Defn. If a basic solution is feasible, it is a basic feasible solution (bfs).

Convex combinations and extreme points.

Defn. Let $x, y \in \mathbb{R}^n$. A convex combination of $x$ and $y$ is a point (i.e., vector) of the form

$$z = \lambda x + (1-\lambda) y,$$

where $\lambda$ is a scalar in the interval $[0, 1]$. If $0 < \lambda < 1$, $z$ is a strict convex combination.

- Note that the set of all convex combinations of $x$ and $y$ forms the line segment joining $x$ and $y$.

Defn. A set $S \subseteq \mathbb{R}^n$ is convex if for all $x, y \in S$, all convex combinations of $x$ and $y$ are also in $S$.

Example: Not convex →

Problem 4 on first problem set asks you to prove that the feasible region of an LP is convex.
Defn. An extreme point of a convex set $S$ is a point $x \in S$ such that if $x$ is expressed as a strict convex combination of two points $y$ and $z$ in $S$, i.e., if

$$x = \lambda y + (1-\lambda)z$$

for some $0 < \lambda < 1$,

then $y = z$.

- Intuitively: An extreme point does not lie between two other points in the set.
- An extreme point of a polytope is a corner.

[P&S Thm 2.47]

Theorem. A solution $x \in \mathbb{R}^n$ is an extreme point of the feasible region of an LP (in std form) if and only if it is a basic feasible solution.

Proof. By reordering the components of $x$ (and the columns of the coefficient matrix $A$), we may assume WLOG that the first $r$ components of $x$ are nonzero and the remaining $n-r$ components are zero:

$$x_i = 0 \text{ for } 1 \leq i \leq r,$$

$$x_i = 0 \text{ for } r+1 \leq i \leq n.$$

Then $x$ is a bfs if and only if the first $r$ columns of $A$ are linearly independent.
[bfs ⇔ extreme point of feas. region]

(bfs ⇒ extreme point.) Suppose x is a bfs, then the first r columns \( A_1, \ldots, A_r \) of A are lin. independent. Let x be written as a strict convex combination of two points y and z in the feasible region:

\[
x = \lambda y + (1-\lambda)z
\]

for some \( \lambda \in (0, 1) \). Since x, y, and z are all feasible solutions, we know that

\[
Ax = b \quad Ay = b \quad Az = b \\
x \geq 0 \quad y \geq 0 \quad z \geq 0.
\]

For \( r+1 \leq i \leq n \), we have

\[
\frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0} = \frac{x_i}{0}.
\]

So \( y_i = z_i = 0 = x_i \). Therefore,

\[
A_1 x_1 + \ldots + A_r x_r = Ax = b \\
A_1 y_1 + \ldots + A_r y_r = Ay = b \\
A_1 z_1 + \ldots + A_r z_r = Az = b.
\]

But by assumption \( A_1, \ldots, A_r \) are lin indep, so the representation of b as a lin combo is unique, and hence \( x = y = z \). So x is an extreme point of the feasible region. \( \blacksquare \)
(Not bfs ⇒ not extreme point.) Suppose \( x \) is not a bfs, so the first \( r \) columns \( \{A_1, \ldots, A_r\} \) of \( A \) are not linearly independent. Then there exist scalars \( \alpha_1, \ldots, \alpha_r \), not all zero, such that

\[
A_1 \alpha_1 + \cdots + A_r \alpha_r = 0.
\]

Let \( \alpha \) be the vector \( \alpha = [\alpha_1, \ldots, \alpha_r, 0, \ldots, 0]^T \), so that \( A\alpha = 0 \). Choose \( \lambda > 0 \) small enough that \( \lambda |\alpha_i| \leq x_i \) for all \( 1 \leq i \leq r \).

[E.g., \( \lambda = \min \frac{x_i}{|\alpha_i|} : 1 \leq i \leq r, \alpha_i \neq 0 \).]

Let \( w = x + \lambda \alpha \) and \( \overline{w} = x - \lambda \alpha \).

