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

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

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

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

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

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)

$\begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_k & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix}$

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

$C = D \Sigma D^T$

Proof:

$C^{-1} = D \Sigma^+ D^T \quad \Sigma^+ = \begin{bmatrix} \lambda_1^{-1} & & & & \\ & \lambda_2^{-1} & & & \\ & & \lambda_k^{-1} & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix}$

(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. $A - BC^+ B^T \succcurlyeq 0$

4W1

$$g(x) = \min_y \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Let $h(y) = x^T A x + x^T B y + y^T B^T x + y^T C y$

$$\nabla h(y) = 2Cy + 2B^T x = 0.$$

$$\Rightarrow \qquad y = -C^+ B^T x$$

Thus, we have

$$g(x) = x^T A x + x^T B(-C^+ B^T x) + (-C^+ B^T x)^T B^T x$$
$$+ (-C^+ B^T x)^T C(-C^+ B^T x)$$
$$= x^T (A - B C^+ B^T) x$$

A.20

3. Operations that preserve convexity

Composition:

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

$$f''(x) = h''\big(g(x)\big)g'(x)^2 + h'\big(g(x)\big)g''(x)$$

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 $x \notin dom\ h$, h is convex

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

3. Operations that preserve convexity

Show that $h\big(g(\theta x + (1-\theta)y)\big) \le \theta h\big(g(x)\big) + (1-\theta)h(g(y))$

for the case that g, h are convex, and $\tilde{h}$ is nondecreasing

(1) g is convex

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

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

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

(3) h is convex

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

(4) From (2) & (3)

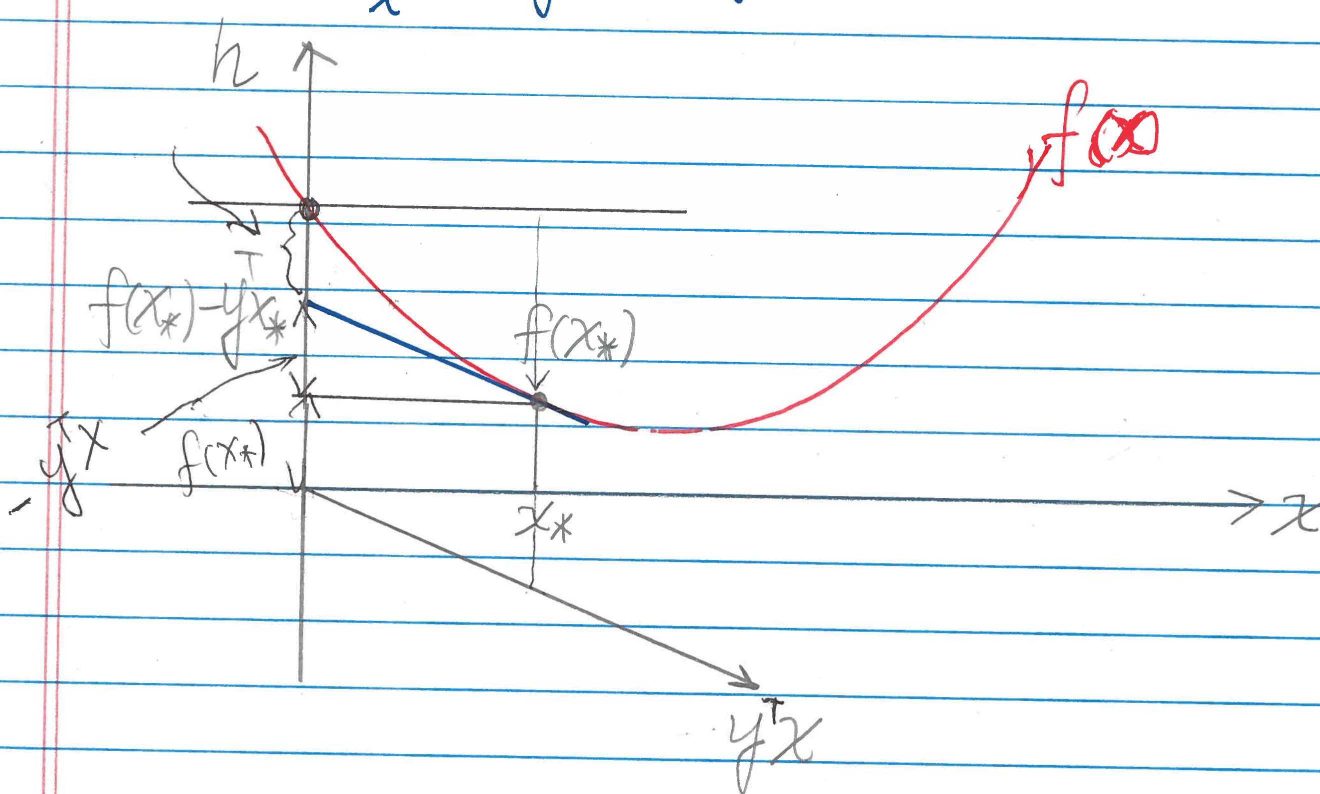
$$h\big(g(\theta x + (1-\theta)y)\big) \le \theta h\big(g(x)\big) + (1-\theta)h\big(g(y)\big)$$

