UNIVERSIDAD CENTROCCIDENTAL
“LISANDRO ALVARADO”

Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas

OPTIMIZACIÓN GLOBAL VIA SUB-ESTIMACIÓN LINEAL A TROZOS

Trabajo especial de grado presentado por:

Br. Yulimar González Rodríguez

Como requisito final

para obtener el título de Licenciado

en Ciencias Matemáticas

Área de conocimiento: Matemática Aplicada

Director: Javier Hernández Benítez.

Barquisimeto, Venezuela octubre 2008

Le dedico con mucho amor y cariño

a mis queridos padres que me dieron la vida

y la mejor herencia "una carrera para mi futuro"...

Agradecimientos

Primero y antes que nada quiero darle gracias a Dios, por protegerme, por iluminarme el camino, por darme fuerzas para seguir con esta meta y por haber puesto en mi camino a aquellas personas maravillosas que han sido mi soporte y compañía durante todo mi periodo universitario.

A mis padres, por ser los mejores, por su cariño, comprensión, amor y apoyo incondicional... por motivarme siempre a seguir adelante. Por lo que soy... Los Amo. Esta meta culminada es de ustedes.

A mi papi Salvador, por ser tan especial y lindo conmigo, por siempre ayudarme, apoyarme, motivarme de que si se puede, por ser uno de mis ejemplos a seguir, porque gracias a ti pude empezar y culminar; siempre estuviste conmigo ; gran parte de lo que soy es gracias a ti... Te Amo mi novio Bello...

A mis hermanas Yohana y Yanet por ser mi gran ejemplo a seguir en todo, por ser tan maravillosas conmigo.

A mi tutor Dr. Javier Hernández por compartir parte de sus conocimientos conmigo, ayudarme con la realización de este trabajo de grado y por siempre transmitirme su tranquilidad.

A la Flia Acosta, Sr. Henri, Sra. Arelis, Carlos y Elsa por ser personas que me brindaron todo el apoyo, colaboración y cariño sin ningún interés, de verdad gracias porque me ayudaron cuando más lo necesité.

A mi querido abuelo Domingo, porque mientras te tuve vivo me apoyaste, me ayudaste muchísimo … y desde que te fuiste tus recuerdos me han dado fuerzas y muchas alegrías; Abuelo hoy tu niña de ojos pardos que no los tiene azules está culminando una de sus grandes metas con mucho orgullo…

A la universidad Centroccidental “Lisandro Alvarado” y a mis profesores Ebner, Abelardo, Eibar, Miguel, Jhonny, Ismael, Mario, Alexander, Pedro Adames, Romulo Castillo, Javier, Luz, Romulo Bervins, gracias a todos ustedes aprendí muchas cosas que me van a servir para ser una gran profesional como lo son y lo fueron ustedes…

A mis grandes amigos Manuel, Elifer y Nelson por ayudarme mucho en la parte académica y enseñarme tantas cosas que me sirvieron durante toda mi carrera…

A AsoEM, a todos mis compañeros de mi promoción, Luis Freytez, Dicmar, Maria, Marilyn, Tony, Joel, Toñeco, Liseth, Jorge C., Jorge M., Maria L., Jesus, Adriana, Luis, Wilennis, Maxiel, Mayggret, Lina, David, Memo, Ronald, Gregorio, Mario, Reyvi, Elizabeth por hacer de mi vida universitaria muy agradable.

A toda mi familia, Abuela Maura, Abuelo Quiterio, mi abuela Antonia, a mis tíos, tías y a todos mis primos muchas gracias por estar tan pendientes de mi…

A mi querido abuelo Salvador, suegros, tíos, tías, primos, compadres, por estar tan pendientes de mí y porque han sido personas muy especiales conmigo…

Y a todas aquellas personas que se me olvidó nombrar, pero colaboraron para que este gran logro se realizara, *MUCHAS GRACIAS.*

OPTIMIZACIÓN GLOBAL VIA SUB-ESTIMACIÓN LINEAL A TROZOS

Br. Yulimar González Rodríguez

RESUMEN

Dada una función en $\mathbb{R}^n$ con varios mínimos locales, la finalidad es encontrar el mínimo global de dicha función. Para ello se puede aproximar la función con una función lineal a trozos o una función cuadrática convexa a trozos mediante puntos de muestra, cuyo mínimo es una aproximación del mínimo global de la función original.

Índice general

Índice de figuras

Capítulo 1

Preliminares

1.1. El problema de optimización

Un problema de optimización trata de tomar una decisión óptima para maximizar (ganancias, velocidad, eficiencia, etc) o minimizar (costos, tiempo, riesgos, error, etc) un criterio determinado. Las restricciones significan que no cualquier decisión es posible.

La optimización o también denominada programación matemática intenta dar respuesta a un tipo general de problemas de la forma:

$$\text{mín } f(x), \quad x \in C \subseteq \mathbb{R}^n, \tag{1.1}$$

donde $x = (x_1, x_2, \ldots, x_n)^t$ es un vector y representa variables de decisión, $f : C \to \mathbb{R}$ es llamada función objetivo y representa o mide la calidad de las decisiones (usualmente números enteros o reales) y C es el conjunto de restricciones o región factible del problema.

Algunas veces es posible expresar la región factible C como solución de un sistema de igualdades o desigualdades. En esos casos, el problema de optimización se formula de la forma siguiente

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ s.\,a. \quad & h_i(x) \leq 0, \quad i \in \mathcal{I}, \\ & g_j(x) = 0, \quad j \in \mathcal{E}, \end{aligned} \tag{1.2}$$

donde $h_i, g_j : \mathbb{R}^n \to \mathbb{R}$ y $\mathcal{I}$ y $\mathcal{E}$ son conjuntos de índices.

