# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

Facultad de Ciencias


# ESTIMACION NO LINEAL PARA MODELOS DE REGRESION

## T E S I S

Que para obtener el título de

A C T U A R I O

p r e s e n t a


MARTIN SALVADOR ARCEO FRANCO



México, D. F 1988

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 está 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.

INDICE

# INTRODUCCION



EL OBJETIVO DE ESTE TRABAJO ES PRESENTAR UN ESBOZO TEORICO DE DIFERENTES METODOS DE ESTIMACION PARA PARAMETROS EN MODELOS DE REGRESION NO LINEAL,OCUPANDONOS ESPECIFICAMENTE DE TRES DE ESTIMACION PUNTUAL Y OTROS DOS PARA LA ESTIMACION POR REGIONES DE CONFIANZA.

CON EL PROPOSITO DE DAR SUFICIENTE CLARIDAD AL TRABAJO,A LO LARGO DEL MISMO,HEMOS TRATADO DE PLANTEAR COMPARACIONES ENTRE LA PROBLEMATICA RELATIVA A LOS MODELOS FUNCIONALES LINEALES Y LA QUE ES PROPIA A LOS NO LINEALES,ASI COMO DETALLAR Y JUSTIFICAR AQUELLOS PUNTOS QUE CONSIDERAMOS MAS IMPORTANTES.CABE ACLARAR QUE NUESTRO TRABAJO FUE ENFOCADO CON MAS ENFASIS A LA ESTIMACION PUNTUAL.

LA EXPOSICION ESTA COMPUESTA DE NUEVE CAPITULOS,UN APARTADO DE CONCLUSIONES Y TRES APENDICES DE APOYO.

DENTRO DE SUS GENERALIDADES,EL TRABAJO MUESTRA UN CONSTANTE REPLANTEAMIENTO DEL MODELO FUNCIONAL,ASI COMO EL DAR POR HECHO QUE PARA EFECTOS DE ESTIMACION EL USO DE LA FUNCION DE MINIMOS CUADRADOS ES CONVENIENTE.

EN CUANTO A LA CONSTANTE REDEFINICION DEL MODELO FUNCIONAL,DIREMOS QUE FUE CONSIDERADO ASI CONVENIENTE PARA QUE, EN CADA CASO,SE EVITASEN CONFUSIONES EN CUANTO A NOTACION Y EXPOSICION DE LOS METODOS.POR LO QUE RESPECTA A LA FUNCION DE MINIMOS CUADRADOS,EL APENDICE III TRATA EN FORMA SOMERA SOBRE LA CONVENIENCIA DE SU USO PARA EFECTOS DE ESTIMACION.

LOS CAPITULOS 1 Y 2,TIENEN COMO FINALIDAD PRINCIPAL,SITUAR AL LECTOR DENTRO DE LA PROBLEMATICA DE ESTIMACION PARAMETRICA DE MODELOS DE REGRESION NO LINEAL,MARCANDO LA RAZONES POR LAS QUE ES NECESARIO BUSCAR ALTERNATIVAS DISTINTAS O HACER MODIFICACIONES A LAS UTILIZADAS PARA EL CASO LINEAL.

EN EL CAPITULO 3,SE PRESENTA EL PRIMER METODO ITERATIVO DE ESTIMACION PUNTUAL,QUE HACE USO DE LA EXPANSION EN SERIES DE TAYLOR HASTA GRADO UNO DEL MODELO FUNCIONAL. ASI MISMO,SE COMENTAN PORMENORES DE LA ELECCION DE UN VECTOR INICIAL,CARACTERISTICO DE TODOS LOS PROCEDIMIENTOS ITERATIVOS DE ESTIMACION.

EN EL CUARTO CAPITULO,SE ESTRUCTURA EN FORMA GENERAL EL PLANTEAMIENTO DE LOS "MODELOS ITERATIVOS ACEPTABLES",DANDO UNA PRUEBA GENERAL DE CONVERGENCIA PARA TODOS AQUELLOS METODOS QUE PERTENECEN A ESTA FAMILIA.

AUNQUE EL METODO DEL DESCENSO POR LA PENDIENTE MAXIMA FORMA PARTE DE LOS ANTERIORES,ES DETALLADO EN FORMA SEPARADA EN EL CAPITULO 5,BUSCANDO DEJAR ANTECEDENTE QUE CUALQUIERA DE LOS PROCEDIMIENTOS ITERATIVOS ACEPTABLES PUEDE SURGIR EN FORMA NATURAL E INDEPENDIENTE DEL RESTO DE SU CLASE.

EN EL CAPITULO 6 APARECE EL METODO DE MARQUARDT,TAMBIEN ITERATIVO ACEPTABLE,QUE DE ALGUNA FORMA MEJORA LOS DOS ANTERIORES.

LA SEGUNDA PARTE DEL TRABAJO INICIA CON EL CAPITULO 7,MISMO QUE PRETENDE CLARIFICAR LA IDEA GEOMETRICA DE LA FUNCION DE MINIMOS CUADRADOS,TANTO PARA EL CASO LINEAL COMO PARA EL NO LINEAL,DEJANDO CLARAS SUS DIFERENCIAS.

LOS CAPITULOS 8 Y 9,FUNDAMENTALMENTE EXPONEN LOS METODOS DE ESTIMACION POR REGIONES DE CONFIANZA QUE HEMOS SELECCIONADO PARA PRESENTAR EN ESTE COMPENDIO;FINALMENTE INCLUIMOS LAS CONCLUSIONES SOBRE LO QUE FUE PLASMADO EN LOS NUEVE CAPITULOS.

# 1.- MODELOS FUNCIONALES

UNA DE LAS SITUACIONES MAS COMUNES EN EL ANALISIS ESTADISTICO,ES AQUELLA DONDE PARA ESTUDIAR UN FENOMENO SE PROPONE UN MODELO FUNCIONAL CON UNA VARIABLE RESPUESTA QUE DEPENDE DE CIERTAS VARIABLES CONTROLADAS DEFINIDAS O EXPERIMENTALES.

LA VARIABLE DEPENDIENTE CUNTIFICA LA RESPUESTA DEL FENOMENO BAJO OBSERVACION Y LAS VARIABLES INDEPENDIENTES REPRESENTAN LOS FACTORES CONTROLABLES QUE INTERVIENEN EN EL MISMO.

ES INDUDABLE QUE EN CUALQUIER FENOMENO INTERVIENEN UN GRAN NUMERO DE FACTORES QUE NO PODEMOS CONTROLAR,POR LO QUE EL MODELO FUNCIONAL QUE SE PROPONGA DEBE CONTEMPLAR ESTA SITUACION.EXPLICITAMENTE,EL MODELO PUEDE SER REPRESENTADO POR LA SIGUIENTE ECUACION CONOCIDA COMO EL MODELO POBLACIONAL:

(1.1) $$Y=f(\Phi;\theta)+\varepsilon$$

DADO QUE NO SE PUEDE TRABAJAR CON TODA LA POBLACION SE ELIJE UNA MUESTRA ALEATORIA Y LA ECUACION (1.1) TOMA LA FORMA DE LAS SIGUIENTES ECUACIONES CONOCIDAS COMO EL MODELO MUESTRAL:

(1.2) $$Y_u=f(\Phi_u;\theta)+\varepsilon_u \qquad u=1,2,\ldots,n$$

DONDE $\Phi_u$ REPRESENTA UN VECTOR CUYAS COMPONENTES SON FUNCIONES DE k VARIABLES CONTROLADAS;$\theta$ ES UN VECTOR p-DIMENSIONAL DE PARAMETROS DESCONOCIDOS;$f(\Phi_u;\theta)$ REPRESENTA APROXIMADAMENTE EL EFECTO DEL VECTOR $\Phi_u$ SOBRE LA VARIABLE RESPUESTA $Y_u$;$\varepsilon_u$ ES UNA VARIABLE ALEATORIA QUE CUANTIFICA EL EFECTO DE LOS FACTORES NO CONTROLADOS O BIEN DEL ERROR DE APROXIMACION DE $f(\Phi_u;\theta)$ A $Y_u$.

USUALMENTE LA FUNCION $f(\Phi_u;\theta)$,PUEDE REPRESENTAR ADECUADAMENTE EL EFECTO MENCIONADO SOBRE $Y_u$ EN FORMA LINEAL CON RESPECTO A LOS PARAMETROS.EXPRESAMENTE:

(1.3) $$Y_u=\theta_1\phi_{u1}+\theta_2\phi_{u2}+\ldots+\theta_p\phi_{up}+\varepsilon_u \qquad u=1,2,\ldots,n$$

O BIEN,EN FORMA MATRICIAL:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & . . . & \phi_{1p} \\ \phi_{21} & \phi_{22} & . . . & \phi_{2p} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \phi_{n1} & \phi_{n2} & . . . & \phi_{np} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdot \\ \cdot \\ \cdot \\ \theta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix}$$

(1.4) $Y = X\theta + \varepsilon \quad ; \quad X = (\phi_{ij})$

$\phi_{ij}$=VALOR QUE TOMA LA j-ESIMA FUNCION DE LAS VARIABLES CONTROLADAS EN LA i-ESIMA OBSERVACION

CUANDO SUPONEMOS ADEMAS QUE LA MATRIZ X TIENE RANGO COMPLETO Y QUE LAS VARIABLES $\varepsilon_u$ SON NO CORRELACIONADAS E IDENTICAMENTE DISTRIBUIDAS COMO NORMALES CON $E(\varepsilon_u)=0$, $VAR(\varepsilon_u)=\sigma^2$ $\forall$ u, ES DECIR:

(1.5) RANGO(X)=p ; $\varepsilon \sim \mathfrak{n}(0,\sigma^2)$ $(p < n)$

EL ESTIMADOR PARA $\theta$ QUE MINIMIZA LA FUNCION DE MINIMOS CUADRADOS:

(1.6) $S(\theta) = \varepsilon^t \varepsilon = \Sigma_{u=1}{}^{n} (Y_u - f(\Phi_u;\theta))^2$

COINCIDE CON EL ESTIMADOR MAXIMO VEROSIMIL Y ES:

(1.7) $\hat{\theta} = (X^t X)^{-1} X^t Y$

RESULTA RELEVANTE PREGUNTARSE DE (1.5) POR QUE $p < n$ ,LA RESPUESTA NO ES MUY COMPLICADA;SI SUCEDIERA $p > n$ ,Y X TUVIESE RANGO COMPLETO,ENTONCES RANGO(X)=n Y RANGO($X^t X$)=n PERO COMO $X^t X$ ES DE $p \times p$,NO EXISTIRIA $(X^t X)^{-1}$.SI SUCEDIERA QUE p=n,EL MODELO SERIA HIPERSENSIBLE.

CABE SENALAR QUE EXISTEN MODELOS FUNCIONALES QUE SI BIEN EN UN PRINCIPIO NO TIENEN UNA EXPRESION LINEAL EN LOS PARAMETROS,PUEDEN ADQUIRIRLA MEDIANTE TRANSFORMACIONES CONVENIENTES,POR EJEMPLO:

$$Y_u=\exp(\theta_1+\theta_2\Phi_u^2+\varepsilon_u) \qquad u=1,2,\ldots,n$$

APLICANDO LOGARITMO NATURAL A AMBOS MIEMBROS DE LA ECUACION SE TIENE:

$$\log(Y_u)=\theta_1+\theta_2\Phi_u^2+\varepsilon_u \qquad u=1,2,\ldots,n$$

SI DEFINIMOS:

$$Y_u'=\log(Y_u)$$

TENEMOS:

$$Y_u'=\theta_1+\theta_2\Phi_u^2+\varepsilon_u \qquad u=1,2,\ldots,n$$

QUE RESULTA SER UN MODELO LINEAL EN LOS PARAMETROS.A ESTA CLASE DE MODELOS FUNCIONALES SE LES CONOCE COMO MODELOS INTRINSICAMENTE LINEALES.

AUNQUE EL CAMPO DE LOS MODELOS LINEALES E INTRINSECAMENTE LINEALES PUEDE ABARCAR UN GRAN NUMERO DE PROBLEMAS PRACTICOS,EXISTEN MUCHAS SITUACIONES EN LAS QUE UN MODELO FUNCIONAL NO LINEAL (CONSIDERANDO COMO TAL A UNO QUE NO SEA NI LINEAL NI INTRINSECAMENTE LINEAL),SE ADECUA MAS A LA NATURALEZA DE LOS DATOS QUE SE POSEEN PARA EL FENOMENO QUE SE ESTA ESTUDIANDO. POR EJEMPLO,CUANDO LA VARIABLE RESPUESTA SURGE COMO LA SOLUCION DE UNA ECUACION DIFERENCIAL,EL MODELO FUNCIONAL CORRESPONDIENTE PODRIA TOMAR LA FORMA:

$$f(\Phi_u:\theta)=\theta_1+\theta_2\exp(\Phi_u\theta_3) \qquad u=1,2,\ldots,n$$

O CUANDO SABEMOS QUE LA RESPUESTA ES PERIODICA PERO DESCONOCEMOS EL PERIODO,UN MODELO FUNCIONAL PERTINENTE PODRIA SER:

$$f(\Phi_u;\theta)=\theta_1+\theta_2\cos(\theta_4\Phi_u)+\theta_3\mathrm{sen}(\theta_4\Phi_u) \qquad u=1,2,\ldots,n$$

O BIEN,

$$Y_u=\theta_1+\theta_2\cos(\theta_4\Phi_u)+\theta_3\mathrm{sen}(\theta_4\Phi_u)+\varepsilon_u \qquad u=1,2,\ldots,n$$

PARA ESTIMAR LOS PARAMETROS EN ESTE TIPO DE PROBLEMAS PODRIAMOS RECURRIR A MINIMIZAR LA FUNCION $S(\theta)$,COMO FUE DEFINIDA EN (1.6),Y OBTENER EL ESTIMADOR DE MINIMOS CUADRADOS,DE MANERA ANALOGA AL CASO LINEAL.

RESULTA PUES INTERESANTE, ANALIZAR LOS PROBLEMAS QUE SE ENFRENTAN PARA PONER EN PRACTICA ESTE METODO CUANDO SE PRESENTA UN PROBLEMA NO LINEAL.

## 2.- MINIMOS CUADRADOS EN EL CASO NO LINEAL



POR RESULTAR CONVENIENTE PARA LA EXPOSICION QUE DESARROLLAREMOS, HAREMOS ALGUNOS CAMBIOS A LA NOTACION QUE SE UTILIZA EN EL MODELO FUNCIONAL PARA MINIMOS CUADRADOS EN EL CASO LINEAL.

SUPONGAMOS QUE TENEMOS n OBSERVACIONES DE UN FENOMENO BAJO ESTUDIO Y QUE POSTULAMOS UN MODELO MUESTRAL DE LA SIGUIENTE FORMA:

(2.1) $$Y_u = f(\xi_u;\theta) + \varepsilon_u \qquad u = 1, 2, \ldots, n.$$

DONDE:

$$\xi_u = \begin{bmatrix} \xi_{1u} \\ \xi_{2u} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{ku} \end{bmatrix}^t$$ ; VECTOR DE k FUNCIONES DE LAS VARIABLES CONTROLADAS

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdot \\ \cdot \\ \cdot \\ \theta_p \end{bmatrix}^t$$ ; VECTOR DE p PARAMETROS DESCONOCIDOS

$\varepsilon_u$ ; VARIABLE ALEATORIA NORMAL DISTIBUIDA CON $E(\varepsilon_u)=0$ Y $VAR(\varepsilon_u)=\sigma^2$ $u=1, 2, \ldots, n$ Y PARA $i \neq j$ $\varepsilon_i$ Y $\varepsilon_j$ SON NO CORRELACIONADAS. A $\varepsilon_u$ SE LE CONOCE COMO EL ERROR PARA LA u-ESIMA OBSERVACION.

CON LAS SUPOSICIONES ANTERIORES CONCLUIMOS QUE SI:

IGUALANDO A CERO TENEMOS:

(2.6) $\Sigma_{u=1}^{n}(Y_u - f(\xi_u;\theta))\ [\delta f(\xi_u;\theta)/\delta\theta_i\ |\theta = \hat{\theta}] = 0 \quad i=1,2,\ldots,p$

DONDE,

$[\delta f(\xi_u;\theta)/\delta\theta_i\ |\theta = \hat{\theta}\ ]$ DENOTA EL GRADIENTE DE $f(\xi_u;\theta)$ EVALUADO EN $\theta = \hat{\theta}$

HASTA AQUI,LA TEORIA DESARROLLADA ES IDENTICA QUE EN EL CASO LINEAL,PERO EN ESTE PUNTO CABE HACER LA SIGUIENTE OBSERVACION;CUANDO EL MODELO FUNCIONAL ES LINEAL EN LOS PARAMETROS,LA EXPRESION DEL GRADIENTE ES MUY SIMPLE DADO QUE NO INVOLUCRA FUNCIONES DE $\theta$,POR EJEMPLO,SI:

$$f(\xi;\theta) = \theta_1\xi_1+\theta_2\xi_2+\ldots+\theta_p\xi_p$$

ENTONCES:

(2.7) $\delta f(\xi;\theta)/\delta\theta_i = \xi_i \qquad \forall\ i=1,2,\ldots,p$

ESTA SITUACION FACILITA EL PROCESO DE ESTIMACION.

DESAFORTUNADAMENTE NO SUCEDE LO MISMO PARA EL CASO DE LOS MODELOS FUNCIONALES NO LINEALES;PENSEMOS POR EJEMPLO EN UNA FUNCION DE ESTE TIPO RELATIVAMENTE SENCILLA COMO LA QUE APARECE A CONTINUACION:

(2.8) $f(\xi) = \exp(-\theta\xi)$

$\Rightarrow \delta f/\delta\theta = -\theta\exp(-\theta\xi)$

SUSTITUYENDO ESTE RESULTADO EN (2.6),TENEMOS:

(2.9) $\Sigma_{u=1}^{n} Y_u\xi_u \exp(-\hat{\theta}\xi_u) - \Sigma_{u=1}^{n}\xi_u \exp(-2\hat{\theta}\xi_u) = 0$

ES MUY CLARO EL ALTO GRADO DE DIFICULTAD QUE SE TENDRIA PARA TRATAR DE DESPEJAR $\hat{\theta}$ EN LA ECUACION (2.9).EVIDENTEMENTE,ESTA DIFICULTAD SE ACENTUA AUN MAS CUANDO LOS MODELOS FUNCIONALES NO LINEALES SON MAS COMPLICADOS E INVOLUCRAN MAS PARAMETROS.

A ESTO, SE DEBE AÑADIR EL HECHO DE QUE EN EL CASO NO LINEAL,LA FUNCION $S(\theta)$ PUEDE TENER VARIOS MINIMOS LOCALES ADEMAS DE UNO O TAMBIEN VARIOS MINIMOS GLOBALES,COMO MAS ADELANTE VEREMOS.

HA SIDO NECESARIO ENTONCES,QUE SE BUSQUEN ALTERNATIVAS PARA

LA ESTIMACION NO LINEAL DE PARAMETROS,QUE SUPEREN LOS OBSTACULOS QUE SE PRESENTAN CUANDO SE TRATA DE APLICAR EL METODO DE MINIMOS CUADRADOS EN FORMA DIRECTA.

EN EL PRESENTE,EXPONDREMOS TRES METODOS ITERATIVOS QUE PUEDEN UTILIZARSE COMO ALTERNATIVAS PARA LA ESTIMACION PUNTUAL DEL VECTOR PARAMETRICO $\theta$.NOS REFERIMOS A UN METODO ITERATIVO COMO AQUEL QUE SE REPITE TANTAS VECES SEA NECESARIO HASTA QUE,CON BASE EN ALGUN CRITERIO DEFINIDO,SE CONSIDERE QUE LOS RESULTADOS FINALMENTE ENCONTRADOS CONVERGEN A UNA SOLUCION O BIEN DIVERGEN.

ES NECESARIO MENCIONAR QUE ESTOS TRES METODOS TIENEN LA CARACTERISTICA DE UTILIZAR UN PUNTO INICIAL $\theta^0$,SOBRE CUYA ELECCION E IMPORTANCIA,HEMOS CONSIDERADO PERTINENTE TRATAR INMEDIATAMENTE DESPUES DE EXPONER EL PRIMERO DE ELLOS.

LOS METODOS SON:

-LINEALIZACION

-DESCENSO POR LA PENDIENTE MAXIMA

-METODO DE MARQUARDT

ESTE METODO ESTA BASADO EN LA EXPANSION EN SERIES DE TAYLOR HASTA GRADO UNO DE LA FUNCION $f(\xi;\theta)$, PARTIENDO DE UN PUNTO INICIAL $\theta^0$ CONVENIENTEMENTE ELEGIDO.

EL PROCEDIMIENTO ES EL SIGUIENTE:

SUPONGAMOS QUE HEMOS ELEGIDO UN PUNTO INICIAL $\theta^0=(\theta_1{}^0,\theta_2{}^0,\ldots,\theta_p{}^0)$, PODEMOS DECIR QUE SI $\theta$ ESTA CERCANO A $\theta^0$, APROXIMADAMENTE SUCEDE:

$$(3.1)\quad f(\xi_u;\theta) = f(\xi_u;\theta^0) + \Sigma_{i=1}{}^p\ [\delta f(\xi_u;\theta)/\delta\theta_i\ |\theta=\theta^0]\ (\theta_i-\theta_i{}^0)\qquad \forall\ u=1,2,\ldots,n$$

OBSERVEMOS QUE EL SEGUNDO MIEMBRO DE LA ECUACION (3.1), ES LA SERIE DE TAYLOR HASTA GRADO UNO PARA LA FUNCION $f(\xi_u;\theta)$ ALREDEDOR DE $\theta^0$.

NUEVAMENTE HAREMOS ALGUNOS CAMBIOS A LA NOTACION POR ASI CONSIDERARLO CONVENIENTE.

DEFINIMOS:

$$(3.2)\qquad \begin{aligned} f_u{}^0 &= f(\xi_u;\theta^0)\\ \beta_i{}^0 &= (\theta_i-\theta_i{}^0) \qquad\qquad \forall\ u=1,2,\ldots,n\\ Z^0{}_{iu} &= [\delta f(\xi_u;\theta)/\delta\theta_i\ |\theta=\theta^0] \end{aligned}$$

ASI $f_u{}^0$ ES EL VALOR DE LA FUNCION $f(\xi_u;\theta)$ SUSTITUYENDO EL VALOR DEL VECTOR QUE SE DESEA ESTIMA $\theta$, POR $\theta^0$, $\beta_i{}^0$ REPRESENTA LA DIFERENCIA ENTRE LA i-ESIMA COMPONENTE DEL VECTOR PARAMETRICO POR ESTIMAR Y LA CORRESPONDIENTE DEL PREVIAMENTE ELEGIDO. ES IMPORTANTE NOTAR QUE EL ESTIMAR $\beta_i{}^0$, EQUIVALDRIA A ESTIMAR $\theta_i$. POR ULTIMO, $Z^0{}_{iu}$ ES EL VECTOR DE DERIVADAS PARCIALES DE LA FUNCION $f(\xi_u;\theta)$ CON RESPECTO A $\theta$ EVALUADO EN $\theta^0$.

AHORA BIEN, EL MODELO FUNCIONAL QUE ESTAMOS CONSIDERANDO ES:

$$(3.3)\qquad Y_u = f(\xi_u;\theta)+\varepsilon_u \qquad\qquad u=1,2,\ldots,n$$

Y DE (3.1) TENEMOS:

$$(3.4)\quad -f(\xi_u;\theta^0) = -f(\xi_u;\theta) + \Sigma_{i=1}^{p} [\delta f(\xi_u;\theta)/\delta\theta_i \,|\theta=\theta^0] \,(\theta_i-\theta_i^{\,0})$$

$u=1,2,\ldots,n$

DE DONDE SUMANDO (3.3) Y (3.4) SE TIENE:

$$Y_u-f(\xi_u;\theta^0) = \Sigma_{i=1}^{p} [\delta f(\xi_u;\theta)/\delta\theta_i \,|\theta=\theta^0] \,(\theta_i-\theta_i^{\,0}) + \varepsilon_u$$

$u=1,2,\ldots,n$

FINALMENTE, SI EN LA ECUACION ANTERIOR SUSTITUIMOS LO DEFINIDO EN (3.2) TENEMOS:

$$(3.5)\qquad Y_u-f_u^{\,0} = \Sigma_{i=1}^{p}\beta_i^{\,0}Z^0_{iu} + \varepsilon_u \qquad u=1,2,\ldots,n$$

ECUACION QUE PODEMOS ESCRIBIR EN FORMA MATRICIAL COMO SIGUE:

$$(3.6)\qquad \mathbb{Y}-\mathbb{F}^0 = \mathbb{Z}^0\mathbb{B}^0$$

DONDE:

$$\mathbb{Y}-\mathbb{F}^0 = \begin{bmatrix} Y_1-f_1^{\,0} \\ Y_2-f_2^{\,0} \\ \cdot \\ \cdot \\ \cdot \\ Y_n-f_n^{\,0} \end{bmatrix} \qquad \mathbb{B} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix}$$

$$\mathbb{Z}^0 = \begin{bmatrix} Z^0_{11} & Z^0_{21} & \cdot & \cdot & \cdot & Z^0_{p1} \\ Z^0_{12} & Z^0_{22} & \cdot & \cdot & \cdot & Z^0_{p2} \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ Z^0_{1n} & Z^0_{2n} & \cdot & \cdot & \cdot & Z^0_{pn} \end{bmatrix}$$

OBSERVEMOS QUE LA ECUACION (3.6) ES LINEAL EN EL PARAMETRO $\mathbb{B}^0$, DE TAL SUERTE QUE SI LA MATRIZ $\mathbb{Z}^0$ ES DE RANGO COMPLETO, EL ESTIMADOR DE MINIMOS CUADRADOS (Y MAXIMO VEROSIMIL) PARA $\mathbb{B}^0$ ES:

$$(3.7)\qquad \hat{\mathbb{B}}^0 = [(\mathbb{Z}^0)^t\mathbb{Z}^0]^{-1}(\mathbb{Z}^0)^t(\mathbb{Y}-\mathbb{F}^0)$$

ES DECIR, EL VECTOR $\hat{\mathbb{B}}^0$ MINIMIZA LA FUNCION:

$$(3.8) \qquad S'(\theta) = \sum_{u=1}^{n}\left( Y_u - f(\xi_u;\theta^0) - \sum_{i=1}^{p} \beta_i^0 Z_{iu}^0 \right)^2$$ 

CON RESPECTO A $\beta_i^0 = \theta_i - \theta_i^0 \qquad i=1,2,\dots,p$

AHORA BIEN, SI $\hat{\beta}_i^0$ ES ES ESTIMADOR DE MINIMOS CUADRADOS PARA $\beta_i$, TENEMOS ENTONCES QUE POR EL PRINCIPIO DE INVARIANZA SUCEDE:

$$(3.9) \qquad \hat{\beta}_i^0 = \hat{\theta}_i - \theta_i^0 \qquad i=1,2,\dots,p$$

Y DE LA ECUACION (3.9) PODEMOS DESPEJAR $\hat{\theta}_i$ EL ESTIMADOR DE $\theta_i$ EN LA PRIMERA ITERACION.

