Linear and Convex Optimization Homework 09

Qiu Yihang, 2021/11/26-11/

1.(a) ***Solution:***

The constrained quadratic program can be rewritten as the following unconstrained problem.

$$\min_{x_1} g(x_1) = \frac{1}{2}x_1^2 + \frac{1}{2}x_1(1-x_1) + \frac{1}{4}(1-x_1)^2 - x_1 - \frac{3}{2}(1-x_1) = \frac{1}{4}x_1^2 + \frac{1}{2}x_1 - \frac{5}{4}$$

Let the solution of the problem above be $x_1^\star$.

Then we have

$$\nabla g(x_1^\star) = \frac{1}{2}x_1^\star + \frac{1}{2} = 0. \Rightarrow x_1^\star = -1$$

Thus, the solution of the original quadratic program problem is

$$\boldsymbol{x}^\star = (-1,1).$$

■

(b) ***Solution:***

Lagrangian function is

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

Let the solution of the problem be $\boldsymbol{x}^\star$.

We have

$$\begin{cases} \nabla\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{x}^\star,\lambda^\star) = (x_1^\star + x_2^\star - 1 + \lambda^\star, x_1^\star + 2x_2^\star - 3 + 2\lambda^\star) = \boldsymbol{0} \\ \nabla\mathcal{L}_\lambda(\boldsymbol{x}^\star,\lambda^\star) = x_1^\star + 2x_2^\star - 1 = 0 \end{cases} \Rightarrow \begin{cases} \lambda^\star = 1 \\ x_1^\star = -1 \\ x_2^\star = 1 \end{cases}$$

Thus, the solution of the original quadratic program problem is

$$\boldsymbol{x}^\star = (-1,1).$$

The corresponding Lagrangian multiplier $\lambda^\star$ is 1. ■

2.(a) ***Solution:***

Lagrangian function is

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

Thus, the Lagrange condition for this problem is

$$\begin{cases} \nabla\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{x},\boldsymbol{\lambda}) = \boldsymbol{0} \\ \nabla\mathcal{L}_\lambda(\boldsymbol{x},\boldsymbol{\lambda}) = \boldsymbol{0} \end{cases} \Rightarrow \begin{cases} \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{g} + \boldsymbol{A}^T\boldsymbol{\lambda} = \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0} \end{cases}$$

■

(b) ***Solution:***

Since $\boldsymbol{Q} \succ \boldsymbol{0}$, $\boldsymbol{Q}^{-1}$ exists.

$$\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{g} + \boldsymbol{A}^T\boldsymbol{\lambda}^\star = \boldsymbol{0} \Rightarrow -\boldsymbol{Q}^{-1}\boldsymbol{A}^T\boldsymbol{\lambda}^\star = \boldsymbol{x}^\star + \boldsymbol{Q}^{-1}\boldsymbol{g} \quad (\star)$$

$$\Rightarrow -\boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{A}^T\boldsymbol{\lambda}^\star = \boldsymbol{A}\boldsymbol{x}^\star + \boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{g} = \boldsymbol{b} + \boldsymbol{A}\boldsymbol{Q}^{-1}\boldsymbol{g}$$

Now we prove $\boldsymbol{AQ}^{-1}\boldsymbol{A}^T$ is invertible.

$$\boldsymbol{Q} \succ \boldsymbol{0} \Rightarrow \boldsymbol{Q}^{-1} \succ \boldsymbol{0} \Rightarrow \forall \boldsymbol{z} \in \mathbb{R}^{\boldsymbol{n}}, \boldsymbol{z}^T\boldsymbol{Q}^{-1}\boldsymbol{z} > 0$$

$$\Rightarrow \forall \boldsymbol{y} \in \mathbb{R}^{\boldsymbol{n}}, (\boldsymbol{A}^T\boldsymbol{y})^T\boldsymbol{Q}^{-1}(\boldsymbol{A}^T\boldsymbol{y}) > 0 \quad \text{(Since } \operatorname{rank}\boldsymbol{A}^T = \operatorname{rank}\boldsymbol{A} = k\text{)}$$

$$\Rightarrow \forall \boldsymbol{y} \in \mathbb{R}^{\boldsymbol{n}}, \boldsymbol{y}^T\boldsymbol{AQ}^{-1}\boldsymbol{A}^T\boldsymbol{y} > 0$$

$$\Rightarrow \boldsymbol{AQ}^{-1}\boldsymbol{A}^T \text{ is invertible.}$$

Thus,

$$\boldsymbol{\lambda}^\star = -(\boldsymbol{AQ}^{-1}\boldsymbol{A}^T)^{-1}(\boldsymbol{b} + \boldsymbol{AQ}^{-1}\boldsymbol{g})$$

