Math 164-1: Optimization

Instructor: Alpár R. Mészáros

Midterm, October 16, 2015 – Solutions

Exercise 1 (12 points)**.**

Let $\Omega \subset \mathbb{R}^3$, $\Omega := \left\{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ -1/2 \le x_i \le 1/2, \ i \in \{1,2,3\}\right\}$ be a unit cube in $\mathbb{R}^3$. Let us define the function $f : \mathbb{R}^3 \to \mathbb{R}$ as $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3x_1$. We consider now the optimization problem

$$\max_{x\in\Omega} f(x) \tag{1}$$

Answer the following questions, showing full justification!

(1) Why does a maximizer exist in Problem 1?

(2) Why is the function f of class $C^2(\mathbb{R}^3)$? Compute the gradient $\nabla f(x)$ and the Hessian $D^2f(x)$ for an arbitrary point $x \in \mathbb{R}^3$.

(3) Select all the candidates from the interior of Ω that could be local maximizers of f. Decide whether the selected candidates are local maximizers or not.

(4) Consider the point $y = (1/2, 0, 0) \in \Omega$. Is y an interior point of Ω? Characterize all the feasible directions $e = (e_1, e_2, e_3) \in \mathbb{R}^3$ at y (a picture could help). Is y a (strict) local maximizer of f?

(5) Consider the point $z = (1/2, 1/2, 1/2) \in \Omega$. Is z an interior point of Ω? Characterize all the feasible directions $e = (e_1, e_2, e_3) \in \mathbb{R}^3$ at z. Is z a (strict) local maximizer of f? What is the value of f at z?

(6) Is z from (5) a global maximizer of f? Determine another point $t \in \Omega$, $t \ne z$ which is a global maximizer of f.

Solution

(1) Since f is continuous on $\mathbb{R}^3$ (as a second order polynomial) and Ω is a compact set (closed and bounded) in $\mathbb{R}^3$, Weierstrass' theorem ensures the existence of minimizers and maximizers of f over Ω.

(2) Since f is a second order polynomial defined on $\mathbb{R}^3$, it is of class $C^\infty(\mathbb{R}^3)$, hence also $C^2(\mathbb{R}^3)$. Moreover $\nabla f(x) = (x_2 + x_3;\ x_1 + x_3;\ x_1 + x_2)$ and

$$D^2f(x) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

are continuous functions on $\mathbb{R}^3$, which proves once again that $f \in C^2(\mathbb{R}^3)$.

(3) The candidates $x = (x_1, x_2, x_3) \in int(\Omega)$ for maximizers should satisfy the first order optimality condition $\nabla f(x) = 0$. This is equivalent to $x = (0, 0, 0) = 0$ which is clearly an interior point of Ω. Since the Hessian $D^2f(0, 0, 0)$ is indefinite (for $e = (e_1, e_2, e_3) \in \mathbb{R}^3, e \ne 0$ one has $e \cdot D^2f(0)e = 2(e_1e_2 + e_2e_3 + e_3e_1)$ which is negative for instance for $(-1, 1, 0)$ and positive for $(1, 1, 0)$. On the other hand the eigenvalues of $D^2f(0)$ are $-1, -1$ and 2 and the change of sign for the eigenvalues also shows that the matrix is indefinite), the interior point is neither a local minimizer nor a local maximizer.

(4) $y \in \partial\Omega$. By definition $e = (e_1, e_2, e_3) \in \mathbb{R}^3$ is a feasible direction at y if there exists ε_0 such that $y + \varepsilon e \in \Omega$ for all $\varepsilon \in [0, \varepsilon_0]$. By this, if e is a feasible direction one should have $y + \varepsilon e \in \Omega$ for $\varepsilon > 0$ small enough. This means that one should have

$$(1/2 + \varepsilon e_1, \varepsilon e_2, \varepsilon e_3) \in \Omega.$$

This clearly implies (since $\varepsilon > 0$ and using the definition of Ω) that $e_1 \leq 0$ and $e_2, e_3 \in \mathbb{R}$.

