HW Policy

I encourage collaborations on homeworks, projects and even the takehome 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.
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.

Contents

1 HW1 (Due 3.2) Stuff last semester

Exercise 1.1 (The geometric meaning of matrices)**.** *Let S be the unit circle in $\mathbb{R}^2$. Let A be a 2 by 2 matrix, which would represent some linear map in the standard basis. Then $A(S)$ should be an ellipse.*

1. *Say $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Find the eigenvalues of A and the lengths of the semi-major axis and semi-minor axis of the ellipse $A(S)$.*
2. *Say $A = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$. Find the eigenvalues of A. Are they the lengths of the semi-major axis and semi-minor axis of the ellipse $A(S)$? (Draw it.)*

Exercise 1.2 (Block matrix manipulation)**.** *We say a square matrix M is positive definite if it is symmetric, and $\boldsymbol{v}^{\mathrm{T}} M \boldsymbol{v} > 0$ for all $\boldsymbol{v} \neq \mathbf{0}$.*

Suppose $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ where A, D are square and A is invertible. We also assume that all entries are real. Show that M is positive definite if and only if $B^T = C$ and both A and $D - CA^{-1}B$ are positive definite. (Hint: Perform $M = S\Sigma S^T$ where Σ is block diagonal.)

Exercise 1.3 (Subspace Geometry)**.** *Pick any matrix $A \in M_n(\mathbb{R})$. We use $\operatorname{Ker}(A)$ to denote the kernel (or zero space) of A.*

1. *Suppose $A^5 = 0$, show that $\dim \operatorname{Ker}(A)$ is at least $\frac{n}{5}$. Give an example of such A with $\dim \operatorname{Ker}(A) = \frac{n}{5}$ when n is a multiple of 5. (Hint: Try to show $\dim \operatorname{Ker}(AB) \leq \dim \operatorname{Ker}(A) + \dim \operatorname{Ker}(B)$. If you are hardcore, try to show $\dim \operatorname{Ker}(AB) = \dim(\operatorname{Ker}(A) \cap \operatorname{Ran}(B)) + \dim \operatorname{Ker}(B)$)*
2. *Suppose $AA^T AAA^T = 0$. Show that $\dim \operatorname{Ker}(A)$ is at least $\frac{n}{2}$. Give an example of such A with $\dim \operatorname{Ker}(A) = \frac{n}{2}$ when n is a multiple of 2. (Hint: Fundamental theorem of linear algebra)*

Exercise 1.4 (Spectral theorem)**.** *Let A be a real symmetric matrix. We define the matrix $R_t = (A - tI)^{-1}$ whenever t is NOT an eigenvalue of A. Show that for each eigenvalue λ of A, the matrix $\lim_{t \to \lambda} (\lambda - t) R_t$ is the orthogonal projection to the eigenspace of A for λ. (Here when you take the limit for matrices, you can just take the limit of each entry.)*

Exercise 1.5 (Dynamic system)**.** *Give a sequence $a_{n+1} = 3a_n + 2^n + n^2$, find a 5×5 matrix A, such that*

$$A \begin{bmatrix} a_n \\ 2^n \\ n^2 \\ n \\ 1 \end{bmatrix} = \begin{bmatrix} a_{n+1} \\ 2^{n+1} \\ (n+1)^2 \\ n+1 \\ 1 \end{bmatrix}.$$

2 HW2 (Due 3.9) Complex Structures; Subspaces

Exercise 2.1. *We would like to find a real $n \times n$ matrix A such that $A^2 = -I$.*

1. *For each even number n, find a real solution.*
2. *If odd n, show that there is no real solution.*

Exercise 2.2. *Suppose $A^2 = -I$ for a real $n \times n$ matrix A. For each vector $\boldsymbol{v} \in \mathbb{R}^n$, we define $\mathrm{i}\boldsymbol{v}$ to mean $A\boldsymbol{v}$. For any $n \times n$ matrix B, we say it is complex linear if $B(k\boldsymbol{v}) = kB\boldsymbol{v}$ for any complex number $k \in \mathbb{C}$.*

1. *Show that B is complex linear if and only if $AB = BA$.*
2. *If X also satisfies $X^2 = -I$, then must X be complex linear? Prove or provide a counter example.*
3. *For $n = 2$, pick any A such that $A^2 = -I$, and pick two distinct C such that $CA = -AC$ and $C^2 = I$.*
4. *(Read only) Consider an $n \times n$ real matrix C such that $CA = -AC$ and $C^2 = I$. This C is called a **complex conjugate operator**. Then such C must be diagonalizable, must have only eigenvalues 1 and -1, and its eigenspaces for 1 and -1 have the same dimension. The eigenspace for 1 is the space of "real vectors" while the eigenspace for -1 is the space of "imaginary vectors". As you can see, the "real part" and "imaginary part" of a vector is NOT defined by the complex structure A alone. In particular, for abstract arguments, it might be a good idea to AVOID arguments that split complex things into real parts and imaginary parts.*

Exercise 2.3 (adapted from Gilbert Strang 9.3.11-15)**.** *Take the permutation matrix P that sends $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ to $\begin{bmatrix} b \\ c \\ d \\ a \end{bmatrix}$. Let F_4 be the 4×4 Fourier matrix.*

1. *Compute $P\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $P\begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}$.*
2. *Show that $PF_4 = F_4 D$ for some diagonal matrix D. Find all eigenvalues and eigenvectors of P.*
3. *Let $C = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix}$. Compute $C\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $C\begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}$.*
4. *Write C as a polynomial of P. Find the eigenvalues and eigenvectors of C.*
5. *(Read Only) Matrices such as C are called circulant matrices. Just like periodic functions can be simplified using Fourier series, circulant matrices can be simplified using Fourier matrices. One can compute $C\boldsymbol{v}$ using fast Fourier transform, and it will be slightly faster than computing $C\boldsymbol{v}$ directly.*

Exercise 2.4. *Prove or find counter examples.*

1. *For four subspaces, if any three of them are linearly independent, then the four subspaces are linearly independent.*

2. *If subspaces V_1, V_2 are linearly independent, and V_1, V_3, V_4 are linearly independent, and V_2, V_3, V_4 are linearly independent, then all four subspaces are linearly independent.*

3. *If V_1, V_2 are linearly independent, and V_3, V_4 are linearly independent, and $V_1 + V_2, V_3 + V_4$ are linearly independent, then all four subspaces are linearly independent.*

Exercise 2.5. *Let V be the space of complex polynomials of degree at most 3. Define $L : V \to V$ such that $L(p) = \frac{\mathrm{d}}{\mathrm{d}x}(x^3 p(\frac{1}{x}))$. This is a linear map.*

1. *Show that $\{a + bx^2 : a, b \in \mathbb{C}\}$ and $\{ax + bx^3 : a, b \in \mathbb{C}\}$ are L-invariant subspaces.*

2. *Find the matrix of L under basis $1, x^2, x+x^3, x-x^3$, which should be block diagonal. (If you don't know the meaning of "the matrix of a linear map under a given basis", it is in the lecture notes Chapter 4, and more specifically example 4.6.4. In this problem, we require the same basis for domain and codomain, since both are the same space.)*

3. *(Read Only) This problem serves to illustrate the following geometric intuition to block diagonalize matrices in general: Find an invariant decomposition of the domain, then pick basis for each invariant subspace in this decomposition. Then the corresponding matrix under this basis would be block diagonal.*

3 HW3 (Due 3.16) More subspaces

Exercise 3.1. *Let V be the space of $n \times n$ real matrices. Let $T : V \to V$ be the transpose operation, i.e., T sends A to A^{T} for each $A \in V$. Find a non-trivial T-invariant decomposition of V, and find the corresponding block form of T.*

(Here we use real matrices for your convenience, but the statement is totally fine for complex matrices and conjugate transpose.)

