# COMP 9602: Convex Optimization

## Duality (I)

Dr. C. Wu

Department of Computer Science
The University of Hong Kong

# Roadmap

| | |
|---|---|
| Theory | convex set<br>convex function<br>standard forms of optimization problems, quasi-convex optimization<br>linear program, integer linear program<br>quadratic program<br>geometric program<br>semidefinite program<br>vector optimization<br>duality |
| Algorithm | unconstrained optimization<br>equality constrained optimization<br>interior-point method<br>localization methods<br>subgradient method<br>decomposition methods |

# Duality

- Optimization in standard form (not necessarily convex)

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

D=dom $f_0 \cap$ dom $f_1 \cap \ldots \cap$ dom $f_m \cap$ dom $h_1 \cap \ldots \cap$ dom $h_p$

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

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

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

  - Lagrange multipliers (Lagrangian dual variables):

    $\lambda_i$ is Lagrange multiplier associated with $f_i(x) \le 0, i = 1, \ldots, m$

    $\nu_i$ is Lagrange multiplier associated with $h_i(x) = 0, i = 1, \ldots, p$

  - augment $f_0(x)$ by weighted sum of constraint functions (penalty functions)

# Lagrange dual

**Lagrange dual function:** $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$,

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

- Lagrange dual problem

$$\max g(\lambda, \nu)$$

subject to: $\lambda \succeq 0$

Note that there is no positive constraint on $\nu$

# LP and its dual

- Standard form LP

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

- Lagrangian:

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

- Lagrange dual function

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu)$$

- Dual problem

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

# Lagrange dual and conjugate function

**Optimization with affine inequality and equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b, \quad Cx = d \end{array}$$

**Lagrangian dual function**

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathbf{dom}\, f_0} \left( f_0(x) + (A^T\lambda + C^T\nu)^T x - b^T\lambda - d^T\nu \right) \\ &= -f_0^*(-A^T\lambda - C^T\nu) - b^T\lambda - d^T\nu \end{aligned}$$

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

(Slide 21, Convex Function notes)

- simplifies derivation of dual if conjugate of $f_0$ is kown

$$f_0(x) = \sum_{i=1}^n x_i \log x_i, \qquad f_0^*(y) = \sum_{i=1}^n e^{y_i - 1}$$

# Primal and Lagrange dual problems

- Primal problem

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

convex or non-convex

- x is the primal variable, $x \in \mathbf{R}^n$
- optimal value p*

- Lagrange dual problem

$$\max g(\lambda, \nu)$$

subject to: $\lambda \succeq 0$

a convex optimization problem

- $\lambda$ , $\nu$ : dual variables $\lambda \in \mathbf{R}^m \quad \nu \in \mathbf{R}^p$
- optimal value d*

# Properties of Lagrange dual

**Lagrange dual function:** $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$,

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

Property 1: $g(\lambda, \nu)$ is concave on $(\lambda, \nu)$ even if the original problem is not convex

# Example: the max cut problem

- Consider an undirected graph $G = (V, E)$. For $S \subset V$, the capacity of $S$ is the number of edges connecting a node in $S$ to a node not in $S$. Find $S \subset V$ with maximum capacity.
  - adjacency matrix of the graph
    $$Q_{ij} = \begin{cases} 1 & \text{if } \{i, j\} \in E \\ 0 & \text{otherwise} \end{cases}$$
  - a cut decided by a vector $x \in R^n$
    $$x_i = \begin{cases} 1 & \text{if } i \in S \\ -1 & \text{otherwise} \end{cases}$$
  - capacity of the cut
    $$c(x) = \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{n} (1 - x_i x_j) Q_{ij}$$

( $1 - x_i x_j = 2$ if $\{i, j\}$ is in the cut set)

# Example: the max cut problem (cont'd)

- Primal problem (NP complete even if $Q \succeq 0$)

$$\begin{array}{ll} \text{minimize} & x^T Q x \\ \text{subject to} & x_i \in \{-1, 1\} \quad \text{for all } i = 1, \ldots, n \end{array}$$

the maximum cut is $$c_{\text{max}} = \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n} Q_{ij} - \frac{1}{4}p^\star$$

$p^\star$: the optimal value of primal problem

<=>

$$\boxed{\begin{array}{ll} \text{minimize} & x^T Q x \\ \text{subject to} & x_i^2 - 1 = 0 \end{array}}$$

# Example: the max cut problem (cont'd)

- Dual problem (SDP)

  let $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$

  the Lagrangian is

$$L(x, \lambda) = x^T Q x - \sum_{i=1}^{n} \lambda_i (x_i^2 - 1) = x^T (Q - \Lambda) x + \mathbf{trace}\, \Lambda$$

  the dual is

$$\begin{array}{ll} \text{maximize} & \mathbf{trace}\, \Lambda \\ \text{subject to} & Q - \Lambda \succeq 0 \end{array}$$

# Properties of Lagrange dual (cont'd)

Property 2: $g(\lambda, \nu) \leq f_0(x)$ for every primal feasible x and dual feasible $(\lambda, \nu)$

Remarks:

(1) $f_0(x) - g(\lambda, \nu)$: duality gap for $(x, \lambda, \nu)$, which is alway non-negative

$p^* - d^*$: optimal duality gap

(2) $d^* \leq p^*$ : weak duality

for any optimization problem: convex or not

can be used to find a non-trivial lower bound on p*, for difficult primal problems

# Example: the max cut problem (cont'd)

- Any dual feasible solution gives a lower bound on primal objective value:

  for any feasible x

$$x^T Q x \geq x^T \Lambda x = \sum_{i=1}^{n} \Lambda_{ii} x_i^2 = \mathbf{trace}\, \Lambda$$

# Strong duality

**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 qualifications**

there exist many types of constraint qualifications

Slater's constraint qualification is a famous one

# Slater's constraint qualification

strong duality holds for a convex problem

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

if it is strictly feasible, *i.e.*,

$$\exists x \in \mathbf{int}\, \mathcal{D}: \qquad f_i(x) < 0, \quad i = 1, \ldots, m, \qquad Ax = b$$

Can be refined: linear inequalities do not need to hold with strict inequality

# Strong duality theorem for LP

- If an LP (in any form) has an optimal solution x*, then the dual also has an optimal solution y*, and $c^T x^* = b^T y^*$

**Primal**

$$\min c^T x$$

subject to:

$$Ax = b$$
$$x \succeq 0$$

**Dual**

$$\max b^T y$$

subject to:

$$A^T y \preceq c$$

| | | Primal | | |
|---|---|---|---|---|
| Dual | | has optimum | unbounded | infeasible |
| | has optimum | Y | N | N |
| | unbounded | N | N | Y |
| | infeasible | N | Y | Y |

- Reference
  - Chapter 5.1 - 5.2, Convex Optimization.

- Acknowledgement
  - Some materials are extracted from the slides created by Prof. Stephen Boyd for the textbook
  - Some materials are extracted from the lecture notes of Convex Optimization by Prof. Wei Yu at the University of Toronto