18.06 Professor Edelman Quiz 3 December 3, 2012

Grading

1

2

3

4

Your PRINTED name is: ____________________

Please circle your recitation:

1	T 9	2-132	Andrey Grinshpun	2-349	3-7578	agrinshp
2	T 10	2-132	Rosalie Belanger-Rioux	2-331	3-5029	robr
3	T 10	2-146	Andrey Grinshpun	2-349	3-7578	agrinshp
4	T 11	2-132	Rosalie Belanger-Rioux	2-331	3-5029	robr
5	T 12	2-132	Geoffroy Horel	2-490	3-4094	ghorel
6	T 1	2-132	Tiankai Liu	2-491	3-4091	tiankai
7	T 2	2-132	Tiankai Liu	2-491	3-4091	tiankai

1 (16 pts.)

a) (4 pts.) Suppose C is $n \times n$ and positive definite. If A is $n \times m$ and $M = A^T CA$ is not positive definite, find the smallest eigenvalue of M. (Explain briefly.)

Solution. The smallest eigenvalue of M is 0.

The problem only asks for brief explanations, but to help students understand the material better, I will give lengthy ones.

First of all, note that $M^T = A^T C^T A = A^T CA = M$, so M is symmetric. That implies that all the eigenvalues of M are real. (Otherwise, the question wouldn't even make sense; what would the "smallest" of a set of complex numbers mean?)

Since we are assuming that M is *not* positive definite, at least one of its eigenvalues must be nonpositive. So, to solve the problem, we just have to explain why M cannot have any negative eigenvalues. The explanation is that M is **positive semidefinite**. That's the buzzword we were looking for.

Why is M positive semidefinite? Well, note that, since C is positive definite, we know that for every vector y in $\mathbb{R}^n$

$$y^T Cy \geqslant 0,$$

with equality if and only if y is the zero vector. Then, for any vector x in $\mathbb{R}^m$, we may set $y = Ax$, and see that

$$x^T Mx = x^T A^T CAx = (Ax)^T C(Ax) \geqslant 0. \tag{*}$$

Since M is symmetric, the fact that $x^T Mx$ is always non-negative means that M is positive semidefinite. Such a matrix never has negative eigenvalues. Why? Well, if M did have a negative eigenvalue, say $\lambda < 0$, with a corresponding eigenvector $v \neq 0$, then

$$v^T Mv = v^T(\lambda v) = \lambda v^T v = \lambda \|v\|^2 < 0,$$

which would contradict $(*)$ above, which is supposed to hold for *every* x in $\mathbb{R}^m$. $\square$

Remark: Some students wrote that M is similar to C, but this is totally false. In the given problem, if $m \neq n$, then M and C don't even have the same dimensions, so they cannot possibly be similar. (Remember that two $n \times n$ matrices A and B are similar if there is an invertible $n \times n$ matrix M such that $A = M^{-1}BM$, which isn't usually the same thing as M^TBM, unless M is an orthogonal matrix.)

b) (12 pts.) If A is symmetric, which of these four matrices are necessarily positive definite? A^3, $(A^2+I)^{-1}$, $A+I$, e^A. (Explain briefly.)

Solution. The answer is that $(A^2+I)^{-1}$ and e^A have to be positive definite, but A^3 and $A+I$ don't.

The key is to use **the $Q\Lambda Q^{-1}$ factorization**. Let me remind you what that is. Since A is symmetric, there is an *orthonormal basis* of $\mathbb{R}^n$ (if A is an $n \times n$ matrix) consisting of eigenvectors $q_1, q_2, \ldots, q_n$ of A, and the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all *real*. Form an $n \times n$ matrix Q whose columns are these n eigenvectors $q_1, q_2, \ldots, q_n$, and let Λ be a diagonal $n \times n$ matrix whose diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$, so that $A = Q\Lambda Q^{-1}$. (In case you're wondering, it would also be correct to write $A = Q\Lambda Q^T$. Since Q is an orthogonal matrix, $Q^{-1} = Q^T$.)

Note that all four matrices we are asked to discuss are symmetric. So the question of positive definiteness is just a question about the positivity of their eigenvalues.

