Questions marked with (*) are a bit more complex, you might want to skip them at the first read. Hints available in the next page.

4.1. BONUS PROBLEM. Consider the function $u\colon (x,y) \mapsto x^{\sin(y)}$, defined for $(x,y) \in (0,\infty) \times \mathbb{R} \subset \mathbb{R}^2$. Compute $\partial_x u$ and $\partial_y u$.

4.2. Connected graphs. Let $U \subset \mathbb{R}^n$ be open and connected and let $f \in C^1(U, \mathbb{R}^m)$. Show that its graph

$$\Gamma_f := \{(x, f(x)) : x \in U\}$$

is a connected subset of $\mathbb{R}^n \times \mathbb{R}^m$.

4.3. *p*–norms. For $p \geq 1$ and $x \in \mathbb{R}^n$ define the $p-$norm of x as

$$|x|_p := \Big(\sum_{i=1}^n |x_i|^p\Big)^{1/p}.$$

1. For $n = 2$ and $p = 1, 2, 10$ sketch the sets $\{x \in \mathbb{R}^2 : |x|_p \leq 1\}$.
2. For a given $x \in \mathbb{R}^n$, compute the limit $|x|_\infty := \lim_{p\to\infty} |x|_p$.
3. Using an appropriate inequality that you have seen in class, prove that

$$\Big(\sum_{i=1}^n a_i^{p-1} b_i\Big)^2 \leq \Big(\sum_{i=1}^n a_i^p\Big)\Big(\sum_{i=1}^n a_i^{p-2} b_i^2\Big),$$

   whenever a_i, b_i are n-tuples of positive numbers.
4. Fix $x, y \in \mathbb{R}^n$ and consider the function $f\colon [0,1] \to \mathbb{R}$ defined as

$$f(t) := |tx + (1-t)y|_p = \Big(\sum_{i=1}^n |tx_i + (1-t)y_i|^p\Big)^{1/p}. \tag{1}$$

   Show that f is convex. You may assume that the coordinates of x and y are all strictly positive and use the inequality of the previous point.
5. Deduce from the previous point that the triangular inequality holds, i.e.,

$$|x+y|_p \leq |x|_p + |y|_p \quad \text{for all } x, y \in \mathbb{R}^n.$$

6. What happens for $p \in (0,1)$?

4.4. *p*–means. For $x \in \mathbb{R}^n$ with positive coordinates and $p \neq 0$ define the p-mean as

$$\mu_p(x) := \Big(\frac{x_1^p + \ldots + x_n^p}{n}\Big)^{1/p}.$$

1. Compute the limits $p \to \pm\infty, p \to 0$ and define accordingly

$$\mu_{-\infty}(x), \quad \mu_0(x), \quad \mu_{+\infty}(x).$$

2. For any n-tuple of numbers $a_i > 0$ show that

$$\sum_{i=1}^{n} \frac{a_i \log(a_i)}{a_1 + \ldots + a_n} \geq \log\Big(\frac{a_1 \ldots + a_n}{n}\Big).$$

3. For a fixed x, show that the function $f\colon \mathbb{R} \to \mathbb{R}$, given by $f(t) := \mu_t(x)$, is continuous and increasing.

4. Prove the Arithmetic-Geometric inequality and Arithmetic-Quadratic inequality:

$$n(x_1 x_2 \cdots x_n)^{1/n} \leq x_1 + \ldots + x_n, \qquad (x_1 + \ldots + x_n)^2 \leq n(x_1^2 + \ldots + x_n^2).$$

5. (*) Is f continuously differentiable in the whole $\mathbb{R}$?

4.5. All norms are equivalent in $\mathbb{R}^n$. Let $|\cdot|$ denote the standard Euclidean norm in $\mathbb{R}^n$ and let $f\colon \mathbb{R}^n \to [0,\infty)$ be another norm (that is a function satisfying the properties of Definition 9.91).

