Honours Algebra - Week 2 - Linear Mappings as Matrices

Antonio León Villares

January 2022

Contents

Based on the notes by Iain Gordon, Sections 1.7 - 1.8 & 2.1 - 2.2

1 Linear Mappings

1.1 The Morphisms

- **What is a homomorphism?**
  - let V, W be vector spaces
  - let $\underline{v}_1, \underline{v}_2 \in V$ and $\lambda \in \mathbb{F}$
  - a **homomorphism of $\mathbb{F}$ - vector spaces** is mapping of the form:
    $$f : V \to W$$
    such that:
    $$f(\underline{v}_1 + \underline{v}_2) = f(\underline{v}_2) + f(\underline{v}_2)$$
    $$f(\lambda \underline{v}_1) = \lambda f(\underline{v}_1)$$
  - this is also known as a **linear mapping** (or a F-linear mapping)
- **What is an isomorphism?**
  - a **bijective homomorphism**
- **What are isomorphic vector spaces?**
  - vector spaces for which an **isomorphism** exists between the two
- **What is an endomorphism?**
  - a **homomorphism** from a vector space to **itself**
- **What is an automorphism?**
  - an **isomorphism** from a vector space to **itself**
- **What is a fixed point of a mapping?**
  - given a mapping $f : X \to X$, a **fixed point** of f is a point $x \in X$ sent to itself under f
  - the set of all fixed points is:
    $$X^f = \{x | x \in X, f(x) = x\}$$

1.1.1 Examples

- projections are linear mappings:
  $$pr_i : (\lambda_1, \lambda_2, \ldots, \lambda_n) \to \lambda_i$$
- squaring ($\lambda \to \lambda^2$) is **not** linear (except for $\mathbb{F} = \mathbb{Z}_2$)
- **projection mappings**:
  $$(V \oplus W) \to W \qquad (V \oplus W) \to V$$
  are linear
- the **canonical injections**:
  $$v \to (v, 0) \qquad w \to (0, w)$$
  are linear

- the bijective map defining a linear combination of basis elements:

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) \to \sum_{i=1}^{n} \lambda_i \underline{v}_i$$

  is linear

- the **automorphisms** of a vector space V form a subgroup of its permutation group (known as **general linear group** or **automorphism group** of V)

1.1.2 Exercises (TODO)

1. **Show that a composition of homomorphisms is a homomorphism.**

   Define homomorphisms $f : V \to W$ and $f : U \to V$. Consider the composition $f \circ g : U \to W$.

   Consider $\lambda \underline{u}_1 + \underline{u}_2$. Then:

$$\begin{aligned}
&(f \circ g)(\lambda \underline{u}_1 + \underline{u}_2) \\
=&f(g(\lambda \underline{u}_1 + \underline{u}_2)) \\
=&f(g(\lambda \underline{u}_1) + g(\underline{u}_2)) \\
=&f(g(\lambda \underline{u}_1)) + f(g(\underline{u}_2)) \\
=&f(\lambda g(\underline{u}_1)) + f(g(\underline{u}_2)) \\
=&\lambda(f \circ g)(\underline{u}_1) + (f \circ g)(\underline{u}_2)
\end{aligned}$$

2. **Show that if $f : V \to W$ is an isomorphism, then $f^{-1} : W \to V$ is also an isomorphism**

   We know that since f is an isomorphism, it is a bijection, so its inverse f^{-1} exists. We want to show that:

$$f^{-1}(\lambda \underline{w}_1 + \underline{w}_2) = \lambda f^{-1}(\underline{w}_1) + f^{-1}(\underline{w}_2)$$

   We know that:

$$\lambda \underline{w}_1 + \underline{w}_2 = f(f^{-1}(\lambda \underline{w}_1 + \underline{w}_2))$$

   Moreover,

$$f(\lambda f^{-1}(\underline{w}_1) + f^{-1}(\underline{w}_2)) = \lambda f(f^{-1}(\underline{w}_1)) + f(f^{-1}(\underline{w}_2)) = \lambda \underline{w}_1 + \underline{w}_2$$

   In other words:

$$f(\lambda f^{-1}(\underline{w}_1 + \underline{w}_2)) = f(\lambda f^{-1}(\underline{w}_1) + f^{-1}(\underline{w}_2)) \implies f^{-1}(\lambda \underline{w}_1 + \underline{w}_2) = \lambda f^{-1}(\underline{w}_1) + f^{-1}(\underline{w}_2)$$

3. **Show that the image of a vector subspace under a homomorphism is again a vector subspace. Moreover, show that the preimage of a vector susbspace under a homomorphism is a vector subspace.**

4. **Consider a vector space V, with the set of its endomorphisms $End(V)$. Show that $V^f \subseteq V$ is a vector subspace.**

5. **Show that, given vector spaces $V_1, V_2, \ldots, V_n, W$ and homomorphisms $f_i : V_i \to W$, we can define a new homomorphism:**

$$f : V_1 \oplus V_2 \oplus \ldots \oplus V_n \to W$$

via:

$$f(v_1, v_2, \ldots, v_n) = \sum_{i=1}^{n} f_i(v_i)$$

Given the above, we can define a bijection:

$$Hom(V_1, W) \times \ldots \times Hom(V_1, W) \to Hom(V_1 \oplus \ldots \oplus V_n, W)$$

6. **Show that, given vector spaces $W_1, W_2, \ldots, W_n, V$ and homomorphisms $g_i : V \to W_i$, we can define a new homomorphism:**

$$g : V \to W_1 \oplus W_2 \oplus \ldots \oplus W_n$$

via:

$$g(v) = (g_1(v), \ldots, g_n(v))$$

Given the above, we can define a bijection:

$$Hom(V, W_1) \times \ldots \times Hom(V, W_n) \to Hom(V, W_1 \oplus \ldots \oplus W_n)$$

7. **Let $X = \mathbb{R}^2$ be a vector space over $\mathbb{F} = \mathbb{R}$. Determine the fixed point set of the following functions:**
   (a) $f(a, b) \to (a, b)$
   For this, $X^f = X$.
   (b) $f(a, b) \to (b, a)$
   For this, the fixed point set is defined by all points for which $a = b$, so:

$$X^f = \{(a, a) | a \in \mathbb{F}\}$$

   Notice, in this case, X^f is the diagonal line through the origin, and f is a function which *reflects* points in X about X^f.
   (c) $f(a, b) \to (-b, a)$
   Notice, if $a = -b$ and $b = a$, this implies that $a = -a = b = -b$. In particular, we must then have:

$$X^f = (0, 0)$$

   Notice, X^f is the origin, and f defines a 90^o anticlockwise rotation.

8. **How many vector subspaces are there in $\mathbb{R}^2$ that are sent to themselves under the reflection $(x, y) \to (x, -y)$? Which vector subspaces in $\mathbb{R}^3$ are sent to themselves by the reflection $(x, y, z) \to (x, y, -z)$?**

1.2 Complementary Subspaces

- **What is a complementary subspace?**
  - consider a vector space V with subspaces V_1, V_2
  - V_1 and V_2 are **complementary subspaces** if we can define a bijection:

$$f : V_1 \times V_2 \to V$$

    via:

$$f(\underline{v}_1, \underline{v}_2) = \underline{v}_1 + \underline{v}_2$$

1.2.1 Examples

- consider $V = \mathbb{R}^2$ defined over $\mathbb{F} = \mathbb{R}$. Then the subspaces:

$$V_1 = \langle 1, 0 \rangle = \{(\mu, 0) | \mu \in \mathbb{F}\}$$

$$V_2 = \langle b, c \rangle = \{(\lambda b, \lambda c) | \lambda \in \mathbb{F}\}$$

are **complementary**. We can see if the addition mapping:

$$((\mu, 0), (\lambda b, \lambda c)) \to (\mu + \lambda b, \lambda c)$$

Notice, if $c = 0$, this won't be bijective (since $(\mu + \lambda b, \lambda c)$ can be generated using many different combinations of $\mu, \lambda b$). Hence, V_1 and V_2 won't be complementary subspaces (in fact, we would have that $V_1 = V_2$ - they are the same line).