- $A^3 = (Q\Lambda Q^{-1})^3 = Q\Lambda^3 Q^{-1}$, so A^3 **is *similar* to** Λ^3, and these two matrices have the same eigenvalues. But Λ^3 is just the diagonal matrix whose diagonal entries are ${\lambda_1}^3, {\lambda_2}^3, \ldots, {\lambda_n}^3$. Do these numbers all have to be positive? Of course not. For example, we could have $A = \Lambda = 0$, the zero matrix. Then $A^3 = \Lambda^3 = 0$, which isn't positive definite.

- Before we discuss $(A^2+I)^{-1}$, let's check that this actually makes sense, i.e., that A^2+I is really invertible. Well,

$$A^2 + I = (Q\Lambda Q^{-1})^2 + I = Q(\Lambda^2 + I)Q^{-1}.$$

  Now $\Lambda^2 + I$ is a diagonal matrix whose diagonal entries ${\lambda_1}^2 + 1$, ${\lambda_2}^2 + 1$, $\ldots$, ${\lambda_n}^2 + 1$ are all nonzero, so $\Lambda^2 + I$ really is invertible. Then $A^2 + I$, which is similar to $\Lambda^2 + I$, must also be invertible, and in fact we can write down its inverse:

$$(A^2 + I)^{-1} = Q(\Lambda^2 + 1)^{-1}Q^{-1}.$$

  Now $(A^2 + I)^{-1}$ **is *similar* to** $(\Lambda^2 + 1)^{-1}$, and these two matrices have the same eigenvalues, namely $({\lambda_1}^2 + 1)^{-1}$, $({\lambda_2}^2 + 1)^{-1}, \ldots,$ $({\lambda_n}^2 + 1)^{-1}$. These eigenvalues are all positive, because $(\lambda^2 + 1)^{-1} > 0$ for any real number λ. So $(A^2 + I)^{-1}$ is positive definite.

- $A + I = Q(\Lambda + I)Q^{-1}$, so $A + I$ **is *similar* to** $\Lambda + I$, and these two matrices have the same eigenvalues, namely $\lambda_1 + 1$, $\lambda_2 + 1, \ldots,$ $\lambda_n + 1$. Do these numbers all have to be positive? Of course not. For example, we could have $A = -I$. Then $A + I = 0$, which isn't positive definite.

- Finally, we have e^A. Note that

$$e^A = e^{Q\Lambda Q^{-1}} = \sum_{k=0}^{\infty} \frac{1}{k!}(Q\Lambda Q^{-1})^k = \sum_{k=0}^{\infty} \frac{1}{k!} Q\Lambda^k Q^{-1} = Q\left(\sum_{k=0}^{\infty} \frac{1}{k!}\Lambda^k\right) Q^{-1} = Qe^{\Lambda}Q^{-1},$$

  so e^A **is *similar* to** e^{Λ}. But e^{Λ} is just the diagonal matrix with diagonal entries $e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n}$, which are all positive, because $e^{\lambda} > 0$ for all real λ. So the eigenvalues of e^A are all positive, and e^A must be positive definite.

You see, diagonalization allows us to reduce a problem about matrices to a problem about real numbers. The general philosophy is this: If A is similar to a diagonal matrix to Λ, then often some expression[1] in A is similar to the same expression in Λ, and the expression in

[1] Here I mean a polynomial (e.g., A^3 or $A + I$; think of I as being akin to the constant 1), a rational function (e.g., $(A^2 + I)^{-1}$), or a convergent power series (e.g., e^A) in the variable A alone. We do *not* allow expressions involving A^T in addition to A, or anything more complicated than that.

Λ can be computed just by plugging in the diagonal entries one by one. So the question basically comes to this: which of the functions λ^3, $(\lambda^2+1)^{-1}$, $\lambda+1$, e^λ is everywhere positive (i.e., positive for all real λ)? Of course, your solution should explain *why* it comes to this. $\square$

Remarks: (i) Some students thought that A must itself be positive definite. Some even wrote a "proof" that all symmetric matrices are positive definite! Please disabuse yourself of this notion. Positive definite matrices (at least the ones with real entries) are required to be symmetric, but there are lots of symmetric matrices that aren't positive definite: for example, 0 and $-I$. (ii) Some students discussed only the matrices that *are* necessarily positive definite, and didn't write anything at all about A^3 and $I+A$. A complete solution should convince people that it is correct. And in order to convince people that "$(A^2+I)^{-1}$ and e^A" is the correct answer, one should explain *both* why these two matrices are necessarily positive definite, *and* why the other two aren't.

