# CS 545 Homework 1

Fall 2023

**Name:** Hongbo Zheng

**NetID:** *hongboz2*

**Date:** Wednesday, September 13$^{\text{th}}$ 11:59 P.M.

## Problem 1. Some simple linear algebra

1. $\text{vec}(a \cdot b^\intercal) = b \otimes a$, where $a$ and $b$ are both vectors

   Suppose $a \in \mathbb{R}^m$, $\dim(a) = m \times 1$, and $b \in \mathbb{R}^n$, $\dim(b) = n \times 1$ are two column vectors

$$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_m \end{bmatrix} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} \qquad b^\intercal = \begin{bmatrix} b_1, b_2, b_3, \dots, b_n \end{bmatrix}$$

$$a \cdot b^\intercal = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_m \end{bmatrix} \begin{bmatrix} b_1, b_2, b_3, \dots, b_n \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & a_2b_3 & \dots & a_2b_n \\ a_3b_1 & a_3b_2 & a_3b_3 & \dots & a_3b_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_mb_1 & a_mb_2 & a_mb_3 & \dots & a_mb_n \end{bmatrix} \in \mathbb{R}^m \tag{1}$$

$$\text{vec}(ab^\intercal) = \begin{bmatrix} a_1b_1 \\ a_2b_1 \\ a_3b_1 \\ \vdots \\ a_mb_1 \\ a_1b_2 \\ a_2b_2 \\ a_3b_2 \\ \vdots \\ a_mb_2 \\ \vdots \\ a_mb_n \end{bmatrix} = \begin{bmatrix} b_1a \\ b_2a \\ b_3a \\ \vdots \\ b_na \end{bmatrix} \tag{2}$$

$$b \otimes a = \begin{bmatrix} b_1a \\ b_2a \\ b_3a \\ \vdots \\ b_na \end{bmatrix} \tag{3}$$

$$\text{vec}(ab^\intercal) = \begin{bmatrix} b_1a \\ b_2a \\ b_3a \\ \vdots \\ b_na \end{bmatrix} = b \otimes a \tag{4}$$

2. $\text{vec}(A)^\intercal \cdot \text{vec}(B) = \text{tr}(A^\intercal \cdot B)$

Suppose matrix $A$ and matrix $B$ both are a $m \times n$ square matrix $\in \mathbb{R}^m$

$$\text{vec}(A) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \\ a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \\ \vdots \\ a_{mn} \end{bmatrix} \in \mathbb{R}^{mn},\ \dim(\text{vec}(A)) = mn \times 1 \qquad \text{vec}(B) = \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \\ b_{12} \\ b_{22} \\ \vdots \\ b_{m2} \\ \vdots \\ b_{mn} \end{bmatrix} \in \mathbb{R}^{mn},\ \dim(\text{vec}(B)) = mn \times 1$$

$\text{vec}(A)^\intercal = \begin{bmatrix} a_{11}, a_{21}, \ldots, a_{m1}, a_{12}, a_{22}, \ldots, a_{m2}, \ldots, a_{mn} \end{bmatrix} \in \mathbb{R}$, $\dim(\text{vec}(A)^\intercal) = 1 \times mn$

$$\text{vec}(A)^\intercal \, \text{vec}(B) = \begin{bmatrix} a_{11}, a_{21}, \ldots, a_{m1}, a_{12}, a_{22}, \ldots, a_{m2}, \ldots, a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \\ b_{12} \\ b_{22} \\ \vdots \\ b_{m2} \\ \vdots \\ b_{mn} \end{bmatrix} = \sum_j^n \sum_i^m a_{ij} b_{ij} \tag{5}$$

$$\begin{aligned} \text{tr}(A^\intercal B) &= \text{tr}\left( \begin{bmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{bmatrix} \begin{bmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \ldots & b_{mn} \end{bmatrix} \right) \\ &= \quad a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + \ldots + a_{1m}b_{m1} \\ &\quad + a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} + \ldots + a_{2m}b_{m2} \\ &\quad + a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} + \ldots + a_{3m}b_{m3} \\ &\qquad \qquad \qquad \ldots \\ &\quad + a_{n1}b_{1n} + a_{n2}b_{2n} + a_{n3}b_{3n} + \ldots + a_{nm}b_{mn} \\ &= \sum_{i=1}^m a_{1i}b_{i1} + \sum_{i=1}^m a_{2i}b_{i2} + \sum_{i=1}^m a_{3i}b_{i3} + \ldots + \sum_{i=1}^m a_{ni}b_{in} \\ &= \sum_{j=1}^n \sum_{i=1}^m a_{ji} b_{ij} \end{aligned}$$

