

Midterm test - Solutions

*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).

(a) $x_k = \frac{1}{k!}$, (b) $x_k = 1 + (0.3)^{2^k}$, (c) $x_k = 2^{-k}$, (d) $x_k = 1/k$
[4 marks]

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

(3) Determine, with justification, which of the following sets is convex.

- (a) $\{(x, y) : x > 1, y > \log(x)\}$;
 (b) $\{(x, y) : x > 0, y > 0, xy < 1\}$;
 (c) $\{(x, y, 1) : x^2 + y^2 \leq 2\}$;
 (d) $\{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_\infty + \|\mathbf{x}\|_1 \leq 1\}$.
 [4 marks]

(4) Consider the following linear programming problem

$$\begin{aligned} &\text{maximize} && x_1 + 2x_2 \\ &\text{subject to} && x_1 \geq 0 \\ &&& x_1 \leq 1 \\ &&& x_1 + 2x_2 \leq 2 \\ &&& x_1 + x_2 \geq 1 \end{aligned}$$

- (a) Sketch the feasible set and determine the vertices of the polyhedron of feasible points from the diagram or from the inequalities;
 (b) Find an optimizer and the optimal value;
 [4 marks]

(5) Consider the function

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Formulate Newton's method for finding a minimizer. By inspecting the gradient, show that $(1, 1)^\top$ is the only local minimizer of this function.
[4 marks]

Solution (1)

(a) The sequence converges to 0. We have

$$x_{k+1} = \frac{1}{(k+1)!} = \frac{1}{k+1} x_k,$$

so that $\lim_{k \rightarrow \infty} \frac{x_{k+1}}{x_k} = 0$. The convergence is *superlinear*.

(b) The sequence converges to 1. We have

$$|x_{k+1} - 1| = (0.3)^{2^{k+1}} = \left((0.3)^{2^k} \right)^2 = |x_k - 1|^2,$$

so that the convergence is quadratic.

(c) The sequence converges to 0. Moreover,

$$x_{k+1} = \frac{1}{2^{k+1}} = \frac{1}{2} x_k,$$

so that the sequence converges linearly.

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

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

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

Solution (2) The gradient is

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

At $\mathbf{x}_0 = (0, 1)^\top$, $\nabla f(0, 1) = (0, 2)^\top$. The direction \mathbf{p} is a descent direction, if $\langle \nabla f(\mathbf{x}_0), \mathbf{p} \rangle < 0$. In our case, $\langle \nabla f(\mathbf{x}_0), \mathbf{p} \rangle = -2 < 0$. The optimal step length along \mathbf{p} is the minimizer of

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

The minimum is the minimum of the quadratic function $\alpha^2 + 1 - \alpha$. Computing the derivative and setting it to zero, $2\alpha - 1 = 0$, we get the optimal step length $\alpha = 1/2$.

Solution (3)

(a) This set is not convex, as becomes clear by looking at the region that this set defines. One way to show this is to find a line segment between two points in the set that contains a point not in the set. A natural choice would be to take the tangent line at a point, say $(2, \log(2))^\top$ (a point not in the set) and take two points

slightly to the left and to the right of this point along the tangent. The tangent line is given by the equation

$$y = f(x) = \frac{1}{2}x + (\log(2) - 1).$$

For example, $(1.9, f(1.9))$ and $(2.1, f(2.1))$ are two points on the set (check the differences $f(1.9) - \log(1.9)$ and $f(2.1) - \log(2.1)$), for which the line segment joining them contains $(2, \log(2))$. It follows that the set is not convex.

- (b) This set is not convex, as should become clear when looking at it as the region under a hyperbola. For example, the line segment joining $(0.1, 9)^\top$ and $(9, 0.1)^\top$ contains the point $(4, 5.1)^\top$, which is not part of the set.
- (c) This set is convex: it describes a circle of radius 2 suspended at height 1 in the z direction. To see this more formally, consider points $\mathbf{p} = (x_1, y_1, 1)^\top$ and $\mathbf{q} = (x_2, y_2, 1)^\top$ in the set, and consider a convex combination $\lambda\mathbf{p} + (1 - \lambda)\mathbf{q}$, with $\lambda \in (0, 1)$. Then

