PURDUE UNIVERSITY · CS 51500
MATRIX COMPUTATIONS

HOMEWORK
Jun Cheng
September 18, 2015

Problem 0: Homework checklist

✓I didn't talk with any one about this homework.
✓Source-code are included at the end of this document.

Problem 1

1. $\begin{bmatrix}0 & 5\\0 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}5 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}$
2. $\begin{bmatrix}5 & 0\\2 & 0\end{bmatrix} = \begin{bmatrix}0.92 & -0.37\\0.37 & 0.92\end{bmatrix}\begin{bmatrix}5.4 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
3. $\begin{bmatrix}5 & -5\\2 & -2\\0 & 0\end{bmatrix} = \begin{bmatrix}0.93 & -0.37 & 0\\0.37 & 0.93 & 0\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}7.6 & 0\\0 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}0.707 & -0.707\\0.707 & 0.707\end{bmatrix}$
4. $\begin{bmatrix}2 & 0\\0 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}2 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$

Problem 2:

1. For example, The matrix

$$\boldsymbol{A} = \begin{bmatrix}5 & -5\\2 & -2\end{bmatrix} \tag{1}$$

From Matlab SVD we can get:

$$\boldsymbol{A} = \begin{bmatrix}5 & -5\\2 & -2\end{bmatrix} = \begin{bmatrix}-0.93 & -0.37\\-0.37 & 0.93\end{bmatrix}\begin{bmatrix}7.6 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}-0.707 & -0.707\\0.707 & -0.707\end{bmatrix} \tag{2}$$

But I got:

$$\boldsymbol{A} = \begin{bmatrix}5 & -5\\2 & -2\end{bmatrix} = \begin{bmatrix}0.93 & -0.37\\0.37 & 0.93\end{bmatrix}\begin{bmatrix}7.6 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}0.707 & -0.707\\0.707 & 0.707\end{bmatrix} \tag{3}$$

That means matrix A has two different decomposition, but in both cases they have the same singular value.

2.

3. Suppose $\boldsymbol{A}$ is a matrix with eigenvalues $\{\sigma_i\}$
The SVD of $\boldsymbol{A}$ is $\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^T$ where $\boldsymbol{U}$ and $\boldsymbol{V}$ are both orthogonal matrix.

$$\begin{aligned}\boldsymbol{A}^{-1} &= (\boldsymbol{U}\Sigma\boldsymbol{V}^T)^{-1} && (4)\\ &= (\boldsymbol{V}^T)^{-1}\Sigma^{-1}\boldsymbol{U}^{-1} && (5)\\ &= (\boldsymbol{V}^T)^T\Sigma^{-1}\boldsymbol{U}^T && (6)\\ &= \boldsymbol{V}\Sigma^{-1}\boldsymbol{U}^T && (7)\end{aligned}$$

Then $\boldsymbol{A}^{-1}$ is a matrix with eigenvalues $\{\frac{1}{\sigma_i}\}$

$$\|\boldsymbol{A}^{-1}\| = max\{\frac{1}{\sigma_i}\} \tag{8}$$
$$= \frac{1}{\sigma_{min}} \tag{9}$$
$$\tag{10}$$

4. The SVD of $\boldsymbol{Q}$ is $\boldsymbol{Q} = \boldsymbol{U}\Sigma\boldsymbol{V}^T = \boldsymbol{Q}\boldsymbol{I}\boldsymbol{I}^T$ where $\boldsymbol{I}$ is an identity matrix and it is also an orthogonal matrix.
   In particular, the singular value are all 1.

Problem 3:

1. If $m < n$, we can use $\boldsymbol{A}^T = (\boldsymbol{U}\Sigma\boldsymbol{V}^T)^T = \boldsymbol{V}\Sigma\boldsymbol{U}^T$ is still SVD.

2. $\boldsymbol{A} = diag(\sigma_1, \sigma_2, ...\sigma_n, 0, ...)$
   Then the best diagonal rank k approximation to $\boldsymbol{A}$ is
   $\boldsymbol{A}_k = diag(\sigma_1, \sigma_2, ...\sigma_k, 0, ...)$
   We aussume $k < n$ since we want to do the lower rank approximation.
   Then

$$\|\boldsymbol{A} - \boldsymbol{A}_k\|_2 = \|diag(\sigma_1, \sigma_2, ...\sigma_n, 0, ...) - diag(\sigma_1, \sigma_2, ...\sigma_k, 0, ...)\|_2 \tag{11}$$
$$= \|diag(0, .., \sigma_{k+1}, ...\sigma_n, 0, ...)\|_2 \tag{12}$$
$$= max\{\sigma_{k+1}, ...\sigma_n\} \tag{13}$$

   The small least answer could be σ_n which is the smallest singular value of $\boldsymbol{A}$.

3.

$$\|\boldsymbol{A} - \boldsymbol{A}_n\| = \|\sum_{i=1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T - \sum_{i=1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T\| \tag{14}$$
$$= 0 \tag{15}$$
$$\tag{16}$$

4.

$$\|\boldsymbol{A} - \boldsymbol{A}_R\| = \|\sum_{i=1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T - \sum_{i=1}^{R}\sigma_i\mathbf{u}_i\mathbf{v}_i^T\| \tag{17}$$
$$= \|\sum_{i=R+1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T\| \tag{18}$$
$$= max\{\sigma_i\} \tag{19}$$
$$\tag{20}$$

   When they have the same set of singular values, Then

$$\|\boldsymbol{A} - \boldsymbol{A}_R\| = 0 \tag{21}$$
$$\tag{22}$$

5.

$$\|\boldsymbol{A} - \boldsymbol{A}_k\| = \|\sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T - \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T\| \tag{23}$$

$$= \|\sum_{i=k+1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T\| \tag{24}$$

$$= max\{\sigma_{k+1}, ..., \sigma_n\} \tag{25}$$

$$= \sigma_{k+1} \tag{26}$$

σ_{k+1} is the largest singular value of $\boldsymbol{A} - \boldsymbol{A}_k$

6. Based on the definition of 2-norm, for any vectors:

$$\|\boldsymbol{A}\| = \sup_{x \neq 0} \frac{\|\boldsymbol{A}x\|}{\|x\|} \tag{27}$$

We know

$$\|(\boldsymbol{A} - \boldsymbol{B})x\| < \sigma_{k+1} \tag{28}$$

Then

$$\sigma_{k+1} > \sup_{x \neq 0} \frac{\|(\boldsymbol{A} - \boldsymbol{B})x\|}{\|x\|} \tag{29}$$

for any vector x.
Therefore

$$\|(\boldsymbol{A} - \boldsymbol{B})x\| < \sigma_{k+1}\|x\| \tag{30}$$

7. For a vector $\mathbf{w}$ in the null-space of $\boldsymbol{B}$

$$\boldsymbol{B}\mathbf{w} = 0 \tag{31}$$

Therefore

$$\|(\boldsymbol{A} - \boldsymbol{B})\mathbf{w}\| = \|\boldsymbol{A}\mathbf{w}\| < \sigma_{k+1}\|\mathbf{w}\| \tag{32}$$

8. For a vector $\mathbf{x} \in span(\mathbf{v}_1, ..., \mathbf{v}_{k+1})$,

$$\|\boldsymbol{A}\| = \sup_{x \neq 0} \frac{\|\boldsymbol{A}\mathbf{z}\|}{\|\mathbf{z}\|} \tag{33}$$

$$= max\{\sigma_1, \sigma_2, ..., \sigma_k\} \tag{34}$$

$$\leq \sigma_{k+1} \tag{35}$$

$$\tag{36}$$

Then

$$\|\boldsymbol{A}\mathbf{z}\| \leq \sigma_{k+1}\|\mathbf{z}\| \tag{37}$$

9. The dimension of the space where $\boldsymbol{A}\mathbf{w}$ is bounded above is $(n - k)$, and the dimension of the space where $\boldsymbol{A}\mathbf{z}$ is bounded above is $(k + 1)$. Because the sum of the dimensions of those two spaces is more than n, there must be a non-zero vector which belongs to two of the subspaces. And this is a contradiction.

10. Yes, I did.

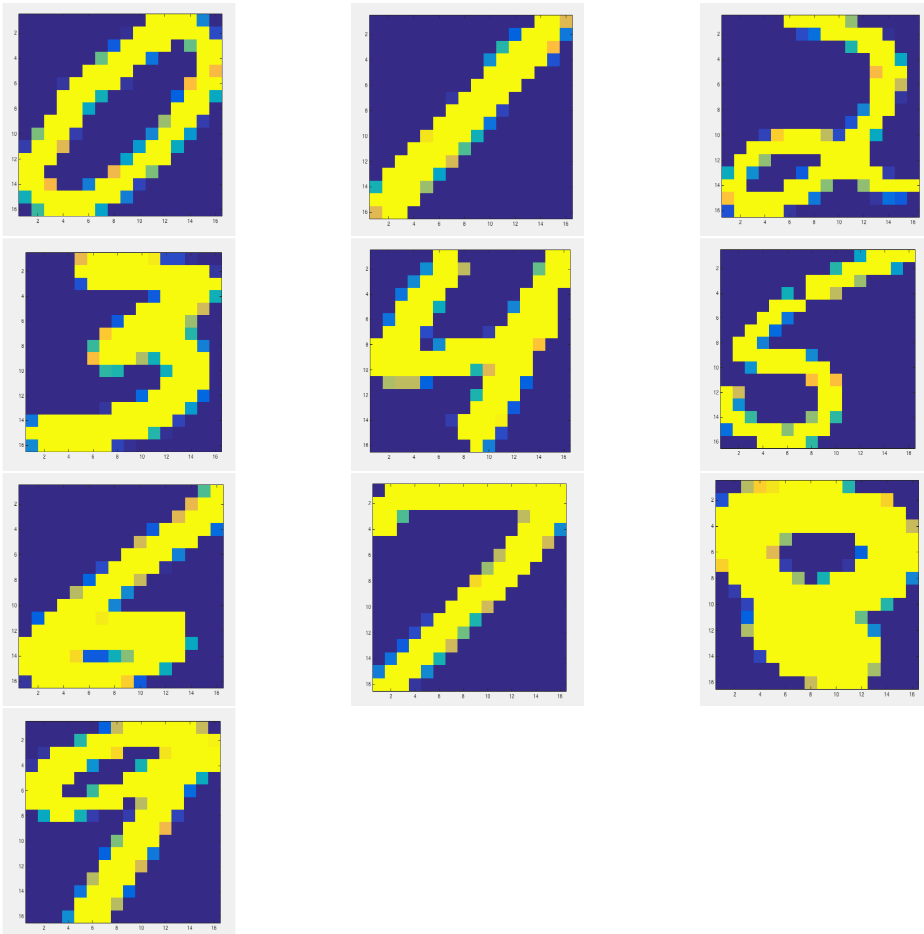

Figure 1: Sample plots for each individual digits from 0 to 9.

Problem 4:

1. The data is stored as a 3 dimension array $256 \times 1100 \times 10$.
   For each digit, there are 1100 images. Those images include 256 pixels with shape $16by16$
   And the sample images are shown as below.

2. To make sum of x to be zero, then

$$\sum_{i=0}^{256} x_i = \sum_{i=0}^{256} f_i + 256\gamma \tag{38}$$
$$= 0 \tag{39}$$

   Then

$$\gamma = -\frac{\sum_{i=0}^{256} f_i}{256} \tag{40}$$

   That means we just need to get summation of all values of all 1100 pictures for each digit, and then divided by 256.

3. Totally there are 1100×10 images. For each image we can get the average of those 256 pixels, and then subtract this average from the data image.

4. For each image matrix, we can get the singular values from Matlab SVD