

## § 4 投影梯度法

该方法是J.B.Rosen于1960年针对求解线性约束的非线性规划问题提出的.

1961年,他又将此方法推广到求解非线性约束的情况,后来又得到了进一步的发展,最终成为求解非线性规划问题的一类重要方法.

考虑含一般线性约束的非线性规划问题

$$\min f(x), x \in R^n$$

$$s.t. Ax \leq b \quad (\text{LNP1})$$

$$Ex = e$$

1960年，Rosen 将求解无约束非线性规划的最速下降法推广到(LNP1)中， 具体而言：

当迭代点是可行域的内点时，将目标函数负梯度作为搜索方向，当迭代点在可行域边界上时，将目标函数的负梯度在这些边界面的交集上的投影作为搜索方向。

无论哪种情况，所得搜索方向都是下降可行方向。然后在可行域内沿此方向一维搜索后得新的迭代点。



**定义1:** 设 $V$ 为数域 $P$ 上的线性空间,  $V_1, V_2$ 是 $V$ 的子空间, 且 $V = V_1 \oplus V_2$ . 那么对任意 $x \in V$ , 有 $x = x_1 + x_2, x_1 \in V_1, x_2 \in V_2$ .

并称 $x_1$ 为 $x$ 在 $V_1$ 上的**正交投影**,

称 $x_2$ 为 $x$ 在 $V_2$ 上的**正交投影**.

**定义2.** 称 $n \times n$ 矩阵 $P$ 为投影矩阵, 如果 $P = P^T$ , 且 $PP = P$ .

**引理1.** 设 $P$ 为 $n \times n$ 矩阵

(1)若 $P$ 为投影矩阵, 则 $P$ 是半正定的;

(2) $P$ 为投影矩阵, 当且仅当 $I - P$ 是投影矩阵,  
其中 $I$ 为 $n$ 阶单位阵;

(3)设 $P$ 为投影矩阵, 令 $Q = I - P$ , 则

$$L = \{Px \mid x \in R^n\} \text{ 与 } L^\perp = \{Qx \mid x \in R^n\}$$

为互相正交的线性子空间, 并且任一点 $x \in R^n$

可唯一地表示为 $x = p + q$ ,  $p \in L, q \in L^\perp$

显然  $p$  称为  $x$  在  $L$  上的正交投影.



(1) 设  $x \in R^n$  为任一向量, 因为  $P$  为投影矩阵, 所以  $x^T Px = x^T PPx = x^T P^T Px = \|Px\|^2 \geq 0$  因此  $P$  是半正定的.

(2) 设  $P$  为投影矩阵, 则

$$(I - P)^T = I - P^T = I - P$$

$$\text{且 } (I - P)(I - P) = I - P - P + PP = I - P$$

因此  $I - P$  是投影矩阵.

反之, 设  $I - P$  是投影矩阵,

由上面的结论知  $P = I - (I - P)$  亦是投影矩阵.



(3) 任取  $L$  中的两元素  $Px, Py \in L$ ,

对任意实数  $\lambda \in R$ , 有  $Px + Py = P(x + y) \in L$

$$\lambda(Px) = P(\lambda x) \in L$$

所以  $L$  对加法和数乘运算封闭, 另外, 易验证

$L$  上的加法和数乘运算也满足8条运算规律,

因此  $L$  是线性空间. 同理可证  $L^\perp$  也是线性空间.

自然他们都是  $R^n$  的线性子空间.

由于  $P^T Q = PQ = P(I - P) = P - PP = O$

所以  $(Px)^T (Qy) = x^T P^T Qy = 0, \forall x, y \in R^n$

所以  $L$  与  $L^\perp$  正交.

由于对任意的  $x \in R^n$ , 有

$$x = Ix = (P + Q)x = Px + Qx = p + q$$

其中  $p \in L, q \in L^\perp$ .

下面再证明此表示式是唯一的.

设有  $p, p' \in L, q, q' \in L^\perp$ , 使  $x = p + q, x = p' + q'$

则有  $p - p' = q - q'$

且  $p - p' \in L, q - q' \in L^\perp$

由于  $L$  与  $L^\perp$  正交, 所以  $p - p' = q - q' = 0$ .

即  $p = p', q = q'$ .

**引理2.** 设 $x$ 为问题(LNP1)的一个可行解, 且使 $A_1x = b_1, A_2x < b_2$ , 而 $A^T = (A_1^T, A_2^T)$ ,  $b^T = (b_1^T, b_2^T)$ . 设 $f(x)$ 在点 $x$ 可微,  $M^T = (A_1^T, E^T)$ 满秩, 则

(1)  $P = I - M^T (MM^T)^{-1} M$ 为投影矩阵;

(2) 若 $P\nabla f(x) \neq 0$ , 则 $d = -P\nabla f(x)$ 是 $x$ 点的一个下降可行方向.



**定理1.** 设 $x$ 为问题(LNP1)的一个可行解, 且使 Page 9

$A_1 x = b_1, A_2 x < b_2$ , 而  $A^T = (A_1^T, A_2^T), b^T = (b_1^T, b_2^T)$ .

