Practice 18.06 Exam 3 questions

List of potential topics:

Material from exams 1 and 2. Eigenvalues and eigenvectors, characteristic polynomials and nullspaces of $A - \lambda I$. Similar matrices and diagonalization. Complex vs. real linear algebra, adjoints vs. transposes. Hermitian ($A^H = A$), anti-Hermitian ($A^H = -A$), and unitary matrices ($A^H = A^{-1}$), and their eigenvalues/eigenvectors. Markov matrices. Linear recurrences $\mathbf{x}_{n+1} = A\mathbf{x}_n$ and powers of matrices. Differential equations $\frac{d\mathbf{x}}{dt} = A\mathbf{x}$ and matrix exponentials. Hermitian operators on functions, eigenfunctions, Fourier series, and equations written in terms of these (exponentials, inverses, etc. of operators) [see online handouts]. Positive-definite and positive-semidefinite matrices. The singular value decomposition (SVD) $A = U\Sigma V^H$ and the pseudoinverse $A^+ = V\Sigma^+U^H$.

Definitely not on exam 3: finite-difference approximations, sparse matrices and iterative methods, non-diagonalizable matrices, generalized eigenvectors, or Jordan forms. Also *not* fast Fourier transforms or the discrete Fourier transform (which were on the original syllabus but were skipped).

Key ideas: for an eigenvector, any complicated matrix or operator acts just like a number λ, and we can do anything we want (inversion, powers, exponentials...) using that number. To act on an arbitrary vector, we expand that vector in the eigenvectors (in the usual case where the eigenvectors form a basis), and then treat each eigenvector individually. Finding eigenvectors and eigenvalues is complicated, though, so we try to infer as much as we can about their properties from the structure of the matrix/operator (Hermitian, Markov, etcetera).

The actual exam will be four or five questions, so this is about two to three (hard) exams worth of potential questions. (Some of these questions were rejected because they were a bit too hard/long for an exam.)

Problem 1

Suppose A is a square matrix with $A^H = u^2 A$, where u is some complex number with $|u| = 1$.

(a) Show that $B = zA$ is Hermitian for some complex number z. What is z?

(b) What can you conclude about the eigenvalues and eigenvectors of A?

Solution:

(a) If $B = zA$, then $B^H = \bar{z}A^H = \bar{z}u^2A = (\bar{z}u^2/z)B$. Then, to make $B^H = B$, we must have $z/\bar{z} = u^2$, which is solvable since $|z/\bar{z}| = 1 = |u^2|$. The magnitude of z is arbitrary, so let us choose $|z| = 1$ in which case $\bar{z} = 1/z$ and thus we have $z^2 = u^2$ and hence $z = u$. That is, $B = uA$ works.

(b) The eigenvectors with distinct eigenvalues are orthogonal, and the eigenvectors form a basis (A is diagonalizable since B is). The eigenvalues of B are real because it is Hermitian, so the eigenvalues of $A = B/u = \bar{u}B$ are real numbers multiplied by $\bar{u}$.

Problem 2

In an ordinary eigenproblem we solve $A\mathbf{x} = \lambda\mathbf{x}$ to find the eigenvectors $\mathbf{x}$ and eigenvalues λ, and if $A = A^H$ (Hermitian) we find that λ is real and eigenvectors with different eigenvalues are orthogonal.

Now, suppose that instead we are looking for solutions of $A\mathbf{x} = \lambda B\mathbf{x}$ where we have matrices A and B on both sides of the equation. Suppose that both A and B are Hermitian, and B is positive-definite.

(a) Show that the "eigenvalues" λ in $A\mathbf{x} = \lambda B\mathbf{x}$ are real. (Hint: take the dot product of both sides with $\mathbf{x}$.)

(b) Show that two solutions $A\mathbf{x}_1 = \lambda_1 B\mathbf{x}_1$ and $A\mathbf{x}_2 = \lambda_2 B\mathbf{x}_2$ with $\lambda_1 \neq \lambda_2$ satisfy $\mathbf{x}_1 \cdot (B\mathbf{x}_2) = 0$. (Hint: take the dot product of both sides of one equation with $\mathbf{x}_1$.)

