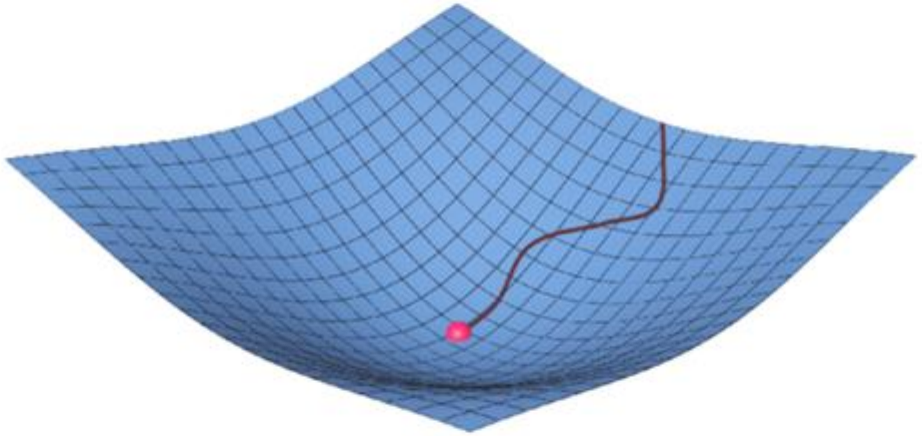
Optimization



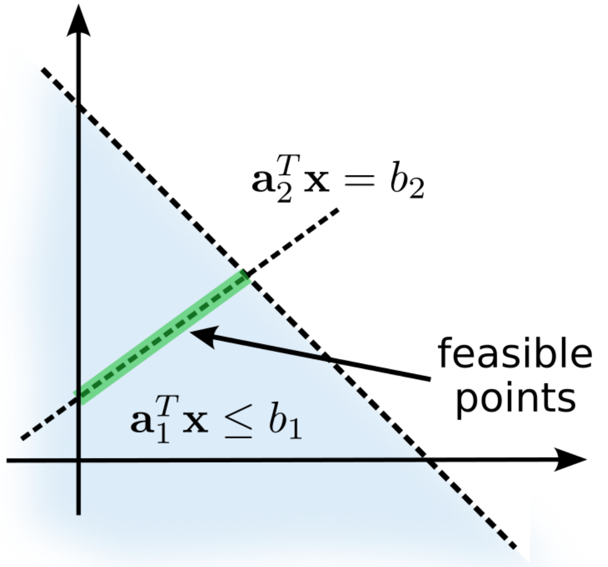
Paul J. Atzberger

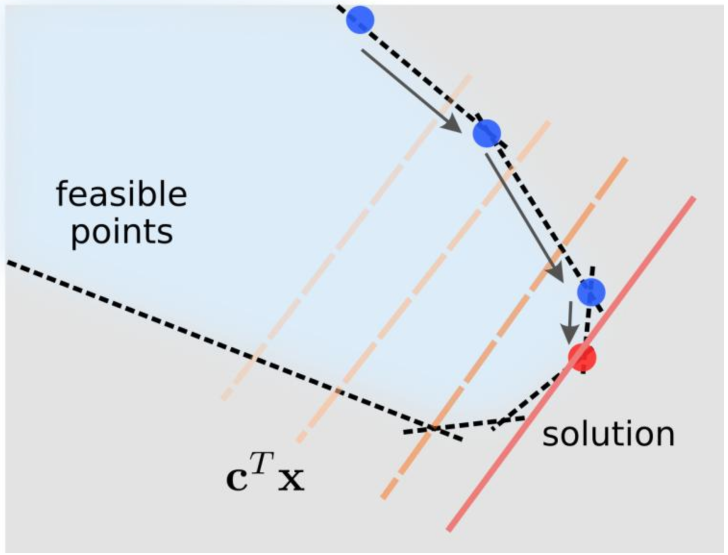
University of California Santa Barbara (UCSB)

Optimization with Linear Constraints

Linear Programming

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{subject} & A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0. \end{array}$$





Motivations

Linear constraints arise naturally in many problems in economics, science, engineering, and statistics.

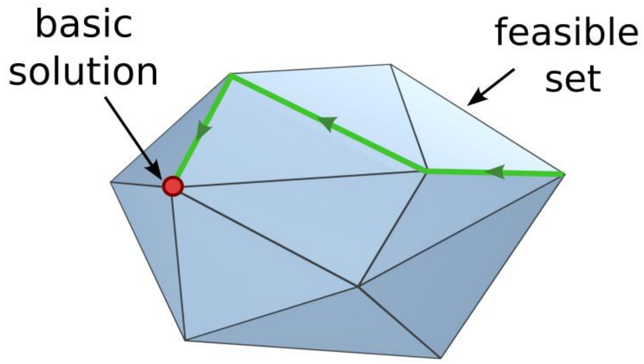
Efficient methods available for solving many linear programs in practice.

Linear objective functions provide useful class for modeling and analysis.

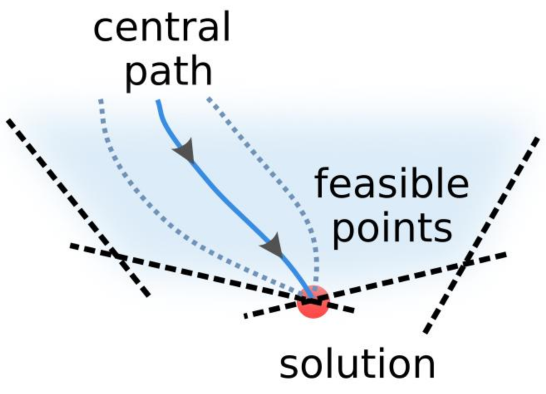
Applications

- Science and Engineering
- Economics, Business Decisions
- Planning Logistics, Transportation
- Machine learning, AI
- Statistics and Data Analysis

Simplex Method



Interior Point Methods



Business Logistics: Manufacturing

factory	porcelain	glass	available operation hours
A	70	30	1500
B	60	50	1000
C	55	45	750
profit per unit	$4000	$2000	

Task: maximize profits by planning production at factories for plates.

Factory production: A: (x_1,x_2) B: (x_3,x_4), C: (x_5,x_6).

Linear Programming Problem

$$\min \quad 4(x_1 + x_3 + x_5) + 2(x_2 + x_4 + x_6)$$

subject

$$7x_1 + 3x_2 \leq 150$$
$$6x_3 + 5x_4 \leq 100$$
$$5.5x_5 + 4.5x_6 \leq 75$$

Machine Learning: Linear Classification

Task: Learn hyperplane that separates the data into two classes.

Data: $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, with features x, labels y.

Example: $x \in \mathbb{R}^N$, $y \in \{-1,+1\}$, with x=image, y = +1→ Apple, y = -1→ Orange.

Linear classifiers

$$\mathcal{H} = \{h \mid h(x) = \text{sign}(w^T x + b),\ w \in \mathbb{R}^N,\ b \in \mathbb{R}\}$$

we require classification with

$$y_i(w^T x + b) \geq 1$$

Linear Programming Problem

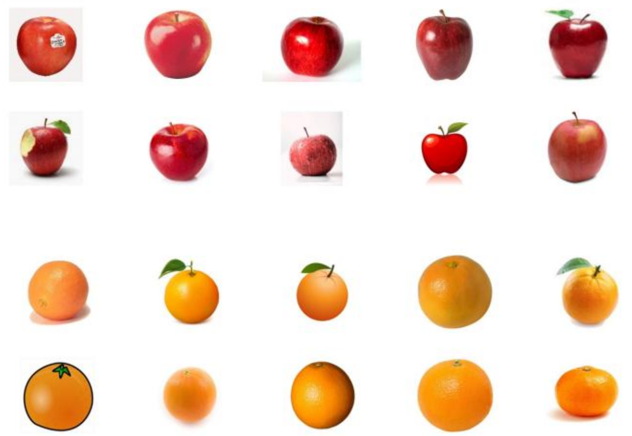
$$\min_{\mathbf{w},b,\xi} \sum_{i=1}^{m} \xi_i$$