设 $f(x)$ 在点 $x$ 可微,  $M^T = (A_1^T, E^T)$ 满秩, 则令

$$P = I - M^T (MM^T)^{-1} M$$

$$w = -(MM^T)^{-1} M \nabla f(x), \quad w^T = (u^T, v^T)$$

设  $P \nabla f(x) = 0$ ,

(1) 若  $u \geq 0$ , 则 $x$ 是一个K-T点;

(2) 若  $u \not\geq 0$ , 令  $u_j$  是一个负分量,  $\tilde{M}^T = (\tilde{A}_1^T, E^T)$

其中  $\tilde{A}_1$  是由  $A_1$  中去掉第  $j$  行后得到的矩阵, 令

$$\tilde{P} = I - \tilde{M}^T (\tilde{M}\tilde{M}^T)^{-1} \tilde{M}, \quad d = -\tilde{P} \nabla f(x)$$

则  $d$  是点  $x$  的一个下降可行方向.



## 二. 投影梯度法算法

考虑含一般线性约束的非线性规划问题

$$\min f(x), x \in R^n$$

$$s.t. Ax \leq b \quad (\text{LNP1})$$

$$Ex = e$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{pmatrix}$$



$$E = \begin{pmatrix} \mathbf{a}_{m+1,1} & \mathbf{a}_{m+1,2} & \cdots & \mathbf{a}_{m+1,n} \\ \mathbf{a}_{m+2,1} & \mathbf{a}_{m+2,2} & \cdots & \mathbf{a}_{m+2,n} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{a}_{m+l,1} & \mathbf{a}_{m+l,2} & \cdots & \mathbf{a}_{m+l,n} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{m+1}^T \\ \mathbf{a}_{m+2}^T \\ \vdots \\ \mathbf{a}_{m+l}^T \end{pmatrix}$$

$$B = \begin{pmatrix} \mathbf{b} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_{m+l} \end{pmatrix},$$

$$J_k = J(\mathbf{x}^k) = \{j \mid \mathbf{a}_j^T \mathbf{x}^k = B_j\},$$

$M_k$  表示以  $\mathbf{a}_j^T$  ( $j \in J_k$ ) 为行所组成的矩阵



- Step1. 选取(LNP1)的一个可行解 $x^1$ , 给定误差精度 $\varepsilon > 0$ , 令 $k = 1$ ;
- Step2. 计算 $\nabla f(x^k)$ ,  $J_k$ , 若 $\|\nabla f(x^k)\| < \varepsilon$ , 则 $x^k$ 为近似的K-T点, 计算结束; 否则,
- (1) 若 $J_k = \phi$ , 令 $P = I$ ,
  - (2) 若 $J_k \neq \phi$ , 令 $P = I - M_k^T (M_k M_k^T)^{-1} M_k$
- Step3. 若 $P\nabla f(x^k) \neq 0$ , 令 $d_k = -P\nabla f(x^k)$ , 转step 4;



若  $P\nabla f(x^k) = \mathbf{0}$ , 则令

$$w = -(M_k M_k^T)^{-1} M_k \nabla f(x^k) = \begin{pmatrix} u \\ v \end{pmatrix}$$

(1) 若  $u \geq \mathbf{0}$ , 则  $x^k$  为  $K$ - $T$  点, 计算结束;

(2) 若  $u$  有某个分量  $u_j < \mathbf{0}$ , 令  $\tilde{M}_k$  是在  $M_k$  中去掉  $u_j$  对应的第  $j$  行而得到的矩阵.

$$\text{令 } \tilde{P} = I - \tilde{M}_k^T (\tilde{M}_k \tilde{M}_k^T)^{-1} \tilde{M}_k, \quad d_k = -\tilde{P} \nabla f(x^k)$$

转 Step 4.



$$\lambda_{Max} = \begin{cases} \min \left\{ \frac{B_i - a_i^T x^k}{a_i^T d_k} \mid i \notin J_k, a_i^T d_k > 0 \right\} \\ +\infty, \text{ 若对任意 } i, a_i^T d_k \leq 0 \end{cases}$$

线搜索求解  $\begin{cases} \min f(x^k + \lambda d_k) \\ s.t. \ 0 \leq \lambda \leq \lambda_{Max} \end{cases}$  得最优步长  $\lambda_k$ ,

令  $x^{k+1} = x^k + \lambda_k d_k$ ,  $k = k + 1$ , 转Step2.

**备注:** 这里求步长的方法与Zoutenijk可行方向法中所用的方法相同, 只是写法不同而已.



# 例1：用投影梯度法求解

$$\min f(x) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$$

$$s.t. \quad x_1 + x_2 \leq 2$$

$$x_1 + 5x_2 \leq 5$$

$$-x_1 \leq 0, -x_2 \leq 0$$

初始点取为  $x^1 = (0, 0)^T$ .

