Math, Problem Set #3, Spectral Theory

Instructor: John Van de Berghe

Due Monday, July 10 at 8:00am

Homework: 2, 4, 6, 8, 13, 15, 16, 18, 20, 24, 25, 27, 28, 31, 32, 33, 36, 38 at the end of Chapter 4 of Humpherys et al. (2017)

Solutions

4.2 The matrix $D = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. Solving $det(D - \lambda I) = 0$, we get $\lambda^3 = 0$ and see that the only eigenvalue is 0. Solving $(D - \lambda I)x = 0$, we see that the eigenspace consists of all vectors of the form $\begin{bmatrix} k \\ 0 \\ 0 \end{bmatrix}$, where $k \in \mathbb{R}$. The eigenvalue 0 has algebraic multiplicity 3, and geometric multiplicity is 1 because the eigenspace is spanned by 1 nonzero eigenvector.

4.4 The charcteristic polinomiyl of any 2×2 matrix is given by:
$p(\lambda) = \lambda^2 - tr(A)\lambda + det(A)$

i) for any Hermetian 2×2 matrix $A = \begin{bmatrix} a & \bar{b} \\ b & a \end{bmatrix}$ the eigenvalues can be found by solving the characteristic equation

$$
\begin{aligned}
0 &= \lambda^2 - 2(a+c)\lambda + ac - b\bar{b} \\
\lambda_{1,2} &= \frac{a+c}{2} \pm \sqrt{\left(\frac{a+c}{2}\right)^2 - ac + b\bar{b}} \\
\Rightarrow \lambda_{1,2} \in \mathbb{R} \quad \text{if} \quad 0 &< \left(\frac{a+c}{2}\right)^2 - ac + b\bar{b} \\
0 &< \left(\frac{a+c}{2}\right)^2 - ac + b\bar{b} \\
0 &< \frac{1}{4}(a^2 + c^2 + 2ac) - ac + b\bar{b} \\
0 &< \frac{a^2 + c^2 - 2ac}{4} + b\bar{b} \\
0 &< \frac{1}{4}(a-c)^2 + b\bar{b}
\end{aligned}
$$

ii)For all eigenvalues of any skew-Hermetian 2×2 matrix $A = \begin{bmatrix} 0 & b-c \\ c-b & 0 \end{bmatrix}$ being imaginary the argument is similar to the argument in part i). From the characteristic polynomial it follows that

$$\lambda_{1,2} = \pm\sqrt{(c-b)(b-c)}$$
$$\Rightarrow \lambda_{1,2} \in \mathbb{I} \quad \text{if} \quad 0 < (b-c)(c-b)$$

The second inequality holds for both possible cases $b < c$ and $c < b$.

4.6 For any triangular matrix A it is true that $det(A) = trace(A)$ so the pcharacteristic polynomial collapses to $0 = (a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda)...(a_{nn} - \lambda) \Leftrightarrow a_{11} = \lambda_1,\ a_{22} = \lambda_2,\ a_{33} - \lambda_3, \quad \ldots \quad a_{nn} = \lambda_n$

4.8 i) let S be the span of V, then for S to be a basis of V, the elementsin S must be linearly independent, that is, there exists no such a, b, c and d that: $a\sin(x) + b\cos(x) + c\sin(2x) + d\cos(2x) = 0 \quad \forall x \in \mathbb{R}$

ii)denote the derivative operator as D such that

$$\begin{bmatrix} \cos x \\ -\sin x \\ 2\cos(2x) \\ -2\sin(2x) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \end{bmatrix}}_{=D} \begin{bmatrix} \sin x \\ \cos x \\ \sin(2x) \\ \cos(2x) \end{bmatrix}$$

iii $W_1 = \{\sin x, \cos x\}$ and $W_1 = \{\sin(2x), \cos(2x)\}$ are two invariant subspaces of S.

4.13 First find the eigenvalues and the corresponding eigenvectors:

$$\begin{aligned} 0 &= |A - \lambda I| \\ 0 &= (0.8 - \lambda)(0.6 - \lambda) - 0.2 \times 0.4 \\ 0 &= \lambda^2 - 1.4\lambda + 0.04 \\ \lambda_{1,2} &= 0.7 \pm \sqrt{0.049 - 0.04} = 0.7 \pm 0.3 \\ \lambda_1 &= 1,\ \lambda_2 = 0.4 \end{aligned}$$

To find the eigenvectors I solve

$$0 = (A - \lambda_1 I)v_i$$
$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.4 \\ 0.2 & -0.4 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} \Rightarrow \quad v_{11} = 2v_{21}$$

and for the second one:

$$0 = (A - \lambda_2 I)v_i$$
$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.4 \\ 0.2 & 0.2 \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} \Rightarrow \quad v_{12} = -v_{22}$$

so we have $\Sigma_1 = \{(2,\ 1)^T$ and $\Sigma_1 = \{(1,\ -1)^T$ which gives the transition matrix $P = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}$ and its inverse $P^{-1} = \frac{1}{3}\begin{bmatrix} -1 & -1 \\ -1 & 2 \end{bmatrix}$.

