MA283: Linear Algebra

70% Exam
30% Continuous Assessment (Homework)
10% Optional Project (Bonus)

Robert Davidson

Contents

1 Review of Matrix Algebra

Fields

- A field F is a set where addition, subtraction, multiplication and division (by nonzero elements) satisfy the usual algebraic properties. Common fields include $\mathbb{R}$ and $\mathbb{C}$
- We write $\mathbb{F}^p$ for the vector space of all p vectors with entries in $\mathbb{F}$.
- We'll cheat and treat any ordered list of p elements of $\mathbb{F}$ as an element of $\mathbb{F}^p$.
- For example, in $\mathbb{R}^3$, we might consider $(1, 2, 3)$ as coordinates, a row vector, or a column vector with 3 real entries.

Matrices Over a Field

- An $m \times n$ matrix over a field $\mathbb{F}$ is an array of m rows and n columns of elements from $\mathbb{F}$.
- When $m = n$, we write $M_n(\mathbb{F})$, otherwise we write $M_{m\times n}(\mathbb{F})$.

Addition and Scalar Multiplication

- Two matrices of the same size $m \times n$ can be added entrywise
- The $m \times n$ matrix has all entries equal to zero and acts as the additive identity (adding it to any matrix does not change the matrix)
- Multiplying a matrix by a scalar means multiplying each entry by that scalar
- The set of all $m \times n$ matrices over $\mathbb{F}$ is a vector space over $\mathbb{F}$

Linear Combinations

- A linear combination of vectors $v_1, v_2, \ldots, v_k$ in a vector space V with coefficients $a_1, a_2, \ldots, a_k \in \mathbb{F}$ is defined as:
$$a_1v_1 + a_2v_2 + \ldots + a_kv_k$$
- In particular, matrices themselves can be combined linearly, (e.g. $2A - 3B$)

Row and Column Vectors

- A column vector is a matrix with one column
- A row vector is a matrix with one row

Matrix-Vector Multiplication

- If A is $m \times n$ matrix and v is an n-entry column vector, the product Av is defined by taking a linear combination of the columns of A with the entries of v as coefficients.
- The result Av is an m-entry column vector.
- For a row vector u with m entries, and an $m \times n$ matrix Am the product uA a row vector in $\mathbb{F}^n$ formed by the linear combination of the rows of A with the entries of u as coefficients.

Matrix-Matrix Multiplication

- If A is a $m \times p$ and B is a $p \times n$ matrix, the product AB is defined only when the inner dimensions match (p)
- To find each column of AB, multiply A with the corresponding column vector of B.
- In entrywise form:
$$(AB)_{ij} = A_{i,1}B_{1,j} + A_{i,2}B_{2,j} + \ldots + A_{i,p}B_{p,j} = \sum_{k=1}^{p} A_{i,k}B_{k,j}$$

Dot Product and Orthogonality

- For two p-entry vectors, $u, v \in \mathbb{F}^p$, their dot product is:
$$u \cdot v = \sum_{k=1}^{p} u_kv_k$$
- Vectors are **orthogonal** if their dot product is zero.
- If $\mathbb{F} = \mathbb{R}$, this means the vector are perpendicular.
- In matrix multiplication, the entry $(AB)_{ij}$ can be viewed as the dot product of Row i with Column j of B.

Matrices and Tables

Lets consider the table that gives the numbers of Maths M, Physics P and Chemistry C students in each of the 3 years of a course:

Year	M	P	C
2015	50	100	70
2016	60	80	80
2017	70	90	90

$$\text{A} = \begin{bmatrix} 50 & 100 & 70 \\ 60 & 80 & 80 \\ 80 & 70 & 70 \end{bmatrix}$$

Each student of M, P, C must also take a course in X and Y. We can represent the credits they earn as a matrix:

Subject	X	Y
M	10	0
P	15	15
C	20	10

$$\text{B} = \begin{bmatrix} 10 & 0 \\ 15 & 15 \\ 20 & 10 \end{bmatrix}$$

The total number of credits earned each year can be found by the matrix product AB:

$$AB = \begin{bmatrix} 50 & 100 & 70 \\ 60 & 80 & 80 \\ 80 & 70 & 70 \end{bmatrix} \begin{bmatrix} 10 & 0 \\ 15 & 15 \\ 20 & 10 \end{bmatrix} = \begin{bmatrix} 50 \cdot 10 + 100 \cdot 15 + 70 \cdot 20 & 50 \cdot 0 + 100 \cdot 15 + 70 \cdot 10 \\ 60 \cdot 10 + 80 \cdot 15 + 80 \cdot 20 & 60 \cdot 0 + 80 \cdot 15 + 80 \cdot 10 \\ 80 \cdot 10 + 70 \cdot 15 + 70 \cdot 20 & 80 \cdot 0 + 70 \cdot 15 + 70 \cdot 10 \end{bmatrix}$$

