Exercise 4.1. Suppose A is a symmetric $(n \times n)$ matrix. Consider the function:

$$\begin{aligned} f \ : \ \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto xAx^t. \end{aligned}$$

(a) Show that f is differentiable at all points $p \in \mathbb{R}^n$, with:

$$Df(p) = 2pA$$

(b) Find:

$$\operatorname{Hess} f(p).$$

Solution: (a) Fix $p \in \mathbb{R}^n$. We compute:

$$\begin{aligned} f(p+h) - f(p) - Df(p)[h] &= (p+h)A(p+h)^t - pAp^t - 2pAh^t \\ &= pAp^t + pAh^t + hAp^t + hAh^t - pAp^t - 2pAh^t \\ &= hAh^t, \end{aligned}$$

where we have used that A is symmetric to deduce $hAp^t = pAh^t$. Now, recall from an example in the lecture notes that for any matrix A there exists a constant K such that:

$$\left\| Ax^t \right\| \leq K \left\| x \right\|$$

for all $x \in \mathbb{R}^n$. Applying the Cauchy-Schwartz inequality we have:

$$\left\| hAh^t \right\| \leq \left\| h \right\| \left\| Ah^t \right\| \leq K \left\| h \right\|^2.$$

Thus, we conclude:

$$\frac{\left\| f(p+h) - f(p) - Df(p)[h] \right\|}{\left\| h \right\|} \leq K \left\| h \right\| \to 0,$$

as $h \to 0$, thus we have that f is differentiable with derivative $Df(p) = 2pA$.

(b) If we write $A = (A_{ij})_{i,j=1}^n$, then we can write:

$$Df(p)[h] = \sum_{j=1}^n D_j f(p) h^j = 2 \sum_{i,j=1}^n p^i A_{ij} h^j$$

where $p = (p^1, \dots p^n)$, $h = (h^1, \dots, h^n)$. We deduce that:

$$D_j f(p) = 2 \sum_{i=1}^n p^i A_{ij}$$

Taking a further derivative, we conclude:

$$D_i D_j f(p) = 2A_{ij}.$$

Thus

$$\operatorname{Hess} f(p) = 2A.$$

Please send any corrections to d.cheraghi@imperial.ac.uk
Questions marked with $*$ are optional

Exercise 4.2. Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ given by:

$$f(x, y, z) = xy^2 + x^2 + xze^y.$$

(a) Compute the first and second partial derivatives. Observe the properties of the second partial derivative.

(b) Write the terms of the Taylor expansion of f at zero up to and including the second-order terms.

(c) Without computation, write the same Taylor expansion up to and including the fourth-order terms.

Solution: (a) We have

$$D_1 f = y^2 + 2x + ze^y, \quad D_2 f = 2xy + xze^y, \quad D_3 f = xe^y.$$

Furthermore,

$$D_1 D_1 f = 2, \quad D_2 D_1 f = 2y + ze^y, \quad D_3 D_1 f = e^y,$$
$$D_1 D_2 f = 2y + ze^y, \quad D_2 D_2 f = 2x + xze^y, \quad D_3 D_2 f = xe^y,$$
$$D_1 D_3 f = e^y, \quad D_2 D_3 f = xe^y, \quad D_3 D_3 f = 0.$$

(b) Thus by the general formula for the Taylor expansion, with $h_1 = x$, $h_2 = y$, $h_3 = z$,

$$\begin{aligned} f(x, y, z) &= \sum_{\alpha, |\alpha| \leq 2} D^\alpha f(0) \frac{h^\alpha}{\alpha!} + R_3 \\ &= f(0) + \sum_{j=1}^{3} D_j f(0) h_j + \sum_{j=1}^{3} D_j D_j f(0) \frac{h_j^2}{2!} + \sum_{j<k, j,k=1}^{3} D_j D_k f(0) h^j h^k + R_3 \\ &= x^2 + xz + R_3 \end{aligned}$$

(c)

$$f(x, y, z) = xy^2 + x^2 + xz(1 + y + y^2/2) + R_5$$

Exercise 4.3 (*)**.** Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by:

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \dfrac{xy^3 - x^3 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

(a) Show that:

$$D_1 f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \dfrac{y^3 - 3x^2 y}{x^2 + y^2} - \dfrac{2x\left(xy^3 - x^3 y\right)}{\left(x^2 + y^2\right)^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

and

$$D_2 f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \dfrac{3y^2x - x^3}{x^2+y^2} - \dfrac{2y\left(xy^3 - x^3y\right)}{\left(x^2+y^2\right)^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0), \end{cases}$$

and show that these functions are both continuous at $(0,0)$.

