The many facets of mathematical

Optimization

François Glineur

Fourth Benelux Mathematical Olympiad, April 22 2012

A plan ?

What is optimization ?

Is optimization useful ?

Is optimization easy ?

Is optimization always hard ?

Can optimization help with your diet ?

Can we trust optimization ?

Can optimization beat computer chip makers ?

1.
What is Optimization ?

Modelling → Decision

1. Model a real-word situation mathematically
2. Choose the **best** feasible decision

Decision	→	vector of variables
Best	→	objective fonction
Feasible	→	constraints

Variables+objective+constraints → optimisation

A mathematical **tool** for decision making

Mathematical formulation

Variables : $x \in \mathbb{R}^n$

Objective function : $f : \mathbb{R}^n \mapsto \mathbb{R} : x \mapsto f(x)$

Constraints : $D \subseteq \mathbb{R}^n$

$$\min_{x \in \mathbb{R}^n} f(x) \text{ such that } x \in D$$

A few simple examples

Find rectangular box of a given volume V with the minimum surface

$$min_{(a,b,c)\in R^3}\ 2ab + 2bc + 2ca\ \ s.t.\ abc = V$$

Find the diameter of a shape S in the plane

$$max_{x\in R^2, y\in R^2}\ \|x - y\|\ \ s.t.\ x \in S, y \in S$$

Find the fastest path to your destination (in your GPS device) – *binary variables*

Terminology

Applied Mathematics

$\supset$ Operations Research

= Quantitative methods for decision making

= *Management Science*

= *Industrial Engineering*

$\supset$ Mathematical Programming

= Optimization = Optimisation

= Mathematical Optimization

= ~~Optimalisation~~

Optimization : a vast domain

- Deterministic or stochastic problem
- Certain or uncertain data
- Single or multiple objective
- Explicit or black-box characteristics
- Continuous or discrete
- Linear or non-linear
- Constrained or unconstrained
- etc.

Presentation will focus on a small part

2.
Is Optimization useful ?

A wide variety of applications

- Planning and scheduling

  Production, timetabling, crew scheduling

- Design

  Structural design, network design

- Economics and finance

  Portfolio optimization, Nash equilibria

- Location and transport

  Facility location, routing, tours

INFORMS Edelman award 2008

Dutch railways

In 2006 : Schedule optimisation for passengers, trains and personnel

After one year :

- Delays reduced by 17%
- Passengers increased by 10%
- Profit increased by € 40M

Engineering applications

- Airbus : structural optimization of wings (min weight for given strength)
- Radiotherapy : dose optimization (modulate a moving beam)
- Spectrum optimization in telecommunications (your next modem)
- Astrophysics : optimize diffraction masks (planet detection)
- And much much more …

