Summary: Simplex method for standard form

Start with $\mathcal{B}$ and $\mathcal{N}$ such that $B$ is nonsingular. Our initial vertex is $Bx_B = b$, $x_N = 0$.

Solve $B^T\pi = c_B$;

Compute $z_N = c_N - N^T\pi$;

$[z_N]_s = \min(z_N)$;

if $[z_N]_s \geq 0$ then stop;

Solve $Bp_B = -a_{\nu_s}$;

$\sigma_i = \begin{cases} 
\frac{[x_B]_i}{-[p_B]_i} & \text{if } [p_B]_i < 0; \\
-\infty & \text{if } [p_B]_i \geq 0;
\end{cases}$

$\sigma_t = \min\{\sigma_i\}; \quad \alpha = \sigma_t$;

if $\alpha = +\infty$ then stop;

Exchange index $\beta_t$ of $\mathcal{B}$ with index $\nu_s$ of $\mathcal{N}$;

$x_B \leftarrow x_B + \alpha p_B; \quad [x_B]_t \leftarrow \alpha;$
Getting Feasible for Standard Form
How do we get a feasible basic solution of $Ax = b$ (i.e. $x_B \geq 0$)?

Example:

$$A = \begin{pmatrix} 2 & 3 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

The basis $B = \{1, 2\}$ defines a feasible basic solution $x_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Can we find a feasible basis (i.e., vertex) in a systematic way?
\[2x_1 + 3x_2 + x_3 + x_4 = 5\]
\[x_1 + 2x_2 + x_3 - x_5 = 3\]
\[x_1, \ x_2, \ x_3, \ x_4, \ x_5 \geq 0\]

Main Idea:

- Add \( m = 2 \) new variables that define an “obvious” basis.
- Drive the new basic variables out of the basis using the simplex method.
Add positive shifts $x_6$ and $x_7$:

\[
\begin{align*}
2x_1 + 3x_2 + x_3 + x_4 + x_6 &= 5 \\
x_1 + 2x_2 + x_3 - x_5 + x_7 &= 3 \\
x_1, x_2, x_3, x_4, x_5, x_6, x_7 &\geq 0
\end{align*}
\]

The basis $B = \{6, 7\}$ defines the basic solution $x_B = \begin{pmatrix} x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$.

The nonbasic set is $N = \{1, 2, 3, 4, 5\}$ with $x_N = 0$. 
Minimize the “sum of infeasibilities”

\[
\begin{align*}
\text{minimize} \quad \mathbf{x} & \in \mathbb{R}^7 \\
2x_1 + 3x_2 + x_3 + x_4 & + x_6 = 5 \\
x_1 + 2x_2 + x_3 & - x_5 + x_7 = 3 \\
x_1, x_2, x_3, x_4, x_5, x_6, x_7 & \geq 0
\end{align*}
\]

- This is an LP in standard form (called the **phase-1 LP**)
- If a feasible point exists, both \(x_6\) and \(x_7\) will be zero (i.e., nonbasic) at the end of phase 1.
  - two other variables will be basic \(\Rightarrow\) phase 1 solution is an initial feasible basic solution for phase 2
A quick note...

If the first constraint had been of the form

\[ 2x_1 + 3x_2 + x_3 + x_4 = -5 \]

we would have included the shift \( x_6 \) as

\[ 2x_1 + 3x_2 + x_3 + x_4 - x_6 = -5 \quad \text{with} \quad x_6 \geq 0 \]

but we still use \(+ x_6\) in the sum of infeasibilities (the objective).
Define shifted constraints

\[ Ax + Dv = b \iff (A \ D) \begin{pmatrix} x \\ v \end{pmatrix} = b \]

with

\[ D = \begin{pmatrix} \text{sign}(b_1) \\ \text{sign}(b_2) \\ \vdots \\ \text{sign}(b_m) \end{pmatrix} \]

\[ \text{sign}(b_i) = \begin{cases} -1 & \text{if } b_i < 0 \\ 1 & \text{if } b_i \geq 0 \end{cases} \]

The initial basis is \( \mathcal{B} = \{ n + 1, \ n + 2, \ldots, \ n + m \} \).
In matrix notation, the phase-1 LP is:

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

with

\[
\bar{x} = \begin{pmatrix} x \\ v \end{pmatrix}, \quad \bar{c} = \begin{pmatrix} 0 \\ e \end{pmatrix}, \quad \bar{A} = \begin{pmatrix} A & D \end{pmatrix} \quad \text{and} \quad \bar{b} = b
\]

where \( e \) is the vector of ones.
Duality Theory
We have defined two forms of the simplex method.

