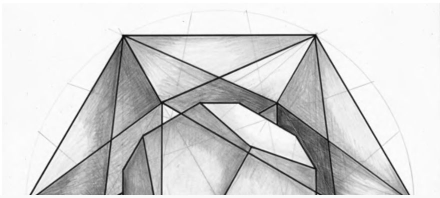
LINEAR OPTIMIZATION

Master Spécialisé OSE 2023 – Mines Paris - PSL

Sophie Demassey (CMA, Mines Paris - PSL)
sophie.demassey@minesparis.psl.eu `http://sofdem.github.io`

Overview



INTRODUCTION

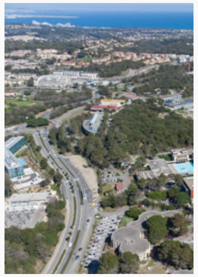
DECISION IS OPTIMIZATION

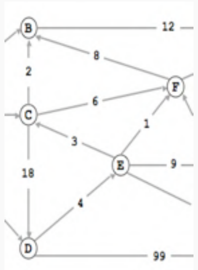
select the **best** of all **possible** alternatives – the solutions –
regarding a quantitative criterion – the objective.

time:	path with minimum travel duration, schedule with minimum total lateness
space:	path with minimum travel distance, layout with minimum wasted space
money:	design with minimum cost, operation with maximum profit
goods:	design with maximum production, operation with minimum energy consumption
choice:	work schedule for maximum satisfaction
quantity:	state of minimum potential energy (equilibrium)

MODELING FOR SOLVING

A mathematical optimization model is
an abstract representation of the problem solutions,
not explicitly as a list, a dataset, but implicitly as
relationships between **unknowns (real-valued) functions** over **(real-valued) variables**





$$\min z = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} d_{ij} x_{ij}$$

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} x_{ij} = 1, \quad \forall i \in N$$

$$\sum_{\substack{i=1 \\ i \neq j}}^{n} x_{ij} = 1, \quad \forall j \in N$$

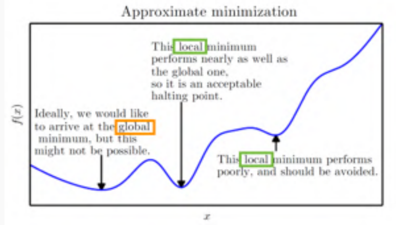
$$\min \{ f(x) \mid g(x) \leq 0,\ x \in \mathbb{R}^n \}$$

with $f : \mathbb{R}^n \to \mathbb{R}$ in the objective: the function to minimize
and $g : \mathbb{R}^n \to \mathbb{R}^m$ in the constraints: the relations to satisfy.

SOLUTIONS: THEORY VS PRACTICE

feasibility ?
- models are **approximate** (*e.g., abstract routes*)
- data are **uncertain** (*e.g., forecast travel times*)
- data are **truncated** (*floating-point numerical errors*)

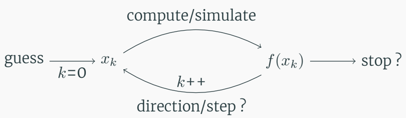
optimality ?
- finite time complexity $\neq$ **reachable** (*e.g.* 2^{90} *operations*)
- provable within a **gap tolerance** ($f(x) \leq f(y) + \epsilon,\ \forall y$)
- provable **locally** vs globally ($f(x) \leq f(y),\ \forall y \in V(x)$)

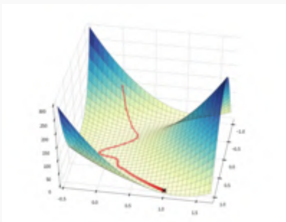


SOLVING METHODS

analytical methods come from a **provable theory**, e.g.:
- *min* $x^2 - 4x + 3, x \in [0, 5]$ (*Fermat, derivative*)
- shortest path in a graph (*Dijkstra, Bellman*)

numerical methods **evaluate** $f(x_k)$ **iteratively** at trial points (x_k)
1st- or 2nd-order methods if driven by $f'(x_{k-1})$ or $f''(x_{k-1})$
derivative-free otherwise





DIFFERENT TECHNIQUES FOR DIFFERENT CLASSES OF MODELS

- **with** or without constraints
- **single** or multiple objectives
- **fixed** or uncertain data
- **analytic** or logic or graphic models
- **linear** or convex or nonconvex functions
- **smooth** or nonsmooth functions
- **continuous** or **discrete** decisions

APPLICATIONS

operational research : operation, design and plan (routing, scheduling, packing, cutting, rostering, allocating) of physical/economical systems in logistics, energy, finance, etc.

optimal control : command $u(t)$ to optimize trajectory $x(t)$ s.t. $x'(t) = g(x(t), u(t))$

machine learning : find a best model/data match (e.g. a linear fit)

artificial intelligence : machines decide when they don't dream of electric sheeps

game theory : multiple players, conflicting goals, best respective strategies

MATHEMATICAL PROGRAMMING

programming = **planning** (military/industrial) operations

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{subject to } g(x) \geq 0 \\ &x \in \mathbb{R}^n \end{aligned}$$

- x: the decision variables
- $f : \mathbb{R}^n \to \mathbb{R}$: the objective function. Note: maximize $f \equiv -$ minimize $(-f)$
- $g : \mathbb{R}^n \to \mathbb{R}^m$: the constraints. Note: $g(x) \leq 0 \equiv -g(x) \geq 0$

solution/assignment $X \in \mathbb{R}^n$
feasible solution $X \in g^{-1}(\mathbb{R}^m_+)$
optimal solution $X \in \arg\min\{f(x) : g(x) \geq 0, x \in \mathbb{R}^n\}$

LINEAR PROGRAM

a mathematical program $\min \{f(x) | g(x) \geq 0, x \in \mathbb{R}^n\}$
with **linear** functions in constraints and objective:
$\min \{c^T x | Ax + b \geq 0, x \in \mathbb{R}^n\}, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$

Example: $n = 3$, $m = 2$,

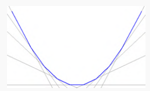
$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, c = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 5 & 3 & -2 \\ 1 & 1 & 1 \end{pmatrix}, b = \begin{pmatrix} -4 \\ 1 \end{pmatrix}$$

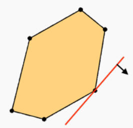
$$\begin{aligned} &\min x_1 \\ \text{s.t.}\ & 5x_1 + 3x_2 - 2x_3 \geq 4 \\ & x_1 + x_2 + x_3 \geq -1 \\ & x_1, x_2, x_3 \in \mathbb{R} \end{aligned}$$

- This is the "**and**": feasible solutions (x_1, x_2, x_3) satisfy **all** constraints
- $x \mapsto 5x^2$, $(x, y) \mapsto 3xy$ are not linear (but quadratic)

HOW RELEVANT IS LP ?

- **broad applicability:**
  format for practical decision problems,
  approximation for convex problems,
  basis for nonconvex/logic problems
  (with discrete variables)
- **easy to solve:**
  polynomial-time algorithms,
  efficient practical algorithms
  (e.g. restart, partial model),
  nice properties: strong duality





EX 1: NUCLEAR WASTE MANAGEMENT

A company eliminates nuclear wastes of 2 types A and B, by applying a sequence of 3 processes I, II and III in any order. The processes I, II, III, have limited availability, respectively: 450h, 350h, and 200h per month. The unit processing times depend on the process and waste type, as reported in the following table:

process	I	II	III
waste A	1h	2h	1h
waste B	3h	1h	1h

The profit for the company is 4000 euros to eliminate one unit of waste A and 8000 euros to eliminate one unit of waste B.

Objective: maximize the profit.

HOW TO MODEL ?

1. decision variables: what a solution is made of ?
2. constraints: what is a feasible solution ?
3. objective: what is an optimal solution ?
4. check the units or convert
5. check LP format (linear, continuous, non-strict inequalities) or reformulate

EX 1: NUCLEAR WASTE MANAGEMENT – LP MODEL

- decision variables ?
  - x_A, x_B the fraction of units of waste of type A or B to process each month
- constraints and objective ?
  - definition domain of the variables (nonnegative)
  - limited availability (in h/month) for each process
  - maximize revenue (in keuros)

$$\begin{aligned} \max\ & 4x_A + 8x_B \\ \text{s.t.}\ & x_A + 3x_B \leq 450 \\ & 2x_A + x_B \leq 350 \\ & x_A + x_B \leq 200 \\ & x_A, x_B \geq 0 \end{aligned}$$

EX 2: PETROLEUM DISTILLATION

The two crude petroleum problem [Ralphs]

A petroleum company distills crude imported from Kuwait (9000 barrels available at 20€ each) and from Venezuela (6000 barrels available at 15€ each), to produce gasoline (2000 barrels), jet fuel (1500 barrels), and lubricant (500 barrels) in the following proportions:

	gasoline	jet fuel	lubricant
Kuwait	0.3	0.4	0.2
Venezuela	0.4	0.2	0.3

(first column reads: *producing 1 unit of gasoline requires 0.3 units of crude from Kuwait and 0.4 from Venezuela*)

Objective: minimize the production cost.

EX 2: PETROLEUM DISTILLATION – LP MODEL

- decision variables ?
  - x_K, x_V the quantity (in thousands of barrels) to import from Kuwait or from Venezuela
- constraints and objective ?
  - availability for each crude, distillation balance for each product, production costs

$$
\begin{aligned}
\min\ & 20x_K + 15x_V \\
\text{s.t.}\ & 0.3x_K + 0.4x_V \geq 2 \\
& 0.4x_K + 0.2x_V \geq 1.5 \\
& 0.2x_K + 0.3x_V \geq 0.5 \\
& 0 \leq x_K \leq 9 \\
& 0 \leq x_V \leq 6
\end{aligned}
$$

NOTE ON MODELLING

linearly equivalent formulations:

$$
\begin{array}{ll}
\max f & -\min(-f) \\
ax \leq b & -ax \geq -b \\
ax = b & ax \geq b \text{ and } ax \leq b \\
ax \leq b & ax + s = b \text{ and } s \geq 0 \\
x \in \mathbb{R} & x = y - z,\ y \geq 0,\ z \geq 0
\end{array}
$$

LINEAR PROGRAM IN STANDARD FORM

equality constraints and **nonnegative** variables:

$$
\begin{aligned}
\min\ & c^T x \\
\text{s.t.}\ & Ax = b \\
& x \geq 0
\end{aligned}
$$

$$
\begin{aligned}
\min\ & \sum_{j=1}^{n} c_j x_j \\
\text{s.t.}\ & \sum_{j=1}^{n} a_{ij} x_j = b_i, & \forall i = 1, \dots, m \\
& x_j \geq 0 & \forall j = 1, \dots, n
\end{aligned}
$$

with $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$

REDUCTION TO STANDARD FORM

Every linear program

$$\min\{c^T x | Ax \geq b, x \in \mathbb{R}^n\}$$

can be transformed into an equivalent problem in standard form

$$\min\{d^T y | Ey = f, y \in \mathbb{R}^p_+\}$$

$$
\begin{aligned}
\min\ & x_1 \\
\text{s.t.}\ & 5x_1 - 3x_2 \geq 4 \\
& x_1 + x_2 \geq -1 \\
& x_1, x_2 \in \mathbb{R}
\end{aligned}
$$

$$
\begin{aligned}
\min\ & (x_1^+ - x_1^-) \\
\text{s.t.}\ & 5(x_1^+ - x_1^-) - 3(x_2^+ - x_2^-) - z_1 = 4 \\
& (x_1^+ - x_1^-) + (x_2^+ - x_2^-) - z_2 = -1 \\
& x_1^+, x_1^-, x_2^+, x_2^-, z_1, z_2 \geq 0
\end{aligned}
$$

