ECE 403/503 Solutions to Assignment 1

Problems for this assignment: 1.1, 1.2, 1.3, and 1.4.

1.1 (2 points) Consider a data matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}$$

(a) Compute the eigenvalues of $\boldsymbol{AA}^T$ without any software.

Answer:

First, compute

$$\boldsymbol{AA}^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 9 \\ 9 & 14 \end{bmatrix}$$

Next, compute get and solve the characteristic equation

$$\det\left(\lambda \boldsymbol{I} - \boldsymbol{AA}^T\right) = \det \begin{bmatrix} \lambda - 6 & -9 \\ -9 & \lambda - 14 \end{bmatrix} = \lambda^2 - 20\lambda + 3 = 0$$

to obtain

$$\lambda_{1,2} = \frac{20 \pm \sqrt{400 - 12}}{2} = 10 \pm \sqrt{97} = 19.8489,\ 0.1511$$

(b) Compute the nonzero eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ with minimum amount of software: compute matrix $\boldsymbol{A}^T\boldsymbol{A}$ by hand, and use MATLAB function **`eig`** to get the eigenvalues.

Answer:

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 8 & 3 \\ 8 & 13 & 5 \\ 3 & 5 & 2 \end{bmatrix}$$

Using **`eig([5 8 3; 8 13 5; 3 5 2])`** yields three eigenvalues of matrix $\boldsymbol{A}^T\boldsymbol{A}$, two of which are nonzero and, up to 4-decimal places, given by

$$\lambda_1 = 19.8489 \text{ and } \lambda_2 = 0.1511.$$

We note that the nonzero eigenvalues of $\boldsymbol{AA}^T$ and $\boldsymbol{A}^T\boldsymbol{A}$ are identical.

(c) Use the results from part (a) or (b) to compute the nonzero singular values of $\boldsymbol{A}$.

Answer:

It is known that the nonzero singular values of $\boldsymbol{A}$ are equal to positive square roots of the nonzero eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$. Hence the results obtained in part (b) yields the singular values of $\boldsymbol{A}$ as

$$\sigma_1 = \sqrt{\lambda_1} = 4.4552 \text{ and } \sigma_2 = \sqrt{\lambda_2} = 0.3888$$

(d) Use MATLAB function **`svd`** to compute the SVD of matrix $\boldsymbol{A}$ (see Eq. (1.3) of the course notes). Then use the first pair of singular vectors $\boldsymbol{u}_1$ = `U(:,1)` and $\boldsymbol{v}_1$ = `V(:,1)` and first singular value σ_1 to construct a rank-1 approximation of $\boldsymbol{A}$ as $A_1 = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^T$. Report the numerical results.

Answer:

Using `A = [1 2 1; 2 3 1]); [U,S,V] = svd(A)`, the SVD of matrix A is found to be

$$\boldsymbol{U} = \begin{bmatrix} -0.5449 & -0.8385 \\ -0.8385 & 0.5449 \end{bmatrix},\ \boldsymbol{\Sigma} = \begin{bmatrix} 4.4552 & 0 & 0 \\ 0 & 0.3888 & 0 \end{bmatrix},\ \text{and}\ \boldsymbol{V} = \begin{bmatrix} -0.4987 & 0.6465 & 0.5774 \\ -0.8092 & -0.1087 & -0.5774 \\ -0.3105 & -0.7551 & 0.5774 \end{bmatrix}$$

Hence

$$\boldsymbol{A}_1 = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^T = 4.4552 \times \begin{bmatrix} -0.5449 \\ -0.8385 \end{bmatrix} \times \begin{bmatrix} -0.4987 & -0.8092 & -0.3105 \end{bmatrix} = \begin{bmatrix} 1.2107 & 1.9646 & 0.7538 \\ 1.8630 & 3.0230 & 1.1600 \end{bmatrix}$$

(e) Compare matrix $\boldsymbol{A}_1$ with the original data matrix $\boldsymbol{A}$ by inspection, and comment on how close they are. An objective evaluation of the approximation error can be made by computing the Frobenius norm of $\boldsymbol{A} - \boldsymbol{A}_1$: the Frobenius norm of a matrix $\boldsymbol{H} = \{h_{i,j}\} \in R^{n \times m}$ is defined by

$$\| \boldsymbol{H} \|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} h_{i,j}^2}$$

and it can be calculated using MATLAB as `norm(H,'fro')`. Use the result from part (d) to compute $\| \boldsymbol{A} - \boldsymbol{A}_1 \|_F$ and report your numerical result.

