SECTION 2: LINEAR TRANSFORMATION

Definitions:

1. A <u>transformation</u> (or function or mapping) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector, x in $\mathbb{R}^n$, an image vector $T(x) \in \mathbb{R}^m$. The set $\mathbb{R}^n$ is called the domain of T, and $\mathbb{R}^m$ is called the codomain of T. For x in $\mathbb{R}^n$ the set of all images T(x) is called the range of T.
2. A <u>transformation (or mapping), T is linear</u> if
   i)$T(u + v) = T(u) + T(v)$ for all $u,v \in$ dom T and
   ii) $T(cu) = cT(u)$ for all $u \in$ dom T & c is a scalar.
   Eg L: $\mathbb{R}^3 \to \mathbb{R}^3$, $L(x,y,z) = (x,y)$ is a linear transformation while $T:\mathbb{R}\to\mathbb{R}$, $T(x) = x + 1$ is not.

Properties of Linear Transformations

Let T be a linear transformation, then

a. $T(0) = 0$
b. $T(cu + dv) = cT(u) + dT(v)$
c. $T(c_1v_1 + \dots + c_pv_p) = c_1T(v_1) + \dots + c_pT(v_p)$ for all vectors $u,v \in$ dom T & scalars c,d.

e.g let L: $\mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation for which

$L(1, 0, 0) = (2, -1)$; $L(0, 1, 0) = (3,1)$ & $L(0, 0, 1) = (-1, 2)$. Find $L(-3, 4, 2)$

Theorem:

Let T: $\mathbb{R}^n\to\mathbb{R}^m$be a linear transformation. Then there exists a unique matrix A such that $T(x) = Ax$ for all x in $\mathbb{R}^n$. In fact, A is the mxn matrix whose jth column is the vector T(ej), where ej is the jth column of the identity matrix of order n.

That is $A = (T(e_1)\ \ T(e_2)\ \dots\ \ T(e_n))$ is called the standard matrix for the linear transformation T.

Example

Find the standard matrix, A, for the linear transformations below

a. $T(x) = 3x, x \in \mathbb{R}^2$ b. L: $\mathbb{R}^3 \to \mathbb{R}^3$, $L\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+y \\ y-z \\ x+z \end{pmatrix}$

Definitions

Let T: U $\to$ V be a linear transformation. Then

1. the <u>kernel of T</u> denoted by ker T, is the set of all vectors in U that are mapped by T into the zero vector of V. in other words ker T = $\{X \in U: T(X) = 0\}$
2. the <u>image of T</u>, denoted Im T, is the set of all vectors in V that are images under T of vectors in U. that is Im T= $\{ Y\in V: Y = T(X)$ for some $X\in U\}$
3. T is called an <u>epimorphism</u> if Im T = V, that is, if T is surjective.
4. T is a <u>monomorphism</u> if it is injective
5. T is an <u>isomorphism</u> if it is both an epimorphism and a monomorphism – that is, if T is bijective.

6. If U is a finite dimensional vector space, the dimension of ker T is called the nullity of T and the dimension of Im T is called the rank of T.

Theorem

Let T: U → V be a linear transformation of U into V. then each of the following is true.

1. Ker T is a subspace of U.
2. T is a monomorphism if and only if ker T = {0}
3. Im T is a subspace of V and T is an epimorphism if and only if Im T = V.
4. If W is a subspace of U, then T(W) is a subspace of V
5. If T is an isomorphism, then T^{-1} : V → U is also an isomorphism.
6. dim U = rank of T + nullity of T
7. The T, linear transformation is completely determined by its effect upon a basis of U.

Note:

1. The rank and nullity of a matrix A are the rank and nullity of the associated linear transformation respectively.

Example:

1. Find ker T and Im T for T: $\mathbb{R}^3 \to \mathbb{R}^3$ defined by T(a,b,c) = (a,b,0).
2. Find the nullity and rank of T
3. Find a basis of Im T and ker T where T(a,b,c) = (2a + b + 3c, 3a – b + c, -4a + 3b + c).
4. Let T:$\mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that T(1,0) = (1,3) and T(0,1) = (-2,0). Find the rule for T, that is, find T(x,y) for any vector (x,y).
5. Let T:$\mathbb{R}^2 \to \mathbb{R}^2$ be defined by T(x,y,z) = (2x-z, 7y+z). Find A.