Solution: 1. Let $p = (x, y)$. If $p \neq 0$, we can differentiate using the quotient rule to find

$$D_1 f(p) = \frac{\partial f}{\partial x} = \frac{y^3 - 3x^2y}{x^2+y^2} - \frac{2x\left(xy^3 - x^3y\right)}{\left(x^2+y^2\right)^2}.$$

Further, note that $f(te_1) = 0$, so that:

$$\lim_{t \to 0} \frac{f(te_1) - f(0)}{t} = 0,$$

thus $D_1 f(0) = 0$.

2. Now, note that $\left|y^2 - 3x^2\right| \leq y^2 + 3x^2 \leq 3(y^2 + x^2)$, thus:

$$\left|\frac{y^3 - 3x^2y}{x^2+y^2}\right| = |y| \left|\frac{y^2 - 3x^2}{x^2+y^2}\right| \leq 3\,|y|$$

Also, note that by Young's inequality $\left|xy^3\right| \leq \frac{1}{2}x^2y^2 + \frac{1}{2}y^4$ and similarly $\left|x^3y\right| \leq \frac{1}{2}x^2y^2 + \frac{1}{2}x^4$, so that:

$$\left|xy^3 - x^3y\right| \leq \left|xy^3\right| + \left|x^3y\right| \leq \frac{1}{2}\left(x^4 + 2x^2y^2 + y^4\right) = \frac{1}{2}\left(x^2+y^2\right)^2.$$

We deduce:

$$\left|\frac{2x\left(xy^3 - x^3y\right)}{\left(x^2+y^2\right)^2}\right| \leq |x|,$$

so that for $p = (x, y)^t \neq 0$, we have:

$$|D_1 f(p)| \leq 3\,|y| + |x| \to 0$$

as $p \to 0$, so that $D_1 f(p)$ is continuous at $p = 0$.

3. Similarly, if $p \neq 0$, we can differentiate using the quotient rule to find

$$D_2 f(p) = \frac{\partial f}{\partial y} = \frac{3y^2x - x^3}{x^2+y^2} - \frac{2y\left(xy^3 - x^3y\right)}{\left(x^2+y^2\right)^2}.$$

Further, note that $f(te_2) = 0$, so that:

$$\lim_{t \to 0} \frac{f(te_2) - f(0)}{t} = 0,$$

thus $D_2 f(0) = 0$.

4. Now, note that $\left|3y^2 - x^2\right| \leq 3y^2 + x^2 \leq 3(y^2 + x^2)$, thus:

$$\left|\frac{3y^2x - x^3}{x^2 + y^2}\right| = |x| \left|\frac{3y^2 - x^2}{x^2 + y^2}\right| \leq 3\,|x|$$

Recalling that:

$$\left|xy^3 - x^3y\right| \leq \frac{1}{2}\left(x^2 + y^2\right)^2.$$

We deduce:

$$\left|\frac{2y\left(xy^3 - x^3y\right)}{\left(x^2 + y^2\right)^2}\right| \leq |y|\,,$$

so that for $p = (x, y) \neq 0$, we have:

$$|D_2 f(p)| \leq 3\,|y| + |x| \to 0$$

as $p \to 0$, so that $D_1 f(p)$ is continuous at $p = 0$.

(b) Show that:

$$\lim_{t \to 0} \frac{1}{t}\left(D_1 f(te_2) - D_1 f(0)\right) = 1$$

and

$$\lim_{t \to 0} \frac{1}{t}\left(D_2 f(te_1) - D_2 f(0)\right) = -1$$

Solution: We have (setting $x = 0$, $y = t$ in the formula for $D_1 f$):

$$D_1 f(te_2) = t, \quad D_1 f(0) = 0,$$

so that:

$$\lim_{t \to 0} \frac{1}{t}\left(D_1 f(te_2) - D_1 f(0)\right) = 1$$

Similarly, we have (setting $x = t$, $y = 0$ in the formula for $D_2 f$):

$$D_2 f(te_1) = -t, \quad D_1 f(0) = 0,$$

so that:

$$\lim_{t \to 0} \frac{1}{t}\left(D_2 f(te_1) - D_2 f(0)\right) = -1$$

(c) Conclude that both $D_2 D_1 f(0)$ and $D_1 D_2 f(0)$ exist, but that:

$$D_2 D_1 f(0) \neq D_1 D_2 f(0)$$

Solution: By definition,

$$D_2 D_1 f(0) = \lim_{t \to 0} \frac{1}{t}\left(D_1 f(te_2) - D_1 f(0)\right),$$

which certainly exists. Similarly,

$$D_1 D_2 f(0) = \lim_{t \to 0} \frac{1}{t}\left(D_2 f(te_1) - D_2 f(0)\right)$$

also exists, but as we've seen above the two are not equal.

