CSE203B Convex Optimization: Lecture 3: Convex Function

CK Cheng
Dept. of Computer Science and Engineering
University of California, San Diego
Outlines

1. Definitions: Convexity, Examples & Views
2. Conditions of Optimality
   1. First Order Condition
   2. Second Order Condition
3. Operations that Preserve the Convexity
   1. Pointwise Maximum
   2. Partial Minimization
4. Conjugate Function
5. Log-Concave, Log-Convex Functions
Outlines

1. Definitions
   1. Convex Function vs Convex Set
   2. Examples
      1. Norm
      2. Entropy
      3. Affine
      4. Determinant
      5. Maximum

3. Views of Functions and Related Hyperplanes
1. Definitions: Convex Function vs Convex Set

Theorem: Given $S = \{x|f(x) \leq b\}$
If function $f(x)$ is convex, then $S$ is a convex set.
Proof: We prove by the definition of convex set.
For every $u, v \in S$, i.e. $f(u) \leq b, f(v) \leq b$,
We want to show that $\alpha u + \beta v \in S$, $\forall \alpha + \beta = 1, \alpha, \beta \geq 0$.
We have

$$f(\alpha u + \beta v) \leq \alpha f(u) + \beta f(v) \quad (f \text{ is convex})$$

$$\leq \alpha b + \beta b \quad (\alpha, \beta \geq 0)$$

$$= (\alpha + \beta) \cdot b = b \quad (\alpha + \beta = 1)$$

Thus $\alpha u + \beta v \in S$

Remark: Convex function $\Rightarrow$ Convex Set

$f(x) \leq b \quad \Rightarrow$ Convex Set

$f(x) \geq b \quad \Rightarrow$ ?
1. Convex Function Definitions: Examples

\( f: R^n \to R \) is convex if \( dom \, f \) is a convex set and

\[
f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)
\]

\( \forall x, y \in dom \, f, 0 \leq \theta \leq 1 \)

Example on \( R \):

**Convex Functions**

Affine: \( ax + b \) on \( R \) for any \( a, b \in R \)

Exponential: \( e^{ax} \) for any \( a \in R \)

Power: \( x^\alpha \) on \( R_+ \) for \( \alpha \geq 1 \) or \( \alpha \leq 0 \)

\( |x|^p \) on \( R \) for \( p \geq 1 \)

**Concave Functions**

Affine: \( ax + b \) on \( R \) for any \( a, b \in R \)

Power: \( x^\alpha \) on \( R_+ \) for \( 0 \leq \alpha \leq 1 \)

Logarithm: \( logx \) on \( R_+ \)
1. Convex Function Definitions: Examples

Example on $\mathbb{R}^n$:

Affine: \[ f(x) = a^T x + b \]
Norms: \[ \|x\|_p = \left( \sum_{i=1}^{n} |x|^p \right)^{1/p} \text{ for } p \geq 1; \]
\[ \|x\|_{\infty} = \max_k |x_k| \]

Example on $\mathbb{R}^{m \times n}$:

Affine: \[ f(X) = tr(A^T X) = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} x_{ij} \]
Spectral (max singular value):
\[ f(X) = \|X\|_2 = \sigma_{max}(X) = (\lambda_{max}(X^T X))^{1/2} \]
1. Convex Function Definitions: Examples

Concave Functions:
Log Determinant: \( f(X) = \log \det X \), \( \text{dom } f = S^n_+ \)
Proof: Let \( g(t) = f(X + tV) \) \( (V \in S^n) \)

\[
g(t) = \log \det (X + tV) = \log \det X + \log \det(I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})
\]

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

\( \lambda_i \): eigenvalue of \( X^{-\frac{1}{2}}VX^{-\frac{1}{2}} \)

\( g \) is concave in \( t \) \( \Rightarrow \) \( f \) is concave
Convex function examples: norm, max, expectation

**Norm:** If \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is a norm and \( 0 \leq \theta \leq 1 \)
\[
f(\theta x + (1 - \theta)y) \leq f(\theta x) + f((1 - \theta)y)
\]
\[
= \theta f(x) + (1 - \theta)f(y)
\]

