第 2 章：数学基础

(4)：凸优化，以及 CVXPY 的介绍

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**大致**可被描述为：定义域 Ω 为凸集，f_0 为凸函数

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 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 $\boldsymbol{a}_i \neq \mathbf{0}$

$\Longrightarrow$ 由于函数 $f_1, \cdots, f_m, h_1, \cdots, h_p$ 皆是凸函数，所以
定义域 Ω 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**

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

不是凸优化问题，因为 f_1 不是凸的、h_1 不是仿射

该问题属于抽象的凸优化问题

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

是凸优化问题

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

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

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。

最优性条件为

$$\nabla f_0(\boldsymbol{x}) = \mathbf{0} \qquad \rightarrow \qquad \underline{\qquad\qquad ? \qquad\qquad}$$

- 若 $\boldsymbol{P}$ 为正定矩阵，则唯一最优解为 $\boldsymbol{x}^* = -\boldsymbol{P}^{-1}\boldsymbol{q}$
- 若 $\boldsymbol{P}$ 的特征值可以为 0（具有 α 个重根），则通过特征分解

$$\boldsymbol{P} = \boldsymbol{QDQ}^{-1}$$

其中 $\boldsymbol{Q}$ 是正交矩阵，$\boldsymbol{D}$ 是含特征值的对角矩阵，$D_{ii} = 0$ $(i = 1, 2, \cdots, \alpha)$。上述优化问题等价为

$$\min_{\boldsymbol{y}} \ \frac{1}{2}\boldsymbol{y}^\top \boldsymbol{D}\boldsymbol{y} + (\boldsymbol{Q}^{-1}\boldsymbol{q})^\top \boldsymbol{y} + r$$

  - 若 $\boldsymbol{Q}^{-1}\boldsymbol{q}$ 的前 α 个分量有非 0 项，优化问题无解（最小值无下界）
    等价于说 $\boldsymbol{q}$ 不在 $\mathrm{Range}(\boldsymbol{P})$
  - 若 $\boldsymbol{Q}^{-1}\boldsymbol{q}$ 的前 α 个分量都为 0，优化问题有无穷多个解
    等价于说 $\boldsymbol{q}$ 在 $\mathrm{Range}(\boldsymbol{P})$

例子 2：只含等式约束问题

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

假定定义域非空，且 f_0 可微。

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

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

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 是一个基于 Python 的解决凸优化问题的数值包。

安装方法：pip3 install cvxpy

基本的使用方法见代码部分

- 简单的例子:
  `https://www.cvxpy.org/tutorial/intro/index.html`
- 封装好的函数:
  `https://www.cvxpy.org/tutorial/functions/index.html`

目录

线性规划的一般形式

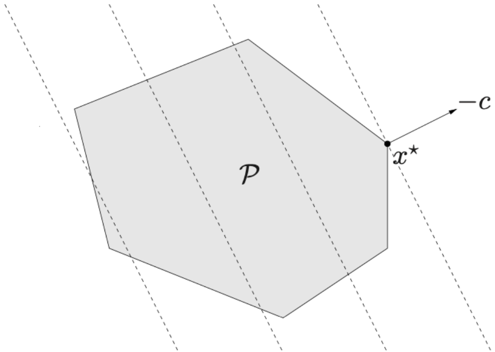
线性规划 (Linear Programming LP)：目标函数和约束函数都是仿射

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

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：

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

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，找到最大的 Euclid 球的问题；最优球的中心被称为**多面体的 Chebyshev 中心**

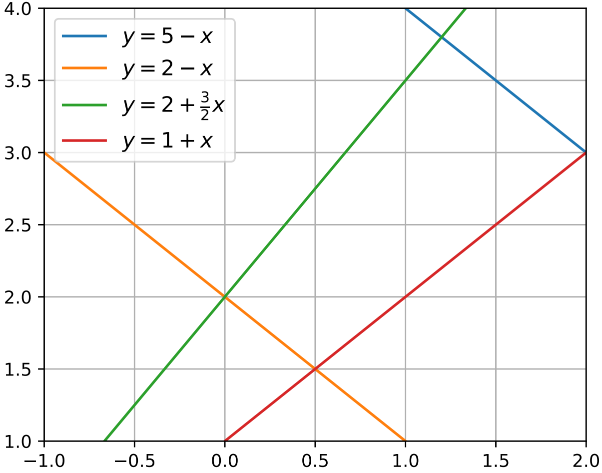
- 多面体：
$$\mathcal{P} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{a}_i^\top \boldsymbol{x} \leqslant b_i, i = 1, \cdots, m \right\}$$
- Euclid 球：
$$\mathcal{B} = \{ \boldsymbol{x}_c + \boldsymbol{u} \mid \|\boldsymbol{u}\|_2 \leqslant r \}$$

