第 2 章：数学基础

(4)：凸优化，以及 CVXPY 和 SciPy 的介绍

授课教师：曹语

课程主页：`https://yucaoyc.github.io/math3806`

背景和目标

我们已经了解了凸集和凸函数

凸优化问题的**直观概念**可大致描述为：在凸的可行域 Ω 内，极小化一个凸函数 f_0。

在本节课，我们将

- 系统化凸优化的一些基本知识
- 介绍一些新例子
- 通过这些例子，介绍 CVXPY，以及 SciPy 的使用方法

练习： $f(\boldsymbol{x}) = e^{\boldsymbol{a}^\top \boldsymbol{x} + b}$ 是否是一个凸函数?

【答案见课内】

目录

凸优化的形式

凸优化问题的基本形式如下：

$$\begin{array}{ll} \text{minimize} & f_0(\boldsymbol{x}) \\ \text{subject to} & f_i(\boldsymbol{x}) \leqslant 0, \quad i = 1, \cdots, m \\ & \boldsymbol{a}_i^\top \boldsymbol{x} - b_i = 0, \quad i = 1, \cdots, p \end{array}$$

此外，我们**要求**：

- 目标函数 f_0 是凸函数；
- 不等式约束函数 f_i $(i = 1, 2, \cdots, m)$ 是凸函数；
- 等式约束函数 $h_i(\boldsymbol{x}) = \boldsymbol{a}_i^\top \boldsymbol{x} - b_i$ 是仿射的，且 $\boldsymbol{a}_i \neq \boldsymbol{0}$

$\Longrightarrow$ 由于凸函数 f_i 的下水平集是凸集，且仿射集也是凸集，所以可行域必然是一个凸集！

对于在一般的凸集上优化凸函数，

- 在这个课本中，我们暂时不一定认定它是凸优化
- 它被称为**抽象的凸优化问题**

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{x}) = x_1^2 + x_2^2 \\ \text{subject to} \quad & f_1(\boldsymbol{x}) = \tfrac{x_1}{1+x_2^2} \leqslant 0 \\ & h_1(\boldsymbol{x}) = (x_1 + x_2)^2 = 0 \end{aligned}$$

不是标准凸优化问题

因为 f_1 不是凸的、h_1 不是仿射（尽管可行域是凸集）

属于“抽象的”凸优化问题

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{x}) = x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1 \leqslant 0 \\ & x_1 + x_2 = 0 \end{aligned}$$

是标准凸优化问题

（由左边等价转化而来）

重要定理 (I)

定理

对于凸优化问题，任意局部最优解也是全局最优解。

证明：见课本 4.2 章，或见课堂

重要定理 (II)

定理 (可微函数 f_0 的最优性准则)

若 f_0 可微，则 $\boldsymbol{x} \in \Omega$ 是（全局）最优解，当且仅当

$$\nabla f_0(\boldsymbol{x})^\top (\boldsymbol{y} - \boldsymbol{x}) \geqslant 0, \qquad \forall \boldsymbol{y} \in \Omega \tag{1}$$

对于无约束问题，最优性条件变成

$$\nabla f_0(\boldsymbol{x}) = \mathbf{0} \tag{2}$$

证明：见课本 4.2 章，或见课堂

思路：考虑 $g(t) = f_0\Big(\boldsymbol{x} + t(\boldsymbol{y} - \boldsymbol{x})\Big)$

例子 1：最小化二次函数

考虑在无约束条件下极小化二次函数：

$$f_0(\boldsymbol{x}) = (1/2)\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + r$$

其中，$\boldsymbol{P}$ 是半正定矩阵（保证曲面开口向上或平的）。

根据极值条件，令梯度为零：

$$\nabla f_0(\boldsymbol{x}) = \mathbf{0} \qquad \Longrightarrow \qquad \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q} = \mathbf{0}$$

- 若 $\boldsymbol{P}$ 为正定矩阵，则唯一最优解为 $\boldsymbol{x}^* = -\boldsymbol{P}^{-1}\boldsymbol{q}$
- 但如果 $\boldsymbol{P}$ 是半正定矩阵，可能存在零特征值，此时 $\boldsymbol{P}$ 不可逆，无法通过上述公式求解最优解，那么解是什么呢?

