

第三章 典型优化问题

修贤超

<https://xianchaoxiu.github.io>

- 3.1 线性规划
- 3.2 最小二乘问题
- 3.3 复合优化问题
- 3.4 随机优化问题
- 3.5 半定规划
- 3.6 矩阵优化
- 3.7 优化模型语言

半定规划

- 半定规划 (SDP) 是线性规划在矩阵空间中的一种推广
- 半定规划问题的标准形式

$$\begin{aligned} \min \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle A_1, X \rangle = b_1 \\ & \dots \\ & \langle A_m, X \rangle = b_m \\ & X \succeq 0 \end{aligned}$$

- 对偶形式

$$\begin{aligned} \min \quad & -b^\top y \\ \text{s.t.} \quad & y_1 A_1 + y_2 A_2 + \dots + y_n A_n \preceq C \end{aligned}$$

LP, SOCP 与 SDP 的比较

■ LP 与 SDP

LP	$\min \quad c^\top x$	SDP	$\min \quad c^\top x$
	$\text{s.t.} \quad Ax \leq b$		$\text{s.t.} \quad \text{diag}(Ax - b) \preceq 0$

■ SOCP 与 SDP

SOCP

$$\begin{aligned} \min \quad & f^\top x \\ \text{s.t.} \quad & \|A_i x + b_i\|_2 \leq c^\top x + d_i, \quad i = 1, \dots, m \end{aligned}$$

SDP

$$\begin{aligned} \min \quad & f^\top x \\ \text{s.t.} \quad & \begin{bmatrix} (c_i^\top x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^\top & c_i^\top x + d_i \end{bmatrix} \succeq 0, \quad i = 1, \dots, m \end{aligned}$$

应用举例：二次约束二次规划问题的半定规划松弛

- 设 A_i 为 $n \times n$ 对称矩阵, 考虑二次约束二次规划问题

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^\top A_0 x + 2b_0^\top x + c_0 \\ \text{s.t.} \quad & x^\top A_i x + 2b_i^\top x + c_i \leq 0, \quad i = 1, 2, \dots, m \end{aligned}$$

- 对任意 $x \in \mathbb{R}^n$ 以及 $A \in \mathcal{S}^n$, 有恒等式

$$x^\top A x = \text{Tr}(x^\top A x) = \text{Tr}(A x x^\top) = \langle A, x x^\top \rangle$$

- 原始问题等价于

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \langle A_0, X \rangle + 2b_0^\top x + c_0 \\ \text{s.t.} \quad & \langle A_i, X \rangle + 2b_i^\top x + c_i \leq 0, \quad i = 1, 2, \dots, m \\ & X = x x^\top \end{aligned}$$

应用举例：二次约束二次规划问题的半定规划松弛

■ 进一步地

$$\begin{aligned}x^\top A_i x + 2b_i^\top x + c_i &= \left\langle \begin{pmatrix} A_i & b_i \\ b_i^\top & c_i \end{pmatrix}, \begin{pmatrix} X & x \\ x^\top & 1 \end{pmatrix} \right\rangle \\ &= \langle \overline{A_i}, \overline{X} \rangle, \quad i = 0, 1, \dots, m\end{aligned}$$

■ 半定规划松弛可以写成

$$\begin{aligned}\min \quad & \langle \overline{A_0}, \overline{X} \rangle \\ \text{s.t.} \quad & \langle \overline{A_i}, \overline{X} \rangle \leq 0, \quad i = 1, 2, \dots, m \\ & \overline{X} \succeq 0 \\ & \overline{X}_{n+1, n+1} = 1\end{aligned}$$

■ 约束 $X = xx^\top$ 松弛成半正定约束 $X \succeq xx^\top$ (等价于 $\overline{X} \succeq 0$)

应用举例：最大割问题的半定规划松弛

- 最大割问题是找到节点集合 V 的一个子集 S 使得 S 与它的补集 \bar{S} 之间相连边的权重之和最大化
- 令 $x_j = 1, j \in S$ 和 $x_j = -1, j \in \bar{S}$, 则

$$\begin{aligned} \max \quad & \frac{1}{2} \sum_{i < j} (1 - x_i x_j) w_{ij} \\ \text{s.t.} \quad & x_j \in \{-1, 1\}, j = 1, 2, \dots, n \end{aligned}$$

