In the following, we write $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_{++} = (0, \infty)$. So $\mathbb{R}_n^+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0\}$ and $\mathbb{R}_n^{++} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_i > 0\}$.

As usual, no taking partial derivative, no quoting nonsense like Jacobi formula; everything here requires nothing more than definition and chain rule. You may freely quote results from previous homework and lecture notes.

1. Let $A_0, A_1, \ldots, A_n \in S^m$ and $\Omega = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succ 0\}$.
   (a) Find the gradient of $f: \Omega \to \mathbb{R}$, 
   $$f(x) = \det(A_0 + x_1 A_1 + \cdots + x_n A_n).$$
   (b) Find the Hessian of $f: \Omega \to \mathbb{R}$, 
   $$f(x) = \log \det(A_0 + x_1 A_1 + \cdots + x_n A_n).$$
   Recall that we have already found the gradient of this function in the lectures.
   (c) Find the gradient and Hessian of $f: \Omega \to \mathbb{R}$, 
   $$f(x) = \text{tr}((A_0 + x_1 A_1 + \cdots + x_n A_n)^{-1}).$$
   (d) Find the gradient of $f: \Omega \to \mathbb{R}$, 
   $$f(x) = (Bx + c)^\top(A_0 + x_1 A_1 + \cdots + x_n A_n)^{-1}(Bx + c)$$
   where $B \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^m$.

2. Decide which of the following sets are convex. Prove your answers.
   - $\text{GL}(n) = \{X \in \mathbb{R}^{n \times n} : \det(X) \neq 0\}$,
   - $S_n^+ = \{X \in S^n : X \succ 0\}$,
   - $\Omega_1 = \{x \in \mathbb{R}^n : A_0 + x_1 A_1 + \cdots + x_n A_n \succ 0\}$,
   - $\Omega_2 = \{X \in \mathbb{R}^{m \times n} : X^\top A X + B^\top X + X^\top B + C \succ 0\}$,
   where $\Omega_1$ is as defined in Problem 1 and $\Omega_2$ is as defined in Homework 2, Problem 4(f).

3. Compute the Hessians of the following functions and decide if they are convex, concave, or neither on their respective domains.
   (a) $f: \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}$ defined by 
   $$f(x, y) = \frac{x^2}{y}.$$ 
   (Hint: Write $\nabla^2 f(x, y)$ as a rank-1 matrix).
   (b) $f: \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}$ defined by 
   $$f(x, y) = \frac{x^\top x}{y}.$$ 
   (c) $f: \mathbb{R}^n \times S_n^+ \to \mathbb{R}$ defined by 
   $$f(x, Y) = x^\top Y^{-1} x.$$
(d) $f : \mathbb{S}_{++}^n \rightarrow \mathbb{R}$ defined by
\[
f(X) = \log \det(X) - \log \text{tr}(X).
\]

(e) $f : \Omega \rightarrow \mathbb{R}$ defined by
\[
f(x) = \|Ax + b\|^2_{c^T x + d}
\]
where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}$, and $\Omega = \{x \in \mathbb{R}^n : c^T x + d > 0\}$.

4. (a) Find the Hessian of the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$. Show that for any $v \in \mathbb{R}^n$,
\[
v^T \nabla^2 f(x)v = \frac{1}{(e^{x_1} + \cdots + e^{x_n})^2} \left[ \sum_{i=1}^n e^{x_i} \left( \sum_{i=1}^n v_i^2 e^{x_i} \right) - \left( \sum_{i=1}^n v_i e^{x_i} \right)^2 \right].
\]
Hence or otherwise, deduce that $f$ is a convex function.

(b) Find the Hessian of the function $g : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$ defined by $g(x) = (x_1 \cdots x_n)^{1/n}$. Show that for any $v \in \mathbb{R}^n$,
\[
v^T \nabla^2 g(x)v = -\frac{g(x)}{n^2} \left[ n \sum_{i=1}^n \frac{v_i^2}{x_i} - \left( \sum_{i=1}^n \frac{v_i}{x_i} \right)^2 \right]
\]
Hence or otherwise, deduce that $g$ is a concave function.

(c) Find the Hessian of the function $h : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$ defined by
\[
h(x) = \frac{1}{1/x_1 + \cdots + 1/x_n}.
\]
By emulating what we did in the previous two parts or otherwise, decide if $h$ is convex, concave, or neither on $\mathbb{R}_{++}^n$.

(d) Find the Hessian of the function $\varphi : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$ defined by $\varphi(x) = \log h(x)$. Decide if $\varphi$ is convex, concave, or neither on $\mathbb{R}_{++}^n$.

5. (a) Show that the negative log function $-\log : \mathbb{R}_{++} \rightarrow \mathbb{R}$ is strictly convex, i.e.,
\[
\log(t x + (1 - t) y) > t \log x + (1 - t) \log y
\]
for any $x, y \in \mathbb{R}_{++}$ and any $t \in (0, 1)$.

(b) Prove the generalized arithmetic-geometric mean inequality
\[
a^t b^{1-t} \leq t a + (1 - t) b
\]
for any $a, b \in \mathbb{R}_+$ and $t \in [0, 1]$ (note that $t = 1/2$ gives us the usual arithmetic-geometric mean inequality). Deduce the Hölder inequality: for $p > 1$ and $1/p + 1/q = 1$,
\[
\sum_{i=1}^n |x_i y_i| \leq \left( \sum_{i=1}^n |x_i|^p \right)^{1/p} \left( \sum_{i=1}^n |y_i|^q \right)^{1/q}
\]
for any $x, y \in \mathbb{R}^n$ (note that $p = q = 2$ gives us the Cauchy–Schwartz inequality).

(c) Show that $(\sin \theta)^{\sin \theta} < (\cos \theta)^{\cos \theta}$ for all $\theta \in (0, \pi/4)$.