Theorem 14. Let A be an n x n matrix. Then the following are equivalent.

(1) A is invertible.
(2) The columns of A form a basis of $\mathbb{R}^n$.
(3) For every $b \in \mathbb{R}^n$, there is exactly one solution to the system Ax = b.
(4) For every $b \in \mathbb{R}^n$, there is at most one solution to the system Ax = b.
(5) For every $b \in \mathbb{R}^n$, there is at least one solution to the system Ax = b.
(6) The only solution to Ax = 0 is 0.
(7) The columns of A are linearly independent.
(8) The columns of Aspan $\mathbb{R}^n$.
(9) Nullity (A) = 0.
(10) rank(A) =n.
(11) A is left-invertible.
(12) A is right-invertible.
(13) T_A is injective.
(14) T_A is surjective.
(15) A^T is invertible.

Proof. We have seen that (4) ⇔ (6) ⇔ (7) ⇔ (9) ⇔ (11) ⇔ (13) and

(5) ⇔ (8) ⇔(10) ⇔ (12) ⇔ (14) and (1) ⇔ (3) ⇔ (2) and clearly(2) ⇒ (7), (8). The rank-nullity theorem fills in the gap by proving (9) ⇔ (10); for if null(A) = 0, then rank(A) = n – 0 = n by rank-nullity, similarly if rank(A) = n then null(A) = n – n = 0.

Finally, check out (1) ⇔ (15).

Example. The following are all linear transformations.

(1) The identity map id :$\mathbb{R}^n \to \mathbb{R}^n$ which takes x $\in \mathbb{R}^n$ to itself.
(2) The map which rotates a point x in the plane $\mathbb{R}$ around the origin by a fixed angle θ.
(3) The map which takes a given x $\in \mathbb{R}^n$ to Ax $\in \mathbb{R}^m$, where A is a fixed mxn matrix.

It is an important fact that the last example is completely representative of all linear transformations.

Theorem 2. If T: $\mathbb{R}^n \to \mathbb{R}^m$is a linear transformation, then there is some mxn matrix A such that for all x$\in \mathbb{R}^n$,

$$T(x) = Ax$$

Moreover such an A is unique; that is there is exactly one mxn matrix A such that the above holds for all x $\in \mathbb{R}^n$.

Proof. We will make use of the following vectors in $\mathbb{R}^n$, which we know:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} e_2 = \begin{pmatrix} 0 \\ 1 \\ . \\ . \\ . \\ 0 \end{pmatrix} \quad \cdots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 1 \end{pmatrix}$$

These are called the standard basis vectors for $\mathbb{R}^n$. now let $a_i = T(e_i)$ for each $i \leq n$. then each a_i is a column vector in $\mathbb{R}^m$. Let A be the mxn matrix whose ith column (from left to right) isa_i. We llshow that this A works, ie, $\forall x \in \mathbb{R}^n$ T(x) = Ax

To see this, let x $\in \mathbb{R}^m$ be given. Then

$$x = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} x_1 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ . \\ . \\ . \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ . \\ x_n \end{pmatrix}$$

$$= x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

Hence, by Fact 4,

$$T(x) = T(x_1e_1 + x_2e_2 + \cdots + x_ne_n = x_1T(e_1) + x_2T(e_2) + \cdots + x_nT(e_n)$$

$$= x_1a_1 + x_2a_2 + \cdots + x_na_n = Ax$$

Since $x \in \mathbb{R}_n$ was given, this proves that there is a matrix A which implements T. To show uniqueness, ie that there is at most one such A, suppose both A and B are mxn matrices which implement T. Then

$$\forall x \in \mathbb{R}^n \quad Ax = T(x) = Bx$$

Now we look again at the standard basis vectors; we have $Ae_i = T(e_i) = Be_i$ for each $i \leq n$. notice that Ae_i and Be_i are the i^{th} columns of A and B respectively. Hence the i^{th} column of A is equal to the i^{th} column of B for each $i \leq n$; it follows then that A = B. ⊡

Definition 5. The standard matrix of a linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the unique mxn matrix A such that T(x) = Ax for all $x \in \mathbb{R}_n$.

