Math, Problem Set #3, Spectral Theory

Francesco Furno

Joint Work with Ildebrando Magnani

OSM Lab

Due Monday, July 10th

Exercise 2

From the previous homework], we know that we can represent the derivative operator as matrix

$$D = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, the roots (eigenvalues) of the characteristic polynomial of D are given by

$$\det(D - \lambda I) = \det \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{bmatrix} = 0$$

that is $-\lambda^3 = 0$. Hence, the unique eigenvalue of D is $\lambda = 0$, with algebraic multiplicity three and geometric multiplicity given by $Dim(Null(D))$. Since the matrix representation of the derivative operator maps vectors of the form $[a \quad b \quad c]$ (where $a, b, c \in R$) to vectors of the form $[b \quad 2c \quad 0]$, $p = [a \quad b \quad c] \in Null(D) \iff b = c = 0$. Thus, $Null(D) = Span(e_1)$ and $Dim(Null(D)) = 1$.

Exercise 4

(i)

Let A be an Hermitian matrix, i.e.,

$$\begin{bmatrix} a & x + iy \\ x - iy & b \end{bmatrix}.$$

By (4.3) we obtain that the characteristic polynomial of A is given by

$$p(\lambda) = \lambda^2 - (a + b)\lambda + ab - x^2 - y^2.$$

Therefore,

$$\Delta(p(\lambda)) = (a+b)^2 - 4(ab - x^2 - y^2) = a^2 + b^2 + 2ab - 4ab + 4x^2 + 4y^2 = (a-b)^2 + 4x^2 + 4y^2 \geq 0,$$

i.e., the discriminant is positive and the roots of the characteristic polynomial are all real.

(ii)

Let A be an Hermitian matrix, i.e.,

$$\begin{bmatrix} ix & a+ib \\ -a+ib & iy \end{bmatrix}.$$

By (4.3) we obtain that the characteristic polynomial of A is given by

$$p(\lambda) = \lambda^2 - (ix + iy)\lambda - xy + a^2 + b^2$$

Therefore,

$$\Delta(p(\lambda)) = -x^2 - y^2 + 2xy - 4a^2 - 4b^2 = -(x-y)^2 - 4(a^2+b^2) < 0$$

i.e., the discriminant is negative and the roots of the characteristic polynomial are all imaginary.

Exercise 6

Let A be an upper triangular matrix, i.e.,

$$A = \begin{bmatrix} \lambda_1 & * & * & ... & * \\ 0 & \lambda_2 & * & ... & * \\ 0 & 0 & \lambda_3 & ... & * \\ 0 & 0 & 0 & ... & * \\ 0 & 0 & ... & 0 & \lambda_n \end{bmatrix}$$

Then, λ is an eigenvalue of $A \iff (A - \lambda I)$ is singular. Thus,

$$A - \lambda I = \begin{bmatrix} \lambda_1 - \lambda & * & * & ... & * \\ 0 & \lambda_2 - \lambda & * & ... & * \\ 0 & 0 & \lambda_3 - \lambda & ... & * \\ 0 & 0 & 0 & ... & * \\ 0 & 0 & ... & 0 & \lambda_n - \lambda \end{bmatrix}$$

This matrix is again upper triangular. An upper triangular matrices is singular if and only if the product of its diagonal entries equals zero. Hence, λ is an eigenvalue of A if and only if it satisfies

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda) = 0.$$

Each root of this equation is an eigenvalue of A. This implies that $\lambda = \lambda_i$ for all $i = 1, 2, 3, ..., n$.

Exercise 8

(i)

We know that every orthonormal set is also linearly independent. We showed in Exercise (3.8) that the set S is orthonormal, and hence linearly independent. Since $V = Span(S)$, we have that S is a basis for V.

(ii)

The matrix representation D of the derivative operator is completely determined by the map it operates on the basis S. In particular, it has to map the vector of coefficients $[a, \quad b, \quad c, \quad d]$ into the vector $[-b, \quad a, \quad -2d, \quad 2c]$. Thus,

$$D = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

(iii)

Clearly, we can pick $Span\{sin(x), cos(x)\}$ and $Span\{sin(2x), cos(2x)\}$.

Exercise 13

We need to find an eigenbasis of A. Thus,

$$\det(A - \lambda I) = (0.8 - \lambda)(0.6 - \lambda) - 0.08 = \lambda^2 - 1.4\lambda + 0.4 = 0$$

The eigenvalues of A are $\{1, 0.4\}$, and with some calculations we also find that the corresponding eigenvectors are $[1, \quad 2]^T$ and $[1, -1]^T$. Thus, let $P = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}$. Then,

$$P^{-1}AP = D = \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix}$$

where $P^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$.

Exercise15

Since A is semisimple it is also diagonalizable. Hence $A = PDP^{-1}$, where $Disdiagonal$. Then,

$$f(A) = P^{-1}(a_0 + a_1 D + ... + a_n D)P = P^{-1}D'P$$

where D' is still diagonal with entries $f(\lambda_i)$, i.e., $f(\lambda_i)$ are the eigenvalues of $f(A)$.

