

CSE203B Convex Optimization: Lecture 3: Convex Function

CK Cheng

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

Convex Problems and Convex Functions

$$\min_x f_0(x), x \in R^n$$

Subject to

$$f_i(x) \leq 0, i = 1, \dots, m$$

$$h_i(x) = 0, i = 1, \dots, p$$

We use functions to define the objective function and qualify the constraints. For a convex problem, the desired features:

Objectives $f_0(x)$: Convex function

Constraints:

$f_i(x)$: Convexity is a sufficient condition for a convex set

$h_i(x)$: Linear equation

Outlines

Examples: Approximation and Mutual Entropy

1. Definition
 1. Convex Function
 2. Examples of Convex Functions
 3. Views of Functions (**Hyperplanes**)
2. Conditions of Convex Functions
 1. First Order Condition
 2. Second Order Condition
3. Operations that Preserve the Convexity
 1. Pointwise Maximum (**Intersection of Sets**)
 2. Partial Minimization
4. Conjugate Function (**Hyperplanes**)
5. Log-Concave, Log-Convex Functions

Examples: Approximation

Given a set of numbers, $x_i \in R, i = 1, \dots, m$,
calculate functions

- $f_1(x_1, x_2, \dots, x_m) = \max_i x_i$
- $f_2(x_1, x_2, \dots, x_m) = \log \sum_i e^{x_i/\gamma}$
- $f_3(x_1, x_2, \dots, x_m) = \sum_i x_i e^{x_i/\gamma} / \sum_i e^{x_i/\gamma}$
- ...

Examples: Entropy and Kullback-Leibler Divergence

Negative Entropy:

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

Kullback-Leibler:

$$D_{kl}(u, v) = \sum_i u_i \log \left(\frac{u_i}{v_i} \right) - u_i + v_i,$$

where $u, v \in R_{++}^n$.

Outlines

1. Definition

1. Convex Function Definition and Relation to Convex Set

2. Examples

1. Norms
2. Entropy
3. Linear functions
4. Determinant of matrices
5. Maximum of functions

3. Views of Functions and Related Hyperplanes

1.1. Convex Function Definition

Definition: $f: R^n \rightarrow R$ is convex if $dom f$ is a convex set and

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

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

1.1. Definition: Convex Functions vs. Convex Sets

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

$$\begin{aligned} f(\alpha u + \beta v) &\leq \alpha f(u) + \beta f(v) \quad (f \text{ is convex}) \\ &\leq \alpha b + \beta b \quad (\alpha, \beta \geq 0) \\ &= (\alpha + \beta) \cdot b = b \quad (\alpha + \beta = 1) \end{aligned}$$

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

Remark: Convex function \Rightarrow Convex Set

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

$$f(x) \geq b \quad \Rightarrow ?$$

1.2. Convex Function Examples

$f: R^n \rightarrow R$ 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

Linear: $ax + b$ on R for any $a, b \in R$

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

Power: x^α 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^α on R_{++} for $0 \leq \alpha \leq 1$

Logarithm: $\log x$ on R_{++}

1.2. Convex Function Examples

Concave Functions:

Log Determinant: $f(X) = \log \det X$, $\text{dom } f = S_{++}^n$

Proof: Let $g(t) = f(X + tV)$ ($V \in S^n$)

$$\begin{aligned} g(t) &= \log \det (X + tV) = \log \det X + \log \det (I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}}) \\ &= \log \det X + \sum_{i=1}^n \log(1 + t\lambda_i) \end{aligned}$$

λ_i : eigenvalue of $X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$

g is concave in $t \Rightarrow f$ is concave

1.2. Convex Function Examples

Example on R^n :

Linear function: $f(x) = a^T x + b$

Norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$;

$$\|x\|_\infty = \max_k |x_k|$$

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

Linear function: $f(X) = \text{tr}(A^T X) = \sum_{i=1}^m \sum_{j=1}^n A_{ij} x_{ij}$

Spectral (max singular value):

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = (\lambda_{\max}(X^T X))^{1/2}$$