Solution

(a) $\mathbf{x} \cdot (A\mathbf{x}) = \lambda \mathbf{x} \cdot (B\mathbf{x})$ but it also $= (A\mathbf{x}) \cdot \mathbf{x} = \bar{\lambda}(B\mathbf{x}) \cdot \mathbf{x} = \bar{\lambda}\mathbf{x} \cdot (B\mathbf{x})$ since A and B are Hermitian (we can move them from one side to the other of the dot product). Therefore, $\lambda \mathbf{x} \cdot (B\mathbf{x}) = \bar{\lambda}\mathbf{x} \cdot (B\mathbf{x})$ and hence $\lambda = \bar{\lambda}$ is real [$\mathbf{x} \cdot (B\mathbf{x}) > 0$ since B is positive-definite].

(b) $\mathbf{x}_1 \cdot (A\mathbf{x}_2) = \lambda_2 \mathbf{x}_1 \cdot (B\mathbf{x}_2) = (A\mathbf{x}_1) \cdot \mathbf{x}_2 = \lambda_2(B\mathbf{x}_1) \cdot \mathbf{x}_2 = \lambda_2 \mathbf{x}_1 \cdot (B\mathbf{x}_2)$, hence $(\lambda_2 - \lambda_1)\mathbf{x}_1 \cdot (B\mathbf{x}_2) = 0$, hence $\mathbf{x}_1 \cdot (B\mathbf{x}_2) = 0$ since $\lambda_1 \neq \lambda_2$.

Problem 3

True or false: any Markov matrix A is also positive-semidefinite. Explain why if true, or give a counter-example if false.

Solution*:* False. All the eigenvalues of a Markov matrix have $|\lambda| \leq 1$ but complex λ and $\lambda < 0$ are also possible. For example, $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is a real-symmetric Markov matrix with eigenvalues $\lambda = \pm 1$. The entries of a Markov matrix are all non-negative, but that should not be confused with positive-semidefinite.

Problem 4

True or false: Explain why if true, or give a counter-example if false.

(a) Any diagonalizable matrix with real eigenvalues is Hermitian.
(b) The product of two Hermitian matrices is Hermitian.
(c) The product of two unitary matrices is unitary.
(d) The sum of two Hermitian matrices is Hermitian.
(e) The sum of two unitary matrices is unitary.

Solutions:

(a) False. Say $A = S\Lambda S^{-1}$ with Λ real and distinct $\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_n$; A is not Hermitian if S is any invertible matrix with non-orthogonal columns (non-orthogonal eigenvectors).

(b) False, since $(AB)^H = B^H A^H = BA \neq AB$ unless A and B commute. For example, $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ give $AB = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$, which is not Hermitian.

(c) True: $(AB)^H = B^H A^H = B^{-1}A^{-1} = (AB)^{-1}$ if A and B are unitary.

(d) True: $(A+B)^H = A^H + B^H = A + B$ if A and B are Hermitian.

(e) False, since $(A+B)^{-1} \neq A^{-1} + B^{-1}$ in general. For example, I is unitary but $I + I = 2I$ is not.

Problem 5

Cal Q. Luss, a Harvard student, doesn't like the definition of Markov matrices. He suggests instead that we use "Markoffish" matrices: real matrices A whose columns sum to 1 like for Markov matrices, but negative entries are allowed.

(a) Show that Markoffish matrices still have a $\lambda = 1$ eigenvalue (hint: consider the eigenvalues of A^T).

(b) Show that the product of two Markoffish matrices is a Markoffish matrix.

(c) For Markov matrices, from (b) we concluded that $|\lambda| > 1$ eigenvalues were not allowed. Is that still true for Markoffish matrices? Explain why if true, or give a counter-example if false.

Solution

(a) As in class, the fact that the sum of each column is one is equivalent to the statement that $A^T\mathbf{x} = \mathbf{x}$ for $\mathbf{x} = (1, 1, 1, \ldots)^T$. Hence, $\lambda = 1$ is still an eigenvalue, since A and A^T have the same eigenvalues.

(b) We can show this by the same explicit summation argument as in class. Or we can use the fact that A being Markoffish is equivalent to $\mathbf{x}^T A = \mathbf{x}^T$ for the $\mathbf{x}$ from (a). Thus, if we have two Markoffish matrices A and B then $\mathbf{x}^T(AB) = (\mathbf{x}^T A)B = \mathbf{x}^T B = \mathbf{x}^T$ and hence AB is Markoffish.

(c) No, this is no longer true; before, the fact that A^n was Markov meant that it could not blow up and hence $|\lambda| \leq 1$, but now A^n is Markoffish and can blow up: its entries can be arbitrarily large and negative. For example, take the Markoffish matrix $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, whose eigenvalues satisfy $\lambda^2 - 4\lambda + 3 = 0$ and hence $\lambda = 3$ and $\lambda = 1$, one of which is > 1.

