Thm 5.9 Let $T \in \mathcal{L}(V)$, $\dim(V) < \infty$, and assume that the char. poly. of T splits.
Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T. Then
a) T is diagonalizable $\iff$ the multiplicity of λ_i is equal to $\dim(E_{\lambda_i})$ for all i.
b) If T is diagonalizable and β_i is an ordered basis for E_{λ_i} for each i, then
$\beta = \beta_1 \cup \dots \cup \beta_k$ is an ordered basis for V consisting of e.vects of T.

Proof.
For each i, let m_i denote the multiplicity of λ_i, $d_i = \dim(E_{\lambda_i})$, and $n = \dim(V)$.

$\Rightarrow$ Suppose that T is diagonalizable.
Let β be a basis for V consisting of e.vects of T.
For each i, let $\beta_i = \beta \cap E_{\lambda_i}$.
Let $n_i = |\beta_i|$.
Then:
- $n_i \le d_i$ for each i (because β_i is a lin. indep. subset of the subspace E_{λ_i} and $\dim(E_{\lambda_i}) = d_i$).
- $d_i \le m_i$ (by Thm 5.7).
- $\sum_{i=1}^{k} n_i = n$ (because β contains n vectors).
- $\sum_{i=1}^{k} m_i = n$ (because the degree of the char. poly. of T is equal to the sum of the mult. of the eigenvalues, on the one hand, and is equal to $\dim(V) = n$ on the other hand).

Thus:
$$n = \sum_{i=1}^{k} n_i \le \sum_{i=1}^{k} d_i \le \sum_{i=1}^{k} m_i = n.$$
It follows that
$$\sum_{i=1}^{k} (m_i - d_i) = 0.$$

Since $(m_i - d_i) \ge 0$ for all i, we conclude that $m_i = d_i$ for all i.

$\Leftarrow$ Conversely, suppose that $m_i = d_i$ for all i.
For each i, let β_i be an ordered basis for E_{λ_i}, and let $\beta = \beta_1 \cup \dots \cup \beta_k$.
By Thm 5.8, β is lin. indep.
Furthermore, since $d_i = m_i$ for all i by assumption, β contains
$$\sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i = n \quad \text{vectors.}$$

Therefore β is an ordered basis for V consisting of e.vects of V. Hence T is diagz.

This theorem concludes our study of the diagonalization problem. Let's summarize.

Test for diagonalization
Let T be a lin. operator on an n-dim. v.s. V.
Then T is diagonalizable if and only if both of the following conditions hold.
1) The char. polynomial of T splits.
2) For each e.val λ of T, the multiplicity of λ equals $\dim E_\lambda = \dim N(T - \lambda I_V) =$
$= n - \operatorname{rank}(T - \lambda I_V)$.

The same conditions can be used to test if a square matrix is diagz, because A is diagz $\iff$ the operator L_A is diagz.

Example
Let $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \in M_{3\times3}(\mathbb{R})$, and we test its diagonalizability.
The char. poly. $f(t) = \det(A - tI_3) = \det \begin{pmatrix} 3-t & 1 & 0 \\ 0 & 3-t & 0 \\ 0 & 0 & 4-t \end{pmatrix} = (4-t)(3-t)^2 = -(t-4)(t-3)^2$.
This shows that $f(t)$ splits, so condition (1) for diagz. holds.

E. vals:

$\lambda_1 = 4$ — mult. 1

$\lambda_2 = 3$ — mult. 2

Condition (2) is automatically satisfied for λ_1 (as by Thm 5.7, $1 \le \dim(E_{\lambda_1}) \le \text{mult}\,\lambda_1 = 1$)

So only need to check (2) for λ_2.

The matrix

$A - \lambda_2 I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ has rank 2 (the rank of $L_{A-\lambda_2 I_3}$, equivalently the max. number of lin. indep. columns).

$\dim E_{\lambda_2} = 3 - \text{rank}(A - \lambda_2 I_3) = 3 - 2 = 1 \neq 2$, the mult. of λ_2.

Hence A is not diagz.

Example.

Let $A = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$.

$f(t) = \det(A - t I_2) = (t-1)(t-2)$.

Hence $\lambda_1 = 1$, $\lambda_2 = 2$ are the e. vals, both of mult. 1. Thus both conditions (1), (2) are satisfied and T is diagz.

$E_{\lambda_1} = N(L_A - 1 \cdot I_2) = \langle \begin{pmatrix} -2 \\ 1 \end{pmatrix} \rangle$.

$E_{\lambda_2} = N(L_A - 2 \cdot I_2) = \langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rangle$.

Hence $\beta_1 = \{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \}$ is a basis for E_{λ_1}, and $\beta_2 = \{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \}$ is a basis for E_{λ_2}.

By the theorem $\beta = \beta_1 \cup \beta_2 = \{ \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \}$ is a basis for $V = \mathbb{R}^2$ consisting of e. vects. So $[L_A]_\beta$ is a diag. matrix.

Let $Q = \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix}$.

Then $Q^{-1} A Q = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ — diagonal.

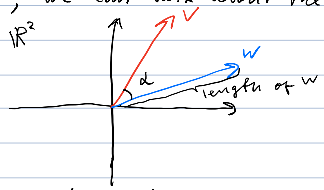
Fact. Let $A \in M_{n \times n}(F)$, let $\gamma = \{v_1, \dots, v_n\}$ be an ord. basis for F^n. Then

$[L_A]_\gamma = Q^{-1} A Q$, where

$$Q = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

Inner products and norms

- In $V = \mathbb{R}^2$, we can talk about the length of a vector, the angle between two vectors, two vectors being orthogonal, etc.



