Policy

This is a takehome midterm. You take it home, do it for a week, and then hand it back. It is due Apr 22th. (So you have a total of 2 weeks.) 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 (Quaternions)**.** *The quaternions are widely applied in computation graphics and computer games, and can also be used to simplify calculations that involves angular stuff or rotations. They are defiend as $a+bi+cj+dk$, where $a,b,c,d \in \mathbb{R}$ and $i^2 = j^2 = k^2 = ijk = -1$. (In particular, the multiplication is defined as $(a+bi+cj+dk)(a'+b'i+c'j+d'k) = (aa'-bb'-cc'-dd') + (ab'+ba'+cd'-dc')i + (ac'+ca'+db'-bd')j + (ad'+da'+bc'-cb')k$. You don't really need to know this formula for this problem though.)*

1. *Recall that we can formulate complex numbers as matrices. Similarly, let us try this for quaternions. Consider matrices of the form*
$$a\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}+b\begin{bmatrix}0&-1&0&0\\1&0&0&0\\0&0&0&-1\\0&0&1&0\end{bmatrix}+c\begin{bmatrix}0&0&-1&0\\0&0&0&1\\1&0&0&0\\0&-1&0&0\end{bmatrix}+d\begin{bmatrix}0&0&0&-1\\0&0&-1&0\\0&1&0&0\\1&0&0&0\end{bmatrix} \text{ where } a,b,c,d \in \mathbb{R}.$$
Show that this gives a model of the quaternions as well, i.e., they satisfy $i^2 = j^2 = k^2 = ijk = -1$.

2. *You may interpret $\begin{bmatrix}a\\b\\c\\d\end{bmatrix} \in \mathbb{R}^4$ as the quaternion $a+bi+cj+dk$. Then for any quaternion q, multiplying q to $a+bi+cj+dk$ from the left gives a linear map $\mathbb{R}^4 \to \mathbb{R}^4 : (a+bi+cj+dk) \mapsto q(a+bi+cj+dk)$. If $q = r+xi+yj+zk$, what is the matrix L_q for this linear map? What is the matrix R_q for the linear map if we multiply q from the right? Do we have $L_qR_q = R_qL_q$?*

3. *The conjugate of a quaternion $q = a+bi+cj+dk$ is $\overline{q} = a-bi-cj-dk$. Show that for a unit quaternion q (i.e., $q\overline{q} = 1$), the matrix $L_qR_{\overline{q}}$ has block form $\begin{bmatrix}1&0\\0&Q\end{bmatrix}$ where Q is an orthogonal matrix. (Hint: The block form is easy. To see that Q is orthogonal, an easy way is to show that $(L_qR_{\overline{q}})^T(L_qR_{\overline{q}}) = I$ by understanding the meaning of the matrices involved.) (In particular, if you interpret a 3D vector (x,y,z) as the quaternion $v = 0+xi+yj+zk$, then the quaternion multiplication $qv\overline{q}$ corresponds to some rotation of v. This quaternion interpretation is currently one of the best way to compute 3D rotations in real life. Many of your games with 3D graphics depend on this.)*

Problem 2 (Drazin Inverse and Differential Equation)**.** *Given an unknown vector-valued function $\boldsymbol{v}(t)$, we know how to solve $\boldsymbol{v}' = A\boldsymbol{v}$ for a constant matrix A. But what if we have $A\boldsymbol{v}' + B\boldsymbol{v} = \mathbf{0}$ for constant matrices? If A is invertible, we can reorganize this into $\boldsymbol{v}' = -A^{-1}B\boldsymbol{v}$ and solve it easily. But what if A is not invertible?*

Here we introduce the Drazin inverse of a matrix. Recall that for any matrix A, according to the ultimate decomposition, we have $A = X\begin{bmatrix}A_R&\\&A_N\end{bmatrix}X^{-1}$ where A_R is invertible and A_N is nilpotent. Then we define $A^{(\mathrm{D})} = X\begin{bmatrix}A_R^{-1}&\\&0\end{bmatrix}X^{-1}$ as the Drazin inverse of A.

1. *Show that if $\begin{bmatrix}R&\\&N\end{bmatrix} = X\begin{bmatrix}R'&\\&N'\end{bmatrix}X^{-1}$ where R,R' are invertible and N,N' are nilpotent, then*

$\begin{bmatrix} R^{-1} & \\ & 0 \end{bmatrix} = X \begin{bmatrix} (R')^{-1} & \\ & 0 \end{bmatrix} X^{-1}$. *(This shows that the Drazin inverse of A is unique and does not depends on the choice of invertible-nilpotent decomposition $A = X \begin{bmatrix} A_R & \\ & A_N \end{bmatrix} X^{-1}$.)*

2. *Show that $AA^{(\mathrm{D})} = A^{(\mathrm{D})}A, A^{(\mathrm{D})}AA^{(\mathrm{D})} = A^{(\mathrm{D})}$, and $A^{(\mathrm{D})}A^{k+1} = A^k$ where k is the smallest integer such that* $\mathrm{Ker}(A^k) = \mathrm{Ker}(A^{k+1})$.