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix},\ A^\intercal = \begin{bmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{bmatrix},\ a_{ji} \text{ in } A^\intercal \text{ is } a_{ij} \text{ in } A$$

$$\text{tr}(A^\intercal B) = \sum_{j=1}^n \sum_{i=1}^m a_{ij} b_{ij} = \text{vec}(A)^\intercal \, \text{vec}(B) \tag{6}$$

3. If $X$ is a real matrix, then the matrix $XX^\intercal$ is positive semi-definite

   Assume $X \in \mathbb{R}^n$, let $y$ be a column vector $\in \mathbb{R}^n$

$$y^\intercal XX^\intercal y = (y^\intercal X)(X^\intercal y) = (y^\intercal X)(X^\intercal y)^\intercal = (y^\intercal X)(y^\intercal X) = \|y^\intercal X\| \geq 0 \tag{7}$$

   Therefore, $XX^\intercal$ is positive semi-definite

## Problem 2. A probability problem

Let $C$ = has breast cancer, $M$ = positive mammogram,

$P(C) = 0.6\%$, $P(M|C) = 90\%$, $P(M|\neg C) = 7\%$

From above, $P(\neg C) = 1 - P(C) = 1 - 0.6\% = 99.4\%$

According to Bayes' Theorem

$$P(C|M) = \frac{P(M|C)P(C)}{P(M)} = \frac{P(M|C)P(C)}{P(M|C)P(C) + P(M|\neg C)P(\neg C)} = \frac{90\% \times 0.6\%}{90\% \times 0.6\% + 7\% \times 99.4\%} \tag{8}$$

$$P(C|M) \approx 7.2019\% \tag{9}$$

## Problem 3. Stats operations using linear algebra

1. (a) Vectorize each image into a column vector $x = \text{vec}(\text{image}) \in \mathbb{R}^{MN}$, $\dim(x) = MN \times 1$

   Let $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_k \end{bmatrix} \in \mathbb{R}^{MN}$ be the matrix that contains the $K$ vectorized gray-scale images of size $M \times N$, $\dim(X) = MN \times K$

   Let $\mathbf{1}_K \in \mathbb{R}^K$ be a column vector that contains $K$ ones, $\dim(\mathbf{1}_K) = K \times 1$

   Average of all $K$ grey-scale images of size $M \times N$ is column vector $\overline{x} = \frac{1}{K}\mathbf{x}\mathbf{1}_K \in \mathbb{R}^{MN}$. $\dim(\overline{x}) = MN \times 1$

   Convert the vector $\overline{x}$ into a $M \times N$ matrix $\overline{X} = \overline{x}^{(M)} \in \mathbb{R}^M$. $\dim(\overline{x}^{(M)}) = M \times N$

$$\overline{X} = \left(\frac{1}{K}\mathbf{x}\mathbf{1}_K\right)^{(M)} \in \mathbb{R}^M \tag{10}$$

   (b) Let $X_i \in \mathbb{R}^M$ be a matrix that represents one gray-scale image with size $M \times N$

   The vectorized top half of the image is obtain with the following equation:

$$\text{vec}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \mathbf{I}_{\frac{1}{2}M} X_i\right) \tag{11}$$

   According to the rule, $A$ is a $K \times L$ matrix, $B$ is a $L \times M$ matrix, $\text{vec}(AB) = (\mathbf{I}_M \otimes A)\,\text{vec}(B)$, let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \mathbf{I}_{\frac{1}{2}M} \in \mathbb{R}^M$, $B = X_i \in \mathbb{R}^M$, $\text{vec}(B) = \text{vec}(X_i) = x_i \in \mathbb{R}^{MN}$

$$\text{vec}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \mathbf{I}_{\frac{1}{2}M} X_i\right) = \left(\mathbf{I}_N \otimes \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \mathbf{I}_{\frac{1}{2}M}\right) x_i \in \mathbb{R}^{MN} \tag{12}$$

