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 Topics in Linear Algebra **1**
1.1 HW1 (Due 3.11) Complex Stuff 1
1.2 HW2 (Due 3.18) Invariant Decomposition 3
1.3 HW3 (Due 3.25) Jordan Canonical Form 4
1.4 HW4 (Due 4.1, this is NOT a joke....) Applications of functions of matrices 5
1.5 HW5 (Due 4.8) 6
1.6 HW6 (Due 5.13) Dual Stuff 7
1.7 HW7 (Due 5.20) Tangent vectors and Cotangent vectors 8
1.8 HW8 (Due 5.27) Multilinear maps 9
1.9 HW9 (Due 6.3) Tensor Calculations 10

Chapter 1

Topics in Linear Algebra

1.1 HW1 (Due 3.11) Complex Stuff

Exercise 1.1.1. *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 solution.*

2. *If odd n, show that there is no real solution.*

Exercise 1.1.2. *Suppose $A^2 = -I$ for a real $n \times n$ matrix A. For each vector $\boldsymbol{v} \in \mathbb{R}^n$, we write $\mathrm{i}\boldsymbol{v}$ 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 1.1.3. *If V is an abstract vector space over $\mathbb{C}$, then for each vector $\boldsymbol{v}$ and each $k \in \mathbb{C}$, obviously $k\boldsymbol{v}$ is well-defined. But as a result, for each vector $\boldsymbol{v}$ and each $k \in \mathbb{R}$, $k\boldsymbol{v}$ must be well-defined. So any complex vector space must also be a real vector space (but NOT vice versa). This gives rise to some very tricky distinctions.*

Given abstract vector spaces V, W over $\mathbb{C}$, we say a map $L : V \to W$ is complex linear if $L(k\boldsymbol{v}) = kL\boldsymbol{v}$ for all $k \in \mathbb{C}$ and $\boldsymbol{v} \in V$. We say it is real linear if $L(k\boldsymbol{v}) = kL\boldsymbol{v}$ for all $k \in \mathbb{R}$ and $\boldsymbol{v} \in V$. Note that it is possible to be real linear but NOT complex linear.

Given a bunch of vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ in a complex vector space, we say they are $\mathbb{R}$-linearly independent if $\sum a_i\boldsymbol{v}_i = \boldsymbol{0}$ for $a_1, \dots, a_k \in \mathbb{R}$ implies all $a_i = 0$. We say they are $\mathbb{C}$-linearly independent if $\sum a_i\boldsymbol{v}_i = \boldsymbol{0}$ for $a_1, \dots, a_k \in \mathbb{C}$ implies all $a_i = 0$. Similarly, we can define $\mathbb{R}$-spanning, $\mathbb{C}$-spanning, $\mathbb{R}$-basis, $\mathbb{C}$-basis and so on.

1. *Consider the map $C : \mathbb{C}^n \to \mathbb{C}^n$ of taking complex conjugates, i.e., $C\begin{bmatrix}1+\mathrm{i}\\2\\3\mathrm{i}\end{bmatrix} = \begin{bmatrix}1-\mathrm{i}\\2\\-3\mathrm{i}\end{bmatrix}$. Is this real linear? Is this complex linear?*

2. Which implies which: $\mathbb{R}$-linear and $\mathbb{C}$-linear.

3. Find an $\mathbb{R}$-basis for $\mathbb{C}^2$ and then find a $\mathbb{C}$-basis for $\mathbb{C}^2$. What is the real dimension of $\mathbb{C}^2$? What is its complex dimension?

4. Which implies which: $\mathbb{R}$-linearly independent and $\mathbb{C}$-linearly independent.

5. Which implies which: $\mathbb{R}$-spanning and $\mathbb{C}$-spanning.

Exercise 1.1.4 (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) One can compute $C\boldsymbol{v}$ using fast Fourier transform, and it will be slightly faster than computing $C\boldsymbol{v}$ directly. The speed gain is very little though, because C is too small. For a larger matrix, the speed gain will be more noticable.

