# CSE203B - Discussion Session

Po-Ya Hsu

01/29/21

# Outline

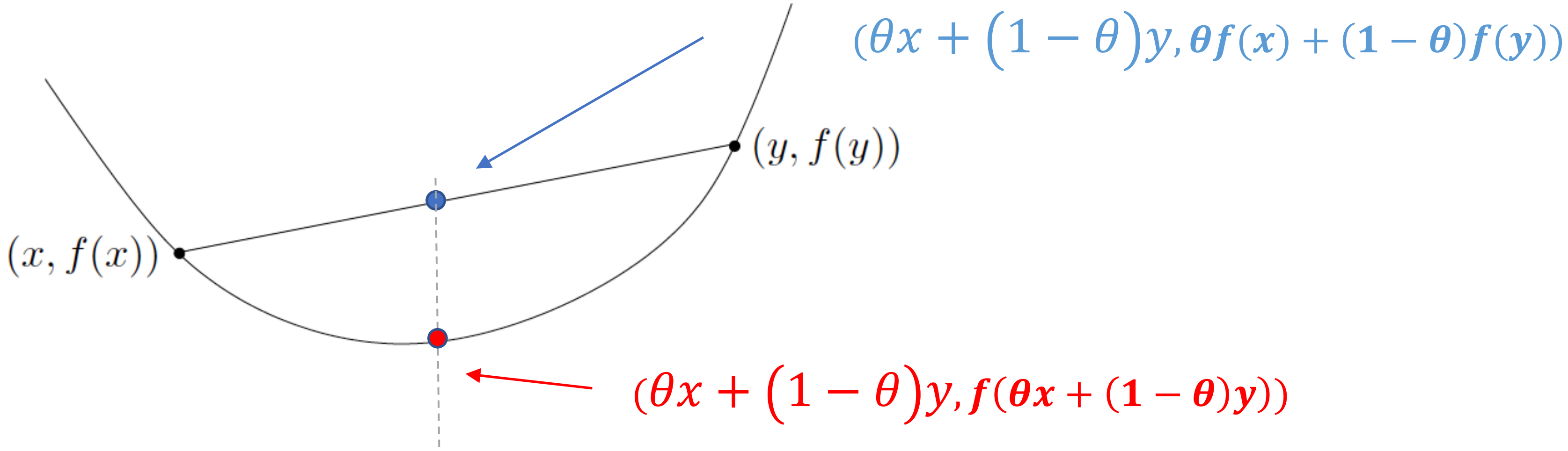
- Convex functions
  - Definition
  - First-order conditions
  - Second-order conditions
  - Operations that preserve convexity
- Conjugate functions
- Quasi-convex functions
- Assignment Hints

# Definition of Convex Functions

- A function $f: R^n \to R$ is convex if dom $f$ is a convex set and if for all $x, y \in \text{dom } f$, and $0 \le \theta \le 1$, we have

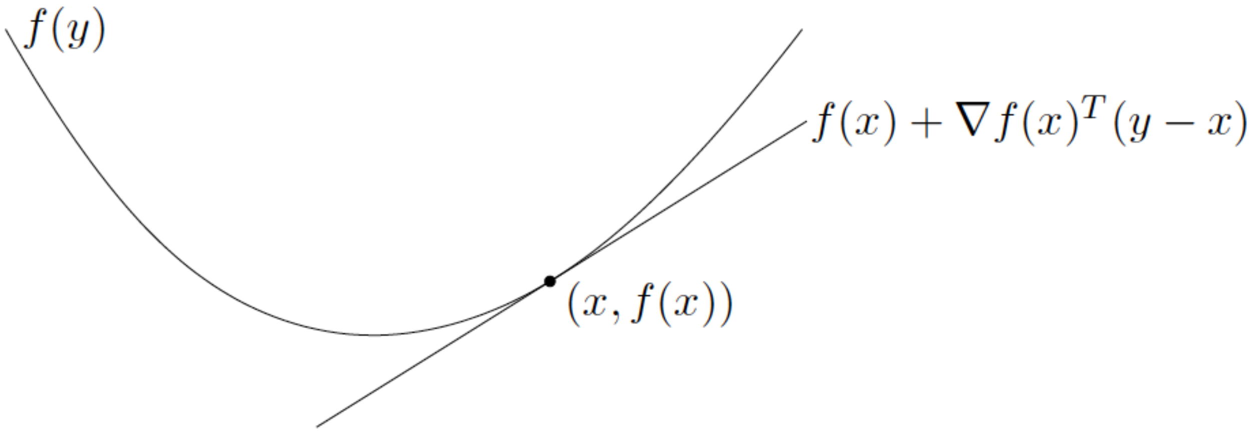
$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y)$$