REDUCTION TO STANDARD FORM (RECIPE)

replace by		
negative variable	$x \leq 0$	$x = -z, z \geq 0$
free variable	y free	$y = y^+ - y^-, y^+, y^- \geq 0$
slack constraint	$Ax \geq b$	$Ax - s = b$, $s \geq 0$
slack constraint	$Ey \leq f$	$Ey + u = f$, $u \geq 0$
maximization	$max\ cx$	$-min(-c)x$

$$\begin{aligned} \max\ & c^T x + d^T y \\ \text{s.t.}\ & Ax \geq b \\ & Ey \leq f \\ & x \leq 0, y\ free \end{aligned}$$

$$\begin{aligned} \min\ & (-c)^T(-z) + (-d)^T(y^+ - y^-) \\ \text{s.t.}\ & A(-z) - s = b \\ & E(y^+ - y^-) + u = f \\ & z, y^+, y^-, s, u \geq 0 \end{aligned}$$

EX: NUCLEAR WASTE MANAGEMENT – LP STANDARD FORM

$$\begin{aligned} \max\ & 4x_A + 8x_B \\ \text{s.t.}\ & x_A + 3x_B \leq 450 \\ & 2x_A + x_B \leq 350 \\ & x_A + x_B \leq 200 \\ & x_A, x_B \geq 0 \end{aligned}$$

$$\begin{aligned} -\min\ & -4x_A - 8x_B \\ \text{s.t.}\ & x_A + 3x_B + s_1 = 450 \\ & 2x_A + x_B + s_2 = 350 \\ & x_A + x_B + s_3 = 200 \\ & x_A, x_B, s_1, s_2, s_3 \geq 0 \end{aligned}$$

EX: PETROLEUM DISTILLATION – LP STANDARD FORM

$$\begin{aligned} \min\ & 20x_K + 15x_V \\ \text{s.t.}\ & 0.3x_K + 0.4x_V \geq 2 \\ & 0.4x_K + 0.2x_V \geq 1.5 \\ & 0.2x_K + 0.3x_V \geq 0.5 \\ & 0 \leq x_K \leq 9 \\ & 0 \leq x_V \leq 6 \end{aligned}$$

$$\begin{aligned} \min\ & 20x_K + 15x_V \\ \text{s.t.}\ & 0.3x_K + 0.4x_V - s_G = 2 \\ & 0.4x_K + 0.2x_V - s_J = 1.5 \\ & 0.2x_K + 0.3x_V - s_L = 0.5 \\ & x_K + s_K = 9 \\ & x_V + s_V = 6 \\ & x_k, x_V, s_G, s_J, s_L, s_K, s_V \geq 0 \end{aligned}$$

LINEAR ALGEBRA REVIEW AND NOTATION (1)

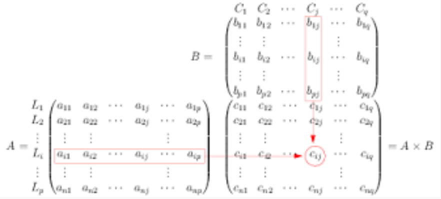
matrix $A \in \mathbb{R}^{m \times n}$ with entry a_{ij} in row $1 \leq i \leq m$, column $1 \leq j \leq n$

transpose $A^T \in \mathbb{R}^{n \times m}$ with $a^T_{ji} = a_{ij}$

(column) vector $a \in \mathbb{R}^n \equiv \mathbb{R}^{n \times 1}$

scalar product $a, b \in \mathbb{R}^n$, $\langle a, b \rangle = a^T b = b^T a = \sum_{j=1}^n a_j b_j$

matrix product $A \in \mathbb{R}^{m \times p}$, $B \in \mathbb{R}^{p \times n}$, $C = AB \in \mathbb{R}^{m \times n}$ with $c_{ij} = \sum_{k=1}^p a_{ik} b_{kj}$.
matrix product is associative $(AB)C = A(BC)$ and $(AB)^T = B^T A^T$



LINEAR ALGEBRA REVIEW AND NOTATION (2)

linear combination $\sum_{i=1}^{p} \lambda_i x^i \in \mathbb{R}^n$
of vectors $x^1, \ldots, x^p \in \mathbb{R}^n$ with scalars $\lambda_1, \ldots, \lambda_p \in \mathbb{R}$

linearly independence $\sum_{i=1}^{p} \lambda_i x^i = 0 \Rightarrow \lambda_1 = \cdots = \lambda_p = 0$

vector-space span $V = \{\sum_{i=1}^{p} \lambda_i x^i \mid \lambda_1, \ldots, \lambda_p \in \mathbb{R}\} \subseteq \mathbb{R}^n$

dimension $dim(V) = p$ if $x^1, \ldots, x^p$ are linearly independent, i.e. form a basis for V

row space of $A \in \mathbb{R}^{m \times n}$ span of the rows $rs_A = \{\lambda^T A, \lambda \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$

column space of $A \in \mathbb{R}^{m \times n}$ span of the columns $cs_A = \{A\lambda, \lambda \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$

rank of $A \in \mathbb{R}^{m \times n}$: $rk_A = dim(rs_A) = dim(cs_A) \leq \min(m, n)$

READING:

to go further:
read [Bertsimas-Tsitsiklis]:
Section 1.1

for the next class:
read [Bertsimas-Tsitsiklis]:
Section 1.5: Linear algebra background

ALGEBRA OF LINEAR PROGRAMMING

A LP in standard form with $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ has $m + n$ constraints:

$$\begin{aligned} \min\ & c^T x & \\ \text{s.t.}\ & Ax = b & (m) \\ & x \geq 0 & (n) \end{aligned}$$

A feasible solution $\equiv$ non-negative coefficients forming b as a linear combination of the columns of A:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

MODELING LPs

HOW TO MODEL ?

1. decision variables: what a solution is made of ?
2. constraints: what is a feasible solution ?
3. objective: what is an optimal solution ?
4. check the units or convert
5. check LP format (linear, continuous, non-strict inequalities) or reformulate

EX 3: NETWORK FLOW

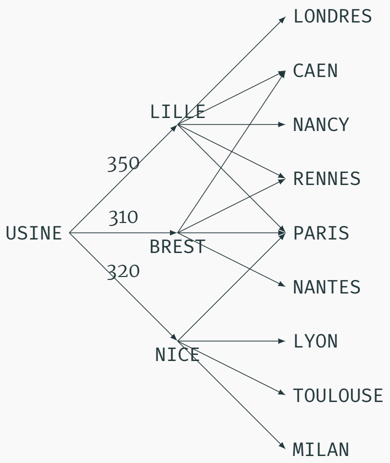
network flow

A company delivers retail stores in 9 cities in Europe from its unique factory *USINE*.
How to manage production and transportation in order to:

- meet the demand of each store,
- not exceed the production limit,
- not exceed the line capacities,
- minimize the transportation costs ?

```
demand = {
    'PARIS': 110,
    'CAEN': 90,
    'RENNES': 60,
    'NANCY': 90,
    'LYON': 80,
    'TOULOUSE': 50,
    'NANTES': 50,
    'LONDRES': 70,
    'MILAN': 70
}
LINES, unitary_cost, capacity = multidict({
    ('USINE','LILLE'): [2.9, 350],
    ('USINE','NICE') : [3.5, 320],
    ('USINE','BREST') : [3.1, 310],
    ('LILLE', 'PARIS') : [1.1, 150],
    ('LILLE','CAEN'): [0.7, 150],
    ('LILLE','RENNES'): [1.0, 150],
    ('LILLE','NANCY'): [1.3, 150],
    ('LILLE','LONDRES'): [1.3, 150],
    ('NICE','LYON'): [0.8, 200],
    ('NICE','TOULOUSE'): [0.2, 110],
    ('NICE','PARIS'): [1.3, 100],
    ('NICE','MILAN'): [1.3, 150],
    ('BREST','NANTES'): [0.9, 150],
    ('BREST','CAEN'):[0.8, 200],
    ('BREST','RENNES'): [0.8, 150],
    ('BREST','PARIS'): [0.9, 100]
})
MAX_PRODUCTION = 900
```


EX 3: GRAPH MODEL



- find a flow on a capacitated directed graph
- flow conservation at each node: IN=OUT

EX 3: LP MODEL

- x_ℓ the quantity of products transported on line $\ell = (i, j) \in$ LINES
- TRANSITS$= \{$LILLE,NICE,BREST$\}$

$$
\begin{aligned}
\min \quad & \sum_{\ell \in \text{LINES}} \text{COST}_\ell x_\ell \\
\text{s.t.} \quad & \sum_{i \in \text{TRANSITS}} x_{(\text{USINE},i)} \leq \text{MAXPROD} \\
& \sum_{i \in \text{TRANSITS}} x_{(i,j)} \geq \text{DEMAND}_j, && \forall j \in \text{STORES} \\
& x_{(\text{USINE},i)} = \sum_{j \in \text{STORES}} x_{(i,j)}, && \forall i \in \text{TRANSITS} \\
& 0 \leq x_\ell \leq \text{CAPACITY}_\ell, && \forall \ell \in \text{LINES}.
\end{aligned}
$$

EX 4: MINIMUM DISTANCE

minimize L^1 and L^∞ norms

Find a solution $x \in \mathbb{R}^n$ of the system of equation $Ax = b$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ of minimum

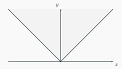
- L^1 norm:

$$\|x\|_1 = \sum_{j=1,\dots,n} |x_j|$$

- L^∞ norm:

$$\|x\|_\infty = \max_{j=1,\dots,n} |x_j|$$

EX 4: LP MODELS $\min \|x\|_1 = \min \sum_j |x_j|$

how to model $|x|$, $x \in \mathbb{R}$? 

variable splitting:

$$|x| = \min\{x^+ + x^- \mid x = x^+ - x^-, x^+, x^- \geq 0\}$$

$$\begin{aligned} \min \sum_{j=1}^{n} (x_j^+ + x_j^-) & \\ \text{s.t.} \quad Ax = b, & \\ x_j = x_j^+ - x_j^-, & \quad \forall j \\ x_j^+, x_j^- \geq 0, & \quad \forall j \end{aligned}$$

supporting plane model:

$$|x| = \max\{x, -x\} = \min\{y \mid y \geq x, y \geq -x\}$$

$$\begin{aligned} \min \sum_{j=1}^{n} y_j & \\ \text{s.t.} \quad Ax = b, & \\ y_j \geq x_j, & \quad \forall j \\ y_j \geq -x_j, & \quad \forall j \end{aligned}$$

Note that $\min \sum |x_j| = \sum \min |x_j|$ because $|x_j| \geq 0$

EX 4: LP MODEL $\min \|x\|_\infty = \min \max_j |x_j|$

- $y \geq |x_j| \iff y \geq x_j \wedge y \geq -x_j$
- $y \geq \max_j |x_j| \iff y \geq x_j \wedge y \geq -x_j \ (\forall j)$

$$\begin{aligned} \min y & \\ \text{s.t.} \quad Ax = b, & \\ y \geq x_j, & \quad \forall j \\ y \geq -x_j, & \quad \forall j \end{aligned}$$