These methods treat the two “generic” constraint formulations:

- **all-inequality form**: \( Ax \geq b \)
- **standard form**: \( Ax = b, \hspace{1em} x \geq 0 \)

Every LP can be converted into one of these two forms.

Which form should we use?

The best choice depends on the *shape* of the general constraint matrix.
Consider the following “base” linear program

\[
\begin{align*}
\text{minimize} & \quad d^T w \\
\text{subject to} & \quad Gw \geq f, \quad w \geq 0
\end{align*}
\]

where \( G \) is \( m \times n \).

This problem may be converted to either all-inequality form or standard form.
The base problem is

\[
\min_{w \in \mathbb{R}^n} d^T w \quad \text{subject to} \quad Gw \geq f, \quad w \geq 0
\]

In all-inequality form, \( c^T x \) is minimized subject to \( Ax \geq b \), where

\[
x = w, \quad c = d, \quad A = \begin{pmatrix} G \\ I_n \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} f \\ 0 \end{pmatrix}
\]

The simplex method for all-inequality form, requires the solution of two \( n \times n \) systems.

The systems involve \( n \) rows gathered from \( \begin{pmatrix} G \\ I_n \end{pmatrix} \).
Alternatively, we may convert the problem to *standard form*.

\[
\begin{align*}
\text{minimize} \quad & d^T w \\
\text{subject to} \quad & Gw \geq f, \quad w \geq 0,
\end{align*}
\]

Introduce \( m \) nonnegative slack variables \( s \) and define

\[
Gw - s = f \quad \text{and} \quad s \geq 0
\]

In this case, we minimize \( c^T x \) subject to \( Ax = b, \ x \geq 0 \) where

\[
\begin{align*}
x &= \begin{pmatrix} w \\ s \end{pmatrix}, \quad c &= \begin{pmatrix} d \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} G & -I_m \end{pmatrix} \quad \text{and} \quad b = f
\end{align*}
\]
The simplex method for constraints in standard form, requires the solution of two $m \times m$ systems.

These systems involve $m$ columns gathered from $(G - I_m)$.
The work associated with a simplex iteration is dominated by the cost of solving the two linear systems.

From the discussion above:

- Use the *all-inequality-form* simplex method if $m > n$, i.e.,

  \[ G = \begin{bmatrix} \end{bmatrix} \]

- Use the *standard-form* simplex method if $m < n$, i.e.,

  \[ G = \begin{bmatrix} \end{bmatrix} \]

If $m \approx n$, then both methods have comparable efficiency.
Example: If $G$ is $50 \times 1000$

- the standard-form basis is $50 \times 50$
- the all-inequality form working set is $1000 \times 1000$.

Example: If $G$ is $1000 \times 50$

- the all-inequality form working set is $50 \times 50$
- the standard-form basis is $1000 \times 1000$. 
An algorithm best suited to the shape of $G$ may not be available.
- e.g., your favorite code may only accept constraints in standard form

- *Duality theory* can turn a bad shape into a good shape
- Duality theory provides the link between all-inequality form and standard form
Problem conversion involves defining a \textit{primal linear program} and converting it into another \textit{dual linear program}.

First, we define the primal problem as an LP in all-inequality form:

\[
\begin{align*}
\text{minimize} & & c^T x \\
\text{subject to} & & Ax \geq b \\
& & x \in \mathbb{R}^n
\end{align*}
\]

where $A$ has $m$ rows.
Primal Problem in All-Inequality Form
First, we define the primal problem as an LP in all-inequality form:

\[
\begin{aligned}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax \geq b
\end{aligned}
\]

where \( A \) has \( m \) rows.

