7. *If $\alpha, \beta \in \mathbb{F}$ and $v \in V$, then $(\alpha + \beta)v = \alpha v + \beta v$ (distributivity of scalar mulatiplication over addition in $\mathbb{F}$).*

8. *$1_{\mathbb{F}} v = v$ for all $v \in V$, where $1_{\mathbb{F}}$ is the multplicative identity element of $\mathbb{F}$.*

When working with examples, we rarely have to worry very much about any of the technical specifications in Definition 1.1.4; they tend to be clearly satisfied. But there is a value in devoting some time to studying each of these axioms carefully. When a class of objects (like vector spaces) is defined in terms of the algebraic operations, it is important to specify the properties that the operations must have in order to qualify for admission to the class. Even if these propoerties are natural and obvious in the examples that motivate the definition, it is still important to note them systematically, they form the rules of the game and a key reference point in the development of the theory.

Exercise 1.1.5. *Check that the axioms are satisfied in the examples that we have met so far.*

1.2 Review of Matrix Algebra

A $m \times n$ matrix over a field $\mathbb{F}$ is an an array of m rows and n columns, whose entries are elements of $\mathbb{F}$. We can take $\mathbb{F}$ to be the field of real numbers. The expression $m \times n$ is referred to as the *size* of a matrix (even though what it really describes is the *shape*). Two matrices can be added together if they have the same size; in this case their sum is obtained by just adding the entries in each position. The $m \times n$ zero matrix is the $m \times n$ matrix whose entries are all zeros. It is the *identity element* for addition of $m \times n$ matrices - this means that addition it to another $m \times n$ matrix has no effect. A matrix can be multiplied by a scalar; this means multiplying each of its entries by that scalar. With these operations of addition and scalar multiplication, the set of $m \times n$ matrices over a field $\mathbb{F}$ is a *vector space* over $\mathbb{F}$.
Notation: We use the notation $M_{m \times n}(\mathbb{F})$ for the vector space of all $m \times n$ matrices over $\mathbb{F}$. When $m = n$, we abbreviate this to $M_n(\mathbb{F})$.

Example 1.2.1. *In $M_{2\times 3}(\mathbb{R})$,*

$$2\begin{pmatrix} 1 & 0 & -1 \\ 2 & -5 & 1 \end{pmatrix} - 3\begin{pmatrix} 2 & 4 & -1 \\ 0 & 1 & -3 \end{pmatrix} = \begin{pmatrix} 2(1)-3(2) & 2(0)-3(4) & 2(-1)-3(-1) \\ 2(2)-3(0) & 2(-5)-3(1) & 2(-3)-3(-3) \end{pmatrix} = \begin{pmatrix} -4 & -12 & 1 \\ 4 & -13 & 3 \end{pmatrix}.$$

We can sometimes also *multiply* matrices, but the way to do this is not as obvious. We begin with a few definitions.

Definition 1.2.2. *Suppose that $v_1, v_2, \ldots, v_k$ are elements of a vector space V over a field $\mathbb{F}$. A $\mathbb{F}$-linear combination (or just linear combination) of $v_1, \ldots, v_k$ is an element of V that has the form $a_1 v_1 + a_2 v_2 + \cdots + a_k v_k$, wher the a_i are elements of $\mathbb{F}$. In this situation the a_i are called the* coefficients *in the linear combination.*

The example above shows a linear combination of two matrices in $M_{2\times 3}(\mathbb{R})$, with coefficients 2 and -3.

Definition 1.2.3. *A* column vector *is a matrix with one column. A* row vector *is a matrix with one row.*

Before defining matrix multiplication in general, we define the product of a matrix with a column vector (when that exists).