**Max function:** \( f(x) = \max_i x_i \), \( x = [x_1, x_2, ..., x_n]^T \)
\[
f(\theta x + (1 - \theta)y) = \max_i (\theta x_i + (1 - \theta)y_i)
\]
\[
\leq \theta \max_i x_i + (1 - \theta) \max_i y_i
\]
\[
= \theta f(x) + (1 - \theta)f(y) \quad \text{for } 0 \leq \theta \leq 1
\]

**Probability:** (Expectation)
If \( f(x) \) is convex with \( p(x) \) a probability at \( x \),
\[ \text{i.e. } p(x) \geq 0, \forall x \text{ and } \int p(x) \, dx = 1 \]
Then \( f(Ex) \leq Ef(x) \),
where \( Ex = \int x \, p(x) \, dx \)
\[
Ef(x) = \int f(x) \, p(x) \, dx
\]
1.3 Views of Functions and Related Hyperplanes

Given $f(x), x \in R^n$, we plot the function in $R^n$ and $R^{n+1}$ spaces.

1. Draw function in $R^n$ space
   Equipotential surface: **tangent plane** $\nabla f (\tilde{x})^T (x - \tilde{x}) = 0$ at $\tilde{x}$

2. Draw function in $R^{n+1}$ space
   2.1 Graph of function: $\{(x, h)|x \in \text{dom } f, h = f(x)\}$
      **hyperplane** $(h = \nabla f (\tilde{x})^T (x - \tilde{x}) + f(\tilde{x}))$
      $$[\nabla f (\tilde{x})^T - 1] (\begin{bmatrix} x \\ h \end{bmatrix} - \begin{bmatrix} \tilde{x} \\ f(\tilde{x}) \end{bmatrix}) = 0$$
      Example: $f(x) = x^2$. We show the hyperplane with $\nabla f (x)$

2.2. Epigraph: $\text{epi } f: \{(x, t)|x \in \text{dom } f, f(x) \leq t\}$
      A function is convex iff its epigraph is a convex set.
      Example: $f(x) = \max\{f_i(x)| i = 1 ... r\}$, $f_i(x)$ are convex.
      Since $\text{epi } f$ is the intersect of $\text{epi } f_i$, $\text{epi } f$ is convex.
      Thus, function $f$ is convex.
2. Conditions of Optimality: First Order Condition

Definition: \( f \) is differentiable if \( \text{dom} f \) is open and
\[
\nabla f(x) \equiv \left( \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)
\]
exists at each \( x \in \text{dom} f \)

Theorem: Differentiable \( f \) with convex domain is convex
iff \( f(y) \geq f(x) + \nabla f(x)^T(y - x), \forall x, y \in \text{dom} f \)

Proof => If \( f \) is convex

Then \( (1 - t)f(x) + tf(y) \geq f((1 - t)x + ty), \forall 0 \leq t \leq 1 \)
\[
t[f(y) - f(x)] \geq f\left(x + t(y - x)\right) - f(x)
\]
\[
f(y) - f(x) \geq \frac{1}{t} \left(f\left(x + t(y - x)\right) - f(x)\right)
\]
\[
= \nabla f(x)(y - x) \quad \text{when } t \rightarrow 0
\]

<= Given \( f(y) \geq f(x) + \nabla f(x)^T(y - x), \forall x, y \in \text{dom} f \)

Let \( z = (1 - t)x + ty \)

where \[
\begin{align*}
f(x) & \geq f(z) + \nabla f(z)^T(x - z) \\
f(y) & \geq f(z) + \nabla f(z)^T(y - z)
\end{align*}
\]

Thus \( (1 - t)f(x) + tf(y) \geq f(z) \)
2. Conditions: Second Order Condition

Definition: \( f \) is twice differentiable if \( \text{dom} f \) is open and the Hessian \( \nabla^2 f(x) \in S^n \)

\[ \nabla^2 f(x)_{ij} \equiv \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n \text{ exists at each } x \in \text{dom} f \]

