1. Question: Classifying critical points (*elementary*)

1.1. True or false? Motivate your answer. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a twice differentiable function with a critical point p_0, whose Hessian matrix at p_0 is

$$\boldsymbol{H}f(p_0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Then:

(a) p_0 cannot be a local maximum

(b) p_0 cannot be a local minimum

(c) p_0 cannot be a saddle point

(d) none of the above.

1.2. True or false? Motivate your answers. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable function and consider its restriction over the square $Q = [0,1] \times [0,1] \subset \mathbb{R}^2$. Then:

(a) If f has a local max / min / saddle at x_0 in Q, then $df(x_0) = 0$

(b) Let $x_0 \in Q$ be a point such that $df(x_0) = 0$, then f has a local max/min/saddle at x_0.

1.3. For each of the following functions, determine their critical points and find those for which the 2nd derivative test applies, determining in such case whether they are local maxima, local minima or saddle points.

(a) $f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = x^3 + y^3 - 3xy,$

(b) $f : \mathbb{R}^3 \to \mathbb{R}, \quad f(x,y,z) = \left(x^3 - 3x - y^2\right) z^2 + z^3,$

(c) $f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = xy^2 - \cos(x).$

Solution:

1.1. (a) p_0 cannot be a local maximum: **TRUE**

(b) p_0 cannot be a local minimum: **FALSE**

(c) p_0 cannot be a saddle point: **FALSE**

(d) none of the above: **FALSE**.

Indeed, denoting $p_0 = (x_0, y_0, z_0)$, the restriction $\varphi(y,z) = f(x_0, y, z)$ has (y_0, z_0) as critical point and $\boldsymbol{H}\varphi(y_0, z_0)$ is positive definite, whence φ has a local minimum at (y_0, z_0). Therefore p_0 cannot be a local maximum for f. But otherwise it is a simple matter to cook up examples where p_0 is either a local minimum or a saddle point: $f(x,y,z) = x^4 + \frac{1}{2}y^2 + z^2$ has in $p_0 = 0$ its absolute minimum (since $f > 0$ in $\mathbb{R}^3 \setminus \{0\}$), while $g(x,y,z) = x^3 + \frac{1}{2}y^2 + z^2$ has in $p_0 = 0$ a saddle point (since, if it where a local minimum, 0 would be a local minimum for the 1-variable function $g(x,0,0) = x^3$, but this latter is a saddle point instead).

1.2. (a) If f has a local max / min / saddle at x_0 in Q, then $df(x_0) = 0$: **FALSE**

(b) Let $x_0 \in Q$ be a point such that $df(x_0) = 0$, then f has a local max/min/saddle at x_0: **TRUE**.

Consider for instance $f(x,y) = x$: clearly $f \leq 1$ and so f has a maximum at each point in $(1,y)$; however $df(1,y) = (1,0) \neq 0$, so (a) is false. However, if $df(x_0) = 0$, this means that f has a min / max / saddle at x_0 in $\mathbb{R}^2$, and thus so it is for the restriction of f to Q.

1.3. (a) The differential of f is

$$df(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = \left(3x^2 - 3y, 3y^2 - 3x\right)$$

so its critical points are the solution to the system

$$\begin{cases} x^2 = y \\ y^2 = x \end{cases}$$

This means that

$$y = x^2 = \left(y^2\right)^2 = y^4 \Leftrightarrow \left(y^3 - 1\right) y = 0 \Leftrightarrow y \in \{1, 0\}$$

For $y = 1$ it follows $x = 1$, For $y = 0$ it follows $x = 0$. So the critical points are $(1,1)$ and $(0,0)$. The Hessian matrix is

$$\boldsymbol{H} f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y \partial y} \end{pmatrix} = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

At the critical points we have

$$\boldsymbol{H} f(1,1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}, \quad \boldsymbol{H} f(0,0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}.$$

Since $\boldsymbol{H} f(1,1)$ has eigenvalues 3 and 9, it is positive definite and thus $(1,1)$ is a local minimum. Since $\boldsymbol{H} f(0,0)$ has eigenvalues 3 and -3, it is indefinite and so $(0,0)$ is a saddle point.

