Lecture 7. Linear Programming

7.1 Linear program and standard form

A linear program may have both equality and inequality constraints. It is easy to rewrite equality constraints by inequalities. Can we rewrite inequality constraints by equalities? In fact, the *standard form* of linear programming only allows equality constraints and a special type of inequalities.

Definition (*Standard form*)

We say that a linear program is in *standard form* if the following are all true:

1. Non-negativity constraints for all variables;
2. All remaining constraints are expressed as equality constraints;
3. The right hand side vector, $\boldsymbol{b}$, is non-negative.

Namely, the *standard form* can be expressed as follows:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s. t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \quad \mathbf{A} \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m_{\geq 0}\,, \\ & \forall\, i = 1, 2, \dots, n,\ x_i \geq 0\,. \end{aligned}$$

Now the question is how to convert a linear program into the standard form? Clearly the third requirement is trivial. We only need to consider the first two requirement.
First, for the second requirement, consider a simple example $x_1 + 2x_2 + x_3 - x_4 \leq 5$. We can add a *slack variable* s_1: $x_1 + 2x_2 + x_3 - x_4 + s_1 = 5$ together with $s_1 \geq 0$. Note that it is equivalent to the original inequality. Suppose we have another example $2x_1 + x_2 - x_3 \geq 1$. Then multiply the inequality by -1 and add a slack variable s_2.
Next we consider the first requirement. If a variable has non-positive constraints, such as $x_4 \leq 0$, we can easily let $y_4 = -x_4$ and substitute $-y_4$ for x_4 in the linear program. If some variable, for example x_3, is unconstrained in sign, we can replace it by $x_3^+ - x_3^-$ and require x_3^+ and x_3^- to be non-negative.

Example

Consider the following linear program:

$$\begin{aligned} \min_{x_1, x_2 \in \mathbb{R}} \quad & -2x_1 - 3x_2 \\ \text{s. t.} \quad & x_1 \le 100 \\ & x_2 \le 200 \\ & x_1 + x_2 \le 160\,. \end{aligned}$$

By add slack variable and splitting free variables, it can be converted into the following standard form:

$$\begin{aligned} \min \quad & -2(x_1^+ - x_1^-) - 3(x_2^+ - x_2^-) \\ \text{s. t.} \quad & x_1^+ - x_1^- + s_1 = 100 \\ & x_2^+ - x_2^- + s_2 = 200 \\ & x_1^+ - x_1^- + x_2^+ - x_2^- + s_3 = 160 \\ & x_1^+, x_1^-, x_2^+, x_2^-, s_1, s_2, s_3 \ge 0\,. \end{aligned}$$

The standard form is useful in algorithm design and analysis, especially in the algorithms based on primal dual method.

7.2 Solving linear programming

We now use the form

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{c}^\mathsf{T} \boldsymbol{x} \\ \text{s. t.} \quad & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b},\ \mathbf{A} \in \mathbb{R}^{m \times n},\ \boldsymbol{b} \in \mathbb{R}^m \\ & \forall i,\ x_i \ge 0 \end{aligned}$$

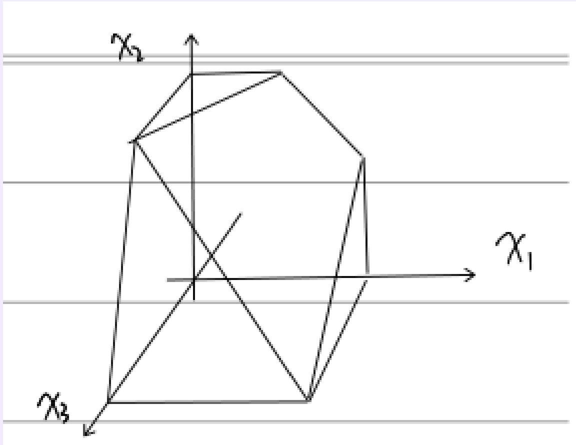
Note that the feasible set is a polyhedron since it is the intersection of some halfspaces.

Example

Consider the following linear program:

$$\begin{aligned}\min_{x\in\mathbb{R}^n} \quad & -x_1 + 6x_2 - 13x_3 \\ \text{s. t.} \quad & x_1 + x_2 + x_3 \leq 400 \\ & x_2 + 3x_3 \leq 600 \\ & x_1 \leq 200 \\ & x_2 \leq 300 \\ & x_1, x_2, x_3 \geq 0\,.\end{aligned}$$

We can sketch its feasible set as follows:



When a linear program achieve its optimal value f^*, intuitively the hyperplane $\{x \mid c^\top x - f^* = 0\}$ passes through a *vertex* of the feasible set. We may use this observation to solve linear programs geometrically.

