4.1 A matrix $A \in M_n(\mathbb{F})$ is said to be nilpotent if $A^k = 0$ for some $k \in \mathbb{N}$. Show that if λ is any eigenvalue of a nilpotent matrix, then $\lambda = 0$

4.2 Let V be the subspace $V = span(1, x, x^2)$ in the inner product space $L^2([0,1], \mathbb{R})$. Let D be the derivative operator $D : V \to V$. Find all the eigenvalues and eigenspaces of D. What are their algebraic and geometric multiplicites?

First of all $D = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

Since D is an upper right triangular matrix its eigenvalues are on its diagonal. So we have $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Now we find the eigenvectors: $D\vec{x} = \lambda\vec{x} \Rightarrow \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ 2x_3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

So x_1 is arbitrary, $x_2 = x_3 = 0$, and an eigenvector for $\lambda_1 = 0$ is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

And so it follows that the eigenspace of D is $span \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

The albegraic multiplicty of λ_1 is 3 and its geometric multiplicity is 1. ♣

4.3

4.4 We know that a matrix $A \in M_n(\mathbb{F})$ is walled Hermitian means that $A^H = A$ and skew-Hermitian if $A^H = -A$. Using exercise 4.3, prove the following:

(i) A Hermitian 2x2 matrix has only real eigenvalues.
This means that, by the definition $A^H = A$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

The determinant of the second matrix would be $(ad - b^2)$. and the $tr(A) = (a + d)$. Having this we can plug this into the characteristic equation $p(\lambda) = \lambda^2 - tr(A)\lambda + det(A)$ to get $p(\lambda) = \lambda^2 - (a + d)\lambda + (ad - b^2)$. In order to solve for the eigenvectors we take the square root of the characteristic equation, getting

$$a + d \pm \frac{\sqrt{(a + d)^2 - 4(ad - b^2)}}{2}$$

$$= a + d \pm \frac{\sqrt{(a)^2 + 2ad + d^2 - 4(ad) + 4b^2)}}{2}$$

$$= a + d \pm \frac{\sqrt{(a - d)^2 + 4b^2)}}{2}$$

1

We know that $(a - d)^2 \geq 0$ and $4b^2 \geq 0$ so there will be no negative values in the square root. Therefore there will be no imaginary numbers, so there will only be real eigenvalues. ∎

(ii) A skew-Hermitian 2 X 2 matrix has only imaginary eigen values:

Proof:

Since A is skew-Hermitian we know that $A^H = -A$. This means that all the elements are equal to their complex conjugate. , or in otherwords $a_{11} = -\bar{a}_{11}$, and this means that all the elements on the diagonal are imaginary numbers because of this property.We alse get $a_{12} = -\bar{a}_{21}$ since we are taking the transpose. So we use x and y to show the real and diagonal elements of a_{21}. Then we have $a_{12} = -\overline{x + iy} = (-1)(x - iy) = -x + iy$. We then create a matrix $A = \begin{bmatrix} ai & -x + iy \\ x + iy & di \end{bmatrix}$ Where $a, d, x, y \in \mathbb{R}$

And so using the characteristic polynomial from 4.3 we get $p(y) = \lambda^2 - tr(A)\lambda + det(A) = \lambda^2 - (ai + di)\lambda - (-x + iy)(x + iy) = 0$

Now we try and solve for the eigenvalues from the characteristic polynomial using the quadratic formula:

$\lambda = \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{(ai + di)^2 - 4(1)(aidi - (-x + iy)(x + iy))}$

$= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{(ai)^2 + 2aidi + (di)^2 - 4(-ad + x^2 + y^2)}$

$= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{-a^2 + 2ad - d^2 - 4x^2 - 4y^2}$

$= \frac{1}{2}(ai + di) \pm \frac{1}{2}\sqrt{(-1)(a^2 - 2ad + d^2) - 4x^2 - 4y^2}$

