

UNIVERSIDAD NACIONAL DE INGENIERIA

FACULTAD DE CIENCIAS

ESCUELA PROFESIONAL DE MATEMATICAS



“ALGORITMO DE DIRECCIONES CONJUGADAS”

T E S I S

**Para Optar el Título Profesional de
Licenciatura en Matemática**

Presentado por :

WILLIAM CARLOS ECHEGARAY CASTILLO

Lima - Perú

1990

INDICE

Introducción	1.
CAPITULO I	
Preliminares	3.
Convexidad	3.
Programa matemático	5.
CAPITULO II	
Algoritmo General de la Programación Matemática	9.
Correspondencia	9.
El teorema de la Convergencia Global	15.
Velocidad de convergencia	20.
Velocidad de convergencia en \mathbb{R}^n	28.
CAPITULO III	
Minimización de una Función en un Intervalo	29.
Introducción	29.
Método de Sección Dorada	29.
Método de Interpolación Cuadrática	34
Convergencia global del método de Interpolación cuadrática.	35
CAPITULO IV	
Los Algoritmo de Descenso más rápido	40
El caso Cuadrático	44
CAPITULO V	
Método de las Direcciones Conjugadas	51
Método de Gram-schmidt	54
Método de Hestenes y Stiefel	58
Aplicación del Método de los Gradientes	
Conjugados a la Minimización no Cuadrática	60
Método de Fletcher-Reeves	61
Convergencia	62
Método de Daniel	67
Método de PARTAN	67

Método de Direcciones Conjugadas que no usa derivadas	68
Método quasi-Newton	72
Construcción de la Inversa	72
Corrección de Rango Uno	73
Método de Fletcher-Powell-Davidon	75
Convergencia	77
Método de los Gradientes Conjugados como Proceso Optimo	79
Cotas para la Convergencia	81
Método Parcial de Gradientes Conjugados	82

INTRODUCCION

El tema principal del presente trabajo es el estudio de una clase de algoritmos de la programación matemática : LOS ALGORITMOS DE DIRECCIONES CONJUGADAS.

En los capítulos I, II y III se presentan los conceptos y fundamentos teóricos para el análisis de tales algoritmos. Cabe destacar la definición del concepto general de un algoritmo iterativo para resolver programas matemáticos. Si $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una correspondencia y $x_0 \in \mathbb{R}^n$ es dado, tal algoritmo genera sucesiones $\{ x_k \}$ con la propiedad que:

$$x_{k+1} \in A(x_k), \quad \forall k \in \mathbb{Z}^+$$

Se introduce el concepto de las correspondencias cerradas y se establecen algunas propiedades fundamentales de esta clase de correspondencias, para demostrar un teorema general de convergencia global para los algoritmos iterativos mencionados. También se analizan los diferentes conceptos de velocidad de convergencia.

En el capítulo IV se analizan los algoritmos de descenso más rápido para resolver el programa matemático:

$$P : \min \{ f(x) / x \in \mathbb{R}^n \}$$

En la descripción usual, estos algoritmos construyen una sucesión $\{ x_k \}$, tal que:

$$x_{k+1} = x_k + \lambda_k t_k, \quad \lambda_k > 0, \quad \forall k$$

Si $\|\cdot\|$ es una norma dada en \mathbb{R}^n , entonces t_k es una dirección de descenso más rápido para f en x_k a la norma dada; es decir: t_k es una solución óptima del programa:

$$P_k : \min \{ \nabla f(x_k)^T t / \|t\| \leq 1 \}$$

Si $\|\cdot\|$ es la norma euclidiana en \mathbb{R}^n , entonces:

$$t_k = - \frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}$$

El hecho de que en, este caso, t_k es la dirección de descenso más rápido de f en x_k en el sentido tradicional de esta clase de algoritmos.

Por último, en el capítulo V, se analizan los algoritmos de las direcciones conjugadas para resolver el programa :

$$P : \min \{ Q(x) \mid x \in \mathbb{R}^n \}$$

donde Q es una función cuadrática :

$$Q(x) = \frac{1}{2} x^T Cx + p^T x$$

$C \in M(n,n)$ es matriz simétrica definida positiva ; $p \in \mathbb{R}^n$.

Se desea encontrar \hat{x} que sea solución óptima de P .

Enseguida se define el concepto de vectores conjugados.

Si se tiene vectores $d_i \neq 0$, $i = 1, 2, \dots, k$ conjugados por pares podemos extender estos vectores a una base $\{ d_1, d_2, \dots, d_n \}$ de vectores conjugados.

Elegimos x_0 arbitrario y calculamos :

$$x_{k+1} = x_k + \lambda_k d_k$$

donde λ_k es solución de :

$$\min \{ Q(x_k + \lambda d_k) \mid \lambda \in \mathbb{R} \}$$

Este algoritmo minimiza a Q en a lo más n pasos.

Sin embargo generalmente no se dispone a priori de una base d_i conjugados.

El algoritmo de gradiente conjugado incluye este algoritmo básico como subalgoritmo pero adicionalmente genera sucesivamente las direcciones conjugadas d_i requeridas.

CAPITULO I

PRELIMINARES

Aquí se resumen algunos conceptos y resultados de la teoría de convexidad y de la programación matemática, que serán usados en el desarrollo del trabajo.

1) CONVEXIDAD:

En esta sección, $\emptyset \neq X$ será un espacio lineal real.

Si $A, B \subset X$; $\lambda, \mu \in \mathbb{R}$, entonces definimos:

$$\lambda A + \mu B = \{ \lambda x + \mu y \mid x \in A, y \in B \}$$

Si M es un subespacio de X , $x_0 \in X$, entonces:

$$V = \{ x_0 \} + M \equiv x_0 + M$$

se llama un **SUBESPACIO AFIN** (un subespacio afín es un espacio trasladado).

Si $\emptyset \neq S \subset X$, entonces el **SUBESPACIO AFIN GENERADO POR S**, denotado por $V(S)$, es el menor subespacio afín que contiene a S .

Si \mathcal{X} es la familia de todos los subespacios afines que contienen a S , entonces:

$$V(S) = \bigcap \{ V \mid V \in \mathcal{X} \}$$

Si $\emptyset \neq K \subset X$, K se llama **CONVEXO**, si:

$$\alpha K + (1-\alpha)K \subset K, \forall \alpha \in [0,1]$$

Notar que \emptyset , los subespacios y los subespacios afines, son conjuntos convexos.

Es conocido que si $K, \emptyset \subset X$ son conjuntos convexos, entonces:

(i) $\alpha K = \{ \alpha y \mid y \in K \}$ es convexo, $\forall \alpha \in \mathbb{R}$.

(ii) $K + \emptyset = \{ k + g \mid k \in K, g \in \emptyset \}$ es convexo.

(iii) Si X es un espacio normado, entonces $K^\circ = \text{int}(K)$ (el interior de K) y \bar{K} (la clausura de K) son convexos.

(iv) La intersección de una familia cualquiera de conjuntos convexos es convexa.

Sea $D \subset X$ un subconjunto convexo.

La función $g : D \subset X \longrightarrow \mathbb{R}$ se llama una **FUNCIÓN CONVEXA**, si: $\forall x, y \in D, \forall \alpha \in [0,1]$ se cumple:

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha)g(y):$$

y se llama una **FUNCIÓN ESTRICTAMENTE CONVEXA**, si: $\forall x, y \in D, x \neq y, \forall \alpha \in [0,1]$ se cumple:

$$g(\alpha x + (1-\alpha)y) < \alpha g(x) + (1-\alpha)g(y)$$

El prototipo de una función convexa es $f(x) = x^T A x$, donde A es una matriz semidefinida positiva y simétrica.

Asumimos conocidas las siguientes propiedades de las funciones convexas:

1) Una combinación positiva de funciones convexas es convexa.

2) $x = \sum_{i=1}^m \alpha_i x_i$, se llama una **COMBINACION CONVEXA** de x_1, x_2, \dots, x_m , si $\alpha_i \geq 0, i = 1, 2, \dots, m$, y además:

...

$$g\left(\sum_{i=1}^m \alpha_i x_i\right) \leq \sum_{i=1}^m \alpha_i g(x_i)$$

3) El conjunto de puntos que satisfacen simultáneamente:

$f_1(x) \leq c_1, \dots, f_m(x) \leq c_m$, donde cada f_i es convexa, es convexo.

4) Sea $\Omega \subset \mathbb{R}^n$ convexo y $f \in C^1(\Omega)$. Entonces se cumple:

f es convexa, si y sólo si,

$$f(y) \geq f(x) + \nabla f(x)^T (y-x), \forall x, y \in \Omega$$

5) Sea $\Omega \subset \mathbb{R}^n$ convexo, $\Omega \neq \emptyset$ y $f \in C^2(\Omega)$. Entonces se cumple:

f es convexa en Ω , si y sólo si,

$$\nabla^2 f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)$$

es semidefinida positiva, $\forall x \in \Omega$.

6) Sea f convexa en Ω convexo, entonces:

$\Gamma_c(f) = \{ x \in \Omega / f(x) \leq c \}$ es convexo, $\forall c \in \mathbb{R}$.

A $\Gamma_c(f)$ se le llama un CONJUNTO DE NIVEL de f .

Si $\emptyset \neq A \subset X$, definimos:

$\text{Co}(A) = \text{CAPSULA CONVEXA DE } A$

= menor conjunto convexo que contiene a A

$\text{Co}(A)$ esta caracterizado por:

(i) $\text{Co}(A)$ es convexo y $\text{Co}(A) \supset A$

(ii) Si D es convexo y $D \supset A$, entonces $D \supset \text{Co}(A)$.

2) PROGRAMA MATEMATICO:

Sea $\emptyset \neq X$ un espacio topológico lineal, $C \subset \Omega \subset X$, $F: \Omega \longrightarrow \mathbb{R}$ continua. Consideramos el siguiente problema:

Determinar $\hat{x} \in C$, tal que:

$$F(x) \geq F(\hat{x}), \forall x \in C \quad (\text{ó tal que: } F(x) \leq F(\hat{x}), \forall x \in C)$$

Escribimos para este problema:

$$P : \min \{ F(x) / x \in C \} \quad (\text{ó } Q : \max \{ F(x) / x \in C \})$$

P (ó Q) se llama un PROGRAMA MATEMATICO, donde F es la funcion objetivo, C es el dominio admisible de P (ó Q),

$x \in C$ es un punto admisible de P (ó Q) ,

$\hat{x} \in C$ con $F(x) \geq F(\hat{x})$, $\forall x \in C$, [ó $F(x) \leq F(\hat{x})$, $\forall x \in C$]
es una solucion optima de P (ó Q) ,

$$\inf (P) = \inf \{ F(x) / x \in C \}$$

$$[\text{ ó } \sup (Q) = \sup \{ F(x) / x \in C \}]$$

es el valor del programa P (ó Q) .

$\hat{x} \in C$ es un minimo local de P , si existe un $\varepsilon > 0$ tal que

$$F(x) \geq F(\hat{x}), \forall x \in \overline{S(\hat{x}, \varepsilon)} \cap C$$

Con respecto a la existencia de soluciones de P , se presentan las siguientes situaciones:

(i) $\inf (P) = +\infty$; ($C \neq \emptyset$)

(ii) $\inf (P) = -\infty$; (en este caso, P no tiene solución óptima)

(iii) $-\infty < \inf (P) < +\infty$ y P no tiene solución óptima.

(iv) $-\infty < \inf (P) < +\infty$ y P tiene solución óptima.

La forma más frecuente del dominio admisible es:

$$C = \{ x \in X / f_j(x) \leq 0, j \in J, g_k(x) = 0, k \in K, x \in D \}$$

donde D se llama dominio fundamental. f_j se llama activa en \hat{x} , si $f_j(\hat{x}) = 0$.

Si C tiene la forma anterior, entonces:

$P : \min \{ F(x) / x \in C \}$ es un PROGRAMA CONVEXO, si D es

convexo, las funciones $F, f_j, j \in J$, son convexas en D , y

las funciones $g_k, k \in K$, son afín lineales.

Sea P un programa convexo; entonces se cumple que:

- 1) El dominio admisible C de P es convexo.
- 2) El conjunto de soluciones óptimas de P es convexo.
- 3) Todo mínimo local de F en C es una solución óptima de P .
- 4) Si F es estrictamente convexa, entonces P tiene a lo más una solución.

Es de notar que, si en P se sustituye \max por \min , el programa resultante no es convexo, pues este programa tiene, por lo general, máximos locales que no son soluciones óptimas.

La FUNCION DE LAGRANGE DE P , se denota por $L(x,u,v)$ y se la define mediante:

$$\begin{aligned} L(x,u,v) &= F(x) + \sum_{j=1}^m u_j f_j(x) + \sum_{k=1}^l v_k g_k(x) \\ &= F(x) + u^T f(x) + v^T g(x) \end{aligned}$$

Asimismo, se hará uso de los siguientes teoremas, cuyas demostraciones son conocidas.

TEOREMA: Sea P un programa convexo con $D = \mathbb{R}^n$, \hat{x} un punto admisible de P ; sean todas las funciones F, f_j, g_k diferenciables en \hat{x} . Si existen constantes $\hat{u}_j \geq 0, \hat{v}_k \in \mathbb{R}$, tal que:

$$\nabla L(\hat{x}, \hat{u}, \hat{v}) = \nabla F(\hat{x}) + \sum_{j=1}^m \hat{u}_j \nabla f_j(\hat{x}) + \sum_{k=1}^l \hat{v}_k \nabla g_k(\hat{x}) = 0$$

entonces \hat{x} es una solución óptima de P .

TEOREMA: Si $f \in C^1(\mathbb{R}^n)$ y f es convexa en \mathbb{R}^n entonces:

\hat{x} es un mínimo absoluto de f , si y sólo si, $\nabla f(\hat{x}) = 0$

CAPITULO II

A) UN ALGORITMO GENERAL DE LA PROGRAMACION MATEMATICA

CORRESPONDENCIA:

DEFINICION: Si X, Y son conjuntos, entonces una CORRESPONDENCIA A de X en Y , que se denota como $A: X \longrightarrow Y$ no es otra cosa que una aplicación $A: X \longrightarrow 2^Y$.

En consecuencia, una correspondencia $A: X \longrightarrow Y$, asocia a todo $x \in X$, un subconjunto $A(x)$ de Y .

Observamos que $A(x) = \emptyset$ es posible.

Al conjunto $\text{dom}(A) = \{ x \in X / A(x) \neq \emptyset \}$ se llama el DOMINIO EFECTIVO DE A.

Observamos también que toda aplicación $\varphi: X \longrightarrow Y$ es un caso especial de una correspondencia de X en Y .

Si $A: X \longrightarrow Y$ es una correspondencia, $D \subset X$, definimos:

$$A(D) = \bigcup_{x \in D} A(x)$$

El conjunto $\text{graf}(A) = \{ (x,y) \in X \times Y / x \in X, y \in A(x) \}$, se llama la GRAFICA DE LA CORRESPONDENCIA A.

DEFINICION: Si $A: X \longrightarrow Y$, $B: Y \longrightarrow Z$ son correspondencias, entonces la correspondencia $BoA: X \longrightarrow Z$ es definida por:

$$(BoA)(x) = \bigcup_{y \in A(x)} B(y)$$

En lo que sigue, todos los espacios X , Y , Z , etc., serán espacios métricos. Para nuestro trabajo, el concepto de una CORRESPONDENCIA CERRADA es muy importante.

DEFINICION: (CORRESPONDENCIA CERRADA)

La correspondencia $A: X \longrightarrow Y$ es CERRADA en $x \in X$, si (i), (ii) implican (iii):

$$(i) x_k \in X, \forall k, x_k \longrightarrow x$$

$$(ii) y_k \in A(x_k), \forall k, y_k \longrightarrow y$$

$$(iii) y \in A(x)$$

■

Si A es cerrada en todo punto $x \in X$, decimos que A es cerrada en X , o simplemente, que A es cerrada.

Ejemplo 1: Si la aplicación $\varphi: X \longrightarrow Y$ es continua en x , entonces φ es cerrada en x , puesto que en este caso (i) y

(ii) toman la forma:

$$(i) x_k \in X, \forall k, x_k \longrightarrow x$$

$$(ii) \varphi(x_k) \longrightarrow y$$

Como φ es continua en x , se debe cumplir: $y = \varphi(x)$. ■

Si la correspondencia $A : X \longrightarrow Y$ es cerrada en x , entonces $A(x)$ es cerrada en y .

Elegimos $x_k = x, \forall k$. Sea $y \in \overline{A(x)}$ cualquiera. Entonces existe una sucesión $\{y_k\} \subset A(x)$ tal que $y_k \longrightarrow y$.

Como A es cerrada en x , se cumple $y \in A(x)$.

Por consiguiente, $\overline{A(x)} \subset A(x)$; entonces $A(x)$ es cerrada en y .

Ejemplo 2: Sea la correspondencia $A : \mathbb{R}_+ \longrightarrow 2^{\mathbb{R}}$ definida por:

$$A(x) = \begin{cases} \{0\} & , \text{ si } x = 0 \\ \langle -x, x \rangle & , \text{ si } x > 0 \end{cases}$$

Como $\langle -x, x \rangle$ es abierto, A no es cerrada en ningún punto $x > 0$. Pero A es cerrada en $x = 0$, puesto que:

$$x_k \longrightarrow 0, x_k > 0, y_k \in A(x_k)$$

$$\text{implica que } y_k \longrightarrow 0 \in A(0) = \{0\}$$
 ■

LEMA 1: La correspondencia $A : X \longrightarrow Y$ es cerrada, si y sólo si, $\text{graf}(A)$ es cerrada en $X \times Y$.

- Demostración:

\Rightarrow Sea $(x, y) \in \overline{\text{graf}(A)}$. Entonces existe una sucesión

$$\{(x_k, y_k)\} \subset \text{graf}(A) \text{ tal que: } (x_k, y_k) \longrightarrow (x, y).$$

Entonces $x_k \longrightarrow x$, $y_k \in A(x_k)$, $\forall k$, $y_k \longrightarrow y$

Como A es cerrada en x , se tiene que $y \in A(x)$.

Entonces $(x,y) \in \text{graf}(A)$.

Luego $\text{graf}(A)$ es cerrada.

⇐) Sea $x \in X$ cualquiera, $x_k \longrightarrow x$, $y_k \in A(x_k)$, $\forall k$,

$y_k \longrightarrow y$.

Entonces $(x_k, y_k) \in \text{graf}(A)$, $(x_k, y_k) \longrightarrow (x, y)$.

Como $\text{graf}(A)$ es cerrada, se tiene $(x, y) \in \text{graf}(A)$

$y \in A(x)$.

Por consiguiente, A es cerrada en x . ■

Ejemplo: Definimos la aplicación $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$, por:

$$\varphi(x) = \begin{cases} 0 & , \text{ si } x = 0 \\ 1/x & , \text{ si } x \neq 0 \end{cases}$$

Como $\text{graf}(\varphi)$ es cerrada, φ es cerrada, pero evidentemente no es continua en 0.

PROPOSICION 1: (COMPOSICION DE LA CORRESPONDENCIA CERRADA)

Sean $A : X \longrightarrow Y$, $B : Y \longrightarrow Z$ correspondencias

tales que A es cerrada en x y B es cerrada en $A(x)$.

Supongamos además que, si $x_k \longrightarrow x$, $y_k \in A(x_k)$, $\forall k$,

entonces la sucesión $\{ y_k \}$ tiene al menos un punto de

acumulación.

Entonces $C = B \circ A$ es cerrada en x .

Demostración:

Sean $x_k \longrightarrow x$, $z_k \longrightarrow z$, con $z_k \in C(x_k)$, $\forall k$.

Debemos probar que $z \in C(x)$.

