MATH 5524 · MATRIX THEORY

Pledged Problem Set 2

Posted Tuesday 25 April 2017. Due by 5pm on Wednesday 3 May 2017.

Complete any four problems, 25 points each. *You are welcome to complete more problems if you like, but specify which four you want to be graded.*

Rules: On this pledged problem set, you are welcome to use course notes (those you have taken, and those posted to the web), books (hard copies or electronic), MATLAB, Mathematica, etc. You may post questions to Piazza, and ask questions of the instructor during office hours. You are *not allowed* to more generally search the web for answers, or discuss the problems with anyone aside from the instructor. Please write out and sign the pledge ("As a Hokie, I will conduct myself with honor and integrity at all times. I will not lie, cheat or steal, nor will I acccept the actions of those who do.") on your assignment. **Pledged problem sets will not be accepted late unless you have made a previous arrangement with the instructor.**

1. (a) Compute the spectral radius and Perron vector (the eigenvector $\mathbf{x} > \mathbf{0}$ associated with $\rho(\mathbf{A})$ having $\sum_{j=1}^n x_j = 1$) for the matrix

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

   with $\alpha + \beta = 1$ and $\alpha, \beta \in [0, 1]$. [Meyer]

   (b) Interpret the matrix $\mathbf{A}$ in part (a) as the transition matrix for a two-state Markov chain. What is $\mathbf{A}^\infty := \lim_{k\to\infty} \mathbf{A}^k$? Describe the limiting steady state, $\mathbf{p}_\infty^* := \mathbf{p}_0^* \mathbf{A}^\infty$. At what rate is this limit reached? Does this depend on the values of α and β?

2. (a) Suppose that $\mathbf{A} \in \mathbb{R}^{n\times n}$ is positive, $\mathbf{A} > \mathbf{0}$. Prove that

$$\min_{1\le j\le n} \sum_{k=1}^n a_{j,k} \le \rho(\mathbf{A}) \le \max_{1\le j\le n} \sum_{k=1}^n a_{j,k},$$

   where $\rho(\mathbf{A})$ denotes the spectral radius of $\mathbf{A}$. [Meyer]

   (b) Computationally verify these bounds by computing

$$\min_{1\le j\le n} \sum_{k=1}^n a_{j,k}, \qquad \rho(\mathbf{A}), \qquad \max_{1\le j\le n} \sum_{k=1}^n a_{j,k}$$

   for matrices with uniform random entries in $[0, 1/n]$, i.e., `A = rand(n)/n`, for one matrix of each of the dimensions `n` $= 16, 64, 256$ and 1024.

3. Recall the Jordan block

$$\mathbf{J}_n = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \in \mathbb{C}^{n\times n},$$

   and, for $z \in \mathbb{C}$, define the vector

$$\mathbf{v}_n = \begin{bmatrix} 1 \\ z \\ z^2 \\ \vdots \\ z^{n-1} \end{bmatrix} \in \mathbb{C}^n.$$

(a) Let $|z| < 1$. Derive a formula (involving only z and n) for

$$\frac{\|\mathbf{J}_n\mathbf{v}_n - z\mathbf{v}_n\|}{\|\mathbf{v}_n\|}.$$

How does this formula behave as $n \to \infty$?
[Notice that such z are ε-pseudoeigenvalues for $\mathbf{J}_n$ for increasingly small values of ε as n increases.]

Now consider the *Toeplitz matrix*

$$\mathbf{A}_n = \begin{bmatrix} a_0 & a_1 & a_2 & & \\ & a_0 & a_1 & \ddots & \\ & & \ddots & \ddots & a_2 \\ & & & a_0 & a_1 \\ & & & & a_0 \end{bmatrix} \in \mathbb{C}^{n\times n}.$$

which has the constant entries on diagonals, and define the function $a(z) = a_0 + a_1 z + a_2 z^2$.

(b) For the vector $\mathbf{v}_n$ specified above, compute $\|\mathbf{A}_n\mathbf{v}_n - a(z)\mathbf{v}_n\|$.

