6.1 梯度投影法

SMaLL

[1] 中国石油大学（华东）
SMaLL 课题组
small.sem.upc.edu.cn
liangxijunsd@163.com

2023

梯度投影法

1. 基于投影方法的可行方向和步长规则
2. 变尺度梯度投影
3. 约束牛顿法

基于投影方法的可行方向

$$\begin{aligned} \min_x \quad & f(x) \\ s.t. \quad & x \in X \end{aligned} \tag{1}$$

其中 f: 凸函数, $X \subseteq \mathbb{R}^n$: 凸集.
梯度投影法是一种可行方向法，具有下面的迭代格式:

$$x^{k+1} = x^k + \alpha^k \left(\bar{x}^k - x^k\right)$$

其中

$$\bar{x}^k = \left[x^k - s^k \nabla f(x^k)\right]^+$$

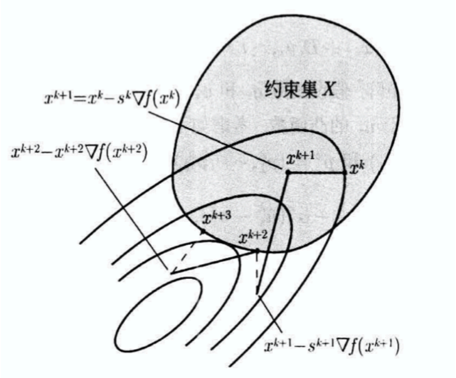
- $[\cdot]^+$ 表示在集合 X 上的投影
- $\alpha^k \in (0, 1]$ 是步长
- s^k 是正数

基于投影方法的可行方向

s^k： 步长.

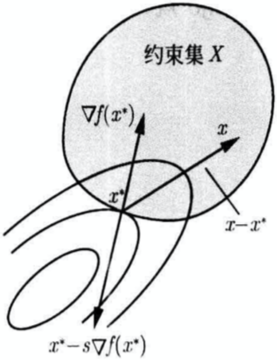
选择 $\alpha^k \equiv 1 \to x^{k+1} = \bar{x}^k$,

$$x^{k+1} = [x^k - s^k \nabla f(x^k)]^+$$



基于投影方法的可行方向

$x^* = [x^* - s\nabla f(x^*)]^+,\ s > 0 \leftrightarrow x^*$ 是驻点



算法停止 $\leftrightarrow$ 遇到驻点时

基于投影方法的可行方向

如果 X 有相对简单的结构，投影运算通常有显式解

例

当约束集是由上下界限定给出的箱式集合，

$$X = \{x \mid \alpha_i \le x_i \le \beta_i, i = 1, \ldots, n\}$$

该集合投影向量 x 的第 i 个分量由下式确定

$$[x]_i^+ = \begin{cases} \alpha_i & \text{若} x_i \le \alpha_i \\ \beta_i & \text{若} x_i \ge \beta_i \\ x_i & \text{否则} \end{cases}$$

步长选择

- **固定步长规则**
  令 s^k 为常数 $s > 0$
  α^k 固定为统一值

$$s^k = s: \text{ 常数}, \quad \alpha^k = 1, \quad k = 0, 1, \ldots$$

- **缩减步长规则**
  α^k 为给定常数且

$$s^k \to 0, \quad \sum_{k=0}^{\infty} s^k = \infty$$

步长选择

- **有限最小化步长规则**
  $s^k = s$: 常数, $k = 0, 1, \ldots$
  α^k 取 $[0, 1]$ 且满足

$$f(x^k + \alpha^k (\bar{x}^k - x^k)) = \min_{\alpha \in [0,1]} f(x^k + \alpha (\bar{x}^k - x^k)) .$$

- **沿可行方向的 Armijo 规则**
  $s^k = s$: 常数, $k = 0, 1, \ldots$
  按 Armijo 规则选取 $\alpha^k \in (0, 1)$
  - 对给定的 β 和 $\sigma \in (0, 1)$
  - 取 $\alpha^k = \beta^{m_k}$, 其中 m_k 是使下式成立的第一个非负整数 m

$$f(x^k) - f(x^k + \beta^m (\bar{x}^k - x^k)) \geq -\sigma \beta^m \nabla f(x^k)' (\bar{x}^k - x^k)$$

步长选择

- **沿投影弧的 Armijo 规则**
  取步长 α^k 为给定值, 令 $\alpha^k = 1, \quad k = 0, 1, \ldots$
  s^k 逐渐地减小直到 Armijo 不等式成立 →

$$\{x^k(s) \mid s > 0\}$$

其中

$$x^k(s) = \left[x^k - s\nabla f(x^k)\right]^+, s > 0$$

  - 选择 $\bar{s} > 0, \beta \in (0, 1)$, 且 $\sigma \in (0, 1)$
  - 设 $s^k = \beta^{m_k}\bar{s}$, 其中 m_k 是使下面的不等式成立的第一个非负整数 m

$$f(x^k) - f(x^k(\beta^m \bar{s})) \geq \sigma \nabla f(x^k)' (x^k - x^k(\beta^m \bar{s}))$$

