More LP Optimality Condition Theory

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

Read Chapters 2.3-2.5, 4.1-4.2

LP in Standard (**Equality**) Form

$$\begin{aligned}
\min \quad & c^T x = \sum_{j=1}^{n} c_j x_j \\
\text{s.t.} \quad & a_1 x = \sum_{j=1}^{n} a_{1j} x_j = b_1 \\
& a_2 x = \sum_{j=1}^{n} a_{2j} x_j = b_2 \\
& \ldots \\
& a_m x = \sum_{j=1}^{n} a_{mj} x_j = b_m \\
& x \geq 0
\end{aligned}$$

⇒

$$\begin{aligned}
\min \quad & c^T x \\
\text{s.t.} \quad & Ax = b, \\
& x \geq 0
\end{aligned}$$

Reduction to Standard Form

- max $\mathbf{c}^T\mathbf{x}$ to min $-\mathbf{c}^T\mathbf{x}$
- Eliminating "free" variables: substitute with the difference of two nonnegative variables

$$x := x' - x'', \quad (x', x'') \geq 0.$$

- Eliminating inequalities: add a slack variable

$$\mathbf{a}^T\mathbf{x} \leq b \implies \mathbf{a}^T\mathbf{x} + s = b, \quad s \geq 0$$

$$\mathbf{a}^T\mathbf{x} \geq b \implies \mathbf{a}^T\mathbf{x} - s = b, \quad s \geq 0$$

Reduction of the Production Problem

$$
\begin{aligned}
\max \quad & x_1 + 2x_2 \\
\text{s.t.} \quad & x_1 \le 1 \\
& x_2 \le 1 \\
& x_1 + x_2 \le 1.5 \\
& x_1 \ge 0 \\
& x_2 \ge 0
\end{aligned}
$$

$\Rightarrow$

$$
\begin{aligned}
\textbf{min} \quad & -x_1 - 2x_2 \\
\textbf{s.t.} \quad & x_1 + x_3 = 1 \\
& x_2 + x_4 = 1 \\
& x_1 + x_2 + x_5 = 1.5 \\
& (x_1, x_2, x_3, x_4, x_5) \ge 0
\end{aligned}
$$

x_3, x_4,and x_5 are called slack variables

We know how to identify corners/extreme-points of the LP feasible region defined all by linear inequalities. What about corners in this LP standard equality form?

How to Identify Corners in LP Equality Form Basic and Basic Feasible Solution

In the LP standard form, select m linearly independent columns, denoted by the variable index set B, from A. Solve

$$A_B \mathbf{x}_B = \mathbf{b}$$

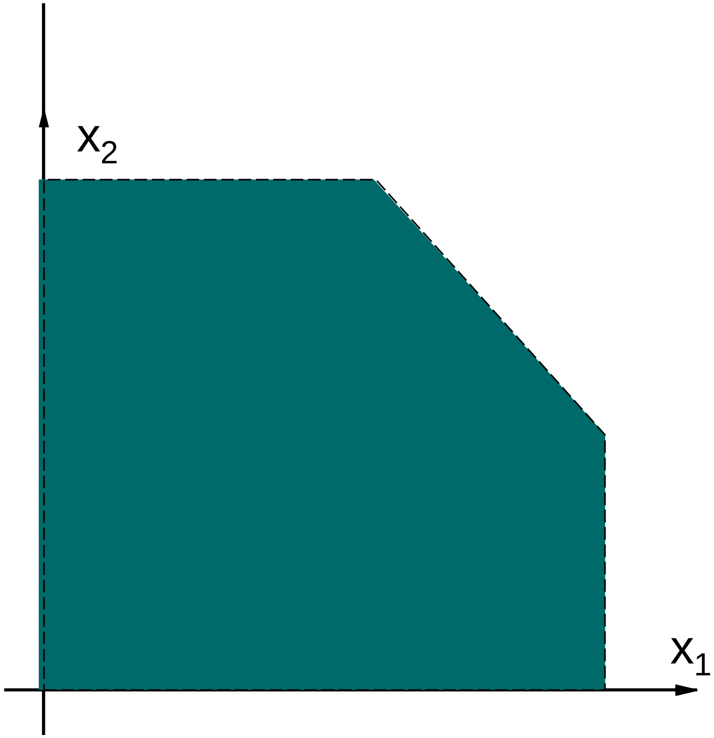
for the dimension-m vector $\mathbf{x}_B$. By setting the variables, $\mathbf{x}_N$, of $\mathbf{x}$ corresponding to the remaining columns of A equal to zero, we obtain a solution $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$.