4.15 Let A be a semi-similar matrix so that there exists a diagonalization of A such that $A = PDP^{-1}$ and D is diagonal. Suppose $(\lambda_i)_{i=1}^n$ are the eigenvalues of A. Then we can find the eigenvalues of the polynomial $f(A)$ by

$$\begin{aligned} f(A) &= a_0 I + a_1 A + a_2 A^2 + \ ... \ + a_n A^n \\ &= a_0 P^{-1} P + a_1 P^{-1} D P + a_2 P^{-1} D^2 P + \ ... \ + a_n P^{-1} D^n P \\ &= P^{-1} f(D) P \\ &= P \begin{bmatrix} a_0 + a_1\lambda_1 + ... + a_n\lambda_1^n & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & a_0 + a_1\lambda_n + ... + a_n\lambda_n^n \end{bmatrix} P^{-1} \end{aligned}$$

Since $f(D)$ is diagonal and the eigenvalues of any diagonal matrix correspond to its diagonal elements we have the eigenvalues of $f(A)$ are given by $f(\lambda_i) = a_0 + a_1\lambda_i + ... + a_n\lambda_i$.

4.16 i) Compute $\lim_{n\to\infty} A^n$ with A and its transition matrix P from exercise 4.13.

$$\begin{aligned} B &= \lim_{n\to\infty} A^n \\ &= \lim_{n\to\infty} P^{-1} D^n P \quad \text{with} \quad D^n = \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} = \begin{bmatrix} 1^n & 0 \\ 0 & 0.4^n \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ B &= \frac{1}{3} \begin{bmatrix} -1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \end{aligned}$$

then $\|A^k - B\|_1 < \varepsilon$

ii) does the same hold for with respect to the frobenius norm?

$$\sqrt{tr(A^k B)} = \sqrt{\left(\frac{2}{3} - \frac{1}{3}\right)^2} = \frac{1}{3}$$

Apparently, this does not hold for the frobenius norm.

What about the $\infty - norm$? $\|A^- kB\|_\infty = 0 < \varepsilon$ because the $\infty - norm$ is basically the supremum of the row sum of the matrix $A - kB$.

iii) Find all eigenvalues of the matrix $C = 3I + 5A + A^3$ using the result of exercise 4.15.

$$\begin{aligned} (\lambda_1)_C =& 3 + 5(\lambda_1)_A + (\lambda_1)_C^3 = 3 + 5 + 1 = 9 \\ (\lambda_2)_C =& 3 + 5(\lambda_2)_A + (\lambda_2)_A^3 = 3 + 2 + 0.4^3 = 5.064 \end{aligned}$$

4.18 Let x be an eigenvector of A withh corresponding eigenvalue λ then because of $det(A) = det(A^T)$ we can write

$$\begin{aligned} A^T x =& \lambda x \\ (A^T x)^T =& (\lambda x)^T \\ x^T A =& \lambda x^T \end{aligned}$$

4.20 Show that if A is Hermetian that B is Hermetian too:
$B^H = (U^H AU)^H = (U^H AU) = B$

4.24 i) We aim to show that $\rho = \rho^H$ because this is true if and only if ρ is a real number. We know $A = A^H$. Then, $\rho = \frac{\langle x, Ax \rangle}{\|x\|^2} = \frac{\langle x, A^H x \rangle}{\|x\|^2} = (\frac{x^H A^H x}{x^H x})^H = (\frac{\langle x, Ax \rangle}{\|x\|^2})^H = \rho^H$.

ii)We aim to show that $-\rho = \rho^H$ because this is true if and only if ρ is an imaginary number. We know $A = A^H$. Then, $-\rho = -\frac{\langle x, Ax \rangle}{\|x\|^2} = \frac{\langle x, -A^H x \rangle}{\|x\|^2} = -(\frac{x^H A^H x}{x^H x})^H = -(\frac{\langle x, Ax \rangle}{\|x\|^2})^H = -\rho^H$.