Dado un punto factible $\bar{x} \in \mathbb{R}$ del problema (1.2) denotamos por $\mathcal{A}(\bar{x})$ al conjunto de todas las restricciones que se cumplen como igualdad que en el futuro llamaremos restricciones activas del punto $\bar{x}$, es decir

$$\mathcal{A}(\bar{x}) = \mathcal{E} \cup \{i \in \mathcal{I} \,|h_i(\bar{x})\}\,. \tag{1.3}$$

Definición 1. *[8] Dado un punto $x^* \in \mathbb{R}$. Diremos que $\mathcal{A}(x^*)$ cumple con la condición de independencia lineal de las restricciones si el conjunto de los gradientes de las restricciones activas $\{\nabla h_i(x^*) \ : \ i \in \mathcal{A}(x^*)\} \cup \left\{\nabla g_j(x^*) \ : \ j \in \mathcal{E}\right\}$ son linealmente independientes.*

Teorema 1 (Condición necesaria de primer orden)**.** *[8] Sea x^* una solución local de* (1.2) *tal que $\mathcal{A}(x^*)$ cumple la condición de independencia lineal de las restricciones.*

Entonces existen vectores λ^*, μ^* llamados multiplicadores de Lagrange, con componentes λ_i^*, $i \in \mathcal{I}$ y μ_j^*, $j \in \mathcal{E}$, tal que las siguientes condiciones se cumplen en (x^*, λ^*, μ^*)

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) &= 0, \\ g_i(x^*) &= 0 \quad \textit{para todo} \;\; i \in \mathcal{E}, \\ h_i(x^*) &\geq 0 \quad \textit{para todo} \;\; i \in \mathcal{I}, \\ \lambda_i^* &\geq 0 \quad \textit{para todo} \;\; i \in \mathcal{I}, \\ \lambda_i^* h_i(x^*) &= 0 \quad \textit{para todo} \;\; i \in \mathcal{I}, \end{aligned} \tag{1.4}$$

donde $\mathcal{L}(x, \lambda, \mu) = f(x) - \sum_{i \in \mathcal{I}} \lambda_i h_i(x) - \sum_{j \in \mathcal{E}} \mu_j g_j(x)$.

Las condiciones necesarias de optimalidad (1.4) son conocidas como condiciones de Karush-Kuhn-Tucker. Un punto (x, λ, μ) que satisface (1.4) es llamado *punto Karush-Kuhn-Tucker*, o en forma abreviada, *punto KKT*.

Veamos un pequeño ejemplo ilustrativo [8],

$$\begin{aligned} &\min (x_1 - 2)^2 + (x_2 - 1)^2 \\ &s.a \quad \begin{cases} {x_1}^2 - x_2 \leq 0 \\ x_1 + x_2 - 2 \leq 0 \end{cases} \end{aligned}$$

Ahora tenemos que,

$$g(x) = (x_1 - 2)^2 + (x_2 - 1)^2 \quad \text{y} \quad C(x) = \begin{bmatrix} C_1(x) \\ C_2(x) \end{bmatrix} = \begin{bmatrix} -{x_1}^2 + x_2 \\ -x_1 - x_2 + 2 \end{bmatrix}$$

La Figura 1.1 muestra los contornos de la función objetivo $g(x)$. También ilustra la región factible C (la región sombreada), que es el conjunto de puntos que cumplen

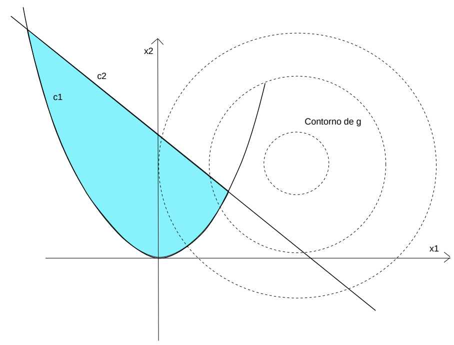

Figura 1.1: Representación geométrica del problema.

todas las limitaciones.

1.2. Matrices definidas positivas

Definición 2. *[1] Sea* $x \in \mathbb{R}^n$ *y* $p \in [1, \infty)$*, la norma p denotada por* $\|x\|_p$*, es la función* $\|\cdot\|_p| : \mathbb{R}^n \longrightarrow [0, +\infty)$ *dada por:* $\|x\|_p = (\sum_{k=1}^n |x|^p)^{\frac{1}{p}}$.

Las matrices simétricas tienen especiales propiedades, en particular, respecto a sus autovalores y autovectores. $\|\cdot\|$ denota la norma matricial inducida.

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

Proposición 1. *[1] Sea* $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ *una matriz simétrica. Entonces*

(a) *Los autovalores de* A *son reales.*

(b) *A tiene un conjunto ortogonal de autovectores no nulos $x_1, x_2, \dots, x_n$.*

(c) *Si $\|x_i\| = 1$ $i = 1, \dots, n$ en (b), entonces*

$$A = \sum_{i=1}^{n} \lambda_i x_i x_i^{\tau}$$

$\lambda_i, \dots, \lambda_n$ son los autovalores de A, asociados a $x_1, \dots, x_n$ respectivamente.

Definición 3. *[1] Sea $A \in \mathbb{M}_n(\mathbb{R})$ simétrica. Diremos que A es definida positiva si $x^T Ax > 0$ $\forall x \in \mathbb{R}^n, x \neq 0$. A es semidefinida positiva si $x^T Ax \geq 0$ $\forall x \in \mathbb{R}^n$.*

Proposición 2. *[1] Sea $A \in \mathbb{M}_{m\times n}(\mathbb{R})$*

(a) *$A^T A$ es semidefinida positiva. $A^T A$ es definida positiva si rang$(A) = n$. En particular, si $m = n$, $A^T A$ es definida positiva si y solo si A es no singular.*

(b) *Si $m = n$ y A es simétrica, A es semidefinida positiva (definida positiva) si y solo si todos sus autovalores son no negativos (positivos).*

(c) *La inversa de una matriz simétrica definida positiva es definida positiva.*

(d) *A es simétrica definida positiva si el det$(\triangle_k) > 0$, $k = 1, \dots, n$ donde*

$$\triangle_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{1k} & \dots & a_{kk} \end{pmatrix}.$$

Definición 4. *Una función cuadrática viene dada por*

$$q(x) = \alpha + c^t x + \frac{1}{2} x^t H x,$$

donde $\alpha \in \mathbb{R}, c \in \mathbb{R}^n$ y $H \in \mathbb{R}^{n\times n}$.