Plug $\boldsymbol{\lambda}^\star$ into $(\star)$, we have

$$\boldsymbol{x}^\star = \boldsymbol{Q}^{-1}\boldsymbol{A}^T(\boldsymbol{AQ}^{-1}\boldsymbol{A}^T)^{-1}(\boldsymbol{b} + \boldsymbol{AQ}^{-1}\boldsymbol{g}) - \boldsymbol{Q}^{-1}\boldsymbol{g}$$

■

(c) ***Solution:***

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}_0\|_2^2 \text{ s.t. } \boldsymbol{Ax} = \boldsymbol{b} \Leftrightarrow \min_{\boldsymbol{x}} \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)^T(\boldsymbol{x} - \boldsymbol{x}_0) \text{ s.t. } \boldsymbol{Ax} = \boldsymbol{b}$$

$$\Leftrightarrow \min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - \boldsymbol{x}_0{}^T\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}_0{}^T\boldsymbol{x}_0 \quad s.t.\ \boldsymbol{Ax} = \boldsymbol{b}$$

From (b) we know

$$\boldsymbol{Q} = \boldsymbol{I}, \boldsymbol{g} = -\boldsymbol{x}_0, c = \frac{1}{2}\boldsymbol{x}_0{}^T\boldsymbol{x}_0,$$

$$\boldsymbol{\lambda}^\star = -(\boldsymbol{AA}^T)^{-1}(\boldsymbol{b} + \boldsymbol{Ax}_0),\ \boldsymbol{x}^\star = \boldsymbol{A}^T(\boldsymbol{AA}^T)^{-1}(\boldsymbol{b} - \boldsymbol{Ax}_0) + \boldsymbol{x}_0$$ ■

When $\boldsymbol{x}_0 = \boldsymbol{0}$, we have $\boldsymbol{\lambda}^\star = -(\boldsymbol{AA}^T)^{-1}\boldsymbol{b}$, $\boldsymbol{x}^\star = \boldsymbol{A}^T(\boldsymbol{AA}^T)^{-1}\boldsymbol{b}$. ■

(d) ***Solution:***

$$\operatorname{dist}(\boldsymbol{x_0}, P) = \min_{\boldsymbol{x}}\|\boldsymbol{x} - \boldsymbol{x}_0\| \text{ s.t. } \boldsymbol{w}^T\boldsymbol{x} = \boldsymbol{b}$$

i.e. $\operatorname{dist}(\boldsymbol{x_0}, P) = \|\boldsymbol{x}^\star - \boldsymbol{x}_0\|$, where $\boldsymbol{x}^\star = \operatorname{argmin}_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}_0\|_2^2 \text{ s.t. } \boldsymbol{w}^T\boldsymbol{x} = \boldsymbol{b}$

From (c) we know

$$\boldsymbol{x}^\star = \boldsymbol{w}(\boldsymbol{w}^T\boldsymbol{w})^{-1}(\boldsymbol{b} - \boldsymbol{w}^T\boldsymbol{x}_0) + \boldsymbol{x}_0 = \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|^2}(\boldsymbol{b} - \boldsymbol{w}^T\boldsymbol{x}_0) + \boldsymbol{x}_0$$

$$\|\boldsymbol{x}^\star - \boldsymbol{x}_0\| = \left\|\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|^2}(\boldsymbol{b} - \boldsymbol{w}^T\boldsymbol{x}_0)\right\| = \frac{\|\boldsymbol{w}^T\boldsymbol{x}_0 - \boldsymbol{b}\|}{\|\boldsymbol{w}\|}$$

i.e.