$= \frac{i}{2}(a + d) \pm \frac{1}{2}\sqrt{(-1)(a - d)^2 - 4x^2 - 4y^2}$

$= \frac{i}{2}(a + d) \pm \frac{1}{2}\sqrt{(-1)((a - d)^2 + 4x^2 + 4y^2)}$

$= i\left(\frac{1}{2}(a + d) \pm \frac{1}{2}\sqrt{(a - d)^2 + 4x^2 + 4y^2}\right)$

We see that $a, d, x, y \in \mathbb{R}$ then $x^2 \geq 0, y^2 \geq 0, (a - d)^2 \geq 0$ because of this the things under the square root are positive and so there will be no imaginary parts created, and we get $\frac{a+d}{2}$ being a real number, but i times the sum of two real numbers is a pure imaginary number. So, the eigenvalues of the matrix are pure imaginary. ∎

Exercise 4.6

Let

$$A = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$$

Compute the transition matrix P such that $P^{-1}AP$ is diagonal.

$det(A - \lambda I) = det\left(\begin{bmatrix} 0.8 - \lambda & 0.4 \\ 0.2 & 0.6 - \lambda \end{bmatrix}\right) = (0.8 - \lambda)(0.6 - \lambda) - 0.8 = 0$

$\lambda^2 - 1.4\lambda + 0, 4 = 0$

At this point we use the quadratic formula to solve for the eigenvalues.

$\lambda = \frac{1.4 \pm \sqrt{(1.4)^2 - 4(0.4)}}{2}$

$= 1.4 \pm \frac{1}{2}\sqrt{1.96 - 1.6}$

$= 0.7 \pm 0.5\sqrt{0.36}$

$= 0.7 \pm 0.3$

$\Rightarrow \lambda_1 = 0.4$ and $\lambda_2 = 1$

We use these eigenvalues to find their respective eigenvectors

$\begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0.4 \begin{bmatrix} x \\ y \end{bmatrix}$ gets us $0.8x + 0.4y = 0.4x$ which gives us $x = -y$ this means

the eigenvector for the first lambda $\lambda_1 = 0.4$ is $\begin{bmatrix} 0.7071 \\ -0.7071 \end{bmatrix}$

Now we find the second eigenvector.

$\begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$ gets us $0.8x + 0.4y = x \Rightarrow x = 2y$ using this we get that the eigen

value for $\lambda_2 = 1$ is $\begin{bmatrix} 0.89 \\ 0.447 \end{bmatrix} \Rightarrow P = \begin{bmatrix} 0.707 & 0.8944 \\ -0.8944 & 0.707 \end{bmatrix} \Rightarrow P^{-1} = \begin{bmatrix} \frac{\sqrt{2}}{3} & \frac{-2\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix} D =$

$\begin{bmatrix} 0.4 & 0 \\ 0 & 1 \end{bmatrix}$

Exercise 4.8 We use matrix A from above.

(i)Compute $\lim\limits_{n\to\infty} A^n$. That is find a matrix B such that for any $\epsilon > 0$ there exists an $N > 0$ with $||A^k - B||_1 < \epsilon$ whenever $k > N$.

$\lim\limits_{n\to\infty} A^n = \lim\limits_{n\to\infty} (PDP^{-1})^n = \lim\limits_{n\to\infty} PD^nP^{-1} = \lim\limits_{n\to\infty} P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = P \begin{bmatrix} \lim\limits_{n\to\infty} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} =$

$P \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = B$

First we show that A^n converges to B using the $|| \cdot ||_1$ norm.

We get $A^n = P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1}$ And so $A^n - B = P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} P^{-1} - P \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} =$

$P \left(\begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) P^{-1} = P \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 0 \end{bmatrix} P^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} (\frac{2}{5})^n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{3} & \frac{-2\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix}$

$= \begin{bmatrix} \frac{1}{\sqrt{2}}(\frac{2}{5})^n & 0 \\ \frac{-1}{\sqrt{2}}(\frac{2}{5})^n & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{3} & \frac{-2\sqrt{2}}{3} \\ \frac{\sqrt{5}}{3} & \frac{\sqrt{5}}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{3}(\frac{2}{5})^n & \frac{-2}{3}(\frac{2}{5})^n \\ \frac{-1}{3}(\frac{2}{5})^n & \frac{2}{3}(\frac{2}{5})^n \end{bmatrix} = \frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$

