

Linear Algebra Qual Solutions

Madison Phelps *

September 9, 2022

*Oregon State University, phelpmad@oregonstate.edu

Contents

1	Eigenvectors and Eigenvalues	3
1.1	Problems	3
2	Determinants	10
2.1	Problems	10
3	Inner Product Spaces	12
3.1	Orthogonal Projections	12
3.1.1	Problems	12
3.2	Positive Definite and Positive Semi-Definite	14
3.2.1	Notes: Real Symmetric Matrices	14
3.2.2	Notes: Real Positive Definite Matrices	14
3.2.3	Problems	16
3.2.4	Berkley Problems in Mathematics	17
3.2.5	Fall 2020 #3	18
4	Rank Nullity Theorem	20
4.1	Problems	20
5	Jordan Canonical Form	22
5.1	Problems	22
5.1.1	Hoffman & Kunze, Sec. 7.3	22

1 Eigenvectors and Eigenvalues

1.1 Problems

Problem #3, Jordan Qual Week 1 - Linear Algebra. Suppose T is a linear operator on V , a real vector space. Prove that if every nonzero vector in V is an eigenvector of T , then T is a multiple of the identity.

Solution:

Assume $\dim(V) = n$. Let $\beta = \{\beta_1, \beta_2, \dots, \beta_n\}$ be a basis for V . Then, each $\beta_i \in V$ is an eigenvector of T with eigenvalue λ_i for $i = 1, \dots, n$. That is, $T\beta_i = \lambda_i\beta_i$.

Set $v = \beta_1 + \beta_2 + \dots + \beta_n \in V$. Then v is an eigenvector of T corresponding to eigenvalue $\lambda \in \mathbb{R}$. Observe,

$$\begin{aligned}\lambda\beta_1 + \lambda\beta_2 + \dots + \lambda\beta_n &= \lambda(\beta_1 + \dots + \beta_n) \\ &= \lambda v \\ &= Tv \\ &= T(\beta_1 + \dots + \beta_n) \\ &= T\beta_1 + \dots + T\beta_n \\ &= \lambda_1\beta_1 + \dots + \lambda_n\beta_n.\end{aligned}\tag{1}$$

If $\lambda = 0$ in line (1), then because β is linearly independent, $\lambda = \lambda_i = 0$ for all $i = 1, \dots, n$. Otherwise, λ is non zero and if we subtract line (2) from line (1) we obtain,

$$(\lambda - \lambda_1)\beta_1 + \dots + (\lambda - \lambda_n)\beta_n = 0.$$

Since β is linearly independent, $\lambda = \lambda_1 = \dots = \lambda_n$. Thus, T only has one eigenvalue such that $Tx = \lambda x$ for all $x \in V$ and writing the matrix representation of T with respect to the arbitrary basis β is

$$[T]_\beta = \begin{bmatrix} \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix}$$

which is a multiple of the $n \times n$ identity matrix. Note that if a nonzero term appeared off of the diagonal, then this would imply that there exists a vector in V that is not an eigenvector, which contradicts our assumption that all vectors in V are eigenvectors. Since β was chosen arbitrarily, every matrix representation of T is of the above form. Therefore, T is a multiple of the identity matrix.

Notes: Well, this problem boils down to thinking about the eigenvalue associated to each eigenvector. If every eigenvector is an eigenvector then there is only one true eigenvalue. We show this by taking an arbitrary basis β and then each basis element is itself an eigenvector which has a corresponding eigenvalue. We use the fact that β is linearly independent to show that all eigenvalues must equal. If we have zero, that's also fine because the sum of the $\lambda_i\beta_i$ is a linear combination equaling zero, which forces each $\lambda_i = 0$. Then, it is still needed to show that T is a multiple of the identity operator, or that T is a multiple of the identity matrix.

Another way to look at this is to take two distinct vectors in V (and linearly independent) to show by contradiction that they must be equal because $(\lambda - \lambda_1)v_1 + (\lambda - \lambda_2)v_2 = 0$. Next you can write the matrix representation of T in its Jordan Canonical Form because T only has one eigenvalue, the Jordan Canonical Form could potentially have 1's on the super-diagonal which correspond to generalized eigenvectors. But, these vectors are still in V which implies that there are no 1's on the super-diagonal. We can conclude the same thing that T is a multiple of the identity.

Problem # 3, Jordan Qual Week 1. Let T be a linear operator on a finite dimensional vector space V . Prove that T is a multiple of the identity if and only if T commutes with every other linear operator on V .

Solution: \implies

Suppose T is a multiple of the identity operator for some λ in the underlying field of V . Since the identity commutes with every other linear operator, T commutes with every other linear operator on V .

Solution: \impliedby

Let β be an arbitrary basis for V and let $A = [T]_\beta$ be the matrix representation of T with respect to the basis β .

Define E to be the $n \times n$ matrix with $E_{ii} = 1$ and all other entries are zero for some $i = 1, \dots, n$. We compute,

$$AE = \begin{bmatrix} 0 & \cdots & 0 & A_{1i} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & A_{2i} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & A_{ii} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & A_{ni} & 0 & \cdots & 0 \end{bmatrix}$$

and

$$EA = \begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ A_{i1} & A_{i2} & \cdots & A_{ii} & \cdots & A_{in} \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

In words, multiplying on the left by E results in the matrix whose i -th column is the i -th column of A and all other entries are 0. Multiplying on the right by E results in the matrix whose i -th row is the i -th row of A . Since A commutes with every matrix, $EA = AE$. Also, $i = 1, \dots, n$ is arbitrary in the above calculation and so, we conclude that for all $i \neq j$, $A_{ij} = 0$. That is, A is a diagonal matrix.

