# Midterm Review for CSE 203B WI 2023

Chester Holtz

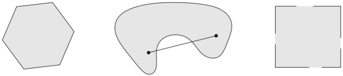
# Logistics

- Released on course website & Piazza:
  `http://cseweb.ucsd.edu/classes/wi23/cse203B-a/`
- No time limit, submission on gradescope
- Released Sunday 2/26 10:00 am PST, due Tuesday 2/28 10:00 am PST
- 2 sections:
  - $\leq$ 10 True/False (with explanation)
  - $\leq$ 5 Written question
  - One programming question

# Overview

- Convex sets
- Convex functions
- Conjugate function
- Lagrangian Dual & graph embedding
- Logistics and other recommended topics

# Convex sets: definition



- A set $S \subseteq \mathbb{R}^d$ is convex if the line segment between any two points in $C$ lies in $C$: for any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$, $\theta x_1 + (1 - \theta)x_2 \in C$

# Convex sets: example

For any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$, $\theta x_1 + (1-\theta) x_2 \in C$

Example: the polytope

$\mathcal{K} = \{x | Ax \leq b\}$ for $x, b \in \mathbb{R}^d$, $A \in \mathbb{R}^{m \times n}$

# Convex sets: example

For any $x_1, x_2 \in C$ and $0 \le \theta \le 1$, $\theta x_1 + (1-\theta)x_2 \in C$

## Example: the polytope

$P = \{x | Ax \le b\}$ for $x, b \in \mathbb{R}^d$, $A \in \mathbb{R}^{m \times n}$

let $x_1, x_2 \in P$ and $0 \le \lambda \le 1$. Then
$A((1-\lambda)x_1 + \lambda x_2) = (1-\lambda)Ax_1 + \lambda Ax_2 \le (1-\lambda)b + \lambda b = b$

Or use a geometric argument: $P$ is an intersection of $m$ half-spaces.

# Convex sets: example

For any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$, $\theta x_1 (1-\theta) x_2 \in C$

## Example: dual cone

$$K = \{(x,t) : ||x||_1 \leq t\} \implies K^* = \{(x,t) : ||x||_\infty \leq t\}$$

# Convex sets: example

For any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$, $\theta x_1 (1-\theta) x_2 \in C$

## Example: dual cone

$$K = \{(x, t) : ||x||_1 \leq t\} \implies K^* = \{(x, t) : ||x||_\infty \leq t\}$$

$$K^* = \{(x, s) : x^\top y + st \geq 0 : (x, t) \in K\}$$

# Convex sets: example

For any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$, $\theta x_1 (1-\theta) x_2 \in C$

## Example: dual cone

$$K = \{(x,t) : ||x||_1 \leq t\} \implies K^* = \{(x,t) : ||x||_\infty \leq t\}$$

$$\begin{aligned}
K^* &= \{(x,s) : x^\top y + st \geq 0 : (x,t) \in K\} \\
&= \{(x,s) : -x^\top y + s \geq 0 : (-x,1) \in K\} \\
&= \{(x,s) : x^\top y \leq s : ||x|| \leq 1\} \\
&= \{(x,s) : \max_{||x|| \leq 1} x^\top y \leq s\} \\
&= \{(x,s) : ||y||_\infty \leq s\}
\end{aligned}$$

# Convex sets: example

For any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$, $\theta x_1 (1 - \theta) x_2 \in C$

## Example: dual cone 2

Set of PSD matrices. Claim: $Y | tr(X^\top Y) \geq 0, \forall X \geq 0 \iff Y \geq 0$

# Convex sets: example

For any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$, $\theta x_1 (1-\theta) x_2 \in C$

## Example: dual cone 2

Set of PSD matrices. Claim: $Y | tr(X^\top Y) \geq 0, \forall X \geq 0 \iff Y \geq 0$

Suppose $Y \notin \mathbb{S}_+$. Show $Y \notin (\mathbb{S}_+)^*$

Suppose $Y \in \mathbb{S}_+$. Show $Y \in (\mathbb{S}_+)^*$

# Convex sets: example



## Example: dual cone 2

Set of PSD matrices $\mathcal{K} = \mathbb{S}_+$. Claim:
$Y | tr(X^\top Y) \geq 0, \forall X \geq 0 \iff Y \geq 0$

Suppose $Y \notin \mathbb{S}_+$. The, $q^\top Y q = Tr(qq^\top Y) < 0$. So, take $X = qq^\top$. Then, $y \notin (\mathbb{S}_+)^*$

Suppose $Y \in \mathbb{S}_+$. Show $Y \in (\mathbb{S}_+)^*$

# Convex sets: example



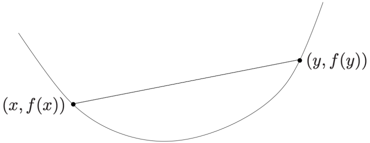
## Example: dual cone 2

Set of PSD matrices $\mathcal{K} = \mathbb{S}_+$. Claim:
$Y | tr(X^\top Y) \geq 0, \forall X \geq 0 \iff Y \geq 0$

