最优化方法

Convex Analysis

最优化方法

课程信息:

- 课程名称：最优化方法
- 开课学院：数学与统计学院
- 主讲老师：刘嘉
- 课时安排：周二 7-8 节（1-12 周）；周五 3-4 节（1-12 周）
- 电子邮件：jialiu@xjtu.edu.cn
- 答疑时间：周二 5-6 节、周三 5-7 节
- 答疑地点：理科楼 332

最优化方法

课程信息:

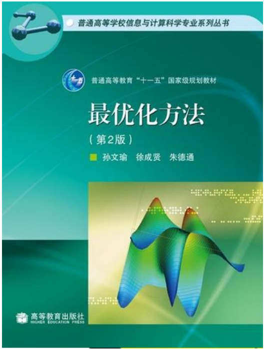
- 课程名称：最优化方法
- 开课学院：数学与统计学院
- 主讲老师：刘嘉
- 课时安排：周二 7-8 节（1-12 周）；周五 3-4 节（1-12 周）
- 电子邮件：jialiu@xjtu.edu.cn
- 答疑时间：周二 5-6 节、周三 5-7 节
- 答疑地点：理科楼 332

考核与成绩:

- 作业成绩 (30%)+ 期末考试 (70% - 闭卷)

教材

- 最优化方法，孙文瑜，徐成贤，朱德通著，高等教育出版社，2004



学校网络教材中心有在线电子版 http://jiaocai.lib.xjtu.edu.cn/

课程大纲

聚焦基本无约束优化和约束优化问题（连续、光滑、确定性）

- 最优化理论简介（凸分析，数学学院，秋上）
- 线性规划（运筹学）
- 一维线性搜索
- 无约束优化算法（重点）
- 最小二乘问题与二次规划
- 约束优化理论（重点）
- 约束优化方法

最优化

什么是最优化？最优化就是在众多可行的方案或方法中找到最好的方案或方法。

- 投资问题
- 运输问题
- 参数拟合问题

应用领域：

- 国防、工农业生产、交通运输、金融
- 贸易、管理、科学研究等

典型运筹学问题举例：生产安排问题

例 1.1.1 某公司生产甲乙两种商品，每件甲商品要耗 A 种原料 3kg，B 种原料 3kg，产值为 180 元；每件乙商品要耗 A 种原料 4.5kg，B 种原料 1.5kg，产值为 150 元．现公司共有 A 种原料 900kg，B 种原料 600kg，试确定甲乙两种商品共生产多少件，使得公司的总产值最大．

典型运筹学问题举例：生产安排问题

例 1.1.1 某公司生产甲乙两种商品，每件甲商品要耗 A 种原料 3kg，B 种原料 3kg，产值为 180 元；每件乙商品要耗 A 种原料 4.5kg， B 种原料 1.5kg，产值为 150 元. 现公司共有 A 种原料 900kg，B 种原料 600kg，试确定甲乙两种商品共生产多少件，使得公司的总产值最大.

建立模型： **变量**：生产甲商品 x_1 件，乙商品 x_2 件，故公司总产值为 $180x_1 + 150x_2$ （**目标函数**）.

约束：考虑到 A,B 两种原材料总量限制，故 x_1, x_2 需满足如下不等式约束条件：

$$\begin{cases} 3x_1 + 4.5x_2 \leq 900 \\ 3x_1 + 1.5x_2 \leq 600 \end{cases}$$

非负约束： $x_1 \geq 0, x_2 \geq 0.$

典型运筹学问题举例：生产安排问题

例 1.1.1 某公司生产甲乙两种商品，每件甲商品要耗 A 种原料 3kg，B 种原料 3kg，产值为 180 元；每件乙商品要耗 A 种原料 4.5kg，B 种原料 1.5kg，产值为 150 元．现公司共有 A 种原料 900kg，B 种原料 600kg，试确定甲乙两种商品共生产多少件，使得公司的总产值最大．

