A CRASH-COURSE ON

GROUPS & VECTOR SPACES

Luke Collins

maths.com.mt/notes

1 Groups

These notes are a brief summary of the main results on groups. The results on number theory (residue groups, Euler's theorem, Fermat's little theorem, etc.) as well as those on lattices (posets, Hasse diagrams) are omitted.

1.1 Basics

Recall that a nonempty set G together with a binary operation $\cdot : G \times G \to G$ is said to form a *group* $(G, \cdot)$ if the following properties are satisfied:

Closure. For all $x, y \in G$, $x \cdot y \in G$.

Associativity. For all $x, y, z \in G$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

Identity. There exists $1 \in G$ such that for all $x \in G$, $1 \cdot x = x \cdot 1 = x$.

Inverses. For all $x \in G$, there exists $x^{-1} \in G$ such that $x \cdot x^{-1} = x^{-1} \cdot x = 1$.

Throughout the notes, we shorten the notation $x \cdot y$ to xy, and we relax the ordered pair notation $(G, \cdot)$, referring to the corresponding structure simply as "the group G".

Certain groups satisfy the following additional property:

Commutativity. For all $x, y \in G$, $x \cdot y = y \cdot x$.

Such groups are called *abelian*.

Examples 1.1. The following are some examples of groups.

(i) The set of integers $\mathbb{Z}$ is a group under the usual addition operation $+$, Moreover, this group is abelian. Similarly, the rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$ are abelian groups under addition.

(ii) The set of integers $\mathbb{Z}$ is *not* a group under the usual multiplication operation $\times$, since only 1 has an inverse. Neither are $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$ under this $\times$, since 0

has no inverse. If we remove zero, we get that $\mathbb{Q} \setminus \{0\}$, $\mathbb{R} \setminus \{0\}$ and $\mathbb{C} \setminus \{0\}$ are abelian groups under $\times$.

(iii) The finite set $\mathbb{Z}_n = \{0, \dots, n-1\}$ is an abelian group under $+$ modulo n, and $\mathbb{Z}_p \setminus \{0\}$ is an abelian group under $\times$ modulo n, when p is prime. When n is not prime,

- Groups which are *abelian*, that is, groups which have the additional property that all of their elements commute: for all $g_1, g_2 \in G$, $g_1 \cdot g_2 = g_2 \cdot g_1$.
- The *cyclic group* $C_n = \{e, x, x^2, \dots, x^{n-1}\}$ with operation

$$(x^a, x^b) \mapsto x^{a+b \bmod n},$$

  which is an abelian group generated by one element x, with the property that $x^n = e$ (we write $C_n = \langle x : x^n = e \rangle$). This group is isomorphic to the groups $(\mathbb{Z}_n, +_{\bmod n})$ and $\mathbb{Z}/n\mathbb{Z}$.
- The *infinite cyclic group* $C = \{e, x^{\pm 1}, x^{\pm 2}, \dots\}$ is the group generated by the elements $\{x, x^{-1}\}$, and $x^n \neq e$ for all $n \in \mathbb{N}$. We write $C = \langle x, x^{-1} \rangle$. This group is isomorphic to $(\mathbb{Z}, +)$.
- The *dihedral group* $D_n = \langle r, t : r^n = t^2 = e \wedge rt = tr^{-1} \rangle$, which represents the symmetries (rotations r and reflections t) of a regular polygon on n vertices. In general, D_n has $2n$ elements, and is non-abelian for $n \geqslant 3$.
- The *symmetric group* S_n, which is the set of permutations on n elements, or equivalently, the set of bijections from a set X with $|X| = n$ onto itself. In general, $|S_n| = n!$.
- The *alternating group* A_n, which is the set of *even permutations* on n elements. An even permutation is a permutation made up of an even number of transpositions, i.e. it can be written as a product of an even number of 2-cycles. A_n is a subgroup of S_n. In particular, $|A_n| = |S_n|/2 = n!/2$.