具体例子：P 存在零特征值的情况

考虑在 $\mathbb{R}^2$ 中的无约束二次规划，设半正定矩阵 $\boldsymbol{P} = \left[\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right]$。
此时 $\mathrm{Range}(\boldsymbol{P})$ 仅在 x_1 轴方向上。目标函数为 $f_0(\boldsymbol{x}) = \frac{1}{2}x_1^2 + q_1x_1 + q_2x_2$：

- **情况一：$\boldsymbol{q} \notin \mathbf{Range}(\boldsymbol{P})$**
  例如 $\boldsymbol{q} = \left[\begin{smallmatrix}0\\ -1\end{smallmatrix}\right]$。目标函数变为 $f_0(\boldsymbol{x}) = \frac{1}{2}x_1^2 - x_2$。
  当 $x_2 \to +\infty$ 时，$f_0(\boldsymbol{x}) \to -\infty$，即极小值无下界，问题无解。
- **情况二：$\boldsymbol{q} \in \mathbf{Range}(\boldsymbol{P})$**
  例如 $\boldsymbol{q} = \left[\begin{smallmatrix}-1\\ 0\end{smallmatrix}\right]$。目标函数变为 $f_0(\boldsymbol{x}) = \frac{1}{2}x_1^2 - x_1$。
  当 $x_1 = 1$ 时达到最小值 $-\frac{1}{2}$。由于 x_2 在原函数中不产生影响且可以任意取值，因此存在无穷多个解：$\boldsymbol{x}^\star = \left[\begin{smallmatrix}1\\ c\end{smallmatrix}\right], \forall c \in \mathbb{R}$。

例子 2：只含等式约束问题

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

假定定义域非空，且 f_0 可微。令最优解为 x^*。

该优化问题通过方程(1)可以转变成：存在 $\boldsymbol{\nu}$ 使得

$$\begin{aligned} \nabla f_0(\boldsymbol{x}^*) + \boldsymbol{A}^\top \boldsymbol{\nu} &= \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x}^* &= \boldsymbol{b} \end{aligned}$$

该形式就是拉格朗日（Lagrange）乘子最优性条件（后面章节会再涉及）。

目录

CVXPY

CVXPY

CVXPY 是一个强大的基于 Python 的解决凸优化问题的开源工具包。它允许用户以符合数学直觉的形式表达问题。

安装方法：`pip install cvxpy`

基本的使用方法见随堂代码部分

- 官方快速入门教程：
  `https://www.cvxpy.org/tutorial/intro/index.html`
- Python 封装函数文档：
  `https://www.cvxpy.org/tutorial/functions/index.html`

SciPy 的优化模块

scipy.optimize

SciPy 是 Python 的一个强大的科学计算库，其中包含的优化模块 (`scipy.optimize`)，提供了许多优化算法的实现框架，包括：

- 局部极小值求解
- 全局优化
- 线性规划
- 最小二乘法与曲线拟合等

【更多演示见随堂 Jupyter Notebook 代码】

目录

线性规划的一般形式

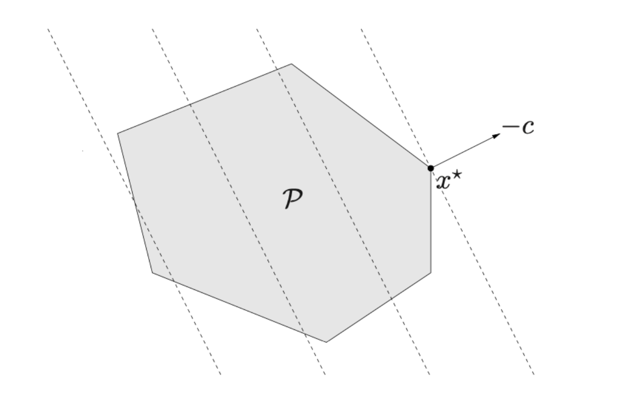
线性规划 (Linear Programming LP)：目标函数和约束函数都是仿射

