Professor: *S. Ortiz-Latorre*

Review of Basic Results on Linear Programming

---

1 Linear Programming

Linear programming (**LP**) is about solving optimization problems where the objective function and the constraints are linear. The optimization problem can be finding a maximum or a minimum and the constraints can be given by equalities and/or inequalities. In what follows most inequalities will be vector inequalities, that is, the inequalities hold componentwise. All (**LP**) problems can be written in the following standard form

Primal Problem (P)

$$\begin{aligned} \max J(x) &= \max c^T x \\ \text{subject to } Ax &\leq b, \\ x &\geq 0, \end{aligned}$$

where $x \in \mathbb{R}^n, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

We will use the following notation:

- **Objective function**: It is the function J to be optimized. In this case the linear function $J(x) = c^T x$.
- **Feasible set/solution**: $x \in \mathbb{R}^n$ is a feasible solution if satisfies the constraints, i.e., $Ax \leq b, x \geq 0$. The feasible set F_P is the convex set defined by all feasible solutions, i.e,
$$F_P := \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}.$$

- **Optimal solution**: $\hat{x} \in F_P$ such that

$$J(\hat{x}) = c^T \hat{x} = \max\left\{c^T x : Ax \leq b, x \geq 0\right\}.$$

- **Optimal value**: It is the value (finite) of the objective function at an optimal solution , i.e., $J(\hat{x})$.

There are three different cases regarding the problem (**P**):

1. There exists an optimal solution (or many) and only one optimal value.
2. $F_P = \emptyset$, then the optimal value is set to $-\infty$. We say that the problem is not feasible.
3. The problem is unbounded. There exists a sequence $\{x_k\}_{k\geq 1} \subseteq F_P$ such that $J(x_k) \to_{k\to\infty} \infty$.

2 Reduction to the standard form

We have the following rules:

- "min" $\longrightarrow$"max": $\min J(x) = -\max J(-x)$.
- "$\geq$" $\longrightarrow$"$\leq$": Multiply the equation by -1.
- "$=$" $\longrightarrow$"$\leq$": Write as two inequalities using "$\leq$" and "$\geq$". Then apply the previous point to the inequality with "$\geq$".
- "Free variables"$\longrightarrow$"Restricted variables": Write $x = x^+ - x^-$,where $x^+ = \max(0, x) \geq 0$ and $x^- = -\min(0, x) \geq 0$ and rewrite the other constraints and the objective function in terms of x^+ and x^-.

A general (iterative) method to solve LP problems is the simplex method (Dantzig, 1947). In the simplex method the constraints must be in equality form. We can go from "$\leq$" to "$=$" by introducing the so called slack variables $w := b - Ax$, then the problem (**P**) can be written as

$$\begin{aligned} &\max J(x) \\ \text{subject to } &w = b - Ax, \\ &w \geq 0, \\ &x \geq 0. \end{aligned}$$

Example 1. Consider the (**LP**) problem

$$\boxed{\begin{aligned} &\max J(x) = 3x_1 + 2x_2 \\ \text{subject to } &-x_1 + 3x_2 \leq 12, \\ &x_1 + x_2 \leq 10, \\ &2x_1 - x_2 \leq 10 \\ &x_1 \geq 0, \quad x_2 \geq 0 \end{aligned}}$$

Here, $c = (3,2)^T$, $b = (12,8,10)^T$, $A = \begin{pmatrix} -1 & 3 \\ 1 & 1 \\ 2 & -1 \end{pmatrix}$. Moreover, one can prove that the optimal solution is $\hat{x} = (6,2)^T$and the optimal value is$\hat{J} = J(\hat{x}) = 22$. Any point in F_P is a lower bound for $\hat{J}$. To find an upper bound we can try conic linear combinations (i.e., linear combinations with non-negative scalars) of the constraints. In this way the inequalities are not reversed. For example, consider

$$\begin{aligned} 2 \cdot (-x_1 + 3x_2) &\leq 2 \cdot 12, \\ 1 \cdot (x_1 + x_2) &\leq 1 \cdot 8, \\ 3 \cdot (2x_1 - x_2) &\leq 3 \cdot 1, \end{aligned}$$

