# The fundamental theorem of linear programming

Michael Tehranchi

June 8, 2017

This note supplements the lecture notes of Optimisation. The statement of the fundamental theorem of linear programming and the proof of weak duality is examinable. The proof of strong duality and the existence of optimisers is not.

## 1 Statement and proof

Given dimensions $m, n \geq 1$, let $A$ be a $m \times n$ matrix, column vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Let

$$p = \sup\{c^\top x : \ x \in \mathbb{R}^n, \ Ax \leq b, \ x \geq 0\}$$

and

$$d = \inf\{b^\top \lambda : \ \lambda \in \mathbb{R}^m, \ A^\top \lambda \geq c, \ \lambda \geq 0\}$$

with the convention that $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

**Theorem.**(Fundamental theorem of linear programming)

> *Weak duality.* $p \leq d$.
>
> *Strong duality.* If either $p > -\infty$ or $d < +\infty$ then $p = d$.
>
> *Existence of optimisers.* If both $p > -\infty$ and $d < \infty$, then there exist vectors $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^m$ such that $p = c^\top x = b^\top y = d$, satisfying
>
> > *primal feasibility.* $Ax \leq b, \ x \geq 0$,
> > *dual feasibility.* $A^\top \lambda \geq c, \ \lambda \geq 0$,
> > *complementary slackness.* $\lambda^\top (Ax - b) = 0 = x^\top (A^\top \lambda - c)$.

Before giving the proof, we pause to make some observations.

*Remark.* Two consequences of weak duality are that

$$d = -\infty \text{ implies } \{x : \ Ax \leq b, x \geq 0\} = \emptyset.$$

and

$$p = +\infty \text{ implies } \{\lambda : \ A^\top \lambda \geq c, \lambda \geq 0\} = \emptyset.$$

*Remark.* Strong duality fails only in the case where both $p = -\infty$ and $d = +\infty$; that is, when both

$$\{x : \ Ax \leq b, x \geq 0\} = \emptyset \text{ and } \{\lambda : \ A^\top \lambda \geq c, \lambda \geq 0\} = \emptyset.$$

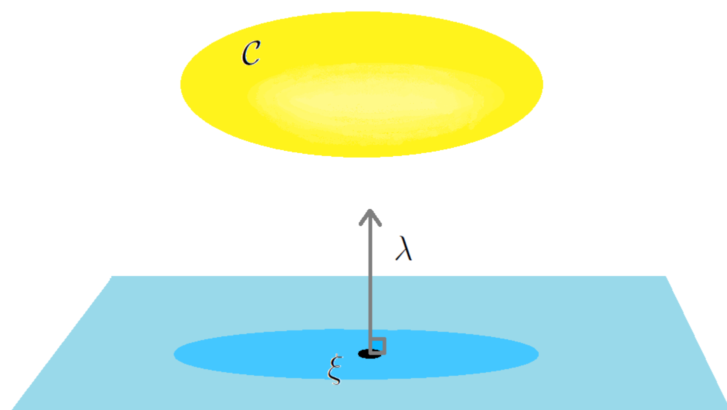
An example of this situation is when $A = 0$ is the $m \times n$ matrix of zeros, $b = (-1, \ldots, -1)^\top$ and $c = (+1, \ldots, +1)^\top$.

Our proof of the fundamental theorem of linear programming will use the following basic results of convex analysis:

**Theorem.** (Separating hyperplane theorem) Given a closed convex set $\mathcal{C} \subseteq \mathbb{R}^k$ and a point $\xi \in \mathbb{R}^k$. If $\xi$ is not in $C$, then there exists a vector $\lambda \in \mathbb{R}^k$ and a constant $\delta > 0$ such that

$$\lambda^\top (x - \xi) \geq \delta \text{ for all } x \in \mathcal{C}.$$

The separating hyperplane theorem is illustrated below. The details of the proof are given in the next section.




The next theorem seems rather obvious.

