Lec16-mwf

Sunday, February 15, 2009 1:48 PM

So far we have covered:

- Convex sets and functions
- Convex optimization problems
- Optimality conditions
- Optimization algorithms

Duality

Sunday, February 01, 2009 1:18 PM

- Convert a convex program (primal problem) with a set of variables (primal varibles) to another convex program (dual problem) with another set of variables (dual variables)

- Under very general conditions, the two problems are equivalent!

Benefit of Studying the Dual Problem

① Reveal addition structure of the optimal solution (KKT condition)

  - The dual variables often have interesting engineering interpretations ("price")

  - The original problem may be decomposed into sub-problems when we convert to the dual

    - distributed solutions.

② The constraint set of the dual problem is often simpler

  - easy to compute projection

Basic set-up of duality

Sunday, February 01, 2009 1:23 PM

- The primal problem

$$\begin{aligned} \min \quad & f_0(x) & & \\ \text{sub to} \quad & f_i(x) \le 0 & i=1,\cdots,m & \quad \lambda \ge 0 \\ & h_i(x) = 0 & i=1,\cdots,p & \quad \nu \in \mathbb{R} \end{aligned}$$

$\uparrow$ multipliers

- Lagrangian

$$L(x,\lambda,\nu) = f_0(x) + \underbrace{\sum_{i=1}^{m} \lambda_i f_i(x)}_{\substack{\text{can be viewed} \\ \text{as the penalty} \\ \text{if } f_i(x) > 0}} + \sum_{i=1}^{p} \nu_i h_i(x)$$

  - all plus sign when we use the standard form.

  - $\lambda_i \ge 0 \Rightarrow L$ is convex in x.

- Minimize the Lagrangian over prime variables

$$g(\lambda,\nu) = \min_{x \in D} L(x,\lambda,\nu)$$

$$D = \operatorname{dom} f_0 \cap \operatorname{dom} f_i \cap \operatorname{dom} h_i$$

- Note: the constraints are eliminated!

- Dual problem

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

$$\text{sub to} \quad \lambda \geq 0$$

no condition of ν.

<u>Remarks</u>:

① We may view the additional terms in $L(x, \lambda, \nu)$ as a penalty term when x is not feasible (i.e, x does not satisfy the original constraints).

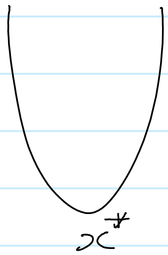
Not quite though ⇒ They also disturb the value of $f_0(\cdot)$ <u>within</u> the feasible set!

② When we compute $g(\lambda, \nu)$, we use an <u>unconstrained</u> minimization over x.

③ $g(\lambda, \nu)$ is a concave function of λ & ν

- It is a minimization of a set of linear functions of (λ, ν).

④ The dual problem is also a convex problem



primal problem

dual problem

⑤ The constraint set of the dual problem is a quadrant set

$\Rightarrow$ Projection becomes trivial.

(Q) When/Why does solving the dual also solve the primal?

- Easy to see that

$$g(\lambda, \nu) \leq L(x^*, \lambda, \nu) \leq f_0(x^*) \quad \forall \lambda \geq 0$$

$$\Rightarrow \max_{\lambda \geq 0} g(\lambda, \nu) \leq f_0(x^*)$$

- The optimal value of dual is always no greater than the optimal value of the primal.

$\Rightarrow$ weak duality

- Under light conditions, they are in fact equal.

$\Rightarrow$ strong duality

<u>Consequence</u> of strong duality

- There exists (λ^*, ν^*) such that

$$g(\lambda^*, \nu^*) = L(x^*, \lambda^*, \nu^*) = f_0(x^*)$$
$$\| $$
$$\min_x L(x, \lambda^*, \nu^*)$$

$\Rightarrow$ We may recover x^* from (λ^*, ν^*) by minimizing $L(x, \lambda^*, \nu^*)$ over x.

