Linear Algebra Notes

Mario L. Gutierrez Abed

Diagonalization

EIGENVALUES AND EIGENVECTORS

Let $T \in \mathcal{L}(V^n)$. Then we want to know if there exist a basis β for V such that $[T]_\beta$ is a diagonal matrix. If there exists such a basis, how do we find it?

a) Let $T \in \mathcal{L}(V^n)$. Then T is said to be diagonalizable if $\exists$ a basis β for V such that $[T]_\beta$ is a diagonal matrix.
b) Let $A \in M_{n\times n}(\mathbb{F})$. Then A is diagonalizable if L_A is diagonalizable.
Alternatively (from theorem 2.3), we can say that a square matrix A is diagonalizable if $\exists$ an invertible Q such that $A = Q^{-1}\, D\, Q$, where D is a diagonal matrix, i.e. A is similar to a diagonal matrix.

Motivations for eigenvalues and eigenvectors:

1) Let $\beta = \{v_1, \ \ldots, \ v_n\}$ be a basis for V such that $T \in \mathcal{L}(V)$ is diagonalizable with respect to β, i.e.

$$[T]_\beta = D = \begin{pmatrix} \alpha_{11} & 0 & 0 & 0 & 0 \\ 0 & \alpha_{22} & 0 & 0 & 0 \\ 0 & 0 & . & 0 & 0 \\ 0 & 0 & 0 & . & 0 \\ 0 & 0 & 0 & 0 & \alpha_{\text{nn}} \end{pmatrix}$$

Then we have

$$\begin{aligned} T\, v_1 &= \alpha_{11}\, v_1 + 0 + 0 + \ldots + 0 \\ T\, v_2 &= 0 + \alpha_{22}\, v_2 + 0 + \ldots + 0 \\ &\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\ T\, v_n &= 0 + 0 + \ldots + 0 + \alpha_{\text{nn}}\, v_n \end{aligned}$$

This implies that $T\, v_j = \alpha_{\text{jj}}\, v_j$ for $1 \le j \le n$.

2) Recall T-invariant subspaces of V (a subspace $W \subseteq V$ is said to be T-invariant if $T(w) \subseteq W \ \ \forall \ w \in W$)

We use the same basis $\beta = \{v_1, \ ..., v_n\}$. Then we look at the following nontrivial T-invariant subspaces of dimension 1:

$w_1 = \text{span}(v_1)$
$w_2 = \text{span}(v_2)$
… … … … … ..
$w_n = \text{span}(v_n)$

Thus for $w_j = \text{span}(v_j)$ we have $T\,v_j = \alpha_j\,v_j$ (same equation that we arrived to on part 1).

Definition:
i) Let $T \in \mathcal{L}(V)$. A nonzero vector $v \in V$ is called an eigenvector of T if $\exists$ a scalar λ such that $T\,v = \lambda\,v$. Then λ is an eigenvalue of T corresponding to v.
ii) Let $A \in M_{n\times n}(\mathbb{F})$. Then, a nonzero vector $v \in \mathbb{F}^n$ is an eigenvector of L_A, i.e. $A\,v = \lambda\,v$, and λ is the eigenvalue of A corresponding to v.

• Theorem:
Let $T \in \mathcal{L}(V)$. Then T is diagonalizable iff $\exists$ an ordered basis β of Vconsisting of only eigenvectors (β is often called an eigenbasis).

Note: To diagonalize an operator $T \in \mathcal{L}(V)$, we must find an eigenbasis for V. In order to achieve this, we first compute the eigenvalues.

• Theorem:
Let $A \in M_{n\times n}(\mathbb{F})$. Then $\lambda \in \mathbb{F}$ is an eigenvalue of A iff $\det(A - \lambda\,I_n) = 0$.

Proof:
(($\Rightarrow$) and ($\Leftarrow$))
Suppose λ is an eigenvalue of A. Then $\exists$ a nonzero vector $v \in V$ such that $A\,v = \lambda\,v$.
$\Longleftrightarrow A\,v - \lambda\,v = 0$
$\Longleftrightarrow A\,v - \lambda\,I_n\,v = 0$
$\Longleftrightarrow (A - \lambda\,I_n)\,v = 0$
$\Longleftrightarrow L_A - \lambda\,I_n$ is not injective (since v is assumed to be nonzero)
$\Longleftrightarrow A - \lambda\,I_n$ is not invertible
$\Longleftrightarrow \det(A - \lambda\,I_n) = 0$ ■

Definition: Let $A \in M_{n\times n}(\mathbb{F})$. The polynomial $\text{char}(A) = f(t) = \det(A - t\,I_n)$ is called the characteristic polynomial of A. Thus, the eigenvalues of A are the zeroes/roots of char(A).

