MA1522 - Linear Algebra

Sherisse Tan Jing Wen

November 25, 2024

Contents

1 Linear Systems **3**
1.1 Elementary Row Operations . . . 4
1.2 Solutions . . . 4
1.2.1 Gaussian and Gauss-Jordan Elimination . . . 5
1.2.2 Unknown Coefficients . . . 5
1.2.3 Solution Sets . . . 6

2 Matrices **7**
2.1 Matrix Operations . . . 7
2.2 Determinant and Adjoint . . . 8
2.2.1 Finding Determinant . . . 9
2.3 Inverse of a Matrix . . . 10
2.4 Elementary Matrices . . . 10
2.5 LU Factorization . . . 11

3 Vector Spaces **12**
3.1 Linear Combination . . . 12
3.1.1 Linear Independence . . . 12
3.2 Linear Span . . . 13
3.3 Subspace . . . 14
3.4 Basis . . . 14
3.4.1 Coordinates Relative to a Basis . . . 15
3.5 Dimensions . . . 15
3.6 Transition Matrices . . . 16
3.6.1 Example . . . 16
3.7 Column / Row / Null Space . . . 17
3.7.1 Basis for Col(A), Row(A), Null(A) . . . 18
3.8 Rank . . . 18

4 Orthogonality **20**
4.1 Formulas . . . 20
4.2 Orthogonal and Orthonormal . . . 21
4.3 Gram-Schmidt Process . . . 22
4.4 QR Factorization . . . 23
4.5 Orthogonal Projection . . . 23
4.6 Least Square Approximation . . . 24
4.6.1 Examples . . . 25

5 Eigenanalysis **27**
5.1 Diagonalization . . . 28
5.2 Orthogonally Diagonalizable . . . 29
5.3 Markov Chain . . . 30
5.3.1 Full Examples . . . 30
5.4 Singular Value Decomposition . . . 32

6 Linear Transformation **35**

7 Appendix **38**
7.1 Equivalent Statements . . . 38
7.2 Matlab . . . 39

Chapter 1

Linear Systems

Definition 1.0.1 A Linear Equation, with n variables is an equation of the form:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ and $b \in \mathbb{R}$

Definition 1.0.2 A Linear **System** is made up of x Linear Equations

We say that a Linear System is Consistent $\iff$ $\exists$ a solution to the Linear System s.t. $a_1x_1+a_2x_2+\cdots+a_nx_n = b$

In contrast, we then say that a Linear System is Inconsistent $\rightarrow$

- $a_1 = a_2 = \ldots = a_n = 0$, BUT $b \neq 0$, that is, the last column is a Pivot Column, OR
- Not all Linear Equations can be satisfied by values of $x_1, x_2, \ldots, x_n$

We can then classify Linear Systems as 1 of the following:

1. Homogenous Linear System: $b_1 = b_2 = \cdots = b_n = 0$
2. Non-Homogenous Linear System: Not all $b_1 = b_2 = \cdots = b_n = 0$

Notice that for a Homogenous Linear System, we will always have a trivial solution where $x_1 = x_2 = \ldots = x_n = 0$. Thus, a Homogenous Linear System is **always** consistent.

Theorem 1.0.1 A Homogenous Linear System has a non-trivial solution $\iff$ it has infinitely many solutions

Theorem 1.0.2 – Solutions to Homogenous and Non-Homogenous Linear Systems Let v_1, v_2 be solutions to the Linear System $Ax = b$, and u be a solution to $Ax = 0$.

Then,

- $v_1 + u$ is also a solution to $Ax = b$
- $v_1 - v_2$ is a solution to $Ax = 0$

Generally, for Non-Square Matrices:

- **m** $>$ **n** (# columns $>$ # rows): The Linear System has no unique solution (infinitely many solutions) / is inconsistent.
- **n** $>$ **m** (# rows $>$ # columns): The Linear System is inconsistent.

Definition 1.0.3 Another way to look at Linear Systems, is to consider them in Matrix Form.

Specifically, we define a general Augmented Matrix as follows:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right]$$

This allows us to also represent Matrices using Matrix Multiplication:

$$AB = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} a_1 B \\ a_2 B \\ \vdots \\ a_m B \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_n \end{bmatrix}$$

1.1 Elementary Row Operations

We can then perform **Elementary Row Operations** (EROs) of 3 kinds:

1. Interchanging: $R_i \leftrightarrow R_j$
2. Rotation: $R_i - R_j$
3. Scaling: aR_i

where $a \neq 0$ and $i \neq j$.

We then say that 2 Augmented Matrices are **Row Equivalent** if they can be obtained from each other by a series of EROs.

Theorem 1.1.1 If two augmented matrices are Row Equivalent → their systems have the same solution.

Matlab Functions

- `addRow(matrix, target, factor, source)`
- `multiplyRow(matrix, factor, target)`
- `divideRow(matrix, factor, target)`
- `swapRow(matrix, target, source)`

1.2 Solutions

A systematic way to obtain the solution(s) to a Linear System, is to

1. Express the Linear System in Augmented Matrix form
2. By performing EROs, row-reduce the Augmented Matrix to Row-Echelon Form (REF) or Reduced Row-Echelon Form (RREF)
3. Decide if the System is Consistent, if Consistent, proceed with the next step
4. (REF) Perform back-subsitution to obtain the unique solution
5. (RREF) Read off the Augmented Matrix to obtain the unique OR general solution

Definition 1.2.1 An Augmented Matrix is in Row-Echelon Form (REF) $\iff$

1. All Zero Rows are at the bottom of the Augmented Matrix
2. The first non-zero entry (***Leading Entry***) in each row is further to the right as we move down the Matrix

Definition 1.2.2 An Augmented Matrix is in Reduced Row-Echelon Form (RREF) $\iff$

1. All Leading Entries are 1
2. In each column containing a Leading Entry, ***(Pivot Column)***, all entries other than the Leading Entry is 0

Suppose the RREF of an Augmented Matrix contains non-pivot columns, we refer to the variable represented by that column as a ***Free Variable***. The presence of a Free Variable is indicative of a General Solution in terms of said Free Variable, and thus, infinitely many solutions to the Linear System.

Theorem 1.2.1 Let v be a particular solution to $Ax = b$, that is $Av = b$, and u be a particular solution to the Homogenous Linear System $Ax = 0$, that is $Au = 0$ with the same coefficient matrix A. Then,

$$v + u \text{ is also a solution to } Ax = b$$

Theorem 1.2.2 Let v_1 and v_2 be solutions to the Linear System $Ax = b$. Then, $v_1 - v_2$ is a solution to the Homogenous Linear System $Ax = 0$ with the same Coefficient Matrix.

1.2.1 Gaussian and Gauss-Jordan Elimination

Gaussian Elimination is used to reduce an Augmented Matrix to REF. The steps are as follows:

1. Bring a row with a non-zero leading entry of the left-most column possible to the top of the matrix
2. For each column below the left-most column found above, make all entries **below** it in the same column to 0 by adding multiples of the top row
3. Repeat the above 2 steps for each subsequent row in the matrix

Gauss-Jordan Elimination is used to reduce an Augmented Matrix **already in REF** into RREF. Typically, this involves performing Gaussian Elimination first, before performing the following steps:

1. Multiply each row by constants to make all Leading Entries 1
2. Starting from the last row, add multiples of the current row to the rows above to make all other entries in the same column 0

Example

Gaussian Elimination:

$$\left[\begin{array}{ccc|c} 0 & 3 & 9 & 0 \\ 1 & 2 & -3 & 0 \\ 4 & 1 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 0 & 3 & 9 & 0 \\ 4 & 1 & 0 & 0 \end{array}\right] \xrightarrow{R_3 - 4R_1} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 0 & 3 & 9 & 0 \\ 0 & -7 & 12 & 0 \end{array}\right] \xrightarrow{R_3 + \frac{7}{3}R_2} \left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 0 & 3 & 9 & 0 \\ 0 & 0 & 33 & 0 \end{array}\right]$$

Gauss-Jordan Elimination:

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 0 & 3 & 9 & 0 \\ 0 & 0 & 33 & 0 \end{array}\right] \xrightarrow{\frac{1}{3}R_2} \frac{1}{33}R_3 \left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \xrightarrow{R_2 - 3R_3} R_1 + 3R_3 \left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \xrightarrow{R_1 - 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

1.2.2 Unknown Coefficients

For Linear Systems with unknown coefficients, it is important to note that we cannot perform EROs without care. For example, we will not be able to perform the $\frac{1}{a}$ ERO as $a = 0$ is possible, in which case, we will have a Zero Division Error.

Therefore, for Linear Systems (or Augmented Matrices) with Unknown Coefficients, we can have multiple cases, each with different solutions, and different number of solutions.

Typically, the cases we consider are those where the Leading Entry in REF contains an unknown coefficient. Note that for this, we will have to be able to convert the Augmented Matrix to REF **without** performing illegal operations given any value of the Unknown Coefficient.

This also indicates that the Gaussian Elimination, and Gauss-Jordan Elimination, do not have to be followed strictly if we are just looking to get the solution(s) of the system.

Example

The Linear System $Ax = b$ is given by: $\left[\begin{array}{ccc|c} 1 & a & 2 & 1 \\ 1 & 2a & 3 & 1 \\ 1 & a & a+3 & 2a^2+1 \end{array}\right] \xrightarrow{EROs} \left[\begin{array}{ccc|c} 1 & a & 2 & 1 \\ 0 & a & 1 & 0 \\ 0 & 0 & a+1 & 2a^2-2 \end{array}\right]$.

We have 2 cases: $a = 0$ and $a = -1$,

$$\text{Let } a = 0, \quad \left[\begin{array}{ccc|c} 1 & a & 2 & 1 \\ 0 & a & 1 & 0 \\ 0 & 0 & a+1 & 2a^2-2 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

$\therefore$ The system has no solution when $a = 0$ since we cannot fulfill $z = 0$ and $z = -2$ simultaneously.

$$\text{Let } a = -1, \quad \left[\begin{array}{ccc|c} 1 & a & 2 & 1 \\ 0 & a & 1 & 0 \\ 0 & 0 & a+1 & 2a^2-2 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$\therefore$ The system has infinitely many solutions when $a = -1$ since there is a non-pivot column on the 3rd column.

$$\text{Finally, let } a \neq -1 \text{ and } a \neq 0, \quad \left[\begin{array}{ccc|c} 1 & a & 2 & 1 \\ 0 & a & 1 & 0 \\ 0 & 0 & a+1 & 2a^2-2 \end{array}\right] \xrightarrow{EROs} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3-2a \\ 0 & 1 & 0 & -\frac{2a-2}{a} \\ 0 & 0 & 1 & 2a-2 \end{array}\right]$$

$\therefore$ The system has exactly 1 solution when $a \neq -1$ and $a \neq 0$.

1.2.3 Solution Sets

Definition 1.2.3 – Terminologies The following are some Terminologies to know of:

- **Solution Set**: The set of all solutions to a Linear System
- **General Solution**: The expression that gives the entire Solution Set

Converting Solutions Sets from Implicit to Explicit

Let the Implicit Expression of a Solution Set be given by:

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + s\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \middle| s, t \in \mathbb{R} \right\}$$

Then, to obtain the Explicit Expression of the Solution Set:

$$\left.\begin{array}{l} x = 1 - 2s + t \quad y = 2 + s \quad z = t - 1 \\ s = y - 2 \\ t = z + 1 \end{array}\right\} \quad \begin{array}{l} x = 1 - 2(y-2) + (z+1) \\ x + 2y - z = 6 \end{array}$$

Thus, we can also express the same Solution Set in the following manner: $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| x + 2y - z = 6 \right\}$.

