Homework Assignment 6

Matthew Tiger

March 28, 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 $\boldsymbol{x}^*$.

b. The two problems to consider are:

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

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

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

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

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

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

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

Since $\boldsymbol{x}^\top P\boldsymbol{x} > 0$ 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

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

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

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

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

- $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^\top$.
- For all $\boldsymbol{y} \in T(\boldsymbol{x}^*)$, $\boldsymbol{y} \neq \boldsymbol{0}$, we have that $\boldsymbol{y}^\top \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)\boldsymbol{y} < 0$, 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:

$$\begin{array}{rl} \text{maximize} & x_1x_2 \\ \text{subject to} & x_1^2 + 4x_2^2 = 1. \end{array}$$

Solution. Lagrange's Theorem prescribes how to find the local extremizers for the optimization problem. Let $f(\boldsymbol{x}) = x_1x_2$ 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 $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\mathsf{T}} D\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}^{\mathsf{T}}$, we formulate the system of equations

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

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

$$\begin{aligned} x_2 + 2\lambda x_1 &= 0 \\ x_1 + 8\lambda x_2 &= 0 \\ x_1^2 + 4x_2^2 &= 1. \end{aligned}$$

From the first two equations, we see that $x_2 = -2\lambda x_1$ and $x_1 = -8\lambda x_2$. These equations in conjunction show that either $x_1 = 0$, $x_2 = 0$, or $\lambda = \pm 1/4$. If $\boldsymbol{x}$ is a feasible point, then we can't have that $x_1 = 0$ nor $x_2 = 0$. Thus, we must have that $\lambda = \pm 1/4$. In that case, from the first equation, we have that $x_2 = \mp x_1/2$. Substituting this into the third equation yields that

$$x_1^2 + 4(x_1^2/4) = 2x_1^2 = 1$$

which implies that $x_1 = \pm 1/\sqrt{2}$. Thus, $x_1 = \pm 1/\sqrt{2}$ and $x_2 = \mp 1/(2\sqrt{2})$ are the only solutions that satisfy the above system. That is, the extremizers to the optimization problem are

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

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 $\boldsymbol{x}^{(3)}$. Inspecting graphically, we can see that these are the maximizers and minimizers of the optimization problem. $\square$

Problem 3. Consider the problem

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

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 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

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

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 $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^\top$ 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:

$$\begin{aligned} 2 + \lambda x_2 &= 0 \\ 3 + \lambda x_1 &= 0 \\ x_1 x_2 &= 6. \end{aligned}$$

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

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

Since $f(\boldsymbol{x}^{(1)}) = -16$ and $f(\boldsymbol{x}^{(2)}) = 8$, we have that $\boldsymbol{x}^{(1)}$ is a possible local minimizer and $\boldsymbol{x}^{(2)}$ is a possible local maximizer.

b. The second-order sufficient conditions allow us to specify whether the possible local minimizers and maximizers (1) 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 (1), it is true that for all $\boldsymbol{y} \in T(\boldsymbol{x}^{(i)})$, $\boldsymbol{y} \neq \boldsymbol{0}$, we have

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

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

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

$$\boldsymbol{y}^\top \left(\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} + 1 \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \right) \boldsymbol{y} = \boldsymbol{y}^\top \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \boldsymbol{y} = 2y_1y_2 > 0.$$

Note that $\boldsymbol{y} \in T(\boldsymbol{x}^{(1)})$ if

$$D\boldsymbol{h}(\boldsymbol{x}^{(1)})\boldsymbol{y} = \nabla \boldsymbol{h}(\boldsymbol{x}^{(1)})^\top \boldsymbol{y} = \begin{bmatrix}-2 - 3\end{bmatrix} \begin{bmatrix}y_1\\ y_2\end{bmatrix} = 0$$

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 $2y_1y_2 = (-4/3)y_1^2 < 0$. Thus, this solution is not a strict local minimizer.

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

$$\boldsymbol{y}^\top \left(\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} - 1 \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \right) \boldsymbol{y} = \boldsymbol{y}^\top \begin{bmatrix}0 & -1\\ -1 & 0\end{bmatrix} \boldsymbol{y} = -2y_1y_2 < 0.$$

Note that $\boldsymbol{y} \in T(\boldsymbol{x}^{(2)})$ if

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

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 $2y_1y_2 = (-4/3)y_1^2 < 0$. Thus, this solution a strict local maximizer.

c.

$\square$

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

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

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

Solution. □

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. □