1.2 Lagrange's Theorem and Normal Subgroups

Suppose G is a group, and $H \subseteq G$. Then H is said to be a *subgroup* of G, written $H \leqslant G$, if it forms a group in its own right. The simplest way to show that a subset $H \subseteq G$ is a subgroup is called the one-step test:

Theorem 1.2 (One-step test for subgroups)**.** *Let G be a group, and let H be a subset of G. Then $H \leqslant G$ if and only if for all $x, y \in H$, $xy^{-1} \in H$.*

Proof. If H is a subgroup, then $y^{-1} \in H$ for any $y \in H$. Moreover, for any $x \in H$, $xy^{-1} \in H$ by closure. This proves the 'only if' part. For the converse, associativity is hereditary. If we take an arbitrary $h \in H$, then we have $e = h \cdot h^{-1} \in H$, so H contains the identity. Consequently for any $y \in H$, $y^{-1} = ey^{-1} \in H$, so

H has inverses. Finally, for any $x, y \in H$, we have $y^{-1} \in H$, and consequently $xy = x(y^{-1})^{-1} \in H$. This proves closure. □

Every element of a finite group G satisfies $g^n = e$ for some $n \in \mathbb{N}$ (*Proof*: otherwise the subgroup $\langle g \rangle$ is infinite). The smallest such n is called the *order of* g, denoted $|g|$ (or $o(g)$ by Herstein and other authors).

Definition 1.3 (Coset)**.** If G is a group, $H \leqslant G$ and $g \in G$, then:

(i) the *left coset* of H with respect to g is the set $g \cdot H = \{g \cdot h : h \in H\}$, and

(ii) the *right coset* of H with respect to g is the set $H \cdot g = \{h \cdot g : h \in H\}$.

A *coset* is some left or right coset of a subgroup in G. We often relax the notation $g \cdot H$ to gH, and similarly $H \cdot g$ to Hg for right cosets.

Let us define the relation $g_1 \sim g_2$ on G for some $H \leqslant G$ by $g_1 \in g_2 H$. One can easily show that this is an equivalence relation, and its equivalence classes are the distinct left cosets of H. Thus by the properties of equivalence relations, we get that:

- Any two left cosets are either identical or disjoint, and
- the distinct left cosets of H in G give a partition of G.

One may similarly define an equivalence relation for right cosets, and the same results hold when we substitute 'left' for 'right'. Other important facts about cosets are given in exercise 1.4.

Exercise 1.4 (Cosets and Lagrange's theorem)**.**

1. Show that all cosets (left or right) of a subgroup $H \leqslant G$ have the same size as the set H.

   [Hint: Show that for all $g \in G$, the function $\phi_g \colon H \to gH$ defined by $\phi_g \colon h \mapsto gh$ is a bijection. Define a similar bijection for right cosets.]

2. (**Lagrange's Theorem**) Prove that given a group G and $H \leqslant G$, then $|H|$ divides $|G|$.

   [Hint: Use the result of question 1 above, and the fact that distinct left cosets of H in G partition G.]

3. Let G be a finite group. Show that:

   (i) For all $g \in G$, the order $|g|$ of g divides $|G|$.
   [Hint: Consider the subgroup $\langle g \rangle \leqslant G$.]

   (ii) If $|G|$ is prime, then G must be cyclic.

   (iii) For any $g \in G$, $g^{|G|} = e$.

   (iv) If $g \in G$ and $g^n = e$, then $|g|$ divides n.

(v) For any $g \in G$, $|g| = |g^{-1}|$.

4. Show that if $H, K \leqslant G$, then $H \cap K \leqslant G$.

5. Show that there are only two groups of order 4.

In Lagrange's theorem, the positive integer $|G|/|H|$ is called the *index of H in G*, denoted $[G : H]$ or $i_G(H)$.

Note that in general, the set of left cosets and the set of right cosets are not the same.[1] Any two left cosets or any two right cosets are either equal or disjoint, but this is not necessarily true for *any* two cosets (i.e. we cannot say that a left coset and a right coset are either equal or disjoint).

Definition 1.5 (Quotient)**.** Let A be a set, and let $\sim \subseteq A \times A$ be an equivalence relation on A. Then the *quotient of A by $\sim$*, denoted by $A/\sim$ or $\frac{A}{\sim}$, is the set of equivalence classes of $\sim$.

For the equivalence class we defined previously, the quotient $G/\sim$ is the set of distinct left cosets of H. We will denote this set by G/H instead of $G/\sim$, and call it the *quotient of G by H*.

Definition 1.6 (Normal Subgroup)**.** Let G be a group and let $N \leqslant G$. Then N is said to be a *normal subgroup of G*, denoted $N \trianglelefteqslant G$, if $Ng = gN$ for any $g \in G$.