(c) If $|z| < 1$, how does $\|\mathbf{A}_n\mathbf{v}_n - a(z)\mathbf{v}_n\|$ behave as $n \to \infty$?
What does this imply about $\|(a(z)\mathbf{I} - \mathbf{A}_n)^{-1}\|$ as $n \to \infty$, for $|z| < 1$?

(d) Take $a_0 = 0$, $a_1 = 2$, and $a_2 = 1$. Plot the curve

$$a(\mathbb{T}) := \{a(\mathrm{e}^{i\theta}) : \theta \in [0, 2\pi)\}$$

and indicate the area $\{a(z) : |z| < 1\}$.
You can plot this curve in MATLAB via

```
T = exp(linspace(0,2i*pi,500));
aT = a0*T.^0 + a1*T.^1 + a2*T.^2;
plot(real(aT),imag(aT),'-');
```

(e) For dimensions $n = 25, 50, 100$ and 200, construct $\mathbf{A}_n$ and plot the eigenvalues of $\mathbf{A}_n + \mathbf{E}$ for 50 random matrices $\mathbf{E}$ of norm 10^{-3}:

```
E = randn(n); E = 1e-3*E/norm(E);
```

(f) [optional] The same theory extends (with a different form of $\mathbf{v}$) for banded Toeplitz matrices,

$$\mathbf{A}_n = \begin{bmatrix} a_0 & a_1 & \cdots & a_p & & \\ a_{-1} & a_0 & \ddots & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \ddots & a_p \\ a_{-m} & \ddots & \ddots & \ddots & \ddots & \vdots \\ & \ddots & \ddots & \ddots & \ddots & a_1 \\ & & a_{-m} & \cdots & a_{-1} & a_0 \end{bmatrix} \in \mathbb{C}^{n\times n}$$

with

$$a(z) = \sum_{j=-m}^{p} a_j z^j.$$

Repeat parts (d) and (e) for the values

$$a_{-1} = -1, \quad a_0 = a_1 = a_2 = a_3 = 1,$$

which gives the *Grcar matrix*.

4. Suppose $\mathbf{A}$ is an adjacency matrix for an undirected graph with n vertices, i.e.,

$$a_{j,k} = \begin{cases} 1, & \text{an edge connects vertex } j \text{ to vertex } k; \\ 0, & \text{otherwise.} \end{cases}$$

In particular, the diagonal of $\mathbf{A}$ is zero.

In class we noted that the number of paths of length m from node j to node k equals the (j,k) entry of $\mathbf{A}^m$. Inspired by this method of path-counting, Estrada and Rodríguez-Velázquez (2005) proposed a way to gauge the importance of each vertex as a weighted sum of the number of paths from a given vertex back to itself. In particular, the *subgraph centrality* (or *Estrada index*) of vertex j is given by

$$(\mathrm{e}^{\mathbf{A}})_{j,j}.$$

The higher this value, the more central a vertex is regarded. Estrada and Hatano (2007) extend this notion to measure the *communicability* of vertices j and k as

$$(\mathrm{e}^{\mathbf{A}})_{j,k}.$$

(This work is surveyed and extended in a paper by Estrada and D. J. Higham, *SIAM Review*, 2010.)

We now wish to apply these ideas to real data. Analyzing a variety of media sources, Valdis Krebs has produced a graph of the 9/11 terrorist network, including both the hijackers and their accomplices; see `http://www.orgnet.com/hijackers.html` for details. We have written a routine `sept11.m`, that creates an adjacency matrix for Krebs's graph, given an arbitrary vertex numbering. The graph contains 69 vertices, 19 of which correspond to the hijackers. (These nodes are identified at the top of `sept11.m`, with the pilots singled out.)

(a) Compute the subgraph centrality for all the vertices, and present a table showing the results for the top thirty vertices. Each row of this table should correspond to a vertex, and include:

   i. subgraph centrality rank (in order, 1,…, 30);
   ii. vertex number, j;
   iii. subgraph centrality value, $(\mathrm{e}^{\mathbf{A}})_{j,j}$;
   iv. number of edges connected to vertex j;

(b) In your table, clearly identify the four pilots (nodes 36, 44, 50, 65; see `sept11.m` for details).

