Lagrangian Duality Theory

Yinyu Ye

Department of Management Science and Engineering

Stanford University

Stanford, CA 94305, U.S.A.

http://www.stanford.edu/˜yyye

Lagrangian Function

We consider the general constrained optimization:

$$
\text{(GCO)} \qquad \begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & c_i\mathbf{x}) \quad (\le, =, \ge) \quad 0, \; i = 1, ..., m, \end{array}
$$

For Lagrange Multipliers.

$$
Y := \{y_i \quad (\le, '\text{free}', \ge) \quad 0, \; i = 1, ..., m\},
$$

the Lagrangian Function is given by

$$
L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T \mathbf{c}(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^{m} y_i c_i(\mathbf{x}), \; \mathbf{y} \in Y.
$$

Toy Example Again

$$\begin{aligned} \text{minimize} \quad & (x_1-1)^2+(x_2-1)^2 \\ \text{subject to} \quad & x_1+2x_2-1 \le 0, \\ & 2x_1+x_2-1 \le 0. \end{aligned}$$

$$L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x}) = f(\mathbf{x}) - \sum_{i=1}^{2} y_i c_i(\mathbf{x}) =$$

$$= (x_1-1)^2+(x_2-1)^2 - y_1(x_1+2x_2-1) - y_2(2x_1+x_2-1), \; (y_1;y_2) \le \mathbf{0}$$

where

$$\nabla L_x(\mathbf{x},\mathbf{y}) = \begin{pmatrix} 2(x_1-1)-y_1-2y_2 \\ 2(x_2-1)-2y_1-y_2 \end{pmatrix}$$

Lagrangian Relaxation Problem

For given multipliers $\mathbf{y} \in Y$:

$$
\begin{array}{lll}
(LRP) & \inf & L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}^T \mathbf{c}(\mathbf{x}) \\
 & \text{s.t.} & \mathbf{x} \in R^n.
\end{array}
$$

Again, $\mathbf{y}_i$ can be viewed as a penalty parameter to penalize constraint violation $c_i(\mathbf{x}), i = 1, ..., m$.

In the example, for given $(y_1; y_2) \le \mathbf{0}$, the LRP is:

$$
\begin{array}{ll}
\inf & (x_1 - 1)^2 + (x_2 - 1)^2 - y_1(x_1 + 2x_2 - 1) - y_2(2x_1 + x_2 - 1) \\
\text{s.t.} & (x_1; x_2) \in R^2,
\end{array}
$$

and it has a close form solution $\mathbf{x}$:

$$
x_1 = \frac{y_1 + 2y_2}{2} + 1 \quad \text{and} \quad x_2 = \frac{2y_1 + y_2}{2} + 1
$$

with the minimal value $= -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2$.

Minimal Value Function as the Dual Objective

For any $\mathbf{y} \in Y$, define the the minimal value function (including unbounded from below) and the Lagrangian Dual Problem (LDP):

$$\begin{array}{llll} \phi(\mathbf{y}) := & \inf_{\mathbf{x}} & L(\mathbf{x},\mathbf{y}), & \text{s.t.} \quad \mathbf{x} \in R^n. \\ (LDP) & \sup_{\mathbf{y}} & \phi(\mathbf{y}), & \text{s.t.} \quad \mathbf{y} \in Y. \end{array}$$

Theorem 1 *The Lagrangian dual objective $\phi(\mathbf{y})$ is a concave function.*