1. Expressing x in a basis and using the "abstract" properties that f must have, show that there is a constant $C_1 > 0$ such that

$$f(x) \leq C_1 |x| \text{ for all } x \in \mathbb{R}^n.$$

2. Show that f is continuous in $\mathbb{R}^n$ (with respect to the standard distance of $\mathbb{R}^n$!).

3. Show that there is a number $c_2 > 0$ such that

$$f(x) \geq c_2 \text{ for all } |x| = 1.$$

4. Conclude that $f(x) \geq c_2|x|$ for all $x \in \mathbb{R}^n$.

5. Show that if $\tilde{f}$ is yet another norm, then there is $C > 0$ such that

$$C^{-1} f(x) \leq \tilde{f}(x) \leq C f(x) \text{ for all } x \in \mathbb{R}^n.$$

4.6. Hilbert Schmidt norm of the composition. Take two linear functions $\phi\colon \mathbb{R}^d \to \mathbb{R}^n$ and $\psi\colon \mathbb{R}^n \to \mathbb{R}^m$, and denote with Φ, Ψ the matrices that represent them in the canonical basis'. Recall that the linear map $\psi \circ \phi\colon \mathbb{R}^d \to \mathbb{R}^m$ is represented in these basis' by the matrix $\Psi \cdot \Phi$. Show that

$$\|\Psi \cdot \Phi\| \leq \|\Psi\| \|\Phi\|,$$

where $\|\cdot\|$ is the Hilbert-Schmidt norm of a matrix (see 10.1.3 in the notes).

4.7. Mean value for vector-valued functions. Let $f \in C^1(\mathbb{R}, \mathbb{R}^m)$ for $m > 1$. Is it true that there is $t \in [0, 1]$ such that

$$f(1) - f(0) = Df_t(1) = \begin{bmatrix} f_1'(t) \\ \vdots \\ f_m'(t) \end{bmatrix}?$$

Prove it or provide a counterexample.

4.8. A directional derivative vanish. Let $u \in C^1(\mathbb{R}^n)$ and $\nu \in \mathbb{R}^n$. Show that

1. If $\partial_1 u \equiv 0$ then "u does not depend on x_1", more rigorously: there exists a unique function $v \in C^1(\mathbb{R}^{n-1})$ such that

$$u(x_1, \dots, x_n) = v(x_2, \dots, x_n) \text{ for all } x \in \mathbb{R}^n. \tag{2}$$

2. If $\partial_\nu u \equiv 0$ and $\nu \cdot e_1 \neq 0$ then "u is a function of $n-1$ variables", more rigorously: there exists a unique function $w \in C^1(\mathbb{R}^{n-1})$ such that

$$u(x_1, \dots, x_n) = w(x_2 - \tfrac{x_1 \nu_2}{\nu_1}, \dots, x_n - \tfrac{x_1 \nu_n}{\nu_1}) \text{ for all } x \in \mathbb{R}^n.$$

3. (*) What can we conclude if we assume only that $\partial_1 u = 0$ in an open connected subset $U \subset \mathbb{R}^n$?

Hints:

4.3.3 Use Cauchy–Schwartz and the fact that $a_i^{p-1}b_i = a_i^{p/2} \cdot a_i^{(p-2)/2}b_i$.

4.3.4 Show that $f''(t) \geq 0$, it is not the simplest derivative, but if you get it right the inequality in 4.2.3 will be just what you need to prove that it $f'' \geq 0$.

4.3.5 Write the convexity inequality between x, y and the middle point $(x+y)/2$. To get the general case from the one with positive coordinates observe the following. For $x \in \mathbb{R}^n$ denote with $\hat{x} := (|x_1|, \ldots, |x_n|)$, then

$$|x+y|_p \leq |\hat{x}+\hat{y}|_p, \quad |\hat{x}|_p = |x|_p, \quad \text{for all } x, y \in \mathbb{R}^n.$$

4.4.2 Combine the concavity of the $\log(\cdot)$ and the Cauchy–Schwarz inequality. You need to use the following little generalisation of the concavity inequality: for any concave function $f\colon \mathbb{R} \to \mathbb{R}$:

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n) \geq \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n)$$
$$\text{for all } x_i \in \mathbb{R}, 0 \leq \lambda_i \leq 1 \text{ with } \lambda_1 + \ldots + \lambda_n = 1.$$

4.4.3 It might be more convenient to work with $\log f(t)$, be careful with the computation, once again the inequality in 4.3.2 will be just what you need to prove that $f' \geq 0$.