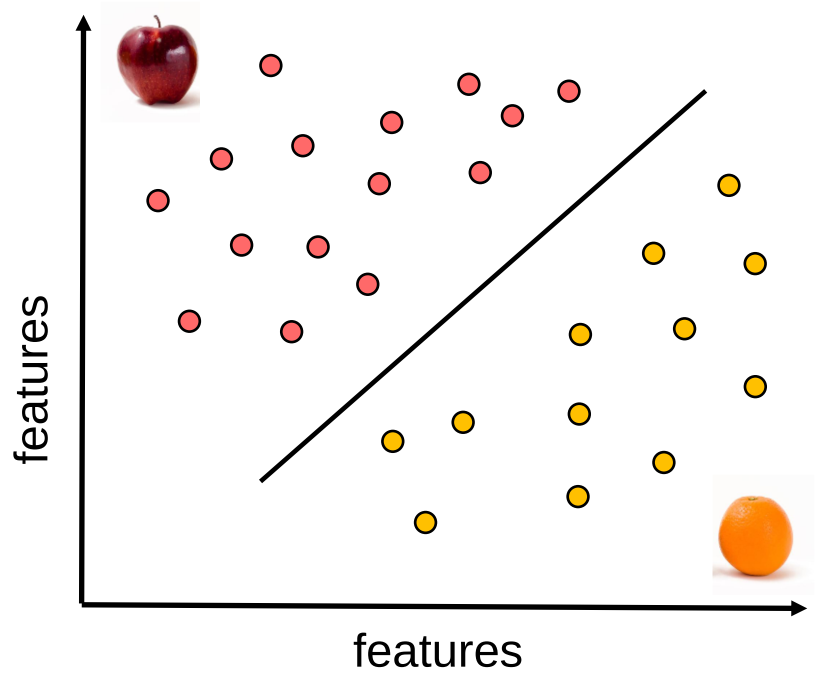
subject

$$y_i\left(\mathbf{w}^T \mathbf{x}_i + b\right) \geq 1 - \xi_i,$$

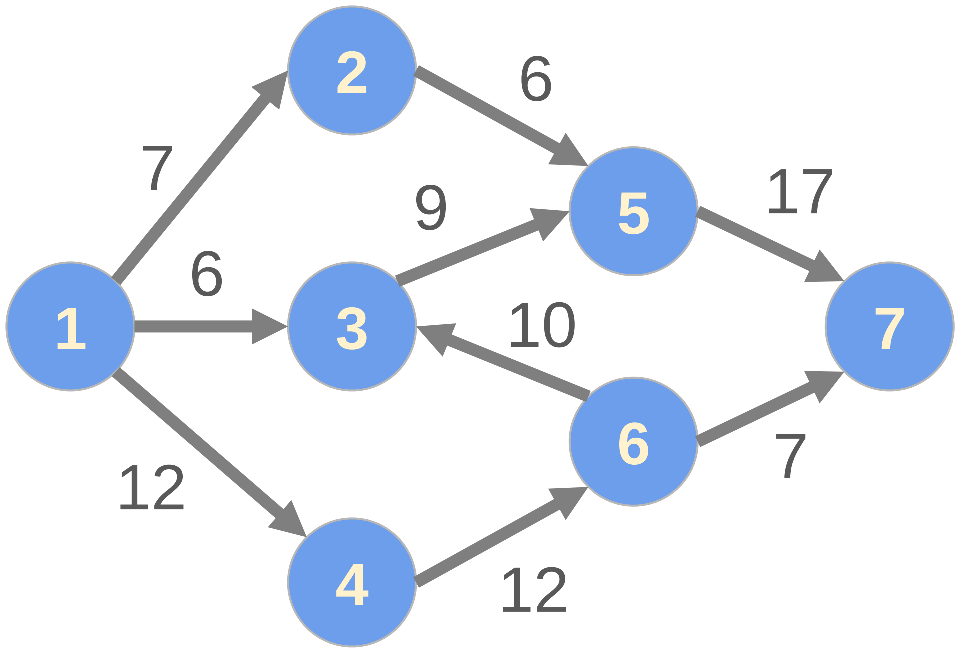
$$\xi \geq 0.$$

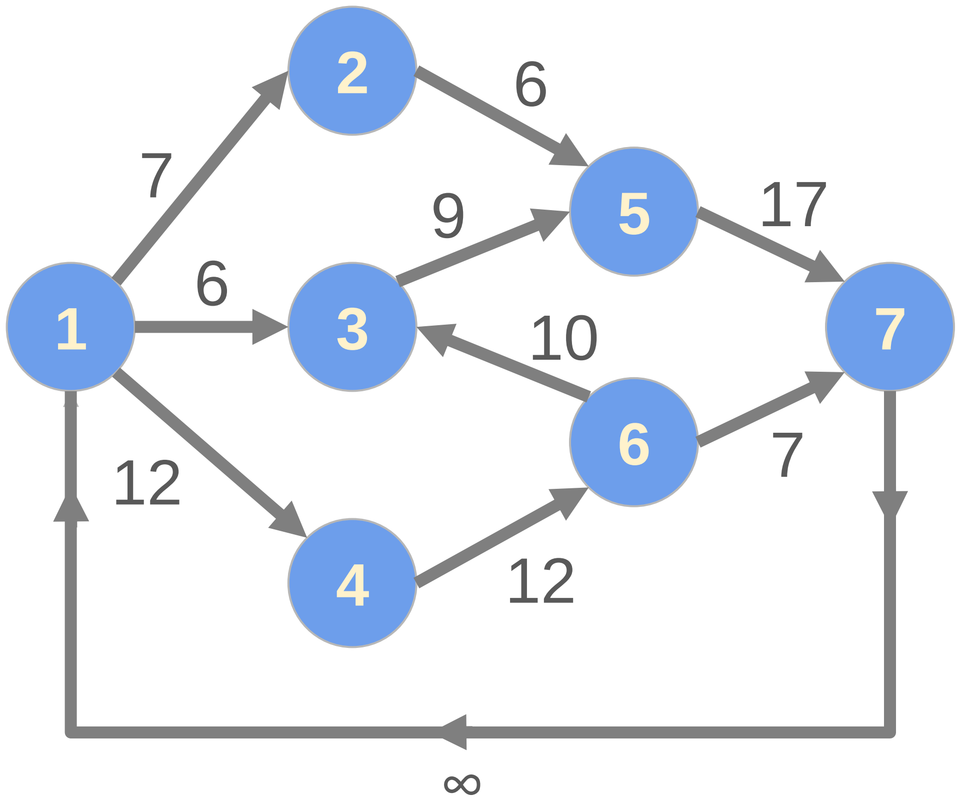
example images





Network Transport Capacity





Task: Determine the maximum network flow possible from 1→7.

- Decision variables are x_{ij} for amount to send from node i → j.

- We need to decide how much we send out along each edge, x_{ij} .

Constraints:

- Edges can only sustain the shown amounts k_{ij} from i → j.

- The amount going out of each note can not exceed the amount coming in.

Formulation: Add an extra edge with $k_{71} = \infty$.

Linear Programming Problem

$$\max \; x_{71}$$

subject

$$\sum_j x_{ij} \leq \sum_k x_{ki}, \;\; i = 1, \ldots, 7$$

$$x_{ij} \leq k_{ij}$$

Business: Supply-Chain Transportation Costs

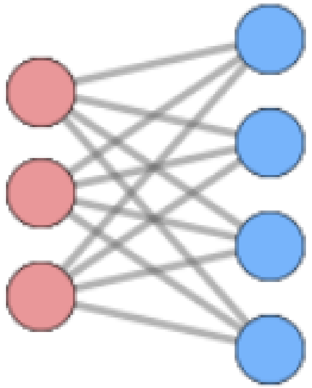
	Retailer A	Retailer B	Retailer C	Retailer D	SUPPLY (si)
Supplier A	10	25	10	5	250
Supplier B	12	30	18	23	450
Supplier C	5	40	22	15 (c_{ij})	300
DEMAND (d_j)	400	200	250	150	1000




Task: Determine how much to ship from each supplier to satisfy the retailer order demand.

