Optimización no diferenciable

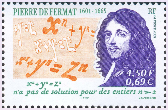
Emilio Carrizosa

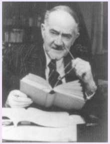
Dpto. de Estadística e Investigación Operativa
Universidad de Sevilla

Diciembre de 2010

- Introducción a la optimización no diferenciable.
- Cálculo subdiferencial.
- Condiciones de optimalidad.
- Algoritmos.
- Problemas de optimización de distancias.

El problema de Fermat-Weber





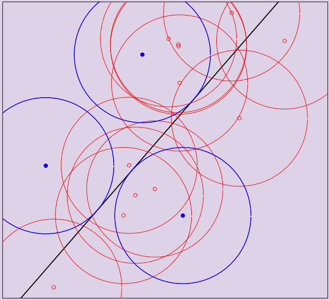
Sea $A \subset \mathbb{R}^n$, finito (usuarios). Para cada $a \in A$, sea $\omega_a > 0$ la demanda de a.

$$\min_{x \in \mathbb{R}} f(x) := \sum_{a \in A} \omega_a \|x - a\|$$

Regresión ortogonal (ℓ_∞)

Sea $A \subset \mathbb{R}^n$. Se busca el hiperplano H que minimice el máximo de las distancias euclídeas de los puntos de A a H.

$$\min_{(u,b)\in\mathbb{R}^n\times\mathbb{R},\ u\neq 0} \ \max_{a\in A} \frac{|u^\top a + b|}{\|u\|}$$



Referencias básicas I

J.-B. Hiriart-Urruty y C. Lemaréchal.
Convex analysis and minimization algorithms I.
Springer, 1993.

H. Tuy.
Convex analysis and global optimization.
Kluwer, 1998.

Convexidad

Conjunto convexo

$S \subset \mathbb{R}^n$ se dice convexo si, para cualesquiera $x, y \in S$, el segmento de extremos x, y está en S :

$$x, y \in S \Rightarrow (1-\lambda)x + \lambda y \in S \quad \forall \lambda \in [0,1].$$

Operaciones que preservan la convexidad

- Dado $C \subset \mathbb{R}^n$: convexo, son convexos:
  - $int(C), cl(C)$.
  - $A \cdot C + \mu := \{Ac + \mu : c \in C\}$, con $A \in \mathbb{R}^{m \times n}$, $\mu \in \mathbb{R}^m$.
- Si $\{C_i\}_{i \in I}$ es una colección de convexos en $\mathbb{R}^n$, entonces son convexos los siguientes conjuntos:
  - $\cap_{i \in I} C_i$.
  - $\prod_{i \in I} C_i$.
  - $\sum_{i \in I} C_i$.

Cierre convexo

Dado $X \subset \mathbb{R}^n$, definimos su cierre convexo $conv(X)$ como el menor convexo que contiene a X : $conv(X) = \bigcap_{C \supset X,\, C:\text{convexo}} C$

Ejemplo

Para $E \subset \mathbb{R}^n$: finito no vacío,
$conv(E) = \{\sum_{e \in E} \lambda_e e : \ \lambda_e \geq 0 \, \forall e \in E, \sum_{e \in E} \lambda_e = 1\}$

Teorema de Carathéodory

Sea $X \subset \mathbb{R}^n$. Para cada $x \in conv(X)$, $\exists X^* \subset X$, $|X^*| \leq n+1$ tal que $x \in conv(X^*)$.

Conos

$K \subset \mathbb{R}^n$, $K \neq \emptyset$ se dice cono si $K = \mathbb{R}_+ \cdot K$:

$$x \in K \Rightarrow \lambda x \in K \qquad \forall \lambda \in \mathbb{R}_+.$$

Todo cono contiene al origen.

Conos convexos

Un cono $K \subset \mathbb{R}^n$ es convexo sii $K + K = K$.

