ECE/MATH 520, Spring 2008

Homework Problems 7

Solutions (version: May 5, 2008, 10:36)

19.6b,c Consider the problem

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

a. Use the Lagrange multiplier 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.

Ans.: b. We have $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{O}$, and

$$\boldsymbol{H}(\boldsymbol{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

First, consider the point $\boldsymbol{x}^{(1)} = [3, 2]^T$, with corresponding Lagrange multiplier $\lambda^{(1)} = -1$. We have

$$\boldsymbol{L}(\boldsymbol{x}^{(1)}, \lambda^{(1)}) = -\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and

$$T(\boldsymbol{x}^{(1)}) = \{\boldsymbol{y} : [2, 3]\boldsymbol{y} = \boldsymbol{0}\} = \{\alpha[-3, 2]^T : \alpha \in \mathbb{R}\}.$$

Let $\boldsymbol{y} = \alpha[-3, 2]^T \in T(\boldsymbol{x}^{(1)})$, $\alpha \neq 0$. We have

$$\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^{(1)}, \lambda^{(1)})\boldsymbol{y} = 12\alpha^2 > 0.$$

Therefore, by the SOSC, $\boldsymbol{x}^{(1)} = [3, 2]^T$ is a strict local minimizer.

Next, consider the point $\boldsymbol{x}^{(2)} = -[3, 2]^T$, with corresponding Lagrange multiplier $\lambda^{(2)} = 1$. We have

$$\boldsymbol{L}(\boldsymbol{x}^{(2)}, \lambda^{(2)}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

and

$$T(\boldsymbol{x}^{(2)}) = \{\boldsymbol{y} : -[2, 3]\boldsymbol{y} = \boldsymbol{0}\} = \{\alpha[-3, 2]^T : \alpha \in \mathbb{R}\} = T(\boldsymbol{x}^{(1)}).$$

Let $\boldsymbol{y} = \alpha[-3, 2]^T \in T(\boldsymbol{x}^{(2)})$, $\alpha \neq 0$. We have

$$\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^{(2)}, \lambda^{(2)})\boldsymbol{y} = -12\alpha^2 < 0.$$

Therefore, by the SOSC, $\boldsymbol{x}^{(2)} = -[3, 2]^T$ is a strict local maximizer.

c. Note that $f(\boldsymbol{x}^{(1)}) = 8$, while $f(\boldsymbol{x}^{(2)}) = -16$. Therefore, $\boldsymbol{x}^{(1)}$, although a strict local minimizer, is not a global minimizer. Likewise, $\boldsymbol{x}^{(2)}$, although a strict local maximizer, is not a global maximizer.

19.11b,c Consider the problem:

$$\begin{array}{ll} \text{minimize} & x_1 x_2 - 2x_1, \qquad x_1, x_2 \in \mathbb{R} \\ \text{subject to} & x_1^2 - x_2^2 = 0. \end{array}$$

a. Apply the Lagrange multiplier theorem directly to the problem to show that if a solution exists, it must be either $[1, 1]^T$ or $[-1, 1]^T$.

b. Use the second-order necessary conditions to show that $[-1, 1]^T$ cannot possibly be the solution.

c. Use the second-order sufficient conditions to show that $[1, 1]^T$ is a strict local minimizer.

Ans.: b. Consider the point $\boldsymbol{x}^* = [-1, 1]^T$. The corresponding Lagrange multiplier is $\lambda^* = -1/2$. The Hessian of the Lagrangian is

$$\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.$$

The tangent plane is given by

$$T(\boldsymbol{x}^*) = \{\boldsymbol{y} : [-2, -2]\boldsymbol{y} = 0\} = \{[a, -a]^T : a \in \mathbb{R}\}.$$