- Concave functions: $-f$ is convex



# First-order Conditions

- Suppose $f$ is differentiable ($dom\ f$ is open and $\nabla f$ exists at each point in $dom\ f$), then $f$ is convex iff $\mathrm{dom}\ f$ is convex and for all $x, y \in dom\ f$
$$f(y) \geq f(x) + \nabla f(x)^T(y - x)$$
- Strict convexity: $f(y) > f(x) + \nabla f(x)^T(y - x), x \neq y$
- Concave functions: $f(y) \leq f(x) + \nabla f(x)^T(y - x)$



# Second Order Condition

- Suppose $f$ is twice differentiable ($dom\ f$ is open and its Hessian exists at each point in $dom\ f$), then $f$ is convex iff $\mathrm{dom}\ f$ is convex and for all $x, y \in dom\ f$

$$\nabla^2 f(x) \succcurlyeq 0 \text{ (the Hessian is positive semidefinite)}$$

- Strict convexity: $\nabla^2 f(x) \succ 0$
- Concave functions: $\nabla^2 f(x) \preccurlyeq 0$

# Example of Convex Functions

- Quadratic over linear function

$$f(x,y) = \frac{x^2}{y}, \text{for } y > 0$$

Its gradient $\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{2x}{y} \\ \frac{-x^2}{y^2} \end{bmatrix}$

Hessian $\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \frac{2}{y^3} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix} \succcurlyeq 0 \Rightarrow$ convex

*Positive semidefinite*? $\begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^{\mathrm{T}}$, for any $u \in R^2, u^T(vv^T)u = (v^T u)^T (v^T u) = \left||v^T u|\right|_2^2 \geq 0.$

# Epigraph

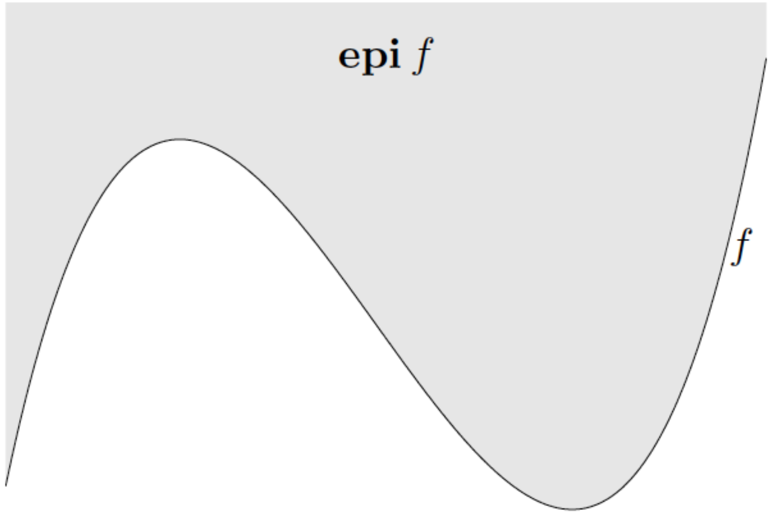
- $\alpha$*-sublevel* set of $f: R^n \rightarrow R$

$$C_\alpha = \{x \in dom\ f \mid f(x) \leq \alpha\}$$

  sublevel sets of a convex function are convex for any value of $\alpha$.

- Epigraph of $f: R^n \rightarrow R$ is defined as

$$\textbf{epi}\ f = \{(x,t) | x \in dom\ f, f(x) \leq t\} \subseteq \boldsymbol{R^{n+1}}$$



