The double root at $t=1$ was part of the problem.

THM. 2.7.4: If A has eigenvectors $\bar{v}_1,\dots,\bar{v}_k$ with distinct eigenvalues $\lambda_1,\dots,\lambda_k$ (i.e. $\lambda_i \neq \lambda_j$ for $i\neq j$) then $\bar{v}_1,\dots,\bar{v}_k$ are lin. independent.

In particular, if $\chi_A(t) = (t-\lambda_1)\cdots(t-\lambda_n)$ has no repeated roots (i.e. $\lambda_i \neq \lambda_j \; \forall\, 1\le i<j\le n$) then A is diagonalizable, since the λ_i-eigenvectors $\bar{v}_i$ give an eigenbasis for A.

proof: If $\bar{v}_1,\dots,\bar{v}_k$ are lin. dependent, write down a nontrivial lin. dependence $c_1\bar{v}_1 + c_2\bar{v}_2 + \dots + c_k\bar{v}_k \overset{(*)}{=} \bar{0}$ that has the fewest nonzero coefficients, and assume $c_1 \neq 0$ (by re-indexing, if needed). We'll get a contradiction by manufacturing one with fewer nonzero coefficients:

Apply A to $(*)$ to get $c_1A\bar{v}_1 + c_2A\bar{v}_2 + \dots + c_kA\bar{v}_k = \bar{0}$

i.e. $c_1\lambda_1\bar{v}_1 + c_2\lambda_2\bar{v}_2 + \dots + c_k\lambda_k\bar{v}_k = \bar{0}$

and subtract $\lambda_1(*)$ i.e. $c_1\lambda_1\bar{v}_1 + c_2\lambda_1\bar{v}_2 + \dots + c_k\lambda_1\bar{v}_k = \bar{0}$

giving $0 + \underbrace{c_2(\lambda_2-\lambda_1)}_{\neq 0}\bar{v}_2 + \dots + \underbrace{c_k(\lambda_k-\lambda_1)}_{\neq 0,}\bar{v}_k = \bar{0}$

these new coefficients $c_i(\lambda_i-\lambda_1)$ are nonzero $\Leftrightarrow$ the old c_i was nonzero

Fewer total! ∎

11/23/2016 >

REMARK: This already shows most matrices A are diagonalizable, since $\chi_A(t)$ usually has distinct roots. In fact, with a little more work, one can write down a polynomial in the entries (a_{ij}) of A that must vanish to get repeated roots in $\chi_A(t)$.

e.g. for $n=2$ $A=\begin{bmatrix} a & b \\ c & d\end{bmatrix}$ has repeated roots for $\chi_A(t) = \det\begin{bmatrix} t-a & -b \\ -c & t-d\end{bmatrix} = t^2 - \underbrace{(a+d)}_{B}t + \underbrace{ad-bc}_{C}$

$\Leftrightarrow 0 = B^2 - 4C = (a+d)^2 - 4(ad-bc)$

$= a^2 + 2ad + d^2 - 4ad + 4bc$

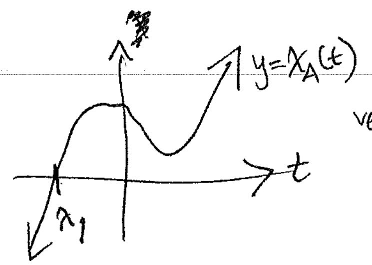
$= (a-d)^2 + 4bc$

(106)

It is also somewhat rare for $n\times n$ matrices A with entries in $\mathbb{R}$ to have all their eigenvalues $\lambda_i \in \mathbb{R}$ rather than $\mathbb{C}$, e.g. for $n=3$, $\chi_A(t) = \det\begin{bmatrix} t-a_{11} & -a_{12} & \cdots \\ -a_{21} & t-a_{22} & \\ \vdots & & t-a_{33}\end{bmatrix} = t^3 - at^2 + bt + c$

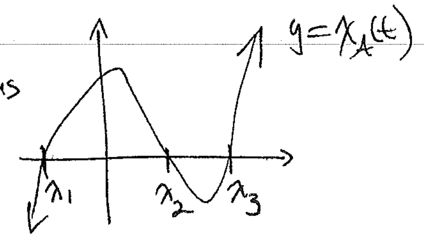
is a cubic, so it has at least one real root λ_1 (with $\mathbb{R}$ coefficients)

but that might be all:



one real eigenvalue λ_1, two complex conjugate $\lambda_2, \lambda_3 = x \pm iy$

versus



3 real eigenvalues

In this way, Symmetric real matrices $A = A^T$ are very special (and extremely important!)

THM 3.7.14 (Spectral theorem for symmetric matrices)

Symmetric real matrices $A = A^T$

(i) have only real eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$

(ii) always have an orthonormal eigenbasis $\bar{v}_1, \ldots, \bar{v}_n$ for $\mathbb{R}^n$

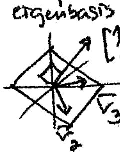
and hence can be diagonalized by an orthogonal matrix $P = \begin{bmatrix} | & & | \\ \bar{v}_1 & \cdots & \bar{v}_n \\ | & & | \end{bmatrix}$,

i.e. $P^TP = I_n$, so $P^T = P^{-1}$ and $P^{-1}AP = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$.