- f es una función cuadrática estrictamente convexa si H es simétrica definida positiva.
- f es una función cuadrática estrictamente cóncava si H es simétrica definida negativa.

1.3. Conjuntos convexos, funciones convexas

Definición 5. *[6] C es un conjunto convexo si para cada par $(x, y) \in C \times C$ se cumple que $\alpha x + (1 - \alpha)y \in C$, $\forall \alpha \in [0, 1]$. La expresión $\alpha x + (1 - \alpha)y$ es el segmento de recta que une a x con y, es decir: C es convexa si para cada par de puntos de él, el segmento que los une está totalmente incluido en C.*

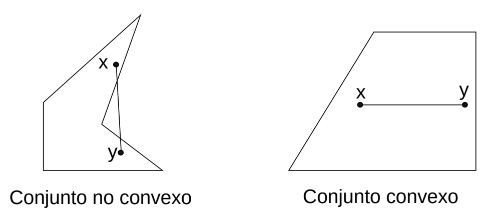

Figura 1.2: Conjunto convexo y no-convexo

Definición 6. *[6] Sea $C \subseteq \mathbb{R}^n$ un conjunto convexo no vacío. Una función $f : C \to \mathbb{R}$ se llama convexa en C cuando, para todo par $(x, y) \in C \times C$ y para todo $\alpha \in [0, 1]$, se tiene*

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y). \tag{1.5}$$

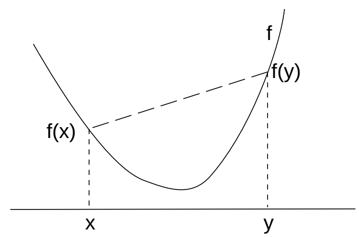

Figura 1.3: Función convexa

Una función cóncava es lo opuesto de una función convexa, es decir, f es una función cóncava entonces $(-f)$ es una función convexa. Así su definición sería,
Sea C un conjunto convexo no vacío en $\mathbb{R}^n$. Una función $f : C \to \mathbb{R}$ se llama cóncava en C cuando, para todo par $(x, y) \in C \times C$ y para todo $\alpha \in [0, 1]$, se tiene

$$f(\alpha x + (1 - \alpha)y) \geq \alpha f(x) + (1 - \alpha)f(y). \tag{1.6}$$

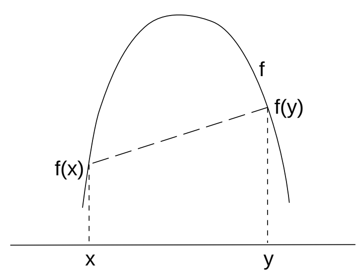

Figura 1.4: Función cóncava

Definición 7. *[8] Sea $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ una función, esta tiene un mínimo global en $x^* \in C$ si se cumple*

$$f(x^*) \leq f(x), \quad \forall x \in C \tag{1.7}$$

y de forma similar sería, Sea $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ una función, $x^* \in C$ es un mínimo local de f si existe una vecindad $B(x^*)$ de x^* tal que

$$f(x^*) \leq f(x), \quad \forall x \in B(x^*) \tag{1.8}$$

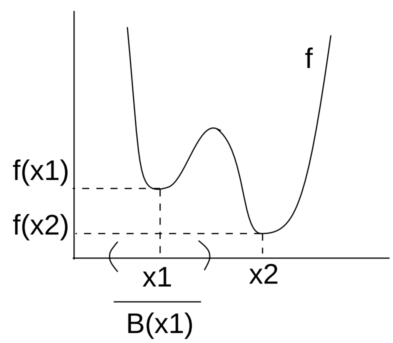

Figura 1.5: Mínimo local (x_1) en una vecindad de x_1 y el mínimo global x_2 de la función.

Hay funciones que el gradiente en un punto no es fácil de encontrar o simplemente no existe, por ello surge el supergradiente para funciones cóncavas y el subgradiente para las funciones convexas. Así podríamos definirlas como [4]:

- Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función cóncava, el vector $\partial f(x)$ es un supergradiente de f en x, con $x \in \mathbb{R}^n$, si satisface lo siguiente:

$$f(y) - f(x) \leq \partial f(x)(y - x) \text{ para cualquier } y \in \mathbb{R}^n. \tag{1.9}$$

- Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa, el vector $\partial f(x)$ es un subgradiente de f en x, con $x \in \mathbb{R}^n$, si:

$$f(x) - f(y) \geq \partial f(x)(x - y) \text{ para cualquier } y \in \mathbb{R}^n. \tag{1.10}$$

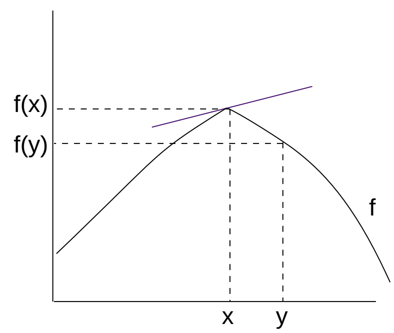

Figura 1.6: Supergradiente de f

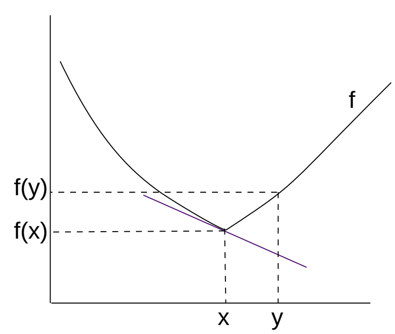

Figura 1.7: Subgradiente de f

Capítulo 2

Sub-estimador cuadrático.

