Math 164 (Lec 1): Optimization

Instructor: Alpár R. Mészáros

Midterm, October 26, 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 lose 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 **26 points**, which means that there are **6 "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.
- The problems are not necessarily ordered with respect to easiness!
- I wish you success!

Problem	Score
Exercise 1	
Exercise 2	
Exercise 3	
Total	

Exercise 1 (7 points)**.**

(1) Let us consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as

$$f(x_1, x_2) = \sin(x_1)\cos(x_2).$$

Find all the local minimizers and maximizers of f on the set $S = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 - \cos(x_1) = 0 \text{ and } x_2 - \sin(x_1) = 0\}$. Find the global minimizers and maximizers of f on S as well. Are they unique? Justify your answer!

Hint: it is easier to begin describing the geometry of S and solve the problem without Lagrange multipliers.

(2) Let $\Omega := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \in [\pi/4, 5\pi/4] \text{ and } \cos(x_1) \le x_2 \le \sin(x_1)\}$. Consider moreover $g : \mathbb{R}^2 \to \mathbb{R}$ defined as

$$g(x_1, x_2) = 1 - x_2^2.$$

Find the global minimizers and maximizers of g on Ω. Are they unique? Justify your answers!

Hint: it is more complicated to use the KKT techniques, than to use Lagrange multipliers (+ dealing separately with the interior case). Nevertheless, first understand the geometry of Ω and the structure of g.

Solutions

(1) The feasible set S geometrically represents the points in the intersection of the graphs of sin and cos defined on the whole $\mathbb{R}$. Hence it is easy to write S as

$$S = \left\{ \left(\frac{\pi}{4} + 2k\pi, \frac{\sqrt{2}}{2} \right) : k \in \mathbb{Z} \right\} \cup \left\{ \left(\frac{5\pi}{4} + 2k\pi, -\frac{\sqrt{2}}{2} \right) : k \in \mathbb{Z} \right\}.$$

So, to find the extremizers of f on S, one has to evaluate it in the discrete points of S. Clearly,

$$f(\pi/4 + 2k\pi, \sqrt{2}/2) = \sqrt{2}/2 \cos(\sqrt{2}/2) > 0$$

and

$$f(5\pi/4 + 2k\pi, -\sqrt{2}/2) = -\sqrt{2}/2 \cos(-\sqrt{2}/2) = -\sqrt{2}/2 \cos(\sqrt{2}/2) < 0,$$

for all $k \in \mathbb{Z}$. This shows that all the points $\left\{ \left(\frac{\pi}{4} + 2k\pi, \frac{\sqrt{2}}{2} \right) : k \in \mathbb{Z} \right\}$ are global maximizers (hence there is no uniqueness), while the points $\left\{ \left(\frac{5\pi}{4} + 2k\pi, -\frac{\sqrt{2}}{2} \right) : k \in \mathbb{Z} \right\}$ are all global minimizers (hence there is no uniqueness). This are trivially local maximizers and local minimizers, respectively.

(2) A figure helps you to see that on the interval $[\pi/4, 5\pi/4]$ the graph of cos is below of the graph of sin, hence Ω is just the region between these two graphs. Observe moreover that g is independent of the first variable, hence one has to only understand in what range x_2 is varying. Actually, the maximal range is $x_2 \in [-1, 1]$ (the endpoint of this intervals are achieved at the points $(\pi, -1)$ and $(\pi/2, 1)$ of Ω).

Clearly, $g(x_1, x_2) > g(\pi, -1) = g(\pi/2, 1) = 0$, for any other $(x_1, x_2) \in \Omega$, so these points are the only global minimizers (uniqueness does not hold). One the other hand $g(x_1, x_2) < g(x_1^*, 0) = 1$, where $x_1^* \in [\pi/2, \pi]$ and (x_1, x_2) is any other point from Ω. Hence $(x_1^*, 0)$ are all global maximizers (uniqueness does not hold).

