Chapter 5 Eigenvalues, Eigenvectors, and Diagonalization

5.1 Eigenvalues and Eigenvectors

Definitions Let T be a linear operator on R^n. A *nonzero* vector $\boldsymbol{v}$ in R^n is called an **eigenvector** of T if $T(\boldsymbol{v})$ is a multiple of $\boldsymbol{v}$; that is, $T(\boldsymbol{v})=\lambda\boldsymbol{v}$ for some scalar λ. The scalar λ is called the **eigenvalue** of T that corresponds to $\boldsymbol{v}$.

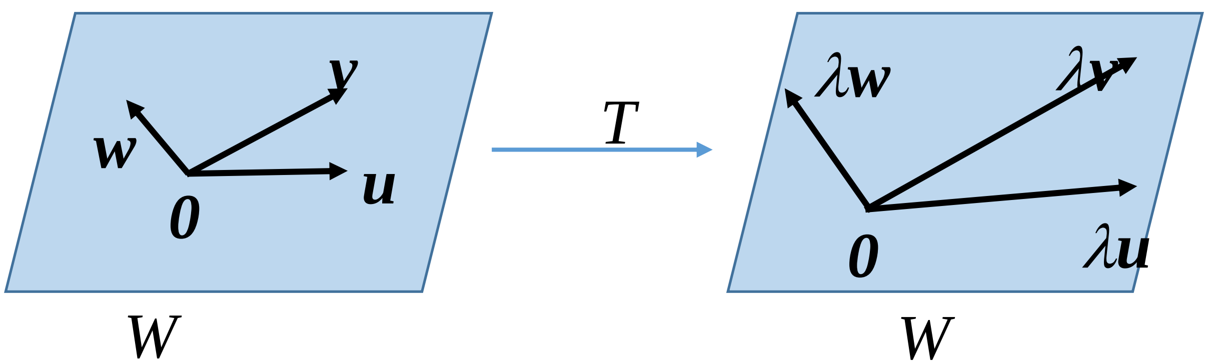
Definitions Let A be an $n\times n$ matrix. A *nonzero* vector $\boldsymbol{v}$ in R^n is called an **eigenvector** of A if $A\boldsymbol{v}=\lambda\boldsymbol{v}$ for some scalar λ. The scalar λ is called the **eigenvalue** of A that corresponds to $\boldsymbol{v}$.

The eigenvectors and corresponding eigenvalues of a linear operator are the same as those of its standard matrix.

- An eigenvector ***v*** of a matrix A is associated with exactly one eigenvalue.
- In contrast, if ***v*** is an eigenvector of A corresponding to eigenvalue λ, then every nonzero multiple of ***v*** is also an eigenvector of A corresponding to λ.

Let A be an $n \times n$ matrix with eigenvalue λ. The eigenvectors of A corresponding to λ are the nonzero solutions of $(A-\lambda I_n)\boldsymbol{x}=0$.

- The set of solutions of $(A-\lambda I_n)\boldsymbol{x}=0$ is called the **eigenspace of *A* corresponding to the eigenvalue λ**.
  - ✓ It is just the null space of $(A-\lambda I_n)$, the kernel of $(A-\lambda I_n)$, a subspace of R^n.
- If λ is an eigenvalue of a linear operator T on R^n, the set of vectors $\boldsymbol{v}$ in R^n such $T(\boldsymbol{v})=\lambda\boldsymbol{v}$ is called the **eigenspace** of T corresponding to λ.



W is the eigenspace of *T* corresponding to eigenvalue λ.

5.2 The characteristic polynomial

- If λ is an eigenvalue of A, there must be a nonzero vector $\boldsymbol{v}$ in R^n such that $A\boldsymbol{v}=\lambda\boldsymbol{v}$.
  - ✓ For an $n\times n$ matrix A and the homogeneous system of linear equations $(A-\lambda I_n)\boldsymbol{x}=0$ to have nonzero solutions, the rank of $A-\lambda I_n$ must be less than n. By the Invertible Matrix Theorem, $A-\lambda I_n$ is then not invertible, so its determinant must be zero.

The eigenvalues of a square matrix A are the values of t that satisfy

$$\det(A-tI_n)=0.$$

- The equation det $(A-\lambda I_n)=0$ is called the **characteristic equation** of A.
- det $(A-\lambda I_n)$ is called the **characteristic polynomial** of A, a polynomial of degree n.

Let R be the reduced row echelon form of an $n\times n$ matrix A.

- The characteristic polynomial of A is not usually equal to the characteristic polynomial of R!
- In general, the eigenvalues of A and R are not the same!
- The eigenvectors of A and R are not usually the same!

The eigenvalues of an upper triangular or low triangular matrix are its diagonal entries.