EX 4: NORMS AND DISTANCES

- $\min |x| = \min\{y \geq 0 \mid y \geq x$ AND $y \geq -x\}$ is a linear program
  but NOT $\max |x| = \max\{x, -x\} = \max\{y \geq 0 \mid y = x$ OR $y = -x\}$
- we will see how to formulate disjunctions using binary (0/1) variables
  e.g. to formulate $\max \|x\|_1$ and $\max \|x\|_\infty$ as I(nteger)LPs
- modeling $\|x\|_p = (\sum_j |x_j|^p)^{1/p}$ for $p \geq 2$ usually requires nonlinear functions

EX 4: DATA FITTING

data fitting [Bertsimas-Tsitsiklis]

Given m observations – data points $a_i \in \mathbb{R}^n$ and associate values $b_i \in \mathbb{R}$, $i = 1..m$ – predict the value of any point $a \in \mathbb{R}^n$ according to a linear regression model ?

a best linear fit is a function :

$$b(a) = a^T x + y, \text{ for chosen } x \in \mathbb{R}^n, y \in \mathbb{R}$$

minimizing the residual/prediction error $|b(a_i) - b_i|$, globally over the dataset $i = 1..m$, e.g:

Least Absolute Deviation or L_1-regression:

$$\min \sum_i |b(a_i) - b_i|$$

EX 4: DATA FITTING – LAD REGRESSION (1)

supporting planes

$$\begin{aligned} & min \sum_i d_i \\ \text{s.t. } & d_i \geq \sum_j a_{ij}x_j + y - b_i, && \forall i \\ & d_i \geq -(\sum_j a_{ij}x_j + y - b_i), && \forall i \\ & d \in \mathbb{R}^m, x \in \mathbb{R}^n, y \in \mathbb{R} \end{aligned}$$

sparse supporting planes

$$\begin{aligned} & min \sum_i d_i \\ \text{s.t. } & r_i = \sum_j a_{ij}x_j + y - b_i, && \forall i \\ & d_i \geq r_i, && \forall i \\ & d_i \geq -r_i, && \forall i \\ & r, d \in \mathbb{R}^m, x \in \mathbb{R}^n, y \in \mathbb{R} \end{aligned}$$

Second model is better for many algorithms: larger (more variables and constraints) but its constraint matrix is less dense (more zeros)

EX 4: DATA FITTING – LAD REGRESSION (2)

variable splitting

$$\begin{aligned} & min \sum_i d_i^+ + d_i^- \\ \text{s.t. } & d_i^+ - d_i^- = \sum_j a_{ij}x_j + y - b_i, && \forall i \\ & d_i^+, d_i^- \geq 0, && \forall i \\ & x \in \mathbb{R}^n, y \in \mathbb{R} \end{aligned}$$

dual model (see later)

$$\begin{aligned} & max \sum_i b_i z_i \\ \text{s.t. } & \sum_i a_{ij} z_i = 0, && \forall j \\ & \sum_i z_i = 0, \\ & z_i \in [-1, 1], && \forall i \end{aligned}$$

Both models are equivalent by strong duality (see later) but the second one has much fewer variables and non-bound constraints. The best algorithms for LAD regression (Barrodale-Roberts) are special purpose simplex methods (see later) for dense matrices and absolute values.

READING:

to go further:

read [Bertsimas-Tsitsiklis]:
Sections 1.2, 1.3, 1.4

for the next class:

read [Bertsimas-Tsitsiklis]:
Section 2.1: Polyhedra and convex sets

GEOMETRY AND ALGEBRA

EX 5: DOORS & WINDOWS

A factory made of 3 workshops produces doors and windows. The workshops A, B, C are open 4, 12 and 18 hours a week, respectively. Assembling one door occupies workshop A for 1 hour and workshop C for 3 hours and the door is sold 3000 euros. Assembling one window occupies workshops B and C for 2 hours each and a window is sold 5000 euros. How to maximize the revenue ?

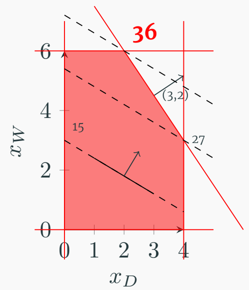
EX 5: LP DOORS & WINDOWS

- decision variables ?
  - x_D, x_W (fractional) number of doors and windows produced a day
- constraints and objective?
  - availability of each workshop (in hours/day), nonnegativity of the variables
  - maximize revenue (in keuros)

$$\begin{aligned} \max\ & 3x_D + 5x_W \\ \text{s.t.}\ & x_D \leq 4 \\ & 2x_W \leq 12 \\ & 3x_D + 2x_W \leq 18 \\ & x_D, x_W \geq 0 \end{aligned}$$

GRAPHICAL REPRESENTATION (EX: DOORS & WINDOWS)

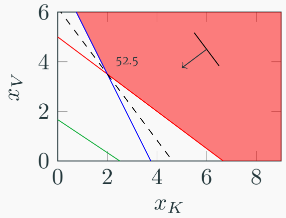
$$\begin{aligned} \max\ & 3x_D + 5x_W \\ \text{s.t.}\ & x_D \leq 4 \\ & 2x_W \leq 12 \\ & 3x_D + 2x_W \leq 18 \\ & x_D, x_W \geq 0 \end{aligned}$$



- solution space $\mathbb{R}^2$
- linear constraint $\equiv$ halfspace, ex: $\{x \in \mathbb{R}^2 \mid 3x_D + 2x_W \leq 18\}$
- feasible region $\equiv$ intersection of a finite number of halfspaces $\triangleq$ polyhedron
- objective: $z = 3x_D + 5x_W$, optimum: move the line up $z \nearrow$ until unfeasible
- optimum solution: $2x^*_W = 12$ and $3x^*_D + 2x^*_V = 18 \Rightarrow x^*_W = 6, x^*_D = 2, z^* = 36$

GRAPHICAL REPRESENTATION (EX: PETROLEUM DISTILLATION)

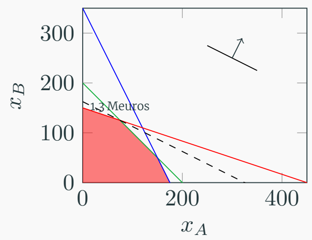
$$\begin{aligned} \min\ & 20x_K + 15x_V \\ \text{s.t.}\ & 3x_K + 4x_V \geq 20 \\ & 4x_K + 2x_V \geq 15 \\ & 2x_K + 3x_V \geq 5 \\ & 0 \leq x_K \leq 9 \\ & 0 \leq x_V \leq 6 \end{aligned}$$



- constraint $2x_K + 3x_V \geq 5$ is redundant
- constraints $3x_K + 4x_V \geq 20$ and $4x_K + 2x_V \geq 15$ are active/binding at the optimum $(2, 3.5)$ but not constraints $x_K \geq 0$ or $x_V \leq 6$

GRAPHICAL REPRESENTATION (EX: NUCLEAR WASTE)

$$\begin{aligned} \max\ & 4x_A + 8x_B \\ \text{s.t.}\ & x_A + 3x_B \leq 450 \\ & 2x_A + x_B \leq 350 \\ & x_A + x_B \leq 200 \\ & x_A, x_B \geq 0 \end{aligned}$$

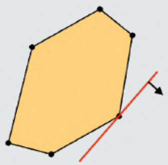


GEOMETRY OF LINEAR PROGRAMMING

- the feasible region is defined as a polyhedron
- thus it is **convex** (intersection of convex regions)

where are the optimal solutions ?

intuition: the optimum of a linear function on a polyhedron is reached at a "*corner point*" (under conditions of existence)



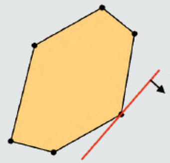
idea: **solving an LP = evaluate the corner points progressively**

CHARACTERIZING THE CORNER POINTS

Theorem [BT 2.3]

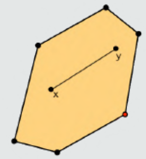
A nonempty polyhedron $\mathcal{P} = \{x \in \mathbb{R}^n | Ax \geq b\}$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and a feasible solution $\hat{x} \in \mathcal{P}$, then these are equivalent: $\hat{x}$ is a

vertex



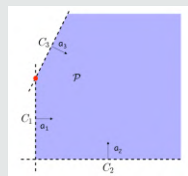
$\exists c \in \mathbb{R}^n, \forall x \in \mathcal{P} \setminus \{\hat{x}\}$, $c^T\hat{x} < c^T x$

extreme point



$\hat{x} = \lambda x + (1-\lambda)y$, $x, y \in \mathcal{P} \Rightarrow \lambda = 0$

basic feasible solution



$\exists n$ linearly independent rows a_i in A s.t. $a_i x = b_i$

vertices and extreme points are model-independent; their number $\leq \binom{m}{n}$ is **finite** but large and not known a priori

CHARACTERIZING THE CORNER POINTS (PROOF)

Theorem [BT 2.3]

$\hat{x} \in \mathcal{P} = \{x \in \mathbb{R}^n | Ax \geq b\}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ is either none or all together:

vertex	extreme point	basic feasible solution
$\exists c \in \mathbb{R}^n, \forall x \in \mathcal{P} \setminus \{\hat{x}\}$, $c^T\hat{x} < c^T x$	$\hat{x} = \lambda x + (1-\lambda)y$, $x, y \in \mathcal{P} \Rightarrow \lambda = 0$	$\exists n$ linearly independent rows a_i in A s.t. $a_i x = b_i$

Proof:

- $\hat{x}$ vertex $\Rightarrow$ xpoint: $\exists c, \forall x, y \in \mathcal{P} \setminus \{\hat{x}\}$, $c^T\hat{x} < c^T x$ and $c^T\hat{x} < c^T y$ then $c^T\hat{x} < \lambda c^T x + (1-\lambda)c^T y$, $\forall 0 \leq \lambda \leq 1$, then $\hat{x} \neq \lambda x + (1-\lambda)y$
- $\hat{x}$ not basic $\Rightarrow$ not xpoint: let $I = \{i | a_i\hat{x} = b_i\}$ then $rk(a_I^T) < n$ then $\exists d \in \mathbb{R}^n$, $a_I^T d = 0$. Let $x = \hat{x} + \epsilon.d$ and $y = \hat{x} - \epsilon.d$ then $\hat{x} = \frac{x+y}{2}$ and $x, y \in \mathcal{P}$: $a_i^T x = a_i^T y = b_i$ if $i \in I$, otherwise $a_i^T\hat{x} > b_i$ then $a_i^T x > b_i$ and $a_i^T y > b_i$ for ϵ small enough.
- $\hat{x}$ basic feasible $\Rightarrow$ vertex: let $c = \sum_{i \in I} a_i$ then $c^T\hat{x} = \sum_{i \in I} b_i \leq c^T x \; \forall x \in \mathcal{P}$, and equality holds only for $\hat{x}$ the unique solution of system $a_I^T x = b_I$.