# Link between convex sets and convex functions

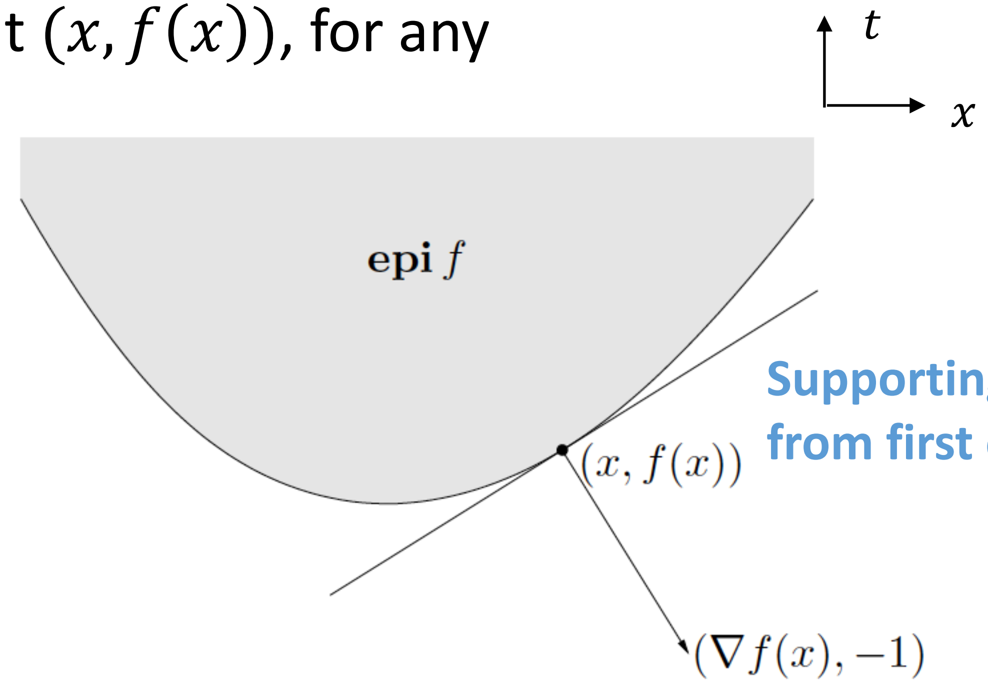
- A function is convex **iff** its epigraph is a convex set.
- Consider a convex function $f$ and $x, y \in dom\ f$

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

**epi *f*** First order condition for convexity

- The hyperplane supports **epi *f*** at $(x, f(x))$, for any

$$(y, t) \in \mathbf{epi}\ f \Rightarrow$$
$$\nabla f(x)^T(y - x) + f(x) - t \leq 0$$
$$\Rightarrow \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^T \left( \begin{bmatrix} y \\ t \end{bmatrix} - \begin{bmatrix} x \\ f(x) \end{bmatrix} \right) \leq 0$$




Supporting hyperplane, derived from first order condition

# Operations that preserve convexity

*Practical methods for establishing convexity of a function*

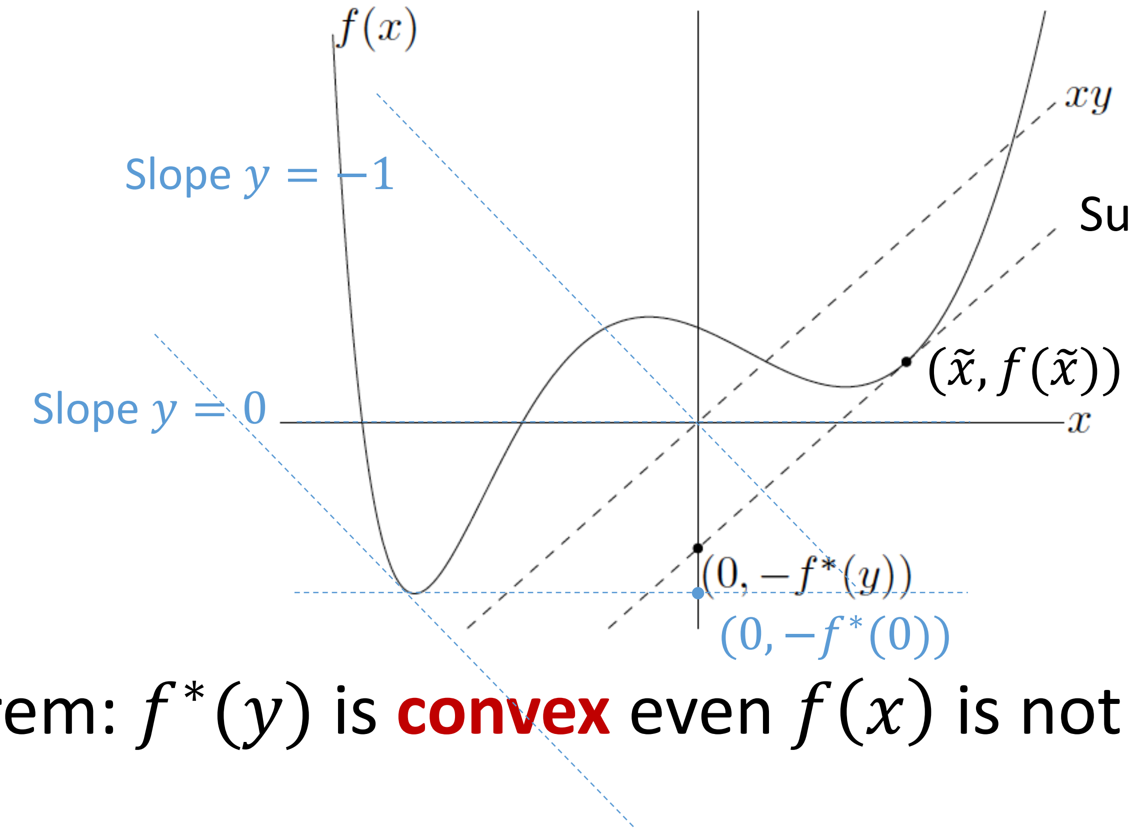
- Verify definition ( often simplified by restricting to a line)
- For twice differentiable functions, show $\nabla^2 f(x) \succcurlyeq 0$
- Show that $f$ is obtained from simple convex functions by operations that preserve convexity (Ref. Chap. 3.2)
  - Nonnegative weighted sum
  - Composition with affine function
  - Pointwise maximum and supremum
  - Composition
  - Minimization
  - Perspective

