LINEAR OPTIMIZATION

MASTER SPÉCIALISÉ OSE 2024 – MINES PARIS - PSL

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

OVERVIEW

introduction

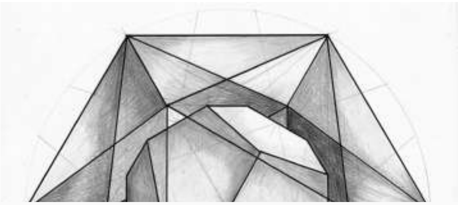
modeling LPs

geometry and algebra

the simplex methods

duality

sensitive analysis



INTRODUCTION

DECISION IS OPTIMIZATION

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

time : min travel duration, min lateness schedule
space : min travel distance, min wasted space layout
money : min cost design, max profit operation
goods : max production, min energy consumption
choice : max satisfaction
quantity : min potential energy (equilibrium)

DECISION FOR CLIMATE

optimize to help decarbonize

- **better processes** : minimize consumption, maximize utility
- **new technologies** : makes decision (problems) harder

Ex : PV, heat pumps, insulating materials for residential heat : how to choose, size, arrange, plan, manage them ? which criteria : heating needs, budget, efficiency, emissions, lifespan ?

hard decision making requires decision aid

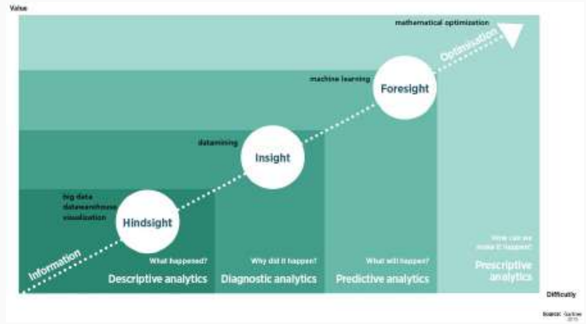
- **strategic** (design/long-term) or **operational** (control/short-term)
- **large-scale** (e.g. European electric system) or **small-scale** (e.g. water heater)
- imperfect knowledge : **complex** dynamics, **uncertain** forecasts
- **CPU intensive**

DECISION SUPPORT

Operations Research (1940)

IA/deep learning (2010)

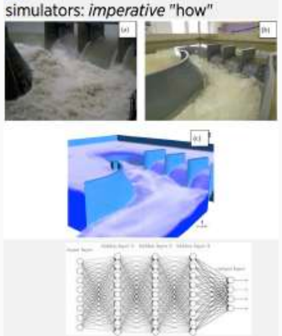
Business Analytics (2000)

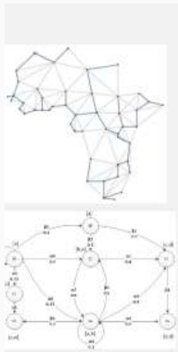


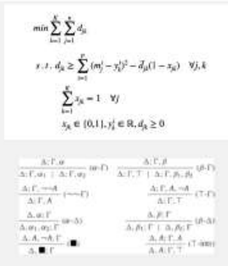
MODELS

Decision feasibility and value are observed through a model of the system/process

physical and virtual/numerical models
simulators: *imperative* "how"







conceptual models
formulation: *declarative* "what"

formalized from the dynamics knowledge or learned from solution samples

OPTIMIZATION MODELS

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 functions** over **variables**

$$\min \{\ f(x)\ :\ g_i(x) \leq 0\ \forall i \in \{1,\dots,m\},\ x \in \mathbb{R}^n\ \}$$

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.

ACCURACY & APPROXIMATION

decision making

problem solving

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

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

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

concrete problem $\longrightarrow$ abstract model $\xrightarrow{\text{solve}}$ model solution $\longrightarrow$ practical decision

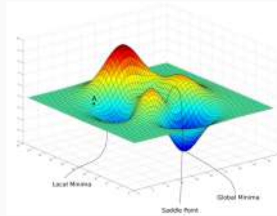
PROBLEM SOLVING : THEORY VS PRACTICE

solve a model $\neq$ solve a problem

- uncertain (forecast) and imprecise (truncated) data
- approximate (simplified) dynamics/constraints
- conceptual objective

solve a model $\min f(x) : g(x) \leq 0$ **?**

- feasible within a tolerance gap : $g(x) \leq \epsilon$
- optimal within a tolerance gap : $f(x) \leq \min f + \epsilon$
- optimal local vs global
- theoretic vs practical guarantees : high complexity, slow convergence, limited time



DECISION PRESCRIPTIVE TOOLS

- **mathematical optimization** : algorithms to compute a solution :

  $$x^* \in \arg\min \{\ f(x)\ :\ g(x) \leq 0,\ x \in \mathbb{R}^n\ \}, \quad f, g_i : \mathbb{R}^n \to \mathbb{R}$$

  The solution can be exact or approximate : $f(\tilde{x}) \approx \min f$ or $g(\tilde{x}) \leq \epsilon$
- **simulation** : evaluate a given decision x w.r.t. a model of the system/process, checking feasibility $g(x) \leq 0$ and computing value $f(x)$
- **simulation-optimization** or **black-box optimization** : iterative simulation of decisions $x_1, x_2, \ldots, x_N \in \mathbb{R}^n$ often searched heuristically

  $$\tilde{x} \in \arg\min \{\ f(x_k)\ :\ g(x_k) \leq 0, k \in \{1, \ldots, N\}\}$$
- **machine learning** : learn a numerical approximate model from samples of the system/process $(\tilde{f}, \tilde{g}) \approx (f, g)$ or, directly, of the best decisions $\mathcal{M}(f, g) \approx x^*$

SOLVING METHODS

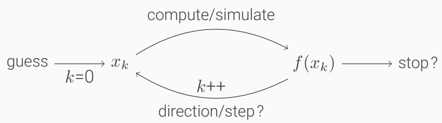
analytical methods come from a **provable theory**, e.g. :

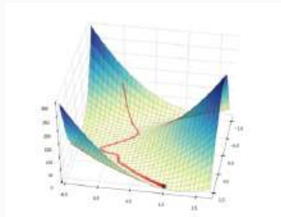
- $min\ x^2 - 4x + 3, x \in [0, 5]$ *(Fermat, derivative)*
- shortest path in a graph *(Dijkstra, Bellman)*

numerical methods **evaluate** $f(x_k)$ **iteratively** at trial points (x_k)