收敛速度

例如，当目标函数 f 是二次函数

$$f(x) = \frac{1}{2}x' Qx - b' x,$$

其中 Q 是正定的.
设 x^* 为 X 上 f 的唯一最小解, 考虑给定步长的情况.

$$\begin{aligned}
\|x^{k+1} - x^*\| &= \left\|[x^k - s\nabla f(x^k)]^+ - [x^* - s\nabla f(x^*)]^+\right\| \\
&\leq \|(x^k - s\nabla f(x^k)) - (x^* - s\nabla f(x^*))\| \\
&= \|(I - sQ)(x^k - x^*)\| \\
&\leq \max\{|1 - sm|, |1 - sM|\} \|x^k - x^*\|
\end{aligned}$$

其中 m 和 M 分别是 Q 的最小和最大的特征值.

梯度投影法

1. 基于投影方法的可行方向和步长规则
2. 变尺度梯度投影
3. 约束牛顿法

变尺度梯度投影

第 k 次迭代, 设 H^k 是一个正定矩阵，并考虑由

$$x = \left(H^k\right)^{-1/2} y$$

那么问题可以写做

$$\begin{aligned} &\text{minimize } h^k(y) \equiv f\left(\left(H^k\right)^{-1/2} y\right) \\ &\text{subject to } y \in Y^k, \end{aligned}$$

其中 Y^k 是集合

$$Y^k = \left\{ y \mid \left(H^k\right)^{-1/2} y \in X \right\}.$$

变尺度梯度投影

该问题的梯度投影迭代形式如下

$$y^{k+1} = y^k + \alpha^k \left(\bar{y}^k - y^k\right)$$

其中

$$\bar{y}^k = \left[y^k - s^k \nabla h^k \left(y^k\right)\right]^+$$

$\bar{y}^k$ 可以定义为使表达式最小化的向量

$$\left\|y - y^k + s^k \nabla h^k \left(y^k\right)\right\|^2 = \left(s^k\right)^2 \left\|\nabla h^k \left(y^k\right)\right\|^2 + 2 s^k \nabla h^k \left(y^k\right)' \left(y - y^k\right) + \left\|y - y^k\right\|^2$$

over $y \in Y^k$.

变尺度梯度投影

忽略表达式 $(s^k)^2 \|\nabla h^k(y^k)\|^2$ 并除以 $2s^k$，

$$\bar{y}^k = \arg\min_{y \in Y^k} \left\{ \nabla h^k(y^k)'(y - y^k) + \frac{1}{2s^k}\|y - y^k\|^2 \right\}$$

通过变化

$$x = (H^k)^{-1/2} y, \quad x^k = (H^k)^{-1/2} y^k, \quad \bar{x}^k = (H^k)^{-1/2} \bar{y}^k$$

$$\nabla h^k(y^k) = (H^k)^{-1/2} \nabla f(x^k)$$

迭代可以写成

$$x^{k+1} = x^k + \alpha^k(\bar{x}^k - x^k)$$

其中

$$\bar{x}^k = \arg\min_{x \in X} \left\{ \nabla f(x^k)'(x - x^k) + \frac{1}{2s^k}(x - x^k)' H^k (x - x^k) \right\}$$

梯度投影法

1. 基于投影方法的可行方向和步长规则

2. 变尺度梯度投影

3. 约束牛顿法

约束牛顿法

f 是二阶连续可微函数.

考虑变尺度梯度投影方法中的矩阵 $H^k = \nabla^2 f(x^k)$,

$$x^{k+1} = x^k + \alpha^k (\bar{x}^k - x^k)$$

$$\bar{x}^k = \arg\min_{x \in X} \left\{ \nabla f(x^k)' (x - x^k) + \frac{1}{2s^k} (x - x^k)' \nabla^2 f(x^k) (x - x^k) \right\}$$

$s^k = 1 \to f$ 在 x^k 的二阶泰勒展开 (约束牛顿法)

梯度投影法: $\bar{x}^k = \left[x^k - s^k \nabla f\left(x^k\right)\right]^+$

$$\bar{x}^k = \operatorname{argmin}_{x \in X} \left\{ \nabla f\left(x^k\right)' \left(x - x^k\right) + \frac{1}{2s^k} \left(x - x^k\right)' I \left(x - x^k\right) \right\}$$

$\bar{x}^k = \operatorname{argmin}_{x \in X}$ linear approx. of $f +$ proximal term $(\frac{1}{2s^k}||x - x^k||^2)$