最优化理论与方法笔记

叶胥达

2021 年 1 月 11 日

目录

1 无约束优化基本理论

1.1 基本概念

考虑无约束最优化问题

$$\min_{x\in\mathbb{R}^n} f(x)$$

- 一阶必要条件: $\nabla f(x^*) = 0$
- 二阶必要条件: $\nabla^2 f(x^*)$ 半正定
- 二阶充分条件: $\nabla^2 f(x^*)$ 正定

我们一般使用迭代格式求解最优化问题:

$$x_{k+1} = x_k + \alpha_k d_k \tag{1.1}$$

其中 x_k 为迭代点, d_k 为下降方向, α_k 为步长. 下降方向选取的最基本准则是: 沿着方向 d_k 应当使得目标函数值下降, 即

$$g_k^{\mathrm{T}} d_k < 0 \tag{1.2}$$

于是 $f(x_k + \alpha d_k)$ 在 α 很小时是关于步长 α 的减函数.

1.2 收敛速度

收敛速度是用于刻画算法效率的重要指标.

定义 1.1 (Q 收敛速度) 设存在 $p \geqslant 1$ 和 $q > 0$ 使得

$$\varlimsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} = q = Q_p\{x_k\}$$

则称 $\{x_k\}$ 具有 p 阶收敛速度.

收敛速度的一些例子:

- $p = 1$, $q \in (0, 1)$, 线性收敛:
$$x_k = 1 + \frac{1}{2^k}$$
- $p = 1$, $q = 0$, 超线性收敛:
$$x_k = k^{-k}$$
- $p = 1$, $q = 1$, 次线性收敛:
$$x_k = \frac{1}{k}$$
- $p > 1$, $q < +\infty$, 超线性收敛

定理 1.1 (超线性收敛的刻画) 设 $x_k \to x^*$, 记 $h_k = x_k - x^*$ 且 $s_k = x_{k+1} - x_k$, 则 $\{x_k\}$ 超线性收敛等价于下列四式中任意一个:

(i) $h_{k+1} = o(\|h_k\|)$;

(ii) $h_{k+1} = o(\|s_k\|)$;

(iii) $s_k = -h_k + o(\|h_k\|)$;

(iv) $s_k = -h_k + o(\|s_k\|)$

定义 1.2 (R 收敛速度) 给定 $p \geqslant 1$, 设迭代点序列 $\{x_k\}$ 满足

$$R_p\{x_k\} = \begin{cases} \overline{\lim\limits_{k\to\infty}} \|x_k - x^*\|^{\frac{1}{k}} \\ \overline{\lim\limits_{k\to\infty}} \|x_k - x^*\|^{\frac{1}{p^k}} \end{cases}$$

R 收敛速度不依赖于范数的选取.

定理 1.2 若 $\{x_k\}$ 具有 p 阶 Q 收敛速度, 则它也有 p 阶 R 收敛速度. 特别地, 若 $p=1$, 则

$$Q_1\{x_k\} \geqslant R_1\{x_k\}$$

设 $q = Q_1\{x_k\}$, 则对任意 $\varepsilon > 0$, 当 k 充分大时总有

$$\|x_{k+1} - x^*\| \leqslant (q+\varepsilon)\|x_k - x^*\|^p$$

当 $p=1$ 时, 有

$$\|x_k - x^*\| \leqslant (q+\varepsilon)^{k-k_0}\|x_{k_0} - x^*\|$$

因此

$$\|x_k - x^*\|^{\frac{1}{k}} \leqslant (q+\varepsilon)^{1-\frac{k_0}{k}}\|x_{k_0} - x^*\|^{\frac{1}{k}}$$

令 $k \to \infty$, 有

$$R_1\{x_k\} = \lim_{k\to\infty} \|x_k - x^*\|^{\frac{1}{k}} \leqslant q + \varepsilon$$

令 $\varepsilon \to 0$, 得到 $R_1\{x_k\} \leqslant Q_1\{x_k\}$. 在 $p > 1$ 时, 类似可以得到

$$\|x_k - x^*\| \leqslant (q+\varepsilon)^{s(k-k_0)}\|x_{k_0} - x^*\|^{p^{k-k_0}}$$

其中

$$s(m) = \sum_{i=0}^{m-1} p^i = \frac{p^m - 1}{p-1}$$

于是

$$\|x_k - x^*\|^{\frac{1}{p^k}} \leqslant (q+\varepsilon)^{\frac{p^{-k_0} - p^{-k}}{p-1}}\|x_{k_0} - x^*\|^{-p^{k_0}}$$

令 $k \to \infty$, 有

$$R_p\{x_k\} = \lim_{k\to\infty} \|x_k - x^*\|^{\frac{1}{p^k}} \leqslant (q+\varepsilon)^{\frac{p^{-k_0}}{p-1}}\|x_{k_0} - x^*\|^{-p^{k_0}}$$

即 $R_p\{x_k\}$ 是一个有限值.

2 线搜索准则

在得到下降方向 d_k 后, 一个关键的步骤是在 d_k 上求步长 α, 使得目标函数

$$f(x_k + \alpha d_k)$$

下降到较低的水平. 为简单起见, 我们去掉下标 k 并讨论 $f(x + \alpha d)$ 的优化问题. 进退法是用来确定搜索区间的一种简单策略, 它输出一个区间 $[a, b]$, 使得一定有极小点位于此区间中.

关于 $f(x + \alpha d)$ 的优化问题可以简化为关于一维函数

$$\varphi(\alpha) = f(x + \alpha d) \tag{2.1}$$

的优化, 此时

$$\varphi'(\alpha) = d^{\mathrm{T}} g(x + \alpha d) \tag{2.2}$$

2.1 精确线搜索准则

所谓的精确线搜索就是求 $\varphi(\alpha)$ 在指定区间上的最小值, 即求步长 α_k 使得

$$\alpha_k = \arg\min_{\alpha} \varphi(\alpha) \tag{2.3}$$

黄金分割 (0.618 方法) 可以方便地求单峰函数的最小值.

Algorithm 1: 0.618 方法

输入: 初始区间 $[a, b]$

输出: 极小点 $\theta \in [a, b]$

给定一个小常数 ε_0, 计算常数 $\tau = \frac{\sqrt{5}-1}{2}$

while *true* **do**

  计算试探函数点 $\lambda = a + (1-\tau)(b-a)$, $\mu = a + \tau(b-a)$

  if $\varphi(\lambda) > \varphi(\mu)$ **then**

    (左边函数值大, 极小点在 $[\lambda, b]$)

    $a = \lambda$, $\lambda = \mu$, $\mu = a + \tau(b-a)$

    if $b - a < \varepsilon_0$ **then**

      $\theta = \mu$

      return

    end

  else

    (右边函数值大, 极小点在 $[a, \mu]$)

    $b = \mu$, $\mu = \lambda$, $\lambda = a + (1-\tau)(b-a)$

    if $b - a < \varepsilon_0$ **then**

      $\theta = \lambda$

      return

    end

  end

end

黄金分割法仅利用了函数值的信息, 没有用到导数信息, 具有一阶收敛速度.

精确线搜索方法的收敛性分析:

定理 2.1 (下降量的估计) 设步长 α_k 由精确线搜索得到, 且

$$\left\|\nabla^2 f(x_k + \alpha d_k)\right\| \leqslant M$$

对一切 $\alpha > 0$ 成立. 则有

$$f_k - f_{k+1} \geqslant \frac{1}{2M}\|g_k\|^2 \cos^2\theta_k$$

其中 θ_k 是下降方向 d_k 和负梯度方向 $-g_k$ 的夹角.

定理 2.2 (精确线搜索的收敛性) 设 $\nabla f(x)$ 在水平集 $L = \{x \in \mathbb{R}^n : f(x) \leqslant f(x_0)\}$ 上存在且

一致连续. 若精确线搜索产生的迭代序列满足

$$\theta_k \leqslant \frac{\pi}{2} - \mu, \quad \text{对某个 } \mu > 0$$

则或者对某个 k 有 $g_k = 0$, 或者 $f_k \to -\infty$, 或者 $g_k \to 0$.

[Proof: 袁亚湘 p63-64]

反证法, 若结论不成立, 则存在一个无穷集 K 使得

$$\|g_k\| \geqslant \varepsilon, \quad \forall k \in K$$

因此利用夹角条件可以给出 $-g_k^{\mathrm{T}} d_k$ 的下界估计:

$$-g_k^{\mathrm{T}} d_k = \cos\theta_k \|g_k\| \|d_k\| \geqslant \varepsilon_1 \|d_k\|, \quad \forall k \in K$$

下面利用 Taylor 展开对每步下降量 $f_k - f_{k+1}$ 作出估计: 对任何 $\alpha \geqslant 0$, 有

$$f(x_k + \alpha d_k) = f(x_k) + \alpha d_k^{\mathrm{T}} g(\xi_k)$$

其中 $\xi_k \in [x_k, x_k + \alpha d_k]$. 因此

$$\begin{aligned} f_{k+1} - f_k &\leqslant f(x_k + \alpha d_k) - f_k \\ &= \alpha d_k^{\mathrm{T}} g(\xi_k) \\ &= \alpha d_k^{\mathrm{T}} g_k + \alpha d_k^{\mathrm{T}} [g(\xi_k) - g_k] \\ &\leqslant -\alpha \varepsilon_1 \|d_k\| + \alpha d_k^{\mathrm{T}} [g(\xi_k) - g_k] \end{aligned}$$

由于 $g(x) = \nabla f(x)$ 一致连续, 故存在 $\bar{\alpha} > 0$ 使得

$$\alpha \|d_k\| \leqslant \bar{\alpha} \Longrightarrow |g(\xi_k) - g_k| \leqslant \frac{\varepsilon_1}{2}$$

此时

$$f_{k+1} - f_k \leqslant -\alpha \varepsilon_1 \|d_k\| + \alpha \|d_k\| \frac{\varepsilon_1}{2} = -\frac{\alpha}{2} \varepsilon_1 \|d_k\|$$

特别地, 我们可以取 $\alpha \|d_k\| = \bar{\alpha}$, 从而得到

$$f_{k+1} - f_k \leqslant -\frac{\bar{\alpha}}{2} \varepsilon_1, \quad \forall k \in K$$

即每一步函数的下降量都至少是一个常数. 但 $\sum_{k=1}^{\infty} (f_k - f_{k+1})$ 为有限值, 矛盾!

精确线搜索的特点是比任何其它步长选择都更好.

2.2 插值法

插值法的基本思想是用一个简单的函数 $q(\alpha)$(常常是多项式函数) 对目标函数 $\varphi(\alpha)$ 进行逼近, 然后用 $q(\alpha)$ 的极小点对 $\varphi(\alpha)$ 的极小点进行近似. 由于 $q(\alpha)$ 要用来刻画 $\varphi(\alpha)$ 的极小点, 应当至少使用二次多项式, 即

$$q(\alpha) = a\alpha^2 + b\alpha + c \tag{2.4}$$

来逼近 $\varphi(\alpha)$. 该函数的极小点为

$$\alpha^* \approx \bar{\alpha} = -\frac{b}{2a} \tag{2.5}$$

一点二次插值法 在给定的点 α_1 处, 使用 $\varphi(\alpha_1), \varphi'(\alpha_1), \varphi''(\alpha_1)$ 对 $\varphi(\alpha)$ 进行插值, 得到

$$a = \frac{1}{2}\varphi''(\alpha_1), \quad b = \varphi'(\alpha_1) - \varphi''(\alpha_1)\alpha_1$$

因此计算所得 $q(\alpha)$ 的极小点为

$$\bar{\alpha} = -\frac{b}{2a} = \alpha_1 - \frac{\varphi'(\alpha_1)}{\varphi''(\alpha_1)}$$

相应的迭代格式为

$$\alpha_{k+1} = \alpha_k - \frac{\varphi'(\alpha_k)}{\varphi''(\alpha_k)}$$

两点二次插值 I 在区间 $[\alpha_1, \alpha_2]$ 上给定 $\varphi(\alpha_1), \varphi'(\alpha_1), \varphi(\alpha_2)$ 的值, 则 $q(\alpha)$ 的参数为

$$a = \frac{\varphi_1 - \varphi_2 - \varphi_1'(\alpha_1 - \alpha_2)}{-(\alpha_1 - \alpha_2)^2}$$

$$b = \varphi_1' + 2 \cdot \frac{\varphi_1 - \varphi_2 - \varphi_1'(\alpha_1 - \alpha_2)}{(\alpha_1 - \alpha_2)^2} \cdot \alpha_1$$

计算所得 $q(\alpha)$ 的极小点为

$$\bar{\alpha} = \alpha_1 - \frac{(\alpha_1 - \alpha_2)\varphi_1'}{2\left[\varphi_1' - \frac{\varphi_1 - \varphi_2}{\alpha_1 - \alpha_2}\right]}$$

相应的迭代格式为

$$\alpha_{k+1} = \alpha_k - \frac{(\alpha_k - \alpha_{k-1})\varphi_k'}{2\left[\varphi_k' - \frac{\varphi_k - \varphi_{k+1}}{\alpha_k - \alpha_{k+1}}\right]}$$

两点二次插值 II 在区间 $[\alpha_1, \alpha_2]$ 上给定 $\varphi'(\alpha_1), \varphi'(\alpha_2)$ 的值, 极小点为

$$\bar{\alpha} = \alpha_1 - \frac{\alpha_1 - \alpha_2}{\varphi_1' - \varphi_2'}\varphi_1'$$

相应的迭代格式为

$$\alpha_{k+1} = \alpha_k - \frac{\alpha_k - \alpha_{k-1}}{\varphi_k' - \varphi_{k-1}'}\varphi_{k-1}'$$

2.3 单调非精确线搜索方法

我们在本节介绍单调非精确线搜索方法, 这里单调的含义是保证目标函数在每步迭代之后单调下降, 即 $f(x_k+\alpha_k d_k) \leqslant f(x_k)$. 非精确意味着步长 α_k 不需要被精确计算, 只需满足特定的不等式条件即可. 我们仍然记 $\varphi(\alpha)=f(x_k+\alpha d_k)$, 则 $\varphi(0)=f_k$, $\varphi'(0)=g_k^{\mathrm{T}}d_k$.

Armijo 准则

$$\varphi(\alpha) \leqslant \varphi(0)+\rho\varphi'(0)\alpha$$

其中 $\rho \in (0,1)$.

Goldstein 准则

$$\begin{aligned}\varphi(\alpha) &\leqslant \varphi(0)+\rho\varphi'(0)\alpha\\ \varphi(\alpha) &\geqslant \varphi(0)+(1-\rho)\varphi'(0)\alpha\end{aligned}$$

其中 $\rho \in (0,\frac{1}{2})$.

Wolfe 准则

$$\begin{aligned}\varphi(\alpha) &\leqslant \varphi(0)+\rho\varphi'(0)\alpha\\ \varphi'(\alpha) &\geqslant \sigma\varphi'(0)\end{aligned}$$

其中 $1>\sigma>\rho>0$.

强 Wolfe 准则

$$\begin{aligned}\varphi(\alpha) &\leqslant \varphi(0)+\rho\varphi'(0)\alpha\\ |\varphi'(\alpha)| &\leqslant -\sigma\varphi'(0)\end{aligned}$$

当 $\sigma \to 0$ 时, 强 Wolfe 准则接近精确线搜索. 是当 $\sigma \leqslant \rho$ 时, 满足 Wolfe 准则的点可能不存在.

引理 2.1 (非精确线搜索的性质) 假设 $\varphi'(0)<0$, 则

(1) 当 α 充分小时, Armijo 成立; Goldstein(2) 和 Wolfe(2) 不成立.

(2) 若 $\varphi(\alpha)$ 有下界, 则 α 充分大时, Goldstein(2) 成立; Armijo 不成立.

定理 2.3 (非精确线搜索步长的存在性) 假设 $\varphi'(0)<0$ 且 $\varphi(\alpha)$ 有下界, 则存在 $\alpha>0$ 满足 Goldstein 或强 Wolfe 准则.

由于 α 充分小时 Armijo 准则成立, 故可设 α_0 是使得 Armijo 准则成立的最大 α. 下面证明, 存在 $\alpha \in (0, \alpha_0]$ 使得 Goldstein(2) 或 Wolfe(2) 成立.

Goldstein 准则 若 $\forall \alpha \in (0, \alpha_0]$ 都有

$$\varphi(\alpha) < \varphi(0) + (1-\rho)\varphi'(0)\alpha$$

取 $\alpha = \alpha_0$, 有

$$\varphi(\alpha_0) - \varphi(0) = \rho\varphi'(0)\alpha_0 < (1-\rho)\varphi'(0)\alpha_0$$

这与 $\rho \in (0, \frac{1}{2})$ 矛盾!

强 Wolfe 准则 若 $\forall \alpha \in (0, \alpha_0]$ 都有

$$|\varphi'(\alpha)| > -\sigma\varphi'(0)$$

由 Armijo 准则和 Lagrange 中值定理, 存在 $\alpha \in (0, \alpha_0]$ 使得

$$\rho\varphi'(0)\alpha_0 = \varphi(\alpha_0) - \varphi(0) = \varphi'(\alpha)\alpha_0$$

故 $\varphi'(\alpha) = \rho\varphi'(0) \geqslant \sigma\varphi'(0)$. 矛盾!

定理 2.4 ($\boldsymbol{f(x)}$ 下降量的下界估计) 设 $f(x)$ 连续可微, 且 $g(x)$ 满足 Lipschitz 条件

$$\|g(x) - g(y)\| \leqslant L\|x - y\| \tag{2.6}$$

若 $f(x_k + \alpha d_k)$ 在 $\alpha > 0$ 时有下界, 则对满足 Wolfe 准则的 α_k, 有

$$f(x_k) - f(x_k + \alpha_k d_k) \geqslant \frac{\rho(1-\sigma)}{L}\|g_k\|^2\cos^2\theta_k \tag{2.7}$$

证明 令 $s_k = \alpha_k d_k$ 为下降步, 则只需证明:

$$f(x_k) - f(x_k + s_k) \geqslant \frac{\rho(1-\sigma)}{L}\|g_k\|^2\cos^2\theta_k$$

由 Wolfe 准则有

$$s_k^{\mathrm{T}} g(x_k + s_k) \geqslant \sigma s_k^{\mathrm{T}} g_k$$

因此

$$(\sigma - 1)s_k^{\mathrm{T}} g_k \leqslant s_k^{\mathrm{T}}(g(x_k + s_k) - g_k) \leqslant L\|s_k\|^2$$

而夹角条件给出

$$(\sigma - 1)s_k^{\mathrm{T}} g_k \geqslant (1-\sigma)\|s_k\|\|g_k\|\cos\theta_k$$

因此

$$\|s_k\| \geqslant \frac{1-\sigma}{L}\|g_k\|\cos\theta_k$$

由 Armijo 准则,

$$f(x_k)-f(x_k+s_k) \geqslant -\rho g_k^{\mathrm{T}} s_k \geqslant \rho\|g_k\|\|s_k\|\cos\theta_k \geqslant \frac{\rho(1-\sigma)}{L}\|g_k\|^2\cos^2\theta_k$$

定理 2.5 (非精确线搜索的收敛性：Wolfe) 设 $g(x)$ 在水平集 $L(x_0) := \{x : f(x) \leqslant f(x_0)\}$ 上满足 Lipschitz 条件, 且 d_k 与 $-g_k$ 的夹角 θ_k 满足

$$0 \leqslant \theta_k \leqslant \frac{\pi}{2}-\mu$$

则对于 Wolfe 准则, 或者对某个 k 有 $g_k=0$, 或者 $f_k \to -\infty$, 或者 $g_k \to 0$.

[Proof: 袁亚湘 p105-106]

定理 2.6 (非精确线搜索的收敛性：Goldstein) 设 $g(x)$ 在水平集 $L(x_0) := \{x : f(x) \leqslant f(x_0)\}$ 一致连续, 且 d_k 与 $-g_k$ 的夹角 θ_k 满足

$$0 \leqslant \theta_k \leqslant \frac{\pi}{2}-\mu$$

则对于 Goldstein 准则, 或者对某个 k 有 $g_k=0$, 或者 $f_k \to -\infty$, 或者 $g_k \to 0$.