- Decision variables are x_{ij} for amount to send from supplier i → retailer j.

- The cost of each transport route is c_{ij}.



Constraints:

- Our shipping must meet demand, $\sum_i x_{ij} = d_j$.

- We must ship all of our supply, $\sum_j x_{ij} = s_i$.

Linear Programming Problem

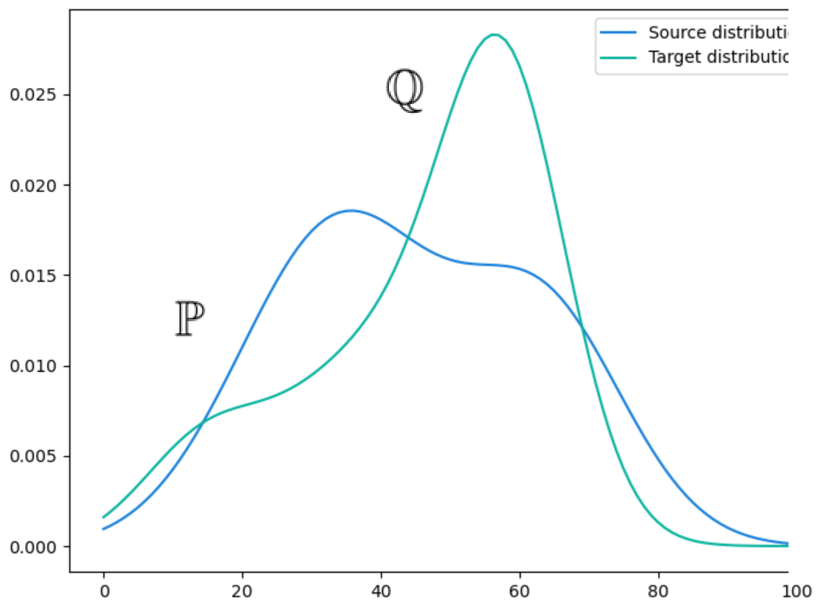
$$\min \sum_{i,j} c_{ij} x_{ij}$$

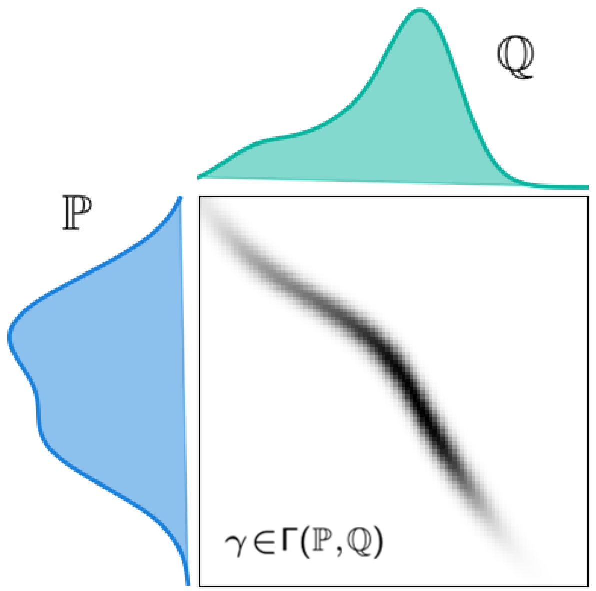
subject

$$\sum_i x_{ij} = d_j, \quad \sum_j x_{ij} = s_i,$$

$$x_{ij} \geq 0.$$

Wasserstein Distance: Probability Theory and Statistics





Task: Determine the "best" alignment between two probability distributions.

We use the transport cost associated with the align to define a "distance" between probability distributions P, Q.

Wasserstein Distance

$$W_c(\mathbb{P},\mathbb{Q}) = \inf_{\gamma \in \Gamma(\mathbb{P},\mathbb{Q}),(X,Y)\sim\gamma} \mathbb{E}_\gamma\left[c(X,Y)\right]$$
$$= \int\int c(x,y)\gamma(x,y)dxdy,$$

where $\gamma \in \Gamma(\mathbb{P},\mathbb{Q})$ is a joint probability distribution with marginal $\int \gamma(x,y)dy = \mathbb{P}(x)$ and $\int \gamma(x,y)dx = \mathbb{Q}(y)$.

In the discrete case, we have

$$\mathbb{P}(x) = \sum_i p_i \delta(x - x_i)$$
$$\mathbb{Q}(x) = \sum_i q_i \delta(x - x_i)$$

Wasserstein Distance

$$W_c(\mathbb{P},\mathbb{Q}) = \inf_{\gamma \in \Gamma(\mathbb{P},\mathbb{Q}),X,Y\sim\gamma} \mathbb{E}_\gamma\left[c(X,Y)\right]$$
$$= \sum_{ij} c(x_i,x_j)\gamma(x_i,x_j)$$

Let $c_{ij} = c(x_i,x_j)$, and $p_i=p(x_i)$, $q_j=q(x_j)$, $x_{ij}=\gamma(xi,xj)$.
This can be reformulated as a linear program.

Linear Programming Problem

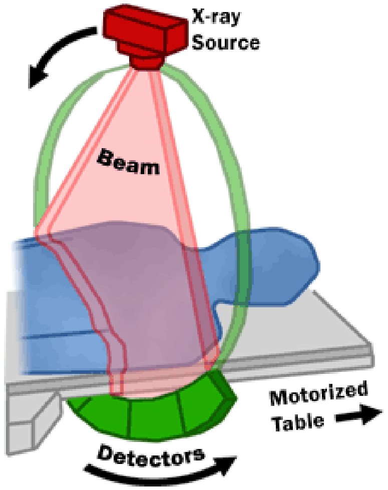
$$\min \sum_{i,j} c_{ij} x_{ij}$$

subject

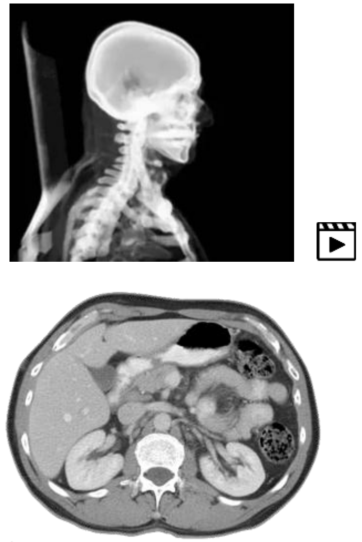
$$\sum_i x_{ij} = p_j, \quad \sum_j x_{ij} = q_i,$$
$$0 \leq x_{ij} \leq 1.$$

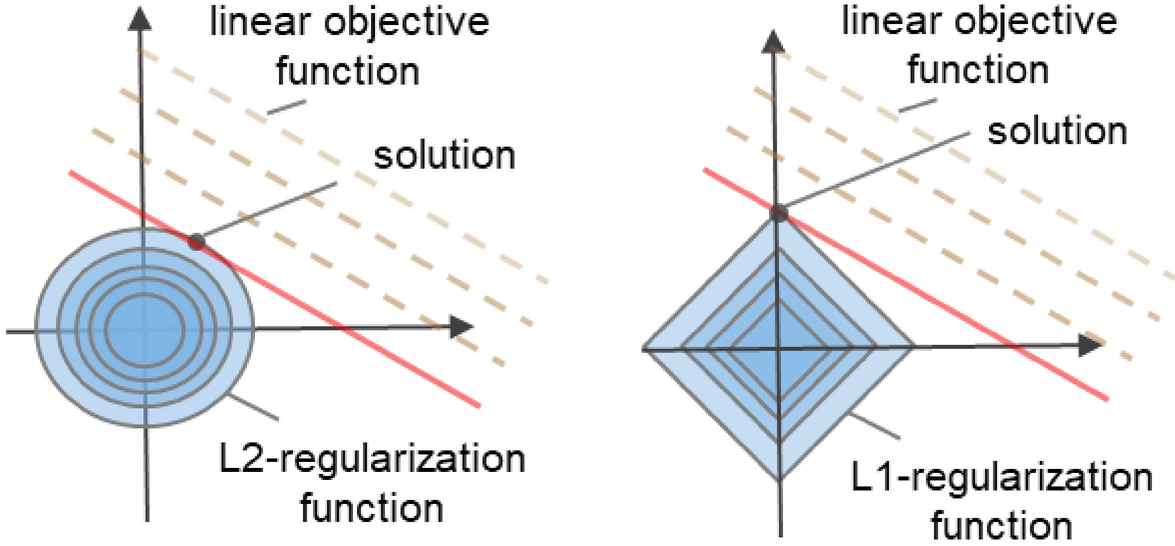
Compressed Sensing: L1-Reconstruction of Sparse Signals

fda.gov



fda.gov





Task: Determine the "best sparse" reconstruction of x satisfying the under-determined linear system Ax = b.

