第7章 二次规划

定义7.1.1.

目标函数为二次函数且约束为线性约束的最优化问题称为**二次规划**(quadratic programming, QP)。

二次规划的代表性应用，比如：

- 马科维茨投资组合风险收益模型;

$$\max \quad \boldsymbol{x}^T\mu - \kappa\boldsymbol{x}^T G\boldsymbol{x} \tag{7.1.1a}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} x_i = 1, \boldsymbol{x} \geq 0, \tag{7.1.1b}$$

- 经典约束优化方法中二次子问题的求解，如增广Lagrange方法、序列二次规划方法等。

7-1 消去法与Lagrange方法

二次规划

二次规划问题一般数学模型：

$$\min \quad q(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x} + \boldsymbol{g}^T\boldsymbol{x} \tag{7.1.2a}$$

$$\text{s.t.} \quad \boldsymbol{a}_i^T\boldsymbol{x} = b_i, \qquad i \in \mathcal{E} \tag{7.1.2b}$$

$$\boldsymbol{a}_i^T\boldsymbol{x} \geq b_i, \qquad i \in \mathcal{I} \tag{7.1.2c}$$

其中：G是n阶对称矩阵，$\boldsymbol{g}$，$\boldsymbol{x}$和$\{\boldsymbol{a}_i\}$，$i \in \mathcal{E} \cup \mathcal{I}$都是$n$维向量。可行域可能为空集，我们仅考虑可行域非空情形。

二次规划问题分类：

- **凸二次规划**: Hesse矩阵半正定，问题有全局解。
- **严格凸二次规划**: Hesse矩阵正定，问题有唯一全局解。
- **非凸二次规划**：Hessen矩阵不定，问题可能有多个稳定点和局部极小点。

直接消去法

在**直接消去法**中, 对相关的量$A, \boldsymbol{x}, G$与$\boldsymbol{g}$作如下分块。

$$A = \begin{bmatrix} A_B \\ A_N \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{bmatrix}, \quad G = \begin{bmatrix} G_{BB} & G_{BN} \\ G_{NB} & G_{NN} \end{bmatrix}, \quad \boldsymbol{g} = \begin{bmatrix} \boldsymbol{g}_B \\ \boldsymbol{g}_N \end{bmatrix}$$

这里A_B为m阶非奇异方阵, $\boldsymbol{x}_B \in \mathbb{R}^m$, $\boldsymbol{x}_N \in \mathbb{R}^{n-m}$.
于是等式约束(7.1.2b)改写成:

$$A_B^T \boldsymbol{x}_B + A_N^T \boldsymbol{x}_N = b, \tag{7.1.3}$$

因A_B^{-1}存在, 此时由方程(7.1.3)可改写为:

$$\boldsymbol{x}_B = A_B^{-T}(\boldsymbol{b} - A_N^T \boldsymbol{x}_N). \tag{7.1.4}$$

直接消去法

将(7.1.4)代入$q(\boldsymbol{x})$, 则等式约束二次规划问题(7.1.2)可转化为如下无约束优化问题：

$$\min_{\boldsymbol{x}_N\in\mathbb{R}^{n-m}} \frac{1}{2}\boldsymbol{x}_N^T\hat{G}\boldsymbol{x}_N+\hat{\boldsymbol{g}}^T\boldsymbol{x}_N+\hat{\boldsymbol{c}}, \tag{7.1.5}$$

其中

$$\begin{aligned}
\hat{G} &= G_{NN}-G_{NB}A_B^{-T}A_N^T-A_NA_B^{-1}G_{BN}\\
&\qquad +A_NA_B^{-1}G_{BB}A_B^{-T}A_N^T,\\
\hat{\boldsymbol{g}} &= \boldsymbol{g}_N-A_NA_B^{-1}\boldsymbol{g}_B+(G_{NB}-A_NA_B^{-1}G_{BB})A_B^{-1}\boldsymbol{b},\\
\hat{\boldsymbol{c}} &= \frac{1}{2}\boldsymbol{b}^TA_B^{-1}G_{BB}A_B^{-T}\boldsymbol{b}+\boldsymbol{g}_B^TA_B^{-T}\boldsymbol{b}.
\end{aligned}$$