DEFINAMOS PARA ESTA PRIMERA ITERACION $\theta_i^1 = \hat{\theta}_i \ \forall\ i=1,2,\dots,p$. EL SIGUIENTE PASO DEL PROCEDIMIENTO ES REPETIR EL RAZONAMIENTO HACIENDO QUE LOS ESTIMADORES $\theta_i^1$ $i=1,2,\dots,p$, SEAN UTILIZADOS EN EL PAPEL QUE JUGARON LOS $\theta_i^0$ $i=1,2,\dots,p$, Y ASI SUCESIVAMENTE.

EN NOTACION MATRICIAL, Y DEFINIENDO COMO $\theta^j$ AL ESTIMADOR OBTENIDO PARA $\theta$ EN LA j-ESIMA ITERACION, PODEMOS ESCRIBIR:

$$(3.10) \qquad \begin{aligned} \theta^j &= \theta^{j-1} + \hat{\mathbb{B}}^{j-1} \\ &= \theta^{j-1} + [(Z^{j-1})^t Z^{j-1}]^{-1} (Z^{j-1})^t (\mathbb{Y} - \mathbb{F}^{j-1}) \end{aligned}$$

DONDE:

$Z^{j-1} = (Z_{iu}^{j-1})$ Y $Z_{iu}^{j-1}$ ES LA DERIVADA PARCIAL DE LA FUNCION $f(\xi;\theta)$ CON RESPECTO A $\theta_i$ EVALUADA EN $\theta^{j-1}$ Y EN EL u-ESIMO ENSAYO

$\mathbb{F}^{j-1} = (f_1^{j-1}, f_2^{j-1}, \dots, f_n^{j-1})^t$ Y $f_u^{j-1}$ ES LA FUNCION $f(\xi;\theta)$ EVLUADA EN $\theta^{j-1}$ Y EN EL u-ESIMO ENSAYO

$\theta^{j-1} = (\theta_1^{j-1}, \theta_2^{j-1}, \dots, \theta_p^{j-1})$ Y $\theta_i^{j-1}$ ES EL ESTIMADOR DEL i-ESIMO PARAMETRO EN LA (j-1)-ESIMA ITERACION DEL PROCESO

EL PROCESO DEBE CONTINUAR HASTA QUE LOS ESTIMADORES ENCONTRADOS CONVERJAN A UNA SOLUCION; ESTO ES, HASTA QUE CON BASE A UN CRITERIO CONVENIENTEMENTE ELEGIDO, LOS RESULTADOS ITERATIVAMENTE OBTENIDOS SE ESTABILICEN. N.R. DRAPPER Y H. SMITH [6], PROPONEN CONTINUARLO HASTA QUE PARA LAS ITERACIONES SUCESIVAS (j-1) Y j, SE TENGA:

(3.11) $\quad | (\theta_i^{\,j} - \theta_i^{\,j-1})/\theta_i | \quad < \delta \quad i=1,2,\ldots,p$

CON $\delta > 0$ SUFICIENTEMENTE PEQUENA (POR EJ. $\delta = 1 \times 10^{-6}$).

ES COVENIENTE TAMBIEN,EVALUAR LA FUNCION S($\theta$) EN CADA ITERACION PARA VER SI EFECTIVAMENTE VA DECRECIENDO EN FORMA REGULAR A MEDIDA QUE EL PROCESO AVANZA.

EL METODO DE LINEALIZACION PUEDE PRESENTAR CIERTAS DESVENTAJAS EN ALGUNOS PROBLEMAS ESPECIFICOS.LOS INCONVENIENTES PUEDEN SER TALES COMO:

A) LA CONVERGENCIA PUEDE SER MUY LENTA.ESTO ES,PUEDE SUCEDER QUE SEA NECESARIO UN NUMERO MUY GRANDE DE ITERACIONES ANTES QUE LOS ESTIMADORES ENCONTRADOS CONVERJAN,AUN CUANDO EL VALOR DE LA FUNCION S($\theta$) DECREZCA CONSISTENTEMENTE.

B) PUEDE SUCEDER TAMBIEN QUE LOS ESTIMADORES OBTENIDOS EN LAS DIFERENTES ITERACIONES PROVOQUEN QUE LOS RESPECTIVOS VALORES DE LA FUNCION S($\theta$) OSCILEN CONSIDERABLEMENTE ES DECIR,QUE LOS VALORES DE S($\theta$) AUMENTEN Y DISMINUYAN DE MANERA IRREGULAR A MEDIDA QUE EL PROCESO AVANZA.NO OBSTANTE LO ANTERIOR LA SOLUCION PODRIA ESTABILIZARSE EVENTUALMENTE.

C) SERIA POSIBLE TAMBIEN QUE NO EXISTIERA CONVERGENCIA Y QUE LA FUNCION S($\theta$) SE INCREMENTARA ITERACION A ITERACION PARECIENDO NO TENER COTA.

DE CUALQUIER FORMA,ES EN MUCHOS CASOS PRACTICOS UN METODO DE ESTIMACION PARAMETRICA PARA MODELOS NO LINEALES CONVENIENTE.

EN RESUMEN,PODEMOS DECIR QUE EL METODO DE LINEALIZACION PARA MODELOS DE REGRESION NO LINEALES TRANSFORMA EL PROBLEMA DE ENCONTRAR UN MINIMO DE S($\theta$),EN UNA SERIE DE PROBLEMAS LINEALES CON EL MISMO OBJETIVO,PARTIENDO DE UN VALOR INICIAL $\theta^0$.

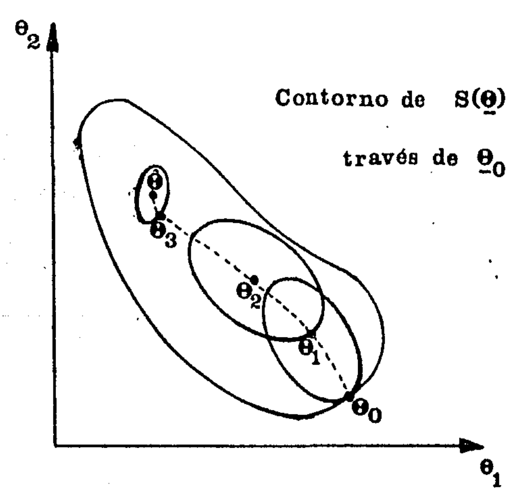
AHORA BIEN,MIENTRAS QUE EN EL CASO LINEAL LAS CURVAS DE NIVEL DE LA FUNCION S($\theta$) TIENEN UNA FORMA ELIPTICA,HECHO QUE PERMITE AFIRMAR QUE S($\theta$) POSEE SOLO UN MINIMO LOCAL Y POR CONSIGUIENTE SOLO UN MINIMO GLOBAL,EN EL CASO NO LINEAL ESTAS PUEDEN TENER UNA FORMA MUY IRREGULAR POR LO QUE PUEDE SUCEDER QUE S($\theta$) TENGA VARIOS

MINIMOS LOCALES Y TAL VEZ MAS DE UN MINIMO GLOBAL.

EL METODO DE LINEALIZACION REEMPLAZA LA FORMA IRREGULAR DE LAS CURVAS DE NIVEL DE $S(\theta)$ EN CURVAS DE NIVEL ELIPTICAS CARACTERISTICAS AL CASO LINEAL.POR TRABAJAR CON $S'(\theta)$.

LO QUE PRETENDE ESTE METODO ES ENCONTRAR EN ALGUN PASO ITERATIVO EL VERDADERO VALOR DE $\hat{\theta}$ QUE SEA UN MINIMO GLOBAL DE $S(\theta)$.RESULTA INTUITIVAMENTE EVIDENTE LA IMPORTANCIA DE UNA ELECCION ACERTADA DEL PUNTO $\theta^0$ YA QUE,ENTRE OTRAS COSAS,PODRIA ELEGIRSE POR EJEMPLO,UN PUNTO QUE CONDUJERA A UN MINIMO LOCAL DE $S(\theta)$ QUE DISTASE MUCHO DEL VALOR DEL MINIMO O MINIMOS GLOBALES PARA ESTA FUNCION.

A CONTINUACION MOSTRAMOS UNA GRAFICA QUE DESCRIBE LO SENALADO ANTERIORMENTE CUANDO $\theta = (\theta_1, \theta_2)$.




ABRIREMOS AQUI UN PARENTESIS PARA TRATAR CON MAS DETALLE LA IMPORTANCIA QUE TIENE LA ELECCION DEL PUNTO INICIAL $\theta^0$ ASI COMO DE ALGUNAS SUGERENCIAS PARA OBTENERLO.

CABE SENALAR QUE CUALQUIER METODO ITERATIVO REQUIERE DE LA ELECCION DE UN PUNTO INICIAL,POR TANTO LO QUE SE COMENTE SUBSECUENTEMENTE SOBRE EL,SERA VALIDO TANTO PARA EL METODO EXPUESTO CON ANTERIORIDAD COMO PARA LOS OTROS DOS DE LOS QUE NOS OCUPAREMOS POSTERIORMENTE

COMO FUE SENALADO CON ANTERIORIDAD, CUANDO SE TRABAJA CON MODELOS NO LINEALES LA FUNCION $S(\theta)$ PUEDE TENER VARIOS MINIMOS LOCALES ADEMAS DE UNO O VARIOS MINIMOS GLOBALES, POR LO QUE UNA ELECCION POCO CONVENIENTE DE $\theta^0$ PODRIA CONDUCIR A UN PUNTO ESTACIONARIO NO DESEADO O BIEN A UNA CONVERGENCIA EXCESIVAMENTE LENTA DEL METODO.

SOBRE COMO ELEGIR $\theta^0$ DRAPPER Y SMITH [6] SUGIEREN:

A) SI SON p PARAMETROS LOS INVOLUCRADOS EN EL MODELO, ELEGIR p ECUACIONES DEL MODELO FUNCIONAL MUESTRAL, SUSTITUYENDO LOS RESPECTIVOS VALORES DEL VECTOR DE VARIABLES CONTROLADA $\xi$ E IGNORANDO EL ERROR EN CADA UNA DE ELLAS. POSTERIORMENTE, RESOLVER p ECUACIONES SIMULTANEAS PARA $\theta$.

EVIDENTEMENTE, NO SIEMPRE ES POSIBLE QUE ESTA OPCION TENGA EXITO PARA OBTENER $\theta^0$ DADA LA POSIBLE COMPLEJIDAD DE LAS ECUACIONES RESULTANTES, PERO EN MUCHOS CASOS PUEDEN HACERSE MANIPULACIONES CONVENIENTES PARA TRATAR DE HACER LOS DESPEJES NECESARIOS.

B) EN CIERTOS CASOS, UNA REVISION UNA DEL MODELO FUNCIONAL SIN CONSIDERAR AL ERROR, PUEDE SUGERIR UNA TRANSFORMACION CONVENIENTE PARA HACER UNA ESTIMACION INICIAL. POR EJEMPLO, SI EL MODELO ES $Y=\theta_1 exp(-\theta_2 t)+\varepsilon$, NO CONSIDERANDO AL ERROR Y APLICANDO log EN AMBOS MIEMBROS DE LA ECUACION TENDRIAMOS $\log Y=\log\theta_1-\theta_2 t$, EL PROBLEMA ENTONCES PODRIA SER CONSIDERADO COMO LINEAL Y PODRIAMOS OBTENER UN PUNTO CONVENIENTE $\theta^0=(\theta_1{}^0,\theta_2{}^0)$ COMO VECTOR INICIAL.

C) PODRIA ELEGIRSE UNA RED DE PUNTOS EN EL ESPACIO PARAMETRICO ($\theta$-ESPACIO) Y EVALUAR $S(\theta)$, LO QUE NOS DARIA UNA IDEA DE SU GRAFICA Y AL MISMO TIEMPO NOS SUGERIRIA UN VALOR INICIAL $\theta^0$.

D) PARA UN CASO PARTICULAR, EL VALOR DE $\theta^0$ PODRIA SER SUGERIDO POR UN INVESTIGADOR CON EXPERIENCIA EN PROBLEMAS DE ESE TIPO.

LOS METODOS DEL DESCENSO POR LA PENDIENTE MAXIMA Y EL DE MARQUARTDT PERTENECEN A LOS DENOMINADOS "METODOS ITERATIVOS ACEPTABLES". PARA SU EXPOSICION SEGUIREMOS EL ORDEN DESCRITO EN LA INTRODUCCION DE ESTE TRABAJO POR LAS RAZONES AHI MISMO REFERIDAS.

EN ESTE CAPITULO,ESTUDIAREMOS EL ESQUEMA GENERAL DE LA FAMILIA A LA QUE PERTENECEN LOS ALGORITMOS DEL DESCENSO POR LA PENDIENTE MAXIMA Y EL DE MARQUARDT. LOS METODOS QUE CONFORMAN ESTA FAMILIA SON CONOCIDOS BAJO EL TERMINO GENERICO DE "METODOS ITERATIVOS ACEPTABLES".

AUNQUE DAREMOS UNA PRUEBA DE CONVERGENCIA PARA EL CASO GENERAL DE ESTOS METODOS,EN EL CAPITULO 5 PROPORCIONAMOS UNA QUE SERA EXCLUSIVA PARA EL PROCEDIMIENTO QUE SE DESCRIBE EN EL MISMO, Y QUE PERTENECE A LA CLASE DE LOS ITERATIVOS ACEPTABLES.

LOS METODOS ITERATIVOS ACEPTABLES SON TALES QUE,PARA UNA FUNCION OBJETIVO,DIGAMOS $S(\theta)$,POSEEN LA PROPIEDAD SIGUIENTE:

(4.1) $\qquad S(\theta^{i+1}) < S(\theta^{i}) \qquad i=0,1,2,\ldots$

ESTOS METODOS TIENEN ADEMAS EL ESQUEMA SIGUIENTE:

(4.2) 1.- CON $i=0$,DEBE SUGERIRSE EXTERNAMENTE AL METODO UN PUNTO INICIAL $\theta^{0}$

2.- SE DETERMINA UNA DIRECCION $\nu_i$ SEGUN EL RESULTADO OBTENIDO EN LA i-ESIMA ITERACION

3.- SE DETERMINA UN ESCALAR $\rho_i$,TAL QUE EL PASO:

$$\sigma_i = \rho_i \nu_i$$

SEA ACEPTABLE;ESTO ES,SE TOMA:

$$\theta^{i+1} = \theta^{i} + \rho_i \nu_i$$

TAL QUE LA DESIGUALDAD (4.1) SE CUMPLA

4.- SE PRUEBA,CON UN CRITERIO PREVIAMENTE DEFINIDO,SI DEBE DETENERSE EL ALGORITMO (CONVERGENCIA);SI NO,SE SE INCREMENTA i Y SE REGRESA A 2.SI SE HA DETERMINADO CONVERGENCIA SE HACE: $\theta^{i+1} = \hat{\theta}$

TODOS LOS METODOS ACEPTABLES DIFIEREN EN LA FORMA EN QUE SON DEFINIDOS $\nu_i$ Y $\rho_i$.

CONSIDERESE LA i-ESIMA ITERACION DE UN PROCEDIMIENTO DE MINIMIZACION. SUPONGASE QUE PARTIENDO DE $\theta^i$ CONTINUAMOS A LO LARGO DE UNA DIRECCION $\nu$, GENERANDO UN RAYO DEFINIDO COMO:

$$(4.3) \qquad \theta(\rho) = \theta^i + \rho\nu \qquad ( \rho >= 0 )$$

ES EVIDENTE QUE A LO LARGO DE ESTE RAYO LA FUNCION OBJETIVO VARIA DE ACUERDO A LA VARIACION DE $\rho$ EXCLUSIVAMENTE, TENEMOS:

$$(4.4) \qquad \Psi_{i\,v}(\rho) = S(\theta(\rho)) = S(\theta^i + \rho\nu)$$

DERIVANDO CON RESPECTO A $\rho$ :

$$(4.5) \qquad \delta(\Psi_{i\,v})/\delta\rho = (\delta S/\delta\theta)^t \ \delta\theta/\delta\rho = (\delta S/\delta\theta)^t \ \nu$$

DENOTAREMOA $q(\theta)=\delta S/\delta\theta$ , ASI LA $\alpha$-ESIMA COMPONENTE DE $q(\theta)$ ES, $\delta S/\delta\theta_\alpha$ . DENOTAREMOS TAMBIEN $q_i = \delta S/\delta\theta|\theta=\theta^i$.

TENEMOS ENTONCES:

$$(4.6) \qquad \Psi'_{i\,v} = \delta(\Psi_{i\,v})/\delta\rho|\rho=o = (q_i)^t\nu$$

Y SUPONDREMOS QUE $\quad q_i \neq 0$

ENTONCES, $\Psi'_{i\,v}$ ES LA DERIVADA DIRECCIONAL DE S RELATIVA A $\nu$ EN $\theta^i$.

SI $\Psi'_{i\,v}$ ES NEGATIVA, ENTONCES $S(\theta)$ DECRECE EN VALOR CUANDO A PARTIR DE $\theta^i$ NOS MOVEMOS EN DIRECCION DE $\nu$ Y, SI $\rho$ ES UN NUMERO POSITIVO SUFICIENTEMENTE PEQUENO, EL PASO $\rho\nu$ ES ACEPTABLE PUES $S(\theta^i+\rho\nu) < S(\theta^i)$. SI POR EL CONTRARIO, $\Psi'_{i\,v}$ ES MAYOR O IGUAL A CERO, ENTONCES NO PUEDE EXISTIR NINGUN VALOR POSITIVO DE $\rho$ PARA EL QUE $\rho\nu$ SEA ACEPTABLE. LLAMAREMOS A $\nu$ UNA DIRECCION ACEPTABLE SI SUCEDE QUE $\quad \Psi'_{i\,v} < 0$.

EL SIGUIENTE RESULTADO PROPORCIONA UNA CONDICION NECESARIA Y SUFICIENTE PARA QUE UNA DIRECCION $\nu$ SEA ACEPTABLE.

TEOREMA 1

UNA DIRECCION $\nu$ ES ACEPTABLE SI Y SOLO SI EXISTE UNA MATRIZ POSITIVA DEFINIDA $\mathbb{R}$ TAL QUE:

$$\nu = -\mathbb{R}q_i$$

DEMOSTRACION:

$\Rightarrow$) SUPONGAMOS QUE EXISTE UNA MATRIZ POSITIVA DEFINIDA TAL QUE $\nu = -\mathbb{R}q_i$

ENTONCES, $(q_i)^t\mathbb{R}q_i > 0$, DE DONDE $-(q_i)^t\mathbb{R}q_i < 0$

AHORA BIEN, DE (4.6) SE TIENE:

$$\Psi'_{i\nu} = (q_i)^t\nu = (q_i)^t(-\mathbb{R}q_i) = -(q_i)^t\mathbb{R}q_i < 0$$

$$\Rightarrow \quad \Psi'_{i\nu} < 0$$

POR LO TANTO, $\nu$ ES ACEPTABLE .

$\Leftarrow$) SUPONGAMOS AHORA QUE $\nu$ ES ACEPTABLE, ENTONCES $\Psi'_{i\nu} < 0$ O BIEN, $(q_i)^t\,\nu < 0$.

DEFINIMOS:

$$\mathbb{R} = \{I-[q_i(q_i)^t/(q_i)^tq_i]-[\nu(\nu)^t/(\nu)^tq_i]\}$$

ENTONCES,

$$\mathbb{R}q_i = [q_i - q_i(q_i)^tq_i/(q_i)^tq_i - \nu(\nu)^tq_i/(\nu)^tq_i]$$
$$= -\nu$$

ASI,

$$\mathbb{R}q_i = -\nu$$

ADEMAS, UTILIZANDO LA IGUALDAD ANTERIOR Y LA HIPOTESIS,

$$(q_i)^t\mathbb{R}q_i = (q_i)^t(-\nu) = -(q_i)^t\nu > 0$$

POR TANTO, $\mathbb{R}$ ES POSITIVA DEFINIDA Y $-\mathbb{R}q_i=\nu$ COMO QUERIA MOSTRARSE.

UN METODO DE MINIMIZACION EN EL QUE LAS DIRECCIONES SON OBTENIDAS SEGUN LO SUGERIDO POR EL TEOREMA 1, ES CONOCIDO COMO UN "METODO GRADIENTE ACEPTABLE". CUANDO LA MATRIZ $\mathbb{R}$ NO ES POSITIVA DEFINIDA SE CONOCE SIMPLEMENTE COMO UN "METODO GRADIENTE".

LA ECUACION BASICA DE CUALQUIER METODO GRADIENTE EN LA i-ESIMA ITERACION ES:

$$(4.7) \qquad \theta^{i+1} = \theta^i - \rho_i\mathbb{R}q_i$$

CUANDO SE HA SELECCIONADO UN METODO DE MINIMIZACION, A CUALQUIERA LE GUSTARIA ESTAR EN POSIBILIDAD DE PROBAR QUE CONVERGE AL VERDADERO MINIMO DE LA FUNCION OBJETIVO. DESAFORTUNADAMENTE, LAS PRUEBAS DE CONVERGENCIA REQUIEREN DE CIERTAS SUPOSICIONES SOBRE LA NATURALEZA DE LA FUNCION OBJETIVO CUYA VALIDEZ ES DIFICIL DE VERIFICAR EN UN PROBLEMA DADO. MAS AUN, LA EXISTENCIA DE UNA PRUEBA DE CONVERGENCIA NO ES GARANTIA DE UN COMPORTAMIENTO DESEADO EN LA PRACTICA. UN METODO PUEDE CONVERGER EN TEORIA, PERO PARA UN PROBLEMA DADO, PUEDE REQUERIR DE UN NUMERO EXAGERADO DE ITERACIONES Y CALCULOS ANTES DE CONVERGER. LOS METODOS ITERATIVOS ACEPTABLES NO SON LA EXCEPCION, SIN EMBARGO PODEMOS DECIR LO SIGUIENTE:

SI $S^i$ DENOTA EL VALOR DE $S(\theta^i)$ Y SI EN CADA ITERACION SELECCIONAMOS UN PUNTO ACEPTABLE, ENTONCES LA SUCESION $\{S^i\}_{i=0}^{\infty} = \{S^0, S^1, S^2, \ldots,\}$ ES MONOTONA DECRECIENTE. AHORA BIEN, SI LOS VALORES DE LA FUNCION OBJETIVO POSEEN UNA COTA INFERIOR ENTONCES LA SUCESION DEBE CONVERGER A UN LIMITE $S^{\infty}$. SI LA SUCESION $\{\theta^i\}$ ESTA ACOTADA, ES DECIR EXISTE $M>0$ TAL QUE $(\theta^i)^t \theta^i < M \quad \forall i$, ENTONCES EXISTE AL MENOS UN PUNTO LIMITE DE LA MISMA. SE SIGUE ENTONCES QUE SI S ES CONTINUA $S(\theta^{\infty})=S^{\infty}$, DONDE $\theta^{\infty}$ ES EL PUNTO LIMITE DE $\{\theta^i\}$. PUEDE SUCEDER QUE LA RAZON DE CONVERGENCIA SEA MUY LENTA, DE TAL FORMA QUE PUDIESE PARECER QUE LA SUCESION NO CONVERGE.

UN PUNTO ESTACIONARIO DE LA FUNCION OBJETIVO ES AQUEL $\theta$ TAL QUE $q(\theta)=0$, ASI, SI $\theta^i$ ES ESTACIONARIO $q_i=0$ Y DE LA ECUACION (4.7) SE DESPRENDE QUE $\theta^i=\theta^j \quad \forall\ j>=i$. DE LO ANTERIOR SE CONCLUYE QUE LO MAS QUE PODEMOS PROBAR DE UN METODO GRADIENTE ES QUE CONVERGE A UN PUNTO ESTACIONARIO QUE BIEN PODRIA SER UN PUNTO SILLA. LA CONVERGENCIA AL VERDADERO MINIMO PUEDE SER GARANTIZADA SOLO SI PUEDE MOSTRARSE QUE LA FUNCION OBJETIVO NO TIENE OTROS PUNTOS ESTACIONARIOS. LA ESPERANZA ES QUE POR LO MENOS SE CONVERJA A UN MINIMO LOCAL Y QUE ESTE NO ESTE MUY ALEJADO DEL MINIMO O MINIMOS

EL SIGUIENTE TEOREMA FORMALIZA LO QUE HEMOS COMENTADO EN LOS PARRAFOS ANTERIORES.