Define E to be the matrix with $E_{11} = 1$ and E_{ii} for some $i = 1, \dots, n$, i.e.,

$$E = \begin{bmatrix} 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Since A is diagonal, the eigenvalues are along the diagonal of the matrix. Suppose that one entry along the diagonal is different. Specifically, suppose the i -th entry along the diagonal equals λ_i and every

other entry equals λ such that $\lambda \neq \lambda_i$. We compute,

$$EA = \begin{bmatrix} \lambda & 0 & \cdots & 0 & \lambda_i & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

and

$$AE = \begin{bmatrix} \lambda & 0 & \cdots & 0 & \lambda & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

In words, because A diagonal, multiplying E by A on the right scales the i -th column of E by the A_{ii} entry in A . Then, multiplying E by A on the left scales the i -th row by the entry A_{ii} . Since A commutes with every matrix, $EA = AE$ implies that $\lambda = \lambda_i$ which contradicts our assumption that $\lambda \neq \lambda_i$. Therefore, every entry along the diagonal of A equals λ , and so, A is a multiple of the identity matrix, where A is the matrix representation of the linear operator T with respect to an arbitrary basis β .

Notes: Well, again. Really good problem to look at. Yes, it was not a qual problem but it also points out a tool that we can use whenever we see “for every” something about the problem does ”BLANK” and how we can sometimes use that to our advantage. Also, this problem is a good way to fuss with notation. Of course on the quals you should really label the pictures but at least now you have some notation to fall back on. For example,

- we define E to be a matrix
- entries in E are given by E_{ij} to denote the entry in the i -th row and j -th column of E , i.e., the (i, j) -th entry
- columns of E are denoted by E_j
- while rows of E are denoted by E^i

Just stick with the above notation and redefine E if you are using the matrix in different ways. They will be okay with redefining the matrix because they are mathematicians.

Matrix multiplication has always fucked me up. So, with this one you really had to think about what the matrix E is doing in the first part vs in the second part. Actually, we are talking about the same thing just in the first part E is the diagonal matrix whereas in the second part, A is the diagonal matrix. Think about how a row vector times a matrix is a row vector and a matrix times a column vector is a column vector. So,

$$\begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} \begin{bmatrix} * \\ * \\ * \\ * \\ * \end{bmatrix} = * \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} + * \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} + * \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} + * \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix}$$

In words, multiplying a matrix by a column vector means that each entry of the column vector scales the corresponding column in the matrix by that entry. That sounds weird, the first entry of the column vector scales the first column of A , the second entry in the column vector scales the second column of A , and continues. We can think about the column vector rotating 90 degrees and imposing into the matrix A and then collapsing horizontally to create the resulting column vector $Ax = b$.

On the other side, if we multiply A by a row vector, we are multiplying A on the left by the row vector and similarly we have

$$\begin{aligned}
 \begin{bmatrix} * & * & * & * \end{bmatrix} \begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} &= * \begin{bmatrix} \text{---} \end{bmatrix} \\
 &+ * \begin{bmatrix} \text{---} \end{bmatrix} \\
 &+ * \begin{bmatrix} \text{---} \end{bmatrix} \\
 &+ * \begin{bmatrix} \text{---} \end{bmatrix} \\
 &= \text{the row vector}
 \end{aligned}$$

So, what this does is that the row vector is rotating 90 degrees and then imposing itself on the rows of A to scale each row by the entries in the row vector. Then, we add together by collapsing downwards (vertically) to compute the resulting row vector of $x^T A = b^T$.

So, when we are working with diagonal matrices we do the same thing because all entries off the diagonal are 0 the resulting computation of A times a diagonal matrix on the left will scale the rows by each diagonal entry whereas if we multiply A by the diagonal matrix on the right we will scale each column of A by the corresponding diagonal entry.

Super super cool stuff.

Problem #2, Summer 2021. Let A be an $n \times n$ real matrix. Suppose $\text{tr}(AX) = 0$ for any $n \times n$ real matrix X with $\text{tr}(X) = 0$. Prove that $A = \lambda I_{n \times n}$ for some $\lambda \in \mathbb{R}$.

Solution:

Let $i = 2, \dots, n$ be arbitrary and define E to be the matrix where $E_{11} = 1$ and $E_{ii} = -1$ and otherwise all entries of E are equal to 0. Then,

$$\text{tr}(E) = \sum_{k=1}^n E_{kk} = 1 + 0 + \dots + 0 + (-1) + 0 + \dots + 0 = 0.$$

We compute,

$$\begin{aligned} AE &= \begin{bmatrix} & & & & & \\ & & & & & \\ & & & & & \\ A_1 & A_2 & \cdots & A_i & \cdots & A_n \\ & & & & & \end{bmatrix} \begin{bmatrix} 1 & & & & & \\ & 0 & & & & \\ & & \ddots & & & \\ & & & 0 & & \\ & & & & -1 & \\ & & & & & 0 \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{bmatrix} \\ &= \begin{bmatrix} & & & & & \\ & & & & & \\ & & & & & \\ A_1 & 0 & \cdots & 0 & -A_i & 0 & \cdots & 0 \\ & & & & & & & \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & & & & & & & \\ & 0 & & & & & & \\ & & \ddots & & & & & \\ & & & 0 & & & & \\ & & & & -A_{ii} & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{bmatrix} \end{aligned}$$

where the boxes denote the entries above and below the diagonal and A_k denotes the k -th column of A . Since $\text{tr}(E) = 0$, $\text{tr}(AE) = 0$ which implies that $A_{11} - A_{ii} = 0$, i.e., $A_{11} = A_{ii}$ for all $i = 2, \dots, n$. Let $\lambda \in \mathbb{R}$. Then, the entries along the diagonal of A are all equal to λ .