We can represent the result as a table:

Year	X credits	Y credits
2015	3400	2200
2016	3400	2000
2017	3250	1750

$$\text{A} = \begin{bmatrix} 3400 & 2200 \\ 3400 & 2000 \\ 3250 & 1750 \end{bmatrix}$$

Linear Transformations

Let m and n be positive integers, A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$, denoted $T : \mathbb{R}^n \to \mathbb{R}^m$, is a function that satisfies the following properties:

- $T(u + v) = T(u) + T(v)$
- $T(\lambda u) = \lambda T(u)$

$\forall u, v \in \mathbb{R}^n$ and scalars $\lambda \in \mathbb{R}$

When $T : \mathbb{R}^3 \to \mathbb{R}^2$, if we know T applied to the three standard basis vectors of $\mathbb{R}^3$:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

we can form a 2×3 matrix A whose columns are exactly these image,then $T(v) = Av$ for any column vector $v \in \mathbb{R}^3$.

Composition of Linear Transformations

- If $T : \mathbb{R}^n \to \mathbb{R}^p$ and $S : \mathbb{R}^p \to \mathbb{R}^m$, then the composition $(S \circ T)(v) = S(T(v))$
- If T is represented by a $p \times n$ matrix A and S by an $m \times p$ matrix B then the composition $S \circ T$ is represented by the matrix product BA.
- Also $(AB)C = A(BC)$
- Composing transformations is only possible if the codomain of the first transformation matches the domain of the second transformation, that is:

$$A \in M_{m \times n} \quad B \in M_{p \times m} \quad \Rightarrow \quad AB \in M_{m \times n}$$

The $n \times n$ Identity Matrix

I_n has 1s on the main diagonal:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The identity matrix is the **neutral element** for multiplication:

$$A \cdot I_n = A \quad \text{and} \quad I_n \cdot B = B \quad (\text{where } A \text{ has } n \text{ columns and } B \text{ has } n \text{ rows})$$

The is interpreted as the **identity transformation** on $\mathbb{R}^n$, so composing with it has no effect on any linear map.

Invertible (Non-Singular) Matrices

A square $n \times n$ matrix A has an inverse A^{-1} if there exists another $n \times n$ matrix such that:

$$AB = I_n \quad \text{and} \quad BA = I_n$$

If A has an inverse, we say it is **unique**; there cannot be two different inverses for the same matrix.

Not all matrices are invertible. A key fact (explained later) is that:

$$A \text{ is invertible} \quad \Leftrightarrow \quad \text{the determinant} \neq 0$$

Transpose of a Matrix

FGor a $m \times n$ matrix A, the transpose A^T is the $n \times m$ matrix obtained by turning the rows of A into the columns of A^T:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

2 Systems of linear equations

2.1 Linear equations and Solution Sets

A linear equation in the variables x and y is an equation of the form

$$2x + y = 3$$

If we replace x and y with some numbers, the statement **becomes true or false**.

Definition 2.1: Solution to a linear equation

A pair, $(x_0, y_0) \in \mathbb{R}$, is a solution to an linear equation if setting $x = x_0$ and $y = y_0$ **makes the equation true.**

Definition 2.2: Solution set

The **solution set** is the set of all solutions to a linear equation.

$$a_1X_1 + a_2X_2 + \ldots + a_nX_n = b \quad \text{where } a_i, b \in \mathbb{R}$$

is an **affine hyperplane** in $\mathbb{R}^n$; geometrically resembles a copy of $\mathbb{R}^{n-1}$ inside $\mathbb{R}^n$.

2.2 Elementary Row Operations

To solve a system of linear equations we associate an **augmented matrix** to the system of equations. For example:

$$\begin{array}{rcrcrcr} x & + & 2y & - & z & = & 5 \\ 3x & + & y & - & 2z & = & 9 \\ -x & + & 4y & + & 2z & = & 0 \end{array} \quad \Rightarrow \quad \left[\begin{array}{rrr|r} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ -1 & 4 & 2 & 0 \end{array}\right]$$

To solve, we can perform the following **Elementary Row Operations (EROs)**:

1. Multiply a row by a non-zero constant.
2. Add a multiple of one row to another row.
3. Swap two rows.

The goal of these operations is to transform the augmented matrix into **row echelon form** (REF) or **reduced row echelon form** (RREF).