On the other hand $\nabla f(y) = (0, 1/2, 1/2)$ and for all feasible directions $e \in \mathbb{R}^3$ described in the above line one should have the first order necessary optimality condition $\nabla f(y) \cdot e \leq 0$, i.e. $1/2e_2 + 1/2e_3 \leq 0$. By the characterization of the feasible directions at y, e_2 and e_3 can be arbitrary, so the condition does not hold (if $e_2 = 1, e_3 = 0$ one has $1/2e_2 + 1/2e_3 = 1/2 \geq 0$). Hence y is not a local maximizer of f over Ω.

(5)-(6) $z \in \partial\Omega$. Using the definition of the feasible directions $e = (e_1, e_2, e_3) \in \mathbb{R}^3$ at z (as in the previous point), one should have $z + \varepsilon e \in \Omega$ for $\varepsilon > 0$ small. This is equivalent to say that $(1/2 + \varepsilon e_1, 1/2 + \varepsilon e_2, 1/2 + \varepsilon e_3) \in \Omega$ which by the definition of Ω and by the fact that $\varepsilon > 0$ implies that e is a feasible direction iff $e_1 \leq 0, e_2 \leq 0$ and $e_3 \leq 0$.

On the other hand $\nabla f(z) = (1, 1, 1)$ and the first order optimality condition for z (to be a local maximizer) is $\nabla f(z) \cdot e \leq 0$ for all feasible directions $e \in \mathbb{R}^3$. This is equivalent to $e_1 + e_2 + e_3 \leq 0$, which by the above characterization holds true. Hence z is a good candidate for a local minimizer.

It is clear that $f(z) = 3/4$. On the other hand $f(x) = x_1x_2 + x_2x_3 + x_3x_1 \leq |x_1||x_2| + |x_2||x_3| + |x_3||x_1| \leq 3/4$ for all $x = (x_1, x_2, x_3) \in \Omega$, with other words $3/4$ is an upper bound for f on Ω. And since $f(z) = 3/4$ (i.e. the upper bound is achieved) this implies that z is a global maximizer.

Taking a neighborhood point of z in Ω, this can be described as $(1/2 - \varepsilon_1, 1/2 - \varepsilon_2, 1/2 - \varepsilon_3)$ for some $\varepsilon_1, \varepsilon_2, \varepsilon_3 \geq 0$ small reals (not all of them 0 at once). It is clear that $3/4 = f(z) > f(1/2 - \varepsilon_1, 1/2 - \varepsilon_2, 1/2 - \varepsilon_3) = (1/2 - \varepsilon_1)(1/2 - \varepsilon_2) + (1/2 - \varepsilon_2)(1/2 - \varepsilon_3) + (1/2 - \varepsilon_3)(1/2 - \varepsilon_1)$, just by simple comparison, which means that z is also a strict local maximizer of f.

By the very same reasoning $t = (-1/2, -1/2, -1/2) \in \Omega$ for instance satisfies all the properties as z, $t \neq z$ and $f(t) = f(z) = 3/4$, which means that t is a global maximizer as well, hence z is not a strict (or unique) global maximizer.

Exercise 2 (8 points).

Let $\Omega \subset \mathbb{R}^2$, $\Omega := \left\{x = (x_1, x_2) \in \mathbb{R}^2 \; : \; x_1^2 + x_2^2 \le 3\right\}$. We define the function $f : \mathbb{R}^2 \to \mathbb{R}$ as $f(x_1, x_2) = x_1 x_2^2$. We wish to study the local minimizers and maximizers of f on Ω. Answer the following questions, showing full justification!

(1) Find all the candidates for local minimizers and maximizers in the interior of Ω.

(2) Study the points selected in (1) and decide whether they are (strict) local minimizers, (strict) local maximizers or neither of them.

