最优化第十一次作业

张晋 15091060

2017 年 11 月 23 日

5.3 (a)

$$设 \quad q(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T G\boldsymbol{x} - \boldsymbol{b}\boldsymbol{x} \tag{1}$$

由题意可知:

$$\nabla q(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x}) = G\boldsymbol{x} - \boldsymbol{b} \tag{2}$$

$$\nabla^2 \boldsymbol{q}\,(\boldsymbol{x}) = G \tag{3}$$

$$G\boldsymbol{s} = \lambda\boldsymbol{s} \tag{4}$$

$$\boldsymbol{g}(\boldsymbol{x}^{(0)}) = \boldsymbol{g}^{(0)} = G\boldsymbol{x}^{(0)} - \boldsymbol{b} \tag{5}$$

$$\boldsymbol{g}(\boldsymbol{x}^{(\star)}) = \boldsymbol{g}^{\star} = G\boldsymbol{x}^{\star} - \boldsymbol{b} = \boldsymbol{0} \tag{6}$$

$$\boldsymbol{g}^{(0)} = \boldsymbol{g}^{(0)} - \boldsymbol{g}^{(\star)} = G(\boldsymbol{x}^{(0)} - \boldsymbol{x}^{\star}) = G\mu\boldsymbol{s} = \mu\lambda\boldsymbol{s} \tag{7}$$

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} - \frac{-\boldsymbol{g}^{(0)^T}\boldsymbol{g}^{(0)}}{\boldsymbol{g}^{(0)^T}\boldsymbol{G}\boldsymbol{g}^{(0)}} \tag{8}$$

$$= \boldsymbol{x}^{(0)} - \frac{\mu^2\lambda^2\boldsymbol{s}^T\boldsymbol{s}}{\mu^2\lambda^2\lambda\boldsymbol{s}^T\boldsymbol{s}}\mu\lambda\boldsymbol{s} \tag{9}$$

$$= \boldsymbol{x}^{(0)} - \mu\boldsymbol{s} \tag{10}$$

$$= \boldsymbol{x}^{\star} \tag{11}$$

(b) 若 $G = \lambda\boldsymbol{I}$, 则对于 $\boldsymbol{x}^{(0)} - \boldsymbol{x}^{\star} = \lambda\boldsymbol{t}$, 易见 $\boldsymbol{t}$ 为 G 相对于 λ 的特征向量，由 (1) 可知对任意的初始点 $\boldsymbol{x}^{(0)}$, 其沿最速下降方向进行一次精确线搜索后达到最优解

(c)

$$g^{(0)} = G\left(x^{(0)} - x^*\right) \tag{12}$$
$$= G\sum_{i=1}^{m}\mu_i \boldsymbol{s}_i \tag{13}$$
$$= \sum_{i=1}^{m}\mu_i\lambda_i\boldsymbol{s}_i \tag{14}$$

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} - \frac{-\boldsymbol{g}^{(0)^T}\boldsymbol{g}^{(0)}}{\boldsymbol{g}^{(0)^T}\boldsymbol{G}\boldsymbol{g}^{(0)}} \tag{15}$$
$$= \boldsymbol{x}^{(0)} - \frac{\sum_{1=1}^{m}\sum_{j=1}^{m}\mu_i\mu_j\lambda_i\lambda_j\boldsymbol{s}_i^T\boldsymbol{s}_i}{\sum_{1=1}^{m}\sum_{j=1}^{m}\mu_i\mu_j\lambda_i\lambda_j^2\boldsymbol{s}_i^T\boldsymbol{s}_i}\sum_{i=1}^{m}\mu_i\lambda_i s_i \tag{16}$$
$$= \boldsymbol{x}^{(0)} - \frac{\sum_{1=1}^{m}\mu_i^2\lambda_i^2\boldsymbol{s}_i^T\boldsymbol{s}_i}{\sum_{1=1}^{m}\mu_i^2\lambda_i^3\boldsymbol{s}_i^T\boldsymbol{s}_i}\sum_{i=1}^{m}\mu_i\lambda_i\boldsymbol{s}_i \tag{17}$$
$$= \boldsymbol{x}^{(0)} - \left(\sum_{i=1}^{m}\frac{1}{\lambda_i}\right)\left(\sum_{i=1}^{m}\mu_i\lambda_i\boldsymbol{s}_i\right) \tag{18}$$

