Developing the Simplex Method

- Basic feasible solutions for LPs in standard form
- Developing the simplex method

1 Basic feasible solutions for LPs in standard form

Recall the definition of a polyhedron, and a basic feasible solution:

- $P \subseteq \mathbb{R}^n$ is a **polyhedron**, if it can be expressed as $P = \{x \in \mathbb{R}^n : Fx \geq g\}$ for some matrix F and vector g.

  (*We use F, g instead of A, b because we want to use A, b in the description of an LP in standard form.*)
- x is a **basic feasible solution (bfs)** of P if $x \in P$ and $\{f_i : f_i^T x = g_i\}$ contains n linearly independent vectors.

We also recall the standard form LP:

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. We assume A has m linearly independent rows (i.e., $rank(A) = m$). It is possible to show that this assumption is without loss of generality: see section 2.3 (page 57).

The feasible region of the LP is $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$. Note that the feasible region is a polyhedron: just take $F = \begin{bmatrix} A \\ -A \\ I_n \end{bmatrix}, g = \begin{bmatrix} b \\ -b \\ \vec{0} \end{bmatrix}$, where I_n is the $n \times n$ identity matrix, and $\vec{0}$ is an n-dimensional 0-vector. So when is x a basic feasible solution of the feasible region?

- x must be in the feasible region $\Rightarrow$ x must satisfy $Ax = b, x \geq 0$
- $\{f_i : f_i^T x = g_i\}$ must contain n linearly independent vectors.
  - we get m linearly independent vectors in $\{a_i : a_i^T x = b_i\}$;
  - we need to have at least $n - m$ more constraints that hold at equality from $I_n x \geq \vec{0} \Rightarrow x_j = 0$ for at least $n - m$ indices j;
  - to check that we satisfy the linear independence requirement, let e_j^T be the j-th row of I_n. Then $\{a_i : a_i^T x = b_i\} \cup \{e_j : e_j^T x = 0\}$ must contain n linearly independent vectors.

    Consider the matrix $F_=$ that has $\{a_i^T : a_i^T x = b_i\} \cup \{e_i^T : e_i^T x = 0\}$ as its rows. Assume without loss of generality that the first m rows are $a_1^T, \ldots, a_m^T$. The matrix has n columns, and we need to verify that the row rank of this matrix is n. Then, the column rank of this matrix must also be n, so the columns must be linearly independent.

    What do the columns of $F_=$ look like?

    * Consider j such that $x_j > 0$. The j-th column is just $\begin{bmatrix} A_j \\ \vec{0} \end{bmatrix}$.

* Consider j such that $x_j = 0$. Then e_j^T is a row of the matrix $F_=$, say row $k > m$. The j-th column of $F_=$ is similar to the one for j with $x_j > 0$, except that the k-th entry is 1. Since this is the only column that has a non-zero in the k-th entry, this column must be linearly independent of the other columns.

So, we need that the columns of $F_=$ corresponding to the entries j with $x_j > 0$ are linearly independent, which thus means that the columns A_j for j such that $x_j > 0$ must be linearly independent.

We can now easily prove the following theorem (Theorem 2.4 in the book):

Theorem 1. *Consider an LP in standard form. Then x is a basic feasible solution if and only if $Ax = b, x \geq 0$ and there exists indices $B(1), \ldots, B(m)$ such that*

(a) The columns $A_{B(1)}, \ldots, A_{B(m)}$ are linearly independent;

(b) If $j \neq$ [†] $B(1), \ldots, B(m)$, then $x_j = 0$.

Proof. ($\Rightarrow$) Suppose x is a basic feasible solution. Let $B(1), \ldots, B(k)$ be the indices j such that $x_j > 0$. We already saw that $k \leq m$, and that the columns $A_{B(1)}, \ldots, A_{B(k)}$ must be linearly independent. If $k < m$, then we can find $m - k$ additional columns $A_{B(k+1)}, \ldots, A_{B(m)}$ such that $A_{B(1)}, \ldots, A_{B(m)}$ are linearly independent (this is because A has rank m – see Theorem 1.3(b) in Section 1.5).

($\Leftarrow$) The columns of $F_=$ as defined above are linearly independent, and therefore $F_=$ has n linearly independent rows, so x is a basic feasible solution.

