Conjugate Functions, Optimality Conditions, and Duality

Pontus Giselsson

Learning goals

- Able to derive conjugate formulas
- Able to prove convexity of conjugate
- Know that biconjugate is convex envelope
- Know when $\partial f = (\partial f^*)^{-1}$ holds
- Know Fenchel-Young's inequality and when equality holds
- Know that strong convexity and smoothness are dual properties
- Formulate Fenchel dual problem in general and for examples
- Derive dual problem with primal-dual optimality conditions
- Able to recover primal solution
- Be aware of inf-sup interpretation in derivation of dual problem

Conjugate Functions

Conjugate functions

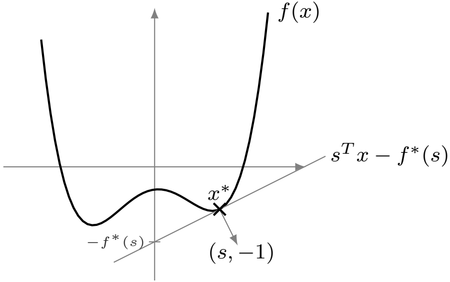
- The conjugate function of $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is defined as

$$f^*(s) := \sup_x \left(s^T x - f(x) \right)$$

- Implicit definition via optimization problem

Conjugate interpretation

- Conjugate $f^*(s)$ defines affine minorizer to f with slope s:



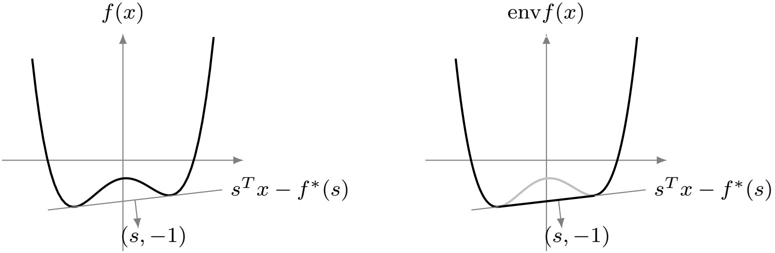
  where $f^*(s)$ decides the constant offset to have support at x^*
- "Affine minorizor generator: Pick slope s, get offset for support"
- Why? Consider $f^*(s) = \sup_x \left(s^Tx - f(x)\right)$ with maximizer x^*:

$$
\begin{aligned}
f^*(s) = s^Tx^* - f(x^*) \quad &\Leftrightarrow \quad f^*(s) \geq s^Tx - f(x) \text{ for all } x \\
&\Leftrightarrow \quad f(x) \geq s^Tx - f^*(s) \text{ for all } x
\end{aligned}
$$

- Support at x^* since $f(x^*) = s^Tx^* - f^*(s)$

Consequence

- Conjugate of f and $\mathrm{env} f$ are the same, i.e., $f^* = (\mathrm{env} f)^*$



- Functions have same supporting affine functions
- Epigraphs have same supporting hyperplanes

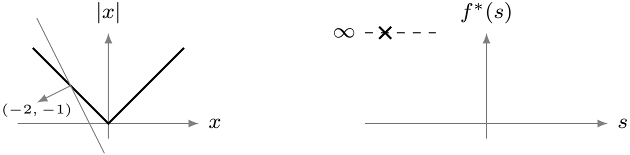
Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s = -2$ $f^*(s)$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis



Slope, $s = -2$ $\quad$ $f^*(s) \to \infty$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s = -1$ $f^*(s)$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s = -1$ $f^*(s) = 0$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s = -0.5$ $\quad f^*(s)$

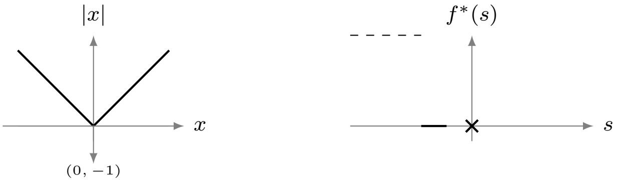
Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s = -0.5$ $\quad f^*(s) = 0$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis



Slope, $s = 0$ $f^*(s)$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s =0$ $f^*(s) = 0$

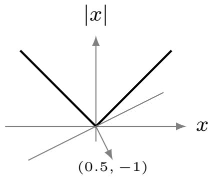
Conjugate function – Example

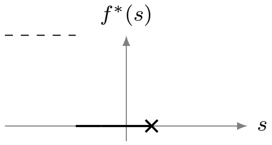
- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s =0.5$ $\qquad f^*(s)$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis





Slope, $s = 0.5$ $\quad f^*(s) = 0$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s = 1$ $\qquad$ $f^*(s)$

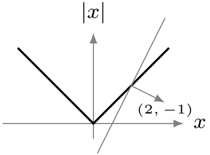
Conjugate function – Example

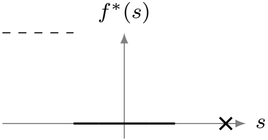
- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

Slope, $s = 1$ $\qquad$ $f^*(s) = 0$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis

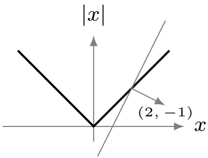


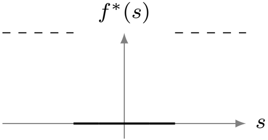


Slope, $s = 2$ $\qquad$ $f^*(s)$

Conjugate function – Example

- Compute conjugate of $f(x) = |x|$
- For given slope s: $-f^*(s)$ is point that crosses $|x|$-axis





Slope, s =2 $\quad f^*(s) \to \infty$

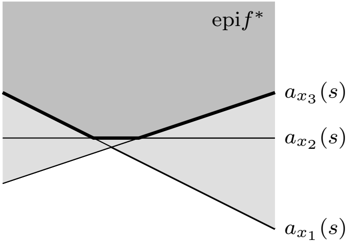
- Conjugate is $f^*(s) = \iota_{[-1,1]}(s)$

Conjugate function properties

- Let $a_x(s) := s^T x - f(x)$ be affine function parameterized by x:

$$f^*(s) = \sup_x a_x(s)$$

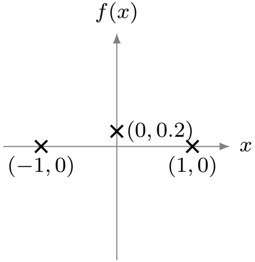
  is supremum of family of affine functions
- Epigraph of f^* is intersection of epigraphs of (below three) a_x



- f^* convex: epigraph intersection of convex halfspaces $\mathrm{epi}\, a_x$
- f^* closed: epigraph intersection of closed halfspaces $\mathrm{epi}\, a_x$

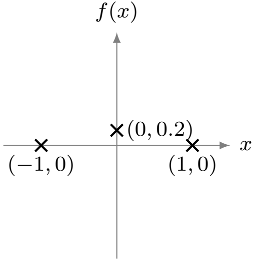
Draw the conjugate

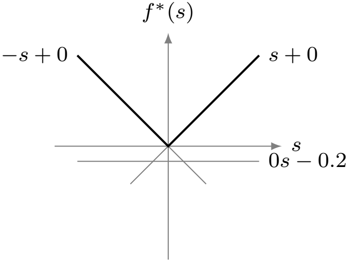
- Draw conjugate of f ($f(x) = \infty$ outside points)



Draw the conjugate

- Draw conjugate of f ($f(x) = \infty$ outside points)



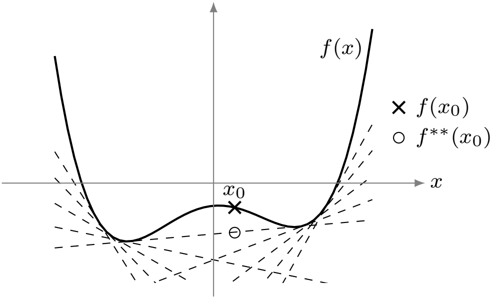


- Draw all affine $a_x(s)$ and select for each s the max to get $f^*(s)$

$$\begin{aligned} f^*(s) &= \sup_x (sx - f(x)) = \max(-s - 0, 0s - 0.2, s - 0, xs - \infty) \\ &= \max(-s, -0.2, s) = |s| \end{aligned}$$