(b) The differential of f is

$$df(x,y,z) = \left(3\left(x^2 - 1\right) z^2, -2yz^2, 2z\left(x^3 - 3x - y^2\right) + 3z^2\right)$$

so the critical points are the solutions to

$$\begin{cases} 3\left(x^2 - 1\right) z^2 = 0, \\ -2yz^2 = 0, \\ 2z\left(x^3 - 3x - y^2\right) + 3z^2 = 0. \end{cases}$$

Clearly, every point with $z = 0$ is a solution to the system. When $z \neq 0$, we have the system

$$\begin{cases} \left(x^2 - 1\right) = 0, \\ y = 0, \\ 2\left(x^3 - 3x - y^2\right) + 3z = 0. \end{cases}$$

whose solutions are then $\left(1, 0, \frac{4}{3}\right)$ and $\left(-1, 0, -\frac{4}{3}\right)$. To sum up the critical points are

$$(x, y, 0) \text{ for every } x, y \in \mathbb{R}, \quad (1, 0, 4/3), \quad (-1, 0, -4/3).$$

We need to determine their type. The Hessian of f is

$$\boldsymbol{H} f(x,y,z) = \begin{pmatrix} 6xz^2 & 0 & 6z\left(x^2 - 1\right) \\ 0 & -2z^2 & -4yz \\ 6z\left(x^2 - 1\right) & -4yz & 2\left(x^3 - 3x - y^2\right) + 6z \end{pmatrix},$$

hence

$$\boldsymbol{H} f(1, 0, 4/3) = \begin{pmatrix} 6(4/3)^2 & 0 & 0 \\ 0 & -2(4/3)^2 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

which is clearly indefinite, so $(1,0,4/3)$ is a saddle point; then

$$\boldsymbol{H} f(-1,0,-4/3)=\left(\begin{array}{ccc} -6(4/3)^2 & 0 & 0 \\ 0 & -2(4/3)^2 & 0 \\ 0 & 0 & -4 \end{array}\right),$$

which is clearly negative definite, so $(-1,0,-4/3)$ is a local maximum; finally

$$\boldsymbol{H} f(x, y, 0)=\left(\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2\left(x^3-3 x-y^2\right) \end{array}\right),$$

is has eigenvalues equal to 0, so the second derivative test does not apply in this case.

(c) The differential of f is

$$d f(x, y)=\left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right)=\left(y^2+\sin (x), 2 x y\right).$$

so the critical points are the solutions to the system

$$\begin{cases} y^2+\sin (x)=0, \\ 2 x y=0. \end{cases}$$

From the 2 nd equation it follows that $x=0$ or $y=0$. If $x=0$, the 1 st equation yields $y=0$. If $y=0$ the 1 st equation yields $x=k \pi$ with $k \in \mathbb{Z}$. The set of critical points is then $\{(k \pi, 0) \mid k \in \mathbb{Z}\}$. The Hessian is

$$\boldsymbol{H} f(x, y)=\left(\begin{array}{cc} \frac{\partial^2 f}{\partial x \partial x} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y \partial y} \end{array}\right)=\left(\begin{array}{cc} \cos (x) & 2 y \\ 2 y & 2 x \end{array}\right).$$

At $(k \pi, 0) \in \mathbb{R}^2$ with $k \in \mathbb{Z}$ we get

$$\boldsymbol{H} f(k \pi, 0)=\left(\begin{array}{cc} \cos (k \pi) & 0 \\ 0 & 2 k \pi \end{array}\right)=\left(\begin{array}{cc} (-1)^k & 0 \\ 0 & 2 k \pi \end{array}\right),$$

which is diagonal, and hence we deduce that

$k>0$	even	$\Rightarrow$	both EV positive	$\Rightarrow$ local minima
$k>0$	odd	$\Rightarrow$	EV with different sign	$\Rightarrow$ saddle point
$k=0$		$\Rightarrow$	one EV vanishes	$\Rightarrow$ cannot conclude
$k<0$	even	$\Rightarrow$	EV with different sign	$\Rightarrow$ saddle point
$k<0$	odd	$\Rightarrow$	both EV negative	$\Rightarrow$ local maxima.

