1. Least Squares Hyperplanes and Orthogonal Least Squares Planes

a) Create an artificial data set of 1000 points in $\mathbb{R}^4$ by typing the following Matlab commands

```matlab
x1(1:500,1) = rand(500,1);
x1(501:1000,1) = 2+rand(500,1);
x2(1:500,1) = rand(500,1);
x2(501:1000,1) = 2+rand(500,1);
x3(1:1000,1) = rand(1000,1)/5;
$\epsilon$ = 1;
x4 = 0.25*x1+1.3*x2-1.2*x3+23+$\epsilon$*randn(1000,1);
```

Use SVD decomposition implemented in Matlab (do not use the command `regress`) to find the equation of the least squares approximation of the form $x_4 = a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3 + a_0$ (that is, specify the coefficients $a_0$, $a_1$, $a_2$ and $a_3$) and the corresponding averaged $l_2$ error, that it minimizes:

$$
\left( \frac{1}{N} \sum_{i=1}^{N} (x_4)_i - (a_1 \cdot (x_1)_i + a_2 \cdot (x_2)_i + a_3 \cdot (x_3)_i + a_0))^2 \right)^{\frac{1}{2}}.
$$

Next, change $\epsilon = \frac{1}{10}$ and report your answer in that case.

b) Use the data set created above (apply both $\epsilon = 1$ and $\epsilon = \frac{1}{10}$ for all of the following questions) to find a 3-dimensional plane which minimizes the orthogonal $l_2$ error (we refer to it as best $l_2$ 3-plane). Express the plane in the form $c + Sp(v_1, v_2, v_3)$, where $\{v_1, v_2, v_3\}$ is an orthonormal set of vectors (do not confuse the notation above for the columns $x_1, x_2, x_3, x_4$ with the different notation used in class for the $N$ rows of the data matrix, where here $N = 1000$).

c) Form the data matrix $A$, whose rows are the data vectors (in $\mathbb{R}^4$) minus the vector $c$ found above. Then form a $4 \times 3$ matrix $P$, whose columns are $v_1, v_2, v_3$. Multiplying $A$ by $P$, we obtain a matrix with the coordinates of the shifted data points projected onto the best (orthogonal) $l_2$ 3-plane (which passes through $c$). Plot the vectors projected on this plane (i.e. the rows of $A \cdot P$; use the command `plot3`). Similarly plot the projection onto the best (orthogonal) $l_2$ 2-plane (use the command `plot`). Does the outcome make sense to you?

d) In http://www.ics.uci.edu/~mlearn/databases/iris/iris.data you may find the Iris data. It lists information on 150 Iris flowers. It includes three Iris species: Setosa, Versicolour, and Virginica (50 flowers per class). Each flower is characterized by five attributes: sepal length in centimeters,
sepal width in centimeters, petal length in centimeters, petal width in centimeters, class (Setosa, Versicolour, and Virginica). In the attached file to the homework (Iris_data.mat) the data was saved as Matlab file, separated to the 3 classes. In each class a flower is represented by a four dimensional data point according to the first four attributes. Project that data on the best (orthogonal) \( l_2 \)-2-plane and plot the projected points, where each class is distinguished (use e.g. 'o', 'x', '+'). Which class is easily separated from the others using that projection?

2. Image compression

Choose your favorite image of size at least 512 \( \times \) 512. Apply SVD compression with 3, 10, 20 and 40 top singular values and vectors. Plot your original image and the “compressed” ones. Also make a table with relative errors and compression ratios (i.e., ratios between sizes of the compressed SVD nonzero components to the sizes of the original matrices) for each of the images.

3. Properties of Convolution

You may solve only 5 out the following 6 subquestions.

a) Show that if \( v = \{v_i\}_{i \in \mathbb{Z}} \) and \( u = \{u_i\}_{i \in \mathbb{Z}} \) are two vectors in \( \ell_1(\mathbb{Z}) \), then their convolution \( u \ast v \) is also in \( \ell_1(\mathbb{Z}) \).

b) Show that if \( v = \{v_i\}_{i \in \mathbb{Z}} \) and \( u = \{u_i\}_{i \in \mathbb{Z}} \) are two probability vectors in \( \ell_1(\mathbb{Z}) \), that is their elements are positive and sum to 1, then their convolution \( u \ast v \) is also a probability vector. (If you can, give an interpretation of those probabilities).

c) Show that if \( v = \{v_i\}_{i \in \mathbb{Z}} \in \ell_1(\mathbb{Z}) \) and \( u = \{u_i\}_{i \in \mathbb{Z}} \) is of period \( N \), that is for all \( i \in \mathbb{Z} \): \( u_i = u_{i+N} \), then their convolution is well defined and is also of period \( N \).

d) Show that the convolution of signals in \( \ell_1(\mathbb{Z}) \) is commutative. That is, if \( v = \{v_i\}_{i \in \mathbb{Z}} \) and \( u = \{u_i\}_{i \in \mathbb{Z}} \) are in \( \ell_1(\mathbb{Z}) \), then \( u \ast v = v \ast u \).