So, $||A^n - B||_1 = ||\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}||_1 = |\frac{2}{3}(\frac{2}{5})^n| + |\frac{-2}{3}(\frac{2}{5})^n| + |\frac{1}{3}(\frac{2}{5})^n| + |\frac{-1}{3}(\frac{2}{5})^n| =$

$6|\frac{1}{3}(\frac{2}{5})^n| = 2|(\frac{2}{5})^n|$

So for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow 2|(\frac{2}{5})^n| < \epsilon$
So for all $\epsilon > 0$ we will be able to choose $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow$

$||A^n - B||_1 < \epsilon$, as desired. ∎

We now will show that $||A^n - B||_\infty$ converges to zero, using an infinity norm
Proof:
This is very similar to the previous proof, so we proceed where we proceed using the ∞-norm.

So have $||A^n - B||_\infty = ||\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}||_\infty = |\frac{2}{3}(\frac{2}{5})^n|$

For each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow |\frac{2}{3}(\frac{2}{5})^n| < \epsilon$

Hence for each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow ||A^n - B||_1 < \epsilon$, as desired. ∎

We now prove the convergence of A^n to B using the 2-norm.
Proof:

We have $||A^n - B||_2 = ||\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}||_2$

$= (|\frac{1}{3}(\frac{2}{5})^n(1)|^2 + |\frac{1}{3}(\frac{2}{5})^n(-1)|^2 + |\frac{2}{3}(\frac{2}{5})^n(1)|^2 + |\frac{-2}{3}(\frac{2}{5})^n(1)|^2)^{\frac{1}{2}} = (10|\frac{1}{3}(\frac{2}{5})^n(1)|^2)^{\frac{1}{2}} = \sqrt{10}|\frac{1}{3}(\frac{2}{5})^n(1)|$

For each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow \sqrt{10}|\frac{1}{3}(\frac{2}{5})^n(1)| < \epsilon$

Hence for each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $n > N \Rightarrow ||A^n - B||_2 < \epsilon$, as desired. ∎

3e We now prove the convergence of A^n to B using the Frobenius norm.

We have $||A^n - B||_{Frobenius} = ||\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}||_{Frobenius}$

$= \sqrt{tr\left(\left(\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}\right)\left(\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}\right)^H\right)}$

$= \sqrt{tr\left(\frac{1}{3}(\frac{2}{5})^n\frac{1}{3}(\frac{2}{5})^n \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}\right)}$

$= \sqrt{\frac{1}{9}(\frac{2}{5})^{2n} tr\left(\begin{bmatrix} 5 & -5 \\ -5 & 5 \end{bmatrix}\right)} = = \frac{1}{3}(\frac{2}{5})^n\sqrt{10}$

For each $\epsilon > 0$ we can choose an $N \in \mathbb{N}$ large enough so that $\frac{1}{3}(\frac{2}{5})^n \sqrt{10} < \epsilon$ ∎

We see that it converges regardless of the norm used.

(ii)Using the matrix
$$A = \begin{bmatrix} .8 & .4 \\ .2 & .6 \end{bmatrix}$$

we find all the eigenvalues of the matrix $3 + 5A + A^3$ So we get,

$$\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} + \begin{bmatrix} .512 & .064 \\ .008 & .216 \end{bmatrix}$$

$$= \begin{bmatrix} 7.512 & 2.064 \\ 1.008 & 6.216 \end{bmatrix}$$

Finding the eigenvalues,

$$= \begin{bmatrix} 7.512 - \lambda & 2.064 \\ 1.008 & 6.216 - \lambda \end{bmatrix}$$

we get the characteristic polynomial $\lambda^2 - 13.728\lambda + 44.6141$. Using the quadratic equation,

$$\lambda = \frac{13.728 \pm \sqrt{13.728^2 - 4(44.6141)}}{2}$$

we get the eigenvalues $\lambda_1 = 8.445$ and $\lambda = 5.2827$

4.10

For any semisimple matrix $A \in M_n(\mathbb{F})$, let $p(x)$ be the characteristic polynomial of A. Prove that $p(A) = 0$.