1.2. Convex function examples: norm, max, expectation

Norm: If $f: R^n \rightarrow R$ is a norm and $0 \leq \theta \leq 1$

$$\begin{aligned} f(\theta x + (1 - \theta)y) &\leq f(\theta x) + f((1 - \theta)y) && \text{triangle inequality} \\ &= \theta f(x) + (1 - \theta)f(y) && \text{scalability} \end{aligned}$$

Max function: $f(x) = \max_i x_i$, $x = [x_1, x_2, \dots, x_n]^T$

$$\begin{aligned} f(\theta x + (1 - \theta)y) &= \max_i (\theta x_i + (1 - \theta)y_i) \\ &\leq \theta \max_i x_i + (1 - \theta) \max_i y_i \\ &= \theta f(x) + (1 - \theta)f(y) \quad \text{for } 0 \leq \theta \leq 1 \end{aligned}$$

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 $f(Ex) \leq Ef(x)$,

$$\text{where } Ex = \int x p(x) dx$$

$$Ef(x) = \int f(x) p(x) dx$$

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 $(h = \nabla f(\tilde{x})^T (x - \tilde{x}) + f(\tilde{x}))$

$$[\nabla f(\tilde{x})^T \quad -1] \left(\begin{bmatrix} x \\ h \end{bmatrix} - \begin{bmatrix} \tilde{x} \\ f(\tilde{x}) \end{bmatrix} \right) = 0$$

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: $f(x) = \max\{f_i(x) | i = 1 \dots r\}$, $f_i(x)$ are convex.

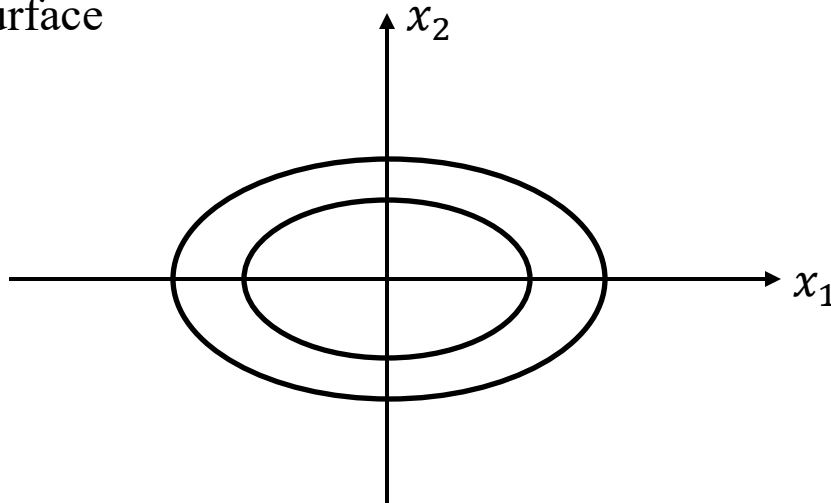
Since $\text{epi } f$ is the intersect of $\text{epi } f_i$, $\text{epi } f$ is convex.

Thus, function f is convex.

1.3 Views of Functions: Example

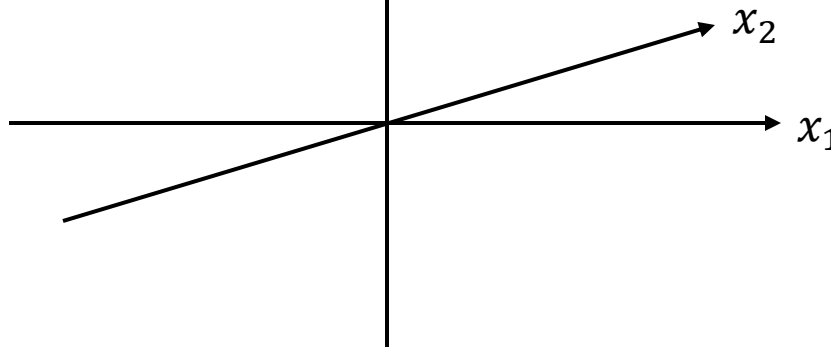
$$\text{Ex. } f(x_1, x_2) = ax_1^2 + bx_2^2, a, b > 0.$$

