Policy

This is a takehome midterm. You take it home, do it for a week, and then hand it back. It is due Apr 20th. (So you have a total of 2 weeks.) This exam is hard, so please START EARLY.

I encourage collaborations on this midterm. However, you must obey the following rule:

1. You MUST each hand in your own work individially in your own words.
2. You MUST understand everything you wrote. (Say you copied your friend's WRONG answer without thinking, and that will most likely be in violation of this rule.)
3. You need to write down the names of your collaborator, if any.
4. Failure to comply rule 2 and rule 3 will be treated as plagiarism.
5. Collaboration with people not in this class (such as a math grad student) is not forbidden but not recommended. If you choose to, then write down their names as well.

We have a total of six problems and 11 points each, and a total of 66 points. Full credit is 60 points, and if you get more than 60 points, then your score is simply 60 points. (So you have some room for mistakes.)

The Midterm

Problem 1 (Pokémon Battle!). *You are in a Pokémon battle, and your rattata is fighting Ash's Pikachu. You will need to attack three times to fully defeat Pikachu. However, your rattatta is paralyzed. Each round, you have $\frac{3}{4}$ chance to attack successfully, and $\frac{1}{4}$ chance to be effected by the paralysis status and therefore unable to attack. Let us study this battle with linear algebra.*

1. *(2pt) Suppose after the battle go on for t rounds, the chance we still need 3, 2 or 1 successful attacks to win is p_3, p_2 and p_1. And the chance that you have already won is p_0. Find a matrix A such that $A^t \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} p_3 \\ p_2 \\ p_1 \\ p_0 \end{bmatrix}$.(A is called the Markov matrix of this dynamical system. It describes the evolution of the battle.)*
2. *(2pt) Find the Jordan canonical form of A and the minimal polynomial of A.*
3. *(2pt) Suppose we want to block diagonalize A into a 3×3 block and a 1×1 block, which Sylvester's equation do I need to solve? Find the equation and solve it.*
4. *(2pt) Find X such that $X^{-1}AX$ is in Jordan canonical form.*
5. *(3pt) Note that $\mathbf{e}_4^{\mathrm{T}}(A^t - A^{t-1})\mathbf{e}_1$ is the probability that Pikachu is defeated using exactly t rounds. (You do not need to prove this.) Therefore the expexted number of rounds needed to defeat Pikachu is $\mathbf{e}_4^{\mathrm{T}}(\sum_{t=1}^{\infty} tA^{t-1})(A-I)\mathbf{e}_1$. Power series calculation then shows that $\sum_{t=1}^{\infty} tx^{t-1} = -\frac{1}{(1-x)^2}$, which means the expected value should be something like $\mathbf{e}_4^{\mathrm{T}}(A-I)^{-1}\mathbf{e}_1$. However, $A - I$ is not invertible! What should we do? Well, note that $(A^t - A^{t-1})\mathbf{e}_1 = (A^t - A^{t-1})(\mathbf{e}_1 - \mathbf{e}_4)$, and A has an invariant subspace V containing $\mathbf{e}_1 - \mathbf{e}_4$ on which $A - I$ is invertible. For this problem, find this invariant subspace V, and compute $\mathbf{e}_4(A-I)^{-1}(\mathbf{e}_1 - \mathbf{e}_4)$ where $(A-I)^{-1}$ is interpreted as the inverse of $A - I$ restricted to the invariant subspace V. (This is the expected number of rounds we need to defeat pikachu.)*

Answer:

1. $A = \begin{bmatrix} \frac{1}{4} & & & \\ \frac{3}{4} & \frac{1}{4} & & \\ & \frac{3}{4} & \frac{1}{4} & \\ & & \frac{3}{4} & 1 \end{bmatrix}$.
2. The Jordan canonical form is $\begin{bmatrix} \frac{1}{4} & 1 & & \\ & \frac{1}{4} & 1 & \\ & & \frac{1}{4} & \\ & & & 1 \end{bmatrix}$. The minimal polynomial is $(x - \frac{1}{4})^3(x-1)$.
3. The Sylvester's equation is $X \begin{bmatrix} \frac{1}{4} & & \\ \frac{3}{4} & \frac{1}{4} & \\ & \frac{3}{4} & \frac{1}{4} \end{bmatrix} - \begin{bmatrix} 1 \end{bmatrix} X = \begin{bmatrix} 0 & 0 & \frac{3}{4} \end{bmatrix}$. The solution is $X = \begin{bmatrix} -1 & -1 & -1 \end{bmatrix}$.

