Math, Problem Set #3, Spectral Theory

Name: Benjamin Lim

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.

4.4i Let x be an arbitrary nonzero eigenvector, and $lambda$ be its eigenvalue. First note that $x^H Ax = x^H = \lambda x^H x$. Then note that $x^H Ax = x^H A^H x = (Ax)^H x = (\lambda x)^H x = x^H \lambda^H x = \lambda^H x^H x$. Since $x \neq 0$, it follows that $\lambda = \lambda^H$. Because the hermitian conjugate of a scalar is simply its conjugate and the scalar equals its conjugate, it follows that the eigenvalue λ is a real number.

4.4ii Let x be an arbitrary nonzero eigenvector, and λ be its eigenvalue. First note that $\langle Ax, x\rangle = \langle \lambda x, x\rangle = \bar{\lambda}\langle x, x\rangle$. Then note that $\langle Ax, x\rangle = \langle x, A^H x\rangle = \langle x, -Ax\rangle = \langle x, -\lambda x\rangle = -\lambda\langle x, x\rangle$. Since x is nonzero, we know that $= \bar{\lambda} = \lambda$ for any arbitrary eigenvalue λ. This can only be true if λ is purely imaginary with no real component.

4.6 We want to prove that the diagonal entries of any triangular matrix are its eigenvalues. Without loss of generality, suppose that the n x n matrix A is upper triangular, with entries a_{ij}. To find the eigenvalues, we need to set the determinant of the following matrix to 0 and solve for the eigenvalues λ: $\begin{bmatrix} \mathrm{a}_{1,1}-\lambda & \cdots & \cdots & \cdots & \mathrm{a}_{1,n} \\ & \mathrm{a}_{2,2}-\lambda & \cdots & \cdots & \mathrm{a}_{2,n} \\ & \ddots & \cdots & \cdots & \vdots \\ & & \ddots & \cdots & \vdots \\ & & & \ddots & \vdots \\ & & & & \mathrm{a}_{n,n}-\lambda \end{bmatrix}$.

This reduces to the characteristic polynomial equation $(a_{1,1}-\lambda)(a_{2,2}-\lambda)\cdot\cdots\cdot(a_{n-2,n-2}-\lambda)det(\begin{bmatrix} d_{n-1,n-1}-\lambda & d_{n-1,n} \\ 0 & d_{n,n} \end{bmatrix}) = (a_{1,1}-\lambda)(a_{2,2}-\lambda)\cdot\cdots\cdot(a_{n-2,n-2}-\lambda)(a_{n-1,n-1}-\lambda)(a_{n,n}-\lambda) = 0$ because the 0's in the lower diagonal nullify the off-diagonal terms at each step. It is then easy to see that the roots to the equation above (eignevalues) are simply $a_{1,1}, a_{2,2}, ..., a_{n,n}$ which are the diagonal terms of the matrix A.

4.8i To show that S is a basis for V, we have to show that the functions in S are linearly independent. That is, $a_1 sinx + a_2 cosx + a_3 sin(2x) + a_4 cos(2x) = 0$ only if $a_1 = a_2 = a_3 = a_4 = 0$. To solve for the coefficients, we let $x = 0, \frac{\pi}{2}, \frac{\pi}{3}, \pi$ to get 4 equations to solve for the 4 unknowns. This results in the 4 equations:

$a_1 - a_4 = 0$

$a_2 + a_4 = 0$

$a_4 - a_2 = 0$

$\frac{\sqrt{3}}{2}a_1 - 0.5a_2 + \frac{\sqrt{3}}{2}a_3 + 0.5a_4 = 0$. Solving these, we obtain $a_2 = a_4 = a_1 = a_3 = 0$. Therefore, the elements of the set S are linearly independent. Since V is also the span of the set S, it follows that S forms a basis for V.