2. Question: Second derivative test for $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$ *(elementary)*

2.1. Prove the following statement. (***A bit more advanced, you may also just use this statement for the following question for now.***)

Let $A=\left(\begin{array}{cc} a & b \\ b & c \end{array}\right)$ *be a symmetric* 2×2 *matrix, where* $a, b, c \in \mathbb{R}$. *Then* A *is positive definite if* $a>0$ *and* $a c-b^2>0$.A *is negative definite if* $a<0$ *and* $a c-b^2>0$.

2.2. Show that if $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$ has a critical point $x_0 \in A$ and we let

$$\Delta = \frac{\partial^2 f}{\partial x_1 \partial x_1} \cdot \frac{\partial^2 f}{\partial x_2 \partial x_2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2$$

be evaluated at x_0, then

(a) $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 > 0$ imply f has a local minimum at x_0.

(b) $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 < 0$ imply f has a local maximum at x_0.

(c) $\Delta < 0$ implies x_0 is a saddle point of f.

Solution:

2.1. Let $v = (x, y)^\top$ be an arbitrary non-zero vector. Then

$$\begin{aligned} v^\top A v &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2 \\ &= ax^2 + 2bxy + \frac{b^2y^2}{a} - \frac{b^2y^2}{a} + cy^2 \\ &= a\left(x + \frac{by}{a}\right)^2 + \left(c - \frac{b^2}{a}\right) y^2 \end{aligned}$$

If A is to be positive definite, then $v^\top A v > 0$ for all v. In particular, for $y = 0$, we must have $ax^2 > 0$, which implies $a > 0$. Also, when $x = -(b/a)y$, $\left(c - b^2/a\right) y^2 > 0$ implies $ac - b^2 > 0$. If A is to be negative definite, then $v^\top A v < 0$ for all v. In particular, for $y = 0$, we must have $ax^2 < 0$, which implies $a < 0$. Also, when $x = -(b/a)y$, $\left(c - b^2/a\right) y^2 < 0$ implies $a\left(c - b^2/a\right) > 0$, which yields $ac - b^2 > 0$.

2.2. First, we note that $\Delta = \det(\boldsymbol{H} f)$.

(a) We know that f has a local minimum at x_0 if x_0 is a critical point of f such that $\boldsymbol{H} f(x_0)$ is positive definite. Thus we need only show that $\boldsymbol{H} f(x_0)$ is positive definite if $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 > 0$.
Since the Hessian matrix $\boldsymbol{H} f$ clearly satisfies the conditions of the statement from **2.1**, it follows that $\boldsymbol{H} f(x_0)$ is positive definite if $\Delta = ac - b^2 > 0$ and $\partial^2 f/\partial x_1 \partial x_1 = a > 0$.

(b) Since the Hessian matrix clearly satisfies the conditions of the lemma, it follows that $\boldsymbol{H} f(x_0)$ is negative definite if $\Delta = ac - b^2 > 0$ and $\partial^2 f/\partial x_1 \partial x_1 = a < 0$.

(c) If x_0 is not a local max or min, then it must be a saddle point. Similarly, if it is not the case that $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 > 0$ or $\Delta > 0$ and $\partial^2 f/\partial x_1 \partial x_1 < 0$, then it must be true that $\Delta < 0$. Therefore, $\Delta < 0$ implies x_0 is a saddle point of f.

3. Question: Taylor Polynomial and Taylor-Series *(elementary)*

3.1. Show that the Taylor Series generated by the function $f(x) := e^x$ at $x_0 = 0$ converges to $f(x)$ for every value of x.

3.2. Use your result form **3.1** to prove that $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

In fact, it turns out that e^x is analytic in the sense that the Taylor series $\sum_{k=0}^{\infty} \frac{e^{x_0}}{k!}(x - x_0)^k$ converges to e^x for all $x_0 \in \mathbb{R}$. However, even a converging Taylor series does not necessarily converge to its corresponding function $f(x)$.

3.3. Find the Taylor series approximation of

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

at $x_0 = 0$. How accurate is the k th degree Taylor approximation?

Hint: You may use the fact that $\lim_{x\to\infty} \frac{e^{-\frac{1}{x^2}}}{x^n} = 0 \ \forall n \in \mathbb{N}$.

