BCD&ADMM 收敛理论

罗自炎

zyluo@bjtu.edu.cn

AI4M优化方向暑期班 北京大学
2024.7

课程参考资料

课程参考资料

- 刘浩洋，户将，李勇锋，文再文，**最优化：建模、算法与理论**，高教出版社，2020
- Jérôme Bolte, Shoham Sabach, Marc Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, *Math. Program., Ser. A* (2014) 146:459-494
- Maryam Fazel, Ting Kei Pong, Defeng Sun, and Paul Tseng, Hankel matrix rank minimization with applications to system identification and realization, *SIAM J. Matrix Anal. Appl.* 34 (2013) 946-977

Lecture 1：最优化简介

最优化简介

最优化的定义

Q：什么是最优化？

A1： 在所有可能中挑选最好的【通俗解释】

informs https://www.informs.org/

Operations Research:
The Science of Better

Better Choices,
Better Life

最优化简介

最优化问题无处不在!

- ✓ 早在18世纪，著名数学家欧拉曾说：“宇宙万物无不与最小化或者最大化原理有关系.”
- ✓ 最优化具有基础性、实用性，一直为学术界和实际部门所重视.

经济金融：最大利润、最小风险
交通运输：列车运行图、物流
信息科学：数据挖掘、图像处理
生命科学：DNA 序列、蛋白质折叠
工程力学：最大载重、最优结构
军事国防：排兵布阵、后勤保障

任何决策问题都是最优化问题！

最优化简介

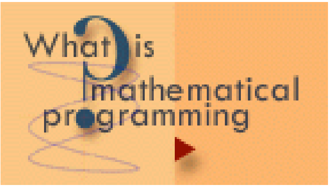
最优化的定义

Q：什么是最优化？

A2： 根据国际数学优化学会(MOS)定义，最优化问题是指在一定**约束条件**下**极大化或极小化**某一**实值函数**的问题，其**变量**可能是实值的或整数的.

Mathematical Optimization Society

http://www.mathopt.org/



In a mathematical optimization (or programming) problem, one seeks to minimize or maximize a real function of real or integer variables, subject to constraints on the variables.

最优化简介

最优化的数学模型

目标函数

$$\min \ \boxed{f(x)}$$

$$s.t.\ \boxed{x} \in \boxed{\Omega}$$

决策变量

可行域or决策集

无约束优化

$$\min_{x\in R^n} \ f(x)$$

约束优化

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & c_i(x) \geq 0, \quad i \in \boldsymbol{I} \\ & c_j(x) = 0, \quad j \in \boldsymbol{\varepsilon} \end{aligned}$$

约束函数

最优化简介

最优解概念

- **全局最优解：** 设$f(x)$为目标函数，Ω为可行域，$x^* \in \Omega$. 若 $\forall x \in \Omega$，有 $f(x) \geq f(x^*)$，则称x^*为极小化问题 $\min_{x\in\Omega} f(x)$ 的全局最优解，记为 $x^* \in \mathbf{argmin}\{f(x): x \in \Omega\}$.

  the argument of minimum

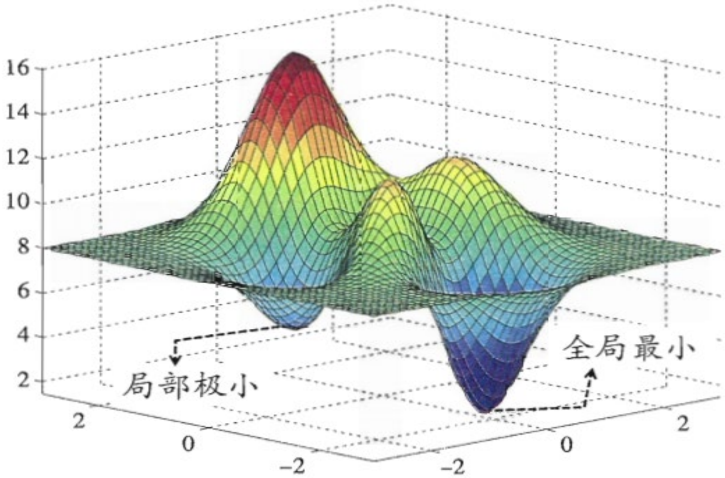
  ✓ 全局最优解也称为整体最优解, 英文为*Global optimal solution*