which added give

$$5x_1 + 4x_2 \leq 62.$$

Since

$$J(x) = 3x_1 + 2x_2 \leq 5x_1 + 4x_2 \leq 62,$$

we obtain the upper bound $J(\hat{x}) \leq 62$. We can use this procedure to get the best upper bound. Take $y_1, y_2, y_3 \geq 0$ and compute

$$\begin{aligned} y_1 \cdot (-x_1 + 3x_2) &\leq y_1 \cdot 12, \\ y_2 \cdot (x_1 + x_2) &\leq y_2 \cdot 8, \\ y_3 \cdot (2x_1 - x_2) &\leq y_3 \cdot 1, \end{aligned}$$

which added yield

$$(-y_1 + y_2 + y_3)\, x_1 + (3y_1 + y_2 - y_3)\, x_2 \leq 12y_1 + 8y_2 + y_3.$$

Since $J(x) = 3x_1 + 2x_2$, we take y_1, y_2 and y_3 such that

$$-y_1 + y_2 + 2y_3 \geq 3 \qquad \text{and} \qquad 3y_1 + y_2 - y_3 \geq 2.$$

Then,

$$J(x) = 3x_1 + 2x_2 \leq (-y_1 + y_2 + y_3)\, x_1 + (3y_1 + y_2 - y_3)\, x_2 \leq 12y_1 + 8y_2 + y_3.$$

Finally, to get the best upper bound we can solve the followin (**LP**) problem

$$\boxed{\begin{aligned} \min J(x) &= 12y_1 + 8y_2 + y_3 \\ \text{subject to } -y_1 + y_2 + 2y_3 &\geq 3, \\ 3y_1 + y_2 - y_3 &\geq 2, \\ y_1 \geq 0, \quad y_2 \geq 0, \quad y_3 &\geq 0. \end{aligned}}$$

3 Duality

The previous example justifies the introduction of the dual problem of a (**LP**) .

Definition 2. Given the (**LP**) problem (**P**) we define its dual (**D**) as

Dual Problem (**D**)

$$\begin{aligned} \min J(y) &= \min b^T y \\ \text{subject to } A^T y &\geq c, \\ y &\geq 0, \end{aligned}$$

where $y \in \mathbb{R}^m, b \in \mathbb{R}^m, A^T \in \mathbb{R}^{n\times m}$ and $c \in \mathbb{R}^n$.

Remark 3. We have that

- The dual problem of a (**LP**) problem is also a (**LP**) problem.
- The dual problem provides upper bounds for the optimal value of the primal problem.
- (**D**) is sometimes easier to solve than (**P**) .
- Good implementations of the simplex algorithm solve simultaneously (**P**) and (**D**).

Lemma 4. *The dual of* (**D**) *is* (**P**).

Proof. We can write

$$\min \left\{ b^T y : A^T y \geq c, y \geq 0 \right\} = -\max \left\{ (-b)^T y : -A^T y \leq -c \right\}.$$

The problem on the right hand side of the previous equation is in standard form, so we can take its dual to get

$$-\min \left\{ (-c)^T x : \left(A^T\right)^T x \geq b, x \geq 0 \right\},$$

which written in standard form is

$$\max = \left\{ c^T x : Ax \leq b, x \geq 0 \right\}.$$

□

Sometimes it is convenient to find the dual of a (**LP**) problem without finding first its standard form. We assume that we have a (**LP**) problem in the form of a generalised primal problem ($\mathbf{P_g}$) (this means that we have a primal problem with some constraints that are equalities and only R variables are restricted), i.e.,

Generalized Primal Problem ($\mathbf{P_g}$)

$$\max J(x) = \max c^T x$$

$$\text{subject to } \sum_{j=1}^{n} a_{ij} x_j \le b_i, \qquad i \in I,$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \qquad i \in E,$$

$$x_j \ge 0, \qquad j \in R$$

where $x \in \mathbb{R}^n, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, $R \subseteq \{1, \ldots, n\}$, $I, E \subseteq \{1, \ldots, m\}$ and

$$I \cap E = \emptyset, \qquad I \cup E = \{1, \ldots, m\}.$$

Using the following primal-dual correspondence

	In ($\mathbf{P_g}$)	In ($\mathbf{D_g}$)	
I	Inequality constraints	Restricted variables	R
E	Equality constraints	Free variables	F
R	Restricted variables	Inequality constraints	I
F	Free variables	Equality constraints	E

we can find its associated generalised dual problem ($\mathbf{D_g}$) (this means a dual problem with some constraints that are equalities and only some variables which are restricted), i.e,

