Convex Programs, Duality, and Optimality Conditions

Yi Zhang

Outline

- Duals of norms, cones, and functions
- Convex programs and duality
- Strong duality and Slater’s conditions
- KKT optimality conditions

Outline

- Duals of norms, cones, and functions
- Convex programs and duality
- Strong duality and Slater's conditions
- KKT optimality conditions

Dual …, dual …, dual ...

- Dual norms
- Dual cones
- Dual functions (i.e., ***conjugate*** functions)
- Lagrange duals

Dual norms

- The dual norm $\|\ \|_*$ of $\|\ \|$ is:

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

  - Examples

$$\sup\{z^T x \mid \|x\|_2 \leq 1\} = \|z\|_2.$$

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

Dual norms

- The dual norm $\| \|_*$ of $\| \|$ is:

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

  - Examples

$$\sup\{z^T x \mid \|x\|_2 \leq 1\} = \|z\|_2.$$

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

- Warnings
  - A norm $\| \|$ is a function
  - The dual norm is ***not*** the dual of this function
  - Will discuss later …

Dual cones

- A set *C* is a ***cone***, if for every $x \in K$ and $\theta \geq 0$

  We have $\theta x \in K$

  - E.g.,: second order cone

$$K = \{(x,t) \in \mathbf{R}^{n+1} \mid \|x\|_2 \leq t\}$$

Dual cones

- A set *C* is a ***cone***, if for every $x \in K \text{ and } \theta \geq 0$

  We have $\theta x \in K$

  - E.g.,: second order cone

$$K = \{(x,t) \in \mathbf{R}^{n+1} \mid \|x\|_2 \leq t\}$$

- The ***dual cone*** *K** of a cone *K:*

$$K^* = \{y \mid y^T x \geq 0 \text{ for all } x \in K\}$$

  - E.g.,

$$K = \mathbf{R}^n_+ \colon\ K^* = \mathbf{R}^n_+$$

$$K = \{(x,t) \mid \|x\|_1 \leq t\} \colon\ K^* = \{(x,t) \mid \|x\|_\infty \leq t\}$$

Conjugate (dual) function

- The ***conjugate*** *f** of a function *f* is:

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

  - Also called the *dual* of function *f*
  - Use *conjugate* function to avoid confusion

Conjugate (dual) function

- The ***conjugate*** *f** of a function *f* is:

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

  - E.g., given $f(x) = -\log x$

$$\begin{aligned} f^*(y) &= \sup_{x>0} (xy + \log x) \\ &= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

Conjugate (dual) function

- The ***conjugate*** *f** of a function *f* is:

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

- E.g., given $f(x) = (1/2)x^T Q x$, where $Q \in \mathbf{S}^n_{++}$

$$\begin{aligned} f^*(y) &= \sup_x (y^T x - (1/2)x^T Q x) \\ &= \frac{1}{2} y^T Q^{-1} y \end{aligned}$$

Conjugate (dual) function

- The ***conjugate*** *f** of a function *f* is:

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

- Dual norm $\|\ \|_*$ is not the dual of $\|\ \|$
  - Norm is a function: $f(x) = \|x\|$
  - Dual norm: $\|z\|_* = \sup\{z^T x \mid \|x\| \le 1\}.$
  - Conjugate:

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

$$= \begin{cases} 0 & \|y\|_* \le 1 \\ \infty & \text{otherwise} \end{cases}$$

Outline

- Duals of norms, cones, and functions
- Convex programs and duality
  - Lagrange dual function
  - Lagrange dual function and conjugate functions
  - Lagrange dual problem
- Strong duality and Slater's conditions
- KKT optimality conditions

Lagrange dual function

- A *general* optimization problem with $x \in \mathbf{R}^n$

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

  - With optimal value *p**
- The ***Lagrange L***

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

  - With $\lambda_i \ge 0$

Lagrange dual function

- The ***Lagrange dual function***

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right) \end{aligned}$$

  - With $\lambda_i \geq 0$
- $g(\lambda, \nu)$ is a lower bound: $g(\lambda, \nu) \leq p^\star$
  - For ***any*** feasible point $\tilde{x}$

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \leq L(\tilde{x}, \lambda, \nu) \leq f_0(\tilde{x})$$

Lagrange dual function

- Example: standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0 \end{array}$$

  - The *Lagrange*

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x + \nu^T (Ax - b) - \lambda^T x \\ &= -b^T \nu + (c + A^T \nu - \lambda)^T x \end{aligned}$$

  - The *Lagrange dual function*

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Outline

- Duals of norms, cones, and functions
- Convex programs and duality
  - Lagrange dual function
  - Lagrange dual function and conjugate functions
  - Lagrange dual problem
- Strong duality and Slater's conditions
- KKT optimality conditions

Lagrange dual and conjugate

- Problem: $$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d. \end{array}$$

$$\begin{aligned} g(\lambda, \nu) &= \inf_x \left(f_0(x) + \lambda^T (Ax - b) + \nu^T (Cx - d) \right) \\ &= -b^T \lambda - d^T \nu + \inf_x \left(f_0(x) + (A^T \lambda + C^T \nu)^T x \right) \\ &= -b^T \lambda - d^T \nu - \sup_x \left((-A^T \lambda - C^T \nu)^T x - f_0(x) \right) \\ &= -b^T \lambda - d^T \nu - f_0^*(-A^T \lambda - C^T \nu) \end{aligned}$$