Example:

Find the eigenvalues of $A = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$.

Solution:

char $(A) = \det(A - t\, I_n)$

Now solving for the determinant and setting it equal to 0 we get the eigenvalues :

$$\det\left[\begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}\right] = \det\begin{pmatrix} 3-t & 2 \\ -1 & -t \end{pmatrix} = 0$$

$\Longrightarrow (3-t)(-t) + 2 = t^2 - 3\,t + 2 = 0$

$\Longrightarrow t = 1,\ 2 \Longrightarrow \lambda = 1,\ 2$ (eigenvalues)

Side Note: Remember the following algebraic identity:

$$x^n - c^n = (x - c)\left(x^{n-1} + c\, x^{n-2} + c^2\, x^{n-3} + \ldots + c^{n-1}\right).$$

• Theorem:

Let $T \in \mathcal{L}(V^n)$, with $\lambda_1, \ \ldots, \lambda_m$ distinct eigenvalues of T and $v_1, \ \ldots, v_m$ the corresponding eigenvectors of T, with $m \leq n$. Then, $\{v_1, \ \ldots, v_m\}$ is linearly independent.

** Proof on Axler pg 79 and also on Friedberg pg 261 **

• Corollary:

If $m = n$, then $\{v_1, \ \ldots, v_m\} = \{v_1, \ \ldots, v_n\}$ is an eigenbasis, and thus T is diagonalizable. (**Note: The converse of this corollary is not true, i.e. just because T is diagonalizable, it doesn't guarantee that $\{v_1, \ \ldots, v_n\}$ is an eigenbasis. **)

** Proof on Friedberg pg 261-262 **

• Theorem:

Let $A \in M_{n\times n}(\mathbb{F})$. Then

a) char (A) is a polynomial of degree n with leading coefficients $= (-1)^n$.

b) A has at most n eigenvalues.

• Theorem:

Let $T \in \mathcal{L}(V)$, and λ an eigenvalue of T. Then $v \in V$ is an eigenvector corresponding to λ iff $v \in \mathcal{N}(T - \lambda\, I)$ and $v \neq 0$.

Proof:

((⇒) and (⇐))

Suppose v is an eigenvector of T. Then by definition we have that $v \neq 0$ and

$$T\, v = \lambda\, v \Longleftrightarrow T\, v - \lambda\, v = 0 \Longleftrightarrow (T - \lambda\, I)\, v = 0 \Longleftrightarrow v \in \mathcal{N}(T - \lambda\, I)\,. \qquad \blacksquare$$

Example:

Find the eigenvalues and associated eigenvectors of $T : P_2(\mathbb{R}) \longrightarrow P_2(\mathbb{R})$ defined by $f(x) \longmapsto f(x) + (x+1)\, f'(x)$.

Solution:

Let β be the standard basis for $P_2(\mathbb{R})$, i.e. $\beta = \{1,\ x,\ x^2\}$. Then let's compute $[T]_\beta$:

$\hookrightarrow T(1) = 1 + (x+1) \cdot 0 = 1 = \ 1\,(1) + 0\,(x) + 0\,(x^2)$

$\hookrightarrow T(x) = x + (x+1) \cdot 1 = 2\,x + 1 = \ 1\,(1) + 2\,(x) + 0\,(x^2)$

$\hookrightarrow T(x^2) = x^2 + (x+1) \cdot 2\,x = 3\,x^2 + 2\,x = \ 0\,(1) + 2\,(x) + 3\,(x^2)$

Hence $[T]_\beta = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$.

We let $[T]_\beta = A$. Then we have $\mathrm{char}([T]_\beta) = \mathrm{char}(A)$.

Now let's solve for the eigenvalues :

$$\mathrm{char}(A) = \det(A - t\, I_n) = \det\begin{pmatrix} 1-t & 1 & 0 \\ 0 & 2-t & 2 \\ 0 & 0 & 3-t \end{pmatrix} = 0$$

$\mathrm{char}(A) = (1-t)\,(2-t)\,(3-t) = 0$.

Solving for t we have :

$t = 1,\ 2,\ 3 \Longrightarrow \lambda = 1,\ 2,\ 3.$ (eigenvalues of T) ✓

Now we want to find the eigenvectors associated with each eigenvalue :

$\hookrightarrow$ For $\lambda = 1$:

We're looking for vectors v such that $(T - t\, I_n)\, v = 0 \ \cong \ (A - \lambda\, I_n)\, \mathcal{M}_\beta(v) = 0$.

$$A - (1)\, I_n = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

Now we solve

$$(A-(1)\,I_n)\,x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \longrightarrow x_2,\ x_3 = 0 \ \text{ and } \ x_1 \ \text{ is a free variable}$$

Then the set of all solutions is

$\left\{\gamma\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} : \gamma \in \mathbb{R}\right\}$. Hence $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ is an eigenvector of A associated with the eigenvalue $\lambda = 1$.