Example for Solution Space

Let $V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \middle| \begin{cases} x_1 + x_2 + x_3 + 2x_4 = 0 \\ 3x_1 + 2x_2 + 3x_3 + 3x_4 = 0 \\ x_1 + 2x_2 + x_3 + 5x_4 = 0 \end{cases} \right\}$. Then, the Solution Space (Null Space) is given by:

$$\text{RREF of V: } \left(\begin{array}{cccc|c} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \rightarrow \quad x_1 + s - t = 0 \text{ and } x_2 + 3t = 0$$

$\therefore$ The Basis for V is given by: $\left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -3 \\ 1 \end{pmatrix} \right\}$.

Chapter 2

Matrices

Typically, we denote a Matrix with m rows and n columns as m x n matrices, and we can also subscript them: A_{mn} for a matrix A.

There are a variety of different types of matrices, some of which are important enough to get a name:

- Column/Row Vectors: n x 1 / 1 x n matrices
- Zero Matrix: All entries are 0
- Square Matrix: $m = n$
- Diagonal Matrix: All entries besides those along the diagonal is equal to 0, $\forall a \in A_{ij}$, where $i \neq j, a = 0$
- Scalar Matrix: A subset of Diagonal Matrix, where all Diagonal Entries are equal to each other
- Identity Matrix: A subset of Scalar Matrix, where all Diagonal Entries are 1
- Upper / Lower Triangular Matrix: A subset of Square Matrix, where all Entries below / above the Diagonal is 0
- (Strictly) Upper / Lower Triangular Matrix: A subset of Upper / Lower Triangular Matrix, where all Diagonal Entries are 0
- Symmetric Matrix: $\forall a \in A, a_{ij} = a_{ji}$

We can also define Matrix Equality, where two Matrices are equal $\iff$ they have the same size, and each of their corresponding entries are equal. Note that Matrix Equality is different from Row Equivalence.

2.1 Matrix Operations

Similar to normal scalar operations, we can perform Addition, Subtraction and Multiplication between 2 or more Matrices. However, we cannot perform Division.

Definition 2.1.1 Matrix Addition

Let $A = (a_{ij})_{mxn}$ and $B = (b_{ij})_{mxn}$. Then, $A + B = C$ is the m x n matrix whose (i, j)-th entry is defined as follows

$$C_{ij} = a_{ij} + b_{ij}$$

Note that we can also define Matrix Subtraction equivalently

Definition 2.1.2 Matrix Multiplication

Let $A = (a_{ij})_{mxp}$ and $B = (b_{ij})_{pxn}$. Then, $AB = C$ is the m x n matrix whose (i, j)-th entry is defined as follows

$$C_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj}$$

Here, we refer to $C = AB$ as the pre-multiplication of A to B, or equivalently, the post-multiplication of B to A

We also define the following Properties for Matrix Multiplication:

- $A(BC) = (AB)C$
- $(A(B_1 + B_2)) = AB_1 + AB_2$
- $c(AB) = (cA)B = A(cB)$
- $A0 = 0A = 0$
- $AI = IA = A$

Definition 2.1.3 Transpose of a Matrix

The transpose of a Matrix A of size m x n is the n x m matrix $A^T = B$ whose (i, j)-th entry is defined as follows

$$b_{ij} = a_{ji}$$

We also define the following Properties for Transpose Matrices:

- $(A^T)^T = A$
- A is symmetric $\iff A = A^T$
- $c(A)^T = cA^T$
- $(A + B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

2.2 Determinant and Adjoint

Geometrically, the Determinant of a matrix is the area, for $\mathbb{R}^2$, or volume, for $\mathbb{R}^3$ enclosed by the vector lines. This also helps us understand how applying EROs affect the Determinant of a matrix.

Definition 2.2.1 The (i, j)-th Matrix Minor, M_{ij} of a matrix A is the submatrix of A obtained by deleting the i-th row and the j-th column of A

Definition 2.2.2 The (i, j)-th cofactor of a matrix A is the (real) number given by

$$(-1)^{i+j} det(M_{ij})$$

Definition 2.2.3 Let A be a n x n Square Matrix. Then, $adj(A)$ is the n x n Square Matrix whose (i, j)-th entry is the (j, i)-th cofactor of A

General Properties of Determinant and Adjoint:

- $det(A) = det(A^T)$
- $\forall c \in \mathbb{R}, det(cA) = c^n det(A)$, where A is a Square Matrix of order n
- $\forall c \in \mathbb{R}, adj(cA) = c^{n-1} adj(A)$
- $det(AB) = det(A)det(B)$, where A and B are Square Matrices of the same order

Additionally, if A is an invertible matrix of order n, then we can also conclude that:

- $adj(A)$ is Invertible
- $adj(A)^T = adj(A^T)$
- $det(A^{-1}) = [det(A)]^{-1}$
- $A^{-1} = \frac{1}{det(A)} adj(A)$
- $det(I + AB) = det(I + BA)$, assuming B is also a Square Matrix of order n

2.2.1 Finding Determinant

Definition 2.2.4 There are a number of base cases and quick calculation definitions for Determinant of a Matrix, as described below:

($n = 1$)

$$(a) \rightarrow det(A) = a$$

($n = 2$)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow det(A) = ad - bc$$

(Triangular Matrices)

$$\text{Let R be the RREF of A, then } det(R) = \prod_{k=1}^{n} a_{ii}$$

Theorem 2.2.1 If the Square Matrix A has a Zero Row $\rightarrow det(A) = 0$

Theorem 2.2.2 If two matrices A and B are Row-Equivalent $\rightarrow det(A) = 0 \iff det(B) = 0$

Cofactor Expansion

Cofactor Expansion is mainly useful when there are many 0s in a single row/column, such that we can ignore many Matrix Minors, OR when there are unknowns in A.

Additionally, note the formula for Adjoint. If the question also requires finding Adjoint, it may be faster to simply use Cofactor Expansion rather than Reduction to compute the Matrix Minor at the same time.

Using Cofactor Expansion, the determinant of a matrix A is defined as:

$$\begin{aligned} det(A) &= a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} = \sum_{k=1}^{n} a_{ik}A_{ik} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj} = \sum_{k=1}^{n} a_{kj}A_{kj}. \end{aligned}$$

Reduction

Reduction is mainly useful when the matrix is relatively large, and has little zero entries.

Performing EROs on a matrix alters its determinant based on a set of rules:

- Swapping ($R_i \leftrightarrow R_j$): $det(B) = -det(A)$
- Scaling (cR_i): $det(B) = c\ det(A)$
- Rotation ($R_i + cR_j$): $det(B) = det(A)$

Then, suppose R is the RREF of A, $det(R) = det(E_k) \ldots det(E_2)det(E_1)det(A)$ and thus,

$$det(A) = \frac{det(R)}{det(E_k) \ldots det(E_2)det(E_1)}$$

To find $det(R)$, remember that the determinant of a Triangular Matrix is the product of all its diagonal entries.

2.3 Inverse of a Matrix

Definition 2.3.1 A n x n square matrix A is invertible $\iff \exists$ a matrix B s.t. $AB = I_n = BA$.

- B is also referred to as the Inverse of the matrix A, aka. A^{-1}
  - Geometrically, B is the matrix which reverses the Linear Transformation done on the Original Matrix by A, that is B applies the reverse of the Linear Transformation applied by A
- If A is a Square Matrix, and is not Invertible, $\rightarrow A$ is Singular

Theorem 2.3.1 If A is Invertible $\rightarrow$ Its inverse, $B = A^{-1}$, is unique. That is,

$$C_1A = C_2A \rightarrow C_1 = C_2 \text{ AND } AB_1 = AB_2 \rightarrow B_1 = B_2$$

We have the following properties of Invertible Matrices:

- $(cA)^{-1} = \frac{1}{c}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$
- $(A^{-1})^{-1} = A$
- $A^{-k} = (A^{-1})^k = (A^k)^{-1}$
- $(A_1A_2 \dots A_k)^{-1} = A_k^{-1} \dots A_2^{-1}A_1^{-1}$
- If matrices A and B are invertible matrices of order n, then AB is invertible with inverse: $(AB)^{-1} = B^{-1}A^{-1}$

Algorithm to Compute Inverse

Since the Inverse of a Matrix is unique, we get that $AX = I_n$. Therefore, we are essentially solving:

$$(A \mid I) \xrightarrow{RREF} (I \mid A^{-1})$$

which is the equivalent of premultiplying the matrix B, which is the Product of the Elementary Matrices that reduces A to I_n, to I_n ($BI_n = B = A^{-1}$). Specifically,

$$\begin{aligned} BA = I_n &\rightarrow (E_k \dots E_2E_1)A \\ BI_n = A^{-1} &\rightarrow (E_k \dots E_2E_1)I_n \\ B = A^{-1} &= (E_k \dots E_2E_1). \end{aligned}$$

Additionally, note that if the system, $(I \mid A^{-1})$, is inconsistent, then we can conclude that A cannot be row-reduced to I_n, and thus, A does not have an inverse, and A is not invertible.

2.4 Elementary Matrices

Definition 2.4.1 A Square Matrix E of order n is an Elementary Matrix if it can be obtained from the Identity Matrix I_n by performing a single ERO

- Performing EROs on a matrix A is equivalent to pre-multiplying the corresponding Elementary Matrix to A.
- The Inverse of an Elementary Matrix is the Elementary Matrix corresponding to the reverse ERO
- Every Elementary Matrix is Invertible

Suppose B is Row Equivalent to A,

$$A \xrightarrow{r_1}\xrightarrow{r_2} \dots \xrightarrow{r_k} B$$

and let E_i be the Elementary Matrix corresponding to the r_i-th Row Operation for $i = 1, 2, \dots, k$. Then,

$$\begin{aligned} B &= E_k \dots E_2E_1A \\ A &= (E_k \dots E_2E_1)^{-1}B \\ A &= E_1^{-1}E_2^{-1} \dots E_k^{-1}B \end{aligned}$$

Theorem 2.4.1 If $E_1 E_2 \ldots E_k$ are all Elementary Matrices corresponding to EROs of the type $R_i + cR_j$ for $i > j \rightarrow L = E_1 E_2 \ldots E_k$ must be a Unit Lower Triangular Matrix

Theorem 2.4.2 If $E_1 E_2 \ldots E_k$ follows Gaussian Elimination **strictly**, and L is a Unit Lower Triangular Matrix, then the (i, j)-th entry of L is given by $-c_l$ for each Row Operation $r_l = R_i + cR_j$ for $i > j$ and $c_l \in \mathbb{R}$.

2.5 LU Factorization

When we reduce a matrix A to REF, we can express A in terms of two other matrices, L and U, where

- L is the product of Elementary Matrices
- U is the REF of A

We can thus arrange $Ax = b$ into the following equations:

$$\left.\begin{array}{l} A = LU \\ Ax = b \end{array}\right\} LUx = b \xrightarrow{Ux=y} Ly = b.$$

Theorem 2.5.1 The determinant of LU Factorisable Matrices A is thus given by:

$$det(A) = det(L)det(U) = det(U)$$

LU Factorization is mainly helpful in scenarios whereby A may be large and we have to compute $Ax = b$ multiple times, for different b.

Chapter 3

Vector Spaces

Definition 3.0.1 A (real) n-vector is a collection of n ordered real numbers.