Answer:

By inspection the rank-1 approximation $\boldsymbol{A}_1$ is reasonably close to matrix $\boldsymbol{A}$. Objectively we compute the Frobenius norm of $\boldsymbol{A} - \boldsymbol{A}_1$ was found to be

$$\| \boldsymbol{A} - \boldsymbol{A}_1 \|_F = 0.3888$$

Note that the Frobenius norm of $\boldsymbol{A}$ is equal to $\| \boldsymbol{A} \|_F = 4.4721$ which implies a rather small relative approximation error

$$\frac{\| \boldsymbol{A} - \boldsymbol{A}_1 \|_F}{\| \boldsymbol{A} \|_F} = 0.0869$$

1.2 (2.5 points) Download image `building256.mat` from the course web site, then in MATLAB normalize the image to `A = building256/255;`

(a) Compute the SVD of $\boldsymbol{A}$ (see Eq. (1.3) of the course notes) using MATLAB function `svd`.

Answer:

Invoke MATALAB and execute `[U,S,V] = svd(A);`

(b) Denote the ith left- and right-singular vectors of $\boldsymbol{A}$ by $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$, respectively. Construct five rank-q approximations of $\boldsymbol{A}$ with $q = 1, 2, 3, 4, 5$ as follows:

$$\boldsymbol{A}_k = \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T \qquad \text{for } k = 1, 2, 3, 4, 5$$

Answer:

In MATLAB execute

```
A1 = S(1,1)*U(:,1)*V(:,1)';
A2 = A1 + S(2,2)*U(:,2)*V(:,2)';
A3 = A2 + S(3,3)*U(:,3)*V(:,3)';
```

```
A4 = A3 + S(4,4)*U(:,4)*V(:,4)';
A5 = A4 + S(5,5)*U(:,5)*V(:,5)';
```

(c) Compute the relative approximation errors $e_k = \dfrac{\| \boldsymbol{A}_k - \boldsymbol{A} \|_F}{\| \boldsymbol{A} \|_F}$ for $k = 1, 2, 3, 4, 5$ and report the numerical results.

Answer:

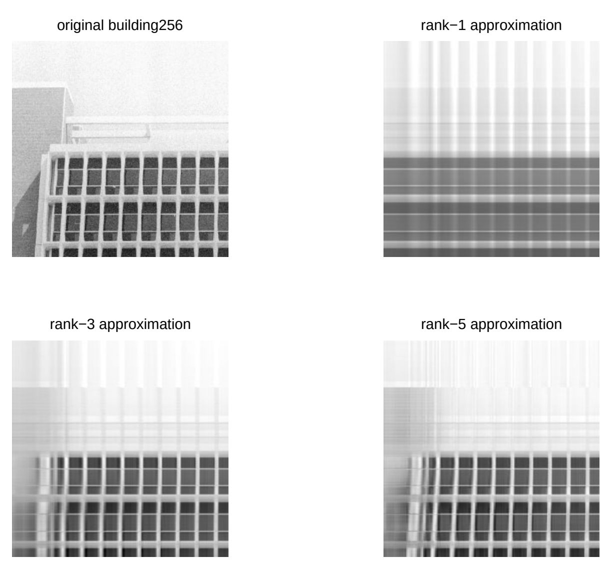
The five approximation errors are summarized as follows:

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} 0.1742 \\ 0.1145 \\ 0.0828 \\ 0.0713 \\ 0.0648 \end{bmatrix}$$

(d) Plot images **A1**, **A3**, and **A5** together with the original **A** in a single figure for comparison using the code below:

```
figure(1)
subplot(221)
imshow(A)
title('original building256')
subplot(222)
imshow(A1)
title('rank-1 approximation')
subplot(223)
imshow(A3)
title('rank-3 approximation')
subplot(224)
imshow(A5)
title('rank-5 approximation')
```

Answer:



(e) Image **`building256`** (hence matrix A) is of size 256×256, hence you need to save $256^2 = 65,536$ numbers to keep the image. Alternatively, if you happen to like what image **`A5`** offers, how many numbers do you need to save to keep **`A5`**? Report your numerical result.