Exercise 3.2. *Let $p(x)$ be any polynomial, and define $p(A)$ in the obvious manner. E.g., if $p(x) = x^2+2x+3$, then $p(A) = A^2 + 2A + 3I$. We fix some $n \times n$ matrix A.*

1. *If $AB = BA$, show that $\mathrm{Ker}(B), \mathrm{Ran}(B)$ are both A-invariant subspaces.*
2. *Prove that $Ap(A) = p(A)A$ for all polynomials $p(x)$.*
3. *Conclude that $N_\infty(A - \lambda I), R_\infty(A - \lambda I)$ are both A-invariant for any $\lambda \in \mathbb{C}$.*
4. *If $AB = BA$, show that $N_\infty(A - \lambda I), R_\infty(A - \lambda I)$ are both B-invariant for any $\lambda \in \mathbb{C}$. (This fact is used implicitly in many places.)*

Exercise 3.3. *Given a linear map $L : V \to W$, and a subspace $V' \subseteq V$, we define the pushforward subspace of V' as $L(V') = \{L\boldsymbol{v} : \boldsymbol{v} \in V'\}$.*

Given a linear map $L : V \to W$, and a subspace $V' \subseteq V$, and a subspace $W' \subseteq W$ such that $L(V') \subseteq W'$, then we can restrict the domain of L to V' and the codomain of L to W'. This way, we get a new linear map $L' : V' \to W'$, which is operationally the same as L, but with smaller domain and codomain.

1. *Show that $\mathrm{Ker}(L') = V' \cap \mathrm{Ker}(L)$.*
2. *Show that $\mathrm{Ran}(L') = L(V')$.*
3. *Show that $L(V') = \dim V' - \dim(V' \cap \mathrm{Ker}(L))$.*
4. *You don't truly understand an inequality, until you understand the meaning of the gap. From last semester, we know that given any $m \times n$ matrix A and $n \times d$ matrix B, then $\mathrm{rank}(AB) \le \mathrm{rank}(B)$. To understand the gap here, describe the subspace whose dimension is $\mathrm{rank}(B) - \mathrm{rank}(AB)$. (Hint: $\mathrm{Ran}(AB) = A(\mathrm{Ran}(B))$.)*
5. *(Read only) Compare this with HW 1.3, where we try to understand the consequence of $A^5 = 0$ or $AA^{\mathrm{T}}AAA^{\mathrm{T}} = 0$. For example, the rank of $AA^{\mathrm{T}}AAA^{\mathrm{T}}$ and the rank of $A^{\mathrm{T}}AAA^{\mathrm{T}}$ differ by the dimension of the space $\mathrm{Ker}(A) \cap \mathrm{Ran}(A^{\mathrm{T}}AAA^{\mathrm{T}}) \subseteq \mathrm{Ker}(A) \cap \mathrm{Ran}(A^{\mathrm{T}}) = \{\mathbf{0}\}$. So $AA^{\mathrm{T}}AAA^{\mathrm{T}} = 0$ if and only if $A^{\mathrm{T}}AAA^{\mathrm{T}} = 0$, and so on.*

Exercise 3.4. *Given a linear transformation $A : \mathbb{C}^n \to \mathbb{C}^n$, consider the decomposition $\mathbb{C}^n = N_\infty(A - \lambda I) \oplus R_\infty(A - \lambda I)$. Let A_N be the restriction of A with domain and codomain $N_\infty(A)$, and let A_R be the restriction of A with domain and codomain $R_\infty(A)$.*

1. *Suppose the subspace $V \subseteq \mathbb{C}^n$ is A-invariant, and the restriction of A with domain and codomain V is nilpotent. Show that for any $\boldsymbol{v} \in V$, we can find some positive integer k such that $\boldsymbol{v} \in \mathrm{Ker}(A^k)$.*
2. *Suppose the subspace $V \subseteq \mathbb{C}^n$ is A-invariant, and the restriction of A with domain and codomain V is invertible. Show that for any $\boldsymbol{v} \in V$, and for all positive integer k, we always have $\boldsymbol{v} \in \mathrm{Ran}(A^k)$.*

4 HW4 (Due 3.23) Jordan Canonical Form

Exercise 4.1. *Find a basis in the following vector space so that the linear map involved will be in Jordan normal form. Also find the Jordan normal form.*

1. *$V = \mathbb{C}^2$ is a 4 dimensional real vector space, and $A : V \to V$ that sends $\begin{bmatrix} x \\ y \end{bmatrix}$ to $\begin{bmatrix} \overline{x} - \Re(y) \\ (1+i)\Im(x) - y \end{bmatrix}$ is a real linear map. (Here $\overline{x}$ means the complex conjugate of a complex number x, and $\Re(x), \Im(x)$ means the real part and the imaginary part of a complex number x.)*

2. *$V = P_4$, the real vector space space of all real polynomials of degree at most 4. And $A : V \to V$ is a linear map such that $A(p(x)) = p'(x) + p(0) + p'(0)x^2$ for each polynomial $p \in P_4$.*

3. $A = \begin{bmatrix} & & & a_1 \\ & & a_2 & \\ & a_3 & & \\ a_4 & & & \end{bmatrix}$. *Be careful here. Maybe we have many possibilities for its Jordan normal form depending on the values of a_1, a_2, a_3, a_4.*

Exercise 4.2. *A partition of integer n is a way to write n as a sum of other positive integers, say $5 = 2+2+1$. If you always order the summands from large to small, you end up with a dot diagram, where each column represent an integer: $\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \end{bmatrix}$. Similarly, $7 = 2+4+1$ should be represented as $\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \\ \cdot & & \\ \cdot & & \end{bmatrix}$. Again, note that we always first re-order the summands from large to small.*

1. *If the Jordan normal form of an $n \times n$ nilpotent matrix A is $diag(J_{a_1}, J_{a_2}, ..., J_{a_k})$, then we have a partition of integer $n = a_1 + ... + a_k$. However, we also have a partition of integer $n = [\dim \mathrm{Ker}(A)] + [\dim \mathrm{Ker}(A^2) - \dim \mathrm{Ker}(A)] + [\dim \mathrm{Ker}(A^3) - \dim \mathrm{Ker}(A^2)] + ...$, where we treat the content of each bracket as a positive integer. Can you find a relation between the two dot diagrams? (Don't be shy, try some examples to give yourself some hints.)*

2. *A partition of integer $n = a_1 + ... + a_k$ is called self-conjugate if, for the matrix $A = diag(J_{a_1}, J_{a_2}, ..., J_{a_k})$, the two dot diagrams you obtained above are the same. Show that, for a fixed integer n, the number of self-conjugate partition of n is equal to the number of partition of n into distinct odd positive integers.(Hint: The best prove is to establish a one-to-one correspondence. For a self-conjugate dot diagram, count the total number of dots that are either in the first column or in the first row or in both. Is this always odd?)*

