无约束优化习题解答

2021 年 5 月 23 日

Exercise 1　写出基于Wolfe-Powell准则的非精确一维搜索算法中插值多项式$p^{(1)}(t), p^{(2)}(t)$**的具体表达式。**

由

$$\begin{cases} p^{(1)}(a_1) = \varphi_1 \\ p^{(1)\prime}(a_1) = \varphi_1' \\ p^{(1)}(\alpha) = \varphi \end{cases}$$

可设

$$p^{(1)}(t) = \varphi_1 + \varphi_1'(t - a_1) + A(t - a_1)^2$$

向上式代入$p^{(1)}(\alpha) = \varphi$得

$$\varphi_1 + \varphi_1'(\alpha - a_1) + A(\alpha - a_1)^2 = \varphi \;\Rightarrow\; A = \frac{\varphi - \varphi_1 + \varphi_1'(a_1 - \alpha)}{(\alpha - a_1)^2}$$

由

$$\begin{cases} p^{(2)\prime}(a_1) = \varphi_1' \\ p^{(2)}(\alpha) = \varphi \\ p^{(2)\prime}(\alpha) = \varphi' \end{cases}$$

可设

$$p^{(2)}(t) = \varphi + \varphi'(t - \alpha) + C(t - \alpha)^2$$

于是

$$p^{(2)\prime}(a_1) = \varphi' + 2C(a_1 - \alpha) = \varphi_1' \;\Rightarrow\; C = \frac{\varphi_1' - \varphi'}{2(a_1 - \alpha)}$$

Excercise 2　证明基于Goldstein准则的非精确一维搜索算法的全局收敛性。　Goldstein准则：

$$\varphi(\alpha) \leqslant \varphi(0) + \rho\alpha\varphi'(0) \tag{1}$$

$$\varphi(\alpha) \geqslant \varphi(0) + (1 - \rho)\alpha\varphi'(0) \tag{2}$$

设$\forall k$，$g^{(k)}=\nabla f(x^{(k)})\neq 0$和$f(x^{(k)})$有下界，则$f(x^{(k)})-f(x^{(k+1)})\to 0$，由(1)，$-{g^{(k)}}^\top s^{(k)}\to 0$。

（反证）若$g^{(k)}\to 0$不成立，则$\exists\varepsilon>0$和子列$\left\{x^{(k)}\right\}_{k\in K}$使得$\left\|g^{(k)}\right\|\geqslant\varepsilon$，从而由

$$-{g^{(k)}}^\top s^{(k)}=\left\|g^{(k)}\right\|\left\|s^{(k)}\right\|\cos\theta_k\geqslant\varepsilon\left\|s^{(k)}\right\|\sin\mu$$

以及对$\forall k$, $\theta_k\leqslant\frac{\pi}{2}-\mu$，得$\left\|s^{(k)}\right\|\to 0$。
又由

$$f(x^{(k)}+s^{(k)})=f(x^{(k)})+\nabla f(x^{(k)})^\top s^{(k)}+o\left(\left\|s^{(k)}\right\|\right)$$

得

$$\lim_{k\to\infty}\frac{f(x^{(k)})-f(x^{(k+1)})}{-\nabla f(x^{(k)})^\top s^{(k)}}=1$$

由式(2)得

$$\frac{f(x^{(k)})-f(x^{(k+1)})}{-\nabla f(x^{(k)})^\top s^{(k)}}\leqslant 1-\rho<1$$

矛盾！所以$g^{(k)}=\nabla f(x^{(k)})\to 0, k\to\infty$。

Exercise 3 将非线性方程组求根$F(x)=0$的牛顿迭代，用于求最优化问题$\min_{x\in\mathbb{R}^n}f(x)$，给出相应的迭代格式并说明理由。 牛顿迭代的原理是取$F(x)$在$x^{(k)}$处的一阶Taylor展开做近似并令其为0

$$F(x^{(k)})+J_F(x^{(k)})(x-x^{(k)})=0$$

解的迭代格式为

$$x^{(k+1)}=x^{(k)}-\left(J_F(x^{(k)})\right)^{-1}F(x^{(k)})$$

对无约束问题，求解$\nabla f(x)=0$，即有迭代格式

$$x^{(k+1)}=x^{(k)}-\left(\nabla^2 f(x^{(k)})\right)^{-1}\nabla f(x^{(k)})$$

Exercise 4 证明对称秩一拟牛顿法具有遗传性和二次终止性。 对于二次函数$f(x)=\frac{1}{2}x^\top Gx+c^\top x$，

$$\nabla f(x)=Gx+c,\ \nabla^2 f(x)=G,$$

拟牛顿法中$y^{(k)}=g^{(k+1)}-g^{(k)}=G_k s^{(k)}$，正割条件为$H_{k+1}y^{(k)}=s^{(k)}$。

遗传性 使用归纳法。

1. $k=1$时，由正割条件，直接成立。

2. 假设遗传性对于k成立，即$H_k y^{(l)} = s^{(l)}, l = 0, 1, \ldots, k-1$。对于$k+1$和$l = 0, 1, \ldots, k-1$，由对称秩一校正公式

$$H_{k+1} y^{(l)} = H_k y^{(l)} + \frac{(s^{(k)} - H_k y^{(k)})(s^{(k)} - H_k y^{(k)})^\top y^{(l)}}{(s^{(k)} - H_k y^{(k)})^\top y^{(k)}}$$