Then, $\mathbf{x}$ is said to be a basic solution to (LP) with respect to the basic variable set B. The variables in $\mathbf{x}_B$ are called basic variables, those in $\mathbf{x}_N$ are nonbasic variables, and A_B is called a basis.

If a basic solution $\mathbf{x}_B \geq \mathbf{0}$, then $\mathbf{x}$ is called a basic feasible solution, or BFS. Note that A_B and $\mathbf{x}_B$ follow the same index order in B.

Two BFS are adjacent if they differ by exactly one basic variable.

BS of the Production Problem in Equality Form



$$
\begin{array}{cccccc}
x_1 & & +x_3 & & & = 1 \\
& x_2 & & +x_4 & & = 1 \\
x_1 & +x_2 & & & +x_5 & = 1.5 \\
(x_1, & x_2, & x_3, & x_4, & x_5) & \geq 0
\end{array}
$$

Basis	3,4,5	1,4,5	3,4,1	3,2,5	3,4,2	1,2,3	1,2,4	1,2,5
Feasible?	√	√		√		√	√	
x_1 , x_2	0, 0	1, 0	1.5, 0	0, 1	0, 1.5	.5, 1	1, .5	1, 1

BFS and Corner Point Equivalence Theorem

Theorem *Consider the feasible region in the standard LP form. Then, a basic feasible solution and a corner (extreme) point are equivalent; the formal is algebraic and the latter is geometric. Moreover, Two corners are neighboring if exact one variable difference in basis*

- Feasible directions of an BFS: an increasing direction of the nonbasic variables (they equal 0 right now).
- Extreme feasible direction: the increasing direction of a nonbasic variable x_j: $\mathbf{x}_B = (A_B)^{-1}\mathbf{b} - (A_B)^{-1}\mathbf{a}_j x_j$
- Optimality test: No improving (extreme) feasible direction exists

Feasible Directions at a BFS and Optimality Test

- Recall at a BFS: $A_B x_B + A_N x_N = b$, and $x_B \geq 0$ and $x_N = 0$.

Thus we can express x_B in terms of x_N,

$$x_B = (A_B)^{-1}b - (A_B)^{-1}A_N x_N .$$

Reduced Objective

Then, $c^T x = c^T_B x_B + c^T_N x_N = (c^T_N - c^T_B (A_B)^{-1} A_N)\, x_N + c^T_B (A_B)^{-1} b$

- Note that increase any one variable of x_N is an extreme feasible direction. Thus, for the BFS to be optimal, any (extreme) feasible direction must be an ascent direction, or

$$(c^T_N - c^T_B (A_B)^{-1} A_N) \geq 0$$

is necessary and sufficient for the current BFS being optimal!

- This vector is called the reduced cost coefficient vector or reduced gradient vector from the current BFS. Note that reduced cost coefficients for basic variables are all zeros.

The Shadow-Price and Reduced Cost Vectors

We first introduce and compute an intermediate shadow-price/multiplier vector:

$$y^T = c^T_B (A_B)^{-1} \text{ , or } y^T A_B = c^T_B,$$

by solving a system of linear equations.

Then we compute reduced cost $r^T = c^T - y^T A$, where r_N is the reduced cost vector for nonbasic variables (and $r_B = 0$ always).

If one of r_N is negative, then an improving (extreme) feasible direction is find by increasing the corresponding nonbasic variable value.

In the LP production example, suppose the basic variable set $B = \{3, 4, 5\}$.

$$\begin{array}{lllllll} \min & -x_1 & -2x_2 & & & & \\ \text{s.t.} & x_1 & & +x_3 & & & = 1 \\ & & x_2 & & +x_4 & & = 1 \\ & x_1 & +x_2 & & & +x_5 & = 1.5 \\ & x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0. \end{array}$$

$$c_N = \begin{pmatrix} -1 \\ -2 \end{pmatrix}, c_B = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, A_B = I, A_N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix},$$

$$A_B^{-1} = I, y^T = (0 \quad 0 \quad 0), \quad r_N^T = (-1 \quad -2).$$

Thus, increasing either x_1 and x_2 is a feasible and improving direction and the variable is called the incoming basic variable…

In the LP production example, suppose the basic variable set $B = \{1, 2, 3\}$.