Exercise 16

(i), (ii)

Using proposition (4.3.10), we know that

$$\lim_{n\to\infty} A^n = \lim_{n\to\infty} PD^nP^{-1} = \lim_{n\to\infty} \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1^n & 0 \\ 0 & 0.4^n \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{2}{3} \end{bmatrix}$$

The answer is not dependent on the type of norm that we use.

(iii)

Using Theorem (4.3.12) we know that the eigenvalues of $f(A) = 3I + 5A + A^3$ are $f(1) = 9$ and $f(0.4) = 1.2 + 2 + 0.064 = 3.264$.

Exercise 18

We know that λ is an eigenvalue of $A \iff$ there exist a non-zero column vector x such that $Ax = \lambda x$, or (since A and A^T have the same characteristic polynomial) $A^T x = \lambda x$. Taking the transpose of both sides in the last equation, we get $x^T A = \lambda x^T$. Thus, λ is an eigenvalue of $A \iff$ there exist a non-zero row vector x such that $x^T A = \lambda x^T$.

Exercise 20

By assumption, we have that A is Hermitian, i.e., $A^H = A$ and that A is orthonormally similar to B, i.e., there exists an orthonormal matrix Q such that $A = Q^H BQ \iff B = QAQ^H$. Then,

$$B^H = [(QA)Q^H]^H = Q(QA)^H = QA^H Q^H = QAQ^H = B$$

Therefore, B is also Hermitian.

Exercise 24

If A is an Hermitian matrix, then by the First Spectral Theorem (and Corollary) we know that A is orthonormally diagonalizable and has all real eigenvalues, i.e., $D = Q^H AQ$, where D is a real diagonal matrix of eigenvalues of A. Let the vector $x = Qc$, where Q is an eigenbasis for C^n. Then,

$$\rho(x) = \frac{\langle x, Ax \rangle}{\|x\|^2} = \frac{c^H Q^H AQc}{c^H c} = \frac{c^H Dc}{c^H c} = \frac{\sum \lambda_i c_i^2}{\sum c_i^2}$$

which is real, since all the eigenvalues of A Hermitian matrix are real. Given that all eigenvalues of a Skew Hermitian matrix are all imaginary, we can go over the same proof for A Skew Hermitian to show that in this case the Rayleigh Quotient takes only imaginary values.

Exercise 25

(i)

Let $X = [\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}]$. Then, since the columns of X are orthonormal eigenvectors of A, we have that

$$\sum_{i=1}^{n} \mathbf{x_i x_i^H} = I$$

(ii)

Left-multiplying the previous equation, we get

$$AI = A = \sum_{i=1}^{n} A\mathbf{x_i}\mathbf{x_i^H} = \sum_{i=1}^{n} \lambda_i \mathbf{x_i}\mathbf{x_i^H}$$

Exercise 27

We have that $A = [a_{ij}]$ is positive definite if it is Hermitian and $\langle x, Ax\rangle > 0$ and real valued. Now let's assume by contradiction that $a_{jj} \leq 0$ and imaginary for some $0 \leq j \leq n$. Then, if e_j is the j^{th} standard basis element, we obtain

$$\langle e_j, Ae_j\rangle = e_j^H[a_{1j}, ..., a_{nj}]^H = a_{jj} \leq 0$$

This is clearly a contradiction. Thus, all the diagonal elements of A must be positive and real.

Exercise 28

By the previous Exercise, we have that the diagonal entries of a positive semidefinite matrix have to be all non-negative and real. Thus, for $A = [a_{ij}]$ and $B = [b_{ij}]$ positive semidefinite, by Cauchy-Schwarts we have

$$0 \leq |Tr(AB)| = Tr(AB) \leq \|diag(A)\|_2 \|diag(B)\|_2 \; =$$

$$= \sqrt{\sum_{i=1}^{n} a_{ii}^2 \sum_{i=1}^{n} b_{ii}^2} \leq \sqrt{(\sum_{i=1}^{n} a_{ii})^2 (\sum_{i=1}^{n} b_{ii})^2} = Tr(A)Tr(B)$$

Exercise 31

(i)

$$\|A\|_2 = \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{x\neq 0} \frac{\|U\Sigma V^H x\|_2}{\|x\|_2} = \sup_{x\neq 0} \frac{\|\Sigma V^H x\|_2}{\|x\|_2} =$$

(letting $y = V^H x$)

$$= \sup_{y\neq 0} \frac{\|\Sigma y\|_2}{\|Vy\|_2} = \frac{(\sum_{i=1}^{r} \sigma_i^2 y_i^2)^{1/2}}{(\sum_{i=1}^{r} y_i^2)^{1/2}} \leq \sigma_1$$

Thus, for $y = e_1$,$\|\Sigma y\|_2 = \sigma_1$ and the supremum is attained.

(ii)

Same proof, just with $\|\Sigma^{-1} y\|_2 = (\sum_{i=1}^{r} (\frac{1}{\sigma_i^2}) y_i^2)^{1/2}$.