Theorem: Twice Differentiable \( f \) with convex domain is convex

iff \( \nabla^2 f(x) \succeq 0, \forall x \in \text{dom} f \)

Proof: Using Lagrange remainder, we can find a \( z \)

\[ f(x + t(y - x)) = f(x) + \nabla f(x)^T t(y - x) + \frac{1}{2} t^2 (y - x)^T \nabla^2 f(z)(y - x), \]

\( \forall 0 \leq t \leq 1, z \) is between \( x \) and \( x + t(y - x) \)

Since the last term is always positive by assumption, the first order condition is satisfied.
2. Conditions: Second Order Condition

Example: Negative Entropy:
\[ f(x) = x \log x, x \in R_+ \]
\[ f'(x) = \frac{x}{x} + \log x = 1 + \log x, f''(x) = \frac{1}{x} \]
Since \( x \in R_+ \), \( f''(x) > 0 \Rightarrow f(x) \) is convex

Show the plot of \( x \log x \)

Remark:
- 1\(^{\text{st}}\) order condition can be used to design and prove the property of opt. algorithms.
- 2\(^{\text{nd}}\) order condition implies the 1\(^{\text{st}}\) order condition
- 2\(^{\text{nd}}\) order condition can be used to prove the convexity of the functions.
2. Conditions: Examples

- Quadratic Function: \( f(x) = \frac{1}{2} x^T P x + q^T x + r, P \in S^n \)
  \[ \nabla f(x) = P x + q, \quad \nabla^2 f(x) = P \]

- Least Square: \( f(x) = \|Ax - b\|_2^2 \)
  \[ \nabla f(x) = 2A^T (Ax - b), \quad \nabla^2 f(x) = A^T A \]

- Quadratic over linear: \( f(x, y) = \frac{x^2}{y}, \quad y > 0 \)
  \[ \nabla f(x, y) = \left( \frac{2x}{y}, -\frac{x^2}{y^2} \right)^T \]
  \[ \nabla^2 f(x) = \begin{bmatrix}
  2 & -2x \\
  \frac{2}{y} & -\frac{2x}{y^2} \\
  -\frac{2}{y^2} & -\frac{2x^2}{y^3}
\end{bmatrix} = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y & -x \end{bmatrix} \]
2. Conditions: Examples

- Log-sum-exp: \( f(x) = \log \sum_{k=1}^{n} e^{x_k} \) (Smooth max of softmax function)

\[
\nabla^2 f(x) = \frac{1}{1^T z} \text{diag}(z) - \frac{1}{1^T z} zz^T, \quad z_k = e^{x_k}
\]

\[
v^T \nabla^2 f(x) v = \frac{1}{(1^T z)^2} \left[ (\sum_{i=1}^{n} z_i)(\sum_{i=1}^{n} v_i^2 z_i) - (\sum_{i=1}^{n} v_i z_i)^2 \right] \geq 0,
\]

for all \( v \in \mathbb{R}^n \) (Cauchy-Schwarz inequality)

Thus, \( f(x) \) is a convex function

Cauchy-Schwarz inequality: \([ (a^T a)(b^T b) \geq (a^T b)^2, \quad a_i = \sqrt{z_i}, \quad b_i = v_i \sqrt{z_i} ] \)

Proof 1: Let \( z = a - \frac{a^T b}{b^T b} b \), or \( a = z + \frac{a^T b}{b^T b} b \)

We have

\[
a^T a = z^T z + \frac{(a^T b)^2}{(b^T b)^2} b^T b \geq \frac{(a^T b)^2}{(b^T b)^2} b^T b = \frac{(a^T b)^2}{b^T b}
\]

Proof 2: By induction
3. Operations that preserve convexity

• Nonnegative multiple: $\alpha f$, where $\alpha \geq 0$, $f$ is convex
• Sum: $f_1 + f_2$, where $f_1$, and $f_2$ are convex
• Composition with affine function: $f(Ax + b)$, where $f$ is convex