2.2.1 REF and Strategy

We say a matrix is in **row echelon form** (REF) if:

- The first non zero entry in each row is a 1 (called the **leading 1**).
- If a column has a leading 1, then all entries below it are 0.
- The leading 1 in each row is to the right of the leading 1 in the previous row.
- All rows of 0s are at the bottom of the matrix.

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

Example of REF

We have produced a new system of equations. This is easily solved by back substitution.

Concept 2.1: Stategy for Obtaining REF

- Get a 1 as the top left entry
- Use this 1 to clear the entries below it
- Move to the next column and repeat
- Continue until all leading 1s are in place
- Use back substitution to solve the system

2.2.2 Row Reduced Echelon Form

A matrix is in **reduced row echelon form** (RREF) if:

- It is in REF
- The leading 1 in each row is the only non-zero entry in its column.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

Example of RREF

2.3 Leading variables and free variables

We'll start by an example:

$$\begin{array}{rcrcrcrcr} x_1 & - & x_2 & - & x_3 & + & 2x_4 & = & 0 \\ 2x_1 & + & x_2 & - & x_3 & + & 2x_4 & = & 8 \\ x_1 & - & 3x_2 & + & 2x_3 & + & 7x_4 & = & 2 \end{array} \Rightarrow \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 0 \\ 2 & 1 & -1 & 2 & 8 \\ 1 & -3 & 2 & 7 & 2 \end{array}\right]$$

Solving this system of equations, we get:

$$\text{RREF:} \quad \left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & 4 \\ 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right] \Rightarrow \begin{array}{rcrcr} x_1 & + & 2x_4 & = & 4 \\ x_2 & - & x_4 & = & 2 \\ x_3 & + & x_4 & = & 2 \end{array} \Rightarrow \begin{array}{rcl} x_1 & = & 4 - 2x_4 \\ x_2 & = & 2 + x_4 \\ x_3 & = & 2 - x_4 \end{array}$$

This RREF tells us how the **leading variables** (x_1, x_2, x_3) depend on the **free variable** (x_4). The free variable can take any value in $\mathbb{R}$. We write the solution set as:

$$x_1 = 4 - 2t, \quad x_2 = 2 + t, \quad x_3 = 2 - t, \quad x_4 = t \quad \text{where } t \in \mathbb{R}$$

$$(x_1, x_2, x_3, x_4) = (4 - 2t, 2 + t, 2 - t, t); \quad t \in \mathbb{R}$$

Definition 2.3: Leading and Free Variables

- **Leading variable** : A variable whose columns in the RREF contain a leading 1
- **Free variable** : A variable whose columns in the RREF do not contain a leading 1

2.4 Consistent and Inconsistent Systems

Consider the following system of equations:

$$\begin{array}{rcrcrcr} 3x & + & 2y & - & 5z & = & 4 \\ x & + & y & - & 2z & = & 1 \\ 5x & + & 3y & - & 8z & = & 6 \end{array} \Rightarrow \left[\begin{array}{ccc|c} 3 & 2 & -5 & 4 \\ 1 & 1 & -2 & 1 \\ 5 & 3 & -8 & 6 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & -2 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right] \quad \text{(REF)}$$

We can see the last row of the REF is:

$$0x + 0y + 0z = 1$$

This equation clearly has no solution, and hence the system has no solutions. We say the system is **inconsistent**. Alternatively, we say the system is **consistent** if it has at least one solution.

2.5 Possible Outcomes when solving a system of equations

- The system may be **inconsistent** (no solutions) - i.e:

$$[0\ 0\ \dots\ 0 \mid a] \quad a \neq 0$$

- The system may be **consistent** which occurs if:
  - **Unique Solutions** each column (aside from the rightmost) contains a single leading 1. - i.e:

$$\begin{bmatrix} 1 & 0 & 0 & \mid & 4 \\ 0 & 1 & 0 & \mid & 3 \\ 0 & 0 & 1 & \mid & -2 \end{bmatrix}$$

  - **Infinitely many solutions** at least one variable does not appear as a leading 1 in any row, making it a free variable - i.e:

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

2.6 Elementary Row Operations as Matrix Transformations

Elementary row operations may be interpreted as **matrix multiplication**. To see this, first we introduce the **Identity matrix:**

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The I_m Identity matrix is an $m \times m$ matrix with 1s on the diagonal and 0s elsewhere. We also introduce the $E_{i,j}$ matrix which has 1 in the (i, j) position and 0s elsewhere. For example:

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

Then:

$$I_3 + 4E_{1,2} = \begin{bmatrix} 1 & 4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Performing a row operation on A is the same as multiplying A by an appropriate matrix E on the left. These matrices are called **elementary matrices**. They are **always invertible,** and **their inverses are also elementary matrices**. The statement:

"*every matrix can be reduced to RREF through EROs*"

is equivalent to saying that

"*for every matrix A with m rows, there exists a $m \times m$ matrix B which is a product of elementary matrices such that BA is in RREF.*"

2.6.1 Multiplying a Row by a Non-Zero Scalar

When multiplying row i of matrix A by a scalar $\alpha \neq 0$, we can use the matrix:

$$I_m + (\alpha - 1)E_{i,i}$$

This works because it modifies only the (i, i) entry of the identity matrix to be α while keeping all other entries unchanged. When multiplied with A, it scales row i by α and leaves all other rows intact.

Example: If $\alpha = 5$ and $i = 2$, then:

$$I_3 + 4E_{2,2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad (I_3 + 4E_{2,2})A = \begin{bmatrix} 1 & 2 & 3 \\ 20 & 25 & 30 \\ 7 & 8 & 9 \end{bmatrix}$$

2.6.2 Switching Two Rows

To swap rows i and k, we use:

$$S = I_m + E_{i,k} + E_{k,i} - E_{i,i} - E_{k,k}$$

This works by:

- Removing the 1's at positions (i, i) and (k, k) from the identity matrix
- Adding 1's at positions (i, k) and (k, i)

Example: Swapping rows 1 and 3:

$$S = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad SA = \begin{bmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}$$

2.6.3 Adding a Multiple of One Row to Another

To replace row k with row $k + \alpha\times$ row i, use:

$$I_m + \alpha E_{k,i}$$

This adds α times row i to row k while leaving all other rows unchanged because:

- For any row $j \neq k$, the corresponding row in this matrix is just the standard basis row
- Row k becomes the sum of the standard basis row k plus α times the standard basis row i

Example: Adding 3 times row 1 to row 2:

$$I_3 + 3E_{2,1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad (I_3 + 3E_{2,1})A = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 11 & 15 \\ 7 & 8 & 9 \end{bmatrix}$$

Example 2.1

Write the inverse of an elementary matrix and show it is an elementary matrix.

Multiplying a row by a nonzero scalar:

- **Operation:** Multiply row i by $\alpha \neq 0$.
- **Elementary Matrix:** $E = I_m + (\alpha - 1)E_{i,i}$
- **Inverse:** To reverse the operation, multiply row i by $1\backslash\alpha$. Hence, the inverse is

$$E^{-1} = I_m + \left(\frac{1}{\alpha} - 1\right) E_{i,i} = I_m + \frac{1-\alpha}{\alpha} E_{i,i}.$$

Swapping two rows:

- **Operation:** Swap rows i and k.
- **Elementary Matrix:** $S = I_m - E_{i,i} - E_{k,k} + E_{i,k} + E_{k,i}$
- **Inverse:** Since swapping the same two rows twice returns them to their original positions,

$$S^{-1} = S.$$

Adding a multiple of one row to another:

- **Operation:** Add α times row i to row k.
- **Elementary Matrix:** $E = I_m + \alpha\, E_{k,i}$
- **Inverse:** To undo the operation, subtract α times row i from row k. Therefore,

$$E^{-1} = I_m - \alpha\, E_{k,i}.$$

Example 2.2

Prove that every invertible matrix in $M_n(\mathbb{R})$ is a product of elementary matrices.

Let A be an invertible matrix in $M_n(\mathbb{R})$. Since A is invertible, we can use Gaussian elimination to transform A into the identity matrix I_n.

Let $E_1, E_2, \ldots, E_k$ be the elementary matrices corresponding to the row operations used in the elimination process. Then, we have:

$$\begin{aligned} \text{Multiplying a row by a scalar:} \quad & I_n + (\alpha - 1)E_{i,i} \\ \text{Swapping two rows:} \quad & I_n + E_{i,k} + E_{k,i} - E_{i,i} - E_{k,k} \\ \text{Adding a multiple of one row to another:} \quad & I_n + \alpha\, E_{k,i} \end{aligned}$$

Applying these in sequence to A gives:

$$E_k \cdots E_2 E_1 A = I_n$$

Since $E_k \cdots E_2 E_1 = I_n$, we can multiply both sides by $(E_k \cdots E_2 E_1)^{-1}$ on the left to obtain:

$$A = (E_k \cdots E_2 E_1)^{-1} I_n = (E_k \cdots E_2 E_1)^{-1}$$

Using the property:

$$(E_k \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1},$$

we can express A as a product of elementary matrices:

$$A = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$$

Since each E_i is an elementary matrix, its inverse is also an elementary matrix. Therefore, A can be expressed as a product of elementary matrices.