In order to do these we first note that A is a semisimple matrix, so it diagonalizable. This means that A can be written as $A = PDP^{-1}$, where D is a diagonal matrix. Plugging in A into the characteristic equation we get,

$$p(A) = a_1 A + a_2 A^2 + \dots + a_n A^n$$

and now substituting in $A = PDP^{-1}$

$$a_1(PDP^{-1}) + a_2(PDP^{-1})^2 + \dots + a_n(PDP^{-1})^n$$

by proposition 4.2.9 we can expand the diagonal matrices with the eignvalues on the diagonal giving us

$$= a_1 P \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} P^{-1} + a_2 P \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix} P^{-1} + \dots + a_n P \begin{pmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_n^n \end{pmatrix} P^{-1}$$

factoring out the P and P^{-1} we get

$$= P \left(a_1 \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} + a_2 \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix} + \dots + a_n \begin{pmatrix} \lambda_1^n & & 0 \\ & \ddots & \\ 0 & & \lambda_n^n \end{pmatrix} \right) P^{-1}$$

adding the matrices together we get

$$= P \begin{pmatrix} a_1\lambda_1 + ... + a_n\lambda_1^n & & 0 \\ & \ddots & \\ 0 & & a_1\lambda_n + ... + a_n\lambda_n^n \end{pmatrix} P^{-1}$$

and since $p(\lambda_i) = 0$ for all characteristic equations and for $i = 1, ..., n$ we get

$$p(A) = P \begin{pmatrix} a_1\lambda_1 + ... + a_n\lambda_1^n & & 0 \\ & \ddots & \\ 0 & & a_1\lambda_n + ... + a_n\lambda_n^n \end{pmatrix} P^{-1}$$

$$= P \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix} P^{-1} = \mathbf{0} \qquad\qquad \blacksquare$$

4.12

Give an example to show that the power method may fail if there are two distinct dominant eigenvalues of the same magnitude. What happens if the dominant eigenvalue has two linearly independent eigenvectors?

To show an example of the power method breaking, let's start with the following matrix: $\begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix}$

This matrix has eigenvalues of $\lambda_1 = 2$ and $\lambda_2 = -2$. These two clearly have the same absolute value, so the power method should have problems. For reference, I've also calculated the eigenvalues, so we have an idea of what it's supposed to converge to.

$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} -3 \\ 4 \end{bmatrix}$

Are the eigenvectors corresponding to the eigenvalues of 2 and -2, respectively. Now, using the power method with the initial guess of $[1, 1]^T$ we get:

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{x}_3 = A\mathbf{x}_2 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 20 \\ -8 \end{bmatrix} = 4 \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

$$\mathbf{x}_4 = A\mathbf{x}_3 = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 20 \\ -8 \end{bmatrix} = \begin{bmatrix} 16 \\ 16 \end{bmatrix} = 16 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\vdots$$

As we can see, there is an alternating pattern in the vectors and it clearly dosen't converge to one of our analytic eigenvectors. So clearly the power method doesn't always work when there are two dominant eigenvalues.

Now if we examine what happens when there one dominant eigenvalue with two eigen vectors. We'll show a trivial example: $\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$

Which has only one eigenvalue, 5. It also has two eigenvectors from this, which are: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Starting with our same initial guess, we get:

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 5 \\ 5 \end{bmatrix} = 25 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\vdots$$

Which rather than oscillating around two vectors, it only stays on the one vector. So, long story short, only one dominant eigenvalue with one eigenvector will work. ■

4.16

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. ∎

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 $tr(D_1D_2) \geq tr(UD_1U^HVD_2V^H)$

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

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

$$= tr\begin{pmatrix} v_{11}^2 u_{11}^2 d_{11}\delta_{11} & & 0 \\ & \ddots & \\ 0 & & v_{nn}^2 u_{nn}^2 d_{nn}\delta_{nn} \end{pmatrix}$$

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 $tr(AB) \leq tr(D_1D_2)$.

