Optimization Problems

Lecture 5, Convex Optimization

National Taiwan University

March 25, 2021

Table of contents

Conjugate functions

Conjugate functions

Let $f : \mathbf{R}^n \to \mathbf{R}$. The function $f^* : \mathbf{R}^n \to \mathbf{R}$, defined as

inner product

$$f^*(y) = \sup_{x \in \textbf{dom } f} \left(y^T x - f(x) \right),$$ 在 dom f 內找 x, s.t. sup($y^T x - f(x)$)

is called the **conjugate** of the function f. The domain of f^* is

$$\textbf{dom } f^* = \left\{ y \in \mathbf{R}^n \;\middle|\; \exists z \in \mathbf{R} \text{ s.t. } \forall x \in \textbf{dom } f,\ y^T x - f(x) < z \right\}.$$ 要有上界

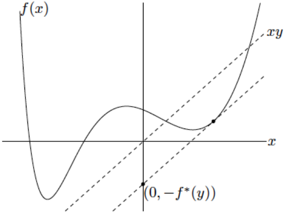
不論 f convex or not, f^* convex.

Example:

$f : \mathbf{R}^1 \to \mathbf{R}$, $f^* : \mathbf{R}^1 \to \mathbf{R}$

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

$\Rightarrow -f^*(y) = \min(f(x) - xy)$



Conjugate functions

- A conjugate function

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

is always convex.

  - $\because$ it is the pointwise supremum of a family of convex (indeed, affine) functions of y. $\because$ x and $f(x)$ are argument in y
  - This is true whether or not f is convex.
  - Note that when f is convex, the subscript $x \in \mathbf{dom}\ f$ is not necessary since $y^T x - f(x) = -\infty$ for $x \notin \mathbf{dom}\ f$.

Conjugate Functions – Examples for $f : \mathbf{R} \to \mathbf{R}$

$f: R \to R$, $f^*: R \to R$, $f^*(y) = \sup_{x \in \mathrm{dom} f} (xy - f(x)) = \sup (x(y-a)-b)$ $\begin{cases} y > a, & y \notin \mathrm{dom} f^* \\ y < a, & y \notin \mathrm{dom} f^* \end{cases}$ ($\because$ value $= \infty$)

- Affine function. $f(x) = ax + b$. The function $yx - ax - b$ is bounded if and only if $y = a$. Therefore **dom** $f^* = \{a\}$, and $f^*(a) = -b$. dom $f^* = \{-b\}$
- Negative logarithm. $f(x) = -\log x$, with **dom** $f = \mathbf{R}_{++}$. The function $xy + \log x$ is unbounded above if $y \geq 0$ and reaches its maximum at $x = -1/y$ otherwise. Therefore, **dom** $f^* = \{y \mid y < 0\} = -\mathbf{R}_{++}$ and $f^*(y) = -\log(-y) - 1$ for $y < 0$.
- Exponential. $f(x) = e^x$. The function $xy - e^x$ is unbounded above if $y < 0$. It can be shown that **dom** $f^* = \mathbf{R}_+$ and

$$f^*(y) = \begin{cases} y \log y - y, & y > 0 \\ 0, & y = 0 \end{cases}.$$

2. $f^*(y) = \sup\limits_{x \in R_{++}} (xy + \log x) \Rightarrow \begin{cases} y \geq 0: & \text{not in dom } f^*(y) \\ y < 0: & \text{differentiate: } y + \frac{1}{x} = 0, \\ & x = -y^{-1} \end{cases}$

3. $f^*(y) = \sup\limits_{x \in R} (xy - f(x)) \Rightarrow \begin{cases} y < 0: & \text{unbounded } (x \text{ 取} < 0, \text{as well}) \\ y \geq 0: & e^x \text{ 跑更快} \end{cases}$

Conjugate Functions – Examples for $f : \mathbf{R} \to \mathbf{R}$

$y-1-\log x = 0, \quad x = e^{y-1}$

- Negative entropy. $f(x) = x \log x$, with **dom** $f = \mathbf{R}_+$ (and $f(0) = 0$). The function $xy - x \log x$ is bounded above on $\mathbf{R}_+$ for all y, hence **dom** $f^* = \mathbf{R}$. It attains its maximum at $x = e^{y-1}$, and substituting we find $f^*(y) = e^{y-1}$.
- Inverse. $f(x) = 1/x$ on $\mathbf{R}_{++}$. For $y > 0$, $yx - 1/x$ is unbounded above. For $y = 0$, this function has supremum 0; for $y < 0$, the supremum is attained at $x = (-y)^{-1/2}$. Therefore we have $f^*(y) = -2(-y)^{1/2}$, with **dom** $f^* = -\mathbf{R}_+$.

