COLUMBIA UNIVERSITY

Math S2500
Analysis and Optimization
Summer 2016

Midterm II
06.16.2016

Instructor: S. Ali Altug

Name and UNI: ________________________________

<table>
<thead>
<tr>
<th>Question:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points</td>
<td>15</td>
<td>20</td>
<td>30</td>
<td>10</td>
<td>75</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Instructions:

- There are 4 questions on this exam.
- Please write your NAME and UNI on top of EVERY page.
- Unless otherwise is explicitly stated SHOW YOUR WORK in every question.
- Please write neatly, and put your final answer in a box.
- No calculators, cell phones, books, notebooks, notes or cheat sheets are allowed.
1. (a) (10 points) Maximize 

\[ f(x) = e^x + y + z, \quad \text{subject to } \begin{cases} x + y + z = 1 \\ x^2 + y^2 + z^2 = 1. \end{cases} \]

(b) (5 points) Roughly, how much does the answer of the first part change if we change the constraints to \(x + y + z = 1.03\) and \(x^2 + y^2 + z^2 = 0.97?\) (Hint: Linear approximations! It may also be useful to remember that \(\lambda_i = \frac{\partial f^*}{\partial b_i}\), where \(f^*\) is the value function.)

Solution:

(a) The Lagrangian corresponding to the problem is 

\[ \mathcal{L}(x, y, z, \lambda_1, \lambda_2) = e^x + y + z - \lambda_1(x + y + z - 1) - \lambda_2(x^2 + y^2 + z^2 - 1). \]

The partial derivative and the constraint conditions are given by 

\[ \begin{align*} 
\mathcal{L}_x &= e^x - \lambda_1 - 2x\lambda_2 = 0, \\
\mathcal{L}_y &= 1 - \lambda_1 - 2y\lambda_2 = 0, \\
\mathcal{L}_z &= 1 - \lambda_1 - 2z\lambda_2 = 0, \\
x + y + z &= 1, \\
x^2 + y^2 + z^2 &= 1. 
\end{align*} \]

Then, 

\( (2) - (3) \Rightarrow \lambda_2(z - y) = 0 \Rightarrow \lambda_2 = 0 \) or \( z = y. \)

- \( \lambda_2 = 0.\) Substituting this in (3) implies \( \lambda_1 = 1,\) and substituting this into (1) then implies that \( x = 0.\) Substituting \( x = 0 \) in the constraints (4) and (5) implies that \( z(z - 1) = 0, \) and therefore either \( z = 0, y = 1 \) or \( z = 1, y = 0.\) So the candidate points are \( (x^*, y^*, \lambda_1^*, \lambda_2^*) = (0, 1, 0, 1, 0) \) or \( (0, 0, 1, 1, 0).\)

- \( z = y.\) Substituting this in the constraints gives \( y(3y - 2) = 0 \) and therefore \( y = z = 0, x = 1, \lambda_1 = 1, \lambda_2 = \frac{e - 1}{2} \) or \( y = z = \frac{2}{3}, x = -\frac{1}{3}, \lambda_1 = \frac{1 + 2e^{-1/3}}{3}, \lambda_2 = \frac{3(1 - e^{-1/3})}{4}.\) So the candidate points are \( (x^*, y^*, \lambda_1^*, \lambda_2^*) = (1, 0, 0, 1, \frac{e - 1}{2}) \) or \( (-\frac{1}{3}, \frac{2}{3}, \frac{1 + 2e^{-1/3}}{3}, \frac{1 - e^{-1/3}}{2}).\)

The corresponding values of the objective function are 

\( (x^*, y^*, \lambda_1^*, \lambda_2^*) \leftrightarrow f(x^*, y^*) \)

\( (0, 1, 0, 1, 0) \leftrightarrow 2, \)

\( (0, 0, 1, 1, 0) \leftrightarrow 2, \)

\( (1, 0, 0, 1, \frac{e - 1}{2}) \leftrightarrow e, \)

\( (-\frac{1}{3}, \frac{2}{3}, \frac{1 + 2e^{-1/3}}{3}, \frac{1 - e^{-1/3}}{2}) \leftrightarrow e^{-1/3} + \frac{4}{3}. \)

Therefore the maximum occurs at \( (1, 0, 0) \) and its value is \( e. \)

