Diagonalizability

(Sections 3.3 and 8.2)

Wish: every matrix is diagonal!

We'd wish this because finding determinants is easy for diagonal matrices, any two diagonal matrices commute with respect to matrix multiplication, and more.

Clearly this isn't a realistic wish, so what can we wish for?

Diagonalizability

Let A be an $n \times n$ matrix. We say A is **diagonalizable** if there is an invertible $n \times n$ matrix S and a diagonal $n \times n$ matrix D with

$$\boxed{A = SDS^{-1}}$$

Q: Is every $n \times n$ matrix diagonalizable?

A: NO! not even for 2×2 matrices...

Example 1: Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Suppose A was diagonalizable,

$A = SDS^{-1}$ where $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$.

Calculating the eigenvalues of A,

$$\det\left(A - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

$$= \det\left(\begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix}\right)$$

$$= \lambda^2.$$

Setting this equal to zero, we see that $\lambda = 0$ is the only eigenvalue of A.

But

$$\lambda^2 = \det(A - \lambda I_2) = \det(SDS^{-1} - \lambda I_2)$$

$$= \det(SDS^{-1} - \lambda S \cdot S^{-1})$$

$$= \det(SDS^{-1} - \lambda S I_2 S^{-1})$$

$$= \det(S(D - \lambda I_2)S^{-1})$$

$$= \det(S)\det(D - \lambda I_2)\det(S^{-1})$$

$$= \cancel{\det(S)}\ \det(D - \lambda I_2)\ \frac{1}{\cancel{\det(S)}}$$

$$= \det(D - \lambda I_2)$$

$$= \det\left(\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right)$$

$$= \det\left(\begin{bmatrix} d_1 - \lambda & 0 \\ 0 & d_2 - \lambda \end{bmatrix}\right)$$

$$= (d_1 - \lambda)(d_2 - \lambda)$$

So $\lambda^2 = (d_1 - \lambda)(d_2 - \lambda)$,

which implies $d_1 = d_2 = 0$.

Then $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$,

and so if

$$A = SDS^{-1}, \text{ then}$$

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = S \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} S^{-1}$$

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which is not true!

This shows that $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

is not diagonalizable!

Q: When is a matrix diagonalizable?

Unfortunately, there is no really good answer to this question, in that all conditions are seemingly as hard as diagonalizability to check.

Linear Independence, Bases, and Dimension

A collection of vectors S in a vector space V is said to be **linearly independent** if no vector in S is a linear combination of other vectors in S. Written formally, if $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n \in S$ and there are scalars $c_1, c_2, \ldots, c_n$ with

$$\boxed{c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n = 0_V},$$

then $\underline{c_1 = c_2 = \cdots = c_n = 0}$

$(c_i = 0 \text{ for all } 1 \le i \le n)$

A subset $\mathcal{B}$ of V is called a basis for V if both

1) $\mathcal{B}$ is linearly independent

2) $\mathrm{span}(\mathcal{B}) = V$.

Finally, we can define the dimension of V to be the number of elements in a basis, denoted $\underline{\dim(V)}$.
If V has no finite basis, we set $\dim(V) = \infty$.

Aside: (summing up spans) We now have the terminology to say:

1) If $\vec{v} = 0_V$, then $\mathrm{span}(\vec{v})$ is a point, and zero-dimensional.

2) If $\vec{v} \neq 0_V$, then $\mathrm{span}(\vec{v})$ is a line, and one-dimensional.

3) If $\vec{v} \neq 0_V \neq \vec{w}$, then $\mathrm{span}(\{\vec{v}, \vec{w}\})$ is either a line (one-dimensional) or a plane, and two-dimensional.

4) Given three nonzero vectors $\vec{v}$, $\vec{w}$, and $\vec{x}$, then $\mathrm{span}(\{\vec{v}, \vec{w}, \vec{x}\})$ is

either a line (one-dimensional),

a plane (two-dimensional), or

three-dimensional.

Eigenspaces

If A is an $n \times n$ matrix and λ is an eigenvalue of A, define the eigenspace $E_\lambda \subseteq \mathbb{R}^n$ to be

$$\boxed{E_\lambda = \{ \vec{x} \in \mathbb{R}^n \mid A\vec{x} = \lambda \vec{x} \}}$$

Note: here, $\vec{x} = \vec{0}$ is allowed, and E_λ is a subspace of $\mathbb{R}^n$.

Criteria for Diagonalizability

Let A be an $n \times n$ matrix.

Then A is diagonalizable precisely when either of the following two conditions hold:

1) There is a basis $\mathcal{B}$ for $\mathbb{R}^n$ consisting of eigenvectors of A.

2) For all eigenvalues γ of A,

$$\underbrace{\dim(E_\gamma)}_{\text{geometric multiplicity of } \gamma} = \text{the number of factors of } (\lambda - \gamma) \text{ in } \underbrace{\det(A - \lambda I_n)}_{\text{algebraic multiplicity of } \gamma}$$

But... neither of these conditions are particularly easy to check. We want a condition that is easy to check, and implies diagonalizability.

