Convex sets and simplex method

Things left to be proved:
- The correspondence between corner point solutions and basic solutions
- Why it is enough just consider the corner points?

The matrix form of LP model

A general LP model in the standard form is

Max \( z = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \)

Subject to

\[ a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n \leq b_1 \]
\[ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \leq b_2 \]
\[ \vdots \]
\[ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n \leq b_m \]

and

\[ x_1, x_2, \ldots, x_n \geq 0 \]

After we added the slack variables, it becomes

Max \( z = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \)

Subject to

\[ a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n + x_{n+1} = b_1 \]
\[ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n + x_{n+2} = b_2 \]
\[ \vdots \]
\[ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n + x_{m+m} = b_m \]

and

\[ x_1, x_2, \ldots, x_n, x_{n+1}, \ldots, x_{m+m} \geq 0 \]

Let
The system can be expressed as

\[
\begin{align*}
\text{Max} & \quad z = cx \\
\text{s.t.} & \quad Ax = b \\
& \quad x \geq 0
\end{align*}
\]

**Definition:** A set \( S \) is convex if, for any two points, \( x_1, x_2 \in S \), and \( \alpha \in [0, 1] \) imply that \( \alpha x_1 + (1 - \alpha) x_2 \in S \).

An equivalent definition would be:

**Definition:** A set \( S \) is convex if, for any positive integer \( p \) and \( p \) points, \( x_1, x_2, \ldots, x_p \in S \) and 
\( \alpha_1, \alpha_2, \ldots, \alpha_p \in [0, 1] \), \( \alpha_1 + \alpha_2 + \cdots + \alpha_p = 1 \) imply that 
\( \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p \in S \).

\( \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p \) is called a convex combination of \( x_1, x_2, \ldots, x_p \). If none of \( \alpha_1, \alpha_2, \ldots, \alpha_p \) is 0 or 1, then it is a strict convex combination of \( x_1, x_2, \ldots, x_p \).

**Definition:** \( x \) is a basic feasible solution if it satisfies the constraints, and the columns of \( A \) that correspond to the positive components of \( x \) are linearly independent.

**Theorem:** The set \( S = \{ x : Ax = b, x \geq 0 \} \) is a convex set. (4.5-3)

**Definition:** A corner point or extreme point of a convex set \( S \) is a point \( x \) such that \( x \) is not a strict convex combination of any two other points in \( S \).

**Theorem:** A point \( x \) of the set \( S = \{ x : Ax = b, x \geq 0 \} \) is a corner point of \( S \) if and only if it is a basic feasible solution.

**Reminder:** \( Ax = b \) can be understood as \( \sum x_i A_i = b \). Here \( A_i \) is the \( i \)-th column of \( A \).

**Proof of the theorem:** Without loss of generality, we may assume that the components of \( x \) are zero except for the first \( p \) components, namely...
\[ x = \begin{bmatrix} x \\ 0 \end{bmatrix} \]

where \( x = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} \)

We also denote the first \( p \) columns of \( A \) by \( \bar{A} \), namely \( \bar{A} = \begin{bmatrix} A_1 & \cdots & A_p \end{bmatrix} \). We know that \( \bar{A}x = Ax = b \).

(\( \Rightarrow \)) Suppose that \( A_1, \ldots, A_p \) are linearly dependent, then there are \( w_1, \ldots, w_p \) not all zero such that \( \sum w_i A_i = 0 \) or if we let \( w = \begin{bmatrix} w_1 \\ \vdots \\ w_p \end{bmatrix} \), we have \( \bar{A}w = 0 \). Since \( x_1 > 0, \ldots, x_p > 0 \), we may choose a \( \delta \) small enough so that \( x_i \pm \delta w_i > 0 \) for all \( i = 1, \ldots, p \). (We can pick \( \delta < \min \left\{ \frac{x_i}{|w_i|} \right\} \).

Let \( y = \bar{x} + \delta \bar{w}, z = x - \delta \bar{w}, y = \begin{bmatrix} y \\ 0 \end{bmatrix}, z = \begin{bmatrix} z \\ 0 \end{bmatrix} \) then we have \( y \geq 0, z \geq 0 \). Also we have \( A_y = \bar{A}y = \bar{A}(x + \delta \bar{w}) = b, Az = \bar{A}z = \bar{A}(x + \delta \bar{w}) = b \). Therefore \( y, z \in S \). Also we have \( \frac{1}{2}y + \frac{1}{2}z = x \), a contradiction to the fact that \( x \) is a corner point.

(\( \Leftarrow \)) Suppose that \( x \) is not a corner point. Then there are distinct points \( y, z \) in \( S \) such that \( x = ay + (1 - a)z \) for some \( 0 < a < 1 \). Since \( y, z \geq 0 \), the last \( n - p \) components of \( y, z \) must be zero. Then \( w = x - y \) has last \( n - p \) components being zero as well. Therefore \( Aw = \bar{A}w = 0 \).

\( w \neq 0 \Rightarrow w \neq 0 \). That implies that the first \( p \) columns of \( A \) are linearly dependent.

**Theorem:** Let \( S = \{ x : Ax = b, x \geq 0 \} \). If \( S \) is bounded, then every point in \( S \) is a convex combination of some corner points \( x_1, \ldots, x_n \).

[Proof] Let \( x \) be a point in \( S \). We show that \( x \) is a convex combination of some corner points of \( S \) by using induction and an idea similar to the one used in the previous proof.

