Math 115AH Homework 7

Brendan Connelly

Tuesday, May 28, 2024

1 Reflection Transformation

Consider the linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by reflection across the line

$$L = \{(x, y, z) \in \mathbb{R}^3 \mid x = y,\, z = 0\}.$$

Concretely, this equation is given as follows: for any vector $v \in \mathbb{R}^3$, there exist unique vectors $x \in L$ and y perpendicular to L such that $v = x + y$. Then $T(v) = x - y$. In words, T does not change the component of x parallel to L, and negates the component of v orthogonal to L. Let $\beta = \{(1, 1, 0),\ (1, -1, 0),\ (0, 0, 1)\}$.

a. Check that $(1, 1, 0)$ spans L and that $(1, -1, 0)$ and $(0, 0, 1)$ are orthogonal to $(1, 1, 0)$ (and therefore are orthogonal to the entire line L).

First, let's check that $(1, 1, 0)$ spans L. We have that

$$\text{span}\{(1, 1, 0)\} \Leftrightarrow \{(a, a, 0) \mid a \in \mathbb{R}\} \Leftrightarrow \{(x, y, z) | \, x = y,\, z = 0\} \Leftrightarrow L$$

Using the geometrically defined dot product, we can check that $(1, -1, 0)$ and $(0, 0, 1)$ are orthogonal to $(1, 1, 0)$. We see

$$(1, -1, 0) \cdot (1, 1, 0) = 1 - 1 = 0 \qquad \text{and} \qquad (0, 0, 1) \cdot (1, 1, 0) = 0$$

Thus, we have shown that $(1, 1, 0)$ spans L and that $(1, -1, 0)$ and $(0, 0, 1)$ are orthogonal to $(1, 1, 0)$.

b. Compute $[T]_\beta$.

We can see that by definition, $T(v) = x - y$ where x is the component parallel to L and y the component orthogonal. Thus, because $(1, 1, 0)$ is on the line entirely, $T(1, 1, 0) = (1, 1, 0)$. And because the other two vectors are not, they wil take the role of y in our definition and thus, $T(1, -1, 0) = -(1, -1, 0)$ and $T(0, 0, 1) = -(0, 0, 1)$. Thus, we can represent this transformation of the basis vectors in β coordinates as keeping the first one and negating the second two. This produces

$$[T]_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

c. Let $S = \{(1, 0, 0),\ (0, 1, 0),\ (0, 0, 1)\}$ be the standard basis for $\mathbb{R}^3$. Compute $Q^{-1} = [I_{\mathbb{R}^3}]_\beta^S$.

This one follows obviously from the coordinates of β itself. We can just stack β's vectors as columns and then converting them to standard coordinates leaves them untouched.

$$Q^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

d. Compute $Q = ([I_{\mathbb{R}^3}]_S^\beta$.

We can do row operations to find the inverse reasonably easily.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

e. Recall that the change of basis formula states that $[T]_S = Q^{-1}[T]_\beta Q$. Use the change of basis formula and the previous items to compute $[T]_S$.

We therefore want to do the following, plugging in the matrices we found in the previous sections.

$$\begin{aligned} [T]_S &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

□

---

2 Determinants of triangular matrices

Recall that a $n \times n$ matrix A with entires in F is *upper-triangular* if $A_{ij} = 0$ whenever $i > j$. A is *lower triangular* if $A_{ij} = 0$ whenever $j > i$.
a. Using induction, prove a general formula for the characteristic polynomial of an upper-triangular matrix.

I claim that for an upper triangular matrix, $f_A(t) = \prod_{i=1}^{n}(t - A_{ii})$.
Base Case: (n = 1)
Let's check the base case. This would just be a single entry matrix which by definition has a determinant of itself.

$$\det([t - A_{11}]) = t - A_{11} = \prod_{i=1}^{1}(t - A_{ii})$$

We can also check the $n = 2$ base case. This by definition is

$$\det \begin{pmatrix} t - A_{11} & A_{12} \\ 0 & t - A_{22} \end{pmatrix}$$

This, we can see, has $f_2(t) = \prod_{i=1}^{2}(t - A_{ii})$
We have thus shown our base case.
Inductive Step: Assume $f_{n-1}(t) = \prod_{i=1}^{n-1}(t - A_{ii})$.
Now, let's consider $f_n(t) = \det(tI_n - A) = \det \begin{pmatrix} D & B \\ 0 & C \end{pmatrix}$ where D is some $n - 1 \times n - 1$ square upper-triangular matrix and C is 1×1 square matrix equal to the A_{nn} entry of $(tI_n - A)$. B is the rightmost column and 0 represents the row of $n - 1$ zeros from being upper triangular. By the the stated property in discussion, we have

$$\begin{aligned} f_n(t) &= \det(tI_n - A) \\ &= \det \begin{pmatrix} D & B \\ 0 & C \end{pmatrix} \\ &= \det(D)\det(C) && \text{from discussion} \\ &= \left(\prod_{i=1}^{n-1}(t - A_{ii})\right)\det(t - A_{nn}) && \text{inductive step} \\ &= \prod_{i=1}^{n}(t - A_{ii}) && \text{properties of products} \end{aligned}$$

We have thus shown our inductive and base case. Therefore, we have proven our argument. □

b. Using properties of matrices given in class, give a formula for the characteristic polynomial of a lower-triangular matrix.

