Linear Algebra

Murti V. Salapaka
Electrical Engineering Department
University Of Minnesota, Twin Cities
murtis@umn.edu

Preliminary Notation

- Column and row vectors

  ⋆ A column vector x is a $n-$tuple of real or complex numbers

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

  ⋆ A row vector

$$x = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}, \; x_i \in C, \; R$$

- $m \times n$ matrix is the following array

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \ a_{ij} \in C, \ R$$

short hand notation is $A = (a_{ij})$

It is useful to view A as a row vector of columns

$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$, where $a_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}$

View A as a collection of row vectors

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}, a_i = \begin{bmatrix} a_{i1} & \dots & a_{in} \end{bmatrix}$$

- Upper and Lower triangular matrices

  $A = (a_{ij})$ is the upper triangular if $a_{ij} = 0, \; j < i$

- Transpose: A^T denotes the transpose of a matrix A whose elements are $A^T = (a_{ji})$ if $A = (a_{ij})$

- Conjugate Transpose:

  $A^* = (\overline{a_{ji}})$ if $A = (a_{ij})$

- Symmetric:

  A matrix A is Symmetric if $A = A^T$

- Hermitian:

  A matrix A is Hermitian if $A = A^*$

- Multiplication:

$$C = AB,\ A \in R^{m \times p},\ B \in R^{p \times n}$$

$$C_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = (i^{th}\ row\ of\ A)[j^{th}\ column\ of\ B]$$

Fact:

$If\ C = AB,\ then\ C^* = B^* A^*$

Suppose $B \in R^{p \times n}$

$$B = \begin{bmatrix} b_1 & \ldots & b_n \end{bmatrix}$$

$$AB = \begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_n \end{bmatrix}$$

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$$

$$AB = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_m B \end{bmatrix}$$

Orthogonal, Unitary Matrices, Linear Independence

Definition 1.

- *Orthogonal Matrix:*

  A is Orthogonal if $A^T A = AA^T = I$

- *Unitary matrix:*

  A *is unitary if* $A^* A = AA^* = I$

- *Orthogonal Vectors:*

  Given two column vectors $x, \ y$ *with n-element, they are said to be Orthogonal if* $x^* y = 0$*.*
  They are Orthonormal if $x^* y = 0, \ x^* x = 1, \ y^* y = 1$

- *Linear Independence:*

  Given a set of column vectors with element each denoted by $x^1, \ x^2, \ \ldots, \ x^m$*.*

They are said to be independent if

$$\sum_{i=1}^{m} c_i x^i = 0 \Rightarrow c_i = 0,\ c_i \in R$$

. *If* $x^1,\ x^2, \ldots, x^m$ *are not independent then they are said to be dependent.*

Determinants

- *Determinants:*

  Suppose

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \; then \; \det(B) \triangleq ad - bc$$

Suppose $A \in R^{n \times n}$ *then let* A_{ij} *be defined as the* $(n-1) \times (n-1)$ *matrix obtained by deleting the* i^{th} *row and the* j^{th} *column.* A_{ij} *is called the co-factor associated with* a_{ij}*. Determinant of* A *denoted by* $det(A)$ *is defined by*

$$det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij}).$$

Properties of Determinants

- *It can be shown that*

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} det(A_{ij}).$$

- *If any two rows or any two columns of* A *are the same then* $det(A) = 0.$

- *If any two rows (columns) are interchanged then the sign of the determinant changes but the magnitude remains same.*

- *If an row (column) is scaled by* α *then the* det *also gets scaled by* α.

- $deta(A) = det(A^T)$ *and* $det(A^*) = \overline{det(A)}$.

- *If a scalar multiple if a particular row (column) is added to another row (column) then the determinant remains unchanged.*

- $det(AB) = det(A)det(B)$. *(Not easy to prove).*

Matrix Inverse

- *Inverse of a matrix: Suppose* $A \in R^{n \times n}$ *and there exists a matrix* X *such that*
$$AX = XA = I.$$
Then X *is the inverse of* A. *An inverse of a matrix* A *is denoted by* A^{-1}.

If A is invertible then

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix}$$

and if D is invertible then

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BD^{-1}C & 0 \\ C & D \end{bmatrix}.$$

Simultaneous Equations

Consider the following set of n equations in n unknowns $x_1, x_2, \ldots, x_n$.

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & = & b_n \end{array}$$

Another way of representing this set of equations is

$$Ax = b,\ x := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},\ b := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \text{ and } A := (a_{ij})$$

Gaussian Elimination

Theorem 1. *Consider the equation*

$$Ax = b$$

where x *and* b *are vectors in* R^n *and* $A \in R^{n \times n}$. A *and* b *are known and* x *is the solution to be determined. Then* $Ax = b$ *admits a unique solution* x^* *if* $det(A) \neq 0$.

Proof: $Ax = b$ is a notation for

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{1n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & = & b_n \end{array} \tag{1}$$

Note that there is at least one i such that $a_{i1} \neq 0$ (as $det(A) \neq 0$). Without loss of generality assume that $a_{11} \neq 0$. Perform the following operation: multiply the first row by $-a_{i1}/a_{11}$ and add it to the i^{th} row for $i = 2, \ldots, n$. Replace the i^{th} row with this row. This leads us to the following set of equations

$$\begin{array}{ccccccccc} a_{11}^{(1)} x_1 & + & a_{12}^{(1)} x_2 & + & \ldots & + & a_{1n}^{(1)} x_n & = & b_1^{(1)} \\ 0 & + & a_{22}^{(1)} x_2 & + & \ldots & + & a_{1n}^{(1)} x_n & = & b_2^{(1)} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & + & a_{n2}^{(1)} x_2 & + & \ldots & + & a_{nn}^{(1)} x_n & = & b_n^{(1)} \end{array} \quad (2)$$

with the first row unchanged. The above set of equations can be denoted by $A^{(1)}x = b^{(1)}$ where $A^{(1)}) = (a_{ij}^{(1)})$. Note that $det(A) = det(A^{(1)})$ and therefore $det(A^{(1)}) \neq 0$. Using this fact we can assert that there is at least one $i \in \{2, \ldots, n\}$ such that $a_{i2} \neq 0$. Without loss of generality assume that $a_{22}^{(1)} \neq 0$.

Perform the following operation: multiply the second row by $-a_{i2}^{(1)}/a_{22}^{(1)}$ and add it to the i^{th} row for $i = 3, \ldots, n$ to obtain. Replace the i^{th} for rows

$i = 3, \ldots, n$ with the new rows. This leads us to the following set of equations

$$\begin{array}{ccccccccc} a_{11}^{(2)}x_1 & + & a_{12}^{(2)}x_2 & + & \ldots & + & a_{1n}^{(2)}x_n & = & b_1^{(2)} \\ 0 & + & a_{22}^{(2)}x_2 & + & \ldots & + & a_{1n}^{(2)}x_n & = & b_2^{(2)} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & + & 0 & + & \ldots & + & a_{nn}^{(2)}x_n & = & b_n^{(2)} \end{array} \tag{3}$$

with the first two rows same as in (2). These iterations can be continued to obtain

$$\begin{array}{ccccccccc} a_{11}^{*}x_1 & + & a_{12}^{*}x_2 & + & \ldots & + & a_{1n}^{*}x_n & = & b_1^{*} \\ 0 & + & a_{22}^{*}x_2 & + & \ldots & + & a_{1n}^{*}x_n & = & b_2^{*} \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ 0 & + & 0 & + & \ldots & + & a_{nn}^{*}x_n & = & b_n^{*} \end{array} \tag{4}$$

with $a^* \neq 0$. Thus we have the unique solution

$$\begin{array}{lll} x_n & = & b_n^*/a_{nn}^* \\ x_{n-1} & = & \frac{b_{n-1}^* - a_{(n-1)n}^* x_n}{a_{(n-1)(n-1)}^*} \\ \vdots & \vdots & \vdots \\ x_1 & = & \frac{b_1^* - a_{12}^* x_2 - a^* 13 x_3 \ldots a_{1n}^* x_n}{a_{11}^*} \end{array}$$

∎

The method used to obtain the solution of $Ax = b$ in the proof above is called the *Gaussian Elimination* method.

Theorem 2. *If $A \in R^{n \times n}$, then $det(A) \neq 0$, if and only if A^{-1} exists.*

Proof: Let e_i denote a column vector with 1 in the i^{th} position, a zero otherwise.

$$e^i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $e^i_k = \delta_{ik}$. From Theorem 1, $Ax = e^i$ has a unique solution x^i. Let

$$\begin{aligned} X &= \begin{bmatrix} x^1 & x^2 & \dots & x^n \end{bmatrix} \\ AX &= A \begin{bmatrix} x^1 & x^2 & \dots & x^n \end{bmatrix} \\ &= \begin{bmatrix} Ax^1 & Ax^2 & \dots & Ax^n \end{bmatrix} \\ &= \begin{bmatrix} e^1 & e^2 & \dots & e^n \end{bmatrix} \\ &= I \end{aligned}$$

If A^{-1} exists,then $\exists\ X$ such that

$$\begin{aligned} AX &= I \\ \det(AX) &= \det(I) \\ \det(A)\det(X) &= \det(I) = 1 \\ &\Rightarrow \det(A) \neq 0 \end{aligned}$$

Therefore, A^{-1} exists $\Leftrightarrow \det(A) \neq 0$. ∎

Theorem 3. *Suppose A is a $n \times n$ real or complex matrix, then the following are equivalent:*

1. $\det(A) \neq 0$

2. $\exists$ *a matrix* A^{-1} *such that* $A^{-1}A = AA^{-1} = I$

3. $AX = b$ *has a unique solution for every* $b \in R^n$

4. $AX = 0$ *has the only solution* $X = 0$

5. The rows and columns of A *are independent.*

Proof: We have shown that $1 \Leftrightarrow 2$ from Theorem 2. We have also shown $2 \Leftrightarrow 3$ and $3 \Leftrightarrow 4$

To show $4 \Rightarrow 5$: Assume that scalars $c_1, c_2, \ldots c_n$ such that

$$\begin{aligned}
c_1 a_1 + c_2 a_2 + \ldots + c_n a_n &= 0 \\
\begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} &= 0 \\
Ac &= 0 \\
c &= 0 \; (from \; 4) \\
c_i &= 0 \text{ for all } i
\end{aligned}$$

Thus $a_1, a_2 \ldots a_n$ are independent

To show $5 \Rightarrow 4$:

Note that $\det(A) = \det(A^T) \Rightarrow$ rows of A are independent $\Rightarrow$ columns of A^T are independent $\Rightarrow \det(A^T) \neq 0 \Rightarrow \det(A) \neq 0 \Rightarrow 4$ holds which follows from equivalence of 1 and 4 ∎

Eigenvectors and Eigenvalues

Definition 2. *Eigenvectors and Eigenvalues: Given a square matrix $A \in R^{n\times n}$, (or $A \in C^{n\times n}$,) $\lambda \in C$ is an eigenvalue of A if there exists a vector $x \neq 0 \in C^n$ such that*

$$Ax = \lambda x.$$

Such a vector x is called an eigenvector of A associated with eigenvalue λ.

Theorem 4. *Let $A \in R^{n\times n}$. Then the following statements hold.*

1. *λ is an eigenvalue of A if and only if $det(A - \lambda I) = 0$.*

2. *If λ is an eigenvalue of A then λ^m is an eigenvalue of A^m.*

Proof: (1) Suppose λ is an eigenvalue of A. Then from the definition there exists a $x \neq 0$ such that $Ax = \lambda x$ or in other words there exists $x \neq 0$ such that $(A - \lambda I)x = 0$. From Theorem 3 it follows that $det(A - \lambda I) = 0$.

Suppose $det(A - \lambda I) = 0$. Then from Theorem 3 it follows that there exists $x \neq 0$ such that $(A - \lambda I)x = 0$ which implied there exists $x \neq 0$ such that $Ax = \lambda x$. Thus λ is an eigenvalue of A.

This proves (1).

(2) Suppose λ is an eigenvalue of A. Then from the definition there exists a $x \neq 0$ such that $Ax = \lambda x$. Multiplying this equality by A on both sides we have $A^2x = \lambda Ax = \lambda^2 x$. Thus $A^2x = \lambda^2 x$. Repeating this step m times we have $A^m x = \lambda^m x$. This proves the theorem. ∎

Theorem 5. *Let* $p(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \alpha_m\lambda^m$ *and* $p(A) = \alpha_0 I + \alpha_1 A + \ldots + \alpha_m A^m$*.If* λ_0 *is an eigenvalue of* A *then* $p(\lambda_0)$ *is an eigenvalue of* $p(A)$.