$$G\boldsymbol{x}^{(1)} - \boldsymbol{b} = Gx^{(1)} - Gx^{(0)} \tag{19}$$
$$= G(x^{(1)} - x^{(0)}) \tag{20}$$
$$= -G\left(\sum_{i=1}^{m}\frac{1}{\lambda_i}\right)\left(\sum_{i=1}^{m}\mu_i\lambda_i\boldsymbol{s}_i\right) \tag{21}$$
$$= \left(\sum_{i=1}^{m}\frac{1}{\lambda_i}\right)\left(\sum_{i=1}^{m}\mu_i\lambda_i^2\boldsymbol{s}_i\right) \tag{22}$$

由于 $\boldsymbol{s}_i$ 线性无关，故式 (22) 必然不等于 0，即证

5.4 由于其为精确线搜索，故有:$\nabla f\left(x^{(k+1)}\right)^T p\left(k\right) = 0$，即 $\boldsymbol{g}^{(k+1)^T}\boldsymbol{p}^{(k)} = -\boldsymbol{p}^{(k+1)^T}\boldsymbol{p}^{(k)} = 0$

$$G = \begin{bmatrix} 20 & 0 \\ 0 & 2 \end{bmatrix}$$

得 $\lambda_1 = 20, \lambda_2 = 2, \quad (\frac{\lambda_1 - \lambda_n}{\lambda_1 + x_n})^2 = 81/121 = 0.66942$

```
第 45 次迭代结果:
[ -1.19740783e-05   1.19740783e-04]
r=0.669421
第 46 次迭代结果:
[  9.79697313e-06   9.79697313e-05]
r=0.669421
第 47 次迭代结果:
[ -8.01570529e-06   8.01570529e-05]
r=0.669421
第 48 次迭代结果:
[  6.55830433e-06   6.55830433e-05]
r=0.669421
第 49 次迭代结果:
[ -5.36588536e-06   5.36588536e-05]
r=0.669421
第 50 次迭代结果:
[  4.39026984e-06   4.39026984e-05]
r=0.669421
第 51 次迭代结果:
[ -3.59203896e-06   3.59203896e-05]
```

图 1: 输出

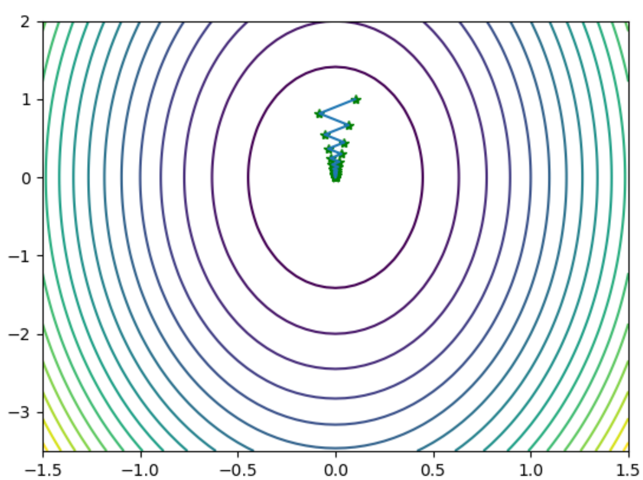

图 2: 图像

5.6

$$G = \begin{bmatrix} 10 & -9 \\ -9 & 10 \end{bmatrix}, \qquad \lambda_1 = 19, \lambda_2 = 1, (\frac{\lambda_1 - \lambda_2}{\lambda_1 + x_2})^2 = 0.81$$