Recall that a $n \times n$ matrix A with entires in F is *upper-triangular* if $A_{ij} = 0$ whenever $i > j$. A is *lower triangular* if $A_{ij} = 0$ whenever $j > i$. Taking the transpose would send an uppper-triangular matrix A to one where $A_{ji} = 0$ whenever $i > j$. This is the exact definition of a lower triangular matrix. We have proven the determinant in closed under transposition, i.e. $A = A^t$ for arbitrary square matrix A. Thus, we have the same formula.

$$f_A(t) = \prod_{i=1}^{n}(t - A_{ii})$$

c. Textbook 5.2 #9.

We want to prove that the characteristic polynomial for some linear operator T with some basis β splits if $[T]_\beta$ is upper-triangular. Our formula for the characteristic polynomial involves $[T]_\beta$. If we just suppose $A = [T]_\beta$ where A is the matrix we worked with in our previous problem, we have that $[T]_\beta = \prod_{i=1}^{n}(t - A_{ii})$. This directly implies that the characteristic polynomial of T is the product of linear terms and thus splits. Part B of this textbook question has been trivially proven by the previous parts of this question; the characteristic polynomial would by the formula be the product of linear terms and thus split. □

---

3 Similarity and Determinants

Suppose that A and B are both square matrices with entries in a field F.
a. Prove that if A is similar to B, then $\det A = \det B$. (Hint: use the definition of similarity and basic facts about determinants stated in class/ in section 4.4 of the textbook.)

We have that for some $Q \in M_{n\times n}(F)$ invertible, $A = Q^{-1}BQ$. We know then that

$$\begin{aligned}
\det(A) &= \det(Q^{-1}BQ) && \text{def. of similarity}\\
&= \det(Q^{-1})\det(B)\det(Q) && \text{prop. of determinants}\\
&= \frac{1}{\det(Q)}\det(B)\det(Q) && \text{determinant of inverse}\\
&= \det(B) && \text{field axioms}
\end{aligned}$$

b. Are the following two matrices A and B with entries in $\mathbb{R}$ similar?

$$A = \begin{pmatrix} 0 & 0 & 1\\ 0 & 1 & 1\\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 0\\ 0 & 4 & 0\\ 0 & 0 & 1 \end{pmatrix}$$

Using the fact that we will prove in the next subpart–that similar matrices have the same characteristic polynomial and thus same eigenvalues–we will calculate the eigenvalues of A. We can compare these values with the eigenvalues we can read off of the diagonal matrix B, those being, $2, 4, 1$. Consider

$$\begin{aligned}
f_A(t) &= \det\begin{pmatrix} t & 0 & -1\\ 0 & t-1 & -1\\ 0 & -1 & t \end{pmatrix}\\
&= t\det\begin{pmatrix} t-1 & -1\\ -1 & t \end{pmatrix}\\
&= t(t^2 - t - 1)
\end{aligned}$$

We can see that $t = 2, 4, 1$ are not roots to $t(t^2 - t - 1)$ and thus, the two matrices are $\boxed{\text{not similar}}$. Additionally, not the row of zeros in A. This means that $\det(A) = 0$ while from the product of the entries $\det(B) = 8 \implies A \not\sim B$

c. Prove that if A is similar to B, then $\det(tI_n - A) = \det(tI_n - B)$.

Let's again start with $A = Q^{-1}BQ$ for some $Q \in M_{n\times n}(F)$ invertible by the definition of similarity. Now consider

$$\begin{aligned}
\det(tI_n - A) &= \det(tI_n - Q^{-1}BQ) && \text{substitution} \\
&= \det(tQ^{-1}I_nQ - Q^{-1}BQ) && \text{def. and prop. of identity and inverse matrix} \\
&= \det(Q^{-1}(I_nt - B)Q) && \text{distributivity} \\
&= \det(Q^{-1})\det(I_nt - B)\det(Q) && \text{prop. of determinatnts} \\
&= \det(I_nt - B) && \text{determinants of inverses}
\end{aligned}$$

Thus, we have shown that if A is similar to B, then $\det(tI_n - A) = \det(tI_n - B)$. □

d. Let V be an n-dimensional vector space and let $T : V \to V$ be a linear transformation. Prove that the characteristic polynomial of T does not depend on the choice of matrix representation: if β and γ are two bases for V, then $f_{[T]_\beta}(t) = f_{[T]_\gamma}(t)$.

