Concavidad y Optimización

Abelardo Jordán Liza

Maestría en Matemáticas Aplicadas
PUCP

Lima, Setiembre 09, 2023

Definición

Sea $\emptyset \neq C \subset E$ *un conjunto convexo y* $f : C \to \mathrm{R}$ *una función. Se dice que:*

(a) f *es* **cóncava** *en* C, *si* $\forall x, y \in C$, $\forall t \in [0,1]$, *se cumple*
$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y) \tag{1}$$

(b) f *es* **estrictamente cóncava** *en* C, *si* $\forall x, y \in C$,*con* $x \neq y$, $\forall t \in]0,1[$, *se cumple*
$$f(tx + (1-t)y) > tf(x) + (1-t)f(y) \tag{2}$$

(c) f *es* **convexa** *en* C, *si* $\forall x, y \in C$, $\forall t \in [0,1]$, *se cumple*
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \tag{3}$$

(d) f *es* **estrictamente convexa** *en* C, *si* $\forall x, y \in C$,*con* $x \neq y$, $\forall t \in]0,1[$, *se cumple*
$$f(tx + (1-t)y) < tf(x) + (1-t)f(y) \tag{4}$$

Las definiciones que siguen, se dan para funciones definidas en C un subconjunto convexo no vacío, en particular en $\mathbb{R}$.

Definición

Una función $f : C \to \mathbb{R}$, se llama:

(a) **casicóncava**, *si*

$$x, y \in C, t \in [0,1] \Rightarrow f(tx + (1-t)y) \geq \text{mín}\{f(x), f(y)\} \tag{5}$$

(b) **casiconvexa**, *si*

$$x, y \in C, t \in [0,1] \Rightarrow f(tx + (1-t)y) \leq \text{máx}\{f(x), f(y)\} \tag{6}$$

(c) **estrictamente casicóncava**, *si*

$$x, y \in C, x \neq y, t \in]0,1[\Rightarrow f(tx + (1-t)y) > \text{mín}\{f(x), f(y)\} \tag{7}$$

(d) **estrictamente casiconvexa**, *si*

$$x, y \in C, x \neq y, t \in]0,1[\Rightarrow f(tx + (1-t)y) < \text{máx}\{f(x), f(y)\} \tag{8}$$

Particularmente, las funciones de $\mathbb{R}$ en $\mathbb{R}$ que son monótonas, son ejemplos de funciones que son tanto casicóncavas como casiconvexas. Es fácil probar que toda función cóncava es casicóncava y que toda función convexa es casiconvexa, mas lo recíproco en general no es cierto. Además, , si f es casicóncava, entonces $-f$ es casiconvexa, y viceversa.

Dados un subconjunto no vacío C de $\mathbb{R}^n$ y una función $f : C \to \mathbb{R}$, y un número real λ. se definen los conjuntos:

Definición

$$L_\lambda(f) := \{x \in C : f(x) = \lambda\} \tag{9}$$

"conjunto de nivel λ de f".

$$S_\lambda(f) := \{x \in C : f(x) \leq \lambda\} \tag{10}$$

"conjunto de nivel inferior λ de f".

$$S^\lambda(f) := \{x \in C : f(x) \geq \lambda\} \tag{11}$$

"conjunto de nivel superior λ de f".

Dado un conjunto no vacío C y una función $f : C \to \mathbb{R}$. Se presentan las siguientes propiedades:

Proposición

(i) *f es casi-cóncava, si y solo si, para cada $\lambda \in \mathbb{R}$ se cumple que $S^{\lambda}(f)$ es un subconjunto convexo de C.*

(ii) *f es casi-convexa, si y solo si, para cada $\lambda \in \mathbb{R}$ se cumple que $S_{\lambda}(f)$ es un subconjunto convexo de C.*

Corolario

Sea C un conjunto convexo y $f : C \to \mathbb{R}$. Entonces:

(i) *Si f es convexa, entonces sus conjuntos de nivel inferior, son conjuntos convexos.*

(ii) *Si f es cóncava, entonces sus conjuntos de nivel superior, son conjuntos convexos.*

Para funciones diferenciables cóncavas y convexas, existe una caracterización. Como también cuando éstas son diferenciables de segundo orden.

Proposición

Sea f una función continuamente diferenciable en un conjunto abierto $\Omega \subset \mathbb{R}^n$ y C un subconjunto convexo no vacío de Ω. Entonces

(i) *f es convexa en C, si y solo si*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \,, \qquad \forall x, y \in C. \tag{12}$$

(ii) *f es estrictamente convexa, si y solo si*

$$f(y) > f(x) + \langle \nabla f(x), y - x \rangle \,, \qquad \forall x, y \in C, x \neq y.$$

(iii) *f es cóncava en C, si y solo si:*

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle \,, \qquad \forall x, y \in C. \tag{13}$$

(iv) *f es estrictamente cóncava, si $\cdots$*

Proposición

Sean $C \subset \mathbb{R}^n$ *un conjunto abierto y convexo, y* $f : C \to \mathbb{R}$ *una función de clase* C^2*.*

(i) f *es convexa si y solo si,* $Hf(x)$ *es positiva semidefinida para cada* $x \in C$*.*

(ii) *Si* $Hf(x)$ *es positiva definida para cada* $x \in C$ *entonces* f *es estrictamente convexa.*

Proposición

Sean $C \subset \mathbb{R}^n$ un conjunto abierto y convexo, y $f : C \to \mathbb{R}$ una función de clase C^2.