Proof: λ_0 is an eigenvalue, then $\exists x \neq 0$ such that $Ax = \lambda x$

$$\begin{aligned} p(A)x &= (\alpha_0 x + \alpha_1 Ax + \alpha_2 A^2 x + \ldots + \alpha_m A^m x) \\ &= (\alpha_0 x + \alpha_1 \lambda x + \alpha_2 \lambda^2 x + \ldots + \alpha_m \lambda^m x) \\ &= (\alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2 + \ldots + \alpha_m \lambda^m) x \\ &= p(\lambda) x \end{aligned}$$

∎

Theorem 6. *Suppose A is an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and A is nonsingular($\det(A) \neq 0) then \lambda_1^{-1}, \lambda_2^{-1}, \ldots \lambda_n^{-1}$ are eigenvalues of A^{-1}.*

Proof: If λ is an eigenvalue, then $\exists x \neq 0$ such that

$$\begin{aligned} Ax &= \lambda x \\ \Rightarrow \quad x &= \lambda A^{-1} x \\ \Rightarrow \quad \tfrac{1}{\lambda} x &= A^{-1} x \end{aligned}$$

Thus λ^{-1} is an eigenvalue of A^{-1} ∎

Theorem 7. *If A is an $n \times n$ matrix then A and A^T have the same eigenvalues. If A is an $n \times n$ matrix with eigenvalue λ then A^* has an eigenvalue $\overline{\lambda}$.*

Proof: Note that $\lambda \in C$ is an eigenvalue of A if and only if

$$\begin{array}{rlcl} & det(A - \lambda I) & = & 0 \\ \Leftrightarrow & \overline{det(A - \lambda I)} & = & 0 \\ \Leftrightarrow & det(\overline{A - \lambda I}) & = & 0 \\ \Leftrightarrow & det(\overline{A - \lambda I})^T & = & 0 \\ \Leftrightarrow & det(A^* - \overline{\lambda} I) & = & 0 \end{array}$$

Theorem 8. *If A is a $n \times n$ matrix then*

$$det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_1) \ldots (\lambda - \lambda_n)$$

where $\lambda_i, \ i = 1, \ldots n$ are eigenvalues of A. Thus $det(A) = \Pi_{i=1}^n \lambda_i$.

Proof: The proof follows from the fact that $det(\lambda I - A)$ is a n^{th} order polynomial and thus will have n roots. From Theorem 4 it follows that λ_i, $i = 1, \ldots, n$ are all roots of polynomial $det(\lambda I - A)$. This proves the theorem. ∎

Theorem 9. *If A is a Hermitian matrix then all its eigenvalues are real.*

Proof: Note that $A = A^*$. If λ is an eigenvalue of A then there exists a vector $x \neq 0$ such that $Ax = \lambda x$.

$$
\begin{array}{lll}
 & Ax & = \lambda x \\
\Rightarrow & x^* A x & = \lambda x^* x \\
\Rightarrow & (x^* A x)^* & = (\lambda x^* x)^* \\
\Rightarrow & x^* A^* x^* & = \overline{\lambda} x^* x \\
\Rightarrow & x^* A x^* & = \overline{\lambda} x^* x \\
\Rightarrow & \frac{x^* A x^*}{x^* x} & = \overline{\lambda} \\
\Rightarrow & \overline{\lambda} & = \lambda
\end{array}
$$

This proves that λ is real.

Definition 3. *A* $n \times n$ *matrix* A *is said to be*

1. *positive definite if* $x^* A x > 0,$ *for all* $x \neq 0.$

2. *positive semi-definite if* $x^* A x \geq 0,$ *for all* $x.$

3. *negative definite if* $x^* A x < 0,$ *for all* $x \neq 0.$

4. *negative semi-definite if* $x^* A x \leq 0,$ *for all* $x \neq 0.$

Theorem 10. *A* $n \times n$ *is a Hermitian matrix. Then*

1. *all its eigenvalues are positive if* A *is positive definite*

2. *all its eigenvalues are non-negative if* A *is positive semi-definite*

3. *all its eigenvalues are negative if A is negative definite*

4. *all its eigenvalues are non-negative if A is negative semi-definite*

Proof: We will prove (1). Let $x \neq 0$ and $Ax = \lambda x$. As A is hermitian λ is real. Note that

$$\begin{array}{rcl} x^*Ax & = & \lambda x^*x \\ \Rightarrow \frac{x^*Ax^*}{x^*x} & = & \lambda \\ \Rightarrow \lambda & > & 0 \end{array}$$

The last step follows as $x^*Ax > 0,$ and $x^*x > 0.$ ∎

General Vector Spaces

Definition 4. *A linear Vector Space is a collection of objects called vectors with two operations,*$" + "$ *and* $"."$ *defined between two vectors and a vector and scalar respectively which satisfy*

1. $x, y \in V \Rightarrow x + y \in V$

2. $(x + y) + z = x + (y + z) \; \forall x, y, z \in V$

3. $x + y = y + x \; \forall x, y \in V$

4. *There is an element* 0 *called the zero vector such that*
$\underbrace{0}_{scalar} .x = \underbrace{0}_{vector} \; \forall x \in V$

5. $1.x = x \; \forall x \in V$

6. $\alpha(\beta.x) = (\alpha\beta).x$ *where* α, β *are scalars and* $x \in V$

7. $(\alpha + \beta).x = \alpha x + \beta x,\ \alpha, \beta$ *are scalars and* $x \in V$

8. $\alpha.(x + y) = \alpha x + \beta y,\ \alpha$ *is a scalar and* $x, y \in V$

Example 1.

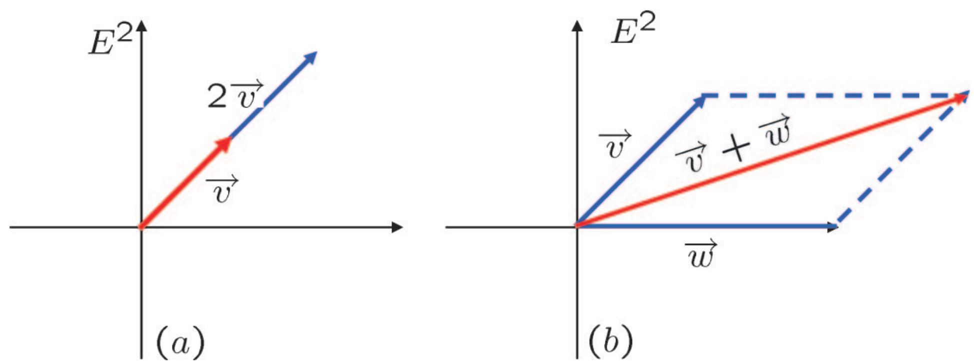

Figure 1: (a) Scalar multiplication (b) vector addition.

Example 2. *Let scalars be the real numbers and* $V = R^n$*.*

$$\{x : x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad x_i \in R, \} \qquad '+' : \; V \times V \to V$$

$$x + y := \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$x_i \in R, \; y_i \in R, \;\; \alpha . x = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

$$(x+y)+z = \begin{bmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1+y_1+z_1 \\ \vdots \\ x_n+y_n+z_1 \end{bmatrix}$$

similarly, $x+(y+z) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1+z_1 \\ \vdots \\ y_n+z_n \end{bmatrix} = \begin{bmatrix} x_1+y_1+z_1 \\ \vdots \\ x_n+y_n+z_1 \end{bmatrix}$

Example 3. *Let*

$$V := \{\textit{set of all polynomials of order less than or equal to } n\}$$

and the scalars be the real numbers. The vector addition operation is defined as follows: if

$$p(t) = p_0 + p_1 t + \ldots + p_n t^n \textit{ and } q(t) = q_0 + q_1 t + \ldots + q_n t^n$$

then

$$(p+q)(t) := (p_0+q_0) + (p_1+q_1)t + \ldots + (p_n+q_n)t^n \textit{ and}$$

$$(\alpha p)(t) = \alpha p_0 + \alpha p_1 t + \ldots + \alpha p_n t^n.$$

Then V *with the* R *as the scalars satisfies all the properties of a vector space.*

Definition 5. Linear Independence: *Let V be a vector space and let $v_1, v_2, \ldots, v_n$ be vectors in V. If $\sum_{i=1}^{n} c_i v_i \Rightarrow c_i = 0$ where $c_1, c_2, \ldots, c_n$ are scalars, then we say $v_1, v_2, \ldots, v_n$ are independent.*

Definition 6. Linear Combination: *Suppose V is a vector space and $v_1, v_2, \ldots, v_n$ are any vectors in V. Then $V = \sum_{i=1}^{n} c_i v_i$ is said to be a Linear Combination of the vectors $v_1, v_2, \ldots, v_n$.*

Definition 7. Subspace: *Suppose X is a vector space. If $V \subset X$ and V is a vector space, then V is said to be a Subspace of X.*

Example 4.

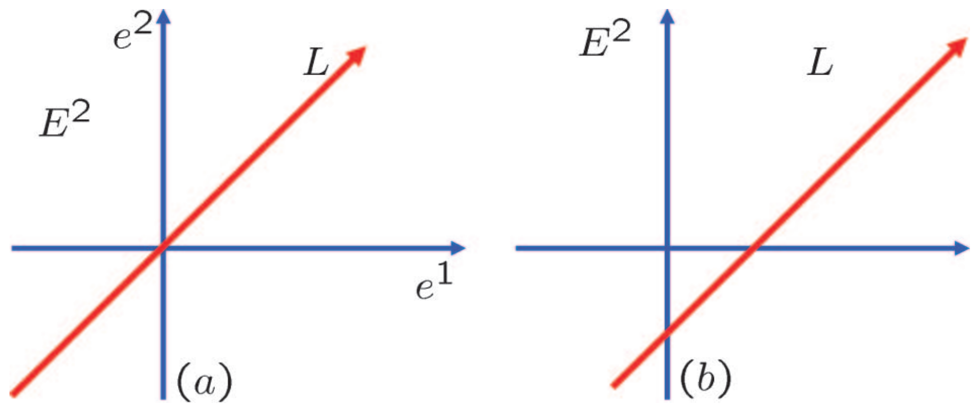

Figure 2: (a) L is a subspace (b) L is not a subspace.

Definition 8. Span: *Let X be a vector space and let $x_1, x_2, x_3, \ldots, x_n$ be vectors in X. Span$(x_1, \ldots, x_n)$ is the set of all linear combination of vectors $x_1, x_2, \ldots, x_n$.*

$$\text{Span}(x_1, \ldots, x_n) = \{x \in X, \quad X = \sum_{i=1}^{n} c_i x_i \text{ , where } c_i \text{ are scalars } \}$$

Example 5.

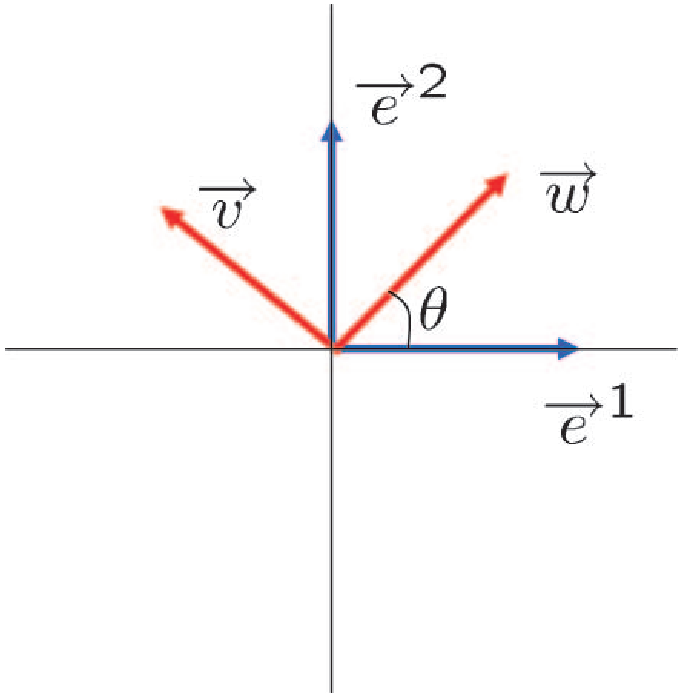

Figure 3: Span$(e^1, e^2) = E^2$, *and Span*$(v, w) = E^2$

Definition 9. Basis: *Let X be a vector space. Then a set of independent vectors $x_1, x_2, \ldots, x_n$ are said to be a Basis if Span$(x_1, \ldots, x_n) = X$.*

Example 6. • e_1, e_2 *is a Basis for E^2.*

- X = *all polynomials of degree* $\leq n$.

  $X = \{\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_n t^n | \alpha_i \in R\}$

  $\{1, t, \ldots, t^n\}$ *forms a Basis for* X.