Assume for the moment that the problem is feasible with a bounded optimal value of \( c^T x^* \).
The optimality conditions (in active-set form) are:

\[ Ax^* \geq b \]

\[ c = A_a^T \lambda^*_a, \quad \lambda^*_a \geq 0 \]

Equivalently, in complementary slackness form:

\[ Ax^* \geq b \]

\[ c = A^T \lambda^*, \quad \lambda^* \geq 0 \]

\[ \lambda^* \cdot (Ax^* - b) = 0 \]

\( \lambda^* \) is the “scattered” \( \lambda^*_a \)
Suppose the Lagrange multipliers are viewed as the *dual variables*. The optimality conditions for our primal LP

\[ A^T \lambda^* = c \quad \text{and} \quad \lambda^* \geq 0 \]

can be viewed as *constraints on the dual variables*

\[ A^T \lambda = c \quad \text{and} \quad \lambda \geq 0. \]
• Any \( x \) that satisfies \( Ax \geq b \) is \textit{primal-feasible}, and

• Any \( \lambda \) that satisfies \( A^T\lambda = c \) and \( \lambda \geq 0 \) is \textit{dual-feasible}.

If we have such \( x \) and \( \lambda \), then

\[
\lambda^T(Ax - b) = \lambda^T (Ax - b) \geq 0
\]

Some rearrangement gives

\[
\lambda^T(Ax - b) = \lambda^T Ax - \lambda^T b = c^T x - b^T \lambda \geq 0.
\]

Thus, given any primal-feasible point \( x \)

\[
\implies c^T x \geq b^T \lambda.
\]

for all \( \lambda \) that are dual-feasible.
The *smallest value* of $c^T x$ is $c^T x^*$. It follows that

$$c^T x^* \geq b^T \lambda$$

for all $\lambda$ such that $A^T \lambda = c$, and $\lambda \geq 0$.

The *largest value* of $b^T \lambda$ is defined by the linear program

$$\begin{align*}
\text{maximize} & \quad b^T \lambda \\
\text{subject to} & \quad A^T \lambda = c, \quad \lambda \geq 0.
\end{align*}$$

This is called the *dual linear program*. 
The dual linear program is

\[
\begin{align*}
\text{maximize} & \quad b^T \lambda \\
\text{subject to} & \quad A^T \lambda = c, \quad \lambda \geq 0.
\end{align*}
\]

The constraints of this dual problem are in \textit{standard form}.

The dual constraint matrix is the \textit{transpose} of the primal constraint matrix.
For the moment we assume that the dual problem is feasible.

We have shown the following result.

**Result (Weak duality)**

If $x^*$ is a feasible bounded solution of the primal and $\lambda^*$ is a solution of the dual, then $c^T x^* \geq b^T \lambda^*$.

The property $c^T x^* \geq b^T \lambda^*$ is called **weak duality**, $c^T x^* - b^T \lambda^*$ is called the **duality gap**.

The obvious question is: does $c^T x^* = b^T \lambda^*$?
Result (Strong duality)

If $x^*$ is a feasible bounded solution of the primal, then there exists a feasible bounded solution $\lambda^*$ of the dual with $b^T\lambda^* = c^Tx^*$. 
Proof: Apply the all-inequality simplex method to the primal LP.

If Bland’s rule is used, the simplex method must either declare the problem unbounded, or terminate at a bounded solution $x^*$. By assumption, the primal LP is bounded, so the simplex method must terminate with a solution $x^* = x_\ell$ and working-set matrix $A_\ell$ and final working set be $\mathcal{W}_\ell = \{w_1, w_2, \ldots, w_n\}$ such that

$$A_\ell x_\ell = b_\ell, \quad A_\ell^T \lambda_\ell = c \quad \text{with} \quad \lambda_\ell \geq 0$$

We also know that $\lambda_\ell$ can be scattered to form $\lambda^*$ such that

$$(Ax^* - b) \cdot \lambda^* = 0, \quad \lambda^* \geq 0.$$
Thus, 

$$0 = (Ax^* - b)^T \lambda^* = (Ax^*)^T \lambda^* - b^T \lambda^* = c^T x^* - b^T \lambda^*$$

$$\implies c^T x^* = b^T \lambda^*$$

Notice that $b^T \lambda^* = c^T x^* \geq b^T \lambda$ for all dual feasible $\lambda$.