Let $\boldsymbol{y} \in T(\boldsymbol{x}^*)$, $\boldsymbol{y} \neq \boldsymbol{0}$. Then, $\boldsymbol{y} = [a, -a]^T$ for some $a \neq 0$. We have $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \lambda^*)\boldsymbol{y} = -2a^2 < 0$. Hence, SONC does not hold in this case, and therefore $\boldsymbol{x}^* = [-1, 1]^T$ cannot be local minimizer. In fact, the point is a strict local maximizer.

c. Consider the point $\boldsymbol{x}^* = [1, 1]^T$. The corresponding Lagrange multiplier is $\lambda^* = 1/2$. The Hessian of the Lagrangian is

$$\boldsymbol{L}(\boldsymbol{x}^*, \lambda^*) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

The tangent plane is given by

$$T(\boldsymbol{x}^*) = \{\boldsymbol{y} : [2, -2]\boldsymbol{y} = 0\} = \{[a, a]^T : a \in \mathbb{R}\}.$$

Let $\boldsymbol{y} \in T(\boldsymbol{x}^*)$, $\boldsymbol{y} \neq \boldsymbol{0}$. Then, $\boldsymbol{y} = [a, a]$ for some $a \neq 0$. We have $\boldsymbol{y}^T \boldsymbol{L}(\boldsymbol{x}^*, \lambda^*)\boldsymbol{y} = 2a^2 > 0$. Hence, by the SOSC, the point $\boldsymbol{x}^* = [1, 1]^T$ is a strict local minimizer.

19.15b Consider the sequence $\{x_k\}$, $x_k \in \mathbb{R}$, generated by the recursion

$$x_{k+1} = ax_k + bu_k, \quad k \geq 0 \qquad (a, b \in \mathbb{R},\ a, b \neq 0)$$

where $u_0, u_1, u_2, \ldots$ is a sequence of "control inputs," and the initial condition $x_0 \neq 0$ is given. We wish to find values of control inputs u_0 and u_1 such that $x_2 = 0$, and the average input energy $(u_0^2 + u_1^2)/2$ is minimized. Denote the optimal inputs by u_0^* and u_1^*.

a. Find expressions for u_0^* and u_1^* in terms of a, b, and x_0.

b. Use the second-order sufficient conditions to show that the point $\boldsymbol{u}^* = [u_0^*, u_1^*]^T$ in part a is a strict local minimizer.

Ans.: b. The Hessians of f and h are $\boldsymbol{F}(\boldsymbol{u}) = \boldsymbol{I}_2$ (2×2 identity matrix) and $\boldsymbol{H}(\boldsymbol{u}) = \boldsymbol{O}$, respectively. Hence, the Hessian of the Lagrangian is $\boldsymbol{L}(\boldsymbol{u}^*, \lambda^*) = \boldsymbol{I}_2$, which is positive definite. Therefore, $\boldsymbol{u}^*$ satisfies the SOSC, and is therefore a strict local minimizer.

20.9 Consider a square room, with corners located at $[0, 0]^T$, $[0, 2]^T$, $[2, 0]^T$, and $[2, 2]^T$ (in $\mathbb{R}^2$). We wish to find the point in the room that is closest to the point $[3, 4]^T$.

a. Guess which point in the box is the closest point in the room to the point $[3, 4]^T$.

b. Use the second-order sufficient conditions to prove that the point you have guessed is a strict local minimizer.

Hint: Minimizing the distance is the same as minimizing the square distance.

Ans.: a. By inspection, we guess the point $[2, 2]^T$ (drawing a picture may help).

b. We write $f(\boldsymbol{x}) = (x_1 - 3)^2 + (x_2 - 4)^2$, $g_1(\boldsymbol{x}) = -x_1$, $g_2(\boldsymbol{x}) = -x_2$, $g_3(\boldsymbol{x}) = x_1 - 2$, $g_4(\boldsymbol{x}) = x_2 - 2$, $\boldsymbol{g} = [g_1, g_2, g_3, g_4]^T$. The problem becomes