直接消去法

情形1： 如果矩阵$\hat{G}$正定，则问题(7.1.5)的最优解为$\boldsymbol{x}_N^* = -\hat{G}^{-1}\hat{\boldsymbol{g}}$. 代入(7.1.4) 即可确定对应的$\boldsymbol{x}_B^*$，从而找到问题(7.1.2)的最优解：

$$\boldsymbol{x}^* = \begin{bmatrix} \boldsymbol{x}_B^* \\ \boldsymbol{x}_N^* \end{bmatrix} = \begin{bmatrix} A_B^{-T}\boldsymbol{b} + A_B^{-T}A_N^T\hat{G}^{-1}\hat{\boldsymbol{g}} \\ -\hat{G}^{-1}\hat{\boldsymbol{g}} \end{bmatrix}, \tag{7.1.6}$$

相应的最优Lagrange乘子$\boldsymbol{\lambda}^*$可由下式确定

$$A\boldsymbol{\lambda}^* = \boldsymbol{g} + G\boldsymbol{x}^*. \tag{7.1.7}$$

只须考虑(7.1.7)前m行即可给出$\boldsymbol{\lambda}^*$的表达式：

$$\boldsymbol{\lambda}^* = A_B^{-1}(\boldsymbol{g}_B + G_{BB}\boldsymbol{x}_B^* + G_{BN}\boldsymbol{x}_N^*).$$

直接消去法

情形2： 如果矩阵$\hat{G}$半正定, 不难证明当

$$(I - \hat{G}\hat{G}^{\dagger})\hat{\boldsymbol{g}} = 0 \tag{7.1.8}$$

时,问题(7.1.5)有下界, 且其最优解为

$$\boldsymbol{x}_N^* = -\hat{G}^{\dagger}\hat{\boldsymbol{g}} + (I - \hat{G}^{\dagger}\hat{G})\tilde{\boldsymbol{x}},$$

这里$\tilde{\boldsymbol{x}} \in \mathbb{R}^{n-m}$为任意向量, $\hat{G}^{\dagger}$表示$\hat{G}$的广义逆矩阵.

若(7.1.8)不成立, 则可以推出问题(7.1.5)无下界, 从而原问题(7.1.2)亦无下界。

直接消去法

情形3： 如果矩阵$\hat{G}$存在负的特征值, 则问题(7.1.5)无下界, 原问题(7.1.2)不存在有限最优解。

直接消去法优缺点：

- 等式二次规划问题的直接消去法简单明了;
- 计算A_B的逆涉及多个矩阵相乘, 且A_B近于奇异时，可能由于计算误差导致不稳定。

直接消去法可推广为广义消去法。

算例分析

例 7.1.1 考虑等式约束二次规划问题：

$$\begin{aligned} \min \quad & x_1^2 - x_1x_2 + x_2^2 - x_2x_3 + x_3^2 + 2x_1 - x_2 \\ \text{s.t.} \quad & 3x_1 - x_2 - x_3 = 0 \\ & 2x_1 - x_2 - x_3 = 0 \end{aligned}$$

用直接消去法求其最优解。

解： 写出目标函数中的G, g如下：

$$G = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, g = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, A^T = \begin{bmatrix} 3 & -1 & -1 \\ 2 & -1 & -1 \end{bmatrix}$$

选择$B = \{1, 2\}, N = \{3\}$, 则$x_1 = 0, x_2 = -x_3$, 代入目标函数得:

$$\hat{q}(\boldsymbol{x}_N) = 3x_3^2 + x_3$$

其极小点$x_3^* = -\frac{1}{6}$, 从而得到：$\boldsymbol{x}^* = (0, \frac{1}{6}, -\frac{1}{6})^T, \boldsymbol{\lambda}^* = (\frac{5}{6}, -\frac{1}{3})^T$

广义消去方法

广义消去法也称为**零空间方法**。假定A是列满秩的，将$\Re^n$分成为两个互补的子空间：

$$\mathbb{R}^n = R(A) \oplus N(A^T)$$

其中$R(A)$是由A的列向量生成的像空间，$N(A^T)$是A^T的零空间。

设：

$$\begin{aligned} Y &= [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m] \in \mathbb{R}^{n\times m} \\ Z &= [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n-m}] \in \mathbb{R}^{n\times(n-m)} \end{aligned}$$

其中$\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$是$R(A)$中一组线性无关向量,$\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n-m}$是$N(A^T)$中一组线性无关向量。