$$\begin{array}{llllllll} \min & -x_1 & -2x_2 & & & & \\ \text{s.t.} & x_1 & & +x_3 & & & = 1 \\ & & x_2 & & +x_4 & & = 1 \\ & x_1 & +x_2 & & & +x_5 & = 1.5 \\ & x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0. \end{array}$$

$$c_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, c_B = \begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}, A_B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}, A_N = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$A_B^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}, \quad y^T = (0 \;\; -1 \;\; -1), \quad r_N^T = (1 \;\; 1).$$

Thus, this BFS is optimal

Summary of BFS Optimality Test/Condition

When a BFS with basis B, $\mathbf{x}_B$, is optimal?

$$\mathbf{x}_B = (A_B)^{-1}\mathbf{b} \geq 0, \ \mathbf{x}_N = 0$$

$$\mathbf{r}^T = \mathbf{c}^T - \mathbf{y}^T A \geq 0$$

where the shadow-price/multiplier vector $\boldsymbol{y}^T = \mathbf{c}_B^T (A_B)^{-1}$.

Moreover $\quad$ OV = $\mathbf{c}^T \boldsymbol{x} = \mathbf{c}_B^T \boldsymbol{x}_B = \mathbf{c}_B^T (A_B)^{-1} \boldsymbol{b} = \boldsymbol{y}^T \boldsymbol{b}$

The existence of such a shadow-price/multiplier vector $\boldsymbol{y}$ is served as a certificate of the optimality of corner feasible solution $\boldsymbol{x}$. Such a $\mathbf{y}$ is also called optimal shadow-price vector.

Dos this optimal test/condition apply to any feasible solution $\boldsymbol{x}$?

The Optimality Condition Theorem

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b, \\ & x \geq 0 \end{array}$$

Theorem A feasible solution x in the LP standard equality form is optimal if and only if there is an optimal shadow-price vector y such that:

$$\left\{ (\mathbf{x}, \mathbf{y}) \in (R^n, R^m) : \begin{array}{ll} \mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} = 0 & \\ A\mathbf{x} & = \mathbf{b}, \mathbf{x} \geq 0 \\ A^T \mathbf{y} & \leq \mathbf{c} \end{array} \right\},$$

This is a system of linear inequalities and equations. Thus it is easy to verify whether or not a pair **(x, y)** is optimal by a computer.

Sketch Proof of The Optimality Condition Theorem

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \ge 0 \end{aligned}$$

Consider any vector $\mathbf{y}$ who satisfies

$$A^T \mathbf{y} \le \mathbf{c}.$$

Then for any feasible solution $\mathbf{x}$ in the LP standard equality form, we must have

$$\mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} = \mathbf{c}^T \mathbf{x} - (A\mathbf{x})^T \mathbf{y} = \mathbf{c}^T \mathbf{x} - \mathbf{y}^T (A\mathbf{x}) = (\mathbf{c}^T - \mathbf{y}^T A)\mathbf{x} \ge 0.$$

That is, the value $\mathbf{b}^T \mathbf{y}$ is a lower bound on any feasible objective value $\mathbf{c}^T \mathbf{x}$.

Thus, if $\mathbf{c}^T \mathbf{x} = \mathbf{b}^T \mathbf{y}$, $\mathbf{c}^T \mathbf{x}$ must be the minimal among all possible feasible solution $\mathbf{x}$.
(Of course, $\mathbf{b}^T \mathbf{y}$ must be maximal among all possible $\mathbf{y}$ such that $A^T \mathbf{y} \le \mathbf{c}$, which is called the dual program; more on this later.)

An Equivalent Optimality Condition

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b, \\ & x \geq 0 \end{array}$$

A feasible solution $\boldsymbol{x}$ in the LP standard equality form is optimal if and only if there are vectors $(\boldsymbol{y}, \boldsymbol{r})$ such that:

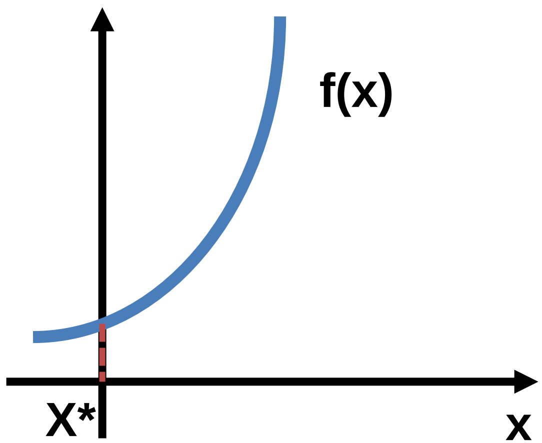
$$\left\{ (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{r}) \in (R^n, R^m, R^n) : \begin{array}{ll} \mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{y} = 0 \text{ or } \boldsymbol{r}^T \boldsymbol{x} = \mathbf{0} & \\ A\mathbf{x} & = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \\ A^T \mathbf{y} + \mathbf{r} & = \mathbf{c}, \mathbf{r} \geq \mathbf{0} \end{array} \right\},$$