Thus, $\lambda^*$ must be optimal for the dual problem.

$$\implies c^T x^* = b^T \lambda^*, \text{ with } x^* \text{ and } \lambda^* \text{ the primal-dual solutions.} \quad \blacksquare$$
The primal and dual LPs are given by:

(P): minimize \( c^T x \)
subject to \( A x \geq b \)

(D): maximize \( b^T \lambda \)
subject to \( A^T \lambda = c, \ \lambda \geq 0 \)

If \((x^*, \lambda^*)\) is the solution to (P), then \(\lambda^*\) satisfies \(A^T \lambda = c, \ \lambda \geq 0\), and is dual optimal with \(c^T x^* = b^T \lambda^*\).

Basic idea:

Solve the dual problem and construct the primal solution from the dual solution.

How do we find \(x^*\)?
The primal and dual LPs are given by:

\[(P)\]: \[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad A x \geq b
\end{align*}
\]

\[(D)\]: \[
\begin{align*}
\text{minimize} & \quad -b^T \lambda \\
\text{subject to} & \quad A^T \lambda = c, \quad \lambda \geq 0
\end{align*}
\]

If \((x^*, \lambda^*)\) is the solution to \((P)\), then \(\lambda^*\) satisfies \(A^T \lambda = c, \quad \lambda \geq 0\), and is \textit{dual optimal} with \(c^T x^* = b^T \lambda^*\).

\[\text{Basic idea:}\]

Solve the dual problem and construct the primal solution from the dual solution.

\[\text{How do we find } x^*?\]
minimize \[ -b^T \lambda \]
subject to \[ A^T \lambda = c, \quad \lambda \geq 0 \]

The optimality conditions for the standard-form problem are

(A) \[ A^T \lambda^* = c, \quad \lambda^* \geq 0, \]

(B) \[ -b = A\pi^* + z^*, \quad z^* \geq 0, \]

(C) \[ z^* \cdot \lambda^* = 0, \]

where \( \pi^* \) and \( z^* \) are the multipliers for \( A^T \lambda = c \) and \( \lambda \geq 0 \).
The optimality conditions for the standard-form problem are

(A) \( A^T \lambda^* = c, \lambda^* \geq 0, \)

(B) \(-b = A \pi^* + z^*, z^* \geq 0,\)

(C) \(z^* \cdot \lambda^* = 0,\)

where \(\pi^*\) and \(z^*\) are the multipliers for \(A^T \lambda = c\) and \(\lambda \geq 0.\)

However, notice that

\[(B) \implies z^* = A(-\pi^*) - b \geq 0\]

If \(x^* = -\pi^*,\) then

\[z^* = Ax^* - b \geq 0 \quad \text{and} \quad z^* \cdot \lambda^* = (Ax^* - b) \cdot \lambda^* = 0\]

\(\implies \) the optimal primal \(x\) is the negative of the optimal dual \(\pi^*.\)

\(\implies \) the optimal primal residuals \(z^*\) are the dual reduced costs.
How is the basic set from the dual standard-form LP related to the primal all-inequality LP?

Strong duality implies that

Final dual basic set = Final primal working set

\[ A^T \downarrow B \triangleq A^T_\ell \implies A \downarrow \rightarrow A_\ell \]

\[ \Rightarrow A_\ell = B^T \] defines an optimal working-set matrix for the primal.
Summary

(P): \begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad A x \geq b
\end{align*}

(D): \begin{align*}
\text{maximize} & \quad b^T \lambda \\
\text{subject to} & \quad A^T \lambda = c, \quad \lambda \geq 0
\end{align*}

Solution \quad x^* \quad \lambda^* \\
Multipliers \quad \lambda^* \quad \pi\text{-values} \quad \pi^* \\
Residual \quad r^* \quad \text{Reduced costs} \quad z^* \\
Working set \quad A_\ell \quad \text{Basis} \quad B

The primal and dual solutions satisfy the following:

\[ x^* = -\pi^* \]

primal multipliers \quad \lambda^* = \text{dual solution} \quad \lambda^* \quad (\lambda^*_\ell = \lambda^*_B) 

\[ r^* = z^* \]

\[ A_\ell = B^T \]