Observe that if G is abelian, every subgroup H is normal, since

$$gH = \{gh : h \in H\} = \{hg : h \in H\} = Hg.$$

The converse is not necessarily true however; that is, a normal subgroup is not necessarily abelian. Indeed, G is a normal subgroup of G for any G. This does not make G abelian!

Normal subgroups are nice because the quotient G/H is independent of whether we work with left cosets or right cosets.

We have the following theorems about normal subgroups.

- If $N \trianglelefteqslant G$, then the quotient G/N forms a group under the product

$$\cdot : G/N \times G/N \to G/N,$$

  defined by $aN \cdot bN = abN$.

- Let G be a group, and let $N \leqslant G$. The following are equivalent:

[1] For example, take the symmetries of the triangle, $D_3 = \{e, r, r^2, t, tr, tr^2\}$, where $r^3 = t^2 = e$ and $rt = tr^{-1}$. Then $H = \langle t \rangle = \{e, t\}$ is a subgroup, and its left cosets are $rH = \{r, rt\} = \{r, tr^{-1}\} = \{r, tr^2\}$ and $r^2H = \{r^2, r^2t\} = \{r^2, rtr^{-1}\} = \{r^2, tr^{-1}r^{-1}\} = \{r^2, tr\}$, but its right cosets are $Hr = \{r, tr\}$ and $Hr^2 = \{r, tr^2\}$. Thus $\{H, rH, r^2H\} \neq \{H, Hr, Hr^2\}$.

(i) $N \trianglelefteq G$,

(ii) For all $g \in G$, $gN = Ng$,

(iii) For all $g \in G$, $g^{-1}Ng \subseteq N$,

(iv) For all $g \in G$, $g^{-1}Ng = N$,

(v) For all $g \in G$, there exists $g' \in G$ such that $gN = Ng'$,

(vi) For all $a, b \in G$, $aNbN = abN$, where $aNbN = \{an_1bn_2 : n_1, n_2 \in N\}$.

Example 1.7. Consider the integers under addition $G = (\mathbb{Z}, +)$. As a subgroup, consider the multiples of 5, $H = 5\mathbb{Z}$. Since the group G is abelian, it follows immediately that any subgroup is normal. Now the quotient $G/H = \mathbb{Z}/5\mathbb{Z}$ is the set

$$\{\{0, \pm5, \pm10, \dots\}, \{\pm1, \pm6, \dots\}, \dots \{\pm4, \pm9, \dots\}\} = \{5\mathbb{Z}, 1+5\mathbb{Z}, \dots, 4+5\mathbb{Z}\},$$

and $(3 + 5\mathbb{Z}) + (4 + 5\mathbb{Z}) = (3 + 4) + 5\mathbb{Z} = 2 + 5\mathbb{Z}$, for example.

1.3 Homomorphisms and Isomorphisms

Let $(G, \cdot)$ and $(H, *)$ be two groups with products $\cdot$ and $*$ respectively. A homomorphism is a function $\phi : G \to H$ such that for all $g_1, g_2 \in G$,

$$\phi(g_1 \cdot g_2) = \phi(g_1) * \phi(g_2).$$

A homomorphism *preserves the structure of multiplication* of the group G in the group H.

Examples 1.8. Let $G = \mathbb{Z}$ under addition, and $H = \{1, -1\}$ under multiplication. Define the function ϕ by $\phi(n) = (-1)^n$. This defines a homomorphism:

$$\phi(m + n) = (-1)^{m+n} = (-1)^m \times (-1)^n = \phi(m) \times \phi(n).$$

Another example, inspired somewhat by example 1.7, is the following. If we map each $a \in \mathbb{Z}$ to $a \bmod 5$ in $H = \{0, 1, 2, 3, 4\}$ under $+_{\bmod 5}$, we get a homomorphism:

$$\phi(a + b) = (a + b) \bmod 5 = (a \bmod 5) +_{\bmod 5} (b \bmod 5) = \phi(a) +_{\bmod 5} \phi(b).$$

Definition 1.9 (Isomorphism)**.** Let G, H be two groups, and let $\phi\colon G \to H$ be a homomorphism. Then ϕ is said to be an *isomorphism* if it is injective.

If $\phi(G) = H$, then G and H are said to be *isomorphic*, written $G \simeq H$.