1st- or 2nd-order methods if driven by $f'(x_k)$ or $f''(x_k)$ *(simplex, gradient)*

derivative-free otherwise *(metaheuristics, branch-and-bound)*





DIFFERENT ALGORITHMS 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 OF MATHEMATICAL OPTIMIZATION

- **operational research** : operation, design and plan (routing, scheduling, packing, cutting, rostering, allocating) of physical/economical systems in logistics, energy, finance, etc.
- **prospective** : long-term vision on large systems
- 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

Definition : mathematical program

minimize $f(x)$ subject to $g(x) \geq 0$ $x \in \mathbb{R}^n$	maximize $f(x)$ subject to $g(x) \leq 0$ $x \in \mathbb{R}^n$

minimize or maximize under constraints $\leq, \geq$ or $=$, but never $>$ or $<$

- 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 : $g(x) \leq 0 \equiv -g(x) \geq 0$

solutions	$\mathbb{R}^n$
feasible solutions	$\{X \in \mathbb{R}^n : g(X) \geq 0\}$
optimal solutions	$\arg\min\{f(x) : g(x) \geq 0, x \in \mathbb{R}^n\}$

LINEAR PROGRAM

a mathematical program $\min\{f(x)|g(x) \geq 0, x \in \mathbb{R}^n\}$ with linear/affine functions f, g : $f(x) = c^T x, g(x) = Ax - b$ where $c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Definition : linear program (LP)

$$\begin{aligned} \min\ & c^T x \\ \text{s.t.}\ & Ax \geq b \\ & x \in \mathbb{R}^n \end{aligned}$$

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

LINEAR PROGRAM : AN EXAMPLE

$f(x) = c^T x, g(x) = Ax - b$ with $c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Example with $n = 3$ variables, $m = 2$ constraints

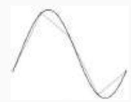
$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, c = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 5 & 3 & -2 \\ 1 & 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

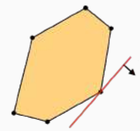
$$\begin{aligned} \text{min } & x_1 \\ \text{s.t. } & 5x_1 + 3x_2 - 2x_3 \geq 4 \\ & x_1 + x_2 + x_3 \geq -1 \\ & x_1, x_2, x_3 \in \mathbb{R} \end{aligned}$$

- (x_1, x_2, x_3) is feasible iff it satisfies **EVERY** constraints
- $x \mapsto 5x^2$, $(x, y) \mapsto 3xy$ are not linear (but quadratic)

HOW RELEVANT IS LP ? (COURSE MOTIVATION)

- **many applications :**
  format for practical decision problems,
  approximation for convex problems,
  basis for nonconvex/logic problems
  (with discrete/integer variables)
- **easy to solve :**
  polynomial-time algorithms,
  efficient practical algorithms,
  efficient off-the-shelf solvers,
  nice properties (geometry, 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 ? (may require additional variables)
3. objective : what is an optimal solution ? (may require add variables/constraints)
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)

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

Note on modelling

linearly equivalent formulations :

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

Linear program in standard form

Definition : LP in standard form

only **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

Proposition : reduction

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 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

(e.g. : *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

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

Ex : petroleum distillation – LP standard form

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

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

How to solve my LP ?

- LPs are smooth convex optimization problems and many algorithms apply
- LP solvers (`gurobi, cplex, glpk, mosek...`) are software or libraries with efficient implementations of dedicated algorithms (e.g. simplex, interior point)
- to solve an LP : call the solver with input A, b, c (no algorithm to implement)
- formats for input data (depending of the solver) :
  - text format (`lp`),
  - modelling langage (`gams, ampl`)
  - library (`pyomo,matlab`),
  - solver API (`gurobipy`)

Gurobi and the python API

GUROBI OPTIMIZATION

`gurobi + python = gurobipy`

- Gurobi is a commercial solver, freely available for students and academics
- a trial version of `gurobipy` limited to small-size models is available from Google Colab
- code examples as Jupyter Notebook can be edited and executed :
  `https://www.gurobi.com/jupyter_models/`

Linear algebra review and notation (1)

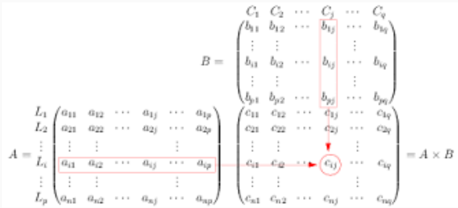
matrix $A \in \mathbb{R}^{m\times n}$ with entry a_{ij} in row $1 \leq i \leq m$, column $1 \leq j \leq n$

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

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 $(AB)^T = B^TA^T$



LINEAR ALGEBRA REVIEW AND NOTATION (2)

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

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

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 $rs_A = \{\lambda^T A, \lambda \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$

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

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 : 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 3 : 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)

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

EX 4 : NETWORK FLOW

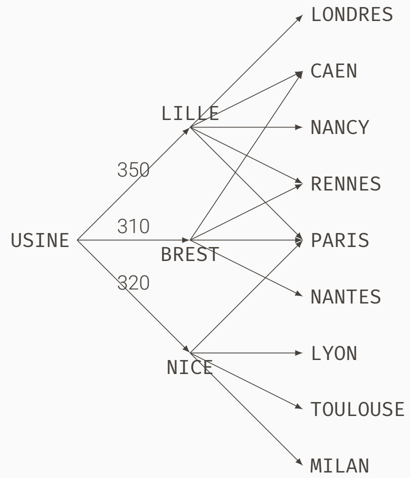
network flow

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 4 : GRAPH MODEL



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

EX 4 : LP MODEL

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

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

EX 5 : 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 5 : LINEARIZE THE ABSOLUTE VALUE

