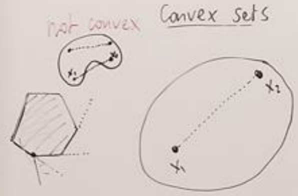
Pick up Midterm Exam: 1 – 2pm

Introduction to Convex Optimization

Convex sets



• intersection of convex sets is convex

$\alpha x_1 + (1-\alpha) x_2$ with $\alpha \in [0,1]$

convex functions

$f: C \to \mathbb{R}$ (domain of f; C convex)

pure definition: $f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$

Properties and results for convex functions

Properties

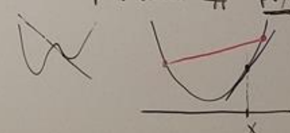
• For a convex function, any local minimizer of f is a global minimizer. (P1)

• If $f \in \mathcal{F}^1(C)$, then any critical point of f is a global minimizer. (P2)

Results about the characterization of convex fcn's

(R1) • Let $f \in C^1(C, \mathbb{R})$ with C convex set. Then:

- f is convex iff $f(y) \geq f(x) + (\nabla f)(x)^T (y-x)$ $\forall x, y \in C$.



(R2) • $f \in C^2$ is convex on an open set $C \subset \mathbb{R}^n$ (convex set)

$\Leftrightarrow (Hf)(x) \succeq 0 \quad \forall x \in C$.

Proof of Property P2:

Let x^* be a critical point. Since $f \in \mathcal{F}^1(C)$, we have from (R1):

$$f(y) \geq f(x^*) + \underbrace{\underbrace{(\nabla f)(x^*)^T}_{=0} (y - x^*)}_{=0} \qquad \forall y \in C$$

$$\Rightarrow \underbrace{\forall y \in C \quad f(y) \geq f(x^*)}$$

x^* is a global minimizer of f.

Proof of Result (R2):

"$\Rightarrow$": Let $f \in \mathcal{F}^2(C)$. Since C is an open set, then for any $x \in C$ there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset C$

Consider an arbitrary $z \in B_\varepsilon(x)$ and consider the Taylor expansion

$$f(z) = \underline{f(x) + (\nabla f)(x)^T (z-x)} + \frac{1}{2}(z-x)^T (Hf)(x)(z-x) + o(\|z-x\|^2).$$

From convexity of f on C, we have (R1): $f(z) \geq f(x) + (\nabla f)(x)^T (z-x)$, which means that

$$\frac{1}{2}(z-x)^T (Hf)(x)(z-x) + o(\|z-x\|^2) \geq 0.$$

By dividing by $\|z-x\|^2$ and defining $s := \frac{z-x}{\|z-x\|}$, we obtain: $\frac{1}{2} s^T (Hf)(x) s + \frac{o(\|z-x\|^2)}{\|z-x\|^2} \geq 0$

By taking $z \to x$: $s^T (Hf)(x) s \geq 0 \quad \forall s \in \mathbb{R}^n$ with $\|s\| = 1$

$(Hf)(x) \succeq 0$.

"$\Leftarrow$": Since C is convex, $\forall x, y \in C$ the line segment $[x_1, x_2]$ lies fully in C.

This implies that there is $z \in [x, y] \subset C$ such that

$$f(y) = f(x) + (\nabla f)(x)^T (y - x) + \underbrace{\frac{1}{2}(y-x)^T \underbrace{(Hf)(z)}_{(Hf)(z) \succeq 0 \text{ since } z \text{ is a point in } C} (y-x)}_{\geq 0}.$$

Extended Law of the Mean

$$\Rightarrow f(y) \geq f(x) + (\nabla f)(x)^T (y - x)$$

(R1) characterization of convex functions! ■

Theorem: If $(Hf)(x) \succ 0$ for all $x \in C$, *then* f is strictly convex. $\leftarrow$ *But*: strict convexity generally doesn't imply $(Hf)(x) \succ 0$. *Counterexample*: $f(x) = x^4$ but $f''(x) = 12x^2$