Math 164-1: Optimization

Instructor: Alpár R. Mészáros

Midterm, November 16, 2015

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.
- **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. You may loose points in the lack of justification of your answers.
- Theorems from the lectures 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.
- I wish you success!

Problem	Score
Exercise 1	
Exercise 2	
Exercise 3	
Total	

Exercise 1 (15 points).

(1) Describe Newton's method in 1D to approximate the roots and local optima (minimizers/maximizers) of C^1, respectively C^2 functions.

(2) We aim now to use Newton's method to approximate the maximizer of $f : [-\pi/2, \pi/2] \to \mathbb{R},\ \ f(x) = \cos(x)$. What are the optimizers of this function and what is the optimal value? Are the optimizers local or global? What about the uniqueness of the optimizers? Why is this problem equivalent to look for roots of the function $g : [-\pi/2, \pi/2] \to \mathbb{R},\ \ g(x) = \sin(x)$?

(3) Initiate the algorithm (for g) with $x_0 \in [-\pi/2, \pi/2]$ such that $x_0 > 0$. Show that $x_1 < 0$. Show that in general $x_k \cdot x_{k+1} \leq 0$, for all $k \in \mathbb{N}$.

Hint: study the sign and growth properties of the function $h_1 : [-\pi/2, \pi/2] \to \mathbb{R},\ \ h_1(x) = x - \tan(x)$.

(4) Show that if $x_k \in [-\pi/4, \pi/4]$ one has that $|x_{k+1}| < |x_k|$.

Hint: study the sign and growth properties of the function $h_2 : [-\pi/4, \pi/4] \to \mathbb{R},\ \ h_2(x) = 2x - \tan(x)$.

(5) Show that the algorithm converges for all $x_0 \in [-\pi/4, \pi/4]$.

(6) Show that the order of convergence (if it exists) is at least 2. Is the interval in (5) optimal, i.e. could we choose x_0 outside of this interval (but of course not outside of $[-\pi/2, \pi/2]$) and still have the convergence? Justify your answer!

Hint for the convergence proof and order of convergence: a possible way is to use a second order (exact, i.e. with reminder term) Taylor expansion for $\sin(0)$ around x_k, then use the construction of the sequence x_k and try to give an upper bound for the term $\frac{|\sin(\xi_k)|}{2|\cos(x_k)|}$, where ξ_k is between 0 and x_k. Other correct proofs are also accepted!

Solution:

(1) For finding roots of a nonlinear equation like $f(x) = 0$, where $f \in C^1(\mathbb{R})$ reads as follows. Pick an initial guess x_0 (close enough to the root) and construct the sequence $(x_k)_{k\geq 0}$ with the recursive relation

$$x_{k+1} = x_k - f(x_k)/f'(x_k),$$

provided $f'(x_k) \neq 0$, otherwise the tangent line is already parallel to the $x-$axes, so there it is not possible to construct the next iteration. Draw a picture for illustration.

Local extrema of $f \in C^2$ function (without constraints) satisfy the first order necessary optimality condition, that is $f'(x) = 0$. Hence the problem reduces to find the roots of this nonlinear equation. By the previous reasoning the sequence can be constructed (after choosing an initial guess x_0) as

$$x_{k+1} = x_k - f'(x_k)/f''(x_k),$$

provided $f''(x_k) \neq 0$.

(2) The unique global maximizer is clearly $x^* = 0$ (with the maximal value $f(0) = 1$) that is an interior point satisfying both the first order necessary ($-\sin(0) = 0$) and second order sufficient ($-\cos(0) = -1 < 0$) conditions. At both boundary point f is 0, so these cannot be global maximizers, they are both local minimizers and are not unique. Since x^* is an interior point, one has that $f'(x^*) = -\sin(x^*) = 0$, hence the problem is equivalent finding the root of sin on the given interval.

(3) The construction of Newton's algorithm yields

$$x_{k+1} = x_k - \tan(x_k).$$

So let us follow the hint and study the properties of h_1. Since $h_1'(x) = 1-1/\cos^2(x) \leq 0,\ \forall x \in (-\pi/2, \pi/2)$, the function is strictly decreasing. This means in particular that $0 = h_1(0) > h_1(x_0) = x_1$, for all $x_0 \in (0, \pi/2)$.

By the same idea, using the fact that h_1 is an odd function, if $x_0 < 0$ one has that $0 < h_1(x_0) = x_1$. This holds true for any x_k and x_{k+1} two consecutive iterations, hence $x_k \cdot x_{k+1} < 0$.

(4) In the previous point we have shown that x_k and x_{k+1} have opposite signs. Without the loss of generality, let us suppose that $x_k > 0$. The inequality to be shown reads as

$$-x_k < x_{k+1} < x_k,$$

and the second inequality holds true by the construction and (3). Hence it remains to show that $-x_k < x_k - \tan(x_k)$, i.e. $0 < 2x_k - \tan(x_k)$. To show this inequality, as the hint suggests, we study the properties of h_2. In particular $h_2'(x) = 2 - 1/\cos^2(x)$. Since on the interval $[-\pi/4, \pi/4]$ $\cos^2(x) \geq 1/2$, one has that $h_2'(x) \geq 0$ on this interval, hence it is strictly increasing. This implies in particular that $0 = h_2(0) < h_2(x)$ for all $x \in (0, \pi/4)$, and setting $x = x_k > 0$ we have proved the inequality that we wanted.