- every value $d \in \mathbb{R}$ can be decomposed as $d = u - v$ with $u \geq 0$ and $v \geq 0$
- in an infinite way, e.g. :

$$-4 = 4 - 8 = 1000 - 1004 = 2.7 - 6.7 = 0 - 4 = \cdots$$

- but only one decomposition minimizes $u + v$: $(u,v) = \begin{cases} (d,0) & \text{si } d \geq 0 \\ (0,-d) & \text{si } d \leq 0. \end{cases}$
- and the minimum value is precisely the absolute value : $|d| = \min_{(u,v)\geq 0}\{u + v : d = u - v\}$
- $\min_d \|d\|_1 = \min_d \sum_i |d_i|$, positive independent terms, thus $\min$ and $\sum$ can be exchanged :

$$\min_d \sum_i |d_i| = \sum_i \min_{d_i} |d_i| = \sum_i \min_{d_i,u_i,v_i} \{u_i + v_i : d_i = u_i - v_i\} = \min_{d,u,v} \sum_i \{u_i + v_i : d_i = u_i - v_i\}.$$

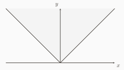
EX 5 : LP MODELS $\min \|x\|_1 = \min \sum_j |x_j|$

Two different ways to model $|x|$, $x \in \mathbb{R}$

1. variable splitting :

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

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

2. supporting plane model : 

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

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

EX 5 : LP model $\min \|x\|_\infty = \min \max_j |x_j|$

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

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

EX 5 : 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 5 : 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 5 : 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 5 : DATA FITTING – LAD REGRESSION (2)

variable splitting

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

dual model (see later)

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

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

EXERCISE : SIZING PV PANELS

sizing PV panels

how to equip two roofs with PV panels, respectively 4m and 6m long, to maximize the total power with an installation budget limited to 18ke, knowing that, one linear meter of PV installed is :

- 3ke on roof 1 for 150Wp peak power
- 2ke on roof 2 for 250Wp peak power

$$\begin{aligned} \text{max } & 150x_1 + 250x_2 \\ \text{s.t. } & x_1 \leq 4 \\ & x_2 \leq 6 \\ & 3x_1 + 2x_2 \leq 18 \\ & x_1, x_2 \geq 0 \end{aligned}$$

linear program (see ex : doors and windows)
x_1, x_2 : installed length (in meters)
constraints : maximal length, budget
objective : maximize production

Graphical representation (ex : doors & windows)

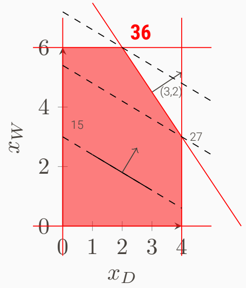
max $3x_D + 5x_W$

s.t. $x_D \le 4$

$x_W \le 6$

$3x_D + 2x_W \le 18$

$x_D, x_W \ge 0$



- solution space $\mathbb{R}^2$
- linear constraint $\equiv$ halfspace, ex : $\{x \in \mathbb{R}^2 \mid 3x_D + 2x_W \le 18\}$
- 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 : $x_W^* = 6$ and $3x_D^* + 2x_V^* = 18 \Rightarrow x_W^* = 6, x_D^* = 2, z^* = 36$

Graphical representation (ex : petroleum distillation)

min $20x_K + 15x_V$

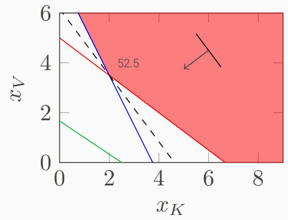
s.t. $3x_K + 4x_V \ge 20$

$4x_K + 2x_V \ge 15$

$2x_K + 3x_V \ge 5$

$0 \le x_K \le 9$

$0 \le x_V \le 6$



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

Graphical representation (ex : nuclear waste)

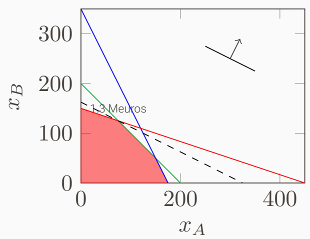
max $4x_A + 8x_B$

s.t. $x_A + 3x_B \le 450$

$2x_A + x_B \le 350$

$x_A + x_B \le 200$

$x_A, x_B \ge 0$

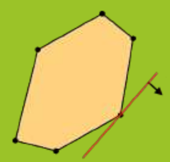


Geometry of linear programming

- the feasible region is a polyhedron = intersection of half-planes
- intuition : a linear function on a polyhedron reaches its min at a "*corner*"
- idea for solving an LP : **evaluate the corners progressively**

The primal simplex algorithm

1. find a first *corner* if exists
2. choose a feasible descent direction along an *edge*
3. if no direction, STOP : the corner is optimal
4. select the corner in this direction and goto step 2

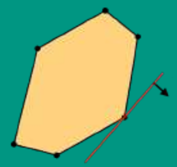


What is a corner?

Theorem : vertex = extreme point = basic feasible solution

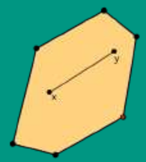
A nonempty polyhedron $\mathcal{P} = \{x \in \mathbb{R}^n | Ax \geq b\}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and a feasible solution $\hat{x} \in \mathcal{P}$, then these are equivalent : $\hat{x}$ is a

vertex



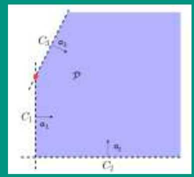
$\exists c \in \mathbb{R}^n, \forall x \in \mathcal{P} \setminus \{\hat{x}\}$, $c^\top \hat{x} < c^\top x$

extreme point



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

basic (feasible) solution



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

corners are associated to invertible submatrices of A and associated null slack variables : $a_i x + s_i = b_i$, $s_i = 0$; their number $\leq \binom{m}{n}$ is **finite** but large and not known a priori

Vertex, extreme point, and basic solution (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 :

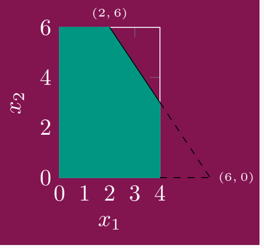
- $\hat{x}$ vertex $\Rightarrow$ xpoint : $\exists c, \forall x, y \in \mathcal{P} \setminus \{\hat{x}\}, c^T \hat{x} < c^T x$ and $c^T \hat{x} < c^T y$ then $c^T \hat{x} < \lambda c^T x + (1-\lambda) c^T y$, $\forall 0 \leq \lambda \leq 1$, then $\hat{x} \neq \lambda x + (1-\lambda) y$
- $\hat{x}$ not 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 $a_I^T x = b_I$.