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

where $r_i \in \mathbb{R}$ for $i = 1, \ldots, n$

We can visualise vectors as either

1. An arrow with the tail at the origin, and the head at the coordinates denoted by v, OR
2. A point in the Euclidean n-space at the coordinates denoted by v

Definition 3.0.2 The Euclidean n-space $\mathbb{R}^n$ is the collection of all n-vectors.

$$\mathbb{R}^n = \left\{ v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \middle| v_i \in \mathbb{R} \right\}$$

- The solution set to a Linear System $Ax = b$ is a subset in $\mathbb{R}^n$

3.1 Linear Combination

Definition 3.1.1 Let $u_1, u_2, \ldots, u_n$ be vectors in $\mathbb{R}^n$. Then, a linear combination of these vectors, v can be expressed as:

$$v = c_1u_1 + c_2u_2 + \cdots + c_nu_n$$

Algorithm to check Linear Combination

To check if a vector v is a Linear Combination of another vector(s) $u_1, u_2, \ldots, u_n$, check if the Linear System $Ax = v$ where $A = (u_1, u_2, \ldots, u_n)$ is consistent.

3.1.1 Linear Independence

Definition 3.1.2 Let $S = u_1, u_2, \ldots, u_n \subseteq \mathbb{R}^n$. Then, S is Linearly Independent $\iff$

$$c_1u_1 + c_2u_2 + \ldots c_nu_n = 0 \text{has only the trivial solution}$$

More intuitively, S is Linearly Independent $\iff$ none of the vectors in S are Linear Combinations of some other vector(s) in S.

We can also say that a vector v, which is a Linear Combination of some other vector(s) in S, is Redundant.

Algorithm to check Linear Independence

To check if $A = u_1, u_2, \ldots, u_n \subseteq \mathbb{R}^n$ is Linearly Independent, check if the RREF of the Linear System $Ax = 0$ has any non-pivot columns.

- If there are non-pivot columns $\rightarrow$ there is a General Solution $\rightarrow$ infinitely many solutions $\rightarrow$ A is Linearly Dependent.

General Properties of Linear Independence, given $S = \{u_1, u_2, \ldots, u_k\} \subseteq \mathbb{R}^n$:

- $0 \in S \rightarrow S$ is linearly dependent
- If $k > n \rightarrow S$ is linearly dependent
- S is Linearly Independent $\iff$ S spans $\mathbb{R}^n$

Theorem 3.1.1 Let S_1, S_2 be finite subsets of $\mathbb{R}^n$ such that $S_1 \subseteq S_2$,

- S_1 is linearly dependent $\rightarrow S_2$ is linearly dependent
- S_2 is linearly independent $\rightarrow S_1$ is linearly independent

Theorem 3.1.2 Linear Independence Theorem

Let V be a subspace of $\mathbb{R}^n$ and $S = \{u_1, u_2, \ldots, u_k\}$ is a Linearly Independent subset of $V, S \subseteq V$. Then, $\exists$ a set $T \subseteq S$ such that T is a Basis for V.

3.2 Linear Span

Definition 3.2.1 The Linear Span of $S = u_1, u_2, \ldots, u_n$ is the subset of $\mathbb{R}^n$ containing all Linear Combinations of $u_1, u_2, \ldots, u_n$.

$$span u_1, u_2, \ldots, u_n = span(S) = \{\ c_1u_1, c_2u_2, \ldots, c_nu_n \mid c_1, c_2, \ldots, c_n \in \mathbb{R}\}$$

A Linear Span S should also satisfy the following properties:

1. S contains the Zero Vector
2. S is closed under Scalar Multiplication
3. S is closed under Vector Addition

that is, a Linear Span contains the Zero Vector, and is Closed under Linear Combination.

General Properties of Linear Span, given $S = u_1, u_2, \ldots, u_k \subseteq \mathbb{R}^n$:

- If $k < n \rightarrow span(S) \neq \mathbb{R}^n$
- The Zero Vector in $\mathbb{R}^n, 0 \in span(S) \forall S$

Theorem 3.2.1 Two Linear Spans, S and T, are equal $\iff$ $span(S) \subseteq span(T) \land span(T) \subseteq span(S)$, that is, $\forall u \in T, u$ is a Linear Combination of vectors in $S \land \forall v \in S, v$ is a Linear Combination of vectors in T.

Algorithm to check Linear Spans and Subsets

- To check if a vector v is in $span(S), S = \{u_1, u_2, \ldots, u_n\}$, we check if v is a Linear Combination of the vectors in S
- To check if $span(S) \subseteq span(T)$, we can check if $(S \mid T)$ is consistent
- To check if $span(S) = \mathbb{R}^n$, we can check if $(S \mid x)$, where $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ is consistent

Theorem 3.2.2 Let $S = \{u_1, u_2, \ldots, u_k\} \subseteq \mathbb{R}^n$ and $A = (v_1 \ldots v_k) \rightarrow Ax = v$ is consistent $\iff$ $v \in span(S)$

Theorem 3.2.3 Spanning Set Theorem

Let $S = \{u_1, u_2, \ldots, u_k\}$ be a subset of vectors in $\mathbb{R}^n$ and let $V = span(S)$. Suppose $V \neq 0$, then $\exists$ a subset of S that is a basis for V

Generally, there is no relationship between Linear Independence and Spanning a Vector Space. However, if $S \subseteq \mathbb{R}^n$ contains exactly n vectors $\rightarrow$ Linear Independence is equivalent to spanning $\mathbb{R}^n$

3.3 Subspace

Definition 3.3.1 A subset $V \in \mathbb{R}^n$ is a subspace $\iff$ it satisfies the following properties:

1. V contains the Zero Vector
2. V is closed under Scalar Multiplication
3. V is closed under Vector Addition

Geometrically, a subset of the Vector Space (typically $\mathbb{R}^n$) is a Subspace of said Vector Space $\iff$ all Linear Combinations of vectors in the Subspace are also in the Subspace (Closure under Linear Combination) and the Origin is in the Subspace.

Notice the similarity between Subspaces and Linear Spans, specifically, that both contain the Zero Vector, and are closed under Linear Combinations.

Theorem 3.3.1 The Linear Span of any set of vectors is a valid subspace of $\mathbb{R}^n$.

That is, a subset $V \subseteq \mathbb{R}^n$ is a Subspace $\iff V = span(S)$ for some finite set $S = \{u_1, u_2, \ldots, u_k\}$

To show V is a Subspace, either:

1. Find a Spanning Set, that is a Set S s.t. $V = span(S)$, OR
2. Show that V satisfies the 3 properties

To show V is not a Subspace:

1. Show that V does not satisfy any of the 3 properties

Definition 3.3.2 – Solution and Affine Space The following are definitions for Solution and Affine Space

- **Solution Space**: The Solution Set $V = \{u | Au = b\}$ to a Homogenous Linear System $Ax = b$
- **Affine Space**: The Solution Set $V = \{u | Au = b\}$ to a Non-homogenous Linear System $Ax = b$

3.4 Basis

Definition 3.4.1 Let V be a subspace of $\mathbb{R}^n$. The Set $S = \{u_1, u_2, \ldots, u_k\} \subseteq V \rightarrow$

- S spans V, that is, $span(S) \subseteq V$, AND
- S is Linearly Independent

- A subspace V can have > 1 Basis
- Basis is a Maximal Linearly Independent subset of its subspace V
- Basis is a Minimal Spanning Set for its subspace V

Theorem 3.4.1 Let $S = \{u_1, u_2, \ldots, u_k\} \subseteq \mathbb{R}^n$ and $A = (u_1 u_2 \ldots u_k)$ be the matrix whose columns are vectors in S. Then,

S is a Basis for $\mathbb{R}^n \iff k = n \wedge A$ is Invertible

Equivalent Methods to check for Basis, given V is a k-dimensional subspace of $\mathbb{R}^n, dim(V) = k$:

1. Suppose S is a Linearly Independent subset of V containing k vectors $\rightarrow S$ is a basis for V.
2. Suppose S contains k vectors such that $V \subseteq span(S) \rightarrow S$ is a basis for V

3.4.1 Coordinates Relative to a Basis

Note that typically, we talk about coordinates in terms of the Standard Basis, that is where the unit vector lies along each of the planes. Given that the basis for a subspace is not unique, we can then imagine that the same point in the subspace has different coordinates depending on what basis we are looking at. Coordinates Relative to a Basis then allows us to describe these points relative to each Basis a subspace has.

Definition 3.4.2 Let $S = \{u_1, u_2, \ldots, u_k\}$ be a Basis for V, a subspace of $\mathbb{R}^n$. Then,

$$\forall \text{ vectors } v \in V, v = c_1u_1 + c_2u_2 + \cdots + c_ku_k$$

We can then define the Coordinates of each vector v relative to the basis S as the Coefficients of the Linear Combination,

$$[v]_s = x = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

General Properties for Coordinates Relative to a Basis, given V is a Subspace of $\mathbb{R}^n$ and B is a Basis for V:

- $\forall$ vectors $u, v \in V, u = v \iff [u]_B = [v]_B$
- $\forall v_1, v_2, \ldots, v_m \in V, [c_1v_1 + c_2v_2 + \cdots + c_mv_m]_B = c_1[v_1]_B + c_2[v_2]_B + \cdots + c_m[v_m]_B$

Now, additionally suppose B contains k vectors, that is $|B| = k$, then:

- $v_1, v_2, \ldots, v_m$ is Linearly Independent / Dependent $\iff$ $[v_1]_B, [v_2]_B, \ldots, [v_m]_B$ is Linearly Independent / Dependent (respectively)
- $\{v_1, v_2, \ldots, v_m\}$ spans V $\iff$ $\{[v_1]_B, [v_2]_B, \ldots, [v_m]_B\}$ spans $\mathbb{R}^n$
- If $S \subseteq V$ with $m > k \to S$ is Linearly Dependent
- If $S \subseteq V$ with $m < k \to S$ cannot span V

Algorithm to Computing Coordinates Relative to a Basis

To find the Coordinates of a vector v relative to a basis S, we have to solve $v = c_1u_1 + c_2u_2 + \cdots + c_ku_k$, and find the coefficients $c_1u_1 + c_2u_2 + \cdots + c_ku_k$. This is equivalent to solving the Linear System:

$$(u_1u_2 \ldots u_k \mid v)$$

3.5 Dimensions

Definition 3.5.1 Let V be a subspace for $\mathbb{R}^n$. Then, the dimension of $V, dim(V)$, is the number of vectors in any basis of V.

Therefore, we can also conclude that $dim(S) = dim(T) \to$ both S and T are Basis for the same subspace $V \subseteq \mathbb{R}^n$

Let A be a m x n matrix. Then, the dimension of its solution Space, $V = \{u \in \mathbb{R}^n | Au = 0\}$ is given by:

$$\begin{aligned} dim(V) &= \text{the number of non-pivot columns in the RREF of } A \\ &= \text{the number of parameters in the General Solution of } A \end{aligned}$$

General Properties of Dimensions, given U and V be a subspace of $\mathbb{R}^n$:

- If $U \subseteq V \to dim(U) \leq dim(V)$
- If $U \subsetneq V \to dim(U) < dim(V)$

3.6 Transition Matrices

Similar to the underlying concept of 'Coordinates Relative to a Basis', Transition Matrices tell us how to convert relative coordinates to those of a different Basis.

Definition 3.6.1 Let V be a subspace of $\mathbb{R}^n$. Suppose $S = \{u_1, \ldots, u_k\}$ and $T = \{v_1, \ldots, v_k\}$ are both Basis for the same subspace V. Then, the Transition Matrix from T to S, P is given by:

$$P = ([v_1]_S\ [v_2]_S\ \ldots\ [v_k]_S),$$

and the Inverse of the Transition Matrix, P^{-1} is the Transition Matrix from S to T.