2 (30 pts.)

Let $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

a) (6 pts.) What are the eigenvalues of A ? (Explain briefly.)

This matrix is upper triangular. For such a matrix, the determinant is the product of the diagonal entries. Using this observation, if we try to compute $|A - xI|$, we find $-x^3$. This implies that the only eigenvalue is 0 with multiplicity 3.

b) (6 pts.) What is the rank of A?

It is clear that the last two columns of A are pivot columns. Therefore, the rank is 2.

c) (6 pts.) What are the singular values of A?

The singular values of A are the square roots of the eigenvalues of $A^T A$.

$$A^T A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

We have $|xI - A^T A| = x((x-1)(x-2) - 1) = x(x^2 - 3x + 1)$

The roots are 0, $\frac{3+\sqrt{5}}{2}$, $\frac{3-\sqrt{5}}{2}$. Therefore, the singular values of A are 0, $\sqrt{\frac{3+\sqrt{5}}{2}}$ and $\sqrt{\frac{3-\sqrt{5}}{2}}$.

d) (6 pts.) What is the Jordan form of A? (Explain briefly.)

In general, the Jordan form has zeroes everywhere except on the diagonal where you put the eigenvalues on the second diagonal where you have 1 and 0. Note that the matrix A as it is is not in Jordan normal form because you have a 1 in the upper right corner. There are 3 possibilities for what the Jordan form can be. One with two ones over the diagonal and two with one one and one zero. To determine which is the actual Jordan form, you can look at the rank. We know that A has rank 2 and the Jordan form is similar to A so it must have rank 2 as well. Therefore, the only possibility is :

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

e) (6 pts.) Compute in simplest form e^{tA}.

We can use the series expression for e^{tA}. In general this is an infinite sum which is unpleasant but in this particular case, the powers of A quickly become zero. Indeed, we have :

$$A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$A^3 = 0$$

Therefore, we have :

$$e^{tA} = I + tA + \frac{t^2}{2}A^2 = \begin{pmatrix} 1 & t & t + t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$$

3 (28 pts.)

We are told that A is 2×2, symmetric, and Markov and one of the real eigenvalues is y with $-1 < y < 1$.

a) (7 pts.) What is this matrix A in terms of y?

We have a symmetric matrix, hence $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Also, it is Markov, so we want $a+b=1$ and $b+c=1$, with all entries non-negative. So $a=c$ and we have $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$.

Now, we want the eigenvalues of this matrix to be y and 1 (recall that Markow matrices ALWAYS have 1 as an eigenvalue, with the all-ones vector as the corresponding eigenvector). But we know the eigenvalues of A satisfy $det(A-\lambda I)=0$, or $(a-\lambda)^2-b^2=0$ or $a-\lambda=\pm b$. So $\lambda_1 = a+b = 1$ and $\lambda_2 = a-b = y$ (since $b \geq 0$). Using $a+b=1$ into $a-b=y$ we get $2a-1=y$ or $a=(y+1)/2$, and then $b=1-a=(1-y)/2$. So we have found our symmetric Markov matrix with eigenvalues 1 and y: $A = \begin{pmatrix} (1+y)/2 & (1-y)/2 \\ (1-y)/2 & (1+y)/2 \end{pmatrix}$.

b) (7 pts.) Compute the eigenvectors of A.

An easy way to find the eigenvector corresponding to the eigenvalue 1 is to recall we have a symmetric Markov matrix, so columns add to 1 but rows too, hence the constant vector will be an eigenvector. So for $\lambda_1 = 1$ we have $v_1 = (1/2\ 1/2)^T$. And for $\lambda_2 = y$, we find a vector in the nullspace of $A - yI = \begin{pmatrix} (1-y)/2 & (1-y)/2 \\ (1-y)/2 & (1-y)/2 \end{pmatrix}$. This is easy, we find $v_2 = (1/2\ \ -1/2)^T$.

c) (7 pts.) What is A^{2012} in simplest form?

