Math 164-1: Optimization

Instructor: Alpár R. Mészáros

First Midterm, April 20, 2016

Name (use a pen):

Student ID (use a pen):

Signature (use a pen):

Rules:

- Duration of the exam: **50 minutes**.
- By writing your name and signature on this exam paper, you attest that you are the person indicated and will adhere to the UCLA Student Conduct Code.
- You may use either a pen or a pencil to write your solutions. However, if you use a pencil I will withheld your paper for **two** weeks after grading it.
- **No** calculators, computers, cell phones (all the cell phones should be turned off during the exam), notes, books or other outside material are permitted on this exam. If you want to use scratch paper, you should ask for it from one of the supervisors. Do not use your own scratch paper!
- Please justify all your answers with mathematical precision and write rigorous and clear proofs. You may loose points in the lack of justification of your answers.
- Theorems from the lectures and homework may be used in order to justify your solution. In this case state the theorem you are using.
- This exam has 3 problems and is worth **20 points**. Adding up the indicated points you can observe that there are **27 points**, which means that there are **7 "bonus" points**. This permits to obtain the highest score 20, even if you do not answer some of the questions. On the other hand nobody can be bored during the exam. All scores higher than 20 will be considered as 20 in the gradebook.
- I wish you success!

Problem	Score
Exercise 1	
Exercise 2	
Exercise 3	
Total	

Exercise 1 (13 points).

Let us consider the rhombus $\Omega \subset \mathbb{R}^2$ bounded by the straight lines $y = x$, $y = -x$, $y = \sqrt{2} + x$ and $y = \sqrt{2} - x$. We consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given as $f(x, y) = x^2 + y^2 - \sqrt{2}y + 9/2$. We aim to find all the local minimizers and maximizers of f on Ω. Note that by Weierstrass' theorem (Ω is compact and f is continuous on Ω), both global minimizers and maximizers exist.

(1) Select all the candidates from the interior of Ω both for local minimizers and maximizers.

(2) Using second order sufficient conditions decide whether the selected points in (1) are local minimizers, maximizers or neither of them.

(3) Compute the feasible directions at $(0, 0)$. Using first and second order necessary conditions, decide whether $(0, 0)$ is a good candidate for local maximizer of f on Ω or not. Compute the value of f at $(0, 0)$ and compare it with the values of f at the other three vertices of Ω.

(4) Find the global maximizers of f on the side of Ω where $y = x$. Compute the value of f at these points. *Hint:* do not use Lagrange multipliers!

(5) Characterize and draw the following level sets of f: $LS_5 := \{(x, y) \in \mathbb{R}^2 : \ f(x, y) = 5\}$ and $LS_6 := \{(x, y) \in \mathbb{R}^2 : \ f(x, y) = 6\}$.

(6) Find all the global maximizers and minimizers of f on Ω. *Hint:* you may use some of the geometrical properties of f – from the previous point – and point (3).

Solutions

(1) For interior points the first order necessary condition reads as $\nabla f(x, y) = 0$ which is equivalent to $(x, y) = (0, \sqrt{2}/2)$, and clearly this is an interior point of Ω.

(2) Clearly $D^2 f(x, y) = 2I_2$ for all $(x, y) \in \mathbb{R}^2$ (where I_2 is the identity matrix in $\mathbb{R}^{2\times 2}$) which is a positive definite matrix, the point $(x, y) = (0, \sqrt{2}/2)$ selected in (1) is an interior point (at which ∇f vanishes), hence by the second order sufficient condition it is a strict local minimizer of f on Ω.

(3) Feasible directions at $(0, 0)$ are those vectors $e = (e_1, e_2) \in \mathbb{R}^2$ for which $e_2 \geq e_1$ and $e_2 \geq -e_1$ (hence they lie in the intersection of the two half-planes), from where one obtains that $e_2 \geq 0$ and $-e_2 \leq e_1 \leq e_2$.

At $(0, 0)$ the first order necessary condition for local maximizer reads as $\nabla f(0, 0) \cdot e \leq 0$ for all e feasible directions at $(0, 0)$. This is equivalent to $(0, -\sqrt{2}) \cdot e \leq 0$ which is $-\sqrt{2}e_2 \leq 0$, which holds true, since for the feasible directions at $(0, 0)$ one has $e_2 \geq 0$.