Optimization Problem

$$\min \; \|\mathbf{x}\|_1$$

subject

$$A\mathbf{x} = \mathbf{b}.$$

$\|x\|_1 = |x_1| + |x_2| + \ldots + |x_n|$. Seeks x_i with bias toward zero components.

For simplicity, above is for the noiseless case (can extend for noise).

Linear Programming Problem

$$\min_{\mathbf{x}^+,\mathbf{x}^-,\mathbf{y}} \sum_i y_i$$

subject

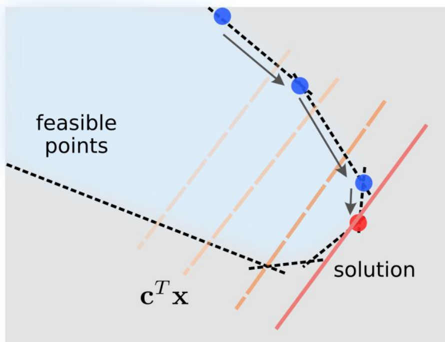
$$A\left(\mathbf{x}^+ - \mathbf{x}^-\right) = \mathbf{b}$$

$$y_i - \left(x_i^+ - x_i^-\right) \geq 0, \;\; y_i + \left(x_i^+ - x_i^-\right) \geq 0$$

$$x_i^+ \geq 0, \;\; x_i^- \geq 0, \;\; y_i \geq 0.$$

This yields $x_i = x_i^+ - x_i^-$, $y_i \geq \max(x_i, -x_i) = |x_i|$.

Linear Programming: Primal and Dual Problems



Primal LP Problem

$$\begin{array}{ll} \max & \mathbf{c}^T\mathbf{x} \\ \text{subject} & A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0. \end{array}$$

Lagrangian for LP

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{s}) = \mathbf{c}^T\mathbf{x} - \boldsymbol{\lambda}^T(A\mathbf{x} - \mathbf{b}) - \mathbf{s}^T\mathbf{x}.$$

KKT Conditions

$$\nabla_{\mathbf{x}}\mathcal{L} = 0 \Rightarrow A^T\boldsymbol{\lambda} + \mathbf{s} = \mathbf{c}.$$

$$\nabla_{\boldsymbol{\lambda}}\mathcal{L} = 0 \Rightarrow A\mathbf{x} = \mathbf{b}.$$

$$\nabla_{\mathbf{s}}\mathcal{L} \leq 0 \Rightarrow \mathbf{x} \geq 0, \mathbf{s} \geq 0, \quad x_i s_i = 0.$$

For solution $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{s}^*)$, we have $(\mathbf{s}^{*,T}\mathbf{x} = 0)$,

$$\mathbf{c}^T\mathbf{x}^* = \left(A^T\boldsymbol{\lambda}^* + \mathbf{s}^*\right)^T\mathbf{x}^* = (A\mathbf{x}^*)^T\boldsymbol{\lambda}^* + \mathbf{s}^{*,T}\mathbf{x}^* = \mathbf{b}^T\boldsymbol{\lambda}^*.$$

KKT sufficient for showing $\mathbf{x}^*$ is a minimizer. For example, let $\bar{\mathbf{x}}$ be any feasible point $A\bar{\mathbf{x}} = \mathbf{b}$, $\bar{\mathbf{x}} \geq 0$, then $(\mathbf{s}^{*,T}\bar{\mathbf{x}} \geq 0)$,

$$\mathbf{c}^T\bar{\mathbf{x}} = \left(A^T\boldsymbol{\lambda}^* + \mathbf{s}^*\right)^T\bar{\mathbf{x}} = (A\bar{\mathbf{x}})^T\boldsymbol{\lambda}^* + \mathbf{s}^{*,T}\bar{\mathbf{x}} \geq \mathbf{c}^T\mathbf{x}^*$$

$$\Rightarrow \mathbf{c}^T\bar{\mathbf{x}} \geq \mathbf{c}^T\mathbf{x}^*.$$

Dual LP Problem

$$\begin{array}{ll} \max & \mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject} & A^T\boldsymbol{\lambda} \leq \mathbf{c}. \end{array}$$

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{s}) &= \mathbf{c}^T\mathbf{x} - \boldsymbol{\lambda}^T(A\mathbf{x} - \mathbf{b}) - \mathbf{s}^T\mathbf{x} \\ &= \mathbf{b}^T\boldsymbol{\lambda} - \mathbf{x}^T\left(A^T\boldsymbol{\lambda} + \mathbf{s} - \mathbf{c}\right) \\ &= \mathbf{x}^T\left(\mathbf{c} - A^T\boldsymbol{\lambda} - \mathbf{s}\right) + \mathbf{b}^T\boldsymbol{\lambda}. \end{aligned}$$

$\nabla_{\mathbf{x}}\mathcal{L} = \mathbf{c} - A^T\mathbf{x} - \mathbf{s} = 0$, implies

$$q(\boldsymbol{\lambda}, \mathbf{s}) = \inf_{\mathbf{x} \in \mathbb{R}^n} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{s})$$

$$= \begin{cases} \mathbf{b}^T\boldsymbol{\lambda}, & \text{if } (\mathbf{c} - A^T\boldsymbol{\lambda} - \mathbf{s}) = 0 \\ -\infty, & \text{if } (\mathbf{c} - A^T\boldsymbol{\lambda} - \mathbf{s}) \neq 0. \end{cases}$$

$$\max_{\boldsymbol{\lambda}, \mathbf{s},\ \mathbf{s} \geq 0} q(\boldsymbol{\lambda}, \mathbf{s}) = \begin{array}{ll} \max & \mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject} & A^T\boldsymbol{\lambda} - \mathbf{c} + \mathbf{s} = 0,\ \mathbf{s} \geq 0. \end{array}$$

$$= \begin{array}{ll} \max & \mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject} & A^T\boldsymbol{\lambda} - \mathbf{c} \leq 0. \end{array}$$

Solving KKT gives an optimal solution to both the Primal and Dual LP Problems.