Example of extreme points

ex : doors & windows

$$\begin{aligned} \max\ & 3x_1 + 5x_2 \\ \text{s.t. } & x_1 \leq 4 \\ & x_2 \leq 6 \\ & 3x_1 + 2x_2 \leq 18 \\ & x_1, x_2 \geq 0 \end{aligned}$$

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



- $n = 2$ variables (dimension), $m = 5$ constraints (edges)
- rows 2 and 3 are lin. independent, active at $(2, 6)$ feasible : vertex
- rows 5 and 3 are lin. independent, active at $(6, 0)$ unfeasible : basic solution
- rows 2 and 5 do not intersect (lin. dependent)

Existence of optima and extreme points

Theorem : existence of an extreme point [BT 2.6]

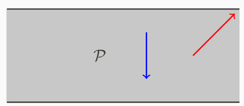
a nonempty $\mathcal{P} = \{x \in \mathbb{R}^n | Ax \geq b\}$ has at least one extreme point

$\iff$ 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]

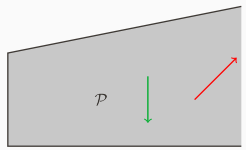
Minimizing a linear function over $\mathcal{P}$ having at least one extreme point, then : either optimal cost is $-\infty$, or an extreme point is optimal.



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 $(0, 0, 0, 0)$

HOW TO FIND A FIRST CORNER ?

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 $m < n$ 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. $x_j = 0, \forall j \notin \beta$.

Submatrix A_β is **invertible** : its columns form a **basis** of $\mathbb{R}^m$ with basic variables $(x_j)_{j \in \beta}$.

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 $\mathbb{R}^m$: $Ax = A_{|\beta} x_{|\beta} = b$

- the resulting basic solution x is **feasible** iff $x_j \geq 0\ \forall j$ (i.e. $x_{|\beta} = A_{|\beta}^{-1} b \geq 0$)

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 $A_j, j \in \beta \subset \{1, \dots, n\}$ 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.

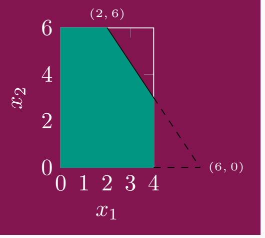
Example of basic solutions in standard form

check $\beta = (1, 2, 3)$, $\beta = (1, 3, 4)$, and $\beta = (1, 3, 5)$

$$\begin{aligned} \max\ & 3x_1 + 5x_2 \\ \text{s.t. } & x_1 + x_3 = 4 \\ & x_2 + x_4 = 6 \\ & 3x_1 + 2x_2 + x_5 = 18 \\ & x \geq 0 \end{aligned}$$

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

$n = 5$ variables,
$m = 3$ lin. indep. rows



- $\beta = (1, 2, 3)$: $A_\beta = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 3 & 2 & 0 \end{pmatrix}$ invertible; fix $x_4 = x_5 = 0$ solve $A_\beta \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 18 \end{pmatrix}$: $x_1 + x_3 = 4$, $x_2 = 6$, $3x_1 + 2x_2 = 18$: $x = (2, 6, 2, 0, 0) \geq 0$ **feasible**
- $\beta = (1, 3, 4)$: A_β invertible, fix $x_2 = x_5 = 0$ solve $x_1 + x_3 = 4$, $x_4 = 6$, $3x_1 = 18$: $x = (6, 0, -2, 6, 0) \not\geq 0$ **unfeasible**
- $\beta = (1, 3, 5)$ **not a base** : A_β is not invertible (cannot have $x_2 = x_4 = 0$ and $x_2 + x_4 = 6$)

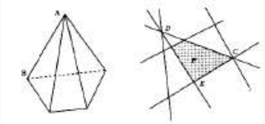
How to find a next corner ?

Definition : degeneracy and adjacency

Let $\mathcal{P} = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$ with $m < n$ linear independent rows $A \in \mathbb{R}^{m \times n}$; let $\beta \subset \{1, \dots, n\}$ defines a basis with associated basic solution x

- x is degenerate if $x_j = 0$ for some basic variable $j \in \beta$
- two bases β and β' are adjacent if they differ by 1 column

- a degenerate basic solution has different bases and more than n active constraints



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

- non-degenerate adjacent bases correspond to adjacent vertices along an edge of $\mathcal{P}$
- move to an adjacent vertex by swapping a basic and a non-basic column

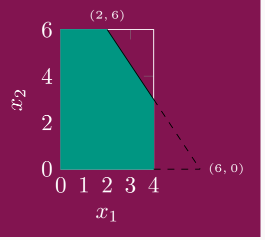
Example of degeneracy and adjacency

check point $(2, 6)$ and add constraint $3x_1 + x_2 \leq 12$

$$\begin{aligned} \max\ & 3x_1 + 5x_2 \\ \text{s.t. } & x_1 + x_3 = 4 \\ & x_2 + x_4 = 6 \\ & 3x_1 + 2x_2 + x_5 = 18 \\ & x \geq 0 \end{aligned}$$

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

$n = 5$ variables,
$m = 3$ lin. indep. rows



- $(2, 6)$ lies on edges $3x_1 + 2x_2 = 18$ (nonβ $x_5 = 0$), $x_2 = 6$ ($x_4 = 0$) but $x_1, x_2, x_3 > 0$
- bases $(1, 2, 3)$ and $(1, 3, 4)$ are **adjacent**, as they differ by 1 column : the associated vertices $(2, 6)$ and $(6, 0)$ both satisfy $x_5 = 0$ and lie on $3x_1 + 2x_2 = 18$
- when adding **redundant** constraint $3x_1 + x_2 \leq 12$, vertex $(2, 6)$ becomes degenerate :
- add column x_6 and row $(3, 1, 0, 0, 0, 1)$, then the basic solution $(2, 6, 2, 0, 0, 0)$ corresponds to 3 bases : $(1, 2, 3, 4)$, $(1, 2, 3, 5)$, $(1, 2, 3, 6)$

