Math 110 Linear Algebra
Midterm 2 Review
October 28, 2017

Material

Material covered on the midterm includes:

- All lectures from Thursday, Sept. 21st to Tuesday, Oct. 24th
- Homeworks 9 to 17
- Quizzes 5 to 9

Sections in the book related to this material are

- Section 2.5, Change of Coordinate Matrices
- Section 3.1, Elementary Matrices
- Sections 3.3 and 3.4, Systems of Linear Equations
- Sections 4.1 to 4.4, Determinants
- Section 5.1, Eigenvalues and Eigenvectors
- Section 5.2, Diagonalization
- Section 5.4, Invariant Subspaces and the Cayley Hamilton Theorem
- Section 7.1, Jordan Canonical Form, esp. Jordan Blocks and Generalized Eigenvectors

In particular it is a good idea to go back through the homework assignments and read through them with an eye towards understanding what is meant by the hints, and especially by the sections titled "Think". Now is the time to assemble your understanding of the material into a more cohesive whole, so think about general ideas as well as computations.

This document will review some small subset of the material from the past 5 weeks, with an emphasis on applying concepts to some particular problems. The material presented here is in no way comprehensive, so make sure to review what you need to from other course materials.

Best of luck studying!

Change of Coordinates and Diagonalization

Recall that if $T : V \to W$ is a linear transformation and $\beta = \{\beta_1, \dots, \beta_n\}$ and $\gamma = \{\gamma_1, \dots, \gamma_m\}$ are ordered bases of V and W respectively, then $[T]_\beta^\gamma$ is the **matrix of T** with respect to the bases β and γ, whose i-th column is the vector $[T(\beta_i)]_\gamma$, the coefficients of $T(\beta_i)$ when represented as a linear combination of the vectors in γ.

Now suppose that $V = W$, so that all operations are happening within the same vector space. Then $Q = [\mathrm{Id}]_\beta^\gamma$ is the **change of coordinates matrix** from basis β to basis γ. It has the useful properties that

- $Q[v]_\beta = [v]_\gamma$ for any vector $v \in V$
- $[T]_\beta = Q^{-1}[T]_\gamma Q$ for any linear operator $T : V \to V$

A linear operator $T : V \to V$ is called **diagonalizable** if there is an ordered basis β such that $[T]_\beta$ is diagonal. Note that this is the case iff there exists a basis of V consisting of eigenvectors of T, in which case $[T]_\beta$ is diagonal with diagonal entries equal to the eigenvalues of T.

Another important characterization of diagonalizable operators is the following. $T : V \to V$ is diagonalizable iff

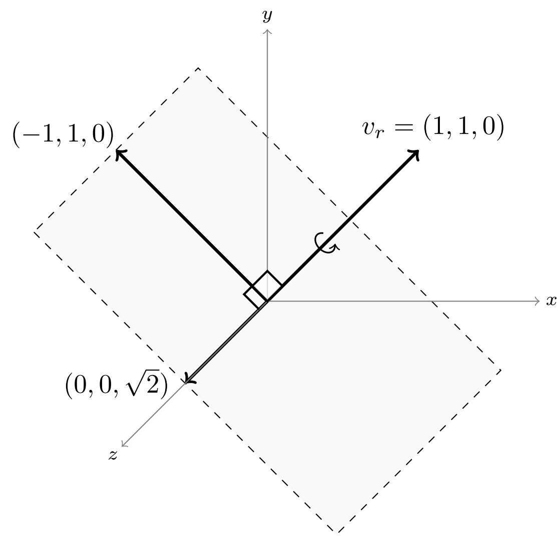
- The characteristic polynomial of T, given by $\mathrm{char}_T(t) = \det(T - t\,\mathrm{Id})$, splits into a product of linear factors
- Each eigenspace E_λ corresponding to an eigenvalue λ has dimension (or geometric degree) equal to the algebraic degree of λ in $\mathrm{char}_T(t)$.

In this case, an eigenbasis for T can be formed by taking a union of bases for each of the eigenspaces E_λ.

Problem 1.

Let T be the transformation in $\mathbb{R}^3$ which is a rotation about the axis spanned by $v_r = (1,1,0)$ by $90°$ counterclockwise. Find the matrix $[T]_E$ for T in terms of the standard basis E. Is T diagonalizable? Is T diagonalizable if considered as a linear operator on $\mathbb{C}^3$ over the field $\mathbb{C}$?

Solution. Begin by representing T in terms of a basis which is related to the geometry of the transformation.