□

Some more notation and terminology:

- We call x a **basic solution** if x satisfies the conditions (a) and (b) in the theorem, and $Ax = b$ (but not necessarily $x \geq 0$).
- If x is a basic solution or a basic feasible solution, then $x_{B(1)}, \ldots, x_{B(m)}$ are called the **basic variables**, and the remaining variables are called **nonbasic**. The columns $A_{B(1)}, \ldots, A_{B(m)}$ are called the **basis**.
- We use $\mathbf{B}$ to denote the following matrix: $\mathbf{B} = \begin{bmatrix} A_{B(1)} & A_{B(2)} & \ldots & A_{B(m)} \end{bmatrix}$. This is called the basis matrix, and its columns are the basic columns.
- We use x_B to denote the following vector: $x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}$.

  In general, a vector $d \in \mathbb{R}^n$ has $d_B = \begin{bmatrix} d_{B(1)} \\ \vdots \\ d_{B(m)} \end{bmatrix}$

Finally, before we start thinking about algorithms, consider two questions:

- **Do all bases give different solutions?**

  We can suspect the answer is "no" by looking at the proof of the theorem: if we have $k < m$ indices j for which $x_j > 0$, then we fix only k basic columns, and we can choose the remaining $m - k$ columns as long as we ensure that the columns are linearly independent (and, so, possibly, there is more than one choice).

  The answer is indeed "no", see example 2.5 in the book. This is called degeneracy:

[†] there was a typo here in the first version of these notes

Definition 1. *Consider a standard form polyhedron* $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$*, and let* x *be a basic solution. Let* m *be the number of linearly independent rows of* A*. Then* x *is a degenerate basic solution, if more than* $n - m$ *of the components of* x *are zero.*

- **We know there is an optimal bfs if the feasible region is bounded and non-empty. What happens if the feasible region is unbounded?**

  It is possible to show that if the feasible region is in standard form and it is non-empty, then either, there is a basic feasible solution that is optimal, or the problem is unbounded; i.e., the optimal value is $-\infty$. We will not prove this now, but we will indirectly prove this by showing the correctness of the simplex algorithm.

2 Developing the simplex method

Consider an LP in standard form. By the theorem in the previous section, we know how to express the basic feasible solutions. We also know that if the feasible region is non empty and bounded, then we can find an optimal solution by checking all the basic feasible solutions, which we can do by checking all bases:

- Let $opt = \infty$, $x^* = [\ ]$;
- Repeat:
  - Choose m linearly independent columns from A, say $A_{B(1)}, \ldots, A_{B(m)}$;
  - Let $x_j = 0$ for all $j \neq B(1), \ldots, B(m)$;
  - Determine x_B by solving $\mathbf{B}x_B = b$;
  - If $x_B \geq 0$, and $c^T x < opt$, then $opt \leftarrow c^T x, x^* \leftarrow x$.

This method is not fully general, because it only works for bounded feasible regions. We could also wonder if there is a more intelligent way than simply checking all bases?

Suppose we have some basic feasible solution x. Can we determine a new basis to check, so that the new basic feasible solution is guaranteed not to be worse than our current basic feasible solution?

2.1 Choosing which nonbasic variable to increase

Example:

$$\begin{array}{llllll} \min & -2x_1 & +x_2 & -3x_3 & +x_4 & \\ \text{s.t.} & x_1 & +x_2 & +x_3 & +3x_4 & = 2 \\ & 2x_1 & & +x_3 & +4x_4 & = 2 \\ & x_1, & x_2, & x_3, & x_4 & \geq 0 \end{array}$$

Choose the first two columns of A as the first basis:

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}, \mathbf{B}^{-1} = \begin{bmatrix} 0 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{bmatrix}.$$

Set $x_3 = x_4 = 0$ and solve $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{B}^{-1} \begin{bmatrix} 2 \\ 2 \end{bmatrix} \Rightarrow x_1 = 1, x_2 = 1^\dagger$ to find the bfs.

It seems like a good idea to make x_3 a basic variable, since it has objective coefficent -3. If we increase x_3 by θ, how should we change x_1, x_2 to ensure the solution remains feasible? Let $\theta d_1, \theta d_2$ be the amount by which we should change x_1 and x_2 if we increase x_3 by θ.

†there was a typo here in the first version of these notes

We need

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 0 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1+\theta d_1 \\ 1+\theta d_2 \\ 0+\theta \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

which is the same as

$$\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} \theta d_1 \\ \theta d_2 \end{bmatrix} = -\theta \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

So, we need that $\mathbf{B} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = - \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. This gives $d_1 = -3/2, d_2 = 1/2$.

We thus find that if we increase x_3 by θ, we need to change x_1 by $-\frac{3}{2}\theta$ and x_2 by $\frac{1}{2}\theta$, in order to remain feasible. So the objective changes by $-2(-\frac{3}{2}\theta) + 1(\frac{1}{2}\theta) - 3\theta = \frac{1}{2}\theta$. So, in fact, the objective value increases! Can we generalize what we have learnt?