema 4.11**

Si las v. a.  $\hat{l}(x)$  y  $\hat{k}(x)$  son independientes,  $P(A) = 1 - \alpha/2$ , y  $P(E_1) = \alpha/2m$ ; entonces el número  $t$  en (4.60) es la solución positiva de la ecuación

$$t = \sqrt{[(2/\alpha)^{2/(n-(m+1)(m+2)/2)} - 1]m}$$

y el número  $t_1$  es la solución positiva de

$$t = \sqrt{[(2m/\alpha)^{2/(n-(m+1)(m+2)/2)} - 1]m}$$

**Demostración.**

Por los teoremas 4.5 y 4.6, se sabe que los números  $t$  y  $t_1$  tienen la forma

$$t = \sqrt{[(1/\alpha)^{2/m} - 1]m}$$

donde  $\alpha$  es la significancia y  $m$  los grados de libertad correspondientes.

Si en esta fórmula se sustituye la significancia y los grados de libertad de los conjuntos  $A$  y  $E_1$ , se tiene lo que afirma el teorema.

**Teorema 4.12**

Si el diseño cumple las propiedades de simetría (4.32), (4.33) y (4.34); entonces  $\hat{\ell}(\mathbf{x})$  y  $\hat{k}(\mathbf{x})$  son independientes si y sólo si el diseño cumple la relación

$$n(\sum x_{1j}^4 + (m-1)\sum x_{1j}^2 x_{kj}^2) - (m+4)(\sum x_{1j}^2)^2 = 0, \quad (4.62)$$

**Demostración.**

Como  $\hat{\ell}(\mathbf{x})$  y  $\hat{k}(\mathbf{x})$  son v. a. normales, entonces  $\text{Cov}(\hat{\ell}(\mathbf{x}), \hat{k}(\mathbf{x})) = 0$  si y sólo si  $\hat{\ell}(\mathbf{x})$  y  $\hat{k}(\mathbf{x})$  son independientes.

$$\text{Cov}(\hat{\ell}(\mathbf{x}), \hat{k}(\mathbf{x})) = \text{Cov}(\hat{\beta}_0, \hat{\mathbf{b}}) + (2\text{Cov}(\hat{\beta}_0, \hat{\mathbf{B}}) + V(\hat{\mathbf{b}})/2)^T \mathbf{x} + \text{Cov}(\hat{\mathbf{b}}^T \mathbf{x}, \hat{\mathbf{B}} \mathbf{x})$$

Pero se sabe que el diseño cumple las tres condiciones de simetría, por lo que

$$\text{Cov}(\hat{\ell}(\mathbf{x}), \hat{k}(\mathbf{x})) = (2\text{Cov}(\hat{\beta}_0, \hat{\mathbf{B}}) + V(\hat{\mathbf{b}})) \mathbf{x} = \text{diag}(2\text{Cov}(\hat{\beta}_0, \hat{\beta}_{11}) + V(\hat{\beta}_1)/2) \mathbf{x}$$

este término es cero, independientemente del vector  $\mathbf{x}$ , si y sólo si

$$2\text{Cov}(\hat{\beta}_0, \hat{\beta}_{11}) + V(\hat{\beta}_1)/2 = 0.$$

Al sustituir (4.36) y (4.37) en esta ecuación se sigue que

$$\frac{2\sum x_{1j}^2}{(n(m-1)\sum x_{1j}^2 x_{kj}^2 - m(\sum x_{1j}^2)^2 + n\sum x_{1j}^4)} + \frac{1}{2\sum x_{1j}^2} = 0$$

•

$$n(\sum x_{1j}^4 + (m-1)\sum x_{1j}^2 x_{kj}^2) - (m+4)(\sum x_{1j}^2)^2 = 0$$

y esto implica lo que afirma el teorema.

**Teorema 4.13**

Si el diseño cumple las propiedades de simetría (4.32), (4.33) y (4.34), entonces  $V(\hat{\ell}(\mathbf{x}))$  es diagonal y que  $\hat{\ell}(\mathbf{x})$  y  $\hat{k}(\mathbf{x})$  son independientes, si y sólo si los puntos del diseño cumplen las relaciones:

$$\sum x_{ij}^4 = \frac{\sqrt{(m+1)^2+16} - (m-1)}{2} \sum x_{ij}^2 x_{kj}^2 \quad (4.63)$$

$$(\sum x_{ij}^2)^2 = \frac{n(\sqrt{(m+1)^2+16} + (m-1))}{2(m+4)} \sum x_{ij}^2 x_{kj}^2 \quad (4.64)$$

**Demstración.**

Suponga que  $V(\hat{\ell}(\mathbf{x}))$  es diagonal y que  $\hat{\ell}(\mathbf{x})$  y  $\hat{k}(\mathbf{x})$  son independientes, entonces se cumplen simultaneamente (4.50) y (4.62).

si de (4.62) se despeja  $(\sum x_{ij}^2)$  y se sustituye en (4.50) y llega a

$$\sum x_{ij}^4 = \frac{\sqrt{(m+1)^2+16} - (m-1)}{2} \sum x_{ij}^2 x_{kj}^2$$

y si en (4.62) se sustituye este término, y se despeja a  $(\sum x_{ij}^2)^2$  se llega a

$$(\sum x_{ij}^2)^2 = \frac{n(\sqrt{(m+1)^2+16} + (m-1))}{2(m+4)} \sum x_{ij}^2 x_{kj}^2$$

que es lo que afirma el teorema.

Una manera de tener un diseño que cumpla las propiedades que se establecen en el teorema 4.13 es la siguiente:

- i) Escoja 3 subconjuntos de  $R^n$ ,  $D_1$ ,  $D_2$  y  $D_3$ , que cumplan las propiedades (4.32), (4.33) y (4.34).
- ii) Obtenga el conjunto  $D_{\alpha\delta} = D_1 \cup \alpha D_2 \cup \delta D_3$  más las repeticiones en el centro.
- iii) Encuentre los valores de  $\alpha$  y  $\delta$  que permiten que  $D_{\alpha\delta}$  cumpla las condiciones (4.63) y (4.64).

Por ejemplo, para dos factores se tiene el siguiente diseño que cumple las condiciones (4.32), (4.33), (4.34), (4.62), (4.63) y (4.64).

$$\begin{aligned} \text{si } D_1 = D_2 &= \{(1, 1), (1, -1), (-1, 1), (-1, -1)\} \\ \text{y } D_3 &= \{(1, 0), (-1, 0), (0, 1), (0, -1)\} \end{aligned}$$

$$D_{\alpha\delta} = D_1 \cup \alpha D_2 \cup \delta D_3 = \{(1, 1), (1, -1), (-1, 1), (-1, -1), \alpha(1, 1),$$

$\alpha(1, -1), \alpha(-1, 1), \alpha(-1, -1), \delta(1, 0), \delta(-1, 0), \delta(0, 1), \delta(0, -1)\}$   
además 13 repeticiones en el centro.

Los valores de  $\alpha$  y  $\delta$  se encuentran observando que:

$$\sum_{i,j}^4 x_{ij}^4 = 4(1 + \alpha^4) + 2\delta^4; \quad \sum_{i,j}^2 x_{ij}^2 = 4(1 + \alpha^4); \quad \sum_{i,j}^2 x_{ij}^2 = 4(1 + \alpha^2) + 2\delta^2.$$

luego sustituyendo en (4.63) y (4.64) estos terminos, recordando que  $m = 2$  y  $n = 25$ ; con lo que se tiene

$$4(1 + \alpha^4) + 2\delta^4 = 8(1 + \alpha^4)$$

$$(4(1 + \alpha^2) + 2\delta^2)^2 = 50(1 + \alpha^4)$$

La solución de este sistema de ecuaciones es:

$$\alpha^2 = 8 \pm \sqrt{63}, \quad \text{y} \quad \delta^4 = 2(1 + \alpha^4)$$

con lo que finalmente se puede encontrar los puntos del diseño.

Intervalo de confianza para el valor individual de  $\eta_{op}$ .

