1 Basic feasible solutions

Let’s suppose we are solving a general linear program in equational form:

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b \\
& \quad x \geq 0.
\end{align*}
\]

Here, \(A\) is an \(m \times n\) matrix, \(b \in \mathbb{R}^m\), and \(c \in \mathbb{R}^n\). Today, we will assume that the rows of \(A\) are linearly independent. (If not, then either the system \(Ax = b\) has no solutions, or else some of the equations are redundant. In the first case, we just forget about analyzing such a linear program; in the second case, we can begin by deleting the redundant rows.)

We’ve informally said that a \textit{basic feasible solution} is one in which “as many of the variables as possible” are 0. This is not quite precise: in some cases (due to degeneracy) it’s possible to have unusually many 0 values, and we don’t want this to mess with our definition. Instead we make the definition as follows.

Choose some \(m\)-tuple of columns (or of variables) \(B\) to be basic. We want \(B\) to be ordered, because our tableaux look slightly different when the basic variables are chosen in a different order. For convenience, we let \(N\) be the \((n - m)\)-tuple of nonbasic variables: those that are not in \(B\).

We can split up vectors and matrices into the basic and nonbasic part. For example, if \(x = (x_1, x_2, x_3, x_4, x_5)\), \(B = (2, 4)\), and \(N = (1, 3, 5)\), we have \(x_B = (x_2, x_4)\) and \(x_N = (x_1, x_3, x_5)\). This can also be done with \(A\) and \(c\): we can write the objective function as

\[
c^T x = c_B^T x_B + c_N^T x_N
\]

and the system of equations \(Ax = b\) as

\[
A_B x_B + A_N x_N = b.
\]

To get a basic solution, we want to choose \(B\) so that \(A_B\) (an \(m \times m\) matrix) is invertible. This is always possible if the rows of \(A\) are linearly independent. Not every choice of \(B\) will work: for example, in 2 dimensions, if two of the sides of the feasible region are parallel lines, they never intersect.

Now set \(x_N = 0\), and \(x_B = A_B^{-1} b\). This satisfies \(Ax = b\). If, additionally, we have \(x_B \geq 0\) (we always have \(x_N \geq 0\), because \(x_N = 0\)), we call \(x\) a \textit{basic feasible solution}.

---

1This document comes from the Math 482 course webpage: [https://faculty.math.illinois.edu/~mlavrov/courses/482-fall-2019.html](https://faculty.math.illinois.edu/~mlavrov/courses/482-fall-2019.html)
2 Other notions of corner points

There are other notions of “corner point” besides a basic feasible solution. We say that

- A **vertex** of a set $S \subseteq \mathbb{R}^n$ is a point $x \in S$ such that some linear function $\alpha^T x$ is strictly minimized at $x$: $\alpha^T x < \alpha^T y$ for any $y \in S, y \neq x$.

- An **extreme point** of a set $S \subseteq \mathbb{R}^n$ is a point $x \in S$ that does not lie between any other points of $S$. Formally, if $x$ is an extreme point if, whenever $x \in [y, y']$ for $y, y' \in S$, either $x = y$ or $x = y'$.

In other words, if $x$ can be written as $t y + (1 - t) y'$ for $y, y' \in S$ and $0 \leq t \leq 1$, either $x = y$ (and we can set $t = 1$) or $x = y'$ (and we can set $t = 0$).

When the set we’re considering is $F = \{x \in \mathbb{R}^n : A x = b, x \geq 0\}$, the feasible region of a linear program, all three notions—basic feasible solution, vertex, and extreme point—are the same. This is what we’ll try to prove today.

2.1 From basic feasible solutions to vertices

**Proposition 2.1.** Any basic feasible solution is a vertex of the feasible region.

*Proof.* Take any choice of basic and nonbasic variables $(B, N)$ for which setting $x_N = 0$ produces a basic feasible solution. Define $\alpha$ by

\[
\alpha_i = \begin{cases} 
1 & i \in N, \\
0 & i \in B.
\end{cases}
\]

Then $a^T x$ is the sum of the nonbasic variables in $x$.

Since $x_N \geq 0$, $a^T x$ is minimized exactly when we set $x_N = 0$. And that’s exactly the basic feasible solution corresponding to $(B, N)$. \qed

2.2 From vertices to extreme points

**Proposition 2.2.** Any vertex of a set $S \subseteq \mathbb{R}^n$ is also an extreme point of $S$. (In particular, any basic feasible solution is also an extreme point of the feasible region.)

