(148) The key tool is that eigenvalues are roots of $p_A(t)$:

<u>DEF'N</u>: Given $\begin{array}{ccc} V & \xrightarrow{T} & V \\ \uparrow & & \uparrow \\ F^n & \xrightarrow{A} & F^n \end{array}$, its <u>characteristic</u> <u>polynomial</u> is

$$p_A(t) := \det(tI_n - A)$$
$$= \det\begin{bmatrix} t-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & t-a_{22} & & \\ \vdots & & \ddots & \\ -a_{n1} & & & t-a_{nn} \end{bmatrix}$$
$$= t^n - \underbrace{(a_{11}+\dots+a_{nn})}_{\text{Trace}(A)} t^{n-1} + \dots \cdots (-1)^n \det A$$

a polynomial in t

12/12/2018

<u>NOTE</u>: $p_A(t)$ depends only on T, since if we change bases, we get $A' = P^{-1}AP$ with $\det(tI_n - P^{-1}AP)$

$$= \det(P^{-1}(tI_n)P - P^{-1}AP)$$
$$= \det(P^{-1}\cdot(tI_n - A)\cdot P)$$
$$= \det(P^{-1})\cdot \det(tI_n - A) \det(P)$$
$$= \det(tI_n - A)$$

PROPOSITION: $\lambda \in F$ is an <u>eigenvalue</u> for T (say with an eigenvector $v \neq 0$) $Tv = \lambda v$ $\iff$ $t = \lambda$ is a <u>root</u> of $p_A(t) = 0$ (and $v \in \ker(\lambda I_n - A)$)

<u>proof</u>: $Tv = \lambda v$ for some $v \neq 0$ in V

$\iff Av = \lambda v$ —"—

$\iff \lambda v - Av = \underline{0}$ —"—

$\iff (\lambda I_n - A)v = \underline{0}$ —"—

$\iff \lambda I_n - A$ is noninvertible in $F^{n\times n}$

$\iff \det(\lambda I_n - A) = 0$ i.e. $t = \lambda$ is a root of $\det(tI - A) =: p_A(t)$ ■

Since $p_A(t) = \det(tI_n - A)$ is a nonconstant polynomial,
$\qquad\qquad = t^n - \mathrm{Tr}(A)t^{n-1} + \dots$

when working with $F = \mathbb{C}$, the Fundamental Theorem of Algebra
(sketch proof in Artin §15.10)
says we will always have at least one eigenvalue λ_1, and an associated eigenvector $v_1 \in \ker(\lambda_1 I_n - A) \neq \{0\}$. This lets us prove...

THEOREM: Working with $F=\mathbb{C}$, Every
(PROP 4.6.1)

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ \uparrow & & \uparrow \\ \mathbb{C}^n & \xrightarrow{A} & \mathbb{C}^n \end{array}$$

can be triangularized, i.e.

$\exists P \in GL_n(\mathbb{C})$ with $P^{-1}AP = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

or equivalently, $\exists$ a basis $\mathcal{B} = (v_1, \dots, v_n)$ for V

with T having matrix $A' = \begin{array}{c|c} & \begin{array}{ccc} v_1 & \cdots & v_n \end{array} \\ \begin{array}{c} v_1 \\ \vdots \\ v_n \end{array} & \begin{bmatrix} \lambda_1 & & * \\ & \lambda_2 & \\ 0 & & \ddots\ \lambda_n \end{bmatrix} \end{array}$

proof: Induct on n, with base case $n=1$ being trivial: $A = [\lambda_1]$

In the inductive step, choose λ_1 a root of $p_A(t) = \det(tI_n - A)$ and $v_1 \in \ker(\lambda_1 I_n - A) - \{0\}$, so $T(v_1) = \lambda_1 v_1$.

Extending v_1 arbitrarily to a basis $(v_1, v_2, \dots, v_n)$ for V, the matrix for T looks like

$$A = \begin{array}{c|c} & \begin{array}{cccc} v_1 & v_2 & & v_n \end{array} \\ \begin{array}{c} v_1 \\ v_2 \\ \vdots \\ v_n \end{array} & \left[\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ 0 & & \hat{A} & \\ \vdots & & & \\ 0 & & & \end{array}\right] \end{array}$$

for some $\hat{A} \in \mathbb{C}^{(n-1)\times(n-1)}$.