Biconjugate

- Biconjuate $f^{**} := (f^*)^*$ is conjugate of conjugate
- For every x, it is largest value of all affine minorizers



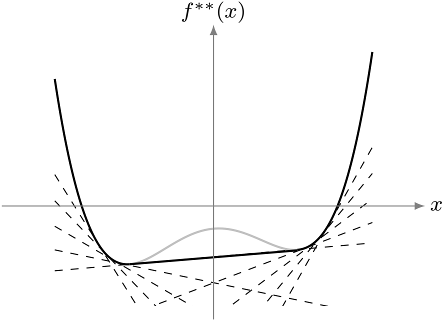
- Why?: Biconjugate

$$f^{**}(x) = \sup_s (x^T s - f^*(s)),$$

  - is pointwise supremum of affine functions $a_s(x) := x^T s - f^*(s)$
  - $\{a_s(x)\}_{s \in \mathbb{R}^n}$ are all supporting affine minorizers to f with slope s

Biconjugate and convex envelope

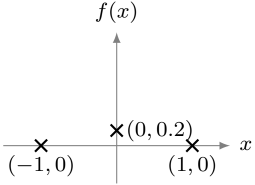
- Biconjugate is closed convex envelope



- $f^{**} \leq f$ and $f^{**} = f$ if and only if f (closed and) convex

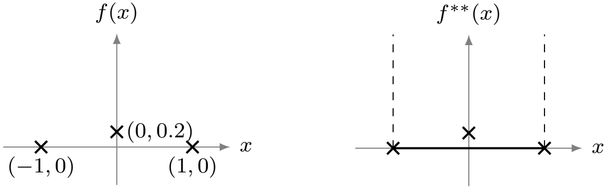
Biconjugate – Example

- Draw the biconjugate of f $\big(f(x) = \infty$ outside points$\big)$



Biconjugate – Example

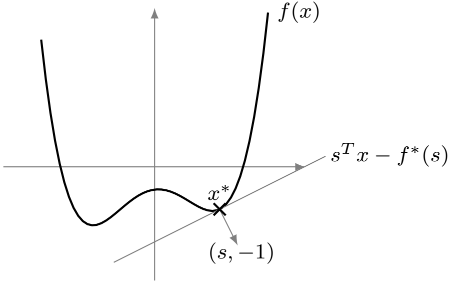
- Draw the biconjugate of f ($f(x) = \infty$ outside points)



- Biconjugate is convex envelope of f
- We found before $f^*(s) = |s|$, and now $(f^*)^*(x) = \iota_{[-1,1]}(x)$
- Therefore also $\iota^*_{[-1,1]}(s) = |s|$
  (since $f^* = (\mathrm{env} f)^* = (f^{**})^* =: f^{***}$)

Fenchel Young's equality

- Going back to conjugate interpretation:



- Fenchel's inequality: $f(x) \geq s^T x - f^*(s)$ for all x, s
- Fenchel-Young's equality and equivalence:

$$\boxed{f(x^*) = s^T x^* - f^*(s) \text{ holds if and only if } s \in \partial f(x^*)}$$

A subdifferential formula

$\boxed{\text{Assume } f \text{ closed convex, then } \partial f(x) = \mathrm{Argmax}_s (s^T x - f^*(s))}$

- Since $f^{**} = f$, we have $f(x) = \sup_s (x^T s - f^*(s))$ and

$$\begin{aligned} s^* \in \operatorname*{Argmax}_s (x^T s - f^*(s)) \quad &\iff \quad f(x) = x^T s^* - f^*(s^*) \\ &\iff \quad s^* \in \partial f(x) \end{aligned}$$