For a linear operator *T*:

- The characteristic equation of the standard matrix of *T* is called the **characteristic equation** of *T*.
- The characteristic polynomial of the standard matrix of *T* is called the **characteristic polynomial** of *T*.
- The characteristic polynomial of a linear operator *T* on R^n is a polynomial of degree *n*, whose roots are the eigenvalues of *T*.

The multiplicity of an eigenvalue:

If λ is an eigenvalue of an $n \times n$ matrix M, then the largest positive integer k such that $(t-\lambda)^k$ is a factor of the characteristic polynomial of M is called the **multiplicity** of λ.

- For matrix $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$, $\det(A - tI_3) = -(t+1)^2(t-3)$.
- For matrix $B = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$, $\det(B - tI_3) = -(t+1)(t-3)^2$.

Theorem 5.1

Let λ be an eigenvalue of a matrix A. The dimension of the eigenspace of A corresponding to λ is less than or equal to the multiplicity of λ.

(*Geometric multiplicity* $\leq$ *Algebraic multiplicity*)

The eigenvalue of similar matrices

Matrices A and B are similar if there exists an invertible matrix P such that $B=P^{-1}AP$.

By Theorem 3.4 we have

$$\det(B - tI_n) = \det(P^{-1}AP - tP^{-1}I_nP)$$
$$= \det(P^{-1}(A - tI_n)P)$$
$$= (\det P^{-1})\,[\det(A - tI_n)]\,(\det P)$$
$$= \det(A - tI_n)$$

Similar matrices have the same characteristic polynomial and hence have the same eigenvalues and multiplicities. In addition, their eigenspaces corresponding to the same eigenvalue have the same dimension.

Complex eigenvalues ($R \Rightarrow C$; $R^n \Rightarrow C^n$; Self-study)

5.3 Diagonalization of Matrices

Definition An $n \times n$ matrix A is called diagonalizable if $A = PDP^{-1}$ for some diagonal $n \times n$ matrix D and some invertible $n \times n$ matrix P.

- Because $A = PDP^{-1}$ can be written as $D = P^{-1}AP$, a diagonalizable matrix A is *similar* to a diagonal matrix D.
  - ✓ The eigenvalues of A are the diagonal entries of D.
- Every diagonal matrix is diagonalizable.
- Not every matrix is diagonalizable.

Theorem 5.2

An $n \times n$ matrix A is diagonalizable iff there is a basis for R^n consisting of eigenvectors of A.

Furthermore, $A=PDP^{-1}$, where D is a diagonal matrix and P is an invertible matrix iff the columns of P are a basis for R^n consisting of eigenvectors of A and the diagonal entries of D are the eigenvalues corresponding to the respective columns of P.

- Proof ?
- The matrices P, a modal matrix for A, and D, a spectral matrix for A, such that $PDP^{-1} = A$ in Theorem 5.2 are not unique.

Theorem 5.3

A set of eigenvectors of a square matrix that correspond to distinct eigenvalues is linearly independent.

- Proof ?
- Thus, an $n \times n$ matrix having n distinct eigenvalues must have n linearly independent eigenvectors.
- If the bases for distinct eigenspaces are combined, then the resulting set is linearly independent.

Every $n \times n$ matrix having n distinct eigenvalues is diagonalizable.

Algorithm for Matrix Diagonalization

Let A be a diagonalizable $n \times n$ matrix. Combining bases for each eigenspace of A forms a basis $\mathcal{B}$ for R^n consisting of eigenvectors of A. Therefore, if P is the matrix whose columns are the vectors in $\mathcal{B}$ and D is the diagonal matrix whose diagonal entries are eigenvalues of A corresponding to the respective columns of P, then $A = PDP^{-1}$.

Test for a Diagonalizable Matrix Whose Characteristic Polynomial is Known

An $n \times n$ matrix A is diagonalizable iff both of the following conditions are true:

1. The total number of eigenvalues of A, when each eigenvalue is counted as often as its multiplicity, is equal to n.
2. For each eigenvalue λ of A, the dimension of the corresponding eigenspace, which is $n - \text{rank}(A - \lambda I_n)$, is equal to the multiplicity of λ.

(Item 2 states that *Geometric multiplicity = Algebraic multiplicity for each eigenvalue.*)

- By Theorem 5.1, the eigenspace corresponding to an eigenvalue of multiplicity 1 must have dimension 1. Hence condition (2) need be checked only for eigenvalues of multiplicity greater than 1.

5.4* Diagonalization of linear operators

A linear operator on R^n is defined to be **diagonalizable** if there is a basis for R^n consisting of eigenvectors of the operator.

- *A linear operator is diagonalizable iff its standard matrix is diagonalizable*.

Let T be a linear operator on R^n for which there is a basis $\mathcal{B}=\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_\mathrm{n}\}$ consisting of eigenvectors of T.

