Speed dates with optimization problems

10 problems in 10 minutes

tommyod @ GitHub

February 15, 2019

Introduction

Understand $\leftrightarrow$ **Search** $\leftrightarrow$ Solve

- Efficient algorithm exist for many problems.
- Implementations of these algorithms are also available.
- The remaining difficulty is often having enough prior knowledge to recognize a problem.

Figure: Johann Bernoulli, circa 1740

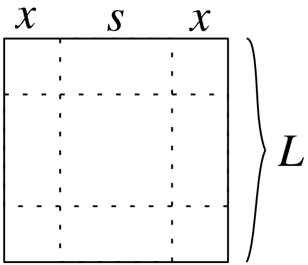
(1) The paper box problem

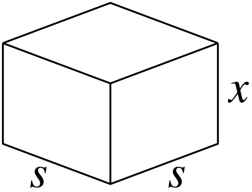
Introduction

Given a square piece of paper, cut the paper at a location x to maximize the volume of the resulting box.

This problem can be formulated as

$$\begin{aligned} \text{minimize} \quad & V(x,s) = s^2 x \\ \text{subject to} \quad & 2x + s = L \\ & s, x \geq 0. \end{aligned}$$

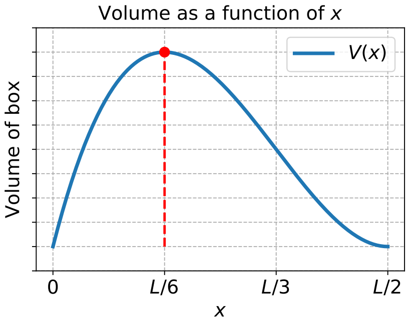




(1) The paper box problem

Problem instance

$$\begin{aligned} \text{minimize} \quad & V(x) = (L-2x)^2 x \\ \text{subject to} \quad & x \geq 0 \\ & x \leq L/2 \end{aligned}$$



Problem generalization

Given a smooth function $f : \mathbb{R} \to \mathbb{R}$.

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & x \geq a \\ & x \leq b \end{aligned}$$

Solved using differentiation.

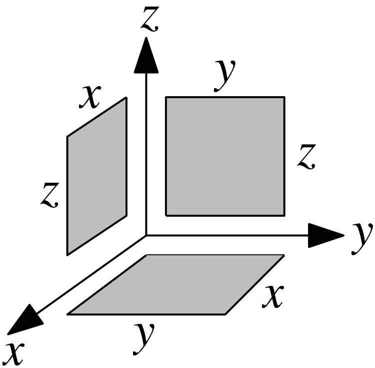
(2) The industrial box problem

Introduction

Given a square meter price for the material of a box with no lid, construct the cheapest box having unit volume.

This problem may be formulated as

$$
\begin{array}{rl}
\text{minimize} & P(x,y,x) = xy + 2yz + 2xz \\
\text{subject to} & V(x,y,z) = xyz = 1 \\
& x,y,z \geq 0.
\end{array}
$$

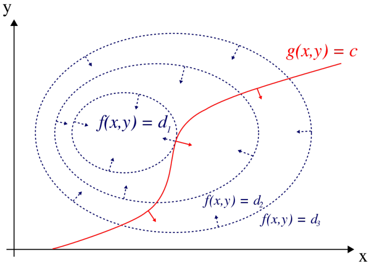


(2) The industrial box problem

Problem instance

Construct a *Lagrange function* $L(x,y,z,\lambda)$, solve the equations

$$
\begin{aligned}
L_x &= y + 2z + \lambda yz &&= 0 \\
L_y &= x + 2z + \lambda xz &&= 0 \\
L_z &= 2(x+y) + \lambda xy &&= 0 \\
L_\lambda &= xyz - 1 &&= 0
\end{aligned}
$$



Problem generalization

Given a smooth function $f : \mathbb{R}^n \to \mathbb{R}$.

$$
\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & g_i(\mathbf{x}) = 0 \\
& i = 1, 2, \ldots
\end{aligned}
$$

Introduction

A company wants equal exposure to 4 sub-populations. Given 2 advertisement channels and their associated reach units views/dollar, spend the money.