In particular, as pictured above, the vector v_r itself is fixed by the transformation, and the vectors $v_1 = (-1,1,0)$ and $v_2 = (0,0,\sqrt{2})$ are contained in the plane perpendicular to v_r which is fixed by the rotation. Further, these two vectors are perpendicular inside of the fixed plane, so that v_1 maps to v_2 under the rotation, and v_2 maps to v_1. This gives a very simple form for the transformation in terms of the basis $\beta = \{v_r, v_1, v_2\}$:

$$\beta = \{(1,1,0), (-1,1,0), (0,0,\sqrt{2})\}, \qquad [T]_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

To find $[T]_E$, we will need to apply the change of coordinates formula.

$$[T]_E = [\mathrm{Id}]_\beta^E \, [T]_\beta \, [\mathrm{Id}]_E^\beta$$

In particular, the change of coordinates matrix $[\mathrm{Id}]_\beta^E$ can be formed by using the column vectors for the vectors in β as the columns of a matrix, so

$$[\mathrm{Id}]_\beta^E = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

Further, we have $[\mathrm{Id}]_E^\beta = \left([\mathrm{Id}]_\beta^E\right)^{-1}$, so a short computation using row reduction gives

$$[\mathrm{Id}]_E^\beta = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

Finally, we are able to compute

$$\begin{aligned}
[T]_E &= [\mathrm{Id}]_\beta^E \, [T]_\beta \, [\mathrm{Id}]_E^\beta \\
&= \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & -\sqrt{2} \\ -1 & 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1/2 & 1/2 & \sqrt{2}/2 \\ 1/2 & 1/2 & -\sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \end{pmatrix}
\end{aligned}$$

At this point, it is good to check that the matrix $A = [T]_E$ which we computed actually does what we want: $Av_r = v_r$, $Av_1 = v_2$, and $Av_2 = -v_1$. In fact this, is the case.

For diagonalizability, we start by computing the characteristic polynomial of T. Notice that we can compute the characteristic polynomial with the determinantal formula using the matrix of T with respect to any basis we want. With that in mind, the $[T]_\beta$ is a much simpler matrix, so it will probably make our computation a little easier. We have

$$\mathrm{char}_T(\lambda) = \det([T]_\beta - \lambda I_3) = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = -(\lambda-1)(\lambda^2+1)$$

The polynomial λ^2+1 has no real roots, so it doesn't factor into a product of linear terms with real coefficients. In particular, this means that $\mathrm{char}_T(\lambda)$ doesn't split over $\mathbb{R}$, and thus the operator is not diagonalizable.

However, thinking of T as a linear operator over $\mathbb{C}$, the story is different. We have

$$\mathrm{char}_T(\lambda) = -(\lambda-1)(\lambda^2+1) = -(\lambda-1)(\lambda+i)(\lambda-i)$$

Thus the characteristic polynomial *does* split over $\mathbb{C}$, and in fact it has three distinct eigenvalues. This already implies that T is diagonalizable over $\mathbb{C}$ since the geometric multiplicity of an eigenvalue is always at least 1.

We can compute the eigenvectors associated to these eigenvalues by finding a basis for the null spaces of the matrices $B_\lambda = [T]_\beta - \lambda I_3$. We have

$$B_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \qquad B_i = \begin{pmatrix} 1-i & 0 & 0 \\ 0 & -i & -1 \\ 0 & 1 & -i \end{pmatrix}, \qquad B_{-i} = \begin{pmatrix} 1+i & 0 & 0 \\ 0 & i & -1 \\ 0 & 1 & i \end{pmatrix}$$

After row reducing, we find reduced row echelon forms

$$B_1 \sim \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B_i \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -i \\ 0 & 0 & 0 \end{pmatrix}, \qquad B_{-i} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & i \\ 0 & 0 & 0 \end{pmatrix}$$

This gives a single eigenvector u_λ spanning each eigenspace,

$$[u_1]_\beta = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad [u_i]_\beta = \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}, \qquad [u_{-i}]_\beta = \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}$$

Thus $\gamma = \{u_1, u_i, u_{-i}\}$ is an eigenbasis of T, and

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

is a diagonal matrix. Notice that because we computed the null space of the matrix $[T-\lambda I]_\beta$, the eigenvectors we obtained are represented using coordinates in terms of β. To diagonalize $[T]_E$, we will need to do one last computation to find the right change of coordinates matrix. Since we've expressed γ as vectors in terms of β, we have

$$[\mathrm{Id}]_\gamma^\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 1 \\ 0 & 1 & i \end{pmatrix}$$

Since we already have a change of coordinates matris from β to E, we compute

$$Q = [\mathrm{Id}]_\gamma^E = [\mathrm{Id}]_\beta^E \, [\mathrm{Id}]_\gamma^\beta = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 1 \\ 0 & 1 & i \end{pmatrix} = \begin{pmatrix} 1 & -i & -1 \\ 1 & i & 1 \\ 0 & \sqrt{2} & \sqrt{2}i \end{pmatrix}$$