(i) *f es convexa si y solo si, $Hf(x)$ es positiva semidefinida para cada $x \in C$.*

(ii) *Si $Hf(x)$ es positiva definida para cada $x \in C$ entonces f es estrictamente convexa.*

Para(i): Si f es de clase C^2 en C, entonces para cada $x, y \in C$ existe $x_y \in [x\ y]$ tal que

$$f(x) = f(y) + \langle \nabla f(y), x - y \rangle + \frac{1}{2} \langle x - y, Hf(x_y)(x - y) \rangle (*)$$

y tomando como dato que $Hf(x_y)$ es positiva semidefinida:

$$f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle.$$

Recíprocamente, suponga que para algún $x \in C$, existe $d \neq 0$ tal que $\langle d, Hf(x)d < 0$, entonces en (*) por la continuidad de Hf se sigue que para $\epsilon > 0$ suficientemente pequeño con $y = x + \epsilon d$, se obtiene una contradicción respecto a la convexidad de f.

Dada un subconjunto no vacío C de un e.n. E y $f : C \to \mathbb{R}$ una función. Se definen:

(i) El **epigrafo** de f, es un subconjunto de $E \times \mathbb{R}$ dado por

$$epi(f) := \{(x, \alpha) \in C \times \mathbb{R} \ : \ f(x) \leq \alpha\}$$

Dada un subconjunto no vacío C de un e.n. E y $f : C \to \mathbb{R}$ una función. Se definen:

(i) El **epigrafo** de f, es un subconjunto de $E \times \mathbb{R}$ dado por

$$epi(f) := \{(x, \alpha) \in C \times \mathbb{R} \ : \ f(x) \leq \alpha\}$$

(i) El **hipografo** de f, es un subconjunto de $E \times \mathbb{R}$ dado por

$$hip(f) := \{(x, \alpha) \in C \times \mathbb{R} \ : \ f(x) \geq \alpha\}$$

Dada un subconjunto no vacío C de un e.n. E y $f : C \to \mathbb{R}$ una función. Se definen:

(i) El **epigrafo** de f, es un subconjunto de $E \times \mathbb{R}$ dado por

$$epi(f) := \{(x, \alpha) \in C \times \mathbb{R} \ : \ f(x) \leq \alpha\}$$

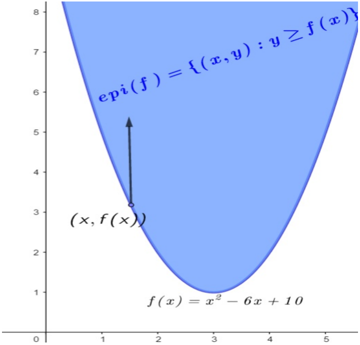
(i) El **hipografo** de f, es un subconjunto de $E \times \mathbb{R}$ dado por

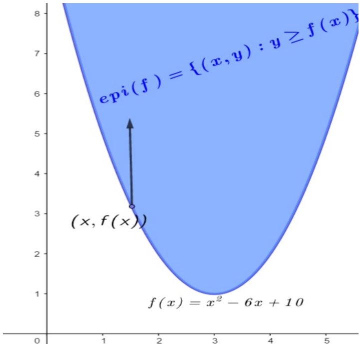
$$hip(f) := \{(x, \alpha) \in C \times \mathbb{R} \ : \ f(x) \geq \alpha\}$$

Propiedad: Si adicionalmente C es un conjunto convexo, entonces se cumple:

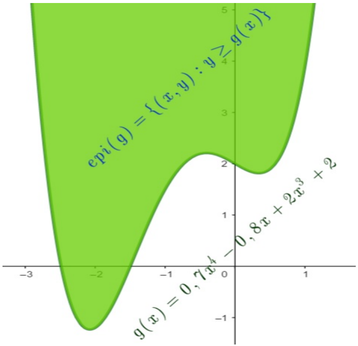
(a) f es una función convexa si y solo si $epi(f)$ es un subconjunto convexo de $E \times \mathbb{R}$.

(b) f es una función cóncava si y solo si $hip(f)$ es un subconjunto convexo de $E \times \mathbb{R}$.

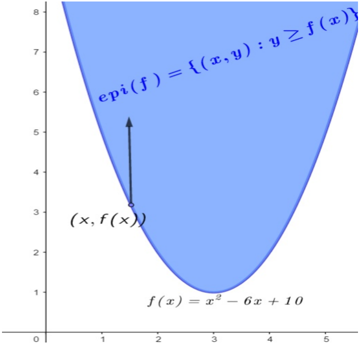

$epi(f) = \{(x,y) : y \geq f(x)\}$
$(x, f(x))$
$f(x) = x^2 - 6x + 10$

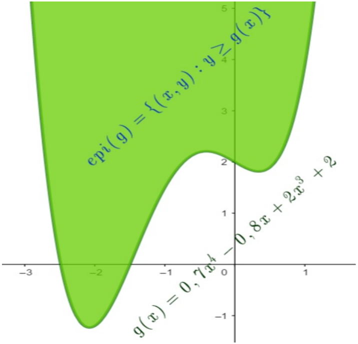


$epi(f) = \{(x,y) : y \geq f(x)\}$
$(x, f(x))$
$f(x) = x^2 - 6x + 10$



$epi(g) = \{(x,y) : y \geq g(x)\}$
$g(x) = 0,7x^4 - 0,8x + 2x^3 + 2$





Note que $epi(f)$ es un conjunto convexo de $\mathbb{R}^2$ mas $epi(g)$ no es un conjunto convexo de $\mathbb{R}^2$.