# Conjugate Function

- Given function $f: R^n \to R$ , the conjugate function

$$f^*(y) = \sup_{x \in dom\, f} y^T x - f(x)$$

  - The **dom** $f^*$ consists $y \in R^n$ for which $\sup_{x \in dom\, f} y^T x - f(x)$ is **bounded**




Supporting hyperplane: if $f$ is convex in that domain

$(\tilde{x}, f(\tilde{x}))$ where $\nabla f^T(\tilde{x}) = y$ if $f$ is differentiable

Theorem: $f^*(y)$ is **convex** even $f(x)$ is not convex.

***(Proof with pointwise supremum)***

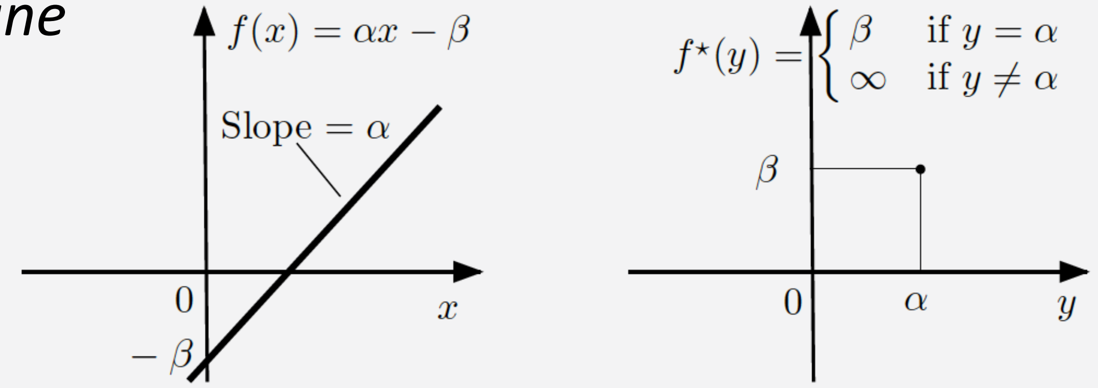
***$y^T x - f(x)$ is affine function in $y$***