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}. \end{array}$$

We now check the SOSC for the point $\boldsymbol{x}^* = [2, 2]^T$. We have two active constraints: g_3, g_4. Regularity holds, since $\nabla g_3(\boldsymbol{x}^*) = [1, 0]^T$ and $\nabla g_4(\boldsymbol{x}^*) = [0, 1]^T$. We have $\nabla f(\boldsymbol{x}^*) = [-2, -4]^T$. We need to find a $\boldsymbol{\mu}^* \in \mathbb{R}^4$, $\boldsymbol{\mu}^* \geq \boldsymbol{0}$, satisfying FONC. From the condition

$\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$, we deduce that $\mu_1^* = \mu_2^* = 0$. Hence, $Df(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}^T$ if and only if $\boldsymbol{\mu}^* = [0, 0, 2, 4]^T$. Now,

$$\boldsymbol{F}(\boldsymbol{x}^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \qquad [\boldsymbol{\mu}^* \boldsymbol{G}(\boldsymbol{x}^*)] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence

$$\boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\mu}^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

which is positive definite on $\mathbb{R}^2$. Hence, SOSC is satisfied, and $\boldsymbol{x}^*$ is a strict local minimizer.

20.10 Consider the *quadratic programming* problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \end{array}$$

where $\boldsymbol{Q} = \boldsymbol{Q}^T > 0$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, and $\boldsymbol{b} \geq \boldsymbol{0}$. Find all points satisfying the KKT condition.

Ans.: The KKT condition is

$$\begin{aligned} \boldsymbol{x}^T \boldsymbol{Q} + \boldsymbol{\mu}^T \boldsymbol{A} &= \boldsymbol{0}^T \\ \boldsymbol{\mu}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) &= 0 \\ \boldsymbol{\mu} &\geq \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} &\leq \boldsymbol{0}. \end{aligned}$$

Postmultiplying the first equation by $\boldsymbol{x}$ gives

$$\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{\mu}^T \boldsymbol{A} \boldsymbol{x} = 0.$$

We note from the second equation that $\boldsymbol{\mu}^T \boldsymbol{A}\boldsymbol{x} = \boldsymbol{\mu}^T \boldsymbol{b}$. Hence,

$$\boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{\mu}^T \boldsymbol{b} = 0.$$

Since $\boldsymbol{Q} > 0$, the first term is nonnegative. Also, the second term is nonnegative because $\boldsymbol{\mu} \geq \boldsymbol{0}$ and $\boldsymbol{b} \geq \boldsymbol{0}$. Hence, we conclude that both terms must be zero. Because $\boldsymbol{Q} > 0$, we must have $\boldsymbol{x} = \boldsymbol{0}$.

Aside: Actually, we can deduce that the only solution to the KKT condition must be $\boldsymbol{0}$, as follows. The problem is convex; thus, the only points satisfying the KKT condition are global minimizers. However, we see that $\boldsymbol{0}$ is a feasible point, and is the only point for which the objective function value is 0. Further, the objective function is bounded below by 0. Hence, $\boldsymbol{0}$ is the only global minimizer.

20.14 Solve the following optimization problem using the second-order sufficient conditions:

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1^2 - x_2 - 4 \leq 0 \\ & x_2 - x_1 - 2 \leq 0. \end{aligned}$$

See Figure 21.1 for a graphic illustration of the problem.

Ans.: Write $f(\boldsymbol{x}) = x_1^2 + x_2^2$, $g_1(\boldsymbol{x}) = x_1^2 - x_2 - 4$, $g_2(\boldsymbol{x}) = x_2 - x_1 - 2$, and $\boldsymbol{g} = [g_1, g_2]^T$. We have $\nabla f(\boldsymbol{x}) = [2x_1, 2x_2]^T$, $\nabla g_1(\boldsymbol{x}) = [2x_1, -1]^T$, $\nabla g_2(\boldsymbol{x}) = [-1, 1]^T$, and $D^2 f(\boldsymbol{x}) = \text{diag}[2, 2]$. We compute

$$\nabla f(\boldsymbol{x}) + \boldsymbol{\mu}^T \nabla \boldsymbol{g}(\boldsymbol{x}) = [2x_1 + 2\mu_1 x_1 - \mu_2, 2x_2 - \mu_1 + \mu_2]^T.$$