$z_k \in C(x_k) = \bigcup_{y \in A(x_k)} B(y)$, $\forall k$, entonces podemos elegir:

$y_k \in A(x_k)$, $\forall k$, tal que $z_k \in B(y_k)$, y luego, por hipótesis, existe $y \in Y$, y , una subsucesión $\{y_{k_i}\}$ con

$y_{k_i} \longrightarrow y$.

Como A es cerrada en x , tenemos $y \in A(x)$.

Debido a que $y_{k_i} \longrightarrow y$, $z_{k_i} \in B(y_{k_i})$, $z_{k_i} \longrightarrow z$, y como B

es cerrada en y , entonces:

$$z \in B(y) \subset B(A(x)) = \bigcup_{y \in A(x)} B(y) = C(x)$$

$$\Rightarrow z \in C(x).$$

Luego: $C = B \circ A$ es cerrada en x . ■

COROLARIO 1: Sean $A : X \longrightarrow Y$ una correspondencia cerrada

en x y $B : Y \longrightarrow Z$ una correspondencia cerrada en $A(x)$ e

Y compacto. Entonces $C = B \circ A$ es cerrada en x .

Demostración:

Si $x_k \longrightarrow x$, $y \in A(x_k)$, $\forall k$.

Como Y es compacto, entonces $\exists y \in Y$ una subsucesión

$\{y_{k_i}\}$ con $y_{k_i} \longrightarrow y$.

Entonces por la proposición anterior se tiene que:

$C = B \circ A$ es cerrada en x . ■

Corolario 2: Sean $A : X \longrightarrow Y$ una aplicación continua en x y $B : Y \longrightarrow Z$ una aplicación cerrada en $A(x)$.

Entonces $C = B \circ A$ es cerrada en x .

-Demostración:

Como A es continua en x , A es cerrada en x y

$x_k \longrightarrow x$ implica que $y_k = A(x_k) \longrightarrow y = A(x)$

Entonces por la proposición anterior se tiene que:

C es cerrada en x ■

Consideremos, ahora, problemas de optimización de la forma:

$$P : \min \{ F(x) / x \in X \subset \mathbb{R}^n \}$$

La mayor parte de los algoritmos para resolver P , se dejan describir de la siguiente forma:

Se elige $x_0 \in \mathbb{R}^n$, se construye una correspondencia

$A: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ y una sucesión $\{x_k\} \subset \mathbb{R}^n$ tal que:

$$x_{k+1} \in A(x_k), \forall k \in \mathbb{Z}_+ \quad (1)$$

A tal correspondencia A se le llama un ALGORITMO y una sucesión con la propiedad (1) se llama SUCESION GENERADA por el algoritmo A .

Se elige un subconjunto $D \subset X$, que se llamará CONJUNTO DE SOLUCIONES. Generalmente D será el conjunto de todos los

puntos $x \in X$, que satisfacen una determinada condición de optimidad necesaria o suficiente.

$D = \{x \in X \mid \text{se satisface una determinada condición de optimidad necesaria o suficiente}\}$.

Así, por ejemplo, si $\forall x, y \in X$, existe la derivada direccional $F'(x, y-x)$, se elige con frecuencia:

$$D = \{x \in X \mid F'(x, y-x) \geq 0, \forall y \in X\}$$

La función $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ continua, se llamará una **FUNCIÓN DE DESCENSO** para A y D , si se cumple:

i) Si $x \notin D$, $y \in A(x)$, entonces $\Phi(y) < \Phi(x)$.

ii) Si $x \in D$, $y \in A(x)$, entonces $\Phi(y) \leq \Phi(x)$.

Con frecuencia, se **elige** una de las siguientes funcionales Φ como una función de descenso:

$$\Phi(x) = F(x), \quad \Phi(x) = \|\nabla F(x)\|, \quad \Phi(x) = \text{dist}(x, D).$$

B) EL TEOREMA DE LA CONVERGENCIA GLOBAL

Sea $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ un algoritmo, D un conjunto de soluciones, y Φ una función de descenso para A y D .

Sea $\{x_k\}$ una sucesión generada por A . Asumimos que A es cerrada para todo $x \in \mathbb{R}^n \setminus D$ y $\{x_k\} \subset E$, donde E es compacto; entonces todo punto de acumulación x^* de $\{x_k\}$ se encuentra en D .

Demostración:

Sea $\{x_{k_j}\}$ una subsucesión convergente ($\lim_{j \rightarrow \infty} x_{k_j} = x$)

Como Φ es una función continua de descenso, se tiene:

$$\lim_{j \rightarrow \infty} \Phi(x_{k_j}) = \Phi(x) \quad \dots\dots(*)$$

$$\Phi(x) \leq \Phi(x_{k_{j+1}}) \leq \Phi(x_{k_j}), \quad \forall k \quad \dots\dots(**)$$

De (*) y (**) obtenemos:

$$\lim_{k \rightarrow \infty} \Phi(x_k) = \Phi(x)$$

Probaremos que $x \in D$:

Supongamos que $x \in \mathbb{R}^n \setminus D$, y consideremos la sucesión

$\{x_{k_{j+1}}\} \subset E$. Como E es compacto, existe una subsucesión

$$\{x_{\bar{k}_{j+1}}\} \subset \{x_{k_{j+1}}\} / \lim_{\bar{k}_j \rightarrow \infty} x_{\bar{k}_{j+1}} = \bar{x}$$

Se cumple entonces que:

$$x_{\bar{k}_j} \longrightarrow x$$

$$x_{\bar{k}_{j+1}} \in A(x_{\bar{k}_j}), \quad \forall \bar{k}_j$$

$$x_{k_j+1} \longrightarrow \bar{x}$$

Como A es cerrada en $\mathbb{R}^n \setminus D$, $\bar{x} \in A(x)$; y así $\Phi(\bar{x}) < \Phi(x)$.

Debido a la continuidad de Φ tenemos:

$$\Phi(x) = \lim_{j \rightarrow \infty} \Phi(x_{k_j}) = \lim_{k \rightarrow \infty} \Phi(x_k) = \lim_{k_j \rightarrow \infty} \Phi(x_{k_j+1}) = \Phi(\bar{x})$$

$$\rightarrow \Phi(x) = \Phi(\bar{x}) \quad (\text{contradicción})$$

Por tanto: $x \in D$. ■

Este teorema será usado para establecer la convergencia de varios algoritmos. En muchos aspectos, la condición del teorema de que A es cerrada fuera de D ; es la condición más importante. Para ver esto, consideramos el siguiente ejemplo:

Ejemplo 1:

Consideremos en \mathbb{R} , el algoritmo:

$$A(x) = \begin{cases} \frac{x-1}{2} + 1 & , x \geq 1 \\ \frac{x}{2} & , x < 1 \end{cases} \quad \text{y } D = \{0\}$$

En este caso, una función de descenso es $\Phi(x) = |x|$

Sea $\{x_k\}$ una sucesión / $x_k > 1, \forall k$.

Probaremos que $\lim_{k \rightarrow \infty} x_k = 1$:

$$x_k > 1, \forall k$$

$$\Rightarrow x_{k+1} = A(x_k) = \frac{x_k - 1}{2} + 1 = \frac{x_k + 1}{2} < x_k, \forall k \dots (*)'$$

$$\Rightarrow 1 < x_{k+1} < x_k, \forall k$$

Luego $\{x_k\}$ es una sucesión monótona decreciente y acotada,

por tanto, existe $l = \lim_{k \rightarrow \infty} x_k$, y por $(*)'$, $l = 1$, la

cual no es una solución, dado que $1 \notin D$.

Esto se debe a que A no es cerrada en $\mathbb{R} \setminus \{0\}$; puesto que si $x_k < 1$, $x_k \rightarrow 1$, entonces:

$$A(x_k) \rightarrow 1/2 \neq 1 = A(1)$$

Ejemplo 2:

Sobre la recta real X , con $D \neq \emptyset$, $\Phi(x) = e^{-x}$,
 $A(x) = x + 1$. Todas las condiciones del teorema de convergencia global se cumplen, salvo la primera.

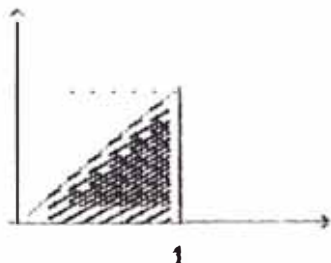
Para cualquier punto inicial x_0 , la sucesión $\{x_k\}$ generada por A diverge a $+\infty$. Pero el teorema exige que $\{x_k\} \subset E$, donde E es compacto. En este caso, $\{x_k\}$ diverge, entonces la sucesión no puede tener una subsucesión convergente.

Ejemplo 3:

Consideremos el algoritmo A dado explícitamente en $X = [0,1]$ por:

$$A(x) = \begin{cases} [0, x) & , 0 < x \leq 1 \\ 0 & , x = 0 \end{cases}$$

cuya gráfica es :



Si elegimos $D = \{0\}$, entonces $\Phi(x) = x$ es una función de descenso para A y D ; por que, para $x \neq 0$ todos los puntos en $A(x)$ son menores que x ; es decir,

$$\Phi(y) = y < x, \forall y \in A(x) = [0, x)$$

La sucesión definida por :

$$x_0 = 1, \quad x_{k+1} = x_k - \frac{1}{2^{k+2}}$$

satisface $x_{k+1} \in A(x_k), \forall k$, dado que $x_{k+1} < x_k$.

En consecuencia $\{x_k\}$ es una sucesión generada por el algoritmo A .

Pero :

$$x_{k+1} = 1 - \frac{1}{4} \cdot \frac{1 - (\frac{1}{2})^{k+1}}{\frac{1}{2}}$$

$$\Rightarrow \lim_{k \rightarrow \infty} x_k = \frac{1}{2} \notin D = \{0\}.$$

Por tanto, $\{x_k\}$ no tiene punto de acumulación en D . Esto es debido a que A no es cerrada fuera de D .

VELOCIDAD DE CONVERGENCIA

Aparte de la convergencia global de un algoritmo, hay dos aspectos importantes: su velocidad de convergencia y su estabilidad numérica.

Trataremos brevemente la primera.

El estudio de la velocidad de convergencia es mucho más importante que el análisis de la convergencia global y no es posible resumir los principales resultados en un solo teorema, como el caso de la convergencia global.

Sin embargo, existe una teoría que permite predecir con certeza la eficiencia relativa de una amplia clase de algoritmos.

Vamos a introducir varios conceptos que sirven para medir la velocidad de convergencia.

Orden de convergencia:

Consideremos una sucesión de números reales $\{r_k\}$ que converge a r^* .

Definición:

Sea $\lim_{k \rightarrow \infty} r_k = r^*$. El orden de convergencia de $\{r_k\}$ es

$$p = \sup \left\{ q \geq 0 \mid \lim_{k \rightarrow \infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^q} < \infty \right\} \quad (1)$$

Nota:

Para sucesiones $\{r_k\}$, generadas por el algoritmo con

$\lim_{k \rightarrow \infty} r_k = r^*$, se cumplen generalmente a) o b):

a) $r_k \neq r^*$, $\forall k \in \mathbb{N}$.

b) $r_k = r^*$ \Rightarrow $r_l = r^*$, $\forall l \geq k$.

Por esta razón (1) está bien definido, si adicionalmente consideramos a la expresión $0/0$ como un número finito.

Cabe destacar que el orden de convergencia, así como los otros conceptos de velocidad de convergencia, dependen sólo de las propiedades de la sucesión que se cumplen en el límite (para $k \rightarrow \infty$).

Diremos que el orden de convergencia depende sólo de la "cola de la sucesión". En este lenguaje, el orden de convergencia mide la calidad de convergencia de la peor parte de la cola.

Si el orden de convergencia de $\{r_k\}$ es p y:

$$\lim_{k \rightarrow \infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p} = \beta < \infty$$

existe (este caso se presenta generalmente), entonces se cumple asintóticamente:

$$|r_{k+1} - r^*| \cong \beta |r_k - r^*|^p$$

Esto significa que r_k se aproxima a r^* , tanto más rápidamente, cuanto mayor es p .

Ejemplo 1:

Si $0 < a < 1$, entonces la sucesión $\{r_k = a^k\}$ converge a 0, con orden 1; puesto que $\frac{a^{k+1}}{a^k} = a$ y para $p > 1$, se

cumple :

$$\lim_{k \rightarrow \infty} \frac{a^{k+1}}{a^{kp}} = \lim_{k \rightarrow \infty} a^{k(1-p)+1} = a^{-\infty} = 0$$

Ejemplo 2:

La sucesión $\{ r_k = a^{2^k} \}$ converge a cero, con orden 2, si $0 < a < 1$.

Veamos:

$$\lim_{k \rightarrow \infty} \frac{a^{2^{k+1}}}{a^{p2^k}} = \lim_{k \rightarrow \infty} a^{(2-p)2^k} = \begin{cases} 0, & 0 \leq p < 2 \\ 1, & p = 2 \\ +\infty, & p > 2 \end{cases}$$

entonces el orden de convergencia es $p = 2$.

La convergencia lineal:

Muchos algoritmos tienen orden de convergencia $p = 1$. Por esta razón, analizamos este caso con más detalle.

Definición:

$$\text{Si } \lim_{k \rightarrow \infty} r_k = r^* \text{ y } \lim_{k \rightarrow \infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|} = \beta < 1,$$

decimos que $\{ r_k \}$ converge linealmente a r^* con razón de convergencia β . ■

Podemos afirmar que una sucesión que converge linealmente con razón β tiene una cola, tal que $|r_k - r^*|$ converge al menos tan rápidamente a cero, como la sucesión geométrica $c\beta^k$ para cierto $c > 0$.

Por esta razón, se habla también de convergencia geométrica en lugar de convergencia lineal.

Si comparamos dos sucesiones que convergean linealmente con razones $0 < \beta_1 < \beta_2 < 1$, entonces la sucesión con razón β_1 converge más rápidamente que la otra.

Definición:

$$\text{Si } \lim_{k \rightarrow \infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|} = 0, \text{ diremos que la convergencia de}$$

$\{ r_k \}$ es superlineal. ■

Si $\{r_k\}$ tiene orden de convergencia $p > 1$, entonces $\{r_k\}$ converge de manera superlineal:

Veamos :

$$\text{Sea } \lim_{k \rightarrow \infty} r_k = r^*, \quad p > 1 \text{ y } \lim_{k \rightarrow \infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p} = c < \infty$$

entonces $\exists k_0 \in \mathbb{N}$ tal que:

$$\begin{aligned} \frac{|r_{k+1} - r^*|}{|r_k - r^*|^p} &\leq 2c, \quad \forall k \geq k_0 \\ \Rightarrow \frac{|r_{k+1} - r^*|}{|r_k - r^*|} &\leq 2c |r_k - r^*|^{p-1} \\ \Rightarrow \lim_{k \rightarrow \infty} \frac{|r_{k+1} - r^*|}{|r_k - r^*|} &= 0. \blacksquare \end{aligned}$$

De otro lado, es posible que una sucesión que converge de manera superlineal, tenga orden de convergencia $p = 1$.

Ejemplo 3:

La sucesión $\{r_k = \frac{1}{k}\}$ converge a cero con orden $p = 1$.

Veamos:

$$\lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k^p} = \lim_{k \rightarrow \infty} \frac{k^p}{k+1} = \lim_{k \rightarrow \infty} \frac{k^{p-1}}{1 + \frac{1}{k}} = \begin{cases} 0 & , 0 \leq p < 1 \\ 1 & , p = 1 \\ \infty & , p > 1 \end{cases}$$

el orden de convergencia es $p = 1$ pero la convergencia no es lineal.

Ejemplo 4:

La sucesión $\{r_k = (\frac{1}{k})^k\}$ tiene orden de convergencia 1 y converge a cero, de manera superlineal.

Veamos:

$$0 < \frac{r_{k+1}}{r_k} = \frac{k^k}{(k+1)^{k+1}} = \left(\frac{1}{k+1}\right) \left(\frac{k}{k+1}\right)^k \leq \frac{1}{k+1}$$

$$\Rightarrow \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k} = 0.$$

Sea $p > 1$. Entonces:

$$\frac{r_{k+1}}{r_k^p} = \frac{k^{kp}}{(k+1)^{k+1}} = \exp[kp \ln(k) - (k+1) \ln(k+1)] =$$

$$= \exp(k \ln(k) \left[p - \frac{k+1}{k} \frac{\ln(k+1)}{\ln(k)} \right])$$

$$= \exp[k \ln(k) \psi(k)],$$

donde $\psi(k) = p - \frac{k+1}{k} \frac{\ln(k+1)}{\ln(k)} = 1 \Rightarrow \lim_{k \rightarrow \infty} \psi(k) = p - 1 > 0$

$$\Rightarrow \lim_{k \rightarrow \infty} k \ln(k) \psi(k) = +\infty.$$

$$\Rightarrow \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k^p} = +\infty.$$

$\Rightarrow r_k \rightarrow \infty$ con orden de convergencia $p = 1$ y converge de manera superlineal. ■

Tasas medias de convergencia :

Las definiciones del párrafo precedente se refieren a una sola iteración (a un solo paso), puesto que establecen cotas para el progreso que se logra en una sola iteración (pasando de k a $k+1$).

Vamos ahora a definir un concepto que mide el promedio del progreso por paso que se logra en un gran número de pasos (iteraciones).

Definición :

Consideremos una sucesión $\{ r_k \}$ que converge a

r^* .

El orden medio P_m de convergencia de $\{r_k\}$ se define:

$$P_m = \inf \left\{ p > 1 \mid \overline{\lim}_{k \rightarrow \infty} |r_k - r^*|^{1/p^k} = 1 \right\}$$

Si no existe $p > 1$, tal que este limite sea 1, definimos $P_m = \infty$. ■

Ejemplo 5:

Para la sucesión $\{r_k = a^{2^k}\}$, $a \in \langle 0,1 \rangle$, se cumple:

$$\lim_{k \rightarrow \infty} r_k = \lim_{k \rightarrow \infty} a^{2^k} = 0.$$

$$\Rightarrow |r_k|^{1/p^k} = a^{(2/p)^k}, \quad \lim_{k \rightarrow \infty} \left(\frac{2}{p}\right)^k = \begin{cases} +\infty, & 1 < p < 2 \\ 1, & p = 2 \\ 0, & p > 2 \end{cases}$$

$$\Rightarrow \lim_{k \rightarrow \infty} |r_k|^{1/p^k} = \begin{cases} 0, & 1 < p < 2 \\ a, & p = 2 \\ 1, & p > 2 \end{cases}$$

$$P_m = \inf \left\{ p > 1 \mid \overline{\lim}_{k \rightarrow \infty} |r_k|^{1/p^k} = 1 \right\} = 2$$

$$\Rightarrow P_m = 2. \quad \blacksquare$$

Ejemplo 6:

Para la sucesión $\{r_k = a^k\}$, $a \in \langle 0,1 \rangle$ se cumple:

$$\lim_{k \rightarrow \infty} r_k = 0.$$

$$\Rightarrow |r_k|^{1/p^k} = a^{k/p^k}, \quad \lim_{k \rightarrow \infty} \frac{k}{p^k} = 0, \quad \forall p > 1.$$

$$\Rightarrow \lim_{k \rightarrow \infty} |r_k|^{1/p^k} = a^0 = 1, \quad p > 1$$

$$\Rightarrow P_m = 1. \quad \blacksquare$$

Observamos que para los ejemplos 5 y 6 se cumple : $p = P_m$.
 Como en el párrafo precedente el caso $P_m = 1$ es el más interesante.