EXISTENCE OF OPTIMA AND EXTREME POINTS

Theorem: existence of an extreme point [BT 2.6]

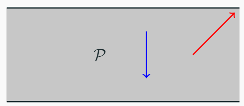
nonempty $\mathcal{P} = \{x \in \mathbb{R}^n | Ax \geq b\}$, $A \in \mathbb{R}^{m \times n}$ has at least one extreme point

$\iff$ it has no line: $\forall x \in \mathcal{P}, d \in \mathbb{R}^n$, $\{x + \theta d | \theta \in \mathbb{R}\} \not\subseteq \mathcal{P}$

$\iff$ A has n linearly independent rows

Theorem: existence of an optimal solution [BT 2.8]

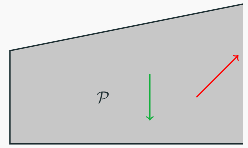
Minimizing a linear function over $\mathcal{P}$ having at least one extreme point, then: either optimal cost is $-\infty$, or an extreme point is optimal.



unbounded

∞ optima / 0 vertex

∞ optima including 1 vertex



EXISTENCE OF EXTREME POINTS (PROOF)

Theorem: existence of an extreme point [BT 2.6]

nonempty $\mathcal{P} = \{x \in \mathbb{R}^n | Ax \geq b\}$, $A \in \mathbb{R}^{m \times n}$ has at least one extreme point

$\iff$ it has no line: $\forall x \in \mathcal{P}, d \in \mathbb{R}^n$, $\{x + \theta d | \theta \in \mathbb{R}\} \not\subseteq \mathcal{P}$

$\iff$ A has n linearly independent rows

Proof:

- no line $\Rightarrow$ xpoint: let $x \in \mathcal{P}$ "of rank k", i.e. $I = \{i | a_i x = b_i\}$ has k lin. indep. rows, if not basic then $k < n$ and $\exists d$, $a_I^T d = 0$. The line (x, d) satisfies $a_I^T(x + \theta d) = b_i$ and it intersects the border of $\mathcal{P}$, i.e. $\exists \hat{\theta}, j \notin I$ s.t. $a_j^T(x + \hat{\theta} d) = b_j$, then $a_j^T d \neq 0$, then $x' = x + \hat{\theta} d \in \mathcal{P}$ is of rank $k+1$. Repeat until reaching n.
- $(a_i)_{i \in I}$ linearly independent $\Rightarrow$ no line: if $\mathcal{P}$ contains a line $x + \theta d$ with $d \neq 0$ then $a_i(x + \theta d) \geq b_i \; \forall \theta$ then $a_i d = 0 \; \forall i \in I$ then $d = 0$.

EXISTENCE OF OPTIMA (PROOF)

Theorem: existence of an optimal solution [BT 2.8]

Minimizing a linear function over $\mathcal{P}$ having at least one extreme point, then: either optimal cost is $-\infty$, or an extreme point is optimal.

Proof:

- let $x \in \mathcal{P}$ of rank $k < n$, then $\exists d$, $a_I^T d = 0, \forall i \in I = \{i | a_i x = b_i\}$. Assume $c^T d \leq 0$ (or use $-d$) then line (x, d) intersects the border of $\mathcal{P}$ at some $x' = x + \theta d \in \mathcal{P}$ of rank $k+1$ (see previous proof). If $c^T d = 0$ then $c^T x' = c^T x$. If $c^T d \leq 0$ then assume $\theta > 0$ (or optimal cost=$-\infty$), then $c^T x' < c^T x$. Repeat until reaching rank n, i.e. a basic feasible solution.
- let x^* be a basic feasible solution of $\mathcal{P}$ of minimum cost, then $c^T x^* \leq c^T x \; \forall x \in \mathcal{P}$

OPTIMA AND EXTREME POINTS (EXERCISE)

show that:

- $\mathcal{P} = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$ is nonempty and has no extreme point
- $(x, y) \mapsto 5(x + y)$ has a finite optimum on $\mathcal{P}$
- $\min\{5(x + y) \mid (x, y) \in \mathcal{P}\}$ has an optimal solution which is an extreme point (not of $\mathcal{P}$)

answer: put in standard form
$\min\{5(x^+ - x^- + y^+ - y^-) \mid x^+ - x^- + y^+ - y^- = 0,\ x^+, x^-, y^+, y^- \geq 0\}$ reaches its optimum at $(0, 0, 0, 0)$

CONSTRUCTING A BASIC SOLUTION

Theorem: basic solution for standard form [BT 2.4]

A nonempty polyhedron **in standard form** $\mathcal{P} = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ with m linear independent rows $A \in \mathbb{R}^{m \times n}$: $x \in \mathbb{R}^n$ is a basic solution iff $Ax = b$ and there exists m linear independent columns $A_j, j \in \beta \subset \{1, \ldots, n\}$ s.t. $x_j = 0, \forall j \notin \beta$.

The columns $A_j, j \in \beta$ is a **basis** of $\mathbb{R}^m$ and form an invertible basis matrix $A_{|\beta} \in \mathbb{R}^{m \times m}$; $x_j, j \in \beta$ are the basic variables

Algorithm: find a basic solution

1. pick m linear independent columns $A_j, j \in \beta \subset \{1, \ldots, n\}$
2. fix $x_j = 0, \forall j \notin \beta$
3. solve the system of m equations in $\mathbb{R}^m$: $A_{|\beta} x_{|\beta} = b$
4. the resulting basic solution x is feasible iff $x_{|\beta} = A_{|\beta}^{-1} b \geq 0$

BASIC SOLUTION FOR STANDARD FORM (PROOF)

Theorem: basic solution for standard form [BT 2.4]

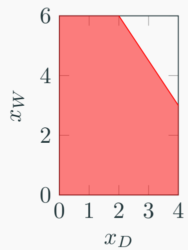
A nonempty polyhedron **in standard form** $\mathcal{P} = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ with m linear independent rows $A \in \mathbb{R}^{m \times n}$: $x \in \mathbb{R}^n$ is a basic solution iff $Ax = b$ and there exists m linear independent columns $A_j, j \in \beta \subset \{1, \ldots, n\}$ s.t. $x_j = 0, \forall j \notin \beta$.

Proof:

- $\Leftarrow$: let $x \in \mathbb{R}^n$ and β as in the statement, then $A_{|\beta} x_{|\beta} = Ax = b$ and $x_{|\beta} = A_{|\beta}^{-1} b$ is uniquely determined, then $span(A_{|\beta}) = \mathbb{R}^n$ (otherwise $\exists d$, $A_{|\beta} d = 0$ and $A_{|\beta} y = b$ would have many solutions $x_{|\beta} + \theta d$)
- $\Rightarrow$: let x basic and $I = \{i | x_i \neq 0\}$, then the active constraints ($Ax = b$ and $x_i = 0\ \forall i \notin I$) forms a system with an unique solution (otherwise for two solutions x^1 and x^2 then $d = x^1 - x^2$ would be orthogonal, i.e. not in the span=$\mathbb{R}^n$) then $A_{|I} x_{|I} = b$ has a unique solution and then $A_{|I}$ has lin. ind. columns. Since A has m lin. ind. rows then there exist $m - |I|$ columns lin. ind. with $A_{|I}$ and, by def of I, $x_i = 0$ for any other column i.

BASIC SOLUTIONS (EX: LP DOORS & WINDOWS)

$$\begin{aligned} \max\ & 3x_D + 5x_W \\ \text{s.t.}\ & x_D + s_1 = 4 \\ & 2x_W + s_2 = 12 \\ & 3x_D + 2x_W + s_3 = 18 \\ & x_D, x_W, s_1, s_2, s_3 \geq 0 \end{aligned}$$



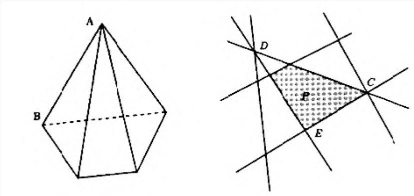
$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 3 & 2 & 0 & 0 & 1 \end{pmatrix}$$

$$\beta_1 = (3, 4, 5), \beta_2 = (1, 2, 5), \beta_3 = (1, 4, 5), \beta_4 = (1, 2, 3)$$

$$x_{\beta_1} = (0, 0, 4, 12, 18), x_{\beta_2} = (4, 6, 0, 0, -6), x_{\beta_3} = (4, 0, 0, 12, 6), x_{\beta_4} = (2, 6, 2, 0, 0)$$

DEGENERACY

one basis defines one unique basic solution
but one basic solution may correspond to different bases, when it is
degenerate $\iff$ more than n active constraints
$\iff$ some basic variables are set to 0.



basic nonfeasible degenerate ? D
basic feasible nondegenerate ? B and E
basic feasible degenerate ? A and C

EX 6: CAPACITY PLANNING

capacity planning [Bertsimas-Tsitsiklis]

find a least cost electric power capacity expansion plan:

- planning horizon: the next $T \in \mathbb{N}$ years
- forecast demand (in MW): $d_t \geq 0$ for each year $t = 1, \dots, T$
- existing capacity (oil-fired plants, in MW): $e_t \geq 0$ available for each year t
- options for expanding capacities: (1) coal-fired plant and (2) nuclear plant
  - lifetime (in years): $l_j \in \mathbb{N}$, for each option $j = 1, 2$
  - capital cost (in euros/MW): c_{jt} to install capacity j operable from year t
  - political/safety measure: share of nuclear should never exceed 20% of available capacity

EX 6: LP MODEL

- decision variables, x_{jt}: installed capacity (in MW) of type $j = 1, 2$ starting at year $t = 1, \dots, T$
- constraints, each year: total capacity meets the demand + nuclear share
- implied variables, y_{jt} available capacity (in MW) $j = 1, 2$ for year t

$$\min \sum_{t=1}^{T} \sum_{j=1}^{2} c_{jt} x_{jt}$$

$$\begin{aligned}
\text{s.t. } & y_{jt} - \sum_{s=\max\{1, t-l_j+1\}}^{t} x_{js} &= 0, \quad & \forall j = 1, 2, t = 1, \dots, T \\
& y_{1t} + y_{2t} - u_t &= d_t - e_t, \quad & \forall t = 1, \dots, T \\
& 8y_{2t} - 2y_{1t} + v_t &= 2e_t, \quad & \forall t = 1, \dots, T \\
& x_{jt} \geq 0, y_{jt} \geq 0, u_t \geq 0, v_t \geq 0 & & \forall j = 1, 2, t = 1, \dots, T
\end{aligned}$$

EX: BASIC SOLUTION (CAPACITY PLANNING)

$$\min \sum_{t=1}^{T} \sum_{j=1}^{2} c_{jt} x_{jt}$$

$$\begin{aligned}
\text{s.t. } & y_{jt} - \sum_{s=\max\{1, t-l_j+1\}}^{t} x_{js} &= 0, \quad & \forall j = 1, 2, t = 1, \dots, T \\
& y_{1t} + y_{2t} - u_t &= d_t - e_t, \quad & \forall t = 1, \dots, T \\
& 8y_{2t} - 2y_{1t} + v_t &= 2e_t, \quad & \forall t = 1, \dots, T \\
& x_{jt} \geq 0, y_{jt} \geq 0, u_t \geq 0, v_t \geq 0 & & \forall j = 1, 2, t = 1, \dots, T
\end{aligned}$$

$$\begin{pmatrix} L & 0 & I & 0 & 0 & 0 \\ 0 & L & 0 & I & 0 & 0 \\ 0 & 0 & I & I & -I & 0 \\ 0 & 0 & -2I & 8I & 0 & I \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \\ u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ d-e \\ 2e \end{pmatrix}$$

$n = 6T$ variables, $m = 4T$, A has linearly independent rows;
I: identity matrix, L: lower triangular matrix of 1s and 0s basic solution $(0, 0, 0, 0, e-d, 2e)$ is feasible iff $e_t \geq d_t, \forall t$,
degenerate ($4T > n - m$ zeros), other basis e.g (x_1, x_2, u, v)

EX: BASIC SOLUTION AND DEGENERACY (CAPACITY PLANNING)