Ejercicio 1

Cono normal

Dados $X \subset \mathbb{R}^n$, $x \in X$, el cono normal a X en x es el cono convexo cerrado $N_X(x)$,

$$N_X(x) = \left\{ p \in \mathbb{R}^n : p^\top (y - x) \leq 0 \quad \forall y \in X \right\}.$$

Ejemplo

Para $X = \{x \in \mathbb{R}^n : \ Ax \leq b\}$,

$$N_X(x) = \left\{ A^\top u : \ u^\top (Ax - b) = 0, \, u \geq 0 \right\}.$$

Ejercicio 2

Separación

Dados $C \subset \mathbb{R}^n$, $C \neq \emptyset$, convexo, cerrado, y $x \in \mathbb{R}^n \setminus C$, existe $u \in \mathbb{R}^n$ tal que

$$u^\top y < u^\top x \qquad \forall y \in C.$$

Teorema minimax

Para convexos cerrados $X, Y \subset \mathbb{R}^n$, con Y : compacto,

$$\min_{x \in X} \max_{y \in Y} x^\top y = \max_{y \in Y} \min_{x \in X} x^\top y$$

Polar

Dado $X \subset \mathbb{R}^n$, se define su polar X° como

$$X^\circ = \left\{ p : \; p^\top x \leq 1 \, \forall x \in X \right\}.$$

Ejemplos

- Para $x_0 \in \mathbb{R}^n$, $x_0 \neq 0$, $\{x_0\}^\circ$ es un semiespacio cerrado.
- Para $X = \{x : \|x\|_2 \leq c\}$, $X^\circ = \{x : \|x\|_2 \leq \frac{1}{c}\}$.

Ejercicio 3

Polar

Propiedades

- Dado $X \subset \mathbb{R}^n$, X° es un conjunto convexo, cerrado, con $0 \in X^\circ$.
- Dados $X \subset Y \subset \mathbb{R}^n$, $X^\circ \supset Y^\circ$.
- Sea X convexo, cerrado con $0 \in X$. Entonces
  - $(X^\circ)^\circ = X$.
  - $0 \in int(X)$ sii X° : acotado.

Ejercicio 4

Convexidad

Función convexa

Epigrafo

El epigrafo $epi(f)$ de una función $f : S \longrightarrow R$ se define como

$$epi(f) = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : x \in S, t \geq f(x)\}.$$

Sea $S \subset \mathbb{R}^n$: convexo. $f : S \longrightarrow \mathbb{R}$ se dice convexa si $epi(f)$ es convexo. Equivalentemente, f es convexa sii

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y).$$

Funciones $\mathcal{C}^2$ convexas

Cualquier función $\mathcal{C}^2$ en $\mathbb{R}^n$ con matriz hessiana semidefinida positiva en todo $\mathbb{R}^n$. Por ejemplo

- $f(x) = a^\top x + b$.
- $f(x) = x^\top Qx + b^\top x + c$, con Q : semidefinida positiva (como caso particular, la identidad).

Álgebra de funciones convexas

- Dada f : convexa sobre el convexo $S \subset \mathbb{R}^n$,
  - $\lambda f + \mu$, con $\lambda \geq 0, \mu \in \mathbb{R}$: convexa
  - $f(Ax + b)$: convexa, con $A \in R^{m \times n}, b \in \mathbb{R}^m$
- Dadas f_1, f_2, convexas sobre el convexo $S \subset \mathbb{R}^n$, son funciones convexas:
  - $f_1 + f_2$.
  - $\max\{f_1, f_2\}$.

Sean $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, convexa en el convexo S, y $g : \mathbb{R} \longrightarrow \mathbb{R}$: convexa no decreciente. Entonces $g \circ f$: convexa en S.

Ejercicio 5

Calibradores

Sea $B \subset \mathbb{R}^n$, convexo, compacto, con $0 \in int(B)$. Definimos el calibrador γ_B con bola unidad B como

$$\gamma_B(x) = \min\{t \geq 0 : \; x \in tB\}.$$

Propiedades

- $\gamma_B(x) \geq 0 \, \forall x \in \mathbb{R}^n$, $\gamma_B(x) = 0$ sii $x = 0$.
- $\gamma_B(\lambda x) = \lambda \gamma_B(x) \, \forall x \in \mathbb{R}^n, \lambda \in \mathbb{R}_+$.
- $\gamma_B(x + y) \leq \gamma_B(x) + \gamma_B(y) \, \forall x, y \in \mathbb{R}^n$.
- γ_B es convexa en $\mathbb{R}^n$.

Ejercicio 6

Calibradores

γ :

$\gamma(x) \geq 0 \, \forall x \in \mathbb{R}^n$
$\gamma(x) = 0$ sii $x = 0$
$\gamma(\lambda x) = \lambda \gamma(x) \, \forall x \in \mathbb{R}^n, \lambda \in \mathbb{R}_+$
$\gamma(x + y) \leq \gamma(x) + \gamma(y) \, \forall x, y \in \mathbb{R}^n$

convexos compactos, con 0 en su interior

γ

$\{x : \gamma(x) \leq 1\}$

Calibradores simétricos: normas

B : simétrico respecto a 0 sii γ_B : norma en $\mathbb{R}^n$.

