Lecture 3: Lagrangian duality

Michael Patriksson

2010-09-06

The Relaxation Theorem

- Problem: find

$$f^* = \operatorname*{infimum}_{\mathbf{x}} \; f(\mathbf{x}), \tag{1a}$$
$$\text{subject to } \mathbf{x} \in S, \tag{1b}$$

  where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a given function and $S \subseteq \mathbb{R}^n$
- A *relaxation* to (1a)–(1b) has the following form: find

$$f_R^* = \operatorname*{infimum}_{\mathbf{x}} \quad f_R(\mathbf{x}), \tag{2a}$$
$$\text{subject to} \quad \mathbf{x} \in S_R, \tag{2b}$$

  where $f_R : \mathbb{R}^n \mapsto \mathbb{R}$ is a function with $f_R \leq f$ on S and $S_R \supseteq S$

Relaxation example (maximization)

- Binary knapsack problem:

$$\begin{aligned} z^* = \underset{\mathbf{x}\in\{0,1\}^4}{\text{maximize}} \quad & 7x_1 + 4x_2 + 5x_3 + 2x_4 \\ \text{subject to} \quad & 3x_1 + 3x_2 + 4x_3 + 2x_4 \quad \leq \quad 5 \end{aligned}$$

  - Optimal solution: $\mathbf{x}^* = (1, 0, 0, 1)$, $z^* = 9$
- Continuous relaxation:

$$\begin{aligned} z^*_{\mathrm{LP}} = \underset{\mathbf{x}\in[0,1]^4}{\text{maximize}} \quad & 7x_1 + 4x_2 + 5x_3 + 2x_4 \\ \text{subject to} \quad & 3x_1 + 3x_2 + 4x_3 + 2x_4 \quad \leq \quad 5 \end{aligned}$$

  - Optimal solution: $\mathbf{x}^*_R = (1, \frac{2}{3}, 0, 0)$, $z^*_R = 9\frac{2}{3} > z^*$
- $\mathbf{x}^*_R$ is *not feasible* in the binary problem

The relaxation theorem

1. [relaxation] $f_R^* \leq f^*$
2. [infeasibility] *If* (2) *is infeasible, then so is* (1)
3. [optimal relaxation]
   If the problem (2) *has an optimal solution* $\mathbf{x}_R^* \in S$ *for which*

$$f_R(\mathbf{x}_R^*) = f(\mathbf{x}_R^*),$$

   then $\mathbf{x}_R^*$ *is an optimal solution to* (1) *as well*

- *Proof portion.* For 3., note that

$$f(\mathbf{x}_R^*) = f_R(\mathbf{x}_R^*) \leq f_R(\mathbf{x}) \leq f(\mathbf{x}), \qquad \mathbf{x} \in S$$

Lagrangian relaxation, I

- Consider the optimization problem:

$$f^* = \operatorname*{infimum}_{\mathbf{x}} f(\mathbf{x}), \tag{3a}$$
$$\text{subject to } \mathbf{x} \in X, \tag{3b}$$
$$g_i(\mathbf{x}) \leq 0, \qquad i = 1, \dots, m, \tag{3c}$$

  where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $g_i : \mathbb{R}^n \mapsto \mathbb{R}$ $(i = 1, 2, \dots, m)$ are given functions, and $X \subseteq \mathbb{R}^n$
- Here we assume that

$$-\infty < f^* < \infty, \tag{4}$$

  that is, that f is bounded from below and that the problem has at least one feasible solution
- The form of (3) represents our decision on the Lagrangian relaxation

Lagrangian relaxation, II

- For a vector $\boldsymbol{\mu} \in \mathbb{R}^m$, we define the *Lagrange function*

$$L(\mathbf{x}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^{m} \mu_i g_i(\mathbf{x}) = f(\mathbf{x}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}(\mathbf{x})$$

- We call the vector $\boldsymbol{\mu}^* \in \mathbb{R}^m$ a *Lagrange multiplier* if it is non-negative and if $f^* = \inf_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*)$ holds

Lagrange multipliers and global optima

- *Let* $\boldsymbol{\mu}^*$ *be a Lagrange multiplier.*
  Then, $\mathbf{x}^*$ *is an optimal solution to*

  $$f^* = \inf\{ f(\mathbf{x}) \,|\, \mathbf{x} \in X, g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \},$$

  if and only if it is feasible and

  $$\mathbf{x}^* \in \arg\min_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*), \qquad \textit{and} \qquad \mu_i^* g_i(\mathbf{x}^*) = 0, \; i = 1, \ldots, m$$