(b) The question is asking how much \( f^*(b) \) changes when when \( b^* = (b_1, b_2) \) changes from \( (1, 1) \) to \( (1.03, 0.97). \) The linear approximation to \( f^*(b) \) around \( b^* \) is given by 

\[ f^*(b) \sim f^*(b^*) + \frac{\partial f^*(b)}{\partial b_1}(b_1 - b_1^*) + \frac{\partial f^*(b)}{\partial b_2}(b_2 - b_2^*) \]

\[ = f^*(b^*) + \lambda_1^*(b_1 - b_1^*) + \lambda_2^*(b_2 - b_2^*) \]

\[ = f^*(b^*) + 0.03 + \frac{e - 1}{2}(-0.03) \]

\[ = f^*(b^*) + 0.03(1 - \frac{e}{2}). \]

Therefore the answer changes roughly by \( 0.03(1 - \frac{e}{2}). \)
2. Let \( n \geq 1 \) be an integer, let \( a_{ij} \in \mathbb{R} \) be such that \( a_{ij} = a_{ji} \) and let

\[
Q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j.
\]

We will find the maximum and the minimum of \( Q(x) \) subject to the constraint

\[
x_1^2 + x_2^2 + \cdots + x_n^2 \leq 1.
\]

(Note the domain defined by the constraint is a closed ball in \( \mathbb{R}^n \) so that the function necessarily has a maximum and a minimum subject to this constraint by the extreme value theorem.)

(a) (5 points) Let us denote the matrix associated to \( Q(x) \) by \( M_Q \). i.e. \( M_Q = (a_{ij})_{i,j=1}^{n} \). Show that the stationary points of \( Q(x) \) that are in the interior of the ball are all in the kernel of \( M_Q \). i.e. Show that

\[
x \text{ is a stationary point } \Rightarrow M_Q x = 0.
\]

(b) (3 points) Using part (a) show that for any stationary point \( x \) that is in the interior of the domain defined by the constraint \( Q(x) = 0 \).

For the rest of the question Let \( g(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2 \). Then the boundary of the domain is

\[
g(x_1, x_2, \ldots, x_n) - 1 = 0.
\]

(c) (5 points) Show that the equations for the Lagrange multipliers are exactly

\[
(M_Q - \lambda 1_{n \times n}) x = 0,
\]

where \( 1_{n \times n} \) denotes the \( n \times n \) identity matrix.

(d) (2 points) Using the constraint \( g(x) = 1 \) show that any solution to the Lagrange multiplier is indeed an eigenvector of \( M_Q \) (Recall that an eigenvector is \( x \) such that \( M_Q x = \lambda x \), and that \( x \) has to be non-zero!).

(e) (5 points) Using parts (b) and (d) and the spectral theorem conclude that the maximum is attained at the largest eigenvalue of \( M_Q \) (Recall that the spectral theorem tells us that the eigenvalues of a symmetric matrix are real!).

(f) (0 points) For the curious student: More is true along these lines. In the sense that, if we want to maximize \( Q \) subject to \( g(x) \leq 1 \) and that \( x \) is orthogonal to the vector maximizing the form we find that the maximum is the second largest eigenvalue of \( M_Q \). If we keep going like this, projecting on the orthogonal complement of the first \( k \) vectors, we find that the maximum is the \((k+1)\)st eigenvalue. What we are doing is nothing but Gram-Schmidt orthogonalization!

**Solution:**

(a) Differentiating with respect to each of the variables and using the fact that \( a_{ij} = a_{ji} \) we get that the stationary points are the solutions to the following system of equations:

\[
\frac{\partial Q}{\partial x_1} = 2 \sum_{j=1}^{n} a_{1j} x_j = 0,
\]

\[
\frac{\partial Q}{\partial x_2} = 2 \sum_{j=1}^{n} a_{2j} x_j = 0,
\]

\[
\vdots
\]

\[
\frac{\partial Q}{\partial x_n} = 2 \sum_{j=1}^{n} a_{nj} x_j = 0.
\]

The above equations can be written as \( M_Q x = 0 \), which means \( x \) is in the kernel of \( M_Q \).
(b) By part (a) if \( x \) is a stationary point inside the domain it has to satisfy \( MQx = 0 \). But note that \( Q(x) = \langle x, MQx \rangle = 0 \).

