EIGENVALUES AND EIGENVECTORS

1. Diagonalizable linear transformations and matrices

Recall, a matrix, D, is *diagonal* if it is square and the only non-zero entries are on the diagonal. This is equivalent to $D\vec{e}_i = \lambda_i \vec{e}_i$ where here $\vec{e}_i$ are the standard vector and the λ_i are the diagonal entries. A linear transformation, $T : \mathbb{R}^n \to \mathbb{R}^n$, is *diagonalizable* if there is a basis $\mathcal{B}$ of $\mathbb{R}^n$ so that $[T]_{\mathcal{B}}$ is diagonal. This means $[T]$ is similar to the diagonal matrix $[T]_{\mathcal{B}}$. Similarly, a matrix $A \in \mathbb{R}^{n\times n}$ is *diagonalizable* if it is similar to some diagonal matrix D. To *diagonalize* a linear transformation is to find a basis $\mathcal{B}$ so that $[T]_{\mathcal{B}}$ is diagonal. To *diagonalize* a square matrix is to find an invertible S so that $S^{-1}AS = D$ is diagonal.

Fix a matrix $A \in \mathbb{R}^{n\times n}$ We say a vector $\vec{v} \in \mathbb{R}^n$ is an *eigenvector* if

(1) $\vec{v} \neq 0$.
(2) $A\vec{v} = \lambda\vec{v}$ for some scalar $\lambda \in \mathbb{R}$.

The scalar λ is the *eigenvalue associated to* $\vec{v}$ or just an *eigenvalue* of A. Geometrically, $A\vec{v}$ is parallel to $\vec{v}$ and the eigenvalue, λ. counts the stretching factor. Another way to think about this is that the line $L := \mathrm{span}(\vec{v})$ is left invariant by multiplication by A.

An *eigenbasis* of A is a basis, $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_n)$ of $\mathbb{R}^n$ so that each $\vec{v}_i$ is an eigenvector of A.

Theorem 1.1. *The matrix A is diagonalizable if and only if there is an eigenbasis of A.*

Proof. Indeed, if A has eigenbasis $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_n)$, then the matrix

$$S = \begin{bmatrix} \vec{v}_1 & | & \cdots & | & \vec{v}_n \end{bmatrix}$$

satisfies

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

where each λ_i is the eigenvalue associated to $\vec{v}_i$. Conversely, if A is diagonalized by S, then the columns of S form an eigenbasis of A. □

EXAMPLE: The the standard vectors $\vec{e}_i$ form an eigenbasis of $-I_n$. Their eigenvalues are -1. More generally, if D is diagonal, the standard vectors form an eigenbasis with associated eigenvalues the corresponding entries on the diagonal.

EXAMPLE: If $\vec{v}$ is an eigenvector of A with eigenvalue λ, then $\vec{v}$ is an eigenvector of A^3 with eigenvalue λ^3.

EXAMPLE: 0 is an eigenvalue of A if and only if A is not invertible. Indeed, 0 is an eigenvalue $\iff$ there is a non-zero $\vec{v}$ so $A\vec{v} = \vec{0}$ true $\iff$ $\vec{v} \in \ker A$ so $\ker A$ is non-trivial $\iff$ A not invertible.

EXAMPLE: If $\vec{v}$ is an eigenvector of Q which is orthogonal, then the associated eigenvalue is ± 1. Indeed,

$$||\vec{v}|| = ||Q\vec{v}|| = ||\lambda\vec{v}|| = |\lambda|||\vec{v}||$$

as $\vec{v} \neq 0$ dividing, gives $|\lambda| = 1$.

EXAMPLE: If $A^2 = -I_n$, then there are no eigenvectors of A. To see this, suppose $\vec{v}$ was an eigenvector of A. Then $A\vec{v} = \lambda\vec{v}$. As such

$$-\vec{v} = -I_n\vec{v} = A^2\vec{v} = \lambda^2\vec{v}$$

That is, $\lambda^2 = -1$. There are no real numbers whose square is negative, so there is no such $\vec{v}$. This means A has no real eigenvalues (it does have have a comples eigenvalues – see Section 7.5 of the textbook. This is beyond scope of this course).