(c) We might alternatively have ranked the vertices by their edge counts, rather than subgraph centrality. Does your table reveal any anomalous vertices that are ranked more highly than their edge counts would suggest, or vertices with large edge counts that have relatively low centrality scores?

(d) Now use `imagesc` in MATLAB (or similar) to visualize the magnitude of the entries in $\mathrm{e}^{\mathbf{A}}$. This allows you to visually inspect the *communicability* of all vertices with one another at once. Designate the rows and columns corresponding to the hijackers of each plane. You may do this by hand, or using some slick command along the lines of, e.g.,

```
fill([0 70 70 0],[31.5 31.5 36.5 36.5],'w','facealpha',.3,'edgealpha',.3)
```

as you like.

(e) Interpret your plot in part (d). Were the different crews of hijackers highly connected with one another? Do you notice anything about the crew of United Flight 93, which crashed in Pennsylvania as a result of passenger intervention?

5. Suppose that $\mathbf{A} \in \mathbb{R}^{n\times n}$ with $\mathbf{A} > \mathbf{0}$ for parts (a)–(c).

   (a) Argue, by way of similarity transformation, that the result of Problem 2(a) can be adapted to

   $$\min_{1\le j\le n} \frac{1}{x_j}\sum_{k=1}^{n} a_{j,k}x_k \le \rho(\mathbf{A}) \le \max_{1\le j\le n} \frac{1}{x_j}\sum_{k=1}^{n} a_{j,k}x_k,$$

   for *any* $\mathbf{x} > \mathbf{0}$. (Note that $\mathbf{x}$ need not be the Perron vector.)

   (b) Show that $\rho(\mathbf{A})$ is the only eigenvalue of $\mathbf{A}$ that has an eigenvector with all entries positive.

   (c) Prove the "max min" characterization of $\rho(\mathbf{A})$ for positive matrices:

   $$\rho(\mathbf{A}) = \max_{\mathbf{x}>\mathbf{0}} \min_{1\le j\le n} \frac{1}{x_j}\sum_{k=1}^{n} a_{j,k}x_k = \min_{\mathbf{x}>\mathbf{0}} \max_{1\le j\le n} \frac{1}{x_j}\sum_{k=1}^{n} a_{j,k}x_k.$$

   (d) Show by example that the result of part (b) need not hold if we only require $\mathbf{A}$ and its eigenvector to be *nonnegative*. That is, build $\mathbf{A} \ge \mathbf{0}$ with eigenvalue $\lambda \ne \rho(\mathbf{A})$ that has a nonnegative eigenvector.

   [Horn and Johnson]

6. (a) Suppose that $\mathbf{A} \in \mathbb{R}^{n\times n}$ is nonnegative, $\mathbf{A} \ge \mathbf{0}$. Prove that conditions (i) and (ii) below are equivalent. [Varga]

   (i) $\alpha > \rho(\mathbf{A})$.

   (ii) $\alpha\mathbf{I} - \mathbf{A}$ is invertible and $(\alpha\mathbf{I} - \mathbf{A})^{-1} \ge \mathbf{0}$.

   (b) The finite difference discretization of the differential equation $-u''(x) = f(x)$ for $x \in [0,1]$ with $u(0) = u(1) = 0$ on a grid of points $x_0, \ldots, x_{n+1}$ with $x_j = jh$ for $h = 1/(n+1)$ leads to a matrix equation of the form

   $$h^{-2}\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}.$$

   Denote this equation as $\mathbf{Au} = \mathbf{f}$ (note that $\mathbf{A}$ includes the h^{-2} factor). Use the fact that

   $$\sigma(\mathbf{A}) = \{h^{-2}(2 + 2\cos(k\pi/(n+1))) : k = 1, \ldots, n\}$$

   (which you do not need to prove) and the result in (a) to show that $\mathbf{A}^{-1} \ge \mathbf{0}$.
   (This $\mathbf{A}$ is an example of an *M-matrix*, an important class of matrices for many applications.)