By our choice of \( \lambda \), we have \( w \geq 0 \) and \( \overline{w} \geq 0 \). Also,

\[
Aw = A(x + \lambda \alpha) = Ax + \lambda A\alpha = b + 0 = b,
\]

\[
A\overline{w} = A(x - \lambda \alpha) = Ax - \lambda A\alpha = b - 0 = b.
\]

So \( w \) and \( \overline{w} \) are feasible solutions, and \( w \neq x \) and \( \overline{w} \neq x \) because \( \lambda > 0 \) and \( \alpha \neq 0 \).

But \( x = \frac{1}{2} w + \frac{1}{2} \overline{w} \), so \( x \) is not an extreme point of the feasible region. \( \blacksquare \)
Theorem. Let $x$ be a feasible solution to an LP. Then either there exists a bfs whose objective value is at least as good as that of $x$, or else the LP is unbounded.

Sketch of proof via example. (Farmer Brown)

\[ \begin{align*}
\text{max} & \quad 40p + 120w \\
\text{s.t.} & \quad p + w + s_1 = 100 \\
& \quad p + 4w + s_2 = 160 \\
& \quad 10p + 20w + s_3 = 1100 \\
& \quad p \geq 0, \quad w \geq 0, \quad s_1 \geq 0, \quad s_2 \geq 0, \quad s_3 \geq 0.
\end{align*} \]

A basic solution will have at most three nonzero components.

\[ x = [20, 35, 45, 0, 200]^T \] is feasible but not basic. (too many nonzero components).

We can decrease the number of nonzeros as follows.

First, find a nonzero solution to the homogeneous system

\[ \begin{align*}
p + w + s_1 &= 0 \\
p + 4w + s_2 &= 0 \\
10p + 20w + s_3 &= 0
\end{align*} \]

having $s_2 = 0$ (to preserve the zero we already have).
Set $s_2 = 0$. Now we have three equations in four unknowns, so one free variable. Set $w = 1$, say, to guarantee a nonzero solution. So we have

\[
\begin{align*}
  p + 1 + s_1 &= 0 \\
  p + 4 &= 0 \\
  10p + 20 + s_3 &= 0
\end{align*}
\]

Solution is $p = -4$, $s_1 = 3$, $s_3 = 20$. So full solution to homogeneous system is $h = [-4, 1, 3, 0, 20]^T$.

By linearity, we can add any scalar multiple of $h$ to $x$, and the resulting vector $y = x + th$ (for $t \in \mathbb{R}$) will also satisfy the constraints $Ay = b$, because

\[
Ay = A(x + th) = Ax + t(Ah) = b + 0 = b.
\]

The resulting objective value will be

\[
c^T y = c^T(x + th) = c^T x + t(c^T h).
\]

Since we want this to be at least as good (i.e., large) as $c^T x$, we want $t(c^T h) \geq 0$. We have $c^T h = 40(-4) + 120(1) = -40$,

so we will take $t < 0$. (Alternatively, use $-h$ instead of $h$ and take $t > 0$.)

Now, how large in magnitude can $t$ be?

\[
y = x + th = [20 - 4t, 35 + t, 45 + 3t, 0, 200 + 20t]^T.
\]

The components of $y$ cannot be negative, so the second, third, and fifth components impose the constraints

\[
\begin{align*}
  35 + t &\geq 0 \\
  45 + 3t &\geq 0 \\
  200 + 20t &\geq 0
\end{align*}
\]

\[\implies\]

\[
\begin{align*}
  t &\geq -35 \\
  t &\geq -15 \\
  t &\geq -10
\end{align*}
\]
1 June.

(Note that first component imposes no constraint if \( t < 0 \).)

The strongest of these three constraints is \( t \geq -10 \).

Taking \( t = -10 \), then, we get

\[
y = x + th = [20 - 4(-10), 35 - 10, 45 + 3(-10), 0, 200 + 20(-10)]^T
\]

\[
= [60, 25, 15, 0, 0]^T.
\]

This solution \( y \)

- satisfies constraints, because we began with feasible solution \( x \) and added a solution to the homogeneous system;
- has all nonnegative components;
- has objective value at least as large as that of \( x \), because \( t(c^T h) \geq 0 \), hence is feasible.

