



Universidad Nacional
Autónoma de México

4
Z E J
EST
EST

Facultad de Ciencias

Un algoritmo para regresión
ortogonal no-lineal

Tesis

Que para obtener el título de

ACTUARIO

presenta

Claudia Alquicira Esquivel

México D.F.



1995

FACULTAD DE CIENCIAS
SECCION ESCOLAR

FALLA DE ORIGEN

**TESIS CON
FALLA DE ORIGEN**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

M. en C. Virginia Abrín Batule
Jefe de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis:

"UN ALGORITMO PARA REGRESION ORTOGONAL NO-LINEAL"

realizado por ALQUICIRA ESQUIVEL CLAUDIA

con número de cuenta 8630855-1 , pasante de la carrera de ACTUARIA

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
Propietario DR. PABLO BARRERA SANCHEZ
Propietario DR. JESUS LOPEZ ESTRADA
Propietario DR. HUMBERTO MADRID DE LA VEGA
Suplente DR. DARIO MORENO OSORIO
Suplente DRA. REBECA ROMERO VAZQUEZ

[Handwritten signatures on lines]
Humberto Madrid
Dario Moreno
Rebeca Romero



Consejo Departamental de Matemáticas

FACULTAD DE CIENCIAS
M. en C. ALEJANDRO BRAVO MOJICA
COORDINADOR DE LICENCIATURA
MATEMÁTICAS

A mis padres
Rafa y Tere.

A mis hermanos
Rosy, Feli, Tere y Mary.

A mis sobrinos
Paola, Anita y Rafa.

Por motivarme a seguir adelante,
quiero agradecer a mi profesora
de primaria. Gracias

Profa. Hilda Martínez del Valle.

Por enseñarme que el mundo no
es como pensaba y por darme tu
apoyo en los diferentes retos que
se me han presentado. Gracias

Dr. Pablo Barrera Sánchez.

Indice

Introducción.	v
I Problema de ajuste de parámetros del modelo lineal.	1
I.1 Planteamiento del problema de ajuste ordinario de parámetros del modelo lineal.	1
I.2 Solución del problema de ajuste de parámetros ordinario del modelo lineal.	4
I.3 Planteamiento del problema de ajuste ortogonal de parámetros del modelo lineal.	5
I.4 Solución del problema de ajuste ortogonal de parámetros del modelo lineal.	8
I.4.1 Análisis de la solución del ajuste ortogonal lineal.	10
I.5 Ejemplo	13
II Problema de ajuste de parámetros del modelo múltiple.	15
II.1 Planteamiento del problema de ajuste ordinario de parámetros del modelo múltiple.	15
II.2 Solución del problema del ajuste ordinario de parámetros al modelo múltiple	19
II.3 Planteamiento del problema de ajuste ortogonal de parámetros del modelo múltiple.	21
II.4 Solución del problema de ajuste ortogonal de parámetros del modelo múltiple.	25
II.4.1 Ejemplos.	33

III Métodos de solución para el problema de mínimos cuadrados no-lineales.	39
III.1 Método de Gauss-Newton.	40
III.2 Método de Gauss-Newton Modificado.	41
III.3 Método de Levenberg-Marquardt.	41
III.3.1 Solución del problema estructurado del problema de mínimos cuadrados lineales.	47
III.3.2 Actualización del radio de la región de confianza.	51
III.3.3 El parámetro de Levenberg-Marquardt.	53
III.3.4 Resultados teóricos.	63
IV Problema de ajuste ortogonal de parámetros del modelo no-lineal.	65
IV.1 Planteamiento del problema de ajuste ortogonal del modelo lineal general.	65
IV.2 Solución del problema de ajuste ortogonal del modelo lineal general.	68
IV.3 Ajuste ortogonal de parámetros del modelo no-lineal.	74
V O D R P A C K	77
V.1 Especificando la tarea.	78
V.2 Ajuste ortogonal a modelos implícitos.	79
V.3 Valores iniciales.	80
V.4 Pesos.	81
V.5 Valores por omisión.	81
V.6 E j e m p l o s.	82
A Producto de Kronecker.	91
B Propiedades de las funciones cuadráticas.	93
C Métodos de región de confianza.	103

D Cálculos adicionales.	107
D.1 Cálculo del cociente de reducción.	107
D.2 Cálculos del ajuste de la cuadrática $q(\theta)$	109
D.3 Cálculo de la derivada de $\phi(\lambda)$	110
Bibliografía y referencias	115

Introducción.

El problema de ajuste de un modelo a un conjunto de datos con errores en las observaciones tiene una rica historia y una considerable literatura. Sin embargo, los problemas donde también existen errores en las variables independientes han recibido poca atención.

Los errores en las variables independientes siempre ocurren, pero generalmente son ignorados en las técnicas de mínimos cuadrados lineales y no-lineales. Aunque cabe mencionar, que si estos errores son relativamente pequeños con respecto a los de las observaciones de la variable dependiente entonces ignorarlos usualmente no degrada seriamente la precisión de los parámetros estimados.

En algunas áreas, las técnicas de medición son tan precisas que los errores en las variables independientes son significativos comparados con los de la variable dependiente. Un ejemplo es la calibración de dispositivos electrónicos, medidores de flujo y calorímetros, otra clase de ejemplos proviene del ajuste de curvas y superficies.

El objetivo de esta tesis es estudiar un algoritmo eficiente para el problema de ajuste de un modelo a un conjunto de datos, considerando que todas las variables están sujetas a error.

En el capítulo I desarrollamos el ajuste de parámetros al modelo lineal considerando solamente errores en las ordenadas (ajuste ordinario) y paralelamente las ideas del ajuste considerando errores en ambas variables (ajuste ortogonal), ésto lo hacemos con el objeto de introducir al lector a las dificultades del problema y de como superarlas, para hacer más fácil de entender el caso general.

En el capítulo II generalizamos las ideas del capítulo I tanto para el ajuste ordinario como para el ajuste ortogonal al modelo multiple, cuya presentación es nueva.

En el capítulo III consideramos el problema ordinario de mínimos cuadrados no-lineales; en primer lugar tratamos brevemente el método de Gauss-Newton y Gauss-Newton Modificado y posteriormente el de Levenberg-Marquardt en forma extensa por ser la parte medular del algoritmo del ajuste ortogonal no-lineal.

En el capítulo IV desarrollamos las ideas del ajuste ortogonal del mode-

lo no-lineal, primero consideramos un ejemplo para ilustrar las ideas que permiten simplificar el problema y convertirlo a un problema ordinario de mínimos cuadrados no-lineales, así el caso general no-lineal se sigue de inmediato.

En el capítulo V presentamos una breve reseña de un paquete de subrutinas en FORTRAN, ODRPACK, que realiza ajustes tanto ordinarios como ortogonales usando el algoritmo descrito en este trabajo, con él resolvimos algunos ejemplos que presentamos al final de este capítulo. Este paquete fue hecho por Paul Boggs [6] y sus colaboradores en el National Institute of Standards and Technology y está disponible en la INTERNET (netlib).

Finalmente presentamos algunos apéndices que complementan el material desarrollado en este trabajo.

El proyecto que nos deja este material es realizar un programa que sea de uso fácil para las personas que lo quieran aplicar a casos sencillos.

Capítulo I

Problema de ajuste de parámetros del modelo lineal.

De los posibles modelos que pueden ser ajustados a un conjunto de datos, el más común es el modelo lineal; ya sea por simplicidad o porque es una buena elección de ajuste para los datos. En nuestro caso quisimos discutir este tipo de ajuste como una introducción al problema de ajuste de parámetros del modelo no-lineal.

El problema de ajuste del modelo lineal lo plantearemos en base a dos ideas, una considerando las distancias verticales y otra considerando las distancias ortogonales de los datos al modelo, obtendremos la solución en ambos casos, para que finalmente comparemos similitudes y diferencias; así como establecer conclusiones.

I.1 Planteamiento del problema de ajuste ordinario de parámetros del modelo lineal.

Un problema clásico en el área de Estadística es el ajuste de parámetros a una recta bajo ciertas suposiciones, nosotros trataremos este problema en los aspectos numéricos. Nuestro objetivo para este efecto, es ajustar al conjunto de datos coordinados, de la forma :

$$\{(x_i, y_i) : \text{para } i = 1, \dots, n\} \quad (i.1.1)$$

al modelo lineal:

$$m(x; \bar{\beta}) = \beta_0 + \beta_1 x \quad (i.1.2)$$

donde $x \in \mathbb{R}$ y $\bar{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$.

Para desarrollar las ideas del planteamiento de este problema, nos apoyaremos en las ideas gráficas, por lo que pensemos que la recta que ajusta a los datos (i.1.1) es $\ell = \beta_0 + \beta_1 x$ y denotemos con ε_i a la distancia sobre las ordenadas (distancia vertical) del i -ésimo dato a la recta ℓ y que gráficamente lo observamos así:

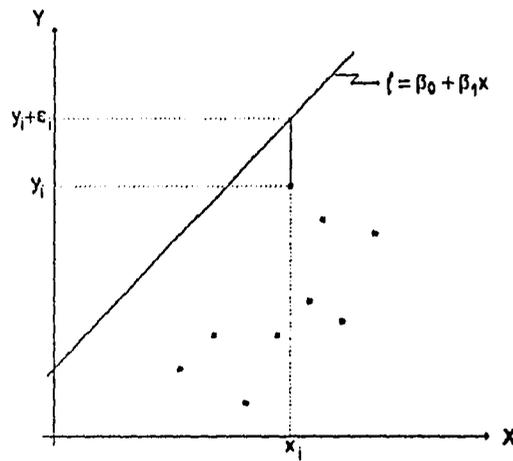


Figura I.1: Idea gráfica del ajuste ordinario a una recta.

Si consideramos el conjunto coordinado $\{(x_i, \ell_i) : \text{para } i = 1, \dots, n\}$ donde $\ell_i = \beta_0 + \beta_1 x_i$, es clara la siguiente igualdad:

$$\varepsilon_i = y_i - \ell_i \quad \text{para } i = 1, \dots, n \quad (i.1.3)$$

de esta igualdad obtenemos:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad \text{para } i = 1, \dots, n \quad (i.1.4)$$

a partir de este momento utilizamos notación vectorial, por lo que denotamos:

$$\bar{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad \bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \text{y} \quad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (i.1.5)$$

así la expresión escalar (i.1.4) la escribimos en forma vectorial como:

$$\bar{y} = \beta_0 \bar{1} + \beta_1 \bar{x} + \bar{\epsilon}$$

el simbolo $\bar{1}$ representa el vector unitario, cuyas entradas son el escalar uno, nuestro propósito es calcular β_0 y β_1 de tal manera que las distancias, ϵ_i , sean lo más pequeñas posibles. Uno de los posibles caminos para obtener la recta que mejor ajusta a los datos es minimizar la norma al cuadrado del vector $\bar{\epsilon}$, es decir, planteamos el problema con restricciones como :

$$\min_{\beta, \bar{\epsilon}} \frac{1}{2} \|\bar{\epsilon}\|^2$$

$$\text{sujeto a } \bar{y} = \beta_0 \bar{1} + \beta_1 \bar{x} + \bar{\epsilon}$$

de la restricción despejamos $\bar{\epsilon}$ y sustituimos, así obtenemos el siguiente problema de minimización sin restricciones:

$$\min_{\beta} \frac{1}{2} \|\bar{y} - (\beta_0 \bar{1} + \beta_1 \bar{x})\|^2 \quad (i.1.6)$$

el factor un medio es sólo para eliminar constantes al momento de derivar para resolver este problema, lo que haremos en la siguiente sección.

En general, el vector $\bar{\epsilon}$ se le conoce como residuo, pero lo denotaremos como $\bar{r}(\beta)$:

$$\bar{r}(\beta) = \bar{y} - (\beta_0 \bar{1} + \beta_1 \bar{x}) \quad (i.1.7)$$

aunque es una función que depende del vector $\bar{\beta}$, algunas veces sólo la escribiremos como \bar{r} , bajo estos términos otra manera de escribir el problema (i.1.6) es:

$$\min_{\beta} \frac{1}{2} \|\bar{r}(\beta)\|^2$$

en las secciones subsecuentes el término de residuo nos será de gran utilidad, en especial para fijar ideas.

I.2 Solución del problema de ajuste de parámetros ordinario del modelo lineal.

Recordemos, que para encontrar un mínimo de una función, obtenemos los puntos críticos, derivando dicha función e igualando su derivada a cero, en este caso como es una función de \mathbb{R}^n a \mathbb{R} calculamos el gradiente de la función :

$$f(\bar{\beta}) = \frac{1}{2} \| \bar{y} - (\beta_0 \bar{\mathbf{1}} + \beta_1 \bar{\mathbf{x}}) \|^2 \quad (i.2.1)$$

entonces el problema (i.1.6) se simplifica a:

$$\min_{\bar{\beta}} f(\bar{\beta}) \quad (i.2.2)$$

si calculamos las derivadas parciales de la función f con respecto a los parámetros β_0 y β_1 , obtenemos.

$$\frac{\partial f}{\partial \beta_0} = -\bar{\mathbf{1}}^t \cdot [\bar{y} - (\beta_0 \bar{\mathbf{1}} + \beta_1 \bar{\mathbf{x}})]$$

$$\frac{\partial f}{\partial \beta_1} = -\bar{\mathbf{x}}^t \cdot [\bar{y} - (\beta_0 \bar{\mathbf{1}} + \beta_1 \bar{\mathbf{x}})]$$

las igualamos a cero, con esto obtenemos un sistema de ecuaciones lineales para los parámetros de la forma:

$$\beta_0 \bar{\mathbf{1}}^t \cdot \bar{\mathbf{1}} + \beta_1 \bar{\mathbf{1}}^t \cdot \bar{\mathbf{x}} = \bar{\mathbf{1}}^t \cdot \bar{y} \quad (i.2.3)$$

$$\beta_0 \bar{\mathbf{x}}^t \cdot \bar{\mathbf{1}} + \beta_1 \bar{\mathbf{x}}^t \cdot \bar{\mathbf{x}} = \bar{\mathbf{x}}^t \cdot \bar{y} \quad (i.2.4)$$

estas ecuaciones reciben el nombre de ecuaciones normales del problema de mínimos cuadrados lineales (i.2.2); despejamos β_0 de la ecuación (i.2.3), obtenemos:

$$\beta_0 = \frac{\bar{\mathbf{1}}^t \cdot \bar{y} - \beta_1 \bar{\mathbf{1}}^t \cdot \bar{\mathbf{x}}}{n} \quad (i.2.5)$$

como $\bar{\mathbf{1}} \in \mathbb{R}^n$ entonces $\bar{\mathbf{1}}^t \cdot \bar{\mathbf{1}} = n$; sustituyendo β_0 en la ecuación (i.2.4) tenemos el desarrollo:

$$\frac{(\bar{\mathbf{1}} \cdot \bar{y} - \beta_1 \bar{\mathbf{1}}^t \cdot \bar{\mathbf{x}})(\bar{\mathbf{x}} \cdot \bar{\mathbf{1}})}{n} + \beta_1 \bar{\mathbf{x}}^t \bar{\mathbf{x}} = \bar{\mathbf{x}}^t \cdot \bar{y}$$

$$(\bar{\mathbf{1}} \cdot \bar{y} - \beta_1 \bar{\mathbf{1}}^t \cdot \bar{\mathbf{x}})(\bar{\mathbf{x}} \cdot \bar{\mathbf{1}}) + n \beta_1 \bar{\mathbf{x}}^t \bar{\mathbf{x}} = n \bar{\mathbf{x}}^t \cdot \bar{y}$$

de aquí despejamos el parámetro β_1 :

$$\beta_1 = \frac{n\bar{x}^t \cdot \bar{y} - (\bar{1}^t \cdot \bar{y})(\bar{x}^t \cdot \bar{1})}{n\bar{x}^t \cdot \bar{x} - (\bar{1}^t \cdot \bar{x})(\bar{1}^t \cdot \bar{x})} \quad (i.2.5)$$

otra forma de escribir estas expresiones es en términos de las medias cuya definición es:

Definición : Sea $\bar{u} \in \mathbb{R}^l$, definimos como la media de \bar{u} como:

$$\mu_u = \frac{\bar{1}^t \cdot \bar{u}}{\ell}.$$

Una forma natural de extender esta definición es:

Definición : Sea $\bar{v}, \bar{u} \in \mathbb{R}^l$, definimos como la media de \bar{v} y \bar{u} como:

$$\mu_{vu} = \frac{\bar{v}^t \cdot \bar{u}}{\ell}.$$

Con estas definiciones, las expresiones para β_0 y β_1 son :

$$\beta_0 = \mu_y - \beta_1 \mu_x \quad (i.2.6)$$

$$\beta_1 = \frac{\mu_{xy} - \mu_x \mu_y}{\mu_{x^2} - \mu_x^2}. \quad (i.2.7)$$

I.3 Planteamiento del problema de ajuste ortogonal de parámetros del modelo lineal.

Como en la sección I.1 suponemos que al conjunto de datos :

$$\{(x_i, y_i) : \text{para } i = 1, \dots, n\} \quad (i.3.1)$$

les queremos ajustar un modelo lineal de la forma:

$$m(x; \beta) = \beta_0 + \beta_1 x \quad (i.3.2)$$

donde $x \in \mathbb{R}$ y $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$.

Habíamos dicho que nuestro propósito es ajustar una recta a los datos (i.3.1), pero ahora consideramos la distancia ortogonal del dato i a la recta, la cual la denotamos como d_i , imaginemos que la gráfica correspondiente es:

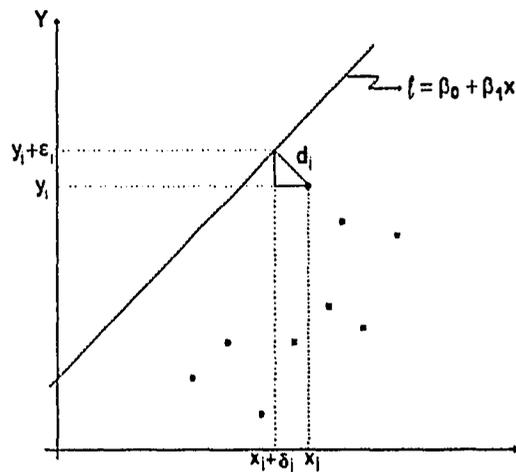


Figura I.2: Idea gráfica del ajuste ortogonal a una recta.

como podemos ver en la gráfica, ϵ_i es la distancia sobre la ordenada y δ_i es la distancia sobre la abscisa del dato i a la recta, si utilizamos el teorema de Pitágoras encontramos la siguiente expresión para d_i :

$$d_i^2 = \epsilon_i^2 + \delta_i^2 \quad \text{para } i = 1, \dots, n \quad (i.3.3)$$

Hemos mencionado que preferimos manejar vectores que escalares por facilidad y para evitar uso de índices confusos, además de la notación usada en (i.1.5), construimos con la misma idea los vectores $\bar{\mathbf{d}}$ y $\bar{\boldsymbol{\delta}}$ como:

$$\bar{\mathbf{d}} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \quad \text{y} \quad \bar{\boldsymbol{\delta}} = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{bmatrix}$$

como nuestro objetivo es encontrar la recta que mejor ajuste a los datos, en este caso también lo haremos en el sentido que las distancias ortogonales sean mínimas, por lo que minimizamos la norma del vector $\bar{\mathbf{d}}$ planteando el problema

como:

$$\min_{\bar{\beta}, \bar{\epsilon}, \bar{\delta}} \frac{1}{2} \|\bar{\mathbf{d}}\|^2 \quad (i.3.4)$$

$$\text{sujeto a } \bar{\mathbf{y}} = \beta_0 \bar{\mathbf{1}} + \beta_1 (\bar{\mathbf{x}} + \bar{\delta}) + \bar{\epsilon}$$

como $\bar{\epsilon}$ es el vector de parámetros que depende en forma lineal, lo despejamos de la restricción (i.3.4) y lo sustituimos en (i.3.3) para obtener:

$$d_i^2 = (y_i - [\beta_0 + \beta_1(x_i + \delta_i)])^2 + \delta_i^2 \quad \text{para } i = 1, \dots, n$$

consecuentemente (i.3.4) lo transformamos en un problema de minimización sin restricciones de la forma:

$$\min_{\bar{\beta}, \bar{\delta}} \frac{1}{2} \left\{ \|\bar{\mathbf{y}} - [\beta_0 \bar{\mathbf{1}} + \beta_1 (\bar{\mathbf{x}} + \bar{\delta})]\|^2 + \|\bar{\delta}\|^2 \right\} \quad (i.3.5)$$

observamos que resolver este problema de minimización implica obtener $(n + 2)$ parámetros, es decir, el número de parámetros será tan grande como el número de datos, lo cual complica considerablemente el problema.

En este caso también construimos la función:

$$\bar{\mathbf{r}}_\delta(\bar{\beta}, \bar{\delta}) = \bar{\mathbf{y}} - [\beta_0 \bar{\mathbf{1}} + \beta_1 (\bar{\mathbf{x}} + \bar{\delta})]$$

la cual la llamaremos como residuo ortogonal y así escribimos el planteamiento del problema (i.3.5) como:

$$\min_{\bar{\beta}, \bar{\delta}} \frac{1}{2} \left\{ \|\bar{\mathbf{r}}_\delta(\bar{\beta}, \bar{\delta})\|^2 + \|\bar{\delta}\|^2 \right\}$$

al vector $\bar{\mathbf{r}}_\delta$, lo podemos escribir en términos del residuo ordinario $\bar{\mathbf{r}}$ (definido en la sección I.1) de la siguiente forma:

$$\bar{\mathbf{r}}_\delta = \bar{\mathbf{r}} - \beta_1 \bar{\delta}$$

por lo que otra forma de escribir el problema (i.3.7) es:

$$\min_{\bar{\beta}} \frac{1}{2} \left\{ (\bar{\mathbf{r}} - \beta_1 \bar{\delta})' (\bar{\mathbf{r}} - \beta_1 \bar{\delta}) + \bar{\delta}' \bar{\delta} \right\}$$

esta expresión del planteamiento del problema la utilizaremos en la siguiente sección, para desarrollar un método de solución para este problema.

I.4 Solución del problema de ajuste ortogonal de parámetros del modelo lineal.

Para resolver el problema de ajuste de parámetros a una recta via distancias ortogonales, seguiremos el mismo procedimiento de la sección I.2. En este caso la función a minimizar es:

$$f_{\delta}(\eta) = \frac{1}{2} \{ (\bar{\mathbf{r}} - \beta_1 \bar{\delta})' (\bar{\mathbf{r}} - \beta_1 \bar{\delta}) + \bar{\delta}' \bar{\delta} \}$$

donde $\eta = \begin{pmatrix} \beta_1 \\ \bar{\delta} \end{pmatrix}$.

Por lo que el problema que resolvemos es:

$$\min_{\eta} f_{\delta}(\eta) \quad (i.4.1)$$

calculamos las derivadas parciales de la función f_{δ} como:

$$\frac{\partial f_{\delta}}{\partial \beta_0} = (\beta_1 \bar{\delta} - \bar{\mathbf{r}})' \cdot \bar{\mathbf{I}}$$

$$\frac{\partial f_{\delta}}{\partial \beta_1} = (\beta_1 \bar{\delta} - \bar{\mathbf{r}})' \cdot (\bar{\mathbf{x}} + \bar{\delta})$$

$$\frac{\partial f_{\delta}}{\partial \delta} = \beta_1 (\beta_1 \bar{\delta} - \bar{\mathbf{r}}) + \bar{\delta}$$

a estas derivadas parciales las igualamos a cero, así obtenemos el siguiente sistema de $(n + 2)$ ecuaciones para los parámetros β_0 , β_1 y $\bar{\delta}$:

$$(\beta_1 \bar{\delta} - \bar{\mathbf{r}})' \cdot \bar{\mathbf{I}} = 0 \quad (i.4.2)$$

$$(\beta_1 \bar{\delta} - \bar{\mathbf{r}})' \cdot (\bar{\mathbf{x}} + \bar{\delta}) = 0 \quad (i.4.3)$$

$$\beta_1 (\beta_1 \bar{\delta} - \bar{\mathbf{r}}) + \bar{\delta} = 0 \quad (i.4.4)$$

si aplicamos la transpuesta y multiplicamos por el vector uno a (i.4.4), además utilizamos la ecuación (i.4.2) obtenemos la siguiente relación:

$$\bar{\delta}' \cdot \bar{\mathbf{I}} = 0 \quad (i.4.5)$$

si ahora , de la ecuación (i.4.4) despejamos el vector $\bar{\delta}$ obtenemos:

$$\bar{\delta} = \frac{\beta_1}{\beta_1^2 + 1} \bar{r} \quad (i.4.6)$$

con esta igualdad y a partir de (i.4.5) concluimos que la suma de los residuos es igual a cero, que es un resultado muy conocido en Estadística. Observamos que podemos despejar β_0 de (i.4.2) y esto se nos facilita por la ecuación (i.4.5), así obtener:

$$\beta_0 = \frac{1}{n} \left(\bar{y}^t \cdot \bar{1} - \beta_1 \bar{x}^t \cdot \bar{1} - \beta_1 \bar{\delta}^t \cdot \bar{1} \right)$$

que escrita en términos de las medias queda como:

$$\beta_0 = \mu_y - \beta_1 \mu_x \quad (i.4.7)$$

que es la misma relación que obtuvimos en el ajuste ordinario de parámetros lineal, a diferencia de este, el cálculo de β_1 es diferente. Con esto, nos resta obtener una expresión para el parámetro β_1 , esto lo haremos de la ecuación (i.4.3), de la cual obtenemos:

$$\beta_1(\bar{x} + \bar{\delta})^t(\bar{x} + \bar{\delta}) - \bar{y}^t(\bar{x} + \bar{\delta}) + \beta_0 \bar{1}^t \cdot \bar{x} = 0$$

si ahora utilizamos (i.4.6) y (i.4.7) para simplificar esta igualdad, obtenemos una ecuación cuadrática para β_1 de la forma:

$$\frac{n}{(\beta_1^2 + 1)^2} [(\mu_{xy} - \mu_x \mu_y) \beta_1^2 + (\mu_x^2 - \mu_x^2 + \mu_y^2 - \mu_y^2) \beta_1 - \mu_{xy} + \mu_x \mu_y] = 0$$

como n es distinta de cero implica que β_1 satisface:

$$(\mu_{xy} - \mu_x \mu_y) \beta_1^2 + (\mu_x^2 - \mu_x^2 + \mu_y^2 - \mu_y^2) \beta_1 - \mu_{xy} + \mu_x \mu_y = 0$$

esta ecuación tiene siempre solución real, para observar esto, haremos las siguientes asignaciones:

$$\begin{aligned} a &= \mu_{xy} - \mu_x \mu_y \\ b &= \mu_x^2 - \mu_x^2 + \mu_y^2 - \mu_y^2 \end{aligned}$$

con esto la ecuación la escribimos como:

$$a\beta_1^2 + b\beta_1 - a = 0$$

calculamos el discriminante Δ de esta ecuación de segundo grado:

$$\Delta = b^2 + 4a^2$$

como siempre es positivo, obtenemos dos raíces reales; pero ahora nos enfrentamos a otro problema pues no sabemos cual de las dos soluciones minimiza las distancias ortogonales, por lo que decidimos sustituir las expresiones de β_0 y $\bar{\delta}$ en la función objetivo de (i.4.1) y así obtener una función de β_1 de la forma:

$$f(\beta_1) = \frac{1}{1 + \beta_1^2} \bar{\mathbf{r}}' \bar{\mathbf{r}} \quad (i.4.9)$$

donde $\bar{\mathbf{r}}$ depende tanto de β_0 como de β_1 por lo que sustituimos (i.4.8) en la expresión del residuo (i.1.7), por lo que obtenemos:

$$\bar{\mathbf{r}} = \hat{\mathbf{y}} - \hat{\mathbf{x}}\beta_1$$

que sólo depende de un parámetro, por lo que resolvemos el problema:

$$\min_{\beta_1} \frac{1}{1 + \beta_1^2} \bar{\mathbf{r}}' \bar{\mathbf{r}}.$$

otro camino que podemos tomar es graficar la función (i.4.9) y las dos soluciones de la cuadrática para saber cual corresponde al mínimo. Esto lo veremos en la sección I.5 mediante un ejemplo pero antes haremos un análisis de la solución en términos del coeficiente de correlación utilizado en Estadística.

I.4.1 Análisis de la solución del ajuste ortogonal lineal.

En esta sección analizamos la solución de la ecuación de segundo grado de β_1 obtenida anteriormente, que es:

$$(\mu_{xy} - \mu_x \mu_y) \beta_1^2 + (\mu_{x^2} - \mu_x^2 + \mu_y^2 - \mu_{y^2}) \beta_1 - \mu_{xy} + \mu_x \mu_y = 0 \quad (i.4.1.1)$$

con respecto a la correlación de la variables x y y , su definición en términos de nuestra notación es:

$$\rho = \frac{\mu_{xy} - \mu_x \mu_y}{\sqrt{\mu_{x^2} - \mu_x^2} \sqrt{\mu_{y^2} - \mu_y^2}} \quad (i.4.1.2)$$

para evitar confusiones, a la solución de β_1 del ajuste ordinario lineal, la denotamos como:

$$\overset{\circ}{\beta}_1 = \frac{\mu_{xy} - \mu_x \mu_y}{\mu_x^2 - \mu_x^2} \quad (i.4.1.3)$$

con esta notación escribimos (i.4.1.1) en términos de ϱ y $\overset{\circ}{\beta}_1$ como:

$$\beta_1^2 + \left(\frac{1}{\overset{\circ}{\beta}_1} - \frac{\overset{\circ}{\beta}_1}{\varrho^2} \right) \beta_1 - 1 = 0 \quad (i.4.1.4)$$

de esta ecuación concluimos la siguiente propiedad:

Prop: Los puntos críticos corresponden a dos rectas ortogonales entre sí.

