Homework Assignment 6

Matthew Tiger

April 11, 2016

Problem 1. a. Where is the assumption "$\boldsymbol{x}^*$ is regular" essential in the proof of the results of section: Lagrange Multipliers?

b. In the example on page 49 (Example 20.8 in *An Introduction to Optimization*) explain in what way is (P_0) equivalent to (P_1).

c. State the SOSC Theorem on p. 51 (Theorem 20.5 p. 474 in the book) for $\boldsymbol{x}^*$ a local maximizer.

Solution. a. The assumption that $\boldsymbol{x}^*$ is regular is essential in the proof of the Lagrange Multipliers Theorem in applying the results of Theorem 20.1, i.e. assuming that $\boldsymbol{y} \in T(\boldsymbol{x}^*)$ if and only if there exists a differentiable curve in S passing through $\boldsymbol{x}^*$ with derivative $\boldsymbol{y}$ at $\boldsymbol{x}^*$ for $\boldsymbol{x}^*$ a regular point.

b. The two problems to consider are:

$$(P_0) \quad \begin{array}{ll} \text{maximize} & \frac{\boldsymbol{x}^\mathsf{T} Q \boldsymbol{x}}{\boldsymbol{x}^\mathsf{T} P \boldsymbol{x}} \\ \text{subject to} & Q = Q^\mathsf{T} \geq 0 \\ & P = P^\mathsf{T} > 0. \end{array} \qquad (P_1) \quad \begin{array}{ll} \text{maximize} & \boldsymbol{x}^\mathsf{T} Q \boldsymbol{x} \\ \text{subject to} & \boldsymbol{x}^\mathsf{T} P \boldsymbol{x} = 1. \end{array}$$

Note that if P is positive semi-definite and Q is positive definite, then $\boldsymbol{x}^\mathsf{T} Q \boldsymbol{x} \geq 0$ and $\boldsymbol{x}^\mathsf{T} P \boldsymbol{x} > 0$ for every $\boldsymbol{x}$. Consequently

$$\frac{\boldsymbol{x}^\mathsf{T} Q \boldsymbol{x}}{\boldsymbol{x}^\mathsf{T} P \boldsymbol{x}} \geq 0$$

for every $\boldsymbol{x}$. From problem (P_0) we see that if $\boldsymbol{x}$ is a solution to the problem, then so is $t\boldsymbol{x}$ for any $t \neq 0$. Note that

$$\frac{(t\boldsymbol{x})^\mathsf{T} Q (t\boldsymbol{x})}{(t\boldsymbol{x})^\mathsf{T} P (t\boldsymbol{x})} = \frac{t^2 \boldsymbol{x}^\mathsf{T} Q \boldsymbol{x}}{t^2 \boldsymbol{x}^\mathsf{T} P \boldsymbol{x}} = \frac{\boldsymbol{x}^\mathsf{T} Q \boldsymbol{x}}{\boldsymbol{x}^\mathsf{T} P \boldsymbol{x}}$$

showing that the above remark is true. Adding the additional constraint to problem (P_0) that $\boldsymbol{x}^\mathsf{T} P \boldsymbol{x} = 1$ removes the multiplicity of the solutions and transforms the original problem into problem (P_1). To see this, if the constraint $\boldsymbol{x}^\mathsf{T} P \boldsymbol{x} = 1$ is satisfied then for any non-zero scalar multiple of $\boldsymbol{x}^*$ we have that

$$(t\boldsymbol{x})^\mathsf{T} P (t\boldsymbol{x}) = t^2 \boldsymbol{x}^\mathsf{T} P \boldsymbol{x} = \boldsymbol{x}^\mathsf{T} P \boldsymbol{x}$$

Since $\boldsymbol{x}^\mathsf{T} P \boldsymbol{x} > 0$ we must have that $t = 1$ removing the multiplicity of the solutions and the problems are equivalent.

c.