$$
\begin{array}{rll}
\text{minimize} & \boldsymbol{c}^\top \boldsymbol{x} + d & \\
\text{subject to} & \boldsymbol{G}\boldsymbol{x} \preceq \boldsymbol{h} & (\text{不等式约束}) \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} & (\text{等式约束})
\end{array} \tag{3}
$$

其中 $\boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{G} \in \mathbb{R}^{m\times n}$, $\boldsymbol{h} \in \mathbb{R}^m$, $\boldsymbol{A} \in \mathbb{R}^{p\times n}$, $\boldsymbol{b} \in \mathbb{R}^p$, $d \in \mathbb{R}$

练习：线性规划的可行集的几何名字叫什么？

线性规划的几何示意图



标准形式线性规划：

$$
\begin{aligned}
\text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\
\text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \succeq \boldsymbol{0}
\end{aligned} \tag{4}
$$

引理

任意线性规划问题，即公式(3)，可以等价转变为标准形式，即公式(4)。

例子 1：多面体的 Chebyshev 中心

对于如下的多面体，求内接于其中的最大 Euclid 球的问题；最优球的中心被称为**多面体的 Chebyshev 中心**

- 多面体：

$$\mathcal{P} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}_i^\top \boldsymbol{x} \leqslant b_i, i = 1, \cdots, m \right\}$$

- Euclid 球：

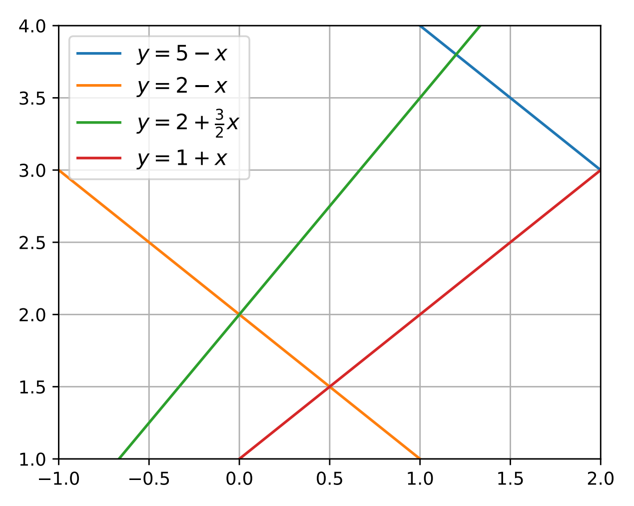
$$\mathcal{B} = \{ \boldsymbol{x}_c + \boldsymbol{u} \mid \|\boldsymbol{u}\|_2 \leqslant r \}$$

该问题可以被改写成如下线性规划问题：

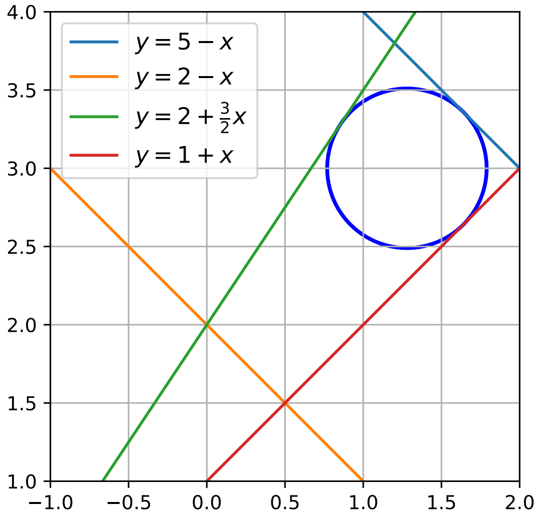
$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & \boldsymbol{a}_i^\top \boldsymbol{x}_c + r \left\|\boldsymbol{a}_i\right\|_2 \leqslant b_i, \quad i = 1, \cdots, m \end{array}$$