Theorem 3.6.1 Let V be a subspace of $\mathbb{R}^n$. Suppose $S = \{u_1, \ldots, u_k\}$ and $T = \{v_1, \ldots, v_k\}$ are both Basis for the same subspace V, and P is the Transition Matrix from T to S. Then,

$$\forall w \in V, [w]_s = P[w]_T$$

Algorithm to find Transition Matrix

Let $S = \{u_1, \ldots, u_k\}$ and $T = \{v_1, \ldots, v_k\}$ are both Basis for the same subspace V in $\mathbb{R}^n$. Then, to find the Transition Matrix P from T to S,

$$(S \mid T) \to (u_1\ u_2\ \ldots\ u_k | v_1\ v_2\ \ldots\ v_k) \to \left(\begin{array}{c|cccc} I_k & [v_1]_S & [v_2]_S & \ldots & [v_k]_S \\ 0_{(n-k)xk} & 0 & 0 & \ldots & 0 \end{array}\right) \to \left(\begin{array}{c|c} I_k & P \\ 0_{(n-k)xk} & 0_{(n-k)xk} \end{array}\right)$$

3.6.1 Example

Given vectors in 1 basis as linear combinations of vectors in the other basis,

Let $S = \{u_1, u_2, u_3\}$ and $T = \{v_1, v_2, v_3\}$ be 2 bases for a Vector Space V where:

$$v_1 = u_1 + u_2, \quad v_2 = -u_1 + 2u_3, \quad v_3 = u_1 + u_2 + u_3$$

Then, the Transition Matrix from T to S is given by:

$$[v_1]_S = (1\ \ 1\ \ 0)^T \quad [v_2]_S = (-1\ \ 0\ \ 2)^T \quad [v_3]_S = (1\ \ 1\ \ 1)^T$$

$$\begin{aligned} \therefore \text{Transition Matrix from } T \text{ to } S &= P \\ &= ([v_1]_S\ \ [v_2]_S)\ \ [v_3]_S) \\ &= \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \end{aligned}$$

Given the Transition Matrix and 1 Basis,

Given $P = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 4 & 1 \\ 3 & 8 & -2 \end{pmatrix}$ and $S = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \\ 1 & 3 & 4 \end{pmatrix}$, where S and $T = \{v_1, v_2, v_3\}$ is a Basis for V and P is the Transition Matrix from T to S.

$$\begin{aligned} T = SP &= \begin{pmatrix} 1 & 2 & -2 \\ 1 & 0 & 3 \\ 1 & 2 & 0 \\ 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 4 & 1 \\ 3 & 8 & -2 \end{pmatrix} \\ &= \begin{pmatrix} -1 & -7 & 8 \\ 10 & 25 & -4 \\ 5 & 9 & 4 \\ 19 & 45 & -3 \end{pmatrix} \end{aligned}$$

3.7 Column / Row / Null Space

Definition 3.7.1 – Column and Row Space Given an m x n matrix A: $\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$,

- The Row Space of $A, Row(A)$ is the subspace of $\mathbb{R}^n$ spanned by the rows of A

$$Row(A) = span\{(a_{11} \;\; a_{12} \; ... \; a_{1n}), (a_{21} \;\; a_{22} \; ... \; a_{2n}), (a_{m1} \;\; a_{m2} \; ... \; a_{mn})\}$$

- The Column Space of $A, Col(A)$ is the subspace of $\mathbb{R}^m$ spanned by the columns of A

$$\text{Col}(A) = \text{span}\left\{ \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \right\}$$

General Properties of Column / Row Space:

- $Col(AB) \subseteq Col(A)$

Theorem 3.7.1 Row Operations preserve Row Space

Suppose A and B are Row Equivalent Matrices, then

$$Row(A) = Row(B)$$

Theorem 3.7.2 Row Operations preserve Linear Relations between Columns

Let $A = (a_1 \; a_2 \; \dots \; a_n)$ and $B = (b_1 \; b_2 \; \dots \; b_n)$ be Row-Equivalent m x n matrices. Then,

$$c_1a_1 + c2a_2 + \cdots + c_na_n = 0 \iff c_1b_1 + c2b_2 + \cdots + c_nb_n = 0$$

- Linear Independence of Column Vectors in a matrix A is not affected by EROs
- Rank of a matrix A is not affected by EROs

However,

- Row Operations do not preserve Column Space
- Row Operations do not preserve Linear Relations between Rows

Definition 3.7.2 – Null Space The Null Space of a m x n matrix A is the Solution Space to the Homogenous Linear System $Ax = 0$.

$$Null(A) = \{v \in \mathbb{R}^n \mid Av = 0\}$$

Geometrically, the Null Space of A refers to the set of vectors which become 0 when the Linear Transformation denoted by A is applied on said vectors.

Definition 3.7.3 The nullity of A is the dimension of the null space of A, denoted as

$$nullity(A) = dim(Null(A))$$

Theorem 3.7.3 Let A be a m x n matrix $\rightarrow Null(A) = Null(A^TA)$

General Properties of Null Space, given two Square Matrices, A and B of the same order:

- $Null(B)$ is a Subspace of $Null(AB)$
- $nullity(A) + nullity(B) \geq nullity(AB)$

- If A and B are Row-Equivalent Matrices, then A and B have the same Null Space

3.7.1 Basis for Col(A), Row(A), Null(A)

The Basis for the Column/Row Space can be found by identifying the Linearly Independent Column/Row vectors.

Let A be the original m x n matrix, and R be the RREF of A:

- Basis for Column Space: Columns of A corresponding to the Pivot Columns in R
- Basis for Row Space: Non-zero rows of R
- Basis for Null Space: Basis of the Solution Space of A

Example

Let $A = \begin{pmatrix} 1 & 2 & 8 & -3 & 1 & 1 \\ 1 & 2 & 8 & -3 & 0 & 4 \\ 1 & 3 & 11 & -4 & 3 & -4 \\ 1 & 1 & 5 & -2 & 1 & 0 \end{pmatrix}$, and thus the RREF of A is given by $R = \begin{pmatrix} 1 & 0 & 2 & -1 & 0 & 2 \\ 0 & 1 & 3 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Then, the Basis for the Row Space, the Column Space and the Null Space of A is given as follows:

$$\text{Basis for Row Space of } A = \left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 3 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -3 \end{pmatrix} \right\}$$

$$\text{Basis for Column Space of } A = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \\ 2 \end{pmatrix} \right\}$$

$$\text{Basis for Null Space of } A = \left\{ \begin{pmatrix} -2 \\ -1 \\ 0 \\ 0 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

Additionally, to extend the Basis of the Row Space of A to $\mathbb{R}^6$, we can add the following vectors:

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

, each vector containing a 1 at each non-pivot column.

3.8 Rank

Definition 3.8.1 The rank of the m x n matrix A is the dimension of its column / row space.

$$rank(A) = dim(Col(A)) = dim(Row(A))$$

Here, note the following to see that $dim(Col(A)) = dim(Row(A))$ is always true:

$$\begin{aligned} dim(Col(A)) &= \text{number of pivot columns in RREF of A} \\ &= \text{number of leading entries in RREF of A} \\ &= \text{number of linearly independent columns of A} \\ &= \text{number of non-zero rows in RREF of A } = dim(Row(A)) \end{aligned}$$

General Properties of Rank, given m x n matrix A and a n x p matrix B:

- $rank(A) = rank(A^T)$
- $rank(A) = 0 \iff A = 0$
- $rank(AB) \leq min\{rank(A), rank(B)\}$
- A and B are Row-Equivalent $\rightarrow rank(A) = rank(B)$
- If A and B are the same size, $rank(A + B) \leq rank(A) + rank(B)$

Theorem 3.8.1 The Linear System $Ax = b$ is consistent $\iff$ the rank of A is equal to the rank of the Augmented Matrix $(A \mid B)$

$$rank(A) = rank((A \mid B))$$

Theorem 3.8.2 Rank-Nullity Theorem

Let A be a m x n matrix, then

$$rank(A) + nullity(A) = n$$

Definition 3.8.2 – Full Rank A m x n matrix A is Full Rank $\rightarrow rank(A) = min\{m, n\}$

Chapter 4

Orthogonality

4.1 Formulas

Given two vectors $u = (u_1, u_2, \dots, u_n)$ and $v = (v_1, v_2, \dots, v_n)$:

Definition 4.1.1 – Dot Product The Dot Product, aka. Inner Product, of u and v is:

$$u \cdot v = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

- If u and v are Column / Row Vectors, we can also view Dot Products as: $u \cdot v = u^Tv$

Definition 4.1.2 – Norm The Norm, aka. Length, of v is given by:

$$||v|| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

- We call a vector x as a Unit Vector $\iff ||x|| = 1$, that is, it has a length of 1
- We can normalise a vector x by the following formula:

$$y = \frac{x}{||x||}$$

  where y is the normalised unit vector of x.

Definition 4.1.3 – Distance between 2 vectors The distance between u and v is given by:

$$d(u, v) = ||u - v|| = \sqrt{\sum_{i=1}^{n}(u_i - v_i)^2}$$

Definition 4.1.4 – Angle between 2 Vectors The angle between 2 vectors, u and v, is given by:

$$\theta = cos^{-1}(\frac{u \cdot v}{||u||||v||})$$

where $0° \leq \theta \leq 180°$

General Properties:

- $|u \cdot v| \leq ||u||||v||$ (Cauchy-Schwarz Inequality)
- $||u + v|| \leq ||u|| + ||v||$ (Triangle Inequality)
- $d(u, w) \leq d(u, v) + d(v, w)$ (Triangle Inequality)
- $\forall c \in \mathbb{R},\ ||cv|| = |c|\ ||v||$
- $v \cdot v \geq 0 \land v \cdot v = 0 \iff v = 0$

4.2 Orthogonal and Orthonormal

Definition 4.2.1 – Orthogonal Vectors Two vectors, u and $v \in \mathbb{R}^n$ are Orthogonal $\iff$

$$u \cdot v = 0$$

that is, 2 vectors are orthogonal if they are Perpendicular to each other.

Definition 4.2.2 – Orthonormal Vectors Two vectors $u, v \in \mathbb{R}^n$ are called orthonormal if and only if

$$\begin{aligned} u \cdot v = 0 \quad & \text{(is orthogonal)} \\ \|u\| = 1 \quad \text{and} \quad \|v\| = 1 \quad & \text{(is a Unit Vector).} \end{aligned}$$

Alternatively, a set of vectors $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ is orthonormal if

$$v_i \cdot v_j = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

Orthogonal / Orthonormal Sets

For a set $S = (u_1, u_2, \ldots, u_n)$ to be Orthogonal / Orthonormal,

$$\forall u, v \in S, u \text{ and } v \text{ are Orthogonal / Orthonormal to each other.}$$

that is, the vectors in S are Pairwise Orthogonal / Orthonormal.

Theorem 4.2.1 Let $S = (u_1, u_2, \ldots, u_n)$ be an Orthogonal / Orthonormal Set, where $\forall u \in \mathbb{R}^n, u \neq 0$. Then, S is Linearly Independent.

Definition 4.2.3 – Orthogonal to a Subspace Let V be a subspace of $\mathbb{R}^n$ and $S = \{u_1, u_2, \ldots, u_k\}$ be a Spanning Set for $V, span(S) = V$.

Then, a vector w is Orthogonal to the subspace $V \iff w \cdot u_i = 0 \;\; \forall i = 1, \ldots, k$.

Definition 4.2.4 – Orthogonal Complement Let V be a subspace of $\mathbb{R}^n$. The Orthogonal Complement of V is the set of all vectors that are orthogonal to V and is denoted as

$$V^{\perp} = \{w \in \mathbb{R}^n | w \cdot v = 0 \;\; \forall v \in V\}$$

Theorem 4.2.2 Let V be a subspace of $\mathbb{R}^n$ and $S = \{u_1, u_2, \ldots, u_k\}$ be a Spanning Set for $V, span(S) = V$. Additionally, Let $A = (u_1 \;\; u_2 \;\; \ldots \;\; u_k)$.