However, note that I am using the convention $\begin{bmatrix} I & \\ X & I \end{bmatrix}\begin{bmatrix} A & \\ & B \end{bmatrix}\begin{bmatrix} I & \\ -X & I \end{bmatrix} = \begin{bmatrix} A & \\ C & B \end{bmatrix}$. If the student do it the other way around, then $-X$ will be the correct answer.

4. $X = \begin{bmatrix} 0 & 0 & 16 & 0 \\ 0 & 12 & 0 & 0 \\ 9 & 0 & 0 & 0 \\ -9 & -12 & -16 & 1 \end{bmatrix}$. The third column can be any $\boldsymbol{v}_3$ whose first coordinate is non-zero and all coordinates add up to 0. Then $\boldsymbol{v}_2 = (A - \frac{1}{4}I)\boldsymbol{v}_3$ and $\boldsymbol{v}_1 = (A - \frac{1}{4}I)\boldsymbol{v}_2$.

5. Using basis $X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}$, then $AX = XA'$ where $A' = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 \\ 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix}$. Now $(A' - I)^{-1} = -\frac{4}{3}\begin{bmatrix} 1 & & \\ 1 & 1 & \\ 1 & 1 & 1 \end{bmatrix}$. So $(A - I)^{-1}(\boldsymbol{e}_1 - \boldsymbol{e}_4) = (A - I)^{-1}X\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = X(A' - I)^{-1}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = -\frac{4}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} -\frac{4}{3} \\ -\frac{4}{3} \\ -\frac{4}{3} \\ 4 \end{bmatrix}$.

So the answer is $\boldsymbol{e}_4^{\mathrm{T}}\begin{bmatrix} -\frac{4}{3} \\ -\frac{4}{3} \\ -\frac{4}{3} \\ 4 \end{bmatrix} = 4$.

Problem 2 (Matrix Square Root)**.** *Given a matrix A, we would like to find all square roots of A, i.e., matrices B such that $B^2 = A$. (Some of the questions here concerns Hermitian matrices. You may use the spectral theorem of Hermitian matrices freely.)*

1. *(1pt) Suppose A is $n \times n$ and diagonalizable with distinct eigenvalues. How many square roots of A are there?*
2. *(2pt) Describe all square roots of the 2×2 identity matrix.*
3. *(2pt) Describe all square roots of $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.*
4. *(2pt) Suppose A is a Hermitian positive-definite matrix. (I.e., $A^* = A$ and all eigenvalues are positive.) Show that there is a UNIQUE Hermitian positive-definite matrix B such that $B^2 = A$. We shall write this B as $A^{\frac{1}{2}}$ from now on. (Note that there might be other non-Hermitian square roots. If you fail to see how, then you might be doing this problem wrong.)*
5. *(2pt) Suppose A is a Hermitian positive-definite matrix. Show that $(A^{\frac{1}{2}})^{-1} = (A^{-1})^{\frac{1}{2}}$.*
6. *(2pt) Find an example of Hermitian positive-definite matrices A, B such that $(A^{\frac{1}{2}}B^{\frac{1}{2}})^2 \neq AB$.*

Answer:

1. If B has repeated eigenvalue, then $A = B^2$ must also have repeated eigenvalue. Since A has distinct eigenvalue, this implies that B has distinct eigenvalue, and therefore B is diagonalizable. Once diagonalized, each eigenvalue of B must corresponds to the square root of the eigenvalue of A in the corresponding location.

   Now each non-zero eigenvalue has two square roots, while each zero eigenvalue has only one square root. So if the matrix is invertible, the answer is 2^n. If the matrix is not invertible, since there are at most one zero-eigenvalue, the answer is 2^{n-1}.

2. $\pm I$, and all matrices with eigenvalue $1, -1$, i.e., trace zero and determinant -1, i.e., matrices of the form $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ such that $a^2 + bc = 1$.