$$\operatorname{dist}(\boldsymbol{x_0}, P) = \frac{\|\boldsymbol{w}^T\boldsymbol{x}_0 - \boldsymbol{b}\|}{\|\boldsymbol{w}\|}.$$

■

3. ***Solution:***

The Lagrangian function is

$$\mathcal{L}(\boldsymbol{x}, \lambda) = x_1x_2 + \lambda(x_1^2 + 4x_2^2 - 1).$$

Let the solution be $\boldsymbol{x}^\star$, we have

$$\begin{cases}\nabla\mathcal{L}_{x_1} &= 0\\ \nabla\mathcal{L}_{x_2} &= 0\\ \nabla\mathcal{L}_{\lambda} &= 0\end{cases} \Rightarrow \begin{cases} x_2 + 2\lambda x_1 = 0\\ x_1 + 8\lambda x_2 = 0\\ x_1^2 + 4x_2^2 - 1 = 0\end{cases}$$

$$\Rightarrow (1)\begin{cases}\lambda = \frac{1}{4}\\ x_1 = \frac{1}{2}\sqrt{2}\\ x_2 = -\frac{1}{4}\sqrt{2}\end{cases} \quad (2)\begin{cases}\lambda = -\frac{1}{4}\\ x_1 = \frac{1}{2}\sqrt{2}\\ x_2 = \frac{1}{4}\sqrt{2}\end{cases} \quad (3)\begin{cases}\lambda = \frac{1}{4}\\ x_1 = -\frac{1}{2}\sqrt{2}\\ x_2 = \frac{1}{4}\sqrt{2}\end{cases} \quad (4)\begin{cases}\lambda = -\frac{1}{4}\\ x_1 = -\frac{1}{2}\sqrt{2}\\ x_2 = -\frac{1}{4}\sqrt{2}\end{cases}$$

Considering $(1) x_1x_2 = -1/4, (2) x_1x_2 = 1/4, (3) x_1x_2 = -1/4, (4) x_1x_2 = 1/4,$

we know (1) and (4) are global maximum while (2) and (3) are global minimum.

■

4.(a) ***Solution:***

The Lagrangian function is

$$\mathcal{L}(\boldsymbol{x}, \lambda) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \lambda(\|\boldsymbol{x}\|_2^2 - 1) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \lambda(\boldsymbol{x}^T\boldsymbol{x} - 1).$$

For solution $\boldsymbol{x}^\star$ and its corresponding Lagrangian multiplier $\lambda^\star$, we have

$$\begin{cases}\nabla\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{x}^\star, \lambda^\star) = \boldsymbol{0}\\ \nabla\mathcal{L}_{\lambda}(\boldsymbol{x}^\star, \lambda^\star) = 0\end{cases} \Rightarrow \begin{cases}\boldsymbol{A}\boldsymbol{x}^\star + \lambda^\star\boldsymbol{x}^\star = \boldsymbol{0}\\ (\boldsymbol{x}^\star)^T\boldsymbol{x}^\star - 1 = 0\end{cases}$$

Since $\boldsymbol{A}\boldsymbol{x}^\star = -\lambda^\star\boldsymbol{x}^\star$, we know $\boldsymbol{x}^\star$ is an eigenvector of $\boldsymbol{A}$ associated to $-\lambda^\star$.

Let $-\lambda^\star = \lambda_i$, which is an eigenvalue of $\boldsymbol{A}$.

Plug λ_i into the original function, we have

$$(\boldsymbol{x}^\star)^T\boldsymbol{A}\boldsymbol{x}^\star = (\boldsymbol{x}^\star)^T\lambda_i\boldsymbol{x}^\star = \lambda_i(\boldsymbol{x}^\star)^T\boldsymbol{x}^\star = \lambda_i \cdot 1 = \lambda_i. (i \in \{1,2,\dots,n\})$$

Since $\boldsymbol{x}^\star$ is the solution of Problem (1) and $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$, the optimal value must be λ_1.

Thus, $\boldsymbol{x}^\star$ is the eigenvector of $\boldsymbol{A}$ associated to λ_1.