Para encontrar el intervalo de confianza de  $\eta_{op}$  se encuentra una región en  $R^{m-1}$  que contiene a  $(\mathbf{x}, \eta(\mathbf{x}))$  con  $(1 - \alpha)100\%$  de confianza, dado por el conjunto:

$$\{(\mathbf{x}, y) \mid |\hat{\eta}(\mathbf{x}) - y| \leq (k\hat{V}(\hat{\eta}(\mathbf{x})))^{1/2} \text{ para todo } \mathbf{x} \in R^m\} \quad (4.65)$$

**Lema 4.26**

Una condición suficiente para que  $(\mathbf{x}, \eta(\mathbf{x}))$  esté en el conjunto (4.65) con  $(1 - \alpha)100\%$  de confianza es que

$$k = (m+1)(m+2)F_{(m+1)(m+2)/2; n-(m+1)(m+2)/2} / 2$$

**Demostración.**

De acuerdo al lema 4.5, la relación

$$|\hat{\eta}(\mathbf{x}) - \eta(\mathbf{x})| \leq (k\hat{V}(\hat{\eta}(\mathbf{x})))^{1/2} \quad \text{para todo } \mathbf{x} \in \mathbb{R}^m$$

es equivalente a  $k \geq \epsilon^T W(W^T W)^{-1} W^T \epsilon / \sigma^2$ . Y como  $W(W^T W)^{-1} W^T / \sigma^2$  es idempotente de rango  $(m+1)(m+2)/2$  la variable  $\epsilon^T W(W^T W)^{-1} W^T \epsilon / \sigma^2$  se distribuye como una  $\chi^2$  con  $(m+1)(m+2)/2$  grados de libertad, y por consecuencia

$$\frac{\epsilon^T W(W^T W)^{-1} W^T \epsilon}{(m+1)(m+2)\sigma^2/2} \sim F$$

con  $(m+1)(m+2)/2$  y  $n - (m+1)(m+2)/2$  grados de libertad.

con lo que se sigue que

$$P((m+1)(m+2)F_{\alpha} / 2 \geq \epsilon^T W(W^T W)^{-1} W^T \epsilon / \sigma^2) = 1 - \alpha$$

implica que  $k = (m+1)(m+2)F_{(m+1)(m+2)/2; n-(m+1)(m+2)/2} / 2$

y queda demostrado el lema.

Así cuando  $2k = (m+2)(m+1)F_{(m+2)(m+1)/2; n-(m+2)(m+1)/2}$  se tiene que:

$$\eta(\mathbf{x}) \in (\hat{\eta}(\mathbf{x}) - \sqrt{k\hat{V}(\hat{\eta})}, \hat{\eta}(\mathbf{x}) + \sqrt{k\hat{V}(\hat{\eta})})$$

para todo  $\mathbf{x} \in \mathbb{R}^m$ , con un  $100(1 - \alpha)\%$  de confianza.

Ahora considere que  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  y  $\mathbf{x}_3$  son los puntos donde

$$\hat{\eta}(\mathbf{x}) - \sqrt{k\hat{V}(\hat{\eta})}, \quad \eta(\mathbf{x}) \quad \text{y} \quad \hat{\eta}(\mathbf{x}) + \sqrt{k\hat{V}(\hat{\eta})}$$

alcanzan su valor máximo, respectivamente, entonces

$$\max_{\mathbf{x}} \{ \hat{\eta}(\mathbf{x}) - \sqrt{kV(\hat{\eta})_{op}} \} \leq \eta_{op} \leq \max_{\mathbf{x}} \{ \hat{\eta}(\mathbf{x}) + \sqrt{kV(\hat{\eta})} \}$$

con  $(1-\alpha)100\%$  de confianza. Esto da origen a un intervalo de confianza conservador para  $\eta_{op}$  independiente de la posición de  $\mathbf{x}_{op}$ .

**Definición.**

El intervalo de confianza para  $\eta_{op}$  es el conjunto

$$\{ y \mid \max_{\mathbf{x}} \{ \hat{\eta}(\mathbf{x}) - \sqrt{kV(\hat{\eta})_{op}} \} \leq y \leq \max_{\mathbf{x}} \{ \hat{\eta}(\mathbf{x}) + \sqrt{kV(\hat{\eta})} \} \}$$

es el intervalo de confianza para  $\eta_{op}$ .

## 5. CONCLUSIONES Y COMENTARIOS FINALES

El trabajo se presenta en tres partes: En la primera parte se revisaron, de manera general, los métodos y los diseños utilizados para encontrar el óptimo; en la segunda parte se propusieron siete variantes al método de ascenso por pendiente máxima y, posteriormente, se exploró su efectividad mediante una serie de simulaciones; y por último, en la tercera parte se dieron los elementos para obtener las regiones de confianza para  $\bar{x}_{op}$  y  $\eta_{op}$  y se propuso una; posteriormente, analizando estas regiones de confianza se determinan algunas de las condiciones experimentales que optimizan las estimaciones.

Los resultados obtenidos de las simulaciones son muy claros al señalar que la séptima variante al método de ascenso por pendiente máxima es la más efectiva en los casos considerados; y ésta es una de las aportaciones de este trabajo.

Las modificaciones que se hicieron al método de pendiente máxima y que dieron lugar a la séptima variante fueron:

- i) Un nuevo criterio para elegir la longitud de paso.
- ii) Una nueva manera de encontrar el centro de los subsecuentes diseños.
- iii) La incorporación de toda la información disponible al momento de obtener los estimadores.
- iv) Un proceso para elegir, en cada etapa, un diseño que abarca mayor área que los anteriores, donde aún es razonable estimar a la función de respuesta con una función cuadrática.
- v) Y la determinación de un criterio para finalizar la búsqueda del óptimo; dicho criterio está relacionado al hecho que  $\bar{x}_{op}$  pertenezca o no al conjunto de extrapolación S; esto equivale a pedir que  $\bar{x}_{op}$  se encuentre "cercano" al diseño.

Se utilizaron 3 criterios para calificar la bondad de los diferentes métodos de búsqueda.

- i) La desviación media estimada; tanto de  $\bar{x}_{op}$  como de  $\eta_{op}$ .
- ii) El tamaño de muestra requerido para tener las estimaciones.
- iii) Y el tiempo de cómputo consumido.

Los resultados son:

Con respecto a la exactitud: La séptima variante proporciona estimadores mucho más exactos y precisos. Esto se puede resumir de la siguiente manera: la desviación media empírica de  $\hat{x}_{op}$ , con cualquier otro método, es al menos 100 veces más grande que la desviación media empírica del mismo estimador obtenido con la séptima variante. Esto se observó aun en los casos en que el óptimo real se encuentra lejos del punto de inicio; o bien, cuando las curvas de nivel de la función de respuesta son elipses alargadas.

Con respecto al tamaño de la muestra: En todos los casos considerados la séptima variante requirió de pocos datos para obtener los estimadores.

Con respecto al tiempo de cómputo: La séptima variante tardó más que otros; pero, cabe mencionar que con este método es necesario invertir, en cada etapa, la matriz asociada a los datos; y también es necesario guardar en memoria la información de todos los diseños y esto aumenta el tiempo del proceso computacional; sin embargo, hay que recordar que existen algoritmos para invertir matrices que reducen el tiempo de cómputo y que mejoran la exactitud de los redondeos; además de que el problema computacional es cada día menos importante debido a la enorme velocidad de los equipos.

Por otro lado, al analizar las regiones de confianza para  $\hat{x}_{op}$ ; en particular las regiones en (4.45) y (4.48), se observa que la frontera de estas regiones de confianza dependen de una forma cuadrática de  $\hat{b}$ , y de las varianzas de las coordenadas del vector  $\hat{\beta}$ .

Cuando  $\hat{b}$  esté más cerca del origen (y por consecuencia  $\hat{x}_{op} = \hat{B}^{-1}\hat{b}$  esté más cerca de él) o cuando las varianzas de los elementos de  $\hat{\beta}$  sean menores, las regiones de confianza tendrán menor medida. De aquí se

desprende que hay dos aspectos que reducen el tamaño de la región de confianza: el primero es, que  $\hat{x}_{op}$  esté cerca del origen; y el segundo, que la varianza de los estimadores de los coeficientes del modelo sea pequeña.

Cuando  $\hat{x}_{op} \in S$ ,  $\hat{x}_{op}$  se encuentra "cerca" del centro del diseño y por lo tanto su norma es "pequeña".

Cuando el área que abarca el diseño es mayor y cuando se utiliza una muestra más grande, la varianza de los estimadores disminuye.

Estos dos aspectos fueron considerados en la séptima variante, por lo que de cierta manera se justifica el porque la precisión empírica de las estimaciones con este método resultaron ser pequeñas.

Los resultados proporcionados por las simulaciones no sólo permiten sugerir el mejor método de búsqueda del óptimo, sino que también sugieren algunas acciones que se pueden tomar antes de iniciar dicha búsqueda (al menos después de una primera estimación de la función de segundo grado).

Por ejemplo: los resultados de la simulación parecen indicar que con un diseño de simetría "regular" con respecto al cero, como el factorial  $2^n$  o el simplex, se obtienen mejores estimaciones cuando las curvas de nivel de  $\hat{\eta}(x)$  son círculos concéntricos al óptimo.

Con esta información se puede conjeturar que si se aplica una transformación que convierta las curvas de nivel de  $\hat{\eta}(x)$  de elipsoides "alargados" a unos más cercanos a círculos, y en este espacio transformado se elige un diseño de simetría "regular", se podrían mejorar las estimaciones de  $\hat{x}_{op}$  y de  $\hat{\eta}$ .