Proof: $\nabla_x^2 f(Ax + b) = A^T \nabla_y^2 f(y | y = Ax + b) A$

E.g. $f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x_i)$,

$\text{dom } f = \{x | a_i^T x < b_i, i = 1, \ldots, m\}$

$f(x) = \|Ax + b\|$ (if $f$ is twice differentiable)
3. Operations that preserve convexity

- Pointwise maximum: \( f(x) = \max\{f_1(x), \ldots, f_r(x)\} \), \( f_i \) are convex

- Pointwise supremum:
  \[
g(x) = \sup_{y \in C} f(x, y), \quad \text{where } f(x, y) \text{ is convex in } x \text{ and } C \text{ is an arbitrary set}
  \]

Examples

- \( S_c(x) = \sup_{y \in C} y^T x, \) for an arbitrary set \( C \)
- \( f(x) = \sup_{y \in C} ||x - y||, \) for an arbitrary set \( C \)
- \( \lambda_{max}(X) = \sup_{\|y\|_2=1} y^T X y, \) \( X \in S^n \)
3. Operations that preserve convexity: Dual norm

Example:

\[ f(x) = \max_{\|y\|_2 \leq 1} y^T x \]

\[ f(x) = \max_{\|y\|_1 \leq 1} y^T x \]

\[ f(x) = \max_{\|y\|_p \leq 1} y^T x \]
3. Operations that preserve convexity: max function

Theorem: Pointwise maximum of convex functions is convex
Given \( f(x) = \max\{f_1(x), f_2(x)\} \), where \( f_1 \) and \( f_2 \) are convex and \( \text{dom } f = \text{dom}\{f_1\} \cap \text{dom}\{f_2\} \) is convex, then \( f(x) \) is convex.

Proof: For \( 0 \leq \theta \leq 1 \), \( x, y \in \text{dom } f \)
\[
f(\theta x + (1 - \theta)y) \\
= \max\{f_1(\theta x + (1 - \theta)y), f_2(\theta x + (1 - \theta)y)\} \\
\leq \max\{\theta f_1(x) + (1 - \theta)f_1(y), \theta f_2(x) + (1 - \theta)f_2(y)\} \\
\leq \theta \max\{f_1(x), f_2(x)\} + (1 - \theta)\max \{f_1(y), f_2(y)\} \\
= \theta f(x) + (1 - \theta)f(y)
\]
i.e. \( f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y) \)
Thus, function \( f(x) \) is convex.
3. Operations that preserve convexity: minimization

Theorem: Partial minimization

If \( g(x, y) \) is convex in \( x \) and \( y \), and a set \( C \) is convex

Then \( f(x) = \min_{y \in C} g(x, y) \) is convex.

Proof: Let \( y_1 \in \{ y | \min_{y \in C} g(x_1, y) \} \) and \( y_2 \in \{ y | \min_{y \in C} g(x_2, y) \} \),

we can write

\[
\theta f(x_1) + (1 - \theta)f(x_2) \\
= \theta g(x_1, y_1) + (1 - \theta)g(x_2, y_2) \\
\geq g(\theta x_1 + (1 - \theta)x_2, \theta y_1 + (1 - \theta)y_2) \quad (g \text{ is convex}) \\
\geq \min_{y \in C} g(\theta x_1 + (1 - \theta)x_2, y) \quad (C \text{ is convex}) \\
= f(\theta x_1 + (1 - \theta)x_2)
\]

i.e. we have \( \theta f(x_1) + (1 - \theta)f(x_2) \geq f(\theta x_1 + (1 - \theta)x_2) \)

Therefore, \( f(x) = \min_{y \in C} g(x, y) \) is convex.
3. Operations that preserve convexity

Examples for Partial Minimization

Given $f(x, y) = [x^T \ y^T] \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} [x]$

$x \in \mathbb{R}^n, y \in \mathbb{R}^m, A \in S_+^n, C \in S_+^m, \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \in S_+^{n+m}$

Let $g(x) = \min_y f(x, y) = x^T (A - BC^+B^T)x,$

$C^+: \textbf{pseudo inverse} \text{ of matrix } C. \ (\textbf{Drazin inverse, or generalized inverse})$

We can claim that function $g(x)$ is convex.