- **局部最优解：** 设$f(x)$为目标函数，Ω为可行域，$x^* \in \Omega$. 若存在正实数$\varepsilon > 0$与x^*的邻域 $\mathrm{N}_\varepsilon(x^*) \coloneqq \{x: \|x - x^*\| < \varepsilon\}$，使得 $\forall x \in \Omega \cap \mathrm{N}_\varepsilon(x^*)$，有 $f(x) \geq f(x^*)$，则称x^*为极小化问题 $\min_{x\in\Omega} f(x)$ 的一个局部最优解；若$\forall x \in \Omega \cap \mathrm{N}_\varepsilon(x^*)$且 $x \neq x^*$，有$f(x) > f(x^*)$，则称x^*为极小化问题 $\min_{x\in\Omega} f(x)$ 的一个严格局部最优解.

  ✓ 局部最优解英文为*Local optimal solution*

最优化简介

最优解概念



Global minimizer

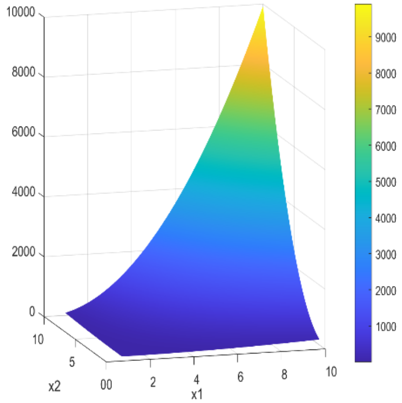
Local minimizer

最优化简介

最优值概念

- 若优化问题存在最优解，则该优化问题的目标函数在最优解处的取值称为**最优值**，相应的优化问题为**最优值可达（attainable）**。
- 若目标函数在可行域有下界，但无最优解，则称目标函数在可行域上的下确界为优化问题的最优值，此时**最优值不可达**。例如：$f(x) = x_1^2 + (1 - x_1 x_2)^2$的最优值为0，但只有在$x_1 = x_2^{-1}$且$x_2 \to \infty$时才能达到。

鉴于此，我们将最优值记为 $f^* = \inf_{x \in \Omega} f(x)$



最优化简介

最优化主要研究内容

理论：

- 一阶/二阶最优性条件
- 对偶理论
- 鞍点问题
- 灵敏度分析
- 复杂性

算法：

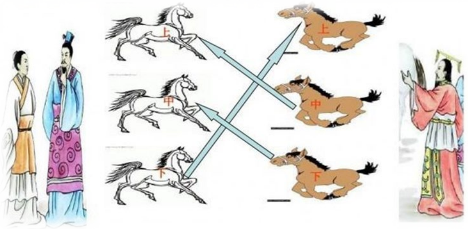
- 最速下降方法
- 共轭梯度方法
- 牛顿方法
- 拟牛顿方法
- 罚函数方法
- 复合优化算法

应用：

- 建立模型
- 理论分析
- 编程计算
- 解决实际问题

最优化简介

优化思想起源



- 公元前340年田忌赛马：“今以君之下驷与彼上驷，取君上驷与彼中驷，取君中驷与彼下驷”

- 《史记一高祖本纪》中有一段话 ： “夫运筹策帷帐之中，决胜于千里之外，吾不如子房”

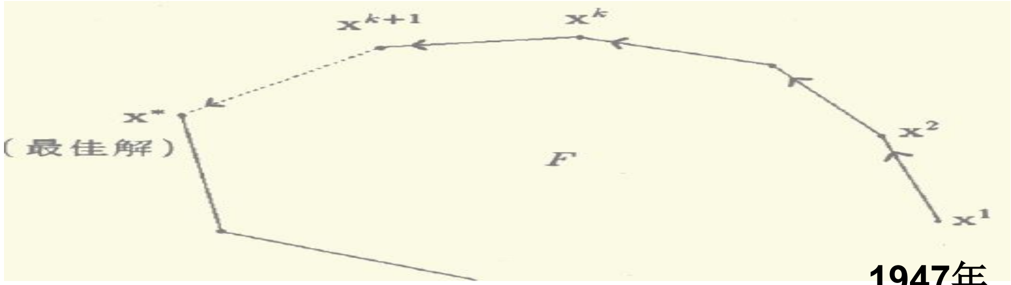
- 18世纪Euler,Lagrange等对与力学相关的极值问题或变分问题统一处理方法的研究

