1) Let $\mathcal{H} := \{ h : [n] \rightarrow [k] \}$ be a family of pairwise independent hash functions and let $\mathcal{S} = \{ \sigma : [n] \rightarrow \{-1, +1\} \}$ be a family of 4-wise independent hash functions. Sample $h \sim \mathcal{H}$ and $\sigma \sim \mathcal{S}$, the countsketch matrix $S_{h,\sigma} \in \mathbb{R}^{k \times n}$ is defined as follows: The matrix has only one non-zero entry in every column $j$ at $S_{h(j),j}$ and that is equal to $\sigma(j)$. Note that the matrix has only $n$ nonzero entries altogether. Because of this for any matrix $A \in \mathbb{R}^{n \times d}$, $S_{h,\sigma}A$ can be computed in time $O(nnz(A))$ where as usual $nnz(A)$ is the number of non-zero entries of $A$. The countsketch matrix has numerous applications in approximate matrix multiplication, regression, etc. These applications mostly use the JL-property of countsketch matrix that we prove in this exercise.

a) Show that for any vector $x \in \mathbb{R}^n$, $\mathbb{E} [\| S_{h,\sigma} x \|^2] = \| x \|^2$. Here the randomness is over the randomness of $h, \sigma$.

b) Show that $\mathbb{E} [\| Sx \|^2] \leq (1 + 2/k) \| x \|^2$.

c) Use these to conclude that for any unit norm $x$, $\mathbb{E} [\| Sx \|^2 - 1^2] \leq 2/k$. In other words, the matrix $S$ can be seen as a dimension reduction matrix to $\mathbb{R}^k$ that preserve the norm of any unit norm vector $x$ with a constant probability.

2) A Hadamard matrix $H \in \{-1, +1\}^{n \times n}$ is a matrix where the inner product of every (not equal) two rows of $H$ are zero, i.e., for any $i \neq j$ we have $\sum_k H_{i,k} H_{j,k} = 0$. For example,

$$H_2 = \begin{pmatrix} +1 & -1 \\ +1 & +1 \end{pmatrix}.$$  

In general, if $A$ is a $k \times k$ Hadamard matrix you can construct $2k \times 2k$ Hadamard matrix by putting

$$\begin{pmatrix} A & -A \\ A & A \end{pmatrix}.$$  

a) Let $H$ be an $n \times n$ Hadamard matrix. Prove that all singular values of $H$ are equal to $\sqrt{n}$.

b) Let $A \in [-1, +1]^{n \times n}$ matrix, i.e., every entry of $A$ is in the range $[-1, +1]$ such that at most $n^2/8$ entries of $A$ are different from $H$. Use the following theorem to prove that $A$ has rank at least $\Omega(n)$.

**Theorem 3.1** (Hoffman-Wielandt Inequality). Let $A, B \in \mathbb{R}^{n \times n}$ with singular values $\sigma_1 \geq \cdots \geq \sigma_n$ and $\sigma'_1 \geq \cdots \geq \sigma'_n$. Then,

$$\sum_{i=1}^n |\sigma_i - \sigma'_i|^2 \leq \| A - B \|_F^2.$$  

3) For a vector $u \in \mathbb{R}^n$, we write $u \otimes u$ to denote the vector in $\mathbb{R}^{n^2}$ where for any $1 \leq i, j \leq n$, $(u \otimes u)_{i,n+j} = u_i \cdot u_j$.

a) Show that for any pair of vectors $u, v \in \mathbb{R}^n$,

$$\langle u \otimes u, v \otimes v \rangle = \langle u, v \rangle^2.$$  

b) Let $A \in \mathbb{R}^{n \times n}$ be a PSD matrix, and let $B \in \mathbb{R}^{n \times n}$ be the matrix where $B_{i,j} = A_{i,j}^2$. Prove that $B$ is PSD.

**Hint:** Use part (a) and that any matrix $A$ is PSD iff it can be written as $A = CC^T$ for some matrix $C \in \mathbb{R}^{n \times k}$, for some integer $k$. 

3-1
4) In this problem we discuss a fast algorithm for approximately estimating the low rank approximation (up to an additive error) with respect to the Frobenius norm.

a) Let \( A \in \mathbb{R}^{m \times n} \) and suppose we want to estimate \( Av \) for a vector \( v \in \mathbb{R}^n \). Here is a randomized algorithm for this task. Choose the \( i \)-th column of \( A \), \( A_i \), with probability

\[
p_i = \frac{\|A_i\|^2}{\|A\|^2_F}
\]

and let \( X = A_i v_i / p_i \). Show that \( E[X] = Av \). Calculate \( \text{Var}(X) = E[\|X\|^2] - \|EX\|^2 \). Note that, with this definition, \( E[X] \) is a vector whereas \( \text{Var}(X) \) is a number.

b) Next, we use a similar idea to approximate \( A \). For \( 1 \leq i \leq s \) let \( X_i = \frac{A_j}{\sqrt{p_j}} \) with probability \( p_j \) where \( 1 \leq j \leq n \). Let \( X \in \mathbb{R}^{m \times s} \) and let \( X_i \) be the \( i \)-th columns of \( X \). Note that \( XX^T = \sum_{i=1}^s X_i X_i^T \). Show that

\[
EXX^T = AA^T.
\]

Show that \( E\|XX^T - AA^T\|_F^2 \leq \frac{1}{s}\|A\|_F^4 \).

5) Run a low-rank approximation on the “jecond.jpg” file in the website. How many singular values do you need to use to get a relatively good approximation of the image?

6) **Extra Credit:** Let \( P = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d \) be a set of points of norm 1. For \( \sigma > 0 \), let \( G_\sigma \in \mathbb{R}^{n \times n} \) be the Gaussian kernel on \( P \), i.e.,

\[
G_\sigma(i, j) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\|p_i - p_j\|^2/2\sigma}.
\]

Prove that \( G_\sigma \succeq 0 \).