Exercise 4.4. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as $f(x, y) = e^x \sin(y)$.

a) Compute the degree 1 and degree 2 Taylor polynomial of f near the point $(x_0, y_0) = (0, \pi/2)$ and use those to approximate the value of f at $(x_1, y_1) = (0, \pi/2 + 1/4)$. Compare your results with the values you obtain from a calculator.

Solution: For all $x, y \in \mathbb{R}$ we have

$$\begin{aligned} D_1 f(x,y) &= e^x \sin(y), & D_1 D_1 f(x,y) &= e^x \sin(y), & D_2 D_1 f(x,y) &= e^x \cos(y) \\ D_2 f(x,y) &= e^x \cos(y), & D_2 D_2 f(x,y) &= -e^x \sin(y), & D_1 D_2 f(x,y) &= e^x \cos(y). \end{aligned}$$

Evaluating the above expressions at $(x_0, y_0) = (0, \pi/2) \in \mathbb{R}^2$, we get $f(x_0, y_0) = 1$ as well as

$$\begin{aligned} D_1 f(0,\pi/2) &= 1, & D_1 D_1 f(0,\pi/2) &= 1, & D_2 D_1 f(0,\pi/2) &= 0 \\ D_2 f(0,\pi/2) &= 0, & D_2 D_2 f(0,\pi/2) &= -1, & D_1 D_2 f(0,\pi/2) &= 0. \end{aligned}$$

The Taylor polynomials $T_1 f$ and $T_2 f$ of degree 1 and 2, respectively, are therefore

$$T_1 f(x,y) = f(x_0,y_0) + D_1 f(x_0,y_0) \cdot (x - x_0) + D_2 f(x_0,y_0) \cdot (y - y_0) = 1 + x$$

and

$$\begin{aligned} T_2 f(x,y) = T_1 f(x,y) + \frac{1}{2}\Big[& D_1 D_1 f(x_0,y_0) \cdot (x-x_0)^2 + D_2 D_2 f(x_0,y_0) \cdot (y-y_0)^2 \\ & + 2 D_1 D_2 f(x_0,y_0) \cdot (x-x_0)(y-y_0) \Big] \\ = {} & 1 + x + \frac{1}{2}x^2 - \frac{1}{2}(y - \pi/2)^2. \end{aligned}$$

At the point $(x_1, y_1) = (0, \pi/2 + 1/4)$, these yield the approximations

$$T_1 f(0, \pi/2 + 1/4) = 1, \quad T_2 f(0, \pi/2 + 1/4) = 1 - 1/2(1/4)^2 = 31/32 = 0.96875.$$

The approximation by T_2 is very good as the actual value (using a high precision calculator) is

$$f(0, \pi/2 + 1/4) \approx 0.96891.$$

b) How precise is the degree 1 approximation in the closed ball of radius 1/4 around (x_0, y_0). Find a rigorous upper bound for the approximation error.

Solution: Let B denote the ball of radius 1/4 about (x_0, y_0), that is $B_{1/4}(x_0, y_0)$. By Theorem 1.14, the remainder term $R_2 = f - T_1 f$ can be expressed as

$$\begin{aligned} R_2(x,y) = \frac{1}{2}\Big[& D_1 D_1 f(x_r, y_r) \cdot (x - x_0)^2 \\ & + D_2 D_2 f(x_r, y_r) \cdot (y - y_0)^2 + 2 D_1 D_2 f(x_r, y_r) \cdot (x - x_0)(y - y_0) \Big] \end{aligned}$$

for some (x_r, y_r) such that x_r lies in the interval $[x_0, x]$ when $x > x_0$ and in the interval $[x, x_0]$ when $x \le x_0$, and similarly for y_r. In particular, for all $(x, y) \in B_{1/4}(x_0, y_0)$, this gives $|x_r - x_0| \le 1/4$ and $|y_r - y_0| \le 1/4$. Moreover, by part a) for all $(x, y) \in \mathbb{R}^2$

we have $|D_1D_1f(x,y)| \leq e^x$, $|D_1D_2f(x,y)| \leq e^x$, and $|D_2D_2f(x,y)| \leq e^x$, using $|\sin(x)| \leq 1$ and $|\cos(x)| \leq 1$. Overall, this gives for all $(x,y) \in B_{1/4}(x_0,y_0)$,

$$|R_2(x,y)| \leq 4\left(\tfrac{1}{2}e^{\frac{1}{4}} \cdot (\tfrac{1}{4})^2\right) = \tfrac{1}{8}e^{\frac{1}{4}} \approx 0.1605.$$

Even the (relatively crude) first-order approximation is off by at most about 16% of the value at (x_0,y_0) in $B_{1/4}(x_0,y_0)$.