Esta transformación debe estar en función de los datos de un muestreo piloto. ( $\hat{B} = -R^T A^2 R$ , con R una matriz ortogonal, A una matriz diagonal y  $T = AR$  la transformación)

Debe notarse que para llevar a cabo las simulaciones se consideró únicamente el caso en que la matriz A es invertible; si ocurre que la matriz  $\hat{B}$  proporciona evidencia de ser singular se recomienda hacer un análisis anterior, por ejemplo: de componentes principales, para reducir el número de factores en la ecuación y tener una matriz no

singular.

Para terminar se recomienda para trabajos futuros:

Realizar mayor número de simulaciones variando la función de respuesta, aumentando el número de factores y cambiando la forma de  $\eta(\mathbf{x})$ .

Investigar sobre las características de la matriz  $V(\hat{l}(\mathbf{x}))$ , para determinar condiciones que permitan invertirla fácilmente aunque no sea diagonal.

Efectuar algunas simulaciones, usando un diseño con la transformación previa, sugerida anteriormente, para ver si se mejoran los resultados.

Explorar las propiedades estadísticas de la transformación antes mencionada.

Buscar otros diseños, por ejemplo, rotatables que cumplan las condiciones que se enumeraron para obtener las regiones de confianza.

Aplicar lo desarrollado en este trabajo a casos reales, para constatar su eficacia, y para determinar qué limitantes presenta su aplicación a problemas particulares.

Investigar sobre las propiedades del conjunto  $S$ , para determinar más concretamente las regiones de predictibilidad de los diseños experimentales.

## BIBLIOGRAFIA

- Atkinson A. C. y Donev A. M.** (1989), "The construction of exact D-optimum experimental Design with Application to Blocking Response Surface Design", *Biometrika*, 76, 3, 515-526.
- Box G. E. P. y Hunter J. S.** (1954), "A Confidence Region for the Solution of a Set of Simultaneous Equations With an Application to Experimental Design", *Blometrika*, 41, 190-199.
- Box G. E. P. y Hunter J. S.** (1957), "Multi-Factor Experimental Design for Exploring Response Surfaces", *The Annals of Mathematical Statistics*, 28, 195-241.
- Box G. E. P. y Wilson, K. B.** (1951), "On the Experimental Attainment of Optimum Conditions", *Journal of the Royal Statistical Society, Ser. B*, 13, 1-45.
- Carter Jr. W. H., Chinchilli V. M. Campbell E. D. y Wampler G. L.** (1984), "Confidence Interval about the Response at the Stationary Point of a Response Surface, with an Application to Preclinical Cancer Therapy", *Biometrics*, 40, 1125-1130.
- Carter Jr. W. H., Chinchilli V. M. Myers R. H. y Campbell E. D.** (1986), "Confidence Interval and an Improved Ridge Analysis of Response Surfaces", *Technometrics*, Vol. 28, No. 4, 339-346.
- Cochran W. G. & Cox G. M.** (1980), *Diseños Experimentales*, Editorial Trillas, México.
- Davies O. L.** (1954), *The Design and Analysis of Industrial Experimentos*, New York: Hafner Press.
- Fieller E. C.** (1944) "A fundamental Formula in the Statistics of Biological Assay, and some Applications", *Quarterly Journal of Pharmacology*, Vol XVII, No. 2, 117-123.
- Myers R. H. y Khuri A. I.** (1979), "A New Procedure for Steepest Ascent", *Communications in Statistics, Part A-Theory and Methods*, 8(14), 1359-1376.
- Myers, R. H., Khuri A. I. y Carter Jr. W. H.** (1989) "Response Surface Methodology: 1966-1988", *Technometrics*, Vol. 31, No. 2, 137-157.
- Noble B.** (1969), *Applied Linear Algebra*, Prentice-Hall, Inc.
- Notz W.** (1982), "Minimal Point Second Order Design", *Journal of Statistical Planning and Inference*, 6, 47-58.
- O'Reilly F. J.** (1975), "On a Criterion for extrapolation in Normal Regression", *The Annals of Statistics*, Vol. 3, No. 1, 219-222.
- Rao C. R.** (1973), *Linear Statistical Inference and its Applications*, 240. New York, Wiley.
- Searle S. R.** (1971), *Linear Models*, John Wiley & Sons, New York.

## APENDICE I

### DIAGRAMAS DE BLOQUE DE LOS METODOS ANALIZADOS.

En los diagramas de bloque que se presentan se utilizan los siguientes términos.

$CME = (Y-\hat{Y})^T(Y-\hat{Y})/(n-m-1)$  el cuadrado medio del error, o residuo.

$CMR = (\hat{Y}-\bar{Y})^T(\hat{Y}-\bar{Y})/m$  el cuadrado medio de la regresión.

En estas condiciones, el cociente  $CMR/CME$  bajo la hipótesis  $H_0: \beta_i=0$ , para toda  $i = 1, m$ ; tiene una distribución F con  $m$  y  $n-m-1$  grados de libertad. Cuando este cociente es "grande" existe evidencia que para alguna  $i$ ,  $\beta_i \neq 0$ ; y por lo tanto, al menos uno de los factores influye en la respuesta.

$CMEP = (Y(0)-\bar{Y}(0))^T(Y(0)-\bar{Y}(0))/(n_1-1)$  el cuadrado medio del error puro; en esta expresión  $n_1$  indica el número de repeticiones en el centro del diseño.

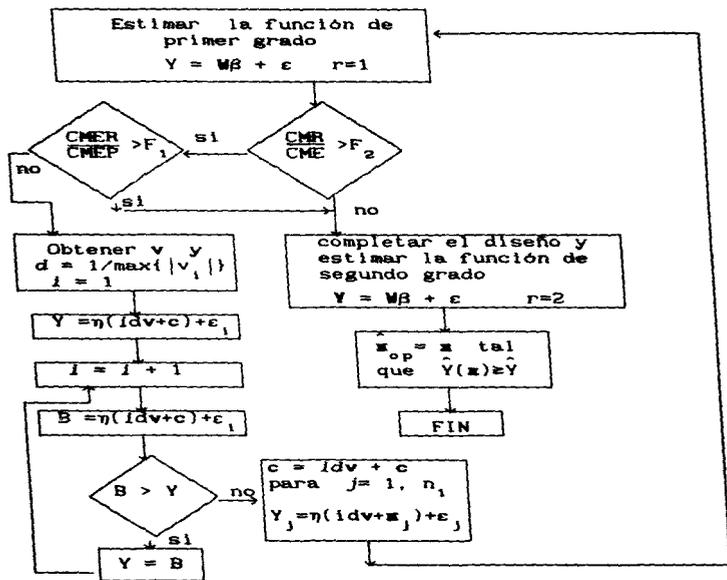
$CMER = [(Y-\hat{Y})^T(Y-\hat{Y}) - (Y(0)-\bar{Y}(0))^T(Y(0)-\bar{Y}(0))]/(n-m-n_1)$  el cuadrado medio del error por falta de ajuste.

El cociente  $CMER/CMEP$  bajo la hipótesis que los datos se ajustan bien al modelo, tiene una distribución F con  $n-m-n_1$  y  $n_1-1$  grados de libertad. Si este cociente es "grande", entonces existe evidencia de falta de ajuste y de la presencia de curvatura.

$F_1$  y  $F_2$  son los valores correspondientes de la distribución F con una significancia  $\alpha$ .

DIAGRAMA DE BLOQUE DEL MÉTODO DE ASCENSO POR PENDIENTE  
MAXIMA

se comienza con  $c = 0$ . ( $c$  es el centro del diseño).



$n_1$  es el número de puntos del  
diseño de primer orden.

Diagrama de la primera y segunda variante.  
 La explicaciones de \* y \*\*, coinciden con las del diagrama del método de ascenso por pendiente máxima, la doble raya (==) indica el flujo de la segunda variante.

se comienza con  $c = 0$ . (c el centro del diseño).