2. Characteristic Equaiton

One of the hardest (computational) problems in linear algebra is to determine the eigenvalues of a matrix. This is because, unlike everything else we have considered so far, it is a *non-linear* problem. That being said, it is still a tractable problem (especially for small matrices).

To understand the approach. Observe that if λ is an eigenvalue of A, then there is a non-zero $\vec{v}$ so that$A\vec{v} = \lambda\vec{v}$. That is,

$$A\vec{v} = (\lambda I_n)\vec{v} \iff (A - \lambda I_n)\vec{v} = \vec{0} \iff \vec{v} \in \ker(A - \lambda I_n).$$

From which we conclude that $A - \lambda I_n$ is not invertible and so $\det(A - \lambda I_n) = 0$. In summary,

Theorem 2.1. *λ is an eigenvalue of A if and only if*

$$\det(A - \lambda I_n) = 0.$$

The equation $\det(A - \lambda I_n) = 0$ is called the *characteristic equation* of A.

EXAMPLE: Find the eigenvalues of

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}.$$

The characteristic equation is

$$\det(A - \lambda I_2) = \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & 2-\lambda \end{bmatrix} = \lambda^2 - 4\lambda - 5 = (\lambda + 1)(\lambda - 5)$$

Hence, the eigenvalues are $\lambda = -1$ and $\lambda = 5$. To find corresponding eigenvectors we seek non-trivial solutions to

$$\begin{bmatrix} 2-(-1) & 3 \\ 3 & 2-(-1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0} \text{ and } \begin{bmatrix} 2-(5) & 3 \\ 3 & 2-(5) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{0}$$

By inspection the non-trivial solutions are

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Hence,

$$\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}$$

So we have diagonalized A.

EXAMPLE: Find eigenvalues of

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

So

$$\det(A - \lambda I_3) = \det \begin{bmatrix} 1-\lambda & -2 & 3 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{bmatrix} = (1-\lambda)(2-\lambda)(3-\lambda)$$

Hence, eigenvalues are $1, 2, 3$.

This example is a special case of a more general phenomena.

Theorem 2.2. *If M is upper triangular, then the eigenvalues of M are the diagonal entries of M.*

EXAMPLE: When $n = 2$, the eigenvalues of

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

so the characteristic equation is

$$\det(A - \lambda I_2) = \det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} = \lambda^2 - (a+d)\lambda + (ad - bc) = \lambda^2 - \mathrm{tr} A\lambda + \det A$$

Using the quadratic formula we have the following:

(1) When $tr(A)^2 - 4\det A > 0$, then two distinct eigenvalues
(2) When $tr(A)^2 - 4\det A = 0$, exactly one eigenvalue $\frac{1}{2} trA$.
(3) When $tr(A)^2 - 4\det A < 0$, then no (real) eigenvalues.

3. Characteristic Polynomial

As we say for a 2×2 matrix, the characteristic equation reduces to finding the roots of an associated quadratic polynomial. More generally, for a $n \times n$ matrix A, the characteristic equation $\det(A - \lambda I_n) = 0$ reduces to finding roots of a degree n polynomial o fthe form

$$f_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} (\mathrm{tr} A) \lambda^{n-1} + \cdot + \det A$$

this is called the *characteristic polynomial of A*. To see why this is true observe that if P is the diagonal pattern of $A - \lambda I_n$, then

$$\mathrm{prod}(P) = (a_{11} - \lambda) \cdots (a_{nn} - \lambda) = (-1)^n \lambda^n + (-1)^{n-1} (trA) \lambda^{n-2} + R_P(\lambda)$$

where $R_P(\lambda)$ is a polynomial of degree at most $n - 2$ that depends on P (and of course also on A and λ). If P is some other pattern of $A - \lambda I_n$, then at least two entries are not on the diagonal. Hence,

