CS257 Linear and Convex Optimization Homework 8

Xiangyuan Xue (521030910387)

1. (a) Substitute the constraint $x_1 = 1 - 2x_2$ into the quadratic function

$$\begin{aligned} f(1-2x_2, x_2) &= (1-2x_2)^2 + (1-2x_2)x_2 + x_2^2 - (1-2x_2) - 3x_2 \\ &= 3x_2^2 - 4x_2 \\ &= 3(x_2 - \frac{2}{3})^2 - \frac{4}{3} \end{aligned}$$

Thus, the minimum value is

$$f^* = -\frac{4}{3}$$

where the optimal solution is

$$x_1^* = -\frac{1}{3}, x_2^* = \frac{2}{3}$$

(b) Write down the Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda) = x_1^2 + x_1x_2 + x_2^2 - x_1 - 3x_2 + \lambda(x_1 + 2x_2 - 1)$$

By the Lagrange condition

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x_1} = 2x_1 + x_2 + \lambda - 1 = 0 \\ \dfrac{\partial \mathcal{L}}{\partial x_2} = x_1 + 2x_2 + 2\lambda - 3 = 0 \\ \dfrac{\partial \mathcal{L}}{\partial \lambda} = x_1 + 2x_2 - 1 = 0 \end{cases}$$

which yields

$$x_1^* = -\frac{1}{3}, x_2^* = \frac{2}{3}, \lambda^* = 1$$

Since the Hessian matrix

$$\nabla^2 f(\boldsymbol{x}) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \succ \boldsymbol{O}$$

The problem is convex, so the solution is the optimal solution.

2. Since f is continuous and the feasible set is compact, the global minimum must exist. Let

$$h(x_1, x_2) = x_1^2 + \frac{1}{8}x_2^2 - 1$$

then

$$\nabla h(\boldsymbol{x}) = \begin{pmatrix} 2x_1 \\ \frac{1}{4}x_2 \end{pmatrix} \neq \mathbf{0}$$

Thus, any x^* will be regular point of h. Write down the Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda) = x_1x_2 + x_1^2 + \lambda(x_1^2 + \frac{1}{8}x_2^2 - 1)$$

By the Lagrange condition

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x_1} = 2(1+\lambda)x_1 + x_2 = 0 \\ \dfrac{\partial \mathcal{L}}{\partial x_2} = x_1 + \dfrac{1}{4}\lambda x_2 = 0 \\ \dfrac{\partial \mathcal{L}}{\partial \lambda} = x_1^2 + \dfrac{1}{8}x_2^2 - 1 = 0 \end{cases}$$

which yields 4 solutions

$$\begin{cases} x_1 = \dfrac{1}{\sqrt{3}} \\ x_2 = -\dfrac{4}{\sqrt{3}} \\ \lambda = 1 \end{cases}, \quad \begin{cases} x_1 = -\dfrac{1}{\sqrt{3}} \\ x_2 = \dfrac{4}{\sqrt{3}} \\ \lambda = 1 \end{cases}, \quad \begin{cases} x_1 = \dfrac{\sqrt{2}}{\sqrt{3}} \\ x_2 = \dfrac{2\sqrt{2}}{\sqrt{3}} \\ \lambda = -2 \end{cases}, \quad \begin{cases} x_1 = -\dfrac{\sqrt{2}}{\sqrt{3}} \\ x_2 = -\dfrac{2\sqrt{2}}{\sqrt{3}} \\ \lambda = -2 \end{cases}$$

where the first 2 solutions are the minimum. Therefore, the minimum value is $f^* = -1$, where the optimal solutions are $x^* = \left(\frac{1}{\sqrt{3}}, -\frac{4}{\sqrt{3}}\right)^T$ and $x^* = \left(-\frac{1}{\sqrt{3}}, \frac{4}{\sqrt{3}}\right)^T$.