   Let $\mathbf{1}_K \in \mathbb{R}^K$ be a column vector that contains $K$ ones, $\dim(\mathbf{1}_K) = K \times 1$, and let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_k \end{bmatrix}$ which contains all the vectorized grey-scale images. The average of the vectorized top half of the image is:

$$t = \frac{1}{K}\left(\mathbf{I}_N \otimes \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \mathbf{I}_{\frac{1}{2}M}\right)\mathbf{x}\mathbf{1}_K \in \mathbb{R}^{MN} \tag{13}$$

   Similarly, the average of the vectorized bottom half of the image is:

$$b = \frac{1}{K}\left(\mathbf{I}_N \otimes \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \otimes \mathbf{I}_{\frac{1}{2}M}\right)\mathbf{x}\mathbf{1}_K \in \mathbb{R}^{MN} \tag{14}$$

   Let $\mathbf{1}_{MN}$ be a column vector with $MN$ ones, the average value of the vector $\bar{t}$ is

$$\bar{t} = \frac{1}{\frac{1}{2}MN}\mathbf{1}_{MN}^{\intercal} t \in \mathbb{R} \tag{15}$$

   Similarly, the average value of the vector $\bar{b}$ is

$$\bar{b} = \frac{1}{\frac{1}{2}MN}\mathbf{1}_{MN}^{\intercal} b \in \mathbb{R} \tag{16}$$

   The $2 \times 2$ covariance matrix of the average of the top half of all the images with the average of the bottom half is:

$$C = \begin{bmatrix} \text{var}(t) & \text{cov}(t, b) \\ \text{cov}(t, b) & \text{var}(b) \end{bmatrix} = \begin{bmatrix} (t - \bar{t})^{\intercal}(t - \bar{t}) & (t - \bar{t})^{\intercal}(b - \bar{b}) \\ (t - \bar{t})^{\intercal}(b - \bar{b}) & (b - \bar{b})^{\intercal}(b - \bar{b}) \end{bmatrix} \tag{17}$$

2. Let $X \in \mathbb{R}^K$ be the matrix that contains all the images. $\dim(X) = M \times N \times 3 \times K$

   (a) Let $\mathbf{1}_K$ be a column vector with $K$ ones in it, similarly, $\mathbf{1}_3$ is a column vector with 3 ones in it.

   The mean image $\overline{X}$ $(M \times N)$ matrix over all channels is:

$$\overline{X} = \left(\left(\left(\left(\frac{1}{K}\mathbf{1}_K^\intercal \otimes \frac{1}{3}\mathbf{1}_3^\intercal\right) \otimes \mathbf{I}_N\right) \otimes \mathbf{I}_M\right) \operatorname{vec}(\mathbf{X})\right)^{(M)} \tag{18}$$

   $\frac{1}{3}\mathbf{1}_3^\intercal$ is used to calculate the average across all 3 channels, and $\frac{1}{K}\mathbf{1}_K^\intercal$ is to used to calculate the average across all $K$ images. The following kronekcer product with $\mathbf{I_N}$ and $\mathbf{I}_M$ are used to scale the matrix to fit the size of the vectorized input $\operatorname{vec}(\mathbf{X})$. Finally, we perform vec-transpose like in prblem 3.1 (a).

   (b) Let $\mathbf{1}_K$ be a column vector with $K$ ones in it.

   The mean image $\overline{X}$ $(M \times N)$ matrix of only the red channel is:

$$\overline{X}_R = \left(\left(\left(\left(\frac{1}{K}\mathbf{1}_K^\intercal \otimes \begin{bmatrix}1 & 0 & 0\end{bmatrix}\right) \otimes \mathbf{I}_N\right) \otimes \mathbf{I}_M\right) \operatorname{vec}(\mathbf{X})\right)^{(M)} \tag{19}$$

   $\begin{bmatrix}1 & 0 & 0\end{bmatrix}$ is used to only take the red channel from the 3 channels, and zero out the green and blue channels. $\frac{1}{K}\mathbf{1}_K^\intercal$ is to used to calculate the average across all $K$ images. The following kronekcer product with $\mathbf{I_N}$ and $\mathbf{I}_M$ are used to scale the matrix to fit the size of the vectorized input $\operatorname{vec}(\mathbf{X})$. Finally, we perform vec-transpose like in prblem 3 (a).