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

variables x_{jt} installed capacity (in MW) of type $j = 1, 2$ at year $t = 1, \dots, T$
objective minimize the installation costs
constraints each year, demand satisfaction + nuclear share
implied variables y_{jt} available capacity (in MW) $j = 1, 2$ for year t

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

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

EX : BASIC SOLUTION (CAPACITY PLANNING)

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

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

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

$n = 6T$ 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)

reformulate by dropping the redundant variables y_1 and y_2, find a basic solution, and give conditions of degeneracy (assume that $T - l_j + 1 \leq 1$ and constant $e_t = E \geq 0 \; \forall t$)

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

$$\begin{aligned}
& \sum_{s=1}^{t} x_{1s} + \sum_{s=1}^{t} x_{2s} - u_t = d_t - E, && \forall t = 1, \dots, T \\
& 8 \sum_{s=1}^{t} x_{2s} - 2 \sum_{s=1}^{t} x_{1s} + v_t = 2E, && \forall t = 1, \dots, T \\
& x_{jt} \geq 0, u_t \geq 0, v_t \geq 0 && \forall j = 1, 2, t = 1, \dots, T
\end{aligned}$$

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

- basic solution $(0, 0, E-d, 2E)$: feasible iff $E \geq d_t, \forall t$, degenerate iff $\exists t, E = 0$ or $E = d_t$
- basis (x_1, v) and suppose that $D_t = d_t - d_{t-1} > 0 \; \forall t$ with $d_0 = E$ then the basic solution $(D, 0, 0, 2d)$ is feasible nondegenerate (full coal scenario)
- question : under which condition can we improve the cost by installing nuclear at $t = 1$?

SUMMARY

- the feasible set of an LP is a polyhedron $\mathcal{P}$
- if $\mathcal{P}$ is nonempty and bounded, then LP has a basic optimal solution
- we can solve LP by enumerating all basic solutions : move along the edges of $\mathcal{P}$ by taking adjacent bases
- next lesson : the primal simplex algorithm improves the basic solution cost at each iteration (if non-degenerate)

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

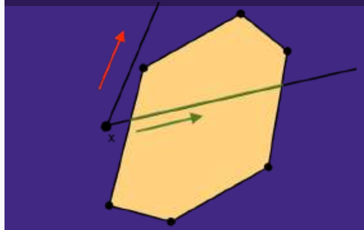
- $\min c^\top x$ over $\mathcal{P} = \{Ax = b, x \geq 0\}$, $A \in \mathbb{R}^{m\times n}$, $rk(A) = m$ reaches its optimum at a **basic feasible solution**
- a **basis** $\beta \subseteq \{1, \ldots, n\}$ is made of m linearly independent columns of A and the associated basic solution is : $x_\beta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

instead of visiting the basic solutions randomly, the **primal simplex method** selects the next **adjacent** basic solution such that it is **feasible** and of **better cost**.

Feasible descent direction

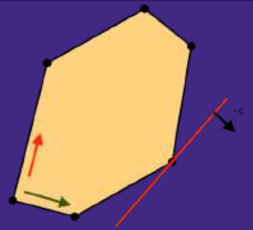
minimize $c^\top x$ over $\mathcal{P} = \{x \in \mathbb{R}^n | Ax = b, x \geq 0\}$, and some point $x \in \mathbb{R}^n$

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



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

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



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

if d is a feasible descent direction, then there is a feasible solution $x' = x + \theta d$ strictly improving upon x since $c^\top x' = c^\top x + \theta . c^\top d < c^\top x$

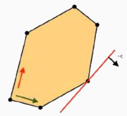
Basic descent direction

min $\{c^\top x : Ax = b, x \geq 0\}$, x 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 $Ad = 0$

is a feasible direction (if x nondegenerate) and $d_\beta = -A_\beta^{-1} A_{j'}$:



$$\begin{cases} Ad = 0 \;\Rightarrow\; A(x + \theta d) = Ax = b \\ x_j > 0 \,\forall j \in \beta \Rightarrow \exists \theta > 0 : x_\beta + \theta d_\beta \geq 0 \end{cases}$$

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

$\bar{c}_{j'} = c^\top d = c_{j'} - c_\beta^\top A_\beta^{-1} A_{j'}$

- $\bar{c}_{j'} = c^\top d = c^\top (x + d) - c^\top x$ is the cost deviation between solutions x and $x + d$
- d is a **descent direction** iff $\bar{c}_{j'} < 0$
- the reduced cost of a basic variable $j \in \beta$ is always 0 : $\bar{c}_j = c_j - c_\beta^\top A_\beta^{-1} A_j = c_j - c_\beta^\top e_j = 0$

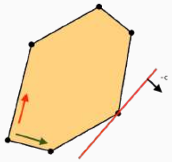
Step length θ

β basis of x feasible nondegenerate, d feasible direction to $j' \notin \beta$ s.t. $c^\top d = \bar{c}_{j'} < 0$
look for the largest value $\theta > 0$ such that $x' = x + \theta d$ remains feasible, i.e. $x' \geq 0$:

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 : constraint $x_{j''} \geq 0$ becomes active



Step length θ (proof)

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

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 $\beta' = \beta \cup \{j'\} \setminus \{j''\}$:

Proof :

- $d \geq 0 \;\Rightarrow\; x + \theta d \in \mathcal{P} \,\forall \theta > 0$ and $c^\top (x + \theta d) \searrow$ when $\theta \nearrow$
- x nondegenerate $\Rightarrow x_{j''} > 0 \Rightarrow \theta > 0$
- $x' \in \mathcal{P} \iff x_j + \theta d_j \geq 0 \,\forall j \iff x_j + \theta d_j \geq 0 \,\forall j \in \beta : d_j < 0$ (since $Ax' = Ax = b$)
- $A_\beta^{-1} A_j = e_j, \forall j \in \beta \setminus \{j''\}$, 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 descent direction

check basis $(1, 2)$ **and find basic descents**

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

- $m = 2, n = 4, rk(A) = 2, \beta = (1, 2)$ forms a basis
- $x = (1, 1, 0, 0)$ feasible nondegenerate : $x_j > 0\ \forall j \in \beta$
- basic direction $j = 3$: $d_3 = 1, d_4 = 0, Ad = \binom{d_1 + d_2 + 1}{2d_1 + 3} = 0 \Rightarrow d_\beta = \binom{d_1}{d_2} = \binom{-3/2}{1/2}$
- is a descent direction : $\bar{c} = c^\top d = 2(-3/2) + (1/2) + 1 = -3/2 < 0$
- step length : $x' = x + \theta d \geq 0 \Rightarrow x'_1 = 1 - (3/2)\theta \geq 0 \Rightarrow \theta \leq 2/3$
- $x' = (0, 4/3, 2/3, 0)$ basic feasible $\beta' = (2, 3)$, $c^\top x' = c^\top x + \theta \bar{c}_3 = c^\top x - 1$

When stops the algorithm ?

Theorem : optimality condition [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 $c_{\neg\beta} \geq 0$:

$A_\beta d_\beta + A_{\neg\beta} y_{\neg\beta} = Ad = Ay - Ax = b - b = 0 \Rightarrow d_\beta = -A_\beta^{-1} A_{\neg\beta} y_{\neg\beta} \Rightarrow$

$c^\top y - c^\top x = c_\beta^\top d_\beta + c_{\neg\beta}^\top y_{\neg\beta} = (c_{\neg\beta}^\top - c_\beta^\top A_\beta^{-1} A_{\neg\beta}) y_{\neg\beta} = \bar{c}_{\neg\beta} y_{\neg\beta} \geq 0$

($\Leftarrow$) 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.)

check basis $(2, 3)$

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

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

The primal 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 primal 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 bounded by C_n^m

in case of degeneracy : apply techniques (ex : fixed order subscripts) to avoid cycling on the same vertex

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

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)

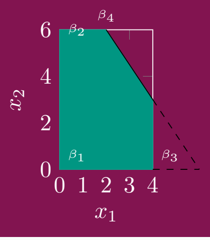
(see [Bertsimas-Tsitsiklis] Section 3.3 for details)

EX : SIMPLEX ALGORITHM

start at $\beta_1 = (3,4,5)$

$$\begin{aligned} &\min_{x\geq 0} -3x_1 - 5x_2 \\ &\text{s.t. } x_1 + x_3 = 4 \\ &\quad x_2 + x_4 = 6 \\ &\quad 3x_1 + 2x_2 + x_5 = 18. \end{aligned}$$

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



- $x = (0,0,4,6,18)$ is feasible nondegenerate
- let $d_1 = 0, d_2 = 1$ and $Ad = 0$: $d = (0,1,0,-1,-2)$, $\bar{c} = c^\top d = -5 < 0 \Rightarrow$ descent
- find the largest $\theta > 0$ s.t. $x + \theta d = (0,\theta,4,6-\theta,18-2\theta) \geq 0$, i.e. $\theta = \min(6/1, 18/2) = 6$: new basis $\beta = (2,3,5)$ and solution $x + \theta d = (0,6,4,0,6)$
- next : $d = (1,0,-1,0,-3)$, $\bar{c} = -3$ descent $x + \theta d = (\theta, 6, 4-\theta, 0, 6-3\theta) = (2,6,2,0,0)$
- next : $d = (2/3,-1,-2/3,1,0)$, $\bar{c} = 3$ optimum $x = (2,6,2,0,0)$, $cx = -36$

READING :

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

A constrained nonlinear convex problem

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

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

lagrangian relaxation (general optimization)

$$P : z = min\ c(x) \quad \text{s.t. } g(x) = 0, \quad x \geq 0$$

$$P_u : z_u = min\ c(x) + u^\top g(x) \quad \text{s.t. } x \geq 0$$

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

dual problem

find the tightest (greater) lower bound z_u of z :

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

- weak duality $d \leq z$ always holds (by definition)
- strong duality $d = z$ may hold if exists x optimal for some P_u and feasible for P

Specific properties of LP duality

- if P is an LP then D is also an LP and the dual of D is the primal P
- constraints/variables of P correspond to variables/constraints of D
- strong duality always holds for LP
- if P is unbounded then D is unfeasible, and conversely
- primal simplex : computes solutions in the dual space, stops when dual feasible
- dual simplex : computes solutions in the primal space, stops when primal feasible
- sensitive analysis : how to recover feasibility in the primal or in the dual space

The dual linear program

Theorem : the dual of an LP is an LP

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

$$(D) : max\ u^\top b \quad \text{s.t. } u^\top A \leq c^\top$$

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 of an LP ?

primal/dual correspondence

min	max
cost vector c	RHS vector b
matrix A	matrix $A^\top$
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 $u^\top A_j = c_j$
nonnegative variable $x_j \geq 0$	constraint $u^\top A_j \leq c_j$

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

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

equivalent forms of (P) give equivalent forms of (D)

EX 7 : STEEL FACTORY

steel factory

A factory produces 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

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

EX : DUAL MODEL (STEEL FACTORY)

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

WEAK DUALITY

Theorem [BT 4.3]

- if x is feasible for P (min) and u is feasible for D (max) then : $u^\top b \leq cx$
- 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 $u^\top b = cx$ then x is optimal for P and u is optimal for D

Proof :

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

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 : $u^\top b = c^\top x$

Proof :

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

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

COMPLEMENTARY SLACKNESS

Theorem [BT 4.5]

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

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

Proof :

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

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

EXERCISE : OPTIMALITY WITHOUT SIMPLEX

show that $\beta = (1, 3)$ is an optimal basis

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

EXERCISE : OPTIMALITY WITHOUT SIMPLEX

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

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

- $\beta = \{1, 3\} \Rightarrow x_2 = 0, x_1 = 3/3 = 1, x_3 = (8 - 5)/3 = 1$
- $x = (1, 0, 1), x \geq 0 \Rightarrow$ 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
- basic dual solution $u = c_\beta^\top A_\beta^{-1}$ feasible $\iff$ reduced cost $\bar{c} = c^\top - u^\top A \geq 0$

OPTIMALITY CONDITIONS

Theorem : Karush-Kuhn-Tucker optimality conditions in LP

x is optimal for $P = min\{c^\top x | Ax = b, x \geq 0\}$ iff 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 : $u^\top A \leq c^\top$
4. complementary slackness : $x_j > 0 \Rightarrow u^\top A_j = c_j$

