MATHEMATICAL PROGRAMMING

Sophie Demassey (CMA, Mines ParisTech/PSL)
sophie.demassey@mines-paristech.fr
`http://sofdem.github.io`

Overview

introduction to optimisation

OSE: demandez le programme

introduction to linear programming

linear programming models

geometry and algebra of linear programming

the simplex method

duality

sensitive analysis

INTRODUCTION TO OPTIMISATION

DECISION IS OPTIMISATION

Among possible alternatives, the solutions, select the best one regarding a quantitative criterion, the objective, e.g.:

time : minimize travel or project duration

space : minimize distance, maximize occupation

money : minimize cost, maximize profit of operation/design

goods : minimize consumption, maximize production/transaction

equilibrium : minimize potential energy

implicit description ($\neq$ a list) of the solutions as math relationships between unknowns, the variables

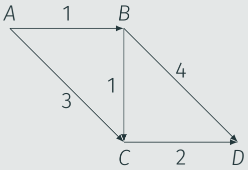
solutions $\equiv$ value-to-variable assignments satisfying the relationships

implicit description ($\neq$ a list) of the solutions as math relationships between unknowns, the variables

solutions $\equiv$ value-to-variable assignments satisfying the relationships

ex: routes of minimal duration

- a list of sequences of locations ($ABCD4, ABD5, ACD5$)
- a graph + restrictions



- a logical/arithmetic model

$$\min d_1 + d_2 + d_3$$
$$x_1x_2d_1 \in \{AB1, AC3\},$$
$$x_2x_3d_2 \in \{BC1, BD4, CD2\},$$
$$x_3x_4d_3 \in \{CD2, DD0\}$$

RECOGNIZE THE SOLUTIONS, MOST OF THE TIME...

feasibility

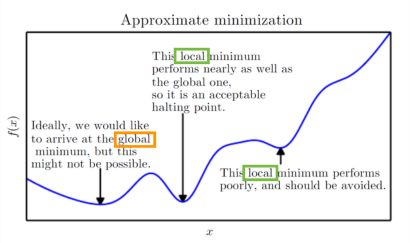
- models are approximate (routes ?)
- data are uncertain (duration ?)
- numeric computations are not exact (π ?)

feasibility

- models are approximate (routes ?)
- data are uncertain (duration ?)
- numeric computations are not exact (π ?)

optimality

- not provable in polynomial time
- provable within an error gap $\pm\epsilon$
- provable locally:



analytical methods results 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*)

SOLUTION METHODS

analytical methods results 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 assignments iteratively:

$$\begin{tikzcd}
\text{guess} \arrow[r, "k{=}0"'] & x_k \arrow[rr, bend left, "\text{evaluate}"] & & M(x_k) \arrow[ll, bend left, "k{+}{+}"] \arrow[r] & \text{stop}
\end{tikzcd}$$

- evaluate $M(x_k)$? simple arithmetic or numeric simulation
- direction/step length from x_k to x_{k+1} ? 0/1st/2nd order methods
- halting condition ?

operational research : system operation and design (transport, scheduling, packing, cutting, rostering, allocation,...)

graph theory : specialization to systems as graphs

optimal control : command $u(t)$ to optimize a trajectory $x(t)$ under dynamic constraints $x'(t) = c(x(t), u(t))$

machine learning : find a better match from a huge set of data

artificial intelligence : logic programming for robot decision

game theory : find the best strategies for multiple players with respective strategies

SOME CLASSES OF PROBLEMS

- deterministic problems / with uncertain data
- one / multiple objectives
- constrained / unconstrained
- analytic / logic models
- linear / convex / nonconvex models
- continuous / discrete values

OSE: demandez le programme

lectures & practice

(integer) linear prog.	8	SD, JPM	oct.
(integer) nonlinear prog.	3	WO, SD	oct-nov.
stochastic prog.	9	WO	jan.
prospective modelling	8	NM, EA, SS	nov.
machine learning	10	VR	nov-feb.

application & project

MILP: power plant provision	6	SD	oct
MINLP: water network operation	4	SD	nov-dec.
LP: market equilibrium	2	JPM	nov

Edi Assoumou (EA), Sophie Demassey (SD), Welington de Oliveira (WO), Nadia Maïzi (NM), Jean-Paul Marmorat (JPM), Valérie Roy (VR), Sandrine Selosse (SS).

agenda

1.introduction	27/09		
2.models	1/10	TP gurobipy	[Chap 1]*
3.geometry	3/10	TP models	[Chap 2]
4.simplex	3/10		[Chap 3]
5.duality	4/10	TP duality	[Chap 4]
6.sensitivity analysis	8/10	TP analysis	[Chap 5]
7.integer models	9/10	TP integer	[Chap 10]
8.integer solutions	10/10		[Chap 11]

*of [BERTSIMAS-TSITSIKLIS]: Bertsimas, Dimitris, and John N. Tsitsiklis. Introduction to linear optimization. Vol. 6, Athena Scientific, 1997.

INTRODUCTION TO LINEAR PROGRAMMING

programming = planning (military/industrial) operations

programming = planning (military/industrial) operations

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

- x: the decision variables
- $f : \mathbb{R}^n \to \mathbb{R}$: the objective function
- $g : \mathbb{R}^n \to \mathbb{R}^m$: the constraints

programming = planning (military/industrial) operations

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

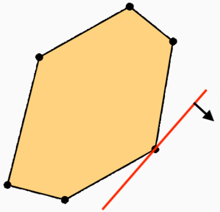
- x: the decision variables
- $f : \mathbb{R}^n \to \mathbb{R}$: the objective function
- $g : \mathbb{R}^n \to \mathbb{R}^m$: the constraints

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

- *min* $f(x)$, s.t $g(x) \geq 0, x \in \mathbb{R}^n$
- with linear constraints and objective: ex: $n = 3$, $m = 2$:

$$g(x) = \begin{pmatrix} 5x_1 + 3x_2 - 2x_3 - 4 \\ x_1 + x_2 + x_3 + 1 \end{pmatrix} = \begin{pmatrix} 5, & 3, & -2 \\ 1, & 1, & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} -4 \\ 1 \end{pmatrix}$$

- many optimization problems can be modeled this way (possibly with approximation or discrete variables)
- the restricted format makes the solution easy



conjunction of equality constraints linear in the nonnegative variables:[†]

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

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

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

[†]with $\mathbb{R}^n \equiv \mathbb{R}^{n \times 1}$ and $A^T \in \mathbb{R}^{n \times m}$ the transpose matrix

REDUCTION TO STANDARD FORM

replace by

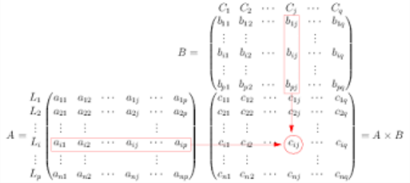
negative variable	$x \leq 0$	$x = -z, z \geq 0$
free variable	y free	$y = y^+ - y^-, y^+, y^- \geq 0$
slack constraint	$Ax \geq b$	$Ax - s = b$, $s \geq 0$
slack constraint	$Ey \leq f$	$Ey + u = f$, $u \geq 0$
maximisation	$max\ cx$	$-min(-c)x$