Given cost of product $f(x)$

cost of production $-y^T x$ $(x \geq 0)$

We $\min_x f(x) - y^T x$

or $\max_x y^T x - f(x)$



Let $g(x) = y^T x - f(x)$

$\nabla g(x) = y - \nabla f(x)$

4. Conjugate Functions

The setting of conjugate functions starts from the following problem (which may not be convex)

$$\min f(x)$$

subject to

$$x \le 0$$

We convert to a function of y

$$\inf_x f(x) - y^T x$$

shadow price, or Lagrange multiplier

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

$[\theta + (1-\theta)] f(x)$

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

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

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

$$\le \sup_x \left(\theta y_1^T x \cancel{+} - \theta f(x)\right) + \sup_x ((1-\theta) y_2^T x - (1-\theta) f(x))$$

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

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

$\max(V+U) \le \max V + \max U = 7$ (3, 4)

$V = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $U = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$

$\max(V+U) = 5$

4. Conjugate Functions

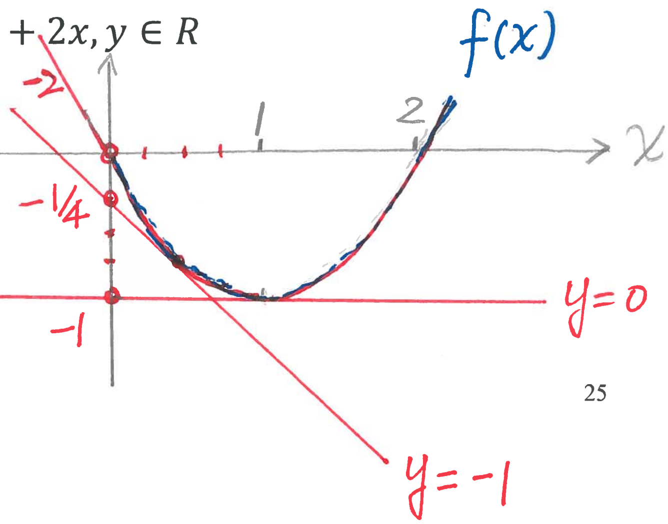
Suppose we have a pair $\bar{x}, \bar{y}$, such that $f^*(\bar{y}) = \bar{y}^T\bar{x} - f(\bar{x})$, we can show that $\bar{y} = \nabla_x f(\bar{x})$ (exercise 3.40)