3. *Calculate $(\boldsymbol{a}\boldsymbol{b}^*)^{(\mathrm{D})}$ for non-zero vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{C}^n$. (Hint: Harder if you use brute force calculation with the definition. Easier if you can guess it out right, and prove that you have the right guess.)*

4. *For fixed A, show that we can find a polynomial $p(x)$ such that $A^{(\mathrm{D})} = p(A)$. (Again, note that for different A this polynomial may be different.) (Hint: Suppose $A_R^{-1} = q(A_R)$ for a polynomial $q(x)$.)*

5. *If $AB = BA$, show that $\mathrm{e}^{-A^{(\mathrm{D})}Bt}AA^{(\mathrm{D})}\boldsymbol{v}_0$ is a solution to $A\boldsymbol{v}' + B\boldsymbol{v} = \mathbf{0}$ for any constant vector $\boldsymbol{v}_0$. (This is one of the results in a paper by Campbell, Meyer and Rose in 1967.)*

Problem 3 (Sherman-Morrison-Woodburry Formula)**.** *The famous Sherman-Morrison-Woodburry formula states that, for any $m \times m$ invertible matrix X, $n \times n$ invertible matrix Y, $m \times n$ matrix A and $n \times m$ matrix B, we have $(X - AYB)^{-1} = X^{-1} + X^{-1}A(Y^{-1} - BX^{-1}A)^{-1}BX^{-1}$. This can be proven using block eliminations on $\begin{bmatrix} X & -A \\ B & Y \end{bmatrix}$. Well, I always find that proof annoying. So let us not do that, and try to find some alternative proofs. To simplify, you can easily see that it is enough to establish the special case $(I_m - AB)^{-1} = I_m + A(I_n - BA)^{-1}B$.*

Consider the function $f(x) = (1-x)^{-1}$. This function has a Taylor expansion $f(x) = 1 + x + x^2 + \dots$ for all $|x| < 1$ over the complex numbers. (This is also the sum of the geometric series, which you should have learned about in high school.)

1. *For any $m \times n$ matrix A and $n \times n$ matrix B, suppose $I_m - AB$ is invertible and all eigenvalues of AB have absolute value less than 1. Write $(I_m - AB)^{-1}$ as the sum of a series of matrices.*

2. *Using above idea, deduce the formula $(I_m - AB)^{-1} = I_m + A(I_n - BA)^{-1}B$, when $I_m - AB$ and $I_n - BA$ are invertible and all eigenvalues of AB and BA have absolute value less than 1. (Btw, note that AB and BA always have the same non-zero eigenvalues. You don't need this fact though.)*

3. *Oops, unfortunately, the method above does not always work. It has many annoying requirements on eigenvalues. Let us now forget about Taylor expansion. In general, prove that $Ap(BA) = p(AB)A$ for all polynomials $p(x)$. (Hint: try $p(x) = x$ first.)*

4. *Show that for any function $g(x)$ and any two square matrix X, Y, we can find a polynomial $p(x)$ such that $p(X) = g(X)$ and $p(Y) = g(Y)$ simultaneously. (Hint: block matrix.)*

5. *Show that $Af(BA) = f(AB)A$ as long as $f(AB)$ and $f(BA)$ are defined.*

6. *Verify that $f(AB) = I_m + Af(BA)B$, using the identity above.*

Problem 4 (Equations of matrices)**.** *Let N be the $n \times n$ nilpotent Jordan block.*

1. *Show that the solutions to the Sylvester's equation $NX - XN = 0$ are exactly the polynomials of N.*

2. *Suppose $Y = \mathrm{e}^N$. Show that $Y, Y - I, (Y - I)^2, \dots, (Y - I)^{n-1}$ are linearly independent in the space of matrices, and they span the space of matrices made of polynomials of N. (Consequently, N is a polynomial of Y.)*

3. *Find all solutions X to the matrix equation $\mathrm{e}^X = \mathrm{e}^N$. (Make sure to consider COMPLEX matrices X.)*

4. *Find real matrices A, B such that $AB \neq BA$ but $\mathrm{e}^A = \mathrm{e}^B$. (Hint: for complex 1×1 matrices, try to find $x \neq y \in \mathbb{C}$ such that $\mathrm{e}^x = \mathrm{e}^y$.)*