  $\{1, 1+t, 1+t+t^2, \ldots, 1+t, \ldots, +t^n\}$ *also forms a Basis.*

- $X = \{$*polynomials or order* $\leq 3\}$

  $\{1, 1+t, t^2, t, t^3\}$ *is not a Basis (note that* $1 - (1+t) + t = 0$ *and therefore not independent).*

Definition 10. Finite Dimensional Vector Space: X *a Vector Space is said to be finite dimensional if it has a Basis which has a finite number of elements. Any Vector Space that is not finite dimensional is said to be Infinite Dimensional Vector Space.*

Example: $X = \{$ all polynomials of any degree $\}$ (infinite dimensional)

Dimension is unique

Theorem 11. *Let X be a Vector Space. Suppose $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ are two set of basis Vectors for X, then $n = m$.*

Proof: Assume without loss of generality that $m > n$. As $\{x_j\}j = 1^n$ is a basis there exist constants $a_{ji}, \, i = 1, \ldots, m$ such that

$$y_i = \sum_{j=1}^{n} a_{ji} x_j.$$

Let

$$A := \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & & & \\ a_{j1} & a_{j2} & \dots & a_{jm} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nm} \\ \vdots & & & \\ 0 & 0 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & 0 \end{bmatrix}_{m \times m}$$

$det(A) = 0$. It follows from Theorem 3 that $\exists$ an $\alpha \in R^m, \;\; \alpha \neq 0$ such that $A\alpha = 0$.

Let us consider the linear combination

$$\sum_{i=1}^{m} \alpha_i y_i, \quad \alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} \in R^m \begin{aligned} &= \sum_{i=1}^{m} \alpha_i (\sum_{j=1}^{n} a_{ji} x_j) \\ &= \sum_{j=1}^{n} (\sum_{i=1}^{m} a_{ji} \alpha_i) x_j \\ &= 0 \end{aligned}$$

Thus we have shown that there exist scalars $\alpha_1, \ldots, \alpha_m$ and $\sum_{i=1}^{m} \alpha_i y_i = 0 \Rightarrow$ $y_1, y_2, \ldots, y_m$ are not linearly independent. This is a contradiction to the fact that $\{y_i\}_{i=1}^{m}$ are independent and thus $m = n$. ∎

Definition 11. Dimension of a Finite Dimensional Vector Space: *The Dimension of a Finite Dimensional Vector Space is the number of vectors in any basis of the vector space.*

Suppose V_1, V_2 *are subspaces of a Vector space* V, *then*

$$V_1 \cap V_2 = \{v \in V : v \in V_1 \; and \; v \in V_2\}$$

$$V_1 + V_2 = \{v \in V : v = v_1 + v_2, \; v_1 \in V_1, \; v_2 \in V_2\}$$

$V_1 + V_2$ *is called the* Direct Sum *of* V_1 *and* V_2 *if* $V_1 \cap V_2 = \{0\}$. *The notation* $V_1 \oplus V_2$ *is used to denote a Direct Sum.*

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$$

$$\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2).$$

Independent vectors extended to form a basis

Theorem 12. *Let V be a n dimensional vector space and let $v_i,\ i = 1, \ldots, m$ be independent vectors with $m < n$. Then there exist n independent vectors $\hat{v}_i,\ i = 1, \ldots, n$ such that $\hat{v}_i = v_i$ for $i = 1, \ldots, m$.*

Proof: Let

$$V_0 := span\{v_1, \ldots, v_2\}.$$

Let $\hat{v}_{m+1} \in V$ such that $\hat{v}_{m+1} \notin V_0$. Such a vector exists from Theorem 11 and as $m < n$. Let

$$V_1 := span\{v_1, \ldots, v_2, \hat{v}_{m+1}\}.$$

Clearly $dim(V_2) = m + 1$. Continuing the above process till we obtain

$$V_(n - m) := span\{v_1, \ldots, v_2, \hat{v}_{m+1}, \ldots, \hat{v}_n\}.$$

From Theorem 11 these set of vectors has to form a basis for V. The theorem follows by defining $\hat{v}_i := v_i$ for $i = 1, \ldots, m$. ∎

Coordinates

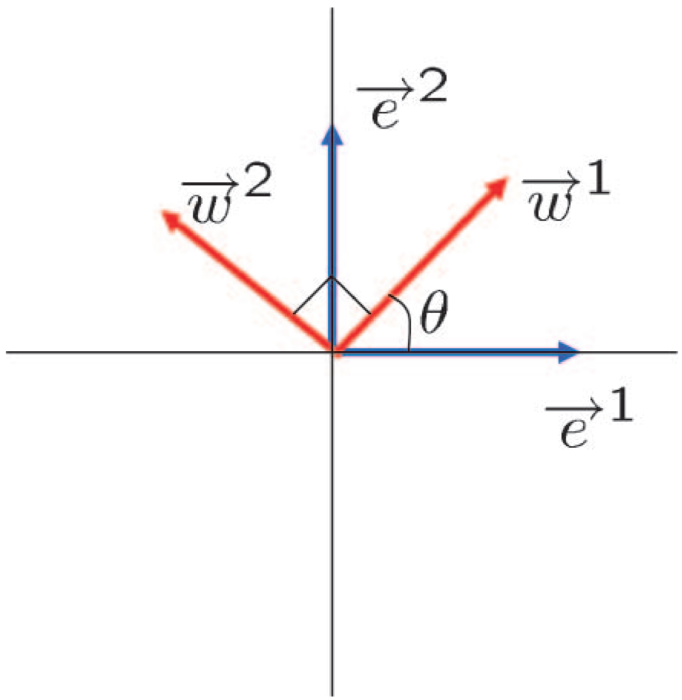

Figure 4: The coordinates of w^1 in the basis e^1 and e^2 is $[\cos\theta \quad \sin\theta]'$

Note that

$$\overrightarrow{w}^1 = \cos(\theta)\overrightarrow{e}^1 + \sin(\theta)\overrightarrow{e}^2.$$

We say

$$\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

are the coordinates of $\overrightarrow{w}^1$ in the bases $\overrightarrow{e}^1$ and $\overrightarrow{e}^2$.

Note that $\overrightarrow{w}^1$ and $\overrightarrow{w}^2$ also forms a basis for E^2. As

$$\overrightarrow{w}^1 = 1\overrightarrow{w}^1 + 0\overrightarrow{w}^2.$$

And thus the coordinates of $\overrightarrow{w}^1$ in the basis $\overrightarrow{w}^1$ and $\overrightarrow{w}^2$ is

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Let V be a vector space and suppose $\{v_i\}_{i=1}^n$ be a set of basis vectors. Then any vector $v \in V$ can be written as

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Then $\alpha_1, \ldots, \alpha_n$ are coordinates in the basis $\{v_i\}_{i=1}^n$ and

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

is the coordinate vector in the basis $\{v_i\}_{i=1}^n$. Note that if

$$v = \sum_{i=1}^{n} \hat{\alpha}_i v_i$$

then

$$0 = \sum_{i=1}^{n} (\hat{\alpha}_i - \alpha_i) v_i$$

and as $v_1, \ldots, v_n$ are independent it follows that $(\hat{\alpha}_i - \alpha_i) = 0$ for $i = 1, \ldots, n$. Thus $\alpha_i = \hat{\alpha}_i$ for all $i = 1, \ldots, n$. This implies that coordinates are well defined.

Example 7. *Suppose*

$$V = \{\textit{all polynomials with degree less than or equal to } n\}.$$

Note that the polynomials $1, t, t^2, \ldots, t^n$ *forms a basis for* V. *Suppose* p *is a polynomial given by*

$$p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2, \ldots, \alpha_n t^n.$$

The coordinate vector of p in the basis $\{t^i\}_{i=0}^n$ is

$$\begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

One can check that

$$\{1, 1+t, 1+t+t^2, \ldots, 1+t+t^2+\ldots+t^n\},$$

is also a set of basis vectors for V. Note that

$$\begin{aligned} p(t) &= \alpha_0 + \alpha_1 t + \alpha_2 t^2, \ldots, \alpha_n t^n \\ &= \alpha_0 + \alpha_1(1+t-1) + \alpha_2(1+t+t^2-(1+t)) + \ldots \\ &\quad +\alpha_n(1+t+\ldots+t^n-(1+t+\ldots+t^{n-1})) \\ &= (\alpha_0-\alpha_1) + (\alpha_1-\alpha_2)(1+t) + \\ &\quad \ldots + (\alpha_{n-1}-\alpha_n)(1+t+\ldots+t^{n-1}) + \alpha_n(1+t+t^2+\ldots+t^n) \end{aligned}.$$

Thus in new basis the coordinate vector is

$$\begin{bmatrix} \alpha_0 - \alpha_1 \\ \alpha_1 - \alpha_2 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{bmatrix}.$$

Linear Operator

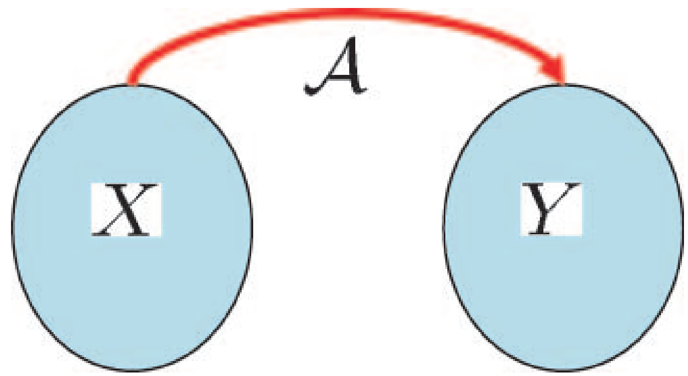

Figure 5: A map

Let X and Y be vector spaces. $\mathcal{A}$ a mapping from X to Y which assigns a vector $\mathcal{A}x \in Y$ for every vector $x \in X$ is a linear operator if

$$\mathcal{A}(\alpha_1 x^1 + \alpha_2 x^2) = \alpha_1 \mathcal{A} x_1 + \alpha_2 \mathcal{A} x_2 \text{ for all } x^1, x^2 \in X \text{ and } \alpha_1, \alpha_2 \text{ scalars.}$$

Example 8. *Suppose V is the set of all polynomials of degree less than or equal to n. Suppose W is the set of all polynomials of degree less than or* *equal to* $n-1$. *Let* $\mathcal{A}: V \to W$ *be defined by*

$$\mathcal{A}p := \frac{dp}{dt}.$$

Note that for every $p \in V$, $\mathcal{A}p \in W$. *Also note that*

$$\begin{aligned} \mathcal{A}(\alpha p + \beta q) &= \tfrac{d(\alpha p+\beta q)}{dt} \\ &= \alpha\tfrac{dp}{dt} + \beta\tfrac{dq}{dt} \\ &= \alpha\mathcal{A}p + \beta\mathcal{A}q \end{aligned}$$

proving that $\mathcal{A}$ *is a linear operator.*

Example 9. *Suppose* $V = R^n$ *and* $W = R^m$. *Suppose* $\mathcal{A}: V \to W$ *id defined* *by*

$$\mathcal{A}x := \overbrace{\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{mn} \end{bmatrix}}^{=:A} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

where $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. *Its evident that*

$$\begin{aligned} \mathcal{A}(\alpha_1 x^1 + \alpha_2 x^2) = A(\alpha_1 x^1 + \alpha_2 x^2) &= \alpha_1 A x_1 + \alpha_2 A x^2 \\ &= \alpha_1 \mathcal{A} x^1 + \alpha_2 \mathcal{A} x^2 \end{aligned}$$

Thus $\mathcal{A}$ *is linear.*

Matrix Representation of a Linear Operator

Suppose $\mathcal{A} : V \to W$ is a linear operator. Suppose $\{v_1, \ldots, v_n\}$ is a basis for V and $\{w_1, \ldots, w_2\}$ is a basis for W. Suppose $v \in V$ and suppose $\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$ is its coordinate vector in the basis given. That is

$$v = \sum_{j=1}^{n} \alpha_j v_j.$$

Note that $\mathcal{A}v_j \in W$. Let the coordinate vector of $\mathcal{A}v_j$ be $\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$ for $j = 1, \ldots, n$. That is

$$\mathcal{A}v_j = \sum_{i=1}^{m} a_{ij} w_i, \ j = 1, \ldots, n.$$

Note that

$$\begin{aligned}
\mathcal{A}v &= \mathcal{A}(\textstyle\sum_{j=1}^{n} \alpha_j v_j) \\
&= \textstyle\sum_{j=1}^{n} \mathcal{A}(\alpha_j v_j) \\
&= \textstyle\sum_{j=1}^{n} \alpha_j \mathcal{A}(v_j) \\
&= \textstyle\sum_{j=1}^{n} \sum_{i=1}^{m} \alpha_j a_{ij} w_i \\
&= \textstyle\sum_{i=1}^{m} \underbrace{\left(\displaystyle\sum_{j=1}^{n} a_{ij}\alpha_j\right)}_{:=\beta_i} w_i \\
&=: \textstyle\sum_{i=1}^{m} \beta_i w_i
\end{aligned}$$

where we have defined $\beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j, i = 1, \dots m$. Thus the coordinate vector of $\mathcal{A}v$ is

$$\beta := \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix},$$

where

$$\beta = A\alpha,$$

with $A = (a_{ij})$.

