CX 4640 Homework 5

Wenqi He

December 4, 2017

1

1.1 first-order optimality condition

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2 \\ 2x_2 \\ -2x_3 + 4 \end{pmatrix}, \quad \nabla f(\boldsymbol{x}^*) = \begin{pmatrix} 2 \times 2.5 - 2 \\ 2 \times -1.5 \\ -2 \times -1 + 4 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \\ 6 \end{pmatrix}$$

$$\boldsymbol{J}_g^T = \nabla g = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

The first-order optimality condition is

$$\nabla f(\boldsymbol{x}^*) + \boldsymbol{J}_g^T(\boldsymbol{x}^*)\boldsymbol{\lambda} = \begin{pmatrix} 3 \\ -3 \\ 6 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \lambda = 0$$

$\lambda^* = -3$ satisfies the condition.

1.2 second-order optimality condition

$$\boldsymbol{H}_f = \boldsymbol{J}_{\nabla f} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

$$\boldsymbol{H}_g = \boldsymbol{J}_{\nabla g} = \boldsymbol{0}$$

$$\boldsymbol{B}(\boldsymbol{x}^*, \lambda^*) = \boldsymbol{H}_f + \lambda \boldsymbol{H}_g = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} + \boldsymbol{0} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Now find the null space of $\boldsymbol{J}_g$:

$$\boldsymbol{J}_g \boldsymbol{x} = \begin{pmatrix} 1 & -1 & 2 \end{pmatrix} \boldsymbol{x} = 0$$

The solution is $x_1 = x_2 - 2x_3$, where x_2 and x_3 are free variables. Now we can construct a $\boldsymbol{Z}$ whose column space is the null space of $\boldsymbol{J}_g$:

$$\boldsymbol{Z} = \begin{pmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\boldsymbol{B}^T \boldsymbol{Z} \boldsymbol{B} = \begin{pmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ -4 & 6 \end{pmatrix}$$

$$\boldsymbol{v}^T \boldsymbol{B}^T \boldsymbol{Z} \boldsymbol{B} \boldsymbol{v} = 4v_1^2 - 8v_1 v_2 + 6v_2^2 = (2v_1 - 2v_2)^2 + 2v_2^2 \geq 0$$

Since $\boldsymbol{B}^T \boldsymbol{Z} \boldsymbol{B}$ is positive definite, the point $(\boldsymbol{x}^*, \lambda^*)$ satisfies the second-order optimality condition.

2

(a)

$$\begin{aligned}
\nabla f(\boldsymbol{x}) &= \frac{1}{2} \partial^k (x_i A^i_j x^j) - \partial^k (x_i b^i) + \partial^k c \\
&= \frac{1}{2} \big((\partial^k x_i) A^i_j x^j + x_i A^i_j (\partial^k x_j) \big) - (\partial^k x_i) b^i \\
&= \frac{1}{2} \big((\partial^k x_i) A^i_j x^j + (\partial^k x_j) A^j_i x^i \big) - (\partial^k x_i) b^i \\
&= \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}
\end{aligned}$$

$$\boldsymbol{H}_f(\boldsymbol{x}) = \boldsymbol{J}_{\nabla f} = \partial_k (A^i_j x^j) + \partial_k b^i = A^i_j (\partial_k x^j) = \boldsymbol{A}$$

Using Newton's method:

$$\boldsymbol{H}_f(\boldsymbol{x}_0) \boldsymbol{s}_0 = -\nabla f(\boldsymbol{x}_0)$$

$$\boldsymbol{A} \boldsymbol{s}_0 = -\boldsymbol{A} \boldsymbol{x}_0 + \boldsymbol{b}$$

After the first iteration:

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 + \boldsymbol{s}_0$$

$$\begin{aligned}
\nabla f(\boldsymbol{x}_1) &= \boldsymbol{A}\boldsymbol{x}_1 - \boldsymbol{b} = \boldsymbol{A}(\boldsymbol{x}_0 + \boldsymbol{s}_0) - \boldsymbol{b} \\
&= \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{A}\boldsymbol{s}_0 - \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{b} - \boldsymbol{b} \\
&= \boldsymbol{0}
\end{aligned}$$

(b)

Using the steepest descent method,

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - \alpha \nabla f(\boldsymbol{x}_0),$$

where α minimizes $f(\boldsymbol{x})$ along the direction of negative gradient. From (a),

$$\nabla f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$$

The fact that $\boldsymbol{x}^*$ is the solution means that

$$\nabla f(\boldsymbol{x}^*) = \boldsymbol{A}\boldsymbol{x}^* + \boldsymbol{b} = \boldsymbol{0}$$

The fact that $\boldsymbol{x}_0 - \boldsymbol{x}^*$ is an eigenvector of $\boldsymbol{A}$ means that there exists some λ such that

$$\boldsymbol{A}(\boldsymbol{x}_0 - \boldsymbol{x}^*) = \lambda(\boldsymbol{x}_0 - \boldsymbol{x}^*)$$

$$\boldsymbol{A}\boldsymbol{x}_0 = \lambda(\boldsymbol{x}_0 - \boldsymbol{x}^*) + \boldsymbol{A}\boldsymbol{x}^*$$

Plug this result in the update function:

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - \alpha(\boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{b})$$

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - \alpha\big(\lambda(\boldsymbol{x}_0 - \boldsymbol{x}^*) + \boldsymbol{A}\boldsymbol{x}^* + \boldsymbol{b}\big)$$

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - \alpha\lambda(\boldsymbol{x}_0 - \boldsymbol{x}^*)$$

When $\alpha = \lambda^{-1}$, f reaches a critical point,

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - (\boldsymbol{x}_0 - \boldsymbol{x}^*) = \boldsymbol{x}^*$$

$$\nabla f(\boldsymbol{x}_1) = \boldsymbol{0}$$

Therefore the method with the given starting point converges in one iteration.