We essentially have already proven the foundation for this statement. $f_{[T]_\beta}(t) = f_{[T]_\gamma}(t)$ is equivalent to stating that the the matrices $[T]_\beta$ and $[T]_\gamma$ are similar by the previous part of the this question. We by definition have that they are similar as there exists a $Q = [\mathbb{1}_{F^n}]_\gamma^\beta$ and a $Q^{-1} = [\mathbb{1}_{F^n}]_\beta^\gamma$. This comes from the definition of a basis and coordinate representations–that the identity matrix is invertible. Thus, because $[T]_\beta$ and $[T]_\gamma$ are similar, the characteristic polynomial is preserved by part c. Thus, $f_{[T]_\beta}(t) = f_{[T]_\gamma}(t)$. □

4 Textbook 5.1 Question 2(a),(c): Characteristic Polynomial Computation

a. Let's plug in our standard basis vectors to compute $[T]_S$. We see that by definition, $T(1, 0) = (2, 5)$ and that $T(0, 1) = (-1, 3)$. Thus,

$$[T]_S = \begin{pmatrix} 2 & -1 \\ 5 & 3 \end{pmatrix}$$

Then,

$$f_T(t) = \det\begin{pmatrix} t-2 & 1 \\ -5 & t-3 \end{pmatrix}$$

This produces

$$f_T(t) = (t-2)(t-3) + 5 = t^2 - 5t + 11$$

C. Let's again plug in our standard basis vectors to compute $[T]_S$. To take a little less time and space to write up, we will then immediately put the results in the columns of said matrix. This gives us

$$[T]_S = \begin{pmatrix} 1 & 0 & -1 & 0 \\ -1 & 1 & 0 & 1 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

Then, let's calculate

$$\begin{aligned}\det(tI_n - [T]_S) &= \det\begin{pmatrix} t-1 & 0 & 1 & 0 \\ 1 & t-1 & 0 & -1 \\ -1 & -1 & t & 1 \\ 0 & 0 & 1 & t \end{pmatrix} \\ &= t\cdot\det\begin{pmatrix} t-1 & 0 & 1 \\ 1 & t-1 & 0 \\ -1 & -1 & t \end{pmatrix} - \det\begin{pmatrix} t-1 & 0 & 0 \\ 1 & t-1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \\ &= t\cdot\big((t-1)(t^2-t) - 1 + t - 1\big) - (t-1)(t-2) \\ &= t(t^3 - 2t^2 + 2t - 2) - (t^2 - 3t + 2) \\ &= t^4 - 2t^3 + t^2 + t - 2\end{aligned}$$

---

5 Textbook 5.1 Question 3(a),(f): Coordinate Representation and Eigenbasis Calculation

a. Let's check the first vector in the formula $T(a, b) = (10a - 6b, 17a - 10b)$. Plugging in

$$T\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}-2\\-3\end{pmatrix}$$

We see that this does not equal some $\lambda \in F$ scaling $\langle 1, 2\rangle$ and thus it is not an eigenvector.

$$[T]_\beta$$

is not a basis of eigenvectors, something we found before even finding the matrix itself.

Plugging in our second vector, $T(2, 3) = (2, 4)$. We can see that each vector results in a scaled version of the opposite basis vector. This means that

$$[T]_\beta$$

will be an anti-diagonal matrix with the scaling of the opposite vector in each component. This produces

$$[T]_\beta = \begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}$$

f. Let's evaluate each vector in β from our formula

$$T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -7a - 4b + 4c - 4d & b \\ -8a - 4b + 5c - 4d & d \end{pmatrix}$$

So we have

$$T\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -3 & 0 \\ -3 & 0 \end{pmatrix} = \boxed{-3}\cdot\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

$$T\begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix} = \boxed{1}\cdot\begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix}$$

$$T\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} = \boxed{1}\cdot\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$$

$$T\begin{pmatrix}-1 & 0\\ 0 & 2\end{pmatrix} = \begin{pmatrix}-1 & 0\\ 0 & 2\end{pmatrix} = \boxed{1}\cdot\begin{pmatrix}-1 & 0\\ 0 & 2\end{pmatrix}$$

We can see that every basis vector is mapped to some scaled version of itself, by a factor that is boxed above. This means that $[T]_\beta$ will be an eigenbasis and will be easy to create simply by tacking the eigenvalues on the diagonal–something that we have shown in class. Thus, we have,

$$[T]_\beta = \begin{pmatrix}-3 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

