MA106

Linear Algebra Revision Guide

Written by Shriti Somaiya and Jonathan Elliott

Contents

Introduction

This revision guide for MA106 Linear Algebra has been designed as an aid to revision, not a substitute for it. Linear Algebra is a fairly abstract theoretical course, and this guide should contain most of the theory. However, being able to apply the theorems is also important, since it tests your understanding.

Disclaimer: Use at your own risk. No guarantee is made that this revision guide is accurate or complete, or that it will improve your exam performance. Use of this guide *will* increase entropy, contributing to the heat death of the universe. Contains no GM ingredients. Your mileage may vary. All your base are belong to us.

Authors

Originally by Shriti Somaiya, edited by Dave Taylor.
Revised in 2007 by J. A. Elliott (j.a.elliott@warwick.ac.uk), with additions by David McCormick (d.s.mccormick@warwick.ac.uk).
Further revisions carried out by Jess Lewton & Guy Barwell (r.g.barwell@warwick.ac.uk) 2012.
Any corrections or improvements should be entered into our feedback form at http://tinyurl.com/WMSGuides (alternatively email revision.guides@warwickmaths.org).
Originally based upon lectures given by Derek Holt at the University of Warwick, 2001 and 2006, checked against subsequent courses.

History

First Edition: 2001.
Second Edition: May 16, 2007.
Current Edition: January 18, 2020.

1 Introduction

Linearity pervades mathematics: *linear algebra* is that branch of mathematics concerned with the study of vectors, vector spaces, linear maps, and systems of linear equations, and is the language with which we talk about linearity. It has extensive applications in the natural sciences and the social sciences, since nonlinear models can often be approximated by linear ones.

Linear algebra originated from the theoretical study of the solutions of sets of simultaneous linear equations. Using techniques from linear algebra, problems about systems of linear equations can be reduced to equivalent problems about matrices. For instance

$$\begin{cases} 2x + y = 1 \\ x - 3y = 2 \end{cases} \quad \text{is equivalent to} \quad \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

1.1 Number Systems and Fields

In order to talk about such problems in as general a setting as possible, we fix a definite starting point and, assuming nothing else, work from there. Our starting point will be number systems, using the term as a vague intuitive idea rather than giving any formal definition. The most used number systems in mathematics are:

$$\underset{\text{natural numbers}}{\mathbb{N}} \subset \underset{\text{integers}}{\mathbb{Z}} \subset \underset{\text{rational numbers}}{\mathbb{Q}} \subset \underset{\text{real numbers}}{\mathbb{R}} \subset \underset{\text{complex numbers}}{\mathbb{C}}$$

A perhaps less well-known example is that of the algebraic numbers $\mathbb{A} \subset \mathbb{C}$, i.e. those numbers which are solutions of polynomials with rational coefficients: $\sqrt{3}, i \in \mathbb{A}$, but $e, \pi \notin \mathbb{A}$.

1.2 Axioms for Number Systems

The term "axiom" has a variety of meanings in mathematics. Sometimes it is taken to mean a self-evident undeniable truth; in other situations it simply refers to anything that is assumed without proof when developing some theory (e.g. linear algebra). Here it is the latter.

Definition 1.1. A number system K is said to be a *field* if it satisfies the following ten axioms:
(A1) $\alpha + \beta = \beta + \alpha$, for all $\alpha, \beta \in K$ (commutativity of addition)
(A2) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$, for all $\alpha, \beta, \gamma \in K$ (associativity of addition)
(A3) There exists $0 \in K$, such that $\alpha + 0 = \alpha$, for all $\alpha \in K$ (existence of zero element)
(A4) For each $\alpha \in K$ there exists $-\alpha \in K$ such that $\alpha + (-\alpha) = 0$ (existence of additive inverses)
(M1) $\alpha\beta = \beta\alpha$, for all $\alpha, \beta \in K$ (commutativity of multiplication)
(M2) $(\alpha\beta)\gamma = \alpha(\beta\gamma)$, for all $\alpha, \beta, \gamma \in K$ (associativity of multiplication)
(M3) There exists $1 \in K$ such that $\alpha 1 = \alpha$, for all $\alpha \in K$ (existence of identity element)
(M4) For each $\alpha \in K \setminus \{0\}$ there exists $\alpha^{-1} \in K$ such that $\alpha\alpha^{-1} = 1$ (existence of multiplicative inverses)
(D) $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$, for all $\alpha, \beta, \gamma \in K$ (distributivity of multiplication over addition)
(ND) $1 \neq 0$ (non-degeneracy[1])

Recall the definition of a group:

Definition 1.2. Let G be a set and let $*$ be a binary operation on G (a map that takes any two elements of G and returns an element of G). We say that the pair $(G, *)$ is a group if
(G0) The set G is closed with respect to the operation $*$, i.e. if $\alpha, \beta \in G$ then $\alpha * \beta \in G$. (Strictly speaking, this is part of the definition of the binary operation, but is often included anyway.)
(G1) $(\alpha * \beta) * \gamma = \alpha * (\beta * \gamma)$, for all $\alpha, \beta, \gamma \in G$ (associativity)
(G2) There exists $1 \in G$ such that $\alpha * 1 = 1 * \alpha = \alpha$, for all $\alpha \in G$ (existence of identity)
(G3) For each $\alpha \in G$ there exists $\alpha^{-1} \in G$ such that $\alpha * \alpha^{-1} = \alpha^{-1} * \alpha = 1$ (existence of inverses)

It is common practice to refer to "the group G" with the operation implicit. G is said to be *abelian* (or *commutative*) if additionally

- $\alpha * \beta = \beta * \alpha$, for all $\alpha, \beta \in G$.

