Advanced Linear and Nonlinear Programming

Polyhedron Theory

Eg:

minimize \(-2x_1 + x_2\)

subject to

\begin{align*}
  x_1 + 3x_3 + x_5 &= 12 \\
  x_1 + x_4 + x_5 &= 6 \\
  2x_1 + 3x_5 + x_7 &= 15 \\
  x_1, x_2, x_3, x_4, x_5 &\geq 0.
\end{align*}

This is equivalent to

\begin{align*}
  \text{min} & \quad -2x_1 + x_2 \\
  \text{s.t.} & \quad x_1 + 3x_3 \leq 12 \\
 & \quad x_1 \leq 6 \\
 & \quad 2x_1 + 3x_5 \leq 15 \\
 & \quad x_1, x_2 \geq 0.
\end{align*}

Graph (d):
The optimal point is one of the extreme points; pull the line normal to $-2x_1 + x_2 = 0$ as far as possible.

So optimal point is $x_1 = 6, x_2 = 0$.

In the original variables this corresponds to the point $x_1 = 6, x_2 = 0, x_3 = 6, x_4 = 0, x_5 = 3$.

So have three variables which are nonzero, and two which are zero.

Now number of equations is also three. In general,

Look at other extreme points.
$x_1 = 6, x_2 = 1$; $x_3 = 3, x_4 = 6, x_5 = 0$.
So again have three nonzero and two zeros.

$x_1 = 0, x_2 = 0$; $x_3 = 12, x_4 = 6, x_5 = 0$; again, three nonzero.
$x_1 = 3, x_2 = 3$; $x_3 = 0, x_4 = 3, x_5 = 0$; two zeros.
$x_1 = 0, x_2 = 4$; $x_3 = 0, x_4 = 6, x_5 = 3$.

Suggest: could find extreme points by
(a) setting two variables to zero, then
(b) solving the three equations in the remaining three variables.

E.g. (i) set $x_1, x_3$ to 0, solve $3x_5 = 12$ $x_4 = 6$
$3x_5 + x_5 = 15$
$6, 0, x_5 (x_3 = 0), x_4 = 4, (x_3 = 0), x_4 = 6, x_5 = 3$.

(ii) set $x_2, x_4$ to 0, solve $x_1 + 3x_2 = 12$
$x_5 = 6$
$2x_1 + 3x_2 + x_5 = 15$
$6, 3, 0, x_1 = 6, x_2 = 2, (x_3 = 0), (x_4 = 0), x_5 = 3$ Not feasible.
So suggest algorithm:

For each pair of variables \( x_i, x_j \)
- Set \( x_i, x_j = 0 \). Solve three equations in three remaining variables.
- If solution is feasible (i.e., nonnegative), calculate objective function value.
- Pick best feasible point.

Works in general: (partially equations are linear independent)
- If have \( m \) equations in \( n \) nonnegative variables,
  - set each set of \( n-m \) variables to zero, solve for remaining variables.

Note: have \( (n^m) \) steps to this algorithm, so finite, but large.

In graph, feasible region clearly the dimension \( 2 \).

Also true of graphical problem in five dimensions:
- feasible region would be a two dimensional slice with five edges, corresponding to each \( x_i = 0 \).
Standard form linear program:

\[
\begin{align*}
\text{min} & \quad e^T x \\
\text{s.t.} & \quad A x = b \\
& \quad x \geq 0.
\end{align*}
\]

(A is an \(m \times n\) matrix, \(C\) and \(x\) are \(n\)-vectors, and \(b\) is an \(m\)-vector.)

This formulation is general.

E.g.: \[
\begin{align*}
\text{min} & \quad g^T x \\
\text{s.t.} & \quad H x \leq b \\
& \quad x \text{ unrestricted in sign.}
\end{align*}
\]

Equivalent to \[
\begin{align*}
\text{min} & \quad g^T x \\
\text{s.t.} & \quad H x + I s = b \\
& \quad s \geq 0
\end{align*}
\]

\(I = \text{max \ identity \ matrix \ } [1 \ 0; 0 \ 1]\)

Equivalent to \[
\begin{align*}
\text{min} & \quad g^T u - g^T v \\
\text{s.t.} & \quad H u - H v + I s = b \\
& \quad u, v, s \geq 0
\end{align*}
\]

\(\text{(make substitution } x = u - v, \ u, v \geq 0).\)
Some definitions:

Given $x, y \in \mathbb{R}^n$, $\{x, y\}$ denotes the line segment joining $x$ and $y$.

$\{x, y\} = \{z \in \mathbb{R}^n : z = \lambda x + (1-\lambda)y, \quad 0 \leq \lambda \leq 1\}$

is the set of all convex combinations of $x$ and $y$.

A set $C \subseteq \mathbb{R}^n$ is convex if the line segment joining any two points in $C$ is also in $C$.

Examples:

- Convex:
  ![Convex example]

- Not convex:
  ![Not convex example]

Subspaces: A point $z \in \mathbb{R}^n$ is a linear combination of $x$ and $y$ if $z = \lambda_1 x + \lambda_2 y$ for some $\lambda_1, \lambda_2 \in \mathbb{R}$.

A set $S$ is called a subspace of $\mathbb{R}^n$ if every linear combination of points in $S$ is also in $S$.

Note: If $S$ is nonempty, it must contain the origin.

Subspaces are straight lines, planes, ... which contain the origin.

Affine sets: A point $z \in \mathbb{R}^n$ is an affine combination of $x$ and $y$ if $z = \lambda x + (1-\lambda)y$ for some $\lambda \in \mathbb{R}$ (no restriction on $\lambda$).

A set $F \subseteq \mathbb{R}^n$ is affine if every affine combination of points in $F$ is also in $F$.

Also called linear manifolds, planes.

![Affine example]
Halfspace: \( \{ x \in \mathbb{R}^n : a^T x \leq \alpha \} \), \( a, \alpha \) given.

Polyhedron: Intersection of a finite number of halfspaces. The feasible region of a linear programming problem is a polyhedron.

Hyperplane: \( \{ x \in \mathbb{R}^n : a^T x = \alpha \} \), \( a, \alpha \) given.

The hyperplane \( H = \{ x \in \mathbb{R}^n : a^T x = \alpha \} \) supports the convex set \( C \) if (a) either \( a^T x \leq \alpha \) or \( a^T x \geq \alpha \) for all \( x \in C \) and (b) \( a^T x = \alpha \) for some \( x \in C \).

For \( C \subseteq \mathbb{R}^n \), convex set, the point \( x \) in \( C \) is an extremal point of \( C \) if it cannot be expressed as a convex combination of two other distinct points in \( C \).
The feasible region of a linear programming problem

\[ K = \{ x \in \mathbb{R}^n : A x = b, x \geq 0 \} \]

is a \textit{polyhedron}, i.e. the intersection of a finite number of closed half spaces.

Recall: The vector \( x^1, \ldots, x^n \in \mathbb{R}^n \) are \textit{linearly independent} if

\[ \sum_{i=1}^n \lambda_i x^i = 0 \quad \text{implies} \quad \lambda_i = 0, \quad i = 1, \ldots, n. \]

\[ \begin{bmatrix} 1 \\ 0 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \]

can be only \[ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 2 \\ 2 \end{bmatrix} \]

\[ \longleftarrow \text{Linearly independent}. \]

\[ \text{Let } K = \text{polyhedron}. \text{ Then } a \text{ face of } K \text{ is either} \]

(i) \( K \text{ itself}\)

(ii) \( K \text{ empty set} \)

(iii) \( K \text{ and a supporting hyperplane } H, K \cap H \)

\[ \text{Diagram of polyhedron.} \]
Dimension: \( S \) is a subspace of \( \mathbb{R}^n \) if \( \{ x, x', \ldots, x_s \} \in S \), \( a, b \in \mathbb{R} \)
\( a x' + b x \) is always \( S \).
For the dimension of \( S = \text{dim}(S) \) is the size of the maximal linearly independent set of vectors in \( S \).

