Treceava sesión

Análisis Convexos - CM3E2

Jonathan Munguia[1]

[1]Facultad de Ciencias
Universidad Nacional de Ingeniería

25 de mayo de 2021

Outline

Definición 1 (Función fuertemente convexa)

Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es fuertemente convexa de coeficiente (o módulo) $\alpha > 0$, si para todo $t \in [0, 1]$ y para todo $x, y \in \mathbb{R}^n$, se cumple

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\alpha}{2}t(1-t)\|x-y\|^2.$$

Ver Definición 2.5 de [2].

Proposición 1

Sea $a \in \mathbb{R}^n$ fijo arbitario. Entonces, $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es fuertemente convexa de coeficiente $\mu > 0$ si y solo si la función $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ definida por

$$g(x) = f(x) - \frac{\mu}{2}\|x-a\|^2 \quad \forall x \in \mathbb{R}^n,$$

es convexa.

Demostración

Ver Proposición 2.23 de [2] o Lema 2.20 de [4]. Sin pérdidad de generalidad veamos el caso $a = 0$.

$\Longrightarrow)$ Dados $x, y \in \mathrm{dom}(f)$, $t \in [0,1]$, se tiene

$$\begin{aligned} g(tx+(1-t)y) &= f(tx+(1-t)y) - \frac{\mu}{2}\|tx+(1-t)y\|^2 \\ &\leq tf(x)+(1-t)f(y) - \frac{\mu}{2}\left[t(1-t)\|y-x\|^2 + \|tx+(1-t)y\|^2\right] \\ &= tf(x)+(1-t)f(y) - \frac{\mu}{2}\left[t\|x\|^2 + (1-t)\|y\|^2\right] \\ &= tg(x)+(1-t)g(y). \end{aligned}$$

$$\cancel{t(1-t)}\left[\|x\|^2 - \cancel{2\langle x,y\rangle} + \|y\|^2\right]$$

$$+ t^2\|x\|^2 + 2\cancel{t(1-t)\langle x,y\rangle} + (1-t)^2\|y\|^2$$

$$t(1-t)\|x\|^2 + t(1-t)\|y\|^2 + t^2\|x\|^2 + (1-t)^2\|y\|^2$$

$$= t\|x\|^2 + (1-t)\|y\|^2$$

$\Longleftarrow$) Dados $x, y \in \mathrm{dom}(g)$ y $t \in [0,1]$, se tiene

$$
\begin{aligned}
f(tx+(1-t)y) &= g(tx+(1-t)y) + \frac{\mu}{2}\|tx+(1-t)y\|^2 \\
&\leq tg(x) + (1-t)g(y) + \frac{\mu}{2}\|tx+(1-t)y\|^2 \\
= tf(x) + (1-t)f(y) - \frac{\mu}{2}\Big[t\|x\|^2 &+ (1-t)\|y\|^2 - \underbrace{\|tx+(1-t)y\|^2}\Big] \\
= tf(x) + (1-t)f(y) &- \frac{\mu}{2}t(1-t)\Big[\underbrace{\|x\|^2+\|y\|^2-2\langle x,y\rangle}\Big] \\
&= tf(x) + (1-t)f(y) - \frac{\mu}{2}t(1-t)\|x-y\|^2.
\end{aligned}
$$

$$\langle \nabla f(x) - \mu x - \nabla f(y) + \mu y, x - y \rangle \geq 0$$

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

Proposición 2

Sea $C \subset \mathbb{R}^n$ convexo, abierto, no vacío y $f : C \to \mathbb{R}$ diferenciable. Las siguientes afirmaciones son equivalentes

convexa

i) f *es fuertemente convexa con coeficiente* $\mu > 0$.

ii) $\langle \nabla f(x) - \nabla f(y), x - y \rangle \overset{\geq 0}{\geq} \mu \|x - y\|^2$ *para todo* $x, y \in C$.

iii) $f(y) \geq \underbrace{f(x) + \langle \nabla f(x), y - x \rangle}_{} + \frac{\mu}{2}\|y - x\|^2$ *para todo* $x, y \in C$.