e) Show that the convolution of signals in \( \ell_1(\mathbb{Z}) \) is associative. That is, if \( v, u, w \in \ell_1(\mathbb{Z}) \), then \( (u \ast v) \ast w = u \ast (v \ast w) \).

f) Let \( p_1(x) = \sum_{i=1}^{11} i \cdot x^i \) and \( p_2(x) = \sum_{i=1}^{9} i^2 \cdot x^i \). By only using the command \texttt{conv} in Matlab, find \( p_1(x) \cdot p_2(x) \) (print your Matlab output and explain how it is related to \( p_1(x) \cdot p_2(x) \)).

4. Properties of Correlation

If \( v = \{v_i\}_{i \in \mathbb{Z}} \) and \( u = \{u_i\}_{i \in \mathbb{Z}} \) are two vectors in \( \ell_1(\mathbb{Z}) \), we denote by \( \diamond \) their correlation and by \( \tilde{u} \) and \( \tilde{v} \) their reflection with respect to zero (\( \tilde{u}_i = u_{-i} \)).

a) Show that \( u \diamond v = \tilde{v} \diamond \tilde{u} \).

b) Show that \( u \diamond v = u \ast \tilde{v} = \tilde{v} \ast u \) and \( u \ast v = u \diamond \tilde{v} = \tilde{v} \diamond u \).

5. Questions from Textbook
Solve problems 3.1, 3.6 (the solution on the textbook webpage is not sufficiently clearly, I suggest you consider a particular case of 2-bit image, where 0 is obtained with frequency \( p \) and 1 with frequency \( 1 - p \) and show what you get by histogram equalization and then conclude on other discrete cases), 3.7, 3.11, 3.14, 3.21, 3.24, 3.29.

6. Bonus Problem: Best Low Rank Approximation of a Matrix

You do not need to submit this problem, but you will get bonus points if you solve it correctly (there is no partial credit).

If \( A \in \mathbb{R}^{m \times n} \) is a matrix with rank \( r \) and singular value decomposition \( A = \sum_{j=1}^{r} \sigma_j u_j v_j^T \) and if \( 0 \leq \nu \leq r \), then we denote

\[
A_{\nu} := \sum_{j=1}^{\nu} \sigma_j u_j v_j^T.
\]

You have used this \( \nu \)-rank approximation before in order to compress an image.

a) Prove that

\[
\|A - A_{\nu}\|_2 = \inf_{B \in \mathbb{R}^{m \times n} \atop \text{rank}(B) \leq \nu} \|A - B\|_2 = \sigma_{\nu+1}.
\]

b) Prove that

\[
\|A - A_{\nu}\|_F = \inf_{B \in \mathbb{R}^{m \times n} \atop \text{rank}(B) \leq \nu} \|A - B\|_F = \sqrt{\sum_{i=\nu+1}^{r} \sigma_i^2}.
\]

7. Bonus Problem: Geometric Interpretation of PCA

You do not need to submit this problem, but you will get bonus points if you solve it correctly (there is no partial credit).

Let \( \{x_i\}_{i=1}^{m} \) denote a set of \( m \) data points in \( \mathbb{R}^p \). If \( V \) is a \( d \)-dimensional plane in \( \mathbb{R}^p \) \( (d < p) \), we denote the \( l_2 \) averaged distance of the data set from \( V \) by \( \text{dist}_2(\{x_i\}_{i=1}^{m}, V) \). That is,

\[
\text{dist}_2(\{x_i\}_{i=1}^{m}, V) = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \text{dist}(x_i, V)^2}.
\]

Show that a minimizer of this distance among all \( d \)-planes in \( \mathbb{R}^p \) is

\[
V = \bar{c} + Sp\{v_1, \ldots, v_d\}, \tag{1}
\]

where \( \bar{c} \) is the center of mass (mean) of the data points and \( v_1, \ldots, v_d \) are the top \( d \) principal vectors (i.e. the top \( d \) right vectors of the centered data matrix). This \( d \)-plane is unique if and only if the principal values satisfy \( \sigma_d < \sigma_{d+1} \).
Guide: Represent any $d$-plane $V$ in the form of equation (1), where $c$ is an arbitrary vector in $\mathbb{R}^p$ and $\{v_1, \ldots, v_d\}$ is an arbitrary orthonormal system in $\mathbb{R}^p$; then express $\text{dist}_2(\{x_i\}_{i=1}^m, V)$ as a function of $c, v_1, \ldots, v_d$. Next, minimize $\text{dist}_2(\{x_i\}_{i=1}^m, V)$ as a function of $c$ for any fixed $v_1, \ldots, v_d$. The last step of specifying $Sp\{v_1, \ldots, v_d\}$ can be done in different ways, I will let you be creative.