**ECE 6270, Spring 2021**

**Homework #2 – SOLUTIONS**

1. Prepare a one paragraph summary of what we talked about in the last week of class. I do not want just a bulleted list of topics, I want you to use complete sentences and establish context (Why is what we have learned relevant? How does it connect with other classes?). The more insight you give, the better.

   Solution: Make sure that what is written is in coherent, complete sentences.

2. Provide feedback to your peers on Homework #1 in Canvas.

3. A function $f(\boldsymbol{x}) : \mathbb{R}^N \to \mathbb{R}$ is *concave* if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$,

$$f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \;\geq\; \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y}), \quad \text{for all} \;\; 0 \leq \theta \leq 1.$$

   Give a simple yet rigorous argument that

$$f(\boldsymbol{x}) \text{ is concave} \quad \Leftrightarrow \quad -f(\boldsymbol{x}) \text{ is convex.}$$

   Solution: Suppose $f(\boldsymbol{x})$ is concave. Then

$$f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \geq \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y})$$

   and thus

$$\begin{aligned} -f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) &\leq -\theta f(\boldsymbol{x}) - (1-\theta) f(\boldsymbol{y}) \\ &= \theta(-f(\boldsymbol{x})) + (1-\theta)(-f(\boldsymbol{y})), \end{aligned}$$

   and so $-f(\boldsymbol{x})$ is convex. By a similar reasoning we have that if $-f(\boldsymbol{x})$ is convex then

$$-f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta(-f(\boldsymbol{x})) + (1-\theta)(-f(\boldsymbol{y})),$$

   and thus

$$f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \geq \theta(f(\boldsymbol{x})) + (1-\theta)(f(\boldsymbol{y})),$$

   and hence $f(\boldsymbol{x})$ is concave.

4. Recall that a norm is a function $\|\cdot\| : \mathbb{R}^N \to \mathbb{R}$ which obeys

   - $\|\boldsymbol{x}\| \geq 0$ and $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$.
   - $\|a\boldsymbol{x}\| = |a| \, \|\boldsymbol{x}\|$ for all $\boldsymbol{x} \in \mathbb{R}^N$ and scalars $a \in \mathbb{R}$.
   - $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$.

   (a) Suppose that $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$, where $\|\cdot\|$ denotes any valid norm. Prove that $f(\boldsymbol{x})$ is convex (using only the properties above).
   Solution: For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ and any $\theta \in [0,1]$ we have

$$\begin{aligned} \|\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}\| &\leq \|\theta\boldsymbol{x}\| + \|(1-\theta)\boldsymbol{y}\| \\ &= |\theta|\|\boldsymbol{x}\| + |1-\theta|\|\boldsymbol{y}\| \\ &= \theta\|\boldsymbol{x}\| + (1-\theta)\|\boldsymbol{y}\|, \end{aligned}$$

   which is precisely the definition of convexity.

(b) Can $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ ever be *strictly* convex? If so, give an example of such a norm. If not, provide a proof that no norm can be strictly convex.

Solution: No norm can be strictly convex. To see why, consider any $\boldsymbol{x}$ and $\boldsymbol{y}$ that are co-linear, meaning that $\boldsymbol{y} = \alpha\boldsymbol{x}$ for some $\alpha \in \mathbb{R}$ with $\alpha \neq 1$. Note that for such $\boldsymbol{x}$ and $\boldsymbol{y}$ we have that for any $\theta \in (0,1)$,

$$\|\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}\| = \|\theta\boldsymbol{x} + (1-\theta)\alpha\boldsymbol{x}\| = (\theta + (1-\theta)\alpha)\|\boldsymbol{x}\|. \tag{1}$$

We also have

$$\theta\|\boldsymbol{x}\| + (1-\theta)\|\boldsymbol{y}\| = \theta\|\boldsymbol{x}\| + (1-\theta)\|\alpha\boldsymbol{x}\| = (\theta + (1-\theta)\alpha)\|\boldsymbol{x}\|. \tag{2}$$

Thus, (1) and (2) are equal, but for $\|\cdot\|$ to be *strictly* convex, the former would have to be strictly less than the latter, and hence $\|\cdot\|$ cannot be strictly convex. Note, however, that $\|\cdot\|_2^2$ *is* strictly convex.