**Theorem.** (Linear image of a closed orthant is closed) Let $M$ be a $k \times h$ matrix, and

$$\mathcal{C} = \{Mx : \ x \in \mathbb{R}^h, \ x \geq 0\} \subseteq \mathbb{R}^k.$$

Then $\mathcal{C}$ is a closed and convex set.

The convexity of $\mathcal{C}$ follows directly from the definition. In the case where the columns of the matrix $M$ are linearly independent, the closedness of $\mathcal{C}$ is easy to show. However, while the proof in the general case is not extremely difficult, it does require more than a few lines. See section 3.

*Proof of weak duality.* If either $p = -\infty$ or $d = +\infty$, there is nothing to show since the inequality holds trivially. Therefore, we may suppose $p > -\infty$ and $d < +\infty$. Fix $x \geq 0$ such that $Ax \leq b$ and $\lambda \geq 0$ such that $A^\top \lambda \geq c$. Then

$$\begin{aligned} c^\top x &\leq c^\top x + \lambda^\top (b - Ax) \\ &= x^\top (c - A^\top \lambda) + b^\top \lambda \\ &\leq b^\top y. \end{aligned}$$

The conclusion now follows by taking the supremum over such all $x$ and infimum over all such $\lambda$. $\square$

*Proof of strong duality.* As remarked above, strong duality is implied by weak duality in the cases $p = +\infty$ and $d = -\infty$. Hence, we may suppose $-\infty < p \leq d < +\infty$. In particular, we are assuming that both sets $\{x : Ax \leq b, x \geq 0\}$ and $\{\lambda : A^\top \lambda \geq c, \lambda \geq 0\}$ are not empty.

Fix $\varepsilon > 0$. Note that

$$\{x : \ Ax \leq b, c^\top x = p + \varepsilon, x \geq 0\} = \emptyset.$$

That is to say, the point

$$\xi = \begin{pmatrix} b \\ p + \varepsilon \end{pmatrix}$$

is not a member of the closed convex set

$$\mathcal{C} = \left\{ \begin{pmatrix} Ax + z \\ c^\top x \end{pmatrix} : x \geq 0, z \geq 0 \right\}$$

Define an auxiliary function $L$ by

$$L(x, z, \lambda, \mu) = \lambda^\top (Ax + z - b) + \mu(c^\top x - p - \varepsilon)$$

By the separating hyperplane theorem, there is a vector $\binom{\lambda}{\mu} \in \mathbb{R}^{m+1}$ such that

$$L(x, z, \lambda, \mu) > 0$$

for all $x \geq 0$ and $z \geq 0$.

Letting $x \geq 0$ be such that $Ax \leq b$ and setting $z = b - Ax \geq 0$, we have

$$L(x, z, \lambda, \mu) = \mu(c^\top x - p - \varepsilon) > 0$$

Since $c^\top x \leq p < p + \varepsilon$ for all such $x$ by the definition of $p$, we conclude that $\mu < 0$. By homogeneity, we may take $\mu = -1$.

Again, the separating hyperplane theorem implies that

$$L(x, z, \lambda, -1) = z^\top \lambda + x^\top (A^\top \lambda - c) + p + \varepsilon - b^\top \lambda > 0$$

for all $x \geq 0$ and $z \geq 0$. In order for the inequality to hold for all $z \geq 0$, we must have $\lambda \geq 0$. Similarly, for the inequality to hold for all $x \geq 0$, we must have $A^\top \lambda \geq c$.

Note by the definition of $d$ that $b^\top \lambda \geq d$. Setting $x = 0$ and $z = 0$ implies

$$0 < L(0, 0, \lambda, -1) = p + \varepsilon - b^\top \lambda \leq p + \varepsilon - d.$$

and hence

$$d < p + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary and since we have $d \geq p$ by weak duality, we conclude that $d = p$ as desired. $\square$

*Proof of the existence of optimisers.* This time we would like to show that the point

$$\xi = \begin{pmatrix} b \\ p \end{pmatrix}$$

*is* a member of the set

$$\mathcal{C} = \left\{ \begin{pmatrix} Ax + z \\ c^\top x \end{pmatrix} : x \geq 0, z \geq 0 \right\}.$$