[1] This condition is simply to exclude the trivial set $\{0\}$ from being a field.

Using the idea of a group we can summarise the definition of a field as follows.

Definition 1.3. A number system K is said to be a *field* if:
- K is an abelian group under addition;
- $K \setminus \{0\}$ is an abelian group under multiplication;
- Multiplication on K distributes over addition;
- $1 \neq 0$.

2 Vector Spaces

In applied mathematics vectors are often thought of geometrically, perhaps representing some physical quantity, such as velocity or momentum; in such cases a vector is considered as something with magnitude and direction. In pure mathematics, on the other hand, vectors can be treated entirely algebraically and this is how vectors encountered in linear algebra should be thought of: as mathematical objects that obey certain rules. One advantage of the algebraic approach is that it is just as easy to study vectors in n dimensions as it is to study them in two or three.

Definition 2.1. A *vector space* over a field K is a set V together with two basic operations, known as vector addition and scalar multiplication, such that the following axioms hold:

(V0) The set V is closed under vector addition and scalar multiplication. That is, if $\mathbf{v}, \mathbf{w} \in V$ and $\alpha \in K$ then $\mathbf{v} + \mathbf{w} \in V$ and $\alpha\mathbf{v} \in V$. (As in the definition of a group, this axiom is actually part of the definition of the operations themselves but is included as a reminder.)
(V1) With respect to the operation of vector addition, V is an abelian group.
(V2) $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$, for all $\alpha \in K, \mathbf{v}, \mathbf{w} \in V$
(V3) $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$, for all $\alpha, \beta \in K, \mathbf{v} \in V$
(V4) $(\alpha\beta)\mathbf{v} = \alpha(\beta\mathbf{v})$, for all $\alpha, \beta \in K, \mathbf{v} \in V$
(V5) $1\mathbf{v} = \mathbf{v}$, for all $\mathbf{v} \in V$ (where 1 is the identity scalar in K)

The elements of K are called scalars and the elements of V are called vectors. Often, but not always, Greek letters and boldface letters are used for these, respectively. Note that both K and V have zero elements, and these are distinct. The zero scalar is 0_K (sometimes just written "0") and the zero vector is $\mathbf{0}_V$ (sometimes $\mathbf{0}$).

It is usually not important what field K actually is. Throughout this course it is safe to assume that $K = \mathbb{R}$, but in later courses there are times when it is necessary to have $K = \mathbb{C}$ (e.g. to find the Jordan Canonical Form of a matrix – see MA251 Algebra I: Advanced Linear Algebra).

Using the axioms of vector spaces it is possible to prove some obvious properties of vectors and scalars, such as
- $\alpha\mathbf{0}_V = \mathbf{0}_V$, for all $\alpha \in K$.
- $0_K\mathbf{v} = \mathbf{0}_V$, for all $\mathbf{v} \in V$.
- $-(\alpha\mathbf{v}) = (-\alpha)\mathbf{v} = \alpha(-\mathbf{v})$, for all $\alpha \in K, \mathbf{v} \in V$.

2.1 Subspaces

Another important definition is that of a vector subspace.

Definition 2.2. Let V be a vector space and let $W \subset V$ be non-empty. We say that W is a (*vector* or *linear*) *subspace* of V if W is itself a vector space with respect to the same operations as those on V.

If $W \neq V$ then we say that W is a *proper subspace* of V.

Note that since W is a subset of V most of the properties of V are carried over to W so we only really need to check that W is closed with respect to the relevant operations. This is summed up in the following proposition.

Proposition 2.3. Let V be a vector space over a field K and let $W \subset V$ be non-empty. If for all $\mathbf{v}, \mathbf{w} \in W$ and $\alpha \in K$ we have $\mathbf{v} + \mathbf{w} \in W$ and $\alpha\mathbf{v} \in W$ then W is a subspace of V.

For any given vector space V, the sets V and $\{\mathbf{0}_V\}$ are always automatically subspaces of V, which we refer to as "trivial subspaces". Note that every subspace of V must contain the zero vector $\mathbf{0}_V$.

Proposition 2.4. If W_1 and W_2 are subspaces of a vector space V then $W_1 \cap W_2$ and $W_1 + W_2 = \{\mathbf{w}_1 + \mathbf{w}_2 \mid \mathbf{w}_1 \in W_1, \mathbf{w}_2 \in W_2\}$ are both subspaces of V.