If one wants to use the Lagrangian theory instead, first the candidates for local minimizers and maximizers from the interior of Ω are those points for which $\nabla g(x_1, x_2) = 0$. This condition gives the points $(x_1^*, 0)$, where $x_1^* \in (\pi/2, \pi)$. For all these points, the value of g is 1. So these all will be global maximizers, by definition.

The boundary consists of two pieces: $x_2 = \sin(x_1)$, and $x_2 = \cos(x_1)$, where $x_1 \in [\pi/4, 5\pi/4]$. So, on the first piece

$$g(x_1, x_2) = 1 - \sin^2(x_1) = \cos^2(x_1),$$

which has minimal value at $x_1 = \pi/2$ (corresponding to the point $(\pi/2, 1) \in \Omega$) which is 0. And it has maximal value at $x_1 = \pi$ (corresponding to the point $(\pi, 0) \in \Omega$) which is 1.

One the other piece of the boundary, one can write

$$g(x_1, x_2) = 1 - \cos^2(x_1) = \sin^2(x_1),$$

which has maximal value at $x_1 = \pi/2$ (corresponding to the point $(\pi/2, 0) \in \Omega$) which is 1. And it has minimal value at $x_1 = \pi$ (corresponding to the point $(\pi, -1) \in \Omega$) which is 0.

Gluing together the information from the interior and from the boundary cases, one obtains exactly the same solutions as with the first method.

Exercise 2 (Aquathlon problem – 9 points)**.**

John is doing aquathlon, which is a sport combining running and swimming. The start is situated on the left riverside and the finish is situated on the right riverside of the same river. Hence, he needs to run along the river on the riverside, cross the river by swimming, then possibly run further on the other side. The sides of the river are parallel straight lines, and the width of the river is uniform everywhere, and it is $d > 0$ meters. It is also allowed to immediately cross the river by swimming, or run only on the left side and arrive to the finish by swimming. The length of each riverside (where the race takes place) measures $L = 500$ meters.

We suppose that,

$$d < \sqrt{3}L, \tag{width}$$

meaning that the river is not that wide, compared to the length of the race.

Knowing that John can run with a constant uniform speed of $v_r = 2\ m/s$, and he can swim with a constant uniform speed of $v_s = 1\ m/s$, he wants to minimize the time that he needs for this combined race. We also assume that he will swim along a straight line from one riverside to the other. Help John to find an optimal strategy, i.e. tell him how far to run on the left side, then in which angle to start swimming and then to run on the other side, in order to achieve the shortest possible time.

(1) Introducing some variables if it is necessary, write an objective function and a feasible set, where the objective function should be minimized.

Hint: one knows that the constant speed is given by ratio of the distance travelled and the duration of the time.

(2) Show that if there is an optimal strategy from the interior of the feasible set, then that is depending only on the angle in which John needs to start swimming. Deduce from this the non-uniqueness of the optimal strategy.

(3) Show that there exists an optimal strategy and describe all of them (for which distance should John run on the left side, then in which angle needs he to start swimming) and show that these strategies are independent of the width d of the river. Compute the corresponding optimal time as well.

(4) Explain, how the condition (width) enters into the optimization problem. What happens if this condition is not fulfilled?

Solutions

(1) There are several (equivalent) ways to model this problem mathematically. Let us use the following variables: denote by $a \in [0, L]$ the distance traveled on the left side of the river, and by $\theta \in [0, \pi/2)$ the angle in which John starts to swim towards to the right side. We do not include here $\theta = \pi/2$ which would mean swimming parallel to the riversides, because in that case it would be smarter to just run (since $v_r > v_s$). It is also clear that one does not have to include swimming with negative angles, since that would be clearly not optimal. These two variables immediately determine how much is left to run on the right side, i.e. $L - a - d\tan(\theta)$. It is clear that one should impose $L - a - d\tan(\theta) \geq 0$, otherwise John would arrive beyond the finish point on the right side. So this already determines the feasible set of the problem, as

$$\Omega := \{(a, \theta) \in [0, L] \times [0, \pi/2) : L - a - d\tan(\theta) \geq 0\}.$$