[Proof: 袁亚湘 p103-104]

如果在某个无穷集 K 上有 $\|g_k\| \geqslant \varepsilon_1$, 则由夹角条件可以得到

$$-g_k^{\mathrm{T}} s_k = \|g_k\|\|s_k\|\cos\theta_k \geqslant \varepsilon_1\|s_k\|$$

由 Armijo 准则,

$$f_{k+1} \leqslant f_k + \rho g_k^{\mathrm{T}} s_k$$

故单步下降量满足

$$f_k - f_{k+1} \geqslant -\rho g_k^{\mathrm{T}} s_k \geqslant \rho\varepsilon_1\|s_k\|, \qquad \forall k \in K$$

由于 $\lim\limits_{k\to\infty}(f_k-f_{k+1})=0$, 故

$$\lim_{k\in K}\|s_k\|=0$$

即相邻迭代点的间距会在 $k \in K$ 时收敛到 0. 因此当 $k \in K$ 且 k 充分大时, 可以进行 Taylor 展开(进行更精细的估计)

$$f_{k+1}-f_k = s_k^{\mathrm{T}} g(\xi_k) = s_k^{\mathrm{T}} g_k + o(\|s_k\|)$$

由 Goldstein 准则,

$$f_{k+1} - f_k \geqslant (1-\rho) g_k^{\mathrm{T}} s_k$$

因此

$$\rho g_k^{\mathrm{T}} s_k \geqslant o(\|s_k\|)$$

矛盾! 因此 Goldstein 线搜索准则在夹角条件下必定收敛.

关于证明的评注: 在上面的证明中, 一个关键的步骤是利用 Taylor 展开进行更精细的估计. 由于所有非精确线搜索方法都需要用到 $f(x)$ 的一阶导数信息, 因此进行 Taylor 展开是必要的.

3 Newton 型方法

3.1 基本 Newton 方法

基本 Newton 方法的下降方向 d_k^N 由方程

$$G_k d_k^N = -g_k \tag{3.1}$$

给出. 当迭代格式由

$$x_{k+1} = x_k + d_k^N \tag{3.2}$$

给出时, 相应的方法称为基本 Newton 方法. 关于基本 Newton 方法的评注:

- 基本 Newton 方法可以看作是在每个点 x_k 处用一个二次函数

$$m_k(p) = f_k + g_k^{\mathrm{T}} p + p^{\mathrm{T}} G_k p \tag{3.3}$$

  来近似, 因此上述函数的极小点在 $p^* = -G_k^{-1} g_k$ 处取到. 因此下降方向取为

$$d_k^N = -G_k^{-1} g_k \tag{3.4}$$

- Newton 方法的一大优势是可以达到二阶收敛速度.

定理 3.1 (基本 Newton 方法的收敛性) 设 $f \in C^2$, 且 f 的 Hesse 矩阵满足 Lipschitz 条件. 若 x_0 充分接近 x^*, 且 G^* 正定, 则基本 Newton 方法具有 2 阶收敛速度.

直观上, 上述结果可以使用 Taylor 展开证明. 要证明 2 阶收敛, 只需证明

$$\|x_{k+1} - x^*\| \leqslant C\|x_k - x^*\|^2 \tag{3.5}$$

对某个常数 $C > 0$ 成立. 由于 $x_{k+1} = x_k + d_k^N$, 且 $d_k^N = -G_k^{-1} g_k$, 故

$$x_{k+1} - x^* = x_k - x^* - G_k^{-1} g_k \tag{3.6}$$

而

$$\begin{aligned} -g_k &= g(x^*) - g(x_k) \\ &= G_k(x^* - x_k) + O(\|x_k - x^*\|^2) \end{aligned}$$

故

$$-G_k^{-1} g_k = x^* - x_k + O(\|x_k - x^*\|^2) \tag{3.7}$$

故

$$\|x_{k+1} - x^*\| = O(\|x_k - x^*\|^2) \tag{3.8}$$

在实际中, 也可以把 Newton 方向与线搜索方法 (精确线搜索, Wolfe 准则) 结合起来使用. Newton 方法的优缺点:

1. 优点: 二次终止性, 二阶收敛速度
2. 缺点: 全局收敛性不好, 每步需要求解线性方程组

当 Hesse 矩阵 G 不是正定的情况下, Newton 方向 $d_k^N = -G_k^{-1} g_k$ 不一定是下降方向.

取定 $\eta > 0$ 是一个小常数, θ_k 为 d_k^N 与 $-g_k$ 的夹角, 我们可以使用下面的策略选择下降方向:

$$d_k = \begin{cases} d_k^N, & \cos\theta_k \geqslant \eta \\ -d_k^N, & \cos\theta_k \leqslant -\eta \\ -g_k, & \text{else} \end{cases}$$

当 d_k^N 不是下降方向时, 选取 $-d_k^N$ 下降即可. 基本 Newton 方法也可以收敛到鞍点.

定理 3.2 (基本 Newton 方法的鞍点收敛性) 考虑迭代格式 $x_{k+1} = x_k + \alpha_k d_k^N$, 其中 G_k 非奇异, $\alpha_k = \mathrm{sgn}[g_k^{\mathrm{T}} G_k^{-1} g_k]$. 假定

(1) $\{x_k\}$ 收敛到 x^*;

(2) $g_k^{\mathrm{T}} d_k \neq 0$

则 x^* 是 $f(x)$ 的稳定点, 且非局部极小点.

在上面的定理中, α_k 的取法是使得方向 $d_k = \alpha_k d_k^N$ 一定是下降方向, 即 $g_k^{\mathrm{T}} d_k^N < 0$. 若 x^* 是局部极大点, 则在 x^* 恒有

$$y^{\mathrm{T}} G(x) y < 0, \quad \forall x \in B(x^*), \quad y \neq 0$$

因此存在 $\eta > 0$, 使得
$$\lambda_{\max}(G(x)) \leqslant -\eta, \quad \forall x \in B(x^*)$$
因此
$$\begin{aligned} f(x_k - d_k^N) &\leqslant f(x_k) - g_k^{\mathrm{T}} d_k^N + \frac{1}{2}(d_k^N)^{\mathrm{T}} G_k d_k^N + o(\|d_k^N\|^2) \\ &= f_k + \frac{3}{2}(d_k^N)^{\mathrm{T}} G_k d_k^N + o(\|d_k^N\|^2) \\ &\leqslant f_k - \frac{3}{2}\eta\|d_k^N\|^2 + o(\|d_k^N\|^2) \end{aligned}$$
因此当 k 充分大时, 应有 $f(x_{k+1}) \leqslant f(x_k)$, 因此其收敛点 x^* 不可能是局部极大点, 矛盾!

Newton 方法是可能使得迭代收敛到鞍点的.

3.2 修正 Newton 方法

在实际应用 Newton 方法时, Hesse 矩阵 G 可能不是正定的, 给实际应用带来了很大的困难. 我们可以将其做一定的修正, 使其成为正定的矩阵.

3.2.1 特征值修正方法

设 G 的所有特征值为 $\lambda_1, \cdots, \lambda_n$, 特征向量为 $u_1, \cdots, u_n$, 则 G 可以表示为
$$G = \sum_{i=1}^{n} \lambda_i u_i u_i^{\mathrm{T}} \tag{3.9}$$
若取 $\beta_i = \min(|\lambda_i|, \delta)$, 则一个合理的对 G 进行修正的方式是
$$\bar{G} = \sum_{i=1}^{n} \beta_i u_i u_i^{\mathrm{T}} \tag{3.10}$$
然而, 在实际计算中求解每步迭代矩阵 G 的特征值既困难也不稳定, 因此一般不会被采用.

3.2.2 修正 Cholesky 分解

在本节我们考虑对一般的对称矩阵 $G \in \mathbb{R}^{n\times n}$ 的修正 Cholesky 分解. 对于正定对称的情形, Cholesky 分解一般由下式给出:
$$G = LDL^{\mathrm{T}} \tag{3.11}$$

但是当 G 不正定时, 上述算法可能数值不稳定, 因此我们考虑修正 Cholesky 分解

$$G + E = L^{\mathrm{T}} D L$$

其中 L 为单位下三角矩阵, D 为正定对角矩阵, E 为对角修正矩阵. $L, D, E \in \mathbb{R}^{n\times n}$ 的计算由以下算法给出:

Algorithm 2: 修正 Cholesky 分解

输入: 对称矩阵 $G \in \mathbb{R}^{n\times n}$, 参数 $\delta, \beta > 0$

输出: $L, D \in \mathbb{R}^{n\times n}$

```
for j = 1, ··· , n do
    for i = j + 1, ··· , n do
        g_ij = g_ij − Σ_{s=1}^{j−1} d_s l_is l_js
    end
    θ_j = max_{j<i⩽n} |g_ij|
    d_j = max(|g_jj|, θ_j^2/β^2, δ)
    for i = j + 1, ··· , n do
        l_ij = g_ij / d_j
        g_ii = g_ii − d_j l_ij^2
    end
end
```

上述算法的关键在于 d_{jj} 的选取保证 D 是对称正定的, 因此计算 $l_{ij} = g_{ij}/d_j$ 时不会发生数值不稳定性. 参数 β 应当选取满足

$$\beta^2 = \max\{\gamma, \xi/\sqrt{n^2-1}\}$$

其中 γ 是矩阵 $A \in \mathbb{R}^{n\times n}$ 的对角元的最大模, ξ 是 A 的非对角元的最大模. 将修正 Cholesky 分解应用于 Hessian 矩阵 G, 得到

$$G + E = LDL^{\mathrm{T}}$$

于是下降方向由

$$d_k = -(LDL^{\mathrm{T}})^{-1} g_k = -L^{-\mathrm{T}} D^{-1} L^{-1} g_k$$

给出. 为了避免算法收敛到鞍点, 引入如下的负曲率方法. 设对角矩阵 D, E 由

$$D = \mathrm{diag}(d_1, \cdots, d_n), \qquad E = \mathrm{diag}(e_1, \cdots, e_n)$$

给出.

Algorithm 3: Gill-Murray 负曲率方向

输入: 修正 Cholesky 分解 $G+E=LDL^{\mathrm{T}}$

输出: 负曲率方向 $d \in \mathbb{R}^n$

计算 $\psi_j = d_j - e_j,\ j=1,\cdots,n$

求下标 t 使得 $\psi_t = \min\limits_{1 \leqslant j \leqslant n} \psi_j$

if $\psi_t < 0$ **then**

  令 1_t 是仅有第 t 个元素为 1 的单位向量

  令 d 为 $L^{\mathrm{T}} d = 1_t$ 的解

  输出负曲率方向 $d \in \mathbb{R}^n$

else

  不输出结果

end

可以证明上述算法的输出一定是一个负曲率方向. 事实上, 容易计算

$$d^{\mathrm{T}} G d + e_t = d^{\mathrm{T}} L D L^{\mathrm{T}} d = 1_t^{\mathrm{T}} D 1_t = d_t$$

因此

$$d^{\mathrm{T}} G d = d_t - e_t = \psi_t < 0$$

即 $d \in \mathbb{R}^n$ 是负曲率方向.

根据, 在 β 的选取方式下, 若 G 有负特征值, 则一定存在一个 t 是的 $\psi_t < 0$.

最后, 我们使用如下的稳定 Newton 算法来实现对 Hesse 矩阵 $G \in \mathbb{R}^{n \times n}$ 的修正:

Algorithm 4: Gill-Murray 稳定 Newton 方法

输入: 梯度 $g \in \mathbb{R}^n$, Hesse 矩阵 $G \in \mathbb{R}^{n\times n}$, 允许误差 $\varepsilon > 0$

输出: 下降方向 $d \in \mathbb{R}^n$

对 G 进行修正 Cholesky 分解 $G + E = LDL^{\mathrm{T}}$

(用于判断使用 Newton 方向还是负曲率方向)

令 $c = true$

if $\|g_k\| \leqslant \varepsilon$ **then**

  求解 Gill-Murray 负曲率方向 d_k

  如果不输出结果, 令 $c = true$

  (输出负曲率方向)

  如果成功输出, 令 $c = false$. 令 $g^{\mathrm{T}} d > 0$, 令 $d = -d$

end

if c **then**

  (输出 Newton 方向)

  求解下降方向 $LDL^{\mathrm{T}} d = -g$ 得到下降方向 d.

end

3.3 非精确 Newton 算法

基本 Newton 算法要求在每一步求解方程

$$G_k d_k^N = -g_k \tag{3.12}$$

如果 G_k 的维数太大, 那么线性方程组的求解会变得困难. 在实际中我们可以对该方程进行近似求解. 如果近似的下降方向为 d_k, 那么 d_k 对 Newton 方向 d_k^N 的逼近程度可以用残量

$$r_k = G_k d_k + g_k \tag{3.13}$$

来刻画. 如果残量满足

$$\|r_k\| \leqslant \eta_k \|g_k\| \tag{3.14}$$

对某个小的常数 η_k 成立, 则收敛速度可以得到保证.

定理 3.3 (非精确 Newton 局部收敛) 设 G^* 正定, 且下降方向 d_k 满足残量

$$r_k = G_k d_k + g_k, \qquad \|r_k\| \leqslant \eta_k \|g_k\|$$

且 $\eta_k \leqslant \eta$, $\eta \in (0, 1)$. 则对于迭代更新格式

$$x_{k+1} = x_k + d_k$$

若 x_0 充分接近 x^*, 则 x_k 线性收敛到 x^*.

我们从 $g(x)$ 的 Taylor 展开计算 g_{k+1} 的估计.

$$\begin{aligned} g_{k+1} - g_k &= G_k(x_{k+1} - x_k) + O(\|x_{k+1} - x_k\|^2) \\ &= G_k d_k + O(\|d_k\|^2) \\ &= r_k - g_k + O(\|d_k\|^2) \end{aligned}$$

因此

$$g_{k+1} = r_k + O(\|d_k\|^2) \tag{3.15}$$

由于

$$d_k = G_k^{-1}(r_k - g_k) \tag{3.16}$$

因此 $d_k = O(\|g_k\|)$. 因此我们得到

$$\|g_{k+1}\| \leqslant \eta\|g_k\| + o(\|g_k\|) \tag{3.17}$$

最后, 我们注意到从 g_k 的线性收敛速度可以得到 x_k 的收敛性. 由于

$$g_k - g^* = G^*(x_k - x^*) + O(\|g_k\|) \tag{3.18}$$

因此 $\|x_k - x^*\| = O(\|g_k\|)$. 故迭代序列 x_k 具有线性收敛速度. 由上面的定理证明也可以得到

定理 3.4 (非精确 Newton 超线性收敛) 若非精确 Newton 方法产生的迭代序列 x_k 收敛到 x^*, 并且 $\eta_k \to 0$, 则该方法具有超线性收敛速度, 即

$$\lim_{k\to\infty} \frac{\|g_{k+1}\|}{\|g_k\|} = 0 \tag{3.19}$$

事实上, 上述极限结果可以由不等式

$$\|g_{k+1}\| \leqslant \eta_k\|g_k\| + o(\|g_k\|) \tag{3.20}$$

直接得到.

3.4 全局非精确 Newton 方法

在全局非精确 Newton 方法中, 除了基本的判定准则 $\|r_k\| \leqslant \eta_k\|g_k\|$ 之外, 我们会加上一条额外的额外的准则.

定义 3.1 (全局非精确 Newton 方向) 给定常数 $t \in (0,1)$ 并考虑迭代格式 $x_{k+1} = x_k + d_k$, 则 d_k 称为非精确 Newton 方向, 如果

$$\begin{cases} \|r_k\| \leqslant \eta_k \|g_k\| \\ \|g_{k+1}\| \leqslant \big(1 - t(1-\eta_k)\big)\|g_k\| \end{cases} \tag{3.21}$$

容易看出, 满足上述条件的下降步 d_k 是存在的. 事实上, 当 k 充分大时, 有

$$\|g_{k+1}\| \leqslant \eta_k \|g_k\| + o(\|g_k\|) \tag{3.22}$$

因此 (3.21) 对一切 $t < 1$ 都存在解. 来使得

$$ared_k(d_k) \geqslant t \cdot pred_k(d_k) \tag{3.23}$$

定义 3.2 设当前迭代点为 x_k, 下降步为 $d \in \mathbb{R}^n$, 则实际下降量 (actual reduction) 定义为

$$ared_k(d) = \|g_k\| - \|g(x_k + d)\| \tag{3.24}$$

而预计下降量 (predicted reduction) 定义为

$$pred_k(d) = \|g_k\| - \|g_k + G_k d\| \tag{3.25}$$

当条件 (3.21) 满足, 即 d 是全局非精确 Newton 方向时,

$$pred_k(d) \geqslant (1-\eta_k)\|g_k\| \tag{3.26}$$

$$ared_k(d) \geqslant t(1-\eta_k)\|g_k\| \tag{3.27}$$

因此全局非精确 Newton 方法对实际和预计的下降量都提出了要求.

在非精确 Newton 方法中, 对参数 $\{\eta_k\}$ 的一些评注:

1. η_k 是刻画求解线性方程组 $G_k d_k = -g_k$ 的精度. η_k 越小, 线性方程组的求解误差就越小.
2. 如果 η_k 取得太小, 则不仅会增加求解线性方程组的复杂度, 还会降低数值解的有效性.
3. 增加一条对 d_k 的约束主要是为了保证 $ared_k(d)$ 有有效的下降.

我们考虑 2 种选择 η_k 的方式.

1. 选择

   $$\eta_k = \frac{\|g_k - g_{k-1} - G_{k-1}d_{k-1}\|}{\|g_{k-1}\|} \tag{3.28}$$

   或

   $$\eta_k = \frac{\big|\|g_k\| - \|g_{k-1} + G_{k-1}d_{k-1}\|\big|}{\|g_{k-1}\|} \tag{3.29}$$

注意到有不等式

$$\big|\|a\|-\|b\|\big| \leqslant \|a-b\|$$

故第一种方式得到的 η_k 更大. η_k 描述了 $g_{k-1}+G_{k-1}d_{k-1}$ 对 g_k 的近似程度. 使用第一种方式可以得到很好的局部收敛性:

定理 3.5 若 x_0 充分接近 x^*, 且 $\{\eta_k\}$ 由上面的式子决定, 则 $x_k \to x^*$, 且

$$\|x_{k+1}-x^*\| \leqslant \beta\|x_k-x^*\|\|x_{k-1}-x^*\| \tag{3.30}$$

这一定理意味着 $\{x_k\}$ 具有超线性的收敛速度, 且收敛阶为

$$\frac{\sqrt{5}+1}{2} \approx 1.618 > 1$$

2. 设 $\gamma \in [0,1]$ 且 $\alpha \in (1,2]$, 选择

$$\eta_k = \gamma\left(\frac{\|g_k\|}{\|g_{k-1}\|}\right)^{\alpha}, \quad k=1,2,\cdots$$

它反映了 $\|g_k\|$ 的下降的速度.

在实际更新 $\{\eta_k\}$ 时, 我们也会对 η_k 添加一些安全保护机制, 以防止 η_k 的减小速度太快.

$$\eta_k = \max\{\eta_k, \eta_{k-1}^{\frac{\sqrt{5}+1}{2}}\}, \qquad \eta_k = \max\{\eta_k, \gamma\eta_{k-1}^{\alpha}\}$$

Algorithm 5: Inexact Newton Backtracking Method

输入: x_0, $\eta_{\max} \in [0,1)$, $t \in (0,1)$, $0 < \theta_{\min} < \theta_{\max} < 1$

输出: 迭代点序列 $\{x_k\}$

for $k=1,2,\cdots$ **do**

  选择初始 η_k, 并选择下降步 d_k 满足

$$\|g_k + G_k d_k\| \leqslant \eta_k \|g_k\|$$

  while $\|g(x_k+d_k)\| > (1-t(1-\eta_k))\|g_k\|$ **do**

    选择 $\theta \in [\theta_{\min}, \theta_{\max}]$

    令 $d_k = \theta_k d_k$, $\eta_k = 1-\theta(1-\eta_k)$

  end

  令 $x_{k+1} = x_k + d_k$

end

4 拟 Newton 算法

4.1 拟 Newton 算法简介

拟 Newton 算法的主要目标是将 Newton 方法的思想用于求解大规模的无约束优化问题. 修正 Newton 算法和拟 Newton 算法的区别在于:

- 修正 Newton 算法主要处理 Hesse 矩阵 $G \in \mathbb{R}^{n\times n}$ 不正定的情形, 它通过在 G 上加入额外的正数项以使得 G 正定, 从而在求解线性方程组时不会产生数值不稳定. 由于修正 Newton 算法本质上仍然需要求解线性方程组, 因此只适用于中小规模的优化问题. 通过合适的迭代格式可以到达超线性甚至二阶收敛.