2.0.1. Sub-estimador cuadrático.

Definición 8 (Sub-estimador cuadrático)**.** *[7] Sea $f : \mathbb{R}^n \to \mathbb{R}$, y consideremos la data (x^j, y^j) con $x^j \in \mathbb{R}^n$ y $y^j = f(x^j), j = 1, \ldots, m$, un sub-estimador cuadrático estrictamente convexo de la función f es una función de la forma*

$$q(x) = \alpha + c^t x + \frac{1}{2} x^t H x,$$

donde $\alpha \in \mathbb{R}, c \in \mathbb{R}^n$ y $H \in \mathbb{R}^{n \times n}$ es una matriz simétrica definida positiva, que cumple

$$q(x^j) \leqslant y^j$$

Dado un conjunto $\{x^1, x^2, \ldots, x^m\}$, se puede obtener un sub-estimador cuadrático resolviendo el siguiente problema de ajuste:

$$\begin{aligned}
&\min_{\alpha,c,H} \sum_{j=1}^{m} \left(y^j - \left(\alpha + c^t x^j + \frac{1}{2}(x^j)^t H x^j \right) \right) \\
&s.a \qquad \alpha + c^t x^j + \tfrac{1}{2}(x^j)^t H x^j \leqslant y^j, \qquad j = 1, \ldots, m, \\
&\qquad\qquad H,\ S.D.P.
\end{aligned}$$

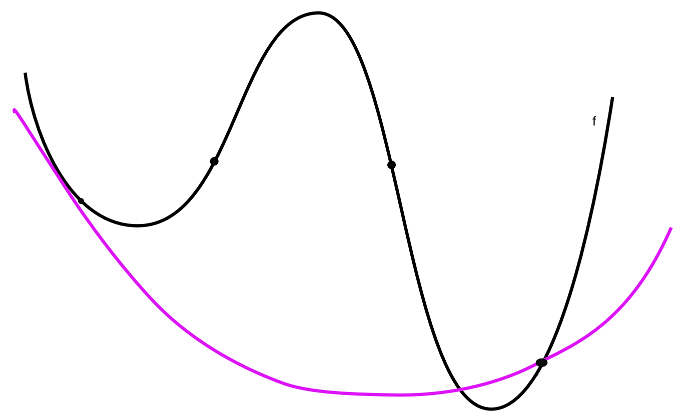

Figura 2.1: Sub-estimador cuadrático

Aproximar una función mediante un sub-estimador cuadrático puede ser impreciso (ver [4]) y además la condición definida positiva de la matriz H es complicado obtener y en consecuencia q puede no ser un buen sub-estimador. Es por ello que se propone estudiar el sub-estimador cuadrático a trozos.

2.0.2. Sub-estimador cuadrático a trozos.

El problema que estamos interesados en resolver es encontrar el mínimo global de una función $f : \mathbb{R}^n \to \mathbb{R}$. Consideremos la data:

$$f(x^j) = y^j,\ j = 1, \ldots, m. \tag{2.1}$$

Definición 9 (Sub-estimador cuadrático a trozos). *[5] Una función convexa a trozos es de la forma:*

$$q(x; p) := \min_{1 \leqslant i \leqslant \ell} \alpha^i + (c^i)^t x + \frac{1}{2} x^t H_i x, \tag{2.2}$$

donde $\alpha^i \in \mathbb{R}$, $c^i \in \mathbb{R}^n$ y $H_i \in \mathbb{R}^{n \times n}$ es una matriz simétrica definida positiva y p^i representa las variables (α^i, c^i, H_i) como sigue:

$$p^i := \begin{bmatrix} \alpha^i \\ c^i \\ H_i \end{bmatrix},\ i = 1, 2, \ldots, l, \quad p := \begin{bmatrix} p^1 \\ \vdots \\ p^\ell \end{bmatrix}$$

Nuestra aproximación para un sub-estimador de $f(x)$ se obtiene resolviendo el siguiente problema de minimización.

$$\begin{aligned} &\min_{\alpha, c, H} \sum_{j=1}^{m} \left(y^j - \left(\min_{1 \leqslant i \leqslant \ell} \alpha^i + (c^i)^t x^j + \frac{1}{2} (x^j)^t H_i x^j \right) \right) \\ s.a \qquad &\min_{1 \leqslant i \leqslant \ell} (\alpha^i + (c^i)^t x^j + \frac{1}{2} (x^j)^t H_i x^j) \leqslant y_j, \quad j = 1, \ldots, m, \\ &H_i \ \text{ S.D.P.} \end{aligned}$$

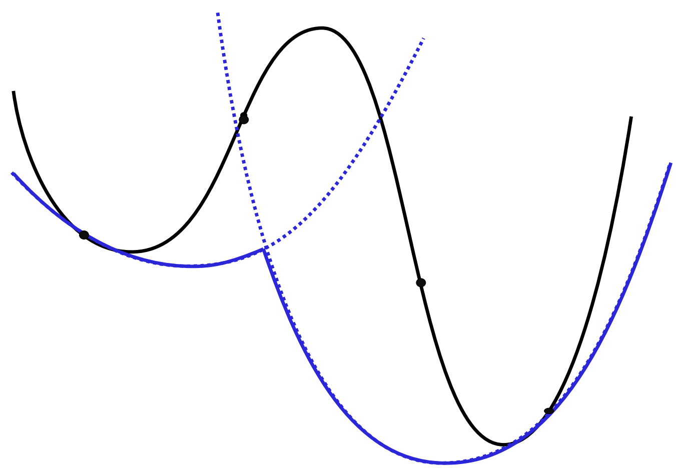

Figura 2.2: Sub-estimador cuadrático a trozos con $\ell = 2$

Capítulo 3

Sub-estimación lineal a trozos

El problema que estamos interesados en resolver es encontrar el mínimo global de una función $f : \mathbb{R}^n \to \mathbb{R}$, dados m valores de la función f

$$f(x^k) = y^k, k = 1, ..., m.$$

Definición 10 (Sub-estimador lineal a trozos)**.** *[4] Una función lineal a trozos viene dada por:*

$$p(x) = \alpha + c^t x + \|Ax + b\|_1, \tag{3.1}$$

donde $\alpha \in \mathbb{R},\ c \in \mathbb{R}^n,\ A \in \mathbb{R}^{\ell \times n}$ y $b \in \mathbb{R}^\ell$, donde ℓ es el número de funciones lineales que generan esta sub-estimación lineal a trozos.
p debe cumplir:

$$p(x^k) \leqslant y^k$$

Así el problema de encontrar un sub-estimador de f está dado por el siguiente problema de minimización

$$
\begin{aligned}
&\min_{\alpha,c,A,b} \sum_{k=1}^{m} \left(y^k - \left(\alpha + c^t x^k + \|Ax^k + b\|_1\right)\right) \\
s.a. \qquad & \alpha + c^t x^k + \|Ax^k + b\|_1 \leqslant y^k, \quad k = 1, \ldots, m.
\end{aligned}
\tag{3.2}
$$

