Solutions to Problem Sheet 1

Solution (1.1)

(a) The gradient is given by

$$\nabla f = \begin{pmatrix} 3x^2 - 12y \\ -12x + 24y^2 \end{pmatrix}$$

Therefore

$$3x^2 - 12y = 0$$
$$-12x + 24y^2 = 0$$

The second equation implies $x = 2y^2$, therefore

$$3y^4 - 12y = 0 \Rightarrow y(y^3 - 1) = 0$$

Therefore $y_1 = 0, x_1 = 0$ or $y_2 = 1, x_2 = 2$. **(2)**

The Hessian is

$$\nabla^2 f(x, y) = \begin{pmatrix} 6x & -12 \\ -12 & 48y \end{pmatrix}$$

Its determinante is det$\nabla^2 f = 288xy - 122$. **(2)**
Therefore

- $(0, 0)$: det$\nabla^2 f(0, 0) = -122$ and f_{xx} at $x = 0$ is zero, hence it is a saddle point **(1)**
- $(2, 1)$: det$\nabla^2 f(2, 1) = 12$ and $f_{xx}(2, 1) > 0$; hence it is a local minimum. **(1)**

(b) The gradient is given by **(1)**

$$\nabla f(x, y) = \begin{pmatrix} \cos(x)\cos(y) \\ -\sin(x)\sin(y) \end{pmatrix}$$

The optimality condition implies that (each row corresponding to the respective component of the equation system

$$x = \frac{\pi}{2} \quad \text{or} \quad x = \frac{3\pi}{2} \quad \text{or} \quad y = \frac{\pi}{2} \quad \text{or} \quad y = \frac{3\pi}{2}$$
$$x = 0 \quad \text{or} \quad x = \pi \quad \text{or} \quad y = 0 \quad \text{or} \quad y = \pi.$$

This gives the following critical points **(2)**

$$P_1 = (\frac{\pi}{2}, 0),\ P_2 = (\frac{\pi}{2}, \pi),\ P_3 = (\frac{3\pi}{2}, 0),\ P_4 = (\frac{3\pi}{2}, \pi)$$
$$P_5 = (0, \frac{\pi}{2}),\ P_6 = (0, \frac{3\pi}{2})\ P_7 = (\pi, \frac{\pi}{2}), P_8 = (\pi, \frac{3\pi}{2}).$$

The Hessian is **(1)**

$$\nabla^2 f(x,y) = \begin{pmatrix} -\sin(x)\cos(y) & -\cos(x)\sin(y) \\ -\cos(x)\sin(y) & -\sin(x)\cos(y) \end{pmatrix}$$

The respective matrices are given by **(2)**

$$P_1 : \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{maximum}$$

$$P_2 : \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{minimum}$$

$$P_3 : \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{minimum.}$$

$$P_4 : \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{maximum}$$

$$P_5 : \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \qquad \text{saddle}$$

$$P_6 : \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{saddle}$$

$$P_7 : \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{saddle}$$

$$P_8 : \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \qquad \text{saddle}$$

Solution (1.2)

(a) The function $f(x) = x^4$ has a strict minimum at $x = 0$, but the second derivative satisfies $f''(0) = 0$. **(1)**

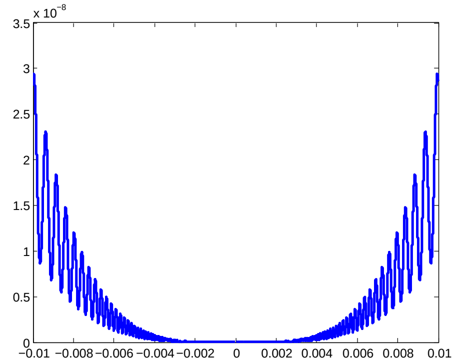
(b) We construct a function that has a strict minimizer x^*, but such that every open neighourhood U of x^* contains other local minimizers. One such function is

$$f(x) = \begin{cases} x^4(\cos(1/x) + 2) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0. \end{cases}$$

