Math 344 Homework 4.4

Chris Rytting

November 13, 2015

4.19

We know that there exists an eigenbasis on T. Then let

$$T = PDP^{-1}$$

Let $Q = PV^{-1}$, wehre V is the matrix of orthonormal basis vectors. Then Q is orthonormal, $QV = P$, and $QVDV^{-1}Q^{-1}$ is an orthonormal basis of T.

4.20

We have the following

$$A^H = A \quad A = U^H B^H U = A^H = U^H B U$$

and thus we have that $B^H = B$, which is the desired result.

4.21 (i)

Let $V \in \Sigma_{\lambda_i}(S)$. Note that

$$Sx = \lambda x$$
$$TSx = T\lambda x$$
$$STx = \lambda Tx$$

and thus

$$TV \in \Sigma_{\lambda_i}$$

as desired

4.21 (ii)

S is simple, so it has n distinct eigenvectors. Therefore, the dimension of each eigenspace is 1. Now let

$$v \in \Sigma_{\lambda_i}(S)$$

and by part (i)

$$\implies Tv \in \Sigma_{\lambda_i}(S)$$

but it's one dimensional, so

$$Tv = a_i v$$

for some $a_i \in \mathbb{F}$. Therefore, we have $\Sigma_{a_i}(T)$ for each λ_i, and by proposition 4.4.7, we have that T is semi-simple.

Exercise 4.22 (i)

By proposition 3.7.12, If V is a finite-dimensional inner product space, then we have

$$(i).If\ S \in V,\ then\ S^{**} = S$$
$$(ii).If\ S, T \in V,\ then (ST)^* = T^* S^*$$

Therefore we have that

$$(TT^*)^* = T^{**}T^* = TT^*$$

So TT^* is self-adjoint. We also have that

$$\langle \mathbf{w}, TT^*\mathbf{w} \rangle = \langle T^*\mathbf{w}, T^*\mathbf{w} \rangle \geq 0$$

Exercise 4.22 (ii)

Because TT^* is normal, it is orthonoramlly similar to a diagonal matrix D. Thus,

$$TT^* = UDU^{-1}$$

Thus, define $S = U\sqrt{D}U^{-1}$, where $\sqrt{D}$ is the elementwise square root of D.

$$S^2 = U\sqrt{D}U^{-1}U\sqrt{D}U^{-1} = U\sqrt{U}^2 U{-1} = UDU^{-1} = TT^*$$

To show that it is self-adjoint, Note

$$\begin{aligned} SS &= TT^* \\ S &= TT^* S^{-1} \\ S^* &= (TT^* S^{-1})^* \\ S^* &= S^{-1*} TT^* \\ S^* S^* &= TT^* = SS \end{aligned}$$

Which implies that $S = S^*$

Exercise 4.22 (iii)

It is sufficent to show that $(S^{-1}T)^*S^{-1}T = I$

$$\begin{aligned}(S^{-1}T)^*S^{-1}T &= T^*(S^{-1})^*S^{-1}T \\ &= T^*S^{-1}S^{-1}T \\ &= T^*(S^2)^{-1}T \\ &= T^*(TT^*)^{-1}T \\ &= T(T^*)^{-1}T^{-1}T \\ &= I\end{aligned}$$

4.23

Given an nxn matrix we define the Rayleigh quotient:

$$p(x) := \frac{\langle x, Ax \rangle}{||x||^2}$$

Show that the Rayleigh Quotient can only take on real values for Hermitian matrices and only imaginary values for skew-Hermitian matrices.

Proof:
We assume that A is hermitian. We use an important identity of Hermitian matrices, that $A^H = A$. We use this identity and the properties of inner products giving us $\langle x, Ax \rangle = x^H Ax = x^H A^H x = (x^H Ax)^H = \overline{\langle x, Ax \rangle}$. This means that the complex conjugate of the inner product is equal to the inner product. Which implies that all elements of the inner product are real. Wince the deonominator of the Rayleigh quotient is a norm, by definition the norm is a real number, meaning that the denominator is also a real number. A real number divided by a real number is real, as desired. ■