1.2 HW2 (Due 3.18) Invariant Decomposition

Exercise 1.2.1. *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 1.2.2. *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.)

Proof. Let V_+ be the space of symmetric matrices, and let V_- be the space of skew-symmetric matrices. Then since $A = \frac{A+A^{\mathrm{T}}}{2} + \frac{A-A^{\mathrm{T}}}{2}$, we see that $V = V_+ + V_-$. Furtheremore, they obviously have trivial intersection. (If $A = A^{\mathrm{T}} = -A$, then $A = 0$.)

So $V = V_+ \oplus V_-$. Invariance is easy to check so I skip that.

Finally, since this is an invariant decomposition, T is block diagonal. It is the identity map on V_+ and it is negation on V_-, so its block form would be $\begin{bmatrix} I & \\ & -I \end{bmatrix}$. ☐

Exercise 1.2.3. *Consider an $m \times n$ real matrix A treated as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. Note that $\mathbb{R}^n = \mathrm{Ran}(A^{\mathrm{T}}) \oplus \mathrm{Ker}(A)$ and $\mathbb{R}^m = \mathrm{Ran}(A) \oplus \mathrm{Ker}(A^{\mathrm{T}})$. Suppose the block form for A according to these decompositions is $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. Find the rank of linear maps $A_{11}, A_{12}, A_{21}, A_{22}$.*

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

Proof. Pick any $\boldsymbol{v} \in \mathrm{Ker}(A)$, then $A\boldsymbol{v} = \mathbf{0} + \mathbf{0}$ is the decomposition for $\mathbb{R}^m = \mathrm{Ran}(A) \oplus \mathrm{Ker}(A^{\mathrm{T}})$ always. So A_{21} and A_{22} are both the zero map. They have zero rank.

Pick any $\boldsymbol{v} \in \mathrm{Ran}(A^{\mathrm{T}})$, then $A\boldsymbol{v} = A\boldsymbol{v} + \mathbf{0}$ is the decomposition for $\mathbb{R}^m = \mathrm{Ran}(A) \oplus \mathrm{Ker}(A^{\mathrm{T}})$ always. So we see that the A_{12} portion is always zero. It has zero rank.

Finally, I claim that the rank of A_{11} is the rank of A. To see that, note that $\mathrm{Ker}(A) \cap \mathrm{Ran}(A^{\mathrm{T}})$ is zero. So A will kill nothing (other than zero) in $\mathrm{Ran}(A^{\mathrm{T}})$. So $A_{11} : \mathrm{Ran}(A^{\mathrm{T}}) \to \mathrm{Ran}(A)$ has zero kernel, and thus injective. But its domain and codomain both have dimension $\mathrm{rank}(A)$. So A_{11} is bijective with rank $\mathrm{rank}(A)$. ☐

Exercise 1.2.4.

1. *If V is an A-invariant subspace, show that A has an eigenvector in V.*
2. *Suppose $AB = BA$ for two complex square matrices A, B. Then show that A, B has a common eigenvector. (Hint: show that an eigenspace for A is B-invariant.)*

Exercise 1.2.5. *Let D be the derivative operator on the space of smooth functions (i.e., infinitely differentiable functions). Find $N_\infty(D)$. Also, prove or find counter example: $N_\infty(D - I)$ is spanned by e^x. Here I is the identity operator.*

1.1小问！！！i的虚部是1不是i，虚部不虚！

1.3 HW3 (Due 3.25) Jordan Canonical Form

Exercise 1.3.1. *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 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 1.3.2. *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?*
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: 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 and uniformly in the interval $[-1, 1]$. What are the probabilities that its Jordan canonical form corresponds to the partitions $4 = 4, 4 = 3 + 1, 4 = 2 + 2, 4 = 2 + 1 + 1, 4 = 1 + 1 + 1 + 1$?*
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.*

1.4 HW4 (Due 4.1, this is NOT a joke....) Applications of functions of matrices

Exercise 1.4.1. *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.*

Exercise 1.4.2. *Suppose* $A = \begin{bmatrix} B & I \\ & B \end{bmatrix}$. *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? Prove or give counter-example.