We explain the construction of this function:

1. Start out with $g(x) = \cos(1/x) + 2$ for $x \neq 0$ and $g(0) = 1$. This function has minimizers $x_0 = 0$ and $x_k = 1/(\pi(2k+1)$ for $k \geq 0$, with values $g(x_k) = 1$ at all minimizers. Therefore, any open interval around 0 contains (infinitely many) local minimizers x_k other than $x_0 = 0$.
2. Multipy x^4 to the function: $f(x) = x^4 g(x)$. This ensures that $f(0) = 0$ and $f(x) > 0$ for $x \neq 0$. There are still local minima in every neighbourhood of 0. To see this, compute the derivative

$$f'(x) = x^2(4x\cos(1/x) + \sin(1/x) + 8x). \tag{1}$$



Set $z_m = 1/(\pi/2 + m\pi)$ for $m > 0$. Since $\sin(1/z_m) = \sin(\pi/2 + m\pi) = 1$ for m even and -1 for m odd, and for m sufficiently large the contribution of the other terms is negligible (as the z_m become arbitrary small), the derivative (1) changes signs between successive z_m. Since $f'(x)$ is continuous, it has roots between any z_m and z_{m+1} for large enough m, and these correspond to maxima and minima of f.

The function is in $C^2(\mathbb{R})$. For $x \neq 0$ this is clear, and to verify this at $x = 0$, one shows that the right and left limits as $x \to 0$ of $f'(x)$ and $f''(x)$ coincide (they are in fact 0).

Note the subtle point that one minimizer x^* can have local minimizers that are arbitrary close: while each open interval I surrounding x^* has another local minimizer $\tilde{x}$, every such $\tilde{x}$ has an interval $\tilde{I}$ surrounding it where this $\tilde{x}$ is the only minimizer! **(4)**

Note: there is no unique way of solving this, and any sensible function with the desired properties will do the job. A solution that is based on the right idea but that is otherwise incomplete is worth two points.

Solution (1.3) The general procedure is as follows: we first make an educated guess as to whether the function could be convex or not. If we think it is not convex, then it is enough to find a *counterexample*: find points in S for which the line segment joining them is not completely contained in S. If we think it is convex, then we can show that for any two points the line segment joining them is in S.

(a) This set is not convex: take $\boldsymbol{x} = (1,0,0)^\top$ and $\boldsymbol{y} = (-1,0,0)^\top$, then $\frac{1}{2}\boldsymbol{x} + \frac{1}{2}\boldsymbol{y} = \boldsymbol{0} \notin S$.

(b) This set is convex: if $\boldsymbol{x}, \boldsymbol{y} \in S$, then $1 \leq x_1 - x_2 < 2$ and $1 \leq y_1 - y_2 < 2$, and

$$\lambda x_1 + (1-\lambda)y_1 - \lambda x_2 - (1-\lambda)y_2 = \lambda(x_1 - x_2) + (1-\lambda)(y_1 - y_2) < \lambda 2 + (1-\lambda)2 = 2,$$

with the same argument giving the lower bound.

(c) This set is convex. In fact, S is the unit ball of the 1-norm

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|.$$

Given $\boldsymbol{x}, \boldsymbol{y} \in S$,

$$\|\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|_1 \leq \lambda\|\boldsymbol{x}\|_1 + (1-\lambda)\|\boldsymbol{y}\|_1 \leq \lambda \cdot 1 + (1-\lambda) \cdot 1 = 1.$$

(d) This set is convex. Here, one needs to show that convex combinations preserve symmetry and positive definiteness of a matrix. The symmetry is clear. As for the positive definiteness, let $\boldsymbol{x} \neq \mathbf{0}$ be given. Then

$$\boldsymbol{x}^\top(\lambda\boldsymbol{A} + (1-\lambda)\boldsymbol{B})\boldsymbol{x} = \lambda\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x} + (1-\lambda)\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x} \geq 0,$$

which shows that positive definiteness is also preserved.