Utilizamos la fórmula general para ecuaciones cuadráticas para obtener las dos soluciones de (i.4.1.4):

$$\beta_1 = \frac{-(\varrho^2 - \overset{\circ}{\beta}_1^2) \pm \sqrt{(\varrho^2 - \overset{\circ}{\beta}_1^2)^2 + 4 \overset{\circ}{\beta}_1^2 \varrho^4}}{2 \overset{\circ}{\beta}_1 \varrho^2}$$

si sólo trabajamos con el radical para simplificar obtenemos:

$$\begin{aligned} (\varrho^2 - \overset{\circ}{\beta}_1^2) + 4 \overset{\circ}{\beta}_1^2 \varrho^4 &= \varrho^4 - 2\varrho^2 \overset{\circ}{\beta}_1^2 + \overset{\circ}{\beta}_1^4 + 4 \overset{\circ}{\beta}_1^2 \varrho^4 + 2\varrho^2 \overset{\circ}{\beta}_1^2 - 2\varrho^2 \overset{\circ}{\beta}_1^2 \\ &= (\varrho^2 + \overset{\circ}{\beta}_1^2)^2 + 4\varrho^2 \overset{\circ}{\beta}_1^2 (\varrho^2 - 1) \end{aligned}$$

con esta simplificación las soluciones de (i.4.1.4) están dadas por:

$$\beta_1 = \frac{-(\varrho^2 - \overset{\circ}{\beta}_1^2) \pm (\varrho^2 + \overset{\circ}{\beta}_1^2) \sqrt{1 + \frac{4\varrho^2 \overset{\circ}{\beta}_1^2 (\varrho^2 - 1)}{(\varrho^2 + \overset{\circ}{\beta}_1^2)^2}}}{2 \overset{\circ}{\beta}_1 \varrho^2} \quad (i.4.1.5)$$

como observamos esta solución a simple vista es muy difícil de analizar por lo que hacemos la siguiente aproximación:

$$\sqrt{1 + \frac{4\varrho^2 \overset{\circ}{\beta}_1^2 (\varrho^2 - 1)}{(\varrho^2 + \overset{\circ}{\beta}_1^2)^2}} \approx 1 - \varepsilon \quad (i.4.1.6)$$

lo haremos pensando en la función:

$$\Psi(\xi) = \sqrt{1 + \xi}$$

donde

$$\xi = \frac{4\rho^2 \overset{\circ}{\beta}_1 (\rho^2 - 1)}{(\rho^2 + \overset{\circ}{\beta}_1)^2}$$

si suponemos que $|\xi| \ll 1$, obtenemos la siguiente aproximación de primer orden:

$$\varepsilon \approx -\frac{\xi}{2}$$

por lo que sustituyendo y simplificamos obtenemos:

$$\varepsilon \approx \frac{2\rho^2 \overset{\circ}{\beta}_1 (1 - \rho^2)}{(\rho^2 + \overset{\circ}{\beta}_1)^2} \quad (i.4.1.7)$$

si esto lo sustituimos en (i.4.1.5) tenemos que la primera raíz tomando el signo positivo es:

$$\beta_1^+ \approx \frac{\overset{\circ}{\beta}_1}{\rho^2} - \frac{\rho^2 + \overset{\circ}{\beta}_1}{2 \overset{\circ}{\beta}_1 \rho^2} \varepsilon \quad (i.4.1.8)$$

y la segunda raíz tomando el signo negativo es:

$$\beta_1^- \approx -\frac{1}{\overset{\circ}{\beta}_1} + \frac{\rho^2 + \overset{\circ}{\beta}_1}{2 \overset{\circ}{\beta}_1 \rho^2} \varepsilon \quad (i.4.1.9)$$

que sustituyendo (i.4.1.7) en estas dos soluciones obtenemos:

$$\beta_1^+ \approx \frac{\overset{\circ}{\beta}_1}{\rho^2} + \frac{(\rho^2 - 1) \overset{\circ}{\beta}_1}{\rho^2 + \overset{\circ}{\beta}_1}$$

$$\beta_1^- \approx -\frac{1}{\overset{\circ}{\beta}_1} - \frac{(\rho^2 - 1) \overset{\circ}{\beta}_1}{\rho^2 + \overset{\circ}{\beta}_1}$$

ahora analizaremos estas dos soluciones de acuerdo a los valores de ρ^2 :

- Si $\rho^2 \approx 1$ entonces

$$\beta_1^+ \approx \overset{\circ}{\beta}_1$$

que es la solución del ajuste ordinario lineal, que en realidad es lo que nosotros esperábamos y la otra solución es:

$$\beta_1^- \approx -\frac{1}{\overset{\circ}{\beta}_1}$$

que corresponde a la recta ortogonal.

I.5 Ejemplo

Una vez realizada la teoría del ajuste de parámetros del modelo lineal tanto ordinario como ortogonal, la ilustraremos con un ejemplo, es decir, consideraremos los siguientes datos:

Tabla I.1: Datos para el ajuste ortogonal y ordinario a una recta.

x	0	0.9	1.8	2.6	3.3	4.4	5.2	6.1	6.5	7.4
y	15.9	5.4	4.4	4.6	3.5	3.7	2.8	2.8	2.4	1.5

para ajustarlos al modelo lineal:

$$m(x; \bar{\beta}) = \beta_0 + \beta_1 x$$

como habíamos mencionado antes, como primer paso calculamos las dos raíces de la cuadrática, las cuales las graficamos con la función (i.4.9), como lo muestra la siguiente figura:

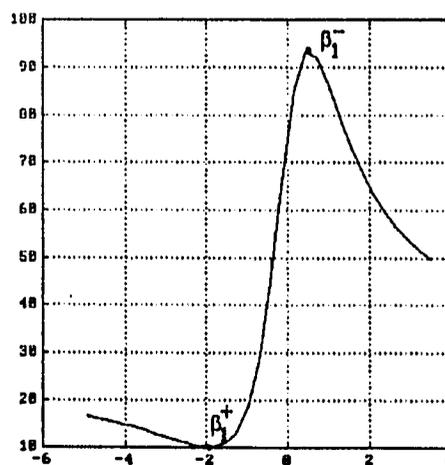


Figura I.3: Mínimo y máximo de $f(\beta_1) = \frac{1}{1+\beta_1^2} \bar{F}'\bar{r}$.

aquí nos damos cuenta cual raíz corresponde al mínimo y cual al máximo, por lo que podremos seguir dos caminos para obtener dos soluciones a este proble-

ma, una via regresión ortogonal y otra por regresión ordinaria, las cuales las utilizamos para obtener las siguiente gráfica:

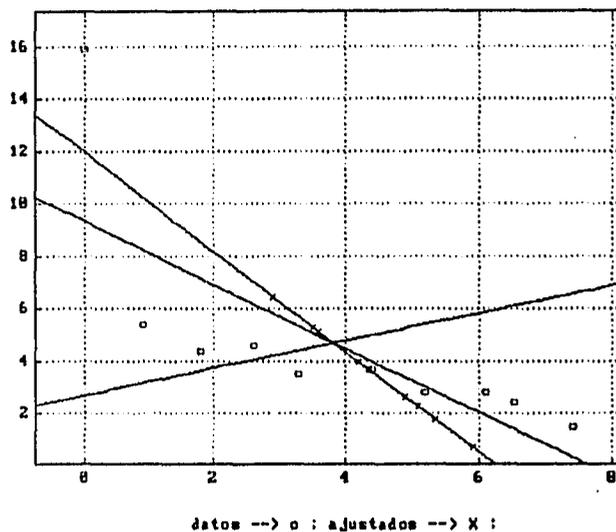


Figura I.4: Rectas obtenidas por ajuste ortogonal y ordinario.

aquí también graficamos la recta correspondiente al máximo.

Capítulo II

Problema de ajuste de parámetros del modelo múltiple.

En el capítulo anterior mostramos el problema de ajuste de parámetros a la recta, ahora trataremos el caso general, en el cual ajustaremos p parámetros, tanto para el caso ordinario como para el ortogonal.

II.1 Planteamiento del problema de ajuste ordinario de parámetros del modelo múltiple.

Vamos a generalizar las ideas del ajuste ordinario a la recta, al modelo general de $(p + 1)$ parámetros.

Pero antes de esto, lo haremos para tres parámetros y posteriormente con la experiencia ya obtenida se nos hará natural generalizarlo sin necesidad de hacerlo paso por paso, para esto suponemos que el conjunto de n datos es de la forma:

$$\{(x_{i1}, x_{i2}, y_i); \text{ para } i = 1 \dots n\} \quad (ii.1.1)$$

le queremos ajustar un modelo definido como:

$$m(x_1, x_2; \bar{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (ii.1.2)$$

donde $x_1, x_2 \in \mathbb{R}$ y $\bar{\beta}$ es un vector cuyas entradas son los parámetros a estimar, de la forma:

$$\bar{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

y además, sea p un plano en \mathbb{R}^3 que ajusta a los datos (ii.1.1), lo cual gráficamente lo podemos ver como:

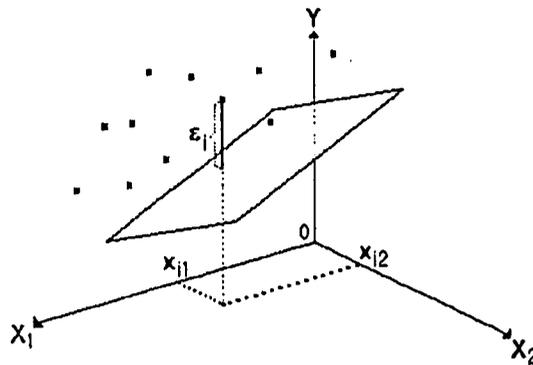


Figura II.1: Idea gráfica del ajuste ordinario a un plano.

en la gráfica denotamos con ε_i a la distancia vertical del dato i al plano p , es decir:

$$\varepsilon_i = y_i - p_i$$

donde

$$p_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$$

es el punto sobre el plano p correspondiente al punto (x_{i1}, x_{i2}) .

Nos gustaría que p sea el plano que ajuste mejor a los datos en el sentido que las distancias verticales de los datos al plano sean mínimas, por lo que el planteamiento es:

$$\begin{aligned} \min_{\beta, \bar{\varepsilon}} \quad & \frac{1}{2} \|\bar{\varepsilon}\|^2 \\ \text{sujeto a} \quad & \bar{y} = \beta_0 \bar{1} + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \bar{\varepsilon} \end{aligned} \quad (\text{ii.1.3})$$

construimos el vector $\bar{\varepsilon}$ como:

$$\bar{\varepsilon} = \begin{bmatrix} y_1 - p_1 \\ y_2 - p_2 \\ \vdots \\ y_n - p_n \end{bmatrix}$$

es equivalente resolver el problema sin restricciones:

$$\min_{\beta} \quad \frac{1}{2} \|\bar{y} - (\beta_0 \bar{1} + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2)\|^2 \quad (\text{ii.1.4})$$

el vector residuo en este caso lo escribiremos como:

$$\bar{r}(\bar{\beta}) = \bar{y} - (\beta_0 \bar{1} + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2) \quad (ii.1.5)$$

donde

$$\bar{x}_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} \quad \text{y} \quad \bar{x}_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}$$

si observamos (ii.1.5), el residuo lo podemos escribir de la forma:

$$\bar{r}(\bar{\beta}) = \bar{y} - \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

si denotamos:

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix}$$

finalmente el residuo lo escribimos en forma matricial como:

$$\bar{r}(\bar{\beta}) = \bar{y} - X\bar{\beta}$$

por lo que (ii.1.4) lo escribimos como:

$$\min_{\bar{\beta}} \frac{1}{2} (\bar{y} - X\bar{\beta})' (\bar{y} - X\bar{\beta}) \quad (ii.1.6)$$

o que es lo mismo:

$$\min_{\bar{\beta}} \frac{1}{2} \|\bar{y} - X\bar{\beta}\|^2 \quad (ii.1.7)$$

Caso general.

Con lo anterior podemos generalizar fácilmente para $(p + 1)$ parámetros, es decir, en este caso el conjunto de datos es de la forma:

$$\{(x_{i1}, x_{i2}, \dots, x_{ip}, y_i); \text{ para } i = 1 \dots n\}$$

y el modelo lineal general es:

$$m(x_1, x_2, \dots, x_p; \bar{\beta}) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

en este caso el planteamiento es el mismo que (ii.1.6) pero no $\bar{\beta}$ y X , ahora serán de la forma:

$$\bar{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad \text{y} \quad X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}$$

para el mejor manejo de esta matriz la escribimos de la forma:

$$X = [\bar{1} \quad M]$$

por lo que la matriz M queda definida como:

$$M = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (\text{ii.1.8})$$

con esta notación el planteamiento del problema lo escribimos como:

$$\min_{\bar{\beta}} \frac{1}{2} \left[\bar{y} - (\beta_0 \bar{1} + M \hat{\beta}) \right]^t \left[\bar{y} - (\beta_0 \bar{1} + M \hat{\beta}) \right] \quad (\text{ii.1.9})$$

donde $\hat{\beta}$ es el vector de parámetros excepto β_0 de la forma:

$$\hat{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$$

de aquí es fácil observar que el vector residuo en forma matricial es de la forma:

$$\bar{r} = \bar{y} - (\beta_0 \bar{1} + M \hat{\beta}) \quad (\text{ii.1.10})$$

y consecuentemente en forma vectorial es:

$$\bar{r} = \bar{y} - \left(\beta_0 \bar{1} + \sum_{j=1}^p \beta_j \bar{x}_j \right) \quad (\text{ii.1.11})$$

lo cual nos permitirá un mejor manejo de ellos en secciones posteriores.

II.2 Solución del problema del ajuste ordinario de parámetros al modelo múltiple .

Para resolver el problema del ajuste ordinario del modelo múltiple utilizamos el planteamiento (ii.1.9):

$$\min_{\hat{\beta}} \frac{1}{2} \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right]^t \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right]$$

donde M es una matriz de $n \times p$.

Calculamos el gradiente de la función :

$$f(\hat{\beta}) = \frac{1}{2} \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right]^t \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right]$$

que es un vector de la forma:

$$\nabla f(\hat{\beta}) = \begin{bmatrix} \bar{1}^t \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right] \\ M^t \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right] \end{bmatrix}$$

si lo igualamos a cero obtenemos el siguiente sistema de ecuaciones:

$$\bar{1}^t \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right] = 0 \tag{ii.2.1}$$

$$M^t \left[\bar{y} - (\beta_0 \bar{1} + M\hat{\beta}) \right] = 0$$

como ya habíamos mencionado estas son las ecuaciones normales correspondientes al caso ordinario múltiple, de la primera ecuación observamos que la suma de los residuos es igual a cero y además también de esta despejamos β_0 y así obtener:

$$\beta_0 = \frac{1}{n} \left[\bar{1}^t \bar{y} - \bar{1}^t M \hat{\beta} \right]$$

que en términos de las medias queda escrito como:

$$\beta_0 = \mu_y - \bar{\mu}^t \hat{\beta} \tag{ii.2.2}$$

donde $\bar{\mu}$ es el vector de las medias de los vectores $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_p$ es decir:

$$\bar{\mu} = \begin{bmatrix} \mu_{x_1} \\ \mu_{x_2} \\ \vdots \\ \mu_{x_p} \end{bmatrix}$$

si sustituimos el valor de β_0 en la segunda ecuación de (ii.2.1) y simplificando, obtenemos:

$$M^t \left[\bar{y} - \mu_y \bar{\mathbf{1}} - (M - \bar{\mathbf{1}} \cdot \bar{\mu}^t) \hat{\beta} \right] = 0 \quad (\text{ii.2.3})$$

a partir de este sistema es claro denotar:

$$\tilde{y} = \bar{y} - \mu_y \bar{\mathbf{1}} \quad (\text{ii.2.4})$$

es decir estamos denotando al vector menos su media esto nos ayudará a simplificar (ii.2.3) y fórmulas posteriores, análogamente lo haremos para:

$$\tilde{x}_k = \bar{x}_k - \mu_{x_k} \bar{\mathbf{1}} \quad \text{para } k = 1, \dots, p \quad (\text{ii.2.5})$$

y finalmente definimos a la matriz \tilde{M} como:

$$\tilde{M} = [\tilde{x}_1 | \tilde{x}_2 | \dots | \tilde{x}_p] \quad (\text{ii.2.6})$$

con esta notación escribimos el sistema de ecuaciones (ii.2.3) como:

$$M^t (\tilde{y} - \tilde{M} \hat{\beta}) = 0 \quad (\text{ii.2.7})$$

observamos la necesidad de escribir este sistema en términos de las medias, lo cual lo haremos a partir de la relación:

$$M = \tilde{M} + \bar{\mathbf{1}} \bar{\mu}^t \quad (\text{ii.2.8})$$

entonces ahora (ii.2.7) es de la forma:

$$(\tilde{M} + \bar{\mathbf{1}} \bar{\mu}^t)^t (\tilde{y} - \tilde{M} \hat{\beta}) = 0$$

pero si verificamos que se cumplen las siguientes igualdades:

$$\bar{\mathbf{1}}^t \tilde{y} = 0$$

$$\bar{\mathbf{1}}^t \tilde{M} = 0$$

entonces podemos escribir este sistema simplemente como:

$$\tilde{M}^t (\tilde{y} - \tilde{M} \hat{\beta}) = 0 \quad (\text{ii.2.9})$$

que corresponden a las últimas p ecuaciones normales escritas en términos de las medias, de estas ecuaciones obtenemos el vector $\hat{\beta}$ y una alternativa para hacer esto es utilizar la factorización QR de la matriz \tilde{M} , es decir:

$$\tilde{M} = QR$$

donde $Q'Q = I$ y R es una matriz triangular superior, esto lo sustituimos en (ii.2.9) y obtenemos el siguiente desarrollo:

$$\tilde{M}'\tilde{M}\hat{\beta} = \tilde{M}'\tilde{y}$$

$$(QR)'(QR)\hat{\beta} = (QR)'\tilde{y}$$

$$R'Q'QR\hat{\beta} = R'Q'\tilde{y}$$

$$R'R\hat{\beta} = R'Q'\tilde{y}$$

$$R\hat{\beta} = Q'\tilde{y}$$

finalmente, resolvemos este sistema equivalente a (ii.2.9), por sustitución hacia atrás, ya que R es una matriz triangular superior, así evitamos calcular la inversa de \tilde{M} .

II.3 Planteamiento del problema de ajuste ortogonal de parámetros del modelo múltiple.

Plantearemos el problema de ajuste ortogonal para el modelo múltiple, es decir para $p(n+1)+1$ parámetros, el porqué de este número lo veremos más adelante.

Como en el caso ordinario múltiple primero lo haremos para tres parámetros y posteriormente generalizaremos las ideas, por lo que el modelo que utilizamos es:

$$m(x_1, x_2; \beta) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad (ii.3.1)$$

lo ajustamos al conjunto de datos de la forma:

$$\{(x_{i1}, x_{i2}, y_i) \text{ para } i = 1 \dots n\} \quad (ii.3.2)$$

para graficar esto suponemos que $p = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ es un plano que ajusta a los datos:

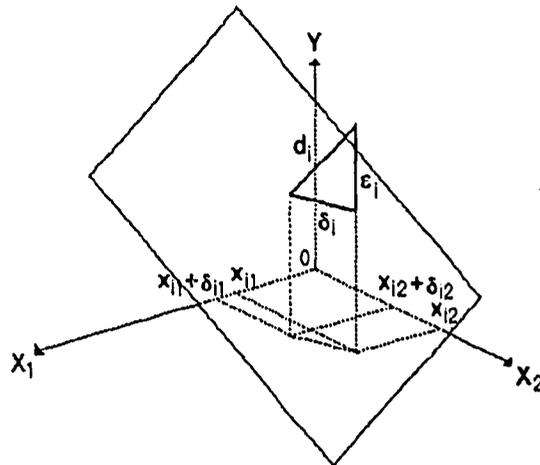


Figura II.2: Idea gráfica del ajuste ortogonal a un plano.

en este caso usamos la idea de distancia ortogonal del dato i al plano p , la cual la representamos con d_i .

Como ya hemos mencionado, minimizaremos el cuadrado las distancias d_i , las cuales las calculamos como:

$$d_i^2 = \delta_i^2 + \epsilon_i^2 \quad \text{para } i = 1, \dots, n \quad (\text{ii.3.3})$$

de acuerdo a la gráfica, la distancia δ_i^2 es:

$$\delta_i^2 = \delta_{i1}^2 + \delta_{i2}^2 \quad \text{para } i = 1, \dots, n \quad (\text{ii.3.4})$$

puesto que dos veces aplicamos el teorema de Pitágoras, por lo que sustituimos (ii.3.4) en (ii.3.3) obtenemos:

$$d_i^2 = \delta_{i1}^2 + \delta_{i2}^2 + \epsilon_i^2 \quad \text{para } i = 1, \dots, n \quad (\text{ii.3.5})$$

con los elementos que hemos adquirido, el planteamiento del problema con restricciones es :

$$\min_{\bar{\beta}, \bar{\epsilon}, \bar{\delta}_1, \bar{\delta}_2} \frac{1}{2} \| \bar{\mathbf{d}} \|^2$$

sujeto a $\bar{y} = \beta_0 \bar{\mathbf{1}} + \beta_1 (\bar{\mathbf{x}}_1 + \bar{\delta}_1) + \beta_2 (\bar{\mathbf{x}}_2 + \bar{\delta}_2) + \bar{\epsilon}$

donde

$$\bar{\mathbf{d}} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \quad \bar{\boldsymbol{\varepsilon}} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad \text{y} \quad \bar{\delta}_j = \begin{bmatrix} \delta_{1j} \\ \delta_{2j} \\ \vdots \\ \delta_{nj} \end{bmatrix} \quad \text{para } j = 1, 2.$$

que es equivalente resolver el problema sin restricciones:

$$\min_{\eta} \frac{1}{2} \left\{ \|\bar{\mathbf{y}} - [\beta_0 \bar{\mathbf{1}} + \beta_1(\bar{\mathbf{x}}_1 + \bar{\delta}_1) + \beta_2(\bar{\mathbf{x}}_2 + \bar{\delta}_2)]\|^2 + \|\bar{\delta}_1\|^2 + \|\bar{\delta}_2\|^2 \right\} \quad (\text{ii.3.6})$$

donde

$$\eta = \begin{pmatrix} \bar{\beta} \\ \bar{\delta}_1 \\ \bar{\delta}_2 \end{pmatrix}$$

para simplificar notación.

El residuo ortogonal en este caso es:

$$\bar{r}_\delta = \bar{\mathbf{y}} - [\beta_0 \bar{\mathbf{1}} + \beta_1(\bar{\mathbf{x}}_1 + \bar{\delta}_1) + \beta_2(\bar{\mathbf{x}}_2 + \bar{\delta}_2)]$$

este residuo lo escribimos en términos del residuo ordinario $\bar{\mathbf{r}}$ de la forma:

$$\bar{r}_\delta(\eta) = \bar{\mathbf{r}}(\bar{\beta}) - (\beta_1 \bar{\delta}_1 + \beta_2 \bar{\delta}_2)$$

así el planteamiento del problema de ajuste a un plano, por medio de distancias ortogonales, con notación vectorial es:

$$\min_{\eta} \frac{1}{2} \left\{ \bar{\mathbf{r}}_\delta' \bar{\mathbf{r}}_\delta + \bar{\delta}_1' \bar{\delta}_1 + \bar{\delta}_2' \bar{\delta}_2 \right\}.$$

Caso general.

El objetivo que perseguimos al plantear el problema del modelo lineal para tres parámetros es que de manera natural generalizemos para $(p+1)$ parámetros, es decir, ahora queremos ajustar el modelo múltiple de la forma:

$$m(x_1, x_2, \dots, x_p; \bar{\beta}) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

al conjunto de datos:

$$\{ (x_{i1}, x_{i2}, \dots, x_{ip}, y_i) ; \quad \text{para } i = 1 \dots n \}$$

lo cual lo logramos resolviendo el problema de minimización:

$$\min_{\eta} \frac{1}{2} \left\{ \bar{r}_\delta' \bar{r}_\delta + \sum_{j=1}^p \bar{\delta}_j' \bar{\delta}_j \right\} \quad (ii.3.7)$$

donde η es el vector de todos los parámetros a calcular de la forma:

$$\eta = \begin{bmatrix} \bar{\beta} \\ \bar{\delta}_1 \\ \vdots \\ \bar{\delta}_p \end{bmatrix}$$

y a su vez los vectores que componen este vector son:

$$\bar{\beta} = \begin{bmatrix} \beta_0 \\ \hat{\beta} \end{bmatrix} \quad \text{y} \quad \bar{\delta}_j = \begin{bmatrix} \delta_{1j} \\ \delta_{2j} \\ \vdots \\ \delta_{nj} \end{bmatrix} \quad \text{para} \quad j = 1, \dots, p.$$

donde $\hat{\beta}$ lo definimos en (ii.1.9) por lo que tenemos un total de $p(n+1) + 1$ parámetros que calcular, en este caso se incrementa pn con respecto al caso ordinario, pues en este calculamos $(p+1)$ parámetros.

También la expresión del residuo ortogonal la generalizamos de la siguiente forma:

$$\bar{r}_\delta = [\bar{r} - \sum_{j=1}^p \beta_j \bar{\delta}_j] \quad (ii.3.8)$$

donde \bar{r} es de la forma (ii.1.11), con todos estos elementos podemos ya buscar una solución a este problema, tratando de obtener un procedimiento parecido al de la sección I.4 que reduzca el trabajo que implica resolver este problema.

II.4 Solución del problema de ajuste ortogonal de parámetros del modelo múltiple.