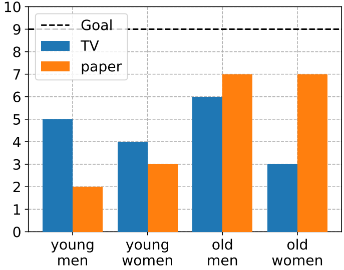
$$\begin{pmatrix}5\\4\\6\\3\end{pmatrix} x_1 + \begin{pmatrix}2\\3\\7\\7\end{pmatrix} x_2 = \begin{pmatrix}9\\9\\9\\9\end{pmatrix} + \begin{pmatrix}e_1\\e_2\\e_3\\e_4\end{pmatrix}$$

This problem can be formulated as

$$\text{minimize} \quad \sum_{i=1}^{4} e_i^2 = \mathbf{e}^T \mathbf{e} = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 .$$

Views per unit of money

$$A = \begin{array}{c} \\ \text{y}\,♀ \\ \text{y}\,♂ \\ \text{o}\,♀ \\ \text{o}\,♂ \end{array} \begin{array}{c} \begin{array}{cc} \text{tv} & \text{paper} \end{array} \\ \begin{pmatrix} 5 & 2 \\ 4 & 3 \\ 6 & 7 \\ 3 & 7 \end{pmatrix} \end{array}$$



(3) The advertisement problem

Problem instance

$$\text{minimize} \quad \mathbf{e}^T\mathbf{e} = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$$

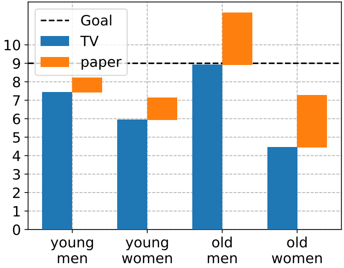
Problem generalization

Minimizing $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ is a *least squares problem*, solved analytically by the equation

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

Solution

$$\begin{pmatrix}5\\4\\6\\3\end{pmatrix}\mathbf{1.5} + \begin{pmatrix}2\\3\\7\\7\end{pmatrix}\mathbf{0.4} = \begin{pmatrix}9\\9\\9\\9\end{pmatrix} + \begin{pmatrix}-0.7\\-1.8\\2.8\\-1.7\end{pmatrix}$$



(4) The *constrained* advertisement problem

Introduction

Same as before, but constrained by a budget of 1 unit of money.

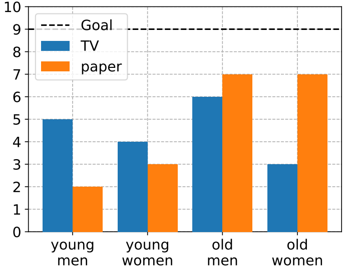
$$\begin{pmatrix}5\\4\\6\\3\end{pmatrix} x_1 + \begin{pmatrix}2\\3\\7\\7\end{pmatrix} x_2 = \begin{pmatrix}9\\9\\9\\9\end{pmatrix} + \begin{pmatrix}e_1\\e_2\\e_3\\e_4\end{pmatrix}$$

This problem can be formulated as

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{4} e_i^2 = \mathbf{e}^T\mathbf{e} = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \\ \text{subject to} \quad & x_1 + x_2 = 1 \end{aligned}$$

Views per unit of money

$$A = \begin{array}{c} \\ \text{y}\,♀ \\ \text{y}\,♂ \\ \text{o}\,♀ \\ \text{o}\,♂ \end{array} \begin{array}{c} \begin{array}{cc} \text{tv} & \text{paper} \end{array} \\ \begin{pmatrix} 5 & 2 \\ 4 & 3 \\ 6 & 7 \\ 3 & 7 \end{pmatrix} \end{array}$$