- 拟 Newton 算法主要考虑将类似于 Newton 算法的思想应用于大规模的优化问题中, 本质上不需要利用二阶导数, 和基本 Newton 方法的步骤差别较大. 只能达到一阶收敛速度.

在拟 Newton 算法中, 我们希望用一个低秩或容易计算的矩阵近似 Hesse 矩阵 G_k. 例如, 可以用矩阵 B_k 对 G_k 进行近似, 或用 H_k 对 G_k^{-1} 进行近似, 即

$$G_k \approx B_k, \quad G_k^{-1} \approx B_k^{-1}$$

按照 Newton 算法的思想, 我们在每个迭代点 x_k 处用二次函数

$$m_k(p) = f_k + g_k^{\mathrm{T}} p + \frac{1}{2} p^{\mathrm{T}} B_k p \tag{4.1}$$

来近似目标函数, 则我们希望 $\nabla m_{k+1}(-s_k) = g_k$, 即上一步的梯度为 g_k. 它给出

$$g_{k+1} - B_{k+1} s_k = g_k \Longrightarrow B_{k+1} s_k = y_k \tag{4.2}$$

其中 $s_k = x_{k+1} - x_k$, $y_{k+1} - y_k$.

定理 4.1 (割线方程) 在 x_{k+1} 处选择的拟 Newton 矩阵 $B_{k+1}(H_{k+1}) \in \mathbb{R}^{n\times n}$ 应当满足

$$B_{k+1} s_k = y_k \Longleftrightarrow H_{k+1} y_k = s_k \tag{4.3}$$

其中 $s_k = x_{k+1} - x_k$, $y_{k+1} - y_k$.

割线方程是拟 Newton 算法应满足的最基本的准则. 具有代表性的拟 Newton 算法包括:

1. DFP 算法:

$$B_{k+1} = (I - \rho_k y_k s_k^{\mathrm{T}}) B_k (I - \rho_k s_k y_k^{\mathrm{T}}) + \rho_k y_k y_k^{\mathrm{T}}$$

   其中 $\rho_k = 1/y_k^{\mathrm{T}} s_k$. 可以等价写为

$$H_{k+1} = H_k - \frac{H_k y_k y_k^{\mathrm{T}} H_k}{y_k^{\mathrm{T}} H_k y_k} + \frac{s_k s_k^{\mathrm{T}}}{y_k^{\mathrm{T}} s_k}$$

2. BFGS 算法:

$$H_{k+1} = (I - \rho_k s_k y_k^{\mathrm{T}}) H_k (I - \rho_k y_k s_k^{\mathrm{T}}) + \rho_k s_k s_k^{\mathrm{T}}$$

可以等价写为

$$B_{k+1} = B_k - \frac{B_k s_k s_k^{\mathrm{T}} B_k}{s_k^{\mathrm{T}} B_k s_k} + \frac{y_k y_k^{\mathrm{T}}}{y_k^{\mathrm{T}} s_k}$$

DFP 算法可以由下面的最优化问题得到:

定理 4.2 (DFP 的推导) 给定对称矩阵 $B_0 \in \mathbb{R}^{n\times n}$ 和向量 $s, y \in \mathbb{R}^n$, 设对称正定矩阵 $W \in \mathbb{R}^{n\times n}$ 满足 $Wy = s$. 则约束优化问题

$$\begin{aligned} \min_{B} \quad & \|B - B_0\|_W \\ \text{s.t.} \quad & B = B^{\mathrm{T}},\ Bs = y \end{aligned}$$

有唯一的最优解

$$B = (I - \rho y s^{\mathrm{T}}) B_0 (I - \rho s y^{\mathrm{T}}) + \rho y y^{\mathrm{T}}$$

其中 $\rho = 1/y^{\mathrm{T}} s$, 而 $\|\cdot\|_W$ 是加权的 Frobenius 范数:

$$\|A\|_W = \left\| W^{\frac{1}{2}} A W^{\frac{1}{2}} \right\|_F, \qquad \forall A \in \mathbb{R}^{n\times n}$$

由于此问题为凸优化问题, 因此 B 是约束优化问题的局部极小点当且仅当 B 是 KKT 点. 容易计算问题的 Lagrange 函数为

$$L(B, \lambda, \Lambda) = \left\| W^{\frac{1}{2}} (B - B_0) W^{\frac{1}{2}} \right\|_F^2 - \lambda^{\mathrm{T}} (Bs - y) - \mathrm{Tr}\big[\Lambda, (B - B^{\mathrm{T}})\big]$$

因此

$$\frac{\partial L}{\partial B} = W(B - B_0)W - \lambda s^{\mathrm{T}} - (\Lambda - \Lambda^{\mathrm{T}}) = 0$$

即 KKT 条件转化为

$$W(B - B_0)W = \lambda s^{\mathrm{T}} + (\Lambda - \Lambda^{\mathrm{T}})$$

因此

$$2W(B - B_0)W = \lambda s^{\mathrm{T}} + s \lambda^{\mathrm{T}}$$

代入 $\lambda = W\mu$, $s = Wy$, 得到

$$2(B - B_0) = \mu y^{\mathrm{T}} + y \mu^{\mathrm{T}} \tag{4.4}$$

下面我们通过 (4.4) 来得到 B 的表达式. 在 (4.4) 两边同乘 s, s^{T}, 得到

$$\mu = \rho (2I - \rho y s^{\mathrm{T}})(y - B_0 s) \tag{4.5}$$

故

$$\frac{1}{2}(\mu y^{\mathrm{T}} + y\mu^{\mathrm{T}}) = -\rho B_0 s y^{\mathrm{T}} - \rho y s^{\mathrm{T}} B_0 + \rho^2 y s^{\mathrm{T}} B_0 s y^{\mathrm{T}} + \rho y y^{\mathrm{T}} \tag{4.6}$$

因此

$$B = B_0 + \frac{1}{2}(\mu y^{\mathrm{T}} + y\mu^{\mathrm{T}}) = (I - \rho y s^{\mathrm{T}}) B_0 (I - \rho s y^{\mathrm{T}}) + \rho y y^{\mathrm{T}}$$

原命题成立. 利用 Sherman-Morrison 公式可以得到对应的 H_k 的更新公式.

类似地, 可以得到 BFGS 公式的推导公式:

定理 4.3 (BFGS 的推导) 给定对称矩阵 $H_0 \in \mathbb{R}^{n\times n}$ 和向量 $s, y \in \mathbb{R}^n$, 设对称正定矩阵 $G \in \mathbb{R}^{n\times n}$ 满足 $Gs = y$. 则约束优化问题

$$\begin{aligned} \min_{H} \quad & \|H - H_0\|_G \\ \text{s.t.} \quad & H = H^{\mathrm{T}},\ Hy = s \end{aligned}$$

有唯一的最优解

$$H = (I - \rho s y^{\mathrm{T}}) H_0 (I - \rho y s^{\mathrm{T}}) + \rho s s^{\mathrm{T}}$$

其中 $\rho = 1/y^{\mathrm{T}} s$, 而 $\|\cdot\|_W$ 是加权的 Frobenius 范数:

$$\|A\|_G = \left\| G^{\frac{1}{2}} A G^{\frac{1}{2}} \right\|_F, \quad \forall A \in \mathbb{R}^{n\times n}$$

由 BFGS 方法得到的矩阵 H_{k+1} 可以保持正定性.

4.2 有限内存 BGFS 方法

有限内存 BFGS 方法 (L-BFGS) 主要依赖于下面的 H_k 的更新公式:

$$H_{k+1} = (I - \rho_k s_k y_k^{\mathrm{T}}) H_k (I - \rho_k y_k s_k^{\mathrm{T}}) + \rho_k s_k s_k^{\mathrm{T}} \tag{4.7}$$

若定义 $V_k = I - \rho_k y_k s_k^{\mathrm{T}}$, 则更新公式成为

$$H_{k+1} = V_k^{\mathrm{T}} H_k V_k + \rho_k s_k s_k^{\mathrm{T}} \tag{4.8}$$

有限内存 BFGS 方法的思想是仅使用前几步的信息估计第 k 步的矩阵 H_k. 简单来说, L-BFGS 可归结为如下问题:

定义 4.1 (L-BFGS) 给定 $H_0 \in \mathbb{R}^{n\times n}$, $g \in \mathbb{R}^n$, 向量 $\{s_0, s_1, \cdots, s_{k-1}\}$ 和 $\{y_0, y_1, \cdots, y_{k-1}\}$, $g \in \mathbb{R}^n$, 希望计算 $H_k g$.

即用前 k 步的向量 s_i, y_i 估计矩阵-向量乘积 $H_k g$. 实际中, 初始矩阵可以取为 $H_0 = \gamma_0 I$,

$$\gamma_0 = \frac{s_0^{\mathrm{T}} y_0}{y_0^{\mathrm{T}} y_0}$$

4.2.1 Two-Loop Recursion

定义 $\{q_i\}_{i=0}^k \subset \mathbb{R}^n$ 如下:

$$q_k = g, \quad q_i = (I - \rho_i y_i s_i^{\mathrm{T}}) q_{i+1}, \quad i = k-1, \cdots, 0$$

则由更新公式 (4.8) 可以得到

$$H_{i+1} q_{i+1} = V_i^{\mathrm{T}} H_i V_i q_{i+1} + \rho_i s_i s_i^{\mathrm{T}} q_{i+1} = V_i^{\mathrm{T}} H_i q_i + \alpha_i s_i$$

其中我们定义 $\alpha_i = \rho_i s_i^{\mathrm{T}} q_{i+1}$, $i = 0, 1, \cdots, k-1$. 再定义 $\{z_i\}_{i=0}^k \subset \mathbb{R}^n$ 如下:

$$z_i = H_i q_i, \quad i = 0, 1, \cdots, k$$

则

$$z_{i+1} = (I - \rho_i s_i y_i^{\mathrm{T}}) z_i + \alpha_i s_i = z_i + s_i(\alpha_i - \beta_i)$$

其中我们定义 $\beta_i = \rho_i y_i^{\mathrm{T}} z_i$. 因此我们可以用如下的算法计算 $H_i g$:

Algorithm 6: Two-loop recursion for $H_k g$

输入: $H_0 \in \mathbb{R}^n$, $g \in \mathbb{R}^n$, $s_0, \cdots, s_{k-1}, y_0, \cdots, y_{k-1} \in \mathbb{R}^n$

输出: $H_k g \in \mathbb{R}^n$

令 $q_k = g$

for $i = k-1, \cdots, 0$ **do**

  $\alpha_i = \rho_i s_i^{\mathrm{T}} q_{i+1}$

  $q_i = q_{i+1} - \alpha_i y_i$

end

$z_0 = H_0 q_0$

for $i = 0, 1, \cdots, k-1$ **do**

  $\beta_i = \rho_i y_i^{\mathrm{T}} z_i$

  $z_{i+1} = z_i + s_i(\alpha_i - \beta_i)$

end

输出 $z_k = H_k g$

4.2.2 Compact Representation

我们将导出 BFGS 更新公式的压缩表示. 定义矩阵

$$S_k = [s_0, \cdots, s_{k-1}] \in \mathbb{R}^{n \times k}, \quad Y_k = [y_0, \cdots, y_{k-1}] \in \mathbb{R}^{n \times k}$$

以及 $R_k \in \mathbb{R}^{k\times k}$ 满足

$$(R_k)_{i,j} = \begin{cases} s_{i-1}^{\mathrm{T}} y_{j-1}, & i \leqslant j \\ 0, & \text{else} \end{cases}$$

则有如下引理成立:

引理 4.1 对于 $V_i = I - \rho_i y_i s_i^{\mathrm{T}} (i = 0, 1, \cdots, k-1)$, 有

$$V_0 V_1 \cdots V_{k-1} = I - Y_k R_k^{-1} S_k^{\mathrm{T}} \tag{4.9}$$

该引理的证明可以直接由数学归纳法得到. 当 $k = 1$ 时, 结论显然成立. 设结论对 k 成立, 则只需证明 $k+1$ 的情形, 即

$$(I - Y_k R_k^{-1} S_k^{\mathrm{T}}) V_k = I - Y_{k+1} R_{k+1}^{-1} S_{k+1}^{\mathrm{T}} \tag{4.10}$$

这等价于

$$Y_{k+1} R_{k+1}^{-1} S_{k+1}^{\mathrm{T}} = Y_k R_k^{-1} S_k^{\mathrm{T}} - \rho_k Y_k R_k^{-1} S_k^{\mathrm{T}} y_k s_k^{\mathrm{T}} + \rho_k y_k s_k^{\mathrm{T}} \tag{4.11}$$

下面我们只需证明 (4.11). 注意到 R_{k+1} 具有分块矩阵的形式

$$R_{k+1} = \begin{bmatrix} R_k & S_k^{\mathrm{T}} y_k \\ 0 & s_k^{\mathrm{T}} y_k \end{bmatrix} \tag{4.12}$$

因此 R_{k+1}^{-1} 也是分块对角矩阵

$$R_{k+1}^{-1} = \begin{bmatrix} R_k^{-1} & -\rho_k R_k^{-1} s_k^{\mathrm{T}} y_k \\ 0 & \rho_k \end{bmatrix} \tag{4.13}$$

故

$$\begin{aligned} Y_{k+1} R_{k+1}^{-1} S_{k+1}^{\mathrm{T}} &= \begin{bmatrix} Y_k & y_k \end{bmatrix} \begin{bmatrix} R_k^{-1} & -\rho_k R_k^{-1} s_k^{\mathrm{T}} y_k \\ 0 & \rho_k \end{bmatrix} \begin{bmatrix} S_k^{\mathrm{T}} \\ s_k^{\mathrm{T}} \end{bmatrix} \\ &= Y_k R_k^{-1} S_k^{\mathrm{T}} - \rho_k Y_k R_k^{-1} S_k^{\mathrm{T}} y_k s_k^{\mathrm{T}} + \rho_k y_k s_k^{\mathrm{T}} \end{aligned}$$

故 (4.11) 成立. 由数学归纳法, 命题得证.

定理 4.4 设 H_0 对称正定且 $\{s_i, y_i\}_{i=0}^{k-1}$ 满足 $s_i^{\mathrm{T}} y_i > 0$, 则 BFGS 更新的第 k 步由下式给出:

$$H_k = H_0 + \begin{bmatrix} S_k & H_0 Y_k \end{bmatrix} \begin{bmatrix} R_k^{-\mathrm{T}} (D_k + Y_k^{\mathrm{T}} H_0 Y_k) R_k^{-1} & -R_k^{-\mathrm{T}} \\ -R_k^{-1} & 0 \end{bmatrix} \begin{bmatrix} S_k^{\mathrm{T}} \\ Y_k^{\mathrm{T}} H_0 \end{bmatrix} \tag{4.14}$$

定理的证明参考文献. B_k 的更新也有相应的压缩表示:

定理 4.5 设 B_0 对称正定且 $\{s_i, y_i\}_{i=0}^{k-1}$ 满足 $s_i^{\mathrm{T}} y_i > 0$, 则 BFGS 更新的第 k 步由下式给出:

$$B_k = B_0 - \begin{bmatrix} B_0 S_k & Y_k \end{bmatrix} \begin{bmatrix} S_k^{\mathrm{T}} B_0 S_k & L_k \\ L_k^{\mathrm{T}} & -D_k \end{bmatrix}^{-1} \begin{bmatrix} S_k^{\mathrm{T}} B_0 \\ Y_k^{\mathrm{T}} \end{bmatrix} \tag{4.15}$$

其中 $L_k \in \mathbb{R}^{k \times k}$ 由下式给出:

$$(L_k)_{i,j} = \begin{cases} s_{i-1}^{\mathrm{T}} y_{j-1}, & i > j \\ 0, & \text{else} \end{cases}$$

5 信赖域方法

信赖域方法 (trust region) 主要用于求解无约束问题

$$\min_{x \in \mathbb{R}^n} f(x) \tag{5.1}$$

信赖域方法在迭代点 x_k 处用二次函数 $m_k(p)$ 对目标函数 $f(x_k + p)$ 进行近似. 根据 $f(x)$ 在 x_k 处的 Taylor 展开,

$$f(x_k + p) = f_k + g_k^{\mathrm{T}} p + \frac{1}{2} p^{\mathrm{T}} \nabla^2 f(x_k) p + o(\|p\|^2) \tag{5.2}$$

其中 $f_k = f(x_k)$, $g_k = \nabla f(x_k)$, $p \in \mathbb{R}^n$ 表示该步信赖域的下降步. 因此 $m_k(x)$ 可取为

$$m_k(p) = f_k + g_k^{\mathrm{T}} p + \frac{1}{2} p_k^{\mathrm{T}} G_k p_k \tag{5.3}$$

其中 $G_k \in \mathbb{R}^{n \times n}$ 是 Hesse 矩阵 $\nabla^2 f(x_k)$.

5.1 信赖域方法的基本框架

信赖域方法通过求解一系列信赖域子问题来解原优化问题. 在每个 x_k 处, 信赖域子问题形如

$$\min_{p \in \mathbb{R}^n} \ m_k(p) = f_k + g_k^{\mathrm{T}} p + \frac{1}{2} p^{\mathrm{T}} G_k p \tag{5.4}$$

$$\text{s.t.} \ \ \|p\| \leqslant \Delta_k \tag{5.5}$$

其中 $\Delta_k > 0$ 为信赖域半径, 它刻画了二次函数 $m_k(p)$ 对目标函数 $f(x_k + p)$ 的逼近有多精确. 我们用下面的实际下降量 (actual reduction) 和预测下降量 (predicted reduction) 的比值 ρ_k 来定量的刻画逼近的精确性:

$$\rho_k = \frac{f(x_k) - f(x_k + p)}{m_k(0) - m_k(p)} \tag{5.6}$$

我们可以利用ρ_k 的大小来决定信赖域半径 Δ_k 的更新方式:

- 若 $\rho_k \approx 1$, 则上述逼近是精确的, 可以增大信赖域半径.
- 若 ρ_k 接近 0 或为负值, 则应该减少 Δ_k 的值.
- 若 $\rho_k < 0$, 则该步的结果应当被舍弃.

注意: 老师的讲稿上使用的符号是

$$g_k(d) = f_k + g_k^{\mathrm{T}} d + \frac{1}{2} d^{\mathrm{T}} G_k d$$

其中 $G_k \in \mathbb{R}^{n\times n}$ 是 Hesse 矩阵, $d \in \mathbb{R}^n$ 是下降步.

Algorithm 7: 信赖域方法

输入: 最大半径 $\hat{\Delta} > 0$, 初始半径 $\Delta_0 > 0$, $\eta \in [0, 1/4)$

输出: 迭代点序列 $\{x_k\}_{k\geqslant 0}$

for $k = 1, 2, \cdots$ **do**

  近似求解信赖域子问题 (5.5) 得到下降步 p_k;

  按照 (5.6) 计算下降量比值 ρ_k;

  if $\rho_k < \frac{1}{4}$ **then**

    $\Delta_{k+1} = \frac{1}{4}\Delta_k$

  else

    if $\rho_k > \frac{3}{4}$ 且 $\|p_k\| = \Delta_k$ **then**

      $\Delta_{k+1} = \min(2\Delta_k, \hat{\Delta})$

    else

      $\Delta_{k+1} = \Delta_k$

    end

  end

  if $\rho_k > 0$ **then**

    $x_{k+1} = x_k + p_k$

  else

    $x_{k+1} = x_k$

  end

end

关于此算法的一些评注:

- 在上述算法中, 信赖域半径 Δ_k 增加当且仅当 p_k 达到了边界 $B(0, \Delta_k)$.
- 信赖域算法可以保证每一步目标函数值至少是下降的. 如果目标函数上升, 即 $\rho_k \leqslant 0$, 则该步会被直接拒绝掉.
- 算法对常数 $\frac{1}{4}, \frac{3}{4}$ 不敏感.

5.2 Cauchy 点及性质

去掉子问题中的下标 k 得到子问题

$$\min_{p\in\mathbb{R}^n} m(p) = p + g^{\mathrm{T}}p + \frac{1}{2}p^{\mathrm{T}}Gp, \quad \text{s.t. } \|p\| \leqslant \Delta \tag{5.7}$$

一种简单的下降步 p 的取法是, 在负梯度方向 g 极小化 $m(p)$, 所得的结果即为 Cauchy 点 p^C.

