答案

2022 年 11 月 30 日

问题1 解:

$$r(x) = \begin{pmatrix} x_1^3 - x_2 - 1 \\ x_1^2 - x_2 \end{pmatrix};$$

由定义

$$J(x) = [\nabla r_1(x), \nabla r_2(x)]^\top = \begin{pmatrix} 3x_1^2 & -1 \\ 2x_1 & -1 \end{pmatrix};$$

$$\nabla f(x) = J(x)^\top r(x) = \begin{pmatrix} 3x_1^5 - 3x_1^2 x_2 - 4x_1^2 + x_2 \\ 2x_1^4 - 2x_1 x_2 - x_1^2 - 2x_1 + x_2 \end{pmatrix};$$

$$\nabla S(x) = \sum_{i=1}^{2} r_i(x) \nabla^2 r_i(x) = \begin{pmatrix} 6x_1^4 - 6x_1 x_2 + 2x_1^2 - 6x_1 - 2x_2 & 0 \\ 0 & 0 \end{pmatrix}.$$

问题2

证明.

$$2q(d_i) = d_i^T J^T J d_i + 2d_i^T J^T r + r^T r. \quad (*)$$

由已知条件

$$d_i^T J^T J d_i + \nu_i d_i^T d_i = -d_i^T J^T r,$$

因此(*)转为

$$\begin{aligned} 2q(d_i) &= -\nu_i d_i^T d_i + d_i^T J^T r + r^T r \\ &= -\nu_i r^T J (J^T J + v_i I)^{-2} J^T r + r^T J (J^T J + v_i I)^{-1} J^T r + r^T r, \end{aligned}$$

考虑对矩阵进行正交分解即:$(J^TJ+\nu_iI)^{-1} = Q^T\Lambda_iQ$,其中$\Lambda_I$为对角矩阵，其第$k$个对角元素为$\frac{1}{t_k+\nu_i}$($t_k$为$J^TJ$的第$k$个特征值.)于是上式可进一步转换为

$$\begin{aligned}2q(d_i) =& -\nu_ir^TJQ^T\Lambda_i^2QJ^Tr + r^TJQ^T\Lambda_iQJ^Tr + r^Tr\\ =& r^TJQ^T(-\nu_i\Lambda_i^2+\Lambda_i)QJ^Tr + r^Tr,\end{aligned}$$

注意矩阵$\nu_i\Lambda_i^2+\Lambda_i$的特征值为$\frac{t_i}{(t_i+\nu_i)^2}$为$\nu_i$的减函数,故有$2q(d_2) \le 2q(d_1)$.

□

问题3 解：令$g(x) = x_1-1$,构造拉格朗日函数$\mathcal{L}(x,\lambda) = f(x)+\lambda g(x)$,由KKT条件

$$\nabla(f(x^*)+\lambda^*g(x^*)) = 0; x_1^*-1 \ge 0; \lambda^* \le 0; \lambda^*g(x^*) = 0,$$

即:

$$\begin{pmatrix}1+\lambda^*\\ -2x_2^*\end{pmatrix} = 0.$$

由上述方程及$\lambda^*g(x^*) = 0$,解得$\lambda^* = -1, x^* = (1,0)$为唯一的KT点。但由f(x)的表达式，显然$x^* = (1,0)^\top$不是极小点。

问题4 解:优化问题对应的拉格朗日函数为

$$L(x,\lambda) = x_1^2 + x_2^2 - \lambda(\frac{x_1^2}{4} + x_2^2 - 1),$$

该问题在任意一点的线性化可行方向锥为

$$\mathcal{F}(x) = \{(d_1,d_2)|\frac{x_1}{4}d_1 + x_2d_2 = 0\},$$

因为只有一个等式约束且对应函数的梯度非0，故LICQ条件成立，且在KKT点对(x,λ)处有$\mathcal{F}_1(x) = \mathcal{F}(x)$,经计算可以得到四个KKT对:

$$(x^T,\lambda) = (2,0,-4),(-2,0,-4),(0,1,-1)(0,-1,-1).$$

对于第一个KKT点$y = (2,0,-4)$计算可得

$$\nabla_{xx}^2L(y) = \begin{bmatrix}0 & 0\\ 0 & -6\end{bmatrix}, \mathcal{F}(y) = \{(d_1,d_2)|d_1 = 0\},$$

取方向d=(0,1)则

$$d^T\nabla_{xx}^2L(y)d = -6 \le 0,$$

,同样对于第三个KKT点$z=(0,1,-1)$有

$$\nabla_{xx}^2 L(z)=\begin{bmatrix}\frac{3}{2} & 0\\ 0 & 0\end{bmatrix},\quad \mathcal{F}(z)=\{(d_1,d_2)|d_2=0\},$$

经验证z为局部极小点,也就说前两个KKT点不是极小解，后两个KKT点为严格局部极小解.