Generalized Dual Problem ($\mathbf{D_g}$)

$$\min J(y) = \min b^T y$$

$$\text{subject to } \sum_{i=1}^{m} a_{ij} y_i \ge c_j, \qquad j \in R,$$

$$\sum_{i=1}^{m} a_{ij} y_i = c_j, \qquad i \in F,$$

$$y_i \ge 0, \qquad i \in I$$

where $x \in \mathbb{R}^n, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$,

where R,$F \subseteq \{1, \ldots, n\}$ are such that
and

$$R \cap F = \emptyset, \qquad R \cup F = \{1, \ldots, n\}.$$

Theorem 5 (Duality)**.** *Let $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.*

1. (Weak duality) If x is feasible for (**P**) and y is feasible for (**D**), then

$$c^T x \leq \left(A^T y\right)^T x = y^T A x = \left((xA)^T y\right)^T = (xA)^T y \leq b^T y.$$

   Moreover:

   (a) If (**P**) is unbounded $\Longrightarrow$ (**D**) is not feasible.

   (b) If (**D**) is unbounded $\Longrightarrow$ (**P**) is not feasible.

   (c) If $c^T \hat{x} = b^T \hat{y}$ with $\hat{x}$ feasible for (**P**) and $\hat{y}$ feasible for (**D**), then $\hat{x}$ must solve (**P**) and $\hat{y}$ must solve (**D**).

2. (Strong duality) If either (**P**) or (**D**) has a finite optimal value, then so does the other, the optimal values coincide, and optimal solutions for both (**P**) and (**D**) exist.

4 Convex analysis

Definition 6. A set $A \subset \mathbb{R}^n$ is convex if one has that $\lambda x + (1 - \lambda) y \in A$, for all $x, y \in A$ and $\lambda \in (0, 1)$.

Definition 7. An hyperplane with normal vector $a \neq 0 \in \mathbb{R}^n$ and level α is the set

$$H_{a,\alpha} = \left\{x \in \mathbb{R}^n : a^T x = \alpha\right\}.$$

Every hyperplane $H_{a,\alpha}$ is the intersection of the halfspaces

$$\begin{aligned} H^-_{a,\alpha} &= \left\{x \in \mathbb{R}^n : a^T x \leq \alpha\right\}, \\ H^+_{a,\alpha} &= \left\{x \in \mathbb{R}^n : a^T x \geq \alpha\right\}. \end{aligned}$$

Definition 8. Let S and T be two sets in $\mathbb{R}^n$. We say that $H_{a,\alpha}$ strongly separates S and T if there exists $\varepsilon > 0$ such that $S \subseteq H^-_{a,\alpha-\varepsilon}$ and $T \subseteq H^+_{a,\alpha+\varepsilon}$ or viceversa.

Theorem 9 (Separating Hyperplane Theorem)**.** *Let S and T be two disjoint, non-empty, closed, convex sets in $\mathbb{R}^n$ and one of them is compact. Then, there exists an hyperplane $H_{a,\alpha}$ that strongly separates S and T.*

Corollary 10. *Let S be a non-empty, closed, convex set in $\mathbb{R}^n$ and such that $0 \notin S$. Then, there exist $a \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}_{++}$ such that*

$$a^T x \geq \alpha > 0, \qquad x \in S.$$

Corollary 11. *Let V be a linear subspace of $\mathbb{R}^n$ and let K be a non-empty, compact, convex set in $\mathbb{R}^n$,such that $K \cap V = \emptyset$. Then, there exists $a \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}_{++}$ such that*

$$\begin{aligned} a^T x &= 0, \qquad x \in V, \\ a^T y &\geq \alpha > 0, \qquad y \in K. \end{aligned}$$

5 Linear algebra

Definition 12. Given $A \in \mathbb{R}^{m \times n}$, we can consider the following fundamental linear subspaces:

- $\mathrm{col}\,(A)$: The *column space* of A, it contains all linear combinations of the columns of A.
- $\mathrm{null}\,(A)$: The *null space* of A, it contains all solutions to the system $Ax = 0$.
- $\mathrm{col}\left(A^T\right)$: The *row space* of A, it contains all linear combinations of the rows of A, (or columns of A^T).
- $\mathrm{null}\left(A^T\right)$: The *left null space* of A^T, it contains all solutions to the system $A^T y = 0$.