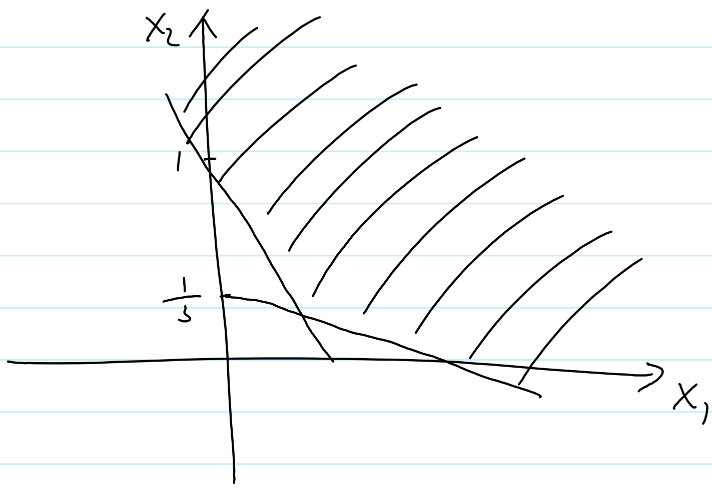
(some complication though ...)

15

Simple example and separable problems

Tuesday, February 22, 2011 3:23 PM

$$\begin{aligned} \min \quad & x_1^2 + 9x_2^2 & \\ \text{sub to} \quad & 2x_1 + x_2 \geq 1 & \lambda_1 \\ & x_1 + 3x_2 \geq 1 & \lambda_2 \end{aligned}$$



The Lagrangian is

$$L(x_1, x_2, \lambda_1, \lambda_2) = x_1^2 + 9x_2^2 + \lambda_1(1 - 2x_1 - x_2) + \lambda_2(1 - x_1 - 3x_2)$$

$$= x_1^2 - (2\lambda_1 + \lambda_2)x_1 + 9x_2^2 - (\lambda_1 + 3\lambda_2)x_2 + \lambda_1 + \lambda_2$$

Minimize the Lagrangian over x_1 & x_2. Since there are no longer coupling constraints, in this case x_1 & x_2 can be optimized separately $\rightarrow$ "seperable problems".

$$\frac{\partial L}{\partial x_1} = 2x_1 - (2\lambda_1 + \lambda_2) = 0$$

$$\frac{\partial L}{\partial x_2} = 18x_2 - (\lambda_1 + 3\lambda_2) = 0$$

$$\Rightarrow \quad x_1 = \frac{2\lambda_1 + \lambda_2}{\cdot}$$

$$x_2 = \frac{\lambda_1 + 3\lambda_2}{18}$$

$$g(\lambda_1, \lambda_2) = -\frac{(2\lambda_1 + \lambda_2)^2}{4} - \frac{(\lambda_1 + 3\lambda_2)^2}{36} + \lambda_1 + \lambda_2$$

The dual problem is

$$\max \quad g(\lambda_1, \lambda_2)$$

$$\text{sub to} \quad \lambda_1 \geq 0, \quad \lambda_2 \geq 0$$

$\Leftarrow$ constraint is a lot simpler (for using gradient projection)

To solve the dual problem: skip

$$\frac{\partial g}{\partial \lambda_1} = -\frac{2(2\lambda_1 + \lambda_2)}{4} \cdot 2 - \frac{2(\lambda_1 + 3\lambda_2)}{36} + 1$$

$$= -(2\lambda_1 + \lambda_2) - \frac{\lambda_1 + 3\lambda_2}{18} + 1$$

$$= -\frac{37}{18}\lambda_1 - \frac{7}{6}\lambda_2 + 1$$

$$\frac{\partial g}{\partial \lambda_2} = -\frac{2(2\lambda_1 + \lambda_2)}{4} - \frac{2(\lambda_1 + 3\lambda_2)}{36} \cdot 3 + 1$$

$$= -\frac{2\lambda_1 + \lambda_2}{2} - \frac{\lambda_1 + 3\lambda_2}{6} + 1$$

$$= -\frac{7}{6}\lambda_1 - \lambda_2 + 1$$

Suppose all constraints are not binding. We can set $\frac{\partial g}{\partial \lambda_1} = 0$ & $\frac{\partial g}{\partial \lambda_2} = 0$. We have

$$\begin{cases} \frac{37}{18}\lambda_1 + \frac{7}{6}\lambda_2 = 1 \\ \frac{49}{36}\lambda_1 + \frac{7}{6}\lambda_2 = \frac{7}{6} \end{cases}$$

$$\Rightarrow \quad \frac{25}{36}\lambda_1 = -\frac{1}{6} \quad \Rightarrow \quad \lambda_1 = -\frac{6}{25}$$

$$\lambda_2 = 1 - \frac{7}{6}\lambda_1 = \frac{32}{25}$$

However, this violates the constraint on $\lambda_1 \geq 0$.