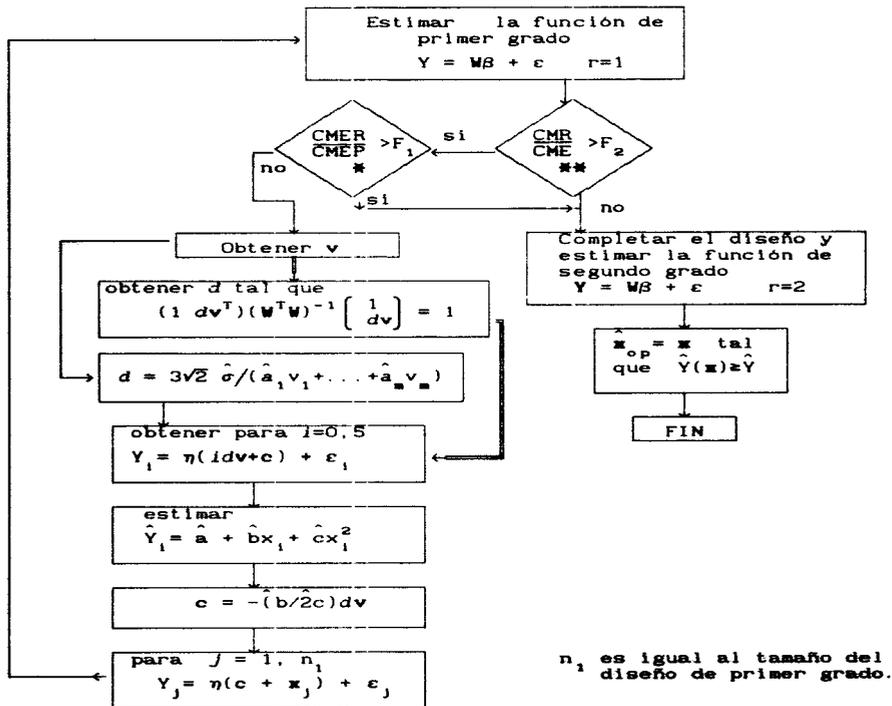


Diagrama de la tercer variante  
 En esta variante la matriz  $W$  se forma al incorporar la nueva información conforme se vaya obteniendo.

se comienza con  $c = 0$ .

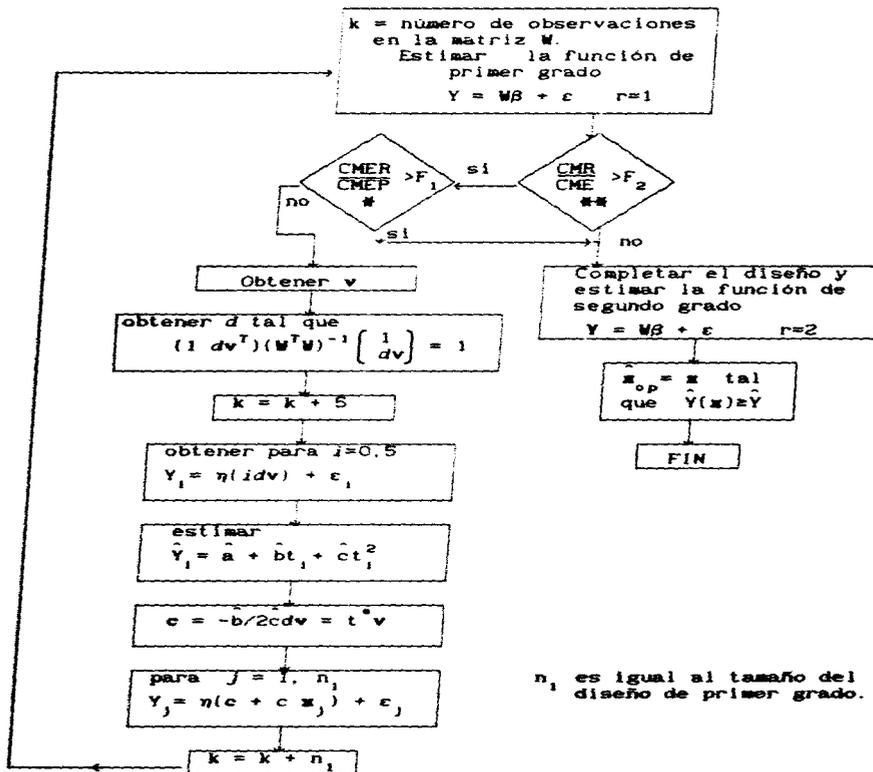


Diagrama de la cuarta variante.

Esta variante modifica la salida del método de ascenso por pendiente máxima, condicionandola a que  $\hat{x}_{op}$  este en la región de extrapolación.

se comienza con  $c = 0$ .

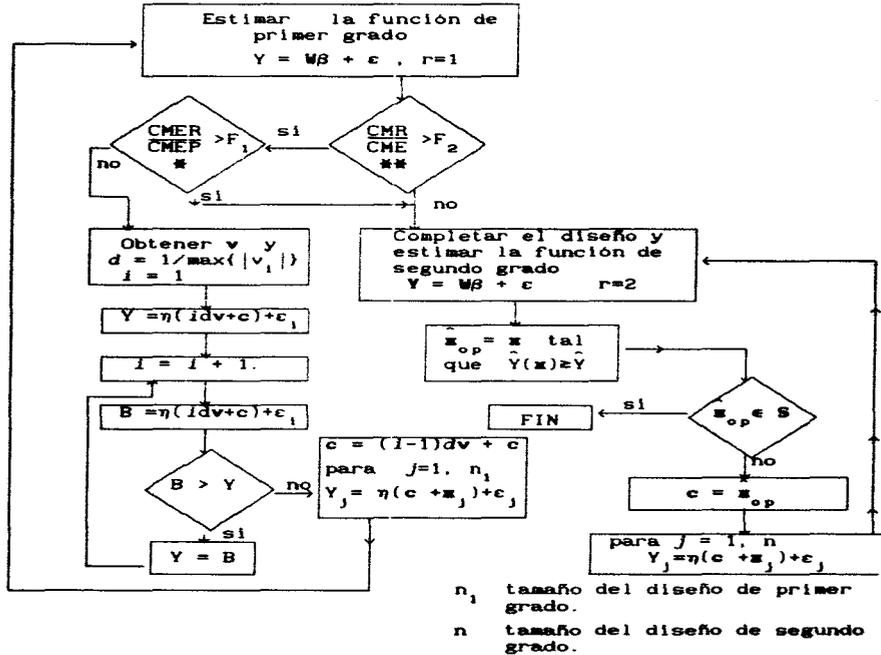


Diagrama de la quinta y sexta variante.  
 La explicaciones de \* y \*\* coinciden con las del diagrama del método de ascenso por pendiente máxima, la doble raya (==) indica el flujo de la sexta variante.

se comienza con  $c = 0$ . ( $c$  el centro del diseño).

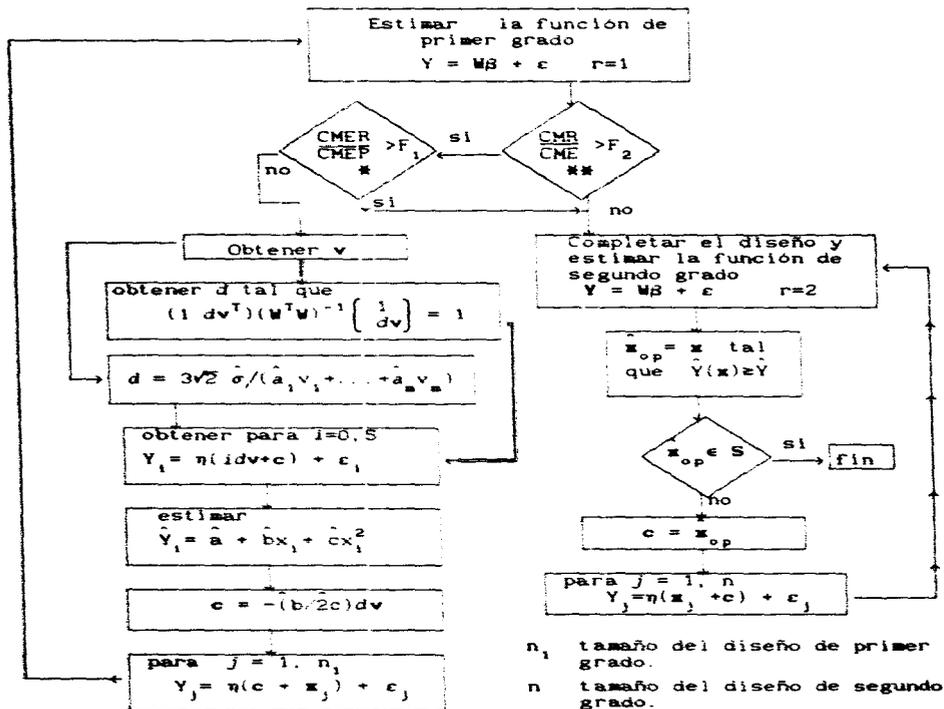
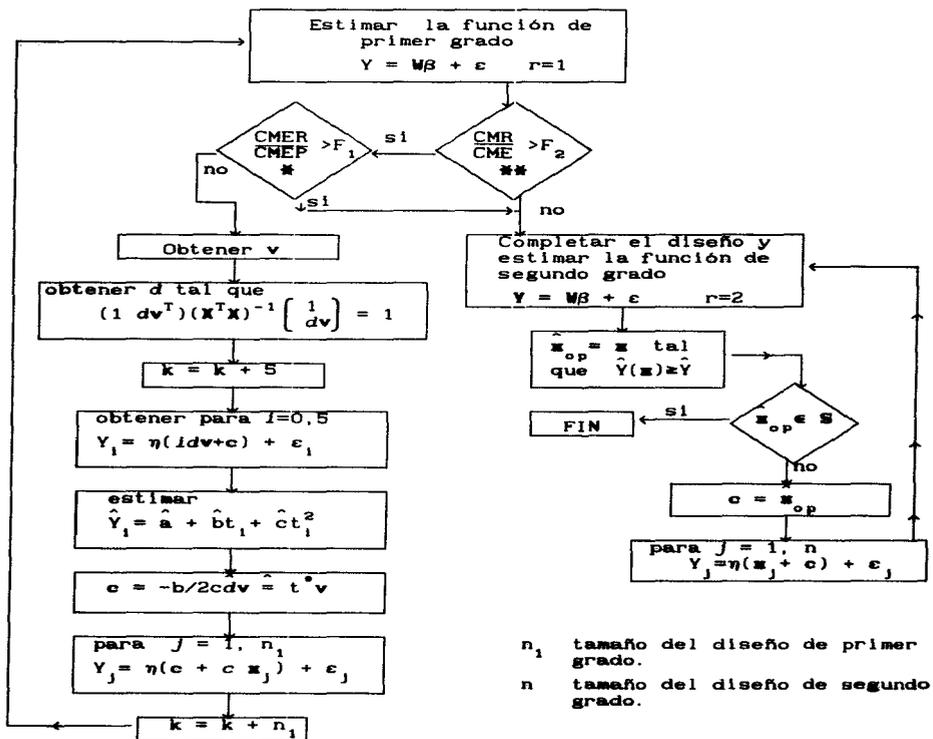