Theorem 1 (***Second-Order Sufficient Conditions***). Suppose that $f, \boldsymbol{h} \in \mathcal{C}^2$ with $f : \mathbb{R}^n \to \mathbb{R}$ and $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$. Let $l(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \lambda_1 h_1(\boldsymbol{x}) + \lambda_2 h_2(\boldsymbol{x}) + \ldots \lambda_m h_m(\boldsymbol{x})$ be the Lagrangian function. Let

$$
\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\boldsymbol{x}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\boldsymbol{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\boldsymbol{x}) \end{bmatrix}
$$

be the Hessian matrix of f at $\boldsymbol{x}$ and

$$
\boldsymbol{H_k}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 h_k}{\partial x_1^2}(\boldsymbol{x}) & \frac{\partial^2 h_k}{\partial x_2 \partial x_1}(\boldsymbol{x}) & \cdots & \frac{\partial^2 h_k}{\partial x_n \partial x_1}(\boldsymbol{x}) \\ \frac{\partial^2 h_k}{\partial x_1 \partial x_2}(\boldsymbol{x}) & \frac{\partial^2 h_k}{\partial x_2^2}(\boldsymbol{x}) & \cdots & \frac{\partial^2 h_k}{\partial x_n \partial x_2}(\boldsymbol{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 h_k}{\partial x_1 \partial x_n}(\boldsymbol{x}) & \frac{\partial^2 h_k}{\partial x_2 \partial x_n}(\boldsymbol{x}) & \cdots & \frac{\partial^2 h_k}{\partial x_n^2}(\boldsymbol{x}) \end{bmatrix}
$$

be the Hessian matrix of h_k at $\boldsymbol{x}$ for $k = 1, \ldots, m$. Define

$$
\boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{F}(\boldsymbol{x}) + \lambda_1 \boldsymbol{H_1}(\boldsymbol{x}) + \cdots + \lambda_m \boldsymbol{H_m}(\boldsymbol{x})
$$

to be the Hessian Matrix of $l(\boldsymbol{x}, \boldsymbol{\lambda})$ with respect to $\boldsymbol{x}$.

Suppose there exists a point $\boldsymbol{x}^* \in \mathbb{R}^n$ and $\boldsymbol{\lambda^*} \in \mathbb{R}^m$ such that

- $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda^*}^\mathsf{T} D\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^\mathsf{T}$.
- For all $\boldsymbol{y} \in T(\boldsymbol{x}^*)$, $\boldsymbol{y} \neq \boldsymbol{0}$, we have that $\boldsymbol{y}^\mathsf{T} \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\boldsymbol{y} < 0$, i.e. $\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ is negative definite on $T(\boldsymbol{x}^*)$.

Then $\boldsymbol{x}^*$ is a strict local maximizer of f subject to $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$.

$\square$

Problem 2. Find local extremizers for the following optimization problem:

$$\text{maximize} \quad x_1 x_2$$
$$\text{subject to} \quad x_1^2 + 4x_2^2 = 1.$$

Solution. Lagrange's Theorem prescribes how to find the possible local extremizers for the optimization problem. Let $f(\boldsymbol{x}) = x_1 x_2$ and $h(\boldsymbol{x}) = x_1^2 + 4x_2^2 - 1$. Note that we then have that

$$\nabla f(\boldsymbol{x})^\mathsf{T} = \begin{bmatrix} x_2 & x_1 \end{bmatrix} \quad \text{and} \quad \nabla h(\boldsymbol{x})^\mathsf{T} = \begin{bmatrix} 2x_1 & 8x_2 \end{bmatrix}.$$

Since for every feasible $\boldsymbol{x}$, the Jacobian of $\boldsymbol{h}$ is of rank 1, i.e. of full rank, every feasible point is a regular point. Using the Lagrange condition $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\mathsf{T}} Dh(\boldsymbol{x}^*) = \boldsymbol{0}^\mathsf{T}$, we formulate the system of equations

$$Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\mathsf{T}} Dh(\boldsymbol{x}^*) = \begin{bmatrix} x_2 + 2\lambda x_1 & x_1 + 8\lambda x_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} = \boldsymbol{0}^\mathsf{T}$$

for $\lambda \in \mathbb{R}$. Thus, an extremizer of the optimization problem must satisfy the following system of equations

$$x_2 + 2\lambda x_1 = 0$$
$$x_1 + 8\lambda x_2 = 0$$
$$x_1^2 + 4x_2^2 = 1.$$

Solving the above system leads to four distinct solutions. Thus, the possible extremizers to the optimization problem are