(c) The Lagrangian corresponding to the problem is

\[
L(x, \lambda) = Q(x) = \lambda(x_1^2 + x_2^2 + \cdots + x_n^2 - 1).
\]

Exactly as in part (a) we have

\[
\frac{\partial L}{\partial x_1} = 2 \sum_{j=1}^{n} a_{1j}x_j - 2\lambda x_1 = 0, \\
\frac{\partial L}{\partial x_2} = 2 \sum_{j=1}^{n} a_{2j}x_j - 2\lambda x_2 = 0, \\
\vdots \\
\frac{\partial L}{\partial x_n} = 2 \sum_{j=1}^{n} a_{nj}x_j - 2\lambda x_n = 0.
\]

Which can be written as \((MQ - \lambda I_{n \times n})x = 0\).

(d) By part (c) we see that \( x \) satisfies \( MQx = \lambda x \), and the constraint \( g(x) = 1 \) tells us that \( x \neq 0 \). Therefore \( x \) is an eigenvector.

(e) The maximum is going to be attained either in the interior or the boundary of the domain. By part (b) the only way the maximum is attained in the interior is if \( x \in \ker(MQ) \) in which case \( x \) is an eigenvector (with eigenvalue \( \lambda = 0 \)) otherwise it is attained on the boundary for some eigenvalue \( \lambda \). Now, by the spectral theorem we know that \( \lambda \in \mathbb{R} \). Hence, if there is an eigenvalue \( \lambda > 0 \) then the maximum is attained on the boundary and necessarily have to be an eigenvalue by part (d), and therefore has to be the largest eigenvalue. Otherwise the largest value is attained in the interior, which by part (b) is 0.
3. In this question we will solve
\[
\max \{3x^2 + 6y\}, \quad \text{subject to } \begin{cases} x^2 + y^2 \leq \frac{1}{4} \\
y \geq 0 \end{cases}.
\]

(a) (3 points) Bring the problem into the standard form and write the corresponding Lagrangian \( L(x, y, \lambda_1, \lambda_2) \).

(b) (5 points) Write the Kuhn-Tucker conditions for the problem.

(c) (3 points) Show that there are no solutions \((x^*, y^*, \lambda_1^*, \lambda_2^*)\) when both of the constraints (*) and (**) are active.

(d) (3 points) Show that there are no solutions when (*) is inactive and (**) is active.

(e) (6 points) If (*) is active and (**) is inactive show that the only solution to the K.-T. conditions is given by \((x^*, y^*, \lambda_1^*, \lambda_2^*) = (0, \frac{1}{2}, 6, 0)\).

(f) (3 points) Show that there are no solutions when both constraints are inactive.

(g) (6 points) Parts (c) to (f) shows that the only possible solution to our problem is the point \((x^*, y^*, \lambda_1^*, \lambda_2^*) = (0, \frac{1}{2}, 6, 0)\). In order to show that this indeed solves the maximum problem show that the Lagrangian that you wrote in part (a) is indeed concave when \(\lambda_1 = 6\) and \(\lambda_2 = 0\).

(h) (1 point) From part (g) conclude that the maximum value is 3.

Solution:

(a) The standard form of the problem is
\[
\max \{3x^2 + 6y\}, \quad \text{subject to } \begin{cases} x^2 + y^2 \leq \frac{1}{4} \\
y \geq 0 \end{cases}.
\]

The Lagrangian is
\[
L(x, y, \lambda_1, \lambda_2) = 3x^2 + 6y - \lambda_1(x^2 + y^2 - \frac{1}{4}) + \lambda_2 y.
\]

(b) The K.-T. conditions are
\[
\frac{\partial L(x, y, \lambda_1, \lambda_2)}{\partial x} = 6x - 2x\lambda_1 = 0, \quad (6)
\]
\[
\frac{\partial L(x, y, \lambda_1, \lambda_2)}{\partial y} = 6 - 2y\lambda_1 + \lambda_2 = 0 \quad (7)
\]
\[
\lambda_1 \geq 0 \quad \text{and} \quad x^2 + y^2 < \frac{1}{4} \Rightarrow \lambda_1 = 0, \quad (8)
\]
\[
\lambda_2 \geq 0 \quad \text{and} \quad y > 0 \Rightarrow \lambda_2 = 0. \quad (9)
\]

(c) If both (1) and (2) are active then we need to have \(y = 0\) which implies that \(\lambda_2 = -6\) but we need to have \(\lambda_2 \geq 0\) by the K.-T. conditions. Therefore we do not have any solutions in this case.

(d) If (1) is inactive and (2) and active we have \(x^2 + y^2 < \frac{1}{4}\) and \(y = 0\). By the K.-T. conditions for (1) we have \(\lambda_1 = 0\). Substituting this in (7) gives \(\lambda_2 = -6\) a contradiction.