4.5.2 Recall that from the definition it follows that $|f(x) - f(y)| \leq f(x-y)$... And that Lipschitz functions are always continuous.

4.5.3 Apply Weierstrass Theorem to f on $S := \{x \in \mathbb{R}^n : |x| = 1\}$ and use that f, being a norm, is nondegenerate.

4.5.5 If f is equivalent to $|\cdot|$ and $\tilde{f}$ is equivalent to $|\cdot|$ it follows by transitivity that f is equivalent to $\tilde{f}$...

4.6 First recall that $\|M\|^2$ is the sum of the squares of the entries of M. Notice that $(\Psi \cdot \Phi)^i_j$ (ith row and jth column) is the scalar product of Ψ^i and Φ_j which are vectors of $\mathbb{R}^n$. Apply Cauchy-Schwartz to each of them.

4.7 Try $f(t) = (\sin(2\pi t), \cos(2\pi t))$...

4.8.2 Apply the mean value theorem to $u(x + t\nu)$...

4.8.3 Consider the domain $U := \{(x, y) : x^2 < y < x^2 + 1\}$ and the function

$$u(x,y) := \begin{cases} \max\{0, y-2\}^2 & \text{if } x \geq 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

4. Solutions

Solution of 4.1: $\partial_x u = \sin(y)x^{\sin(y)-1}$ and $\partial_y u = \log(x)\cos(y)x^{\sin(y)}$.

Solution of 4.2: We claim that Γ_f is in fact path connected. Let $(x_0, f(x_0))$ and $(x_1, f(x_1)$ be two points in Γ_f. Since $U \subset \mathbb{R}^n$ is open and connected it is path connected, so there is $\gamma\colon [0,1] \to U$ such that $\gamma(0) = x_0, \gamma(1) = x_1$. Then the path

$$[0,1] \ni t \mapsto (\gamma(t), f(\gamma(t))$$

connects $(x_0, f(x_0))$ with $(x_1, f(x_1))$.

Solution of 4.3:

1.
2. $|x|_\infty = \max\{|x_1|, \ldots, |x_n|\}$, let's prove it. Assume that $|x_1| \geq |x_j|$ for all j's, then

$$|x_1| \leq |x|_p = |x_1|\Big(1 + \tfrac{|x_n|^p}{|x_1|^p} + \ldots + \tfrac{|x_n|^p}{|x_1|^p}\Big)^{1/p} \leq |x_1|(1 + 1 + \ldots + 1)^{1/p} = |x_1|n^{1/p},$$

and $|x_1|n^{1/p} \to |x_1|$ as $p \to \infty$.

3. This is Cauchy-Schwarz inequality in disguise:

$$\Big(\sum_{i=1}^n a_i^{p-1}b_i\Big)^2 = \Big(\sum_{i=1}^n \underbrace{a_i^{\frac{p}{2}-1}b_i}_{=:x_i}\,\underbrace{a_i^{\frac{p}{2}}}_{=:y_i}\Big)^2 \leq \Big(\sum_{i=1}^n y_i^2\Big)\Big(\sum_{i=1}^n x_i^2\Big) = \Big(\sum_{i=1}^n a_i^p\Big)\Big(\sum_{i=1}^n a_i^{p-2}b_i^2\Big),$$

notice that we did not use $p \geq 1$, any $p \in \mathbb{R}$ would have worked.

4. We compute

$$f'(t) = \tfrac{1}{p}f(t)^{1-p}\frac{d}{dt}\Big(\sum_{i=1}^n |tx_i + (1-t)y_i|^p\Big) = f(t)^{1-p}\sum_{i=1}^n |tx_i + (1-t)y_i|^{p-1}(x_i - y_i),$$

and then

$$\begin{aligned} f''(t) = (1-p)f(t)^{-p}f'(t)\sum_{i=1}^n &|tx_i + (1-t)y_i|^{p-1}(x_i - y_i) \\ &+ (p-1)f(t)^{1-p}\sum_{i=1}^n |tx_i + (1-t)y_i|^{p-2}(x_i - y_i)^2, \end{aligned}$$

