1 无导数优化中关于二次函数 H^0 半范数的小思考

关于 Powell [1] 提出的扩展的对称 Broyden 修正，也即通过求解

$$
\begin{aligned}
&\min_{Q_k \in Q} \left\|\nabla^2 Q_k - \nabla^2 Q_{k-1}\right\|_{\mathrm{F}}^2 + \sigma \left\|\nabla Q_k(x_0) - \nabla Q_{k-1}(x_0)\right\|_2^2 \\
&s.t. Q_k(x) = f(x), \quad x \in \mathcal{I}_k
\end{aligned}
$$

来得到第 k 步迭代的模型 Q_k，文献 [2] 通过研究 H^1 半范数与扩展的对称 Broyden 修正的关系，从几何角度给出了参量 σ 的选取方式.

具体来说，设 x_0 为 $\mathbb{R}^n$ 中一点, r 为一正数, 且

$$
B_2^r(x_0) = \{x \in \mathbb{R}^n \big| \|x - x_0\| \leq r\},
$$

那么对任意 $Q \in \mathcal{Q}$, 我们有

$$
|Q|^2_{H^1(B_2^r(x_0))} = V_2 r^n [\frac{r^2}{n+2} \|\nabla^2 Q\|_F^2 + \|\nabla Q_1(x_0)\|_2^2].
$$

其中 V_2 是 $\mathbb{R}^n$ 中 $\mathcal{L}_2$ 单位的体积. 知下面的 $P_1(\sigma)$ 问题和 $P_2(r)$ 问题是等价的.

$$
\begin{aligned}
P_1(\sigma) : &\min_{Q \in \mathcal{Q}} \|\nabla^2 Q\|_F^2 + \sigma \|\nabla Q(x_0)\|_2^2 \\
&s.t.\ Q(x) = F(x), x \in \mathcal{I}\ (\text{插值系统}) \\
P_2(r) : &\min_{Q \in \mathcal{Q}} |Q|_{H^1(B_2^r(x_0))} \\
&s.t.\ Q(x) = F(x), x \in \mathcal{I}.
\end{aligned}
$$

在文献 [2] 中着重分析了利用 Sobolev 空间的 H^1 半范数进行的问题转化: 若 $\sigma > 0$, 则最小范数插值问题 $P_1(\sigma)$ 等价于问题 $P_2(r)$, 其中 $r = (n+2)/\sigma$. 通过转化为极小化球上的 H^1 半范数问题，从另一个角度解释了扩展的对称 Broyden 修正中参数 σ 的选取方式.

文献 [2] 告诉我们，对于二次函数 $Q(x) = \frac{1}{2} x^T B x + g x + c$，其在球 $B_2^r = \{x \in \mathbb{R}^n : \|x - x_0\|_2 \leq r\}$ 上的 H^0 半范数为

$$
|Q|^2_{H^0(B_2^r)} = V_2 r^n [\frac{r^4 (2\|B\|_F^2 + Tr^2 B)}{4(n+2)(n+4)} + \frac{r^2 (cTrB + \|g\|_2^2)}{n+2} + c^2].
$$

在此基础上继续，我们可以得到上面括号内为:

$$\frac{r^4}{2(n+2)(n+4)}\|B\|_F^2+\frac{r^2}{(n+2)}\|g\|_2^2+\frac{r^4}{4(n+2)(n+4)}Tr^2B+\frac{r^2\cdot c}{(n+2)}TrB+c^2$$

上式后三项可以配方为:

$$\begin{aligned}
&(\frac{r^2}{2\sqrt{(n+2)(n+4)}}TrB+\sqrt{\frac{n+4}{n+2}}c)^2-\frac{2c^2}{n+2}\\
=&\frac{1}{n+2}(\frac{r^2}{2\sqrt{n+4}}TrB+\sqrt{n+4}c)^2-\frac{2c^2}{n+2}\\
=&\frac{n+4}{n+2}(\frac{r^2}{2(n+4)}TrB+c)^2-\frac{2c^2}{n+2}\\
=&\frac{r^2(n+4)}{n+2}(\frac{r}{2(n+4)}TrB+\frac{c}{r})^2-\frac{r^2}{n+2}\frac{2c^2}{r^2}\\
=&\frac{r^2}{n+2}\big((n+4)(\frac{r}{2(n+4)}TrB+\frac{c}{r})^2-\frac{2c^2}{r^2}\big)\\
=&\frac{r^2}{n+2}\big(\frac{n+4}{r^2}(\frac{r^2}{2(n+4)}TrB+c)^2-\frac{2c^2}{r^2}\big).
\end{aligned}$$

进一步我们可以得到

$$|Q|^2_{H^0(B_2^r)}=V_2\frac{r^{n+2}}{(n+2)}[\frac{r^2}{2(n+4)}\|B\|_F^2+\|g\|_2^2+\frac{n+4}{r^2}(\frac{r^2}{2(n+4)}TrB+c)^2-\frac{2}{r^2}c^2].$$

