Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimisation I Spring 2017

Solutions to exercise set 1

1 **a)** Remembering that, by definition,

$$\liminf_{k\to\infty} y_k = \lim_{k\to\infty} \inf_{n\geq k} y_n,$$

this inequality follows from three properties: (1) the infimum—think minimum—of two sets A and B is greater than or equal to the sum of the infimum of each set:

$$\inf(A+B) \geq \inf A + \inf B,$$

where $A+B = \{a+b : a\in A \text{ and } b \in B\}$; (2) the limit operation is linear:

$$\lim_{k\to\infty}(a_k + Cb_k) = \lim_{k\to\infty} a_k + C \lim_{k\to\infty} b_k;$$

and (3) taking limits preserve non-strict inequalites ($\leq$ or $\geq$): if $a_k \leq b_k$ for all $k\in\mathbb{N}$, then $\lim_k a_k \leq \lim_k b_k$, provided the limits exist. (Try to prove these properties if they are not clear to you.) Hence,

$$\begin{aligned}\liminf_{k\to\infty} y_k + \liminf_{k\to\infty} z_k &= \lim_{k\to\infty}\left(\inf_{m\geq k} y_m + \inf_{n\geq k} z_n\right)\\ &\leq \lim_{k\to\infty}\left(\inf_{\ell\geq k}(y_\ell + z_\ell)\right) = \liminf_{k\to\infty}(y_k+z_k),\end{aligned}$$

where ℓ is just a common index.

Strict inequality occurs for example if $y_k = (-1)^k$ and $z_k = (-1)^{k+1}$. Then $y_k + z_k = 0$ for all k, which yields

$$\liminf_{k\to\infty} y_k = -1 = \liminf_{k\to\infty} z_k \qquad \text{and} \qquad \liminf_{k\to\infty}(y_k+z_k) = 0.$$

b) For any k it is true that

$$y_k^i \leq \sup_{i\in I} y_k^i$$

for all i by definition of supremum, and $\sup_{i\in I} y_k^i$ is just a real sequence indexed by k—let us call it x_k to make notation easier. Since $y_k^i \leq x_k$ for all k (and i), it follows that

$$\liminf_{k\to\infty} y_k^i \leq \liminf_{k\to\infty} x_k.$$

(Try to prove this yourself, or ask for help.) As this inequality holds for all i, we can take the supremum and obtain

$$\sup_{i\in I}\liminf_{k\to\infty} y_k^i \leq \liminf_{k\to\infty} x_k = \liminf_{k\to\infty}\Big(\sup_{i\in I} y_k^i\Big).$$

2 Let $x \in \mathbb{R}^n$ be arbitrary, and $(x_k)_{k\in\mathbb{N}} \subset \mathbb{R}^n$ be any sequence converging to x. We must show that $\liminf_{k\to\infty} f(x_k) \geq f(x)$, and this is a direct consequence of exercise 1 b):

$$\liminf_{k\to\infty} f(x_k) = \liminf_{k\to\infty}\Big(\sup_{i\in I} f_i(x_k)\Big) \geq \sup_{i\in I}\Big(\liminf_{k\to\infty} f_i(x_k)\Big) \geq \sup_{i\in I} f_i(x) = f(x).$$

Indeed, the first inequality comes from exercise 1 b) with $y_k^i = f_i(x_k)$, while the latter is a result of the lower semi-continuity of the f_i's (and properties of supremum).

3 Note first that a finite sum of lower semi-continuous (l.s.c.) functions is l.s.c., and furthermore, that continuous functions especially are l.s.c. (try to prove both facts).

a) The polynomial $x \mapsto x^4 - 20x^3$ is continuous, and hence, also l.s.c. Moreover, $x \mapsto \sin(kx)$ is continuous (and l.s.c.) for all $k \in \mathbb{N}$, so that $\sup_{k\in\mathbb{N}} \sin(kx)$ is l.s.c. by exercise 2. Thus f is l.s.c.