The objective function is composed by the "three times lengths", while running on the left side, swimming in the river, and running on the right side, i.e. $f : \Omega \to \mathbb{R}$,

$$f(a, \theta) = \frac{a}{v_r} + \frac{d}{v_s \cos(\theta)} + \frac{L - a - d\tan(\theta)}{v_r} = \frac{L - d\tan(\theta)}{v_r} + \frac{d}{v_s \cos(\theta)},$$

since the time is the ratio of the distance and the speed.

(2) From the second equation in the definition of f one observes immediately, that it is independent of the variable a. So if there is an optimal strategy, meaning that John runs for a distance a on the left

side, then swims starting with and angle θ, another optimal strategy would be to run a bit further, or less, and start swimming with the very same angle. This basically means that the optimal strategy is depending only on the angle in which he starts to swim. This is because his speed is the same on both sides of the river. If that would not be the case, this invariance would not be present in the problem.

This in particular means that if there is an optimal strategy, there are infinitely many, so for sure one does not have uniqueness.

(3) We examine now the possible candidates for the optimal strategy. Let us work first in the interior of Ω, meaning that $a \in (0, L)$, $\theta \in (0, \pi/2)$ and $L - a - d\tan(\theta) > 0$. So by FONC, any local minimizer (a, θ) should satisfy $\nabla f(a, \theta) = 0$. Since f is independent of a, this means only $\partial_\theta f(a, \theta) = 0$, or equivalently

$$-\frac{d}{v_r \cos^2(\theta)} + \frac{d\sin(\theta)}{v_s \cos^2(\theta)} = 0,$$

from where one obtains

$$v_s = v_r \sin(\theta), \text{ or } \sin(\theta) = \frac{v_s}{v_r} = \frac{1}{2},$$

which immediately implies, since $\theta \in (0, \pi/2)$, that $\theta = \pi/6$. And one observes also that this strategy is independent of the width of the river, hence that term simplified above.

Since the function is just of one variable, one can check SONC and SOSC conditions just using

$$\partial^2_{\theta\theta} f(a, \theta) = d\left(-\frac{2\sin(\theta)}{v_r \cos^3(\theta)} + \frac{1}{v_s \cos(\theta)} + \frac{2\sin^2(\theta)}{v_s \cos^3(\theta)}\right) = d\left(-\frac{\sin(\theta)}{\cos^3(\theta)} + \frac{1}{\cos(\theta)} + \frac{2\sin^2(\theta)}{\cos^3(\theta)}\right),$$

which implies that

$$\partial^2_{\theta\theta} f(a, \pi/6) = d\left(-\frac{4}{3\sqrt{3}} + \frac{2}{\sqrt{3}} + \frac{4}{3\sqrt{3}}\right) = \frac{2d}{\sqrt{3}} > 0,$$

so actually $\pi/6$ is a strict local minimizer of $f(a, \cdot)$.

The optimal time corresponding to this strategy is

$$f(a, \pi/6) = \frac{500 - d\frac{\sqrt{3}}{3}}{2} + \frac{2d}{\sqrt{3}} = 250 + d\frac{\sqrt{3}}{2}.$$

Now it remained to examine the boundary cases. The boundary case, when $a = 0$ is similar, this produces the same optimal strategy. The other type of boundary cases consist of situations, when John arrives by swimming to the finish. This means that $L - a - d\tan(\theta) = 0$, hence he need to start swimming at an angle that is strictly smaller than θ. Let us see how the value of the objective function changes. Since we optimized this function independently of a on the interval $\theta \in (0, \pi/2)$, clearly the objective function achieves its global minimizer at $\theta = \pi/6$. Which means that taking angles that are smaller, would result worse objective function value.