Symmetric and Orthogonal matrices

A matrix $A \in M_n(\mathbb{R})$ is said to be **symmetric** if

$$\boxed{A = A^t}$$

A matrix $O \in M_n(\mathbb{R})$ is said to be **orthogonal** (NOT usually an "orthogonal projection") if

$$\boxed{O^t O = I_n}, \text{ i.e., }$$

$$O^t = O^{-1}.$$

Orthogonal Diagonalizability

If A is an $n \times n$ symmetric matrix, then there is a diagonal matrix $D \in M_n(\mathbb{R})$ and an $n \times n$ orthogonal matrix O with

$$\boxed{A = ODO^t}$$

So symmetric matrices are not only diagonalizable, you can choose the invertible matrix to be orthogonal.

Example 2: Let $A = \begin{bmatrix} 1 & 3 \\ 3 & -4 \end{bmatrix}$.

Then A is symmetric. Orthogonally diagonalize A.

Solution: The diagonal matrix D will have the eigenvalues of A on the diagonal, so let's calculate the eigenvalues:

$$\det(A - \lambda I_2)$$

$$= \det\left(\begin{bmatrix} 1 & 3 \\ 3 & -4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right)$$

$$= \det\left(\begin{bmatrix} 1-\lambda & 3 \\ 3 & -4-\lambda \end{bmatrix} \right)$$

$$= (1-\lambda)(-4-\lambda) - 9$$

$$= \lambda^2 + 3\lambda - 13$$

Set equal to zero:

$$0 = \lambda^2 + 3\lambda - 13$$

Use quadratic formula.

$$\lambda = \frac{-3 \pm \sqrt{9-4(-13)}}{2}$$

$$\boxed{\lambda = \frac{-3 \pm \sqrt{61}}{2}}$$

According to Wolfram Alpha,

the eigenvectors are

$$\begin{bmatrix} \frac{5 \pm \sqrt{61}}{6} \\ 1 \end{bmatrix}$$

for $\lambda_1 = \frac{-3 + \sqrt{61}}{2}$, $\vec{v}_1 = \begin{bmatrix} \frac{5+\sqrt{61}}{6} \\ 1 \end{bmatrix}$

for $\lambda_2 = \frac{-3 - \sqrt{61}}{2}$, $\vec{v}_2 = \begin{bmatrix} \frac{5-\sqrt{61}}{6} \\ 1 \end{bmatrix}$.

Then

$$A = \begin{bmatrix} 1 & 3 \\ 3 & -4 \end{bmatrix} = O D O^t$$

where

$$D = \begin{array}{c} \begin{array}{cc} \lambda_1 & \end{array} \\ \begin{bmatrix} \frac{-3+\sqrt{61}}{2} & 0 \\ 0 & \frac{-3-\sqrt{61}}{2} \end{bmatrix} \\ \begin{array}{cc} & \lambda_2 \end{array} \end{array}$$

and

$$O = \begin{array}{c} \begin{array}{cc} \vec{v}_1 & \vec{v}_2 \end{array} \\ \begin{bmatrix} \frac{5+\sqrt{61}}{6} & \frac{5-\sqrt{61}}{6} \\ 1 & 1 \end{bmatrix} \end{array},$$

right?

$$O^t O = \begin{bmatrix} \frac{5+\sqrt{61}}{6} & 1 \\ \frac{5-\sqrt{61}}{6} & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{5+\sqrt{61}}{6} & \frac{5-\sqrt{61}}{6} \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1+\left(\frac{5+\sqrt{61}}{6}\right)^2 & * \\ * & 1+\left(\frac{5-\sqrt{61}}{6}\right)^2 \end{bmatrix}$$

where

$$* = \left(\frac{5+\sqrt{61}}{6}\right) \cdot \left(\frac{5-\sqrt{61}}{6}\right) + 1$$

So

$$* = \frac{1}{36}\left(25 - 61\right) + 1$$

$$* = \frac{-36}{36} + 1 = 0.$$

Then

$$O^t O = \begin{bmatrix} 1 + \left(\frac{5+\sqrt{61}}{6}\right)^2 & 0 \\ 0 & 1 + \left(\frac{5-\sqrt{61}}{6}\right)^2 \end{bmatrix}$$

$$\neq I_2$$

So O is NOT orthogonal!

Easy fix:

replace $\vec{v}_1$ and $\vec{v}_2$ with unit vectors

$$\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|_2}, \quad \vec{w}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|_2}.$$

Then

$$A = O D O^t$$

with D as before and

$$O = \begin{bmatrix} \vec{w}_1 & \vec{w}_2 \end{bmatrix}.$$

Observations: (eigenvalues and eigenvectors)

If you can diagonalize $A = SDS^{-1}$,

then:

1) The diagonal entries of D are the eigenvalues of A, appearing as many times as their associated factor occurs in $\det(A - \lambda I_n)$.

2) The columns of S will be eigenvectors for A, corresponding to the eigenvalue in the same column of D. i.e., if

$$S = \begin{bmatrix} \vec{s}_1 & \vec{s}_2 & \cdots & \vec{s}_n \end{bmatrix}$$

$$\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n$$

eigenvectors