If $c \neq 0$, then for any point (x, y), we can find unique λ satisfying $\lambda c = y$, and unique μ such that $\mu + \lambda b = x$. In other words, V_1 and V_2 are complementary when $c \neq 0$ (in other words, when the lines aren't parallel).

1.2.2 Exercises

1. **Show that the bijection defining complementary subspaces is an isomorphism:**

$$f : V_1 \oplus V_2 \to V$$

Here $V_1 \oplus V_2$ is an *internal direct sum* (not to be confused with *external direct sum*: the IDS refers to an operation on subspaces, whilst the EDS is more generally applicable to vector spaces).

1.3 Theorem: Classification of Vector Spaces by their Dimension

> *Let $n \in \mathbb{N}$. A **vector space** over a field $\mathbb{F}$ is **isomorphic** to $\mathbb{F}^n$ **if and only if** it has dimension n.*
> *In other words, for finite dimensional vector spaces, up to isomorphism, all that "matters" is its dimension. [Theorem 1.7.7]*

Let V be vector spaces over $\mathbb{F}$.

1. ($\Longleftarrow$): say V has dimension n. Then, it has a basis:

$$E = \{\underline{a}_1, \ldots, \underline{a}_n\}$$

We know that, for basis elements, the following is a bijective map (Theorem 1.5.11):

$$f : \mathbb{F}^n \to V$$

given by:

$$(\alpha_1, \ldots, \alpha_n) \to \sum_{i=1}^{n} \alpha_i \underline{a}_i$$

Moreover, this is a linear map, so in particular, it defines an isomorphism, as required.

2. ($\Longrightarrow$): assume there exists an isomorphism:

$$f : \mathbb{F}^n \to V$$

We know that $\mathbb{F}^n$ has a basis of n elements:

$$E = \{\underline{e}_1, \ldots, \underline{e}_n\}$$

Notice, since f is a bijection, it suffices to show that $f(E)$ is a basis for V, since then V will have a basis of n elements, as required. Hence, we need to show that $f(E)$ is:

- *a generating set for* V: pick $\underline{v} \in V$. Since f is a bijection, $\exists!\underline{x} \in F^n$ such that $f(\underline{x}) = \underline{v}$. Moreover, we can write $\underline{x}$ in terms of the basis elements of $\mathbb{F}^n$:

$$\underline{x} = \sum_{i=1}^{n} \alpha_i \underline{e}_i$$

Hence, using the linearity of the homomorphism:

$$f(\underline{x}) = \underline{v} \implies \sum_{i=1}^{n} \alpha_i f(\underline{e}_i) = \underline{v}$$

In other words, the set $f(E)$ is generating.

- *linearly independent*: since E is a basis, it is linearly independent, so:

$$\sum_{i=1}^{n} \alpha_i \underline{e}_i = \underline{0}$$

only if $\alpha_i = 0, \forall i \in [1, n]$. But then, applying f means that:

$$\sum_{i=1}^{n} \alpha_i f(\underline{e}_i) = f(\underline{0}) = \underline{0}$$

In other words, the elements of $f(E)$ are also linearly independent.

Hence, $f(E)$ is a basis for V. Moreover, it contains n elements, so $dimV = n$, as required.

1.4 Lemma: Linear Mappings and Bases

Define the set of all ***homomorphisms*** *between vector spaces* V, W *as:*

$$Hom_{\mathbb{F}}(V, W) \subseteq Maps(V, W)$$

Let B *be a* ***basis*** *for* V*. We can define a* ***bijection****:*

$$Hom_{\mathbb{F}}(V, W) \to Maps(B, W)$$

via:

$$f \to f_B$$

where f_B *is* f*, with its domain restricted to* B*.*
This means that any homomorphism can be defined by the values it takes at a basis. [Lemma 1.7.8]

Proof. Let Φ define the bijection.

Recall, Φ is injective if:

$$\Phi(f) = \Phi(g) \implies f = g$$

Let f, g be linear mappings $f, g : V \to W$. If $\forall \underline{v} \in B$ we have:

$$f(v) = g(v)$$

then we must have that, $\forall \lambda_i \in \mathbb{F}, v_i \in B$:

$$\sum_{i=1}^{n} \lambda_i f(v_i) = \sum_{i=1}^{n} \lambda_i g(v_i)$$

Since f, g are homomorphisms, then:

$$f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = g\left(\sum_{i=1}^{n} \lambda_i v_i\right)$$

In other words, if f and g are equal for each element in the basis B, then they must be equal for any element in V. Hence, Φ must be injective.

For surjectivity, we need to ensure that for each element $g \in Maps(B, W)$, we have at least one other element $\bar{g} \in Hom_{\mathbb{F}}(V, W)$, such that:

$$\Phi(g) = \bar{g}$$

Indeed, take any $g : B \to W$. We can extend it to a homomorphism of the form $\bar{g} : V \to W$. Notice, any element $v \in V$ can be written as:

$$v = \sum_{i=1}^{n} \lambda_i v_i$$

where each $v_i \in B$. In other words, the mapping:

$$\bar{g}(v) = \sum_{i=1}^{n} \lambda_i g(v_i)$$

is clearly a homomorphism. In other words, we can map any element $\bar{g} \in Hom_{\mathbb{F}}(V, W)$ to $g \in Maps(B, W)$ using the above.

□

1.4.1 Exercises (TODO)

1. **Let V, W be vector spaces over a field $\mathbb{F}$. Show that $Hom_{\mathbb{F}}(V, W)$ is a vector subspace of the set of all mappings $Maps(V, W)$. Moreover, show that its vector space structure is given similarly to the free vector space. Show that:**

   $$dim Hom_{\mathbb{F}}(V, W) = (dim V)(dim W)$$

   where I am using the convention $0 \times \infty = 0$

2. **Let V be a finite dimensional vector space, and let U be a proper vector subspace. Show that there exists at least one (and in fact many different) vector subspace(s) of V complementary to U. If you're brave, try to do this also for not necessarily finite dimensional vector spaces**

1.5 Proposition: Left and Right Inverses

1. *Every injective homomorphism:*

$$f : V \to W$$

has a ***left inverse*** $g : W \to V$ *such that:*

$$g \circ f = 1_V$$

2. *Every surjective homomorphism:*

$$f : V \to W$$

has a ***right inverse*** $g : W \to V$ *such that:*

$$f \circ g = 1_W$$

[Proposition 1.7.9]

Proof.