解：因为本题没有等式约束，所以  $M = A_1$ ,  $w = u$

$$\nabla f(x) = (4x_1 - 2x_2 - 4, 4x_2 - 2x_1 - 6)^T$$

下面进行第一次迭代：



# 1. 确定搜索方向

在 $x^1$ 点处,  $\nabla f(x^1) = (-4, -6)^T$ ,  $J_1 = \{3, 4\}$

$$M_1 = A_1 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, P = I - M_1^T (M_1 M_1^T)^{-1} M_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$P \nabla f(x^1) = (0, 0)^T, w = -(M_1 M_1^T)^{-1} M_1 \nabla f(x^1) = (-4, -6)^T = u$$

取 $u_4 = -6$ , 从 $M_1$ 中去掉与 $u_4$ 对应的行得到

$$\tilde{M}_1 = \tilde{A}_1 = (-1, 0), \tilde{P} = I - \tilde{M}_1^T (\tilde{M}_1 \tilde{M}_1^T)^{-1} \tilde{M}_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

此时取搜索方向为:

$$d_1 = -\tilde{P} \nabla f(x^1) = -\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -4 \\ -6 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$$



## 2. 进行线搜索求步长

$$x = x^1 + \lambda d_1 = (0, 6\lambda)^T$$

所以  $f(x^1 + \lambda d_1) = 72\lambda^2 - 36\lambda$       注意到  $A_2 d_1 = \begin{pmatrix} 6 \\ 30 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$\lambda_M = \min\{2/6, 5/30\} = 1/6$$

$$b_2 - A_2 x^1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

求解线搜索问题

$$\begin{cases} \min f(x^1 + \lambda d_1) = 72\lambda^2 - 36\lambda \\ s.t. \quad 0 \leq \lambda \leq 1/6 \end{cases}$$

得最优解  $\lambda_1 = 1/6$ ,

于是有  $x^2 = x^1 + \lambda_1 d_1 = (0, 1)^T$ .



备注：上页中求步长上界的方法与上一节相同 Page 18

$$\text{注： } b_2 - A_2 x^1 = \begin{pmatrix} 2 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

$$A_2 d_1 = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 6 \\ 30 \end{pmatrix}$$

这里的 $A_2$ 是指无效约束对应的系数矩阵。

可见  $A_2 d_1 > 0$ ,

$$\text{所以 } \lambda_M = \min \left\{ \frac{2}{6}, \frac{5}{30} \right\} = \frac{1}{6}$$

## 1. 确定搜索方向

在 $x^2$ 点处,  $\nabla f(x^2) = (-6, -2)^T$ ,  $J_2 = \{2, 3\}$

$$M_2 = \begin{pmatrix} 1 & 5 \\ -1 & 0 \end{pmatrix}, \quad P = I - M_2^T (M_2 M_2^T)^{-1} M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$P \nabla f(x^2) = (0, 0)^T, \quad u = -(M_2 M_2^T)^{-1} M_2 \nabla f(x^2) = (2/5, -28/5)^T$$

因为 $u_3 = -28/5 < 0$ , 所以从 $M_2$ 中去掉行 $(-1, 0)$ , 得到

$$\tilde{M}_2 = (1, 5), \quad \tilde{P} = I - \tilde{M}_2^T (\tilde{M}_2 \tilde{M}_2^T)^{-1} \tilde{M}_2 = \begin{pmatrix} 25/26 & -5/26 \\ -5/26 & 1/26 \end{pmatrix}$$

$$d_2 = -\tilde{P} \nabla f(x^2) = (70/13, -14/13)^T, \quad \text{不妨取 } d_2 = (5, -1)^T.$$



## 2. 进行线搜索

$$x = x^2 + \lambda d_2 = (5\lambda, 1 - \lambda)^T$$

$$\text{所以 } f(x^2 + \lambda d_2) = 62\lambda^2 - 28\lambda - 4$$

$$\lambda_M = \min\{1/4, 1\} = 1/4$$

注意到  $A_2 d_2 = \begin{pmatrix} 4 \\ 1 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$b_2 - A_2 x^2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

求解线搜索问题

$$\begin{cases} \min f(x^2 + \lambda d_2) = 62\lambda^2 - 28\lambda - 4 \\ \text{s.t. } 0 \leq \lambda \leq 1/4 \end{cases}$$

得最优解  $\lambda_2 = 7/31$ ,

$$x^3 = x^2 + \lambda_2 d_2 = (35/31, 24/31)^T$$



第三次迭代,  $x^3 = (35 / 31, 24 / 31)^T$

在 $x^3$ 点处,  $\nabla f(x^3) = (-32 / 31, -160 / 31)^T$ ,  $J_3 = \{2\}$

$$M_3 = (1, 5),$$

$$P = I - M_3^T (M_3 M_3^T)^{-1} M_3 = \frac{1}{26} \begin{pmatrix} 25 & -5 \\ -5 & 1 \end{pmatrix}$$

$$P \nabla f(x^3) = (0, 0)^T$$

$u = -(M_3 M_3^T)^{-1} M_3 \nabla f(x^3) = 32 / 31 > 0$ , 所以 $x^3$ 是 $K$ - $T$ 点.

又因为 $f(x)$ 是凸函数, 因此 $x^3$ 也是问题的全局最优解.



# 本节结束