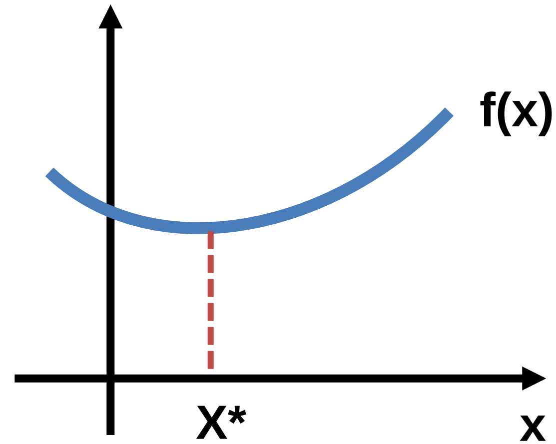
Since $\boldsymbol{r}^T\boldsymbol{x} = \sum_{j=1}^{n} r_j x_j$ and all entries are nonnegative, the condition implies that $r_j x_j = 0$ for all *j*; that is, for each *j*, at least one of r_j and x_j is 0. We often call this a complementarity property: two nonnegative vectors **r** and ***x*** are complementary to each other.

Physical Explanation of Complementarity Condition

Complementarity or Complementary-Slackness Phenomenon typically occurs when optimization with inequality constraints.

Consider min $f(x)$, s.t. $x \geq 0$





Two possible Scenarios:

$x^*= 0$ & $f'(0) \geq 0$ or $x^*>0$ & $f'(x^*)=0$

In both cases, the complementarity condition holds:
first, the derivative at the minimizer must be **nonnegative**;
second, it must be zero if the minimizer is in the interior of the constraint set, that is, the product of the derivative and the slack value mut be zero

Interpretation of *y*: Shadow Price Vector of RHS *b*

Given a BFS in the LP standard form with basis A_B

$$\mathbf{x}_B = (A_B)^{-1}\mathbf{b} > \mathbf{0}, \qquad \mathbf{x}_N = \mathbf{0},$$

so that small change in **b** does not change the optimal basis and the shadow price vector remains:

$$\mathbf{y}^T = c_B^T(A_B)^{-1}$$

At optimality, the OV is a function of **b:**

$$\mathbf{c}^T\mathbf{x} = \mathbf{c}_B^T\mathbf{x}_B = \mathbf{c}_B^T(A_B)^{-1}\mathbf{b} = \mathbf{y}^T\mathbf{b}.$$

Thus, when **b** is changed to **b**+Δ**b**, then the new OV

$$OV_+ = \mathbf{c}_B^T\mathbf{x}_B = \mathbf{c}_B^T(A_B)^{-1}(\mathbf{b}+\Delta\mathbf{b}) = \mathbf{y}^T(\mathbf{b}+\Delta\mathbf{b}) = OV + \boxed{\mathbf{y}^T\Delta\mathbf{b}}$$

when the basis is unchanged.

=Net Change

*OV(**b**)* is a convex function of ***b***

and $\nabla OV(\mathbf{b}) = \mathbf{y}^*$

$$OV(b) := \min \quad c^T x \quad \text{s.t.} \quad Ax = b, \quad x \geq 0$$

Recall in the LP production example, the BFS with $B = \{1, 2, 3\}$ is optimal with $\mathbf{x} = (½, 1, ½, 0, 0)^T$ and $\mathbf{y} = (0, -1, -1)^T$

$$
\begin{array}{llllllll}
\min & -x_1 & -2x_2 & & & & \\
\text{s.t.} & x_1 & & +x_3 & & & = 1 \\
 & & x_2 & & +x_4 & & = 1 \\
 & x_1 & +x_2 & & & +x_5 & = 1.5 \\
 & x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0.
\end{array}
$$

The current OV= -2.5

- If b_1 is increased or decreased a little, does OV change?
- If b_2 is increased or decreased a little, does OV change? How much?
- If b_3 is increased or decreased a little, does OV change? How much?

This is called sensitivity analyses and an economical interpretation of $\mathbf{y}$

.