- Existence of Left Inverse For Injective Mappings
  - we begin by noting that $f(V)$ is a subspace of W
  - in particular, by Exercise 2 above, we can find a subspace U of W which is **complementary** to $f(V)$
    - in exercise 2 we would require $f(V)$ to be a proper subspace
    - if it isn't, then $f(V) = W$, so f would be surjective, and so an isomorphism
    - isomorphisms are bijective, and so, have a left inverse
  - since U and $f(V)$ are complementary, we know that $\forall w \in W$, we have unique $u \in U, f(v) \in f(V)$ such that:

$$w = u + f(v)$$

  - moreover, by injectivity of f, $f(v)$ is uniquely produced by $v \in V$
  - hence, we can define a mapping $g : W \to V$ such that:

$$g(w) = v, \qquad w = u + f(v)$$

  (Apparently this mapping then shows that $g(f(v)) = v$, as required)

- Existence of Right Inverse For Surjective Mappings
  - we can pick a basis $B \subseteq W$

– using the fact that f is surjective, we can define a mapping of sets:

$$\bar{g} : B \to V$$

such that:

$$f(\bar{g}(b)) = b$$

(we can think of $\bar{g}$ as mapping basis elements to enough elements of V such that f can send them back to the basis elements)

– by (1.4), we know that there exists a bijection between $Hom_{\mathbb{F}}(V, W)$ and $Maps(B, W)$. In particular, we can find $g \in Hom_{\mathbb{F}}(V, W)$, such that for $b \in B$, we have $g(b) = \bar{g}(b)$

– thus, $\forall b \in B$, we have:

$$f(g(b)) = b$$

– notice, if $w \in W$, then we can write:

$$w = \sum_{i=1}^{n} \alpha_i b_i$$

– but then, using the fact that f, g are homomorphisms:

$$\begin{aligned} f(g(w)) &= f\left(g\left(\sum_{i=1}^{n} \alpha_i b_i\right)\right) \\ &= f\left(\sum_{i=1}^{n} \alpha_i g(b_i)\right) \\ &= \sum_{i=1}^{n} \alpha_i f(g(b_i)) \\ &= \sum_{i=1}^{n} \alpha_i b_i \\ &= w \end{aligned}$$

– thus, we have found a right inverse, as required

□

2 The Rank-Nullity Theorem

2.1 Images and Kernels

- **What is the image of a homomorphism?**
  - let f be a homomorphism $f : V \to W$
  - the subset $f(V) \subseteq W$ is the **image** of f
  - we denote it $im(f) = f(V)$
  - $im(f)$ is a **subspace** of W
- **What is the kernel of a homomorphism?**
  - the **preimage** of $\underline{0} \in W$
  - in other words:

$$ker(f) = \{\underline{v} \mid \underline{v} \in V, f(\underline{v}) = 0\}$$

  - $ker(f)$ is a **subspace** of V

2.1.1 Examples

- if $f(a,b) = (0,0)$, then:
$$ker(f) = \mathbb{R}^2 \qquad im(f) = \{(0,0)\}$$
- if $f(a,b) = (b-a, a-b)$, then:
$$ker(f) = \{(a,a) \mid a \in \mathbb{R}\}$$
$$im(f) = \{(c,-c) \mid c \in \mathbb{R}\}$$
- if $f(a,b) = (-b-a, a-b)$, then:
$$ker(f) = \{(0,0)\}$$
(since we require $-b-a = 0 = a-b \implies a = -a \implies a = 0$)
$$im(f) = \mathbb{R}^2$$
(since we obtain a system
$$-b-a = x, x \in \mathbb{R}$$
$$a-b = y, y \in \mathbb{R}$$
which has unique solutions:
$$2a = y-x \implies a = \frac{y-x}{2}$$
$$b = \frac{y-x}{2} - y = \frac{-y-x}{2}$$
)

2.2 Lemma: Injectivity and the Kernel

> *A homomorphism is injective* ***if and only if*** *its kernel* ***only*** *contains* $\underline{0}$. *[Lemma 1.8.2]*

Proof. We prove in both directions.

- ($\implies$): assume that f is an injective homomorphism. Then, since $f(\underline{0}) = \underline{0}$, no other element of V will be mapped to $\underline{0}$ by f, so:
$$ker(f) = \{\underline{0}\}$$
- ($\impliedby$): assume that $ker(f) = \{\underline{0}\}$. Further, assume that $f(\underline{v}_1) = f(\underline{v}_2)$. Then:
$$\begin{aligned} & f(\underline{v}_1) - f(\underline{v}_2) = \underline{0} \\ \implies & f(\underline{v}_1 - \underline{v}_2) = \underline{0} \text{ (by linearity of } f) \\ \implies & \underline{v}_1 - \underline{v}_2 \in ker(f) \\ \implies & \underline{v}_1 - \underline{v}_2 = 0 \\ \implies & \underline{v}_1 = \underline{v}_2 \end{aligned}$$
Hence, it follows that f is injective.

□

2.3 Theorem: Rank-Nullity Theorem

- **What is the rank of a homomorphism?**
  - the dimension of $im(f)$
- **What is the nullity of a homomorphism?**
  - the dimension of $ker(f)$

Let $f : V \to W$ be a homomorphism. Then:

$$dimV = dim(ker(f)) + dim(im(f))$$

[Theorem 1.8.4]

Proof. We begin by noticing that if V is a finitely generated vector space, then:

- since $ker(f)$ is a subspace of V, then $dim(ker(f)) \leq dim(V)$
- if E is a generating set of V, then $f(E)$ is a generating set for $f(V) = im(f)$. Hence, $im(f)$ must also be finitely generated.

We first note that, if the rank or nullity of f are infinite,we can immediately see that the $dimV = \infty$ (by the work above).

We now consider having a finite rank and nullity. Then, we can define the basis of $ker(f)$:

$$A = \{v_1, \ldots, v_r\}$$

and of $im(f)$:

$$B = \{w_1, \ldots, w_s\}$$

Our aim is to show that there exists some basis E of V, such that:

$$|E| = r + s$$

To do this, define $\bar{w}_i \in V$, such that:

$$f(\bar{w}_i) = w_i$$

We know that such $\bar{w}_i$ exist, since $w_i \in im(f)$ (they are basis vectors of $im(f)$). We claim that:

$$E = \{v_1, \ldots, v_r, \bar{w}_1, \ldots, \bar{w}_s\}$$

Hence, we need to show that:

- $\langle E \rangle = V$

- since B is a basis for $f(V)$, we know that we can find $\alpha_i \in \mathbb{F}$ and $v \in V$ such that:

$$f(v) = \sum_{i=1}^{s} \alpha_i w_i$$

- now consider:

$$f\left(v - \sum_{i=1}^{s} \alpha_i \bar{w}_i\right)$$

Applying linearity:

$$f\left(v - \sum_{i=1}^{s} \alpha_i \bar{w}_i\right) = f(v) - \sum_{i=1}^{s} \alpha_i f(\bar{w}_i) = 0$$

where we have used the fact that $f(\bar{w}_i) = w_i$.

- we thus know that:

$$v - \sum_{i=1}^{s} \alpha_i \bar{w}_i \in ker(f)$$

- then, since A is a basis for $ker(f)$, we can write:

$$v - \sum_{i=1}^{s} \alpha_i \bar{w}_i = \sum_{i=1}^{r} \beta_i v_i \implies v = \sum_{i=1}^{r} \beta_i v_i + \sum_{i=1}^{s} \alpha_i \bar{w}_i$$

- thus, we have shown that if $v \in V$, it can be generated by E

- E is linearly independent
  - assume that we have α_i, β_i such that:

$$\sum_{i=1}^{r} \beta_i v_i + \sum_{i=1}^{s} \alpha_i \bar{w}_i = 0$$

  - applying f:

$$f\left(\sum_{i=1}^{r} \beta_i v_i + \sum_{i=1}^{s} \alpha_i \bar{w}_i\right) = f(0) \implies \sum_{i=1}^{r} \beta_i f(v_i) + \sum_{i=1}^{s} \alpha_i w_i = 0$$

  - since $v_i \in A, v_i \in ker(f)$, we have that:

$$\sum_{i=1}^{r} \beta_i f(v_i) = 0$$

  - moreover, since w_i are part of the basis B, they are linearly independent, so if $\sum_{i=1}^{s} \alpha_i w_i = 0$, then $\alpha_i = 0$
  - if each of the α_i are 0, then we have:

$$\sum_{i=1}^{r} \beta_i v_i = 0$$

  - again, the v_i are part of A, so they are linearly independent, and so, $\beta_i = 0$
  - thus, the elements in E are linearly independent

Thus, E is a basis for V, and so:

$$dimV = |E| = r + s$$

□

2.3.1 Exercises (TODO)

1. **Show that two subspaces U, W of a vector space V are complementary if and only if:**

   - $V = U + W$
   - $U \cap W = \{\underline{0}\}$

   Recall, U, W are complementary if and only if the following bijection exists:

   $$\phi : (u, w) \to u + w, \qquad u \in U, w \in W$$

   We claim that:

   - ϕ is surjective **if and only if** $U + W = V$
   - ϕ is injective **if and only if** $U \cap W = \underline{0}$

   If ϕ is surjective, then $\forall v \in V, \exists u \in U, w \in W$ such that:

   $$\phi(u, w) = u + w = v$$

   Hence, for any element in $v \in V$, we can find elements in U, W which generate v, so:

   $$U + W = V$$

   Similarly, if $U + W = V$, then $\forall v \in V, \exists u \in U, w \in W$ such that:

   $$v = u + w$$

   But then, $u + w$ is nothing else but $\phi(u, w)$, so ϕ maps to every element in V, and so, ϕ is surjective.

