8. Conjugate functions

- closed functions
- conjugate function
Closed set

a set $C$ is closed if it contains its boundary:

$$x^k \in C, \quad x^k \to \bar{x} \implies \bar{x} \in C$$

operations that preserve closedness

- the intersection of (finitely or infinitely many) closed sets is closed
- the union of a finite number of closed sets is closed
- inverse under linear mapping: $\{x \mid Ax \in C\}$ is closed if $C$ is closed
Image under linear mapping

the image of a closed set under a linear mapping is not necessarily closed

**example** \((C\) is closed, \(AC = \{Ax \mid x \in C\}\) is open):

\[
C = \{(x_1, x_2) \in \mathbb{R}_+^2 \mid x_1x_2 \geq 1\}, \quad A = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad AC = \mathbb{R}_{++}
\]

**sufficient condition:** \(AC\) is closed if

- \(C\) is closed and convex
- and \(C\) does not have a recession direction in the nullspace of \(A\). i.e.,

\[
Ay = 0, \quad \hat{x} \in C, \quad \hat{x} + \alpha y \in C \quad \forall \alpha \geq 0 \implies y = 0
\]

in particular, this holds for any \(A\) if \(C\) is bounded
Closed function

**definition:** a function is closed if its epigraph is a closed set

**examples**

- \( f(x) = - \log(1 - x^2) \) with \( \text{dom } f = \{x \mid |x| < 1\} \)
- \( f(x) = x \log x \) with \( \text{dom } f = \mathbb{R}_+ \) and \( f(0) = 0 \)
- indicator function of a closed set \( C: f(x) = 0 \) if \( x \in C = \text{dom } f \)

**not closed**

- \( f(x) = x \log x \) with \( \text{dom } f = \mathbb{R}_{++} \), or with \( \text{dom } f = \mathbb{R}_+ \) and \( f(0) = 1 \)
- indicator function of a set \( C \) if \( C \) is not closed
Properties

**sublevel sets:** $f$ is closed if and only if all its sublevel sets are closed

**minimum:** if $f$ is closed with bounded sublevel sets then it has a minimizer

**common operations on convex functions that preserve closedness**

- **sum:** $f + g$ is closed if $f$ and $g$ are closed (and $\text{dom } f \cap \text{dom } g \neq \emptyset$)

- **composition with affine mapping:** $f(Ax + b)$ is closed if $f$ is closed

- **supremum:** $\sup_{\alpha} f_{\alpha}(x)$ is closed if each function $f_{\alpha}$ is closed
Outline

- closed functions
- conjugate function
Conjugate function

the **conjugate** of a function $f$ is

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

$f^*$ is closed and convex even if $f$ is not

**Fenchel’s inequality**

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

(extends inequality $x^T x/2 + y^T y/2 \geq x^T y$ to non-quadratic convex $f$)
Quadratic function

\[ f(x) = \frac{1}{2} x^T A x + b^T x + c \]

strictly convex case \((A \succ 0)\)

\[ f^*(y) = \frac{1}{2} (y - b)^T A^{-1} (y - b) - c \]

general convex case \((A \succeq 0)\)

\[ f^*(y) = \frac{1}{2} (y - b)^T A^\dagger (y - b) - c, \quad \text{dom } f^* = \text{range}(A) + b \]
negative entropy

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

negative logarithm

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

matrix logarithm

\[ f(X) = - \log \det X \quad (\text{dom } f = S^{++}_n) \quad f^*(Y) = - \log \det(-Y) - n \]

Conjugate functions
Indicator function and norm

**Indicator** of convex set $C$: conjugate is support function of $C$

$$f(x) = \begin{cases} 0 & x \in C \\ +\infty & x \notin C \end{cases} \quad f^*(y) = \sup_{x \in C} y^T x$$

**Norm**: conjugate is indicator of unit dual norm ball

$$f(x) = \|x\| \quad f^*(y) = \begin{cases} 0 & \|y\|_* \leq 1 \\ +\infty & \|y\|_* > 1 \end{cases}$$

