Lesson 12: Linear programming

Theme: Linear programming.

1 Linear programming (LP)

An LP instance is defined as follows.

Input: $(m \times n)$-matrix A with real entries, and $\bar{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$, $\bar{b} = (b_1, \ldots, b_m) \in \mathbb{R}^m$.
Task: Find $\bar{x} = (x_1, \ldots, x_n) \geqslant (0, \ldots, 0)$ that

$$\begin{array}{ll} \text{minimizes} & c_1 x_1 + \cdots + c_n x_n \\ \text{subject to} & A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \geqslant \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \end{array}$$

We will write $(A, \bar{b}, \bar{c})$ to denote the above LP instance.

The function $f(\bar{x}) = \bar{c} \cdot \bar{x}$ is called the objective function of LP, and the inequalities $A\bar{x}^t \geqslant \bar{b}^t$ and $\bar{x} \geqslant 0$ are called the constraints of LP. Any $\bar{x}$ that satisfies the constraints is called a feasible solution, or simply, a solution. A solution $\bar{x}$ such that $f(\bar{x})$ is minimal is called an optimal solution.

2 LP duality

The dual of an LP instance (in Section 1) is defined with the same input, but with the task to find $\bar{y} = (y_1, \ldots, y_m) \geqslant (0, \ldots, 0)$ that

$$\begin{array}{ll} \text{maximizes} & b_1 y_1 + \cdots + b_m y_m \\ \text{subject to} & A^t \bar{y}^t \leqslant \bar{c}^t \end{array}$$

Similar to above, any $\bar{y}$ that satisfies the constraints $A^t \bar{y}^t \geqslant \bar{c}^t$ and $\bar{y} \geqslant 0$ is called a solution, and a solution $\bar{y}$ with maximal $\bar{b} \cdot \bar{y}$ is called an optimal solution.

Theorem 12.1 (LP weak-duality) *If $\bar{x}_0$ is a solution to an LP instance $(A, \bar{b}, \bar{c})$, and $\bar{y}_0$ is a solution to its dual, then $\bar{x}_0 \cdot \bar{c} \geqslant \bar{y}_0 \cdot \bar{b}$. Hence, if $\bar{x}_0 \cdot \bar{c} = \bar{y}_0 \cdot \bar{b}$, then $\bar{x}_0$ and $\bar{y}_0$ are the optimal solutions for $(A, \bar{b}, \bar{c})$ and its dual, respectively.*

Proof. We have:

$$\bar{c}\ \bar{x}_0^t \quad \geqslant \quad \bar{y}_0\ A\ \bar{x}_0^t \quad \geqslant \quad \bar{y}_0\ \bar{b}^t$$

The first inequality comes from $\bar{y}_0\ A \leqslant \bar{c}$, while the second from $A\ \bar{x}_0^t \geqslant \bar{b}^t$ and $\bar{y}_0 \geqslant 0$. ■

Theorem 12.2 (LP strong-duality) *If $\bar{x}_0$ is an optimal solution to an LP instance $(A, \bar{b}, \bar{c})$, and $\bar{y}_0$ is an optimal solution to its dual, then $\bar{x}_0 \cdot \bar{c} = \bar{y}_0 \cdot \bar{b}$.*

Proof. Let $\delta = \bar{c} \cdot \bar{x}_0$, i.e., the minimal possible value for the objective function of $(A, \bar{b}, \bar{c})$.

We want to show that there is $\bar{y} \geqslant 0$ such that $A^t \bar{y}^t \leqslant \bar{c}^t$ and $\bar{y} \cdot \bar{b} \geqslant \delta$, which is equivalent to:

$$\begin{pmatrix} A^t \\ -\bar{b} \end{pmatrix} \bar{y}^t \leqslant (\bar{c}, -\delta)^t$$
$$\bar{y} \geqslant 0$$

Suppose there is no such $\bar{y}$. By Farkas' lemma (version 2), there is $\bar{x}, z \geqslant 0$ such that

$$(A \mid -\bar{b}^t)\ (\bar{x}, z)^t \ \geqslant\ 0 \qquad \text{and} \qquad (\bar{c}, -\delta) \cdot (\bar{x}, z)\ <\ 0.$$

which is equivalent to:

$$A\ \bar{x}^t \ -\ z\bar{b}^t \ \geqslant\ 0 \qquad \text{and} \qquad \bar{c} \cdot \bar{x} \ -\ \delta z \ <\ 0.$$

There are two cases.

- When $z > 0$.

  Since $A\bar{x}^t - z\bar{b}^t \geqslant 0$, we have $A\bar{x}^t \geqslant z\bar{b}^t$, and so, $A(\bar{x}/z)^t \geqslant \bar{b}^t$. Therefore, $(\bar{x}/z)$ is a solution to $(A, \bar{b}, \bar{c})$. Moreover, $\bar{c} \cdot (\bar{x}/z) < \delta$, which contradicts the minimality of δ.

- When $z = 0$.