   Now assume that ϕ is injective. Furthermore, lets assume that $U \cap W \neq \{\underline{0}\}$. Then, we can find an element $a \in U \cap W$, with $a \neq \underline{0}$. But then:

   $$\phi(a, -a) = a - a = \underline{0}$$

   (this is well defined, since a is in both U and W). Since $\phi(0, 0) = \underline{0}$, then clearly ϕ can't be injective, a contradiction. Hence, if ϕ is injective, then $U \cap W = \underline{0}$.

   Now assume that $U \cap W = \{\underline{0}\}$. ϕ will be injective if $\phi(u, w) = \underline{0}$ is only possible when $u = w = \underline{0}$ (since we know that $\phi(0, 0) = \underline{0}$). Consider $u, w \neq 0$ with $\phi(u, w) = 0$. Then:

   $$u + w = 0 \implies u = -w$$

   This means that $u \in W$ and $W \in U$. Hence, $u, w \in U \cap W$. But $U \cap W = \{\underline{0}\}$, so $u = w = \underline{0}$, as required.

   Hence, ϕ is a bijection (and so U, W are complementary) if and only if $U * W = V$ **and** $U \cap W = \{\underline{0}\}$

2. **Show that two subspaces U, W of a vector space V are complementary if and only if:**

   - $V = U + W$
   - $dim U + dim W \leq dim V$

3. **Show that the kernel of a non-zero linear mapping $V \to F$ is a hyperplane, in the sense that together with another vector, the hyperplane and the vector generate V.**

4. **Let**

   $$\phi : V \to V$$

   be an endomorphism of a finitely dimensional vector space V. Show that:

   $$ker(\phi \circ \phi) = ker(\phi)$$

if and only if:

$$V = ker(\phi) \oplus im(\phi)$$

We begin by noting that $V = ker(\phi) \oplus im(\phi)$ is equivalent to saying that $ker(\phi)$ and $im(\phi)$ are **complementary**, and so, by the previous exercise, it is true if and only if:

- $V = U + W$
- $U \cap W = \{\underline{0}\}$

We proceed with the proof:

- $(1 \implies 2)$: assume that
$$ker(\phi \circ \phi) = ker(\phi)$$
  - define $U = ker(\phi)$ and $W = im(\phi)$
  - pick $v \in U \cap W$
  - since $v \in U$, we know that $\phi(v) = 0$
  - since $v \in W$, we know that $\exists \bar{v} \in V$ such that:
$$\phi(\bar{v}) = v$$
  - now consider:
$$\phi^2(\bar{v}) = \phi(\phi(\bar{v})) = \phi(v) = 0$$
  - in other words, $\bar{v} \in ker(\phi \circ \phi)$
  - but by assumption, $ker(\phi \circ \phi) = ker(\phi)$, so $\bar{v} \in ker(\phi)$
  - hence, $\phi(\bar{v}) = 0$
  - but since $\phi(\bar{v}) = v$ it follows that $v = 0$
  - hence, any element $v \in U \cap W$ must be 0, so:
$$U \cap W = \{0\}$$
  - by the equivalence outlined at the start, we must then have $V = ker(\phi) \oplus im(\phi)$
- $(2 \implies 1)$: assume that
$$V = ker(\phi) \oplus im(\phi)$$
  - notice, if $v \in ker(\phi)$, then:
$$\phi \circ \phi(v) = \phi(0) = 0$$
    so any element in $ker(\phi)$ is also in $ker(\phi \circ \phi)$, so:
$$ker(\phi) \subseteq ker(\phi \circ \phi)$$
  - pick $v \in ker(\phi \circ \phi)$. then:
$$\phi^2(v) = \phi(\phi(v)) = 0$$
    In other words, $\phi(v) \in ker(\phi)$
  - notice, $\phi(v) \in ker(\phi) = U$, but also $\phi(v) \in im(f) = W$, so we have that:
$$\phi(v) \in U \cap W$$
  - by assumption, we know that $V = ker(\phi) \oplus im(\phi)$ implies that $U \cap W = \{0\}$, so it follows that:
$$\phi(v) = 0$$
  - hence, if $v \in ker(\phi \circ \phi)$, then also $v \in ker(\phi)$, so:
$$ker(\phi \circ \phi) \subseteq ker(\phi)$$

– in other words, if $V = ker(\phi) \oplus im(\phi)$, then:

$$ker(\phi \circ \phi) = ker(\phi)$$

as required

What this is saying is that, if we have an idempotent endomorphism, we can decompose the vector space using the kernel and image of the endomorphism.

5. **An element f is idempotent if $f^2 = f$. By the previous exercise, the idempotent endomorphisms of V correspond uniquely to a decomposition of V into a direct product of complementary subspaces. Show that:**

$$f \to (im(f), ker(f))$$

leads to a bijection:

$$\{f \mid f \in End(V), f^2 = f\} \to \{(I, K) \mid (I, K) \in \mathcal{P}(V)^2, I, K \subseteq V, I \oplus K = V\}$$

3 Linear Mappings and Matrices

3.1 Theorem: Assigning Matrices to Linear Mappings

*Let $\mathbb{F}$ be a **field**, and let $m, n \in \mathbb{N}$.*
*There exists a **bijection** between:*

- *the space of homomorphisms $\mathbb{F}^m \to \mathbb{F}^n$*
- *the set of $n \times m$ matrices with entries in $\mathbb{F}$*

via:

$$M : Hom_{\mathbb{F}}(\mathbb{F}^m, \mathbb{F}^n) \to Mat(n \times m; \mathbb{F})$$

$$M : f \to [f]$$

*We call $[f]$ the **representing matrix** of the mapping f.*
The columns of $[f]$ are given by applying f to the standard basis vectors $\underline{e}_i, i \in [1, m]$ of $\mathbb{F}^m$:

$$[f] := (f(\underline{e}_1) \mid \ldots \mid f(\underline{e}_m))$$

[Theorem 2.1.1]

Proof. This uses (1.4), using $V = \mathbb{F}^m, W = \mathbb{F}^n$. We see that the homomorphism f is determined by what it does to the basis elements of $V = \mathbb{F}^m$

□

3.1.1 Examples

- the **identity matrix** is defined by the identity homomorphism given by $id_{\mathbb{F}^m}(\underline{e}_i) = \underline{e}_i$:

$$\mathbb{I} = [id_{\mathbb{F}^m}] = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

  Each column is just the elements of the standard basis $\underline{e}_i$. Conciseley, $\mathbb{I}_{ij} = \delta_{ij}$, the Kronecker Delta.