3. Note that if $B^2 = A$, then B must also be nilpotent, and its JCF must have a single Jordan block. So it is easy to see that $\mathrm{Ker}(B) = \mathrm{Ran}(B^2) = \mathrm{Ran}(A)$ is spanned by $\boldsymbol{e}_1$, and $\mathrm{Ker}(B^2) = \mathrm{Ker}(A)$ is spanned by $\boldsymbol{e}_1, \boldsymbol{e}_3$. So $B\boldsymbol{e}_1 = \mathbf{0}$, $B\boldsymbol{e}_3 \in \mathrm{Ker}(B)$ is a multiple of $\boldsymbol{e}_1$, and $B\boldsymbol{e}_2 \in \mathrm{Ker}(B^2)$ is a linear combination of $\boldsymbol{e}_1, \boldsymbol{e}_3$. So $B = \begin{bmatrix} 0 & a & b \\ 0 & 0 & 0 \\ 0 & c & 0 \end{bmatrix}$ such that $bc = 1$.

4. Since A is diagonalizable, its eigenspaces form an invariant decomposition of the domain. So change of basis gives $A = \begin{bmatrix} A_1 & & \\ & \dots & A_k \end{bmatrix}$ where each A_i is the restriction of A to an eigenspace. Since $AB = BA$, and each eigenspace of A is also B-invariant, this implies that $B = \begin{bmatrix} B_1 & & \\ & \dots & B_k \end{bmatrix}$. So we have $B_i^2 = A_i$. So we only need to do this block by block. So WLOG, as may assume that $A = \lambda I$. If $B^2 = A$, then all eigenvalues of B must be square root of λ. Since B is required to be positive definite, all eigenvalues of B must be $\sqrt{\lambda}$. Finally, since B is Hermitian, it is also diagonalizable, so we must have $B = \sqrt{\lambda} I$.

5. $(A^{\frac{1}{2}})^{-1}$ is easily verified to be Hermitian and positive-definite. Now $((A^{\frac{1}{2}})^{-1})^2 = ((A^{\frac{1}{2}})^2)^{-1} = A^{-1}$. So indeed $(A^{\frac{1}{2}})^{-1} = (A^{-1})^{\frac{1}{2}}$.

6. Note that if $(A^{\frac{1}{2}}B^{\frac{1}{2}})^2 = AB$, then $A^{\frac{1}{2}}B^{\frac{1}{2}}A^{\frac{1}{2}}B^{\frac{1}{2}} = A^{\frac{1}{2}}A^{\frac{1}{2}}B^{\frac{1}{2}}B^{\frac{1}{2}}$, and thus $A^{\frac{1}{2}}B^{\frac{1}{2}} = B^{\frac{1}{2}}A^{\frac{1}{2}}$. This implies that $AB = BA$. So pick any Hermitian positive definite A, B such that $AB \neq BA$ will do. Say $A = \begin{bmatrix} 1 & \\ & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$.

Problem 3 (Geometric Mean). *Given two positive real numbers a, b, their arithmetic mean is $\frac{a+b}{2}$, while their geometric mean is $\sqrt{ab}$. A famous inequality is $\sqrt{ab} \leq \frac{a+b}{2}$.*

We now attempt to derive a similar idea for matrices. Given two positive definite Hermitian matrices A, B (so that their eigenvalues are positive real numbers), their arithmetic mean is easily $\frac{1}{2}(A+B)$. But what should be their geometric mean? In last problem, you have see that $A^{\frac{1}{2}}B^{\frac{1}{2}}$ does NOT work, because it might not even be a square root of AB. So we need to do something else.

Given two positive definite Hermitian matrices A, B, let us define the matrix $A\#B$ as a positive definite Hermitian matrix X such that $XB^{-1}X = A$. We shall see that such X is unique.

(BTW, the idea of geometric mean of matrices have some interesting applications. For example, in computational graphics, given two motions described by linear transformations A, B, how to find a "mid-point" between the motion? Then $A\#B$ would be a good candidate.)