- Then, $T(\boldsymbol{v}_i)=\lambda_i\boldsymbol{v}_i$ for each i, where λ_i is the eigenvalue corresponding to $\boldsymbol{v}_i$.
- Thus $[T(\boldsymbol{v}_i)]_{\mathcal{B}}=\lambda_i\boldsymbol{e}_i$ for each i, and so

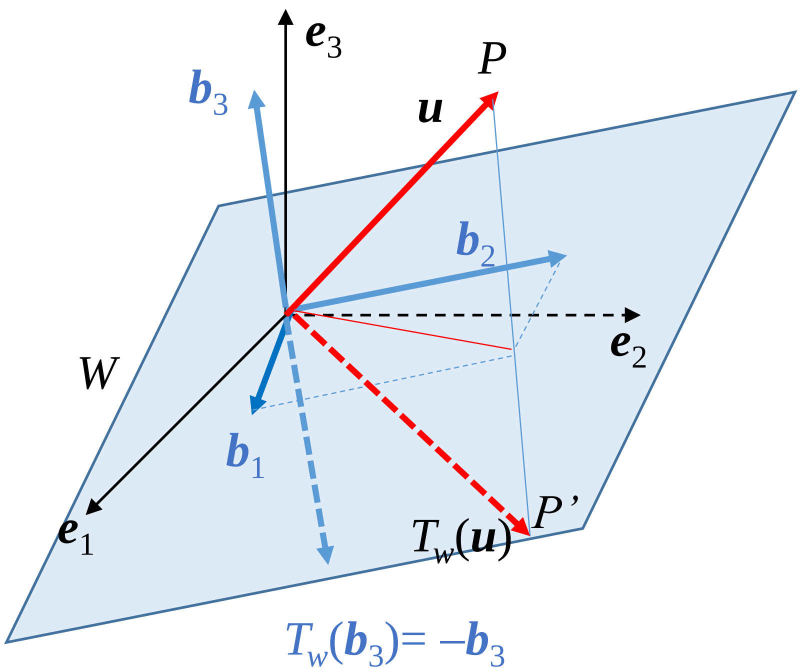
$[T]_{\mathcal{B}}=[\ [T(\boldsymbol{v}_1)]_{\mathcal{B}}\ \ [T(\boldsymbol{v}_2)]_{\mathcal{B}}\ \ \ldots\ [T(\boldsymbol{v}_n)]_{\mathcal{B}}\ ]=[\lambda_1\boldsymbol{e}_1\ \lambda_2\boldsymbol{e}_2\ \ \ldots\ \lambda_n\boldsymbol{e}_n\ ]$
is a diagonal matrix.

A linear operator T on R^n is diagonalizable iff there is a basis $\mathcal{B}$ for R^n such that $[T]_{\mathcal{B}}$, the $\mathcal{B}$-matrix of T, is a diagonal matrix. Such a basis $\mathcal{B}$ must consist of eigenvectors of T.

By Theorem 4.12, the $\mathcal{B}$-matrix of T is given by that $[T]_{\mathcal{B}} = B^{-1}AB$, where B is the matrix whose columns are the vectors in $\mathcal{B}$ and A is the standard matrix of T. So, let $\mathcal{B}$ consist of eigenvectors of T. Then $[T]_{\mathcal{B}}$ is the diagonal matrix whose diagonal entries are ……

A Reflection Operator

- Let W be a 2-dimensional subspace of $\mathcal{R}^3$, a plane containing the origin.
- T_W: $\mathcal{R}^3 \to \mathcal{R}^3$, defined as follows: For a vector $\boldsymbol{u}$ in $\mathcal{R}^3$ with end point P, drop a perpendicular from P to W, and extend this perpendicular an equal distance to the point P' on the other side of W. Then $T_W(\boldsymbol{u})$ is the vector with endpoint P'.



- The set $\mathcal{B}=\{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$ is linearly independent and hence a basis for $\mathcal{R}^3$.
- $T_W(\boldsymbol{b}_1) = \boldsymbol{b}_1$, $T_W(\boldsymbol{b}_2) = \boldsymbol{b}_2$, $T_W(\boldsymbol{b}_3) = -\boldsymbol{b}_3$.
  - ✓ $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$ are eigenvectors of T_W with corresponding eigenvalues 1, 1, and –1, respectively
  - ✓ $[T_W(\boldsymbol{b}_1)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $[T_W(\boldsymbol{b}_2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $[T_W(\boldsymbol{b}_3)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$, and

$$[T_W]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

5.5* Applications of Eigenvalues

Markov chain

- States
- Transition matrix
- Regular (All are *recurrent* states in one class)

Theorem 5.4

If A is a *regular* $n \times n$ transition matrix and $\boldsymbol{p}$ is a probability vector in R^n, then

a) 1 is an eigenvalue of A;

b) there is a unique probability vector $\boldsymbol{v}$ of A that is also an eigenvector corresponding to eigenvalue 1;

c) The vectors $A^m\boldsymbol{p}$ approach $\boldsymbol{v}$ for m=1, 2, 3, ….

Google searches (PageRank algorithm) ?