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

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

LINEAR ALGEBRA REVIEW AND NOTATION (1)

- **matrix** $A \in \mathbb{R}^{m\times n}$ with entry a_{ij} in row $i \leq m$, column $j \leq n$
- **transpose** $A^T \in \mathbb{R}^{n\times m}$ with $a^T_{ji} = a_{ij}$
- **vector** $\equiv$ column vector: $\mathbb{R}^n \equiv \mathbb{R}^{n\times 1}$
- **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 $(AB)^T = A^T B^T$



LINEAR ALGEBRA REVIEW AND NOTATION (2)

affine combination $\lambda_0 + \sum_{i=1}^{p} \lambda_i x^i$ of vectors $x^1, \dots, x^p \in \mathbb{R}^n$ with scalars $\lambda_0 \lambda_1, \dots, \lambda_p \in \mathbb{R}$

linear combination affine combination with $\lambda_0 = 0$

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

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

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

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

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

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

Reading:

to go further:

read [Bertsimas-Tsitsiklis]:

Section 1.1

for the next class:

read [Bertsimas-Tsitsiklis]:

Section 1.5: Linear algebra background

LINEAR PROGRAMMING MODELS

Verre

Une ligne de production de portes et fenêtres est composée de 3 postes de traitement. Les postes 1 (fabrication de portes), 2 (fabrication de fenêtres) et 3 (finitions) ouvrent respectivement 4 heures, 12 heures et 18 heures par semaine et ne peuvent traiter chacun, qu'un seul élément à la fois. La fabrication d'une porte occupe le poste 1 pendant 1h et le poste 3 pendant 3 heures. La fabrication d'une fenêtre occupe le poste 2 pendant 2h et le poste 3 pendant 2h. Avec un prix de vente de 3000 euros par porte et 5000 euros par fenêtre, comment la fabrique peut-elle maximiser ses revenus ?

ex 1: modèle PL

- variables de décision ?
    - x_P, x_F le nombre (fractionnaire) de portes et de fenêtres produites par jour
- contraintes ?
    - occupation des postes

$$
\begin{aligned}
\max\ & 3000x_P + 5000x_F \\
\text{s.t. }\ & x_P \leq 4 \\
& 2x_F \leq 12 \\
& 3x_P + 2x_F \leq 18 \\
& x_P, x_F \geq 0
\end{aligned}
$$

The two crude petroleum problem [Ralphs]

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

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

(first entry reads *producing 1 unit of gasoline requires 0.3 units of crude from Kuwait*)

Objective: minimize production cost.

ex 2: LP model

- decision variables ?
  - x_K, x_V the quantity (in thousands of barrel) to import from Kuwait or from Venezuela
- constraints ?
  - production and availability

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

steel factory

A factory produces steel in coils (*bobines*), tapes (*rubans*), and sheets (*tôles*) in the limit, respectively, of 6000 tons, 4000 tons and 3500 tons every week. The selling prices are, respectively, 25 euros, 30 euros and 2 euros per ton of product. The production involves two stages, heating (*réchauffe*) and rolling (*laminage*). These two mills are respectively available 35 hours and 40 hours a week. The amount of products (in tons) that can be processed in 1 hour by a mill is as follows:

	heating	rolling
coils	200	200
tapes	200	140
sheets	200	160

The factory wants to maximize its profit.

EX 3: 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

$$\max 25x_C + 30x_T + 2x_S$$

s.t.

$$\frac{x_C}{200} + \frac{x_T}{200} + \frac{x_S}{200} \leq 35$$
$$\frac{x_C}{200} + \frac{x_T}{140} + \frac{x_S}{160} \leq 40$$
$$0 \leq x_C \leq 6000$$
$$0 \leq x_T \leq 4000$$
$$0 \leq x_S \leq 3500$$

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 time to process one unit of waste is:

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

The unit profit for the company is 4000 euros for waste A and 8000 euros for waste B.

Objective: maximize the profit.

ex 4: LP model

- decision variables ?
    - x_A, x_B the number of units of waste of type A or B to process each month
- constraints ?
    - availability and operation

$$\begin{aligned} \max\ & 4000x_A + 8000x_B \\ \text{s.t.}\ & x_A + 3x_B \leq 450 \\ & 2x_A + x_B \leq 350 \\ & x_A + x_B \leq 200 \\ & x_A, x_B \geq 0 \end{aligned}$$

network flow

A company delivers every week retail stores in 9 cities in Europe from its unique factory *USINE*.

How to organize the production/ transportation in order to:

meet the demand of each store,

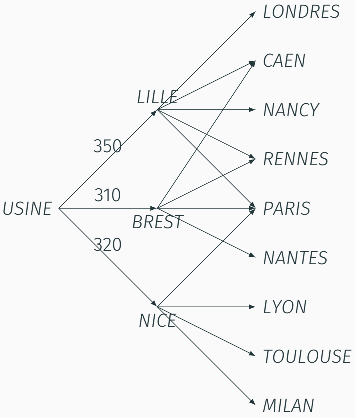
meet the maximum production,

meet the capacity of each line,

minimize the transportation cost ?

```
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
```

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



- $x_l, l \in LINES$ the quantity of products transported on line l
- $TRANSITS = \{LILLE, NICE, BREST\}$

$$
\begin{aligned}
\min \quad & \sum_{l \in LINES} UNITCOST_l x_l \\
\text{s.t.} \quad & \sum_{l \in LINES('USINE', TRANSITS)} x_l \leq MAX\_PROD, \\
& \sum_{l \in LINES(TRANSITS, i)} x_l \geq DEMAND_i, && \forall i \in STORES \\
& \sum_{l \in LINES('USINE', i)} x_l = \sum_{l \in LINES(i, STORES)} x_l, && \forall i \in TRANSITS \\
& 0 \leq x_l \leq CAPACITY_l, && \forall l \in LINES
\end{aligned}
$$

capacity planning [Bertsimas-Tsitsiklis]

find a least cost electric power capacity expansion plan:

- planning horizon: the next $T \in \mathbb{N}$ years, $t = 1, \ldots, T$
- forecast demand: $d_t \in \mathbb{R}_+$ MW for each year $t = 1, \ldots, T$
- existing capacity (oil-fired plants): $e_t \in \mathbb{R}_+$ MW available for each year t
- 2 alternatives for expanding capacities: (1) coal-fired plant, (2) nuclear plant
  - lifetime: l_j years, for each alternative $j = 1, 2$
  - capital cost: c_{jt} euros/MW installed in capacity j to be operational from year t
  - for political and safety reasons: no more than 20% of capacity should be nuclear