广义消去方法

不失一般性，可以选择Y, Z, 使得$[Y, Z]$非奇异, 且满足：

$$A^T Y = I, \quad A^T Z = 0 \tag{7.1.9}$$

令:

$$\boldsymbol{x} = Y\boldsymbol{x}_y + Z\boldsymbol{x}_z \tag{7.1.10}$$

等式约束$A^T\boldsymbol{x} = \boldsymbol{b}$可写为：

$$A^T Y\boldsymbol{x}_y + A^T Z\boldsymbol{x}_z = \boldsymbol{b}$$

结合(7.1.9)得$\boldsymbol{x}_y = \boldsymbol{b}$, 因此原问题的可行点可表示为：

$$\boldsymbol{x} = Y\boldsymbol{b} + Z\boldsymbol{x}_z \tag{7.1.11}$$

广义消去方法

将(7.1.11)代入目标函数得到无约束优化问题：

$$\min \frac{1}{2}\boldsymbol{x}_z^T Z^T G Z \boldsymbol{x}_z + (\boldsymbol{g} + GY\boldsymbol{b})^T Z\boldsymbol{x}_z + \frac{1}{2}(2\boldsymbol{g} + GY\boldsymbol{b})^T Y\boldsymbol{b} \qquad (7.1.12)$$

这里$Z^T GZ \in \mathbb{R}^{(n-m)\times(n-m)}$称为**简约(投影)Hesse矩阵**。若该矩阵正定，则以上问题的唯一解满足：

$$(Z^T GZ)\boldsymbol{x}_z = -Z^T(\boldsymbol{g} + GY\boldsymbol{b}) \qquad (7.1.13)$$

利用Cholesky分解可求解以上方程组得到$\boldsymbol{x}_z^*$，于是相应最优解与最优Lagrange乘子分别为：

$$\boldsymbol{x}^* = Y\boldsymbol{b} + Z\boldsymbol{x}_z^*, \quad \boldsymbol{\lambda}^* = Y^T(G\boldsymbol{x}^* + \boldsymbol{g})$$

广义消去方法

接下来介绍选取Y, Z满足(7.1.9)的方法。不同的Y, Z选取方法得到不同的广义消去法。比如, QR分解:

$$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix} = [Q_1, Q_2]\begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 R$$

其中$Q \in \mathbb{R}^{n\times n}$是正交矩阵, $R \in \mathbb{R}^{m\times m}$是非奇异上三角矩阵, $Q_1 \in \mathbb{R}^{n\times m}$, $Q_2 \in \mathbb{R}^{n\times(n-m)}$.

取:

$$Y = Q_1 R^{-T}, \quad Z = Q_2$$

即满足条件(7.1.9).

用广义消去方法解**例 7.1.1.**

解： 对矩阵A作QR分解：

$$A=\begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{2}{\sqrt{22}} & 0 \\ -\frac{1}{\sqrt{11}} & -\frac{3}{\sqrt{22}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{11}} & -\frac{3}{\sqrt{22}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} \sqrt{11} & \frac{8}{\sqrt{11}} \\ 0 & \sqrt{\frac{2}{11}} \\ 0 & 0 \end{bmatrix}$$

取：

$$Y=\begin{bmatrix} 1 & -1 \\ 1 & -\frac{3}{2} \\ 1 & -\frac{3}{2} \end{bmatrix},\qquad Z=\frac{\sqrt{2}}{2}\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

算例分析

解(续)：令$\boldsymbol{x} = Yx_y + Zx_z$，由约束与$Y, Z$性质知：$x_y = 0$，故可行点为$\boldsymbol{x} = Zx_z$，代入目标函数得：

$$\min \frac{3}{2}x_z^2 + \frac{\sqrt{2}}{2}x_z$$

该问题的最优解为：$x_z^* = -\frac{\sqrt{2}}{6}$。

于是原问题的最优解与相应的拉格朗日乘子分别为：

$$\boldsymbol{x}^* = Zx_z^* = \left(0, \frac{1}{6}, -\frac{1}{6}\right)^T,$$

$$\boldsymbol{\lambda}^* = \left(\frac{5}{6}, -\frac{1}{3}\right)^T$$

Lagrange方法

考虑Lagrange函数为：

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x} + \boldsymbol{g}^T\boldsymbol{x} - \boldsymbol{\lambda}^T(A^T\boldsymbol{x} - \boldsymbol{b}), \tag{7.1.14}$$