Question

Is this observation always true?

Fundamental Theorem of Linear Programming

Theorem (*Fundamental theorem of linear programming*)

Suppose a linear program has an optimal solution. Then there exists an optimal solution at a *vertex* (*extreme point*).

First, we define what a vertex and an extreme point are.

Definition (*Vertex*)

A point $x \in \mathbb{R}^n$ is a vertex, if at least n linearly independent constraints are tight at x.

Definition (*Extreme point*)

A point $x \in \mathbb{R}^n$ is called an extreme point of a convex set C, if there does not exist $u \neq v \in C$ and $\theta \in (0,1)$ such that $x = \theta u + \bar{\theta} v$. In other words, x cannot be expressed by a convex combination of other points in C.

An important fact is that, these two types of points are equivalent for polyhedra.

Proposition

For any polyhedron $P = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$, $\boldsymbol{x} \in P$ is a extreme point if and only if $\boldsymbol{x}$ is a vertex.

Proof

Suppose $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1^\top \\ \vdots \\ \boldsymbol{a}_m^\top \end{pmatrix} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} = \begin{pmatrix} b_1^\top \\ \vdots \\ b_m^\top \end{pmatrix} \in \mathbb{R}^m$.

- **Sufficiency:** Suppose $\boldsymbol{x} \in P$ is a vertex, then there exists n indices $i_1, i_2, \ldots, i_n$ such that $\tilde{\boldsymbol{A}} = \begin{pmatrix} \boldsymbol{a}_{i_1}^\top \\ \vdots \\ \boldsymbol{a}_{i_n}^\top \end{pmatrix} \in \mathbb{R}^{n \times n}$ and $\tilde{\boldsymbol{b}} = \begin{pmatrix} b_{i_1}^\top \\ \vdots \\ \boldsymbol{b}_{i_n}^\top \end{pmatrix}$ satisfying $\tilde{\boldsymbol{A}}\boldsymbol{x} = \tilde{\boldsymbol{b}}$. The n constraints are linearly independent, so $\tilde{\boldsymbol{A}}$ is invertible. Assume x is not an extreme point, that is, there exists $\boldsymbol{u} \neq \boldsymbol{v}$ and $\theta \in (0,1)$ such that $\boldsymbol{x} = \theta\boldsymbol{u} + \bar{\theta}\boldsymbol{v}$. After substituting we have $\theta(\tilde{\boldsymbol{A}}\boldsymbol{u}) + \bar{\theta}(\tilde{\boldsymbol{A}}\boldsymbol{v}) = \tilde{\boldsymbol{b}}$. Note that $\boldsymbol{u}, \boldsymbol{v} \in P$. So it holds that $\tilde{\boldsymbol{A}}\boldsymbol{u} \leq \tilde{\boldsymbol{b}}$ and $\tilde{\boldsymbol{A}}\boldsymbol{v} \leq \tilde{\boldsymbol{b}}$. Thus, we conclude that $\tilde{\boldsymbol{A}}\boldsymbol{u} = \tilde{\boldsymbol{A}}\boldsymbol{v} = \boldsymbol{b}$. Since $\tilde{\mathbf{A}}$ is invertible, it yields $\boldsymbol{u} = \boldsymbol{v}$, which leads to a contradiction.
- **Necessity:** Assume $\boldsymbol{x}$ is an extreme point but not a vertex in P. Let $I = \{i \mid \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\}$. Since $\boldsymbol{x}$ is not a vertex, there does not exist linearly independent n constraints that are tight at $\boldsymbol{x}$. Hence there exists $\boldsymbol{d} \neq \mathbf{0} \in \mathbb{R}^n$, such that $\boldsymbol{a}_i^\top \boldsymbol{d} = 0$ for all $i \in I$.

Let $\boldsymbol{u} = \boldsymbol{x} + \varepsilon \boldsymbol{d}, \boldsymbol{v} = \boldsymbol{x} - \varepsilon \boldsymbol{d}$. We argue that $\boldsymbol{u}, \boldsymbol{v} \in P$ for some sufficiently small $\varepsilon > 0$ as follows.

- $\forall\, i \in I$, note that $\boldsymbol{a}_i^\top \boldsymbol{u} = \boldsymbol{a}_i^\top \boldsymbol{v} = \boldsymbol{a}_i^\top \boldsymbol{x} = b_i$ (since $\boldsymbol{a}_i^\top \boldsymbol{d} = 0$).
- $\forall\, j \notin I$, we have $\boldsymbol{a}_j^\top \boldsymbol{x} < b_j$. Then there exists $\varepsilon_j > 0$ such that $\boldsymbol{a}_j^\top \boldsymbol{u} \leq b_j$ and $\boldsymbol{a}_j^\top \boldsymbol{v} \leq b_j$.