---

6 Textbook 5.2 Question 3(a), (b), (f)

Let's consider $T : P_3(\mathbb{R}) \to P_3(\mathbb{R})$ defined by $T(f(x)) = f'(x) + f''(x)$. Let's plug in our favorite basis vectors for a polynomial to calculate $[T]_S$ where our basis for $P_3(\mathbb{R})$ will as usual be $S = \{1, x, x^2, x^3\}$.

$$T(1) = 0 \implies [T(1)]_S = \vec{0}$$

$$T(x) = 1 \implies [T(x)]_S = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$$

$$T(x^2) = 2x + 2 \implies [T(x^2)]_S = \begin{pmatrix}2\\2\\0\\0\end{pmatrix}$$

$$T(x^3) = 3x^2 + 6x \implies [T(x^3)]_S = \begin{pmatrix}0\\6\\3\\0\end{pmatrix}$$

Thus, we have that

$$[T]_S = \begin{pmatrix}0 & 1 & 2 & 0\\ 0 & 0 & 2 & 6\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{pmatrix}$$

This, because it is upper triangular and has no elements on the diagonal, by our formula from earlier, $f_T(t) = t^4$. This means that the only eigenvalue is 0 and has an associated algebraic multiplicity of 4. The only way $[T]_S$ would have a geometric multiplicity of four would be if it had a four dimensional kernel. But, that would mean that every vector is in the kernel which only happens with the zero matrix. $[T]_S$ is obviously not the matrix of all zeros and is thus not diagonalizable. □

b. Let's use $S = \{1, x, x^2\}$ and calculate the transformation on the basis vectors. We have

$$T(1) = x^2 \qquad T(x) = x \qquad T(x^2) = 1$$

We can again transform this into the standard coordinate representation in $\mathbb{R}^3$ in the obvious way. And plugging in, we see that we have

$$[T]_S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Calculating $f_T(t)$ with our standard formula, expanding on the top row produces

$$f_T(t) = t^2(t-1) - 1(t-1) = (t^2-1)(t-1) = (t-1)^2(t+1) \implies t = -1, 1$$

Let's consider $t = 1$. It has an algebraic multiplicity of two and thus should have a geometric multiplicity of two to match if it is to be diagonalizable. Consider

$$\ker \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

We can add the first and third row, cancelling one, giving us a rank of one, meaning a two dimensional kernel. It can spanned by the vectors $\{(1,0,1),(0,1,0)\}$ which should be obvious from the matrix reduction described. This means we are good for $t = 1$, the algebraic multiplicity and geometric multiplicity match.

For $t = -1$, we have

$$\ker \begin{pmatrix} -1 & 0 & -1 \\ 0 & -2 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

Cancelling out one of the first and third row, we are left with a matrix with rank two, meaning a one-dimensional kernel. This matches the geometric multiplicity and can thus be spanned by $\{(1,0,-1)\}$.

Therefore, we can have a basis $\gamma = \{1+x^2, x, 1-x^2\}$. Therefore,

$$[T]_\gamma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

f. In this subpart, we will work with the transpose transformation from $M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$. We have come up with a $[T]_S$ in a previous homework–where $S = \{E_{11}, E_{12}, E_{21}, E_{22}\}$–and found this to be

$$[T]_S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This will have a characteristic polynomial equal to $(t+1)(t-1)^3$ by simple inspection. This produces $\lambda_1, \lambda_2 = -1, 1$, with algebraic multiplies of one and three respectively. By definition, the geometric multiplicity will match for λ_1. Plugging in λ_2 we have

$$\ker \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Cancelling out one of the middle rows, this does indeed have a rank of one, so geometric multiplicity of three. We can take the vectors that span this–being E_{λ_2}–combined with those that span the E_{λ_1} eigenspace to form a basis. We can quickly find that one vector by considering

$$\ker \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Now, $\gamma = \{E_{11}, E_{12} + E_{21}, E_{22}, E_{12} - E_{21}\}$, where that final vector corresponds to E_{λ_1}. This produces.

$$[T]_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

---

7 Textbook 5.2 Question 8: Proof on Diagonalizability Conditions

> Suppose that $A \in M_{n\times n}(F)$ has distinct eigenvalues λ_1, λ_2 and that $\dim(E_{\lambda_1}) = n - 1$. Prove that A is diagonalizable.