Thus the method to obtain the matrix representation of a linear operator *given a basis* $\{v_j\}_{j=1}^n$ *of the domain space* V *and a basis* $\{w_i\}_{i=1}^m$ *of the range space* W is to follow the steps below:

1. Obtain the coordinates of $\mathcal{A}v_j$ in the basis $\{w_i\}_{i=1}^m$. Let $\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$ be the coordinate vector for v_j.

2. The coordinate vector of $\mathcal{A}v$ is $\beta = A\alpha$ if α is the coordinate vector of v in the basis $\{v_j\}_{j=1}^n$.

Example 10. *Consider*

$$V = \{ \textit{ all polynomials of degree } \leq n\}$$

and

$$W = \{ \text{ all polynomials of degree } \leq n-1\}.$$

Let $\mathcal{A} : V \to W$ *be defined by*

$$\mathcal{A}p = \frac{dp}{dt}.$$

Let $(1, t, t^2, \ldots, t^n)$ *be the basis for* V *and let* $(1, t, \ldots, t^{n-1})$ *be the basis for* W*.*

$$\mathcal{A}v_j = \sum_{i=1}^{m} a_{ij} w_i.$$

Note that $v_j = t^{j-1}$*,* $w_i = t^{i-1}$ *Thus*

$$\begin{aligned} \mathcal{A}v_j &= \tfrac{dv_j}{dt} \\ &= (j-1)t^{j-2} \\ &= \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} a_{ij} t^{i-1} \end{aligned}$$

This implies that

$$(j-1)t^{j-2} = \sum_{i=1}^{m} a_{ij} t^{i-1}$$

and thus

$$\begin{array}{rcll} a_{ij} & = & 0 & if \;\; i \neq (j-1) \\ & = & (j-1) & if \;\; i = (j-1) \end{array}$$

Let p *in* V *be given by*

$$p = \alpha_0 1 + \alpha_1 t + \ldots + \alpha_n t^n$$

which has coordinate vector

$$\begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Then $\mathcal{A}v$ has coordinates $\beta = A\alpha$ where $A = (a_{ij})$ where

$$\begin{aligned} a_{ij} &= 0 \qquad\qquad if \;\; i \neq (j-1) \\ &= (j-1) \;\; if \;\; i = (j-1) \end{aligned}$$

Example 11. $V = R^n$, $W = R^m$

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

to be the basis for R^n and a similar basis for R^m.

Let $\mathcal{A} : R^n \to R^m$ be defined by

$$\mathcal{A}v = \overbrace{\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}}^{v}$$

$\beta = A\alpha$.

Composition of Linear Operators

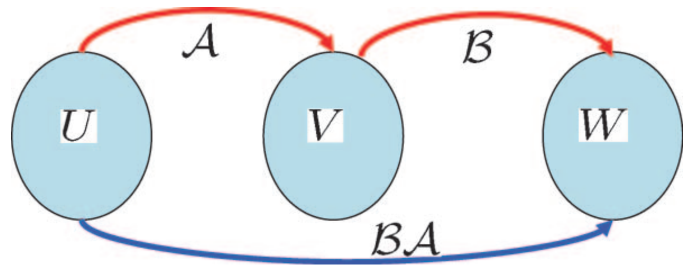

Figure 6: Composition of two operators

Theorem 13. *Suppose U, V and W are vector spaces with bases $\{u_1, \ldots, u_n\}$, $\{v_1, \ldots, v_m\}$ and $\{w_1, \ldots, w_q\}$ respectively. $\mathcal{A} : U \to V$ and $\mathcal{B} : V \to W$ are linear operators with matrix representations A and B respectively in the bases given. Then the matrix representation of the linear operator $\mathcal{BA} : U \to W$ has a matrix representation BA with $\{u_1, \ldots, u_n\}$, and $\{w_1, \ldots, w_q\}$ as bases for U and W respectively.*

Change of basis

$\mathcal{A} : V \to W$ is a linear operator, then the matrix representation of $\mathcal{A}$ depends on the basis of V and W.

Example 12. *$V = R^3$, $W = R^3$ and $\mathcal{A} : V \to W$ is defined by $\mathcal{A}v = Av$ where $A = (a_{ij})$.*

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$w_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ w_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \ w_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

$$\mathcal{A}v_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} = \sum_{i=1}^{3} \alpha_{i1} w_i = \begin{bmatrix} \alpha_{11} \\ 2\alpha_{21} \\ 3\alpha_{31} \end{bmatrix}$$

. Thus the coordinate vector of $\mathcal{A}v_1$ *is given by*

$$\begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix} = \begin{bmatrix} a_{11} \\ \frac{1}{2}a_{21} \\ \frac{1}{3}a_{31} \end{bmatrix}$$

$$\mathcal{A}v_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = \sum_{i=1}^{3} \alpha_{i2} w_i = \begin{bmatrix} \alpha_{12} \\ 2\alpha_{22} \\ 3\alpha_{32} \end{bmatrix}$$

Thus the coordinate vector of $\mathcal{A}v_2$ *is given by*

$$\begin{bmatrix} \alpha_{12} \\ \alpha_{22} \\ \alpha_{32} \end{bmatrix} = \begin{bmatrix} a_{12} \\ \frac{1}{2}a_{22} \\ \frac{1}{3}a_{32} \end{bmatrix}$$

$$\mathcal{A}v_3 = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = \sum_{i=1}^{3} \alpha_{i3} w_i = \begin{bmatrix} \alpha_{13} \\ 2\alpha_{23} \\ 3\alpha_{33} \end{bmatrix}$$

Thus the coordinate vector of $\mathcal{A}v_3$ is given by

$$\begin{bmatrix} \alpha_{13} \\ \alpha_{23} \\ \alpha_{33} \end{bmatrix} = \begin{bmatrix} a_{13} \\ \frac{1}{2}a_{23} \\ \frac{1}{3}a_{33} \end{bmatrix}$$

Matrix Representation of $\mathcal{A}$ is given by

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \frac{1}{2}a_{21} & \frac{1}{2}a_{22} & \frac{1}{2}a_{23} \\ \frac{1}{3}a_{31} & \frac{1}{3}a_{32} & \frac{1}{3}a_{33} \end{bmatrix}$$

Suppose V is a vector space with two sets of basis vectors given by $\{v_i\}_{i=1}^{n}$ and $\{\hat{v}_i\}_{i=1}^{n}$. Suppose the coordinate vector of a vector $v \in V$ in the bases $\{v_i\}_{i=1}^n$ and $\{\hat{v}_i\}_{i=1}^n$ is given by

$$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \text{ and } \hat{\alpha} = \begin{bmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_n \end{bmatrix}$$

respectively. Suppose

$$\hat{v}_j = \sum_{i=1}^n q_{ij} v_i$$

Note that

$$\begin{aligned} v &= \textstyle\sum_{j=1}^n \hat{\alpha}_j \hat{v}_j \\ &= \textstyle\sum_{j=1}^n \hat{\alpha}_j \sum_{i=1}^n q_{ij} v_i \\ &= \textstyle\sum_{i=1}^n (\sum_{j=1}^n q_{ij} \hat{\alpha}_j) v_i \end{aligned}$$

Therefore we have

$$\alpha = Q\hat{\alpha} \text{ where } Q = (q_{ij}).$$

Lemma 1. *The matrix Q above is invertible.*

Proof: Suppose Q is not invertible. Then from Theorem 3 it follows that there exists $\hat{\alpha} \neq 0$ such that $Q\hat{\alpha} = 0$. This implies that

$$\sum_{j=1}^{n} q_{ij}\hat{\alpha}_j = 0 \text{ for all } i = 1, \ldots, n.$$

Consider

$$\begin{aligned} \textstyle\sum_{j=1}^{n} \hat{\alpha}_j \hat{v}_j &= \textstyle\sum_{j=1}^{n} \hat{\alpha}_j (\sum_{i=1}^{n} q_{ij} v_i) \\ &= \textstyle\sum_{i=1}^{n} \underbrace{(\displaystyle\sum_{j=1}^{n} q_{ij}\hat{\alpha}_j)}_{=0} v_i \\ &= 0. \end{aligned}$$

This implies there exists $\hat{\alpha} \neq 0$ such that $\sum_{j=1}^{n} \hat{\alpha}_j \hat{v}_j = 0$. This would imply that $\{\hat{v}_j\}$ is not an independent set. This is a contradiction. ∎

Example 13. $V = R^3$, *with basis* $\overbrace{(e_1, e_2, e_3)}^{v_1, v_2, v_3}$ *and* $\overbrace{(e_1, \frac{1}{2}e_2, \frac{1}{3}e_3)}^{\hat{v}_1, \hat{v}_2, \hat{v}_3}$

$$\begin{aligned} \hat{v}_1 &= (1)v_1 + (0)v_2 + (0)v_3 \\ \hat{v}_2 &= (0)v_1 + (\tfrac{1}{2})v_2 + (0)v_3 \\ \hat{v}_3 &= (0)v_1 + (0)v_2 + (\tfrac{1}{3})v_3 \\ Q &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix} \end{aligned}$$

If α *is the coordinate vector in* (e_1, e_2, e_3), *then the coordinate vector in* $(e_1, \frac{1}{2}e_2, \frac{1}{3}e_3)$ *is* $Q^{-1}\alpha$.

Theorem 14. *Suppose* $\mathcal{A} : V \to W$ *is a linear operator from vector space* V *to vector space* w. *Furthermore, suppose* $(v_1, v_2, \ldots, v_n)$, $(\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n)$ *forms two sets of basis for* V *with the associated change of basis matrix* Q. *Also,* *suppose* $(w_1, \ldots, w_m)$ *and* $(\hat{w}_1, \ldots, \hat{w}_m)$ *form basis for* W *with change of basis matrix* T*. Let* A *be the matrix representation of* $\mathcal{A}$ *in the basis* $(v_1, \ldots, v_n)$ *for* V *and* $(w_1, \ldots, w_m)$ *for* W*. Let* B *be the matrix representation of* $\mathcal{A}$ *in the basis* $(\hat{v}_1, \ldots, \hat{v}_n)$ *for* V *and* $(\hat{w}_1, \ldots, \hat{w}_m)$ *for* W*. Then,* $B = PAQ$*,* $P = T^{-1}$*.*

Proof: Suppose α is the coordinate vector of $v \in V$ in the basis $(v_1, \ldots, v_n)$. Let $\hat{\alpha}$ be the coordinate vector in the basis $(\hat{v}_1, \ldots, \hat{v}_n)$. Then

$$\alpha = Q\hat{\alpha}.$$

Suppose β is the coordinate vector of $\mathcal{A}v$ in the basis $(w_1, w_2, \ldots, w_m)$. Then

$$\beta = A\alpha.$$

Suppose $\hat{\beta}$ is the coordinate vector of $\mathcal{A}v$ in the basis $(\hat{w}_1, \ldots, \hat{w}_m)$. Then

$$\beta = T\hat{\beta}.$$

$\beta = T\hat{\beta} \Rightarrow \hat{\beta} = T^{-1}\beta = T^{-1}A\alpha = \underbrace{T^{-1}AQ}_{B}\,\hat{\alpha}$. Therefore, $B = T^{-1}AQ$. ∎

Example: $C : \overbrace{R^3}^{V} \rightarrow \overbrace{R^3}^{W}$

$$
\begin{aligned}
C &= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \\
\Rightarrow Cv &= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_3 \end{bmatrix} \; where \; v = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_3 \end{bmatrix}
\end{aligned}
$$

Let (e_1, e_2, e_3) be a basis for V and W. Then we have argued earlier that the matrix representation in these basis vectors is

$$A = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

Let another set of basis vector for V and W be (e_1, e_2, e_3) and $(e_1, \frac{1}{2}e_2, \frac{1}{3}e_3)$.

$B = T^{-1}AQ$.