Para este caso, se define como razon media de convergencia a

$$\beta_m = \overline{\lim}_{k \rightarrow \infty} |r_k - r^*|^{1/k}.$$

Para la sucesión geométrica $r_k = ca^k$, $0 < a < 1$, se cumple:

$$\lim_{k \rightarrow \infty} r_k = 0$$

Entonces:

$$\begin{aligned} |r_k|^{1/k} &= (ca^k)^{1/k} = c^{1/k} a, \\ \Rightarrow \lim_{k \rightarrow \infty} |r_k|^{1/k} &= a \lim_{k \rightarrow \infty} c^{1/k} = a. \\ &\rightarrow \beta_m = a. \blacksquare \end{aligned}$$

Si $\beta_m < 1$ [$\beta_m = 0$], hablamos de convergencia lineal [superlineal] en promedio. Nos restringiremos mayormente a los conceptos que se refieren a un solo paso, pues para algoritmos iterativos es natural comparar un paso con el próximo. Además, si las sucesiones se comportan bien y en las definiciones, los límites existen, entonces generalmente se cumplirá: $p = P_m$, $\beta = \beta_m$.

VELOCIDAD DE CONVERGENCIA EN \mathbb{R}^n

Sea $\{x_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ una sucesión convergente. Las propiedades de convergencia de tal función, se define con respecto a una función determinada, que convierte $\{x_k\}$ en una sucesión numérica. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función continua dada, entonces $x_k \rightarrow x^*$ implica que $f(x_k) \rightarrow f(x^*)$, y las propiedades de convergencia de x_k se obtienen analizando la convergencia de $f(x_k)$ a $f(x^*)$.

La función usada de este modo, para medir la convergencia de $\{x_k\}$, se llama función de error.

Si $\{x_k\}$ es una sucesión generada por un algoritmo para resolver un programa matemático:

$$P : \min \{ F(x) / x \in S \}$$

Se elige con frecuencia la función objetivo F como función de error. Si $\|\cdot\|$ es una norma determinada, se usa también $g(x) = \|x - x^*\|^2$, ó , $f(x) = \|x - x^*\|$, como función de error.

Generalmente, el orden de convergencia de una sucesión no depende de la función de error elegida. Pero, en el caso de la convergencia lineal, la razón de convergencia β , sí depende de la función de error.

Proposición:

Sean f y g dos funciones de error, con $g(x^*) = f(x^*) = 0$, tales que existen constantes $0 < a_1 \leq a_2$ de manera que:

$$a_1 g(x) \leq f(x) \leq a_2 g(x), \forall x \in \mathbb{R}^n \dots(*)$$

Si la sucesión $\{x_n\}$ converge linealmente a x^* con razón media β_m con respecto a una de las dos funciones, entonces lo hace también con respecto a la otra función.

-Demostración:

De (*) se obtiene:

$$\frac{1}{a_2} f(x) \leq g(x) \leq \frac{1}{a_1} f(x), \forall x \in \mathbb{R}^n \dots(**)$$

(**) es simétrica con respecto a f y g .

Supongamos que $\lim_{k \rightarrow \infty} x_k = x^*$ linealmente con razón media β_m respecto a f .

Entonces:

$$\beta_m = \overline{\lim}_{k \rightarrow \infty} |f(x_k) - f(x^*)|^{1/k} = \overline{\lim}_{k \rightarrow \infty} |f(x_k)|^{1/k} \quad (f(x^*) = 0)$$

$$\begin{aligned}
\beta_m &= \overline{\lim}_{k \rightarrow \infty} |f(x_k)|^{1/k} \leq \overline{\lim}_{k \rightarrow \infty} |a_2 g(x_k)|^{1/k} = \\
&= \lim_{k \rightarrow \infty} |a_2|^{1/k} \overline{\lim}_{k \rightarrow \infty} |g(x_k)|^{1/k} = \overline{\lim}_{k \rightarrow \infty} |g(x_k)|^{1/k}, \\
\beta_m &= \overline{\lim}_{k \rightarrow \infty} |f(x_k)|^{1/k} \geq \overline{\lim}_{k \rightarrow \infty} |a_1 g(x_k)|^{1/k} = \\
&= \lim_{k \rightarrow \infty} |a_1|^{1/k} \overline{\lim}_{k \rightarrow \infty} |g(x_k)|^{1/k} = \overline{\lim}_{k \rightarrow \infty} |g(x_k)|^{1/k} \\
&\Rightarrow \beta_m = \overline{\lim}_{k \rightarrow \infty} |g(x_k)| \quad \blacksquare
\end{aligned}$$

CAPITULO III

MINIMIZACION DE UNA FUNCION EN UN INTERVALO

Introducción:

Sean $a, b, t \in \mathbb{R}^n$, $t \neq 0$ dados, $I = [a,b]$ ó $I = \{ z / z = a + \lambda t, \lambda \geq 0 \}$

Analizaremos, en este capítulo, algoritmos para resolver el programa:

$$P : \min \{ f(x) / x \in I \}$$

donde $f : I \longrightarrow \mathbb{R}$ es una función continua.

Si $I = [a,b]$, y sea $g : [0,1] \longrightarrow \mathbb{R}$ definida por $g(\lambda) = f(\lambda a + (1-\lambda)b)$.

Entonces P es equivalente a:

$$\tilde{P} : \min \{ g(\lambda) / \lambda \in [0,1] \}$$

Por esta razón, sin pérdida de generalidad, podemos asumir que $I \subset \mathbb{R}$.

Es muy importante, disponer de algoritmos eficientes para resolver el programa P , porque muchos algoritmos de programación no lineal incluyen la solución de tales programas unidimensionales, como subpaso.

Analizaremos brevemente dos algoritmos para resolver P :

El método de la sección dorada.

El método de la interpolación cuadrática.

EL METODO DE LA SECCION DORADA

Definición: (Funciones Unimodales)

Una función $f : I = [a,b] \longrightarrow \mathbb{R}$ es unimodal en I , si es continua y tiene un único mínimo local. ■

Nota : Como $I = [a,b]$ es compacto, este único mínimo local es al mismo tiempo la única solución óptima de P .

Ejemplos :

1. Toda función $g : I = [a,b] \longrightarrow \mathbb{R}$ continua y estrictamente convexa es unimodal.

Veamos:

Supongamos que existen x_1 y x_2 en I ($x_1 \neq x_2$) tales que sean mínimos locales, es decir $g(x_1) = g(x_2) \leq g(x)$

Como g es estrictamente convexa en I , entonces:

$$g\left(\frac{1}{2}(x_1 + x_2)\right) < \frac{1}{2}g(x_1) + \frac{1}{2}g(x_2) = g(x_1)$$

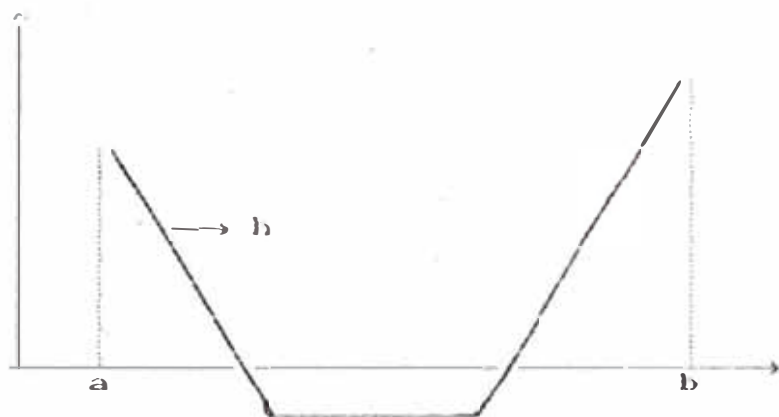
Debido a que $g(x_1) \leq g(x)$, $\forall x \in I$,

tomemos $x = \frac{1}{2}(x_1 + x_2)$

$$\rightarrow g(x_1) \leq g\left(\frac{1}{2}(x_1 + x_2)\right) < g(x_1) \quad (\text{contradicción})$$

Luego g posee un único mínimo local $\Rightarrow g$ es unimodal ■

2. Una función $h : I \rightarrow \mathbb{R}$ continua y convexa no es necesariamente unimodal:



h es continua
y convexa
pero no es
UNIMODAL

3. Existen funciones unimodales que no son convexas.



g es unimodal
pero no es
convexa

Es fácil ver que si $f : I \longrightarrow \mathbb{R}$ es unimodal en I , entonces f es unimodal en todo subintervalo cerrado $J \subset I$.

Sea $f : [a,b] \longrightarrow \mathbb{R}$ unimodal, $a < x_1 < x_2 < b$. Entonces el valor de $f(x_1)$, $f(x_2)$ permite establecer en cual de los dos subintervalos $[a,x_2]$, $[x_1,b]$ se encuentran x^* de P .

Lema 1:

- a) Si $f(x_1) < f(x_2)$, entonces $x^* \in [a,x_2]$
- b) Si $f(x_1) > f(x_2)$, entonces $x^* \in [x_1,b]$
- c) Si $f(x_1) = f(x_2)$, entonces $x^* \in [x_1,x_2]$.

Demostración :

a) Sea $f(x_1) < f(x_2)$ y supongamos que $x^* \notin [a,x_2]$.

Como f es continua, entonces toma su mínimo en $[a,x_2]$ en un punto $\tilde{x} \in [a,x_2]$.

Por consiguiente, f tiene en I al menos dos mínimos locales \tilde{x} y x^* , esto es una contradicción debido a que f es unimodal.

En consecuencia, $x^* \in [a,x_2]$.

b) La demostración es análoga a la de a).

c) Sea $f(x_1) = f(x_2)$ y supongamos que $x^* \notin [x_1,x_2]$.

Como f es unimodal en $J = [x_1,x_2]$, toma su mínimo en J en un punto $\tilde{x} \in [x_1,x_2]$.

Por consiguiente f tiene en I al menos dos mínimos locales x^* y \tilde{x} , (contradicción)

■

En el caso donde $f : I \longrightarrow \mathbb{R}$ es unimodal, se deduce del LEMA 1, el siguiente ESQUEMA DE UN ALGORITMO PARA RESOLVER P :

Si $I_0 = [a, b]$ y x^* es la solución óptima de P , se construye una sucesión de intervalos $I_k = [a_k, b_k]$, tal que se cumple:

a) $x^* \in I_k, \forall k \in \mathbb{Z}_+$

b) $I_{k+1} \subset I_k, \forall k \in \mathbb{Z}_+$

c) $\ell(I_{k+1}) = b_{k+1} - a_{k+1} < \ell(I_k) = b_k - a_k, \forall k \in \mathbb{Z}_+$

Si $I_k = [a_k, b_k]$ es dado y $x^* \in I_k$, entonces elegimos x_k^1 ,

x_k^2 tales que $a_k < x_k^1 < x_k^2 < b_k$.

Si $f(x_k^1) < f(x_k^2)$, entonces $x^* \in [a_k, x_k^2]$ y elegimos $I_{k+1} = [a_k, x_k^2]$

Si $f(x_k^1) \geq f(x_k^2)$, entonces $x^* \in [x_k^1, b_k]$ y elegimos $I_{k+1} = [x_k^1, b_k]$

La sucesión $\{I_k\}$, que se obtiene de este modo, tiene evidentemente las propiedades a), b) y c).

Si los puntos x_k^1, x_k^2 se eligen adecuadamente, se logra que para una constante $0 < \gamma < 1$, se cumplen:

$$\ell(I_{k+1}) = \gamma \ell(I_k), \forall k \in \mathbb{Z}_+ \dots\dots\dots 1)$$

En este caso las sucesiones $\{a_k\}, \{b_k\}$ convergen linealmente a la única solución óptima x^* de P , con tasa γ .

Consideramos la ecuación :

$$z^2 + z - 1 = 0 \dots\dots\dots 2)$$

que tiene raíces :

$$z_1 = -\frac{1}{2} + \frac{1}{2}\sqrt{5} \quad \text{y} \quad z_2 = -\frac{1}{2} - \frac{1}{2}\sqrt{5}$$

elegimos $\gamma = z_1 \approx 0.62$ y

$$x_k^1 = b_k - \gamma(b_k - a_k) \dots\dots\dots 3)$$

$$x_k^2 = a_k + \gamma(b_k - a_k) \dots\dots\dots 4)$$

con lo cual se cumple :

$$a_k < x_k^1 < x_k^2 < b_k .$$

Si $f(x_k^1) < f(x_k^2)$, entonces :

$$I_{k+1} = [a_k, x_k^2] \text{ y } \ell(I_{k+1}) = x_k^2 - a_k = \gamma(b_k - a_k) = \gamma \ell(I_k).$$

Si $f(x_k^1) \geq f(x_k^2)$, entonces :

$$I_{k+1} = [x_k^1, b_k] \text{ y } \ell(I_{k+1}) = b_k - x_k^1 = \gamma(b_k - a_k) = \gamma \ell(I_k).$$

Por consiguiente se cumple : $\ell(I_{k+1}) = \gamma \ell(I_k)$, $\forall k \in \mathbb{Z}_+$.

El método que acabamos de describir se llama, el ALGORITMO D—
LA SECCION DORADA .

Este algoritmo tiene la siguiente ventaja adicional :

En cada iteración es necesario evaluar f en un solo punto.

Esto es así, por que se cumple :

$$x_{k+1}^1 = x_k^2 \text{ ó } x_{k+1}^2 = x_k^1, \forall k \quad 5)$$

Demostracion :

Si $I_{k+1} = [a_k, x_k^2]$, entonces se cumple por 1),2),3) y 4):

$$\begin{aligned} x_{k+1}^2 &= a_k + \gamma(x_k^2 - a_k) = a_k + \gamma^2(b_k - a_k) = \\ &= a_k + (1 - \gamma)(b_k - a_k) = b_k - \gamma(b_k - a_k) = x_k^1 \\ &\Rightarrow x_{k+1}^2 = x_k^1, \forall k \end{aligned}$$

Si $I_{k+1} = [x_k^1, b_k]$, entonces se deduce de 1),2),3) y 4) :

$$\begin{aligned} x_{k+1}^1 &= b_k - \gamma(b_k - x_k^1) = b_k - \gamma^2(b_k - a_k) = \\ &= b_k - (1 - \gamma)(b_k - a_k) = a_k + \gamma(b_k - a_k) = x_k^2 \\ &\Rightarrow x_{k+1}^1 = x_k^2, \forall k \quad \blacksquare \end{aligned}$$

METODO DE INTERPOLACION CUADRATICA

El método a menudo más útil en la minimización de una función en un intervalo, es el de la INTERPOLACION CUADRATICA, a través de tres puntos dados. Este tiene la ventaja de no requerir información de ninguna derivada.

La base para la interpolación cuadrática es aproximar f , la función cuyo mínimo es buscado, por una función cuadrática Φ , dada por: $\Phi(x) = a + bx + cx^2$.

Supongamos que evaluamos f en tres puntos x_1, x_2, x_3 tales que: $x_1 < x_2 < x_3$.

Haciendo: $\Phi(x_i) = f(x_i)$, $i = 1, 2, 3$; podemos calcular a , b , c .

El mínimo de Φ (si existe) puede hallarse analíticamente resolviendo $\Phi'(x) = 0$; y como una primera aproximación del mínimo de f , obtenemos el punto \hat{x} , dado por: $\hat{x} = -\frac{b}{2c}$, suponiendo que $c > 0$. Si $c < 0$, entonces Φ es una función cuadrática cóncava, que toma su máximo \hat{x} . En este caso, el punto \hat{x} es inútil como aproximación a un mínimo local de f .

Una condición que asegura que $c > 0$ es:

$$f(x_1) > f(x_2) \quad \dots \dots \dots (1)$$

$$f(x_2) < f(x_3)$$

(1) nos indica también que f tiene un mínimo local en el intervalo $\langle x_1, x_3 \rangle$

Para simplificar la descripción del algoritmo de la iteración cuadrática, asumimos que $f(\hat{x}) \neq f(x_2)$ y $\hat{x} \neq x_2$. (El caso donde $f(\hat{x}) = f(x_2)$ será tratado más adelante).

Distinguimos dos casos:

1) $\hat{x} \in \langle x_1, x_2 \rangle$

Si $f(\hat{x}) < f(x_2)$, entonces la terna $(x'_1, x'_2, x'_3) = (x_1, \hat{x}, x_2)$ satisface (1). En consecuencia minimizamos el polinomio cuadrático de interpolación por los puntos $f(x'_i)$, $i = 1, 2, 3$.

Si $f(\hat{x}) > f(x_2)$, entonces $(x'_1, x'_2, x'_3) = (\hat{x}, x_2, x_3)$ tiene

la propiedad (1) y repetimos el paso inicial para (x'_1, x'_2, x'_3) .

2) $\hat{x} \in (x_2, x_3)$

Si $f(\hat{x}) < f(x_2)$, usaremos en el próximo paso la terna

$(x'_1, x'_2, x'_3) = (x_2, \hat{x}, x_3)$. De otro modo, usaremos la terna

$(x'_1, x'_2, x'_3) = (x_1, x_2, \hat{x})$.

Mencionamos el siguiente resultado :

Si f es de clase C^2 y \hat{x} es un mínimo local (estricto) de f con $f''(\hat{x}) > 0$ y si (x_1^k, x_2^k, x_3^k) son los puntos de iteración, entonces las sucesiones $\{x_1^k\}$, $\{x_2^k\}$, $\{x_3^k\}$ convergen a \hat{x} , con el orden de convergencia $p = 1.3$, si la terna inicial (x_1^0, x_2^0, x_3^0) se encuentra en una vecindad suficientemente pequeña \mathcal{U} de \hat{x} .

A continuación, analizamos la convergencia global del algoritmo de interpolación cuadrática.

LA CONVERGENCIA GLOBAL DEL METODO DE INTERPOLACION CUADRATICA

Cabe destacar que los algoritmos descritos en este párrafo, convergen sólo si se elige x_0 en una vecindad suficientemente pequeña \mathcal{U} de una solución óptima (o punto estacionario) x^* . Decimos, por esta razón, que sólo convergen localmente

Si x_0 no es elegido adecuadamente, entonces generalmente no es posible predecir el comportamiento de $\{x_k\}$, y es posible que la sucesión $\{x_n\}$ diverja o "se escape", porque ningún punto de iteración cae en una vecindad suficientemente pequeña \mathcal{U} de x^* .

Para lograr la convergencia global, será necesario combinar ingeniosamente los diferentes métodos y/o modificarlos ligeramente.

El punto crucial para lograr la convergencia global es el TEOREMA DE CONVERGENCIA GLOBAL ; y para este teorema es decisivo disponer de una función de descenso.

En consecuencia, se deberá combinar y/o modificar los algoritmos, de modo que se obtenga un algoritmo de descenso, para una función de descenso adecuada.

Vamos a analizar la convergencia global del algoritmo de interpolación cuadrática. Asumimos que f es unimodal y de clase C^2 , y que no existen puntos x con $f'(x) = f''(x) = 0$.

Se inicia el algoritmo buscando tres puntos x_1, x_2, x_3 tales que: $x_1 < x_2 < x_3$, y $f(x_1) > f(x_2)$, $f(x_2) < f(x_3)$.

En este caso, la función cuadrática de interpolación es convexa y toma su mínimo en $[x_1, x_3]$.

Modificamos el algoritmo de tal modo que en cada iteración k se trabaje con tres puntos: x_1^k, x_2^k, x_3^k tales que:

$$x_1^k < x_2^k < x_3^k, f(x_2^k) \leq f(x_1^k), f(x_2^k) \leq f(x_3^k) \quad \dots (1)$$

Supongamos que (x_1^k, x_2^k, x_3^k) tenga la propiedad (1).

Determinamos el mínimo x^* de la función cuadrática de interpolación $q_k(x)$.

Distinguimos varios casos:

a) $x_1^k < x^* < x_2^k$

Si $f(x^*) < f(x_2^k)$, entonces podemos elegir:

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) = (x_1^k, x^*, x_2^k)$$

Si $f(x^*) \geq f(x_2^k)$, entonces: $x^* < x_2^k < x_3^k, f(x_2^k) \leq f(x^*)$.