On the other hand having $\theta = \pi/6$ will imply also that $L - d\tan(\theta) = L - \frac{d}{\sqrt{3}} > 0$, the the condition (width), so if that condition would be violated, the optimal strategy would not be $\pi/6$ anymore. This answers also (4).

To summarize, thanks to the condition (width), the above determined optimal strategies exist, moreover they can described as follows: run for an arbitrary distance a on the left side such that $0 \leq a \leq L - d\tan(\pi/6)$, then start swimming at the angle $\pi/6$, then run further on the right side, if you have not arrived yet to the finish.

Exercise 3 (10 points).

Let us consider the functions $f_\delta, g_\delta : \mathbb{R}^2 \to \mathbb{R}$ defined as

$$f_\delta(x_1, x_2) = 2x_1^2 + 2x_1x_2 + \delta x_2^2 - 2x_1 - x_2, \text{ and } g_\delta(x_1, x_2) = 2x_1^2 + 2x_1x_2 + \delta x_2^2$$

where $\delta \in \mathbb{R}$ is a given parameter. We aim to study some optimization problems involving f_δ and g_δ for different values of the parameter δ.

(1) Show that for any $\delta > 1/2$, f_δ has a unique minimizer on $\mathbb{R}^2$. Compute this minimizer in terms of δ. *Hint:* write down first and second order conditions for f_δ and study when $D^2 f_\delta$ becomes positive definite.

(2) Write f_δ as $\frac{1}{2}x^\top Q_\delta x - b \cdot x$, for some $Q_\delta \in \mathbb{R}^{2\times 2}$ and $b \in \mathbb{R}^2$ to be determined. Let $\delta > 1/2$ and $x^0 = (0 \;\; 0)^\top$, perform one iteration (compute x^1) of the gradient descent algorithm for f_δ with optimal step-size initialized with x^0. Compare x^1 to the optimal solution computed in (1). Is x^1 closer than x^0?

(3) Show that $g_{1/2}$ has infinitely many global minimizers, that lie on a line. Determine this line, and show that it corresponds to the eigenspace generated by the eigenvector corresponding to the smallest eigenvalue of $D^2 g$.

(4) Show that g_0 does not have either local minimizers or local maximizers.

Solution

We first write f_δ and g_δ as

$$f(x_1, x_2) = \frac{1}{2}x^\top Q_\delta x - b \cdot x, \text{ and } g_\delta = \frac{1}{2}x^\top Q_\delta x,$$

where

$$Q_\delta = \begin{pmatrix} 4 & 2 \\ 2 & 2\delta \end{pmatrix} \quad \text{and} \quad b = (2 \;\; 1)^\top.$$

Observe that Q_δ is a symmetric matrix.

(1) All the candidates for local minimizers should satisfy FONC, hence $Q_\delta x = b$. Moreover, $D^2 f(x) = Q_\delta$ and let us study whether this matrix is positive definite or not. We compute its eigenvalues (λ_1 and λ_2) for instance, which are roots of the equation $(4 - \lambda)(2\delta - \lambda) - 4 = 0$, or equivalently

$$\lambda^2 - (4 + 2\delta)\lambda + 8\delta - 4 = 0,$$

which should have real roots, because Q_δ is symmetric. Indeed,

$$\lambda_{1,2} = \frac{4 + 2\delta \pm \sqrt{(4 + 2\delta)^2 - 4(8\delta - 4)}}{2} = \frac{4 + 2\delta \pm 2\sqrt{(\delta - 2)^2 + 4}}{2}, \tag{1}$$

and from the first equation one can see immediately, that if $8\delta - 4 > 0$, i.e. $\delta > 1/2$ then the term under the square root is always smaller than the one in the front of it, hence both eigenvalues are positive, which means also that Q_δ is positive definite, hence invertible, so the only candidate will be a strict global minimizer of f_δ on $\mathbb{R}^2$. The same can we achieved using Sylvester's criterion on the leading principal minors. This minimizer is given by