线性规划模型：

$$
\begin{aligned}
\max_{x_1,x_2} \quad & z = 180x_1 + 150x_2 \\
\text{s.t.} \quad & 3x_1 + 4.5x_2 \le 900 \\
& 3x_1 + 1.5x_2 \le 600 \\
& x_1 \ge 0, x_2 \ge 0
\end{aligned}
$$

该模型的数学性质，求解方法？ 最优化理论与方法

典型运筹学问题举例：生产安排问题

线性规划模型：

$$
\begin{aligned}
\max_{x_1,x_2} \quad & z = 180x_1 + 150x_2 \\
\text{s.t.} \quad & 3x_1 + 4.5x_2 \leq 900 \\
& 3x_1 + 1.5x_2 \leq 600 \\
& x_1 \geq 0, x_2 \geq 0
\end{aligned}
$$

Matlab 程序：

```
cvx_begin
variables x1 x2
maximize 182*x1+150*x2
subject to
3*x1+4.5*x2<=900;
3*x1+1.5*x2<=600;
cvx_end
```

典型运筹学问题举例：生产安排问题

cvx 求解结果：

```
Calling SDPT3 4.0: 2 variables, 0 equality constraints
------------------------------------------------------------


 num. of constraints =  1
 dim. of linear var  =  3
*******************************************************************
   SDPT3: Infeasible path-following algorithms
*******************************************************************
 version  predcorr  gam  expon  scale_data
    NT       1      0.000   1        0
it pstep dstep pinfeas dinfeas  gap      prim-obj      dual-obj    cputime
-------------------------------------------------------------------
 0|0.000|0.000|4.5e+00|1.2e+00|1.4e+03| 6.066667e+02  0.000000e+00| 0:0:00| chol  1  1
 1|1.000|1.000|1.5e-06|3.7e-03|1.1e+02| 7.229749e+01 -3.768628e+01| 0:0:00| chol  1  1
 2|0.989|0.989|4.4e-08|4.1e-04|1.2e+00| 7.884486e-01 -4.008792e-01| 0:0:00| chol  1  1
 3|0.989|0.989|2.0e-08|4.1e-05|1.3e-02| 8.664335e-03 -3.415978e-03| 0:0:00| chol  1  1
 4|0.989|0.989|5.4e-09|4.1e-06|1.5e-04| 9.521934e-05  6.136119e-05| 0:0:00| chol  1  1
 5|0.989|0.989|2.0e-11|4.2e-07|1.6e-06| 1.046442e-06  1.056424e-05| 0:0:00| chol  1  1
 6|0.995|0.999|8.3e-12|2.9e-10|2.3e-08| 1.680754e-08  1.335095e-09| 0:0:00| chol  1  1
 7|0.997|1.000|2.7e-14|1.1e-12|3.3e-10| 2.150869e-10 -8.082235e-11| 0:0:00|
```

典型运筹学问题举例：生产安排问题

cvx 求解结果：

```
primal objective value =  2.15086914e-10
dual   objective value = -8.08223508e-11
gap := trace(XZ)       = 3.26e-10
relative gap           = 3.26e-10
actual relative gap    = 2.96e-10
rel. primal infeas (scaled problem)   = 2.72e-14
rel. dual     "        "       "      = 1.10e-12
rel. primal infeas (unscaled problem) = 0.00e+00
rel. dual     "        "       "      = 0.00e+00
norm(X), norm(y), norm(Z) = 1.0e+00, 8.1e-11, 4.5e+01
norm(A), norm(b), norm(C) = 2.0e+00, 2.0e+00, 4.6e+01
Total CPU time (secs)  = 0.45
CPU time per iteration = 0.06
termination code       =  0
DIMACS: 2.7e-14  0.0e+00  1.2e-12  0.0e+00  3.0e-10  3.3e-10
-------------------------------------------------------------------

------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +42300
```