Proof:

(1) $f(x, y)$ is convex
(2) $y \in \mathbb{R}^m$ where $\mathbb{R}^m$ is a convex non-empty set
(3) Therefore, $g(x)$ is convex, i.e. $A - BC^+B^T \succeq 0$
3. Operations that preserve convexity

Composition:

Given \( g: \mathbb{R}^n \rightarrow \mathbb{R} \) and \( h: \mathbb{R} \rightarrow \mathbb{R} \), we set \( f(x) = h(g(x)) \).

- \( f \) is convex if \( g \) convex, \( h \) convex, \( \tilde{h} \) nondecreasing
- \( g \) concave, \( h \) convex, \( \tilde{h} \) nonincreasing

- \( f \) is concave if \( g \) convex, \( h \) concave, \( \tilde{h} \) nonincreasing
- \( g \) concave, \( h \) concave, \( \tilde{h} \) nondecreasing

Proof: for \( n=1 \)

\[
 f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)
\]

Ex1: \( \exp g(x) \) is convex if \( g \) is convex

Ex2: \( 1/g(x) \) is convex if \( g \) is concave and positive

Note that we set \( \tilde{h}(x) = \infty \) if \( x \notin \text{dom } h \), \( h \) is convex

\[
\tilde{h}(x) = -\infty \text{ if } x \notin \text{dom } h, \ h \text{ is concave}
\]
3. Operations that preserve convexity

Show that $h(g(\theta x + (1 - \theta)y)) \leq \theta h(g(x)) + (1 - \theta)h(g(y))$
for the case that $g, h$ are convex, and $\tilde{h}$ is nondecreasing

(1) $g$ is convex

\[ g(\theta x + (1 - \theta)y) \leq \theta g(x) + (1 - \theta)g(y) \]

(2) $h$ is nondecreasing: From (1), we have

\[ h(g(\theta x + (1 - \theta)y)) \leq h(\theta g(x) + (1 - \theta)g(y)) \]

(3) $h$ is convex

\[ h(\theta g(x) + (1 - \theta)g(y)) \leq \theta h(g(x)) + (1 - \theta)h(g(y)) \]

(4) From (2) & (3)

\[ h(g(\theta x + (1 - \theta)y)) \leq \theta h(g(x)) + (1 - \theta)h(g(y)) \]
The setting of conjugate functions starts from the following problem (which may not be convex)

\[ \min f(x) \]
subject to
\[ x \leq 0 \]

We convert to a function of \( y \)

\[ \inf_x f(x) - y^T x \]

The conjugate function is

\[ f^*(y) = \sup_x y^T x - f(x) \]

In the class, we interchange \( \min \) and \( \inf \); \( \max \) and \( \sup \) to simplify the notation.
4. Conjugate Functions

Given $f : R^n \rightarrow R$, we have $f^* : R^n \rightarrow R$

$$f^*(y) = \sup_{x \in \text{dom} f} y^T x - f(x); \quad (-f^*(y) = \min_{x \in \text{dom} f} -y^T x + f(x))$$

Constraint: $y \in R^n$ for which the supremum is finite (bounded)

$f^*(y)$ is called the conjugate of function $f$

Theorem: $f^*(y)$ is convex (pointwise maximum)

Proof: $f^*(\theta y_1 + (1 - \theta)y_2) = \sup_x (\theta y_1 + (1 - \theta)y_2)^T x - f(x)$$

$$\leq \sup_x (\theta y_1^T x + \theta f(x)) + \sup_x ((1 - \theta)y_2^T x - (1 - \theta)f(x))$$

$$= \theta f^*(y_1) + (1 - \theta)f^*(y_2)$$

Remark: $f^*(y)$ is convex even if $f(x)$ is not convex
4. Conjugate Functions

Suppose we have a pair $\bar{x}$, $\bar{y}$, such that $f^*(\bar{y}) = \bar{y}^T\bar{x} - f(\bar{x})$, we can show that $\bar{y} = \nabla_x f(\bar{x})$ (exercise 3.40)
And the supporting hyperplane: $\bar{y}^T x - h = f^*(\bar{y})$

$$\begin{bmatrix} \bar{y}^T & -1 \end{bmatrix} \begin{bmatrix} x \\ h \end{bmatrix} = f^*(\bar{y})$$