Then, the Orthogonal Complement of V is in the Null Space of A^T:

$$V^{\perp} = w \in Null(A^T)$$

Definition 4.2.5 A n x n Square matrix A is Orthogonal if $A^T = A^{-1}$, and thus, $A^T A = I_n = AA^T$

The following statements are thus equivalent:

- A is an Orthogonal Matrix of Order n
- The Columns of A form an Orthonormal Basis for $\mathbb{R}^n$
- The Rows of A form an Orthonormal Basis for $\mathbb{R}^n$
- $A^{-1} = A^T$
- $A^T A = I_n$
- $AA^T = I_n$

General Properties:

- $Row(A)^{\perp} = Null(A)$
- If 2 vectors are orthogonal, they are not in the span of each other
- $(V^{\perp})^{\perp} = V$

Orthogonal / Orthonormal Basis

Definition 4.2.6 Let V be a subspace of $\mathbb{R}^n$ and S be a Basis for V

- S is an Orthogonal Basis if S is an Orthogonal Set
- S is an Orthonormal Basis if S is an Orthonormal Set

Theorem 4.2.3 – Coordinates Relative to an Orthogonal Basis Let $S = \{u_1, u_2, \dots, u_k\}$ be an Orthogonal Basis for a Subspace V of $\mathbb{R}^n$. Then,

$$\forall v \in V, \quad v = \left(\frac{v \cdot u_1}{||u_1^2||}\right) u_1 + \left(\frac{v \cdot u_2}{||u_2^2||}\right) u_2 + \cdots + \left(\frac{v \cdot u_k}{||u_k^2||}\right) u_k$$

Theorem 4.2.4 – Coordinates Relative to an Orthonormal Basis Let $S = \{u_1, u_2, \dots, u_k\}$ be an Orthonormal Basis for a Subspace V of $\mathbb{R}^n$. Then,

$$\forall v \in V, \quad v = (v \cdot u_1)u_1 + (v \cdot u_2)u_2 + \cdots + (v \cdot u_k)u_k$$

Thus, the coordinates of v relative to S are:

$$\begin{array}{ccc} & S \text{ is Orthogonal:} & S \text{ is Orthonormal:} \\ [v]_S = & \begin{pmatrix} \frac{v \cdot u_1}{\|u_1\|^2} \\ \frac{v \cdot u_2}{\|u_2\|^2} \\ \vdots \\ \frac{v \cdot u_k}{\|u_k\|^2} \end{pmatrix} & \begin{pmatrix} v \cdot u_1 \\ v \cdot u_2 \\ \vdots \\ v \cdot u_k \end{pmatrix} \end{array}$$

4.3 Gram-Schmidt Process

The Gram-Schmidt Process is a method to convert a **Linearly Independent** Set S into an Orthogonal Set, which we can then also convert into an Orthonormal Set by Normalizing all vectors in S.

Here, note that if S is also a Basis for a subspace V, then the resulting Orthogonal / Orthonormal Set is also a Basis for V.

The Steps for the Gram-Schmidt Process to convert a Linearly Independent Set $S = \{u_1, u_2, \dots, u_k\}$ into an Orthogonal Set $V = \{v_1, v_2, \dots, v_k\}$ is as follows:

1. Let $v_1 = u_1$
2. Let $v_{i+1} = u_{i+1} - u'_{i+1}$, where u'_{i+1} is the projection of u_{i+1} onto the Span of S

More Concretely,

$$
\begin{aligned}
v_1 &= u_1, \\
v_2 &= u_2 - \frac{u_2 \cdot v_1}{\|v_1\|^2} v_1, \\
v_3 &= u_3 - \frac{u_3 \cdot v_1}{\|v_1\|^2} v_1 - \frac{u_3 \cdot v_2}{\|v_2\|^2} v_2, \\
v_4 &= u_4 - \frac{u_4 \cdot v_1}{\|v_1\|^2} v_1 - \frac{u_4 \cdot v_2}{\|v_2\|^2} v_2 - \frac{u_4 \cdot v_3}{\|v_3\|^2} v_3, \\
&\vdots \\
v_i &= u_i - \sum_{j=1}^{i-1} \frac{u_i \cdot v_j}{\|v_j\|^2} v_j, \quad \text{for } i = 1, 2, \ldots, k.
\end{aligned}
$$

The Set $V = \{v_1, v_2, \ldots, v_k\}$ is thus an Orthogonal Set (of Non-zero vectors), where each vector v_i in this process is orthogonal to all preceding vectors $v_1, v_2, \ldots, v_{i-1}$.

To convert V into an Orthonormal Set, V', normalize all vectors within the Set V, such that:

$$V' = \left\{ \frac{v_1}{||v_1}, \frac{v_2}{||v_2}, \ldots, \frac{v_k}{||v_k} \right\}$$

V' is thus an Orthonormal Set s.t. $span(V') = span(S)$

Matlab

To perform the Gram-Schmdit Process on a matrix A, we can use the following functions:

```
orthogonal_matrix = gsp(A)
orthonormal_matrix = gsp_norm(A)
```

4.4 QR Factorization

Theorem 4.4.1 Suppose A is a m x n matrix with Linearly Independent columns. Then, A can be decomposed into two Matrices, Q and R, where

- Q is a m x n matrix s.t. the columns of Q form an Orthonormal Set and thus, $Q^T Q = I_n$
- R is an Invertible Upper Triangular Matrix with **positive** diagonal entries

Therefore, the following equations hold:

$$A^T A = (QR)^T QR = R^T Q^T QR = R^T R$$
$$Q^T QR = Q^T A = R$$

- R is the Transition Matrix from A to Q

Algorithm to decompose A into its QR Factorization

Specifically, let A be a m x n matrix with Linearly Independent Columns, to compute Q and R,

1. Perform the Gram-Schmidt Process on the columns of A to obtain the **Orthonormal** Set $\{q_1, q_2, \ldots, q_n\}$
2. Let $Q = (q_1 \ \ q_2 \ \ \ldots \ \ q_n)$
3. Compute $R = Q^T A = Q^T QR$

4.5 Orthogonal Projection

Theorem 4.5.1 – Orthogonal Projection Theorem Let V be a subspace of $\mathbb{R}^n$. Then, every vector $w \in \mathbb{R}^n$ can be decomposed uniquely as the sum:

$$w = w_p + w_n$$

where w_n is Orthogonal to V and w_p is a vector in V.

Then, the Orthogonal Projection of w onto the subspace V, w_p is given by:

$$w_p = \frac{w \cdot u_1}{u_1 \cdot u_1}u_1 + \frac{w \cdot u_2}{u_2 \cdot u_2}u_2 + \cdots + \frac{w \cdot u_k}{u_k \cdot u_k}u_k$$

where $S = \{u_1, u_2, \ldots, u_k\}$ is an Orthogonal Basis for V.

Equivalently, let $A = (v_1 \;\; v_2 \;\; ... \;\; v_n)$, then the Projection w_p is also given by:

$$p = A(A^T A)^{-1} A^T w$$

Notice that we can extend the subspace V from $\mathbb{R}^n$ to $\mathbb{R}^{n+1}$ by adding $w_n = w - w_p$ to V

Note:-

We can also solve for the projection of a vector b onto a subspace $V = span\{a_1, a_2, \ldots, a_n\}$ without having an Orthogonal Basis by using Least Square Approximation.

4.6 Least Square Approximation

Theorem 4.6.1 – Best Approximation Theorem Let V be a subspace of $\mathbb{R}^n$ and a vector $w \in \mathbb{R}^n$. Let w_p be the Projection of w onto V.

Then, w_p is the vector in V closest to w. That is,

$$||w - w_p|| \leq ||w - v|| \;\; \forall v \in V$$

Definition 4.6.1 – Least Square Approximation Let A be a m x n matrix and a vector $b \in \mathbb{R}^m$. A vector $u \in \mathbb{R}^n$ is a Least Square Solution of $Ax = b$ if $\forall v \in \mathbb{R}^n$,

$$||Au - b|| \leq ||Av - b||$$

Theorem 4.6.2 By the Best Approximation Theorem, the vector $b' = Au$ in $Col(A)$ closest to b is thus the Projection of b onto $Col(A)$. Therefore,

- u is a Least Square Solution to $Ax = b \iff Au$ is the projection of b onto $Col(A)$
- u is a Least Square Solution to $Ax = b \iff u$ is a solution to $A^T Ax = A^T b$

- The Least Square Solution may not be unique, however, for any choice of Least Square Solution u, the **Projection** Au is unique.
- The set of Least Square Solutions is an Affine Space

Theorem 4.6.3 – Number of Least Squares Solution Whether a Linear System has a Unique Least Squares Solution or has Infinitely Many Least Squares Solutions can be given by the following:

- **Unique**: $M^T M$ is Invertible
- **Infinitely Many**: $M^T M$ is NOT Invertible

Theorem 4.6.4 Suppose $Ax = b$ is consistent, then u is a Least Square Solution to $Ax = b \iff u$ is a solution to $Ax = b$.

Algorithm to find a Least Squares Solution to $\mathbf{Ax = b}$

The Steps involved are as follows:

1. Find an Orthogonal / Orthonormal Basis for V, the column space of A
2. Find the Projection p of b onto V (See the Algorithm below)
3. Solve the Linear System $Ax = p$

Algorithm to find Orthogonal Projection using Least Square Approximation

Another method to find the Projection of a vector b onto a subspace is given as follows:

1. Write a matrix $A = (a_1, a_2, \ldots, a_n)$
2. Find a Least Squares Solution, u, to $Ax = b$, that is, a solution to $A^T Ax = A^T b$
3. Let the Projection of the vector b onto V be $p = Au$

Matlab

To find the Least Square Solution u, and further, the Projection p of b onto the Column Space of A, for a given system $Ax = b$,

```
rref_system = lss(A, b)
p = A * u
```

where u is the solution for `rref_system`.

4.6.1 Examples

Least Square Approximation

Let $z = ax^2 + by + c$ be a function in variables x, y for some fixed real numbers a, b, c. Suppose that we have measured 4 data points for the values (x, y, z) as follows:

$$(1,2,0) \quad (0,1,1) \quad (-1,0,1) \quad (1,1,1)$$

We can then find the Best Approximation of this function (such that the sum of the squares of the errors of the function is minimal) by using the Least Squares Method as follows:

Subsituting the 4 Data Points into the function, we can contruct the matrix as follows:

Let $z = ax^2 + by + c$ be a function in variables x, y for some fixed real numbers a, b, c. Suppose that we have measured 4 data points for the values (x, y, z) as follows:

$$(1,2,0), \ (0,1,1), \ (-1,0,1), \ (1,1,1).$$

We can find the best approximation of this function (such that the sum of the squares of the errors of the function is minimal) by using the Least Squares Method.