数学模型

数学模型一般形式为

$$
\begin{aligned}
\min_{x} \quad & f(x) \\
\text{s.t.} \quad & c_i(x) = 0, \quad i = 1, 2, \cdots, m, \\
& c_i(x) \geq 0, \quad i = m+1, \cdots, p,
\end{aligned}
\tag{1.1.6}
$$

其中

- 决策变量 $x = (x_1, x_2, \cdots, x_n)^\top \in R^n$;
- 目标函数 f: $R^n \to R^1$ (连续、可微);
- 约束函数 c_i : $R^n \to R^1 (i = 1, 2, \cdots, p)$ (等式、不等式)

缩写 min 对应 minimize （极小化）s.t. 对应 subject to（受约束）

优化问题分类

- 约束条件：无约束优化（不含约束条件）与约束优化
- 连续性：连续优化与离散优化（组合优化）
- 目标与约束的线性性：线性规划与非线性规划
- 光滑性：光滑优化与非光滑优化
- 优化目标数量：单目标与多目标
- 变量与时间关系：静态与动态
- 参数的随机性：确定性优化与随机优化

最优化的主要研究内容

最优化理论

- 优化问题的数学性质（凸性、光滑性）
- 最优性条件（约束规范，一阶条件，二阶条件）
- 最优解和最优值的连续性
- 参数的扰动分析，定性与定量稳定性

最优化算法

- 算法设计（无约束、约束、组合优化）
- 算法的收敛性
- 算法的收敛速度
- 算法的稳定性与泛化能力

非线性优化

例 1.1.2 某公司生产甲乙两种商品，每件甲商品要耗 A 种原料 3kg, B 种原料 3kg, 效用为 $u_1(x) = x^{1/2}$; 每件乙商品要耗 A 种原料 4.5kg，B 种原料 1.5kg, 效用为 $u_2(x) = x^{1/2}$. 现公司共有 A 种原料 900kg, B 种原料 600kg, 试确定甲乙两种商品共生产多少件, 使得公司的总产值最大.

效用可以为非线性的函数!

$$
\begin{aligned}
\max_{x_1,x_2} \quad & (x_1)^{1/2} + (x_2)^{1/2} \\
\text{s.t.} \quad & 3x_1 + 4.5x_2 \leq 900 \\
& 3x_1 + 1.5x_2 \leq 600 \\
& x_1 \geq 0, x_2 \geq 0
\end{aligned}
$$

非线性优化

组合优化

例 1.1.3 某公司生产甲乙两种商品，每件甲商品要耗 A 种原料 3kg，B 种原料 3kg，产值为 180 元；每件乙商品要耗 A 种原料 4.5kg，B 种原料 1.5kg，产值为 150 元．现公司共有 A 种原料 900kg，B 种原料 600kg，试确定甲乙两种商品共生产多少件，使得公司的总产值最大．

件数为整数！

$$
\begin{aligned}
\max_{x_1,x_2} \quad & z = 180x_1 + 150x_2 \\
\text{s.t.} \quad & 3x_1 + 4.5x_2 \le 900 \\
& 3x_1 + 1.5x_2 \le 600 \\
& x_1 \ge 0, x_2 \ge 0 \\
& x_1 \in \mathbb{Z}_+, x_2 \in \mathbb{Z}_+.
\end{aligned}
$$

整数规划，离散优化，组合优化

随机优化

例 1.1.4 某公司生产甲乙两种商品，每件甲商品要耗 A 种原料 3kg，B 种原料 3kg，产值 V_1 随机；每件乙商品要耗 A 种原料 4.5kg，B 种原料 1.5kg，产值 V_2 随机；V_1 和 V_2 服从某种分布 P;. 现公司共有 A 种原料 900kg，B 种原料 600kg，试确定甲乙两种商品共生产多少件，使得公司的总产值最大.