Definition 1.10 (Kernel)**.** Let $\phi : G \to H$ be a homomorphism, and suppose $\tilde{e} \in H$ is the identity in H. The *kernel of* ϕ, denoted $\ker \phi$, is the following subset of G,

$$\ker \phi = \{g \in G : \phi(g) = \tilde{e}\},$$

i.e. all members of G mapped to the identity element of H.

Exercise 1.11 (Important properties of homomorphisms)**.** Suppose $(G, \cdot)$ and $(H, *)$ are groups, and let $\phi\colon G \to H$ be a homomorphism. Show that:

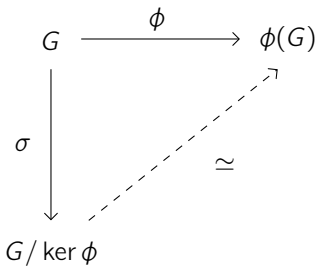
1. If $e \in G$ and $\tilde{e} \in H$ are the identities, then $\phi(e) = \tilde{e}$.
2. For all $g \in G$, $\phi(g^{-1}) = [\phi(g)]^{-1}$.
3. $\phi(G) \leqslant H$.
4. $\ker\phi \trianglelefteq G$.
5. ϕ is injective if and only if $\ker\phi = \{e\}$.

Now let G be a group, and suppose we have a normal subgroup $N \trianglelefteq G$ (think of $5\mathbb{Z} \trianglelefteq \mathbb{Z}$). The *natural homomorphism*, usually denoted σ, is the homomorphism $\sigma\colon G \to G/N$, defined by $\sigma\colon g \mapsto Ng$. Furthermore, its kernel consists of the elements of N, that is, $\ker\phi = N$. This homomorphism has an important role in the so-called *first isomorphism theorem*:

Theorem 1.12 (First Isomorphism Theorem)**.** *Let $\phi\colon G \to H$ be a homomorphism from the group G to the group H. Then the image of G under ϕ is isomorphic to $G/\ker\phi$, i.e.*

$$\phi(G) \simeq \frac{G}{\ker\phi}.$$

This theorem is often visualised with the following commutative diagram:



Example 1.13. Again, let us think of the recurring example of $\mathbb{Z}$ and $5\mathbb{Z}$. First in example 1.7 we saw that $\mathbb{Z}/5\mathbb{Z} = \{5\mathbb{Z}, \ldots, 4 + 5\mathbb{Z}\}$, and we later saw in examples 1.8 that $\phi\colon \mathbb{Z} \to \{0, \ldots, 4\}$ is a homomorphism from the group $\mathbb{Z}$ to $(\{0, \ldots, 4\}, +_{\bmod 5})$. What is $\ker\phi$? It is not hard to see that it contains precisely the multiples of 5, $5\mathbb{Z}$, since these are mapped to 0 by ϕ.

What the first isomorphism is asserting is that these two groups are isomorphic, which makes sense. Indeed, if we simply remove the suffix "$+ 5\mathbb{Z}$" from each element of the quotient group, their behaviour is identical.

Perhaps one of the more surprising results we encounter in modern group theory is *Cayley's theorem*:

Theorem 1.14 (Cayley's Theorem)**.** *Every group G is isomorphic to a subgroup of S_n, that is, to some group of permutations.*

In other words, all the groups which one may construct from the four axioms are equivalent (in the isomorphic sense) to a group of permutations. This is quite a disappointment, since historically, the axioms were formulated as an attempt to generalise the notion of a group of permutations (which was already an extensively studied mathematical object). But as we can see from the theorem, this was a failed attempt. Not all hope is lost however—the axiomatic viewpoint of group theory allows us to think much more abstractly about groups. This is after all, the reason that groups are still presented axiomatically, and not simply as substructures of S_n.

2 Vector Spaces

Here we provide a summary of the important definitions and results on vector spaces covered in the course.

2.1 Basics

A *vector space* $V(F)$, where F is a field, is an abelian group under vector addition $+$, with identity $\mathbf{0}$. We also have closure under scalar multiplication, i.e. the unary operation $\mathbf{v} \mapsto \lambda\mathbf{v}$, where $\lambda \in F$.

A *linear combination* of the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n$ in $V(F)$ is a sum of the form $\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$ for $\alpha_1, \ldots, \alpha_n \in F$.

The set $A = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\} \subseteq V$ is said to be *linearly dependent* if there exist $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$, not all zero, such that $\sum_{i=1}^{n} \alpha_i\mathbf{v}_i = \mathbf{0}$. Otherwise, we say A is *linearly independent*.