7. *Banded matrices* arise in many applications; for example, problems 3 and 6 on this problem set involve such matrices. We say that that $\mathbf{A}$ has *bandwidth* b provided

   $$a_{j,k} = 0 \quad \text{whenever } |j-k| > b.$$

   Thus, a *tridiagonal* matrix like the one in question 6(b) has bandwidth $b = 1$.

   Banded matrices can be stored efficiently: an $n \times n$ matrix with bandwidth b needs only $\mathcal{O}(nb)$ storage, rather than $\mathcal{O}(n^2)$ storage for a general (*dense*) matrix. Unfortunately, for many functions f, the matrix $f(\mathbf{A})$ is *dense*, even when $\mathbf{A}$ is banded.

In this problem, you will derive simple versions of some more sophisticated results that show how the entries of $f(\mathbf{A})$ must decay rapidly away from the main diagonal. (This idea has been suggested in connection with the construction of *preconditioners* that approximate $f(\mathbf{A}) = \mathbf{A}^{-1}$.)

Throughout this problem, you may use without proof the following fact (due to Bernstein in 1912). Suppose $f : \mathbb{C} \to \mathbb{C}$ is analytic in an ellipse in $\mathbb{C}$ containing the interval $[a, b]$. Then there exist constants $C > 0$ and $\rho \in [0, 1)$ such that

$$\min_{p \in \mathcal{P}_m} \max_{x \in [a,b]} |f(x) - p(x)| \leq C\rho^m,$$

where $\mathcal{P}_m$ denotes the set of all polynomials of degree m or less.

Important: throughout this problem, assume $\mathbf{A} \in \mathbb{C}^{n \times n}$ *is Hermitian positive definite.*

(a) Suppose $\mathbf{A}$ is tridiagonal (bandwidth $b = 1$), and let $m \geq 1$ be a positive integer. What entries in $\mathbf{A}^m$ must be zero?

(b) Again let $b = 1$. Given some position (j, k) (with $j, k \in \{1, \ldots, n\}$), what is the largest value of m for which $(p(\mathbf{A}))_{j,k} = 0$ for all polynomials $p \in \mathcal{P}_m$?

(c) For the m identified in part (b), explain why (for $b = 1$)

$$|(f(\mathbf{A}))_{j,k}| \leq \min_{p \in \mathcal{P}_m} \|f(\mathbf{A}) - p(\mathbf{A})\|$$

and hence there exist $C > 0$ and $\rho \in [0, 1)$ such that

$$|(f(\mathbf{A}))_{j,k}| \leq C\rho^m.$$

(d) Explain how you can change the bound in part (c) to accommodate Hermitian positive definite $\mathbf{A}$ with generic bandwidth $b \geq 1$.

(e) Compute the entries of $\mathbf{A}^{-1}$ for the tridiagonal matrices

$$\mathbf{A}_1 = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}, \qquad \mathbf{A}_2 = \begin{bmatrix} 4 & -1 & & \\ -1 & 4 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 4 \end{bmatrix}.$$

of dimension $n = 32$. Do not print out all your results, but find some way to visualize them. (For example, show a `surf` plot of the entries of $|\mathbf{A}^{-1}|$, or a `semilogy` plot showing the largest magnitude entry on each diagonal.) How do the two examples differ?

(f) Compute $f(\mathbf{A}) = \mathrm{e}^{\mathbf{A}}$ for the matrix $\mathbf{A}_1$ from part (e), and visualize the entries as in part (e).

(g) Interpret your results from parts (e) and (f) in light of the bound in (c). Be as specific as possible, and be sure to explain the behavior of $\mathbf{A}_1^{-1}$ versus $\mathbf{A}_2^{-1}$ in part (e) and $\mathbf{A}_1^{-1}$ versus $\mathrm{e}^{\mathbf{A}_1}$ in part (f).