具体例子

练习： 请将如下的四条线构成的多面体写成标准形式



优化结果



目录

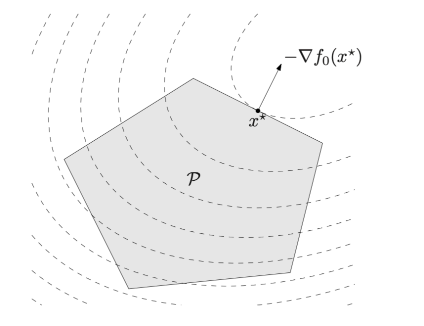
二次规划 (Quadratic Programming, QP)

是指具有如下形式的优化问题：

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^\top \boldsymbol{x} + r \\
\text{subject to} & \boldsymbol{G}\boldsymbol{x} \preceq \boldsymbol{h} \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}
\end{array}
$$

其中 $\boldsymbol{P} \in \mathbf{S}_+^n$ 是半正定矩阵

几何示意图：



二次约束二次规划（QCQP）

$$
\begin{aligned}
\text{minimize} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_0 \boldsymbol{x} + \boldsymbol{q}_0^\top \boldsymbol{x} + r_0 \\
\text{subject to} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}_i \boldsymbol{x} + \boldsymbol{q}_i^\top \boldsymbol{x} + r_i \leqslant 0, \quad i = 1, \cdots, m \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}
\end{aligned}
$$

其中 $\boldsymbol{P}_i \in \mathbf{S}_+^n$ 是半正定矩阵

例子 1：最小二乘问题

一般的**最小二乘问题（常见于线性回归等问题）**

$$
\begin{aligned}
\min_{\boldsymbol{c}} \ f_0(\boldsymbol{c}) = \|\boldsymbol{A}\boldsymbol{c}-\boldsymbol{b}\|^2 &= (\boldsymbol{A}\boldsymbol{c}-\boldsymbol{b})^\top(\boldsymbol{A}\boldsymbol{c}-\boldsymbol{b}) \\
&= \boldsymbol{c}^\top\left(\boldsymbol{A}^\top\boldsymbol{A}\right)\boldsymbol{c} - 2\boldsymbol{b}^\top\boldsymbol{A}\boldsymbol{c} + \boldsymbol{b}^\top\boldsymbol{b}
\end{aligned}
$$

其中 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 是矩阵，$\boldsymbol{c} \in \mathbb{R}^n$ 是优化变量，$\boldsymbol{b} \in \mathbb{R}^m$ 是给定的数据

该例子就是一个**（无约束的）二次规划**，其中 $\boldsymbol{P} = 2\boldsymbol{A}^\top\boldsymbol{A}$, $\boldsymbol{q} = -2\boldsymbol{A}^\top\boldsymbol{b}$。