The *linear span* of a finite set $A = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of vectors is the set

$$\text{span}(A) = \{\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n : \alpha_1, \ldots, \alpha_n \in F\}.$$

If A is not finite, then

$$\text{span}(A) = \Big\{ \sum_{\mathbf{v} \in A'} \alpha_{\mathbf{v}}\mathbf{v} : A' \subseteq A \text{ is finite, and } \alpha_{\mathbf{v}} \in F \text{ for all } \mathbf{v} \in A \Big\},$$

i.e. $\text{span}(A)$ is the set of all finite linear combinations of vectors in A. If the linear span of a set B is the whole space V, we say that B *spans* V.

An important result which allows us to define dimension is the following.

Theorem 2.1 (Steinitz Replacement Theorem)**.** *If $A \subseteq V$ is a linearly independent set and B is a finite set such that* $\text{span}(B) = V$*, then $|A| \leqslant |B|$.*

An immediate consequence of the Steinitz replacement theorem: if the finite sets B_1 and B_2 are both linearly independent and span V, then $|B_1| = |B_2|$ (since by Steinitz, $|B_1| \leqslant |B_2|$ and $|B_2| \leqslant |B_1|$).

A finite set B which is linearly independent and spans the space V is called a *basis* for V. The unique number $|B|$ for any basis B of V is called the *dimension* of V, denoted by $\dim V$. We say that V is *n-dimensional*.

These ideas can be extended to bases which are infinite, but we consider only *finite dimensional vector spaces* here, i.e. vector spaces where bases are finite, and $\dim V \in \mathbb{N}$.

2.2 Linear Maps and Matrices

Definition 2.2 (Linearity)**.** Let U and V be finite dimensional vector spaces over a field F, and let $\Lambda\colon U \to V$ be a map. Then Λ is said to be *linear* if for all $\mathbf{x}, \mathbf{y} \in U$ and $\alpha \in F$,

$$\Lambda(\mathbf{x}+\mathbf{y}) = \Lambda(\mathbf{x}) + \Lambda(\mathbf{y}) \qquad \text{and} \qquad \Lambda(\alpha\mathbf{x}) = \alpha\,\Lambda(\mathbf{x}).$$

Now this is a bit of a long theorem, but be sure to read it and understand what it is saying. The proof is easier than the statement!

Theorem 2.3 (Linear map $\Leftrightarrow$ matrix)**.** *Let U and V be vector spaces with $\dim U = n$ and $\dim V = m$, and let $\Lambda\colon U \to V$ be a linear map. Let $B_U = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis for U, and let $B_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ be a basis for V.*

For each basis vector $\mathbf{u}_i$, compute the image $\Lambda(\mathbf{u}_i)$ and write it in terms of B_V as $\Lambda(\mathbf{u}_i) = \lambda_{1i}\mathbf{v}_1 + \cdots + \lambda_{mi}\mathbf{v}_m$. Now let $\mathbf{x} \in U$, and write $\mathbf{x} = \alpha_1\mathbf{u}_1 + \cdots + \alpha_n\mathbf{u}_n$ in terms of B_U. Then the image $\Lambda(\mathbf{x})$ is given by

$$\Lambda(\mathbf{x}) = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} | & & | \\ \Lambda(\mathbf{u}_1) & \cdots & \Lambda(\mathbf{u}_n) \\ | & & | \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

In other words, the entries $\beta_1, \ldots, \beta_m$ of the resulting column vector of the matrix product above are the coefficients of the image

$$\Lambda(\mathbf{x}) = \beta_1\mathbf{v}_1 + \cdots + \beta_m\mathbf{v}_n$$

in terms of the basis B_V.

In other words, every linear map Λ from an n-dimensional vector space to an m-dimensional one corresponds to an $m \times n$ matrix, denoted $[\Lambda]$, whose columns are the images of the basis vectors of the domain under Λ.

We therefore use terms about linear maps and matrices interchangeably. For example, the "kernel" of a matrix $\mathbf{M}$ is the kernel of the corresponding linear map $\mathbf{v} \mapsto \mathbf{M}\mathbf{v}$.

Example 2.4. Instead of the usual maps from $\mathbb{R}^m$ to $\mathbb{R}^n$, let us look at a more interesting one which maybe explains this result a bit better.