  Then, $A\bar{x}^t \geqslant 0$ and also, $\bar{c} \cdot \bar{x} < 0$. Furthermore, for every $\xi > 0$,

  $$A(\bar{x}_0 + \xi\bar{x}) \ =\ A\bar{x}_0 + \xi A\bar{x} \ \geqslant\ A\bar{x}_0 \ \geqslant\ \bar{b}.$$

  So, $\bar{x}_0 + \xi\bar{x}$ is also a solution to $(A, \bar{b}, \bar{c})$. Now,

  $$(\bar{x}_0 + \xi\bar{x}) \cdot \bar{c} \ =\ \bar{x}_0 \cdot \bar{c} + \xi(\bar{x} \cdot \bar{c}) \ =\ \delta + \xi(\bar{x} \cdot \bar{c}).$$

  Since $\bar{x} \cdot \bar{c} < 0$, this contradicts the minimality of δ.

Therefore, there is some $\bar{y} \geqslant 0$ such that $A^t\bar{y}^t \leqslant \bar{c}^t$ and $\bar{y} \cdot \bar{b} \geqslant \delta$. Since $\bar{y}_0$ is an optimal solution, $\bar{y}_0 \cdot \bar{b} \geqslant \bar{y} \cdot \bar{b} \geqslant \delta$. By weak duality, $\bar{y}_0 \cdot \bar{b} \leqslant \delta$, which implies $\bar{y}_0 \cdot \bar{b} = \delta = \bar{x}_0 \cdot \bar{c}$. ■

Corollary 12.3 (Complementary slackness) *If $\bar{x}_0 = (x_{0,1}, \ldots, x_{0,n})$ and $\bar{y}_0 = (y_{0,1}, \ldots, y_{0,m})$ are optimal solutions to an LP instance $(A, \bar{b}, \bar{c})$ and its dual, respectively, then we have the following.*

- *$\bar{c}$ and $\bar{y}_0 A$ agree on the coordinates in $\{i \mid x_{0,i} \neq 0\}$.*
- *$A\bar{x}_0^t$ and $\bar{b}^t$ agree on the coordinates in $\{j \mid y_{0,j} \neq 0\}$.*

In particular, if $\bar{x}_0 > 0$ and $\bar{y}_0 > 0$, then $A\bar{x}_0^t = \bar{b}^t$ and $A^t\bar{y}_0^t = \bar{c}^t$.

Proof. As in the proof of weak duality, $\bar{c}\ \bar{x}_0^t \ \geqslant\ \bar{y}_0\ A\ \bar{x}_0^t \ \geqslant\ \bar{y}_0\ \bar{b}^t$, and by strong duality, $\bar{c}\ \bar{x}_0^t \ =\ \bar{y}_0\ \bar{b}^t$, and hence, all the inequalities become equalities. In particular, $(\bar{c} - \bar{y}_0 A) \cdot \bar{x}_0 = 0$.

Since $\bar{c} \geqslant \bar{y}_0 A$, we have $\bar{c} - \bar{y}_0 A \geqslant 0$. Furthermore, $\bar{x}_0 \geqslant 0$. Therefore, $\bar{c}$ and $\bar{y}_0 A$ must agree on the coordinates in $\{i \mid x_{0,i} \neq 0\}$. The second item can be proved in a similar manner. ■

References

[1] V. Chvátal. *Linear Programming.* W. H. Freeman, 1983.

[2] T. Cormen, C. Leiserson, R. Rivest, and C. Stein. *Introduction to Algorithms (3. ed.).* MIT Press, 2009.

Appendix

For more details about LP, see, for example, [1, 2].

Theorem 12.4 (Farkas' lemma, version 1) *Let $\bar{x}$ and $\bar{y}$ be vectors of variables, and M and $\bar{a}$ are matrix and vectors of appropriate length. Exactly one of the following systems of linear inequalities has a solution, but not both.*

- $M\ \bar{y}^t = \bar{a}^t$ *and* $\bar{y} \geqslant 0$.
- $M^t\ \bar{x}^t \geqslant 0$ *and* $\bar{a} \cdot \bar{x} < 0$.

Proof. That it is impossible for both of them to have solution follows from the following.

$$M\bar{y}^t = \bar{a}^t \quad \Rightarrow \quad \bar{x}M\bar{y}^t = \bar{x}\bar{a}^t \quad < \quad 0$$

On the other hand, $\bar{x}M\bar{y}^t \geqslant 0\bar{y}^t = 0$, which implies $0 < 0$.

To show that one of them has a solution, suppose that $M\bar{y}^t = \bar{a}^t$ and $\bar{y} \geqslant 0$ has no solution. Define the space:

$$\Pi \quad := \quad \{M\bar{y}^t \mid \bar{y} \geqslant 0\}$$

Then, $\bar{a} \notin \Pi$. Let $\bar{p}$ be the point in Π with the minimal distance to $\bar{a}$ (among all the points in Π). Since the space Π is convex, for every $\bar{y} \geqslant 0$,

$$(\bar{p} - \bar{a}) \cdot (M\bar{y}^t - \bar{p}) \quad \geqslant \quad 0 \tag{1}$$

$$(\bar{p} - \bar{a})M(\bar{y}^t - \bar{u}^t) \quad \geqslant \quad 0 \tag{2}$$

where $\bar{u} \geqslant 0$ is such that $\bar{u}M^t = \bar{p}$. If there is $\bar{y} \geqslant 0$ such that $(\bar{p} - \bar{a})M(\bar{y}^t - \bar{u}^t) < 0$, there will be a point in Π strictly closer to $\bar{a}$ than $\bar{p}$, which contradicts the assumption that $\bar{p}$ is the closest point to $\bar{a}$ from among all the points in Π.