Induction on the number of positive components in \( x \). Obviously that if that number is zero then \( x \) is a BF solution therefore a corner point itself. Suppose that \( x \) is not a corner point, then the columns of \( A \) corresponding to the positive components of \( x \) are linearly dependent. To simplify the notations, we may assume that these are the first \( r \) columns of \( A \):

\[ x = \begin{bmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} \]

There are \( p_1, p_2, \ldots, p_r \) not all of them zero such that \( p_1A_1 + p_2A_2 + \cdots + p_rA_r = 0 \). Let
We have, and

\[ \mathbf{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \]

then

\[ A\mathbf{p} = p_1 \mathbf{A}_1 + p_2 \mathbf{A}_2 + \cdots + p_r \mathbf{A}_r + \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}. \]

For all \( \varepsilon > 0 \), we have

\[ A(\mathbf{x} + \varepsilon \mathbf{p}) = A\mathbf{x} + \varepsilon A\mathbf{p} = \mathbf{b} + \mathbf{0} = \mathbf{b} \]

So \( \mathbf{x} + \varepsilon \mathbf{p} \) is also in \( S \) if it satisfies the nonnegative constraints. It is the same for \( \mathbf{x} - \varepsilon \mathbf{p} \).

Since \( x_i > 0 \) for \( i = 1, \ldots, r \). So we can choose \( \varepsilon_1 \) small enough such that \( x_i + \varepsilon_1 p_i \geq 0 \) for all \( i \) and we can choose \( \varepsilon_1 > 0 \) such that for one or more \( i \) \( (1 \leq i \leq r) \), \( x_i + \varepsilon_1 p_i = 0 \) (otherwise \( S \) would be unbounded). Similarly we can choose \( \varepsilon_2 > 0 \) such that \( x_i - \varepsilon_2 p_i \geq 0 \) for all \( i \) and \( x_i - \varepsilon_2 p_i = 0 \) for at least one \( i \). (Remember that both \( \mathbf{x} + \varepsilon_1 \mathbf{p} \) and \( \mathbf{x} - \varepsilon_2 \mathbf{p} \) are in \( S \).)

Let \( \mathbf{y} = \mathbf{x} + \varepsilon_1 \mathbf{p} \) and \( \mathbf{z} = \mathbf{x} - \varepsilon_2 \mathbf{p} \). Easy to see that

\[ \mathbf{x} = \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2} \mathbf{y} + \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} \mathbf{z} \]

so \( \mathbf{x} \) is a convex combination of \( \mathbf{y} \) and \( \mathbf{z} \). Both \( \mathbf{y} \) and \( \mathbf{z} \) has at least one less positive components than \( \mathbf{x} \). By the inductive hypothesis, \( \mathbf{y} \) and \( \mathbf{z} \) are convex combinations of the corner points of \( S \):

\[ \mathbf{y} = \sum \alpha_i \mathbf{y}_i \]

\[ \mathbf{z} = \sum \beta_i \mathbf{z}_i \]

where \( \sum \alpha_i = \sum \beta_i = 1, \alpha_i \geq 0, \beta_i \geq 0 \) for all \( i \) and \( \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{z}_1, \mathbf{z}_2, \ldots \) are corner points of \( S \). Let \( \gamma_1 = \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2} \) and \( \gamma_2 = \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} \). We have

\[ \mathbf{x} = \gamma_1 \mathbf{y} + \gamma_2 \mathbf{z} \]

\[ = \sum \gamma_1 \alpha_i \mathbf{y}_i + \sum \gamma_2 \beta_i \mathbf{z}_i \]

Since \( \sum \gamma_1 \alpha_i = \sum \gamma_2 \beta_i = 1 \) and \( \gamma_1 \alpha_i \geq 0, \gamma_2 \beta_i \geq 0 \) for all \( i \), \( \mathbf{x} \) is a convex combination of the corner points.

**Theorem:** If a linear programming problem has an optimal solution, then it has an optimal solution on a corner point of the feasible region.

**Proof:** Suppose that \( \mathbf{z}^* = \mathbf{c}\mathbf{x} \) is an optimal solution. If \( \mathbf{x} \) is a corner point, then we are done. Suppose that \( \mathbf{x} \) is not a corner point, then there are corner points of the feasible region such that there are \( \mathbf{x}_i \) and \( \alpha_i \) that \( \sum \alpha_i \mathbf{x}_i = \mathbf{x} \) and \( \sum \alpha_i = 1 \). Then

\[ \mathbf{c}\mathbf{x} = \mathbf{c}(\sum \alpha_i \mathbf{x}_i) = \sum \alpha_i \mathbf{c}\mathbf{x}_i \leq \sum \alpha_i \mathbf{z}^* \leq \mathbf{z}^*. \]

Since \( \mathbf{z}^* = \mathbf{c}\mathbf{x} \), every step here is an equality. So \( \mathbf{c}\mathbf{x}_i = \mathbf{z}^* \) for all \( i \). Also this shows that \( \mathbf{x} \) is only optimal if it is a convex combination of optimal corner point solutions. That is a part of 4-5.4.

**Theorem:** The feasible region of a linear programming problem has a finite number of corner points.

**Proof:** Suppose that there are \( n \) equations and \( n + m \) variables. Then there are at most \( \binom{m+n}{n} \) ways to choose that \( n \)-linearly independent columns of \( \mathbf{A} \).

**Theorem:** If a corner point has no adjacent corner point that has better objective function value, then
there is no other corner point with better objective function value.

Proof: Clear from the row zero of the simplex tableau.