En esta sección desarrollaremos las ideas para resolver el problema (ii.3.7), el cual escribimos a continuación:

$$\min_{\eta} \frac{1}{2} \left\{ \bar{r}_\delta^t \bar{r}_\delta + \sum_{j=1}^p \bar{\delta}_j^t \bar{\delta}_j \right\} \quad (ii.4.1)$$

construyamos la función $G : \mathbb{R}^{p(n+1)+1} \rightarrow \mathbb{R}^{n(p+1)}$ de la siguiente forma:

$$G(\eta) = \begin{bmatrix} \bar{r}_\delta(\eta) \\ \bar{\delta}_1 \\ \vdots \\ \bar{\delta}_p \end{bmatrix}$$

donde $\bar{r}_\delta(\eta)$ la denotamos en (ii.3.8), con la definición de esta función, (ii.4.1) lo escribimos en una forma más compacta:

$$\min_{\eta} \frac{1}{2} \| G(\eta) \|^2 \quad (ii.4.2)$$

seguimos el mismo procedimiento, calculamos el gradiente de la función :

$$f(\eta) = \frac{1}{2} \| G(\eta) \|^2$$

de la siguiente forma:

$$\nabla f = J_G^t G$$

donde J_G es el jacobiano de la función G , que al calcularlo obtenemos una matriz de la forma:

$$J_G = \begin{bmatrix} -\bar{1} & -(\bar{x}_1 + \bar{\delta}_1) & \cdots & -(\bar{x}_p + \bar{\delta}_p) & -\beta_1 I_n & -\beta_2 I_n & \cdots & -\beta_p I_n \\ 0_{n \times 1} & 0_{n \times 1} & \cdots & 0_{n \times 1} & I_n & 0_n & \cdots & 0_n \\ 0_{n \times 1} & 0_{n \times 1} & \cdots & 0_{n \times 1} & 0_n & I_n & \cdots & 0_n \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{n \times 1} & 0_{n \times 1} & \cdots & 0_{n \times 1} & 0_n & 0_n & \cdots & I_n \end{bmatrix}$$

al multiplicar este jacobiano por la función G obtenemos:

$$\nabla f = \begin{pmatrix} -\bar{\Gamma}'\bar{r}_\delta \\ -(\bar{x}_1 + \bar{\delta}_1)'\bar{r}_\delta \\ -(\bar{x}_2 + \bar{\delta}_2)'\bar{r}_\delta \\ \vdots \\ -(\bar{x}_p + \bar{\delta}_p)'\bar{r}_\delta \\ -\beta_1\bar{r}_\delta + \bar{\delta}_1 \\ -\beta_2\bar{r}_\delta + \bar{\delta}_2 \\ \vdots \\ -\beta_p\bar{r}_\delta + \bar{\delta}_p \end{pmatrix}$$

como ya tenemos el gradiente, el siguiente paso a seguir es igualarlo al vector cero, con esto obtenemos un sistema de ecuaciones, el cual lo escribiremos sustituyendo \bar{r}_δ por (ii.3.8), por consiguiente el sistema de $p(n+1) + 1$ ecuaciones resultante es:

$$\bar{\Gamma}'\left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r}\right] = 0 \quad (ii.4.3)$$

$$\left. \begin{aligned} (\bar{x}_1 + \bar{\delta}_1)'\left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r}\right] &= 0 \\ (\bar{x}_2 + \bar{\delta}_2)'\left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r}\right] &= 0 \\ \vdots \\ (\bar{x}_p + \bar{\delta}_p)'\left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r}\right] &= 0 \end{aligned} \right\} \quad (ii.4.4)$$

$$\left. \begin{aligned} \beta_1\left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r}\right] + \bar{\delta}_1 &= 0 \\ \beta_2\left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r}\right] + \bar{\delta}_2 &= 0 \\ \vdots \\ \beta_p\left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r}\right] + \bar{\delta}_p &= 0 \end{aligned} \right\} \quad (ii.4.5)$$

colocamos una numeración por bloques de sistemas de ecuaciones que serán importantes para nosotros, es interesante observar que el bloque (ii.4.5) es un sistema de $(n \times p)$ ecuaciones, que tal vez a simple vista no podamos identificar. Para reducir el número de parámetros a calcular seguiremos un procedimiento parecido a I.4. Para esto de (ii.4.5) tomemos sólo la l -ésima ecuación para evitar confusiones pues para las demás ecuaciones de este bloque es análogo:

$$\beta_l \left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r} \right] + \bar{\delta}_l = 0$$

si la multiplicamos por el vector $\bar{\mathbf{1}}^t$ y por la ecuación (ii.4.3) obtenemos:

$$\bar{\mathbf{1}}^t \cdot \bar{\delta}_l = 0 \quad (\text{ii.4.6})$$

también obtenemos una expresión de β_0 en términos de las medias, si la despejamos de la ecuación (ii.4.3) y adicionalmente utilizamos la igualdad (ii.4.6), con esto obtenemos:

$$\beta_0 = \mu_y - \sum_{j=1}^p \beta_j \mu_{x_j} \quad (\text{ii.4.7})$$

si la escribimos en forma vectorial queda como:

$$\beta_0 = \mu_y - \bar{\mu}^t \hat{\beta}$$

donde $\bar{\mu}$ lo definimos en (ii.2.2); como β_0 quedo expresado en términos de los parámetros $\beta_1, \beta_2, \dots, \beta_p$, nos sugiere tratar de escribir a las $\bar{\delta}_j$'s en términos de estos parámetros, por lo que escribimos (ii.4.5) de la forma:

$$\begin{aligned} \beta_1 \sum_{j=1}^p \beta_j \bar{\delta}_j + \bar{\delta}_1 - \beta_1 \bar{r} &= 0 \\ \beta_2 \sum_{j=1}^p \beta_j \bar{\delta}_j + \bar{\delta}_2 - \beta_2 \bar{r} &= 0 \\ &\vdots \\ \beta_p \sum_{j=1}^p \beta_j \bar{\delta}_j + \bar{\delta}_p - \beta_p \bar{r} &= 0 \end{aligned}$$

en el siguiente paso desarrollamos las sumas con el fin de tener una mayor claridad de la estructura de la matriz asociada a este sistema de ecuaciones

lineales con respecto a los parámetros $\bar{\delta}_1, \bar{\delta}_2, \dots, \bar{\delta}_p$:

$$\begin{aligned} (\beta_1^2 + 1)\bar{\delta}_1 + \beta_1\beta_2\bar{\delta}_2 + \beta_1\beta_3\bar{\delta}_3 + \dots + \beta_1\beta_p\bar{\delta}_p &= \beta_1\bar{r} \\ \beta_2\beta_1\bar{\delta}_1 + (\beta_2^2 + 1)\bar{\delta}_2 + \beta_2\beta_3\bar{\delta}_3 + \dots + \beta_2\beta_p\bar{\delta}_p &= \beta_2\bar{r} \\ \beta_3\beta_1\bar{\delta}_1 + \beta_3\beta_2\bar{\delta}_2 + (\beta_3^2 + 1)\bar{\delta}_3 + \dots + \beta_3\beta_p\bar{\delta}_p &= \beta_3\bar{r} \\ \vdots & \\ \beta_p\beta_1\bar{\delta}_1 + \beta_p\beta_2\bar{\delta}_2 + \beta_p\beta_3\bar{\delta}_3 + \dots + (\beta_p^2 + 1)\bar{\delta}_p &= \beta_p\bar{r} \end{aligned}$$

escrito así no nos queda duda que el sistema de ecuaciones anterior, lo escribimos en forma matricial como:

$$\begin{bmatrix} (\beta_1^2 + 1)I_n & \beta_1\beta_2I_n & \beta_1\beta_3I_n & \dots & \beta_1\beta_pI_n \\ \beta_2\beta_1I_n & (\beta_2^2 + 1)I_n & \beta_2\beta_3I_n & \dots & \beta_2\beta_pI_n \\ \beta_3\beta_1I_n & \beta_3\beta_2I_n & (\beta_3^2 + 1)I_n & \dots & \beta_3\beta_pI_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_p\beta_1I_n & \beta_p\beta_2I_n & \beta_p\beta_3I_n & \dots & (\beta_p^2 + 1)I_n \end{bmatrix} \begin{bmatrix} \bar{\delta}_1 \\ \bar{\delta}_2 \\ \bar{\delta}_3 \\ \vdots \\ \bar{\delta}_p \end{bmatrix} = \begin{bmatrix} \beta_1\bar{r} \\ \beta_2\bar{r} \\ \beta_3\bar{r} \\ \vdots \\ \beta_p\bar{r} \end{bmatrix}$$

utilizamos la definición del producto de Kronecker o Tensorial (ver apéndice A) el sistema es equivalente a:

$$\begin{bmatrix} (\beta_1^2 + 1) & \beta_1\beta_2 & \beta_1\beta_3 & \dots & \beta_1\beta_p \\ \beta_2\beta_1 & (\beta_2^2 + 1) & \beta_2\beta_3 & \dots & \beta_2\beta_p \\ \beta_3\beta_1 & \beta_3\beta_2 & (\beta_3^2 + 1) & \dots & \beta_3\beta_p \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \beta_p\beta_1 & \beta_p\beta_2 & \beta_p\beta_3 & \dots & (\beta_p^2 + 1) \end{bmatrix} \otimes I_n \begin{bmatrix} \bar{\delta}_1 \\ \bar{\delta}_2 \\ \bar{\delta}_3 \\ \vdots \\ \bar{\delta}_p \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{bmatrix} \otimes \bar{r} \quad (ii.4.8)$$

si denotamos

$$\bar{\delta} = \begin{bmatrix} \bar{\delta}_1 \\ \bar{\delta}_2 \\ \vdots \\ \bar{\delta}_p \end{bmatrix}$$

entonces (ii.4.8), lo escribimos como:

$$\left[(I_p + \hat{\beta}\hat{\beta}^t) \otimes I_n \right] \bar{\delta} = \hat{\beta} \otimes \bar{r} \quad (ii.4.9)$$

Una vez que obtenemos esta expresión recordemos que nuestro objetivo es despejar los vectores $\bar{\delta}_j$'s en términos de los parámetros β_j para $j = 1, \dots, p$, para esto calculamos la inversa de la matriz $(I_p + \hat{\beta}\hat{\beta}^t)$ que intuimos tenga una

estructura similar, por lo cual la escribimos como:

$$(I_p + \hat{\beta}\hat{\beta}^t)^{-1} = I_p - \bar{\mathbf{u}}\bar{\mathbf{u}}^t$$

donde $\bar{\mathbf{u}} \in \mathbb{R}^p$; como estamos calculando la inversa entonces se cumple la siguiente igualdad:

$$(I_p + \hat{\beta}\hat{\beta}^t)(I_p - \bar{\mathbf{u}}\bar{\mathbf{u}}^t) = I_p \quad (\text{ii.4.10})$$

si realizamos las operaciones y hacemos la siguiente asignación:

$$\bar{\mathbf{u}} = \gamma\hat{\beta} \quad \text{donde } \gamma \in \mathbb{R},$$

obtenemos la siguiente igualdad:

$$\hat{\beta}\hat{\beta}^t[1 - \gamma^2(1 + \|\hat{\beta}\|^2)] = 0_p$$

como $\hat{\beta} \neq 0$ por lo tanto obtenemos:

$$\gamma^2 = \frac{1}{1 + \|\hat{\beta}\|^2}$$

por lo que finalmente obtenemos la inversa deseada:

$$(I_p + \hat{\beta}\hat{\beta}^t)^{-1} = I_p - \frac{\hat{\beta}\hat{\beta}^t}{1 + \|\hat{\beta}\|^2}$$

multiplicamos por esta matriz inversa a (ii.4.9) y así despejar $\bar{\delta}$:

$$\begin{aligned} \bar{\delta} &= \left[\left(I_p - \frac{\hat{\beta}\hat{\beta}^t}{1 + \|\hat{\beta}\|^2} \right) \otimes I_n \right] (\hat{\beta} \otimes \mathbf{F}) \\ &= \left(\hat{\beta} - \frac{\hat{\beta}\hat{\beta}^t\hat{\beta}}{1 + \|\hat{\beta}\|^2} \right) \otimes \mathbf{F} \\ &= \left(1 - \frac{\|\hat{\beta}\|^2}{1 + \|\hat{\beta}\|^2} \right) (\hat{\beta} \otimes \mathbf{F}) \end{aligned}$$

por lo que finalmente llegamos a la expresión:

$$\bar{\delta} = \frac{1}{1 + \|\hat{\beta}\|^2} (\hat{\beta} \otimes \bar{r})$$

de aquí obtenemos una expresión para $\bar{\delta}_k$, en términos de $\hat{\beta}$:

$$\bar{\delta}_k = \frac{\beta_k}{1 + \|\hat{\beta}\|^2} \bar{r} \quad \text{para } k = 1 \dots p \quad (ii.4.11)$$

una vez obtenidas estas expresiones, nuestro problema se reduce a resolver el sistema (ii.4.4), este sistema de ecuaciones depende tanto del parámetro $\bar{\delta}$ como de $\bar{\beta}$, por lo que lo transformamos para que sólo dependa de $\hat{\beta}$, para esto consideremos la k -ésima ecuación:

$$(\bar{x}_k + \bar{\delta}_k)' \left[\sum_{j=1}^p \beta_j \bar{\delta}_j - \bar{r} \right] = 0$$

basta simplificar esta sola ecuación, puesto que para las demás es análogo, por lo que sustituimos (ii.4.11) y simplificamos para obtener:

$$\left[(1 + \|\hat{\beta}\|^2) \bar{x}_k + \beta_k \bar{r} \right]' \bar{r} = 0 \quad \text{para } k = 1, \dots, p$$

con esto el sistema de ecuaciones (ii.4.4) lo escribimos en forma matricial como:

$$\left[(1 + \|\hat{\beta}\|^2) M' + \hat{\beta} \bar{r}' \right] \bar{r} = 0 \quad (ii.4.12)$$

donde M es la matriz definida en (ii.1.8).

Ahora sólo nos falta sustituir el parámetro β_0 por su expresión, el cual está involucrado en el residuo ordinario \bar{r} , para esto utilizamos (ii.2.4), (ii.2.5) y (ii.2.6) para escribir el residuo como:

$$\bar{r} = \bar{y} - \bar{M} \hat{\beta} \quad (ii.4.13)$$

así el sistema de ecuaciones (ii.4.12), en estos términos es:

$$\left[(1 + \|\hat{\beta}\|^2) M + (\bar{y} - \bar{M} \hat{\beta}) \hat{\beta}' \right]' (\bar{y} - \bar{M} \hat{\beta}) = 0 \quad (ii.4.14)$$

bueno, pero este sistema nos gustaría escribirlo en términos de \tilde{M} , para esto utilizamos (ii.2.8), para transformar finalmente las ecuaciones (ii.4.14) a:

$$(1 + \|\hat{\beta}\|^2)\tilde{M}'(\tilde{y} - \tilde{M}\hat{\beta}) + \hat{\beta}(\tilde{y} - \tilde{M}\hat{\beta})'(\tilde{y} - \tilde{M}\hat{\beta}) = 0 \quad (ii.4.15)$$

que es equivalente a:

$$(1 + \|\hat{\beta}\|^2)\tilde{M}'(\tilde{y} - \tilde{M}\hat{\beta}) + \hat{\beta}\|(\tilde{y} - \tilde{M}\hat{\beta})\|^2 = 0 \quad (ii.4.16)$$

$$\left[(1 + \|\hat{\beta}\|^2)\tilde{M} + (\tilde{y} - \tilde{M}\hat{\beta})\hat{\beta}' \right]' (\tilde{y} - \tilde{M}\hat{\beta}) = 0$$

Si observamos detenidamente (ii.4.16) y suponemos que:

$$\|\bar{r}\|^2 \approx 0 \quad (ii.4.17)$$

entonces obtenemos:

$$\tilde{M}'(\tilde{y} - \tilde{M}\hat{\beta}) \approx 0$$

que son parecidas a las ecuaciones normales escritas en (ii.2.9), por lo que interpretaremos que la solución del sistema (ii.4.16) es parecida a la del ajuste ordinario, la cual denotaremos como $\hat{\beta}$. Con esta primera observación podemos establecer un primer método de solución.

Método 1.

- Calcular la solución $\hat{\beta}_s$ del sistema de ecuaciones

$$(1 + \|\hat{\beta}_s\|^2)\tilde{M}'(\tilde{y} - \tilde{M}\hat{\beta}_s) + \hat{\beta}_s\|(\tilde{y} - \tilde{M}\hat{\beta}_s)\|^2 = 0$$

tomando como punto inicial $\hat{\beta}_s$.

- Obtener β_0 con la expresión

$$\beta_0 = \mu_y - \bar{\mu}'\hat{\beta}_s$$

- Obtener δ con la expresión

$$\bar{\delta} = \frac{1}{1 + \|\hat{\beta}_s\|^2}(\hat{\beta}_s \otimes \bar{r})$$

Bueno pero si no se cumple la suposición (ii.4.17), tal vez no estemos cerca del mínimo y con el método anterior no obtengamos este. Por lo que una alternativa que se nos ocurre es utilizar la expresión (ii.4.13) del residuo ordinario $\bar{\mathbf{r}}$ y sustituir (ii.4.11) en la función objetivo:

$$f(\eta) = \frac{1}{2} \left\{ \bar{\mathbf{r}}_s^t \bar{\mathbf{r}}_s + \sum_{j=1}^p \bar{\delta}_j^t \bar{\delta}_j \right\}$$

de nuestro problema original (ii.4.1), para que sólo dependa del vector $\hat{\beta}$, esto lo haremos por partes, primero sustituimos (ii.4.11) en el residuo ortogonal, con esto obtenemos:

$$\bar{\mathbf{r}}_s = \frac{1}{1 + \|\hat{\beta}\|^2} \bar{\mathbf{r}}$$

y si ahora también sustituimos (ii.4.11) en la suma restante, obtenemos:

$$\sum_{j=1}^p \bar{\delta}_j^t \bar{\delta}_j = \frac{\|\hat{\beta}\|^2}{1 + \|\hat{\beta}\|^2} \bar{\mathbf{r}}^t \bar{\mathbf{r}}$$

y con estas dos igualdades obtenemos la función que depende del vector de parámetros $\hat{\beta}$, de la forma :

$$f(\hat{\beta}) = \frac{1}{2(1 + \|\hat{\beta}\|^2)} \bar{\mathbf{r}}^t \bar{\mathbf{r}} \quad (\text{ii.4.18})$$

con esto transformamos nuestro problema de minimización (ii.4.1) a:

$$\min_{\hat{\beta}} \frac{1}{2(1 + \|\hat{\beta}\|^2)} \bar{\mathbf{r}}^t \bar{\mathbf{r}}$$

con esto podemos establecer el siguiente método alternativo.

Método 2.

- Calcular la solución $\hat{\beta}_m$ de

$$\min_{\hat{\beta}} \frac{1}{2(1 + \|\hat{\beta}\|^2)} \bar{\mathbf{r}}^t \bar{\mathbf{r}}$$

tomando como punto inicial $\hat{\beta}^0$.

- Obtener β_0 con la expresión

$$\beta_0 = \mu_y - \bar{\mathbf{x}}^t \hat{\beta}_m$$

- Obtener δ con la expresión

$$\bar{\delta} = \frac{1}{1 + \|\hat{\beta}_m\|^2} (\hat{\beta}_m \otimes \bar{\mathbf{r}})$$

con esto nos queda claro que reducimos nuestro problema a calcular p parámetros.

II.4.1 Ejemplos.

Ilustremos la teoría desarrollada del ajuste de parámetros del modelo múltiple por medio de dos ejemplos aplicando el **Método 1** y el **Método 2** de la sección anterior.

Decidimos obtener los datos aleatoriamente por medio de la $N(0, 1)$, para ambos ejemplos y ajustarlos al modelo:

$$m(x_1, x_2; \bar{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

es decir, a un plano, pues en el caso múltiple para $p = 2$ aún tenemos la ventaja de poder gráficar.

Ejemplo 1 :

En este primer ejemplo consideramos veintitrés datos, escritos en la siguiente tabla:

Tabla II.1: Datos obtenidos aleatoriamente por medio de la $N(0,1)$.

i	x_1	x_2	y	i	x_1	x_2	y
1	0.8089	0.2893	5.5458	13	0.7558	0.7065	8.1262
2	0.9317	0.5374	8.4060	14	0.4622	0.7417	6.0289
3	0.6516	0.5144	6.9491	15	0.9514	0.0191	5.7857
4	0.2152	0.1034	3.5739	16	0.6327	0.8860	8.0434
5	0.6796	0.4140	6.0816	17	0.4393	0.5250	5.0179
6	0.9089	0.5767	6.8848	18	0.8247	0.4633	6.2497
7	0.2501	0.8766	5.5556	19	0.6890	0.0652	5.8154
8	0.8609	0.4400	7.8886	20	0.7022	0.7134	6.4503
9	0.4713	0.7297	5.5797	21	0.9871	0.4889	7.9168
10	0.5060	0.8693	6.6863	22	0.9544	0.6677	8.3642
11	0.6004	0.7156	6.3187	23	0.8513	0.6820	6.7133
12	0.8176	0.8007	8.0027				

como primer paso graficamos la superficie correspondiente a la función (ii.4.18), que observamos como:

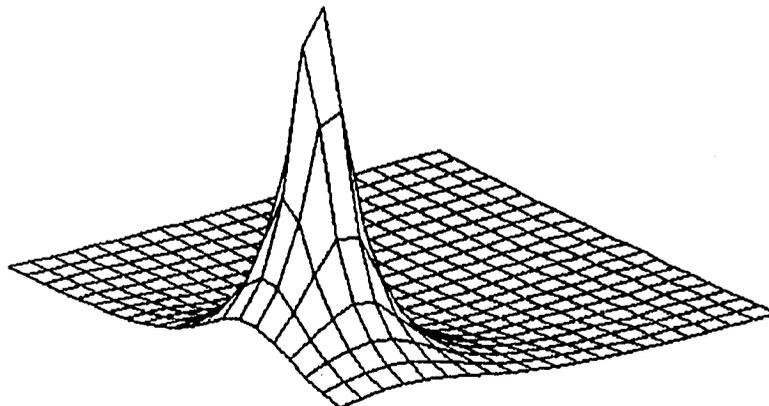


Figura II.3: Gráfica de la función $f(\hat{\beta}) = \frac{1}{2(1 + \|\hat{\beta}\|^2)} \bar{F}'\bar{F}$.

aquí observamos claramente un máximo, pero no distinguimos el mínimo, aunque éste existe. Graficamos las curvas de nivel de la función (ii.4.18) y las soluciones obtenidas de los dos métodos, así como la solución del ajuste ordinario:

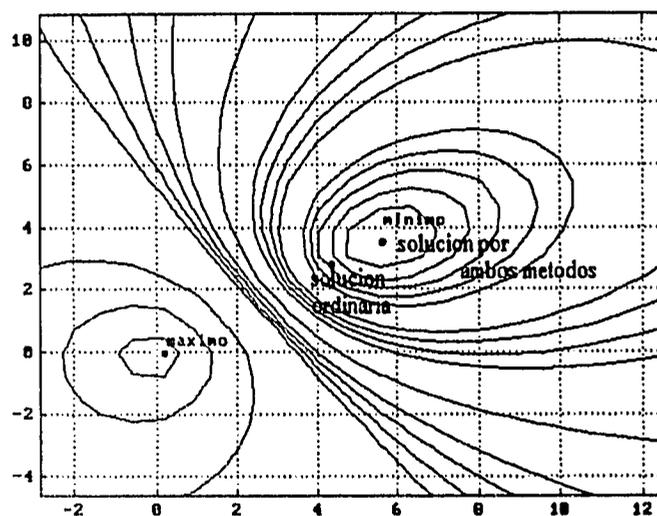


Figura II.4: Curvas de nivel de la función $f(\hat{\beta}) = \frac{1}{2(1 + \|\hat{\beta}\|^2)} \bar{\mathbf{r}}^T \bar{\mathbf{r}}$.

en este ejemplo tanto la solución correspondiente al **Método 1** como la solución del problema de minimización del **Método 2** coinciden, por lo que en ambas resuelven satisfactoriamente el problema.

Ejemplo 2 :

El conjunto de datos de este ejemplo lo escogimos para mostrar el caso en

que el Método 1 no obtenemos un mínimo sino un máximo:

Tabla II.2: Datos obtenidos aleatoriamente por medio de la $N(0, 1)$.

i	x ₁	x ₂	y	i	x ₁	x ₂	y
1	0.5097	0.2618	1.1416	11	0.7671	0.0342	2.6285
2	0.8258	0.7116	5.2413	12	0.0617	0.5238	2.9150
3	0.8343	0.3784	4.0072	13	0.5193	0.0316	1.9954
4	0.7264	0.2006	2.8553	14	0.9291	0.7532	3.6687
5	0.3294	0.2326	2.7810	15	0.2205	0.5916	2.7801
6	0.8021	0.4361	1.5660	16	0.2046	0.8229	1.9193
7	0.6318	0.4695	0.9206	17	0.0890	0.3296	3.2755
8	0.0453	0.8612	5.3043	18	0.6997	0.2321	2.8307
9	0.0414	0.9772	4.2728	19	0.7065	0.5365	1.2293
10	0.1533	0.4445	1.8073	20	0.4946	0.4306	0.2687

graficamos la función (ii.4.18) utilizando las datos de la Tabla II.2:

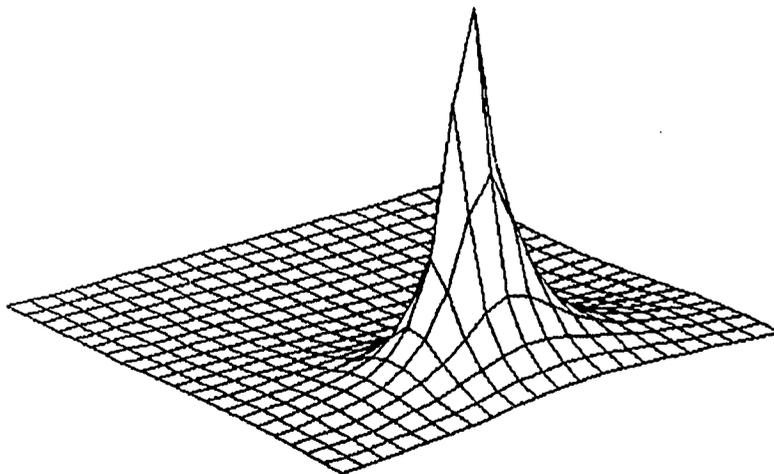


Figura II.5: Gráfica de la función $f(\hat{\beta}) = \frac{1}{2(1 + \|\hat{\beta}\|^2)} \bar{F}^t \bar{F}$.

aquí también observamos el máximo, pero no el mínimo, por lo que graficamos sus curvas de nivel con las soluciones de ambos métodos y la solución ordinaria,

lo cual observamos como:

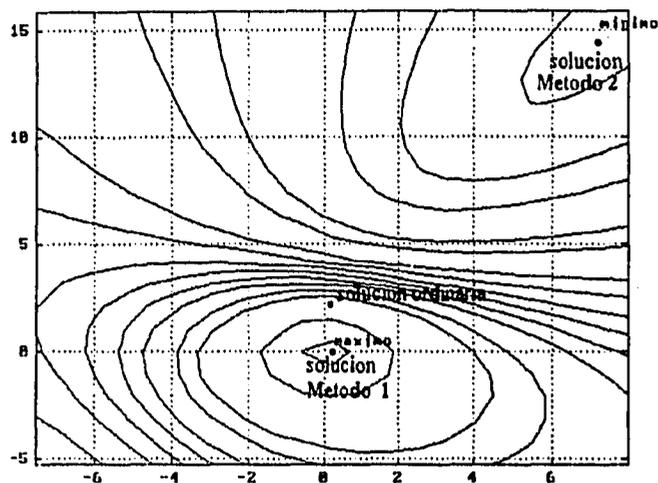


Figura II.6: Curvas de nivel de la función $f(\hat{\beta}) = \frac{1}{2(1 + \|\hat{\beta}\|^2)} \mathbf{F}^T \mathbf{F}$.

observamos que tanto la solución ordinaria como la solución del **Método 1** se aproximan al máximo de la función (ii.4.18), mientras tanto la solución del **Método 2** se aproxima al mínimo, por lo que recomendamos utilizar el **Método 2**, pues nos asegura que se aproxima al mínimo.

Capítulo III

Métodos de solución para el problema de mínimos cuadrados no-lineales.

El problema central que nos interesa tratar en este capítulo es el problema de mínimos cuadrados no-lineales, para resolverlo existe una extensa colección de métodos, nosotros revisaremos los métodos de Gauss-Newton, Gauss-Newton Modificado y Levenberg-Marquardt. De los dos primeros sólo daremos sus ideas principales, como antecedente del tercero, puesto que este último lo desarrollaremos con más detalle para utilizarlo como parte del método de mínimos cuadrados no-lineales con distancias ortogonales que trataremos en el siguiente capítulo.

Consideremos la función no-lineal $G : \mathbb{R}^l \rightarrow \mathbb{R}^m$ de la forma:

$$G(\bar{\eta}) = \begin{bmatrix} g_1(\bar{\eta}) \\ g_2(\bar{\eta}) \\ \vdots \\ g_m(\bar{\eta}) \end{bmatrix}$$

con $\bar{\eta} \in \mathbb{R}^l$ y g_i una función no-lineal para $i = 1, \dots, m$; definida esta función planteamos el problema de mínimos cuadrados no-lineales como:

$$\min_{\bar{\eta}} \frac{1}{2} \| G(\bar{\eta}) \|^2 \quad (3.1)$$

La idea general de los métodos para resolver (3.1) es construir una sucesión $\{\bar{\eta}_k\}$ con la característica que converja a la solución $\bar{\eta}^*$ de (3.1).

III.1 Método de Gauss-Newton.

En este método para construir la sucesión $\bar{\eta}_k$, suponemos que conocemos el elemento k -ésimo de esta sucesión y a partir de este elemento obtener $\bar{\eta}_{k+1}$, para esto aproximamos localmente la función G , por la parte lineal del desarrollo de Taylor, en una vecindad de $\bar{\eta}_k$, es decir:

$$G(\bar{\eta}) \approx G(\bar{\eta}_k) + G'(\bar{\eta}_k)(\bar{\eta} - \bar{\eta}_k) \quad (iii.1.1)$$

donde G' es el jacobiano de la función G . De aquí resulta natural seguir la siguiente notación para lo que resta del capítulo:

$$G(\bar{\eta}_k) = G_k$$

$$G'(\bar{\eta}_k) = J_{G_k}$$

$$\bar{s} = \bar{\eta} - \bar{\eta}_k$$

con la aproximación (iii.1.1) la iteración de este método es:

$$\bullet \bar{\eta}_{k+1} = \bar{\eta}_k + \bar{s}_k$$

donde \bar{s}_k es la solución al problema:

$$\min_{\bar{s}} \frac{1}{2} \| G_k + J_{G_k} \bar{s} \|^2 \quad (iii.1.2)$$

observamos que las correspondientes ecuaciones normales de este problema de mínimos cuadrados lineales son:

$$(J_{G_k}^t J_{G_k}) \bar{s} = -J_{G_k}^t G_k.$$

La convergencia de este método local es usualmente lineal y es mejor a medida que $\| G \|^2$ es pequeña en la solución, como estamos construyendo una sucesión que converja a la solución de un problema de minimización es lógico pedir que se cumpla la siguiente desigualdad :

$$\| G_k \|^2 \geq \| G_{k+1} \|^2$$

pero en la práctica se observó que en algunos casos esta desigualdad no se satisfacía, por lo que modificaron este método para lograr una disminución del valor de la norma de la función G en cada iteración.

III.2 Método de Gauss-Newton Modificado.

Como lo habíamos dicho, en esta sección explicaremos las modificaciones al método de Gauss-Newton para evitar su divergencia, introduciendo un algoritmo de Búsqueda de Línea, transformando la iteración de Newton a:

- $J_{G_k}^t J_{G_k} \bar{s} = -J_{G_k}^t G_k$
- $\bar{\eta}_{k+1} = \bar{\eta}_k + \alpha_k \bar{s}_k$

donde α_k se elige de tal manera que satisfaga las condiciones de Wolfe-Powell [14], las cuales son:

1. $\phi(\alpha_k) \leq \phi(0) + \sigma_1 \phi'(0) \alpha_k$
2. $|\phi(\alpha_k)| \leq \sigma_2 |\phi'(0)|$

donde $\phi(\alpha) = \frac{1}{2} \|G(\bar{\eta}_k + \alpha_k \bar{s}_k)\|^2$ y en la práctica los valores que se usan son:

$$\sigma_1 = 10^{-4} \quad \text{y} \quad \sigma_2 = 0.9$$

III.3 Método de Levenberg-Marquardt.