diagrama de la septima variante.

se comienza con  $c = 0$ .  
 $k$  es el número de observaciones en la matriz  $W$ .



## APENDICE II

### RESULTADOS DE LA SIMULACIÓN

En cada tabla se indica:

- 1) La matriz de la forma cuadrática que define la función de respuesta. La matriz A.
- 2) El punto donde se localiza el óptimo. El punto  $x_{op}$ .
- 3) El diseño utilizado.
- 4) La desviación estandar del error de observación,  $\sigma$ .
- 5) Y el método empleado.

Los resultados que se presentan en cada tabla son:

- 1) La variación empírica promedio de la respuesta máxima dada por:

$$1^{\circ} \quad DM(\hat{\eta}_{op}) = \sum |\hat{\eta}_{op} - \eta_{op}| / 500$$

- 2) La razón de la variación empírica promedio de la respuesta máxima dada por:

$$2^{\circ} \quad RV = \sum |\hat{\eta}_{op} - \eta_{op}| / 500\sigma$$

- 3) La variación empírica promedio del punto óptimo, dada por:

$$3^{\circ} \quad DM(\hat{x}) = \sum |\hat{x}_{op} - x_{op}| / 500$$

- y 4) El tamaño de la muestra promedio para converger, dado por:

$$4^{\circ} \quad \bar{N} = \sum N / 500.$$

La simulación se efectuó en la computadora Burroughs-800 de la UNAM. (actualmente la A-12); cada corrida tuvo un tiempo límite de 10 minutos de proceso de cómputo, si en ese tiempo, no llegaba a encontrar los estimadores, se consideró que el proceso no convergía.

TABLA DE RESULTADOS 1.

$$Y = -10.000 + 800x_1 + 800x_2 - 2x_1^2 - 2x_2^2 + \epsilon$$

$$x_{op} = (200, 200)$$

diseño factorial.

1° DM( $\hat{\eta}_{11}$ )    2° RV  
3° DM( $\hat{x}_{11}$ )    4°  $\bar{N}$

$$A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

$\sigma$	M1	M2	M2a	M2b	M2c	M2T	M2aT	M2bT	M2cT
.1	.00017	456.0	437.9	454.9	158.4	.0124	.00282	.00352	4.7E-9
	.00054	1441	1382	1438	500.9	.1247	.00891	.01113	1.4E-8
	.00834	13.4	13.14	13.26	12.46	.0333	.327	.0339	.00006
	30.07	11.0	11.0	11.0	11.0	35.7	44.4	44.6	20.0
.2	.00025	289.3	986.2	899.0	589.0	.0697	.00755	.00892	4.7E-8
	.00127	1446.7	4901.2	4497.0	2945.0	.0345	.037	.0446	2.3E-8
	.00980	18.5	18.6	18.42	18.8	.0449	.0507	.0512	.00009
	30.02	21.3	11.6	11.6	11.8	42.07	46.6	47.60	20.08
.3	.00043	1198.0	1395.0	4652.0	1293.0	.0852	.0073	.00814	5.3E-8
	.00146	4919.0	4652.0	3995.6	4318.0	.0285	.0244	.0271	1.3E-7
	.01304	16.14	23.13	22.04	20.59	.0535	.0510	.0522	.00001
	30.20	54.47	13.88	14.00	12.67	62.31	47.94	51.65	20.18
.4	.00051	12.38	2310.0	1716.2	1539.1	.0115	.0112	.0119	4.8E-8
	.00127	32.08	5775.2	4290.5	3846.6	.0288	.0280	.0292	9.5E-8
	.0140	19.94	29.09	25.45	19.86	.0646	.0595	.0637	.00014
	30.07	71.25	15.45	15.9	13.0	86.9	46.9	52.3	20.27
.5	.00081	72.7	1945.0	2422	89.3	.0186	.0223	.0199	5.7E-8
	.00162	145	3890	4845	17.3	.0308	.0372	.0333	9.3E-8
	.01834	17.35	26.93	29.94	16.17	.0729	.0834	.0865	.00017
	30.27	85.54	17.86	17.65	14.46	95.56	47.11	55.15	20.47
.6	.00096	3.27	3393	3502	82.4	.0205	.0252	.0224	6.7E-8
	.00161	5.4	5655	5837	17.6	.0299	.0360	.0321	9.5E-8
	.01954	18.58	35.84	35.48	19.09	.0719	.0981	.0846	.00016
	30.42	100.3	20.12	21.32	14.47	110.2	49.5	57.3	20.46
.7	.00104	9.97	3446	3175	1432	.0247	.0248	.0259	7.6E-8
	.00142	14.25	4923	4536	19.20	.0311	.0309	.0324	9.5E-8
	.02034	18.4	35.08	34.55	22.68	.0721	.0959	.0946	.00018
	31.41	116.5	20.0	22.52	15.6	135.7	44.67	61.74	20.49
.8	.00130	.0130	4695	5155	.1254	.0151	.0178	.0167	7.6E-8
	.00163	.016	5869	6443	.1752	.0304	.0356	.0334	8.8E-8
	.0227	15.25	41.9	40.82	18.25	.0646	.0758	.0745	.00020
	30.08	133	25.5	26.36	16.4	92.37	46.17	54.40	20.49
.9	.00126	237.5	6586.7	7181.3	259.5	.0165	.0431	.0285	7.6E-8
	.00140	263.9	7318.7	7979.3	273.8	.0329	.0479	.0317	8.5E-8
	.02224	25.53	46.33	50.51	25.75	.0761	.1173	.0945	.00020
	31.30	114.0	24.68	23.36	15.63	140.25	46.35	65.28	20.49
1.0	.00173	57.79	5816	6211	157.45	.0279	.0397	.0299	6.9E-7
	.00173	57.79	5816.3	6211.4	157.45	.0279	.0397	.0299	6.9E-7
	.02564	20.26	45.95	47.29	20.42	.0867	.1140	.09957	.00022
	32.07	131.14	24.8	29.89	16.42	156.21	46.86	68.24	20.51

1\*  
2\*  
3\*  
4\*

TABLA DE RESULTADOS 2.