$f(x_2^k) \leq f(x_3^k)$. Por consiguiente, podemos elegir:

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) = (x^*, x_2^k, x_3^k)$$

b) $x_2^k < x^* < x_3^k$

Si $f(x^*) < f(x_2^k)$, podemos elegir:

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) = (x_2^k, x^*, x_3^k),$$

puesto que: $f(x^*) < f(x_2^k) \leq f(x_3^k)$.

Si $f(x^*) \geq f(x_2^k)$, podemos elegir:

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) = (x_1^k, x_2^k, x^*),$$

puesto que: $f(x_1^k) \geq f(x_2^k)$, $f(x_2^k) \leq f(x^*)$.

Para analizar la convergencia global del algoritmo, definimos la aplicación: $A : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, por

$A(x_1, x_2, x_3) = (x_1', x_2', x_3')$, donde (x_1', x_2', x_3') es la terna que se obtiene por el algoritmo partiendo de (x_1, x_2, x_3) .

Sin embargo, falta definir la aplicación A también para el caso donde al menos dos de los tres puntos x_1, x_2, x_3 , coinciden; puesto que en el análisis precedente omitimos el caso $x^* = x_2^k$.

Analicemos este caso:

Si $f'(x^*) > 0$, elegimos como nueva terna:

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) = (x_1^k, x_2^k, x_2^k).$$

En el próximo paso, determinamos el polinomio:

$q_{k+1}(x) = ax^2 + bx + c$, tal que:

$$q_{k+1}(x_1^k) = f(x_1^k), \quad q_{k+1}(x_2^k) = f(x_2^k), \quad q_{k+1}'(x_2^k) = f'(x_2^k)$$

Si \tilde{x} es el mínimo de q_{k+1} ,

$$f(\tilde{x}) = q_{k+1}(\tilde{x}) < f(x_2^k) = q_{k+1}(x_2^k),$$

se analizará, en el próximo paso la terna: $(x_1^k, \tilde{x}, x_2^k)$.

Si $f'(x^*) < 0$, elegimos:

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) = (x_2^k, x_2^k, x_3^k) \text{ y } q_{k+1}(x) \text{ tal que:}$$

$$q_{k+1}(x_2^k) = f(x_2^k), \quad q_{k+1}'(x_2^k) = f'(x_2^k), \quad q_{k+1}(x_3^k) = f(x_3^k)$$

Con esto queda establecido cómo se define $A(x_1, x_2, x_2)$ y $A(x_1, x_1, x_2)$ con $x_1 \neq x_2$.

Si $x^* = x_2^k$, $f'(x^*) = 0$ y $f''(x^*) \geq 0$, entonces:

$$(x_1^{k+1}, x_2^{k+1}, x_3^{k+1}) = (x_2^k, x_2^k, x_2^k).$$

(Como f es unimodal, $f'(x^*) = 0$, y, $f''(x^*) < 0$ no se presenta).

De este modo, la aplicación A está bien definida. A es también continua, puesto que la interpolación depende de manera continua de los datos.

Si x^{**} es el mínimo de f , elegimos como conjunto de soluciones $X = \{(x^{**}, x^{**}, x^{**})\}$.

Finalmente, definimos $z(x) = f(x_1) + f(x_2) + f(x_3)$; z es una función de descenso para A : cuando se aplica A , entonces uno de los valores $f(x_1)$, $f(x_2)$, ó, $f(x_3)$ es reemplazado por $f(x^*)$, donde x^* es el mínimo del polinomio cuadrático.

Si $f(x_i)$ es sustituido, entonces $f(x_i) > f(x^*)$, $i = 1, 2, 3$. Esto es consecuencia de la unimodalidad de f y de la definición de A .

Se cumple evidentemente:

$$A(x^{**}, x^{**}, x^{**}) = (x^{**}, x^{**}, x^{**}) \text{ y}$$

$$z(A(x^{**}, x^{**}, x^{**})) = z(x^{**}, x^{**}, x^{**})$$

Como todas las ternas se encuentran en el intervalo inicial, se cumplen todas las hipótesis del Teorema de la Convergencia Global.

consecuencia, el algoritmo modificado converge globalmente al mínimo x^{**}

Cabe resaltar, que el algoritmo modificado tiene el mismo orden de convergencia que el algoritmo original, porque en una vecindad suficientemente pequeña de x^{**} , los dos algoritmos coinciden.

CAPITULO IV

LOS ALGORITMOS DE DESCENSO MAS RAPIDO

Sea $\|\cdot\|$ una norma cualquiera en \mathbb{R}^n y $\beta > 0$ una constante dada. Consideremos los programas:

$$J(x) : \min \{ \nabla f(x)^T t \mid \|t\| \leq 1 \}$$

$$J(x,t) : \min \{ f(x+\lambda t) \mid \lambda \geq 0, \lambda \|t\| \leq \beta \}$$

Sea $R(x)$ el conjunto de soluciones óptimas de $J(x)$ y $Q(x) = \{x\} \times R(x)$. Sea, finalmente, $S(x,t)$ el conjunto de soluciones óptimas de $J(x,t)$.

$$Q : \mathbb{R}^n \longrightarrow \mathbb{R}^n \times \mathbb{R}^n \quad \text{y} \quad S : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

son correspondencias.

El algoritmo definido por la correspondencia

$$H = S \circ Q : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

se llama un ALGORITMO DE DESCENSO MAS RAPIDO.

Si $X^* = \{x \in \mathbb{R}^n \mid \nabla f(x) = 0\}$, entonces f es una función de descenso para H y X^* .

Si $y \in H(x)$, entonces, por la definición de $J(x,t)$ se cumple evidentemente : $f(y) \leq f(x)$.

Si $x \in \mathbb{R}^n \setminus X^*$, entonces $\inf J(x) < 0$, y para todo $t \in R(x)$ se cumple que : $t \neq 0, \nabla f(x)^T t < 0$.

De esto se deduce que : $f(y) < f(x), \forall y \in H(x)$.

Usando el teorema de la convergencia global, vamos a demostrar el siguiente :

TEOREMA : Si $\{x_k\}$ es una sucesión generada por el algoritmo H , entonces todo punto de acumulación de $\{x_k\}$ se encuentra en X^* .

Demostración :

De acuerdo al teorema de la convergencia global, basta probar que H es cerrada en $\mathbb{R}^n \setminus X^*$.

a) Q es cerrada en \mathbb{R}^n :

Sea $\lim_{k \rightarrow \infty} x_k = x$, $t_k \in R(x_k)$, $\forall k$, $\lim_{k \rightarrow \infty} t_k = \bar{t}$.

Si $\|t\| \leq 1$, entonces:

$$\nabla f(x_k)^T t_k \leq \nabla f(x_k)^T t, \quad \|t_k\| \leq 1, \quad \forall k$$

Como $f \in C^1(\mathbb{R}^n)$, obtenemos de este paso, para $k \rightarrow \infty$:

$$\nabla f(x)^T \bar{t} \leq \nabla f(x)^T t, \quad \|t\| \leq 1.$$

Por consiguiente $(x, \bar{t}) \in Q(x)$, y Q es cerrada en x.

b) S es cerrada en todo punto (x, t) con $t \neq 0$:

Sea $\lim_{k \rightarrow \infty} (x_k, t_k) = (x, t)$, $z_k \in S(x_k, t_k)$, $\forall k$, y

$$\lim_{k \rightarrow \infty} z_k = \bar{z}$$

Como $t \neq 0$, podemos asumir que $t_k \neq 0$, $\forall k$.

Se cumple entonces:

$$z_k = x_k + \lambda_k t_k, \quad \lambda_k = \frac{\|z_k - x_k\|}{\|t_k\|} \geq 0, \quad \lambda_k \|t_k\| \leq \beta$$

De esto se obtiene para $k \rightarrow \infty$:

$$\lim_{k \rightarrow \infty} \lambda_k = \frac{\|z - x\|}{\|t\|} = \bar{\lambda} \geq 0, \quad \bar{z} = x + \bar{\lambda} t,$$

$$\bar{\lambda} \|t\| \leq \beta$$

Por consiguiente, \bar{z} es admisible para $J(x, t)$.

Sea $y = x + \mu t$ un punto admisible cualquiera de $J(x, t)$.

Entonces existe una sucesión $y_k = x_k + \mu_k t_k$ tal que y_k

es admisible para $J(x_k, t_k)$, $\forall k$, $\lim_{k \rightarrow \infty} y_k = y$

[si $\mu \|t\| = \beta$, elegimos $\mu_k = \frac{\beta}{\|t_k\|}$; si $\mu \|t\| < \beta$

entonces existe un k_0 tal que $\mu \|t_k\| \leq \beta$, $\forall k \geq k_0$, y

podemos elegir $\mu_k = \mu$, $\forall k \geq k_0$].

Se cumple entonces:

$$f(z_k) \leq f(y_k), \quad \forall k$$

Como $f \in C^1(\mathbb{R}^n)$, obtenemos de esto:

$$f(\bar{z}) \leq f(y)$$

Por consiguiente, $\bar{z} \in S(x,t)$ y S es cerrada en (x,t) .

c) H es cerrada en $\mathbb{R}^n \setminus X^*$:

Sea $x \in \mathbb{R}^n \setminus X^*$, $\lim_{k \rightarrow \infty} x_k = x$, $(x_k, t_k) \in Q(x_k)$, $\forall k$.

Entonces $t \neq 0$, $\forall (x,t) \in Q(x)$, y de acuerdo a b), S es cerrada en $Q(x)$.

Si $B = \{ t \mid \|t\| \leq 1 \}$, entonces $Q(x_k) \subset \{x_k\} \times B$, $\forall k$.

De esto y la compacidad de B se deduce que existe una subsucesión (\bar{x}_k, \bar{t}_k) que converge.

Por consiguiente, de acuerdo a la proposición de la composición de la correspondencia cerrada, H es cerrada en $\mathbb{R}^n \setminus X^*$.

Si $x \notin X^*$, entonces toda solución óptima t de $J(x)$ se llama una DIRECCION DE DESCENSO MAS RAPIDO para f en x .

Observación:

A pesar que los algoritmos de descenso más rápido convergen globalmente, generalmente no son muy eficientes, porque la velocidad de su convergencia es baja.

De otro lado, existen algoritmos clásicos, como el método de Newton, que convergen localmente muy rápidamente pero no tienen la propiedad de la convergencia global.

¿ Qué significa convergencia local en este caso ?

Generalmente significa lo siguiente:

Si \bar{x} es un punto estacionario, que satisface además alguna condición de regularidad [por ejemplo : $\nabla^2 f(\bar{x})$ tiene inversa], entonces la sucesión $\{x_k\}$ generada por el algoritmo, converge a \bar{x} , si

el punto inicial x_0 se encuentra en una vecindad suficientemente pequeña de \bar{x} .

Por su baja velocidad de convergencia, los métodos de descenso más rápido deberían aplicarse en los casos donde no es posible encontrar una buena aproximación x_0 a una solución óptima de :

$$P : \min \{ f(x) / x \in \mathbb{R}^n \}$$

Además, es conveniente combinar un método de descenso más rápido con otro método cuya velocidad de convergencia local es alta.

En este caso se aplica el método de descenso, hasta que se cumpla algún criterio adecuado [por ejemplo: $\|\nabla f(x_k)\| < \epsilon$] que nos permite cambiar al método más veloz que no converge globalmente, porque, con alta probabilidad, nos encontramos en una región donde este método converge a una solución óptima de P ó al menos a un mínimo local de f .

EL CASO CUADRÁTICO

Sea $C = C^T \in M(n,n)$ una matriz definida positiva,

$$f(x) = \frac{1}{2} x^T C x + p^T x .$$

Como C es definida positiva, se cumple, para sus valores propios λ_i , $1 \leq i \leq n$:

$$0 < a = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = A$$

Como f es estrictamente convexa, se cumple:

x^* es solución óptima de:

$$P : \min \{ f(x) / x \in \mathbb{R}^n \} \iff 0 = \nabla f(x^*) = Cx^* + p$$

$$\iff Cx^* = -p \quad \text{ó} \quad x^* = C^{-1}p$$

Consideramos la función :

$$E(x) = \frac{1}{2} (x - x^*)^T C (x - x^*)$$

se cumple entonces:

$$E(x) = \frac{1}{2} x^T C x - x^T C x^* + \frac{1}{2} (x^*)^T C x^*$$

$$= \frac{1}{2} x^T C x + p^T x + \frac{1}{2} (x^*)^T C x^*$$

$$E(x) = f(x) + \frac{1}{2} (x^*)^T C x^*$$

En consecuencia, la función E se distingue de la función sólo por una constante aditiva.

Para muchos propósitos será más conveniente considerar la minimización de E en lugar de la de f .

Si $g_k = \nabla f(x_k) = Cx_k + p$, entonces la iteración del método del gradiente tiene la forma:

$$x_{k+1} = x_k - \alpha_k g_k$$

donde α_k es la solución óptima de :

$$I_k : \min \{ h(\alpha) = f(x_k - \alpha g_k) / \alpha \geq 0 \}$$

En este caso podemos calcular α_k de manera explícita.

Se cumple:

$$\begin{aligned} h(\alpha) &= \frac{1}{2} (x_k - \alpha g_k)^T C (x_k - \alpha g_k) + p^T (x_k - \alpha g_k) \\ &= f(x_k) + \frac{1}{2} (g_k^T C g_k) \alpha^2 - g_k^T (C x_k + p) \alpha \end{aligned}$$

Como $h(\alpha)$ es estrictamente convexa se obtiene α_k de:

$$\begin{aligned} 0 = h'(\alpha_k) &= (g_k^T C g_k) \alpha_k - g_k^T (C x_k + p) \\ &= (g_k^T C g_k) \alpha_k - g_k^T g_k \end{aligned}$$

Por consiguiente se cumple:

$$\begin{aligned} \alpha_k &= \frac{g_k^T g_k}{g_k^T C g_k} \quad \dots\dots(1a) \\ x_{k+1} &= x_k - \frac{g_k^T g_k}{g_k^T C g_k} g_k \quad \dots\dots(1) \end{aligned}$$

Vamos a analizar la velocidad de la convergencia de la iteración (1).

Lema 1: Para la iteración (1) se cumple:

$$E(x_{k+1}) = \left[1 - \frac{(g_k^T g_k)^2}{(g_k^T C g_k)(g_k^T C^{-1} g_k)} \right] E(x_k) \quad \dots(\star)$$

Demostracion :

Como $x^* = C^{-1} p$, se cumple :

$$\begin{aligned} g_k^T C^{-1} g_k &= (C x_k + p)^T C^{-1} (C x_k + p) \\ &= x_k^T C x_k + 2p^T x_k + p^T C^{-1} p \quad \dots(2) \\ &= 2f(x_k) + p^T C^{-1} C C^{-1} p \\ &= 2f(x_k) + (x^*)^T C x^* = 2 E(x_k) \end{aligned}$$

Poniendo $y_k = x_k - x^*$, obtenemos, usando (2):

$$\frac{E(x_k) - E(x_{k+1})}{E(x_k)} =$$

$$\begin{aligned}
&= \frac{\frac{1}{2} (y_k^T C y_k) - \frac{1}{2} (y_k - \alpha_k g_k)^T C (y_k - \alpha_k g_k)}{\frac{1}{2} g_k^T C^{-1} g_k} \\
&= \frac{2\alpha_k g_k^T C y_k - \alpha_k^2 g_k^T C g_k}{g_k^T C^{-1} g_k} \quad (3)
\end{aligned}$$

De (3), (1a) y de $g_k = Cx_k + p = C(x_k - x^*) = Cy_k$, obtenemos:

$$\frac{E(x_k) - E(x_{k+1})}{E(x_k)} = \frac{(g_k^T g_k)^2}{(g_k^T C g_k)(g_k^T C^{-1} g_k)}$$

y de esto obtenemos (*).

■

Vamos a establecer una cota superior para el factor de la derecha en (*).

TEOREMA : (DESIGUALDAD DE KANTOROVICH)

Si $C = C^T \in M(n,n)$ es definida positiva, y a (respectivamente, A) es el menor (respectivamente, el mayor) valor propio de C , entonces se cumple:

$$\frac{(x^T x)^2}{(x^T C x)(x^T C^{-1} x)} \geq \frac{4aA}{(a + A)^2}$$

Demostración :

Sean $\lambda_i, 1 \leq i \leq n$ los valores propios, tales que

$$a = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = A$$

Existe entonces una base ortonormal $B = \{d_1, d_2, \dots, d_n\}$ de \mathbb{R}^n , tal que d_i es un vector propio de C asociado a $\lambda_i, 1 \leq i \leq n$.

Si $x \in \mathbb{R}^n$ es cualquiera, podemos escribir :

$$\begin{aligned}
x &= \sum_{i=1}^n x_i d_i, & x^T x &= \left(\sum_{i=1}^n x_i d_i \right)^T \left(\sum_{j=1}^n x_j d_j \right) \\
& & &= \sum_{i,j=1}^n x_i x_j d_i^T d_j = \sum_{i=1}^n x_i^2,
\end{aligned}$$

puesto que $d_i^T d_j = \delta_{ij}$.

Similarmente se obtiene de $Cd_i = \lambda_i d_i$ ó $C^{-1}d_i = \frac{1}{\lambda_i} d_i$:

$$\begin{aligned}
x^T C x &= \left(\sum_{j=1}^n x_j d_j \right)^T C \left(\sum_{i=1}^n x_i d_i \right) \\
&= \left(\sum_{j=1}^n x_j d_j \right)^T \left(\sum_{i=1}^n \lambda_i x_i d_i \right) = \\
&= \sum_{i,j=1}^n \lambda_i x_i x_j d_j^T d_i = \sum_{i,j=1}^n \lambda_i x_i x_j \delta_{ij} = \\
x^T C x &= \sum_{i=1}^n \lambda_i x_i^2 \\
x^T C^{-1} x &= \sum_{i=1}^n \frac{x_i^2}{\lambda_i}
\end{aligned}$$

Por consiguiente se cumple :

$$\frac{(x^T x)^2}{(x^T C x)(x^T C^{-1} x)} = \frac{\left(\sum_{i=1}^n x_i^2 \right)^2}{\left(\sum_{i=1}^n \lambda_i x_i^2 \right) \left(\sum_{i=1}^n \frac{x_i^2}{\lambda_i} \right)} \quad \dots (+)$$

$$\text{Sea } \xi_i = \frac{x_i^2}{\sum_{j=1}^n x_j^2}, \quad i = 1, 2, \dots, n$$

Entonces (+) se deja escribir como:

$$\frac{(x^T x)^2}{(x^T C x)(x^T C^{-1} x)} = \frac{\left(\sum_{i=1}^n \xi_i \lambda_i \right)^{-1}}{\sum_{i=1}^n \xi_i \frac{1}{\lambda_i}} = \frac{\Phi(\xi)}{\Psi(\xi)}$$

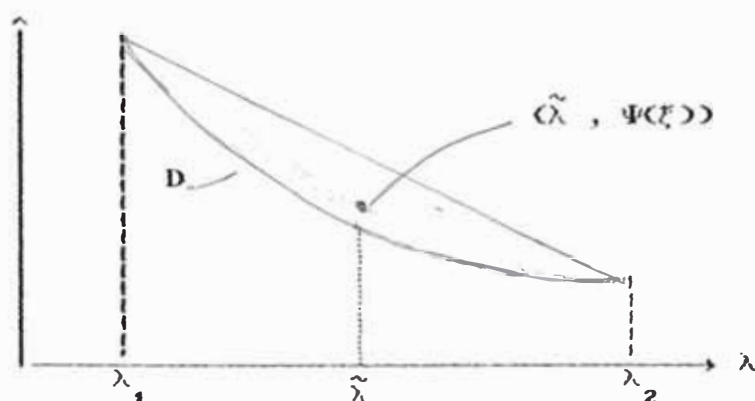
$\tilde{\lambda} = \sum_{i=1}^n \xi_i \lambda_i$ es un combinación convexa de $\lambda_1, \lambda_2, \dots, \lambda_n$.