Since, sin is a bounded function, it follows that f grows like x^4 as $|x| \to \infty$. In particular, f is coercive, and we deduce the existence of a global minimiser on $\mathbb{R}$ by Theorem 9 in the note "Minimisers of Optimization Problems."

b) Since g is continuous, it is l.s.c. Coercivity fails, because $g(x) \to 0$ as $x \to -\infty$. However, g is bounded from below, and strictly increasing for sufficiently large x (hence, these points cannot be minimisers). As, for example, $g(-1) < 0 = g(-\infty)$, it follows that g attains a global minimiser in $\mathbb{R}$.

c) Again, h is continuous and therefore also l.s.c. It is not coercive, which can be seen by letting $\|(x_1, x_2)\| \to \infty$ with $x_1 \equiv 0$, yielding $h(x_1, x_2) \equiv 0$ for this particular sequence. Due to quadratic terms, h is bounded from below by 0, and we observe that all points on the x_1-axis are global minimisers with function value equal to 0.

4 Effectively, the exponential map $x \mapsto \exp(-x^2)$ is always positive, decreasing, and goes to 0 as $|x| \to \infty$ much faster than any polynomial goes to infinity. Thus for f to have a global minimum, we need that $x^2 - 2x + c \leq 0$ for some (finite) $x \in \mathbb{R}$. This is also sufficient, as f is continuous and thus will have a minimum in a compact interval containing that particular x. By factoring,

$$x^2 - 2x + c = \left(x - 1 - \sqrt{1-c}\right)\left(x - 1 + \sqrt{1-c}\right) \leq 0,$$

and so the factors need to have opposite sign, or at least one be equal to 0. This leads to

$$1 - \sqrt{1-c} \leq x \leq 1 + \sqrt{1-c}$$

(the other alternative is impossible), from which we conclude that $c \leq 1$ is necessary and sufficient.

5 We aim to apply Proposition 7 from "Minimisers of Optimization Problems."

Taking for granted that $\|\cdot\|_{\mathrm{F}}$ is a norm on $\mathbb{R}^{n\times n}$ (exercise: show this), it follows that $\|\cdot\|_{\mathrm{F}}$ is continuous (all norms are continuous; prove it if you want). Moreover, the determinant function $\det\colon \mathbb{R}^{n\times n} \to \mathbb{R}$ is continuous because $\det A$ equals just a linear combination of products of the elements of A. Thus the map

$$f\colon \Omega \to \mathbb{R} \qquad \text{given by} \qquad f(A) = \|A\|_{\mathrm{F}} + \frac{1}{\det A},$$

where $\Omega = \{A \in \mathbb{R}^{n\times n} : \det A > 0\}$, is continuous (in particular, l.s.c.).

But Ω is an open, unbounded set, and so we need to find a compact subset $K \subset \Omega$ where global minimiser(s) in Ω ought to exist. Since $f(I_{n\times n}) = n+1$ for the identity matrix $I_{n\times n}$, it suffices to consider $A \in \Omega$ with

$$\|A\|_{\mathrm{F}} \leq n+1 \qquad \text{and} \qquad \frac{1}{\det A} \leq n+1.$$

Defining

$$K = \left\{A \in \mathbb{R}^{n\times n} : \|A\|_{\mathrm{F}} \leq n+1 \text{ and } \det A \geq (n+1)^{-1}\right\},$$

which is a compact set (why?), we conclude that f admits a global minimum in $K \subset \Omega$.

6 **a)** Routine differentiation yields

$$\nabla f(x) = \begin{bmatrix} -400x_1(x_2 - x_1^2) - 2(1-x_1) \\ 200(x_2 - x_1^2) \end{bmatrix}$$

and

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

b) Since f is defined in terms of quadratic terms, it is bounded from below by 0. Moreover $f(x) = 0$ if and only if $x_2 = x_1^2$ and $1 - x_1 = 0$, which means that $x_1 = 1$ and $x_2 = 1$. Therefore $(1,1)$ is the unique minimiser of f.

Another argument: $(1,1)$ is the only stationary/critial point of f, that is, the only point for which $\nabla f = 0$. Moreover, the Hessian

$$\nabla^2 f(1,1) = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

has eigenvalues $501 \pm \sqrt{250601} > 0$, and so f is symmetric positive definite at $(1,1)$. Hence, $(1,1)$ is a strict local minimiser by Theorem 2.4 in N&W. As f is coercive (check it), $(1,1)$ must be a global minimiser, and in fact, *the* minimiser since there are no other stationary points.