- decision variables ?
  - x_{jt}: capacity $j = 1, 2$ installed in MW to be operational from year t
- constraints ?
  - each year: total capacity meets the demand + nuclear rate
- available capacity $j = 1, 2$ for year t:
  $y_{jt} = \sum_{s=\max\{1,t-l_j+1\}}^{t} x_{js}$

- decision variables ?
  - x_{jt}: capacity $j = 1, 2$ installed in MW to be operational from year t
- constraints ?
  - each year: total capacity meets the demand + nuclear rate
- available capacity $j = 1, 2$ for year t:
  $y_{jt} = \sum_{s=\max\{1,t-l_j+1\}}^{t} x_{js}$

$$
\begin{aligned}
\min\ & \sum_{t=1}^{T} \sum_{j=1}^{2} c_{jt} x_{jt} \\
\text{s.t.}\ & y_{jt} = \sum_{s=\max\{1,t-l_j+1\}}^{t} x_{js}, && \forall j = 1, 2, t = 1, \dots, T \\
& e_t + y_{1t} + y_{2t} \geq d_t, && \forall t = 1, \dots, T \\
& 0.8 y_{2t} \leq 0.2 e_t + 0.2 y_{1t}, && \forall t = 1, \dots, T \\
& x_{jt} \geq 0, y_{jt} \geq 0, && \forall j = 1, 2, t = 1, \dots, T
\end{aligned}
$$

EX 6: IN STANDARD FORM

$$
\begin{aligned}
\min \; & \sum_{t=1}^{T} \sum_{j=1}^{2} c_{jt} x_{jt} \\
\text{s.t.} \; & y_{jt} - \sum_{s=\max\{1, t-l_j+1\}}^{t} x_{js} = 0, && \forall j = 1, 2, t = 1, \dots, T \\
& y_{1t} + y_{2t} - s_t^d = d_t - e_t, && \forall t = 1, \dots, T \\
& 0.8 y_{2t} - 0.2 y_{1t} + s_t^n = 0.2 e_t, && \forall t = 1, \dots, T \\
& x_{jt} \geq 0, y_{jt} \geq 0, s_t^d \geq 0, s_t^n \geq 0 && \forall j = 1, 2, t = 1, \dots, T
\end{aligned}
$$

$n = 6T$ variables, $m = 4T$ constraints.

minimize $\|\|_1$ and $\|\|_\infty$

1. Find a solution $x \in \mathbb{R}^n$ of the system of equation $Ax = b$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ with minimal 1-norm:

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

2. Find a solution $x \in \mathbb{R}^n$ of the system of equation $Ax = b$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ with minimal ∞-norm:

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

EX 7: LP MODELS $\min \|x\|_1 = \min \sum_j |x_j|$

- decision variables ?
    - the entries of $x \in \mathbb{R}^n$ and their absolute values $|x_j|$, $j = 1..n$
- how to *linearize* $|x|$, $x \in \mathbb{R}$?

EX 7: LP MODELS $\min \|x\|_1 = \min \sum_j |x_j|$

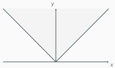
- decision variables ?
  - the entries of $x \in \mathbb{R}^n$ and their absolute values $|x_j|$, $j = 1..n$
- how to *linearize* $|x|, x \in \mathbb{R}$?
  1. $|x| = \min\{y + z \mid y, z \geq 0, x = y - z\}$,

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

partial linearization ! $min_{a \in A} \sum_j min_{b_j \in B_a^j} b_j = min_{a \in A, b_j \in B_a^j} \sum_j b_j$

EX 7: LP MODELS $\min \|x\|_1 = \min \sum_j |x_j|$

- decision variables ?
  - the entries of $x \in \mathbb{R}^n$ and their absolute values $|x_j|$, $j = 1..n$
- how to *linearize* $|x|$, $x \in \mathbb{R}$?
  1. $|x| = \min\{y + z \mid y, z \geq 0, x = y - z\}$,
  2. $|x| = \max\{x, -x\} = \min\{y \geq 0 \mid y \geq x, y \geq -x\}$



$$(1) \min \sum_{j=1}^{n} y_j + z_j$$

$$\begin{aligned} \text{s.t.}\quad & x_j = y_j - z_j, && \forall j \\ & Ax = b, && \\ & y_j, z_j \geq 0, && \forall j \end{aligned}$$

$$(2) = \min \sum_{j=1}^{n} y_j$$

$$\begin{aligned} \text{s.t.}\quad & y_j \geq x_j, && \forall j \\ & y_j \geq -x_j, && \forall j \\ & Ax = b, && \\ & y_j \geq 0, && \forall j \end{aligned}$$

EX 7: LP MODEL $\min \|x\|_\infty = \min \max_j |x_j|$

- $|x_j| = \min\{y_j \geq 0 \mid y_j \geq x_j, y_j \geq -x_j\}$
- $\max_j |x_j| = \min\{y \geq 0 \mid y \geq y_j, y_j \geq x_j, y_j \geq -x_j \ \forall j\}$

$$
\begin{aligned}
\min\ & y & \\
\text{s.t.}\ & y \geq x_j, & \forall j \\
& y \geq -x_j, & \forall j \\
& Ax = b, & \\
& y \geq 0, & \forall j
\end{aligned}
$$

EX 7: NORMS

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

data fitting [Bertsimas-Tsitsiklis]

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

data fitting

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

A "best" linear fit:

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

with $x \in \mathbb{R}^n$, $y \geq 0$ minimizing the *prediction error* (or *residual*) $|b(a_i) - b_i|\ \forall i = 1..m$, e.g.:

1. minimizing the largest residual: $max_i |b(a_i) - b_i|$
2. minimizing the sum of the residuals: $\sum_i |b(a_i) - b_i|$

1. minimize the largest residual: $min\ max_i|a_i * x + y - b_i|$
2. minimize the sum of the residuals: $min\ \sum_i |a_i * x + y - b_i|$

$$
\begin{aligned}
(1) & min\ z \\
\text{s.t. } & z \geq a_i x + y - b_i, && \forall i \\
& z \geq b_i - a_i x - y, && \forall i \\
& z \geq 0, && \forall i \\
& x \in \mathbb{R}^n, y \in \mathbb{R}
\end{aligned}
$$

$$
\begin{aligned}
(2)\ \min & \sum_i z_i \\
\text{s.t. } & z_i \geq a_i x + y - b_i, && \forall i \\
& z_i \geq b_i - a_i x - y, && \forall i \\
& z_i \geq 0, && \forall i \\
& x \in \mathbb{R}^n, y \in \mathbb{R}
\end{aligned}
$$

to go further:

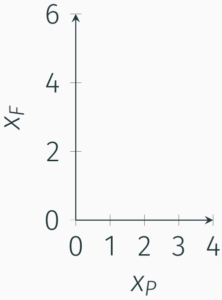
read [Bertsimas-Tsitsiklis]:
Sections 1.2, 1.3, 1.4

for the next class:

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

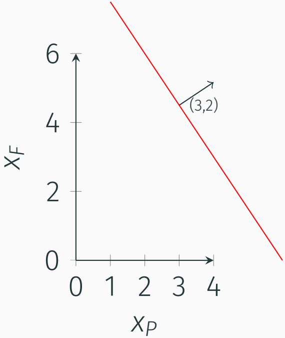
GEOMETRY AND ALGEBRA OF LINEAR PROGRAMMING

$$
\begin{aligned}
\max\ & 3000x_P + 5000x_F \\
\text{s.t.}\ & x_P \leq 4 \\
& 2x_F \leq 12 \\
& 3x_P + 2x_F \leq 18 \\
& x_P, x_F \geq 0
\end{aligned}
$$



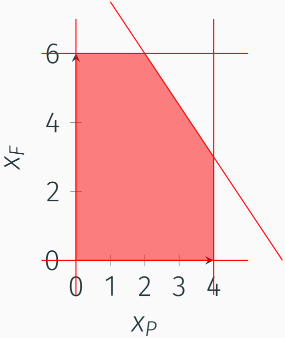
- solution space $\mathbb{R}^2$

$$\begin{aligned} \max\ & 3000x_P + 5000x_F \\ \text{s.t.}\ & x_P \leq 4 \\ & 2x_F \leq 12 \\ & 3x_P + 2x_F \leq 18 \\ & x_P, x_F \geq 0 \end{aligned}$$



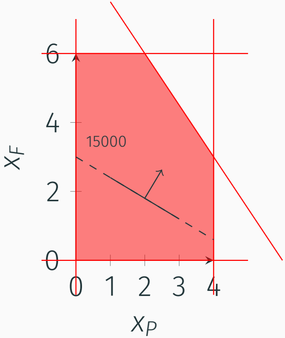
- solution space $\mathbb{R}^2$
- linear constraints $\equiv$ halfspaces: $\{x \in \mathbb{R}^2 \mid 3x_P + 2x_F \leq 18\}$

$$
\begin{aligned}
\max\ & 3000x_P + 5000x_F \\
\text{s.t.}\ & x_P \leq 4 \\
& 2x_F \leq 12 \\
& 3x_P + 2x_F \leq 18 \\
& x_P, x_F \geq 0
\end{aligned}
$$



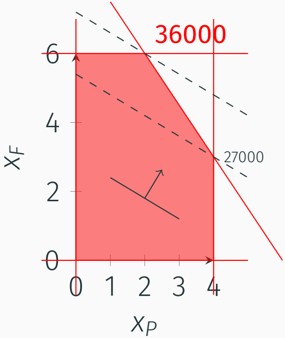
- solution space $\mathbb{R}^2$
- linear constraints $\equiv$ halfspaces: $\{x \in \mathbb{R}^2 \mid 3x_P + 2x_F \leq 18\}$
- feasible region $\equiv$ intersection of a finite number of halfspaces $\triangleq$ polyhedron in $\mathbb{R}^2$

$$\begin{aligned} \max\ & 3000x_P + 5000x_F \\ \text{s.t.}\ & x_P \leq 4 \\ & 2x_F \leq 12 \\ & 3x_P + 2x_F \leq 18 \\ & x_P, x_F \geq 0 \end{aligned}$$



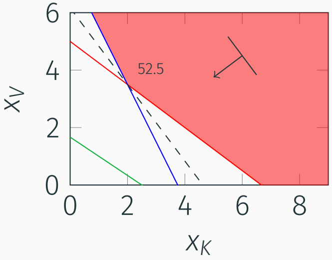
- solution space $\mathbb{R}^2$
- linear constraints $\equiv$ halfspaces: $\{x \in \mathbb{R}^2 \mid 3x_P + 2x_F \leq 18\}$
- feasible region $\equiv$ intersection of a finite number of halfspaces $\triangleq$ polyhedron in $\mathbb{R}^2$
- objective: $z = 3000x_P + 5000x_F$

$$\begin{aligned} \max\ & 3000x_P + 5000x_F \\ \text{s.t.}\ & x_P \leq 4 \\ & 2x_F \leq 12 \\ & 3x_P + 2x_F \leq 18 \\ & x_P, x_F \geq 0 \end{aligned}$$



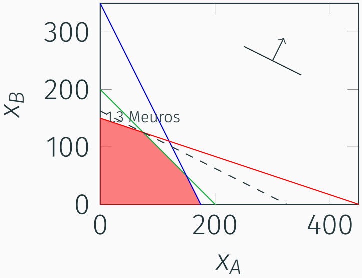
- solution space $\mathbb{R}^2$
- linear constraints $\equiv$ halfspaces: $\{x \in \mathbb{R}^2 \mid 3x_P + 2x_F \leq 18\}$
- feasible region $\equiv$ intersection of a finite number of halfspaces $\triangleq$ polyhedron in $\mathbb{R}^2$
- objective: $z = 3000x_P + 5000x_F$
- optimum: move the line up as long as feasible

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

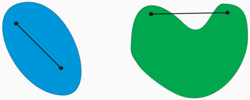


- constraint $0.2x_K + 0.3x_V \geq 0.5$ is redundant
- constraints $0.3x_K + 0.4x_V \geq 2$ and $0.4x_K + 0.2x_V \geq 1.5$ are active/binding at the optimum $(2, 3.5)$ but not constraints $x_k \geq 0$ or $x_V \leq 6$

$$\begin{aligned} \max\ & 4000x_A + 8000x_B \\ \text{s.t.}\ & x_A + 3x_B \leq 450 \\ & 2x_A + x_B \leq 350 \\ & x_A + x_B \leq 200 \\ & x_A, x_B \geq 0 \end{aligned}$$

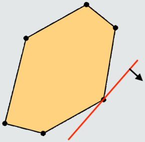


- the feasible region is defined as a polyhedron
- thus it is convex (intersection of convex regions)



where are the optimal solutions ?

intuition: the optimum of a linear function cx on a polyhedron $\mathcal{P} = \{x \in \mathbb{R}^n | Ax \geq b\}$ is reached at a "*corner point*" (under conditions of existence)



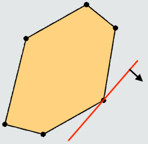
idea: solving an LP = evaluate corner points

Theorem

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

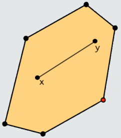
$\hat{x}$ is a

vertex



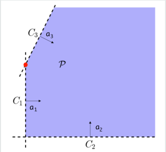
$\exists c, \forall x \in \mathcal{P}, c^T\hat{x} < c^T x$

extreme point



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

basic feasible solution



$\exists n$ linearly independent rows a_i, $i \in I$ s.t. $a_i x = b_i$

vertices and extreme points are independent of the model of $\mathcal{P}$

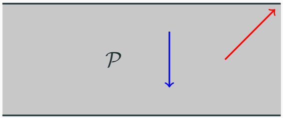
their number is finite, at most $\binom{m}{n}$, but probably large