γ :

$\gamma(x) \geq 0 \, \forall x \in \mathbb{R}^n$
$\gamma(x) = 0$ sii $x = 0$
$\gamma(\lambda x) = |\lambda|\gamma(x) \, \forall x \in \mathbb{R}^n$
$\gamma(x+y) \leq \gamma(x) + \gamma(y) \, \forall x, y \in \mathbb{R}^n$

convexos compactos, simétricos respecto a 0 con 0 en su interior

γ

$\{x : \gamma(x) \leq 1\}$

Ejemplos de normas

Norma de componentes ordenadas

Sea $c : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ tal que, para cada $x \in \mathbb{R}^n$, $c(x)$ es la reordenación de $(|x_1|, \ldots, |x_n|)$ con $c_1(x) \geq c_2(x) \geq \ldots \geq c_n(x)$. La función

$$x \longmapsto \gamma(x) = \sum_j \omega_j c_j(x)$$

es una norma siempre que $\omega_1 \geq \ldots \geq \omega_n \geq 0$, $\omega_1 > 0$.

Casos particulares:

ω_1	ω_2	$\ldots$	ω_n	γ
1	0	$\ldots$	0	ℓ_∞
1	1	$\ldots$	1	ℓ_1
λ	μ	$\ldots$	μ	$\ell_{1,\infty}$

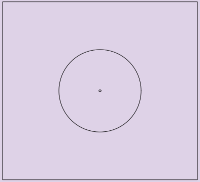
Calibrador sesgado

Calibrador sesgado (skewed gauge)

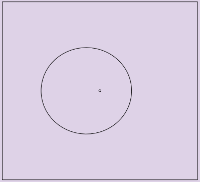
Dado $p \in \mathbb{R}^n$, $\|p\|_2 < 1$, definamos γ como

$$\gamma(x) = \|x\|_2 + p^\top x.$$

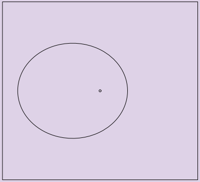
$p = (0, 0)$



$p = (0.3, 0)$



$p = (0.5, 0)$



Calibrador sesgado I

M. Cera, J.A. Mesa, F.A. Ortega, F. Plastria
"Locating a central hunter on the plane"
JOTA 136 (2008) 155–166.

P. Chaudhuri
"On a geometric notion on quantiles for multivariate data"
Journal of the American Statistical Association 91 (1996) 862–872.

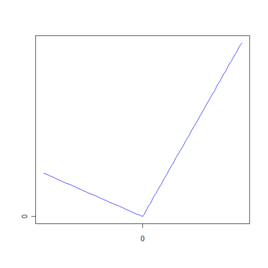
F. Plastria
"On destination optimality in asymmetric distance Fermat-Weber problems"
Annals of Operations Research 40 (1992) 355–369.

Calibrador sesgado

Calibrador sesgado en $\mathbb{R}$:

Dado $p \in \mathbb{R}, |p| < 1$, definamos γ como

$$\gamma(x) = |x| + px = \begin{cases} (1+p)x, & \text{si } x > 0, \\ -(1-p)x, & \text{si } x \leq 0 \end{cases}$$

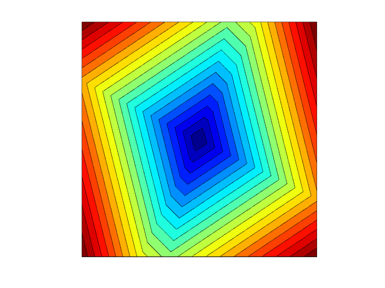


Calibrador poliédrico

Sea $E \subset \mathbb{R}^n$: finito, con $0 \in conv(E)$, $afin(E) = \mathbb{R}^n$ (o, equivalentemente, $0 \in int(conv(E))$. Entonces el calibrador γ,

$$\gamma(x) = \max_{e \in E} e^\top x = \max_{e \in conv(E)} e^\top x$$

se llama calibrador poliédrico.