Lagrange dual and conjugate

- Example: maximum entropy

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d. \end{array} \qquad \begin{array}{ll} \text{minimize} & f_0(x) = \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

$$g(\nu) \quad = \quad -b^T\lambda - d^T\nu - f_0^*(-A^T\lambda - C^T\nu)$$

Lagrange dual and conjugate

- Example: maximum entropy

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d. \end{array} \qquad \begin{array}{ll} \text{minimize} & f_0(x) = \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

$$g(\nu) \quad = \quad -b^T\lambda - d^T\nu - f_0^*(-A^T\lambda - C^T\nu)$$

$$f_0(x) = \sum_{i=1}^n x_i \log x_i$$

$$f_0^*(y) = \sum_{i=1}^n e^{y_i - 1}$$

($u \log u$ ➔ e^{v-1})

Lagrange dual and conjugate

- Example: maximum entropy

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d. \end{array} \qquad \begin{array}{ll} \text{minimize} & f_0(x) = \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

$$\begin{aligned} g(\nu) &= -b^T\lambda - d^T\nu - f_0^*(-A^T\lambda - C^T\nu) \\ &= -b^T\lambda - \nu - e^{-\nu-1}\sum_{i=1}^n e^{-a_i^T\lambda} \end{aligned}$$

$$f_0(x) = \sum_{i=1}^n x_i \log x_i$$

$$f_0^*(y) = \sum_{i=1}^n e^{y_i - 1}$$

$$(\; u \log u \;\rightarrow\; e^{v-1} \;)$$

Outline

- Duals of norms, cones, and functions
- Convex programs and duality
  - Lagrange dual function
  - Lagrange dual function and conjugate functions
  - Lagrange dual problem
- Strong duality and Slater's conditions
- KKT optimality conditions

The (Lagrange) dual problem

- The primal problem

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

- The (Lagrange) dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

  - where

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

The (Lagrange) dual problem

- Example: standard LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array}$$

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0 \end{array}$$

Outline

- Duals of norms, cones, and functions
- Convex programs and duality
- Strong duality and Slater’s conditions
- KKT optimality conditions

Weak and strong duality, and slater's conditions

- $$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$ $$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$
  - Primal optimal value *p**
  - Dual optimal value *d**
- Weak duality: *d** <= *p**
  - Always holds (convex or non-convex problems)
  - Why ? … $g(\lambda, \nu) \le p^\star$
    - For any feasible point $\tilde{x}$

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \le L(\tilde{x}, \lambda, \nu) \le f_0(\tilde{x})$$

Strong duality and slater's conditions

- Strong duality: *d** = *p**
  - Not true in general
  - Usually (but not always) true for convex prog.
- Slater's conditions for convex prog.
  - ***Sufficient*** for strong duality

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

  - There is a ***strictly*** feasible point → strong duality
    - Relax: linear inequalities need not hold strictly

Strong duality and slater's conditions

- Example: LP in inequality form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

  - Slater's conditions … no "strict" required
    - Any primal feasible point → strong duality
    - Any dual feasible point → strong duality
  - i.e., either primal or dual feasible → strong duality

Outline

- Duals of norms, cones, and functions
- Convex programs and duality
- Strong duality and Slater's conditions
- KKT optimality conditions

KKT optimality conditions

- minimize $f_0(x)$
  subject to $f_i(x) \le 0, \quad i = 1, \ldots, m$
  $h_i(x) = 0, \quad i = 1, \ldots, p$

  maximize $g(\lambda, \nu)$
  subject to $\lambda \succeq 0$

- KKT conditions
  - Primal constraints: $f_i(x) \le 0, \quad h_i(x) = 0$
  - Dual constraints: $\lambda \succeq 0$
  - Complimentary slackness: $\lambda_i f_i(x) = 0$
  - Vanished gradients of L w.r.t x

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0$$

KKT optimality conditions

- KKT conditions are
  - Necessary for strong duality
  - Sufficient & necessary for strong duality for convex prog.
- Very useful
  - Slater's conditions are sufficient conditions for convex prog., and hard to check
  - KKT conditions are S&N for convex prog., and easy to check

KKT Optimality Conditions

- QP with linear constraints

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & Ax = b, \end{array}$$

- KKT conditions

$$Ax^\star = b, \qquad Px^\star + q + A^T \nu^\star = 0,$$

- Obtain opt. solution by solving a linear system

$$\left[\begin{array}{cc} P & A^T \\ A & 0 \end{array}\right] \left[\begin{array}{c} x^\star \\ \nu^\star \end{array}\right] = \left[\begin{array}{c} -q \\ b \end{array}\right].$$

KKT optimality conditions

- Several examples in HW5
  - Non-convex prog.
  - Convex prog.
  - SOCP
  - QCQP

Summary

- Duals of norms, cones, and functions
  - Dual norm is not dual function of the norm
  - Dual func. is the indicator func. of dual norm
- Convex programs and duality
  - Dual func. is not Lagrange dual func.
  - But closely related
- Strong duality and Slater's conditions
  - Sufficient for strong duality in convex prog.
- KKT optimality conditions
  - Sufficient & necessary for strong duality in convex.prog.