(Not part of the problem. But if you are curious, feel free to find out the exact relation between the two minimal polynomials.)

Exercise 1.4.3. *Suppose $AB = BA$.*

1. *Show that A, B can be simultaneously triangularized. I.e., you can find invertible C such that CAC^{-1}, CBC^{-1} are both upper triangular.*

2. *Can they always be simultaneously put into Jordan normal form? Prove or provide counter example.*

Exercise 1.4.4. *Note that if a linear transformation A has minimal polynomial x^2+1, then $A(A(\boldsymbol{x}))+\boldsymbol{x} = \mathbf{0}$ for any input $\boldsymbol{x}$. This problem is not about linear algebra, but about satisfying some curiosity. Suppose we have a real function $f : \mathbb{R} \to \mathbb{R}$ such that $f(f(x)) + x = 0$ for all input x, and f is continuous, what can you say about f?*

这道题的意义：
注意到当f(A)well-defined时，我们对A采用Jordan分解后对jordan进行f(J)和利用可对角矩阵来逼近A，得到f(A)。两个结果是必然相同的，但是如果不well-defined，算出的结果很大概率不同。所以这个题的题意就是希望通过后者来验证结果会不同，然后不well-defined

1.5 HW5 (Due 4.8)

Exercise 1.5.1. *Let* $J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. *We have a function* $f(x) = x|x|$. *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* $\lim f(A_t)$.
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 1.5.2. *Compute the following*

1. *Find the derivative of* $\sin(tA)$ *as a function of* t.
2. *For the formula* $f(\begin{bmatrix} 2A & A \\ & 2A \end{bmatrix}) = \begin{bmatrix} f(2A) & B \\ & f(2A) \end{bmatrix}$, *what is the block matrix* B *in terms of* f *and* A?
3. *Prove or find counter example: The derivative to* $f(A + tB)$ *as a function of* t *at* $t = 0$ *is* $f'(A)B$.

1.5.2的意义：
注意到我们还是有jordan和可对角逼近两种算法，但是这里用jordan也不太理智，因为对分块矩阵去做jordan大概率会遇见我们没有学过的内容，所以还是采用了对角化逼近的方法，

1.6 HW6 (Due 5.13) Dual Stuff

Exercise 1.6.1. *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_2, p_3 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 1.6.2. *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.*

Exercise 1.6.3. *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. If we use row vector, then the evaluation of ∇f at $\boldsymbol{v}$ is purely linear, and no inner product structure is needed.

But if we write ∇f as a column vector, then we would have to do a dot product between ∇f and $\boldsymbol{v}$, which now requires an inner product structure. The "flipping" of a row vector to a column vector requires an identification of V and V^*, which requires an inner product structure.)

Exercise 1.6.4. *Consider a linear map $L : V \to W$ and its dual map $L^* : W^* \to V^*$. 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* $\mathrm{Ker}(L)$.

Exercise 1.6.5. *On the space $\mathbb{R}^n$, we fix a symmetric positive-definite matrix A, and define $(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^{\mathrm{T}} A \boldsymbol{w}$.*

1. *Show that this is an inner product.*
2. *The Riesz map from V^* to V would send a row vector $\boldsymbol{v}^{\mathrm{T}}$ to what?*
3. *The inverse of the Riesz 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?*
5. *The adjoint of the Riesz map from V to V^* would send a vector $\boldsymbol{v}$ to what?*

1.7 HW7 (Due 5.20) Tangent vectors and Cotangent vectors

Exercise 1.7.1 (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 $f(x, y, z) = a_0 + a_1x + a_2y + a_3z + a_4x^2 + a_5xy + a_6xz + a_7y^2 + \dots$, and things should converge always.

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):

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

(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. Suppose $\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$, 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. Suppose $\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$, 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. Suppose $\boldsymbol{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$, 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 are exactly derivations, i.e., things that satisfy the Leibniz rule.)