5. The $\alpha$-sublevel set of a function $f : \mathbb{R}^N \to \mathbb{R}$ is the set $S_\alpha = \{\boldsymbol{x} \ : \ f(\boldsymbol{x}) \leq \alpha\}$.

(a) Suppose $f$ is convex. Show that $S_\alpha$ is convex for all $\alpha \in \mathbb{R}$.

Solution: Let $\alpha \in \mathbb{R}$, and take $\boldsymbol{x}, \boldsymbol{y} \in S_\alpha$. Note that $f(\boldsymbol{x}) \leq \alpha$, and $f(\boldsymbol{y}) \leq \alpha$. Let $\theta \in [0,1]$. Because $f$ is convex,

$$\begin{aligned} f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) &\leq \theta f(\boldsymbol{x}) + (1-\theta)f(\boldsymbol{y}) \\ &\leq \theta\alpha + (1-\theta)\alpha \\ &= \alpha \end{aligned}$$

so $\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y} \in S_\alpha$. This holds for all such $\boldsymbol{x}$, $\boldsymbol{y}$, and $\theta$, so $S_\alpha$ is convex.

(b) Suppose $f$ is convex. Show that the set of global minimizers of $f$ is a convex set.

Solution: Let $\boldsymbol{x}^\star$ be any global minimum. The set of global minimizers is exactly the sublevel set $S_\alpha$ for $\alpha = f(\boldsymbol{x}^\star)$.

(c) Recall that the unit ball of a norm is the set $\mathcal{B} = \{\boldsymbol{x} : \|\boldsymbol{x}\| \leq 1\}$. Show that the unit ball of any norm must be convex.

Solution: The unit ball $\mathcal{B}$ is simply the sublevel set $S_1$ for the function $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$.

(d) *Optional:* Suppose $S_\alpha$ is convex for all $\alpha \in \mathbb{R}$. Is $f$ convex? Prove or find a counterexample.

Solution: Not necessarily. For example, any function that is monotonically increasing has convex sublevel sets, but may or may not be convex.

6. (a) Let $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ be convex functions on $\mathbb{R}^N$. Show that

$$f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$$

is convex.

Solution: Since $f_1(\boldsymbol{x})$ is convex, we have that for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ and for any $\theta \in [0,1]$, we have that

$$f_1(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta f_1(\boldsymbol{x}) + (1-\theta)f_1(\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + (1-\theta)f(\boldsymbol{y}),$$

where the last inequality follows since $f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\} \geq f(\boldsymbol{x})$. By the exact same argument, we have

$$f_2(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y}),$$

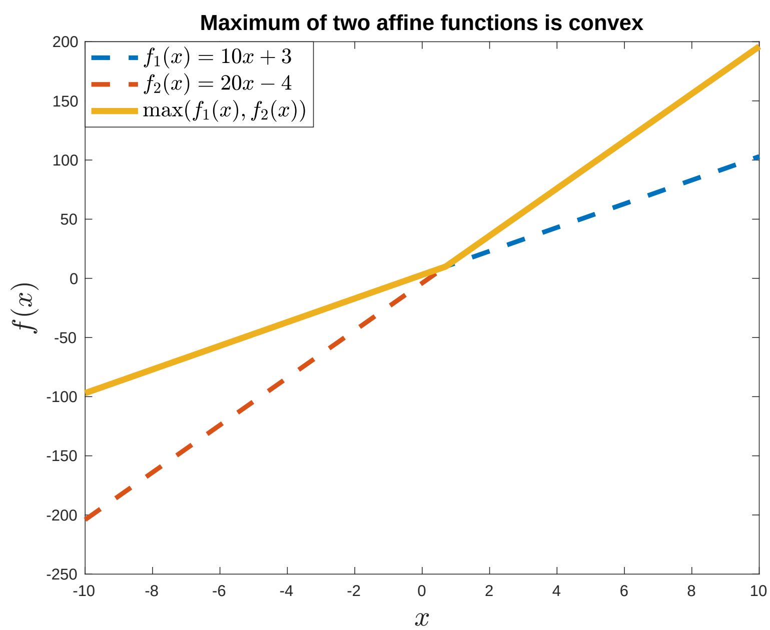
If *both* $f_1(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y})$ and $f_1(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y})$ are less than $\theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y})$, then we in fact have that