Calibrador poliédrico

Sea $E \subset \mathbb{R}^n$: finito, y sea $\gamma(x) = \max_{e \in E} e^\top x$. Entonces γ es un calibrador poliédrico sii se satisface

- el siguiente problema de optimización lineal tiene valor óptimo 0

$$\begin{array}{ll} \max & \sum_{e \in E} \lambda_e \\ \text{s.a.} & 0 = \sum_{e \in E} \lambda_e e \\ & \lambda_e \geq 0 \qquad\qquad \forall e \in E. \end{array}$$

- Para cada $i = 1, 2, \ldots, n$, $\delta \in \{-1, 1\}$, el siguiente problema de optimización lineal tiene valor óptimo estrictamente positivo

$$\begin{array}{ll} \min & z \\ \text{s.a.} & z \geq e^\top x \quad \forall e \in E \\ & x_i = \delta \\ & x \in \mathbb{R}^n. \end{array}$$

Dual de un calibrador

Calibrador dual

Sea γ_B un calibrador con bola unidad B. Se define su dual γ_B° como

$$\gamma_B^\circ(x) = \max_{p \in B} p^\top x$$

Dual de un calibrador poliédrico

$\gamma(x) = \max_{e \in E} e^\top x$:

$$\begin{aligned}
\gamma^\circ(x) &= \max_{\gamma(u) \leq 1} u^\top x \\
&= \max\{u^\top x : \; e^\top u \leq 1 \, \forall e \in E\} \\
&= \min\{\sum_{e \in E} \lambda_e : \; \sum_{e \in E} \lambda_e e = x, \; \lambda_e \geq 0 \, \forall e \in E\}
\end{aligned}$$

Desigualdad de Cauchy-Schwarz

γ_B° es un calibrador, con bola unidad B° :

$$\gamma_B^\circ = \gamma_{B^\circ}$$

Ejercicio 8

$$((\gamma_B)^\circ)^\circ = \gamma_{B^{\circ\circ}} = \gamma_B$$

$$\gamma_B(x) = ((\gamma_B)^\circ)^\circ(x) = \max_{u \in B^\circ} u^\top x = \max_{u \neq 0} \frac{u^\top x}{\gamma_B^\circ(u)}$$

Cauchy-Schwarz para la norma euclídea ...

$$\|x\|_2 \|u\|_2 \geq x^\top u \ \forall x, u$$

Funciones convexas, continuidad y diferenciabilidad

Sea f : convexa en el convexo abierto $S \subset \mathbb{R}^n$.

- f es continua en S.
- El conjunto de puntos de no diferenciabilidad de f tiene medida de Lebesgue 0.

Ejercicio 9

Subdiferenciabilidad

Sea f : convexa en el convexo $S \subset \mathbb{R}^n$, y $x_0 \in S$: punto de diferenciabilidad de f. Entonces

$$f(x) \geq f(x_0) + \nabla f(x_0)^\top (x - x_0) \qquad \forall x \in S.$$

Sea f : convexa en el convexo $S \subset \mathbb{R}^n$. Dado $x_0 \in S$, el subdiferencial $\partial f(x_0)$ de f en x_0 se define como

$$\partial f(x_0) = \left\{ p : \ f(x) \geq f(x_0) + p^\top (x - x_0) \, \forall x \right\}.$$

$p \in \partial f(x_0)$ sii la función afín $x \longmapsto f(x_0) + p^\top (x - x_0)$ es una minorante de f.

$\partial f(x)$ es un conjunto convexo, cerrado, no vacío en cada $x \in int(S)$

Subdiferenciales y derivadas direccionales

Sea f : convexa en el convexo abierto $S \subset \mathbb{R}^n$. Para $x_0 \in S$, $d \in \mathbb{R}^n$, $d \neq 0$, siempre existe $\nabla f(x_0; d)$, la derivada direccional de f en x_0 en la dirección d, y satisface

$$\nabla f(x_0; d) = \max_{p \in \partial f(x_0)} p^\top d$$

Cálculo subdiferencial

Si f es convexa diferenciable en x_0, entonces

$$\partial f(x_0) = \{\nabla f(x_0)\}$$

Sean $f_1, f_2, \ldots, f_k$: convexas en el convexo abierto $S \subset \mathbb{R}^n$. Para $x \in S$,

- $\partial\left(f_1(x) + f_2(x) + \ldots + f_k(x)\right) = \partial f_1(x) + \partial f_2(x) + \ldots + \partial f_k(x)$.
- $\partial \max_{1 \leq i \leq k} f_i(x) = conv(\bigcup_{i \in A(x)} \partial f_i(x)$, con $A(x) = \{j : f_j(x) = \max_{1 \leq i \leq k} f_i(x)\}$