4.25 i) Suppose A is a normal matrix with eigenvalues $(\lambda_1, ..., \lambda_n)$ and orthonormal eigenvectors $(x_1, ..., x_n)$ which forms an eigenbasis for the space. We know that I is an identity matrix if and only if $Ix = x$ for any vector $x \in \mathbb{F}$. Since we have an eigenbasis, x can be rewritten as $x = \alpha_n x_1 + ... + \alpha_n x_n$. Then, $(x_1 x_1^H + ... + x_n x_n^H)x = (x_1 x_1^H + ... + x_n x_n^H)\alpha_n x_1 + ... + \alpha_n x_n = (\alpha_1 x_1 x_1^H x_1 + ... + \alpha_n x_n x_n^H x_n) = \alpha_n x_1 + ... + \alpha_n x_n$ because the vectors are orthonormal so $x_j^H x_i = 0 \forall j \neq i$ and $x_j^H x_i = 1 \forall j = i$.

ii) Because $(x_1 x_1^H + ... + x_n x_n^H) = I$, we can write $A = A(x_1 x_1^H + ... + x_n x_n^H) = (Ax_1 x_1^H + ... + Ax_n x_n^H) = (\lambda_1 x_1 x_1^H + ... + \lambda_n x_n x_n^H)$, as desired.

4.27 Prove that the diagonal elements of any positive definite matrix A are positive and real:

From the definition of pos. def. of A we know that the inner product of A and any vector x is positive and real, that is $\langle x, Ax \rangle > 0 \, \forall x$ and further that $0 < \langle x, Ax \rangle = \overline{\langle x, Ax \rangle}$. Let e_i be the vector of zeroes with a one at the ith position, then $\langle e_i, Ae_i \rangle = e_i^T Ae_i = a_{ii}$ where a_{ii} is the ith diagonal element of A.

4.28 Show $0 \leq tr(AB) \leq tr(A)tr(B)$ for A, B being positive semidefinite:

From the definition of positive semidefinite matrices we know that $0 \leq \langle x, Ax \rangle$. Following the argument in 4.27 we have that the diagonal elements of A and B are non-negative, i.e. $a_{ii}, b_{jj} \geq 0 \, \forall i$. Then it follows that $tr(AB) =$

$\sum_{i=j}^{n} a_{ii} b_{jj} \geq 0$. For the second inequality consider

$$\begin{aligned}
tr(A)tr(B) &= \sum_{i=1}^{n} a_{ii} \sum_{j=1}^{n} b_{jj} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ii} b_{jj} \\
&= \sum_{i=1}^{n} \sum_{j \neq i}^{n} a_{ii} b_{jj} + \sum_{i=j}^{n} a_{ii} b_{jj} \\
&= \sum_{i=1}^{n} \sum_{j \neq i}^{n} a_{ii} b_{jj} + tr(AB) \geq tr(AB)
\end{aligned}$$

4.31 Assume $A \in M_{m\times}(\mathbb{F})$ is of rank r prove that

i) $\|A\|_2 = \sigma_1$ with σ_1 being the largest singular value of A.

$$\begin{aligned}
\|A\|_2 = sup\frac{\|Ax\|_2}{\|x\|_2} &= sup\frac{\|U\Sigma V^H x\|_2}{\|x\|_2} \quad \text{because U is othonormal} \\
&= sup\frac{\|\Sigma V^H x\|_2}{\|x\|_2} \qquad \text{define}\, y = V^H x \\
&= sup\frac{\|\Sigma y\|_2}{\|Vy\|_2} \qquad \text{because V is orthonormal} \\
&= sup\frac{\|\Sigma y\|_2}{\|y\|_2} = sup\frac{\sqrt{\sigma_1^2 y_1^2 + ... + \sigma_r^2 y_r^2}}{\sqrt{y_1^2 + ... y_r^2}} \\
&= \frac{\sigma_1 \sqrt{y_1^2 + ... y_r^2}}{\sqrt{y_1^2 + ... y_r^2}} = \sigma_1
\end{aligned}$$

ii) if A is invertible than $\|A^{-1}\|_2 = \frac{1}{\sigma_r}$ where σ_r is the smallest singular value of A.