From what we proved in discussion, $\dim(E_{\lambda_2}) \geq 1$ and $\dim(E_{\lambda_2}) \leq n_2 = 1$. This implies that $\dim(E_{\lambda_2}) = 1$. Since the sum of the geometric multiplicities is thus $n - 1 + 1 = n$, we have that the sum of the algebraic multiplicity is also n because that this is the maximum degree of the characteristic polynomial. Because the sum of the geometric multiplicities match match with the algebraic multiplicites–and correspond with each other–we have that the characteristic polynomial must split because the degree of the irreducible $g(t)$ in the formula for the characteristic polynomial must be zero. Thus, by proving $f_T(t)$ splits and proving that the algebraic and geometric multiplicities match, we have shown diagonalizability. □

---

8 Eigenspace with Quotient Spaces

> Let V be a vector space over F and $T\colon V \to V$ linear with eigenvalues $\lambda_1, \ldots, \lambda_n$, and such that $\lambda_i \neq \lambda_j$ for $i \neq j$. (You should assume that T has no other eigenvalues.)
> **a.** Let $E = E_{\lambda_1}$ denote the λ_1 eigenspace. Let V/E denote the quotient vector space. Prove that the function $S\colon V/E \to V/E$ defined for $[v] \in V/E$ by $S([v]) = [T(v)]$ is well-defined and linear.

Well-defined

Suppose $[v] = [v'] \in V/E$. We want to show that for these arbitrary two representations of an element in V/E produce the same output when S is applied. Thus, we want to show that $S([v]) = S([v'])$. Note, because $[v] = [v']$, we can say that $v - v' = w \in E$ by the definition of equivalence classes and similarity. Let's consider

$$\begin{aligned}
S([v]) &= [T(v)] && \text{def of } S \\
&= [T(v' + w)] && \text{substitution} \\
&= [T(v') + T(w)] && \text{linearity} \\
&= [T(v')] + [T(w)] && \text{def of modulo addition} \\
&= [T(v')] + [\lambda_1 w] && \text{applying } T \\
&= [T(v')] + \lambda_1 [w] && \text{def of scalar multiplication} \\
&= [T(v')] && w \text{ is in } E \\
&= S([v']) && \text{def of } S
\end{aligned}$$

Thus, we have shown two arbitrary representations of the same element produce the same result, showing that S is well defined.

Linearity
Let's suppose $[v], [u] \in V/E$ and that $\alpha \in F$. Consider

$$\begin{aligned}
S([v] + \lambda[u]) &= S([v + \lambda u]) && \text{def. of modulo addition and scalar multiplication} \\
&= [T(v + \lambda u)] && \text{def. of } S \\
&= [T(v) + \lambda T(u)] && \text{linearity} \\
&= [T(v)] + \lambda[T(u)] && \text{def. of modulo addition and scalar multiplication} \\
&= S([v]) + \lambda S([u]) && \text{def. of } S
\end{aligned}$$

Thus, we have shown linearity. □

b. Prove that $\lambda_2, \ldots, \lambda_n$ are eigenvalues of S.

Suppose that each of $\lambda_2, \ldots, \lambda_n$ has associated eigenvectors $v_2, \ldots, v_n$ where these are eigenvectors for T. We can say that their respective equivalence classes may be eigenvectors for S because eigenvectors are linearly independent and thus won't collapse to zero in the quotient space. Consider for arbitrary $2 \leq i \leq n$.

$$\begin{aligned}
S([v_i]) &= [T(v_i)] && \text{def. of } S \\
&= [\lambda_i v_i] && \text{def. of eigenvector} \\
&= \lambda_i [v_i] && \text{def. of multiplication}
\end{aligned}$$

Thus, for each eigenvector $v_i \in V$ for T, $[v_i] \in V/E$ is an eigenvector for S with the same associated eigenvalue λ_i by the definition of an eigenvector/value for $2 \leq i \leq n$. □

---

9 Linear Independence of Eigenvectors

Let A be an $n \times n$ square matrix with n distinct eigenvalues λ_i, $1 \leq i \leq n$. Suppose we have nonzero eigenvectors v_i, $1 \leq i \leq n$, such that $Av_i = \lambda_i v_i$. Show that the set $\beta = \{v_1, v_2, \ldots, v_n\}$ is linearly independent.

Let's prove this by induction on n.
Base Case: (n = 1)
Consider a matrix $A = [A_{11}]$ and a set $\beta = v_1$. Because v_1 is a non-zero eigenvector, it is by definition linearly independent.
Inductive Case:
Assume that A be an $n-1 \times n-1$ square matrix with n distinct eigenvalues λ_i, $1 \leq i \leq n-1$. Suppose we have nonzero eigenvectors v_i, $1 \leq i \leq n-1$, such that $Av_i = \lambda_i v_i$. We have that the set $\beta = \{v_1, v_2, \ldots, v_{n-1}\}$ is linearly independent.
Now, let A be an $n \times n$ square matrix with n distinct eigenvalues λ_i, $1 \leq i \leq n$. Suppose we have nonzero eigenvectors v_i, $1 \leq i \leq n$, such that $Av_i = \lambda_i v_i$. We want to show that the set $\beta = \{v_1, v_2, \ldots, v_n\}$ is linearly independent. Consider the following where a_i are scalars.