Sea $D = \left\{ \left(\lambda, \frac{1}{\lambda} \right) / \lambda_1 \leq \lambda \leq \lambda_n \right\}$.

Entonces $\left(\tilde{\lambda}, \Psi(\xi) \right) = \sum_{i=1}^n \xi_i \left(\lambda_i, \frac{1}{\lambda_i} \right) \in \text{co}(D)$.

Como $h(\lambda) = \frac{1}{\lambda}$ es convexa en $[\lambda_1, \lambda_n]$, se cumple :

$$\Psi(\xi) = \sum_{i=1}^n \xi_i \frac{1}{\lambda_i} = \sum_{i=1}^n \xi_i h(\lambda_i) \geq h\left(\sum_{i=1}^n \xi_i \lambda_i\right) = \left(\sum_{i=1}^n \xi_i \lambda_i\right)^{-1} = \Phi(\xi).$$



Como $\text{fr}(\text{co}(D)) = D \cup [(\lambda_1, \frac{1}{\lambda_1}), (\lambda_n, \frac{1}{\lambda_n})]$, se cumple :

$$\begin{aligned} \Psi(\xi) &\leq \frac{\lambda_n - \tilde{\lambda}}{\lambda_n - \lambda_1} \cdot \frac{1}{\lambda_1} + \frac{\tilde{\lambda} - \lambda_1}{\lambda_n - \lambda_1} \cdot \frac{1}{\lambda_n} = \frac{\lambda_n^2 - \lambda_1^2 - \tilde{\lambda}(\lambda_n - \lambda_1)}{(\lambda_n - \lambda_1)\lambda_1\lambda_n} = \\ &= \frac{\lambda_n + \lambda_1 - \tilde{\lambda}}{\lambda_1\lambda_n} \end{aligned}$$

Por consiguiente se cumple :

$$\begin{aligned} \frac{\Phi(\xi)}{\Psi(\xi)} &\geq \min_{\lambda_1 \leq \tilde{\lambda} \leq \lambda_n} \frac{\lambda_1\lambda_n}{\tilde{\lambda}(\lambda_n + \lambda_1 - \tilde{\lambda})} = \frac{\lambda_1\lambda_n}{\max_{\lambda_1 \leq \tilde{\lambda} \leq \lambda_n} [-(\tilde{\lambda})^2 + \tilde{\lambda}(\lambda_n + \lambda_1)]} = \\ &= \frac{\lambda_1\lambda_n}{\max_{\lambda_1 \leq \tilde{\lambda} \leq \lambda_n} \varphi(\tilde{\lambda})} = \alpha. \end{aligned}$$

Como $\varphi'(\tilde{\lambda}) = -2\tilde{\lambda} + (\lambda_n + \lambda_1) = 0 \Leftrightarrow \tilde{\lambda} = \frac{\lambda_1 + \lambda_n}{2}$, se

cumple:

$$\frac{\Phi(\xi)}{\Psi(\xi)} \geq \alpha = \frac{\lambda_1\lambda_n}{(\lambda_1 + \lambda_n)^2} = \frac{4\lambda_1\lambda_n}{(\lambda_1 + \lambda_n)^2} = \frac{4aA}{(a+A)^2}, \quad \forall \xi \in \mathbb{R}^n$$

Teorema :

Para todo $x_0 \in \mathbb{R}^n$, la sucesión $\{x_k\}$ obtenida por (1) converge al unico mínimo x^* de f .

Para $E(x) = \frac{1}{2} (x - x^*)^T C (x - x^*)$ se cumple además :

$$E(x_{k+1}) \leq \left(\frac{\Lambda - a}{\Lambda + a}\right)^2 E(x_k)$$

Demostracion :

Por el Lema 1 y la desigualdad de Kantorovich se cumple :

$$E(x_{k+1}) = \left[1 - \frac{(g_k^T g_k)^2}{(g_k^T C g_k)(g_k^T C^{-1} g_k)} \right] E(x_k)$$

$$= \left(1 - \frac{\Phi(\xi)}{\Psi(\xi)}\right) E(x_k) \leq \left[1 - \frac{4a\Lambda}{(a + \Lambda)^2}\right] E(x_k) = \left(\frac{\Lambda - a}{\Lambda + a}\right)^2 E(x_k).$$

Como $0 \leq \left(\frac{\Lambda - a}{\Lambda + a}\right)^2 < 1$, se obtiene de esto que $\lim_{k \rightarrow \infty} x_k = x^*$ ■

Observaciones :

Si usamos $E(x)$ como función de error, entonces por el teorema precedente, la sucesión $\{x_k\}$ generada por el método del gradiente converge de manera lineal a la solución óptima de x^* con tasa de convergencia $\beta \leq \left(\frac{\Lambda - a}{\Lambda + a}\right)^2$.

Se ha demostrado que la cota $\left(\frac{\Lambda - a}{\Lambda + a}\right)^2$ para β es exacta, esto significa que para ciertos ejemplos se cumple :

$$\beta = \left(\frac{\Lambda - a}{\Lambda + a}\right)^2.$$

Por esta razón decimos que la tasa de convergencia del método del gradiente es $\left(\frac{\Lambda - a}{\Lambda + a}\right)^2$.

Observamos finalmente que esta tasa depende solamente del cociente $r = \frac{\Lambda}{a}$, puesto : $\left(\frac{\Lambda - a}{\Lambda + a}\right)^2 = \left(\frac{r - 1}{r + 1}\right)^2$.

En consecuencia la convergencia se vuelve más lenta, cuando r crece.

El cociente $r = \frac{A}{a}$ que es el único mínimo que caracteriza la convergencia, se llama con frecuencia la medición de C .

CAPITULO V

METODO DE LAS DIRECCIONES CONJUGADAS

Los métodos de las direcciones conjugadas se consideran un método que se encuentra entre el método de Descenso más rápido y el método de Newton. El cual fue motivado con el deseo de acelerar la convergencia típicamente baja asociada al Descenso más rápido y a la vez evita los requerimientos asociados a la evaluación, almacenamiento e inversión del Hessiano del método de Newton.

Sea $C \in M(n,n)$ una matriz simétrica definida positiva.

Sea $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$.

Definamos : $Q(x) = \frac{1}{2} x^T C x + p^T x$.

Sean d_1, d_2, \dots, d_k ; $k \leq n$ vectores linealmente independientes en \mathbb{R}^n , sea $x_0 \in \mathbb{R}^n$.

Consideramos el subespacio afín :

$$L = x_0 + \text{Lin} \{ d_1, d_2, \dots, d_k \}$$

Si $D = (d_1, d_2, \dots, d_k) \in M(n,k)$ entonces :

$$L = \{ x = x_0 + D\lambda \mid \lambda \in \mathbb{R}^k \}$$

Consideremos el programa :

$$P : \min \{ Q(x) \mid x \in L \}$$

LEMA 1:

\hat{x} es una solución óptima de P si y solamente si $\hat{x} \in L$, $\nabla Q(\hat{x})^T d_i = 0$, $i = 1, 2, \dots, k$.

DEMOSTRACION :

$$\forall x \in L \text{ se cumple : } x = x_0 + D\lambda$$

$$\rightarrow Q(x) = Q(x_0 + D\lambda) = \frac{1}{2}(x_0 + D\lambda)^T C(x_0 + D\lambda) + p^T(x_0 + D\lambda)$$

$$\Rightarrow Q(x) = \frac{1}{2}x_0^T C x_0 + p^T x_0 + \frac{1}{2}\lambda^T D^T C x_0 + \frac{1}{2}x_0^T C D \lambda + \frac{1}{2}\lambda^T D^T C D \lambda + p^T D \lambda$$

$$\Rightarrow Q(x) = Q(x_0) + x_0^T C D \lambda + \frac{1}{2} \lambda^T D^T C D \lambda + p^T D \lambda = F(\lambda)$$

Veamos que $D^T C D$ sea definida positiva :

Sea $y \in \mathbb{R}^k$, $y \neq 0$, entonces $y^T D^T C D y = (Dy)^T C (Dy)$,

haciendo $z = Dy \in \mathbb{R}^n$, necesariamente $z \neq 0$.

Supongamos que $z = 0$, debido a que D está formado por los vectores d_i linealmente independiente, se obtiene que la única solución es $y = 0$ (contradicción).

Y como C es definida positiva concluye que $z^T C z > 0$,

$\rightarrow D^T C D$ es definida positiva.

Luego F es estrictamente convexa en \mathbb{R}^k y $\hat{\lambda}$ minimiza F en \mathbb{R}^k si y solo si $\nabla F(\hat{\lambda}) = 0$

$$\nabla F(\hat{\lambda}) = x_0^T C D + p^T D + (\hat{\lambda})^T D^T C D = 0$$

$$\iff [(x_0^T C + p^T) + (\hat{\lambda})^T D^T C] D = 0$$

$$\iff [C(x_0 + D\hat{\lambda}) + p^T]^T D = 0$$

$$\Rightarrow \hat{x} = x_0 + D\hat{\lambda} \in L$$

$\Rightarrow x$ es solución óptima de P .

$$\iff \nabla Q(\hat{x})^T D = 0$$

$$\iff \nabla Q(\hat{x})^T d_i = 0 \quad , \quad i = 1, 2, \dots, k$$

Definición :

Los vectores $d_1, d_2 \in \mathbb{R}^n$ son conjugados (C-conjugados) si $d_1^T C d_2 = 0$ □

LEMA 2:

Sean $d_i \neq 0$, $i = 1, 2, \dots, k$; $k \leq n$, $d_i^T C d_j = 0$, $\forall i \neq j$.

Entonces estos vectores son linealmente independientes.

DEMOSTRACION :

Sean $\alpha_1, \alpha_2, \dots, \alpha_k$ tales que :

$$\sum_{j=1}^k \alpha_j d_j = 0 \quad \dots (*)$$

Veamos que los $\alpha_i = 0 \quad \forall i = 1, 2, \dots, k$:

en (*) :

$$d_i^T C \sum_{j=1}^k \alpha_j d_j = 0$$

$$\Rightarrow \sum_{j=1}^k \alpha_j d_i^T C d_j = 0. \text{ Como } d_i^T C d_j = 0 \quad \forall i \neq j.$$

se tiene :

$$\alpha_i d_i^T C d_i = 0, \forall i = 1, 2, \dots, k. \text{ Además } d_i^T C d_i > 0, \forall i = 1, 2, \dots, k$$

$$\Rightarrow \alpha_i = 0, \forall i = 1, 2, \dots, k.$$

Luego $\{ d_1, d_2, \dots, d_k \}$ son linealmente independientes. ■

LEMA 3:

Si $d_i \neq 0, i = 1, 2, \dots, k, k < n$ son conjugados por pares, entonces este sistema se deja extender a una base de vectores que son conjugados por pares.

DEMOSTRACION :

Dado que $k < n \Rightarrow \exists$ un vector $h_{k+1} \neq 0$ que no depende linealmente de d_1, d_2, \dots, d_k .

Definamos :

$$d_{k+1} = h_{k+1} - \sum_{i=1}^k \beta_i d_i \quad \dots \quad (**)$$

Determinemos los β_i de tal manera que d_{k+1} sea conjugado a d_1, d_2, \dots, d_k :

de (**):

$$d_{k+1}^T C d_j = h_{k+1}^T C d_j - \sum_{i=1}^k \beta_i d_i^T C d_j = 0, j = 1, 2, \dots, k$$

$$\Rightarrow h_{k+1}^T C d_j - \sum_{i=1}^k \beta_i d_i^T C d_j = h_{k+1}^T C d_j - \beta_j d_j^T C d_j = 0, j = 1, 2, \dots, k$$

Como $d_j^T C d_j > 0 \forall j = 1, 2, \dots, k$ se tiene :

$$\beta_j = \frac{h_{k+1}^T C d_j}{d_j^T C d_j}, j = 1, 2, \dots, k.$$

reemplazando β_j en (**):

$$d_{k+1} = h_{k+1} - \sum_{i=1}^k \left(\frac{h_{k+1}^T C d_i}{d_i^T C d_i} \right) d_i$$

$$\Rightarrow d_{k+1} \text{ es conjugado s } d_1, d_2, \dots, d_k.$$

Como $h_{k+1} \notin \text{Lin} \{ d_1, d_2, \dots, d_k \}$ y $h_{k+1} \neq 0$

$\Rightarrow d_1, d_2, \dots, d_k, d_{k+1}$ son linealmente independientes.

Si $k + 1 < n$ se repite el mismo procedimiento.

METODO DE GRAM-SCHMIDT

Proposición :

Si los vectores $\{ p_1, p_2, \dots, p_k \}$ son independientes, se puede construir vectores mutuamente conjugados $\{ d_1, d_2, \dots, d_k \}$ tal que :

$$\text{Lin} \{ d_1, d_2, \dots, d_k \} = \text{Lin} \{ p_1, p_2, \dots, p_k \}$$

Demostración :

Elijamos $d_1 = p_1$

Sean d_1, d_2, \dots, d_{i-1} vectores conjugados , ya construidos, tales que :

$$\text{Lin} \{ d_1, d_2, \dots, d_{i-1} \} = \text{Lin} \{ p_1, p_2, \dots, p_{i-1} \}$$

Elijamos :

$$d_i = p_i - \sum_{\ell=1}^{i-1} \alpha_{i\ell} d_\ell, \quad i \geq 2 \quad (***)$$

Debemos determinar los coeficientes $\alpha_{i\ell}$, tales que :

$$d_i^T C d_j = 0, \quad j = 1, 2, \dots, i-1 :$$

De (***) tenemos :

$$\begin{aligned} d_i^T C d_j &= p_i^T C d_j - \sum_{\ell=1}^{i-1} \alpha_{i\ell} d_\ell^T C d_j = 0 \\ \Rightarrow p_i^T C d_j - \alpha_{ij} d_j^T C d_j &= 0, \quad j = 1, 2, \dots, i-1 \\ \Rightarrow \alpha_{ij} &= \frac{p_i^T C d_j}{d_j^T C d_j}, \quad j = 1, 2, \dots, k. \end{aligned}$$

reemplazando α_{ij} en (***) :

$$d_i = p_i - \sum_{\ell=1}^{i-1} \left(\frac{p_i^T C d_\ell}{d_\ell^T C d_\ell} \right) d_\ell$$

Veamos ahora que :

$$\text{Lin} \{ p_1, p_2, \dots, p_i \} = \text{Lin} \{ d_1, d_2, \dots, d_i \}$$

$$\text{Sea } x \in \text{Lin} \{ p_1, \dots, p_i \} \Rightarrow x = y + \lambda p_i$$

$$y \in \text{Lin} \{ p_1, p_2, \dots, p_{l-1} \} = \text{Lin} \{ d_1, d_2, \dots, d_{l-1} \}$$

$$\lambda p_l \in \text{Lin} \{ d_1, d_2, \dots, d_{l-1}, d_l \}$$

$$\Rightarrow x \in \text{Lin} \{ d_1, d_2, \dots, d_{l-1}, d_l \}$$

$$\Rightarrow \text{Lin} \{ p_1, p_2, \dots, p_l \} \subset \text{Lin} \{ d_1, d_2, \dots, d_l \}$$

Analogamente se verifica la otra contencion.

$$\text{Luego : } \text{Lin} \{ d_1, d_2, \dots, d_k \} = \text{Lin} \{ p_1, p_2, \dots, p_k \}$$

Proposicion 2:

Si $\{ d_1, d_2, \dots, d_n \}$ es una base de vectores conjugados, entonces :

$$C^{-1} = \sum_{l=1}^n \frac{d_l d_l^T}{d_l^T C d_l} = D$$

Demostracion :

$$\text{Sea } x = \sum_{k=1}^n \alpha_k d_k \in \mathbb{R}^n \text{ cualquiera}$$

$$\Rightarrow Cx = \sum_{k=1}^n \alpha_k C d_k$$

$$\Rightarrow DCx = D \sum_{k=1}^n \alpha_k C d_k = \sum_{l=1}^n \frac{d_l d_l^T}{d_l^T C d_l} \sum_{k=1}^n \alpha_k C d_k$$

$$= \sum_{l=1}^n \sum_{k=1}^n \alpha_k d_l \frac{d_l^T C d_k}{d_l^T C d_l} \quad \text{Como } d_1, d_2, \dots, d_n \text{ son conjugados}$$

$$\Rightarrow DCx = \sum_{l=1}^n \alpha_l d_l \frac{d_l^T C d_l}{d_l^T C d_l}$$

$$\Rightarrow DCx = \sum_{l=1}^n \alpha_l d_l = x, \forall x \in \mathbb{R}^n.$$

$$\Rightarrow DCx = x, \forall x \in \mathbb{R}^n$$

$$\rightarrow DC = I, \quad \rightarrow D = C^{-1}$$

Observacion :

Observación

$$C^{-1} = \sum_{i=1}^n \frac{d_i d_i^T}{d_i^T C d_i}$$

es una representación de C^{-1} como suma de matrices de rango 1

Consideremos nuevamente el programa :

$$P : \min \{ Q(x) = \frac{1}{2} x^T C x + p^T x \mid x \in \mathbb{R}^n \}$$

Sea (d_0, d_1, \dots, d_n) una base de vectores conjugados.

Elijamos x_0 , y :

$$x_{k+1} = x_k + \lambda_k d_k, \quad \lambda_k = - \frac{(C x_k + p)^T d_k}{d_k^T C d_k},$$

$$k = 0, 1, \dots, n-1.$$

Veamos cómo obtenemos λ_k :

Se elige x_{k+1} de tal manera que minimice Q en

$L = \{ x = x_k + \lambda d_k \mid \lambda \in \mathbb{R} \}$, y luego, por el Lema 1 :

$$\nabla Q(x_{k+1})^T d_k = 0$$

$$\Rightarrow \nabla Q(x_k + \lambda_k d_k)^T d_k = [C(x_k + \lambda_k d_k) + p]^T d_k = 0$$

$$\Rightarrow (C x_k + p)^T d_k + \lambda_k d_k^T C d_k = 0$$

$$\Rightarrow \lambda_k = - \frac{(C x_k + p)^T d_k}{d_k^T C d_k}$$

■

TEOREMA 1 :

Para $0 \leq j \leq k$ se cumple $\nabla Q(x_{k+1})^T d_j = 0$. En consecuencia x_{k+1} minimiza Q en el subespacio afín:

$$L = x_0 + \text{Lin} \{ d_0, d_1, \dots, d_k \}$$

Demostración :

1) Si $j=k$

$$\Rightarrow \nabla Q(x_{k+1})^T d_k = 0.$$

2) Si $j < k$

$$\begin{aligned} \Rightarrow \nabla Q(x_{k+1}) &= Cx_{k+1} + p \\ &= C(x_{j+1} + \sum_{i=1}^k \lambda_i d_i) + p \\ &= (Cx_{j+1} + p) + \sum_{i=1}^k \lambda_i Cd_i \end{aligned}$$

$$\Rightarrow \nabla Q(x_{k+1}) = \nabla Q(x_{j+1}) + \sum_{i=1}^k \lambda_i Cd_i$$

$$\Rightarrow \nabla Q(x_{k+1})^T d_j = \nabla Q(x_{j+1})^T d_j + \sum_{i=1}^k \lambda_i d_i^T Cd_j$$

$$\Rightarrow \nabla Q(x_{k+1})^T d_j = 0.$$

Luego $\nabla Q(x_{k+1})^T d_j = 0 \forall 0 \leq j \leq k$.