We proceed similar to part i)

$$\begin{aligned}
\|A^{-1}\|_2 = sup\frac{\|(U\Sigma V^H)^{-1} x\|_2}{\|x\|_2} &= sup\frac{\|(V^H)^{-1}\Sigma^{-1}U^{-1}x\|}{\|x\|_2} \quad \text{because V is othonormal} \\
&= sup\frac{\|\Sigma^{-1}U^{-1}x\|_2}{\|x\|_2} \qquad \text{define}\, y = U^{-1}x \\
&= sup\frac{\|\Sigma^{-1} y\|_2}{\|yU\|_2} \qquad \text{because U is orthonormal} \\
&= sup\frac{\|\Sigma^{-1} y\|_2}{\|y\|_2} = sup\frac{\sqrt{\left(\frac{1}{\sigma_1}\right)^2 y_1^2 + ... + \left(\frac{1}{\sigma_r}\right)^2 y_r^2}}{\sqrt{y_1^2 + ... y_r^2}} \\
&= \frac{1}{\sigma_r}\frac{\sqrt{y_1^2 + ... y_r^2}}{\sqrt{y_1^2 + ... y_r^2}} = \frac{1}{\sigma_r}
\end{aligned}$$

iii)

iv) If U and V are orthonormal then $\|UAV\|_2 = \|A\|_2$

$$\begin{aligned}\|UAV\|_2 &= \frac{\|UAVx\|_2}{\|x_2\|} \quad \text{with } y = Vx \\ &= \frac{\|UAy\|_2}{\|V^H y\|} = \frac{\|Ay\|_2}{\|y\|} = \|A\|_2\end{aligned}$$

3.32 Assume U and V are orthonormal and $rank(A) = r$.
i) $\|UAV\|_F = \sqrt{tr((UAV)^H UAV)} = \sqrt{tr(V^H A^H U^H UAV)} = \sqrt{tr(V^H V A^H A)} = \sqrt{tr(A^H A)} = \|A\|_F$
ii)$\|A\|_F = \|U\Sigma V\|_F = \sqrt{tr((U\Sigma V)^H U\Sigma V)} = \sqrt{tr(V^H \Sigma^H U^H U\Sigma V)} = \sqrt{tr(\Sigma^H \Sigma)} = \sqrt{(\sigma_1^2 + \sigma_2^2 + ... + \sigma_r^2)}$

4.33 Show that $\|A\|_2 = sup_{\|x\|=1, \|y\|=1}|y^H Ax|$. From exercise 4.31 we know that $\|A\|_2 = \sigma_1$ where σ_1 is the largest singular value of A.

$$\begin{aligned}sup_{\|x\|_2=1, \|y\|_2=1}|y^H Ax| &= sup_{\|x\|=1, \|y\|=1}|y^H U\Sigma V^H x| \\ &= sup_{\|x\|_2=1, \|y\|_2=1}|aU^H U\Sigma V^H Vb|\end{aligned}$$

(define $y^H = a^H U^H$, $x = Vb$ since V and U ar orthonormal $\|y\| = \|a\|, \|b\| = \|x\|$)

$$= sup_{\|a\|_2=1, \|a\|_2=1}|a\Sigma b|$$

$$= sup_{\|a\|_2=1, \|a\|_2=1}|\sum_{i=1}^{r} a_i \sigma_i b_i|$$

assume $\quad a = b = \begin{pmatrix}1 & 0\cdots & 0\end{pmatrix}^T$ to get the supremum:

$$= sup_{\|a\|_2=1, \|a\|_2=1}|\sum_{i=1}^{r} a_i \sigma_i b_i| = \sigma_1$$

4.36 Example of a 2×2 matrix with non-zero determinante and its eigenvalues being different from its singular values: $A = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$. Then, $A^H A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
Then, the eigenvalues of A are i, but the singular value of A is 1, and the determinant is -1, which is nonzero.

4.38 Proove properties of the Moore-Penrose Pseudoinverse $A^\dagger = V_1 \Sigma_1^{-1} U_1^H$ of the matrix $A = U_1 \Sigma_1 V_1^H$ where U_1 and V_1 are orthonormal:
i) $AA^\dagger A = U_1\Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H = U_1 \Sigma_1 \Sigma_1^{-1} \Sigma_1 V_1^H = U_1 \Sigma_1 V_1^H = A$
ii) $A^\dagger A A^\dagger = V_1 \Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H = V_1 \Sigma_1^{-1} \Sigma_1 \Sigma_1^{-1} U_1^H = V_1 \Sigma_1^{-1} U_1^H = A^\dagger$
iii) $(AA^\dagger)^H = (U_1 \Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H)^H = I^H$ and the Hermetian of the identity is just the identity its self so $(AA^\dagger)^H = I^H = I = AA^\dagger$

iv) $(A^\dagger A)^H = (V_1 \Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H)^H = (V_1 \Sigma_1^{-1} \Sigma_1 V_1^H)^H = (V_1 V_1^H)^H = I^H = I$

v)

vi)