Norwegian University of Science and Technology
Department of Mathematical Sciences

Solutions to exercise set 2

1 **a)** $\nabla f(x,y,z) = (4x+y-6, x+2y+z-7, y+2z-8)$, so critial/stationary point(s) of f, that is, points for which $\nabla f = 0$, must satisfy the system

$$\begin{bmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 8 \end{bmatrix}.$$

The unique solution is $(x,y,z) = (6,6,17)/5$.

b) Since f is twice continuously differentiable—in fact, smooth—around $(6,6,17)/5$, this point will be a strict local minimum provided the Hessian of f is symmetric positive definite (SPD) there. Now, for every (x,y,z),

$$\nabla^2 f = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

All three eigenvalues of $\nabla^2 f$ are positive, and thus it is SPD.

c) Function f is strictly convex because $\nabla^2 f$ is SPD. Hence, local minimisers are global minimisers, and we conclude that $(6,6,17)/5$ is a (*the*, in fact) global minimum.

2 Let $u \in \mathbb{R}^n$ be arbitrary. Using the hint, note first that stationary points of f_u are solutions to the equation $\nabla f(x) = u$, because $\nabla f_u(x) = \nabla f(x) - u$. Thus it suffices to show that f_u has a critical point, and in particular, we look for a global minimum, which is guaranteed to exist provided f_u is coercive (it is continuously differentiable and therefore lower semi-continuous). Now, f is certainly more than coercive: it grows superlinearly—faster than linear in the sense of the norm—to $+\infty$ as $\|x\| \to \infty$. (Indeed,

$$f(x) = \frac{f(x)}{\|x\|} \cdot \|x\| \to +\infty \cdot (+\infty) = +\infty \qquad \text{as} \qquad \|x\| \to \infty.)$$

Moreover, from Cauchy–Schwarz' inequality we have $u^\top x \le \|u\|\|x\|$, and so

$$f_u(x) = \frac{f(x)}{\|x\|}\|x\| - u^\top x \ge \left(\frac{f(x)}{\|x\|} - \|u\|\right)\|x\|.$$

Since $f(x)/\|x\| \to +\infty$ as $\|x\| \to \infty$, the first part in the parenthesis will eventually be larger than $\|u\|$, no matter which u we consider. Therefore $f_u(x) \to +\infty$ as $\|x\| \to \infty$, and f_u is coercive.

3 We utilise that twice continuously differentiable functions are convex if and only if their Hessian matrix is (symmetric) positive semi-definite. Routine calculations yield that

$$\nabla^2 f(x,y) = \frac{\mathrm{e}^{x+y}}{(\mathrm{e}^x + \mathrm{e}^y)^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

The exponential terms are positive, and the constant matrix has eigenvalues 0 and 2. As such, $\nabla^2 f$ is positive semi-definite, and f is convex.

4 On $\mathbb{R}_{>0} = (0, \infty)$, the map $-\log x$ is strictly convex because $(-\log x)'' = x^{-2} > 0$. Thus

$$-\log(\alpha x + \beta y) \leq \alpha(-\log x) + \beta(-\log y) = -\log\bigl(x^\alpha y^\beta\bigr),$$

(where $\beta = 1 - \alpha$) or

$$\log\bigl(x^\alpha y^\beta\bigr) \leq \log(\alpha x + \beta y),$$

with equality if and only if $x = y$. Exponentiating each side gives the arithmetic–geometric mean inequality.

5 Denote by $C \subseteq \mathbb{R}^n$ the set of minimisers of f, and let $x, y \in C$ and $\alpha \in (0,1)$. Then $f(x) = \min f = f(y)$ by assumption, and convexity of f yields that

$$f\left(\alpha x + (1-\alpha)y\right) \leq \alpha f(x) + (1-\alpha) f(y) = \min f.$$

Hence, by definition of $\min f$, we must have $f\left(\alpha x + (1-\alpha)y\right) = \min f$. In other words, $\alpha x + (1-\alpha)y \in C$, so C is convex.

6 Suppose that f has two global minimisers x and y. Since the set of minimisers is convex, it follows that $\alpha x + (1-\alpha)y$ is also a global minimiser for any $\alpha \in (0,1)$. But

$$f\left(\alpha x + (1-\alpha)y\right) < \alpha f(x) + (1-\alpha) f(y) = \min f$$

by strict convexity of f, which contradicts the fact that x and y are global minimisers. Therefore f has at most one global minimiser.

Every exponential map $f\colon x \mapsto a^x$ with $a > 0$ and $a \neq 1$ is strictly convex, but admits no local (and thus no global) minimiser on $\mathbb{R}$. Indeed, since $f'(x) = a^x \ln a$, it follows that f is strictly increasing for $a > 1$ and strictly decreasing for $0 < a < 1$. Moreover, strict convexity is a consequence of $f''(x) = a^x (\ln a)^2 > 0$.

7 Geometrically, the idea is as follows: consider a sphere of some radius R around the global minimum x^*. Let $C \subset \mathbb{R}^n \times \mathbb{R}$ be the circular convex cone with vertex $(x^*, f(x^*))$, and that touches the minimum of f on the sphere. Note that this minimum is greater than $f(x^*)$ by strict convexity of f. Then $(x, f(x))$ lies within C whenever $\|x - x^*\| \geq R$ because f is convex, and so f blows up for large x (since the cone "blows up").

Concretely, let $R = 1$ and c be the minimum of f on the sphere $\{y : \|y - x^*\| = 1\}$. Let x be any point outside the sphere, that is, $\|x - x^*\| \geq 1$. Then we can consider a corresponding point y on the sphere as the convex combination

$$y = x^* + \frac{1}{\|x - x^*\|}(x - x^*) = \left(1 - \frac{1}{\|x - x^*\|}\right) x^* + \frac{1}{\|x - x^*\|} x.$$

By convexity of f we have

$$f(y) \leq \left(1 - \frac{1}{\|x - x^*\|}\right) f(x^*) + \frac{1}{\|x - x^*\|} f(x),$$

and since $f(y) \geq c$, this yields

$$f(x) \geq f(x^*) + \big(c - f(x^*)\big)\|x - x^*\|.$$

Observing that $c > f(x^*)$ by strict convexity of f (why?), we may let $\|x\| \to \infty$ and conclude that $f(x) \to +\infty$, as desired.