Profesor: Fernando Roldán
Ayudantes: Cristofer Mamani y Fernando Muñoz
Departamento de Ingeniería Matemática
Facultad de Ciencias Físicas y Matemáticas
Universidad de Concepción

Optimización II (5225565)

Tarea 2

Fecha de entrega: Lunes 06 de Octubre de manera individual.

P1) Decimos que $f\colon \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ es *fuertemente convexa* con constante $\mu > 0$ si

$$(\forall \lambda \in [0,1]) \quad f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda) f(y) - \frac{\mu}{2}\lambda(1-\lambda)\|x-y\|^2.$$

Sea $f\colon \mathbb{R}^n \to \mathbb{R}\cup\{+\infty\}$ y $\mu > 0$.

(i) Demuestre que si f es μ-fuertemente convexa entonces es estricatmente convexa y por lo tanto convexa.

(ii) Demuestre que f es μ-fuertemente convexa si y solo si $(f - \frac{\mu}{2}\|\cdot\|^2)$ es convexa.

(iii) Demuestre que si f es propia, μ-fuertemente convexa y semicontinua inferior entonces f es coerciva. **Hint:** Utilice una función afín $g\colon \mathbb{R}^n \to \mathbb{R}$ tal que $f - \frac{\mu}{2}\|x\|^2 \geq g$ (justifique su existencia).

(iv) Demuestre que si f es propia, μ-fuertemente convexa y semicontinua inferior, entonces $\arg\min_{\mathbb{R}^n} f$ es no vacío. Más aún, muestre que f tiene un único mínimo.

Solución

(i) Debido a que $\mu > 0$ y $\lambda \in]0,1[$, para $x \neq y$ se tiene que

$$\begin{aligned} f\big(\lambda x + (1-\lambda)y\big) &\leq \lambda f(x) + (1-\lambda) f(y) - \frac{\mu}{2}\lambda(1-\lambda)\|x-y\|^2 \\ &< \lambda f(x) + (1-\lambda) f(y) \\ &\leq \lambda f(x) + (1-\lambda) f(y). \end{aligned}$$

(ii) Notemos que $f - \frac{\mu}{2}\|\cdot\|^2$ es convexa si y solo si

$$\begin{aligned} & f\big(\lambda x + (1-\lambda)y\big) - \frac{\mu}{2}\|\lambda x + (1-\lambda)y\|^2 \leq \lambda\left(f(x) - \frac{\mu}{2}\|x\|^2\right) + (1-\lambda)\left(f(y) - \frac{\mu}{2}\|y\|^2\right) \\ \Leftrightarrow\ & f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda) f(y) + \frac{u}{2}\left(\|\lambda x + (1-\lambda)y\|^2 - \lambda\|x\|^2 - (1-\lambda)\|y\|^2\right). \end{aligned}$$

Adicionalmente

$$\begin{aligned} \|\lambda x + (1-\lambda)y\|^2 - \lambda\|x\|^2 - (1-\lambda)\|y\|^2 &= \lambda^2\|x\|^2 - 2\lambda(1-\lambda)x^\top y + (1-\lambda)^2\|y\|^2 - \lambda\|x\|^2 - (1-\lambda)\|y\|^2 \\ &= \lambda(\lambda-1)\|x\|^2 - 2\lambda(1-\lambda)x^\top y + (1-\lambda)\lambda\|y\|^2 \\ &= \lambda(1-\lambda)\|x-y\|^2, \end{aligned}$$

de donde se sigue el resultado.

(iii) Debido a que f propia, μ-fuertemente convexa y semicontinua inferior, se sigue que $f - \frac{\mu}{2}\|\cdot\|^2$ es propia, convexa y semicontinua inferior ($\|\cdot\|$ tiene dominio $\mathbb{R}^n$ y es continua). Por lo tanto, existe una función afín $g\colon \mathbb{R}^n \to \mathbb{R}\colon x \mapsto a^\top x + \alpha$ para $a \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ tal que $f - \frac{\mu}{2}\|\cdot\|^2 \geq g$. Esto último debido a que esta clase de funciones son el supremo de sus minorantes afines (Teorema visto en clases). Ahora, para todo $x \in \mathbb{R}^n$,

$$f(x) - \frac{\mu}{2}\|x\|^2 \geq a^\top x + \alpha \Leftrightarrow f(x) \geq \frac{\mu}{2}\|x - a/\mu\|^2 - \frac{1}{2\mu}\|a\|^2 + \alpha,$$

de donde se concluye que f es coerciva.

(iv) Como f es propia, semicontinua inferior y coerciva, el resultado de sigue por el teorema de Weirstrass. La unicidad viene de que f es estrictamente convexa.

P2) Sea $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ una función propia y convexa. Suponga que f es localmente acotada en $x_0 \in \mathbb{R}^n$, es decir, existe $\delta > 0$ y $(m, M) \in \mathbb{R}^2$ tal que para cada $x \in B(x_0, \delta)$ se cumple que

$$m \leq f(x) \leq M.$$

Demuestre que f es Lipschitz continua en $B(x_0, \delta/2)$ con constante $\frac{2(M-m)}{\delta}$, esto es:

$$x_2, x_1 \in B(x_0, \delta/2) \Rightarrow |f(x_2) - f(x_1)| \leq \frac{2(M-m)}{\delta}\|x_1 - x_2\|.$$

Hint: Considere el vector $y = x_2 + \frac{\delta}{2}\frac{x_2 - x_1}{\|x_2 - x_1\|}$.