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

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



(1) $x^* = \frac{y}{2} + 1$

(2) $f^*(y) = \frac{y^2}{4} + y + 1$

y	x^*	$f(x^*)$	$y^T x^* - f(x^*)$
0	1	-1	-1
-1	1/2	-3/4	-1/4
-2	0	0	0

4. Conjugate Functions

One way to view conjugate function

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

$\nabla_x y^T x - f(x) = y - \nabla f(x) = 0$

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: $f(x) = ax + b,\ x \in R$

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

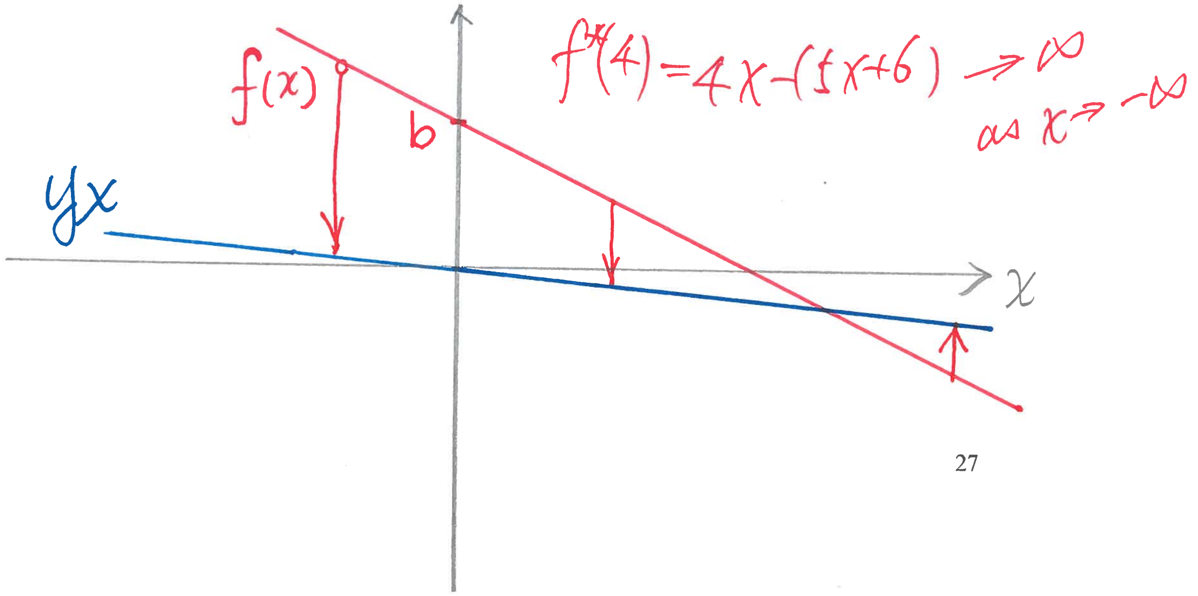
(1) If $y \neq a, f^*(y) = \infty$

(2) If $y = a, f^*(y) = -b \rightarrow dom\ f^* = a, f^*(y) = -b$

eg $f(x) = 5x + 6$

$f^*(6) = 6x - (5x+6) \to \infty$ as $x \to \infty$

$f^*(4) = 4x - (5x+6) \to \infty$ as $x \to -\infty$.



4. Conjugate Functions: Examples (single variable)

Ex: $f(x) = -logx,\ \ x \in \mathrm{R}_{++}$

$$f^*(y) = \sup_{x \in R_{++}} yx + logx$$

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

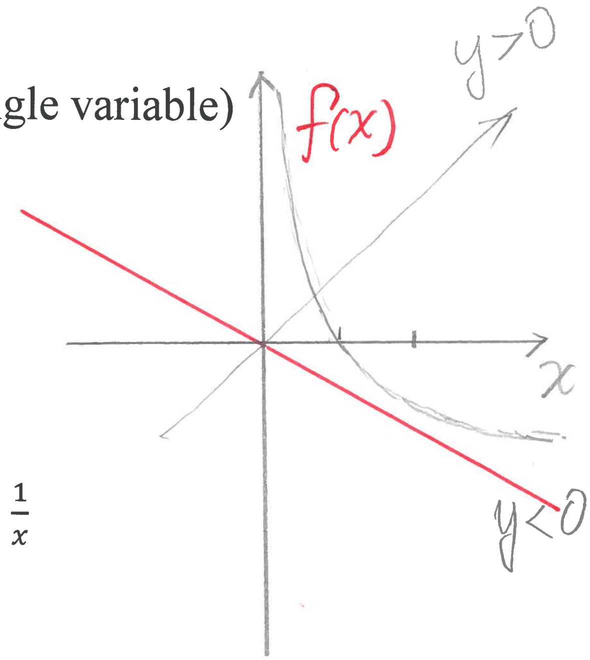
(2) If $y < 0,\ f^*(y) = \max_{x \in R_{++}} xy + logx$

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

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

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

$\rightarrow dom\ f^* = -R_{++},\ f^*(y) = -1 - \log(-y)$



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 $g(x) = xy - e^x \rightarrow g'(x) = y - e^x$