Note that $W_1 + W_2$ is not the same as $W_1 \cup W_2$, which may not even be a subspace. For example, the lines $W_1 = \{(\alpha, 0) \mid \alpha \in \mathbb{R}\}$ and $W_2 = \{(0, \alpha) \mid \alpha \in \mathbb{R}\}$ are both subspaces of $\mathbb{R}^2$, but their union is not a subspace as it is not closed (e.g. $(1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2$).

Definition 2.5. Two subspaces W_1 and W_2 of a vector space V are said to be *complementary* if $W_1 \cap W_2 = \{\mathbf{0}_V\}$ and $W_1 + W_2 = V$. This is equivalent to saying that each vector $\mathbf{v} \in V$ can be written uniquely as $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$ where $\mathbf{w}_1 \in W_1$ and $\mathbf{w}_2 \in W_2$.

2.2 Examples of Vector Spaces

The most obvious example of a vector space is K^n (sometimes written $V_n(K)$), where the vectors are n-tuples of elements of K. That is,

$$K^n = \{(\alpha_1, \alpha_2, \ldots, \alpha_n) \mid \alpha_1, \alpha_2, \ldots, \alpha_n \in K\}.$$

For instance, if $K = \mathbb{R}$ then $\mathbb{R}^n$ is just n-dimensional space (e.g. $\mathbb{R}^2 = \{(\alpha, \beta) \mid \alpha, \beta \in \mathbb{R}\}$ is the set of ordered pairs, representing points in the plane). Vector addition and scalar multiplication are defined in the obvious way.

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n), \qquad \lambda(\alpha_1, \ldots, \alpha_n) = (\lambda\alpha_1, \ldots, \lambda\alpha_n).$$

The zero vector is $\mathbf{0} = (0, 0, \ldots, 0)$.

Examples of non-trivial subspaces of $\mathbb{R}^n$ include lines, planes, etc. up to $(n-1)$-dimensional hyperplanes through the origin. For $n = 3$, for instance, a line through the origin is a set of the form $\{\lambda\mathbf{v} \mid \lambda \in \mathbb{R}\}$ for some direction vector $\mathbf{v}$.

The set of all polynomials with coefficients in K and degree less than or equal to some fixed natural number n is a vector space, $K[x]_{\leq n}$ (sometimes written $P_n(K)$), where vector addition and scalar multiplication are defined as expected. In fact the set of *all* polynomials with coefficients in K (and unlimited degree) is a vector space, $K[x]$. However, the set of all polynomials with coefficients in K and degree *exactly* n is not a vector space as it is not closed under vector addition. For any natural number n, the vector space $K[x]_{\leq n}$ is a subspace of $K[x]$.

As an example from analysis, the set of all real-valued functions on some set $A \subset \mathbb{R}$ is a vector space, with vector addition and scalar multiplication defined by

$$(f + g)(x) = f(x) + g(x) \qquad (\lambda f)(x) = \lambda f(x)$$

The set of continuous real-valued functions defined on A, which we denote $C^0(A)$, is a subspace of this vector space.

3 Linear Independence, Spanning Sets and Bases

An important idea in linear algebra is that of the dimension of a vector space. Geometrically, the dimension can be thought of as the number of different "coordinates" (e.g. $\mathbb{R}^3$ is 3-dimensional as it can be described by an x-, a y- and a z-coordinate). This intuitive interpretation works well for relatively simple vector spaces, such as $\mathbb{R}^n$, but is somewhat less useful for more complicated examples, including spaces of polynomials or functions.

It is possible to define dimension of a vector space in a purely algebraic way using the notion of a "basis". There are some important preliminary definitions.

Definition 3.1. A *linear combination* of a set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a vector space V over a field K is any sum

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n$$

where $\lambda_1, \ldots, \lambda_n$ are scalars (possibly zero) in K.

Definition 3.2. A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a vector space V over a field K are said to be *linearly independent* if none of them is a linear combination of the others. This is the same as saying that

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n = \mathbf{0}_V \implies \lambda_1 = \cdots = \lambda_n = 0_K.$$

If the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are not linearly independent then we say that they are *linearly dependent.*

Lemma 3.3. A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ are linearly dependent if and only if for some $\mathbf{v}_r$ either $\mathbf{v}_r = \mathbf{0}_V$ or $\mathbf{v}_r$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$.

Definition 3.4. A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a vector space V over a field K are said to *span* V if every $\mathbf{v} \in V$ can be written in at least one way as a linear combination of vectors in the set. That is, if for all $\mathbf{v} \in V$ there exist scalars $\lambda_1, \ldots, \lambda_n \in K$, such that

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n.$$

Definition 3.5. A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a vector space V are said to form a *basis* for V if they are linearly independent and span V.

Proposition 3.6. If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis for the vector space V then every $\mathbf{v} \in V$ can be written as a *unique* linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. That is,

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_n \mathbf{v}_n$$

where the scalars $\lambda_1, \ldots, \lambda_n$ are uniquely determined by $\mathbf{v}$.