(5) and (6) Since the sequence of positive real numbers $(|x_k|)_{k\geq 0}$ is decreasing by (4) for any initial guess $x_0 \in [-\pi/4, \pi/4]$ and it is bounded from below, it is convergent. By the continuity of the absolute value this means that $(x_k)_{k\geq 0}$ and $(x_{k+1})_{k\geq 0}$ converge to the same limit, to some $\ell \in [-\pi/4, \pi/4]$. By the continuity of the tangent function, we can pass to the limit in the recursive relation to obtain that

$$\ell = \ell - \tan(\ell),$$

from where one obtains that $\ell = 0$.

To obtain a rate of convergence for the convergence, one has to perform a finer analysis. We so the same Taylor's expansion technique as during the lectures, i.e.

$$0 = \sin(0) = \sin(x_k) - \cos(x_k)x_k - \frac{1}{2}\sin(\xi_k)x_k^2,$$

where ξ_k is a real between 0 and x_k. Dividing by $\cos(x_k) \neq 0$ the equality and using the recursive relation for x_k and x_{k+1} one obtains (passing also to absolute values) that

$$|x_{k+1} - 0| = \frac{1}{2}\frac{|\sin(\xi_k)|}{|\cos(x_k)|}|x_k - 0|^2.$$

Since $x_k \in [-\pi/4, \pi/4]$ (and ξ_k is between 0 and x_k) one has that

$$|\sin(\xi_k)| \leq \frac{\sqrt{2}}{2} \quad \text{and} \quad \cos(x_k) \geq \frac{\sqrt{2}}{2},$$

hence

$$\frac{1}{2}\frac{|\sin(\xi_k)|}{|\cos(x_k)|} \leq \frac{1}{2}.$$

Iterating, one obtains that $|x_{k+1}| \leq (1/2)^{2^k-1}x_0^2$, hence one has the convergence to 0 of $(x_k)_{k\geq 0}$ and the order of convergence, if it exists, it is at least 2.

From this idea, it is clear that one can increase slightly the interval for the initial guess, and still have convergence. One has to ensure only that

$$\frac{|\sin(\xi_k)|}{\cos(x_k)} < 2,$$

which can be achieved for a larger interval than $[-\pi/4, \pi/4]$, since both functions are continuous.

Exercise 2 (7 points)**.**

We aim to solve numerically the following system of linear equations for $x = [x_1 \ \ x_2]^T \in \mathbb{R}^2$:

$$\begin{bmatrix} 0 & 1 \\ 2 & 0 \\ 1 & 2 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \tag{E}$$

We use the notation $Ax = b$ for (E) in the followings.

(1) Why does a solution for (E) exist? We claim that finding a solution of (E) (if there is any) is equivalent to the problem of finding a minimizer of the function

$$\mathbb{R}^2 \ni x \mapsto \frac{1}{2}\|Ax - b\|^2. \tag{F}$$

Why is this the case?

(2) We will use the conjugate gradient method to solve numerically the system (E). Show that the function in (F) is a quadratic one, generated by a positive definite, symmetric matrix from $\mathbb{R}^{2\times 2}$.

(3) Using the initial guess $x^0 = [0 \ \ 0]^T$, develop the steps of the conjugate gradient algorithm and show that it converges in at most 2 steps. Check if you have found indeed a solution of the system (E). If it is the case, is it unique? Why?

Solution:

(1) Just by simple correspondence the only solution of the system is $(x_1, x_2) = (1, 1)$. In particular the solution exists. Since the function defined in (F) is nonnegative, its minimum (i.e. 0) is attained whenever $Ax = b$, i.e. if one has found a solution to the system.

(2) and (3) Actually these points are not considered for our current midterm, so I don't want to confuse you with the solution.

Exercise 3 (4 points)**.**

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a C^1 convex function. We suppose that the function has a unique minimizer and construct the following algorithm to approximate it:

$$x^{k+1} = x^k - \alpha_k B \nabla f(x^k),$$

where

$$B := \begin{bmatrix} b & 0 \\ 0 & 1 \end{bmatrix}$$

with $b \in \mathbb{R}$ and

$$\alpha_k := \text{argmin}_{\alpha \in \mathbb{R}} f(x^k - \alpha B \nabla f(x^k)). \tag{1}$$

At some x^k during the algorithm, let us suppose to obtain $\nabla f(x^k) = [1 \;\; 2]^\top$. What is the maximal range for $b \in \mathbb{R}$ in this case that implies that $\alpha_k \geq 0$ (where α_k is given in (1))? For all b in the found range, is the matrix B positive definite?

Hint: use the fact that the graph of a convex function lies always above its tangent plane at any point (in particular at x^k) and use the construction of x^{k+1} and α_k.

Solution:

First, by the fact that f is convex and C^1 on $\mathbb{R}^2$ one has that

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x), \;\; \forall x, y \in \mathbb{R}^2.$$

Secondly, at x^k one has that $B \nabla f(x^k) = [b \;\; 2]^\top$. Now writing the above convexity inequality for $y = x^{k+1}$ and x^k one has that

$$f(x^{k+1}) \geq f(x^k) - \alpha_k [1 \;\; 2] B [1 \;\; 2]^\top = f(x^k) - \alpha_k (b + 4).$$

On the other hand by the construction of α_k one has that $f(x^{k+1}) - f(x^k) \leq 0$, which with the previous inequality implies that one should have necessarily that $\alpha_k (b + 4) \geq 0$. Hence the maximal range for b that implies $\alpha_k \geq 0$ is $b \geq 4$.

Since the eigenvalues of B are b and 1, if one chooses $b \in [-4, 0]$, B is not positive definite.