# Prerequisites from Linear Algebra

Mark Wildon

October 17, 2006

It is asssumed that the reader is familiar with the linear and bilinear algebra from the second year core algebra course. The main 'extra' that is needed is the idea of a quotient vector space; this will be familiar to those who have done B2a algebra, but maybe not to others, so I summarise it below. It will also be useful to know the statement of Jordan normal form. We really only need one result from bilinear algebra, with which I think everyone will be familiar, but it's repeated below just to make sure.

Highly recommended for alternating or additional reading is Halmos' book, *Finite-Dimensional Vector Spaces*.

## 1 Quotient Spaces and isomorphism theorems

Suppose that $W$ is a subspace of the vector space $V$. A *coset of* $W$ is a set of the form

$$v + W = \{v + w : w \in W\}.$$

It is important to realise that unless $W = 0$, each coset will have many different labels; in fact, $v + W = v' + W$ if and only if $v - v' \in W$.

The *quotient space* $V/W$ is the set of all cosets of $W$. This becomes a vector space, with zero element $0 + W = W$, if addition is defined by

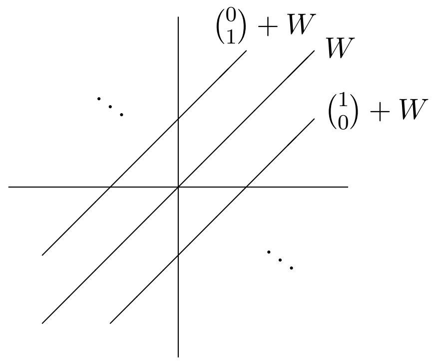
$$(v + W) + (v' + W) = (v + v') + W \quad \text{for } v, v' \in V$$

and scalar multiplication by

$$\lambda(v + W) = \lambda v + W \quad \text{for } v, v' \in V,\ \lambda \in F.$$

One must check that these operations are *well-defined*; that is, they do not depend on the choice of labelling elements. Suppose for instance that $v + W = v' + W$. Then, since $v - v' \in W$, we have $\lambda v - \lambda v' \in W$ for any scalar $\lambda$, so $\lambda v + W = \lambda v' + W$.

The following diagram shows the elements of $\mathbb{R}^2/W$, where $W$ is the subspace of $\mathbb{R}^2$ spanned by $\binom{1}{1}$.




The cosets $\mathbb{R}^2/W$ are all the translations of the line $W$. One can choose a standard set of coset representatives by picking any line through 0 (other than $W$) and looking at its intersection points with the cosets of $W$; this gives a geometric interpretation of the isomorphism $\mathbb{R}^2/W \cong \mathbb{R}$.

It is often useful to consider quotient spaces when attempting a proof by induction on the dimension of a vector space. In this context, it can be useful to know that if $v_1, \ldots, v_k$ are vectors in $V$ such that the cosets $v_1 + W, \ldots, v_k + W$ form a basis for the quotient space $V/W$, then $v_1, \ldots, v_k$, together with any basis for $W$, forms a basis for $V$.

We can now state the isomorphism theorems for vector spaces.

**Theorem 1.1.** (a) *Let $V$ and $W$ be vector spaces and let $x : V \to W$ be a linear map. Then* $\ker x$ *is a subspace of $V$,* $\operatorname{im} x$ *is a subspace of $W$, and*

$$V/\ker x \cong W.$$

*Now let $U$ and $W$ be subspaces of $V$.* (b) $(U+W)/W \cong U/(U \cap W)$. (c) *The quotient space $W/U$ is a subspace of $V/U$ and* $(V/U)/(W/U) \cong V/W$.

*Proof.* (a) Define a map $\phi : V/\ker x \to \operatorname{im} x$ by

$$\phi(v + \ker x) = x(v).$$

This map is well-defined since if $v + \ker x = v' + \ker x$ then $v - v' \in \ker x$, so $\phi(v + \ker x) = x(v) = x(v') = \phi(v' + \ker x)$. It is routine to check that $\phi$ is linear, injective, and surjective, so it gives the required isomorphism.