(3) Characterize the feasible directions $e = (e_1, e_2) \in \mathbb{R}^2$ and $d = (d_1, d_2) \in \mathbb{R}^2$ at the points $x = (1, \sqrt{2})$ and $y = (-1, \sqrt{2})$ and determine whether x and y satisfy the first and second order necessary optimality conditions for maximizers and minimizers respectively!

(4) Is x a (unique) global maximizer? Is y a (unique) global minimizer?

Solution

(1) Since $f \in C^2(\mathbb{R}^2)$ (it is a third order polynomial), for any point $x = (x_1, x_2) \in \mathbb{R}^2$ one has $\nabla f(x) = (x_2^2; \;\; 2x_1x_2)$ and

$$D^2 f(x) = \begin{bmatrix} 0 & 2x_2 \\ 2x_2 & 2x_1 \end{bmatrix}$$

Since we are looking for interior points as candidates, the first order necessary condition for both minimizers and maximizers is

$$\nabla f(x^*) = 0,$$

and from here the possible candidates are $(x_1^*, 0)$, where $x_1^* \in (-\sqrt{3}, \sqrt{3})$.

(2) Calculating the Hessian matrix in the above points one has

$$D^2 f(x) = \begin{bmatrix} 0 & 0 \\ 0 & 2x_1^* \end{bmatrix},$$

which is neither positive definite nor negative definite, hence one cannot uses SOSC. So, let us use the definition of local extremizers, to check whether the selected candidates are local maximizers or minimizers. First, take $x_1^* < 0$ and take $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$ small in absolute value such that $-\sqrt{3} < x_1^* + \varepsilon_1 < 0$. Now compute

$$f(x_1^* + \varepsilon_1, \varepsilon_2) = (x_1^* + \varepsilon_1)\varepsilon_2^2 \le 0 = f(x_1^*, 0),$$

hence the points $(x_1^*, 0)$, when $x_1^* \in (-\sqrt{3}, 0)$ are local maximizer that are not strict. In a similar way one can show that all the points $(x_1^*, 0)$, where $x_1^* \in (0, \sqrt{3})$ are local minimizers that are not strict!

The point $(0, 0)$ is also selected, however this is neither local minimizer, not local maximizer. To see this it is enough to use the definition and compare $0 = f(0, 0)$ with $f(\varepsilon_1, \varepsilon_2) = \varepsilon_1 \varepsilon_2^2$, where $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$ with small absolute value. Since $\varepsilon_1 \varepsilon_2^2 > 0$ if $\varepsilon_1 > 0$ and $\varepsilon_1 \varepsilon_2^2 < 0$ if $\varepsilon_1 < 0$, these facts contradict to the definition of both local minimizers and maximizers.

(3) We use the definition of a feasible direction $e \in \mathbb{R}^2$ at the given points $x, y \in \partial\Omega$. There exists $\varepsilon_0 > 0$ such that $x + \varepsilon e \in \Omega$ for all $\varepsilon \in [0, \varepsilon_0]$. By the definition of Ω this means that

$$(1 + \varepsilon e_1)^2 + (\sqrt{2} + \varepsilon e_2)^2 \le 3.$$

Chose $\varepsilon > 0$. The above inequality is then equivalent to

$$2e_1 + 2\sqrt{2}e_2 + \varepsilon(e_1^2 + e_2^2) \le 0.$$

Since $\varepsilon > 0$ could be arbitrary small, we should have

$$e_1 + \sqrt{2}e_2 \le 0,$$

and this inequality characterizes the feasible directions at x. Using similar computations the feasible directions $d = (d_1, d_2) \in \mathbb{R}^d$ at y are characterized by the inequality

$$-d_1 + \sqrt{2}d_2 \le 0.$$

We have moreover $\nabla f(x) = (2;\ \ 2\sqrt{2}), \nabla f(y) = (2;\ \ -2\sqrt{2})$ and

$$D^2 f(x) = \begin{bmatrix} 0 & 2\sqrt{2} \\ 2\sqrt{2} & 2 \end{bmatrix}, \quad D^2 f(y) = \begin{bmatrix} 0 & 2\sqrt{2} \\ 2\sqrt{2} & -2 \end{bmatrix}.$$