Then with these choices of D and Q, we finally can diagonalize $[T]_E$ by

$$Q^{-1}[T]_E Q = D$$

Determinants

Recall that for a square $n \times n$ matrix A, we can consider the **determinant** of A in a number of different ways which can be useful theoretically or computationally in different contexts. In particular, we've looked at rook placements, cofactor expansion, and row reduction as methods for working with determinants. In particular, we know the following properties of determinants. If A and B are $n \times n$ matrices, then

- Determinants are multiplicative: $\det(AB) = \det(A)\det(B)$
- Determinants are preserved by similarity: If $A \sim B$, then $\det(A) = \det(B)$
- Determinants characterize invertibility: A is invertible iff $\det(A) \neq 0$, and if A is invertible, then $\det(A^{-1}) = 1/\det(A)$
- Determinants are preserved by transposes: $\det(A^t) = \det(A)$

In particular, because determinants are preserved by similarity, for a linear transformation $T : V \to V$ where V is a finite-dimensional vector space, we can define $\det(T) = \det([T]_\beta)$, where β is *any* basis of V. This makes sense (is "well-defined") since for different bases β and γ, the matrices $[T]_\beta$ and $[T]_\gamma$ are similar by the change of coordinates formula, so the determinant doesn't depend on which basis you choose.

Also recall how determinants interact with elementary row operations:

- Adding a multiple of one row of a matrix to another doesn't change the determinant.
- Multiplying a row of a matrix by a scalar α multiplies the determinant by α.
- Swapping two rows of a matrix changes the sign of the determinant.

The same properties also hold for the corresponding elementary column operations.

Finally, recall a favorite important characteristic of determinants: The determinant of a matrix can be thought of as a *polynomial*, with the coordinates of the matrix as the variables. This comes from the rook placement definition of determinants. For instance:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

This is a polynomial with variables a, b, c, and d. In general, for an $n \times n$ determinant, the polynomial will have $n!$ terms, each of which is a product of n different variables from the entries of the matrix.

Problem 2.

Compute the determinant $\det(A)$, where

$$A = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & -1 & -1 & 4 \\ -4 & 5 & -10 & -6 \\ 3 & -2 & 10 & -1 \end{pmatrix}$$

Solution. We row reduce A, using only the operation of adding a scalar multiple of one row to another, to form a new matrix B which is upper triangular. This matrix will have the same determinant of A, and the determinant of an upper triangular matrix is equal to the product of its diagonal entries. We compute

$$A \sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -5 & 2 \\ 0 & 1 & -2 & -2 \\ 0 & 1 & 4 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -5 & 2 \\ 0 & 0 & 3 & -4 \\ 0 & 0 & 9 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -5 & 2 \\ 0 & 0 & 3 & -4 \\ 0 & 0 & 0 & 6 \end{pmatrix} = B$$

Thus the determinant of A is given by the product of the diagonal entries of B, or $1 \cdot 1 \cdot 3 \cdot 6 = 18$.

Problem 3.

Let N be an $n \times n$ matrix which is **nilpotent**, i.e. for which there is some integer $k \geq 0$ such that $N^k = 0$.

1. Prove that $N - tI_n$ is invertible, with inverse

$$(N - tI_n)^{-1} = (-t^{-k})(N^{k-1} + tN^{k-1} + \cdots + t^{k-2}N + t^{k-1}I_n)$$

2. Use this fact to prove that $\mathrm{char}_N(t) = (-1)^n t^n$.
3. Prove that in fact, $N^n = 0$.

Solution. For the first part, we simply multiply:

$$\begin{aligned}
&(N - tI_n) \cdot (-t^{-k})\big(N^{k-1} + tN^{k-1} + \cdots + t^{k-2}N + t^{k-1}I_n\big) \\
&= (-t^{-k})\big(N^k + tN^{k-1} + \cdots + t^{k-1}N \\
&\qquad\qquad - tN^{k-1} - \cdots - t^{k-1}N - t^k I_n\big) \\
&= (-t^{-k}) \cdot (N^k - t^k I_n) = (-t^{-k}) \cdot (-t^k I_n) = I_n
\end{aligned}$$

For the second part, first multiply the equation $I_n = (N - tI_n) \cdot (N - tI_n)^{-1}$ on both sides by $-t^k$, and then use properties of determinants to compute as follows.

$$\begin{aligned}
&(-t^k)^n = (-1)^n t^{nk} = \det(-t^k I_n) \\
&= \det\Big((N - tI_n) \cdot \big(N^{k-1} + tN^{k-1} + \cdots + t^{k-2}N + t^{k-1}I_n\big)\Big) \\
&= \det(N - tI_n) \det\big(N^{k-1} + tN^{k-1} + \cdots + t^{k-2}N + t^{k-1}I_n\big)
\end{aligned}$$

Now note that both of the determinants at the end of this computation are polynomials in t since the entries of each matrix are themselves polynomials in t. In particular, this means that $\det(N - tI_n)$ divides $(-1)^n t^{nk}$. Since the characteristic polynomial has degree n and leading coefficient $(-1)^n$, the only possibility is that $\mathrm{char}_N(t) = \det(N - tI_n) = (-1)^n t^n$.

