CSE 203B: Convex Optimization
Week 3 Discuss Session

Xinyuan Wang
01/31/2020
Contents

• **Convex functions** (Ref. Chap.3)
  • Definition
  • First order condition
  • Second order condition
  • Epigraph
  • Operations that preserve convexity

• Conjugate functions
Definition of Convex Functions

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

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

sometimes called Jensen’s inequality

• Review the proof in class: necessary and sufficiency

• Strict convexity: $f(\theta x + (1-\theta)y) < \theta f(x) + (1 - \theta) f(y)$, $x \neq y$, $0 < \theta < 1$

• Concave functions: $-f$ is convex
Example: convexity of functions with definition

3.4 [RV73, page 15] Show that a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for every line segment, its average value on the segment is less than or equal to the average of its values at the endpoints of the segment: For every $x, y \in \mathbb{R}^n$,

$$
\int_0^1 f(x + \lambda(y - x)) d\lambda \leq \frac{f(x) + f(y)}{2}.
$$

i. Necessary: suppose $f$ is convex, then $f$ satisfies the basic inequality

$$f(x + \lambda(y - x)) \leq f(x) + \lambda(f(y) - f(x))$$

for $0 \leq \lambda \leq 1$. Integrating both sides from 0 to 1 on $\lambda$

$$
\int_0^1 f(x + \lambda(y - x)) d\lambda \leq \int_0^1 f(x) + \lambda(f(y) - f(x)) d\lambda = \frac{f(x) + f(y)}{2}
$$

ii. Sufficiency: suppose $f$ is not convex, then there exists $0 \leq \tilde{\lambda} \leq 1$ and

$$f\left(x + \tilde{\lambda}(y - x)\right) > f(x) + \tilde{\lambda}(f(y) - f(x))$$

there exist $\alpha, \beta \in [0,1]$ so that in the interval $[\alpha, \beta]$ the above inequality always holds, which contradicts the given inequality.
Restriction of a convex function to a line

• A function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if and only if the function $g: \mathbb{R} \rightarrow \mathbb{R}$,
  \[
g(t) = f(x + tv), \quad \text{dom } g = \{t \mid x + tv \in \text{dom } f\}
\]
is a convex on its domain for $\forall x \in \text{dom } f, v \in \mathbb{R}^n$.

• The property can be useful to check the convexity of a function

Example: Prove $f(X) = \log \det X$, $\text{dom } f = S^n_{++}$ is concave.
Restriction of a convex function to a line

Example: Prove \( f(X) = \log \det X \), \( \text{dom } f = S_{++}^n \) is concave.

Consider an arbitrary line \( X = Z + tV \), where \( Z \in S_{++}^n, V \in S^n \). Define \( g(t) = f(Z + tV) \) and restrict \( g \) to the interval values of \( t \) for \( Z + tV > 0 \). We have

\[
g(t) = \log \det(Z + tV) = \log \det\left(Z^{1/2}(I + tZ^{-1/2}VZ^{-1/2})Z^{1/2}\right)
\]

\[
= \sum_{i=1}^{n} \log(1 + t\lambda_i) + \log \det Z
\]

where \( \lambda_i \) are the eigenvalues of \( Z^{-1/2}VZ^{-1/2} \). So we have

\[
g'(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1 + t\lambda_i}, \quad g''(t) = -\sum_{i=1}^{n} \frac{\lambda_i^2}{(1 + t\lambda_i)^2} \leq 0
\]

\( f(X) \) is concave. For more practice, see Exercise 3.18.
First-order Condition

- Suppose $f$ is differentiable ($\text{dom } f$ is open and $\nabla f$ exists at $\forall x \in \text{dom } f$), then $f$ is convex iff $\text{dom } f$ is convex and for all $x, y \in \text{dom } f$
  \[ f(y) \geq f(x) + \nabla f(x)^T (y - x) \]

- Review the proof in class: necessary and sufficiency
- Strict convexity: $f(y) > f(x) + \nabla f(x)^T (y - x), x \neq y$
- Concave functions: $f(y) \leq f(x) + \nabla f(x)^T (y - x)$

Proof of first-order condition: chap 3.1.3 with the property of restricting $f$ to a line.
Second Order Condition