产值为随机变量！

$$\begin{aligned} \max_{x_1, x_2} \quad & \mathbb{E}_P[V_1 x_1 + V_2 x_2] \\ \text{s.t.} \quad & 3x_1 + 4.5x_2 \le 900 \\ & 3x_1 + 1.5x_2 \le 600 \\ & x_1 \ge 0, x_2 \ge 0 \end{aligned}$$

随机优化，鲁棒优化，分布式鲁棒优化

问题转换

- 考虑目标函数 $f(x)$ 极大化问题：

$$\max f(x)(\text{极大化}) \to \min -f(x)(\text{极小化})$$

形如：$c_i(x) \leq 0$ 的不等式约束，可同样转换成上述形式的不等式约束 $-c_i(x) \geq 0$

- $a(x) \leq b(x) + c$ 的不等式约束转化为 $h(x) \geq 0$，可通过令

$$h(x) = b(x) - a(x) + c$$

转换成 $h(x) \geq 0$ 的不等式约束形式。

约束优化问题

- 等式约束优化问题 (equality constraint) 仅含等式约束，模型为：

$$\begin{cases} \min & f(x), \\ \text{s.t.} & c_i(x) = 0, \quad i = 1, 2, \cdots, m, \end{cases}$$

- 不等约束优化问题 (inequality constraint) 仅含不等式约束，模型为：

$$\begin{cases} \min & f(x), \\ \text{s.t.} & c_i(x) \geq 0, \quad i = 1, 2, \cdots, m, \end{cases}$$

- 混合约束优化问题既有等式约束，又有不等式约束。

无约束最优化

无约束最优化问题的数学模型为

$$\min \quad f(x), \qquad x \in R^n, \tag{1.1.8}$$

一般简记为 $\min\ f(x)$.

无约束最优化问题是最优化的基础，很多实际的最优化问题本身就是无约束最优化问题；许多约束最优化方法都是通过变换把约束最优化问题转换成无约束最优化问题后，用适当的无约束优化方法求解。

线性规划

线性规划 (Linear Programming) 的数学模型:

$$
\begin{aligned}
\min \quad & c^\top x, \\
\text{s.t.} \quad & A_1 x = b^1,\ A_2 x \geq b^2,
\end{aligned}
\tag{1.1.11}
$$

其中

- $c = (c_1, c_2, \cdots, c_n)^\top$, $b^1 = (b_1, b_2, \cdots, b_m)^\top$, $b^2 = (b_{m+1}, \cdots, b_p)^\top$,
- 约束矩阵:

$$
A_1 = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, A_2 = \begin{bmatrix} a_{m+1,1} & a_{m+1,2} & \cdots & a_{m+1,n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}.
$$

二次规划

二次规划 (Quadratic Programming) 的数学模型:

$$\begin{aligned} \min \quad & q(x) = \frac{1}{2}x^\top Gx + c^\top x + d, \\ \text{s.t.} \quad & A_1 x = b^1, \\ & A_2 x \geq b^2, \end{aligned} \tag{1.1.12}$$

其中 A_1, A_2 的表示同线性规划模型类似，$c = (c_1, c_2, \cdots, c_n)^\top$，$d$ 为纯量，G 为 $n \times n$ 阶对称矩阵

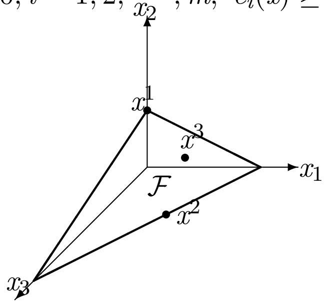
$$G = \begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1n} \\ G_{21} & G_{22} & \cdots & G_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ G_{n1} & G_{n2} & \cdots & G_{nn} \end{bmatrix},$$