$$f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y}) = \max(f_1(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y}), f_2(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y})) \leq \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y}),$$

which shows that $f$ is convex.

(b) Illustrate the above in $\mathbb{R}^1$ by making a sketch with affine functions $f_1(x) = a_1 x + b_1$ and $f_2(x) = a_2 x + b_2$. You may choose $a_1, b_1, a_2, b_2$ to your liking.

Solution: Here is an example.



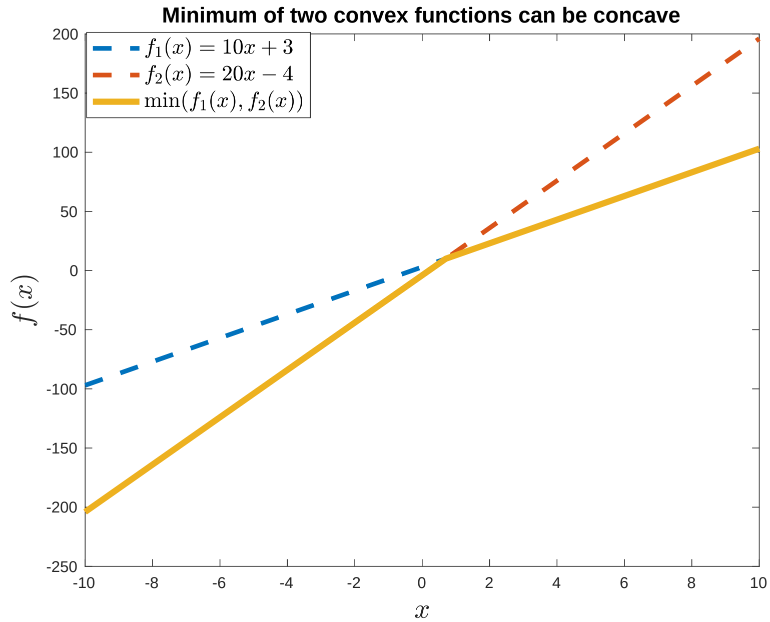

The yellow plot shows the maximum of two affine functions. This maximum is a convex function as any line connecting two points on the yellow plot would be above the yellow plot.

(c) Is it necessarily true that

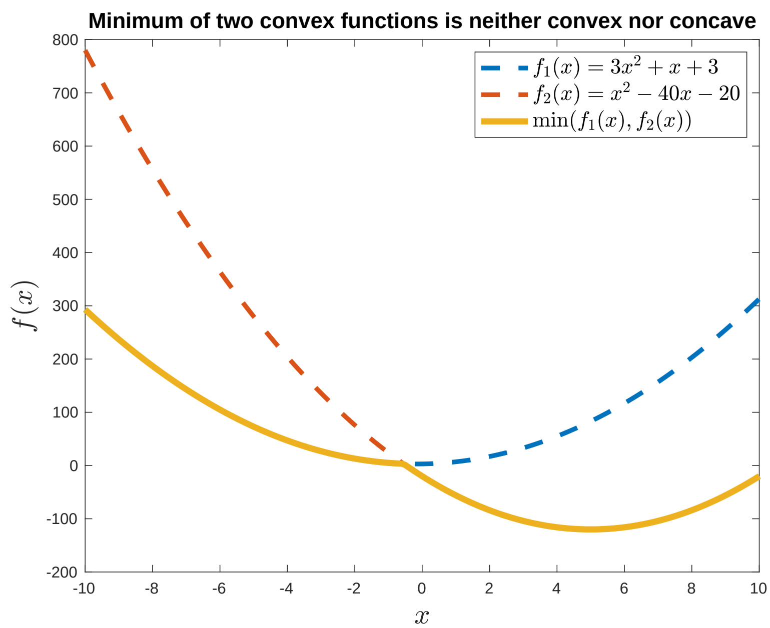
$$f(\boldsymbol{x}) = \min\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}, \quad f_1, f_2 \ \text{ convex},$$

is convex? Sketch an example in $\mathbb{R}^1$ that supports your argument.

