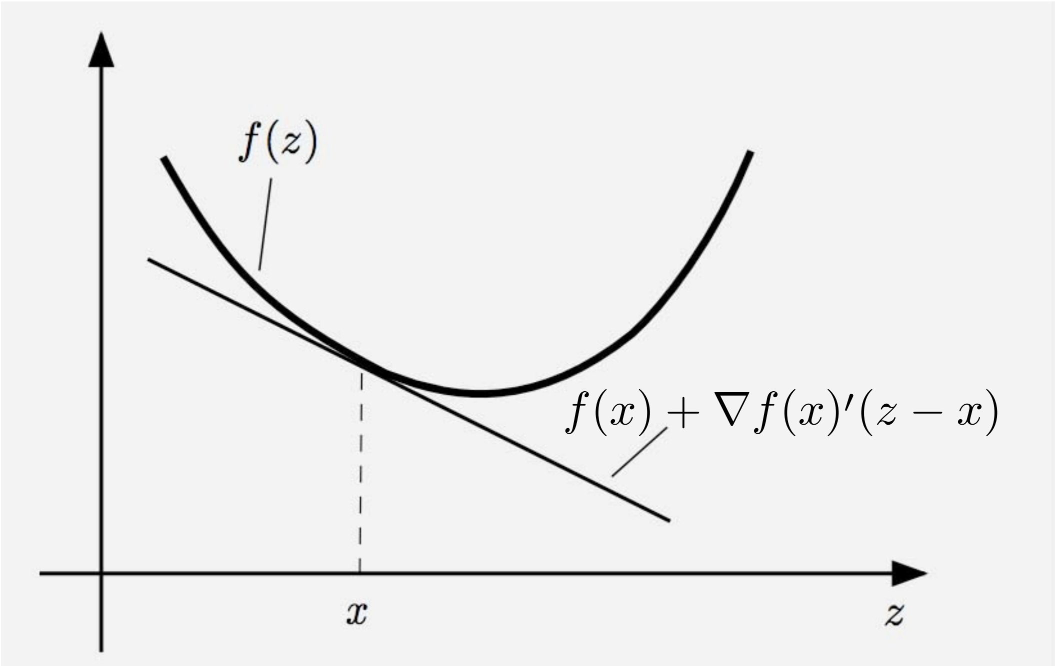
# LECTURE 3

# LECTURE OUTLINE

- Differentiable Convex Functions
- Convex and Affine Hulls
- Caratheodory's Theorem

**Reading:** Sections 1.1, 1.2

# DIFFERENTIABLE CONVEX FUNCTIONS



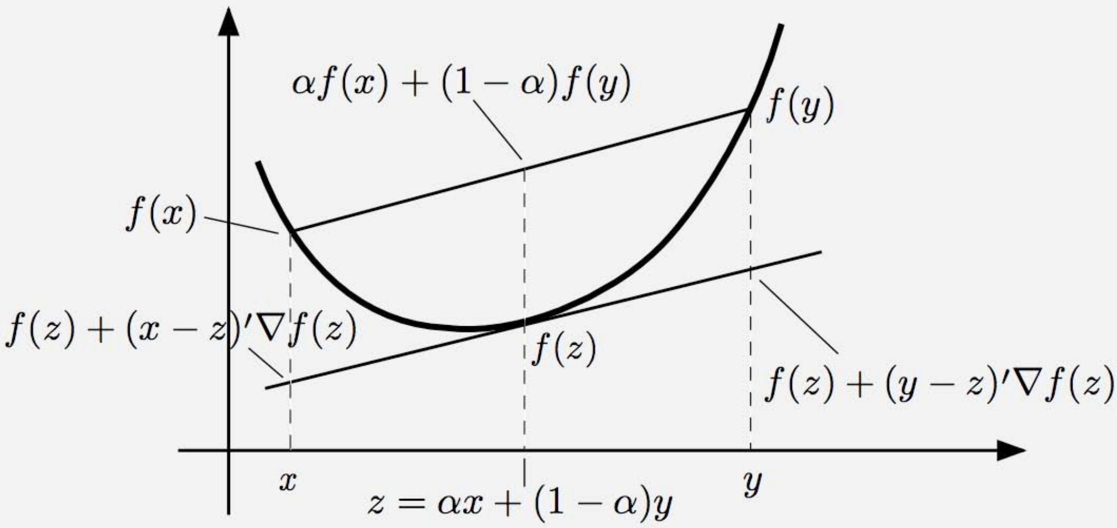

- Let $C \subset \Re^n$ be a convex set and let $f : \Re^n \mapsto \Re$ be differentiable over $\Re^n$.