定义 5.1 (Cauchy 点) 对于信赖域子问题 (5.7), Cauchy 点定义为

$$p^C \equiv -\tau \frac{\Delta}{\|g\|} g \tag{5.8}$$

其中 τ 的取法是

$$\tau = \begin{cases} 1, & \text{如果 } g^{\mathrm{T}}Gg \leqslant 0 \\ \min\left(\dfrac{\|g\|^3}{\Delta g^{\mathrm{T}}Gg}, 1\right), & \text{其它} \end{cases}$$

Cauchy 点是信赖域子问题在一维子空间 (负梯度方向) 上的最优解. 使用 Cauchy 点可以为 $m(p)$ 带来一定的下降量:

引理 5.1 (Cauchy 点的下降量) Cauchy 点 p^C 满足

$$m(0) - m(p^C) \geqslant \frac{1}{2}\|g\| \min\left(\Delta, \frac{\|g\|}{\|G\|}\right) \tag{5.9}$$

该引理的证明可参见 Nocedal 的 Lemma 4.3. 在实际设计线搜索方法时, 我们希望方法达到的下降量应当至少不比 Cauchy 点少.

5.3 信赖域方法的收敛性

信赖域方法的一大好处从它可以得到好的全局收敛性, 而这种收敛性主要

定理 5.1 (信赖域方法的全局收敛性) 设 $\Omega \subset \mathbb{R}^n$ 是有界集, 且 $x_k \in \Omega, \forall k \in B$. 若 $f \in C^2$ 且在迭代过程中 $\|G_k\| \leqslant M$, 则信赖域算法产生一个满足一阶和二阶必要条件的聚点x^∞.

关于此定理的评注:

- 此定理首先要求算法是稳定的: 即迭代点稳定在一个有界的区间当中.
- 对于无约束问题, 一阶必要条件意味着 $g(x^*) = 0$, 二阶必要条件意味着 $G(x^*)$ 半正定.

分 2 种情形讨论: $\inf\limits_k \Delta_k = 0$ 或 $\inf\limits_k \Delta_k > 0$.

- 在第一种情形中, 可以取一个子列 $\{k_i\}$, 使得存在无穷 i, 满足

$$\Delta_{k_i+1} = \frac{1}{4}\Delta_{k_i}, \qquad \Delta_{k_i+1} < \frac{1}{2}\Delta_{k_{i-1}+1}$$

因此存在一个子序列 $K \subset \mathbb{N}$, 使得

$$\rho_k < \frac{1}{4}, \qquad \Delta_{k+1} \to 0, \qquad k \in K \tag{5.10}$$

由于 $\Delta_{k+1} = \frac{1}{4}\Delta_k$, 故 $\Delta_k \to 0$. 由于每一步的下降步 p_k 满足 $\|p_k\| \leqslant \Delta_k$, 故

$$\|p_k\| \to 0, \qquad k \in K \tag{5.11}$$

取 x^∞ 为序列 $\{x_k\}_{k\in K}$ 的聚点. 我们来证明 $g^\infty = 0$ 以及 G^∞ 半正定.

反设 $g^\infty \neq 0$. 由于每一步的下降量至少为 Cauchy 点的下降量, 故当 $k \in K$ 时,

$$m_k(0) - m_k(p_k) \geqslant \frac{1}{2}\|g_k\| \min\left(\Delta_k, \frac{\|g_k\|}{M}\right) \tag{5.12}$$

当 k 充分大时, $g_k \to g^\infty$, 而 $\Delta_k \to 0$, 故

$$m_k(0) - m_k(p_k) \geqslant \frac{1}{2}\|g_k\|\Delta_k \tag{5.13}$$

从而

$$\frac{\|p_k\|^2}{m_k(0) - m_k(p_k)} \leqslant \frac{2\|p_k\|^2}{\Delta_k\|g_k\|} \leqslant \frac{2\|p_k\|}{\|g_k\|} \tag{5.14}$$

当 k 充分大时, $g_k \to g^\infty$, $p_k \to 0$, 因此

$$\frac{\|p_k\|^2}{m_k(0) - m_k(p_k)} \to 0, \qquad k \in K \tag{5.15}$$

但实际下降量满足

$$f(x_k + p_k) - f(x_k) = m_k(0) - m_k(p_k) + o(\|p_k\|^2)$$

故下降量的比

$$r_k = \frac{f(x_k + p_k) - f(x_k)}{m_k(0) - m_k(p_k)} = 1 + o(1), \qquad k \in K$$

故 $r_k \to 1$, 这与 $\Delta_{k+1} = \frac{1}{4}\Delta_k$ 矛盾! 因此必有 $g^\infty = 0$.

再证明 G^∞ 是半正定的. 若否, 则设 G^∞ 的最小负特征值为 λ. 在第 k 步迭代时, 令 v_k 是 G^∞ 的最小负特征向量, 且适当选择 v 的方向来保证

$$v^{\mathrm{T}} g_k \leqslant 0 \tag{5.16}$$

于是

$$m_k(0)-m_k(p_k)\geqslant m_k(0)-m_k(\Delta_k v_k)\geqslant -\frac{1}{2}\Delta_k^2 v_k^{\mathrm{T}} G_k v_k=\frac{1}{2}\Delta_k^2(-\lambda+o(1)) \tag{5.17}$$

再次利用

$$f(x_k+p_k)-f(x_k)=m_k(0)-m_k(p_k)+o(\|p_k\|^2)$$

可以得到

$$r_k=\frac{f(x_k+p_k)-f(x_k)}{m_k(0)-m_k(p_k)}=1+o(1),\qquad k\in K$$

矛盾! 因此 G^∞ 是半正定的.

- 在第二种情形中, 可以取一个子列 K 使得

  $$\rho_k\geqslant 0.25,\quad h_k\geqslant \bar{h}>0,\quad \forall k\in K$$

  恒成立. 由于

  $$f(x_1)-f(x^\infty)=\sum_{k=1}^{\infty}(f(x_k)-f(x_{k+1}))<\infty$$

  故 $\lim\limits_{k\to\infty} f(x_k)-f(x_{k+1})=0$. 由于当 $k\in K$ 时,

  $$\rho_k=\frac{f(x_k)-f(x_k+p_k)}{m_k(0)-m_k(p_k)}\geqslant\frac{1}{4}$$

  故

  $$m_k(0)-m_k(p_k)\leqslant 4\big(f(x_k)-f(x_k+p_k)\big)\to 0,\quad k\in K$$

  令在 x^∞ 处的极限二次函数为

  $$m^\infty(p)=f^\infty+p^{\mathrm{T}}g^\infty+\frac{1}{2}p^{\mathrm{T}}G^\infty p$$

  且 $\bar{p}$ 为 $m^\infty(p)$ 的极小点, 则

  $$m_k(\bar{p})\geqslant m_k(p_k)\geqslant m_k(0)+o(1)=f(x_k)+o(1)$$

  在上式中取 $k\to\infty$, 有

  $$m^\infty(\bar{p})\geqslant f(x^\infty)=f^\infty$$

  这意味着二次函数 $m^\infty(p)$ 在 $B(0,\bar{h})$ 上的最小值就是 f^∞. 这必然意味着 $g^\infty=0$ 以及 G^∞ 半正定.

注: 从定理的证明可以看出, 对任何一个满足条件的聚点 x^∞, 它都满足一阶和二阶必要条件.

定理 5.2 (信赖域方法的二阶收敛性) 若上述定理的聚点 x^∞ 满足 f 的 Hesse 矩阵 G^∞ 正定, 则对原序列 $\{x_k\}$ 有 $\rho_{\to}1$, $x_k \to x^\infty$, $glb(x_k) > 0$, 以及对充分大的 k, 有 $\|p_k\| \leqslant \Delta_k$. 此外, 收敛速度是二阶的.

上面的定理实际上说明了当二阶充分条件满足时, 收敛点 x^∞ 是唯一存在的.

5.4 Levenberg-Marquardt 方法

在介绍信赖域子问题求解的数值方法前, 我们先介绍一些理论结果. 考虑信赖域子问题 (5.7)

$$\min_{p\in\mathbb{R}^n} m(p) = f + g^{\mathrm{T}}p + \frac{1}{2}p^{\mathrm{T}}Gp, \quad \text{s.t. } \|p\| \leqslant \Delta$$

关于信赖域子问题 (5.7) 的最优解, 有如下定理成立:

定理 5.3 (子问题的最优解) 向量 $p^* \in \mathbb{R}^n$ 是子问题 (5.7) 的最优解当且仅当 p^* 为可行点, 且存在 $\lambda \geqslant 0$, 使得

$$\begin{aligned}(G+\lambda I)p^* &= -g \\ \lambda(\Delta - \|p^*\|) &= 0 \\ (G+\lambda I)\ &\text{半正定}\end{aligned} \tag{5.18}$$

我们可以使用利用 KKT 条件对此定理进行证明. 对给定的 $\lambda \in \mathbb{R}$, 定义 Lagrange 函数

$$L(p) = f + g^{\mathrm{T}}p + \frac{1}{2}p^{\mathrm{T}}(G+\lambda I)p = m(p) + \frac{\lambda}{2}\|p\|^2$$

则 $L(p)$ 是关于 p 的二次函数.

- 充分性: 若条件 (5.18) 成立, 则 p^* 是 $L(p)$ 在 $\mathbb{R}^n$ 上的极小点, 因此对任何 $p \in \mathbb{R}^n$, 有

  $$L(p) \geqslant L(p^*) \Longrightarrow m(p) + \frac{\lambda}{2}\|p\|^2 \geqslant m(p^*) + \frac{\lambda}{2}\|p^*\|^2$$

  因此

  $$m(p) \geqslant m(p^*) + \frac{\lambda}{2}(\|p^*\|^2 - \|p\|^2)$$

  下面分 2 种情形讨论.

  - 若 $\lambda = 0$, 则上式直接给出

    $$m(p) \geqslant m(p^*), \quad \forall \|p\| \leqslant \Delta$$

  - 若 $\lambda > 0$, 则 $\|p^*\| = \Delta$, 因此当 $\|p\| \leqslant \Delta$ 时,

    $$m(p) \geqslant m(p^*) + \frac{\lambda}{2}(\|p^*\|^2 - \|p\|^2) \geqslant m(p^*), \quad \forall \|p\| \leqslant \Delta$$

在以上情形中, 均有 $m(p) \geqslant m(p^*)$ 对 $\|p\| \leqslant \Delta$ 恒成立. 因此 (5.18) 给出的 p^* 是子问题的最优解.

- 必要性: 若 p^* 是最优解, 则对任何 $\|p\| \leqslant \Delta$, 有 $m(p) \geqslant m(p^*)$. 考虑约束函数

  $$c(p) = \frac{1}{2}(\Delta^2 - \|p\|^2)$$

  由于约束规范条件满足, 故 p^* 也是 KKT 点, 故存在 $\lambda \geqslant 0$ 使得

  $$\nabla m(p^*) = \lambda \nabla c(p^*) \Longrightarrow (G + \lambda I)p^* = -g$$

  即 $\nabla L(p^*) = 0$. 而互补性条件给出 $\lambda(\Delta - \|p^*\|) = 0$. 最后我们只需证明 $G + \lambda I$ 半正定. 事实上, 对于 Lagrange 函数 $L(p)$, 在 p^* 进行 Taylor 展开有

  $$L(p) = L(p^*) + \frac{1}{2}(p - p^*)^{\mathrm{T}}(G + \lambda I)(p - p^*)$$

  即

  $$m(p) = m(p^*) + \frac{\lambda}{2}(\|p^*\|^2 - \|p\|^2) + \frac{1}{2}(p - p^*)^{\mathrm{T}}(G + \lambda I)(p - p^*)$$

  根据 $m(p) \geqslant m(p^*)$, 有

  $$\frac{\lambda}{2}(\|p^*\|^2 - \|p\|^2) + \frac{1}{2}(p - p^*)^{\mathrm{T}}(G + \lambda I)(p - p^*) \geqslant 0, \quad \forall \|p\| \leqslant \Delta$$

  下面分 2 种情形分析:

  - 若 $\lambda = 0$, 则 $(p - p^*)^{\mathrm{T}} G (p - p^*) \geqslant 0$, 因此在一个半空间上 $v^{\mathrm{T}} G v \geqslant 0$, 故 G 半正定.
  - 若 $\lambda \neq 0$, 则 $\|p^*\| = \Delta$, 则我们取 $\|p\| = \Delta$, 可以得到

    $$\frac{1}{2}(p - p^*)^{\mathrm{T}}(G + \lambda I)(p - p^*) \geqslant 0, \quad \forall \|p\| = \Delta$$

    因此 $v^{\mathrm{T}}(G + \lambda I)v \geqslant 0$ 在一个半空间上成立.

  因此 $G + \lambda I$ 半正定.

关于上述定理的评注:

- 上述定理证明的关键是构造 Lagrange 函数 $L(p)$ 并讨论极小点处的性质. 由于 $L(p)$ 是二次函数, 故其 Taylor 展开式可以直接写出.

- 该定理的内容不完全等于 KKT 条件.

- 方程组

$$\begin{cases} (B+\lambda I)p = -g \\ \|p\| = \Delta \end{cases}$$

中有 $p \in \mathbb{R}^n$ 和 $\lambda \in \mathbb{R}$ 共 $n+1$ 个未知数和 $n+1$ 方程, 因此理论上可以显式求解.

从上述定理可以得到

推论 5.1 若 G 正定且 $\|G^{-1}g\| \leqslant \Delta$, 则极小点一定在 $p^* = -G^{-1}g$ 处取到. 否则, $m(p)$ 的最小值一定可以在 $\|p^*\| = \Delta$ 时取到.

下面给出基于上述定理的 Levenberg-Marquardt 方法:

Algorithm 8: Levenberg-Marquardt 方法

输入: 当前迭代点 x_k 和常数 λ_k

输出: 下步迭代点 x_{k+1} 和常数 λ_{k+1}

1. 分解 $G_k + \lambda_k I$ 并判断其是否正定. 若否, 令 $\lambda_k = 4\lambda_k$ 并重复此步.
2. 求解 LM 方程 $(G_k + \lambda_k I)p_k = -g_k$ 得到下降步 p_k
3. 计算下降量的比值

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)}$$

并利用 ρ_k 的大小决定 λ_k 的更新步骤:

- 若 $\rho_k < 0.25$, $\lambda_{k+1} = 4\lambda_k$
- 若 $\rho_k > 0.75$, $\lambda_{k+1} = \lambda_k/2$
- 其它情形: $\lambda_{k+1} = \lambda_k$

利用 ρ_k 的正负性决定 x_k 的更新步骤:

- 若 $\rho_k \leqslant 0$, $x_{k+1} = x_k$
- 其它情形: $x_{k+1} = x_k + p_k$

最后, 我们说明

$$\|p(\lambda)\| = \left\|(G + \lambda I)^{-1} p\right\|$$

是一个关于 λ 的递减函数.

定理 5.4 设 λ_1 是 $G \in \mathbb{R}^{n \times n}$ 的最小特征值, 则当 $\lambda > -\lambda_1$ 时, $\|p(\lambda)\|$ 关于 λ 单调递减.

由于

$$(G + \lambda I)p(\lambda) = -g$$

故对 λ 求导可得

$$p(\lambda)+(G+\lambda I)p'(\lambda)=0$$

因此

$$p'(\lambda)=-(G+\lambda I)^{-1}p(\lambda)$$

故

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\|p(\lambda)\|=\frac{p(\lambda)}{\|p(\lambda)\|}\cdot p'(\lambda)=-\frac{1}{\|p\|}p^{\mathrm{T}}(G+\lambda I)^{-1}p$$

当 $\lambda>-\lambda_1$ 时, $G+\lambda I$ 正定, 故由 $p\neq 0$, 可得

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\|p(\lambda)\|<0$$

因此 $\|p(\lambda)\|$ 在 $\lambda>-\lambda_1$ 时单调递减.

5.5 信赖域子问题的最优解分析

为简便起见, 假设对称矩阵 $G\in\mathbb{R}^{n\times n}$ 的正交对角化的形式为 $G=Q\Lambda Q^{\mathrm{T}}$, 其中

$$\Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$$

为 G 的所有特征值, 且 $\lambda_1\leqslant\cdots\leqslant\lambda_n$, 而

$$Q=[q_1,\cdots,q_n]\in\mathbb{R}^{n\times n}$$

为正交阵, 且 $q_1,\cdots,q_n$ 为 G 的全部特征向量.

下面我们根据 $G\in\mathbb{R}^{n\times n}$ 的正定性对最优解 p^* 的性质做进一步的分析.

1. 若 $G\in\mathbb{R}^{n\times n}$ 是正定的, 则对任何 $\lambda\geqslant 0$, 方程 $(G+\lambda I)p=-g$ 总存在解

   $$p(\lambda)=-(G+\lambda I)^{-1}g$$

   且 $\|p(\lambda)\|$ 在 $[0,+\infty)$ 严格单调下降. 当 $\lambda=0$ 时, $p(0)$ 恰好为 Newton 下降步

   $$p^N=-G^{-1}g$$

   因此, 信赖域子问题 (5.7) 的最优解必然满足以下情形之一:

   - $\left\|p^N\right\|\leqslant\Delta$, 最优解为 $p^*=p^N$;
   - $\left\|p^N\right\|>\Delta$, 最优解由 $p^*=p(\lambda)=-(G+\lambda I)^{-1}g$, 其中 $\lambda>0$ 使得 $\|p(\lambda)\|=\Delta$.

2. 若 $G \in \mathbb{R}^{n\times n}$ 不是正定的, 则信赖域子问题的最优解一定可以在边界上取到, 即 $\|p^*\| = \Delta$. 由于 λ 应使得 $G+\lambda I$ 半正定, 故 $\lambda \geqslant -\lambda_1$. 下面讨论方程 $(G+\lambda I)p = -g$ 的解的性质. 设 $I = \{i = 1, \cdots, n : \lambda_i = \lambda_1\}$ 为最小特征值 λ_1 的指标集.

   - 若 $q_j^{\mathrm{T}} g$ 对 $j \in I$ 不全为 0, 则方程的解可表示为
   $$p(\lambda) = -(G+\lambda I)^{-1} g = -\sum_{j=1}^{n} \frac{q_j^{\mathrm{T}} g}{\lambda_j + \lambda} q_j$$
   $\|p(\lambda)\|$ 在 $\lambda \to -\lambda_1$ 时取值为 $+\infty$, 即
   $$p_{LS} = \sum_{j=1}^{n} \frac{q_j^{\mathrm{T}} g}{\lambda_j - \lambda_1} q_j$$
   满足 $\|p_{LS}\| = +\infty$. 由于 $\|p(\lambda)\|$ 在 $(-\lambda_1, +\infty)$ 上单调递减趋于 0, 存在唯一的 $\lambda \in (-\lambda_1, +\infty)$ 使得 $\|p(\lambda)\| = \Delta$.

   - 若 $q_j^{\mathrm{T}} g = 0$, $j \in I$ 全为 0, 则方程
   $$(G - \lambda_1 I)p = -g$$
   有解, 且
   $$p_{LS} = -\sum_{j \notin I} \frac{q_j^{\mathrm{T}} g}{\lambda_j - \lambda_1} q_j$$
   是所有可能的解当中范数最小的, 即最小二乘解. 分 2 种情形讨论:
     - 若 $\lambda > -\lambda_1$, 则 $G+\lambda I$ 正定, 且方程的解为
     $$p(\lambda) = -\sum_{j \notin I} \frac{q_j^{\mathrm{T}} g}{\lambda_j + \lambda} q_j$$
     显然, 当 $\lambda \to -\lambda_1$ 时, $p(\lambda) \to p_{LS}$, 因此
     $$\|p(\lambda)\| \in (0, \|p_{LS}\|)$$
     - 若 $\lambda = -\lambda_1$, 则 $G - \lambda_1 I$ 半正定, 且方程的解为
     $$p(\tau) = -\sum_{j \notin I} \frac{q_j^{\mathrm{T}} g}{\lambda_j - \lambda_1} q_j + \sum_{j \in I} \tau_j q_j = p_{LS} + \sum_{j \in I} \tau_j q_j$$
     其中 $\tau_j, j \in I$ 可以是任意常数. 这实际上就是信赖域子问题求解中的hard case.

注意到 q_j 是单位正交的, 故

$$\|p(\tau)\|^2 = \|p_{LS}\|^2 + \sum_{j \in I} \tau_j^2 \geqslant \|p_{LS}\|^2$$

因此 $\|p_{LS}\|$ 的取值范围是

$$\|p(\tau)\| \in [\|p_{LS}\|, +\infty)$$

因此, 我们可以通过判断 $\|p_{LS}\|$ 的值和 Δ 的关系来判断最优解是否属于 hard case.

hard case 的本质特征是: 当信赖域子问题的解

$$p(\lambda) = (G + \lambda I)^{-1} g$$

无法到达边界 $\|p(\lambda)\| = \Delta$ 时, 信赖域方法会因为每次的迭代量太小而变得低效. 解决 hard case 的方法就是在原本的 $p(\lambda)$ 上加上额外的负曲率方向.