As seen in (2), $D^2 f(x, y) = 2I_2$ for all $(x, y) \in \mathbb{R}^2$, and the second order necessary condition (for $(0, 0)$ to be a local maximizer) should be that $e^\top D^2 f(0, 0)e \leq 0$ for all e feasible directions at $(0, 0)$ for which $\nabla f(0, 0) \cdot e = 0$. This is equivalent to $2(e_1^2 + e_2^2) \leq 0$ for all e feasible direction at $(0, 0)$ for which $e_1 = e_2 = 0$, which is a trivial direction and hence the second order necessary condition for $(0, 0)$ to be a local maximizer holds true.

$f(0, 0) = 9/2 = f(\sqrt{2}/2, \sqrt{2}/2) = f(-\sqrt{2}/2, \sqrt{2}/2) = f(0, \sqrt{2})$, thus the values of the function are the same at all the vertices of Ω.

(4) On the side of Ω where $y = x$ the function reduces to a function of one variable, i.e. $g(x) = f(x, x) = 2x^2 - \sqrt{2}x + 9/2$ which has to the optimized on $x \in [0, \sqrt{2}/2]$. In the interior of the interval if one has a local maximizer, one should have that $g'(x) = 0$, which is equivalent to $x = \sqrt{2}/4$, which is an interior point. But the parabola is convex, hence it has a global minimizer at this point ($f(\sqrt{2}/4, \sqrt{2}/4) = 17/4$). Since the parabola g is symmetric on $[0, \sqrt{2}/2]$, one has that the maximum of f is achieved in both endpoints on the line segment $(0, 0)$ and $(\sqrt{2}/2, \sqrt{2}/2)$, where one has computed already the values of f in (3) which was $9/2$.

(5) f can be written as

$$f(x, y) = x^2 + (y - \sqrt{2}/2)^2 + 4.$$

Hence all the level sets of f (for levels greater than 4) are circles around the point $(0, \sqrt{2}/2)$. In particular $LS_5 := \{(x, y) \in \mathbb{R}^2 : \ x^2 + (y - \sqrt{2}/2)^2 = 1\}$ and $LS_6 := \{(x, y) \in \mathbb{R}^2 : \ x^2 + (y - \sqrt{2}/2)^2 = 2\}$.

(6) Using the previous point, having in mind that the level sets of f are circles, the value of f at a point (x, y) represents the distance square of this point from the 'center' $(0, \sqrt{2}/2)$ plus 4, i.e

$$f(x, y) = \|(x, y) - (0, \sqrt{2}/2)\|^2 + 4.$$

Moreover Ω is symmetric w.r.t. $(0, \sqrt{2}/2)$ and f is radially increasing w.r.t. the center (larger is the level at which one considers the function, larger is the value of it), it is clear that its maximum is attained on a level set which has the largest radius still intersecting Ω (i.e. precisely $\sqrt{2}/2$). These points (the global maximizers of f) are precisely the 4 vertex points (which are on $LS_{\frac{9}{2}}$) and the value of the function at these points was computed in (3). By the same reasoning, the global minimum of the function is attained at the local minimizer (since this has distance 0 from the center), computed in (1), which is the 'center'. And the value of f here is 4.

Exercise 2 (8 points)**.**

Let $A \in \mathbb{R}^{n\times n}$, $n \geq 1$, be a symmetric matrix. We would like to find the minimum and the maximum of the so-called *Rayleigh quotient* associated to the matrix A. To do so, we define the Rayleigh quotient as the the function $R : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, defined as $R(x) = \dfrac{x^\top Ax}{\|x\|^2}$. Here $\|x\| := \sqrt{x_1^2 + \cdots + x_n^2}$ denotes the usual euclidean norm on $\mathbb{R}^n$.

(1) Show that $R(cx) = R(x)$ for all $x \in \mathbb{R}^n \setminus \{0\}$ and $c \in \mathbb{R} \setminus \{0\}$.

(2) Using (1), show that looking for the local maximizers and minimizers of R on $\mathbb{R}^n \setminus \{0\}$ is equivalent to look for the local extremizers of R on the set $\Omega := \{x \in \mathbb{R}^n : \ \|x\|^2 = 1\}$.

(3) Using eventually the theory of Lagrange multipliers, show that the possible candidates for the local extremizers of R on Ω are the eigenvectors (normalized to be points in Ω) of A.

(4) Characterize the tangent space $T(x^1)$ and the normal space $N(x^1)$ to Ω at x^1, where $x^1 \in \Omega$ is an eigenvector of A associated to the smallest eigenvalue λ_1 of A.

