25. Lecture 25 7.3 Eigenvectors

We will show that some square matrices A can be factorized $A = SDS^{-1}$,
where D is diagonal (i.e. the entries off the main diagonal are all zeros).
This can be used to compute A^k, for large k, which is useful in the applications.
(If multiplying by A represents the evolution of a system during one time unit then multiplying by A^k represents the evolution after k time units)

Ex Let $D = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$. Compute D^2, D^3 and D^k.

Sol $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 4^2 \end{bmatrix}$, $D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 4^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 4^3 \end{bmatrix}$, $D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix}$.

Ex Let $A = \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$. Compute A^k. Use that $A = SDS^{-1}$; where $D = \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$, $S = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $S^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$. **Sol** We have $A^2 = SDS^{-1}SDS^{-1} = SDIDS^{-1} = SD^2S^{-1}, \ldots$, so

$$A^k = SD^kS^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2\cdot 5^k - 4^k & -5^k + 4^k \\ 2\cdot 5^k - 2\cdot 4^k & -5^k + 2\cdot 4^k \end{bmatrix}.$$

A square matrix A is called **diagonalizable** if it can be written $A = SDS^{-1}$, where D is diagonal and S is invertible. When is A diagonalizable and if it is how do we find D and S? The answer lies in the eigenvalues and eigenvectors. Note that

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

so the columns of S are made out of the eigenvectors of A and the diagonal entries of D are the eigenvalues of A. We can put this to equations together in one matrix equation:

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

i.e.

$$\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1},$$

In general if A is an $n \times n$ matrix with n linearly independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and eigenvalues $\lambda_1, \ldots, \lambda_n$ then

$$A\begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ A\mathbf{v}_1 & \cdots & A\mathbf{v}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \lambda_1\mathbf{v}_1 & \cdots & \lambda_n\mathbf{v}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

and hence

$$A = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}\begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix}^{-1}$$

We have hence proven:

Diagonalization Theorem An $n \times n$ matrix is diagonalizable A if and only if it has n linearly independent eigenvectors.

Ex If possible, diagonalize $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$.

Sol The eigenvalues $\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 1 & 2-\lambda & 1 \\ -1 & 0 & 1-\lambda \end{vmatrix} = (2-\lambda)^2(1-\lambda) = 0.$

Basis for $\lambda = 1$: $\mathbf{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$.

Basis for $\lambda = 2$: $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

Construct $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. $A = PDP^{-1}$.

Ex If possible, diagonalize $A = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{bmatrix}$.

Sol The eigenvalues $\det(A - \lambda I) = (\lambda - 2)^2(\lambda - 4) = 0$.

Basis for $\lambda = 2$: $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Basis for $\lambda = 4$: $\mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$.

There are not three linearly independent eigenvectors so A can not be diagonalized.

Th If $\lambda_1, \dots, \lambda_n$ are distinct eigenvalues of an $n \times n$ matrix A with corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, then $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.
Pf We argue by contradiction. Let $k \le n$ be the smallest integer so that $c_1\mathbf{v}_1 + \dots c_k\mathbf{v}_k = 0$ for some combination with some $c_k \neq 0$. Then multiplying with $(A - \lambda_k I)$ gives $c_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \dots c_k(\lambda_k - \lambda_k)\mathbf{v}_k = 0$, and since $(\lambda_i - \lambda_k) \neq 0$, for $i < k$ it follows that we have a linear combination with few vectors, which contradicts our assumption.

Th If A is symmetric matrix $A^T = A$ then A has n linearly independent Eigenvectors.
We will study diagonalization for symmetric matrices in the next chapter so we postpone the proof.

Th If B is similar to A, i.e. $B = S^{-1}AS$ then A and B have the same characteristic polynomial and hence the same eigenvalues.
Pf Since we can write $I = S^{-1}S = S^{-1}IS$ we get

$$\det(B - \lambda I) = \det\big(S^{-1}AS - \lambda S^{-1}IS\big) = \det\big(S^{-1}(A - \lambda I)S\big)$$
$$= \det S^{-1}\,\det(A - \lambda I)\,\det S = \det(A - \lambda I),$$

by the product rules for determinants: $(\det(CD) = \det C\,\det D)$.

This theorem says something very important; that the eigenvalues does not depend on in which coordinate system we view a linear transformation, and hence describe some fundamental property of the linear transformation.

Def The **geometric multiplicity** of an eigenvalue λ_0 is the dimension of the eigenspace $E_{\lambda_0} = \operatorname{Ker}(A - \lambda_0 I)$. The **algebraic multiplicity** is the integer k such that $p_A(\lambda) = (\lambda - \lambda_0)^k g(\lambda)$, where g is a polynomial with $g(\lambda_0) \neq 0$.

