# Simplex Method

Pramesh Kumar

IIT Madras

November 17, 2023

# Outline

# Recap

By now, we know

- What is a linear program?
- How to model some of the problems as a linear program?
- General concepts such as bounds, maximum, supremum, etc.
- General concepts such as halfspace, polyhedra, convex set, convex hull, extreme point, etc.
- How to write an LP in standard form?
- Reformulate pointwise max/min, linear fractional programs into LP.

# Outline

# A look at LP graphically

Suppose we want to solve the following LP:

$$
\begin{aligned}
Z = \underset{x,y}{\text{maximize}} \quad & x + 3y && \text{(1a)}\\
\text{subject to} \quad & 2x + y \leq 8 && \text{(1b)}\\
& y \leq 4 && \text{(1c)}\\
& x \geq 0 && \text{(1d)}\\
& y \geq 0 && \text{(1e)}
\end{aligned}
$$

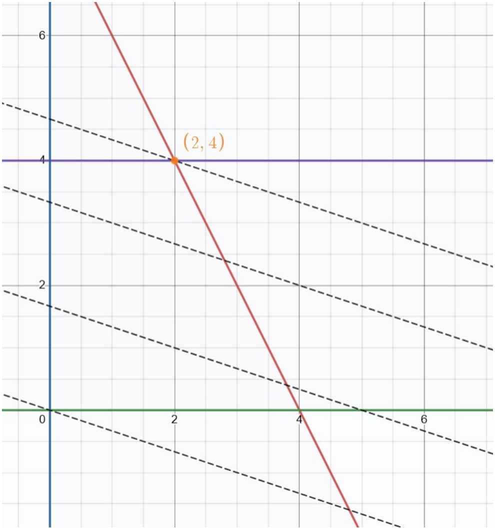


Figure: Optimal solution is obtained at an extreme point (2, 4)

# Outline

# Definitions

Definition (Extreme point) Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ be a polyhedron. We say that $\mathbf{x} \in P$ is an extreme point if we cannot express $\mathbf{x}$ as a convex combination of two other points in $P$.

Definition (Vertex) A face of dimension 0 of a polyhedron $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ is known as a vertex.

Definition (Basic feasible solution) Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ be a polyhedron. We say that $\mathbf{x} \in P$ is a basic feasible solution if there are $n$ linearly independent constraints of $A\mathbf{x} \leq \mathbf{b}$ active at $\mathbf{x}$.

Definition (Exposed solution) Let $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ be a polyhedron. We say that $\mathbf{x} \in P$ is an exposed solution if there exists $\mathbf{c} \in \mathbb{R}^n$ such that $\mathbf{c}^T\mathbf{x} < \mathbf{c}^T\hat{\mathbf{x}}, \forall \hat{\mathbf{x}} \in P \backslash \{\mathbf{x}\}$.

**Theorem**

*Let* $P = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} \leq \mathbf{b}\}$ *be a non-empty polyhedron. Let* $\mathbf{x} \in P$*. The following statements are equivalent.*

1. $\mathbf{x}$ *is an extreme point.*
2. $\mathbf{x}$ *is a vertex.*
3. $\mathbf{x}$ *is a basic feasible solution.*
4. $\mathbf{x}$ *is an exposed solution.*

## Theorem

*Let $P$ be a non-empty polyhedron. Consider LP $\min\{\mathbf{c}^T\mathbf{x} \text{ s.t. } \mathbf{x} \in P\}$. Suppose the LP has at least one optimal solution and $P$ has at least one extreme point. Then, above LP has at least one extreme point of $P$ that is an optimal solution.*

- This means that we need to search only among extreme points to find an optimal solution to LP.
- Since there are finite extreme points, we can just check out the objective value at every extreme point and find out which one attains the optimal value. BUT this doesn't seem practical!

# Outline

# Standard Form of LP

$$
\begin{aligned}
Z = \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x} && \text{(2a)} \\
\text{subject to} \quad & A\mathbf{x} = \mathbf{b} && \text{(2b)} \\
& \mathbf{x} \geq 0 && \text{(2c)}
\end{aligned}
$$

where, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ ($m < n$ fat matrix), $\mathbf{b} \in \mathbb{R}^m$.