根据以上对子问题最优解的分析, 容易得到求解信赖域子问题的通用算法:

Algorithm 9: 信赖域子问题的求解算法

输入: $g \in \mathbb{R}^n$, $G \in \mathbb{R}^{n\times n}$, $\Delta > 0$

输出: 信赖域子问题的最优解 $p^* \in \mathbb{R}^n$

令 (λ_1, q) 为对称阵 $G \in \mathbb{R}^{n\times n}$ 的最小特征对, 其中 λ_1 为最小特征值, $\|q\| = 1$

$a = \max(0, -\lambda_1)$ (λ 的取值下界)

if $\lambda_1 > 0$ **then**

  (G 对称正定)

  $p^N = -G^{-1}g$ (Newton 下降方向)

  if $\|p^N\| \leqslant \Delta$ **then**

    $p^* = p^N$ (无约束最小点)

    return

  end

else if $q^{\mathrm{T}}g = 0$ **then**

  $p_{LS} = -(G - \lambda_1 I)^{-1}g$ 为最小二乘解

  if $\|p_{LS}\|^2 \leqslant \Delta$ **then**

    (hard case)

    $\tau = \sqrt{\Delta^2 - \|p_{LS}\|^2}$

    $p^* = p_{LS} + \tau q$

    return

  end

end

(normal cases)

$b = \|g\|/\Delta + a$ (λ 的取值上界)

在区间 $[a, b]$ 上求方程 $\|p(\lambda)\| = \Delta$ 的解, 输出解 $p^* = p(\lambda)$

上述算法对可能出现的各种情形都进行了处理, 最终信赖域子问题的求解化归为在一个有限区间 $[a, b]$ 上求解非线性方程 $\|p(\lambda)\| = \Delta$.

5.6 数值求解信赖域子问题的算法

基于以上对 hard case 的一些讨论, 我们介绍一些求解信赖域子问题 (5.7) 的数值算法. 注意 LM 方法仅包含对常数 λ_k 的简单修正, 因此没有对 hard case 进行额外的处理.

5.6.1 Hebden 方法

我们讨论非线性函数

$$\varphi(\lambda) = \|p(\lambda)\| - \Delta \tag{5.19}$$

在 $[a,b]$ 上的根 λ 的求解算法, 其中 $p(\lambda) = -(G+\lambda I)^{-1}g$. Hebden 方法的基本思想是用一个有理函数

$$\bar{\varphi}(\lambda) = \frac{\beta}{\alpha+\lambda} - \Delta \tag{5.20}$$

逼近 $\varphi(\lambda)$, 其中 $\alpha, \beta \in \mathbb{R}$. 给定 $\lambda \in \mathbb{R}$, 若取 $\alpha, \beta \in \mathbb{R}$ 使得 $\varphi(\lambda) = \bar{\varphi}(\lambda)$ 且 $\varphi'(\lambda) = \bar{\varphi}'(\lambda)$, 则

$$\alpha = -\lambda - \frac{\varphi(\lambda)+\Delta}{\varphi'(\lambda)}, \qquad \beta = (\varphi(\lambda)+\Delta)(\lambda+\alpha) \tag{5.21}$$

则由 $\bar{\varphi}(\lambda) = 0$ 可以得到

$$\lambda = \lambda - \frac{\varphi(\lambda)+\Delta}{\Delta}\frac{\varphi(\lambda)}{\varphi'(\lambda)} \tag{5.22}$$

显然, 此方程的不动点为 $\varphi(\lambda) = 0$ 的解, 因此我们得到 $\lambda^{(l)}$ 的迭代计算公式:

$$\lambda^{(l+1)} = \lambda^{(l)} - \frac{\varphi(\lambda^{(l)})+\Delta}{\Delta}\frac{\varphi(\lambda^{(l)})}{\varphi'(\lambda^{(l)})} \tag{5.23}$$

相应的数值算法为

Algorithm 10: Hebden 迭代: 在 $[a,b]$ 上求解 $\|p(\lambda)\| = \Delta$

输入: $g \in \mathbb{R}^n$, $G \in \mathbb{R}^{n\times n}$, $\Delta > 0$, 区间 $[a,b]$

输出: $\|p(\lambda)\| = \Delta$ 的解

$\lambda^{(0)} = -1$, $\lambda^{(1)} = \lambda$;

for $l = 1, 2, \cdots$ **do**

  $p_l = -(G+\lambda^{(l)}I)^{-1}g$

  $\varphi(\lambda^{(l)}) = \|p_l\| - \Delta$

  $\varphi'(\lambda^{(l)}) = -\frac{1}{\|p_l\|}p_l^{\mathrm{T}}(G+\lambda^{(l)}I)^{-1}p_l$

  计算下一步迭代点

$$\lambda^{(l+1)} = \lambda^{(l)} - \frac{\varphi(\lambda^{(l)})+\Delta}{\Delta}\frac{\varphi(\lambda^{(l)})}{\varphi'(\lambda^{(l)})}$$

  令 $\lambda^{(l)} = \max(\lambda^{(l)}, 1.01a)$;

end

输出 $p(\lambda) = p^{(l)}$.

5.6.2 More-Sorensen 方法

与 Hebden 方法不同, More-Sorensen 方法考虑用 Newton 迭代算法求解非线性方程

$$\varphi(\lambda) = \frac{1}{\|p(\lambda)\|} - \frac{1}{\Delta} \tag{5.24}$$

的解, 因为 $\varphi(\lambda)$ 在 $\lambda = -\lambda_1$ 处连续, 所以 $\varphi(\lambda)$ 具有良好的正则性. Newton 迭代格式为

$$\lambda^{(l+1)} = \lambda^{(l)} - \frac{\varphi(\lambda^{(l)})}{\varphi'(\lambda^{(l)})} \tag{5.25}$$

容易计算

$$\varphi(\lambda) = \frac{\Delta - \|p\|}{\Delta\|p\|}, \qquad \varphi'(\lambda) = \frac{1}{\|p\|^3} p^{\mathrm{T}}(G + \lambda I)^{-1} p$$

故

$$\frac{\varphi(\lambda)}{\varphi'(\lambda)} = \frac{\Delta - \|p\|}{\Delta} \cdot \frac{\|p\|^2}{p^{\mathrm{T}}(G + \lambda I)^{-1} p} \tag{5.26}$$

因此相应的数值算法为

Algorithm 11: More-Sorensen 迭代

输入: $g \in \mathbb{R}^n$, $B \in \mathbb{R}^{n\times n}$, $\Delta > 0$, $a \geqslant 0$, 满足 $\|p(\lambda)\| \leqslant \Delta$ 的初始点

输出: $\|p(\lambda)\| = \Delta$ 的解

$\lambda^{(0)} = -1$, $\lambda^{(1)} = \lambda$;

for $l = 1, 2, \cdots$ **do**

  如果 $|\lambda^{(l)} - \lambda^{(l-1)}| < 10^{-6}$, 跳出循环;

  做 Cholesky 分解 $G + \lambda^{(l)} I = R^{\mathrm{T}} R$;

  求解 $R^T R p_l = -g$, $R^{\mathrm{T}} q_l = p_l$, 然后计算下一步迭代点

$$\lambda^{(l+1)} = \lambda^{(l)} + \left(\frac{\|p_l\|^2}{\|q_l\|^2}\right)\left(\frac{\|p_l\| - \Delta}{\Delta}\right)$$

  令 $\lambda^{(l)} = \max(\lambda^{(l)}, 1.01a)$;

end

输出 $p(\lambda) = p^{(l)}$.

在上述方法中可以证明

$$\|q_l\|^2 = p_l^{\mathrm{T}}(G + \lambda^{(l)})^{-1} p_l$$

5.6.3 Dogleg 方法

Dogleg 方法是对 Cauchy 点的一个简单改进, 它只能应用于 Hesse 矩阵 $G \in \mathbb{R}^{n\times n}$ 对称正定的情形. 设

$$p^U = -\frac{g^{\mathrm{T}}g}{g^{\mathrm{T}}Gg}g$$

是负梯度方向上的无约束最小点, 而

$$p^N = -G^{-1}g$$

是 Newton 方法给出的最小点, 则 Dogleg 方法在一条折线

$$\tilde{p}(\tau) = \begin{cases} \tau p^U, & 0 \leqslant \tau \leqslant 1 \\ p^U + (\tau - 1)(p^N - p^U), & 1 \leqslant \tau \leqslant 2 \end{cases}$$

上计算目标函数的最小值. 可以证明 Dogleg 折线具有如下性质:

引理 5.2 (Dogleg 折线) 设 $G \in \mathbb{R}^{n\times n}$ 对称正定, 则

(1) $\|\tilde{p}(\tau)\|$ 关于 τ 单调递增;

(2) $m(\tilde{p}(\tau))$ 关于 τ 单调递减.

因此, 为了计算 Dogleg 点, 只需计算 Dogleg 折线与边界 $B(0, \Delta)$ 的交点.

5.6.4 二维子空间极小化方法

二维子空间极小化方法的思想是在 $\mathbb{R}^n$ 的一个二维子空间上求解 $f(x)$ 的最小值. 当 $B \in \mathbb{R}^{n\times n}$ 正定时, 考察由 $g, B^{-1}g$ 生成的二维子空间

$$\min_{p\in\mathbb{R}^n} m(p) = f + g^{\mathrm{T}}p + \frac{1}{2}p^{\mathrm{T}}Bp, \quad \text{s.t. } \|p\| \leqslant \Delta, p \in \mathrm{span}[g, B^{-1}g] \tag{5.27}$$

则 (5.27) 的解可以作为子问题 (5.7) 的近似解. 当 B 非正定时, 取定 $\alpha > 0$ 使得 $B + \alpha I$ 正定, 则相应的极小化问题为

$$\min_{p\in\mathbb{R}^n} m(p) = f + g^{\mathrm{T}}p + \frac{1}{2}p^{\mathrm{T}}Bp, \quad \text{s.t. } \|p\| \leqslant \Delta, p \in \mathrm{span}[g, (B + \alpha I)^{-1}g] \tag{5.28}$$

一般地, 二维子空间方法由如下算法给出:

Algorithm 12: 二维子空间极小化方法

输入: $g \in \mathbb{R}^n$, $B \in \mathbb{R}^{n\times n}$, $\Delta > 0$

输出: 信赖域子问题 (5.7) 的近似解

令 (λ_1, q) 为对称阵 $B \in \mathbb{R}^{n\times n}$ 的最小特征对, 其中 λ_1 为最小特征值, $\|q\| = 1$

取定常数 $\varepsilon_0 = 10^{-6}$

if $\lambda_1 \geqslant \varepsilon_0$ **then**

  (B 对称正定)

  $p^B = -B^{-1}g$ (Newton 下降方向)

  if $\|p^N\| \leqslant \Delta$ **then**

    $p = p^B$ (无约束极小点)

    return

  else

    $S = \mathrm{span}[g, -B^{-1}g]$ (构造二维子空间)

  end

else if $|\lambda_1| < \varepsilon_0$ **then**

  (B 几乎奇异)

  $p = p^C$ (Cauchy 点)

  return

else

  (B 为不定矩阵)

  选择常数 $\alpha = 1.5|\lambda_1| \in (-\lambda_1, -2\lambda_1]$

  $p = -(B + \alpha I)^{-1}g$

  if $\|p\| \leqslant \Delta$ **then**

    $\gamma = -p^{\mathrm{T}}q + \sqrt{(p^{\mathrm{T}}q)^2 + \Delta^2 - \|p\|^2}$

    $p = p + \gamma q$ (hard case 增加的额外项)

    return

  else

    $S = \mathrm{span}\{g, p\}$ (构造二维子空间)

  end

end

在二维子空间 S 上极小化二次函数 $m(p)$, 得到近似的最优解 p^*

在上述算法中, 我们对于 $\|(B + \alpha I)^{-1}g\| \leqslant \Delta$ 使用了类似于 hard case 的处理方式. 当 B 是不定矩阵时, 可以通过在

$$p = -(B + \alpha I)^{-1}g$$

加上负曲率方向 q 来使得 $\|p+\gamma q\|=\Delta$. 容易验证, 此式等价于方程

$$\gamma^2+2\gamma p^{\mathrm{T}}q=\Delta^2-\|p\|^2 \tag{5.29}$$

此方程的解恰好为 $\gamma=-p^{\mathrm{T}}q+\sqrt{(p^{\mathrm{T}}q)^2+\Delta^2-\|p\|^2}$.

在以上算法中, 仍需解决的问题是求解一般形式的二维子空间极小化问题

$$\min_{p\in\mathbb{R}^n} m(p)=f+g^{\mathrm{T}}p+\frac{1}{2}p^{\mathrm{T}}Bp, \quad \text{s.t. } \|p\|\leqslant\Delta, p\in\operatorname{span}[m_1,m_2] \tag{5.30}$$

其中 m_1,m_2 是二维子空间 S 的一个标准正交基. 利用 (m_1,m_2), 下降步 $p\in\mathbb{R}^n$ 可以表示为

$$p=xm_1+ym_2=Mq \tag{5.31}$$

其中 $M=[m_1,m_2]\in\mathbb{R}^{n\times 2}$, $q\in\mathbb{R}^2$. 因此 (5.30) 变为

$$m(q)=f+g^{\mathrm{T}}Mq+\frac{1}{2}q^{\mathrm{T}}M^{\mathrm{T}}BMq=f+U^{\mathrm{T}}q+\frac{1}{2}q^{\mathrm{T}}Gq \tag{5.32}$$

这里我们定义 $U=M^{\mathrm{T}}g\in\mathbb{R}^2$ 和 $G=M^{\mathrm{T}}BM\in\mathbb{R}^2$. 经过以下推导, 二维子空间极小化问题 (5.30) 简化为

$$\min_{q\in\mathbb{R}^2} m(q)=f+U^{\mathrm{T}}q+\frac{1}{2}q^{\mathrm{T}}Gq, \quad \text{s.t. } \|q\|\leqslant\Delta \tag{5.33}$$

它可以转化为一个四次多项式的根问题. 具体的细节这里不再赘述.

6 约束优化理论

6.1 约束优化的基本概念

考虑约束优化问题

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) \\ \text{s.t. } &c_i(x)=0, \quad i\in E \\ &c_i(x)\geqslant 0, \quad i\in I \end{aligned} \tag{6.1}$$

其中 $f(x)$ 是目标函数, $c_i(x)$ 是约束函数. $i\in E$ 时为等式约束, $i\in I$ 时为不等式约束.

定义 6.1 约束优化问题 (6.1) 的可行域为

$$X=\{x\in\mathbb{R}^n : c_i(x)=0, i\in E; c_i(x)\geqslant 0, i\in I\} \tag{6.2}$$

因而约束优化问题转化为 $f(x)$ 在 X 上的最小值的求解.

定义 6.2 若 $x^* \in X$ 满足

$$f(x) \geqslant f(x^*), \qquad \forall x \in X \tag{6.3}$$

则 x^* 是问题的全局极小点(最优解). 如果对 $x \neq x^*$ 都有

$$f(x) > f(x^*), \qquad \forall x \neq x^* \tag{6.4}$$

则 x^* 是问题的严格全局极小点(最优解).

类似地还可以定义局部极小点和严格局部极小点. 在给定的点 $x \in \mathbb{R}^n$ 处, 约束条件 $c_i(x)$ 称为积极的, 如果在此处 $c_i(x) = 0$. 引入定义

定义 6.3 对任何 $x \in \mathbb{R}^n$, 称下标集合

$$\mathcal{A}(x) = E \cup I(x) \tag{6.5}$$

为积极约束集合, 其中

$$I(x) = \{i : c_i(x) \leqslant 0\} \tag{6.6}$$

简而言之, 积极约束描述来优化问题的临界状态: 可行方向的范围受到积极约束的限制.

6.2 一阶必要条件

判断一个点 x^* 是否为局部极小点主要依赖于它在局部的性质, 因此可行方向在最优性条件中起到重要的作用.

定义 6.4 设 $x^* \in X$, $d \in \mathbb{R}^n$ 称为可行方向, 如果存在 $\delta > 0$ 使得

$$x^* + td \in X, \qquad \forall t \in [0, \delta] \tag{6.7}$$

x^* 处的所有可行方向的集合记为 $\mathrm{FD}(x^*, X)$.

沿着可行方向的一整条线段都落在可行域中.

定义 6.5 设 $x^* \in X$, $d \in \mathbb{R}^n$, x^* 称为线性化可行方向, 如果

$$d^{\mathrm{T}} c_i(x) = 0, \quad i \in E \tag{6.8}$$

$$d^{\mathrm{T}} c_i(x) \geqslant 0, \quad i \in I(x^*) \tag{6.9}$$

x^* 处的所有线性化可行方向记为 $\mathrm{LFD}(x^*, X)$.

在线性化可行方向上,

- 等式约束的 $c_i(x)$ 应当不变, 即 $d^{\mathrm{T}} c_i(x) = 0$;
- 有效不等式约束的 $c_i(x)$ 应当增加, 即 $d^{\mathrm{T}} c_i(x) \geqslant 0$;
- 无效不等式约束的 $c_i(x)$ 没有限制.

因此其定义如上所示.

定义 6.6 设 $x^* \in X$, $d \in \mathbb{R}^n$, 如果存在序列 d_k 和 $\delta_k \to 0$ 使得

$$x^* + \delta_k d_k \in X \tag{6.10}$$

且 $d_k \to d$, $\delta_k \to 0$, 则 d 称为 x^* 处的序列可行方向, 记为 $\mathrm{SFD}(x^*, X)$.

三种方式定义的可行方向有如下关系:

引理 6.1 若 $x^* \in X$ 且所有约束函数在 x^* 处可微, 则

$$\mathrm{FD}(x^*, X) \subset \mathrm{SFD}(x^*, X) \subset \mathrm{LFD}(x^*, X) \tag{6.11}$$

可行方向对 $d \in \mathbb{R}^n$ 的约束太强, 通常不会用到. 线性化可行方向对应于线性方程组的解, 在实际中较为常用. 序列可行方向在局部极小点的判断中较为有用.

引理 6.2 (一阶必要条件: 序列可行) 设 x^* 是局部极小点, 若 $f(x), c_i(x)$ 在 x^* 处可微, 则

$$d^{\mathrm{T}} \nabla c_i(x^*) \geqslant 0, \qquad \forall d \in \mathrm{SFD}(x^*, X) \tag{6.12}$$

上述定理揭示了局部极小点处序列可行方向的性质.

引理 6.3 (Farkas 引理) 设 $a_0, a_i (i = 1, \cdots, l), b_i (i = 1, \cdots, l') \in \mathbb{R}^n$, 则线性方程组

$$d^{\mathrm{T}} a_i = 0, \quad i = 1, \cdots, l \tag{6.13}$$

$$d^{\mathrm{T}} b_i \geqslant 0, \quad i = 1, \cdots, l' \tag{6.14}$$

$$d^{\mathrm{T}} a_0 < 0 \tag{6.15}$$

无解当且仅当存在实数 $\lambda_i (i = 1, \cdots, l)$ 和非负实数 $\mu_i (i = 1, \cdots, l')$ 使得

$$a_0 = \sum_{i=1}^{l} \lambda_i a_i + \sum_{i=1}^{l'} \mu_i b_i \tag{6.16}$$

在 Farkas 引理中, $d^{\mathrm{T}} a_i = 0$ 对应于等式约束, $d^{\mathrm{T}} b_i \geqslant 0$ 对应于不等式约束, $d^{\mathrm{T}} a_0 < 0$ 对应于下降方向的条件. 如果可行方向上没有下降方向, 则梯度方向 a_0 一定是约束梯度方向的线性组合, 并且不等式约束上的分量非负. 基于这种原因, 我们常常希望所谓的约束规范条件成立.

定义 6.7 (约束规范条件) 设 $x^* \in X$, 称 x^* 处约束规范条件满足, 如果

$$\mathrm{LFD}(x^*, X) = \mathrm{SFD}(x^*, X) \tag{6.17}$$

注意, 在各种可行方向和约束规范条件的定义中, 我们没有涉及目标函数的性质. 本质上说, 约束规范条件是对定义域性质和约束条件的刻画.