1. *(2pt) Show that if $XB^{-1}X = A$, then we must have $(B^{-\frac{1}{2}}XB^{-\frac{1}{2}})^2 = B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$, and show that $X = B^{\frac{1}{2}}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})^{\frac{1}{2}}B^{\frac{1}{2}}$. (Here $B^{-\frac{1}{2}}$ means the inverse of $B^{\frac{1}{2}}$.) In particular, we see that such positive definite Hermitian X is uniquely determined.*

2. *(2pt) Show that $A\#A = A$ and $(A\#B)^{-1} = A^{-1}\#B^{-1}$.*

3. *(2pt) Show that $A\#B = B\#A$.*

4. *(2pt) Given two Hermitian matrices A, B, we say $A \geq B$ if $A - B$ is positive semidefinite (i.e., all eigenvalues are non-negative). Show that $A\#B \leq \frac{1}{2}(A+B)$. So indeed the geometric mean is at most the algebraic mean. (Hint: Note that the square of any Hermitian matrix must be positive semidefinite, so $((B^{-\frac{1}{2}}AB^{-\frac{1}{2}})^{\frac{1}{2}} + I)^2 \geq 0$. Also recall that M and XMX^* have the same definite-ness for invertible X.)*

5. *(3pt) Show that if* $\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq 0$ *for some Hermitian* X*, then* $X \leq A\#B$*. (So* $A\#B$ *is the largest Hermitian matrix with such a property. Hint: Block LDU.)*

Answer:

1. Direct computation.

2. Direct plug-in to the formula.

3. $XB^{-1}X = A$ iff $B^{-1} = X^{-1}AX^{-1}$ iff $B = XA^{-1}X$.

4. Since $((B^{-\frac{1}{2}}AB^{-\frac{1}{2}})^{\frac{1}{2}} + I)^2 \geq 0$, therefore $B^{\frac{1}{2}}((B^{-\frac{1}{2}}AB^{-\frac{1}{2}})^{\frac{1}{2}} + I)^2 B^{\frac{1}{2}} \geq 0$, and therefore expansion gives $A + 2(A\#B) + B \geq 0$.

5. Block LDU shows that $\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq 0$ if and only if $B - XA^{-1}X \geq 0$. So $(A^{-\frac{1}{2}}XA^{-\frac{1}{2}})^2 \leq A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$. Let us denote $|A^{-\frac{1}{2}}XA^{-\frac{1}{2}}|$ as $((A^{-\frac{1}{2}}XA^{-\frac{1}{2}})^2)^{\frac{1}{2}}$, then it is easy to see that $|A^{-\frac{1}{2}}XA^{-\frac{1}{2}}|$ is simply obtained by taking absolute value on all eigenvalues of $A^{-\frac{1}{2}}XA^{-\frac{1}{2}}$. So we have $|A^{-\frac{1}{2}}XA^{-\frac{1}{2}}| \geq A^{-\frac{1}{2}}XA^{-\frac{1}{2}}$. Since we have $|A^{-\frac{1}{2}}XA^{-\frac{1}{2}}|^2 \leq A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$, I claim that this implies that $|A^{-\frac{1}{2}}XA^{-\frac{1}{2}}| \leq (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\frac{1}{2}}$, which would then gives us the desired result.

   So we need to show that for positive semi-definite X and positive definite Y, if $Y^2 \geq X^2$, then $Y \geq X$. To see this, $Y^2 \geq X^2$ imples that $I \geq Y^{-1}XXY^{-1}$, so all eigenvalues of $Y^{-1}XXY^{-1}$ are at most 1. So the operator norm of XY^{-1} is at most 1. But this implies that all eigenvalues of XY^{-1} are at most one. Hence all eigenvalues of $Y^{-\frac{1}{2}}XY^{-\frac{1}{2}} = Y^{-\frac{1}{2}}(XY^{-1})Y^{\frac{1}{2}}$ are at most one as well. So $I \geq Y^{-\frac{1}{2}}XY^{-\frac{1}{2}}$, which means $Y \geq X$.

Problem 4 (Motion under weird force field)**.** *Suppose a mass point with mass* $m = 1$ *is moving inside* $\mathbb{R}^2$*, under the effect of some weird force field. Let* $\boldsymbol{r}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$ *be the location of the mass point at time* t*. Suppose the force field applies a force to the mass point depending on the location of the mass point, say,* $\boldsymbol{F}(t) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{r}(t)$*, where* $\boldsymbol{F}(t)$ *is the force acting on the mass point at time* t*. We would like to solve this differential equation.*

1. *(3pt) Let* $\boldsymbol{v}(t) = \boldsymbol{r}'(t)$ *be the velocity at time* t*. Then we can think of the status of the mass point as a vector* $\begin{bmatrix} \boldsymbol{r}(t) \\ \boldsymbol{v}(t) \end{bmatrix}$*. Find a constant matrix* A *such that* $\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \boldsymbol{r}(t) \\ \boldsymbol{v}(t) \end{bmatrix} = A \begin{bmatrix} \boldsymbol{r}(t) \\ \boldsymbol{v}(t) \end{bmatrix}$.