Solución Sean $x_1, x_2 \in B(x_0, \delta)$, con $x_1 \neq x_2$ y definamos

$$y := \left(1 + \frac{\delta}{2\left\|x_1 - x_2\right\|}\right)x_2 - \frac{\delta}{2\left\|x_1 - x_2\right\|}x_1 = x_2 + \frac{\delta(x_2 - x_1)}{2\left\|x_2 - x_1\right\|}.$$

Notemos que $\|y - x_2\| = \delta/2 \leq \delta$, es decir $y \in B\left(x_0, \delta\right)$. Reordenando términos

$$x_2 = \frac{2\left\|x_1 - x_2\right\|}{\delta + 2\left\|x_1 - x_2\right\|}y + \frac{\delta}{\delta + 2\left\|x_1 - x_2\right\|}x_1$$

vemos que x_2 es una combinación convexa de x_1 y y. Ya que f es convexa

$$f\left(x_2\right) \leq \frac{2\left\|x_1 - x_2\right\|}{\delta + 2\left\|x_1 - x_2\right\|}f(y) + \frac{\delta}{\delta + 2\left\|x_1 - x_2\right\|}f\left(x_1\right).$$

Lo que implica que

$$\begin{aligned} f\left(x_2\right) - f\left(x_1\right) &\leq \frac{2\left\|x_1 - x_2\right\|}{\delta + 2\left\|x_1 - x_2\right\|}\left(f(y) - f\left(x_1\right)\right) \\ &\leq \frac{2(M-m)}{\delta}\left\|x_1 - x_2\right\| \end{aligned}$$

El resultado se sigue intercambiando los roles de x_1 y x_2.

P3) Sea $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$ una función Gâteaux diferenciable[1] y D un conjunto abierto.

(i) Sea x un mínimo de f en D. Pruebe que $\nabla f(x) = 0$.

[1]Recuerde que la derivada direccional de f en $x \in \mathbb{R}^n$ con dirección $d \in \mathbb{R}^n$ está definida por

$$f'(x; v) = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}.$$

Además, decimos que f es diferenciable en x en el sentido de Gâteaux, Si $f'(x, d)$ existe para cada $d \in \mathbb{R}^n$ y la función $d \mapsto f'(x, d)$ es lineal y continua. Si f es Gâteaux diferenciable, se tiene que $f'(x; v) = v^\top \nabla f(x)$ para todo $v \in \mathbb{R}^n$

(ii) Pruebe que si $D = \mathbb{R}^n$ y f es convexa, $x \in \mathbb{R}^n$ es un mínimo de f si y sólo si, $\nabla f(x) = 0$.

(iii) Proporcione un ejemplo donde $\nabla f(x) = 0$ pero x no sea un mínimo global.

Solución:

(i) Dado que x es un mínimo local de f, para todo $v \in \mathbb{R}^n$ y $t \in]0, +\infty[$ pequeño, se tiene que

$$\frac{f(x+tv) - f(x)}{t} \geq 0.$$

Tomando $t \downarrow 0$ se tiene que, para todo $v \in \mathbb{R}^n$

$$\nabla^\top f(x) \cdot v = f'(x, v) \geq 0.$$

Tomando $v \neq 0$ y $-v$ en esta desigualdad, se concluye que $\nabla f(x) = 0$.

(ii) Si f es convexa se tiene que

$$f(y) \geq f(x) + (\nabla f(x))^\top (y - x).$$

Por lo tanto, si $\nabla f(x) = 0$ se tiene que $f(y) \geq f(x)$, para todo $y \in \mathbb{R}^n$, por lo tanto x es un mínimo global. La otra dirección es directa del punto anterior.

(iii) Consideremos $f\colon \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^4 - 2x^2$. Se tiene que $f'(0) = 0$ pero $f(0) = 0$ mientras que el valor mínimo de f es $f(1) = f(-1) = -1$.

P4) Verifique que la función $f\colon \mathbb{R}^n \to \mathbb{R}$ dada por $f(x, y, z) = x^2 + y^2 + 2xy + (z+1)^2 + 2e^{-z}$ es convexa. Justifique la existencia de mínimos de f y determinelos.

Observemos primero que, f se puede escribir como $f(x, y, z) = (x+y)^2 + (z+1)^2 + 2e^{-z}$. Además, el gradiente de f es $\nabla f(x, y, z) = (2x + 2y, 2x + 2y, 2(z+1) - 2e^{-z})$. Calculando el hessiano de f obtenemos la matriz

$$\nabla^2 f(x, y, z) = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 2 + 2e^{-z} \end{pmatrix}.$$

Para verificar que $\nabla^2 f(x, y, z)$ es semidefinida positiva basta evaluar la forma cuadrática asociada a un vector $(u, v, w) \in \mathbb{R}^3$:

$$(u, v, w)^\top \nabla^2 f(x, y, z)(u, v, w) = 2u^2 + 2v^2 + 4uv + (2 + 2e^{-z})w^2 = 2(u+v)^2 + (2 + 2e^{-z})w^2 \geq 0.$$

Así, $\nabla^2 f(x, y, z)$ es semidefinida positiva para todo (x, y, z), y por el criterio del hessiano f es convexa en $\mathbb{R}^3$. Adicionalmente f es continua y coerciva ya que

$$\lim_{\|(x,y,z)\| \to +\infty} f(x, y, z) = +\infty.$$

Por el teorema de Weierstrass se sigue que f tiene mínimos. Por lo visto en la pregunta anteriores, los mínimos están caracterizados por $\nabla f(x, y, z) = 0$ que en este caso es en el conjunto

$$\{(x, y, z) \in \mathbb{R}^3 : \ x + y = 0, \ z = 0\}.$$

Finalmente, el valor mínimo de f es

$$f(x, -x, 0) = 3.$$