- In a general v.s. V, these notions are not defined. For example, if $V = P_2(\mathbb{R})$, what is the length of a polynomial $3x^2 + 2x - 1$?
- In order to study these notions in general, we introduce an "upgraded" version of vector spaces.

Def

Let V be a v.s. over F (for $F = \mathbb{R}$ or $F = \mathbb{C}$).

An inner product on V is a function that assigns, to every ordered pair of vectors x and y in V, a scalar in F, denoted by $\langle x, y \rangle$, such that for all $x, y, z \in V$ and $c \in F$ the following holds:

a) $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$ } – linearity in the first variable
b) $\langle cx, y \rangle = c \langle x, y \rangle$
c) $\overline{\langle x, y \rangle} = \langle y, x \rangle$ (where "$\overline{}$" denotes complex conjugation). – conjugate symmetry
d) $\langle x, x \rangle > 0$ for $x \neq 0$ – positivity

Remark 1) If $F = \mathbb{R}$, then (c) reduces to $\langle x,y\rangle = \langle y,x\rangle$.
2) It follows from the definition that if $a_1, \dots, a_n \in F$ and $y, v_1, \dots, v_n \in V$, then
$$\left\langle \sum_{i=1}^{n} a_i v_i \, , \, y \right\rangle = \sum_{i=1}^{n} a_i \langle v_i, y\rangle.$$

Example. We define the standard inner product on F^n:
For $x = (a_1, \dots, a_n)$, $y = (b_1, \dots, b_n)$ in F^n, define
$$\langle x,y\rangle = \sum_{i=1}^{n} a_i \overline{b_i}.$$
We can verify that $\langle \cdot, \cdot \rangle$ satisfies the conditions (a) through (d).
For example, if $z = (c_1, \dots, c_n)$, we have for (a)
$$\langle x+z, y\rangle = \sum_{i=1}^{n} (a_i + c_i)\overline{b_i} = \sum_{i=1}^{n} a_i \overline{b_i} + \sum_{i=1}^{n} c_i \overline{b_i} = \langle x,y\rangle + \langle z,y\rangle.$$
For example, for $x = (1+i, 4)$ and $y = (2-3i, 4+5i)$ in $\mathbb{C}^2$, $\langle x,y\rangle = (1+i)(2+3i) + 4\cdot(4-5i) = 15-15i$.

When $F = \mathbb{R}$ the conjugations are not needed, and $\langle x,y\rangle$ gives the dot product from 33A.

Example If $\langle x,y\rangle$ is any inner product on a v.s. V and $r > 0$, we may define another inner product by the rule $\langle x,y\rangle' = r\langle x,y\rangle$. (If $r \le 0$, then (d) would not hold.)

Example
Let $V = C(\mathbb{R})$, the v.s. of real-valued continuous functions on $\mathbb{R}$.
For $f, g \in V$, define
$$\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt.$$

(a) and (b) hold by the basic properties of integration, for example for (a) we have
$\langle f_1 + f_2, g\rangle = \int_0^1 (f_1(t) + f_2(t))\, g(t)\, dt = \int_0^1 f_1(t)g(t)\,dt + \int_0^1 f_2(t)g(t)\,dt = \langle f_1, g\rangle + \langle f_2, g\rangle$.
(c) is clear, and (d) is easy to verify as $\int_0^1 (f(t))^2\,dt > 0$ for any continuous $f \neq 0$.
Thus, $\langle \cdot, \cdot \rangle$ is an inner product on $C(\mathbb{R})$.
Note that similarly, $\langle f,g\rangle' = \int_{-1}^{1} f(t)g(t)\,dt$ gives another inner product on $C(\mathbb{R})$.

Example
Let $A \in M_{m\times n}(F)$. We define the conjugate transpose of A as the $n\times m$ matrix A^* s.t. $(A^*)_{ij} = \overline{A_{ji}}$.
When $F = \mathbb{R}$, then A^* is simply A^t.
For example, if $A = \begin{pmatrix} i & 1+2i \\ 2 & 3+4i \end{pmatrix}$, then $A^* = \begin{pmatrix} -i & 2 \\ 1-2i & 3-4i \end{pmatrix}$.

Consider now $V = M_{n\times n}(F)$, and define $\langle A,B\rangle = tr(B^*A)$ for $A, B \in V$.
This defines an inner product on V, called the Frobenius inner product.
(see page 331, Example 5 for a proof that (a)–(d) hold)

Def A v.s. V over F endowed with a specific inner product is called an inner product space.
If $F = \mathbb{C}$, V is called a complex inner product space.
If $F = \mathbb{R}$, V is called a real inner product space.

Remark. 1) If a v.s. V has an inner product $\langle x,y\rangle$ and W is a subspace of V, then W is also an inner product space when the same function $\langle x,y\rangle$ is restricted to the vectors $x,y \in W$.
As $P_n(\mathbb{R})$ is a subspace of $C(\mathbb{R})$, it follows that $P_n(\mathbb{R})$ can be equipped with (many different inner products.

Thm 6.1 (basic properties of inner products).

Let V be an inner product space. Then for any $x, y, z \in V$ and $c \in F$ we have

a) $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle$

b) $\langle x, cy \rangle = \bar{c} \langle x, y \rangle$

c) $\langle x, 0 \rangle = \langle 0, x \rangle = 0$

d) $\langle x, x \rangle = 0 \iff x = 0$

e) If $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in V$, then $y = z$.

Proof.

(a) $\langle x, y+z \rangle = \overline{\langle y+z, x \rangle} = \overline{\langle y, x \rangle + \langle z, x \rangle} = \overline{\langle y, x \rangle} + \overline{\langle z, x \rangle} = \langle x, y \rangle + \langle x, z \rangle$.

(b) – (e). Exercize.