Theorem 3.7. Any two bases[2] of a vector space contain the same number of vectors. (append the two bases and apply 3.9)

3.1 Dimension

The previous result means that the following is well-defined.

Definition 3.8. The *dimension* of a vector space V is the number of vectors in any basis for V. We write $\dim V$ for the dimension of V. (By convention, $\dim\{\mathbf{0}_V\} = 0$.)

For example, $\dim K^n = n$; any vector can be described uniquely by n coordinates. Any vector space V where $\dim V = n$ for some natural number n is said to be *finite dimensional.* There are also vector spaces with infinite dimension: $K[x]$ has the countably infinite basis

$$1, x, x^2, x^3, \ldots, x^n, \ldots$$

whereas the space of all real-valued functions defined on a set $A \subset \mathbb{R}$ has uncountably infinite dimension. However, this course deals almost exclusively with finite dimensional vector spaces.

Note that a finite dimensional vector space is not necessarily finite. For instance, consider the plane $\mathbb{R}^2$. This has a finite dimension of two, but contains an uncountably infinite number of points (vectors). As long as the field K is infinite then so is the vector space.

Lemma 3.9. Suppose that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w} \in V$ span V and that $\mathbf{w}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span V. In other words, given a spanning set, you can remove any vector that is a linear combination of the others and still have a spanning set; this is called "sifting".

Corollary 3.10. Suppose that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$ span V and that $\dim V = n$ where $r > n$. Then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ contains a proper subset that is a basis for V. That is, any spanning set can be reduced to a basis.

Lemma 3.11. Suppose that V is an n-dimensional vector space and that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$ are linearly independent, where $r < n$. Then there exist vectors $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n \in V$ such that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ forms a basis for V. Thus, any set of linearly independent vectors can be extended to a basis.

[2]The plural of basis is "bases".

Given two subspaces W_1 and W_2 of a vector space V, we can form the subspaces $W_1 + W_2$ and $W_1 \cap W_2$. As any subspace is itself a vector space, we can find the dimension of each subspace. The following theorem tells us how the dimensions of W_1, W_2, $W_1 + W_2$ and $W_1 \cap W_2$ are related.

Theorem 3.12. Suppose that V is a finite-dimensional and W_1, W_2 are two subspaces of V. Then

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

4 Matrices and Linear Maps

4.1 Linear Transformations

Single vector spaces considered in isolation are not very interesting. The main results in linear algebra are concerned with the maps between vector spaces, which are called linear maps (or linear transformations).

Definition 4.1. Let U and V be two vector spaces over the same field K. A *linear map* (or *linear transformation*) is a map $T\colon U \to V$ such that

- $T(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2)$, for all $\mathbf{u}_1, \mathbf{u}_2 \in U$
- $T(\lambda\mathbf{u}) = \lambda T(\mathbf{u})$, for all $\mathbf{u} \in U$ and $\lambda \in K$.

These two conditions can be condensed into one equivalent condition:

$$T(\lambda\mathbf{u}_1 + \mu\mathbf{u}_2) = \lambda T(\mathbf{u}_1) + \mu T(\mathbf{u}_2), \text{ for all } \mathbf{u}_1, \mathbf{u}_2 \in U \text{ and } \lambda, \mu \in K.$$

Proposition 4.2. The following results follow immediately.

- $T(\mathbf{0}_U) = \mathbf{0}_V$.
- $T(-\mathbf{u}) = -T(\mathbf{u})$, for all $\mathbf{u} \in U$.

Linear maps between vector spaces are just one example of structure-preserving maps between algebraic structures. A homomorphism between two groups $(G, *)$ and $(H, \cdot)$ is a map $\phi\colon G \to H$ such that $\phi(g_1 * g_2) = \phi(g_1) \cdot \phi(g_2)$ for every $g_1, g_2 \in G$, i.e. such that in some sense the structure is preserved. A linear map between vector spaces can be thought of as a type of homomorphism.

There are many examples of linear maps between vector spaces. For instance, the embedding $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T\colon (\alpha, \beta) \mapsto (\alpha, \beta, 0)$ is a linear map, as is a rotation about the origin in the plane (i.e. $\mathbb{R}^2$).

However, there are also plenty of examples of maps between vector spaces which are not linear. Consider the translation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ defined by $T\colon (\alpha_1, \alpha_2, \ldots, \alpha_n) \mapsto (\alpha_1 + 1, \alpha_2, \ldots, \alpha_n)$. Since $T(\mathbf{0}) \neq \mathbf{0}$, this cannot be linear.

The following theorem is very important.

Theorem 4.3. A linear map is completely determined by its action on a basis. If two linear maps have the same effect on a basis of the domain then they are the same map.

Now, some more definitions.