*Proof.* Let $x \in S$ be a vertex of $S$, and et $\alpha$ be the vector such that $\alpha^T x < \alpha^T y$ for any $y \in S$ with $y \neq x$.

Suppose, for the sake of contradiction, that $x$ lies on the line segment $[y, y']$ with $y, y' \in S$ and $y, y' \neq x$. This is what it means to *not* be an extreme point.

The $x = t y + (1 - t) y'$, so

\[
\alpha^T x = t(\alpha^T y) + (1 - t)(\alpha^T y') < t(\alpha^T x) + (1 - t)(\alpha^T x) = \alpha^T x,
\]

which is a contradiction.

Therefore $x$ must be an extreme point. \qed
2.3 From extreme points to basic feasible solutions

The last step, going from extreme points to basic feasible solutions, is trickier.

Proposition 2.3. Any extreme point of the feasible region is a basic feasible solution.

Proof. Let \( x \) be any extreme point of the feasible region. Define \( \mathcal{W} = \{ i : x_i \neq 0 \} \) and \( \mathcal{Z} = \{ i : x_i = 0 \} \), by analogy with \( \mathcal{B} \) and \( \mathcal{N} \) for a basic feasible solution.

We can’t really ask whether \( A_W \) is invertible or not, because we have no reason to even think it’s square. But what we can do is ask whether we can find a nonzero \( u \in \mathbb{R}^n \) such that:

\[
\begin{cases}
A_W u_W = 0 \\
u_Z = 0.
\end{cases}
\]

If there is no such \( u \), then we’ll argue that \( x \) is a basic feasible solution. If there is such an \( u \), then we’ll argue that \( x \) actually wasn’t an extreme point, and get a contradiction.

First, suppose there is no such \( u \): whenever \( A_W u_W = 0 \), we have \( u = 0 \). Here, we’ll need some linear algebra. The columns indexed by \( \mathcal{W} \) are \(|\mathcal{W}| \) linearly independent columns, so we know that \(|\mathcal{W}| \leq m \) (because the columns are vectors in \( \mathbb{R}^m \)). Becaus \( A \) has full row rank, we know that we can extend \( \mathcal{W} \) to some \( \mathcal{B} \) (with \(|\mathcal{B}| = m \)) such that the columns indexed by \( \mathcal{B} \) are still linearly independent, and therefore \( \mathcal{B} \) is invertible.

Now, let \( \mathcal{N} \) be the complement of \( \mathcal{B} \). Because we found \( \mathcal{B} \) by starting with \( \mathcal{W} \) and possibly making it bigger, we know that \( \mathcal{N} \) is found by starting with \( \mathcal{Z} \) and possibly making it smaller. Because \( x_Z = 0 \) (that’s how we chose \( \mathcal{Z} \)), we know that \( x_N = 0 \) as well.

Now we’re nearly done. \( A x = A_B x_B + A_N x_N = b \). Since \( x_N = 0 \), we have \( A_B x_B = b \), so \( x_B = A_B^{-1} b \). This is exactly what we wanted from a basic solution.

Second, suppose such a \( u \) exists. Then we have

\[ A u = A_W u_W + A_Z u_i = 0 + 0 = 0. \]

This means that points of the form \( x + t u \) still satisfy the system of equations: that is, \( A(x + t u) = b + t 0 = b \).

Not all points of the form \( x + t u \) are feasible: for some \( t \), we can find a coordinate \( i \) such that \( x_i + t u_i < 0 \). However, we know that such an \( i \) must be in \( \mathcal{W} \), not \( \mathcal{Z} \), because for any \( i \in \mathcal{Z} \), we’d have \( x_i = u_i = 0 \). So \( x > 0 \). This means that when \(|t|\) is sufficiently small, \( x_i + t u_i > 0 \) as well.

Choose a small enough \( t > 0 \) that \( x + t u \) and \( x - t u \) are both feasible: \( x_i + t u_i > 0 \) and \( x_i - t u_i > 0 \) for each \( i \). This is a lot like pivoting: it’s enough to ask that \(|t| < \frac{|x_i|}{|u_i|} \) for each \( i \) such that \( u_i \neq 0 \).

But now, \( x \) lies on the line segment from \( x + t u \) to \( x - t u \). In fact, it’s the midpoint of that line segment. Because \( u \neq 0 \), and \( t > 0 \), it’s distinct from both endpoints. So \( x \) is not an extreme point in this case, contrary to assumption. \( \square \)