下面我可以得到局部极小点满足的一阶必要条件:

定理 6.1 (一阶必要条件: KKT) 设 x^* 是局部极小点, 如果 x^* 处约束规范条件满足, 即

$$\mathrm{SFD}(x^*, X) = \mathrm{LFD}(x^*, X) \tag{6.18}$$

则必存在 λ^* 使得

$$\nabla f(x^*) = \sum_{i \in E \cup I(x^*)} \lambda_i^* \nabla c_i(x^*) \tag{6.19}$$

且

$$\lambda_i^* \geqslant 0, \quad \lambda_i^* c_i(x^*) = 0, \quad i \in I \tag{6.20}$$

证明 根据一阶必要条件和约束优化规范条件,

$$d^{\mathrm{T}} \nabla c_i(x^*) \geqslant 0, \qquad \forall d \in \mathrm{LFD}(x^*, X) \tag{6.21}$$

故下面的线性方程组无解:

$$d^{\mathrm{T}} \nabla c_i(x^*) = 0, \quad i \in E \tag{6.22}$$

$$d^{\mathrm{T}} \nabla c_i(x^*) \geqslant 0, \quad i \in I(x^*) \tag{6.23}$$

$$d^{\mathrm{T}} \nabla f(x^*) < 0 \tag{6.24}$$

因此, 由 Farkas 引理, 存在 $\lambda_i^*(i \in E \cup I(x^*))$ 使得

$$\nabla f(x^*) = \sum_{i \in E} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in I(x^*)} \lambda_i^* \nabla c_i(x^*) \tag{6.25}$$

其中 $\lambda_i^* \geqslant 0, i \in I(x^*)$. 反过来, 我们也可以通过 KKT 条件来得到一阶必要条件:

$$d^{\mathrm{T}} \nabla f(x^*) = \sum_{i \in E} \lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) + \sum_{i \in I(x^*)} \lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) \geqslant 0 \tag{6.26}$$

因此 (6.21) 成立. 综上, KKT 条件等价于 (6.21). 我们可以使用直观的方式来理解 KKT 条件:

- 最优性: $\nabla f(x^*) = \sum_{i \in E \cup I} \lambda_i^* \nabla c_i(x^*)$

- 非负性: $\lambda_i^* \geqslant 0,\, i \in I$
- 互补性: $\lambda_i^* c_i(x^*) = 0,\, i \in I$

一种更直观的方式是利用有效约束改写 KKT 条件:

- 最优性: $\nabla f(x^*) = \sum\limits_{i \in \mathcal{A}(x^*)} \lambda_i^* \nabla c_i(x^*)$
- 非负性: $\lambda_i^* \geqslant 0,\, i \in I(x^*)$

KKT 条件里的 x^* 称为 KKT 点, λ^* 称为 Lagrange 乘子. 本质上, 我们只需要有效约束上的 λ_i^* 就可以判断是否为 KKT 点.

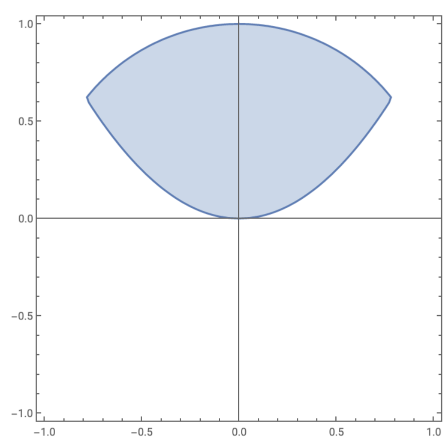
例 1 考虑约束优化问题

$$\min_{x \in \mathbb{R}^n} -x_1 - x_2 \tag{6.27}$$

$$\text{s.t.} \quad x_2 - x_1^2 \geqslant 0 \tag{6.28}$$

$$1 - x_1^2 - x_2^2 \geqslant 0 \tag{6.29}$$

求出该问题的所有 KKT 点.



在该问题中, $f(x) = -x_1 - x_2$, 约束函数

$$c_1(x) = x_2 - x_1^2, \qquad c_2(x) = 1 - x_1^2 - x_2^2$$

于是

$$\nabla f(x) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \nabla c_1(x) = \begin{bmatrix} -2x_1 \\ 1 \end{bmatrix}, \quad \nabla c_2(x) = \begin{bmatrix} -2x_1 \\ -2x_2 \end{bmatrix}$$

因此 KKT 条件给出

$$\begin{cases} -1 = -2x_1\lambda_1 - 2x_1\lambda_2 \\ -1 = \lambda_1 - 2x_2\lambda_2 \end{cases}$$

因此 $\lambda_2 \neq 0$, 否则 $\lambda_1 = -1 < 0$. 下面分别讨论 $\lambda_1 = 0$ 和 $\lambda_1 \neq 0$ 的情形:

- $\lambda_1 = 0$, 则 $x_1 = x_2$, 因此 $x_1 = x_2 = \pm\frac{1}{\sqrt{2}}$. 由于 $\lambda_2 > 0$, 故 $x_1 = x_2 = \frac{1}{\sqrt{2}}$, 因此

$$x^* = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \qquad \lambda^* = \left(0, \frac{1}{\sqrt{2}}\right)$$

- $\lambda_1 \neq 0$, 则两个约束条件都起作用, 有

$$1 = x_1^2 + x_1^4$$

  因此

$$x_1 = \pm\sqrt{\frac{\sqrt{5}-1}{2}}, \qquad x_2 = \frac{\sqrt{5}-1}{2}$$

  但 $x_1(\lambda_1 + \lambda_2) = 1 > 0$, 故 $x_1 > 0$, 从而

$$x_1 = \sqrt{\frac{\sqrt{5}-1}{2}}$$

  因此对应的点 x^* 和 Lagrange 乘子 λ^* 为

$$x^* = (a, a^2), \qquad \lambda^* = \left(\frac{a-1}{2a^2+1}, \frac{2a+1}{2a(2a^2+1)}\right)$$

  其中

$$a = \sqrt{\frac{\sqrt{5}-1}{2}} < 1$$

  因此 $\lambda_1^* < 0$, 故 x^* 不是 KKT 点!

综上, $x^* = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ 是唯一的 KKT 点.

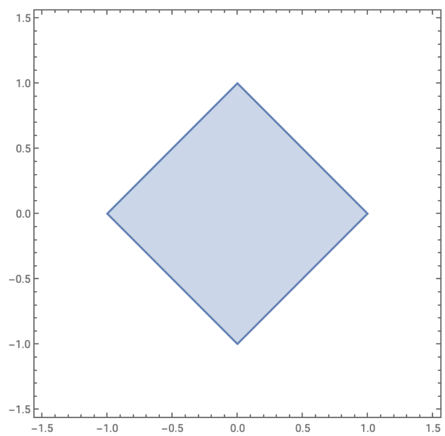
在下面的例子我们将看到, 也可以用 $d^{\mathrm{T}}\nabla f(x^*) \geqslant 0$ 判断给定的点是否为 KKT 点.

例 2 (Nocedal 12.15) Consider the following problem where t is a parameter to be fixed prior to solving the problem:

$$\min_x \left(x_1 - \frac{3}{2}\right)^2 + (x_2 - t)^4, \qquad \text{s.t.} \begin{bmatrix} 1 - x_1 - x_2 \\ 1 - x_1 + x_2 \\ 1 + x_1 - x_2 \\ 1 + x_1 + x_2 \end{bmatrix} \geqslant 0$$

(a) For what values of t does the point $x^* = (1,0)^{\mathrm{T}}$ satisfy the KKT conditions?

(b) Show that when $t = 1$, only the first constraint is active at the solution, and find the solution.



在此问题中, 目标函数为 $f(x) = (x_1 - \frac{3}{2})^2 + (x_2 - t)^4$, 约束函数 $c_1(x), c_2(x), c_3(x), c_4(x)$ 由上述不等式给出. 由于

$$\nabla f(x) = \begin{bmatrix} 2x_1 - 3 \\ 4(x_2 - t)^3 \end{bmatrix}$$

故在可行点 x 处的 KKT 条件为

$$\begin{cases} 2x_1 - 3 = -\lambda_1 - \lambda_2 + \lambda_3 + \lambda_4 \\ 4(x_2 - t)^3 = -\lambda_1 + \lambda_2 - \lambda_3 + \lambda_4 \end{cases}$$

(a) 在 $x^* = (1,0)$ 处, 约束条件 $c_3(x), c_4(x)$ 不起作用, 因此 $\lambda_3 = \lambda_4 = 0$, 于是可得

$$\begin{cases} -1 = -\lambda_1 - \lambda_2 \\ -4t^3 = -\lambda_1 + \lambda_2 \end{cases}$$

因此得到 Lagrange 乘子为

$$\lambda_1 = \frac{1}{2} + 2t^3, \qquad \lambda_2 = \frac{1}{2} - 2t^3$$

根据 $\lambda_1, \lambda_2 \geqslant 0$ 的条件, x^* 是 KKT 点当且仅当 $\lambda_1, \lambda_2 \geqslant 0$, 即 $|t| \leqslant \frac{1}{\sqrt[3]{4}}$.

(b) 分情况讨论:

- 当 $t = 1$ 时, 我们首先证明约束 $c_1(x) = 0$ 必须被满足. 否则, 若 $x \in \mathbb{R}^2$ 使得 $c_1(x) > 0$, 则 $d = (1,1)$ 是线性化可行方向, 但是沿着 d 的方向

$$d^{\mathrm{T}} g(x) = (2x_1 - 3) + 4(x_2 - 1)^3 < 0$$

故 x 一定不是 KKT 点. 在 KKT 点处, 必须有 $c_1(x)=0$ 满足.

- 在 $x=(1,0)$ 处, 线性化可行方向 $d=(-1,1)$ 使得

$$d^{\mathrm{T}}g(x)=-(2x_1-3)+4(x_2-1)^3=-3<0$$

因此 $x=(1,0)$ 不是 KKT 点.

- 在 $x=(0,1)$ 处, 线性化可行方向 $d=(1,-1)$ 使得

$$d^{\mathrm{T}}g(x)=(2x_1-3)-4(x_2-1)^3=-3<0$$

因此 $x=(0,1)$ 不是 KKT 点.

- 因此, KKT 点只能出现在 $c_1(x)=0$ 对应的开线段上. 此时 $c_2(x),c_3(x),c_4(x)$ 均不起作用, 故 $\lambda_2=\lambda_3=\lambda_4=0$, 且由 KKT 条件可得

$$\begin{cases} 2x_1-3=-\lambda_1 \\ 4(x_2-1)^3=-\lambda_1 \\ x_1+x_2=1 \end{cases}$$

其解为 $x_1\approx 0.7281$.

综上, 此问题在 $c_1(x)=0$ 上存在唯一的 KKT 点.

例 3 考虑约束优化问题

$$\min_{x\in\mathbb{R}^2}\ (x_1-2)^2+x_2^2 \tag{6.30}$$

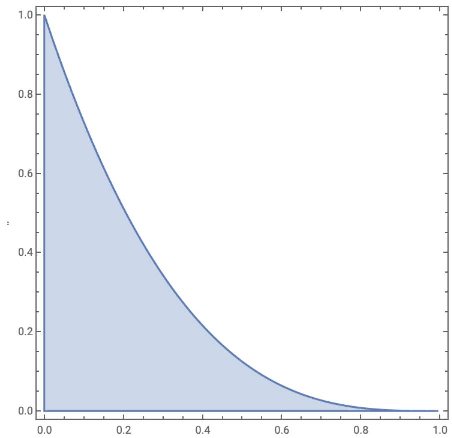
$$\text{s.t. } (1-x_1)^3-x_2\geqslant 0 \tag{6.31}$$

$$x_1,x_2\geqslant 0 \tag{6.32}$$

回答以下问题:

(1) 求出该问题的所有 KKT 点;

(2) 求问题的局部极小点.



目标函数为

$$f(x) = (x_1 - 2)^2 + x_2^2$$

约束函数为

$$c_1(x) = (1 - x_1)^3 - x_2, \quad c_2(x) = x_1, \quad c_3(x) = x_2$$

因此

$$\nabla f(x) = \begin{bmatrix} 2(x_1 - 1) \\ 2x_2 \end{bmatrix}, \quad \nabla c_1(x) \begin{bmatrix} -3(1 - x_1)^2 \\ -1 \end{bmatrix}, \quad \nabla c_2(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \nabla c_3(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

相应的 KKT 条件为

$$\begin{cases} 2(x_1 - 2) = -3\lambda_1(1 - x_1)^2 + \lambda_2 \\ 2x_2 = -\lambda_1 + \lambda_3 \end{cases}$$

由于 $\lambda_1 \geqslant 0$, 故 $\lambda_3 \geqslant 2x_2$. 下面分 3 种情形讨论上述方程的解:

- 若 $x_2 \neq 0$, 则 $\lambda_3 > 0$, 因此 $c_3(x) = x_2 = 0$ 为有效约束. 但这意味着 $x_2 = 0$, 矛盾!
- 因此必须 $x_2 = 0$. 若 $x_1 \in [0, 1)$, 则在 $x^* = (x_1, 0)$ 处, 沿着方向 $d = (1, 0)$ 一定会使目标函数 $f(x)$ 的值下降, 因为

  $$d^{\mathrm{T}} \nabla f(x) = 2(x_1 - 2) < 0$$

  但 $d = (1, 0) \in \mathrm{LFD}(x^*, X)$ 为线性化可行方向, 与 KKT 条件矛盾! 因此 $x_1 \in [0, 1)$ 时, x^* 不可能是 KKT 点.
- 剩下唯一的可能性: $x^* = (1, 0)$. 此时 $c_2(x) = x_1$ 为无效约束, $\lambda_2 = 0$. 但 KKT 条件给出

  $$\begin{cases} -2 = \lambda_2 \\ 0 = -\lambda_1 + \lambda_3 \end{cases}$$

矛盾!

综上, 该问题不存在 KKT 点.

但是, 这显然不意味着这个问题没有局部极小点. 实际上, 除了 $x^* = (1,0)$ 之外, 其它所有点处都有约束规范条件成立, 因此那些点一定不是局部极小点. 因此, 唯一的局部极小点就是 $x^* = (1,0)$, 但是在这个点, 约束规范条件不成立.

下面我们给出使得约束规范条件 (6.17) 成立的若干条件:

定理 6.2 (约束规范条件) 若以下之一成立, 则约束规范条件 $\mathrm{SFD}(x^*, X) = \mathrm{LFD}(x^*, X)$ 成立:

1. 所有 $c_i(x), i \in \mathcal{A}(x^*)$ 是线性函数.
2. LICQ: $\nabla c_i(x^*), i \in \mathcal{A}(x^*)$ 线性无关.

综合以上结果, 我们可以给出完整的一阶最优条件刻画:

定理 6.3 (一阶必要条件) 设 x^* 是局部极小点, 如果 $\nabla c_i(x), i \in \mathcal{A}(x^*)$ 线性无关, 则 x^* 是 KKT 点, 即存在 Lagrange 乘子 λ^* 使得 KKT 条件成立.

6.3 一阶充分条件

我们讨论局部极小点的一阶充分条件.

定理 6.4 设 $x^* \in X$. 如果 $f(x)$ 和 $c_i(x)$ 在 x^* 处可微, 且

$$d^{\mathrm{T}} \nabla f(x^*) > 0, \qquad \forall 0 \neq d \in \mathrm{SFD}(x^*, X) \tag{6.33}$$

则 x^* 是严格局部极小点.

由于 $\mathrm{SFD}(x^*, X) \subset \mathrm{LFD}(x^*, X)$, 因此我们得到以下结果:

推论 6.1 设 $x^* \in X$. 如果 $f(x)$ 和 $c_i(x)$ 在 x^* 处可微, 且

$$d^{\mathrm{T}} \nabla f(x^*) > 0, \qquad \forall 0 \neq d \in \mathrm{LFD}(x^*, X) \tag{6.34}$$

则 x^* 是严格局部极小点.

在实际中, 更加实用的判断极小点的方式是直接验证不等式:

定理 6.5 (实用充分条件) 设 x^* 是约束优化问题的 KKT 点, λ 是相应的 Lagrange 乘子. 如果

$$L(x^*, \lambda) \leqslant L(x, \lambda) \tag{6.35}$$

在 x^* 的某个邻域 $B(x^*, \delta)$ 中成立, 则 x^* 是局部极小点. 若不等式对 $x \in \mathbb{R}^n$ 恒成立, 则 x^* 是全局极小点.

在 KKT 点处, 有 $L(x^*, \lambda) = f(x^*)$. 由

$$L(x, \lambda) = f(x) - \sum_{i=1}^{n} \lambda_i c_i(x) \geqslant L(x^*, \lambda) = f(x^*)$$

可以得到, 当 x 位于可行域中时,

$$f(x) \geqslant f(x^*) + \sum_{i=1}^{n} \lambda_i c_i(x) \geqslant f(x^*)$$

因此 x^* 为局部极小点.

例 4 求解如下的约束优化问题

$$\min x_1^2 + x_2^2 \tag{6.36}$$

$$\text{s.t. } x_1^2 + x_2 - 1 \geqslant 0 \tag{6.37}$$

首先我们求出问题的 KKT 点. 目标函数 $f(x) = x_1^2 + x_2^2$, 约束函数 $c(x) = x_1^2 + x_2 - 1$, 故

$$\nabla f(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla c(x) = \begin{bmatrix} 2x_1 \\ 1 \end{bmatrix}$$

因此 KKT 条件给出

$$\begin{cases} 2x_1 = 2x_1\lambda \\ 2x_2 = \lambda \end{cases}$$

按照 x_1 分情况讨论:

- 如果 $x_1 \neq 0$, 则 $\lambda = 1$, $x_2 = \frac{1}{2}$, 并且约束有效. 此时 $x_1 = \frac{1}{\sqrt{2}}$(或者相应的负值). 故

  $$x^* = \left(\frac{1}{\sqrt{2}}, \frac{1}{2}\right), \quad \lambda^* = 1$$

  它是一个 KKT 点. 为了证明它是局部极小点, 我们证明

  $$L(x, \lambda^*) \geqslant f(x^*)$$

对 $x \in \mathbb{R}^2$ 恒成立. 事实上,

$$
\begin{aligned}
L(x, \lambda^*) &= f(x) - c(x) \\
&= x_2^2 - x_2 + 1 \\
&\geqslant \frac{3}{4} = f(x^*)
\end{aligned}
$$

因此当 x 位于可行域中时

$$f(x) \geqslant f(x^*) + c(x) \geqslant f(x^*)$$

因此 $x^* = (\frac{1}{\sqrt{2}}, \frac{1}{2})$ 是局部极小点.

- 如果 $x_1 = 0$, 则 $x_2 \geqslant 1$, 故 $\lambda_2 > 0$, 为有效约束. 因此 $x_2 = 1$, $\lambda = 2$, 故

  $$x^* = (0, 1), \qquad \lambda^* = 2$$

  它是一个 KKT 点. 为了证明它不是局部极小点, 考虑

  $$x^\varepsilon = (\varepsilon, 1 - \varepsilon^2), \qquad \varepsilon > 0$$

  则 x^ε 位于可行域中, 且 ε 充分小时,

  $$f(x^\varepsilon) = \varepsilon^2 + (1 - \varepsilon^2)^2 < f(x^*)$$

  故 x^* 不是局部极小点.

综上, 该问题的极小点为 $x^* = (\pm\frac{1}{\sqrt{2}}, \frac{1}{2})$.

我们对一阶最优性条件做一总结:

必要	$d^{\mathrm{T}}\nabla f(x^*) \geqslant 0, \forall d \in \mathrm{SFD}(x^*, X)$	$\xLeftrightarrow{\text{约束规范}}$	$d^{\mathrm{T}}\nabla f(x^*) \geqslant 0, \forall d \in \mathrm{LFD}(x^*, X)$: KKT 条件
充分	$d^{\mathrm{T}}\nabla f(x^*) > 0, \forall 0 \neq d \in \mathrm{SFD}(x^*, X)$	$\Longleftarrow$	$d^{\mathrm{T}}\nabla f(x^*) > 0, \forall 0 \neq d \in \mathrm{LFD}(x^*, X)$
实用充分	$L(x^*, \lambda) \leqslant L(x, \lambda), \forall x \in \mathbb{R}^n$		

表 1:: 一阶最优性条件

求 KKT 点的解题策略:

1. 求解 KKT 方程组
2. 验证 x 不是 KKT 点时, 取一线性化可行方向 d 使得 $d^{\mathrm{T}}\nabla f(x) < 0$

验证局部极小点的解题策略:

1. 排除其它所有可能是局部极小点的地方 (适用于 KKT 点只有 1 个的情形)
2. 当约束函数和目标函数简单时, 使用基本不等式证明
3. 证明 $L(x^*, \lambda) \leqslant L(x, \lambda), \forall x \in \mathbb{R}^n$
4. 使用定义证明某一点不是局部极小点

6.4 二阶最优性条件

二阶最优性条件主要适用于已经求出 KKT 点, 但是利用已知的一阶条件不足以判断它是否是最优的情形. 不过, 在实际情形中, 我们可以用其它的方法来判断某个点是否是最优的, 比如目标函数比较简单的情形. 计算 Lagrange 函数的二阶导数实际上还是相对复杂的, 零约束子空间实际上也没那么好求.