  (a) The function $f$ is convex over $C$ iff

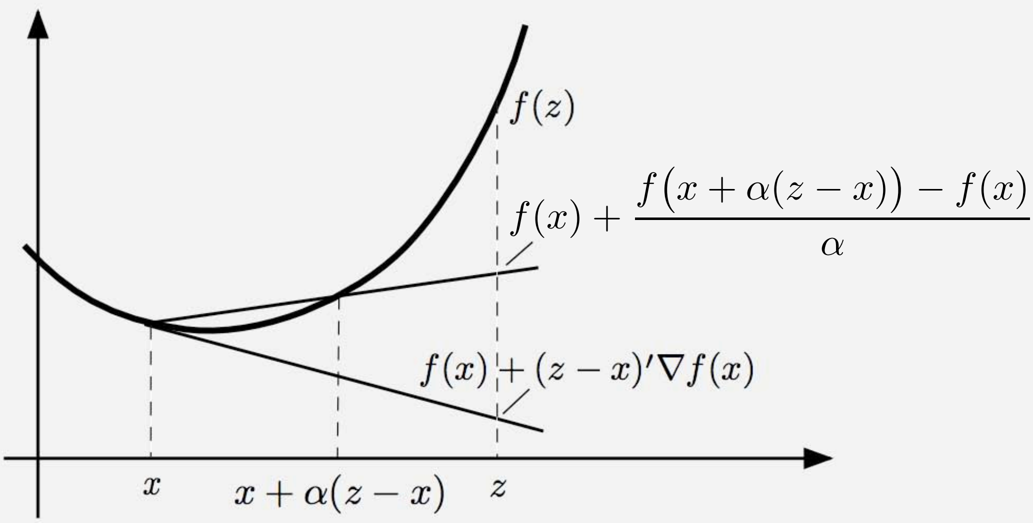
  $$f(z) \geq f(x) + (z - x)'\nabla f(x), \qquad \forall\, x, z \in C$$

  (b) If the inequality is strict whenever $x \neq z$, then $f$ is strictly convex over $C$.

# PROOF IDEAS




(a)




(b)

# OPTIMALITY CONDITION

• Let $C$ be a nonempty convex subset of $\Re^n$ and let $f : \Re^n \mapsto \Re$ be convex and differentiable over an open set that contains $C$. Then a vector $x^* \in C$ minimizes $f$ over $C$ if and only if

$$\nabla f(x^*)'(x - x^*) \geq 0, \qquad \forall\, x \in C.$$

**Proof:** If the condition holds, then

$$f(x) \geq f(x^*) + (x - x^*)'\nabla f(x^*) \geq f(x^*), \quad \forall\, x \in C,$$

so $x^*$ minimizes $f$ over $C$.

Converse: Assume the contrary, i.e., $x^*$ minimizes $f$ over $C$ and $\nabla f(x^*)'(x - x^*) < 0$ for some $x \in C$. By differentiation, we have

$$\lim_{\alpha \downarrow 0} \frac{f\big(x^* + \alpha(x - x^*)\big) - f(x^*)}{\alpha} = \nabla f(x^*)'(x - x^*) < 0$$

so $f\big(x^* + \alpha(x - x^*)\big)$ decreases strictly for sufficiently small $\alpha > 0$, contradicting the optimality of $x^*$. **Q.E.D.**

# PROJECTION THEOREM

- Let $C$ be a nonempty closed convex set in $\Re^n$.

  (a) For every $z \in \Re^n$, there exists a unique minimum of
  $$f(x) = \|z - x\|^2$$
  over all $x \in C$ (called the *projection of $z$ on $C$*).

  (b) $x^*$ is the projection of $z$ if and only if
  $$(x - x^*)'(z - x^*) \le 0, \qquad \forall\, x \in C$$

**Proof:** (a) $f$ is strictly convex and has compact level sets.

(b) This is just the necessary and sufficient optimality condition
$$\nabla f(x^*)'(x - x^*) \ge 0, \qquad \forall\, x \in C.$$

# TWICE DIFFERENTIABLE CONVEX FNS

- Let $C$ be a convex subset of $\Re^n$ and let $f : \Re^n \mapsto \Re$ be twice continuously differentiable over $\Re^n$.