Exercise 1.7.2 (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 $X(fg) = fX(g) + gX(f)$, i.e., 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: just use definitions and calculate directly.)*
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.)*

1.8 HW8 (Due 5.27) Multilinear maps

Exercise 1.8.1 (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 $M : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$. We also think of M as an element of $((\mathbb{R}^2)^*)^{\otimes 3} = (\mathbb{R}^2)^* \otimes (\mathbb{R}^2)^* \otimes (\mathbb{R}^2)^*$.

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

Exercise 1.8.3 (Kronecker product?)**.** *Consider two linear maps $X : V \to V$ and $Y : W \to W$ over finite dimensional spaces. We define $X \otimes Y$ to be the map from $V \times W$ to $V \otimes W$, such that $(\boldsymbol{v}, \boldsymbol{w}) \mapsto X\boldsymbol{v} \otimes Y\boldsymbol{w}$.*

1. *Verify that $X \otimes Y$ is bilinear. Therefore we can think of it as a linear map $X \otimes Y : V \otimes W \to V \otimes W$*
2. *Show that $\operatorname{trace} X \otimes Y = (\operatorname{trace} X)(\operatorname{trace} Y)$.*

Exercise 1.8.4 (Trace as a tensor)**.** *Note that a linear map from $\mathbb{R}^n \to \mathbb{R}^n$ is an element of $\mathbb{R}^n \otimes (\mathbb{R}^n)^*$. Hence $\mathbb{R}^n \otimes (\mathbb{R}^n)^*$ is exactly the space of $n \times n$ matrices.*

The trace map send matrices to numbers. So we have $\operatorname{trace} : \mathbb{R}^n \otimes (\mathbb{R}^n)^* \to \mathbb{R}$. It is also linear. So what space is trace an element of? What are the entries of the tensor trace?

1.9 HW9 (Due 6.3) Tensor Calculations

Exercise 1.9.1 (Quantum Entanglement)**.** *(Optional Background) Quantum physics usually start like this. Given a particual A, its "possible states" are unit vectors $\boldsymbol{v}$ in some (usually infinite dimensional) inner product space H_A. (We use unit vectors, because $\|\boldsymbol{v}\|$ usually means total probability, which should be one.) When we make an observation on A, the observation is encoded by a (usually self-adjoint) linear map $L : H_A \to H_A$. If $L = I$ the identity map, then this means we do not observe at all.*

For simplicity, suppose we have two inner product spaces H_A, H_B. (You may think that they describe the states of two different particles A and B.) Then on their tensor space $H_A \otimes H_B$, we define $(\boldsymbol{v}_1 \otimes \boldsymbol{w}_1, \boldsymbol{v}_2 \otimes \boldsymbol{w}_2) = (\boldsymbol{v}_1, \boldsymbol{v}_2)(\boldsymbol{w}_1, \boldsymbol{w}_2)$ for rank one elements, and extend this bilinearly on all elements of $H_A \otimes H_B$.

For simplicity, we assume that $H_A = H_B = \mathbb{R}^2$.

1. *Show that $(-,-)$ on $H_A \otimes H_B$ is indeed an inner product. (You don't need to prove bilinearity, but prove that it is symmetric and positive definite.)*

2. *Show that $a\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + b\boldsymbol{e}_2 \otimes \boldsymbol{e}_2$ has rank two when a, b are both non-zero. (Rank one states in $H_A \otimes H_B$ are called non-entangled states. And these rank two states here are entangled states.)*

这个结果就是一个坐标，不是内积

3. *Given a pair of linear maps $L_A : H_A \to H_A$ and $L_B : H_B \to H_B$, we define the expected observation result of $L_A \otimes L_B$ to be $(\omega, L_A \otimes L_B(\omega))$. Now let $I_A : H_A \to H_A$ and $I_B : H_B \to H_B$ be the identity maps on H_A and H_B respectively. Suppose $L = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$. Show that if ω is non-entangled, say $\omega = \boldsymbol{v} \otimes \boldsymbol{w}$, then $(\omega, L \otimes I_B(\omega))$ and $(\omega, I_A \otimes L(\omega))$ could be any pair of real numbers. (Here $L_A \otimes I_B$ means that we perform observation L_A on A while we do not observe B at all. As we can see here, the obervation of only A and the observation of only B has no relation to each other.)*