由约束优化一阶必要条件，即KKT条件：

$$\frac{\partial L(\boldsymbol{x}, \boldsymbol{\lambda})}{\partial \boldsymbol{x}} = G\boldsymbol{x} + \boldsymbol{g} - A\boldsymbol{\lambda} = 0, \quad \frac{\partial L(\boldsymbol{x}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = A^T\boldsymbol{x} - \boldsymbol{b} = 0.$$

其矩阵形式为($\boldsymbol{x}^*$, $\boldsymbol{\lambda}^*$)满足：

$$\begin{bmatrix} G & -A \\ -A^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^* \\ \boldsymbol{\lambda}^* \end{bmatrix} = -\begin{bmatrix} \boldsymbol{g} \\ \boldsymbol{b} \end{bmatrix}. \tag{7.1.15}$$

上式中称为**KKT方程组**, 系数矩阵称为**KKT矩阵**。

定理 7.1.1.

假设A为列满秩矩阵，若简约Hesse矩阵Z^TGZ正定，则KKT矩阵非奇异，KKT方程组有唯一解$(\boldsymbol{x}^*,\ \boldsymbol{\lambda}^*)$.

证明： 假设存在向量$\boldsymbol{u} \in \mathbb{R}^n, \boldsymbol{v} \in \mathbb{R}^m$使得:

$$\begin{bmatrix} G & -A \\ -A^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = 0. \tag{7.1.16}$$

因为$A^T\boldsymbol{u} = 0$, 则对上式左乘$[\boldsymbol{u}^T\ \boldsymbol{v}^T]$可得

$$0 = [\boldsymbol{u}^T\ \boldsymbol{v}^T] \begin{bmatrix} G & -A \\ -A^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{bmatrix} = \boldsymbol{u}^T G \boldsymbol{u}.$$

定理7.1.1证明

因为$\boldsymbol{u}$在A^T的零空间, 则对某个$\boldsymbol{w} \in \mathbb{R}^{n-m}$, $\boldsymbol{u} = Z\boldsymbol{w}$成立. 因此，

$$0 = \boldsymbol{u}^T G \boldsymbol{u} = \boldsymbol{w}^T Z^T G Z \boldsymbol{w}.$$

由于矩阵$Z^T G Z$正定，则有$\boldsymbol{w} = 0$, 从而$\boldsymbol{u} = 0$. 将$\boldsymbol{u} = 0$代入(7.1.16)得

$$A\boldsymbol{v} = 0,$$

再由A为列满秩，可得$\boldsymbol{v} = 0$，所以KKT矩阵非奇异，进而KKT方程组有唯一解$(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$。 ■

惯性指数定理

定理 7.1.2 (惯性指数定理).

假设A为列满秩，其秩为m，若简约Hesse矩阵Z^TGZ正定, 则等式约束二次规划问题(7.1.2)的KKT矩阵有n个正的特征值，m个负的特征值，并且没有零特征值。

例 7.1.2 考虑等式约束二次规划问题：

$$\begin{aligned}\min \quad & q(x)=3x_1^2+2x_1x_2+x_1x_3+2.5x_2^2+2x_2x_3+2x_3^2\\ & -8x_1-3x_2-3x_3\\ \text{s.t.} \quad & x_1+x_3=3\\ & x_2+x_3=0.\end{aligned}$$

解： 使用矩阵形式，有

$$G=\begin{bmatrix}6&2&1\\2&5&2\\1&2&4\end{bmatrix},\ \boldsymbol{g}=\begin{bmatrix}-8\\-3\\-3\end{bmatrix},\ A^T=\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix},\ \boldsymbol{b}=\begin{bmatrix}3\\0\end{bmatrix},$$

矩阵G是正定的，齐次方程组$A^T\boldsymbol{x}=0$的基础解系为：

$$\boldsymbol{\eta}_1=(-1,-1,1)^T.$$

Lagrange方法

所以零空间$\mathcal{N}(A^T)$的基矩阵为

$$Z = (-1, -1, -1)^T.$$

容易验证Z^TGZ正定，故KKT矩阵非奇异, 因而KKT方程组有唯一最优解，解该方程组得最优解$\boldsymbol{x}^*$和相应的Lagrange乘子$\boldsymbol{\lambda}^*$为：

$$\boldsymbol{x}^* = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\lambda}^* = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$