4.8ii Noting that $D(sinx) = cosx =$ [0 1 0 0],

$D(cosx) = -sinx =$ [-1 0 0 0],

$D(sin2x) = 2cos2x =$ [0 0 0 2],

$D(cos2x) - 2sin2x =$ [0 0 -2 0]

, the matrix representation of D with respect to the basis $[sinx\ cosx\ sin2x\ cos2x]$ is

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

4.8iii Let the subspaces be $U = \{sinx, cosx\}$ and $W = \{sin2x, cos2x\}$. Then, U and W are complementary subspaces because for any basis $\{u_1, u_2\}$ of U and $\{w_1, w_2\}$ of W, $\{u_1, u_2, w_1, w_2\}$ forms a basis for the space spanned by $V = \{sinx, cosx, sin2x, cos2x\}$. Moreover, U and W are D-invariant because $D(U) = \{cosx, -sinx\} \subset U$ and $D(W) = \{2cos2x, -2sin2x\} \subset W$.

4.13 The matrix A can be diagonalized as $A = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$, where P = $\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$ whose columns are the eigenvectors of A, and the matrix in the middle is a diagonal matrix with the eigenvalues of A as its diagonal entries, and the matrix on the right is P^{-1}.

4.15 Suppose $(\lambda_i)_{i=1}^n$ are eigenvalues of a semisimple matrix A. Then A is diagonalizable and can be rewritten as $A = PDP^{-1}$. Then, $f(A) = a_0 I + a_1 A + ... + a_n A^n = a_0 I + a_1 PDP^{-1} + ... + a_n PD^n P^{-1} = P(a_0 I + a_1 D + a_2 D^2 + ... + a_n D^n)P^{-1} = P \begin{bmatrix} a_0 + a_1\lambda_1 + ... + a_n\lambda_1^n & & \\ & \ddots & \\ & & a_0 + a_1\lambda_n + ... + a_n\lambda_n^n \end{bmatrix} P^{-1}$ because the diagonal entries of D are just the eigenvalues $\lambda_i, i \in \{1, 2, ..., n\}$ of A. As such, we have diagonalized $f(A)$, and so the diagonal entries of the resulting matrix in the middle are the eigenvalues of $f(A)$, where $f(\lambda_i) = a_0 + a_1\lambda_1 + ... + a_n\lambda_n$, as desired.

4.16 i and ii By diagonalization, we can rewrite $A^k = PD^kP^{-1} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix}^k \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} =$

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & 0.4^k \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 2 & -0.4^k \\ 1^k & 0.4^k \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} + \frac{1}{3}0.4^k & \frac{2}{3} - \frac{2}{3}0.4^k \\ \frac{1}{3} - \frac{1}{3}0.4^k & \frac{1}{3} + \frac{2}{3}0.4^k \end{bmatrix}.$$

Then, $B = lim_{k\to\infty}A^k = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$. Then, $A^k - \bar{A} = \begin{bmatrix} \frac{1}{3}0.4^k & -\frac{2}{3}0.4^k \\ -\frac{1}{3}0.4^k & \frac{2}{3}0.4^k \end{bmatrix}$. The norms are as follows:

$$\|A\|_1 = \tfrac{2}{3}0.4^k,\ \|A\|_2 = 0.4^k,\ \|A\|_F = \sqrt{tr(A^HA)} = \sqrt{tr\begin{bmatrix} \frac{1}{9}0.4^{2k} + \frac{1}{9}0.4^{2k} & 0 \\ 0 & \frac{4}{9}0.4^{2k} + \frac{4}{9}0.4^{2k} \end{bmatrix}} =$$

$\sqrt{\frac{10}{9}0.4^{2k}}$, all of which evidently converge to 0. Therefore, the choice of norm does not matter.

4.16 iii Using Theorem 4.3.12, we know $f(x) = 3+5x+x^3$, so $f(0.4) = 3+2+(0.4)^3 = 5+(0.4)^3$.