Ejemplo

Sean $E \subset \mathbb{R}^n$, finito, $\{a_e\}_{e \in E}$, $f(x) = \max_{e \in E} e^\top (x - a_e)$.

$$\partial f(x) = conv\left(\left\{e : \ f(x) = e^\top (x - a_e)\right\}\right)$$

Sea $f(t) = |t - a| = \max\{t - a, a - t\}$. Entonces

$$\partial f(t) = \begin{cases} -1, & \text{si } t < a \\ conv(\{-1, 1\}) = [-1, 1], & \text{si } t = a \\ 1, & \text{si } t > a \end{cases}$$

Composición con funciones afines

Sea f convexa en $\mathbb{R}^m$, y sean $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Sea g la función convexa en $\mathbb{R}^n$ $g(x) = f(Ax + b)$. Entonces

$$\partial g(x) = A^\top \partial f(y)|_{y=Ax+b}$$

Subdiferencial de un calibrador

$$\partial\gamma(x) = \left\{ u \in B^\circ : \; u^\top x = \max_{u \in B^\circ} u^\top x \right\}$$

- $\partial\gamma(0) = B^\circ$
- For $x \neq 0$, $\partial\gamma(x)$ es una cara expuesta de of B° :

$$p \in \partial\gamma(x) \text{ sii } \gamma^\circ(p) = 1, \; p^\top x = \gamma(x).$$

$$\gamma(x) = \max_{p \in \partial\gamma(x)} p^\top x$$

Ejemplos

- $\partial \|x\|_2 = \begin{cases} \frac{1}{\|x\|_2} x, & \text{si } x \neq 0 \\ \{y : \|y\|_2 \leq 1\}, & \text{si } x = 0 \end{cases}$
- Sean $p \in \mathbb{R}^n$, $\|p\|_2 < 1$, $\gamma(x) = \|x\|_2 + p^\top x$. Entonces

$$\partial \gamma(x) = \begin{cases} \frac{1}{\|x\|_2} x + p, & \text{si } x \neq 0 \\ \{y + p : \|y\|_2 \leq 1\}, & \text{si } x = 0 \end{cases}$$

Ejercicio 10

Subdiferenciales de normas monótonas

Sea γ : norma monótona en $\mathbb{R}^n_+$. Dado $x \in R^n_+$,

1. Si $x \in \mathbb{R}^n_{++}$, entonces $\partial\gamma(x) \subset \mathbb{R}^n_+$.
2. Si $x \in \mathbb{R}^n_+$, entonces $\partial\gamma(x) \cap \mathbb{R}^n_+ \neq \emptyset$.
3. Si $p \in \mathbb{R}^n_+$, entonces $\gamma^\circ(p) = \max_{\gamma(x)\leq 1, x \in \mathbb{R}^n_+} p^\top x$.
4. γ° es monótona en $\mathbb{R}^n_+$.

Ejercicio 11

Dual de un calibrador compuesto

Sean $\gamma_1, \ldots, \gamma_k$: calibradores en $\mathbb{R}^{n_1}, \ldots, \mathbb{R}^{n_k}$ y τ : calibrador en $\mathbb{R}^k$, no decreciente en $\mathbb{R}^k_+$. El calibrador compuesto, $\nu(s_1, s_2, \ldots, s_k) = \tau(\gamma_1(s_1), \ldots, \gamma_k(s_k))$, tiene como dual

$$\nu^\circ(s_1, \ldots, s_k) = \tau^\circ(\gamma_1^\circ(s_1), \ldots, \gamma_k^\circ(s_k))$$

El conjunto de soluciones óptimas de un problema convexo

Sea f : convexa sobre el convexo cerrado $S \subset \mathbb{R}^n$. El conjunto de soluciones óptimas del problema $\min_{x \in S} f(x)$ es un convexo cerrado (que puede ser vacío).

Ejercicio 12

Los métodos de optimización de funciones diferenciables pueden no converger (o incluso no estar definidos) en el caso de optimización no diferenciable.

¿Podemos minimizar en $\mathbb{R}$ la función $|x|$ usando el método de Newton?

En cambio, en muchas ocasiones, la no-diferenciabilidad de las funciones implicadas no impide la convergencia empírica a la solución óptima.