Let \( L \subseteq \mathbb{R}^n \) is a linear space of \( x, x', x'' \in L \), \( x, y \in \mathbb{R} \).
\( L \) is an affine space, and \( x^0 \) is \( L \) then,
\( L_5 = \{ x, y \in \mathbb{R}^n \mid y = x - x^0, y \in L \} \) is a subspace of \( \mathbb{R}^n \).

(Exercise: \( L_5 \) does not depend on \( x^0 \))
Then define \( \text{dim}(L) = \text{dim}(L_5) \).

Let \( W \) be a subset of \( \mathbb{R}^n \).
(1) Affine Hull of \( W \)
\( \text{aff}(W) = \{ x \in \mathbb{R}^n \mid x = \sum \alpha_i x_i \text{ with } x_i \in W \} \).
(2) \( \text{aff}(W) \)

(i) \( W = \text{aff}(W) \)
(iv) \( W \subseteq \text{aff}(W) \)

The \( \text{dim}(W) = \text{dim}(\text{aff}(W)) \).

1. \( W \) is a polyhedron:
   A EDGE of \( W \) is a face of dimension 1
   A FACET of \( W \) is a face of dimension \( \text{dim}(W) - 1 \)
   A VERTEX of \( W \) is a face of dimension 0.
   ("\( \text{dim}(\emptyset) = -1 \)"")
Basic feasible solution

Def: Let \( X \) be the set of feasible solutions to \( \{ Ax = b, x \geq 0 \} \).

\( x \in X \) if \( x \) is a basic feasible solution (BFS).

The support of \( x \) is \( J_x = \{ j : x_j > 0 \} \). The set of columns of \( A \) used by \( x \).

The \( x \) is a basic feasible solution (BFS) if the linear program \( \min \{ c^T x : Ax = b, x \geq 0 \} \) has only the set of columns of \( A \) used by \( x \) is linearly independent set.

Eg: Consider the constraints
\[
\begin{align*}
  x_1 - x_3 + 3x_4 - x_5 &= 1 \\
  x_2 + x_3 + 4x_4 + 2x_5 &= 4 \\
  -3x_4 + 3x_5 &= 0
\end{align*}
\]

Let \( x^T = (2, 3, 1, 0, 0) \), so \( x \) is not a BFS.

The \( J_x = \{ 1, 2, 3 \} \).

Columns of \( A \) used by \( x \) are \( [1, 0, 0, 0, 0], [0, 1, 0, 0, 0], [-1, 0, 0, 0, 0] \).

Not linearly independent, so \( x \) is not a BFS.

Consider \( x^T = (1, 4, 0, 0, 0) \), so \( J_x = \{ 1, 4 \} \).

New columns of \( A \) used by \( x \) are \( [0, 1, 0, 0, 0], [0, 0, 0, 0, 0] \).

Linearly independent, so \( x \) is a BFS.

Consider \( x^T = (0, 4, 0, 3, 8) \). The \( J_x = \{ 2, 4, 5 \} \).

New columns of \( A \) used by \( x \) are \( [0, 0, 0, 0, 0], [3, 4, 5, 0, 0], [-4, 0, 0, 0, 0] \).

Linearly independent, so \( x \) is BFS.
Theorem Let \( \bar{x} = \{ x : Ax = b, x \geq 0 \} \). \( \bar{x} \) is a extreme point of \( \bar{x} \) if and only if \( \bar{x} \) is a BFS.

Proof = By contradiction.

Assume \( \bar{x} \) is not a BFS.

Then, the column of \( A \) used by \( \bar{x} \) are linearly dependent, so there exists \( y \neq 0 \) with support \( (y) \leq \text{support} (\bar{x}) \) and \( Ay = 0 \).

Let \( \tilde{x} = \max \{ x_j : x_j \geq 0 \} \). Let \( \bar{x} = \max \{ x_j \} \).

Then \( \bar{x} = \frac{\tilde{x}}{x} \) is feasible, so \( \bar{x} \) is not extreme.

We conclude \( \bar{x} \) is a BFS. Let \( J = \text{support} (\bar{x}) \).

Assume \( \bar{x} \) is an extreme point.

So, \( \bar{x} = \lambda \bar{x} + (1-\lambda) \bar{x} \) for some \( \bar{x}, \bar{x} \neq \bar{x}, \lambda > 0 < 1, \lambda + \lambda = 1 \).

\( \bar{x} \neq \bar{x} \) for all \( j \neq J \). Since \( \bar{x}_j > 0 \) and \( \bar{x}_j = 0 \), must have \( \bar{x}_j = 0 \).

So, \( \text{support} (\bar{x}) \leq J \), \( \text{support} (\bar{x}) \leq J \).

Now, \( A \bar{x} = Ax = b \), so \( A(\bar{x} - \bar{x}) = 0 \).

Since \( \text{support} (\bar{x} - \bar{x}) \leq J \) and \( \bar{x} - \bar{x} \neq 0 \), columns used by \( \bar{x} \) are linearly dependent.

So, \( \bar{x} \) is not a BFS.
Theorem (P) has a finite number of BFSs.

**Part 1.** Number of BFS is \( \leq (n) \).

Theorem 1. (P) has a feasible solution, it is a BFS.

**Proof.**

Case 1: If \( b = 0 \), then \( x = 0 \) is a BFS (in fact, \( x \in \text{OPT} \)).

Case 2: Let \( b \neq 0 \). Let \( x \) be a feasible solution, \( x \neq \text{BFS} \).

Let \( J_x \) be the support of \( x \).

Can find a \( y \) with \( \text{support}(y) \subseteq J_x \).

\( A^T y = 0 \).

Thus, \( y \) has a non-zero component.

Now, \( x + \theta y \) satisfies \( A(x + \theta y) \geq 0 \) for any \( \theta \).

Let \( \theta = \min \left\{ \frac{x_j}{y_j} : x_j > 0, \ y_j < 0 \right\} \).

Then \( x = x + \theta y \) is feasible, and \( \text{support}(x) \neq J_x \).

Continue in this way until find a point with \( \text{support} \) which is a linearly independent set of columns.
Theorem 1: If \( \{P\} \) has an optimal solution, then \( \{P\} \) has an optimal BFS.

Proof: Similar to previous theorem. □

Theorem 2: The set of optimal feasible solutions \( \{P\} \) is a face of the set of feasible solutions.

Proof: Let \( x \) be the set of feasible solutions. If \( k \neq 0 \), then:
- Objective function is unbounded, set of optimal solutions is empty.
- Otherwise, let \( z \) be the optimal value.

Consider \( H = \{x : c^Tx = z\} \).

If \( x \) is feasible, \( c^Tx \geq z \), so \( H \) is a supporting hyperplane.
Thus \( \{P\} \cap H \) is a face of \( H \), and \( \{P\} \) is the set of optimal solutions. □
Picking columns to give basic feasible solutions.

A is m x n. Choose any m columns of A that are linearly independent. We can assume they are the first m columns of A. Denote these columns by B, the other columns by N, so 
\[ A = [B \mid N]. \]

So problem (P):

\[ \text{min } c_B^T x_B + c_N^T x_N \]
\[ \text{s.t. } Bx_B + Nx_N = b \]
\[ x_B, x_N \geq 0. \]

Since B is invertible, constraints are equivalent to
\[ x_B + B^{-1}N x_N = b \]
\[ x_B, x_N \geq 0. \]

So an solution is
\[ x_B = B^{-1}b, \quad x_N = 0. \]

This is the basic solution corresponding to B.

BFS provided \( x_B > 0 \).