下面我们假设在某个 x^* 处, 一阶必要条件满足, 但充分条件不满足, 即:

$$d^{\mathrm{T}}\nabla f(x^*) \geqslant 0, \qquad \forall d \in \mathrm{SFD}(x^*, X) \tag{6.38}$$

$$d^{\mathrm{T}}\nabla f(x^*) = 0, \qquad \exists 0 \neq d \in \mathrm{SFD}(x^*, X) \tag{6.39}$$

则这些条件不足以判断 x^* 是否是局部极小点. 假设约束规范条件成立, 则第一个条件给出 KKT 条件, 因此

$$\nabla f(x^*) = \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* \nabla c_i(x^*) \tag{6.40}$$

于是 d 使得 $d^{\mathrm{T}}\nabla f(x^*) = 0$ 当且仅当

$$\lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) = 0, \qquad \forall i \in I(x^*) \tag{6.41}$$

自然地, 我们引入如下定义

定义 6.8 设 x^* 是 KKT 点且 λ^* 为相应的 Lagrange 乘子. 若 $d \in \mathrm{LFD}(x^*, X)$ 且

$$\lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) = 0, \qquad \forall i \in I(x^*) \tag{6.42}$$

则 d 称为 x^* 处的线性化零约束方向. x^* 处的所有线性化零约束方向的集合记为 $G(x^*, \lambda^*)$.

关于线性化零约束方向 d 的评注:

- d 首先得是一个可行方向;

- d 的定义与目标函数没有直接关系, 间接关系是最优性条件:

$$\nabla f(x^*) = \sum_{i \in \mathcal{E}} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in \mathcal{I}(x^*)} \lambda_i^* \nabla c_i(x^*) \tag{6.43}$$

  一方面, d 可行意味着

$$d^{\mathrm{T}} \nabla c_i(x^*) = 0, \qquad i \in \mathcal{E}$$

  另一方面, d 是线性化零约束方向意味着

$$\lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) = 0, \qquad i \in I(x^*)$$

  因此在 x^* 处沿着 d 的方向 $f(x)$ 的值不改变. 这是它称为线性化零约束方向的原因.

- 在老师的笔记中, 线性化零约束方向被定义为

$$\mathrm{LFD}_1(x^*, X) = \left\{ d \neq 0 : \begin{array}{c} d^{\mathrm{T}} \nabla c(x^*) = 0, i \in E \\ d^{\mathrm{T}} \nabla c(x^*) = 0, i \in I(x^*), \lambda_i^* \neq 0 \\ d^{\mathrm{T}} \nabla c(x^*) \geqslant 0, i \in I(x^*), \lambda_i^* = 0 \end{array} \right\} \tag{6.44}$$

  这两种定义是等价的.

类似地还可以定义序列零约束方向:

定义 6.9 设 x^* 是 KKT 点且 λ^* 为相应的 Lagrange 乘子. 如果存在序列 d_k 和 $\delta_k > 0$ 使得

$$x^* + \delta_k d_k \in X$$

$$\sum_{i=1}^{m} \lambda_i^* c_i(x^* + \delta_k d_k) = 0$$

且 $d_k \to d$ 和 $\delta_k \to 0$, 则称 d 是 x^* 处的序列零约束方向.

在上面的定义中, 上面的等式等价于

$$\lambda_i^* c_i(x^* + \delta_k d_k) = 0, \qquad \forall i \in I$$

事实上, 由于 $x^* + \delta_k d_k$ 是可行点, 故

$$c(x^* + \delta_k d_k) = 0, \qquad i \in \mathcal{E}$$

引理 6.4 设 x^* 是 KKT 点且 λ^* 是相应的 Lagrange 乘子, 则

$$S(x^*, \lambda^*) \subset G(x^*, \lambda^*) \tag{6.45}$$

在等式

$$\sum_{i=1}^{m} \lambda_i^* \frac{c_i(x^* + \delta_k d_k) - c_i(x^*)}{\delta_k} = 0 \tag{6.46}$$

中令 $k \to \infty$, 有

$$\sum_{i=1}^{m} \lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) = 0 \tag{6.47}$$

对于等式约束, 已经有 $d^{\mathrm{T}} \nabla c_i(x^*) = 0$; 对于无效不等式约束, 有 $\lambda_i^* = 0$, 因此上式变成

$$\sum_{i \in I(x^*)} \lambda_i^* d^{\mathrm{T}} \nabla c_i(x^*) = 0 \tag{6.48}$$

由于其中的每一项非负, 故每一项均为 0. 下面我们给出二阶必要条件和充分条件:

定理 6.6 (二阶必要条件) 设 x^* 是局部极小点, λ^* 是 Lagrange 乘子, 则

$$d^{\mathrm{T}} \nabla_{xx}^2 L(x^*, \lambda^*) d \geqslant 0, \quad \forall d \in S(x^*, \lambda^*) \tag{6.49}$$

注意: x^* 是局部极小点时, 由约束规范条件 x^* 是 KKT 点, 故 λ^* 一定存在.

定理 6.7 (二阶充分条件) 设 x^* 是 KKT 点, λ^* 是相应的 Lagrange 乘子. 若

$$d^{\mathrm{T}} \nabla_{xx}^2 L(x^*, \lambda^*) d > 0, \quad \forall 0 \neq d \in G(x^*, \lambda^*) \tag{6.50}$$

则 x^* 是严格极小点.

关于二阶条件我们有以下评注:

1. 为什么是Lagrange 函数 $L(x, \lambda^*)$ 而不是目标函数 $f(x)$? 因为 Lagrange 函数才满足

   $$\nabla_x L(x^*, \lambda^*) = 0 \tag{6.51}$$

   即一阶导数为 0. 对这样的函数考虑二阶导数才合理.

2. 为什么只在零约束方向上考虑二阶导数? 因为在其它方向上, $f(x)$ 沿着 d 的方向一定是递增的, 也就是一阶条件已经够用了, 没必要考虑二阶条件.

6.5 鞍点理论

对给定的约束优化问题 (6.1), 考虑 Lagrange 函数

$$L(x, \lambda) = f(x) - \sum_{i \in E \cup I} \lambda_i c_i(x) \tag{6.52}$$

其中 $x \in \mathbb{R}^n$, $\lambda \in K$, 区域 K 定义为

$$K = \{\lambda : \lambda_i \geqslant 0, i \in I\} \tag{6.53}$$

定义 6.10 $\bar{x} \in \mathbb{R}^n$ 和 $\bar{\lambda} \in K$ 称为 $L(x,\lambda)$ 的鞍点, 如果对任何 $x \in \mathbb{R}^n$ 和 $\lambda \in K$, 都有

$$L(\bar{x},\lambda) \leqslant L(\bar{x},\bar{\lambda}) \leqslant L(x,\bar{\lambda}) \tag{6.54}$$

鞍点理论最重要的结果是: 鞍点一定是最优解.

定理 6.8 (鞍点定理) 设 $(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda)$ 的鞍点, 则 $\bar{x}$ 是约束优化问题 (6.1) 的最优解.

首先证明: $L(\bar{x},\lambda) \leqslant L(\bar{x},\bar{\lambda})$ 恒成立给出 $\bar{x}$ 是可行点. 由 $L(x,\lambda)$ 的表达式, 可得

$$f(\bar{x}) - \sum_{i\in E\cup I} \lambda_i c_i(\bar{x}) \leqslant f(\bar{x}) - \sum_{i\in E\cup I} \bar{\lambda}_i c_i(\bar{x}), \qquad \lambda \in K \tag{6.55}$$

即

$$\sum_{i\in E} \bar{\lambda}_i c_i(\bar{x}) + \sum_{i\in I} \bar{\lambda}_i c_i(\bar{x}) \leqslant \sum_{i\in E} \lambda_i c_i(\bar{x}) + \sum_{i\in I} \lambda_i c_i(\bar{x}) \tag{6.56}$$

当 $i \in E$ 时, 必须 $c_i(\bar{x}) = 0$; 当 $i \in I$ 时, 必须 $c_i(\bar{x}) \geqslant 0$. 因此 $\bar{x}$ 是可行点.

下面考察右端的不等式. 当 x 是可行点时, 由 $L(\bar{x},\bar{\lambda}) \leqslant L(x,\bar{\lambda})$ 可以得到

$$f(\bar{x}) \leqslant f(x) - \sum_{i\in I} \lambda_i c_i(x) \tag{6.57}$$

故

$$f(x) \geqslant f(\bar{x}) + \sum_{i\in I} \lambda_i c_i(x) \geqslant f(\bar{x}) \tag{6.58}$$

因此 $\bar{x}$ 一定是最优解. 鞍点的这一条性质反过来不一定对.

定理 6.9 (鞍点逆定理) 设凸规划问题是强相容的, 则若 x^* 是约束优化问题的最优解, 则存在 λ^* 使得 (x^*,λ^*) 是 Lagrange 函数 $L(x,\lambda)$ 的鞍点, 且 $\lambda_i^* c_i(x^*) = 0$, $i \in I$.

求一个 Lagrange 函数的鞍点的方法:

1. 求出约束优化问题的最小点;
2. 求出最小点的 Lagrange 乘子;
3. 验证相应的 (x^*,λ^*) 是鞍点.

6.6 对偶问题

根据鞍点理论, 最优点的计算可以通过 $L(x,\lambda)$ 的鞍点的计算得到. 更具体地说, 计算分为 2 步:

1. 将 $L(x,\lambda)$ 对 $x \in \mathbb{R}^n$ 取最小值得到 $\theta(\lambda)$;

2. 对 $\theta(\lambda)$ 在定义域 K 取最大值.

我们可以将上述思想说明得更加形式化:

定义 6.11 (对偶问题) 对于约束优化问题

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) \\ &\text{s.t. } c_i(x) = 0,\ i \in E \\ &\quad\ \ c_i(x) \geqslant 0,\ i \in I \end{aligned}$$

设它的 Lagrange 函数为

$$L(x,\lambda) = f(x) - \sum_{i\in E\cup I} \lambda_i c_i(x)$$

其中 $x \in \mathbb{R}^n$, $\lambda \in K$, 且

$$\theta(\lambda) = \min_{x\in\mathbb{R}^n} L(x,\lambda)$$

则相应的对偶问题为

$$\begin{aligned} &\min\ \ \theta(\lambda) \\ &\text{s.t. } \lambda_i \geqslant 0,\ i \in I \end{aligned}$$

我们用一些简单的例子来说明上述思想.

例 5 考虑约束最优化问题

$$\min\ \ x_1 + x_2 \tag{6.59}$$

$$\text{s.t. } x_1^2 + x_2^2 \leqslant 2 \tag{6.60}$$

在此问题中, $f(x) = x_1 + x_2$, $c(x) = 2 - x_1^2 - x_2^2$, 因此 Lagrange 函数为

$$\begin{aligned} L(x,\lambda) &= f(x) - \lambda c(x) \\ &= \lambda(x_1^2 + x_2^2) + x_1 + x_2 - 2\lambda \end{aligned}$$

当 $\lambda > 0$ 时, 该函数的最小值在

$$x_1 = x_2 = -\frac{1}{2\lambda}$$

处取到. 因此

$$\theta(\lambda) = \max_x L(x,\lambda) = -\frac{1}{2\lambda} - 2\lambda \leqslant -2$$

且等号成立当且仅当 $\lambda = \frac{1}{2}$. 因此最优点为 $x^* = (-1,-1)$, 相应的 Lagrange 乘子 $\lambda^* = \frac{1}{2}$.

例 6 考虑约束最优化问题

$$\min\ x_1^2 + x_2^2 \tag{6.61}$$

$$\text{s.t. } x_1 + x_2 \geqslant 4 \tag{6.62}$$

$$x_1, x_2 \geqslant 0 \tag{6.63}$$

$$\tag{6.64}$$

此问题的 Lagrange 函数为

$$L(x,\lambda) = x_1^2 + x_2^2 - \lambda_1(x_1 + x_2 - 4) - \lambda_2 x_1 - \lambda_3 x_2$$

对 x_1, x_2 取最小值, 可以得到

$$\theta(\lambda) = \min_x L(x,\lambda) = -\frac{(\lambda_1+\lambda_2)^2}{4} - \frac{(\lambda_1+\lambda_3)^2}{4} + 4\lambda_1$$

其中 $\lambda_1, \lambda_2, \lambda_3 \geqslant 0$. 下面求 $\theta(\lambda)$ 的最大值.

$$\theta(\lambda) \leqslant -\frac{\lambda_1^2}{2} + 4\lambda_1 \leqslant 8$$

等号成立当且仅当 $\lambda_1 = 4$, $\lambda_2 = \lambda_3 = 0$. 此时

$$x_1 = \frac{\lambda_1 + \lambda_2}{2} = 2, \qquad x_2 = \frac{\lambda_1 + \lambda_3}{2} = 2$$

因此最优点 $x^* = (2,2)$.

例 7 考虑优化问题

$$\min\ x_1$$
$$\text{s.t. } |x_1| + |x_2| \leqslant 1$$

求出该问题的对偶问题并求解之.

该问题的目标函数为 $f(x) = x_1$, 约束函数为

$$c_1(x) = 1 - x_1 - x_2$$
$$c_2(x) = 1 - x_1 + x_2$$
$$c_3(x) = 1 + x_1 - x_2$$
$$c_4(x) = 1 + x_1 + x_2$$

因此 Lagrange 函数为

$$\begin{aligned}L(x,\lambda) &= f(x) - \lambda_1 c_1(x) - \lambda_2 c_2(x) - \lambda_3 c_3(x) - \lambda_4 c_4(x) \\ &= x_1 + \lambda_1(x_1 + x_2 - 1) + \lambda_2(x_1 - x_2 - 1) + \lambda_3(-x_1 + x_2 - 1) + \lambda_4(-x_1 - x_2 - 1) \\ &= x_1(1 + \lambda_1 + \lambda_2 - \lambda_3 - \lambda_4) + x_2(\lambda_1 - \lambda_2 + \lambda_3 - \lambda_4) - \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\end{aligned}$$

定义区域

$$K = \left\{ \lambda \in \mathbb{R}^4 : \begin{array}{r} 1 + \lambda_1 + \lambda_2 - \lambda_3 - \lambda_4 = 0 \\ \lambda_1 - \lambda_2 + \lambda_3 - \lambda_4 = 0 \\ \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geqslant 0 \end{array} \right\}$$

则在 K 上有 $\lambda_4 = \lambda_1 + \frac{1}{2}$, $\lambda_3 = \lambda_2 + \frac{1}{2}$. 因此对偶 Lagrange 函数为

$$\theta(\lambda) = \begin{cases} -\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4, & \lambda \in K \\ -\infty, & \lambda \notin K \end{cases}$$

因此 $\lambda \in K$ 时,

$$\theta(\lambda) \leqslant -1$$

且等号成立当且仅当 $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = \lambda_4 = 0$. 由于 $c_3(x) = c_4(x) = 0$ 为有效约束, 故 $x_1 = -1$, $x_2 = 0$, 即原问题的最优点为 $x^* = (-1, 0)$.

现在一个自然的问题是, 原问题与对偶问题的最优解是否是相同的, 即是否有

$$\min_{x \in X} f(x) = \max_{\lambda \in K} \theta(\lambda)$$

定理 6.10 (弱对偶定理) 设 $x \in X$, $\lambda \in K$, 则

$$f(x) \geqslant \theta(\lambda)$$

按照 $\theta(\lambda)$ 的定义,

$$\theta(\lambda) = \min_{x \in \mathbb{R}^n} \{f(x) - \sum_{i \in E \cup I} \lambda_i c_i(x)\} \leqslant f(x) - \sum_{i \in E \cup I} \lambda_i c_i(x)$$

当 $x \in X$, $\lambda \in K$ 时,

$$\sum_{i \in E \cup I} \lambda_i c_i(x) \geqslant 0$$

因此 $\theta(\lambda) \leqslant f(x)$. 由弱对偶定理可以得到

$$\inf_{x \in X} f(x) \geqslant \sup_{\lambda \in K} \theta(\lambda)$$

定义 6.12 (对偶间隙) 给定约束优化问题 $f(x)$ 和相应的对偶函数 $\theta(\lambda)$, 定义其对偶间隙为

$$\gamma = \inf_{x \in X} f(x) - \sup_{\lambda \in K} \theta(\lambda) \tag{6.65}$$

下面我们的任务是探讨在何种情形下对偶间隙为 0.

定义 6.13 (Slater 条件) 对凸规划问题, 若

$$D = \left\{ x \in \mathbb{R}^n : \begin{array}{l} c_i(x) = 0, \quad i \in E \\ c_i(x) > 0, \quad i \in I \end{array} \right\}$$

非空, 则称约束函数满足 Slater 条件.

定理 6.11 (强对偶定理) 对凸规划问题, 若 Slater 条件满足, 则对偶间隙 $\gamma = 0$.

考虑一个具体的例子: 二次规划.

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x^{\mathrm{T}} G x + g^{\mathrm{T}} x \tag{6.66}$$

$$\text{s.t. } a_i^{\mathrm{T}} x = b_i, \ i \in E \tag{6.67}$$

$$a_i^{\mathrm{T}} x \geqslant b_i, \ i \in I \tag{6.68}$$

此问题的 Lagrange 函数为

$$\begin{aligned} L(x, \lambda) &= f(x) - \lambda^{\mathrm{T}}(Ax - b) \\ &= \frac{1}{2} x^{\mathrm{T}} G x + g^{\mathrm{T}} x - \lambda^{\mathrm{T}}(Ax - b) \\ &= \frac{1}{2} x^{\mathrm{T}} G x + x^{\mathrm{T}}(g - A^{\mathrm{T}}\lambda) + b^{\mathrm{T}}\lambda \end{aligned}$$

其中 $\lambda \in K$, 区域 K 定义为

$$K = \left\{ \lambda : \begin{array}{l} \lambda_i = 0, \ i \in E \\ \lambda_i \geqslant 0, \ i \in I \end{array} \right\}$$

因此对偶函数

$$\theta(\lambda) = \min_{x \in \mathbb{R}^n} L(x, \lambda) = -\frac{1}{2}(-g + A^{\mathrm{T}}\lambda)^{\mathrm{T}} G^{-1} (-g + A^{\mathrm{T}}\lambda) + b^{\mathrm{T}}\lambda$$

原优化问题的对偶函数为

$$\max \ \frac{1}{2}(-g + A^{\mathrm{T}}\lambda)^{\mathrm{T}} G^{-1} (-g + A^{\mathrm{T}}\lambda) + b^{\mathrm{T}}\lambda \tag{6.69}$$

$$\text{s.t. } \lambda_i \geqslant 0, \ i \in I \tag{6.70}$$

下面我们来计算对偶间隙 ($y = -g + A^{\mathrm{T}}\lambda$):

$$f(x) - \theta(\lambda) = \lambda^{\mathrm{T}}(Ax - b) + \frac{1}{2}(x^{\mathrm{T}}Gx + y^{\mathrm{T}}G^{-1}y + 2x^{\mathrm{T}}y) \tag{6.71}$$

其中当 $x \in D$, $\lambda \in K$ 时, $\lambda^{\mathrm{T}}(Ax - b) \geqslant 0$, 因此对偶间隙非负. 可以验证, 对偶间隙为 0 当且仅当

$$\lambda_2^{\mathrm{T}}(Ax_2 - b_2) = 0$$

且

$$A^{\mathrm{T}}\lambda = Gx + g$$

这恰好就是原优化问题对应的 KKT 条件! 对偶理论主要用于线性规划和二次规划.

