(1) In class we defined positive semidefinite (PSD) matrices. A $d \times d$ matrix $A$ is said to be positive definite (PD) if for all $d \times 1$ vectors $x$ that are not the all-zeros vector, we have that $x^\top Ax > 0$. Observe that a positive definite matrix is also positive semi-definite, but a positive semi-definite matrix may not be positive definite.

For each of the following matrices, state if they are (a) positive definite (b) positive semi-definite but not positive definite (c) not positive semi-definite. Justify your answer with a short proof.

(a) (3 points) Let $u$ be a $d \times 1$ unit vector – that is, $\|u\| = 1$. $A = uu^\top$.

For any nonzero vector $x$, we have
$$x^\top Ax = x^\top uu^\top x = (u^\top x)^2 u^\top x = (u, x)^2 \geq 0.$$ 

Now, to show that $A$ is not positive definite, we can pick a unit vector $v$ that is perpendicular to $u$. Then, we have
$$v^\top Av = (u, v)^2 = 0.$$ 
Thus, $A$ is positive semi-definite, but not positive definite.

(b) (3 points) $B = 2I_d$. (Recall $I_d$ is the $d \times d$ identity matrix.)

For any nonzero vector $x$, we have
$$x^\top Bx = x^\top 2I_d x = 2x^\top x = 2\|x\|^2 > 0.$$ 
Note that the equality does not hold as $x$ is nonzero. Thus, $B$ is positive definite.

(c) (4 points) Let $u$ be a $d \times 1$ unit vector – that is, $\|u\| = 1$. $C = uu^\top - \frac{1}{2}I_d$.

Let us pick a unit vector $v$ that is perpendicular to $u$. Then, we have
$$v^\top Cv = v^\top \left( uu^\top - \frac{1}{2}I_d \right) v = v^\top uu^\top v - \frac{1}{2}v^\top v = 0 - \frac{1}{2}\|v\|^2 = -\frac{1}{2}.$$ 
This violates the condition of positive semi-definite. Thus, $C$ is not positive semi-definite.
(2) (5 points) Let \( x \) be a \( d \times 1 \) vector, let \( z_1, \ldots, z_n \) be \( n \) vectors, where each \( z_i \) is \( d \times 1 \), and let \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) be the following multivariate function of \( x \):

\[
f(x) = \sum_{i=1}^{n} e^{x^\top z_i}
\]

Write down the gradient of \( f \) with respect to \( x \). Write down an algorithm that computes this gradient in \( O(nd) \) time when given the vectors \( z_1, \ldots, z_n \).

First, we note that

\[
f(x) = \sum_{i=1}^{n} e^{x^\top z_i} = \sum_{i=1}^{n} e^{z_i^1 x_1 + z_i^2 x_2 + \cdots + z_i^d x_d} = \sum_{i=1}^{n} e^{z_i^1 x_1} e^{z_i^2 x_2} \cdots e^{z_i^d x_d}.
\]

Then, we can compute the partial derivative with respect to \( x_j \) as follows.

\[
\frac{\partial f}{\partial x_j} = \sum_{i=1}^{n} e^{z_i^1 x_1} e^{z_i^2 x_2} \cdots \left( z_{ij} e^{z_{ij} x_j} \right) \cdots e^{z_i^d x_d} \quad \text{(by product rule)}
\]

\[
= \sum_{i=1}^{n} e^{z_i^1 x_1 + z_i^2 x_2 + \cdots + z_{ij} x_j + \cdots + z_i^d x_d}
\]

\[
= \sum_{i=1}^{n} e^{x^\top z_i} z_{ij}.
\]

Thus, the gradient of \( f \) is

\[
\frac{\partial f}{\partial x} = \begin{bmatrix}
\sum_{i=1}^{n} e^{x^\top z_i} z_{i1} \\
\sum_{i=1}^{n} e^{x^\top z_i} z_{i2} \\
\vdots \\
\sum_{i=1}^{n} e^{x^\top z_i} z_{id}
\end{bmatrix} = \sum_{i=1}^{n} e^{x^\top z_i} z_i.
\]

This can be computed in \( O(nd) \) time as follows.

\begin{algorithm}
\textbf{Algorithm 1: Compute the gradient} \\
\textbf{Input:} A \( d \times 1 \) vector \( x \) and \( n \) \( d \times 1 \) vectors \( z_1, z_2, \ldots, z_n \) \\
\textbf{Output:} A \( d \times 1 \) vector representing the gradient \( \frac{\partial f}{\partial x} \)

\begin{enumerate}
\item Let \( g \) be an array of size \( d \), initialized to zeros
\item for \( i = 1 \) to \( n \) do
\item \hspace{1em} \( y \leftarrow x^\top z_i \) \hspace{1em} // This can be done in \( O(d) \) time
\item \hspace{1em} for \( j = 1 \) to \( d \) do
\item \hspace{2em} \( g[j] \leftarrow g[j] + e^y z_{ij} \) \hspace{1em} // This can be done in \( O(1) \) time
\item \hspace{1em} end
\item \hspace{1em} end
\item return \( g \)
\end{enumerate}
\end{algorithm}
(3) (5 points) Given a $d \times d$ matrix $A$, its Frobenius norm is defined as:

$$\|A\|_F = \sqrt{\sum_{i=1}^{d} \sum_{j=1}^{d} A_{ij}^2}$$

Now suppose we are given two $d \times 1$ vectors $x$ and $y$. State whether the following statement is true or false. Justify your answer.

$$x^\top y \leq \|xy^\top\|_F$$

Let $A = xy^\top$. We have

$$\left( x^\top y \right)^2 = \left( \sum_{i=1}^{d} x_i y_i \right)^2$$

$$\leq \left( \sum_{i=1}^{d} x_i^2 \right) \left( \sum_{i=1}^{d} y_i^2 \right)$$  \hspace{1cm} \text{(by Cauchy-Schwarz inequality)}

$$= \sum_{i=1}^{d} \sum_{j=1}^{d} x_i^2 y_j^2$$

$$= \sum_{i=1}^{d} \sum_{j=1}^{d} A_{ij}^2$$  \hspace{1cm} \text{(by definition, } A_{ij} = x_i y_j)$$

$$= \|xy^\top\|_F^2$$

Since $\|xy^\top\|_F$ is always nonnegative, we can conclude that

$$x^\top y \leq \|xy^\top\|_F$$.

This problem is similar to Question 4 in Homework1, where we showed that $\|xy^\top\|_F = \|x\|\|y\|$.