Definition 1.2.4. *Let A be a $m \times n$ matrix and let v be a column vector with n entries. Then the matrix-vector product Av is the column vector obtained by taking the linear combination of the columns of A whose coefficients are the entries of v. It is a column vector with m entries.*

Example 1.2.5. $\begin{pmatrix} 2 & 4 & -1 \\ 0 & 1 & -3 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 0 \end{pmatrix} + (-2)\begin{pmatrix} 4 \\ 1 \end{pmatrix} + 1\begin{pmatrix} -1 \\ -3 \end{pmatrix} = \begin{pmatrix} -3 \\ -5 \end{pmatrix}$

Note that the product Av is only defined when the number of entries in a row of A (i.e. the number of columns in A) is equal to the number of entries in v. In the same way, if u is a row vector with m entries, and A is a $m \times n$ matrix, then the vector-matrix product uA is the linear combination of the rows of A whose coefficients are the entries of u. It is a row vector with n entries (a $1 \times n$ matrix).

Definition 1.2.6. *Let A and B be matrices of size $m \times p$ and $p \times n$ respectively. Write $v_1, \dots, v_n$ for the columns of B. Then the product AB is the $m \times n$ matrices whose columns are $Av_1, \dots, Av_n$.*

Remarks

1. This version of the definition of the matrix product AB emphasizes that we can think about the matrix B as being an arrangement of n column vectors placed side by side. This viewpoint can be quite useful, but maybe not as the only way to think about matrices. But in this situation it allows use to understand matrix multiplication as a straightforward extension of matrix-vector multiplication.

2. Exercise: write down an alternative version of Definition 1.2.6, that emphasises vector-matrix products of the rows of A with the matrix B.

3. If the number of entries in a row of A (the number of columns of A) is not equal to the number of entries in a column of B, then the product AB is not defined.

Matrix products are often presented and explained just in terms of their individual entries. This viewpoint is sometimes convenient and it is quite standard, and it gives us an opportunity to introduce some notation that essential for linear algebra. Suppose that A is a $m \times p$ matrix and B is a $p \times n$ matrix, with entries in a field $\mathbb{F}$. The rows of A are labelled Row 1 through Row m, from top to bottom, and the columns of A are labelled Column 1 through Column p, from left to right (similar story for B). The entry in Row i and Column j of A is denoted A_{ij}. So A_{11} is the entry in the upper left corner of A. Now AB is the product of a $m \times p$ and a $p \times m$ matrix: it is a $m \times n$ matrix. According to Definition 1.2.6, the entry in the the (i, j) position of AB (i.e. Row i and Column j) is the ith entry of the vector Av_j, where the vector v_j is Column j of B. Again according to Definition 1.2.6, this is the ith entry of the linear combination of the columns of A with coefficients from the jth column of B. This is the linear combination of the ith entries of the columns of A (i.e. the entries of Row i of A, with coefficients from Column j of B). It is given by

$$(AB)_{ij} = A_{i1}B_{ij} + A_{i2}B_{2j} + \cdots + A_{ip}B_{pj} = \sum_{k=1}^{p} A_{ik}B_{kj}.$$

It is worth taking some time to get used to the notation in the above line if it is not already familiar.

We note that the expression for $(AB)_{ij}$ above involves the *scalar product* of two vectors with p entries. For a field $\mathbb{F}$, we write $\mathbb{F}^p$ for the vector space of all vectors with p entries from $\mathbb{F}$. Sometimes we will need to specify whether we mean row vectors or column vectors, but for now we will cheat and allow every ordered list of p elements of $\mathbb{F}$ to be considered as an element of $\mathbb{F}^p$, regardless of how it is written. For example we might consider elements of $\mathbb{R}^3$ to be written as coordinates of a point, like $(1, 1, 3)$, or as column vectors with three real entries.

Definition 1.2.7. *Let $u = (a_1, \dots, a_p)$ and $v = (b_1, \dots, b_p)$ be vectors in $\mathbb{F}^p$. Then the* ordinary scalar product *or* dot product *of u and v is the element of $\mathbb{F}$ defined by*

$$u \cdot v = a_1b_1 + a_2b_2 + \cdots + a_pb_p = \sum_{k=1}^{p} a_kb_k.$$

If $u \cdot v = 0$, we say that u and v are *orthogonal* with respect to the scalar product. If $\mathbb{F} = \mathbb{R}$, this means that the vectors u and v are perpendicular in Euclidean space.