满足 $G_{ij} = G_{ji}, \ \forall \ i \neq j.$

可行性

定义：如果点 $x \in R^n$ 满足最优化模型 (1.1.6) 中的所有约束条件就称为可行点 (Feasible Point)。所有可行点的全体称为可行域 (Feasible Region)，

$$\mathcal{F} = \{x \mid c_i(x) = 0, i = 1, 2, \cdots, m, \ \ c_i(x) \geq 0, i = m+1, \cdots, p\}.$$



$$\mathcal{F} = \{x \in R^3 | c_1(x) = 2x_1 + 3x_2 + x_3 - 6 = 0, x_i \geq 0, i = 1, 2, 3\}$$

可行域图示

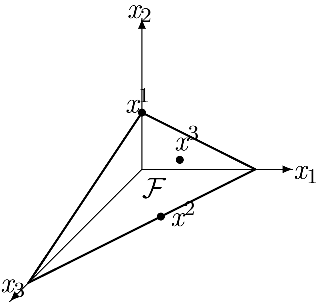


图 1.1.3: 可行域与可行域的边界

注：上图给出了由下述约束条件给出的可行域 $\mathcal{F}$，

$$\mathcal{F} = \{x \in R^3 | c_1(x) = 2x_1 + 3x_2 + x_3 - 6 = 0, x_i \geq 0, i = 1, 2, 3\}$$

对于可行点 x^1，约束 $x_1 \geq 0$ 和 $x_3 \geq 0$ 是有效约束，而 $x_2 \geq 0$ 是无效约束。对于可行点 x^2，则刚好相反，约束 $x_2 \geq 0$ 是有效约束，而 $x_1 \geq 0$ 和 $x_3 \geq 0$ 是无效约束。而对于可行点 x^3，三个不等式约束都是无效约束。图中可行域的边界由粗线表示。

全局最优

定义：一个可行点 $x^* \in \mathcal{F}$ 称为问题 (1.1.6) 的（全局或总体）最优解，如果有

$$f(x^*) \leq f(x), \quad \forall x \in \mathcal{F} \tag{1.1.16}$$

如果上述不等式对所有不同于 x^* 的可行点 x 严格成立，即

$$f(x^*) < f(x), \quad \forall x \in \mathcal{F}, \ x \neq x^* \tag{1.1.17}$$

x^* 称为严格（全局或总体）最优解。

局部最优

定义：对于可行点 x^* 如果存在一个邻域 $\mathcal{N}(x^*) = \{x \mid \|x - x^*\| \leq \delta\}$ 使得成立

$$f(x^*) \leq f(x),\ \forall x \in \mathcal{N}(x^*) \cap \mathcal{F} \tag{1.1.18}$$

则称 x^* 为优化问题 (1.1.6) 的局部最优解，其中 $\delta > 0$ 是一个小的正数，范数 $\|\cdot\|$ 可以是任意向量范数，但一般常用 ℓ_2 范数

$$\|x\| = [\sum_{i=1}^{n} x_i^2]^{\frac{1}{2}}.$$

如果上述不等式 (1.1.18) 对所有 $x \in \mathcal{N}(x^*) \cap \mathcal{F}$，$x \neq x^*$ 严格成立，则称 x^* 为严格局部极小点。

全局与局部最优图示

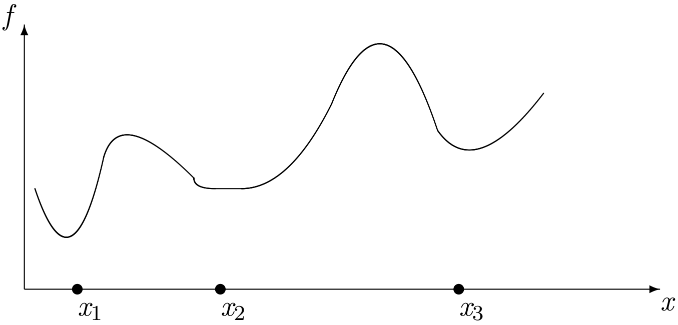
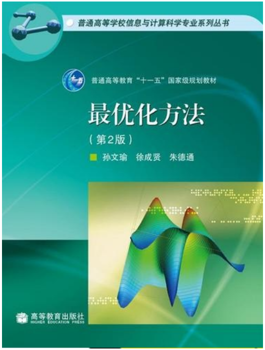