Simplex Method: Canonicalization and Geometry

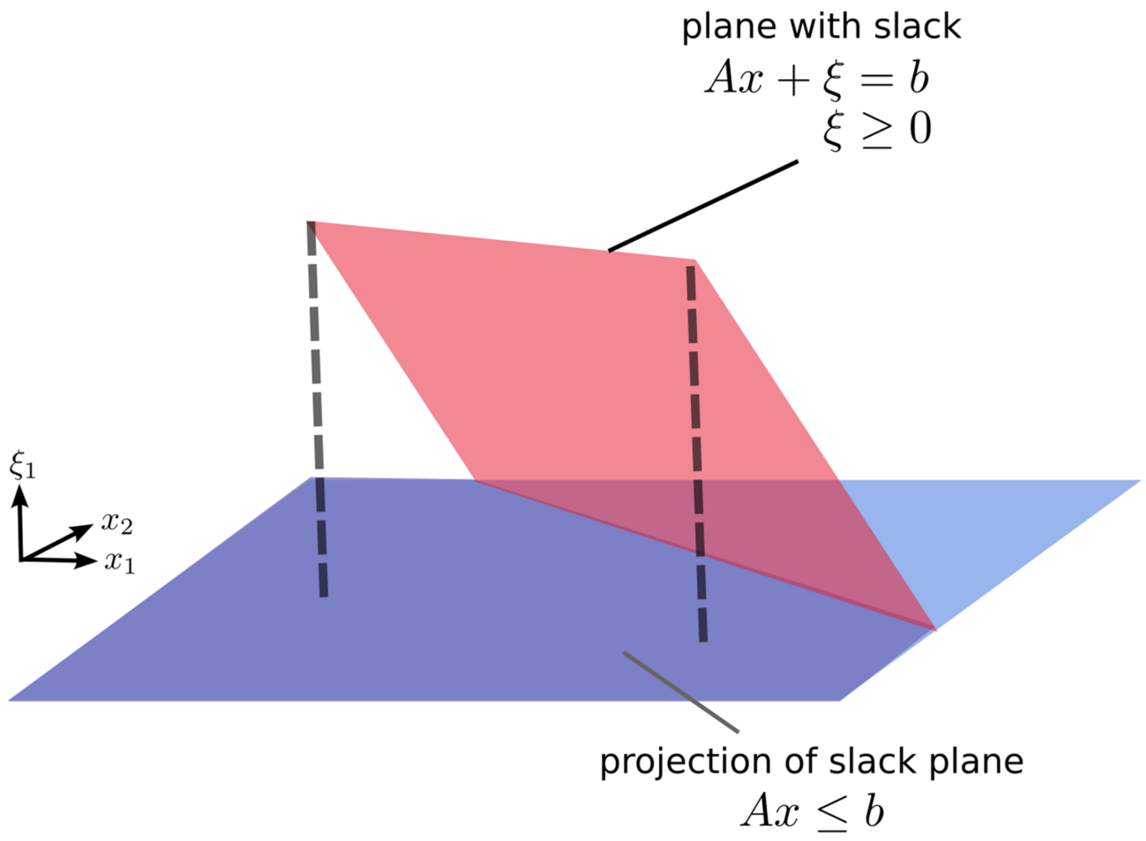
LP Problem

$$\begin{aligned} \min x_1 + x_2 \\ \text{subject } x_1 + 2x_2 \leq 4 \\ x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

Use slack or surplus variables to standardize.

LP Problem (canonicalized)

$$\begin{aligned} \min x_1 + x_2 \\ \text{subject } x_1 + 2x_2 + x_3 = 4 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{aligned}$$



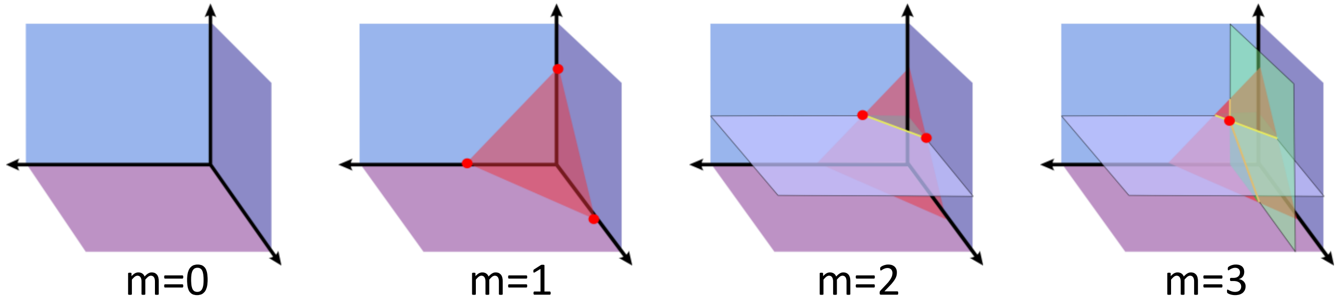
Geometry of having only equality constraints.

Solutions now lie on intersections of hyperplanes in the generalized positive quadrant (x ≥ 0).

Planes arising from slack, project to half-spaces in the original variables.

Provides unified approaches to treat diverse LP problems.

Simplex Method: Basic Feasible Points



Consider the constraints and **basic feasible points**

$$Ax = b,\ x \geq 0$$

As we increase m, we show how these points change (example in 3D).

m=1: triangle (simplex dimension n-1)
m=2: line (simplex dimension n - 2)
m=3: point (simplex dimension n - 3)

Basic feasible points are the vertices.

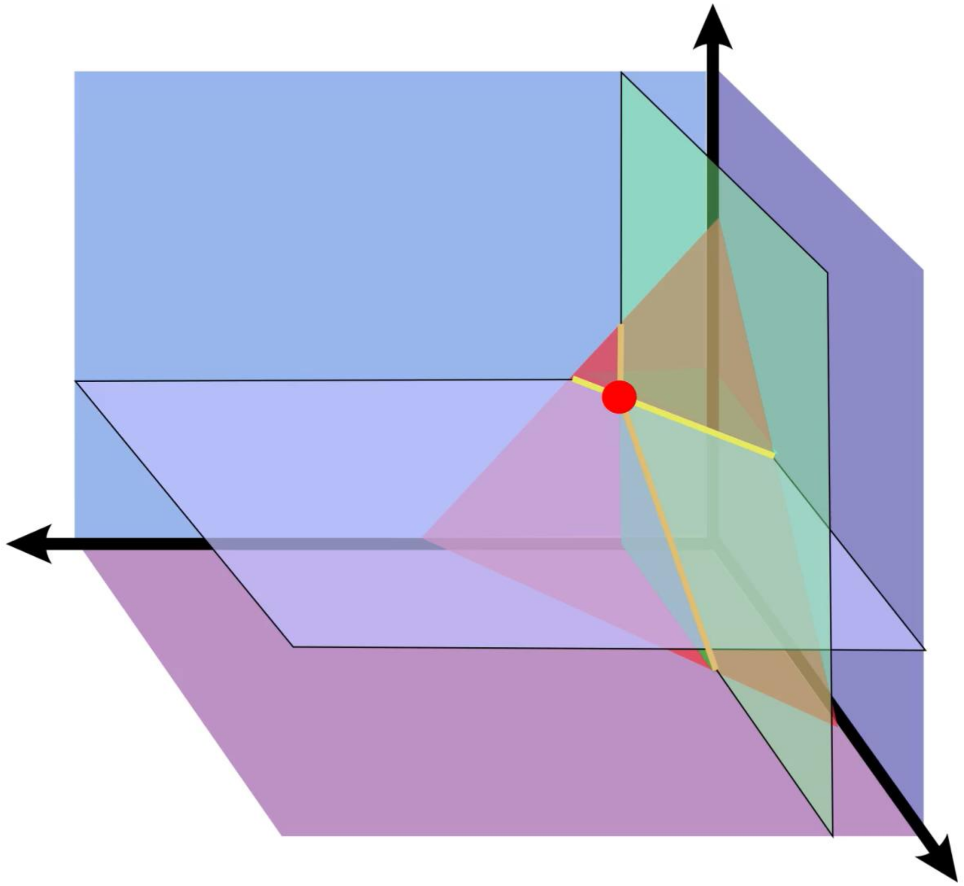
For canonicalized problem, solutions lie in the intersection of hyperplanes in the generalized positive quadrant (x ≥ 0).

Geometry requires they have **at most m non-zero values**

$$x = (x_1, x_2, \ldots, x_m, 0, \ldots, 0)$$

Solution strategy: develop algorithms that search among the basic feasible points.

basic feasible points



$$Ax = b,\ x \geq 0$$

Simplex Method

Def: A *basic feasible point* $\mathbf{x}$ is a point that is feasible and for which there exists a collection of indices $\mathcal{B}$ satisfying the following properties

(i) $|\mathcal{B}| = m$, contains exactly m indices.

(ii) if $i \notin \mathcal{B}$ then $x_i = 0$.

(iii) The $m \times m$ matrix $B = [A_i]_{i\in\mathcal{B}}$ is non-singular.