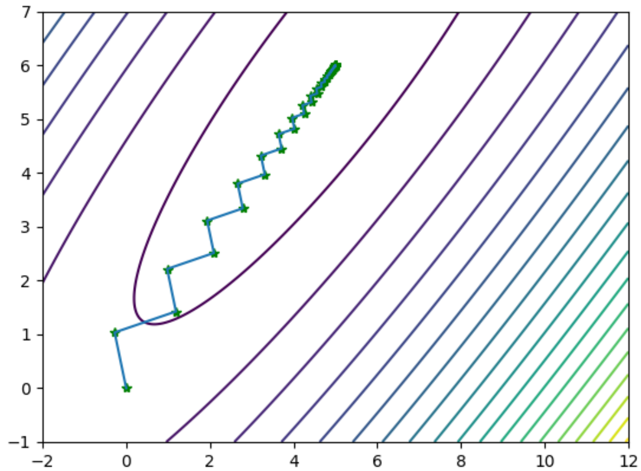

图 3: 初始点为 $(0, 0)$ 的图像

```
第 65 次迭代结果:
[ 4.99910969  5.99916235]
r=0.762255
第 66 次迭代结果:
[ 4.99935689  5.99922826]
r=0.762255
第 67 次迭代结果:
[ 4.99932136  5.99936149]
r=0.762255
第 68 次迭代结果:
[ 4.99950978  5.99941174]
r=0.762255
第 69 次迭代结果:
[ 4.9994827  5.9995133]
r=0.762255
第 70 次迭代结果:
[ 4.99962633  5.9995516 ]
r=0.762256
```

图 4: 初始点为 $(0, 0)$ 的输出

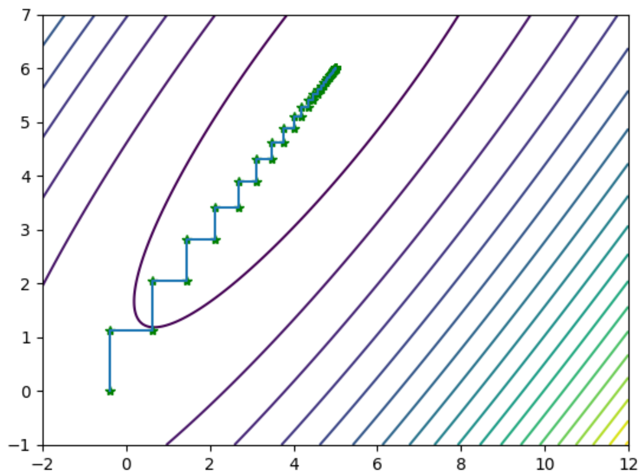

图 5: 初始点为 $(-0.4, 0)$ 的图像

```
第 85 次迭代结果:
[ 4.99922596  5.99930336]
r=0.810000
第 86 次迭代结果:
[ 4.99937303  5.99930336]
r=0.810000
第 87 次迭代结果:
[ 4.99937303  5.99943572]
r=0.810000
第 88 次迭代结果:
[ 4.99949215  5.99943572]
r=0.810000
第 89 次迭代结果:
[ 4.99949215  5.99954294]
r=0.810000
第 90 次迭代结果:
[ 4.99958864  5.99954294]
r=0.810000
```

