MA283: Linear Algebra

 $\begin{array}{c} 70\% \ {\rm Exam} \\ 30\% \ {\rm Continuous \ Assessment \ (Homework)} \\ 10\% \ {\rm Optional \ Project \ (Bonus)} \end{array}$

Robert Davidson

${\bf Contents}$

1	Sys	tems of linear equations
	1.1	Linear equations and Solution Sets
	1.2	Elementary Row Operations
		1.2.1 REF and Strategy
		1.2.2 Row Reduced Echelon Form
	1.3	Leading variables and free variables
	1.4	Consistent and Inconsistent Systems
	1.5	Possible Outcomes when solving a system of equations
	1.6	Elementary Row Operations as Matrix Transformations
		1.6.1 Multiplying a Row by a Non-Zero Scalar
		1.6.2 Switching Two Rows
		1.6.3 Adding a Multiple of One Row to Another
	1.7	EROs and Inverses
	1.8	Review of Matrix Algebra
		1.8.1 Fields
		1.8.2 Matrices Over a Field
		1.8.3 Addidion and Scalar Multiplication
		1.8.4 Linear Combinations
		1.8.5 Row and Column Vectors
		1.8.6 Matrix-Vector Multiplication
		1.8.7 Matrix-Matrix Multiplication
		1.8.8 Dot Product and Orthogonality
		1.8.9 Matrices and Tables
		1.8.10 Linear Transformations
		1.8.11 Composition of Linear Transformations

1 Systems of linear equations

1.1 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 1.1: 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 1.2: Solution set

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

$$a_1X_1 + a_2X_2 + \ldots + a_nX_n = b$$
 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 .

1.2 Elementary Row Operations

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

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).

1.2.1 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**).
- $\bullet\,$ If a column has a leading 1, then all entries below it are 0.
- $\bullet\,$ 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.

Example of REF

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

Concept 1.1: 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

1.2.2 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.

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

Example of RREF

1.3 Leading variables and free variables

We'll start by an example:

Solving this system of equations, we get:

RREF:
$$\begin{bmatrix} 1 & 0 & 0 & 2 & | & 4 \\ 0 & 1 & 0 & -1 & | & 2 \\ 0 & 0 & 0 & 1 & | & 2 \end{bmatrix} \Rightarrow \begin{array}{c} x_1 + 2x_4 = 4 \\ x_2 - x_4 = 2 \\ x_3 + x_4 = 2 \end{array} \Rightarrow \begin{array}{c} x_1 = 4 - 2x_4 \\ x_2 = 2 + x_4 \\ x_3 = 2 - x_4 \\ x_3 = 2 - x_4 \\ x_4 = 2 \end{array}$$

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:

$$x_1 = 4 - 2t$$
, $x_2 = 2 + t$, $x_3 = 2 - t$, $x_4 = t$ 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 1.3: 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

1.4 Consistent and Inconsistent Systems

Consider the following system of equations:

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.

1.5 Possible Outcomes when solving a system of equations

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

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

- $\bullet\,$ The system may be ${\bf consistent}$ which occurs if:
 - Unique Solutions each column (aside from the rightmost) contains a single leading 1. i.e:

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

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

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

1.6 Elementary Row Operations as Matrix Transformations

Elementary row operations may be interpreted as **matrix multiplication**. To see this, first we introduce the **Identity matrix:**[1, 0, 0]

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

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."

1.6.1 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:

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

1.6.2 Switching Two Rows

To swap rows i and k, we use:

$$S = I_m + E_{ik} + E_{ki} - E_{ii} - E_{kk}$$

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:

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

1.6.3 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 1.1

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 $1 \setminus \alpha$. 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 1.2

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

Multiplying a row by a scalar: $I_n + (\alpha - 1)E_{i,i}$

Swapping two rows: $I_n + E_{i,k} + E_{k,i} - E_{i,i} - E_{k,k}$

Adding a multiple of one row to another: $I_n + \alpha E_{k,i}$

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.

1.7 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

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

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

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

where \mathbf{e}_i is the *i*th column of I_n (which has a 1 in the *i*th 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 transform A into I_n . When this is achieved, the augmented portion becomes A^{-1} . Thus, we have

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

Example 1.3

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

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

$$A' = \begin{bmatrix} 3 & 4 & -1 & | & 1 & 0 & 0 \\ 1 & 0 & 3 & | & 0 & 1 & 0 \\ 2 & 5 & -4 & | & 0 & 0 & 1 \end{bmatrix}$$

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:

$$\begin{bmatrix} 1 & 0 & 0 & | & \frac{3}{2} & -\frac{11}{10} & -\frac{6}{5} \\ 0 & 1 & 0 & | & -1 & 1 & 1 \\ 0 & 0 & 1 & | & -\frac{1}{2} & \frac{7}{10} & \frac{2}{5} \end{bmatrix}$$

That is:

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

7

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

1.8 Review of Matrix Algebra

1.8.1 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.

1.8.2 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 $M_{m \times n}(\mathbb{F})$.

1.8.3 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}

1.8.4 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)

1.8.5 Row and Column Vectors

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

1.8.6 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.

1.8.7 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:

$$(AB)_{ij} = A_{i,1}B_{1,j} + A_{i,2}B_{2,j} + \ldots + A_{i,p}B_{p,j} = \sum_{k=1}^{p} A_{i,k}B_{k,j}$$

1.8.8 Dot Product and Orthogonality

• For two p-entry vectors, $u,v\in\mathbb{F}^p,$ their dot product is:

$$u \cdot v = \sum_{k=1}^{p} u_k v_k$$

- 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.

1.8.9 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	Р	С	
2015 2016 2017	50	100	70	
2016	60	80	80	
2017	70	90	90	

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

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	
$^{\mathrm{C}}$	20	10	

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

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

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

1.8.10 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)
- $T(\lambda u) = \lambda T(u)$

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

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 $v \in \mathbb{R}^3$.

1.8.11 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:

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

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