Solution: No. $f(\boldsymbol{x})$ *might* be concave, but in general it will be neither convex nor concave (e.g., consider the minimum of two shifted quadratic functions). Here we provide two examples.




The minimum of two convex functions *can* be concave, as illustrated above. However, it isn't always, as illustrated in the figure below.




7. Recall that we use $\mathbb{S}_+^N$ to denote the set of $N \times N$ matrices that are symmetric and whose eigenvalues are non-negative.

   (a) Consider the function
   $$f(\boldsymbol{X}) = \min_{\boldsymbol{v}:\|\boldsymbol{v}\|_2=1} \boldsymbol{v}^{\mathrm{T}} \boldsymbol{X} \boldsymbol{v}.$$
   Note that $f$ maps any arbitrary (not necessarily symmetric) matrix to a single real number. Show that $f(\boldsymbol{X})$ is concave.

Solution: If $\boldsymbol{X}$ is symmetric, then $f(\boldsymbol{X})$ returns the smallest eigenvalue of $\boldsymbol{X}$ (if $\boldsymbol{X}$ is not symmetric, then it returns some other real number). We will show that $f_M(\boldsymbol{X})$ is concave over the entire space of $N \times N$ matries. We have

$$\begin{aligned} f_M(\theta\boldsymbol{X} + (1-\theta)\boldsymbol{Y}) &= \min_{\|\boldsymbol{v}\|_2=1} \left[\theta\boldsymbol{v}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{v} + (1-\theta)\boldsymbol{v}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{v}\right] \\ &\geq \theta \min_{\|\boldsymbol{v}\|_2=1} \boldsymbol{v}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{v} + (1-\theta)\min_{\|\boldsymbol{v}\|_2=1} \boldsymbol{v}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{v} \\ &= \theta f_M(\boldsymbol{X}) + (1-\theta) f_M(\boldsymbol{Y}), \end{aligned}$$

since it is always true that the pointwise minimum of the sum of two functions is greater than the sum of the pointwise minima.

(b) Use your result from the previous part to show that $\mathbb{S}_+^N$ is a convex set.

Solution: We know that $\boldsymbol{X} \in \mathbb{S}_+^N$ if and only if $\boldsymbol{X}$ is symmetric and it's largest eigenvalue is nonnegative, i.e. $-f_M(\boldsymbol{X}) \leq 0$. The set of symmetric matrices is a linear subspace, and hence is convex, and since $-f_M(\boldsymbol{X})$ is a convex function, $\{\boldsymbol{X} \ : \ -f_M(\boldsymbol{X}) \leq 0\}$ is a convex set. Thus $\mathbb{S}_+^N$ is an intersection of two convex sets, and so is itself convex.

(c) Find a set of convex functions $f_1(\boldsymbol{X}), \ldots, f_M(\boldsymbol{X})$ that map *arbitrary* $N \times N$ matrices to scalars ($f_m(\boldsymbol{X}) : \mathbb{R}^{N\times N} \to \mathbb{R}$) and scalars $b_1, \ldots, b_M$ that specify $\mathbb{S}_+^N$, meaning

$$\boldsymbol{X} \in \mathbb{S}_+^N \quad \Leftrightarrow \quad f_m(\boldsymbol{X}) \leq b_m, \text{ for all } m = 1, \ldots, M.$$

(Note that if $f_m$ is linear, then $f_m$ is both convex and concave, and so $f_m(\boldsymbol{X}) = b_m$ can be implemented using the pair of inequalities $f_m(\boldsymbol{X}) \leq b_m$ and $-f_m(\boldsymbol{X}) \leq -b_m$.) Note that there are multiple valid answers to this problem.