- 只有当 x_i 与 x_j 不同时, 目标函数中 w_{ij} 的系数非零
- 最大割问题是一个离散优化问题, 很难在多项式时间内找到最优解

应用举例: 最大割问题的半定规划松弛

- 令 $W = (w_{ij}) \in \mathcal{S}^n$, 并定义 $C = -\frac{1}{4}(\text{Diag}(W\mathbf{1}) - W)$, 得到

$$\begin{aligned} \min \quad & x^\top C x \\ \text{s.t.} \quad & x_i^2 = 1, \quad i = 1, 2, \dots, n \end{aligned}$$

- 令 $X = xx^\top$, 则最大割问题可以转化为

$$\begin{aligned} \min \quad & \langle C, X \rangle \\ \text{s.t.} \quad & X_{ii} = 1, \quad i = 1, 2, \dots, n \\ & X \succeq 0, \quad \text{rank}(X) = 1 \end{aligned}$$

□ $x_i^2 = 1$ 意味着矩阵 X 对角线元素 $X_{ii} = 1$

□ $X = xx^\top$ 可以用约束 $X \succeq 0$ 和 $\text{rank}(X) = 1$ 等价刻画

应用举例：极小化最大特征值

- 极小化最大特征值问题可表示为

$$\min \quad \lambda_{\max}(A_0 + \sum_i x_i A_i)$$

- 由于 $\lambda_{\max}(A) \leq t \Leftrightarrow A \preceq tI$ ，则极小化最大特征值可以转化为

SDP 形式

$$\begin{array}{ll} \min & z \\ \text{s.t.} & zI - \sum_i x_i A_i \succeq A_0 \end{array}$$

对偶问题形式

$$\begin{array}{ll} \max & \langle A_0, Y \rangle \\ \text{s.t.} & \langle A_i, Y \rangle = 0 \\ & \langle I, Y \rangle = 1 \\ & Y \succeq 0 \end{array}$$

应用举例: 极小化二范数问题

- 令 $A_i \in \mathbb{R}^{m \times n}$, 极小化 $A(x) = A_0 + \sum_i x_i A_i$ 的二范数

$$\min_x \|A(x)\|_2$$

- SDP 形式

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & \begin{pmatrix} tI & A(x) \\ A(x)^\top & tI \end{pmatrix} \succeq 0 \end{aligned}$$

- 约束形式来源于

$$\begin{aligned} \|A\|_2 \leq t &\Leftrightarrow A^\top A \preceq t^2 I, \quad t \geq 0 \\ &\Leftrightarrow \begin{pmatrix} tI & A(x) \\ A(x)^\top & tI \end{pmatrix} \succeq 0 \end{aligned}$$

- 3.1 线性规划
- 3.2 最小二乘问题
- 3.3 复合优化问题
- 3.4 随机优化问题
- 3.5 半定规划
- 3.6 矩阵优化
- 3.7 优化模型语言

矩阵优化的基本形式

■ 矩阵优化问题的形式

$$\min_{X \in \mathcal{X}} \psi(X)$$

□ \mathcal{X} 为特定的矩阵空间

□ $\psi(X) : \mathcal{X} \rightarrow \mathbb{R}$ 为给定的函数，可能是非光滑的

■ 和向量相比，矩阵有许多新的性质，如秩、特征值等

■ 广泛地出现在组合数学、材料科学、机器学习和统计学等

矩阵优化的基本形式

■ 低秩矩阵恢复问题

$$\min_{X \in \mathbb{R}^{m \times n}} \quad \frac{1}{2} \|X - M\|_F^2 + \mu \|X\|_*$$