Proof: For any given two multiply vectors $\mathbf{y}^1 \in Y$ and $\mathbf{y}^2 \in Y$,

$$\begin{aligned} \phi(\alpha\mathbf{y}^1 + (1-\alpha)\mathbf{y}^2) &= \inf_{\mathbf{x}} L(\mathbf{x}, \alpha\mathbf{y}^1 + (1-\alpha)\mathbf{y}^2) \\ &= \inf_{\mathbf{x}}[f(\mathbf{x}) - (\alpha\mathbf{y}^1 + (1-\alpha)\mathbf{y}^2)^T\mathbf{c}(\mathbf{x})] \\ &= \inf_{\mathbf{x}}[\alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{x}) - \alpha(\mathbf{y}^1)^T\mathbf{c}(\mathbf{x}) - (1-\alpha)(\mathbf{y}^2)^T\mathbf{c}(\mathbf{x})] \\ &= \inf_{\mathbf{x}}[\alpha L(\mathbf{x},\mathbf{y}^1) + (1-\alpha)L(\mathbf{x},\mathbf{y}^2)] \\ &\ge \alpha[\inf_{\mathbf{x}} L(\mathbf{x},\mathbf{y}^1)] + (1-\alpha)[\inf_{\mathbf{x}} L(\mathbf{x},\mathbf{y}^2)] \\ &= \alpha\phi(\mathbf{y}^1) + (1-\alpha)\phi(\mathbf{y}^2), \end{aligned}$$

Dual Objective Establishes a Lower Bound

Theorem 2 *(Weak duality theorem) For every $\mathbf{y} \in Y$, the Lagrangian dual function $\phi(\mathbf{y})$ is less or equal to the infimum value of the original GCO problem.*

Proof:

$$\begin{aligned} \phi(\mathbf{y}) &= \inf_{\mathbf{x}} \{f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x})\} \\ &\le \inf_{\mathbf{x}} \{f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x}) \text{ s.t. } \mathbf{c}(\mathbf{x})(\le,=,\ge)\mathbf{0}\ \} \\ &\le \inf_{\mathbf{x}} \{f(\mathbf{x}) :\ \text{ s.t. } \mathbf{c}(\mathbf{x})(\le,=,\ge)\mathbf{0}\ \}. \end{aligned}$$

The first inequality is from the fact that the unconstrained inf value is no greater than the constrained one.

The second inequality is from $\mathbf{c}(\mathbf{x})(\le,=,\ge)\mathbf{0}$ and $\mathbf{y}(\le,' \text{free}', \ge)\mathbf{0}$ imply $-\mathbf{y}^T\mathbf{c}(\mathbf{x}) \le 0$.

The Lagrangian Dual Problem for the Toy Example

$$\begin{array}{ll} \text{minimize} & (x_1-1)^2+(x_2-1)^2 \\ \text{subject to} & x_1+2x_2-1\le 0, \\ & 2x_1+x_2-1\le 0; \end{array}$$

where $\mathbf{x}^* = \left(\frac{1}{3};\ \frac{1}{3}\right)$.

$$\phi(\mathbf{y}) = -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2,\ \mathbf{y}\le\mathbf{0}.$$

$$\begin{array}{ll} \max & -1.25y_1^2 - 1.25y_2^2 - 2y_1y_2 - 2y_1 - 2y_2 \\ \text{s.t.} & (y_1;y_2)\le\mathbf{0}. \end{array}$$

where $\mathbf{y}^* = \left(\frac{-4}{9};\ \frac{-4}{9}\right)$.

The Lagrangian Dual of LP I

Consider LP problem

$$(LP) \quad \text{minimize} \quad \mathbf{c}^T\mathbf{x}$$
$$\text{subject to} \quad A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0};$$

and it conic dual problem

$$(LD) \quad \text{maximize} \quad \mathbf{b}^T\mathbf{y}$$
$$\text{subject to} \quad A^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ \mathbf{s} \geq \mathbf{0}.$$

Let the Lagrangian multipliers be $\mathbf{y}('\text{free}')$ for equalities and $\mathbf{s} \geq \mathbf{0}$ for constraints $\mathbf{x} \geq \mathbf{0}$. Then the Lagrangian function would be

$$L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \mathbf{c}^T\mathbf{x} - \mathbf{y}^T(A\mathbf{x} - \mathbf{b}) - \mathbf{s}^T\mathbf{x} = (\mathbf{c} - A^T\mathbf{y} - \mathbf{s})^T\mathbf{x} + \mathbf{b}^T\mathbf{y}.$$