- The last equivalence is Fenchel-Young

Subdifferential of conjugate

$$\boxed{s \in \partial f(x) \text{ implies that } x \in \partial f^*(s)}$$

- Since $f^{**} \leq f$ and $s \in \partial f(x)$, Fenchel-Young's equality gives:

$$0 = f^*(s) + f(x) - s^T x \geq f^*(s) + f^{**}(x) - s^T x \geq 0$$

  where last step is Fenchel's inequality
- Hence $f^*(s) + f^{**}(x) - s^T x = 0$ and FY $\Rightarrow x \in \partial f^*(s)$
- Apply result above to f^* to get corollary:

$$\boxed{x \in \partial f^*(s) \text{ implies that } s \in \partial f^{**}(x)}$$

Subdifferential of conjugate – Inversion formula

$\boxed{\text{Suppose } f \text{ closed convex, then } s \in \partial f(x) \iff x \in \partial f^*(s)}$

- Using what we have on previous slide and $f^{**} = f$:

$$s \in \partial f(x) \Rightarrow x \in \partial f^*(s) \Rightarrow s \in \partial f^{**}(x) \Rightarrow s \in \partial f(x)$$

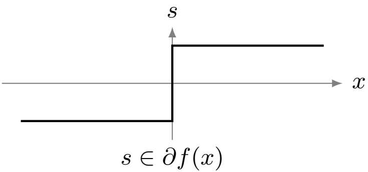
- Another way to write the result is that for closed convex f:

$$\partial f^* = (\partial f)^{-1}$$

(Definition of inverse of set-valued A: $x \in A^{-1}u \iff u \in Ax$)

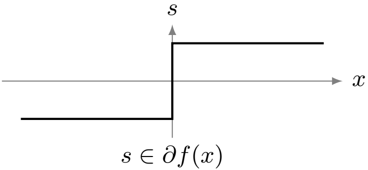
Relation between subdifferentials – Example

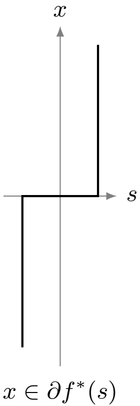
- What is ∂f^* for below ∂f?



Relation between subdifferentials – Example

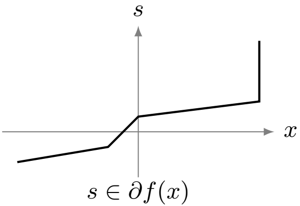
- What is ∂f^* for below ∂f?

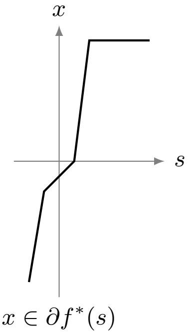




- Since $\partial f^* = (\partial f)^{-1}$, we flip the figure

Another example





- region with slope σ in $\partial f(x)$ $\Leftrightarrow$ region with slope $\frac{1}{\sigma}$ in $\partial f^*(s)$
- Implication: ∂f σ-strong monotone $\Leftrightarrow$ $\partial f^*(s)$ σ-cocoercive?
  (Recall: σ-cocoercivity $\Leftrightarrow$ $\frac{1}{\sigma}$-Lipschitz and monotone)

Cocoercivity and strong monotonicity

$$\boxed{\begin{array}{c}\partial f : \mathbb{R}^n \to 2^{\mathbb{R}^n} \text{ maximal monotone and } \sigma\text{-strongly monotone} \\ \Longleftrightarrow \\ \partial f^* = \nabla f^* : \mathbb{R} \to \mathbb{R} \text{ single-valued and } \sigma\text{-cocoercive}\end{array}}$$