Demostración

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

Si f diferenciable y fuertemente convexa, entonces por la Proposición 1 es equivalente a que $g = f - \frac{\mu}{2}\| \cdot \|^2$ es convexa y diferenciable, por lo tanto verifica la Proposición 3 de la sesión 12. Solo reemplazamos g y su gradiente $\nabla g(x) = \nabla f(x) - \mu x$ y concluimos reordenando.

$$\langle \nabla g(x) - \nabla g(y), x - y \rangle \geq 0$$

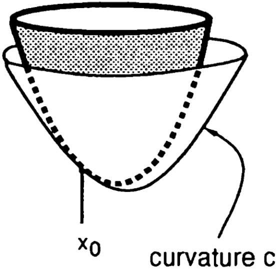
Observación 1

$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$

f es fuertemente convexa con módulo c, cuando está minorada por la función convexa cuadrática

$$x \longmapsto f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \frac{c}{2}\|x - x_0\|^2,$$

cuyo gradiente en x_0 es también $\nabla f(x_0)$. Esta propiedad tangencial, se ilustra a continuación.



f es convexa $\Leftrightarrow$ $\nabla^2 f(x)$ es semidefinida

$\langle \nabla^2 f(x) y, y \rangle \geq 0 \quad \forall y \in \mathbb{R}^n$

Proposición 3

Sea $C \subset \mathbb{R}^n$ convexo, abierto, no vacío y $f : C \to \mathbb{R}$ dos veces diferenciable. Entonces, f es fuertemente convexa con coeficiente $\mu > 0$ si y solo si para todo $x \in C$, la matriz Hessiana $\nabla^2 f(x)$ es definida positiva.

Demostración

Se sabe que $g = f - \frac{\mu}{2}\|\cdot\|^2$ convexa y 2 veces diferenciable, con matriz hessiana $\nabla^2 g(x) = \nabla^2 f(x) - \mu I$. Luego, por la Proposición 4 de la sesión 12, se tiene que es equivalente a que $\nabla^2 g(x)$ sea semidefinida positiva para todo $x \in C$. Se observa

$\langle \mu I y, y \rangle$

$$\begin{aligned} \underline{\forall y \in \mathbb{R}^n \setminus \{0\}} : \quad & \langle \nabla^2 g(x) y, y \rangle \geq 0. \\ & \langle \nabla^2 f(x) y, y \rangle \geq \mu \|y\|^2 > 0. \end{aligned}$$

> 0

Ejemplo 1

Las sgtes funciones son fuertemente convexas:

- $f(x) = x^2 - \cos x$ para $x \in \mathbb{R}$.
- $g(x) = \dfrac{\mu}{2}\|x\|^2$ para $x \in \mathbb{R}^n$.
- $h(x) = \sum_i x_i \log x_i$ para $x \in \Delta^n := \{x \in \mathbb{R}^n \,:\, \sum x_i = 1 \ \wedge \ x_i \geq 0\}$.
- Una función cuadrática $f(x) = \frac{1}{2}x^t Qx + b^t x + c$ es estrictamente convexa si y solo si es fuertemente convexa, donde $b, c \in \mathbb{R}^n$ y $Q \in \mathbb{R}^{n\times n}$.

$$\lim_{\|x\|\to+\infty} f(x) = +\infty \equiv \forall M \in \mathbb{R},\ \exists \delta > 0,\ \|x\| > \delta \Rightarrow f(x) > M$$

Definición 2 (Función coerciva)

Una función $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ se llama 0-coerciva si

$$f(x) \to +\infty \quad \text{siempre que} \quad \|x\| \to +\infty.$$

Se llamará 1-coerciva o solo coerciva si

$$\frac{f(x)}{\|x\|} \to +\infty \quad \text{siempre que} \quad \|x\| \to +\infty.$$

Ver Definición 3.2.5 de [1].

Proposición 4

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. Si los conjuntos de nivel $S_k(f)$ son acotados para todo $k \in \mathbb{N}$, entonces f es 0-coerciva. Si f es 0-coerciva, entonces para todo $t \in \mathbb{R}$, los conjuntos de nivel $S_t(f)$ son acotados.

$\lim_{\|x\| \to +\infty} f(x) = +\infty \equiv \forall M \in \mathbb{R}, \exists \delta > 0, \|x\| > \delta \Rightarrow f(x) > M$

Demostración

Ver Teorema 1.2. de [3].

- Sea $M > 0$, luego $S_M(f)$ es acotado pues $S_M(f) \subset S_k(f)$ para algún $k \in \mathbb{N}$. Por tanto sea $r > 0$ tal que $S_M(f) \subset B_r(0)$, esto es equivalente a

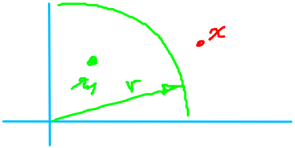
$$\|x\| > r \Longrightarrow f(x) > M.$$

- Sea f 0-coerciva y supongamos que $\exists t \in \mathbb{R}$ tal que $S_t(f)$ no es acotado. Luego, podemos construir una sucesión $(x_n) \subset S_t(f)$ tal que $\|x_n\| > n$, lo cual contradice la coercividad. $(\not\exists \Leftarrow)$ $f(x_n) \leq t$