  (a) If $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$, then $f$ is convex over $C$.

  (b) If $\nabla^2 f(x)$ is positive definite for all $x \in C$, then $f$ is strictly convex over $C$.

  (c) If $C$ is open and $f$ is convex over $C$, then $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$.

**Proof:** (a) By mean value theorem, for $x, y \in C$

$$f(y) = f(x) + (y-x)'\nabla f(x) + \tfrac{1}{2}(y-x)'\nabla^2 f\big(x + \alpha(y-x)\big)(y-x)$$

for some $\alpha \in [0,1]$. Using the positive semidefiniteness of $\nabla^2 f$, we obtain

$$f(y) \geq f(x) + (y-x)'\nabla f(x), \qquad \forall\, x, y \in C$$

From the preceding result, $f$ is convex.

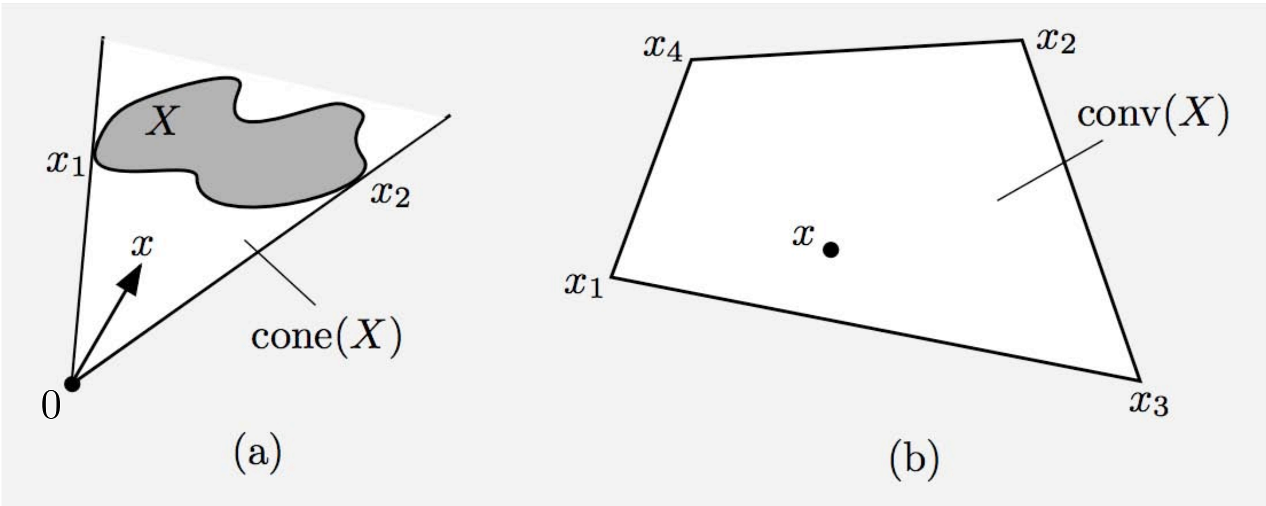
(b) Similar to (a), we have $f(y) > f(x) + (y-x)'\nabla f(x)$ for all $x, y \in C$ with $x \neq y$, and we use the preceding result.

(c) By contradiction ... similar.

# CONVEX AND AFFINE HULLS

- Given a set $X \subset \Re^n$:
- A *convex combination* of elements of $X$ is a vector of the form $\sum_{i=1}^{m} \alpha_i x_i$, where $x_i \in X$, $\alpha_i \geq 0$, and $\sum_{i=1}^{m} \alpha_i = 1$.
- The *convex hull* of $X$, denoted $\mathrm{conv}(X)$, is the intersection of all convex sets containing $X$. (Can be shown to be equal to the set of all convex combinations from $X$).
- The *affine hull* of $X$, denoted $\mathrm{aff}(X)$, is the intersection of all affine sets containing $X$ (an affine set is a set of the form $\overline{x} + S$, where $S$ is a subspace).
- A *nonnegative combination* of elements of $X$ is a vector of the form $\sum_{i=1}^{m} \alpha_i x_i$, where $x_i \in X$ and $\alpha_i \geq 0$ for all $i$.
- The *cone generated by* $X$, denoted $\mathrm{cone}(X)$, is the set of all nonnegative combinations from $X$:
  - It is a convex cone containing the origin.
  - It need not be closed!
  - If $X$ is a finite set, $\mathrm{cone}(X)$ is closed (nontrivial to show!)