- if $m \geq n$ and f is the homomorphism:

$$f : (x_1, \dots, x_m) \to (x_1, \dots, x_n)$$

  (in other words, elements beyond x_n are "ignored), the corresponding matrix will be given by:

$$A_{ij} = \begin{cases} \delta_{ij}, & j \leq n \\ 0, & j > n \end{cases}$$

  so:

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{pmatrix}$$

  (since once $i, j > n$, f maps each $\underline{e}_i$ to the 0 vector)

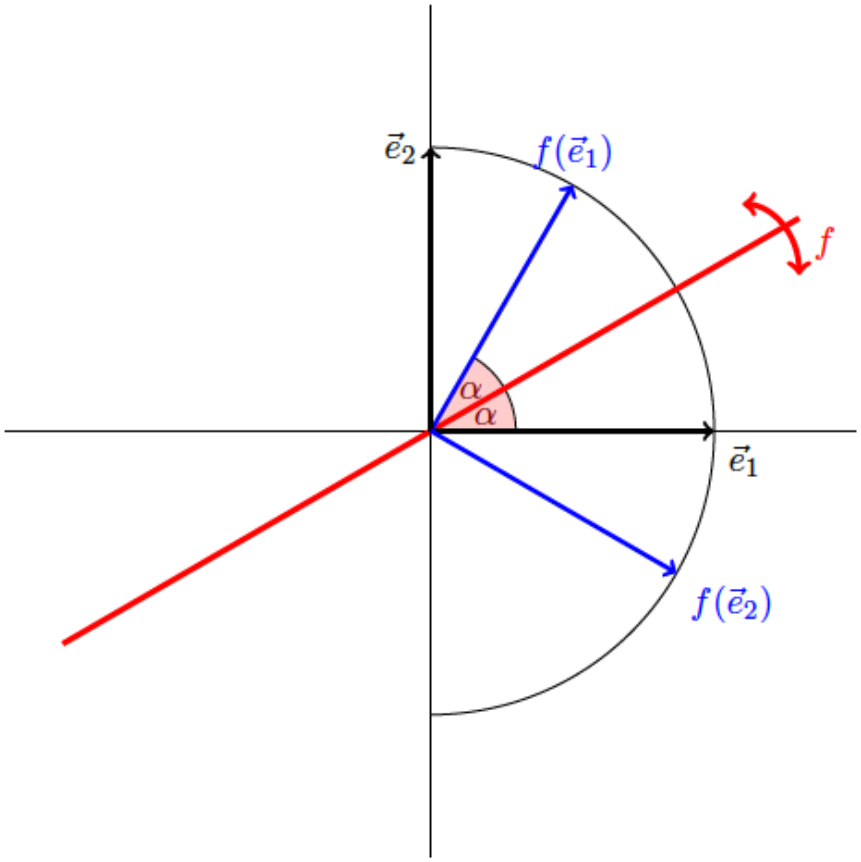
- if $g : (x, y) \to (y, x)$ permutes coordinates in $\mathbb{F}^2$, the corresponding matrix is:

$$[g] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

  More generally, we can define a **permutation matrix** by using a permutation $\pi \in S_n$, such that:

$$P_\pi(\underline{e}_i) = \underline{e}_{\pi(i)}$$

- if f is the reflection about the straight line making an agle α with the x-axis:



Then:

$$[f] = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix}$$

This follows from the fact that, if for example $\underline{e}_1 = (1, 0)$, we can think of such a reflection as a rotation by 2α radians about the origin. Then, the coordinates will just be the coordinates of the circle, after traversing 2α rad. These can be derived, by using a right angle triangle, with unit hypotenuse, and angles $2\alpha, \frac{\pi}{2}, \frac{\pi}{2} - 2\alpha$ Hence:

$$f(\underline{e}_1) = (\cos(2\alpha), \sin(2\alpha))$$

For $\underline{e}_2$, we just rotate in the oppositve direction, and using an angle of $\pi - 2\alpha$:

$$f(\underline{e}_2) = (\cos(\pi - 2\alpha), -\sin(\pi - 2\alpha)) = (\sin(2\alpha), -\cos(2\alpha))$$

3.2 Theorem: Composition of Linear Mappings and Products of Matrices)

- **How do we define multiplication of matrices?**
  - 2 matrices A, B can be multiplied to give a product $A \circ B = AB$ if:

$$A \in Mat(n \times m; \mathbb{F})$$

$$B \in Mat(m \times l; \mathbb{F})$$

– their product $AB \in Mat(n \times l; \mathbb{F})$ is given by:

$$AB_{ik} = \sum_{j=1}^{m} A_{ij} B_{jl}, \qquad i \in [1, n], k \in [1, l]$$

– in other words A_{ik} is given by taking the **dot product** of the ith row of A, and the kth column of B

- **Is matrix multiplication a mapping?**
  – yes, of the form:
  $$Mat(n \times m; \mathbb{F}) \times Mat(m \times l; \mathbb{F}) \to Mat(n \times l; \mathbb{F})$$
  via:
  $$(A, B) \to AB$$

Consider the homomorphisms:

$$g : \mathbb{F}^l \to \mathbb{F}^m$$

$$f : \mathbb{F}^m \to \mathbb{F}^n$$

*Then, the **representing matrix** of $f \circ g$ is the product of the representing matrices of f and g. In other words:*

$$[f \circ g] = [f] \circ [g]$$

[Theorem 2.1.8]

Proof. Lets define the matrices, and the bases of the spaces:

$$A = [f], \qquad \mathbb{F}^m = \langle \{\underline{a}_i \mid i \in [1, m]\} \rangle$$

$$B = [g], \qquad \mathbb{F}^l = \langle \{\underline{b}_j \mid j \in [1, l]\} \rangle$$

$$C = [f] \circ [g], \qquad \mathbb{F}^n = \{\underline{c}_k \mid k \in [1, n]\}$$

Then, by how we define the bijection from homomorphisms to vectors, we know that:

$$f(\underline{a}_i) = A_{*i} = \sum_{k=1}^{n} A_{ki} \underline{c}_k$$

$$g(\underline{b}_j) = B_{*j} = \sum_{i=1}^{m} B_{ij} \underline{a}_i$$

What this is saying is that, for example, for the ith column of $[f]$ (denoted A_{*i}), we are taking an element from the basis of $\mathbb{F}^m$, and mapping it to an element of $\mathbb{F}^n$, by using a linear combination of basis vectors

of $\mathbb{F}^n$. The "coordinates" of the element in $\mathbb{F}^n$ to which we map are precisely the coefficients of this linear combination, which is given by the matrix entries A_{ki}.

Using this, we can write:

$$\begin{aligned}
(f \circ g)(\underline{b}_j) &= f\left(\sum_{i=1}^{m} B_{ij}\underline{a}_i\right) \\
&= \sum_{i=1}^{m} B_{ij} f(\underline{a}_i) \\
&= \sum_{i=1}^{m} B_{ij} \sum_{k=1}^{n} A_{ki}\underline{c}_k \\
&= \sum_{i=1}^{m}\sum_{k=1}^{n} (A_{ki}B_{ij})\underline{c}_k \\
&= \sum_{k=1}^{n} \left(\sum_{i=1}^{m} A_{ki}B_{ij}\right)\underline{c}_k \qquad \textit{(we can switch the sums, since they are finite)} \\
&= \sum_{k=1}^{n} C_{kj}\underline{c}_k \qquad \textit{(by definition of matrix product, } C_{kj} = \sum_{i=1}^{m} A_{ki}B_{ij}\textit{)}
\end{aligned}$$

Notice, this is just giving the jth column of the matrix C, defined by $[f] \circ [g]$, so it follows that:

$$[f \circ g] = [f] \circ [g]$$

□

3.3 Proposition: Calculating With Matrices

Define the following matrices:

- $A, A' \in Mat(n \times m; \mathbb{F})$
- $B, B' \in Mat(m \times l; \mathbb{F})$
- $C \in Mat(l \times k; \mathbb{F})$
- I_m, *the* $m \times m$ *identity matrix*

Then, the following hold for matrix multiplication:

1. $$(A + A')B = AB + A'B$$
2. $$A(B + B') = AB + AB'$$
3. $$I_m B = B$$
4. $$AI_m = A$$
5. $$(AB)C = A(BC)$$

[Proposition 2.1.9]

Proof. Whilst these can be proven from first principles (i.e using all the summation business), it is more elegant to use the bijection between homomorphisms and matrices, alongside the fact that $[f \circ g] = [f] \circ [g]$, and the distributive and associative property of functions.