$$
\begin{aligned}
\sum_{i=1}^{n} a_i v_i &= \vec{0} && \text{initial assumption} \\
(\dagger) \sum_{i=1}^{n-1} a_i v_i + a_n v_n &= \vec{0} && \text{rewrite summation} \\
T\left(\sum_{i=1}^{n-1} a_i v_i + a_n v_n\right) &= \vec{0} && \text{applying } T \\
\sum_{i=1}^{n-1} a_i T(v_i) + a_n T(v_n) &= \vec{0} && \text{linearity} \\
(*) \sum_{i=1}^{n-1} a_i \lambda_i v_i + a_n \lambda_n v_n &= \vec{0} && \text{def. of eigenvector}
\end{aligned}
$$

Now, let's also consider λ_n scaling †, producing

$$\sum_{i=1}^{n-1} \lambda_n a_i v_i + \lambda_n a_n v_n = \vec{0}$$

Finally, let's subtract this from $*$ to produce:

$$\sum_{i=1}^{n-1} a_i \lambda_i v_i + a_n \lambda_n v_n - \left(\sum_{i=1}^{n-1} \lambda_n a_i v_i + \lambda_n a_n v_n\right)$$

This simplifies to

$$\sum_{i=1}^{n-1} (\lambda_i - \lambda_n)(a_i v_i) = \vec{0}$$

We already know that $\lambda_i - \lambda_n \neq 0$ because each eigenvector is distinct. Thus, we can simply apply the inductive step, apply the linear independence of the first $n-1$ vectors, showing that each of $a_i = 0$ for all $1 \leq i \leq n$. Because all of these must be zero, we are left with

$$a_n v_n = \vec{0}$$

This directly implies $a_n = 0$ as well. Thus, we have shown linear independence, that $\sum_{i=1}^{n} a_i v_i = \vec{0} \implies a_i = 0 \; \forall i$.

□

---

10 Degree of Characteristic Polynomial

Recall that, for $A \in M_{n\times n}(F)$,

$$f_A(t) := \det(tI_n - A).$$

We may view f_A as an element in

$$\mathrm{Poly}(F) = \{\sum_{i=1}^{k} a_i t^i \mid k \in \mathbb{Z}, k \geq 0, a_i \in F\}.$$

Recall also that, for $g \in \mathrm{Poly}(F)$, if

$$g(t) = \sum_{i=0}^{k} a_i t^i$$

and $a_k \neq 0$, then

$$\deg(g) := k.$$

a. Let $A \in \mathrm{M}_{n\times n}$. Prove, by induction on n, that $\deg(f_A) = n$.

Base Case: (n = 1)

Consider an arbitrary single-entry matrix $A = [A_{11}]$. Then, by definition, $f_A(t) = t - A_1 1$. This by definition has degree $n = 1$. Thus, we have proven our base case.

Inductive Case

Let's assume $\deg(f_{A_{n-1}}(t)) = n-1$. Thus, we can write $f_{A_{n-1}}(t) = a_0 + a_1 t + \ldots + a_{n-1} t^{n-1}$. Now, let's consider how we would find $f_{A_n}(t)$. We have

$$\begin{aligned}\det(tI_n - A) &= \sum_{j=1}^{n} (-1)^{n+j} (tI_n - A)_{nj} \det(\overline{(tI_n - A)_{nj}}) \\ &= \sum_{j=1}^{n-1} (-1)^{n+j} (tI_n - A)_{nj} \det(\overline{(tI_n - A)_{nj}}) + (tI_n - A)_{nn} \det(\overline{(tI_n - A)_{nn}})\end{aligned}$$

Because there are the most terms with t in $\overline{(tI_n - A)_{nn}}$ of all the j values in the summation and $(tI_n - A)_{nn}$ contains a t term, we can ignore all other aspects of this expansion because they will not contribute to the degree. Let's simply consider

$$(tI_n - A)_{nn} \cdot \det(\overline{(tI_n - A)_{nn}})$$

Here, $(t - A_{nn})$ has degree 1 by definition, and thus, by the inductive hypothesis, it will be multiplied by something with degree $n-1$. We know from class that the degrees will add in this case, leaving us with degree n, hence proving our claim. □

b. Using the previous item and definitions, prove that if V is a vector space over F, $\dim(V) = n$ and $T\colon V \to V$ is linear, then $\deg(f_T) = n$.