Th The geometric multiplicity is less than or equal to the algebraic multiplicity.
Pf Suppose λ_0 is an eigenvalue of an $n \times n$ matrix A with multiplicity m. Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ be a basis for E_{λ_0}. Let S be an invertible matrix with the first m columns consisting of $\mathbf{v}_1, \dots, \mathbf{v}_m$ (You can find such a matrix by letting the remaining columns be a basis for the orthogonal complement of the first m columns.) Then $B = S^{-1}AS$ is similar to A. We compute $B\mathbf{e}_i = S^{-1}AS\mathbf{e}_i = S^{-1}A\mathbf{v}_i = \lambda_0 S^{-1}\mathbf{v}_i = \lambda_0\mathbf{e}_i$. It follows that B is a block matrix of the form $B = \begin{bmatrix} \lambda_0 I & P \\ 0 & Q \end{bmatrix}$, where I in the $m \times m$ identity matrix. By the previous theorem $p_A(\lambda) = p_B(\lambda)$. But because of the block structure of B we have $\det(B - \lambda I) = \det(\lambda_0 I - \lambda I)\det(Q - \lambda I) = (\lambda_0 - \lambda)^m \det(Q - \lambda I)$. Hence $m \le k$, where k is the algebraic multiplicity.

Summary

A square matrix A is called **diagonalizable** if $A = SDS^{-1}$, for some S where D is diagonal.

Th An $n\times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.

If A is an $n \times n$ matrix with n linearly independent eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ and eigenvalues $\lambda_1, \dots, \lambda_n$ then

$$A \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ A\mathbf{v}_1 & \cdots & A\mathbf{v}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \lambda_1\mathbf{v}_1 & \cdots & \lambda_n\mathbf{v}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

and hence

$$A = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix}^{-1}$$

Th If $\lambda_1, \dots, \lambda_n$ are distinct eigenvalues of an $n \times n$ matrix A with corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, then $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.
Pf We argue by contradiction. Let $k \le n$ be the smallest integer so that $c_1\mathbf{v}_1 + \dots c_k\mathbf{v}_k = 0$ for some combination with some $c_k \ne 0$. Then multiplying with $(A - \lambda_k I)$ gives $c_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \dots c_k(\lambda_k - \lambda_k)\mathbf{v}_k = 0$, and since $(\lambda_i - \lambda_k) \ne 0$, for $i < k$ it follows that we have a linear combination with few vectors, which contradicts our assumption.

Th If A is symmetric matrix $A^T{=}A$ then A has n linearly independent Eigenvectors.
We will study diagonalization for symmetric matrices in the next chapter.

Th If B is similar to A, i.e. $B = S^{-1}AS$ then A and B have the same characteristic polynomial and hence the same eigenvalues.

Hence the eigenvalues does not depend on in which coordinate system we view a linear transformation, and hence describe some fundamental property of the linear transformation.

Def The **geometric multiplicity** of an eigenvalue λ_0 is the dimension of the eigenspace $E_{\lambda_0} = \mathrm{Ker}(A - \lambda_0 I)$. The **algebraic multiplicity** is the integer k such that $p_A(\lambda) = (\lambda - \lambda_0)^k g(\lambda)$, where g is a polynomial with $g(\lambda_0) \ne 0$.

Th The geometric multiplicity is less than or equal to the algebraic multiplicity.
Pf Suppose λ_0 is an eigenvalue of an $n\times n$ matrix A with multiplicity m. Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ be a basis for E_{λ_0}. Let S be an invertible matrix with the first m columns consisting of $\mathbf{v}_1, \dots, \mathbf{v}_m$ (You can find such an S by letting the remaining columns be a basis for the orthogonal complement of the first m columns.) Then $B = S^{-1}AS$ is similar to A. We compute $B\mathbf{e}_i = S^{-1}AS\mathbf{e}_i = S^{-1}A\mathbf{v}_i = \lambda_0 S^{-1}\mathbf{v}_i = \lambda_0\mathbf{e}_i$. It follows that B is a block matrix of the form $B = \begin{bmatrix} \lambda_0 I & P \\ 0 & Q \end{bmatrix}$, where I in the $m \times m$ identity matrix. By the previous theorem $p_A(\lambda) = p_B(\lambda)$. But because of the block structure of B we have $\det(B - \lambda I) = \det(\lambda_0 I - \lambda I)\det(Q - \lambda I) = (\lambda_0 - \lambda)^m \det(Q - \lambda I)$. Hence $m \le$ the algebraic multiplicity.