## Problem 4. Signal Processing in Linear Algebra

To construct a matrix $A$, such that the product $Ax$ will produce the vec$(\cdot)$ of the complex spectrogram coefficients, we need construct a Fourier matrix $F$ with a DFT size of 64 and a Hann window vector $h$ with a size of 64. Then, we need to transform the Hann window vector $h$ into a diagonal matrix $H$ with size of $64 \times 64$, and finally multiply these two matrix together to get matrix $A$.

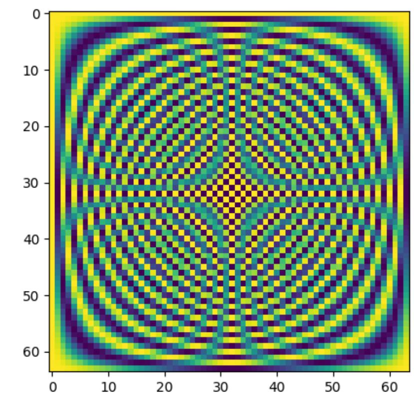

Figure 1: The fourier matrix F with DFT size of 64.

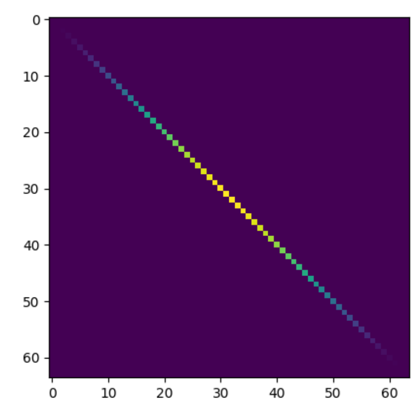

Figure 2: The Hann window diagonal matrix with a size of $64 \times 64$.

The transformation matrix $A = F \times H$

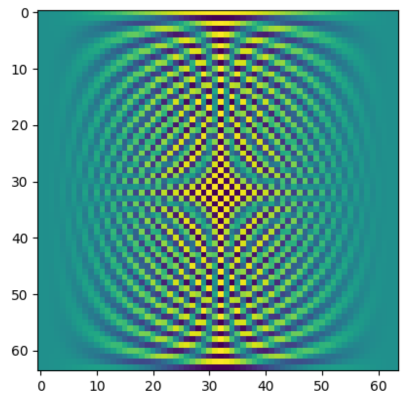


Figure 3: The real part of the matrix $A$.

For a input sound vector $x \in \mathbb{R}^n$, we will have to stack the sub-matrix $\frac{(n-64)}{32} + 1$ times. For example, if $n = 192$, we will have to stack the sub-matrix $\frac{192-64}{32} + 1 = 5$ times, the real part of the matrix $A$ will be the following, with a hop size of 32.

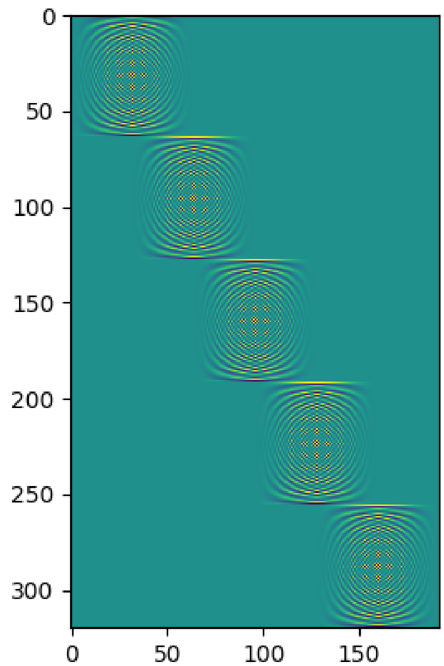


Figure 4: The real part of the matrix $A$ with a input vector $x$ size of 192.