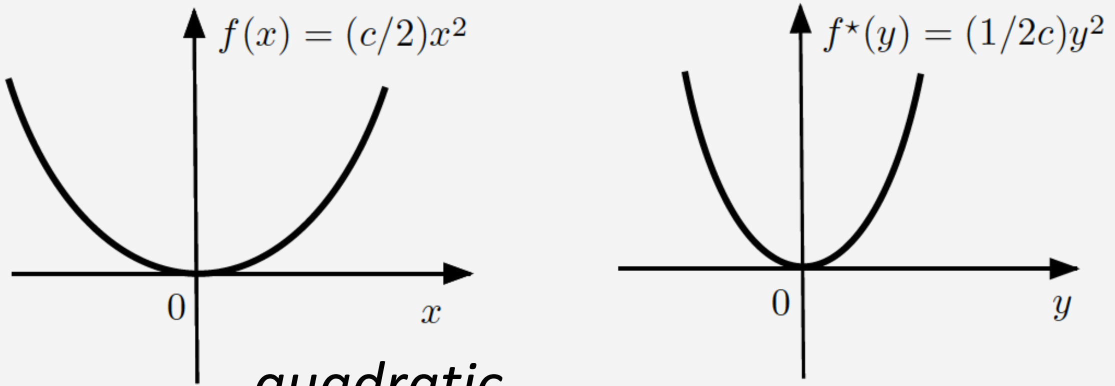
# Examples of Conjugates

- Derive the conjugates of $f: R \to R$

$$f^*(y) = \sup_{x \in dom\, f} \; yx - f(x)$$

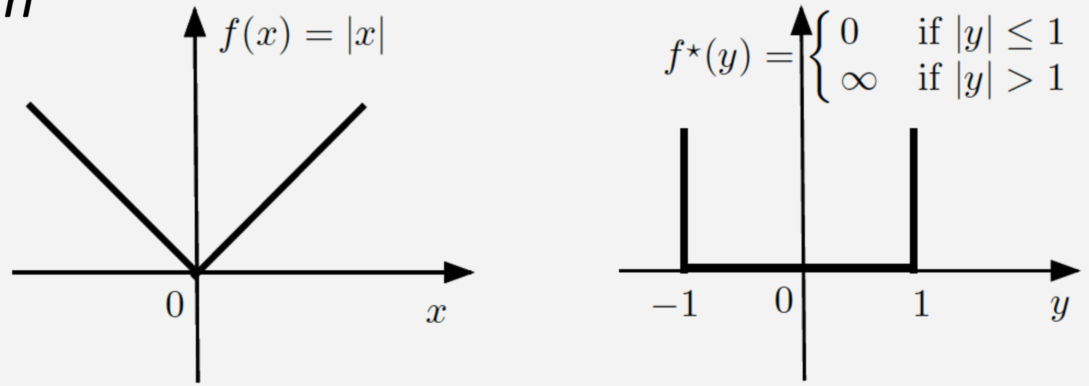
*Affine*




*quadratic*

*Norm*




***See more examples in chap 3.3.1***

# Quasi-convex Functions

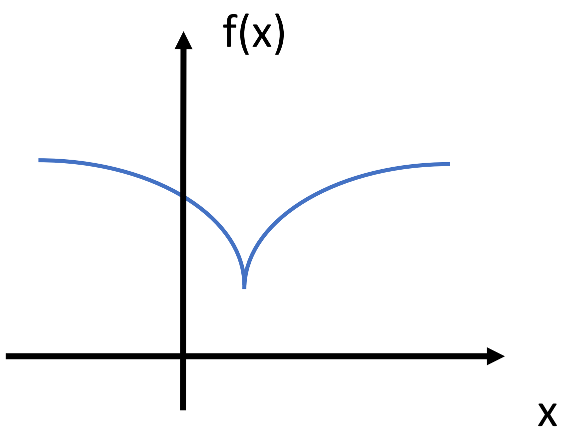
- A function $f: R^n \rightarrow R$ is quasi-convex if its domain and all its sublevel sets

$$S_\alpha = \{x \epsilon \; dom \; f \mid f(x) \leq \alpha\}, \alpha \epsilon \mathcal{R}$$

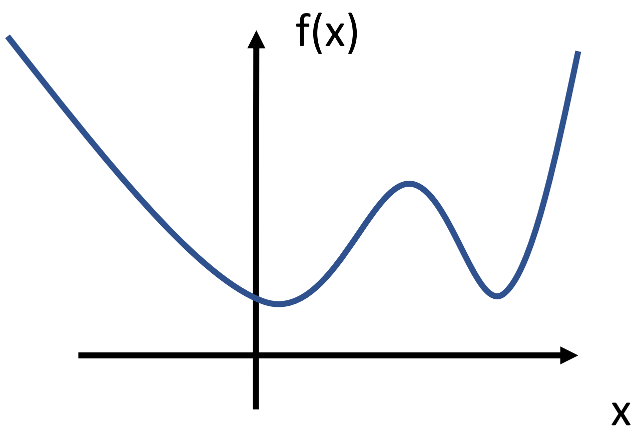
are convex.

- Another way to define a quasi-convex function: a function $f: S \rightarrow \mathcal{R}$ defined on a convex subset $S$ is quasi-convex if for all $x, y \epsilon S, 0 \leq \lambda \leq 1$, we have
- $f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}$

# Examples




Quasi-convex




Not Quasi-convex

# Assignment - Entropy

- 1) Feel comfortable to use any properties of a convex function.
- 2) Use the definition of the conjugate function directly. You should arrive at the result similar to sum of exponentials.

# Assignment – KL Divergence

- 1) Use the definition of KL divergence (e.g., for KL(p,q) sum $p_i \log(p_i / q_i)$ for the six events)
- 2) Plug in the number from the table directly.
- 3) Give two examples. One for P=Q; the other for P $\neq$Q

# Assignment – Piecewise Conjugate Function

- Discuss each region
- Your answer should include both bounded and unbounded cases in the expression of y

# Assignment – Dual Norm

- $\max(x^T y) = \max(\frac{x^T y}{||y||_p})$
- Try to solve $\nabla_y \max\left(\frac{x^T y}{\left||y|\right|_p}\right) = 0$