Propiedades

En esta parte, sea C un subconjunto no vacío y convexo de un espacio normado (como puede ser $\mathbb{R}^n$)

- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces para toda constante no negativa α, la función αf definida por $(\alpha f)(x) = \alpha f(x)$ para cada $x \in C$, también es convexa(cóncava).

Propiedades

En esta parte, sea C un subconjunto no vacío y convexo de un espacio normado (como puede ser $\mathbb{R}^n$)

- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces para toda constante no negativa α, la función αf definida por $(\alpha f)(x) = \alpha f(x)$ para cada $x \in C$, también es convexa(cóncava).
- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces la función $-f$ es cóncava(convexa).

En esta parte, sea C un subconjunto no vacío y convexo de un espacio normado (como puede ser $\mathbb{R}^n$)

- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces para toda constante no negativa α, la función αf definida por $(\alpha f)(x) = \alpha f(x)$ para cada $x \in C$, también es convexa(cóncava).
- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces la función $-f$ es cóncava(convexa).
- Si $f_1, \cdots, f_m$ son funciones convexas en C, entonces la función $f = f_1 + \cdots + f_m$ es convexa.

En esta parte, sea C un subconjunto no vacío y convexo de un espacio normado (como puede ser $\mathbb{R}^n$)

- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces para toda constante no negativa α, la función αf definida por $(\alpha f)(x) = \alpha f(x)$ para cada $x \in C$, también es convexa(cóncava).
- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces la función $-f$ es cóncava(convexa).
- Si $f_1, \cdots, f_m$ son funciones convexas en C, entonces la función $f = f_1 + \cdots + f_m$ es convexa.
- Si $f_1, \cdots, f_m$ son funciones cóncavas en C, entonces la función $f = f_1 + \cdots + f_m$ es cóncava.

En esta parte, sea C un subconjunto no vacío y convexo de un espacio normado (como puede ser $\mathbb{R}^n$)

- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces para toda constante no negativa α, la función αf definida por $(\alpha f)(x) = \alpha f(x)$ para cada $x \in C$, también es convexa(cóncava).
- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces la función $-f$ es cóncava(convexa).
- Si $f_1, \cdots, f_m$ son funciones convexas en C, entonces la función $f = f_1 + \cdots + f_m$ es convexa.
- Si $f_1, \cdots, f_m$ son funciones cóncavas en C, entonces la función $f = f_1 + \cdots + f_m$ es cóncava.
- Si $f_1, \cdots, f_m$ son funciones convexas en C, entonces la función $f = \max\limits_{i=1,m} f_i$ definida por

$$f(x) = \max_{i=1,m} f_i(x) \, , \qquad\qquad x \in C$$

también es convexa.

En esta parte, sea C un subconjunto no vacío y convexo de un espacio normado (como puede ser $\mathbb{R}^n$)

- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces para toda constante no negativa α, la función αf definida por $(\alpha f)(x) = \alpha f(x)$ para cada $x \in C$, también es convexa(cóncava).
- Si $f : C \to \mathbb{R}$ es una función convexa(cóncava), entonces la función $-f$ es cóncava(convexa).
- Si $f_1, \cdots, f_m$ son funciones convexas en C, entonces la función $f = f_1 + \cdots + f_m$ es convexa.
- Si $f_1, \cdots, f_m$ son funciones cóncavas en C, entonces la función $f = f_1 + \cdots + f_m$ es cóncava.
- Si $f_1, \cdots, f_m$ son funciones convexas en C, entonces la función $f = \max\limits_{i=1,m} f_i$ definida por

$$f(x) = \max_{i=1,m} f_i(x) , \qquad x \in C$$

también es convexa.

- Si $f_1, \cdots, f_m$ son funciones cóncavas en C, entonces la función $f = \min\limits_{i=1,m} f_i$ definida por

$$f(x) = \min_{i=1,m} f_i(x) , \qquad x \in C$$

también es cóncava.

Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable[1]. Se dice que f es fuertemente convexa de módulo $c > 0$ si satisface

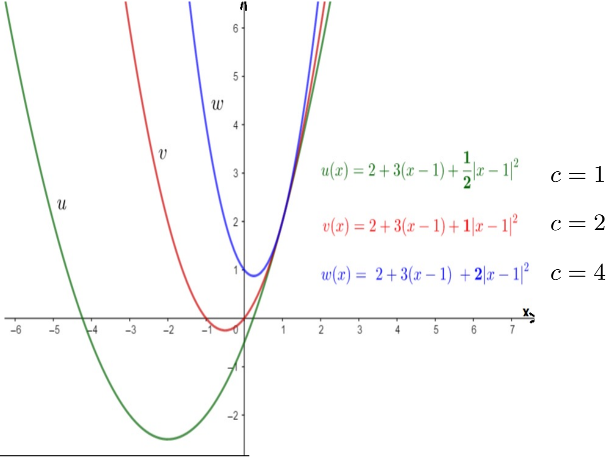
$$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} c \|x - x_0\|^2, \qquad \forall x, x_0 \in \mathbb{R}^n.$$

[1] $\mathbb{R}^n$ puede ser reemplazado por un convexo abierto

Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable[1]. Se dice que f es fuertemente convexa de módulo $c > 0$ si satisface

$$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} c \|x - x_0\|^2, \qquad \forall x, x_0 \in \mathbb{R}^n.$$

¿Cómo se comportan las funciones del lado derecho de la inecuación previa?