# CARATHEODORY'S THEOREM




- Let $X$ be a nonempty subset of $\Re^n$.
  (a) Every $x \neq 0$ in cone($X$) can be represented as a positive combination of vectors $x_1, \ldots, x_m$ from $X$ that are linearly independent (so $m \leq n$).
  (b) Every $x \notin X$ that belongs to conv($X$) can be represented as a convex combination of vectors $x_1, \ldots, x_m$ from $X$ with $m \leq n + 1$.

# PROOF OF CARATHEODORY'S THEOREM

(a) Let $x$ be a nonzero vector in cone$(X)$, and let $m$ be the smallest integer such that $x$ has the form $\sum_{i=1}^{m} \alpha_i x_i$, where $\alpha_i > 0$ and $x_i \in X$ for all $i = 1, \ldots, m$. If the vectors $x_i$ were linearly dependent, there would exist $\lambda_1, \ldots, \lambda_m$, with

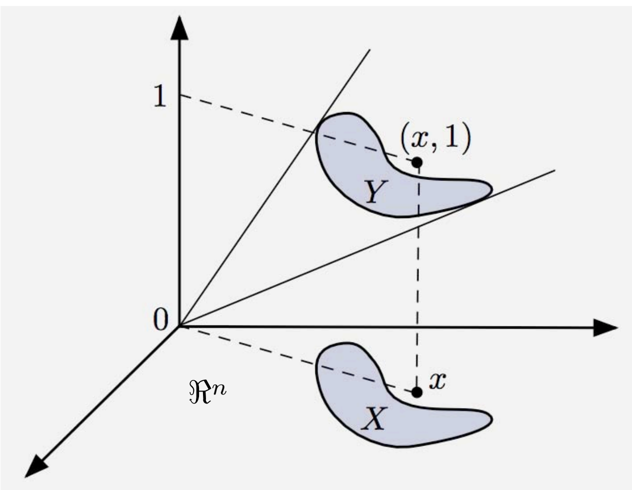
$$\sum_{i=1}^{m} \lambda_i x_i = 0$$

and at least one of the $\lambda_i$ is positive. Consider

$$\sum_{i=1}^{m} (\alpha_i - \overline{\gamma}\lambda_i) x_i,$$

where $\overline{\gamma}$ is the largest $\gamma$ such that $\alpha_i - \gamma\lambda_i \geq 0$ for all $i$. This combination provides a representation of $x$ as a positive combination of fewer than $m$ vectors of $X$ – a contradiction. Therefore, $x_1, \ldots, x_m$, are linearly independent.

(b) Use "lifting" argument: apply part (a) to $Y = \{(x, 1) \mid x \in X\}$.



# AN APPLICATION OF CARATHEODORY

- The convex hull of a compact set is compact.

**Proof:** Let $X$ be compact. We take a sequence in $\mathrm{conv}(X)$ and show that it has a convergent subsequence whose limit is in $\mathrm{conv}(X)$.

By Caratheodory, a sequence in $\mathrm{conv}(X)$ can be expressed as $\left\{\sum_{i=1}^{n+1} \alpha_i^k x_i^k\right\}$, where for all $k$ and $i$, $\alpha_i^k \geq 0$, $x_i^k \in X$, and $\sum_{i=1}^{n+1} \alpha_i^k = 1$. Since the sequence

$$\left\{(\alpha_1^k, \ldots, \alpha_{n+1}^k, x_1^k, \ldots, x_{n+1}^k)\right\}$$

is bounded, it has a limit point

$$\left\{(\alpha_1, \ldots, \alpha_{n+1}, x_1, \ldots, x_{n+1})\right\},$$

which must satisfy $\sum_{i=1}^{n+1} \alpha_i = 1$, and $\alpha_i \geq 0$, $x_i \in X$ for all $i$.

The vector $\sum_{i=1}^{n+1} \alpha_i x_i$ belongs to $\mathrm{conv}(X)$ and is a limit point of $\left\{\sum_{i=1}^{n+1} \alpha_i^k x_i^k\right\}$, showing that $\mathrm{conv}(X)$ is compact. **Q.E.D.**

- Note that the convex hull of a closed set need not be closed!