Let:

- $[f] = A$
- $[f'] = A'$
- $[g] = B$
- $[g'] = B'$
- $[h] = C$
- $[id_{\mathbb{F}^m}] = I_m$

1. $[(f + f')] \circ [g] = [(f + f') \circ g] = [f \circ g + f' \circ g] = [f \circ g] + [f' \circ g]$

2. use similar logic as above
3. $[id_{\mathbb{F}^m}] \circ [g] = [id_{\mathbb{F}^m} \circ g] = [g]$
4. use similar logic as above
5. $([f] \circ [g]) \circ [h] = [f \circ g] \circ [h] = [(f \circ g) \circ h] = [f \circ (g \circ h)] = f \circ [g \circ h] = f \circ ([g] \circ [h])$

□

3.4 Remark: Assigning Linear Mappings to Matrices

Above we discussed how the mapping:

$$M : Hom_{\mathbb{F}}(\mathbb{F}^m, \mathbb{F}^n) \to Mat(n \times m; \mathbb{F})$$

produces a ***representing matrix*** *from a homomorphism.*
We can easily define the inverse transformation, by thinking of applying a matrix $A \in Mat(n \times m; \mathbb{F})$ *to an element of* $\mathbb{F}^m$*, and producing an element of* $\mathbb{F}^n$*:*

$$(A\circ) : \mathbb{F}^m \to \mathbb{F}^n$$

Indeed, this is our old intuition of "multiplying a vector by a matrix":

$$A\underline{x} = \underline{b}, \underline{x} \in \mathbb{F}^m, \underline{b} \in \mathbb{F}^n$$

(here we are taking liberty, since technically for this to be applicable we should use $Mat(m \times 1; \mathbb{F})$ *instead of* $\mathbb{F}^m$*, and* $Mat(n \times 1; \mathbb{F}$ *instead of* $\mathbb{F}^n$*). This then allows us to define the inverse of* M*:*

$$M^{-1} : Mat(n \times m; \mathbb{F}) \to Hom_{\mathbb{F}}(\mathbb{F}^m, \mathbb{F}^n)$$

via:

$$A \to (A\circ)$$

In other words, to each matrix, we associate a linear map. [Remark 2.1.10]

3.4.1 Exercises (TODO)

1. **Let** $f : \mathbb{R}^2 \to \mathbb{R}^2$ **be the reflection:**

   $$(x, y) \to (x, -y)$$

   Show that:

   $$\{g \mid g \in Hom_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2), f \circ g = g \circ f\}$$

   is a subspace of $Hom_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R}^2)$**, and give a basis of this subspace.**

2. **Define the *transpose* of a matrix via:**

   $$(A^T)_{ij} = A_{ji}$$

   Show that:

   - $(A^T)^T = A$
   - $(AB)^T = B^T A^T$

4 Properties of Matrices

4.1 Invertibility of a Matrix

- **When is a matrix invertible?**
  - when it has **both** a **left** and **right** inverse
  - in other words, A is invertible **if and only if** $\exists B, C$ such that:

$$AB = CA = \mathbb{I}$$

  - notice, a matrix may have a left **or** a right inverse, but said matrix won't be invertible. For example:

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mathbb{I}$$

    but $\begin{pmatrix} 1 & 0 \end{pmatrix}$ doesn't have a left inverse, so it isn't invertible

- **How does the isomorphism defining the matrix define its invertibility?**
  - consider a matrix A, defined as a homomorphism:

$$a : \mathbb{F}^n \to \mathbb{F}^m$$

  - by (3.2), we can "translate" the requirements for matrix invertibility to be in terms of function composition
  - hence, A is invertible if and only if we can find b, c such that:

$$a \circ b = id_{\mathbb{F}^k}, \qquad b : \mathbb{F}^k \to F^n$$

$$c \circ a = id_{\mathbb{F}^n}, \qquad c : \mathbb{F}^m \to F^l$$

  - notice, the identity function is an isomorphism, so $a \circ b$ and $c \circ a$ must be isomorphisms. Hence:
    - $\mathbb{F}^m$ is isomorphic to $\mathbb{F}^k$, and since vector spaces are characterised by their dimension [(1.3)], we must have $m = k$
    - $\mathbb{F}^n$ is isomorphic to $\mathbb{F}^l$, and since vector spaces are characterised by their dimension [(1.3)], we must have $n = l$
  - notice that since the compositions are isomorphism, we must then again have that $n = m$, so in particular, if A is invertible, $n = m$ - A must be a square matrix

- **Are all matrices invertible?**
  - only **square** matrices are invertible
  - this is summarised by the following set of equivalences:
    1. There exists a square matrix B such that:

$$BA = \mathbb{I}$$

    2. There exists a square matrix C such that:

$$AC = \mathbb{I}$$

    3. The square matrix A is invertible

Proof. To show equivalence, we show that 3 implies 1 and 2, and that 1 and 2 each imply 3.

* $\Big(\textcircled{3} \implies \textcircled{1}, \textcircled{2}\Big)$: this is from the definition of matrix invertibility
* $\Big(\textcircled{1} \implies \textcircled{3}\Big)$: lets assume that A has a left inverse. Then, it must be the case that a (where $a : \mathbb{F}^n \to \mathbb{F}^n$) has a left inverse under function composition, so a is injective. Then, $ker(a) = \{0\}$, so by rank nullity theorem:

  $$dim(\mathbb{F}^n) = dim(im(a))$$

  Since $im(a)$ is a vector subspace of $\mathbb{F}^n$, it follows by Remark 1.6.9 that $im(a) = \mathbb{F}^n$, so in particular a is surjective. Hence, a must be an isomorphism, so A must be invertible.
* $\Big(\textcircled{2} \implies \textcircled{3}\Big)$: lets assume that A has a right inverse. Then, it must be the case that a (where $a : \mathbb{F}^n \to \mathbb{F}^n$) has a right inverse under function composition, so a is surjective. By rank nullity theorem:

  $$dim(\mathbb{F}^n) = dim(im(a)) + dim(ker(a))$$

  Since $dim(im(a)) = dim(\mathbb{F}^n)$, we must have that $dim(ker(a)) = 0$, so it follows that a is also injective. Hence, a must be an isomorphism, so A must be invertible.