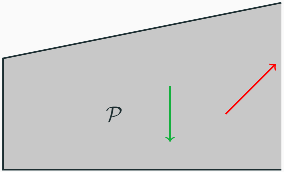
Theorem: existence of an extreme point

A nonempty polyhedron $\mathcal{P}$ has at least one extreme point
$\iff$ it has no line: $\forall x \in \mathcal{P}, d \in \mathbb{R}^n,\ \{x + \theta d | \theta \in \mathbb{R}\} \notin \mathcal{P}$
$\iff$ A has n linearly independent rows

Theorem: existence of an optimal solution

A minimization LP over a nonempty polyhedron $\mathcal{P}$. Either optimal cost is $-\infty$ or there exists an optimal solution which is an extreme point (e.g. for the standard form)





unbounded ∞ optima / 0 vertex optima including 1 vertex show that:

- $\mathcal{P} = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$ is nonempty and has no extreme point
- $5(x + y)$ 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

show that:

- $\mathcal{P} = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$ is nonempty and has no extreme point
- $5(x + y)$ 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

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 in $(0, 0, 0, 0)$

Theorem

A nonempty polyhedron **in standard form** $\mathcal{P} = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ with m linear independent rows $A \in \mathbb{R}^{m \times n}$ (i.e. no redundant constraints): $x \in \mathbb{R}^n$ 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. $x_j = 0, \forall j \notin \beta$.

Find a basic (perhaps not feasible) 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 $\mathbb{R}^m$: $A_{|\beta} x_{|\beta} = b$

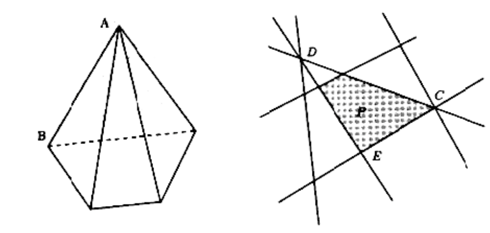
The columns A_j, $j \in \beta$ is a basis of $\mathbb{R}^m$ and form an invertible basis matrix $A_{|\beta} \in \mathbb{R}^{m \times m}$; x_j, $j \in \beta$ are the basic variables

2 different basic solutions correspond to 2 different bases but a basic solution is degenerate

$\iff$ more than n active constraints

$\iff$ more that $n - m$ variables to 0 (in standard form)

$\implies$ different bases



basic nonfeasible degenerate ?

basic feasible nondegenerate ?

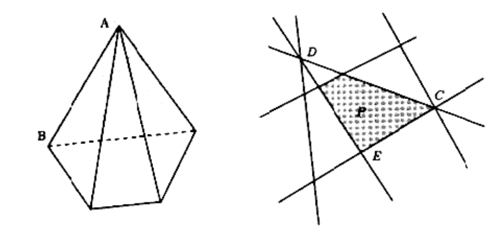
basic feasible degenerate ?

2 different basic solutions correspond to 2 different bases but a basic solution is degenerate

$\iff$ more than n active constraints

$\iff$ more that $n - m$ variables to 0 (in standard form)

$\implies$ different bases



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

EX: BASIC SOLUTION (CAPACITY PLANNING)

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} \sum_{j=1}^{2} c_{jt} x_{jt} \\
\text{s.t.} \quad & y_{jt} - \sum_{s=\max\{1, t-l_j+1\}}^{t} x_{js} = 0, && \forall j = 1, 2, t = 1, \ldots, T \\
& y_{1t} + y_{2t} - s_t^d = d_t - e_t, && \forall t = 1, \ldots, T \\
& 0.8 y_{2t} - 0.2 y_{1t} + s_t^n = 0.2 e_t, && \forall t = 1, \ldots, T \\
& x_{jt} \geq 0, y_{jt} \geq 0, s_t^d \geq 0, s_t^n \geq 0 && \forall j = 1, 2, t = 1, \ldots, T
\end{aligned}
$$

$$
\begin{pmatrix}
L & 0 & I & 0 & 0 & 0 \\
0 & L & 0 & I & 0 & 0 \\
0 & 0 & I & I & -I & 0 \\
0 & 0 & -0.2I & 0.8I & 0 & I
\end{pmatrix}
\begin{pmatrix}
x_1 \\ x_2 \\ y_1 \\ y_2 \\ s^d \\ s^n
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ d - e \\ 0.2e
\end{pmatrix}
$$

$n = 6T$ variables, $m = 4T$ linearly independent rows

EX: BASIC SOLUTION (CAPACITY PLANNING)

$$
\begin{aligned}
\min \; & \sum_{t=1}^{T} \sum_{j=1}^{2} c_{jt} x_{jt} \\
\text{s.t.} \; & y_{jt} - \sum_{s=\max\{1,t-l_j+1\}}^{t} x_{js} && = 0, && \forall j = 1, 2, t = 1, \ldots, T \\
& y_{1t} + y_{2t} - s_t^d && = d_t - e_t, && \forall t = 1, \ldots, T \\
& 0.8 y_{2t} - 0.2 y_{1t} + s_t^n && = 0.2 e_t, && \forall t = 1, \ldots, T \\
& x_{jt} \geq 0, y_{jt} \geq 0, s_t^d \geq 0, s_t^n \geq 0 && && \forall j = 1, 2, t = 1, \ldots, T
\end{aligned}
$$

$$
\begin{pmatrix}
L & 0 & I & 0 & 0 & 0 \\
0 & L & 0 & I & 0 & 0 \\
0 & 0 & I & I & -I & 0 \\
0 & 0 & -0.2I & 0.8I & 0 & I
\end{pmatrix}
\begin{pmatrix}
x_1 \\ x_2 \\ y_1 \\ y_2 \\ s^d \\ s^n
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ d - e \\ 0.2e
\end{pmatrix}
$$

basic solution $(0, 0, 0, 0, e - d, 0.2e)$ is feasible iff $e_t \geq d_t, \forall t$, degenerate ($4T > n - m$ zeros), other basis e.g (x_1, x_2, s^d, s^n)

EX: BASIC SOLUTION (CAPACITY PLANNING)

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

EX: BASIC SOLUTION (CAPACITY PLANNING)

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

$$
\begin{aligned}
\min\ & \sum_{t=1}^{T}\sum_{j=1}^{2} c_{jt}x_{jt} \\
\text{s.t.}\ & \sum_{s=\max\{1,t-l_1+1\}}^{t} x_{1s} + \sum_{s=\max\{1,t-l_2+1\}}^{t} x_{2s} - s_t^d = d_t - e_t, && \forall t = 1,\dots,T \\
& 0.8 \sum_{s=\max\{1,t-l_2+1\}}^{t} x_{2s} - 0.2 \sum_{s=\max\{1,t-l_1+1\}}^{t} x_{1s} + s_t^n = 0.2e_t, && \forall t = 1,\dots,T \\
& x_{jt} \geq 0, s_t^d \geq 0, s_t^n \geq 0 && \forall j = 1,2, t = 1,\dots,T
\end{aligned}
$$

$$
\begin{pmatrix} L & L & -I & 0 \\ -0.2L & 0.8L & 0 & I \end{pmatrix}
\begin{pmatrix} x_1 \\ x_2 \\ s^d \\ s^n \end{pmatrix}
=
\begin{pmatrix} d-e \\ 0.2e \end{pmatrix}
$$

