Midterm test

Closed book. Attempt all questions. Calculators permitted. 13:00-13:50
Please write your name and student identity number on the front page.

(1) Determine the order of convergence of each of the following sequences (if they converge at all). You may assume $k \geq 1$.

$$\text{(a)}\;\; x_k = \frac{1}{\sqrt{k}}, \quad \text{(b)}\;\; x_k = 1 + (0.2)^{3^k}, \quad \text{(c)}\;\; x_k = k^{-k}, \quad \text{(d)}\;\; x_k = 1$$

[4 marks]

(2) Consider the function on $\mathbb{R}^2$, $f(\boldsymbol{x}) = (2x_1 + x_2^2)^2$. Show that $\boldsymbol{p} = (-1, 0)^\top$ is a descent direction at $\boldsymbol{x}_0 = (0, 1)^\top$, and find a step length α that minimizes $f(\boldsymbol{x}_0 + \alpha p)$.

[4 marks]

(3) Determine, with justification, which of the following functions is convex ($\ln(x)$ refers to the natural logarithm).

(a) $f(x) = \ln(x)$ for $x > 0$;

(b) $f(x) = \dfrac{1}{x}$ for $x > 0$;

(c) $f(x, y, z) = z^2 - x^2 - y^2$ for $x \in \mathbb{R}$;

(d) $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_1 + \|\boldsymbol{x}\|_\infty$.

You may used criteria for convexity seen in the lecture and problem sessions. [4 marks]

(4) Consider the following linear programming problem

$$\begin{aligned} \text{maximize} \quad & x_1 - x_2 \\ \text{subject to} \quad & x_1 \leq 1 \\ & x_2 \leq 2 \\ & 2x_1 + x_2 \geq 2 \end{aligned}$$

(a) Determine the vertices of the polyhedron of feasible points;

(b) Find an optimizer and the optimal value;

(c) Write down the dual to this problem.

[4 marks]

(5) Consider the function

$$f(x, y) = \sqrt{1 + x^2 + y^2}$$

By computing the gradient and Hessian, show that this function is convex and determine the unique minimum. Write down the form of one iteration of Newton's method for this function. [4 marks]

Solution (1)

(a) The sequence converges to 0. We have the identity

$$x_{k+1} = \frac{1}{\sqrt{k+1}} = \sqrt{\frac{k}{k+1}}\frac{1}{\sqrt{k}} = \sqrt{\frac{k}{k+1}}x_k,$$

which means that for any fixed constant $c < 1$ there is a k such that $1 > \sqrt{k/(k+1)} > c$, and therefore $x_{k+1} > cx_k$. It follows that the sequence does not converge linearly (or to any higher order).

(b) The sequence converges to 1. We have

$$|x_{k+1} - 1| = (0.2)^{3^{k+1}} = \left((0.2)^{3^k}\right)^3 = |x_k - 1|^3,$$

so that the convergence is cubic.

(c) The sequence converges to 0. We can write $k^{-k} = 2^{-k\log(k)}$. For $k \geq 1$,

$$x_{k+1} = \frac{1}{(k+1)^{k+1}} = \frac{1}{(k+1)^k \cdot (k+1)} = \frac{1}{2^{k\log(k+1)}(k+1)} \leq \frac{1}{k+1}x_k,$$

which shows that the sequence converges superlinearly.

(d) The sequence converges to 1. Moreover, $|x_{k+1} - 1| = 0 \leq M|x_k - 1|^p$ for any $M > 0$ and $p > 0$, so this sequence convergences to any order.

Solution (2) The gradient is

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4(2x_1 + x_2^2) \\ 4x_2(2x_1 + x_2^2) \end{pmatrix}.$$

At $\boldsymbol{x}_0 = (0,1)^\top$, $\nabla f(0,1) = (4,4)^\top$. The direction $\boldsymbol{p} = (-1,0)^\top$ is a descent direction, if $\langle \nabla f(\boldsymbol{x}_0), \boldsymbol{p}\rangle < 0$. In our case, $\langle \nabla f(\boldsymbol{x}_0), \boldsymbol{p}\rangle = -4 < 0$. The optimal step length along $\boldsymbol{p}$ is the minimizer of

$$f(\boldsymbol{x}_0 + \alpha\boldsymbol{p}) = (1 - 2\alpha)^2.$$

Computing the derivative and setting it to zero, $2\alpha - 1 = 0$, we get the optimal step length $\alpha = 1/2$.

Solution (3)

(a) The function is not convex. The derivative is $1/x$, which for $x > 0$ is positive. The second derivative is $-1/x^2 < 0$.

(b) The function is convex. The second derivative is $2/x^3 > 0$.