$$Y = -2500 + 100x_1 + 100x_2 - 2x_1^2 - 2x_2^2 + \epsilon$$

$$x_{op} = (25, 25)$$

1° DM( $\mu$ )      2° RV

Diseño simplex

3° DM( $x_{op}$ )      4°  $\bar{N}$

$$A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M2T	M2aT	M2bT	M2cT	
.1	.00010	11.51	20.59	15.74	11.98	.00395	.00345	.00414	1.3E-6	1*
	.00103	115.1	205.9	167.4	119.8	.03951	.03453	.04149	1.3E-5	2*
	.00125	2.32	2.73	2.48	2.38	.0372	.0365	.0376	.00073	3*
	29.07	13.16	10.54	10.43	10.23	27.31	35.84	35.37	16.24	4*
.2	.00015	10.89	34.43	36.7	16.01	.00941	.00941	.00831	3.0E-6	
	.00077	54.45	172.3	183.5	80.14	.0475	.0470	.0432	1.5E-5	
	.00690	2.31	3.74	3.89	2.462	.0502	.0577	.0541	.00103	
	30.02	20.82	13.95	15.16	31.04	39.31	39.59	39.97	16.66	
.3	.00023	10.18	91.55	66.77	13.24	.01457	.01602	.04391	4.5E-6	
	.00076	33.54	305.2	222.4	44.1	.4857	.0534	.0796	1.5E-5	
	.00367	2.13	5.65	4.94	2.44	.0686	.0764	.0904	.00135	
	28.20	24.04	18.57	17.25	13.33	32.76	41.75	43.53	17.01	
.4	.00036	10.24	150.1	88.88	11.05	.01701	.02247	.01901	6.6E-6	
	.00095	33.51	375.8	221.0	27.64	.04251	.05621	.04753	1.6E-5	
	.00414	2.134	7.109	5.890	2.228	.07411	.08840	.08094	.00162	
	30.47	27.72	22.84	22.36	14.33	35.49	47.84	48.52	17.20	
.5	.00061	8.54	241.7	163.8	8.12	.01854	.03714	.03338	1.0E-5	
	.00122	19.08	483.4	326.6	16.24	.03708	.07431	.06664	2.0E-5	
	.00634	1.94	9.00	6.99	1.783	.0788	.1145	.1184	.00181	
	30.27	31.54	30.24	30.12	15.05	36.57	56.99	53.81	17.40	
.6	.00097	8.67	244.7	173.4	11.58	.0246	.0462	.0463	8.5E-6	
	.00161	14.4	407.8	287.6	19.3	.0413	.0771	.0771	1.4E-5	
	.01064	1.834	8.702	7.260	2.564	.0917	.1297	.1281	.00180	
	30.40	29.86	29.35	29.57	15.30	37.45	60.50	57.41	17.28	
.7	.00095	7.25	657.5	351.2	6.57	.0259	.0600	.0539	1.3E-5	
	.00140	12.08	823.1	439.4	10.9	.0432	.0750	.0674	1.6E-5	
	.00934	1.51	11.84	10.52	1.10	.0981	.1476	.1382	.00214	
	29.47	34.28	41.44	39.69	16.29	37.68	75.11	73.43	17.20	
.8	.00091	14.7	7709.5	710.12	8.48	.0612	.0634	.0786	1.7E-5	
	.00143	18.42	8566.1	790.56	10.6	.0778	.0737	.0873	1.9E-5	
	.01283	2.36	23.0	13.4	1.49	.11364	.1560	.1568	.00245	
	30.80	32.68	46.51	47.17	16.29	49.20	74.70	76.58	17.46	
.9	.00106	14.3	7709.2	710.1	7.92	.0596	.06633	.0786	1.7E-5	
	.00135	15.1	8556.5	790.4	8.83	.0662	.0737	.0873	1.9E-5	
	.01234	2.36	23.01	13.42	1.49	.1129	.1564	.1636	.00245	
	31.37	32.68	46.51	47.17	16.29	49.20	74.70	76.58	17.46	
1.0	.00123	9.58	970.1	1403.7	8.87	.892	.0700	.0627	1.7E-5	
	.00123	9.58	970.1	1403.7	8.87	.0689	.0700	.0627	1.7E-5	
	.01574	1.57	15.27	13.77	1.986	.1171	.1558	.1433	.00256	
	30.01	34.94	50.69	50.14	16.62	39.64	76.74	87.92	17.54	

TABLA DE RESULTADOS 3.

$$Y = -160000 + 800x_1 + 800x_2 - 2x_1^2 - 2x_2^2 + \varepsilon$$

$$x_{op} = (200, 200)$$

1° DM( $\hat{\mu}$ )      2° RV

Diseño simplex

3° DM( $\hat{\sigma}$ )      4°  $\bar{n}$

$$A = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M2T	M2aT	M2bT	M2cT	
.1	.00017	1154	1218	1295	1097	.00456	.00500	.00560	4.8E-9	1° 2° 3° 4°
	.00173	11542	12184	12955	10972	.04561	.05003	.05609	4.8E-8	
	.00026	20.12	21.04	21.70	20.29	.0413	.0419	.0458	.00008	
	30.07	23.40	9.77	9.99	9.72	49.32	41.69	40.68	16.12	
.2	.00026	1288	2703	2357	1088	.00572	.01001	.01081	5.2E-8	
	.00127	64.43	13517	11788	5432	.0286	.0501	.0574	2.6E-7	
	.00090	21.54	31.18	29.00	19.11	.0494	.0594	.0613	.00015	
	30.02	89.03	13.51	13.42	13.24	99.99	46.58	46.84	16.84	
.3	.00043	1088	35785	3287	1021	.0168	.01082	.01721	5.2E-8	
	.00146	3631	12260	10955	3436	.0564	.0662	.0574	1.7E-7	
	.01387	18.08	35.89	35.20	16.16	.0705	.0849	.0783	.00017	
	30.20	120.8	20.44	18.57	13.32	118.2	49.86	53.94	16.90	
.4	.00081	1162	52551	54388	12174	.01897	.02287	.02241	5.2E-8	
	.00128	29055	13139	13594	30431	.04742	.05621	.05613	1.5E-7	
	.01414	21.16	44.72	44.93	21.50	.07971	.09350	.09094	.00018	
	30.47	112.5	22.20	20.66	13.59	133.0	56.82	59.12	17.10	
.5	.00081	888.1	83387	70.98	1102	.01892	.02494	.03138	1.1E-7	
	.00162	1916	10676	14192	2204	.03984	.04981	.06264	2.1E-7	
	.01834	16.04	46.11	47.81	19.73	.0893	.0989	.01030	.00022	
	30.27	137.3	26.16	26.01	14.67	151.02	60.28	65.42	17.18	
.6	.00087	823.0	13068	97244	913	.02014	.0345	.0448	1.0E-7	
	.00161	1536	21781	16207	1502	.03023	.0573	.0746	1.7E-7	
	.01854	18.86	63.96	57.81	15.23	.0921	.1127	.1201	.00029	
	30.40	139.4	29.59	30.01	15.21	153.36	64.86	63.79	17.10	
.7	.00108	781	10318	10966	1136	.0390	.0493	.0433	1.1E-7	
	.00189	1115	14740	15666	1622	.0621	.0700	.0619	2.3E-7	
	.02034	13.45	60.03	61.82	20.68	.1031	.1344	.1249	.00026	
	31.47	172.8	38.29	41.66	15.30	148.81	71.85	68.33	17.30	
.8	.00131	888	11000	16585	800.1	.0264	.0408	.0463	2.7E-7	
	.00163	1234	13750	20731	1000	.0335	.0511	.0579	3.3E-7	
	.02283	18.09	61.8	70.33	11.05	.09254	.1207	.1268	.00030	
	30.80	149.5	36.28	40.57	15.75	180.5	84.48	76.09	17.34	
.9	.00128	883	19603	12618	812	.0362	.06523	.0512	2.4E-7	
	.00140	1103	21781	14020	902	.0402	.0724	.0613	2.7E-7	
	.02234	18.30	78.44	67.81	14.17	.1199	.1515	.1453	.00029	
	31.37	166.3	43.32	46.73	15.75	168.16	86.48	79.07	17.54	
1.0	.00173	788	17190	19000	796	.0321	.0690	.0534	1.9E-7	
	.00173	788	17190	19000	796	.0321	.0690	.0534	1.9E-7	
	.02574	12.86	72.12	76.74	14.50	.1131	.0157	.1403	.00031	
	32.01	166.9	56.52	42.44	16.01	189.7	91.83	96.92	17.32	

TABLA DE RESULTADOS 4.

$$Y = -1295 + 6.48x_1 + 6.48x_2 - 1.0081x_1^2 + 1.983x_1x_2 - 1.0081x_2^2 + \epsilon$$

$x_{op} = (200, 200)$

1° DM( $\bar{y}$ )      2° RV

Diseño factorial.