3. *Suppose a 4 by 4 matrix A is nilpotent and upper trianguler, and all (i,j) entries for $i < j$ are chosen randomly, independently and uniformly in the interval $[-1,1]$. What are the probabilities that its Jordan canonical form corresponds to the partitions $4 = 4, 4 = 3+1, 4 = 2+2, 4 = 2+1+1, 4 = 1+1+1+1$? (Hint: What are the possible probabilistic distribution for the rank of A?)*

4. *(NOT part of the HW.) If you want a challenge, show that the number of partitions of n into distinct parts is the same as the number of partitions of n into odd parts. Perferably you should do this via some construction of one-to-one correspondence.*

5. *(NOT part of the HW.) As a side remark, two matrices A, B are similar in GL_n if $A = CBC^{-1}$ for some invertible C. But in physics sometimes we are interested in the case when two matrices A, B are similar in SO_n, i.e., if $A = CBC^{-1}$ for some orthogonal C with determinant 1. You may also require C to be symplectic or whatever. The similarity classes in each case usualy corresponds with some special kind of partitions of integers (although they no longer necessarily be related to Jordan normal forms). Partitions of integers also connect with physics DIRECTLY by providing an estimate for the density of energy levels for a heavy nucleus. (I don't really know how, so don't ask me.) Curiously,*

many properties of partitions of integers are still OPEN PROBLEMS of mathematics. We don't really know enough about them.

Exercise 4.3. *Suppose we have a complex matrix* $A = \begin{bmatrix} B & I \\ & B \end{bmatrix}$. *We know the characteristic polynomial of* A *is just the square of the characteristic polynomial of* B. *Is the minimal polynomial of* A *the square of minimal polynoial of* B*? Let us investigate this.*

1. *Show that for any square matrices* X, Y*, the matrices* $\begin{bmatrix} X & & I & \\ & Y & & I \\ & & X & \\ & & & Y \end{bmatrix}$ *and* $\begin{bmatrix} X & I & & \\ & X & & \\ & & Y & I \\ & & & Y \end{bmatrix}$ *are similar.*
2. *Suppose* $B = \begin{bmatrix} 3 & \\ & 4 \end{bmatrix}$, *find the minimal polynomial for* B *and for* A.
3. *Suppose* $B = \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix}$, *find the rank of* A, A^2, A^3, A^4, *and deduce the Jordan canonical form* J *of* A. *(You are not required to find the* X *such that* $A = XJX^{-1}$. *Finding the* J *is enough.)*
4. *Suppose* $B = \begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix}$, *find the minimal polynomial for* B *and for* A.
5. *Guess (no need to prove) the general relation between the minimal polynomial of* A *and of* B. *(The proof is not too bad, just lenthy. It is the usual proof: first prove it for nilpotent Jordan blocks, then for* λ*-Jordan blocks, then for matrices whose eigenvalues are all the same, and finally for an arbitrary matrix.)*

5 HW5 (Due 3.30) Sylvester's equation

Exercise 5.1. *Consider the matrix* $A = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 4 \end{bmatrix}$.

1. *Find a matrix B such that* $BAB^{-1} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 4 \end{bmatrix}$.
2. *Find a basis for the subspace $V_3 + V_4$, where V_λ is the eigenspace of A for the eigenvalue λ.*
3. *Find the minimal polynomial of A.*

Exercise 5.2. *In class we see that for Sylvester's equations $AX - XB = C$, if A, B have no common eigenvalue, then there is always a unique solution. What if A, B have common eigenvalues? Let us take an extreme case, and assume that $A = B$. So we are looking at an equation $AX - XA = C$ for constant $n \times n$ matrices A, C. Let V be the space of $n \times n$ matrices, and consider the linear map $L : V \to V$ such that $L(X) = AX - XA$.*

1. *Show that $L(X) = 0$ always has infinitely many solutions.*
2. *Show that L satisfy the Leibniz rule (or the product rule for derivatives), i.e., $L(XY) = L(X)Y + XL(Y)$. (This is an indication of the usefulness of this map in physics.)*
3. *Suppose $L(X)X = XL(X)$. Show that L satisfy the chain rule, i.e., for any polynomial $p(x)$, we have $L(p(X)) = L(X)p'(X)$, where $p'(x)$ is the derivative of $p(x)$.*
4. *Find an example of matrices A, X such that $L(X)X \neq XL(X)$. Verify that $L(X^2) = L(X)X + XL(X) \neq L(X)(2X)$. (So chain rule is still true, but we need to address non-commutativity.)*
5. *Show that $L(X) = I$ has no solution by choosing $p(x)$ in the last subproblem to be the minimal polynomial. (Maybe you have seen another proof of this fact before. To show that $AX - XA = I$ has no solution, you can also just take trace on both sides. Hence now you have two proofs, and the new one is more insightful with potential meanings in physics.)*
6. *If A is diagonalizable with distinct eigenvalues, find* $\dim \operatorname{Ker}(L)$.
7. *Find a 3×3 matrix A such that* $\operatorname{Ran}(L)$ *is exactly the subspace of matrices with zero entries on the diagonal. (Hint: the range of L here is as large as possible. So pick an A that makes the kernel of L as small as possible.)*

Exercise 5.3. *Note that any linear transformation must have at least one eigenvector. (You may try to prove this yourself, but it is not part of this homework.) You may use this fact freely in this problem.*

Fix any two $n \times n$ square matrices A, B. Suppose $AB = BA$.

1. *If W is an A-invariant subspace, show that A has an eigenvector in W.*
2. *Show that* $\operatorname{Ker}(A - \lambda I)$ *is always B-invariant for all $\lambda \in \mathbb{C}$. (Hint: HW problem in the past.)*
3. *Show that A, B has a common eigenvector. (Hint: Last two sub-problems.)*
4. *Show that we can find invertible X_1, such that* $A = X_1 \begin{bmatrix} a_1 & * \\ & A_1 \end{bmatrix} X_1^{-1}, B = X_1 \begin{bmatrix} b_1 & * \\ & B_1 \end{bmatrix} X_1^{-1}$, *and $A_1 B_1 = B_1 A_1$. (Hint: Use the common eigenvector.)*
5. *Show that A, B can be simultaneously triangularized, i.e., we can find invertible X such that $A = XR_AX^{-1}, B = XR_BX^{-1}$ for some upper triangular R_A, R_B. (Hint: Use induction. Note that A, B can be simultaneously triangularized, but might be impossible to simultaneously "Jordanize".)*

6 HW6 (Due 4.6) Functions of Matrices

Exercise 6.1. *Let* $J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. *We have a function* $f(x) = x|x|$, *and we would like to define* $f(J)$. *Note that as a real function,* $f(x)$ *is everywhere differentiable. (However, as a complex function, it is not differentiable.)*

1. *Let* $A_t = \begin{bmatrix} 1 & 1 \\ 0 & 1+t \end{bmatrix}$. *Note that* $\lim A_t = J$. *Find* $\lim f(A_t)$. *(Since* A_t *are all diagonalizable for* $t \neq 0$, *therefore* $f(A_t)$ *is defined via diagonalization and then apply* f *to the diagonal entries. I.e.,* $f(X \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} X^{-1}) = X \begin{bmatrix} f(d_1) & & \\ & \ddots & \\ & & f(d_n) \end{bmatrix} X^{-1}$.*)*

2. *Let* $A_t = \begin{bmatrix} 1 & 1 \\ -t^2 & 1 \end{bmatrix}$. *Note that* $\lim A_t = J$. *Find* $\lim f(A_t)$. *Is* $f(J)$ *well-defined? (No credit but fun to think about: Why is real differentiability not enough?)*

3. *No credit, for fun challenge problem: If all* A_t *have real eigenvalues and* $\lim A_t = J$, *would* $\lim f(A_t)$ *always converge to the same matrix? If always converge to the same matrix, find this matrix. If they could converge to different things, given two sequences with different limits.*

Exercise 6.2. *Let us take another look at the famous Sherman-Morrison-Woodbury formula, which claims that* $(I - AB)^{-1} = I + A(I - BA)^{-1}B$. *Here* A *is any* $m \times n$ *matrix and* B *is any* $n \times m$ *matrix, and we assume that* $I + AB, I + BA$ *are both invertible.*

1. *Let* $f(x) = \frac{1}{1-x}$. *Find the Taylor expansion of* $f(x)$ *at* $x = 0$ *and find the radius of convergence.*

2. *Suppose all eigenvalues of* AB *and* BA *have absolute value less than* 1. *Express* $(I-AB)^{-1}, (I-BA)^{-1}$ *as power series of* AB *and* BA.