EX: BASIC SOLUTION (CAPACITY PLANNING)

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

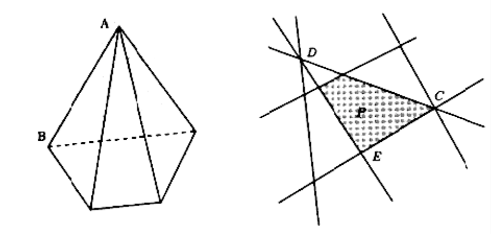
$$
\begin{aligned}
\min\ & \sum_{t=1}^{T} \sum_{j=1}^{2} c_{jt} x_{jt} \\
\text{s.t.}\ & \sum_{s=\max\{1,t-l_1+1\}}^{t} x_{1s} + \sum_{s=\max\{1,t-l_2+1\}}^{t} x_{2s} - s_t^d = d_t - e_t, && \forall t = 1, \dots, T \\
& 0.8 \sum_{s=\max\{1,t-l_2+1\}}^{t} x_{2s} - 0.2 \sum_{s=\max\{1,t-l_1+1\}}^{t} x_{1s} + s_t^n = 0.2 e_t, && \forall t = 1, \dots, T \\
& x_{jt} \geq 0, s_t^d \geq 0, s_t^n \geq 0 && \forall j = 1, 2, t = 1, \dots, T
\end{aligned}
$$

$$
\begin{pmatrix} L & L & -I & 0 \\ -0.2L & 0.8L & 0 & I \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ s^d \\ s^n \end{pmatrix} = \begin{pmatrix} d - e \\ 0.2e \end{pmatrix}
$$

basic solution $(0, 0, e - d, 0.2e)$ is feasible iff $e_t \geq d_t, \forall t$, degenerate iff $\exists t$, $e_t = 0$ or $e_t = d_t$

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



- 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

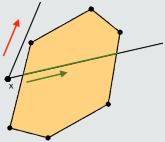
THE SIMPLEX METHOD

- optimal solutions of $\min_{x \in \mathcal{P}} cx$, $\mathcal{P} = \{Ax = b, x \geq 0\}$, $A \in \mathbb{R}^{m \times n}$, $rk(A) = m$ are basic feasible solutions
- to construct a basic solution: choose a basis $\beta \subseteq \{1, \ldots, n\}$ of m linearly independent columns of A, then solve $x_\beta = A_\beta^{-1} b$, $x_{\neg\beta} = 0$
- to move to an adjacent basic solution: replace one basic column: $\beta' = \beta \cup \{j'\} \setminus \{j''\}$
- the solutions may coincide if degenerate (if $x_{j'} = 0$)

the simplex method: move between adjacent basic feasible solutions while the cost decreases

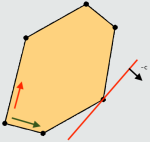
feasible direction from $x \in \mathbb{R}^n$

$d \in \mathbb{R}^n$ 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 $cd < 0$



feasible improving direction leads to $x' = x + \theta d \in \mathcal{P}, cx' < cx$

FEASIBLE IMPROVING BASIC DIRECTION

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

j'th basic direction

$d \in \mathbb{R}^n$: $d_{j'} = 1$, $d_j = 0, \forall j \notin \beta \cup \{j'\}$, $d_\beta = -A_\beta^{-1} A_{j'}$

feasible direction if x nondegenerate:

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

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

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

- $\bar{c}_{j'} = cd$ is the cost increase $c^T x' - c^T x$ when $x'_{j'} = \theta = 1$
- d is an improving direction iff $\bar{c}_{j'} < 0$
- for $j \in \beta$, then $\bar{c}_j = c_j - c_\beta^T A_\beta^{-1} A_j = c_j - c_\beta^T e_j = 0$

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

- $m = 2$, $n = 4$, $rk(A) = 2$
- $\beta = \{1, 2\}$ is a basis

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

- $m = 2$, $n = 4$, $rk(A) = 2$
- $\beta = \{1, 2\}$ is a basis
- $x = (1, 1, 0, 0)$ feasible nondegen.
- basic direction $j = 3$: $d_3 = 1$, $d_4 = 0$, $Ad = 0 \implies d_\beta = \binom{-3/2}{1/2}$
- improving direction: $\bar{c} = c^T d = -3/2 < 0$

STEP LENGTH θ

Let x be a nondegenerate basic feasible solution of basis β, and $j' \notin \beta$ of feasible improving direction d, i.e. $\bar{c}_{j'} < 0$

Theorem

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 basic feasible of basis $\beta' = \beta \cup \{j'\} \setminus \{j''\}$:
j' enters the basis, j'' exits the basis.

- $x' \in \mathcal{P}$ iff $x_j + \theta d_j \geq 0, \forall j \in \beta$ and $\theta \nearrow \implies c(x + \theta d) \searrow$
- step length is the largest value $\leq x_j/(-d_j), \forall j \in \beta, d_j < 0$
- β' is a basis: $A_\beta^{-1} A_j = e_j, \forall j \in \beta \setminus \{j''\}$, and $A_\beta^{-1} A_{j'} = -d_\beta$ has a nonzero j'' component $\implies \{A_j, j \in \beta'\}$ are linear independent
- $\theta > 0$ since x nondegenerate ($x_\beta > 0$)
- several possible entering columns j''

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

- $\beta = \{1, 2\}$ is a basis: $x = (1, 1, 0, 0)$ feasible nondegen.
- basic feasible improving direction $j = 3$: $d = (-3/2, 1/2, 1, 0)$
- $x_\beta + \theta d_\beta \geq 0 \implies x'_1 = 1 - (3/2)\theta \geq 0 \implies \theta \leq 2/3$
- $x' = (0, 4/3, 2/3, 0)$ basic feasible solution $\beta' = \{2, 3\}$, $cx' = cx + \theta\bar{c} = cx - 1$

Theorem

Let x be a basic feasible solution of basis β and $\bar{c} \in \mathbb{R}^n$ the corresponding 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 $\bar{c} \geq 0$

($\Rightarrow$) for any $y \in \mathcal{P}$: $d = y - x \implies Ad = 0 \implies$
$d_\beta = -\sum_{j \notin \beta} A_\beta A_j y_j \implies c^T(y - x) = \sum_{j \notin \beta} \bar{c}_j y_j \geq 0$ if $\bar{c} \geq 0$
($\Leftarrow$) if x nondegenerate and $\bar{c}_j < 0$, then j is nonbasic and of feasible improving direction, then x nonoptimal

example: basic improving direction (continuation)