Problem 6

Consider the vector space of real functions $f(x)$ on $x \in [0, 1]$ with $f(0) = f(1) = 1$, and define the dot product $f \cdot g = \int_0^1 x\, f(x)\, g(x)\, dx$.

(a) Is the second-derivative operator d^2/dx^2 still Hermitian under this inner product? Why or why not?

(b) Show that the operator A defined by $Af = \frac{1}{\sqrt{x}} \frac{d^2}{dx^2}[\sqrt{x} f(x)]$ is Hermitian under this inner product.

(c) What can you conclude about the eigenfunctions and eigenvalues of A?

(d) Show that $-A$ is positive definite ($f \cdot (-Af) > 0$ for all $f \neq 0$). What does this tell you about the eigenvalues of A?

(e) What does your answer to (d) tell you about the solution to $\frac{\partial f}{\partial t} = \frac{1}{\sqrt{x}} \frac{\partial^2}{\partial x^2}[\sqrt{x} f(x, t)] = Af$ with some initial condition $f(x, 0) = g(x)$, as $t \to \infty$? (Hint: expand $g(x)$ in the eigenfunctions, as in class, and write a series solution for $f(x, t)$. You can assume that the eigenfunctions form a basis for the space.)

Solution

(a) No. If we integrate by parts twice in $\int xfg''$, as in classwe get $\int (xf)''g \neq \int xf''g$.

(b) Plugging in, we find $f \cdot Ag = \int \sqrt{x} f(\sqrt{x} g)''$. Integrating by parts twice as above, we get $\int (\sqrt{x} f)'' \sqrt{x} g = (Af) \cdot g$ and hence A is Hermitian.

(c) The eigenvalues must be real as usual, and eigenfunctions for distinct eigenvalues must be orthogonal. [For this operator, you can actual solve analytically for the eigenfunctions $\sin(n\pi x)/\sqrt{x}$ and the eigenvalues $-(n\pi)^2$. However, this is *not* necessary to solve the problem.]

(d) Integrating by parts once: $f \cdot (-Af) = \int \sqrt{x} f(\sqrt{x} f)'' = \int |(\sqrt{x} f)'|^2 \geq 0$. It only $= 0$ if $\sqrt{x} f(x)$ is a constant, but to satisfy the boundary conditions $f(0) = f(1) = 0$ we must therefore have $f(x) = 0$. Hence,

$f \cdot (-Af) > 0$ for all $f \neq 0$, and $-A$ is positive-definite. Hence the eigenvalues of $-A$ are positive, and hence the eigenvalues of A are negative.

(e) Call the eigenfunctions $f_n(x)$ and the corresponding eigenvalues $\lambda_n < 0$. Then, we write $g(x) = \sum_n c_n f_n(x)$ for some coefficients $c_n = f_n \cdot g / \|f_n\|$. We have the equation $\frac{\partial f}{\partial t} = Af$, so formally the solution is $f(x,t) = e^{At} f(x,0) = e^{At} g(x)$ as for the diffusion equation in class. The exponential of A is a bit weird, but as usual we know that it is just a number when it acts on an eigenfunction:

$$f(x,t) = e^{At} g(x) = e^{At} \sum_n c_n f_n(x) = \sum_n c_n e^{At} f_n(x) = \sum_n c_n e^{\lambda_n} f_n(x).$$

But since all the λ_n eigenvalues are negative, every term goes exponentially to zero and hence $f(x,t) \to 0$ as $t \to \infty$.

Problem 7

Suppose that you have a system of differential equations $B^{-1} \frac{d\mathbf{x}}{dt} = -B^H \mathbf{x}$ with some initial condition $\mathbf{x}(0)$, for some invertible matrix B.

(a) Find $\mathbf{x}(t)$ for $B = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{x}(0) = \begin{pmatrix} 2 \\ -4 \end{pmatrix}$.

(b) What happens to your $\mathbf{x}(t)$ from (a) as $t \to \infty$?

(c) Argue that your answer from (b) is true in general, for *any* invertible $n \times n$ matrix B and any initial condition $\mathbf{x}(0)$.