Proof:
We assume that A is skew-Hermitian. An important property regarding skew-hermitian's is that $A^H = -A$. So from there we can also use the properties of inner products and say $\langle x, Ax \rangle = x^H Ax = -x^H A^H x = -(x^H Ax)^H = -\overline{\langle x, Ax \rangle}$. Since this innerproduct is equal to the its negation after a hermitian, entails that there are only imaginary parts. So this pure imaginary number can be written as bi where $b \in \mathbb{R}$. The denominator in the Rayleigh quotient is a norm, which by definition produces only real numbers. We see that the Rayleigh quotient $= \frac{bi}{a} = i\frac{b}{a}$ and $\frac{b}{a} \in \mathbb{R}$ so the Rayleigh quotient of a skew-hermitian only takes on imaginary numbers. ■

Exercise 4.24 (i)

A is normal with eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ and corresponding orthonormal eigenvectors$\{x_1, x_2, ..., x_n\}$. We can use the decomposition $A = UDU^{-1}$. U is the orthonormal matrix of the eigenvectors and therefore $U^H = U^{-1}$ and that$U^H U = UU^H$. We know that $UU^{-1} = UU^H = U^H U$ so $U^H U = I$. Showing that $U^H U = x_1 x_1^H + x_2 x_2^H + ... + x_n x_n^H$ will prove that $I = x_1 x_1^H + x_2 x_2^H + ... + x_n x_n^H$.

Since the vectors of U are orthonormal when we multiple any of the columns of U against any of the rows of U^H which do not correspond to the same eigenvector this results in a zero norm, and when they do correspond this returns $x_i x_i^H$, showing that $U^H U = x_1 x_1^H + x_2 x_2^H + ... + x_n x_n^H$, thus proving that $I = x_1 x_1^H + x_2 x_2^H + ... + x_n x_n^H$.

Exercise 4.24 (ii)

Using the same decomposition as in (i), we can express A in the following manner:

$$\begin{aligned} A &= U^H D U \\ &= \left[\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}\right]^{\mathbf{H}} \left[\lambda_1, \lambda_2, ..., \lambda_n\right] \left[\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}\right] \end{aligned}$$

where the columns of U are the columns $\{\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}\}$, the diagonal of D consists of the eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_n\}$, and the rows of U^H consist of $\{\mathbf{x_1^H}, \mathbf{x_2^H}, ..., \mathbf{x_n^H}\}$. If we multiply these matrices, we get

$$A = \begin{bmatrix} \lambda_1 \mathbf{\bar{x}_{11} x_{11}} + \lambda_2 \mathbf{\bar{x}_{21} x_{21}} + ... + \lambda_n \mathbf{\bar{x}_{n1} x_{n1}} & ... & \lambda_1 \mathbf{\bar{x}_{1n} x_{11}} + \lambda_2 \mathbf{\bar{x}_{2n} x_{21}} + ... + \lambda_n \mathbf{\bar{x}_{nn} x_{n1}} \\ \vdots & \ddots & \vdots \\ \lambda_1 \mathbf{\bar{x}_{11} x_{1n}} + \lambda_2 \mathbf{\bar{x}_{21} x_{2n}} + ... + \lambda_n \mathbf{\bar{x}_{n1} x_{nn}} & ... & \lambda_1 \mathbf{\bar{x}_{1n} x_{1n}} + \lambda_2 \mathbf{\bar{x}_{2n} x_{2n}} + ... + \lambda_n \mathbf{\bar{x}_{nn} x_{nn}} \end{bmatrix}$$

$$= \lambda_1 \mathbf{x_1 x_1^H} + \lambda_2 \mathbf{x_2 x_2^H} + ... + \lambda_n \mathbf{x_n x_n^H}$$

4.20 Assume $A, B \geq 0$. Prove that $0 \leq tr(AB) \leq tr(A)tr(B)$ and use this to prove that $||\cdot||_F$ is a matrix norm.