For (b) consider the composite of the inclusion map $U \to U + W$ with the quotient map $U + W \to (U + W)/W$. This gives us a linear map $U \to (U + W)/W$. Under this map, $x \in U$ is sent to $0 \in (U + W)/W$ if and only if $x \in W$, so its kernel is $U \cap W$. Now apply part (a). Part (c) can be proved similarly; we leave this to the reader. $\square$

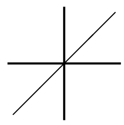
## 2 Interlude: The Diagonal Fallacy

Consider the following (fallacious) argument. Let $V$ be a 2-dimensional vector space, say with basis $v_1$, $v_2$. Let $x : V \to V$ be the linear map whose matrix with respect to this basis is

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

We claim that if $U$ is an $x$-invariant subspace of $V$; that is, $x(U) \subseteq U$, then either $U = 0$, $U = \mathrm{Span}\{v_1\}$, or $U = V$. Clearly each of these subspaces is invariant under $x$, so we only need to prove that there are no others. But since $x(v_2) = v_1$, $\mathrm{Span}\{v_2\}$ is not $x$-invariant. (QED?)

Here we committed the *diagonal fallacy*: We assumed that an arbitrary subspace of $V$ would contain one of our chosen basis vectors. This assumption is very tempting — which perhaps explains why it is so often made[1] — but it is nonetheless totally unjustified. I suspect one reason why people end up committing this error is that they get confused with a *good* way to use linearity, namely the fact that one can save time and space by only defining a linear map on elements of a given basis.



## 3 Jordan Normal Form

Let $V$ be a finite-dimensional complex vector space and let $x : V \to V$ be a linear map. The exercise below outlines a proof that one can always find a basis of $V$ in which $x$ is represented by an upper triangular matrix. For many purposes, this result is sufficient. For example, since the eigenvalues of a matrix in upper triangular form are its diagonal entries, it implies that a nilpotent map may be represented by a strictly upper triangular matrix, and so nilpotent maps have trace 0.

**Exercise 3.1.** *Let $V$ be an $n$-dimensional vector space where $n \geq 1$, and let $x : V \to V$ be a linear map.*

(i) *Show that $x$ has an eigenvector, $v$ say.*

(ii) *Let $U = \mathrm{Span}\{v\}$. Show that $x$ induces a linear transformation $\bar{x} : V/U \to V/U$. By induction, we know that there exists a basis $\{v_1 + U \dots v_{n-1} + U\}$ of $V/U$ in which $\bar{x}$ has a upper triangular matrix. Prove that $\{v, v_1, \dots, v_{n-1}\}$ is a basis of $V$ and that the matrix of $x$ in this basis is upper triangular.*

[1] The author's personal record is hearing the diagonal fallacy committed in three tutorials in a row, on two different courses. After reading Kafka's *The Penal Colony* an unpalatable but probably highly successful to the problem occured to him. (See marginal diagram.)

Sometimes, however, one needs the full strength of Jordan normal form. A general matrix in Jordan normal form looks like

$$\begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_r \end{pmatrix},$$

where each $A_i$ is a *Jordan block matrix* $J_t(\lambda)$ for some $t \in \mathbb{N}$ and $\lambda \in \mathbb{C}$:

$$J_t(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}_{t \times t.}$$

Any linear transformation of a complex vector space can be represented by a matrix in Jordan normal form. One can successfully use Jordan normal form without knowing anything about how to prove this; that said, if you really want a proof you might see `www.maths.ox.ac.uk/~wildon/JNF.pdf`.

# 4 Bilinear Algebra

**Definition 4.1.** *A* bilinear form *on a vector space $V$ is a map*

$$(-,-) : V \times V \to F$$

*such that*

$$\begin{aligned} (\lambda_1 v_1 + \lambda_2 v_2, w) &= \lambda_1 (v_1, w) + \lambda_2 (v_2, w), \\ (v, \mu_1 w_1 + \mu_2 w_2) &= \mu_1 (v, w_1) + \mu_2 (v, w_2), \end{aligned}$$

*for all $v_i, w_i \in V$ and $\lambda_i, \mu_i \in F$.*

For example, if $F = \mathbb{R}$ and $V = \mathbb{R}^n$, then the usual dot product is a bilinear form on $V$.

Given a subset $U$ of a vector space $V$, we set

$$U^\perp = \{v \in V : (u, v) = 0 \text{ for all } u \in U\}.$$

This is always a subspace of $V$. We say that the form $(-,-)$ is *non-degenerate* if $V^\perp = \{0\}$.

There is an important connection between bilinear forms and dual spaces. Let $\varphi : V \to V^\star$ be the linear map defined by $\varphi(v) = (-, v)$. That is, $\varphi(v)$ is the linear map sending $u \in V$ to $(u, v)$. If $(-,-)$ is non-degenerate, then $\ker \varphi = 0$, so by dimension counting, $\varphi$ is an isomorphism from $V$ to $V^\star$. Hence every element of $V^\star$ is of the form $(-, v)$ for a unique $v \in V$; this is a special case of the *Riesz representation theorem.*