so $f'' \geq 0$ if and only if

$$\frac{f'(t)}{f(t)^{1-p}}\sum_{i=1}^n |tx_i + (1-t)y_i|^{p-1}(x_i - y_i) \leq f(t)^p\sum_{i=1}^n |tx_i + (1-t)y_i|^{p-2}(x_i - y_i)^2$$

which setting $a_i := |tx_i + (1-t)y_i|, b_i := x_i - y_i$, and substituting f, f' becomes

$$\Big(\sum_{i=1}^n a_i^{p-1}b_i\Big)^2 \leq \Big(\sum_{i=1}^n a_i^p\Big)\Big(\sum_{i=1}^n a_i^{p-2}b_i^2\Big),$$

which we proved at the previous point. Notice that we proved that

$$f''(t) = (p-1) \times \{\text{something } \geq 0\}$$

so we proved that if $p < 1$ then f is concave.

5. If the coordinates of x, y are non-negative we use the convexity inequality on f and get
$$\Big|\frac{x+y}{2}\Big|_p = f(1/2) \leq \frac{f(0)+f(1)}{2} = \frac{|x|_p + |y|_p}{2},$$
simplifying the factors 2 we get the triangle inequality.

   In the general case we use the trick in the hint and find
$$|x+y|_p \leq |\hat{x}+\hat{y}|_p \leq |\hat{x}|_p + |\hat{y}|_p = |x|_p + |y|_p.$$

6. Since f is concave for $p < 1$, one has the converse inequality:
$$|x+y|_p \geq |x|_p + |y|_p.$$

Solution of 4.4:

1. This is similar to 4.3.2: $\mu_{+\infty}(x) = \max\{x_1, \dots, x_n\}$, let's prove it. Assume that $x_1 \geq x_j$ for all j's, then
$$x_1 \geq \mu_p(x) = x_1\Big(1 + \tfrac{x_n^p}{x_1^p} + \ldots + \tfrac{x_n^p}{x_1^p}\Big)^{1/p} n^{-1/p} \geq x_1 n^{-1/p},$$
and $n^{-1/p} \to 1$ as $p \to \infty$.

   Noticing that $1/\mu_p(x) = \mu_{-p}(1/x)$ we find
$$\mu_{-\infty}(x) = \lim_{p\to-\infty} \mu_p(x) = \lim_{q\to+\infty} (\mu_{-q}(1/x))^{-1} = \Big(\max_i 1/x_i\Big)^{-1} = \Big(\frac{1}{\min_i x_i}\Big)^{-1} = \min_i x_i.$$

   Finally $\mu_0(x) = (x_1 \cdots x_n)^{1/n}$. Recall
$$y^t = e^{t\log y} = 1 + t\log(y) + O(t^2) \quad \text{ as } t \to 0, \text{ with } y > 0 \text{ fixed.}$$

   Thus taking a log we find
$$\begin{aligned}\log \mu_p(x) &= \frac{1}{p}\log\frac{x_1^p + \ldots + x_n^p}{n} = \frac{1}{p}\log\Big(\frac{1+\log(x_1)p + \ldots + 1 + \log(x_n)p + O(p^2)}{n}\Big)\\ &= \frac{1}{p}\log\Big(1 + \frac{p}{n}\{\log(x_1) + \ldots + \log(x_n)\} + O(p^2)\Big)\\ &= \frac{1}{p}\log\Big(1 + \frac{p}{n}\log(x_1\cdots x_n) + O(p^2)\Big) = \frac{1}{n}\log(x_1\cdots x_n) + O(p).\end{aligned}$$

2. Using the convexity inequality of the hint with the convex function $x \mapsto x\log(x)$ and
$$x_i := a_i, \quad \lambda_i := \frac{1}{n}, \text{ for all } i = 1, \ldots, n,$$
we find that
$$\frac{1}{n}\sum_{i=1}^{n} a_i \log(a_i) \geq \frac{a_1 + \ldots + a_n}{n} \log\Big(\frac{a_1 + \ldots + a_n}{n}\Big),$$
which is what we wanted up to reshuffling the terms.