The Lagrangian Dual of LP II

Now consider the Lagrangian dual objective

$$\phi(\mathbf{y},\mathbf{s}) = \inf_{\mathbf{x}\in R^n} L(\mathbf{x},\mathbf{y},\mathbf{s}) = \inf_{\mathbf{x}\in R^n} [(\mathbf{c} - A^T\mathbf{y} - \mathbf{s})^T\mathbf{x} + \mathbf{b}^T\mathbf{y}].$$

If $(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}) \neq \mathbf{0}$, then $\phi(\mathbf{y},\mathbf{s}) = -\infty$. Thus, in order to maximize $\phi(\mathbf{y},\mathbf{s})$, the dual must maintain a constraint $(\mathbf{c} - A^T\mathbf{y} - \mathbf{s}) = \mathbf{0}$. This constraint, together with the sign constraint $\mathbf{s} \geq \mathbf{0}$, establish the Lagrangian dual problem:

$$\begin{array}{lll} (LDP) & \text{maximize} & \mathbf{b}^T\mathbf{y} \\ & \text{subject to} & A^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ \mathbf{s} \geq \mathbf{0}. \end{array}$$

which is identical to the conic dual of LP.

Lagrangian Strong Duality Theorem

Theorem 3 *Let (GCO) be a convex minimization problem and the infimum* f^* *of (GCO) be finite, and the suprermum of (LDP) be* ϕ^**. In addition, let (GCO) have an interior-point feasible solution with respect to inequality constraints, that is, there is* $\hat{\mathbf{x}}$ *such that all inequality constraints are strictly held. Then,* $f^* = \phi^*$*, and (LDP) admits a maximizer* $\mathbf{y}^*$ *such that*

$$\phi(\mathbf{y}^*) = f^*.$$

Furthermore, if (COP) admits a minimizer $\mathbf{x}^*$*, then*

$$y_i^* c_i(\mathbf{x}^*) = 0, \; \forall i = 1, ..., m.$$

The assumption of "interior-point feasible solution" is called Constraint Qualification condition, which was also needed as a condition to prove the strong duality theorem for general Conic Linear Optimization.

Note that the problem would be a convex minimization problem if all equality constraints are hyperplane or affine functions $c_i(\mathbf{x}) = \mathbf{a}_i\mathbf{x} - b_i$, all other level sets are convex.

An Example where the Constraint Qualification Failed

Consider the problem

$$
\begin{array}{lll}
\min & & x_1 \\
\text{s.t.} & x_1^2 + (x_2 - 1)^2 - 1 \leq 0, & (y_1 \leq 0) \\
 & x_1^2 + (x_2 + 1)^2 - 1 \leq 0, & (y_2 \leq 0)
\end{array}
$$

$$\mathbf{x}^* = (0;\ 0).$$

$$L(\mathbf{x}, \mathbf{y}) = x_1 + y_1(x_1^2 + (x_2 - 1)^2 - 1) + y_2(x_1^2 + (x_2 + 1)^2 - 1).$$

$$\phi(\mathbf{y}) = \frac{-1 - (y_1 - y_2)^2}{y_1 + y_2}.$$

Although there is no duality gap, but the dual does not admit a (finite) maximizer...

Proof of Lagrangian Strong Duality Theorem with Constraints $\mathbf{c}(\mathbf{x}) \geq \mathbf{0}$

Consider the convex (?) set

$$C := \{(\kappa; \mathbf{s}) : \ \exists \mathbf{x} \text{ s.t. } f(\mathbf{x}) \leq \kappa, \ -\mathbf{c}(\mathbf{x}) \leq \mathbf{s}\}.$$

Then, $(f^*; \mathbf{0})$ is on the closure of C. From the supporting hyperplane theorem, there exists $(y_0^*; \mathbf{y}^*) \neq \mathbf{0}$ such that

$$y_0^* f^* \leq \inf_{(\kappa;\mathbf{s}) \in C} (y_0^* \kappa + (\mathbf{y}^*)^T \mathbf{s}).$$

First, we show $\mathbf{y}^* \geq \mathbf{0}$, since otherwise one can choose some $(0; \mathbf{s} \geq \mathbf{0})$ such that the inequality is violated. Secondly, we show $y_0^* > 0$, since otherwise one can choose $(\kappa \to \infty; \mathbf{0})$ if $y^* < 0$, or $(0; \mathbf{s} = -\mathbf{c}(\hat{\mathbf{x}}) < \mathbf{0})$ if $y^* = 0$ (where $\mathbf{y}^* \neq \mathbf{0}$), such that the above inequality is violated.