图 1.1.4: 全局极小点与局部极小点

注： x_1, x_2, x_3 均为局部极小点， 其中 x_1 为全局极小点

教材

- 最优化方法，孙文瑜，徐成贤，朱德通著，高等教育出版社，2004



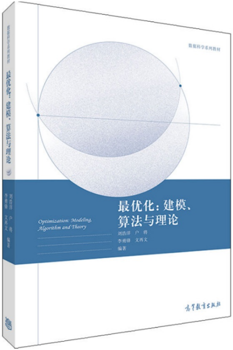
学校网络教材中心有在线电子版 http://jiaocai.lib.xjtu.edu.cn/

参考书籍

- 参考书 1：最优化理论与方法，袁亚湘，孙文瑜，科学出版社
- 参考书 2：数值最优化方法，高立，北京大学出版社

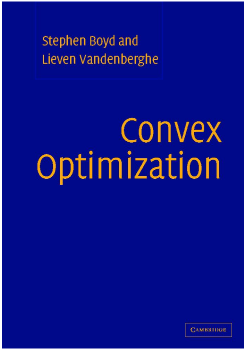
参考书籍

参考书 3：工程优化方法及其应用, 张可村，李换琴 著，2007
参考书 4：最优化：建模、算法与理论, 刘浩洋，户将，李勇锋，文再文 著，2021



参考书籍

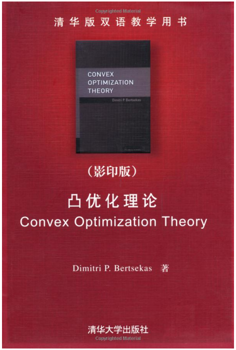
参考书 5：Convex Optimization, Stephen P. Boyd and Lieven Vandenberghe, cambridge university press，2004.

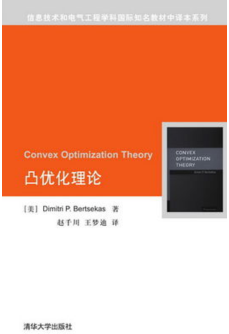


Stephen P. Boyd：斯坦福大学电子工程与信息系统实验室教授

参考书籍

优化理论：凸优化理论 (convex optimization theory)(影印版)，Dimitri P.Bertsekas，清华大学出版社，2011





Dimitri P.Bertsekas：麻省理工电子工程与计算机系教授
中译本：赵千川王梦迪等译. —北京：清华大学出版社，2013

参考书籍

优化理论 2：非线性最优化基础, [日]Masao，Fukushima 著，林贵华译

优化理论 3：变分分析 [Variational Analysis] [美]R.Tyrrell Rockafellar，[美]Roger J-B Wets 著

课程参考资料

- 北大文再文凸优化课件：
  https://bicmr.pku.edu.cn/ wenzw/opt-2021-fall.html
- 斯坦福大学 Boyd 教授和加州大学 Vandenberghe 教授凸优化课件：
  https://web.stanford.edu/ boyd/cvxbook/
- Vandenberghe 教授多门课程课件：
  http://www.seas.ucla.edu/ vandenbe/
- Boyd 课程录像：
  https://www.youtube.com/watch?v=McLq1hEq3UY

课程需要用到的上机软件

- 北大文再文凸优化课件：
  https://bicmr.pku.edu.cn/ wenzw/opt-2021-fall.html
- 斯坦福大学 Boyd 教授和加州大学 Vandenberghe 教授凸优化课件：
  https://web.stanford.edu/ boyd/cvxbook/
- Vandenberghe 教授多门课程课件：
  http://www.seas.ucla.edu/ vandenbe/
- Boyd 课程录像:
  https://www.youtube.com/watch?v=McLq1hEq3UY