Assumption: $A$ has linearly independent rows (full rank $m$).

What if the assumption is not satisfied?

- Either there are redundant constraints, which one can remove or inequations are not consistent (in which case there is no feasible solution).

## Proposition

*Consider an LP in standard form* $\min\{\mathbf{c}^T\mathbf{x} \text{ s.t. } A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq 0\}$*. Then,* $\hat{\mathbf{x}}$ *is a basic solution iff* $\exists\ m$ *columns of* $A$ *(denoted as* $A^{B(1)}, ..., A^{B(m)}$*) such that*

1. *columns* $A^{B(1)}, ..., A^{B(m)}$ *are linearly independent*
2. $\hat{x}_i = 0$,*if* $i \notin \{B(1), ..., B(m)\}$

To find a basic solution $\hat{\mathbf{x}}$,

1. Select $m$ independent columns of $A$, i.e., $A^{B(1)}, ..., A^{B(m)}$
2. Set $\hat{x}_i = 0$,if $i \notin \{B(1), ..., B(m)\}$ (non-basic variables)
3. Solve $A\hat{\mathbf{x}} = \mathbf{b}$ for the remaining variables $\hat{x}_{B(1)}, ..., \hat{x}_{B(m)}$ (basic variables)

If $\mathbf{x}_{B(i)} \geq 0, \forall i = 1, ..., m$, then $\hat{\mathbf{x}}$ is a basic feasible solution.

# Outline

# A first look at the Simplex Method

Consider the previous LP but in standard form:

$$
\begin{aligned}
Z = \underset{x_1,x_2}{\text{minimize}} \quad & -x_1 - 3x_2 && && \text{(3a)}\\
\text{subject to} \quad & 2x_1 + x_2 \quad +s_1 && = 8 && \text{(3b)}\\
& x_2 \quad + s_2 && = 4 && \text{(3c)}\\
& x_1, \quad x_2, \quad s_1, \quad s_2 && \geq 0 && \text{(3d)}
\end{aligned}
$$

Our strategy is to find a feasible solution $(x_1, x_2, s_1, s_2)$ and proceed to another feasible solution $(\bar{x}_1, \bar{x}_2, \bar{s}_1, \bar{s}_2)$ which is better in the sense that

$$-\bar{x}_1 - 3\bar{x}_2 < -x_1 - 3x_2$$

By repeating this procedure multiple times, we shall eventually arrive at an optimal solution.

# A first look at the Simplex Method

$$s_1 = 8 - 2x_1 - x_2 \tag{4}$$
$$s_2 = 4 \qquad\quad - x_2 \tag{5}$$
$$Z = \quad - x_1 - 3x_2 \tag{6}$$

- To begin with we set $x_1 = x_2 = 0$ and we find $s_1 = 8, s_2 = 4$. The solution $(0, 0, 8, 4)$ yields $Z = 0$.
- To decrease the objective value, we know that we should increase the value of $x_1, x_2$, but how much? We should keep in mind that by increasing the value of $x_1, x_2$, we should not make $s_1, s_2$ negative.
- Assuming that we only increase the value of $x_1$ and keep $x_2 = 0$, (4) implies that $s_1 = 8 - 2x_1 - x_2 \geq 0 \implies x_1 \leq 4$.
- Similarly, (5) implies that $s_2 = 4 - x_2 \geq 0 \implies 4 \geq 0$.
- We choose $x_1 = 4, x_2 = 0$, which gives us $Z = -4$ (an improvement in the objective value).
- We had all the variables taking zero values on the r.h.s. and rest on the l.h.s. By changing the value of $x_2$, we have changed that pattern. Let's rearrange.