Problem instance

$$\begin{aligned} \text{minimize} \quad & \mathbf{e}^T\mathbf{e} = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \\ \text{subject to} \quad & \mathbf{x}^T\mathbf{1} = 1 \end{aligned}$$

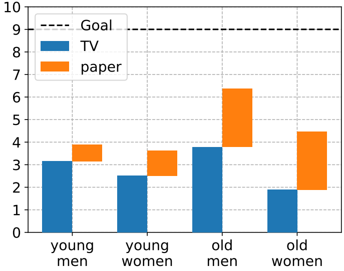
Problem generalization

Constrained least squares problem, solved by Lagrange multipliers and linear algebra.

$$\begin{pmatrix} 2\mathbf{A}^T\mathbf{A} & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{z}} \end{pmatrix} = \begin{pmatrix} 2\mathbf{A}^T\mathbf{b} \\ \mathbf{d} \end{pmatrix}$$

Solution

$$\begin{pmatrix} 5 \\ 4 \\ 6 \\ 3 \end{pmatrix} \mathbf{0.6} + \begin{pmatrix} 2 \\ 3 \\ 7 \\ 7 \end{pmatrix} \mathbf{0.4} = \begin{pmatrix} 9 \\ 9 \\ 9 \\ 9 \end{pmatrix} + \begin{pmatrix} -5.1 \\ -5.4 \\ -2.6 \\ -4.5 \end{pmatrix}$$



(5) The worker-assigment problem

Introduction

Assign 4 workers to 4 tasks, given a matrix C specifying to which degree workers enjoy each task.

This amounts to specifying X with entries in $X_{ij} \in \{0,1\}$, i.e.

$$X = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

The above yields a satisfaction of

$$6+4+8+1 = 19.$$

Problem data

$$C = \begin{array}{r c} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{r} \text{ole} \\ \text{åse} \\ \text{dag} \\ \text{lise} \end{array} & \begin{pmatrix} 5 & 6 & 1 & 6 \\ 4 & 5 & 0 & 1 \\ 1 & 2 & 6 & 8 \\ 7 & 2 & 1 & 1 \end{pmatrix} \end{array}$$

Solution space growth

n	digits in $n!$
1	1
10	7
25	26
50	65

Problem instance

$$\begin{aligned} \text{minimize} \quad & -\sum_i \sum_j C_{ij} X_{ij} \\ \text{subject to} \quad & \sum_i X_{ij} = 1 \text{ for every } j \\ & \sum_j X_{ij} = 1 \text{ for every } i \end{aligned}$$

Solution

$$C_{ij}\widehat{X}_{ij} = \begin{array}{r c} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{r} \text{ole} \\ \text{åse} \\ \text{dag} \\ \text{lise} \end{array} & \begin{pmatrix} 5 & 6 & 1 & \mathbf{6} \\ 4 & \mathbf{5} & 0 & 1 \\ 1 & 2 & \mathbf{6} & 8 \\ \mathbf{7} & 2 & 1 & 1 \end{pmatrix} \end{array}$$

Problem generalization

This is the *assignment problem*, solved by the Hungarian algorithm. Solved in $\mathcal{O}(n^3)$ time, not $\mathcal{O}(n!)$.

$$-\sum_i \sum_j C_{ij}\widehat{X}_{ij} = 6 + 4 + 6 + 7 = 23$$

(6) The diet problem

Introduction

Minimize the total cost of the diet, subject to the dietary constraints.

$$\begin{aligned} \text{minimize} \quad & p_1x_1 + p_2x_2 + p_3x_3 \\ \text{subject to} \quad & 16x_1 + 5x_2 + 12x_3 \geq 100 \\ & 150x_1 + 100x_2 + 40x_3 \geq 2000 \\ & 150x_1 + 100x_2 + 40x_3 \leq 2500 \end{aligned}$$

“Minimize cost, but get 100 grams of protein, and between 2000 and 2500 calories.”

Problem data

Food	Price	Protein	Calories
x_1 eggs	p_1	16	150
x_2 bread	p_2	5	100
x_3 milk	p_3	12	40

The numbers above are fictitious.