$$\boldsymbol{x}^{(1)} = \begin{bmatrix} 1/\sqrt{2} \\ 1/(2\sqrt{2}) \end{bmatrix}, \ \boldsymbol{x}^{(2)} = \begin{bmatrix} 1/\sqrt{2} \\ -1/(2\sqrt{2}) \end{bmatrix}, \ \boldsymbol{x}^{(3)} = \begin{bmatrix} -1/\sqrt{2} \\ 1/(2\sqrt{2}) \end{bmatrix}, \ \boldsymbol{x}^{(4)} = \begin{bmatrix} -1/\sqrt{2} \\ -1/(2\sqrt{2}) \end{bmatrix}$$

where the Lagrange multiplier for $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(4)}$ is $\lambda = -1/4$ while the Lagrange multiplier for $\boldsymbol{x}^{(2)}$ and $\boldsymbol{x}^{(3)}$ is $\lambda = 1/4$. Since $f(\boldsymbol{x}^{(1)}) = f(\boldsymbol{x}^{(4)}) = 1/4$ and $f(\boldsymbol{x}^{(2)}) = f(\boldsymbol{x}^{(3)}) = -1/4$, we have that the possible maximizers of the problem are located at $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(4)}$ while the possible minimizers of the problem are located at $\boldsymbol{x}^{(2)}$ and $\boldsymbol{x}^{(3)}$.

The Second Order Sufficient Conditions prescribe how to check if these extremizers are strict local extremizers. We have already shown that these extremizers satisfy the first condition, i.e. the Lagrange condition. We must now show that for $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(i)})$ for $i = 1, \ldots, 4$ we have that

$$\boldsymbol{y}^\mathsf{T}(D^2 f(\boldsymbol{x}^{(i)}) + \boldsymbol{\lambda}^{(i)\mathsf{T}} D^2 h(\boldsymbol{x}^{(i)}))\boldsymbol{y} = \boldsymbol{y}^\mathsf{T}\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \boldsymbol{\lambda}^{(i)\mathsf{T}} \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}\right)\boldsymbol{y} > (<)0 \qquad (1)$$

for $\boldsymbol{x}^{(i)}$ to be a strict local minimum (maximum).

Note that if $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{x})$ then

$$Dh(\boldsymbol{x})\boldsymbol{y} = \begin{bmatrix} 2x_1 & 8x_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 2(x_1 y_1 + 4x_2 y_2) = 0.$$

We wish to see if $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(4)}$ are strict local maximizers. Start with $\boldsymbol{x}^{(1)} = \left[1/\sqrt{2} \quad 1/(2\sqrt{2})\right]^{\mathsf{T}}$. The Lagrange multiplier associated to this solution is $\boldsymbol{\lambda}^{(1)} = -1/4$. Then $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(1)}) = \{\boldsymbol{y} \mid \boldsymbol{y} = [y_1 \quad -y_1/2]\}$. From this we see that (1) becomes

$$\begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix}^{\mathsf{T}} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - 1/4 \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \right) \begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix} = \begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} -1/2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix} = -2y_1^2 < 0$$

showing indeed that $\boldsymbol{x}^{(1)}$ is a strict local maximizer.

Now let's check $\boldsymbol{x}^{(4)} = \left[-1/\sqrt{2} \quad -1/(2\sqrt{2})\right]^{\mathsf{T}}$. The Lagrange multiplier associated to this solution is $\boldsymbol{\lambda}^{(4)} = -1/4$. Then $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(4)}) = \{\boldsymbol{y} \mid \boldsymbol{y} = [y_1 \quad -y_1/2]\}$. From this we see that (1) becomes

$$\begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix}^{\mathsf{T}} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - 1/4 \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \right) \begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix} = \begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} -1/2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ -y_1/2 \end{bmatrix} = -2y_1^2 < 0$$

showing indeed that $\boldsymbol{x}^{(4)}$ is a strict local maximizer.