Taking $\varepsilon = \min_{j \notin I} \varepsilon_j > 0$, we have $\boldsymbol{u} \neq \boldsymbol{v} \in P$. Hence $\boldsymbol{x}$ is an extreme point, which leads to a contradiction.

However, a polyhedron may not contain any vertex.

Proposition

A polyhedron $P = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ has extreme points if and only if P does not contain a line, and $P \neq \emptyset$.

Proof

- **Necessity:** Assume there exists a line $\ell \subseteq P$ and $\ell = \{\boldsymbol{x} = \boldsymbol{u} + t\boldsymbol{v} \mid t \in \mathbb{R}\}$ for some $\boldsymbol{u} \neq \boldsymbol{v}$.
  1. If $\boldsymbol{x}$ is on the line ℓ, it's obvious that $\boldsymbol{x}$ cannot be an extreme point.
  2. If $\boldsymbol{x}$ is not on the line ℓ, we claim that $\boldsymbol{x} + \boldsymbol{v}$ and $\boldsymbol{x} - \boldsymbol{v}$ are also in P. For any $\theta \in (0, 1)$ and $t \in \mathbb{R}$, $\theta \boldsymbol{x} + \bar{\theta}(\boldsymbol{u} + t\boldsymbol{v}) \in P$ by convexity. Let $t = \frac{1}{\bar{\theta}}$. It follows that $\theta \boldsymbol{x} + \bar{\theta}\boldsymbol{u} + \boldsymbol{v} \in P$. Thus, since P is closed,

     $$\lim_{\theta \to 1} (\theta \boldsymbol{x} + \bar{\theta}\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{x} + \boldsymbol{v} \in P.$$

     Similarly, we have $\boldsymbol{x} - \boldsymbol{v} \in P$. Hence, $\boldsymbol{x}$ cannot be an extreme point.
- **Sufficiency:** If P contains no extreme point, then there exists $\boldsymbol{d} \neq \boldsymbol{0}$ such that $\boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}$. Thus, for all $\boldsymbol{x} \in P$, we have $\boldsymbol{A}(\boldsymbol{x} + t\boldsymbol{d}) = \boldsymbol{A} + t\boldsymbol{A}\boldsymbol{d} = \boldsymbol{A}\boldsymbol{d} \leq \boldsymbol{b}$ for all $t \in \mathbb{R}$, which gives that $\{\boldsymbol{y} = \boldsymbol{x} + t\boldsymbol{d} \mid t \in \mathbb{R}\} \subseteq P$.

We now prove the fundamental theorem of linear programming.
Let P be the feasible set of a linear programming

$$
\begin{aligned}
\min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s. t.} \quad & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b},\ \mathbf{A}\in\mathbb{R}^{m\times n},\ \boldsymbol{b}\in\mathbb{R}^m \\
& \forall\, i,\ x_i \ge 0
\end{aligned}
$$

and Q be the set of optimal solution. Assume $Q \neq \emptyset$. Since $\boldsymbol{x} \ge 0$, there is no line in P and thus no line in Q. Note that Q is also a polyhedron since $Q = P \cap \{\boldsymbol{x} \mid \boldsymbol{c}^\top \boldsymbol{x} = f^*\}$ where f^* is the optimal value of the objective function. By the above proposition, we know that Q has an extreme point $\boldsymbol{x}^*$. Now it suffices to show that $\boldsymbol{x}^*$ is also an extreme point in P.
Suppose $\boldsymbol{x}^*$ is not an extreme point in P, then there exists $\boldsymbol{u}, \boldsymbol{v} \in P$ and $\theta \in (0,1)$ such that $\boldsymbol{x}^* = \theta\boldsymbol{u} + \bar{\theta}\boldsymbol{v}$. We have

$$
\boldsymbol{c}^\top \boldsymbol{x}^* = \boldsymbol{c}^\top(\theta\boldsymbol{u} + \bar{\theta}\boldsymbol{v}) = \theta(\boldsymbol{c}^\top\boldsymbol{u}) + \bar{\theta}(\boldsymbol{c}^\top\boldsymbol{v}) \ge \theta(\boldsymbol{c}^\top\boldsymbol{x}^*) + \bar{\theta}(\boldsymbol{c}^\top\boldsymbol{x}^*) = \boldsymbol{c}^\top\boldsymbol{x}
$$

since $\boldsymbol{x}^*$ is an optimal solution. Thus, $\boldsymbol{c}^\top\boldsymbol{u} = \boldsymbol{c}^\top\boldsymbol{v} = \boldsymbol{c}^\top\boldsymbol{x}^*$. It implies that $\boldsymbol{u}, \boldsymbol{v} \in Q$, which leads to a contradiction.