Exercise: reformulate by dropping the redundant variables y_1 and y_2, find a basic solution, and give conditions of degeneracy

$$
\begin{aligned}
\min\ & \sum_{t=1}^{T}\sum_{j=1}^{2} c_{jt}x_{jt} \\
\text{s.t. } & \sum_{s=\max\{1,t-l_1+1\}}^{t} x_{1s} + \sum_{s=\max\{1,t-l_2+1\}}^{t} x_{2s} - u_t = d_t - e_t, && \forall t = 1,\dots,T \\
& 8\sum_{s=\max\{1,t-l_2+1\}}^{t} x_{2s} - 2\sum_{s=\max\{1,t-l_1+1\}}^{t} x_{1s} + v_t = 2e_t, && \forall t = 1,\dots,T \\
& x_{jt} \geq 0, u_t \geq 0, v_t \geq 0 && \forall j = 1,2, t = 1,\dots,T
\end{aligned}
$$

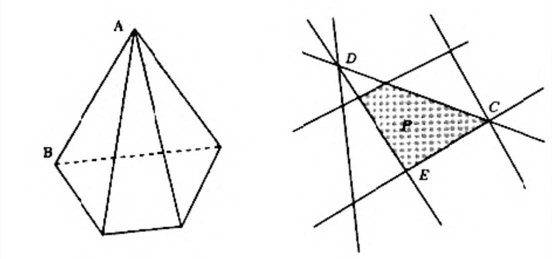
$$
\begin{pmatrix} L & L & -I & 0 \\ -2L & 8L & 0 & I \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ u \\ v \end{pmatrix} = \begin{pmatrix} d-e \\ 2e \end{pmatrix}
$$

basic solution $(0, 0, e-d, 2e)$ is feasible iff $e_t \geq d_t, \forall t$,
degenerate iff $\exists t$, $e_t = 0$ or $e_t = d_t$

MOVING TO ANOTHER BASIC SOLUTION

Adjacency

- two basic solutions x and y are adjacent if there exists $n-1$ linearly independent constraints active at x and y
- the line segment between 2 adjacent basic feasible solutions is an edge of $\mathcal{P}$
- (nondegenerate) adjacent basic feasible solutions correspond to adjacent bases (in standard form), i.e. that share $m-1$ columns



SUMMARY

- the feasible set of an LP is a polyhedron $\mathcal{P}$
- if $\mathcal{P}$ is nonempty and bounded, then (i) there exists an optimal solution which is an extreme point
- if $\mathcal{P}$ is unbounded, then either (i), or (ii) there exists an optimal solution but no extreme point (not in standard form), or (iii) the optimal cost is infinite
- if (i) then the LP can be solved in a finite (probably exponential) number of steps by evaluating all extreme points

Instead of complete enumeration: the simplex algorithm moves along the edges of $\mathcal{P}$ while **improving** the objective

READING:

to go further:

read [Bertsimas-Tsitsiklis]:
Sections 2.2, 2.3, 2.4, 2.5, 2.6

for the next class:

read [Bertsimas-Tsitsiklis]:
Section 1.6: Algorithms and operation count

THE SIMPLEX METHODS

REVIEW

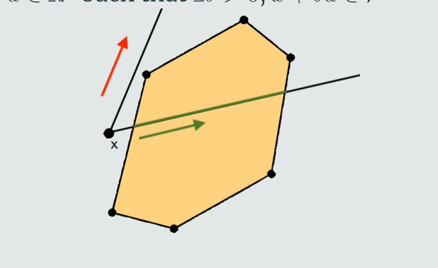
- min cx over $\mathcal{P} = \{Ax = b, x \geq 0\}$, $A \in \mathbb{R}^{m \times n}$, $rk(A) = m$ reaches its optimum at a **basic feasible solution**
- a **basis** $\beta \subseteq \{1, \ldots, n\}$ is made of m linearly independent columns of A and the associated basic solution is: $x_\beta = A_\beta^{-1} b$, $x_{\neg\beta} = 0$
- adjacent basic solutions share $m-1$ basic variables: $\beta' = \beta \cup \{j'\} \setminus \{j''\}$
- adjacent basic solutions may coincide if degenerate (if $x_{j'} = x_{j''} = 0$)

the simplex method goes from a basic feasible solution to an adjacent one as the cost decreases

FEASIBLE IMPROVING DIRECTION

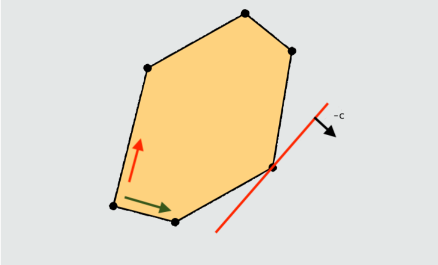
feasible direction from $x \in \mathbb{R}^n$

$d \in \mathbb{R}^n$ such that $\exists \theta > 0$, $x + \theta d \in \mathcal{P}$



improving direction from $x \in \mathbb{R}^n$

$d \in \mathbb{R}^n$ such that $c^T d < 0$



following a feasible improving direction d with a step $\theta > 0$ leads to a feasible solution $x' = x + \theta d \in \mathcal{P}$ of better cost $c^T x' = c^T x + \theta . c^T d < cx$

FEASIBLE IMPROVING BASIC DIRECTION

Let x be a basic feasible solution of basis β, and $j' \notin \beta$:

the j'th basic direction

$d \in \mathbb{R}^n$: $d_{j'} = 1, d_j = 0, \forall j \notin \beta \cup \{j'\}$, and $Ad = 0$ (i.e. $d_\beta = -A_\beta^{-1} A_{j'}$)

is a feasible direction if x nondegenerate:

- $x_\beta > 0 \Rightarrow \exists \theta > 0, x_\beta + \theta d_\beta \geq 0 \Rightarrow x + \theta d \geq 0$
- $Ad = A_\beta d_\beta + A_{j'} = 0 \Rightarrow \forall \theta > 0,\ A(x + \theta d) = Ax = b$

reduced cost of nonbasic variable $x_{j'}$

$\bar{c}_{j'} = c^T d = c_{j'} - c_\beta^T A_\beta^{-1} A_{j'}$

- $\bar{c}_{j'} = c^T d = c^T x' - c^T x$ is the cost deviation when $\theta = 1$ and $x' = x + d$
- d is an **improving direction** iff $\bar{c}_{j'} < 0$
- the reduced cost of a basic variable $j \in \beta$ is always 0: $\bar{c}_j = c_j - c_\beta^T A_\beta^{-1} A_j = c_j - c_\beta^T e_j = 0$

EXAMPLE: BASIC IMPROVING DIRECTION

$$\begin{aligned} min_{x\geq 0}\ & 2x_1 + x_2 + x_3 + x_4 \\ \text{s.t.}\ & x_1 + x_2 + x_3 + x_4 = 2 \\ & 2x_1 + 3x_3 + 4x_4 = 2 \end{aligned}$$

- $m = 2$, $n = 4$, $rk(A) = 2$
- $\beta = \{1, 2\}$ is a basis
- $x = (1, 1, 0, 0)$ feasible nondegenerate ($x_j > 0\,\forall j \in \beta$)
- basic direction $j = 3$: $d_3 = 1$, $d_4 = 0$, $Ad = \begin{pmatrix} d_1 + d_2 + 1 \\ 2d_1 + 3 \end{pmatrix} = 0 \Rightarrow d_\beta = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} -3/2 \\ 1/2 \end{pmatrix}$
- improving direction: $\bar{c} = c^T d = 2(-3/2) + (1/2) + 1 = -3/2 < 0$

STEP LENGTH θ

β basis of x feasible nondegenerate, d feasible direction to $j' \notin \beta$ s.t. $c^T d = \bar{c}_{j'} < 0$

Theorem [BT 3.2]

if $d \geq 0$ then the LP is unbounded, otherwise
if $j'' \in argmin\{-x_j/d_j, j \in \beta, d_j < 0\}$ and $\theta = -x_{j''}/d_{j''}$ then $x' = x + \theta d$ is a basic feasible solution of basis $\beta' = \beta \cup \{j'\} \setminus \{j''\}$: j' enters the basis, j'' exits the basis.

θ is the highest value s.t. $x' \in \mathcal{P}$, i.e. s.t. one (or more) new active constraint $x'_{j''} \geq 0$

Proof:

- by construction, $Ax' = Ax + \theta Ad = Ax = b$ then
  $x' \in \mathcal{P} \iff x_j + \theta d_j \geq 0\,\forall j \iff x_j + \theta d_j \geq 0\,\forall j \in \beta : d_j < 0.$
- $\theta > 0$ since x nondegenerate ($x_\beta > 0$)
- if $d \geq 0$ then $x + \theta d \in \mathcal{P}\ \forall \theta > 0$ and $c(x + \theta d) \searrow$ when $\theta \nearrow$
- $A_\beta^{-1} A_j = e_j, \forall j \in \beta \setminus \{j''\}$, and $A_\beta^{-1} A_{j'} = -d_\beta$ has a nonzero j'' component $\Rightarrow \{A_j, j \in \beta'\}$ are linear independent $\Rightarrow \beta'$ is a basis

EXAMPLE: BASIC IMPROVING DIRECTION (CONT.)

$$\begin{aligned} min_{x\geq 0}\ & 2x_1 + x_2 + x_3 + x_4 \\ \text{s.t.}\ & x_1 + x_2 + x_3 + x_4 = 2 \\ & 2x_1 + 3x_3 + 4x_4 = 2 \end{aligned}$$

- $\beta = \{1, 2\}$ is a basis: $x = (1, 1, 0, 0)$ feasible nondegenerate
- basic feasible improving direction $j = 3$: $d = (-3/2, 1/2, 1, 0)$, $\bar{c}_3 = c^T d = -3/2$
- $x' = x + \theta d \geq 0 \Rightarrow x'_1 = 1 - (3/2)\theta \geq 0 \Rightarrow \theta \leq 2/3$
- $x' = (0, 4/3, 2/3, 0)$ basic feasible solution $\beta' = \{2, 3\}$, $cx' = cx + \theta \bar{c}_3 = cx - 1$

OPTIMALITY CONDITION

Theorem [BT 3.1]

Let x be a basic feasible solution of basis β and $\bar{c} \in \mathbb{R}^n$ the vector of reduced costs.