En la sección anterior describimos como modificar el método de Gauss-Newton para asegurar el decrecimiento de la norma de la función G en cada iteración. En esta sección desarrollaremos dos ideas, las cuales a primera vista pensaríamos que no tienen relación alguna, pero concluiremos que son dos formas de abordar el mismo problema.

En primer lugar retomemos el problema de mínimos cuadrados no-lineales que nos interesa:

$$\min_{\bar{\eta}} \frac{1}{2} \|G(\bar{\eta})\|^2 \quad (\text{iii.3.1})$$

al resolver este problema por el método de Gauss-Newton, en cada iteración resolvemos un sistema lineal de la forma:

$$(J_{G_k}^t J_{G_k}) \bar{s} = -J_{G_k}^t G_k$$

un problema común que se puede presentar al resolver este sistema, es que la matriz $J_{G_k}^t J_{G_k}$ sea singular o casi singular, para evitar estos casos, corregimos los elementos de la diagonal de esta matriz sumándoles un número positivo λ y así resolveremos el sistema:

$$(J_{G_k}^t J_{G_k} + \lambda I) \bar{s} = -J_{G_k}^t G_k \quad (iii.3.2)$$

se observó que en la práctica funcionaba, aunque no hemos dicho como elegir el parámetro λ .

Para tratar de entender la elección de λ , consideremos a \bar{s} como una función de λ de la siguiente forma:

$$\bar{s}(\lambda) = -(J_{G_k}^t J_{G_k} + \lambda I)^{-1} g_k$$

donde

$$g_k = J_{G_k}^t G_k$$

es el gradiente de la función $\frac{1}{2} \| G(\bar{\eta}) \|^2$.

A partir del planteamiento anterior podemos analizar los siguientes dos puntos:

1. Si $\lambda = 0$ entonces \bar{s} es la solución de Gauss-Newton, es decir tenemos:

$$\bar{s}(0) = -(J_G^t J_G)^{-1} g$$

la cual la denotamos como \bar{s}_{GN} .

2. Ahora si dividimos a (iii.3.2) entre λ obtenemos:

$$\left(\frac{J_G^t J_G}{\lambda} + I \right) \bar{s} = \frac{-g}{\lambda}$$

entonces cuando $\lambda \gg 1$, tenemos la siguiente aproximación

$$s(\lambda) \approx \frac{-g}{\lambda}$$

por lo que intuitivamente observamos que \bar{s} tiende a cero, si λ tiende a infinito.

Con estas dos observaciones podemos decir que $\bar{s}(\lambda)$ es una función decreciente y a medida que λ crece, $\bar{s}(\lambda)$ se aproxima a $-\frac{g}{\lambda}$, con esto construimos la gráfica de $\bar{s}(\lambda)$ como:

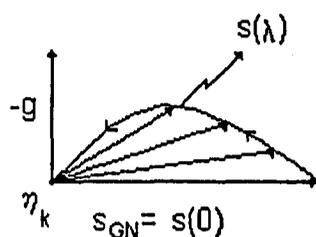


Figura III.1: Gráfica de la función $\bar{s}(\lambda)$.

así que dado λ obtenemos una $\bar{s}(\lambda)$, por lo que la iteración quedaría como:

- $\bar{\eta}_{k+1} = \bar{\eta}_k + s_k(\lambda)$.

Para desarrollar la segunda idea, recordemos que el método de Gauss-Newton tiene convergencia local, pues utilizamos la aproximación lineal (iii.1.1), lo que nos sugiere incorporar la idea de región de confianza, es decir, si en una vecindad dada es aceptable dicha aproximación a la función, así el planteamiento del problema, nos queda:

$$\min_{\|\bar{s}\|^2 \leq \Delta^2} \frac{1}{2} \|G_k + J_{G_k} \bar{s}\|^2 \quad (\text{iii.3.3})$$

donde Δ es el radio de la región de confianza. Este problema gráficamente lo veremos así:

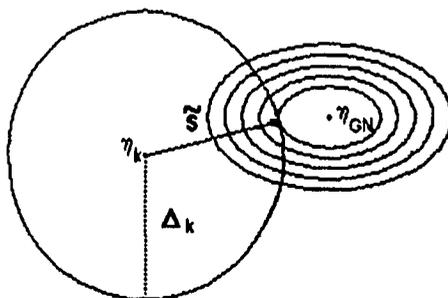


Figura III.2: Idea gráfica del problema de minimización con región de confianza.

entonces dado el valor de Δ obtenemos un valor de \bar{s} , es decir, obtenemos una función del radio de la región de confianza la cual denotamos como $\bar{s}(\Delta)$, por lo que en este caso la iteración queda como:

- $\bar{\eta}_{k+1} = \bar{\eta}_k + \bar{s}_k(\Delta)$.

Algunas propiedades de \bar{s} , que podemos observar son:

1. Si $\Delta \geq \| \bar{s}_{GN} \|$ entonces $\bar{s}(\Delta) = \bar{s}_{GN}$, donde \bar{s}_{GN} es la solución por Gauss-Newton.
2. Si Δ tiende a cero entonces $\bar{s}(\Delta)$ también tiende a cero.
3. $\bar{s}'(\Delta) \approx -c\Delta g$ para $0 < \Delta \ll 1$ y c es una constante.

entonces la gráfica de \bar{s} es de la forma:

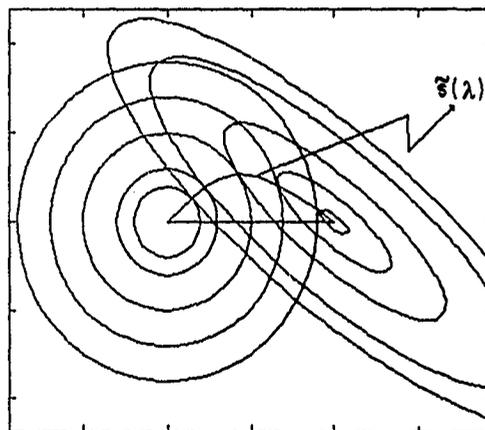


Figura III.3: Gráfica de la función $\bar{s}(\lambda)$.

Para observar que relación existe entre \bar{s} y \bar{s} , resolveremos el problema (iii.3.3) por el método de Multiplicadores de Lagrange, por lo que definimos a la función L como:

$$L(\bar{s}, \bar{\lambda}) = \frac{1}{2} \| G_k + J_{G_k} \bar{s} \|^2 + \frac{\bar{\lambda}}{2} (\| \bar{s} \|^2 - \Delta^2) \quad \text{con } \bar{\lambda} \geq 0$$

calculamos los gradientes con respecto \bar{s} y $\bar{\lambda}$, obtenemos:

$$\nabla_{\bar{s}} L = J_{G_k}^t G_k + J_{G_k}^t J_{G_k} \bar{s} + \bar{\lambda} \bar{s}$$

$$\nabla_{\bar{\lambda}} L = \|\bar{s}\|^2 - \Delta^2$$

igualamos los gradientes al vector cero y despejamos, para obtener las condiciones de optimalidad:

$$\begin{aligned} (J_{G_k}^t J_{G_k} + \bar{\lambda} I) \bar{s}(\bar{\lambda}) &= -J_{G_k}^t G_k \\ \|\bar{s}(\bar{\lambda})\|^2 &= \Delta^2 \end{aligned} \tag{iii.3.4}$$

de este sistema de ecuaciones, concluimos que $\bar{\lambda}$ depende de Δ , por lo que lo escribimos como:

$$\bar{s}(\bar{\lambda}) = \bar{s}(\bar{\lambda}(\Delta))$$

si elegimos $\lambda = \bar{\lambda}$ entonces observaremos que (iii.3.4) es el sistema (iii.3.2) que obtuvimos para $\bar{s}(\lambda)$, así podemos concluir que \bar{s} y \bar{s} representan la misma función pero con diferente parametrización, puesto que la recorren en distinta dirección, por lo tanto:

$$\bar{s}(\lambda) = \bar{s}(\lambda(\Delta))$$

así, dado el radio de la región de confianza, queda determinado el parámetro λ , que anteriormente no habíamos especificado como elegirlo.

Si suponemos \bar{s}^* es una solución de (iii.3.3) entonces $\bar{s}^* = \bar{s}(\lambda^*)$ para alguna $\lambda^* \geq 0$, es decir :

$$\bar{s}(\lambda^*) = -(J_{G_k}^t J_{G_k} + \lambda^* I)^{-1} J_{G_k}^t G_k$$

para esta solución tenemos dos posibilidades:

1. Si $\lambda^* > 0$ y $\|\bar{s}(\lambda)\|^2 = \Delta^2$ entonces \bar{s}^* es la única solución.
2. Si $\lambda^* = 0$ y $\|\bar{s}(0)\|^2 \leq \Delta^2$, lo cual nos quiere decir que \bar{s}^* es la solución por Gauss-Newton, es decir $\bar{s}^* = \bar{s}_{GN}$.

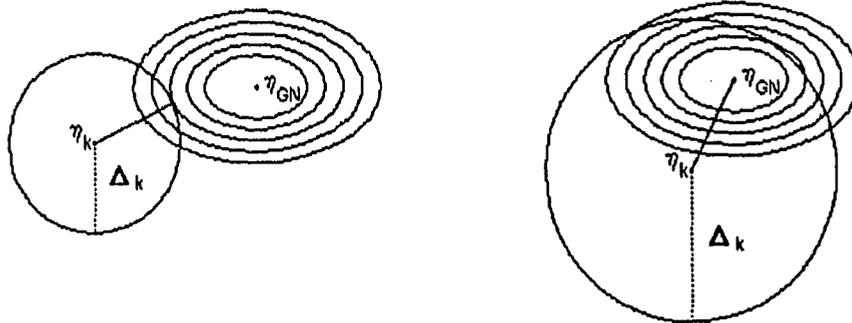


Figura III.4: 1. $\lambda^* = 0$ y $\|\bar{s}(0)\|^2 \leq \Delta^2$. Figura III.5: 2. $\lambda^* > 0$ y $\|\bar{s}(\lambda)\|^2 = \Delta^2$.

A partir de las ideas que hemos planteado, podemos establecer el siguiente algoritmo para resolver el problema :

$$\min_{\bar{\eta}} \frac{1}{2} \|G(\bar{\eta})\|^2$$

ALGORITMO:

1. Dada $\Delta_k > 0$ encontrar $\lambda \geq 0$ tal que

$$\|\bar{s}(\lambda)\|^2 = \Delta^2$$

2. Calcular $\bar{s}(\lambda)$ de

$$(J_{G_k}^t J_{G_k} + \lambda I) \bar{s}(\lambda) = -J_{G_k}^t G_k$$

3. Si $\|G(\bar{\eta}_k)\|^2 \geq \|G(\bar{\eta}_k + \bar{s}_k)\|^2$

entonces $\bar{\eta}_{k+1} = \bar{\eta}_k + \bar{s}_k$ y evaluar $J_{G_{k+1}}$

sino $\bar{\eta}_{k+1} = \bar{\eta}_k$ y $J_{G_{k+1}} = J_{G_k}$

4. Elegir Δ_{k+1}

Este algoritmo es muy general, por lo que en las siguientes subsecciones explicaremos como calcular en forma eficiente y robusta cada una de sus partes. La teoría general que se generó a partir de estas ideas se encuentra en los apéndices B y C.

III.3.1 Solución del problema estructurado del problema de mínimos cuadrados lineales.

En esta sección discutiremos como resolver el sistema lineal de la forma:

$$(J_G^t J_G + \lambda D^t D) \bar{s} = -J_G^t G \quad (iii.3.1.1)$$

un camino sencillo, para resolver este sistema de ecuaciones es por medio de la factorización de Cholesky, pero si $\lambda = 0$ y J_G es muy cercana a una matriz singular, cabe la posibilidad de obtener overflows o underflows al realizar el producto $J_G^t J_G$, por lo que un camino alternativo es observar que (iii.3.1.1) son las ecuaciones normales del problema:

$$\min_{\bar{s}} \left\| \begin{bmatrix} J_G \\ \lambda^{\frac{1}{2}} D \end{bmatrix} \bar{s} + \begin{bmatrix} G \\ 0 \end{bmatrix} \right\|^2 \quad (iii.3.1.2)$$

el cual resolveremos en dos etapas siguiendo las ideas de Golub [16], pero considerando pivoteo.

En la primera etapa, calcularemos la factorización QR con pivoteo de J_G , obtenemos:

$$J_G \Pi = QR \quad (iii.3.1.3)$$

donde:

- Q es una matriz ortogonal, es decir que $Q^t Q = I$.
- R es una matriz trapezoidal de $m \times \ell$ de la forma

$$R = \begin{matrix} & \begin{matrix} \kappa & \ell - \kappa \end{matrix} \\ \begin{matrix} \kappa \\ m - \kappa \end{matrix} & \begin{pmatrix} T & S \\ 0 & 0 \end{pmatrix} \end{matrix}$$

y T es una matriz triangular superior no singular de rango κ .

- Π es la matriz obtenida del producto de las permutaciones elementales, que realizamos en el proceso de realizar la factorización QR del jacobiano, por lo que Π es una matriz ortogonal.

Si suponemos $\lambda = 0$ y además utilizamos (iii.3.1.3) el problema (iii.3.1.2) nos quedará :

$$\min_{\bar{s}} \| (Q' J_G \Pi)(\Pi' \bar{s}) + Q' G \|^2 \quad (\text{iii.3.1.4})$$

para simplificar sólo tomemos de (iii.3.1.4) la siguiente parte :

$$(Q' J_G \Pi)(\Pi' \bar{s}) + Q' G = \begin{bmatrix} T & S \\ 0 & 0 \end{bmatrix} (\Pi' \bar{s}) + Q' G \quad (\text{iii.3.1.5})$$

para multiplicar por bloques haremos las siguientes asignaciones:

$$\Pi' \bar{s} = \begin{bmatrix} y \\ z \end{bmatrix} \quad \text{y} \quad Q' G = \begin{bmatrix} c \\ d \end{bmatrix}$$

sustituímos en (iii.3.1.5) para obtener:

$$\begin{aligned} (Q' J_G \Pi)(\Pi' \bar{s}) + Q' G &= \begin{bmatrix} T & S \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} \\ &= \begin{bmatrix} Ty + Sz \\ 0 \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} \end{aligned}$$

entonces el problema de mínimos cuadrados (iii.3.1.4) lo transformamos a:

$$\min_{y,z} \left\| \begin{bmatrix} Ty + Sz \\ 0 \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} \right\|^2$$

o que es lo mismo y más claro:

$$\min_{y,z} \left\{ \| Ty + (Sz + c) \|^2 + \| d \|^2 \right\}$$

como tenemos un problema de minimización y $\| d \|^2$ es una constante es equivalente resolver :

$$\min_{y,z} \| Ty + (Sz + c) \|^2$$

ahora, sabemos que T es una matriz triangular superior y no singular; por lo que dado z implica que :

$$y = -T^{-1}(Sz + c)$$

como T es una matriz de orden κ , tenemos un sistema lineal cuadrado, entonces las soluciones son de la forma:

$$\bar{s} = \Pi \begin{bmatrix} -T^{-1}(Sz + c) \\ z \end{bmatrix}$$

y $z \in \mathbb{R}^{l-\kappa}$; por lo que obtenemos una infinidad de soluciones, en particular asignaremos cero a z , por lo que $\bar{s}(0)$ tiene la forma:

$$\begin{aligned}\bar{s}(0) &= \Pi \begin{bmatrix} -T^{-1}c \\ 0 \end{bmatrix} \\ &= \Pi \begin{bmatrix} -T^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} \\ &= \Pi \begin{bmatrix} -T^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^t G\end{aligned}$$

si denotamos a la pseudoinversa de J_G como:

$$J_G^- = \Pi \begin{bmatrix} -T^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^t$$

entonces finalmente escribimos a $\bar{s}(0)$ como:

$$\bar{s}(0) = J_G^- Q^t G.$$

esto es todo lo que se refiere al caso $\lambda = 0$, por lo que ahora nos corresponde analizar para el caso $\lambda > 0$, por lo que resolveremos el problema:

$$\min_{\bar{s}} \left\| \begin{bmatrix} J_G \\ \lambda^{\frac{1}{2}} D \end{bmatrix} + \begin{bmatrix} G \\ 0 \end{bmatrix} \right\|^2 \quad (iii.3.1.6)$$

si utilizamos (iii.3.1.3) y como habíamos mencionado antes (iii.3.1.6) equivale resolver el problema:

$$\min_{\bar{s}} \left\| \begin{bmatrix} Q^t & 0 \\ 0 & \Pi^t \end{bmatrix} \begin{bmatrix} J_G \\ \lambda^{\frac{1}{2}} D \end{bmatrix} \Pi \Pi^t \bar{s} + \begin{bmatrix} Q^t & 0 \\ 0 & \Pi^t \end{bmatrix} \begin{bmatrix} G \\ 0 \end{bmatrix} \right\|^2 \quad (iii.3.1.7)$$

ya que la matriz $\begin{bmatrix} Q^t & 0 \\ 0 & \Pi^t \end{bmatrix}$ es una matriz ortogonal. Reducimos (iii.3.1.7) a la forma:

$$\min_{\bar{s}} \left\| \begin{bmatrix} Q^t J_G \Pi \\ \lambda^{\frac{1}{2}} \Pi^t D \Pi \end{bmatrix} \Pi^t \bar{s} + \begin{bmatrix} Q^t G \\ 0 \end{bmatrix} \right\|^2 \quad (iii.3.1.8)$$

para facilitar el manejo de (iii.3.1.8) denotaremos:

$$P = \begin{bmatrix} Q^t J_G \Pi \\ \lambda^{\frac{1}{2}} \Pi^t D \Pi \end{bmatrix}$$

aquí vale la pena que analicemos la estructura de esta matriz:

$$P = \begin{bmatrix} R' \\ 0 \\ D_\lambda \end{bmatrix}$$

donde:

- R' es una matriz de orden ℓ de la forma:

$$R' = \begin{bmatrix} T & S \\ 0 & 0 \end{bmatrix}.$$

- T es la misma que ya habíamos descrito.
- D_λ es una matriz de orden ℓ de la forma

$$D_\lambda = \lambda^{\frac{1}{2}} \Pi' D \Pi.$$

como observamos R' es una matriz triangular, si consideramos al jacobiano singular o casi singular, esta matriz también lo es.

En la segunda etapa calculamos la factorización QR de la matriz P lo cual lo podemos hacer por medio de una secuencia de rotaciones de Givens, obtenemos:

$$P = W \begin{bmatrix} R_\lambda \\ 0 \end{bmatrix}$$

donde:

- W es una matriz ortogonal.
- R_λ es una matriz triangular superior no singular de orden ℓ .

Con lo anterior el problema (iii.3.1.8) queda:

$$\min_{\bar{s}} \left\| \begin{bmatrix} R_\lambda \\ 0 \end{bmatrix} \Pi' \bar{s} + W \begin{bmatrix} Q'G \\ 0 \end{bmatrix} \right\|^2 \quad (\text{iii.3.1.9})$$

si denotamos:

$$W \begin{bmatrix} Q'G \\ 0 \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$

y sustitúmos en (iii.3.1.9) tenemos:

$$\min_{\bar{s}} \left\{ \| R_{\lambda} \Pi^t \bar{s} + c \|^2 + \| d \|^2 \right\} \quad (\text{iii.3.1.9})$$

puesto que $\| d \|^2$ no involucra a \bar{s} , es equivalente resolver el problema:

$$\min_{\bar{s}} \| R_{\lambda} \Pi^t \bar{s} + c \|^2 \quad (\text{iii.3.1.10})$$

por lo que la solución al problema (iii.3.1.6) cuando $\lambda > 0$ es:

$$\bar{s}(\lambda) = -\Pi R_{\lambda}^{-1} c$$

lo que debemos hacer notar es que en caso de que λ cambie, sólo calcularemos la segunda etapa.

III.3.2 Actualización del radio de la región de confianza.

La actualización del radio Δ dependerá si la aproximación lineal (iii.3.2) de la función $\| G(\bar{\eta} + \bar{s}) \|^2$ es buena en la región de confianza determinada por Δ , una forma de medir esto es por medio del **cociente de reducción**, que está definido como el cociente de la **reducción real** entre la **reducción esperada**, es decir:

$$\rho(\bar{s}) = \frac{\| G(\bar{\eta}) \|^2 - \| G(\bar{\eta} + \bar{s}) \|^2}{\| G(\bar{\eta}) \|^2 - \| G(\bar{\eta}) + J_G(\bar{\eta}) \bar{s} \|^2}$$

Notamos fácilmente que si G es una función lineal entonces $\rho(\bar{s}) = 1$ para toda \bar{s} y si $J_G(\bar{\eta})^t G(\bar{\eta}) \neq 0$ entonces $\rho(\bar{s})$ tiende a uno si la norma de \bar{s} tiende a cero.

Para calcular eficientemente $\rho(\bar{s})$ podemos escribirlo como:

$$\rho = \frac{1 - \frac{\| G(\bar{\eta} + \bar{s}) \|^2}{\| G(\bar{\eta}) \|^2}}{\frac{\| J_G(\bar{\eta}) \bar{s} \|^2}{\| G(\bar{\eta}) \|^2} + 2\lambda \frac{\| Ds \|^2}{\| G(\bar{\eta}) \|^2}} \quad (\text{iii.3.2.2})$$

donde el desarrollo para llegar a esta expresión lo explicamos en el apéndice D.

En (iii.3.2.2), observamos que si $\| G(\bar{\eta} + \bar{s}) \|^2 \geq \| G(\bar{\eta}) \|^2$ entonces $\rho(\bar{s}) \leq 0$, pero en nuestro algoritmo no aceptaremos valores negativos, por lo que en esta

situación asignaremos $\rho = 0$. Entonces el rango de valores de ρ es el intervalo $[0, 1]$.

El criterio de actualización que generalmente se utiliza en la práctica es el siguiente:

- Si $\rho \geq \frac{3}{4}$ entonces incrementamos el valor de Δ , multiplicándola por una constante mayor que uno.
- Si $\rho < \frac{3}{4}$ y $\rho > \frac{1}{4}$ no cambiamos el valor de Δ .
- Si $\rho \leq \frac{1}{4}$ entonces decrementamos el valor de Δ .

Para decrementar el radio Δ , seguiremos la propuesta de Fletcher, la cual nos sugiere calcular el mínimo de la función cuadrática $q(\theta)$ que ajusta a los puntos $\delta(0)$, $\delta'(0)$ y $\delta(1)$ donde la función δ está definida como:

$$\delta(\theta) = \frac{1}{2} \| G(\bar{\eta} + \theta) \|^2$$

que realizando los cálculos, localizados en el apéndice D, obtenemos la siguiente expresión para q :

$$q(\theta) = \frac{1}{2} \| G \|^2 + (\bar{s}^t J_G^t G) \theta + \left(\frac{1}{2} \| G_+ \|^2 - \frac{1}{2} \| G \|^2 - \bar{s}^t J_G^t G \right) \theta^2$$

si denotamos el mínimo de esta cuadrática como μ , su expresión correspondiente es:

$$\mu = \frac{\frac{1}{2} \tau}{\frac{1}{2} \left[1 - \frac{\| G_+ \|^2}{\| G \|^2} \right] + \tau}$$

donde $\tau = - \left(\frac{\| J_G^t \|^2}{\| G \|^2} + \lambda \frac{\| D_s \|^2}{\| G \|^2} \right)$.

Entonces para decrementar el radio Δ , seguimos el siguiente criterio:

- Si $\mu \in [\frac{1}{10}, \frac{1}{2}]$ entonces decrementamos Δ multiplicándola por μ .
- Si $\mu \notin [\frac{1}{10}, \frac{1}{2}]$ entonces multiplicamos a Δ por el extremo más cercano de este intervalo a μ .

III.3.3 El parámetro de Levenberg-Marquardt.

Calculamos el parámetro de Levenberg-Marquardt, como un número positivo λ que satisface:

$$\| Ds(\lambda) \| - \Delta = 0$$

si denotamos a esta función como $\phi(\lambda)$, no es necesario calcular la raíz exacta de esta función puesto que éste no es nuestro objetivo final, sino que sólo queremos una aproximación de ésta, para lo cual calcularemos λ tal que :

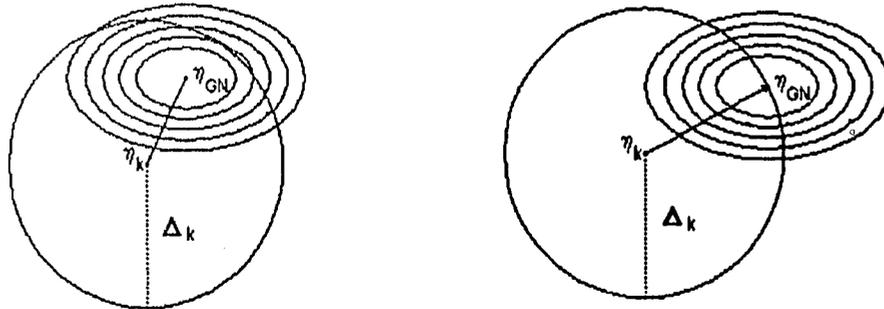
$$|\phi(\lambda)| \leq \sigma \Delta \quad (\text{iii.3.3.1})$$

donde $\sigma \in (0, 1)$ es el error relativo en $\| Ds(\lambda) \|$, en la práctica $\sigma = 10^{-1}$ produce buenos resultados, para obtener una aproximación de la raíz de (iii.3.3.1).

Aquí es importante considerar un caso especial de $\phi(\lambda)$.

Caso Especial.

Si $\phi(0) \leq 0$ entonces las gráficas correspondientes son:



Figuras III.6 y III.7: La solución de Newton está en la región de confianza, si $\phi(0) \leq 0$.

donde $\bar{\eta}_{GN}$ es la solución de Newton. Si $\bar{\eta}_{GN}$ pertenece a la región de confianza entonces el valor del parámetro λ es igual a cero, por lo que a partir de este

momento suponemos que λ es mayor que cero, para este caso la gráfica es:

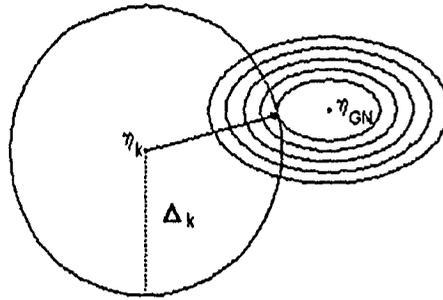


Figura III.8: Si $\lambda > 0$ la solución de Newton no pertenece a la región de confianza.

Es de gran ayuda saber como es la gráfica de una función, en este caso la de $\phi(\lambda)$, para esto recordemos que en la sección III.3 concluimos que $\bar{s}(\lambda)$ es una función decreciente y como demostramos en el apéndice D, que la derivada de $\phi(\lambda)$ es siempre negativa podemos concluir que $\phi(\lambda)$ es una función continua estrictamente decreciente convexa para $\lambda \geq 0$ y $\phi(\lambda)$ se aproxima a $-\Delta$ cuando λ tiende a infinito, por lo que existe una única $\lambda^* > 0$ tal que $\phi(\lambda^*) = 0$, como se muestra en la siguiente gráfica:

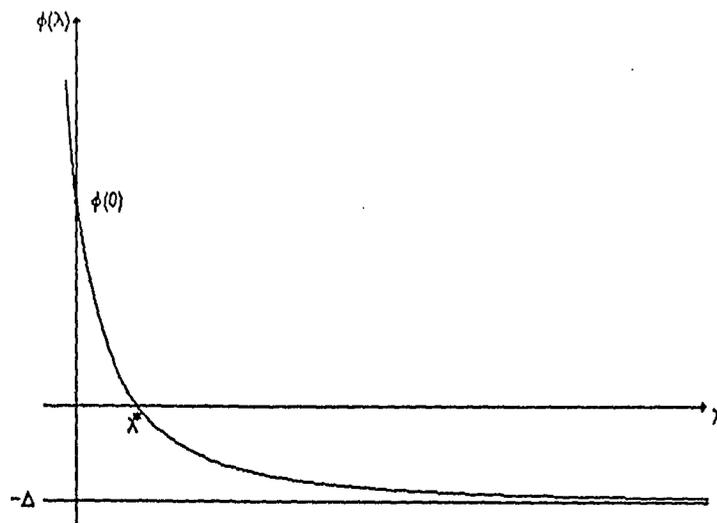


Figura III.9: Gráfica de la función $\phi(\lambda) = \|Ds(\lambda)\| - \Delta$.

por lo que ahora uno de nuestros problemas es determinar una aproximación a λ^* , para esto es usual que, dado λ_0 , un valor inicial, utilizar el método de Newton para generar una sucesión λ_k que converja a λ^* . Pero por la estructura particular de este problema nos permite derivar una iteración mucho más eficiente, para construir esta sucesión, la cual fué propuesta por Hebden (1973).

Cálculo de la aproximación del cero de la función $\phi(\lambda)$.