From the previous example, we have

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}, Q = I.$$

$$B = T^{-1}A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ 2c_{21} & 2c_{22} & 2c_{23} \\ 3c_{31} & 3c_{32} & 3c_{33} \end{bmatrix}.$$

Equivalence and Similarity Transformations

Definition 12. • **Equivalence Transformation:** *If A and B are $m \times n$ matrices and P and Q are nonsingular $m \times m$ and $n \times n$ matrices respectively. Then A and B are equivalent if $B = PAQ$. It immediately follows that if A and B are two matrix representation of a linear operator $\mathcal{A} : V \to W$ then A and B are equivalent.*

• **Similarity Transformation:** *If A and B are $m \times m$ matrices, $Q \in R^{m \times m}$ is invertible, then A and B are similar if $B = Q^{-1}AQ$.*

Theorem 15. *If $\mathcal{A} : V \to V$ be a linear operator with a matrix representation A in the basis $(v_1, \ldots, v_n)$ and B in the basis $(\hat{v}_1, \ldots, \hat{v}_n)$. Then A and B are similar.*

Proof: We know from Theorem 14 that $B = T^{-1}AQ$. T is the basis transformation between $(w_1, \ldots, w_n) \to (\hat{w}_1, \ldots, \hat{w}_n)$. $T = Q \Rightarrow B = Q^{-1}AQ$. ∎

Definition 13. Range of a Linear Operator $\mathcal{A}$: *Let $\mathcal{A}$ be a linear operator from vector space V to vector space W.*

Range$(\mathcal{A})$ = $\{w \in W$ *such that* $\exists v \in V$ *with* $\mathcal{A}v = w\}$

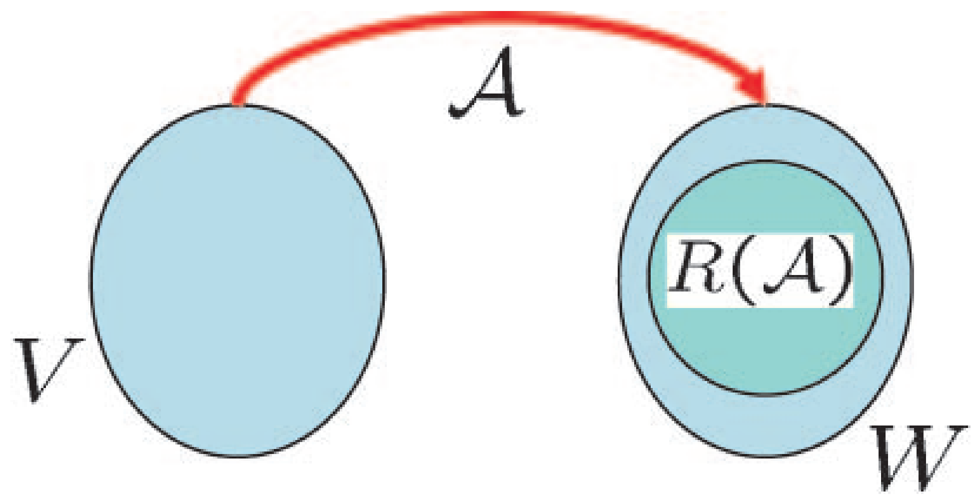

Figure 7: Range of a operator

Range$(\mathcal{A}) \subset W$.

Theorem 16. *If $(v_1, \ldots, v_n)$ is a basis for a vector space and $\mathcal{A} : V \to W$ where W is a vector space with $\mathcal{A}$ is linear, then* *span*$(\mathcal{A}v_1, \mathcal{A}v_2, \ldots, , \mathcal{A}v_n) = Range(\mathcal{A})$

Proof: To prove that $Range(\mathcal{A}) \subset span\{\mathcal{A}v_1, \mathcal{A}v_2, \ldots, , \mathcal{A}v_n\}$

Let $w \in Range(\mathcal{A})$

From definition, it follows that $\exists v \in V$ such that $w = \mathcal{A}v$

$v \in V \Rightarrow \exists(\alpha_1, \alpha_2, \ldots, \alpha_n)$ such that $v = \sum_{i=1}^{n} \alpha_i v_i$

$$\begin{aligned}
w = \mathcal{A}v &= \mathcal{A}(\sum_{i=1}^{n} \alpha_i v_i) \\
&= \sum{}_i = 1^n \alpha_i (Av_i) \\
&\Rightarrow w \in span(\{\mathcal{A}v_1, \mathcal{A}v_2, \ldots, , \mathcal{A}v_n\} \\
&\Rightarrow Range(\mathcal{A} \subset span\{\mathcal{A}v_1, \mathcal{A}v_2, \ldots, , \mathcal{A}v_n\}
\end{aligned}$$

Suppose we have $w \in span\{\mathcal{A}v_1, \mathcal{A}v_2, \ldots, , \mathcal{A}v_n\}$. Then

$$\exists(\beta_1, \ldots, \beta_n) \text{ such that } w = \sum_{i=1}^{n} \beta_i \mathcal{A}v_i = \mathcal{A}(\sum_{i=1}^{n} \beta_i v_i) = \mathcal{A}v$$

where $v \in V$

$w \in Range(\mathcal{A})$

$span(\mathcal{A}v_1, \ldots, \mathcal{A}v_1) \subset Range(\mathcal{A})$

Therefore,

$$span(\mathcal{A}v_1, \ldots, \mathcal{A}v_1) = Range(\mathcal{A})$$

We can show that $Range(\mathcal{A})$ is a vector space. ∎

Definition 14. $Rank(\mathcal{A})$**:** *Suppose $\mathcal{A}$ is a linear operator from vector space V to vector space W. Then $Rank(\mathcal{A}) = dim(Range(\mathcal{A}))$.*

Example 14.

$$V = \{\textit{set of polynomials of order } \leq 2\} \textit{ and } W \equiv V$$

$\mathcal{A} : V \rightarrow W$ *be the operator defined by*

$$\mathcal{A}V = \frac{dv}{dt}.$$

Note that

$$Range(\mathcal{A}) = \{\textit{all polynomials with degree } \leq 1\} \textit{ and}$$

$$Rank(\mathcal{A}) = dim\{Range(\mathcal{A})\} = 2.$$

$1, t, t^2$ *forms a basis for V*

$$
\begin{aligned}
Range(\mathcal{A}) &= span\{\mathcal{A}(1), \mathcal{A}(t), \mathcal{A}(t^2)\} \\
&= span\{0, 1, 2t\} \\
&= span\{1, 2t\}
\end{aligned}
$$

Example 15. *$Rank(\mathcal{A})$: Suppose A is a $m \times n$ matrix, then $Rank(\mathcal{A})$ = number of independent columns of A.*

Theorem 17. *Suppose $\mathcal{A} : V \to W$ is a linear operator and A is the matrix representation of $\mathcal{A}$ in the basis $(v_1, \ldots, v_n)$ for V and $(w_1, \ldots, w_n)$ for W. Then $Rank(A) = Rank(\mathcal{A})$.*

Proof: Suppose that $dim(Range(\mathcal{A})) = r = rank(\mathcal{A})$. Then, there should be r independent vectors $\mathcal{A}v_1, \mathcal{A}v_2, \ldots, \mathcal{A}v_n$ which follows from Theorem 16.

Let us assume without loss of generality that only $\mathcal{A}v_1, \mathcal{A}v_2, \ldots, \mathcal{A}v_r$ are independent.

The matrix A was defined by the following

$$\mathcal{A}v_j = \sum_{i=1}^{m} a_{ij} w_i.$$

Consider a linear combination of the first r columns of A

$$c_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + c_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + c_3 \begin{bmatrix} a_{1r} \\ a_{2r} \\ \vdots \\ a_{mr} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{r} a_{1j} c_j \\ \sum_{j=1}^{r} a_{2j} c_j \\ \vdots \\ \sum_{j=1}^{r} a_{mj} c_j \end{bmatrix}$$

Suppose $\exists\ c_1, c_2, \ldots, c_r$ such that

$$c_1 a_1 + c_2 a_2 + \ldots + c_r a_r = 0.$$

that is

$$\sum_{j=1}^{r} a_{ij} c_j = 0; \quad i = 1, 2, \ldots, m$$

Consider the linear combination

$$\begin{aligned} c_1\mathcal{A}v_1 + c_2\mathcal{A}v_2 + \ldots + c_1\mathcal{A}v_r &= c_1\sum_{i=1}^{m} a_{i1}w_i + c_2\sum_{i=1}^{m} a_{i2}w_i + \ldots + c_r\sum_{i=1}^{m} a_{ir}w_i \\ &= \sum_{j=1}^{r} c_j \sum_{i=1}^{m} a_{ij}w_i \\ &= \sum_{i=1}^{m} (\sum_{j=1}^{r} c_j a_{i1})w_i \\ &= 0 \end{aligned}$$

Because $\mathcal{A}v_1, \mathcal{A}v_2, \ldots, \mathcal{A}v_r$ are independent, it follows that $c_j = 0, \quad j = 1, 2, \ldots, r$.

In summary, if $c_1a_1 + c_2a_2 + \ldots + c_ra_r = 0$ then $c_j = 0 \; \forall j = 1, 2, \ldots, r$.

We have shown that $a_1, a_2, \ldots, a_r$ are independent.

Therefore, $Rank(A) \geq r = Rank(\mathcal{A})$. The proof that $Rank(\mathcal{A}) \geq Rank(A)$ follows similarly. ∎

Theorem 18. *If A and B are two matrix representations of the linear operator $\mathcal{A}$, then $Rank(A) = Rank(B)$.*

Proof: Note that $Rank(A) = Rank(\mathcal{A}) = Rank(B)$.

∎

In particular, let A be a $m \times n$ matrix.

P and Q are nonsingular $m \times m$, $n \times n$ matrices respectively. Then,

$$\begin{aligned} Rank(A) &= Rank(PA) \\ &= Rank(AQ) \\ &= Rank(PAQ) \end{aligned}$$

$\mathcal{A} : R^n \to R^m$ and $\mathcal{A}v = A\alpha$.

A, PA, AQ, PAQ are all matrix representations of $\mathcal{A}$

Null Space

Definition 15. Null Space: *Suppose V and W are vector spaces and $\mathcal{A}$ a linear operator from $V \to W$. Then,*

$$Null(\mathcal{A} = \{v \in V | \mathcal{A}v = 0\}$$

Note that $Null(\mathcal{A}) \subset V$ and $Range(\mathcal{A}) \subset W$.

Example 16.

$$V = \{\textit{Vector space of all polynomials of degree } \leq 2\}.$$

Let $W \equiv V, \;\; \mathcal{A}v = \frac{dv}{dt}$. *Then*

$$Null(\mathcal{A}) = \{\textit{all constants}\}$$

and

$$Basis(Null(\mathcal{A})) = 1.$$

Rank Nullity Theorem

Theorem 19. *Suppose V and W are vector spaces, and $dim(V) = n$. $\mathcal{A} : V \to W$ be a linear operator. Then*

$$dim(Null(\mathcal{A})) + dim(Range(\mathcal{A})) = n.$$

Proof: Suppose $dim(Null(\mathcal{A})) = n$. Therefore, $\exists$ independent vectors $v_1, v_2, \ldots, v_n$ such that

$$\mathcal{A}v_1 = \mathcal{A}v_2 = \ldots = \mathcal{A}v_n = 0.$$

Because $dim(V) = n$, $v_1, v_2, \ldots, v_n$ forms a basis for V. Thus, given any vector $v \in V$,

$$v = \sum_{i=1}^{n} \alpha_i v_i, \ \mathcal{A}v = \sum_{i=1}^{n} \alpha_i \mathcal{A} v_i = 0.$$

Thus,

$$Range(\mathcal{A}) = \{0\}.$$

Suppose $dim(null(\mathcal{A})) = q < n$. Then there exist independent vector $v_1, v_2, \dots, v_q$ such that

$$\mathcal{A}v_1 = \mathcal{A}v_2 = \dots = \mathcal{A}v_q = 0.$$

From Theorem 12 one can extend the basis to $v_1, v_2, \dots, v_q, v_{q+1}, \dots, v_n$. We will show that $\mathcal{A}v_{q+1}, \dots, \mathcal{A}v_n$ are independent. Note that

$$\sum_{q+1}^{n} c_i \mathcal{A} v_i = 0.$$

Then we have

$$\mathcal{A}(\sum_{q+1}^{n} c_i v_i) = 0$$

$$\sum_{q+1}^{n} c_i v_i \in null(\mathcal{A}).$$

As $v_{q+1}, \ldots, v_n$ are independent it follows that $c_i = 0 \ \forall \ i = q+1, \ldots, n$. Thus we have shown that

$$\mathcal{A}v_{q+1}, \ldots, \mathcal{A}v_n$$

are independent.

Suppose $w \in Range(\mathcal{A})$. Let $v \in V$ then $v = \sum_{i=1}^{n} \alpha_i v_i$. It follows that

$$\mathcal{A}v = \sum_{i=1}^{n} \alpha_i \mathcal{A}v_i = \sum_{i=q+1}^{n} \alpha_i \mathcal{A}v_i.$$

$\{\mathcal{A}v_{q+1}, \mathcal{A}v_{q+2}, \ldots, \mathcal{A}v_n\}$ is a basis for $Range(\mathcal{A})$. Thus

$$Range(\mathcal{A}) = span\{\mathcal{A}v_{q+1}, \ldots, \mathcal{A}v_n\}.$$

Thus

$$dim(Range(\mathcal{A})) = n - q$$

and it follows that

$$dim(Range(\mathcal{A}) + dim(null(\mathcal{A})) = n.$$

■

Theorem 20. *Let B and C be $m \times n$ and $n \times p$ matrices with $rank(B) = b$ and $rank(C) = c$. Then*

$$rank(BC) \leq min(b, c).$$

Proof: Note that $Range(BC) \subset RangeB$. Indeed, suppose there exists a y such that $BCy = z$ with $z \in Range(BC)$. It follows that $By' = z$ with $y' = Cy$.