$$\mathrm{prod}(P) = R_P(\lambda)$$

for some $R_P(\lambda)$ that is a polynomial of degree at most $n - 2$ that depends on P. Hence,

$$f_A(\lambda) = \det(A - \lambda I_n) = (-1)^n \lambda^n + (-1)^{n-1} (\mathrm{tr} A) \lambda^{n-1} + R(\lambda)$$

where $R(\lambda)$ has degree at most $n - 2$. Finally, $f_A(0) = \det(A)$ and so the constant term of f_A is $\det A$ as claimed.

EXAMPLE: If n is odd, there is always at least one real eigenvalue. Indeed, in this case

$$\lim_{n\to\pm\infty} f_A(\lambda) = \lim_{n\to\infty} -\lambda^3 = \mp\infty$$

That is for $\lambda >> 1$, $f_A(\lambda) < 0$ and for $\lambda << -1$, $f_A(\lambda) > 0$ and so by the intermediate value theorem from calculus (and the fact that polynomials are continuous), $f_A(\lambda_0) = 0$ for some λ_0. This may fail when n is even.

An eigenvalue λ_0 has *algebraic multiplicity* k if

$$f_A(\lambda) = (\lambda_0 - \lambda)^k g(\lambda)$$

where g is a polynomial of degree $n-k$ with $g(\lambda_0) \neq 0$. Write $\mathrm{almu}(\lambda_0) = k$ in this case.

EXAMPLE: If

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

then $f_A(\lambda) = (2-\lambda)^2(1-\lambda)(-1-\lambda)$ and so $\mathrm{almu}(2) = 2$, while $\mathrm{almu}(1) = \mathrm{almu}(-1) = 1$. Strictly speaking, $\mathrm{almu}(0) = 0$, as 0 is not an eigenvalue of A and it is sometimes convenient to follow this convention.

We say an eigenvalue, λ, is *repeated* if $\mathrm{almu}(\lambda) \geq 2$.

Algebraic fact, counting algebraic multiplicity, a $n \times n$ matrix has at most n real eigenvalues. If n is odd, then there is at least one real eigenvalue. The fundamental theorem of algebra ensures that, counting multiplicity, such a matrix always has exactly n complex eigenvalues.

We conclude with a simple theorem

Theorem 3.1. *If $A \in \mathbb{R}^{n\times n}$ has eigenvalues $\lambda_1, \dots, \lambda_n$ (listed counting multiplicity):*

(1) $\det A = \lambda_1 \lambda_2 \cdots \lambda_n$.
(2) $trA = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Proof. If the eigenvalues are $\lambda_1, \dots, \lambda_n$, then we have the complete factorization of $f_A(\lambda)$

$$f_A(\lambda) = (\lambda_1 - \lambda)\dots(\lambda_n - \lambda)$$

This means $\det(A) = f_A(0) = \lambda_1 \dots \lambda_n$. $(\lambda_1 - \lambda)\dots(\lambda_n - \lambda) = (-1)^n\lambda^n + (-1)^{n-1}(\lambda_1 + \dots + \lambda_n)\lambda^{n-1} + R(\lambda)$ where $R(\lambda)$ of degree at most $n-2$. Comparing the coefficient of the λ^{n-2} term gives the result. ☐

4. Eigenspaces

Consider an eigenvalue λ of $A \in \mathbb{R}^{n\times n}$. We define the *eigenspace* associated to λ to be

$$E_\lambda = \ker(A - \lambda I_n) = \{\vec{v} \in \mathbb{R}^n : A\vec{v} = \lambda\vec{v}\} \subset \mathbb{R}^n.$$

Observe that $\dim E_\lambda \geq 1$. All *non-zero* elements of E_λ are eigenvectors of A with eigenvalue λ.