Sabemos que \bar{s} satisface las ecuaciones (iii.3.1.1), es decir:

$$(J_G^t J_G + \lambda D^t D) \bar{s} = -J_G^t G$$

de éstas tenemos el siguiente desarrollo:

$$D^t (D^{-t} J_G^t J_G D^{-1} + \lambda I) D \bar{s} = -J_G^t G$$

multiplicamos por D^{-t} :

$$(D^{-t} J_G^t J_G D^{-1} + \lambda I) D \bar{s} = -D^{-t} J_G^t G$$

si denotamos como $\tilde{J} = J_G D^{-1}$ nos queda:

$$(\tilde{J}^t \tilde{J} + \lambda I) D \bar{s} = -\tilde{J}^t G \quad (\text{iii.3.3.2})$$

calculamos la descomposición de valores singulares de \tilde{J} ,

$$\tilde{J} = U \Sigma V^t \quad (\text{iii.3.3.3})$$

donde:

- U y V son matrices ortogonales.
- Σ es una matriz diagonal, la cual la escribimos como:

$$\Sigma = \text{diag}\{\sigma_i : \sigma_i \text{ es el } i\text{-ésimo valor singular de } \tilde{J} \text{ para } i = 1, \dots, \ell\}$$

sustituimos \tilde{J} por su descomposición singular en (iii.3.3.2), tenemos:

$$[(V \Sigma U^t)(U \Sigma V^t) + \lambda I] D \bar{s} = -V \Sigma U^t G$$

$$[(V \Sigma^2 V^t) + \lambda I] D \bar{s} = -V \Sigma U^t G$$

$$V(\Sigma^2 + \lambda I)V^t D \bar{s} = -V \Sigma U^t G$$

multiplicamos por V^t y denotamos $z = U^t G$:

$$(\Sigma^2 + \lambda I)V^t D\bar{s} = -\Sigma z$$

multiplicamos por la inversa de $(\Sigma^2 + \lambda I)$:

$$V^t D\bar{s} = -(\Sigma^2 + \lambda I)^{-1} \Sigma z$$

$$V^t D\bar{s} = -(\Sigma^2 + \lambda I)^{-1} \begin{bmatrix} \sigma_1 z_1 \\ \sigma_2 z_2 \\ \vdots \\ \sigma_\ell z_\ell \end{bmatrix}$$

donde z_i son los elementos del vector z , simplificando obtenemos:

$$V^t D\bar{s} = \begin{bmatrix} \frac{\sigma_1 z_1}{\sigma_1^2 + \lambda} \\ \frac{\sigma_2 z_2}{\sigma_2^2 + \lambda} \\ \vdots \\ \frac{\sigma_\ell z_\ell}{\sigma_\ell^2 + \lambda} \end{bmatrix}$$

finalmente obtenemos la norma de este vector como:

$$\|D\bar{s}\|^2 = \|V^t D\bar{s}\|^2 = \sum_{i=1}^{\ell} \left[\frac{\sigma_i z_i}{\sigma_i^2 + \lambda} \right]^2$$

por lo que concluiremos que:

$$\phi(\lambda) = \left[\sum_{i=1}^{\ell} \left[\frac{\sigma_i z_i}{\sigma_i^2 + \lambda} \right]^2 \right]^{\frac{1}{2}} - \Delta$$

de esta igualdad, es natural proponer la siguiente aproximación a la función ϕ :

$$\phi(\lambda) \approx \frac{a}{b + \lambda} - \Delta$$

donde $(b + \lambda) \neq 0$ y $\bar{\phi}(\lambda) = \frac{a}{b + \lambda} - \Delta$, calculamos las constantes a y b suponiendo conocida λ_k y además las siguientes igualdades:

$$\bar{\phi}(\lambda_k) = \phi(\lambda_k)$$

$$\bar{\phi}'(\lambda_k) = \phi'(\lambda_k)$$

de estas condiciones obtenemos el sistema:

$$\frac{a}{b + \lambda_k} - \Delta = \phi(\lambda_k) \quad (\text{iii.3.3.4})$$

$$-\frac{a}{(b + \lambda_k)^2} = \phi'(\lambda_k) \quad (\text{iii.3.3.5})$$

de (iii.3.3.4) obtenemos:

$$a = (\phi(\lambda_k) + \Delta)(b + \lambda_k)$$

y de (iii.3.3.5) :

$$-a = \phi'(\lambda_k)(b + \lambda_k)^2$$

de estas igualdades obtenemos la siguiente ecuación:

$$\gamma[\phi'(\lambda_k)\gamma + \phi(\lambda_k) + \Delta] = 0$$

donde $\gamma = b + \lambda_k$, como $\gamma \neq 0$ es equivalente resolver:

$$\phi'(\lambda_k)\gamma + \phi(\lambda_k) + \Delta = 0$$

de esta ecuación obtenemos los valores para a y b :

$$a = -\frac{[\phi(\lambda_k) + \Delta]^2}{\phi'(\lambda_k)}$$

$$b = -\frac{[\phi(\lambda_k) + \Delta]}{\phi'(\lambda_k)} - \lambda_k$$

sustituimos estos valores en el modelo $\tilde{\phi}(\lambda)$ tenemos:

$$\tilde{\phi}(\lambda) = \frac{\frac{[\phi(\lambda_k) + \Delta]^2}{\phi'(\lambda_k)}}{\frac{[\phi(\lambda_k) + \Delta]}{\phi'(\lambda_k)} + \lambda_k - \lambda} - \Delta$$

puesto que queremos una aproximación a λ^* , tal que $\phi(\lambda^*) = 0$, entonces λ_{k+1} tiene que satisfacer $\tilde{\phi}(\lambda_{k+1})$ por lo que tenemos:

$$\frac{\frac{[\phi(\lambda_k) + \Delta]^2}{\phi'(\lambda_k)}}{\frac{[\phi(\lambda_k) + \Delta]}{\phi'(\lambda_k)} + \lambda_k - \lambda_{k+1}} - \Delta = 0$$

$$\frac{[\phi(\lambda_k) + \Delta]^2}{\Delta \phi'(\lambda_k)} = \frac{[\phi(\lambda_k) + \Delta]}{\phi'(\lambda_k)} + \lambda_k - \lambda_{k+1}$$

$$\begin{aligned}
\lambda_{k+1} &= \lambda_k - \frac{[\phi(\lambda_k) + \Delta]^2}{\Delta \phi'(\lambda_k)} + \frac{[\phi(\lambda_k) + \Delta]}{\phi'(\lambda_k)} \\
&= \lambda_k - \left[\frac{\phi(\lambda_k) + \Delta}{\phi'(\lambda_k)} \right] \left[\frac{\phi(\lambda_k) + \Delta}{\Delta} - 1 \right] \\
&= \lambda_k - \left[\frac{\phi(\lambda_k) + \Delta}{\Delta} \right] \left[\frac{\phi(\lambda_k)}{\phi'(\lambda_k)} \right].
\end{aligned}$$

Finalmente para calcular los elementos de la sucesión $\{\lambda_k\}$, suponemos que conocemos λ_k entonces el siguiente elemento lo calculamos como:

$$\lambda_{k+1} = \lambda_k - \left[\frac{\phi(\lambda_k)}{\phi'(\lambda_k)} \right] \left[\frac{\phi(\lambda_k) + \Delta}{\Delta} \right]$$

es muy parecida a la iteración de Newton, corregida por un valor, lo que nos permite una convergencia más rápida, además si estamos cerca de la solución implica que $\phi(\lambda_k) \approx 0$, por lo que es muy aproximada a la iteración de Newton.

Como ya hemos explicado como calcular la iteración del parámetro λ , para asegurarnos de su convergencia hacemos una iteración protegida, por lo que ponemos la condición a λ_k de pertenecer al intervalo $[\ell_k, u_k]$, es decir, asignar un límite superior u_k y un límite inferior ℓ_k al elemento λ_k , en caso que λ_k no pertenezca al intervalo, entonces la calculamos como la media geométrica de los extremos de los intervalos; antes de que escribamos este algoritmo; veremos como inicializar los extremos de este intervalo. Para el límite inferior supodremos que J_G es no singular, entonces $\phi'(0)$ está definida y por la convexidad de $\phi(\lambda)$ aseguramos que:

$$\ell_0 = -\frac{\phi(0)}{\phi'(0)}$$

es una cota inferior de λ^* , tal que $\phi(\lambda^*) = 0$, lo cual observamos en la siguiente

gráfica:

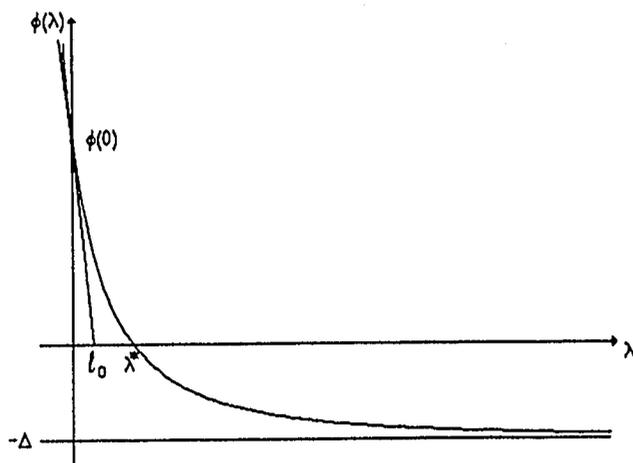


Figura III.10: Gráfica de la cota inferior inicial, ℓ_0 .

pero si J_G es singular entonces $\ell_0 = 0$. Para obtener el límite superior u_0 inicial, la calculamos como:

$$u_0 = \frac{\| (J_G D^{-1})^t G \|}{\Delta_0}$$

pero no vemos a simple vista el porqué de esta igualdad, por lo que la explicaremos a continuación, de la expresión (iii.3.3.2) obtenemos:

$$\phi(\lambda) = \left\| (\tilde{J}^t \tilde{J} + \lambda I)^{-1} \tilde{J}^t G \right\| - \Delta \quad (\text{iii.3.3.6})$$

usamos nuevamente (iii.3.3.3) la descomposición de valores singulares de \tilde{J} obtenemos la siguiente igualdad:

$$(\tilde{J}^t \tilde{J} + \lambda I)^{-1} = V^t (\Sigma^2 + \lambda I)^{-1} V$$

obtenemos la norma de esta matriz para obtener una desigualdad que nos será de gran utilidad:

$$\begin{aligned} \left\| (\tilde{J}^t \tilde{J} + \lambda I)^{-1} \right\| &= \left\| V^t (\Sigma^2 + \lambda I)^{-1} V \right\| \\ &\leq \| V^t \| \left\| (\Sigma^2 + \lambda I)^{-1} \right\| \| V \| \\ &= \left\| (\Sigma^2 + \lambda I)^{-1} \right\| \\ &\leq \max_{i=1, \dots, \ell} \left\{ \frac{1}{\sigma_i^2 + \lambda} \right\} \\ &\leq \frac{1}{\lambda} \end{aligned}$$

finalmente obtenemos una cota superior, es decir, tenemos:

$$\|(\tilde{J}^t \tilde{J} + \lambda I)^{-1}\| \leq \frac{1}{\lambda}$$

después acotamos (iii.3.3.6) de la siguiente forma:

$$\begin{aligned} \phi(\lambda) &\leq \|(\tilde{J}^t \tilde{J} + \lambda I)^{-1}\| \| \tilde{J}^t G \| - \Delta \\ &\leq \frac{1}{\lambda} \| \tilde{J}^t G \| - \Delta = \phi_a(\lambda) \end{aligned} \quad (iii.3.3.7)$$

como queremos una cota superior de $\phi(\lambda^*)$, de la desigualdad (iii.3.3.7) observamos:

- $0 = \phi(\lambda^*) \leq \phi_a(\lambda^*)$
- $\phi(u_0) \leq \phi_a(u_0) = 0$

las cual las observamos en la siguiente gráfica:

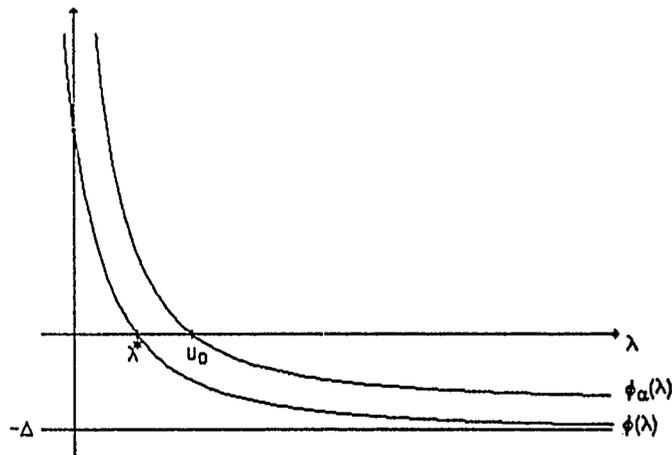


Figura III.11: Gráfica de las funciones de $\phi(\lambda)$ y $\phi_a(\lambda)$.

observamos que el valor de la función ϕ en una cota superior de λ^* debe ser negativa o igual a cero, esto se refleja en (iii.3.3.7), de la siguiente forma:

$$\phi(\lambda) \leq 0$$

por lo tanto:

$$\frac{1}{u_0} \| \tilde{J}^t G \| - \Delta = 0$$

de aquí obtenemos la expresión a la que queríamos llegar:

$$u_0 = \frac{\|\tilde{J}^t G\|}{\Delta}$$

Con estos elementos ya podemos escribir el algoritmo para calcular el parámetro λ de Leverberg-Marquardt:

ALGORITMO :

Dadas $\lambda_0, u_0, \ell_0, \Delta, \sigma = 10^{-1}$.

Para $j = 0, 1, \dots$, converja

1. Si $\lambda_j \notin (\ell_j, u_j)$ entonces

$$\lambda_j = \max \left\{ 10^{-3} u_j, (\ell_j u_j)^{\frac{1}{2}} \right\}$$

2. Evaluar $\phi(\lambda_j)$ y $\phi'(\lambda_j)$

Actualizar u_j

Si $\phi(\lambda_j) < 0$ entonces

$$u_{j+1} = \lambda_k$$

sino

$$u_{j+1} = u_j$$

Actualizar ℓ_j como:

$$\ell_{k+1} = \max \left\{ \ell_j, \lambda_j - \frac{\phi(\lambda_j)}{\phi'(\lambda_j)} \right\}$$

3. Obtener λ_{j+1} como:

$$\lambda_{j+1} = \lambda_j - \left[\frac{\phi(\lambda_j)}{\phi'(\lambda_j)} \right] \left[\frac{\phi(\lambda_j) + \Delta}{\Delta} \right]$$

4. Si $|\phi(\lambda_{j+1})| > \sigma \Delta$ entonces

ir al paso 1.

sino TERMINAMOS

En el paso 1., como ya habíamos mencionado, cuidamos que $\lambda^i \in [\ell_k, u_k]$, en caso de que no suceda esto, se reemplaza el valor de λ_k por el punto $(\ell_k u_k)^{\frac{1}{2}}$, el

cual está sesgado en dirección a ℓ_k , pero para prever el caso de valores $\ell_k \ll 1$ y en particular $\ell_k = 0$; λ_k le asignamos el valor de $10^{-3}u_k$.

En el paso 2. para actualizar u_k y ℓ_k , observamos la siguiente gráfica:

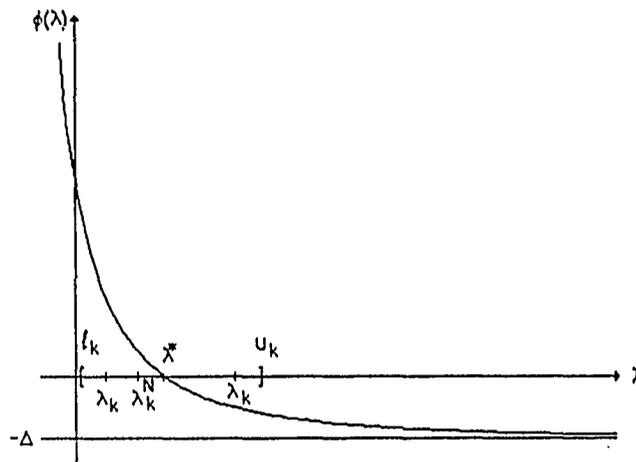


Figura III.12: Actualización del intervalo $[\ell_k, u_k]$.

Primer caso

Como $\lambda \in [\ell_k, u_k]$ en este caso consideramos que $\phi(\lambda_k) \leq 0$ lo que implica la desigualdad:

$$\lambda^* < \lambda_k < u_k$$

entonces u_{k+1} toma el valor de λ_k .

Segundo caso

Si $\phi(\lambda_k) > 0$ entonces actualizamos ℓ_k tomando el valor que esté más cerca de λ^* entre λ_k y la iteración de Newton que en este caso es:

$$\lambda_k^N = \lambda_k - \frac{\phi(\lambda_k)}{\phi'(\lambda_k)}$$

No es muy difícil mostrar que este algoritmo siempre genera una sucesión que converge cuadráticamente a λ^* . En la práctica sólo dos iteraciones son requeridas aproximadamente para satisfacer el criterio de parada cuando $\sigma = 10^{-4}$.

III.3.4 Resultados teóricos.

El siguiente algoritmo nos lo propone Moré [18], de acuerdo a los detalles teóricos que explicamos en las secciones anteriores:

ALGORITMO (iii.3.4.1):

Dada $\sigma \in (0, 1)$.

1. Si $\|D_k J_{G_k}^- G_k\| \leq (1 + \sigma)\Delta_k$ asignar $\lambda_k = 0$ y $\bar{s}_k = J_{G_k}^- G_k$ en el otro caso determinar $\lambda_k > 0$ tal que si

$$\begin{bmatrix} J_{G_k} \\ \lambda_k^{\frac{1}{2}} D_k \end{bmatrix} \bar{s}_k = - \begin{bmatrix} G_k \\ 0 \end{bmatrix}$$

entonces

$$(1 - \sigma)\Delta_k \leq \|D_k \bar{s}_k\| \leq (1 + \sigma)\Delta_k.$$

2. Calcular el cociente de reducción ρ_k .
3. Si $\rho_k \leq 0.0001$ asignar $\eta_{k+1} = \eta_k$ y $J_{G_{k+1}} = J_{G_k}$.
4. Si $\rho_k > 0.0001$ asignar $\eta_{k+1} = \eta_k + \bar{s}_k$ y calcular $J_{G_{k+1}}$.
5. Si $\rho_k \leq \frac{1}{4}$ asignar a Δ_{k+1} un punto del intervalo

$$\left[\frac{1}{10\Delta_k}, \frac{1}{2}\Delta_k \right].$$

6. Si $\rho_k \in \left[\frac{1}{4}, \frac{3}{4} \right]$ y $\lambda_k = 0$ o si $\rho_k \geq \frac{3}{4}$ asignar

$$\Delta_{k+1} = 2 \|D_k \bar{s}_k\|.$$

7. Actualizar D_{k+1} .

La prueba de los resultados de convergencia es un tanto largo y por esta razón no será incluida.

Teorema : Sea $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función continua diferenciable en \mathbb{R}^n y $\{\eta_k\}$ una sucesión generada por el algoritmo (iii.3.4.1) entonces

$$\liminf_{k \rightarrow +\infty} \|(J_{G_k} D_k^{-1})^t G_k\|^2 = 0. \quad (\text{iii.3.4.2})$$

Este resultado garantiza que eventualmente el tamaño del gradiente sea pequeño a medida que k tienda a infinito.

Por supuesto, si $\{J_{G_k}\}$ es acotada entonces (iii.3.4.2) implica el siguiente resultado que es más estándar.

$$\liminf_{k \rightarrow +\infty} \|J_{G_k}^t G_k\|^2 = 0 \quad (\text{iii.3.4.3})$$

Además podemos demostrar que si G' es uniformemente continua entonces:

$$\lim_{k \rightarrow +\infty} \|J_{G_k}^t G_k\|^2 = 0 \quad (\text{iii.3.4.4})$$

Powell y Osborne, también han obtenido resultados de convergencia global para sus versiones del algoritmo de Levenberg-Marquardt. Powell expone un algoritmo general para minimización sin restricciones la cual es un caso especial de (iii.3.4.1) con $\sigma = 0$ y $\{D_k\}$ constante. Para este caso Powell obtiene (iii.3.4.3) bajo la suposición que $\{J_{G_k}\}$ es acotada.

El algoritmo de Osborne controla directamente $\{\lambda_k\}$ en vez de $\{\Delta_k\}$ y permite que $\{D_k\}$ se elija por (6.1) y (6.2). Para este caso él demuestra (iii.3.4.4) bajo la suposición que $\{J_{G_k}\}$ y $\{\lambda_k\}$ son acotadas.

Capítulo IV

Problema de ajuste ortogonal de parámetros del modelo no-lineal.

Ya hemos estudiado el desarrollo del ajuste de parámetros por medio de distancias ortogonales para el caso lineal, ahora discutiremos este ajuste para funciones no-lineales, del cual también aprovecharemos su estructura para disminuir la misma cantidad de parámetros que en el lineal.

Primero consideramos el ajuste al modelo lineal general, del que pensamos, que tal vez podamos aun explotar su estructura de manera más eficiente que en el no-lineal y si esto no es posible, nos servirá como una introducción a la solución de este problema.

IV.1 Planteamiento del problema de ajuste ortogonal del modelo lineal general.

El modelo lineal general está definido como:

$$m(x; \beta) = \beta_0 \varphi_0(x) + \beta_1 \varphi_1(x) + \cdots + \beta_p \varphi_p(x) \quad (iv.1.1)$$

donde $\varphi_i : \mathbb{R} \rightarrow \mathbb{R}$ son funciones no-lineales con respecto a x . Nuestro problema es ajustar el conjunto de datos:

$$\{(x_i, y_i) : \text{para } i = 1, \dots, n\} \quad (iv.1.2)$$

al modelo (iv.1.1), resolvemos este problema minimizando la suma de los cuadrados de las distancias ortogonales del dato i al modelo m , que gráficamente lo

observamos así:

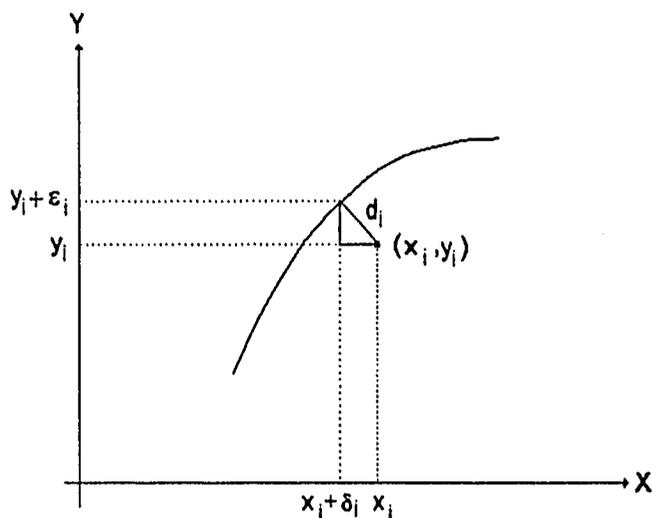


Figura IV.1: Idea gráfica del ajuste ortogonal no-lineal.

Si seguimos la misma notación del capítulo II, entonces d_i es la distancia del punto (x_i, y_i) a otro punto $(x_i + \delta_i, y_i + \epsilon_i)$ sobre la curva m , por lo que tenemos la siguiente igualdad, en términos vectoriales:

$$\|\bar{\mathbf{d}}\|^2 = \|\bar{\boldsymbol{\epsilon}}\|^2 + \|\bar{\boldsymbol{\delta}}\|^2 \quad (iv.1.3)$$

así el planteamiento del problema de ajuste ortogonal del modelo lineal general lo planteamos de la siguiente manera:

$$\min_{\bar{\boldsymbol{\beta}}, \bar{\boldsymbol{\delta}}} \frac{1}{2} \|\bar{\mathbf{d}}\|^2 \quad (iv.1.4)$$

$$\text{sujeto a } \bar{\mathbf{y}} = \beta_0 \bar{\varphi}_0 + \beta_1 \bar{\varphi}_1 + \dots + \beta_p \bar{\varphi}_p - \bar{\boldsymbol{\epsilon}}$$

donde las funciones $\bar{\varphi}_i$ para $i = 0, \dots, p$, las consideraremos que sólo dependen del vector de parámetros $\bar{\boldsymbol{\delta}}$, puesto que el vector de las abscisas de los datos $\bar{\mathbf{x}}$ está predeterminado, por lo que las definimos de la siguiente manera:

$$\bar{\varphi}_i(\bar{\boldsymbol{\delta}}) = \begin{bmatrix} \varphi_i(x_1 + \delta_1) \\ \varphi_i(x_2 + \delta_2) \\ \vdots \\ \varphi_i(x_n + \delta_n) \end{bmatrix}$$

entonces $\bar{\varphi}_i$ es una función con dominio y rango en \mathbb{R}^n .

Utilizando la igualdad (iv.1.3), podemos transformar el planteamiento (iv.1.4) a un problema sin restricciones, de la siguiente forma:

$$\min_{\bar{\eta}} \frac{1}{2} \left\{ \left\| \beta_0 \bar{\varphi}_0 + \beta_1 \bar{\varphi}_1 + \cdots + \beta_p \bar{\varphi}_p - \bar{y} \right\|^2 + \left\| \bar{\delta} \right\|^2 \right\} \quad (iv.1.5)$$

donde $\bar{\eta} = \begin{pmatrix} \beta \\ \bar{\delta} \end{pmatrix}$. En este caso el residuo ortogonal es:

$$\bar{r}_\delta(\bar{\eta}) = \beta_0 \bar{\varphi}_0 + \beta_1 \bar{\varphi}_1 + \cdots + \beta_p \bar{\varphi}_p - \bar{y}$$

si definimos a la matriz A como:

$$A(\delta) = [\bar{\varphi}_0(\delta) | \bar{\varphi}_1(\delta) | \cdots | \bar{\varphi}_p(\delta)] \quad (iv.1.6)$$

entonces el residuo lo escribimos de manera más compacta:

$$\bar{r}_\delta(\bar{\eta}) = A(\delta) \bar{\beta} - \bar{y} \quad (iv.1.7)$$

por lo que un planteamiento equivalente a (iv.1.5) es:

$$\min_{\bar{\eta}} \frac{1}{2} \left\{ \left\| A(\delta) \bar{\beta} - \bar{y} \right\|^2 + \left\| \bar{\delta} \right\|^2 \right\}$$

observamos que si fijamos $\bar{\delta}$ entonces tenemos un problema de mínimos cuadrados lineales, es decir, resolveríamos un ajuste ordinario de parámetros, por lo que claramente se refleja la estructura del modelo lineal general.

Con la observación anterior, pensamos que dado $\bar{\delta}_0$ podríamos determinar un óptimo para $\bar{\beta}_0$, con estos dos linearizar la ecuación cuadrática, alrededor de $(\bar{\beta}_0, \bar{\delta}_0)$, corregir $\bar{\delta}_1 = \bar{\delta}_0 + \Delta \bar{\delta}_0$ y repetir el proceso, esto lo vamos hacer de manera muy simple usando el método de Levenberg-Marquardt, así como la estructura del jacobiano de la función G .

De acuerdo a la definición de norma de un vector, podemos construir la función $G: \mathbb{R}^{n+p} \rightarrow \mathbb{R}^{2n}$, de la siguiente forma:

$$G(\bar{\eta}) = \begin{bmatrix} \bar{r}_\delta(\bar{\eta}) \\ \bar{\delta} \end{bmatrix} \quad (iv.1.8)$$

y así escribir el planteamiento del problema (iv.1.5) como:

$$\min_{\bar{\eta}} \frac{1}{2} \left\| G(\bar{\eta}) \right\|^2$$

entonces tenemos un problema tradicional de mínimos cuadrados no-lineales, pero con el defecto que tenemos $n + p$ parámetros, como consecuencia del ajuste ortogonal, pues en el caso ordinario sólo resolveríamos un problema de mínimos cuadrados lineales con p parámetros, es decir estos se incrementan tanto como el número de datos, que en la práctica suele ser muy grande, por lo que explicaremos, en la siguiente sección, como resolver este problema reduciendo el número de parámetros a p .

IV.2 Solución del problema de ajuste ortogonal del modelo lineal general.

Para resolver el problema de ajuste ortogonal al modelo lineal general:

$$\min_{\bar{\eta}} \frac{1}{2} \| G(\bar{\eta}) \|^2 \quad (iv.2.1)$$

nos basaremos en el método de Levenberg-Marquardt. De acuerdo a la notación utilizada en el capítulo tres, la k -ésima iteración de este método aplicada a (iv.2.1) es:

- Suponemos conocida $\bar{\eta}_k$.
- Obtenemos \bar{s}_k , solución del problema:

$$\min_{\| D\bar{s} \|^2 \leq \Delta^2} \frac{1}{2} \| G_k + J_{G_k} \bar{s} \|^2 \quad (iv.2.2)$$

donde $\bar{s} = \bar{\eta} - \bar{\eta}_k$ y D es una matriz diagonal de escalamiento.

- $\bar{\eta}_{k+1} = \bar{\eta}_k + \bar{s}_k$.

por el procedimiento del método, \bar{s}_k , satisface el siguiente sistema de ecuaciones:

$$(J_{G_k}^t J_{G_k} + \lambda D_k^t D_k) \bar{s}(\lambda) = -J_{G_k}^t G_k \quad (iv.2.3)$$

donde $\lambda > 0$, como la notación podría llegar a ser muy pesada por el subíndice de la iteración, k , lo omiteremos en la notación de las matrices pero lo dejaremos en la de los vectores para recordar que estamos utilizando información del paso anterior.