## A first look at the Simplex Method

$$x_1 = 4 - \frac{1}{2}s_1 - \frac{1}{2}x_2 \tag{7}$$

$$s_2 = 4 \qquad\qquad - x_2 \tag{8}$$

$$Z = -4 + \frac{1}{2}s_1 - \frac{5}{2}x_2 \tag{9}$$

- Increasing the value of $s_1$ will not help in reducing the objective value further, but increasing $x_2$ will.
- Keeping $s_1 = 0$, we increase $x_2$.
- (7) implies $x_1 = 4 - s_1 - \frac{1}{2}x_2 \geq 0 \implies x_2 \leq 8$
- (8) implies $s_2 = 4 - x_2 \implies x_2 \leq 4$. $x_2 = 4$ is the best we can do.
- New solution is $x_1 = 2, x_2 = 4, s_1 = 0, s_2 = 0$ and the new objective value is $Z = -14$.

# A first look at the Simplex Method

We again write the system by writing $x_2$ on LHS

$$x_2 = 4 - s_2 \tag{10}$$

$$x_1 = 2 - \frac{1}{2}s_1 + \frac{1}{2}s_2 \tag{11}$$

$$Z = -14 + \frac{1}{2}s_1 + \frac{5}{2}s_2 \tag{12}$$

- Increasing the value of $s_1$ and $s_2$ will not help in reducing the objective value further. Therefore, we have arrived at an optimal solution $(2, 4, 0, 0)$.

# A first look at the Simplex Method

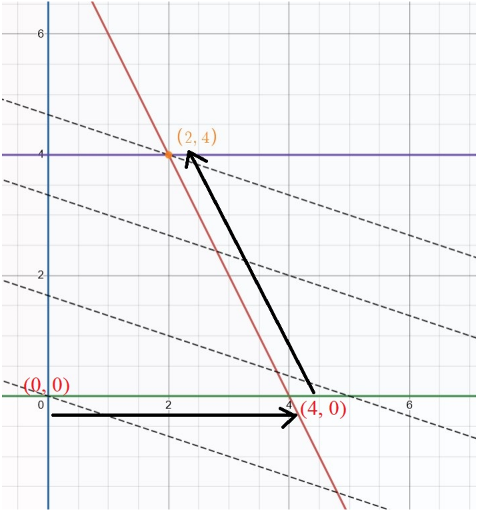


Figure: Observe the direction of exploring extreme points

# Feasible direction

For a given basic feasible solution (BFS), let

- $N$: index set of non-basic variables (ones on the r.h.s.)
- $B$: index set of basic variables (Basis) (ones on the l.h.s.)
- $A_B$ : $m \times m$ matrix (invertible) corresponding to $m$ columns of basic variables
- $A_N$ : $m \times (n-m)$ matrix corresponding to $n-m$ columns of non-basic variables
- Set of equations at every step of the simplex method is called a dictionary.

The value of basic variables is $\hat{\mathbf{x}}_B = A_B^{-1}\mathbf{b}$.

Definition (Feasible direction) Given $\hat{\mathbf{x}}$ (feasible), we say $\mathbf{d} \in \mathbb{R}^n$ is a feasible direction if $\hat{\mathbf{x}} + \theta\mathbf{d}$ is feasible for some $\theta > 0$.

- As we know $\hat{x}_j = 0, \forall j \in N$, we try to pick some $j^{'} \in N$ and increase the value of $\hat{x}_{j'}$ while keeping $\hat{x}_j = 0, \forall j \neq j^{'}$.
- Essentially, we are looking for $\mathbf{d} \in \mathbb{R}^n$ such that $d_j = 1,$ for some $j \in N$ and $d_k = 0$ for $k \in N \backslash \{j\}$ and moving from $\hat{\mathbf{x}}$ to $\hat{\mathbf{x}} + \theta\mathbf{d}, \theta > 0$, we have $A(\hat{\mathbf{x}} + \theta\mathbf{d}) = \mathbf{b}$.

## Feasible direction

- $A(\hat{\mathbf{x}} + \theta \mathbf{d}) = \mathbf{b} \implies A\mathbf{d} = 0$ (since $A\hat{\mathbf{x}} = \mathbf{b}$).
- $A\mathbf{d} = 0 \implies A_B \mathbf{d}_B + A_N \mathbf{d}_N = 0 \implies A_B \mathbf{d}_B + A_j = 0$
- $\implies \boxed{\mathbf{d}_B = -A_B^{-1} A_j}$.
- $\mathbf{d} = \begin{bmatrix} -A_B^{-1} A_j; & 0; & \dots & 1; & 0; & \dots & 0 \end{bmatrix}$. We call this $\mathbf{d}$ as the $j^{\text{th}}$ basic direction.