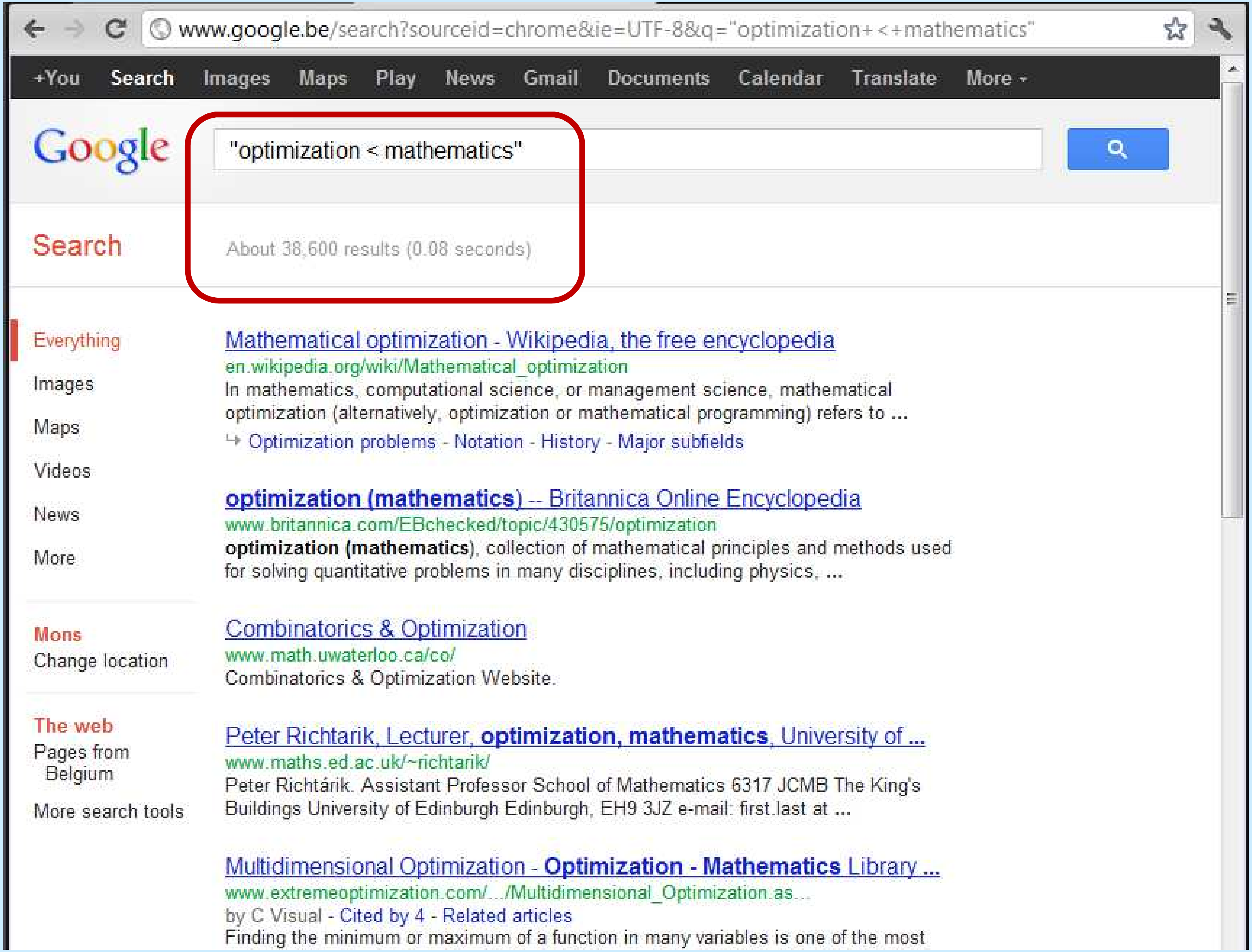
Optimization is important

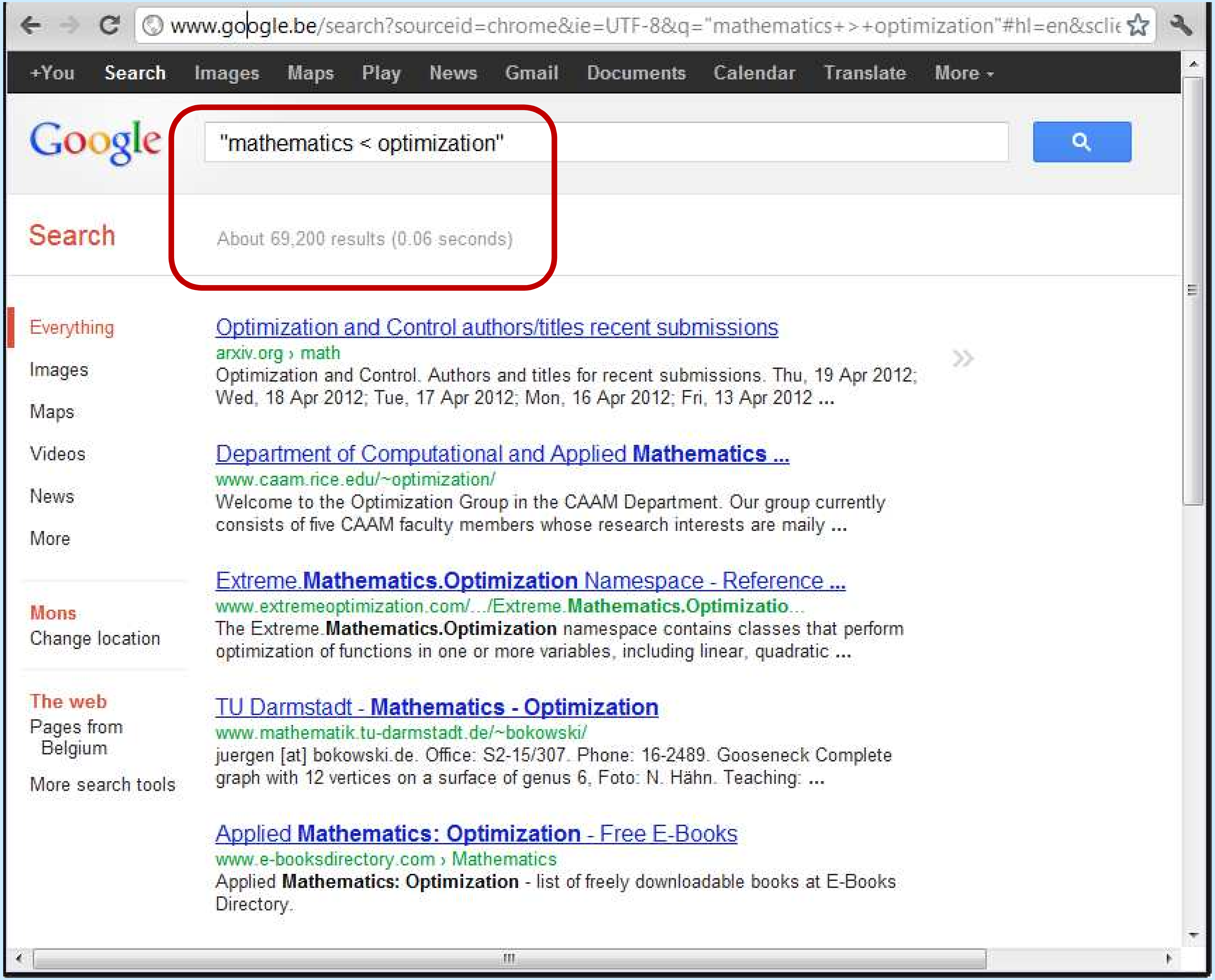
Mathematics

Versus

Optimization

Let us ask the modern *Oracle* …


www.google.be/search?sourceid=chrome&ie=UTF-8&q="optimization+<+mathematics"
+You Search Images Maps Play News Gmail Documents Calendar Translate More
Google
"optimization < mathematics"
Search
About 38,600 results (0.08 seconds)
Everything
Images
Maps
Videos
News
More
Mons
Change location
The web
Pages from Belgium
More search tools
Mathematical optimization - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Mathematical_optimization
In mathematics, computational science, or management science, mathematical optimization (alternatively, optimization or mathematical programming) refers to ...
Optimization problems - Notation - History - Major subfields
optimization (mathematics) -- Britannica Online Encyclopedia
www.britannica.com/EBchecked/topic/430575/optimization
optimization (mathematics), collection of mathematical principles and methods used for solving quantitative problems in many disciplines, including physics, ...
Combinatorics & Optimization
www.math.uwaterloo.ca/co/
Combinatorics & Optimization Website.
Peter Richtarik, Lecturer, optimization, mathematics, University of ...
www.maths.ed.ac.uk/~richtarik/
Peter Richtárik. Assistant Professor School of Mathematics 6317 JCMB The King's Buildings University of Edinburgh Edinburgh, EH9 3JZ e-mail: first.last at ...
Multidimensional Optimization - Optimization - Mathematics Library ...
www.extremeoptimization.com/.../Multidimensional_Optimization.as...
by C Visual - Cited by 4 - Related articles