We use the FONC to find critical points. Rewriting $\nabla f(\boldsymbol{x}) + \boldsymbol{\mu}^T \nabla \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$, we obtain

$$x_1 = \frac{\mu_2}{2 + 2\mu_1}, \qquad x_2 = \frac{\mu_1 - \mu_2}{2}.$$

We also use $\boldsymbol{\mu}^T \boldsymbol{g}(\boldsymbol{x}) = 0$ and $\boldsymbol{\mu} \geq \boldsymbol{0}$, giving

$$\mu_1(x_1^2 - x_2 - 4) = 0, \qquad \mu_2(x_2 - x_1 - 2) = 0.$$

The vector $\boldsymbol{\mu}$ has two components; therefore, we try four different cases.

Case 1: ($\mu_1 > 0$, $\mu_2 > 0$) We have

$$x_1^2 - x_2 - 4 = 0, \qquad x_2 - x_1 - 2 = 0.$$

We obtain two solutions: $\boldsymbol{x}^{(1)} = [-2, 0]^T$ and $\boldsymbol{x}^{(2)} = [3, 5]^T$. For $\boldsymbol{x}^{(1)}$, the two FONC equations give $\mu_1 = \mu_2$ and $-2(2 + 2\mu_1) = \mu_1$, which yield $\mu_1 = \mu_2 = -4/5$. This is not a legitimate solution since we require $\boldsymbol{\mu} \geq \boldsymbol{0}$. For $\boldsymbol{x}^{(2)}$, the two FONC equations give $\mu_1 - \mu_2 = 10$ and $3(2 + 2\mu_1) = \mu_2$, which yield $\boldsymbol{\mu} = [-16/5, -66/5]$. Again, this is not a legitimate solution.

Case 2: ($\mu_1 = 0$, $\mu_2 > 0$) We have

$$x_2 - x_1 - 2 = 0, \qquad x_1 = \frac{\mu_2}{2}, \qquad x_2 = -\frac{\mu_2}{2}.$$

Hence, $x_1 = -x_2$, and thus $\boldsymbol{x} = [-1, 1]$, $\mu_2 = -2$. This is not a legitimate solution since we require $\boldsymbol{\mu} \geq \boldsymbol{0}$.

Case 3: ($\mu_1 > 0$, $\mu_2 = 0$) We have

$$x_1^2 - x_2 - 4 = 0, \qquad x_1 = 0, \qquad x_2 = \frac{\mu_1}{2}.$$

Therefore, $x_2 = -4$, $\mu_1 = -8$, and again we don't have a legitimate solution.

Case 4: ($\mu_1 = 0$, $\mu_2 = 0$) We have $x_1 = x_2 = 0$, and all constraints are inactive. This is a legitimate candidate for the minimizer. We now apply the SOSC. Note that since the candidate is an interior point of the constraint set, the SOSC for the problem is equivalent to the SOSC for unconstrained optimization. The Hessian matrix $D^2 f(\boldsymbol{x}) = \text{diag}[2, 2]$ is symmetric and positive definite. Hence, by the SOSC, the point $\boldsymbol{x}^* = [0, 0]^T$ is the strict local minimizer (in fact, it is easy to see that it is a global minimizer).

21.2 Show that $f(\boldsymbol{x}) = x_1 x_2$ is a convex function on $\Omega = \{[a, ma]^T : a \in \mathbb{R}\}$, where m is any given nonnegative constant.