Exercise 4.5. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by:

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y-xy \\ x^2 \end{pmatrix}$$

Determine the set of points in $\mathbb{R}^2$ such that f is invertible near those points, and compute the derivative of the inverse map.

Solution: The derivative is

$$Df = \begin{pmatrix} 1-y & 1-x \\ 2x & 0 \end{pmatrix}.$$

We have $\det Df = 2x(x-1)$ which is zero if $x=0$ or $x=1$ for any y. Thus, for any $(x,y) \in \mathbb{R}^2$ such that $x \notin \{0,1\}$, the function is invertible on a ball around $(x,y) \in \mathbb{R}^2$, and the derivative of the inverse is

$$Df^{-1} = (Df)^{-1} = \frac{1}{2x(x-1)} \begin{pmatrix} 0 & x-1 \\ -2x & 1-y \end{pmatrix}.$$

Exercise 4.6. (a) Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable in a neighbourhood of the origin, and $f'(0) = 0$. Give an example to show that f may nevertheless be bijective.

[Hint: Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f : x \mapsto x^3$.]

Solution: The function $f : x \mapsto x^3$ is strictly monotone increasing and continuous, hence it is bijective. On the other hand $f'(0) = 0$.

(b) Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is bijective, differentiable at the origin, and $\det Df(0) = 0$. Show that f^{-1} is not differentiable at $f(0)$.

[Hint: Assume that f^{-1} is differentiable at $f(0)$ and apply the chain rule to $\iota = f^{-1} \circ f = f \circ f^{-1}$ to derive a contradiction.]

Solution: Assume that f^{-1} is differentiable at $f(0)$ and let us apply the chain rule to differentiate $\iota = f^{-1} \circ f$ at 0. We find

$$\iota = Df^{-1}(f(0)) \circ Df(0).$$

Similarly, applying the chain rule to differentiate $\iota = f \circ f^{-1}$ at $f(0)$, we have:

$$\iota = Df(f^{-1}(f(0))) \circ Df^{-1}(f(0)) = Df(0) \circ Df^{-1}(f(0)).$$

We conclude that $Df(0)$ has both a left and right inverse and thus is invertible, however $\det Df(0) = 0$. This contradicts the assumption that f^{-1} is differentiable at $f(0)$.

Exercise 4.7. The non-linear system of equations

$$\begin{aligned} e^{xy}\sin(x^2-y^2+x) &= 0 \\ e^{x^2+y}\cos(x^2+y^2) &= 1 \end{aligned}$$

admits the solution $(x,y)=(0,0)$. Prove that there exists $\varepsilon>0$ such that for all (ξ,η) with $\xi^2+\eta^2<\varepsilon^2$, the perturbed system of equations

$$\begin{aligned} e^{xy}\sin(x^2-y^2+x) &= \xi \\ e^{x^2+y}\cos(x^2+y^2) &= 1+\eta \end{aligned}$$

has a solution $(x(\xi,\eta),y(\xi,\eta))$ which depends continuously on (ξ,η).

Solution: Let us define the maps

$$f^1(x,y)=e^{xy}\sin(x^2-y^2+x), \qquad f^2(x,y)=e^{x^2+y}\cos(x^2+y^2),$$

for $(x,y)\in\mathbb{R}^2$. Consider the map $f:\mathbb{R}^2\to\mathbb{R}^2$ defined as

$$f(x,y)=\begin{pmatrix} f^1(x,y) \\ f^2(x,y) \end{pmatrix}=\begin{pmatrix} e^{xy}\sin(x^2-y^2+x) \\ e^{x^2+y}\cos(x^2+y^2) \end{pmatrix}.$$

Then we have $f(0,0)=(0,1)$. We aim to employ the Inverse Function Theorem.