By the definition of $p$, there is a sequence $(x_n)_n$ such that $Ax_n \leq b,\ x_n \geq 0$ and $c^\top x_n \to p$. Letting $z_n = b - Ax_n \geq 0$, we see that

$$\begin{pmatrix} Ax_n + z_n \\ c^\top x_n \end{pmatrix} \to \begin{pmatrix} b \\ p \end{pmatrix}.$$

Since $\mathcal{C}$ is closed, there must exist a point $\binom{x}{z} \in \mathbb{R}^{n+m}$ with $x \geq 0, z \geq 0$ such that

$$\begin{pmatrix} Ax + z \\ c^\top x \end{pmatrix} = \begin{pmatrix} b \\ p \end{pmatrix}.$$

meaning $Ax \leq b, x \geq 0$ and $c \top x = p$.

The proof of the existence of $\lambda \geq 0$ such that $A^\top \lambda \geq c$ and $b^\top \lambda = d$ is analogous. $\square$

## 2 Proof of the separating hyperplane theorem

We now fill in some of the details used in the proof of the fundamental theorem of linear programming.

**Theorem.** (Projection onto a closed, convex set) Let $\mathcal{C} \subseteq \mathbb{R}^k$ be closed and convex, and suppose $\xi \in \mathbb{R}^k$ is not in $\mathcal{C}$. Then there exists a point $x^* \in \mathcal{C}$ such that

$$\|x^* - \xi\| \leq \|x - \xi\| \text{ for all } x \in \mathcal{C}.$$

*Proof.* Let $\delta = \inf_{x\in\mathcal{C}} \|x - \xi\|^2$. Let $(x_n)_n$ be a sequence in $\mathcal{C}$ such that $\|x_n - \xi\|^2 \to \delta$. Applying the parallelogram law $\|a+b\|^2 + \|a-b\|^2 = 2\|a\|^2 + 2\|b\|^2$ we have

$$\begin{aligned}\|x_m - x_n\|^2 =& 2\|x_m - \xi\|^2 + 2\|x_n - p\|^2 - 4\left\|\tfrac{1}{2}(x_m + x_n) - \xi\right\|^2 \\ \leq& 2\|x_m - \xi\|^2 + 2\|x_n - \xi\|^2 - 4\delta \to 0\end{aligned}$$

as $m, n \to \infty$, where we have used the convexity of $\mathcal{C}$ to assert that $\frac{1}{2}(x_m + x_n) \in \mathcal{C}$ and hence $\left\|\frac{1}{2}(x_m + x_n) - \xi\right\|^2 \geq \delta$. We have established that the sequence $(x_n)_n$ is Cauchy, and thus converges to some point $x^*$. Since $\mathcal{C}$ is closed, we have $x^* \in C$ as claimed. □

*Proof of the separating hyperplane theorem.* Let $x^*$ be the point in $\mathcal{C}$ closest to $\xi$. Let $\lambda = x^* - \xi$ and $\delta = \|\lambda\|^2$. Fix a point $x \in \mathcal{C}$ and $0 < \theta < 1$, and note that the point $(1-\theta)x^* + \theta x$ is in $\mathcal{C}$ by convexity. Then

$$\begin{aligned}0 =& \|x^* - \xi\|^2 - \delta \\ \leq& \|(1-\theta)x^* + \theta x - \xi\|^2 - \delta \\ =& \|\theta(x - x^*) + \lambda\|^2 - \delta \\ =& \theta^2\|x - x^*\|^2 + 2\theta\lambda^\top(x - x^*).\end{aligned}$$

By first dividing by $\theta$ and then taking the limit as $\theta \downarrow 0$ in the above inequality, we conclude $\lambda^\top(x - x^*) \geq 0$. Hence

$$\lambda^\top(x - \xi) = \lambda^\top(x - x^*) + \lambda^\top(x^* - \xi) \geq \delta$$

as desired. □

# 3 Proof that a linear image of a closed orthant is closed

Consider the set

$$\mathcal{C} = \{Mx : x \in \mathcal{X}\} \subseteq \mathbb{R}^k$$

where $\mathcal{X} \subseteq \mathbb{R}^h$ and $M$ is a $k \times h$ matrix. In this section we will show that $\mathcal{C}$ is closed when $\mathcal{X} = \{x \in \mathbb{R}^h : \ x \geq 0\}$ is the non-negative orthant.