Now if we denote that number by n_5, compute and report $\frac{256^2}{n_5}$ as the "compression ratio" .

Answer:

To keep image **`A5`**, we need to save 5 pairs of weighted (in order to absorb the singular value) singular vectors and hence a total of $n_5 = 5\times 2\times 256 = 2,560$ numbers. This achieves a compression ratio

$$\frac{256^2}{n_5} = \frac{256^2}{5\times 2\times 256} = 25.6$$

1.3 (3 points) Let A be a data matrix of size $d \times m$, and the compact version of the SVD of A of rank r be given by

$$\boldsymbol{A} = \boldsymbol{U}_r \boldsymbol{S}_r \boldsymbol{V}_r^T \tag{A1.1}$$

where $\boldsymbol{S}_r$ is a diagonal matrix with diagonal elements $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$, $\boldsymbol{U}_r = \begin{bmatrix}\boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r\end{bmatrix}$ and $\boldsymbol{V}_r = \begin{bmatrix}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r\end{bmatrix}$ are composed of the first r columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, which are associated with the r nonzero singular values, respectively. See Eq. (1.6) of the course notes.

(a) Show that if we have already computed the left-singular vectors $\{\boldsymbol{u}_i, i = 1, 2, \ldots, r\}$, then the right-singular vectors can be found by using $\boldsymbol{v}_i = \sigma_i^{-1}\boldsymbol{A}^T\boldsymbol{u}_i$ for $i = 1, 2, \ldots, r$.

Answer:

Since $\{\boldsymbol{u}_i, i = 1, 2, \ldots, r\}$ are the left-singular vectors of $\boldsymbol{A}$, by definition we have

$$\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{u}_i = \sigma_i^2\boldsymbol{u}_i \tag{A1.2}$$

To show $\boldsymbol{v}_i = \sigma_i^{-1}\boldsymbol{A}^T\boldsymbol{u}_i$ are right-singular vectors, we compute

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v}_i = \boldsymbol{A}^T\boldsymbol{A}\left(\sigma_i^{-1}\boldsymbol{A}^T\boldsymbol{u}_i\right) = \sigma_i^{-1}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{u}_i = \sigma_i^{-1}\boldsymbol{A}^T\sigma_i^2\boldsymbol{u}_i = \sigma_i^2\left(\sigma_i^{-1}\boldsymbol{A}^T\boldsymbol{u}_i\right) = \sigma_i^2\boldsymbol{v}_i$$

which shows $\boldsymbol{v}_i$ is an eigenvector of $\boldsymbol{A}^T\boldsymbol{A}$. In addition, we compute the inner product

$$\boldsymbol{v}_i^T\boldsymbol{v}_j = \left(\sigma_i^{-1}\boldsymbol{A}^T\boldsymbol{u}_i\right)^T\left(\sigma_j^{-1}\boldsymbol{A}^T\boldsymbol{u}_j\right) = \sigma_i^{-1}\sigma_j^{-1}\boldsymbol{u}_i^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{u}_j = \sigma_i^{-1}\sigma_j^{-1}\boldsymbol{u}_i^T\sigma_j^2\boldsymbol{u}_j = \sigma_i^{-1}\sigma_j\boldsymbol{u}_i^T\boldsymbol{u}_j = \begin{cases}1 & \text{if } i = j\\ 0 & \text{if } i \ne j\end{cases}$$

which shows $\{\boldsymbol{v}_i, i = 1, 2, \ldots, r\}$ are orthonormal. Hence they are right singular vectors of $\boldsymbol{A}$.