Now, let E be the matrix with $E_{kk} = 0$ for all $k = 1, \dots, n$ and $E_{ij} = 1$ for some $i \neq j$ and all other entries are equal to 0. In other words, the j -th column of E is a column vector of all zeros except in

the i -th row for $i \neq j$. We compute,

$$\begin{aligned}
 AE &= \begin{bmatrix} & & & & \\ & & & & \\ A_1 & \cdots & A_i & \cdots & A_n \\ & & & & \\ & & & & \end{bmatrix} \begin{bmatrix} & & & & \\ & & & & \\ \mathbf{0} & \cdots & \mathbf{0} & E_j & \mathbf{0} \cdots \mathbf{0} \\ & & & & \\ & & & & \end{bmatrix} \\
 &= \begin{bmatrix} & & & & \\ & & & & \\ \mathbf{0} & \cdots & \mathbf{0} & AE_j & \mathbf{0} \cdots \mathbf{0} \\ & & & & \\ & & & & \end{bmatrix}
 \end{aligned}$$

and compute AE_j (the matrix A times the column vector E_j)

$$AE_j = 0 \cdot A_1 + \cdots + 0 \cdot A_{i-1} + 1 \cdot A_i + 0 \cdot A_{i+1} + \cdots + 0 \cdot A_n = A_i$$

which means that the j -th column of AE equals the i -th column of A , since E_j is the column vector with 1 in the i -th slot for $i \neq j$. That is,

$$A_{1i} = (AE)_{1j}, \dots, A_{ij} = (AE)_{jj}, \dots, \text{ and } A_{1n} = (AE)_{nj}.$$

By the assumption that $\text{tr}(AE) = 0$ for any matrix with $\text{tr}(E) = 0$, we conclude that $A_{ij} = 0$. Since i and j with $i \neq j$ were arbitrary in the above calculation, $A_{ij} = 0$ for all $i \neq j$.

Therefore, $A = \lambda I_{n \times n}$ for some $\lambda \in \mathbb{R}$.

Notes: Welp. I did that. Time to take a break.

2 Determinants

2.1 Problems

Problem #4, Spring 2022. Let a and b be $n \times 1$ column vectors with entries from a field F .

- (a) Verify the $n = 3$ case of the identity $\det(I + ab^t) = 1 + a^t b$.
 (b) Suppose A is an $n \times n$ nonsingular matrix. Assuming the identity in (a), prove that

$$\det(A + ab^t) = (1 + b^t A^{-1} a) \det(A).$$

- (c) Suppose $M = (m_{ij})$ a an $n \times n$ matrix such that

$$m_{ij} = \begin{cases} 1 + d_i & \text{if } i = j, \\ 1 & \text{if } i \neq j, \end{cases} \quad \text{where } d_i \neq 0 \text{ for all } i.$$

Prove that

$$\det(M) = \prod_{i=1}^n d_i \left(1 + \sum_{i=1}^n \frac{1}{d_i} \right).$$

Solution: (a)

USE CRAMER'S RULE and everything cancels. If you have time, type it up. If not, just remember that for a three by three matrix you adjoin the first two columns of A to the end of A and multiple the diagonal entries from top to bottom then, starting at the entry in the first row second column and multiple these entries together, repeat with the last entry in the first column. Do the same but in reverse by multiplying diagonally from the bottom to the top. Add the things you multiplied downwards and subtract the things you multiplied upwards. This will be the determinant.

Solution: (b)

We compute

$$\begin{aligned} \det(A + ab^t) &= \det(A + (AA^{-1})ab^t) && [AA^{-1} = I] \\ &= \det(A(I + A^{-1}ab^t)) \\ &= \det(A) \det(I + (A^{-1}a)b^t) && [\det(CD) = \det(C) \det(D)] \\ &= \det(A) \det\left((I + (A^{-1}a)b^t)^t\right) && [\det(C^t) = \det(C)] \\ &= \det(A) \det(I + b(A^{-1}a)^t) \\ &= (1 + b^t(A^{-1}a)) \det(A). && [\text{Apply part (a) with "a" := b, "b" := } A^{-1}a] \end{aligned}$$

Solution: (c)

Let a be the column vector with $a_i = 1$ for all $i = 1, \dots, n$. Then, aa^t is the $n \times n$ matrix with every entry equal to 1. Let A be the diagonal matrix with entries $A_{ii} = d_i$ for all $i = 1, \dots, n$. Because

$d_i \neq 0$, A is a nonsingular matrix whose inverse, A^{-1} , is also diagonal with $(A^{-1})_{ii} = \frac{1}{d_i}$ for all $i = 1, \dots, n$. Observe that we can write $M = A + aa^t$. By part (b) have that,

$$\det(M) = \det(A + aa^t) = (1 + a^t A^{-1} a) \det(A).$$

So, we compute $\det(A)$ and $a^t A^{-1} a$. Since A is diagonal, the determinant is the product of the eigenvalues (which lie on the diagonal) meaning that,

$$\det(A) = \prod_{i=1}^n A_{ii} = \prod_{i=1}^n d_i,$$

and

$$a^t A^{-1} a = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1/d_1 & & & \mathbf{0} \\ & 1/d_2 & & \\ & & \ddots & \\ & & & \ddots \\ \mathbf{0} & & & & 1/d_n \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1/d_1 \\ 1/d_1 \\ \vdots \\ 1/d_1 \end{bmatrix} = \sum_{i=1}^n 1/d_i.$$

Altogether, the above computations result in

$$\det(M) = \det(A + aa^t) = (1 + a^t A^{-1} a) \det(A) = \prod_{i=1}^n d_i \left(1 + \sum_{i=1}^n \frac{1}{d_i} \right).$$

Notes: This problem is really about looking at it and seeing how the parts can fit together. How you can use the parts to construct a way to solve the last part. Like Tyler keeps on telling me, these problems with multiple parts are really connected and I should think about how to use them to my advantage. This is getting better.