Objective function value = \( c_B^T x_B = c_B^T B^{-1}b. \)

\[ \text{min } 2x_1 + 2x_2 - x_3 + 3x_4 \]
\[ \text{s.t. } \begin{align*}
  x_1 + 2x_2 - x_3 &+ 2x_4 = 2 \\
  x_2 + x_3 &+ 2x_4 = 1 \\
  x_i \geq 0.
\end{align*} \]

Take first 2 columns of A: 
\[ B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} -1 & -2 \\ 1 & 0 \end{bmatrix} \]

\[ x_B = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad x_N = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} \]

\[ c_B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad c_N = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \]

\[ B^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}, \quad \text{so } B^{-1}b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \]

\[ x_B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \]

So get bfs \( x_1 = 0, x_2 = 1, (x_3 = 0, x_4 = 0) \), Value: \[ c_B^T B^{-1}b = 2 \]
Take column 1 & 3 of A: \( B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \quad N = \begin{bmatrix} 2 & -2 \\ 1 & 0 \end{bmatrix} \)

\[
x_8 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad x_N = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} \quad c_8 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad c_N = \begin{bmatrix} 2 \\ -1 \end{bmatrix}^T
\]

\[
B^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{so} \quad B^{-1}b = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \geq 0
\]

So get basic solution \( x_1 = 3, x_3 = 1, (x_2 = 0, x_4 = 0) \).

Value: \( \begin{bmatrix} 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \end{bmatrix} = 5 \).

\[
\sqrt{\text{Take column 3 \& 4 of A: } \quad B = \begin{bmatrix} -1 & -2 \\ 1 & 0 \end{bmatrix} \quad N = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \}
\]

\[
x_8 = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} \quad x_N = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} \quad c_8 = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \quad c_N = \begin{bmatrix} 2 \\ 3 \end{bmatrix}
\]

\[
B^{-1} = \frac{1}{2} \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}, \quad \text{so} \quad B^{-1}b = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}
\]

So get basic solution \( x_3 = 2, x_4 = -3, (x_1 = 0, x_2 = 0) \), which is not admissible.

\[
\sqrt{\text{Take column 1 \& 4 of A: } \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \quad N \text{ is basic} \}
\]

\[
\sqrt{\text{Take column 2 \& 3 of A: } \quad B = \begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix} \quad N = \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \}
\]

\[
x_8 = \begin{bmatrix} x_3 \\ x_6 \end{bmatrix} \quad x_N = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} \quad c_8 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad c_N = \begin{bmatrix} 2 \\ 3 \end{bmatrix}
\]

\[
B^{-1} = \frac{1}{3} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \text{so} \quad B^{-1}b = \frac{1}{3} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \geq 0
\]

So get \( \frac{1}{3} \quad x_3 = 0, x_4 = 1, (x_1 = 0, x_2 = 0) \)

Value: \( \begin{bmatrix} -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2 \).
Notice that the basic using column 1 or 2 and using column 3 or 2 give the same BFS. This is an example of a Degenerate BFS.

Let \( \hat{x} \) be a BFS of (P). \( \hat{x} \) is an Degenerate BFS of (P) if exactly \( m \) of the components of \( \hat{x} \) are positive. If less than \( m \) of the component of \( \hat{x} \) are positive, \( \hat{x} \) is a Degenerate BFS.

In the example, let \( \hat{x} = [2, 0, 0, 0]^T \).

Notice that \( A\hat{x} = [0] \), and \( \hat{x} \geq 0 \).

It follows that if \( \hat{x} \) is feasible then \( \hat{x} + \theta \hat{x} \) is feasible for any \( \theta \geq 0 \).

Def. Let \( \hat{x} \in \mathbb{R}^n \), \( \hat{x} \neq 0 \). The ray generated by \( \hat{x} \) is the set 
\[
\{ x : x = \lambda \hat{x}, \lambda \geq 0 \} 
\]

If \( \hat{x} \in \mathbb{R}^n \), the set \( \{ x : x = \hat{x} + \theta \hat{x}, \theta \geq 0 \} \) is the half-line through \( \hat{x} \) parallel to the ray generated by \( \hat{x} \).

Def. A Homogeneous Solution corresponding to (P) is a vector \( \hat{y} \) satisfying
\[
A \hat{y} = 0,
\]
\[
\hat{y} \geq 0.
\]

If \( \hat{x} \) is feasible for (P) and \( \hat{y} \) is homogeneous solution then
\( \hat{x} + \theta \hat{y} \) is feasible for (P) for all \( \theta \geq 0 \).
Finding the dimension of a polyhedron

\[ \mathcal{V} = \{ x : Ax = b, x \geq 0 \} \]

\( A \) is an \( m \times n \) matrix, \( b \) is a vector, \( x \) is an \( n \)-vector.

Now, from linear algebra, \( \dim \{ x : Ax = b \} \leq n - \text{rank}(A) \).

So get upper bound on \( \dim \{ \mathcal{V} \} \).

How do we get a lower bound?

Find a point in \( \mathcal{V} \) and find some linearly independent directions out of \( \mathcal{V} \) which keep you in \( \mathcal{V} \).

\[
\begin{align*}
x_1 + x_2 - x_3 &= 1 \\
3x_1 + x_3 - 4x_1 &= 2 \\
6x_1 - x_3 + 2x_2 &= 2 \\
x_i &\geq 0.
\end{align*}
\]

The \( x = [1, 0, 0, 0, 0, 0]^T \) is feasible.

Can we find two linearly independent directions out of \( \mathcal{V} \) which keep us feasible?

Two other possible points \( x = [0, 1, 1, 2, 0, 0]^T \), \( x = [1, 1, 1, 1, 1]^T \).

There are two degrees of freedom. So \( \dim \{ \mathcal{V} \} \geq 2 \).
This may not be enough to pin down the dimension.

Eg

\[ U = \{ x \in \mathbb{R}^3 : x_1 + x_2 = 0, \ x_2 \geq 0 \} \]

Then \( A = [0 \ 1 \ 0] \), so \( \text{rank}(A) = 1 \), so \( \text{dim}(U) \leq 1 \).

But \( U = \{ [x \ 0 \ 0] \} \), so \( \text{dim}(U) = 0 \).

What do we need in addition?

\( x_1 + x_2 = 0, \ x_2 \geq 0 \Rightarrow \text{must have} \ x_1 = 0 \text{ and} \ x_2 = 0 \).

So really, equality constraints for this problem are

\[ x_1 = 0, \ x_2 = 0 \]

Rank of this constraint matrix is 2.

\[ \text{dim}(U) = 2 - 2 = 0 \]

**Theorem** Let \( U \) be the feasible region of (P). If (P) has a nondegenerate bfs, then \( \text{dim}(U) = n - m \).

\( \text{dim}(U) \leq n - m \): (P) has a nondegenerate bfs, so there are \( n \) basic variables, with values 0.

Consider an extreme point of \( U \) with \( x = [x_1, \ldots, x_n, 0, \ldots, 0] \) where \( x_1 > 0, \ldots, x_k > 0 \)

\( \text{Since columns} \ 1, \ldots, k \ \text{of} \ A \ \text{form a basis for} \ M_n^k \),

\( \text{column} \ (m+1), \ldots, \text{say} \) \( y \) is a linear combination of these columns.

So we can find a vector \( y = [y_1, \ldots, y_n, 1, 0, \ldots, 0] \)

with \( Ay = 0 \).

For small enough \( \delta > 0 \),

\( \delta x + \delta y > 0 \).

So get non-zero direction of the form:

\[ \begin{bmatrix} y_1^* & y_2^* & \ldots & y_n^* \\ y_1 & y_2 & \ldots & y_n \\ 0 & 0 & \ldots & 0 \end{bmatrix} \]

So \( \text{dim}(U) \geq n - m \).
All deg, dim < n-1?

No: e

\[ X_1 - x_1 = 0 \]
\[ x_1, x_2 > 0. \]

Also:

If there exists

Allow combination of \( a_1, \ldots, a_n \) \( \Sigma a_i \) where \( \Sigma a_i = 1 \).