Solution: Take $M = N(N-1)+1$. We know that $\boldsymbol{X}$ is symmetric if and only if $X_{i,j} = X_{j,i}$ for all $i, j$. This introduces $\binom{N}{2} = \frac{N(N-1)}{2}$ pairs of affine (so convex) constraints on $\boldsymbol{X}$ of the form

$$\begin{aligned} X_{i,j} - X_{j,i} &\leq 0 \\ -X_{i,j} + X_{j,i} &\leq 0, \quad i = 1, \ldots, N, \ i < j \leq N. \end{aligned}$$

We have already seen that $-f(\boldsymbol{X})$ from part (a) is convex, and so $-f(\boldsymbol{X}) \leq 0$ serves as the final convex constraint.

**Note:** An alternative, and perhaps more elegant, solution is to replace the first $N(N-1)$ functions above with the function $f_1(\boldsymbol{X}) = \|\boldsymbol{X} - \boldsymbol{X}^T\|$ where $\|\cdot\|$ denotes your favorite matrix norm. Trivially, $\boldsymbol{X}$ is symmetric $\Leftrightarrow \boldsymbol{X} - \boldsymbol{X}^T = \boldsymbol{0}$ $\Leftrightarrow f_1(\boldsymbol{X}) = \|\boldsymbol{X} - \boldsymbol{X}^T\| = 0 \Leftrightarrow \|\boldsymbol{X} - \boldsymbol{X}^T\| \leq 0$. So $f_1(\boldsymbol{X}) \leq 0$ takes care of all $N(N-1)$ of the affine constraints above, and as we have just proven above, any norm is a convex function.

8. Compute the first and second derivatives of the following functions (remember the product and chain rules).

(a) $f(x) = ax^2 + bx + c$, where $a, b, c$ are constants.

Solution: $f'(x) = 2ax + b$ and $f''(x) = 2a$.

(b) $f(x) = \sum_{m=1}^{M} \log(1 + e^{-a_m x})$, where $a_1, \ldots, a_M$ are constants.
Solution: If we set $g(x) = \log(x)$ and $h(x) = 1 + e^{-a_m x}$, together with the facts that

$$g'(x) = \frac{1}{x} \qquad h'(x) = -a_m e^{-a_m x},$$

we have that

$$f'(x) = \sum_{m=1}^{M} \frac{1}{1 + e^{-a_m x}} \cdot \left(-a_m e^{-a_m x}\right) = -\sum_{m=1}^{M} \frac{a_m}{e^{a_m x} + 1}$$

Using the chain rule again we have

$$f''(x) = \sum_{m=1}^{M} \frac{a_m^2 e^{a_m x}}{(e^{a_m x} + 1)^2}.$$

9. Compute the gradient and Hessian matrix of the following functions. Note that $\boldsymbol{x}$ is a vector in $\mathbb{R}^N$ in all the problems below.

(a) $f(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} + \boldsymbol{b}^T \boldsymbol{x} + c$, where $\boldsymbol{A}$ is an $N \times N$ symmetric matrix (i.e., $\boldsymbol{A} = \boldsymbol{A}^T$),$\boldsymbol{b}$ is an $N \times 1$ vector, and $c$ is a scalar.
Solution: The $j^{\text{th}}$ element of the gradient is given by

$$\begin{aligned}
\frac{\partial}{\partial x_j} f(\boldsymbol{x}) &= \frac{\partial}{\partial x_j} \left( \sum_{n=1}^{N} \sum_{m=1}^{N} A_{m,n} x_m x_n + \sum_{n=1}^{N} b_n x_n + c \right) \\
&= \frac{\partial}{\partial x_j} \left( \sum_{m=1}^{N} A_{m,j} x_m x_j \right) + \frac{\partial}{\partial x_j} \left( \sum_{n=1}^{N} A_{j,n} x_n x_j \right) + \frac{\partial}{\partial x_j} \left( \sum_{n=1}^{N} b_n x_n \right) \\
&= \sum_{m=1}^{N} A_{m,n} x_m + \sum_{n=1}^{N} A_{m,n} x_n + b_j
\end{aligned}$$