Entonces por el Lema 1, x_{k+1} minimiza a Q en el subespacio

afín : $L = x_0 + \text{Lin} \{ d_0, d_1, \dots, d_k \}$ ■

CONSECUENCIAS :

El método propuesto resuelve el programa P a los más en n pasos :

$$\nabla Q(x_n)^T d_j = 0, \quad j = 0, 1, \dots, n$$

$$\Rightarrow \nabla Q(x_n) = 0$$

$$\Rightarrow \nabla Q(x_m) = 0 \text{ para algún } m \leq n.$$

Veamos :

Sabemos que por el teorema anterior tenemos:

$$\nabla Q(x_n)^T d_j = 0, \quad j = 0, 1, \dots, n-1$$

$$\begin{aligned} \nabla Q(x_n)^T d_n &= [C(x_{n-1} + \lambda_{n-1} d_{n-1}) + p]^T d_n \\ &= (Cx_{n-1} + p)^T d_n + \lambda_{n-1} d_{n-1}^T Cd_n \\ &= \nabla Q(x_{n-1})^T d_n + \lambda_{n-1} d_{n-1}^T Cd_n \\ &= 0. \end{aligned}$$

luego : $\nabla Q(x_n)^T d_j = 0, \quad \forall j = 0, 1, \dots, n$ ■

Como generalmente no se conoce una base de vectores

conjugados, modificamos el método de minimización precedente de tal modo que las direcciones d_i , a largo de las cuales se minimiza Q , se generan sucesivamente.

Ponemos $p_i = -\nabla Q(x_i)$, $0 \leq i \leq n$.

Para $k < n$ hacemos la siguiente hipótesis de inducción :

Las direcciones conjugadas d_0, \dots, d_k han sido generadas por el método de Gram-Schmidt, tal que :

$$\text{Lin} \{ p_0, p_1, \dots, p_k \} = \text{Lin} \{ d_0, d_1, \dots, d_k \} \quad \dots(*)$$

Sean x_1, x_2, \dots, x_{k+1} los puntos de iteración generados por el método precedente.

El por el teorema anterior y la hipótesis de inducción (*) se cumple :

$$p_{k+1}^T d_j = p_{k+1}^T p_j = 0, \quad 0 \leq j \leq k.$$

Si $p_{k+1} = 0$, entonces x_{k+1} minimiza Q y el algoritmo termina.

Si $p_{k+1} \neq 0$, entonces $p_{k+1} = \nabla Q(x_{k+1}) \notin \text{Lin} \{ d_0, d_1, \dots, d_k \}$.

En consecuencia podemos generar d_{k+1} por el método de Gram-Schmidt, usando d_j , $0 < j \leq k$ y p_{k+1} obtenemos :

$$d_{k+1} = p_{k+1} - \sum_{j=1}^k \left(\frac{p_{k+1}^T C d_j}{d_j^T C d_j} \right) d_j \quad \dots(+)$$

Vamos a demostrar que para generar d_{k+1} se usa solamente p_{k+1} y d_k :

Si $j < k$, entonces se cumple para un $\lambda_j > 0$

$$0 = p_{k+1}^T (p_j - p_{j+1}) = \lambda_j p_{k+1}^T C d_j$$

Por consiguiente :

$$p_{k+1}^T C d_j = 0 \quad \forall j < k$$

De esto y de (+) obtenemos :

$$d_{k+1} = p_{k+1} - \frac{p_{k+1}^T C d_k}{d_k^T C d_k} d_k$$

De estas consideraciones se obtiene el método de los gradientes conjugados de HESTENES Y STIEFEL :

Inicio :

Se elige x_0 y $d_0 = -\nabla Q(x_0)$.

Iteración General :

Sean x_k y d_k ya construidos :

Calculamos:

$$\alpha) x_{k+1} = x_k + \lambda_k d_k ; \quad \lambda_k = - \frac{\nabla Q(x_k)^T d_k}{d_k^T C d_k}$$

(x_{k+1} minimiza Q en la recta $x = x_k + \lambda d_k$)

$$\beta) d_{k+1} = -\nabla Q(x_{k+1}) + \beta_k d_k ; \quad \beta_k = \frac{\nabla Q(x_{k+1})^T C d_k}{d_k^T C d_k}$$

(β_k es elegido de modo que $d_{k+1}^T C d_k = 0$) .

TEOREMA 2 :

- a) $d_j^T C d_\ell = 0$, $\forall \ell < j \leq n$.
- b) $\nabla Q(x_j)^T \nabla Q(x_\ell) = 0$, $\forall \ell < j \leq n+1$.
- c) $\nabla Q(x_j)^T d_\ell = 0$, $\forall \ell < j \leq n$.
- d) $\nabla Q(x_j)^T C d_\ell = 0$, $\forall \ell < j-1 \leq n$.

DEMOSTRACION :

a) i) $d_1^T C d_0 = 0$. (por la construcción de d_1)

ii) Hipótesis de inducción :

Para un k con $1 \leq k < n$ se cumple :

$$d_\ell^T C d_j = 0 \quad , \quad \forall \ell < j \leq k \quad \dots (+)$$

iii) Veamos que para $k+1$ se cumple :

$$d_{k+1}^T C d_\ell = 0 \quad , \quad \forall 0 \leq \ell < k+1 .$$

De (+) implica :

$$\nabla Q(x_j)^T \nabla Q(x_\ell) = 0 \quad , \quad \forall \ell < j \leq k+1$$

$$\nabla Q(x_j)^T d_\ell = 0 \quad , \quad \forall \ell < j \leq k .$$

$$\nabla Q(x_j)^T C d_\ell = 0 \quad , \quad \forall \ell < j-1 \leq k .$$

Veamos :

Por teorema 1 se cumple :

$$\nabla Q(x_j)^T d_\ell = 0, \quad \forall \ell < j \leq k+1.$$

Por $\beta)$ del algoritmo se cumple :

$$\nabla Q(x_\ell) = -d_\ell + \beta_{\ell-1} d_{\ell-1}$$

$$\Rightarrow \nabla Q(x_j)^T \nabla Q(x_\ell) = -\nabla Q(x_j)^T d_\ell + \beta_{\ell-1} \nabla Q(x_j)^T d_{\ell-1},$$

$\forall \ell < j \leq k+1.$

De esto se obtiene :

$$\nabla Q(x_j)^T C d_\ell = \nabla Q(x_j)^T \left[\frac{\nabla Q(x_{\ell+1}) - \nabla Q(x_\ell)}{\lambda_\ell} \right] = 0$$

$$\Rightarrow \nabla Q(x_j)^T C d_\ell = 0, \quad \forall \ell < j-1 \leq k.$$

Luego es suficiente demostrar que $iii)$ se verifica :

Por construcción se cumple :

$$d_{k+1}^T C d_k = 0$$

Para $\ell < k$ se cumple por la recursión y la hipótesis de inducción de (+) :

$$d_{k+1} = -\nabla Q(x_{k+1}) + \beta_k d_k$$

$$\Rightarrow d_{k+1}^T C d_\ell = -\nabla Q(x_{k+1})^T C d_\ell + \beta_k d_k^T C d_\ell = 0$$

$$\Rightarrow d_{k+1}^T C d_\ell = 0, \quad \forall \ell < k+1$$

con lo cual el teorema se verifica. ■

Aplicación del método de los gradientes conjugados a la

minimización de funciones no cuadráticas.

Las fórmulas de recursión $\alpha)$ y $\beta)$ del algoritmo de HESTENES y STIEFEL, se dejan transformar de diferentes maneras de modo que serán definidas también para funciones no cuadráticas.

De este modo se obtiene una clase de métodos que coinciden con el método de HESTENES y STIEFEL que se aplica a una función cuadrática convexa.

EL METODO DE FLETCHER REEVES

Como
$$Cd_k = \frac{\nabla Q(x_{k+1}) - \nabla Q(x_k)}{\lambda_k}$$

y
$$\beta_k = \frac{\nabla Q(x_{k+1})^T Cd_k}{d_k^T Cd_k}, \quad d_k = -\nabla Q(x_k) + \beta_{k-1} d_{k-1}$$

Entonces β_k toma la forma siguiente :

$$\beta_k = \frac{\|\nabla Q(x_{k+1})\|^2}{\|\nabla Q(x_k)\|^2}$$

El método de Fletcher-Reeves para implementar al algoritmo del gradiente conjugado para problemas no cuadrático; no requiere de la evaluación $F(x_k)$, pero si requiere de una búsqueda lineal.

Sea $F \in C^1(\mathbb{R}^n)$.

EL ALGORITMO DE FLETCHER-REEVES

Inicio :

Se elige x_0 y $d_0 = -\nabla F(x_0)$.

Iteración General :

Sean d_k y x_k ya construidos .

α) Se determina una solución λ_k del programa :

$$\min \{ g_k(\lambda) = F(x_k + \lambda d_k) \mid \|\lambda d_k\| \leq \gamma \}$$

donde γ es dado

y calculamos :

$$x_{k+1} = x_k + \lambda_k d_k$$

β) se calcula :

$$\beta_k = \frac{\|\nabla F(x_{k+1})\|^2}{\|\nabla F(x_k)\|^2}$$

$$d_{k+1} = -\nabla F(x_{k+1}) + \beta_k d_k$$

LA CONVERGENCIA DE LOS ALGORITMOS

Hipótesis: $F \in C^1(\mathbb{R}^n)$, $\inf_{x \in \mathbb{R}^n} F(x) = \alpha > -\infty$

$\bar{X} = \{ x / F(x) \leq F(x_0) \}$ es compacto

Consideramos una clase de algoritmos iterativos de minimización, cuya k -ésima iteración tiene la siguiente forma:

a) Si x_i es dado, se calcula una dirección d_i :

b) $x_{i+1} = x_i + \lambda_i d_i$, donde:

$$\lambda_i = \inf \{ \lambda > 0 / \nabla F(x_i + \lambda d_i)^T d_i = 0 \}$$

Observamos que los algoritmos de direcciones conjugadas (en especial, el algoritmo de Fletcher - Reeves) se dejan ejecutar de este modo.

Teorema:

Hipótesis:

i) $\nabla F(x_k) \neq 0$, $\forall k$

ii) $r_k = - \frac{\nabla F(x_k)^T d_k}{\|\nabla F(x_k)\| \|d_k\|} > 0$, $\forall k$

iii) $\sum_{k=0}^{\infty} r_k^2 = \infty$

iv) $\exists M \in (0, \infty)$ tal que:

$$\|\nabla F(x) - \nabla F(y)\| \leq M \|x - y\|$$

Tesis:

$\{x_k\}$ tiene al menos un punto de acumulación que es estacionario.

Demostración:

Por la hipótesis ii), se tiene $\{x_k\} \subset \bar{X}$, $\forall k$.

Como \bar{X} es compacto, entonces $\{x_k\}$ tiene al menos un punto de acumulación.

Supongamos que $\liminf_k \|\nabla F(x_k)\| > 0 \Rightarrow \exists \varepsilon > 0$ tal que:

$$\|\nabla F(x_k)\| \geq \varepsilon, \forall k; \dots \dots \dots (*)$$

de la hipótesis ω) tenemos que:

$$\nabla F(x_k)^T d_k = -r_k \|\nabla F(x_k)\| \|d_k\|, \forall k \dots\dots\dots (**)$$

Por otro lado:

$$\begin{aligned} \nabla F(x_k + \lambda d_k)^T d_k &= \nabla F(x_k)^T d_k + [\nabla F(x_k + \lambda d_k) - \nabla F(x_k)]^T d_k \\ \rightarrow \nabla F(x_k + \lambda d_k)^T d_k &\leq \nabla F(x_k)^T d_k + \|\nabla F(x_k + \lambda d_k) - \nabla F(x_k)\| \|d_k\| \end{aligned}$$

Por la hipótesis ω) y (**):

$$\begin{aligned} \nabla F(x_k + \lambda d_k)^T d_k &\leq \nabla F(x_k)^T d_k + \|\nabla F(x_k + \lambda d_k) - \nabla F(x_k)\| \|d_k\| \leq \\ &\leq -r_k \|\nabla F(x_k)\| \|d_k\| + M \lambda \|d_k\|^2 \leq \\ &\leq -r_k \varepsilon \|d_k\| + M \lambda \|d_k\|^2 \quad [\text{por } (*)] \end{aligned}$$

$$\Rightarrow \nabla F(x_k + \lambda d_k)^T d_k \leq -r_k \varepsilon \|d_k\| + M \lambda \|d_k\|^2, \forall k \quad (**')$$

Como $\nabla F(x_k + \lambda_k d_k)^T d_k = 0$, por la desigualdad anterior tiene:

$$0 \leq -r_k \varepsilon \|d_k\| + M \lambda_k \|d_k\|^2, \quad \|d_k\| > 0$$

entonces: $\lambda_k \geq \frac{\varepsilon r_k}{M \|d_k\|} = \mu_k \dots\dots (**')$

De otro lado:

$$F(x_{k+1}) - F(x_k) \leq \int_0^{\lambda_k} \nabla F(x_k + \lambda d_k)^T d_k \, d\lambda$$

Además:

$$\nabla F(x_k + \lambda d_k)^T d_k < 0, \forall \lambda \in [0, \lambda_k]$$

De (**') y (**') se tiene:

$$\begin{aligned} F(x_{k+1}) - F(x_k) &\leq \int_0^{\mu_k} \nabla F(x_k + \lambda d_k)^T d_k \, d\lambda \leq \\ &\leq \int_0^{\mu_k} [-r_k \varepsilon \|d_k\| + M \lambda \|d_k\|^2] \, d\lambda = -\frac{(r_k \varepsilon)^2}{M} \end{aligned}$$

$$\Rightarrow F(x_{k+1}) - F(x_k) \leq -\frac{(r_k \varepsilon)^2}{M}$$

$$\rightarrow \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_k)] = F(x_n) - F(x_0) \leq -\frac{\varepsilon^2}{M} \sum_{k=0}^{n-1} r_k^2$$

$$\begin{aligned} \Rightarrow \lim_{n \rightarrow \infty} F(x_n) &\leq \lim_{n \rightarrow \infty} \left[F(x_0) - \frac{\varepsilon^2}{M} \sum_{k=0}^{n-1} r_k^2 \right] = \\ &= F(x_0) - \frac{\varepsilon^2}{M} \sum_{k=0}^{\infty} r_k^2 = -\infty \\ \Rightarrow \lim_{n \rightarrow \infty} F(x_n) &= -\infty. \quad (\Rightarrow \Leftarrow) \end{aligned}$$

Dado que $F(x_n) \leq F(x_0)$, $\forall n$ en \bar{X} , $\{x_n\} \subset \bar{X}$, \bar{X} compacto,

$F \in C^1(\bar{X})$.

Luego: $\liminf_n \|\nabla F(x_n)\| = 0$.

$\Rightarrow \{x_n\}$ tiene al menos un punto de acumulación que es estacionario.

TEOREMA :

Si \bar{X} es compacto y $\|\nabla F(x) - \nabla F(y)\| \leq M \|x - y\|$, $\forall x, y \in \bar{X}$, entonces la sucesión $\{x_k\}$ generada por el método de Fletcher-Reeves tiene al menos un punto estacionario.

Demostración :

$$\text{Si } g_k = \nabla F(x_k) \text{ y } S_k = -\frac{d_k}{\|g_k\|}$$

entonces la fórmula :

$$d_k = -g_k + \frac{\|g_k\|^2}{\|g_{k-1}\|^2} d_{k-1}$$

se convierte en:

$$S_k = -\frac{g_k}{\|g_k\|^2} + S_{k-1} \quad \dots (*)$$

Como $S_0 = -\frac{g_0}{\|g_0\|^2}$, obtenemos de (*):

$$S_k - S_0 = \sum_{i=1}^k (S_i - S_{i-1}) = -\sum_{i=1}^k \frac{g_i}{\|g_i\|^2}$$

$$\Rightarrow S_k = S_0 - \sum_{i=1}^k \frac{g_i}{\|g_i\|^2} = - \frac{g_0}{\|g_0\|^2} - \sum_{i=1}^k \frac{g_i}{\|g_i\|^2}$$

$$\Rightarrow S_k = - \sum_{i=0}^k \frac{g_i}{\|g_i\|^2}$$

Como $g_k^T d_{k-1} = 0 = g_k^T \frac{d_{k-1}}{\|g_{k-1}\|^2} = g_k^T S_{k-1} = 0$,

de (*) obtenemos:

$$g_k^T S_k = - \frac{g_k^T g_k}{\|g_k\|^2} + g_k^T S_{k-1}$$

$$\Rightarrow g_k^T S_k = -1 \quad \dots (**)$$

$$\|S_k\|^2 = \left(- \frac{g_k}{\|g_k\|^2} + S_{k-1} \right)^T \left(- \frac{g_k}{\|g_k\|^2} + S_{k-1} \right) =$$

$$= \frac{g_k^T g_k}{\|g_k\|^4} + S_{k-1}^T S_{k-1}$$

$$\|S_k\|^2 = \frac{1}{\|g_k\|^2} + \|S_{k-1}\|^2$$

$$\Rightarrow \|S_k\|^2 - \|S_{k-1}\|^2 = \frac{1}{\|g_k\|^2}, \quad \|S_0\|^2 = \frac{1}{\|g_0\|^2}$$

$$\rightarrow \sum_{i=1}^k [\|S_i\|^2 - \|S_{i-1}\|^2] = \|S_k\|^2 - \|S_0\|^2 = \sum_{i=1}^k \frac{1}{\|g_i\|^2}$$

$$\Rightarrow \|S_k\|^2 = \sum_{i=1}^k \frac{1}{\|g_i\|^2} \quad \dots (***)$$

Por otro lado tenemos:

$$r_k = - \frac{g_k^T d_k}{\|g_k\| \|d_k\|} = - \frac{g_k^T}{\|g_k\|} \left(\frac{d_k}{\|g_k\|^2} \right) \frac{1}{\frac{\|d_k\|}{\|g_k\|^2}} =$$

$$= \frac{\sigma_k^2 S_k}{\|g_k\| \|S_k\|}$$

De (**) y (***) tenemos:

$$r_k = \frac{1}{\|g_k\| \left(\sum_{i=0}^k \frac{1}{\|g_i\|^2} \right)^{1/2}} > 0$$

Supongamos que $\liminf \|g_k\| > 0 \Rightarrow \exists \varepsilon > 0$ tal que :

$$\|g_k\| \geq \varepsilon, \forall k$$

Como X es compacto y $(x_k) \subset X, \forall k$,

$\Rightarrow \exists M > 0$ tal que : $\|g_k\| \leq M, \forall k$.

$$\text{Luego : } \left(\sum_{i=0}^k \frac{1}{\varepsilon^2} \right)^{1/2} \geq \left(\sum_{i=0}^k \frac{1}{\|g_i\|^2} \right)^{1/2}$$

$$\Rightarrow \left(\sum_{i=0}^k \frac{1}{\|g_i\|^2} \right)^{1/2} \leq \left(\sum_{i=0}^k \frac{1}{\varepsilon^2} \right)^{1/2} \leq \frac{(k+1)^{1/2}}{\varepsilon}$$

$$\rightarrow \left(\sum_{i=0}^k \frac{1}{\|g_i\|^2} \right)^{1/2} \leq \frac{(k+1)^{1/2}}{\varepsilon}$$

$$\Rightarrow \beta_k = \|g_k\| \left(\sum_{i=0}^k \frac{1}{\|g_i\|^2} \right)^{1/2} \leq M \frac{(k+1)^{1/2}}{\varepsilon}$$

$$\Rightarrow r_k = \frac{1}{\beta_k} \geq \frac{\varepsilon}{M(k+1)^{1/2}}$$

$$\rightarrow \sum_{k=0}^{\infty} r_k^2 \geq \frac{\varepsilon^2}{M^2} \sum_{k=0}^{\infty} \frac{1}{k+1} = \infty$$

$$\Rightarrow \sum_{k=0}^{\infty} r_k^2 = \infty$$

De acuerdo al teorema precedente, \exists un punto de acumulación de $\{x_k\}$, digamos \bar{x} , tal que $\nabla F(\bar{x}) = 0$, lo que implica que

$\liminf_k \|g_k\| = 0$. Esto contradice el hecho que :

$\liminf_k \|g_k\| > 0$.