Now since $\exists$ an isomorphism $\mathcal{M}_\beta : P_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$, we can compute the inverse

$\mathcal{M}_\beta^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1 + 0\,x + 0\,x^2 = \underset{\text{Eigenvector of } T}{\underline{1}}$.

$\bullet\!\!\rightarrow$ For $\lambda = 2$:

$$A-(2)\,I_n = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

Then

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Longrightarrow x_1,\ x_2 \text{ are free variables and } x_3 = 0\ .$$

Thus the solution set is $\left\{\gamma\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} : \gamma \in \mathbb{R}\right\}$. Hence $x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ is an eigenvector of A associated with $\lambda = 2$,

and by the isomorphism discussed above we have that $1 + x$ is an eigenvector of T .

$\bullet\!\!\rightarrow$ For $\lambda = 3$:

$$A-(3)\,I_n = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Then

$$\begin{pmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow \quad x_3 \ = \tfrac{1}{2}\,x_2 \ \text{ and } \ x_2 = 2\,x_1$$

Thus the solution set is $\left\{\gamma\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} : \gamma \in \mathbb{R}\right\}$. Hence

$x = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ is an eigenvector of $A \Longrightarrow 1 + 2x + x^2$ is an eigenvector of T.

Thus T has three eigenvalues and three basis eigenvectors associated to these eigenvalues. ✓

** Recalling lemma previously stated **
Let $T \in \mathcal{L}(V)$, with $\lambda_1, \ldots, \lambda_m$ distinct eigenvalues of T and $v_1, \ldots, v_m$ the corresponding eigenvectors of T. Then, $\{v_1, \ldots, v_m\}$ is linearly independent.

By the lemma above we have that $\beta_E = \{1, 1 + x, 1 + 2x + x^2\}$ is linearly independent; thus this is an eigenbasis for T and this in turn means that T is diagonalizable.

Then we have that $[T]_{\beta_E}$, which is the matrix representation of T with respect to the eigenbasis, is calculated as follows:

- $T(1) = 1$
- $T(1 + x) = 2x + 2 = 0(1) + 2(1 + x) + 0(1 + 2x + x^2)$
- $T(1 + 2x + x^2) = (1 + 2x + x^2) + (x + 1)(2 + 2x)$
  $= (x + 1)^2 + 2(x + 1)^2$
  $= 3(x + 1)^2 = 0(1) + 0(1 + x) + 3(1 + 2x + x^2)$

Hence,

$[T]_{\beta_E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, and we can see that the diagonal entries are the eigenvalues. ✡

Definition: The spectral radius of a square matrix or a bounded linear operator is the supremum of the absolute values of the elements in its spectrum, which is sometimes denoted by ρ. That is, let $\lambda_1, \ldots, \lambda_n$ be the (real or complex) eigenvalues of a matrix $A \in M_{n\times n}(\mathbb{F})$. Then its spectral radius $\rho(A) = \max_i |\lambda_i|$.

DIAGONALIZABILITY

Definition: A polynomial $f(t)$ in $P(\mathbb{F})$ is said to split over $\mathbb{F}$ if $\exists$ scalars $c, a_1, \ldots, a_n \in \mathbb{F}$ such that $f(t) = c(x - a_1)(x - a_2)\ldots(x - a_n)$. (beware: the $(x - a_i)$ terms have to be linear)
Examples of polynomials that don't split over $\mathbb{R}$:
- $x^2 + 1$ (not factorable)
- $x^3 - 1 = (x - 1)(x^2 + x + 1)$ (it factors but not into linear terms, hence it doesn't split over $\mathbb{R}$)

• Theorem:
Let $T \in \mathcal{L}(V)$ be diagonalizable. Then char(T) splits over $\mathbb{F}$.

Proof:
Suppose T is diagonalizable. Then $\exists$ an eigenbasis β_E for V such that $[T]_{\beta_E} = D$ is a diagonal matrix.
Suppose that

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & . & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & \lambda_n \end{pmatrix},$$

and let $f(t)$ be the characteristic polynomial of T.
Then,

$$f(t) = \det(D - t\,I_n) = \det\begin{pmatrix} \lambda_1 - t & 0 & 0 & 0 \\ 0 & . & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & \lambda_n - t \end{pmatrix} = \underbrace{(-1)^n}_{c}\Big(t - \underbrace{\lambda_1}_{a_1}\Big)\Big(t - \underbrace{\lambda_2}_{a_2}\Big)\ldots\Big(t - \underbrace{\lambda_n}_{a_n}\Big). \quad \checkmark$$

Thus char(T) splits over $\mathbb{F}$. ■

Definition: Let λ be an eigenvalue of $T \in \mathcal{L}(V)$. Then the algebraic multiplicity of λ is the largest positive integer k for which $(t - \lambda)^k$ is a factor of char(T), i.e. the algebraic multiplicity of λ is the number of times λ appears as a root of char(T).