1. Equipotential surface



2. $\{(x, h) | x \in \text{dom } f, h = f(x)\}$

f



3. Epigraph

2. Conditions of Convexity: First Order Condition

Definition: f is differentiable if $dom f$ is open and

$$\nabla f(x) \equiv \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T \text{ exists at each } x \in dom f$$

Theorem: Differentiable f with convex domain is convex

$$\text{iff } f(y) \geq f(x) + \nabla f(x)^T (y - x), \forall x, y \in dom f$$

Proof \Rightarrow If f is convex

$$\text{Then } (1 - t)f(x) + tf(y) \geq f((1 - t)x + ty), \forall 0 \leq t \leq 1$$

$$t[f(y) - f(x)] \geq f(x + t(y - x)) - f(x)$$

$$f(y) - f(x) \geq \frac{1}{t} (f(x + t(y - x)) - f(x))$$

$$= \nabla f(x)(y - x) \quad \text{when } t \rightarrow 0$$

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

$$\text{Let } z = (1 - t)x + ty$$

$$\text{where } \begin{cases} f(x) \geq f(z) + \nabla f(z)^T (x - z) \\ f(y) \geq f(z) + \nabla f(z)^T (y - z) \end{cases}$$

$$\text{Thus } (1 - t)f(x) + tf(y) \geq f(z)$$

2. Conditions: Second Order Condition

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

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

Theorem: Twice Differentiable f with convex domain is convex
iff $\nabla^2 f(x) \succeq 0, \forall x \in dom f$

Proof: Using Lagrange remainder, we can find a z
 $f(x + t(y - x))$

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

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

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:

- 1st order condition can be used to design and prove the property of opt. algorithms.
- 2nd order condition implies the 1st order condition
- 2nd order condition can be used to prove the convexity of the functions.

2. Conditions: Examples

- Quadratic Function: $f(x) = \frac{1}{2}x^T Px + q^T x + r, P \in S^n$
 $\nabla f(x) = Px + q, \nabla^2 f(x) = P$
- Least Square: $f(x) = \|Ax - b\|_2^2$
 $\nabla f(x) = 2A^T(Ax - b), \nabla^2 f(x) = A^T A$
- Quadratic over linear: $f(x, y) = \frac{x^2}{y}, y > 0$

$$\nabla f(x, y) = \left(\frac{2x}{y}, -\frac{x^2}{y^2} \right)^T,$$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix} = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y & -x \end{bmatrix}$$

2. Conditions: Examples

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

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

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

for all $v \in R^n$ (Cauchy-Schwarz inequality)

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

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

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

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

$$f(x) = \|Ax + b\| \quad (\text{if } f \text{ is twice differentiable})$$

3. Operations that preserve convexity

- Pointwise maximum: $f(x) = \max\{f_1(x), \dots, f_r(x)\}$, f_i are convex
- Pointwise supremum:
$$g(x) = \sup_{y \in C} f(x, y),$$
 where $f(x, y)$ is convex in x and C 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 $y_1 \in \{y \mid \min_{y \in C} g(x_1, y)\}$ and $y_2 \in \{y \mid \min_{y \in C} g(x_2, y)\}$,

we can write

$$\begin{aligned} & \theta f(x_1) + (1 - \theta)f(x_2) \\ &= \theta g(x_1, y_1) + (1 - \theta)g(x_2, y_2) \\ &\geq g(\theta x_1 + (1 - \theta)x_2, \theta y_1 + (1 - \theta)y_2) \text{ (*g is convex*)} \\ &\geq \min_{y \in C} g(\theta x_1 + (1 - \theta)x_2, y) \text{ (*C is convex*)} \\ &= f(\theta x_1 + (1 - \theta)x_2) \end{aligned}$$

