Section 6.4 Symmetric Matrices

We cannot diagonalize every $n \times n$ A: $A = X \Lambda X^{-1}$ because $\mathbb{R}^n$ might not have a basis of eigenvectors.

Helps with calculating A^N, for example; $A^N = X \Lambda^N X^{-1}$.

But if S is symmetric, $S = S^T$, we can always diagonalize.

We can say even more:

"Spectral Theorem"

1. Eigenvalues of symmetric S are all real numbers.
2. S can be diagonalized, even if there are repeated eigenvalues.
3. Eigenvectors with different eigenvalues are orthogonal.

From 2 and 3, $\mathbb{R}^n$ has an orthonormal basis of eigenvectors:

- If eigenvalues are all different, eigenvectors are already orthogonal, so just need to rescale to get unit vectors.
- If an eigenvalue is repeated, can use Gram-Schmidt process to get an orthonormal basis for each eigenspace.

This means $S = Q \Lambda Q^T$

For orthogonal matrix, inverse = transpose

matrix of orthonormal eigenvectors $\longrightarrow$ orthogonal matrix

diagonal matrix of real eigenvalues

Example (Problem 6.4.7)

$$S = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{bmatrix}$$

Eigenvalues: $\det(S-\lambda I) = \begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & -1-\lambda & -2 \\ 2 & -2 & -\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -1-\lambda & -2 \\ -2 & -\lambda \end{vmatrix} + 2\begin{vmatrix} 0 & -1-\lambda \\ 2 & -2 \end{vmatrix}$

$= (1-\lambda)(\lambda^2+\lambda-4) + 4(1+\lambda) = \cancel{\lambda^2}+\lambda-\cancel{4}-\lambda^3-\cancel{\lambda^2}+4\lambda+\cancel{4}+4\lambda$

$= 9\lambda-\lambda^3 = \lambda(9-\lambda^2) = \lambda(3-\lambda)(3+\lambda) = 0 \longrightarrow \lambda = 0, 3, -3$

Eigenvectors for $\lambda=0$: Solve $S\vec{x}=\vec{0}$

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{bmatrix} \xrightarrow[-\text{Row } 2]{\text{Row}3-2\text{Row}1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & -2 & -4 \end{bmatrix} \xrightarrow{\text{Row } 3+2\text{Row}2} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{array}{l} x_1 = -2x_3 \\ x_2 = -2x_3 \\ x_3 \text{ free} \end{array}$$

So $\vec{x} = x_3 \begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}$

$\lambda=3$: Solve $(S-3I)\vec{x}=\vec{0}$

$$\begin{bmatrix} -2 & 0 & 2 \\ 0 & -4 & -2 \\ 2 & -2 & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 2 & -2 & -3 \end{bmatrix} \xrightarrow[2\text{Row }1]{\text{Row } 3 -} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & -2 & -1 \end{bmatrix} \xrightarrow[+2\text{Row}2]{\text{Row } 3} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

So $\begin{array}{l} x_1 = x_3 \\ x_2 = -\frac{1}{2}x_3 \\ x_3 \text{ free} \end{array} \longrightarrow \vec{x} = x_3 \begin{bmatrix} 1 \\ -1/2 \\ 1 \end{bmatrix} = \left(\frac{1}{2}x_3\right)\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$

$\lambda=-3$: Solve $(S+3I)\vec{x}=\vec{0}$

$$\begin{bmatrix} 4 & 0 & 2 \\ 0 & 2 & -2 \\ 2 & -2 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1 \\ 2 & -2 & 3 \end{bmatrix} \xrightarrow{\text{Row } 3-2\text{Row}1} \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1 \\ 0 & -2 & 2 \end{bmatrix} \xrightarrow[+2\text{Row}2]{\text{Row}3} \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

So $\begin{array}{l} x_1 = -\frac{1}{2}x_3 \\ x_2 = x_3 \\ x_3 \text{ free} \end{array} \longrightarrow \vec{x} = x_3 \begin{bmatrix} -1/2 \\ 1 \\ 1 \end{bmatrix} = \left(\frac{1}{2}x_3\right)\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$

Basis for $\mathbb{R}^3$ of eigenvectors: $\begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$ orthogonal, but not orthonormal.