2. *(3pt) Find* X, J *such that* $A = XJX^{-1}$ *is the Jordan decomposition for* A*.*

3. *(2pt) For fixed* t*, find a polynomial* $p(x)$ *of degree* 3 *such that* $p(t) = \mathrm{e}^t, p'(t) = \mathrm{e}^t, p(-t) = \mathrm{e}^{-t}, p'(-t) = \mathrm{e}^{-t}$*.*

4. *(3pt) Say* $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ *in the last sub-problem. Calculate the entries of* e^{At} *involving the constants* a_0, a_1, a_2, a_3*. (We use* a_0, a_1, a_2, a_3 *so that your answer is not absurdly long and complicated. We do not require you to solve this differential equation, but at this stage, you should know how to obtain a formula for the solution to this differential equation.)*

Answer:

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

2. $X = \begin{bmatrix} 2 & 1 & 2 & -1 \\ 0 & -4 & 0 & 4 \\ -2 & 1 & 2 & 1 \\ 0 & 4 & 0 & 4 \end{bmatrix}$ and $J = \begin{bmatrix} -1 & 1 & & \\ & -1 & & \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}$.

3. The coefficient for x^3 is $\frac{t-1}{4t^3}\mathrm{e}^t + \frac{t+1}{4t^3}\mathrm{e}^{-t}$. The coefficient for x^2 is $\frac{1}{4t}\mathrm{e}^t - \frac{1}{4t}\mathrm{e}^{-t}$. The coefficient for x is $\frac{-t+3}{4t}\mathrm{e}^t + \frac{-t-3}{4t}\mathrm{e}^{-t}$. Finally, the constant term is $\frac{2-t}{4}\mathrm{e}^t + \frac{2+t}{4}\mathrm{e}^{-t}$.

4. $\mathrm{e}^{tA} = p(tA)$ since e^x and $p(x)$ takes the same value and same derivative at eigenvalues of tA. So
$$\mathrm{e}^{At} = a_0 I + a_1 At + a_2 A^2t^2 + a_3 A^3 t^3 = \begin{bmatrix} a_0 + a_2t^2 & a_2t^2 & a_1t + a_3t^3 & a_3t^3 \\ 0 & a_0 + a_2t^2 & 0 & a_1t + a_3t^3 \\ a_1t + a_3t^3 & a_1t + 2a_3t^3 & a_0 + a_2t^2 & a_2t^2 \\ 0 & a_1t + a_3t^3 & 0 & a_0 + a_2t^2 \end{bmatrix}.$$

Problem 5 (Generalized Cayley-Hamilton). *Given any square matrix A, if $p_A(x) = \det(xI - A)$, then we know $p_A(A) = 0$. This is the Cayley-Hamilton theorem. However, in a letter to Sylvester, Cayley claimed that he in fact know a more generalized version of this result. Suppose we have two square matrices A, B such that $AB = BA$. Let us define a function $f(x, y) = \det(xA - yB)$ (so the inputs are two complex numbers, and the output is a complex number). The generalized Cayley-Hamilton claim that $f(B, A) = 0$. As you can see, the original Cayley-Hamilton is the special case when $A = I$.*

Alas, Cayley did not really provide any proof in that letter. So let us prove it ourselves.

1. *(2pt) An absent-minded student Yilong Yang used the following argument: If we plug in $x = B, y = A$ directly into the definition of f, then we have $f(B, A) = \det(BA - AB)$. And since $AB = BA$, we see that $f(B, A) = 0$. Unfortunately, this argument is wrong. Could you describe what's wrong with it?*

2. *(2pt) Verify that $f(B, A) = 0$ when $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$.*

3. *(3pt) If A is invertible, show that $f(x, y)$ is a multiple of $y^n p_C(y^{-1}x)$, where A, B are $n \times n$ matrices, $C = A^{-1}B$, and p_C is the characteristic polynomial of C.*

4. *(2pt) If A is invertible, prove that $f(B, A) = 0$.*

5. *(2pt) Show that there is a sequence of invertible matrices A_t such that $A = \lim A_t$, and $A_tB = BA_t$ for all A_t. (We do not require you to fill out the rest of the proof of the generalized Cayley-Hamilton, but you should be able to imagine what the proof looks like.)*

Answer:

1. $f(B, A)$ is supposed to be a matrix, but the argument by Yilong have $f(B, A) = \det(BA - AB)$ a scalar, which is wrong.