Second guess : $\lambda_1 = 0 \qquad \frac{\partial g}{\partial \lambda_2} = 0$

$$\Rightarrow \quad \begin{cases} \lambda_1^* = 0 \\ \lambda_2^* = 1 \end{cases}$$

This is the optimal point for the dual problem.

- Let us verify it strong duality holds.
- Suppose strong-duality holds. Then, we can find the optimal primal solution by minimizing $L(x, \lambda_1^*, \lambda_2^*)$:

Hence, $\quad x_1^* = \frac{2\lambda_1 + \lambda_2}{2} = \frac{1}{2}$

$$x_2^* = \frac{\lambda_1 + 3\lambda_2}{18} = \frac{1}{6}$$

Note that this is feasible for the primal problem:

$$\left.\begin{matrix} 2x_1^* + x_2^* = \frac{7}{6} > 1 & \lambda_1^* = 0 \\ x_1^* + 3x_2^* = 1 & \lambda_2^* > 0 \end{matrix}\right\} \text{complementary slackness (later)}$$

$$x_1^* + 3x_2^* = 1 \qquad \lambda_2^* > 0$$ slackness
(later)

Finally, $x_1^{*2} + 9x_2^{*2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$$g(\lambda_1^*, \lambda_2^*) = -\frac{1}{4} - \frac{1}{4} + 1 = \frac{1}{2}$$

The duality gap is indeed zero!

Example: separable problems - skip

Wednesday, February 04, 2009 10:13 AM

$$\min \quad \sum_{i=1}^{n} x_i \log x_i$$

$$\text{sub to} \quad \sum_i a_{ik} x_i \le b_k, \quad k = 1, 2, \cdots, K \qquad \lambda_k$$

$$\sum x_i = 1 \qquad \nu$$

$$x_i \ge 0 \quad (\text{implicit constraint})$$

$$L(x, \lambda, \nu) = \sum_{i=1}^{n} x_i \log x_i + \sum_{k=1}^{K} \lambda_k \left(\sum_i a_{ik} x_i - b_k \right) + \nu \left(\sum x_i - 1 \right)$$

$$= \sum_{i=1}^{n} \left(x_i \log x_i + \sum_{k=1}^{K} \lambda_k a_{ik} x_i + \nu x_i \right)$$

$$- (\lambda^T b + \nu)$$

- Let $a_i = [a_{ik}]_{k=1,\cdots,K}$. Then $a_i^T \lambda = \sum_{k=1}^{K} \lambda_k a_{ik}$.

- Note that the first term is separable for each i

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

$$= \sum_{i=1}^{n} \min \left[x_i \log x_i + a_i^T \lambda x_i + \nu x_i \right]$$

$$- (\lambda^T b + \nu)$$

- For each i,

$$\frac{\partial}{\partial x_i} = \log x_i + 1 + a_i^T \lambda + \nu = 0$$

$$\Rightarrow \quad x_i^* = e^{-a_i^T\lambda - \nu - 1}$$

$$\therefore \quad g(\lambda, \nu) = -\sum_{i=1}^{n} x_i^* - (\lambda^T b + \nu)$$

$$= -\sum_{i=1}^{n} e^{-a_i^T\lambda - \nu - 1} - (\lambda^T b + \nu)$$

concave.

$$\Rightarrow \quad \max \; g(\lambda, \nu)$$

$$\text{sub to} \quad \lambda \geq 0$$

- Can minimize further for ν (Boyd p228)

- This type of convex problem is said to have separable objective functions.

  Although the constraint in the primal problem is coupled, duality helps us to decouple them in the Lagrangian.

  Hence, the minimization of the Lagrangian can be carried out independently for each variable.

- Useful for developing distributed solutions.

35