The sums above represent an inner product between the $j^{\text{th}}$ row and $j^{\text{th}}$ column of $\boldsymbol{A}$. As a vector, this can be represented more compactly as

$$\nabla f(\boldsymbol{x}) = (\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}) \boldsymbol{x} + \boldsymbol{b}.$$

Since $\boldsymbol{A}$ is symmetric, $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$ and we can write this as

$$\nabla f(\boldsymbol{x}) = 2 \boldsymbol{A} \boldsymbol{x} + \boldsymbol{b}.$$

To compute the Hessian, we can compute the mixed partial derivative of the equation above to obtain

$$\begin{aligned}
\frac{\partial^2}{\partial x_i \partial x_j} &= \frac{\partial}{\partial x_i} \left( \sum_{m=1}^{N} A_{m,j} x_m + \sum_{n=1}^{N} A_{j,n} x_n + b_j \right) \\
&= A_{i,j} + A_{j,i}.
\end{aligned}$$

The Hessian matrix can be written more compactly as

$$\nabla^2 f(\boldsymbol{x}) = \boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}} = 2 \boldsymbol{A}.$$

(b) $f(\boldsymbol{x}) = \sum_{m=1}^{M} \log(1 + e^{-\boldsymbol{a}_m^T \boldsymbol{x}})$, where $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M$ are $N \times 1$ vectors.
Solution: Considering a single term in the sum, the $j^{\text{th}}$ element in the gradient is given by

$$\frac{\partial}{\partial x_j}\left(\log(1 + e^{-\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}})\right).$$

Recalling that $\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x} = \sum_{n=1}^{N} \boldsymbol{a}_{m,n} \boldsymbol{x}_n$, where $\boldsymbol{a}_{m,n}$ denotes the $n^{\text{th}}$ entry in $\boldsymbol{a}_m$, using the chain rule (twice) we have that

$$\frac{\partial}{\partial x_j}\left(\log(1 + e^{-\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}})\right) = \frac{-\boldsymbol{a}_{m,j} e^{-\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}}}{1 + e^{-\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}}} = -\frac{\boldsymbol{a}_{m,j}}{e^{\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}} + 1}.$$

Summing over $m$, in matrix notation we have

$$\nabla f(\boldsymbol{x}) = -\sum_{m=1}^{M} \frac{\boldsymbol{a}_m}{e^{\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}} + 1}.$$

You could also compute this using the multivariate chain rule, but you can do this as an exercise. Computing the second partial derivatives we have

$$\frac{\partial^2}{\partial x_i \partial x_j} = \sum_{m=1}^{M} \frac{\boldsymbol{a}_{m,j} \boldsymbol{a}_{m,i} e^{\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}}}{(e^{\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}} + 1)^2},$$

and so in matrix notation the Hessian is given by

$$\nabla^2 f(\boldsymbol{x}) = \sum_{m=1}^{M} \frac{e^{\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}}}{(e^{\boldsymbol{a}_m^{\mathrm{T}} \boldsymbol{x}} + 1)^2} \boldsymbol{a}_m \boldsymbol{a}_m^{\mathrm{T}}.$$

10. Determine whether the following functions are convex, concave, or neither.

(a) $f(x) = e^{x^2}$ on $\operatorname{dom} f = \mathbb{R}$.
Solution: We have

$$f'(x) = 2xe^{x^2}$$
$$f''(x) = 4x^2 e^{x^2} + 2e^{x^2} = 2e^{x^2}(2x^2 + 1) > 0.$$

Hence, $f(x)$ is convex.