TEOREMA 2

SUPONGAMOS QUE $S(\theta)$ TIENE PRIMERAS DERIVADAS CONTINUAS Y DIFERENCIABLES. SEA $S^0=S(\theta^0)$, Y SEA $\mathbb{D}$ EL CONJUNTO DE TODOS LOS PUNTOS $\theta$ TALES QUE $S(\theta) <= S^0$. DEFINIMOS LA SUCESION DE PUNTOS $\theta^1, \theta^2, \ldots$ POR:

$$\theta^{i+1} = \theta^i - \rho_i \mathbb{R}_i q_i$$

ASUMIMOS ADEMAS LO SIGUIENTE:

1.- EXISTE UN NUMERO M TAL QUE NINGUN EIGENVALOR DEL HESSIANO $H(\theta)$ EXCEDE A M EN VALOR ABSOLUTO $\forall\ \theta \in \mathbb{D}$

2.- TODAS LAS MATRICES $\mathbb{R}_i$ SON POSITIVAS DEFINIDAS Y SUS EIGENVALORES CAEN ENTRE DOS NUMEROS POSITIVOS $0 < \beta < \alpha$

3.- TODOS LOS $\rho_i$ SON ELEGIDOS DE TAL FORMA QUE:

$$MIN(\tau, \alpha\ \mu^i) <= \rho^i <= \mu^i$$

DONDE $\tau$ ES UNA CONSTANTE POSITIVA, $\alpha$ ES OTRA CONSTANTE QUE SATISFACE $0<\alpha<1$ Y $\mu_i$ ES EL MAS PEQUENO NO NEGATIVO VALOR DE $\rho$ TAL QUE $S(\theta^i - \rho\mathbb{R}_i q_i)$ ES UN PUNTO ESTACIONARIO. $S(\theta^i - \rho\mathbb{R}_i q_i)$ COMO FUNCION DE $\rho$.

SEA $\theta^*$ UN PUNTO LIMITE DE LA SUCESION $\{\theta^i\}$. ENTONCES $\theta^*$ ES UN PUNTO ESTACIONARIO DE S. ES DECIR $q^* = q(\theta^*) = 0$

COMENTARIO: TAL PUNTO LIMITE (NO NECESARIAMENTE UNICO) DEBE EXISTIR SI $\mathbb{D}$ ES ACOTADO.

DEMOSTRACION:

TENEMOS QUE LA SUCESION $\{S^i\}=\{S(\theta^i)\}$ ES MONOTONA DECRECIENTE DADA LA CONTINUIDAD DE S, TENEMOS TAMBIEN QUE:

$$S^* = S(\theta^*) <= S(\theta^i) \qquad \forall\ i=1,2,\ldots$$

SUPONGAMOS QUE $\theta^*$ NO ES UN PUNTO ESTACIONARIO DE S. ENTONCES, $\|q^*\| = a > 0$

POR LA CONTINUIDAD DE LA FUNCION q Y DADO QUE $\theta^*$ ES UN PUNTO LIMITE, PARA $\varepsilon=a/2$ DEBE EXISTIR $n_0 \in \mathbb{N}$ TAL QUE:

$$| \ \|q^*\| - \|q^{no}\| \ | < a/2$$

$$\Rightarrow \quad \|q^*\| - a/2 < \|q^{no}\|$$

(4.8) $\Rightarrow$ $a/2 < \|q^{no}\|$

ADEMAS $a/2 < 3a$ ,POR LO QUE:

(4.9) $\|q^{no}\| < 3a - \|q^*\| = 2a$

DE (4.8) Y (4.9) CONCLUIMOS QUE $\exists$ $n_o \in \mathbb{N}$ TAL QUE:

(4.10) $a/2 <= \|q^{no}\| <= 2a$

POR LA CONTINUIDAD DE LA FUNCION $S(\theta)$ Y DADO QUE $\theta^*$ ES UN PUNTO LIMITE, PARA:

$$\varepsilon = \text{MIN} \{ (2-\alpha)\alpha a^2\beta^2/256\gamma^2 M \ , \ a^2\beta^2/256\gamma^2 M \ , \ a^2\beta\tau/16 \} > 0$$

EXISTE $n_1 \in \mathbb{N}$ TAL QUE:

$$| S^{n1} - S^* | <= \varepsilon$$

(4.11) $\Rightarrow S^{n1} <= S^* + \text{MIN}\{(2-\alpha)\alpha a^2\beta^2/256\gamma^2 M, a^2\beta^2/256\gamma^2 M, a^2\beta\tau/16\}$

SEA $n_2 = \text{MAX} \{n_o, n_1\}$

(4.12) $\Rightarrow$ $a/2 <= \|q^{n2}\| <= 2a$

$$S^{n2} <= S^* = \text{MIN}\{(2-\alpha)\alpha a^2\beta^2/256\gamma^2 M, a^2\beta^2/256\gamma^2 M, a^2\beta\tau/16\}$$

CONSIDEREMOS LA FUNCION:

(4.13) $\Psi(\rho) = S(\theta^{n2} - \rho \, \mathbb{R}_{n2} \, q^{n2})$

TENEMOS ENTONCES QUE:

(4.14) $\delta\Psi/\delta\theta|\rho=o = (\delta S/\delta\theta) \ (\delta\theta/\delta\rho) = -(q^{n2})^t \ \mathbb{R}_{n2} \ q^{n2}$

(4.15) $\Rightarrow \delta\Psi/\delta\rho|\rho=o = -(q^{n2})^t \ \mathbb{R}_{n2} \ q^{n2} <= -\beta \ \|q^{n2}\|^2 <= -\beta a^2/4$

RESULTADO QUE SE SIGUE DEL APENDICE I Y DE (4.12) ADEMAS,

(4.16) $0 <= |\delta^2\Psi/\delta\rho^2| = |(\delta\theta/\delta\rho)^t \ (\delta^2 S/\delta\rho^2) \ (\delta\theta/\delta\rho)|$

$$<= M \ \| \mathbb{R}_{n2} q^{n2}\|^2 <= 4a^2\gamma^2 M$$

QUE TAMBIEN SE SIGUE DEL TEOREMA 1 DEL APENDICE I Y DE (4.12).

AHORA BIEN, DE (4.15) Y (4.16), PODEMOS CONCLUIR QUE :

(4.17) $\delta\Psi/\delta\rho <= -\beta a^2/4 + 4a^2\gamma^2 M\rho$

YA QUE SEGUN (4.16),

$$\delta^2\Psi/\delta\rho^2 <= 4a^2\gamma^2 M$$

ENTONCES, $\delta\Psi/\delta\rho <= 4a^2\gamma^2 M\rho + C$

ASI, $\delta\Psi/\delta\rho <= C$

Y POR (4.15) SABEMOS QUE $C=-\beta a^2/4$ CUMPLE ESTA CONDICION, POR LO QUE (4.17) ES VERDADERA.

SEA $\rho=\mu^{n0}$, ENTONCES $\delta\Psi/\delta\rho=0$, Y ASI DE (4.17) OBTENEMOS:

$$0 <= -\beta a^2/4 + 4a^2\gamma^2 M\mu^{n2}$$

(4.18) $\Rightarrow \beta/16\gamma^2 M <= \mu^{n2}$

$\Rightarrow \alpha\beta/16\gamma^2 M <= \alpha\mu^{n2}$

INTEGRANDO LA ECUACION (4.17) SE TIENE:

(4.19) $\Psi(\rho) <= -\beta a^2\rho/4 + 2a^2\gamma^2 M\rho^2 + C$

Y COMO $\Psi(0) = S^{n2}$, SE SIGUE DE (4.19) QUE:

$S^{n2} <= C$ , POR LO QUE :

(4.20) $\Psi(\rho) <= S_{n2} - \beta a^2\rho/4 + 2a^2\gamma^2 M\rho^2$

SUPONGAMOS QUE $\rho^{n2}$ HA SIDO ESCOGIDA DE TAL FORMA QUE $\alpha\mu^{n2} <= \rho^{n2} <= \mu^{n2}$, DADO QUE $\Psi(\rho)$ ES MONOTONA DECRECIENTE PARA $0 <= \rho <= \mu^{n2}$, TENEMOS DE LA ECUACION (4.20) QUE:

$S^{n2+1} = \Psi(\rho^{n2}) <= \Psi(\alpha\mu^{n2}) <= \Psi(\alpha\beta/16\gamma^2 M)$

$<= S^{n2} - (\beta a^2/4)(\alpha\beta/16\gamma^2 M) + 2a^2\gamma^2 M(\alpha\beta/16\gamma^2 M)^2$

(4.20) $= S^{n2} - [(2-\alpha)\alpha a^2\beta^2/128\gamma^2 M\ ]$

DE (4.12) OBTENEMOS QUE:

(4.21)
$$S^{n2} <= S^* + (2-\alpha)\alpha a^2\beta^2/256\gamma^2 M$$
$$S^{n2} - (2-\alpha)\alpha a^2\beta^2/128\gamma^2 M <= S^* - (2-\alpha)\alpha a^2\beta^2/256\gamma^2 M$$

UTILIZANDO (4.20) Y (4.21) CONCLUIMOS ENTONCES QUE:

$$S^{n2+1} <= S^* - (2-\alpha)\alpha a^2\beta^2/256\gamma^2 M < S \quad !$$

PUES $\{S^i\}$ ES CRECIENTE

LA OTRA ALTERNATIVA ES ELEGIR $\tau <= \rho^{n2} <= \mu^{n2}$, PERO CON UN RAZONAMIENTO SIMILAR SE TENDRIA:

(4.22) $S^{n2+1} <= S^* - \beta a^2\tau/4 + 2a^2\gamma^2 M\ \tau^2$

AQUI HABRIA DOS SUBCASOS:

a) $\tau <= \beta/16\gamma^2 M$ EN CUYO CASO,

$$S^{n2+1} <= S^{n2} - a^2\tau(\beta/4 - 2\gamma^2 M\tau)$$

$$<= S^{n2} - a^2\tau\ [\beta/4 - 2\gamma^2 M(\beta/16\gamma^2 M)]$$
$$= S^{n2} - a^2\beta\tau/8 <= S^* - a^2\beta\tau/16 < S^* \quad !$$

b) $\beta/16\gamma^2 M <= \tau <= \rho^{n2} <= \mu^{n2}$ ENTONCES,

$$S^{n2+1} = \Psi(\rho^{n2}) <= \Psi(\tau) <= \Psi(\beta/16\gamma^2 M)$$
$$<= S^{n2} - (\beta a^2/4)(\beta/16\gamma^2 M) + 2a^2\gamma^2 M(\beta/16\gamma^2 M)^2$$
$$= S^{n2} - (a^2\beta^2/128\gamma^2 M) <= S^* - (a^2\beta^2/256\gamma^2 M) < S^* \quad !$$

POR LO TANTO $q^*=0$

POR LO TANTO $\theta^*$ ES UN PUNTO ESTACIONARIO DE $S(\theta)$ L.Q.Q.D.

LA IDEA DE ESTE METODO ES TRABAJAR CON LA FUNCION $S(\theta)$,E IR AVANZANDO HACIA EL VALOR DE $\theta$ QUE LA MINIMICE,EN DIRECCION DEL -GRAD $S(\theta)$ Y PARTIENDO DE UN PUNTO INICIAL $\theta^0$.

DEFINAMOS ENTONCES:

$\theta^0=(\theta_1{}^0,\theta_2{}^0,\ldots,\theta_p{}^0)$ ,UN PUNTO EN EL ESPACIO PARAMETRICO $\mathbb{R}^p$

(5.1) $Z^0=(Z_1{}^0,Z_2{}^0,\ldots,Z_p)$

CON $Z_i{}^0= -\lambda(\delta S/\delta\theta_i)|\theta = \theta^0$

DE TAL FORMA QUE:

(5.2) $Z^0 = -\lambda\mathrm{GRAD}(S(\theta))$ EVALUADO EN $\theta = \theta^0$

$\lambda$,ES UN FACTOR POSITIVO DE PROPORCIONALIDAD.

(POR EJ., $\lambda > 0$ TAL QUE $\|\lambda Z^0\| = 1$)

AHORA,PARA DOS VECTORES $\Omega$ Y $\Delta$ EN $\mathbb{R}^p$ DEFINAMOS LA FUNCION:

(5.3) $g(t) = S(\Omega +t\Delta)$ ; $t\in\mathbb{R}$

EN ESTAS CONDICIONES,

(5.4) $dg(t)/dt = [\mathrm{GRAD}\ S|\Omega+t\Delta]\cdot \Delta$

SI CONSIDERAMOS EN LA ECUACION (5.3) A LOS VECTORS $\theta^0$ Y $Z^0$ TENEMOS QUE:

(5.5) $g(t) = S(\theta^0+tZ^0)$ $t\in\mathbb{R}$

OBSERVEMOS QUE ESTA FUNCION $g$ TOMA LOS VALORES DE LA FUNCION S A LO LARGO DE SU VECTOR GRADIENTE DESDE EL PUNTO $\theta^0$

DERIVANDO LA FUNCION (5.5) CON RESPECTO A $t$,OBTENEMOS:

(5.6) $dg(t)/dt = [\mathrm{GRAD}\ S|\theta=\theta^0+tZ^0]\cdot Z^0$

Y AL EVALUAR ESTA ECUACION EN $t=0$ RESULTA:

$$dg(0)/dt = [\mathrm{GRAD}\ S|\theta=\theta^0]\cdot Z^0$$
$$= (\delta S/\delta\theta_1|\theta=\theta^0,\delta S/\delta\theta_2|\theta=\theta^0,\ldots,\delta S/\delta\theta_P|\theta=\theta^0)\cdot$$
$$\cdot(-\lambda\delta S/\delta\theta_1|\theta=\theta^0,-\lambda\delta S/\delta\theta_2|\theta=\theta^0,\ldots,-\lambda\delta S/\delta\theta_p|\theta=\theta^0)$$

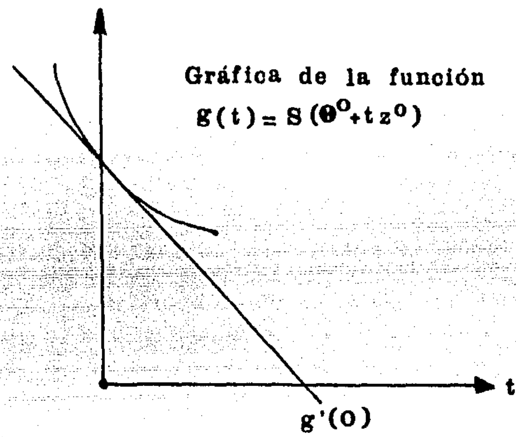
POR LO TANTO:

(5.7) $dg(0)/dt = -\lambda\ \Sigma_{i=1}{}^p(\delta S/\delta\theta_i\ |\theta=\theta^0)^2 < 0$ ,PUES $\lambda > 0$

ASI, LA DERIVADA DE LA FUNCION $g(t)$ EN $t=0$ RESULTA SER NEGATIVA, ES DECIR, LA FUNCION ES DECRECIENTE EN $t=0$, LUEGO ENTONCES DEBE EXISTIR $t > 0$ TAL QUE :

(5.8) $\qquad g(t) < g(0)$

GEOMETRICAMENTE, PARA $\theta^0$ Y $Z^0$ LA FUNCION $g(t)$ CERCA DEL CERO DEBE TENER UNA GRAFICA SIMILAR A LA SIGUIENTE:




CON LA $t$ QUE CUMPLE LA CONDICION (5.8), PODEMOS DEFINIR:

(5.9) $\qquad \theta^1 = \theta^0 + tZ^0$

COMO NUESTRO NUEVO PUNTO PARA COMENZAR LA ITERACION.

CABE HACER NOTAR QUE :

$$g(t) < g(0)$$

$$\Rightarrow S(\theta^0+tZ^0) = g(t) < g(0) = S(\theta^0)$$

$$\Rightarrow \qquad S(\theta^0+tZ^0) < S(\theta^0)$$

Y COMO $\theta^1 = \theta^0 + tZ^0$, SE TIENE:

(5.10) $\qquad S(\theta^1) < S(\theta^0)$

DE ESTA MANERA, SIGUIENDO EL PROCESO ITERATIVO OBTENDRIAMOS UNA SUCESION DE PUNTOS $\theta^0, \theta^1, \ldots$ TALES QUE:

(5.11) $\qquad S(\theta^{k+1}) < S(\theta^k)$

Y ASI, BAJO CIERTAS CONDICIONES ADECUADAS ESTA SUCESION TENDERIA A UN PUNTO ESTACIONARIO DE LA FUNCION S, MISMO QUE TOMARIAMOS COMO $\hat{\theta}$.

UNA PREGUNTA RELEVANTE ES: COMO ENCONTRAR $t > 0$ TAL QUE $g(t) < g(0)$ ? A CONTINUACION PROPONDREMOS UN METODO QUE PUEDE RESOLVER ESTA CUESTION.

LA DETERMINACION DE LA $t$ MENCIONADA PUEDE REALIZARSE POR MEDIO DE ENSAYOS, PODEMOS OBSERVAR PRIMERO EL VALOR DE LA INTERSECCION DE $g'(0)$ CON EL EJE DE LAS $t$'s ,SI ESTE NO SATISFACE LA DESIGUALDAD $g(t) < g(0)$ ENTONCES ES DEMASIADO GRANDE, PODEMOS ENTONCES TOMAR LA MITAD DE ESTE VALOR Y OBSERVAR LA RELACION $g(t), g(0)$ Y ASI SUCESIVAMENTE.

UNA CARACTERISTICA IMPORTANTE DE LO ANTERIOR, ES QUE PODEMOS DIBUJAR UN ESBOZO DE LA GRAFICA DE $g(t)$ Y DESPUES DE ALGUNOS ENSAYOS NORMALMENTE ES POSIBLE LOCALIZAR UNA $t$ QUE ESTE CERACNA A UN MINIMO DE $g(t)$.

EN EL CASO QUE TOMARAMOS $t$ PRECISAMENTE LA RAIZ POSITIVA MAS PEQUENA DE LA ECUACION $g'(t) = 0$, EL PROCESO TENDRIA LA SIGUIENTE INTERPRETACION GEOMETRICA:

INICIAMOS EN UN PUNTO $\theta^0$ Y DETERMINAMOS LA DIRECCION, A PARTIR DE $\theta^0$, EN LA CUAL LA FUNCION $S(\theta)$ DECRECE MAS RAPIDAMENTE, QUE SEGUN LA TEORIA DEL CALCULO DIFERENCIAL ESTA DADA POR $-\mathrm{GRAD}S(\theta^0)$, CONTINUAMOS EN ESA DIRECCION HASTA ENCONTRAR UN PUNTO SOBRE LA RECTA $\theta^0+tZ^0$ DONDE LA FUNCION $S(\theta)$ TOMA UN MINIMO, ENTONCES PARAMOS Y TOMAMOS OTRA DIRECCION DEL METODO DE LA PENDIENTE MAXIMA Y CONTINUAMOS.

ES IMPORTANTE DESTACAR QUE LAS DIRECCIONES DE $Z^K$ Y $Z^{K+1}$ SON ORTOGONALES, LO QUE PUEDE SER MOSTRADO DE LA SIGUIENTE MANERA:

DEFINIMOS, $\quad F(t) = \theta^k + tZ^k \;;\;\; F'(t) = Z^k$

$$\Rightarrow \; g(t) = S(\theta^k+tZ^k) = S(F(t)) = (S\circ F)(t)$$

$$\Rightarrow g'(t^k) = 0 = (S\circ F)'(t^k) = \mathrm{GRAD}\; S(F(t^k)) \cdot Z^k$$

$$\Rightarrow \qquad 0 = \mathrm{GRAD}\; S(t^{k+1}) \cdot Z^k$$

POR LO TANTO:

$$Z^k \perp Z^{k+1} \qquad \text{L.Q.Q.D.}$$

EN CUANTO A LA CONVERGENCIA DE ESTE METODO, A CONTINUACION PROPONEMOS UNA DEMOSTRACION DE LA MISMA.

SUPONGAMOS QUE LA FUNCION $S(\theta_1,\theta_2,\dots,\theta_p)$ ESTA DEFINIDA Y QUE $\delta S/\delta\theta_i$ $(i=1,2,\dots,p)$ , SON CONTINUAS EN LOS PUNTOS DE $\mathbb{D}^0$ Y EN $Fr(\mathbb{D})$, DONDE $\mathbb{D}$ ES UNA REGION. LLAMEMOS $\mathfrak{C}$ A LA LINEA QUEBRADA QUE COMIENZA EN $\theta^0$ EN DIRECCION DE LA PENDIENTE MAXIMA DE DESCENSO HASTA $Fr(\mathbb{D})$ O BIEN, EL PUNTO $\theta^1$ DEFINIDO POR LA MINIMA RAIZ $t$ DE : $g'(t) = GRAD\ S(\theta^0+tZ^0)\cdot Z^0$ ; SI ESTO ULTIMO SUCEDE, LA DIRECCION DE LA PENDIENTE MAXIMA DE DESCENSO DESCIENDE A PARTIR DE $\theta^1$ HASTA ENCONTRAR LA $Fr(\mathbb{D})$ O BIEN LA SIGUIENTE APROXIMACION $\theta^2$ DETERMINADA DE MANERA ANALOGA Y ASI SUCESIVAMENTE.

LA FUNCION $S(\theta)$ ES ENTONCES ESTRICTAMENTE DECRECIENTE A LO LARGO DE $\mathfrak{C}$, EXISTIENDO TRES POSIBILIDADES:

a) EL CAMINO $\mathfrak{C}$ PUEDE IR A LO LARGO DE LA $Fr(\mathbb{D})$,

b) EL CAMINO $\mathfrak{C}$ PUEDE TERMINAR EN UN PUNTO ESTACIONARIO DE LA FUNCION $S(\theta)$ ES DECIR, DONDE LA DIRECCION DEL METODO DE LA PENDIENTE MAXIMA NO EXISTE,

c) EL PROCESO PODRIA CONTINUAR INDEFINIDAMENTE.

AHORA BIEN, LA PRIMERA POSIBILIDAD QUEDARIA EXCLUIDA EN EL CASO DE QUE $S(\theta^0) < S(\theta)\quad \forall\ \theta \in Fr(\mathbb{D})$. EN CUANTO A LA SEGUNDA POSIBILIDAD, ESTA CONSTITUIRIA EL CASO TRIVIAL ES DECIR, CON ESTA TENDRIAMOS QUE EL METODO CONVERGE. A CONTINUACION MOSTRAREMOS QUE LA TERCERA POSIBILIDAD NO PUEDE DARSE.

DADO QUE $\bar{\mathbb{D}}$ ES UN COMPACTO LA SUCESION $\theta^0,\theta^1,\dots$ TIENE UNA SUBSUCESION QUE CONVERGE; SEA $\theta^\infty$ UN PUNTO LIMITE DE DICHA SUCESION, LO QUE ES CIERTO ES QUE:

$$(5.12)\qquad S(\theta^\infty) < S(\theta^k)\qquad k=1,2,\dots$$

VAMOS A DEMOSTRAR QUE $\theta^\infty$ ES UN PUNTO ESTACIONARIO DE LA FUNCION S.