- if $\bar{c}_j \geq 0\,\forall j \notin \beta$ then x is **optimal**
- if x is optimal and nondegenerate then $\bar{c} \geq 0$

Proof:

($\Rightarrow$) for any $y \in \mathcal{P}$, let $d = y - x$ and $c_{\neg\beta} \geq 0$:
$A_\beta d_\beta + A_{\neg\beta} y_{\neg\beta} = Ad = Ay - Ax = b - b = 0 \Rightarrow d_\beta = -A_\beta^{-1} A_{\neg\beta} y_{\neg\beta} \Rightarrow$
$c^T y - c^T x = c_\beta^T d_\beta + c_{\neg\beta}^T y_{\neg\beta} = (c_{\neg\beta}^T - c_\beta^T A_\beta^{-1} A_{\neg\beta}) y_{\neg\beta} = \bar{c}_{\neg\beta} y_{\neg\beta} \geq 0$
($\Leftarrow$) if x nondegenerate and $\bar{c}_j < 0$, then j is nonbasic and of feasible improving direction, then x nonoptimal

EXAMPLE: BASIC IMPROVING DIRECTION (CONT.)

$$\begin{aligned} min_{x\geq 0}\ & 2x_1 + x_2 + x_3 + x_4 \\ \text{s.t. } & x_1 + x_2 + x_3 + x_4 = 2 \\ & 2x_1 + 3x_3 + 4x_4 = 2 \end{aligned}$$

- note that optimum ≥ 2 since $cx = x_1 + 2$, $\forall x$ feasible
- $\beta = \{2,3\}$ is a basis with $x = (0, 4/3, 2/3, 0)$ nondegenerate
- basic directions are not improving:
  - $j = 1$: $d = (1, -1/3, -2/3, 0)$ and $\bar{c}_1 = cd = 1 \geq 0$
  - $j = 4$: $d = (0, 1/3, -4/3, 1)$ and $\bar{c}_4 = cd = 0 \geq 0$
- then x is optimal

THE SIMPLEX METHOD (SIMPLE CASE)

steps	howto:
1. get a basis β	find m linearly independent columns
2. get a basic **feasible** x	$x_{\neg\beta} = 0$, $x_\beta = A_\beta^{-1}b$ **if** $x_\beta \geq 0$
halt condition (optimality)	$\bar{c} = c - c_\beta^T A_\beta^{-1} A \geq 0$ **if nondegenerate**
3. find an improving direction	any $j' \notin \beta$ s.t. $\bar{c}_{j'} < 0$ **if nondegenerate**
halt condition (unboundness)	$d_\beta = -A_\beta^{-1} A_{j'} \geq 0$
4. find the largest step length	any $j'' \in argmin\{-x_j/d_j \mid j \in \beta, d_j < 0\}$
5. update the basis	$\beta := \beta \cup \{j'\} \setminus \{j''\}$
6. goto 2	$x := x - (x_{j''}/d_{j''})d$

THE SIMPLEX METHOD

convergence [BT 3.3]

if $\mathcal{P} \neq \emptyset$ and every basic feasible solution is nondegenerate then the simplex method terminates after a finite number of iterations with either an optimal basis β or with some direction $d \geq 0$, $Ad = 0$, $c^T d < 0$, and the optimal cost is $-\infty$

Proof:

- cx decreases at each iteration, all x are basic feasible solutions, the number of basic feasible solutions is finite

PIVOTING RULES

- choice of the entering column $j' \notin \beta$ s.t. $\bar{c}_{j'} < 0$, e.g.:
  - largest cost decrease per unit change: min $\bar{c}_j$
  - largest cost decrease: min $\theta\bar{c}_j$
  - smallest subscript: min j
- choice of the exiting column $j'' \in argmin\{-x_j/d_j \mid j \in \beta, d_j < 0\}$
- trade-off between computation burden and efficiency, e.g. compute a subset of reduced costs

IN CASE OF DEGENERARY ?

- if x is degenerate with $x_j = 0$ and $d_j < 0$ for some $j \in \beta$, then $\theta = 0$: the basis changes but not the basic feasible solution
- a sequence of basis changes may lead to a cost reducing feasible direction or it may **cycle**
- to avoid cycles and ensure convergence: select the smallest subscript pivoting rules for both entering and exiting columns (see [Bertsimas-Tsitsiklis] Section 3.4 for details)

THE INITIAL BASIC FEASIBLE SOLUTION ?

- if $\mathcal{P} = \{Ax \leq b, x \geq 0\}$, then we directly get a basis from the slack variables: $\mathcal{P} = \{Ax + Is = b, x \geq 0, s \geq 0\}$
- if the problem is already in standard form $min\{cx, Ax = b, x \geq 0\}$, then we can first solve the auxiliary LP:

$$min\{1.y, Ax + Iy = b, x \geq 0, y \geq 0\}$$

if optimum is 0 then we get a feasible basic solution for the original LP, otherwise it is unfeasible (see [Bertsimas-Tsitsiklis] Section 3.5 for details)

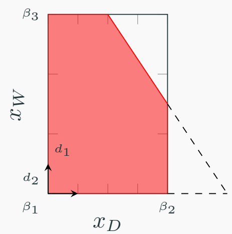
IMPLEMENTATIONS

- each iteration involves costly arithmetic operations:
  - computing $u^T = c_\beta^T A_\beta^{-1}$ or $A_\beta^{-1} A_j$ takes $O(m^3)$ operations
  - computing $\bar{c}_j = c_j - u^T A_j$ for all $j \notin \beta$ takes $O(mn)$ operations
- revised simplex: update matrix $A^{-1}_{\beta \cup \{j''\} \setminus \{j'\}}$ from A_β^{-1} in $O(mn)$
- full tableau: maintain and update the $m \times (n+1)$ matrix $A_{\beta^{-1}}(b|A)$
- specific data structures for sparse (many 0 entries in A) vs. dense matrices
- in theory, complexity is exponential in the worst case: the LP may have 2^n extreme points and the simplex method visits them all
- in practice, sophisticated implementations of the simplex method perform often better than polynomial-time algorithms (interior point/barrier, ellipsoid) and have additional features (duality, restart)

(see [Bertsimas-Tsitsiklis]Section 3.3 for details)

EX: SIMPLEX ALGORITHM (LP DOORS & WINDOWS)

$$\begin{aligned} \min \quad & -3x_D - 5x_W \\ \text{s.t.} \quad & x_D + s_1 = 4 \\ & 2x_W + s_2 = 12 \\ & 3x_D + 2x_W + s_3 = 18 \\ & x_D, x_W, s_1, s_2, s_3 \geq 0 \end{aligned}$$



- start at $\beta_1 = (3, 4, 5)$: $x_{\beta_1} = (0, 0, 4, 12, 18)$ (feasible nondegenerate)
- $d_1 = (1, 0, -1, 0, -3)$, $\bar{c}_1 = -3$, and $d_2 = (0, 1, 0, -2, -2)$, $\bar{c}_2 = -5$ both improving
- choose $j' = 1$: $\theta = \min(4/1, 18/3) = 4$, $j'' = 3$, $\beta_2 = (1, 4, 5)$, $x_{\beta_2} = (4, 0, 0, 12, 6)$
- or choose $j' = 2$: $\theta = \min(12/2, 18/2) = 6$, $j'' = 4$, $\beta_3 = (2, 3, 5)$, $x_{\beta_3} = (0, 6, 4, 0, 6)$

DUALITY

DUALITY: MOTIVATION

$$P : z = min \{x^2 + y^2 \mid x + y = 1\} \qquad \text{(not linear, still convex)}$$

- unconstrained smooth convex optimization is easy: zero of the derivative
- penalization methods: $P_u:\ z_u = min\ x^2 + y^2 + u(1 - x - y)$
  relax the constraints and penalize the violations with price/multiplier $u \in \mathbb{R}$
- provides a lower bound $z_u \leq z$:
  (x, y) feasible for $P \Rightarrow$ feasible for P_u and $z_u \leq x^2 + y^2 + u(1 - x - y) = x^2 + y^2$
- P_u is a relaxation of P
- the optimal solution of P_u is $(u/2, u/2)$: $\nabla c(x, y) = 0$ iff $(2x - u, 2y - u) = 0$
- for $u = 1$: $(1/2, 1/2)$ is both optimal for P_1 and feasible for P,
  thus it is optimal for P: $1/2 = z_1 \leq z \leq (1/2)^2 + (1/2)^2 = 1/2$

LAGRANGIAN MULTIPLIERS

$$P : z = min\ c^T x \quad \text{s.t. } Ax = b,\ x \geq 0$$

$$P_u : z_u = min\ c^T x + u^T(b - Ax) \quad \text{s.t. } x \geq 0$$

with multipliers $u \in \mathbb{R}^m$

- lagrangian problems P_u, $u \in \mathbb{R}^m$ provide lower bounds $z_u \leq z$
- dual problem: find the tightest (greater) lower bound

$$D : d = max_{u \in \mathbb{R}^m} z_u$$

- if x is optimal for some P_u and satisfies $Ax = b$ then x is optimal for P and $d = z$

THIS CLASS: PROPERTIES OF LP DUALITY

- if P is an LP then D is also an LP and $z = d$ when finite (strong duality)
- the dual of D is P and the constraints of P correspond to the variables of D (and vice versa)
- the primal simplex algorithm also computes solutions in the dual space and stops when the basis is dual feasible
- the dual simplex algorithm also computes solutions in the primal space and stops when the basis is primal feasible
- sensitive analysis / restart when problem changes: check how to recover feasibility in the primal or in the dual space

DUAL LINEAR PROGRAM

Theorem

- the dual of a linear program is a linear program:

$$(P): min\ c^T x \quad \text{s.t. } Ax = b, x \geq 0 \qquad\qquad (D): max\ u^T b \quad \text{s.t. } u^T A \leq c^T$$

- the dual of D is the primal P
- equivalent forms of P give equivalent forms of D

Proof:

- $z_u = min_{x \geq 0} c^T x + u^T(b - Ax) = u^T b + min_{x \geq 0}(c^T - u^T A)x$
- $z_u = \begin{cases} u^T b & \text{if } (c^T - u^T A) \geq 0 \\ -\infty & \text{otherwise} \end{cases}$

HOW TO BUILD THE DUAL ?

primal/dual correspondence

min	max
cost vector c	RHS vector b
matrix A	matrix A^T
constraint $a_i x = b_i$	free variable $u_i \in \mathbb{R}$
constraint $a_i x \geq b_i$	nonnegative variable $u_i \geq 0$
free variable $x_j \in \mathbb{R}$	constraint $u^T A_j = c_j$
nonnegative variable $x_j \geq 0$	constraint $u^T A_j \leq c_j$

$$\begin{aligned} P: min\ & c^T x + d^T y \\ \text{s.t. } & Ax = b && (u) \\ & Dx + Ey \geq f && (v) \\ & x \geq 0 \end{aligned}$$

$$\begin{aligned} D: max\ & u^T b + v^T f \\ \text{s.t. } & A^T u + D^T v \leq c && (x) \\ & E^T v = d && (y) \\ & v \geq 0 \end{aligned}$$

EX 7: STEEL FACTORY

steel factory

A factory can produce steel in coils (*bobines*), tapes (*rubans*), and sheets (*tôles*) every week up to 6000 tons, 4000 tons and 3500 tons, respectively. The selling prices are 25, 30, and 2 euros, respectively, per ton of product. Production involves two stages, heating (*réchauffe*) and rolling (*laminage*). These two mills are available up to 35 hours and 40 hours a week, respectively. The following table gives the number of tons of products that each mill can process in 1 hour:

	heating	rolling
coils	200	200
tapes	200	140
sheets	200	160

The factory wants to maximize its profit.

EX 7: LP MODEL

- decision variables ?
  - x_C, x_T, x_S the quantity (in tons) of weekly produced coils, tapes and sheets
- constraints ?
  - mill occupation
  - maximum production

$$\begin{aligned} P: \text{max } & 25x_C + 30x_T + 2x_S \\ \text{s.t.} \\ & \frac{x_C}{200} + \frac{x_T}{200} + \frac{x_S}{200} \leq 35 && (heating) \\ & \frac{x_C}{200} + \frac{x_T}{140} + \frac{x_S}{160} \leq 40 && (rolling) \\ & 0 \leq x_C \leq 6000 && (coils) \\ & 0 \leq x_T \leq 4000 && (tapes) \\ & 0 \leq x_S \leq 3500 && (sheets) \end{aligned}$$

EX: DUAL MODEL (STEEL FACTORY)

$$D: \min 35u_H + 40u_R + 6000u_C + 4000u_T + 3500u_S$$

s.t.

$$\frac{u_H}{200} + \frac{u_R}{200} + u_C \geq 25 \qquad (coils)$$
$$\frac{u_H}{200} + \frac{u_R}{140} + u_T \geq 30 \qquad (tapes)$$
$$\frac{u_H}{200} + \frac{u_R}{160} + u_S \geq 2 \qquad (sheets)$$
$$u \geq 0$$

WEAK DUALITY

Theorem [BT 4.3]

- if x is feasible for P (min) and u is feasible for D (max) then: $u^T b \leq cx$
- if the optimal cost of P is $-\infty$ then D is unfeasible
- if the optimal cost of D is $+\infty$ then P is unfeasible
- if $u^T b = cx$ then x is optimal for P and u is optimal for D

Proof:

- if P in standard form: $Ax = b, x \geq 0$ and $u^T A \leq c^T$, then $u^T b = u^T Ax \leq cx$.
- in any form: if (x, u) primal-dual feasible then by construction $u^T(Ax - b) \geq 0$ and $(c^T - u^T A)x \geq 0$, then $u^T b \leq u^T Ax \leq cx$.