考虑函数 $h(X) = \|X\|_*$ 的次微分

$$\partial h(X) = \{UV^\top + W \mid \|W\|_2 \leq 1, U^\top W = 0, WV = 0\}$$

■ 主成分分析问题

$$\min_{X \in \mathbb{R}^{p \times d}} \quad \psi(X) = -\text{Tr}(X^\top A A^\top) \quad \text{s.t.} \quad X^\top X = I_d$$

考虑目标函数的微分

$$\nabla \psi(X) = -2AA^\top X$$

应用举例：非负矩阵分解

- 给定矩阵 $A = [a_1, a_2, \dots, a_n] \in \mathbb{R}^{d \times n}$, 将其分解成非负基矩阵 $X \in \mathbb{R}^{d \times p}$ 和非负系数矩阵 $Y \in \mathbb{R}^{p \times n}$ 的乘积, 即

$$A = XY$$

- 由于观测含有噪声, 原始数据矩阵 A 和分解 XY 不会完全吻合, 应考虑

$$\begin{aligned} \min_{X \in \mathbb{R}^{d \times p}, Y \in \mathbb{R}^{p \times n}} \quad & \frac{1}{2} \|A - XY\|_F^2 \\ \text{s.t.} \quad & X \geq 0, Y \geq 0 \end{aligned}$$

- 本质上是将高维空间中的数据在一个低维空间中表示
- 和主成分分析模型类似, 但会得到比主成分分析模型更有实际意义的解

应用举例：非负矩阵分解

- 根据具体应用的不同，还可以考虑带正则项的非负矩阵分解模型

$$\begin{aligned} \min_{X \in \mathbb{R}^{d \times p}, Y \in \mathbb{R}^{p \times n}} \quad & \frac{1}{2} \|A - XY\|_F^2 + \alpha r_1(X) + \beta r_2(Y) \\ \text{s.t.} \quad & X \geq 0, \quad Y \geq 0 \end{aligned}$$

- $r_1(X)$ 和 $r_2(Y)$ 是正则项
- $\alpha, \beta > 0$ 是用来权衡拟合项和正则项的正则化参数
- 如果基向量的线性无关性，取 $r_1(X) = \|X^\top X - I\|_F^2$
- 如果每一个观测值都可以用少数几个基向量来表示，取 $r_2(Y) = \|Y\|_1$

Algorithms for non-negative matrix factorization

D Lee, HS Seung - Advances in neural information ..., 2000 - proceedings

... nonnegativity is a useful constraint for matrix factorization that can learn data [4, 5]. The nonnegative ... for learning the optimal nonnegative factor

☆ 保存 ↯ 引用 被引用次数: 12260 相关文章 所有 36 个版本 ↯↯

Learning the parts of objects by non-negative matrix fact

DD Lee, HS Seung - nature, 1999 - nature.com

... an algorithm for non-negative matrix factorization that is able to ... Non- distinguished from the ... When non-negative matrix factorization is imple

☆ 保存 ↯ 引用 被引用次数: 16498 相关文章 所有 17 个版本

应用举例：相关系数矩阵估计

- 给定对称矩阵 $C \in \mathcal{S}^n$ 和非负对称权重矩阵 $H \in \mathcal{S}^n$ ，求解一个秩小于等于 p 的相关系数矩阵 X ，使得在结合了权重矩阵的某种度量下最小化

$$\begin{aligned} \min_{X \succeq 0} \quad & \frac{1}{2} \|H \odot (X - C)\|_F^2 \\ \text{s.t.} \quad & X_{ii} = 1, \quad i = 1, 2, \dots, n \\ & \text{rank}(X) \leq p \end{aligned}$$

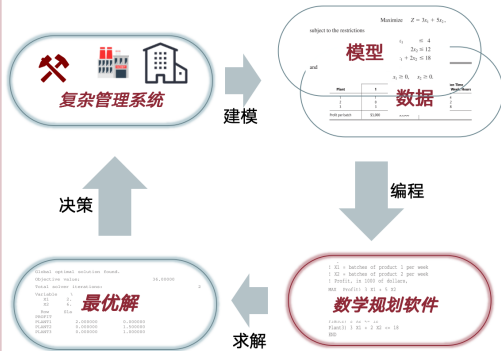
- 将 $\text{rank}(X) \leq p$ 表示为 $X = V^\top V$ ，其中 $V = [V_1, V_2, \dots, V_n] \in \mathbb{R}^{p \times n}$ ，得到

$$\begin{aligned} \min_{V \in \mathbb{R}^{p \times n}} \quad & \frac{1}{2} \|H \odot (V^\top V - C)\|_F^2 \\ \text{s.t.} \quad & \|V_i\|_2 = 1, \quad i = 1, 2, \dots, n \end{aligned}$$