- σ-strong monotonicity: for all $u \in \partial f(x)$ and $v \in \partial f(y)$

$$(u - v)^T(x - y) \geq \sigma\|x - y\|_2^2 \tag{1}$$

  or equivalently for all $x \in \partial f^*(u)$ and $y \in \partial f^*(v)$
- ∂f^* is single-valued:
  - Assume $x \in \partial f^*(u)$ and $y \in \partial f^*(u)$, then lhs of (1) 0 and $x = y$
- ∇f^* is σ-cocoercive: plug $x = \nabla f^*(u)$ and $y = \nabla f^*(v)$ into (1)
- That ∂f^* has full domain follows from Minty's theorem

Duality correspondance

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. Then the following are equivalent:

(i) f is closed and σ-strongly convex
(ii) ∂f is maximally monotone and σ-strongly monotone
(iii) ∇f^* is σ-cocoercive
(iv) ∇f^* is maximally monotone and $\frac{1}{\sigma}$-Lipschitz continuous
(v) f^* is closed convex and satisfies descent lemma (is $\frac{1}{\sigma}$-smooth)

where $\nabla f^* : \mathbb{R}^n \to \mathbb{R}^n$ and $f^* : \mathbb{R}^n \to \mathbb{R}$

Comments:

- (i) $\Leftrightarrow$ (ii) and (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v): Previous lecture
- (ii) $\Leftrightarrow$ (iii): This lecture
- Since $f = f^{**}$ the result holds with f and f^* interchanged
- Full proof available on course webpage

Proximal operator

- Recall: $\text{prox}_{\gamma g}(z) = \text{argmin}_x(g(x) + \frac{1}{2\gamma}\|x - z\|_2^2)$

$$\boxed{\text{Assume } g \text{ closed convex, then } \text{prox}_{\gamma g}(z) \text{ is 1-cocoercive}}$$

- The function $r = \gamma g + \frac{1}{2}\|\cdot\|_2^2$ is 1-strongly convex and

$$\begin{aligned}
\text{prox}_{\gamma g}(z) &= \text{argmin}(g(x) + \tfrac{1}{2\gamma}\|x - z\|_2^2) \\
&= \text{argmax}(-\gamma g(x) - \tfrac{1}{2}\|x - z\|_2^2) \\
&= \text{argmax}(z^T x - (\tfrac{1}{2}\|x\|_2^2 + \gamma g(x))) \\
&= \text{argmax}(z^T x - r(x)) \\
&= \nabla r^*(z)
\end{aligned}$$

  where we have used the subgradient formula for r^*
- Therefore $\text{prox}_{\gamma g} = \nabla r^*$ is 1-cocoercive

Moreau decomposition

$$\boxed{\text{Assume } f \text{ closed convex, then } \operatorname{prox}_f(z) + \operatorname{prox}_{f^*}(z) = z}$$

- When f scaled by $\gamma > 0$, it becomes

$$z = \operatorname{prox}_{\gamma f}(z) + \operatorname{prox}_{(\gamma f)^*}(z) = \operatorname{prox}_{\gamma f}(z) + \gamma \operatorname{prox}_{\gamma^{-1} f^*}(\gamma^{-1} z)$$

  (since $\operatorname{prox}_{(\gamma f)^*} = \gamma \operatorname{prox}_{\gamma^{-1} f^*} \circ \gamma^{-1}\mathrm{Id}$)
- Don't need to know f^* to compute $\operatorname{prox}_{\gamma f^*}$!

Optimality Conditions and Duality

Composite optimization problem

- Consider *primal* composite optimization problem

$$\text{minimize}\, f(Lx) + g(x) \tag{2}$$

  where f, g closed convex
- We will derive primal-dual optimality conditions and dual problem

Primal optimality condition

Let $f : \mathbb{R}^m \to \overline{\mathbb{R}}$, $g : \mathbb{R}^n \to \overline{\mathbb{R}}$, $L \in \mathbb{R}^{m \times n}$ with f, g closed convex and assume CQ, then:

$$\text{minimize } f(Lx) + g(x) \tag{1}$$

is solved by $x \in \mathbb{R}^n$ if and only if x satisfies

$$0 \in L^T \partial f(Lx) + \partial g(x) \tag{2}$$

- CQ implies subdifferential calculus with equality:

$$0 \in L^T \partial f(Lx) + \partial g(x) = \partial((f \circ L)(x) + g(x))$$

which by Fermat's rule is equivalent to x solution to (1)