Luego : $\liminf_k \|g_k\| = 0$. ■

Metodo de Daniel :

Sea $F \in C^2(\mathbb{R}^n)$; ahora sustituimos $\nabla^2 F(x_{k+1})$ en lugar de C en β_k (en el método de Hestenes-Stiefel).

INICIO :

Sea x_0 dado y $d_0 = -\nabla F(x_0)$.

ITERACION GENERAL :

Sean x_k , d_k , ya construidos :

1) Se determina :

$$\lambda_k = \inf_{\lambda \geq 0} \{ \nabla F(x_k + \lambda d_k)^T d_k = 0 \}$$

y se calcula :

$$x_{k+1} = x_k + \lambda_k d_k$$

2) Se calcula $\nabla^2 F(x_{k+1})$ y

$$d_{k+1} = -\nabla F(x_{k+1}) + \left(\frac{\nabla F(x_{k+1})^T \nabla^2 F(x_{k+1}) d_k}{d_k^T \nabla^2 F(x_{k+1}) d_k} \right) d_k$$

Metodo PARTAN :

En los primeros experimentos con el método de descenso más rápido, se notó que la trayectoria de descenso era altamente zig-zageante, progresando poco hacia la solución. También se notó que en dos dimensiones, el punto solución cae sobre la recta que conecta a los puntos zig-zageantes.

Estas observaciones constituyeron la motivación para acelerar el método del gradiente, que consiste en completar el círculo tomando dos pasos de descenso más rápido y luego hacer la búsqueda lineal que conecta el punto inicial y el punto obtenido después de dos pasos de gradiente.

El método de las TANGENTES PARALELAS (PARTAN) se desarrolló como un intento de extender la idea anterior hacia un esquema de aceleración que involucra todos los pasos previos. El desarrollo original se basó grandemente en una propiedad geométrica especial de las tangentes a los contornos de una función cuadrática, pero ahora, el método se reconoce como una implementación particular del método de gradientes conjugados.

ALGORITMO PARTAN :

El método construye dos sucesiones $\{ x_k \}$, $\{ \xi_k \}$ de puntos de iteración:

a) $x_1 = x_0 - \lambda_0 \nabla F(x_0)$, donde :

$$g_0(\lambda_0) = \min_{\lambda \geq 0} \{ g_0(\lambda) = F(x_0 - \lambda \nabla F(x_0)) \}$$

b) para $k \geq 1$ se procede de la siguiente manera:

$$\xi_k = x_k - \eta_k \nabla F(x_k) \text{ donde:}$$

$$g_k(\eta_k) = \min \{ g_k(\eta) = F(x_k - \eta \nabla F(x_k)) \}$$

$$x_{k+1} = x_{k-1} - \lambda_k (\xi_k - x_{k-1}) \text{ , donde:}$$

$$h_k(\lambda_k) = \min \{ h_k(\lambda) = F(x_{k-1} - \lambda(\xi_k - x_{k-1})) \}$$

METODO DE DIRECCIONES CONJUGADAS QUE NO USA DERIVADA

Sea $Q(x) = \frac{1}{2} x^T C x + p^T x$, donde $C = C^T > 0$.

El método se basa en la siguiente proposición :

PROPOSICION :

Sean q_1, q_2, \dots, q_k , $k \leq n$, vectores linealmente independientes. Asumimos que los puntos x_0 , x_1 minimizan

$Q(x)$ en los subespacios afines :

$$L_a = \left\{ x \in \mathbb{R}^n \mid x = a + \sum_{i=1}^n \lambda_i q_i \right\}$$

$$L_b = \left\{ x \in \mathbb{R}^n \mid x = b + \sum_{i=1}^n \lambda_i q_i \right\}$$

Entonces se cumple:

$$(x_0 - x_1)^T C q_i = 0 \quad i = 1, 2, \dots, k$$

Demostración :

Como x_0, x_1 minimizan $Q(x)$, entonces se cumple:

$$\left. \begin{aligned} \nabla Q(x_0)^T q_i &= (Cx_0 + p)^T q_i = 0 \\ \nabla Q(x_1)^T q_i &= (Cx_1 + p)^T q_i = 0 \end{aligned} \right\} \quad i = 1, 2, \dots, k$$

$$\Rightarrow (x_0 - x_1)^T C q_i = 0 \quad , \quad i = 1, 2, \dots, k$$

En base a la proposición anterior podemos construir , iterativamente una base de direcciones conjugadas como sigue:

-Sean las direcciones conjugadas d_1, d_2, \dots, d_k , $k < n$ ya construidos . Si

$$L_k = \text{Lin} \{ d_1, d_2, \dots, d_k \}$$

$$M_k = x_0 + L_k$$

determinamos la solución x_k de :

$$\min \{ Q(x) \mid x \in M_k \}$$

Elegimos ahora un punto $z_k \in M_k$, y determinamos la solución óptima x_{k+1} de :

$$\min \{ Q(x) \mid x \in z_k + L_k \}$$

Entonces para $d_{k+1} = x_{k+1} - x_k$ se cumple :

$$d_{k+1}^T C d_i = 0 \quad , \quad i = 1, 2, \dots, k$$

EL ALGORITMO :

Sea $B = \{ a_1, a_2, \dots, a_n \}$ una base de \mathbb{R}^n .

Paso Inicial :

Se elige x_0 y determina la solución λ_1 de:

$$\min \{ Q(x) \mid x = x_0 + \lambda a_1, \lambda \in \mathbb{R} \}$$

$$y \quad x_1 = x_0 + \lambda_1 a_1, \quad d_1 = a_1$$

Paso General :

Sean dados los vectores mutuamente conjugados d_1, d_2, \dots, d_i , los puntos de iteración x_0, x_1, \dots, x_i y

$$B_i = (a_{j_1}, a_{j_2}, \dots, a_{j_{n-i}}, d_1, d_2, \dots, d_i) = (b_1, b_2, \dots, b_n)$$

donde los a_{j_k} son vectores de la base inicial B.

a) La generación de d_{i+1} :

Para $k = 1, 2, \dots, n-i+1$ se calcula los puntos auxiliares ξ_i^k como sigue :

$$\xi_i^1 = x_i$$

Para $k > 1$ se elige :

$$\xi_i^k = \xi_i^{k-1} + \eta_{k-1} b_{k-1}.$$

donde η_{k-1} minimiza la función :

$$\tilde{q}_{k-1}(\eta) = Q(\xi_i^{k-1} + \eta b_{k-1})$$

Si $\eta_k = 0$, $k = 1, 2, \dots, n-i$, entonces x_i es óptimo.

En otro caso se elige :

$$d_{i+1} = \xi_i^{n-i+1} - x_i$$

b) Se determina la solución λ_{i+1} de :

$$\min_{\lambda} (q_{i+1}(\lambda) = Q(x_i + \lambda d_{i+1}) / \lambda \in \mathbb{R})$$

$$y \quad x_{i+1} = x_i + \lambda_{i+1} d_{i+1}$$

c) Se forma el sistema B_{i+1} sustituyendo en B_i un vector b_k , $k \leq n-i$, para el cual $\eta_k \neq 0$, por d_{i+1} .

TEOREMA :

El método minimiza Q en a lo más n iteraciones .

Demosttracin :

Asumimos que los vectores d_1, d_2, \dots, d_i , $i < n$ sean mutuamente conjugados y que:

$B_i = (a_{j_1}, a_{j_2}, \dots, a_{j_{r-i}}, d_1, d_2, \dots, d_i)$ sea una base.

Vamos a demostrar que si x_i no es óptimo, que $d_{i+1}^T C d_j = 0$ $j < i + 1$ y que B_{i+1} es una base.

De acuerdo con el teorema 1 x_i minimiza Q en el subespacio afín:

$$L_i = \{ x / x = x_0 + \sum_{k=1}^i \nu_k d_k \}$$

Si $\eta_k = 0$, $k \leq n - i$, entonces $\xi_i^{n+1} \in L_i$.

Como $Q(\xi_i^{n+1}) \leq Q(x_i) = \min_{x \in L_i} Q(x)$, se cumple:

$$\xi_i^k = x_i, \quad k = 1, 2, \dots, n + 1.$$

Además se cumple:

$$\nabla Q(x_i)^T b_j = 0, \quad j = 1, 2, \dots, n \quad \dots (+)$$

Si fuera j_0 el primer índice con $\nabla Q(x_i)^T b_{j_0} \neq 0$, se

cumpliría $\eta_{j_0} \neq 0$ y $\xi_i^{j_0+1} = x_i + \eta_{j_0} b_{j_0} \neq x_i$.

Como B_i es una base, (+) implica que $\nabla Q(x_i) = 0$. Por consiguiente x_i es óptimo.

Si $\eta_k \neq 0$ para algún $k \leq n - i$ y $x = x_i + \sum_{k=1}^{n-i} \eta_k b_k$, entonces

ξ_i^{n+1} se encuentra en el subespacio afín:

$$\tilde{L}_i = \{ x / x = \tilde{x} + \sum_{k=1}^i \nu_k d_k \}$$

$$\Rightarrow \xi_i^{n+1} \text{ minimiza a } Q \text{ en } \tilde{L}_i.$$

Por consiguiente de acuerdo con la proposición precedente,

$$d_{i+1} = \xi_i^{n+1} - x_i = \sum_{k=1}^n \eta_k b_k \neq 0 \text{ es conjugado con } d_k, \quad \forall k \leq i$$

Si en la formación de B_{i+1} el vector b_j es sustituido por d_{i+1} , entonces $\eta_j \neq 0 \Rightarrow B_{i+1}$ es una base.

Como $B_1 = (a_1, a_2, \dots, a_n)$ es una base, queda demostrado el teorema. ■

MÉTODOS QUASI-NEWTON

Los métodos quasi-Newton que construyen una aproximación al Hessiano inverso, son analíticamente los métodos más sofisticados para resolver problemas sin restricciones, y representan la culminación del desarrollo de los algoritmos, mediante el análisis detallado del problema cuadrático. Como podría esperarse, las propiedades de convergencia de estos métodos son algo más difíciles de hallar que aquellas de métodos más simples.

CONSTRUCCION DE LA INVERSA :

La idea fundamental que existe en los métodos quasi-Newton es tratar de construir el Hessiano inverso o una aproximación de él, usando información obtenida conforme progresa el proceso de descenso. Luego, esta aproximación se usa en cada etapa para definir la siguiente dirección de descenso. Idealmente, las aproximaciones convergen a la inversa del Hessiano en el punto solución, y todo el método se comporta de modo parecido al método de Newton.

Vamos a construir el Hessiano inverso a partir de la información del gradiente obtenida en varios puntos.

Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función, $f \in C^2(\mathbb{R}^n)$.

Sean x_{k+1} , x_k dos puntos.

$$g_k = \nabla f(x_k)$$

$$g_{k+1} = \nabla f(x_{k+1})$$

$$p_k = x_{k+1} - x_k$$

Luego :

$$g_{k+1} - g_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

entonces por el teorema de valor medio :

$$g_{k+1} - g_k = \nabla f(x_{k+1}) - \nabla f(x_k) = \nabla^2 f(x_k + \theta(x_{k+1} - x_k))(x_{k+1} - x_k)$$

para $\theta \in]0,1[$.

$$g_{k+1} - g_k = \nabla^2 f(x_k + \theta p_k) p_k = F(x_k + \theta p_k) p_k \quad (1)$$

para algún $\theta \in (0,1)$.

si F es constante, tenemos :

$$q_k = g_{k+1} - g_k = F p_k \quad (2)$$

Vemos que al evaluar $\nabla f = F$ en x_k , x_{k+1} da información acerca de F .

Sean n direcciones $\{p_0, p_1, \dots, p_{n-1}\}$ L.I. y los

correspondientes $\{q_0, q_1, \dots, q_{n-1}\}$, definamos :

$P = (p_0, p_1, \dots, p_{n-1})$, $Q = (q_0, q_1, \dots, q_{n-1})$ luego obtenemos :

$$FP = Q. \text{ Como } \exists \text{ la inversa de } P \text{ (} p_i \text{ son L.I. } \forall i \text{)}$$

$$\Rightarrow F = QP^{-1}$$

Queremos construir aproximaciones S_k a F^{-1} basados en los K primeros pasos (de un proceso de descenso) de tal manera que si F es constante debe satisfacer (2) , es decir :

$$S_{k+1} q_j = p_j \quad 0 \leq j \leq k \quad (3)$$

Luego de n pasos L.I. se tendría $S_n = F^{-1}$.

CORRECCION DE RANGO UNO

S_k es una aproximación de F^{-1} .

F , F^{-1} son simétricos , entonces S_k es simétrico.

Definamos :

$$S_{k+1} = S_k + a_k z_k z_k^T \quad (4)$$

donde el vector z_k , y la constante a_k , define una matriz de rango (la cual actualiza la aproximación de la inversa), se escoge de tal manera que (3) se satisfice.

De (3) y (4), para $j = k$ obtenemos:

$$p_k = S_{k+1} q_k = S_k q_k + a_k z_k z_k^T q_k \quad (5)$$

De (5) obtenemos:

$$q_k^T p_k - q_k^T S_k q_k = a_k (z_k^T q_k)^T z_k^T q_k \quad (6)$$

De (5) obtenemos:

$$z_k = \frac{p_k - S_k q_k}{a_k z_k^T q_k} ; \quad (z_k^T q_k = q_k^T z_k)$$

De (4) y de la última relación, se tiene:

$$S_{k+1} = S_k + \frac{(p_k - S_k q_k)(p_k - S_k q_k)^T}{a_k (q_k^T z_k)^2}$$

De esto y de (6), obtenemos finalmente:

$$S_{k+1} = S_k + \frac{(p_k - S_k q_k)(p_k - S_k q_k)^T}{q_k^T (p_k - S_k q_k)}$$

TEOREMA :

Sea F una matriz simétrica fija, y sean dados los vectores p_0, p_1, \dots, p_k .

Definamos:

$$q_j = F p_j, \quad j = 0, \dots, k$$

Iniciando con cualquier matriz simétrica S_0 ,

$$\text{sea : } S_{j+1} = S_j + \frac{(p_j - S_j q_j)(p_j - S_j q_j)^T}{q_j^T (p_j - S_j q_j)}$$

Entonces: $p_j = S_{k+1} q_j$, para $j \leq k$.

Demostración :

Sabemos que se cumple: $p_k = S_{k+1} q_k$

Para $j < k$ se cumple para $y_k = \frac{p_k - S_k q_k}{q_k^T (p_k - S_k q_k)}$

$$\begin{aligned} S_{k+1} q_j &= S_k q_j + \frac{(p_k - S_k q_k)(p_k - S_k q_k)^T}{q_k^T (p_k - S_k q_k)} q_j = \\ &= S_k q_j + (p_k^T q_j - q_k^T S_k q_j) y_k \end{aligned}$$

De la hipótesis de inducción obtenemos :

$$S_{k+1} q_j = p_j + (p_k^T q_j - q_k^T p_j) y_k \quad \dots (*)$$

pero $q_j = Fp_j \quad j = 0, 1, \dots, k$

$$\Rightarrow q_k^T p_j = p_k^T F p_j = p_k^T q_j$$

reemplazando en (*) tenemos:

$$S_{k+1} q_j = p_j \quad \forall j < k$$

Luego : $S_{k+1} q_j = p_j \quad \forall j \leq k$

METODO DE FLETCHER-POWELL-DAVIDON

El más antiguo, y ciertamente, uno de los esquemas más ingeniosos esquemas para construir el Hessiano inverso, fue propuesto originalmente por Davidon y desarrollado posteriormente por Fletcher y Powell. Tiene la fascinante y deseable propiedad que, para una función objetivo cuadrática, genera simultáneamente las direcciones del método gradientes conjugados y construye a la vez el Hessiano inverso. En cada paso, se actualiza el Hessiano inverso mediante la suma de dos matrices simétricas de rango uno, y por tanto, este esquema se denomina un procedimiento corrección de rango dos. También suele llamarse un método de métrica variable, nombre propuesto originalmente por Davidon.

Sea S_0 una matriz definida positiva simétrica

x_0 cualquier punto dado en \mathbb{R}^n .

Iniciamos con $k = 0$

PAso 1. - Hacemos $d_k = -S_k g_k$.

PAso 2. - Determinamos α_k de manera que sea solución del programa : $\min \{ f(x_k + \alpha d_k) \mid \alpha \geq 0 \}$ y calculamos :

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k \\ p_k &= \alpha_k d_k \\ g_{k+1} &= \nabla f(x_{k+1}) \end{aligned}$$

PASO 9. - hacemos :

$$q_k = g_{k+1} - g_k$$

$$S_{k+1} = S_k + \frac{p_k p_k^T}{p_k^T q_k} - \frac{S_k q_k q_k^T S_k}{q_k^T S_k q_k}$$

Incrementamos k y regresamos al paso (1).

Proposición :

Si S_k es definida positiva , entonces S_{k+1} es definida positiva.

Demostración :

Sea $x \in \mathbb{R}^n \setminus \{0\}$ cualquiera, tenemos :

$$x^T S_{k+1} x = x^T S_k x + \frac{x^T p_k p_k^T x}{p_k^T q_k} - \frac{(x^T S_k q_k)(q_k^T S_k x)}{q_k^T S_k q_k}$$

$$= x^T S_k x + \frac{(x^T p_k)^2}{p_k^T q_k} - \frac{(x^T S_k q_k)^2}{q_k^T S_k q_k} \quad \dots (*)$$

Como S_k es definida positiva existe una matriz simétrica $S_k^{1/2}$

con $S_k^{1/2} S_k^{1/2} = S_k$,

definamos $a = S_k^{1/2} x$, $b = S_k^{1/2} q_k$

Entonces (*) se convierte en:

$$x^T S_{k+1} x = \frac{(a^T a)(b^T b)}{b^T b} + \frac{(x^T p_k)^2}{p_k^T q_k} \quad (**)$$

también tenemos :

$$p_k^T q_k = p_k^T g_{k+1} - p_k^T g_k = - p_k^T g_k$$

como : $p_k^T g_{k+1} = 0$

Dado que x_{k+1} minimiza f a lo largo de p_k .

Luego por la definición de p_k se tiene :

$$p_k^T q_k = \alpha_k g_k^T S_k g_k$$

Luego en (**):

$$x^T S_{k+1} x = \frac{(a^T a)(b^T b) - (a^T b)^2}{b^T b} + \frac{(x^T p_k)^2}{\alpha_k g_k^T S_k g_k}$$

Los dos términos de la relación anterior son mayores que cero tenemos que probar que para $x \neq 0$, los dos términos no se anulan simultáneamente.

El primer término se anula si y solo si $a = rb$.

Como $S_k^{1/2} x = a = rb = r S_k^{1/2} q_k \Rightarrow x = r q_k$, $r \neq 0$

$$\Rightarrow p_k^T x = r p_k^T q_k = r \alpha_k g_k^T S_k g_k$$

$$\Rightarrow x^T S_{k+1} x > 0, \forall x \in \mathbb{R}^n \setminus \{0\}$$

$\Rightarrow S_{k+1}$ es definida positiva.

Convergencia:

Sea f una función cuadrática, con Hessiano $F = \nabla^2 f$ constante.