3. (a) Write down the Lagrangian

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{g}^T\boldsymbol{x} + c + \boldsymbol{\lambda}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})$$

The Lagrange condition is

$$\begin{cases} \nabla_{\boldsymbol{x}}\mathcal{L} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{g} + \boldsymbol{A}^T\boldsymbol{\lambda} = \mathbf{0} \\ \nabla_{\boldsymbol{\lambda}}\mathcal{L} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} = \mathbf{0} \end{cases}$$

(b) Since $\boldsymbol{Q} \succ \boldsymbol{O}$, $\boldsymbol{Q}^{-1} \succ \boldsymbol{O}$, for any $\boldsymbol{y} \neq \mathbf{0}$ we have

$$\boldsymbol{y}^T\boldsymbol{Q}^{-1}\boldsymbol{y} > 0$$

Since $\operatorname{rank}(\boldsymbol{A}) = k$, for any $\boldsymbol{x} \neq \mathbf{0}$ we have $\boldsymbol{A}^T\boldsymbol{x} \neq \mathbf{0}$, thus

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{A}^T\boldsymbol{x} > 0$$

which yields $\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{A}^T \succ \boldsymbol{O}$ and hence is invertible.

From the first equation we have

$$\boldsymbol{x} = -\boldsymbol{Q}^{-1}\left(\boldsymbol{g} + \boldsymbol{A}^T\boldsymbol{\lambda}\right)$$

Substitute it into the second equation

$$-\boldsymbol{A}\boldsymbol{Q}^{-1}\left(\boldsymbol{g} + \boldsymbol{A}^T\boldsymbol{\lambda}\right) = \boldsymbol{b}$$

which yields

$$\boldsymbol{\lambda}^* = -\left(\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{A}^T\right)^{-1}\left(\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{g} + \boldsymbol{b}\right)$$

Substitute it into the initial expression

$$\boldsymbol{x}^* = \boldsymbol{Q}^{-1}\boldsymbol{A}^T\left(\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{A}^T\right)^{-1}\left(\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{g} + \boldsymbol{b}\right) - \boldsymbol{Q}^{-1}\boldsymbol{g}$$

(c) Let $\boldsymbol{Q} = \boldsymbol{I}$, $\boldsymbol{g} = -\boldsymbol{x}_0$, $c = \frac{1}{2}\boldsymbol{x}_0^T\boldsymbol{x}_0$, and we have

$$\mathrm{Proj}_S(\boldsymbol{x}_0) = \boldsymbol{x}^* = \boldsymbol{x}_0 - \boldsymbol{A}^T\left(\boldsymbol{A}\boldsymbol{A}^T\right)^{-1}\left(\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}\right)$$

(d) Let $\boldsymbol{Q} = \boldsymbol{I}$, $\boldsymbol{g} = -\boldsymbol{x}_0$, $c = \frac{1}{2}\boldsymbol{x}_0^T\boldsymbol{x}_0$, $\boldsymbol{A} = \boldsymbol{x}^T$ and we have

$$\mathrm{Proj}_P(\boldsymbol{x}_0) = \boldsymbol{x}^* = \boldsymbol{x}_0 - \boldsymbol{w}\left(\boldsymbol{w}^T\boldsymbol{w}\right)^{-1}\left(\boldsymbol{w}^T\boldsymbol{x}_0 - \boldsymbol{b}\right) = \boldsymbol{x}_0 - \frac{\boldsymbol{w}^T\boldsymbol{x}_0 - \boldsymbol{b}}{\|\boldsymbol{w}\|^2}\boldsymbol{w}$$

Therefore, the distance between $\boldsymbol{x}_0$ and P is

$$\mathrm{d}(\boldsymbol{x}_0, P) = \|\mathrm{Proj}_P(\boldsymbol{x}_0) - \boldsymbol{x}_0\| = \left\|\frac{\boldsymbol{w}^T\boldsymbol{x}_0 - \boldsymbol{b}}{\|\boldsymbol{w}\|^2}\boldsymbol{w}\right\| = \frac{\left|\boldsymbol{w}^T\boldsymbol{x}_0 - \boldsymbol{b}\right|}{\|\boldsymbol{w}\|}$$