For a basic feasible point $x=(x_B,x_N)$, construct the triple **(x,λ,s)** to check KKT for optimality.

$$A\mathbf{x} = \mathbf{b} \;\Rightarrow\; B\mathbf{x}_B + N\mathbf{x}_N = \mathbf{b} \;\Rightarrow\; B\mathbf{x}_B = \mathbf{b} \;\Rightarrow\; \mathbf{x}_B = B^{-1}\mathbf{b}$$

$$A^T\boldsymbol{\lambda} + \mathbf{s} = \mathbf{c}$$

$$B^T\boldsymbol{\lambda} + \mathbf{s}_B = \mathbf{c}_B, \quad N^T\boldsymbol{\lambda} + \mathbf{s}_N = \mathbf{c}_N$$

$$\mathbf{x}^T\mathbf{s} = \mathbf{x}_B^T\mathbf{s}_B + \mathbf{x}_N^T\mathbf{s}_N = 0$$

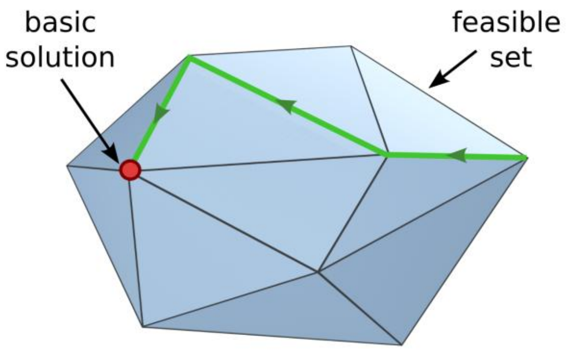
$$B^T\boldsymbol{\lambda} = \mathbf{c}_B, \quad N^T\boldsymbol{\lambda} + \mathbf{s}_N = \mathbf{c}_N$$

$$\boldsymbol{\lambda} = B^{-T}\mathbf{c}_B, \quad \mathbf{s}_N = \mathbf{c}_N - N^T\boldsymbol{\lambda} = \mathbf{c}_N - \left(B^{-1}N\right)^T\mathbf{c}_B$$

Reduction in the objective function for non-optimal point is

$$\mathbf{c}^T\mathbf{x} = \mathbf{b}^T\boldsymbol{\lambda} + \mathbf{s}^T\mathbf{x} = \mathbf{b}^T\boldsymbol{\lambda} + \mathbf{s}_N^T\mathbf{x}_N$$

Simplex Method



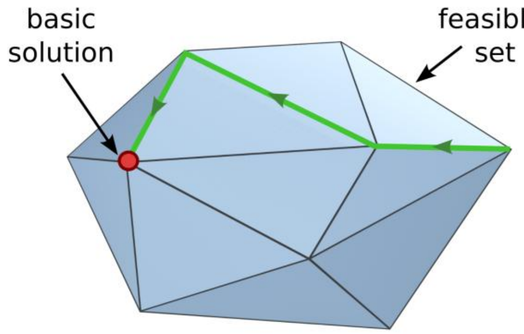
KKT Conditions

$$\begin{aligned} A^T\boldsymbol{\lambda} + \mathbf{s} &= \mathbf{c} \\ A\mathbf{x} - \mathbf{b} &= 0 \\ \mathbf{x} &\geq 0 \\ \mathbf{s} &\geq 0 \\ s_i x_i &= 0. \end{aligned}$$

Simplex Method: Example

Example

$$\min\ -4x_1 - 2x_2 \quad \text{subject to}$$
$$x_1 + x_2 + x_3 = 5,$$
$$2x_1 + (1/2)x_2 + x_4 = 8,$$
$$x \geq 0.$$



Iterations of Simplex Method

$$\mathcal{B} = \{3, 4\}$$
$$x_B = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad s_N = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$$

$$\mathcal{B} = \{3, 1\} \text{ and } \mathcal{N} = \{4, 2\};$$
$$x_B = \begin{bmatrix} x_3 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \lambda = \begin{bmatrix} 0 \\ -3/2 \end{bmatrix}, \quad s_N = \begin{bmatrix} s_4 \\ s_2 \end{bmatrix} = \begin{bmatrix} 3/2 \\ -5/4 \end{bmatrix}$$

$$\mathcal{B} = \{2, 1\} \text{ and } \mathcal{N} = \{4, 3\}$$
$$x_B = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 4/3 \\ 11/3 \end{bmatrix}, \quad \lambda = \begin{bmatrix} -5/3 \\ -2/3 \end{bmatrix}, \quad s_N = \begin{bmatrix} s_4 \\ s_3 \end{bmatrix} = \begin{bmatrix} 7/3 \\ 5/3 \end{bmatrix}$$

Wright 1999

solution: $c^T x = -41/3 \qquad s_N \geq 0$

Simplex Method Steps (Non-Degenerate Case):

Start with $\mathcal{B}$, B, $\mathbf{x}_B = B^{-1}\mathbf{b} \geq 0$, $\mathbf{x}_N = 0$.

1. Solve for KKT triple $(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{s})$, $B^T\boldsymbol{\lambda} = \mathbf{c}_B$, $\mathbf{s}_N = \mathbf{c}_N - N^T\boldsymbol{\lambda}$.
2. **If** $\mathbf{s}_N \geq 0$ then
   halt: (KKT triple is valid and $\mathbf{x}_B$ is optimal) ■.
3. Determine index $q \in \mathcal{B}$ with most negative $s_q < 0$, (entering index is q).
4. Solve for $\mathbf{v}$ in $B\mathbf{v} = A_q$.
5. **If** $v_i \leq 0$ for all i then
   halt: (LP is unbounded) ■.
6. Compute $p = \arg\min_{i\ v_i>0}\ x_{B,i}/v_i$, the $x_q^+ = x_{B,p}/v_p$ (exiting index is p).
7. Construct new basis set $\mathcal{B}^+ = (\mathcal{B} \setminus \{p\}) \bigcup \{q\}$, matrix B^+, and basic feasible point $\mathbf{x}_B^+$.
8. Repeat from step 1.

Tableau notation

	a_1	a_2	$\cdots$	a_n	b
	y_{11}	y_{12}	$\cdots$	y_{1n}	y_{01}
	$\vdots$			$\vdots$	$\vdots$
	y_{m1}	y_{m2}	$\cdots$	y_{mn}	y_{0m}
r^T	s_1	s_2	$\cdots$	s_n	$-\mathbf{c}^T\mathbf{x}_B$

Used for bookeeping key terms.
(details on the next slides)

Simplex Method: Tableau Notation

We construct the following tableau matrix for a basic feasible point $x=(x_B,x_N)$ and $\mathscr{B}$.

System Tableau

	a_1	a_2	$\cdots$	a_n	b
	a_{11}	a_{12}	$\cdots$	a_{1n}	b_1
	$\vdots$			$\vdots$	$\vdots$
	a_{m1}	a_{m2}	$\cdots$	a_{mn}	b_m
c^T	c_1	c_2	$\cdots$	c_n	0

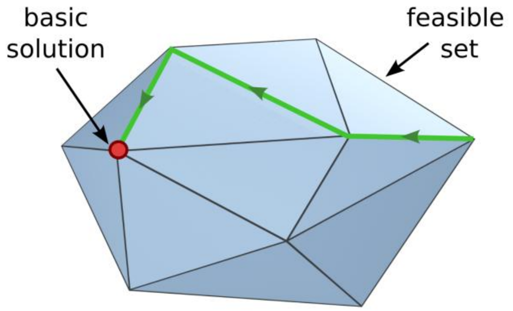
$$\rightarrow \begin{bmatrix} A & \mathbf{b} \\ \mathbf{c}^T & 0 \end{bmatrix} = \begin{bmatrix} B & N & \mathbf{b} \\ \mathbf{c}_B^T & \mathbf{c}_N^T & 0 \end{bmatrix}$$ multiply on left,

$$\begin{bmatrix} B^{-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B & N & \mathbf{b} \\ \mathbf{c}_B^T & \mathbf{c}_N^T & 0 \end{bmatrix} = \begin{bmatrix} I_m & B^{-1}N & B^{-1}\mathbf{b} \\ \mathbf{c}_B^T & \mathbf{c}_N^T & 0 \end{bmatrix}$$

$$= \begin{bmatrix} I_m & [y_{ji}]_{i\notin\mathcal{B}} & [y_{j0}] \\ \mathbf{c}_B^T & \mathbf{c}_N^T & 0 \end{bmatrix}$$

$$\begin{bmatrix} I_m & 0 \\ -\mathbf{c}_B^T & 1 \end{bmatrix} \begin{bmatrix} I_m & B^{-1}N & B^{-1}\mathbf{b} \\ \mathbf{c}_B^T & \mathbf{c}_N^T & 0 \end{bmatrix} = \begin{bmatrix} I_m & B^{-1}N & B^{-1}\mathbf{b} \\ 0^T & \mathbf{c}_N^T - \mathbf{c}_B^T B^{-1}N & -\mathbf{c}_B^T B^{-1}\mathbf{b} \end{bmatrix}$$