Example:
Let char(T) $= 5\,(x-3)^2\,(x-4)\,(x+10)^5$. In this case we have roots $3,\ 4,$ and -10 with multiplicity $2,\ 1,$ and 5, respectively.

Note: If T is diagonalizable, each eigenvalue of T must occur on the diagonal as many times as its multiplicity.

Definition: Let $T \in \mathcal{L}(V)$ and λ be an eigenvalue of T.
Then define

$$E_\lambda = \{x \in V : T\,x = \lambda\,x\} = \mathcal{N}(T - \lambda\,I).$$

E_λ is called the eigenspace with respect to λ, i.e. it's the subspace that contains all the eigenvectors that correspond to λ.

• Theorem:

If λ has multiplicity m, then $1 \leq \dim(E_\lambda) \leq m$.

• Theorem:

Let $T \in \mathcal{L}(V)$ and let $\lambda_1, \ \ldots, \ \lambda_k$ be distinct eigenvalues of T. For each $i = 1, \ \ldots, \ k$, let S_i be a finite linearly independent subset of the eigenspace E_{λ_i}. Then $S = S_1 \cup S_2 \cup \ldots. \cup S_k = \bigcup_{i=1}^{k} S_i$ is a linearly independent subset of V.

$\lambda_1 \longrightarrow E_{\lambda_1} = \mathcal{N}(T - \lambda_1 I) \ \supseteq \ S_1$

$\lambda_2 \longrightarrow E_{\lambda_2} = \mathcal{N}(T - \lambda_2 I) \ \supseteq \ S_2$

....

....

$\lambda_k \longrightarrow E_{\lambda_k} = \mathcal{N}(T - \lambda_k I) \ \supseteq \ S_k$

• Theorem:

T is diagonalizable iff

a) $\operatorname{char}(T)$ splits over $\mathbb{F}$

b) $\operatorname{multiplicity}(\lambda_i) = \dim(E_{\lambda_i}) = \operatorname{nullity}(T - \lambda_i I) \ \forall$ eigenvalues λ_i.

Alternatively,

$$\operatorname{multiplicity}(\lambda_i) = \dim(V) - \operatorname{rank}(T - \lambda_i I) = \dim(V) - \operatorname{rank}([T - \lambda_i I]_\beta),$$

where $\dim(V) = n$ given that $[T]_\beta$ is an $n \times n$ matrix (see example below).

Example:

We have an operator $T \in P_2(\mathbb{R})$, which is defined by $T(f(x)) = f(1) + f'(0)\,x + [f'(0) + f''(0)]\,x^2$. Is T diagonalizable?

Solution:

Let us choose the standard basis for $P_2(\mathbb{R})$, $\beta = \{1, \ x, \ x^2\}$. Then

$\bullet\!\!\rightarrow T(1) = 1$

$\bullet\!\!\rightarrow T(x) = 1 + x + x^2$

$\bullet\!\!\rightarrow T(x^2) = 1 + 2\,x^2$

Hence $[T]_\beta = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$.

Now we have

$$\operatorname{char}(T) = \det\left(\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} - \begin{pmatrix} t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & t \end{pmatrix}\right) = \det\begin{pmatrix} 1-t & 1 & 1 \\ 0 & 1-t & 0 \\ 0 & 1 & 2-t \end{pmatrix} = (1-t)(1-t)(2-t) = 0$$

Hence we have roots $t = 1,\ 2,$ with multiplicity 2 and 1, respectively.
We can see that char(T) splits over $\mathbb{F}$. Now we only need to check the multiplicities to determine whether T is diagonalizable :

We have mult(1) = 2 = 3 − rank($A - 1\, I_3$)

where rank($A - 1\, I_3$) = rank$\left(\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$ = rank$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ = 1. So we have confirmed that mult(1) = 2 = 3 − 1 = 2 . ✓

Now we check with the other eigenvalue. We have mult(2) = 1 = 3 − rank($A - 2\, I_3$)

where rank($A - 2\, I_3$) = rank$\left(\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right)$ = rank$\begin{pmatrix} -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ = 2 . So we have confirmed that mult(2) = 1 = 3 − 2 = 1 . ✓

Hence, since both conditions a) and b) from the theorem above hold, we have that T is diagonalizable.