

Math 218D Problem Session

Week 10

1. Some simple examples

- a) The matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ has characteristic polynomial $(\lambda - 1)^2$, the only eigenvalue is $\lambda_1 = 1$, the λ_1 -eigenspace is \mathbf{R}^2 with basis $\{(1, 0), (0, 1)\}$, the matrix is diagonal and diagonalizable.
- b) The matrix $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ has characteristic polynomial $(\lambda - 2)(\lambda + 2)$, the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -2$, the λ_1 -eigenspace is $\text{Span}\{(1, 0)\}$ and the λ_2 -eigenspace is $\text{Span}\{(0, 1)\}$, the matrix is diagonal and diagonalizable.
- c) The matrix $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ has characteristic polynomial λ^2 , the only eigenvalue is $\lambda_1 = 0$, the λ_1 -eigenspace is \mathbf{R}^2 , the matrix is diagonal and diagonalizable.
- d) The matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has characteristic polynomial $(\lambda - 1)(\lambda + 1)$, the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -1$, the λ_1 -eigenspace is $\text{Span}\{(1, 1)\}$ and the λ_2 -eigenspace is $\text{Span}\{(1, -1)\}$, the matrix is not diagonal but is diagonalizable.
- e) The matrix $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ has characteristic polynomial $\lambda(\lambda - 2)$, the eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 2$, the λ_1 -eigenspace is $\text{Span}\{(1, -1)\}$ and the λ_2 -eigenspace is $\text{Span}\{(1, 1)\}$, the matrix is not diagonal but is diagonalizable.
- f) The matrix $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ has characteristic polynomial $(\lambda - 2)^2$, the only eigenvalue is $\lambda_1 = 2$, the λ_2 -eigenspace is $\text{Span}\{(1, 0)\}$, the matrix is neither diagonal nor diagonalizable.
- g) The matrix $\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$ has characteristic polynomial $\lambda^2 - 4\lambda + 5$. Since $4^2 - 4 \cdot 5 < 0$, this polynomial has no real root. This means it has no real eigenvalues, and cannot be diagonalized via real matrices. It is not diagonal.

2. A 2×2 diagonalization

Consider the matrix $A = \begin{pmatrix} -1 & 2 \\ -3 & 4 \end{pmatrix}$.

a) The characteristic polynomial is $\lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2)$

b) The two eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$.

c) A $\lambda_1 = 1$ eigenvector is $(1, 1)$.

d) A $\lambda_2 = 2$ eigenvector is $(2, 3)$.

e) $D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}^{-1}$.

f) It is not hard to "guess" that $(1, 2) = -(1, 1) + (2, 3)$, i.e. $c_1 = -1$, $c_2 = 1$. If you already computed the inverse $C^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}$, you could also do $(c_1, c_2) = C^{-1}(1, 2) = (3, -1) + 2(-2, 1) = (-1, 1)$.

This means that $A^n(1, 2) = A^n(-(1, 1) + (2, 3)) = -\lambda_1^n(1, 1) + \lambda_2^n(2, 3) = -(1, 1) + 2^n(2, 3)$.

g) When n is very large, the ratio $\frac{\|A^{n+1}(1,2)\|^2}{\|A^n(1,2)\|^2} = \frac{(2 \cdot 2^{n+1} + 1)^2 + (3 \cdot 2^{n+1} + 1)^2}{(2 \cdot 2^n + 1)^2 + (3 \cdot 2^n + 1)^2}$ is approximately 4 (the +1's are negligible compared to the large 2^n terms). This means that the ratio $\frac{\|A^{n+1}(1,2)\|}{\|A^n(1,2)\|}$ is approximately 2.

h) For any n , $\|A^{n+1}(1, 1)\|/\|A^n(1, 1)\|$ is not just approximately, but exactly, equal to 1.

i) If you were given a random vector w , you would expect $\|A^{n+1}w\|/\|A^n w\|$ to be approximately 2 when n is very large - most vectors are not in the $\lambda_1 = 1$ eigenspace, and for any vector not in that eigenspace, the same logic as in g) would apply.