其中

$$\begin{aligned}(s^{(k)} - H_k y^{(k)})^\top y^{(l)} =& s^{(k)^\top} y^{(l)} - y^{(k)^\top} H_k y^{(l)} \\ =& s^{(k)^\top} y^{(l)} - s^{(l)^\top} y^{(k)} \\ =& s^{(k)^\top} G s^{(l)} - s^{(l)^\top} G s^{(k)} \\ =& 0\end{aligned}$$

二次终止性 （这里需要假设$s_0, \ldots, s_{n-1}$线性无关）由于

$$H_n y^{(l)} = H_n G s^{(l)} = s^{(l)},\ l = 0, 1, \ldots, n-1$$

而$s_0, \ldots, s_{n-1}$是$\mathbb{R}^n$的一组基，因此 $H_n G = I \Rightarrow H_n = G^{-1}$。而对二次函数，$x - G^{-1} \nabla f(x) = x - G^{-1}(Gx + c) = -G^{-1} c$直接得到全局最优解，因此

$$x^{(n+1)} = x^{(n)} - H_n g^{(k)}$$

为全局最优解。

Exercise 5　利用秩一校正的求逆公式（Sherman-Morrison定理），由$H_{k+1}^{(DFP)}$推导$B_{k+1}^{(DFP)}$。

$$(A + uv^\top)^{-1} = A^{-1} - \frac{A^{-1} u v^\top A^{-1}}{1 + v^\top A^{-1} u}$$

$$H_{k+1}^{(DFP)} = H_k + \frac{s^{(k)} s^{(k)^\top}}{s^{(k)^\top} y^{(k)}} - \frac{H_k y^{(k)} y^{(k)^\top} H_k}{y^{(k)^\top} H_k y^{(k)}}$$

为方便书写，忽视所有角标k。记$M = H + \frac{ss^\top}{s^\top y}$，则有

$$M^{-1} = H^{-1} - \frac{H^{-1} s s^\top H^{-1}}{s^\top y + s^\top H^{-1} s} = B - \frac{B s s^\top B}{s^\top y + s^\top B s} \tag{3}$$

$$B_{k+1}^{(DFP)} = \left(H_{k+1}^{(DFP)}\right)^{-1} = M^{-1} + \frac{M^{-1} H y y^\top H M^{-1}}{y^\top H y - y^\top H M^{-1} H y} \tag{4}$$

将(3)代入(4)的RHS第二项并展开，得

$$\frac{M^{-1}Hyy^\top HM^{-1}}{y^\top Hy - y^\top HM^{-1}Hy} = \frac{\left(B - \frac{Bss^\top B}{s^\top y + s^\top Bs}\right) Hyy^\top H \left(B - \frac{Bss^\top B}{s^\top y + s^\top Bs}\right)}{y^\top Hy - y^\top H \left(B - \frac{Bss^\top B}{s^\top y + s^\top Bs}\right) Hy}$$

利用$BH = I$，

$$y^\top Hy - y^\top HM^{-1}Hy = y^\top Hy - y^\top HBHy + \frac{y^\top HBss^\top BHy}{s^\top y + s^\top Bs} = \frac{\left(y^\top s\right)^2}{s^\top y + s^\top Bs}$$

$$\begin{aligned}\frac{M^{-1}Hyy^\top HM^{-1}}{y^\top Hy - y^\top HM^{-1}Hy} &= \frac{yy^\top(s^\top y + s^\top Bs)}{(y^\top s)^2} - \frac{ys^\top B}{y^\top s} - \frac{Bsy^\top}{y^\top s} + \frac{Bss^\top B}{s^\top y + s^\top Bs} \\ &= \left(1 + \frac{s^\top Bs}{y^\top s}\right)\frac{yy^\top}{y^\top s} - \frac{ys^\top B + Bsy^\top}{y^\top s} + \frac{Bss^\top B}{s^\top y + s^\top Bs}\end{aligned}$$

将上式和式(3)代入(4)即有

$$B_{k+1}^{(DFP)} = B_k + \left(1 + \frac{s^{(k)\top} B_k s^{(k)}}{y^{(k)\top} s^{(k)}}\right) \frac{y^{(k)} y^{(k)\top}}{y^{(k)\top} s^{(k)}} - \frac{B_k s^{(k)} y^{(k)\top} + y^{(k)} s^{(k)\top} B_k}{y^{(k)\top} s^{(k)}}$$