$\quad P^{-1}AP = P^TAP$

EXAMPLE (from section): $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2\end{bmatrix}$ has eigenbasis for $\mathbb{R}^3$

$\begin{bmatrix}1\\1\\1\end{bmatrix} = \bar{v}_1$ spans the 4-eigenspace



$\bar{v}_2, \bar{v}_3$ any orthonormal basis for $\begin{bmatrix}1\\1\\1\end{bmatrix}^\perp$ = the 1-eigenspace

proof: First prove (i), by considering an eigenvalue $\lambda \in \mathbb{C}$ for A, that is, some root λ of $\chi_A(t) = \det(tI - A)$,

and any associated eigenvector $\vec{v} \neq \vec{0}$ in $\ker(\lambda I - A) \subset \mathbb{C}^n$,

so $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} x_1 + iy_1 \\ \vdots \\ x_n + iy_n \end{bmatrix}$.

We compute in 2 ways the scalar

2nd way: complex conjugation! $[\bar{v}_1 \cdots \bar{v}_n] A \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = [\bar{v}_1 \cdots \bar{v}_n] \bar{A}^T \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \left(\begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix}^T \bar{A}^T\right)\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \left(\bar{A}\begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix}\right)^T \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \left(\bar{\lambda}\begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix}\right)^T \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$

$A^T = A$, and $\bar{A} = A$ since A is real.

$= \bar{\lambda} \sum_{i=1}^n \bar{v}_i v_i$

$= \bar{\lambda} \sum_{i=1}^n (x_i^2 + y_i^2)$

1st way: $[\bar{v}_1 \cdots \bar{v}_n] A \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = [\bar{v}_1 \cdots \bar{v}_n] \cdot \lambda \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \lambda [\bar{v}_1 \cdots \bar{v}_n] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \lambda \sum_{i=1}^n (x_i^2 + y_i^2)$

Hence $\bar{\lambda} = \lambda$, i.e. $\lambda \in \mathbb{R}$.

(107)

For (ii), we'll prove it by induction on n.

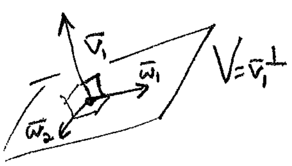
Start with the existence of an eigenvalue $\lambda \in \mathbb{C}$ as a root of $\chi_A(t)$, which we just showed has $\lambda \in \mathbb{R}$. Note that this means $\lambda I_n - A$ also has $\mathbb{R}$ entries, so picking an eigenvector $\bar{v} \in \ker(\lambda I_n - A) - \{\bar{0}\}$ can be done with $\bar{v} \in \mathbb{R}^n$ (don't need $\mathbb{C}^n$!).

We can also replace $\bar{v}$ with $\bar{v}_1 := \frac{\bar{v}}{\|\bar{v}\|}$ of unit length, and call $\lambda_1 := \lambda$.

To get the rest of the orthonormal eigenbasis $\bar{v}_2, \ldots, \bar{v}_n$, we will work inside $V := \bar{v}_1^{\perp} = \{\bar{x} \in \mathbb{R}^n : \bar{x} \cdot \bar{v}_1 = 0\}$.

First note that A restricts to a linear map $V \xrightarrow{A} V$ since if $\bar{x} \in \bar{v}_1^{\perp}$ then

$$(A\bar{x}) \cdot \bar{v}_1 = (A\bar{x})^T \bar{v}_1 = \bar{x}^T A^T \bar{v}_1 = \bar{x}^T A \bar{v}_1$$
$$= \bar{x}^T(\lambda \bar{v}_1) = \lambda \bar{x}^T \bar{v}_1$$
$$= \lambda\, \bar{x} \cdot \bar{v}_1$$
$$= \lambda \cdot 0 = 0.$$



Now we can pick <u>any orthonormal</u> basis $\bar{w}_2, \bar{w}_3, \ldots, \bar{w}_n$ for V (pick $\bar{w}_2$ of unit length in V, then $\bar{w}_3$ of unit length in $V \cap \bar{w}_2^{\perp} = \mathrm{span}(\bar{v}_1, \bar{w}_2)^{\perp}$, etc.) and then write down the matrix $B = (b_{ij})_{\substack{i=2,\ldots,n \\ j=2,\ldots,n}}$

expressing $\underset{\text{in basis } (\bar{w}_2,\ldots,\bar{w}_n)}{V} \xrightarrow{A} \underset{\text{basis } (\bar{w}_2,\ldots,\bar{w}_n)}{V}$

as follows: $\underset{\text{for } j=2,\ldots,n}{A\bar{w}_j} = \sum_{i=2}^{n} b_{ij}\, \bar{w}_i$.

We claim that B is again <u>symmetric</u>, i.e. $B^T = B$:

By orthonormality,

$$b_{ij} = (A\bar{w}_j) \cdot \bar{w}_i = (A\bar{w}_j)^T \bar{w}_i = \bar{w}_j^T A^T \bar{w}_i \overset{A^T = A}{=} \bar{w}_j^T A \bar{w}_i$$

same!

$$b_{ji} = (A\bar{w}_i) \cdot \bar{w}_j = \bar{w}_j \cdot (A\bar{w}_i) = \bar{w}_j^T A \bar{w}_i$$

REMARK: We'll use Spectral Thm in §3.5 to analyze quadratic forms. But it's also very useful in defining the <u>singular value decomposition</u> (SVD) of a rectangular matrix

$$\underset{m\times n}{X} = \underset{m\times m}{U} \left[\begin{array}{c|c} \begin{matrix}\sigma_1 & 0 \\ 0 & \sigma_r\end{matrix} & 0 \\ \hline 0 & 0 \end{array}\right] \underset{n\times n}{V^T}$$

U, V <u>orthogonal</u> matrices, $\sigma_i \geq 0$, via $A = X^T X$, $B = XX^T$

Hence by induction on n (since B is $(n-1)\times(n-1)$, symmetric) we can find an orthonormal eigenbasis $\bar{v}_2, \ldots, \bar{v}_n$ for B on V, which gives an orthonormal eigenbasis $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ for A on $\mathbb{R}^n$ ∎