3. *Suppose all eigenvalues of* AB *and* BA *have absolute value less than* 1. *Compute directly using the power series above, to see that* $(I - AB)^{-1} = I + A(I - BA)^{-1}B$.

Exercise 6.3. *Let us figure out the derivative of* $f(A + tI)$ *as a function of* t.

1. *Suppose* A *is a single Jordan block. Show that* $\frac{\mathrm{d}}{\mathrm{d}t}(f(A + tI)) = f'(A + tI)$.

2. *Use Jordan canonical form and the last sub-problem to show that* $\frac{\mathrm{d}}{\mathrm{d}t}(f(A+tI)) = f'(A+tI)$ *is true in general.*

3. *Prove or find counter example: For a differentiable function* f, *the derivative to* $f(A + tB)$ *as a differentiable function of* t *at* $t = 0$ *is* $f'(A)B$.

Exercise 6.4. *Suppose we want to compute* $f(\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix})$. *The output should be something like* $\begin{bmatrix} f(2A) & * \\ & f(2A) \end{bmatrix}$. *But what is the formula for the upper right block?*

1. *Suppose* $f(x) = x^3$, *find the formula for* $f(\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix})$.

2. *Suppose* A *is* 1×1, *find the formula for* $f(\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix})$.

3. *Let* $M_t = \begin{bmatrix} 2A & A \\ & 2A + tI \end{bmatrix}$ *for* $t \neq 0$. *Can you find* X_t *such that* $M_t = X_t \begin{bmatrix} 2A & \\ & 2A + tI \end{bmatrix} X_t^{-1}$*?*

4. *Find* $f(\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix})$ *by taking limits of* $f(M_t)$.

7 HW7 (Due 4.27) Commuting matrices

Exercise 7.1. *Let* $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$*, and consider the Sylvester's equation* $AX - XA^{\mathrm{T}} = B$ *for some constant matrix* B*.*

1. *Write out a matrix* M *such that* $M\mathrm{vec}(X) = \mathrm{vec}(B)$*. Again, recall that* $\mathrm{vec}(\begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix}) = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$*.*
2. *When* $B = \begin{bmatrix} 0 & -4 \\ 4 & 0 \end{bmatrix}$*, find all* X *that solves this Sylvester's equation.*
3. *Find all possible* B *such that the Sylvester's equation* $AX - XA^{\mathrm{T}} = B$ *has solutions.*

Exercise 7.2. *Let* $A = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$*,* $B = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ & 0 & 1 & 0 & 1 \\ & & 0 & 0 & 0 \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$*.*

1. *Show that* $AB = BA$*.*
2. *Suppose we can find a matrix* C *and functions* f, g *such that* $A = f(C), B = g(C)$*. Show that all eigenvectors of* C *must be multiples of* $\boldsymbol{e}_1$*. (So the Jordan canonical form of* C *has a single Jordan block.)*
3. *Show that it is impossible to find a matrix* C *and functions* f, g *such that* $A = f(C), B = g(C)$*. (Hint: Check rank of* A, B*.)*
4. *Show that if* A_1, A_2, B_1, B_2 *such that* $A = A_1 \otimes A_2, B = B_1 \otimes B_2$*, and* $A_iB_i = B_iA_i$ *for* $i = 1, 2$*, then this is the trivial Kronecker tensor decomposition. I.e., we must have* A_i, B_i *to be some scalar multiples of* A, B *for some* i*. (So this commutative behavior is different from any commutative behavior we talked about in class.)*
5. *Is it possible to find* X *such that* A, B *are both in Jordan canonical form? (Hint: consider the kernel of* A^2 *and* B^2*.)*
6. *Find* X *such that* $X^{-1}BX$ *is in Jordan canonical form, while* $X^{-1}AX$ *is upper triangular.*

Exercise 7.3. *Let* V *be the space of real polynomials of degree less than* 3*. Which of the following is a dual vector? Prove it or show why not.*

1. $p \mapsto \mathrm{ev}_5((x+1)p(x))$*.*
2. $p \mapsto \lim_{x\to\infty} \frac{p(x)}{x}$*.*
3. $p \mapsto \lim_{x\to\infty} \frac{p(x)}{x^2}$*.*
4. $p \mapsto p(3)p'(4)$*.*
5. $p \mapsto \deg(p)$*, the degree of the polynomial* p*.*

8 HW8 (Due 5.11) Dual Space and Dual Basis

Exercise 8.1. *Let V be the space of real polynomials of degree less than n. So $\dim V = n$. Then for each $a \in \mathbb{R}$, the evaluation ev_a is a dual vector.*

For any real numbers $a_1, \dots, a_n \in \mathbb{R}$, consider the map $L : V \to \mathbb{R}^n$ such that $L(p) = \begin{bmatrix} p(a_1) \\ \vdots \\ p(a_n) \end{bmatrix}$.

1. *Write out the matrix for L under the basis $1, x, \dots, x^{n-1}$ for V and the standard basis for $\mathbb{R}^n$. (Do you know the name for this matrix?)*
2. *Prove that L is invertible if and only if $a_1, \dots, a_n$ are distinct. (If you can name the matrix L, then you may use its determinant formula without proof.)*
3. *Show that $\mathrm{ev}_{a_1}, \dots, \mathrm{ev}_{a_n}$ form a basis for V^* if and only if all $a_1, \dots, a_n$ are distinct.*
4. *Set $n = 3$. Find polynomials p_{-1}, p_0, p_1 such that $p_i(j) = \delta_{ij}$ for $i, j \in \{-1, 0, 1\}$.*
5. *Set $n = 4$, and consider $\mathrm{ev}_{-2}, \mathrm{ev}_{-1}, \mathrm{ev}_0, \mathrm{ev}_1, \mathrm{ev}_2 \in V^*$. Since $\dim V^* = 4$, these must be linearly dependent. Find a non-trivial linear combination of these which is zero.*