Thus $z \in Range(B)$. Thus it follows that $Range(BC) \subset Range B$. Thus we can conclude that $dim(Range(BC) \leq dim(Range(B)) = b$.

Suppose $V \in Null(C)$. Then $Cv = 0$ and therefore $BCv = 0$. Therefore $Null(C) \subset Null(BC)$. This implies that $dim(Null(BC)) \geq dim(Null(C))$. Also note that

$$\begin{aligned} p &= dim(Null(C)) + dim(Range(C)) \\ &= dim(Null(BC)) + dim(Range(BC)). \end{aligned}$$

Since $dim(Null(BC)) \geq dim(Null(C))$ it follows that

$$rank(BC) = dim(Range(BC)) = p - dim(Null(BC)) \leq p - dim(Null(C)) = dim(R$$

Thus

$$rank(BC) \leq min(b, c)$$

$\blacksquare$

Theorem 21. *Let A be a $m \times n$ matrix of rank r then A can be written as $A = BC$ where B is a $m \times r$ matrix of rank r and c is a $r \times n$ matrix of rank r.*

Proof: Let $A : R^n \to R^m$ has rank r implies that there exist vectors $v_1, v_2, \ldots, v_r$ which forms a basis for $Range(A)$. Now note that

$$A = [a_1 \; a_2 \; \ldots \; a_n]$$

and $a_i \in Range(A)$. Therefore c_i represents the coordinate vector of a_i in the basis $v_1, \ldots v_r$ then we have

$$a_i = \sum_{j=1}^{r} c_{ji} v_j.$$

Thus

$$A = [a_1 \; a_2 \; \ldots a_n] = B[c_1 \; c_2 \ldots c_n]$$

where $B = [v_1\ v_2\ \ldots\ v_r]$. As $v_1, \ldots, v_r$ are linearly independent it follows that B has rank r. Note that $r = rank(A) \leq rank(C)$. However C has only r rows and thus $rank(C) = r$.

∎

Theorem 22. *Suppose $A \in R^{m \times n}$. Consider the equation*

$$A\alpha = \beta \tag{5}$$

where $\alpha \in R^n$ and $\beta \in R^m$. Then (5) has a solution if and only if $\beta \in Range(A)$. If a solution exists then it is unique if and only if $Null(A) = \{0\}$.

Proof: We will prove only the second part of the theorem. Suppose $NullA = \{0\}$. If α_1 and α_2 are two elements such that $A\alpha_1 = A\alpha_2$ then $A(\alpha_1 - \alpha_2) = 0$ and therefore $\alpha_1 - \alpha_2 = 0$. Thus $\alpha_1 = \alpha_2$. Thus the solution to $A\alpha = b$ is unique.

Suppose $Null(A) \neq \{0\}$. Then there exists $\alpha_1 \neq 0$ such that $A\alpha_1 = 0$. Suppose $A\alpha = \beta$ then $A(\alpha + \alpha_1) = \beta$ and therefore the solution is not unique.

∎

Definition 16. *Let $\mathcal{A} : V \to W$ be a linear operator with V and W are vector spaces.*

$\mathcal{A}$ is said to be right invertible if there exist a map $\mathcal{A}^{-R} : W \to V$ such that $\mathcal{A}\mathcal{A}^{-R} = I_w$ where I_w is the identity transformation on W.

$\mathcal{A}$ is said to be left invertible if there exist a map $\mathcal{A}^{-l} : W \to V$ such that $\mathcal{A}^{-l}\mathcal{A} = I_v$ where I_v is the identity transformation on V.

$\mathcal{A}$ is invertible if it has both right and left inverses.

Theorem 23. *Let $\mathcal{A} : V \to V$ where $\mathcal{A}$ is linear and V is a vector space.*

1. *If there exists a unique right inverse to $\mathcal{A}$ then $\mathcal{A}$ is invertible.*

2. If there exists a unique left inverse to $\mathcal{A}$ then $\mathcal{A}$ is invertible.

Proof: (1) Suppose $\mathcal{A}^{-R}$ is the right inverse of $\mathcal{A}$. Note that

$$\mathcal{A}(\mathcal{A}^{-R} + \mathcal{A}^{-R}\mathcal{A} - I) = \mathcal{A}\mathcal{A}^{-R} + \mathcal{A}\mathcal{A}^{-R}\mathcal{A} - \mathcal{A} = I + \mathcal{A} - \mathcal{A} = I.$$

As the right inverse is unique it follows that

$$\mathcal{A}^{-R} + \mathcal{A}^{-R}\mathcal{A} - I = \mathcal{A}^{-R}.$$

Thus

$$\mathcal{A}^{-R}\mathcal{A} = I$$

and thus $\mathcal{A}^{-R}$ is the left inverse of $\mathcal{A}$. Thus $\mathcal{A}$ is invertible.

(2) follows in a similar way as (1). ∎

Definition 17. **Onto and into:** *$\mathcal{A} : V \to W$ is onto if $Range(\mathcal{A}) = W$. If $\mathcal{A}$ is such that $\mathcal{A}\alpha_1 = \mathcal{A}\alpha_2$ implies that $\alpha_1 = \alpha_2$ for any pair $\alpha_1,\ \alpha_2 \in V$ then $\mathcal{A}$ is into.*

Example 17. *Let $\mathcal{A} : R^2 \to R$ be defined by*

$$\mathcal{A}v = Av$$

where

$$A = \begin{pmatrix} 1 & 2 \end{pmatrix}.$$

Notice that $Range(\mathcal{A}) = R$. Indeed if $\alpha \in R$ then

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

and this $\mathcal{A}$ is onto.

Now we will find a right inverse to A. Consider the equation

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 1.$$

and thus $\beta_1 + 2\beta_2 = 1$ *Thus any* $(\beta_1, \beta_2)^T$ *is a right inverse if* β_1, β_2 *satisfy* $\beta_1 + 2\beta_2 = 1$. *Evidently there are infinite number of right inverses.*

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

is a right inverse.

$$\begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

is a right inverse too.

Example 18. $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathcal{A} : R^1 \to R^2.$

Then $\mathcal{A}$ is one to one. ($\because Null(A) = 0$)

$\mathcal{A}^{-l}$ is a left inverse if $\begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = I$ where $\mathcal{A}^{-l} \equiv \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix}$

$\begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix}$ is a left inverse if $\alpha_1 + 2\alpha_2 = 1$

Again this has infinite solutions and thus there are infinite left inverses for $\mathcal{A}$.

Theorem 24. *Consider $\mathcal{A} : V \to W$ where $dim(V) = n,\ dim(W) = m$. Then $\mathcal{A}$ is one to one if and only if $m \geq n$ and the rank of any matrix representation of $\mathcal{A}$ is n. In particular, if $n = m$ then $rank(\mathcal{A}) = n$ only if $\mathcal{A}$ is non singular.*

Proof: Note that from Theorem 19 it follows that

$$dim(N(\mathcal{A})) + dim(R(\mathcal{A})) = n.$$

If $m < n$, then $dim(N(\mathcal{A})) = n - dim(R(\mathcal{A})) \geq (n - dim(W)) = (n - m) > 0$. Therefore, if $m < n$, then $\mathcal{A}$ is not one to one as $N(\mathcal{A}) \neq \{0\}$.

$m \geq n$ and $rank(\mathcal{A}) = n$. Then $dim(N(\mathcal{A})) = \{0\})$. Therefore $\mathcal{A}$ is $1 - 1$. ∎

Theorem 25. *Let $\mathcal{A} : V \to W$ be a linear operator where V and W are vector spaces. Then*

1. *$\mathcal{A}$ is right invertible if and only if $\mathcal{A}$ is onto.*

2. *$\mathcal{A}$ is left invertible if and only if $\mathcal{A}$ is one to one.*

Proof: Suppose $\mathcal{A}$ is onto. Then given any $w \in W$ there exists $v \in V$ such that $\mathcal{A}v = w$ (note that v is not unique).

Define $\mathcal{A}^{-R}w := v$ where v is any vector that satisfies $\mathcal{A}v = w$. Then it follows that $\mathcal{A}(\mathcal{A}^{-R}w) = \mathcal{A}v = w$.

Suppose $\mathcal{A}$ is not onto, then $\exists w^1$, such that $w^1 \notin Range(\mathcal{A})$

Suppose $\exists$ a right inverse operator $\mathcal{A}^{-R}$ Then for the given $w^1 \in W$, $\mathcal{A}(\mathcal{A}^{-R}w^1) = w^1$.

Then with $v = \mathcal{A}^{-R}w^1$, we have $\mathcal{A}v = w^1$. Thus, $w^1 \in Range(\mathcal{A})$ and we have a contradiction.

This proves (1). (2) is left as an exercise. ∎

Eigenvalues and Eigenvectors of operators

Definition 18. *Let $\mathcal{A}$ be a linear operator from V to W where V and W are of the same dimension n. Then λ, a scalar is called an eigenvalue if $\mathcal{A}v = \lambda v$ for some $v \neq 0, v \in V \quad v$ is the eigenvector associated with λ.*

Theorem 26. *Let $\mathcal{A} : V \to V$ be a linear operator, and let V be n-dimensional. Then all matrix representations of $\mathcal{A}$ have the same n-eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Moreover, these eigenvalues are precisely the eigenvalues of $\mathcal{A}$.*

Theorem 27. *Similar matrices have the same characteristic polynomial and therefore they have the same eigenvalues. Moreover, if $\hat{A} = P^{-1}AP$ and V is an eigenvector of A, then Pv is an eigenvector of $\hat{A}$. A and $\hat{A}$ are both matrix representations of the linear operator $\mathcal{A}$ defined by $\mathcal{A}v = Av$.*

Inner Product Spaces

Definition 19. Inner Product: (V, s) *is a vector space* V *with scalar being* s*. An inner product on* (V, s) *is a function* $<,>: (V, s) \times (V, s) \rightarrow s$ *which has the following properties:*

1. $< v, v > \ \geq\ 0$ *for all* $v \in V$ *and* $< v, v >= 0$ *only if* $v = 0$*.*

2. $< v, w >=< w, v > \quad v, w \in V, s \equiv R$

$< v, w >= \overline{< w, v >} \qquad v, w \in V, s \equiv C$

3. $< \alpha v, w >= \overline{\alpha} < v, w > \qquad v, w \in V, \alpha \in s.$

4. $< v_1 + v_2, w >=< v_1, w > + < v_2, w >, \qquad v_1, v_, w \in V.$

Inner Product Spaces

Definition 20.

(V, s) is a vector space with an inner product defined is called an inner product space.

Example 19. *Let* $(V, s) \equiv (R^2, R)$

$$< v_1, v_2 >:= (v_1)^T v_2 = \sum_{i=1}^{2} v_1(i) v_2(i)$$

where $v_1 = \begin{bmatrix} v_1(1) \\ v_1(2) \end{bmatrix}$ $v_2 = \begin{bmatrix} v_2(1) \\ v_2(2) \end{bmatrix}$

$<, >$ *is indeed an inner product on* (R^2, R)

Orthogonal and orthonormal vectors

Definition 21. (V, s) *be an inner product space. Then two non-zero vectors* $v_1, v_2, v_3, \ldots, v_n$ *are orthogonal if* $< v_i, v_j >= 0$ *if* $i \neq j,\ j = 1, 2, \ldots, n$. *They are orthonormal if in addition* $< v_i, v_i >= 1$ *for* $i = 1, 2, \ldots, n$.

Orthogonal complements

Definition 22. *suppose* X *is an inner product space and* V *and* W *are subspaces of* X*, then,* V *and* W *are said to be orthogonal complements of one another if* $V \oplus W = X$ *and* $< v, w >= 0 \;\; \forall v \in V, w \in W.$

Example 20. $X \equiv R^2$

NOTE;DRAW FIGURE Let

$$V = \{v : v = \alpha(\begin{array}{c} 1 \\ 0 \end{array}), \alpha \in R\},$$

and

$$W = \{w : w = \beta(\begin{array}{c} 0 \\ 1 \end{array}), \beta \in R\}.$$

Then

$$V \cap W = \{0\},$$

$$V \oplus W \equiv R^2 = \{v : v = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \alpha, \beta \in R\}.$$

Also, for $v \in V, w \in W,$

$$< v, w >= (\begin{array}{c} \alpha \\ 0 \end{array})^T (\begin{array}{c} 0 \\ \beta \end{array}) = (\alpha, 0)(\begin{array}{c} 0 \\ \beta \end{array}) = 0.$$

Thus V *and* W *are orthogonal complements.*

$X \equiv R^3$*, then*

$$V = \{ \begin{bmatrix} \alpha \\ 0 \\ 0 \end{bmatrix} |\alpha \in R\}, \text{ and } W = \{ \begin{bmatrix} 0 \\ \alpha \\ 0 \end{bmatrix} |\alpha \in R\}$$

are not orthogonal complements.