EXAMPLE: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ has repeated eigenvalue 1. Clearly,

$$E_1 = \ker(A - I_2) = \ker(0_{2\times 2}) = \mathbb{R}^2$$

Similarly, the matrix $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ has one repeated eigenvalue 1. However,

$$\ker(B - I_2) = \ker \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} = \mathrm{span}(\begin{bmatrix} 1 \\ 0 \end{bmatrix}).$$

Motivated by this example, define the *geometric multiplicity* of an eigenvalue λ of $A \in \mathbb{R}^{n\times n}$ tobe

$$\mathrm{gemu}(\lambda) = \mathrm{null}(A - \lambda I_n) = n - \mathrm{rank}(A - \lambda I_n) \geq 1.$$

5. Diagonalizable Matrices

We are now ready to give a computable condition that will allow us to determine an answer to our central question in this part of the course: When is $A \in \mathbb{R}^{n\times n}$ diagonalizable?

Theorem 5.1. *A matrix $A \in \mathbb{R}^{n\times n}$ is diagonalizable if and only if the sum of the geometric multiplicities of all of the eigenvalues of A is n.*

EXAMPLE: For which k is the following diagonalizable

$$\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}?$$

As this is upper triangular, the eigenvalues are 1 with $\mathrm{almu}(1) = 2$ and 2 with $\mathrm{almu}(2) = 1$. It is not hard to see that $\mathrm{gemu}(1) = 1$ when $k \neq 0$ and $\mathrm{gemu}(1) = 2$ when $k = 0$. We always have $\mathrm{gemu}(2) = 1$ Hence, according to the theorem the matrix is diagonalizable only when it is already diagonal (that is $k = 0$) and is otherwise not diagonalizable.

To prove this result we need the following auxiliary fact

Theorem 5.2. *Fix a matrix $A \in \mathbb{R}^{n\times n}$ and let $\vec{v}_1, \ldots, \vec{v}_s$ be a set of vectors formed by concatenating a basis of each non-trivial eigenspace of A. This set is linearly independent (and so $s \leq n$.)*

To explain what I mean by concatenating. Suppose $A \in \mathbb{R}^{5\times 5}$ has exactly three distinct eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 3$ and $\lambda_3 = 4$ If $\mathrm{gemu}(2) = 2$ and

$$E_2 = span(\vec{a}_1, \vec{a}_2)$$

while $\mathrm{gemu}(3) = \mathrm{gemu}(4) = 1$ and

$$E_3 = span(\vec{b}_1) \text{ and } E_4 = span(\vec{c}_1),$$

then their concatenation is the set of vectors

$$(\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4) = (\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{c}_1).$$

According to the auxiliary theorem this list is linearly independent. As $s = 4 < 5 = n$, if this was the complete set of eigenvalues, A would not be diagonalizable by the main theorem. It would be if we had omitted one eigenvalue from our list.

Let us now prove the auxiliary theorem.

Proof. If the vectors $\vec{v}_i$ are not linearly independent, then they at least one is redundant. Let $\vec{v}_m$ be the first redundant vector on the list That is for some $1 \leq m \leq s$ we can write

$$\vec{v}_m = \sum_{i=1}^{m-1} c_i \vec{v}_i$$

and cannot do this for any smaller m. This means $\vec{v}_1, \ldots, \vec{v}_{m-1}$ are linearly independent.

Let λ_m be the eigenvalue associated to $\vec{v}_m$. Observe, there must be some $1 \leq k \leq m-1$ so that $\lambda_k \neq \lambda_m$ and $c_k \neq 0$ as otherwise we would have a non-trivial linear relation for of a set of linearly independent vectors in E_{λ_m} (which is impossible). Clearly,

$$\vec{0} = (A - \lambda_m I_n)\vec{v}_m = \sum_{i=1}^{m-1} c_i(\lambda_i - \lambda_m)\vec{v}_i$$

Hence, $c_i(\lambda_i - \lambda_m) = 0$ for each i. This contradicts, $\lambda_k \neq \lambda_m$ and $c_k \neq 0$ and proves the claim. ☐

The main theorem follows easily form this. Indeed, the hypotheses gives n lin indep vectors all which are eigenvectors of A. That is, an eigenbasis of A.