SUPONGAMOS QUE SUCEDE LO CONTRARIO; DEFINIMOS:

$$(5.13)\qquad \begin{aligned} s(\theta) &= \| \text{GRAD } S(\theta) \| \\ z(\theta) &= - \text{GRAD } S(\theta)/s(\theta) \end{aligned}$$

DADO QUE $\theta^{\infty}$ NO ES UN PUNTO ESTACIONARIO DE S, ENTONCES:

$$(5.14)\qquad s(\theta^{\infty}) > 0$$

POR LO QUE PODEMOS CONSTRUIR UNA VECINDAD $U$ DE $\theta^{\infty}$ TAL QUE:

$$(5.15)\qquad \| \text{GRAD } S(\theta) - \text{GRAD } S(\theta^{\infty}) \| < \varepsilon\, s(\theta^{\infty}) \quad \forall\, \theta \in U$$

CON, $\varepsilon > 0$

COMO CONSECUENCIA DE LA DESIGUALDAD DEL TRIANGULO, OBTENEMOS DE (5.15) LO SIGUIENTE:

$$(5.16)\qquad \begin{aligned} &\big|\, \|\text{GRAD } S(\theta)\| - \|\text{GRAD } S(\theta^{\infty})\| \,\big| <= \|\text{GRAD } S(\theta) - \text{GRAD } S(\theta^{\infty})\| \\ &\qquad < \varepsilon\, s(\theta^{\infty}) \\ &\Rightarrow \quad |s(\theta) - s(\theta^{\infty})| < \varepsilon\, s(\theta^{\infty}) \end{aligned}$$

ADEMAS:

$$\begin{aligned}
{}^{*}\ &\|\text{GR } S(\theta)/s(\theta) - \text{GR } S(\theta^{\infty})/s(\theta^{\infty})\| = \\
&= \| \text{GR } S(\theta)\ s(\theta^{\infty}) - \text{GR } S(\theta^{\infty})\ s(\theta)\| \,/\, s(\theta)s(\theta^{\infty}) = \\
&= \| \text{GR } S(\theta)\ s(\theta^{\infty}) - \text{GR } S(\theta)\ s(\theta) + \text{GR } S(\theta)\ s(\theta) - \\
&\qquad - \text{GR } S(\theta^{\infty})\ s(\theta)\| \,/\, s(\theta)s(\theta^{\infty}) <= \\
&<= \| \text{GR } S(\theta)\ (\ s(\theta^{\infty}) - s(\theta)\ )\| \,/\, s(\theta)s(\theta^{\infty}) + \\
&\qquad + \| s(\theta)\ (\text{GR } S(\theta) - \text{GR } S(\theta^{\infty})\ \| \,/\, s(\theta)s(\theta^{\infty}) = \\
&= \| \text{GR } S(\theta)\|\ |s(\theta^{\infty}) - s(\theta)| \,/\, s(\theta)s(\theta^{\infty}) + \\
&\qquad + |s(\theta)|\ \| \text{GR } S(\theta) - \text{GR } S(\theta^{\infty})\| \,/\, s(\theta)s(\theta^{\infty}) = \\
&= s(\theta)\ |s(\theta^{\infty}) - s(\theta)| \,/\, s(\theta)s(\theta^{\infty}) + \\
&\qquad + s(\theta)\ \| \text{GR } S(\theta) - \text{GR } S(\theta^{\infty})\| \,/\, s(\theta)s(\theta^{\infty}) = \\
&= \{\ |s(\theta^{\infty}) - s(\theta)| + \| \text{GR } S(\theta) - \text{GR } S(\theta^{\infty})\ \|\ \}/s(\theta^{\infty})
\end{aligned}$$

AHORA BIEN, POR (5.15) Y (5.16) TENEMOS QUE:

$$|s(\theta^{\infty}) - s(\theta)| + \| \text{GR } S(\theta) - \text{GR } S(\theta^{\infty})\| < 2\varepsilon\ s(\theta^{\infty})$$

POR LO QUE:

$$\| \text{GR } S(\theta)/s(\theta) - \text{GR } S(\theta^{\infty})/s(\theta^{\infty})\| < 2\varepsilon\ s(\theta^{\infty})/s(\theta^{\infty}) = 2\varepsilon$$

Y COMO, $\| \text{GR } S(\theta)/s(\theta) - \text{GR } S(\theta^{\infty})/s(\theta^{\infty})\| = \| z(\theta) - z(\theta^{\infty})\|$

$$= \| z - z^{\infty}\|$$

* USAMOS AQUI GR $S$ PARA DESIGNAR GRAD $S$

CONCLUIMOS QUE $\forall\ \theta \in U$ SUCEDE:

(5.17) $\quad \| z - z^{\infty}\| < 2\varepsilon$

POR OTRO LADO, SI $\theta^{k} \in U$ , $\theta^{k+1} \notin U$ , SUJETO A QUE $\varepsilon$ SEA SUFICIENTEMENTE PEQUENO, PUES DE LO CONTRARIO GRAD $S(\theta^{k})$ Y $S(\theta^{k})$ ESTARIAN MUY CERCANOS Y POR CONSECUENCIA, GRAD $S(\theta^{k})/s(\theta^{k})$ Y GRAD $S(\theta^{k+1})/s(\theta^{k+1})$ LO ESTARIAN, CONTRADICIENDO ENTONCES LO MOSTRADO ANTERIORMENTE, ES DECIR QUE ESTAS DOS ULTIMAS DIRECCIONES SON ORTOGONALES.

CONSIDEREMOS AHORA, UN SECTOR CANONICO $W$ DE $U$ CUYO ANGULO DE APERTURA SEA TAL QUE:

(5.18) $\quad \mathrm{COS}(\varphi) = (\theta - \theta^{\infty})/\|\theta - \theta^{\infty}\| \cdot z^{\infty} > \varepsilon \quad \forall\ \theta \in W$

PODEMOS MOSTRAR QUE $\forall\ \theta \in W$ ,

(5.19) $\quad S(\theta) < S(\theta^{\infty})$

POR MEDIO DEL SIGUIENTE RAZONAMIENTO:

SUPONIENDO QUE SUCEDE LO CONTRARIO A LO ESTIPULADO EN (5.19), EXISTIRIA ENTONCES UNA SUCESION $\{\theta^{n}\} \to \theta^{\infty}$ CON $\theta^{n} \neq \theta^{\infty}$ Y $S(\theta^{n}) > S(\theta^{\infty})\ \forall\ n$, ADEMAS EL VECTOR $(\theta^{n}-\theta^{\infty})/\|\theta^{n}-\theta^{\infty}\|$ ESTARIA CONVERGIENDO A UN VECTOR UNITARIO $u^{0} \in W$ , SE TIENE ENTONCES:

$$0 <= \lim_{n \to \infty}\ [S(\theta^{n}) - S(\theta^{\infty})]/\|\theta^{n}-\theta^{\infty}\| =$$

$$= \{S\ [\theta^{\infty} + \|\theta^{n}-\theta^{\infty}\|\ (\theta^{n}-\theta^{\infty})/\|\theta^{n}-\theta^{\infty}\|\ ] - S(\theta^{\infty})\}\ /\ \|\theta^{n}-\theta^{\infty}\|$$

$$\Rightarrow 0 <= \lim_{t \to 0}\ [S(\theta^{\infty}+tu^{0}) - S(\theta^{\infty})]/t = \mathrm{GRAD}_{u^{0}}\ S(\theta^{\infty})$$

$$= \mathrm{GRAD}\ S(\theta^{\infty})\ \ u^{0}$$

$$= -s(\theta^{\infty})z^{\infty}\ \ u^{0}$$

Y DADO QUE $u^{0} \in W$ ENTONCES, $z^{\infty}\ u^{0} > \varepsilon > 0$ , $-s(\theta^{\infty})z^{\infty}u^{0} < 0$

SE TIENE POR TANTO:

$$0 <= \lim_{n \to \infty}\ [S(\theta^{n})-S(\theta^{\infty})]/\|\theta^{n}-\theta^{\infty}\| = -s(\theta^{\infty})z^{\infty}\ u^{0} < 0$$

$$\Rightarrow 0 < 0 \quad !$$

POR TANTO:

$$S(\theta) < S(\theta^{\infty}) \quad \forall\ \theta \in W$$

AHORA BIEN, POR LA CONTINUIDAD DE LA FUNCION (GRAD S) PODEMOS ASEGURAR QUE EXISTE UNA SUBVECINDAD $\mathcal{V}$ DE $\mathcal{W}$ TAL QUE $\forall\ \theta \in \mathcal{V}$ EL RAYO EN LA DIRECCION DE $z$ INTERSECTA A $\mathcal{W}$. PUESTO QUE $\theta^{\infty}$ ES UN PUNTO LIMITE EXISTE $N \in \mathbb{N}$ TAL QUE $\theta^{N} \in \mathcal{V}$, DE TAL FORMA QUE EL RAYO EN DIRECCION DE $z^{N}$ TIENE UN VECTOR $\theta^{v} \in \mathcal{W}$, ADEMAS $\theta^{v}$ NO PUEDE IR MAS ALLA DE $\theta^{N+1}$, PUESTO QUE $\theta^{N+1} \notin \mathcal{V}$. CON TODO ESTO Y DADO QUE LA FUNCION S ES MONOTONA DECRECIENTE EN $\mathfrak{C}$ SE SIGUE:

$$S(\theta^{N+1}) < S(\theta^{v}) < S(\theta^{\infty}) \quad !$$

DADO QUE SEGUN (5.12) ES NO PUEDE DARSE

POR LO TANTO:

$$s(\theta^{\infty}) = 0$$

POR LO TANTO:

$\theta^{\infty}$ ES UN PUNTO ESTACIONARIO DE S L.Q.Q.D.

POR OTRO LADO, ESTE METODO TAMBIEN PRESENTA DESVENTAJAS Y LOS AUTORES SEÑALAN COMO LA MAS RELEVANTE A LA SIGUIENTE:

- AUNQUE TEORICAMENTE EL PROCEDIMIENTO CONVERGE, EN LA PRACTICA PUEDE SUCEDER QUE LA CONVERGENCIA SEA MUY LENTA. NO OBSTANTE QUE EN LAS PRIMERAS ITERACIONES SE OBSERVEN PROGRESOS RAPIDOS, EN LAS ITERACIONES POSTERIORES PODRIAN OBSERVARSE SOLO PEQUENAS REDUCCIONES EN EL VALOR DE LA FUNCION $S(\theta)$.

BASANDOSE EN LAS VENTAJAS Y DESVENTAJAS DE LOS METODOS DE LINEALIZACION Y DEL DESCENSO POR LA PENDIENTE MAXIMA, MARQUARDT HACE UN ANALISIS DEL PROBLEMA DESDE SU FUNDAMENTO.

REDEFINIENDO COMO VECTOR DE CORRECCION ASOCIADO AL METODO DE LINEALIZACION $\delta_t$, AL DEFINIDO COMO $\mathbb{B}$ EN LA ECUACION (3.6), Y VECTOR DE CORRECCION ASOCIADO AL METODO DEL DESCENSO POR LA PENDIENTE MAXIMA $\delta_g$, AL MENOS GRADIENTE DE LA FUNCION $S(\theta)$, MARQUARDT SE PERCATA DE DOS PUNTOS IMPORTANTES, A SABER:

a) CUALQUIER METODO BASADO EN LA MINIMIZACION DE LA FUNCION $S(\theta)$, DEBERA TENER UN VECTOR ASOCIADO CUYA DIRECCION ESTE DENTRO DE UN RANGO $[0^{o},90^{o}]$ DEL $-$GRAD $S(\theta)$. DE OTRA FORMA, LOS VALORES DE $S(\theta)$ PODRAN TENDER A CRECER MAS QUE A REDUCIRSE SOBRE DICHO VECTOR DE CORRECCION.

b) DEBIDO A LA FORMA DE LA SUPERFICIE $S(\theta)$, $\delta_t$ ESTA A CASI $90^{o}$ DE $\delta_g$. DE HECHO EN UN GRAN NUMERO DE EXPERIMENTOS, MARQUARDT ENCONTRO QUE EL ANGULO $\Gamma$ ENTRE $\delta_t$ Y $\delta_g$ ERA TAL QUE: $80^{o} < \Gamma < 90^{o}$

CON ESTAS OBSERVACIONES, MARQUARDT CONCLUYE QUE CUALQUIER METODO QUE MEJORE TANTO AL DE LINEALIZACION COMO AL DE DESCENSO POR LA PENDIENTE MAXIMA, DEBERA TENER UN VECTOR DE CORRECCION QUE DE ALGUNA MANERA INTERPOLE ENTRE $\delta_t$ Y $\delta_g$.

ANTES DE EXPLICAR EL ALGORITMO DE MARQUARDT, PLANTEAREMOS ALGUNOS RESULTADOS QUE CONFORMAN LAS BASES TECNICAS DEL MISMO.

TEOREMA 1

SEA $\lambda >= 0$ Y SEA $\partial_o$ UN VECTOR QUE SATISFACE LA ECUACION:

(6.1) $\quad (Z^t Z + \lambda I)\delta_o = \mathbf{g}$

DONDE $Z$ ES LA MATRIZ DEFINIDA COMO EN (3.6) Y,

$$\mathbf{g} = [\Sigma_{u=1}{}^{n} \; (Y_u - f_u) \; \delta f_u / \delta\theta_j \;] \qquad j=1,2,\ldots,p$$

ENTONCES, $\delta_o$ MINIMIZA $S'(\theta)$ COMO FUE DEFINIDA EN (3.8). LO

En caso de tener una renta anual "Ra" y una tasa efectiva anual "i", el valor presente de la anualidad decreciente ordinaria vencida denotado por "DA" será:

$$DA = Ra(n - a_{\overline{n|}}i)/i \quad . \; . \; . \quad (2.42)$$

Si tuvieramos una tasa nominal "i'", la ecuación anterior se convierte en:

$$DA = (Ra/p)(mn - a_{\overline{mn|}}i')/i' \quad . \; . \; . \quad (2.43)$$

Donde:

$i' = (i\^m)/m$

m: (m = p) convertibilidad de la tasa nominal

iii) Montos de Anualidades Anticipadas Crecientes o Decrecientes.

Basémonos en una anualidad anticipada con "n" pagos periódicos de $P, P + Q, P + 2Q, \; . \; . \; . \;, P + (n-2)Q, P + (n-1)Q$, para determinar el monto de esta anualidad anticipada, tomamos como punto de valuación el período "n", resultando la siguiente ecuación de valor:

```
  PAGO
     P    P + Q  . . .     P+(n-1)Q P+nQ
     !-----!-----!-----!-----!-----!-----* P. de Valuación
  PERIODO
     0     1     2   . . .       n-1     n
```

Sea X el monto valuado a una tasa de interés por período "i", teniendo la siguiente igualdad:

$$X = (P+(n-1)Q)(1+i)+(P+(n-2)Q)(1+i)^2+. \; . \; .+(P+Q)(1+i)^{(n-1)}+P(1+i)^n \quad (2.44)$$

Multiplicando la ecuación anterior por (1+i)

$$(1+i)X=(P+(n-1)Q)(1+i)^2+(P+(n-2)Q)(1+i)^3+.+(P+Q)(1+i)^n+P(1+i)^{(n+1)} \quad (2.45)$$

7. Pagos Anticipados.

El acreditado podrá efectuar pagos anticipados de capital durante toda la vigencia del crédito, ajustandose a lo siguiente: a) cada pago anticipado deberá ser cuando menos por el equivalente a 10 erogaciones netas; y b) el pago anticipado habrá de efectuarse el día en que deba pagarse la erogación neta.

Todo pago anticipado se aplicara a reducir el saldo insoluto del crédito.

8. Seguros.

La institución acreedora podra cobrar mensualmente al acreditado, por concepto de los seguros relativos al crédito y a la vivienda de que se trate, intereses en adición a los previstos en el punto 5 anterior, a una tasa máxima del 0.75% anual sobre el saldo insoluto del crédito. Esta tasa de interés adicional no deberá considerarse para determinación de las tasas señaladas en ese punto 5.

Todas las cantidades que el acreditado debe cubrir a la institución acreedora por concepto de los seguros relativos al credito y a la vivienda de que se trate, no serán objeto de financiamiento.

9. Consultas.

Para la resolución de cualquier consulta relacionada con este nuevo regimen, los interesados podrán acudir a la institución de crédito de su preferencia o bien escribir a la oficina de estudios, autorizaciones y consultas de Banca Central del Banco de Mexico.

---

Por ser una vivienda de tipo 4 situada en la ciudad de Queretaro, le corresponde la Zona II. El salario mínimo mensual al 1o. de febrero de 1984 fue de $ 20,400.00, el factor que se le aplicó de 1.90, por lo tanto, la renta mensual que pago durante 1985 fue de $ 38,760.00. El interés que se le aplicó fue de 1.90, por lo tanto la renta mensual que pago durante 1985 fue de $ 38,760.00, el interés que se le aplicó en ese año fue del 25 % anual.

Para el período de 1985 el incremento del salario mínimo fue del 56 %, el 70 % de ese incremento es el 39.2 %, por lo tanto la renta mensual para 1986 fue de $ 53,953.92, el interés que se aplicó a la deuda fue del 33.4 %, debido a que el 15 % del aumento del salario mínimo es 8.4 % el cual se sumó al 25 % anterior.

## RESERVA DE PRIMAS NETAS PERIODICAS DE SEGUROS BASICOS

**************************************************

ANTES DE PROCEDER A REALIZAR EL CALCULO DE LA RESERVA ES CONVENIENTE HACER LA DEFINICION DE LO QUE SE ENTIENDE POR "RESERVA".

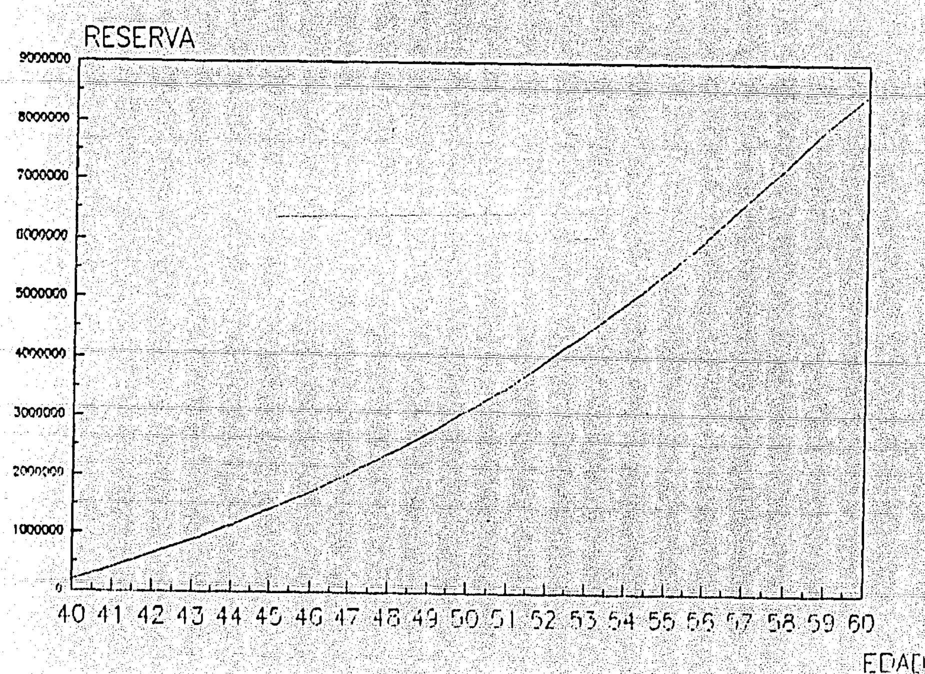
EN BASE A LA SUPOSICION DE UNA TASA DE MORTALIDAD , UN INTERES FIJO Y DEPENDIENDO DEL TIPO DE SEGURO , ES COMO SE LLEVA A CABO EL CALCULO DE LAS PRIMAS NETAS PERIODICAS, LAS CUALES EN CONJUNTO DEBEN SER SUFICIENTES PARA HACER FRENTE A TODAS LAS RECLAMACIONES POR MUERTE QUE VAYAN SURGIENDO.

SE HA VISTO YA COMO DETERMINAR, PARA SEGUROS BASICOS, LA PRIMA NETA UNICA; SIN EMBARGO, DADA UNA SUMA ASEGURADA Y UNA EDAD DE ADQUISICION DEL SEGURO, LA PRIMA NETA UNICA PUEDE RESULTAR (Y DE HECHO LO ES) DEMASIADO GRANDE DE TAL FORMA QUE SU PAGO AFECTARIA LA ECONOMIA DEL ASEGURADO.

SE PODRIA PENSAR EN CAMBIAR LA COBERTURA DEL SEGURO, POR UNA SERIE DE COBERTURAS TEMPORALES A UN AÑO Y PAGAR AL INICIO DE CADA PERIODO LA PRIMA CORRESPONDIENTE A LA COBERTURA ANUAL. ASI UN SEGURO TEMPORAL A 10 AÑOS PODRIA SER SUSTITUIDO POR DIEZ POLIZAS DE SEGUROS TEMPORALES A UN AÑO, ADQUIRIDAS CADA UNA AL INICIO DE CADA AÑO SUCESIVO.

EN EL ANALISIS GRAFICO PARA UN SEGURO DOTAL ENTRARAN EN CONSIDERACION QUE LA RESERVA ES GENERADA DE DOS FUENTES, UNA PROVENIENTE DE UN SEGURO TEMPORAL, (PARA LA CUAL YA SE REALIZO EL ANALISIS) Y LA OTRA DE UN SEGURO DOTAL PURO EL CUAL FUNCIONA COMO UN FONDO DE AHORRO. DE IGUAL MANERA QUE EL ANTERIOR HACIENDO UN EJERCICIO BASADO EN LA TABLA DE EXPERIENCIA MEXICANA (62-67) CON UN INTERES DEL 8.5 % PARA UN SEGURO DOTAL MIXTO EMITIDO A EDAD 40 CON PERIODO DE COBERTURA DE 20 AÑOS Y CON PAGO DE PRIMAS IGUAL AL PERIODO DE COBERTURA **(VER TABLA NO. 2 DEL APENDICE)** TENDREMOS LA SIGUIENTE GRAFICA.

# RESERVA DOTAL



EN EL CASO DOS ESTOS ESTAN SUJETOS A UN NUMERO ESTIMADO DE SEGUROS VENDIDOS PARA OBTENER LOS GASTOS PROPORCIONALES POR CADA UNIDAD DE PRIMA PAGADA.

EN LA PRACTICA VEAMOS COMO SE PUEDE LLEVAR A CABO ESTE CALCULO.

PARA LOS GASTOS ADMINISTRATIVOS SE ANALIZAN LOS BALANCES DE LA COMPAÑIA DE LOS ULTIMOS 5 AÑOS A LA FECHA, Y EN BASE A ELLO SE HACE UN ESTIMADO DEL GASTO POR CONCEPTO.

SI CONVENIMOS EN DENOTAR POR $Y_{p,i}$ EL GASTO ADMINISTRATIVO A RECARGAR POR EL $p$-ÉSIMO CONCEPTO CORRESPONDIENTE AL I-ÉSIMO PERIODO DE EXHIBICION DE PRIMAS, DE FORMA TAL QUE

$$Y_{p,i+1} = Y_{p,i}\,(1 + f_{i+1})$$

DONDE $f_{i+1}$ ES LA INFLACION ESTIMADA POR EL i+1 - ESIMO PERIODO.

PUEDE CALCULARSE ENTONCES LOS VALORES DE $Y_{p,i}$ PARA $i \in \overline{1,n}$, SIENDO n EL PERIODO DE COBERTURA DEL SEGURO DADO UN $Y_{p,0}$ QUE PUEDE SER DETERMINADO EN RELACION A LOS BALANCES DE LA COMPAÑIA EN AÑOS PREVIOS CON VALORES DEFLACTADOS A UNA FECHA DETERMINADA.

DE ESTA MANERA, SI $T_{p,k}$ REPRESENTA EL GASTO DEFLACTADO EFECTUADO EN EL P-ÉSIMO PERIODO CORRESPONDIENTE AL K-ÉSIMO AÑO, ENTONCES

$$Y_{p,0} = \frac{1}{L}\left[\sum_{k=0}^{L-1} \frac{T_{p,L-k}}{\beta_{L-k}}\right]$$

HAGAMOS EL SIGUIENTE EJERCICIO NUMERICO PARA ANALIZAR EL ASSET SHARE DE UN PLAN DE SEGURO PROPUESTO PARA DIFERENTES EDADES, EL CUAL NO OTORGA DIVIDENDOS NI PRESTAMOS.

**SUPOSICIONES**

**TIPO DE SEGURO:** TEMPORAL A 20 AÑOS.

**PAGO DE PERIODO DE PRIMAS:** PRIMAS NETAS ANUALES. (CONSTANTES)

**EDAD DE EMISION:** 25 AÑOS.

**SUMA ASEGURADA:** 1,000,000.00

**NUMERO DE ASEGURADOS AL INICIO DEL PLAN:** 10,000.

**TABLA DE MORTALIDAD PARA CALCULOS DEL ASSET SHARE:** CSO 1958.

**TABLA DE MORTALIDAD PARA CALCULOS DE LA PRIMA NETA:** EXPERIENCIA MEXICANA 62-67

**INTERES AL CUAL SE REINVIERTE LA RESERVA:** 65% CADA AÑO.

**METODO DE RESERVA MODIFICADO:** UN AÑO PRELIMINAR COMPLETO (FPT)

**GASTOS DE ADQUISICION:** 30% EL PRIMER AÑO Y

0% LOS AÑOS POSTERIORES.