(b) $f(x) = \log(1 + e^x)$ on $\operatorname{dom} f = \mathbb{R}$.
Solution: We have

$$f'(x) = \frac{e^x}{1 + e^x}$$
$$f''(x) = \frac{(1 + e^x)e^x - (e^x)e^x}{(1 + e^x)^2} = \frac{e^x}{(1 + e^x)^2} > 0.$$

Hence, $f(x)$ is convex.

(c) $f(x_1, x_2) = x_1 x_2$ on $\mathbf{dom}\, f = \mathbb{R}^2_{++}$.
Solution: We have

$$\nabla f(x_1, x_2) = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$
$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The eigenvalues of the Hessian are 1 and -1. Hence, $f(x_1, x_2)$ is neither convex nor concave.

(d) $f(x_1, x_2) = 1/x_1 x_2$ on $\mathbf{dom}\, f = \mathbb{R}^2_{++}$.
Solution: We have

$$\nabla f(x_1, x_2) = \begin{bmatrix} -\frac{1}{x_1^2 x_2} \\ \\ -\frac{1}{x_1 x_2^2} \end{bmatrix}$$
$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_1^3 x_2} & \frac{1}{x_1^2 x_2^2} \\ \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{bmatrix}$$

The Sylvester's criterion (`https://en.wikipedia.org/wiki/Sylvester's_criterion`) states that a matrix $M$ is positive semidefinite if and only if all the principal minors of M are nonnegative.

Here, for the Hessian we have

$$D_1 = \frac{2}{x_1^3 x_2} \geq 0$$
$$D_2 = \frac{2}{x_1 x_2^3} \geq 0$$
$$D_3 = \frac{2}{x_1^3 x_2}\left(\frac{2}{x_1 x_2^3}\right) - \frac{1}{x_1^2 x_2^2}\left(\frac{1}{x_1^2 x_2^2}\right) = \frac{3}{x_1^4 x_2^4} \geq 0$$

The Hessian is positive semidefinite, hence $f(x_1, x_2)$ is convex.

(e) $f(x_1, x_2) = x_1/x_2$ on $\mathbf{dom}\, f = \mathbb{R}^2_{++}$.
Solution: We have

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{1}{x_2} \\ \\ -\frac{x_1}{x_2^2} \end{bmatrix}$$
$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & -\frac{1}{x_2^2} \\ \\ -\frac{1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix}$$

Consider $x_1 = 2$ and $x_2 = 1$. We have

$$\nabla^2 f(2, 1) = \begin{bmatrix} 0 & -1 \\ -1 & 4 \end{bmatrix}$$

i. Let $\boldsymbol{a} = \begin{bmatrix} 3 & 4 \end{bmatrix}^T$. We have

$$\boldsymbol{a}^T \ \nabla^2 f(2,1) \ \boldsymbol{a} = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 40 > 0$$

ii. Let $\boldsymbol{a} = \begin{bmatrix} 5 & 1 \end{bmatrix}^T$. We have

$$\boldsymbol{a}^T \ \nabla^2 f(2,1) \ \boldsymbol{a} = \begin{bmatrix} 5 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \end{bmatrix} = -6 < 0$$

The Hessian is neither positive semidefinite nor negative semidefinite. Hence $f(x_1, x_2)$ is neither convex nor concave.

(f) $f(x_1, x_2) = x_1^2/x_2$ on $\mathrm{dom}\, f = \mathbb{R}^2_{++}$.
Solution: We have

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{2x_1}{x_2} \\ \\ -\frac{x_1^2}{x_2^2} \end{bmatrix}$$

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2^2} \\ \\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix}$$

Using the Sylvester's criterion stated above, the principal minors of the Hessian are

$$D_1 = \frac{2}{x_2} \geq 0$$

$$D_2 = \frac{2x_1^2}{x_2^3} \geq 0$$

$$D_3 = \frac{2}{x_2}\left(\frac{2x_1^2}{x_2^3}\right) - \left(-\frac{2x_1}{x_2^2}\right)\left(-\frac{2x_1}{x_2^2}\right) = 0$$

Using the Sylvester's criterion stated above, we have that the Hessian is positive semidefinite. Hence $f(x_1, x_2)$ is convex.