Primal-dual optimality condition 1

- Introduce *dual* variable $\mu \in \partial f(Lx)$, then optimality condition

$$0 \in L^T \underbrace{\partial f(Lx)}_{\mu} + \partial g(x)$$

is equivalent to

$$\begin{aligned} \mu &\in \partial f(Lx) \\ -L^T \mu &\in \partial g(x) \end{aligned}$$

- This is a necessary and sufficient primal-dual optimality condition
- (*Primal-dual* since involves primal x and dual μ variables)

Primal-dual optimality condition 2

- Primal-dual optimality condition

$$\begin{aligned} \mu &\in \partial f(Lx) \\ -L^T\mu &\in \partial g(x) \end{aligned}$$

- Using subdifferential inverse:

$$\mu \in \partial f(Lx) \qquad \Longleftrightarrow \qquad Lx \in \partial f^*(\mu)$$

gives equivalent primal dual optimality condition

$$\begin{aligned} Lx &\in \partial f^*(\mu) \\ -L^T\mu &\in \partial g(x) \end{aligned}$$

Dual optimality condition

- Using subdifferential inverse on other condition

$$-L^T\mu \in \partial g(x) \qquad \iff \qquad x \in \partial g^*(-L^T\mu)$$

gives equivalent primal dual optimality condition

$$\begin{aligned} Lx &\in \partial f^*(\mu) \\ x &\in \partial g^*(-L^T\mu) \end{aligned}$$

- This is equivalent to that:

$$0 \in \partial f^*(\mu) - L\underbrace{\partial g^*(-L^T\mu)}_{x} = \partial f^*(\mu) + \partial(g^* \circ -L^T)(\mu)$$

which is a dual optimality condition since it involves only μ

Dual problem

- The dual optimality condition (for solving primal problem)

$$0 \in \partial f^*(\mu) + \partial(g^* \circ -L^T)(\mu) \tag{1}$$

is sufficient optimality condition for dual problem:

$$\min_{\mu} f^*(\mu) + g^*(-L^T\mu) \tag{2}$$

- If contraint qualification holds on *dual* problem (2):

$$\mathrm{relint}\,\mathrm{dom}(g^* \circ -L^T) \cap \mathrm{relint}\,\mathrm{dom} f^* \neq \emptyset,$$

which we call CQ-D, we have equivalence also in last step

Equivalence not needed in last step since (2) *is solved via* (1)*, which has solution.*

Optimality conditions – Summary

- Assume f, g closed convex and that CQ holds
- Problem $\min_x f(Lx) + g(x)$ is solved by x iff

$$0 \in L^T \partial f(Lx) + \partial g(x)$$

- Primal dual necessary and sufficient optimality conditions:

$$\begin{cases} \mu \in \partial f(Lx) \\ -L^T\mu \in \partial g(x) \end{cases} \qquad \begin{cases} Lx \in \partial f^*(\mu) \\ -L^*\mu \in \partial g(x) \end{cases}$$

$$\begin{cases} \mu \in \partial f(Lx) \\ x \in \partial g^*(-L^T\mu) \end{cases} \qquad \begin{cases} Lx \in \partial f^*(\mu) \\ x \in \partial g^*(-L^T\mu) \end{cases}$$

- Dual optimality condition

$$0 \in \partial f^*(\mu) + \partial(g^* \circ -L^T)(\mu) \tag{1}$$

solves dual problem $\min_\mu f^*(\mu) + g^*(-L^T\mu)$
- If CQ-D holds, all dual problem solutions satisfy (1)

Solving the primal

- We of course want to solve primal problem

$$\underset{x}{\text{minimize}}\, f(Lx) + g(x)$$

- Can be solved via primal, primal-dual, or dual optimality condition
- In this course consider only solving via primal or dual condition:

$$0 \in \partial f^*(\mu) + \partial(g^* \circ -L^T)(\mu)$$

- Why solve dual? Sometimes easier to solve than primal
- Caveat: Only useful if primal solution can be obtained from dual