最优化简介

线性规划与单纯形方法

George Dantzig

1914.11.8--2005.5.13



1947年

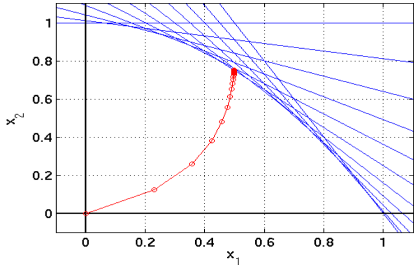
- 线性规划之父
- 师从著名统计学家Jerzy Neyman

一个人的潜能是难以预料的，成功的障碍往往来自于心理上的畏难情绪；一定要相信自己，保持积极的态度。

最优化简介

最优化的发展

- **线性整数规划：** Dantzig，Fulkerson和Johnson在1950年研究旅行商问题；Gomory提出的割平面方法则奠定了现代整数规划算法的基础.
- **非线性规划：** 1939年 Karush, 1951年Kuhn&Tucker提出了约束最优化问题必要条件—KKT条件，标志着现代非线性规划理论研究的开端.
- **计算复杂性：** 1970年，Victor Klee与George Minty给出实例证明了单纯形方法不是多项式时间的，而是指数级 $O(2^n)$
  - 椭球法---L.G.Khachian 1979 $O(n^6L^2)$
  - 内点法---Karmarkar 1984 $O(n^{3.5}L^2)$
  - Nesterov和Nemirovski等推广到凸优化问题(90年代）.



最优化简介

最优化的发展

- 向量优化
- 互补与均衡问题
- 组合优化
- 半定规划
- 鲁棒优化
- 随机优化
- 稀疏优化
- 统计优化
- 张量与多项式优化
- 非光滑优化

···

最优化简介

■ **稀疏性：** 是指蕴含在高维数据中可降维或可压缩的结构特征.

元素稀疏性

$$x = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

向量ℓ_0范数: $\|x\|_0 = 2$

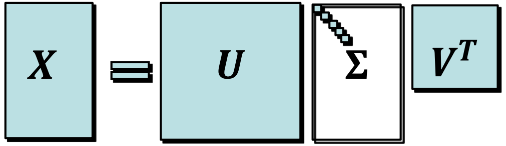
组稀疏性

$$X = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 3 & \cdots & n \\ 0 & 0 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & 0 \\ -2 & 5 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

矩阵$\ell_{q,0}$范数: $\|X\|_{q,0} = 2$

矩阵低秩性

$$X = U\Sigma V^T$$



矩阵秩函数:

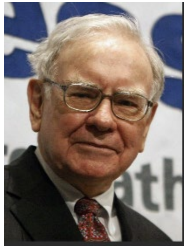
$$\mathrm{rank}(X) = \|\mathrm{diag}(\Sigma)\|_0$$

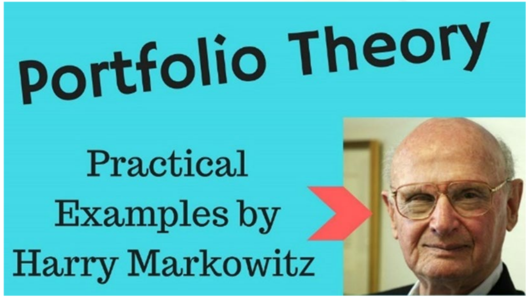
最优化简介

➢ 金融：稀疏投资组合

优化模型：
$$\begin{aligned} &\min\ x^T V x \\ &\text{s.t.}\ \ r^T x = \mu \\ &\qquad\ e^T x = 1 \\ &\qquad\ \|x\|_0 \le s \end{aligned}$$

“如果你投资股票建议最好不要超过20只”





其中：

r：期望回报率向量

V：投资组合的协方差矩阵

x：投资比例向量

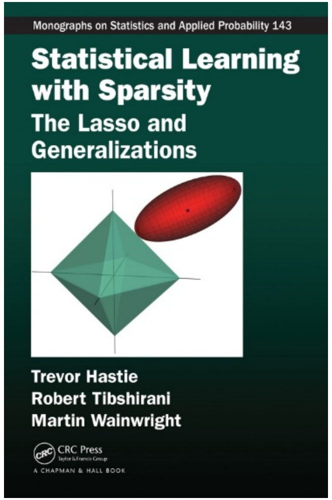
μ：期望的投资组合总回报

e：全1向量

[1] Gao J J, Li D. Optimal cardinality constrained portfolio selection, Operational Research, 2013, 61(3): 745-761.