□

- **How do we denote the inverse matrix?**
  - if a^{-1} is the inverse of the mapping a defining a matrix A, we can denote the inverse of A via:

    $$[a^{-1}] = A^{-1}$$

4.1.1 Exercises

1. **Show that the *general linear group of* 2×2, $GL(2; \mathbb{F}_2)$, and S_3 are isomorphic.**
2. **Is the group $GL(1; \mathbb{F}_p)$ abelian? What is its order?**
3. **What is the center of $GL(n; \mathbb{F})$? What is the order of the center of $GL(n; \mathbb{F}_p)$?**

4.2 Elementary Row Operations as Matrices

- **What is the basis matrix?**
  - a matrix E_{ij}, such that:

    $$(E_{ij})_{lm} = \begin{cases} 1, & l = i, m = j \\ 0, & otherwise \end{cases}$$

- **How can row addition be represented as a matrix?**
  - we use the matrix:

    $$\mathbb{I} + \lambda E_{ij}$$

    to add λ times the jth row to the ith row

- as an example, if we want to add the 2 times the 3rd row to the 1st column of:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix}$$

we apply the above matrix:

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1+2(3) & 2+2(2) & 3+2(1) \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix}$$

- from this we see that E_{ij} does nothing but "pick up" the jth row of the matrix to which it is applied on, and puts it in the ith row of the resulting matrix
- notice, $\mathbb{I} + \lambda E_{ij}$ is **invertible** (since $(\mathbb{I} + \lambda E_{ij})(\mathbb{I} - \lambda E_{ij}) = \mathbb{I}$) so applying it is reversible, thus preserving the solution

- **How can row swap be represented as a matrix?**
  - define the matrix P_{ij} which swaps row i with row j
  - P_{ij} will just be the identity matrix, with its own ith and jth rows swapped
  - this can be thought as a homomorphism $\mathbb{F}^m \to \mathbb{F}^m$
  - again, $P_{ij}P_{ij} = \mathbb{I}$, so the operation is reversible

4.3 Theorem: Elementary Matrices as Building Blocks

- **What is an elementary matrix?**
  - a (square) matrix differing from $\mathbb{I}$ in **at most** one entry
- **When are elementary matrices invertible?**
  - so long as the change to I doesn't change a 1 by a 0, the elementary matrix is invertible

---

*Every **square** matrix with entries in a field can be written as a **product** of **elementary matrices**. [Theorem 2.2.3]*

(This here is a nice work through; in this proof, they choose to apply right multiplications, when, as far as I am concerned, applying left EROs should be sufficient)

Proof. Notice, we can represent any permutation matrix as a **product** of elementary matrices, namely:

$$P_{ij} = diag(1, \dots, 1, -1, 1, \dots, 1)(\mathbb{I} + E_{ij})(\mathbb{I} - E_{ji})(\mathbb{I} + E_{ij})$$

where the -1 in $diag(1, \dots, 1, -1, 1, \dots, 1)$ is at the jth diagonal entry.

Instead of getting all theoretical, lets use an example, with a 3×3 matrix. For example, if we want to swap the first and second rows:

$$\begin{aligned}
&\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

If we then apply this to our matrix above:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

The second thing to notice is that if an elementary matrix is invertible, then its inverse will also be an elementary matrix. We have already seen that for example $(\mathbb{I} + \lambda E_{ij})(\mathbb{I} - \lambda E_{ij}) = \mathbb{I}$. Moreover, a product of invertible matrices is also invertible (if A, B are invrtible, the inverse of AB is $B^{-1}A^{-1}$. Notice, any elementary matrix can be constructed by chaining row swaps and row additions. Row swaps are invertible (since they are a product of invertible matrices). Hence, any elementary matrix must be invertible.

With all this in mind, we can give the proof. Take an arbitrary matrix A. Then:

1. We can find invertible, elementary matrices $S_1, S_2, \dots, S_t$, such that:

   $$S_t \dots S_1 A$$

   is in row echelon form (this is just Gaussian elimination through EROs)

2. We can then reduce the columns, by applying a set of invertible, elementary matrices $T_1, \dots, T_s$ to the **right** of the resulting matrix:

   $$S_t \dots S_1 A T_1 \dots T_s$$

3. by reducing rows and columns, we can obtain a diagonal matrix by simply using elementary matrices:

   $$D = diag(1, \dots, 1, 0, \dots, 0)$$

4. D can be expressed using elementary matrices (apparently non-invertible) as well, so overall:

$$S_t \dots S_1 A T_1 \dots T_s = D_1 \dots D_k$$

Hence, we can express A as a product of elementary matrices:

$$A = S_1^{-1} \dots S_t^{-1} D_1 \dots D_k T_s^{-1} \dots T_1^{-1}$$

□

4.4 The Smith Normal Form

- **When is a matrix in Smith normal Form?**
  - a matrix is in **Smith Normal Form** if all the non-zero entries are along the main diagonal, and the non-zero entries are consecutive 1s

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

  - alternatively, we can think of SNF as being an identity matrix, padded with 0s

4.5 Theorem: Transforming a Matrix into SNF

> *Given a matrix $A \in Mat(n \times m; \mathbb{F}$, there exist **invertible** matrices P, Q such that PAQ is in **Smith Normal Form**. [Theorem 2.2.5]*

Proof. We can find invertible elementary matrices $S_1, \dots, S_t$ such that

$$S_1 \dots S_t A$$

is in row echelon form (these matrices perform eleemntary row operations to achieve this)

Once in row echeleon form, we can apply more matrices $T_1, \dots, T_s$, such that:

$$S_1 \dots S_t A T_1 \dots T_s$$

is brought to Smith Normal Form.

We can then just define:

$$P = S_1 \dots S_t$$
$$Q = T_1 \dots T_s$$

□

4.6 Matrix Rank

- **What is the column rank of a matrix?**
  - the dimension of the space generated by the columns of a matrix
- **What is the row rank of a matrix?**
  - the dimension of the space generated by the rows of a matrix
- **When does a matrix have full rank?**
  - when both the column and row ranks are equal, and they are equal to the smallest number, out of the number of rows or columns

4.7 Theorem: Column and Row Rank

> *The **column rank** and the **row rank** of a matrix are **equal**. [Theorem 2.2.8]*

Proof. Notice, applying elementary row operations won't change the column rank of a matrix (and if we transpose it, EROs won't change the row rank). This is because elementary row operations are linear combinations of the columns/row vectors, so this won't change the number of linear independent vectors which are part of the matrix rows/columns. In particular, this means that any matrix A has the same column and row rank as its corresponding SNF matrix. But a matrix in SNF has the same row and column ranks (since this is given by the number of LiD vectors, and these are given by the number of 1s present along the rows/columns, and this number is the same). □

4.7.1 Exercises (TODO)

1. **Find a 3×3 matrix with all entries non-zero, but with rank 2.**

4.8 Inverting Matrices

- **How can you invert a matrix?**
  - the idea is to write an augmented matrix of the form:

$$\left(A \;\middle|\; \mathbb{I}\right)$$

  - then, apply EROs, such that we obtain:

$$\left(\mathbb{I} \;\middle|\; (S_1 S_2 \dots S_t)\mathbb{I}\right)$$

  - then, the inverse will be:

$$A^{-1} = S_1 S_2 \dots S_t$$

4.8.1 Exercises (TODO)

1. **Show that:**
$$rank(A+B) \leq rank(A) + rank(B)$$

2. **Show that:**
$$rank(AB) \leq min\{rank(A), rank(B)\}$$

3. **Let $r = rank(B)$, where B is such that:**
$$B \in Mat(n \times r; \mathbb{F})$$
Show that there exists $E \in Mat(r \times n; \mathbb{F}$ such that:
$$EB = 1_r$$
Similarly let $C \in Mat(r \times n; \mathbb{F})$ where $rank(C) = r$. Then there exists $H \in Mat(n \times r; \mathbb{F})$ such that $CH = 1_r$.

4. **Let $A \in Mat(m \times n; \mathbb{F})$ have rank r. Show that $A = BC$, where:**
   - $B \in Mat(m \times r; \mathbb{F})$
   - $C \in Mat(r \times n; \mathbb{F})$

   and $rank(B) = rank(C) = r$.
   1. **Is this decomposition unique?**
   2. **Hence, or otherwise, show that A can be decomposed as a sume of r rank 1 matrices.**

5 Workshop

1. **True or False. If $f : F^4 \to F^2$ is a linear map such that:**
$$\ker(f) = \{(x_1, x_2, x_3, x_4)^T \mid x_1 = 2x_3, x_2 = 4x_4\}$$
then f is surjective.