We claim that $\bar{x} = \bar{p} - \bar{a}$ is a solution for $M^t\bar{x}^t \geqslant 0$ and $\bar{x} \cdot \bar{a} < 0$. From Equation 2,

$$\bar{x}M(\bar{y}^t - \bar{u}^t) \quad \geqslant \quad 0 \qquad\qquad \text{for every } \bar{y} \geqslant 0$$

If $\bar{y} = e_i$, i.e., the unit vector with the i-th coordinate being 1, then the i-th coordinate of $\bar{x}M \geqslant 0$. Thus,

$$M^t\bar{x}^t \quad = \quad \bar{x}M \quad \geqslant \quad 0.$$

Now, we have to show that $\bar{x} \cdot \bar{a} < 0$. Note that by plugging in $\bar{y} = 0$ in Equation 2,

$$0 \quad \geqslant \quad (\bar{a} - \bar{p})M(0 - \bar{u}^t) \quad = \quad (-\bar{x}) \cdot (-M\bar{u}^t) \quad = \quad \bar{x} \cdot \bar{p} \tag{3}$$

Thus,

$$\bar{x} \cdot \bar{a} \quad = \quad \bar{x} \cdot (\bar{p} - \bar{x}) \quad = \quad \bar{x} \cdot \bar{p} - \bar{x} \cdot \bar{x} \quad \geqslant \quad \bar{x} \cdot \bar{p} \quad \geqslant \quad 0$$

The last inequality comes from Inequatity (3), and the second last from $\bar{x} \cdot \bar{x} \geqslant 0$. ■

Corollary 12.5 (Farkas' lemma, version 2) *Exactly one of the following systems of linear inequalities has a solution, but not both.*

- $M\ \bar{y}^t \leqslant \bar{a}^t$ *and* $\bar{y} \geqslant 0$.
- $M^t\ \bar{x}^t \geqslant 0$, $\bar{a} \cdot \bar{x} < 0$ *and* $\bar{x} \geqslant 0$.

Proof. That it is impossible for both of them to have solution is as before. Now, note that by adding extra variables $\bar{z}$, the constraint $M\bar{y}^t \leqslant \bar{a}^t$ and $\bar{y} \geqslant 0$ is equivalent to $M\bar{y}^t + \bar{z}^t = \bar{a}^t$ and $\bar{y}, \bar{z} \geqslant 0$, which, in turn, can be written as:

$$\big(M \mid I\big)(\bar{y}, \bar{z})^t \ = \ \bar{a}^t \qquad \text{and} \qquad \bar{y}, \bar{z} \geqslant 0$$

where I is the identity matrix. If it has no solution, by Farkas' lemma above, the following system has solution:

$$\begin{pmatrix} M^t \\ I \end{pmatrix} \bar{x}^t \ \geqslant \ 0 \qquad \text{and} \qquad \bar{a} \cdot \bar{x} < 0,$$

where the left hand side is equivalent to $M^t\bar{x}^t \geqslant 0$ and $I\bar{x}^t \geqslant 0$. ∎

Geometric interpretation of Farkas' lemma (version 1). It can be shown that Theorem 12.4 is equivalent to the Theorem 12.6 below.

Recall that a half-space S can be defined by a vector $\bar{c}$, where $S = \{\bar{p} \mid \bar{c} \cdot \bar{p} \geqslant 0\}$. The *cone* defined by vectors $\bar{u}_1, \ldots, \bar{u}_m$ is the following set.

$$\text{cone}(\bar{u}_1, \ldots, \bar{u}_m) \ \ := \ \ \{\lambda_1 \bar{u}_1 + \cdots + \lambda_m \bar{u}_m \mid \lambda_1, \ldots, \lambda_m \geqslant 0\}.$$

Theorem 12.6 (Farkas' lemma, intuitive version) *Let* $\bar{u}_1, \ldots, \bar{u}_m$ *and* $\bar{a}$ *be vectors such that* $\bar{a}$ *lies outside the cone defined by* $\bar{u}_1, \ldots, \bar{u}_m$. *Then, there is a half-space* S *separating* $\bar{u}_1, \ldots, \bar{u}_m$ *from* $\bar{a}$, *i.e.,* $\bar{u}_1, \ldots, \bar{u}_m \in S$, *but* $\bar{a} \notin S$.

The equivalence follows immediately by taking the column vectors of M as $\bar{u}_1^t, \ldots, \bar{u}_m^t$, and a solution for $M^t\bar{x}^t \geqslant 0$ and $\bar{a} \cdot \bar{x} < 0$ as the vector that defines the half-space S separating $\bar{u}_1, \ldots, \bar{u}_m$ from $\bar{a}$.