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

\[ \nabla^2 f(x) \succeq 0 \text{ (positive semidefinite)} \]

• Review the proof in class: necessary and sufficiency

• Strict convexity: $\nabla^2 f(x) > 0$

• Concave functions: $\nabla^2 f(x) \preceq 0$
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} \\ -x^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} \geq 0 \Rightarrow \text{convex} \]

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

More examples see chap. 3.1.5
Epigraph

• $\alpha$-sublevel set of $f: \mathbb{R}^n \to \mathbb{R}$

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

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

• Epigraph of $f: \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\text{epi } f = \{(x, t) \mid x \in \text{dom } f, f(x) \leq t\} \subseteq \mathbb{R}^{n+1}$$

• A function is convex iff its epigraph is a convex set.
Relation between convex sets and convex functions

• A function is convex iff its epigraph is a convex set.

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

\[ epif \]

• Consider a convex function \( f \) and \( x, y \in \text{dom } f \)

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

First order condition for convexity

• The hyperplane supports \( epif \) at \((x, f(x))\), for any

\[ (y, t) \in \text{epi } f \Rightarrow \nabla f(x)^T (y - x) + f(x) - t \leq 0 \]

\[ \Rightarrow \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^T \begin{bmatrix} y \\ t \end{bmatrix} - \begin{bmatrix} x \\ f(x) \end{bmatrix} \leq 0 \]
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 $\nabla^2 f(x) \succeq 0$
- 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
Pointwise Supremum

• If for each $y \in U, f(x, y): \mathbb{R}^n \to \mathbb{R}$ is convex in $x$, then function
  \[ g(x) = \sup_{y \in U} f(x, y) \]

  is convex in $x$.

• Epigraph of $g(x)$ is the intersection of epigraphs with $f$ and set $U$
  \[ \text{epi} \ g = \cap_{y \in U} \text{epi} \ f(\cdot, y) \]

  knowing $\text{epi} \ f(\cdot, y)$ is a convex set ($f$ is a convex function in $x$ and regard $y$ as a const), so $\text{epi} \ g$ is convex.

• An interesting method to establish convexity of a function: the pointwise supremum of a family of affine functions. (ref. Chap 3.2.4)
Contents

• Convex functions (Ref. Chap.3)
  • Definition
  • First order condition
  • Second order condition
  • Epigraph
  • Operations that preserve convexity

• Conjugate functions
Conjugate Function

- Given function $f: \mathbb{R}^n \to \mathbb{R}$, the conjugate function
  $$f^*(y) = \sup_{x \in \text{dom } f} y^T x - f(x)$$

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

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

Proof with pointwise supremum:

$y^T x - f(x)$ is affine function in $y$.
Examples of Conjugates

- Derive the conjugates of \( f: R \to R \)

\[
f^*(y) = \sup_{x \in \text{dom } f} yx - f(x)
\]

Affine

\( f(x) = \alpha x - \beta \)

Norm

\( f(x) = |x| \)

See more examples in chap 3.3.1
Properties of Conjugates

(1) \( f(x) + f^*(y) \geq x^Ty \)

Fenchel’s inequality. Thus, in the above example
\[
x^Ty \leq \frac{1}{2}x^TQx + \frac{1}{2}y^TQ^{-1}y, \quad \forall x, y \in \mathbb{R}^n, Q \in S_{++}^n
\]

(2) \( f^{**} = f \), if \( f \) is convex & \( f \) is closed (i.e. \( \text{epi} \, f \) is a closed set)

(3) If \( f \) is convex & differentiable, \( \text{dom} \, f = \mathbb{R}^n \)

For \( \max y^T x - f(x) \), we have \( y = \nabla f(x^*) \)

Thus, \( f^*(y) = x^T \nabla f(x^*) - f(x^*), \quad y = \nabla f(x^*) \)
Conjugate Function

• An economic interpretation

  \( f(x) \): cost to produce product with quantity \( x \)

  \( y \): market price

  \( f^*(y) \): optimal profit at given price \( y \)

• Conjugate function and the Lagrange dual function (will be covered in later class)

  • Optimization problem with linear constraints