Muchos problemas de optimización que involucran funciones convexas no diferenciables pueden reescribirse como problemas de optimización de funciones convexas diferenciables (en ocasiones aumentando considerablemente la dimensión del problema).

Regresión ℓ_1

G. Appa y C. Smith. "On L1 and Chebyshev estimation." *Mathematical Programming* 5 (1973) 73–87.

$$\min_{\beta \in \mathbb{R}^n} \sum_{j=1}^{p} |y_j - \beta^\top x_j|$$

$$\begin{array}{lll}
\min & \sum_{j=1}^{p} \left(d_j^+ + d_j^-\right) & \\
\text{s.a.} & y_j - \beta^\top x_j = d_j^+ - d_j^- & \forall j = 1, 2, \ldots, p \\
& d_j^+, d_j^- \geq 0 & \forall j = 1, 2, \ldots, p \\
& \beta \in \mathbb{R}^n. &
\end{array}$$

problema lineal !!!

El problema del 1-centro en $\mathbb{R}^n$

$$\min_{x\in\mathbb{R}^n} \max_{a\in A} \|x-a\|_2{}^2$$

$$\begin{array}{ll} \min & z \\ \text{s.a.} & (x-a)^\top(x-a) \leq z \quad \forall a \in A \\ & x \in \mathbb{R}^n, z \geq 0. \end{array}$$

Condiciones de optimalidad. Problemas sin restricciones

$$\min_{x \in \mathbb{R}^n} f(x)$$

x^* : óptimo sii $0 \in \partial f(x^*)$

Caso diferenciable: $\partial f(x^*) = \nabla f(x^*)$

x^* : óptimo sii $0 \in \nabla f(x^*)$

Ejercicio 13

Condiciones de optimalidad. Problemas con restricciones

$$\min_{x \in S} f(x)$$

x^* : óptimo sii $0 \in \partial f(x^*) + N_S(x^*)$

Aplicación. El problema de Fermat-Weber en la recta

$$\min_{x\in\mathbb{R}} f(x) := \sum_{a\in A} \omega_a |x-a|$$

sin pérdida de generalidad, $\sum_{a\in A} \omega_a = 1$

Condiciones de optimalidad:

$$\begin{array}{rcl} 0 & \in & \partial f(x) \\ & = & \sum_{a\in A} \omega_a \partial |x-a| \\ & = & \sum_{a\in A, a<x} \omega_a - \sum_{a\in A, a>x} \omega_a + \sum_{a\in A, a=x} \omega_a [-1,1]. \end{array}$$

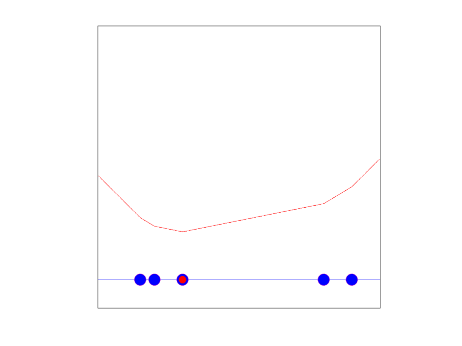
para $x = \{a_0\} \in A$:

$$\begin{array}{rcl} \sum_{a\in A, a\geq a_0} \omega_a & \geq & \frac{1}{2} \\ \sum_{a\in A, a\leq a_0} \omega_a & \geq & \frac{1}{2} \end{array}$$

Aplicación. El problema de Fermat-Weber en la recta

$$\min_{x \in \mathbb{R}} f(x) := \sum_{a \in A} \omega_a |x - a|$$

Solución óptima: mediana(s) de A.



Aplicación. El problema de Fermat-Weber en la recta (calibrador sesgado)

$$\min_{x\in\mathbb{R}} f(x) := \sum_{a\in A} \omega_a \left(|x-a| + p(x-a)\right)$$

sin pérdida de generalidad, $\sum_{a\in A} \omega_a = 1$

para $x = \{a_0\} \in A$:

$$\begin{array}{rcl} \sum_{a\in A, a\leq a_0} \omega_a & \geq & \frac{1-p}{2} \\ \sum_{a\in A, a\geq a_0} \omega_a & \geq & \frac{1+p}{2}. \end{array}$$