[1] $\mathbb{R}^n$ puede ser reemplazado por un convexo abierto

Observación

¿Cuál es la implicancia geométrica de la gráfica de f fuertemente convexa?

Si además f es de clase C^2, entonces $Hf(x)$ es positiva definida para cada $x \in \mathbb{R}^n$. En consecuencia, f es estrictamente convexa.

Proposición

f es fuertemente convexa de módulo $c > 0$ si, y solo si, la función $f - \frac{1}{2}c\| \ \|^2$ es convexa.

Ejemplo:

Observación

¿Cuál es la implicancia geométrica de la gráfica de f fuertemente convexa?

Si además f es de clase C^2, entonces $Hf(x)$ es positiva definida para cada $x \in \mathbb{R}^n$. En consecuencia, f es estrictamente convexa.

Proposición

f es fuertemente convexa de módulo $c > 0$ si, y solo si, la función $f - \frac{1}{2}c\| \ \|^2$ es convexa.

Ejemplo:Para cada $\alpha > 0$, la función $f(x) = \alpha\|x\|^2$ es fuertemente convexa. En general toda función de la forma $f(x) = \alpha\|x - x_0\|^2$.

Proposición

Sea C convexo abierto y $f : C \to \mathbb{R}$ fuertemente convexa de clase C^2. Para cada $x_0 \in C$, los conjuntos de nivel inferior $\{x \in C : f(x) \leq f(x_0)\}$ son convexos y compactos.

Proposición

Sea C convexo abierto y $f : C \to \mathbb{R}$ fuertemente convexa de clase C^2. Para cada $x_0 \in C$, los conjuntos de nivel inferior $\{x \in C : f(x) \leq f(x_0)\}$ son convexos y compactos.

Dem: Fijamos $x_0 \in C$, el conjunto $L := \{x \in C : f(x) \leq f(x_0)\}$ es convexo y cerrado. Para cada $x \in L$, $f(x) - f(x_0) \leq 0$.

Proposición

Sea C convexo abierto y $f : C \to \mathbb{R}$ fuertemente convexa de clase C^2. Para cada $x_0 \in C$, los conjuntos de nivel inferior $\{x \in C : f(x) \leq f(x_0)\}$ son convexos y compactos.

Dem: Fijamos $x_0 \in C$, el conjunto $L := \{x \in C : f(x) \leq f(x_0)\}$ es convexo y cerrado. Para cada $x \in L$, $f(x) - f(x_0) \leq 0$.
De la convexidad fuerte, existe $c > 0$ tal que

$$\begin{aligned} f(x) - f(x_0) &\geq \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2} c \|x - x_0\|^2 \\ &\geq -\|\nabla f(x_0)\| \|x - x_0\| + \frac{1}{2} c \|x - x_0\|^2 \end{aligned}$$

Para todo $x \neq x_0$, se sigue que $\|x - x_0\| \leq \dfrac{2}{c} \|\nabla f(x_0)\|$. De esto se concluye que para $x \in L$, $\|x\|$ es acotado.

Proposición

Sea C convexo abierto y $f : C \to \mathbb{R}$ fuertemente convexa de clase C^2. Para cada $x_0 \in C$, los conjuntos de nivel inferior $\{x \in C : f(x) \leq f(x_0)\}$ son convexos y compactos.

Dem: Fijamos $x_0 \in C$, el conjunto $L := \{x \in C : f(x) \leq f(x_0)\}$ es convexo y cerrado. Para cada $x \in L$, $f(x) - f(x_0) \leq 0$.
De la convexidad fuerte, existe $c > 0$ tal que

$$\begin{aligned} f(x) - f(x_0) &\geq \langle \nabla f(x_0), x - x_0 \rangle + \frac{1}{2}c\|x - x_0\|^2 \\ &\geq -\|\nabla f(x_0)\|\|x - x_0\| + \frac{1}{2}c\|x - x_0\|^2 \end{aligned}$$

Para todo $x \neq x_0$, se sigue que $\|x - x_0\| \leq \dfrac{2}{c}\|\nabla f(x_0)\|$. De esto se concluye que para $x \in L$, $\|x\|$ es acotado.

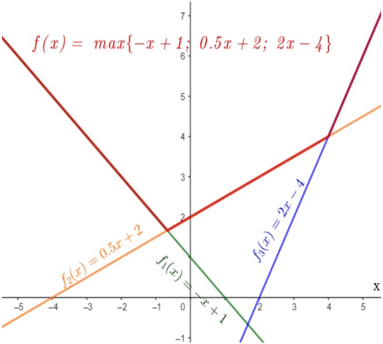
Corolario

$f : \mathbb{R}^n \to \mathbb{R}$ con las condiciones previas, entonces existe un único $\overline{x} \in \mathbb{R}^n$ donde f alcanza su valor mínimo global.

Observación

la función $f(x) = x^4$, $x \in \mathbb{R}$ *es estrictamente convexa y tiene un único minimizante en* $\overline{x} = 0$, *mas* f *no es fuertemente convexa, pues caso contrario debe existir un* $c > 0$ *tal que la función* $g(x) = f(x) - \dfrac{c}{2}x^2$ *sea convexa en* $\mathbb{R}$, *y como* $g''(x) = 12x^2 - c$ *y esto no es no-positivo en* $\mathbb{R}$, *se sigue que* g *no es convexa y en consecuencia* f *no es fuertemente convexa.*

- Toda función lineal $f : \mathbb{R} \to \mathbb{R}$ ($f(x) = ax + b$) es convexa y cóncava a la vez (no es estrictamente en ningún caso).
- Para constantes $a > 0$ y $x_0 \in \mathbb{R}$, la función $f(x) = a|x - x_0|$ es convexa, mientras que para $a < 0$, la función $g(x) = a|x - x_0|$ es cóncava.
- Para $i = 1, \cdots, m$, dados $a_i \in \mathbb{R}, b_i \in \mathbb{R}$,
  la función $f(x) = \max\limits_{i=1,m} (ax_i + b_i)$
  es convexa.



