# MATH 5524 · MATRIX THEORY

## Problem Set 3

Posted Wednesday 15 March 2017. Due Wednesday 22 March 2017.

Complete any four problems, 25 points each.

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

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

Recall from class that if $\mathbf{X} \in \mathbb{C}^{m\times n}$ and $\mathbf{Y} \in \mathbb{C}^{n\times m}$, then $\mathrm{tr}(\mathbf{XY}) = \mathrm{tr}(\mathbf{YX})$.

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. This problem concerns *generalized eigenvalue problems*. Let $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$. The matrix pair $(\mathbf{A}, \mathbf{B})$ is said to have the eigenvalue $\lambda \in \mathbb{C}$ if there exists some nonzero $\mathbf{v} \in \mathbb{C}^n$ such that

$$\mathbf{Av} = \lambda \mathbf{Bv}.$$

   We then write $\lambda \in \sigma(\mathbf{A}, \mathbf{B})$.

   (a) Show that if $\mathbf{A}$ and $\mathbf{B}$ are Hermitian positive definite, then all the eigenvalues of $(\mathbf{A}, \mathbf{B})$ are *positive real numbers.*

   (b) What happens if $\mathbf{A}$ is Hermitian positive definite but $\mathbf{B}$ is only Hermitian positive *semidefinite* but not positive definite? (Trying an example with $n = 2$ might help you build your intuition.)

   (c) Suppose $\mathbf{A}$ is Hermitian positive definite but $\mathbf{B}$ is Hermitian and invertible (not necessarily positive definite). Prove that the eigenvalues $\lambda \in \sigma(\mathbf{A}, \mathbf{B})$ must be real.

   (d) Construct an example where $\mathbf{A}$ and $\mathbf{B}$ are both Hermitian and invertible, but there exist eigenvalues $\lambda \in \sigma(\mathbf{A}, \mathbf{B})$ that are not real. (You should work the example out by hand; do not simply give numerical evidence.)

2. 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 $\mathbf{Ax} = \mathbf{0}$, then $\mathbf{x} = \mathbf{0}$.

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

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

(c) Show that $-\mathbf{BK}^{-1}\mathbf{B}^*$ 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.]

3. 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, $\mathbf{A} - \mathbf{BQ}$. The simplest solution arises when the mismatch is minimized in the *Frobenius norm*:

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

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 properties of the trace to show that for any unitary $\mathbf{Q}$,

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

It follows that $\|\mathbf{A}-\mathbf{BQ}\|_{\mathrm{F}}$ 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 $\|\mathbf{A} - \mathbf{BQ}\|_{\mathrm{F}}$?

4. 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/`]

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

$$\mathbf{x}''(t) = -\mathbf{K}\mathbf{x}(t) - \mathbf{D}\mathbf{x}'(t),$$

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

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

Define

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{D} \end{bmatrix} \in \mathbb{C}^{2n\times 2n}.$$

Prove that if $\lambda$ is an eigenvalue of $\mathbf{A}$ *with a nonzero 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 $\lambda$.

[Falk; Lancaster]

6. Throughout this semester we have considered the standard (Euclidean) inner product $\mathbf{y}^*\mathbf{x}$, where $\mathbf{y}^*$ denotes the conjugate-transpose of $\mathbf{y}$. This is but one example of a much broader class of potential inner products; these functions play an essential role in applications, where often the inner product provides a way to measure the energy in a system. This problem asks you to investigate how the usual inner product relates to these more sophisticated alternatives.

   (a) Let $\mathbf{H} \in \mathbb{C}^{n\times n}$ be a Hermitian positive definite matrix, and define the $\mathbf{H}$-inner product to be

   $$\langle \mathbf{x}, \mathbf{y}\rangle_{\mathbf{H}} := \mathbf{y}^*\mathbf{H}\mathbf{x}.$$

   Prove that the $\mathbf{H}$-inner product satisfies the fundamental axioms required of all inner products:

      i. $\langle \mathbf{x}, \mathbf{x}\rangle_{\mathbf{H}} \geq 0$ for all $\mathbf{x} \in \mathbb{C}^n$ and $\langle \mathbf{x}, \mathbf{x}\rangle_{\mathbf{H}} = 0$ if and only if $\mathbf{x} = \mathbf{0}$;
      ii. $\langle \mathbf{x}, \mathbf{y}\rangle_{\mathbf{H}} = \overline{\langle \mathbf{y}, \mathbf{x}\rangle_{\mathbf{H}}}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ (complex conjugate symmetry);
      iii. $\langle \mathbf{x}+\mathbf{y}, \mathbf{z}\rangle_{\mathbf{H}} = \langle \mathbf{x}, \mathbf{z}\rangle_{\mathbf{H}} + \langle \mathbf{y}, \mathbf{z}\rangle_{\mathbf{H}}$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^n$;
      iv. $\langle \alpha\mathbf{x}, \mathbf{y}\rangle_{\mathbf{H}} = \alpha\langle \mathbf{x}, \mathbf{y}\rangle_{\mathbf{H}}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$.