Let us check now the first and second order optimality conditions for x and y. For all feasible directions $e \in \mathbb{R}^2$ and $d \in \mathbb{R}^2$ at the points x and y respectively one has

$$\nabla f(x) \cdot e = 2(e_1 + \sqrt{2}e_2) \le 0, \text{ and } \nabla f(x) \cdot d = 2(d_1 - \sqrt{2}d_2) = -2(-d_1 + \sqrt{2}d_2) \ge 0,$$

where we used the characterization of e and d. Hence x satisfies the f.o.n.c. for local maximizers, while y satisfies the f.o.n.c. for minimizers.

On the other hand

$$e^T D^2 f(x) e = 4\sqrt{2}e_1 e_2 + 2e_2^2 \le 0$$

for all feasible directions for which $\nabla f(x) \cdot e = 0$. Indeed, this last condition says that we have to consider only the feasible directions $e = (e_1, e_2)$ such that $e_1 + \sqrt{2}e_2 = 0$. And

$$4\sqrt{2}e_1 e_2 + 2e_2^2 = 3\sqrt{2}e_1 e_2 + \sqrt{2}e_2(e_1 + \sqrt{2}e_2) = 3\sqrt{2}e_1 e_2 \le 0,$$

where the last inequality is coming from the fact that e_1 and e_2 have opposite signs. Thus x satisfies the second order necessary condition for maximizers.

Doing the same calculation and argument for y, here we need to take only those feasible directions $d = (d_1, d_2)$ for which $\nabla f(y) \cdot d = 0$ as well, i.e. $d_1 = \sqrt{2}d_2$. In this case

$$d^T D^2 f(y) d = 4\sqrt{2}d_1 d_2 - 2d_2^2 = 3\sqrt{2}d_1 d_2 + \sqrt{2}d_2(d_1 - \sqrt{2}d_2) \le 0,$$

since d_1 and d_2 have the same sings. Thus y satisfies also the second order necessary condition for the minimizers.

(4) Is is clear that $f(x) = 2$ and $f(y) = -2$. On the other hand by the definition of f, i.e. $f(x_1, x_2) = x_1 x_2^2$, for $x_1 > 0$ and by the constraint set, i.e. $x_1^2 + x_2^2 \le 3$, i.e. $x_2^2 \le 3 - x_1^2$, one has that

$$f(x_1, x_2) = x_1 x_2^2 \le x_1(3 - x_1^2) = 3x_1 - x_1^3.$$

The last expression clearly takes its maximum (on the set $[0, \sqrt{3}]$) at $x_1 = 1$ and its value is 2. To see this it is enough to look at the sign of the derivative of the expression $3x_1 - x_1^3$, i.e. $3 - 3x_1^2 \ge 0$ for $x_1 \in [0, 1]$ and negative for $x_1 \in (1, \sqrt{2}]$, which proves that the expression $3x_1 - x_1^3$ has its maximum at $x_1 = 1$. So by this we have found that

$$f(x_1, x_2) \le 2, \ \ \forall (x_1, x_2) \in \Omega.$$

And as $f(x) = 2$, one has that the point x is a global maximizer of f on Ω. It is not unique, since $f(1, -\sqrt{2}) = 2$.

By a similar argument one can show that $f(x_1, x_2) \ge -2, \ \ \forall (x_1, x_2) \in \Omega$, and since $f(y) = -2$, one can conclude that y is a global minimizer. However since $f(-1, -\sqrt{2}) = -2$, it is not unique neither.

Exercise 3 (5 points)**.**

Imagine that we have 3 events that have probability p_1, p_2 and p_3 respectively (these are real numbers in $[0, 1]$), moreover we know that $p_3 = 1 - p_1 - p_2$ and $p_3 \geq 0$.