- la función $f(x) = ax^2 + bx + c$ es estictamente convexa si $a > 0$ y es estrictamente cóncava si $a < 0$.
- La función $f(x) = e^{ax}$ es convexa para cualquier constante real a.
- La función $f(x) = \max\{0, x\}$ es convexa en $\mathbb{R}$.
- La función $f(x) = x^p$ es cóncava en $[0, +\infty)$ para $0 < p < 1$.
- La función $f(x) = log(x)$ es cóncava en $(0, +\infty)$
- La función $f(x) = \begin{cases} x\log(x), & \text{si } x > 0 \\ 0, & \text{si } x = 0 \end{cases}$ es convexa en $[0, +\infty)$

- Para todo $a \in \mathbb{R}^n$ y $b \in \mathbb{R}$, la función $F(x) = a'x + b$ es cóncava y convexa a la vez.
- Para constantes $a_1, \cdots, a_n$, la función $F(x) = \displaystyle\sum_{i=1}^{n} e^{a_i x_i}$ es convexa en $\mathbb{R}^n$.

- Para todo $a \in \mathbb{R}^n$ y $b \in \mathbb{R}$, la función $F(x) = a'x + b$ es cóncava y convexa a la vez.
- Para constantes $a_1, \cdots, a_n$, la función $F(x) = \sum_{i=1}^{n} e^{a_i x_i}$ es convexa en $\mathbb{R}^n$.
- la función $F(x) = \log(\prod_{i=1}^{n} x_i)$ es cóncava en $\mathbb{R}^n_{++}$

- Para todo $a \in \mathbb{R}^n$ y $b \in \mathbb{R}$, la función $F(x) = a'x + b$ es cóncava y convexa a la vez.
- Para constantes $a_1, \cdots, a_n$, la función $F(x) = \sum_{i=1}^{n} e^{a_i x_i}$ es convexa en $\mathbb{R}^n$.
- la función $F(x) = \log(\prod_{i=1}^{n} x_i)$ es cóncava en $\mathbb{R}^n_{++}$
- La función $F(x) = \sum_{i=1}^{n}(-x_i \log(x_i)) = -\sum_{i=1}^{n}(x_i \log(x_i))$ es cóncava en $\mathbb{R}^n_+$ con el convenio $0.\infty = 0$.

Definición

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *una función convexa y propia, y* $x \in dom(f)$. *Un vector* z *tal que*

$$f(y) \geq f(x) + \langle z, y - x \rangle, \qquad \forall y \in R^n \tag{14}$$

se llama ***un subgradiente de*** f ***en*** x.

Definición

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *una función convexa y propia, y* $x \in dom(f)$. *Un vector* z *tal que*

$$f(y) \geq f(x) + \langle z, y - x \rangle, \qquad \forall y \in R^n \tag{14}$$

se llama un subgradiente de f *en* x.*El conjunto de los vectores subgradientes de* f *en* x *se llama el subdiferencial de* f *en* x *y se denota por* $\partial f(x)$.

Definición

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *una función convexa y propia, y* $x \in dom(f)$*. Un vector* z *tal que*

$$f(y) \geq f(x) + \langle z, y - x \rangle, \qquad \forall y \in R^n \tag{14}$$

se llama un subgradiente de f *en* x*.El conjunto de los vectores subgradientes de* f *en* x *se llama el subdiferencial de* f *en* x *y se denota por* $\partial f(x)$*.*
Si $x \notin dom(f)$*, se define* $\partial f(x) = \emptyset$*.*

Definición

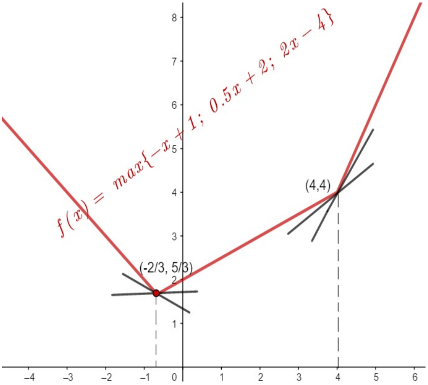
Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ una función convexa y propia, y $x \in dom(f)$. Un vector z tal que

$$f(y) \geq f(x) + \langle z, y - x \rangle, \qquad \forall y \in R^n \tag{14}$$

se llama un subgradiente de f en x.El conjunto de los vectores subgradientes de f en x se llama el subdiferencial de f en x y se denota por $\partial f(x)$.
Si $x \notin dom(f)$, se define $\partial f(x) = \emptyset$.

NOTA: Si $x \in dom(f)$ esto no garantiza que $\partial f(x) \neq \emptyset$. Por ejemplo, considere la función $f(x) = \begin{cases} 1, & \text{si } x = 0 \\ x^2, & \text{si } 0 < x \leq 1 \\ +\infty, & \text{o.c.} \end{cases}$ Si existe $z \in \mathbb{R}$ tal que $z \in \partial f(0)$, entonces $x^2 \geq 1 + zx$, $\forall x \in dom(f)$. Esto genera una contradicción para $x = 0$.