STRONG DUALITY

Theorem [BT 4.4]

if a linear programming problem has an optimal solution, so does its dual and their respective optima are equal: $u^T b = cx$

Proof:

- let x an optimal solution of $P = min\{c^T x | Ax = b, x \geq 0\}$ of basis β
- x optimal then the reduced costs are all nonnegative $\bar{c}^T = c^T - c_\beta^T A_\beta^{-1} A \geq 0$
- let $u^T = c_\beta^T A_\beta^{-1}$ then u is feasible for $D = max\{u^T b | u^T A \leq c^T\}$
- $u^T b = c_\beta^T A_\beta^{-1} b = c_\beta^T x_\beta = c^T x$ then u is optimal for D

At optimality: the primal reduced costs $\bar{c}^T$ are the dual slacks $c^T - u^T A$

COMPLEMENTARY SLACKNESS

Theorem [BT 4.5]

let x feasible for P and u feasible for D then they are optimal iff

$$u_i(a_i^T x - b_i) = 0 \quad \forall i \text{ row of } P$$
$$(c_j - u^T A_j)x_j = 0 \quad \forall j \text{ row of } D.$$

Proof:

- (x, u) primal(min)-dual(max) feasible then $u_i(a_i x - b_i) \geq 0$ and $(c_j - u^T A_j)x_j \geq 0$
- $c^T x - u^T b = \sum_j (c_j - u^T A_j)x_j + \sum_i u_i(a_i x - b_i)$ sum of nonnegative terms is zero iff all terms are zero

Either a constraint is binding at the optimum or the dual variable is zero

Exercise: Optimality Without Simplex

$$P: min\ 13x_1 + 10x_2 + 6x_3$$
$$\text{s.t. } 5x_1 + x_2 + 3x_3 = 8$$
$$3x_1 + x_2 = 3$$
$$x_1, x_2, x_3 \geq 0$$

show that the basic solution of P of basis $\beta = \{1, 3\}$ is feasible nondegenerate and optimal using the complementary slackness theorem

Exercise: Optimality Without Simplex

$P: min\ 13x_1 + 10x_2 + 6x_3$	$D: max\ 8u_1 + 3u_2$
s.t. $5x_1 + x_2 + 3x_3 = 8$	s.t. $5u_1 + 3u_2 \leq 13$
$3x_1 + x_2 = 3$	$u_1 + u_2 \leq 10$
$x_1, x_2, x_3 \geq 0$	$3u_1 \leq 6$

- $\beta = \{1, 3\} \Rightarrow x_2 = 0$, $x_1 = 3/3 = 1$, $x_3 = (8 - 5)/3 = 1$
- $x = (1, 0, 1)$, $x \geq 0 \Rightarrow$ feasible, $x_j > 0, \forall j \in \beta \Rightarrow$ nondegenerate
- P in standard form $\Rightarrow$ first C.S. is always condition satisfied
- let u satisfying second C.S. condition, i.e. $5u_1 + 3u_2 = 13$ and $3u_1 = 6$
- $u = (2, 1)$ is feasible for D since $u_1 + u_2 = 3 \leq 10$
- C.S. theorem $\Rightarrow x$ and u are optimal with cost 19
- $u = c_\beta^\top A_\beta^{-1}$ basic dual solution: feasible $\iff \bar{c}_2 = c_2^T - u^T A_2 \geq 0$ (reduced cost)

Optimality Conditions

Theorem

x is optimal for $P = min\{c^T x | Ax = b, x \geq 0\}$ if exists $u \in \mathbb{R}^m$ s.t. (x, u) satisfies:

1. primal feasibility: $Ax = b$
2. primal feasibility: $x \geq 0$
3. dual feasibility: $u^T A \leq c^T$
4. complementary slackness: $x_j > 0 \Rightarrow u^T A_j = c_j$

- basic feasible solutions always satisfy 1,2 and 4 with $u^T = c_\beta^T A_\beta^{-1}$ ($x_j > 0 \Rightarrow j \in \beta$ and $\bar{c}_j = c_j^T - u^T A_j = 0$).
- Condition 3 is the halting condition $\bar{c} \geq 0$ of the simplex algorithm
- if x is degenerate then solutions u of condition 4 may not be unique
- these are the KKT necessary and sufficient conditions on $l(x, u, v) = c^T x + u^T(b - Ax) - vx$: exists $(u, v) \in \mathbb{R}^{m \times n}$ s.t. $Ax = b$ (primal), $x \geq 0$ (primal), $\nabla l_{u,v}(x) = c - (u^\top A + v) = 0$ (stationarity), $v \geq 0$ (dual), $x^\top v = 0$ (CS)

Dual Simplex

for $P = min\{cx | Ax = b, x \geq 0\}$ and $D = max\{u^T b | u^T A \leq c\}$

- a basis β determines basic solutions for P and D: $x_\beta = A_\beta^{-1} b$ and $u^T = c_\beta^T A_\beta^{-1}$
- if both are feasible, then both are optimal (according to C.S. since $u^T(Ax - b) = 0$ and $(c^T - u^T A)x = (c_\beta^T - u^T A_\beta)x_\beta = 0$)
- simplex algorithm maintains primal feasibility ($x_\beta \geq 0$) while trying to achieve dual feasibility ($\bar{c}^T = c^T - u^T A \geq 0$)
- dual simplex algorithm maintains dual feasibility ($\bar{c} \geq 0$) while trying to achieve primal feasibility ($x_\beta \geq 0$)
- examples of usage: after modifying b or adding a new constraint to P, run the dual simplex starting from the feasible dual solution $c_\beta^T A_\beta^{-1}$

Farka's lemma and unfeasibility

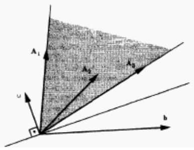
theorem

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Exactly one of the following holds:

1. $\exists x \in \mathbb{R}^n, x \geq 0, Ax = b$ ($\mathcal{P} = \min_{x\geq 0}\{cx : Ax = b\}$ is feasible)
2. $\exists u \in \mathbb{R}^m$, $u^T A \geq 0$ and $u^T b < 0$ (xor b can be separated from $\{Ax, x \geq 0\}$ by a plane)

Proof:

($1 \Rightarrow \neg 2$) if $x \in \mathcal{P}$ and $u^T A \geq 0$ then $u^T b = u^T Ax \geq 0$
($\neg 1 \Rightarrow 2$) if $P : max\{0|Ax = b, x \geq 0\}$ is unfeasible then $D : min\ \{u^T b | u^T A \geq 0\}$ is either unbounded or unfeasible. Since $u = 0$ is feasible for D, then (2) holds.



if b is not in the cone $\{Ax, x \geq 0\}$ spanned by the columns of A then a separating hyperplane $\{x \in \mathbb{R}^m | u^T x = 0\}$ exists

Interior-Point Methods (applied to LP)

- idea: iterate on primal and dual feasible solutions until achieving complementary slackness
- disturbed KKT conditions: x is optimal for $P = min\{c^T x | Ax = b, x \geq 0\}$ if exists $(u, v) \in \mathbb{R}^{m\times n}$ s.t. $Ax = b$ (primal), $x \geq 0$ (primal), $Au + v = c$ (stationarity), $v \geq 0$ (dual), $x^\top v = 1/t$ (quasi-CS)
- this are the KKT conditions for the centered problem $P_t = min\{tc^T x + \phi(x) | Ax = b\}$ where the barrier function $\phi(x) = -\sum_j log(x_j)$ is a smooth approximation of the indicator function for $x \geq 0$
- barrier method: solve P_t with the Newton method for increasing t (fix $v = x^{-1}/t$)
- primal-dual interior-point method: update (x, u, v) at each iteration

reading:

to go further:

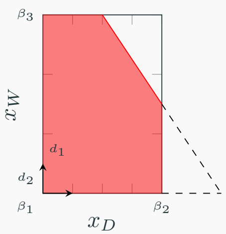
read [Bertsimas-Tsitsiklis]:
Sections 4.1, 4.2, 4.5, 4.6, 4.7

for the next class:

read [Bertsimas-Tsitsiklis]:
Section 4.4: Optimal dual variables as marginal costs

ex: simplex algorithm (LP doors & windows)

$$\begin{aligned} \min\ & -3x_D - 5x_W \\ \text{s.t.}\ & x_D + s_1 = 4 \\ & 2x_W + s_2 = 12 \\ & 3x_D + 2x_W + s_3 = 18 \\ & x_D, x_W, s_1, s_2, s_3 \geq 0 \end{aligned}$$



- $\beta_1 = (3, 4, 5)$:

SENSITIVE ANALYSIS

GOAL OF SENSITIVE ANALYSIS

models of real-world decision problems are often approximated:

- they rely on forecast/inaccurate data: a model is more reliable if its solutions are less sensitive to changes in the data
- they have incomplete knowledge of the problem: a model is more robust if its solutions are less sensitive to additions of variables/constraints

how to evaluate the sensitivity of an optimal solution of $P : min\{cx \mid Ax = b, x \geq 0\}$ to **one local change** in A, b or c without having to simulate every possible changes by solving from scratch the LP again and again ?

THE CORE IDEA

- let P in standard form $P : min\{cx \mid Ax = b, x \geq 0\}$
- when the simplex method stops with an optimal solution, it returns an optimal basis β and feasible primal and dual solutions x and u such that:

$$x = (x_\beta, x_{\neg\beta}) = (A_\beta^{-1}b, 0)$$

$$x_\beta \geq 0 \qquad \text{primal feasibility}$$

$$u^T = c_\beta^T A_\beta^{-1}$$

$$\bar{c}^T = c^T - u^T A \geq 0 \qquad \text{dual feasibility}$$

- when the problem changes, check how these conditions are affected

ADDING A NEW VARIABLE/COLUMN

- new variable x_{n+1} and column (c_{n+1}, A_{n+1}): like assuming $n+1 \notin \beta$ (with $x_{n+1} = 0$)
- β remains a basis and $x_\beta = A_\beta^{-1}b$, $x_{\neg\beta\cup\{n+1\}} = 0$ is primal feasible
- it remains optimal if $u^T = c_\beta^T A_\beta^{-1}$ is dual feasible, i.e.:

$$\bar{c}_{n+1} = c_{n+1} - c_\beta^T A_\beta^{-1} A_{n+1} \geq 0$$

  and the optimal value $c_\beta x_\beta$ does not change
- otherwise the $n+1$-th direction is improving and we must run additional iterations of the **primal** simplex algorithm from β to reach an optimal basis

EXAMPLE: ADDING A VARIABLE

$\beta = \{1,3\}$ optimal basis $x^T = (1,0,1)$, $u^T = (2,1)$ primal-dual feasible, $opt = 19$

$P: min\ 13x_1 + 10x_2 + 6x_3 + \delta x_4$	$D: max\ 8u_1 + 3u_2$
s.t. $5x_1 + x_2 + 3x_3 + x_4 = 8$	s.t. $5u_1 + 3u_2 \leq 13$
$3x_1 + x_2 + x_4 = 3$	$u_1 + u_2 \leq 10$
$x_1, x_2, x_3, x_4 \geq 0$	$3u_1 \leq 6$
	$u_1 + u_2 \leq \delta$