Consider the set of quadratic polynomials, $U = \{ax^2+bx+c : a, b, c \in \mathbb{R}\}$. Check that this forms a vector space under usual addition and scalar multiplication of polynomials. Similarly, consider the set of linear polynomials $V = \{ax + b : a, b \in \mathbb{R}\}$.

Verify that $B_U = \{1, x, 2x^2 - 1\}$ is a basis for U, and that $B_V = \{2, 1 - x\}$ is a basis for V. Now consider the linear map $\Lambda\colon U \to V$, defined by

$$\Lambda(ax^2 + bx + c) = (a - 2b + c)x + (a + b - 2c).$$

Verify that this is linear. How do we represent Λ as a matrix? First, we find the images of the basis vectors B_U. These are $\Lambda(1) = x - 2$, $\Lambda(x) = -2x + 1$, $\Lambda(2x^2-1) = x+4$. We have to write these in terms of B_V though, as the theorem states. If we mess around with comparing coefficients, we get that

$$\begin{aligned}\Lambda(1) = x - 2 &= -\tfrac{1}{2}(2) - 1(1 - x)\\ \Lambda(x) = -2x + 1 &= -\tfrac{1}{2}(2) + 2(1 - x)\\ \Lambda(2x^2 - 1) = x + 4 &= \tfrac{5}{2}(2) - 1(1 - x)\end{aligned}$$

In other words, in terms of B_V, we have

$$\Lambda(1) = \begin{pmatrix} -1/2 \\ -1 \end{pmatrix}, \qquad \Lambda(x) = \begin{pmatrix} -1/2 \\ 2 \end{pmatrix}, \qquad \Lambda(2x^2 - 1) = \begin{pmatrix} 5/2 \\ -1 \end{pmatrix}.$$

Thus the matrix representation $[\Lambda]$ is

$$[\Lambda] = \begin{pmatrix} -1/2 & -1/2 & 5/2 \\ -1 & 2 & -1 \end{pmatrix}.$$

So if we have the polynomial

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = 1 + 2x + 5(2x^2 - 1)$$

(in terms of B_U), its image under Λ is

$$[\Lambda]\mathbf{x} = \begin{pmatrix} -1/2 & -1/2 & 5/2 \\ -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 11 \\ -2 \end{pmatrix}$$

where the image is in terms of B_V, i.e. $\Lambda(\mathbf{x}) = 11(2) - 2(1 - x) = 2x + 20$. One can check that this is correct by finding $\Lambda(1 + 2x + 5(2x^2 - 1))$ directly from the definition of Λ.

Proposition 2.5. *If the columns of a matrix* $\mathbf{M}$ *are linearly independent, then* $\ker \mathbf{M} = \{\mathbf{0}\}$.

2.3 Subspaces

A subset U of the vector space V is a subspace if U itself is a vector space. To prove that $U \leqslant V$, i.e. that U is a subspace of V, we must have:

- Closure and inverse law under vector addition $(+)$
- Closure under scalar multiplication $(\lambda \cdot)$

Alternatively, we can use the following result: $U \leqslant V(F)$ if and only if for all $\mathbf{x}, \mathbf{y} \in U$ and $\alpha, \beta \in F$, $\alpha\mathbf{x} + \beta\mathbf{y} \in U$.

2.4 Change of Basis

Suppose $B_\mathrm{e} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $B_\mathrm{f} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ are two bases of a vector space V. The *transition matrix* from B_f to B_e, denoted ${}_\mathrm{e}\mathbf{P}_\mathrm{f}$, is an $n \times n$ which changes the representation of vector $\mathbf{v} = \beta_1\mathbf{f}_1 + \cdots + \beta_n\mathbf{f}_n$ in terms of B_f into the representation $\mathbf{v} = \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n$ in terms of B_e.

The transition matrix ${}_\mathrm{e}\mathbf{P}_\mathrm{f}$ is given by

$$ {}_\mathrm{e}\mathbf{P}_\mathrm{f} = \begin{pmatrix} | & & | \\ \mathbf{f}_1 & \cdots & \mathbf{f}_n \\ |_\mathrm{e} & & |_\mathrm{e} \end{pmatrix}, $$

where the columns are the basis vectors $\mathbf{f}_i$ written in terms of the basis B_e.