Definition 4.4. Let U and V be vector spaces and let $T\colon U \to V$ be a linear map. The *image* of T, written $\operatorname{im} T$, is the set of vectors $\mathbf{v} \in V$ such that $\mathbf{v} = T(\mathbf{u})$ for some $\mathbf{u} \in U$. That is,

$$\operatorname{im} T = \{T(\mathbf{u}) \mid \mathbf{u} \in U\}.$$

Definition 4.5. Let U and V be vector spaces and let $T\colon U \to V$ be a linear map. The *kernel* (or *nullspace*) of T, written $\ker T$, is the set of vectors $\mathbf{u} \in U$ such that $T(\mathbf{u}) = \mathbf{0}_V$. That is,

$$\ker T = \{\mathbf{u} \in U \mid T(\mathbf{u}) = \mathbf{0}_V\}.$$

Proposition 4.6. The kernel and image of a linear map $T\colon U \to V$ are subspaces of U and V, respectively.

Definition 4.7. The *rank* of a linear map $T\colon U \to V$ is the dimension of its image, i.e. $\operatorname{rank} T = \dim(\operatorname{im} T)$.

Definition 4.8. The *nullity* of a linear map $T\colon U \to V$ is the dimension of its kernel, i.e. $\operatorname{nullity} T = \dim(\ker T)$.

The dimensions of the kernel and image of a linear map between vector spaces are closely related. The next theorem tells us how.

Theorem 4.9 (Dimension Theorem)**.** Let U and V be finite-dimensional vector spaces over a field K and let $T\colon U \to V$ be a linear map. Then

$$\operatorname{rank} T + \operatorname{nullity} T = \dim U.$$

Proposition 4.10. Let U and V be vector spaces with $\dim U = \dim V = n$ and let $T\colon U \to V$ be a linear map. Then the following are equivalent.

- T is surjective
- $\operatorname{rank} T = n$
- $\operatorname{nullity} T = 0$
- T is injective
- T is bijective

Definition 4.11. If U and V are vector spaces with $\dim U = \dim V$ a linear map $T\colon U \to V$ is said to be *non-singular* if it is surjective and *singular* if it is not. Equivalently, a map T is singular if $\ker T \neq 0$.

Linear maps can be combined in several ways. Let U, V and W be vector spaces and let $T_1\colon U \to V$, $T_2\colon U \to V$ and $T_3\colon V \to W$ be linear maps. Then the following are also linear maps:

- $T_1 + T_2\colon U \to V$, defined as $(T_1 + T_2)(\mathbf{u}) = T_1(\mathbf{u}) + T_2(\mathbf{u})$, for all $\mathbf{u} \in U$.
- $\lambda T_1\colon U \to V$, defined as $(\lambda T_1)(\mathbf{u}) = \lambda T_1(\mathbf{u})$, for all $\mathbf{u} \in U$ and fixed $\lambda \in K$.
- $T_3 \circ T_2\colon U \to W$, defined as $(T_3 \circ T_2)(\mathbf{u}) = T_3(T_2(\mathbf{u}))$, for all $\mathbf{u} \in U$.

4.2 Matrices

Matrices are combinatorial structures which represent linear transformations. That is, the effect of multiplying a column vector by a matrix gives the same result as applying the corresponding linear transformation to that column vector, and vice versa. For example, the map $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T\colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + y \\ 2x - y \end{pmatrix}$$

is described by the matrix $\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$, since

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ 2x - y \end{pmatrix}$$

The basic matrix operations are straightforward. Addition and scalar multiplication are carried out term by term. Slightly more complicated is the multiplication of two matrices. Note that two matrices can only be multiplied together if the second has the same number of rows as the first has columns.

Using the notation

$$(\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$

the product of $A = (\alpha_{ij})$ and $B = (\beta_{ij})$ is $AB = C$ where $C = (\gamma_{ij})$ and

$$\gamma_{ij} = \sum_{k=1}^{n} \alpha_{ik} \beta_{kj}.$$

Proposition 4.12. The following "laws of matrices" always hold:
- $A + B = B + A$ where A and B are $m \times n$ matrices.
- $(A + B)C = AC + BC$ where A and B are $l \times m$ matrices and C is an $m \times n$ matrix.
- $A(B + C) = AB + AC$ where A is an $l \times m$ matrix and B and C are $m \times n$ matrices.
- $(\lambda A)B = \lambda(AB) = A(\lambda B)$ where A is an $l \times m$ matrix, B is an $m \times n$ matrix and $\lambda \in K$.
- $(AB)C = A(BC)$ where A is an $l \times m$ matrix, B is an $m \times n$ matrix and C is an $n \times p$ matrix.

Note that matrix multiplication is *not* commutative, i.e. in general it is not true that $AB = BA$.

Definition 4.13. The *transpose* of an $m \times n$ matrix A, written A^t, is the $n \times m$ matrix whose (i, j)th entry is the (j, i)th entry of A.

Proposition 4.14. For an $m \times n$ matrix A and a $n \times p$ matrix B, $(AB)^t = B^t A^t$.

Definition 4.15. The *zero matrix* has every entry equal to 0. Sometimes the $m \times n$ zero matrix is written $0_{m \times n}$.