例子 2：多面体间距离

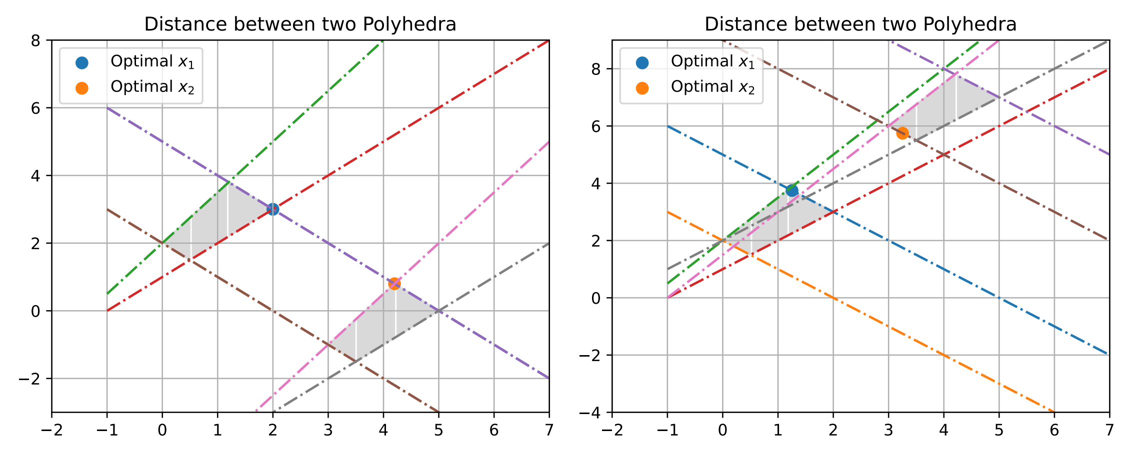
考虑两个多面体

$$\mathcal{P}_1 = \{\boldsymbol{x} : \ \boldsymbol{A}_1\boldsymbol{x} \preceq \boldsymbol{b}_1\}$$
$$\mathcal{P}_2 = \{\boldsymbol{x} : \ \boldsymbol{A}_2\boldsymbol{x} \preceq \boldsymbol{b}_2\}$$

优化最小距离可以被改写成如下的二次规划问题：

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2^2 \\ \text{subject to} & \boldsymbol{A}_1\boldsymbol{x}_1 \preceq \boldsymbol{b}_1, \quad \boldsymbol{A}_2\boldsymbol{x}_2 \preceq \boldsymbol{b}_2 \end{array}$$

优化结果



Distance between two Polyhedra
Optimal x_1
Optimal x_2
Distance between two Polyhedra
Optimal x_1
Optimal x_2

例子 3：Markowitz 投资组合优化

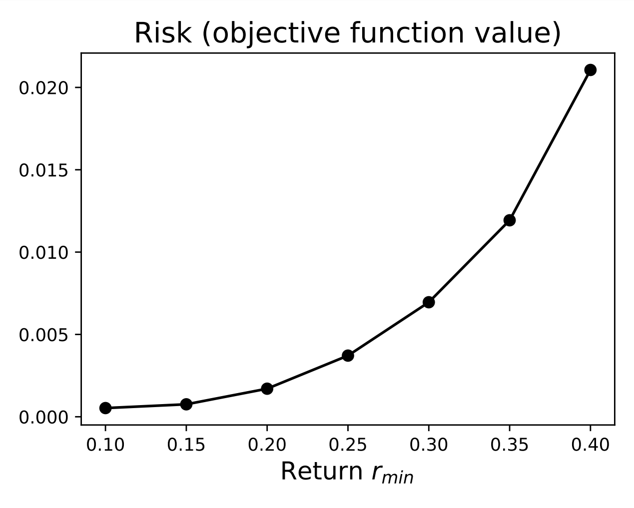
- x_i 表示持有资产 i 的数量
  - 资产 i 的多头对应 $x_i > 0$
  - 资产 i 的空头对应 $x_i < 0$
- p_i 表示资产相对价格变动；$\boldsymbol{p} \in \mathbb{R}^n$ 为随机变量，均值为 $\bar{\boldsymbol{p}}$ ，协方差为 Σ，并且假设已知 $\bar{\boldsymbol{p}}$ 和 Σ
- 投资总回报为 $r = \boldsymbol{p}^\top \boldsymbol{x}$

Markowitz 引入的经典的投资组合优化问题：

$$\begin{array}{ll} \text{minimize} & \boldsymbol{x}^\top \Sigma \, \boldsymbol{x} \\ \text{subject to} & \bar{\boldsymbol{p}}^\top \boldsymbol{x} \geqslant r_{\min} \\ & \boldsymbol{1}^\top \boldsymbol{x} = 1, \quad \boldsymbol{x} \succeq 0 \end{array}$$