Simplex Method



Canonical Tableau

	a_1	a_2	$\cdots$	a_n	b
	y_{11}	y_{12}	$\cdots$	y_{1n}	y_{01}
	$\vdots$			$\vdots$	$\vdots$
	y_{m1}	y_{m2}	$\cdots$	y_{mn}	y_{0m}
r^T	s_1	s_2	$\cdots$	s_n	$-\mathbf{c}^T\mathbf{x}_B$

Canonical Tableau

$$\begin{bmatrix} I_m & B^{-1}N & B^{-1}\mathbf{b} \\ 0^T & \mathbf{s}_N^T & -\mathbf{c}_B^T B^{-1}\mathbf{b} \end{bmatrix}$$

$$\mathbf{s}_N^T = \mathbf{c}_N^T - \mathbf{c}_B^T B^{-1}N$$

Update to new basis

$$A_q = \sum_{i=1}^m y_{iq}A_i = \sum_{i=1,i\neq p}^m y_{iq}A_i + y_{pq}A_p \Rightarrow A_p = \frac{1}{y_{pq}}A_q - \sum_{i=1,i\neq p}^m \frac{y_{iq}}{y_{pq}}A_i$$

$$\mathbf{u} = \sum_{i=1}^m u_iA_i = \sum_{i=1,i\neq p}^m u_iA_i + u_pA_p = \sum_{i=1,i\neq p}^m \left(u_i - \frac{y_{iq}}{y_{pq}}u_p\right)A_i + \frac{u_p}{y_{pq}}A_q$$

Update to new basis

$$y_{ij}^+ = y_{ij} - \frac{y_{iq}}{y_{pq}}y_{pj}, \quad \text{if } i \neq p$$

$$y_{pj}^+ = \frac{y_{pj}}{y_{pq}}, \quad \text{if } i = p,$$

$$\mathbf{y}_i = B^{-1}A_i$$
$$\mathbf{y}_0 = B^{-1}\mathbf{b}$$

Simplex Method Example: Canonical Tableau

Update

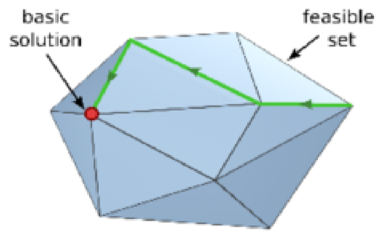
$$y_{ij}^{+} = y_{ij} - \frac{y_{iq}}{y_{pq}} y_{pj}, \quad \text{if } i \neq p$$

$$y_{pj}^{+} = \frac{y_{pj}}{y_{pq}}, \quad \text{if } i = p,$$

$$\mathbf{y}_0 = \mathbf{x}_B = B^{-1}\mathbf{b}, \quad \mathbf{y}_i = B^{-1}A_i$$

$$s_N^T = c_N^T - c_B^T B^{-1} N.$$

Simplex Method



Example

$$\min \quad -7x_1 - 6x_2$$

subject

$$2x_1 + x_2 + x_3 = 3$$
$$x_1 + 4x_2 + x_4 = 4$$
$$x_1, x_2, x_3, x_4 \geq 0.$$

initial-canonical

	a_1	a_2	a_3	a_4	b
	2	1	1	0	3
	1	4	0	1	4
r^T	−7	−6	0	0	0

$\mathcal{B} = \{3, 4\}$, $q = 1$

$\frac{y_{i0}}{y_{iq}} = [\frac{3}{2}, 4]$, $i = [1, 2]$, $p = 1$

→

iteration 1

	a_1	a_2	a_3	a_4	b
	1	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{3}{2}$
	0	$\frac{7}{2}$	$-\frac{1}{2}$	1	$\frac{5}{2}$
r^T	0	$-\frac{5}{2}$	$\frac{7}{2}$	0	$\frac{21}{2}$

$\mathcal{B} = \{1, 4\}$, $q = 2$

$\frac{y_{i0}}{y_{iq}} = [3, \frac{5}{7}]$, $i = [1, 2]$, $p = 2$

→

iteration 2

	a_1	a_2	a_3	a_4	b
	1	0	$\frac{4}{7}$	$-\frac{1}{7}$	$\frac{8}{7}$
	0	1	$-\frac{1}{7}$	$\frac{2}{7}$	$\frac{5}{7}$
r^T	0	0	$\frac{22}{7}$	$\frac{5}{7}$	$\frac{86}{7}$

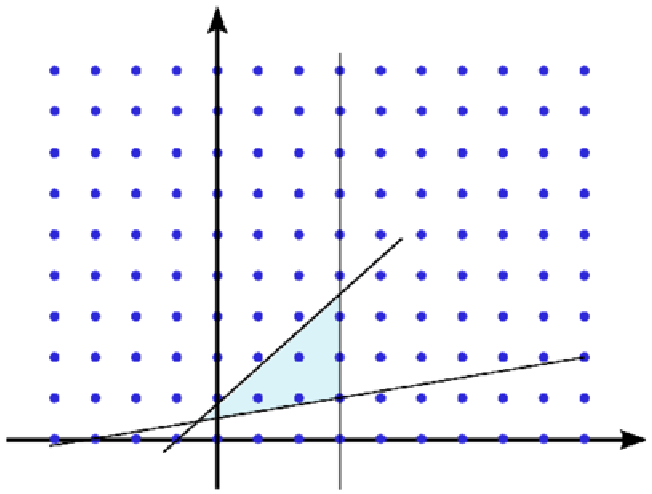
$\mathcal{B} = \{1, 2\}$,

(final)

→

solution:
$\mathbf{x} = [\frac{8}{7}, \frac{5}{7}, 0, 0].$

Simplex Method Example: Canonical Tableau



Update

$$y_{ij}^{+} = y_{ij} - \frac{y_{iq}}{y_{pq}} y_{pj}, \quad \text{if } i \neq p$$

$$y_{pj}^{+} = \frac{y_{pj}}{y_{pq}}, \quad \text{if } i = p,$$

$$s_N^T = c_N^T - c_B^T B^{-1} N.$$

Example

$$\begin{aligned} \text{min} \quad & -2x_1 - 3x_2 \\ \text{subject} \quad & x_1 + x_3 = 3 \\ & x_1 - 6x_2 + x_4 = -3 \\ & -9x_1 + 8x_2 + x_5 = 1 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{aligned}$$

initial-canonical

	a_1	a_2	a_3	a_4	a_5	b
	1	0	1	0	0	3
	1	−6	0	1	0	−3
	−9	8	0	0	1	1
r^T	−2	−3	0	0	0	0

$\mathcal{B} = \{3, 4, 5\}$, $q = 2$

$\frac{y_{i0}}{y_{iq}} = [\frac{1}{8}]$, $i = [3]$, $p = 3$

→

iteration 1

	a_1	a_2	a_3	a_4	a_5	b
	1	0	1	0	0	3
	$-\frac{23}{4}$	0	0	1	$\frac{3}{4}$	$-\frac{9}{4}$
	$-\frac{9}{8}$	1	0	0	$\frac{1}{8}$	$\frac{1}{8}$
r^T	$-\frac{43}{8}$	0	0	0	$\frac{3}{8}$	$\frac{3}{8}$

$\mathcal{B} = \{3, 4, 2\}$, $q = 1$

$\frac{y_{i0}}{y_{iq}} = [3]$, $i = [1]$, $p = 1$

→

iteration 2

	a_1	a_2	a_3	a_4	a_5	b
	1	0	1	0	0	3
	0	0	$\frac{23}{4}$	1	$\frac{3}{4}$	15
	0	1	$\frac{9}{8}$	0	$\frac{1}{8}$	$\frac{7}{2}$
r^T	0	0	$\frac{43}{8}$	0	$\frac{3}{8}$	$\frac{33}{2}$