Aplicación. El problema de Fermat-Weber en un segmento

$$\min_{x \in S} f(x) := \sum_{a \in A} \omega_a |x - a|$$

$$S = [s_1, s_2],\ s_1 < s_2,\ s_1, s_2 \notin A.\ \sum_{a \in A} \omega_a = 1$$

- $N_S(x) = \begin{cases} \mathbb{R}_-, & \text{si } x = s_1 \\ \{0\}, & \text{si } s_1 < x < s_2 \\ \mathbb{R}_+, & \text{si } x = s_2. \end{cases}$
- Condiciones de optimalidad en $x \in (s_1, s_2)$: x : mediana de A
- Condiciones en x_{s_1} : $\sum_{a > x} \omega_a \leq \frac{1}{2}$
- Condiciones en x_{s_2} : $\sum_{a < x} \omega_a \leq \frac{1}{2}$

El problema de Fermat-Weber euclídeo en $\mathbb{R}^n$

$$\min_{x \in \mathbb{R}^n} f(x) := \sum_{a \in A} \omega_a \|x - a\|_2$$

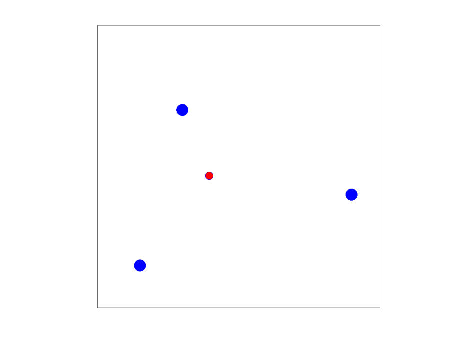
Algoritmo de Weiszfeld. E. Weiszfeld. "Sur le point pour lequel la somme des distances de n points donnés est minimum". *Tohoku Math Journal* 43 (1937) 335.386.

1. Comprobar la condición de optimalidad en todos los puntos de A. Si se verifica en algún punto, PARAR
2. Partiendo de x_0, aplicar el esquema de punto fijo

$$x = \sum_{a \in A} \frac{\omega_a / \|x - a\|_2}{\sum_{b \in A} \omega_b / \|x - b\|_2} a$$

El problema de Fermat

$$\min_{x \in \mathbb{R}^2} \|x-a\|_2 + \|x-b\|_2 + \|x-c\|_2$$



Ejercicio 14

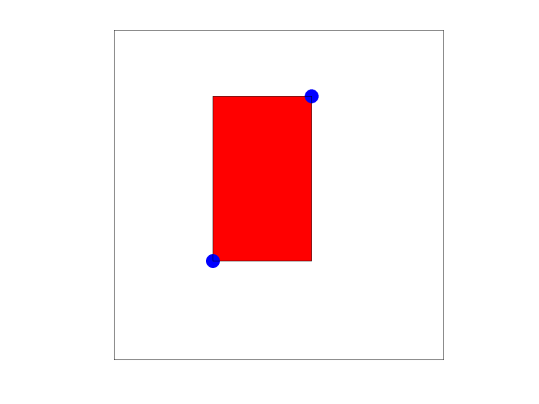
El problema de Fermat-Weber problem en $\mathbb{R}^n$ (ℓ_1)

$$\min_{x \in \mathbb{R}^n} \sum_{a \in A} \omega_a \|x - a\|_1$$

$$\sum_{j=1}^{n} \min_{x_j \in \mathbb{R}} \sum_{a \in A} \omega_a |x_j - a_j|$$

El problema de Fermat-Weber en $\mathbb{R}^n$.

$\gamma : \ell_1, \ldots$



El problema del centro en $\mathbb{R}^n$

Sea $A \subset \mathbb{R}^n$, finito. Sea γ : norma

$$\min_{x \in \mathbb{R}^n} f(x) := \max_{a \in A} \gamma(x - a)$$

Condición de optimalidad en x^* :

$$0 \in conv\left(\{\partial\gamma(x^* - a) : \ f(x^*) = \gamma(x^* - a)\}\right)$$

Carathéodory:

Si x^* : óptima, entonces existe $A^* \subset A$, con cardinal $|A^*| \leq n + 1$ tal que x^* : óptima para

$$\min_{x \in \mathbb{R}^n} f(x) := \max_{a \in A^*} \gamma(x - a)$$

El problema del centro (euclídeo) en $\mathbb{R}^2$

Algoritmo de Elzinga-Hearn

J. Elzinga y D.W. Hearn.
"Geometrical solutions for some minimax location problems".
Transportation Science 6 (1972) 379–394.

D.W. Hearn y J. Vijay.
"Efficient algorithms for the (weighted) minimum circle problem".
Operations Research 30 (1982) 777–795.