Conjugate Functions – Examples for $f : \mathbf{R}^n \to \mathbf{R}$

Q : PSD

- Strictly convex quadratic function. Consider $f(x) = \frac{1}{2}x^T Qx$, with $Q \in \mathbf{S}^n_{++}$. The function $y^T x - \frac{1}{2}x^T Qx$ is bounded above as a function of x for all y. It attains its maximum at $x = Q^{-1}y$, so 必有上界 ∵ Q is PSD

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

$$f^*(y) = \sup_{x \in \mathrm{dom}\, x} \left(y^T x - \underbrace{\tfrac{1}{2}x^T Q x}_{\geq 0} \right)$$

$\Downarrow$ x後面 [illegible] 有上界

$\nabla(\cdot) = y - Qx = 0,$

$x = Q^{-1}y$

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

- Log-sum-exp function. Consider

$$f(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right).$$

Then, $f^*(y) = \sum_{i=1}^{n} y_i \log y_i$ with

$$\mathbf{dom}\ f^* = \left\{ y \mid \mathbf{1}^T y = 1, y \succeq 0 \right\}.$$

Conjugate Functions – Examples for $f : \mathbf{R}^n \to \mathbf{R}$

- Norm. Let $||\cdot||$ be a norm on $\mathbf{R}^n$, with dual norm $||\cdot||_*$. We will show that the conjugate of $f(x) = ||x||$ is

$$f^*(y) = \begin{cases} 0, & ||y||_* \leq 1 \\ \infty, & \text{otherwise} \end{cases},$$

  i.e., the conjugate of a norm is the indicator function of the dual norm unit ball.
- The definition of the dual norm of a given norm is defined in the following pages.

Introduction to Dual Norms (1/3)

- Let $||\cdot||$ be a norm on $\mathbf{R}^n$. The associated **dual norm**, denoted $||\cdot||_*$, is defined as

$$||z||_* = \sup\left\{z^T x \mid ||x|| \leq 1\right\}.$$

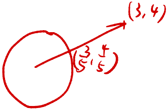
- It can be shown that

$$||z||_* = \sup\left\{|z^T x| \mid ||x|| \leq 1\right\}$$

max value 會出現在 bound 上

and

$$||z||_* = \sup_{x\neq 0} \frac{z^T x}{||x||}.$$



- A dual norm is also a norm (why?).
  - *Hint*: $||u+v||_* = \sup\left\{(u+v)^T x \mid ||x|| \leq 1\right\} \leq \sup\{u^T x \mid ||u|| \leq 1\} + \sup\{v^T x \mid ||v|| \leq 1\}$

Introduction to Dual Norms (2/3)

- From the definition of dual norm we have the inequality

$$z^T x \leq ||x|| \ ||z||_*,$$

  for all x and z.
- The dual of the dual norm is the original norm: we have $||x||_{**} = ||x||$ for all x.
  - *Hint*: $||x||_{**} = \sup_{z \neq 0} \frac{x^T z}{||z||_*}$
- The dual of the Euclidean norm is the Euclidean norm, since $\sup \left\{ z^T x \mid ||x||_2 \leq 1 \right\} = ||z||_2$.
  - This follows from the Cauchy-Schwarz inequality.
  - For nonzero z, the value of x that maximizes $z^T x$ over $||x||_2 \leq 1$ is $z/||z||_2$.