First, we subsitute the data points into the function $z = ax^2 + by + c$ and obtain the following equations:

For $(1,2,0)$:

$$z = 0 \Rightarrow a(1^2) + b(2) + c = 0 \quad \Rightarrow \quad a + 2b + c = 0$$

For $(0,1,1)$:

$$z = 1 \Rightarrow a(0^2) + b(1) + c = 1 \quad \Rightarrow \quad b + c = 1$$

For $(-1,0,1)$:

$$z = 1 \Rightarrow a((-1)^2) + b(0) + c = 1 \quad \Rightarrow \quad a + c = 1$$

For $(1,1,1)$:

$$z = 1 \Rightarrow a(1^2) + b(1) + c = 1 \quad \Rightarrow \quad a + b + c = 1$$

We then construct the Augmented Matrix for $Ax = b$ (not the same b as the variable) as follows:

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Then, we can find the Least Square Approximation as per normal by solving $A^T Ax = A^T b$:

$$\left(\begin{array}{ccc|c} 3 & 3 & 3 & 2 \\ 3 & 6 & 4 & 2 \\ 3 & 4 & 4 & 3 \end{array}\right) \xrightarrow{RREF} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{1}{3} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{2} \end{array}\right)$$

Therefore, the Least Square Approximation of the function is $u = \begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{2} \\ \frac{3}{2} \end{pmatrix} = -\frac{1}{3}x^2 - \frac{1}{2}y + \frac{3}{2}$.

Orthogonal Projection and Least Square Solution

Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix}$ and $b = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Then, the Least Squares Solution of $Ax = b$ is given by solving $A^T Ax = A^T b$:

$$\left(\begin{array}{ccc|c} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \end{array}\right)$$

Therefore, the Least Squares Solution $= u = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

To find the Projection p of b onto the Column Space of A, we solve $p = Au$:

$$p = Au = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \\ 0 \end{pmatrix}$$

The smallest possible value of $||Av - b||$ among all vectors $v \in \mathbb{R}^3$ is given by $||Au - b|| = ||p - b||$:

$$\left\|\begin{pmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}\right\| = \left\|\begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix}\right\| = \sqrt{\left(-\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}}$$

Finally, given that the 3 columns of A form an Orthogonal Set, extend this set to an Orthogonal Basis for $\mathbb{R}^4$.

$$\begin{aligned} w &= w_p + w_n, \text{ where } w_n \text{ is orthogonal to } A \\ \therefore w_n &= w - w_p \\ &= b - p \\ &= \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} \\ 1 \\ \frac{1}{2} \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \\ 0 \end{pmatrix} \end{aligned}$$

$\therefore$ The extended Orthogonal Basis for $\mathbb{R}^4$ is: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \\ 0 \end{pmatrix} \right\}$

Chapter 5

Eigenanalysis

Definition 5.0.1 – Eigenvalues and Eigenvectors Let A be a square matrix of order n. Then,

- $\lambda \in \mathbb{R}$ is an **Eigenvalue** of A if there is a vector $v \in \mathbb{R}^n$ s.t. $Av = \lambda v$ where $v \neq 0$.
- v is called the **Eigenvector** associated with λ.

Definition 5.0.2 – Characteristic Polynomial Let A be a square matrix of order n. Then,

the **Characteristic Polynomial** of A, denoted as $char(A)$, is the degree n polynomial:

$$det(xI - A)$$

Theorem 5.0.1 – Finding Eigenvalues Let A be a square matrix of order n. Then,

the **Eigenvalues** of A are precisely all the roots to the Characteristic Equation:

$$det(xI - A) = 0$$

Definition 5.0.3 – Algebraic Multiplicity Let λ be an Eigenvalue of A. Then,

the **Algebraic Multiplicity** of λ, r_λ, is the largest integer power of the solution to the Characteristic Equation, that is:

$$det(xI - A) = (x - \lambda)^{r_\lambda} p(x)$$

Theorem 5.0.2 – Eigenvalues of Triangular Matrices Let A be a Triangular Matrix, then

- The Eigenvalues of A are its diagonal entries
- The Algebraic Multiplicity of each Eigenvalue is the number of times it appears in the diagonal entries of A

Definition 5.0.4 – Eigenspace Let A be a square matrix of order n. Then,

the **Eigenspace** associated to an Eigenvalue λ of A is given by:

$$E_\lambda = \{v \in \mathbb{R}^n | Av = \lambda v\} = Null(\lambda I - A)$$

Definition 5.0.5 – Geometric Multiplicity Further, the **Geometric Multiplicity** of an Eigenvalue λ is the Dimension of its Eigenspace:

$$dim(E_\lambda) = nullity(\lambda I - A)$$

Note that since $\lambda I - A$ is Singular by definition, $dim(E_\lambda) \geq 1$.

General Properties of Eigenvalues and Eigenvectors:

- Vectors from different Eigenspaces are Linearly Independent
- The Characteristic Polynomial of A is equal to the Characteristic Polynomial of A^T, that is, A and A^T have the same Eigenvalues
- The Geometric Multiplicity of A is equal to the Geometric Multiplicity of A^T
- The Geometric Multiplicity of an Eigenvalue λ of a square matrix A is no greater than the Algebraic Multiplicity, that is:
$$1 \leq dim(E_\lambda) \leq r_\lambda$$
- Suppose B is the resulting matrix of the x-th power of a matrix A, then the Eigenvalues of B are the x-th power of the Eigenvalues of A with the same corresponding Eigenvectors
- If $AB = BA$ and v is an Eigenvector of B, then Av is also an Eigenvector of B
- If λ is an Eigenvalue of A, then $c\lambda$ is an Eigenvalue for cA assuming $c \neq 0$.

Theorem 5.0.3 v is an Eigenvector associated with the Eigenvalue 0 $\iff$ v is in the Null Space of A, and thus, the Null Space of A^TA

Therefore, given two vectors v_1 and v_2 which are Least Square Solutions to $Ax = b$, where $b \neq 0$, their difference is in the Null Space of A^TA, and is thus, $v_1 - v_2$ is the Eigenvector associated with 0

5.1 Diagonalization

Definition 5.1.1 – Diagonalizable Let A be a square matrix of order n, with n Linearly Independent Eigenvectors. Then,

A is **Diagonalizable** into 3 Matrices, with an Invertible matrix P, s.t.

$$P^{-1}AP = D$$
$$A = PDP^{-1}$$

where D is a Diagonal Matrix.

Theorem 5.1.1 – Powers of Diagonalizable Matrices Suppose A is Diagonalizable, and can thus be decomposed as such: $A = PDP^{-1}$. Then,

$$A^k = PD^kP^{-1} = P \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix} P^{-1}$$

Equivalent Statements for Diagonalizability, given a Square Matrix A of order n:

1. A is Diagonalizable
2. There exists a Basis $\{u_1, u_2, \dots, u_n\}$ of $\mathbb{R}^n$ of the Eigenvectors of A
3. The Characteristic Polynomial of A splits into Linear Factors s.t. the Geometric Multiplicity of each Eigenvalue is equal to the Algebraic Multiplicity:
$$dim(E_{\lambda_i}) = r_{\lambda_i}$$

Therefore, to show that a Square Matrix A of order n is not diagonalizable, show either of the following:

1. $det(xI - A)$ does not split into Linear Factors
2. There exists an Eigenvalue λ s.t. $dim(E_\lambda) < r_\lambda$

Notice here that we do not require n Eigenvalues, but rather n Linearly Independent Eigenvectors, and each Eigenvalue can have more than 1 Eigenvector

Algorithm to Diagonalize a Matrix

Let A be a Diagonalizable Square Matrix of order n, then, to diagonalize A:

1. Compute the Eigenvalues of A
2. For each Eigenvalue λ_i of A, find the Basis S_{λ_i} for the Eigenspace. That is, find a Basis for the Solution Space of the Following Linear System:
$$(\lambda_i I - A)x = 0$$
3. Let $S = \{u_1, u_2, \ldots, u_n\}$ be a basis for $\mathbb{R}^n$ consisting of the Eigenvectors of A
4. Let $P = (u_1 \;\; u_2 \;\; \ldots \;\; u_n)$ and $D = diag(\lambda_1 \;\; \lambda_2 \;\; \lambda_n)$
5. Then, $A = PDP^{-1}$

Matlab

To perform Diagonalization with Matlab, use the `diagg` function as follows:

```
[P, D] = diagg_sym(matrix, size)
```

This will provide:

- `poly`: The characteristic polynomial
- `sad`: The RREF of $\lambda I_{size} - matrix$

5.2 Orthogonally Diagonalizable

Definition 5.2.1 – Orthogonally Diagonalizable A **Symmetric** Matrix A of order n is Orthogonally Diagonalizable if
$$A = PDP^T$$
for some Orthogonal Matrix P and a Diagonal Matrix D.

Theorem 5.2.1 – Spectral Theorem Let A be a n x n Square Matrix, then

A is Orthogonally Diagonalizable $\iff$ A is Symmetric

Equivalent Statements for Orthogonally Diagonalizable, given a Square Matrix A of order n:

1. A is Orthogonally Diagonalizable
2. There exists an Orthonormal Basis $\{u_1, u_2, \ldots, u_n\}$ of $\mathbb{R}^n$ of Eigenvectors of A
3. A is a Symmetric Matrix
4. A is a Square Matrix

Theorem 5.2.2 – Eigenspaces of a Symmetric Matrix is Orthogonal If A is a Symmetric Matrix, then the Eigenspaces of A are orthogonal to each other.

$\therefore$ The Eigenvectors associated with distinct Eigenvalues are orthogonal to each other.

The above thus allows us to forego applying Gram-Schmidt Process to Orthogonally Diagonalise a Matrix IF said Matrix is of Order n and has n distinct Eigenvalues.

Algorithm to Orthogonally Diagonalize a Matrix

Let A be an Orthogonally Diagonalizable Symmetric Square Matrix of order n, then, to diagonalize A:

1. Compute the Eigenvalues of A
2. For each Eigenvalue λ_i of A, find the Basis S_{λ_i} for the Eigenspace. That is, find a Basis for the Solution Space of the Following Linear System:
$$(\lambda_i I - A)x = 0$$
3. Let $S = \{u_1, u_2, \ldots, u_n\}$ be a basis for $\mathbb{R}^n$ consisting of the Eigenvectors of A
4. Then, apply the Gram-Schmidt Process on S to obtain an Orthonormal Basis for $\mathbb{R}^n$, $T = \{v_1, v_2, \ldots, v_n\}$
5. Let $P = (v_1 \;\; v_2 \;\; \ldots \;\; v_n)$ be an Orthogonal Matrix, and $D = diag(\lambda_1 \;\; \lambda_2 \;\; \ldots \;\; \lambda_n)$
6. Then, $A = PDP^T$

Note that the only difference between this, and the algorithm is simply Diagonalise a matrix, is that we apply the Gram-Schmidt Process on S **and** normalize the resulting Orthogonal Matrix.

Matlab

To perform Orthogonal Diagonalization with Matlab, we use yet another custom function that uses `diagg`, then normalizes each vector in the resulting `S` matrix to get `P`

```
[S, D] = orthoDiagg(matrix, size)
```

5.3 Markov Chain

A Markov Chain of vectors in $\mathbb{R}^n$ describe a system which has n fixed outcomes, and the probability of each of these outcomes in a particular state x, is dependent solely on the previous state $x-1$.

Definition 5.3.1 The following definitions are all relevant for Markov Chains:

1. **Probability Vector**: A vector $v = (v_i)_n$ with **Non-negative** coordinates that add up to 1
2. **Stochastic Matrix**: A Square Matrix whose columns are all Probability Vectors
3. **Markov Chain**: A sequence of Probability Vectors: $x_0, x_1, \ldots, x_k$ together with a Stochastic Matrix P s.t.
$$x_1 = Px_0, \quad x_2 = Px_1, \quad \ldots, \quad x_k = Px_{k-1}$$
4. **Steady-State Vector / Equilibrium Vector**: The probability vector in a Stochastic Matrix associated with the Eigenvalue 1

Algorithm to Compute Equilibrium Vector

Let P be a m x n Stochastic Matrix, then:

1. Find an Eigenvector u associated with Eigenvalue $= 1$; That is, find a non-trivial solution to $(I-P)x = 0$
2. Then, the Equilibrium Vector v is given by the following:
$$v = \frac{1}{\sum_{k=1}^{n} u_k} u$$

Matlab

The relevant Matlab functions for Markov Chain are as follows:

- To get the state vector after n days: `markchain(matrix, n, start)`, given the Stochastic Matrix *matrix*, and the initial state vector *start*
- To compute the Equilibrium vector given a Stochastic Matrix *matrix*:

```
chareqn(matrix, 1, size)
ssv(e1)
```

  where $e1$ is given by the Eigenvector associated with Eigenvalue 1.