(5) Assuming that one can order the eigenvalues (with possible multiplicities) of A as $\lambda_1 < \lambda_2 \leq \cdots \leq \lambda_{n-1} < \lambda_n$, find the global minimizers and maximizers of R on Ω. Compute the minimum and the maximum of R.

Hint: you may plug in the candidates from (3) into R to find the global extremizers. Another possibility is to work with second order sufficient conditions from the Lagrangian theory.

Solutions

(1) By direct computation one checks $R(cx) = \dfrac{(cx)^\top A(cx)}{\|cx\|^2} = \dfrac{c^2x^\top Ax}{c^2\|x\|^2} = R(x)$.

(2) From (1) we see that R is depending only on the direction of the vector x and not on its length, in particular for all $x \in \mathbb{R}^n \setminus \{0\}$ one has that $R((1/\|x\|)x) = R(x)$ and $\|x/\|x\|\| = 1$, hence it is enough to look for the local minimizers and maximizers of R on the unit sphere Ω.

(3) By (2) one has that optimizing R on $\mathbb{R}^n \setminus \{0\}$ is equivalent to optimize it on Ω. Hence the local extremizers of R (and the optimal values) are the same as for the function $P : \mathbb{R}^n \to \mathbb{R}, \ \ P(x) = x^\top Ax$ on Ω. The first order necessary Lagrangian condition says that if $x \in \Omega$ is a local extremizer of P on Ω, then there exists a constant $\lambda \in \mathbb{R}$ such that $\nabla P(x) = \lambda\nabla h(x)$, where $h : \mathbb{R}^n \to \mathbb{R}$, $h(x) = \|x\|^2 - 1$. Computing the gradients, one has that

$$2Ax = 2\lambda x,$$

which is clearly the condition for $x \in \Omega$ to be the eigenvector of A associated to the eigenvalue $\lambda \in \mathbb{R}$ of A.

(4) By definition $T(x^1) = \{y \in \mathbb{R}^n : \ \nabla h(x^1) \cdot y = 0\} = \{y \in \mathbb{R}^n : \ 2x^1 \cdot y = 0\}$, which is the $(n-1)-$dimensional hyperplane in $\mathbb{R}^n$ orthogonal to x^1. Since $T(x^1)$ and $N(x^1)$ are orthogonal complements, one has that $N(x^1)$ is the line passing through 0, which has direction x^1.

(5) Since any local extremizer of P (hence of R) is an eigenvector of A (in Ω), let us compute $P(x^i)$, where x^i is any eigenvector associated to λ_i ($i \in \{1, \dots, n\}$.)

$$P(x^i) = (x^i)^\top Ax^i = (x^i)^\top \lambda_i x^i = \lambda_i.$$

This implies in particular that $P(x^i) \leq P(x^j)$, if $i \leq j$ and in particular $P(x^1) < P(x^j) < P(x^n)$ for all $j \in \{2, \dots, n-1\}$. We know that one can chose a set of n eigenvectors that form an orthonormal basis of $\mathbb{R}^n$, hence every $x \in \mathbb{R}^n$ can be written as linear combination of these vectors. Thus, if $x \in \mathbb{R}^n$ with $\|x\| = 1$, one has that

$$P(x) = x^\top Ax = c_1^2\lambda_1 + \dots c_n^2\lambda_n,$$

where c_i is the coefficient in the front of x^i, while writing x in the orthonormal basis of the eigenvectors. In particular $c_1^2 + \cdots + c_n^2 = 1$. Hence it is clear that $P(x^1) = \lambda_1 < P(x) < \lambda_n = P(x^n)$, for all $\|x\| = 1$, hence clearly the eigenvectors corresponding to λ_1 are strict global minimizers of P (hence of R) and the eigenvectors corresponding to λ_n are strict global maximizers of P (hence of R) on Ω. The minimal and maximal values of P and R are λ_1 and λ_n respectively.

If one wants to work with the second order sufficient Lagrangian conditions instead (which is more difficult), one has to consider the matrices

$$L(x^i, \lambda_i) = 2A - 2\lambda_i I_n,$$

and check the sign of $y^\top L(x^i, \lambda_i)y$ for all $y \in T(x^i)$, where $I_n \in \mathbb{R}^{n\times n}$ denotes the identity matrix. Let us check it for $i = 1$ (we forget about the positive coefficient 2).

$$y^\top L(x^i, \lambda_i)y = y^\top Ay - \lambda_1\|y\|^2.$$

Since $y \in T(x^1)$, one has that y is orthogonal to x^1, hence while writing y in the orthonormal basis formed by eigenvectors of A one does not have a component of x^1, i.e. $y = c_2x^2 + \dots c_nx^n$ (with $c_2^2 + \dots c_n^2 = \|y\|^2$) which implies that

$$y^\top L(x^i, \lambda_i)y = y^\top Ay - \lambda_1\|y\|^2 = c_2^2\lambda_2 + \dots c_n^2\lambda_n - \lambda_1\|y\|^2 > \|y\|^2(\lambda_2 - \lambda_1) > 0,$$

thus by the second order sufficient condition all the eigenvectors associated to λ_1 are strict local minimizers of P (hence of R) on Ω. The fact that they are global follows from the same arguments as in the first approach.