Definition 4.16. The $n \times n$ *identity matrix* is a square $n \times n$ matrix (α_{ij}) where $\alpha_{ii} = 1$ for $1 \leq i \leq n$ and $\alpha_{ij} = 0$ if $i \neq j$. In other words, $\alpha_{ij} = \delta_{ij}$, the Kronecker delta.

Let U and V be vector spaces and let $T\colon U \to V$ be a linear map. For each choice of basis for U and V there is exactly one matrix that represents T. The product of two matrices representing linear transformations corresponds to the composition of the linear transformations, provided the multiplication is carried out in the correct order.

5 Elementary Operations and the Rank of a Matrix

Recall that the rank of a linear map $T\colon U \to V$ is the dimension of its image, i.e. $\operatorname{rank} T = \dim(\operatorname{im} T)$.

5.1 Row and Column Operations

Definition 5.1. The *row rank* of a matrix A is the dimension of the vector space spanned by the vectors that make up its rows.

Definition 5.2. The *column rank* of a matrix A is the dimension of the vector space spanned by the vectors that make up its columns.

Lemma 5.3. If $T\colon U \to V$ is a linear map represented by a matrix A then

$$\operatorname{rank} T = \operatorname{row\,rank} A = \operatorname{column\,rank} A.$$

In order to determine the rank of a matrix it is often easier to row- or column-reduce it. This is done via a series of operations which do not change the rank and results in a matrix whose rank can be read off without any thought.

Let A be an $m \times n$ matrix. There are three *elementary row operations* and three *elementary column operations*.

(R1) Add a multiple of one row to another different row.
(R2) Interchange two different rows.
(R3) Multiply a row by a *non-zero* scalar.

The column operations (C1)–(C3) are defined analogously to the row operations above.

Lemma 5.4. Applying any row operation (R1)–(R3) or any column operation (C1)–(C3) to a matrix A does not change the rank of A.

All row and column operations can be represented by *elementary matrices* whose effect on another matrix is then to apply that row or column operation. To obtain the matrix of a row or column operation simply apply the row or column operation to the identity matrix, giving the elementary matrix required. The effect of applying the same row or column operation to a matrix is then found by premultiplying that matrix by the elementary matrix just found. Make sure you know how to express a matrix in terms of elementary matrices.

By repeated application of the elementary row operations a matrix can be reduced to *upper echelon form*, where the leftmost non-zero element of each row is equal to 1:

$$\begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 & \cdot \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The rank of the original matrix is then equal to the number of non-zero rows in upper echelon form.

It is possible to further reduce a matrix in upper echelon form to *row-reduced echelon form* where the leftmost non-zero entry in each row is a 1 and is the only non-zero entry in that column, for example:

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The above matrix has a rank equal to 3.

By applying column operations, we can further reduce a matrix in row-reduced echelon form to *Smith normal form*, which has the first s entries on the leading diagonal equal to 1 and all other entries zero, for example:

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

Proposition 5.5. Let s be the number of non-zero rows in the Smith normal form of a matrix A. Then both the row rank of A and the column rank of A are equal to s.

5.2 Linear Equations and Row Operations

Given a system of linear equations such as:

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n &= \beta_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n &= \beta_2 \\ &\vdots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \cdots + \alpha_{nn}x_n &= \beta_n \end{aligned}$$

we can set $A = (\alpha_{ij})$ and write $A\mathbf{x} = \mathbf{b}$, where $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and $\mathbf{b} = (\beta_1, \beta_2, \ldots, \beta_n)^T$. By forming the *augmented matrix*

$$A' = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & \beta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} & \beta_n \end{array}\right).$$

we can apply elementary row operations to this matrix to reduce the system of equations to a simpler form and hence find the solution. Note that we cannot apply elementary column operations since they change the system.

6 Determinants and Inverses

When does a linear map between two vector spaces have an inverse? That is, given a linear map $T\colon U \to V$, what conditions do we require for there to be a map $T^{-1}\colon V \to U$ such that $TT^{-1} = I_V$ and $T^{-1}T = I_U$? Essentially, to be invertible, it must be injective and surjective, which means that it must be non-singular.

Theorem 6.1. A linear map T is invertible if and only if T is non-singular. In particular, if T is invertible then $m = n$, so only square matrices can be invertible.

We can calculate the inverse of a matrix by row-reduction, and the row-reduced form of any invertible $n \times n$ matrix is I_n. However, this is tedious to compute. *Determinants* exist to allow us to quickly check whether a given matrix (or the linear map which it represents) is invertible. They are only defined for square matrices, since a non-square matrix will necessarily either have non-zero nullity (and hence not be surjective) or its transpose will; in both these cases the matrix clearly cannot have an inverse.

Geometrically speaking the determinant of a linear map or the matrix representing a linear map is related to the scaling factor of the map. The determinant is equal to the area of the image of a unit square under that transformation. If the determinant is zero then some unit square is mapped to a region of zero area, and therefore some non-zero regions will be shrunk to lines. Such a map is certainly non-invertible.

Theorem 6.2. Let A be a matrix. A is invertible if and only if $\det A \neq 0$.

