PURDUE UNIVERSITY · CS 51500
MATRIX COMPUTATIONS

HOMEWORK
Jun Cheng
September 19, 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 decompositions, but in both cases they have the same singular value.

2. Suppose $\mathbf{f}$ and $\mathbf{g}$ are two vectors:

$$\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^T = \mathbf{f}'\Sigma\mathbf{g}' \tag{4}$$

where $\mathbf{f}'$ and $\mathbf{g}'$ are two orthogonal vectors.

$$\mathbf{f}' = \frac{\mathbf{f}}{\|\mathbf{f}\|} \tag{5}$$

$$\mathbf{g}' = \frac{\mathbf{g}^T}{\|\mathbf{g}\|} \tag{6}$$

$$\Sigma = \boldsymbol{I}\|\mathbf{f}\|\|\mathbf{g}\| \tag{7}$$

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.

$$\begin{aligned}
\boldsymbol{A}^{-1} &= (\boldsymbol{U}\Sigma\boldsymbol{V}^T)^{-1} && (8)\\
&= (\boldsymbol{V}^T)^{-1}\Sigma^{-1}\boldsymbol{U}^{-1} && (9)\\
&= (\boldsymbol{V}^T)^T\Sigma^{-1}\boldsymbol{U}^T && (10)\\
&= \boldsymbol{V}\Sigma^{-1}\boldsymbol{U}^T && (11)
\end{aligned}$$

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

$$\begin{aligned}
\|\boldsymbol{A}^{-1}\| &= max\{\frac{1}{\sigma_i}\} && (12)\\
&= \frac{1}{\sigma_{min}} && (13)\\
& && (14)
\end{aligned}$$

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

$$\begin{aligned}
\|\boldsymbol{A} - \boldsymbol{A}_k\|_2 &= \|diag(\sigma_1, \sigma_2, ...\sigma_n, 0, ...) - diag(\sigma_1, \sigma_2, ...\sigma_k, 0, ...)\|_2 && (15)\\
&= \|diag(0, .., \sigma_{k+1}, ...\sigma_n, 0, ...)\|_2 && (16)\\
&= max\{\sigma_{k+1}, ...\sigma_n\} && (17)
\end{aligned}$$

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

3.

$$\begin{aligned}
\|\boldsymbol{A} - \boldsymbol{A}_n\| &= \|\sum_{i=1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T - \sum_{i=1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T\| && (18)\\
&= 0 && (19)\\
& && (20)
\end{aligned}$$

4.

$$\begin{aligned}
\|\boldsymbol{A} - \boldsymbol{A}_R\| &= \|\sum_{i=1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T - \sum_{i=1}^{R}\sigma_i\mathbf{u}_i\mathbf{v}_i^T\| && (21)\\
&= \|\sum_{i=R+1}^{n}\sigma_i\mathbf{u}_i\mathbf{v}_i^T\| && (22)\\
&= max\{\sigma_i\} && (23)\\
& && (24)
\end{aligned}$$

When they have the same set of singular values, Then

$$\|\boldsymbol{A} - \boldsymbol{A}_R\| = 0 \tag{25}$$
$$\tag{26}$$

5.

$$\begin{aligned}
\|\boldsymbol{A} - \boldsymbol{A}_k\| &= \|\sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T - \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T\| && (27)\\
&= \|\sum_{i=k+1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T\| && (28)\\
&= max\{\sigma_{k+1}, ..., \sigma_n\} && (29)\\
&= \sigma_{k+1} && (30)
\end{aligned}$$

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:

$$\|\boldsymbol{A}\| = \sup_{x \neq 0} \frac{\|\boldsymbol{A}x\|}{\|x\|} \tag{31}$$

We know

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

Then

$$\sigma_{k+1} > \sup_{x \neq 0} \frac{\|(\boldsymbol{A} - \boldsymbol{B})x\|}{\|x\|} \tag{33}$$

for any vector x.
Therefore

$$\|(\boldsymbol{A} - \boldsymbol{B})x\| < \sigma_{k+1}\|x\| \tag{34}$$

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

$$\boldsymbol{B}\mathbf{w} = 0 \tag{35}$$

Therefore

$$\|(\boldsymbol{A} - \boldsymbol{B})\mathbf{w}\| = \|\boldsymbol{A}\mathbf{w}\| < \sigma_{k+1}\|\mathbf{w}\| \tag{36}$$

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

$$\begin{aligned}
\|\boldsymbol{A}\| &= \sup_{x \neq 0} \frac{\|\boldsymbol{A}\mathbf{z}\|}{\|\mathbf{z}\|} && (37)\\
&= max\{\sigma_1, \sigma_2, ..., \sigma_k\} && (38)\\
&\leq \sigma_{k+1} && (39)\\
& && (40)
\end{aligned}$$

Then

$$\|\boldsymbol{A}\mathbf{z}\| \leq \sigma_{k+1}\|\mathbf{z}\| \tag{41}$$

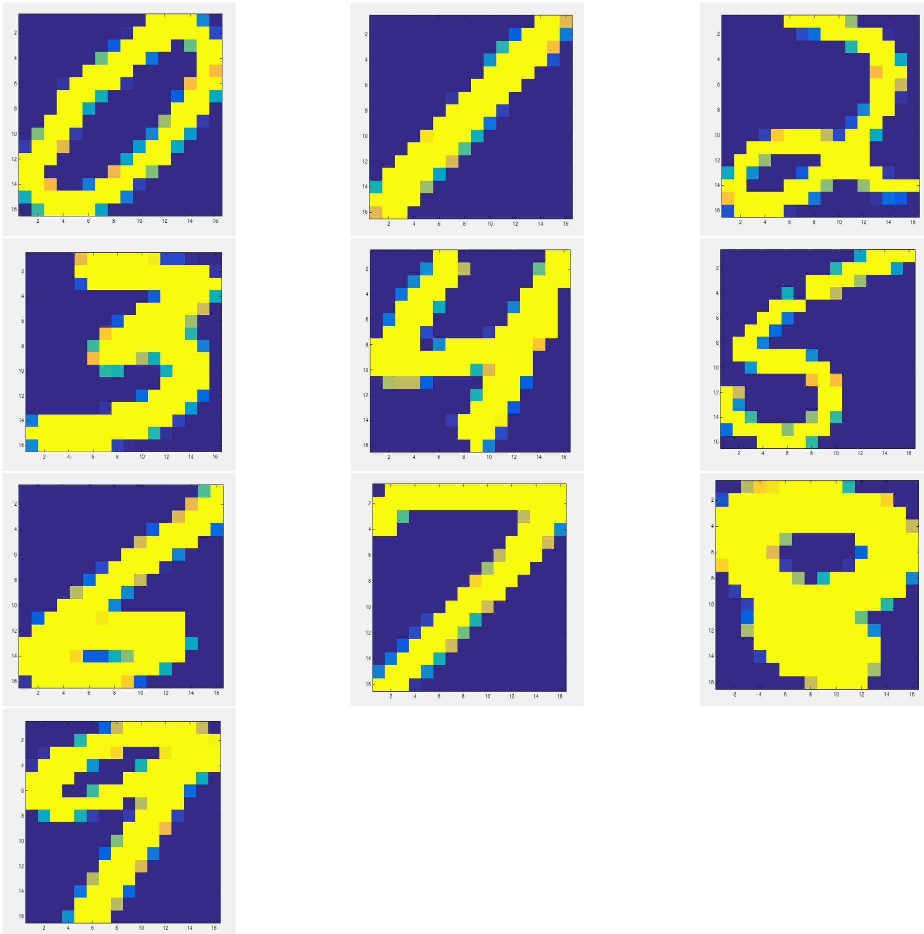

Figure 1: Sample plots for each individual digits from 0 to 9.

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.

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 $16 by 16$
   And the sample images are shown as below.

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

$$\begin{aligned} \sum_{i=0}^{256} x_i &= \sum_{i=0}^{256} f_i + 256\gamma \quad (42) \\ &= 0 \quad (43) \end{aligned}$$

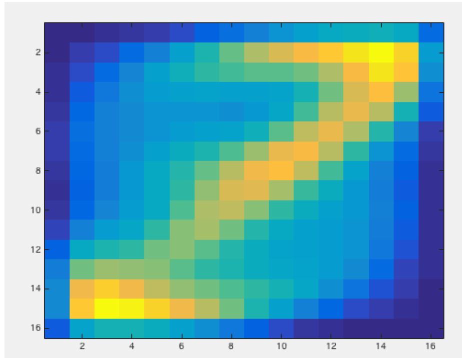

Figure 2: Plot of the leading singular vector $\mathbf{u}_1$ for problem4.4

Then

$$\gamma = -\frac{\sum_{i=0}^{256} f_i}{256} \tag{44}$$

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.

```
%% Subtract the mean from each image
residual =double(data) - repmat(mean(data),256, 1,1)
```

4. For each image matrix, we can get the singular values from Matlab SVD. The largest singular value is 58194.51. This is the global maximum singular value.

```
%% Reshape the matrix
m1 = reshape(residual, [256,11000])
[u,s, v] = svd(m1)
```

5. The leading singular vector $\mathbf{u}_1$ is shown as in Figure 2.

6. The dominant eigenvalues for 10 matrices are:
36414.5244628648
24385.2163297518
27837.9507648218
23367.3792733199
25695.6406646778
27847.3481015609
26588.9227446172
26005.8518124203
26587.9255579203
33791.0691296690

All the leading singular vectors are shown as in Figure 3.

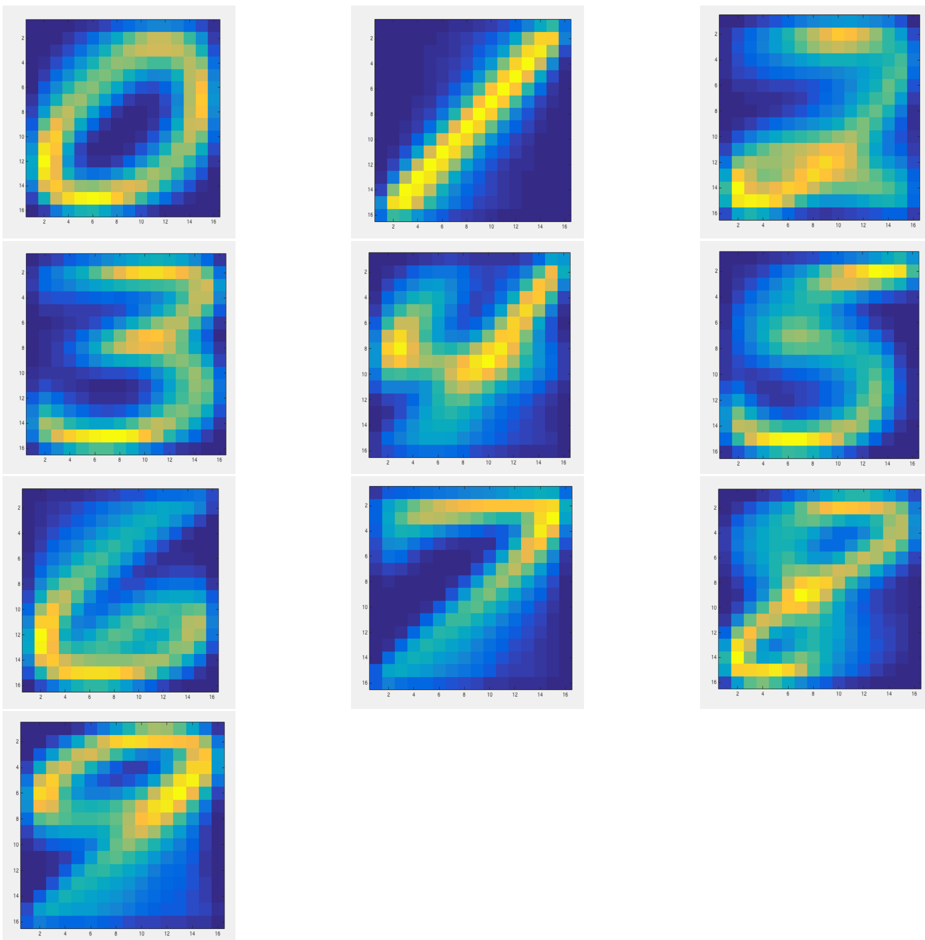

Figure 3: Plot of leading singular vectors for those 10 matrices

```
mm = zeros(1,10);
leadingV = zeros(256,10);

for i=1:10
    [u,s,¬] = svd(residual(:,:,i));
    mm(i)=s(1,1);
    leadingV(:,i)=u(:,1);
end
%% For example, plot the leading singular vector for digit 2.
imagesc(reshape(leadingV(:, 2), [16,16]))
```

The key difference is that the picture in number 5 include all features of digit 0 to 9. But in this problem we can see that all the images contain its own digit features.