Also, throughout thinking about this problem I drew pictures to help me visualize what in the hell these matrices are doing because on one side of things, they are matrices and on the other, they are numbers. So, I tried using row and column blocks with a matrix to see how I can use the equality in part (a) to help me see what I needed to do to show the equality in part (b). They really messed me up with using a and b because in the second part, the roles are switched... and of course, this makes it harder because its not staring you in the face on “what to choose” to be “clever” and demolish the test like the demagorgans. God, I love stranger things. Anyways, what did you learn? You learned that drawing pictures is helpful. You learned to rely on the things you knew and also, I always forget what to add and subtract in Cramer’s rule. But, at the end of the day you know what you want. So, whatever pops out should be that, right? So, if you’re getting a negative when you should be getting a positive, then you should probably reverse the roles. Also, your scratch work for part (b) was all over the place. Yet, one really good thing that you did is that you didn’t completely scratch your idea when something didn’t instantly work. Just like he says, try and figure out what when wrong and how you can fix the problem to find the solution. I was choosing the wrong a and b which led to working with $(A^{-1})^t$ which we know nothing about. So, you’re going on the right track, but your using the wrong thing to get you there. Take a moment, think about what works and what doesn’t then think about how you can fix these things to get where you want.

Good job.

3 Inner Product Spaces

3.1 Orthogonal Projections

3.1.1 Problems

Problem #5, Fall 2020. Let W be a subspace of an inner product space $(V, \langle \cdot, \cdot \rangle)$ and $P : V \rightarrow V$ be a linear transformation. Prove that P is an orthogonal projection onto W if and only if P is both idempotent and self-adjoint.

Recall: P is

- *idempotent* if $P^2 = P$;
- *self-adjoint* if $\forall x, y \in V \langle x, Py \rangle = \langle Px, y \rangle$;
- *orthogonal projection onto W* if for any $x \in W$, $P(x) = x$ and for any $y \perp W$, $P(y) = 0$.

Solution: \implies

Since P is an orthogonal projection, $P^2 = P$ by definition of a projection operator, meaning P is idempotent. So, we will show that P is also self-adjoint. Let $V = W \oplus W^\perp$ (V is an inner product space with $\text{range } P = W$ and $\text{null } P = W^\perp$). Let $x, y \in V$ be arbitrary and then decompose the vectors as $x = x_1 + x_2$ and $y = y_1 + y_2$ where $x_1, y_1 \in W$ and $x_2, y_2 \in W^\perp$. We compute

$$\begin{aligned} \langle Px, y \rangle &= \langle P(x_1 + x_2), y_1 + y_2 \rangle \\ &= \langle P(x_1) + P(x_2), y_1 + y_2 \rangle \\ &= \langle x_1 + 0, P(y_1) + y_2 \rangle \\ &= \langle x_1, P(y_1) \rangle + \langle x_1, y_2 \rangle \\ &= \langle x_1, P(y_1) \rangle \tag{3} \\ &= \langle x_1, P(y_1) \rangle + \langle x_2, 0 \rangle + \langle 0, P(y_2) \rangle \\ &= \langle x_1 + x_2, P(y_1) + P(y_2) \rangle \\ &= \langle x, Py \rangle \tag{4} \end{aligned}$$

where we have used the properties of inner products and specifically on line (3) we used that $x_1 \in W$ and $y_2 \in W^\perp$ which implies that $\langle x_1, y_2 \rangle$ equals zero, and on line (4) we use the property that $\langle 0, x \rangle = \langle x, 0 \rangle = 0$ with $P(y) = 0$ for $y \in W^\perp$.

Solution: \Leftarrow (By Contraposition)

Suppose P is not an orthogonal projection and let $\text{range } P = W$. Then there exists some $y \in V$ that is not orthogonal to W with $Py = 0$. Since 0 is orthogonal to every vector, $y \neq 0$. Choose any $x \in W$ with $x \notin \text{null } P$. Then $Pw = w \neq 0$. We compute,

$$0 = \langle 0, x \rangle = \langle Py, x \rangle \quad \text{and} \quad \langle y, Px \rangle > 0$$

because $y, Px \neq 0$. Thus $\langle Py, x \rangle \neq \langle y, Px \rangle$ which shows that P is not self-adjoint.

Also, if P is not a projection, then $P^2 \neq P$ because there exists some $w \in \text{range } P = W$ such that $Pw = v$ for $v \neq w$ and so, $P^2(w) \neq P(w)$. That is, if P is not a projection then P is not idempotent.

Therefore, by contraposition if P is idempotent and self-adjoint then P is an orthogonal projection.

Notes: Well... yeah. You did work too hard in the beginning of this problem. But you chose it because you needed to work on this particular skill. The trick was to add by zero and then use proof by contraposition instead of a direct proof. Also, drawing pictures will help you out. Understanding what a projection is will also help because projections are defined to be idempotent. So, use this. Also, be able to prove it if you need to. For example, let $x = x_1 + x_2$ with $x_1 \in W$ and $x_2 \in W^\perp$. Then, $P(x_1) = x_1$ and $P(x_2) = 0$ and

$$\begin{aligned} Px &= P(x_1 + x_2) \\ &= P(x_1) + P(x_2) \\ &= P(P(x_1)) + 0 \\ &= P^2(x_1) + P(0) \\ &= P^2(x_1) + P(P(x_2)) \\ &= P^2(x_1) + P^2(x_2) \\ &= P^2(x_1 + x_2) \\ &= P^2(x) \end{aligned}$$

for all $x \in V = W \oplus W^\perp$.

Note to self, try and work out what these questions mean in words and try and draw pictures of what happens. Then, if that doesn't go anywhere do symbol mashing but symbol mashing normally gets you into trouble.