The fundamental theorem of linear programming gives us an algorithm to solve linear programs by enumerating all ($\le \binom{m}{n}$) vertices of the feasible set.

However, consider the n-dimensional cube $[0,1]^n$. Only $2n$ constraints produce a polyhedron of 2^n vertices.

7.3 Simplex method

We now introduce a (usually) efficient algorithm: the simplex method. We remark here that the simplex method is not a polynomial-time algorithm. However, it runs fast except for some artificially designed cases.

The key idea is that when we find a vertex of the feasible set, move from the current vertex to a "better" neighbor, where two vertices are neighbors if they share $n-1$ tight constraints.
Assume $\boldsymbol{x} = \mathbf{0}$ is a feasible set. Then we can start from the origin point $\boldsymbol{x} = \mathbf{0}$. It is clear that there are n neighbors of the origin, and each of them has $n-1$ zero coordinates.
How can we know whether a neighbor is "better"? Note that our goal is to compute $\min \sum c_i x_i$. If $x_i > 0$, the objective function is "better" as long as $c_i < 0$. To this end, we can choose i such that $c_i < 0$ and increase x_i to $x_i = r$ until some

constraint $\sum_k a_{jk}x_k \le b_j$ is tight. Now there are n constraints tight:

$$\begin{cases} x_1 = \ldots = x_{i-1} = x_{i+1} = \ldots = x_n = 0\,, \\ a_{j1}x_1 + \ldots + a_{jn}x_n \le b_j\,. \end{cases}$$

Then we shift coordinates so that $(0, \ldots, 0, x_i = r, 0, \ldots, 0)$ becomes the origin point. It suffices to let $y_1 = x_1, \ldots, y_{i-1} = x_{i-1}$, $y_i = b_j - \sum_k a_{jk}x_k$, $y_{i+1} = x_{i+1}, \ldots, y_n = x_n$.

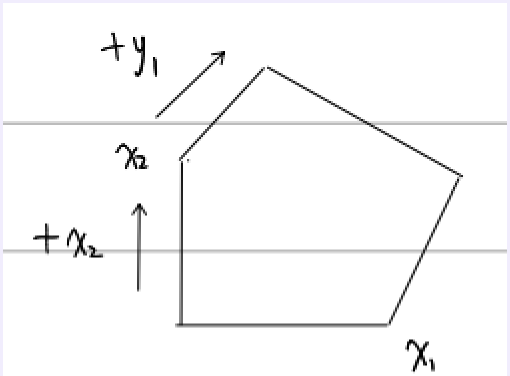
Definition (*Neighbor*)

Two vertices are neighbors if they share $n-1$ tight constraints.

Example

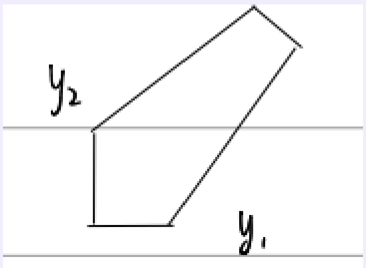
Consider the following linear program:

$$\begin{aligned} \min_{x_1, x_2 \in \mathbb{R}} \quad & -2x_1 - 5x_2 \\ \text{subject to} \quad & 2x_1 - x_2 \le 4 \\ & x_1 + 2x_2 \le 9 \\ & x_2 - x_1 \le 3 \\ & x_1 \ge 0 \\ & x_2 \ge 0 \end{aligned}$$

- Suppose we start from $(0, 0)$. Since the coefficient for x_2 is negative, we could increase x_2. When $x_1 = 0$, the constraints are $x_2 \ge -4$, $x_2 \le 9/2$, and $x_2 \le 3$. To make the objective function as small as possible, we should increase the value of x_2 until some constraint is tight. Then we increase x_2 to 3 and arrive at $(0, 3)$. Now the constraint $x_2 - x_1 \le 3$ is tight. 