$$(x_1^\delta \;\; x_2^\delta)^\top = Q_\delta^{-1} b = (1/2 \;\; 0).$$

(2) The gradient descent algorithm with optimal step size (that can be used in our situation, since $Q_\delta > 0$) reads as

$$x^{k+1} = x^k - \alpha_k(Q_\delta x^k - b), \quad \text{with } \alpha_k = \frac{\|Q_\delta x^k - b\|^2}{(Q_\delta x^k - b)^\top Q_\delta (Q_\delta x^k - b)},$$

provided, we did not reach yet that the optimizer with x^k, hence $Q_\delta x^k \neq b$. Let x^0 be the zero vector. And compute

$$\alpha_0 = \frac{\|b\|^2}{b^\top Q_\delta b} = \frac{5}{24+2\delta},$$

so

$$x^1 = \frac{5}{24+2\delta}(2 \quad 1)^\top.$$

The initial distance square between x^0 and the optimal solution is $1/4$. Now, let us compute

$$\|x^1 - (1/2 \quad 0)^\top\|^2 = \frac{25}{(24+2\delta)^2} + \frac{(2+\delta)^2}{(24+2\delta)^2} = \frac{\delta^2+4\delta+29}{(24+2\delta)^2},$$

which is clearly much more smaller than $1/4$.

(3) Clearly, for $\delta = 1/2$, using the formula (1), one sees that the eigenvalues of $Q_{1/2}$ will be $\lambda_1 = 0$ and $\lambda_2 = 5$, hence this matrix is only positive semi-definite, which means that one cannot use for instance SOSC to decide, whether the candidates for the minimizers of $g_{1/2}$ are indeed minimizers or not.

The candidates, in this case as well have to satisfy FONC, meaning that

$$Q_{1/2}x = 0.$$

More precisely, this means that they solve the linear system

$$2x_1 + x_2 = 0, \tag{2}$$

so they lie on a line. Actually all these points satisfying the above equation are eigenvectors associated to $\lambda_1 = 0$, so this is the eigenspace associated to this eigenvalue (it is also the kernel of $Q_{1/2}$.) By the fact that $Q_{1/2}$ is positive semi-definite (or by Rayleigh's inequality), one has that

$$g_{1/2}(x) = x^\top Q_{1/2}x \geq 0,$$

on the other hand $g_{1/2}(x_1, x_2) = 0$, for all vectors characterized by (2), hence all these will be global minimizers.

(4) Let us assume that (a, b) is a local extremizer of $g_0(x_1, x_2) = 2x_1^2 + 2x_1x_2$. Then, there should exists a small ball around this point, in this (a, b) achieves the best value. Take an arbitrary direction $e = (e_1, e_2)$, and compute the directional derivative of the function at (a, b) this direction, that is

$$\partial_e g_0(a, b) = \nabla g_0(a, b) \cdot e = e_1(4a + 2b) + 2ae_2.$$

Now suppose that $a \neq 0$. Then choose $e_1 = 0$ and choosing $e_2 = a$ and $e_2 = -a$, one finds that the function is increasing in one of the directions and decreasing in the other, hence this point cannot be either a local minimizer or a maximizer.

If $a = 0$ and $b \neq 0$, choose e_2 arbitrary and $e_1 = b$ and $e_2 = -b$, which gives the same contradiction.

This means that the only possibility for local extremizer is $(0, 0)$, where all the directional derivatives vanish, so we need to use the definition. For this, take $\varepsilon \neq 0$ to be a small number and compute

$$g_0(\varepsilon, \varepsilon) = 4\varepsilon^2 > 0, \quad \text{and} \quad g_0(\varepsilon, -2\varepsilon) = -2\varepsilon^2 < 0,$$

hence $(0, 0)$ cannot be a local extremizer either.