Algoritmo de Elzinga-Hearn

1. Seleccionar $A_0 \subset A$, de cardinal $|A_0| = 3$. Hacer $k := 1$
2. Encontrar c_k, el centro de mínimo radio cubriendo A_k, y anotar en r_k su radio
3. Hallar $b_k \in A$, el punto más alejado en A de c_k. Si $\|b_k - c_k\| \leq r_k$, `PARAR:` c_k es óptimo.
4. Para cada subconjunto A_k^* de cardinal 3 de $A_k \cup \{b_k\}$, $b_k \in A_k^*$, encontrar el centro del disco de mínimo radio que cubre a A_k^*, y definir A_{k+1} como el A_k^* que da el mayor de los tres radios.
5. hacer $k := k + 1$ y volver al Paso 2.

Algoritmos. Cuesta de mayor pendiente

1. Seleccionar $x_1 \in \mathbb{R}^n$, hacer $k = 1$.
2. Si $0 \in \partial f(x_k)$, `FIN`
3. Hallar d_k resolviendo $\min_{\|d\|=1} \max_{s \in \partial f(x_k)} s^\top d$
4. Hallar un paso $t_k > 0$ y $x_{k+1} := x_k + t_k d_k$ tal que $f(x_{k+1}) < f(x_k)$
5. Hacer $k := k + 1$ y volver al Paso 2.

...que no necesariamente converge

Algoritmos convergentes

- Métodos de punto interior
- Bundle methods
- ...

Algoritmos convergentes

J.-B. Hiriart-Urruty y C. Lemaréchal.
Convex analysis and minimization algorithms I.
Springer, 1993.

Yu. Nesterov y A. Nemirovskii.
Polinomial-time interior-point methods in convex programming.
SIAM, 1994.

G. Xue y Y. Ye.
"An efficiente algorithm for minimizing a sum of p-norms".
SIAM Journal on Optimization 10 (1997) 551..579.

1. Prueba que cono $K \subset \mathbb{R}^n$ es convexo sii $K + K = K$.
2. Calcula el cono normal en un punto $x \in X$ de los siguientes conjuntos X:
    1. $X = \{x \in \mathbb{R}^n : c^\top x = c_0\}$
    2. $X = \{x : x^\top x \leq 1\}$
3. Calcula el polar de los siguientes conjuntos:
    1. $X = \{x \in \mathbb{R}^n : \ x_i \geq 0 \, \forall i\}$.
    2. $X = \{x \in \mathbb{R}^n : \ x_i \geq 1 \, \forall i\}$.
4. Demuestra con ejemplos que es necesario que X sea convexo, cerrado y que $0 \in X$ para que $(X^\circ)^\circ = X$.
5. Sea $f : S \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, convexa en el convexo S, y $g : \mathbb{R} \longrightarrow \mathbb{R}$: cóncava. Estudia la convexidad en S de la función $g \circ f$ suponiendo que g es
    1. no creciente
    2. no decreciente

6. Prueba que si γ es un calibrador en $\mathbb{R}^n$, entonces su epigrafo es un cono convexo. ¿Es cierto el recíproco?
7. Escribe en la forma $\max_{e \in E} x^\top e$ las normas ℓ_1 y ℓ_∞
8. Sea A una matriz $\mathbb{R}^{n \times n}$ invertible. Halla la norma dual de $\gamma(x) = \sqrt{x^\top A x}$.
9. Construye un convexo $S \subset \mathbb{R}^n$ y una función f, convexa sobre S pero no continua en todos los puntos de S.
10. Calcula el subdiferencial $\partial \|x\|$ de la norma en $\mathbb{R}^n$ $\|\cdot\|$ en x para
    1. $\|\cdot\| : \ell_\infty$
    2. $\|\cdot\| : \ell_1$
11. Construye una norma que no sea monótona en $\mathbb{R}^n_+$.

12. Encuentra un ejemplo (o demuestra que no existe) de problema de minimización con función objetivo convexa y región factible convexa cerrada
    1. sin solución óptima
    2. con infinitas soluciones óptimas
    3. con exactamente dos soluciones óptimas
13. Sea f : convexa. Prueba que x^* : es solución óptima del problema $\min_{x\in\mathbb{R}^n} f(x)$ sii $0 \in \partial f(x^*)$
14. Resuelve analíticamente el problema de Fermat (minimizar la suma de distancias ℓ_2 a 3 puntos fijos del plano con la norma ℓ_2).