Solving the primal from the dual

- Assume f, g closed convex and CQ
- Assume optimal dual μ known: $0 \in \partial f^*(\mu) + \partial(g^* \circ -L^T)(\mu)$
- Optimal primal x must satisfy any and all primal-dual conditions:

$$\begin{cases} \mu \in \partial f(Lx) \\ -L^T\mu \in \partial g(x) \end{cases} \qquad \begin{cases} Lx \in \partial f^*(\mu) \\ -L^*\mu \in \partial g(x) \end{cases}$$

$$\begin{cases} \mu \in \partial f(Lx) \\ x \in \partial g^*(-L^T\mu) \end{cases} \qquad \begin{cases} Lx \in \partial f^*(\mu) \\ x \in \partial g^*(-L^T\mu) \end{cases}$$

- If one of these uniquely characterizes x, then must be solution:
  - ∂g^* is differentiable at $-L^T\mu$ for dual solution μ
  - ∂f^* is differentiable at dual solution μ and L invertible
  - $\ldots$

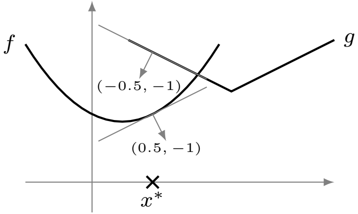
A dual problem interpretation

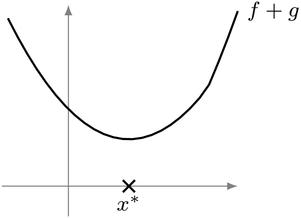
- Let $L = I$, consider dual problem $\min_\mu f^*(\mu) + g^*(-\mu)$
- Given CQ-D, μ is solution to dual if and only if

$$\begin{cases} \mu \in \partial f(x) \\ -\mu \in \partial g(x) \end{cases}$$

where x is a primal solution (x^* in figure below)

- "Dual problem searches subgradients of f and g that sum to 0"





- To solve primal, must find corresponding point x^*

Fenchel duality - A minmax formulation

- Write the problem $\min_x f(Lx) + g(x)$ on equivalent form

$$\begin{array}{ll} \text{minimize} & f(y) + g(x) \\ \text{subject to} & Lx = y \end{array}$$

- Equivalent formulation with indicator functions:

$$\text{minimize} \quad f(y) + g(x) + \iota_{\{0\}}(Lx - y)$$

where the indicator function is defined as

$$\iota_{\{0\}}(Lx - y) = \begin{cases} 0 & \text{if } Lx - y = 0 \\ \infty & \text{else} \end{cases}$$

Reformulation

- We can show (an exercise) that:

$$\iota_{\{0\}}(x,y) = \sup_{\mu} \mu^T (Lx - y)$$

(this μ is the same as the μ in previous dual formulation)

- Therefore problem is equivalent to

$$\inf_{x,y} \left(f(y) + g(x) + \sup_{\mu} \mu^T (Lx - y) \right)$$

or equivalently

$$\inf_{x,y} \sup_{\mu} \big(f(y) + g(x) + \mu^T (Lx - y) \big)$$

Fenchel weak duality

- We always have:

$$
\begin{aligned}
&\inf_x (f(Lx) + g(x)) \\
&\quad = \inf_{x,y} \sup_\mu (f(y) + g(x) + \mu^T(Lx - y)) \\
&\quad \geq \sup_\mu \inf_{x,y} \big(f(y) + g(x) + \mu^T(Lx - y)\big) \\
&\quad = \sup_\mu -\big(\sup_{x,y} \big(-f(y) - g(x) + \mu^T(-Lx + y)\big)\big) \\
&\quad = \sup_\mu \big(- \big(\sup_y \big(y^T\mu - f(y)\big) + \sup_x \big(x^T(-L^T\mu) - g(x)\big)\big)\big) \\
&\quad = \sup_\mu \big(-f^*(\mu) - g^*(-L^T\mu)\big),
\end{aligned}
$$