Si observamos el sistema de ecuaciones (iv.2.3) podemos concluir que son las ecuaciones normales correspondientes al siguiente problema de mínimos cuadrados no-lineales:

$$\min_s \left\| \begin{bmatrix} J_G \\ \lambda^{\frac{1}{2}} D \end{bmatrix} \bar{s} + \begin{bmatrix} G_k \\ 0 \end{bmatrix} \right\|^2 \quad (iv.2.4)$$

la idea principal es la explotación adecuada de la estructura del jacobiano de G , J_G , el cual, lo calculamos a partir de (iv.1.8) y así obtener:

$$J_G = \begin{matrix} & \begin{matrix} p & n \end{matrix} \\ \begin{matrix} n \\ n \end{matrix} & \begin{pmatrix} \frac{\partial \bar{r}_i(\bar{\eta})}{\partial \beta} & \frac{\partial \bar{r}_i(\bar{\eta})}{\partial \delta} \\ \frac{\partial \bar{\delta}}{\partial \beta} & \frac{\partial \bar{\delta}}{\partial \delta} \end{pmatrix} \end{matrix}$$

si consideramos los bloques de izquierda a derecha y de arriba hacia abajo, entonces cada bloque está definido como:

- El primer bloque es fácil de calcular a partir de (iv.1.7), puesto que esta expresión depende linealmente de $\bar{\beta}$, por lo tanto obtenemos:

$$\frac{\partial \bar{r}_i(\bar{\eta})}{\partial \beta} = A$$

donde la matriz A la definimos en (iv.1.6).

- El segundo bloque es una matriz diagonal, que denotaremos con V , la cual está definida como:

$$V = \text{diag} \left\{ \sum_{j=0}^p \beta_j \frac{\partial \varphi_j(x_i + \delta)}{\partial \delta_i}; \text{ para } i = 1, \dots, n \right\}$$

por lo tanto obtenemos:

$$\frac{\partial \bar{r}_i(\bar{\eta})}{\partial \delta} = V.$$

- El tercer bloque es la matriz cero de $n \times p$, puesto que derivamos $\bar{\delta}$ con respecto a $\bar{\beta}$, es decir:

$$\frac{\partial \bar{\delta}}{\partial \beta} = 0.$$

- El cuarto bloque es una matriz identidad de orden n , ya que estamos derivando $\bar{\delta}$ con respecto de él mismo, así obtenemos:

$$\frac{\partial \bar{\delta}}{\partial \bar{\delta}} = I_n .$$

Con esta explicación la estructura de J_G es de la forma:

$$J_G = \begin{matrix} & p & n \\ n & \begin{pmatrix} A & V \\ 0 & I_n \end{pmatrix} \\ n & \end{matrix} \quad (iv.2.5)$$

de acuerdo a esta estructura, es natural dividir a los vector \bar{s} y G_k de la siguiente forma:

$$\bar{s} = \begin{pmatrix} \bar{\beta}_s \\ \bar{\delta}_s \end{pmatrix} \quad y \quad G_k = \begin{pmatrix} \bar{r}_k \\ \bar{\delta}_k \end{pmatrix} \quad (iv.2.6)$$

donde $\bar{\beta}_s$ y $\bar{\delta}_s$ son de la forma:

$$\begin{aligned} \bar{\beta}_s &= \bar{\beta} - \bar{\beta}_k \quad y \quad \bar{\beta} \in \mathbb{R}^p \\ \bar{\delta}_s &= \bar{\delta} - \bar{\delta}_k \quad y \quad \bar{\delta} \in \mathbb{R}^n \end{aligned}$$

a la matriz D también la dividimos de acuerdo a las mismas dimensiones de la siguiente manera:

$$D = \begin{matrix} & p & n \\ p & \begin{pmatrix} \lambda^{\frac{1}{2}} S & 0 \\ 0 & \lambda^{\frac{1}{2}} T \end{pmatrix} \\ n & \end{matrix} \quad (iv.2.7)$$

entonces sustituyendo las igualdades (iv.2.5), (iv.2.6) y (iv.2.7) en el problema (iv.2.4), lo transformamos a:

$$\min_{\bar{\beta}_s, \bar{\delta}_s} \frac{1}{2} \left\| \begin{bmatrix} A & V \\ 0 & I_n \\ \lambda^{\frac{1}{2}} S & 0 \\ 0 & \lambda^{\frac{1}{2}} T \end{bmatrix} \begin{bmatrix} \bar{\beta}_s \\ \bar{\delta}_s \end{bmatrix} + \begin{bmatrix} \bar{r}_k \\ \bar{\delta}_k \\ 0 \\ 0 \end{bmatrix} \right\|^2$$

cuyas ecuaciones normales de este problema son:

$$\begin{bmatrix} A^t & 0 & \lambda^{\frac{1}{2}} S^t & 0 \\ V^t & I_n & 0 & \lambda^{\frac{1}{2}} T^t \end{bmatrix} \begin{bmatrix} A & V \\ 0 & I_n \\ \lambda^{\frac{1}{2}} S & 0 \\ 0 & \lambda^{\frac{1}{2}} T \end{bmatrix} \begin{bmatrix} \bar{\beta}_s \\ \bar{\delta}_s \end{bmatrix} = - \begin{bmatrix} A^t & 0 & \lambda^{\frac{1}{2}} S^t & 0 \\ V^t & I_n & 0 & \lambda^{\frac{1}{2}} T^t \end{bmatrix} \begin{bmatrix} \bar{r}_k \\ \bar{\delta}_k \\ 0 \\ 0 \end{bmatrix} \quad (iv.2.8)$$

como ya hemos mencionado D es una matriz de escalamiento no singular, no necesariamente diagonal; pero nosotros vamos a suponer que lo es para el resto del capítulo, por lo que denotaremos:

$$S^t S = S^2 \quad \text{y} \quad T^t T = T^2 .$$

Al utilizar esta notación y realizar las operaciones por bloques en (iv.2.8) obtenemos el siguiente sistema de ecuaciones no-lineales:

$$(A^t A + \lambda S^2) \bar{\beta}_s + A^t V \bar{\delta}_s = -A^t \bar{r}_k \quad (\text{iv.2.9})$$

$$V^t A \bar{\beta}_s + (V^t V + I_n + \lambda T^2) \bar{\delta}_s = -(V^t \bar{r}_k + I_n \bar{\delta}_k) \quad (\text{iv.2.10})$$

Si $P = V^t V + I_n + \lambda T^2$ entonces observamos que de (iv.2.10) podemos despejar $\bar{\delta}_s$ en términos de $\bar{\beta}_s$, de la siguiente forma:

$$\bar{\delta}_s = -P^{-1}(V^t A \bar{\beta}_s + V^t \bar{r}_k + I_n \bar{\delta}_k) \quad (\text{iv.2.11})$$

sustituimos $\bar{\delta}_s$ en las ecuaciones (iv.2.9) para obtener un sistema sólo en términos de $\bar{\beta}_s$:

$$[A^t(I_n - VP^{-1}V^t)A + \lambda S^2] \bar{\beta}_s = -A^t \bar{r}_k + A^t VP^{-1}(V^t \bar{r}_k + I_n \bar{\delta}_k)$$

si observamos cuidadosamente, podemos afirmar que son las ecuaciones normales del problema de minimización:

$$\min_{\bar{\beta}_s} \frac{1}{2} \left\| \begin{bmatrix} (I_n - VP^{-1}V^t)^{\frac{1}{2}} A \\ \lambda^{\frac{1}{2}} S \end{bmatrix} \bar{\beta}_s + \begin{bmatrix} (I_n - VP^{-1}V^t)^{\frac{1}{2}} [-\bar{r}_k + VP^{-1}(V^t \bar{r}_k + I_n \bar{\delta}_k)] \\ 0 \end{bmatrix} \right\|^2 \quad (\text{iv.2.12})$$

para poder reducir este problema, analicemos la matriz P de la siguiente forma:

- $V^t V$ es una matriz diagonal de la forma:

$$V^t V = \text{diag} \left\{ \left[\sum_{j=0}^p \beta_j \frac{\partial \varphi_j(x_i + \delta_i)}{\partial \delta_i} \right]^2 ; \quad \text{para } i = 1, \dots, n \right\}$$

esto lo concluimos a partir de (iv.2.5)

- T es una matriz diagonal de orden n , la denotamos como:

$$T = \text{diag} \{ \tau_i : \text{ para } i = 1, \dots, n \}$$

puesto que T es un bloque de la matriz de escalamiento D .

Con ésto llegamos a la conclusión que P es una matriz diagonal de la forma:

$$P = \text{diag} \left\{ \left[\sum_{j=0}^p \beta_j \frac{\partial \varphi_j(x_i + \delta_i)}{\partial \delta_i} \right]^2 + 1 + \lambda \tau_i^2; \text{ para } i = 1, \dots, n \right\}$$

donde denotamos a γ_i como:

$$\gamma_i = \sum_{j=0}^p \beta_j \frac{\partial \varphi_j(x_i + \delta_i)}{\partial \delta_i}.$$

Con esto es inmediata la inversa de la matriz P :

$$P^{-1} = \text{diag} \left\{ \frac{1}{\gamma_i^2 + 1 + \lambda \tau_i^2}; \text{ para } i = 1, \dots, n \right\}$$

ya es fácil deducir la forma de la siguiente matriz:

$$(I_n - VP^{-1}V^t) = \text{diag} \left\{ \frac{1}{\omega_i + 1}; \text{ para } i = 1, \dots, n \right\}$$

donde $\omega_i = \frac{\gamma_i^2}{1 + \lambda \tau_i^2}$ y como λ es positiva, no hay problemas para obtener la raíz cuadrada de los elementos de esta matriz:

$$(I_n - VP^{-1}V^t)^{\frac{1}{2}} = \text{diag} \left\{ \left[\frac{1}{\omega_i + 1} \right]^{\frac{1}{2}}; \text{ para } i = 1, \dots, n \right\}$$

a esta matriz la denotamos con la letra M , es decir:

$$M = (I_n - VP^{-1}V^t)^{\frac{1}{2}}$$

por lo tanto su inversa es también una matriz diagonal de la forma:

$$M^{-1} = \text{diag} \left\{ [\omega_i + 1]^{\frac{1}{2}}; \text{ para } i = 1, \dots, n \right\}$$

la importancia de la matriz M radica que a partir de esta matriz podemos simplificar la expresión del problema (iv.2.12), para lograr esto, si denotamos a $E = I_n + \lambda T^2$, podemos verificar fácilmente que se cumplen las siguientes igualdades:

$$\begin{aligned} \text{a) } VP^{-1} &= M^2VE^{-1} \\ \text{b) } P^{-1}V^t &= E^{-1}V^tM^2 \end{aligned} \quad (\text{iv.2.13})$$

por lo que escribimos el problema (iv.2.12) en términos de la matriz M , de la forma:

$$\min_{\bar{\beta}_s} \frac{1}{2} \left\| \begin{bmatrix} MA \\ \lambda^{\frac{1}{2}} S \end{bmatrix} \bar{\beta}_s + \begin{bmatrix} M^{-1}[-\bar{r}_k + VP^{-1}(V^t \bar{r}_k + I_n \bar{\delta}_k)] \\ 0 \end{bmatrix} \right\|^2 \quad (iv.2.14)$$

aun simplificamos un poco más la expresión de este problema utilizando las igualdades (iv.2.13) y obtenemos la siguiente equivalencia:

$$M^{-1}[-\bar{r}_k + VP^{-1}(V^t \bar{r}_k + I_n \bar{\delta}_k)] = M[-\bar{r}_k + VE^{-1}I_n \bar{\delta}_k]$$

que si observamos (iv.2.14) es obvio cual término simplificamos, por lo que lo podemos escribir como:

$$\min_{\bar{\beta}_s} \frac{1}{2} \left\| \begin{bmatrix} MA \\ \lambda^{\frac{1}{2}} S \end{bmatrix} \bar{\beta}_s + \begin{bmatrix} -M[\bar{r}_k - VE^{-1}I_n \bar{\delta}_k] \\ 0 \end{bmatrix} \right\|^2 \quad (iv.2.15)$$

esta estructura ya es muy conocida por nosotros, hemos llegado a un problema de mínimos cuadrados lineales para $\bar{\beta}_s$, el cual lo observaremos más fácilmente, si denotamos como:

$$\tilde{J} = MA \quad y \quad \tilde{y} = M[-\bar{r}_k + VE^{-1}I_n \bar{\delta}_k]$$

entonces escribimos a (iv.2.15) de la siguiente forma:

$$\min_{\bar{\beta}_s} \frac{1}{2} \left\| \begin{bmatrix} \tilde{J} \\ \lambda^{\frac{1}{2}} S \end{bmatrix} \bar{\beta}_s + \begin{bmatrix} \tilde{y} \\ 0 \end{bmatrix} \right\|^2.$$

Resolver este problema implica obtener la factorización QR de \tilde{J} y eliminar $\lambda^{\frac{1}{2}} S$ para obtener la solución $\bar{\beta}_s^*$ y con esta solución obtener $\bar{\delta}_s^*$ a partir de:

$$\bar{\delta}_s^* = -E[V^t M^2(\bar{r}_k + A\bar{\beta}_s^* - VE^{-1}I\delta_k) + I_n \delta_k]$$

esta última igualdad la obtenemos utilizando las igualdades (iv.2.13) en (iv.2.11).

Por lo que finalmente la $(k+1)$ -ésima iteración queda definida como:

$$\begin{aligned} \bar{\beta}_{k+1} &= \bar{\beta}_s^* + \bar{\beta}_k \\ \bar{\delta}_{k+1} &= \bar{\delta}_s^* + \bar{\delta}_k \end{aligned}$$

con esto cumplimos nuestro objetivo de resolver sólo para p parámetros.

IV.3 Ajuste ortogonal de parámetros del modelo no-lineal.

Ahora trataremos el ajuste ortogonal del modelo no-lineal de la forma:

$$f(x; \bar{\beta})$$

es decir, es el caso más general donde la función es no-lineal también con respecto al vector de parámetros $\bar{\beta}$, por lo que para nosotros es inmediato el planteamiento del problema:

$$\min_{\bar{\eta}} \frac{1}{2} \| G(\bar{\eta}) \|^2 \quad (iv.3.1)$$

donde la función G tiene la misma estructura que en los casos anteriores, es decir:

$$G(\bar{\eta}) = \begin{bmatrix} \bar{r}_\delta(\bar{\eta}) \\ \bar{\delta} \end{bmatrix} \quad \text{y} \quad \bar{\eta} = \begin{pmatrix} \bar{\beta} \\ \bar{\delta} \end{pmatrix}$$

pero en este caso, el vector residual ortogonal queda definido como:

$$\bar{r}_\delta(\bar{\eta}) = \begin{bmatrix} f(x_1 + \delta_1; \bar{\beta}) - y_1 \\ f(x_2 + \delta_2; \bar{\beta}) - y_2 \\ \vdots \\ f(x_n + \delta_n; \bar{\beta}) - y_n \end{bmatrix}$$

si resolvemos el problema (iv.3.1) por el método de Levenberg-Marquardt, entonces la iteración de este método está definida como:

- Suponemos conocida $\bar{\eta}_k$.
- Obtenemos \bar{s}_k , solución del problema:

$$\min_{\|D\bar{s}\|^2 \leq \Delta^2} \frac{1}{2} \| G_k + J_{G_k} \bar{s} \|^2 \quad (iv.3.2)$$

donde $\bar{s} = \bar{\eta} - \bar{\eta}_k$ y D es una matriz diagonal de escalamiento.

- $\bar{\eta}_{k+1} = \bar{\eta}_k + \bar{s}_k$.

Por el desarrollo de este método sabemos que resolver (iv.3.2) es equivalente a resolver el siguiente sistema de ecuaciones:

$$(J_{G_k}^T J_{G_k} + \lambda D_k^T D_k) \bar{s}(\lambda) = -J_{G_k}^T G_k$$

para $\lambda \geq 0$. Hasta este momento este desarrollo es igual que el del caso lineal general, incluso la estructura del jacobiano de G es la misma que la del caso anterior, es decir:

$$J_G = \begin{matrix} & p & n \\ n & \begin{pmatrix} A & V \\ 0 & I_n \end{pmatrix} \\ n & \begin{pmatrix} p & n \\ \frac{\partial f(x+\delta; \beta)}{\partial \beta} & \frac{f(x+\delta; \beta)}{\partial \delta} \\ \frac{\partial \bar{\delta}}{\partial \beta} & \frac{\partial \bar{\delta}}{\partial \delta} \end{pmatrix} \end{matrix}$$

calculando cada uno de estos bloques observamos que los dos últimos son iguales al del jacobiano del caso lineal general, por lo que sólo definimos los dos primeros bloques:

- A es una matriz de $n \times p$ que depende sólo de los parámetros δ y β , el elemento A_{ij} es de la forma:

$$A_{ij} = \frac{\partial f(x_i + \delta_i; \beta)}{\partial \beta_j}.$$

- V es una matriz diagonal de $n \times n$ de la forma:

$$V = \text{diag} \left\{ \frac{\partial f(x_i + \delta_i; \beta)}{\partial \delta_i}; \text{ para } i = 1, \dots, n \right\}.$$

Observamos que la matriz J_G tiene la misma estructura del jacobiano (iv.2.5), por lo que no es necesario repetir todo el desarrollo de la sección anterior para reducir el número de parámetros a calcular, sólo escribiremos brevemente el algoritmo para resolver el problema (iv.3.1).

ALGORITMO.

1. Suponemos conocida $\bar{\beta}_k$ y $\bar{\delta}_k$.

2. Obtenemos la solución $\bar{\beta}_s^*$ del problema:

$$\min_{\bar{\beta}_s} \frac{1}{2} \left\| \begin{bmatrix} MA \\ \lambda \frac{1}{2} S \end{bmatrix} \bar{\beta}_s + \begin{bmatrix} M[-\bar{r}_k + VE^{-1}I_n \bar{\delta}_k] \\ 0 \end{bmatrix} \right\|^2$$

3. Asignamos $\bar{\beta}_{k+1} = \bar{\beta}_k + \bar{\beta}_s^*$.

4. Calculamos $\bar{\delta}_s^*$ de la expresión:

$$\bar{\delta}_s = -E[V^t M^2 (\bar{r}_k + A \bar{\beta}_s - VE^{-1} I \bar{\delta}_k) + I_n \bar{\delta}_k]$$

5. Asignamos $\bar{\delta}_{k+1} = \bar{\delta}_k + \bar{\delta}_s^*$.

Capítulo V

ODRPACK

En las primeras secciones de este capítulo haremos una breve referencia a una colección portátil de subrutinas escritas en ANSI FORTRAN 77, para el ajuste ortogonal de un modelo a un conjunto de datos, ODRPACK. Inicialmente, su diseño fue para el caso en que todas las variables están sujetas a errores significativos, en el cual se implementó un algoritmo eficiente para resolver el problema de ajuste ortogonal ponderado [3] y [5], en el cual se minimiza la suma de los cuadrados de las distancias ortogonales ponderadas de cada dato a la curva descrita por el modelo, pero también puede ser utilizada para resolver el problema ordinario de mínimos cuadrados en donde todos los errores son atribuidos a las observaciones de la variable independiente.

Las características más importantes de ODRPACK son:

- Es fácil de usar, provee dos niveles de control para los cálculos, puede opcionalmente imprimir reportes, no tiene restricciones de espacio adicionales a los de la computadora en la cual se trabaje.
- Resuelve tanto problemas implícitos como explícitos e incluso para datos multirespuesta, es decir, datos donde cada observación es multidimensional.
- Las derivadas necesarias (jacobianos) se aproximan numéricamente, si éstas no son proporcionadas por el usuario y en el caso contrario provee un procedimiento para verificarlas.
- Se pueden realizar análisis ponderados o no.
- Algunos parámetros se pueden considerar como constantes, fijando sus

valores, de esta forma el usuario puede examinar los resultados obtenidos, al estimar los parámetros restantes del modelo general sin reescribir la subrutina del modelo.

- Tanto la matriz de covarianza como las estimaciones de los parámetros del modelo pueden ser proporcionados opcionalmente.
- El algoritmo de escalamiento de ODRPACK compensa automáticamente problemas pobremente escalados.
- El algoritmo de región de confianza de Levenberg-Marquardt implementado en ODRPACK, requiere un esfuerzo computacional del mismo orden, que se requiere para el problema ordinario de mínimos cuadrados.
- El código es portátil y fácil de usar con otras bibliotecas de subrutinas en FORTRAN.

Aunque ODRPACK tiene estas características sólo nos referiremos brevemente a aquéllas que consideramos importantes para este trabajo. De estas características cabe mencionar, que sólo trataremos el ajuste ortogonal implícito y no lo haremos para el caso de multirespuesta por ser la notación pesada a comparación con la que hemos estado trabajando, en lo que se refiere a la matriz de covarianza no la utilizaremos ya que el objetivo de esta tesis es hacia el aspecto numérico y no hacia el estadístico.

En la última sección, utilizaremos este paquete para dar algunos ejemplos de ajuste ortogonal.

V.1 Especificando la tarea.

El usuario tiene la opción de especificar las características de cada problema para ser resuelto, como las que a continuación se describen:

- Si el ajuste es ortogonal implícito, explícito o es un ajuste ordinario de mínimos cuadrados.
- Si los jacobianos necesarios serán aproximados por ODRPACK usando diferencias finitas centrales o diferencias finitas hacia adelante, o si el

usuario proporciona el código para calcularlos y si éste es el caso, también tiene la posibilidad de indicarle que deben ser verificados.

- Si la matriz de covarianza V_β y la desviación estándar σ_β deben ser calculadas a partir de valor estimado de $\bar{\beta}$, y si los jacobianos deben ser recalculados en la solución.
- Si los errores $\bar{\delta}_i$ para $i = 1, \dots, n$ han sido inicializados por el usuario.
- Si el ajuste es una reinicialización, es decir, si el ajuste usará información guardada, para continuar de un punto previamente encontrado.
- Si los datos son multirespuesta o no.
- Si algunos de los parámetros β_i para $i = 1, \dots, p$ y δ_i para $i = 1, \dots, n$ debe ser considerados con valores fijos, permitiendo al usuario examinar los resultados obtenidos para el modelo general determinado por el resto de los parámetros, sin la necesidad de reescribir rutina del modelo.
- Si debe realizarse un análisis ponderado o no.
- Si los parámetros $\bar{\beta}$ y $\bar{\delta}_i$ para $i = 1, \dots, n$ son escalados para compensar los casos donde las magnitudes de sus valores existe una gran variación.

V.2 Ajuste ortogonal a modelos implícitos.

En el problema de ajuste ortogonal unidimensional a un modelo implícito, no hay distinción de una variable dependiente, por lo que los datos deben satisfacer:

$$m(x_i + \delta_i; \bar{\beta}) = 0 \quad \text{para } i = 1, \dots, n$$

para $x_i \in \mathbb{R}$, cuyo planteamiento queda determinado como:

$$\min_{\bar{\beta}, \bar{\delta}} \sum_{i=1}^n \omega_{\delta_i} \delta_i^2 \quad (v.1.1)$$

$$\text{sujeto a } m(x_i + \delta_i; \bar{\beta}) = 0 \quad \text{para } i = 1, \dots, n$$

donde $\omega_{\delta_i} \in \mathbb{R}$ para $i = 1, \dots, n$ denota los pesos, los cuales son números positivos. Las restricciones en (v.1.1) no son consideradas lineales en δ_i , por lo que

no pueden ser eliminadas, para transformar el problema a uno sin restricciones, como en el caso explícito, por lo tanto el problema de ajuste ortogonal a modelos implícitos se debe resolver de manera diferente.

ODRPACK encuentra la solución a este tipo de problema usando el método clásico de la función cuadrática de penalización y una sucesión de problemas de minimización sin restricciones son solucionados para una sucesión de valores del parámetro de penalización que tienden a infinito.

V.3 Valores iniciales.

El usuario debe proporcionar los valores iniciales de los parámetros $\bar{\beta}$ y δ_i para $i = 1, \dots, n$, los usuarios familiarizados con el problema ordinario de mínimos cuadrados no-lineales saben de la importancia de obtener un buen valor inicial para $\bar{\beta}$, en este caso tiene la misma importancia, pues buenas aproximaciones iniciales pueden causar un decrecimiento significativo en el número de iteraciones requeridas para encontrar una solución, y tal vez con una mala aproximación inicial podría no converger a la solución.

Cuando el usuario no tiene conocimiento de cuales serían los puntos iniciales convenientes, se le recomienda realizar un análisis preliminar para obtenerlos.

Los usuarios que no provean información sobre la escala se les recomienda dar la aproximación inicial diferente de cero para alguna β_i , pues un valor cero podría causar un escalamiento incorrecto.

En el ajuste ortogonal también es importante obtener una buena aproximación a los errores δ_i para $i = 1, \dots, n$, el valor por omisión de ODRPACK es inicializarlos como cero, lo cual no causa problemas como para el valor inicial de $\bar{\beta}$, pero observamos que si inicializamos de esta forma equivale a resolver un problema ordinario de mínimos cuadrados en la primera iteración, es decir, es como si pensamos que sólo la variable dependiente contiene errores, esto puede ser conveniente en algunos casos pero en otros no. Un camino, para saber si es conveniente o no esta inicialización, es graficar los datos y la curva descrita por el modelo y $\bar{\beta}$ inicial, en su caso determinar los mejores valores iniciales para estos parámetros.

Por ejemplo, si la gráfica muestra que la distancia vertical del dato (x_i, y_i) a

la curva está más distante que la correspondiente distancia ortogonal, entonces probablemente $\bar{\delta}_i$ debe no ser inicializado con cero. Esto podría ocurrir cerca de una asíntota, de un máximo o de un mínimo local, en cada caso, es apropiado inicializar $\bar{\delta}_i$ como la distancia horizontal del dato a la curva.

V.4 Pesos.

Los pesos pueden ser utilizados para eliminar observaciones del análisis, para compensar varianzas diferentes o correlaciones en las variables y y x , o simplemente para modificar el efecto de varios errores ϵ y δ en el ajuste.

Pero dado que nosotros no tenemos experiencia en el ajuste ponderado, recomendamos a las personas que estén interesadas sobre este tema consulte *User's reference guide for ODRPACK - software for weighted orthogonal distance regression*, [6].

V.5 Valores por omisión.

La disponibilidad de valores por omisión es una característica versátil que nos proporciona este paquete de subrutinas, pues nos proporciona la posibilidad de no involucrarnos con la mayoría de sus argumentos, a menos que el problema a resolver requiera de otros valores. Para indicar que se van a utilizar estos valores por omisión se debe asignar a los argumentos correspondientes un valor negativo. En el caso de arreglos se asigna al primer elemento un valor negativo, por lo que sólo se utilizará este escalar en lugar del arreglo lo que permite minimizar espacio y eliminar la necesidad de declarar estos arreglos.

Los valores por omisión de ODRPACK han sido considerados razonables para un gran número de problemas.

V.6 Ejemplos.

En esta sección presentamos una serie de ejemplos sobre ajuste ortogonal, de los cuales obtuvimos su solución con el paquete ODRPACK [6].

Realizamos una versión especial en FORTRAN con este paquete, para ajustar datos a un polinomio de la forma:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p$$

de esta versión obtuvimos la solución de los dos primeros ejemplos.

1. Ajuste ortogonal lineal ponderado.

Aunque en el capítulo I hemos ilustrado el problema de ajuste ortogonal lineal con un ejemplo, aquí quisimos mostrar otro ejemplo de este tipo, pero ahora realizando un ajuste ponderado, para esto, utilizamos los datos del artículo de York [29], los cuales los escribimos a continuación con sus respectivos pesos:

Tabla V.1 Datos para el ajuste ortogonal ponderado.

i	x	y	px	py
1	0.0	5.9	1 000.0	1.0
2	0.9	5.4	1 000.0	1.8
3	1.8	4.4	500.0	4.0
4	2.6	4.6	800.0	8.0
5	3.3	3.5	200.0	20.0
6	4.4	3.7	80.0	20.0
7	5.2	2.8	60.0	70.0
8	6.1	2.8	20.0	70.0
9	6.5	2.4	1.8	100.0
10	7.4	1.5	1.0	500.0

El modelo que le ajustamos a estos datos es de la forma :

$$y = \beta_0 + \beta_1 x$$

iniciaremos las iteraciones con el punto $\bar{\beta} = [2.5, 1.5]^t$, por lo que la gráfica inicial es:

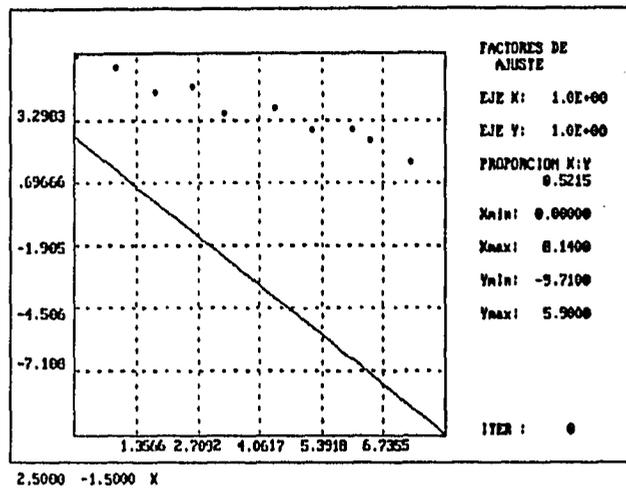


Figura V.1: Gráfica de los datos y del modelo en el punto inicial.

