Linear programming: Duality and interior point method

2012-10-9

Hao Wu

LP problem for resource allocation

For a typical maximizing LP problem like the following (with 3 variables and 2 constraints):

$$
\begin{aligned}
max \quad & c_1x_1 + c_2x_2 + c_3x_3 \\
s.t. \quad & a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \leq b_1 \\
& a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \leq b_2 \\
& x_1, x_2, x_3 \geq 0
\end{aligned}
$$

Economical interpretation of the problem:

- x_j: unit of production of product $j, j = 1, 2, 3$. Unknown to be obtained.
- c_j: profit per unit of product $j, j = 1, 2, 3$.
- a_{ij}: unit of material i ($i = 1, 2$) required to produce 1 unit of product j.
- b_i: unit of available material $i, i = 1, 2$.

The goal is to maximize the profit, subject to the material constraints.

Now assume a buyer consider to buy our entire inventory of materials but not sure how to price the materials, but s/he knows that we will only do the business if selling the materials yields higher return than producing the product.

Buyer's business strategy: for us producing one unit **less** of product j will save:

- a_{1j} unit of material 1, and
- a_{2j} unit of material 2.

Buyer want to compute the unit prices of materials to **minimize** his/her cost, subject to the constraints that we will do business (that we will not make less money). Assume the unit price for the materials are y_1 and y_2, the buyer will face the following optimization problem (called **Resource valuation problem**):

$$
\begin{aligned}
min \quad & b_1 y_1 + b_2 y_2 \\
s.t. \quad & a_{11} y_1 + a_{21} y_2 \geq c_1 \\
& a_{12} y_1 + a_{22} y_2 \geq c_2 \\
& a_{13} y_1 + a_{23} y_2 \geq c_3 \\
& y_1, y_2 \geq 0
\end{aligned}
$$

Duality

The buyer's LP problem is called the "**dual**" problem of the original problem, which is called the "**primal** problem".

In matrix notation, if the **primal** LP problem is:

$$\begin{aligned} max \quad & cx \\ s.t. \quad & Ax \leq b, x \geq 0 \end{aligned}$$

The corresponding dual problem is:

$$\begin{aligned} min \quad & b^T y \\ s.t. \quad & A^T y \geq c^T, y \geq 0 \end{aligned}$$

Or to express in the canonical form (a maximization problem with $\leq$ constraints):

$$\begin{aligned} max \quad & -b^T y \\ s.t. \quad & -A^T y \leq -c^T, y \geq 0 \end{aligned}$$

Dual is the "negative transpose" of the primal. It's easy to see, the dual of the dual problem is the primal problem.

Duality in non-canonical form

What if the primal problem doesn't fit into the canonical form, e.g., with $\geq$ or $=$ constraints, unrestricted variable, etc.The general rules of converting are:

- The variable types of the dual problem is determined by the constraints types of the primal:

Primal (max) constraints	Dual (min) variable
$\leq$	≥ 0
$\geq$	≤ 0
$=$	unrestricted

- The constraints types of the dual problem is determined by the variable types of the primal:

Primal (max) variable	Dual (min) constraints
≥ 0	$\geq$
≤ 0	$\leq$
$unrestricted$	$=$

Examples of duality in non-canonical form

If the primal problem is:

$$\begin{aligned}
max \quad & 20x_1 + 10x_2 + 50x_3 \\
s.t. \quad & 3x_1 + x_2 + 9x_3 \leq 10 \\
& 7x_1 + 2x_2 + 3x_3 = 8 \\
& 6x_1 + x_2 + 10x_3 \geq 1 \\
& x_1 \geq 0, x_2\ unrestricted, x_3 \leq 0
\end{aligned}$$

The dual problem is:

$$\begin{aligned}
min \quad & 10y_1 + 8y_2 + y_3 \\
s.t. \quad & 3y_1 + 7y_2 + 6y_3 \geq 20 \\
& y_1 + 2y_2 + y_3 = 10 \\
& 9y_1 + 3y_2 + 10y_3 \leq 50 \\
& y_1 \geq 0, y_2\ unrestricted, y_3 \leq 0
\end{aligned}$$

Weak duality

Weak duality Theorem: *the objective function value of the primal problem (max) at any feasible solution is always less than or equal to the objective function value of the dual problem (min) at any feasible solution.*

So if $(x_1, \ldots, x_n)$ is a feasible solution for the primal problem, and $(y_1, \ldots, ym)$ is a feasible solution for the dual problem, then $\sum_j c_j x_j \leq \sum_i b_i y_i$.