Allow spaces as you like, but do not off-set,

\[ \lambda_1, \ldots, \lambda_n \in \mathbb{R} \]
\[ \Sigma \lambda_i = 1 \]
\[ \Sigma a_i \lambda_i = L. \]
\[ x_1 + x_2 \leq 1 \]
\[ x_1 + x_3 = 1 \]
\[ x_1 + x_4 = 1 \]
\[ x_1 + x_5 \geq 1 \]
\[ x_1 + x_6 = 0 \]

\[ x_1 \geq 0 \]

\[ \dim = 2 = n - m \]

Even though all vertices are degenerate.

Vertices:
- A: \((0, 1, 0, 1, 0, 1)\)
- B: \((0, 0, 1, 1, 1, 0)\)
- C: \((1, 0, 0, 1, 0, 1)\)
Consumer has food of type \( j = 1, \ldots, n \) available on market at price \( c_j \) per unit.

Consumer has nutritional requirements for nutrient \( i = 1, \ldots, m \) (vitamin, mineral, etc.) required daily.

\( \alpha_{ij} \) = amount of nutrient \( i \) contained in one unit of foodstuff \( j \).

Solve:

\[
\begin{align*}
\max & \quad \sum_{j=1}^{n} c_j x_j \\
\text{subject to} & \quad \sum_{j=1}^{n} \alpha_{ij} x_j \geq b_i, \quad \forall i, \quad x_j \geq 0, \quad x_j \geq 0.
\end{align*}
\]

Find purchase amount \( x_j \) to minimize cost and meet nutritional requirements.

Related problem for producer of nutrient pills.

Produce pills of type \( i = 1, \ldots, m \).
Set unit price \( y_i \) for pills so that consumers buy pills.

\( y_i \geq 0 \), \( i = 1, \ldots, m \).

Consider \( j \)th food type.

Consumer has \( i \) options:
1. purchase one unit at cost \( c_j \), giving \( x_j \) units
2. purchase the corresponding nutrient from producer at cost

\( y_i \alpha_{ij} \).

To make pill price attractive, producer chooses \( y_i \) so that

\( \sum y_i \alpha_{ij} \leq c_j \)

Do this for all \( j \) to limit purchase to pills only.
So, in order to maximize profit from pills, we can write:

\[ \max \sum b_i y_i \]

subject to:

\[ \sum a_{ij} y_i \leq c_j \quad 1 \leq j \leq n \]

\[ y_i \geq 0 \quad 1 \leq i \leq m. \]

i.e.,

\[ \text{max } b^T y \text{ s.t. } A^T y \leq c \text{ } y \geq 0. \]

Notice that the producer receives \( b^T y \) from consumers.

So, we would like \( b^T y \) to be just less than \( c^T x \).

If \( \max b^T y = \max c^T x \), consumer will be indifferent.

In fact, at solution, optimal values agree.

In addition, optimal solution will have \( y = 0 \) if consumer cannot satisfy requirement for nutrient \( f_i \).

Can get a dual for any problem:

\[ \begin{align*}
\text{min } & \quad c^T x + d^T v + \rho^T u \\
\text{s.t. } & \quad A^T x + D^T v + \rho^T u = b \\
& \quad F^T x + G^T v + Hu \geq 0 \\
& \quad K^T x + LV + Pu \leq 0 \\
& \quad x \geq 0, \quad v \leq 0, \quad u \leq 0 \end{align*} \]

\[ \begin{align*}
\text{max } & \quad b^T y + g^T z + \phi^T w \\
\text{s.t. } & \quad A^T y + F^T z + k^T w \leq d \\
& \quad D^T y + G^T z + L^T w \geq d \\
& \quad C^T y + H^T z + P^T w = d \\
& \quad z \geq 0, \quad w \leq 0 \end{align*} \]
General duality correspondence

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

\[
\begin{align*}
\text{min} & \quad b^T y \\
\text{subject to} & \quad A^T y \leq c
\end{align*}
\]
Another derivation of duality (Chvátal)

\[ \min z = 3x_1 + 5x_2 + 2x_3 \]

s.t. \[ 2x_1 + 3x_2 + x_3 = \#3 \]
\[ x_1 - x_2 + 3x_3 = \#4 \]
\[ x_i \geq 0. \]

Can we get a lower bound on the optimal value?

Now, \[ \frac{3}{2}(2x_1 + 3x_2 + x_3) = 3x_1 + \frac{9}{2}x_2 + \frac{3}{2}x_3 \leq 3x_1 + 5x_2 + 2x_3 \text{ since } x_3 \geq 0 \]

So \[ z \geq \frac{3}{2}(2x_1 + 3x_2 + x_3) = \frac{9}{2}. \]

Can we bound the constraints:
\[ y_1(2x_1 + 3x_2 + x_3) - y_2(x_1 - x_2 + 3x_3) = 3y_1 + 4y_2 \]
for any feasible \( x \), for any chain of \( y \).

Thus, \[ (2y_1 + y_2)x_1 + (3y_1 - y_2)x_2 + (y_1 + 3y_2)x_3 = 3y_1 + 4y_2 \]

Gives us a lower bound on \( z^* \) of \( 3y_1 + 4y_2 \), provided:
\[ 2y_1 + y_2 \leq 3 \]
\[ 3y_1 - y_2 \leq 5 \]
\[ y_1 + 3y_2 \leq 2 \]
To get best possible lower bound:

Try to maximize $3y_1 + 4y_c$.

Get dual problem:

$$\begin{align*}
\text{max} & \quad 3y_1 + 4y_c \\
\text{st.} & \quad 2y_1 + y_c \leq 3 \\
& \quad 3y_1 - y_c \leq 5 \\
& \quad y_1 + 3y_c \leq 2
\end{align*}$$

Optimal dual solution: $y_1 = 1.4$, $y_c = \frac{1}{3}$, value $= 5$.

So $x = (1, 0, 1)$ is optimal.
Weak Duality Theorem

If \( x \) is primal feasible and \( y \) is dual feasible then \( c^T x \geq b^T y \).

\( \begin{align*}
\text{If } & \quad (P) \text{-feasible and } (D) \text{-feasible, and } c^T x \geq b^T y, \\
& \quad \text{then } (P) \text{-optimal implies } (D) \text{-optimal.}
\end{align*} \)

Proof (by contradiction):

Assume \( (P) \) is feasible and \( (D) \) is feasible, and \( c^T x \geq b^T y \) for some \( x \) and \( y \).

Then, there exists \( x \) and \( y \) such that \( c^T x \geq b^T y \) by hypothesis.

Provide a counterexample of optimality:

If someone gives you an \( x \) and \( y \) and claims they are optimal for \( (P) \) and \( (D) \) respectively, you do as check:

(i) \( x \) is \( (P) \)-feasible,
(ii) \( y \) is \( (D) \)-feasible,
(iii) \( c^T x \geq b^T y \).

Strong Duality Theorem

Suppose \( (P) \) is feas., and \( (P) \) has finite optimal value. Then \( (D) \) is feasible and there exist solutions \( x^*, y^* \) feasible for \( (P), (D) \) with \( c^T x^* = b^T y^* \).

Proof: Since \( (P) \) has an optimal solution, \( (D) \) also has an optimal solution (by proof of weak).
By linear algebra exercise, if \((P)\) has an optimal solution \(x^*\), we can assume \(x^*\) were a subset of the \(B\) columns, and these columns are linearly independent. Call these columns \(B\).

So have \[
\begin{align*}
&\text{max } b^T y \\
\text{s.t. } &b_x^T x + N x_N = b \\
&x_B, x_N \geq 0
\end{align*}
\]

No define \(y^* = B^{-T} c_B\).

Then \(b^T y^* = c_B\).

Moreover \(N^T y^* = N^T B^{-T} c_B\).