Proof:

In this proof we will use Proposition 4.4.5 and we also will make use of the fact that in general $tr(CM) = tr(MC)$ for two square matrices M and N. So we can write $A = UD_1U^H$ and $B = VD_2V^H$ where U and V are orthonormal matrices and D_1 and D_2 are diagonal matrices and where d_{ij} and δ_{ij} are the (i, j) entries of D_1 and D_2 respectively. So it follows that $tr(A)tr(B) = tr(UD_1U^H)tr(VD_2V^H) = tr(U^HUD_1)tr(V^HVD_2) = tr(D_1)tr(D_2) = \sum_{i=1}^{n} d_{ii} = \sum_{i=1}^{n} \delta_{ii} \geq \sum_{i=1}^{n} d_{ii}\delta_{ii} = tr(D_1D_2)$.

At this point we need to show that $tr(D_1D_2) \geq tr(UD_1U^HVD_2V^H)$

$$tr(UD_1U^HVD_2V^H) = tr\left(\begin{pmatrix} u_{11}d_{11} & & 0 \\ & \ddots & \\ 0 & & u_{nn}d_{nn} \end{pmatrix} U^H \begin{pmatrix} v_{11}\delta_{11} & & 0 \\ & \ddots & \\ 0 & & v_{nn}\delta_{nn} \end{pmatrix} V^H\right)$$

$$= tr\left(\left(\begin{array}{ccc} u_{11}^2 d_{11} & & 0 \\ & \ddots & \\ 0 & & u_{nn}^2 d_{nn} \end{array}\right)\left(\begin{array}{ccc} v_{11}^2 \delta_{11} & & 0 \\ & \ddots & \\ 0 & & v_{nn}^2 \delta_{nn} \end{array}\right)\right)$$

$$= tr\left(\begin{array}{ccc} v_{11}^2 u_{11}^2 d_{11}\delta_{11} & & 0 \\ & \ddots & \\ 0 & & v_{nn}^2 u_{nn}^2 d_{nn}\delta_{nn} \end{array}\right)$$

And since U and V are orthonormal $u_{ij} \leq 1$ and $v_{ij} \leq 1$.
Hence $u_{ii}^2 v_{ii}^2 d_{ii}\delta_{ii} \leq d_{ii}\delta_{ii}$ which implies $tr(AB) \leq tr(D_1 D_2)$.
Furthermore, Theorem 4.4.4 states that positive semi-definite matrices have nonnegative eigenvalues so we have $d_{ii} \geq 0$ and $\delta_{ii} \geq 0$ which gives $0 \leq tr(AB)$. At this point the proof is complete. ■

Now we show that the Frobenius is a norm.

Proof:

Positivity:

Proposition 4.4.7 shows that AA^H is positive semi-definite. $||A||_F = \sqrt{tr(AA^H)} = \sqrt{tr(AA^H)}$. I is also positive semi-definite. So by our proof of the first part of this problem we know that the trace of the product of to positive semidefinite matrices is nonnegative.

Scale Preservation:

$||\alpha A||_F = \sqrt{tr(\alpha A \alpha A^H)} = \sqrt{\alpha^2 tr(AA^H)} = \sqrt{\alpha^2}\sqrt{tr(AA^H)} = |\alpha|||A||_F$.

Triangle Inequality:

$||A+B||_F = \sqrt{tr((A+B)(A+B)^H)} = \sqrt{tr((A+B)(A^H+B^H))} = \sqrt{tr(AA^H + AB^H + BA^H + BB^H}$
$= \sqrt{tr(AA^H) + tr(BB^H) + 2tr(AB^H)}$

So $||A+B||^2 = tr(AA^H) + tr(BB^H) + 2tr(AB^H)$

$||A||_F + ||B||_F = \sqrt{tr(AA^H)} + \sqrt{tr(BB^H)}$

$(||A||_F + ||B||_F)^2 = ||A||_F^2 + 2||A||_F||B||_F + ||B||_F^2 = tr(AA^H) + 2\sqrt{tr(AA^H)tr(BB^H)} + tr(BB^H)$

In order to show that the triangle inequality holds we only need to show that $2\sqrt{tr(AA^H)tr(BB^H)} \geq 2tr(AB^H)$, or equivalently that $tr(AA^H)tr(BB^H) \geq tr(AB^H)tr(AB^H)$

$tr(AA^H)tr(BB^H) = ||A||_F^2||B||_F^2$ and by Cauchy-Scharz we have
$||A||_F^2||B||_F^2 \geq \langle A, B\rangle^2 = tr(AB^H)^2$

So the Triangle Inequality holds. ■