Solution:

3.1. Since $\frac{\partial}{\partial x} e^x = e^x$, $f(x)$ is infinitely differential on $(-\infty, \infty)$. Using the Taylor Polynomial generated by $f(x) = e^x$ at $x_0 = 0$ and Taylor's theorem, we have

$$e^x = \sum_{k=0}^{\infty} \frac{e^0}{k!} (x-0)^k + R_n^{x_0}(x)(x) \overset{e^0=1}{=} 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + R_n^{x_0}(x)(x)$$

where $R_n^{x_0}(x)(x) = \frac{e^c}{(n+1)!} x^{n+1}$ for some c between 0 and x. Recall that e^x is an increasing function, so;

$$\begin{aligned} x > 0: &\quad 0 < c < x \Longrightarrow e^0 < e^c < e^x \Longrightarrow 1 < e^c < e^x \\ x < 0: &\quad x < c < 0 \Longrightarrow e^x < e^c < e^0 \Longrightarrow e^x < e^c < 1 \\ x = 0: &\quad e^x = 1, x^{n+1} = 0 \Longrightarrow R_n^{x_0}(x)(x) = 0. \end{aligned}$$

And, therefore,

$$\begin{aligned} x > 0: &\quad |R_n^{x_0}(x)(x)| = \left| \frac{e^c x^{n+1}}{(n+1)!} \right| \leq \frac{e^x x^{n+1}}{(n+1)!} \xrightarrow{n\to\infty} 0 \\ x \leq 0: &\quad |R_n^{x_0}(x)(x)| = \left| \frac{e^c x^{n+1}}{(n+1)!} \right| \leq \frac{|x|^{n+1}}{(n+1)!} \xrightarrow{n\to\infty} 0. \end{aligned}$$

It follows that $\lim_{n\to\infty} R_n^{x_0}(x)(x) = 0$ for all x, so the series converges to e^x on $(-\infty, \infty)$. Thus, $\forall x \in (-\infty, \infty)$

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

3.2. This immediately follows from the fact that $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ for all $x \in (-\infty, \infty)$, including $x = 1$.

3.3. It turns out that $f(x)$ is infinitely differentiable at $x = 0$. For $x \neq 0$, we can use the chain rule to compute its derivative explicitly,

$$f'(x) = \frac{2}{x^3} \cdot e^{-\frac{1}{x^2}}, \quad f''(x) = \frac{4 - 6x^2}{x^6} \cdot e^{-\frac{1}{x^2}}, \quad f'''(x) = \frac{24x^4 - 36x^2 + 8}{x^9} \cdot e^{-\frac{1}{x^2}}, \quad \ldots$$

In general, the nth derivative will be a rational functions up to order x^{-3n} times $e^{-\frac{1}{x^2}}$. If we take the limit as $x \to 0$, each of the terms above is 0, as $\lim_{x\to\infty} \frac{e^{-\frac{1}{x^2}}}{x^{3n}} = 0 \ \forall n \in \mathbb{N}$ by the given hint, so

$$f^{(n)}(0) = 0 \quad \text{for all } n \geq 0.$$

This means that the Taylor series for $f(x)$ is the constant function 0 . For any $k > 0$, the k th degree Taylor polynomial $P_k^{x_0}(x) = 0$, which implies that

$$R_k^{x_0}(x) = f(x) - P_k^{x_0}(x) = f(x) \neq 0$$

unless $x = 0$. Therefore, the Taylor polynomial approximation is completely useless for computing $f(x)$.

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

$$f(x) = P_n^{x_0}(x) + \frac{f^{(n+1)}(c)}{(n+1)!} (x - x_0)^{n+1} .$$

The term $R_n^{x_0}(x)^{x_0}(x) := \frac{f^{(n+1)}(c)}{(n+1)!} (x - x_0)^{n+1}$

If you have any questions or feedback, please feel free to contact me via E-mail at hannah.kuempel@stat.uni-muenchen.de!!

Also, thank you to the authors of the book *Mathematics for Machine Learning* as well as Prof. Özlem Imamoglu, and Anthony Varilly whose exercises this sheet was inspired by.