ODRPACK obtiene la solución $\bar{\beta} = [5.4799099, -0.480533241]^t$, en seis iteraciones, por lo que la gráfica del ajuste es:

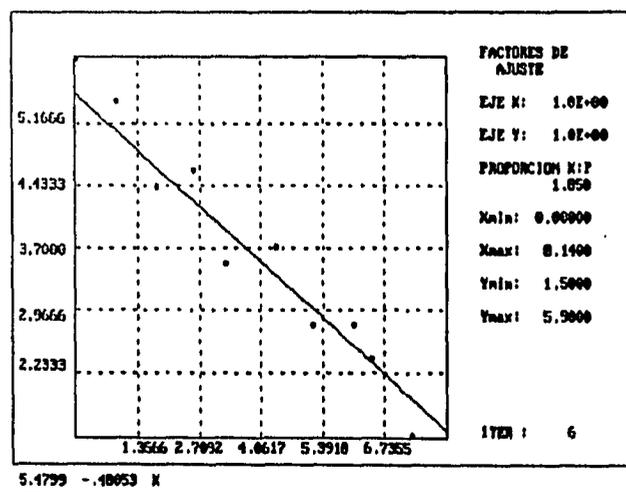


Figura V.2: Gráfica del ajuste ortogonal lineal ponderado.

2. Ajuste ortogonal a un polinomio cúbico.

La importancia de la familia de los polinomios, nos motivó para presentar, como segundo ejemplo el ajuste a un polinomio de tercer grado de la forma:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

para construir los datos, primero particionamos el intervalo $[-8, 12]$ en 15 partes iguales de la siguiente forma:

$$x_i = -8 + \left(\frac{20}{15}\right) * i \quad \text{para } i = 0 \dots 15$$

después evaluamos las ordenadas como:

$$y_i = 53 - 49x_i - 2x_i^2 + x_i^3 \quad \text{para } i = 0 \dots 15$$

tanto las ordenadas como las abscisas las perturbamos con un número aleatorio de una distribución uniforme de rango $[-0.5, 0.5]$, así obtuvimos los siguientes datos:

Tabla V.2: Datos obtenidos por un polinomio cúbico.

i	x	y	i	x	y
1	-7.8187	-223.7248	9	3.1221	-51.9440
2	-6.7809	15.9525	10	4.0561	-67.2103
3	-5.4456	68.7835	11	4.9815	-139.5274
4	-4.0003	112.0990	12	7.1500	-98.1697
5	-3.0191	127.9403	13	7.9088	26.7540
6	-1.2461	62.7074	14	8.9752	273.1805
7	0.3456	44.4293	15	10.7316	531.6133
8	1.4234	-60.8309	16	11.7521	870.0335

El ajuste que realizamos no fue ponderado y lo iniciamos con el punto:

$$\bar{\beta} = [65.9, -43.6, -2.7, 1.2]^t$$

entonces la iteración inicial gráficamente la vemos así:

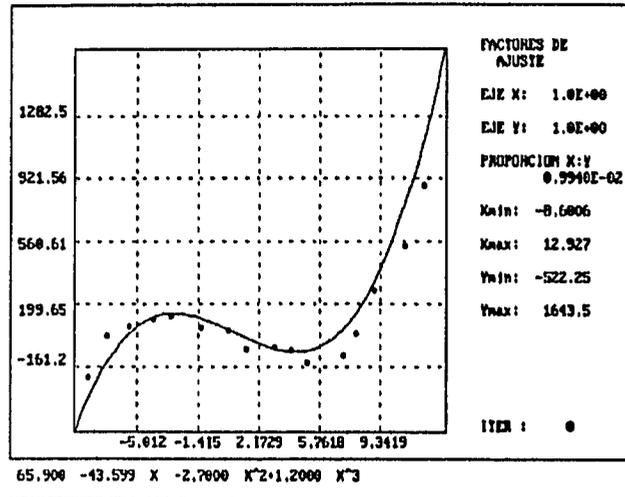


Figura V.3: Gráfica de los datos y del modelo en el punto inicial.

ODRPACK obtiene en treinta iteraciones, la siguiente solución:

$$\bar{\beta} = [38.5613368, -47.5090224, -2.74540397, 1.02546682]^t$$

por lo que la gráfica final del ajuste es:

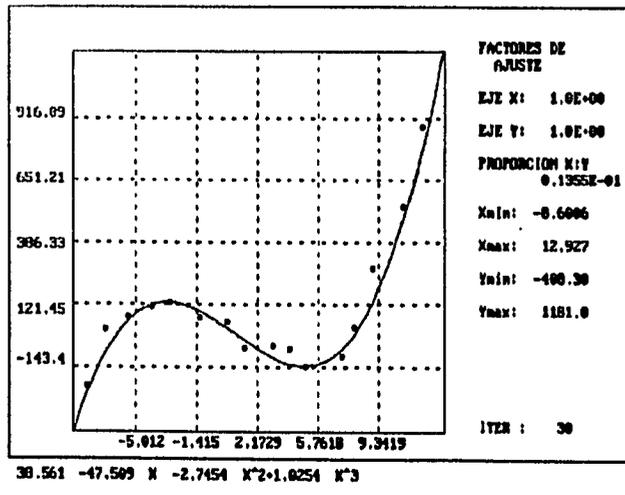


Figura V.4: Gráfica del ajuste ortogonal al polinomio $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$.

3. Ajuste ortogonal implícito.

Para ejemplificar el ajuste ortogonal a un modelo implícito, consideramos el segundo ejemplo de la guía del usuario de ODRPACK [6], en el cual se tomaron las siguientes 20 observaciones de una radiografía de una prótesis de cadera:

Tabla V.3: Datos obtenidos de una radiografía de prótesis de cadera

i	x	y	i	x	y
1	0.50	-0.12	11	1.34	-5.97
2	1.20	-0.60	12	0.90	-6.32
3	1.60	-1.00	13	-0.28	-6.44
4	1.86	-1.40	14	-0.78	-6.44
5	2.12	-2.54	15	-1.36	-6.41
6	2.36	-3.36	16	-1.90	-6.25
7	2.44	-4.00	17	-2.50	-5.88
8	2.36	-4.75	18	-2.88	-5.50
9	2.06	-5.25	19	-3.18	-5.24
10	1.74	-5.64	20	-3.44	-4.86

dada la naturaleza del problema, se ajustó el modelo cuadrático de la forma:

$$\beta_3(y - \beta_1)^2 + 2\beta_4(y - \beta_1)(x - \beta_2) + \beta_5(x - \beta_2)^2 - 1 = 0 \quad (v.2.1)$$

donde y representa la distancia vertical y x la distancia horizontal al origen.

Decidimos utilizar el siguiente punto inicial:

$$\bar{\beta} = [-0.1, 0.1, 0.1, 0.1, 0.1]^t$$

por ser un punto lejano de la solución, esto lo podemos ver si evaluamos este punto en (v.2.1) y denotamos como:

$$\hat{x} = x - \beta_2, \quad \hat{y} = y - \beta_1$$

y además realizamos algunas operaciones, llegamos a la siguiente ecuación:

$$\hat{y} = \hat{x} \pm 10$$

por lo que nuestra primer función que ajusta a nuestros datos es un par de rectas paralelas, que gráficamente observamos así:

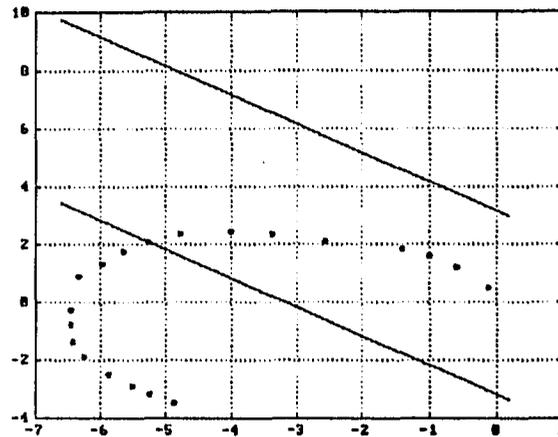


Figura V.5: Gráfica inicial del ajuste es un par de rectas.

como observamos están muy lejos de ajustar satisfactoriamente los datos. La solución que obtenemos en veintiocho iteraciones es:

$$\bar{\beta} = [-0.9993808, -2.9310485, 0.0875730, 0.0162299, 0.0797538]^t$$

Al evaluar en en este punto la gráfica resulta ser:

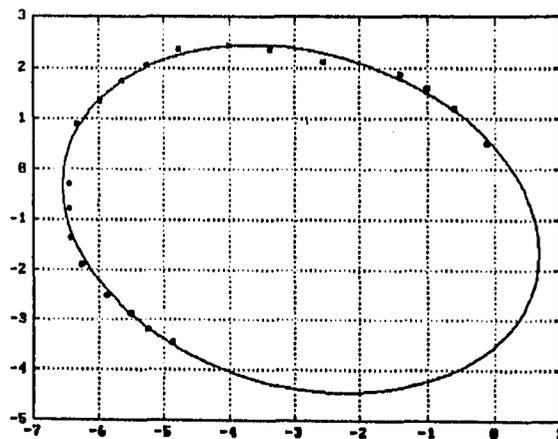


Figura V.6: La gráfica final del ajuste es una elipse.

4. Ajuste ortogonal a una función de Kowalik y Osborne.

Por último presentaremos el ajuste ortogonal implícito de una función de Kowalik y Osborne definida como:

$$\frac{\beta_1 x(x + \beta_2)}{x(x + \beta_3) + \beta_4} - y = 0$$

al siguiente conjunto de once datos:

Tabla V.4: Datos para ajustar la función de Kowalik y Osborne.

i	x	y
1	4.0000	0.1957
2	2.0000	0.1947
3	1.0000	0.1735
4	0.5000	0.1600
5	0.2500	0.0844
6	0.1670	0.0627
7	0.1250	0.0456
8	0.1000	0.0342
9	0.0833	0.0323
10	0.0714	0.0235
11	0.0625	0.0246

En este caso tomamos como punto inicial:

$$\bar{\beta} = [-0.25, 0.39, 0.415, 0.39]^t$$

por lo que gráfica correspondiente a estos parámetros iniciales es:

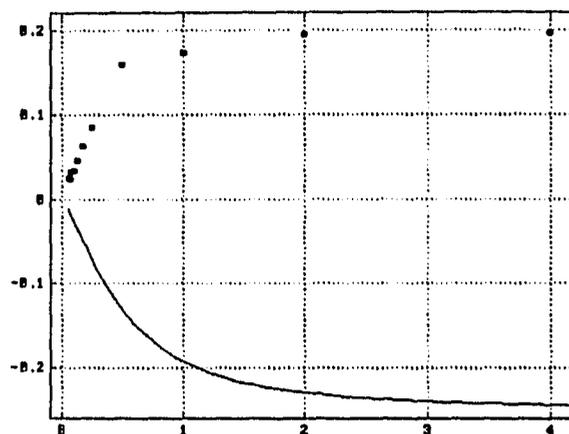


Figura V.7: Gráfica de los datos y del modelo en el punto inicial.

obtuvimos como aproximación final del óptimo, al vector:

$$\bar{\beta} = [0.193132119, 0.179413870, 0.118492054, 0.130645862]^t$$

cuya gráfica es:

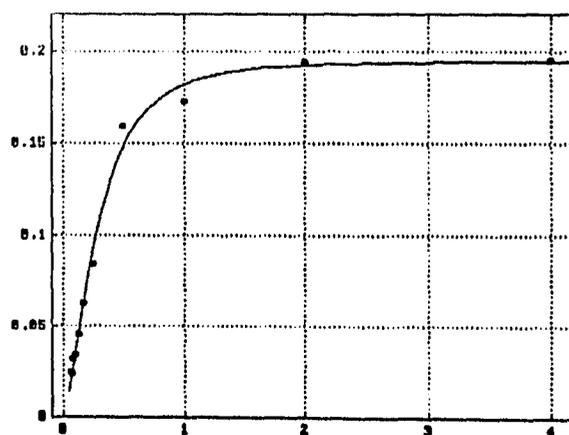


Figura V.8: Gráfica final del ajuste.

Apéndice A

Producto de Kronecker.

El producto de Kronecker es una operación tensorial la cual es una generalización de la multiplicación de una matriz por un escalar. Si tenemos dos matrices $\mathbf{A}_{m \times n}$ y $\mathbf{B}_{p \times q}$, el producto de Kronecker de \mathbf{A} por \mathbf{B} , $\mathbf{A} \otimes \mathbf{B}$, es una nueva matriz con mp renglones y nq columnas, las cuales se obtienen reemplazando cada elemento b_{ij} de la matriz \mathbf{B} por el siguiente arreglo:

$$b_{ij}\mathbf{A} = \begin{bmatrix} b_{ij}a_{11} & b_{ij}a_{12} & \cdots & b_{ij}a_{1n} \\ b_{ij}a_{21} & b_{ij}a_{22} & \cdots & b_{ij}a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{ij}a_{m1} & b_{ij}a_{m2} & \cdots & b_{ij}a_{mn} \end{bmatrix}$$

Este producto generalmente no es conmutativo:

$$\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$$

Por ejemplo, si la matriz \mathbf{B} es :

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

el producto de Kronecker de la matriz \mathbf{A} por la matriz \mathbf{B} es:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} b_{11}\mathbf{A} & b_{12}\mathbf{A} \\ b_{21}\mathbf{A} & b_{22}\mathbf{A} \end{bmatrix}$$

En particular el producto de Kronecker de la matriz identidad de orden n I_n por la de orden m I_m , es igual a la identidad de dimensión nm :

$$I_n \otimes I_m = I_{nm}$$

Similarmente, el producto de Kronecker de una matriz diagonal por otra diagonal es una matriz diagonal.

El producto de Kronecker puede ser combinado con productos matriciales convencionales y entonces tendríamos las siguientes propiedades, suponemos que las dimensiones son compatibles:

$$1. \quad (ABC) \otimes I = (A \otimes I)(B \otimes I)(C \otimes I).$$

$$2. \quad (A \otimes B \otimes C)(D \otimes E \otimes F) = (AD) \otimes (BE) \otimes (CF).$$

$$3. \quad (A \otimes B \otimes C)^{-1} = A^{-1} \otimes B^{-1} \otimes C^{-1}.$$

$$4. \quad (A \otimes B \otimes C)^t = A^t \otimes B^t \otimes C^t.$$

Apéndice B

Propiedades de las funciones cuadráticas.

Las funciones cuadráticas juegan un papel muy importante en el desarrollo de algoritmos para resolver problemas de optimización. Por ejemplo sabemos que en una vecindad de un mínimo local de una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$, al aplicar el método de Newton necesitamos que el paso w_k sea el que minimice el modelo local cuadrático:

$$\Psi_k(w) = \nabla f(x_k)^t w + \frac{1}{2} w^t \nabla^2 f(x_k) w \quad (b.1)$$

de la reducción esperada en f , es decir:

$$f(x_k + w) - f(x_k) \approx \nabla f(x_k)^t w + \frac{1}{2} w^t \nabla^2 f(x_k) w$$

donde $w = x - x_k$, por lo que es importante entender las propiedades de las funciones cuadráticas y proveer algoritmos numéricos estables para minimizarlas, el siguiente resultado describe completamente la minimización sin restricciones de las funciones cuadráticas.

Lema B.1 Sea $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$ una función cuadrática de la forma:

$$\Psi(w) = g^t w + \frac{1}{2} w^t B w \quad (b.3)$$

donde $g \in \mathbb{R}^n$, B es una matriz simétrica de orden n .

1. La función Ψ tiene un mínimo si y sólo si B es positiva semidefinida y g pertenece al rango de B .
2. La función Ψ tiene un único mínimo si y sólo si B es positiva definida.

3. Si B es positiva semidefinida entonces cada solución de la ecuación:

$$Bp = -g$$

es un mínimo global.

Prueba

1. \Leftarrow) Primero suponemos que B es positiva semidefinida y g pertenece al rango de B ; entonces existe un vector p que satisface el siguiente sistema:

$$Bp = -g \tag{b.4}$$

si evaluamos en $(p + w)$ la función cuadrática (b.3) y además utilizamos el hecho que coincide con su desarrollo de Taylor, obtenemos la siguiente igualdad:

$$\Psi(p + w) = \Psi(p) + \frac{1}{2}w^t Bw + (Bp + g)^t w$$

al utilizar (b.4) llegamos a:

$$\Psi(p + w) = \Psi(p) + \frac{1}{2}w^t Bw$$

como sabemos que $B \geq 0$, es inmediata la siguiente desigualdad:

$$\Psi(p + w) \geq \Psi(p) \quad \forall w \in \mathbb{R}^n \tag{b.5}$$

por lo tanto p es un mínimo de Ψ .

1. \Rightarrow) Si ahora suponemos que p es un mínimo de Ψ entonces sabemos que:

$$\nabla \Psi(p) = Bp + g = 0$$

y que el Hessiano de Ψ :

$$\nabla^2 \Psi(p) = B$$

es positiva semidefinida.

Para establecer 1. y 2. nota que (b.5) se cumple si $Bp = -g$ y B es positiva semidefinida y que la desigualdad estricta sucede para $w \neq 0$ si y sólo si B es positiva definida. ♣

Para obtener el mínimo de Ψ , como primera opción calculamos la factorización de Cholesky de B , esta factorización existe si y sólo si B es positiva semidefinida y en este caso se construye una matriz triangular superior R tal que:

$$B = R^t R$$

si un elemento negativo de la diagonal es encontrado durante el proceso de la factorización, entonces B no es positiva semidefinida y por el Lema B.1 muestra que la cuadrática Ψ no tiene mínimo, en el caso contrario si los elementos de la diagonal de R son mayores que cero entonces el mínimo es calculado resolviendo el sistema:

$$Bp = -g$$

que equivale a resolver los sistemas triangulares:

$$R^t v = -g$$

$$R^t p = v$$

Pero si los elementos de la diagonal de R son positivos y por lo menos un elemento es igual a cero, entonces B es positiva semidefinida y singular, en esta situación podría ser posible calcular una solución p , pero desde el punto de vista numérico este cálculo es inestable, porque bajo pequeñas perturbaciones pueden transformar a B a una matriz positiva definida o a una matriz indefinida.

Sabemos que en una vecindad de un mínimo local de f , el Hessiano es positivo definido y por el Lema B.1 muestra que el modelo local cuadrático (b.1) tiene un único mínimo, por lo que este caso el mínimo del modelo local cuadrático es un razonable paso para el algoritmo de minimización. Sin embargo, lejos de un mínimo local el Hessiano $\nabla^2 f(x)$ podría tener un eigenvalor negativo y por el Lema B.1 nos dice que este modelo local cuadrático no tiene mínimo, en efecto, este modelo no está acotado inferiormente. Hay muchas opciones para superar esta dificultad, una posibilidad es modificar el modelo cuadrático, sumando una matriz positiva semidefinida $E(x)$ tal que:

$$S = \nabla^2 f(x) + E(x)$$

sea positivo definido. Cuando B es reemplazada por la matriz S , el cálculo del paso se puede realizar como ya se describió anteriormente.

Otra posibilidad es restringir la región en la cual suponemos que el modelo local cuadrático es apropiado. Localmente el modelo provee una excelente aproximación a la reducción esperada de f , por lo que es razonable restringir Ψ sobre la región:

$$\{w : \|w\| \leq \Delta \text{ y } w \in \mathbb{R}^n\}$$

para alguna $\Delta > 0$ y calcular el paso, como el mínimo de Ψ en esta región, esta idea da origen a los métodos de Región de Confianza que trataremos en el apéndice C. El siguiente resultado caracteriza las soluciones globales del problema de minimización de la función cuadrática sobre esta región restringida.

Lema B.2 Sea $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$ una función cuadrática de la forma

$$\Psi(w) = g^t w + \frac{1}{2} w^t B w$$

y dada $\Delta > 0$. Un punto $p \in \mathbb{R}^n$ es la solución del problema:

$$\min_{\|w\| \leq \Delta} \Psi(w) \tag{b.7}$$

si y sólo si $\|p\| \leq \Delta$ y existe $\lambda \geq 0$ tal que

$$(B + \lambda I)p = -g; \quad \lambda(\Delta - \|p\|) = 0 \tag{b.8}$$

con $(B + \lambda I)$ positiva semidefinida.

Prueba.

Supongamos que λ y p satisfacen (b.8) con $\|p\| \leq \Delta$ y $(B + \lambda I)$ positiva semidefinida, entonces por el Lema B.1 implica que p minimiza la función cuadrática:

$$\hat{\Psi}(w) = g^t w + \frac{1}{2} w^t (B + \lambda I) w$$

entonces $\hat{\Psi}(w) \geq \hat{\Psi}(p)$ lo cual implica las siguientes desigualdades :

$$\begin{aligned} g^t w + \frac{1}{2} w^t (B + \lambda I) w &\geq g^t p + \frac{1}{2} p^t (B + \lambda I) p \\ \Psi(w) + \frac{\lambda}{2} w^t w &\geq \Psi(p) + \frac{\lambda}{2} p^t p \\ \Psi(w) &\geq \Psi(p) + \frac{\lambda}{2} (p^t p - w^t w) \end{aligned} \tag{b.9}$$

hasta aquí dejamos esta desigualdad y ahora trabajaremos con la siguiente igualdad, la cual tenemos como hipótesis:

$$\begin{aligned}\lambda(\Delta - \|p\|) &= 0 \\ \lambda(\Delta - \|p\|)(\Delta + \|p\|) &= 0 \\ \lambda(\Delta^2 - \|p\|^2) &= 0 \\ \lambda\Delta^2 &= \lambda\|p\|^2\end{aligned}$$

esta igualdad la utilizamos en la desigualdad (b.9) y así obtener:

$$\frac{\Psi(w) \geq \Psi(p) + \lambda}{2(\Delta^2 - \|w\|^2)}$$

por lo que si $\|w\|^2 \leq \Delta^2$ entonces podemos finalmente podemos concluir:

$$\Psi(w) \geq \Psi(p)$$

por lo tanto p es la solución del problema (b.7).

Si ahora suponemos que p es solución de (b.7), esto implica que p cumple:

$$\|p\| \leq \Delta$$

por lo que analizaremos dos casos.

Primer caso :

Si suponemos que $\|p\| < \Delta$ entonces p es un mínimo del problema sin restricciones:

$$\min_w \Psi(w)$$

por lo que p satisface el siguiente sistema de ecuaciones :

$$Bp = -g$$

con B positiva definida, esta igualdad la podemos escribir como:

$$(B + 0I)p = -g$$

es decir, considerar $\lambda = 0$ y como $\|p\| < \Delta$ es inmediata la igualdad:

$$\lambda(\Delta - \|p\|) = 0$$

Segundo caso:

Si $\|p\| = \Delta$ entonces p debe también resolver el problema:

$$\min_{\|w\|=\Delta} \Psi(w)$$

un camino posible, para resolver este problema es por medio del método de Lagrange el cual garantiza la existencia de λ , tal que p y λ es un punto crítico de la siguiente función:

$$\frac{L(w, \lambda) = \Psi(w) + \lambda}{2(w^t w - \Delta^2)}$$

ahora si calculamos el gradiente de la función L , obtenemos:

$$\nabla L = \begin{bmatrix} g + (B + \lambda I)w \\ \frac{1}{2}(w^t w - \Delta^2) \end{bmatrix}$$

como sabemos que p es solución de (b.7) implica que el gradiente evaluado en este punto es igual a cero, por lo que tenemos el siguiente sistema de ecuaciones:

$$\begin{aligned} g + (B + \lambda I)p &= 0 \\ \frac{1}{2}(p^t p - \Delta^2) &= 0 \end{aligned}$$

por lo tanto para esta λ y p se cumplen las siguientes igualdades:

$$\begin{aligned} (B + \lambda I)p &= -g \\ \lambda(\Delta - \|p\|) &= 0 \end{aligned}$$

que queríamos, más aun como p es una solución de (b.7) tenemos que (b.9) es válido para esta λ y p , pues $\|w\| = \|p\|$.

Para demostrar que $(B + \lambda I)$ es positiva definida, lo haremos a partir de la siguiente desigualdad:

$$\Psi(w) \geq \Psi(p) \quad \text{si } \|w\| \leq \Delta \quad (\text{b.10})$$

que sabemos que se cumple por hipótesis, si sustituimos el valor del gradiente g de (b.8) en esta desigualdad, y realizando operaciones obtenemos que:

$$\frac{1}{2}(w - p)^t (B + \lambda I)(w - p) \geq 0$$

de esta desigualdad concluimos que $(B + \lambda I)$ es positiva semidefinida.

Para mostrar que $\lambda \geq 0$, utilicemos el hecho que la desigualdad (b.9):

$$\frac{\Psi(w) \geq \Psi(p) + \lambda}{2(p^t p - w^t w)}$$

es válida para toda $w \in \mathbb{R}^n$, ahora si suponemos que λ es no positiva entonces de esta desigualdad obtenemos:

$$\Psi(w) \geq \Psi(p) \quad \text{para} \quad \|w\| \geq \|p\|$$

como p es solución de (b.7), concluimos que p es un mínimo sin restricciones de Ψ , lo que implica que $\lambda = 0$; por lo tanto $\lambda \geq 0$ como se quería.

♣

Una complicación inmediata, que se presenta al tratar de resolver (b.7) se debe a la restricción no-lineal, por lo que en general no podemos utilizar un método directo.

En efecto, si $g = 0$ y B tiene un eigenvalor negativo entonces una solución p a (b.7) debe ser un eigenvector de norma Δ correspondiente al eigenvalor más pequeño de B , por consiguiente un método general para resolver (b.7) debe resolver un problema de eigenvalor simétrico en este caso especial, puesto que nos interesa soluciones globales de (b.7), podría parecer que un algoritmo que resuelve (b.7) es limitado y bastante costoso; sin embargo, ahora mostraremos que existe un algoritmo que produce una solución cercana a la optima de (b.7) en todos los casos y en pocas iteraciones.

No es difícil probar que (b.7) no tiene soluciones con $\|p\| = \Delta$ si y sólo si B es positiva definida y $\|B^{-1}g\| < \Delta$; si (b.7) tiene una solución en la frontera de $\{w : \|w\| \leq \Delta\}$, entonces el Lema B.2 muestra que es razonable esperar que la ecuación:

$$\|p_\alpha\| = \Delta \tag{b.11}$$

donde

$$p_\alpha = -(B + \alpha I)^{-1}g$$

tiene una solución $\alpha^* = \lambda \geq 0$ en $(-\lambda_1, \infty)$, donde $\lambda_1 \leq 0$ es el eigenvalor más pequeño de B , como (b.11) es un problema de encontrar el cero de una función escalar en α que puede ser resuelto, por ejemplo por el método de Newton, pero el inconveniente que surge es que en cada evaluación de p_α requiere la solución

de un sistema de ecuaciones lineales, por lo que es importante observar la estructura de esta función para realizarlo en pocas evaluaciones.

Para resolver (b.11), Reinsch [21], [22] y Hebden [17] independientemente observaron que podían tomar ventaja del hecho que la función $\|p_\alpha\|^2$ es una función racional en α con polos de segundo orden en un subconjunto de los eigenvalores negativos de la matriz simétrica B . Para ver esto consideramos la descomposición de valores singulares:

$$B = Q\Lambda Q^t$$

con $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ y $QQ^t = I$ y observa que:

$$\|p_\alpha\|^2 = \|Q(\Lambda + \alpha I)^{-1}Q^t g\|^2 = \sum_1^n \frac{\gamma_j^2}{(\lambda_j + \alpha)^2} \quad (b.12)$$

donde γ_i es el i -ésimo componente de $Q^t g$. Una vez conocida la forma de la funcional (b.12) muestra que el método de Newton podría no ser muy eficiente, si es aplicado a la función:

$$\varphi_1(\alpha) = \|p_\alpha\| - \Delta$$

una razón para esto es que φ_1 tiene un polo en $-\lambda_1$ y por esto el método de Newton tiende a realizarse deficientemente cuando la solución de (b.11) esta cerca de $-\lambda_1$. Reinsch y Hebden sugirieron que es más eficiente aplicar el método de Newton a la función:

$$\varphi_2(\alpha) = \frac{1}{\Delta} - \frac{1}{\|p_\alpha\|}$$

Esta función no tiene polos, y es casi lineal cerca de la solución de (b.11). La iteración de Newton aplicada para encontrar un cero de φ_2 toma la siguiente forma:

ALGORITMO b.13

1. Dadas λ_0 y $\Delta > 0$
2. Para $k = 0, 1, \dots$ hasta converger

(a) Obtener la factorización

$$B + \lambda_k I = R_k^t R_k.$$

(b) Resolver el sistema

$$R_k^t R_k p_k = -g.$$

(c) Resolver el sistema

$$R_k^t q_k = p_k.$$

(d) Calcular λ_{k+1} de

$$\lambda_{k+1} = \lambda_k + \left[\frac{\|p_k\|}{\|q_k\|} \right]^2 \left[\frac{\|p_k\| - \Delta}{\Delta} \right].$$

Si ciertas precauciones son tomadas, entonces esta iteración básica puede ser usada para resolver (b.7) en la mayoría de los casos. Sin embargo cuando B es indefinida, existen casos en los cuales la ecuación (b.11) no tiene soluciones en $(-\lambda_1, \infty)$ y entonces el algoritmo (b.13) falla. Esto sucede por ejemplo cuando $g = 0$ y B es indefinida. Podría suceder también cuando $g \neq 0$, como vemos en el siguiente ejemplo. Si

$$B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad g = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

entonces $\lambda_1 = -1$ y si $\alpha > 1$ entonces $\|p_\alpha\|^2 < \frac{1}{2}$, en este ejemplo g es ortogonal al eigenspacio de B correspondiente al eigenvalor más pequeño, definido como

$$S_1 = \{z : Bz = \lambda_1 z, z \neq 0\}$$

y en general esto sucede siempre que (b.11) no tenga soluciones en $(-\lambda_1, \infty)$. Para ver esto es suficiente notar que si g es no ortogonal a S_1 , entonces $\gamma_1 \neq 0$ en (b.12) y entonces

$$\lim_{\alpha \rightarrow -\lambda_1} \|p_\alpha\| = \infty; \quad \lim_{\alpha \rightarrow +\infty} \|p_\alpha\| = 0$$

por lo que pequeñas perturbaciones de g conducen a un γ_1 no cero. Sin embargo en muchos casos g es casi ortogonal a S_1 , y en estos casos un algoritmo basada completamente en el método de Newton podría requerir un número grande de iteraciones, lo cual no es aceptable ya que en cada iteración se requiere la factorización de una matriz.