（注：此模型加上 $\boldsymbol{x} \succeq 0$ 约束表示不允许做空）

优化结果


Risk (objective function value)
0.020
0.015
0.010
0.005
0.000
0.10
0.15
0.20
0.25
0.30
0.35
0.40
Return r_{min}

目录

几何规划

单项式：对于函数 $f:\Omega\to\mathbb{R}$, 定义域 $\Omega=\mathbb{R}^n_{++}$，且满足

$$f(\boldsymbol{x})=cx_1^{a_1}x_2^{a_2}\cdots x_n^{a_n},\qquad c>0, a_i\in\mathbb{R}$$

正项式：单项式的和

$$f(\boldsymbol{x})=\sum_{k=1}^{K}c_kx_1^{a_{1k}}x_2^{a_{2k}}\cdots x_n^{a_{nk}},\qquad c_k>0$$

几何规划 (GP)

$$\begin{array}{ll}\text{minimize} & f_0(\boldsymbol{x})\\ \text{subject to} & f_i(\boldsymbol{x})\leqslant 1,\quad i=1,\cdots,m\\ & h_i(\boldsymbol{x})=1,\quad i=1,\cdots,p\end{array}$$

$f_0,\cdots,f_m$ 为正项式，$h_1,\cdots,h_p$ 为单项式；定义域 $\Omega=\mathbb{R}^n_{++}$ 隐含约束 $\boldsymbol{x}\succ 0$

凸形式的几何规划

令 $y_i = \log x_i$, 则上述问题可以等价为:

$$
\begin{aligned}
\text{minimize} \quad & \tilde{f}_0(\boldsymbol{y}) = \log\left(\textstyle\sum_{k=1}^{K_0} e^{\boldsymbol{a}_{0k}^\top \boldsymbol{y} + b_{0k}}\right) \\
\text{subject to} \quad & \tilde{f}_i(\boldsymbol{y}) = \log\left(\textstyle\sum_{k=1}^{K_i} e^{\boldsymbol{a}_{ik}^\top \boldsymbol{y} + b_{ik}}\right) \leqslant 0, \quad i = 1, \cdots, m \\
& \tilde{h}_i(\boldsymbol{y}) = \boldsymbol{w}_i^\top \boldsymbol{y} + c_i = 0, \quad i = 1, \cdots, p
\end{aligned}
$$

由于 $\tilde{f}_i$ 是凸函数, $\tilde{h}_i$ 是仿射函数，因此该问题是一个凸优化问题。

注：前面练习已经通过 Hessian 矩阵说明了 $e^{\tilde{f}_i}$ 是凸函数；而证明 $\tilde{f}_i$ 本身是凸函数将作为课后作业。

例子 1：最小化表面积

假设我们需要设计某个箱子，其体积固定为 V；我们希望最小化材料（即该盒子的表面积）。优化问题可以设置如下：

$$\begin{aligned} \text{minimize} \quad & 2\big(x_1x_2 + x_1x_3 + x_2x_3\big) \\ \text{subject to} \quad & \boldsymbol{x} \succ \mathbf{0} \quad \text{(长宽高必须为正)} \\ & x_1x_2x_3 = V \end{aligned}$$

该问题可以等价变成凸优化问题：

$$\begin{aligned} \text{minimize} \quad & \log\left(2e^{y_1+y_2} + 2e^{y_1+y_3} + 2e^{y_2+y_3}\right) \\ \text{subject to} \quad & y_1 + y_2 + y_3 = \log(V) \end{aligned}$$

目录

总结

- 能够复述凸优化问题的定义
- 能够复述出两个重要定理的内容
- 能写出简单的优化代码
- 对于线性规划问题、二次规划问题、几何规划问题这三类问题，掌握一般形式、以及了解上述具体例子

阅读作业 & 参考资料：

- 课本第 4.1 - 4.5 章