$$\begin{aligned} & (\lambda x_1 + (1 - \lambda)x_2)^2 + (\lambda y_1 + (1 - \lambda)y_2)^2 \\ &= \lambda^2(x_1^2 + x_2^2) + (1 - \lambda)^2(y_1^2 + y_2^2) + 2\lambda(1 - \lambda)(x_1x_2 + y_1y_2) \\ &\stackrel{(*)}{\leq} \lambda^2(x_1^2 + x_2^2) + (1 - \lambda)^2(y_1^2 + y_2^2) + 2\lambda(1 - \lambda)\sqrt{x_1^2 + y_1^2} \cdot \sqrt{x_2^2 + y_2^2} \\ &\leq \lambda^2(x_1^2 + x_2^2) + (1 - \lambda)^2(y_1^2 + y_2^2) + 4\lambda(1 - \lambda) \\ &= (\lambda 2 + (1 - \lambda)2)^2 = 4, \end{aligned}$$

so that $\lambda\mathbf{p} + (1 - \lambda)\mathbf{q}$ also is in that set. You can also argue with the convexity of the circle and then use the property that a linear image of a convex set is convex.

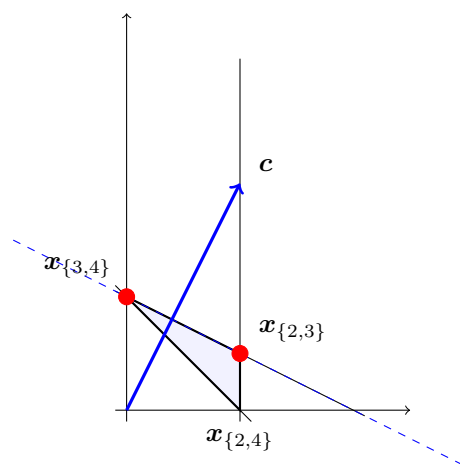
- (d) This set is convex. For $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$ with \mathbf{x} and \mathbf{y} such that $\|\mathbf{x}\|_\infty + \|\mathbf{x}\|_1 \leq 1$ and $\|\mathbf{y}\|_\infty + \|\mathbf{y}\|_1 \leq 1$ we get, using the norm properties,

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

This shows that the set is convex.

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

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



- (a) To determine the vertices, we can use the diagram to reduce the number of minors we need to consider. From the diagram we see that P is determined by equations 2, 3, 5. The minors are

$$\mathbf{A}_{\{2,3\}} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{A}_{\{2,4\}} = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}, \quad \mathbf{A}_{\{3,4\}} = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix},$$

and the subsets of the vector \mathbf{b} are

$$\mathbf{b}_{\{2,3\}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{b}_{\{2,4\}} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{b}_{\{3,4\}} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Solving the three systems $\mathbf{A}_I \mathbf{x} = \mathbf{b}_I$, we get the solutions (which may also be read off the diagram)

$$\mathbf{x}_{\{2,3\}} = (1, 0.5)^\top, \quad \mathbf{x}_{\{2,4\}} = (1, 0)^\top, \quad \mathbf{x}_{\{3,4\}} = (0, 1)^\top.$$

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

$$\langle \mathbf{c}, \mathbf{x}_{\{2,3\}} \rangle = 2, \quad \langle \mathbf{c}, \mathbf{x}_{\{2,4\}} \rangle = 1, \quad \langle \mathbf{c}, \mathbf{x}_{\{3,4\}} \rangle = 2.$$

As the picture shows, the optimal value is attained at two of the vertices.

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

$$\nabla f(x_1, x_2) = \begin{pmatrix} -400x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}, \quad (1)$$

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}.$$

Newton's method starts with a point $(x_1^{(0)}, x_2^{(0)})$, and then for every $k \geq 0$, first solves the system of equations

$$\begin{pmatrix} 1200(x_1^{(k)})^2 - 400x_2^{(k)} + 2 & -400x_1^{(k)} \\ -400x_1^{(k)} & 200 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = \begin{pmatrix} -400x_1^{(k)}(x_2^{(k)} - (x_1^{(k)})^2) - 2(1 - x_1^{(k)}) \\ 200(x_2^{(k)} - (x_1^{(k)})^2) \end{pmatrix}$$

and then computes

$$(x_1^{(k+1)}, x_2^{(k+1)}) = (x_1^{(k)}, x_2^{(k)}) + (\Delta x_1, \Delta x_2).$$

Since $\nabla f(1, 1) = \mathbf{0}$, the point $(1, 1)^\top$ is a stationary point. Looking at the gradient, we see that if $(x_1, x_2)^\top$ is a stationary point, then necessarily $x_2 = x_1^2$ (from the second equation) and $x_1 = 1$ (from the first equation). This gives only the choice $(1, 1)^\top$ for stationary point. The function itself satisfies $f(1, 1) = 0$ and is positive for all other values $(x_1, x_2)^\top$, from which we deduce that $(1, 1)$ is indeed a local (and global) minimum and not a saddle point or maximum.