Now \((P)\) is equivalent to

\[
\begin{align*}
&\text{max } c_B^T x_B + c_N^T x_N \\
\text{s.t. } &x_B + B^{-1} N x_N = b \\
x_B, x_N \geq 0.
\end{align*}
\]
Strong Duality Theorem

The following are mutually exclusive and exhaustive possibilities for (P) and (D):

1. (P) and (D) infeasible
2. (P) feas., (D) infeasible (and unbounded in obj. value)
3. (D) feas., (P) infeasible (and unbounded in obj. value)
4. (P), (D) feas. and nos. same.

Proof from strong duality theorem:

(P) \text{ unbounded} \implies (D) \text{ infeas., } (D) \text{ infeas.} \implies (P) \text{ unbounded}

From strong duality theorem:

(P) \text{ unbounded} \implies (D) \text{ feas., } (D) \text{ of same optimal value}

Simplifying Slackness

A pair of primal and dual feasible solution are optimal for the respective problem in a primal-dual pair of LPs iff, whenever these variables make a slack variable in one problem strictly positive, the value of the associated non-negative variable in the other is zero.

\text{min } c^T x = \text{max } b^T y \\
A x = b \quad A^T y \leq c \\
x \geq 0

Example: \quad c_i < c_j \implies x_i = 0 \quad (c_i = 10 \text{th column of } A)
There are two cases:

\[ \hat{x}, \hat{y} \text{ optimal } \iff (c_i - a_i \hat{y}) < 0 \Rightarrow x = 0 \quad \text{and} \quad (c_i = \text{ i-th column of } A). \]

1. \( \hat{x}, \hat{y} \) optimal \( \iff \hat{x} \cdot (c_i - a_i \hat{y}) = 0 \quad \text{by } \text{ Feasibility}. \)

\( \Rightarrow \) Know \( c_i - a_i \hat{y} \geq 0, \hat{x} \geq 0 \) by Feasibility,

\( \Rightarrow \hat{x} \cdot (c_i - a_i \hat{y}) = 0 \quad \text{by } \text{ Feasibility}. \)

\( \Rightarrow 0 = \hat{x} \cdot (c_i - a_i \hat{y}) = \hat{x}^T(c_i - a_i \hat{y}) = \hat{x}^T A \hat{y} = c^T \hat{x} - b^T \hat{y} = 0 \quad \text{since } \hat{x}, \hat{y} \text{ optimal pair}. \)

\( \Rightarrow \hat{x} \cdot (c_i - a_i \hat{y}) = 0 \quad \text{by } \text{ Feasibility} \Rightarrow 0 = c^T \hat{x} - b^T \hat{y} \Rightarrow c, y \text{ optimal.} \]

**Simplex:** (P)-feasible \( \iff \text{C.S.} \), work to (D). Feasible

**Interior:** (P)-feasible, (D) Feasible, work to C.S.

Exactly one of the following systems is feasible:

1. \( \begin{align*}
Ax & = b \\
x & \geq 0
\end{align*} \tag{I} \)

2. \( \begin{align*}
A^Ty & \leq 0 \\
b^Ty & > 0
\end{align*} \tag{II} \)

**Proof:** Consider dual pair of LPS:

\[ \begin{align*}
\text{min} & \quad c^T x \\
\text{ subject to } & \quad Ax - b \geq 0, \quad x \geq 0 \quad \text{(P)} \tag{D}
\end{align*} \]

(I) holds \( \iff \) (P) feasible \( \iff \) (P) has optimal value \( 0 \Rightarrow \) (D) L, optimal value \( 0 \Rightarrow \) (II) does not hold.

(I) does not hold \( \iff \) (P) infeasible \( \iff \) (P) infeasible \( \iff \) (D) unbounded or infeasible

Now \( y = 0 \) always feasible for (D), \( \Rightarrow \) (D) unbounded \( \iff \) \( (I) \) holds

\( \text{Note: not enough to show } (I) \iff \text{not } (II), (II) \iff \text{not } (I). \)
General illustration:

\[ A = \begin{pmatrix} 1 & 3 & 3 \\ 3 & 2 & 1 \end{pmatrix} \]

\[ C = \{ x \mid x = Ax, x \geq 0 \} \]

\[ \{ y : A^T y < 0 \} = \mathbb{R}^3 \]

\[ H_1 : \{ y : A^T y = 0 \} \]

\[ H_3 \]

If \( b \in C \) then (I) holds, (II) does not hold.

If \( b \not\in C \), then (I) does not hold, and we must have \( b \in C^0 \) so (II) holds.
Simplex Algorithm

Standard form LP

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

Where \( A \) is \( m \times n \), \( c \) is \( n \)-vector, \( b \) is \( m \)-vector.

Assume WLOG \( A \) has rank \( m \).

Pick any \( m \) linearly independent columns of \( A \). WLOG, can assume they are the first \( m \) in columns. Denote these columns by \( B \), denote remaining columns by \( N \). Then \( A = [B|N] \).

The (P) becomes:

\[
\begin{align*}
\min & \quad c_B^T x_B + c_N^T x_N \\
\text{subject to} & \quad B x_B + N x_N = b \\
& \quad x_B, x_N \geq 0
\end{align*}
\]

Since \( B \) is invertible, constraints are equivalent to

\[
B x_B + N x_N = b \Rightarrow B^{-1} B x_B + B^{-1} N x_N = B^{-1} b
\]

or

\[
x_B = B^{-1} b - B^{-1} N x_N.
\]

So, given \( x_N \), one can get a solution to \( A x = b \) by taking

\[
x_B = B^{-1} b - B^{-1} N x_N.
\]

E.g., if \( x_N = 0 \), then \( x_B = B^{-1} b \).

This is the BASIC solution corresponding to \( B \).

It is a BFS if \( 0^{-1} b \geq 0 \).

Objective function value is \( c_B^T x_B = c_B^T B^{-1} b \).
Substitute $x_B = B^{-1}b - B^{-1}Nx_N$ into objective function:

\[ f(x) = c^T x \]

\[ = c^T B^{-1}b + (c_N^T - c_B^T B^{-1}N)x_N \]

\[ = x_N \quad x_B + B^{-1}N x_N = B^{-1}b \]

\[ x_B, x_N \geq 0. \]

**Def:** The vector of reduced costs corresponding to the basis $B$ is the vector $c_N^T - c_B^T B^{-1}N$.

**Basic variable:** variable for which corresponding column of $A$ belongs to $B$.

**Nonbasic variable:** remaining variable.

**Theorem:** If the vector of reduced costs is nonnegative, then the point $x_B = B^{-1}b, x_N = 0$ is optimal, provided $B^{-1}b \geq 0$ (i.e., provided $B^{-1}$ is feasible).

**Proof:** For $x_B = B^{-1}b, x_N = 0$, we have objective function value $z = c_B^T B^{-1}b$.

For any other $x$, we have objective function value

\[ z = c_B^T B^{-1}b + \sum_{x_N} c_N^T x_N < c_B^T B^{-1}b. \]
Theorem: If \( x_B = B^{-1}b, x_N = 0 \) is primal optimal, then
\[
y = B^{-1}c_B \text{ is dual optimal.}
\]
Proof: Dual is max \( b^T y \) subject to \( A^T y \leq c \) and \( N^T y = c_N \).

Theorem: If some \( z_k \) is negative, then the point \( x_B = B^{-1}b, x_N = 0 \)
1) is not optimal, provided \( B^{-1}b > 0 \).

Proof: Key: set \( x_j = x_j \) for all \( j \in \mathbb{R} \setminus \{k\} \), at zero. Increase \( x_k \) from zero ("pivot on \( x_k \)).")

Compute the equation using \( x_B \), since \( B \) has full rank.

New obj fn value is
\[
z = c_B^T B^{-1}b + \sum_{j \neq k} c_j x_j = c_B^T B^{-1}b + \bar{c}_k x_k < c_B^T B^{-1}b.
\]

Handout Now:
Assume \( z_k < 0 \). How large can \( x_k \) become?

Note: \( x_B = B^{-1}b - B^{-1}N x_N \), need \( x_B \geq 0 \).