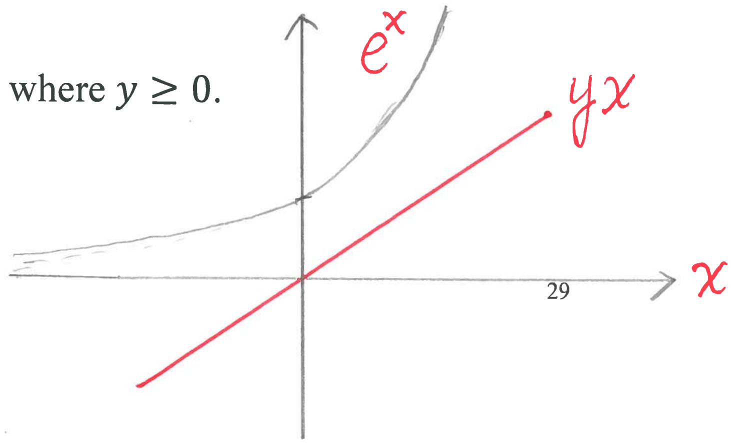
If $g'(x) = 0$, then $x = logy$

Thus $f^*(y) = ylogy - y$

(3) $y = 0 : f^*(y) = 0 \rightarrow dom\ f^* = R_+,\ f^*(y) = ylogy - y$

Therefore, we have

$$f^*(y) = ylogy - y, \text{ where } y \geq 0.$$



4. Conjugate Functions

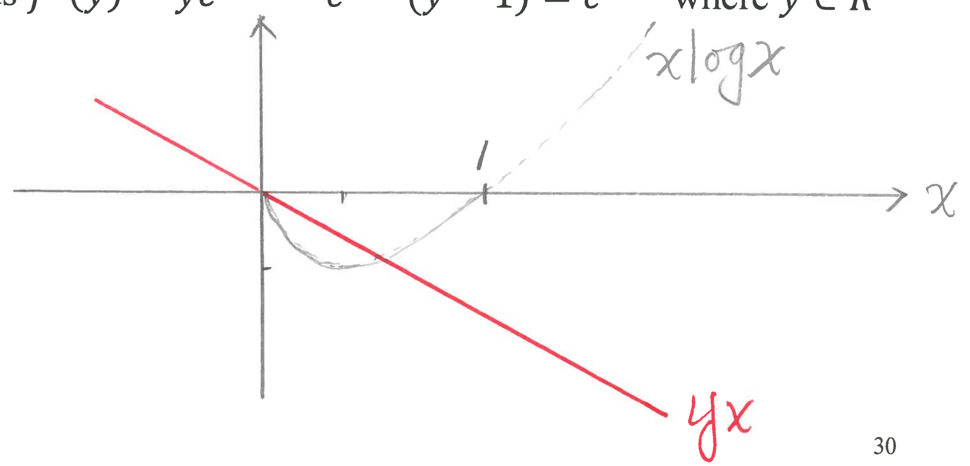
Ex: $f(x) = xlogx,\ x \in R_+,\ f(0) = 0$

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

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

Suppose $g'(x) = 0$, we have $y = 1 + logx$ 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 $\nabla^2 f^*(\bar{y}) = \left(\nabla^2 f(\bar{x})\right)^{-1}$ (exercise 3.40)

$Q^{-1} = (Q)^{-1}$

4. Conjugate Functions

$f^*(y) = \max_x x^T y - f(x)$

$\geq x^T y - f(x)\ \forall x$

Basic Properties

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

Fenchel's inequality. Thus, in the above example

$\Rightarrow f(x) + f^*(y) \geq x^T y\ \forall x$

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

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

(3) If *f* is convex & differentiable, $dom\ f = R^n$

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

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

h

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) $1^T y \neq 1$ → unbounded

(2) $y_i < 0$ → unbounded

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

5. Log-Concave, Log-Convex Functions

Log function : $logf(x)$, $f: R^n \to R, f(x) > 0, \forall x \in dom\ f$

Suppose f is twice differentiable, $dom\ f$ is convex.

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

Then

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

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

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

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

5. Log-Concave, Log-Convex Functions

$f: R^n \to R, \qquad f(x) > 0, \forall x \in dom\, f$

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

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

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

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

$a > 0$: log-concave

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

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

$f(x) = \frac{1}{\sqrt{(2\pi)^n det\Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})}$: log-concave

5. Log-Concave, Log-Convex Functions

Properties

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

$f(x)\nabla^2 f(x) \geq \nabla f(x) \nabla f(x)^T,\ \forall x \in dom\, f$: log-convex

$f(x)\nabla^2 f(x) \leq \nabla f(x) \nabla f(x)^T,\ \forall x \in dom\, f$: 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