In conclusion, the solution $\boldsymbol{x}^\star$ to Problem (1) is the eigenvector of $\boldsymbol{A}$ associated to λ_1 and the optimal value is λ_1. ■

(b) i) ***Proof:***

The Lagrangian function is

$$\begin{aligned}\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \mathcal{L}(\boldsymbol{x}, \lambda_{(1)}, \lambda_{(2)}) &= \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \lambda_{(1)}(\|\boldsymbol{x}\|_2^2 - 1) + \lambda_{(2)}\boldsymbol{v}_1^T\boldsymbol{x}\\ &= \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \lambda_{(1)}(\boldsymbol{x}^T\boldsymbol{x} - 1) + \lambda_{(2)}\boldsymbol{v}_1^T\boldsymbol{x}.\end{aligned}$$

For solution $\boldsymbol{x}^\star$ and its corresponding Lagrangian multiplier $\boldsymbol{\lambda}^\star = (\lambda_1^\star, \lambda_2^\star)$, we have

$$\begin{cases}\nabla\mathcal{L}_{\boldsymbol{x}}(\boldsymbol{x}^\star, \boldsymbol{\lambda}^\star) = \boldsymbol{0}\\ \nabla\mathcal{L}_{\boldsymbol{\lambda}}(\boldsymbol{x}^\star, \boldsymbol{\lambda}^\star) = \boldsymbol{0}\end{cases} \Rightarrow \begin{cases}\boldsymbol{A}\boldsymbol{x}^\star + \lambda_1^\star\boldsymbol{x}^\star + \lambda_2^\star\boldsymbol{v}_1 = \boldsymbol{0}\\ (\boldsymbol{x}^\star)^T\boldsymbol{x}^\star - 1 = 0\\ \boldsymbol{v}_1^T\boldsymbol{x}^\star = \boldsymbol{0}\end{cases}$$

Thus, exist $c_0 = -\lambda_1^\star, c_1 = -\lambda_2^\star$ s.t.

$$\boldsymbol{A}\boldsymbol{x}^\star = -\lambda_1^\star\boldsymbol{x}^\star - \lambda_2^\star\boldsymbol{v}_1 = c_0\boldsymbol{x}^\star + c_1\boldsymbol{v}_1.$$

■

ii) ***Proof:***

From (b)(i), we have

$$\boldsymbol{v}_1^T\boldsymbol{A}\boldsymbol{x}^\star = c_0\boldsymbol{v}_1^T\boldsymbol{x}^\star + c_1\boldsymbol{v}_1^T\boldsymbol{v}_1 = c_1\boldsymbol{v}_1^T\boldsymbol{v}_1.$$

Meanwhile, since $\boldsymbol{A}$ is symmetric, we have

$$\boldsymbol{v}_1^T\boldsymbol{A} = \boldsymbol{v}_1^T\boldsymbol{A}^T = (\boldsymbol{A}\boldsymbol{v}_1)^T = (\lambda_1\boldsymbol{v}_1)^T = \lambda_1\boldsymbol{v}_1^T \Rightarrow \boldsymbol{v}_1^T\boldsymbol{A}\boldsymbol{x}^\star = \lambda_1\boldsymbol{v}_1^T\boldsymbol{x}^\star = \boldsymbol{0}.$$

Thus,

$$c_1\boldsymbol{v}_1^T\boldsymbol{v}_1 = \boldsymbol{0}.$$

Since $\boldsymbol{v}_1 \neq \boldsymbol{0}$, $\boldsymbol{v}_1^T\boldsymbol{v}_1 = \|\boldsymbol{v}_1\|_2^2 > \boldsymbol{0}$.

Therefore,

$$c_1 = 0. \qquad \blacksquare$$

iii) ***Proof:***

For the solution $\boldsymbol{x}^\star$ and its corresponding Lagrangian multiplier $\boldsymbol{\lambda}^\star = (\lambda_1^\star, \lambda_2^\star)$, from (b)(i) and (b)(ii) we have

$$\begin{cases} \boldsymbol{A}\boldsymbol{x}^\star = c_0\boldsymbol{x}^\star \\ (\boldsymbol{x}^\star)^T\boldsymbol{x}^\star = 1 \\ \boldsymbol{v}_1^T\boldsymbol{x}^\star = \boldsymbol{0} \end{cases}$$

Thus, $\boldsymbol{x}^\star$ is an eigenvector of $\boldsymbol{A}$ associated to c_0.