We first consider the case there the columns of $M$ are linearly independent. In this case, the map sending $x$ to $Mx$ is a bijection from $\mathbb{R}^h$ to the range of $M \subseteq \mathbb{R}^k$ with a continuous inverse sending $y$ to $(M^\top M)^{-1}M^\top y$. In particular, the set $\mathcal{C}$ is closed if and only if the set $\mathcal{X}$ is closed. So theorem really is obvious in this case.

*Remark.* It is routine linear algebra to conclude that the square matrix $M^\top M$ is invertible when the columns of $M$ are linearly independent. Indeed, if $M^\top Mx = 0$ for some $x \in \mathbb{R}^h$, we have then $0 = x^\top M^\top Mx = \|Mx\|^2$ and hence $Mx = 0$. But if the columns of $M$ are linearly independent, this implies $x = 0$.

To start to see why things are more subtle in the general case, note that the set $\mathcal{C}$ is not necessarily closed even if $\mathcal{X}$ is closed and convex. For example, let

$$\mathcal{X} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \ |x_1| \leq \sqrt{1 - \frac{1}{x_2}}, \ x_2 \geq 1 \right\} \subseteq \mathbb{R}^2$$

which is closed and convex. Consider the $1 \times 2$ matrix $M = (1 \ \ 0)$. In this case, the linear

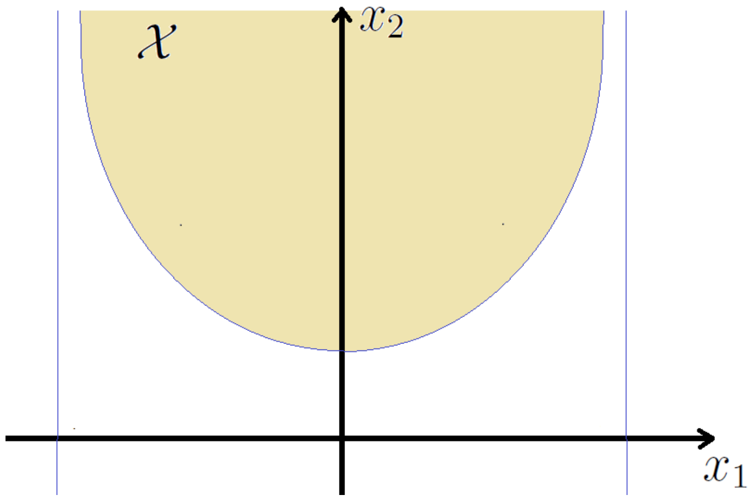



image $\{Mx : x \in \mathcal{X}\} = \{x_1 : -1 < x_1 < 1\}$ is open! We will see that we must exploit the extra structure of the non-negative orthant $\{x : \ x \geq 0\} \subseteq \mathbb{R}^h$ in our proof.

**Theorem.** (Carathéodory's theorem) Let $M$ be a $k \times h$ matrix with columns $m_1, \ldots, m_h$. Fix a point $x \in \mathbb{R}^h$ with $x \geq 0$, and let

$$y = Mx.$$

There is a set

$$B = \{i_1, \ldots, i_r\} \subseteq \{1, \ldots, h\}$$

and a point $x_B \in \mathbb{R}^r$ with $x_B > 0$ such that

$$y = M_B x_B$$

where $M_B$ is a $k \times r$ matrix with linearly independent columns $m_{i_1}, \ldots, m_{i_r}$.