7 可行方向法

在前面的章节我们已经讨论过约束优化问题的基本理论: KKT 条件. 在本节我们讨论求解约束优化问题的基本算法. 有一类简单的罚函数方法, 它通过增加罚函数将约束问题转化为无约束问题. 例如, 对于等式约束问题

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & c_i(x) = 0, i \in E \end{aligned}$$

则相应的罚函数为

$$p(x, \sigma) = f(x) + \sigma \sum_{i \in E} c_i^2(x) \tag{7.1}$$

当 $\sigma \to +\infty$ 时, $p(x, \sigma)$ 的极小点收敛到约束优化问题的极小点.

7.1 可行方向法的一般性质

可行方向法的基本思想是每一步迭代都必须在可行方向 $\mathrm{FD}(x, D)$ 上进行. 考虑

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & x \in D \end{aligned}$$

其中 D 是由等式约束和不等式约束围成的区域:

$$D = \left\{ x \in \mathbb{R}^n : \begin{array}{ll} c_i(x) = 0, & i \in E \\ c_i(x) \geqslant 0, & i \in I \end{array} \right\}$$

在每一个迭代点 x_k, 我们希望找到一个可以使得函数值下降的可行方向, 即找到 $d \in \mathbb{R}^n$ 使得

$$d^{\mathrm{T}} \nabla f(x_k) < 0, \qquad d \in \mathrm{FD}(x_k, D)$$

这样我们就可以在 d 的方向上进行线搜索, 并且使得函数值严格下降.

Algorithm 13: 可行方向法

输入: 定义域 D, 初始点 $x_0 \in D$

输出: 迭代点序列 $\{x_k\}$

for $k = 0, 1, \cdots$ **do**

找到一个可行下降方向 $d_k \in \mathbb{R}^n$ 使得

$$d_k^{\mathrm{T}} \nabla f(x_k) < 0, \qquad d_k \in \mathrm{FD}(x_k, D)$$

如果不存在, 则迭代停止;

在 d_k 上进行 (精确) 线搜索得到步长 α_k;

更新迭代点 $x_{k+1} = x_k + \alpha_k d_k$;

end

在迭代中, 要求 $f(x_k)$ 的值严格下降. 值得注意的是, 这样的可行下降方向并非总是存在. 例如在非线性等式约束

$$c(x_1, x_2) = x_1^2 - x_2 = 0$$

上, 可行方向 $\mathrm{FD}(x, D) = \emptyset$. 幸运的是, 对于凸优化问题, 除全局极小点外, 可行下降方向总是可以找到的.

引理 7.1 (凸优化存在可行下降方向) 设定义域 D 是凸集, 且目标函数 $f(x)$ 在 D 上为凸函数, 则在 x 处存在可行下降方向当且仅当 x 不是局部极小点.

当 x 是局部极小点时, 它显然不存在可行下降方向, 否则沿着此方向函数值减少. 当 x 不存在可行下降方向时, 我们来验证它是局部极小点. 若否, 则存在 $\bar{x} \neq x$, 且 $f(\bar{x}) < f(x)$. 令方向 $d = \bar{x} - \bar{x}$, 则必有 $d^{\mathrm{T}} \nabla f(x) < 0$, 即沿着 d 的方向函数值严格下降. 但是 $d \in \mathrm{FD}(x, D)$, 故 d 为可行下降方向, 矛盾!

7.2 变量消去法

变量消去法的基本思想是用非基本变量 x_N 表示基本变量 x_B. 对于线性约束的情形, 设约束条件为 $Ax = b$, 其中 $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. 设

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}, \qquad A = \begin{bmatrix} A_B & A_N \end{bmatrix}$$

且 $A_B \in \mathbb{R}^{m\times n}$ 非奇异, 则基本变量 x_B 可以表为

$$x_B = A_B^{-1}(b - A_N x_N) \tag{7.2}$$

因此原约束优化问题成为了关于 x_N 的无约束优化问题.

我们可以将上面的思想应用于一般的非线性约束中. 考虑等式约束

$$c(x) = \begin{bmatrix} c_1(x) \\ \vdots \\ c_m(x) \end{bmatrix} = 0$$

并且假设 x 可以分解为

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} \in \mathbb{R}^n$$

其中基本变量 $x_B \in \mathbb{R}^m$, 非基本变量 $x_N \in \mathbb{R}^{n-m}$, 则约束条件可以写为

$$c(x_B, x_N) = 0 \tag{7.3}$$

假设我们可以唯一地从中解出

$$x_B = \varphi(x_N) \tag{7.4}$$

则原约束问题转化为无约束问题

$$\tilde{f}(x_N) := f(\varphi(x_N), x_N)$$

的优化. 下面的任务是计算 $\tilde{f}(x_N)$ 关于 x_N 的梯度, 即所谓的简约梯度.

引理 7.2 (简约梯度) 对于无约束优化函数 $\tilde{f}(x_N) := f(\varphi(x_N), x_N)$, 它的简约梯度为

$$\tilde{g}(x_N) = \underbrace{\nabla_N f(x)}_{(n-m)\times 1} - \underbrace{\nabla_N c(x)}_{(n-m)\times m} \underbrace{(\nabla_B c(x))}_{m\times m}^{-1} \underbrace{\nabla_B f(x)}_{m\times 1} \in \mathbb{R}^{(n-m)\times 1} \tag{7.5}$$

在上面的表达式中, 约束条件 $c(x)$ 视为行向量, 而所有的梯度均为列向量. 首先考虑关于自由变量 x_N 的方程

$$c(\varphi(x_N), x_N) = 0$$

对 x_N 取导数, 有

$$(\nabla\varphi(x_N))^{\mathrm{T}} \nabla_B c(x) + \nabla_N c(x) = 0 \Longrightarrow (\nabla\varphi(x_N))^{\mathrm{T}} = -\nabla_N c(x)(\nabla_B c(x))^{-1}$$

因此 $\tilde{f}(x_N)$ 的导数为

$$\begin{aligned}\tilde{g}(x_N) &= \nabla_N f(\varphi(x_N), x_N) \\ &= (\nabla\varphi(x_N))^{\mathrm{T}}\nabla_N f(x) + \nabla_B f(x) \\ &= \nabla_N f(x) - \nabla_N c(x)(\nabla_B c(x))^{-1}\nabla_B f(x)\end{aligned}$$

得证. 因此, 下降方向可以选为

- 最速下降方向: $\tilde{d} = -\tilde{g}(x_N)$
- 拟 Newton 方向: $\tilde{d} = -B^{-1}\tilde{g}(x_N)$

在迭代点 x_N 处, 我们对目标函数 $\tilde{f}(x_N + \alpha d)$ 进行线搜索, 实际上是在对

$$f(\varphi(x_N + \alpha d), x_N + \alpha d)$$

进行线搜索. 因此, 每一步需要求解一个非线性方程

$$c(x_B, x_N + \alpha d) = 0$$

来得到 $x_N = \varphi(x_N + \alpha d)$.

除了直接做变量分离

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$$

来进行变量消去外, 我们也可以考虑一般形式

$$x = \underbrace{Y}_{n\times m} x_y + \underbrace{Z}_{n\times(n-m)} x_z, \quad x_y \in \mathbb{R}^m, \quad x_z \in \mathbb{R}^{n-m}$$

其中 $AY \in \mathbb{R}^{m\times m}$ 非奇异, $AZ = 0$, 且 Z 的列线性无关. 因此可以得到

$$Ax = b = AYx_y \Longrightarrow x_y = (AY)^{-1}b$$

因此变量 x 可以用自由变量 x_z 线性表出:

$$x = Y(AY)^{-1}b + Zx_z, \quad x_z \in \mathbb{R}^{n-m}$$

Y, Z 可以用下面的方法选取: 做 QR 分解

$$A^{\mathrm{T}}P = \begin{bmatrix} Y & Z \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix}$$

我们在本节考虑线性不等式约束其中 P 是任意可逆矩阵.

7.3 Rosen 投影梯度法

在本节我们考虑不等式约束优化问题

$$\begin{aligned} &\min \ f(x) \\ &\text{ s.t. } Ax \geqslant b \end{aligned}$$

7.3.1 投影矩阵及其性质

定义 7.1 (投影矩阵) 矩阵 $P \in \mathbb{R}^{n\times n}$ 称为投影矩阵, 如果 $P^2 = P$, 称为正交投影矩阵, 如果 $P^2 = P$ 且 $P = P^{\mathrm{T}}$. 对给定的矩阵 $P \in \mathbb{R}^{n\times n}$, 定义

$$R(P) = \{Px : x \in \mathbb{R}^n\}, \qquad N(P) = \{x \in \mathbb{R}^n : Px = 0\}$$

分别为其值域和零空间.

例 8 设 $e \in \mathbb{R}^n$ 为任意单位向量, 则

$$P = I - ee^{\mathrm{T}} \tag{7.6}$$

是正交投影矩阵.

定理 7.1 (正交投影的性质) 若 $P \in \mathbb{R}^{n\times n}$ 为正交投影矩阵, 则

(i) P 半正定;

(ii) $R(P)$ 与 $N(P)$ 互为正交补. 对任意 $x \in \mathbb{R}^{n\times n}$, 存在唯一的 $p \in R(P)$ 和 $q \in N(P)$ 使得

$$x = p + q$$

(iii) $I - P$ 也是正交投影, 且 $R(I-P) = N(P)$, $N(I-P) = R(P)$.

(i) 由于 $P = P^{\mathrm{T}}$, 故 P 是一个对称矩阵. 由于

$$x^{\mathrm{T}}Px = \|Px\|^2$$

故 P 半正定.

(ii) 先证明 $R(P)$ 与 $N(P)$ 相互正交. 任取 $p \in R(P)$ 和 $q \in N(P)$, 有

$$p = Pz, \qquad Pq = 0$$

因此

$$p^{\mathrm{T}}q = z^{\mathrm{T}}Pq = 0$$

故 $R(P)$ 和 $N(P)$ 正交. 由于

$$\dim R(P) = \operatorname{rank}(P), \qquad \dim N(P) = n - \operatorname{rank}(P)$$

故 $\dim R(P) + \dim N(P) = n$, 从而 $R(P)$ 与 $N(P)$ 互为正交补.

(iii) 先证 $R(I-P) = N(P)$. 任取 $x \in \mathbb{R}^n$, 有

$$(I-P)x = x - Px \in N(P) \Longrightarrow R(I-P) \subset N(P)$$

任取 $z \in N(P)$, 有

$$z = (I-P)z \in R(I-P) \Longrightarrow N(P) \subset R(I-P)$$

因此 $R(I-P) = N(P)$. 取正交补得到 $N(I-P) = R(P)$. 由于 P 是半正定的, 可以得到

推论 7.1 (投影下降方向) 设 $g \in \mathbb{R}^n$ 是迭代点 x 处的梯度, 则

$$d = -Pg$$

满足 $g^{\mathrm{T}} d \leqslant 0$.

根据以上讨论, 正交投影矩阵 P 的作用可以看作是在 $R(P)$ 上的正交投影.

引理 7.3 (零空间正交投影) 设 $A \in \mathbb{R}^{q \times n}$ 且 A 为行满秩矩阵, 则

$$P = I - A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A \tag{7.7}$$

是从 $\mathbb{R}^n$ 在 $N(A)$ 上的正交投影.

给定 $x \in \mathbb{R}^n$, 我们计算 x 在 $N(A)$ 上的正交投影. 它可转化为约束优化问题

$$\begin{aligned} \min_z \ & \frac{1}{2}\|z - x\|^2 \\ \text{s.t.} \ & Az = 0 \end{aligned}$$

其中 $Az = 0$ 等价于 $z \in N(A)$, 而 $\|z - x\|^2$ 的最小值刻画了投影的性质. 由于此问题为凸优化问题, 因此极小点等价于 KKT 点. 该问题的 Lagrange 函数为

$$L(z, \lambda) = \frac{1}{2}\|z - x\|^2 - \lambda^{\mathrm{T}} Az$$

在 KKT 点处

$$\frac{\partial L}{\partial z} = z - x - A^{\mathrm{T}}\lambda = 0$$

由于 $Az=0$, 故

$$A(z-x)=AA^{\mathrm{T}}\lambda \Longrightarrow \lambda=-(AA^{\mathrm{T}})^{-1}Ax$$

因此

$$z=x+A^{\mathrm{T}}\lambda=x-A^{\mathrm{T}}(AA^{\mathrm{T}})Ax=Px$$

故 $Px=z$, 即 P 是 $N(A)$ 上的正交投影.

7.3.2 投影梯度矩阵构造可行下降方向

在不等式约束优化问题中, 投影矩阵起到重要的作用:

定理 7.2 (投影可行下降方向 I) 设 $\bar{x}\in D$,

(1) 对任何投影矩阵 P, 若 $Pg(\bar{x})\neq 0$, 则

$$\bar{d}=-Pg(\bar{x})$$

是 $\bar{x}$ 处的下降方向.

(2) 在不等式约束条件下, 设 $A_q\in\mathbb{R}^{q\times n}$ 是由有效约束对应的行向量构成的矩阵, 且设 A_q 是行满秩的, 则

$$P_q=I-A_q^{\mathrm{T}}(A_qA_q^{\mathrm{T}})^{-1}A_q \tag{7.8}$$

为正交投影矩阵, 且 $\bar{d}=-P_qg(\bar{x})$ 为可行下降方向.

关于上述定理我们有以下评注:

1. 在沿着下降方向 $\bar{d}$ 时, 起作用约束条件没有被破坏, 因为 $\bar{d}\in N(A_q)$. 因此, 算法适用于在给定的起作用约束条件下求 $f(x)$ 的最小值.

2. 算法有一个重要的应用条件: $P_qg(\bar{x})\neq 0$. 如果此条件不成立, 则说明 $P_qg(\bar{x})=0$, 因此

$$g(\bar{x})=A_q^{\mathrm{T}}(A_qA_q^{\mathrm{T}})^{-1}A_qg(\bar{x}) \tag{7.9}$$

若我们定义

$$\lambda=(A_qA_q^{\mathrm{T}})^{-1}A_qg(\bar{x})$$

则上述等式可以改写为

$$g(\bar{x})=A_q^{\mathrm{T}}\lambda=\sum_{i=1}^{q}\lambda_ia_i \tag{7.10}$$

其中 $a_i\in\mathbb{R}^n$ 是 A_q 的第 i 行, 即 $g(\bar{x})$ 是 A_q 的行向量的线性组合! 这恰好对应于我们熟知的 KKT 条件.

3. 在 $g(\bar{x})$ 的展开 (7.10) 中, 注意 $a_i, i=1,\cdots,q$ 是线性无关的, 因为 A_q 是行满秩的. 因此, 展开式中的系数 λ_i 是唯一的.

4. 可以直接验证 $P = I - A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}A$ 满足 $AP = PA^{\mathrm{T}} = 0$.

定理 7.3 (投影可行下降方向 II) 设 $\bar{x} \in D$ 且 $P_q g(\bar{x}) = 0$. 定义

$$\lambda = (A_q A_q^{\mathrm{T}})^{-1} A_q g(\bar{x}) \in \mathbb{R}^q \tag{7.11}$$

则有

(1) 若 $\lambda \geqslant 0$, 则 $\bar{x}$ 为 KKT 点;

(2) 若 $\lambda_i < 0$, 则从约束条件中去掉 i 得到 $A_{q-1} \in \mathbb{R}^{(q-1)\times n}$, 则投影矩阵

$$P_{q-1} = I - A_{q-1}^{\mathrm{T}}(A_{q-1}A_{q-1}^{\mathrm{T}})^{-1}A_{q-1} \tag{7.12}$$

使得 $\bar{d} = -P_{q-1}g(\bar{x})$ 为可行下降方向.

为简便起见, 假定 $i = q$, 即第 q 个约束被去除了. 设

$$A_q = \begin{bmatrix} A_{q-1} \\ a_q^{\mathrm{T}} \end{bmatrix} \in \begin{bmatrix} \mathbb{R}^{(q-1)\times n} \\ \mathbb{R}^{1\times n} \end{bmatrix}, \qquad \lambda = \begin{bmatrix} \mu \\ \lambda_q \end{bmatrix} \in \begin{bmatrix} \mathbb{R}^{q-1} \\ \mathbb{R} \end{bmatrix}$$

则由$P_q g(\bar{x}) = 0$可知

$$g(\bar{x}) = A_{q-1}^{\mathrm{T}}\mu + \lambda_q a_q \tag{7.13}$$

因此方向 $\bar{d} = -P_{q-1}g(\bar{x})$ 可以表示为

$$\bar{d} = -\lambda_q P_{q-1} a_q \tag{7.14}$$

为了证明 $\bar{d}$ 是可行下降方向, 需要

1. $\bar{d}$ 是可行方向. 由于只有 q 一个约束条件被破坏, 所以只需要

$$a_q^{\mathrm{T}}\bar{d} \geqslant 0 \Longleftrightarrow \lambda_q a_q^{\mathrm{T}} P_{q-1} a_q \leqslant 0 \tag{7.15}$$

   由于 $\lambda_q \geqslant 0$, P_{q-1} 是半正定的, 上式显然成立.

2. $\bar{d}$ 是下降方向, 即证明 $P_{q-1}g(\bar{x}) \neq 0$. 若 $P_{q-1}g(\bar{x}) = 0$, 则 $\bar{g}$ 是 A_{q-1} 的行向量的线性组合, 这与唯一展开 (7.13) 矛盾!

因此我们得到如下的求解线性不等式约束问题

$$\begin{aligned} &\min\ f(x) \\ &\text{s.t. } Ax \geqslant b \end{aligned}$$

的投影梯度算法:

Algorithm 14: 投影梯度法 (线性不等式约束)

for $k = 0, 1, \cdots$ **do**
  确定有效约束 $A_q \in \mathbb{R}^{q\times n}$
  计算正交投影矩阵 $P_q \in \mathbb{R}^{n\times n}$
  if $P_q g(x_k) \neq 0$ **then**
    (投影可行下降方向 I: 给定约束下的优化)
    取可行下降方向 $d_k = -P_q g(x_k)$
  else
    (投影可行下降方向 II: 减少一个约束条件)
    计算 Lagrange 乘子 $\lambda = (A_q A_q)^{\mathrm{T}} A_q g(x_k) \in \mathbb{R}^q$
    if $\lambda \geqslant 0$ **then**
      x_k 是 KKT 点
      return
    else
      设下标 i 使得 $\lambda_i < 0$
      从有效约束 A_q 中去掉第 i 行得到 $A_{q-1} \in \mathbb{R}^{(q-1)\times n}$
      计算正交投影矩阵 $P_{q-1} \in \mathbb{R}^{n\times n}$
      取可行下降方向 $d_k = -P_{q-1} g(x_k)$
    end
  end
  在下降方向 d_k 上进行线搜索得到步长 α_k
  $x_{k+1} = x_k + \alpha_k d_k$
end

7.4 Zoutendijk 可行方向

考虑含有等式约束和不等式约束的优化问题

$$
\begin{aligned}
\min\ & f(x) \\
\text{s.t. } & a_i^{\mathrm{T}} x = b_i,\ i \in E \\
& a_i^{\mathrm{T}} x \geqslant b_i,\ i \in I
\end{aligned}
$$

Zoutendijk 子问题由下给出:

定义 7.2 (Zoutendijk 子问题) 在给定的迭代点 x_k 处, 考虑线性规划问题

$$
\begin{aligned}
\min\ & g_k^{\mathrm{T}} d \\
\text{s.t. } & A_E d = 0 \quad (i \in E) \\
& A_{I_k} d \geqslant 0 \quad (i \in I(x_k)) \\
& \|d_k\|_\infty \leqslant 1
\end{aligned}
$$

Motzkin 定理是 Farkas 引理的一个推广.

引理 7.4 (Motzkin 定理) 设 $m_1, m_2 \geqslant 0$, $m_3 > 0$ 为整数. 再设 $A_1 \in \mathbb{R}^{m_1 \times n}$, $A_2 \in \mathbb{R}^{m_2 \times n}$, $\mathbb{R}^{m_3 \times n}$, 则下面的两个线性系统中有且仅有一个有解:

$$A_1 x = 0, \quad A_2 x \geqslant 0, \quad A_3 x > 0 \tag{7.16}$$

$$A_1^{\mathrm{T}} a_1 + A_2^{\mathrm{T}} a_2 + A_3^{\mathrm{T}} a_3 = 0, \quad a_2 \geqslant 0, \quad a_3 \neq 0 \tag{7.17}$$

利用上述定理可以给出 Zoutendijk 子问题的可解性:

定理 7.4 设 $\bar{x} \in D$, 则 $\bar{x}$ 是极小点当且仅当 Zoutendijk 子问题的最优值为 0,