Turn into unit vectors to get an orthonormal basis:

$$\vec{x}_1 = \frac{1}{3}\begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \frac{1}{3}\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}, \quad \vec{x}_3 = \frac{1}{3}\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$$

Now we can diagonalize S:

$$S = Q\Lambda Q^T = \underbrace{\begin{bmatrix} -2/3 & 2/3 & -1/3 \\ -2/3 & -1/3 & 2/3 \\ 1/3 & 2/3 & 2/3 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}}_{\text{check: } \begin{bmatrix} 0 & 2 & 1 \\ 0 & -1 & -2 \\ 0 & 2 & -2 \end{bmatrix}}\begin{bmatrix} -2/3 & -2/3 & 1/3 \\ 2/3 & -1/3 & 2/3 \\ -1/3 & 2/3 & 2/3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{bmatrix} = S \quad \checkmark$$

Now let's prove some parts of the Spectral Theorem:

Orthogonal Eigenvectors: Say $\vec{x}_1$ and $\vec{x}_2$ are eigenvectors.

So $\begin{cases} S\vec{x}_1 = \lambda_1 \vec{x}_1 \\ S\vec{x}_2 = \lambda_2 \vec{x}_2 \end{cases}$ What is $\vec{x}_1 \cdot \vec{x}_2$?

$$(S\vec{x}_1)\cdot\vec{x}_2 = (\vec{x}_1^T S^T)\vec{x}_2 \underset{\uparrow}{=} \vec{x}_1^T(S\vec{x}_2) = \vec{x}_1\cdot(S\vec{x}_2)$$

S is symmetric

$$(S\vec{x}_1)\cdot\vec{x}_2 = \lambda_1(\vec{x}_1\cdot\vec{x}_2), \qquad \vec{x}_1\cdot(S\vec{x}_2) = \lambda_2(\vec{x}_1\cdot\vec{x}_2)$$

$$(\lambda_1 - \lambda_2)(\vec{x}_1\cdot\vec{x}_2) = 0$$

Only two possibilities: $\lambda_1 = \lambda_2$ or $\vec{x}_1\cdot\vec{x}_2 = 0$

So if $\vec{x}_1$ and $\vec{x}_2$ are eigenvectors for <u>different</u> eigenvalues ($\lambda_1 \neq \lambda_2$), then $\vec{x}_1 \perp \vec{x}_2$. ✓

<u>Real</u> <u>Eigenvalues</u> Say $S\vec{x} = \lambda\vec{x}$ Might also be complex...

Could λ be a <u>complex</u> number $\lambda = a + ib$? We want to show $b = 0$.

Complex conjugate of λ: $\bar{\lambda} = a - ib$.

Note: To say $b=0$ is the same as to say $\bar{\lambda} = a - i0 = a + i0 = \lambda$.

So actually we need to show $\lambda = \bar{\lambda}$.

<u>Claim</u>: $\bar{\lambda}$ is <u>also</u> an eigenvalue of S. Why?

Start with $S(\vec{u} + i\vec{v}) = (a+ib)(\vec{u} + i\vec{v})$

Apply complex conjugation: $\bar{S}(\vec{u} - i\vec{v}) = (a - i\vec{b})(\vec{u} - i\vec{v})$

<u>same</u> as S, since S is <u>real</u>.

So $\bar{\lambda} = a - ib$ <u>is</u> an eigenvalue, has eigenvector $\vec{u} - i\vec{v}$.

Now let's look at another dot product:

$$(\vec{u} - i\vec{v})^T (S(\vec{u} + i\vec{v})) \underset{\uparrow}{=} ((\vec{u} - i\vec{v})^T S^T)(\vec{u} + i\vec{v})$$