Theorem 3. Let A be an mxn matrix. Then the following are equivalent:

(1) For any choice of constants $b_1, \ldots, b_m$, the system of liner equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\;\;. \\ &\;\;. \\ &\;\;. \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

has atleast one solution.

(2) For every $b \in \mathbb{R}^m$, there is some $s \in \mathbb{R}^n$ such that As = b

(3) $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is surjective.

(4) A is right-invertible.

Proof. The equivalence of (1), (2) and (3) is just a recapitulation of the definitions in play. The real work is in proving they're equivalent to (4).

Let's prove that (2) is equivalent to (4).

Firstly let's assume (2) and prove (4). Consider the standard basis vectors in $\mathbb{R}^m$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} e_2 = \begin{pmatrix} 0 \\ 1 \\ . \\ . \\ . \\ 0 \end{pmatrix} \quad \cdots \quad e_m = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 1 \end{pmatrix}$$

By our assumption, there are vectors $s_1, \ldots, s_m \in \mathbb{R}^m$ such that $As_1 = e_1, \ldots, As_m = e_m$. Now let B be the matrix with columns $s_1, \ldots, s_m$, in that order. Then B isnxm, and for any i, j, $B_{ij} = (s_j)_i$, the ith entry of column vector s_j. Then for all $i, j \leq m$, we have

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{n} A_{ik}\,(s_j)_k = (As_j)_i = (e_j)_i = I_{ij}$$

So AB = I.

Now assume (4) and let B be a right inverse of A. let $b \in \mathbb{R}^m$ be given. Then

A(Bb) = (AB)b = Ib = b and so there is some $s \in \mathbb{R}^n$ such that As = b, namely s = Bb.

So far, we've been working exclusively with real numbers. Soon we'll have to use complex numbers. Everything we have done up to now works for complex numbers too. Recall the following

Fact 10. Let z = a + bi be a nonzero complex number (so at least one of a and b is nonzero). Then

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2} = \frac{a - bi}{a^2 + b^2}$$

In particular, if z lies on the unit circle (so z =cosθ + isinθ) then

$$\frac{1}{z} = \bar{z}$$

Example. Find the rank and nullity of the matrix A below. Find bases for row(A), null(A) and col(A). A =

$$\begin{pmatrix} 1+2i & -1 & 0 \\ 1-2i & i & 3 \\ 0 & 2 & 1 \end{pmatrix}$$

Example.

(1) Let p and q be points in $\mathbb{R}^2$ (ie the plane). Show that {p, q} is linearly independent if and only if the points 0, p and q are not colinear.
(2) Let p, q, r ∈ $\mathbb{R}^3$. Show that {p, q, r} is linearly independent if and only if the points 0, p, q, and r are not coplanar.

Definition. $M_n(\mathbb{C})$ is the set of n x n matrices with entries from $\mathbb{C}$. We will identify a matrix A ∈ $M_n(\mathbb{C})$ with its sequence of rows, $\rho_1, \dots, \rho_n \in \mathbb{C}^n$.

A multilinear map is a function T: $M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ such that

(1) For all $\rho_1, \dots, \rho_n \in \mathbb{C}^n$ and $\sigma \in \mathbb{C}$,
 $T(\rho_1, \dots,\rho_i + \sigma, \dots, \rho_n) = T(\rho_1, \dots, \rho_i, \dots, \rho_n) + T(\rho_1, \dots,\sigma, \dots, \rho_n)$
(2) For all $\rho_1, \dots, \rho_n \in \mathbb{C}^n$ and $t \in \mathbb{C}$
 $T(\rho_1, \dots, t\rho_i, \dots, \rho_n) = tT(\rho_1, \dots, \rho_i, \dots, \rho_n)$

In other words, T is a linear map on its i^{th} argument when all others are fixed.

A multilinear map T is alternating if in addition we have

$T(\rho_1, , \dots, \rho_j \dots, \rho_i, \dots, \rho_n) = -T(\rho_1, \dots, \rho_i, \dots, \rho_j \dots, \rho_n)$

whenever $\rho_1, \ldots, \rho_n \in \mathbb{C}^n$ and $i < j$.

Lemma 17. Suppose $T: M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ is an alternating multilinear map, and $A \in M_n(\mathbb{C})$ is some matrix with two rows which are the same, or a row of all zeroes. Then $T(A) = 0$.