Furthermore, $x_B$ can recovered from $y$ by the formula

$$x_B = (M_B^\top M_B)^{-1} M_B^\top y$$

when $B$ is not empty.

*Proof.* We will construct the basis $B$ by the following algorithm:

(1) If the columns of $M$ are linearly independent, we set simply take $B = \{i : x_i > 0\}$.

(2) Suppose the columns of $M$ are linearly dependent, so there exist constants $\alpha_1, \ldots, \alpha_h$ such that

$$\alpha_1 m_1 + \ldots + \alpha_h m_h = 0,$$

and without loss of generality we may assume at least one of coefficients $\alpha_i$ is strictly positive. Notice that for any real $\lambda$ we have

$$\begin{aligned} y &= x_1 m_1 \ldots + x_h m_h \\ &= x_1 m_1 \ldots + x_h m_h - \lambda(\alpha_1 m_1 + \ldots + \alpha_h m_h) \\ &= (x_1 - \lambda\alpha_1) m_1 + \ldots + (x_h - \lambda\alpha_h) m_h. \end{aligned}$$

Now let

$$\lambda = \min\left\{\frac{x_i}{\alpha_i} : \alpha_i > 0\right\}.$$

For this choice of $\lambda$ we have

$$x_i - \lambda\alpha_i \geq 0 \text{ for all } i$$

and there exists at least one $i^*$ such that

$$x_{i^*} - \lambda\alpha_{i^*} = 0.$$

Notice that we have shown that

$$y = \hat{M}\hat{x}$$

where $\hat{M}$ is the $k \times h - 1$ matrix

$$\hat{M} = \begin{pmatrix} \vdots & & \vdots & \vdots & & \vdots \\ m_1 & \ldots & m_{i^*-1} & m_{i^*+1} & \ldots & m_h \\ \vdots & & \vdots & \vdots & & \vdots \end{pmatrix}$$

and $\hat{x} \in \mathbb{R}^{h-1}$ is given by

$$\hat{x} = \begin{pmatrix} x_1 - \lambda\alpha_1 \\ \vdots \\ x_{i^*-1} - \lambda\alpha_{i^*-1} \\ x_{i^*+1} - \lambda\alpha_{i^*+1} \\ \vdots \\ x_{h+1} - \lambda\alpha_h \end{pmatrix}.$$

Now return to step (1) but now with $\hat{M}$ playing the role of $M$ and $\hat{x}$ that of $x$.

In each iteration of the algorithm, the number of columns of the matrix $M$ is reduced by one. The algorithm terminates when the remaining columns are linearly independent. $\square$

*Proof that the linear image of a closed orthant is closed.* Let $y$ be a limit point of $\mathcal{C}$. That means there is a sequence $(x_n)_n$ in $\mathbb{R}^h$ with $x_n \geq 0$ such that $Mx_n = y_n \to y$. We must show that there is a point $x \in \mathbb{R}^h$ with $x \geq 0$ such that $y = Mx$.

By Carathéodory's theorem there is a sequence of bases $B_n$ of cardinality $r_n$ and vectors $x_{B_n,n}$ of dimension $r_n$ with $x_{B_n,n} > 0$ such that

$$M_{B_n} x_{B_n,n} = y_n.$$

Since there are only a finite number of such bases, one basis $B$ must appear an infinite number of times in the sequence $(B_n)_n$. Therefore, we can pass to a subsequence where all the bases $B_n = B$ are equal, so that

$$M_B x_{B,n} = y_n.$$

Hence we have the convergence

$$x_{B,n} \to \xi$$

where

$$\xi = (M_B^\top M_B)^{-1} M_B^\top y.$$

Since $x_{B,n} > 0$ for each $n$, we have $\xi \geq 0$. Create a new vector $x \in \mathbb{R}^h$ by

$$x_i = \begin{cases} \xi_s & \text{if } i = i_s \in B \\ 0 & \text{otherwise} \end{cases}$$

Note that

$$M x_n \to M x \text{ and } x \geq 0.$$

This shows that the set $\mathcal{C}$ is closed. $\square$