2.7 EROs and Inverses

Elementary Row Operations can be used to find the inverse of a square matrix. Consider a square matrix $A \in M_n(\mathbb{F})$ (that is, an $n \times n$ matrix over a field $\mathbb{F}$). If A is invertible, let

$$A^{-1} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}$$

be its inverse, where each $\mathbf{v}_i$ is the ith column of A^{-1}. By definition of the matrix inverse, we have

$$A\,A^{-1} = A\begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{bmatrix} = \begin{bmatrix}A\mathbf{v}_1 & A\mathbf{v}_2 & \cdots & A\mathbf{v}_n\end{bmatrix} = I_n,$$

the $n \times n$ identity matrix. This implies that

$$A\mathbf{v}_i = \mathbf{e}_i, \quad \text{for each } i = 1, 2, \ldots, n,$$

where $\mathbf{e}_i$ is the ith column of I_n (which has a 1 in the ith row and 0 everywhere else). In other words, each column $\mathbf{v}_i$ of A^{-1} is the unique solution to the linear system

$$A\mathbf{v}_i = \mathbf{e}_i.$$

To find A^{-1} effectively, we form the augmented matrix $[A \mid I_n]$ and apply EROs to trasnform A into I_n. When this is achieved, the augmented portion becomes A^{-1}. Thus, we have

$$\mathrm{RREF}\big([\,A \mid I_n\,]\big) = \big[\,I_n \mid A^{-1}\big].$$

Example 2.3

Find A^{-1} if $A = \begin{bmatrix} 3 & 4 & -1 \\ 1 & 0 & 3 \\ 2 & 5 & -4 \end{bmatrix}$.

We form a 3×6 matrix $A' = [A \mid I_3]$:

$$A' = \left[\begin{array}{ccc|ccc} 3 & 4 & -1 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & 5 & -4 & 0 & 0 & 1 \end{array}\right]$$

We apply the following EROs to A':

- $R_1 \leftrightarrow R_2$
- $R_2 \to R_2 - 3R_1$
- $R_3 \to R_3 - 2R_1$
- $R_3 \to R_3 + R - 2$
- $R_3 \leftrightarrow R_2$
- $R_3 \to R_3 - 4R_2$
- $R_3 \times (-\frac{1}{10})$
- $R_1 \to R_1 - 3R_3$

To obtain:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{2} & -\frac{11}{10} & -\frac{6}{5} \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{array}\right]$$

That is:

$$A^{-1} = \begin{bmatrix} \frac{3}{2} & -\frac{11}{10} & -\frac{6}{5} \\ -1 & 1 & 1 \\ -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{bmatrix}$$

It is easily checked that $AA^{-1} = I_3$.

3 Spanning sets, bases and dimensions

3.1 Vector Spaces

A **vector space V** over $\mathbb{F}$ is a non empty set of objects equipped with an addition operation and whose elements can be multiplied by scalars in $\mathbb{F}$, subject to the following axioms:

1. $u + v = v + u, \quad \forall\, u, v \in \mathbf{V}$
2. $(u + v) + w = u + (v + w), \quad \forall\, u, v, w \in \mathbf{V}$
3. $\exists\, 0_{\mathbf{V}}$, so that $0_v + v = v, \quad \forall\, v \in \mathbf{V}$
4. $\exists\, -v \in \mathbf{V}$, so that $v + (-v) = 0_{\mathbf{V}}, \quad \forall\, v \in \mathbf{V}$
5. $\alpha(\beta v) = \alpha\beta(v), \quad \forall\, \alpha, \beta \in \mathbb{F}, v \in \mathbf{V}$
6. $(\alpha + \beta)v = \alpha v + \beta v, \quad \forall\, \alpha, \beta \in \mathbb{F}, v \in \mathbf{V}$
7. $1v = v, \quad \forall\, v \in \mathbf{V}$

In the definitions axioms above, the field $\mathbb{F}$ can be replaced with any other field, such as $\mathbb{R}$ or $\mathbb{C}$.

Examples of vector spaces over $\mathbb{R}$

- The space $M_{m\times n}(\mathbb{R})$ of $m \times n$ with real entries.
- The space of all polynomials with real coefficients
- The set of complex numbers is a vector space over $\mathbb{R}$.

Consider the space **V** consisting of all **symmetric** 2×2 matrices in $M_2(\mathbb{R})$ with **trace zero**.