# Change in the objective value

- Change in the objective value: $\mathbf{c}^T[\mathbf{x}+\theta\mathbf{d}] - \mathbf{c}^T\mathbf{x} = \theta\mathbf{c}^T\mathbf{d}$
- Rate of change in the objective function.

$$\begin{aligned} \mathbf{c}^T\mathbf{d} = & \quad \mathbf{c}_B^T\mathbf{d}_B + \textstyle\sum_{k\in N} c_k d_k && (13)\\ = & \quad -\mathbf{c}_B^T A_B^{-1} A_j + c_j && (14) \end{aligned}$$

- $\boxed{\bar{c}_j = c_j - \mathbf{c}_B^T A_B^{-1} A_j}$ is called reduced cost of variable $j \in N$.
- Note: For basic variable $\bar{c}_{B(i)} = c_{B(i)} - \mathbf{c}_B^T A_B^{-1} A_{B(i)} = c_{B(i)} - \mathbf{c}_B^T e_i = 0$.

## Theorem

*Let $\hat{\mathbf{x}}$ be BFS. If $\bar{\mathbf{c}} \geq 0$, then $\hat{\mathbf{x}}$ is optimal.*

This gives us a stopping criterion for the simplex method. If all the reduced costs are non-negative, then we stop. This means we cannot find a neighboring solution which is better than the current solution.

## Changing the basis

- If we have, $\bar{c}_j < 0$ for some $j \in N$, then by bringing the $j^{\text{th}}$ non-basic variable (entering variable) into the basis will further decrease the objective value. So, we want to go in the $j^{\text{th}}$ basic direction $\mathbf{d}$. But how much should we go in that direction? i.e., $\theta = ?$
- Let $\theta^*$ be the largest value of $\theta$ s.t. $\hat{\mathbf{x}} + \theta^* \mathbf{d}$ is still feasible.
  1. $d_k \geq 0, \forall k \in \{1, ..., n\}$. Then, $\hat{\mathbf{x}} + \theta \mathbf{d} \geq 0, \forall \theta \geq 0$ and $A(\hat{\mathbf{x}} + \theta \mathbf{d}) = \mathbf{b}$ (since $A\mathbf{d} = 0$). Therefore, $\hat{\mathbf{x}} + \theta \mathbf{d} \geq 0, \forall \theta$. But we travel in this direction, the objective function will keep reducing since $\bar{c}_j < 0$. Therefore, LP is unbounded.
  2. $d_k < 0$ for some $k \in \{1, ..., n\}$. Then, we want $\hat{x}_k + \theta^* d_k \geq 0, \forall k$ s.t. $d_k < 0 \implies \theta^* \leq -\frac{\hat{x}_k}{d_k}, \forall k$ s.t. $d_k < 0$.

$$\boxed{\theta^* = \min_{\{k \in B : d_k < 0\}} \left\{ -\frac{\hat{x}_k}{d_k} \right\}}$$

# Changing the basis

## Proposition

*Let $y = \hat{\mathbf{x}} + \theta\mathbf{d}$ as computed above. Then, $y$ is a BFS.*

If we don't have degenerate vertices, i.e, $\hat{x}_{B(i)} > 0, \forall i$. Then, $\theta^* > 0$ using above formula and change in the objective value $\theta^* \bar{c}_j < 0$.