We wish to now see if $\boldsymbol{x}^{(2)}$ and $\boldsymbol{x}^{(3)}$ are strict local minimizers. Start with $\boldsymbol{x}^{(2)} = \left[1/\sqrt{2} \quad -1/(2\sqrt{2})\right]^{\mathsf{T}}$. The Lagrange multiplier associated to this solution is $\boldsymbol{\lambda}^{(2)} = 1/4$. Then $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(2)}) = \{\boldsymbol{y} \mid \boldsymbol{y} = [y_1 \quad y_1/2]\}$. From this we see that (1) becomes

$$\begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix}^{\mathsf{T}} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 1/4 \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \right) \begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} 1/2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix} = 2y_1^2 > 0$$

showing indeed that $\boldsymbol{x}^{(2)}$ is a strict local minimizer.

Now let's check $\boldsymbol{x}^{(3)} = \left[-1/\sqrt{2} \quad 1/(2\sqrt{2})\right]^{\mathsf{T}}$. The Lagrange multiplier associated to this solution is $\boldsymbol{\lambda}^{(3)} = 1/4$. Then $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(3)}) = \{\boldsymbol{y} \mid \boldsymbol{y} = [y_1 \quad y_1/2]\}$. From this we see that (1) becomes

$$\begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix}^{\mathsf{T}} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + 1/4 \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \right) \begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} 1/2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_1/2 \end{bmatrix} = 2y_1^2 > 0$$

showing indeed that $\boldsymbol{x}^{(3)}$ is a strict local minimizer. $\qquad\qquad\square$

Problem 3. Consider the problem

$$\text{minimize} \quad 2x_1 + 3x_2 - 4, \quad x_1, x_2 \in \mathbb{R}$$
$$\text{subject to} \quad x_1 x_2 = 6.$$

a. Use Lagrange's theorem to find all possible local minimizers and maximizers.

b. Use the second-order sufficient conditions to specify which points are strict local minimizers and which are strict local maximizers.

c. Are the points in part b global minimizers or maximizers? Explain.

Solution.   a. We use Lagrange's theorem to find the possible extremizers of this optimization problem. To start, say $f(\boldsymbol{x}) = 2x_1 + 3x_2 - 4$ and $h(\boldsymbol{x}) = x_1 x_2 - 6$. Then

$$\nabla f(\boldsymbol{x})^\mathsf{T} = \begin{bmatrix} 2 & 3 \end{bmatrix} \quad \text{and} \quad \nabla \boldsymbol{h}(\boldsymbol{x})^\mathsf{T} = \begin{bmatrix} x_2 & x_1 \end{bmatrix}.$$

Since the Jacobian of $\boldsymbol{h}$ is of full rank only if $x_1 \neq 0$ and $x_2 \neq 0$, we have that for every feasible point the Jacobian of $\boldsymbol{h}$ is of full rank, i.e. every feasible point is regular. The Lagrange condition states that the possible extremizers $\boldsymbol{x}^*$ satisfy $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\mathsf{T}} D\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^\mathsf{T}$ or

$$\begin{bmatrix} 2 + \lambda x_2 & 3 + \lambda x_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

The Lagrange condition in conjunction with the condition that $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$ yields the following system of equations:

$$2 + \lambda x_2 = 0$$
$$3 + \lambda x_1 = 0$$
$$x_1 x_2 = 6.$$

From the first two equations, we see that $x_1 = -3/\lambda$ and $x_2 = -2/\lambda$. Substituting into the third equation yields that $6 = 6/\lambda^2$ or that $\lambda = \pm 1$. Thus, the possible extremizers of the optimization problem are

$$\boldsymbol{x}^{(1)} = \begin{bmatrix} -3 \\ -2 \end{bmatrix}, \ \boldsymbol{x}^{(2)} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}. \tag{2}$$

b. The second-order sufficient conditions allow us to specify whether the possible extremizers (2) are strict local minimizers or maximizers. We have already shown that these solutions satisfy the first condition, i.e. these solutions satisfy the Lagrange condition. We need to show that for the solutions $\boldsymbol{x}^{(i)}$, $i = 1, 2$, in (2), it is true that for all $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(i)})$, $\boldsymbol{y} \neq \boldsymbol{0}$, we have

