Week 10: Recall from last week

1. If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, an eigenvector of T is a non-zero $v \in \mathbb{R}^n$ with the property that $T(v) = \lambda v$ for some $\lambda \in \mathbb{R}$, called the eigenvalue of T to which v corresponds.
2. If a basis $\mathcal{B}$ of $\mathbb{R}^n$ consists entirely of eigenvectors of T, then the matrix of T with respect to $\mathcal{B}$ is diagonal, with the corresponding eigenvalues as the diagonal entries.
3. Eigenvectors of T that correspond to distinct eigenvalues are linearly independent in $\mathbb{R}^n$. It follows that T can have at most n distinct eigenvalues. If it has n distinct eigenvalues, then it is diagonalizable.
4. If the matrix of T with respect to the standard basis is A, then the matrix of T with respect to another basis $\mathcal{B}$ is $P^{-1}AP$, where the columns of P are the basis elements of $\mathcal{B}$.

A note about (non)-diagonalizability

For $A \in M_n(\mathbb{F})$, it does not always happen that $\mathbb{F}^n$ has a basis consisting of eigenvectors of A.

Examples

1. The matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is diagonalizable in $M_2(\mathbb{C})$ but not in $M_2(\mathbb{R})$.
2. The matrix $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable even over $\mathbb{C}$.

Example Find a matrix P with $P^{-1}AP$ diagonal, where $A = \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$

To answer this, we need to find two linearly independent eigenvectors of A. These are non-zero solutions of

$$\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} \Longrightarrow \begin{array}{rcl} 2x+2y & = & \lambda x \\ x+3y & = & \lambda y \end{array} \Longrightarrow \begin{array}{rcl} 0 & = & (\lambda-2)x-2y \\ 0 & = & -x+(\lambda-3)y \end{array} \Longrightarrow \begin{bmatrix} \lambda-2 & -2 \\ -1 & \lambda-3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So we are looking for non-zero solutions $\begin{bmatrix} x \\ y \end{bmatrix}$ of the system

$$\begin{bmatrix} \lambda-2 & -2 \\ -1 & \lambda-3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

These can occur only if the coefficient matrix is non-invertible. If it is invertible, the only solution is $x = y = 0$.
A 2×2 matrix is non-invertible if and only if its determinant is 0.

$$\det \begin{bmatrix} \lambda-2 & -2 \\ -1 & \lambda-3 \end{bmatrix} = (\lambda-2)(\lambda-3) - (-2)(-1) = \lambda^2 - 5\lambda + 4.$$

The Characteristic Polynomial of a 2×2 matrix

The characteristic polynomial of A is

$$\det(\lambda I - A) = \lambda^2 - 5\lambda + 4 = (\lambda - 4)(\lambda - 1).$$

The eigenvalues of A are the solutions of the characteristic equation $\det(\lambda I - A) = 0$, 1 and 4. The eigenspace of A corresponding to $\lambda = 1$ is the set of all solutions of the system

$$\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1 \begin{bmatrix} x \\ y \end{bmatrix} \Longrightarrow \begin{bmatrix} 1-2 & -2 \\ -1 & 1-3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This is the nullspace of the matrix $1I - A = \begin{bmatrix} -1 & -2 \\ -1 & -2 \end{bmatrix}$, which is

$$\left\{ \begin{bmatrix} -2t \\ t \end{bmatrix}, t \in \mathbb{R} \right\}.$$

An eigenvector of A for $\lambda = 1$ is any non-zero element of this space, for example $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

The characteristic polynomial of $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is

$$\begin{aligned} \det(\lambda I_2 - A) = (\lambda - a)(\lambda - d) - (-c)(-b) &= \lambda^2 - (a+d)\lambda + (ad - bc) \\ &= (\lambda - \lambda_1)(\lambda - \lambda_2). \end{aligned}$$

The sum of the eigenvalues λ_1 and λ_2 is $a + d$, the trace of A.
The product of the eigenvalues λ_1 and λ_2 is $ad - bc$, the determinant of A.

The eigenspace of A corresponding to λ_1 is the nullspace of the matrix $\lambda_1 I_2 - A$. Its non-zero elements are the eigenvectors of A corresponding to λ_1.

Section 3.5.1: The Determinant (a digression)

For any 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \underbrace{\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}}_{\text{adj}(A)} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = (\underbrace{ad - bc}_{\det(A)})I_2.$$

- If $ad - bc = 0$, then A is not invertible.
- If $ad - bc \neq 0$, then the equation shows that
  $A^{-1} = \dfrac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
  The matrix $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ is the adjugate of A.
- The matrix A has an inverse if and only if $ad - bc \neq 0$. This means that the number $ad - bc$ tells us whether or not the columns of A form a basis of $\mathbb{F}^2$ (or $\mathbb{R}^2$). The number $ad - bc$ is the determinant of A.

The 3×3 case

The version of the above equation for a 3×3 matrix $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \underbrace{\begin{bmatrix} ei - fh & -bi + ch & bf - ce \\ -di + fg & ai - cg & -af + cd \\ dh - eg & -ah + bg & ae - bd \end{bmatrix}}_{\text{adj}(A)} = \underbrace{(aei - afh - bdi + bfg + cdh - ceg)}_{\det(A)} I_3.$$

Definitions

- The minor $M_{i,j}$ of the (i,j)-entry of A is the determinant of the 2×2 matrix that remains when Row i and Column j are deleted from A.
- The cofactor $C_{i,j}$ is either $M_{i,j}$ or $-M_{i,j}$, according to $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$
- The adjugate of A is has $C_{j,i}$ in the (i,j)-position. It is the transpose of the matrix of cofactors of A.
- The determinant A can be found by multiplying each entry of any chosen row or column by its own cofactor, and adding the results.