则以下两个最小化问题等价:

$$\begin{aligned}
P_3(\sigma):&\min_{Q\in\mathcal{Q}}\|B\|_F^2+\sigma\|g\|_2^2+\frac{1}{2}(TrB+\sigma c)^2-\frac{\sigma^2}{n+4}c^2\\
&s.t.\ Q(x)=F(x),x\in\mathcal{I}\\
P_4(r):&\min_{Q\in\mathcal{Q}}|Q|_{H^0(B_2^r(x_0))}\\
&s.t.\ Q(x)=F(x),x\in\mathcal{I}\\
\text{即}:&\min_{Q\in\mathcal{Q}}\frac{r^2}{2(n+4)}\|B\|_F^2+\|g\|_2^2+\frac{n+4}{r^2}(\frac{r^2}{2(n+4)}TrB+c)^2-\frac{2}{r^2}c^2\\
&s.t.\ Q(x)=F(x),x\in\mathcal{I}
\end{aligned}$$

令 $\frac{2(n+4)}{r^2}=\sigma$ 即可验证. 对比问题 $P_1(\sigma)$ 和 $P_3(\sigma)$，知 $P_3(\sigma)$ 比 $P_1(\sigma)$ 多了两项:$\frac{1}{2}(TrB+\sigma c)^2$ 和 $-\frac{\sigma^2}{n+4}c^2$，某种意义上来说，有理由猜测 $P_1(\sigma)$ 问题没有充分

体现 TrB 和 c，也即二次函数 Q 的 Hessian 矩阵的迹以及常数项. 这同相对应的问题 $P_2(r)$ 和 $P_4(r)$ 分别是基于 H^1 半范数 (因求导而缺失了部分信息) 和 H^0 半范数的相吻合.

基于此，或许可以尝试将项 $\frac{1}{2}(TrB+\sigma c)^2$ 和 $-\frac{\sigma^2}{n+4}c^2$ 加入到原无导数优化迭代时确定模型函数 Q 所作的扩展的对称 Broyden 修正中. 也就是通过极小化

$$\begin{aligned} P_3(\sigma):& \min_{Q\in\mathcal{Q}} \|B\|_F^2+\sigma\|g\|_2^2+\frac{1}{2}(TrB+\sigma c)^2-\frac{\sigma^2}{n+4}c^2 \\ & s.t.\ Q(x)=F(x), x\in\mathcal{I} \end{aligned}$$

来确定模型函数 Q 的 B, g, c.

在实际计算中选择

$$r=\max\left\{M\Delta_k, \max_{y\in\mathcal{I}_k}\|y-x_k\|_2\right\}, \sigma=\frac{2(n+4)}{r^2}, x_0=x_k.$$

以上分析仅是一种思路雏形，数值实验有待进一步尝试.

$Q(x)=\frac{1}{2}x^\top Bx+g^\top x+c \quad \Lambda=B_r(x_0)=\{x\in R^n \mid \|x-x_0\|\leq r\}$

$B\in R^{n\times n}, \quad g\in R^n, x\in R^n, c\in R$

Frobenins 范数: $A=(a_{ij})_{nxn}. \quad \|A\|_F=\left(\sum_{ij}(a_{ij})^2\right)^{\frac{1}{2}}$

$$\begin{aligned} (Q)_{H^0(B_r)}^2 &= \int_{B_r(x_0)}(Q(x))^2dx=\int_{B_r(x_0)}\left(\frac{1}{2}x^\top Bx+g^\top x+c\right)^2dx \\ &= \int_{B_r(0)}\left(\frac{1}{2}(x+x_0)^\top B(x+x_0)+g^\top(x+x_0)+c\right)^2dx \\ &= \int_{Br}(0)\left(\frac{1}{2}\left(x^\top Bx+x^\top Bx_0+x_0^\top B_x+x_0^\top Bx_0\right)+g^\top x\right. \end{aligned}$$

$\left.+g^\top x_0+C\right)^2dx$

$=\int_{B_r}(0)\left(\frac{1}{4}\left(x^\top B_x+2x^\top B_{x_0}+x_0^TBx_0\right)^2+\left(g^\top x\right)^2\right.$

$+\left(g^\top x_0\right)^2+c^2+\left(x^\top Bx+2x^\top Bx_0+x_0^\top Bx_0\right)\left(g^\top x\right)$

$+\left(x^\top Bx+2x^\top Bx_0+x_0^\top B_{x_0}\right)\left(g^\top x_0+c\right)+2g^\top x\left(g^\top x_0+c\right)$

$+2(g^\top x_0)\,dx$

注意 (上面倒数第二行是)$+2(g^\top x_0)\,dx$ 还是 $+2cg^\top x_0)\,dx$ 打的是 $+2(g^\top x_0)\,dx$

先设 $x_0=0$: 则 $$\begin{aligned}(1) &= \int_{Br}(0)\frac{1}{4}\left(x^{\top}B_x\right)^2 dx=\frac{1}{4}I\left(2\|B\|_F^2+\operatorname{Tr}^2 B\right)\\ &+\frac{1}{4}(J-3I)\|D\|_F^2.\end{aligned}$$

其中 $I=\int_{B_r(0)}x_1^2x_2^2dx, \quad J=\int_{B_r(0)}x_1^4dx \;\; k=\int_{B_r(0)}x_1^2dx \quad \|D\|_F=\left(\sum_i b_{ii}^2\right)^{\frac{1}{2}}$