Now let us divide both sides by y_0^* and let $\mathbf{y}^* := \mathbf{y}^*/y_0^*$, we have

$$f^* \leq \inf_{(\kappa;\mathbf{s}) \in C} (\kappa + (\mathbf{y}^*)^T \mathbf{s}) = \inf_{\mathbf{x}} (f(\mathbf{x}) - (\mathbf{y}^*)^T \mathbf{c}(\mathbf{x})) = \phi(\mathbf{y}^*) \leq \phi^*.$$

Then, from the weak duality theorem, we must have $f^* = \phi^*$.

If (GCO) admits a minimizer $\mathbf{x}^*$, then $f(\mathbf{x}^*) = f^*$ so that

$$f(\mathbf{x}^*) \leq \inf_{\mathbf{x}}(f(\mathbf{x}) - (\mathbf{y}^*)^T\mathbf{c}(\mathbf{x})) \leq f(\mathbf{x}^*) - (\mathbf{y}^*)^T\mathbf{c}(\mathbf{x}^*) = f(\mathbf{x}^*) - \sum_i^m y_i^* c_i(\mathbf{x}^*),$$

which implies that

$$\sum_i^m y_i^* c_i(\mathbf{x}^*) \leq 0.$$

Since $y_i^* \geq 0$ and $c_i(\mathbf{x}^*) \geq 0$ for all i, it must be true $y_i^* c_i(\mathbf{x}^*) = 0$ for all i.

More on Lagrangian Duality

Consider the constrained problem again

$$
\begin{array}{lll}
(GCO) & \inf & f(\mathbf{x}) \\
 & \text{s.t.} & \mathbf{c}_i(\mathbf{x}) \ (\le, =, \ge) \ 0, \ i = 1, ..., m, \\
 & & \mathbf{x} \in \Omega \subset R^n.
\end{array}
$$

Typically, Ω is a simple set such as the cone

$$\Omega = R^n_+ = \{\mathbf{x} : \ \mathbf{x} \ge \mathbf{0}\}$$

or the box

$$\Omega := \{\mathbf{x} : \ -\mathbf{e} \le \mathbf{x} \le \mathbf{e}.\}$$

Lagrangian Relaxation Problem

Lagrangian Function:

$$L(\mathbf{x},\mathbf{y},\mathbf{s}) = f(\mathbf{x}) - \mathbf{y}^T\mathbf{c}(\mathbf{x}),\ \mathbf{y} \in Y;$$

and let the dual objective be

$$\begin{array}{rll} \phi(\mathbf{y}) := & \inf_{\mathbf{x}} & L(\mathbf{x},\mathbf{y}) \\ & \text{s.t.} & \mathbf{x} \in \Omega. \end{array}$$

Theorem 4 *The Lagrangian dual function $\phi(\mathbf{y})$ is a concave function.*

Theorem 5 *(Weak duality theorem) For every $\mathbf{y} \in Y$, the Lagrangian dual function value $\phi(\mathbf{y})$ is less or equal to the infimum value of the original CO problem.*

The Lagrangian Dual Problem

$$(LDP) \quad \sup \quad \phi(\mathbf{y})$$
$$\text{s.t.} \quad \mathbf{y} \in Y.$$

would called the Lagrangian dual of the original CO problem:

Theorem 6 *(Strong duality theorem) Let (GCO) be a convex minimization problem, the infimum f^* of (GCO) be finite, and the suprermum of (LDP) be ϕ^*. In addition, let (GCO) have an interior-point feasible solution with respect to inequality constraints, that is, there is $\hat{\mathbf{x}}$ such that all inequality constraints are strictly held. Then, $f^* = \phi^*$, and (LDP) admits a maximizer $\mathbf{y}^*$ such that*

$$\phi(\mathbf{y}^*) = f^*.$$

Furthermore, if (GCO) admits a minimizer $\mathbf{x}^*$, then

$$y_i^* c_i(\mathbf{x}^*) = 0, \ \forall i = 1, ..., m.$$

The Lagrangian Dual of LP with Bound Constraints

Consider

$$
\begin{array}{lll}
(LP) & \text{minimize} & \mathbf{c}^T\mathbf{x} \\
 & \text{subject to} & A\mathbf{x} = \mathbf{b},\ -\mathbf{e} \le \mathbf{x} \le \mathbf{e}\ (\|\mathbf{x}\|_\infty \le 1);
\end{array}
$$

Let the Lagrangian multipliers be $\mathbf{y}$ for equalities. Then the Lagrangian dual objective would be

$$\phi(\mathbf{y}) = \inf_{-\mathbf{e}\le\mathbf{x}\le\mathbf{e}} L(\mathbf{x},\mathbf{y}) = \inf_{-\mathbf{e}\le\mathbf{x}\le\mathbf{e}} [(\mathbf{c} - A^T\mathbf{y})^T\mathbf{x} + \mathbf{b}^T\mathbf{y}].$$

If $(\mathbf{c} - A^T\mathbf{y})_j \le 0$, then $x_j^* = 1$; otherwise $x_j^* = -1$. Thus

$$
\begin{array}{lll}
(LDP) & \text{maximize} & \mathbf{b}^T\mathbf{y} + \|\mathbf{c} - A^T\mathbf{y}\|_1 \\
 & \text{subject to} & \mathbf{y} \in R^m.
\end{array}
$$

The Conic Duality vs. Lagrangian Duality I

Consider SOCP problem

$$(SOCP) \quad \text{minimize} \quad \mathbf{c}^T\mathbf{x}$$
$$\text{subject to} \quad A\mathbf{x} = \mathbf{b},\ x_1 - \|\mathbf{x}_{-1}\|_2 \geq 0;$$

and it conic dual problem

$$(SOCD) \quad \text{maximize} \quad \mathbf{b}^T\mathbf{y}$$
$$\text{subject to} \quad A^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ s_1 - \|\mathbf{s}_{-1}\|_2 \geq 0.$$

Let the Lagrangian multipliers be $\mathbf{y}$ for equalities and $s \geq 0$ for the single constraint $s_1 \geq \|\mathbf{s}_{-1}\|_2$. Then the Lagrangian function would be

$$L(\mathbf{x}, \mathbf{y}, s) = \mathbf{c}^T\mathbf{x} - \mathbf{y}^T(A\mathbf{x} - \mathbf{b}) - s(x_1 - \|\mathbf{x}_{-1}\|_2) = (\mathbf{c} - A^T\mathbf{y})^T\mathbf{x} - s(x_1 - \|\mathbf{x}_{-1}\|_2) + \mathbf{b}^T\mathbf{y}.$$

The Conic Duality vs. Lagrangian Duality II

Now consider the Lagrangian dual objective

$$\phi(\mathbf{y}, s) = \inf_{\mathbf{x} \in R^n} L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \inf_{\mathbf{x} \in R^n} [(\mathbf{c} - A^T\mathbf{y})^T\mathbf{x} - s(x_1 - \|\mathbf{x}_{-1}\|_2) + \mathbf{b}^T\mathbf{y}].$$

The objective function of the problem may not be differentiable so that the classical optimal condition theory do not apply. Consequently, it is difficult to write a clean form of the Lagrangian dual problem.

On the other hand, many nonlinear optimization problems, even they are convex, are difficult to transform them into CLP problems (especially to construct the dual cones). Therefore, each of the duality form, Conic or Lagrangian, has its own pros and cons.