$$\boldsymbol{y}^\mathsf{T} \left(\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}^{(i)}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}^{(i)}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}^{(i)}) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}^{(i)}) \end{bmatrix} + \lambda \begin{bmatrix} \frac{\partial^2 h}{\partial x_1^2}(\boldsymbol{x}^{(i)}) & \frac{\partial^2 h}{\partial x_2 \partial x_1}(\boldsymbol{x}^{(i)}) \\ \frac{\partial^2 h}{\partial x_1 \partial x_2}(\boldsymbol{x}^{(i)}) & \frac{\partial^2 h}{\partial x_2^2}(\boldsymbol{x}^{(i)}) \end{bmatrix} \right) \boldsymbol{y} > (<) \, 0$$

to show that $\boldsymbol{x}^{(i)}$ is a strict local minimizer (maximizer).

Take the solution $\boldsymbol{x}^{(1)} = \begin{bmatrix} -3 & -2 \end{bmatrix}^\mathsf{T}$. The Lagrange multiplier associated to this solution was $\lambda = 1$. We must now check that for any $\boldsymbol{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^\mathsf{T} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(1)})$, we have that

$$\boldsymbol{y}^\mathsf{T} \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + 1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) \boldsymbol{y} = \boldsymbol{y}^\mathsf{T} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{y} = 2y_1 y_2 > (<) \, 0$$

for $\boldsymbol{x}^{(1)}$ to be a strict local minimizer (maximizer). Note that $\boldsymbol{y} \in T(\boldsymbol{x}^{(1)})$ if

$$D\boldsymbol{h}(\boldsymbol{x}^{(1)})\boldsymbol{y} = \triangledown \boldsymbol{h}(\boldsymbol{x}^{(1)})^\mathsf{T}\boldsymbol{y} = \begin{bmatrix} -2 & -3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

From this equation we see that for $\boldsymbol{y} \in T(\boldsymbol{x}^{(1)})$, we have that $y_2 = (-2/3)y_1$ and that $2y_1 y_2 = (-4/3)y_1^2 < 0$. Thus, this solution is a strict local maximizer.

Now take the solution $\boldsymbol{x}^{(2)} = \begin{bmatrix} 3 & 2 \end{bmatrix}^\mathsf{T}$. The Lagrange multiplier associated to this solution was $\lambda = -1$. We must now check that for any $\boldsymbol{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^\mathsf{T} \neq \boldsymbol{0} \in T(\boldsymbol{x}^{(2)})$, we have that

$$\boldsymbol{y}^\mathsf{T} \left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} - 1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) \boldsymbol{y} = \boldsymbol{y}^\mathsf{T} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \boldsymbol{y} = -2y_1 y_2 > (<) \, 0.$$

for $\boldsymbol{x}^{(1)}$ to be a strict local minimizer (maximizer). Note that $\boldsymbol{y} \in T(\boldsymbol{x}^{(2)})$ if

$$D\boldsymbol{h}(\boldsymbol{x}^{(2)})\boldsymbol{y} = \triangledown \boldsymbol{h}(\boldsymbol{x}^{(2)})^\mathsf{T}\boldsymbol{y} = \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

From this equation we see that for $\boldsymbol{y} \in T(\boldsymbol{x}^{(2)})$, we have that $y_2 = (-2/3)y_1$ and that $-2y_1 y_2 = (4/3)y_1^2 > 0$. Thus, this solution is a strict local minimizer.

c. The strict local minimizer $\boldsymbol{x}^{(2)}$ is not a global minimizer and the strict local maximizer $\boldsymbol{x}^{(1)}$ is not a global maximizer. To see why this is true, take note that if $\boldsymbol{x}^*$ is a global minimizer (maximizer), then $f(\boldsymbol{x}^*) \leq (\geq) f(\boldsymbol{x})$ for all $\boldsymbol{x}$ such that $h(\boldsymbol{x}) = \boldsymbol{0}$, i.e. $\boldsymbol{x}$ is feasible. Also take note that $f(\boldsymbol{x}^{(2)}) = 8$ and $f(\boldsymbol{x}^{(1)}) = -16$.

If $\boldsymbol{x}^{(2)}$ were a global minimizer, then $f(\boldsymbol{x}^{(2)}) = 8 \leq f(\boldsymbol{x})$ for all feasible $\boldsymbol{x}$. However, $\boldsymbol{x}^{(1)}$ is feasible, yet $f(\boldsymbol{x}^{(2)}) \nleq -16 = f(\boldsymbol{x}^{(1)})$ showing that $\boldsymbol{x}^{(2)}$ cannot be a global minimizer.