$\exists r > 0, \forall x \in S_t(f) : \|x\| \leq r$

$\forall\ r = n \in \mathbb{N}, \exists x_n \in S_t(f)$ tq $\|x_n\| > n$

$x_1 \notin B_r[0]$
$\Rightarrow f(x_1) > M$



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

Observación 2

Si $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ es 0-coerciva y $f(x_1) < +\infty$ para algún $\exists x_1 \in \mathbb{R}^n$ (i.e f propia). Entonces, por la coercividad, para $M = f(x_1)$, existe $r > 0$ tal que $x \notin B_r[0] \to f(x) > M = f(x_1)$. Entonces hemos encontrado r suficientemente grande tal que

$\|x\| > r$

$\inf f(x)$ $\quad x \in B_r[0]$

$$\min_{x \in \mathbb{R}^n} f(x) = \min_{\underbrace{x \in B_r[0] \cap S_M(f)}_{\text{compacto}}} f(x).$$

$f(x) > f(x_1)$

acotado

Luego, por Weiertrass, solo basta verificar que $B_r[0] \cap S_M(f)$ sea compacto. Por ejemplo, tomando f sci.

Prop 2.2 de [2]

$S_M(f) = \{x \mid f(x) \leq M\}$

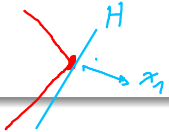
$S_M(f) \subset B_r[0]$

$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle$

Teorema 1

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa, sci y propia. Entonces, existen $\widehat{x}_0, x_1 \in \mathbb{R}^n$ tal que $f(\widehat{x}_0)$ es finito y

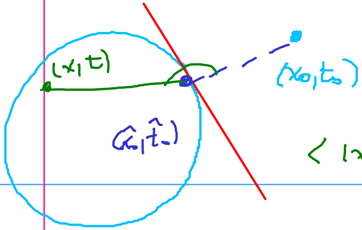
$$f(x) \geq f(\widehat{x}_0) + \langle x - \widehat{x}_0, x_1 \rangle \quad \forall x \in \mathbb{R}^n.$$



Demostración

Como f es propia, sea x_0 tal que $f(x_0) < +\infty$, luego $(x_0, f(x_0)) \in \operatorname{epi}(f)$ y como es sci, entonces $\operatorname{epi}(f)$ es convexo, cerrado no vacío. Dado que es propia elegimos t_0 tal que $-\infty < t_0 < f(x_0)$ asi que $(x_0, t_0) \notin \operatorname{epi}(f)$. Luego, por el teorema de la proyección, existe $(\widehat{x}_0, \widehat{t}_0) \in \operatorname{epi}(f)$ con

$$\langle (x_0, t_0) - (\widehat{x}_0, \widehat{t}_0), (x, t) - (\widehat{x}_0, \widehat{t}_0) \rangle \leq 0 \quad \forall (x, t) \in \operatorname{epi}(f). \tag{1}$$



$$\langle (x,t) - (\hat{x}_0,\hat{t}_0),\ (x_0,t_0) - (\hat{x}_0,\hat{t}_0) \rangle \le 0$$

$$\Rightarrow \langle (x-\hat{x}_0,\ t-\hat{t}_0),\ (x_0-\hat{x}_0,\ t_0-\hat{t}_0) \rangle \le 0$$

$\langle x_0 - \hat{x}_0, x - \hat{x}_0 \rangle + (t_0 - \hat{t}_0)(t - \hat{t}_0) \leq 0$

$\underset{f(\hat{x}_0)}{\shortparallel}$

cont...

Reescribimos la desigualdad

$$\langle x_0 - \widehat{x}_0, x - \widehat{x}_0 \rangle + (t_0 - \widehat{t}_0)(t - \widehat{t}_0) \leq 0\,,\ \ t \in \mathbb{R}\ y\ f(x) \leq t. \qquad (2)$$

Tomando $x = x_0$ y $t = f(x_0) + \lambda$ para $\lambda \geq 0$ conduce a

$$\underbrace{\|x_0 - \widehat{x}_0}_{+}\|^2 + (\underbrace{t_0 - \widehat{t}_0}_{\leq 0})(\underbrace{f(x_0) + \lambda - \widehat{t}_0}_{+}) \leq 0. \qquad (3)$$

Luego, considerando λ suficientemente grande, se infiere que $t_0 \leq \widehat{t}_0$. Veamos ahora que

$$t_0 < \widehat{t}_0 \quad \wedge \quad f(\widehat{x}_0) = \widehat{t}_0. \qquad (4)$$

$(\hat{x}_0, \hat{t}_0) \in \text{epi} f$

$f(\hat{x}_0) \leq \hat{t}_0$

$$\langle x_0 - \hat{x}_0, x - \hat{x}_0 \rangle + (t_0 - \hat{t}_0)(t - \hat{t}) \leq 0$$

$f(x)$ $f(\hat{x}_0)$

cont...

