Lecture 21: 5.6 Similarity transformations.
We have seen that if an $n \times n$ matrix A has n linearly independent eigenvectors and S is the matrix with the eigenvectors as its columns then $S^{-1}AS$ is diagonal. Two matrices A and B are called **similar** if they are related by a **similarity transformation** $B = M^{-1}AM$. Similar matrices have the same eigenvalues.
The question is if we can transform a general matrix into a simple form in this way. Similarity transformations show up naturally by changes of variables in differential equations or changes of basis in which a linear transformation is expressed.

$$\text{If}\quad \mathbf{u} = M\mathbf{v} \quad\text{then}\quad \frac{d\mathbf{u}}{dt} = A\mathbf{u} \quad\text{becomes}\quad M\frac{d\mathbf{v}}{dt} = AM\mathbf{v} \quad\text{or}\quad \frac{d\mathbf{v}}{dt} = M^{-1}AM\mathbf{v}$$

$$\text{If } \mathbf{u}_k = M\mathbf{v}_k \text{ then } \mathbf{u}_{n+1} = A\mathbf{u}_n \text{ becomes } M\mathbf{v}_{n+1} = AM\mathbf{v}_n \text{ or } \mathbf{v}_{n+1} = M^{-1}AM\mathbf{v}_n$$

Expressing a linear transformation in terms of different bases.
Ex Let L be the line in $\mathbf{R}^2$ that is spanned by the vector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$.
Let $\mathbf{T}$ be the linear transformation that projects any vector orthogonally onto L. Find the matrix A representing $\mathbf{T}$ in the standard basis $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
Sol We pick a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ in which we will represent the linear transformation, i.e. we write both $\mathbf{x}$ and $\mathbf{T}(\mathbf{x})$ in terms of the basis:

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 = \mathbf{x} \to \mathbf{T}(\mathbf{x}) = d_1\mathbf{b}_1 + d_2\mathbf{b}_2.$$

The transformation taking, $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, the coordinates of $\mathbf{x}$ in the $\mathcal{B}$ basis, to $\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$, the coordinates of $\mathbf{T}(\mathbf{x})$, is a linear so it is given by multiplication with a matrix B

$$B\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}.$$

In particular if we choose $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ parallel to the line and $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ perpendicular to the line then $\mathbf{T}(\mathbf{b}_1) = \mathbf{b}_1$ and $\mathbf{T}(\mathbf{b}_2) = 0$. Hence $\mathbf{T}(c_1\mathbf{b}_1 + c_2\mathbf{b}_2) = c_1\mathbf{b}_1$, i.e.

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

We can similarly express T in the standard basis

$$x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = \mathbf{x} \to \mathbf{T}(\mathbf{x}) = y_1\mathbf{e}_1 + y_2\mathbf{e}_2.$$

The matrix for A for $\mathbf{T}$ in the standard coordinates

$$A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

is more complicated but one can calculate it from B:

$$\begin{array}{ccc} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & \xrightarrow{A} & \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\ {\scriptstyle P}\big\uparrow & & \big\uparrow{\scriptstyle P} \\ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} & \xrightarrow{B} & \begin{bmatrix} c_1 \\ 0 \end{bmatrix} \end{array},$$

where P is the change of basis matrix such that $\mathbf{x}_1\mathbf{e}_1 + x_2\mathbf{e}_2 = c_1\mathbf{b}_1 + c_2\mathbf{b}_2$ i.e.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = P \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

so $P = [\,\mathbf{b}_1 \;\; \mathbf{b}_2\,] = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}$ and $P^{-1} = \dfrac{1}{10}\begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix}$. Hence

$$A = PBP^{-1} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{10} \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}$$

Triangular form with a unitary M.

Schur's lemma There is a unitary matrix U such that $U^{-1}AU = T$ is triangular.

Proof Every matrix has at least one eigenvalue λ_1 and for this eigenvalue we pick an orthonormal eigenvector $\mathbf{u}_1$. Then pick there other vectors so $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ are orthonormal and set $U_1 = [\,\mathbf{u}_1\,\mathbf{u}_2\,\mathbf{u}_3\,\mathbf{u}_4\,]$. Then $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$ and in general $A\mathbf{u}_i = \sum_{j=1}^4 c_{ij}\mathbf{u}_j$ so

$$AU_1 = [\,A\mathbf{u}_1\,A\mathbf{u}_2\,A\mathbf{u}_3\,A\mathbf{u}_4\,] = U_1 \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix},$$

and hence

$$U_1^{-1}AU_1 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & & & \\ 0 & & A_2 & \\ 0 & & & \end{bmatrix}.$$

Similarly, the 3×3 matrix A_2 has an eigenvalue λ_2 and we can form M_2 such that

$$M_2^{-1}A_2M_2 = \begin{bmatrix} \lambda_2 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

Hence if $U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2 & \\ 0 & & & \end{bmatrix}$ then $U_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2^{-1} & \\ 0 & & & \end{bmatrix}$ and

$$U_2^{-1}U_1^{-1}AU_1U_2 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & & & \\ 0 & & M_2^{-1}A_2M_2 & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}.$$

Continuing in this way we get

$$U_4^{-1}U_3^{-1}U_2^{-1}U_1^{-1}AU_1U_2U_3U_4 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & \lambda_2 & * & * \\ 0 & 0 & \lambda_3 & * \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix} = T.$$

and the lemma follows with $U = U_1U_2U_3U_4$.

Diagonalizing Symmetric and Hermitian Matrices.
Spectral Theorem If A is a real and symmetric $n \times n$ matrix then it has an orthonormal set of n eigenvectors and hence can be diagonalized by an orthogonal matrix U: $U^{-1}AU = D$, or $A = UDU^{-1}$, where D is diagonal and $U^{-1} = U^T$.
Proof It follows from Schur's lemma that $U^TAU = T$ is upper triangular. However $T^T = (U^TAU)^T = U^TA^T(U^T)^T = U^TAU = T$, since $A^T = A$, so T is diagonal.

Ex Diagonalize $A = \begin{bmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{bmatrix}$ with an orthogonal transformation.

Sol A is symmetric so it can be diagonalized by an orthogonal transformation. The eigenvalues are $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 5$. The eigenspace corresponding to eigenvalue -1; $(A+I)\mathbf{x}=\mathbf{0}$ satisfy $x_1 + 2x_2 - x_3 = 0$ so $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ form a basis for the eigenspace corresponding to $\lambda = -1$. We can apply the Gram-Schmidt process to obtain an orthonormal basis. Let

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{p}_2 = \mathbf{v}_2 \cdot \mathbf{u}_1\, \mathbf{u}_1 = -\sqrt{2}\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{\mathbf{v}_2 - \mathbf{p}_2}{\|\mathbf{v}_2 - \mathbf{p}_2\|} = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

The eigenspace corresponding to $\lambda_3 = 5$ is spanned by $\mathbf{v}_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$ and we set

$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$$

Hence $A = UDU^T$ where

$$U = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix}, \qquad D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

Normal matrices $NN^H = N^HN$ are exactly those that have a complete set of orthonormal eigenvectors and the triangular factorization $T = U^TNU$ is diagonal.