Exercise 8.2. *Find the following dual basis.*

1. *Let V be the space of real polynomials of degree less than 3. Find a basis of V whose dual basis in V^* is $\alpha_0, \alpha_1, \alpha_2$ such that $\alpha_i(p) = \int_{-1}^{1} x^i p(x)\,\mathrm{d}x$.*
2. *Let V be the space of real functions spanned by $1, \sin(x), \cos(x), \sin(2x), \cos(2x)$. Using the basis $1, \sin(x), \cos(x), \sin(2x), \cos(2x)$ for V, describe the corresponding dual basis for V in terms of integrals. (Hint: For various integers a, b, consider $f(x) \mapsto \int_0^{2\pi} \sin(ax)\sin(bx)\,\mathrm{d}x$ and $f(x) \mapsto \int_0^{2\pi} \sin(ax)\cos(bx)\,\mathrm{d}x$ and $f(x) \mapsto \int_0^{2\pi} \cos(ax)\cos(bx)\,\mathrm{d}x$.)*
3. *Consider V with basis $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$. Then it has a dual basis for V^*, which in turn has a dual basis for V^{**}. We can call this the "double dual basis" for V^{**}. Find this double dual basis in terms of $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$.*

Exercise 8.3. *Fix a differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$, and fix a point $\boldsymbol{p} \in \mathbb{R}^2$. For any vector $\boldsymbol{v} \in \mathbb{R}^2$, then the directional derivative of f at $\boldsymbol{p}$ in the direction of $\boldsymbol{v}$ is defined as $\nabla_{\boldsymbol{v}} f := \lim_{t \to 0} \frac{f(\boldsymbol{p}+t\boldsymbol{v})-f(\boldsymbol{p})}{t}$. Show that the map $\nabla f : \boldsymbol{v} \mapsto \nabla_{\boldsymbol{v}}(f)$ is a dual vector in $(\mathbb{R}^2)^*$, i.e., a row vector. Also, what are its "coordinates" under the standard dual basis?*

(Remark: In calculus, we write ∇f as a column vector for historical reasons. By all means, from a mathematical perspective, the correct way to write ∇f is to write it as a row vector, as illustrated in this problem. (But don't annoy your calculus teachers though.... In your calculus class, you use whatever notation your calculus teacher told you.)

(Extra Remark: If we use row vector, then the evaluation of ∇f at $\boldsymbol{v}$ is purely linear, and no inner product structure is needed, which is sweet. But if we HAVE TO write ∇f as a column vector (for historical reason), then we would have to do a dot product between ∇f and $\boldsymbol{v}$, which now requires an inner product structure. That is an unnecessary dependence on an extra structure that actually should have no influence.)

Exercise 8.4 (What is a derivative)**.** *The discussions in this problem holds for all manifolds M. But for simplicities sake, suppose $M = \mathbb{R}^3$ for this problem.*

Let V be the space of all analytic functions from M to $\mathbb{R}$. Here analytic means $f(x, y, z)$ is a infinite polynomial series (its Taylor expansion) with variables x, y, z. Approximately $f(x, y, z) = a_0 + a_1 x + a_2 y + a_3 z + a_4 x^2 + a_5 xy + a_6 xz + a_7 y^2 + \dots$, and things should converge always for analytic functions.

Then a dual vector $v \in V^*$ is said to be a "derivation at $\boldsymbol{p} \in M$" if it satisfy the following Leibniz rule (or product rule) at $\boldsymbol{p}$:

$$v(fg) = f(\boldsymbol{p})v(g) + g(\boldsymbol{p})v(f).$$

(Note the similarity with your traditional product rule $(fg)'(x) = f(x)g'(x) + g(x)f'(x)$.)
Prove the following:

1. *Constant functions in V must be sent to zero by all derivations at any point.*

2. *Let $x, y, z \in V$ be the coordinate function, i.e.,* $x(\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}) = p_1, y(\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}) = p_2, z(\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}) = p_3$. *Fix a point* $\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$, *then for any derivation v at $\boldsymbol{p}$, then we have* $v((x-p_1)f) = f(\boldsymbol{p})v(x)$, $v((y-p_2)f) = f(\boldsymbol{p})v(y)$ *and* $v((z - p_3)f) = f(\boldsymbol{p})v(z)$.

3. *Let $x, y, z \in V$ be the coordinate function. Fix a point* $\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$, *then for any derivation v at $\boldsymbol{p}$, then we have* $v((x - p_1)^a(y - p_2)^b(z - p_3)^c) = 0$ *for any non-negative integers a, b, c such that $a + b + c > 1$.*

4. *Let $x, y, z \in V$ be the coordinate function. Fix a point* $\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$, *then for any derivation v at $\boldsymbol{p}$,* $v(f) = \frac{\partial f}{\partial x}(\boldsymbol{p})v(x) + \frac{\partial f}{\partial y}(\boldsymbol{p})v(y) + \frac{\partial f}{\partial z}(\boldsymbol{p})v(z)$. *(Hint: use the Taylor expansion of f at $\boldsymbol{p}$.)*

5. *Any derivation v at $\boldsymbol{p}$ must be exactly the directional derivative operator $\nabla_{\boldsymbol{v}}$ where* $\boldsymbol{v} = \begin{bmatrix} v(x) \\ v(y) \\ v(z) \end{bmatrix}$.

(Remark: So, algebraically speaking, tangent vectors at $\boldsymbol{p}$ are exactly derivations at $\boldsymbol{p}$, i.e., "dual functions" that satisfy the Leibniz rule at $\boldsymbol{p}$.)

Exercise 8.5. *Consider a linear map $L : V \to W$ and its dual map $L^* : W^* \to V^*$. Assume that V, W are finite dimensional. Prove the following.*

1. $\mathrm{Ker}(L^*)$ *is exactly the collection of dual vectors in W^* that kills* $\mathrm{Ran}(L)$.

2. $\mathrm{Ran}(L^*)$ *is exactly the collection of dual vectors in V^* that kills* $\mathrm{Ker}(L)$.

9 HW9 (Due 5.18) Dual Maps

Exercise 9.1 (What is a vector field)**.** *The discussions in this problem holds for all manifolds M. But for simplicities sake, suppose $M = \mathbb{R}^3$ for this problem. Let V be the space of all analytic functions from M to $\mathbb{R}$ as usual.*

We say $X : V \to V$ is a vector field on X if $X(fg) = fX(g) + gX(f)$, and X is linear. I.e., the key property is the Leibniz rule again!

Prove the following:

1. *Show that $X_{\boldsymbol{p}} : V \to \mathbb{R}$ such that $X_{\boldsymbol{p}}(f) = (X(f))(\boldsymbol{p})$ is a derivation at $\boldsymbol{p}$. (Hence X is indeed a vector field, since it is the same as picking a tangent vector at each point.)*

2. *Note that each f on M induces a covector field $\mathrm{d}f$. Then at each point $\boldsymbol{p}$, the cotangent vector $\mathrm{d}f$ and the tangent vector X would evaluate to some number. So $\mathrm{d}f(X)$ is a function $M \to \mathbb{R}$. Show that $\mathrm{d}f(X) = X(f)$, i.e., the two are the same. (Hint: You may use $\mathrm{d}f|_p(X_p) = X_p(f)$ as the definition.)*

3. *If $X, Y : V \to V$ are vector fields, then note that $X \circ Y : V \to V$ might not be a vector field. (Leibniz rule might fail.) However, show that $X \circ Y - Y \circ X$ is always a vector field.*

4. *On a related note, show that if A, B are skew-symmetric matrices, then $AB - BA$ is still skew-symmetric. (Skew-symmetric matrices actually corresponds to certain vector fields on the manifold of orthogonal matrices. So this is no coincidence.)*