Proof. Suppose A has two rows which are the same, ie $\rho_i = \rho_j$ for $i \neq j$. Then $T(\rho_1, , \ldots, \rho_i \ldots, \rho_j , \ldots, \rho_n) = -T(\rho_1, \ldots, \rho_j, \ldots, \rho_i \ldots, \rho_n) = -T(\rho_1, , \ldots, \rho_i \ldots, \rho_j , \ldots, \rho_n)$. The only real (or complex) number t satisfying $t = -t$ is $t = 0$.

Now suppose A has a zero row in the ith place. Then by linearity, $T(\rho_1, , \ldots, 0_{1xn} , \ldots, \rho_n) = T(\rho_1, , \ldots, 0.0_{1xn} , \ldots, \rho_n)=0$. $T(\rho_1, , \ldots, 0_{1xn} , \ldots, \rho_n) = 0$.

Theorem 15. Suppose $T: M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ is an alternating multilinear map, and A and B are n x n matrices such that B is the result of applying a single row operation to A. Then T(A) and T(B) are related in the following way depending on the row operation in question:

$i \neq j$ $\quad \rho_i \leftrightarrow \rho_j \quad T(B) = -T(A)$

$i \neq j$ $\quad \rho_i \longrightarrow \rho_i + \lambda\rho_j \quad T(B) = T(A)$

$\quad \rho_i \longrightarrow \lambda\rho_i \quad T(B) = \lambda T(A)$

Proof. The swap and scaling cases are simply part of the definition of an alternating multilinear map. The row-combination case follows from the lemma above.

$T(\rho_1, \ldots, \rho_i + \lambda\rho_j ,\ldots, \rho_j , \ldots, \rho_n) = T(\rho_1,\ldots, \rho_i,\ldots, \rho_j ,\ldots, \rho_n) + \lambda T(\rho_1,\ldots, \rho_j\ldots, \rho_j ,\ldots, \rho_n)$

$= T(\rho_1, , \ldots, \rho_i \ldots, \rho_j , \ldots, \rho_n)$

Theorem 17: If $S, T: M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ are alternating multilinear maps, and $S(I) = T(I)$, then $S = T$.

Proof. We have already seen that if A is not invertible then S and T both send A to zero. If A is invertible, then it is row-reducible to I. Theorem 15 (along with a routine induction) shows that

$T(I) = (-1)^k m_1 \ldots m_t T(A) \qquad S(I) = (-1)^k m_1 \ldots m_t S(A)$

where k is the number of swaps used, t is the number of scaling operations used, and $m_1, \ldots, m_t$ are the scaling factors. Hence if $S(I) = T(I)$, then

$S(A) = (-1)^k \frac{1}{m_1} \cdots \frac{1}{m_t} S(I) = (-1)^k \frac{1}{m_1} \cdots \frac{1}{m_t} T(I) = T(A)$

Definition. The determinant $\det: M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ is the unique alternating multilinear map such that $\det(I) = 1$.

Theorem 18. Let A be a square matrix. Then the following are equivalent:

(1) A is invertible.
(2) $\det(A) \neq 0$.

Example. Find the determinants of the following matrices.

$A = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}, B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 6 & 10 \\ 6 & 11 & 17 \end{pmatrix}$, D = $\begin{pmatrix} -2 & 2 & 3 \\ -1 & 1 & 3 \\ 2 & 0 & -1 \end{pmatrix}$ Theorem 19. Let A, B $\in M_n(\mathbb{C})$. Then det(AB) = det(A)det(B).

Definition. A permutation matrix is an n x n matrix P such that

(1) every entry in P is either 0 or 1 and
(2) there is exactly one 1 in each row and column of P.

If P is a permutation matrix then there is an associated permutation of [n] = {1,…, n}:

$$\pi(i) = j \Leftrightarrow P_{ij} = 1$$

We write $P = P_\pi$. Note that det(P) = ±depending on the number of swaps needed to reduce P to I. (Which it turns out is the number of transpositions needed to produce π.)

For convenience, a row vector in $\mathbb{C}^n$ is used to describe a permutation π of [n]. specifically, in the ith entry of the row vector we write π(i); eg (312) denotes the permutation π such that π(1) = 3, π(2) = 1 and π(3) = 2.

Definition. perm(n) = {π: [n] ⟶ [n] | π is a bijection}.