In a very similar fashion we argue that $\boldsymbol{x}^{(1)}$ is not a global maximizer. For if $\boldsymbol{x}^{(1)}$ were a global maximizer, then $f(\boldsymbol{x}^{(1)}) = -16 \geq f(\boldsymbol{x})$ for all feasible $\boldsymbol{x}$. However, $\boldsymbol{x}^{(2)}$ is feasible, yet $f(\boldsymbol{x}^{(1)}) \ngeq 8 = f(\boldsymbol{x}^{(2)})$ showing that $\boldsymbol{x}^{(1)}$ cannot be a global maximizer. $\qquad\square$

Problem 4. Consider the problem of minimizing a general quadratic function subject to a linear constraint:

$$\begin{aligned}\text{minimize} \quad & \tfrac{1}{2}\boldsymbol{x}^\mathsf{T}Q\boldsymbol{x} - \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + d \\ \text{subject to} \quad & A\boldsymbol{x} = \boldsymbol{b},\end{aligned}$$

where $Q = Q^\mathsf{T} > 0$, $A \in \mathbb{R}^{m \times n}$ with $m < n$, $\text{rank}A = m$ and d a constant. Derive a closed form solution to the problem.

Solution. From the above problem, we wish to minimize $f(\boldsymbol{x}) = \tfrac{1}{2}\boldsymbol{x}^\mathsf{T}Q\boldsymbol{x} - \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + d$ subject to $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{b} - A\boldsymbol{x} = \boldsymbol{0}$. We need to identify the possible minimizers of this optimization problem.

In order to proceed, we form the Lagrange function as

$$\begin{aligned}l(\boldsymbol{x}, \boldsymbol{\lambda}) &= f(\boldsymbol{x}) + \boldsymbol{\lambda}^\mathsf{T}\boldsymbol{h}(\boldsymbol{x}) \\ &= \frac{1}{2}\boldsymbol{x}^\mathsf{T}Q\boldsymbol{x} - \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + d + \boldsymbol{\lambda}^\mathsf{T}(\boldsymbol{b} - A\boldsymbol{x}).\end{aligned}$$

By Lagrange's Theorem, the possible minimizers satisfy

$$D_x l(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = D_x f(\boldsymbol{x}^*) + \boldsymbol{\lambda}^* D_x \boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^\mathsf{T}. \tag{3}$$

Note that

$$\begin{aligned}D_x\left(\frac{1}{2}\boldsymbol{x}^\mathsf{T}Q\boldsymbol{x}\right) &= \frac{1}{2}D_x\left(\sum_{i=1}^{n}\sum_{j=1}^{n}q_{ij}x_i x_j\right) \\ &= \frac{1}{2}\left[\sum_i\sum_j q_{ij}\frac{\partial}{\partial x_1}\left[x_i x_j\right], \quad \ldots, \quad \sum_i\sum_j q_{ij}\frac{\partial}{\partial x_n}\left[x_i x_j\right]\right] \\ &= \frac{1}{2}\left[\sum_i\sum_j q_{ij}\left[\frac{\partial x_i}{\partial x_1}x_j + \frac{\partial x_j}{\partial x_1}x_i\right], \quad \ldots, \quad \sum_i\sum_j q_{ij}\left[\frac{\partial x_i}{\partial x_n}x_j + \frac{\partial x_j}{\partial x_n}x_i\right]\right] \\ &= \left[x_1\sum_i\sum_j q_{ij}, \quad \ldots, \quad x_n\sum_i\sum_j q_{ij}\right] \\ &= \boldsymbol{x}^\mathsf{T}Q\end{aligned}$$

since $\frac{\partial x_i}{\partial x_k} = 0$ if $i \neq k$. Therefore, we have that the FONC (3) become

$$D_x l(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \boldsymbol{x}^{*\mathsf{T}}Q - \left(\boldsymbol{\lambda}^{*\mathsf{T}}A + \boldsymbol{c}^\mathsf{T}\right) = \boldsymbol{0}^\mathsf{T}$$

and by Lagrange's Theorem, the only possible minimizer is

$$\boldsymbol{x}^* = Q^{-1}A^\mathsf{T}\boldsymbol{\lambda}^* + Q^{-1}\boldsymbol{c}. \tag{4}$$