Only entry in \( x_N \) which is negative is \( x_k \), so

\[
x_D = B^{-1}b - B^{-1}a_k x_k
\]

Hence
\[
\begin{pmatrix}
X_B \\
X_D
\end{pmatrix}
= \begin{pmatrix}
b_I \\
b_m
\end{pmatrix} - \begin{pmatrix}
\bar{a}_{1k} \\
\bar{a}_{mk}
\end{pmatrix} x_k
\]

\( \bar{b} = B^{-1}b \)
\( \bar{a}_{k} = B^{-1}a_k = \bar{a}_k \)
\( x_B = (x_{B1}, \ldots, x_{Bm}) \)

Need to choose \( x_k \) so negative \( x_B \geq 0 \).
(i) \( \bar{a}_i x_i \leq 0 \quad \forall i, \quad 1 \leq i \leq m. \)

Now, \( \bar{x}_0 = \bar{b} - \bar{a}^T \bar{x}_u \geq \bar{b} \geq 0. \)

So, we decrease \( x_u \) arbitrarily.

So, obj value \( z = c^T \bar{B}^{-1} b - \bar{a}^T \bar{x}_u \) is unbounded below.

Get \( \bar{x}_u : \bar{x}_B = -\bar{a}^T, \quad x_u = 1, \quad x_j = 0 \quad j \in \mathbb{R} \setminus \{u\}. \)

Ex: \[ \begin{align*}
-3x_1 + 2x_2 \\
-4x_1 + x_3 \\
-2x_2 - x_3 + x_4 \\
\end{align*} \]

\[ \begin{align*}
x_B &= \begin{bmatrix} x_1 \\ x_3 \end{bmatrix}, \\
\bar{B} &= \begin{bmatrix} 1 & 0 \end{bmatrix}, \\
\bar{b} &= \bar{b} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \\
\bar{a}^T &= \bar{B}^{-1} = \begin{bmatrix} 2 & 1 \end{bmatrix}. \]

So get \( \bar{x}_u : \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = -\bar{a}^T = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad x_2 = 1, \quad x_3 = 0 \)

So \( x^* = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 2 \end{bmatrix}. \)
(ii) \( a_i c_i > 0 \) for some \( i \):

\[
\begin{align*}
\min_{x_1, x_2, x_3, x_4} & \quad -3x_1 + 2x_3 \\
x_1 & \quad -x_2 + x_3 = 3 \\
10x_1 & \quad -x_2 + x_4 = 5 \\
x_1, x_2, x_3, x_4 & \geq 0.
\end{align*}
\]

Keep \( x_3 \) at \( x_3^* \). Increase \( x_2 \).

So
\[
\begin{align*}
x_1 & = 3 + x_2 \\
x_4 & = 5 - 10x_2.
\end{align*}
\]

So need \( x_2 \leq \frac{5}{10} \) in order to maintain \( x_4 \geq 0 \).

In general: If \( x_k = 0 \):
\[
x_i = \frac{b_i}{a_k} - \frac{a_i c_i}{a_k} \Delta \geq 0.
\]

So need \( \Delta \leq \frac{b_i}{a_k} \).

So we can \( \Delta = \min \left\{ \frac{b_i}{a_k} : \frac{b_i}{a_k} > 0 \right\} \).

Say minimum achieved by \( \frac{b_i}{a_k} \), so \( \Delta = \frac{b_r}{a_k} \).

So get \( x_r = \frac{b_r}{a_k} - b_k \frac{b_r}{a_k} = 0 \).

For \( i \in B \setminus \{r\} \), \( x_i = \frac{b_i - a_i c_i b_r}{a_k} \geq 0 \).

Column \( a_r \) replaces column \( \frac{a_k}{a_r} \) in basis.

New Basis:
\[
\begin{align*}
x_k & = \frac{b_r}{a_k} \\
x_{2k} & = \frac{b_k - a_k c_k}{a_k} x_i = b_c - a_i c_i b_r \frac{c_k}{a_k} \text{ while} \\
x_r = c_k \\
x_{j} & = 0 \quad \forall \ j \in R \setminus \{k\} \} \text{ Nonbasic}.
\end{align*}
\]
In example:

Take \( x_1 = \frac{1}{2} \), \( x_4 \) leave \( b > 0 \).
\( x_1 = 3 - (-1) \cdot \frac{1}{2} = 3 \frac{1}{2} \).
\( x_7 = x_4 = 0 \).

Now \( B = \begin{bmatrix} 1 & -1 \\ 0 & 10 \end{bmatrix} \quad B^{-1}b = \begin{bmatrix} \frac{3}{4} + \frac{1}{4} \\ \frac{1}{4} \end{bmatrix} \) (value of \( b \) on \( \mathbf{w} \),

\[ N = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \]

New reduced cost \( = \begin{bmatrix} \frac{3}{4} + \frac{1}{4} & -3 \end{bmatrix} \begin{bmatrix} 10 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \)
\[ = \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{16} & -3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \end{bmatrix} - \frac{1}{16} \begin{bmatrix} 3 & -3 \end{bmatrix} \]
\[ = \begin{bmatrix} \frac{17}{16} & \frac{3}{8} \end{bmatrix} \geq 0 \]

So new point is optimal.

Value: \( c_0^T B^{-1} b = \begin{bmatrix} 0 & -3 \end{bmatrix} \begin{bmatrix} \frac{3}{4} + \frac{1}{4} \end{bmatrix} = -\frac{3}{2} \)

Problem is equivalent to:

\[
\min -\frac{3}{2}
\]
\[\text{s.t. } \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} + B^{-1} \begin{bmatrix} -1 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} + \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}
\]
\[ x_1 \geq 0 \]
Notes

(1) The entering variable in Step 2c is not specified. Many different rules have been proposed. E.g:
(i) Take variable with largest reduced cost. (Dantzig's original negative gradient method)
(ii) Take variable whose ratio results in the largest increase in the objective function. (Dantzig's best improvement rule)
(iii) Steepest edge rule (Dover rule). Pick edge which makes most acute angle with C. (Cohon (1973), Goldberg & Reid (1977)) (see Cherno, p.115)

(i) is easiest to calculate. (ii) & (iii) take more work per iteration, but usually result in fewer iterations.

In large problems, expensive to calculate all reduced costs. So just calculate a subset and choose entering variable from that subset. This is called PARTIAL PRICING.

(2) Getting an initial solution.
Given a problem
\[ \begin{align*}
\text{min } \mathbf{c}^\top \mathbf{x} \\
\text{subject to } \mathbf{A} \mathbf{x} = \mathbf{b}, \; \mathbf{x} \geq \mathbf{0}
\end{align*} \]

problem
\[ \begin{align*}
\text{min } \mathbf{c}^\top \mathbf{s} \\
\text{subject to } \mathbf{A} \mathbf{x} + \mathbf{s} = \mathbf{b}, \; (\mathbf{b} \geq 0). \text{ Have initial bfs} \\
\mathbf{x}_0 = 0, \; \mathbf{s}_0 = \mathbf{b} \\
\text{for the problem, and solve Phase I problem gives} \\
\mathbf{x}_0 \leq 0, \; \mathbf{s}_0 = \mathbf{b} \text{ for original problem, provided it is feasible.}
\end{align*} \]

PHASE I brute force. Also, what does it mean algebraically to say that the feasible region for the dual problem is unbounded?
(3) Step (2b)(ii) involves pivoting a leaving variable if there is a tie in the minimum ratio test. In the next iteration, the basis variable with the smallest value will be set to zero, so one element of B⁻¹ b will be zero. This means the pivot in Step 2(b)(ii) at the next iteration may also be 0, so the next incoming variable could come in at value zero, and thus the basis will change but not the point. This would be a degenerate pivot. This may lead to CYCLING - i.e., going through a sequence of pivots only to return to the initial basis.

See handout (Note: need 37 eps, 36 variables to cycle) or non-optimum vertex.