S is symmetric

$$(\vec{u} - i\vec{v})^T (S(\vec{u} + i\vec{v})) = \lambda(\vec{u} - i\vec{v})^T(\vec{u} + i\vec{v}) = \lambda(\vec{u}^T\vec{u} - \cancel{i\vec{v}^T\vec{u}} + \cancel{i\vec{u}^T\vec{v}} - i^2\vec{v}^T\vec{v}) = \lambda(\|\vec{u}\|^2 - (-1)\|\vec{v}\|^2) = \lambda(\|\vec{u}\|^2 + \|\vec{v}\|^2)$$

$$((\vec{u} - i\vec{v})^T S^T)(\vec{u} + i\vec{v}) = (S(\vec{u} - i\vec{v}))^T(\vec{u} + i\vec{v}) = \bar{\lambda}(\vec{u} - i\vec{v})^T(\vec{u} + i\vec{v})$$

also equals

$$\bar{\lambda}\,\underbrace{(\|\vec{u}\|^2 + \|\vec{v}\|^2)}$$

not 0 since $\vec{u}$ and $\vec{v}$ can't <u>both</u> be $\vec{0}$ (since $\vec{u} + i\vec{v}$ is a non-zero eigenvector)

$$\lambda(\|\vec{u}\|^2 + \|\vec{v}\|^2) = \bar{\lambda}(\|\vec{u}\|^2 + \|\vec{v}\|^2)$$

Cancel $\|\vec{u}\|^2 + \|\vec{v}\|^2$, get $\lambda = \bar{\lambda} \rightarrow \boxed{\lambda \text{ is real}}$

If S has <u>all different</u> eigenvalues, we know we can diagonalize now: $S = Q\Lambda Q^T$.

But we can diagonalize even if eigenvalues are repeated. (I won't prove it.)

<u>Example</u> $S = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$

<u>Eigenvalues</u>: $\begin{vmatrix} -1-\lambda & 1 & 1 \\ 1 & -1-\lambda & 1 \\ 1 & 1 & -1-\lambda \end{vmatrix} = (-1-\lambda)\begin{vmatrix} -1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 1 & -1-\lambda \end{vmatrix} + 1\begin{vmatrix} 1 & -1-\lambda \\ 1 & 1 \end{vmatrix}$

$= (-1-\lambda)(\lambda^2+2\lambda\cancel{+1}\cancel{-1}) - (-1-\lambda-1) + (1+1+\lambda)$

$= -\lambda^2-\lambda^3-2\lambda-2\lambda^2+\lambda+2+\lambda+2 = -\lambda^3-3\lambda^2+4 = 0$

By inspection: $\lambda = 1$ is a root. So can factor out $1-\lambda$:

$(1-\lambda)(\lambda^2+a\lambda+b) = -\lambda^3 + \underbrace{(1-a)}_{-3}\lambda^2 + \underbrace{(a-b)}_{0}\lambda + \underbrace{b}_{4}$

So $a=b=4$: $(1-\lambda)\underbrace{(\lambda^2+4\lambda+4)}_{(\lambda+2)^2} = 0 \longrightarrow \lambda = 1, -2, -2$.

<u>Eigenvectors for $\lambda = 1$</u>:

Solve $(S-I)\vec{x} = \vec{0}$

$$\underbrace{\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}}_{\text{Entries in each row add to } 0}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

one eigenvector is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

<u>Eigenvectors for $\lambda = -2$</u>:

Solve $(S+2I)\vec{x} = \vec{0}$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

So $x_1+x_2+x_3 = 0$ $\rightsquigarrow$

$$\vec{x} = \begin{bmatrix} -x_2-x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

So here is a basis of eigenvectors:

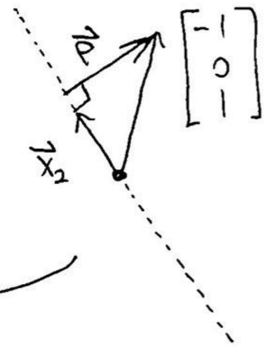
$$\underbrace{\begin{bmatrix}1\\1\\1\end{bmatrix}}_{\lambda=1},\ \underbrace{\begin{bmatrix}-1\\1\\0\end{bmatrix},\begin{bmatrix}-1\\0\\1\end{bmatrix}}_{\lambda=-2}$$

But it's not orthonormal, or even orthogonal.

rescale to unit vector

$$\vec{x}_1=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix},\quad \vec{x}_2=\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix}$$

Use Gram-Schmidt process:



Vector $\vec{e}$ is orthogonal to $\vec{x}_2$, <u>and</u> it's still an eigenvector.