Introduction to Dual Norms (3/3)

- The dual of the ℓ_∞-norm is the ℓ_1-norm:

$$\sup\left\{z^T x \mid ||x||_\infty \leq 1\right\} = \sum_{i=1}^{n} |z_i| = ||z||_1.$$

- The dual of the ℓ_1-norm is the ℓ_∞-norm.
- More generally, the dual of the ℓ_p-norm is the ℓ_q-norm, where q satisfies

$$\frac{1}{p} + \frac{1}{q} = 1,$$

  i.e., $q = p/(p-1)$.
  - *Hint*: Hölder's inequality: $u^T v \leq ||u||_p ||v||_q$.

Conjugate Functions – Examples for $f : \mathbf{R}^n \to \mathbf{R}$

- Come back to the example of the conjugate function of a norm. Let $||\cdot||$ be a norm on $\mathbf{R}^n$, with dual norm $||\cdot||_*$. We now show that the conjugate of $f(x) = ||x||$ is

$$f^*(y) = \begin{cases} 0, & ||y||_* \leq 1 \\ \infty, & \text{otherwise} \end{cases}.$$

- Proof: If $||y||_* > 1$, then by definition of the dual norm, there is a $z \in \mathbf{R}^n$ with $||z|| \leq 1$ and $y^T z > 1$. Taking $x = tz$ and letting $t \to \infty$, we have $y^T x - ||x|| = t(y^T z - ||z||) \to \infty$, which shows that $f^*(y) = \infty$.
- Conversely, if $||y||_* \leq 1$, then we have $y^T x \leq ||x|| ||y||_*$ for all x, which implies for all x, $y^T x - ||x|| \leq 0$. Therefore $x = 0$ is the value that maximizes $y^T x - ||x||$, with maximum value 0.

Quasiconvex functions

不那麼 convex 的 function

Quasiconvex functions

A function $f : \mathbf{R}^n \to \mathbf{R}$ is called quasiconvex if its domain and all its sublevel sets

$$S_\alpha = \{x \in \mathbf{dom}\ f \mid f(x) \leq \alpha\},$$

for $\alpha \in \mathbf{R}$, are convex sets.

Convex set's sublevel set $\Rightarrow$ Convex set

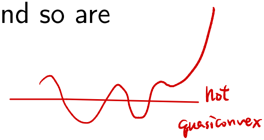
不一定 $\Leftarrow$

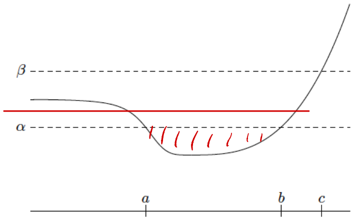
$\Downarrow$

quasiconvex

Convex functions are quasiconvex functions

- For a function on **R**, quasiconvexity requires that each sublevel set be an interval (either a finite-length interval or an infinite interval).
- Convex functions have convex sublevel sets, and so are quasiconvex. But the converse is not true.





f_β

f_α 的 sublevel set

Quasiconcave and quasilinear functions

Quasiconcave and quasilinear functions

- A function is **quasiconcave** if $-f$ is quasiconvex, i.e., every superlevel set $\{x|f(x) \geq \alpha\}$ is convex. $f(x) = \alpha$: level set
- A function that is both quasiconvex and quasiconcave is called **quasilinear**. monotone increasing

- If a function f is quasilinear, then its domain, and every **level set** $\{x \mid f(x) = \alpha\}$ is convex.

同时 convex, concave ⇒ affine.

quasi ——— ⇒ monotone increasing

Quasiconvex functions – Examples

Some examples on $\mathbf{R}$:

- Logarithm. $\log x$ on $\mathbf{R}_{++}$ is quasiconvex (and quasiconcave, hence quasilinear).
- Ceiling function. $\mathrm{ceil}(x) = \inf\{z \in Z \mid z \geq x\}$ is quasiconvex (and quasiconcave).

An example on $\mathbf{R}^n$:

- The **length** of $x \in \mathbf{R}^n$, defined as the largest index of a nonzero component, i.e.,

$$f(x) = \begin{cases} \max\{i \mid x_i \neq 0\} & x \neq 0 \\ 0 & x = 0 \end{cases},$$

is quasiconvex.