Finding the minimum or maximum of a function in many variables is one of the most



www.google.be/search?sourceid=chrome&ie=UTF-8&q="mathematics+>+optimization"#hl=en&sclie
+You
Search
Images
Maps
Play
News
Gmail
Documents
Calendar
Translate
More
Google
"mathematics < optimization"
Search
About 69,200 results (0.06 seconds)
Everything
Images
Maps
Videos
News
More
Mons
Change location
The web
Pages from Belgium
More search tools
Optimization and Control authors/titles recent submissions
arxiv.org › math
Optimization and Control. Authors and titles for recent submissions. Thu, 19 Apr 2012; Wed, 18 Apr 2012; Tue, 17 Apr 2012; Mon, 16 Apr 2012; Fri, 13 Apr 2012 ...
Department of Computational and Applied Mathematics ...
www.caam.rice.edu/~optimization/
Welcome to the Optimization Group in the CAAM Department. Our group currently consists of five CAAM faculty members whose research interests are maily ...
Extreme.Mathematics.Optimization Namespace - Reference ...
www.extremeoptimization.com/.../Extreme.Mathematics.Optimizatio...
The Extreme.Mathematics.Optimization namespace contains classes that perform optimization of functions in one or more variables, including linear, quadratic ...
TU Darmstadt - Mathematics - Optimization
www.mathematik.tu-darmstadt.de/~bokowski/
juergen [at] bokowski.de. Office: S2-15/307. Phone: 16-2489. Gooseneck Complete graph with 12 vertices on a surface of genus 6, Foto: N. Hähn. Teaching: ...
Applied Mathematics: Optimization - Free E-Books
www.e-booksdirectory.com › Mathematics
Applied Mathematics: Optimization - list of freely downloadable books at E-Books Directory.