Fact 11.If A is an n x n matrix, then

$$\det(A) = \sum_{\pi \in perm(n)} \det(P_\pi) \prod_{i=1}^{n} A_{i\pi(i)}$$

example. Say A is a 3 x 3 matrix $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$

The 3! = 6 permutations of [3] are listed below, along with the determinant of the associated permutation matrix:

(1 2 3) 1

(1 3 2) -1

(2 1 3) 1

(2 3

Definition:
The inverse T^{-1} of a linear transformation T: V → V is a mapping satisfying: $TT^{-1} = T^{-1}T = I$, where I is the identity linear transformation. T^{-1} is also a linear transformation.

Theorem

If A is the matrix of the linear transformation T: $\mathbb{R}^n \rightarrow \mathbb{R}^n$ relative to the ordered basis consisting of the unit vectors, then A has an inverse if and only if T has an inverse. Moreover, if A^{-1} exists, then A-1 is the matrix of T^{-1} relative to the ordered bases of unit vectors. Hence to compute $T^{-1}(X)$, we need only compute $A^{-1}X$.

e.g.find the inverse of the linear transformation T: $\mathbb{R}^3 \rightarrow \mathbb{R}^3$defined by T(a + 2c, 2a + b + 2c, c). Find T^{-1}.

Hint: Look for the matrix of T to have B = $\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ and then find B^{-1} as $\begin{pmatrix} 1 & 0 & -2 \\ -2 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ then compute the product $B^{-1}X$ = $\begin{pmatrix} x-2z \\ -2x+y+2z \\ z \end{pmatrix}$where X = $(x,y,z)^T$ and the transpose of this product gives T^{-1}.

T^{-1} (a,b,c) = (a – 2c, -2a + y + 2z, z).

EIGENVALUES & EIGENVECTORS

Definition: A <u>fixed point</u> of an nxn matrix A is a vector x in $\mathbb{R}^n$ such that Ax = x.

Note that every square matrix has atleast one fixed point, namely x = 0. We call this the trivial fixed point of A, viz

Ax =x ⇒ Ax = Ix ⇒ (I – A) x = 0. Solutions of (I – A)x=0 are the fixed points of A.

Theorem

If A is an nxn matrix, then the following statements are equivalent:

a. A has nontrivial fixed points
b. I – A is singular
c. det(I – A) = 0.

Definition: If A is an nxn matrix, then a scalar λ is called an eigenvalue of A if there is a nonzero vector x such that Ax = λx. If λ is an eigenvalue of A, then every nonzero vector s such that Ax = λx is called an eigenvector of A corresponding to λ.

Eg A = $\begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}$, u = $\begin{pmatrix} 6 \\ -5 \end{pmatrix}$& v = $\begin{pmatrix} 3 \\ -2 \end{pmatrix}$. Are u and v eigenvectors of A? Show that 7 is an eigenvalue of A and find its corresponding eigenvectors.

Definition: The set of all solutions of (A –λI)x = 0 is just the null space of the matrix A – λI. So this set is a subspace of $\mathbb{R}^n$ and is called the eigenspace of A corresponding to λ.

That is the eigenspace of λ consists of the zero vector and all the eigenvectors corresponding to λ.

Example

Find a basis for the corresponding eigenspace where eigenvalue of A is 2 given that A = $\begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}$

Theorem

1. If A is a triangular matrix, then the eigenvalues of A are the entries on the main diagonal of A.
2. If λ is an eigenvalue of a matrix A and x is a corresponding eigenvector, and if k is any positive integer, then λ^k is an eigenvalue of A^k and x is a corresponding eigenvector.

Definition

The characteristic equation or characteristic polynomial of a matrix A is defined as det $(A - \lambda I) = 0$ or det $(A - \lambda I)$ respectively, where λ is the eigenvalue and I is the identity matrix.

Note: solving $(A - \lambda I)x = 0$is equivalent to finding all λ such that $A - \lambda I$ is not invertible. That is when its determinant is zero. Thus the eigenvalues of A are the solutions of the characteristic polynomials of A.