$$\vec{e}=\begin{bmatrix}-1\\0\\1\end{bmatrix}-P\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}-1\\0\\1\end{bmatrix}-\frac{\vec{x}_2\vec{x}_2^T}{\underbrace{\vec{x}_2^T\vec{x}_2}_{=1\ (\vec{x}_2\text{ is a unit vector})}}\begin{bmatrix}-1\\0\\1\end{bmatrix}$$

projection onto span($\vec{x}_2$)

$$=\begin{bmatrix}-1\\0\\1\end{bmatrix}-\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix}\frac{1}{\sqrt{2}}\underbrace{\begin{bmatrix}-1&1&0\end{bmatrix}\begin{bmatrix}-1\\0\\1\end{bmatrix}}_{1}=\begin{bmatrix}-1\\0\\1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}-1\\1\\0\end{bmatrix}=\begin{bmatrix}-1/2\\-1/2\\1\end{bmatrix}$$

Finally, $\vec{x}_3=\dfrac{1}{\|\vec{e}\|}\vec{e}=\dfrac{1}{\sqrt{(-\frac{1}{2})^2+(-\frac{1}{2})^2+1^2}}\begin{bmatrix}-1/2\\-1/2\\1\end{bmatrix}=\sqrt{\dfrac{2}{3}}\begin{bmatrix}-1/2\\-1/2\\1\end{bmatrix}=\dfrac{1}{\sqrt{6}}\begin{bmatrix}-1\\-1\\2\end{bmatrix}$

Now we can diagonalize S:

$$\begin{bmatrix}-1&1&1\\1&-1&1\\1&1&-1\end{bmatrix}=\underbrace{\begin{array}{c}\begin{bmatrix}1/\sqrt{3}&-1/\sqrt{2}&-1/\sqrt{6}\\1/\sqrt{3}&1/\sqrt{2}&-1/\sqrt{6}\\1/\sqrt{3}&0&2/\sqrt{6}\end{bmatrix}\\ \vec{x}_1\quad\ \vec{x}_2\quad\ \vec{x}_3\end{array}}_{Q}\underbrace{\begin{bmatrix}1&0&0\\0&-2&0\\0&0&-2\end{bmatrix}}_{\Lambda}\underbrace{\begin{bmatrix}1/\sqrt{3}&1/\sqrt{3}&1/\sqrt{3}\\-1/\sqrt{2}&1/\sqrt{2}&0\\-1/\sqrt{6}&-1/\sqrt{6}&2/\sqrt{6}\end{bmatrix}}_{Q^T=Q^{-1}}$$

Some <u>special</u> symmetric matrices: Suppose A is an $m \times n$ matrix (we don't need $m=n$).

Then $\underset{n\times n}{A^TA}$ and $\underset{m\times m}{AA^T}$ are both symmetric (not usually the same)

What's special about A^TA? For one thing, its eigenvalues are not just <u>real</u> numbers. They are also <u>positive</u> (or 0).

Why? Suppose $A^TA\vec{x} = \lambda\vec{x}$ with $\vec{x} \neq \vec{0}$.

Then $$\underset{(A\vec{x})^T Ax}{\vec{x}^T A^T A \vec{x}} = \underset{\lambda(\vec{x}^T\vec{x})}{\lambda(\vec{x}^T\vec{x})} \rightarrow \lambda = \frac{(A\vec{x})\cdot(A\vec{x})}{\vec{x}\cdot\vec{x}} = \frac{\|A\vec{x}\|^2}{\|\vec{x}\|^2} \geq 0$$

Let's arrange the eigenvalues in decreasing order:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = 0, \dots, \lambda_n = 0$$

$$\vec{v}_1, \vec{v}_2, \dots \vec{v}_r, \underbrace{\vec{v}_{r+1}, \dots, \vec{v}_n}_{\text{Basis for } N(A^TA)} \leftarrow \text{Orthonormal basis of eigenvectors}$$

Since each $\lambda_i \geq 0$, we can take their square roots: $\sigma_i = +\sqrt{\lambda_i}$

called the <u>singular values</u> of A

Also a basis for $N(A)$, because $N(A) = N(A^TA)$.