- 3.1 线性规划
- 3.2 最小二乘问题
- 3.3 复合优化问题
- 3.4 随机优化问题
- 3.5 半定规划
- 3.6 矩阵优化
- 3.7 优化模型语言

优化软件发展

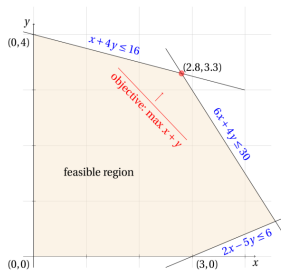
- 现代复杂管理系统决策的主要特点：数据驱动，规模超大，多因素制约，多目标优化
- 数据驱动的决策流程：数据-->规律-->决策
- 数学规划：基于数据驱动的决策，问题建模与求解的核心环节



智能决策中两个经典的关键科学问题：

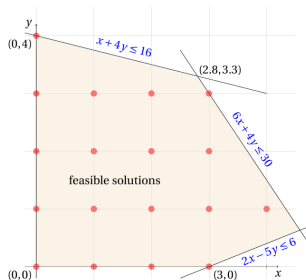
- 1, 如何建模：将复杂决策问题翻译成简洁、有效的数学表达形式
- 2, 如何求解：快速、准确地求解数学模型

优化软件发展



线性规划

- 问题目标：在可行域寻找最优解
- 规模较大：常有数百万变量和约束
- 稀疏性质：大量的稀疏矩阵运算
- 数值问题：受计算机精度影响



整数规划

- 可行域是离散的点
- 是经典的“NP-完全”问题
- 复杂多样的问题结构
- 需要求解多个线性规划问题

二阶锥规划

凸二次规划

凸二次约束规划

.....

优化软件发展

数学理论门槛高

- 理论储备不足，基本理论框架由西方发展起来
- 求解计算模块的相关理论在国内基本空白状态
- 国内理论研究主要聚焦在数学规划在各个领域的定制化应用

软件工程难度大

- 数学规划的软件工程量巨大
- 线性5万行，整数规划百万行
- 相关的系统开发的理论研究国内极少
- 参照MATLAB

领域知识积累难

- Domain Knowledge欠缺
- 例：热启动的30种备选该如何从数千篇文献选取？
- 欧美积累了40多年
- 论文里不会提及，国外求解器也不会公布的不传之秘
- 只能自己一步步摸索，不断试错

技术创新迭代快

- 例如Gurobi研发组每年需要看3到500篇论文，但是只有5到7个技巧显著有用
- 需要自己观察与发展新的理论研究
- 更需要与新一代人工智能算法结合，软硬件结合，寻求“弯道超车”可能性