问题 5

证明. 采用反证法.假设在点x^*处有$\mathcal{T}_x(x^*)\cap\left\{d\mid\nabla f(x^*)^\top d<0\right\}\neq\emptyset$,令$d\in\mathcal{T}_x(x^*)\cap\{d\mid\nabla f(x^*)^\top d\}$,由可行方向的定义,存在$\{t_k\}_{k=1}^{\infty}$和$\{d_k\}_k^{\infty}$使得$x^*=t_kd_k\in\mathcal{X}$,其中$t_k\to 0$,且$d_k\to d$由于$\nabla f(x^*)^\top d<0$,对于充分大的$k$,我们有

$$\begin{aligned}f(x^*+t_kd_k)=&f(x^*)+t_k\nabla f(x^*)^\top d_k+o(t_k)\\ <&f(x^*).\end{aligned}$$

这与x^*的局部极小性矛盾.

□

问题 6 解：这是一个带有简单约束的优化问题，它等价于下面的无约束优化问题

$$min_x x_1+x_1^2$$

显然，它有最优解$x^*=(-\frac{1}{2},\frac{1}{4}),f(x^*)=-\frac{1}{4}$
现在我们用外罚函数法求解．构造外罚函数

$$P(x,\sigma)=x_1+x_2+\sigma(x_2-x_1^2)^2$$

利用解析法求解

$$\tfrac{\partial P}{\partial x_1}=1-4\sigma x_1(x_2-x_1^2),\tfrac{\partial P}{\partial x_2}=1+2\sigma(x_2-x_1^2)$$

令

$$\nabla_x P(x,\sigma)=0$$

得到

$$x_1(\sigma)=-\tfrac{1}{2},x_2(\sigma)=\tfrac{1}{4}-\tfrac{1}{2\sigma}$$

再令$\sigma \to +\infty$，得

$$x(\sigma) = (x_1(\sigma), x_2(\sigma))^T \to x^* = (-\tfrac{1}{2}, \tfrac{1}{4})^T$$

$$P(x(\sigma), \sigma) = -\tfrac{1}{4} - \tfrac{1}{4\sigma} \to f(x^*) = -\tfrac{1}{4}$$

就是原问题的最优解。

问题 7　解：构造障碍函数

$$P(x(r), r) = x_1^2 + 2x_2^2 - rln(x_1 + x_2 - 1)$$

利用解析法，有

$$\nabla_x P(x(r), r) = (2x_1 - \tfrac{r}{x_1+x_2-1}, 4x_2 - \tfrac{r}{x_1+x_2-1}) = 0$$

得

$$x(r) = (\tfrac{1+\sqrt{1+3r}}{3}, \tfrac{1+\sqrt{1+3r}}{6})^T$$

令$r \to 0$，则$x(r) \to x^* = (\frac{2}{3}, \frac{1}{3}), P(x(r), r) \to f(x^*) = \frac{2}{3}$

问题 8　解：增广Lagrange函数为$\phi(x, \lambda, \sigma) = 2x_1^2 + x_2^2 - 2x_1x_2 - \lambda(x_1 + x_2 - 1) + \frac{\sigma}{2}(x_1 + x_2 - 1)^2$
取$\sigma = 2, \lambda^1 = 1$，利用解析法求解

$$min_x \phi(x, 1, 2)$$

得到极小点

$$x^1 = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{4} \end{pmatrix} \tag{1}$$

修正λ，有$\lambda^2 = \lambda^1 - \sigma c(x^1) = \frac{1}{2}$
再解

$$min_x \phi(x, \tfrac{1}{2}, 2)$$

得到x^2。如此继续，一般地，在第k次迭代时，$\phi(x, \lambda^k, 2)$的极小值为

$$x^k = \begin{pmatrix} \frac{1}{6}(2 + \lambda^k) \\ \frac{1}{4}(2 + \lambda^k) \end{pmatrix} \tag{2}$$

易见，当$k \to +\infty$时，$\lambda^k \to \frac{2}{5}$，$x^k = (\frac{2}{5}, \frac{3}{5})^T$，即分别为所求的最优乘子和最优解。