- 3. Some 3×3 characteristic polynomials** Both matrices have characteristic polynomial $\lambda^3 - 2\lambda^2 + \lambda$. This factors as $(\lambda - 1)^2\lambda$, so both polynomials have eigenvalues 1 and 0, with 1 being a repeated eigenvalue. The matrix A is diagonalizable:

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

The matrix B is not, since the 1-eigenspace, $\text{Nul}(B - I)$ is 1-dimensional, and the 0-eigenspace, $\text{Nul}(B)$, is also 1-dimensional. This means you can find at most 2 linearly independent eigenvectors, not the 3 you need for diagonalization.

4. Traces and determinants

Recall that the trace $\text{Tr}(A)$ is the sum of the diagonal entries of A .

- a) For example, for **a**), $\lambda_1 = 1$ and $\lambda_2 = 1$. Therefore $\text{Tr}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 + 1 = 2$, while $\det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot 1 = 1$. For a non-diagonal example, look at **d**) - the eigenvalues are $\lambda_1 = 1, \lambda_2 = -1$, $\text{Tr}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 1 + (-1) = 0$ while $\det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 1 \cdot (-1) = -1$.
- b) For any $n \times n$ matrix, the polynomial $p(\lambda) = \det(A - \lambda I_n)$ can be factored as

$$p(\lambda) = (-1)^n(\lambda - \lambda_1) \cdots (\lambda - \lambda_n).$$

When you set $\lambda = 0$ in $\det(A - \lambda I_n)$, you get $\det(A)$. When you set $\lambda = 0$ in $(-1)^n(\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$, you get $(-1)^n(-\lambda_1) \cdots (-\lambda_n) = \lambda_1 \cdot \lambda_n$. Therefore $\det(A) = \lambda_1 \cdot \lambda_n$.

- c) The determinant $\det(A)$ has another product formula:

$$\det(A) = (-1)^k d_1 \cdots d_n,$$

when the A has REF with pivot entries d_1, \dots, d_n , found using Gaussian elimination w/o row scaling and with k row swaps. Even though this formula looks quite similar to the formula of **b**), eigenvalues and pivots are not at all the same.

An example of a 2×2 matrix where the pivots d_1, d_2 are not the same as the eigenvalues λ_1, λ_2 is given by $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. This matrix has $p(\lambda) = \lambda^2 - \lambda - 1$, hence has eigenvalues $\lambda_1, \lambda_2 = \frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}$. But the REF, with one row swap, is $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, with pivots 1, 1. This gives two different formula for the determinant $\det\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = -1 \cdot 1 = \frac{1+\sqrt{5}}{2} \cdot \frac{1-\sqrt{5}}{2}$.

- d) For any $n \times n$ matrix, we will show that $\text{Tr}(A) = \lambda_1 + \cdots + \lambda_n$. We'll do the same strategy as in **b**), but the details are much trickier.

$p(\lambda)$ -side: If you expand $p(\lambda) = (-1)^n(\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$ into $p(\lambda) = (-1)^n \lambda^n + (?)\lambda^{n-1} + \cdots$, the coefficient of λ^{n-1} is $\lambda_1 + \cdots + \lambda_n$.

For example, $(-1)^3(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) = (-1)^3(\lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 + (\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1)\lambda - \lambda_1\lambda_2\lambda_3)$.

$\det(A - \lambda I)$ -side What is the coefficient of λ^{n-1} for $\det(A - \lambda I)$? Well, you have to think very carefully about the cofactor expansion, or really the formula you get when you do cofactor expansion n times, all the way to 1×1 matrices. The only term in the cofactor expansion which has a possibility of having a λ^{n-1} term is the product $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$, coming from the $(1, 1)$ -cofactor n times.