Examples

Find the eigen values and corresponding eigenvectors of A = $\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$ & B = $\begin{pmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{pmatrix}$

Theorem 4:

1. The characteristic roots of a linear transformation T of V are precisely the characteristic roots of the matrix of T relative to the ordered basis of unit vectors. (Note that if dimV = n, it cannot have more than n characteristic roots)
2. Suppose that a linear transformation T of a vector space V of dimension n has n distinct characteristic roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ in K. If $X_1, X_2, \ldots, X_n$ are characteristic vectors corresponding respectively to these characteristic roots, then $\{X_1, X_2, \ldots, X_n\}$ is a linearly independent set.

Remark 1:

Since $\{X_1, X_2, \ldots, X_n\}$ in Thm4 is a linearly independent set of n vectors and dim V = n, it is a basis of V. the matrix of T relative to the ordered basis $X_1, X_2, \ldots, X_n$ of V is determined from the following equations

$T(X_1) = \lambda_1 X_1$

$T(X_2) = \lambda_2 X_2$

……………………..

$T(X_n) = \lambda_n X_n$

Thus the matrix of T relative to this basis is A = $\begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix}$

Definition: Let A & B be two nxn matrices. A is said to be similar to B if there exists an invertible matrix P such that $P^{-1}AP = B$.

Theorem
If nxn matrices A & B are similar, then they have the same characteristic polynomial and hence the same eigenvalues.

Recal: an nxn matrix D = (d_{ij}) is called a diagonal matrix if d_{ij} =0 whenever i ≠

Definition: Let A be an nxn matrix, then A is diagonalizable if and only if there exists an nxn matrix P such that $P^{-1}AP$ is a diagonal matrix. When such a P exists, we say that P diagonalizes A. j.

Theorem(Condition for diagonalizability).
Let A be an nxn matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $V_1, V_2, \ldots, V_n$ be eigenvectors, with V_j associated with λ_j. Suppose that these eigenvectors are linearly independent and let P be the nxn matrix having V_j as its jth column. Then P is nonsingular and $P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix}$ the diagonal matrix having the eigenvalues of A along its main diagonal, in the same order as the eigenvectors are listed as columns of P.

Examples

1. A = $\begin{pmatrix} -1 & 4 \\ 0 & 3 \end{pmatrix}$ λ= -1, 1; $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ & $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

   Therefore P = $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ $P^{-1}AP = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$ or $v_1 = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ & $v_2 = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$ if we take S = $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, the $S^{-1}AS = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$.

2. A= $\begin{pmatrix} -1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$. Find P such that $P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$

3. A = $\begin{pmatrix} 5 & -4 & 4 \\ 12 & -11 & 12 \\ 4 & -4 & 5 \end{pmatrix}$, observe that λ=1, 1, -3 & we can take 2 vectors for λ=1 since the eigenspace for it has dimension 2, hence P = $\begin{pmatrix} 1 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$

Theorem: Let A be an nxn diagonalizable matrix. Then A has n linearly independent vectors. Further, if $Q^{-1}AQ$ is a diagonal matrix, then the diagonal elements are the eigenvalues of A and the columns of Q are corresponding eigenvectors.

e.g B = $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable. Its eigenvectors are not linearly independent. P-1 does not exists.

Corollary: An nxn matrix with n distinct eigenvalues is diagonalizable.

ORTHOGONAL AND SYMMETRIC MATRICES

Definition:

Let F, G $\in\mathbb{R}^n$. the dot product of F and G denoted F.G is a real number obtained by adding the products of the respective components of these vectors.

Theorem (properties of dot products):

Let F, G, H be vectors and let t be a scalar. Then

a. F.G = G.F
b. (F + G) .H = (F.H) + (G.H)
c. t(F.G) = (tF).G = F.(tG)
d. F.F = $||F||^2$
e. F.F = 0 if and only if F = 0.

Definitions:

1. Two vectors are said to be orthogonal if and only if their dot product is zero.
2. Two vectors U and V are orthonormal if they are orthogonal and each vector is a unit vector ie U.V = 0 and $||V||^2=1$ and $||U||^2=1$
3. A is an orthogonal matrix if and only if $AA^T = A^TA = I_n$

Note

1. If A is orthogonal, then $A^{-1} = A^T$
2. A is orthogonal if and only if AT is orthogonal
3. If A is orthogonal, then |A| = 1 or -1.
4. A real square matrix of order n is orthogonal if and only if the row vectors of A are orthonormal in $\mathbb{R}^n$

   Show that S = {u_1, u_2, u_3} is an orthonormal set. Where $u_1 = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{-1}{\sqrt{5}} \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$,

   $u_3 = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{pmatrix}$

5. Let A be an nxn real matrix. Then A is orthogonal if and only if the column vectors of A are orthonormal in $\mathbb{R}^n$.