By induction, $\exists \hat{P} \in GL_{n-1}(\mathbb{C})$ with $\hat{P}^{-1}\hat{A}\hat{P} = \begin{bmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ for some $\lambda_2, \dots, \lambda_n$

and then $P = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \hat{P} & \\ 0 & & & \end{array}\right]$ has

$$P^{-1}AP = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{P}^{-1} \\ 0 & \end{array}\right] \left[\begin{array}{c|c} \lambda_1 & * \cdots * \\ \hline 0 & \\ \vdots & \hat{A} \\ 0 & \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{P} \\ 0 & \end{array}\right] = \left[\begin{array}{c|c} \lambda_1 & * \cdots * \\ \hline 0 & \\ \vdots & \hat{P}^{-1}\hat{A}\hat{P} \\ 0 & \end{array}\right]$$

$$= \left[\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & \lambda_2 & & * \\ \vdots & & \ddots & \\ 0 & 0 & & \lambda_n \end{array}\right] \blacksquare$$

REMARK: §4.7 proves a much more precise triangular form for $A \in \mathbb{C}^{n\times n}$ called Jordan canonical form with lots more zeroes, and that lets one decide whether $A, A' \in \mathbb{C}^{n\times n}$ have $P^{-1}AP = A'$ for some $P \in GL_n(\mathbb{C})$ or not,
i.e. it's true if and only if A, A' have same Jordan form.

Why should this hold?:

THEOREM (4.6.6): If $p_A(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$ has distinct roots $\lambda_1, \ldots, \lambda_n \in F$ then A is diagonalizable in $F^{n\times n}$.

It's an immediate consequence of this fact:

PROPOSITION (4.6.5): If $v_1, v_2, \ldots, v_r$ are eigenvectors for $V \xrightarrow{T} V$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ (i.e. $\lambda_i \neq \lambda_j$) then they are linearly independent.

Proof: Assume not, and let $c_1 v_1 + c_2 v_2 + \ldots + c_r v_r = \underline{0}$ be some nontrivial linear dependence having the smallest number of nonzero coefficients c_i; assume $c_1 \neq 0$ by re-indexing.

We'll get a contradiction, by applying T:

$$T(c_1 v_1 + c_2 v_2 + \ldots + c_r v_r) = T(\underline{0}) = \underline{0}$$

i.e. $c_1 T(v_1) + c_2 T(v_2) + \ldots + c_r T(v_r) = \underline{0}$

i.e. $\lambda_1 c_1 v_1 + \lambda_2 c_2 v_2 + \ldots + \lambda_r c_r v_r = \underline{0}$

Subtract λ_1 times the original dependence $\lambda_1\left(\sum_{i=1}^{r} c_i v_i\right) = \underline{0}$, giving

$$(\lambda_1 c_1 - \lambda_1 c_1)v_1 + (\lambda_1 c_2 - \lambda_2 c_2)v_2 + \ldots + (\lambda_1 c_r - \lambda_r c_r)v_r = \underline{0}$$

i.e.
$$\underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} c_2 v_2 + \ldots + \underbrace{(\lambda_1 - \lambda_r)}_{\neq 0} c_r v_r = \underline{0}$$

This is a dependence with one fewer nonzero coefficient.
Contradiction ■

What about our last goal theorem? Not so hard, and uses similar ideas to the triangularization result ...

Spectral THEOREM: for symmetric matrices Any symmetric matrix $A^T = A \in \mathbb{R}^{n\times n}$ is diagonalizable, via an <u>orthogonal</u> change-of-basis matrix $P = (P^T)^{-1}$, and has only <u>real</u> eigenvalues,

i.e. $$\underset{\displaystyle P^TAP}{\overset{\|}{P^{-1}AP}} = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \; \lambda_n \end{bmatrix}, \quad \underline{\underline{\lambda_i \in \mathbb{R}}}$$

<u>proof:</u> Induct on n, with base case $n=1$ trivial: $A = [\lambda_1]$.

In the inductive step, regard $A \in \mathbb{C}^{n\times n}$ and pick a root $\lambda_1 \in \mathbb{C}$ for $p_A(t) = \det(tI_n - A)$, and $v_1 = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$ an associated nonzero eigenvector $v \in \ker(\lambda_1 I_n - A)$, so $Av_1 = \lambda_1 v_1$. We'll first show $\underline{\lambda_1 \in \mathbb{R}}$.