Solutions:

(a) We have $\frac{d\mathbf{x}}{dt} = -BB^H \mathbf{x}$, and hence $\mathbf{x}(t) = e^{-BB^H t} \mathbf{x}(0)$. To solve this we must find the eigenvectors of BB^H and expand $\mathbf{x}(0)$ in terms of them. $BB^H = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}$, whose eigenvalues solve $\lambda^2 - 6\lambda + 4 = 0$, and hence $\lambda_\pm = 3 \pm \sqrt{5}$ by the quadratic equation. If we look for eigenvectors $\mathbf{v}$ of the form $\mathbf{v} = \begin{pmatrix} 1 \\ u \end{pmatrix}$, we find $5 + u = \lambda_\pm$ and hence $u = -2 \pm \sqrt{5}$. By inspection, we can then write $\mathbf{x}(0) = \begin{pmatrix} 2 \\ -4 \end{pmatrix} = \begin{pmatrix} 1 \\ -2+\sqrt{5} \end{pmatrix} + \begin{pmatrix} 1 \\ -2-\sqrt{5} \end{pmatrix}$. Hence,

$$\mathbf{x}(t) = e^{-BB^H t} \mathbf{x}(0) = \begin{pmatrix} 1 \\ -2+\sqrt{5} \end{pmatrix} e^{-(3+\sqrt{5})t} + \begin{pmatrix} 1 \\ -2-\sqrt{5} \end{pmatrix} e^{-(3\pm\sqrt{5})t}$$

(b) $3 \pm \sqrt{5} > 0$, so the $\mathbf{x}(t)$ function decays exponentially towards 0.

(c) This is true in general since $BB^H = (B^H)^H (B^H)$ is always positive-definite as shown in class; B has full column rank (it is invertible) so it is not merely semidefinite. This means that the eigenvalues of BB^H are all real and positive, and hence the eigenvalues of $-BB^H$ are negative, and hence $e^{-BB^H t}$ is exponentially decaying for all the eigenvectors.

Problem 8

Suppose A and B are Hermitian $n \times n$ matrices. What can you say about the eigenvalues and eigenvectors of $C = i(AB - BA)$? (Hint: what is C^H?)

Solution: $C^H = -i(B^H A^H - A^H B^H) = -i(BA - AB) = i(AB - BA) = C$, so C is Hermitian and has real eigenvalues, orthogonal eigenvectors for distinct eigenvalues, and is diagonalizable.

Problem 9

Suppose that A is an 6×4 matrix with full column rank (rank $= 4$), and has the SVD

$$A = U\Sigma V^H = U \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \sigma_3 & \\ & & & \sigma_4 \\ & & & 0 \\ & & & 0 \end{pmatrix} V^H.$$

Recall the definition of the pseudo-inverse: $A^+ = V\Sigma^+ U^H$, where Σ^+ is the transpose of Σ with the non-zero entries (σ) inverted ($1/\sigma$). Show (by explicit multiplication etc.) that the pseudoinverse A^+ equals $(A^H A)^{-1} A^H$. [That is, the $A^+\mathbf{b}$ is equivalent to the least-squares solution to $A\mathbf{x} = \mathbf{b}$, as we discussed in class.]

Solution:

Given this A, we find

$$A^H A = V\Sigma^T\Sigma V^H = V \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \sigma_3^2 & \\ & & & \sigma_4^2 \end{pmatrix} V^H,$$

and hence the inverse is given by inverting the eigenvalues σ^2:

$$(A^H A)^{-1} = V \begin{pmatrix} \sigma_1^{-2} & & & \\ & \sigma_2^{-2} & & \\ & & \sigma_3^{-2} & \\ & & & \sigma_4^{-2} \end{pmatrix} V^H V,$$

and hence

$$
\begin{aligned}
(A^H A)^{-1}A^H &= V\begin{pmatrix} \sigma_1^{-2} &&& \\ & \sigma_2^{-2} && \\ && \sigma_3^{-2} & \\ &&& \sigma_4^{-2}\end{pmatrix} V^H V \begin{pmatrix} \sigma_1 &&&&& \\ & \sigma_2 &&&& \\ && \sigma_3 &&& \\ &&& \sigma_4 & 0 & 0\end{pmatrix} U^H \\
&= V\begin{pmatrix} \sigma_1^{-2} &&& \\ & \sigma_2^{-2} && \\ && \sigma_3^{-2} & \\ &&& \sigma_4^{-2}\end{pmatrix} \begin{pmatrix} \sigma_1 &&&&& \\ & \sigma_2 &&&& \\ && \sigma_3 &&& \\ &&& \sigma_4 & 0 & 0\end{pmatrix} U^H \\
&= V\begin{pmatrix} \sigma_1^{-1} &&&&& \\ & \sigma_2^{-1} &&&& \\ && \sigma_3^{-1} &&& \\ &&& \sigma_4^{-1} & 0 & 0\end{pmatrix} U^H = V\Sigma^+ U^H \\
&= A^+.
\end{aligned}
$$