Our remedy:
Bland's left index rule:
In Step 2(b), pick the smallest possible b
In Step 2(b)(ii), pick the smallest possible i.

Then (Bland, 1977) the simplex method does not cycle.

It follows that with this pivot rule, the simplex algorithm terminates either with an optimal solution in Step 2(a) or with a deg in Step 2(b)(i).

(4) Variables with upper bounds.

(5) Handling free variables
An Example of Cycling.

Entering rule: The variable with largest reduced cost enters.

Leaving rule: In the case of ties, the variable with the least index i leaves.

Consider the problem:

\[
\begin{align*}
0. \text{ min } & -53x_2 + 41x_3 + 20x_4 + 20x_5 \\
\text{s.t. } & x_1 - 11x_2 - 5x_3 + 18x_4 + 2x_5 = 0 \\
& 4x_2 + 2x_3 - 8x_4 - x_5 + x_6 = 0 \\
& 11x_2 + 5x_3 - 18x_4 - 2x_5 + x_7 = 0 \\
& x_i \geq 0
\end{align*}
\]

1. \( x_2 \) enters, \( x_6 \) leaves:

\[
\begin{align*}
\text{min } & -\frac{29}{2}x_3 + 98x_4 + \frac{27}{4}x_5 + \frac{53}{4}x_6 \\
& x_1 + \frac{1}{2}x_3 - 4x_4 - \frac{3}{4}x_5 + \frac{11}{4}x_6 = 0 \\
& \frac{1}{2}x_3 - 2x_4 - \frac{1}{4}x_5 + \frac{1}{4}x_6 = 0 \\
& -\frac{1}{2}x_3 + 4x_4 + 3\frac{6}{4}x_5 - \frac{3}{4}x_6 + x_7 = 1
\end{align*}
\]

2. \( x_3 \) enters, \( x_1 \) leaves (on tie break):

\[
\begin{align*}
\text{min } & 29x_1 - 18x_4 - 15x_5 + 93x_6 \\
& 2x_1 + x_3 - 8x_4 - \frac{3}{2}x_5 + \frac{11}{2}x_6 = 0 \\
& 2x_3 - \frac{1}{2}x_5 - \frac{5}{2}x_6 + x_7 = 0 \\
& x_7 = 1
\end{align*}
\]

3. \( x_4 \) enters, \( x_2 \) leaves.
3. $x_4$ enters, $x_2$ leaves:

$$\min \begin{array}{c}
20x_1 + 9x_2 \\
-2x_1 + 4x_2 + x_3 \\
-x_1 + \frac{1}{2}x_2 \\
x_1
\end{array} \begin{array}{c}
\frac{21}{2}x_5 + \frac{14}{2}x_6 \\
+ \frac{1}{2}x_5 - \frac{9}{2}x_6 \\
+ x_4 + 4x_5 - \frac{5}{2}x_6 \\
+ x_7 = 1
\end{array}$$

$x_i \geq 0$

4. $x_5$ enters, $x_3$ leaves (or he breaks):

$$\min \begin{array}{c}
-22x_1 + 93x_2 + 21x_3 \\
-4x_1 + 8x_2 + 2x_3 + x_4 \\
\frac{1}{2}x_1 + \frac{3}{2}x_2 + \frac{1}{2}x_3 + x_4 \\
x_1
\end{array} \begin{array}{c}
-24x_6 \\
x_5 - \frac{9}{2}x_6 \\
x_6 + x_7 = 1
\end{array}$$

$x_i \geq 0$

5. $x_6$ enters, $x_4$ leaves:

$$\min \begin{array}{c}
-10x_1 + 57x_2 + 9x_3 + 24x_4 \\
\frac{1}{2}x_1 - \frac{13}{2}x_2 - \frac{9}{2}x_3 + 9x_4 + x_5 + x_6 \\
x_1
\end{array} \begin{array}{c}
= 0 \\
= 0 \\
+ x_7 = 1
\end{array}$$

$x_i \geq 0$

6. $x_1$ enters, $x_6$ leaves (or he breaks):

$$\min \begin{array}{c}
-53x_1 - 61x_2 + 20x_3 + 20x_4 + 20x_5 \\
x_1 - 11x_2 - 5x_3 + 18x_4 + 2x_5 \\
\frac{11}{2}x_2 + 2x_3 - 8x_4 + x_6 \\
x_5 - 18x_4 - 2x_5
\end{array} \begin{array}{c}
= 0 \\
= 0 \\
+ x_7 = 1
\end{array}$$

$x_i \geq 0$.

So, we return to the initial basis.
Bland's Rule

Step 5 becomes:

So \( x_1 \) enters, \( x_4 \) leaves:

\[
\begin{align*}
\text{min} \quad & 27x_2 - x_3 + 44x_4 + 20x_6 \\
& -4x_2 - 2x_3 + 8x_4 + x_5 - x_6 = 0 \\
& x_1 - 3x_2 - x_3 + 2x_4 + 2x_6 = 0 \\
& 3x_2 + x_3 - 2x_4 - 2x_6 + x_7 = 1 \\
x_i & \geq 0
\end{align*}
\]

So, \( x_3 \) enters, \( x_7 \) leaves:

\[
\begin{align*}
\text{min} \quad & -1 + 36x_2 + 42x_4 + 18x_6 \\
& 2x_2 + 4x_4 + x_5 - 5x_6 + 2x_7 = 2 \\
& x_1 - 3x_2 + x_3 - 2x_4 + x_7 = 1 \\
x_i & \geq 0
\end{align*}
\]

So, finally, we must away from

\( x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = 0, x_7 = 1 \)

to

\( x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 0, x_5 = 2, x_6 = x_7 = 0 \)

which is optimal.
Pictures of feasible regions:

From "A simple iteration" handout:

\[ \begin{align*}
& \min \quad \Gamma = 5x_1 - 2x_4 \\
& \text{s.t.} \\
& \quad x_1 - x_4 = 3 \\
& \quad -3x_1 + 2x_4 = 2 \\
& \quad x_i \geq 0.
\end{align*} \]

Regard \( x_1, x_4 \)

as slack variables.

After one iteration:

\[ \begin{align*}
& \min \quad 3 + 6x_4 + x_7 \\
& \text{s.t.} \\
& \quad x_1 - 6x_2 - 0.5x_7 = 2 \\
& \quad x_4 - 1.5x_1 + 0.5x_7 = 1 \\
& \quad x_i \geq 0
\end{align*} \]
Eg: \[ x_1 + 2x_2 - 3x_3 = 6 \]
\[ x_i \geq 0. \]

If \( x_1 \) is basic:

If \( x_2 \) is basic:

If look along \( x_2 - ax_1 \):

Like looking from infinity for cut along \( x_2 \)-axis.
Dual solution from simplex:

\[ \min \ c_B^T x_B + c_N^T x_N \]
\[ \text{st. } B x_B + N x_N = b \]
\[ x_B, x_N \geq 0 \]

Assume \( B^T b \geq 0 \).

Dual:

\[ \max \ b^T y \]
\[ \text{st. } B^T y \leq c_B \]
\[ N^T y \leq c_N \]

For complementary slackness, want \( B^T y = c_B \), i.e., \( y = B^{-1} c_B \).

For dual feasibility, then need:

\[ N^T y \leq c_N, \text{ i.e., } N^T B^{-1} c_B \leq c_N \]
\[ \text{i.e., } c_N - c_B^T B^{-1} N \geq 0 \]

So dual feasibility is equivalent to nonnegative reduced costs.

Thus, simplex has:

(i) primal feasibility (once Phase I finished)
(ii) complementary slackness (\( y = B^{-1} c_B \))

It works towards dual feasibility.
Handling upper bounds:

\[ 0 \leq x \leq u. \]

Could introduce extra variables:

\[ \text{min } c^T x \]

\[ \text{st. } A x = b \]
\[ x + s = a \]
\[ x, s \geq 0. \]

Instead: modify pivoting rules:

A new basic variable can be either at its upper bound or at its lower bound.