Googlefight : verdict

Mathematics < Optimization : 69 200

versus

Optimization < Mathematics : 38 600

Opt-Art, by R. Bosch

A single loop !

3.
Is Optimization easy ?

Algorithmic black-box optimization

$$f^* = \min_{x \in \mathbb{R}^n} f(x) \text{ such that } x \in D$$

- Rules of the game: the **black box**
  - Give me points in $D : x^{(1)}, x^{(2)}, x^{(3)}, \ldots$
  - I will return the values of $f : f(x^{(1)}), f(x^{(2)}), \ldots$
  - Nothing else is known about the function
- Goal: find a good solution quickly
  - As few evaluations as possible
  - With prescribed accuracy $\varepsilon : f\left(x^{(n)}\right) - f^* < \varepsilon$

Looking at the worst case

- Typical theorem:

  For a given class of functions, any method must take at least some number of evaluations on some function

- Must hold for **any** method

- Main tool for proof: *resisting oracle*

Some games cannot be won

- Typical theorem:

  For a given class of functions, any method must take at least some number of evaluations on some function

- For some functions, game is always lost
  - discontinuous functions
  - functions with very quick variations
  - tiny and complicated domain D

Optimization is hard: the theorem

Among the class of functions with n variables defined on the unit box $\{x \mid 0 \leq x_i \leq 1 \ \forall i\}$ whose variation is not too large in the sense

$$|f(x) - f(y)| < L \, \|x - y\|_\infty$$

for any method there exists a function such that accuracy ε is not reached before

$$\left\lfloor \frac{L}{2\varepsilon} \right\rfloor^n \text{ evaluations}$$

Far-reaching consequences

$$\left\lfloor \frac{L}{2\varepsilon} \right\rfloor^n \text{ evaluations}$$

Consider an easy problem with L=2 and ε = 1%

Some 1-variable problem *must* take 100 iter.

Some 10-variable problem *must* take 10^{20} iter.

Life is too short for such problems!

- Uniform grid search is worst-case optimal
- Result does not really improve for smoother f

3.

Is Optimization always hard ?

Optimization is not always hard

No hope of guaranteeing resolution of optimization problems with 10 variables

Two ways to react:

Ignore theorem	Nullify theorem
Try you luck!	Force your luck! new problem class

A new class: linear optimization

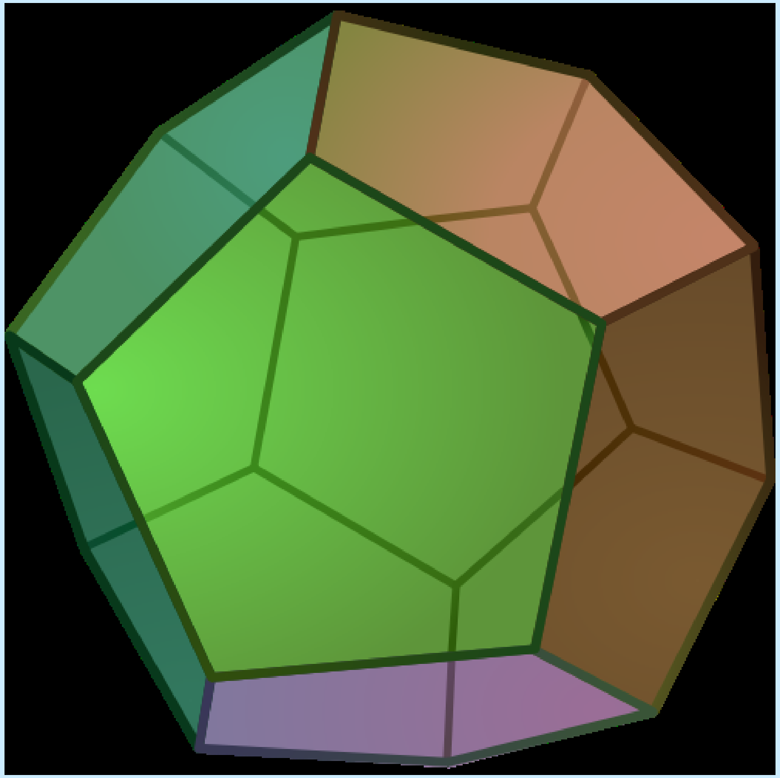
We voluntarily restrict ourselves to a problem class involving only linear functions