- **Trace zero** means that the sum of the diagonal elements is zero.
- **Symmetric** means that the matrix is equal to its transpose.

So a matrix of trace zero has the form:

$$\begin{bmatrix} a & b \\ b & -a \end{bmatrix} \quad \text{where } a, b \in \mathbb{R}$$

Since it takes two real number to specify an element of $\mathbb{V}$, this is another example of a 2-dimensional vector.

3.2 Subspaces

Definition 3.1: Vector Subspaces

Let **V** be a vector space over a field $\mathbb{F}$. A subset **U** is a **subspace** of **V** if **U** is itself a a vector space over $\mathbb{F}$, under the addition and scalar multiplication operations defined on **V**.

Two things need to be checked to confirm that $U \subseteq V$ is a subspace:

1. $\mathbb{U}$ is **closed** under the addition in **V**, i.e. $u_1 + u_2 \in \mathbf{U}$ for all $u_1, u_2 \in \mathbf{U}$.
2. $\mathbb{U}$ is **closed** under scalar multiplication, i.e. $\alpha u \in \mathbb{U}$, whenever $u \in \mathbb{U}$ and $\alpha \in \mathbb{F}$.

Examples of subspaces

1. Let $\mathbb{Q}[x]$ be the set of all polynomials with rational coefficients. Let $P_2 \subseteq \mathbb{Q}[x]$ be the set of all polynomials of degree at most 2. This means $P_2 = \{a_2x^2 + a_1x + a_0 : a_0, a_1, a_2 \in \mathbf{Q}\}$ Then P_2 is a vector subspace of $\mathbb{Q}[x]$. If $f(x)$ and $g(x)$ are rational polynomials of degree at most 2, then also is $f(x) + g(x)$ and $\alpha f(x)$, where $\alpha \in \mathbb{Q}$.
2. The set $\mathbb{C}$ is a vector space over the set of real numbers. Within $\mathbb{C}$, the subset $\mathbb{R}$ is an example of a vector subspace over $\mathbb{R}$. An example of a subset of $\mathbb{C}$ that is not a real vector subset is the unit circle S in the complex plane- this is the set of complex numbers of modulus 1, it consists of all complex numbers of the form a+bi, where $a^2 + b^2 = 1$. This is closed neither under addition nor multiplication by real scalars.
3. The Cartesian plane ($\mathbb{R}^2$) is a real vector space. Within $\mathbb{R}^2$, let $U = \{(a, b) : a > 0, b > 0\}$. Then **U** is closed under addition and under multiplication by positive scalars. It is not a vector subspace of $\mathbb{R}^2$,because it is not closed under multiplication by negative scalars
4. Let v be a fixed non-zero vector $\in \mathbb{R}^3$ and let $v^\perp = \{u \in \mathbb{R}^3 : u^Tv = 0\}$. Then $v^\perp$ is not empty since $0 \in v^\perp$.Suppose $u_1, u_2 \in v^\perp$. If $u \in v^\perp$ and $\alpha \in \mathbb{R}$, then $(\alpha u)^Tv = \alpha u^Tv = 0 = \alpha 0 = 0$. Hence $v^\perp$ is closed under scalar multiplication. Thus $v^\perp$ is a vector subspace of $\mathbb{R}^3$. Note that $v^\perp$ is not all $\mathbb{R}^3$, since $v \notin v^\perp$.

3.3 Span of a set of vectors

Definition 3.2: Span

Let $\mathbf{V}$ be a vector space over a field $\mathbb{F}$, and let S be a non empty subset of $\mathbf{V}$.

The $\mathbb{F}$-linear span, commonly called the **span** of S, denoted $\langle S \rangle$, is the set of all $\mathbb{F}$-linear combinations of the elements of $S \in \mathbf{V}$.

If $S = V$, then S is called a spanning set of V; meaning that every element of $\mathbf{V}$ is a linear combination of the elements of S.

For a subset S of a $\mathbb{F}$-vector space $\mathbf{V}$, the sum of any two linear combinations of S is an element of S, and any scalar multiple of a linear combination of S is also an element of S; hence the following lemma:

Lemma 3.1

For any subset, S, of a vector space, $\mathbf{V}$, the span, $\langle S \rangle$, is a subspace of $\mathbf{V}$.