Suppose $Y \notin \mathbb{S}_+$. The, $q^\top Y q = Tr(qq^\top Y) < 0$. So, take $X = qq^\top$. Then, $y \notin (\mathbb{S}_+)^*$

Suppose $Y \in \mathbb{S}_+$. Take EVD of $X$; $X = \sum_{i=1}^n \lambda_i q_i q_i^\top$, $\lambda_i \geq 0$.
Then $Tr(YX) = Tr(Y \sum_{i=1}^n \lambda_i q_i q_i^\top) = Tr(\sum_{i=1}^n \lambda_i q_i Y q_i^\top) \geq 0$.

# Convex functions: definition

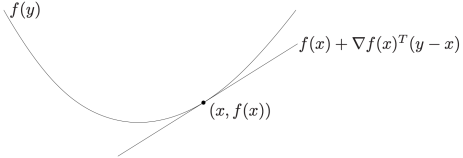



- A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if $\mathrm{dom} f$ is a convex set and if for all $x, y \in \mathrm{dom} f$ and $0 \leq \theta \leq 1$

  $$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) \quad \text{Jensen's inequality}$$

- Concave functions: $-f$ is convex

# Convex functions: first order condition




- If $f$ is differentiable ($\mathrm{dom} f$ is open, $\nabla f$ exists $\forall x \in \mathrm{dom} f$) then $f$ is convex iff $\mathrm{dom} f$ is convex and for all $x, y \in \mathrm{dom}\ \mathrm{f}$

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

# Convex functions: second order condition

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

$$\nabla^2 f \succcurlyeq 0 \quad \text{(positive semidefinite)}$$

# Convex functions: establishing convexity

## By definition

- Show by definition or first-order condition
- For twice-differentiable functions, show $\nabla^2 f \succcurlyeq 0$

## By convexity-preserving operations

- Nonnegative weighted sum
- Composition with affine function / composition with a convex + increasing function
- Pointwise maximum and supremum
- Composition
- Minimization

# Convex functions: establishing convexity

## By definition

- Show by definition or first-order condition
- For twice-differentiable functions, show $\nabla^2 f \succcurlyeq 0$

## By convexity-preserving operations

- Nonnegative weighted sum
- Composition with affine function / convex increasing function
- Pointwise maximum and supremum
- Composition
- Minimization

# Convex functions: examples

powers of absolute value

$f = |x|^p$ is convex with $p > 1$

# Convex functions: examples

powers of absolute value

$f = |x|^p$ is convex with $p > 1$

Pf: Note that the composition of a convex and convex-increasing function is convex. Prove $|\cdot|$ is convex and $x^p$ is convex and increasing.

# Convex functions: examples

log-convex function

$g(x) = log(f(x))$, s.t. $f$ convex.

# Convex functions: examples

quadratic form of inverse

$f : \mathbb{R}^n \times S^n \to \mathbb{R}$, $f(x, Y) = x^T Y^{-1} x$ is convex on $\mathbb{R}^n \times S^n_{++}$

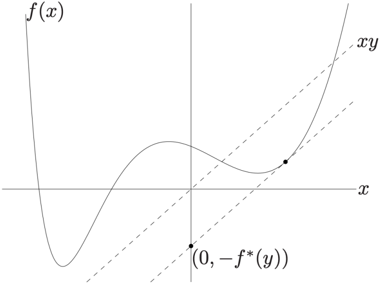
# Convex functions: examples

quadratic form of inverse

$f : \mathbb{R}^n \times S^n \to \mathbb{R}$, $f(x, Y) = x^T Y^{-1} x$ is convex on $\mathbb{R}^n \times S^n_{++}$

Show epigraph of $f$ is a convex set. Express epigraph as an LMI and apply the definiteness conditions of the Schur Complement (appendix 5.5).

# Conjugate function: definition




- Given a function $f : \mathbb{R}^n \to \mathbb{R}$, the conjugate function

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

- $\mathrm{dom} f^*$ consists of $y \in \mathrm{dom} f$ such that $\sup_{y \in \mathrm{dom} f} y^T x - f(x)$ is bounded.
- $f^*(x)$ is convex even if $f(x)$ is not convex

# Conjugate function: example

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

$f(x) = ||x||$, $f^*(x) = \sup_x y^\top x - ||x||$

- What is one way to characterize a norm?

# Conjugate function: example

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

$f(x) = ||x||$

Recall $||z||_* = sup_{||x|| \leq 1} x^\top z$

Case 1: if $||z||_* \leq 1$, then: ?

Case 2: if $||z||_* > 1$, then: ?