- Notice the resemblance to the KKT conditions:
  - If $X = \mathbb{R}^n$ and all functions are convex and in C^1 then "$\mathbf{x}^* \in \arg\min_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*)$" $\Leftrightarrow$ "force equilibrium condition", i.e., the first row of the KKT conditions
  - The second item, "$\mu_i^* g_i(\mathbf{x}^*) = 0$ for all i" $\Leftrightarrow$ complementarity conditions

The Lagrangian dual problem associated with the Lagrangian relaxation

- The *Lagrangian dual function* is

$$q(\boldsymbol{\mu}) = \operatorname*{infimum}_{\mathbf{x} \in X} \; L(\mathbf{x}, \boldsymbol{\mu})$$

- The *Lagrangian dual problem* is to

$$q^* = \operatorname*{maximize}_{\boldsymbol{\mu} \geq \mathbf{0}^m} \; q(\boldsymbol{\mu}) \tag{5}$$

- For some $\boldsymbol{\mu}$, $q(\boldsymbol{\mu}) = -\infty$ is possible. If this is true for all $\boldsymbol{\mu} \geq \mathbf{0}^m$ then

$$q^* = \operatorname*{supremum}_{\boldsymbol{\mu} \geq \mathbf{0}^m} \; q(\boldsymbol{\mu}) = -\infty$$

The Lagrangian dual problem, cont'd

- The *effective domain* of q is $D_q = \{ \boldsymbol{\mu} \in \mathbb{R}^m \mid q(\boldsymbol{\mu}) > -\infty \}$

  [Theorem] D_q is convex, and q is concave on D_q $\square$

- Very good news: The Lagrangian dual problem is always convex!

- Maximize a concave function (even continuous as long as $D_q = \mathbb{R}^m$)

- Need still to show how a Lagrangian dual optimal solution can be used to generate a primal optimal solution

Weak Duality Theorem

Let $\mathbf{x}$ *and* $\boldsymbol{\mu}$ *be* feasible *in*

$$f^* = \inf\{f(\mathbf{x}) \,|\, \mathbf{x} \in X, g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\}$$

and

$$q^* = \max\{\, q(\boldsymbol{\mu}) \,|\, \boldsymbol{\mu} \geq \mathbf{0}^m \,\},$$

respectively. Then,

$$q(\boldsymbol{\mu}) \leq f(\mathbf{x})$$

In particular,

$$q^* \leq f^*$$

If $q(\boldsymbol{\mu}) = f(\mathbf{x})$*, then the pair* $(\mathbf{x}, \boldsymbol{\mu})$ *is optimal in the respective problem and*

$$q^* = q(\boldsymbol{\mu}) = f(\mathbf{x}) = f^*$$

□

Weak Duality Theorem, cont'd

- Weak duality is also a consequence of the Relaxation Theorem: For any $\boldsymbol{\mu} \geq \mathbf{0}^m$, let

$$\begin{aligned} S &= X \cap \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) \leq \mathbf{0}^m \,\}, \\ S_R &= X, \\ f_R &= L(\boldsymbol{\mu}, \cdot) \end{aligned}$$

  Apply the Relaxation Theorem

- If $q^* = f^*$, there is *no duality gap*

- If there exists a Lagrange multiplier vector, then by the weak duality theorem, there is no duality gap

Global optimality conditions

- *The vector* $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ *is a pair of an optimal primal solution and a Lagrange multiplier if and only if*

$$\boldsymbol{\mu}^* \geq \mathbf{0}^m, \quad \textit{(Dual feasibility)} \tag{6a}$$

$$\mathbf{x}^* \in \arg\min_{\mathbf{x} \in X} L(\mathbf{x}, \boldsymbol{\mu}^*), \quad \textit{(Lagrangian optimality)} \tag{6b}$$

$$\mathbf{x}^* \in X, \; \mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}^m, \quad \textit{(Primal feasibility)} \tag{6c}$$

$$\mu_i^* g_i(\mathbf{x}^*) = 0, \; i = 1, \dots, m \quad \textit{(Complementary slackness)} \tag{6d}$$

- If $\exists(\mathbf{x}^*, \boldsymbol{\mu}^*)$ that fulfil (6), then there is a zero duality gap and Lagrange multipliers exist
- Compare with the duality theory of LP!