4. *Suppose $L = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}$, and let $a, b \in \mathbb{R}$ be real numbers such that $a^2 + b^2 = 1$, then calculate $(\omega, L \otimes I_B(\omega))$ and $(\omega, I_A \otimes L(\omega))$ when $\omega = a\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + b\boldsymbol{e}_2 \otimes \boldsymbol{e}_2$. (This is an entangled-state, where observing A is identical to observing B.)*

Exercise 1.9.2 (Change of basis)**.** *Suppose a change of basis happens in V, from the old basis $\mathcal{B}$ to a new basis $\mathcal{C}$. For each $\boldsymbol{v} \in V$, we denote its coordinates in the basis $\mathcal{B}$ as $\boldsymbol{v}_{\mathcal{B}} \in \mathbb{R}^n$, and we similarly define $\boldsymbol{v}_{\mathcal{C}} \in \mathbb{R}^n$.*

Let B be the change of basis matrix such that $\boldsymbol{v}_{\mathcal{C}} = B\boldsymbol{v}_{\mathcal{B}}$. We consider the space of (a,b) tensors over an n-dim vector space V, $\mathcal{T}^a_b(V)$.

1. *Suppose $a = 0, b = 1$. Then $\mathcal{T}^a_b(V) = V^*$. For any $\alpha \in V^*$, let $\alpha_{\mathcal{B}} \in (\mathbb{R}^n)^*$ be the matrix for the linear map $\alpha : V \to \mathbb{R}$ under the basis $\mathcal{B}$ for V, and let $\alpha_{\mathcal{C}}$ be defined similarly. How to express the new coordinates $\alpha_{\mathcal{C}}$ in terms of the old coordinates $\alpha_{\mathcal{B}}$?*

2. *Suppose $a = 2, b = 0$. Then $\mathcal{T}^a_b(V) = V \otimes V$. Find the lineare map $L : (\mathbb{R}^n) \otimes (\mathbb{R}^n) \to (\mathbb{R}^n) \otimes (\mathbb{R}^n)$ such that $L(\boldsymbol{v}_{\mathcal{B}} \otimes \boldsymbol{w}_{\mathcal{B}}) = \boldsymbol{v}_{\mathcal{C}} \otimes \boldsymbol{w}_{\mathcal{C}}$ for all $\boldsymbol{v}, \boldsymbol{w} \in V$.*

3. *Guess the change of basis formula for $\mathcal{T}^a_b(V)$ for generic a, b. Where would $(\boldsymbol{v}_1)_{\mathcal{B}} \otimes \cdots \otimes (\boldsymbol{v}_a)_{\mathcal{B}} \otimes (\alpha^1)_{\mathcal{B}} \otimes \cdots \otimes (\alpha^b)_{\mathcal{B}}$ be sent to?*

Exercise 1.9.3 (Why should the "gradient" be a row vector)**.** *Take a differentiable function $f : \mathbb{R}^3 \to \mathbb{R}$, say $f(x,y,z) = x^2 + y^2 + z^2$. Calculate the gradient of f.*

Now take a change of basis process such that $\boldsymbol{v}_{new} = B(\boldsymbol{v}_{old})$ where $B = \begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}$. Find the function $f_{new} : \mathbb{R}^3 \to \mathbb{R}$ such that $f_{new}(\boldsymbol{v}_{new}) = f(\boldsymbol{v}_{old})$.

Calculate the gradient of f_{new}, and verify that $\nabla f_{new}(a,b,c) = (B^{-1})^{\mathrm{T}}(\nabla f(x,y,z))$ when $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = B \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.