6. Strategy for diagonalizing $A \in \mathbb{R}^{n \times n}$

We have the following strategy for diagonalizing a given matrix:

(1) Find eigenvalues of A by solving $f_A(\lambda) = 0$ (this is a non-linear problem).
(2) For each eigenvalue λ find a basis of the eigenspace E_λ (this is a linear problem).
(3) The A is diagonalizable if and only if the sum of the dimensions of the eigenspaces is n. In this case, obtain an eigenbasis, $\vec{v}_1, \ldots, \vec{v}_n$, by concatenation.
(4) As $A\vec{v}_i = \lambda_i \vec{v}_i$, setting

$$S = \begin{bmatrix} \vec{v}_1 & | & \cdots & | & \vec{v}_n \end{bmatrix}, S^{-1}AS = D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

EXAMPLE: Diagonalize (if possible)

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

We compute

$$f_A(\lambda) = \det(A - \lambda I_3) = \det \begin{bmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 0 \\ 1 & 1 & 1-\lambda \end{bmatrix} = (1-\lambda)\det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$$

$$f_A(\lambda) = (1-\lambda)(\lambda^2 - 4\lambda + 3) = (1-\lambda)(\lambda-3)(\lambda-1) = (1-\lambda)^2(3-\lambda)$$

Hence, eigenvalues are $\lambda = 1$ and $\lambda = 3$. Have $\mathrm{almu}(1) = 2$ and $\mathrm{almu}(3) = 1$. We compute

$$E_1 = \ker(A - I_3) = \ker \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = span(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix})$$

Hence, $\mathrm{gemu}(1) = 2$

Likewise,

$$E_3 = \ker(A - 3I_3) = \ker \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & -2 \end{bmatrix} = span(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix})$$

Hence, $\mathrm{gemu}(3) = 1$. As $\mathrm{gemu}(1)+\mathrm{gemu}(3) = 2+1 = 3$ the matrix is diagonalizable. Indeed, setting

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

We have

$$S^{-1}AS = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

7. Eigenvalues of linear transformations

Fix a linear space V and consider a linear transformation $T : V \to V$. A scalar λ is an *eigenvalue* of T, if

$$T(f) = \lambda f$$

for some nonzero (nonneutral) element $f \in V$. In general we refer to such f as a *eigenvector*. If V is a space of functions, then it is customary to call f an *eigenfunction*, etc. If $\mathcal{B} = (f_1, \ldots, f_n)$ is a basis of V and each f_i is an eigenvector, then we say $\mathcal{B}$ is an eigenbasis of T. If $\mathcal{B}$ is an eigenbasis of T, then T is diagonalizable as

$$[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

EXAMPLE: If $V = C^\infty$, and $D(f) = f'$, then

$$D(e^{kx}) = \frac{d}{dx} e^{kx} = k e^{kx}$$

so each $f_k(x) = e^{kx}$ is an eigenfunction and every scalar $k \in \mathbb{R}$ is an eigenvalue.

EXAMPLE: Consider the map $T : P_2 \to P_2$ given by $T(p) = p(2x + 1)$. Is T diagonalizable? As usual it is computationally more convenient to work in some basis. To that end, let $\mathcal{U} = (1, x, x^2)$ be the usual basis of P_2. As

$$[T(1)]_{\mathcal{U}} = [1]_{\mathcal{U}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$[T(x)]_{\mathcal{U}} = [2x + 1]_{\mathcal{U}} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$[T(x^2)]_{\mathcal{U}} = [4x^2 + 4x + 1]_{\mathcal{U}} = \begin{bmatrix} 1 \\ 4 \\ 4 \end{bmatrix},$$

the associated matrix is

$$A = [T]_{\mathcal{U}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{bmatrix}.$$