TEOREMA:

Si f es cuadrática con Hessiano F definida positiva

Entonces por el método de Fletcher-Powell-Davidon se tiene:

$$p_i^T F p_j = 0, 0 \leq i < j < k \tag{7}$$

$$S_{k+1} F p_i = p_i, \text{ para } 0 \leq i \leq k \tag{8}$$

Demostración:

En el caso cuadrático se tiene:

$$q_k = g_{k+1} - g_k = F p_k$$

también:

$$\begin{aligned} S_{k+1} F p_k &= S_{k+1} q_k = S_k q_k + \frac{p_k p_k^T q_k}{p_k^T q_k} - \frac{S_k q_k q_k^T S_k q_k}{q_k^T S_k q_k} = \\ &= S_k q_k + p_k - S_k q_k = p_k \end{aligned}$$

Probaremos por inducción:

Para $k = 0$ se tiene:

$$S_1 F p_0 = p_0$$

(para $k = 0$, (7) se cumple porque no existe $1 < j \leq 0$)

Supongamos que para $k = r$ se cumple:

$$p_i^T F p_j = 0, \quad 0 \leq i < j < r \quad (7a)$$

$$S_{r+1} F p_i = p_i, \quad \text{para } 0 \leq i \leq r \quad (8a)$$

Veamos para $k = r+1$:

Para $i < r$ se tiene:

$$\begin{aligned} g_{r+1} &= F x_r + p = \\ &= F [x_{i+1} + (x_{i+2} - x_{i+1}) + \dots + (x_r - x_{r-1})] + p \\ &= g_{i+1} + \sum_{l=i+1}^r F p_l \\ \Rightarrow p_i^T g_{r+1} &= p_i^T g_{i+1} + \sum_{l=i+1}^r p_i^T F p_l \end{aligned}$$

Luego, de (7a) y como $p_k^T g_{k+1} = 0$:

$$p_i^T g_{r+1} = p_i^T g_{i+1} = 0, \quad \text{para } 0 \leq i < r+1$$

De (8a):

$$\begin{aligned} p_i^T &= p_i^T S_{r+1} \\ \Rightarrow p_i^T g_{r+1} &= p_i^T S_{r+1} g_{r+1} = 0, \quad 0 \leq i \leq r \quad (9) \end{aligned}$$

Como $p_k = -\alpha_k S_k g_k$, $\alpha_k \neq 0$, en (9):

$$p_i^T F p_{r+1} = 0, \quad 0 \leq i < r+1 \quad (10)$$

De (10) y (8a), obtenemos:

$$\begin{aligned} q_{r+1}^T S_{r+1} F p_i &= q_{r+1}^T p_i = p_{r+1}^T F p_i = 0 \\ S_{r+2} F p_i &= S_{r+1} F p_i + \frac{p_{i+1} p_{i+1}^T F p_i}{p_{r+1}^T q_{r+1}} - \frac{S_{r+1} q_{r+1} q_{r+1}^T S_{r+1} F p_i}{q_{r+1}^T S_{r+1} q_{r+1}} = \\ &= S_{r+1} F p_i = p_i, \quad 0 \leq i < r+1 \\ \Rightarrow S_{r+2} F p_i &= p_i, \quad 0 \leq i < r+1 \end{aligned}$$

Luego, el teorema queda probado. ■

EL METODO DE LOS GRADIENTES CONJUGADOS COMO
PROCESO OPTIMO

Resumimos primero algunas propiedades conocidas del método de los gradientes conjugados:

TEOREMA :

Si el algoritmo de las direcciones conjugadas no termina en x_k , entonces:

a) $\text{Lin} \{ g_0, g_1, \dots, g_k \} = \text{Lin} \{ g_0, Cg_0, \dots, C^k g_0 \}$

b) $\text{Lin} \{ d_0, d_1, \dots, d_k \} = \text{Lin} \{ g_0, Cg_0, \dots, C^k g_0 \}$

c) $d_k^T C d_i = 0, i \leq k-1$

d) $\alpha_k = -\frac{g_k^T g_k}{d_k^T C d_k}$

e) $\beta_k = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$



Consideremos la función cuadrática:

$$F(x) = \frac{1}{2} x^T C x + p^T x$$

donde $C = C^T \in M(n,n)$ definida positiva.

El único mínimo x^* de F en \mathbb{R}^n está caracterizado por:

$$\nabla F(x^*) = Cx^* + p = 0 \quad \text{ó} \quad Cx^* = -p$$

Para $k < n$, sean d_0, \dots, d_k y $\bar{x}_0, \dots, \bar{x}_k, \bar{x}_{k+1}$ las direcciones conjugadas, y los puntos de iteración generados por el método de los gradientes conjugados.

Partimos de lo siguiente:

- Si $\mathcal{B}_{k+1} = x_0 + \text{Lin} \{ d_0, \dots, d_k \}$

entonces \bar{x}_{k+1} minimiza a : $F(x) = \frac{1}{2} x^T C x + p^T x$ en \mathcal{B}_{k+1} .

- $\text{Lin} \{ d_0, d_1, \dots, d_k \} = \text{Lin} \{ g_0, Cg_0, \dots, C^k g_0 \}$

si el algoritmo no termina con x_k .

Sea P_k un polinomio de grado k y

$$x_{k+1} = x_0 + P_k(C)g_0 \quad (1)$$

$$\begin{aligned} \text{entonces: } x_{k+1} - x^* &= x_0 - x^* + P_k(C)g_0 \\ &= x_0 - x^* + P_k(C)(Cx_0 + p) = \\ &= x_0 - x^* + P_k(C)C(x_0 - x^*) = \\ x_{k+1} - x^* &= [I + P_k(C)C](x_0 - x^*) \quad (*) \end{aligned}$$

Consideramos la función de error :

$$E(x) = \frac{1}{2} (x - x^*)^T C (x - x^*)$$

Entonces:

$$\begin{aligned} E(x_{k+1}) &= \frac{1}{2} (x_{k+1} - x^*)^T C (x_{k+1} - x^*) \\ &= \frac{1}{2} x_{k+1}^T C x_{k+1} - (x^*)^T C x_{k+1} + \frac{1}{2} (x^*)^T C x^* \\ &= \frac{1}{2} x_{k+1}^T C x_{k+1} + p^T x_{k+1} + \frac{1}{2} (x^*)^T C x^* \\ &= F(x_{k+1}) + \frac{1}{2} (x^*)^T C x^* \end{aligned}$$

Por (*) se cumple :

$$E(x_{k+1}) = \frac{1}{2} (x_0 - x^*)^T C [I + P_k(C)C]^2 (x_0 - x^*)$$

Problema : Encontrar un polinomio P_k^* de grado de k que minimice $E(x_{k+1})$.

Reescribimos (1) :

$$x_{k+1} = x_0 + \gamma_0 g_0 + \gamma_1 C g_0 + \dots + \gamma_k C^k g_0 \quad \dots (+)$$

Como :

$$\mathcal{B}_{k+1} = \text{Lin} (d_0, d_1, \dots, d_k) = \text{Lin} (g_0, Cg_0, \dots, C^k g_0)$$

el vector $\bar{x}_{k+1} = x_0 + \alpha_0 d_0 + \alpha_1 d_1 + \dots + \alpha_k d_k$, generado por el método de gradientes conjugados tiene precisamente la forma (+).

Como \bar{x}_{k+1} minimiza $F(x)$ en \mathcal{B}_{k+1} , los coeficientes γ_k , determinados por el método de los gradientes conjugados tiene el efecto de minimizar $E(x_{k+1})$.

El método de los gradientes conjugados resuelve nuestro problema de seleccionar un polinomio óptimo

TEOREMA 1: El punto x_{k+1} generado por el método de los gradientes conjugados cumple :

$$E(x_{k+1}) = \min_{P_k} \frac{1}{2} (x_0 - x^*)^T C [I + P_k (C) C] J^2 (x_0 - x^*)$$

donde el mínimo se toma con respecto a los polinomios P_k de grado menor o igual a k .

NOTA : Las relaciones entre α_j , β_j del método de los gradientes conjugados y los coeficientes γ_i de P_k son complejos.

COTAS PARA LA CONVERGENCIA :

Sean e_1, \dots, e_n vectores propios ortonormales de C :

$$C e_i = \lambda_i e_i$$

$$x_0 - x^* = \sum_{i=1}^n \xi_i e_i$$

$$\begin{aligned} \Rightarrow E(x_0) &= \frac{1}{2} (x_0 - x^*)^T C (x_0 - x^*) \\ &= \frac{1}{2} \left(\sum_{i=1}^n \xi_i e_i \right)^T C \sum_{i=1}^n \xi_i e_i \end{aligned}$$

$$E(x_0) = \frac{1}{2} \left(\sum_{i=1}^n \xi_i e_i \right)^T \left(\sum_{i=1}^n \lambda_i \xi_i e_i \right) = \frac{1}{2} \sum_{i=1}^n \lambda_i \xi_i^2$$

Se cumple también:

$$C [I + P_k (C) C] J^2 e_i = \lambda_i [1 + P_k(\lambda_i)] \lambda_i J^2 e_i$$

$$\Rightarrow E(x_{k+1}) = \frac{1}{2} (x_0 - x^*)^T C [I + P_k (C) C] J^2 (x_0 - x^*)$$

$$= \frac{1}{2} \left(\sum_{j=1}^n \xi_j e_j \right)^T \sum_{i=1}^n \lambda_i \xi_i [1 + P_k(\lambda_i)] \lambda_i J^2$$

$$= \frac{1}{2} \sum_{i=1}^n \xi_i^2 \lambda_i [1 + P_k(\lambda_i)] \lambda_i J^2$$

$$\Rightarrow E(x_{k+1}) \leq \left[\max_{\lambda_i} [1 + P_k(\lambda_i)] \lambda_i J^2 \right] \frac{1}{2} \sum_{i=1}^n \lambda_i \xi_i^2$$

$$\Rightarrow E(x_{k+1}) \leq \left[\max_{\lambda_i} [1 + P_k(\lambda_i)] \lambda_i J^2 \right] E(x_0) \quad (++)$$

P_k es un polinomio arbitrario de grado k .

Cada paso del método de gradientes conjugados es al menos tan bueno como una iteración de descenso más rápido.

Sea x_k calculado por el método de gradientes conjugados y x_{k+1} , por el método de descenso más rápido, \bar{x}_{k+1} , calculado por el método de gradientes conjugados.

$$\Rightarrow x_{k+1} = x_k + \alpha_k g_k$$

$$x_k = x_0 + \bar{\gamma}_0 g_0 + \bar{\gamma}_1 C g_0 + \dots + \bar{\gamma}_{k-1} C^{k-1} g_0$$

$$\text{Lin} \{ g_0, g_1, \dots, g_k \} = \text{Lin} \{ g_0, C g_0, \dots, C^k g_0 \}$$

$$\Rightarrow x_{k+1} - x_0 = \sum_{i=1}^k \gamma_i C^i g_0$$

$$\Rightarrow E(x_{k+1}) \leq E(x_k)$$

Si se dispone de información sobre los valores propios, ésta puede usarse para construir polinomios P_k adecuados en (++)).

Ejemplo :

C tiene $m < n$ valores propios distintos, entonces podemos elegir P_{m-1} tal que el polinomio :

$$1 + \lambda P_{m-1}(\lambda)$$

de grado m tenga sus ceros en los valores propios de C

$$\Rightarrow E(x_m) = 0$$

Se requiere sólo de m iteraciones (y no de n) para llegar al óptimo.

METODO PARCIAL DE GRADIENTES CONJUGADOS

Se ejecuta $m + 1 < n$ pasos del método de gradientes conjugados, y se reinicia con un paso de descenso más rápido.

x_k será el punto de iteración después de la realización de k iteraciones de $m+1$ subpasos

$$\begin{cases} m=0 \rightarrow \text{subpaso descenso.} \\ m=n-1 \Rightarrow \text{grad. conj.} \end{cases}$$

Consideramos la iteración:

$$x_{k+1} = x_k + P_k(C)g_k$$

donde P_k es un polinomio de grado m .

Elegimos P_k tal que:

$$E(x_{k+1}) = \frac{1}{2} (x_{k+1} - x^*)^T C (x_{k+1} - x^*)$$

sea mínimo.

Con frecuencia, se tiene lo siguiente:

C tiene m valores propios grandes y $n-m$ valores propios pequeños (en $[a,b]$).

Ejemplo :

Consideramos:

$$\min \frac{1}{2} x^T P x + p^T x$$

$$c^T x = 0$$

Método de penalización:

$$\min \frac{1}{2} x^T P x + p^T x + \mu (c^T x)^2$$

$P = P^T \in M(n,n)$ definida positiva con los valores propios en $[a,A]$, μ grande.

$$(c^T x)^2 = x^T c c^T x$$

El término cuadrático de la función objetivo es $x^T (P + \mu c c^T) x$.

Sea $C = P + \mu c c^T$.

Si $\mu \rightarrow \infty$, entonces los m valores propios de C tienden a ∞ y los restantes se mantienen en $[a,A]$.

Si se aplicara el método de descenso más rápido, la convergencia sería lenta (gobernada por el cociente

$$\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \approx 1, \text{ donde } \lambda_n \text{ es el mayor valor propio de } C \text{ y } \lambda_1$$

es el menor valor propio de C).

La convergencia de los métodos de gradientes conjugados es afectado por el cociente anterior.

Los métodos quasi-Newton aproximan iterativamente la inversa de la matriz Hessiana $\nabla^2 F$ (en el caso cuadrático: $\nabla^2 F = C$), (el método de Fletcher-Powell es quasi-Newton y al mismo tiempo un método de direcciones conjugadas).

Si el factor $\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}$ es grande, entonces $\nabla^2 F$ es mal condicionado y el cálculo de las aproximaciones a $(\nabla^2 F)^{-1}$ se vuelve numéricamente inestable.

Si se aplica el método de gradientes conjugados con $m+1$ pasos de gradientes conjugados, el efecto de los m valores propios es eliminado.

TEOREMA :

HIPOTESIS:

$U = U^T \in M(n,n)$ definida positiva .

U tiene $n-m$ valores propios en $[a,b]$, $a > 0$.

Los restantes valores propios son mayores que b .

TESIS :

Para el método parcial de gradientes conjugados, que se reinicia siempre después de $m+1$ pasos, es válida la estimación :

$$E(x_{k+1}) \leq \frac{(b-a)^2}{(b+a)^2} E(x_k)$$

Demostración :

Se cumple: $E(x_{k+1}) \leq \left[\max_{\lambda_i} \left(1 + \lambda_i P(\lambda_i) \right)^2 \right] E(x_k)$

para todo polinomio P de orden m .

Elegimos P , tal que el polinomio de grado $m + 1$

$q(\lambda) = 1 + \lambda P(\lambda)$ se anula en $\frac{a+b}{2}$ y en los m valores propios de U que son mayores que b .

Como $q(\lambda)$ tiene $m + 1$ ceros reales , $q'(\lambda)$ tiene m ceros reales que se alternan con las raíces de $q(\lambda)$.

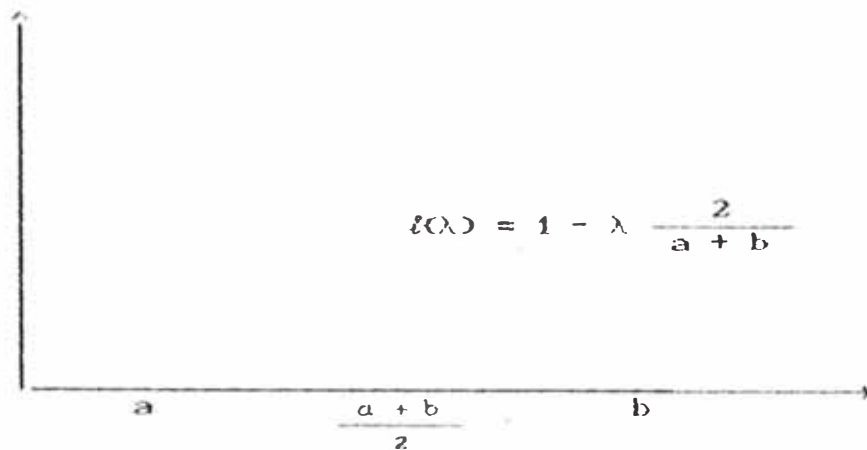
$q''(\lambda)$ tiene $m - 1$ raíces reales que se alternan con las raíces de $q'(\lambda)$.

$q(\lambda)$ no tiene ceros en $\langle -\infty, \frac{a+b}{2} \rangle = I$

$\Rightarrow q''(\lambda)$ no cambia de signo en I .

se verifica que : $q''(0) > 0$.

$\Rightarrow q$ es estrictamente convexa en I .



Consideremos $l(\lambda) = 1 - \lambda \frac{2}{a+b}$ la recta que pasa por los puntos :

$$(0,1) = (0, q(0)) \text{ , y } \left(-\frac{a+b}{2}, 0\right) = \left(-\frac{a+b}{2}, q\left(-\frac{a+b}{2}\right)\right).$$

$$\Rightarrow q(\lambda) \leq 1 - \lambda \frac{2}{a+b} \text{ , } \forall \lambda \in \left[0, \frac{a+b}{2}\right] \quad \dots (1)$$

$$\Rightarrow \frac{q(\lambda) - q\left(-\frac{a+b}{2}\right)}{\lambda - \left(-\frac{a+b}{2}\right)} = \frac{q(\lambda)}{\lambda - \frac{a+b}{2}} \geq \frac{2}{2\lambda - a - b} \left(\frac{a+b-2\lambda}{a+b}\right) = -\frac{2}{a+b} \text{ , } \forall \lambda \in \left[0, \frac{a+b}{2}\right].$$

$$\Rightarrow q'\left(-\frac{a+b}{2}\right) \geq -\frac{2}{a+b} \text{ , } \forall \lambda \in \left[0, \frac{a+b}{2}\right].$$

Consideremos la tangente a $q(\lambda)$ en $\left(-\frac{a+b}{2}, b\right)$.

Como en el caso anterior, obtenemos :

$$q'\left(-\frac{a+b}{2}\right) \geq -\frac{2}{a+b} \text{ , } \forall \lambda \in \left(-\frac{a+b}{2}, b\right]$$

$$\Rightarrow q(\lambda) \geq q'\left(-\frac{a+b}{2}\right)\left(\lambda - \frac{a+b}{2}\right) \geq$$

$$\geq \left(-\frac{2}{a+b}\right)\left(\lambda - \frac{a+b}{2}\right) = 1 - \frac{2}{a+b} \lambda \quad \dots (2)$$

de (1) y (2) se obtiene :

$$|q(\lambda)| = |1 + \lambda P(\lambda)| \leq \left|1 - \frac{2\lambda}{a+b}\right| \text{ , } \forall \lambda \in [a,b]$$

$$\Rightarrow \max |1 + \lambda_i P(\lambda_i)|^2 \leq \left|1 - \frac{2\lambda}{a+b}\right|^2 = \left(\frac{b-a}{b+a}\right)^2$$

$$\Rightarrow E(x_{k+1}) \leq \left(\frac{b-a}{b+a}\right)^2 E(x_k) \quad \blacksquare$$

BIBLIOGRAFIA

I ACTAS DE CONGRESOS :

- (1) Abadie (J.),Ed. Non linear Programming
North Holland, Amsterdam, 1970
- (2) Fletcher (R.),Ed. Optimización
Academic Press, London, 1969.

II LIBROS

- (3) Avriel Non linear programming
Analysis and Methods.
- (4) Blum (E.), Nichtlineare Optimierung.
Oetli (W.) Springer, Berlin , 1975
- (5) Kuenzi (M.R.) Non linear Programming
Rishel (R.W.) Blaisdell, Waltham, Mass 1969
- (6) Sposito (V.A.) Linear and Non linear
Programming.
Iowa State Univ, Press 1975.
- (7) Luenberger Introduction to Linear and non
Linear Programming.
Addison-Wesley P. C. 1973.