Saddle points

- *The vector* $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ *is a pair of an optimal primal solution and a Lagrange multiplier if and only if* $\mathbf{x}^* \in X$, $\boldsymbol{\mu}^* \geq \mathbf{0}^m$, *and* $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ *is a* saddle point *of the Lagrangian function on* $X \times \mathbb{R}^m_+$, *that is,*

$$L(\mathbf{x}^*, \boldsymbol{\mu}) \leq L(\mathbf{x}^*, \boldsymbol{\mu}^*) \leq L(\mathbf{x}, \boldsymbol{\mu}^*), \quad (\mathbf{x}, \boldsymbol{\mu}) \in X \times \mathbb{R}^m_+,$$

  holds

- If $\exists (\mathbf{x}^*, \boldsymbol{\mu}^*)$, equivalent to the global optimality conditions, the existence of Lagrange multipliers, and a zero duality gap

Strong duality for convex programs, introduction

- Convexity of the dual problem comes with very few assumptions on the original, primal problem
- The characterization of the primal–dual set of optimal solutions is also quite easily established
- To establish *strong duality*—sufficient conditions under which there is no duality gap—takes much more
- In particular—as with the KKT conditions—we need regularity conditions (constraint qualifications) and separation theorems

Strong duality theorem

- Consider the problem (3), that is,

$$f^* = \inf\{ f(\mathbf{x}) \,|\, \mathbf{x} \in X, g_i(\mathbf{x}) \leq 0, i = 1, \dots, m \},$$

  where $f : \mathbb{R}^n \mapsto \mathbb{R}$ and g_i $(i = 1, \dots, m)$ are *convex* and $X \subseteq \mathbb{R}^n$ is a *convex* set

- Introduce the following constraint qualification (CQ):

$$\exists \mathbf{x} \in X \text{ with } \mathbf{g}(\mathbf{x}) < \mathbf{0}^m \tag{7}$$

Strong duality theorem

Suppose that $-\infty < f^* < \infty$, *and that the CQ* (7) *holds for the (convex) problem* (3)

(a) *There is no duality gap and there exists at least one Lagrange multiplier* $\boldsymbol{\mu}^*$. *Moreover, the set of Lagrange multipliers is bounded and convex*

(b) *If infimum in* (3) *is attained at some* $\mathbf{x}^*$, *then the pair* $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ *satisfies the global optimality conditions* (6)

(c) *If the functions* f *and* g_i *are in* C^1 *and* X *is open (for example,* $X = \mathbb{R}^n$*) then* (6) *equals the KKT conditions*

If all constraints are linear we can remove the CQ (7)

Example I: An explicit, differentiable dual problem

- Consider the problem to

$$\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \;\; & f(\mathbf{x}) := x_1^2 + x_2^2, \\
\text{subject to} \;\; & x_1 + x_2 \geq 4, \\
& x_j \geq 0, \qquad j = 1, 2
\end{aligned}$$

- Let

$$g(\mathbf{x}) = -x_1 - x_2 + 4$$

and

$$X = \{ (x_1, x_2) \mid x_j \geq 0, \; j = 1, 2 \} = \mathbb{R}^2_+$$

Example I, cont'd

- The Lagrangian dual function is

$$
\begin{aligned}
q(\mu) &= \min_{\mathbf{x} \in X} L(\mathbf{x}, \mu) := f(\mathbf{x}) + \mu(-x_1 - x_2 + 4) \\
&= 4\mu + \min_{\mathbf{x} \geq \mathbf{0}} \{x_1^2 + x_2^2 - \mu x_1 - \mu x_2\} \\
&= 4\mu + \min_{x_1 \geq 0} \{x_1^2 - \mu x_1\} + \min_{x_2 \geq 0} \{x_2^2 - \mu x_2\}, \ \mu \geq 0
\end{aligned}
$$

- For a fixed $\mu \geq 0$, the minimum is attained at $x_1(\mu) = \frac{\mu}{2}, x_2(\mu) = \frac{\mu}{2}$
- Substituting this expression into $q(\mu) \Rightarrow$ $q(\mu) = f(\mathbf{x}(\mu)) + \mu(-x_1(\mu) - x_2(\mu) + 4) = 4\mu - \frac{\mu^2}{2}$
- Note that q is *strictly concave*, and it is differentiable everywhere (since f, g are differentiable and $\mathbf{x}(\mu)$ is unique). More on this later!