Definition

An nxn matrix A is symmetric if and only if $A = A^T$.

Theorems

1. The eigenvalue of a real symmetric matrix are real
2. Let A be a real symmetric matrix, then the eigenvectors associated with distinct eigenvalues are orthogonal.
3. Let A be an nxn real symmetric matrix, then there exists an orthogonal matrix that diagonalizes A.

   Find orthogonal matrix that diagonalizes A = $\begin{pmatrix} 1 & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & 0 \end{pmatrix}$

   Eigen values of A are 2, 2, -1 and get P = $\begin{pmatrix} 1 & 0 & \sqrt{2} \\ 0 & 1 & 0 \\ -\sqrt{2} & 0 & 1 \end{pmatrix}$ which diagonalizes A but not

   orthogonal because the vectors don't have 1 as their norm. Make them to have a norm of 1.
   One advantage of using an orthogonal matrix is that the inverse is easily computed as the transpose.

UNITARY, HERMITIAN AND SKEWHERMITIAN MATRICES

Fact:
Let A be an nxn matrix of complex numbers(some or all of which may be real). Let λ be an eigen value with corresponding eigenvector Z =

Lemma:
If U is a nonsingular nxn complex matrix, then $\overline{U}^{-1} = \overline{U^{-1}}$

Definition:

1. An nxn complex matrix U is unitary if and only if $\overline{U}^{-1}$ = U^t. E.g show that U = $\begin{pmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$ is

   unitary.
2. Complex n-vectors F1, …,Fr form a unitary system if Fi.Fj = 0 for i ≠ j and Fi.Fi=1. If all the components of each Fj are real, then a unitary system is an orthonormal sysem.
3. An nxn complex matrix H is Hermitian if and only if $\overline{H}$ = H^T.
4. An nxn complex matrix S is skew hermitian if and only if $\bar{S}$ = $-S^T$. e.g. S = $\begin{pmatrix} 0 & 8i & 2i \\ 8i & 0 & 4i \\ 2i & 4i & 0 \end{pmatrix}$

Theorem:
1. Let U be an nxn complex matrix. Then U is unitary if and only if its row vectors form a unitary system.

2. Let Z = $\begin{pmatrix} z1 \\ z2 \\ . \\ . \\ . \\ zn \end{pmatrix}$ be a matrix of complex numbers. Then

   a. $\bar{Z}^t H$Z is real if H is hermitian

   b. $\bar{Z}^t H$Z is zero or purely imaginary if H is skewhermitian.

3. The eigen values of a unitary matrix have absolute value 1.
4. The eigenvalues of a hermitianmtrix are real.
5. The eigenvalues of a skew hermitian matrix are zero or purely imaginary.

ORTHOGONAL BASIS, GRAM-SCHMIDT PROCESS, ORTHONORMAL BASIS

Theorem:

Given a basis $\{x_1, \ldots, x_p\}$ for a subspace W of R^n. Define

$v_1 = x_1$

$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$

$$v_3 = x_3 - \frac{x_3 . v_1}{v_1 . v_1} v_1 - \frac{x_3 . v_2}{v_2 . v_2} v_2$$

.

.

$$v_p = x_p - \frac{x_p . v_1}{v_1 . v_1} v_1 - \frac{x_p . v_1}{v_2 . v_2} v_2 - \cdots - \frac{x_p . v_{p-1}}{v_{p-1} . v_{p-1}} v_{p-1}$$

Then $\{v_1, \ldots, v_p\}$ is an orthogonal basis for W.

In addition, Span $\{v_1, \ldots, v_p\}$ = Span $\{x_1, \ldots, x_p\}$. this process is called the Gram-Schmidt process.

Eg

1. the following vectors are linearly independent, construct an orthogonal basis for W by Gram-Schmidt process.

$$x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, x_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

2. Find an orthogonal basis for the column space of the following matrix by Gram-Schmidt process.

$$\begin{pmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \end{pmatrix}$$

3. Find an orthonormal basis of the subspace spanned by the following vectors. $v_1 = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$