Optimization of Communication Systems
EE 669

Chathuranga Weeraddana

March 2019

Lecture 14: Duality [1]

[1] The content in this lecture is based on § 5, ref-3 and S. Boyd's related lecture slides.

Why Duality?

- discuss !

Lagrangian

- standard formulation (not necessarily convex):

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

variables $x \in \mathbb{R}^n$, domain $\mathcal{D}$, optimal value $p^\star$

Lagrangian

- Lagrangian: $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, with $\mathbf{dom}\ L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$

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

  - weighted sum of objective and constraint functions
  - λ_i is Lagrange multiplier associated with $f_i(x) \leq 0$
  - ν_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrangian

- geometrically explain the idea with cost functions :

$$I_i(y) = \begin{cases} 0 & y \leq 0 \\ \infty & \text{otherwise} \end{cases} \qquad i = 1, \ldots, m$$

$$J_i(y) = \begin{cases} 0 & y = 0 \\ \infty & \text{otherwise} \end{cases} \qquad j = 1, \ldots, p$$

- compare above with

$$I_i(y; \lambda_i) = \lambda_i y \qquad i = 1, \ldots, m$$

$$J_i(y; \nu_i) = \nu_i y \qquad j = 1, \ldots, p$$

Lagrange Dual Function

- Lagrange dual function: $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

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

  g is concave, can be $-\infty$ for some (λ, ν)

- Lower bound property: $\lambda \succeq 0 \implies g(\lambda, \nu) \leq p^\star$

- proof?

Least-Norm Solution of Linear Equations

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

Dual function:

- Lagrangian is $L(x, \nu) = x^{\mathsf{T}}x - \nu^{\mathsf{T}}(Ax - b)$
- to minimize L over x, set gradient equals to zero

$$\nabla_x L(x, \nu) = 2x + A^{\mathsf{T}}\nu = 0 \implies x = -(1/2)A^{\mathsf{T}}\nu$$

- plug x in L to obtain g:

$$g(\nu) = L\left(-(1/2)A^{\mathsf{T}}\nu, \nu\right) = -(1/4)\nu^{\mathsf{T}}AA^{\mathsf{T}}\nu - b^{\mathsf{T}}\nu, \text{ concave in } \nu$$

- lower bound property?

Linear Programming

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

Dual function:

- Lagrangian is

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



- L is linear in x, hence

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

g is linear on affine domain $\{(\lambda, \nu) \mid A^{\mathsf{T}} \nu - \lambda + c = 0\}$

- lower bound property?

Quadratic Programming

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

assume $P \in S^n_{++}$

- Lagrangian?
- Dual function?

The Dual Problem

Lagrange dual problem

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

- finds best lower bound on $p^\star$, obtained from Lagrange dual function
- a convex optimization problem; optimal value denoted $d^\star$
- λ, ν are dual feasible if $\lambda \succeq 0$ and $(\lambda, \nu) \in \mathbf{dom}\ g$
- often simplified by making implicit constraint $(\lambda, \nu) \in \mathbf{dom}\ g$ explicit
- write the dual problem for the problems we discussed above

Weak and Strong Duality

weak duality: $d^\star \leq p^\star$

- always holds (for convex and nonconvex problems)
- can be used to find nontrivial lower bounds for difficult problems

strong duality: $d^\star = p^\star$

- does not hold in general
- (usually) holds for convex problems
- conditions that guarantee strong duality in convex problems are called constraint qualiflcations
- there exist many types of const. qualiflcations (e.g., Slater's)

Complementary Slackness

assume strong dualityholds, $x^\star$ is primal optimal, $(\lambda^\star, \nu^\star)$ is dual optimal

$$\begin{aligned} f_0(x^\star) = g(\lambda^\star, \nu^\star) &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^\star f_i(x) + \sum_{i=1}^p \nu_i^\star h_i(x) \right) \\ &\leq f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star f_i(x^\star) + \sum_{i=1}^p \nu_i^\star h_i(x^\star) \\ &\leq f(x^\star) \end{aligned}$$

hence, the two inequalities hold with equality

- consequences?

Complementary Slackness

- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$
- $\lambda_i^\star f_i(x^\star) = 0 \;\; \forall \; i$, (known as complementary slackness)
- from complementary slackness, we have
  $\lambda_i^\star > 0 \implies f_i(x^\star) = 0, \qquad f_i(x^\star) < 0 \implies \lambda_i^\star = 0$

Karush-Kuhn-Tucker (KKT) Conditions

the following four conditions are called KKT conditions (for a problem with difierentiable f_i, h_i):

1. primal constraints: $f_i(x) \leq 0$, $i = 1, \ldots, m$, $h_i(x) \leq 0$, $i = 1, \ldots, p$

2. dual constraints: $\lambda \succeq 0$

3. complementary slackness: $\lambda_i^\star f_i(x^\star) = 0$ $i = 1, \ldots, m$

4. gradient of Lagrangian w.r.t. x vanishes:

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

if strong duality holds and x, λ, ν are optimal, then they must satisfy the KKT conditions

KKT Conditions for Convex Problem

if $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ satisfy KKT for a convex problem, then they are optimal:

- from complementary slackness: $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$
- from 4th condition (and convexity): $g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$
- hence $f_0(\tilde{x}) = g(\tilde{\lambda}, \tilde{\nu})$

KKT Conditions for Convex Problem

if Slater's condition is satisfied:
x is optimal if and only if there exist λ, ν that satisfy KKT conditions

- recall that Slater implies strong duality, and dual optimum is attained

Optimality conditions from KKT for Convex Problems

assume differentiable f_i, h_i:

- optimality condition for unconstrained optimization problems? 
- optimality conditions for equality constrained optimization problems? 
- optimality conditions for constrained optimization problems? 