图 6: 初始点为 $(-0.4, 0)$ 的输出

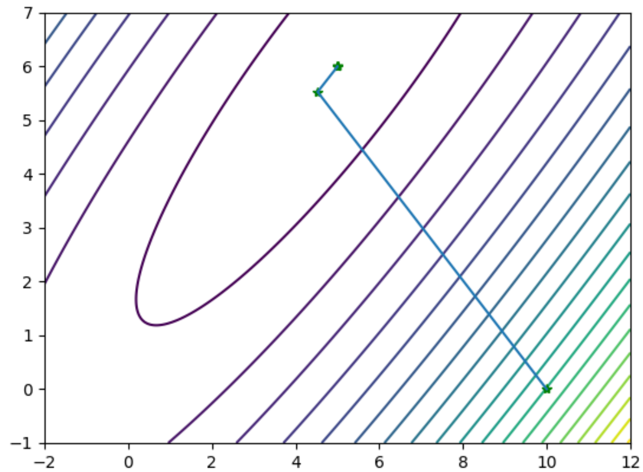

图 7: 初始点为 $(10, 0)$ 的图像

```
第 1 次迭代结果:
[ 4.52619707  5.52643565]
r=0.000390
第 2 次迭代结果:
[ 5.0019511   5.99765868]
r=0.000390
第 3 次迭代结果:
[ 4.99981511  5.99981521]
r=0.000390
第 4 次迭代结果:
[ 5.00000076  5.99999909]
r=0.000391
第 5 次迭代结果:
[ 4.99999993  5.99999993]
r=0.000000
```

图 8: 初始点为 $(10, 0)$ 的输出

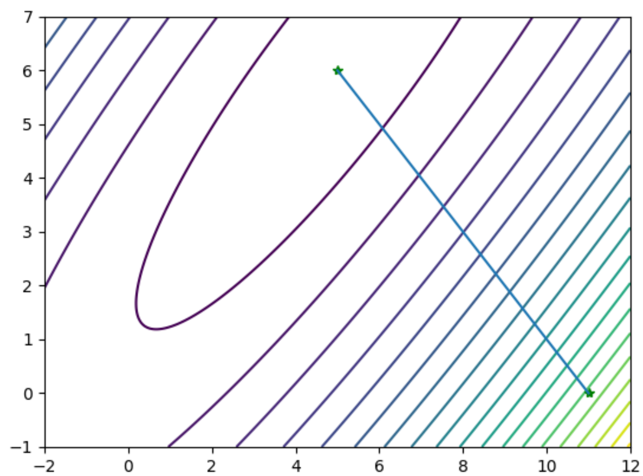

图 9: 初始点为 (11,0) 的图像

```
第 1 次迭代结果:
[ 5.  6.]
r=0.000000
```

图 10: 初始点为 (11,0) 的输出

综上: 第三次迭代时取到序列极限最大值 0.81，根据计算此即是线性收敛因子。

5.11

$$f(x,y) := \frac{1}{2}\left(x^2+y^2\right)e^{x^2-y^2}$$

$$\nabla_{\{x,y\}} f(x,y) = \left[xe^{x^2-y^2} + xe^{x^2-y^2}\left(x^2+y^2\right), ye^{x^2-y^2} - ye^{x^2-y^2}\left(x^2+y^2\right)\right]^T$$

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x^2} \\ \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 4e^{x^2-y^2}x^2 + 2e^{x^2-y^2}\left(x^2+y^2\right)x^2 + e^{x^2-y^2} + e^{x^2-y^2}\left(x^2+y^2\right) \\ -2e^{x^2-y^2}xy\left(x^2+y^2\right) \\ -4e^{x^2-y^2}y^2 + 2e^{x^2-y^2}\left(x^2+y^2\right)y^2 + e^{x^2-y^2} - e^{x^2-y^2}\left(x^2+y^2\right) \end{bmatrix} \tag{23}$$

对于点 $(0,0)^T$，有：

$$\nabla_{\{0,0\}} f(x,y) = [0,0]^T$$

$$G(0,0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

故其是局部极小点。

对于点 $(1,1)^T$，有：

$$\nabla_{\{1,1\}} f(x,y) = [0,0]^T$$

$$G' = G(1,1) = \begin{bmatrix} 11 & -4 \\ -4 & -1 \end{bmatrix}$$

故其 Hessian 非正定. 显然 $\lambda = 3$ 是使得 $G' + \lambda \boldsymbol{I}$ 正定的最小整数由于 $\nabla_{\{1,1\}} f(x,y) = \mathbf{0}$，故 $\boldsymbol{s} = 0, \quad \boldsymbol{x}'' = \boldsymbol{x}' = (1,1)^T$