Since, we have proven that $\deg(f_A) = n$, we can simply set $A = [T]_\beta$ where β is an arbitrary ordered basis. By definition, since $\dim(V) = n$, this basis will have n vectors, and thus, $[T]_\beta \in M_{n\times n}(F)$. Furthermore, we know that the characteristic polynomial is closed to change of basis. Thus any basis should produce the same result and this concept of assigning A to a singular representation is well defined. Therefore, directly following from part A and these ideas, we can say that $\deg(f_T) = n$. □

11 Sum of Subspaces

Let V a vector space over F, let $r \geq 1$ be an integer, and let $W_1, \ldots, W_r \subset V$ be subspaces. The sum of the subspaces is defined to be

$$W_1 + \ldots + W_r = \{w_1 + \ldots + w_r \mid w_i \in W_i\}.$$

Definition (r-fold internal direct sum). We say that the sum $W_1 + \ldots + W_r$ is an (internal) direct sum if the representation of elements in $W_1 + \ldots W_r$ is unique: if $u_1 + \ldots + u_r = u_1' + \ldots + u_r'$ for $u_i, u_i' \in W_i$, then $u_i = u_i'$ for all $1 \leq i \leq r$. If the sum is direct, we write $W_1 \oplus \cdots \oplus W_r$ instead of $W_1 + \cdots + W_r$.

a. Recall that we previously defined $W_1 + W_2$ to be a direct sum if $W_1 \cap W_2 = \{\vec{0}\}$. Prove that $W_1 \cap W_2 = \{\vec{0}\}$ if and only if the definition above with $r = 2$ is satisfied.

This is a biconditional statement. Thus, let's consider each implication separately.

Intersection zero definition implies general definition

Let's assume that $W_1 \cap W_2 = \{\vec{0}\}$. Let's suppose that we have elements $u_1, u_1' \in W_1$ and $u_2, u_2' \in W_2$ such that $u_1 + u_2 = u_1' + u_2'$. Consider

$$\begin{aligned} u_1 + u_2 &= u_1' + u_2' && \text{assumption} \\ u_1 - u_1' &= u_2' - u_2 && \text{reorder and subtract with field axioms} \end{aligned}$$

Observe that $u_1 - u_1' \in W_1$ and that $u_2' - u_2 \in W_2$. Their equality implies that each indvidually equals the zero vectors, i.e. $u_1 - u_1' = u_2' - u_2 = \vec{0} \implies u_1 = u_1'$ and $u_2 = u_2'$. Thus, we have shown one implication.

General definition implies intersection zero definition

Let's asssume that if $u_1 + u_2 = u_1' + u_2'$ for $u_i, u_i' \in W_i$, we have that $u_i = u_i'$. So, let's suppose that $x \in W_1 \cap W_2$. This implies that $x \in W_1$ and $x \in W_2$ by definition. We also know by additive inverses of vectors and closure of the sum of subsapces, $x - x = \vec{0} \in W_1 + W_2$. However, we also know that $\vec{0} - \vec{0} = \vec{0}$. Thus, by the unique representation of a vector in the direct sum of two subspaces, x must be the zero vector. Therefore the intersection must be zero. □

b. Let $V = \mathbb{R}^2$, $W_1 = \text{span}\{(1,0)\}$, $W_2 = \text{span}\{(0,1)\}$, and $W_3 = \{(1,1)\}$. Explain why $W_1 + W_2$, $W_2 + W_3$, and $W_1 + W_3$ are direct sums, but $W_1 + W_2 + W_3$ is not.

We have that

$$W_1 = \{(a,0) \mid a \in \mathbb{R}\} \qquad W_2 = \{(0,b) \mid b \in \mathbb{R}\} \qquad W_3 = \{(c,c) \mid c \in \mathbb{R}\}$$

If we are to set these equal to each other on their own–excuse the abusive notation–we would have something of the forms

$$(a,0) = (0,b) \implies a = b = 0$$

This shows $W_1 \cap W_2 = \vec{0}$

$$(a,0) = (c,c) \implies c = 0 \implies a = 0$$

This shows $W_1 \cap W_3 = \vec{0}$

$$(0,b) = (c,c) \implies c = 0 \implies b = 0$$

This shows $W_2 \cap W_3 = \vec{0}$

Thus, sums of any two are direct sums. However, to show the sum of all three is not a direct sum, we just need two distinct ways to represent it, i.e a counterexample. Let's try

$$(1,0) + (0,1) + (2,2) = (2,0) + (0,2) + (1,1)$$

Above are two different ways to create $(3,3)$ which implies the sum $W_1 + W_2 + W_3$ is not direct.

c. Let $W_1, \ldots, W_r$ be arbitrary subspaces of V (not necessarily with a sum that is direct). Consider the vector space $W_1 \times \ldots \times W_r = \{(w_1, \ldots, w_r) \,|\, w_i \in W_i\}$, with addition and scaling defined component-wise. Prove that the transformation

$$T\colon W_1 \times \ldots \times W_r \to W_1 + \ldots + W_r$$

given by

$$T(v_1, \ldots, v_r) = v_1 + \ldots + v_r$$

is linear.