Example 2.6. Consider the vector space U from example 2.4. An easy basis for this vector space is $B_\mathrm{e} = \{1, x, x^2\}$. We opted to use $B_\mathrm{f} = \{1, x, 2x^2 - 1\}$ there. How do we translate something in terms of B_f into something in terms of B_e? We simply express the basis vectors of B_f in terms of B_e:

$$ \begin{aligned} 1 &= 1 + 0x + 0x^2 \\ x &= 0 + 1x + 0x^2 \\ 2x^2 - 1 &= -1 + 0x + 2x^2. \end{aligned} $$

Putting these as the columns, we get

$$ {}_\mathrm{e}\mathbf{P}_\mathrm{f} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. $$

Indeed, if a polynomial is given as $\mathbf{x}_\mathrm{f} = 3 - 2x + 5(2x^2 - 1)$ in terms of B_f, then

$$ \mathbf{x}_\mathrm{e} = {}_\mathrm{e}\mathbf{P}_\mathrm{f}\mathbf{x}_\mathrm{f} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 10 \end{pmatrix}. $$

Indeed, it is not hard to check that the given polynomial is $-2 - 2x + 10x^2$, upon expansion.

Proposition 2.7 (Change of Basis)**.** *Let B_e and B_f be two bases for V. Then*

$$ {}_f\mathbf{P}_e = {}_e\mathbf{P}_f{}^{-1}. $$

Example 2.8. Consider the map $T\colon \mathbb{R}^3 \to \mathbb{R}^3$. Using the standard basis $B_e = \{(1,0,0),(0,1,0),(0,0,1)\}$ for $\mathbb{R}^3$, T carries out the following:

$$ T(x,y,z) = (x+y, y+z, z+x). $$

Rather than using the standard basis, we wish to interpret vectors in the domain using the basis $B_f = \{(1,0,1),(1,0,-1),(0,1,1)\}$, and vectors in the codomain using basis $B_g = \{(2,0,1),(1,2,3),(1,1,2)\}$. In other words, we want that, for example, $T(1,2,3)$ is interpreted as

$$ T(1(1,0,1) + 2(1,0,-1) + 3(0,1,1)) = T(3,3,2), $$

where now $(3,3,2)$ is equivalent to $(1,2,3)$ but in terms of B_e rather than B_f, and we may apply the definition to get $T(3,3,2) = (6,5,5)$. But now this output in terms of the standard basis too: we want to write it as (α,β,γ) so that it is interpreted with respect to B_g as $\alpha(2,0,1) + \beta(1,2,3) + \gamma(1,1,2)$. Solving some simultaneous equations, we find that in terms of B_g, the resulting vector is $(6,11,-17)$.

Thus what we want is a modified operator T so that $T(1,2,3) = (6,11,-17)$, without having to manually change bases before and after.

We can obtain the desired representation of T in a similar method to that of example 2.4, or else we can do the following. Write T as a matrix in terms of the standard basis first. Let us denote this representation by ${}_e[T]_e$. So

$$ {}_e[T]_e = \begin{pmatrix} | & | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}. $$

Now if we are given a vector $\mathbf{x}$ in terms of the basis B_f, we can transform it into the standard basis by doing $\mathbf{x} \mapsto {}_e\mathbf{P}_f\mathbf{x}$ first, and then apply ${}_e[T]_e$. This composition gives us a new matrix, which we will call ${}_e[T]_f = {}_e[T]_e\,{}_e\mathbf{P}_f$. This now interprets input vectors in terms of B_f, not B_e. Let's work it out:

$$ {}_e[T]_f = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}}_{{}_e\mathsf{P}_f} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 2 & 0 & 1 \end{pmatrix}. $$

But just as we remarked before, this will give us outputs $\mathbf{y}$ in terms of the standard basis, whereas we want them in terms of B_g. If we do $\mathbf{y} \mapsto {}_g\mathbf{P}_e\mathbf{y}$, this gives us the

desired representation. Now

$$ {}_{\mathsf{g}}\mathbf{P}_{\mathsf{e}} = ({}_{\mathsf{e}}\mathbf{P}_{\mathsf{g}})^{-1} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 3 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 3 & -2 \\ -2 & -5 & 4 \end{pmatrix}, $$

so we get the desired representation ${}_{\mathsf{g}}[T]_{\mathsf{f}} = {}_{\mathsf{g}}\mathbf{P}_{\mathsf{e}}\,{}_{\mathsf{e}}[T]_{\mathsf{e}}\,{}_{\mathsf{e}}\mathbf{P}_{\mathsf{f}}$, which when worked out, yields

$$ {}_{\mathsf{g}}[T]_{\mathsf{f}} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & -2 & 5 \\ 1 & 3 & -8 \end{pmatrix}. $$

Indeed, we have

$$ \begin{pmatrix} 0 & 0 & 2 \\ 0 & -2 & 5 \\ 1 & 3 & -8 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 11 \\ -17 \end{pmatrix}, $$

as desired.