Example I, cont'd

- Recall the dual problem

$$q^* = \max_{\mu \geq 0} q(\mu) = \max_{\mu \geq 0} \left(4\mu - \frac{\mu^2}{2} \right)$$

- We have that $q'(\mu) = 4 - \mu = 0 \iff \mu = 4$
  As $4 \geq 0$, this is the optimum in the dual problem!

$\Rightarrow$ $\mu^* = 4$ and $\mathbf{x}^* = (x_1(\mu^*), x_2(\mu^*))^{\mathrm{T}} = (2, 2)^{\mathrm{T}}$

- Also: $f(\mathbf{x}^*) = q(\mu^*) = 8$

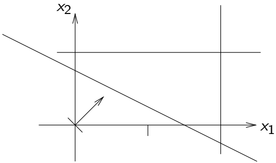
- Here, the dual function is *differentiable*. The optimum $\mathbf{x}^*$ is also unique and automatically given by $\mathbf{x}^* = \mathbf{x}(\mu^*)$

Example II: Implicit non-differentiable dual problem

- Consider the linear programming problem to

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \;\; & f(\mathbf{x}) := -x_1 - x_2, \\ \text{subject to} \;\; & 2x_1 + 4x_2 \leq 3, \\ & 0 \leq x_1 \leq 2, \\ & 0 \leq x_2 \leq 1 \end{aligned}$$

- The optimal solution is $\mathbf{x}^* = (3/2, 0)^{\mathrm{T}}, f(\mathbf{x}^*) = -3/2$

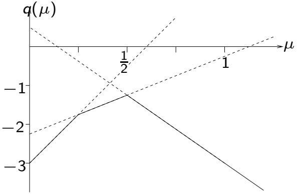


Example II: Lagrangian relax the first constraint

$$L(\mathbf{x}, \mu) = -x_1 - x_2 + \mu(2x_1 + 4x_2 - 3);$$

$$q(\mu) = -3\mu + \min_{0 \le x_1 \le 2} \{(-1 + 2\mu)x_1\} + \min_{0 \le x_2 \le 1} \{(-1 + 4\mu)x_2\}$$

$$= \begin{cases} -3 + 5\mu, & 0 \le \mu \le 1/4, & \Leftrightarrow & x_1(\mu) = 2, x_2(\mu) = 1 \\ -2 + \ \ \mu, & 1/4 \le \mu \le 1/2, & \Leftrightarrow & x_1(\mu) = 2, x_2(\mu) = 0 \\ \quad\ - 3\mu, & 1/2 \le \mu & \Leftrightarrow & x_1(\mu) = x_2(\mu) = 0 \end{cases}$$



$$\mu^* = \tfrac{1}{2},\ q(\mu^*) = -\tfrac{3}{2}$$

Example II, cont'd

- For linear (convex) programs strong duality holds, but how obtain $\mathbf{x}^*$ from μ^*?
- q is non-differentiable at $\mu^* \Rightarrow$ Utilize characterization in (6)
- The subproblem solution set at μ^* is $X(\mu^*) = \{ \binom{2\alpha}{0} \mid 0 \leq \alpha \leq 1 \}$
- Among the subproblem solutions, we next have to find one that is primal feasible as well as complementary
- Primal feasibility means that $2 \cdot 2\alpha + 4 \cdot 0 \leq 3 \iff \alpha \leq 3/4$
- Complementarity means that $\mu^* \cdot (2x_1^* + 4x_2^* - 3) = 0 \iff \alpha = 3/4$, since $\mu^* \neq 0$
- Conclusion: the only primal vector $\mathbf{x}$ that satisfies the system (6) together with the dual solution $\mu^* = 1/2$ is $\mathbf{x}^* = (3/2, 0)^{\mathrm{T}}$
- Observe finally that $f^* = q^*$

A theoretical argument for $\mu^* = 1/2$

- Due to the global optimality conditions, the optimal solution must in this convex case be among the subproblem solutions
- Since x_1^* is not in one of the "corners" of X ($0 < x_1^* < 2$), the value of μ^* must be such that the cost term for x_1 in $L(\mathbf{x}, \mu^*)$ is zero! That is, $-1 + 2\mu^* = 0 \Rightarrow \mu^* = 1/2$!
- A non-coordinability phenomenon—a non-unique subproblem solution means that the optimal solution is not obtained automatically
- In non-convex cases (e.g., integrality constraints) the optimal solution may not be among the points in $X(\boldsymbol{\mu}^*)$ (the set of subproblem solutions at μ^*)
- What do we do then??