Proof:

$$\begin{aligned}
\sum_j c_j x_j &\leq \sum_j \left(\sum_i y_i a_{ij} \right) x_j \\
&= \sum_{ij} y_i a_{ij} x_j \\
&= \sum_i \left(\sum_j a_{ij} x_j \right) y_i \\
&\leq \sum_i b_i y_i .
\end{aligned}$$

Strong duality

- So we know now that at feasible solutions for both, the objective function of the dual problem is always greater or equal.
- A questions is that if there is a difference between the largest primal value and the smallest dual value? Such difference is called the “**Duality gap**”.
- The answer is provided by the following theorem.

Strong Duality Theorem: If the primal problem has an optimal solution, then the dual also has an optimal solution and there is no duality gap.

Economic interpretation: at optimality, the resource allocation and resource valuation problems give the same objective function values. In other words, in the ideal economic situation, to use the materials and generate the products or selling the materials give the same yields.

Sketch of the proof

Recall for the LP problem in standard form: $max\ z = cx,\ s.t.\ Ax \le b, x \ge 0$. Let x^* be an optimal solution. Let B be the columns of A associated with the BV, and R is the set of columns associated with the NBV. We have $c_B B^{-1} a_j - c_j \ge 0$ for $\forall j \in R$.

- Define $y^* \equiv (c_B B^{-1})^T$, we will have $y^{*T} A \ge c$ (why?).
- Furthermore, $y^{*T} \ge 0$ (why?).

Thus y^* is a feasible solution for the dual problem.

How about optimality? We know $y^{*T} b \ge y^{*T} Ax \ge cx$. Want to prove $y^* b = cx^*$. This requires two steps:

1. $y^{*T} b = y^{*T} Ax^* \Leftrightarrow y^{*T}(b - Ax^*) = 0$
2. $y^* Ax^* = cx^* \Leftrightarrow (y^{*T} A - c)x^* = 0$

This can be seen from the optimal tableau:

	z	x_B	x_N	RHS
z	1	0	$c_B B^{-1} N - c_N$	$c_B B^{-1} b$
x_B	0	I	$B^{-1} N$	$B^{-1} b$

An optimal solution $x^* = [x_B^*, 0]^T$, $A = [B, N]$, $c = [c_B, c_N]$:

1. becomes $c_B B^{-1} b - c_B B^{-1} [B, N][x_B^*, 0]^T = c_B B^{-1} b - c_B x_B^* = 0$
2. becomes $(c_B B^{-1}[B, N] - [c_B, c_N])[x_B^*, 0]^T = [0, c_B B^{-1} N - c_N][x_B^*, 0]^T = 0$

□

Complementary Slackness Theorem

The result obtained from proving the strong duality theorem is a theorem itself called "Complementary Slackness Theorem", which states:

If x^* and y^* are feasible solutions of primal and dual problems, then x^* and y^* are both optimal if and only if

1. $y^{*T}(b - Ax^*) = 0$
2. $(y^{*T}A - c)x^* = 0$

This implies that if a primal constraint is not "bounded", its corresponding variables in the dual problem must be 0, and vice versa.

This theorem is useful for solving LP problems, and the foundation of another class of LP solver called "interior problem" method.

Primal/Dual optimality conditions

Given the primal and dual problem with slack/surplus variables added:

Primal:

$$\begin{array}{ll} max & cx \\ s.t. & Ax + w = b, x, w \geq 0 \end{array}$$

Dual:

$$\begin{array}{ll} min & b^T y \\ s.t. & A^T y - z = c^T, y, z \geq 0 \end{array}$$

- The Complementary Slackness Theorem states that at optimal solution, we should have: $x_j z_j = 0, \forall j$, and $w_i y_i = 0, \forall i$.
- To put this in matrix notation, define $X = diag(x)$, which means X is a diagonal matrix with x_j as diagonal elements.
- Define e as a vector of 1's. Dimension of e depends on context.
- Now the complementary conditions can be written as: $XZe = 0$, $WYe = 0$.