3° DM( $\bar{x}_i$ )      4°  $\bar{N}$

$$A = \begin{bmatrix} -1 & 1 \\ .09 & .09 \end{bmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M2T	M2aT	M2bT	M2cT	
.1	.21115	21354	21310	21646	1284				2.7E-6	1°
	2.1145	213542	213107	216469	12844				2.7E-5	2°
	3.5756	466.4	473.	457.6	374.				.00779	3°
	1353	13.0	11.12	11.09	11.1				26.19	4°
.2	.25227	9421	5034	1068599	1867				3.1E-8	
	1.2734	47103	25171	5342998	9337				1.6E-5	
	3.9061	412.79	360.51	986.43	347.6	Para estos			.00879	
	1327	35.52	12.44	12.72	12.64				26.62	
.3	.28017	4652	4105341	694171	1242				2.6E-5	
	9.3632	15501	1.7E+7	2.1E+6	4144	no hubo			8.8E-5	
	4.1047	314.1	1713.53	1054	2243				.01580	
	1311	48.88	13.88	14.89	13.22	Convergencia			26.58	
.4	.32511	1256	51123.9	21262	1409				5.5E-6	
	8.1543	3141	127832	53157.7	3524				1.38E-5	
	4.4164	268.4	556.65	438.7	549.5				.01142	
	1288	74.12	17.84	17.26	14.03				26.42	
.5	.35247	4524	10714	12776	868				6.4E-3	
	7.2484	9248	21428	25552	4343				.01285	
	4.6434	323	412.1	398.2	176.54				.12781	
	1274	86.38	19.24	19.44	15.76				24.988	
.6	.37513	4105	58011	4401	3425				2.6E-3	
	6.2639	6843	96686	7335	5708				.00445	
	4.7234	288.4	539	321	192				.08299	
	12697	136.6	21.44	21.60	14.14				25.597	
.7	.38597	5369	19960	2754	1395				6.7E-4	
	5.5137	7775	28515	39354	1993				9.6E-4	
	4.7984	339.7	441	319	169				.04719	
	1364	97.57	21.44	20.20	15.42				26.12	
.8	.39742	48625	36538	34101	1634				8.2E-3	
	4.8677	60782	45797	4262	2043				.01025	
	4.9233	6649	399	3457	171				.1034	
	1352	540.1	26.24	28.0	17.2				25.57	
.9	.40182	58872	1830	5200	2519				6.0E-3	
	4.4819	66532	2034	2778	2799				6.7E-3	
	4.9534	6872	288	346	156				.16350	
	1401	385.12	28.6	29.48	16.95				25.83	
1.0	.48134	12541	2662	1864	2551				1.3E-2	
	4.8134	12541	2662	1864	2551				1.3E-2	
	5.3284	1255.4	311	303	102				.15054	
	1395	569.4	29.94	32.76	17.74				24.58	

TABLA DE RESULTADOS 5.

$$Y = -630 + 49.59x_1 + 50.40x_2 - 1.008x_1^2 - 7.9838x_1x_2 + \epsilon$$

$$x_{op} = (0,25)$$

1° DM( $\hat{y}$ )      2° RV

Diseño factorial.

3° DM( $x_1$ )      4°  $\bar{N}$

$$A = \begin{bmatrix} -1 & 1 \\ .09 & .09 \end{bmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M3	M3a	M3b	M3c	
.1	.20698	395	726	1905	630				9.2E-4	1° 2° 3° 4°
	2.0698	3962	7264	19059	6301				9.2E-3	
	3.8431	55.43	71.07	143.47	75.4				.16169	
	1297	11.0	11.0	11.0	11.6				20.19	
.2	.24917	2130	1259	395	573				1.8E-3	
	1.2221	1065	6298	1979	2869				7.8E-3	
	3.8261	35.57	87.1	59.54	64.1				.21413	
	1327	13.66	12.48	12.72	12.94				20.24	
.3	.27831	125	245	1504	461				2.3E-3	
	.93130	416	819	5027	1538				7.8E-3	
	4.0797	21.9	45.56	73.39	69.97	Para éstos			.27930	
	1321	15.98	14.82	14.92	13.43				20.11	
.4	.31813	1158	3820	529	473				3.4E-3	
	.79140	2893	9554	1323	1184	no hubo			8.2E-3	
	4.3412	82.83	95.49	62.67	50.24	Convergencia			.31152	
	1297	17.52	16.88	16.52	13.03				20.20	
.5	.35811	438	9654	4476	311				5.0E-3	
	.71351	872	16091	7461	622				1.0E-2	
	4.8645	64.4	143	142	25				.40481	
	1279	19.8	20.24	19.12	16.85				20.42	
.6	.39811	438	9654	4476	261				5.5E-3	
	.51937	732	16091	746	435				9.2E-3	
	4.8212	64.86	143	142	77				.42789	
	1271	20.76	20.36	22.60	16.94				20.36	
.7	.35187	308	387412	3713	274				8.3E-3	
	.51716	437	554161	5305	391				1.2E-2	
	4.8124	41.45	381	110	64				.55289	
	1338	22.38	26.12	27.20	16.72				20.65	
.8	.37441	312	724178	4712	261				7.4E-3	
	.48804	390	905221	5889	326				9.3E-5	
	4.7589	45.83	352	118	167				.4855	
	1317	28.3	26.1	27.5	16.2				20.73	
.9	.37182	188	19059	3805	242				8.8E-3	
	.41278	219	21176	4228	269				9.8E-3	
	4.8514	38.9	180	93	37				.56650	
	1299	24.23	32.6	28.48	17.25				20.77	
1.0	.38112	178	4686	36494	255				9.2E-3	
	.38112	178	4686	36494	255				9.2E-3	
	4.8552	28.8	130	227	111				.52654	
	1304	24.84	34.04	35.76	17.40				20.78	

TABLA DE RESULTADOS 6.

$$Y = -160000 + 8000x_1 + 8000x_2 - 1.0081x_1^2 - 1.0081x_2^2 - 1.9838x_1x_2 + \varepsilon$$

$x_{op} = (200, 200)$       1° DM( $x_1$ )      2° RV

Diseño factorial:

3° DM( $x_2$ )      4° N

$$A = \begin{bmatrix} -.09 & .09 \\ 1 & 1 \end{bmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M2T	M2aT	M2bT	M2cT	
.1	20915	5.4E+5	9.3E+4	9.4E+3	1.5E+5				1.4E-5	1°
	2.0915	5.4E+6	9.3E+5	9.4E+4	1.5E+6				1.4E-5	2°
	3.5606	710	613	402	640				.00612	3°
	1356	11.0	11.0	11.0	11.3				20.19	4°
.2	24527	2.3E+5	7.9E+4	1.4E+5	1.4E+5				1.4E-5	
	12224	1.2E+6	3.9E+5	6.7E+5	6.9E+5	Para	éstos		1.5E-5	
	3.8461	673	673	741	2091				.00944	
	1331	35.92	12.72	12.72	12.64	no			20.04	
.3	27839	9.8E+4	5.1E+4	2.5E+5	1.1E+5				2.9E-5	
	93132	3.2E+5	1.7E+5	8.4E+5	4.1E+5	hubo			1.2E-5	
	4.0857	630	639	1705	1145				.01070	
	1317	50.08	14.82	13.92	13.43	Convergencia			20.21	
.4	317	9.6E+3	4.3E+5	4.5E+5	1.1E+5				3.7E-5	
	79253	2.4E+4	1.1E+6	1.1E+6	2.7E+5				1.2E-5	
	4.3464	517	1172	1128	512				.01232	
	1292	89.72	18.84	18.26	14.03				20.35	
.5	35677	7.2E+6	3.5E+5	3.2E+5	1.0E+5				4.8E-5	
	71354	1.4E+7	7.0E+5	6.5E+5	2.0E+5				1.2E-5	
	4.5774	826	899	1189	556				.01371	
	1277	89.8	17.24	18.12	14.85				20.37	
.6	35583	8.4E+6	3.0E+8	2.4E+4	8.6E+4				5.9E-5	
	51938	1.4E+7	5.0E+8	4.1E+4	1.4E+5				9.8E-5	
	4.6234	1140	9473	505	1387				.01389	
	12727	122.6	25.54	22.60	15.14				20.57	
.7	36187	6.5E+6	3.2E+6	6.7E+5	7.6E+4				5.9E-5	
	51717	9.2E+6	4.6E+6	9.5E+5	1.1E+5				1.0E-5	
	4.6524	918	1769	1044	386				.01559	
	1341	116.8	24.44	23.20	16.42				20.48	
.8	37442	1.3E+5	5.9E+6	4.7E+7	7.8E+4				7.3E-5	
	46807	1.6E+5	7.3E+6	5.9E+7	9.8E+4				1.1E-5	
	4.7683	724	911	4891	584				.0179	
	1312	128.6	25.5	25.5	16.2				20.57	
.9	37152	8.8E+4	5.6E+7	1.8E+6	8.2E+4				8.5E-5	
	41282	9.9E+4	6.3E+7	2.2E+6	9.1E+4				1.4E-6	
	4.8544	616	3019	2055	735				.02050	
	1296	133.6	30.6	29.48	15.25				20.62	
1.0	36834	2.3E+3	8.5E+6	4.2E+7	4.5E+4				9.4E-5	
	36834	2.3E+3	8.5E+6	4.2E+7	4.5E+4				9.4E-6	
	4.8812	463	2883	4005	410				.01734	
	1365	141.4	34.4	27.76	18.40				20.62	

TABLA DE RESULTADOS 7.

$$Y = -630.06 + 4.9x_1 + 50.4x_2 - 1.0081x_1^2 - 1.0081x_2^2 - 1.9838x_1x_2 + c$$

$x_{op} = (0, 25)$

1° DM( $x_1$ )      2° RV

Diseño factorial.