**GASTOS DE ADMINISTRACION:** 10% CADA AÑO.

EL CALCULO DE LAS COLUMNAS DE ESTE EJERCICIO ES EFECTUADO DE LA SIGUIENTE MANERA:

**PRIMER RENGLON DEL ASSET-SHARE: $i = 1$.**

**AÑO:** ES EL AÑO EN EL QUE INICIA EL SEGURO. **0**

**$Z_i$ :** ES EL NUMERO DE POLIZA VIGENTES, EN ESTE CASO ES EL NUMERO DE ASEGURADOS CON LOS QUE INICIA EL PLAN. **10000.**

**$P_i$ :** ES LA CANTIDAD QUE POR CONCEPTO DE POLIZAS SE VA A RECIBIR AL INICIO DEL PLAN.

$P_i$ = prima de tarifa por unidad monetaria x
suma asegurada por poliza x
numero de poliza vigentes en el plan

$$P_i = .00268982 \times 1000000 \times 10000 = 26{,}898{,}193.00$$

**$C_i$ :** ES LA CANTIDAD TOTAL QUE POR CONCEPTO DE COMISIONES SE VA A PAGAR A LOS AGENTES AL INICIO DEL I-ESIMO AÑO.

$C_i$ = porcentaje de comisiones que se paga
a los agentes por unidad monetaria x $P_i$

$$C_i = .6 \times 26898193 = 16138916$$

| EDAD DE EMISION | P L A N | PRIMA NETA | PRIMA INICIAL | PRIMA RENOV. |
|---|---|---|---|---|
| 18 | ORDINARIO VIDA | 0.0032199696 | 0.0010174752 | 0.0032199708 |
| 18 | O.V. 15 PAGOS LIM. | 0.0044294501 | 0.0022597865 | 0.0047036430 |
| 18 | O.V. 25 PAGOS LIM. | 0.0036160723 | 0.0011722159 | 0.0036160724 |
| 18 | O.V. 40 PAGOS LIM. | 0.0032960194 | 0.0008521630 | 0.0032960195 |
| 18 | O.V. 50 PAGOS LIM. | 0.0032397105 | 0.0007958541 | 0.0032397107 |
| 18 | DOTAL 15 A'OS | 0.0338559472 | 0.0316862835 | 0.0341301401 |
| 18 | DOTAL 30 A'OS | 0.0091189908 | 0.0066751344 | 0.0091189909 |
| 18 | DOTAL 40 A'OS | 0.0051462907 | 0.0027024343 | 0.0051462908 |
| 18 | DOTAL 50 A'OS | 0.0037200020 | 0.0012761456 | 0.0037200021 |
| 18 | PRIMA NATURAL | 0.0016967943 | | |
| 30 | ORDINARIO VIDA | 0.0057158037 | 0.0011591056 | 0.0062191354 |
| 30 | O.V. 15 PAGOS LIM. | 0.0076659643 | 0.0031763929 | 0.0082364056 |
| 30 | O.V. 25 PAGOS LIM. | 0.0062893938 | 0.0017327109 | 0.0067927236 |
| 30 | O.V. 40 PAGOS LIM. | 0.0057878452 | 0.0012311623 | 0.0062911750 |
| 30 | O.V. 50 PAGOS LIM. | 0.0057246862 | 0.0011680033 | 0.0062280160 |
| 30 | DOTAL 15 A'OS | 0.0343941996 | 0.0299046282 | 0.0349646409 |
| 30 | DOTAL 30 A'OS | 0.0102390346 | 0.0056823517 | 0.0107423644 |
| 30 | DOTAL 40 A'OS | 0.0067752882 | 0.0022186053 | 0.0072786180 |
| 30 | DOTAL 50 A'OS | 0.0058464399 | 0.0012897570 | 0.0063497697 |
| 30 | PRIMA NATURAL | 0.0022073326 | | |
| 45 | ORDINARIO VIDA | 0.0140249685 | 0.0031033760 | 0.0151254981 |
| 45 | O.V. 15 PAGOS LIM. | 0.0175473837 | 0.0080111806 | 0.0187934084 |
| 45 | O.V. 25 PAGOS LIM. | 0.0147465075 | 0.0050764528 | 0.0158586807 |
| 45 | O.V. 40 PAGOS LIM. | 0.0149703882 | 0.0053003336 | 0.0160825614 |
| 45 | O.V. 50 PAGOS LIM. | 0.0140250601 | 0.0043550055 | 0.0151372333 |
| 45 | DOTAL 15 A'OS | 0.0372221558 | 0.0276859526 | 0.0384681804 |
| 45 | DOTAL 30 A'OS | 0.0158639506 | 0.0061938959 | 0.0169761238 |
| 45 | DOTAL 40 A'OS | 0.0141454531 | 0.0044753984 | 0.0152576263 |
| 45 | DOTAL 50 A'OS | 0.0140256075 | 0.0043555529 | 0.0151377807 |
| 45 | PRIMA NATURAL | 0.0048801665 | | |

## 3.4 Algoritmos empleados.

Los siguientes son los algoritmos empleados para la solución de un problema de flujo en redes a partir de la especialización del método simplex. En el primer algoritmo se describe la solución al sistema $\Pi B = c^B$, que se emplea principalmente para obtener el valor de las variables duales de un problema de flujo en redes, con B la matriz base y $c^B$ el vector básico de costos. El segundo algoritmo muestra la solución al sistema $By = d$, para obtener la columna de cambio en el método simplex, y el algoritmo 3.3 indica los pasos que se siguen en el método simplex especializado en Redes para alcansar la solución óptima basándose en los algoritmos que se mencionaron antes.

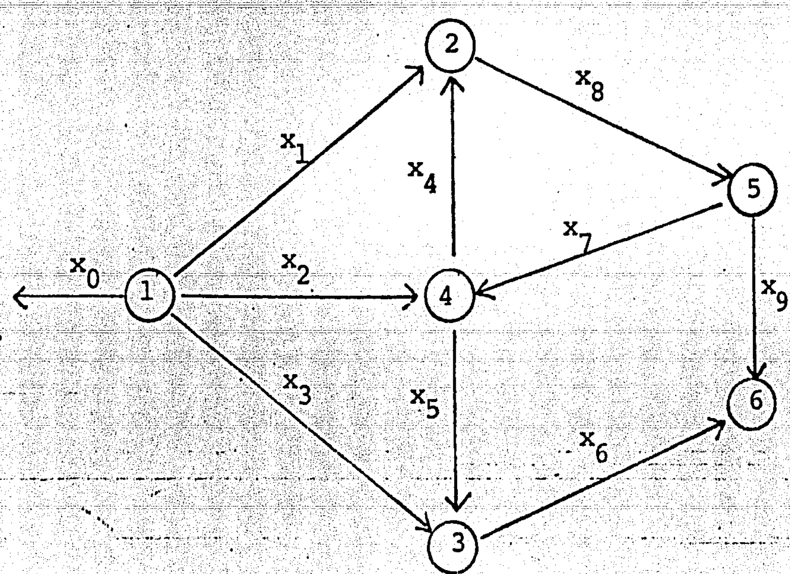
ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

Agregamos un arco raíz $x_0$ y el problema queda como sigue:

$$\begin{array}{llcl}
\text{Min} & Z = x_1 + 4x_2 + 2x_4 + x_6 + 2x_8 + 5x_9 & & \\
\text{S.A} & x_0 + x_1 + x_2 + x_3 & = & 14 \\
 & -x_1 - x_4 + x_8 & = & 5 \\
 & -x_3 - x_5 + x_6 & = & 3 \\
 & -x_2 + x_4 + x_5 - x_7 & = & -12 \\
 & x_7 - x_8 + x_9 & = & -1 \\
 & -x_6 - x_9 & = & -9
\end{array}$$

$x_0 \leq 0$, $x_1 \leq 10$, $x_2 \leq 8$, $x_3 \leq 12$, $x_4 \leq 7$, $x_5 \leq 13$, $x_6 \leq 10$, $x_7 \leq 6$, $x_8 \leq 12$, $x_9 \leq 8$ y $0 \leq x_i$, $i = 0,\ldots,9$.

La gráfica de este problema es



precios no aumentarlos y mantenerles su valor en el período de reajuste.

El grafo de proyección nos evita todo éste trabajo ; simplemente localicemos el valor requerido a Déficit Financiero y observemos los valores de tendencia de los índices para después localizar los valores respectivos al período anterior; Para obtener los diferen--tes valores de forma exácta bastará recurrir a la tabla de la pá--gina 85.

Observemos tambien que el grafo de proyección nos permite ver dentro de que valores se encuentran las tendencias de los dife--rentes índices que , aunque se hubiesen empezado a modificar me--diante reajustes en los precios,éstos tendrían yá un nivel muy alto derivado de la situación en el Déficit Financiero,entonces no solo se debe de contemplar la forma más adecuada de incremento en los precios,sino de buscar la forma de que los precios no se nos disparen tanto , la manera más adecuada sería mediante la im--plantación de medidas de "Freno" a la inflación para que con és--to se inicie un proceso de recuperación económica.

A continuación se propone una forma intuitiva de realizar un reajuste a los valores de los Indices de Precios durante un cierto período de 1988 en base al comportamiento de los mismos durante su correspondiente período inmediato anterior.

## II.2.- TRES AREAS IMPORTANTES PARA LA GRAFICACION.

De todo lo anterior, la memoria se encuentra dividida en varias áreas que son:

a) La memoria de la pantalla, que determina cuáles caracteres se desplegarán en la pantalla.

b) El conjunto de localidades llamado la memoria de los caracteres. En ella se encuentran los códigos de los patrones de los caracteres.

c) El conjunto de localidades conocido como la memoria del color. En esta área se encuentran los códigos de los colores de cada una de las localidades de la pantalla.

Existen otros conjuntos de localidades, pero para efectos de crear gráficas éstos son los más importantes.

### II.2.1.- LA MEMORIA DE LA PANTALLA.

En cualquier localidad de memoria sólo se pueden guardar números entre el 0 y el 255, ya que 255 es el número máximo que se puede expresar en ocho bits. El mismo número puede significar distintas cosas dependiendo de la localidad en que se encuentre.

Cada byte representa una pequeña área de la pantalla. La memoria de la pantalla está constituída por una sucesión de mil bytes, es decir, es una hilera de localidades de memoria que empieza en el byte 0 y termina en el 999, y el procesador la lee como si fuera una página de un libro. El primer byte que es el 0, se encuentra en la esquina superior izquierda de la pantalla. En esa posición se encuentra el corresponda a dicho punto en el byte que vamos a usar como máscara para operar.

Para apagar el bit 7 y que los demás bits permanezcan sin cambio tendremos que colocar un 0 en la posición del bit 7 y un 1 en las demás posiciones, por lo que tendremos 01111111 que en decimal corresponde a 127.

Este número lo podemos obtener de la siguiente manera: primero calculamos el valor decimal correspondiente al bit 7, esto es $2^7 = 128$ y luego debe restarse éste del número máximo que podemos representar en ocho bits (255). Para que permanezcan los demás bits como estaban y solamente se apague el bit 7 tendremos que usar una máscara de 255-128=127.

Si utilizamos AND 127 la operación binaria que efectuaremos será:

```
     bit    7 6 5 4 3 2 1 0

            x x x x x x x x
 AND 127    0 1 1 1 1 1 1 1
           ----------------
            0 x x x x x x x
```

Si deseamos apagar más de un bit, tendremos primero que encontrar el valor decimal de cada uno de los bits que deseamos apagar y posteriormente, la suma de estos valores la utilizamos como argumento para el AND.

Para apagar todos los bits sólo tendremos que poner como argumento del AND un 0 ó simplemente colocar el 0 en la localidad que se desee.

Por ejemplo, la localidad 56320 contiene datos del teclado y de las palancas de juego. Para leer de la palanca de juego minos reales poco frecuente, esto es, un subordinado puede ser especialista en una función (por ejemplo programación), no así su superior, por tanto, el superior debe ser un receptor de conocimientos y darle el crédito al subordinado que se lo merece, i.e., estimular el desarrollo creativo de los subordinados.

b) Voluntar para soltar las riendas: esto es, a los subordinados se le deberá dar responsabilidad plena para hacer su trabajo y tomar sus propias decisiones, como consecuencia se les hace responsables de sus equivocaciones, se los estimula y motiva para que ejerzan la iniciativa y para que acepten la responsabilidad absoluta de que harán el trabajo adecuadamente. Con eso, se les ayuda a ser ellos mismos, individuos más eficientes y empleados dignos de confianza. Al final, se tendrá una organización más eficiente que si se guardara toda la responsabilidad y la autoridad. Al mismo tiempo es preciso no ser un capataz, sino un administrador indulgente.

c) Voluntad para admitir los errores ajenos: el administrador debe estar dispuesto a asumir la responsabilidad plena de sus acciones ó de su fracaso en emprender alguna acción. Deberá también estar dispuesto a asumir toda la responsabilidad de lo que su empresa haga o deje de hacer, sus fracasos y también sus éxitos.

d) Voluntad para confiar en los subordinados: cuando el administrador actua de mala fe, por ejemplo al utilizar su posición

En el fragmento 8 (verso 5) de Parménides leemos la primera negación de una continuidad de la realidad a través de instancias temporales, de las cuales se rechazan expresamente dos, pero a la vez un esbozo del concepto de eternidad al afirmarse la permanencia de la realidad sustraída a la sucesión de instancias, como un presente inmóvil:

– Nunca fue ni será, puesto que es ahora, todo a la vez un continuo."[2]

"El futuro, el pasado y el presente integran el fenómeno del tiempo en la acepción más cotidiana de ésta palabra."[3]

"Al tiempo se le han atribuido aspectos mágicos de poder y de él se dice por ejemplo: 'El tiempo es oro', 'El tiempo todo lo cura', 'El tiempo vuela', 'No hay tiempo para...', 'El tiempo está en nuestra contra'. A Napoleón Bonaparte se atribuye una frase célebre que representa, el enigma, que para el hombre el tiempo sigue significando. Napoleón dijo: 'Pregúntenme sobre cualquier tema menos sobre el tiempo'."[1]

"Supóngase por un momento que en el mundo dejaran de existir de repente los relojes, todos los relojes: los de pared, los de bolsillo, los de pulsera y hasta los de sol. ¿Qué ocurriría entonces? Nadie sabría la hora; ignoraría cuándo tenía que ir a la escuela o al trabajo; la gente concurriría a todas las tareas a horas distintas. Nadie estaría seguro de poder cumplir un compromiso; no se sabría cuándo debería partir un barco, despegar un avión, etc.; en realidad, reinaría el caos. Sin embargo, el mundo así era en un principio, entonces ello no importaba mucho. A nadie le interesaba la hora. No había escuelas a donde ir ni trenes que tomar. Nadie tenía que "llegar a tiempo". No existía el tiempo. Por eso, cuando un hombre salía de su caverna, ni él ni nadie sabía cuándo volvería. Pero, más tarde, sintió la necesidad de decir a los suyos cuándo lo haría y, finalmente encontró una manera de hacerlo."[4]

Einstein, nos dice en su teoría de la relatividad restringida que el tiempo es afectado por la velocidad, es decir que transcurre más lentamente en los sistemas en movimiento. Hay un ejemplo, que manejan ampliamente los escritores de ciencia ficción, y que es el siguiente, si un astronauta hiciera un viaje en una nave espacial a la velocidad de la luz, en un crucero que según el calendario terrestre durara cincuenta años, a su regreso a la Tierra habría "envejecido" solamente cuarenta y dos años, que sería el tiempo transcurrido indicado por el cronómetro de abordo.[5]

"La verdad es que el tiempo es un fenómeno que influye en el hombre y en la sociedad y que bien manejado constituye un factor de los mas importantes en la planeación vital y crecimiento de las personas, grupos y organizaciones."[1]

| Nombre del campo | Concepto | Posición De | Posición A | Longitud | Tipo |
|---|---|---|---|---|---|
| Dia | Fecha de registro o baja (AAMMDD) | 1 | 6 | 6 | N |
| * Serie | Es un campo clave de cursos (Año del curso, # de curso, # de lista) | 7 | 12 | 6 | N |
| * Tipoeqp | Código del tipo de equipo | 13 | 14 | 2 | N |
| * Mvmnt | Tipo de registro<br>0 = Control del sistema<br>1 = Alta<br>2 = Baja | 15 | 15 | 1 | N |
| * Usercode | Clave del usuario | 16 | 21 | 6 | A |
| Nombre | Nombre del usuario | 22 | 51 | 30 | A |
| Calle | Domicilio del usuario | 52 | 81 | 30 | A |
| Tel | Teléfono del usuario | 82 | 91 | 10 | N |
| Dependencia | Dependencia de la UNAM a la que pertenece el usuario | 92 | 94 | 3 | N |
| Carrera | Código de la ocupación o carrera del usuario | 95 | 112 | 18 | A |
| Tiempo | Tiempo límite de uso de la clave | 113 | 116 | 4 | N |
| Tiempgast | Tiempo utilizado de la clave por el usuario | 117 | 120 | 4 | N |

LOG-APARTADOS.- Este archivo está dedicado a llevar la bitácora del sistema y el área de mensajes de servicio.

| Nombre del campo | Concepto | Posición De | Posición A | Longitud | Tipo |
|---|---|---|---|---|---|
| * Dia | Fecha del mensaje (AAMMDD) | 1 | 6 | 6 | N |
| * Logarea | Procedencia del mensaje<br>0 = Control del sistema<br>08 = Control del servicio | 7 | 8 | 2 | N |
| Notas | Espacio para guardar el mensaje | 9 | 308 | 300 | A |

HOJAAPARTADO.- Este archivo está dedicado a llevar un registro de cada uno de los movimientos de reservado, que se realizan en cada una de las salas, simulando el formato de una hoja de apartado.

## B) DISPOSICIONES PRELIMINARES

La Ley del Mercado de Valores define:

### 1. LA OFERTA PUBLICA

Se considera oferta pública la que se haga por algún medio de comunicación masiva o a persona indeterminada para suscribir, enajenar o adquirir títulos o documentos de los denominados valores.

La oferta pública de valores y documentos a que se refiere la Ley del Mercado de Valores, requerirá ser previamente aprobada por la Comisión Nacional de Valores.

### 2. LOS VALORES

Son valores las acciones, obligaciones y demás títulos que se emitan en serie o en masa.

Para saber la Ganacia de Capital que ha de recibir el accionista, es deduciendo la comisión del intermediario la cual está libre de impuesto para el accionista, (que se verá más adelante, en el cálculo de operación de compra-venta de Valores de Renta Variable), habiendo una tabla de comisiones autorizada y expedida por la Bolsa Mexicana de Valores, S.A. de C.V.

2. Rendimiento por pago de dividendos: Es el dividendo que otorga la empresa alos accionistas por utilidades generadas en un ejercicio y pueden ser en efectivo o en especie.

D.2.) Dividendo

Es el reparto que se hace a los accionistas y son por las utilidades que tuvo la empresa en un ejercicio, esta liquidación a los accionistas pueden ser en varias formas, ya sea en dividendo en efectivo y/o en dividendo en Acciones (Capitalización) suscripción, split y canje.

Dividendo en Efectivo:

Es la parte que corresponde a un accionista y que es liquidado en efectivo por la empresa mediante la entrega de un cupón que va adherido al título, ese dividendo en efectivo va con cargo a las utilidades o a la parte de las utilidades que se van a repartir.

ANTERIOR SUCEDE SOBRE LA ESFERA CUYO RADIO $\|\partial\|$ ES TAL QUE: 

$$\|\partial\|^2 = \|\partial_o\|^2$$

DEMOSTRACION:

UTILIZAREMOS EL METODO DE LAGRANGE PARA ENCONTRAR EL MINIMO DE LA FUNCION DEFINIDA POR:

$$(6.2) \qquad U(\partial,\lambda) = S'(\theta) + \lambda(\|\partial\|^2-\|\partial_o\|^2)$$
$$= \Sigma_{u=1}^{n}\ [Y_u-f(\xi_u;\theta)-\Sigma_{i=1}^{p}\partial_i Z_{iu}]^2$$
$$+ \lambda(\|\partial\|^2-\|\partial_o\|^2)$$
$$= \|\ Y - f - Z\partial\ \|^2 + \lambda(\|\partial\|^2 - \|\partial_o\|^2)$$

HACIENDO,

$$(6.3) \qquad \delta U/\delta\partial_1 = \delta U/\delta\partial_2 = \ldots = \delta U/\delta\partial_p = 0 \quad , \quad \delta U/\delta\lambda = 0$$

OBTENEMOS:

$$\delta U/\delta\partial_1 = 0 = -2\Sigma_{u=1}^{n}\ [Z_{1u}(Y_u - f(\xi_u;\theta) - \Sigma_{i=1}^{p}\ \partial_i Z_{iu})] + 2\lambda\partial_1$$
$$\delta U/\delta\partial_2 = 0 = -2\Sigma_{u=1}^{n}\ [Z_{2u}(Y_u - f(\xi_u;\theta) - \Sigma_{i=1}^{p}\ \partial_i Z_{iu})] + 2\lambda\partial_2$$

.

.

.

$$\delta U/\partial_p = 0 = -2\Sigma_{u=1}^{n}\ [Z_{pu}(Y_u - f(\xi_u;\theta) - \Sigma_{i=1}^{p}\ \partial_i Z_{iu})] + 2\lambda\partial_p$$
$$\delta U/\delta\lambda = 0 = \|\partial\|^2 - \|\partial_o\|^2$$

O BIEN:

$$(6.4) \qquad 0 = -[Z^t(Y - f) - Z^t Z\partial] + \lambda\partial$$
$$(6.5) \qquad 0 = \|\partial\|^2 - \|\partial_o\|^2$$

DE LA ECUACION (6.4),CONCLUIMOS QUE PARA $\lambda$ DADO LA SOLUCION $\partial$ DEBE SER TAL QUE:

$$(6.6) \qquad (Z^t Z + \lambda I)\partial = Z^t(Y - f) = g$$

OBSERVAMOS QUE ESTA ECUACION ES IDENTICA A (6.1).SE SIGUE ENTONCES QUE SOBRE LA ESFERA DE RADIO $\|\partial\|$,CON $\|\partial\|^2 = \|\partial_o\|^2$ UN PUNTO ESTACIONARIO ES $\partial_o$.CONCLUIMOS QUE ES UN MINIMO POR EL HECHO DE QUE LA MATRIZ $(Z^t Z + \lambda I)$ ES POSITIVA DEFINIDA.CON LO ANTERIOR QUEDA DEMOSTRADO EL TEOREMA.

TEOREMA 2

SEA $\partial(\lambda)$ LA SOLUCION DE (6.1) PARA UN $\lambda$ DADO. ENTONCES, $\|\partial(\lambda)\|^2$ ES UNA FUNCION CONTINUA DECRECIENTE DE $\lambda$, TAL QUE SI $\lambda \to \infty$ ENTONCES $\|\partial(\lambda)\| \to 0$.

DEMOSTRACION:

DADO QUE LA MATRIZ $Z^tZ$ ES POSITIVA DEFINIDA (RESULTADO 2 DEL APENDICE I), PUEDE SER TRANSFORMADA DE LA SIGUIENTE MANERA:

$$\text{(6.7)} \qquad \mathbb{S}^t \mathbb{A} \mathbb{S} = \mathbb{D} \qquad , \quad Z^t Z = \mathbb{A}$$

DONDE $\mathbb{S}^t\mathbb{S} = I$ Y $\mathbb{D}$ ES UNA MATRIZ DIAGONAL CUYOS ELEMENTOS SON TODOS POSITIVOS, SEGUN EL RESULTADO 3 DEL APENDICE I.

CON LO ANTERIOR, (6.1) ADOPTA LA SIGUIENTE EXPRESION:

$$\text{(6.8)} \qquad (\mathbb{D} + \lambda I)\mathbb{S}^t \partial o = \mathbb{S}^t g$$

DE DONDE:

$$\text{(6.9)} \qquad \partial o = \mathbb{S}(\mathbb{D} + \lambda I)^{-1} \mathbb{S}^t g$$

DEFINIENDO $\nu = \mathbb{S}^t g$, TENEMOS:

$$\text{(6.10)} \qquad \begin{aligned} \|\partial o(\lambda)\|^2 &= [\mathbb{S}(\mathbb{D} + \lambda I)^{-1}\mathbb{S}^t g]^t \; [\mathbb{S}(\mathbb{D} + \lambda I)^{-1}\mathbb{S}^t g] \\ &= g^t \mathbb{S}(\mathbb{D} + \lambda I)^{-1}\mathbb{S}^t \mathbb{S}(\mathbb{D} + \lambda I)^{-1}\mathbb{S}^t g \\ &= \nu^t [(\mathbb{D} + \lambda I)^{-1}]^2 \nu \\ &= \Sigma_{i=1}^{p} \; \nu_i^2/(D_i + \lambda)^2 \end{aligned}$$

Y PUESTO QUE $\lambda >= 0$ CONCLUIMOS QUE SI $\lambda \to \infty$, ENTONCES $\|\partial o\|^2 \to 0$ L.Q.Q.D.

TEOREMA 3

SEA $\Gamma$ EL ANGULO ENTRE $\partial o$ Y $\partial_g$. ENTONCES, $\Gamma$ ES UNA FUNCION CONTINUA MONOTONA DECRECIENTE DE $\lambda$, TAL QUE SI $\lambda \to \infty$, ENTONCES $\Gamma \to 0$. DADO QUE $\partial_g$ ES INDEPENDIENTE DE $\lambda$, SE SIGUE QUE $\partial o$ GIRA HACIA $\partial_g$ A MEDIDA QUE $\lambda \to \infty$

DEMOSTRACION:

OBSERVEMOS PRIMERO QUE:

$$\text{(6.11)} \qquad \partial_g = 2[\; \Sigma_{u=1}^{n} (Y_u - f_u)\; \delta f_u/\delta\partial_j \,]^t \qquad j = 1, 2, \ldots, p$$

$$\text{Y COMO} \quad g = \Sigma_{u=1}^{n} (Y_u - f_u)\; \delta f_u/\delta\partial_j \qquad j = 1, 2, \ldots p$$

TENEMOS QUE :

$$\text{(6.12)} \qquad \partial_g = Cg$$

ES DECIR $\partial_g$ Y $g$ TIENEN LA MISMA DIRECCION, DE DONDE, $\cos\Gamma = \cos(\partial_o, g)$, CON LO QUE:

$$\cos\Gamma = \partial_o^t g / \|\partial_o\|\,\|g\|$$

$$(6.13)\qquad \begin{aligned} &= g^t \mathbb{S}(\mathbb{D} + \lambda I)^{-1}\mathbb{S}^t\, g \,/ (\nu^t[(\mathbb{D} + \lambda I)^2]^{-1}\nu)^{1/2}(g^t g)^{1/2} \\ &= \nu^t(\mathbb{D} + \lambda I)^{-1}\nu / (\nu^t[(\mathbb{D} + \lambda I)^2]^{-1}\nu)^{1/2}(g^t g)^{1/2} = \end{aligned}$$

$$= \{\Sigma_{i=1}^{p}\ \nu_i^2/(D_i + \lambda)\} \,/\, \{[\Sigma_{i=1}^{p}\ \nu_i^2/(D_i + \lambda)^2]^{1/2}\ (g^t g)^{1/2}\} > 0$$

DERIVANDO CON RESPECTO A $\lambda$ Y SIMPLIFICANDO SE TIENE:

$$(6.14)\quad d(\cos\Gamma)/d\lambda = [\Sigma_{i=1}^{p}\nu_i^2/(D_i+\lambda)][\Sigma_{i=1}^{p}\nu_i^2/(D_i+\lambda)^3] \div [\Sigma_{i=1}^{p}\nu_i^2/(D_i+\lambda)^2]^{3/2}(g^t g)^{1/2} - [\Sigma_{i=1}^{p}\nu_i^2/(D_i+\lambda)^2]^2 \div [\Sigma_{i=1}^{p}\nu_i^2/(D_i+\lambda)^2]^{3/2}(g^t g)^{1/2}$$

LA ECUACION ANTERIOR PUEDE SER REESCRITA COMO SIGUE:

$$(6.15)\quad d(\cos\Gamma)/d\lambda = [\Sigma_{i=1}^{p}\nu_i^2\Upsilon_{1i}][\Sigma_{i=1}^{p}\nu_i^2\Upsilon_{3i}] \div [\Sigma_{i=1}^{p}\nu_i^2/(D_i+\lambda)^2]^{3/2}[\Pi_{i=1}^{p}(D_i+\lambda)^2]^2(g^t g)^{1/2} - [\Sigma_{i=1}^{p}\nu_i^2\Upsilon_{2i}] \div [\Sigma_{i=1}^{p}\nu_i^2/(D_i+\lambda)^2]^{3/2}[\Pi_{i=1}^{p}(D_i+\lambda)^2]^2(g^t g)^{1/2}$$

DONDE :

$$\Upsilon_{1i} = \Pi_{i'=1}^{p}\ (D_{i'}+\lambda)$$

$$\Upsilon_{2i} = \Pi_{i'=1}^{p}\ (D_{i'}+\lambda)^2$$

$$\Upsilon_{3i} = \Pi_{i'=1}^{p}\ (D_{i'}+\lambda)^3$$

Y EN TODOS LOS CASOS $i' \neq i$

EL DENOMINADOR DE (6.15) ES CLARAMENTE POSITIVO, POR LO QUE EL SIGNO DE $d(\cos\Gamma)/d\lambda$ ESTA DETERMINADO POR EL NUMERADOR.