Why? If $\vec{x}$ in $N(A)$, the $A\vec{x} = \vec{0}$, so

$A^TA\vec{x} = A^T\vec{0} = \vec{0} \rightarrow \vec{x}$ is <u>also</u> in $N(A^TA)$.

On the other hand, if $A^TA\vec{x} = \vec{0}$, then

$$\underset{(A\vec{x})^T A\vec{x} = \|A\vec{x}\|^2}{\vec{x}^T A^T A \vec{x}} = \vec{x}^T\vec{0} = 0 \rightarrow A\vec{x} = \vec{0}.$$

So $\vec{x}$ is in $N(A)$ as well.

Note that $\dim N(A^TA) = \dim N(A) = n-r$, where $\underset{\uparrow}{r = \text{rank}(A)}$

same as rank of A^TA.

So the non-zero singular values $\sigma_1, \sigma_2, \ldots, \sigma_r$ go along with rank(A)-many orthonormal eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$: $A^TA\vec{v}_i = \sigma_i^2 \vec{v}_i$

For $i = 1, 2, \ldots, r$, let's define $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$ (vectors in $\mathbb{R}^m$ since A is $m \times n$)

What's special about $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r$?

① They are an orthonormal set in $\mathbb{R}^m$:

$$u_i^T \vec{u}_j = \frac{1}{\sigma_i \sigma_j} (Av_i)^T (A\vec{v}_j) = \frac{1}{\sigma_i \sigma_j} \vec{v}_i^T \underbrace{A^TA\vec{v}_j}_{\sigma_j^2 \vec{v}_j}$$

$$= \frac{\sigma_j}{\sigma_i} v_i^T \vec{v}_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

Because $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is orthonormal.

② They form a orthonormal basis of the column space $C(A)$:

- They are linearly independent because they are orthonormal.
- They are in the column space, because $C(A)$ = set of all $A\vec{x}$ for $\vec{x}$ in $\mathbb{R}^n$.
- They are enough for a basis since $\dim C(A) = \text{rank } r$.

③ They are eigenvectors for AA^T!

same eigenvalue λ_i

$$AA^T\vec{u}_i = \frac{1}{\sigma_i} A\underbrace{A^TA\vec{v}_i}_{\sigma_i^2 \vec{v}_i} = \sigma_i A\vec{v}_i = \sigma_i^2 \left(\frac{A\vec{v}_i}{\sigma_i}\right) = \sigma_i^2 \vec{u}_i$$

Now remember one of the big theorems:

$$C(A)^{\perp} = N(A^T) \text{ in } \mathbb{R}^m$$

We can get an orthonormal basis of $\mathbb{R}^m$ by combining

$$\underbrace{\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}}_{\substack{\text{orthonormal basis} \\ \text{of } C(A)\text{, also} \\ \text{eigenvectors for } AA^T}} \text{ with } \underbrace{\{\vec{u}_{r+1}, \ldots, \vec{u}_m\}}_{\substack{\text{orthonormal basis of } N(A^T)\text{, same as} \\ N(AA^T)\text{, so also eigenvectors for } AA^T \\ \text{(with eigenvalue } 0\text{).}}}$$

Conclusion: For any $m \times n$ matrix A, we've shown that we can find orthonormal bases of both $\mathbb{R}^m$ and $\mathbb{R}^n$ that are "good for A":

$\mathbb{R}^m$: $\{\underbrace{\vec{u}_1, \ldots, \vec{u}_r}_{}, \underbrace{\vec{u}_{r+1}, \ldots, \vec{u}_m}_{\text{basis of } N(A^T)}\}$ orthonormal basis of eigenvectors for AA^T

$$AA^T \vec{u}_i = \sigma_i^2 \vec{u}_i$$

$\mathbb{R}^n$: $\{\underbrace{\vec{v}_1, \ldots, \vec{v}_r}_{}, \underbrace{\vec{v}_{r+1}, \ldots, \vec{v}_n}_{\text{basis of } N(A)}\}$ orthonormal basis of eigenvectors for $A^T A$

$$A^T A \vec{v}_i = \sigma_i^2 \vec{v}_i$$

Moreover, $\boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$