- a basic feasible solution x always satisfy 1,2 and 4 with $u^\top = c_\beta^\top A_\beta^{-1}$ $(x_j > 0 \Rightarrow j \in \beta$ and $\bar{c}_j = c_j^\top - u^\top A_j = 0)$.
- 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

ALT ALGORITHM : DUAL SIMPLEX

$(P) : min\{c^\top x : Ax = b, x \geq 0\}$ and $(D) : max\{u^\top b : u^\top A \leq c^\top\}$

- a basis β determines basic solutions for P and D : $x_\beta = A_\beta^{-1} b$ and $u^\top = c_\beta^\top A_\beta^{-1}$
- satisfying complementary slackness : $x_j > 0 \Rightarrow j \in \beta \Rightarrow \bar{c}_j = c_j - u^\top A_j = 0$
- primal simplex algorithm maintains primal feasibility $(x_\beta \geq 0)$ and tries to achieve dual feasibility $(\bar{c}^\top = c^\top - u^\top A \geq 0)$

dual simplex method

- equivalent to solving (D) with the primal simplex
- maintains dual feasibility $(\bar{c} \geq 0)$ and tries to achieve primal feasibility $(x_\beta \geq 0)$

Usage : after modifying b or adding a new constraint to (P), the dual basic solution $u^\top = c_\beta^\top A_\beta^{-1}$ remains feasible : start the dual simplex iterations from this basis

ALT ALGORITHMS : INTERIOR POINT

$(P) : min\{c^\top x : Ax = b, x \geq 0\}$ and $(D) : max\{u^\top b : u^\top A \leq c^\top\}$

KKT : $Ax = b, x \geq 0, v^\top = c^\top - u^\top A \geq 0$, and complementary slackness : $x^\top v = 0$

interior point methods

- iterates on primal feasible x and dual feasible u, v with $x^\top v = n/t$ for increasing t
- KKT with disturbed complementary slackness : $Ax = b, x \geq 0, v \geq 0, x^\top v = n/t$
- = KKT for the centered problem $P^t : min\{tc^\top x + \phi(x) : Ax = b\}$ with barrier function $\phi(x) = -\sum_j log(x_j)$, a smooth approximation of the indicator function $x \geq 0$
- given an interior point $x > 0 : Ax = b$, then P^t can be efficiently solved with Newton method and returns an other interior point $x^t > 0$
- barrier method : at each iteration i, increase $t = t(i) = \mu t(i-1)$, solve P_t with Newton's method starting from $x^{t(i-1)}$ to get (x^t, u^t) and define $v_j^t = 1/tx_j^t$ then (x^t, u^t, v^t) satisfies the disturbed KKT.
- primal-dual interior-point method : update also u, v within inner-loop (Newton) iterations

FARKA'S LEMMA AND UNFEASIBILITY

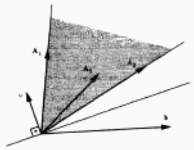
theorem

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

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

Proof :

$(1 \Rightarrow \neg 2)$ if $x \in \mathcal{P}$ and $u^\top A \geq 0$ then $u^\top b = u^\top Ax \geq 0$
$(\neg 1 \Rightarrow 2)$ if $P : max\{0 | Ax = b, x \geq 0\}$ is unfeasible then $D : min\ \{u^\top b | u^\top A \geq 0\}$ 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 $\{x \in \mathbb{R}^m | u^\top x = 0\}$ exists

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

SENSITIVE ANALYSIS

GOAL OF SENSITIVE ANALYSIS

Most LP models of real-world decision problems rely on forecast/inaccurate data and incomplete knowledge

- a model is more reliable if its solutions are less **sensitive to changes in data**
- a model is more robust if its solutions are less **sensitive to addition of variables/constraints**

evaluate the sensitivity of the optimal solution of an LP
to one structural change in the LP
without having to solve the LP again for every possible value change.

THE CORE IDEA

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

$$x = (x_\beta, x_{\neg\beta}) = (A_\beta^{-1}b, 0) \qquad \text{and} \qquad u^\top = c_\beta^\top A_\beta^{-1} \quad \text{satisfying :}$$

$x_\beta \geq 0$	primal feasibility
$\bar{c}^\top = c^\top - u^\top A \geq 0$	dual feasibility

(primal feas. $Ax = b$ and comp. slackness $\bar{c}_\beta = 0$ satisfied by construction of x and u)

- 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})
- equivalent to suppose $n+1$ is non-basic and $x_{n+1} = 0$
- 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 $u^\top = c_\beta^\top A_\beta^{-1}$ is dual feasible, i.e. $n+1$ is not a descent direction :

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

- then, the optimal value $c_\beta^\top x_\beta$ does not change
- otherwise, if $n+1$ is a descent direction : run additional iterations of the **primal** simplex algorithm starting from the primal feasible basis β

EXAMPLE : ADDING A VARIABLE

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

add column $A_4 = (1,1)$: for which cost $c_4 = \delta$ the basis remains optimal ?

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

- $\beta = (1,3)$ remains a basis, $x^\top = (1,0,1,0)$ primal feasible
- $u^\top = (2,1)$ remains feasible iff the dual constraint is satisfied $u_1 + u_2 = 3 \leq \delta$
- the optimal solution x and value 19 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 a given constraint $k = 1, \ldots, m$
- β remains a basis and $u^\top = c_\beta^\top A_\beta^{-1}$ remains dual feasible $(c^\top - u^\top A \geq 0)$
- β remains optimal if the new primal solution $x' = A_\beta^{-1} b'$ is still feasible, i.e :

$$x'_\beta = x_\beta + \delta A_\beta^{-1} e_k \geq 0$$

- then, the optimal cost varies by $\delta u_k = u^\top b' - u^\top b$
- the dual value u_k is the **marginal cost** (or shadow price) per unit increase of b_k
- otherwise, if x' not feasible : run additional iterations of the **dual** simplex algorithm starting from the dual feasible basis β

EXAMPLE : CHANGING b

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

change RHS in the first constraint $b'_1 = b_1 + \delta$

$P : min\ 13x_1 + 10x_2 + 6x_3$	$D : max\ (8+\delta)u_1 + 3u_2$
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, $u^\top$ remains dual feasible
- $x' = (1, 0, 1 + \frac{\delta}{3})$ is feasible iff $1 + \frac{\delta}{3} \geq 0$
- x' remains optimal if $\delta \geq -3$ and the optimum value increases by $u^\top b' - u^\top b = u_1 \delta$
- increasing b_1 by $\delta = 1$ unit induces a marginal (additional) cost $u_1 = 2$