O. L. Mangasarian y R. R. Meyer (1979) [3] probaron que el problema (3.2) tiene los mismos puntos Karush-Kuhn-Tucker que el problema

$$
\begin{aligned}
&\min_{\alpha,c,A,b,s^k} \sum_{k=1}^{m} \left(y^k - \left(\alpha + c^t x^k + e^t|Ax^k + b|\right) + \epsilon e^t s^k\right) \\
s.a \qquad & \alpha + c^t x^k + e^t s^k \leqslant y^k, \; k = 1, \ldots, m \\
& -s^k \leqslant Ax^k + b \leqslant s^k, \; k = 1, \ldots, m,
\end{aligned}
\tag{3.3}
$$

donde $\epsilon \in (0, \bar{\epsilon}]$ para algún $\bar{\epsilon} > 0$ y $s^k \in \mathbb{R}^\ell, \; k = 1, \ldots, m$, y $e \in \mathbb{R}^\ell$ es un vector de unos.

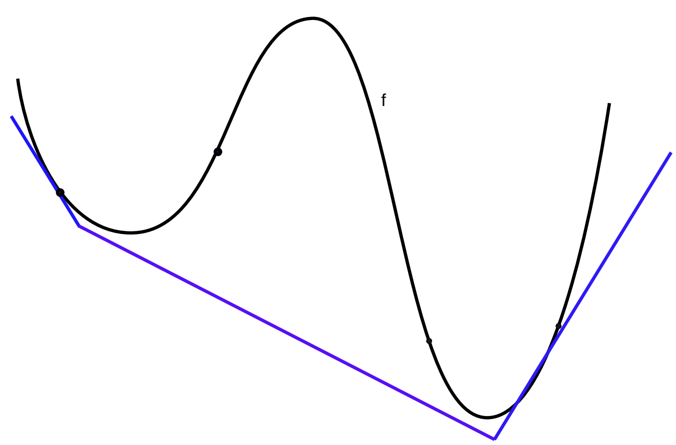

Figura 3.1: Sub-estimador lineal a trozos con $\ell = 3$

Lema 1. *[6] Sea f una función convexa y sea C un conjunto poliédrico convexo no vacío contenido en el dominio de f. Supongamos que C no tiene líneas que van al infinito y que f es acotada sobre C entonces el supremo de f en C es alcanzado en uno de los extremos de C.*

Proposición 3 (Solución de la existencia de un vértice)**.** *[4] Supongamos que la región factible del problema de programación lineal* (3.3) *no tiene líneas rectas que van al infinito en ambas direcciones, entonces existe una solución vértice al problema.*

Demostración. Sabemos que la función objetivo de (3.3) es cóncava y acotada en la región factible, en efecto, del conjunto tenemos

$$-s^k \leq Ax^k + b \leq s^k$$

Por propiedad de valor absoluto y multiplicando por e^t se obtiene,

$$e^t|Ax^k + b| \leq e^t s^k, \quad s^k \geq 0 \tag{3.4}$$

Luego sustituyendo (3.4) en la región factible vemos que esta acotada inferiormente

$$y^k - (\alpha + c^t x^k + e^t|Ax^k + b|) + \epsilon e^t s^k \geq y^k - (\alpha + c^t x^k + e^t s^k) + \epsilon e^t s^k$$

Y esto es equivalente a,

$$y^k - (\alpha + c^t x^k + e^t|Ax^k + b|) + \epsilon e^t s^k \geq y^k - \alpha - c^t x^k - (1 - \epsilon)e^t s^k \tag{3.5}$$

Ahora veamos que la región factible de (3.3) es un conjunto poliédrico no vacío, para esto tomemos, $c = 0$, $A = 0$, $b = 0$, $s^k = 0$ para $k = 1, ..., m$ y $\alpha = \min\limits_{k=1,...,m} y^k$.
tenemos que

$$\alpha + c^t x^k + e^t s^k \leq y^k, \quad k = 1, \ldots, m$$

Luego, sustituimos valores,

$$\alpha + 0x^k + e^t 0 \leq y^k$$

Así,

$$\alpha = \min_{k=1,\dots,m} y^k \leq y^k$$

Y de

$$s^k \leq Ax^k + b \leq s^k$$

Sustituyendo valores,

$$0 \leq 0x^k + 0 \leq 0$$

Por tanto la región factible de (3.3) es un conjunto no vacío.

Por otro lado, el problema matemático (3.3) es equivalente a

$$\begin{aligned} \max_{\alpha,c,A,b,s^k} \; & -\sum_{k=1}^{m} \left(y^k - (\alpha + c^t x^k + e^t |Ax^k + b|) + \epsilon e^t s^k \right) \\ s.a \quad & \alpha + c^t x^k + e^t s^k \leq y^k, \; k = 1, ..., m \\ & -s^k \leq Ax^k + b \leq s^k, \; k = 1, \dots, m \end{aligned} \tag{3.6}$$

la cual cumple que la función objetivo de (3.6) es convexa y es acotada superiormente ya que (3.3) esta acotada inferiormente y la región factible es un conjunto poliédrico convexo. Ahora aplicando el Lema (1) tenemos que existe una solución vértice del problema minimización (3.3). □

Algoritmo de linealización sucesiva (SLA).