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. ∎

4.22

Let V be the subspace $V = span(1, x, x^2)$ in the inner produt space $L^2([0, 1], \mathbb{R})$
Let D be the derivative operator $D : V \rightarrow V$. Write the matrix representation of D with respect to the basis $\{1, x, x^2\}$ and compute its singular value decomposition.

We find that the derivative operator

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

We then find the singular value decomposition, by first finding $D^H D$ and DD^H. We get,

$$D^H D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

and

$$DD^H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We then find the eigenvalues and vectors of DD^H and $D^H D$ respectively

$$\lambda_1 = 4 : \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \lambda_2 = 1 : \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \lambda_3 = 0 : \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and for $D^H D$

$$\lambda_1 = 4 : \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \lambda_2 = 1 : \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \lambda_3 = 0 : \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

We then plug the values for $U : (DD^H)$, $V : (D^H D)$ and Σ(the matrix of the square roots of the eigenvalues) into the equation for singular decomposition $U\Sigma V^H$ and we get

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

and when multiplying these matrices together we get

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

which is our original matrix.

Exercise 4.24

Assume $A \in M_{mxn}(\mathbb{F})$ is of rank r.
(i) We prove that $||A||_F = (\sigma_1^2 + \sigma_2^2 + ... + \sigma_r^2)^{1/2}$.
We note that, by the Frobenius norm,

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

Since A is of rank r and $A \in M_{mxn}(\mathbb{F})$ we can write $A = U\Sigma V^H$ where V and U are orthonormal matrices. So we can continue saying,

$$\sqrt{tr(AA^H)} = \sqrt{tr((U\Sigma V^H)(U\Sigma V^H)^H)} = \sqrt{tr(U\Sigma V^H V \Sigma^H U^H)}$$

$$\sqrt{tr(U\Sigma I\Sigma U^H)} = \sqrt{tr(U^H U\Sigma^2)} = \sqrt{tr(\Sigma^2)} = (\sigma_1^2 + \sigma_2^2 + ... + \sigma_r^2)^{1/2}$$

as desired. ∎

(ii) Since $U \in M_m(\mathbb{F})$ and $V \in M_n(\mathbb{F})$ are orthonormal matrices we prove that $||UAV||_F = ||A||_F$ We first note that, by the Frobenius norm,

$$||UAV||_F = \sqrt{tr((UAV)(UAV)^H)} = \sqrt{tr(UAVV^H A^H U^H)}$$
$$= \sqrt{tr(UAIA^H U^H)} = \sqrt{tr(U^H UAA^H)} = \sqrt{tr(AA^H)} = \sqrt{tr(A^H A)}$$
$$= ||A||_F$$

Exercise 4.26
Prove that $|\det A| = \prod_{i=1}^n \sigma_i$, or in other words, the modulus of the determinant is the product of the singular values. Notice that A can be rewritten as $A = U\Sigma V^H$ so we get

$$|\det(A)| = |\det(U\Sigma V^H)| = |\det(U)\det(\Sigma)\det(V^H)| = 1 \cdot |\det \Sigma| \cdot 1$$

$$= |\sigma_1 \cdot \sigma_2 \cdot \sigma_3 \cdot ... \cdot \sigma_n|$$

since all the singular values are positive we get

$$= \sigma_1 \cdot \sigma_2 \cdot \sigma_3 \cdot ... \cdot \sigma_n$$

as desired. ∎

Exercise 4.28

The good news is, the work for this problem is mostly done from problem 21, we know that:

$$U = \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 0 & 1 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sqrt{15} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad V = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

So, we're clipping the Σ, but there isn't anything else to really clip, so it's going to end up being the transpose. So now using:

$$\begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{15}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 0 & 1 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \end{bmatrix}^H$$

The result is:

$$A^\dagger = \begin{bmatrix} \frac{1}{15} & 0 & \frac{2}{15} \\ \frac{1}{15} & 0 & \frac{2}{15} \\ \frac{1}{15} & 0 & \frac{2}{15} \\ 0 & 0 & 0 \end{bmatrix}$$

Then $A^\dagger A$ is: $\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $A^H A$ is: $\begin{bmatrix} 5 & 5 & 5 & 0 \\ 5 & 5 & 5 & 0 \\ 5 & 5 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

These two are very similar, they just differ by a scalar multiple.