We wil have the **optimality conditions** for the primal/dual problems as:

$$
\begin{aligned}
Ax + w - b &= 0 \\
A^T y - z - c^T &= 0 \\
XZe &= 0 \\
WYe &= 0 \\
x, y, w, z &\geq 0
\end{aligned}
$$

- The first two conditions are simple the constraints for primal/dual problems.
- The next two are complementary slackness.
- The last one is the non-negativity constraint.

Ignoring the non-negativity constraints, this is a set of $2n + 2m$ equations with $2n + 2m$ unknowns, which can be solved using Newton's method. Such approach is called "**primal-dual interior point method**".

Primal-dual interior point method

- The primal-dual interior point method finds the primal-dual optimal solution (x^*, y^*, w^*, z^*) by applying Newton's method to the primal-dual optimality conditions.
- The direction and length of the steps are modified in each step so that the non-negativity condition is strictly satisfied in each interation.

To be specific, define the following function $\mathbf{F} : \mathbb{R}^{2n+2m} \to \mathbb{R}^{2n+2m}$:

$$\mathbf{F}(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{z}) = \begin{bmatrix} A\mathbf{x} + \mathbf{w} - \mathbf{b} \\ A^T\mathbf{y} - \mathbf{z} - \mathbf{c}^T \\ XZe \\ WYe \end{bmatrix}$$

The goal is to find solution for $\mathbf{F} = \mathbf{0}$.

Applying Newton's method, if at iteration k the variables are $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{w}^k, \mathbf{z}^k)$, we obtain a search direction $(\delta\mathbf{x}, \delta\mathbf{y}, \delta\mathbf{w}, \delta\mathbf{z})$ by solving the linear equations:

$$\mathbf{F}'(\mathbf{x}^k, \mathbf{y}^k, \mathbf{w}^k, \mathbf{z}^k)\begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \\ \delta\mathbf{w} \\ \delta\mathbf{z} \end{bmatrix} = -\mathbf{F}(\mathbf{x}^k, \mathbf{y}^k, \mathbf{w}^k, \mathbf{z}^k).$$

Here $\mathbf{F}'$ is the Jacobian. At iteration k, the equations are:

$$\begin{bmatrix} A & \mathbf{0} & \mathbf{I} & -\mathbf{0} \\ \mathbf{0} & A^T & \mathbf{0} & -\mathbf{I} \\ Z & \mathbf{0} & \mathbf{0} & X \\ \mathbf{0} & W & Y & \mathbf{0} \end{bmatrix}\begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \\ \delta\mathbf{w} \\ \delta\mathbf{z} \end{bmatrix} = \begin{bmatrix} -A\mathbf{x}^k - \mathbf{w}^k + \mathbf{b} \\ -A^T\mathbf{y}^k + \mathbf{z}^k + \mathbf{c}^T \\ -X^kZ^ke \\ -W^kY^ke \end{bmatrix}$$

Then the update will be obtained as: $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{w}^k, \mathbf{z}^k) + \alpha(\delta\mathbf{x}, \delta\mathbf{y}, \delta\mathbf{w}, \delta\mathbf{z})$ with $\alpha \in (0, 1]$. α is chosen so that the results from the next iterate is feasible.

It is reasonable to assume that at current iteration, both primal and dual are strictly feasible, then the first two terms on the right hand side are 0.

The algorithm in its current setup is not ideal because often only a small step can be taken before the positivity constraints are violated. A more flexible version is proposed as follow.

- The value of $XZe + WYe$ represents the duality gap.
- Instead of trying to eliminate the duality gap, reducing the duality gap by some factor in each step.

In order word, we replace the complementary slackness by:

$$XZe = \mu_x e$$
$$WYe = \mu_y e$$

When $\mu_x, \mu_y \to 0$ as $k \to \inf$, the solution from this system will converge to the optimal solution of the original LP problem. Easy selections of μ's are $\mu_x^k = (\mathbf{x}^k)^T z/n$ and $\mu_y^k = (\mathbf{w}^k)^T y/m$. Here n and m are dimensions of $\mathbf{x}$ and $\mathbf{y}$ respectively.