5. *Prove that there is no solution X to the equation $\sin(X) = \begin{bmatrix} 1 & 1996 \\ 0 & 1 \end{bmatrix}$. (This is a Putnam competition problem. I'm sure you know the year of the competition....)*

Problem 5 (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. *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. *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. *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. *Suppose A is an $n \times n$ Jordan block with eigenvalue 1. Show that $X_{n-1} = I$.*

Problem 6 (Real Möbius transformations)**.** *A Möbius transformation is a function $f : x \mapsto \frac{ax+b}{cx+d}$. For example, scaling $x \mapsto 2x$, addition $x \mapsto x + 1$, inversion $x \mapsto \frac{1}{x}$ are all Mobius transformations. We may further define $f(\infty) = \frac{a}{c}$ and $f(-\frac{d}{c}) = \infty$, as you can tell by taking limits. So Möbius transformations are actiong on the space $\mathbb{R} \cup \{\infty\}$. In some sense, you are imagining the real line where ∞ and $-\infty$ are treated as the same thing (i.e., they are glued together), and you see that $\mathbb{R} \cup \{\infty\}$ is in fact a big circle.*

(You can skip these materials in this parenthesis. They are irrelavant but maybe of interest: geometrically, consider the circle C on the xy-plane with center $(0, 1/2)$ and radius 1/2. Then imagine that the point $(0, 1)$ on the circle has a ray gun attached. For each point of the circle, the ray gun can shoot that point, and then go through it and intersect with the x-axis somewhere. This would give a 1-to-1 correspondence betwee the real line $\mathbb{R}$ and $C - \{(0, 1)\}$. Now think of ∞ as the point $(0, 1)$, and you have thus build a 1-to-1 correspondence bwteen $\mathbb{R} \cup \{\infty\}$ and the circle C. The Möbius transformations act on $\mathbb{R} \cup \{\infty\}$, so under this correspondence you can imagine them as acting on the circle C. They usually stretch some portions of C while shrinking some other portions. This corresponds to looking at this circle from various perspectives. For example, if you look at the circle from the left, then since the left portion of the circle is closer to your eyes, it appears to be larger, while the right portion of the circle appears to be smaller as it is further away from your eyes. For a more concrete example, consider $x \mapsto 2x$. This means you are moving your perspective to be closer to $(0, 0)$ on the circle, so that points near $(0, 0)$ are more spread out. Everyone is now repelled away from the origin and attracted towards the infinity point $(0, 1)$.)

*For any 2 by 2 **invertible** real matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we can build a corresponding Möbius transformation $f_A : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ such that $f_A(x) = \frac{ax+b}{cx+d}$. Obviously all Möbius transformation arises this way.*

1. *Show that $f_A \circ f_B = f_{AB}$ for any A, B. In particular, if we simply write A to represent the function f_A, then the composition of A and B is the same as the matrix multiplication of A and B. (You may also check $f_{A^{-1}}$ is the inverse function of f_A and so on. This is not required though.)*

2. *Show that for any $k \in \mathbb{R} - \{0\}$, $f_A = f_{kA}$. And conversely, if $f_A = f_B$, then $A = kB$ for some constant k. (So WLOG, to study a Möbius transformation f_A, you can always scale A appropriately and assume that $\det(A) = 1$.)*

3. *Interpret a vector* $\begin{bmatrix} x \\ y \end{bmatrix}$ *as a ratio* $\frac{x}{y}$*. Then show that, under this interpretation,* $A \begin{bmatrix} x \\ y \end{bmatrix}$ *is interpreted exactly as the ratio* $f_A(\frac{x}{y})$*. (As a result, some literature write* A *and* f_A *interchangably and* $\begin{bmatrix} x \\ y \end{bmatrix}$ *and* $\frac{x}{y}$ *interchangably.)*
4. *Show that there are only four kinds of Möbius transformations. There is a kind where* f_A *has two fixed points in* $\mathbb{R} \cup \{\infty\}$ *(a typical example is* $x \mapsto 2x$ *where* 0 *and* ∞ *are the fixed points), a kind where* f_A *has only one fixed point in* $\mathbb{R} \cup \{\infty\}$ *(a typical example is* $x \mapsto x + 1$ *where* ∞ *is the only fixed point), and a kind without any fixed point in* $\mathbb{R} \cup \{\infty\}$ *(a typical example is* $x \mapsto -\frac{1}{x}$*). Finally, there is a kind where everyone is fixed, i.e., the identity function* $x \mapsto x$*.*