3. $f(t)$ is continuous at all $t \neq 0$, since it is given by an analytic formula. At $t = 0$ we showed that the (bilateral) limit $\lim_{t\to 0} f(t) = \lim_{t\to 0} \mu_t(x) = \mu_0(x) = f(0)$ exists finite, so by Analysis I, $f(t)$ is continuous on the whole $\mathbb{R}$. Computing the logarithmic derivative
$$\begin{aligned}(\log f(t))' &= \frac{d}{dt}\frac{1}{t}\{\log(x_1^t + \ldots + x_n^t) - \log(n)\}\\ &= -\frac{1}{t^2}\log\Big(\frac{x_1^t + \ldots + x_n^t}{n}\Big) + \frac{1}{t}\frac{\frac{d}{dt}(x_1^t + \ldots + x_n^t)}{(x_1^t + \ldots + x_n^t)}\\ &= -\frac{1}{t^2}\log\Big(\frac{x_1^t + \ldots + x_n^t}{n}\Big) + \frac{1}{t^2}\frac{\log(x_1^t)x_1^t + \ldots + \log(x_n^t)x_n^t}{(x_1^t + \ldots + x_n^t)},\end{aligned}$$
which is nonnegative by the inequality of the previous point with $a_i = x_i^t$. This computation is rigorous at least for $t \neq 0$, and shows that $f(t)$ is increasing in $(-\infty, 0) \cup (0, +\infty)$, thus (being continuous) it is increasing on the whole $\mathbb{R}$.

4. The AM-GM inequality is equivalent to $f(0) \leq f(1)$ while the $AM - QM$ inequality is equivalent to $f(1) \leq f(2)$. Since we proved that f is increasing both are true.

5. In order to understand if f is of class C^1 we have to understand whether f' extends continuously at $t = 0$. Since $(\log f)' = f'/f$ and $f(0) > 0$ it is the same to show that $(\log f)'$ extends continuously at $t = 0$ so we pick the previous formula and expand
$$\begin{aligned}(\log f(t))' &= \frac{1}{t^2}\Big\{\frac{\log(x_1^t)x_1^t + \ldots + \log(x_n^t)x_n^t}{(x_1^t + \ldots + x_n^t)} - \log\Big(\frac{x_1^t + \ldots + x_n^t}{n}\Big)\Big\}\\ &= \frac{1}{t^2}\Big\{\frac{\sum_{i=1}^n \log(x_i^t) + \log(x_i^t)^2 + O(t^3)}{\sum_{i=1}^n 1 + \log(x_i^t) + \log(x_i^t)^2 + O(t^3)} - \log\Big(1 + \sum_{i=1}^{n} \tfrac{1}{n}\log(x_i^t) + \tfrac{1}{n}\log(x_i^t)^2 + O(t^3)\Big)\Big\}\end{aligned}$$
where we used that for fixed $z > 0$ it holds, as $t \to 0$,
$$\log(z^t)y^t = t\log(z) + t^2\log(z)^2 + O(t^3), \quad z^t = 1 + t\log(z) + t^2\log(z)^2 + O(t^3).$$

Now we short the notation to $y_i := \log(x_i)$ and compute everything

$$\begin{aligned}
(\log f(t))' &= \frac{1}{t^2}\Big\{\frac{\sum_{i=1}^n ty_i + t^2y_i^2 + O(t^3)}{\sum_{i=1}^n 1 + ty_i + t^2y_i^2 + O(t^3)} - \log\Big(1 + \sum_{i=1}^n \tfrac{t}{n}y_i + \tfrac{t^2}{n}y_i^2 + O(t^3)\Big)\Big\} \\
&= \frac{1}{t^2}\Big\{\Big(\sum_{i=1}^n ty_i + t^2y_i^2 + O(t^3)\Big)\Big(\tfrac{1}{n} - \tfrac{1}{n^2}\sum_{i=1}^n ty_i + O(t^2)\Big) \\
&\qquad - \sum_{i=1}^n(\tfrac{t}{n}y_i + \tfrac{t^2}{n}y_i^2) + \frac{1}{2}\Big(\sum_{i=1}^n \tfrac{t}{n}y_i\Big)^2 + O(t^3)\Big\} \\
&= \frac{1}{t^2}\Big\{\sum_{i=1}^n(\tfrac{t}{n}y_i + \tfrac{t^2}{n}y_i^2) - \Big(\sum_{i=1}^n \tfrac{t}{n}y_i\Big)^2 - \sum_{i=1}^n(\tfrac{t}{n}y_i + \tfrac{t^2}{n}y_i^2) + \frac{1}{2}\Big(\sum_{i=1}^n \tfrac{t}{n}y_i\Big)^2 + O(t^3)\Big\} \\
&= -\frac{1}{2}\Big(\tfrac{1}{n}\sum_{i=1}^n y_i\Big)^2 + O(t),
\end{aligned}$$

where we used the expansions

$$\frac{1}{n+t} = \frac{1}{n} - \frac{t}{n^2} + O(t^2), \quad \log(1+t) = t - \frac{t^2}{2} + O(t^3).$$