We compute the first partial derivatives of f, as

$$\begin{aligned} D_1f^1(x,y) &= ye^{xy}\sin(x^2-y^2+x)+(2x+1)e^{xy}\cos(x^2-y^2+x) \\ D_2f^1(x,y) &= xe^{xy}\sin(x^2-y^2+x)-2ye^{xy}\cos(x^2-y^2+x) \\ D_1f^2(x,y) &= 2xe^{x^2+y}\cos(x^2+y^2)-2xe^{x^2+y}\sin(x^2+y^2) \\ D_2f^2(x,y) &= e^{x^2+y}\cos(x^2+y^2)-2ye^{x^2+y}\sin(x^2+y^2) \end{aligned}$$

All these partial derivatives are continuous, so by a theorem in the lectures, f is continuously differentiable. Moreover, we have

$$Df(0,0)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is invertible. Thus, by the Inverse Function Theorem, there exist a neighbourhood $U\subset\mathbb{R}^2$ of $(0,0)$ and a neighbourhood $V\subset\mathbb{R}^2$ of $(0,1)$ such that $f:U\to V$ is a bijection.

Since V is an open neighbourhood of $(0,1)$, there is $\epsilon>0$ such that $B_\epsilon(0,1)\subseteq V$. It follows that all the points $(\xi,1+\eta)$ with $\xi^2+\eta^2<\varepsilon^2$ are elements of V. Thus, the inverse map

$$(x(\xi,\eta),y(\xi,\eta))=f^{-1}(\xi,1+\eta)$$

is well-defined and solves the perturbed system. The continuity of the map f^{-1} implies that $x(\xi,\eta)$ and $y(\xi,\eta)$ each vary continuously in (ξ,η) (see Exercise 1.8(b) on Problem Sheet 1).

Unseen Exercise. Find the minimum of the function $f:\mathbb{R}^3\to\mathbb{R}$ given by:

$$f(x,y,z)=x^4(y^2+x^2)+z^2-4z$$

Solution: Computing the partial derivatives, we have (setting $p = (x, y, z)$)

$$D_1 f(p) = 4x^3y^2 + 6x^5 = x^3(4y^2 + 6x^2)$$
$$D_2 f(p) = 2yx^4$$
$$D_3 f(p) = 2z - 4$$

We see that all partial derivatives are continuous, thus f is everywhere differentiable. If $p_0 = (x_0, y_0, z_0)$ is an extremal point, then $Df(p_0) = 0$. This implies that either

$$(x_0, y_0, z_0) = (0, 0, 2),$$

or

$$(x_0, y_0, z_0) = (0, y, 2),$$

for any value of $y \in \mathbb{R}$. In either of the above cases $f(p_0) = -4$. To see this is a minimum, note that

$$f(p) = x^4(y^2 + x^2) + z^2 - 4z = x^4(y^2 + x^2) + (z - 2)^2 - 4 \geq -4,$$

since the first two terms are manifestly positive.

Unseen Exercise. Let $\Omega = \{(x, y) \in \mathbb{R}^2 : x > 0\}$. Consider the function $f : \Omega \to \mathbb{R}^2$ given by:

$$f : (x, y) = (x \sin y, x \cos y).$$

(a) Show that f is differentiable at all $p = (\xi, \eta) \in \Omega$, with:

$$Df(p) = \begin{pmatrix} \sin \eta & \xi \cos \eta \\ \cos \eta & -\xi \sin \eta \end{pmatrix}.$$

Solution: Let $f^1(x, y) = x \sin y$ and $f^2(x, y) = x \cos y$. We can compute the partial derivatives at p and find

$$D_1 f^1(p) = \sin \eta, \qquad D_2 f^1(p) = \xi \cos \eta,$$
$$D_1 f^2(p) = \cos \eta, \qquad D_2 f^2(p) = -\xi \sin \eta.$$

These are all manifestly continuous functions of p, so we deduce that f is everywhere differentiable and:

$$Df(p) = \begin{pmatrix} \sin \eta & \xi \cos \eta \\ \cos \eta & -\xi \sin \eta \end{pmatrix},$$

by the theorem in the lectures.

(b) Show that $Df(p)$ is invertible for all $p \in \Omega$.

Solution: We have $\det Df(p) = -\xi \neq 0$ for $p = (\xi, \eta) \in \Omega$. Thus $Df(p)$ is invertible for all $p \in \Omega$.

(c) Show that $f : \Omega \to \mathbb{R}^2$ is not injective. Deduce that the restriction to open sets U, V in the inverse function theorem is necessary.

Solution: f is not injective, since (for example) the points $(1, 0)$ and $(1, 2\pi)$ are both mapped to $(0, 1)$ under f. This shows that even for a function whose derivative is globally invertible, we can nevertheless have that the function is not globally injective. Locally (i.e. restricted to small enough open sets) we do recover injectivity.