In order to find a closed-form solution to the problem we must identify $\boldsymbol{\lambda}^*$. Multiplying (9) by A and using the condition that $A\boldsymbol{x}^* = b$ we have that

$$\boldsymbol{b} = AQ^{-1}A^\mathsf{T}\boldsymbol{\lambda}^* + AQ^{-1}\boldsymbol{c}.$$

Since Q is positive definite and A is of full rank, $AQ^{-1}A^\mathsf{T}$ is invertible and

$$\boldsymbol{\lambda}^* = (AQ^{-1}A^\mathsf{T})^{-1}(\boldsymbol{b} - AQ^{-1}\boldsymbol{c}).$$

Therefore, from (4), the possible minimizer to the optimization problem in closed-form is

$$\boldsymbol{x}^* = Q^{-1}A^{\mathsf{T}}(AQ^{-1}A^{\mathsf{T}})^{-1}(\boldsymbol{b} - AQ^{-1}\boldsymbol{c}) + Q^{-1}\boldsymbol{c}. \tag{5}$$

In order to verify that the minimizer (5) is in fact a minimizer, we must check that the Hessian matrix of the Lagrange function is positive definite on the tangent space of $\boldsymbol{x}^*$. We readily see that the Hessian matrix of the Lagrange function is

$$D_x^2 l(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = D_x(\boldsymbol{x}^{*\mathsf{T}}Q - (\boldsymbol{\lambda}^{*\mathsf{T}}A + \boldsymbol{c}^{\mathsf{T}})) = Q.$$

Since Q is positive definite, we have that the Hessian matrix of the Lagrange function is most certainly positive definite on the tangent space of $\boldsymbol{x}^*$ showing that the minimizer (5) is indeed a minimizer by the Second Order Sufficient Conditions. $\qquad\square$

Problem 5. Consider the discrete-time linear system $x_k = 2x_{k-1} + u_k$, $k \geq 1$, with $x_0 = 1$. Find the values of the control inputs u_1 and u_2 to minimize

$$x_2^2 + \frac{1}{2}u_1^2 + \frac{1}{3}u_2^2.$$

Solution. Note that it can easily be proven that the recurrence relation $x_k = 2x_{k-1} + u_k$, $k \geq 1$ has the closed form solution

$$x_n = 2^n x_0 + \sum_{k=1}^{n} 2^{n-k} u_k \quad \text{for } n \geq 1.$$

Thus, our control inputs are subject to the constraints

$$x_1 = 2x_0 + u_1$$
$$x_1 = 4x_0 + 2u_1 + u_2$$

which can be reformulated as

$$2x_0 = x_1 - u_1$$
$$4x_0 = x_2 - 2u_1 - u_2. \tag{6}$$

Forming the vector of decision variables as $\boldsymbol{z}^\mathsf{T} = \begin{bmatrix} x_1 & x_2 & z_1 & z_2 \end{bmatrix}$, we can write the constraints (6) as

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & -1 \end{bmatrix} \boldsymbol{z} = \begin{bmatrix} 2x_0 \\ 4x_0 \end{bmatrix}.$$

Using the same decision variable vector $\boldsymbol{z}$, we can write the object function as

$$x_2^2 + \frac{1}{2}u_1^2 + \frac{1}{3}u_2^2 = \frac{1}{2}\boldsymbol{z}^\mathsf{T} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{2}{3} \end{bmatrix} \boldsymbol{z}.$$

Combining these facts, we can write the original problem as the following quadratic programming problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\boldsymbol{z}^\mathsf{T} Q \boldsymbol{z} \\ \text{subject to} & A\boldsymbol{z} = \boldsymbol{b} \end{array} \tag{7}$$

where the objective is $f(\boldsymbol{z}) = \frac{1}{2}\boldsymbol{z}^\mathsf{T} Q \boldsymbol{z}$, the constraint is $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{b} - A\boldsymbol{z}$, and

$$\boldsymbol{z} = \begin{bmatrix} x_1 \\ x_2 \\ u_1 \\ u_2 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{2}{3} \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & -1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 2x_0 \\ 4x_0 \end{bmatrix}. \tag{8}$$