$$\min_{x \in \mathbb{R}^n} f(x) \ \text{ such that } \ x \in D$$

Linear objective function f

Linear constraints, which gives polyhedron D

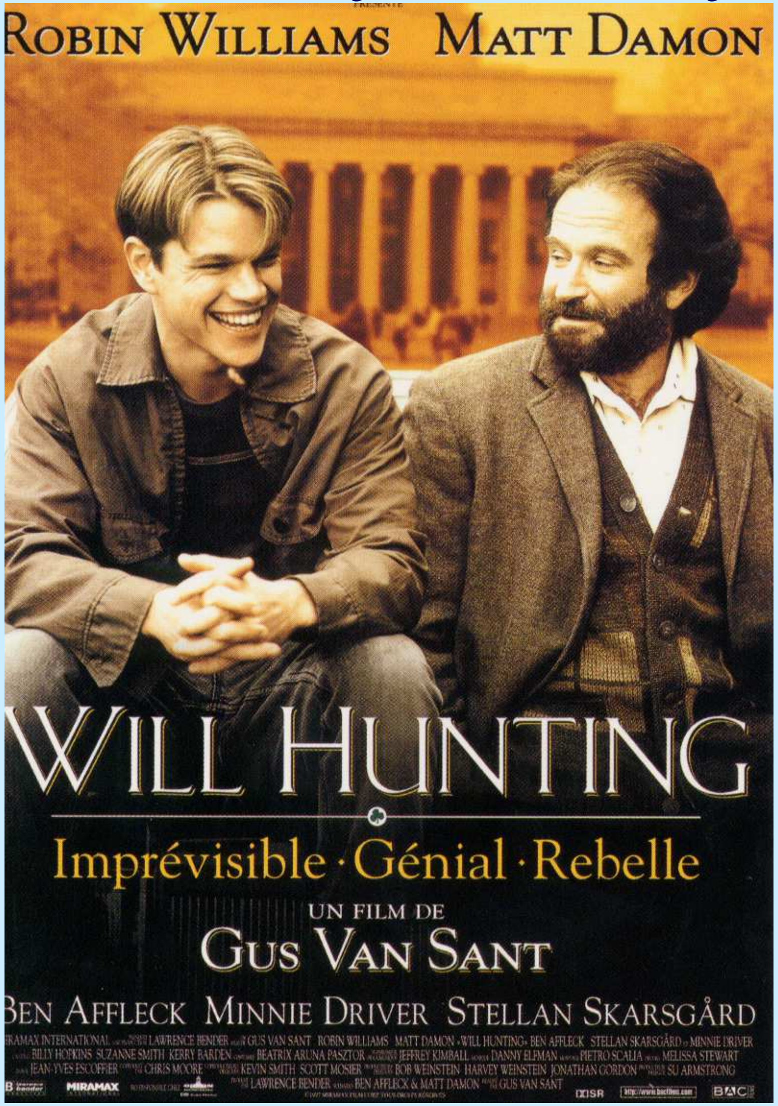
A polyhedron



Application examples:

Production problems, assignment problems, transport problems, flow problems, etc.

The father of linear optimization



Georges B. Dantzig

(1914-2005)

American statistician

Stanford *Professor*

Originated the
« difficult homework »
urban legend

Simplex method

Dantzig's real contribution is not the definition of linear optimization but the invention of the **simplex method** to solve it

- Problem look continuous, but is in reality essentially *combinatorial*
- Optimal solution must lie on *vertex* of D
- Algebraic description of vertices + improving *moves* from vertex to vertex

Algorithmic complexity

An efficient algorithm must

- Solve the problem (at least approximately)
- In a finite amount of time (must eventually stop)
- Or, better: in an amount of time bounded by a function of the problem size

Essential distinction

- polynomial-time complexity (= easy)
- exponential-time complexity (= hard)

Polynomial always beats exponential, eventually

n	n^5	2^n	$n^5/2^n$
20	3.05 seconds	1 second	3.05
40	1.7 minutes	12.1 days	9.3×10^{-5}
60	12.9 minutes	349 centuries	6.7×10^{-10}

General optimization is provable exponential

Could linear optimization be polynomial ?