Problem 10

Suppose that an $m \times n$ matrix has the SVD $A = U\Sigma V^H$. Recall the definition of the pseudo-inverse: $A^+ = V\Sigma^+ U^H$, where Σ^+ is the transpose of Σ with the non-zero entries (σ) inverted ($1/\sigma$). Show that $(A^H)^+ = (A^+)^H$.

Solution:

The adjoint is $A^H = V\Sigma^T U^H$, which immediately tells us the SVD of A^H. Hence $(A^H)^+ = U(\Sigma^T)^+ V^H$. In comparison, $(A^+)^H = U(\Sigma^+)^H V^H$. Clearly, the two are equal if $(\Sigma^T)^+ = (\Sigma^+)^T$, which is obviously true since inverting the singular values ($\Sigma \to \Sigma^+$ means $\sigma \to 1/\sigma$) can clearly be interchanged with transposition (swapping rows and columns) without changing the result.

Problem 11

Consider the vector space of twice differentiable real functions $f(x)$ on $x \in [0,1]$ with $f(0) = f(1) = 1$, and define the dot product $f \cdot g = \int_0^1 f(x)\, g(x)\, dx$ as in class. Now, define a sequence of functions $f_k(x)$ $[k = 0, 1, 2, \ldots]$ by the recurrence relation $Af_{k+1}(x) = f_k(x)$ with $A = \frac{d^2}{dx^2}$.

(a) Suppose that the initial function $f_0(x)$ in the recurrence has the Fourier sine series:

$$
f_0(x) = \frac{4}{\pi^2}\sin(\pi x) - \frac{4}{(3\pi)^2}\sin(3\pi x) + \frac{4}{(5\pi)^2}\sin(5\pi x) - \cdots = \frac{4}{\pi^2}\sum_{\ell=0}^{\infty}\frac{(-1)^\ell}{(2\ell+1)^2}\sin[(2\ell+1)\pi x].
$$

Give an explicit Fourier sine series for $f_k(x)$. [Recall from class: for this vector space and dot product, $\sin(n\pi x)$ is an eigenfunction of the Hermitian operator $\frac{d^2}{dx^2}$, with eigenvalue $-(n\pi)^2$.]

(b) Suppose we replace $\frac{d^2}{dx^2}$ with $c^2\frac{d^2}{dx^2}$ for some real number c. For what values of c (if any) does $\|f_k(x)\|^2$ diverge as $k \to \infty$? How does your answer depend on the initial function $f_0(x)$, if at all?

Solutions:

(a) The solution to this recurrence, like for matrix recurrences, is $f_k(x) = A^{-k} f_0(x)$. To compute A^{-k}, we just multiply the eigenfunctions by λ^{-k} as usual. Hence

$$f_k(x) = (-1)^k \left[\frac{4}{\pi^{2+2k}} \sin(\pi x) - \frac{4}{(3\pi)^{2+2k}} \sin(3\pi x) + \frac{4}{(5\pi)^{2+2k}} \sin(5\pi x) - \cdots \right] = \frac{4(-1)^k}{\pi^{2+2k}} \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell+1)^{2+2k}} \sin[(2\ell+1)\pi x],$$

using the fact that the eigenvalues are $-(n\pi)^2$.

(b) For f_k to diverge, A^{-1} should have an eigenvalue $|\lambda| > 1$, and hence A should have eigenvalues $|\lambda| < 1$. If $A = c^2 \frac{d^2}{dx^2}$, the eigenvalues of A are $-(cn\pi)^2$. This has magnitude less than 1 for $n = 1$ when $|c| < 1/\pi$, for $n = 2$ when $|c| < 1/2\pi$, and so forth. So, f_k diverges when $|c| < 1/\pi$. However, this does depend on our initial function f_0: if the $n = 1$ term is not present, then $|c|$ must be smaller; in general, if the first non-zero eigenfunction in the sine series for f_0 is $n = m$, then we must have $|c| < 1/m\pi$ to make f_k diverge.