We may now observe that if A and B are respectively a $m \times p$ and a $p \times n$ matrix, then the entry in the (i, j)-position of the product AB is the scalar product of Row i of A and Column j of B, both regarded as vectors in $\mathbb{F}^p$. The product AB itself is a table of values of scalar products of Rows of

A with Columns of B. If we write $u_1, \dots, u_m$ for the rows of A (vectors in $\mathbb{F}^p$) and $v_1, \dots, v_n$ for the columns of B (vectors in $\mathbb{F}^p$), then

$$AB = \begin{pmatrix} u_1 \cdot v_1 & u_1 \cdot v_2 & \dots & u_1 \cdot v_n \\ u_2 \cdot v_1 & u_2 \cdot v_2 & \dots & u_2 \cdot v_n \\ \vdots & \vdots & & \vdots \\ u_m \cdot v_1 & u_m \cdot v_2 & \dots & u_m \cdot v_n \end{pmatrix}$$

1.2.1 Two ways to think about a matrix

The definition of matrix multiplication can look a bit obscure, if it is presented purely in terms of how the entries of A and B are combined to produce the entries of AB. It does make sense however, even in very practical contents as in the following example.

Example 1.2.8. One way for a matrix to arise is as a table of data from some "real" (i.e. not just mathematical) process. Remarkably, the operations of matrix algebra can have a meaning even in this context. As an example, let A be the 3×3 matrix formed by the table that gives the numbers of first year Humanities (H), Engineering (E) and Science (S) students in first year at Eigen University, in 2015, 2016 and 2017.

	H	E	S
2015	50	100	70
2016	60	80	80
2017	80	70	70

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

Every fiirst year student at Eigen University takes either Linear Algebra (LA) or Calculus (C) or both. The table below shows the numbers of ECTS credits completed annually in each, by students in each of the three subject areas.

	LA	C
H	10	0
E	15	15
S	20	10

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

Now look at the meaning of the entries of the product AB.

$$\begin{aligned} AB &= \begin{pmatrix} 50 & 100 & 70 \\ 60 & 80 & 80 \\ 80 & 70 & 70 \end{pmatrix} \begin{pmatrix} 10 & 0 \\ 15 & 15 \\ 20 & 10 \end{pmatrix} = \begin{pmatrix} 50(10) + 100(15) + 70(20) & 50(0) + 100(15) + 70(10) \\ 60(10) + 80(15) + 80(20) & 60(0) + 80(15) + 80(10) \\ 80(10) + 70(15) + 70(20) & 80(0) + 70(15) + 70(10) \end{pmatrix} \\ &= \begin{pmatrix} 3400 & 2200 \\ 3400 & 2000 \\ 3250 & 1750 \end{pmatrix}. \end{aligned}$$

The entries in the first column of AB are the total numbers of linear algebra credits taken by first year students in 2015, 2016 and 2017. In the second column are the total numbers of calculus credits in each of the three years. The matrix product AB represents the following table of data

	LA credits	C credits
2015	3400	2200
2016	3400	2000
2017	3250	1750

Another way to interpret matrix multiplication is in terms of *linear transformations*, which are the primary functions between vector spaces that are of interest in linear algebra. For now we will stick to linear transformations between spaces of real column vectors.

Definition 1.2.9. *Let m and n be positive integers. A linear tranformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ that satisfies*

- $T(u + v) = T(u) + T(v)$, *and*
- $T(\lambda v) = \lambda T(v)$,

for all u *and* v *in* $\mathbb{R}^n$, *and all scalars* $\lambda \in \mathbb{R}$.

Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation. Then we can calculate the image under T of any vector $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, if we know the images under T of the *standard basis vectors* $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. From the definition, we have

$$T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = aT\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + bT\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + cT\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = A\begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

where A is the 2×3 matrix that has the images of the three standard basis vectors as its three columns.