5.3.1 Full Examples

Example to compute Equilibrium Vector given a Stochastic Matrix P

Let $P = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 1 & \frac{1}{2} & 0 \end{pmatrix}$,

$$I - P = \begin{pmatrix} 1 & 0 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{3} & 1 & 0 & 0 \\ -\frac{1}{3} & 0 & 1 & -\frac{1}{2} \\ -\frac{1}{3} & -1 & -\frac{1}{2} & 1 \end{pmatrix} \xrightarrow{RREF} \begin{pmatrix} 1 & 0 & 0 & -\frac{9}{10} \\ 0 & 1 & 0 & -\frac{3}{10} \\ 0 & 0 & 1 & -\frac{4}{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore, $u = \begin{pmatrix} \frac{9}{10} \\ \frac{3}{10} \\ \frac{4}{5} \\ 1 \end{pmatrix}$ is the Eigenvector associated with the Eigenvalue $= 1$. Hence, the Equilibrium Vector v of

P is given by:

$$v = \left(\frac{9}{10} + \frac{3}{10} + \frac{4}{5} + 1\right)^{-1} \begin{pmatrix} \frac{9}{10} \\ \frac{3}{10} \\ \frac{4}{5} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{10} \\ \frac{1}{10} \\ \frac{4}{15} \\ \frac{1}{3} \end{pmatrix}$$

Example for Full Markov Chain

Sheldon only patronises 3 stalls in the School Canteen, the Mixed Rice, Noodle and Mala Hotpot stall for lunch everyday. He never buys from the same stall 2 days in a row. If he buys from the mixed rice stall on a certain day, there is a 40% chance he will patronise the Noodle Stall the next day. If he buys from the noodle stall on a certain day, there is a 50% chance he will eat Mala Hotpot the next day. If he eats Mala Hotpot on a certain day, there is a 60% chance he will patronise the Mixed Rice stall the next day.

Now, let a_n, b_n, c_n be the probability that Sheldon patronises the Mixed Rice, Noodles and Mala Hotpot stall for lunch after n days. Let $x_n = \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix}$ be the State Vector, and let P be the Stochastic Matrix.

Then, suppose Sheldon patronises the Mixed Rice, Noodles and Mala Hotpot stalls today, his State Vector is e_1, e_2, e_3 respectively, where $e_i \in \mathbb{R}^3$ is the i-th vector in the Standard Basis. Thus,

$$\begin{pmatrix} 0 \\ 0.4 \\ 0.6 \end{pmatrix} = P \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = Pe_1, \quad \begin{pmatrix} 0.5 \\ 0 \\ 0.5 \end{pmatrix} = P \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = Pe_2, \quad \begin{pmatrix} 0.6 \\ 0.4 \\ 0 \end{pmatrix} = P \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = Pe_3, \;\Rightarrow\; P = \begin{pmatrix} 0 & 0.5 & 0.6 \\ 0.4 & 0 & 0.4 \\ 0.6 & 0.5 & 0 \end{pmatrix}$$

By the definition of Markov Chain, the state vector after n days will be $x_n = P^n x_0$.

Therefore, to compute the state vectors after n days more efficiently, we can diagonalise P as such:

Using 'eig' on Matlab, $\lambda = 1, -0.6 - 0.4$,

$$\lambda = 1, \quad \begin{bmatrix} 1 & -0.5 & -0.6 \\ -0.4 & 1 & -0.4 \\ -0.6 & -0.5 & 1 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.8 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \text{Basis: } \left\{ \begin{pmatrix} 1 \\ 0.8 \\ 1 \end{pmatrix} \right\}$$

$$\lambda = -0.6, \quad \begin{bmatrix} -0.6 & -0.5 & -0.6 \\ -0.4 & -0.6 & -0.4 \\ -0.6 & -0.5 & -0.6 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \text{Basis: } \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$$\lambda = -0.4, \quad \begin{bmatrix} -0.4 & -0.5 & -0.6 \\ -0.4 & -0.4 & -0.4 \\ -0.6 & -0.5 & -0.4 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \text{Basis: } \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$$

Therefore,

$$P = \begin{pmatrix} 1 & 0.5 & 0.6 \\ 0.4 & 0 & 0.4 \\ 0.6 & 0.5 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 0.8 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.6 & 0 \\ 0 & 0 & -0.4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0.8 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}^{-1}$$

This allows us to compute the state vector after n days:

i.e., Suppose Sheldon had Noodles today, then, after 3 days:

$$x_3 = P^3 = \begin{pmatrix} 1 & -1 & 1 \\ 0.8 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1^3 & 0 & 0 \\ 0 & -0.6^3 & 0 \\ 0 & 0 & -0.4^3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0.8 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.38 \\ 0.24 \\ 0.38 \end{pmatrix}$$

We can thus also compute the Equilibrium Vector as such:

$$P^k x_0 = P^k \begin{pmatrix} a \\ b \\ c \end{pmatrix} \to k \to \infty = \begin{pmatrix} 1 & -1 & 1 \\ 0.8 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0.8 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$= \frac{1}{14} \begin{pmatrix} 5 & 5 & 5 \\ 4 & 4 & 4 \\ 5 & 5 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 5(a+b+c) \\ 4(a+b+c) \\ 5(a+b+c) \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 5 \\ 4 \\ 5 \end{pmatrix}$$

5.4 Singular Value Decomposition

Singular Value Decomposition allows us to decompose Non-Square matrices into 3 Matrices:

1. U: Order m Orthogonal Matrix
2. Σ: m x n Matrix
3. V: Order n Orthogonal Matrix

such that, $\forall m$ x n Matrix A,

$$A = U\Sigma V^T$$

Definition 5.4.1 – Singular Values Given the Eigenvalues of a matrix A arranged in decreasing order, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$, the Singular values of A is given by:

$$\sigma_1 \geq \sigma_2 \geq \dots \sigma_n \geq 0$$

where $\sigma_i = \sqrt{\lambda_i}$.

Algorithm to Singular Value Decomposition

Let A be a m x n matrix with $rank(A) = r$,

1. Find the Eigenvalues of $A^T A$, and arrange the Non-zero Eigenvalues in descending order (counting multiplicity)
2. Construct $\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sigma_n \end{bmatrix}$
   - If Σ is not a m x n matrix, fill in the missing rows and columns with 0s
3. Find an Orthogonal Basis for each Eigenspace and let v_i be the **unit vector** associated with each Eigenvalue.
4. Let $V = (v_1 \ \ v_2 \ \ \dots \ \ v_n)$
5. Let $u_i = \frac{1}{\sigma_i} A v_i$, letting $U = (u_1 \ \ u_2 \ \ \dots \ \ u_n)$
   - If U is not a m x m matrix, extend $\{u_1, \dots, u_r\}$ by solving $(u_1 \dots u_r)^T x = 0$ and finding the Orthonormal Basis for the Solution Space
6. Construct $A = U\Sigma V^T$.

Matlab

To perform Singular Value Decomposition with Matlab, we do the following:

```
[U, S, V] = sadvd(matrix)
```

Remember to tranpose V, to construct $A = U\Sigma V^T$.

Example (Extending U)

Let $A = \begin{bmatrix} 4 & 2 & 6 \\ 6 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and thus, $A^T A = \begin{bmatrix} 52 & 8 & 24 \\ 8 & 4 & 12 \\ 24 & 12 & 36 \end{bmatrix}$.

The Eigenvalues of $A^T A$ is: $72, 20, 0$.

First, note that 1 of the Eigenvalues is 0, and thus, we cannot use it as a Eigenvalue. Thus, we will have to extend the Orthonormal Basis U. However, this has no impact on finding V, which we will do first:

$$\lambda = 72, \quad \begin{bmatrix} 20 & -8 & -24 \\ -8 & 68 & -12 \\ -24 & -12 & 36 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & -\frac{4}{3} \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \text{Basis: } \left\{ \begin{pmatrix} \frac{4}{3} \\ \frac{1}{3} \\ 1 \end{pmatrix} \right\}$$

$$\lambda = 20, \quad \begin{bmatrix} -32 & -8 & -24 \\ -8 & 16 & -12 \\ -24 & -12 & -16 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & \frac{5}{6} \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \text{Basis: } \left\{ \begin{pmatrix} -\frac{5}{6} \\ \frac{1}{3} \\ 1 \end{pmatrix} \right\}$$

$$\lambda = 0, \quad \begin{bmatrix} -52 & -8 & -24 \\ -8 & -4 & -12 \\ -24 & -12 & -36 \end{bmatrix} \xrightarrow{rref} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \text{Basis: } \left\{ \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix} \right\}$$

Then, normalising each vector, we get:

$$v_1 = \frac{1}{\sqrt{26}} \begin{pmatrix} \frac{4}{3} \\ \frac{1}{3} \\ 1 \end{pmatrix} \qquad v_2 = \frac{1}{\sqrt{65}} \begin{pmatrix} -\frac{5}{6} \\ \frac{1}{3} \\ 1 \end{pmatrix} \qquad v_3 = \frac{1}{\sqrt{10}} \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix}$$

$$V = (v_1 \; v_2 \; v_3) = \begin{bmatrix} \frac{4}{\sqrt{26}} & -\frac{5}{\sqrt{65}} & 0 \\ \frac{1}{\sqrt{26}} & \frac{2}{\sqrt{65}} & -\frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{26}} & \frac{6}{\sqrt{65}} & \frac{1}{\sqrt{10}} \end{bmatrix}.$$

To find U, we can quickly find the vectors in the Orthonormal Basis in $\mathbb{R}^2$ first.

$$u_1 = \frac{1}{6\sqrt{2}} A v_1 = \frac{1}{\sqrt{13}} \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

$$u_2 = \frac{1}{2\sqrt{5}} A v_2 = \frac{1}{\sqrt{13}} \begin{pmatrix} 2 \\ -3 \\ 0 \end{pmatrix}$$

.

To find the last vector to extend this Orthonormal Basis to $\mathbb{R}^3$ (since A is in $\mathbb{R}^3$),

We solve for $(u_1 \; u_2)^T x = 0$,

$$\left\{ \begin{array}{ccc|c} \frac{3}{\sqrt{13}} & \frac{2}{\sqrt{13}} & 0 & 0 \\ \frac{2}{\sqrt{13}} & -\frac{3}{\sqrt{13}} & 0 & 0 \end{array} \right\} \xrightarrow{rref} \left\{ \begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right\}.$$

Therefore, our last vector is $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Notice that this is easy to eyeball from u_1 and u_2, as both have 0s in Row 3.

Then, our Orthonormal Basis for $\mathbb{R}^3$ is:

$$C = \left\{ \begin{pmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{2}{\sqrt{13}} \\ -\frac{3}{\sqrt{13}} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Performing Gram-Schmidt Process on C,

$$U = \begin{bmatrix} \frac{3}{\sqrt{13}} & \frac{2}{\sqrt{13}} & 0 \\ \frac{2}{\sqrt{13}} & -\frac{3}{\sqrt{13}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore, $A = \begin{bmatrix} \frac{3}{\sqrt{13}} & \frac{2}{\sqrt{13}} & 0 \\ \frac{2}{\sqrt{13}} & -\frac{3}{\sqrt{13}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{6}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{2}{\sqrt{5}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{4}{\sqrt{26}} & -\frac{5}{\sqrt{65}} & 0 \\ \frac{1}{\sqrt{26}} & \frac{2}{\sqrt{65}} & -\frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{26}} & \frac{6}{\sqrt{65}} & \frac{1}{\sqrt{10}} \end{bmatrix}^T$.