4.18 Let x be an eigenvector of A^T. We first claim that the eigenvalues of A and AT are equal. To see why, note that the characteristic equation for A and A^T are $det(A - \lambda I) = 0$ and $det(A^T - \lambda I) = 0$, but $(A^T - \lambda I) = (A - \lambda I)^T$. Therefore, if we let $Q = (A - \lambda I)$, the two equations reduce to $det(Q) = 0$ and $det(Q)^T = 0$. By Theorem 2.8.21, we know that for all matrices Q, $det(Q) = det(Q^T)$. Therefore, the characteristic equations are equal and have the same roots. Therefore, both A and A^T have the same eigenvalues. Let λ be an eigenvalue of A. Then, we know $A^Tx = \lambda x$ by our earlier claim. Taking the transpose of both sides, we obtain $x^TA = \lambda x^T$, as desired.

4.20 Since A is Hermitian, $A = A^H$. Since A is orthonormally similar to B, there exists U such that $B = U^HAU$. $B^H = (U^HAU)^H = U^HA^HU = U^HAU = B$.

4.24i 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, $\rho = \frac{\langle x,Ax\rangle}{\|x\|^2} = \frac{\langle x,A^Hx\rangle}{\|x\|^2} = (\frac{x^HA^Hx}{x^Hx})^H = (\frac{\langle x,Ax\rangle}{\|x\|^2})^H = \rho^H$.

4.24ii 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, $-\rho = -\frac{\langle x, Ax\rangle}{\|x\|^2} = \frac{\langle x, -A^H x\rangle}{\|x\|^2} = -(\frac{x^H A^H x}{x^H x})^H = -(\frac{\langle x, Ax\rangle}{\|x\|^2})^H = -\rho^H$.

4.25i 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 $x_j^H x_i = 1 \forall j = i$.

4.25ii 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 Suppose A is positive definite and $\langle x, Ax\rangle \geq 0 \forall x \neq 0$. Let e_j be the jth standard basis vector. Since $x^T Ax \geq 0$, it follows that $e^T Ae \geq 0$. Then note that $e^T Ae = 0 + ... + 0 + 1 \cdot a_{jj} + 0... + 0 = a_{jj} \geq 0$, where $a_j j$ is the jth diagonal entry of the matrix A. Positive definiteness also implies $A^H = A$, such that $a_{jj} = (e_j)^T Ae_j = \langle e_j, Ae_j\rangle = \langle A^H e_j, e_j\rangle = \langle Ae_j, e_j\rangle = \overline{\langle e_j, Ae_j\rangle} = \overline{(e_j)^T Ae_j} = \overline{a_{jj}}$ which is true if and only if its diagonal entries are real.

4.28 Since A is positive semidefinite, $x^H Ax \geq 0$. From 4.27, we know the diagonal elements of A and B are non-negative. Then it follows that $tr(AB) = \sum_{i=j}^n a_{ii} b_{jj} \geq 0$ as desired for the first inequality. For the second inequality, consider $tr(A)tr(B) = \sum_{i=1}^n a_{ii} \sum_{j=1}^n b_{jj} = \sum_{i=1}^n \sum_{j=1}^n a_{ii} b_{jj} = \sum_{i=1}^n \sum_{j\neq i}^n a_{ii} b_{jj} + \sum_{i=j}^n a_{ii} b_{jj} = \sum_{i=1}^n \sum_{j\neq i}^n a_{ii} b_{jj} + tr(AB) \geq tr(AB)$.

4.31i $\|A\|_2 = sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = sup_{x\neq 0} \frac{\|U\Sigma V^H x\|_2}{\|x\|_2} = sup_{x\neq 0} \| \frac{\Sigma V^H x\|_2}{\|VV^H x\|_2}$ because U and V are orthonormal, such that $\|U\| = 1$ and $VV^H = I$. Letting $y = V^H x$, we obtain $sup_{y\neq 0} \frac{\|\Sigma y\|_2}{\|Vy\|_2} = sup_{y\neq 0} (\frac{\sum_{i=1}^r \sigma_i^2 |y_i|^2}{\sum_{i=1}^r |y_i|^2})^{0.5}$, where σ_i are the singular values of A, decreasing in i, such that σ_i is the largest singular value. Therefore, taking the supremum with respect to y gives us $y = (1\ 0 \cdots 0)^T$, such that $\|A\|_2 = sup_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} = \sigma_1$.