Ex. $f(x) = x^2 - 2x$, $x \in R$

$$f^*(y) = \sup_{x} yx - x^2 + 2x, y \in R$$
4. Conjugate Functions

One way to view conjugate function

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

\( x \): negative slack
\( y \): shadow price (loss) to accommodate the slack
\( f^*(y) \): balance between price slack product \((y^T x)\) and objective function \(f(x)\).

Remark: When \(f^*(y)\) is unbounded, the shadow price \(y\) is not reasonable.
Ex: $f(x) = ax + b, \ x \in R$

$$f^*(y) = \sup_x (yx - ax - b)$$

(1) If $y \neq a, f^*(y) = \infty$

(2) If $y = a, f^*(y) = -b \Rightarrow \text{dom } f^* = a, f^*(y) = -b$
4. Conjugate Functions: Examples (single variable)

Ex: \( f(x) = -\log x, \ x \in R_{++} \)

\[
f^*(y) = \sup_{x \in R_{++}} yx + \log x
\]

(1) If \( y \geq 0 \), \( f^*(y) = \infty \)

(2) If \( y < 0 \), \( f^*(y) = \max_{x \in R_{++}} xy + \log x \)

Let \( g(x) = xy + \log x \), \( g'(x) = y + \frac{1}{x} \)

If \( g'(x) = 0 \), \( x = -\frac{1}{y} \)

Thus, \( f^*(y) = -1 + \log \left(-\frac{1}{y}\right) = -1 - \log(-y) \)

\[\Rightarrow \text{dom } f^* = -R_{++}, \ f^*(y) = -1 - \log(-y)\]
4. Conjugate Functions

Ex: \( f(x) = e^x, \ x \in R \)

\[ f^*(y) = \sup_x xy - e^x \]

(1) \( y < 0 \): \( f^*(y) = \infty \)

(2) \( y > 0 \): Let \( g(x) = xy - e^x \Rightarrow g'(x) = y - e^x \)
   
   If \( g'(x) = 0 \), then \( x = \log y \)
   
   Thus \( f^*(y) = y\log y - y \)

(3) \( y = 0 \): \( f^*(y) = 0 \Rightarrow \text{dom } f^* = R_+, \ f^*(y) = y\log y - y \)

Therefore, we have

\[ f^*(y) = y\log y - y, \ \text{where } y \geq 0. \]
4. Conjugate Functions

Ex: $f(x) = x\log x$, $x \in R_+$, $f(0) = 0$

$$f^*(y) = \sup_x xy - x\log x$$

Let $g(x) = xy - x\log x \rightarrow g'(x) = y - \log x - 1$

Suppose $g'(x) = 0$, we have $y = 1 + \log x$ or $x = e^{y-1}$

Thus $f^*(y) = ye^{y-1} - e^{y-1}(y - 1) = e^{y-1}$ where $y \in R$
4. Conjugate Functions

Ex: \( f(x) = \frac{1}{2} x^T Q x, \ x \in R^n, Q \in S^n_{++} \)

\[ f^*(y) = \sup_x x^T y - \frac{1}{2} x^T Q x \]

Let \( g(x) = x^T y - \frac{1}{2} x^T Q x \Rightarrow \nabla g(x) = y - Q x \)

If \( \nabla g(x) = 0 \), we have \( x = Q^{-1} y \)

Thus, \( f^*(y) = \frac{1}{2} y^T Q^{-1} y \)

Remark: Suppose that \( f^*(\bar{y}) = \bar{y}^T \bar{x} - f(\bar{x}) \) and \( \nabla^2 f(\bar{x}) > 0 \)

