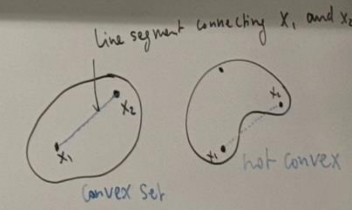
Pickup Midterm Exam: 2:15 – 3:15 pm, Green Hall 1120A

Convex Optimization: Convex Sets / Functions

Def. Convex Set: $C \subset \mathbb{R}^n$ is convex if

$$\forall x_1, x_2 \in C \ \forall \alpha \in [0,1] \quad x = \underbrace{\alpha x_1 + (1-\alpha) x_2}_{\text{Convex Combination}} \overset{!}{\in} C.$$



Def: A convex combination of points $x_1, x_2, \ldots, x_M \in \mathbb{R}^n$ is a linear combination

$$\sum_{i=1}^{M} \alpha_i x_i$$

where $\alpha_1 + \alpha_2 + \ldots + \alpha_M = 1$, $\alpha_i \geqslant 0$.

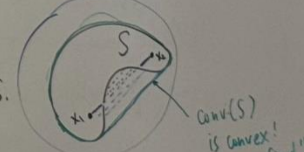
Theorem: Let S be an arbitrary set in $\mathbb{R}^n$. The following two definitions of the convex hull of S are:

(a): The set of all convex combinations of points in S.

(b): The intersection of all convex sets containing S

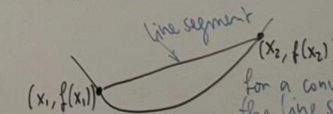
Convex Hull is the smallest convex set that contains S.

Conv(S).



Convex Functions: $f: C \longrightarrow \mathbb{R}$

Function — convex set



(a) f is convex if (by def.)

$$f(\alpha x_1 + (1-\alpha) x_2) \leqslant \alpha f(x_1) + (1-\alpha) f(x_2) \quad \forall x_1, x_2 \in C \quad \forall \alpha \in [0,1]$$

(b) f is strictly convex if (by def.)

$$f(\alpha x_1 + (1-\alpha) x_2) < \alpha f(x_1) + (1-\alpha) f(x_2) \quad \forall x_1, x_2 \in C,\ x_1 \neq x_2 \quad \forall \alpha \in]0,1[\ / \ \alpha \in (0,1)$$

Theorem: Suppose f is a convex function on a convex set $C \subset \mathbb{R}^n$. For any $x_1, \ldots, x_M \in C$ and $\alpha_1, \ldots, \alpha_M \geqslant 0$, $\sum_{i=1}^{M} \alpha_i = 1$, we have

$$f\left(\sum_{i=1}^{M} \alpha_i x_i\right) \leqslant \sum_{i=1}^{M} \alpha_i f(x_i).$$

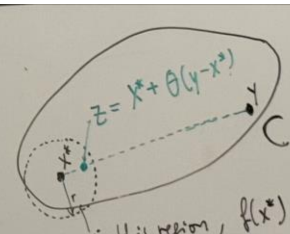
If f is strictly convex on C and $\alpha_i > 0$, then equality holds if and only if $x_1 = \ldots = \ldots = x_M$.

Theorem Let f be a convex function defined on a convex set $C \subset \mathbb{R}^n$. Then, any local minimizer of f is also a global minimizer. If in addition, f is strictly convex, then global minimizer is unique

probably the most important result about convex functions

Proof: Suppose that $x^* \in C$ is a local minimizer. Then

$\exists r>0$ s.t. $f(x) \geqslant f(x^*)$ $\quad \forall x \in B_r(x^*) \cap C$



Our goal is to show that $f(y) \geqslant f(x^*)$ $\forall y \in C$.

To this end, consider $\underbrace{z = x^* + \theta(y-x^*)}_{\substack{x^* + \theta y - \theta x^* \\ = \theta y + (1-\theta)x^*}}$ where $\theta \in (0,1)$ is small enough to have $z \in B_r(x^*) \cap C$

Then $f(x^*) \leq f(z) = f(x^* + \theta(y-x^*)) = f(\theta y + (1-\theta)x^*) \underset{\uparrow}{\leq} \theta f(y) + (1-\theta) f(x^*)$

convexity of the function f

Rearranging the inequality: $\underbrace{\theta f(y) + (1-\theta) f(x^*) - f(x^*) \geqslant 0}_{\theta f(y) - \theta f(x^*) \geqslant 0} \Rightarrow f(y) - f(x^*) \geqslant 0 \Leftrightarrow f(y) \geqslant f(x^*)$.

Notice that the inequality is strict if $y = x^*$ and f is strictly convex. . . . — ☒

Def: $\mathcal{F}^k(C)$ . . . the class of functions that are both convex and C^k

C: convex set

$\mathcal{F}^k_L(C)$ . . . ——— " ——— ——— " ——— and Lipschitz continuous derivatives.

Theorem: Let $f \in C^1$ on a convex set. The f is:

- Convex if and only if $f(y) \geqslant f(x) + (\nabla f)^T(y-x)$ $\forall x,y \in C$
- Strictly convex iff $f(y) > f(x) + (\nabla f)^T(y-x)$ $\forall x,y \in C$, $x \neq y$

Super important inequalities in optimization. They are often used as alternative def's of convexity! (for diff'ble functions)

Proof: "$\Rightarrow$": Suppose $f \in \mathcal{F}^1(C)$ and consider $x, y \in C$ and $\theta \in (0,1]$. Then

$$f(x+\theta(y-x)) = f(\theta y + (1-\theta)x) \leq \theta f(y) + (1-\theta) f(x) \Rightarrow \frac{f(x+\theta(y-x)) - f(x)}{\theta} \leq f(y) - f(x)$$

By taking the limit $\theta \to 0$, we obtain the directional derivative of f along $u := y-x$: $(\nabla f(x))^T(y-x) \leq f(y) - f(x)$

"$\Leftarrow$": Suppose that $f(y) \geqslant f(x) + (\nabla f(x))^T(y-x)$ $\forall x,y \in C$ $\quad (*)$

Let $z = \theta u + (1-\theta)v$ for arbitrary $u, v \in C$, $\theta \in [0,1]$. Using $(*)$ at u and v, we obtain

$$f(u) \geqslant f(z) + (\nabla f(z))^T(u-z) \quad (1)$$

$$f(v) \geqslant f(z) + (\nabla f(z))^T(v-z). \quad (2)$$

Linear combination of inequalities (1), (2): $\theta \times (1) + (1-\theta) \times (2)$

$$\begin{aligned}
\theta f(u) + (1-\theta) f(v) &\geqslant f(z) + (\nabla f(z))^T(\theta(u-z) + (1-\theta)(v-z)) \\
&= f(z) + (\nabla f(z))^T((\theta u + (1-\theta)v) - z) \\
&= f(z) \\
&= f(\theta u + (1-\theta)v)
\end{aligned}$$

original convexity definition. ✓