For the last part, apply the Cayley-Hamilton theorem, which states that a matrix's characteristic polynomial kills the matrix. In this case, this shows us that $(-1)^n N^n = 0$, so in particular, $N^n = 0$.

Invariant and Cyclic Subspaces

Recall that if V is a vector space, $W \leq V$, and $T : V \to V$ is a linear operator on V, we say that W is **T-invariant** if $T(w) \in W$ for any vector $w \in W$. An equivalent statement is that $T(W) \subseteq W$.

We have identified a number of subspaces that end up being T-invariant for a linear operator T, including

- V itself, and the trivial subspace $\{0\}$
- The image $\mathrm{Im}(T)$ of T, and also the image $\mathrm{Im}(T^k)$ of T^k for any $k \geq 1$
- Any eigenspace E_λ of T, especially $E_0 = \ker(T)$
- Any generalized eigenspace K_λ of T

Recall also that if $x \in V$, then the space $W = \mathrm{Span}(\{x, T(x), T^2(x), \ldots\})$ which is the span of all images of x under repeated applications of T is called the **T-cyclic subspace generated by x**. For the rest of this document, denote this subspace by $C_T(x)$, but note that this is not standard notation, so do not use it (in a midterm, say) without first explaining what it means.

Finally, remember that a subspace W is called **T-cyclic** if $W = C_T(x)$ for some vector $x \in W$, in which case x is called a **cyclic generator** of W. One important property of cyclic subspaces is that a T-cyclic subspace is always T-invariant, but not necessarily vice versa.

Problem 4.

Give a linear operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ and a subspace W of V such that W is T-invariant, but not T-cyclic.

Solution. Let $T = 0$ the zero operator which maps every vector to the zero vector, and let $W = \mathbb{R}^2$. W is T-invariant because for any vector $v \in W$, $T(v) = 0 \in W$. However, for any x, we have

$$C_T(x) = \mathrm{Span}(\{x, T(x), T^2(x), \ldots\}) = \mathrm{Span}(\{x, 0, 0, \ldots\}) = \mathrm{Span}(\{x\})$$

so in particular $C_T(x)$ has dimension at most 1. Since $W = \mathbb{R}^2$ has dimension 2, it is impossible that $W = C_T(x)$ for any x, so we see that W can't be expressed in that form, and thus is not T-cyclic.

Problem 5.

Let $V = P_n(\mathbb{R})$ be the vector space of polynomials of degree at most n with real coefficients, and let T be the derivative operator. Prove that V is T-cyclic. What happens if $V = P(\mathbb{R})$ is the space of *all* polynomials, with no restriction on degree?

Solution. $P_n(\mathbb{R})$ is T-cyclic because it can be expressed as $C_T(p)$ where p is the polynomial given by $p(x) = x^n$. To see this, notice that the set $\{p, T(p), T^2(p), \ldots\}$ can be written as

$$\{x^n, nx^{n-1}, n(n-1)x^{n-2}, \ldots, n!x, n!, 0, 0, \ldots\}$$

However, aside from the zero vector, this gives a basis of $P_n(\mathbb{R})$. In particular, the subspace $C_T(p)$ spans the entire space $P_n(\mathbb{R})$, so we have that $P_n(\mathbb{R}) = C_T(p)$, and the space is T-cyclic.

For $V = P(\mathbb{R})$ the story is a little different. Notice first that for a polynomial p of degree k, taking the derivative decreases the degree by 1. In particular, by applying T to p multiple times, you will never get a polynomial which has degree larger than the degree of k which you started with. This is problematic however, because that means that the set

$$\{p, T(p), T^2(p), T^3(p), \ldots\}$$

consists only of polynomials of degree at most k, and so we have

$$C_T(p) \leq P_n(\mathbb{R}) \lneqq P(\mathbb{R})$$

$C_T(p)$ forms a subspace of (in fact is equal to) the polynomials of degree at most n, and thus is a proper subspace of $P(\mathbb{R})$. This means that $P(\mathbb{R})$ can't be the cyclic subspace $C_T(p)$ for any generating polynomial p, so in this case, we see that the space is not T-cyclic.