Exercise 9.2. *On the space $\mathbb{R}^n$, we fix a symmetric positive-definite matrix A, and define an inner product notation $\langle \boldsymbol{v}, \boldsymbol{w} \rangle$ as the scalar $\boldsymbol{v}^{\mathrm{T}} A \boldsymbol{w}$. We use notation $\langle \boldsymbol{v}, - \rangle$ to denote the dual vector sending input $\boldsymbol{x}$ to output $\langle \boldsymbol{v}, \boldsymbol{x} \rangle$. The bra map is the map sending $\boldsymbol{v}$ to the dual vector $\langle \boldsymbol{v}, - \rangle$.*

1. *Show that this is an inner product.*

2. *The Riesz map (inverse map of the bra map) from V^* to V would send a row vector $\boldsymbol{v}^{\mathrm{T}}$ to what?*

3. *The bra map from V to V^* would send a vector $\boldsymbol{v}$ to what?*

4. *The dual of the Riesz map from V^* to V^{**} would send a row vector $\boldsymbol{v}^{\mathrm{T}}$ to what?*

Exercise 9.3 (Partial Integration)**.** *This problem appears in last year's final of this class.*

Let V be the space of real polynomials on two variables x, y with degree at most two. In particular, a typical element of V looks like $p(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$ for some $a, b, c, d, e, f \in \mathbb{R}$. Let W be the space of polynomials on the variable y with degree at most 2. We define a map $S : V \to W$ that sends $p(x, y)$ to $\int_0^1 p(x, y)\,\mathrm{d}x$. (For this integration, we hold y constant.) Then we also have a dual map $S^* : W^* \to V^*$.

1. *Consider $\mathrm{ev}_5 \in W^*$, which sends each $q(y) \in W$ to $q(5)$. What would $S^*(\mathrm{ev}_5)$ send $p(x, y)$ to? Describe it as a formula involving the input $p(x, y)$, then pick any basis and find the coordinates for $S^*(\mathrm{ev}_5)$.*

2. *Consider $\int_0^1 \in W^*$, which sends each $q(y) \in W$ to $\int_0^1 q(y)\,\mathrm{d}y$. What would $S^*(\int_0^1)$ send $p(x, y)$ to? Describe it as a formula involving the input $p(x, y)$, then pick any basis and find the coordinates for $S^*(\int_0^1)$.*

3. *We give W the inner product defined as $\langle q_1(y), q_2(y) \rangle = \int_0^1 q_1(y) q_2(y)\,\mathrm{d}y$. (You don't have to verify that this is indeed an inner product.) Let $B : W \to W^*$ be the bra map. (I.e., $B(q_1) \in W^*$ would send $q_2 \in W$ to $\langle q_1, q_2 \rangle$.) Pick any basis and find the coordinates for $S^*BS(p(x, y))$ when $p(x, y) = 2xy$. Also find the rank of the linear map S^*BS.*

10 HW10 (Due 5.25) Tangents and Cotangents

Exercise 10.1 (Covector fields on a sphere)**.** *Let $M \subseteq \mathbb{R}^3$ be the unit sphere, i.e., it is the collection of points $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that $x^2 + y^2 + z^2 = 1$. Let $x, y, z : M \to \mathbb{R}$ be the coordinate function, i.e., they send input point $p \in M$ to the corresponding coordinates. Then they induces covector fields $\mathrm{d}x, \mathrm{d}y, \mathrm{d}z$ on M.*

1. *At the point $p = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$, describe the tangent space T_pM at p, i.e., what are all possible velocities if we move on the sphere through p? (This should be some subspace of $\mathbb{R}^3$.)*
2. *At the point $p = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$, let $\boldsymbol{v} = \begin{bmatrix} \sqrt{2} \\ -\sqrt{3} \\ 0 \end{bmatrix}$ be a tangent vector at p. Compute $(2\,\mathrm{d}x_p + 3\,\mathrm{d}y_p)(\boldsymbol{v})$.*
3. *At the point $p = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$, we have three covectors $\mathrm{d}x_p, \mathrm{d}y_p, \mathrm{d}z_p$, but the cotangent space to the sphere at p is two dimensional! So at the point p, find a linear relation between $\mathrm{d}x_p, \mathrm{d}y_p, \mathrm{d}z_p$.*
4. *Find all the points $p \in M$ such that $\mathrm{d}x_p, \mathrm{d}y_p$ do not form a basis for the cotangent space at p.*
5. *Let γ be a curve going from the south pole $\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$ to the north pole $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Find $\int_\gamma \mathrm{d}x, \int_\gamma \mathrm{d}y, \int_\gamma \mathrm{d}z$.*
6. *Let γ be a curve going from the point $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$ to the point $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$. Find $\int_\gamma (yz\,\mathrm{d}x + zx\,\mathrm{d}y + xy\,\mathrm{d}z)$.*

Exercise 10.2 (Stereographic projection)**.** *Let S be a unit sphere placed on top of the xy-plane of $\mathbb{R}^3$, i.e., it has center $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and diameter 1, so its bottom is barely touching the xy-plane, i.e., the "ground". Imagine that you are standing on the "north pole" of the sphere, i.e., the point $\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$. You hold a gun in your hand, and for each point $p \in S$ that is not the north pole, you can shoot at it. Now imagine that your sphere S is made of paper, so that your bullet would go through the point p and hit some point $f(p)$ on the ground (xy-plane). This gives a function $f : S - \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \to \mathbb{R}^2$. This map is called the stereographic projection, and it is very useful in complex analysis, projective geometrix, hyperbolic geometry, photography and so on.*

1. *For $p = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, find the coordinates of $f(p)$ in terms of a, b, c.*
2. *At the point $p = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ 1 + \frac{1}{\sqrt{3}} \end{bmatrix}$ on the sphere S, any motion on $S - \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$ through p, f would then map this motion to some motion on $\mathbb{R}^2$. Consequently, any tangent vector (velocity) $\boldsymbol{v}$ at p will be sent to some*

tangent vector at $f(p)$. So we have a map $f_* : T_pS \to T_{f(p)}\mathbb{R}^2$. Let $\boldsymbol{v} = \begin{bmatrix} 1 \\ ? \\ ? \end{bmatrix}$ be a tangent vector in T_pS parallel to the ground, going counter-clockwise. What would the values in the question marks be? Also find $f_*(\boldsymbol{v})$.

3. *Consider the map $g : \mathbb{R}^2 \to \mathbb{R}$ such that $g(x, y) = xy$. Then $\mathrm{d}g$ is a covector field on $\mathbb{R}^2$. Let $f^* : (T_{f(p)}\mathbb{R}^2)^* \to (T_pS)^*$ be the dual map of f_*. Then $f^*(\mathrm{d}g)$ is now a covector field on $S - \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$. Find the row vector $f^*(\mathrm{d}g_{f(p)})$ where p is chosen as the last problem.*

11 HW11 (Due 6.1) Tensors

Exercise 11.1 (Tensor of vectors). *Recall that for $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$, $\boldsymbol{v} \otimes \boldsymbol{w}$ is a bilinear map $V^* \times W^* \to \mathbb{R}$, sending (α, β) to $\alpha(\boldsymbol{v})\beta(\boldsymbol{w})$.*