Thus substituting back the y_i's we proved

$$-2f'(t) = f(t)\Big\{(\log f(0))^2 + O(t)\Big\} \quad \text{as } t \to 0,$$

which proves that f is continuously differentiable also at $t = 0$.

Solution of 4.5: See Theorem 9.107 in the notes.

Solution of 4.6: For $i \in \{1, \ldots, m\}$ we denote with Ψ^i the vector $(\Psi^i_\ell)_{1\le\ell\le n}$ in $\mathbb{R}^n$. For $j \in \{1, \ldots, d\}$ we denote with Φ_j the vector $(\Phi^\ell_j)_{1\le\ell\le n}$ in $\mathbb{R}^n$. Notice that by definition of matrix multiplication

$$(\Psi \cdot \Phi)^i_j = \Psi^i \cdot \Phi_j.$$

Now, by Cauchy-Schwarz inequality in $\mathbb{R}^n$ we have

$$\|\Psi \cdot \Phi\|^2 = \sum_{i=1}^m\sum_{j=1}^d |\Psi^i \cdot \Phi_j|^2 \le \sum_{i=1}^m\sum_{j=1}^d |\Psi^i|^2\,|\Phi_j|^2 = \Big(\sum_{i=1}^m |\Psi^i|^2\Big)\Big(\sum_{j=1}^d |\Phi_j|^2\Big) = \|\Psi\|^2\|\Phi\|^2.$$

Solution of 4.7: It is false. If $f(t) = (\sin(2\pi t), \cos(2\pi t))$ then $f(1) - f(0) = 0 \in \mathbb{R}^m$, but $|f'(t)| = 2\pi|(\cos(2\pi t), -\sin(2\pi t))| = 2\pi \neq 0$.

Solution of 4.8:

1. Set $v(x_2, \ldots, x_n) := u(0, x_2, \ldots, x_n)$ for all $(x_2, \ldots, x_n) \in \mathbb{R}^{n-1}$. By definition of directional derivative, for each fixed $(x_2, \ldots, x_n) \in \mathbb{R}^{n-1}$, we have that

$$\frac{d}{dt}\Big(u(t, x_2, \ldots, x_n) - v(x_2, \ldots, x_n)\Big) = \partial_1 u(t, x_2, \ldots, x_n) = 0,$$

thus $t \mapsto u(t, x_2, \ldots, x_n) - v(x_2, \ldots, x_n)$ is constant. Since it vanishes at $t = 0$ it is constantly zero, proving what we wanted.

2. Following the same reasoning: consider the difference $u(x) - u(x + t\nu)$, freeze the x and prove that t derivative with respect to t is zero. We find that $u(x) = u(x + t\nu)$ for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Now, choosing $-t := x_1/\nu_1$, we make the first coordinate vanish and find the sought formula.

3. Consider the function u and the domain U of the hint. We compute the partial derivatives:
$$\partial_1 u = 0, \quad \partial_2 u = 2\max\{0, y - 2\},$$
which are manifestly continuous, so — by the sufficient condition seen in class — we have $u \in C^1(U)$. On the other hand u cannot be written as a function of y only since $2.5^2 = u(2, 4.5) \neq u(-2, 4.5) = 0$.

   The representation formula (2) holds if U has a special structure, namely it is "convex in the x_1-direction" that is
$$x, y \in U, x - y \in \mathbb{R}e_1, t \in [0, 1] \Rightarrow tx + (1 - t)y \in U.$$
   Geometrically this means that if we slice U with the plane $\{x_1 = c\}$ then we can connect any point in U with this plane following a path aligned with e_1.