$\mathcal{B} = \{1, 4, 2\}$,

(final)

→

solution:

$$\mathbf{x} = \left[3, \frac{7}{2}, 0, 15, 0\right].$$

Simplex Method: Two-Phase Method

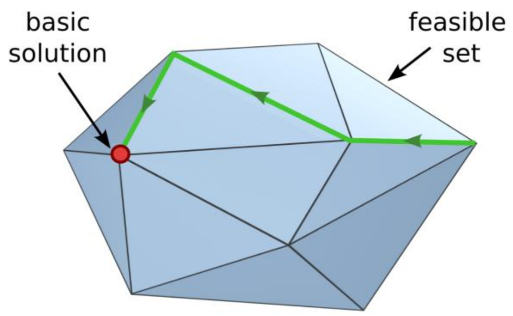
Linear Program

$$\begin{array}{ll} \min & \mathbf{c}^T \mathbf{x} \\ \text{subject} & A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0. \end{array}$$

Artificial Problem

$$\begin{array}{ll} \min & y_1 + y_2 + \cdots + y_m \\ \text{subject} & [A, I_m][\mathbf{x};\mathbf{y}] = \mathbf{b} \\ & [\mathbf{x};\mathbf{y}] \geq 0. \end{array}$$

Simplex Method



Finding initial basic feasible points for LP can be difficult.

Two-Phase Method:

phase I:
- construct an artificial LP that has easy to find initial feasible point.
- minimize the artificial LP problem to find initial feasible point to the original LP.

phase II:
- use solution from phase I for starting value
- minimize the original LP problem

Artificial Problem always has solution

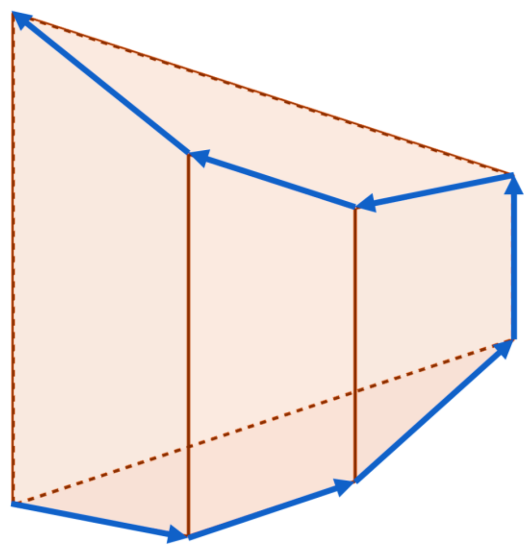
$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{b} \end{bmatrix},\ \mathbf{b} \geq 0.$$

Simplex Method: Exponential Number of Steps Example

Example (Klee-Minty Cube):

$$\begin{aligned} \text{maximize} \quad & x_d \\ \text{subject to} \quad & 0 \leq x_1 \leq 1 \\ & \varepsilon x_1 \leq x_2 \leq 1 - \varepsilon x_1 \\ & \varepsilon x_2 \leq x_3 \leq 1 - \varepsilon x_2 \\ & \vdots \\ & \varepsilon x_{d-1} \leq x_d \leq 1 - \varepsilon x_{d-1}. \end{aligned}$$

Klee-Minty Cube



Shuiberts 2023

$\epsilon = 1/3$, d = 3

Example: d variables and 2d inequality constraints, require $\epsilon \in (0,\frac{1}{2})$.

Simplex Method:

Start at x = 0.

Simplex Method using **Dantzig's rules visits all vertices 2^d!**

Gives **exponential number of steps** in d!

In practice, most problems exhibit convergence in polynomial number of steps in m (typically linear).

Challenge: simplex methods trace the boundary geometry of the feasible set.

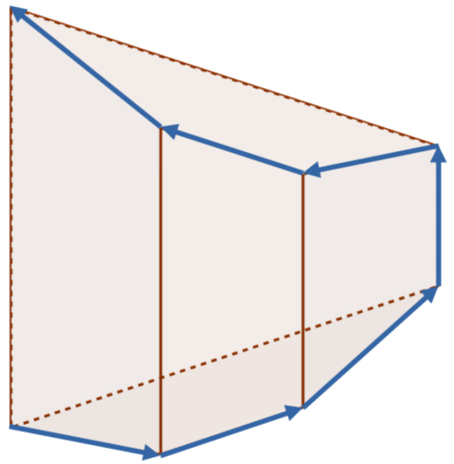
Alternatives: develop methods that approach the solution from outside or inside of the feasible set (avoid the boundary).

Interior Point Methods do this by using penalty methods and central path to approach from inside.

IPMs: can solve LPs in polynomial number of steps.

Solvers for Linear Programming beyond Simplex Methods

Klee-Minty Cube



Shuiberts 2023

$$\begin{aligned} \text{maximize} \quad & x_d \\ \text{subject to} \quad & 0 \leq x_1 \leq 1 \\ & \varepsilon x_1 \leq x_2 \leq 1 - \varepsilon x_1 \\ & \varepsilon x_2 \leq x_3 \leq 1 - \varepsilon x_2 \\ & \quad \vdots \\ & \varepsilon x_{d-1} \leq x_d \leq 1 - \varepsilon x_{d-1}. \end{aligned}$$

Simplex Method in some cases can be inefficient using **Dantzig's rules, since visits all vertices 2^d!**

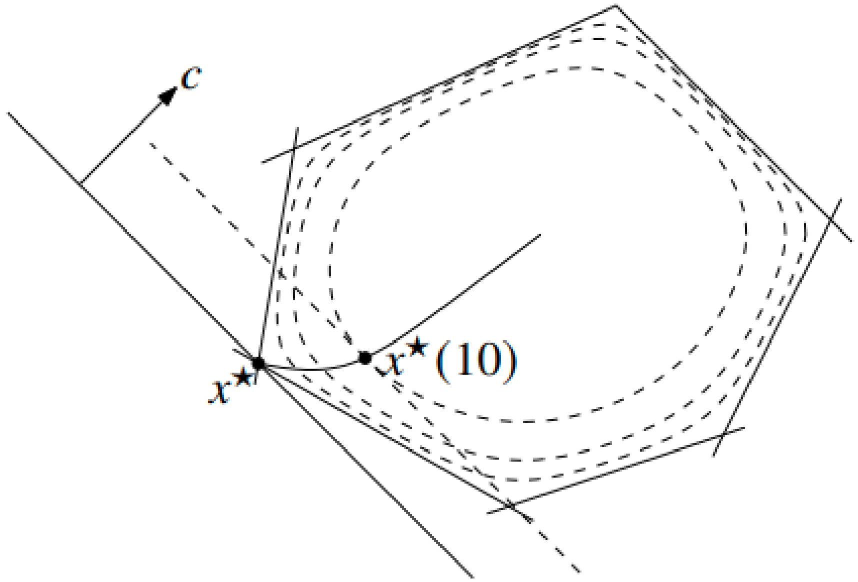
Gives exponential number of steps in d!

For many constraints the geometry of the feasible domain can have complex boundary making vertex traversal inefficient.

Alternative methods for LPs include

- **Ellipsoid Method** (theoretically weak polynomial / but inefficient in practice from ill-conditioning).
- **Interior Point Methods** (uses penalty barrier methods) (future lectures).
- **Primal-Dual Methods**, and others.

Interior Point Methods



Boyd & VandenBerghes

Interior Point Methods: start with an initial feasible point (need not be basic) and optimizes LP + possible penalties. For example, *affine scaling* or

$$\min_{\mathbf{x} \in C} f(\mathbf{x}) + \epsilon g(\mathbf{x}) \qquad g(\mathbf{x}) = \sum_{i=1}^{m} -\log(b_i - a_i \mathbf{x})$$

For good choice of iterative methods, and schedule for penalty ε and accuracy δ , one can show **weak polynomial time complexity.**

Simplex Method and Interior Point Methods central in current software for solving large LP problems.

More on this in upcoming lectures.