该问题可以被改写成如下线性规划问题：

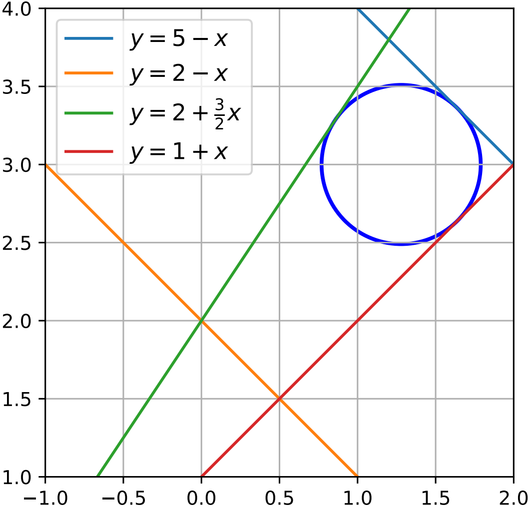
$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & \boldsymbol{a}_i^\top \boldsymbol{x}_c + r \|\boldsymbol{a}_i\|_2 \leqslant b_i, \quad i = 1, \cdots, m \end{array}$$

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

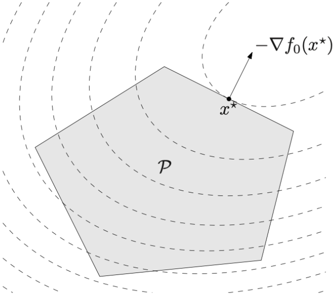
二次规划 (Quadratic Programming, QP)

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

几何示意图：



二次约束二次规划（QCQP）

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

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（又名回归分析）

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

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

练习：请将作业 1 的第一题通过 cvxpy 求解；可下载练习部分的代码

例子 2：多面体间距离

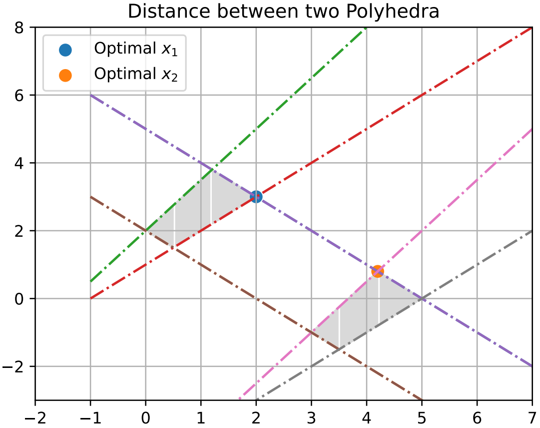
考虑两个多面体

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

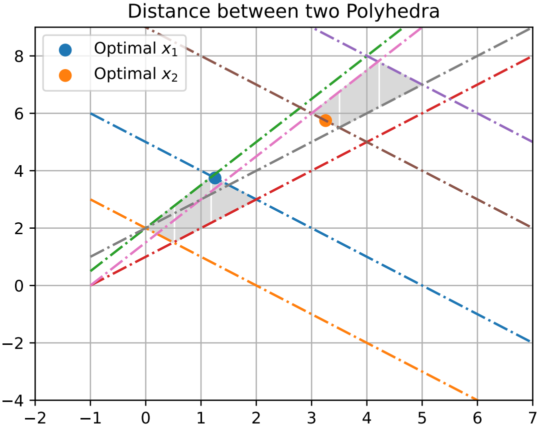
优化最小距离可以被改写成如下的二次规划问题：

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

优化结果



Distance between two Polyhedra
Optimal x_1
Optimal x_2



Distance between two Polyhedra
Optimal x_1
Optimal x_2

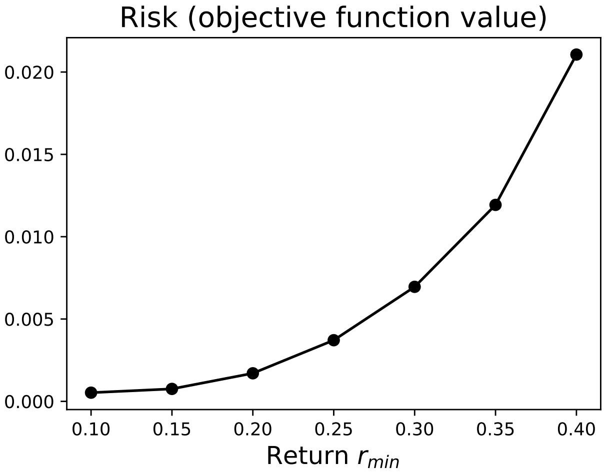
例子 3：Markowitz 投资组合优化

- 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：

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

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

$$f(\boldsymbol{x}) = c x_1^{a_1} x_2^{a_2}\cdots x_n^{a_n}, \qquad c>0,\ a_i\in\mathbb{R}$$

正项式：单项式的和

$$f(\boldsymbol{x}) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}}\cdots x_n^{a_{nk}}, \qquad c_k>0$$

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:

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

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

注：前面练习已经说明了 $e^{\tilde{f}_i}$ 是凸函数；$\tilde{f}_i$ 是凸函数将作为作业。

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：

$$
\begin{aligned}
\text{minimize} \quad & 2\left(x_1x_2 + x_1x_3 + x_2x_3\right) \\
\text{subject to} \quad & \boldsymbol{x} \succeq \mathbf{0} \\
& x_1x_2x_3 = V
\end{aligned}
$$

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，以及掌握证明
- 基于一定的资料，自行写出简单的优化代码
- 掌握线性规划问题的一般形式、几何含义、以及了解上述具体例子
- 掌握二次规划问题一般形式、几何含义、以及了解上述具体例子
- 知道几何规划的定义和如何将几何规划问题变成凸优化问题

阅读作业 & 参考资料:

- 课本第 4.1 - 4.5 章
- 阅读第13页中的两个网站链接的内容