NOTEMOS QUE, $(\Upsilon_{1i})(\Upsilon_{3i}) = (\Upsilon_{2i})^2$ ; EL NUMERADOR PUEDE ENTONCES SER ESCRITO COMO:

$$(6.16)\quad \{\Sigma_{i=1}^{p}[\nu_i(\Upsilon_{1i})^{1/2}]^2\}\{\Sigma_{i=1}^{p}[\nu_i(\Upsilon_{3i})^{1/2}]^2\} - \{\Sigma_{i=1}^{p}[\nu_i(\Upsilon_{1i})^{1/2}][\nu_i(\Upsilon_{3i})^{1/2}]\}^2$$

PERO POR LA DEIGUALDAD DE SCHWARZ (6.16) ES POSITIVO, POR TANTO $d(\cos\Gamma)/d\lambda > 0$, LUGO $\cos\Gamma$ ES CRECIENTE Y POR (6.13) POSITIVO, ASI $0 < (\cos\Gamma) < 1$ . CON LO ANTERIOR $0 < \Gamma < \Pi/2$ (PODRIA SUCEDER TAMBIEN QUE: $-\Pi/2 < \Gamma < 0$ O BIEN, $3\Pi/2 < \Gamma < 2\Pi$, PERO ESTOS CASOS NO LOS CONSIDERAMOS PUES LO QUE FINALMENTE NOS INTERESA ES QUE $\partial_o$ Y $\partial_g$ DISTEN EN MENOS DE $90^o$) . AHORA BIEN, CUANDO

$\lambda \to \infty$, LA MATRIZ $(Z^tZ + \lambda I)$ SE VE DOMINADA POR LA DIAGONAL $\lambda I$. ASI, SE VE DE (6.1) QUE SI $\lambda \to \infty, \partial_0 \to g/\lambda$ ES DECIR EL ANGULO ENTRE $\partial_0$ Y $g$ TIENDE A CERO.

POR TODAS LAS OBSERVACIONES ANTERIORES, SE SIGUE QUE $\Gamma$ ES UNA FUNCION CONTINUA MONOTONA DECRECIENTE DE $\lambda$ TAL QUE SI $\lambda \to \infty \Rightarrow \Gamma \to 0$, QUE ES LO QUE QUERIAMOS MOSTRAR.

NOTEMOS QUE SI $\lambda = 0$ Y LA MATRIZ $Z^tZ$ ES DIAGONAL, $\Gamma = 0$; SI $Z^tZ$ NO ES DIAGONAL, POR LA CONTINUIDAD DE $\Gamma(\lambda)$ AFIRMAMOS QUE, $0 < \Gamma < 90^0$

HABIENDO EXPUESTO ESTOS TRES RESULTADOS, TRATAREMOS AHORA SOBRE EL ESCALAMIENTO DE LA MATRIZ $Z^tZ$ Y LOS VECTORS $\partial_i$ Y $g$.

DEFINIMOS:

(6.17) $$(Z^tZ)^* = (Z_{iu})^* = (Z_{iu}/[(Z_{ii})^{1/2}][(Z_{uu})^{1/2}])$$

(6.18) $$g^* = (g_i^*) = (g_i/(Z_{ii})^{1/2})$$

(6.19) $$\partial_i^* = \partial_i(Z_{ii})^{1/2}$$

AHORA BIEN, SABEMOS QUE $Z^tZ\,\partial_i = g$, POR LO QUE:

(6.20) $$\begin{bmatrix} 1/(Z_{11})^{1/2} & 0 & \cdots & 0 \\ 0 & 1/(Z_{22})^{1/2} & & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot\ \cdot\ 0 & 1/(Z_{pp})^{1/2} \end{bmatrix} Z^tZ\,\partial_i = \begin{bmatrix} 1/(Z_{11})^{1/2} & 0 & \cdots & 0 \\ 0 & 1/(Z_{22})^{1/2} & & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & 0 \\ 0 & \cdot & \cdot\ \cdot\ 0 & 1/(Z_{pp})^{1/2} \end{bmatrix} g$$

DE DONDE:

(6.21) $$(Z^tZ)^*\,\partial_i^* = g^*$$

PODEMOS AHORA EXPONER EL ALGORITMO CUYA CONSTRUCCION PARECE YA MUY CLARA CON LAS CONSIDERACIONES SUGERIDAS HASTA AHORA.

ESPECIFICAMENTE, EN LA j-ESIMA ITERACION DEBE CONSTRUIRSE LA ECUACION:

(6.22) $$(A^{*(j)}+\lambda^{(j)}I)\partial^{*(j)} = g^{*(j)}$$

ESTA ULTIMA DEBE RESOLVERSE PARA $\partial^{*(j)}$ Y, UTILIZANDO (6.19) ENCONTRAMOS $\partial^{(j)}$ ASI, EL NUEVO VECTOR PARA ENSAYO SERA:

(6.23) $$\theta^{(j+1)} = \theta^{(j)} + \partial^{(j)}$$

POR OTRO LADO, ES ESENCIAL ELEGIR $\lambda^{(r)}$ TAL QUE:

(6.24) $S(\theta^{(r+1)}) < S(\theta^{(r)})$ ,YA QUE EL METODO DE MARQUARDT PERTENECE A LOS ITERATIVOS ACEPTABLES.

DEBE ELEGIRSE PUES $\lambda^{(r)}$ SUFICIENTEMENTE GRANDE PARA QUE (6.24) SE CUMPLA, A MENOS QUE $\theta^{(r)}$ SEA YA UN VALOR ESTACIONARIO PARA $S(\theta)$. OBVIAMENTE SE REQUIERE DE ALGUN ALGORITMO PARA LA ELECCION DE DEL VALOR DE $\lambda$ EN LA r-ESIMA ITERACION. NOSOTROS SOLO MENCIONAREMOS QUE MARQUARDT CONSIDERA LO SIGUIENTE:

TOMANDO A S COMO FUNCION DE $\lambda$, DEFINE $\lambda^{(0)} = 10^{-2}$ Y $v > 0$ (EN SUS EXPERIMENTOS ENCONTRO QUE v=10 ERA UN BUEN CANDIDATO) Y CONSIDERANDO ESTAR POR ELEGIR A UN $\lambda^{(r)}$,

i) SI $S(\lambda^{(r-1)}/v) < S(\lambda^{(r-1)})$ ,ENTONCES $\lambda^{(r)} = \lambda^{(r-1)}/v$

ii) SI $S(\lambda^{(r-1)}/v) >= S(\lambda^{(r-1)})$ ,ENTONCES $\lambda^{(r)} = \lambda^{(r-1)}/v^m$

CON m IGUAL AL NUMERO DE VECES NECESARIAS PARA QUE (6.24) SE CUMPLA.

EN CUANTO AL MOMENTO EN EL CUAL DECIDIR QUE EL METODO DEBE DETENERSE, EL MISMO MARQUARDT OPINA QUE ES EL PASO r-ESIMO TAL QUE:

(6.25) $$|\partial^{(r)}{}_j| \,/\, (\sigma+|\theta_j{}^{(r)}|) < \varepsilon \quad \forall j$$

CON, POR EJEMPLO, $\varepsilon = 10^{-5}$ Y $\sigma = 10^{-3}$

POR OTRO LADO, AL PERTENECER A LOS METODOS ITERATIVOS ACEPTABLES LA CONVERGENCIA QUEDA PROBADA EN EL TEOREMA 2 DEL CAPITULO 4.

EN ESTA SECCION ABORDAREMOS DOS CLASES DE ESTIMACION DE REGIONES DE CONFIANZA. CUANDO SEA NECESARIO, RECURRIREMOS A RESULTADOS QUE CORRESPONDEN A LA TEORIA QUE PARA ESTE EFECTO, HA SIDO DESARROLLADA EN LOS MODELOS DE REGRESION LINEAL. 

EN LA PRIMERA CLASE, SUPONIENDO NORMALIDAD E INDEPENDENCIA EN LA DISTRIBUCION DE LOS ERRORES, DESARROLLAREMOS UNA CONSTRUCCION DE REGIONES DE CONFIANZA EXACTAS. EN LA SEGUNDA OBTENDREMOS REGIONES DE CONFIANZA APROXIMADAS, PARA LO CUAL ADEMAS DE NORMALIDAD E INDEPENDENCIA, SE HACE UNA SUPOSICION EXTRA, QUE ES LA DE LINEALIDAD LOCAL DEL MODELO ALREDEDOR DE $f(\xi;\theta)$.

POR CONSIDERARLO CONVENIENTE, ANTES DE LA EXPOSICION DE LOS METODOS DAREMOS UN ESBOZO DE LA GEOMETRIA DE MINIMOS CUADRADOS, TANTO EN EL CASO LINEAL COMO EN EL CASO NO LINEAL.

## 7.- GEOMETRIA DE MINIMOS CUADRADOS

### A) MODELO LINEAL

CONSIDEREMOS EL MODELO LINEAL:

(7.A.1) $$Y = X\theta + \varepsilon$$

Y SUPONGAMOS QUE EL EXPERIMENTO CONSTA DE n OBSERVACIONES.

EL ESPACIO MUESTRAL DEFINIDO POR ESTE MODELO, ES UN ESPACIO n-DIMENSINAL. EL VECTOR DE OBSERVACIONES Y, DEFINE UN VECTOR $\overline{0Y}$ DEL ORIGEN O AL PUNTO Y CON COORDENADAS $Y=(Y_1,Y_2,\dots,Y_n)$. LA MATRIZ X TIENE p VECTORES COLUMNA Y CADA UNO CONTIENE n ELEMENTOS. CADA j-ESIMO VECTOR DE LA MATRIZ X REPRESENTA LOS n VALORES QUE TOMO LA j-ESIMA VARIABLE EXPERIMENTAL DURANTE LA OBSERVACION DEL FENOMENO BAJO ESTUDIO. SI SUPONEMOS QUE LAS VARIABLES EXPERIMENTALES SON INDEPENDIENTES ENTRE SI, ENTONCES LA MATRIZ X ES DE RANGO COMPLETO Y TENEMOS p VECTORES LINEALMENTE INDEPENDIENTES $\overline{0X}_1,\overline{0X}_2,\dots,\overline{0X}_p$, QUE DEFINEN UN SUBESPACIO p-DIMENSIONAL LINEAL LLAMADO EL ESPACIO DE ESTIMACION, MISMO QUE ESTA CONTENIDO DENTRO DEL ESPACIO MUESTRAL. CUALQUIER PUNTO DE ESTE SUBESPACIO PUEDE SER REPRESENTADO

COMO EL PUNTO FINAL DE UN VECTOR QUE ES COMBINACION LINEAL DE LOS VECTORES QUE DEFINEN A ESTE SUBESPACIO,ESTO ES,COMBINACION LINEAL DE LAS COLUMNAS DE X.

SUPONGASE QUE EL VECTOR $X\theta$ DEFINE AL PUNTO T SOBRE EL ESPACIO DE ESTIMACION.LA DISTANCIA CUADRADA DEL PUNTO Y AL PUNTO T ESTA DADA POR:

(7.A.2) $$S(\theta) = (Y-X\theta)^t(Y-X\theta) = (Y-T)^t(Y-T)$$

EL ESTIMADOR PUNTUAL DE $\theta$ OBTENIDO POR EL CRITERIO DE MINIMOS CUADRADOS $\hat{\theta}$ ES TAL QUE EL PUNTO $\hat{Y} = X\hat{\theta}$ SOBRE EL ESPACIO DE ESTIMACION TIENE LA PROPIEDAD DE SER EL QUE MENOS DISTA DE Y ENTRE TODOS LOS DEMAS DE DICHO ESPACIO.

GEOMETRICAMENTE ENTONCES,$\hat{Y}$ RESULTA SER EL PIE DE LA PERPENDICULAR DE Y AL ESPACIO DE ESTIMACION.

EN TERMINOS DE VECTORES,

(7.A.3) $$Y = \hat{Y} + (Y-\hat{Y}) = \hat{Y} + e$$

DE ESTA MANERA TENEMOS QUE EL VECTOR Y PUEDE SER DIVIDIDO EN DOS COMPONENTES ORTOGONALES A SABER: a) $\hat{Y}$,CONTENIDO EN EL ESPACIO DE ESTIMACION Y b) e,EL VECTOR DE RESIDUALES QUE SE ENCUENTRA EN LO QUE SE DENOMINA EL ESPACIO DE ERROR.ESTE ULTIMO ESPACIO ES (n-p) DIMENSIONAL Y ESTA CONTENIDO EN EL ESPACIO MUESTRAL.

PODEMOS MOSTRAR ANALITICAMENTE QUE LOS VECTORES $\hat{Y}$ y e SON ORTOGONALES DE LA SIGUIENTE MANERA:

(7.A.4) $$\begin{aligned}\hat{Y}^t e &= (X\hat{\theta})^t(Y-\hat{Y}) \\ &= (X\hat{\theta})^t(Y-X\hat{\theta}) \\ &= \hat{\theta}^t X^t Y - \hat{\theta}^t X^t X\hat{\theta} \\ &= \hat{\theta}^t(X^t Y - X^t X\hat{\theta})\end{aligned}$$

PERO,COMO FUE SENALADO CON ANTERIORIDAD,SI EL RANGO DE X ES COMPLETO EL ESTIMADOR $\hat{\theta}$ DE $\theta$ OBTENIDO POR MINIMOS CUADRADOS ES TAL QUE, $\hat{\theta} = (X^tX)^{-1}X^tY$, DE DONDE $\hat{\theta}^t(X^tY - X^tX\hat{\theta}) = \hat{\theta}^t(X^tY - X^tY) = 0$, POR LO QUE, DE LA ECUACION (7.A.4) TENEMOS:

(7.A.5) $\hat{Y}^t e = 0$ POR LO TANTO, $\hat{Y} \perp e$ L.Q.Q.D.

NUEVAMENTE, CONSIDEREMOS T CUALQUIER PUNTO EN EL ESPACIO DE ESTIMACION. DADO QUE EL VECTOR $\overline{Y\hat{Y}}$ ES ORTOGONAL A DICHO ESPACIO, POR EL TEOREMA DE PITAGORAS TENEMOS:

$$(7.A.6) \qquad d^2(\overline{YT}) = d^2(\overline{Y\hat{Y}}) + d^2(\overline{\hat{Y}T})$$

O BIEN,

$$(7.A.7) \qquad S(\theta) = S(\hat{\theta}) + d^2(\overline{\hat{Y}T})$$

ASI, EL CONTORNO SOBRE LE ESPACIO MUESTRAL PARA EL QUE $S(\theta)$ = CONSTANTE, DEBE SER TAL QUE $d^2(\overline{\hat{Y}T}) = S(\theta)-S(\hat{\theta})$ = CONSTANTE.

POR OTRO LADO, DADO QUE:

$$(7.A.8) \qquad S(\theta) = (Y-X\theta)^t(Y-X\theta) = Y^tY-2\theta^tX^tY+\theta^tX^tX\theta$$

ENTONCES:

$$\begin{aligned}(7.A.9) \qquad S(\hat{\theta}) &= Y^tY - 2\hat{\theta}X^tY + \hat{\theta}^tX^tX\hat{\theta} \\ &= Y^tY - \hat{\theta}^tX^tY - \hat{\theta}^t(X^tY - X^tX\hat{\theta}) \\ &= Y^tY - \hat{\theta}^tX^tY \\ &= Y^tY - \hat{\theta}^tX^tX\hat{\theta}\end{aligned}$$

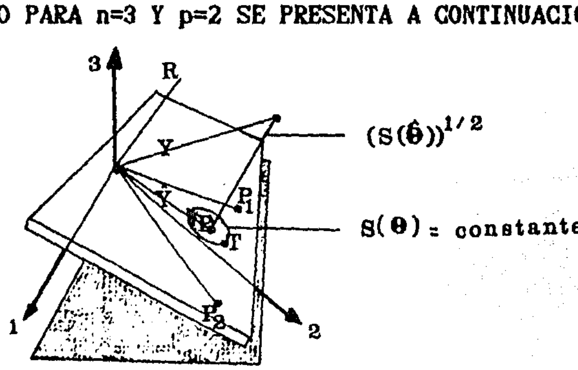
LAS ECUACIONES (7.A.8) Y (7.A.9) IMPLICAN QUE:

$$\begin{aligned}(7.A.10) \qquad S(\theta) - S(\hat{\theta}) &= \theta^tX^tX\theta - 2\theta^tX^tY + \hat{\theta}^tX^tX\hat{\theta} \\ &= (\theta - \hat{\theta})^t\ X^tX\ (\theta - \hat{\theta})\end{aligned}$$

POR EL RESULTADO 2 DEL APENDICE I, LA MATRIZ $X^tX$ ES POSITIVA DEFINIDA, POR LO QUE DE LA ECUACION (7.A.10), CONCLUIMOS QUE $S(\theta) - S(\hat{\theta})$ = CONSTANTE, REPRESENTA UN ELIPSOIDE CENTRADO EN $X\hat{\theta}$.

EN OTRA PALABRAS, $d^2(\overline{\hat{Y}T}) = S(\theta) - S(\hat{\theta})$ = CONSTANTE, CONSISTE EN TODOS LOS PUNTOS $T=X\theta$ EN EL ESPACIO DE ESTIMACION QUE ESTAN SOBRE UN ELIPSOIDE CENTRADO EN $\hat{Y} = X\hat{\theta}$.

UN EJEMPLO GRAFICO PARA n=3 Y p=2 SE PRESENTA A CONTINUACION.



B) MODELO NO LINEAL

CUANDO EL MODELO FUNCIONAL POSTULADO PARA UN FENOMENO ES DE LA FORMA:

(7.B.1) $$Y = f(\xi;\theta) + \varepsilon$$

CON $f$ UNA FUNCION NO LINEAL EN $\theta$ Y TOMAMOS UNA MUESTRA DE TAMANO n, EL ESPACIO MUESTRAL SERA UNO n-DIMENSIONAL QUE CONTENGA AL VECTOR Y. AUNQUE PUEDE HABLARSE DE UN ESPACIO DE ESTIMACION, ESTE NO ESTARA DEFINIDO POR LOS VECTORES COLUMNA DE UNA MATRIZ X EN EL SENTIDO USADO PARA EL CASO DE MODELOS LINEALES, Y PUEDE SER BASTANTE COMPLEJO.

SI TENEMOS p PARAMETROS INDEPENDIENTES $\theta = (\theta_1,\theta_2,\ldots,\theta_p)$, EL ESPACIO DE ESTIMACION DE UN MODELO NO LINEAL, ES UN SUBESPACIO DEL ESPACIO MUESTRAL QUE CONSISTE EN EL CONJUNTO DE PUNTOS DE LA FORMA:

(7.B.2) $$(f(\xi_1;\theta), f(\xi_2;\theta), \ldots, f(\xi_n;\theta))$$

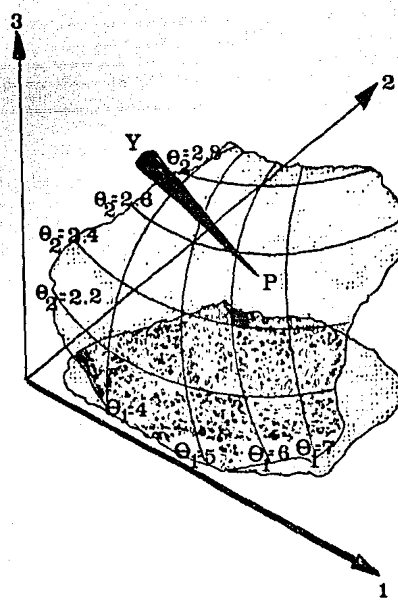
LA SUMA DE CUADRADOS $S(\theta)$ SIGUE REPRESENTANDO LA DISTANCIA CUADRADA DEL PUNTO $Y = (Y_1, Y_2,,,,,Y_n)$ AL PUNTO SOBRE EL ESPACIO DE ESTIMACION DEFINIDO POR $\theta$. DE IGUAL FORMA, LA MINIMIZACION DE $S(\theta)$ CORRESPONDE A ENCONTRAR EL PUNTO SOBRE EL ESPACIO DE ESTIMACION QUE MENOS DISTE DE Y.

LOS CONTORNOS PARA $S(\theta)$ = CONSTANTE NO TIENEN LA FORMA DE ELIPSOIDES, Y SUELEN FORMAR FIGURAS MUY IRREGULARES.

LA FIGURA QUE A CONTINUACION MOSTRAMOS, DESCRIBE UN EJEMPLO QUE INVOLUCRA n=3 OBSERVACIONES $Y_1, Y_2, Y_3$, TOMADAS EN $\xi_1, \xi_2$ Y $\xi_3$ RESPECTIVAMENTE, Y DOS PARAMETROS $\theta_1$ Y $\theta_2$. LAS LINEAS CURVEADAS INDICAN EL SISTEMA DE COORDENADAS DE LOS PARAMETROS EN EL ESPACIO DE ESTIMACION QUE CONSISTE EN TODOS LOS PUNTOS DE LA FORMA:

$$(f(\xi_1;\theta_1,\theta_2), f(\xi_2;\theta_1,O_2), f(\xi_3;,\theta_1\theta_2))$$

AL VARIAR $\theta_1$ Y $\theta_2$, MATENIENDO FIJOS $\xi_1, \xi_2$ Y $\xi_3$.




LOS RESULTADOS GEOMETRICOS ANTERIORES NOS SERAN DE GRAN UTILIDAD PARA LA COMPRENSION DE LOS METODOS DE ESTIMACION QUE A CONTINUACION DETALLAREMOS.

## 8.- REGIONES DE CONFIANZA EXACTAS PARA MODELOS DE REGRESION NO LINEAL

EL METODO QUE PRESENTAMOS ESTA BASADO EN LA OBTENCION DE REGIONES DE CONFIANZA EXACTAS PARA EL CASO LINEAL;POR LO ANTERIOR,COMENZAREMOS POR EXPONER LA FUNDAMENTACION DE ESTAS Y POSTERIORMENTE,VIA CIERTOS ARTIFICIOS CONVENIENTES,HAREMOS USO DE ELLA PARA OBTENER REGIONES EXACTAS PARA EL CASO NO LINEAL.

CONSIDEREMOS ENTONCES,NUEVAMENTE UN MODELO DE LA FORMA $Y = X\theta + \varepsilon$,CON LAS SIGUIENTES SUPOSICIONES:

$$\text{(8.1)} \qquad \begin{array}{l} RAN(X) = p < n \\ \varepsilon \sim N_n\ (0,\sigma^2 I) \end{array}$$

LA ECUACION (8.1) IMPLICA QUE

$$\text{(8.2)} \qquad Y \sim N_n\ (X\theta,\sigma^2 I)$$

BAJO ESTAS CONDICIONES,EL ESTIMADOR DE MINIMOS CUADRADOS Y MAXIMO VEROSIMIL DE $\theta$, $\hat{\theta} = (X^t X)^{-1} X^t Y$,ES TAL QUE:

$$\text{(8.3)} \qquad \hat{\theta} \sim N_p\ (\theta,\sigma^2(X^t X)^{-1})$$

HAREMOS A CONTINUACION UNA SERIE DE DEMOSTRACIONES PARA ENCONTRAR UNA REGION DE CONFIANZA PARA $\theta$.