(c) This function is not convex. The Hessian is given by

$$\begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

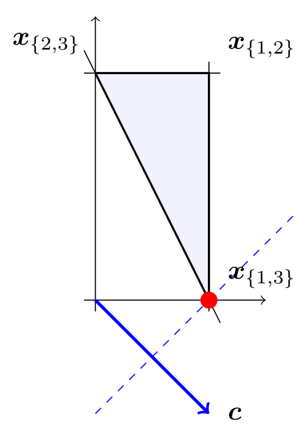
which is not positive definite.

(d) This function is, as the sum of two norms, convex. Precisely, for $\lambda \in [0,1]$,

$$\begin{aligned} \|\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|_1 + \|\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|_\infty &\leq \lambda\|\boldsymbol{x}\|_1 + (1-\lambda)\|\boldsymbol{y}\|_1 + \lambda\|\boldsymbol{x}\|_\infty + (1-\lambda)\|\boldsymbol{y}\|_\infty \\ &\leq \lambda(\|\boldsymbol{x}\|_1 + \|\boldsymbol{x}\|_\infty) + (1-\lambda)(\|\boldsymbol{y}\|_1 + \|\boldsymbol{y}\|_\infty). \end{aligned}$$

Solution (4) The matrix and the vectors associated to this problem are

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -2 & -1 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \ \boldsymbol{c} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$



(a) To determine the vertices, we can read them off the diagram and then verify analytically that these are indeed the vertices. We get:

$$\boldsymbol{x}_{\{1,2\}} = (1,2)^\top, \ \boldsymbol{x}_{\{1,3\}} = (1,0)^\top, \ \boldsymbol{x}_{\{2,3\}} = (0,2)^\top.$$

Multiplying $\boldsymbol{A}$ with each of these vertices we get

$$\boldsymbol{A}\boldsymbol{x}_{\{1,2\}} = \begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix}, \ \boldsymbol{A}\boldsymbol{x}_{\{1,3\}} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \ \boldsymbol{A}\boldsymbol{x}_{\{2,3\}} = \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix}.$$

We see that each of the vertex candidates satisfies all three inequalities, with three as equalities, and we therefore have vertices.

(b) The optimal values can be found among the vertices:

$$\langle \boldsymbol{c}, \boldsymbol{x}_{\{1,2\}}\rangle = -1,\ \langle \boldsymbol{c}, \boldsymbol{x}_{\{1,3\}}\rangle = 1,\ \langle \boldsymbol{c}, \boldsymbol{x}_{\{2,3\}}\rangle = -2.$$

As the picture shows, the optimal value is attained at $\boldsymbol{x}_{\{1,3\}}$ and the optimal value is 1.

(c) The dual of the problem is

$$\text{minimize} \quad \langle \boldsymbol{b}, \boldsymbol{y}\rangle \qquad \text{subject to} \quad \boldsymbol{A}^\top \boldsymbol{y} = \boldsymbol{c},\ \boldsymbol{y} \geq \boldsymbol{0}.$$

Applying this to our data,

$$\begin{aligned} \text{minimize} \quad & y_1 + 2y_2 - 2y_3 \\ \text{subject to} \quad & y_1 - 2y_3 = 1 \\ & y_2 - y_3 = -1 \\ & y_i \geq 0, \quad 1 \leq i \leq 3. \end{aligned}$$

Solution (5) We first compute the gradient and the Hessian of this function.

$$\nabla f(x_1, x_2) = \frac{1}{\sqrt{1+x^2+y^2}} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{1}$$

$$\nabla^2 f(x_1, x_2) = \frac{1}{(1+x^2+y^2)^{3/2}} \begin{pmatrix} 1+y^2 & -xy \\ -xy & 1+x^2 \end{pmatrix}.$$

We have a stationary point at $(0,0)$ which is a minimizer, as the function can never fall below $f(0,0) = 1$. This means that the Hessian is positive definite at $(0,0)$. There are various ways of verifying that the Hessian is positive definite everywhere, and the function therefore convex. One is direct verification:

$$\boldsymbol{v}^\top \nabla^2 f(x,y) \boldsymbol{v} = v_1^2(1+y^2) - 2v_1v_2xy + v_2^2(1+x^2) = v_1^2 + v_2^2 + (v_1y_1 - v_2x_2)^2 > 0.$$

Newton's method starts with a point $(x_{(0)}, y_{(0)})$, and then for every $k \geq 0$, first solves the system of equations

$$\frac{1}{(1+x_{(k)}^2+y_{(k)}^2)^{3/2}} \begin{pmatrix} 1+y_{(k)}^2 & -x_{(k)}y^{(k)} \\ -x_{(k)}y_{(k)} & 1+x_{(k)}^2 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \frac{1}{\sqrt{1+x_{(k)}^2+y_{(k)}^2}} \begin{pmatrix} x_{(k)} \\ y_{(k)} \end{pmatrix},$$

and then computes

$$(x_{(k+1)}, y_{(k+1)}) = (x_{(k)}, y_{(k)}) + (\Delta x, \Delta y).$$