3° DM( $x_2$ )      4° M

$$A = \begin{bmatrix} -.09 & .09 \\ 1 & 1 \end{bmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M2	M2aT	M2bT	M2cT
.1	.20878	6.2E+2	9.3E+3	2.3E+2	6.3E+2				6.9E-4
	.02069	6.2E+3	9.3E+4	2.3E+3	6.3E+3				6.9E-4
	3.5421	80	116	55	81				.61412
	1299	11.27	11.12	11.12	11.0				20.19
.2	.24527	8.9E+3	1.0E+3	1.4E+3	6.0E+2				1.6E-3
	.12224	4.5E+4	5.1E+3	6.9E+3	3.0E+3				8.8E-3
	3.8461	107	77	66	115	Para estos			.19844
	1331	13.92	12.72	12.72	11.64				20.13
.3	.27838	5.6E+2	8.7E+2	2.9E+3	4.8E+2				2.4E-3
	.93132	1.8E+3	2.9E+3	9.8E+3	1.6E+3	no hubo			8.7E-3
	4.0857	73	70	82	84				.27970
	1317	15.08	14.82	14.82	13.43	Convergencia			20.21
.4	.317	2.6E+2	9.8E+3	6.0E+2	4.2E+2				3.4E-6
	.79253	6.2E+2	2.4E+4	1.5E+3	1.1E-3				8.4E-5
	4.3464	43	130	55	36				.31032
	1292	17.72	17.84	16.26	13.03				20.30
.5	.35677	1.0E+4	2.8E+3	2.0E+3	3.5E+2				4.6E-3
	.71354	2.0E+4	5.6E+3	4.0E+3	7.0E+2				9.3E-3
	4.5774	117	91	67	30				.36871
	1277	18.8	17.24	18.12	14.85				20.50
.6	.39983	8.4E+3	1.1E+4	1.2E+3	3.1E+2				6.5E-3
	.51938	1.4E+5	1.9E+4	1.9E+3	5.2E+2				1.1E-2
	4.6234	66	136	65	43				.43789
	12727	20.96	22.54	22.60	16.14				20.47
.7	.38187	5.5E+1	5.7E+3	5.9E+3	2.2E+2				9.3E-3
	.51717	7.8E+1	8.1E+3	7.9E+3	3.5E+2				1.3E-2
	4.6524	26	108	107	32				.59159
	1341	20.88	22.44	23.20	17.42				20.78
.8	.37442	1.3E+2	1.6E+5	1.7E+4	2.8E+2				6.7E-3
	.46807	1.6E+2	2.1E+5	2.2E+4	3.5E+2				8.3E-5
	4.7683	43	253	93	33				.5109
	1312	23.26	26.5	26.3	16.2				20.62
.9	.37152	3.9E+1	1.4E+4	2.0E+4	2.4E+2				7.8E-6
	.41282	4.4E+1	1.5E+4	2.2E+4	2.6E+2				8.6E-6
	4.8544	29	245	208	19				5.650
	1295	24.16	28.6	23.48	17.25				20.86
1.0	.38212	2.3E+1	2.2E+4	6.3E+3	2.4E+2				1.0E-2
	.38212	2.3E+1	2.2E+4	6.3E+3	2.4E+2				1.0E-2
	4.8792	24	187	116	23				.58134
	1308	23.24	34.4	29.76	17.40				20.84

TABLA DE RESULTADOS 8.

$$Y = -160000 + 800x_1 + 800x_2 - 1.0081x_1^2 - 1.0081x_2^2 - 1.9638x_1x_2 + s$$

$x_{op} = (200, 200)$       1° DM(.....)      2° RV

Diseño factorial.

3° DM(x<sub>1</sub>)      4° N

$$A = \begin{bmatrix} -.09 & .09 \\ 1 & 1 \end{bmatrix}$$

$\sigma^2$	M1	M2	M2a	M2b	M2c	M3	M3a	M3b	M3c
.1	.20915	5.4E+5	9.3E+4	9.4E+3	1.5E+5				1.4E-6
	2.0915	5.4E+6	9.3E+5	9.4E+4	1.5E+6				1.4E-5
	3.5606	710	613	402	640				.00612
	1356	11.0	11.0	11.0	11.3				20.19
.2	.24527	2.3E+5	7.9E+4	1.4E+5	1.4E+5				1.4E-6
	.12224	1.2E+6	3.9E+5	6.7E+5	6.9E+5				1.5E-5
	3.8461	673	673	741	2091	Para estos			.00944
	1331	35.92	12.72	12.72	12.64				20.04
.3	.27939	9.8E+4	5.1E+4	2.5E+5	1.1E+5				2.9E-6
	.93132	3.2E+5	1.7E+5	8.4E+5	4.1E+5	no hubo			1.2E-5
	4.0857	630	639	1705	1145				.01070
	1317	50.08	14.82	13.92	13.43	Convergencia			20.21
.4	.317	9.8E+3	4.3E+5	4.5E+5	1.1E+5				3.7E-6
	.79253	2.4E+4	1.1E+6	1.1E+6	2.7E+5				1.2E-5
	4.3464	517	1172	1128	512				.01232
	1292	89.72	18.84	18.26	14.03				20.35
.5	.35677	7.2E+6	3.5E+5	3.2E+5	1.0E+5				4.8E-6
	.71354	1.4E+7	7.0E+5	6.5E+5	2.0E+5				1.2E-5
	4.5774	826	899	1159	556				.01371
	1277	89.8	17.24	18.12	14.85				20.37
.6	.35583	8.4E+6	3.0E+8	2.4E+4	8.6E+4				5.9E-6
	.51938	1.4E+7	5.0E+8	4.1E+4	1.4E+5				9.8E-5
	4.6234	1140	9473	505	1387				.01389
	12727	122.6	25.54	22.60	15.14				20.57
.7	.36187	6.5E+6	3.2E+6	6.7E+5	7.6E+4				5.9E-6
	.51717	9.2E+6	4.6E+6	9.5E+5	1.1E+5				1.0E-5
	4.6524	918	1769	1044	386				.01559
	1341	116.8	24.44	23.20	16.42				20.48
.8	.37442	1.3E+5	5.9E+6	4.7E+7	7.8E+4				7.3E-6
	.46807	1.6E+5	7.3E+6	5.9E+7	9.8E+4				1.1E-5
	4.7683	724	911	4891	584				.0179
	1312	128.6	25.5	25.5	16.2				20.57
.9	.37152	8.9E+4	5.6E+7	1.9E+6	8.2E+4				8.5E-6
	.41282	9.9E+4	6.3E+7	2.2E+6	9.1E+4				1.4E-6
	4.8544	616	3019	2055	735				.02050
	1296	133.6	30.6	29.48	15.25				20.62
1.0	.36834	2.3E+3	8.5E+6	4.2E+7	4.5E+4				9.4E-6
	.36834	2.3E+3	8.5E+6	4.2E+7	4.5E+4				9.4E-6
	4.6812	463	2883	4005	410				.01734
	1365	141.4	34.4	27.76	18.40				20.62

1°  
2°  
3°  
4°

TABLA DE RESULTADOS 9

Tiempo de proceso acumulado.

Cada renglon equivale al tiempo que se requirió para obtener los resultados de cada una de las tablas anteriores.

El tiempo está medido en notacion estandar: (') significa minutos; (") significa segundos.

A y $\bar{x}_{op}$	M1	M2	M2a	M2b	M2c	M3	M3a	M3b	M3c
$\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ (25, 25)	40"		1' 51"		1' 04"		4' 29"		1' 20"
$\begin{pmatrix} -1 & 1 \\ .09 & .09 \end{pmatrix}$ (0, 25)	1' 01"		2' 42"		1' 08"		no hubo convergencia		2' 44"
$\begin{pmatrix} -.09 & .09 \\ 1 & 1 \end{pmatrix}$ (25, 25)	1' 13"		2' 01"		1' 08"		no hubo convergencia		2' 44"
$\begin{pmatrix} -.09 & .09 \\ 1 & 1 \end{pmatrix}$ (0, 25)	2" 00"		1' 51"		1' 04"		4' 29"		2' 44"
$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ (200, 200)	1' 31"		2' 42"		1' 08"		no hubo convergencia		2' 18"
$\begin{pmatrix} 1 & 1 \\ .09 & .09 \end{pmatrix}$ (200, 200)	7' 20"		2' 02"		1' 08"		no hubo convergencia		2' 18"
$\begin{pmatrix} -.09 & .09 \\ 1 & 1 \end{pmatrix}$ (200, 200)	7' 40"		4' 12"		1' 40"		8' 00"		2' 54"