Let's use our typical approach to linearity. Suppose $\lambda \in F$ and that $v = (v_1, \ldots, v_r), u = (u_1, \ldots, u_r) \in W_1 \times \ldots \times W_r$. Consider

$$\begin{aligned}
T(v + \lambda u) &= T((v_1, \ldots, v_r) + \lambda(u_1, \ldots, u_r)) && \text{def of vectors} \\
&= T((v_1 + \lambda u_1, \ldots, v_r + \lambda u_r)) && \text{component-wise addition and multiplication} \\
&= (v_1 + \lambda u_1) + \ldots + (v_r + \lambda u_r) && \text{def. of } T \\
&= v_1 + \ldots + v_r + \lambda(u_1 + \ldots + u_r) && \text{component-wise addition and multiplication} \\
&= T(v) + \lambda T(u) && \text{def. of } T
\end{aligned}$$

Hence, we have shown linearity. □

d. Prove that T is an isomorphism if and only if the sum $W_1 + \ldots + W_r$ is an internal direct sum.

Isomorphism $\implies$ *Direct Sum*

This direction is somewhat trivial. By the definition of an isomorphism, we have that T is both injective and surjective. Therefore, let's suppose that $u_1 + \ldots + u_r = u_1' + \ldots + u_r'$ for the two vectors in $W_1 + \ldots + W_r$. We can see that

$$\iff T(u_1, \ldots, u_r) = T(u_1', \ldots, u_r')$$

By injectivity,

$$\implies (u_1, \ldots, u_r) = (u_1', \ldots, u_r')$$

which implies that $u_i = u_i'$ for all $1 \leq i \leq r$. Thus, we have shown that T being an isomorphism implies that $W_1 + \ldots + W_r$ is an internal direct sum.

Direct Sum $\implies$ *Isomorphism*

In this proof, we will show injectivity and surjectivity, meaning T is an isomorphism. Let's suppose $x = (x_1, \ldots, x_r), y = (y_1, \ldots, y_r) \in W_1 + \cdots + W_r$ such that $T(x) = T(y)$. Consider

$$T(x) = T(y)$$

$$x_1 + \ldots + x_r = y_1 + \ldots + y_r$$

which by the definition of direct sum implies that $x_i = y_i$ for all $1 \leq i \leq r$, showing that $x = (x_1, \ldots, x_r) = (y_1, \ldots, y_r) = y$ by the definition of equality.

Now, for surjectivity, let's suppose that $z \in W_1 + \cdots + W_r$. Then, by the definition of a summation of subspaces, $z = w_1 + \ldots + w_r$ for some $w_i \in W_i$ for all i. Then, we claim that $T(w_1, \ldots, w_r) = w_1 + \ldots + w_r = z$, we can see this is true by the definition of T. Thus, we have proven surjectivity because for all $w_1 + \ldots + w_r$, we have that $T(w_1, \ldots, w_r) = w_1 + \ldots + w_r$. Therefore, T is an isomorphism. □

e. Suppose that V is finite-dimensional, so that $W_1, \ldots, W_r$ are also finite-dimensional vector spaces. Prove that $\dim(W_1 \times \ldots \times W_r) = \dim(W_1) + \ldots + \dim(W_r)$. Using the previous item, conclude that if $W_1 + \cdots + W_r$ is a direct sum, then $\dim(W_1 + \ldots + W_r) = \dim(W_1) + \ldots \dim(W_r)$.

Let's consider a basis for $W_1 \times \ldots \times W_r$. A basis would consist of the individual bases for each subspace, being say $\beta_i = \{v_{i1}, v_{i2}, \ldots, v_{ik}\}$ where k is the specific number of basis vectors a β_i had, possibly different for each i. Now, a basis β for the entire iterative Cartesian product would consist of the union of all the individual basis vectors each in its own component, i.e. in component i. By the component based definition of the Cartesian Product, this directly follows as there is no overlap between how each β_i is used in β. We would thus define

$$\beta = \cup_{i=1}^{r} \left(\cup_{j=1}^{k} (\cdots \vec{0}_{i-1}, v_{ij}, \vec{0}_{i+1}, \cdots) \right)$$

This by definition has the union of all the bases of β_i. This implies that $\dim(W_1 \times \ldots \times W_r) = \dim(W_1) + \ldots + \dim(W_r)$.

Now, because T is an isomorphism as we proved, we can conclude that $\dim(W_1 + \ldots + W_r) = \dim(W_1 \times \ldots \times W_r)$. Since we showed that $\dim(W_1 \times \ldots \times W_r) = \dim(W_1) + \ldots + \dim(W_r)$, we can thus by substitiution say that $\dim(W_1 + \ldots + W_r) = \dim(W_1) + \ldots \dim(W_r)$. □