We have now diagonalized A: $A = S\Lambda S^{-1}$, where columns of S are the eigenvectors and Λ is a diagonal matrix with 1 and y. So we have

$$A^{2012} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & y \end{pmatrix}^{2012} \frac{1}{-1/2} \begin{pmatrix} -1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

so that

$$A^{2012} = \begin{pmatrix} (1+y^{2012})/2 & (1-y^{2012})/2 \\ (1-y^{2012})/2 & (1+y^{2012})/2 \end{pmatrix}$$

d) (7 pts.) What is $\lim_{n\to\infty} A^n$ in simplest form? (Explain Briefly.)

From the above, and the fact that $-1 < y < 1$, we can see clearly that

$$\lim_{n\to\infty} A^n = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

Another way to reason: we know the steady-state is the eigenvector of the dominating eigenvalue, in this case $\lambda_1 = 1$ and so $v_1 = (1/2\ 1/2)^T$. But we are asking for the matrix which will give us this steady-state, no matter what probability vector we start with. And so its column space has to be along the line of v_1, and no bigger. But there is only one vector proportional to v_1 which could also be a column of a Markov matrix, i.e. whose entries sum to 1. So both columns of the answer have to be v_1. (You could also use the fact that the answer should be symmetric too.)

4 (26 pts.)

a) (5 pts.) P is a three by three permutation matrix. List all the possible values of a singular value. (Explain briefly.)

A permutation matrix satisfies $P^T P = I$ which has all ones as eigenvalues, so all the singular values of P are $\sqrt{1} = 1$.

b) (9 pts.) P is a three by three permutation matrix. List all the possible values of an eigenvalue. (Explain briefly.)

We will do part (c) first.

c) (12 pts.) There are six 3×3 permutation matrices. Which are similar to each other? (Explain briefly.)

Let's list the six matrices. There is the identity matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

There are the three transposition matrices:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

There are the two three-cycles:

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

If two matrices have different traces, then they must have different eigenvalues and so are not similar. The trace of the identity is 3, the trace of the transpositions is 1, and the trace of the three cycles is 0.

We first show that all of the transpositions are similar to each other. Every permutation matrix P satisfies

$$P\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

so they all have eigenvalue $\lambda_1 = 1$. Note that each of the transposition matrices has a fixed point and so has a standard basis vector as an eigenvector with eigenvalue $\lambda_2 = 1$. For example,

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Since they all have trace 1, their final eigenvalue λ_3 must be -1 so that $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Thus we have shown that the transposition matrices all have the eigenvalue 1 repeated twice with two linearly independent eigenvectors as well as the eigenvalue -1. Therefore, they are similar as each of their Jordan canonical forms must be

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Finally, we show the two three-cycles are similar to each other. As before, they have eigenvalue $\lambda_1 = 1$ corresponding to the all ones vector. Their other two eigenvalues must satisfy $\lambda_1 + \lambda_2 + \lambda_3 = 0$. Then $\lambda_2 + \lambda_3 = -1$. However, we must have that $|\lambda_2| = |\lambda_3| = 1$ since the permutation matrices are orthonormal. Note that if λ_2, λ_3 were real then they must each be 1 or -1 and it is impossible to have $\lambda_2 + \lambda_3 = -1$. Therefore, they are complex and must satisfy $\lambda_2 = \overline{\lambda_3}$. Then their real parts are the same and must add to -1, so they each have real part $-1/2$. Using that $|\lambda_2| = |\lambda_3| = 1$, we get that one of λ_2, λ_3 must be $1/2 + i\sqrt{3}/2$ and the other must be $1/2 - i\sqrt{3}/2$. Therefore, the three cycles both have the same eigenvalues, namely the three different cubed roots of 1 in the complex plane, and so are similar.

Returning to part (b) of the problem, we have shown that the possible eigenvalues are the square roots of 1 and the cubed roots of 1.