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

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

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

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

那么对任意 $Q\in\mathcal{Q}$, 我们有

$$|Q|^2_{H^1(B_2^r(x_0))}=V_2r^n[\frac{r^2}{n+2}\|\nabla^2Q\|_F^2+\|\nabla Q_1(x_0)\|_2^2].$$

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.

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

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 $Q(x)=\frac{1}{2}x^TBx+gx+c$，其在球 $B_2^r=\{x\in\mathbb{R}^n:\|x-x_0\|_2\le r\}$ 上的 H^0 半范数为

$$|Q|^2_{H^0(B_2^r)}=V_2r^n[\frac{r^4(2\|B\|_F^2+Tr^2B)}{4(n+2)(n+4)}+\frac{r^2(cTrB+\|g\|_2^2)}{n+2}+c^2].$$

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.

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.