困难：投入大，周期长，工作很多时候是对欧美的追赶，很难产生论文

优化软件发展

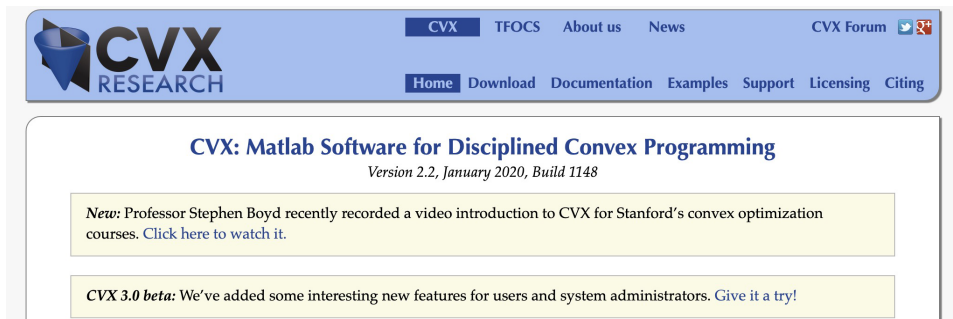
世界著名数学规划求解器研发历程

- 1939年, 苏联诺贝尔经济学奖获得者Leonid Kantorovich发明线性规划
- 1979年芝加哥大学的Charnes 发布Lingo
- 1983年英国爱丁堡大学Ashford创办了XPRESS
- 1987年莱斯大学Bixby创办了CPLEX公司
- 2000年COIN-OR成立, 并发布CLP和CBC
- 2005年德国ZIB发布了开源整数规划SCIP
- 2008年Cplex团队Bixby等离职创办GUROBI
- 2017年, 上财发布中国第一个开源数学规划求解器LEAVES, 2018年中科院 CMIP
- 2019年, 杉数科技发布中国第一个专业数学规划求解器COPT, 此后阿里, 华为等纷纷入局
- 2021年至今, 谷歌, ORACLE, 微软纷纷开始组建自己的数学规划求解器团队

世界著名数学规划求解器一览表

GUROBI	美国与英国三大求解器巨头，累计三十 年以上研发历史和95%以上市场。线性， 整数，非线性各个模块功能齐备				
CPLEX					
XPRESS					
SAS	←世界最大商业统计软件（北卡）				
CVX	←世界著名求解器建模平台(斯坦福大学)				
IPOPT	←著名非线性规划开源软件(卡耐基梅隆)				
Coin-OR	←世界最好开源线性规划（多组织维护）				
Baron	←世界最好非线性规划（多组织维护）				
NEOS	SCIP	←世界最好整数规划(ZIB)			OPTV 华为
CBC	SOPLEX	MOSEK	GLPK	LEAVES 上财	CMIP 中科院
美国	德国	丹麦	俄罗斯	中国	
世界最大免费服务平台（威斯康星大学）			世界最好锥规划（MOSEK）		
世界最好开源组织（多组织维护）					

- CVX 以 MATLAB 为基础的优化模型语言，求解凸优化问题
 - 快速构造和识别凸性
 - 调用已有软件包求解变形后的凸优化问题
 - 包括免费软件 SDPT3 和 SeDuMi 以及商业软件 Gurobi 和 MOSEK 等



The screenshot shows the CVX Research website. The header is blue with the CVX Research logo on the left. Navigation links include CVX, TFOCS, About us, News, CVX Forum, Home, Download, Documentation, Examples, Support, Licensing, and Citing. The main content area has a title 'CVX: Matlab Software for Disciplined Convex Programming' and a subtitle 'Version 2.2, January 2020, Build 1148'. Below this, there are two yellow boxes with news updates. The first box mentions Professor Stephen Boyd's video introduction to CVX for Stanford's convex optimization courses. The second box mentions CVX 3.0 beta and its new features for users and system administrators.

CVX TFOCS About us News CVX Forum

Home Download Documentation Examples Support Licensing Citing

CVX: Matlab Software for Disciplined Convex Programming
Version 2.2, January 2020, Build 1148

New: Professor Stephen Boyd recently recorded a video introduction to CVX for Stanford's convex optimization courses. [Click here to watch it.](#)

CVX 3.0 beta: We've added some interesting new features for users and system administrators. [Give it a try!](#)

■ 考虑优化问题

$$\begin{array}{ll}\min & \|Ax - b\|_2 \\ \text{s.t.} & Cx = d \\ & \|x\|_\infty \leq e\end{array}$$

```
1 m = 20; n = 10; p = 4;
2 A = randn(m,n); b = randn(m,1);
3 C = randn(p,n); d = randn(p,1); e = rand;
4 cvx_begin
5     variable x(n)
6     minimize( norm( A * x - b, 2 ) )
7     subject to
8         C * x == d
9         norm( x, Inf ) <= e
10 cvx_end
```


Q&A

Thank you!

感谢您的聆听和反馈