Keep it up.

3.2 Positive Definite and Positive Semi-Definite

3.2.1 Notes: Real Symmetric Matrices

If a matrix is real symmetric then the matrix is self adjoint and if V is an inner product space then

$$\langle Ax, y \rangle = \langle x, A^* y \rangle = \langle x, A^t y \rangle = \langle x, Ay \rangle.$$

Properties of self-adjoint real matrices over some field F :

- The matrix is symmetric
- Adding and subtracting symmetric matrices preserves the symmetry. That is A and B symmetric, then $A \pm B$ is symmetric.
- Let A and B be symmetric matrices. Then A and B commute, i.e., $AB = BA$ if and only if AB is symmetric: If A and B commute, then we compute

$$(AB)^t = (BA)^t = A^t B^t = AB$$

which shows that AB is symmetric. Conversely, since AB is symmetric and we compute

$$AB = (AB)^t = B^t A^t = BA$$

which shows that A and B commute.

Assuming we are in a real finite dimensional inner product space:

- If A is the matrix representation of a linear operator on a finite dimensional real inner product space where A is symmetric. Then, $A = A^t$, meaning that A is equal to its adjoint (which is the conjugate transpose), i.e., A is self-adjoint. A useful theorem is that A is self-adjoint if and only if there exists an orthonormal basis of eigenvectors. In particular, the orthonormal basis diagonalizes A . Called the **Orthonormal Diagonalizability Theorem**.
- Every eigenvalue is real and the characteristic polynomial splits.

3.2.2 Notes: Real Positive Definite Matrices

Real Positive Definite Matrix: A real matrix is called positive definite if (1) the matrix is symmetric and (2) $x^T Ax > 0$ for all **nonzero** column vectors $x \in \mathbb{R}^n$.

List of equivalences that show that a real matrix is positive definite:

- The matrix is symmetric and all its eigenvalues are positive. ¹

¹If all of the eigenvalues exist and are (real) positive, there exists a full set of eigenvectors that form a basis for \mathbb{R}^n . Also something to keep in mind is that given any basis we can perform the Gram-Schmidt process to turn any basis into an orthogonal basis if we have an inner product space. But, remember we need to be able to use the notion of inner products to do this. Continuing... we can write any vector as a linear combination of the eigenvectors and if x is an eigenvector of A , $Ax = \lambda x$ and $x^T Ax = x^T \lambda x = \lambda x^T x$. Because $\lambda > 0$, $x^T x > 0$ and we conclude that $x^T Ax > 0$ for every eigenvector. Then, we can extend this result to any linear combination being positive and so, for any nonzero vector in \mathbb{R}^n , $x^T Ax > 0$. Sweet.

- If A is congruent to a diagonal matrix, i.e., there exists an invertible matrix P such that we can write $P^T A P = D$ where D is a diagonal matrix.²
- You can use Sylvester's criterion to show that the matrix is positive definite. Sylvester's criterion states that if all of the leading principal minors are positive then the matrix is positive definite.³
- If there exists an invertible matrix B such that $A = B^T B$, then A is positive definite.

Inner products and real symmetric matrices: How to use them

Well, first consider the standard inner product on \mathbb{R}^n which is given by the dot product. So, for all column vectors $x, y \in \mathbb{R}^n$ we define

$$\langle x, y \rangle = x^T \cdot y = x^T y.$$

How can we use this? Well, given a symmetric matrix A define the inner product

$$\langle x, y \rangle_A = \langle Ax, y \rangle$$

which is an inner product. To see this,

- Let $x_1, x_2, y \in \mathbb{R}^n$. We compute,

$$\begin{aligned} \langle x_1 + x_2, y \rangle_A &= \langle A(x_1 + x_2), y \rangle \\ &= \langle Ax_1 + Ax_2, y \rangle \\ &= \langle Ax_1, y \rangle + \langle Ax_2, y \rangle \\ &= \langle x_1, y \rangle_A + \langle x_2, y \rangle_A \end{aligned}$$

- For any scalar $\alpha \in \mathbb{R}$, $(\alpha Ax)^T y = \alpha x^T A y$, meaning that $\alpha \langle x, y \rangle_A = \langle \alpha x, y \rangle_A$.
- $\langle x, x \rangle_A = \langle Ax, x \rangle \leq 0$ for all $x \in \mathbb{R}^n$ and we obtain equality if and only if $x = 0$.
- To see that the inner product is symmetric:

$$\langle x, y \rangle_A = \langle Ax, y \rangle = x^T A y = (Ax)^T y = y^T A x = \langle Ay, x \rangle = \langle y, x \rangle_A$$

where we used that $(BA)^T = A^T B^T$ (the socks and shoes of the transpose) to get the inner equality.

Or, you could simply state that $\langle \cdot, \cdot \rangle_A$ is an inner product because A is a linear operator that respects the linearity of inner products in the first component as well as scalar multiples, and using the property that $(BA)^T = A^T B^T$. Also, if A is not symmetric then this does not hold because the proof hinges on the fact that $A^T = A$. This is a cool trick.

²In particular, because A is self-adjoint and is the matrix representation of some linear operator T on a finite dimensional inner product space there exists an orthonormal basis of eigenvectors that diagonalize A . That is, P is the matrix whose columns equal the orthonormal basis of eigenvectors. On the diagonal, we obtain $x^T A x = \lambda x^T x = \lambda$ since x is orthonormal, $x^T x = 1$.

³The leading principal minors is a fancy word for taking the determinant of the matrix starting in the upper right corner and including one row at a time. That is, the first entry must be positive, then the determinant of the 2×2 matrix formed in the upper right corner must be positive... continue till you conclude that all leading principal minors are positive. Also, note that positive implies that the determinant is not equal to zero.