- β remains a basis, $x^T = (1,0,1,0)$ primal feasible
- $u^T = (2,1)$ remains feasible iff the new constraint is satisfied $u_1 + u_2 = 3 \leq \delta$
- optimal solutions and values do not change when $\delta \geq 3$

CHANGING THE RIGHT HAND SIDE VECTOR

- let $b'_k = b_k + \delta$, i.e. $b' = b + \delta e_k$ for some $k = 1, \dots, m$
- β remains a basis and $u^T = c_\beta^T A_\beta^{-1}$ remains dual feasible $(c^T - u^T A \geq 0)$
- β remains optimal if primal feasibility holds:

$$A_\beta^{-1} b' = A_\beta^{-1}(b + \delta e_k) = x_\beta + \delta h \geq 0$$

  where $h = A_\beta^{-1} e_k$ is the k-th column of A_β^{-1}
  and the optimal cost varies by $\delta u_k = u^T(b + \delta e_k) - u^T b$
- dual value u_k is the **marginal cost** (or shadow price) per unit increase of b_k
- otherwise we must run additional iterations of the **dual** simplex algorithm from β to reach an optimal basis

EXAMPLE: CHANGING b

$\beta = \{1,3\}$ optimal basis $x^T = (1,0,1)$, $u^T = (2,1)$ primal-dual feasible, $opt = 19$

$P: min\ 13x_1 + 10x_2 + 6x_3$	$D: max\ (8 + \delta)u_1 + 3u_2$
s.t. $5x_1 + x_2 + 3x_3 = 8 + \delta$	s.t. $5u_1 + 3u_2 \leq 13$
$3x_1 + x_2 = 3$	$u_1 + u_2 \leq 10$
$x_1, x_2, x_3 \geq 0$	$3u_1 \leq 6$

- β remains a basis, u^T remains dual feasible
- $x^T = (1, 0, 1 + \frac{\delta}{3})$ is feasible iff $1 + \frac{\delta}{3} \geq 0$
- $(1, 0, 1 + \frac{\delta}{3})$ is optimal while $\delta \geq -3$ and the optimum value is $19 + 2\delta$
- increasing b_1 by $\delta = 1$ unit leads to a marginal cost $u_1 = 2$

CHANGING THE COST OF A NON-BASIC VARIABLE

- let $c'_j = c_j + \delta$ for some non-basic $j \notin \beta$
- β remains a basis and $x_\beta = A_\beta^{-1} b \geq 0$ is primal feasible
- β remains optimal if $u^T = c_\beta^T A_\beta^{-1}$ is dual feasible:

$$\bar{c}'_j = (c_j + \delta) - u^T A_j = \bar{c}_j + \delta \geq 0$$

  and the optimal value $c_\beta x_\beta$ does not change
- **reduced cost** $\bar{c}_j$ is the cost reduction value from which j becomes improving
- otherwise j is an improving direction and we must run additional iterations of the **primal** simplex algorithm from β to reach an optimal basis

EXAMPLE: CHANGING c (NON-BASIC)

$\beta = \{1,3\}$ optimal basis $x^T = (1,0,1)$, $u^T = (2,1)$ primal-dual feasible, $opt = 19$

$$P: \min 13x_1 + (10+\delta)x_2 + 6x_3 \quad \text{s.t. } 5x_1 + x_2 + 3x_3 = 8,\ 3x_1 + x_2 = 3,\ x_1, x_2, x_3 \geq 0$$

$$D: \max 8u_1 + 3u_2 \quad \text{s.t. } 5u_1 + 3u_2 \leq 13,\ u_1 + u_2 \leq 10 + \delta,\ 3u_1 \leq 6$$

- β remains a basis, x^T remains primal feasible
- u^T remains feasible iff $u_1 + u_2 = 3 \leq 10 + \delta$
- optimal solutions and values do not change while $\delta \geq -7 = -\bar{c}_2$
- x_2 is profitable if c_2 is below $10 - \bar{c}_2 = 3$

CHANGING THE COST OF A BASIC VARIABLE

- let $c'_j = c_j + \delta$ for some basic $j \in \beta$ and j is the l-th element of β i.e. $c'_\beta = c_\beta + \delta e_l$
- β remains a basis and $x_\beta = A_\beta^{-1} b \geq 0$ is primal feasible
- β remains optimal if $u'^T = c'^T_\beta A_\beta^{-1}$ is dual feasible:

$$\bar{c}'^T_{\neg\beta} = c^T_{\neg\beta} - (c_\beta + \delta e_l)^T A_\beta^{-1} A_{\neg\beta} = \bar{c}^T_{\neg\beta} - \delta e_l^T A_\beta^{-1} A_{\neg\beta}$$
$$= \bar{c}^T_{\neg\beta} - \delta g \geq 0$$

where g is the l-th row of $A_\beta^{-1} A_{\neg\beta}$ (available in the simplex algorithm) and the optimal cost varies by $\delta x_j = (c'^T - c^T)x$

- x_j is the **marginal cost** per unit increase of c_j
- otherwise an improving direction exists and we must run additional iterations of the **primal** simplex algorithm from β to reach an optimal basis

EXAMPLE: CHANGING c (BASIC)

$\beta = \{1,3\}$ optimal basis $x^T = (1,0,1)$, $u^T = (2,1)$ primal-dual feasible, $opt = 19$

$$P: \min (13+\delta)x_1 + 10x_2 + 6x_3 \quad \text{s.t. } 5x_1 + x_2 + 3x_3 = 8,\ 3x_1 + x_2 = 3,\ x_1, x_2, x_3 \geq 0$$

$$D: \max 8u_1 + 3u_2 \quad \text{s.t. } 5u_1 + 3u_2 \leq 13 + \delta,\ u_1 + u_2 \leq 10,\ 3u_1 \leq 6$$

- β remains a basis, x^T remains primal feasible
- $u^T = (2, 1 + \frac{\delta}{3})$ is feasible iff $u_1 + u_2 = 2 + 1 + \frac{\delta}{3} \leq 10$, i.e. $\delta \leq 21$
- and the optimum value increases by $x_1 \delta = \delta$
- x_1 is less profitable than x_2 if c_1 is above $10 + 21 = 31$

ADDING A NEW INEQUALITY CONSTRAINT

- add a violated constraint $a^T_{m+1} x \geq b_{m+1}$; by substitution, assume that $a_{m+1,j} = 0 \, \forall j \notin \beta$
- add a slack variable x_{n+1} and get a new basis $\beta' = \beta \cup \{n+1\}$:

$$A_{\beta'} = \begin{pmatrix} A_\beta & 0 \\ a^T_{m+1} & -1 \end{pmatrix} \quad A^{-1}_{\beta'} = \begin{pmatrix} A_\beta^{-1} & 0 \\ a^T_{m+1} A_\beta^{-1} & -1 \end{pmatrix}$$

- $u^T = (c^T_\beta \ 0) A^{-1}_{\beta'} = (c^T_\beta A^{-1}_\beta \ 0)$ is feasible as the reduced costs are unchanged:

$$\bar{c}'^T = (c^T \ 0) - (c^T_\beta \ 0) A^{-1}_{\beta'} A = (\bar{c}^T \ 0)$$

- we must run additional iterations of the **dual** simplex algorithm to recover primal feasibility
- for an equality constraint, we introduce an artificial variable (as in the two-phase method)

EXAMPLE: ADDING A CONSTRAINT

$\beta = \{1,3\}$ optimal basis $x^T = (1,0,1)$, $u^T = (2,1)$ primal-dual feasible, $opt = 19$

$$\begin{aligned} P: min\ & 13x_1 + 10x_2 + 6x_3 \\ \text{s.t. } & 5x_1 + x_2 + 3x_3 = 8 \\ & 3x_1 + x_2 = 3 \\ & x_1 + x_3 + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

$$\begin{aligned} D: max\ & 8u_1 + 3u_2 + u_3 \\ \text{s.t. } & 5u_1 + 3u_2 + u_3 \leq 13 \\ & u_1 + u_2 \leq 10 \\ & 3u_1 + u_3 \leq 6 \\ & u_3 \leq 0 \end{aligned}$$

- $\beta = \{1,3,4\}$ is a basis, $u^T = (2,1,0)$ is dual feasible
- $x^T = (1,0,1,-1)$ is not primal feasible

CHANGING A NON-BASIC COLUMN

- let $a'_{ij} = a_{ij} + \delta$ for some non-basic $j \notin \beta$
- β remains a basis and $x_\beta = A_\beta^{-1} b \geq 0$ is primal feasible
- β remains optimal if $u^T = c_\beta^T A_\beta^{-1}$ is dual feasible:

$$\begin{aligned} \bar{c'}_j &= c_j - c_\beta^T A_\beta^{-1}(A_j + \delta e_i) \\ &= \bar{c}_j - \delta u_i \geq 0 \end{aligned}$$

  and the optimal value $c_\beta x_\beta$ does not change
- otherwise j is an improving direction and we must run additional iterations of the **primal** simplex algorithm from β to reach an optimal basis

EXAMPLE: CHANGING A_j (NON-BASIC)

$\beta = \{1,3\}$ optimal basis $x^T = (1,0,1)$, $u^T = (2,1)$ primal-dual feasible, $opt = 19$

$$\begin{aligned} P: min\ & 13x_1 + 10x_2 + 6x_3 \\ \text{s.t. } & 5x_1 + (1+\delta)x_2 + 3x_3 = 8 \\ & 3x_1 + x_2 = 3 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

$$\begin{aligned} D: max\ & 8u_1 + 3u_2 \\ \text{s.t. } & 5u_1 + 3u_2 \leq 13 \\ & (1+\delta)u_1 + u_2 \leq 10 \\ & 3u_1 \leq 6 \end{aligned}$$

- β remains a basis, x^T remains primal feasible
- u^T remains feasible iff $(1+\delta)u_1 + u_2 = 3 + \delta \leq 10$
- optimal solutions and values do not change while $\delta \leq 7 = \frac{\bar{c}_2}{u_1}$

CHANGING A BASIC COLUMN

- it's complicated...

APPLICATIONS IN COMPUTING

- parametric simplex method: solve parametric LPs (e.g. with regularization)
- (progressive) column generation: solve LPs with many variables without knowing them a priori
- (progressive) constraint generation: solve LPs with many variables without knowing them a priori
- change variable bounds: e.g. in branch-and-bound

EXERCISE (STEEL FACTORY)

- implement the primal and the dual models of steel factory with Gurobipy
- get the dual optimal values: `Constr.pi`
- get the slack values: `Constr.slack`
- get the reduced costs: `Var.rc`
- how to interpret a zero slack value ?
- how to interpret a non-zero reduced cost ? simulate the change
- how to interpret a non-zero dual value ? simulate the change
- play also with the attributes (see the Gurobi documentation):
  - `Var`: `VBasis, SAObjLow/Up, SALBLow/Up, SAUBLow/Up`
  - `Constr`: `CBasis, SASRHSLow/Up`

EXERCISE (STEEL FACTORY): NOTES

- a zero slack value for a mill: the corresponding dual value is the marginal cost of an extra hour of availability of the mill
- a negative reduced cost for a product (that is not in the solution): how much the unit price of the product have to be raised to make it profitable / the marginal cost of producing 1 unit of the product (if feasible)
- be careful with the signs as the model is not in standard form

READING:

to go further:

read [Bertsimas-Tsitsiklis]:
Section 5.1