・ An alternative approach to prove $\boldsymbol{v}_i = \sigma_i^{-1}\boldsymbol{A}^T\boldsymbol{u}_i$ is to use the compact version of SVD (A1.1) (see (1.6) from the course notes): By transposing (A.11) we obtain

$$\boldsymbol{A}^T = \boldsymbol{V}_r\boldsymbol{S}_r\boldsymbol{U}_r^T$$

then multiplying the above expression by $\boldsymbol{U}_r\boldsymbol{S}_r^{-1}$ from right-hand side and using the fact that the columns of $\boldsymbol{U}_r$ are orthonormal to get

$$\boldsymbol{V}_r = \boldsymbol{A}^T\boldsymbol{U}_r\boldsymbol{S}_r^{-1} = \boldsymbol{A}^T\begin{bmatrix}\sigma_1^{-1}\boldsymbol{u}_1 & \sigma_2^{-1}\boldsymbol{u}_2 & \cdots & \sigma_r^{-1}\boldsymbol{u}_r\end{bmatrix} = \begin{bmatrix}\sigma_1^{-1}\boldsymbol{A}^T\boldsymbol{u}_1 & \sigma_2^{-1}\boldsymbol{A}^T\boldsymbol{u}_2 & \cdots & \sigma_r^{-1}\boldsymbol{A}^T\boldsymbol{u}_r\end{bmatrix}$$

which yields the formula immediately (this approach was provided by Mr. Siyang Liu of Mechanical Engineering Department).

(b) Similarly, show that if we have already computed the right-singular vectors $\{\boldsymbol{v}_i, i = 1, 2, \ldots, r\}$, then the left-singular vectors can be found by using $\boldsymbol{u}_i = \sigma_i^{-1} \boldsymbol{A} \boldsymbol{v}_i$ for $i = 1, 2, \ldots, r$.

Answer:

Since $\{\boldsymbol{v}_i, i = 1, 2, \ldots, r\}$ are the right-singular vectors of $\boldsymbol{A}$, by definition we have

$$\boldsymbol{A}^T \boldsymbol{A} \boldsymbol{v}_i = \sigma_i^2 \boldsymbol{v}_i \tag{A1.3}$$

To show $\boldsymbol{u}_i = \sigma_i^{-1} \boldsymbol{A} \boldsymbol{v}_i$ are left-singular vectors, we compute

$$\boldsymbol{A}\boldsymbol{A}^T \boldsymbol{u}_i = \boldsymbol{A}\boldsymbol{A}^T \left(\sigma_i^{-1} \boldsymbol{A} \boldsymbol{v}_i\right) = \sigma_i^{-1} \boldsymbol{A}\boldsymbol{A}^T \boldsymbol{A} \boldsymbol{v}_i = \sigma_i^{-1} \boldsymbol{A} \sigma_i^2 \boldsymbol{v}_i = \sigma_i^2 \left(\sigma_i^{-1} \boldsymbol{A} \boldsymbol{v}_i\right) = \sigma_i^2 \boldsymbol{u}_i$$

which shows $\boldsymbol{u}_i$ is an eigenvector of $\boldsymbol{A}\boldsymbol{A}^T$. In addition, we compute the inner product

$$\boldsymbol{u}_i^T \boldsymbol{u}_j = \left(\sigma_i^{-1} \boldsymbol{A} \boldsymbol{v}_i\right)^T \left(\sigma_j^{-1} \boldsymbol{A} \boldsymbol{v}_j\right) = \sigma_i^{-1} \sigma_j^{-1} \boldsymbol{v}_i^T \boldsymbol{A}^T \boldsymbol{A} \boldsymbol{v}_j = \sigma_i^{-1} \sigma_j^{-1} \boldsymbol{v}_i^T \sigma_j^2 \boldsymbol{v}_j = \sigma_i^{-1} \sigma_j \boldsymbol{v}_i^T \boldsymbol{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

which shows $\{\boldsymbol{u}_i, i = 1, 2, \ldots, r\}$ are orthonormal. Hence they are left singular vectors of $\boldsymbol{A}$.