(2) $=\int_{B_r(0\rangle}\left(g^{\top}x\right)^2dx=k\|g\|_2^2$

(3) $=0$

(4) $=V_2r^nc^2$

(5) $=\int_{B_r(0)}\left(x^{\top}Bx\right)\left(g^{\top}x\right)dx=0$

(6) $=\int_{B_r(0)}\left(x^{\top}B_X\right)(c)=cKT_rB$

(7) $=\int_{B_r(0)} \quad 2cg^{\top}xdx=0$

(8) $=0$

所以

$\mid Q_{H^0}^2\left(\beta_r(0)\right)=\frac{1}{4}I\left(2\|B\|_F^2+T_r^2B\right)+\frac{1}{4}(J-3I)\ \|D\|_F^2+K\left(\|g\|_2^2+cT_rB\right)+V_2r^nc^2$

注意 (下面公式中不知道 τ 还是啥打成了 τ)

$$\begin{aligned}I &= \frac{\tau\left(\frac{3}{2}\right)^2\tau\left(\frac{n}{2}\right)}{\left(\tau\left(\frac{1}{2}\right)\right)^2\tau\left(\frac{n}{2}+2\right)}\cdot\frac{n}{n+4}v_2r^{n+4}\\ &=\frac{1}{4}\times\frac{1}{\left(\frac{n}{2}+1\right)\left(\frac{n}{2}\right)}\times\frac{n}{n+4}v_2r^{n+4}\\ &=\frac{1}{(n+4)(n+2)}v_2r^{n+4}\end{aligned}$$

$$\begin{aligned}J &= \frac{\tau\left(\frac{5}{2}\right)\tau\left(\frac{1}{2}\right)\tau\left(\frac{n}{2}\right)}{\left(\tau\left(\frac{1}{2}\right)\right)^2\tau\left(\frac{n}{2}+2\right)}\cdot\frac{n}{n+4}v_2r^{n+4}\\ &=\frac{3}{2}\times\frac{1}{2}\times\frac{1}{\left(\frac{n}{2}+1\right)\left(\frac{n}{2}\right)}\cdot\frac{n}{n+4}v_2r^{n+4}\\ &=\frac{3}{(n+4)(n+2)}v_2r^{n+4}=3I\end{aligned}$$

$$\begin{aligned}K &= \frac{\tau\left(\frac{3}{2}\right)\tau\left(\frac{1}{2}\right)\tau\left(\frac{n}{2}\right)}{\left(\tau\left(\frac{1}{2}\right)\right)^2\tau\left(\frac{n}{2}+1\right)}\cdot\frac{n}{n+2}\nu_2r^{n+2}\\ &=\frac{1}{2}\times\frac{2}{n}\times\frac{n}{n+2}v_2r^{n+2}\\ &=\frac{1}{n+2}v_2r^{n+2}\end{aligned}$$

所以 $|Q|_{H^0(B_r(0))}^2=V_2r^n\left(\frac{r^4}{4(n+4)(n+2)}\left(2\|B\|_F^2+\operatorname{Tr}^2B\right)+c^2+\frac{r^2}{n+2}\left(\|g\|_2^2+c\operatorname{Tr}B\right)\right)$

(半范数)

$|Q|^2_{H^1}(B_r(0)) = \int_{B_r(0)} \|B_x + g\|_2^2 \, dx$

$= \int_{B_r(0)} \left(x^\top B^2 x + \|g\|_2^2 \right) dx$

$= K\|B\|_F^2 + \|g\|_2^2 \cdot V_2 r^n$

$= V_2 r^n \left(\frac{r^2}{n+2} \|B\|_F^2 + \|g\|_2^2 \right)$ $|Q|^2_{H^2(B_r(0))} = \int_{B_r(0)} \|B\|_f^2 dx = V_2 r^n \|B\|_F^2$

(半范数)

H^1 范数 $|Q|^2_{H'} = v_2 r^n \left(\left(\frac{r^4}{2(n+4)(n+2)} + \frac{r^2}{n+2} \right) \|B\|_F^2 + \left(\frac{r^2}{n+2} + 1 \right) \|g\|_2^2 + \frac{r^2}{4(n+4)(n+2)} T_r^2 B + c^2 + \right.$

H^2 范数 $|Q|^2_{H^2} = V_2 r^n \left(\left(\frac{r^4}{2(n+4)(n+2)} + \frac{r^2}{n+2} + 1 \right) \|B\|_F^2 + \left(\frac{r^2}{n+2} + 1 \right) \|g\|_2^2 + \frac{r^2}{4(n+4)(n+2)} \operatorname{Tr}^2 B + c^2 + \right.$

参考文献

[1] MJD Powell. Beyond symmetric broyden for updating quadratic models in minimization without derivatives. *Mathematical Programming*, 138(1-2):475–500, 2013.

[2] Zaikun Zhang. Sobolev seminorm of quadratic functions with applications to derivative-free optimization. *Mathematical Programming*, 146(1-2):77–96, 2014.