CAAM 440 · APPLIED MATRIX ANALYSIS

Problem Set 3

Posted Friday 17 February 2012. Due Friday 24 February 2012.
[Late work due Monday 5 March 2012.]

Complete any four problems (1–6), 25 points each.

General definition used in several problems: The *trace* of a square matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is the sum of the diagonal elements:

$$\mathrm{tr}(\mathbf{A}) = \sum_{j=1}^{n} a_{j,j}.$$

0. [No credit: a helpful warm-up if you have not seen the SVD before]

   Determine *by hand calculation*, the full singular value decompositions of the matrices

$$\text{(a)}\ \begin{bmatrix} 3 & 4 \end{bmatrix} \quad \text{(b)}\ \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix} \quad \text{(c)}\ \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{(d)}\ \begin{bmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{(e)}\ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

   [See Trefethen and Bau, problem 4.1]

1. (a) Consider the Hermitian matrix

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & \alpha \end{bmatrix}.$$

   For what values of α is $\mathbf{A}$ positive definite?

   (b) A Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is *diagonally dominant* if the magnitude of each diagonal entry is larger than the sum of the off-diagonal element in the same column (or row), i.e.,

$$|a_{k,k}| > \sum_{j=1, j\neq k}^{n} |a_{j,k}|.$$

   We shall later see that if $\mathbf{A}$ has positive entries on the diagonal, then diagonal dominance implies that $\mathbf{A}$ is positive definite. Show by example that a positive definite matrix need not be diagonally dominant.

2. In class we proved that two diagonalizable matrices $\mathbf{A}, \mathbf{B}$ commute ($\mathbf{AB} = \mathbf{BA}$) if and only if they are simultaneously diagonalizable. Now we investigate other aspects of the products $\mathbf{AB}$ and $\mathbf{BA}$.

   (a) Let $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ be general matrices (not necessarily Hermitian, or even diagonalizable). Prove that $\mathbf{AB}$ and $\mathbf{BA}$ have the same eigenvalues.

   (b) Show that $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ for all $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$.

   (c) Construct 2×2 matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\|\mathbf{AB} - \mathbf{I}\| \ll 1$ but $\|\mathbf{BA} - \mathbf{I}\| \geq 1$. (That is, construct $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{B}$ is a good approximate right inverse to $\mathbf{A}$, but a poor approximate left inverse.) You do not need to rigorously compute the norms – it is enough to observe that all entries of $\mathbf{AB} - \mathbf{I}$ are small, etc. Ideal solutions will give $\mathbf{A}$ and $\mathbf{B}$ in terms of a parameter ε, so that the behavior becomes increasingly extreme as $\varepsilon \to 0$.

3. Many applications give rise to matrices of the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{K} & \mathbf{B}^* \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \in \mathbb{C}^{(n+m)\times(n+m)},$$

where $\mathbf{K} \in \mathbb{C}^{n\times n}$ is *Hermitian positive definite* and $\mathbf{B} \in \mathbb{C}^{m\times n}$ has linearly independent rows, i.e., $\mathcal{N}(\mathbf{B}^*) = \{\mathbf{0}\}$. (Such matrices arise, for example, in constrained optimization, where Lagrange multipliers lead to the matrix $\mathbf{B}$, and in incompressible Stokes fluid flow, where $\mathbf{B}$ comes from the pressure variables and incompressibility condition.)

   (a) Show that $\mathbf{A}$ is nonsingular, i.e., $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$.
   Hint: Show that if $\mathbf{A}\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$.

   (b) Compute the congruence transformation $\mathbf{C}\mathbf{A}\mathbf{C}^*$ for

   $$\mathbf{C} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{B}\mathbf{K}^{-1} & \mathbf{I} \end{bmatrix}.$$

   (c) Show that $-\mathbf{B}\mathbf{K}^{-1}\mathbf{B}^*$ is negative definite.

   (d) Given your answer to part (b), use Sylvester's Law of Inertia to determine the number of positive and negative eigenvalues of $\mathbf{A}$.

[See Benzi, Golub, Liesen (2005) for this and many more details.]

4. A common problem in data analysis requires the alignment of two data sets, stored in the matrices $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$. One seeks a unitary matrix $\mathbf{Q}$ 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*:

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

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

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

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 $\|\mathbf{X}\|_{\mathrm{F}}^2 = \mathrm{tr}(\mathbf{X}^*\mathbf{X})$, where $\mathrm{tr}(\cdot)$ denotes the *trace* of a square matrix.

   (b) Use Problem 2(b) to show that for any unitary $\mathbf{Q}$,

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

   It follows that $\|\mathbf{A} - \mathbf{B}\mathbf{Q}\|_{\mathrm{F}}$ is minimized by the unitary matrix $\mathbf{Q}$ that *maximizes* $\mathrm{Re}(\mathrm{tr}(\mathbf{Q}^*\mathbf{B}^*\mathbf{A}))$.

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

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

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

(d) Show that $\mathbf{Z}$ is unitary, and explain why this implies that $\mathrm{Re}(\mathrm{tr}(\mathbf{\Sigma Z}))$ is maximized when $\mathbf{Z} = \mathbf{I}$. What, then, is the unitary $\mathbf{Q}$ that minimizes $\|\mathbf{A} - \mathbf{BQ}\|_{\mathrm{F}}$?

5. This is a computational problem that follows on from 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 $\mathbf{A}_0, \mathbf{B}_0 \in \mathbb{C}^{n\times 3}$. 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) Solve the orthogonal Procrustes problem, as in Problem 4, to find the unitary matrix $\mathbf{Q}$ that minimizes $\|\mathbf{A} - \mathbf{BQ}\|_{\mathrm{F}}$.

   (f) Produce a new plot (`plot3`, followed by `axis equal`) showing the $\mathbf{A}$ image as dots (`'.'`) and the $\mathbf{BQ}$ 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/`]

6. Typical damped mechanical systems give rise to differential equations of the form

$$\mathbf{x}''(t) = -\mathbf{Kx}(t) - \mathbf{Dx}'(t),$$

   where the stiffness matrix $\mathbf{K} \in \mathbb{C}^{n\times n}$ is Hermitian positive definite with eigenvalues $\kappa_1 \le \kappa_2 \le \cdots \le \kappa_n$, and the damping matrix $\mathbf{D} \in \mathbb{C}^{n\times n}$ is Hermitian positive semidefinite, with eigenvalues $\delta_1 \le \delta_2 \le \cdots \le \delta_n$. We write the differential equation in the form

$$\begin{bmatrix} \mathbf{x}'(t) \\ \mathbf{x}''(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}'(t) \end{bmatrix}.$$

   Call this $2n \times 2n$ matrix $\mathbf{A}$.

   Prove that if λ is an eigenvalue of $\mathbf{A}$ *with a nontrivial imaginary part*, then

$$-\frac{1}{2}\delta_n \le \mathrm{Re}\lambda \le -\frac{1}{2}\delta_1$$

   and

$$\kappa_1 \le |\lambda|^2 \le \kappa_n.$$

   (These bounds *do not* apply to purely real eigenvalues.)
   Draw a sketch showing the region of the complex plane to which these bounds restrict λ.

   [Falk; Lancaster]