We have \( \nabla f^*(\bar{y}) = \bar{x} \) and \( \nabla^2 f^*(\bar{y}) = (\nabla^2 f(\bar{x}))^{-1} \) (exercise 3.40)
4. Conjugate Functions

Basic Properties

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

Fenchel’s inequality. Thus, in the above example

\[
x^T y \leq \frac{1}{2} x^T Q x + \frac{1}{2} y^T Q^{-1} y, \ \forall x, y \in \mathbb{R}^n, Q \in S^n_{++}
\]

(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, \( f^*(y) = x^* T \nabla f(x^*) - f(x^*), \ y = \nabla f(x^*) \)
4. Conjugate Functions

Ex: \( f(x) = \log \sum_{i=1}^{n} e^{x_i} \leftrightarrow f^*(y) = \sum_{i=1}^{n} y_i \log y_i \)

\[ f^*(y) = \sup_{x} y^T x - f(x) = \sup_{x} y^T x - \log \sum_{i=1}^{n} e^{x_i} \]

Let \( g(x) = y^T x - \log \sum_{i=1}^{n} e^{x_i} \)

\[ \frac{\partial g(x)}{\partial x_i} = y_i - \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}} = 0 \]

Thus, \( y_i = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}} \), i.e. \( 1^T y = 1 \)

(1) \( 1^T y \neq 1 \rightarrow \) unbounded

(2) \( y_i < 0 \rightarrow \) unbounded

(3) \( f^*(y) = \sum_{i=1}^{n} y_i \log y_i \), \( y \geq 0, 1^T y = 1 \)
5. Log-Convex, Log-Concave Functions

Log function: $\log f(x), \ f: R^n \rightarrow R, f(x) > 0, \forall x \in \text{dom } f$

Suppose $f$ is twice differentiable, $\text{dom } f$ is convex.

$$\nabla^2 \log f(x) = \frac{1}{f(x)} \nabla^2 f(x) - \frac{1}{f(x)^2} \nabla f(x)\nabla f(x)^T$$

Then

$f$ is log-convex iff $\forall x \in \text{dom } f$

$f(x)\nabla^2 f(x) \geq \nabla f(x)\nabla f(x)^T$

$f$ is log-concave iff $\forall x \in \text{dom } f$

$f(x)\nabla^2 f(x) \leq \nabla f(x)\nabla f(x)^T$
5. Log-Concave, Log-Convex Functions

\( f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) > 0, \forall x \in \text{dom } f \)

Definition: If \( \log f \) is concave, \( f \) is log-concave.

Definition: If \( \log f \) is convex, \( f \) is log-convex.

Ex : \( f(x) = a^T x + b, \text{dom } f = \{x | a^T x + b\} : \text{log-concave} \)

\( f(x) = x^a, \ x \in \mathbb{R}^+, \ a \leq 0 : \text{log-convex} \)

\( a > 0 : \text{log-concave} \)

\( f(x) = e^{ax} : \text{log convex & log-concave} \)

\( f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du : \text{cumulative distribution function of Gaussian density log-concave} \)

\( f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})} : \text{log-concave} \)
5. Log-Convex, Log-Concave Functions

Properties

\[
\nabla^2 \log f(x) = \frac{1}{f(x)} \nabla^2 f(x) - \frac{1}{f(x)^2} \nabla f(x) \nabla f(x)^T
\]

\[
f(x) \nabla^2 f(x) \geq \nabla f(x) \nabla f(x)^T, \quad \forall x \in \text{dom } f : \text{log-convex}
\]

\[
f(x) \nabla^2 f(x) \leq \nabla f(x) \nabla f(x)^T, \quad \forall x \in \text{dom } f : \text{log-concave}
\]
Outlines

1. Definitions: Convexity, Examples & Views
2. Conditions of Optimality
   1. First Order Condition
   2. Second Order Condition
3. Operations that Preserve the Convexity
   1. Pointwise Maximum
   2. Partial Minimization
4. Conjugate Function
5. Log-Concave, Log-Convex Functions