Examples

- **Polynomials over $\mathbb{Q}$**
  $Q[x]$ is the set of all polynomials with rational coefficients, and $P_2 \subset Q[x]$ consists of polynomials of degree at most 2. If $S = \{x^2 + 1,\ x + 1\}$, then
  $$\langle S \rangle \ = \ \{\, a(x^2 + 1) \ + \ b(x + 1) \ : \ a, b \in \mathbb{Q}\}.$$
  All members of $\langle S \rangle$ are degree-$\leq$ 2 polynomials with constant term equal to the sum of the x- and x^2-coefficients. For instance, $x^2 + 2x + 3 \in \langle S \rangle$ but $x^2 + 2x + 4 \notin \langle S \rangle$. Since $\langle S \rangle$ does not include all degree-$\leq$ 2 polynomials in P_2, S is not a spanning set for P_2 over $\mathbb{Q}$.
- **Column vectors in $\mathbb{R}^2$**
  Let
  $$S = \{(3, 1),\ (2, 1),\ (1, -1)\}.$$
  Any vector $(a, b) \in \mathbb{R}^2$ can be written as a linear combination of these three vectors in more than one way. However, $(1, -1)$ itself is a linear combination of $(3, 1)$ and $(2, 1)$, so it is not necessary to span $\mathbb{R}^2$. Hence S has redundant elements and is not a minimal spanning set of $\mathbb{R}^2$.

The second example above motivates the following lemma:

Lemma 3.2

Suppose that $S_1 \subset S$, where $S \subseteq \mathbf{V}$, then
$$\langle S_1 \rangle \subseteq \langle S \rangle$$
if and only if every element of $S \backslash S_1$ is a linear combination of the elements of S_1.

Definition 3.3

- **Finite dimensional**: A vector space that has a finite spanning set
- **Infinite dimensional**: A vector space that has an infinite spanning set

Example of infinite dimensional vector space

- The vector space $\mathbb{R}[x]$ of all polynomials with real coefficients is infinite dimensional. To see this let S be a finite subset of $\mathbb{R}[x]$ and let x^k be the highest power of x in S. Then $x^{k+1} \notin \langle S \rangle$ since x^{k+1} cannot be expressed as a linear combination of the elements of S.
- The set of $\mathbb{R}$ is infinite dimensional as a vector space over the field, $\mathbb{Q}$, of rational numbers.

3.4 Linear independence

Definition 3.4

Let $S \subseteq \mathbf{V}$ with at least two elements.

Then S is linearly independent if **no element of S can be expressed as a linear combination of the other elements of S.**

Equivalently, if no element of S belongs to the span of the other elements of S.

It follows, a subset consisting of a single element is linearly independent if and only if that element is non-zero. The definition above takes a lot of work to check for large sets, the following definition is often more useful:

Definition 3.5

Let S be a non-empty subset of $\mathbf{V}$.

Then S is **linearly independent** if the only linear combination of the elements of S that equals zero is to take all the coefficients to be zero.

Equivalence of the two definitions

Let $S = \{v_1, \dots, v_k\}$ and suppose $v_1 \in \langle v_2, \dots v_k \rangle$. Then:

$$v_1 = \alpha_2 v_2 + \cdots + \alpha_k v_k \quad \Rightarrow \quad 0 = -v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

is an expression for the zero vector as a linear combination of elements of S, whose coefficients are not all zero. On the other hand suppose:

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$$

where not all $c_i = 0$ Then:

$$v_1 = -\frac{c_2}{c_1} v_2 - \cdots - \frac{c_k}{c_1} v_k \quad \Rightarrow \quad v_1 \in \langle v_2, \dots, v_k \rangle$$

Example 3.1

In $\mathbb{R}^3$, let $S = \{[1,2,-1], [-2,3,2], [-3,8,3]\}$. Show that S is linearly independent.

To determine if S is linearly independent, we need to investigate whether the system of equations has solutions other than $(x,y,z) = (0,0,0)$:

$$x \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -2 \\ 3 \\ 2 \end{bmatrix} + z \begin{bmatrix} -3 \\ 8 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad \left[\begin{array}{ccc|c} 1 & -2 & -3 & 0 \\ 2 & 3 & 8 & 0 \\ -1 & 2 & 3 & 0 \end{array}\right]$$

Reducing it to its RREF we get:

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \Longrightarrow \quad \begin{array}{l} x + t = 0 \\ y + 2t = 0 \\ z + t = 0 \end{array} \quad \Longrightarrow \quad (x,y,z) = (-t, -2t, t)$$

Setting $t = 1$ gives:

$$-1 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - 2 \begin{bmatrix} -2 \\ 3 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} -3 \\ 8 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Hence, each of the three elements of S is a linear combination of the other two. So S is not linearly independent (we say it is linearly dependent).