(6) The diet problem

Problem instance

$$
\begin{array}{rl}
\text{minimize} & \begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^T \\
\text{subject to} & \begin{pmatrix} 16 & 5 & 12 \\ 150 & 100 & 40 \\ -150 & -100 & -40 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \geq \begin{pmatrix} 100 \\ 2000 \\ -2500 \end{pmatrix} \\
& x_1, x_2, x_3 \geq 0
\end{array}
$$

Problem generalization

This is the *linear programming* problem. Efficient algorithms exist.

$$
\begin{array}{rl}
\text{minimize} & \mathbf{p}^T \mathbf{x} \\
\text{subject to} & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\
& \mathbf{x} \geq \mathbf{0}
\end{array}
$$

(7) The hotel problem

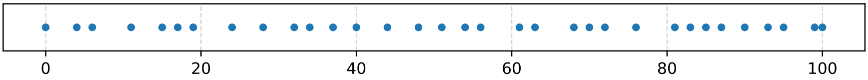
Introduction

We wish to travel 100 units of distance. There are many hotels along the way. Pick hotels to travel ~10 units per day.

Problem instance

Set $M = 10$. Find a sequence $h_1, h_2, \ldots, h_n$ to

$$\text{minimize} \quad \sum_j (M - (h_j - h_{j-1}))^2 .$$



Examples

Traveling from $x = 0$ to $x = 6$ incurs a penalty of $(10 - (6 - 0))^2 = 4^2$.
Traveling from $x = 0$ to $x = 11$ incurs a penalty of $(10 - (11 - 0))^2 = 1^2$.
There are 31 hotels above, and $2^{31} = 2\,147\,483\,648$ possibilities.

(7) The hotel problem

Problem

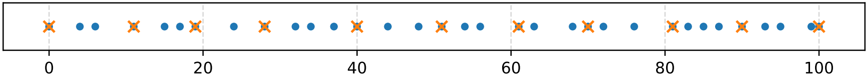
Let $P(j)$ be the minimal penalty at stop j. Realize that

$$P(j) = \min_{0 \leq i < j} \left(P(i) + (M - (h_j - h_i))^2 \right).$$

Solved in $\mathcal{O}(n^2)$ time, not $\mathcal{O}(2^n)$.

Problem generalization

The solution technique is called *dynamic programming*, and depends on an *optimal substructure* property. (1) Identity recursive relationship, (2) initial conditions and (3) solve problems in correct order.



To apply dynamic programming we must (1) identity recursive relationship, (2) initial conditions and (3) solve problems in correct order..

(8) The magnet problem

Introduction

We are given 6 magnets. Choose $x_i \in \{-1, 1\}$ to minimize the total energy

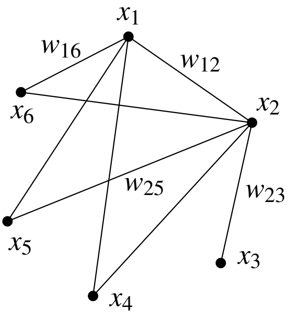
$$E(\mathbf{x}) = w_{12}x_1x_2 + w_{13}x_1x_3 + \cdots + w_{56}x_5x_6.$$

The problem can be formulated as

$$\begin{array}{ll} \text{minimize} & E(\mathbf{x}) = \mathbf{x}^T \mathbf{W} \mathbf{x} \\ \text{subject to} & x_i \in \{-1, 1\}. \end{array}$$

There are 2^{6-1} states, and $E(\mathbf{x})$ is not differentiable. A difficult problem.