The global maximality of the eigenvectors associated to λ_n follow from a similar argument.

Exercise 3 (6 points)**.**

Let us consider the domain $\Omega \subset \mathbb{R}^2$ defined as $\Omega := \{(x,y) \in \mathbb{R}^2 : x \in [1,e];\ 0 \leq y \leq \ln(x)\}$, where e denotes the base of the natural logarithm ln. We aim to find the global minimizers and maximizers of the function $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = e^{xy}$ on Ω. Here also we observe that we are optimizing a continuous function on a compact set, hence the existence of a global minimizer and a global maximizer is given by the Theorem of Weierstrass.

(1) Show that there are no candidates for local extremizers of f in the interior of Ω and deduce that the global extremizers lie on $\partial\Omega$.

(2) Decompose $\partial\Omega$ into three pieces and find the global maximizers and minimizers of f on each of the pieces. Explain why cannot we use the theory of Lagrange multipliers for equality constraints on each piece of the boundary separately! Observe then that to find the extremizers on each piece of the boundary reduces the problem to a 1D problem.

(3) Deduce from (2) the global minimizers and maximizers of f on Ω. Compute the maximal and minimal values as well.

Solutions

(1) For a local extremizer (x,y) in the interior of Ω one should have that $\nabla f(x,y) = 0$, which is equivalent to $(x,y) = (0,0)$. Since this point is outside of Ω one deduces that there are no interior local extremizers of f and all the local ones lie on $\partial\Omega$.

(2) Clearly $\partial\Omega$ decomposes into 3 pieces, $C_1 := \{(x,y) \in \mathbb{R}^2 :\ y = \ln(x), x \in [1,e]\}$, $C_2 := \{(e,y) \in \mathbb{R}^2 :\ 0 \leq y \leq \ln(e) = 1\}$ and $C_1 := \{(x,0) \in \mathbb{R}^2 :\ x \in [1,e]\}$. One cannot use the Lagrangian theory with equality constraints on the 3 pieces of the boundary, because to describe these pieces one needs inequalities as well. This situation can be handled by the so-called KKT theory, which was not the subject of this exam.

Let us optimize f on each C_i, $i \in \{1,2,3\}$. On C_1 the function f can be written as $g_1(x) = e^{x\ln x}$ which has to be optimized on $[1,e]$. Clearly $g_1'(x) = e^{x\ln x}(\ln x + 1) > 0$ for all $x \in [1,e]$, thus g_1 is strictly increasing on this interval, having its global minimizer at $x = 1$, where $g_1(1) = 1$ and its global maximizer at $x = e$, where $g_1(e) = e^e$. Hence the global minimizer of f on C_1 is the point $(1,0)$ and the global maximizer is $(e,1)$ with the corresponding values.

On C_2 f reduces to the function $g_2 = e^{ey}$ which has to be optimized on $[0,1]$. Clearly $g_2'(y) = ee^{ey} > 0$ for all $y \in [0,1]$, which means that g_2 is strictly increasing on $[0,1]$, hence its global minimizer and maximizer are $y = 0$ and $y = 1$ respectively, with the optimal values $g_2(0) = 1$ and $g_2(1) = e^e$. The corresponding global minimizer and maximizer of f are the points $(e,0)$ and $(e,1)$ respectively.

Similarly on C_3 the problem is to optimize $g_3(x) = 1$ on the interval $[1,e]$, for which all the points are both global minimizers and maximizers. The same is true for f on C_3.

(3) By (1) one knows that the global extremizers lie on $\partial\Omega$. In (2) we computed all the global extremizers on each piece of the boundary. Hence one has to select the points from the selected ones in (2) which have the smallest and the biggest value. We see that the global maximizer of f is unique and it is the point $(e,1)$, with maximal value $f(e,1) = e^e$. There are infinitely many global maximizers which are the points $(x,0)$, where $x \in [1,e]$. The minimal value of the function is 1.