CSE 250B: Machine learning **Winter 2019**

Homework 5 — Informative projections and singular value decomposition

1. *Projections.* Let $u_1, u_2 \in \mathbb{R}^d$ be two vectors with $\|u_1\| = \|u_2\| = 1$ and $u_1 \cdot u_2 = 0$. Define U to be the matrix whose columns are u_1 and u_2.

   (a) What are the dimensions of each of the following?
   - U
   - U^T
   - UU^T
   - $u_1 u_1^T$

   (b) What are the differences, if any, between the following four mappings?
   - $x \mapsto (u_1 \cdot x, u_2 \cdot x)$
   - $x \mapsto (u_1 \cdot x)u_1 + (u_2 \cdot x)u_2$
   - $x \mapsto U^T x$
   - $x \mapsto UU^T x$

2. A certain random variable $X \in \mathbb{R}^3$ has mean and covariance as follows:

$$\mathbb{E}X = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \text{cov}(X) = \begin{pmatrix} 5 & -3 & 0 \\ -3 & 5 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

   (a) Consider the direction $u = (1, 1, 1)/\sqrt{3}$. What are the mean and variance of $X \cdot u$?

   (b) The eigenvectors of $\text{cov}(X)$ can be found in the following list; which ones are they?

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

   (c) Find the eigenvalues corresponding to each of the eigenvectors in part (b). Make it clear which eigenvalue belongs to which eigenvector.

   (d) Suppose we used principal component analysis (PCA) to project points X into *two* dimensions. Which directions would it project onto?

   (e) Continuing from part (d), what would be the resulting two-dimensional projection of the point $x = (4, 0, 2)$?

   (f) Continuing from part (e), suppose that starting from the 2-d projection, we tried to reconstruct the original x. What would the three-dimensional reconstruction be, exactly?

3. M is a 2×2 real-valued symmetric matrix with eigenvalues $\lambda_1 = 2, \lambda_2 = -1$ and corresponding eigenvectors

$$u_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

(a) What is M?

(b) What are the eigenvalues of the matrix $M^2 = MM$?

4. *An experiment with PCA.* For this problem, we'll be using the *animals with attributes* data set. Go to

`http://attributes.kyb.tuebingen.mpg.de`

and, under "Downloads", choose the "base package" (the very first file in the list). Unzip it and look over the various text files.

(a) This is a small data set that has information about 50 animals. The animals are listed in `classes.txt`. For each animal, the information consists of values for 85 features: does the animal have a tail, is it slow, does it have tusks, etc. The details of the features are in `predicates.txt`. The full data consists of a 50×85 matrix of real values, in `predicate-matrix-continuous.txt`. Load this real-valued array.

(b) We would like to visualize these animals in 2-d. Show how to do this with a PCA projection from $\mathbb{R}^{85}$ to $\mathbb{R}^2$. Show the position of each animal, and label them with their names.

Python notes: You will need to make the plot larger by prefacing your code with

```
from pylab import rcParams
rcParams['figure.figsize'] = 10, 10
```

(or try a different size if this doesn't seem right).

5. *Singular values versus eigenvalues.* Recall from class that any $p \times q$ matrix M (with $p \le q$, say) can be written in the form:

$$M = \underbrace{\begin{pmatrix} \uparrow & & \uparrow \\ u_1 & \cdots & u_p \\ \downarrow & & \downarrow \end{pmatrix}}_{p \times p \text{ matrix } U} \underbrace{\begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_p \end{pmatrix}}_{p \times p \text{ matrix } \Lambda} \underbrace{\begin{pmatrix} \longleftarrow & v_1 & \longrightarrow \\ & \vdots & \\ \longleftarrow & v_p & \longrightarrow \end{pmatrix}}_{p \times q \text{ matrix } V^T}$$

where $u_1, \ldots, u_p$ are orthonormal vectors in $\mathbb{R}^p$, $v_1, \ldots, v_p$ are orthonormal vectors in $\mathbb{R}^q$, and $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_p \ge 0$ are known as *singular values.* In this problem, we will try to understand these quantities by relating them to eigenvalues and eigenvectors of suitably defined matrices.

(a) What is Mv_i (for $1 \le i \le p$)? Express the answer as simply as possible, in terms of the singular values and vectors of M.

(b) What is $M^T u_i$?

(c) What is $M^T M v_i$? And what is $MM^T u_i$?

(d) Notice that MM^T is a symmetric $p \times p$ matrix and therefore has p real eigenvalues. What are its eigenvalues and eigenvectors?

(e) How do the eigenvalues and eigenvectors of $M^T M$ relate to those of MM^T?

(f) Suppose M has rank k. How would this be reflected in the singular values σ_i?

6. A particular 4×5 matrix M has the following singular value decomposition:

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Find the best rank-2 approximation to M.

7. *Rank-1 matrices.*

   (a) Find the best rank-1 approximation to the matrix:

   $$M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

   If you use the SVD method in Python to solve this problem, you should use the setting `full_matrices` $=$ 0 to get the kind of decomposition we've been discussing (sometimes called the "thin SVD").

   (b) In general, a rank-1 matrix of dimension $p \times q$ is a matrix that can be written in form uv^T, where $u \in \mathbb{R}^p$ and $v \in \mathbb{R}^q$. Do you think this decomposition is unique, or is it in general possible to find a different pair of vectors $a \in \mathbb{R}^P$ and $b \in \mathbb{R}^q$ such that $uv^T = ab^T$?

   (c) Let M be some $p \times q$ matrix whose singular value decomposition is as specified in problem 5. Notice that M can equally be written in the form

   $$M = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_p u_p v_p^T,$$

   that is, as a sum of rank-1 matrices. For $k < p$, let $\widehat{M}$ be the best rank-k approximation to M. Express $\widehat{M}$ as a sum of rank-1 matrices.