Chapter 6

Linear Transformation

Definition 6.0.1 – Linear Transformation A mapping (function) $T : \mathbb{R}^n \to \mathbb{R}^m$ is a Linear Transformation if $\forall u, v \in \mathbb{R}^n$ and Scalars α, β,

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$$

For the above Linear Transformation,

- Domain: $\mathbb{R}^n$
- Codomain: $\mathbb{R}^m$

Specifically,

1. T maps the Origin to the Origin, that is, let $\alpha = \beta = 0$, then
$$T(0) = 0$$
2. If we let $\beta = 0$, then,
$$T(\alpha u) = \alpha T(u)$$
3. If we let $\alpha = \beta = 1$, then,
$$T(u + v) = T(u) + T(v)$$

Therefore, to show that a specific mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is **not** a Linear Transformation, show that any of the above 3 properties are false for said mapping T.

General Properties of Linear Transformation, $T : \mathbb{R}^n \to \mathbb{R}^m$:

- $\{T(v) \mid v \in V\}$ is a vector space for every subspace V of $\mathbb{R}^n$
- $\{v \in \mathbb{R}^n \mid T(v) \in V\}$ is a vector space for every subspace V of $\mathbb{R}^m$

Definition 6.0.2 – Standard Matrix Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a Linear Transformation. Then, the **unique** m x n matrix A s.t.

$$T(u) = Au \quad \forall u \in \mathbb{R}^n$$

is called the Standard Matrix of T.

Specifically,

$$A = (T(e_1) \;\; T(e)2 \;\; \ldots \;\; T(e_n))$$

where $E = \{e_1, e_2, \ldots, e_n\}$ is the Standard Basis for $\mathbb{R}^n$.

We can also say that the Standard Matrix of a Linear Transformation T is the representation of T with respect to the Standard Matrix $A = [T]_E$.

Definition 6.0.3 – Representation of T with respect to a Basis S Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a Linear Transformation and $S = \{u_1, u_2, \ldots, u_n\}$ be a Basis for $\mathbb{R}^n$. Then, the Representation of T with respect to a Basis S, denoted as $[T]_S$ is defined to be the m x n matrix:

$$[T]_S = (T(u_1)\ \ T(u_2)\ \ \dots T(u_n))$$

Definition 6.0.4 – Range Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a Linear Transformation. Then, the **range** of T is the set of all images of T:

$$\begin{aligned} R(T) &= \{T(v) \mid v \in \mathbb{R}^n\} \\ &= span\{T(v_1), \dots, T(v_n)\} \\ &= Col(A) \end{aligned}$$

where A is the Standard Matrix of T.

Definition 6.0.5 – Rank of Linear Transformation Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a Linear Transformation. Then, the **rank** of T is the Dimension of the range of T:

$$rank(T) = dim(R(T))$$

Definition 6.0.6 – Kernel Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a Linear Transformation. Then, the **kernel** of T is the set of all vectors in $\mathbb{R}^n$ that maps to the Zero Vector by T:

$$\begin{aligned} ker(T) &= \{u \in \mathbb{R}^n \mid T(u) = 0\} \\ &= Null(A) \end{aligned}$$

where A is the Standard Matrix of T.

Definition 6.0.7 – Nullity of Linear Transformation Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a Linear Transformation. Then, the **nullity** of T is the Dimension of the Kernel of T:

$$nullity(T) = dim(ker(T))$$

Example: Finding Standard Matrix (if it Exists)

1. $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix} = x\begin{pmatrix} 0 \\ 1 \end{pmatrix} = y\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

   $\therefore$ This is a Linear Transformation with Standard Matrix $= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
2. $T\left(\begin{pmatrix} y \\ x \end{pmatrix}\right) = \begin{pmatrix} x^2 \\ xy \end{pmatrix} = x\begin{pmatrix} x \\ y \end{pmatrix}$

   Let $\alpha = 2$,

   $$2T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \neq \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 2^2 \\ 2*2 \end{pmatrix} = T\left(2\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$$

   $\therefore$ This is **not** a Linear Transformation.
3. Let $R = \{u_1, u_2, u_3\}$ and $T = \{v_1, v_2, v_3\}$ where:

   $$u_1 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \quad u_2 = \begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix} \quad u_3 = \begin{pmatrix} 1 \\ -3 \\ 5 \end{pmatrix}$$

   $$v_1 = \begin{pmatrix} 7 \\ 0 \\ -3 \end{pmatrix} \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \quad v_3 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

   Then, let A be a matrix transformation on $\mathbb{R}^3$ such that:

   $$A(u_1) = v_1 + 3v_2 + 4v_3, \quad A(u_2) = 2v_2 + v_3, \quad A(u_3) = -4v_2 + 3v_3$$

$\therefore$ To find A,

$$A\begin{pmatrix}2\\3\\-1\end{pmatrix}=\begin{pmatrix}7\\0\\-3\end{pmatrix}+3\begin{pmatrix}2\\3\\4\end{pmatrix}+4\begin{pmatrix}-1\\1\\2\end{pmatrix}=\begin{pmatrix}9\\13\\17\end{pmatrix}$$
$$A\begin{pmatrix}0\\-2\\3\end{pmatrix}=2\begin{pmatrix}2\\3\\4\end{pmatrix}+\begin{pmatrix}-1\\1\\2\end{pmatrix}=\begin{pmatrix}3\\7\\10\end{pmatrix}$$
$$A\begin{pmatrix}1\\-3\\5\end{pmatrix}=-4\begin{pmatrix}2\\3\\4\end{pmatrix}+3\begin{pmatrix}-1\\1\\2\end{pmatrix}=\begin{pmatrix}-11\\-8\\-10\end{pmatrix}$$
$$\therefore A\begin{pmatrix}2&0&1\\3&-2&-3\\-1&3&5\end{pmatrix}=\begin{pmatrix}9&3&-11\\13&7&-8\\17&10&-10\end{pmatrix}$$

$$\therefore A=\begin{pmatrix}9&3&-11\\13&7&-8\\17&10&-10\end{pmatrix}\begin{pmatrix}2&0&1\\3&-2&-3\\-1&3&5\end{pmatrix}^{-1}$$
$$=\frac{1}{5}\begin{pmatrix}-122&126&89\\-153&164&121\\-207&221&164\end{pmatrix}$$

4. Let $T:\mathbb{R}^3\to\mathbb{R}^2$ be a Linear Transformation and V be a Basis for $\mathbb{R}^3$ be:

$$u_1=\begin{pmatrix}1\\0\\1\end{pmatrix}\quad u_2=\begin{pmatrix}1\\1\\0\end{pmatrix}\quad u_3=\begin{pmatrix}0\\1\\1\end{pmatrix}$$

Suppose $T(u_1)=\begin{pmatrix}2\\0\end{pmatrix}$, $T(u_2)=\begin{pmatrix}0\\6\end{pmatrix}$, $T(u_3)=\begin{pmatrix}2\\2\end{pmatrix}$.
Then, the Standard Matrix of T is given by:

$$e_1=\begin{pmatrix}1\\0\\0\end{pmatrix}=\frac{1}{2}(u_1+u_2-u_3)\ \rightarrow\ T(e_1)=\frac{1}{2}(T(u_1)+T(u_2)-T(u_3))=\begin{pmatrix}0\\2\end{pmatrix}$$
$$e_2=\begin{pmatrix}0\\1\\0\end{pmatrix}=\frac{1}{2}(u_2+u_3-u_1)\ \rightarrow\ T(e_2)=\frac{1}{2}(T(u_2)+T(u_3)-T(u_1))=\begin{pmatrix}0\\4\end{pmatrix}$$
$$e_3=\begin{pmatrix}0\\0\\1\end{pmatrix}=\frac{1}{2}(u_1+u_3-u_2)\ \rightarrow\ T(e_3)=\frac{1}{2}(T(u_1)+T(u_3)-T(u_2))=\begin{pmatrix}2\\-2\end{pmatrix}$$
$$\therefore T=\begin{pmatrix}0&0&2\\2&4&-2\end{pmatrix}$$

Alternatively,

$$T\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix}=\begin{pmatrix}2&0&2\\0&6&2\end{pmatrix}$$
$$T=\begin{pmatrix}2&0&2\\0&6&2\end{pmatrix}\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix}^{-1}$$
$$=\begin{pmatrix}0&0&2\\2&4&-2\end{pmatrix}$$

Chapter 7

Appendix

7.1 Equivalent Statements

Equivalent Statements of Invertibility

Let A be a Square Matrix of order n, then the following statements are Equivalent:

- A is invertible
- $adj(A)$ is invertible
- A^T is invertible
- (Left Inverse) $\exists$ a matrix B s.t. $BA = I_n$
- (Right Inverse) $\exists$ a matrix B s.t. $AB = I_n$
- The RREF of A is I_n
- A can be expressed as a product of Elementary Matrices
- The Homogenous System $Ax = 0$ has only the trivial solution
- $\forall b$, the sytem $Ax = b$ has a unique solution
- $det(A) \neq 0$
- $A(adj(A)) \neq 0$
- The columns / rows of A are Basis for $\mathbb{R}^n$
- A is full rank, that is, $rank(A) = n$
- $rank(A) = rank(A^T A) = rank(AA^T)$
- $nullity(A) = 0$
- 0 is not **an** Eigenvalue of A
- The Linear Transformation $x \mapsto Ax$ is Bijective, and thus, Injective and Surjective

Equivalent Statements of Full Rank

A is Full Rank, where $rank(A) = n$, the number of columns

- The Rows of A spans $\mathbb{R}^n$, $Row(A) = \mathbb{R}^n$
- The Columns of A are Linearly Independent
- $Null(A) = 0$
- $A^T A$ is an Invertible Matrix of order n
- The Least Square Solution to $Ax = b$ is Unique
- A has a Left Inverse
- The Linear Transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by A is Injective

A is Full Rank, where $rank(A) = m$, the number of rows

- The Columns of A spans $\mathbb{R}^m$, $Col(A) = \mathbb{R}^m$
- The Rows of A are Linearly Independent
- The Linear System $Ax = b$ is consistent $\forall b \in \mathbb{R}^m$
- AA^T is an Invertible Matrix of order m
- A has a Right Inverse
- The Linear Transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by A is Surjective

7.2 Matlab

General Functions:

- For unknowns: `syms x y ...`
- Transpose: `transpose(matrix)`
- Get roots: `solve(eqn)` / `roots(eqn)`
- Power: `mpower(power, matrix)`
- Get Norm: `norm(matrix)`
- Null Space: `null(matrix)`
- Get Determinant: `det(matrix)`
- Get Adjoint: `adjoint(matrix)`
- Get Inverse: `inv(matrix)`
- To get exact values: `sym(matrix)`
- Constructing combined matrices: `[A B]` $\rightarrow$ `C` with `A` and `B` concatenated column-wise
- LU Factorization: `[L, U] = lu(matrix)`
- QR Factorization: `[Q, R] = qr(matrix)`

Shortcuts:

- Comment highlighted section: Ctrl-r
- Uncomment highlighted section: Ctrl-t
- Run file: Ctrl-Enter

Working with Matlab:

- If there is a `*` value, then use `format long`