PODEMOS MOSTRAR QUE:

$$\text{(8.4)} \qquad (\hat{\theta} - \theta)^t\ (X^t X)/\sigma^2\ (\hat{\theta} - \theta) \sim \chi^2_{(p)}$$

PUESTO QUE DE (8.3) SE TIENE,

$$(\hat{\theta} - \theta) \sim N_p(0.\sigma^2(X^t X)^{-1})$$

ADEMAS,

$$[(X^t X)/\sigma^2]\ [\sigma^2(X^t X)^{-1}] = I$$

LUEGO, $[(X^t X)/\sigma^2]\ [\sigma^2(X^t X)^{-1}]$ ES IDEMPOTENTE Y COMO,

$$RAN(X^t X) = RAN(X) = p$$

SE TIENE,POR EL RESULTADO 1 DEL APENDICE II (TOMANDO COMO LA MATRIZ A DE DICHO RESULTADO A LA MATRIZ $X^t X$), QUE:

$$[(\hat{\theta} - \theta)^t \; X^t X \; (\hat{\theta} - \theta)]/\sigma^2 \sim \chi^2_{(p,\lambda)}$$

DONDE ,

$$\lambda = [0^t \; X^t X \; 0]/\sigma^2 \;\; = 0$$

POR TANTO:

$$[(\hat{\theta} - \theta)^t \; X^t X \; (\hat{\theta} - \theta)]/\sigma^2 \sim \chi^2_{(p)} \qquad \text{L.Q.Q.D.}$$

TAMBIEN ES CIERTO QUE:

$$(8.5) \qquad [(Y - X\hat{\theta})^t (Y - X\hat{\theta})] \; 1/\sigma^2 \sim \chi^2_{(n-p)}$$

PARA MOSTRARLO,OBSERVEMOS LO SIGUIENTE:

$$\begin{aligned}[(Y-X\hat{\theta})^t (Y-X\hat{\theta})]1/\sigma^2 &= [Y-X(X^t X)^{-1}X^t Y]^t \; [Y-X(X^t X)^{-1}X^t Y] \; 1/\sigma^2 \\ &= \{[I-X(X^t X)^{-1}X^t]Y\}^t \; \{[I-X(X^t X)^{-1}X^t]Y\} \; 1/\sigma^2 \\ &= [Y^t (I-X(X^t X)^{-1}X^t)^t \; (I-X(X^t X)^{-1}X^t)Y] \; 1/\sigma^2\end{aligned}$$

PERO:

$$\begin{aligned}(I-X(X^t X)^{-1}X^t)^t &= I-(X(X^t X)^{-1}X^t)^t \\ &= I-[(X^t)^t (X(X^t X)^{-1})^t] \\ &= I-X[(X^t X)]^{-1}X^t \\ &= I-X(X^t X)^{-1}X^t\end{aligned}$$

DE TAL FORMA QUE:

$$[(Y-X\hat{\theta})^t (Y-X\hat{\theta})] \; 1/\sigma^2 = [Y^t (I-X(X^t X)^{-1}X^t)(I-X(X^t X)^{-1}X^t)Y] \; /\sigma^2$$

ADEMAS,

$$\begin{aligned}(I-X(X^t X)^{-1}X^t)(I-X(X^t X)^{-1}X^t) &= I-2X(X^t X)^{-1}X^t + X(X^t X)^{-1}X^t X \\ &\qquad (X^t X)^{-1}X^t \\ &= I-2X(X^t X)^{-1}X^t + X(X^t X)^{-1}X^t \\ &= I-X(X^t X)^{-1}X^t\end{aligned}$$

$$\Rightarrow (I-X(X^t X)^{-1}X^t) \text{ ES IDEMPOTENTE,POR LO QUE,}$$

$$[(Y-X\hat{\theta})^t (Y-X\hat{\theta})]/\sigma^2 = [Y^t (I-X(X^t X)^{-1}X^t)Y]/\sigma^2$$

POR LOS RESULTADOS ANTERIORES,LA MATRIZ $(I-X(X^t X)^{-1}X^t)/\sigma^2 \; (\sigma^2 I)$ ES CLARAMENTE IDEMPOTENTE Y COMO $RAN(I-X(X^t X)^{-1}X^t) = n-p$,UTILIZANDO EL RESULTADO 1 DEL APENDICE II TENEMOS QUE:

$$[(Y-X\hat{\theta})^t (Y-X\hat{\theta})]/\sigma^2 = [Y^t (I-X(X^t X)^{-1}X^t)Y]/\sigma^2 \;\sim\; \chi^2_{(n-p,\lambda)}$$

CON,

$$\lambda = [(X\theta)^t \;\; (I-X(X^t X)^{-1}X^t)/\sigma^2 \;\; X\theta \;] \;/\; 2$$

$$= [(X\theta)^t(X\theta)-(X\theta)^tX(X^tX)^{-1}X^tX\theta\ ]\ /2\sigma^2$$

$$= [\theta^tX^tX\theta-\theta^tX^tX(X^tX)^{-1}X^tX\theta\ ]\ /2\sigma^2$$

$$= (\theta^tX^tX\theta-\theta^tX^tX\theta)/2\sigma^2 = 0$$

POR TANTO:

$$[(Y-X\hat{\theta})^t(Y-X\hat{\theta})]/\sigma^2 \sim \chi^2_{(n-p)}$$

ES CONVENIENTE AHORA,REESCRIBIR $(Y-X\hat{\theta})^t(Y-X\hat{\theta})$ DE OTRA FORMA,MISMA QUE DETALLAMOS A CONTINUACION:

$$\begin{aligned}(Y-X\hat{\theta})^t(Y-X\hat{\theta}) &= (Y^t-(X\hat{\theta})^t)(Y-X\hat{\theta})\\ &= Y^tY - Y^tX\hat{\theta} - \hat{\theta}^tX^tY + \hat{\theta}^tX^tX\hat{\theta}\\ &= Y^tY - \hat{\theta}^tX^tY - \hat{\theta}^tX^tY + \hat{\theta}^tX^tX\hat{\theta}\\ &= Y^tY - \hat{\theta}^tX^tY - \hat{\theta}^tX^tY + \hat{\theta}^tX^tX(X^tX)^{-1}X^tY\\ &= Y^tY - \hat{\theta}^tX^tY - \hat{\theta}^tX^tY + \hat{\theta}^tX^tY\\ &= Y^tY - \hat{\theta}^tX^tY\\ &= Y^tY - \hat{\theta}^tX^tX(X^tX)^{-1}X^tY\\ &= Y^tY - \hat{\theta}^tX^tX\hat{\theta}\end{aligned}$$

POR TANTO,

$$(8.6)\qquad (Y-X\hat{\theta})^t(Y-X\hat{\theta}) = Y^tY - \hat{\theta}^tX^tX\hat{\theta}$$

NUESTRO SIGUIENTE PASO ES MOSTRAR QUE LA VARIABLES,

$$(8.7)\ [(\hat{\theta} - \theta)^t\ X^tX\ (\hat{\theta} - \theta)]/\sigma^2 \quad Y \quad [Y^t(I-X(X^tX)^{-1}X^tY]/\sigma^2$$

SON INDEPENDIENTES.

PARA ELLO,NOTEMOS QUE EN LA PRIMERA DE ELLAS LA FUENTE DE VARIACION ES $\hat{\theta}$,POR LO QUE BASTA MOSTRAR INDEPENDENCIA ENTRE $\hat{\theta} = (X^tX)^{-1}X^tY$ Y $Y^t(I-X(X^tX)^{-1}X^t)Y$. PERO,SEGUN EL RESULTADO 4 DEL APENDICE II ESTO SUCEDE SI Y SOLO SI:

$$(X^tX)^{-1}X^t\ (I-X(X^tX)^{-1}X^t) = 0$$

$$\Leftrightarrow \quad (X^tX)^{-1}X^t-(X^tX)^{-1}X^tX(X^tX)^{-1}X^t = 0$$

$$\Leftrightarrow \quad (X^tX)^{-1}X^t - (X^tX)^{-1}X^t = 0$$

$$\Leftrightarrow \quad 0 = 0$$

POR LO TANTO, $[(\hat{\theta} - \theta)^tX^tX(\hat{\theta} - \theta)]/\sigma^2$ Y $[Y^t(I-X(X^tX)^{-1}X^t)Y]/\sigma^2$ SON INDEPENDIENTES.

FINALMENTE UTILIZANDO LOS RESULTADOS (8.4),(8.5),(8.6),(8.7),

Y EL 9 DEL APENDICE II, TENEMOS:

(8.8) $\quad [(\hat{\theta} - \theta)^t X^t X(\hat{\theta} - \theta)/Y^t Y - \hat{\theta}^t X^t X\hat{\theta}\ ]\ (n-p)/p \quad \sim \quad F_{(p,n-p)}$

QUE PUEDE SER UTILIZADA COMO CANTIDAD PIVOTAL PARA ENCONTRAR UNA REGION DE CONFIANZA PARA $\theta$.

CON ESTA BASE, A CONTINUACION TARATAREMOS DE CONSTRUIR REGIONES DE CONFIANZA DENOMINADAS EXACTAS PARA MODELOS NO LINEALES.

POSTULAREMOS EL MODELO FUNCIONAL NO LINEAL COMO SIGUE:

(8.9) $\qquad Y = Z\lambda + \varepsilon$

CON,

$Y \sim$ VECTOR $(n\times 1)$ DE LOS VALORES OBSERVADOS EN LA VARIABLE RESPUESTA

$Z \sim$ MATRIZ $(n\times l)$ QUE INVOLUCRA TANTO A LAS VARIABLES CONTROLADAS COMO A LOS PARAMETROS $\theta_1, \theta_2, \ldots, \theta_p$ EN FORMA NO LINEAL.

SUPONDREMOS, RAN$(Z) = l < n$

$\lambda \sim$ VECTOR $(l\times 1)$ DE CONSTANTES DESCONOCIDAS

$\varepsilon \sim$ VECTOR $(n\times 1)$ DE ERRORES. SUPONDREMOS $\varepsilon \sim N_n\ (0, \sigma^2 I)$

CONSIDEREMOS AHORA, UNA MATRIZ $\mathbb{D}$ $(n\times p)$ ARBITRARIA CON LA UNICA RESTRICCION DE QUE LA MATRIZ AUMENTADA $(Z \mid \mathbb{D})$ $(n\times(l+p))$ SEA TAL QUE RAN$(Z \mid \mathbb{D}) = (l+p) < n$

PODEMOS ENTONCES, REEXPRESAR (8.9) DE LA SIGUIENTE FORMA:

(8.10) $\qquad Y = (Z \mid \mathbb{D})\beta + \varepsilon$

CON,

$$\beta^t = (\lambda^t, 0^t) = (\beta_1{}^t, \beta_2{}^t)$$

EL MODELO (8.10) ES UN MODELO LINEAL EN $\beta$ Y DADO QUE LA MATRIZ ASOCIADA AL MISMO ES DE RANGO COMPLETO SABEMOS, POR LA TEORIA DE REGRESION LINEAL, QUE EL ESTIMADOR DE MINIMOS CUADRADOS ES TAL QUE:

$$(8.11)\quad \hat{\beta}= \begin{bmatrix} Z^tZ & Z^t\mathbb{D} \\ \mathbb{D}^tZ & \mathbb{D}^t\mathbb{D} \end{bmatrix}^{-1} \begin{bmatrix} Z^t \\ \mathbb{D}^t \end{bmatrix} Y$$

O BIEN:

$$\hat{\beta}_1 = (Z^tZ)^{-1}Z^tY - (Z^tZ)^{-1}(Z^t\mathbb{D})(\mathbb{U}^t\mathbb{U})^{-1}\mathbb{U}^tY$$

$$(8.12)\quad \hat{\beta}_2 = (\mathbb{U}^t\mathbb{U})^{-1}\mathbb{U}^tY$$

DONDE, $\mathbb{U}^t = \mathbb{D}^t(I-Z(Z^tZ)^{-1}Z^t)$ ; $\hat{\beta}=(\hat{\beta}_1,\hat{\beta}_2)$

UTILIZANDO EL RESULTADO (8.8) TENEMOS:

$$(8.13)\quad [(\hat{\beta}-\beta)^t(Z\,|\,\mathbb{D})^t(Z\,|\,\mathbb{D})(\hat{\beta}-\beta) \,/\, Y^tY-\hat{\beta}^t(Z\,|\,\mathbb{D})^t(Z\,|\,\mathbb{D})\hat{\beta}] \times$$
$$\times\ [n-(p+l)]/(p+l)] \sim F_{(p+l,\,n-(p+l))}$$

$$(8.14)\quad [\hat{\beta}_2^t\,\mathbb{U}^t\mathbb{U}\,\hat{\beta}_2 \,/\, Y^tY-\hat{\beta}^t(Z\,|\,\mathbb{D})^t(Z\,|\,\mathbb{D})\hat{\beta}] \times$$
$$\times\ [n-(p+l)/p] \sim F_{(p,\,n-(p+l))}$$

LA VARIABLE (8.13) ES UNA CANTIDAD PIVOTAL A PARTIR DE LA CUAL OBTENDRIAMOS REGIONES DE CONFIANZA CONJUNTAS PARA $\lambda_1,\lambda_2,\ldots,\lambda_l,\theta_1,\theta_2,\ldots,\theta_p$.

DE IGUAL FORMA, LA VARIABLE (8.14) PROPORCIONARIA REGIONES DE CONFIANZA PARA $\theta_1,\theta_2,\ldots,\theta_p$ UNICAMENTE,EN CASO DE QUE $\mathbb{D}$ NO DEPENDIERA DE $\lambda$.

HASTA EL MOMENTO,NO HEMOS HECHO NINGUNA RESTRICCION SOBRE LA MATRIZ $\mathbb{D}$ EXCEPTO PEDIR QUE $(Z\,|\,\mathbb{D})$ FUERA DE RANGO COMPLETO.PODEMOS SIN EMBARGO SER MAS ESPECIFICOS SOBRE ESTE PUNTO.

NOTEMOS ANTE TODO QUE SI DESEAMOS SER CONSISTENTES CON LAS CARACTERISTICAS DE LAS REGIONES DE CONFIANZA,SERIA RAZONABLE PEDIR QUE CUANDO $\theta_1,\theta_2,\ldots\theta_p$ COINCIDIERAN CON SU ESTIMADOR DE MINIMOS CUADRADOS,LA VARIABLE (8.14) TOMARA EL VALOR DE CERO.VEAMOS EN QUE FORMA $\mathbb{D}$ DEBERIA ESTAR RELACIONADA CON LAS ECUACIONES DE MINIMOS CUADRADOS PARA $\lambda$ Y $\theta$.

CON UN MODELO COMO EL DE (8.9),TENEMOS QUE:

$$S(\theta;\lambda) = (Y-Z\lambda)^t(Y-Z\lambda) = Y^tY-2\lambda^tZ^tY-\lambda^tZ^tZ\lambda$$
$$= Y^tY-2Y^tZ\lambda-\lambda^tZ^tZ\lambda$$

DE DONDE LAS ECUACIONES DE MINIMOS CUADRADOS PARA $\lambda$ Y $\theta$ ESTAN DADAS POR:

$$(8.15)\quad -2(Z^tY-Z^tZ\lambda) = 0$$

(8.16) $-2\lambda^t[(\delta Z^t/\delta\theta_i)(Y-Z\lambda)] = 0$

LA ECUACION (8.16) PODEMOS REESCRIBIRLA COMO SIGUE:

(8.17) $-2\mathbb{D}^t(Y-Z\lambda) = 0$

TENIENDO A $\mathbb{D}^t$ COMO UNA MATRIZ DE $(p\times n)$ CON COMPONENTES:

(8.18) $D^t_{ih} = \Sigma_{k=1}^{l}\ \lambda_k\ \delta(Z^t_{kh})/\delta\theta_i \quad i=1,2,\ldots,p\ ;h=1,2,\ldots,n$

POR SUPUESTO, SEGUIREMOS PIDIENDO QUE LA MATRIZ DEFINIDA EN (8.18) SEA TAL QUE $(Z\ \vdots\ \mathbb{D})$ SEA DE RANGO COMPLETO.

AHORA BIEN, LA SOLUCION PARA $\lambda$ DE LA ECUACION (8.15) ES:

(8.19) $\hat{\lambda} = (Z^tZ)^{-1}Z^tY$

SUSTITUYENDO ESTE VALOR EN LA ECUACION (8.17) SE TIENE,

$$-2\mathbb{D}^t[Y-Z(Z^tZ)^{-1}Z^tY] = 0$$

(8.20) $\Rightarrow \quad -2\mathbb{D}^t[I-Z(Z^tZ)^{-1}Z^t]Y = 0$

ASI, DE LAS ECUACIONES (8.12) Y (8.20) CONCLUIMOS QUE:

(8.21) $-2\mathbb{U}^tY = 0 \quad \Rightarrow \quad \mathbb{U}^tY = 0$

POR LO QUE,

$$\hat{\beta}_2 = (\mathbb{U}^t\mathbb{U})^{-1}\mathbb{U}^tY = 0$$

Y EN CONSECUENCIA LA VARIABLE (8.14) SERA TAMBIEN IGUAL A CERO CUANDO $\theta_1,\theta_2,\ldots\theta_p$ COINCIDAN CON LOS ESTIMADORES DE MINIMOS CUADRADOS, COMO ERA DESEADO.

AHORA BIEN, ANALIZANDO (8.18) NOS PERCATAMOS QUE SI $\mathbb{D}$ ESTA DEFINIDA DE ESA FORMA, ESTA CONTENDRA AL MENOS UN SUBCONJUNTO DE $\{\lambda_1,\lambda_2,\ldots\lambda_l\}$ DE MODO QUE (8.14) NO ESTARA LIBRE DE $\lambda$. DE CUALQUIER FORMA, PARA CIERTOS MODELOS PODEMOS REDEFINIR $\mathbb{D}$ LIBRE DE $\lambda$ Y QUE POSEA LAS PROPIEDADES HASTA AHORA REQUERIDAS.

SUPONGAMOS QUE $\theta_i$ APARECE EN UNA Y SOLO UNA COLUMNA DE Z, O LO QUE ES LO MISMO, EN UNO Y SOLO UN RENGLON DE $Z^t$. EN ESTAS CONDICIONES, CADA UNA DE LAS ECUACIONES EN (8.16) INVULUCRARA UN SOLO RENGLON DE $Z^t$, DIGAMOS EL j-ESIMO Y ASI, DE $\lambda$ APARECERA SOLO COMO FACTOR $\lambda_j$ QUE CON UNA DEFINICION CONVENIENTE DE $\mathbb{D}$, PUEDE SER REMOVIDO DE TAL FORMA QUE EL i-ESIMO RENGLON DE $\mathbb{D}^t$ PUEDA SER

TOMADO COMO:

$$( \delta(Z^t{}_{j1})/\delta\theta_i , \delta(Z^t{}_{j2})/\delta\theta_i , \ldots , \delta(Z^t{}_{jn})/\delta\theta_i )$$

PARA TENER UNA IDEA MAS CLARA DE LO ANTERIOR, TOMEMOS COMO EJEMPLO EL CASO n=5, p=2, Y 1=2, ENTONCES:

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \quad ; \quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$$

Y,

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \\ Z_{31} & Z_{32} \\ Z_{41} & Z_{42} \\ Z_{51} & Z_{52} \end{bmatrix}$$

SUPONGAMOS QUE LA PRIMERA COLUMNA DE Z SOLO INVOLUCRA AL PARAMETRO $\theta_1$; DE IGUAL FORMA LA SEGUNDA COLUMNA SOLO ,INVOLUCRA AL PARAMETRO $\theta_2$

$$Z^t = \begin{bmatrix} Z_{11} & Z_{21} & Z_{31} & Z_{41} & Z_{51} \\ Z_{12} & Z_{22} & Z_{32} & Z_{42} & Z_{52} \end{bmatrix}$$

$$= \begin{bmatrix} Z^t{}_{11} & Z^t{}_{12} & Z^t{}_{13} & Z^t{}_{14} & Z^t{}_{15} \\ Z^t{}_{21} & Z^t{}_{22} & Z^t{}_{23} & Z^t{}_{24} & Z^t{}_{25} \end{bmatrix}$$

$$\Rightarrow \delta(Z^t)/\delta\theta_1 = \begin{bmatrix} \delta(Z^t{}_{11})/\delta\theta_1 & \delta(Z^t{}_{12})/\delta\theta_1 & \ldots & \delta(Z^t{}_{15})/\delta\theta_1 \\ 0 & 0 & \ldots & 0 \end{bmatrix}$$

$$\Rightarrow \lambda^t \delta(Z^t)/\delta\theta_1 = [\lambda_1\delta(Z^t{}_{11})/\delta\theta_1 \ \lambda_1\delta(Z^t{}_{12})/\delta\theta_1 \ \ldots \ \lambda_1\delta(Z^t{}_{15})/\delta\theta_1]$$

Y DE MANERA ANALOGA,

$$\lambda^t \delta(Z^t)/\delta\theta_2 = [\lambda_2\delta(Z^t{}_{21})/\delta\theta_2 \ \lambda_2\delta(Z^t{}_{22})/\delta\theta_2 \ \ldots \ \lambda_2\delta(Z^t{}_{25})/\delta\theta_2]$$

DEFINIMOS:

$$(\mathbb{D}')^t = \begin{bmatrix} \lambda_1\delta(Z^t{}_{11})/\delta\theta_1 & \lambda_1\delta(Z^t{}_{12})/\delta\theta_1 & \ldots & \lambda_1\delta(Z^t{}_{15})/\delta\theta_1 \\ \lambda_2\delta(Z^t{}_{21})/\delta\theta_2 & \lambda_2\delta(Z^t{}_{22})/\delta\theta_2 & \ldots & \lambda_2\delta(Z^t{}_{25})/\delta\theta_2 \end{bmatrix}$$

$$\mathbb{D}^t = \begin{bmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_2 \end{bmatrix} (\mathbb{D}')^t$$

CON ESTA DEFINICION, $\mathbb{D}$ NO INVOLUCRA AL VECTOR $\lambda$, ADEMAS $(Z \mid \mathbb{D})$ ES DE RANGO COMPLETO SI $(Z \mid \mathbb{D}')$ LO ES, Y LA CONDICION (8.1) DE CONSISTENCIA SE SIGUE CUMPLIENDO.

DE ESTA MANERA LA ESTADISTICA (8.14) ESTARA LIBRE DE $\lambda$ Y POR CONSECUENCIA SERA UNA CANTIDAD PIVOTAL PARA OBTENER REGIONES DE CONFIANZA EXACTAS PARA $\theta$.

EVIDENTEMENTE, EN UN PROBLEMA PRACTICO LA OBTENCION Y, PRINCIPALMENTE LA VISUALIZACION DE REGIONES DE CONFIANZA EXACTAS PARA $\theta$ EN UN MODELO DE REGRESION NO LINEAL, RESULTAN SER MUY COMPLEJAS. ES POR ELLO QUE SURGE LA NECESIDAD DE OBTENER OTRO TIPO DE REGIONES QUE SEAN LO MAS APROXIMADAS POSIBLES A LAS EXACTAS Y REDUZCAN LOS PROBLEMAS QUE ESTAS ULTIMAS PRESENTAN. EXISTEN MODELOS FUNCIONALES NO LINEALES QUE SON SUSCEPTIBLES DE UTILIZARSE PARA TECNICAS DE OBTENCION DE REGIONES DE CONFIANZA APROXIMADAS.

## 9.- REGIONES DE CONFIANZA APROXIMADAS PARA MODELOS DE REGRESION NO LINEAL



EL CONCEPTO "MODERADAMENTE NO LINEAL" LO MANEJAREMOS EN UN PRINCIPIO EN FORMA INTUITIVA.POR SUPUESTO QUE EXISTEN TECNICAS PARA CUANTIFICAR LA NO LINEALIDAD DE UN MODELO Y CRITERIOS PARA SEÑALARLO COMO MODERADAMENTE NO LINEAL O NO.MAS ADELANTE PRESENTAREMOS UNA DE ESTAS TECNICAS.

CABE SENALAR SIN EMBARGO,QUE EL CASO DE NO LINEALIDAD MODERADA ES MUY COMUN.ESTO ES,EN PROBLEMAS PRACTICOS NORMALMENTE SON ELEGIDOS MODELOS FUNCIONALES QUE,SI BIEN SON NO LINEALES,NO SON EXTREMADAMENTE COMPLEJOS,POR LO QUE PUEDEN CONSIDERARSE MODERADAMENTE NO LINEALES.

A CONTINUACION,PLANTEAREMOS LAS SUPOSICIONES DE LAS QUE DESPRENDEREMOS LA EXPOSICION DE ESTE CAPITULO.

(9.1) SEA, $Y_u = f(\xi_u;\theta) + \varepsilon_u \quad u=1,2,\ldots,n$

$f$;FUNCION NO LINEAL SOBRE $\theta$

a)

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix} \sim N(0,\sigma^2 I)$$

b) LOS PARAMETROS $\theta_1,\theta_2,\ldots\theta_p$,TIENEN EL MISMO PESO DENTO DEL MODELO,CON $\theta = (\theta_1,\theta_2,\ldots,\theta_p)$

c) AUNQUE $f(\xi_u;\theta)$ SON NO LINEALES EN $\theta$,LO SON SOLO MODERADAMENTE

SUPONIENDO ENCONTRARSE CON UN PROBLEMA EN LAS CONDICIONES DESCRITAS EN (9.1),BOX Y COUTIE PUBLICARON UN ARTICULO EN 1956 EN DONDE ASEGURAN QUE PUEDE MOSTRARSE QUE LA MATRIZ DE VARIANZAS Y COVARIANZAS DEL ESTIMADOR DE $\theta$ POR MINIMOS CUADRADOS $\hat{\theta}$,EQUIVALE APROXIMADAMENTE (ASINTOTICAMENTE) A UNA MATRIZ (p×p)

$\mathbb{V}^{-1}\sigma^2$,DONDE $2\mathbb{V}$ ES LA MATRIZ CON ELEMENTOS DADOS POR:

(9.2) $\quad S_{ij} = ( \delta^2 S(\theta)/\delta\theta_i\delta\theta_j )|\ \theta = \hat{\theta} \quad (i,j=1,2,\ldots,p)$

DONDE $S(\theta)$ ES LA FUNCION DE MINIMOS CUADRADOS.