Para la función $f(x) = \text{máx}\{-x+1; 0{,}5x+2; 2x-4\}$



$$\partial f(-2/3) = \cdots \,, \qquad\qquad \partial f(4) = \cdots$$

Lema

Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa y $x \in dom(f)$; entonces $z \in \partial f(x)$, si y solamente si,

$$f'(x,d) \geq \langle z, d\rangle, \ \forall d \in \mathbb{R}^n.$$

Teorema

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ una función convexa. Asuma que $x \in int(dom(f))$, entonces $\partial f(x)$ es un conjunto convexo, compacto y no vacío. Además, para cada $d \in \mathbb{R}^n$ se cumple

$$f'(x,d) = \max_{z \in \partial f(x)} \langle z, d\rangle$$

(Teorema 2.74 del texto de Ruszczyński)

En un espacio normado E con norma $\|\ \| : E \to \mathbb{R}$, esta función es convexa, dado que por desigualdad triangular cumple

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\|, \qquad \text{para cada } t \in [0,1], x, y \in E$$

En particular en $\mathbb{R}^n$ toda $p - norma$ con $p \geq 1$ definida por

$$x(x_1, \cdots, x_n) \Rightarrow \|x\|_p := (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

La convexidad de la función norma

En un espacio normado E con norma $\|\ \| : E \to \mathbb{R}$, esta función es convexa, dado que por desigualdad triangular cumple

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\|, \qquad \text{para cada } t \in [0,1], x, y \in E$$

En particular en $\mathbb{R}^n$ toda $p - norma$ con $p \geq 1$ definida por

$$x(x_1, \cdots, x_n) \Rightarrow \|x\|_p := (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

Consideremos la norma euclidiana $\|\ \|$ en $\mathbb{R}^n$ y la función $f(x) = \|x\|, \forall x \in \mathbb{R}^n$.Calcular $\partial f(0)$

La convexidad de la función norma

En un espacio normado E con norma $\| \ \|: E \to \mathbb{R}$, esta función es convexa, dado que por desigualdad triangular cumple

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\|, \qquad \text{para cada } t \in [0,1], x, y \in E$$

En particular en $\mathbb{R}^n$ toda $p - norma$ con $p \geq 1$ definida por

$$x(x_1, \cdots, x_n) \Rightarrow \|x\|_p := (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

Consideremos la norma euclidiana $\| \ \|$ en $\mathbb{R}^n$ y la función $f(x) = \|x\|, \forall x \in \mathbb{R}^n$.Calcular $\partial f(0)$
Para cada $d \in \mathbb{R}^n$, se tiene
$f'(0,d) = \lim_{t \downarrow 0} \dfrac{\|td\|}{t} = \|d\|$

En un espacio normado E con norma $\| \ \| : E \to \mathbb{R}$, esta función es convexa, dado que por desigualdad triangular cumple

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\|, \qquad \text{para cada } t \in [0,1], x, y \in E$$

En particular en $\mathbb{R}^n$ toda $p - norma$ con $p \geq 1$ definida por

$$x(x_1, \cdots, x_n) \Rightarrow \|x\|_p := (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

Consideremos la norma euclidiana $\| \ \|$ en $\mathbb{R}^n$ y la función $f(x) = \|x\|, \forall x \in \mathbb{R}^n$.Calcular $\partial f(0)$
Para cada $d \in \mathbb{R}^n$, se tiene
$f'(0, d) = \lim\limits_{t \downarrow 0} \dfrac{\|td\|}{t} = \|d\|$
Por el lema previo, $z \in \partial f(0)$ si y solo si, $\|d\| \geq \langle z, d \rangle, \ \forall d \in \mathbb{R}^n$ y esto es válido si y solo si $\|z\| \leq 1$. Es decir, $\partial f(0) = \{z \in \mathbb{R}^n : \|z\| \leq 1\}$

En un espacio normado E con norma $\| \ \| : E \to \mathbb{R}$, esta función es convexa, dado que por desigualdad triangular cumple

$$\|tx + (1-t)y\| \leq t\|x\| + (1-t)\|y\|, \qquad \text{para cada } t \in [0,1], x, y \in E$$

En particular en $\mathbb{R}^n$ toda $p - norma$ con $p \geq 1$ definida por

$$x(x_1, \cdots, x_n) \Rightarrow \|x\|_p := (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

Consideremos la norma euclidiana $\| \ \|$ en $\mathbb{R}^n$ y la función $f(x) = \|x\|, \forall x \in \mathbb{R}^n$.Calcular $\partial f(0)$
Para cada $d \in \mathbb{R}^n$, se tiene
$f'(0,d) = \lim_{t \downarrow 0} \dfrac{\|td\|}{t} = \|d\|$
Por el lema previo, $z \in \partial f(0)$ si y solo si, $\|d\| \geq \langle z, d \rangle$, $\forall d \in \mathbb{R}^n$ y esto es válido si y solo si $\|z\| \leq 1$. Es decir, $\partial f(0) = \{z \in \mathbb{R}^n : \|z\| \leq 1\}$
Particularmente en $\mathbb{R}$, la función valor absoluto es tal que

$$\partial(| \ |)(x) = \begin{cases} \{1\}, & \text{si } x > 0 \\ [-1, 1], & \text{si } x = 0 \\ \{-1\}, & \text{si } x < 0 \end{cases}$$

Propiedad:
Sean A una matriz de orden $m \times n$, $b \in \mathbb{R}^m$ y $f : \mathbb{R}^m \to \mathbb{R}$ una función convexa(cóncava). Entonces la función $g(x) = f(Ax - b)$ es convexa(cóncava) en $\mathbb{R}^n$.