This matrix is upper triangular, with distinct eigenvalues $1, 2$ and 4 This means T is also diagonalizable and has the same eigenvalues. We compute (for A)

$$E_1 = \ker(A - I_3) = \ker \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} = \operatorname{span}(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix})$$

$$E_2 = \ker(A - 2I_3) = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{bmatrix} = \operatorname{span}(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix})$$

$$E_3 = \ker(A - 4I_3) = \begin{bmatrix} -3 & 1 & 1 \\ 0 & -2 & 4 \\ 0 & 0 & 0 \end{bmatrix} = \operatorname{span} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Hence, A can be diagonalized by $S = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Going back to T we check

$$T(1) = 1$$

$$T(1 + x) = 1 + (2x + 1) = 2(x + 1)$$

$$T(1 + 2x + x^2) = 1 + 2(2x + 1) + (2x + 1)^2 = 4(1 + 2x + x^2)$$

In particular, $\mathcal{B} = (1, 1 + x, 1 + 2x + x^2)$ is an eigenbasis and

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

8. Eigenvalues and Similarity

There is a close relationship between similar matrices and their eigenvalues. This is clearest for diagonalizable matrices but holds more generally.

Theorem 8.1. *If A is similar to B*

1. $f_A(\lambda) = f_B(\lambda)$
2. $rank(A) = rank(B)$ *and* $null(A) = null(B)$
3. *A and B have the same eigenvalues with the same algebraic and geometric multiplicities (the eigenvectors are in general different)*
4. $trA = trB$ *and* $det(A) = det(B)$

Proof. Proof of claim (1): A is similar to B means $B = S^{-1}AS$ for some invertible S. Hence,

$$f_B(\lambda) = \det(B - \lambda I_n) = \det(S^{-1}AS - \lambda I_n) = \det(S^{-1}(A - \lambda I_n)S)$$
$$= \det(A - \lambda I_n) = f_A(\lambda).$$

Claim (2) was shown in a previous handout (the one on similar matrices).

Claim (3) can be shown as follows: By claim (1), $f_A(\lambda) = f_B(\lambda)$ and so A and B have the same eigenvalues with the same algebraic multiplicities. Furthermore, if λ is an eigenvalue of both A and B, then $A - \lambda I_n$ is similar to $B - \lambda I_n$. Hence, by claim (2)

$$\mathrm{gemu}_A(\lambda) = \mathrm{null}(A - \lambda I_n) = \mathrm{null}(B - \lambda I_n) = \mathrm{gemu}_B(\lambda)$$

so the geometric multiplicities are the same as well. Notice $E_\lambda(A) \neq E_\lambda(B)$ in general. Finally, claim (4) Follows from claim (1) and the observation that the characteristic polynomial encodes the trace and determinant. □

EXAMPLE: The matrix $\begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}$ is not similar to $\begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix}$ as their traces are different.

Even if a matrix is not diagonalizable, it is still similar to one of a list of *canonical* matrices. To do this is full generality is beyond the scope of this course. To illustrate the idea we present the list for 2×2 matrices.

Theorem 8.2. *Any $A \in \mathbb{R}^{2\times 2}$ is similar to one of the following:*

(1) $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ *for $\lambda_1, \lambda_2 \in \mathbb{R}$. This occurs when A is diagonalizable.*

(2) $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$. *This occurs when A has repeated eigenvalue λ with* $\mathrm{gemu}(\lambda) = 1$. *Such A is sometimes called defective.*

(3) $\begin{bmatrix} \lambda_1 & -\lambda_2 \\ \lambda_2 & \lambda_1 \end{bmatrix} = \sqrt{\lambda_1^2 + \lambda_2^2} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ *where $\lambda_2 > 0$ and $\pi > \theta > 0$. This corresponds to an A with no real eigenvalues (in this case the complex eigenvalues of A are $\lambda_1 \pm i\lambda_2$).*