## Theorem

*If all BFS are non-degenerate then the simplex algorithm described above finds an optimal solution or detects unboundedness in finite time.*

# Simplex algorithm

0. We start with a BFS $\hat{\mathbf{x}}$ (with corresponding basis $B$)
1. Compute the reduced costs $\bar{c}_j = c_j - \mathbf{c}_B^T A_B^{-1} A_j, \forall j \in N$
   - If $\bar{c}_j \geq 0 \forall j$, then set $\hat{x}$ as OPTIMAL.
   - Otherwise, choose some $j \in N$ such that $c_j < 0$. If multiple choices available, choose one with the smallest index (Bland's rule)
2. Compute the $j^{\text{th}}$ basic direction
   $\mathbf{d} = \begin{bmatrix} -A_B^{-1} A_j; & 0; & \dots & 1; & 0; & \dots & 0 \end{bmatrix}$.
   - If $\mathbf{d} \geq 0$, then set problem as UNBOUNDED and $Z^* = -\infty$
   - Otherwise, compute $\theta^* = \min_{\{i \in B : d_i < 0\}} \left\{ -\frac{\hat{x}_i}{d_i} \right\}$. If multiple $i$ achieves minimum, select one with the smallest index (Bland's rule).
3. Define the new BFS $\hat{\mathbf{x}} + \theta^* \mathbf{d}$. (New BFS with index $j$ replacing $B(l)$ in the basis, where $B(l) = \operatorname{argmin}_{\{i \in B : d_i < 0\}} \left\{ -\frac{\hat{x}_i}{d_i} \right\}$)
4. Repeat 1-3

# Degeneracy

- It is possible that the objective value does not change from one iteration to another in the simplex method. When will that happen?
- $\theta^* \bar{c}_j = 0 \implies \theta^* = 0$ which implies that $\hat{x}_i = 0$ for some $i \in B$,i.e., there is zero in the current BFS $\hat{\mathbf{x}}$.
- This may result in cycle and the algorithm will not terminate.

Definition (Degeneracy): We call a BFS degenerate if some of its basic variables are zero.

- Also, if $\bar{c}_j < 0$ for multiple $j \in N$. Then, which non-basic index should enter the basis?

Bland's rule: If $\bar{c}_j < 0$ for multiple $j \in N$, choose one with the smallest index for entering the basis. Similarly, if multiple $i \in B : d_i < 0$ attains the minimum in $\left\{ -\frac{\hat{x}_i}{d_i} \right\}$, then choose one with the smallest index to leave the basis.

### Theorem

*Simplex algorithm implemented with Bland's rule will not cycle.*

# Simplex Tableau

- A simpler implementation of the Simplex Method.
- You can solve small sized problems using pen and paper.

| | |
|---|---|
| Reduced cost | -Objective value |
| $A$ | $b$ |

| | |
|---|---|
| $\mathbf{c}^T - \mathbf{c}_B^T A_B^{-1} A$ | -$\mathbf{c}_B^T A_B^{-1} \mathbf{b}$ |
| $A_B^{-1} A$ | $A_B^{-1} b$ |

| | | |
|---|---|---|
| $\mathbf{0}_m$ | $\mathbf{c}_N^T - \mathbf{c}_B^T A_B^{-1} A_N$ | -$\mathbf{c}_B^T A_B^{-1} \mathbf{b}$ |
| $\mathbf{I}_{m \times m}$ | $A_B^{-1} A_N$ | $A_B^{-1} b$ |

If the first row of the LHS (reduced costs) are all non-negative, then we've reached an optimal solution.

# Steps

1. Compute the reduced costs
2. Choose the incoming basis
3. Compute $\theta^*$ and choose the outgoing basis
4. Update the tableau with new basis

Above procedure is called pivoting.
Consider the previous example.

| Basis | -1 | -3 | 0 | 0 | 0 |
|---|---|---|---|---|---|
| $s_1$ | 2 | 1 | 1 | 0 | 8 |
| $s_2$ | 0 | 1 | 0 | 1 | 4 |

(Pivoting Step 1): Choose the incoming basis. Since both $x_1$ and $x_2$ have negative reduced costs, both can be allowed to enter the basis. Let's select $x_1$ (pivot column) using Bland's rule.

(Pivoting Step 2): Choose the outgoing basis. Remember $\theta^* = \min_{\{i \in B : d_i < 0\}} \left\{ -\frac{\hat{x}_i}{d_i} \right\} = \min_{\bar{A}_{ij} > 0} \left\{ \frac{\bar{b}_i}{\bar{A}_{ij}} \right\}$ (index of entering variable is $j$).
This is called the Minimal Ratio Test (MRT). If $\bar{A}_{ij} < 0, \forall i$, then LP is unbounded.