Propiedad:
Sean A una matriz de orden $m \times n$, $b \in \mathbb{R}^m$ y $f : \mathbb{R}^m \to \mathbb{R}$ una función convexa(cóncava). Entonces la función $g(x) = f(Ax - b)$ es convexa(cóncava) en $\mathbb{R}^n$.
Consecuentemente, para una norma $\| \ \|$ en $\mathbb{R}^m$, la función $g(x) = \|Ax - b\|$ es convexa en $\mathbb{R}^n$.

Propiedad:
Sean A una matriz de orden $m \times n$, $b \in \mathbb{R}^m$ y $f : \mathbb{R}^m \to \mathbb{R}$ una función convexa(cóncava). Entonces la función $g(x) = f(Ax - b)$ es convexa(cóncava) en $\mathbb{R}^n$.
Consecuentemente, para una norma $\| \ \|$ en $\mathbb{R}^m$, la función $g(x) = \|Ax - b\|$ es convexa en $\mathbb{R}^n$.
Tarea: Hallar $\partial g(x)$.

Proposición

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ una función convexa propia. Un punto $\overline{x}$ es un mínimo global de f si y solo si,

$$0 \in \partial f(\overline{x}).$$

Proposición

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ una función convexa propia. Un punto $\overline{x}$ es un mínimo global de f si y solo si,

$$0 \in \partial f(\overline{x}).$$

Note que si f es diferenciable en un entorno de $\overline{x}$, entonces la última condición se reduce a $\nabla f(\overline{x}) = 0$.

Proposición

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ una función convexa propia. Un punto $\overline{x}$ es un mínimo global de f si y solo si,

$$0 \in \partial f(\overline{x}).$$

Note que si f es diferenciable en un entorno de $\overline{x}$, entonces la última condición se reduce a $\nabla f(\overline{x}) = 0$.

(Revisar el ejemplo 3.6 del texto de Ruszczyński)

Sean $A = [a_{ij}]$ una matriz constante de orden n, $b = (b_1, \cdots, b_n)'$ un vector constante n-dimensional y c una constante real. Una función cuadrática en las variables $x_1, x_2, \cdots, x_n$ es de la forma

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c \tag{15}$$

Si x denota al vector $(x_1, \cdots, x_n)'$, entonces la función previa se puede escribir como

$$f(x) = x' A x + b' x + c$$

Sean $A = [a_{ij}]$ una matriz constante de orden n, $b = (b_1, \cdots, b_n)'$ un vector constante n-dimensional y c una constante real. Una función cuadrática en las variables $x_1, x_2, \cdots, x_n$ es de la forma

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c \qquad (15)$$

Si x denota al vector $(x_1, \cdots, x_n)'$, entonces la función previa se puede escribir como

$$f(x) = x'Ax + b'x + c$$

Observación: En la práctica se asume que A es una matriz simétrica y que el primer término que describe a f está multiplicado por 1/2. Es decir f toma la forma:

$$f(x) = \frac{1}{2} x'Ax + b'x + c\,, \qquad x \in \mathbb{R}^n \qquad (16)$$

Note que en esta última situación, $\nabla f(x) = Ax + b$ y $Hf(x) = A$.
¿La ecuación $Ax + b = 0$ tiene solución? ¿Cómo es la matriz A en cuestión de definida-semidefinida?

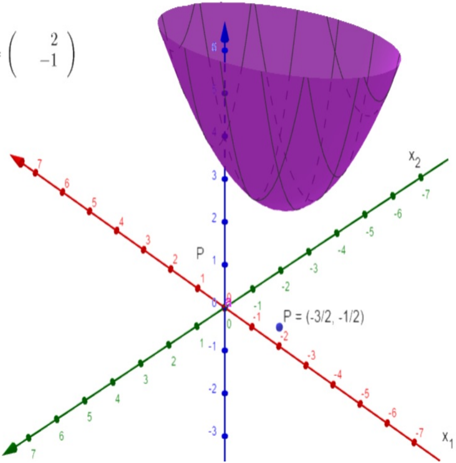
Cómo es una función cuadrática (analíticamente)?

Veamos algunos casos:

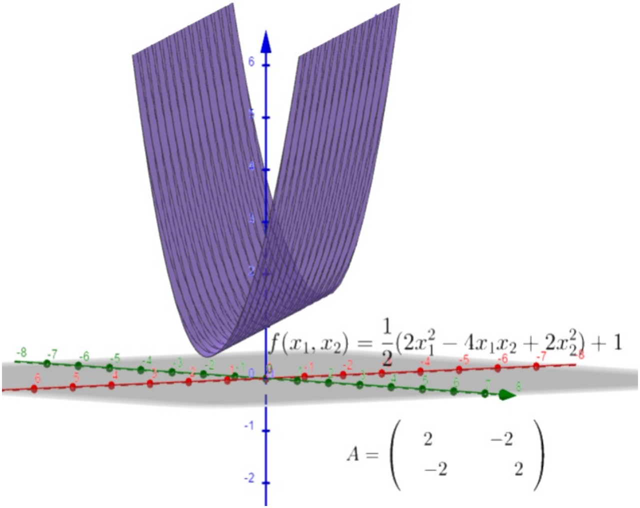
1)

$$f(x) = \frac{1}{2}x'Ax + b'x + 4$$

$$A = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}; b = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$