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

- remark that optimum≥ 2 since $cx = x_1 + 2$, $\forall x$ feasible
- $\beta = \{2, 3\}$ is a basis: $x = (0, 4/3, 2/3, 0)$ nondegenerate
- basic directions:
    - $j = 1$: $d = (1, -1/3, -2/3, 0)$ and $\bar{c}_1 = cd = 1 \geq 0$
    - $j = 4$: $d = (0, 1/3, -4/3, 1)$ and $\bar{c}_4 = cd = 0 \geq 0$
- $\implies$ x optimal

steps	howto:
1. get a basis β	find m lin. indep. columns
2. get a basic feasible x	?
halt condition (optimality)	$\bar{c} = (c_j - c_\beta^T A_\beta^{-1} A_j)_j \geq 0$ if nondeg.
3. find an improving direction	any $j' \notin \beta$ s.t. $\bar{c}_{j'} < 0$ if nondeg.
halt condition (unboundness)	$d_\beta = -A_\beta^{-1} A_{j'} \geq 0$
4. find the largest step length	any $j'' \in argmin_{j \in \beta, d_j < 0} x_j / \lvert d_j \rvert$
5. update the basis	j'' enters, j' exits
6. goto 2.	$x := x - (x_{j''}/d_{j''})d$

convergence

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

- cx decreases at every iteration, all x are basic feasible solutions, the number of basic feasible solutions is finite choice of the entering column $j' \notin \beta$ s.t. $\bar{c}_{j'} < 0$:

- 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 $j'' \in argmin_{j \in \beta, d_j < 0} x_j / |d_j|$

$\rightarrow$ A trade-off between computation burden and efficiency, e.g. compute a subset of reduced costs

in case of degenerary ?

- if x degenerate and $\exists j \in \beta, d_j < 0, x_j = 0$ 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 may cycle
- cycles can be avoided (and convergence ensured) using 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 $\mathcal{P} = \{Ax \leq b, x \geq 0\}$, we can form the basis with the slack variables: $\mathcal{P} = \{Ax + Is = b, x \geq 0, s \geq 0\}$
- if the LP $min\{cx, Ax = b, x \geq 0\}$ is already in standard form, then we can first solve the auxiliary LP:

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

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

- each iteration involves costly arithmetic operations: computing $c_\beta^T A_\beta^{-1}$ (i.e. solving $p^T A_\beta = c_\beta^T$) or $A_\beta^{-1} A_j$ take $O(m^3)$ operations; computing $\bar{c}_j = c_j - pA_j$ 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 $A_{\beta^{-1}}(b|A)$
- the complexity in the worst case remains exponential: the LP may have 2^n extreme points and the simplex method visits them all
- good implementations of the simplex method perform usually very well

(see [BERTSIMAS-TSITSIKLIS]Section 3.3 for details)

to go further:

read [Bertsimas-Tsitsiklis]:

Sections 3.1, 3.2, 3.3

for the next class:

read [Bertsimas-Tsitsiklis]:

Section 1.6: Algorithms and operation count

DUALITY

DUALITY: MOTIVATION

$$P : z = min\ x^2 + y^2$$
$$\text{s.t.} \quad x + y = 1$$

- unconstrained variant: $P_u = z_u = min\ x^2 + y^2 + u(1 - x - y)$
- we relax the constraint and penalize its violation with a price $u \in \mathbb{R}$
- $z_u \leq z$: (x, y) feasible for $P \implies$ feasible for P_u and $z_u \leq x^2 + y^2 + u(1 - x - y) = x^2 + y^2$
- the optimal solution of P_u is $(u/2, u/2)$ (zero of the partial derivative)
- it is also feasible for P iff $u = x + y = 1$
- thus $(1/2, 1/2)$ is optimal for P

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

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

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

- lagrangian problems P_u, $u \in \mathbb{R}^m$: lower bounds $z_u \leq z$
- dual problem $D : d = max_{u \in \mathbb{R}^m} z_u$: the tightest lower bound
- if x is optimal for some P_u and satisfies $Ax = b$ then x is optimal for P and $d = z$

for LPs: $z = d$ (strong duality) and D is a linear program, the dual of which being P

Theorem

- the dual D of $P = min\{cx | Ax = b, x \geq 0\}$ is a linear program:

$$d = max\ u^T b$$
$$\text{s.t.}\quad u^T A \leq c^T$$

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

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

primal/dual correspondence

min	max
equality constraint	free variable
inequality constraint	nonnegative variable
cost vector (c)	RHS vector (b)
matrix (A)	matrix (A)

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

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

EX: DUAL MODEL (STEEL FACTORY)

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

$$
\begin{aligned}
D: \min\ & 35u_H + 40u_R + 6000u_C + 4000u_T + 3500u_S \\
\text{s.t.}\ & \\
& \frac{u_H}{200} + \frac{u_R}{200} + u_C \geq 25 && (\textit{coils}) \\
& \frac{u_H}{200} + \frac{u_R}{140} + u_T \geq 30 && (\textit{tapes}) \\
& \frac{u_H}{200} + \frac{u_R}{160} + u_S \geq 2 && (\textit{sheets}) \\
& u \geq 0
\end{aligned}
$$

theorem

- if x is feasible for P and u is feasible for D 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

$$sign(u_i) = sign(a_i^T x - b_i)$$
$$sign(x_j) = sign(c_j - u^T A_j)$$
$$\implies 0 \leq \sum_{i=1}^{m} u_i(a_i^T x - b_i) + \sum_{j=1}^{n} (c - u^T A_j) x_j = ub - cx$$

theorem

if a linear programming problem has an optimal solution, so does its dual and their respective optima are equal

- let x an optimal solution of $P = min\{cx|Ax = b, x \geq 0\}$ in standard form of basis β
- x optimal then $\bar{c} = c^T - c_\beta^T A_\beta^{-1} A \geq 0$
- let $u = c_\beta^T A_\beta^{-1}$ then u is feasible for $D = max\{u^T b|u^T A \leq c\}$
- $u^T b = c_\beta^T A_\beta^{-1} b = c_\beta x_\beta = cx$ then u is optimal for D
- at optimality: primal reduced costs = dual slacks

complementary slackness

theorem

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

$$u_i(a_i^T x - b^i) = 0 \quad \forall i$$
$$(c_j - u^T A_j)x_j = 0 \forall j.$$

- $u_i(a_i^T x - b^i) \geq 0 \ \forall i$ and $(c_j - u^T A_j)x_j \geq 0 \ \forall j$
- $u^T b - c^T x = \sum_i u_i(a_i^T x - b^i) + \sum_j (c_j - u^T A_j)x_j$
- either a constraint is binding at the optimum or the dual variable is zero

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

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

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

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

- $\beta = \{1, 3\} \implies x_2 = 0, x_1 = 3/3 = 1, x_3 = (8 - 5)/3 = 1$
- $x = (1, 0, 1) \geq 0 \implies$ feasible, nondegenerate
- P in standard form $\implies$ first C.S. condition satisfied
- second C.S. condition: $5u_1 + 3u_2 = 13$ and $3u_1 = 6$
- $u = (2, 1)$ is feasible for D since $u_1 + u_2 \leq 10$
- C.S. $\implies$ x and u are optimal with cost 19

x is optimal for $P = min\{cx|Ax = b, x \geq 0\}$ if $\exists u \in \mathbb{R}^m$:

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

- The basic feasible solutions of the simplex algorithm always satisfy 1,2 and 4 with $u^T = c_\beta^T A_\beta^{-1}$ because $\bar{c}_j = c_j - u^T A_j = 0 \; \forall j \in \beta$. Condition 3 is the halting condition $\bar{c} \geq 0$
- if x degenerate then the solutions u of $u^T A_j = c_j \; \forall j \mid x_j > 0$ may not be unique

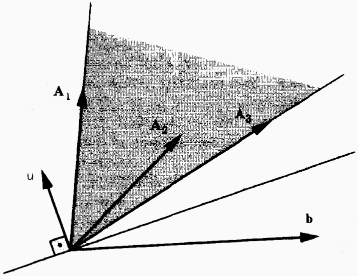
for $P = min\{cx|Ax = b, x \geq 0\}$ and $D = max\{u^T b|u^T A \leq c\}$

- 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
- 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 changing b or adding a new constraint to P, run the dual simplex starting from the feasible dual solution $c_\beta^T A_\beta^{-1}$

theorem

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Exactly one of the following holds:
(1) $\mathcal{P} = \{x \geq 0, Ax = b\} \neq \emptyset$
(2) $\exists u$, $u^T A \geq 0$ and $u^T b < 0$

- ($\Longleftarrow$) $x \in \mathcal{P}$ and $u^T A \geq 0 \implies u^T b = u^T Ax \geq 0$
- ($\implies$) $P = max\{0|Ax = b, x \geq 0\}$ unfeasible $\Rightarrow D = min$ $\{u^T b|u^T A \geq 0\}$ is $-\infty$ or unfeasible. 0 is feasible $\Rightarrow$ (2)



if b is not a linear comb of $(A_j)_j$ then $\exists$ a separating hyperplane $\{z|u^T z = 0\}$

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

SENSITIVE ANALYSIS

GOAL OF SENSITIVE ANALYSIS

- models of real problems are often approximations as the data are not accurate: a model is more reliable if its solutions are less sensitive to changes in the data
- models are often incomplete as constraints/variables are not all known in advance: a model is more robust if its solutions are less sensitive to new additions

how to evaluate the sensitivity of an optimal solution to one local change in A, b or c without having to simulate every possible changes by resolving the LP from scratch ?

- 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:

$$x_\beta = A_\beta^{-1} b \geq 0 \text{ and } \bar{c}^T = c^T - u^T A = c^T - c_\beta^T A_\beta^{-1} A \geq 0$$

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

ADDING A NEW VARIABLE/COLUMN

- adding a new variable $x_{n+1} \iff$ considering $n+1 \notin \beta$
- β remains a basis and $x_\beta = A_\beta^{-1} b$ remains feasible
- they remain optimal if $\bar{c}_{n+1} = c_{n+1} - c_\beta^T A_\beta^{-1} A_{n+1} \geq 0$ and the optimal value $c_\beta x_\beta$ does not change
- otherwise $n+1$ is an improving direction and we must run additional iterations of the primal simplex algorithm from β to reach an optimal basis

CHANGING THE RIGHT HAND SIDE VECTOR

- let $b'_k = b_k + \delta$, i.e. $b = b + \delta e_k$
- β remains a basis and $u^T = c_\beta^T A_\beta^{-1}$ remains dual feasible
- the primal feasibility condition is:

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

  where h is the k-th column of A_β^{-1}
- β remains optimal if

$$max_{i|h_i>0}(-x_\beta(i)/h_i) \leq \delta \leq min_{i|h_i<0}(-x_\beta(i)/h_i)$$

- the optimal cost varies by $u^T b - u^T(b + \delta e_k) = u_k \delta$: i.e. u_k is the marginal cost of changing b_k of one unit
- otherwise we can run additional iterations of the dual simplex algorithm from β to reach an optimal basis

- let $c'_j = c_j + \delta$ with $j \notin \beta$
- β remains a basis, and $x_\beta = A_\beta^{-1} b$ remains feasible
- $u^T = c_\beta^T A_\beta^{-1}$ remains feasible iff
  $\bar{c'}_j = (c_j + \delta) - c_\beta^T A_\beta^{-1} A_j = \bar{c}_j + \delta \geq 0$
- i.e. the basis β remains optimal if $\delta \geq -\bar{c}_j$ 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

- let $c'_j = c_j + \delta$ with $j \in \beta$, j being the l-th element of β
- β remains a basis, and $x_\beta = A_\beta^{-1} b$ remains feasible
- the primal feasibility condition becomes:

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

- i.e. the basis β remains optimal if $\delta g \leq -\bar{c}$ where g is the l-th row of $A_\beta^{-1} A$ (available in the simplex) and the optimal cost varies by δx_j
- otherwise we must run additional iterations of the primal simplex algorithm from β to reach an optimal basis

ADDING A NEW INEQUALITY CONSTRAINT

- adding a new constraint $a_{m+1}^T x \geq b_{m+1}$, by substitution we can assume that $a_{m+1,j} = 0$ for each $j \notin \beta$
- by adding a slack variable x_{n+1}, we get a new basis $\beta' = \beta \cup \{n+1\}$ with

$$A_{\beta'} = \begin{pmatrix} A_\beta & 0 \\ a_{m+1}^T & -1 \end{pmatrix} \quad A_{\beta'}^{-1} = \begin{pmatrix} A_\beta^{-1} & 0 \\ a_{m+1}^T A_\beta^{-1} & -1 \end{pmatrix}$$

  and the reduced costs remain unchanged:

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

- 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)

- let $a'_{ij} = a_{ij} + \delta$ with $j \notin \beta$
- β remains a basis, and $x_\beta = A_\beta^{-1} b$ remains feasible
- $u^T = c_\beta^T A_\beta^{-1}$ remains feasible iff
  $\bar{c}'_j = c_j - c_\beta^T A_\beta^{-1}(A_j + \delta e_i) \geq 0$
- i.e. the basis β remains optimal if $\delta u_i \leq \bar{c}_i$ 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

- it's complicated...

- parametric simplex method
- (progressive) column generation
- (progressive) constraint generation

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 corresponding change
- how to interpret a non-zero dual value ? simulate the corresponding change
- play also with the attributes VBasis, SAObjLow/Up, SALBLow/Up, SAUBLow/Up of Var and CBasis and SASRHSLow/Up of Constr

- a zero slack value/non-zero dual value: the value of an extra hour of availability
- a negative reduced cost: how to change the cost of a non-basic variable to make it an improving direction / the marginal cost of setting the non-basic variable to 1 (if feasible)
- note: the model is not in standard form... be careful with the signs !