The 3×3 case

The version of the above equation for a 3×3 matrix $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ \begin{vmatrix} d & e \\ g & h \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix} = \underbrace{(aei - afh - bdi + bfg + cdh - ceg)}_{\det(A)} I_3.$$

Definitions

- The minor $M_{i,j}$ of the (i,j)-entry of A is the determinant of the 2×2 matrix that remains when Row i and Column j are deleted from A.
- The cofactor $C_{i,j}$ is either $M_{i,j}$ or $-M_{i,j}$, according to $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$
- The adjugate of A is has $C_{j,i}$ in the (i,j)-position. It is the transpose of the matrix of cofactors of A.
- The determinant A can be found by multiplying each entry of any chosen row or column by its own cofactor, and adding the results.

Determinant Properties

1. Each of the definitions above applies to $n \times n$ matrices in general, and gives us a way to recursively define a $n \times n$ determinant, in terms of $(n-1) \times (n-1)$ determinants.
2. The cofactor expansion method is not generally the most efficient way to compute a determinant (it is ok in the 3×3 case). But it can be taken as the definition of a determinant.
3. In some special cases, the determinant is easier to compute. If A is upper or lower triangular, then $\det(A)$ is the product of the entries on the main diagonal of A. If A has a square $k \times k$ block A_1 in the upper left, a square $(n-k) \times (n-k)$ block in the lower right, and only zeros in the lower left $(n-k) \times k$ region, then $\det(A) = \det(A_1)\det(A_2)$.
4. For a pair of $n \times n$ matrices A and B, $\det(AB) = \det(A)\det(B)$. This is the multplicative property of the determinant, or the Cauchy-Binet formula. It is not obvious at all.

The characteristic polynomial of a 3×3 matrix

Example Using cofactor expansion by the first column, we find that the characteristic polynomial of $B = \begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix}$ is

$$\begin{aligned} \det(\lambda I_3 - B) &= \det \begin{bmatrix} \lambda - 5 & -6 & -2 \\ 0 & \lambda + 1 & 8 \\ -1 & 0 & \lambda + 2 \end{bmatrix} \\ &= (\lambda - 5)\left((\lambda + 1)(\lambda + 2) - 0(8)\right) + (-1)\left((-6)(8) - (\lambda + \right. \\ &= (\lambda - 5)(\lambda^2 + 3\lambda + 2) - (2\lambda - 46) \\ &= \lambda^3 - 2\lambda^2 - 15\lambda + 36 \\ &= (\lambda - 3)(\lambda^2 + \lambda - 12) \\ &= (\lambda - 3)(\lambda + 4)(\lambda - 3) \\ &= (\lambda - 3)^2(\lambda + 4) \end{aligned}$$

Algebraic and Geometric Multiplicity

The eigenvalues of B are 3 (occurring twice as a root of the characteristic polynomial), and -4 (occurring once). We say that 3 has algebraic multplicity 2 and -4 has algebraic multiplicity 1 as an eigenvalue of B.

The *geometric multplicity* of each eigenvalue is the dimension of its corresponding eigenspace.

$$3I_3 - B = \begin{bmatrix} -2 & -6 & -2 \\ 0 & 4 & 8 \\ -1 & 0 & 5 \end{bmatrix}$$

The RREF of this matrix is $\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ and the nullspace consists of all vectors $\begin{bmatrix} 5t \\ -2t \\ t \end{bmatrix}$, where $t \in \mathbb{R}$. This is the eigenspace of B corresponding to $\lambda = 3$. It has dimension 1, so 3 has geometric multiplicity 1 as an eigenvector of B.

Algebraic and Geometric Multiplicity

The eigenvalues of B are 3 (occurring twice as a root of the characteristic polynomial), and -4 (occurring once). We say that 3 has algebraic multplicity 2 and -4 has algebraic multiplicity 1 as an eigenvalue of B.
The *geometric multplicity* of each eigenvalue is the dimension of its corresponding eigenspace.

$$3I_3 - B = \begin{bmatrix} -2 & -6 & -2 \\ 0 & 4 & 8 \\ -1 & 0 & 5 \end{bmatrix}$$

The RREF of this matrix is $\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ and the nullspace consists of all vectors $\begin{bmatrix} 5t \\ -2t \\ t \end{bmatrix}$, where $t \in \mathbb{R}$. This is the eigenspace of B corresponding to $\lambda = 3$. It has dimension 1, so 3 has geometric multiplicity 1 as an eigenvector of B.

Algebraic and Geometric Multiplicity

Theorem The geometric multplicity of an eigenvalue is at most equal to its algebraic multiplicity.

Proof: Suppose that μ has geometric multiplicity k as an eigenvalue of the square matrix $A \in M_n(\mathbb{R})$, and let $\{v_1, \dots, v_k\}$ be a basis for the μ-eigenspace of A. Extend this to a basis $\mathcal{B}$ of $\mathbb{R}^n$, and let P be the matrix whose columns are the elements of $\mathcal{B}$. Then the first k columns of $P^{-1}AP$ have μ in the diagonal position and zeros elsewhere. It follows that $\lambda - \mu$ occurs at least k times as a factor of $\det(\lambda I_n - P^{-1}AP)$.

Corollary A matrix is diagonalizable if and only if the geometric multiplicity of each of its eigenvalues is equal to the algebraic multiplicity.