2)


$f(x_1, x_2) = \frac{1}{2}(2x_1^2 - 4x_1x_2 + 2x_2^2) + 1$
$A = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$

3) ¿Qué ocurre con la función $f(x_1, x_2) = \dfrac{1}{2}(2x_1^2 - 4x_1x_2 + 2x_2^2) + x_1$

3) ¿Qué ocurre con la función $f(x_1, x_2) = \dfrac{1}{2}(2x_1^2 - 4x_1x_2 + 2x_2^2) + x_1$

Observación:
Dada la función cuadrática $f(x) = \dfrac{1}{2}x'Ax + b'x + c$ con A una matriz simétrica, b un vector n-dimensional y c constante real.

- Si A es positiva semidefinida(así f es convexa) y $\nabla f(x^*) = 0$, entonces x^* es un mínimo global para f.
- Si A es negativa semidefinida(así f es cóncava) y $\nabla f(x^*) = 0$, entonces x^* es un máximo global para f.

3) ¿Qué ocurre con la función $f(x_1, x_2) = \frac{1}{2}(2x_1^2 - 4x_1x_2 + 2x_2^2) + x_1$

Observación:
Dada la función cuadrática $f(x) = \frac{1}{2}x'Ax + b'x + c$ con A una matriz simétrica, b un vector n-dimensional y c constante real.

- Si A es positiva semidefinida(así f es convexa) y $\nabla f(x^*) = 0$, entonces x^* es un mínimo global para f.
- Si A es negativa semidefinida(así f es cóncava) y $\nabla f(x^*) = 0$, entonces x^* es un máximo global para f.

¿Tal x^* es único ?

Sean $c_1, \cdots, c_n$ constantes.

¿Es la función $F(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} |x_i - c_i|^2$, una función cuadrática ?

Sean $c_1, \cdots, c_n$ constantes.

¿Es la función $F(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} |x_i - c_i|^2$, una función cuadrática ?

Questions

- ¿Cuándo una función cuadrática es fuertemente convexa? ¿La función norma euclidiana al cuadrado es fuertemente convexa?

Sean $c_1, \cdots, c_n$ constantes.

¿Es la función $F(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} |x_i - c_i|^2$, una función cuadrática ?

Questions

- ¿Cuándo una función cuadrática es fuertemente convexa? ¿La función norma euclidiana al cuadrado es fuertemente convexa?
- Si a una función fuertemente convexa se le suma una función convexa ¿ el resultado sigue siendo fuertemente convexa?

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{17}$$

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{17}$$

El conjunto de las direcciones tangente a X en el punto x se denota por $T_X(x)$.

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{17}$$

El conjunto de las direcciones tangente a X en el punto x se denota por $T_X(x)$.
Equivalentemente, $d \in T_X(x)$ si y solo si:
existe una sucesión d^k que converge a d y una sucesión de números reales positivos t_k con $t_k \downarrow 0$ y $x + t_k d^k \in X$.

Sea X un subconjunto no vacío de $\mathbb{R}^n$, el vector $d \in \mathbb{R}^n$ se llama dirección tangente al conjunto X en el punto $x \in X$, si existe una sucesión x^k en X tal que $x^k \to x$, y una sucesión de escalares t_k tal que $t_k \downarrow 0$ y

$$d = \lim_{k \to +\infty} \frac{x^k - x}{t_k} \tag{17}$$

El conjunto de las direcciones tangente a X en el punto x se denota por $T_X(x)$. Equivalentemente, $d \in T_X(x)$ si y solo si:
existe una sucesión d^k que converge a d y una sucesión de números reales positivos t_k con $t_k \downarrow 0$ y $x + t_k d^k \in X$.

Proposición

$T_X(x)$ es un cono cerrado.

Definición

El cono de direcciones factibles en $x \in X$ está dado por

$$K_X(x) := \{d \in \mathbb{R}^n : d = \alpha(y - x), y \in X, \alpha \geq 0\}$$

Definición

El cono de direcciones factibles en $x \in X$ está dado por

$$K_X(x) := \{d \in \mathbb{R}^n : d = \alpha(y - x), y \in X, \alpha \geq 0\}$$

Proposición

Sea X un subconjunto convexo no vacío de $\mathbb{R}^n$ y $x \in X$, entonces

$$T_X(x) = \overline{K_X(x)} \tag{18}$$

Definición

El cono de direcciones factibles en $x \in X$ está dado por

$$K_X(x) := \{d \in \mathbb{R}^n : d = \alpha(y - x), y \in X, \alpha \geq 0\}$$

Proposición

Sea X un subconjunto convexo no vacío de $\mathbb{R}^n$ y $x \in X$, entonces

$$T_X(x) = \overline{K_X(x)} \tag{18}$$

Dem: Sea $d \in K_X(x)$, si $d = 0$ obviamente $d \in T_X(x)$, si $d \neq 0$, entonces existen $y \in X, \alpha > 0$ tales que $d = \alpha(y - x)$. Sean $x^k := x + \frac{1}{k}d \in X$ y cumple $\dfrac{x^k - x}{\frac{1}{k}} = d$, entonces $d \in T_X(x)$. Lo que significa que $K_X(x) \subset T_X(x)$ y así $\overline{K_X(x)} \subset T_X(x)$.
La otra inclusión (ejercicio).