Definition 13. The *rank* of A is the dimension of $\mathrm{col}\,(A)$ or $\mathrm{col}\left(A^T\right)$, i.e.,

$$\mathrm{rank}\,(A) = \dim\left(\mathrm{col}\,(A)\right) = \dim\left(\mathrm{col}\left(A^T\right)\right).$$

Definition 14. Let $S \subseteq \mathbb{R}^n$. We define $S^\perp$, the *orthogonal complement* of S, as the set of vectors in $\mathbb{R}^n$ which are orthogonal to S, that is,

$$S^\perp := \left\{x \in \mathbb{R}^n : x^T y = 0, \quad y \in S\right\}.$$

It is easy to check that $S^\perp$ is a linear subspace, regardless of S being a subspace or not. If S is a linear subspace, then $S \cap S^\perp = \{0\}$.

Proposition 15 (Orthogonal projection)**.** *Let $v \in \mathbb{R}^n$ and let $S \subseteq \mathbb{R}^n$ be a linear subspace. Then there exist unique $x \in S$ and $y \in S^\perp$ such that*

$$v = x + y.$$

We write $\mathbb{R}^n = S \oplus S^\perp$, and we say that $\mathbb{R}^n$ is the direct sum of S and $S^\perp$.

Theorem 16 (Fundamental theorem of linear algebra)**.** *Let $A \in \mathbb{R}^{m \times n}$. Then $\mathrm{col}\,(A)$ is orthogonal to $\mathrm{null}\left(A^T\right)$, and*

$$\mathbb{R}^m = \mathrm{col}\,(A) \oplus \mathrm{null}\left(A^T\right).$$

Moreover, $\mathrm{col}\left(A^T\right)$ is orthogonal to $\mathrm{null}\,(A)$ and

$$\mathbb{R}^n = \mathrm{col}\left(A^T\right) \oplus \mathrm{null}\,(A).$$

Proof. Follows from Proposition 15 and the following equalities

$$\begin{aligned}
\mathrm{col}\,(A)^\perp &= \left\{y \in \mathbb{R}^m : y^T A x = 0, \quad x \in \mathbb{R}^n\right\} \\
&= \left\{y \in \mathbb{R}^m : x^T\left(A^T y\right) = 0, \quad x \in \mathbb{R}^n\right\} \\
&= \left\{y \in \mathbb{R}^m : A^T y = 0\right\} \\
&= \mathrm{null}\left(A^T\right).
\end{aligned}$$

□

Proposition 17 (Fredholm's alternative)**.** *For every matrix $A \in \mathbb{R}^{m\times n}$ and vector $b \in \mathbb{R}^m$, **exactly one** of the following statements is true:*

1. $Ax = b$ has a solution $x \in \mathbb{R}^n$.
2. There exists $0 \neq y \in \mathbb{R}^m$ such that $A^T y = 0$ and $y^T b \neq 0$.

Proof. Suppose $Ax = b$ has a solution. This is equivalent to $b \in \mathrm{col}\,(A)$. Let $y = y_c + y_n \in \mathbb{R}^m$, $y_c \in \mathrm{col}\,(A)$,$y_n \in \mathrm{null}\left(A^T\right)$. Note that

$$A^T y = A^T y_c + A^T y_n = A^T y_c$$

and

$$y^T b = y_c^T b + y_n^T b = y_c^T b.$$

But then, if $A^T y = 0$ we have that

$$A^T y_c = 0 \Leftrightarrow y_c = 0 \Leftrightarrow y_c^T = 0 \Longrightarrow y_c^T b = 0,$$

which also implies that $y^T b = 0$. Therefore, 2. is not true.

Suppose that $Ax = b$ does not have a solution. Note that, in this case, $b \neq 0 \in \mathbb{R}^m$, because for $b = 0$ we always have the solution $x = 0$. Moreover, this is equivalent to $b \notin \mathrm{col}\,(A)$ (i.e., $b \in \mathrm{null}\left(A^T\right)$). Then, $A^T b = 0$ and $b^T b = \|b\|^2 \neq 0$. Hence, we can take $y = b$ and we have that 2. is true. □