4.31ii Suppose A is invertible. Then A^{-1} exists. $\|A^{-1}\|_2 = sup_{x\neq 0}\frac{\|A^{-1}x\|_2}{\|x\|_2} = sup_{x\neq 0}\frac{\|V\Sigma^{-1}U^Hx\|}{\|x\|_2} = sup_{x\neq 0}\frac{\|V\Sigma^{-1}U^Hx\|}{\|UU^Hx\|} = sup_{x\neq 0}\frac{\|\Sigma^{-1}U^Hx\|}{\|UU^Hx\|} = sup_{x\neq 0}\frac{\|V\Sigma^{-1}y\|}{\|Uy\|} = sup_{y\neq 0}(\frac{\sum_{i=1}^r \sigma_i^{-2}|y_i|^2}{\sum_{i=1}^r |y_i|^2})^{0.5}$, where σ_i are the singular values of A, decreasing in i, such that σ_i is the largest singular value. Therefore, taking the supremum with respect to y gives us $y = (1\ 0 \cdots 0)^T$, such that $\|A\|_2 = sup_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} = \sigma_1^{-1}$.

4.32i Suppose U and V are orthonormal. Then, $\|UAV\|_F = \sqrt{tr[(UAV)^H UAV]} = \sqrt{tr[V^HA^HU^HUAV]} = \sqrt{tr[V^HA^HAV]} = \sqrt{tr[VV^HA^HA]} = \sqrt{tr[A^HA]} = \|A\|_F$.

4.32ii $\|A\|_F = \|UAV\|_F = \|UU\Sigma V^HV\|_F = \|UU\Sigma\|_F = \sqrt{tr[(UU\Sigma)^H UU\Sigma]} = \sqrt{tr[\Sigma^HU^HU^HUU\Sigma]} = \sqrt{tr[\Sigma^H\Sigma]} = tr(\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2 \end{bmatrix}) = \sqrt{\sigma_1^2 + ... + \sigma_n^2}$.

4.33 $LHS = \|A\|_2 = \sigma_1$ by 31i.

$RHS = \sup_{\|x\|_2=1, \|y\|_2=1} |y^HAx|$

$= \sup_{\|x\|_2=1, \|y\|_2=1} |y^HU\Sigma V^Hx|$

$= \sup_{\|a\|_2=1, \|b\|_2=1} |a^HU^HU\Sigma V^HVb|$ by rewriting $x = Ua$ and $y = vb$. Note that $\|x\|_2 = \|Ua\|_2 = \|a\|_2$ and $\|y\|_2 = \|Vb\|_2 = \|b\|_2$ because U and V are orthonormal.

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

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

To get the supremum with respect to a and b such that their 2-norms each equal one, we let $a = b = (1, 0, 0, ..., 0)^T$ because the maximum occurs when we put all the weight on the largest eigenvalue, σ_1. Thus, $RHS = (1, 0, ..., 0)\begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2 \end{bmatrix}(1, 0, ..., 0)^T = |\sigma_1| = \sigma_1 = \|A\|_2 = LHS$.

4.36 $A = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$. 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 $A^\dagger = V_1 \Sigma_1^{-1} U_1^H$ 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) $(AA^\dagger)^H = (U_1\Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H)^H = (U_1 \Sigma_1 \Sigma_1^{-1} U_1^H)^H = U_1(\Sigma_1^{-1})^H (\Sigma_1)^H U_1^H = U_1(\Sigma_1^H)^{-1}(\Sigma_1)^H U_1^H = U_1 \Sigma_1 (\Sigma_1)^{-1} U_1^H = U_1 \Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H = AA^\dagger$.

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