Under the new algorithm, at the kth iteration, the Newton equations become:

$$\begin{bmatrix} A & \mathbf{0} & \mathbf{I} & -\mathbf{0} \\ \mathbf{0} & A^T & \mathbf{0} & -\mathbf{I} \\ Z & \mathbf{0} & \mathbf{0} & X \\ \mathbf{0} & W & Y & \mathbf{0} \end{bmatrix} \begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{y} \\ \delta\mathbf{w} \\ \delta\mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -X^k Z^k e + \mu_x^k e \\ -W^k Y^k e + \mu_y^k e \end{bmatrix} \tag{1}$$

This provides the **general primal-dual interior point method** as follow:

1. Choose strictly feasible initial solution $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{w}^0, \mathbf{z}^0)$, and set $k = 0$.
2. **Repeat**:
3. Solve system (1) to obtain the updates $(\delta\mathbf{x}, \delta\mathbf{y}, \delta\mathbf{w}, \delta\mathbf{z})$.
4. Update the solution: $(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}, \mathbf{w}^{k+1}, \mathbf{z}^{k+1}) = (\mathbf{x}^k, \mathbf{y}^k, \mathbf{w}^k, \mathbf{z}^k) + \alpha^k(\delta\mathbf{x}, \delta\mathbf{y}, \delta\mathbf{w}, \delta\mathbf{z})$. α^k is chosen so that all variables are greater than or equal to 0.
5. Set $k = k + 1$.
6. **Until converge**.

The Barrier Problem

The interior point algorithm is closely related to the **Barrier Problem**. Go back to the primal problem:

$$max\ z = cx, s.t.\ Ax + w = b, x, w \geq 0.$$

The non-negativity constraints can be replaced by adding two **barrier terms** in the objective function. The barrier term is defined as $B(\mathbf{x}) = \sum_j \log x_j$, which is finite as long as x_j is positive. Then the primal problem becomes:

$$\begin{aligned} max\ z &= cx + \mu_x B(\mathbf{x}) + \mu_y B(\mathbf{w}) \\ s.t.\ Ax &+ w = b. \end{aligned}$$

The barrier terms make sure x and w won't become negative.

Before trying to solve this problem, we need some knowledge about **Lagrange multiplier**.

The method is Lagrange multiplier is a general algorithm for optimization problems with equality constraints. For example, consider a problem:

$$max\ f(x, y)$$
$$s.t.\ g(x, y) = c$$

We introduce a new variable λ called **Lagrange multiplier** and form the following new objective function :

$$L(x, y, \lambda) = f(x, y) + \lambda[g(x, y) - c]$$

We will then optimize L with respect to x, y and λ using typical method. Note that the condition $\partial L/\partial\lambda = 0$ at optimal solution guarantees that the constraints will be satisfied.

The Barrier Problem (cont.)

Go back to the barrier problem, the Lagrangian for this problem is (using $\mathbf{y}$ as the multiplier):

$$L(\mathbf{x}, \mathbf{y}, \mathbf{w}) = c\mathbf{x} + \mu_x B(\mathbf{x}) + \mu_y B(\mathbf{w}) + \mathbf{y}^T(b - w - Ax).$$

The optimal solution for the problem satisfies:

$$\begin{aligned} c + \mu_x X^{-1}e - A^T\mathbf{y} &= 0 \\ \mu_y W^{-1}e - \mathbf{y} &= 0 \\ b - w - Ax &= 0 \end{aligned}$$

Define new variables $z = \mu_x X^{-1}e$ and rewrite these conditions, we obtain exactly the same set of equations as the relaxed optimality conditions for primal-dual problem (finish!).

Summary

- The major difference between interior point (IP) method and simplex is that IP goes through the middle of the feasible space, and Simplex moves along the edges and only stops at the extreme points.
- IP method has much better performance in large-scale LP problems.
- IP is a general optimization method which also works for non-linear programming problem.