3.2.3 Problems

Problem #1, (from notes above). Let V be a finite dimensional real inner product space. Then A is positive definite if and only if there exists an invertible matrix B such that $A = B^T B$.

\implies Assume A is positive definite.

Let λ be an eigenvalue of A . Because $\langle Ax, x \rangle > 0$ for all nonzero vectors x and

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Ax, x \rangle > 0,$$

we conclude that $\lambda > 0$, meaning that all the eigenvalues of A are positive and real. Given that A is symmetric, A is self-adjoint. So, by the real spectral theorem there exists a full basis of orthonormal eigenvectors given by x_1, \dots, x_n for A corresponding to eigenvalues λ_i for $i \in \{1, \dots, n\}$.

Let P be the matrix whose columns are the orthonormal eigenvectors of A as defined above. Then, P is invertible and we compute $P^T A P$ which results in the matrix whose entries equal $\langle Ax_i, x_j \rangle$ for $i, j \in \{1, \dots, n\}$. If $i = j$, then

$$\langle Ax_i, x_j \rangle = \langle \lambda_i x_i, x_j \rangle = \lambda_i$$

and otherwise for distinct i and j , because the basis of eigenvectors is orthonormal, the columns of P are orthogonal and so, $\langle Ax_i, x_j \rangle = 0$ when $i \neq j$. Therefore, A is congruent to a diagonal matrix whose entries are equal to λ_i . Let D denote the constructed diagonal matrix. Because $\lambda_i > 0$ the square root of each eigenvalue exists and is positive, we may write $D = \sqrt{D}\sqrt{D}$, and note that \sqrt{D} is also a diagonal matrix, meaning that it is equal to its transpose.

Define $B = \sqrt{D}P^{-1}$. Then B is invertible because the entries in the diagonal matrix are nonzero and by multiplying a diagonal matrix on the left scales the rows of P^{-1} by λ_i . Since, P is invertible, P^{-1} is invertible with inverse P . Also, observe that, $P^T A P = D$ implies that $A = (P^{-1})^T D P^{-1}$ and we compute,

$$B^T B = (\sqrt{D}P^{-1})^T \sqrt{D}P^{-1} = (P^{-1})^T \sqrt{D}\sqrt{D}P^{-1} = (P^{-1})^T D P^{-1} = A.$$

Therefore, if A is positive definite, there exists an invertible matrix B such that $A = B^T B$ where B^T is the matrix transpose of B .

\Leftarrow Observe that

$$A^T = (B^T B)^T = B^T (B^T)^T = B^T B = A$$

implies that A is a real symmetric matrix. Also, we compute

$$\langle Ax, x \rangle = \langle B^T Bx, x \rangle = \langle Bx, Bx \rangle \geq 0$$

and see that the strict inequality holds if and only if $Bx \neq 0$. Since we can assume that B is a nonzero matrix (otherwise we are just dealing with the zero matrix) and so, x must be a nonzero vector. Therefore, A is symmetric and $\langle Ax, x \rangle > 0$ for all nonzero vectors $x \in V$.

3.2.4 Berkley Problems in Mathematics

Problem 7.8.9. An $n \times n$ real matrix T is positive definite if T is symmetric and $\langle Tx, x \rangle > 0$ for all nonzero vectors $x \in \mathbb{R}^n$, where $\langle u, v \rangle$ is the standard inner product. Suppose that A and B are two positive definite real matrices.

- (a) Show that there is a basis $\{v_1, \dots, v_n\}$ of \mathbb{R}^n and real numbers $\lambda_1, \dots, \lambda_n$ such that for $1 \leq i, j \leq n$:

$$\langle Av_i, v_j \rangle = \begin{cases} 1 & i = j, \\ 0 & i \neq j \end{cases}$$

and

$$\langle Bv_i, v_j \rangle = \begin{cases} \lambda_i & i = j, \\ 0 & i \neq j \end{cases}$$

- (b) Deduce from part (a) that there is an invertible real matrix U such that $U^T A U = I$ and $U^T B U$ is diagonal.

Solution: (a) Define an inner product as

$$\langle x, y \rangle_A = \langle Ax, y \rangle$$

where $\langle u, v \rangle$ is the standard inner product on \mathbb{R}^n . Then, because A is symmetric and a linear operator $A(\alpha x_1 + \beta x_2) = \alpha Ax_1 + \beta Ax_2$ we conclude that $\langle \cdot, \cdot \rangle_A$ is linear in the first component. To see that $\langle \cdot, \cdot \rangle_A$ has symmetry, we compute

3.2.5 Fall 2020 #3

Problem #3, Fall 2020. A real symmetric $n \times n$ matrix is called positive semi-definite if $x^t Ax \geq 0$ for all $x \in \mathbb{R}^n$. Prove that A is positive semi-definite if and only if $\text{tr}(AB) \geq 0$ for every real symmetric positive definite $n \times n$ matrix B .