Note that for this problem the matrix Q is singular. The Lagrangian associated to this problem is

$$l(\boldsymbol{z}, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{z}^\mathsf{T} Q \boldsymbol{z} + \boldsymbol{\lambda}^\mathsf{T}(\boldsymbol{b} - A\boldsymbol{z}).$$

The Lagrange condition shows that

$$D_z l(\boldsymbol{z}, \boldsymbol{\lambda}) = \boldsymbol{z}^\mathsf{T} Q - \boldsymbol{\lambda}^\mathsf{T} A = \boldsymbol{0}^\mathsf{T} \implies D_z l(\boldsymbol{z}, \boldsymbol{\lambda}) = Q\boldsymbol{z} = A^\mathsf{T}\boldsymbol{\lambda}.$$

Using our definitions of the above matrices we see that for $\boldsymbol{\lambda}^\mathsf{T} = \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix}$ that

$$0 = \lambda_2, \quad x_2 = \frac{\lambda_2}{2}, \quad u_1 = -\lambda_1 - 2\lambda_2, \quad u_2 = -\frac{3}{2}\lambda_2$$

from which we see that the possible minimizer of the problem is

$$\boldsymbol{z}^{*\mathsf{T}} = \begin{bmatrix} x_1 & 0 & -\lambda_1 & 0 \end{bmatrix}.$$

Using the fact that $A\boldsymbol{z}^* = \boldsymbol{b}$ we see that

$$A\boldsymbol{z}^* = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \\ -\lambda_1 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + \lambda_1 \\ 2\lambda_1 \end{bmatrix} = \begin{bmatrix} 2x_0 \\ 4x_0 \end{bmatrix} = \boldsymbol{b}.$$

Thus, $\lambda_1 = 2x_0$ and $x_1 = 0$. Therefore, the possible minimizer is

$$\boldsymbol{z}^{*\mathsf{T}} = \begin{bmatrix} 0 & 0 & -2x_0 & 0 \end{bmatrix}. \tag{9}$$

In order to show that this is in fact a strict local minimizer, we must show that the Hessian matrix of the Lagrange function, $D_z^2 l(\boldsymbol{z}^*, \boldsymbol{\lambda}^*) = Q$, satisfies

$$\text{For all } \boldsymbol{y} \in T(\boldsymbol{z}^*), \, \boldsymbol{y} \neq \boldsymbol{0} \text{ we have, } \boldsymbol{y}^\mathsf{T} Q \boldsymbol{y} > 0.$$

Note that $y \in T(\boldsymbol{z}^*)$ where $T(\boldsymbol{z}^*) = \{\boldsymbol{y} \mid D\boldsymbol{h}(\boldsymbol{z}^*)\boldsymbol{y} = -A\boldsymbol{y} = \boldsymbol{0}\}$. This translates into the null space of the matrix A. Thus, we must have that if $y \in T(\boldsymbol{z}^*)$, then

$$\boldsymbol{y}^\mathsf{T} = \begin{bmatrix} y_1 & y_2 & y_1 & -2y_1 + y_2 \end{bmatrix}.$$

Using this expression we readily see that for $\boldsymbol{y} \neq \boldsymbol{0} \in T(\boldsymbol{z}^*)$,

$$\boldsymbol{y}^\mathsf{T} Q \boldsymbol{y} = y_1^2 + 2y_2^2 + \frac{2}{3}(-2y_1 + y_2)^2 > 0$$

since $y_1 \neq 0$, $y_2 \neq 0$, and $-2y_1 + y_2 \neq 0$ and the SOSCs are satisfied. Thus, the minimizer (9) is in fact a strict local minimizer and the values of the control input to minimize

$$x_2^2 + \frac{1}{2}u_1^2 + \frac{1}{3}u_2^2$$

subject to $x_k = 2x_{k-1} + u_k$, $k \geq 1$, with $x_0 = 1$ are

$$u_1 = -2x_0 = -2 \quad \text{and} \quad u_2 = 0.$$

$\square$