Exercise 4.34 Consider the matrix

$$V_n = \begin{bmatrix} 1 & x_0 & x_0^2 & . & . & . & x_0^n \\ 1 & x_1 & x_1^2 & . & . & . & x_1^n \\ . & . & . & & & & . \\ . & . & & . & & & . \\ . & . & & & . & & . \\ 1 & x_n & x_n^2 & . & . & . & x_n^n \end{bmatrix}$$

Prove that

$$\det(V_n) = \prod_{i<j}(x_j - x_i)$$

First we see that

$$\det(V_n) = \begin{vmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{vmatrix}$$

We take the transpose in order to row reduce since $\det(A) = \det(A^T)$,

$$= \det(V_n) = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_0^n & x_1^n & x_2^n & \dots & x_n^n \end{vmatrix}$$

By using row reduction we can eliminate all the values in the first column except the upper left entry by multiplying the $n-1$ row by x_0^n and subtracting this value from the nth row to get your new n^{th} row with the left-most column reduced to zero getting,

$$= \det(V_n) = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & (x_1 - x_0) & (x_2 - x_0) & \dots & (x_n - x_0) \\ 0 & x_1(x_1 - x_0) & x_2(x_2 - x_0) & \dots & x_n(x_n - x_0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_1^{n-1}(x_1 - x_0) & x_2^{n-1}(x_2 - x_0) & \dots & x_n^{n-1}(x_p - x_0) \end{vmatrix}$$

Since we are taking the determinant, we can take the determinant using the first column so we end up getting

$$= \det(V_n) = \begin{vmatrix} (x_1 - x_0) & (x_2 - x_0) & \dots & (x_n - x_0) \\ x_1(x_1 - x_0) & x_2(x_2 - x_0) & \dots & x_n(x_n - x_0) \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1}(x_1 - x_0) & x_2^{n-1}(x_2 - x_0) & \dots & x_n^{n-1}(x_p - x_0) \end{vmatrix}$$

transposing this matrix and factoring out from each row the common terms and retransposing we get

$$= \det(V_n) = (x_1 - x_0)(x_2 - x_0)\dots(x_n - x_0) \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}$$

Since the matrix we ended us with is similar to the one we started with, performing the same operations iteratively on this matrix will give us

$$= \det(V_n) = \prod_{j=1}^{n}(x_j - x_0)\prod_{j=2}^{n-1}(x_j - x_1) \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_2 & x_3 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_2^{n-2} & x_3^{n-2} & \dots & x_n^{n-2} \end{vmatrix}$$

and we will keep doing this iteratively until we have gone through all the elements of the matrix and we finally get

$$\det(V_n) = \prod_{i<j}(x_j - x_i)$$

as desired.　　　　　　　　　　　　　　　　　　　　　　　　　　　　■

Exercise 4.36

Find all possible Jordan forms for a transformation with characteristic polynomial $(x+5)^2(x-3)^2$.

With this characteristic polynomial we know that $\lambda_1 = 3$ and $\lambda_2 = -5$. These two eigenvalues both have algebraic multiplicity of 2, but we don't know what their geometric multiplicities are. Because we do not know the their geometric multiplicities there are four possibilities for the Jordan Form of these matrices:

The first possibility is that 3 has a geometric multiplicity of 1 and -5 has a geometric multiplicity of 1. Its Jordan Form would be:

$$\begin{bmatrix} -5 & 1 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

The second possibility is that 3 has a geometric multiplicity of 2 and -5 has a geometric multiplicity of 1. Its Jordan Form would be:

$$\begin{bmatrix} -5 & 1 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

The third possibility is that 3 has a geometric multiplicity of 1 and -5 has a geometric multiplicity of 2. Its Jordan Form would be:

$$\begin{bmatrix} -5 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

The fourth possibility is that 3 has a geometric multiplicity of 2 and -5 has a geometric multiplicity of 2. Its Jordan Form would be:

$$\begin{bmatrix} -5 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$