1. *Using definitions, verify that $(k\boldsymbol{v}) \otimes \boldsymbol{w} = \boldsymbol{v} \otimes (k\boldsymbol{w}) = k(\boldsymbol{v} \otimes \boldsymbol{w})$.*
2. *Using definitions, verify that $(\boldsymbol{v}_1 + \boldsymbol{v}_2) \otimes \boldsymbol{w} = \boldsymbol{v}_1 \otimes \boldsymbol{w} + \boldsymbol{v}_2 \otimes \boldsymbol{w}$.*
3. *Set $V = W = \mathbb{R}^2$, and set $\boldsymbol{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Show that $\boldsymbol{v} \otimes \boldsymbol{w} \neq \boldsymbol{w} \otimes \boldsymbol{v}$.*
4. *Set $V = W = \mathbb{R}^3$. For any $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathbb{R}^3$, assume $\boldsymbol{v}_1, \boldsymbol{w}_1$ are linearly independent, and $\boldsymbol{v}_2, \boldsymbol{w}_2$ are linearly independent. Show that $\boldsymbol{v}_1 \otimes \boldsymbol{w}_1 - \boldsymbol{w}_1 \otimes \boldsymbol{v}_1 = \boldsymbol{v}_2 \otimes \boldsymbol{w}_2 - \boldsymbol{w}_2 \otimes \boldsymbol{v}_2$ if and only if the parallelogram made by $\boldsymbol{v}_1, \boldsymbol{w}_1$ and the parallelogram made by $\boldsymbol{v}_2, \boldsymbol{w}_2$ have the same area, span the same subspace, and share the same orientation (clockwise or counter clockwise). (Hint: The parallelogram condition means $\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{w}_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{v}_2 & \boldsymbol{w}_2 \end{bmatrix} A$ for some 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with determinant $ad - bc = 1$.)*
5. *(Read only) One can think of $\boldsymbol{v} \wedge \boldsymbol{w}$ defined as $\boldsymbol{v} \otimes \boldsymbol{w} - \boldsymbol{w} \otimes \boldsymbol{v}$ as a "bivector", i.e., they are representing 2-dim shapes, but we only care about direction, orientation and magnitude (area). When you have any parallelogram in space, but you only care about direction, orientation and magnitude, then you can use a bivector to represent it and perform calculations. For example, when $V = \mathbb{R}^3$ use the dot product, then this induces an inner product on $V \otimes V$, and $\frac{\langle \boldsymbol{v} \wedge \boldsymbol{w}, \boldsymbol{v}' \wedge \boldsymbol{w}' \rangle}{||\boldsymbol{v} \wedge \boldsymbol{w}|| ||\boldsymbol{v}' \wedge \boldsymbol{w}'||}$ is the cosine of the dihedral angle between the two parallelograms.*

Exercise 11.2 (Kronecker product?). *Consider two linear maps $X : \mathbb{R}^2 \to \mathbb{R}^2$ and $Y : \mathbb{R}^3 \to \mathbb{R}^3$ over finite dimensional spaces. Suppose $X = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix}$.*

1. *Write X as a linear combination of the basis elements of the form $\boldsymbol{e}_i^{\mathrm{T}} \otimes \boldsymbol{e}_j$ in $(\mathbb{R}^2)^* \otimes \mathbb{R}^2$. Similarly, write Y as a linear combination of the basis elements of the form $\boldsymbol{e}_i^{\mathrm{T}} \otimes \boldsymbol{e}_j$ in $(\mathbb{R}^3)^* \otimes \mathbb{R}^3$.*
2. *$X \otimes Y$ is an element of $(\mathbb{R}^2)^* \otimes \mathbb{R}^2 \otimes (\mathbb{R}^3)^* \otimes \mathbb{R}^3$. Write it as a linear combination of the basis elements of the form $\boldsymbol{e}_i^{\mathrm{T}} \otimes \boldsymbol{e}_j \otimes \boldsymbol{e}_k^{\mathrm{T}} \otimes \boldsymbol{e}_\ell$ in $(\mathbb{R}^2)^* \otimes \mathbb{R}^2 \otimes (\mathbb{R}^3)^* \otimes \mathbb{R}^3$.*
3. *Note that an element of $(\mathbb{R}^2)^* \otimes \mathbb{R}^2 \otimes (\mathbb{R}^3)^* \otimes \mathbb{R}^3$ is an element of $\mathcal{L}(\mathbb{R}^2 \otimes \mathbb{R}^3, \mathbb{R}^4 \otimes \mathbb{R}^3)$. So $X \otimes Y$ is a linear map from $\mathbb{R}^2 \otimes \mathbb{R}^3$ to $\mathbb{R}^2 \otimes \mathbb{R}^3$. Use the basis $\boldsymbol{e}_1 \otimes \boldsymbol{e}_1, \boldsymbol{e}_1 \otimes \boldsymbol{e}_2, \boldsymbol{e}_1 \otimes \boldsymbol{e}_3, \boldsymbol{e}_2 \otimes \boldsymbol{e}_1, \boldsymbol{e}_2 \otimes \boldsymbol{e}_2, \boldsymbol{e}_2 \otimes \boldsymbol{e}_3$ for $\mathbb{R}^2 \otimes \mathbb{R}^3$, then what is the matrix for $X \otimes Y$? What is the relation between the matrices $X, Y, X \otimes Y$?*

Exercise 11.3 (Elementary layer operations for tensors). *Note that, for "2D" matrices we have row and column operations, and the two kinds of operations corresponds to the two dimensions of the array.*

For simplicity, let M be a $2 \times 2 \times 2$ "3D matrix". Then we have "row layer operations", "column layer operations", "horizontal layer operations". The three kinds corresponds to the three dimensions of the array.

We interpret this as a multilinear map $M : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$. Let $((\mathbb{R}^2)^*)^{\otimes 3}$ be the space of all multilinear maps from $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ to $\mathbb{R}$.

1. *Given $\alpha, \beta, \gamma \in (\mathbb{R}^2)^*$, what is the (i, j, k)-entry of the "3D matrix" $\alpha \otimes \beta \otimes \gamma$ in terms of the coordinates of α, β, γ? Here $\alpha \otimes \beta \otimes \gamma$ is the multilinear map sending $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ to the real number $\alpha(\boldsymbol{u})\beta(\boldsymbol{v})\gamma(\boldsymbol{w})$.*
2. *Let E be an elementary matrix. Then we can send $\alpha \otimes \beta \otimes \gamma$ to $(\alpha E) \otimes \beta \otimes \gamma$. Why can this be extended to a linear map $M_E : ((\mathbb{R}^2)^*)^{\otimes 3} \to ((\mathbb{R}^2)^*)^{\otimes 3}$? (This gives a formula for the "elementary layer operations" on "3D matrices", where the three kinds of layer operations corresponds to applying E to the three arguments respectively.)*

3. *Show that elementary layer operations preserve rank. Here we say M has rank r if r is the smallest possible integer such that M can be written as the linear combination of r "rank one" maps, i.e., maps of the kind $\alpha \otimes \beta \otimes \gamma$ for some $\alpha, \beta, \gamma \in (\mathbb{R}^2)^*$.*

4. *Show that, if some "2D" layer matrix of a "3D matrix" has rank r, then the 3D matrix has rank at least r.*

5. *Let M be made of two layers, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Find its rank.*

6. *(Read only) Despite some practical interests, finding the tensor rank in general is NOT easy. In fact, it is NP-complete just for 3-tensors over finite field. Furthermore, a tensor with all real entries might have different real rank and complex rank.*

Exercise 11.4. *Let M be a $3 \times 3 \times 3$ "3D matrix" whose (i, j, k)-entry is $i + j + k$. We interpret this as a multilinear map $M : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$.*

1. *Let $\boldsymbol{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, then $M(\boldsymbol{v}, \boldsymbol{v}, \boldsymbol{v})$ is a polynomial in x, y, z. What is this polynomial?*