In order to define the determinant of an $n \times n$ matrix we need to recall a few facts about permutations.

Definition 6.3. A *permutation* on the set $X_n = \{1, \ldots, n\}$, where n is a natural number, is a bijection $\phi\colon X_n \to X_n$.

Definition 6.4. The *symmetric group* of order n is the group S_n of permutations from the set X_n where n is a natural number, where the product of two permutations is their composition.

Definition 6.5. A *transposition* is a permutation of X_n that interchanges two numbers i and j and leaves all other numbers fixed; it is written as (i, j).

Definition 6.6. A permutation ϕ is *even*, with $\operatorname{sign}(\phi) = +1$, if ϕ is a composite of an even number of transpositions; ϕ is *odd*, with $\operatorname{sign}(\phi) = -1$, if ϕ is a composite of an odd number of transpositions.

Definition 6.7. Let A be an $n \times n$ matrix over a field K. We define the determinant of A as follows:

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \alpha_{1\sigma(1)} \alpha_{2\sigma(2)} \cdots \alpha_{n\sigma(n)}$$

This definition can look rather complicated so the following two examples show how the determinants of 2×2 and 3×3 matrices are calculated in practice.

Examples 6.8.

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}$$

$$\begin{aligned} \begin{vmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{vmatrix} &= \beta_{11} \begin{vmatrix} \beta_{22} & \beta_{23} \\ \beta_{32} & \beta_{33} \end{vmatrix} - \beta_{12} \begin{vmatrix} \beta_{21} & \beta_{23} \\ \beta_{31} & \beta_{33} \end{vmatrix} + \beta_{13} \begin{vmatrix} \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{vmatrix} \\ &= \beta_{11}(\beta_{22}\beta_{33} - \beta_{23}\beta_{32}) - \beta_{12}(\beta_{21}\beta_{33} - \beta_{23}\beta_{31}) + \beta_{13}(\beta_{21}\beta_{32} - \beta_{22}\beta_{31}). \end{aligned}$$

We now consider the effect of row and column operations on the determinant. First note that $\det(I_n) = 1$.

Proposition 6.9. Let A be an $n \times n$ matrix.
- Applying (R1) to A (i.e. adding a multiple of one row to another) leaves $\det A$ unchanged.
- Applying (R2) to A (i.e. interchanging two rows) multiplies $\det A$ by -1.
- Applying (R3) to A (i.e. multiplying a row by a scalar λ) multiplies $\det A$ by λ.

From this we can compute the determinant of any elementary matrix. We also obtain the following result for the determinant of a product.

Proposition 6.10. For any two $n \times n$ matrices A and B, we have $\det(AB) = \det A \det B$.

6.1 Minors and Cofactors

Definition 6.11. Let A be an $n \times n$ matrix. Let A_{ij} be the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i^{th} row and j^{th} column of A, and let $M_{ij} = \det(A_{ij})$. Then M_{ij} is called the (i,j)th *minor* of A.

Definition 6.12. Let A be an $n \times n$ matrix. Define the (i,j)th *cofactor* of A by $c_{ij} = (-1)^{i+j} M_{ij}$.

Theorem 6.13. For any i with $1 \leq i \leq n$, we have $\det A = \sum_{j=1}^{n} \alpha_{ij} c_{ij}$. Similarly, for any j with $1 \leq j \leq n$, we have $\det A = \sum_{i=1}^{n} \alpha_{ij} c_{ij}$.

Definition 6.14. Let A be an $n \times n$ matrix. Define the *adjoint matrix* $\operatorname{adj}(A)$ of A as the $n \times n$ matrix whose (i,j)th element is the cofactor c_{ji}, i.e. $\operatorname{adj}(A)$ is the transpose of the matrix of cofactors.

Theorem 6.15. $A \operatorname{adj}(A) = \det(A) I_n = \operatorname{adj}(A) A$. Furthermore, if $\det(A) \neq 0$ then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.

7 Change of Basis and Similar Matrices

7.1 Bases

Let $T\colon U \to V$ be a linear map between U and V. To express T as a matrix requires picking a basis $\{\mathbf{e}_i\}$ of U and a basis $\{\mathbf{f}_j\}$ of V. To change between two bases $\{\mathbf{e}_i\}$ and $\{\mathbf{e}'_i\}$ of U, we simply take the identity map $I_U\colon U \to U$ and use the basis $\{\mathbf{e}_i\}$ in the domain and the basis $\{\mathbf{e}'_i\}$ in the codomain; the matrix so formed is the *change of basis matrix* from the basis of $\mathbf{e}_i$s to the $\mathbf{e}'_i$s. Any such change of basis matrix is invertible.

Theorem 7.1. Let A be the matrix of $T\colon U \to V$ with respect to the bases $\{\mathbf{e}_i\}$ of U and a basis $\{\mathbf{f}_j\}$ of V, and let B be the matrix of T with respect to the bases $\{\mathbf{e}'_i\}$ of U and a basis $\{\mathbf{f}'_j\}$ of V. Let P be the change of basis matrix from $\{\mathbf{e}_i\}$ to $\{\mathbf{e}'_i\}$, and let Q be the change of basis matrix from $\{\mathbf{f}_i\}$ to $\{\mathbf{f}'_i\}$. Then $B = QAP^{-1}$.