  which is (concave negative) dual problem from before

- This is called *weak duality*

Fenchel strong duality

Let $f : \mathbb{R}^m \to \overline{\mathbb{R}}$, $g : \mathbb{R}^n \to \overline{\mathbb{R}}$, $L \in \mathbb{R}^{m \times n}$ with f, g closed convex and assume CQ, then:

$$\inf_x (f(Lx) + g(x)) = \max_\mu \left(-f^*(\mu) - g^*(-L^T\mu)\right)$$

- A dual solution exists and optimal values coincide
- Proof steps:
  - Show that solution set to dual is compact under assumption
  - Use Sion's minimax theorem to have equality on previous slide
- Slight generalization useful to show subdifferential calculus rules

Lagrange duality

- Lagrange duality can be derived from Fenchel duality and vice versa
- KKT conditions in Lagrange duality can be derived from optimality conditions in this lecture

Conjugate examples

Conjugate – Example 1

Let $g(x) = \frac{1}{2}x^T Hx + h^T x$ with H positive definite (invertible)

- Gradient satisfies $\nabla g(x) = Hx + h$
- Fermat's rule for $g^*(s) = \sup_x (s^T x - g(x))$:

$$0 = s - \nabla g(x) \quad \Leftrightarrow \quad 0 = Hx + h - s \quad \Leftrightarrow \quad x = H^{-1}(s - h)$$

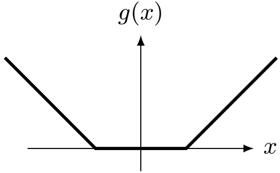
- So

$$\begin{aligned} g^*(s) &= s^T H^{-1}(s-h) - \tfrac{1}{2}(s-h)H^{-1}HH^{-1}(s-h) + h^T H^{-1}(s-h) \\ &= \tfrac{1}{2}(s-h)H^{-1}(s-h) \end{aligned}$$

Conjugate – Example 2

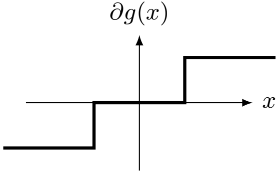
- Consider

$$g(x) = \begin{cases} -x-1 & \text{if } x \le -1 \\ 0 & \text{if } x \in [-1,1] \\ x-1 & \text{if } x \ge 1 \end{cases}$$



- Subdifferential satisfies

$$\partial g(x) = \begin{cases} -1 & \text{if } x < -1 \\ [-1,0] & \text{if } x = -1 \\ 0 & \text{if } x \in (-1,1) \\ [0,1] & \text{if } x = 1 \\ 1 & \text{if } x > 1 \end{cases}$$



Conjugate – Example 2 cont'd

- We use $g^*(s) = sx - g(x)$ if $s \in \partial g(x)$:
  - $x < -1$: $s = -1$, hence $g^*(-1) = -1x - (-x-1) = 1$
  - $x = -1$: $s \in [-1, 0]$ hence $g^*(s) = -s - 0 = -s$
  - $x \in (-1, 1)$: $s = 0$ hence $g^*(0) = 0x - 0 = 0$
  - $x = 1$: $s \in [0, 1]$ hence $g^*(s) = s - 0 = s$
  - $x > 1$: $s = 1$ hence $g^*(1) = x - (x-1) = 1$
- That is

$$g^*(s) = \begin{cases} -s & \text{if } s \in [-1, 0] \\ s & \text{if } s \in [0, 1] \end{cases}$$

- For $s < -1$ and $s > 1$, $g^*(s) = \infty$:
  - $s < -1$: let $x = t \to -\infty$ and $g^*(s) \geq ((s+1)t + 1) \to \infty$
  - $s > 1$: let $x = t \to \infty$ and $g^*(s) \geq ((s-1)t + 1) \to \infty$