2. *Let $\sigma : \{1, 2, 3\} \to \{1, 2, 3\}$ be any bijection. Show that $M(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) = M(\boldsymbol{v}_{\sigma(1)}, \boldsymbol{v}_{\sigma(2)}, \boldsymbol{v}_{\sigma(3)})$. (Hint: brute force works. But alternatively, try find the (i, j, k) entry of the multilinear map M^σ, a map that sends $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ to $M(\boldsymbol{v}_{\sigma(1)}, \boldsymbol{v}_{\sigma(2)}, \boldsymbol{v}_{\sigma(3)})$.)*

3. *Show that the rank r of M is at least 2 and at most 3. (It is actually exactly three, but this is not easy to prove.)*

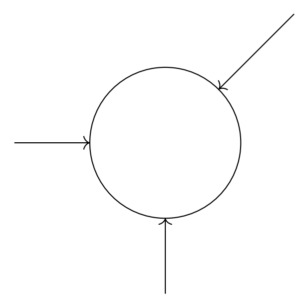
4. *(Read only) Any study of polynomial of degree d on n variables is equivalent to the study of some symmetric d tensor on $\mathbb{R}^n$.*

12 HW12 (Due 6.8) Tensors and Alternating Tensors

Exercise 12.1 (Squeezing a Ping Pong). *Tensors are basically just recording multiple vectors simultaneously, in a multilinear way. This problem applies this idea in the situation of squeezing a Ping Pong ball. The traditional problem is like this: 10 forces are applied on a ping pong ball. Where would the ping pong ball most likely crack? We hereby reduced the problem to a 2-dim version with only three forces to simplify calculation.*

If we apply a force $\boldsymbol{f} \in \mathbb{R}^2$ *on the unit circle at point* $\boldsymbol{p}$*, then we record this information as* $\boldsymbol{p}\otimes\boldsymbol{f} \in \mathbb{R}^2\otimes\mathbb{R}^2$.

Suppose we apply three forces $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{v}$ *where* $\boldsymbol{v} = -\boldsymbol{e}_1 - \boldsymbol{e}_2$*, and the forces are applied to the circle perpendicularly, as shown in the graph below.*



1. *The three forces are recorded via three tensors. Calculate their sum* T *in terms of the standard basis* $\boldsymbol{e}_{11}, \boldsymbol{e}_{12}, \boldsymbol{e}_{21}, \boldsymbol{e}_{22} \in \mathbb{R}^2 \otimes \mathbb{R}^2$*, where* $\boldsymbol{e}_{ij}$ *is the notation form* $\boldsymbol{e}_i \otimes \boldsymbol{e}_j$.

2. *Find two mutually orthogonal unit vector* $\boldsymbol{x}, \boldsymbol{y}$ *such that* T *is a linear combination of* $\boldsymbol{x} \otimes \boldsymbol{x}$ *and* $\boldsymbol{y} \otimes \boldsymbol{y}$*. (So the effect of the three original forces are equivalent to the effect of our two forces here.)*

3. *Let us say that the circle is squeezed such that its shape has changed a tiny bit, into an ellipse. Guess which directions would the long axis and short axis be?*

4. *Find a tensor* $T \in \mathbb{R}^2 \otimes \mathbb{R}^2$ *that cannot be obtained via squeezing the circle perpendicularly, with no matter how many forces. Here squeezing perpendicularly means the forces are always inward. Explain why squeezing perpendicularly cannot produce your tensor. (Hint: Suppose* T *is a sum of tensors for perpendicular squeezing forces. Write* T^{ij} *into a matrix. Is it positive definite, positive semi-definite, negative definite, negative semi-definite or indefinite?)*

Exercise 12.2 (Alternization). *When we have a tensor* τ*, we may try to alternize it. Let* V *be any vector space, then for any* $\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_k \in V^{\otimes k}$*, we define* $Alt(\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_k) = \frac{1}{k!}\sum_\sigma \operatorname{sign}(\sigma)\boldsymbol{v}_{\sigma(1)} \otimes \cdots \otimes \boldsymbol{v}_{\sigma(k)}$*. This is similar to the big formula for determinant. The idea is to permute the orders of* $\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_k$ *in all possible ways, while recording whether this is an even or odd permutation, then take average (where even permutation is added and odd permutation is substracted).*

In this problem, say $k = 3$*. So*

$$Alt(\boldsymbol{u}\otimes\boldsymbol{v}\otimes\boldsymbol{w}) = \frac{1}{6}(\boldsymbol{u}\otimes\boldsymbol{v}\otimes\boldsymbol{w} + \boldsymbol{v}\otimes\boldsymbol{w}\otimes\boldsymbol{u} + \boldsymbol{w}\otimes\boldsymbol{u}\otimes\boldsymbol{v} - \boldsymbol{u}\otimes\boldsymbol{w}\otimes\boldsymbol{v} - \boldsymbol{v}\otimes\boldsymbol{u}\otimes\boldsymbol{w} - \boldsymbol{w}\otimes\boldsymbol{v}\otimes\boldsymbol{u}).$$

1. *Show that* $Alt : V \times V \times V \to V^{\otimes 3}$ *is multilinear.*

2. *Show that for any tensor* $\tau \in V^{\otimes 3}$, $Alt(\tau)$ *is alternating.*

3. *Show that if* τ *is already an alternating tensor, then* $Alt(\tau) = \tau$*. (So Alt is the projection map from the tensor space to the space of alternating tensors.)*

Exercise 12.3. *All the tensor below is over* $\mathbb{R}^n$*. We take standard basis* $e_1, ..., e_n$ *for* $\mathbb{R}^n$ *and the corresponding standard dual basis, and the induced standard basis on all related tensor spaces. Recall that* $\boldsymbol{v} \wedge \boldsymbol{w} = \boldsymbol{v} \otimes \boldsymbol{w} - \boldsymbol{w} \otimes \boldsymbol{v}$*. Compute the following:*

1. $e^1 \otimes e^2(e_1 \otimes e_2)$;
2. $Alt(e^1 \otimes e^2)(e_1 \otimes e_2)$;
3. $e^1 \otimes e^2(Alt(e_1 \otimes e_2))$;
4. $Alt(e^1 \otimes e^2)(Alt(e_1 \otimes e_2))$;
5. $e^1 \wedge e^2(Alt(e_1 \otimes e_2))$;
6. $e^1 \wedge e^2(e_1 \otimes e_2)$;
7. $e^1 \otimes e^2(e_1 \wedge e_2)$;
8. $e^1 \wedge e^2(e_1 \wedge e_2)$;
9. *Using dot product on* $\mathbb{R}^n$ *as inner product, find* $\langle e_1 \otimes e_2, e_1 \otimes e_3\rangle$ *and* $\langle e_1 \wedge e_2, e_1 \wedge e_3\rangle$*; (The inner product of tensors is defined like this:* $\langle \boldsymbol{v}_1 \otimes \boldsymbol{w}_1, \boldsymbol{v}_2 \otimes \boldsymbol{w}_2\rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle\langle \boldsymbol{w}_1, \boldsymbol{w}_2\rangle$ *for "rank one" elements, and we extend this bilinearly to all tensors.)*
10. *Using dot product on* $\mathbb{R}^n$ *as inner product, find* $||e_1 \wedge e_2 \wedge e_3||$ *and* $||Alt(e_1 \otimes e_2 \otimes e_3)||$ *and* $||Alt(e_1 \otimes e_2) \wedge e_3||$.