Orthogonal Subspaces

Definition 23. *V, and W subspaces of inner product space X are orthogonal to each other. If for every $v \in V$, $w \in W$, $< v, w >= 0$.*

If V is a subspace of an inner product space X, then

$$V^{\perp} = \{x \in X| < x, v >= 0, \ \forall v \in V\}.$$

It can be shown that

- $V^{\perp}$ is a subspace of X.
- $V \cap V^{\perp} = 0$
- $V \oplus V^{\perp} = X$.

Adjoint Operator

Definition 24. *Suppose* V *is an inner product space and let* $\mathcal{A}: V \to W$ *be a linear operator, where* W *is also an inner product space. Then the adjoint of the operator* $\mathcal{A}$ *is an operator* $\mathcal{A}^*: W \to V$ *that is defined by*

$$< v, \mathcal{A}^* w >_v = < \mathcal{A} v, w >_w, \; v \in V, w \in W.$$

Example 21. *Let* $V = R^n$ *and* $W = R^m$ *and let* $\mathcal{A}: V \to W$ *be defined by*

$$\mathcal{A} v = A v,$$

where $A = (a_{ij})$. *Let the inner product on* V *and* W *be defined by*

$$< v_1, v_2 >_v = v_1^T v_2 \text{ and } < w_1, w_2 >_w = w_1^T w_2, \; v_1, v_2 \in V \text{ and } w_1, w_2 \in W.$$

Note that

$$< v, A^T w >_v = v^T A^T w = (Av)^T w = < Av, w >_w .$$

Thus the adjoint operator of $\mathcal{A}$ is given by the matrix A^T.

If $V \subset X, \mathcal{A} : V \to W$ then, $N(\mathcal{A})^\perp \subset V$, $N(\mathcal{A}) \subset V$, $R(\mathcal{A}) \subset W$,$R(\mathcal{A})^\perp \subset W$, $Range(\mathcal{A}^*) \subset V$,$N(\mathcal{A}^*) \subset W$, $R(\mathcal{A}^*)^\perp \subset V$, $N(\mathcal{A}^*)^\perp \subset W$

Let V and W be two vector spaces and let $\mathcal{A} : V \to W$ be a linear operator. Then,

- $\mathcal{A}$ is onto if $R(\mathcal{A}) = W$
- $\mathcal{A}$ is one to one if $N(\mathcal{A}) = \{0\}$.

Theorem 28. *The following statements are equivalent:*

1. $N(\mathcal{A}) = \{0\}$

2. If $\mathcal{A}v_1 = \mathcal{A}v_2$, then $v_1 = v_2$.

3. If $v_1 \neq v_2$, then $\mathcal{A}v_1 \neq \mathcal{A}v_2$.

Proof: Suppose $N(\mathcal{A}) = \{0\}$. Also, if v_1, v_2 are such that $\mathcal{A}v_1 = \mathcal{A}v_2$, then

$\mathcal{A}(v_1 - v_2) = 0$

$\Rightarrow (v_1 - v_2) \in N(\mathcal{A})$

$\Rightarrow v_1 - v_2 = 0$

$\therefore v_1 = v_2$

Suppose that $\mathcal{A}v_1 = \mathcal{A}v_2 \Rightarrow v_1 = v_2$

Then if $v \in N(\mathcal{A}), \mathcal{A}v = 0$ is same as

$\mathcal{A}(v - 0) = 0$

$\Rightarrow \mathcal{A}v - \mathcal{A}0 = 0$

$\Rightarrow v = 0$

$\therefore N(\mathcal{A}) = \{0\}$

$\therefore 1 \Leftrightarrow 2$ ∎

Theorem 29. *Let $\mathcal{A}$ be a linear operator from an inner product space V to an inner product space W. Then*

1. $N(\mathcal{A}^*) = [R(\mathcal{A})]^\perp$

2. $[N(\mathcal{A})]^\perp = R(\mathcal{A}^*)$

Proof: (1) Take $w \in [Range(\mathcal{A})]^\perp$ then

$$
\begin{aligned}
&< w, \gamma >_w = 0, \quad \forall \gamma \in Range(\mathcal{A}) \\
&\Rightarrow < w, \mathcal{A}v >_w = 0, \quad \forall v \in V \\
&\Rightarrow < \mathcal{A}v, w >_w = 0, \quad \forall v \in V \\
&\Rightarrow < v, \mathcal{A}^* w >_v = 0, \quad \forall v \in V
\end{aligned}
$$

In particular, $< \mathcal{A}^* w, \mathcal{A}^* w >_v = 0$. $\mathcal{A}^* w = 0$ and thus $w \in Null(\mathcal{A}^*)$. This shows that $Range(\mathcal{A})]^\perp \subset Null(\mathcal{A}^*)$.

Let $w \in N(\mathcal{A}^*)$, then $\mathcal{A}^* w = 0$

$\therefore < v, \mathcal{A}^* w >_v = 0 \quad \forall v \in V$

$\therefore < \mathcal{A}v, w >_w = 0 \quad \forall v \in V$

$\Rightarrow w \in [Range(\mathcal{A})] \perp$

Thus, $(Range(\mathcal{A}))^\perp = N(\mathcal{A}^*) \quad [v \notin [N(A)]^\perp \Leftrightarrow v \notin Range(\mathcal{A}^*)$.

∎

Gram-Schmidt Orthonormalization

Theorem 30. *Let V be a vector space with the inner product $<,>$ defined. Let $v_1, \dots, v_n$ be n independent vectors. Then there exist n orthonormal vectors $e_1, \dots, e_n$ such that*

$$span\{v_1, \dots, v_n\} = span\{e_1, \dots, e_n\}.$$

Proof: Let

$$z_1 := v_1$$

and let

$$e_1 := \frac{z_1}{\|z_1\|}.$$

Let

$$z_2 = v_2 - < v_2, e_1 > e_1 \text{ and } e_2 := \frac{z_2}{\|z_2\|}.$$

Note that

$$
\begin{aligned}
< e_2, e_1 > &= \tfrac{1}{\|z\|_2}[< v_2, e_1 > - < v_2, e_1 >< e_2, e_2 >] \\
&= \tfrac{1}{\|z\|_2}[< v_2, e_1 > - < v_2, e_1 >] = 0
\end{aligned}
$$

Thus $e_2 \perp e_1$. Given $e_1, e_2, \ldots, e_i$ orthonormal define

$$
\begin{aligned}
z_{i+1} &= v_{i+1} - < v_{i+1}, e_1 > e_1 - < v_{i+1}, e_2 > e_2 - \ldots - < v_{i+1}, e_i > e_i \\
&= v_{i+1} - \textstyle\sum_{j=1}^{i} < v_{i+1}, e_j > e_j, \text{ and} \\
e_{i+1} &:= \tfrac{z_{i+1}}{\|z_{i+1}\|}
\end{aligned}
.
$$

Let $k \leq i$ then

$$\begin{aligned} < z_{i+1}, e_k > &= < v_{i+1}, e_k > - \sum_{j=1}^{i} < v_{i+1}, e_j >< e_j, e_k > \\ &= < v_{i+1}, e_k > - \sum_{j=1}^{i} < v_{i+1}, e_j > \delta_{jk} \\ &= < v_{i+1}, e_k > - < v_{i+1}, e_k > \\ &= 0 \text{ and} \\ < e_{i+1}, e_k > &= \frac{< z_{i+1}, e_k >}{\| z_{i+1} \|} = 0. \end{aligned}$$

Thus $< e_{i+1}, e_j >= 0$ for all $j = 1, \ldots, i$. Thus this procedure yields vectors $e_1, \ldots, e_n$ that are orthonormal. Note that e_i is a linear combination of v_j $j = 1, \ldots, n$. Thus

$$span\{e_1, \ldots, e_n\} \subset span\{v_1, \ldots, v_n\}.$$

Note that $e_i,\ i = 1, \ldots, n$ forms an orthonormal set it also forms an independent set. Therefore

$$dim(span\{e_1, \ldots, e_n\}) = dim(span\{v_1, \ldots, v_n\}) = n$$

and thus

$$span\{e_1, \ldots, e_n\} = span\{v_1, \ldots, v_n\}.$$

∎

Theorem 31. *Let A be a $n \times n$ Hermitian matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. There exists a unitary matrix P such that*

$$P^* A P = \Lambda = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}$$

Proof: Let x_1 such that $\|x_1\|_2 = 1$ and $Ax_1 = \lambda_1 x_1$. Let $u_2,\ u_3, \ldots, u_n$ be orthonormal vectors such that $\{x_1, u_2,\ u_3, \ldots, u_n\}$ form an orthonormal set. Let

$$P_1 = [x_1\ u_2\ \ \ldots\ u_n],$$

Then we have that

$$P_1^* P_1 = I.$$

Let $U_1 = [u_2 \; u_3 \ldots \; u_n]$. Note that

$$\begin{aligned}
P_1^* A P_1 &= \begin{bmatrix} x_1^* \\ U_1^* \end{bmatrix} A \begin{bmatrix} x_1 & U_1 \end{bmatrix} \\
&= \begin{bmatrix} x_1^* \\ U_1^* \end{bmatrix} \begin{bmatrix} A x_1 & A U_1 \end{bmatrix} \\
&= \begin{bmatrix} x_1^* \\ U_1^* \end{bmatrix} \begin{bmatrix} \lambda_1 x_1 & A U_1 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & x_1^* A U_1 \\ 0 & U_1^* A U_1 \end{bmatrix}.
\end{aligned}$$

Note that

$$(P_1^* A P_1)^* = \begin{bmatrix} \lambda_1 & x_1^* A U_1 \\ 0 & U_1^* A U_1 \end{bmatrix}^* = \begin{bmatrix} \lambda_1 & 0 \\ U_1^* A^* x_1 & U_1^* A^* U_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ U_1^* A x_1 & U_1^* A U_1 \end{bmatrix}.$$

However

$$(P_1^* A P_1)^* = P_1^* A^* P_1 = P_1^* A P_1 = \begin{bmatrix} \lambda_1 & x_1^* A U_1 \\ 0 & U_1^* A U_1 \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \lambda_1 & 0 \\ U_1^* A x_1 & U_1^* A U_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & x_1^* A U_1 \\ 0 & U_1^* A U_1 \end{bmatrix}.$$

Thus

$$x_1^* A U_1 = 0 = U_1^* A^* x_1 \text{ and}$$

$$P_1^* A P_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & U_1^* A U_1 \end{bmatrix}.$$

Recall that

$$P_1^* P_1 = I.$$

Therefore eigenvalues of $P_1^* A P_1$ are the eigenvalues of A and thus eigenvalues of $U_1^* A U_1 := A_2$ are $\lambda_2, \ldots, \lambda_n$. Let x_2 such that $\|x_2\|_2 = 1$ and $A_2 x_2 = \lambda_2 x_2$. Let $\hat{u}_3,\ \hat{u}_4, \ldots, \hat{u}_n$ be orthonormal vectors such that $\{x_2, \hat{u}_3,\ \ldots, \hat{u}_n\}$ form an orthonormal set. Let $U_2 := [\hat{u}_3\ \ldots\ \hat{u}_n]$. Let

$$Q_2 := [x_2\ U_2] \text{ and } P_2 = \left[\begin{array}{cc} 1 & 0 \\ 0 & Q_2 \end{array}\right].$$

Note that $P_2^* P_2 = I$. Note that

$$Q_2^* A_2 Q_2 = \left[\begin{array}{ll} \lambda_2 & 0 \\ 0 & U_2^* A_2 U_2 \end{array}\right].$$

Note that

$$
\begin{aligned}
(P_1P_2)^*A(P_1P_2) \\
&= P_2^*(P_1^*AP_1)P_2 \\
&= P_2^*\begin{bmatrix} \lambda_1 & 0 \\ 0 & A_2 \end{bmatrix} P_2 \\
&= \begin{bmatrix} 1 & 0 \\ 0 & Q_2^* \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & A_2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & Q_2 \end{bmatrix}. \\
&= \begin{bmatrix} \lambda_1 & 0 \\ 0 & Q_2^*A_2Q_2 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & U_2^*A_2U_2 \end{bmatrix}
\end{aligned}
$$

Continue the argument to obtain

$$P = P_1 P_2 \ldots P_n \text{ and } P^* A P = diag(\lambda_1, \ldots, \lambda_n).$$

Definition 25. *Suppose A and B are matrices such that there exists a P with $P^*P = I$ such that $B = P^*AP$. Then A and B are unitarily similar.*

Theorem 32. *Any $n \times n$ Hermitian matrix A has n orthogonal eigenvectors that form a basis for C^n. In this basis A has a diagonal representation.*

Theorem 33. *If $\mathcal{A}$ is a self adjoint operator on a finite dimensional space V then $\mathcal{A}$ has real eigenvalues and corresponding eigenvectors form a basis for V. In this basis $\mathcal{A}$ has a diagonal representation.*

Theorem 34. *Let A be a $n \times n$ Hermitian matrix with eigenvalues*

$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. *Then for all* $x \in C^n$

$$\lambda_1 x^* x \leq x^* A x \leq \lambda_n x^* x.$$

Proof: Note that $A = A^*$ and that there exists a P such that $P^* A P = \Lambda$ and $P^* P = I$ with Λ diagonal. Thus

$$\begin{aligned} x^* A x &= x^* P \Lambda P^* x \\ &= (P^* x)^* \Lambda \overbrace{(P^* x)}^{:= y} \\ &= y^* \Lambda y \\ &= \textstyle\sum_{i=1}^n \lambda_i y_i^* y_i = \sum_{i=1}^n \lambda_i |y_i|^2 \\ &\leq \lambda_n \|y\|_2^2 \\ &= \lambda_n \|x\|_2^2 \end{aligned}$$

Note that as $y = P^* x,\ y^* y = x^* P P^* x = x^* x$.