Problem



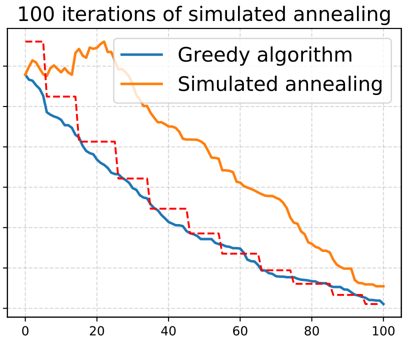
(8) The magnet problem

Problem instance

$$\begin{aligned} \text{minimize} \quad & E(\mathbf{x}) = \mathbf{x}^T \mathbf{W} \mathbf{x} \\ \text{subject to} \quad & x_i \in \{-1, 1\}. \end{aligned}$$

Problem generalization

When we have a (1) non-differentiable function with (2) a wast search space and (3) a notion of neighborhoods, use *simulated annealing* to balance exploitation and exploration.



(9) The egg boiling problem

Introduction

Let b be the boiling time of an egg, c be the cooling time, and s be the amount of salt used. Let $f(b,c,s) : \mathbb{R}^3 \to \mathbb{R}$ be the quality of a boiled egg.

This problem can be formulated as

$$\begin{array}{rl} \text{minimize} & -f(b,c,s) \\ \text{subject to} & b,c,s \geq 0 \end{array}$$

Evaluating $f(b,c,s)$ is expensive.

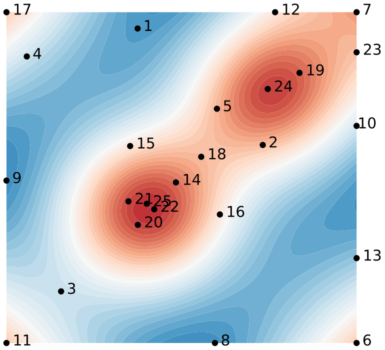
Problem generalization

Given a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ which is expensive to evaluate.

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{a} \leq \mathbf{x} \leq \mathbf{b} \end{aligned}$$

Clever sampling via *bayesian optimization*, which builds a probability distribution over functions. Exploration vs. exploitation.

Bayesian Optimization



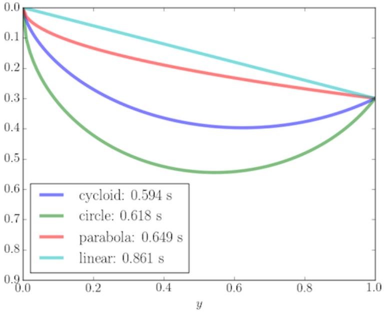
Introduction

A *functional* is a function from a function to real number.

$$\text{functional} : \text{function} \rightarrow \text{real number}$$

Problem

Find the path (i.e. function) minimizing the travel time of a bead.



(10) The brachistochrone problem

Problem generalization

The problem amounts to minimizing a functional. In the space of all functions, find a function minimizing a functional. This is the domain of *calculus of variations*.

Johann Bernoulli solved the problem in 1696. The solution is a cycloid.

References (1/2)

- Strang, Gilbert. *Introduction to Applied Mathematics*. Wellesley, Mass: Wellesley-Cambridge Press, 1986.
  - Chapter 3 – “The brachistochrone problem”
  - Chapter 7 – “The worker-assignement problem”
  - Chapter 8 – “The diet problem”
- Boyd, Stephen, and Lieven Vandenberghe. *Introduction to Applied Linear Algebra*. Cambridge University Press, 2018.
  - Chapter 12 – “The advertisement problem”
  - Chapter 16 – “The constrained advertisement problem”
- Dasgupta, Sanjoy, Christos H. Papadimitriou, and Umesh Virkumar. Vazirani. *Algorithms*. Boston, Mass: McGraw Hill, 2008.
  - Chapter 6 – “The hotel problem”

- Duda, Richard O., Peter E. Hart, and David G. Stork. *Pattern Classification*. 2 edition. New York: Wiley-Interscience, 2000.
  - Chapter 7 – “The magnet problem”
- Jasper Snoek, Hugo Larochelle, Ryan P. Adams. *Practical Bayesian Optimization of Machine Learning Algorithms*. arXiv.org, 2012.
  - “The egg boiling problem”

Thank you for your attention.

Slides, LaTeX source and Python code solving the problems and generating plots:

`github.com/tommyod/10_optimization_problems`