3.5 Characterizations of Linear Independence

Theorem

Let S be a subset of a vector space V. The following are equivalent:

1. S is linearly independent if and only if S is a minimal spanning set of $\langle S \rangle$ – no proper subset of S spans $\langle S \rangle$.
2. S is linearly independent if and only if every element of $\langle S \rangle$ has a unique expression as a linear combination of elements of S.
3. S is linearly independent if and only if every element of $\langle S \rangle$ has unique coordinates with respect to the elements of S.

Definition 3.6: Basis

A **basis** of a vector space V is a spanning set of V that is linearly independent.

Lemma 3.3

If S is a finite spanning set of a vector space V, then S contains a basis of V.

Proof: If S is not linearly independent, then some element $v_1 \in S$ is in the span of $S \setminus \{v_1\}$. Let $S_1 = S \setminus \{v_1\}$, which still spans V. Continue this process, removing elements that are linearly dependent on the remaining ones. Since S is finite, this process terminates with a linearly independent spanning set of V.

Theorem Steinitz Exchange Lemma

Let V be a vector space with a spanning set $S = \{v_1, \dots, v_n\}$. Then any linearly independent subset L of V contains at most n elements.

Proof: Let $L = \{y_1, \dots, y_k\}$ be linearly independent. The key idea is to replace elements of S with elements of L one by one, maintaining a spanning set:

- Express y_1 as a linear combination of elements in S. At least one element, say v_1, must have a non-zero coefficient.
- Replace v_1 with y_1 to get $S_1 = \{y_1, v_2, \dots, v_n\}$, which still spans V.
- Continue this process, replacing v_i with y_i at each step.
- This can continue for at most n steps, so $k \leq n$.

Theorem

If V is a finite-dimensional vector space, then every basis of V has the same number of elements.

Proof: Let B_1 and B_2 be bases of V. Since B_1 is linearly independent and B_2 spans V, we have $|B_1| \leq |B_2|$ by the Steinitz Exchange Lemma. Similarly, $|B_2| \leq |B_1|$. Therefore, $|B_1| = |B_2|$.

Definition 3.7: Dimension

The dimension of a finite-dimensional vector space V, denoted $\dim V$, is the number of elements in any basis of V.

3.6 Properties of Bases in Finite Dimensional Vector Spaces

Let V be a vector space of dimension n over a field F. Recall that a basis of V has two key properties:

- It is a linearly independent set in V
- It spans V

A basis is therefore both a minimal spanning set and a maximal linearly independent set.

Theorem Characterization of Bases by Cardinality

For a finite-dimensional vector space V with $\dim V = n$:

1. Every linearly independent subset with exactly n elements is a basis.
2. Every spanning set with exactly n elements is a basis.

Proof of (1). Let $L = \{v_1, \ldots, v_n\}$ be a linearly independent subset of V.

Suppose L is not a spanning set. Then there exists some $v \in V$ with $v \notin \langle L \rangle$.

This means the set $L' = \{v_1, \ldots, v_n, v\}$ would be linearly independent in V, contradicting Theorem 2.2.6 (which states that no linearly independent set can have more than n elements).

Therefore, L must be a spanning set and thus a basis. □

Proof of (2). Let S be a spanning set of V with n elements.

If S is not linearly independent, then S contains a proper subset that spans V but has fewer than n elements, contradicting Theorem 2.2.6 (which implies all bases have exactly n elements).

Therefore, S must be linearly independent and thus a basis. □

Theorem Basis Extension Theorem

Any linearly independent subset L of V can be extended to a basis of V.

Proof. Let $L = \{v_1, \ldots, v_k\}$. Since L is linearly independent, $k \leq n$ by Theorem 2.2.6.

If $k = n$, then L is already a basis by part (1) above.

If $k < n$, then L does not span V, so there exists $v_{k+1} \in V$ with $v_{k+1} \notin \langle L \rangle$. Then $\{v_1, \ldots, v_k, v_{k+1}\}$ is linearly independent.

We can continue this process, adding elements outside the existing span until we reach n elements, which will form a basis of V. □

Theorem Isomorphism with F^n

Every n-dimensional vector space V over F is isomorphic to the standard vector space F^n.

Sketch. Let $B = \{v_1, \ldots, v_n\}$ be a basis of V over F.

For any $v \in V$, there exists a unique expression $v = a_1 v_1 + \cdots + a_n v_n$.

The mapping $v \mapsto \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ defines a bijective correspondence between V and F^n that preserves vector space operations.

Different bases of V correspond to different isomorphisms with F^n. □

Definition 3.8: Standard Basis

The standard basis of F^n is $\{e_1, \ldots, e_n\}$, where e_i has 1 in position i and 0 in all other positions.