   Notice, the kernel is a 2 dimensional vector subspace, since we can write it as:
$$ker(f) = \left\{ \begin{pmatrix} 2s \\ 4t \\ s \\ t \end{pmatrix} \;\middle|\; s, t \in F \right\}$$
   so clearly it has 2 basis vectors.

   The Rank-Nullity Theorem tells us that:
$$\dim(F^4) = \dim(ker(f)) + \dim(im(f))) \implies \dim(im(f)) = 2$$
   But notice, F^2 is a two-dimensional vector space, and $im(f)$ is also a 2-dimensional vector space, which means that $F^2 = im(a)$. Hence, it follows that f is surjective.

2. **Let:**

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix} \qquad \pi' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 2 & 1 \end{pmatrix} \qquad \pi'' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 3 & 2 \end{pmatrix}$$

Define P_π as the linear mapping:

$$P_\pi(\underline{e}_i) = \underline{e}_{\pi(i)}$$

and define similar mappings $P_{\pi'}, P_{\pi''}$.

(a) **Write the permutations above as a product of disjoint cycles.**

$$\pi = (1)(23)(45)$$
$$\pi' = (1425)(3)$$
$$\pi'' = (152)(34)$$

(b) **Determine the representing matrices of $P_\pi, P_{\pi'}, P_{\pi''}$**

We obtain these by permuting the standard basis vectors, according to each of the permutations:

$$[P_\pi] = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$[P'_\pi] = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$[P''_\pi] = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(c) **Compute $\pi'\pi, \pi''\pi'$ and $\pi''\pi'\pi$**

$$\pi'\pi = (1425)(23)(45) = (14)(235)$$
$$\pi''\pi' = (152)(34)(1425) = (134)$$
$$\pi''\pi'\pi = (152)(34)(1425)(23)(45) = (13245)$$

(d) **Express the permutations as a product of transpositions.**

$$\pi = (1)(23)(45) = (23)(35)$$
$$\pi' = (1425)(3) = (15)(12)(14)$$
$$\pi'' = (152)(34) = (12)(15)(34)$$

3. **Let the linear map $a : F^3 \to F^2$ be given by the matrix:**

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

Determine $ker(a)$, and hence verify the Rank-Nullity Theorem for a.

Let:

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \in ker(a)$$

Then:

$$A\underline{v} = \underline{0}$$

implies that:

$$v_1 + 2v_2 + 3v_3 = 0 \qquad v_1 + v_2 + v_3 = 0$$

which is true if and only if:

$$v_2 + 2v_3 = 0 \implies v_2 = -2v_3$$

Let $s \in F$, such that $v_3 = s$. Then, $v_2 = -2s$ and:

$$v_1 + 2v_2 + 3v_3 = 0 \iff v_1 - 4s + 3s = 0 \implies v_1 = s$$

Hence, the kernel is a one dimension subspace:

$$\ker(a) = \left\{ s \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \;\middle|\; s \in F \right\}$$

The rank is the dimension of the image, which is equal to the number of linearly independent rows in A. We can clearly see that the rows are not multiples, so $im(a) = 2$. Indeed, $dim(F^3) = 3$ and $dim(ker(a)) + dim(im(a)) = 1 + 2 = 3$, as expected by Rank-Nullity.

4. **Let $F \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$, and let $a_0, \ldots, a_n \in F$ be distinct. For $k \in [0, n]$ define:**

$$p_k(x) = \prod_{i=0, i\neq k}^{n} \left(\frac{x - a_i}{a_k - a_i} \right) \in F[x]_{\leq n}$$

(a) **Show that $p_k(a_i) = \delta_{ik}$**

If $j = k$:

$$p_k(a_j) = \prod_{i=0, i\neq k}^{n} \left(\frac{a_j - a_i}{a_k - a_i} \right) = \prod_{i=0, i\neq k}^{n} \left(\frac{a_k - a_i}{a_k - a_i} \right)$$

since $k \neq i$, it follows that this is a product of 1s, so $p_k(a_j) = 1$.
If $j \neq k$:

$$p_k(a_j) = \prod_{i=0, i\neq k}^{n} \left(\frac{a_j - a_i}{a_k - a_i}\right)$$

notice, $i \in [0, n]$, so in particular, $i = j$ at some point, so the product has a 0 term, and so, $p_k(a_j) = 0$.
Thus, $p_k(a_j) = \delta_{jk}$.

(b) **Show that $B = \{p_0(x), \ldots, p_n(x)\}$ forms a basis of $F[x]_{\leq n}$.**

We first show that this is a linearly independent set. Consider:

$$\sum_{i=0}^{n} \lambda_i p_i(x) = 0$$

Evaluating at a_j, we get that:

$$\lambda_j = 0$$

Since $\sum_{i=0}^{n} \lambda_i p_i(x) = 0$ must hold for **any** x, it holds for $x = a_j, \forall j \in [0, n]$ so this means that $\forall j \in [0, n], \lambda_j = 0$, so the set is linearly independent.
Now, notice that the set $\{p_0(x), \ldots, p_n(x)\}$ has $n + 1$ elements, and that $dim(F[x]_{\leq n}) = n + 1$, so it must be the case that the set is indeed a basis.

(c) **Show that an arbitrary polynomial $q(x) \in F[x]_{\leq n}$ may be written as:**

$$q(x) = \sum_{k=0}^{n} q(a_k) p_k(x)$$

Let q be an nth degree polynomial. Since B is a basis, we can write:

$$q(x) = \sum_{k=0}^{n} \lambda_k p_k(x)$$

But if we evaluate at a_i:

$$q(a_i) = \sum_{k=0}^{n} \lambda_k p_k(a_i) = \lambda_i$$

Thus, any polynomial q can be written (uniquely) as:

$$q(x) = \sum_{k=0}^{n} q(a_k) p_k(x)$$

(d) **Deduce that for any $n + 1$ points:**

$$((a_0, c_0), \ldots, (a_n, c_n))$$

with distinct first coordinate, there exists a unique polynomial of degree n through them. This is known as the Lagrange Interpolation Formula.

By the above, if we set $q(a_i) = c_i$, then the polynomial:

$$q(x) = \sum_{k=0}^{n} q(a_k) p_k(x)$$

goes through each of the points, and since the a_i are distinct, this representation is unique.

(e) **Hence, show that:**

$$\sum_{k=0}^{n} \frac{a_k^l}{\prod_{i=0,i\neq k}^{n}(a_k - a_i)} = \begin{cases} 0, & l < n \\ 1, & l = n \end{cases}$$

Consider the polynomial $q(x) = x^l$. By Lagrange Interpolation:

$$q(x) = x^l = \sum_{k=0}^{n} a_k^l p_k(x) = \sum_{k=0}^{n} a_k^l \prod_{i=0,i\neq k}^{n} \left(\frac{x - a_i}{a_k - a_i} \right)$$

Now, consider the coefficient of x^n in the expansion. This is:

$$\sum_{k=0}^{n} \frac{a_k^l}{\prod_{i=0,i\neq k}^{n}(a_k - a_i)}$$

But notice, since we $q(x) = x^l$, whenever $l = n$, we must have:

$$\sum_{k=0}^{n} \frac{a_k^l}{\prod_{i=0,i\neq k}^{n}(a_k - a_i)} = 1$$

and if $l < n$ then:

$$\sum_{k=0}^{n} \frac{a_k^l}{\prod_{i=0,i\neq k}^{n}(a_k - a_i)} = 0$$

as required.