If it is at its upper bound, it can enter if its reduced cost is positive.
- Variable will be decreased.

In a ratio test, make sure no variable exceeds its upper bound.
Eq: \[ m + 2 \quad -2x_2 - 3x_4 = 0 \]
\[ s.t. \quad x_1 + x_2 + x_3 = 4 \]
\[ x_2 - x_3 - 2x_4 = 1 \]
\[ 0 \leq x_1 \leq 10 \]
\[ 0 \leq x_2 \leq 4 \]
\[ 0 \leq x_3 \leq 5 \]
\[ 0 \leq x_4 \leq 1 \]

Have \( b_f \):

Nonbasics: \( x_3 = 0, x_4 = 1 \)

Basis: \( x_1, x_2 \)

Ob: \( f = 5x_1 - 5x_2 + x_3 - 2x_4 = -1 \)

Don't pivot on \( x_4 \) because it is at its upper bound

\( x_3 \) enters at value 1, \( x_2 \) leaves at its upper bound, 4;

\[ \min \quad -2x_2 + x_4 = -2 \]
\[ s.t. \quad x_1 + x_2 - x_4 = 5 \]
\[ -x_1 + x_2 + 2x_4 = 1 \]
\[ 0 \leq x_1 \leq 10 \]
\[ 0 \leq x_2 \leq 4 \]
\[ 0 \leq x_3 \leq 5 \]
\[ 0 \leq x_4 \leq 1 \]

Have \( b_f \):

Nonbasics: \( x_2 = 4, x_4 = 1 \)

Basis: \( x_1, x_3 \)

Ob: \( f = 2 - 2x_1 + 3x_3 = 7 \)

Notice that if we pivot on the second constraint,

\( x_4 \) will decrease in value by the most with value, 2.

So, we'll get \( x_4 = 1 - 2 = -1 \)

Not allowed. So, instead: \( x_4 \) goes to its lower bound.

Equation remains same.

New \( b_f \):

Nonbasics: \( x_2 = 4, x_4 = 0 \)

Basis: \( x_1, x_3 \)

Ob: \( f = 2 - 2x_1 + 3x_3 = 6 \)
After basis pivot
$x = (2, 4, 1, 1)$
$x_4 = 0$

$-x_2 + 2x_4 = -1$
$x_3 = 0$

$x_1 - x_4 = 5$
$x_1 = 0$

After $x_4$ basic
in lower bound
$x = (1, 4, 3, 0)$

$x_2 - x_4 = 5 - x_1 \begin{cases} \leq 5 \\ \geq -5 \end{cases}$

$-x_2 + 2x_4 = -1 - x_3 \begin{cases} \leq -1 \\ \geq -6 \end{cases}$
Modelling free variables.

Free variables are always basic. Do not split into \( x = u - v \).

Eq: \[
\begin{align*}
\text{min} & \quad -2x_3 \\
\text{s.t.} & \quad x_1 + x_2 + 3x_3 = 3 \\
& \quad 2x_1 + x_2 + 4x_3 = 2 \\
& \quad x_1, x_2, x_3 \geq 0
\end{align*}
\]

Col \( 1 \& 2 \) onmain \( 5 \) rel. \( B = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \) \( N = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \).

Then \( B^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \), so \( x_0 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = B^{-1} b = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \).

This is \( \geq 0 \), since it is feasible: \( x_1 = -1, x_2 = 4, (x_3 = 0) \).

\[
B^T N = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}
\]

5. equivalent problem is:

\[
\begin{align*}
\text{min} & \quad -2x_3 \\
\text{s.t.} & \quad x_1 + x_2 + 3x_3 = 3 \\
& \quad 2x_1 + x_2 + 4x_3 = 2 \\
& \quad x_1, x_2, x_3 \geq 0
\end{align*}
\]

\[
\begin{align*}
\text{min} & \quad 4 + x_2 \\
\text{s.t.} & \quad \frac{1}{2}x_1 + x_3 = 2 \\
& \quad x_1, x_2, x_3 \geq 0
\end{align*}
\]

Pivott:

\[
\begin{align*}
\text{min} & \quad -4 + x_2 \\
\text{s.t.} & \quad \frac{1}{2}x_1 + x_3 = 2 \\
& \quad x_1, x_2, x_3 \geq 0
\end{align*}
\]

Optimal: \( x_1 = 3, x_2 = 2, x_3 = 0 \),
so \( x_1 = -1 - x_3 = -3 \).
Dual rays:

Fischer tells us that (P) is infeasible $\iff y^T A_1 y < 0, b_1 y = 0$. A vector $y$ satisfying $A_1^T y < 0, y \geq 0$ is a Dual ray or Dual infeasible solution.

If such a ray exists and if (D) is feasible then the feasible region for (D) is unbounded. If this ray also satisfies $b_1 y > 0$ then (P) is infeasible.

Example:

\[ \begin{align*}
\text{min} & \quad -x_1 \\
\text{in} & \quad 0x_1 = 1 \quad (P) \\
& \quad x_1 \geq 0
\end{align*} \]

Primal ray: $x_1 = 1$. Note: $0x_1 = 0$, $x_1 \geq 0$, $x_1 \neq 0$. Also $-x_1 < 0$ so dual infeasible.

Dual ray: $y_1 = 1$. Note: $0y_1 = 0$, $y_1 \geq 0$. Also $y_1 > 0$ so primal infeasible.

Primal & dual rays provide certificate of infeasibility.

Example:

\[ \begin{align*}
\text{min} & \quad x_1 + 2v_1 = 3 \quad (P) \\
& \quad 3x_1 + 2v_1 = 2 \\
& \quad x_1 \geq 0
\end{align*} \]

Dual ray satisfies:

\[ \begin{align*}
y_1 + 3y_2 & \leq 0 \\
y_1 + 2y_2 & \leq 0 \\
(y_1 + 2y_2 > 0 \text{ for (P) infeasible.}
\end{align*} \]

Example:

\[ \begin{align*}
y_1 = 2, y_2 = -1. \quad \text{Note:} \quad y_1(x_1 + x_2) + y_2(3x_1 + 2v_1) = 2
\end{align*} \]

\[ \begin{align*}
\text{max} & \quad -x_1 + 0x_2 = 4, \quad \text{blatently infeasible.}
\end{align*} \]
Multiple optimal solutions & degeneracy

\[ \begin{align*}
\min & \quad c^T x \\
\text{st.} & \quad A x = b \\
& \quad x \geq 0
\end{align*} \]

Primal degeneracy corresponds to multiple dual optimal solutions. See next page, (61f).

Multiple optimal cases when some reduced cost is zero. Since one then increases the variable.

Graphical motivation for dual:

\[ \begin{align*}
\min & \quad c^T x \\
\text{st.} & \quad A x = b \\
& \quad x \geq 0
\end{align*} \]

Optimal if \( c \) is in here.

\[ \text{I.e., } A^T y = c, \quad y \geq 0, \quad y_i = 0 \text{, then } y_i = 0. \]
Why does primal degeneracy \( \Rightarrow \) dual multiple optima?

Chap. 19: If \( \lambda \) is degenerate?

\[
\begin{align*}
\max & \quad b^T y \\
\text{st.} & \quad B^T y \leq c_B \\
& \quad N^T y \leq c_N.
\end{align*}
\]

Let \( \bar{y} = B^T y \)

So equivalent problem:

\[
\begin{align*}
\max & \quad b^T B^{-T} \bar{y} = (B^{-T} b)^T \bar{y} \\
\text{st.} & \quad \bar{y} \leq c_B \\
& \quad N^T B^{-T} \bar{y} \leq c_N
\end{align*}
\]

So can change one component of \( \bar{y} \) (the one with \( (B^{-T} b)_i = 0 \)) and get new optimal value.

Can change provided all reduced costs \( > 0 \).