The fact that $x^*Ax \geq \lambda_1 x^*x$ is left as an exercise. ∎

Theorem 35. *Let A be a $n \times n$ Hermitian matrix.*

1. *A is positive definite if and only if its eigenvalues are positive.*

2. *A is positive semi-definite if and only if all its eigenvalues are nonnegative.*

3. *A is negative definite if and only if all its eigenvalues are negative.*

4. *A is negative semi-definite if and only if all its eigenvalues are non-positive.*

Definition 26. *Suppose $A : R^n \to R^n$. Then*

$$\|A\|_{2-in} := \max_{\|x\|_2=1} \|Ax\|_2 = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}.$$

Schur's Theorem

Theorem 36. *If A is a $n \times n$ matrix, then there is a unitary matrix P such that $P^*AP = T \ \ (P^*P = 1)$, where T is an upper triangular matrix.*

Theorem 37. *$n \times n$ matrix A is a unitary matrix similar to a diagonal matrix if and only if it commutes with its conjugate transpose ($AA^* = A^*A$).*

Proof: $(\Rightarrow)$ *There exists a P such that $P^*P = I$ and*

$$P^*AP = \Lambda.$$

Thus

$$A = P\Lambda P^* \text{ and } A^* = P\Lambda^* P^*.$$

Thus

$$AA^* = P\Lambda P^* P\Lambda^* P^* = P\Lambda\Lambda^* P^* = P\Lambda^*\Lambda P^* = P\Lambda^* P^* P\Lambda P^* = A^*A.$$

Note that we have used the fact that ss Λ is diagonal $\Lambda\Lambda^* = \Lambda^*\Lambda$. The rest of the proof is left to the reader. ∎

Definition 27. *$n \times n$ matrix A commute with its conjugate transpose is called Normal Matrix.*

A is normal if $AA^* = A^*A$.

Theorem 38. *Let A be a $n \times n$ matrix, then A is similar to a diagonal matrix if and only if A has n independent eigenvectors.*

Proof: *($\Leftarrow$) : Assume there exists n independent eigen vectors $x_1, x_2, \ldots, x_n$. Then*

$$\begin{aligned}
A\overbrace{[x_1, x_2, \ldots, x_n]}^{:=P} &= [\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n] \\
AP &= P\Lambda (P \text{ is invertible }) \\
P^{-1}AP &= \Lambda
\end{aligned}$$

(⇒): There exists P (invertible) such that

$$\begin{aligned} P^{-1}AP &= \Lambda \\ AP &= P\Lambda \\ P &= [p_1, p_2, \ldots, p_n] \\ A[p_1, p_2, \ldots, p_n] &= [p_1\lambda_1, p_2\lambda_2, \ldots, p_n\lambda_n] \\ Ap_i &= \lambda_i p_i \end{aligned}$$

Thus A has $i = n$ independent eigen vectors as P is invertible. ∎

Theorem 39. *If $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct eigen values of A, then the corresponding eigen vectors $x_1, x_2, \ldots, x_m$ are independent.*

Proof: *Suppose to the contrary, $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct but $x_1, x_2, \ldots, x_m$ are dependent. Then*

$\sum_{i=1}^{m} c_i x_i = 0$ *and without loss of generality say* $c_m \neq 0$*. Then*

$$\lambda_1 (\sum_{i=1}^{m} c_i x_i) = 0 \quad (6)$$

$$A (\sum_{i=1}^{m} c_i x_i) = 0 \quad (7)$$

(7) - (6) $\Rightarrow \sum_{i=1}^{m} c_i (A x_i - \lambda_1 x_i) = 0$

$$\Rightarrow \sum_{i=2}^{m} c_i (\lambda_1 - \lambda_i) x_i = 0$$

Multiply by λ_2 and A and subtract each other and follow the same by λ_3 and A $\ldots$. Then

$$c_m(\lambda_1 - \lambda_m)(\lambda_1 - \lambda_m)\ldots(\lambda_{m-1} - \lambda_m) = 0$$

which is a contradiction to our assumption. ∎

Theorem 40. *If a $n \times n$ matrix A has n distinct eigenvalues then A is similar to a diagonal matrix.*

Proof: Follows from the previous two theorems ∎

Theorem 41. *Let A be a $m \times n$ matrix with rank r. Then there exist $m \times m$ unitary matrix P and $n \times n$ unitary matrix Q such that*

$$\Sigma = P^* A Q$$

where Σ is a $m \times n$ matrix with only the first r diagonal elements called the singular values $\sigma_1, \ldots, \sigma_r$ nonzero and rest of the elements zero. The first r singular values are given by

$$\sigma_i = \{\lambda_i(A^*A)\}^{\frac{1}{2}}.$$

Proof: Note that $rank(A^*A) = rank(A) = r$. Let the eigenvalues of A^*A be given by $\lambda_1, \ldots, \lambda_n$ with corresponding eigenvectors $x_1, \ldots, x_n$ that are orthogonal (see Theorem 32). Note that A^*A is Hermitian positive seim-definite and thus all its eigenvalues are non-negative. Define

$$\sigma_i = \{\lambda_i(A^*A)\}^{\frac{1}{2}}.$$

Let

$$y_i = \frac{1}{\sigma_i} A x_i, \;\; i = 1, 2, \ldots, r.$$

Note that

$$
\begin{aligned}
y_i^* y_j &= \frac{1}{\sigma_i \sigma_j}(Ax_i)^*(Ax_j) \\
&= \frac{1}{\sigma_i \sigma_j} x_i^* A^* A x_j \\
&= \frac{1}{\sigma_i \sigma_j} \lambda_j x_i^* x_j \\
&= \frac{\sigma_j}{\sigma_i} \delta_{ij}.
\end{aligned}
$$

$\{y_1, y_2, \ldots, y_r\}$ forms an orthonormal and can be extended to $\{y_1, y_2, \ldots, y_m\}$ to form a orthonormal basis for C^m. Let

$$Q = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \text{ and } P = \begin{bmatrix} y_1 & \ldots & y_m \end{bmatrix}$$

Note that $P^*P = PP^* = Q^*Q = QQ^* = I$. Note that for all

$j = 1, \ldots, n$ and $i = 1, \ldots, r$

$$\begin{aligned} (P^* A Q)_{ij} &= y_i^* A x_j \\ &= \frac{1}{\sigma_i} (A x_i)^* A x_j \\ &= \frac{1}{\sigma_i} x_i^* A^* A x_j \\ &= \frac{\lambda_j}{\sigma_i} x_i^* x_j \\ &= \frac{\lambda_j}{\sigma_i} \delta_{ij} \end{aligned} .$$

Also, if $j = 1, \ldots, r$ and $i = r+1, \ldots, m$ then

$$(P^* A Q)_{ij} = y_i^* A x_j = y_i^* (\sigma_j y_j) = \sigma_j y_i^* y_j = 0.$$

Thus

$$y_i^* A x_j = \sigma_i \delta_{ij} \text{ for all } i = 1, \ldots, r \text{ and } j = 1, \ldots, n.$$

Note that

$$\|A x_j\|_2^2 = x_j^* A^* A x_j = 0, \text{ for all } j = r+1, \ldots, n.$$

Thus

$$Ax_j = 0 \text{ for all } j = r+1, \ldots, n.$$

Thus

$$y_i^* A x_j = \sigma_i \delta_{ij} \text{ for all } i = 1, \ldots, r \text{ and } j = 1, \ldots, r.$$

$$y_i^* A x_j = 0 \text{ for all } i = 1, \ldots, r \text{ and } j = r+1, \ldots, n.$$

$$y_i^* A x_j = 0 \text{ for all } i = r+1, \ldots, m \text{ and } j = 1, \ldots, r.$$

Thus

$$P^* A Q = \left[\begin{array}{ccc|c} \sigma_1 & & & 0 \\ & \ddots & & \vdots \\ & & \sigma_r & 0 \\ \hline 0 & \ldots & 0 & 0 \end{array}\right]$$

∎

Two Induced Norm

Theorem 42. *Let* $A \in R^{n\times n}$. *Then*

$$\|A\|_{2-ind} = \sqrt{\rho(A^* A)}.$$

where $\rho(B)$ *denotes the spectral radius of* B.

Jordan Canonical Form

Consider a matrix of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}_{r \times r}$$

Then J_i is said to be a Jordan block with eigenvalue λ_i and size r. Note that e_1 is the only eigenvector of J_i.

Theorem 43. *A $n \times n$ matrix A is similar to the matrix*

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}$$

where J_i is the Jordan block with eigenvalue λ_i and size $r_i \times r_i$ given by

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}_{r_i \times r_i}$$

and

$$\sum_{i=1}^{p} r_i = n.$$

Definition 28. *The number of Jordan blocks associated with an eigenvalue λ_i is said to be the geometric multiplicity of λ_i. The number of eigenvalues at λ_i is called the algebraic multiplicity of the eigenvalue λ_i.*

Note that from Theorem 43 there exists a invertible matrix P such that

$$P^{-1}AP = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}$$

Thus

$$AP = PJ = \begin{bmatrix} p_1 & p_2 & \dots & p_n \end{bmatrix} \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}.$$

Thus

$$\begin{bmatrix} Ap_1 & Ap_2 & \dots & Ap_{r_1} & \dots & Ap_n \end{bmatrix} = \begin{bmatrix} p_1 & p_2 & \dots & p_{r_1} & \dots & p_n \end{bmatrix} \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}.$$

This implies that

$$\begin{aligned} Ap_1 &= \lambda_1 p_1 \\ (A - \lambda_1 I)p_2 &= p_1 \\ (A - \lambda_1 I)p_3 &= p_2 \\ \vdots \quad &\vdots \quad \vdots \\ (A - \lambda_1 I)p_{r_1} &= p_{r_1 - 1} \end{aligned}$$

p_{r_1} is called the generator. $p_2, \dots, p_{r_1}$ are called generalized eigenvectors.

Definition 29. Y *is an invariant set with respect to* A *if for all* $y \in Y$, $Ay \in Y$.

$S_1 = span\{p_1, p_2, \ldots, p_{r_1}\}$ associated with eigenvalue λ_1 is invariant with respect to A. Similarly S_j, the corresponding set with respect to λ_j is invariant with respect to A for all $j = 1, \ldots, p$

Theorem 44. *Let $A = P^{-1}JP$ be the Jordan decomposition of A with*

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}.$$

If S_i is defined as above then $S_i's$ are invariant with respect to A and

$$C^n = S_1 \oplus S_2 \oplus \ldots \oplus S_p.$$

Cayley Hamilton Theorem

Theorem 45. *The characteristic polynomial associated with matrix A is*

$$f(\lambda) = (\lambda - \lambda_1) \ldots (\lambda - \lambda_n).$$

Then

$$f(A) = 0.$$

Proof: Let the Jordan decomposition be given by

$$J = P^{-1}AP.$$

Thus

$$A^m = PJ^mP^{-1}.$$

Note that

$$f(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \ldots (\lambda - \lambda_p)^{r_p}.$$

Thus it follows that

$$f(A) \quad = \quad P(f(J))P^{-1}$$

where f is any polynomial. Note that

$$\begin{aligned} f(J) &= f\left(\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix}\right) \\ &= (J-\lambda_1 I)^{r_1}(J-\lambda_2 I)^{r_2}\ldots(J-\lambda_p I)^{r_p} \\ &= 0 \end{aligned}$$

Thus $f(A) = 0$. ∎

.

Minimal polynomial

Definition 30. *The minimal polynomial of a square matrix A is the least ordered polynomial $p(\lambda)$ such that $p(A) = 0$.*

Theorem 46. *Suppose A has m distinct eigenvalues. Let t_i be the size of the largest Jordan block of A associated with eigenvalue λ_i. Then the minimum polynomial is given by*

$$\Pi_{i=1}^{m}(\lambda - \lambda_i)^{t_i}.$$