Suppose that we want to find the discrete probability distribution (p_1^*, p_2^*, p_3^*) for these three events in such a way that (p_1^*, p_2^*) maximizes the Shannon-type entropy (a function widely used in information theory, statistical physics and in many other fields as a measure of uncertainty), defined as

$$E(p_1, p_2) := -p_1 \ln p_1 - p_2 \ln p_2,$$

where "ln" denotes the natural logarithm, i.e. of base $e \approx 2.71$. By the structure of the events one has always that

$$p_3^* = 1 - (p_1^* + p_2^*) \tag{2}$$

that is why we do not have to optimize also in p_3.

Since we are dealing with probabilities one has that $p_1, p_2, p_3 \in [0, 1]$ and by the structural condition (2) one has to impose also the constraint that $p_1+p_2 \leq 1$. Setting $\Omega := \{(p_1, p_2) \ : \ p_1, p_2 \in [0, 1], \ p_1 + p_2 \leq 1\}$, find all the interior local maximizers $(p_1^*, p_2^*) \in int(\Omega)$ of E, and determine the optimal value also for p_3^* using formula (2).

What can we say about the boundary points of Ω satisfying $p_1 + p_2 = 1$ and $p_1, p_2 \in (0, 1)$? Are there any, satisfying the first order necessary conditions for maximizers?

Solution

Let us study the optimization problem for (p_1, p_2). If the local maximizers $(p_1^*, p_2*) \in int(\Omega)$, then one should have $\nabla E(p_1^*, p_2^*) = 0$. Since $\nabla E(p_1, p_2) = (-(\ln p_1 + 1); \ \ -(\ln p_2 + 1))$, the above condition gives us that $p_1^* = p_2^* = e^{-1}$. We observe that $0 < p_1^*, p_2^* < 1$ and $p_1^* + p_2^* = 2e^{-1} < 1$, hence (p_1^*, p_2^*) is in the interior of Ω and is the only candidate to be a maximizer of E in $int(\Omega)$.

Since

$$D^2E(p_1, p_2) = \begin{bmatrix} -1/p_1 & 0 \\ 0 & -1/p_2 \end{bmatrix},$$

one has that

$$D^2E(p_1^*, p_2^*) = \begin{bmatrix} -e & 0 \\ 0 & -e \end{bmatrix}$$

which is a negative definite matrix. Thus both the second order necessary and sufficient conditions are fulfilled for

$$(p_1^*, p_2^*) = (e^{-1}, e^{-1})$$

and it is a local maximizer. In this case $p_3^* = 1 - 2/e$.

We should now find the feasible directions at any point on the part of the boundary of Ω where $0 < p_1, p_2 < 1$ and $p_1 + p_2 = 1$. Using the definition, we find that $d = (d_1, d_2) \in \mathbb{R}^2$ is a feasible direction at a point described before if $d_1 + d_2 \leq 0$.

The first order necessary condition at a boundary point like this reads as $\nabla E(p_1, p_2) \cdot d \leq 0$, which is equivalent to

$$-d_1 \ln p_1 - d_1 - d_2 \ln p_2 - d_2 \leq 0.$$

By the characterization of d, one can choose for d for instance $(d_1, -d_1)$, where $d_1 \in \mathbb{R}$ and by definition $p_2 = 1 - p_1$. By this observations the above inequality becomes in this case

$$d_1 \ln(1/p_1 - 1) \leq 0,$$

that clearly does not hold true for all $d_1 \in \mathbb{R}$, hence no point from this piece of $\partial\Omega$ satisfy the first order necessary condition for local maximizers.

In the above argument one had to exclude the point $p_1 = p_2 = 1/2$. However in this situation, taking (d_1, d_2) general feasible directions, one obtains $\nabla E(1/2, 1/2) \cdot d = (d_1 + d_2)(\ln(2) - 1) \geq 0$, hence this cannot be local maximizer either.