Note that $\bar{v}_1 := \begin{bmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_n \end{bmatrix} \in \mathbb{C}^n$ has $\bar{v}_1^T v_1 = v_1 \bar{v}_1^T = \sum_{i=1}^n z_i \bar{z}_i = \sum_{i=1}^n \|z_i\|^2 \underset{(\neq 0)}{>} 0$ since $\underline{v_1 \neq 0}$

Now compute in two ways

$$\bar{v}_1^T A v_1 = \bar{v}_1^T(\lambda v_1) = \lambda_1 \underbrace{(\bar{v}_1^T v_1)}_{\neq 0 \text{ in } \mathbb{R}}$$

since $\bar{A} = A \in \mathbb{R}^{n\times n}$ and $A^T = A$:

$$\bar{v}_1^T A v_1 = (\bar{v}_1^T \bar{A}^T) v_1 = (\bar{A}\bar{v}_1)^T v_1 = (\bar{\lambda} \bar{v}_1)^T v_1 = \bar{\lambda}_1 \underbrace{\bar{v}_1^T v_1}_{\neq 0 \text{ in } \mathbb{R}}$$

comparing these gives $\bar{\lambda} = \lambda$, i.e. $\underline{\underline{\lambda \in \mathbb{R}}}$

Now since $\lambda_1 \in \mathbb{R}$, one can pick $v_1 \in \ker\underbrace{(\lambda_1 I_n - A)}_{\in \mathbb{R}^{n\times n}} - \{0\}$ to have $v_1 \in \mathbb{R}^n - \{0\}$ by row-reduction in $\mathbb{R}$

New consider the subspace $\mathcal{U} := v_1^{\perp} := \{ v \in \mathbb{R}^n : v_1^T v = 0 \} \subset \mathbb{R}^n$,
and we claim $A(\mathcal{U}) \subset \mathcal{U}$: given $v \in \mathcal{U}$ so that $v_1^T v = 0$

then $v_1^T (Av) = (v_1^T A^T) v = (A v_1)^T v = (\lambda v_1)^T v = \lambda \cdot v_1^T v = \lambda \cdot 0 = 0$

It's easy to see that any $\mathbb{R}$-basis $v_2, \dots, v_n$ for $\mathcal{U}$ gives a basis $(v_1, v_2, \dots, v_n)$ for $\mathbb{R}^n$, and then $A(\mathcal{U}) \subset \mathcal{U}$ means that A will have matrix in this basis of the form

$$A = \begin{array}{cc} & \begin{array}{cc} v_1 & v_2 \cdots v_n \end{array} \\ \begin{array}{c} v_1 \\ v_2 \\ \vdots \\ v_n \end{array} & \left[\begin{array}{c|c} \lambda_1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{A} \\ 0 & \end{array} \right] \end{array}$$

for some $\hat{A} \in \mathbb{R}^{(n-1)\times(n-1)}$

We claim that if one choose $v_2, \dots, v_n$ to be any <u>orthonormal basis</u> for $\mathcal{U}$, meaning $v_i^T v_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$, then $\hat{A}^T = \hat{A}$:

one has $\hat{A} = (\hat{a}_{ij})_{\substack{i=2,\dots,n \\ j=2,\dots,n}}$ where $\underset{\text{for } j=2,\dots,n}{A v_j} = \sum_{i=2}^{n} \hat{a}_{ij} v_i$

$$\Downarrow$$

$$v_k^T A v_j = \sum_{i=2}^{n} \hat{a}_{ij} \underbrace{v_k^T v_i}_{=\{1 \text{ if } i=k;\ 0 \text{ else}\}} = \hat{a}_{kj}$$

But $\underbrace{v_k^T A v_j}_{=\hat{a}_{kj}} \underset{A^T = A}{=} v_k^T A^T v_j = (A v_k)^T v_j = v_j^T A v_k = \hat{a}_{jk}$, i.e. $\hat{A}^T = \hat{A}$.

Hence induction applies to $\hat{A}$, so $\exists$ an <u>orthogonal</u> $\hat{P} = (\hat{P}^T)^{-1} \in GL_{n-1}(\mathbb{R})$

with $\hat{P}^{-1} \hat{A} \hat{P} = \begin{bmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ <u>and $\lambda_2, \dots, \lambda_n \in \mathbb{R}$</u>.

so $\underbrace{P = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{P} \\ 0 & \end{array} \right]}_{\text{also } \underline{\text{orthogonal}} \text{ in } GL_n(\mathbb{R})}$ has $P^{-1} A P = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{P}^{-1} \\ 0 & \end{array} \right] \left[\begin{array}{c|c} \lambda_1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{A} \\ 0 & \end{array} \right] \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{P} \\ 0 & \end{array} \right] = \left[\begin{array}{c|c} \lambda_1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & \hat{P}^{-1} \hat{A} \hat{P} \\ 0 & \end{array} \right] = \left[\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \lambda_n \end{array} \right]$

∎