2. $f(x, y) = \det \begin{bmatrix} -2y & x - 3y \\ & -2y \end{bmatrix} = 4y^2$. And thus $f(B, A) = 4A^2 = 0$.

3. $f(x, y) = \det(xA - yB) = \det(A)\det(xI - yC) = \det(A)y^n \det(y^{-1}xI - C)$, which is the desired format.

4. Since $f(x, y) = \det(A)y^n p_C(y^{-1}x)$, therefore $f(B, A) = \det(A)A^n p_C(C)$. But $p_C(C) = 0$ by the regular Cayley-Hamilton theorem.

5. Simply set $A_t = A + tI$, then we are done for t with small enough absolute value.

Problem 6 (Newton's Method). *As we have seen in class,* $\operatorname{sign}(X)$ *is useful to solve certain Sylvester's equations. Here we aim to find a way an approximation to* $\operatorname{sign}(X)$*. Given a matrix A with no purely imaginary eigenvalue, set $X_0 = A$, and set $X_{n+1} = \frac{1}{2}(X_n + X_n^{-1})$. (As a side note, a complex number z is purely imaginary if its real part is zero. In particular, 0 is a purely imaginary number as well. So, if a matrix has no purely imaginary eigenvalue, then it is invertible.)*

1. *(2pt) Show that if X_n has no purely imaginary eigenvalue, then X_{n+1} has no purely imaginary eigenvalue. (So our inductive definition makes sense.)*

2. *(3pt) If A is 1×1, and it is not purely imaginary, show that X_n indeed coverge to* $\operatorname{sign}(A)$. *(This question has little to do with linear algebra....) (Hint: $\frac{f(x)-1}{f(x)+1} = (\frac{x-1}{x+1})^2$ where $f(x) = \frac{1}{2}(x + \frac{1}{x})$.)*

3. *(3pt) If A is diagonalizable and has no purely imaginary eigenvalue, show that X_n indeed coverge to* $\operatorname{sign}(A)$. *(Not part of this problem. But diagonalizable matrices are dense, so you can imagine that this is true in general.)*

4. *(3pt) Suppose A is an $n \times n$ Jordan block with eigenvalue* 1. *Show that* $X_{n-1} = \operatorname{sign}(A)$.

Answer:

1. Set $f(x) = \frac{1}{2}(x + \frac{1}{x})$. If X_n has eigenvalues $\lambda_1, \dots, \lambda_n$, then $X_{n+1} = f(X)$ will have eigenvalues $f(\lambda_1), \dots, f(\lambda_n)$. Hence we only need to show that if x is not purely imaginary, then $f(x)$ is not purely imaginary. Suppose $x = a + bi$ for real a, b. Then $f(x)$ has real part $\frac{1}{2}(a + \frac{a}{a^2+b^2}) = \frac{1}{2}a(1 + \frac{1}{a^2+b^2})$. So if $a \neq 0$, since $1 + \frac{1}{a^2+b^2} > 0$, $f(x)$ must also have non-zero real part. Done.

2. Note that $\frac{X_{n+1}-1}{X_{n+1}+1} = (\frac{X_n-1}{X_n+1})^2$. If X_n has positive real part, then $|\frac{X_n-1}{X_n+1}| < 1$, and hence its limit as $n \to \infty$ is zero. Hence $\lim X_n = 1$.

   If X_n has negative real part, consider $\frac{X_n+1}{X_n-1}$ instead, and by similar logic, we see that $\lim X_n = -1$. Either way we are done.

3. Just diagonalize $A = XDX^{-1}$, and apply $f(x)$ to each diagonal entry of D.

4. If $A = I + aN^k + N^{k+1}p(N)$ for some polynomial $p(x)$, then $A^{-1} = I - aN^k + N^{k+1}q(N)$ for some polynomial $q(x)$. So $f(A) = I + N^{k+1}f(N)$ for some polynomial $f(x)$. So repeating this, we have $X_{n-1} = I + N^n f(N) = I$ for some polynomial $f(x)$.