Algorithmic complexity for linear optimization methods

- Georges Dantzig (1914-2005)

  Proposes in 1947 the simplex method for linear optimization, very efficient in practice but whose algorithmic complexity is unknown

- V. Klee et G. Minty (1972)

  Show that the simplex method is exponential in the worst case

- L. Khachyan (1978)

  Proves that the ellipsoid method (due to Shor) is polynomial in the worst ; it is however very slow in practice!

- N. Karmarkar (1984)

  Proposes a new type of method :interior-point methods, that are both polynomial in the worst case and very efficient in practice (often more than the simplex method)

Optimization before Dantzig: a few pioneers

Simplex method is the start of modern era of optimization

- Leonid Kantorovich (1912-1986)
  Linear optimization en 1939! (Nobel prize in 1975 with Koopmans)
- John Von Neumann (1903-1957)
  Game theory (minimax theorem en 1928, link with linear optimiz.)
- Carl Friedrich Gauss (1777-1855)
  Steepest descent method (non-linear optimization)
- Leonhard Euler (1707-1783)
  Calculus of variations (infinite-dimensional optimization)
- Pierre de Fermat (1601-1665)
  Light: takes minimum-time path (already by Hero of Alexandria)
- Nature !
  In physics: least action principle, minimum energy principle

An example of linear optimization

or

Can optimization help with your diet ?

The diet problem

Eating not just a good breakfast, but an *optimal* breakfast.

Eight foods to choose from:

1. pancakes $= x_1$
2. milk $= x_2$
3. freshly squeezed orange juice $= x_3$
4. scrambled eggs $= x_4$
5. Weetabix, a high-fiber cereal $= x_5$
6. fresh blueberries $= x_6$
7. bacon $= x_7$
8. vitamin/mineral pills $= x_8$

Source: M. Wright








Source: M. Wright

Our goal is to find the cheapest breakfast satisfying

daily fiber ≥ 120, daily calcium ≥ 100, and all x's ≥ 0.

(Can't eat -5 units of a food.)

food	fiber	calcium	cost
pancakes	5	10	3
milk	0	40	1
orange juice	4	5	5
eggs	8	0	3
Weetabix	100	20	10
blueberries	50	20	5
bacon	0	0	10
pills	30	400	2

Source: M. Wright

With this (apparently sensible) formulation, what's the optimal breakfast??

1.14 units of Weetabix and .19 units of vitamin/mineral pills.

Yuck. Not even milk for the Weetabix!

Source: M. Wright

The original diet problem

Compute the optimal daily diet for US soldiers

Satisfy nutritional requirements

Minimize the cost

77 types of food, 9 nutritional requirements

Solved in 1947 by Jack Laderman's team in 120 man-days

Optimal cost found: $39.69 per year

Duality

or

Can we trust optimization ?

Trust an optimal solution ?

Difference between solving

- A system of equations : $0 = f_1(x) = f_2(x) = \ldots$
- An optimization problem: min f(x)

How ?

- For a system of equations:
  simply replace x in $f_1(x)$, $f_2(x)$, …
- For an optimization problem : **how ?**

Generate bounds for linear optimization

Consider the small problem

$$f^* = \max\ x_1 + 2x_2 + 3x_3 \text{ such that } \begin{array}{rcll} x_1 + x_2 & \leq & 1 & (a) \\ x_2 + x_3 & \leq & 2 & (b) \\ x_2 + 3x_3 & \leq & 9 & (c) \\ x_3 & \leq & 3 & (d) \end{array}$$

a. Vector $x = (1, 0, 2)$ is feasible and corresponds to objective value equal to 7

   $\Rightarrow$ lower bound $f^* \geq 7$

b. Combine the constraints according to $(a) + (c)$

   $(x_1 + x_2) + (x_2 + 3x_3) \leq 1 + 9 \Leftrightarrow x_1 + 2x_2 + 3x_3 \leq 10$