3

(a)

$$\nabla f = \begin{pmatrix} 2x \\ 2y \end{pmatrix}, \quad \nabla g = \begin{pmatrix} -3(x-1)^2 \\ 2y \end{pmatrix}$$

Using the method of Lagrange multiplier, we need to solve the equation:

$$\nabla f + \lambda \nabla g = \boldsymbol{0}$$

Or:

$$\begin{aligned} 2x - 3\lambda(x-1)^2 &= 0 \\ 2y + 2\lambda y &= 0 \\ y^2 - (x-1)^3 &= 0 \end{aligned}$$

From the second equation, we must have either $y = 0$ or $\lambda = -1$. If $y = 0$, from the third equation we can conclude that $x = 1$, however this result does not satisfy the first equation. On the other hand, if $\lambda = -1$, the first equation becomes

$$3x^2 - 4x + 3 = 0,$$

which does not have real solutions. Therefore this problem cannot be solved using Lagrange multipliers.

(b)

The penalty function is

$$\phi_\rho(x, y) = x^2 + y^2 + \frac{1}{2}\rho \left(y^2 - (x-1)^3\right)^2$$

$$\nabla \phi_\rho(x, y) = \begin{pmatrix} 2x - 3\rho(y^2 - (x-1)^3)(x-1)^2 \\ 2y + 2\rho(y^2 - (x-1)^3)y \end{pmatrix} = \mathbf{0}$$

From the second equation, if $y \neq 0$

$$\rho(y^2 - (x-1)^3) = -1$$

The first equation $2x + 3(x-1)^2 = 0$ does not have real solutions, therefore y must be 0. Then we have:

$$2x + 3\rho(x-1)^5 = 0$$

$$\lim_{\rho \to \infty} (x-1)^5 = \lim_{\rho \to \infty} -\frac{2x}{3\rho} = 0$$

$$x = 1$$

So $(1, 0)$ is the minimizer.

4

(a)

There are 5 edges, so there must be 5 vertices. They are:

$$\left(\frac{12}{11}, \frac{6}{11}\right), \left(\frac{6}{5}, 0\right), \left(\frac{6}{7}, \frac{6}{7}\right), \left(0, \frac{3}{2}\right), (0, 0)$$

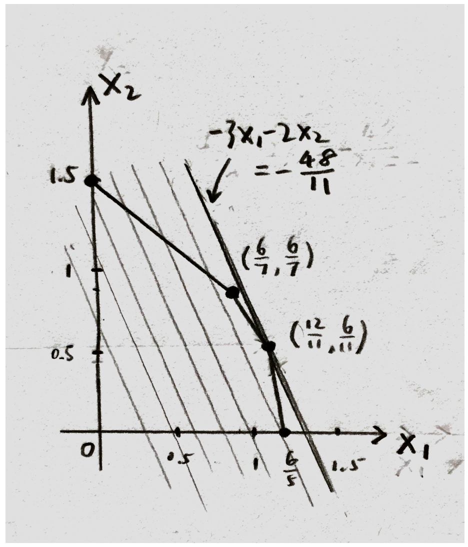
(b)

$$f(\frac{12}{11}, \frac{6}{11}) = -\frac{48}{11}, \quad f(\frac{6}{5}, 0) = -\frac{18}{5}$$

$$f(\frac{6}{7}, \frac{6}{7}) = -\frac{30}{7}, \quad f(0, \frac{3}{2}) = -3, \quad f(0, 0) = 0$$

From the above results, we can see that $(\frac{12}{11}, \frac{6}{11})$ is the minimizer.

(c)



5

(a)

See `mygn.m`
Using Gauss-Newton method, the least squares solution is

$$x_1 = 14.3766$$
$$x_2 = -1.5139$$

.

(b)

See `linear.m`
The result obtained by linear least squares method is

$$x_1 = 8.6350$$
$$x_2 = -1.0967$$

which is different from that of part (a) because the objective here is to minimize

$$\sum_i \big(\log(y_i) - \log(x_1) - x_2 t_i\big)^2 = \sum_i \log^2\left(\frac{y_i}{x_1 e^{x_2 t_i}}\right) = \sum_i \log^2\left(\frac{y_i}{f(t_i, \boldsymbol{x})}\right)$$

whereas Gauss-Newton method in part (a) minimizes

$$\sum_i \left(y_i - f(t_i, \boldsymbol{x})\right)^2$$

6

(a)

See `ueig.m`

(b)

See `ceig.m`

7

(a)

See `qlopt.m`

(b)

See `qnopt.m`

(c)

See `nnopt.m`