Example 1.2.10. *Suppose that the images of the three standard basis vectors under* T *are respectively* $\binom{1}{2}$, $\binom{1}{4}$ *and* $\binom{-2}{3}$. *Then the matrix* A *of* T *is*

$$A = \begin{pmatrix} 1 & 1 & -2 \\ 2 & 4 & 3 \end{pmatrix}.$$

For any vector $v \in \mathbb{R}^3$, *its image under* T *is the matrix-vector product* Av. *For example*

$$T\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -2 \\ 2 & 4 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -3 \\ 8 \end{pmatrix}$$

The matrix A may be considered as a representation of the linear transformation T. Now suppose that $S : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation whose matrix is $B = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$. This means that the images under S of $\binom{1}{0}$ and $\binom{0}{1}$ are respectively the two columns of S. Now the composition $S \circ T$ is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$, so is is represented by a matrix. How dos this matrix depend on A and B. To answer this, think about calculating the image of a vector v under the composition $S \circ T$ (S *after* T).

$$S \circ T(v) = S(Tv) = S(Av) = B(Av) = (BA)v. \tag{1.1}$$

This is saying that the matrix of the transformation $S \circ T$ is the matrix product BA, where B and A are respectively the matrices of S and T. Thus matrix multiplcation may be interpreted as corresponding to composition of linear transformations.

Two things to note about 1.1:

1. It uses the fact that matrix multiplication is *associative*, i.e. $(AB)C = A(BC)$, whenever A, B, C are matrices for which the products AB and BC are defined. It is true but not entirely obvious that this property holds. Something to think about.

2. We have also used the fact that the composition of two linear transformations (when it is defined) is a linear transformation. If $f : \mathbb{R}^n \to R^p$ and $g : \mathbb{R}^q \to \mathbb{R}^m$ are linear transformations, then the composition $f \circ g$ is defined only if $p = q$, and in this case $f \circ g$ is a linear transformation from R^n to $\mathbb{R}^m$. This is equivalent to the statment that the product of a $m \times q$ with a $p \times n$ matrix is defined only if $p = q$, and in this case it is a $m \times n$ matrix.

1.2.2 Some more concepts from matrix algebra

This short section notes some objects and notation that we will need throughout the course.
The $n \times n$ *Identity Matrix*
For a positive integer n, the $n \times n$ identity matrix, denoted I_n, is the $n \times n$ matrix whose entries in the $(1,1),(2,2),\dots,(n,n)$ positions (the positions on the main diagonal) are all 1, and whose entries in all other positions (all off-diagonal positions) are 0. For example

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The special property that I_n has is that is an *identity element* or *neutral element* for matrix multplication. Multiplying another matrix by it has no effect. This means

- If A is any matrix with n rows, then $I_n A = A$, and

- If B is any matrix with n columns, then $BI_n = B$.

- In particular, if C is a $n \times n$ matrix, then $CI_n = I_n C = C$.

Exercise 1.2.11. *Using our intepretations of matrix multiplication so far, explain why the matrix* I_n *has this neutral property. For example, use Definition 1.2.4 to describe what happens when a column vector with* n *entries is multiplied on the left by* I_n*.*
What is the linear transformation that is represented by I_n*? Why is it that composing this linear transformation with any other has no effect?*

The Inverse of a Matrix
Let A be a square matrix of size $n \times n$. If there exists a $n \times n$ matrix B for which $AB = I_n$ and $BA = I_n$, then A and B are called *inverses* (or *multiplicative inverses*) of each other. If it does not already have another name, the inverse of A is denoted A^{-1}. The relationship between A and A^{1} resembles that of two rational numbers that are reciprocals of each other, such as $\frac{5}{31}$ and $\frac{31}{5}$. Their product is the identity element for multiplication (1 in the case of the rational numbers) and so multiplying by one of them reverses the work of multiplying by the other. When applying this general principle in the case of matrices, we need to remember that matrix multiplication is not commutative.

Example 1.2.12. *In* $M_2(\mathbb{Q})$*, the matrices* $\begin{pmatrix} 3 & 2 \\ -5 & -4 \end{pmatrix}$ *and* $\begin{pmatrix} 2 & 1 \\ -\frac{5}{2} & -\frac{3}{2} \end{pmatrix}$ *are inverses of each other.*

Not every square matrix has an inverse. For example the 2×2 matrix $\begin{pmatrix} 3 & 2 \\ -6 & -4 \end{pmatrix}$ does not. (Challenge: prove this without using any knowledge about the determinant of a 2×2 matrix).