   $\Rightarrow$ upper bound $f^* \leq 10$

Finding better lower bounds

$$f^* = \max\ x_1 + 2x_2 + 3x_3 \text{ such that } \begin{aligned} x_1 + x_2 &\leq 1 \quad (a) \\ x_2 + x_3 &\leq 2 \quad (b) \\ x_2 + 3x_3 &\leq 9 \quad (c) \\ x_3 &\leq 3 \quad (d) \end{aligned}$$

a. Lower bound can be improved: find a better vector

$x = (2, -1, 3)$ gives an objective function equal to 9

$\Rightarrow$ lower bound $f^* \geq 9$

To find better vectors we may use the simplex method

Finding better upper bounds

$$f^* = \max\ x_1 + 2x_2 + 3x_3 \text{ such that } \begin{aligned} x_1 + x_2 &\le 1 \quad (a) \\ x_2 + x_3 &\le 2 \quad (b) \\ x_2 + 3x_3 &\le 9 \quad (c) \\ x_3 &\le 3 \quad (d) \end{aligned}$$

a. $x = (2, -1, 3) \Rightarrow$ lower bound $f^* \ge 9$

b. Upper bounds can also be improved:

combining the constraints according to $(a)+(b)+2(d)$

$x_1+x_2+x_2+x_3+2x_3 \le 1+2+2\times 3 \Leftrightarrow x_1+2x_2+3x_3 \le 9$

$\Rightarrow$ upper bound $f^* \le 9 \Rightarrow (2, -1, 3)$ is optimal!

How to find better upper bounds?

$$f^* = \max \; x_1 + 2x_2 + 3x_3 \text{ such that } \begin{array}{rcll} x_1 + x_2 & \le & 1 & (a) \\ x_2 + x_3 & \le & 2 & (b) \\ x_2 + 3x_3 & \le & 9 & (c) \\ x_3 & \le & 3 & (d) \end{array}$$

Finding upper bound is an optimization problem!

Combining the constraints as $y_1(a) + y_2(b) + y_3(c) + y_4(d)$ we should

- choose nonnegative y_1, y_2, y_3, y_4 nonnegative
- satisfying $y_1 = 1, y_1 + y_2 + y_3 = 2, y_2 + 3y_3 + y_4 = 3$
- *minimizing* value of the bound $y_1 + 2y_2 + 9y_3 + 3y_4$

The dual problem

$$f^* = \max\ x_1 + 2x_2 + 3x_3 \text{ such that } \begin{array}{rcll} x_1 + x_2 & \leq & 1 & (a) \\ x_2 + x_3 & \leq & 2 & (b) \\ x_2 + 3x_3 & \leq & 9 & (c) \\ x_3 & \leq & 3 & (d) \end{array}$$

and

$$d^* = \min y_1 + 2y_2 + 9y_3 + 3y_4 \text{ such that } \begin{array}{rcl} y_1 & = & 1 \\ y_1 + y_2 + y_3 & = & 2 \\ y_2 + 3y_3 + y_4 & = & 3 \\ y_1, y_2, y_3, y_4 & \geq & 0 \end{array}$$

form a pair of dual problems!

Duality properties

a. Weak duality

Any feasible vector for the dual problem provides an upper bound for *all* feasible vectors of the problem (and vice-versa!)

b. Strong duality

There *always* exists a feasible dual vector that certifies optimality of the primal solutions

Therefore both problems have the *same* optimal value!

This optimal dual vector is an optimality certificate, can be found by any means (guessing, simplex method, etc.) and can be checked easily and independently

Duality summarized

- Duality is a way to certify quality or optimality of some feasible vectors
- Computing dual certificate is itself a linear optimization problem
- Some methods automatically compute the dual optimal solution (such as the simplex method)
- Dual variables have an economic interpretation (they act as prices for the constraints)

Performance

or

Can Optimization compete with computers?

Progress for linear optimization

In 1947, Laderman solves a small diet problem (77 ingredients x 9 nutriments) in 120 (wo)men-days (with Dantzig's simplex algorithm)

In 2008, linear problems with millions of variables/constraints are routinely solved on standard desktop computer
(provided enough memory is available)

Is it really remarkable ?
Modern computers are so powerful …
Is this a simple consequence of Moore's law, i.e. speed doubles every 1.5 year ?

A 25-year comparison: 1987-2002

R. Bixby considers, for the same problem

- The 1987 version of his solver CPLEX
  on a 1987 computer : 2.5 hours (~10 000 s)

- The same 1987 version CPLEX
  on a 2002 cmputer : 10 seconds

This matches *Moore's law quite well*

A 25-year comparison: 1987-2002

R. Bixby also compares on a larger problem:

- The 1987 version of CPLEX
  on a 2002 computer: 8 hours (~30 000 s)
- The 2002 version of CPLEX
  on a 2002 computer : 30 seconds

Algorithmic progress matches
technological progress: 1000x each

Total gain: 6 orders of magnitude in 15 years

Summary

Optimization is exciting

- Can be used everywhere
- Is actually used everywhere!