Now we prove for any $j \geq 2$, we can find a $\boldsymbol{v}_j$ which is an eigenvector of $\boldsymbol{A}$ associated to λ_j.

CASE 01. $\lambda_j \neq \lambda_1$.

Thus, we have

$$\begin{cases} \boldsymbol{v}_1^T\boldsymbol{A}\boldsymbol{v}_j = \boldsymbol{v}_1^T(\lambda_j\boldsymbol{v}_j) = \lambda_j\boldsymbol{v}_1^T\boldsymbol{v}_j \\ \boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{v}_1 = \boldsymbol{v}_j^T(\lambda_1\boldsymbol{v}_1) = \lambda_1\boldsymbol{v}_j^T\boldsymbol{v}_1 \end{cases} \Rightarrow (\lambda_1 - \lambda_j)\boldsymbol{v}_1^T\boldsymbol{v}_j = \boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{v}_1 - \boldsymbol{v}_1^T\boldsymbol{A}\boldsymbol{v}_j$$

Since $\boldsymbol{A}$ is symmetric and $\boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{v}_1$ is a number,

$$\begin{cases} \boldsymbol{A}^T = \boldsymbol{A} \\ \boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{v}_1 = \left(\boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{v}_1\right)^{\boldsymbol{T}} = \boldsymbol{v}_1^T\boldsymbol{A}^T\boldsymbol{v}_j \end{cases} \Rightarrow (\lambda_1 - \lambda_j)\boldsymbol{v}_1^T\boldsymbol{v}_j = \boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{v}_1 - \boldsymbol{v}_1^T\boldsymbol{A}\boldsymbol{v}_j = \boldsymbol{0}$$

Considering $\lambda_j \neq \lambda_1$,

$$\boldsymbol{v}_1^T\boldsymbol{v}_j = \boldsymbol{0}.$$

CASE 02. $\lambda_j = \lambda_1$.

Let the space containing all eigenvectors associated to λ_1 be S_λ. Then $\dim S_\lambda \geq 2$.

Thus, exists at least a $\boldsymbol{v}_j \in S_\lambda$ s.t. $\boldsymbol{v}_1^T\boldsymbol{v}_j = \boldsymbol{0}$.

Therefore, for any eigenvector $\boldsymbol{v}$ of $\boldsymbol{A}$ which is associated to λ_j $(j \geq 2)$, we can find

$$\boldsymbol{v}^\star = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \text{ s.t. } \begin{cases} \boldsymbol{A}\boldsymbol{v}^\star = \lambda_j\boldsymbol{v}^\star \\ (\boldsymbol{v}^\star)^T\boldsymbol{v}^\star = 1 \\ \boldsymbol{v}_1^T\boldsymbol{v}^\star = \boldsymbol{0} \end{cases}$$

Obviously $\boldsymbol{v}^\star \neq \boldsymbol{v}_1$. Otherwise, $\boldsymbol{v}_1^T\boldsymbol{v}_1 = \boldsymbol{0}$ i.e. $\boldsymbol{v}_1 = \boldsymbol{0}$.

Plug $\boldsymbol{v}^\star$ into the original function,

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{v}^\star)^T\boldsymbol{A}\boldsymbol{v}^\star = (\boldsymbol{v}^\star)^T\lambda_j\boldsymbol{v}^\star = \lambda_j(\boldsymbol{v}^\star)^T\boldsymbol{v}^\star = \lambda_j.$$

Considering $\boldsymbol{x}^\star$ is the solution of Problem (2), $\boldsymbol{x}^\star \neq \boldsymbol{v}_1$, and $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, the optimal value should be λ_2.

Thus, $\boldsymbol{x}^\star$ is eigenvector of $\boldsymbol{A}$ associated to λ_2.

In conclusion, the solution $\boldsymbol{x}^\star$ to Problem (2) is the eigenvector of $\boldsymbol{A}$ associated to λ_2 and the optimal value is λ_2. ∎