Solution:

\Rightarrow Since A is a real symmetric $n \times n$ matrix, A is self-adjoint, and by the real spectral theorem there exists an orthonormal basis v_1, \dots, v_n of eigenvectors corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$ that diagonalizes A . In particular, let U denote the orthogonal matrix whose columns, $U_j = v_j$ for $j \in \{1, \dots, n\}$. Then, $A = U^T D U$ where $U^T = U^{-1}$ and D is the matrix with $\text{diag}(\lambda_1, \dots, \lambda_n)$. Let B be any real symmetric positive semi-definite matrix, and define $E = U B U^T$. Then, E is symmetric positive semi-definite because $E^T = (U B U^T)^T = U B U^T = E$ and for all $x \in \mathbb{R}^n$,

$$\langle Ex, x \rangle = \langle U B U^T x, x \rangle = \langle B U^T x U^T x \rangle = \langle B y, y \rangle \geq 0.$$

Observe that,

$$AB = (U^T D U)(U^T E U) = U^{-1} D U U^{-1} E U = U^{-1} D E U,$$

meaning that AB is similar to DE and because similar matrices have the same trace, it follows that

$$\text{tr}(AB) = \text{tr}(DE) = \sum_{i=1}^n \sum_{j=1}^n D_{ij} E_{ji} = \sum_{i=1}^n \lambda_i E_{ii}$$

(because D is a diagonal matrix). Since A is symmetric positive semi-definite, for each $i \in \{1, \dots, n\}$,

$$\lambda_i = \langle \lambda_i v_i, v_i \rangle = \langle A v_i, v_i \rangle \geq 0$$

and by letting e_1, \dots, e_n denote the standard orthonormal basis of \mathbb{R}^n because E is symmetric positive semi-definite,

$$E_{ii} = \langle E e_i, e_i \rangle \geq 0$$

for all $i \in \{1, \dots, n\}$. Therefore,

$$\text{tr}(AB) = \sum_{i=1}^n \lambda_i E_{ii} \geq 0$$

for every symmetric positive semi-definite matrix B .

\Leftarrow (by contraposition) Suppose A is a real symmetric matrix that is not positive semi-definite. Since A is symmetric, A is self-adjoint and we can let U be the same orthonormal matrix as defined above where $U_j = v_j$ with v_j orthonormal eigenvectors of A . In particular, there exists some $i \in \{1, \dots, n\}$ such that $\lambda_i < 0$ because every symmetric positive semi-definite matrix has non-negative eigenvalues. Let $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ as defined in the forward direction of this proof and write $A = U^T D U$.

Define E to be the matrix with $E_{ii} = 1$ and otherwise all entries equal 0. Then, E is symmetric and positive definite because for any $x \in \mathbb{R}^n$,

$$\langle Ex, x \rangle = x^T E x = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ x_i \\ \vdots \\ 0 \end{bmatrix} = x_i^2 \geq 0.$$

Let $B = U^T E U$. Then, B is symmetric and positive definite (shown in the forward direction). We also have that AB is similar to DE and,

$$\operatorname{tr}(AB) = \operatorname{tr}(DE) = \sum_{j=1}^n \lambda_j E_{jj} = \lambda_i E_{ii} = \lambda_i < 0.$$

Therefore, if A is not positive semi-definite, then there exists a symmetric positive semi-definite matrix B such that $\operatorname{tr}(AB) < 0$.

4 Rank Nullity Theorem

4.1 Problems

Problem #2, Spring 2022.

- (a) Let V be a vector space of dimension 2 over field F . Let T be a linear operator on V . Prove that the null space of T is identical to the image of T if and only if their intersection contains a nonzero vector.
- (b) Let $V = \mathbb{C}$, and $F = \mathbb{R}$. Construct a linear operator T on V such that the null space of T is identical to the image of T . Justify your example.

Solution: (a)

\Rightarrow By the rank-nullity theorem and our assumption that the null space of T is identical to the image of T

$$2\dim V = \text{nullity}T + \text{rank}T = 2\text{nullity}T$$

which implies that $\text{nullity}T = 1$. Thus, there exists a nonzero vector $x \in \text{null}T$ such that $Tx = 0$ and $x \in \text{range}T$. Therefore, there exists a nonzero vector in their intersection.

\Leftarrow (By Contraposition) Suppose the null space of T and the range of T are different. Then by the rank-nullity theorem there are three cases to consider. If $\text{nullity}T = 0$, then $\text{rank}T = 2$ which implies their intersection only contains the zero vector. Similarly, we conclude the same if $\text{nullity}T = 2$ and $\text{rank}T = 0$. If $\text{nullity}T = \text{rank}T = 1$, then there is a nonzero vector $x \in \text{null}T$ that is not in $\text{range}T$ such that $Tx = 0$ and there exists a nonzero vector $y \in \text{range}T$ that is not in the null space. Since the dimension of V is 2, their intersection only contains the zero vector.

Therefore, the null space of T is identical to the image of T if and only if their intersection contains a nonzero vector.

Solution: (b)

Because \mathbb{C} is the vector space over the real numbers, we need two real numbers to write a complex number of the form $x = a + ib$ for $a, b \in \mathbb{R}$ which means that \mathbb{C} is a two-dimensional vector space over the field \mathbb{R} . Define $T : \mathbb{C} \rightarrow \mathbb{C}$ by $Tx = \text{Im}(x)$ (the imaginary part of x). Since $\text{Im}(x) \in \mathbb{R} \subseteq \mathbb{C}$, T is an operator. To see that T is a linear operator, let $x = a_1 + ib_1 = (a_1, b_1)$, $y = a_2 + ib_2 = (a_2, b_2)$, and $\alpha, \beta \in \mathbb{R}$ and we compute

$$\begin{aligned} T(\alpha x + \beta y) &= \text{Im}(\alpha x + \beta y) \\ &= \text{Im}(\alpha a_1 + i(\alpha b_1) + \beta a_2 + i(\beta b_2)) \\ &= \text{Im}(\alpha a_1 + \beta a_2 + i(\alpha b_1 + \beta b_2)) \\ &= \alpha b_1 + \beta b_2 \\ &= \alpha \text{Im}(x) + \beta \text{Im}(y) \\ &= \alpha Tx + \beta Ty. \end{aligned}$$

Now, all complex numbers of the form $x = a$ will get mapped to zero which means that the set of all real numbers is the null space of T , but is also the range of T because the imaginary part of a complex

number is real. Thus, we have constructed a linear operator on \mathbb{C} over \mathbb{R} such that the null space is identical to the range.