Quasiconvex functions – Examples

- Consider $f : \mathbf{R}^2 \to \mathbf{R}$, with $\mathbf{dom}\ f = \mathbf{R}^2_+$ and $f(x_1, x_2) = x_1 x_2$. Then, f is neither convex nor concave since

$$\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

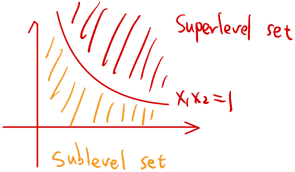
has eigenvalues ± 1 (not definite).

Convex / concave : eigenvalue non-nega / non-posi

- But f is quasiconcave on $\mathbf{R}^2_+$, since the superlevel sets

$$\{x \in \mathbf{R}^2_+ \mid x_1 x_2 \geq \alpha\}$$

are convex sets for all α.

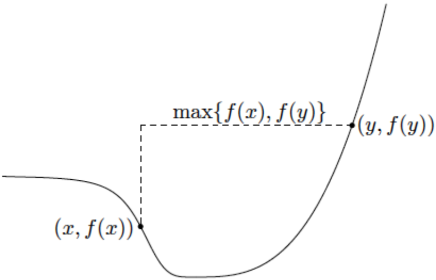


Quasiconvex functions – Basic Properties

Jensen's inequality for quasiconvex functions

A function f is quasiconvex if and only if **dom** f is convex and for any $x, y \in$ **dom** f and $0 \leq \theta \leq 1$,

$$f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}.$$

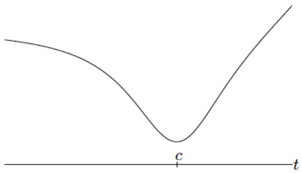


Quasiconvex functions – Basic Properties

Continuous quasiconvex functions on $\mathbf{R}$

A continuous function $f : \mathbf{R} \to \mathbf{R}$ is quasiconvex if and only if at least one of the following conditions holds:

- f is nondecreasing.
- f is nonincreasing.
- There is a point $c \in \mathbf{dom}\ f$ such that for $t \leq c$ (and $t \in \mathbf{dom}\ f$), f is nonincreasing, and for $t \geq c$ (and $t \in \mathbf{dom}\ f$), f is nondecreasing.



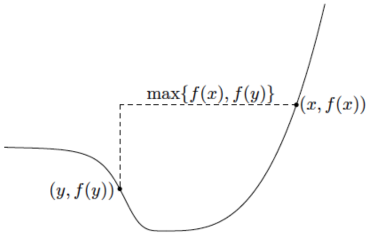
Differentiable quasiconvex functions

First-Order Conditions

Suppose $f : \mathbf{R}^n \to \mathbf{R}$ is differentiable. Then f is quasiconvex if and only if $\mathbf{dom}\ f$ is convex and for all $x, y \in \mathbf{dom}\ f$

$$f(y) \leq f(x) \Rightarrow \nabla f(x)^T (y - x) \leq 0.$$

Proof idea: It suffices to prove the result for a function on $\mathbf{R}$; the general result follows by restriction to an arbitrary line.



Representation via family of convex functions

Representation via family of convex functions

We can always find a family of convex functions $\phi_t : \mathbf{R}^n \to \mathbf{R}$, indexed by $t \in \mathbf{R}$, with

$$f(x) \leq t \iff \phi_t(x) \leq 0,$$

i.e., the t-sublevel set of the quasiconvex function f is the 0-sublevel set of the convex function ϕ_t.

- Evidently ϕ_t must satisfy the property that for all $x \in \mathbf{R}^n$, $\phi_t(x) \leq 0 \Rightarrow \phi_s(x) \leq 0$ for $s \geq t$. This is satisfied if for each x, $\phi_t(x)$ is a nonincreasing function of t, i.e., $\phi_s(x) \leq \phi_t(x)$ whenever $s \geq t$.
- One (straightforward) example:

$$\phi_t(x) = \begin{cases} 0, & f(x) \leq t \\ \infty, & \text{otherwise} \end{cases}.$$

- Another example: if the sublevel sets of f are closed, we can take

$$\phi_t(x) = \mathbf{dist}\,(x, \{z \mid f(z) \leq t\}).$$

We are usually interested in a family ϕ_t with nice properties, such as differentiability.