最优化（运筹学）学术组织

国际性组织（协会）

- MOS - Mathematical Optimization Society
- INFORMS- Institute for Operations Research and the Management Sciences
- IFORS - The International Federation of Operational Research Societies
- SIAM - Society for Industrial and Applied Mathematics

地区性组织（协会）

- EURO - The Association of European Operational Research Societies
- APORS - Asia-Pacific Operations Research Society
- 中国运筹学会 - Operations Research Society of China

运筹学最高奖项

Dantzig Prize（MOS 和 SIAM 颁发）：数学规划领域最高奖项

- 1982 Michael J. D. Powell（无约束优化，信赖域算法），R. T. Rockafellar（凸优化理论，变分分析）
- 1985 Ellis L. Johnson（组合优化），Manfred Padberg（组合优化）
- 1988 Michael J. Todd（非线性优化，半定规划）
- 1991 Martin Gr otschel（组合优化），Arkadi Nemirovskii （凸优化，内点法）
- 1994 Claude Lemaréchal（非光滑优化，bundle method），Roger Wets（变分分析，随机优化）
- 1997 Stephen M. Robinson（非线性优化，随机优化），Roger Fletcher （无约束优化，拟牛顿法）
- 2000 Yurii Nesterov（凸优化，内点法）

运筹学最高奖项

Dantzig Prize（MOS 和 SIAM 颁发）：数学规划领域最高奖项

- 2003 Jong-Shi Pang(均衡优化，变分不等式)，Alexander Schrijver（离散优化）
- 2006 Éva Tardos （算法复杂度理论，离散优化）
- 2009 Gérard Cornuéljols（图论，组合优化）
- 2012 Jorge Nocedal（非线性优化），Laurence Wolsey（离散优化）
- 2015 Dimitri Bertsekas （动态规划）
- 2018 Andrzej Ruszczyński（随机优化），Alexander Shapiro（随机优化）
- 2021 Michel Goemans（组合优化），H dy Attouch（变分分析，非光滑优化）

John von Neumann Theory Prize（INFORMS 颁发）：运筹领域最高奖项

最优化的会议

最优化国际会议

- International Symposium on Mathematical Programming, MOS 举办的运筹学盛会，优化届三大会议之一，三年一次- 2018, 2022-Beijing
- IFORS Conference of the International Federation of Operational Research Societies，国际运筹协会的大会，三年一次，
- SIAM Conference on Optimization (OP17)，三年一次，优化届三大会议之一 Funding, 2017 ，2020

专业会议

- International Conference on Stochastic Programming 2016，三年一次，随机优化会议，2016 ，2019
- ICCOPT - International Conference on Continuous Optimization，世界连续优化会议，三年一次

最优化的会议

国际运筹学会议

- INFORMS 的会议，运筹学的盛会之一，重头戏是其中的 Annual Meeting 和 Analytics Meeting，主要在美国！
- European Conference on Operational Research，欧洲运筹协会的会议，三年两次

国内会议

- 中国运筹学会第十次全国代表大会，会颁布中国运筹各技术奖项，两年一次，2018,2020
- 中国运筹学会数学规划分会学术年会，两年一次，2019，2021

最优化的学术期刊

数学优化杂志 TOP 3

- Mathematical Programming
- SIAM J. Optimization
- Mathematics of Operations Research

运筹学杂志 TOP2

- Management Science
- Operations Research

最优化的学术期刊

数学优化比较好的杂志

- Journal of Optimization Theory and Applications
- Computational Optimization and Applications
- Journal of Global Optimization
- Journal of the Operations Research Society of China （best in China）

运筹学比较好的杂志

- European Journal of Operational Research
- Computers & Operations Research
- Annals of Operations Research