- At $(0, 3)$, we create a new coordinate system. The point (x_1, x_2) in the original coordinate system becomes $(x_1, 3 - (x_2 - x_1))$.
- Let $y_1 = x_1$ and $y_2 = 3 + x_1 - x_2$. We then rewrite the linear programming as follows

$$
\begin{aligned}
\min_{y_1, y_2} \quad & -15 - 7y_1 + 5y_2 \\
\text{subject to} \quad & y_1 + y_2 \leq 7 \\
& 3y_1 - 2y_2 \leq 3 \\
& y_2 \geq 0 \\
& y_1 \geq 0 \\
& y_2 - y_1 \leq 3
\end{aligned}
$$

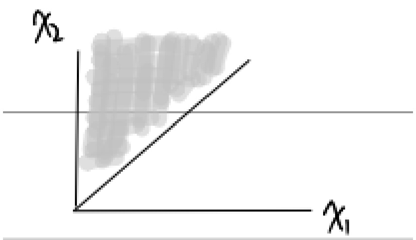


- Now we repeat the above process. We increase y_1 and let $z_1 = 3 - (3y_1 - 2y_2)$, $z_2 = y_2$. The objective function becomes $-22 + \frac{7}{3}z_1 + \frac{1}{3}z_2$.
- Since the coefficients for both z_1 and z_2 are positive, we know that $(z_1, z_2) = (0, 0)$ is an optimal solution. Substituting $\boldsymbol{y}$ and $\boldsymbol{x}$ we have $(y_1, y_2) = (1, 0)$ and $(x_1, x_2) = (1, 4)$.

As shown above, the first step is to choose a variable whose coefficient is negative, and then increase it. What should we do if we have multiple variables who have negative coefficients?

In fact, the following example shows we may encounter some tricky problems if we choose a wrong variable. Consider the linear program with the same objective function as above, but the constraints are

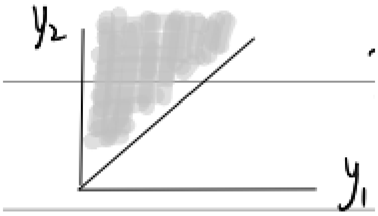
$$
\begin{aligned}
x_1 - x_2 \leq 0\,, \\
x_1, x_2 \geq 0\,.
\end{aligned}
$$



- Suppose in the first step, we choose to increase x_1. Then the tight constraint should be $x_1 - x_2 = 0$ and x_2 cannot increase any more.

- Let $y_1 = x_2 - x_1$ and $y_2 = x_2$. The new constraints are

$$\begin{aligned} y_1 - y_2 \le 0\,, \\ y_1, y_2 \ge 0\,. \end{aligned}$$



- Unfortunately, we choose to increase y_2 and repeat the process again and again......

Clearly it is possible to fail in this case, which is called *degeneracy*. One way to break cycles is to add perturbation in $\boldsymbol{Ax} \le \boldsymbol{b}$, that is, let $b_i' = b_i + \varepsilon_i$ where $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ is an i.i.d.(independent and identically distributed) *Gaussian random variable*.

Now the question is, what if the origin point $\boldsymbol{x} = \mathbf{0}$ is not feasible? If there exists a known feasible solution $\boldsymbol{d}$ such that $\boldsymbol{Ad} \le \boldsymbol{b}$, then let $\boldsymbol{y} = \boldsymbol{d} - \boldsymbol{x}$ and further let $\boldsymbol{y} = \boldsymbol{y}^+ - \boldsymbol{y}^-$ to guarantee that all variables are nonnegative.

However, what if there is no known feasible solutions? The following *two-phase simple method* gives an algorithm to find a feasible solution of a linear program.

- First, convert the linear program into the standard form: $\min \boldsymbol{c}^\top \boldsymbol{x}$ subject to $\boldsymbol{Ax} = \boldsymbol{b}$ and $\boldsymbol{x} \ge \mathbf{0}$, where $\boldsymbol{b} \ge \mathbf{0}$.
- Next, add slack variables $s_1, \dots, s_m$ for constraints. Then the constraints are $\boldsymbol{Ax} + \boldsymbol{s} = \boldsymbol{b}$, $\boldsymbol{x} \ge \mathbf{0}$ and $\boldsymbol{s} \ge \mathbf{0}$.
- Now it is clear that there exists a trivial solution $\boldsymbol{x} = \mathbf{0}$ and $\boldsymbol{s} = \boldsymbol{b}$.
- Finally, we can check whether $\boldsymbol{s} = \mathbf{0}$ is possible by solving the following linear program

$$\begin{aligned} \min_{\boldsymbol{s} \in \mathbb{R}^n} \quad & s_1 + s_2 + \cdots + s_m \\ \text{subject to} \quad & \mathbf{A}\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b}\,,\ \mathbf{A} \in \mathbb{R}^{m \times n},\ \boldsymbol{b} \in \mathbb{R}^m \\ & \forall\, i = 1, 2, \dots, n,\ x_i \ge 0\,; \\ & \forall\, j = 1, 2, \dots, m,\ s_j \ge 0\,. \end{aligned}$$