Denotemos $\theta(z) = \sum_{k=1}^{m} \left(y^k - \left(\alpha + c^t x^k + e^t |Ax^k + b| \right) + \epsilon e^t s^k \right)$ que es la función objetivo del problema (3.3) y su región factible por

$$Z = \left\{ z \in \mathbb{R}^{1+n+\ell+\ell n+\ell+\ell m} \middle/ \begin{array}{l} \alpha + c^t x^k + e^t s^k \leqslant y^k,\ k = 1, \ldots, m \\ -s^k \leqslant Ax^k + b \leqslant s^k,\ k = 1, \ldots, m, \end{array} \right\}$$

entonces (3.3) puede ser escrita como

$$\min_{z \in Z} \theta(z). \tag{3.7}$$

Para el estudio del algoritmo SLA debemos evaluar el supergradiente de la función objetivo (3.3). Primero representaremos a las variables de optimización de (3.3) por $z \in \mathbb{R}^{1+n+\ell+\ell n+\ell+\ell m}$,

$$z = \begin{bmatrix} \alpha \\ c \\ {A_i}^t \\ b \\ s^k \end{bmatrix} \tag{3.8}$$

donde $\alpha \in \mathbb{R},\ c \in \mathbb{R}^n,\ A \in \mathbb{R}^{\ell \times n}, A_i = \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{pmatrix}$ y $b \in \mathbb{R}^{\ell}$, donde ℓ es el número de funciones lineales que generan la sub-estimación lineal a trozos.

Un supergradiente de $\theta(z)$ con respecto a z es dado por,

$$\partial\theta(z) = \sum_{k=1}^{m} - \begin{bmatrix} 1 \\ x^k \\ \partial|A_i x^k + b_i| x^k, \quad i = 1, 2, ..., \ell \\ \partial|A_i x^k + b_i|, \quad i = 1, 2, ..., \ell \\ -\frac{1}{m}\epsilon e \end{bmatrix} \tag{3.9}$$

donde $e \in \mathbb{R}^{\ell m}$, para $i = 1, 2, ..., \ell$ y $k = 1, 2, ..., m$:

$$\partial|A_i x^k + b_i| = \begin{bmatrix} 1 & si & A_i x^k + b_i & > & 0 \\ \in [-1, 1] & si & A_i x^k + b_i & = & 0 \\ -1 & si & A_i x^k + b_i & < & 0 \end{bmatrix} \tag{3.10}$$

Algoritmo 1. ***Algoritmo de linealización sucesiva (SLA).*** *[4] Comienza con un $z^0 \in \mathbb{R}^{1+n+\ell n+\ell+\ell m}$ al azar. Si tenemos z^i, determinar z^{i+1} como una solución vértice del programa lineal:*

$$\min_{z \in Z} \partial\theta(z^i)(z - z^i) \tag{3.11}$$

es decir: $z^{i+1} \in arg\ vertex\ \min_{z \in Z} \partial\theta(z^i)(z - z^i)$.
Parar cuando $\partial\theta(z^i)(z^{i+1} - z^i) = 0$.

Demostraremos que este algoritmo termina en un número finito de pasos de la siguiente manera.

Lema 2 (Condición necesaria de optimalidad de principio mínimo)**.** *[2] Sea $\bar{z}$ una solución local de $\min_{z \in Z} \theta(z)$ donde Z es un conjunto convexo en $\mathbb{R}^n$ y θ es una función cóncava en $\mathbb{R}^n$. Entonces $\bar{z}$ satisface el siguiente principio mínimo*

$$\partial\theta(\bar{z})(z - \bar{z}) \geqslant 0, \quad \forall z \in Z. \tag{3.12}$$

Demostración. Sea $\bar{z}$ un mínimo local, por definición tenemos que,

$$\theta(\bar{z}) \leqslant \theta(y), \quad \forall y \in B(\bar{z}) \cap Z, \tag{3.13}$$

donde $B(\bar{z})$ es una vecindad de $\bar{z}$. Por otro lado tenemos por hipótesis que Z es convexa, entonces para algún $z \in Z$ que no pertenece a $B(\bar{z})$, se deduce que, existe $\lambda \in (0, 1)$ tal que

$$(1-\lambda)\bar{z} + \lambda z \in B(\bar{z}),$$

Por lo tanto

$$\theta(\bar{z}) \leqslant \theta((1-\lambda)\bar{z} + \lambda z) \tag{3.14}$$

En consecuencia

$$0 \leqslant \theta((1-\lambda)\bar{z} + \lambda z) - \theta(\bar{z})$$

Luego, por definición, si $\partial\theta(\bar{z})$ es un supergradiente de θ en $\bar{z}$ se tiene

$$0 \leqslant \theta((1-\lambda)\bar{z} + \lambda z) - \theta(\bar{z}) \leqslant \partial\theta(\bar{z})(((1-\lambda)\bar{z} + \lambda z) - \bar{z})$$

Así nos queda que

$$0 \leqslant \theta((1-\lambda)\bar{z} + \lambda z) - \theta(\bar{z}) \leqslant \partial\theta(\bar{z})(\lambda z - \lambda\bar{z})$$

Entonces

$$0 \leqslant \lambda\partial\theta(\bar{z})(z - \bar{z})$$

Y como $\lambda > 0$, se concluye que,

$$0 \leqslant \partial\theta(\bar{z})(z - \bar{z})$$

□

Teorema 2 (Terminación finita SLA)**.** *[2] Sea $\theta(z)$ una función cóncava en $\mathbb{R}^n$, acotada inferiormente en Z. El SLA genera una sucesión finita de iterados $\{z^1, z^2, \ldots, z^{\bar{i}}\}$ estrictamente decreciente de los valores de la función objetivo: $\{\theta(z^1) > \theta(z^2) >, \ldots, > \theta(z^{\bar{i}})\}$ tal que $z^{\bar{i}}$ satisface la condición necesaria de optimalidad de principio mínimo:*

$$\partial\theta(z^{\bar{i}})(z - z^{\bar{i}}) \geqslant 0, \quad \forall z \in Z \tag{3.15}$$

Demostración. Por hipótesis tenemos que θ es acotada inferiormente en Z, es decir,

$$-\infty < \inf_{z \in Z} \theta(z) \leqslant \theta(z)$$

sumando $-\theta(z^i)$ y aplicando definición de supergradiente obtenemos lo que sigue,

$$-\infty < \inf_{z \in Z} \theta(z) - \theta(z^i) \leqslant \theta(z) - \theta(z^i) \leqslant \partial\theta(z^i)(z - z^i), \quad \forall z \in Z \tag{3.16}$$