Apéndice C

Métodos de región de confianza.

Una modificación importante al método de Newton, es considerar el modelo cuadrático solamente en una región de confianza restringida, esta técnica la mencionamos brevemente en el apéndice B como una motivación para el Lema B.2; su uso para la globalización del método de Newton ha resultado en algoritmos confiables con buenas propiedades de convergencia. En esta sección introducimos las ideas principales de esta aproximación y establecemos algunas de las propiedades básicas de convergencia.

Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^2 . En el método de Newton con estrategia de región de confianza, cada iteración x_k tiene una cota Δ_k tal que:

$$f(x_k + w) \approx f(x_k) + \Psi_k(w)$$

donde $\|w\| \leq \Delta_k$ y

$$\Psi_k(w) = \nabla f(x_k)' w + \frac{1}{2} w' \nabla^2 f(x_k) w$$

es el modelo cuadrático de f en una vecindad de la iteración x_k .

Esto sugiere que podría ser deseable calcular un paso s_k el cual resuelve aproximadamente el problema:

$$\min_{\|w\| \leq \Delta_k} \Psi_k(w) \tag{c.1}$$

si el paso es satisfactorio en el sentido que $x_k + s_k$ produce una reducción suficiente en f , entonces Δ_k puede ser incrementado y si el paso no es satisfactorio entonces Δ_k podría ser decrementado. El siguiente algoritmo expresa estas ideas con más detalle.

ALGORITMO c.2

Sean $0 < \mu < \eta < 1$ y $0 < \gamma_1 < \gamma_2 < 1 < \gamma_3$ constantes dadas.

1. Sean $x_0 \in \mathbb{R}^n$ y $\Delta_0 > 0$ valores iniciales.

2. Para $k = 0, 1, \dots$ hasta converger

(a) Calcular $\nabla f(x_k)$ y $\nabla^2 f(x_k)$.

(b) Determinar una aproximación a la solución s_k del problema (c.1).

(c) Calcular

$$\rho_k = \frac{f(x_k + s_k) - f(x_k)}{\Psi(s_k)}.$$

(d) Si $\rho_k \leq \mu$ entonces disminuir Δ_k por un punto del intervalo

$$[\gamma_1 \Delta_k, \gamma_2 \Delta_k]$$

e ir a b).

(e) Calcular $x_{k+1} = x_k + s_k$.

(f) Si $\rho_k \leq \eta$ entonces disminuir $\Delta_{k+1} \in [\gamma_2 \Delta_k, \Delta_k]$.

(g) Si $\rho_k \geq \mu$ entonces aumentar $\Delta_{k+1} \in [\Delta_k, \gamma_3 \Delta_k]$

Esta es una forma básica del método de Newton con región de confianza, una variante interesante de este algoritmo incluye una matriz de escala para las variables, que con esta variación el subproblema (c.1) es reemplazado por:

$$\min_{\|D_k w\| \leq \Delta_k} \Psi_k(w)$$

donde D_k es una matriz no singular.

No discutiremos esta generalización aquí; sin embargo, es importante notar que todos los resultados presentados aquí se cumplen para esta variante si $\{D_k\}$, cuya condición de la matriz D_k esta uniformemente acotada. Tal modificación puede ser importante en la práctica cuando las unidades de la variables tienen escalas diferentes.

No estamos interesados en resolver el problema (c.1) con gran precisión, sino que nos interesa proveer condiciones relajadas para aceptar una aproximación

de la solución s_k del problema (c.1) las cuales son suficientes para forzar la convergencia de la sucesión $\{x_k\}$ generada por el algoritmo (c.2).

Si Ψ_k^* es el valor óptimo de (c.1) y si la aproximación s_k a la solución de (c.1) satisface:

$$-\Psi_k(s_k) \geq \beta_1 |\Psi_k^*|, \quad \|s_k\| \leq \beta_2 \Delta_k \quad (c.4)$$

para constantes $\beta_1 > 0$ y $\beta_2 > 0$, entonces es posible comprobar que bajo ciertas condiciones de f , la sucesión $\{x_k\}$ es convergente al punto x^* con $\nabla f(x^*) = 0$ y $\nabla^2 f(x^*)$ positiva semidefinida.

No es difícil obtener un vector s_k que satisfaga (c.4), aunque requiere atención de ciertos detalles. Por ejemplo, dado $\sigma \in (0, 1)$ el algoritmo de Moré y Sorensen [19], encuentra un vector s_k tal que:

$$\Psi_k(s_k) - \Psi_k^* \leq \sigma(2 - \sigma) |\Psi_k^*|; \quad \|s_k\| \leq (1 + \sigma)\Delta_k$$

proporciona $\Psi_k^* \neq 0$.

Por supuesto, si $\Psi_k^* = 0$ entonces $\nabla f(x_k) = 0$ y $\nabla^2 f(x_k)$ es positiva semidefinida, por lo que Algoritmo (c.2) termina en x_k .

También es relevante mencionar que si $\sigma = 0.1$ entonces el costo de este algoritmo es muy razonable. En promedio, la solución aproximada de cada problema requiere menos de dos factorizaciones de una matriz simétrica definida positiva de orden n .

La condición (c.4) puede ser expresada en una forma alternativa la cual es más conveniente para las pruebas de convergencia. Si $p_k \in \mathbb{R}^n$ es una solución al problema (c.1) entonces el Lema B.2 implica que existe un parámetro $\lambda_k \geq 0$ tal que:

$$(\nabla^2 f(x) + \lambda_k I) p_k = -\nabla f(x_k); \quad \lambda_k(\Delta_k - \|p_k\|) = 0$$

Ahora, sea $R_k^t R_k$ la factorización de Cholesky de $\nabla^2 f(x) + \lambda_k I$, entonces:

$$|\Psi_k^*| = \frac{1}{2} \left(\|R_k p_k\|^2 + \lambda_k \Delta_k^2 \right) \quad (c.5)$$

esta expresión para Ψ_k^* muestra que si (c.4) se cumple entonces:

$$-\Psi_k(s_k) \geq \frac{1}{2} \beta_1 \left(\|R_k p_k\|^2 + \lambda_k \Delta_k^2 \right) \quad (c.6)$$

y de este modo las iteraciones $\{x_k\}$ generadas por el Algoritmo (c.2) satisface:

$$f(x_k) - f(x_{k+1}) \geq \frac{1}{2} \mu \beta_1 \left(\|R_k p_k\|^2 + \lambda_k \Delta_k^2 \right) \quad (c.7)$$

estas desigualdades son fundamentales para la prueba de nuestro siguiente resultado.

Teorema C.1 Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de clase C^2 sobre un conjunto abierto D y suponemos que el punto inicial x_0 es tal que el conjunto de nivel

$$\Omega = \{x \in D : f(x) \leq f(x_0)\}$$

es compacto. Si la sucesión $\{x_k\}$ es producida por el algoritmo (c.2), donde s_k satisface (c.4) entonces o el algoritmo termina en $x_t \in \Omega$ porque $\nabla f(x_t) = 0$ y $\nabla^2 f(x_t)$ es positiva semidefinida, o $\{x_k\}$ tiene un punto límite x^* en Ω con $\nabla f(x^*) = 0$ y $\nabla^2 f(x^*)$ es positiva semidefinida.

Como la demostración de esta teorema es larga y tediosa no la presentaremos aquí [20].

Los resultados que hemos establecido son solamente un ejemplo de los resultados de convergencia para el Algoritmo (c.2) bajo la suposición (c.4) para s_k . Escribimos a continuación resultados conocidos:

1. La sucesión converge a cero.
2. Si x^* es un punto límite aislado de $\{x_k\}$ entonces $\nabla^2 f(x^*)$ es positiva semidefinida.
3. Si $\nabla^2 f(x^*)$ es no singular para algún punto límite x^* de $\{x_k\}$ entonces $\{x_k\}$ converge a x^* .

De estos resultados, el segundo es característico de la aproximación de la región de confianza y es el único resultado que no se cumple para el método de Newton con búsqueda en la línea.

La diferencia entre las dos aproximaciones es de importancia teórica. Desde el punto de vista práctico, sin embargo, se puede argumentar que una diferencia importante es que con la aproximación con búsqueda en la línea, la búsqueda para el valor de la función menor ocurre en subespacio unidimensional, mientras que con una aproximación de región de confianza no es restringida a un subespacio de menor dimensión.

Apéndice D

Cálculos adicionales.

En este apéndice presentamos desarrollos, que omitimos por claridad en el capítulo III.

D.1 Cálculo del cociente de reducción.

En la sección III.3.2 definimos el cociente de reducción como:

$$\rho(\bar{s}) = \frac{\|G(\eta)\|^2 - \|G(\eta + \bar{s})\|^2}{\|G(\eta)\|^2 - \|G(\eta) + J_G(\eta)\bar{s}\|^2}$$

aquí realizamos las reducciones necesarias para obtener una expresión, en la cual no involucra la diferencia del denominador, para desarrollar esto, primero escribiremos ρ como:

$$\rho = \frac{\|G\|^2 - \|G_+\|^2}{\|G\|^2 - \|G + J_G\bar{s}\|^2} \quad (d.1.1)$$

donde $G_+ = G(\eta + \bar{s})$. Si utilizamos el hecho que \bar{s} satisface (iii.3.1.1)

$$(J_G^t J_G + \lambda D^t D)\bar{s} = -J_G^t G$$

multiplicamos por \bar{s}^t , obtenemos:

$$\bar{s}^t J_G^t J_G \bar{s} + \lambda \bar{s}^t D^t D \bar{s} = -\bar{s}^t J_G^t G$$

que es lo mismo; si utilizamos la notación de norma de un vector, nos queda:

$$\|J_G \bar{s}\|^2 + \lambda \|D \bar{s}\|^2 = -\bar{s}^t J_G^t G \quad (d.1.2)$$

Ahora desarrollamos el denominador de (d.1.1) y además utilizamos (d.1.2), así obtenemos el desarrollo:

$$\begin{aligned}
\|G\|^2 - \|G + J_G \bar{s}\|^2 &= G^t G - (G + J_G \bar{s})^t (G + J_G \bar{s}) \\
&= G^t G - (G^t G + 2\bar{s}^t J_G^t G + \bar{s}^t J_G^t J_G \bar{s}) \\
&= -2\bar{s}^t J_G^t G + \|J_G \bar{s}\|^2 \\
&= 2(\|J_G \bar{s}\|^2 + \lambda \|D\bar{s}\|^2) - \|J_G \bar{s}\|^2 \\
&= \|J_G \bar{s}\|^2 + 2\lambda \|D\bar{s}\|^2
\end{aligned}$$

que finalmente nos queda la igualdad:

$$\|G\|^2 - \|G + J_G \bar{s}\|^2 = \|J_G \bar{s}\|^2 + 2\lambda \|D\bar{s}\|^2 \quad (d.1.3)$$

aquí cumplimos nuestro objetivo de evitar la resta, evitando así errores de redondeo, ya que obtuvimos una expresión positiva puesto que $\lambda \geq 0$.

Reescribimos a (d.1.1) con la expresión (d.1.3) y además de dividimos tanto el denominador como numerador entre $\|G\|^2$, así obtenemos:

$$\rho = \frac{1 - \frac{\|G_+\|^2}{\|G\|^2}}{\frac{\|J_G \bar{s}\|^2}{\|G\|^2} + 2\lambda \frac{\|D\bar{s}\|^2}{\|G\|^2}} \quad (d.1.4)$$

esta es la expresión final, es la que calcularemos, ya que si de (d.1.3) despejamos la $\|G\|^2$ de la siguiente forma:

$$\|G\|^2 = \|J_G \bar{s}\|^2 + 2\lambda \|D\bar{s}\|^2 + \|G + J_G \bar{s}\|^2$$

concluimos fácilmente que se cumplen las siguientes dos desigualdades:

- $\|G\|^2 \geq \|J_G \bar{s}\|^2$
- $\|G\|^2 \geq 2\lambda \|D\bar{s}\|^2$ con $\lambda \geq 0$.

con estas desigualdades concluimos, que el cálculo del denominador de (d.1.4) no provocará algún overflow y además es siempre positivo, aun considerando

los errores de redondeo, y como habíamos mencionado en III.3.2, en el caso $\|G_+\|^2 > \|G\|^2$ asignamos $\rho = 0$, por lo que no calculamos (d.1.4) y por consiguiente evitamos un posible overflow.

D.2 Cálculos del ajuste de la cuadrática $q(\theta)$

En III.3.2 planteamos el problema del ajuste de una cuadrática de la forma:

$$q(\theta) = \alpha_1 + \alpha_2\theta + \alpha_3\theta^2 \quad (d.2.1)$$

a los puntos $\delta(0)$, $\delta'(0)$ y $\delta(1)$, donde la función δ esta definida como:

$$\delta(\theta) = \frac{1}{2} \|G(\eta + \theta)\|^2$$

para que finalmente, calculemos el mínimo μ de $q(\theta)$.

Consideremos que $q(\theta)$ tiene que ajustar a los puntos $\delta(0)$, $\delta'(0)$ y $\delta(1)$, de aquí obtenemos el siguiente sistema de ecuaciones para α_1 , α_2 y α_3 , de la forma:

$$\alpha_1 = \frac{1}{2} \|G\|^2$$

$$\alpha_2 = \bar{s}^t J_G^t G$$

$$\alpha_1 + \alpha_2 + \alpha_3 = \frac{1}{2} \|G_+\|^2$$

observamos que es un sistema de ecuaciones sencillo de resolver, ya que α_1 y α_2 tienen explícitamente sus valores y α_3 queda como:

$$\alpha_3 = \frac{1}{2} \|G_+\|^2 - \frac{1}{2} \|G\|^2 - \bar{s}^t J_G^t G$$

por lo que la cuadrática $q(\theta)$ es de la forma:

$$q(\theta) = \frac{1}{2} \|G\|^2 + (\bar{s}^t J_G^t G)\theta + \left(\frac{1}{2} \|G_+\|^2 - \frac{1}{2} \|G\|^2 - \bar{s}^t J_G^t G\right)\theta^2$$

para calcular el mínimo de esta cuadrática, primero obtenemos su derivada, pero para facilitar lo utilizamos (d.2.1), la cual queda como:

$$q'(\theta) = \alpha_2 + 2\alpha_3\theta$$

igualamos a cero y despejamos para obtener el valor para θ :

$$\theta = -\frac{\alpha_2}{2\alpha_3}$$

es un mínimo con la condición que la segunda derivada sea mayor que cero, como $q''(\theta) = 2\alpha_3$, por lo que si $\alpha_3 > 0$ entonces:

$$\mu = -\frac{\alpha_2}{2\alpha_3}$$

en este caso obtenemos el valor μ explícitamente con el siguiente desarrollo, en el cual utilizamos la igualdad (d.1.2) :

$$\begin{aligned} \mu &= -\frac{s^t J_G^t G}{2\left(\frac{1}{2}\|G_+\|^2 - \frac{1}{2}\|G\|^2 - s^t J_G^t G\right)} \\ &= \frac{\frac{1}{2}s^t J_G^t G}{\frac{1}{2}\left(\|G\|^2 - \|G_+\|^2\right) + s^t J_G^t G} \\ &= \frac{-\frac{1}{2}\left(\|J_G s\|^2 + \lambda\|D_s\|^2\right)}{\frac{1}{2}\left(\|G\|^2 - \|G_+\|^2\right) - \left(\|J_G s\|^2 + \lambda\|D_s\|^2\right)} \end{aligned}$$

si $\tau = -\left(\frac{\|J_G s\|^2}{\|G\|^2} + \lambda\frac{\|D_s\|^2}{\|G\|^2}\right)$ entonces :

$$\mu = \frac{\frac{1}{2}\tau}{\frac{1}{2}\left[1 - \frac{\|G_+\|^2}{\|G\|^2}\right] + \tau}$$

así obtenemos la expresión final para el mínimo μ de la cuadrática $q(\theta)$.

D.3 Cálculo de la derivada de $\phi(\lambda)$

Finalmente discutiremos como calcular $\phi'(\lambda)$, para esto utilizaremos la igualdad (iii.3.3.6), que escribimos a continuación:

$$\phi(\lambda) = \left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\| - \Delta$$

de la igualdad (iii.3.3.2) obtenemos:

$$D\bar{s}(\lambda) = -(\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G$$

calculamos $\phi'(\lambda)$ por medio de (iii.3.3.6), así obtenemos:

$$\phi'(\lambda) = \frac{\partial \left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\|}{\partial \lambda}$$

Ahora calculamos la siguiente derivada para calcular indirectamente $\phi'(\lambda)$:

$$\frac{\partial \left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\|^2}{\partial \lambda} = 2 \left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\| \phi'(\lambda) \quad (d.3.1)$$

calculamos la siguiente función:

$$\left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\|^2 = (\bar{J}^t G)^t (\bar{J}^t \bar{J} + \lambda I)^{-2} \bar{J}^t G$$

calculamos la derivada explícitamente de esta función con respecto a λ ; nos queda:

$$\frac{\partial \left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\|^2}{\partial \lambda} = -2(\bar{J}^t G)^t (\bar{J}^t \bar{J} + \lambda I)^{-3} \bar{J}^t G$$

con esta igualdad y (d.3.1) obtenemos :

$$\begin{aligned} \phi'(\lambda) &= \frac{-(\bar{J}^t G)^t (\bar{J}^t \bar{J} + \lambda I)^{-3} \bar{J}^t G}{\left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\|} \\ &= \frac{-(\bar{J}^t G)^t (\bar{J}^t \bar{J} + \lambda I)^{-1} (\bar{J}^t \bar{J} + \lambda I)^{-1} (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G}{\left\| (\bar{J}^t \bar{J} + \lambda I)^{-1} \bar{J}^t G \right\|} \quad (d.3.2) \\ &= \frac{-[D\bar{s}(\lambda)]^t (\bar{J}^t \bar{J} + \lambda I)^{-1} [D\bar{s}(\lambda)]}{\left\| D\bar{s}(\lambda) \right\|} \end{aligned}$$

Por el momento dejamos este desarrollo y haremos otro que nos servirá para

continuar con el cálculo de la expresión de la derivada de $\phi(\lambda)$:

$$\begin{aligned}
 (\tilde{J}^t \tilde{J} + \lambda I)^{-1} &= [(JD^{-1})^t (JD^{-1}) + \lambda I]^{-1} \\
 &= [D^{-t} (J^t J) D^{-1} + \lambda I]^{-1} \\
 &= [D^{-t} (J^t J + \lambda D^t D) D^{-1}]^{-1} \\
 &= D (J^t J + \lambda D^t D)^{-1} D^t
 \end{aligned}$$

sustituimos esta igualdad en (d.3.2) y obtenemos:

$$\phi'(\lambda) = \frac{-[D^t D \bar{s}(\lambda)]^t (J^t J + \lambda D^t D)^{-1} [D^t D \bar{s}(\lambda)]}{\| D \bar{s}(\lambda) \|}$$

si $q(\lambda) = D \bar{s}(\lambda)$ entonces:

$$\phi'(\lambda) = \frac{-[D^t q(\lambda)]^t (J^t J + \lambda D^t D)^{-1} [D^t q(\lambda)]}{\| q(\lambda) \|} \quad (d.3.3)$$

De acuerdo a lo que trabajamos en la sección III.3.1 obtenemos facilmente las siguientes igualdades:

$$\begin{bmatrix} Q^t & 0 \\ 0 & \Pi^t \end{bmatrix} \begin{bmatrix} J_G \\ \lambda^{\frac{1}{2}} D \end{bmatrix} \Pi = \begin{bmatrix} R^t \\ 0 \\ D_\lambda \end{bmatrix} \quad (d.3.4)$$

$$W^t \begin{bmatrix} R^t \\ 0 \\ D_\lambda \end{bmatrix} = \begin{bmatrix} R_\lambda \\ 0 \end{bmatrix} \quad (d.3.5)$$

todos estos bloques están definidos en dicha sección.

Las igualdades (d.3.4) y (d.3.5) las utilizamos para el siguiente desarrollo:

$$\begin{aligned}
 J_G^t J_G + \lambda D^t D &= [J_G^t \quad \lambda^{\frac{1}{2}} D^t] \begin{bmatrix} J_G^t \\ \lambda^{\frac{1}{2}} D^t \end{bmatrix} \\
 &= [J_G^t \quad \lambda^{\frac{1}{2}} D^t] \begin{bmatrix} Q & 0 \\ 0 & \Pi \end{bmatrix} \begin{bmatrix} Q^t & 0 \\ 0 & \Pi^t \end{bmatrix} \begin{bmatrix} J_G^t \\ \lambda^{\frac{1}{2}} D^t \end{bmatrix} \\
 &= \Pi [R^t \quad 0 \quad D_\lambda^t] \begin{bmatrix} R^t \\ 0 \\ D_\lambda \end{bmatrix} \Pi^t \\
 &= \Pi [R^t \quad 0 \quad D_\lambda^t] W W^t \begin{bmatrix} R^t \\ 0 \\ D_\lambda \end{bmatrix} \Pi^t \\
 &= \Pi [R_\lambda^t \quad 0] \begin{bmatrix} R_\lambda \\ 0 \end{bmatrix} \Pi^t \\
 &= \Pi R_\lambda^t R_\lambda \Pi^t
 \end{aligned}$$

con esto obtenemos finalmente:

$$J_G^t J_G + \lambda D^t D = \Pi R_\lambda^t R_\lambda \Pi^t$$

y como consecuencia de esta igualdad obtenemos:

$$\begin{aligned}
 (J_G^t J_G + \lambda D^t D)^{-1} &= (\Pi R_\lambda^t R_\lambda \Pi^t)^{-1} \\
 &= [(R_\lambda \Pi^t)^t (R_\lambda \Pi^t)]^{-1} \\
 &= (R_\lambda \Pi^t)^{-1} (R_\lambda \Pi^t)^{-t}
 \end{aligned}$$

utilizamos esta igualdad para reducir aun más (d.3.3), obtenemos :

$$\begin{aligned}
 \phi'(\lambda) &= \frac{-[D^t q(\lambda)]^t (R_\lambda \Pi^t)^{-1} (R_\lambda \Pi^t)^{-t} [D^t q(\lambda)]}{\|q(\lambda)\|} \\
 &= \frac{-\| (R_\lambda \Pi^t)^{-t} D^t q(\lambda) \|^2}{\|q(\lambda)\|} \\
 &= \frac{-\| R_\lambda^{-t} \Pi D^t q(\lambda) \|^2}{\|q(\lambda)\|} \\
 &= -\|q(\lambda)\| \left\| R_\lambda^{-t} \left[\frac{\Pi D^t q(\lambda)}{\|q(\lambda)\|} \right] \right\|^2.
 \end{aligned}$$

Bibliografía y referencias.

- [1] Baird, D. C. [1962]. *Experimentation: An introduction to measurement theory and experiment design*, Prentice-Hall, Englewood Cliffs, New Jersey.
- [2] Bartlett, M. S. [1949]. *Fitting a straight line when both variables are subject to error*, *Biometrics*, 5, pp. 207-212.
- [3] Boggs, P. T., Byrd, R. H. and Schnabel, R. B. [1987]. *A stable and efficient algorithm for nonlinear orthogonal distance regression*, *SIAM, J. Sci. Stat. Comput.*, 8, No. 6, pp. 1052-1078.
- [4] Boggs, P. T., Donaldson, J. R., Schnabel, R. B. and Spiegelman, C. H. [1988]. *A computational examination of orthogonal distance regression*, *J. Econometrics*, 38 1/2, pp. 169-201.
- [5] Boggs, P. T., Byrd, R. H., Donaldson, J. R. and Schnabel, R. B. [1989]. *ODRPACK - software for weighted orthogonal distance regression*, *ACM Trans. Math. Software*, 15, pp. 348-364.
- [6] Boggs, P. T., Byrd, R. H., Donaldson, J. R. and Schnabel, R. B. [1989a]. *User's reference guide for ODRPACK - software for weighted orthogonal distance regression*, Internal Report 89-4103, National Institute of Standards and Technology, Gaithersburg, Maryland.
- [7] Boggs, P. T. and Rogers, J. E. [1989]. *The computation and use of the asymptotic covariance matrix for measurement error models*, Internal Report 89-4102, (Revised 1990), National Institute of Standards and Technology, Gaithersburg, Maryland.

- [8] Boggs, P. T. and Rogers, J. E. [1990]. *Orthogonal distance regression*, Contemporary Mathematics, **112**, pp. 183-194.
- [9] Boggs, P. T., Rogers, J. E. and Schnabel, R. B. [1991]. *A second Monte Carlo study of weighted orthogonal distance regression for measurement error models*, Internal Report (in progress), National Institute of Standards and Technology, Gaithersburg, Maryland.
- [10] Britt, H. L. and Luecke, R. H. [1991]. *The estimation of parameters in nonlinear implicit models*, Technometrics, **15**, pp. 233-247.
- [11] Dennis, J. E. and Schnabel, R. B. [1983]. *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall, Englewood Cliffs, New Jersey.
- [12] Donaldson, J. R. and Schnabel, R. B. [1987]. *Computational experience with confidence regions and confidence intervals for nonlinear least squares*, Technometrics, **29**, No. 1, pp. 67-82.
- [13] Draper, N. R. and Smith, H. [1981]. *Applied regression analysis*, John Wiley and Sons, New York, New York.
- [14] Fletcher, R. [1980]. *Practical Methods of Optimization, 1, Unconstrained optimization*, John Wiley, New York.
- [15] Gill, P. E., Murray, W. and Wright, M. H. [1991]. *Numerical Linear Algebra and Optimization, 1*, Addison-Wesley.
- [16] Golub, G. H. and Van Loan, C. F. [1983]. *Matrix computations*, The Johns Hopkins University Press, Baltimore, Maryland.
- [17] Hebden, M. D. [1973]. *An algorithm for minimization using exact second derivatives*, Atomic Energy Research Establishment report T. P. 515, Harwell, England.
- [18] Moré, J. J. [1978]. *The Levenberg-Marquardt algorithm: Implementation and theory*, Proceedings of the Dundee Conference on Numerical Analysis, G. A. Watson, ed., Springer-Verlag, Berlin.

- [19] Moré, J. J. and Sorensen, D. C. [1983]. *Computing a trust region step*, SIAM, J. Sci. Stat. Comput., 4, No. 3, pp. 553-572.
- [20] Moré, J. J. and Sorensen, D. C. [1984]. *Newton's method*, MAA Studies in Mathematics, 24, Studies in Numerical Analysis, Gene H. Golub editor, The Mathematical Association of America.
- [21] Reinsch, C. H. [1967]. *Smoothing by spline functions*, Numer. Math., 10, pp. 177-183.
- [22] Reinsch, C. H. [1971]. *Smoothing by spline functions II*, Numer. Math., 16, pp. 451-454.
- [23] Sachs, W. H. [1976]. *Implicit multifunctional nonlinear regression analysis*, Technometrics, 18, No. 2, pp. 161-173.
- [24] Schnell, D. [1983]. *Maximum likelihood estimation of parameters in the implicit nonlinear errors-in-variables model*, Master's thesis, Iowa State University Department of Statistics, Ames, Iowa.
- [25] Schwetlick, H. and Tiller, V. [1985]. *Numerical methods for estimating parameters in nonlinear models with errors in the variables*, Technometrics, 27, No. 1, pp. 17-24.
- [26] Southwell, W. H. and Tiller, V. [1976]. *Fitting data to nonlinear functions with uncertainties in all measurement variables*, The Computer Journal, 19, No.1, pp. 69-73.
- [27] Stewart, G. W. [1973]. *Introduction to matrix computations*, Academic Press, New York, New York.
- [28] Wolter, K. M. and Fuller, W. A. [1982]. *Estimation of nonlinear errors-in-variables models*, Ann. Statist., 10, pp. 539-548.
- [29] York, D. [1966]. *Least squares fitting of a straight line*, Canadian Journal of Physics, 44, pp. 1079-1086.