DE ESTA FORMA,Y TOMANDO COMO ESTIMADOR DE $\sigma^2$ A LA ESTADISTICA $S(\hat{\theta})/(n-p)$,TENEMOS QUE LA DESVIACION ESTANDAR DE $\hat{\theta}_i$ PUEDE SER ESTIMADA POR:

(9.3) $\quad ( S(\hat{\theta})/(n-p) )^{1/2} ( V^{-1}_{ii} )^{1/2}$

DONDE $V^{-1}_{ii}$ ES EL i-ESIMO ELEMENTO DE LA DIAGONAL DE LA MATRIZ $\mathbb{V}^{-1}$.

ES ASI COMO PROPONEN PARA MUCHAS SITUACIONES PRACTICAS,COMO ADECUADO EL USO DE UNA REGION DE $(1-\alpha)$% APROXIMADA DE CONFIANZA DESCRITA POR LA DESIGUALDAD QUE SIGUE:

(9.4) $\quad 1/2[\Sigma_{i=1}^{p} \Sigma_{j=1}^{p} S_{ij}(\theta_i-\hat{\theta}_i)(\theta_i-\hat{\theta}_j)] <=$

$$<= [S(\hat{\theta})\ F_{(p,n-p)}\ 1-\alpha]\ p/(n-p)$$

ES IMPORTANTE,NOTAR LA ANALOGIA DE LA DESIGUALDAD (9.4) CON LA QUE SE OBTIENE PARA EL MISMO FIN EN EL CASO DE LAS REGIONES DE CONFIANZA EXACTAS EN MODELOS LINEALES.

POR OTRO LADO,EXISTEN MUCHOS AUTORES QUE OBTIENEN REGIONES DE CONFIANZA APROXIMADAS PARA $\theta$ SUPONIENDO LINEALIDAD LOCAL DE LA FUNCION $f$,BASANDOSE EN LA NO LINEALIDAD MODERADA DE LA MISMA.

SABEMOS QUE ALREDEDOR DE UN VALOR FIJO $\theta = \theta^0$ EL POLINOMIO DE TAYLOR DE GRADO UNO PARA $f(\xi;\theta)$ ES TAL QUE:

(9.5) $\quad f(\xi;\theta) \cong f(\xi;\theta^0) + \Sigma_{i=1}^{p} (\delta f(\xi;\theta)/\delta\theta_i\ |\theta = \theta^0)\ (\theta_i-\theta_i{}^0)$

LA SUPOSICION DE LINEALIDAD IMPLICA QUE,AL MENOS LOCALMENTE,EL ESPACIO DE ESTIMACION SEA UN SUBESPACIO p-DIMENSIONAL LINEAL DENTRO DEL ESPACIO MUESTRAL,Y QUE LOS COMPONENTES DE $\theta$ DEFINAN UN SISTEMA CARTESIANO UNIFORME DE COORDENADAS SOBRE DICHA SUPERFICIE.ESTOS SUPUESTOS SON DENOMINADOS RESPECTIVAMENTE:SUPOSICION PLANAR Y SUPOSICION DE COORDENADAS UNIFORMES.

BEALE [3],HACE ESTA SUPOSICION LINEAL LOCAL ALREDEDOR DEL ESTIMADOR DE MINIMOS CUADRADOS $\hat{\theta}$ Y DEFINE UNA REGION DEL $(1-\alpha)$%

APROXIMADA DE CONFIANZA PARA $\theta$,COMO EL CONJUNTO DE $\theta$'s TALES QUE:

$$(9.6)\quad S(\theta)-S(\hat{\theta}) <= \{[1+N\theta\ n(p+2)/(n-p)p]\ [\ (\ S(\hat{\theta})\ F_{(p,n-p)1-\alpha}\ )\ p/(n-p)]\}$$

DONDE $N\theta >= 0$ INDICA LA MEDIDA DE NO LINEALIDAD DEL MODELO.

NOTESE QUE (9.6) ES TAMBIEN MUY SIMILAR A UNA REGION EXACTA PARA EL CASO LINEAL.

DE HECHO,A MEDIDA QUE $N\theta$ DISMINUYE (9.6) TIENE UNA EXPRESION CADA VEZ MAS PARECIDA A LA DE UNA REGION DE CONFIANZA EXACTA PARA UN MODELO LINEAL,LO QUE LE HACE HEREDAR LAS PROPIEDADES DE ESTA ULTIMA.DE IGUAL FORMA, SI $N\theta$ AUMENTA EL NUMERO DE PUNTOS QUE QUEDAN INCLUIDOS EN LA REGION TAMBIEN AUMENTA,Y POR LO TANTO CADA VEZ ES MENOS CONFIABLE.

RESULTA RELEVANTE PREGUNTARSE COMO OBTENER LA MEDIDA DE NO LINEALIDAD $N\theta$ PARA UN MODELO FUNCIONAL DADO,O AL MENOS UNA ESTIMACION DE ELLA.A CONTINUACION PRESENTAREMOS LA CONSTRUCCION DE UNA MEDIDA DE NO LINEALIDAD EMPIRICA QUE DENOTAREMOS POR $\hat{N}\theta$.

PARA ESTE FIN,NOTEMOS QUE EL POLINOMIO DE TAYLOR DE GRADO UNO DE $\hat{\theta}$ PARA $f$ REPRESENTA,ADEMAS DE UNA APROXIMACION LINEAL A LA FUNCION,EL PLANO TANGENTE AL ESPACIO DE ESTIMACION EN $P(\hat{\theta}) = (\ f(\xi_1;\hat{\theta}),f(\xi_2;\hat{\theta}),\ldots,f(\xi_n;\hat{\theta})\ )$,DEFINIDO PARAMETRICAMENTE.

DEFINIREMOS $T(\theta)$ AL PUNTO SOBRE EL PLANO TANGENTE ASOCIADO CON EL PUNTO $P(\theta)$ EN EL ESPACIO DE ESTIMACION,O BIEN,EL PIE DE LA PERPENDICULAR DE $P(\theta)$ AL PLANO.

VAMOS A SUPONER QUE TENEMOS VALORES DE $f(\xi_u;\theta^v)$ PARA UN CONJUNTO DE VECTORES $\theta^v$ CERCANOS A $\hat{\theta}$ CON $w=1,2,\ldots,W$ Y $u=1,2,\ldots,n$.

UNA OPCION SOBRE LA ELECCION DE ESTOS $\theta^v$'s ES ESCOGER UN SUBCONJUNTO DE LOS VECTORES QUE PARA $\theta$ SE HAYAN OBTENIDO EN UN METODO ITERATIVO EN LA BUSQUEDA DE $\hat{\theta}$.

CONSIDEREMOS AHORA:

$$(9.7)\quad Q(\theta) = \Sigma_{v=1}^{W} \| P(\theta^v)-T(\theta^v) \|^2$$
$$= \Sigma_{v=1}^{W}\Sigma_{u=1}^{n} [\, f(\xi_u;\theta^v) - f(\xi_u;\hat{\theta}) -$$
$$- \Sigma_{i=1}^{p} \delta f(\xi_u;\hat{\theta})/\delta\theta_i \,|_{\theta=\hat{\theta}} (\theta^v{}_i-\hat{\theta}_i) \,]^2$$

$Q(\theta)$ REPRESENTA ENTONCES,LA SUMA DE LAS DISTANCIAS AL CUADRADO DE LOS PUNTOS $P(\theta^v)$ SOBRE EL ESPACIO DE ESTIMACION A LOS PUNTOS $T(\theta^v)$ SOBRE EL PLANO TANGENTE EN $P(\hat{\theta})$.ESTO PODRIA TOMARSE COMO UNA ESTIMACION A UNA MEDIDA DE NO LINEALIDAD DEL MODELO.

ES CLARO QUE EL VALOR $Q(\theta)$ DEPENDE TANTO DE UNIDADES COMO DE LA CANTIDAD DE PUNTOS $P(\theta^v)$ CONSIDERADOS,POR LO QUE SE HACE NECESARIO NORMALIZAR ESTA MEDIDA.

PARA RESOLVER ESTE PROBLEMA,PODEMOS CONSIDERAR:

$$(9.8)\quad \Sigma_{v=1}^{W} \| P(\theta^v) - P(\hat{\theta}) \|^2$$

Y TOMAR EL COCIENTE ENTRE (9.7) Y (9.8) COMO SIGUE:

$$(9.9)\quad \hat{N}\theta = [\Sigma_{v=1}^{W}\| P(\theta^v) - T(\theta^v) \|^2]/[\Sigma_{v=1}^{W}\|P(\theta^v) - P(\hat{\theta})\|^2]$$

A ESTA CANTIDAD LE LLAMAREMOS MEDIDA EMPIRICA DE NO LINEALIDAD.

EN CUANTO AL VALOR QUE DEBE TOMAR (9.9) PARA SUPONER QUE UNA REGION DEL $(1-\alpha)\%$ APROXIMADA DE CONFIANZA SEA CONFIABLE,BEALE CONSIDERA,POR EJEMPLO,QUE $\hat{N}\theta$ DEBE SER TAL QUE:

$$(9.10)\qquad \hat{N}\theta < 0.01\ F_{(p,(n-p))1-\alpha}$$

LA OBTENCION DE REGIONES DE CONFIANZA APROXIMADAS HA LLEGADO A SER MUY CONVENCIONAL Y,ADEMAS DE BEALE,HA SIDO SUGERIDA POR AUTORES COMO DRAPER Y SMITH [6].

AHORA BIEN,ESTE PROCEDIMIENTO TIENE VALIDEZ ASINTOTICA Y DESAFORTUNADAMENTE,POR EL COSTO QUE REPRESENTA EFECTUAR UN NUMERO GRANDE DE OBSERVACIONES,LOS MODELOS DE REGRESION NO LINEAL SON AJUSTADOS NORMALMENTE CON MUY POCOS DATOS,DE TAL SUERTE QUE ESTA VALIDEZ ASINTOTICA PUEDE SER UNA GUIA IRREAL PARA UN CASO PRACTICO.

NOS HEMOS OCUPADO EN EL PRESENTE TRABAJO DE LA ESTIMACION PARAMETRICA NO LINEAL TANTO PUNTUAL COMO POR REGIONES DE CONFIANZA Y ES EVIDENTE QUE A LA PRIMERA HEMOS BRINDADO MAYOR ENFASIS FORMAL,PUES EL INTERES DE LA INVESTIGACION SURGIO PRECISAMENTE SOBRE ESTE RENGLON.

POR LAS CARACTERISTICAS QUE PRESENTA EL DESARROLLO DE LOS METODOS DE ESTIMACION PUNTUAL AQUI EXPUESTOS,RESULTA MUY CLARO QUE EL DE LINEALIZACION ES EL MENOS SOFISTICADO EN CUANTO A LOS CALCULOS MATEMATICOS QUE INVOLUCRA.DESAFORTUNADAMENTE Y A PESAR DE LA GRAN VENTAJA DE TRANSFORMAR EL PROBLEMA DE MINIMIZAR LA FUNCION $S(\theta)$ EN UNA SERIE DE PROBLEMAS LINEALES,PRESENTA UNA GRAN DESVENTAJA QUE ES EL NO GARANTIZAR CONVERGENCIA A UN PUNTO ESTACIONARIO DE LA FUNCION DE MINIMOS CUADRADOS.

LOS METODOS ITERATIVOS ACEPTABLES PARECEN RESOLVER ESTE PROBLEMA SATISFACTORIAMENTE,AUNQUE CABE RECALCAR QUE LA CONVERGENCIA A UN PUNTO ESTACIONARIO QUE GARANTIZAN,NO IMPLICA QUE ESTA SEA A UN MINIMO GLOBAL COMO ES DESEABLE.ES AQUI DONDE RADICA PRINCIPALMENTE LA ELECCION ACERTADA DEL PUNTO INICIAL $\theta^{o}$.DE ELEGIRSE ERRONEAMENTE DICHO VECTOR,LOS RESULTADOS SEGURAMENTE NO SERAN SATISFACTORIOS,SEA CUAL SEA EL METODO ELEGIDO.

DENTRO DE ESTA ULTIMA CLASE DE PROCESOS DE ESTIMACION,TAMBIEN EXISTEN ALGUNOS QUE SON PREFERIBLES A OTROS,TAL ES EL CASO DEL METODO DEL DESCENSO POR LA PENDIENTE MAXIMA Y EL DE MARQUARDT.EL PRIMERO PRESENTA EL INCONVENIENTE QUE CARACTERIZA TAMBIEN AL METODO DE LINEALIZACION,Y QUE ES LA LENTITUD EN LA CONVERGENCIA.

MARQUARDT UTILIZA BASES TECNICAS DEL METODO DE LINEALIZACION Y DEL DESCENSO POR LA PENDIENTE MAXIMA,CONSIGUIENDO ENTRE OTRAS COSAS ACELERAR LA CONVERGENCIA Y CONSERAVAR LAS CARACTERISTICAS FAVORABLES A ESTOS DOS PROCEDIMIENTOS QUE DAN ORIGEN A SU

METODO,HACIENDOLO PARECER EL MEJOR DE LOS TRES. 

SIN EMBARGO ES DIFICIL DECIR QUE METODO ES MAS CONVENIENTE UTILIZAR EN CADA PROBLEMA PRACTICO,LA DECISION DEBE SER TOMADA POR EL INVESTIGADOR ANALIZANDO LAS CARACTERISTICAS ASOCIADAS AL PROBLEMA QUE SE LE PRESENTE Y AL COSTO QUE LE REPRESENTE CADA ALTERNATIVA.

EN CUANTO A LAS REGIONES DE CONFIANZA DIREMOS QUE,A PESAR DE QUE EL GRADO DE PRECISION DE LAS EXACTAS ES DE ALGUNA MANERA MAYOR AL DE LAS APROXIMADAS,LAS PRIMERAS PRESENTAN LA ENORME DESVENTAJA DE SER MUY COMPLEJAS EN CUANTO A SU OBTENCION Y VISUALIZACION,MIENTRAS QUE LAS SEGUNDAS SURGEN DE UNA MANERA MAS NATURAL.

AUN A PESAR DE QUE LA PRECISION DE LAS REGIONES DE CONFIANZA APROXIMADAS DEPENDE DEL GRADO DE NO LINEALIDAD DEL MODELO FUNCIONAL QUE EN CADA CASO SE ELIJA,SON MUY UTILIZADAS,PUES POR RAZONES OBVIAS LOS MODELOS DE AJUSTE ELEGIDOS PARA MUCHOS PROBLEMAS PRACTICOS,AUNQUE NO LINEALES,NO SON EXTREMADAMENTE COMPLEJOS.

# APENDICE I

1.- SEA $\mathbb{A}$ UNA MATRIZ CUADRADA Y SEAN $|\lambda_{MIN}|$, $|\lambda_{MAX}|$ LOS EIGENVALORES DE $\mathbb{A}$ MINIMO Y MAXIMO EN VALOR ABSOLUTO RESPECTIVAMENTE, ENTONCES PARA CUALQUIER VECTOR b SE TIENE:

a) $|\lambda_{MIN}| \ \|b\| \ <= \ \|\mathbb{A}b\| \ <= \ |\lambda_{MAX}| \ \|b\|$

b) $|\lambda_{MIN}| \ \|b\|^2 \ <= \ \|b^t \mathbb{A} b\| \ <= \ |\lambda_{MAX}| \ \|b\|^2$

2.- SEA $\mathbb{A}$ UNA MATRIZ DE $n \times m$ CON RAN($\mathbb{A}$) = $m < n$ , ENTONCES $\mathbb{A}^t \mathbb{A}$ ES POSITIVA DEFINIDA Y $\mathbb{A}\mathbb{A}^t$ ES SEMIPOSITIVA DEFINIDA.

3.- SEA $\mathbb{A}$ UNA MATRIZ CUADRADA SIMETRICA Y POSITIVA DEFINIDA, ENTONCES:

$$\mathbb{A} = \mathbb{S}^t \mathbb{D} \mathbb{S}$$

CON $\mathbb{S}^t \mathbb{S} = I$ Y $\mathbb{D}$ UNA MATRIZ DIAGONAL CON ELEMENTOS TODOS POSITIVOS.

A LA MATRIZ $\mathbb{S}^t \mathbb{D} \mathbb{S}$ SE LE CONOCE COMO UNA DESCOMPOSICION ESPECTRAL DE $\mathbb{A}$

## APENDICE II

1.- SI $W \sim N_p(\theta,\Sigma)$, ENTONCES:

$W^t \mathbb{A} W \sim \quad (p,\lambda) \quad \Leftrightarrow \quad \mathbb{A}\Sigma$ ES IDEMPOTENTE

CON $RAN(\mathbb{A}) = p$ Y $\lambda = (\theta^t \mathbb{A} \theta)/2$

2.- SI $W \sim N_p(\theta,\Sigma)$

CUALQUIER SUBCONJUNTO DE $s <= p$ VARIABLES DE $W$, TIENE UNA DISTRIBUCION NORMAL CON EL VECTOR MEDIA Y MATRIZ DE VARIANZAS-COVARIANZAS CORRESPONDIENTES

3.- SI $W \sim \quad (p,\lambda)$ Y $Z \sim \quad (q)$ Y SON INDEPENDIENTES, ENTONCES:

$U = [W/p]/[Z/q] \sim F(p,q,\lambda)$

4.- SEAN $\mathbb{A}$ Y $\mathbb{B}$ MATRICES DE $p \times p$ Y $q \times p$ RESPECTIVAMENTE, SI SI $Y \sim N_p(\mu,\sigma^2 I)$, ENTONCES LA FORMA LINEAL $\mathbb{B}Y$ ES INDEPENDIENTE DE LA FORMA CUADRATICA $Y^t \mathbb{A} Y$, SI $\mathbb{B}\mathbb{A} = 0$

# APENDICE III

## FUNCION DE MINIMOS CUADRADOS

CUANDO ESTABLECEMOS UN MODELO FUNCIONAL PARA ESTUDIAR UN FENOMENO, GENERALMENTE HACEMOS DOS SUPUESTOS FUERTES A SABER:

1) QUE EXISTE UNA FUNCION CON UN NUMERO DE ARGUMENTOS, QUIZA NO FINITO, QUE DETERMINA A UNA VARIABLE RESPUESTA DEL FENOMENO, Y. ES DECIR:

$Y = g(\xi;b)$ DONDE $\xi$ ES UN VECTOR FINITO DE VARIABLES CONTROLADAS Y $b$ UN VECTOR, SEGURAMENTE NO FINITO, DE VARIABLES NO CONTROLADAS.

UNA SUPOSICION ESTRUCTURAL, QUE SE REFIERE A QUE LA FUNCION $g$ PUEDE DESCOMPONERSE ADITIVAMENTE EN DOS FUNCIONES, COMO SIGUE:

$$Y = g(\xi;b) = f(\xi) + \varepsilon(b)$$

CON ESTA SUPOSICIONES, TENEMOS QUE PARA LA FUNCION $f$ CONOCEMOS SUS ARGUMENTOS AUNQUE NO SU EXPRESION NI EL VALOR DE LOS PARAMETROS QUE INVOLUCRA. EN CAMBIO LA FUNCION $\varepsilon$ ES TOTALMENTE ALEATORIA.

PARA LA FUNCION $f$ PUEDE SUGERIRSE UNA EXPRESION, QUE NORMALMENTE SE HACE A TRAVES DE LOS DATOS PROVENIENTES EN UNA MUESTRA ALEATORIA Y POR LOS RESULTADOS OBTENIDOS EN ESTUDIOS HECHOS SOBRE FENOMENOS SIMILARES AL QUE TRATAMOS. EN CUANTO A LA ELECCION DE SUS PARAMETROS, RESULTA RAZONABLE HACERLO BUSCANDO QUE EL EFECTO DE LA FUNCION $\varepsilon(b)$ SEA LO MAS REDUCIDO POSIBLE. BUSCARIAMOS PUES QUE $\varepsilon = Y-f(\xi)$ FUESE LO MAS "PEQUENA" POSIBLE.

RECORDEMOS QUE LO QUE TENEMOS ES UNA MUESTRA ALEATORIA, ES DECIR TENEMOS:

$$\varepsilon_i = Y_i - f_i(\xi) \qquad i=1,2,\ldots,n$$

PODRIAMOS ENTONCES PENSAR POR EJEMPLO EN MINIMIZAR $\sum_{i=1}^{n}\varepsilon_i$ , PERO LO ANTERIOR NO SERIA MUY CONVENIENTE PUES QUIZA $\sum_{i=1}^{n}\varepsilon_i = 0$ Y NOS DARIA UNA IDEA ALEJADA DE LA REALIDAD.

OTRA SUGERENCIA SERIA TRABAJAR CON $\Sigma_{i=1}^{n}|\varepsilon_i|$,PERO RECORDEMOS QUE $\varepsilon_i$ SON VARIABLES ALEATORIAS Y QUE EN UN PROCESO DE OPTIMIZACION VIA LAS TECNICAS DEL CALCULO DIFERENCIAL,HABRIA QUE UTILIZAR DERIVADAS,Y QUE EN EL CASO DEL VALOR ABSOLUTO DE VARIABLES ALEATORIAS RESULTA PROBLEMATICO.

SURGE ENTONCES LA OPCION DE LA FUNCION NORMA,ES DECIR TRATAR DE MINIMIZAR $(\Sigma_{i=1}^{n}|\varepsilon_i|^{p})^{1/p}$.EVIDENTEMENTE $p=1$ QUEDA DESCARTADO,LUEGO ENTONCES Y CON EL FIN DE FACILITAR CALCULOS PODEMOS ELEGIR $p=2$ Y TRABAJAR CON $(\Sigma_{i=1}^{n}|\varepsilon_i|^{2})^{1/2}$ O EQUIVALENTEMENTE,CON $\Sigma_{i=1}^{n}\ \varepsilon_i^{2}$.

ES ASI,COMO LA FUNCION $S(\theta) = \Sigma_{i=1}^{n}\ \varepsilon_i^{2}$ DONDE $\theta$ ES EL VECTOR DE PARAMETROS QUE INVOLUCRA $f$,ES ELEGIDA PARA SER OPTIMIZADA Y ESTIMAR EL VALOR DE $\theta$ QUE SEA MAS CONVENIENTE EN EL SENTIDO DE QUE $\Sigma_{i=1}^{n}\ \varepsilon_i^{2}$ SEA MINIMO.

A LA FUNCION $S(\theta)$ SE LE CONOCE COMO FUNCION DE MINIMOS CUADRADOS Y ES DE USO UNIVERSAL EN PROCESOS DE ESTIMACION PARAMETRICA.

BIBLIOGRAFIA

1.- BATES,DOUGLAS M. Y WATTS,DONALD G. (1980). RELATIVE CURVATURE MEASURES OF NONLINEARITY. J.R. STATIST. SOC. 42 No.1 , p.p. 1-25.

2.- BARD,YONATHAN. (1974). NONLINEAR PARAMETER ESTIMATION. NEW YORK,SAN FRANCISCO AND LONDON. ACADEMIC PRESS.

3.- BEALE, E.M.L. (1960). CONFIDENCE REGIONS IN NON-LINEAR ESTIMATION. J. ROY. STATIST. SOC. B-22, p.p. 41-76.

4.- CURRY,HASKELL B. (1944). THE METHOD OF STEEPEST DESCENT FOR NON-LINEAR MINIMIZATION PROBLEMS. Q. APPL. MATH. 2, p.p. 258-261.

5.- DAHIYA,RAM C. AND GURLAND,JOHN. (1972). PEARSON CHI-SQUARED TEST OF FIT WITH RANDOM INTERVALS. BIOMETRIKA 59, 1, p.p. 147-153.

6.- DRAPER,N.R. AND SMITH,H. (1981). APPLIED REGRESSION ANALYSIS. JOHN WILEY AND SONS, INC.

7.- GALLANT,RONALD A. (1975). TESTING A SUBSET OF PARAMETERS OF A NONLINEAR REGRESSION MODEL. J. AM. STATIST. ASSOC. 70, p.p. 927-932.

8.- GALLANT,RONALD A. (1975). NONLINEAR REGRESSION. J. AM. STATIST. ASSOC. 20, No.2, p.p. 73-81.

9.- GRAYBILL. (1963). INTRODUCTION TO LINEAR STATISTICAL MODELS. VOLUME 1.

10.- GUTTMAN,I. ; PEREYRA,V. AND SCOLNIK,H.D. (1973). LEAST SQUARES ESTIMATION FOR A CLASS OF NON-LINEAR MODELS. TECHNOMETRICS, 15, p.p. 209-218.

11.- HALPERIN,M. (1962). CONFIDENCE INTERVAL ESTIMATION IN NON-LINEAR REGRESSION. APPLIED MATHEMATICS DEPARTMENT SRRC-RR-62-28, SPERRY-RAND RESEARCH CENTER.

12.- MARQUARDT, D.W. (1963). AN ALGORITHM FOR LEAST SQUARES ESTIMATION IN NON-LINEAR PARAMETERS. J. SOC. APPL. MATH. 2, p.p. 431-441.