Exercise 1.2.13. *Prove that a square matrix can have only one inverse.*
(Hint: If both B *and* C *are inverses of the square matrix* A*, think about the product* BAC*.)*

Another important ingredient of matrix algebra that we will need is the *determinant* of a square matrix. The determinant of a matrix in $M_n(\mathbb{F})$ is an element of $\mathbb{F}$ that is defined in a complicated way in terms of the matrix entries (it is not too bad if $n = 2$ or $n = 3$, but in general it is complicated to describe and to calculate). A square matrix has an inverse if and only if its determinant is not zero. We will define the determinant later.

The Transpose

Definition 1.2.14. *The transpose of the* $m \times n$ *matrix* A*, which is denoted* A^T*, is defined to be the* $n \times m$ *matrix which has the entries of Row 1 of* A *in its first column, the entries of Row 2 of* A *in its second column, and so on.*

Example 1.2.15. *If* $A = \begin{pmatrix} 1 & -2 & -3 \\ 2 & 0 & 4 \end{pmatrix}$, *then* $A^T = \begin{pmatrix} 1 & 2 \\ -2 & 0 \\ -3 & 4 \end{pmatrix}$.

For all relevant i and j, the (i,j) entry of A^T is the (j,i) entry of A. If A is $m \times n$, then the products AA^T and A^TA always exist, and they are square matrices of size $m \times m$ and $n \times n$ respectively. Moreover, they are *symmetric*. A square matrix is symmetric if it is equal to its own transpose.

In the above example,

$$AA^T = \begin{pmatrix} 14 & -10 \\ -10 & 20 \end{pmatrix}, \quad A^TA = \begin{pmatrix} 5 & -2 & 5 \\ -2 & 4 & 6 \\ 5 & 6 & 25 \end{pmatrix}.$$

The following lemma describes how the transpose of a matrix product AB depends on the transposes of A and B. The proof is included as an example of how demonstrations in matrix theory often involve the close inspection and description of entries in particular positions. They are fussy and not easy to read, but they often don't involve anything more than careful attention to the details (all the information is contained in the subscripts!). It takes a bit of concentration to get used to arguments like this and to write them (and read them), but you are encouraged to study this one and satisfy yourself that you could produce it yourself if necessary.

Lemma 1.2.16. *Let* A *and* B *be matrices for which the product* AB *is defined. Then* $(AB)^T = B^TA^T$.

The lemma is saying that the transpose of the product A is the product of the transposes of A and B, but in the opposite order. To prove this, we analyse the entry in an arbitrary position (i,j) of $(AB)^T$, noting that this is the (j,i)-entry of AB.

Proof. Suppose that the sizes of A and B are $m \times p$ and $p \times n$ respectively. Choose an arbitrary position (i,j) in $(AB)^T$. The entry in this position is

$$\begin{aligned} (AB)^T_{ij} &= (AB)_{ji} \\ &= \sum_{k=1}^{p} A_{jk}B_{ki} \\ &= \sum_{k=1}^{p} B^T_{ik}A^T_{kj} \\ &= (B^TA^T)_{ij}. \end{aligned}$$

□

If you are thinking about this proof, note that the third last line describes how the (j,i)-entry of AB is assembled from Row j of A and Column i of B. The next line rewrites that in terms of entries of B^T and A^T. Note that $B^T_{ik} = B_{ki}$ and $A^T_{kj} = A_{jk}$.

Exercise 1.2.17. *Use Lemma 1.2.16 to explain why the matrices* AA^T *and* A^TA *are always symmetric.*

1.3 Systems of Linear Equations

Consider the equation

$$2x + y = 3.$$

This is an example of a *linear equation* in the variables x and y. As it stands, the statement "$2x + y = 3$" is neither true nor untrue : it is just a statement involving the abstract symbols x and y. However if we replace x and y with some particular pair of real numbers, the statement will become either true or false. For example