The previous result means that there exist invertible matrices P and Q such that $B = QAP$ (replacing P with P^{-1}).

Definition 7.2. Two $m \times n$ matrices A and B are *equivalent* if there exist invertible matrices P and Q such that $B = QAP$.

Theorem 7.3. Let A and B be $m \times n$ matrices over K. Then the following are equivalent:
- A and B are equivalent.
- A and B represent the same linear map with respect to different bases.
- A and B have the same rank.
- B can be obtained from A by application of elementary row and column operations.

7.2 Similar Matrices

We now consider the special case where $U = V$.

Definition 7.4. Two $n \times n$ matrices A and B over a field K are said to be *similar* if there exists an invertible $n \times n$ matrix P, such that $B = P^{-1}AP$.

Similar matrices are equivalent, but equivalent matrices are not necessarily similar. Not every $n \times n$ matrix is similar to a diagonal matrix[3].

Definition 7.5. Let V be a vector space and let $T\colon V \to V$ be a linear map. An *eigenvector* of T is a non-zero vector $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \lambda \mathbf{v}$ for some $\lambda \in K$.

Definition 7.6. An *eigenvalue* of T is a scalar $\lambda \in K$ such that $T(\mathbf{v}) = \lambda \mathbf{v}$ for some eigenvector $\mathbf{v}$ of T.

[3] An example of this is $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$

Note that because of the correspondence between results about linear maps and matrices that saying $T(\mathbf{v}) = \lambda\mathbf{v}$ is exactly the same as saying $A\mathbf{v} = \lambda\mathbf{v}$, where A is the matrix representing the transformation T.

For an eigenvector $\mathbf{v}$ of a matrix A we have

$$\begin{aligned} A\mathbf{v} &= \lambda\mathbf{v} \\ \implies (A - \lambda I_n)\mathbf{v} &= \mathbf{0}_V \\ \implies \det(A - \lambda I_n) &= 0 \end{aligned}$$

since eigenvectors are non-zero by definition, meaning that the matrix $A - \lambda I_n$ must be singular. Hence, in order to find the eigenvalues of a matrix A we need to solve the following equation

$$\det(A - \lambda I_n) = 0$$

where A is the matrix representing a linear map $T\colon V \to V$ with $\dim V = n$.

Definition 7.7. For an $n \times n$ matrix A, the equation $\det(A - xI_n) = 0$ is called the *characteristic equation* of A, and $\det(A - xI_n)$ is called the *characteristic polynomial* of A. It is a polynomial of degree n in x.

Theorem 7.8. Let A be an $n \times n$ matrix. Then λ is an eigenvalue of A if and only if $\det(A - \lambda I_n) = 0$. That is, the eigenvalues of A are the roots of its characteristic polynomial.

Theorem 7.9. Similar matrices have the same characteristic equation and hence the same eigenvalues.

This is because similar matrices represent the same linear map with respect to different bases. The basis which we use to describe a map does not affect the map itself, so the same map described in two different ways must have the same eigenvalues.

Theorem 7.10. Let V be a vector space and let $T\colon V \to V$ be a linear map. Then the matrix of T is diagonal with respect to some basis of V (or equivalently, any matrix A of T is similar to a diagonal matrix) if and only if V has a basis consisting of eigenvectors of T.

Proposition 7.11. The eigenvalues of an upper triangular matrix are just the entries along the diagonal.

In particular, the determinant of a matrix that has been diagonalised is just the product of the eigenvalues of the matrix.

Theorem 7.12. Let $\lambda_1, \dots, \lambda_r$ be distinct eigenvalues of a linear map $T\colon V \to V$ and let $\mathbf{v}_1, \dots, \mathbf{v}_r$ be the corresponding eigenvectors. Then $\mathbf{v}_1, \dots, \mathbf{v}_r$ are linearly independent.

Corollary 7.13. If the linear map $T\colon V \to V$ has n distinct eigenvalues, where $\dim V = n$, then T is diagonalisable.

An analogous statement holds for the corresponding matrix A. It is important to be aware that the converse of this statement is not true. There are many matrices that can be diagonalised even though they may have repeated eigenvalues.

Closing Remarks

As you can see, there's a reasonable amount of material in Linear Algebra, but it's easy to get confused with all the abstract definitions. Being able to accurately state definitions and key theorems is important (and easy marks in the exam), so learn them well. While a smattering of short proofs might be examined, it's more likely that you'll be asked questions about simple examples to test your understanding, so know how to apply the key theorems as well as what they state. Above all though, practice is the best medicine, especially through past exam questions; the format isn't likely to change all that much compared to previous years. So practise, practise, practise, and good luck in the exam!

This guide would not be possible without our wonderful sponsors: