CMDA 3606 · MATHEMATICAL MODELING II

Problem Set 6

Posted 13 March 2019. Due at 5pm on Thursday, 21 March 2019.

Special Note: Note that Problem 4 is optional; it is included as a 10-point bonus. Even if you don't solve Problem 4, you can still use its results to solve Problem 5.

Basic guidelines: Students may discuss the problems on this assignment, but each student must submit his or her individual writeup and code. (In particular, you *must write up your own individual MATLAB code.*) Students may consult class notes and other online resources for general information; cite all your sources and list those with whom you have discussed the problems.

1. [25 points: 6 points for (a),(c),(d); 7 points for (b)]

   Consider the matrix
$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}.$$

   (a) Compute the eigenvalues and eigenvectors of $\mathbf{A}^T\mathbf{A}$.

   (b) Write down a singular value decomposition, $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{3\times 2}$ and $\mathbf{V} \in \mathbb{R}^{2\times 2}$.

   (c) For this $\mathbf{A}$, use the $\mathbf{U}$, $\boldsymbol{\Sigma}$, and $\mathbf{V}$ matrices to compute the pseudoinverse
$$\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^T.$$

   (d) For this $\mathbf{A}$, use (some or all) of $\mathbf{U}$, $\boldsymbol{\Sigma}$, and $\mathbf{V}$ to compute the projector $\mathbf{A}\mathbf{A}^+$ onto $\mathcal{R}(\mathbf{A})$.

2. [25 points: 7 points each for (a) and (c); 11 points for (b)]

   Continue from Problem 1 with the matrix
$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}.$$

   (a) In MATLAB, form the matrix 5×5 matrix
$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{bmatrix}.$$

   Compute the eigenvalues of $\mathbf{B}$ using `eig`. How do they relate to the singular values of $\mathbf{A}$?

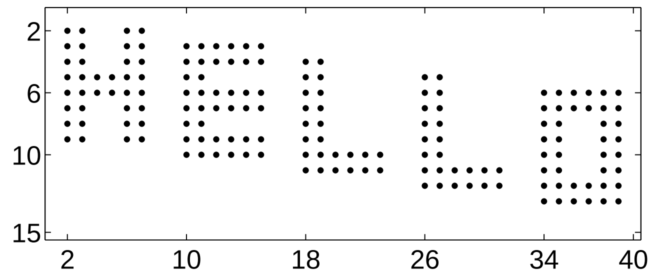
   (b) Consider now a general matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, and the matrix $\mathbf{B}$ as defined in part (a). We wish to show the connection between the eigenvalues of $\mathbf{B}$ and the singular values of $\mathbf{A}$ that you observed in part (a). To prove this connection, use the fact that each eigenvalue ϕ of $\mathbf{B}$ must satisfy
$$\begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \phi \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}.$$

   Recalling how the singular values of $\mathbf{A}$ relate to the eigenvalues of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$, show how ϕ relates to the singular values of $\mathbf{A}$.

   (c) How do the components $\mathbf{x}$ and $\mathbf{y}$ of the eigenvectors of $\mathbf{B}$ in part (b) relate to the left and right singular vectors of $\mathbf{A}$?

3. [25 points: 12 points for (a); 3 points for (b)]

   Recall the example of image compression using the singular value decomposition that we saw in class. Now it is your turn to compress an image, this time a simple bitmap. The routine `hello.m` on the class website constructs a 15×40 matrix $\mathbf{A}$ that is zero everywhere except for ones in the positions marked in the figure below. The upper-left point of the 'H' is in the (2,2) entry, and the bottom-right point of the 'O' is in the (13,39) entry.



   (a) What are the singular values of $\mathbf{A}$? (Use MATLAB's `svd` and `format long` to print 14 digits after the decimal point.) By counting the number of independent rows and columns of $\mathbf{A}$, determine the exact rank of $\mathbf{A}$. Does this agree with your output from `svd`?

   (b) For each k from 1 to rank($\mathbf{A}$), compute the rank-k matrix $\mathbf{A}_k$ that best approximates $\mathbf{A}$. Visualize this matrix using the commands:

```
imagesc(Ak)
colormap(flipud(gray))
```

   Submit print-outs of each of these images. (Please compress them onto one page if you can.)

   [Adapted from Trefethen and Bau, problem 9.3]

4. [10 bonus points: 2 points per part]

   A common problem in data analysis requires the alignment of two data sets, stored in the matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$. One seeks an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$ (i.e., $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$) that rigidly transforms the second data set, minimizing, in some sense, $\mathbf{A} - \mathbf{B}\mathbf{Q}$. The simplest solution arises when the mismatch is minimized in the *Frobenius norm*:

$$\min_{\substack{\mathbf{Q}\in\mathbb{R}^{n\times n}\\ \mathbf{Q}^T\mathbf{Q}=\mathbf{I}}} \|\mathbf{A} - \mathbf{B}\mathbf{Q}\|_{\mathrm{F}},$$

   where, for $\mathbf{X} \in \mathbb{R}^{m\times n}$,

$$\|\mathbf{X}\|_{\mathrm{F}} := \sqrt{\sum_{j=1}^{m}\sum_{k=1}^{n} |x_{j,k}|^2}.$$

   This is called the *orthogonal Procrustes problem*, named for a figure of Greek myth. ['On reaching Attic Corydallus, Theseus slew Sinis's father Polypemon, surnamed Procrustes, who lived beside the road and had two beds in his house, one small, the other large. Offering a night's lodging to travellers, he would lay the short men on the large bed, and rack them out to fit it; but the tall men on the small bed, sawing off as much of their legs as projected beyond it. Some say, however, that he used only one bed, and lengthened or shortened his lodgers according to its measure. In either case, Theseus served him as he had served others.' – Robert Graves, *The Greek Myths*, 1960 ed.]

   (a) Show that $\|\mathbf{X}\|_{\mathrm{F}}^2 = \mathrm{tr}(\mathbf{X}^T\mathbf{X})$, where $\mathrm{tr}(\cdot)$ denotes the *trace* of a square matrix. (The trace is the sum of the diagonal entries.)

(b) Show that $\mathrm{tr}(\mathbf{W}\mathbf{X}) = \mathrm{tr}(\mathbf{X}\mathbf{W})$ for any $\mathbf{X} \in \mathbb{R}^{m\times n}$ and $\mathbf{W} \in \mathbb{R}^{n\times m}$.

(c) Use properties of the trace to show that for any orthogonal $\mathbf{Q} \in \mathbb{R}^{n\times n}$ (i.e., $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$),

$$\|\mathbf{A} - \mathbf{B}\mathbf{Q}\|_\mathrm{F}^2 = \mathrm{tr}(\mathbf{A}^T\mathbf{A}) + \mathrm{tr}(\mathbf{B}^T\mathbf{B}) - 2\,(\mathrm{tr}(\mathbf{Q}^T\mathbf{B}^T\mathbf{A})).$$

(This implies $\|\mathbf{A}-\mathbf{B}\mathbf{Q}\|_\mathrm{F}$ is minimized by the orthogonal matrix $\mathbf{Q}$ that *maximizes* $\mathrm{tr}(\mathbf{Q}^T\mathbf{B}^T\mathbf{A})$.)

(d) Let $\mathbf{B}^T\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \in \mathbb{R}^{n\times n}$ denote the full singular value decomposition of $\mathbf{B}^T\mathbf{A}$. Show that

$$\mathrm{tr}(\mathbf{Q}^T\mathbf{B}^T\mathbf{A}) = \mathrm{tr}(\mathbf{\Sigma}\mathbf{Z}),$$

where $\mathbf{Z} = \mathbf{V}^T\mathbf{Q}^T\mathbf{U}$.

(e) Show that $\mathbf{Z}$ is orthogonal ($\mathbf{Z}^T\mathbf{Z} = \mathbf{I}$), and explain why this orthogonality of $\mathbf{Z}$ implies that $\mathrm{tr}(\mathbf{\Sigma}\mathbf{Z})$ is maximized when $\mathbf{Z} = \mathbf{I}$.

Then explain why this means that $\|\mathbf{A} - \mathbf{B}\mathbf{Q}\|_\mathrm{F}$ is minimized over orthogonal $\mathbf{Q}$ when $\mathbf{Q} = \mathbf{U}\mathbf{V}^T$.

5. [25 points: 3 points each for (a), (b); 4 points each for (c), (d), (f); 7 points for (e)]

This is a computational problem that follows on from Problem 4. (You do not need to solve Problem 4 to do this problem, but this problem will only make sense if you read through the description of Problem 4.)

On the class website you will find two MATLAB data files, `planck.mat` and `cow.mat`. When you `load` each file, you will obtain matrices $\mathbf{A}_0, \mathbf{B}_0 \in \mathbb{R}^{n\times 3}$. Each matrix describes an image in three dimensional space (a bust of Max Planck and a cow, respectively), with each row of the matrix giving the (x, y, z) coordinates of one data point. For example, use `plot3(A0(:,1),A0(:,2),A0(:,3),'k.')` to view an image. The image $\mathbf{A}_0$ should be regarded as the 'exact' image; the $\mathbf{B}_0$ image has been distorted in various ways. Your goal is to manipulate these images in a way that best aligns them, in the sense of the orthogonal Procrustes problem.

For each of the two data files, complete the following steps.

(a) Use `plot3`, followed by `axis equal`, to plot the $\mathbf{A}_0$ image; print this out.

(b) Use `plot3`, followed by `axis equal`, to plot the $\mathbf{B}_0$ image; print this out.

(c) Center the $\mathbf{A}_0$ and $\mathbf{B}_0$ images by subtracting from each point the mean x, y, and z values; call the results $\mathbf{A}_c$ and $\mathbf{B}_c$. (The mean of each column of $\mathbf{A}_c$ and $\mathbf{B}_c$ should be zero.)

(d) Divide $\mathbf{A}_c$ and $\mathbf{B}_c$ each by a scalar, so that the largest magnitude entry in each matrix has magnitude 1. Call these normalized matrices $\mathbf{A}$ and $\mathbf{B}$.

(e) To solve the orthogonal Procrustes problem (see the details given in Problem 4), compute the full SVD $\mathbf{B}^T\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ and set $\mathbf{Q} = \mathbf{U}\mathbf{V}^T$.

(f) Produce a new plot (`plot3`, followed by `axis equal`) showing the $\mathbf{A}$ image as dots (`'.'`) and the $\mathbf{B}\mathbf{Q}$ image as circles (`'o'`). You should see decent overall agreement, despite the noise and distortions that polluted the original $\mathbf{B}_0$ image.

[Unperturbed data was derived from polygonal models available from
`http://www.cs.princeton.edu/gfx/proj/sugcon/models/`]

This problem uses simple toy data sets, but you can imagine how you could apply this technology, for example, to align data from 3d scans of medical or satellite images, taken from different angles or at different times.