CHANGING THE COST OF A NON-BASIC VARIABLE

- let $c'_j = c_j + \delta$ for some non-basic variable $j \notin \beta$
- β remains a basis and $x_\beta = A_\beta^{-1} b \geq 0$ remains primal feasible
- β remains optimal if the basic dual solution $u^\top = c_\beta^\top A_\beta^{-1}$ remains feasible, i.e. j is still not a descent direction :

$$\bar{c}'_j = (c_j + \delta) - u^\top A_j = \bar{c}_j + \delta \geq 0$$

- then, the optimal value $c_\beta^\top x_\beta$ does not change
- the **reduced cost** $\bar{c}_j$ is the cost reduction value from which j becomes profitable
- otherwise, j is a descent direction : run additional iterations of the **primal** simplex algorithm starting from the primal feasible basis β

EXAMPLE : CHANGING c (NON-BASIC)

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

change the non-basic cost c_2 by δ

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

- β remains a basis, x and u are still basic and x remains feasible
- u remains feasible iff $\bar{c}_2 + \delta = (10+\delta) - (u_1 + u_2) \geq 0$, i.e. $\delta \geq -7$
- optimal solutions and values do not change while $\delta \geq -7 = -\bar{c}_2$
- x_2 becomes profitable when its cost is below $10 - \bar{c}_2 = 3$

CHANGING THE COST OF A BASIC VARIABLE

- let $c'_j = c_j + \delta$ for some basic variable $j \in \beta$
- β remains a basis and $x_\beta = A_\beta^{-1} b \geq 0$ remains primal feasible
- β remains optimal iff the new dual basic solution $u'^\top = c'^\top_\beta A_\beta^{-1}$ is feasible :

$$\bar{c}'^\top_{\neg\beta} = \bar{c}^\top_{\neg\beta} - \delta e_j^\top A_\beta^{-1} A_{\neg\beta} \geq 0$$

- then, the optimal cost varies by $\delta x_j = (c'^\top - c^\top)x$
- 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 $x^\top = (1,0,1)$, $u^\top = (2,1)$ primal-dual feasible, $opt = 19$

change the (basic) cost c_1 by δ

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

- β remains a basis, $x^\top$ remains primal feasible
- new dual solution u' solves $5u'_1 + 3u'_2 = 13 + \delta$, $3u'_1 = 6$: $u' = (2, 1 + \frac{\delta}{3})$
- u' is feasible iff $u'_1 + u'_2 = 2 + 1 + \frac{\delta}{3} \leq 10$, i.e. if $\delta \leq 21$
- and the optimum value increases by $x_1\delta = \delta$
- x_1 is less profitable than x_2 if c_1 is above $13 + 21 = 31$

ADDING A NEW INEQUALITY CONSTRAINT

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

$$A_{\beta'} = \begin{pmatrix} A_\beta & 0 \\ a_{m+1}^\top & -1 \end{pmatrix} \quad A_{\beta'}^{-1} = \begin{pmatrix} A_\beta^{-1} & 0 \\ a_{m+1}^\top A_\beta^{-1} & -1 \end{pmatrix}$$

- $u^\top = (c_\beta^\top,\ 0) A_{\beta'}^{-1} = (c_\beta^\top A_\beta^{-1},\ 0)$ is feasible as the reduced costs are unchanged :

$$\bar{c'}^\top = (c^\top,\ 0) - (c_\beta^\top,\ 0) A_{\beta'}^{-1} A = (\bar{c}^\top,\ 0)$$

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

EXAMPLE : ADDING A CONSTRAINT

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

adding constraint $x_1 + x_3 \leq 1$ and slack variable x_4

P	D
$P : min\ 13x_1 + 10x_2 + 6x_3$	$D : max\ 8u_1 + 3u_2 + u_3$
s.t. $5x_1 + x_2 + 3x_3 = 8$	s.t. $5u_1 + 3u_2 + u_3 \leq 13$
$3x_1 + x_2 = 3$	$u_1 + u_2 \leq 10$
$x_1 + x_3 + x_4 = 1$	$3u_1 + u_3 \leq 6$
$x_1, x_2, x_3, x_4 \geq 0$	$u_3 \leq 0$

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

CHANGING A NON-BASIC COLUMN

- let $a'_{ij} = a_{ij} + \delta$ for some constraint i and non-basic variable $j \notin \beta$
- β remains a basis and $x_\beta = A_\beta^{-1} b \geq 0$ is primal feasible
- β remains optimal if $u^\top = c_\beta^\top A_\beta^{-1}$ remains feasible :

$$\begin{aligned} \bar{c'}_j &= c_j - c_\beta^\top A_\beta^{-1}(A_j + \delta e_i) \\ &= \bar{c}_j - \delta u_i \geq 0 \end{aligned}$$

- then, the optimal value $c_\beta^\top x_\beta$ does not change
- otherwise, j becomes a descent direction : run additional iterations of the **primal** simplex algorithm starting from the primal feasible basis β

EXAMPLE : CHANGING A_j (NON-BASIC)

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

changing coefficient in the non-basic column A_2

P	D
$P : min\ 13x_1 + 10x_2 + 6x_3$	$D : max\ 8u_1 + 3u_2$
s.t. $5x_1 + (1+\delta)x_2 + 3x_3 = 8$	s.t. $5u_1 + 3u_2 \leq 13$
$3x_1 + x_2 = 3$	$(1+\delta)u_1 + u_2 \leq 10$
$x_1, x_2, x_3 \geq 0$	$3u_1 \leq 6$

- β remains a basis, $x^\top$ remains primal feasible
- $u^\top$ 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

take advantage of warm-start (feasible primal/dual solutions) in iterative solutions :

- **constraint generation** : generate constraints progressively when they are violated
- **column generation** : generate nonbasic variables progressively when they are profitable
- **branch-and-bound** : update the variable bounds dynamically
- **parametric simplex method** for solving LP with a variable parameter

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