Note: If we had no constraints on the magnitude of \( t \), then we could make \( t \) arbitrarily large in magnitude, thereby increasing the objective value indefinitely; so the LP would be unbounded (and the ray \( x + th \) would be a certificate of this fact).

Small catch: What if \( c^T h = 0 \)?

Well, then we can use either \( h \) or \( -h \), and at least one of these will place constraints on the magnitude of \( t \).
Corollary. If an LP has an optimal feasible solution, then it has an optimal basic feasible solution.

This justifies the geometric observation we made earlier, that we need only consider the corners of the feasible region when seeking an optimal solution (to a bounded LP).
The simplex tableau and pivoting.

Example. (Farmer Brown again)

\[
\begin{align*}
\text{max } & \quad 40p + 120w \\
\text{s.t. } & \quad p + w + s_1 = 100 \\
& \quad p + 4w + s_2 = 160 \\
& \quad 10p + 20w + s_3 = 1100 \\
& \quad p \geq 0, \; w \geq 0, \; s_1 \geq 0, \; s_2 \geq 0, \; s_3 \geq 0.
\end{align*}
\]

Ignoring the objective for now, we can express the constraints with the augmented matrix

\[
\begin{array}{cccccc|c}
 & p & w & s_1 & s_2 & s_3 & \text{RHS} \\
\hline
1 & 1 & 1 & 0 & 0 & 0 & 100 \\
1 & 4 & 0 & 1 & 0 & 0 & 160 \\
10 & 20 & 0 & 0 & 1 & 0 & 1100 \\
\end{array}
\]

This is called a tableau (except that it is missing a row for the objective—later).

P&S write the RHS column on the left-hand side of the tableau, presumably because certain implementation details may become simpler if RHS is column 0.

The coefficient matrix here has an "obvious" basis, having basic variables \(s_1, s_2, s_3\)?. Corresponding basic solution is \(p=0, \; w=0, \; s_1=100, \; s_2=160, \; s_3=1100\).
Suppose we want to choose a different basis, e.g., \( W, s_1, s_3 \). What is the corresponding basic solution? We know we have \( p = 0 \) and \( s_2 = 0 \). To find values of \( W, s_1, \) and \( s_3 \), we solve the resulting system, which we can do via Gauss-Jordan elimination. Observe that the \( s_1 \) and \( s_3 \) columns are already columns of the identity matrix, so if we just make the \( W \) column be \([0,1,0]^T\) then we can read off the solution. We can do this with row operations:

\[
\begin{array}{cccc|c}
p & W & s_1 & s_2 & s_3 & \text{RHS} \\
1 & 1 & 1 & 0 & 0 & 100 \\
1 & 0 & 0 & 1 & 0 & 160 \\
10 & 20 & 0 & 0 & 1 & 1100 \\
\end{array}
\]

Multiply row 2 by \( 1/4 \)

\[
\begin{array}{cccc|c}
p & W & s_1 & s_2 & s_3 & \text{RHS} \\
1 & 1 & 1 & 0 & 0 & 100 \\
1/4 & 0 & 1/4 & 0 & 0 & 40 \\
10 & 20 & 0 & 0 & 1 & 1100 \\
\end{array}
\]

Subtract row 2 from row 1

Subtract \( 20 \) (row 2) from row 3

This operation is called a pivot.
Now the "obvious" basis is $w, s_1, s_3$, and the corresponding basic solution is clearly $p=0$, $w=40$, $s_1=60$, $s_2=0$, $s_3=300$.

Observe:

- This basic solution happens to be feasible: all variables have nonnegative values. (The solution still satisfies the system $Ax=b$ because row operations do not change the set of solutions to a system.)

- One pivot caused a single variable ($w$) to enter the basis, and another single variable ($s_2$) to leave the basis.

- This pivot caused us to move from one corner of the feasible region ($p=0$, $w=0$) to an adjacent corner ($p=0$, $w=40$).

- Pivoting on an entry means using row operations to make that entry 1 and all other entries in its column 0.