Aceptando (4), consideremos cualquier $x \in \mathrm{dom}(f)$ y tomemos $t = f(x)$ en (2) y usando la igualdad en (4), obtenemos

$$\langle x_0 - \widehat{x}_0, x - \widehat{x}_0 \rangle + (t_0 - \widehat{t}_0)(f(x) - f(\widehat{x}_0)) \leq 0.$$

Por la desigualdad estricta en (4), dividimos la desigualdad anterior por $t_0 - \widehat{t}_0 < 0$ y reordenando, se obtiene

$t_0 < \hat{t}_0$

$$f(x) \geq f(\widehat{x}_0) + \langle x - \widehat{x}_0, \underbrace{\frac{x_0 - \widehat{x}_0}{\widehat{t}_0 - t_0}}_{x_1} \rangle \quad \forall x \in \mathrm{dom}(f).$$

cont...

Verifiquemos (4):

- Si $t_0 = \widehat{t}_0$, reemplazando en (3), se obtiene que $x_0 = \widehat{x}_0$. Luego, $(x_0, t_0) = (\widehat{x}_0, \widehat{t}_0) \in \mathrm{epi}(f)$, lo cual es una contradicción.
- Si $f(\widehat{x}_0) < \widehat{t}_0$. Tomando $x = \widehat{x}_0$ y $t = f(\widehat{x}_0)$ en (2), obtenemos $(t_0 - \widehat{t}_0)(f(\widehat{x}_0) - \widehat{t}_0) \leq 0$, lo cual es una contradicción.

Corolario 1

$g(x) \geq g(\hat{x}_0) + \langle x - \hat{x}_0, x_1 \rangle$

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa, sci y propia. Si f es fuertemente convexa con coeficiente μ, entonces f es 0-coerciva.

Demostración

Luego, $g(x) := f(x) - \frac{\mu}{2}\|x\|^2$ es convexa. Es fácil ver que es propia y como la suma de sci's es sci, entonces g es convexa, sci y propia. Aplicando el Teorema 1 para g, se obtiene

$$f(x) \geq f(\widehat{x}_0) + \frac{\mu}{2}\left[\|x\|^2 - \|\widehat{x}_0\|^2\right] + \langle x - \widehat{x}_0, x_1 \rangle,$$

$\langle x, x_1 \rangle - \langle \hat{x}_0, x_1 \rangle$

$\geq -\|x\| \|x_1\|$

por la desigualdad de Cauchy-Schwarz,

$$f(x) \geq f(\widehat{x}_0) + \frac{\mu}{2}\left[\|x\|^2 - \|\widehat{x}_0\|^2\right] - \|x\|\|x_1\| - \langle \widehat{x}_0, x_1 \rangle, \tag{5}$$

lo cual tiende al infinito cuando $\|x\| \to +\infty$.

Corolario 2

Sea $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ convexa, sci y propia y $x_0 \in \mathbb{R}^n$. Entonces $f(x) + \frac{\mu}{2}\|x - x_0\|^2$ es 0-coerciva.

Demostración

Como $\frac{\mu}{2}\|x - x_0\|^2$ es convexa y sci, entonces $h(x) = f(x) + \frac{\mu}{2}\|x - x_0\|^2$ es convexa, sci y además propia. Veamos que h es fuertemente convexa. Como

$$h(x) - \frac{\mu}{2}\|x\|^2 = f(x) - \mu\langle x, x_0\rangle + \frac{\mu}{2}\|x_0\|^2,$$

es la suma de funciones convexa, luego h es fuertemente convexa y se concluye por el Corolario 1

Observación 3

Bajo las mismas condiciones del Corolario 1, pasando a dividir por $\|x\|$ en (5), se obtiene

$$\frac{f(x)}{\|x\|} \geq \frac{f(\widehat{x}_0)}{\|x\|} + \frac{\mu}{2}\|x\| - \frac{\mu}{2}\frac{\|\widehat{x}_0\|^2}{\|x\|} - \|x_1\| - \frac{\langle \widehat{x}_0, x_1 \rangle}{\|x\|},$$

luego tomando $\|x\| \to +\infty$, se concluye que f es 1-coerciva.

Referencias bibliográficas

1. Jean-Baptiste Hiriart-Urruty and Claude Lemarechal. Fundamentals of Convex Analysis, Springer, 1st ed. 2001.
2. Jean-Pierre Crouzeix and Eladio Ocaña. Análisis convexo. Fondo Editorial EDUNI, 2018.
3. https://sites.math.washington.edu/~burke/crs/408/notes/nlp/unoc.pdf
4. http://poisson.phc.dm.unipi.it/~fpmaiale/notes/OTM.pdf

FIN