(see next page)
proof: recall the definition of dual norm:

\[ \|y\|_* = \sup_{\|x\| \leq 1} x^T y \]

to evaluate \( f^*(y) = \sup_x (y^T x - \|x\|) \) we distinguish two cases

- if \( \|y\|_* \leq 1 \), then (by definition of dual norm)
  \[ y^T x \leq \|x\| \quad \forall x \]
  and equality holds if \( x = 0 \); therefore \( \sup_x (y^T x - \|x\|) = 0 \)

- if \( \|y\|_* > 1 \), there exists an \( x \) with \( \|x\| \leq 1 \), \( x^T y > 1 \); then
  \[ f^*(y) \geq y^T (tx) - \|tx\| = t(y^T x - \|x\|) \]
  and r.h.s. goes to infinity if \( t \to \infty \)
The second conjugate

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

- \( f^{**} \) is closed and convex
- from Fenchel’s inequality \((x^T y - f^*(y) \leq f(x) \) for all \( y \) and \( x \)):
  \[ f^{**}(x) \leq f(x) \quad \forall x \]
  equivalently, \( \text{epi } f \subseteq \text{epi } f^{**} \) (for any \( f \))
- if \( f \) is closed and convex, then
  \[ f^{**}(x) = f(x) \quad \forall x \]
  equivalently, \( \text{epi } f = \text{epi } f^{**} \) (if \( f \) is closed convex); proof on next page
proof \( (f^{**} = f \text{ if } f \text{ is closed and convex}) \): by contradiction

suppose \( (x, f^{**}(x)) \notin \text{epi } f \); then there is a strict separating hyperplane:

\[
\begin{bmatrix}
a \\
b \\
\end{bmatrix}^T \begin{bmatrix}
z - x \\
- s + f^{**}(x) \\
\end{bmatrix} \leq c < 0 \quad \forall (z, s) \in \text{epi } f
\]

for some \( a, b, c \) with \( b \leq 0 \) (\( b > 0 \) gives a contradiction as \( s \to \infty \))

- if \( b < 0 \), define \( y = a/(-b) \) and maximize l.h.s. over \((z, s) \in \text{epi } f\):

\[
f^*(y) - y^Tx + f^{**}(x) \leq c/(-b) < 0
\]

this contradicts Fenchel’s inequality

- if \( b = 0 \), choose \( \hat{y} \in \text{dom } f^* \) and add small multiple of \((\hat{y}, -1)\) to \( (a, b) \):

\[
\begin{bmatrix}
a + \epsilon \hat{y} \\
- \epsilon
\end{bmatrix}^T \begin{bmatrix}
z - x \\
- s + f^{**}(x) \\
\end{bmatrix} \leq c + \epsilon \left(f^*(\hat{y}) - x^T\hat{y} + f^{**}(x)\right) < 0
\]

now apply the argument for \( b < 0 \)
Conjugates and subgradients

If \( f \) is closed and convex, then

\[
y \in \partial f(x) \iff x \in \partial f^*(y) \iff x^T y = f(x) + f^*(y)
\]

**Proof:** if \( y \in \partial f(x) \), then \( f^*(y) = \sup_u (y^T u - f(u)) = y^T x - f(x) \)

\[
f^*(v) = \sup_u (v^T u - f(u)) \\
\geq v^T x - f(x) \\
= x^T (v - y) - f(x) + y^T x \\
= f^*(y) + x^T (v - y)
\]

for all \( v \); therefore, \( x \) is a subgradient of \( f^* \) at \( y \) \( (x \in \partial f^*(y)) \)

reverse implication \( x \in \partial f^*(y) \implies y \in \partial f(x) \) follows from \( f^{**} = f \)
Some calculus rules

separable sum

\[ f(x_1, x_2) = g(x_1) + h(x_2) \quad f^*(y_1, y_2) = g^*(y_1) + h^*(y_2) \]

scalar multiplication: (for \( \alpha > 0 \))

\[ f(x) = \alpha g(x) \quad f^*(y) = \alpha g^*(y/\alpha) \]

addition to affine function

\[ f(x) = g(x) + a^T x + b \quad f^*(y) = g^*(y - a) - b \]

infimal convolution

\[ f(x) = \inf_{u+v=x} (g(u) + h(v)) \quad f^*(y) = g^*(y) + h^*(y) \]
References