最优化简介

➢ 稀疏统计学习



- □ 变量选择:

$$\min \frac{1}{N}\|y - X\beta\|^2 + \lambda\|\beta\|_0$$

- □ LASSO模型：

$$\min \frac{1}{N}\|y - X\beta\|^2 + \lambda\|\beta\|_1$$

【高维统计回归、小样本】

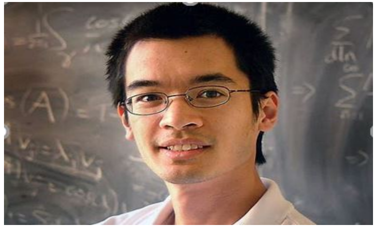
“我们淹没在了信息的海洋里，却渴求着知识”

“尽管世界如此复杂，但有用的信息却非常有限”

“既然无法有效处理稠密问题，倒不如在稀疏问题上寻求有效的处理方法”

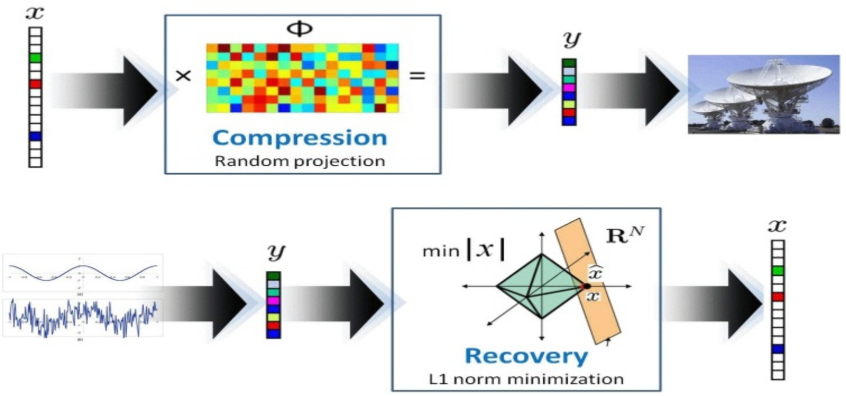
最优化简介

➢ 信号处理：压缩感知



陶哲轩

Donoho　　Candès



如果原始信号具有**稀疏性**的特征，那么通过少量的观测信息就能够恢复原始信号.

$$\min \ \|x\|_0$$
$$\text{s.t.} \ \ Ax = b$$

[2] E. J. Candès, Compressive sampling, International Congress of Mathematicians, Vol. III, 2006, 1433-1452.
[3] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory, 52 (2006), 1289-1306.
[4] E. J. Candès, T.Tao, Decoding by linear programming, IEEE Trans. Inf. Theory, 2005, 51(12): 4203-4215.

最优化简介

- **推荐系统：矩阵填充**

$$\begin{aligned} \min \quad & \operatorname{rank}(X) \\ \text{s.t.} \quad & X_{\Omega} = M \end{aligned}$$

Netflix百万大奖赛

最优化简介

- **Models--the art**: How we choose to represent real problems 【无约束优化\约束优化、凸优化\非凸优化、 光滑优化\非光滑优化、复合优化等】

- **Theory--the science**: What we know about different classes of models; e.g., necessary and sufficient conditions for optimality 【无约束优化最优性、约束优化最优性、对偶理论等】

- **Algorithms--the tools**: How we apply the theory to robustly and efficiently solve powerful models 【无约束优化方法（最速下降法、牛顿法、拟牛顿法、共轭梯度法）；约束优化方法（内点罚函数法、外点罚、ALM）；复合优化方法（APG、BCD、ADMM等）】

Thank You !