# Conjugate function: example

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

$f(x) = ||x||$

Recall $||z||_* = sup_{||x|| \leq 1} x^\top z$

Case 1: if $||z||_* \leq 1$, then: $z^\top (\frac{x}{||x||}) \leq ||z||_* \leq 1$
$\implies z^\top x - ||x|| \leq 0 \implies f^*(z) = sup(z^\top x - ||x|| \leq 0)$

Case 2: if $||z||_* > 1$, then: ?

# Conjugate function: example

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

$f(x) = ||x||$

Recall $||z||_* = sup_{||x|| \leq 1} x^\top z$

Case 1: if $||z||_* \leq 1$, then: $z^\top (\frac{x}{||x||}) \leq ||z||_* \leq 1$
$\implies z^\top x - ||x|| \leq 0 \implies f^*(z) = sup(z^\top x - ||x|| \leq 0)$

Case 2: if $||z||_* > 1$, then: $\exists x$ s.t. $||x|| \leq 1, {}^\top z > 1$
$\implies f^*(z) \geq z^\top (tx) - ||tx|| = t(z^\top x - ||x||) \to \infty$

# Duality

## Primal problem

$$
\begin{aligned}
\min f_0(x) \\
f_i(x) \leq 0 \\
h_i(x) = 0
\end{aligned}
$$

## Lagrange dual function $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$
\begin{aligned}
g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\
&= \inf_{x \in \mathcal{D}} \left( f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)
\end{aligned}
$$

# Duality example: Primal and Dual of a QCQP (homework 4)

$$\min_{X \in \mathbb{R}^{n \times m}} \langle X, LX \rangle + 2\langle B, X \rangle$$
$$\text{s.t. } \langle X, X \rangle \leq r^2$$

- if $n = 1$, the feasible set is the ball in $\mathbb{R}^n$; $K = \{X | X^\top X \leq c\}$
- Step 1: calculate the Lagrangian

# Duality example: Primal and Dual of a QCQP

$$\min_{X \in \mathbb{R}^{n \times m}} \langle X, LX \rangle + 2\langle B, X \rangle$$
$$\text{s.t. } \langle X, X \rangle \leq r^2$$

- $\mathcal{L}(X, \lambda) = \langle X, LX \rangle + 2\langle B, X \rangle + \lambda(X^\top X - r^2) = \langle X, (L + \lambda I)X \rangle + 2\langle B, X \rangle - \lambda r^2$
- Step 2: Lagrange dual function

# Duality example: Primal and Dual of a QCQP

$$\min_{X \in \mathbb{R}^{n \times m}} \langle X, LX \rangle + 2\langle B, X \rangle$$
$$\text{s.t. } \langle X, X \rangle \leq r^2$$

- $\mathcal{L}(X, \lambda) = \langle X, LX \rangle + 2\langle B, X \rangle + \lambda(X^\top X - r^2) = \langle X, (L + \lambda I)X \rangle + 2\langle B, X \rangle - \lambda r^2$
- FOC: $LX + B = -X\lambda$

# Duality example: Primal and Dual of a QCQP

$$\min_{X \in \mathbb{R}^{n \times m}} \langle X, LX \rangle + 2\langle B, X \rangle$$
$$\text{s.t. } \langle X, X \rangle \leq r^2$$

- $\mathcal{L}(X, \lambda) = \langle X, LX \rangle + 2\langle B, X \rangle + \lambda(X^\top X - r^2) = \langle X, (L + \lambda I)X \rangle + 2\langle B, X \rangle - \lambda r^2$
- FOC: $LX + B = -X\lambda$
- $g(\lambda) = -\langle B, (L + \lambda I)^{-1} B \rangle - \lambda r^2$
- Step 4: State dual problem

# Duality example: Primal and Dual of a QCQP

$$\min_{X \in \mathbb{R}^{n \times m}} \langle X, LX \rangle + 2\langle B, X \rangle$$
$$\text{s.t. } \langle X, X \rangle \leq r^2$$

- $\mathcal{L}(X, \lambda) = \langle X, LX \rangle + 2\langle B, X \rangle + \lambda(X^\top X - r^2) = \langle X, (L + \lambda I)X \rangle + 2\langle B, X \rangle - \lambda r^2$
- FOC: $LX + B = -X\lambda$
- $g(\lambda) = -\langle B, (L + \lambda I)^{-1} B \rangle - \lambda r^2$
- $\max_{\lambda \geq 0} g(\lambda)$

# Other

- Definitions and examples
- Characterization of norms: dual norm, p-norms as solutions to optimization problems (examples)
- Characterization of PSD (convex) quadratic forms ($X \geq 0$)
  - $y^\top X y \geq 0$
  - All eigenvalues of $X \geq 0$
  - Unbounded below if $\lambda_{\min}(X) < 0$, otherwise 0.

Good Luck!