For example, when $n = 3$,

$$\det \begin{pmatrix} (a_{11} - \lambda) & a_{12} & a_{13} \\ a_{21} & (a_{22} - \lambda) & a_{23} \\ a_{31} & a_{32} & (a_{33} - \lambda) \end{pmatrix} = (a_{11} - \lambda) \det \begin{pmatrix} (a_{22} - \lambda) & a_{23} \\ a_{32} & (a_{33} - \lambda) \end{pmatrix} \\ - a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} - \lambda \end{pmatrix} \\ + a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ (a_{22} - \lambda) & a_{23} \end{pmatrix}.$$

Both $\det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & (a_{33} - \lambda) \end{pmatrix}$ and $\det \begin{pmatrix} a_{12} & a_{13} \\ (a_{22} - \lambda) & a_{23} \end{pmatrix}$ are degree one polynomials in λ , with no $\lambda^{n-1} = \lambda^2$ term. The first term

$$(a_{11} - \lambda) \det \begin{pmatrix} (a_{22} - \lambda) & a_{23} \\ a_{32} & (a_{33} - \lambda) \end{pmatrix}$$

equals $(a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda) - a_{32}a_{23}$, and only the first part of this, $(a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda)$, can have λ^2 terms.

Back to discussing general n . Since the λ^{n-1} term of $\det(A - \lambda I_n)$ is the same as the λ^{n-1} term of $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$,

$$\det(A - \lambda I_n) = (-1)^{n-1} \lambda^n + (a_{11} + \cdots + a_{nn}) \lambda^{n-1} + \cdots.$$

Conclusion: We then compare the λ^{n-1} -terms on both sides of $\det(A - \lambda I_n) = (-1)^n (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$, which gives

$$a_{11} + \cdots + a_{nn} = \lambda_1 + \cdots + \lambda_n,$$

i.e.

$$\text{Tr}(A) = \lambda_1 + \cdots + \lambda_n.$$

5. Linear independence of eigenvectors

- a) Consider a matrix A with two distinct eigenvalues $\lambda_1 \neq \lambda_2$, with associated eigenvectors v_1 and v_2 . We will show that v_1 is not a scalar multiple of v_2 .

Suppose that they were scalar multiples $v_1 = cv_2$. Note that $c \neq 0$, since the eigenvector v_1 can't be 0. Then $Av_1 = A(cv_2) = cAv_2$. Using the eigenvector equations $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2 v_2$, this becomes $\lambda_1 v_1 = c\lambda_2 v_2$. Substituting $v_1 = cv_2$, this becomes $\lambda_1(cv_2) = c\lambda_2 v_2$. As v_2 and c are not the zero vector/scalar, this implies $\lambda_1 = \lambda_2$, a contradiction.

Therefore v_1 and v_2 are not scalar multiples.

- b) Consider a matrix A with three distinct eigenvalues $\lambda_1, \lambda_2, \lambda_3$, with associated eigenvectors v_1, v_2 and v_3 . We will show that v_1, v_2 , and v_3 are linearly independent.

Suppose they were linearly dependent: we would have an equation

$$av_1 + bv_2 + cv_3 = 0,$$

where at least two of the scalars a, b, c are non-zero. If one of them is zero, we are actually in the situation of a) - we already checked that this was impossible.

Multiplying by A , we obtain another equation

$$\lambda_1 av_1 + \lambda_2 bv_2 + \lambda_3 cv_3 = 0.$$

Now, we may assume that $\lambda_1 \neq 0$ (if it is zero, re-order the eigenvalues - the eigenvalues can't all be zero, since they are 3 distinct numbers). We can subtract λ_1 times the equation $av_1 + bv_2 + cv_3 = 0$ from $\lambda_1 av_1 + \lambda_2 bv_2 + \lambda_3 cv_3 = 0$, to get

$$(\lambda_2 - \lambda_1)bv_2 + (\lambda_3 - \lambda_1)cv_3 = 0.$$

Since all the eigenvalues were distinct, the coefficients $(\lambda_2 - \lambda_1)b$ and $(\lambda_3 - \lambda_1)c$ are both nonzero. Therefore the eigenvectors v_2 and v_3 are scalar multiples of each other. But this is impossible, due to a)!

Since all cases give rise to contradictions, we may conclude that the assumption that v_1, v_2 , and v_3 are linearly dependent is impossible. In other words, any three eigenvectors with distinct eigenvalues must be linearly independent.