2.5 The Dimension Theorem

An important theorem about dimension is the so-called *dimension theorem*.

Definition 2.9 (Image)**.** The *image* of a linear map $\Lambda\colon U \to V$, denoted $\mathrm{Im}(\Lambda)$, is the set $\mathrm{Im}(\Lambda) = \{\Lambda(\mathbf{u}) : u \in U\} \subseteq V$.

Theorem 2.10 (Dimension Theorem)**.** *Let $\Lambda\colon U \to V$ be a linear map defined on an finite dimensional vector space U. Then*

$$ \dim U = \dim(\ker \Lambda) + \dim(\mathrm{Im}\,\Lambda). $$

If $\dim(\ker\Lambda) \neq 0$, then Λ is said to be *singular*.

Example 2.11. Consider the map $T\colon \mathbb{R}^2 \to \mathbb{R}^3$, defined by

$$ T(x, y) = (3x + y, x + y, x + 3y) $$

(standard basis). We wish to determine $\ker T$ and whether or not T is singular. For $\ker T$, we determine which vectors $\mathbf{v}$ send $T(\mathbf{v})$ to $\mathbf{0}$. A simple calculation shows that the only solution to this is $x = y = 0$, hence $\ker T = \{\mathbf{0}\}$ and thus $\dim(\ker T) = 0$ (note that the empty set is a basis for $\{\mathbf{0}\}$). Hence the map T is nonsingular.

Example 2.12. Consider the map $T\colon \mathbb{R}^3 \to \mathbb{R}^3$, where $(x, y, z) \mapsto (3x, y + z, 0)$. Thus

$$ [T] = \begin{pmatrix} | & | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \\ | & | & | \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}. $$

Let the columns of this matrix be denoted $\mathbf{c}_1$, $\mathbf{c}_2$ and $\mathbf{c}_3$. Observe that the third column of $[T]$ can be written as a linear combination of the others ($0\mathbf{c}_1 + 1\mathbf{c}_2$), and hence the set $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ is linearly dependent. The two remaining vectors form a linearly independent set $\{\mathbf{c}_1, \mathbf{c}_2\}$, since it is impossible to find α_i's not all zero such that $\alpha_1\mathbf{c}_1 + \alpha_2\mathbf{c}_2 = \mathbf{0}$.

Now each vector in $\mathbb{R}^3$ can be expressed as $\mathbf{x} = (x, y, z)$, and by matrix multiplication, T maps this vector to $x\mathbf{c}_1 + y\mathbf{c}_2 + z\mathbf{c}_3 \in \operatorname{Im} T$. In general, any vector transformed by T can be expressed as a linear combination of its columns, so the columns of a matrix representing a linear transformation span its image, i.e. $\operatorname{Im} T = \operatorname{span}\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$.

Now if we trim the columns of the matrix to a set of linearly independent vectors (as we have done for $[T]$ obtaining $\{\mathbf{c}_1, \mathbf{c}_2\}$) we obtain a basis for $\operatorname{Im} T$, and hence we can determine the dimension of the image (i.e. the number of vectors in a basis), i.e. the rank. Thus the rank of T is 2, since $|\{\mathbf{c}_1, \mathbf{c}_2\}| = 2$.

In general, the rank of a transformation is the number of linearly independent columns in its matrix representation. It turns out that this result holds if we consider rows instead of columns ("row rank = column rank").

Now by the dimension theorem, we can determine the nullity ($\dim \ker$) of T also:

$$\dim(\ker T) = \dim(\mathbb{R}^3) - \dim(\operatorname{Im} T) = 3 - 2 = 1$$

Hence the kernel of T has dimension 1, i.e. $\ker T \neq \{\mathbf{0}\}$. In fact, one may deduce that any vector of the form $(0, a, -a)$ for any $a \in \mathbb{R}$ is mapped to $\mathbf{0}$ by T.