   (b) In the standard inner product, we often deal with the *conjugate transpose*, $\mathbf{A}^* = \overline{\mathbf{A}}^{\mathrm{T}}$. This definition is inherently tied to our inner product. We define the $\mathbf{H}$-*adjoint* of $\mathbf{A}$ to be the matrix $\mathbf{A}^\sharp$ such that, for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

   $$\langle \mathbf{A}\mathbf{x}, \mathbf{y}\rangle_{\mathbf{H}} = \langle \mathbf{x}, \mathbf{A}^\sharp\mathbf{y}\rangle_{\mathbf{H}}.$$

   Determine a formula for $\mathbf{A}^\sharp$ as a combination of $\mathbf{A}$, $\mathbf{A}^*$, and/or $\mathbf{H}$.

   (c) We defined a matrix to be *Hermitian* (or *self-adjoint*) provided $\mathbf{A}^* = \mathbf{A}$. Similarly, we define $\mathbf{A}$ to be $\mathbf{H}$-*self-adjoint* provided $\mathbf{A}^\sharp = \mathbf{A}$. Determine a simple condition involving $\mathbf{A}$, $\mathbf{A}^*$, and/or $\mathbf{H}$ to test whether the matrix $\mathbf{A}$ is $\mathbf{H}$-self-adjoint.

   (Similar conditions can be defined to test whether $\mathbf{A}$ is $\mathbf{H}$-*unitary* and $\mathbf{H}$-*normal*, but you do not need to derive these.)

   (d) The $\mathbf{H}$-inner product leads to a $\mathbf{H}$-norm for vectors,

   $$\|\mathbf{x}\|_{\mathbf{H}} := \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle_{\mathbf{H}}}$$

   and an induced $\mathbf{H}$-matrix norm for $\mathbf{A} \in \mathbb{C}^{n\times n}$,

   $$\|\mathbf{A}\|_{\mathbf{H}} := \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_{\mathbf{H}}}{\|\mathbf{x}\|_{\mathbf{H}}}.$$

   Suppose you have a function for computing the usual matrix norm induced by the standard Euclidean inner product (e.g., the `norm` command in MATLAB), and you wanted to use this norm routine to compute $\mathbf{H}$-norm of a matrix.

   Determine some matrix $\mathbf{B}$ (which may involve matrices like $\mathbf{A}$, $\mathbf{H}$, $\mathbf{H}^{1/2}$, etc.) such that

   $$\|\mathbf{A}\|_{\mathbf{H}} = \|\mathbf{B}\|,$$

   where the norm on the right is the usual norm we have been dealing with all semester.

   (e) Prove that if $\langle\cdot,\cdot\rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ is any function that obeys the four axioms in part (a), then there exists a Hermitian positive definite matrix $\mathbf{H} \in \mathbb{C}^{n\times n}$ such that for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$,

   $$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{y}^*\mathbf{H}\mathbf{x}.$$

7. This problem combines elements from Problem 1 and Problem 2. Finite element discretizations of the incompressible Navier–Stokes equations often lead to the dynamical system

$$\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}'(t) \\ \mathbf{p}'(t) \end{bmatrix} = \begin{bmatrix} \mathbf{K} & \mathbf{B}^* \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{p}(t) \end{bmatrix}.$$

We can abbreviate the matrices as

$$\mathbf{A} = \begin{bmatrix} \mathbf{K} & \mathbf{B}^* \\ \mathbf{B} & \mathbf{0} \end{bmatrix}, \qquad \mathbf{E} = \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{M} \in \mathbb{C}^{n\times n}$ is Hermitian positive definite, $\mathbf{K} \in \mathbb{C}^{n\times n}$ is invertible (but not necessarily Hermitian), and $\mathbf{B} \in \mathbf{C}^{m\times n}$ is full rank, with $m \le n$.

A first step to understanding the dynamical system is to analyze the associated generalized eigenvalue problem: find nonzero $\mathbf{v} \in \mathbb{C}^{n+m}$ and $\lambda \in \mathbb{C}$ such that $\mathbf{A}\mathbf{v} = \lambda\mathbf{E}\mathbf{v}$. We write $\lambda \in \sigma(\mathbf{A}, \mathbf{E})$.

(a) Prove that $\sigma(\mathbf{A}, \mathbf{E})$ contains (at most) $n - m$ finite numbers, and give a form for the associated eigenvectors.

(b) Since $\mathbf{A}, \mathbf{E} \in \mathbb{C}^{(m+n)\times(m+n)}$, there seem to be $2m$ eigenvalues missing in part (a). Show that "$\lambda = \infty$" is an eigenvalue of $(\mathbf{A}, \mathbf{E})$ with algebraic multiplicity $2m$. (This is equivalent to saying that $0 \in \sigma(\mathbf{A}^{-1}\mathbf{E})$ has algebraic multiplicity $2m$.)

To prove this fact, find $m$ linearly independent vectors $\{\mathbf{v}_j\}_{j=1}^m$ for which $\mathbf{v}_j \in \mathcal{N}(0\mathbf{I} - \mathbf{A}^{-1}\mathbf{E})$. Then find another $m$ linearly independent vectors $\{\mathbf{v}_j^{(1)}\}_{j=1}^m$ (not in the span of $\mathbf{v}_1, \ldots, \mathbf{v}_m$) such that $\mathbf{v}_j^{(1)} \in \mathcal{N}((0\mathbf{I} - \mathbf{A}^{-1}\mathbf{E})^2)$.

(This result essentially shows that the pair $(\mathbf{A}, \mathbf{E})$ has $m$ Jordan blocks of size $2 \times 2$ associated with $\lambda = \infty$.)