Esto se cumple para $\forall\ z^i \in \mathbb{R}^n$, incluso para algún $z^i \notin Z$, que $\partial\theta(z^i)(z - z^i)$ es acotado inferiormente en Z, esto es por (3.16). Por lo tanto el programa lineal (3.15) tiene una solución y es un vértice, digamos z^{i+1}. Esto se debe a que Z es un conjunto poliédrico convexo (Politopo). Se sigue para $i = 1, 2, \ldots,$ que

$$\forall z \in Z : \partial\theta(z^i)(z - z^i) \geqslant \min_{z \in Z} \partial\theta(z^i)(z - z^i) = \partial\theta(z^i)(z^{i+1} - z^i) \tag{3.17}$$

Veamos que para $z^i \in \mathbb{R}^n$ para $i = 1, 2, \ldots,$ tenemos que $\partial\theta(z^i)(z^{i+1} - z^i) \leqslant 0$, en efecto, si definimos $g(z) = \partial\theta(z^i)(z - z^i)$ y tenemos que

$$g(z^{i+1}) = \min_{z \in Z} g(z)$$

Luego,

$$g(z^{i+1}) \leqslant g(z), \ \forall z \in Z$$

Entonces se tiene,

$$g(z^{i+1}) \leqslant g(z^i) = \partial\theta(z^i)(z^i - z^i) = 0, \quad z^i \in Z$$

Por tanto $g(z^{i+1}) = \partial\theta(z^i)(z^{i+1} - z^i) \leqslant 0$. Ahora se nos presentan dos casos. Caso (a): $\partial\theta(z^i)(z^{i+1} - z^i) < 0$, el algoritmo dado no se para en la iteración i, por la definición de supergradiente y la desigualdad estricta del caso (a) se tiene

$$\theta(z^{i+1}) - \theta(z^i) \leqslant \partial\theta(z^i)(z^{i+1} - z^i) < 0$$

Así,

$$\theta(z^{i+1}) \leqslant \theta(z^i) + \partial\theta(z^i)(z^{i+1} - z^i) < \theta(z^i)$$

Por lo tanto,

$$\theta(z^{i+1}) < \theta(z^i), \quad \text{para} \quad i = 1, 2, \ldots,$$

Caso (b): $\partial\theta(z^i)(z^{i+1} - z^i) = 0$, por el algoritmo tenemos que z^i es solución de $\min_{z \in Z} \theta(z)$ y Z es un conjunto convexo, aplicando el Lema (2) tenemos que

$$\forall z \in Z : \quad \partial\theta(z^i)(z - z^i) \geqslant 0 \tag{3.18}$$

Debido que $z^i \in Z$ establecemos $\bar{i} = i$. El punto $z^{\bar{i}}$ por lo tanto cumple con la condición necesaria de optimalidad de principio mínimo con $z^{\bar{i}} = z^i$ y z^i posiblemente puede ser solución global. Puesto que $\{\theta(z^i)\}_1^{\bar{i}}$ es estrictamente decreciente, θ esta acotada inferiormente en Z y Z es un conjunto poliédrico convexo, entonces la sucesión de los z^i es finita (ya que los z^i son vértices de Z). □

Corolario 1 (Terminación finita de SLA)**.** *[4] Supongamos que la región factible Z del programa lineal a trozos* (3.3) *no tiene líneas rectas que van al infinito en ambas direcciones, la SLA genera una sucesión finita de vertices factibles $\{z^1, z^2, \ldots, z^{\bar{i}}\}$ de Z estrictamente decrecientes de los valores de la función objetivo: $\{\theta(z^1), \theta(z^2), \ldots, \theta(z^{\bar{i}})\}$, tal que $\theta(z^{\bar{i}})$ satisface la condición necesaria de optimalidad de principio mínimo:*

$$\partial\theta(z^{\bar{i}})(z - z^{\bar{i}}) \geqslant 0, \quad \forall z \in Z \tag{3.19}$$

Demostración. Sabemos que $\theta(z)$ es cóncava y acotada inferiormente, luego aplicando el teorema (1) tenemos que la SLA genera una sucesión finita de vertices factibles

$\{z^1, z^2, \ldots, z^{\bar{i}}\}$ de Z estrictamente decrecientes de los valores de la función objetivo: $\{\theta(z^1), \theta(z^2), \ldots, \theta(z^{\bar{i}})\}$, tal que $\theta(z^{\bar{i}})$ satisface el mínimo principio de la condición necesaria de optimalidad:

$$\partial\theta(z^{\bar{i}})(z - z^{\bar{i}}) \geq 0, \quad \forall z \in Z. \tag{3.20}$$

□

Una vez terminado el algoritmo SLA, obtenemos un sub-estimador lineal a trozos de la función original. Dado que nuestra meta es encontrar el mínimo global de la función original, buscamos el mínimo global del sub-estimador para llegar a dicha meta. En otras palabras, una vez obtenido el sub-estimador lineal a trozos debemos resolver el siguiente problema

$$\min_{x \in \mathbb{R}^n} \alpha + c^t x + \|Ax + b\|_1. \tag{3.21}$$

El problema (3.21) es equivalente a resolver el siguiente problema de programación lineal

$$\begin{aligned} \min_{x,y} \quad & \alpha + c^t x + e^t y \\ s.a. \quad & -y \leqslant Ax + b \leqslant y, \\ & a^1 \leqslant x \leqslant a^2, \end{aligned}$$

donde a^1, a^2 son las cotas inferior y superior conocidas del vector x.