(e) If (1) is active and (2) is inactive we have \(x^2 + y^2 = \frac{1}{4}\) and \(y > 0\). By the K.-T. conditions this implies that \(\lambda_2 = 0\). Substituting this into (7) implies
\[
3 = \lambda_1 y. \quad (7')
\]

On the other hand (6) implies that \(x(3 - \lambda_1) = 0\), which implies that \(\lambda_1 = 3\) or \(x = 0\).
• $\lambda_1 = 3$. Since we are assuming $\lambda_2 = 0$, by $(7')$ this implies that $y = 1$. But since we need to have $x^2 + y^2 = \frac{1}{4}$, there is no solution in this case.

• $x = 0$. In this case since $x^2 + y^2 = \frac{1}{4}$, $y = \pm \frac{1}{2}$. But we also need to have $y > 0$ therefore $y = \frac{1}{2}$ is the only possibility. Finally by $(7')$ this gives $\lambda_1 = 6$, which is admissible. Therefore the corresponding point is $(0, \frac{1}{2}, 6, 0)$.

(f) When both constraints are inactive we need to have $\lambda_1 = \lambda_2 = 0$ by the K.-T. conditions. Substituting this into (7) gives $6 = 0$ which is a contradiction. Therefore there are no solutions in this case.

(g) The Lagrangian restricted to $\lambda_1 = 6, \lambda_2 = 0$ is

$$L(x, y, 6, 0) = -3x^2 - 6y^2 - 6y + \frac{3}{2}.$$ 

The Hessian of this is

$$L''(x, y, 6, 0) = \begin{pmatrix} -6 & 0 \\ 0 & -12 \end{pmatrix},$$

which is negative definite. Therefore, $L(x, y, 6, 0)$ is concave.

(h) By part (g) the Lagrangian is concave at $\lambda_1 = 6, \lambda_2 = 0$ so it has a unique maximum. By part (e) we know that it is attained at $(x, y) = (0, \frac{1}{2})$. Hence the maximum of our function is $3 \cdot 0^2 + 6 \cdot \frac{1}{2} = 3$. 

4. A very useful characterization of a concave function is the following: \( f \) is concave if and only if the following inequality holds for every \( x_1, x_2, \cdots, x_m \in \mathbb{R}^n \) and every \( \lambda_1, \lambda_2, \cdots, \lambda_m \geq 0 \) with \( \lambda_1 + \lambda_2 + \cdots + \lambda_m = 1 \),

\[
f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_m x_m) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_m f(x_m).
\] (10)

The inequality (10) is known as Jensen’s inequality. Although Jensen’s inequality is straightforward to prove, it has important consequences.

(a) (5 points) Given that \( \ln(x) \) is a concave function (which is obvious from its graph) show that for \( x_1, x_2, \cdots, x_n > 0, \)

\[
\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}.
\]

This is known as the Arithmetic-Geometric mean inequality. (Hint: You may try taking \( \lambda_1 = \lambda_2 = \cdots = \lambda_n = \frac{1}{n} \) in (10).)

(b) (5 points) Using part (a) show that for \( x_1, x_2, \cdots, x_n > 0, \)

\[
\sqrt[n]{x_1 x_2 \cdots x_n} \geq \frac{1}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}.
\]

This is known as the Geometric-Harmonic mean inequality.

**Solution:**

(a) Following the hint, since \( \ln(x) \) is concave, by Jensen’s inequality we have

\[
\ln\left(\frac{1}{n}(x_1 + x_2 + \cdots + x_n)\right) \geq \frac{1}{n} (\ln(x_1) + \ln(x_2) + \cdots + \ln(x_n)).
\]

Exponentiating both sides (note that \( e^x \) is increasing so that exponentiating does not change the inequality) gives

\[
\frac{x_1 + x_2 + \cdots + x_n}{n} \geq e^{\frac{1}{n}(\ln(x_1) + \ln(x_2) + \cdots + \ln(x_n))} = \left(e^{\ln(x_1) + \ln(x_2) + \cdots + \ln(x_n)}\right)^{\frac{1}{n}} = \sqrt[n]{x_1 x_2 \cdots x_n}.
\]

(b) By (a) and that \( x_1, x_2, \cdots, x_n > 0 \) we have

\[
\sqrt[n]{\frac{1}{x_1} \frac{1}{x_2} \cdots \frac{1}{x_n}} \leq \frac{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}{n}
\]

\[
\Rightarrow \sqrt[n]{x_1 x_2 \cdots x_n} \geq \frac{1}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}.
\]