Notes: Well, I think you should keep this because it doesn't really make sense. Or... maybe, its just that the notation is all over the place. You are trying to communicate the correct idea but I don't think that the notation is helping convey this. Come back later and read this and maybe your thoughts will be different. Here's an alternate proof of part (b):

Since \mathbb{C} is a vector space over \mathbb{R} , for any $x \in \mathbb{C}$ we write $x = a + ib = (a, b)$ (any complex number is given by the ordered pair (a, b) for $a, b \in \mathbb{R}$, which means that the vector space of complex numbers over the field \mathbb{R} is a two-dimensional vector space. Define $T : \mathbb{C} \rightarrow \mathbb{C}$ by $(a, b) \mapsto (b, 0)$, meaning T is the mapping that gives the imaginary part of a complex number. The null space of T is the set of all pairs of real numbers of the form $(a, 0)$ which is exactly the range of T . In other words, the null space and the image of T is the one-dimensional space of all real numbers which is a subset of \mathbb{C} . To see that T is a linear operator, let $x = a_1 + ib_1 = (a_1, b_1)$, $y = a_2 + ib_2 = (a_2, b_2)$, and $\alpha, \beta \in \mathbb{R}$ and we compute

$$\begin{aligned} T(\alpha x + \beta y) &= \text{Im}(\alpha x + \beta y) \\ &= \text{Im}(\alpha a_1 + i(\alpha b_1) + \beta a_2 + i(\beta b_2)) \\ &= \text{Im}(\alpha a_1 + \beta a_2 + i(\alpha b_1 + \beta b_2)) \\ &= \alpha b_1 + \beta b_2 \\ &= \alpha \text{Im}(x) + \beta \text{Im}(y) \\ &= \alpha T x + \beta T y. \end{aligned}$$

And with that. Goodnight.

5 Jordan Canonical Form

5.1 Problems

5.1.1 Hoffman & Kunze, Sec. 7.3

Problem #1. Let N_1 and N_2 be 3×3 nilpotent matrices over the field F . Prove that N_1 and N_2 are similar if and only if they have the same minimal polynomial.

Solution: \Rightarrow

Let $k_1 \in \mathbb{N}$ be the smallest positive integer for which N_1 is nilpotent and let k_2 be the smallest positive integer for which N_2 is nilpotent. Since N_1 and N_2 are similar, there exists an invertible matrix Q such that

$$N_1 = QN_2Q^{-1}.$$

Then we compute,

$$N_1^{k_1} = (QN_2Q^{-1})^{k_1} = QN_2^{k_1}Q^{-1}$$

and see that

$$N_2^{k_1} = Q^{-1}N_1^{k_1}Q = 0 = N_2^{k_2}.$$

If $k_1 < k_2$ then k_2 is not the smallest positive integer such that N_2 is nilpotent. If we replace 1 with 2 in our computation above, we obtain that

$$N_1^{k_2} = Q'^{-1}N_2^{k_2}Q' = 0 = N_1^{k_1}$$

and if $k_2 < k_1$, then k_1 is not the smallest positive integer for which N_1 is nilpotent. Therefore, $k_1 = k_2 = k$. Since each matrix is nilpotent, the eigenvalues are all equal to zero and so, the minimal polynomial of N_1 is $m_1(x) = x^k$ and the minimal polynomial of N_2 is $m_2(x) = x^k$. That is, if N_1 and N_2 are similar, then the minimal polynomials are equal.

Solution: \Leftarrow

Suppose the minimal polynomials of N_1 and N_2 are the same. Since each matrix is nilpotent, the eigenvalues are all equal to zero and because F is a field which must contain the zero element, we conclude that each nilpotent matrix contains all its eigenvalues and the characteristic polynomial splits over F . Then N_1 and N_2 are similar to a 3×3 matrix in Jordan form. There are three cases to consider:

- (1) Suppose $m(x) = x$ is the minimal polynomial of N_1 and N_2 . By the Cayley-Hamilton theorem, $m(N_1) = m(N_2) = 0$ which implies that N_1 and N_2 are the zero matrix. Thus, N_1 and N_2 are similar.
- (2) Suppose $m(x) = x^2$ is the minimal polynomial of N_1 and N_2 . Then, N_1 is similar to a 3×3 matrix in the following Jordan form:

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

Since the matrices share the same minimal polynomial and are 3×3 matrices, there is no other matrix in Jordan form corresponding to the minimal polynomial $m(x) = x^2$ since the degree of the minimal polynomial is given by the size of the largest elementary Jordan block. Thus, N_2 is similar to the matrix above and we conclude that N_1 is similar to N_2 .

- (3) Suppose $m(x) = x^3$ is the minimal polynomial of N_1 and N_2 . Since the degree of the minimal polynomial is the size of the largest elementary Jordan block, N_1 is similar to a 3×3 elementary Jordan block, and N_2 is similar to a 3×3 elementary Jordan block. So,

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

is the only matrix in Jordan form with minimal polynomial equaling x^3 , and we conclude that N_1 and N_2 are similar.

Also note that if N_1 and N_2 are 4×4 nilpotent matrices over the field F then the claim does not hold. That is, let A_1 be the matrix with one 2×2 elementary Jordan block and two 1×1 elementary Jordan blocks and let A_2 be the nilpotent matrix with two 2×2 elementary Jordan blocks, i.e.,

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We observe that the minimal polynomial of each nilpotent matrix equals x^2 . In particular, similar matrices have the same Jordan form (up to permutation of blocks) which implies that A_1 is not similar to A_2 because the Jordan forms are different.

Problem #2. Let A and B be $n \times n$ matrices over the field F which have the same characteristic polynomial

$$f(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_k)^{r_k}$$

and the same minimal polynomial. If no r_i for $i \in \{1, \dots, k\}$ is greater than 3, then A and B are similar.