・Like in part (a), alternatively we can show $\boldsymbol{u}_i = \sigma_i^{-1} \boldsymbol{A} \boldsymbol{v}_i$ by using the compact version of SVD (A1.1), namely

$$\boldsymbol{A} = \boldsymbol{U}_r \boldsymbol{S}_r \boldsymbol{V}_r^T$$

By multiplying the above expression by $\boldsymbol{V}_r \boldsymbol{S}_r^{-1}$ from right-hand size and the fact that the columns of $\boldsymbol{V}_r$ are orthonormal, we obtain

$$\boldsymbol{U}_r = \boldsymbol{A} \boldsymbol{V}_r \boldsymbol{S}_r^{-1} = \boldsymbol{A} \begin{bmatrix} \sigma_1^{-1} \boldsymbol{v}_1 & \sigma_2^{-1} \boldsymbol{v}_2 & \cdots & \sigma_r^{-1} \boldsymbol{v}_r \end{bmatrix} = \begin{bmatrix} \sigma_1^{-1} \boldsymbol{A} \boldsymbol{v}_1 & \sigma_2^{-1} \boldsymbol{A} \boldsymbol{v}_2 & \cdots & \sigma_r^{-1} \boldsymbol{A} \boldsymbol{v}_r \end{bmatrix}$$

which yields the formula immediately.

1.4 (2.5 point) Given an input data set consisting of four data points:

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 & \boldsymbol{x}_4 \end{bmatrix} = \begin{bmatrix} 0.6 & 1 & 1.6 & 2 \\ 0.4 & 1.2 & 1.3 & 2.3 \end{bmatrix}$$

(a) Find mean $\boldsymbol{\mu}$ of the dataset.

Answer:

$$\boldsymbol{\mu} = \frac{1}{4} \sum_{i=1}^{4} \boldsymbol{x}_i = \begin{bmatrix} 1.3 \\ 1.3 \end{bmatrix}$$

(b) Construct centralized dataset $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{x}_1 - \boldsymbol{\mu} & \boldsymbol{x}_2 - \boldsymbol{\mu} & \boldsymbol{x}_3 - \boldsymbol{\mu} & \boldsymbol{x}_4 - \boldsymbol{\mu} \end{bmatrix}$

Answer:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{x}_1 - \boldsymbol{\mu} & \boldsymbol{x}_2 - \boldsymbol{\mu} & \boldsymbol{x}_3 - \boldsymbol{\mu} & \boldsymbol{x}_4 - \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} -0.7 & -0.3 & 0.3 & 0.7 \\ -0.9 & -0.1 & 0 & 1 \end{bmatrix}$$

(c) Evaluate covariance matrix $\boldsymbol{C} = \frac{1}{4} \boldsymbol{A}\boldsymbol{A}^T$ and find the eigenvector $\boldsymbol{u}_1$ of matrix $\boldsymbol{C}$ corresponding to its largest eigenvalue using MATLAB function `eig`.

Answer:

$$C = \tfrac{1}{4} AA^T = \begin{bmatrix} 0.29 & 0.34 \\ 0.34 & 0.45 \end{bmatrix}$$

Using **`eig`** or **`eigs`**, the largest eigenvalue and the associated eigenvector are given by

$$\lambda_1 = 0.7224 \text{ and } \boldsymbol{u}_1 = \begin{bmatrix} 0.6181 \\ 0.7861 \end{bmatrix}$$

(d) Use the result from part (c) to compute the first principal component for each data point.

Answer: We use Eq. (1.12) of the course notes with $q = 1$ and $\boldsymbol{u}_1$ obtained from part (c) to compute the first principal components for the four data points as follows:

$$f_1 = \boldsymbol{u}_1^T (\boldsymbol{x}_1 - \boldsymbol{\mu}) = -1.1402$$

$$f_2 = \boldsymbol{u}_1^T (\boldsymbol{x}_2 - \boldsymbol{\mu}) = -0.2640$$

$$f_3 = \boldsymbol{u}_1^T (\boldsymbol{x}_3 - \boldsymbol{\mu}) = 0.1854$$

$$f_4 = \boldsymbol{u}_1^T (\boldsymbol{x}_4 - \boldsymbol{\mu}) = 1.2188$$

END