- Provably difficult … but we can manage

- Many large-scale problems can be solved efficiently in practice
- A lot of research still going on

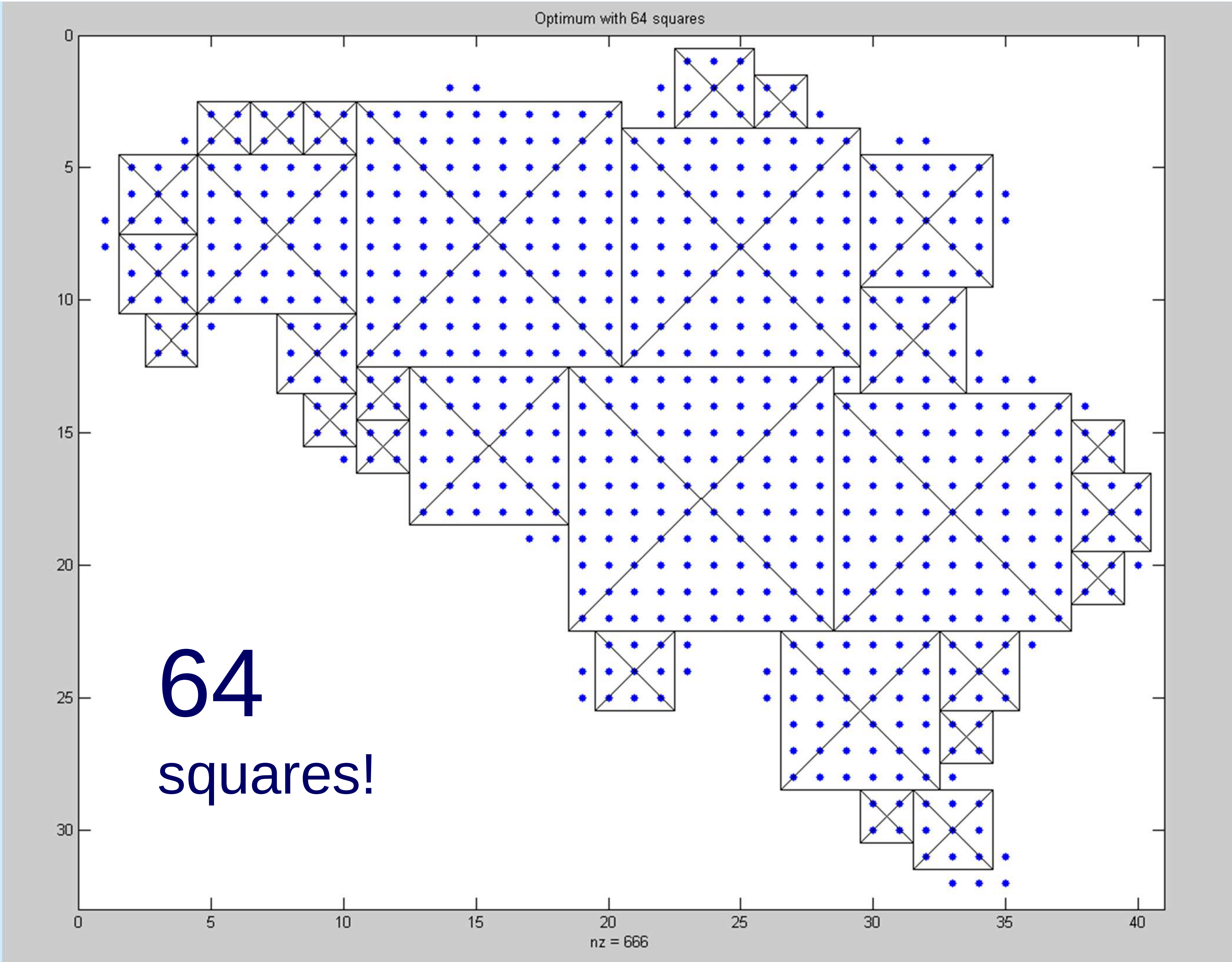
There is a lot more to say …

- Nonlinear Optimization (and convexity)
- Discrete Optimization
- Convex/Structured optimization

- Combinatorial Optimization
- Global Optimization
- Heuristic and meta-heuristic methods
- Average case analysis
- etc.

Thank you for your attention

Can you match
an optimization solver ?


Optimum with 64 squares
0
5
10
15
20
25
30
0
5
10
15
20
25
30
35
40
nz = 666
64
squares!