Exercise 6 共轭梯度法性质定理：设目标函数$f(x) = \frac{1}{2}x^\top Gx + c^\top x$，则采用精确一维搜索的共轭梯度法经$m \leqslant n$步迭代后终止，且对所有的$1 \leqslant k \leqslant m$成立下列关系：

$$d^{(k)\top} Gd^{(j)} = 0,\ j = 0, 1, \ldots, k-1 \tag{6.1}$$

$$g^{(k)\top} g^{(j)} = 0,\ j = 0, 1, \ldots, k-1 \tag{6.2}$$

$$d^{(k)\top} g^{(k)} = -g^{(k)\top} g^{(k)} \tag{6.3}$$

$$\operatorname{span}\{g^{(0)}, g^{(1)}, \ldots, g^{(k)}\} = \operatorname{span}\{g^{(0)}, Gg^{(0)}, \ldots, G^k g^{(0)}\} \tag{6.4}$$

$$\operatorname{span}\{d^{(0)}, d^{(1)}, \ldots, d^{(k)}\} = \operatorname{span}\{g^{(0)}, Gg^{(0)}, \ldots, G^k g^{(0)}\} \tag{6.5}$$

共轭梯度法步骤中得到的条件：

$$g^{(k+1)\top} d^{(k)} = 0 \qquad \text{(精确一维搜索)} \tag{6.6}$$

$$\alpha_k = -\frac{d^{(k)\top} g^{(k)}}{d^{(k)\top} Gd^{(k)}} \qquad \text{(精确一维搜索)} \tag{6.7}$$

$$g^{(k+1)} = g^{(k)} + \alpha_k Gd^{(k)} \qquad \text{(直接展开}\nabla f(x^{(k+1)})\text{)} \tag{6.8}$$

$$\beta_k = \frac{g^{(k+1)\top} g^{(k+1)}}{g^{(k)\top} g^{(k)}} \tag{6.9}$$

$$d^{(k+1)} = -g^{(k+1)} + \beta_k d^{(k)} \tag{6.10}$$

证明. (6.3)的证明

$$\begin{aligned}
{d^{(k)}}^\top g^{(k)} &= -{g^{(k)}}^\top g^{(k)} + \beta_k {d^{(k-1)}}^\top g^{(k)} && \text{(by (6.10))} \\
&= -{g^{(k)}}^\top g^{(k)} + 0 && \text{(by (6.6))} \\
&= -{g^{(k)}}^\top g^{(k)}
\end{aligned}$$

(6.1)与(6.2)的证明 k=1时直接验证可得结论成立若(6.1)与(6.2)对k成立，则对于$k+1$

1) 对(6.8)式左右两边转置后右乘$g^{(j)}$得

$$\begin{aligned}
g^{(k+1)} g^{(j)} =& {g^{(k)}}^\top g^{(j)} + \alpha_k {d^{(k)}}^\top G g^{(j)} \\
=& {g^{(k)}}^\top g^{(j)} + \alpha_k {d^{(k)}}^\top G (d^{(j)} - \beta_{j-1} d^{(j-1)}) \\
=& {g^{(k)}}^\top g^{(j)} + \alpha_k {d^{(k)}}^\top G d^{(j)}
\end{aligned}$$

$j = k$时，将(6.7)代入即可得上式为0；$j < k$时，由归纳假设得上式为0。综上，(6.2)成立。

2) 由(6.10)式，

$$\begin{aligned}
{d^{(k+1)}}^\top G =& (-g^{(k+1)} + \beta_k d^{(k)}) G d^{(j)} \\
=& -g^{(k+1)} G d^{(j)} + \beta_k d^{(k)} G d^{(j)} \\
=& {g^{(k+1)}}^\top (g^{(j)} - g^{(j+1)})/\alpha_k + \beta_k d^{(k)} G d^{(j)}
\end{aligned}$$

$j = k$时，由(6.2)(6.7)(6.9)得上式为0；$j < k$时，由归纳假设得上式为0。综上，(6.1)成立。

(6.4)与(6.5)的证明 由(6.10)式知，存在可逆方阵

$$Q = \begin{bmatrix}
-1 & \beta_0 & 0 & \dots & 0 & 0 \\
0 & -1 & \beta_1 & \dots & 0 & 0 \\
0 & 0 & -1 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \dots & -1 & \beta_{k-1} \\
0 & 0 & 0 & \dots & 0 & -1
\end{bmatrix}$$