The Lagrange Function and Theory

The **Lagrange Function or Lagrangian** was introduced for a constrained optimization problem to make it into a less constrained or unconstrained optimization problem.

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax - b = 0, \\ & x \geq 0 \end{aligned}$$

For LP in the standard equality form:

$$\min_x L(\mathbf{x},\mathbf{y}) = c^T\mathbf{x} - \mathbf{y}^T(A\mathbf{x}-\mathbf{b}), \text{ s.t. } \mathbf{x} \geq \mathbf{0};$$

where multipliers $\mathbf{y}$ in the Lagrange function play the role of penalty weights on equality constraint violations. One can adjust them whenever a constraint is violated at the minimizer of the Lagrange optimization problem (in the text book, we used $\boldsymbol{\lambda}=-\mathbf{y}$)

The goal is to select $\mathbf{y}$ such that the Lagrangian minimization yields an optimal solution $\mathbf{x}$ that is also optimal to the original constrained problem. It turns out such $\mathbf{y}$ must be an optimal shadow-price vector of the original LP problem.

$$
\begin{array}{lllllll}
\min & -x_1 & -2x_2 & & & & \\
\text{s.t.} & x_1 & & +x_3 & & & = 1 \\
 & & x_2 & & +x_4 & & = 1 \\
 & x_1 & +x_2 & & & +x_5 & = 1.5 \\
 & x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0.
\end{array}
$$

For this example, the Lagrange function would be

$L(\boldsymbol{x},\boldsymbol{y}) = -x_1 - 2x_2 - y_1(x_1+x_3-1) - y_2(x_2+x_4-1) - y_3(x_1+x_2+x_5-1.5)$

If we set $\boldsymbol{y}=(0,-1,-1)^T$ then $L(\boldsymbol{x}, (0,-1,-1)^T) = x_4 + x_5 - 2.5$.

Therefore, minimize it subject to each variable to be nonnegative implies $x_4 = 0$ and $x_5 = 0$. Together with the equality constraints $A\boldsymbol{x}=\boldsymbol{b}$ in the original problem they yield the optimal solution for the original optimization problem.

Any other setting of $\boldsymbol{y}$ values in the Lagrangian will not make $\boldsymbol{x}$ possibly feasible to the original problem - either an x_j is negative or $A\boldsymbol{x} \neq \boldsymbol{b}$.

LP Optimality Condition via the Lagrange Function

$$L(\boldsymbol{x},\boldsymbol{y}) = c^T\boldsymbol{x} - \boldsymbol{y}^T(A\boldsymbol{x}-\boldsymbol{b})$$

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax - b = 0, \\ & x \geq 0 \end{aligned}$$

Note the gradient vector with respect to $\boldsymbol{x}$ is

$$\nabla_x L(\boldsymbol{x},\boldsymbol{y}) = c - A^T\boldsymbol{y}$$

which we also called the reduced cost vector $\boldsymbol{r}$.

Theorem A feasible solution $\boldsymbol{x}$ in the LP standard equality form is optimal if and only if there is multiplier vector $\boldsymbol{y}$ such that the gradient vector with respect to $\boldsymbol{x}$ of the Lagrange function is nonnegative and it is complementary to $\boldsymbol{x}$.

In fact, consider any $\min_x f(\boldsymbol{x})$ s.t. $\boldsymbol{x} \geq \boldsymbol{0}$.
$\boldsymbol{x}$ is minimal only if $\nabla f(\boldsymbol{x}) \geq \boldsymbol{0}$ and it is complementary to $\boldsymbol{x}$.

Handling Nonnegative Variables as Constraints

The Lagrange Function or Lagrangian can include variable nonnegativity as part of the constraints so that the variable are all "free". Then for LP in the standard equality form:

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

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax - b = 0, \\ & x \geq 0 \end{array}$$

where multipliers in $\mathbf{y}$ in the Lagrange function are penalty weights on equality constraint violations, and multipliers in $\mathbf{r}$ are penalty weights on inequality constraints $\mathbf{x} \geq \mathbf{0}$. Note that each penalty weight in $\mathbf{r}$ must be nonnegative since we only penalize the corresponding entry in $\mathbf{x}$ who becomes negative but no penalty otherwise.

Theorem A feasible solution $\mathbf{x}$ in the LP standard equality form is optimal if and only if there are multiplier vectors $\mathbf{y}$ and $\mathbf{r} \geq \mathbf{0}$ such that the gradient vector with respect to $\mathbf{x}$ of the Lagrange function is a zero vector and $\mathbf{r}$ is complementary to $\mathbf{x}$.