# Steps

We have $\theta^* = \min\{\frac{8}{2}\}$. The first row (pivot row) attains the minimum. Therefore, $s_1$ is the outgoing basis. The intersection of pivot row and pivot column is the pivot element.

(Step 3): Updating the tableau. First divide the each element of the pivot row by pivot element. Then add or subtract multiples of pivot row to other rows (including the first row) such that the elements in the pivot column (except pivot element) becomes zero. This operation is also performed on the RHS (current BFS).

| Basis | 0 | -5/2 | 1/2 | 0 | 4 |
|---|---|---|---|---|---|
| $x_1$ | 1 | 1/2 | 1/2 | 0 | 4 |
| $s_2$ | 0 | 1 | 0 | 1 | 4 |

(Step 4): Repeat steps 1-3 until all elements in the top row becomes non-negative.

Since only $x_2$ has the negative reduced cost. It enters the basis. Also, from minimal ratio test, we have $\theta^* = \min\{\frac{4}{1/2}, \frac{4}{1}\} = 4$. Therefore, $s_2$ is leaves the basis.

# Steps

| Basis | 0 | 0 | 1/2 | 5/2 | 14 |
|---|---|---|---|---|---|
| $x_1$ | 1 | 1/2 | 1/2 | -1/2 | 2 |
| $x_2$ | 0 | 0 | 0 | 1 | 4 |

Since all the reduces costs are non-negative, current basis is optimal. The optimal solutions is bottom RHS, i.e., $(x_1^*, x_2^*) = (2, 4)$ and the optimal value is the negative of top RHS, i.e., $Z^* = -14$.

# Phase-I simplex method

Two things left

- We don't no how to find an initial BFS. It is straight forward to find an initial BFS if we derived standard form by adding a slack to each constraint (previous example).
- We don't know if LP is infeasible.

Assume $\mathbf{b} \geq 0$ (multiply constraint by -1 if needed).

$$
\begin{aligned}
Z = \underset{\mathbf{x},\mathbf{y}}{\text{minimize}} \quad & \sum_{i=1}^{m} y_i && \text{(15a)} \\
\text{subject to} \quad & A\mathbf{x} + \mathbf{y} = \mathbf{b} && \text{(15b)} \\
& \mathbf{x} \geq 0 && \text{(15c)} \\
& \mathbf{y} \geq 0 && \text{(15d)}
\end{aligned}
$$

We can solve above program using the simplex method. Clearly, $(\mathbf{0}, \mathbf{b})$ is an initial BFS.

## Proposition

*The original LP is feasible iff the Phase-I LP objective value is 0.*

# George Bernard Dantzig

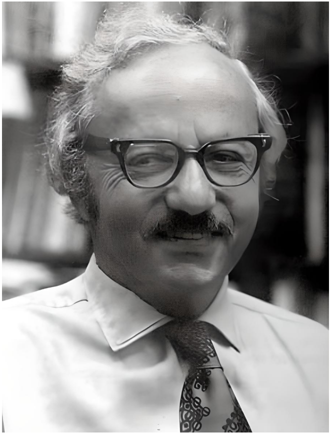

Figure: George Bernard Dantzig invented the Simplex Method

(Picture source: malevus.com)

## Final thoughts

- Simplex method is not a polynomial time algorithm. Klee and Minty (1992) showed an example where it will run exponential iterations (in terms of $n$) to stop. However, the average performance of simplex method is very good and it remains one of the widely used algorithms to solve LP.
- Soviet Union mathematician Khachiyan in 1979 devised a first polynomial-time algorithm for LP. His method is called the ellipsoid method. However, it is observed to be slow in practice.
- Narendra Karmarkar (IITB graduate) devised an interior point method in 1984, which is a polynomial time algorithm and it is quite efficient in practice.
- We'll study LP duality theory in the next lecture.
- If you are interested, read the following topics from the book.
  - Dual Simplex
  - Revised Simplex
  - Sensitivity analysis
  - Network simplex method
  - Primal-dual method