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

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

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

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

C^+ : **pseudo inverse** of matrix C . (**Drazin inverse, or generalized inverse**)

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

Proof:

- (1) $f(x, y)$ is convex
- (2) $y \in R^m$ where R^m is a convex non-empty set
- (3) Therefore, $g(x)$ is convex, i.e. $A - BC^+B^T \succcurlyeq 0$

3. Operations that preserve convexity

Composition:

Given $g: R^n \rightarrow R$ and $h: R \rightarrow R$, 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

$\tilde{h}(x) = -\infty$ if $x \notin \text{dom } h$, h 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)

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

4. Conjugate Functions

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

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

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)

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

$$\begin{aligned} &\leq \sup_x \left(\theta y_1^T x - \theta f(x) \right) + \sup_x \left((1 - \theta) y_2^T x - (1 - \theta) f(x) \right) \\ &= \theta f^*(y_1) + (1 - \theta) f^*(y_2) \end{aligned}$$

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 : $\bar{y}^T x - h = f^*(\bar{y})$

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

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

$$f^*(y) = \sup_x yx - x^2 + 2x, \quad y \in R$$

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.

4. Conjugate Functions: Examples (single variable)

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: $f(x) = -\log x, \quad x \in \mathbb{R}_{++}$

$$f^*(y) = \sup_{x \in \mathbb{R}_{++}} yx + \log x$$

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

(2) If $y < 0, \quad f^*(y) = \max_{x \in \mathbb{R}_{++}} xy + \log x$

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

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

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

$\rightarrow \text{dom } f^* = -\mathbb{R}_{++}, \quad f^*(y) = -1 - \log(-y)$

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) $y > 0 : \text{Let } g(x) = xy - e^x \rightarrow g'(x) = y - e^x$

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: $f(x) = \frac{1}{2}x^T Qx$, $x \in R^n$, $Q \in S_{+++}^n$

$$f^*(y) = \sup_x x^T y - \frac{1}{2}x^T Qx$$

Let $g(x) = x^T y - \frac{1}{2}x^T Qx \rightarrow \nabla g(x) = y - Qx$

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 $\nabla^2 f(\bar{x}) \succ 0$

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

$$x^T y \leq \frac{1}{2} x^T Q x + \frac{1}{2} y^T Q^{-1} y, \quad \forall x, y \in R^n, Q \in S_{++}^n$$

(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, $\text{dom } f = R^n$

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

Thus, $f^*(y) = x^{*T} \nabla f(x^*) - f(x^*)$, $y = \nabla f(x^*)$

4. Conjugate Functions

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

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

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

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

$$\text{Thus, } y_i = \frac{e^{x_i}}{\sum_{i=1}^n e^{x_i}}, \quad \text{i.e. } \mathbf{1}^T y = 1$$

(1) $\mathbf{1}^T y \neq 1 \rightarrow$ unbounded

(2) $y_i < 0 \rightarrow$ unbounded

(3) $f^*(y) = \sum_{i=1}^n y_i \log y_i, \quad y \geq 0, \mathbf{1}^T y = 1$

4. Conjugate Functions: Summary

Given vector y , and conjugate function $f^*(y) = \sup_x y^T x - f(x)$, $x, y \in R^n$

let $x^* = \arg \sup_x y^T x - f(x)$.

For the epigraph of $f(x)$,

we have a **supporting hyperplane**: $h = y^T(x - x^*) + f(x^*)$

We can rewrite as $[y^T \quad -1] \left(\begin{bmatrix} x \\ h \end{bmatrix} - \begin{bmatrix} x^* \\ f(x^*) \end{bmatrix} \right) = 0$

The plane touches the $(n + 1)$ 'th axis at $-f^*(y)$.

5. Log-Concave, Log-Convex 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

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

Definition: If $\log f$ is concave, f is log-concave.

Definition: If $\log f$ is convex, f is log-convex.

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

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

$$a > 0 : \text{log-concave}$$

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

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

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-Concave, Log-Convex 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. Definition: Convexity, Examples & Views
2. Conditions of Convexity
 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