使得$(d^{(0)}, d^{(1)}, \dots, d^{(k)}) Q = (g^{(0)}, g^{(1)}, \dots, g^{(k)})$，所以

$$\mathrm{span}\{d^{(0)}, d^{(1)}, \dots, d^{(k)}\} = \mathrm{span}\{g^{(0)}, g^{(1)}, \dots, g^{(k)}\}$$

下面只需要使用数学归纳法证明(6.4)。

1) $k=0$时，由定义直接得到结论成立。

2) 假设结论对k成立。对于$k+1$，由(6.8)式和归纳假设，

$$g^{(k+1)} = g^{(k)} + \alpha_k G d^{(k)} \in \operatorname{span}\{g^{(0)}, Gg^{(0)}, \dots, G^{k+1}g^{(0)}\}$$

即得

$$\operatorname{span}\{g^{(0)}, g^{(1)}, \dots, g^{(k+1)}\} \subseteq \operatorname{span}\{g^{(0)}, Gg^{(0)}, \dots, G^{k+1}g^{(0)}\}.$$

前面已证$g^{(l)} \perp g^{(j)}$, $0 \leqslant j < l \leqslant k+1$，因此 $\{g^{(0)}, g^{(1)}, \dots, g^{(k+1)}\}$线性无关，

$$\dim\left(\operatorname{span}\{g^{(0)}, g^{(1)}, \dots, g^{(k+1)}\}\right) = k+2$$

因此结论对$k+1$成立，即

$$\operatorname{span}\{g^{(0)}, g^{(1)}, \dots, g^{(k+1)}\} = \operatorname{span}\{g^{(0)}, Gg^{(0)}, \dots, G^{k+1}g^{(0)}\}.$$

□

Exercise 7　证明折线法（信赖域方法）子问题模型的函数单调性。

$$\begin{aligned}
s_C^{(k)} &= -\frac{\left\|g^{(k)}\right\|^2}{g^{(k)\top} B_k g^{(k)}} g^{(k)} \\
s_N^{(k)} &= -B_k^{-1} g^{(k)} \\
x^{(k+1)} &= x^{(k)} + s_C^{(k)} + \lambda(s_N^{(k)} - s_C^{(k)})
\end{aligned}$$

(1) 证明沿着Cauchy点$x_C^{(k+1)}$和牛顿点$x_N^{(k+1)}$的连线，到$x^{(k)}$的距离单调增加。

只要证明$\left(x_C^{(k+1)} - x^{(k)}\right)^\top \left(x_C^{(k+1)} - x_N^{(k+1)}\right) \leqslant 0$，即证明

$$\left(-\frac{g^{(k)\top} g^{(k)}}{g^{(k)\top} B_k g^{(k)}} g^{(k)}\right)^\top \left(B_k^{-1} g^{(k)} - \frac{g^{(k)\top} g^{(k)}}{g^{(k)\top} B_k g^{(k)}} g^{(k)}\right) \leqslant 0$$

只要证明

$$g^{(k)\top} B_k^{-1} g^{(k)} - \frac{g^{(k)\top} g^{(k)}}{g^{(k)\top} B_k g^{(k)}} g^{(k)\top} g^{(k)} \geqslant 0$$

$$\iff \quad g^{(k)\top} B_k^{-1} g^{(k)} g^{(k)\top} B_k g^{(k)} - \left(g^{(k)\top} g^{(k)}\right)^2 \geqslant 0 \qquad (*)$$

即可。

记$S_k = \sqrt{B_k}$, $a^{(k)} = S_k g^{(k)}$, $b^{(k)} = S_k^{-1} g^{(k)}$，则由Cauchy不等式

$$\left(a^{(k)\top} b^{(k)}\right)^2 \leqslant \left(a^{(k)\top} a^{(k)}\right)\left(b^{(k)\top} b^{(k)}\right)$$

式(*)得证。

(2)

$$
\begin{aligned}
h(\lambda) =& q^{(k)}\left(s_C + \lambda(s_N - s_C)\right) \\
=& f(x^{(k)}) + g^\top\left(s_C + \lambda(s_N - s_C)\right) + \frac{1}{2}\left(s_C + \lambda(s_N - s_C)\right)^\top B_k\left(s_C + \lambda(s_N - s_C)\right) \\
h'(\lambda) =& g^\top(s_N - s_C) + s_C^\top B_k(s_N - s_C) + \lambda(s_N - s_C)^\top B_k(s_N - s_C) \\
\leqslant& g^\top(s_N - s_C) + s_C^\top B_k(s_N - s_C) + (s_N - s_C)^\top B_k(s_N - s_C) \\
=& \left(g^\top + s_C^\top B_k + (s_N - s_C)^\top B_k\right)(s_N - s_C) \\
=& (g^\top + s_N^\top B_k)(s_N - s_C) \\
=& 0 \cdot (s_N - s_C) \\
=& 0
\end{aligned}
$$