Ans.: Write $f(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{Q} \boldsymbol{x}$, where

$$\boldsymbol{Q} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Let $\boldsymbol{x}, \boldsymbol{y} \in \Omega$. Then, $\boldsymbol{x} = [a_1, ma_1]^T$ and $\boldsymbol{y} = [a_2, ma_2]^T$ for some $a_1, a_2 \in \mathbb{R}$. By Proposition 21.1, it is enough to show that $(\boldsymbol{y} - \boldsymbol{x})^T \boldsymbol{Q} (\boldsymbol{y} - \boldsymbol{x}) \geq 0$. By substitution,

$$(\boldsymbol{y} - \boldsymbol{x})^T \boldsymbol{Q} (\boldsymbol{y} - \boldsymbol{x}) = m(a_2 - a_1)^2 \geq 0,$$

which completes the proof.

21.10 Consider the problem: minimize $\|\boldsymbol{x}\|^2$ ($\boldsymbol{x} \in \mathbb{R}^n$) subject to $\boldsymbol{a}^T \boldsymbol{x} \geq b$, where $\boldsymbol{a} \in \mathbb{R}^n$ is a nonzero vector, and $b \in \mathbb{R}$, $b > 0$. Suppose $\boldsymbol{x}^*$ is a solution to the problem.

a. Show that the constraint set is convex.

b. Use the Karush-Kuhn-Tucker theorem to show that $\boldsymbol{a}^T \boldsymbol{x}^* = b$.

c. Show that $\boldsymbol{x}^*$ is unique, and find an expression for $\boldsymbol{x}^*$ in terms of $\boldsymbol{a}$ and b.

Ans.: a. Let $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}^T \boldsymbol{x} \geq b\}$, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Omega$, and $\lambda \in [0, 1]$. Then, $\boldsymbol{a}^T \boldsymbol{x}_1 \geq b$ and $\boldsymbol{a}^T \boldsymbol{x}_2 \geq b$. Therefore,

$$\begin{aligned} \boldsymbol{a}^T (\lambda \boldsymbol{x}_1 + (1 - \lambda) \boldsymbol{x}_2) &= \lambda \boldsymbol{a}^T \boldsymbol{x}_1 + (1 - \lambda) \boldsymbol{a}^T \boldsymbol{x}_2 \\ &\geq \lambda b + (1 - \lambda) b \\ &= b \end{aligned}$$

which means that $\lambda \boldsymbol{x}_1 + (1 - \lambda) \boldsymbol{x}_2 \in \Omega$. Hence, Ω is a convex set.

b. Rewrite the problem as

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & g(\boldsymbol{x}) \leq 0 \end{array}$$

where $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ and $g(\boldsymbol{x}) = b - \boldsymbol{a}^T\boldsymbol{x}$. Now, $\nabla g(\boldsymbol{x}) = -\boldsymbol{a} \neq \mathbf{0}$. Therefore, any feasible point is regular. By the Karush-Kuhn-Tucker theorem, there exists $\mu^* \geq 0$ such that

$$\begin{aligned} 2\boldsymbol{x}^* - \mu^*\boldsymbol{a} &= \mathbf{0} \\ \mu^*(b - \boldsymbol{a}^T\boldsymbol{x}^*) &= 0. \end{aligned}$$

Since $\boldsymbol{x}^*$ is a feasible point, then $\boldsymbol{x}^* \neq \mathbf{0}$. Therefore, by the first equation, we see that $\mu^* \neq 0$. The second equation then implies that $b - \boldsymbol{a}^T\boldsymbol{x}^* = 0$.

c. By the first Karush-Kuhn-Tucker equation, we have $\boldsymbol{x}^* = \mu^*\boldsymbol{a}/2$. Since $\boldsymbol{a}^T\boldsymbol{x}^* = b$, then $\mu^*\boldsymbol{a}^T\boldsymbol{a}/2 = \boldsymbol{a}^T\boldsymbol{x}^* = b$, and therefore $\mu^* = 2b/\|\boldsymbol{a}\|^2$. Since $\boldsymbol{x}^* = \mu^*\boldsymbol{a}/2$ then $\boldsymbol{x}^*$ is uniquely given by $\boldsymbol{x}^* = b\boldsymbol{a}/\|\boldsymbol{a}\|^2$.