Ejemplo 1. *Ahora nosotros generaremos una síntesis de una función cuadrática a trozos no convexa como sigue:*

$$y(x) = \min_{j\in\{1,\ldots,r\}} h_j(x), \tag{3.22}$$

donde $h_j(x),\ j = 1,\ldots,r$ *son funciones arbitrarias cuadráticas estrictamente convexas, tales como:*

$$h_j(x) = \beta^j + d^{j^t} x + \frac{1}{2} x^t(0{,}5I + M^{j^t} M^j)x, \quad , j = 1,\ldots,r, \tag{3.23}$$

donde $\beta^j \in \mathbb{R}$*,* $d^j \in \mathbb{R}^n$ *y* $M^j \in \mathbb{R}^{n\times n}$ *son escogidos.*

Algo interesante de la función cuadrática a trozos generada como ya vimos, es que su solución exacta del mínimo global puede ser computado como sigue:

Proposición 4 (Solución exacta del mínimo global de (3.22) y (3.23))**.** *[4] Un mínimo global exacto de* (3.22) *y* (3.23) *viene dado por:*

$$\min_{j\in\{1,\ldots,r\}} \min_{x\in\mathbb{R}^n} h_j(x). \tag{3.24}$$

Más específicamente,

$$\min_{x\in\mathbb{R}^n} y(x) = \min_{j\in\{1,\ldots,r\}} h_j(x^j), \tag{3.25}$$

donde:

$$x^j = (-0{,}5I + M^{j^t} M^j)^{-1} d^j, \quad j \in \{1,\ldots,r\}. \tag{3.26}$$

Demostración. Dado que $h_j,\ j = 1,...,r$ son funciones cuadráticas estrictamente convexas, podemos asegurar la existencia de mínimos de las funciones $h_j,\ j = 1,\ldots,r$. Denotemos por $\bar{x}^j$ el mínimo global de $h_j,\ j = 1,2,\ldots,r$. Entonces,

$$\min_{x\in\mathbb{R}^n} y(x) = \min_{x\in\mathbb{R}^n} \min_{j\in\{1,\ldots,r\}} h_j(x)$$

Aplicando el $\min\limits_{j\in\{1,\ldots,r\}} h_j(x)$

$$\min_{x\in\mathbb{R}^n}\min_{j\in\{1,\ldots,r\}} h_j(x) = \min_{x\in\mathbb{R}^n}\{h_1(x), h_2(x), \ldots, h_r(x)\}$$

Tenemos entonces

$$\min_{x\in\mathbb{R}^n}\min_{j\in\{1,\ldots,r\}} h_j(x) = h_{\bar{j}}(\bar{x}^j)$$

Y esto es equivalente a

$$\min_{x\in\mathbb{R}^n}\min_{j\in\{1,\ldots,r\}} h_j(x) = \min\{h_1(\bar{x}^1), h_2(\bar{x}^2), \ldots, h_r(\bar{x}^r)\}$$

Así

$$\min_{x\in\mathbb{R}^n}\min_{j\in\{1,\ldots,r\}} h_j(x) = \min\{\min_{x\in\mathbb{R}^n} h_1(x), \min_{x\in\mathbb{R}^n} h_2(x), \ldots, \min_{x\in\mathbb{R}^n} h_r(x)\}$$

En consecuencia se tiene que,

$$\min_{x\in\mathbb{R}^n}\min_{j\in\{1,\ldots,r\}} h_j(x) = \min_{j\in\{1,\ldots,r\}}\min_{x\in\mathbb{R}^n} h_j(x)$$

donde $\bar{j} \in \{1, 2, \ldots, r\}$, es tal que $\bar{j} = argmin\{h_1(\bar{x}^1), h_2(\bar{x}^2), \ldots, h_{1r}(\bar{x}^r)\}$, es decir $h_{\bar{j}}(\bar{x}^{\bar{j}}) = \min\limits_{j\in\{1,\ldots,r\}} h_j(x^j)$ Por lo tanto

$$\min_{x\in\mathbb{R}^n}\min_{j\in\{1,\ldots,r\}} h_j(x) = \min_{j\in\{1,\ldots,r\}}\min_{x\in\mathbb{R}^n} h_j(x) \tag{3.27}$$

Veamos donde $h_j(x)$ toma su valor mínimo, para esto veamos para que puntos $x^{\bar{j}}$ la derivada se anula o no existe.

Derivando $h_{\bar{j}}(x^{\bar{j}})$ tenemos,

$$h_{\bar{j}}(x^{\bar{j}})' = d^j + x^{\bar{j}}(0{,}5I + M^{j^t} M^j), \quad , j = 1, \ldots, r, \tag{3.28}$$

Notamos que el único punto que anula a (3.28) es:

$$x^{\bar{j}} = -(0{,}5I + M^{j^t} M^j)^{-1} d^j, \quad , j = 1, \ldots, r, \tag{3.29}$$

□

Capítulo 4

Conclusión

Después de estudiar el trabajo hecho por Mangasarian, Rosen y Thompson (2004), llegamos a las siguientes conclusiones:

- Se ha encontrado un algoritmo que halla el mínimo de una función basada en sub-estimación lineal a trozos. Dicho algoritmo consiste en resolver en cada iteración un problema de programación lineal.
- Otra estrategia que estudiamos consiste en la sub-estimación cuadrática, donde se tiene la dificultad de encontrar una matriz definida positiva, problema que origina usar la sub-estimación lineal a trozos mencionada anteriormente.
- Seria un estudio interesante (en el futuro) aplicar un sub-estimador lineal a trozos que fuese convexa a trozos para que solventara el de la mala estimación que

produce el sub-estimador cuadrático.

Bibliografía

[1] R. A. Horn and C. R Johnson. *Matrix Analysis*. Cambridges University Press, Cambridge, UK, 1999.

[2] O. Mangasarian. Solution of general linear complementarity problems via nondifferentiable concave minimization. *Acta Mathematica Vietnamita*, 22(1):199–205, 1997.

[3] O. Mangasarian and R. R. Meyer. Nonlinear perturbation of linear programs. *SIAM Journal on Control and Optimization*, 17(6):745–752, 1979.

[4] O. L. Mangasarian, J.B Rosen, and M.E. Thompson. Global optimization via piecewise-linear underestimation. *Journal of Global Optimization*, 32:1–9, 2004.

[5] O. L. Mangasarian, J.B Rosen, and M.E. Thompson. Nonconvex piecewise-quadratic underestimation for global minimization. *Journal of Global Optimization*, 34(4):475–488, 2006.

[6] R. T. Rockafellar. *Convex analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[7] J. B. Rosen and R. F. Marcia. Convex quadratic approximation. *Computational Optimization*, 28:173–187, 2004.

[8] S. Wright and Nocedal, J. *Numerical Optimization*. Springer series in operation research. Springer, New York, 1999.