Linear Algebra

Notes by Finley Cooper

27th October 2025

Contents

1 Vector Spaces

1.1 Definitions

For this lecture course, $\mathbb{F}$ will always be field.

Definition. (Vector Space) A $\mathbb{F}$-*vector space* (or a vector space over $\mathbb{F}$) is an abelian group $(V, +, \mathbf{0})$ equipped with a function

$$\begin{aligned} \mathbb{F} \times V &\to V \\ (\lambda, v) &\to v \end{aligned}$$

which we call scalar multiplication such that $\forall v, w \in V, \forall \lambda, \mu \in \mathbb{F}$
(i) $(\lambda + \mu)v = \lambda v + \mu v$
(ii) $\lambda(v + w) = \lambda v + \lambda w$
(iii) $\lambda(\mu v) = (\lambda\mu)v$
(iv) $1 \cdot v = v \cdot 1 = v$

Remember that $\mathbf{0}$ and 0 are not the same thing. 0 is an element in the field $\mathbb{F}$ and $\mathbf{0}$ is the additive identity in V.

For an example consider $\mathbb{F}^n$ n-dimensional column vectors with entries in $\mathbb{F}$. We also have the example of a vector space $\mathbb{C}^n$ which is a complex vector space, but also a real vector space (taking either $\mathbb{C}$ or $\mathbb{R}$ as the underlying scalar field).

We also can see that $M_{m\times n}(\mathbb{F})$ form a vector space with m rows and n columns.
For any non-empty set X, we denote $\mathbb{F}^X$ as the space of functions from X to $\mathbb{F}$ equipped with operations such that:

$$\begin{aligned} f + g \text{ is given by } (f + g)(x) &= f(x) + g(x) \\ \lambda f \text{ is given by } (\lambda f)(x) &= \lambda f(x) \end{aligned}$$

Proposition. For all $v \in V$ we have that $0 \cdot v = \mathbf{0}$ and $(-1) \cdot v = -v$ where $-v$ denotes the additive inverse of v.

Proof. Trivial.

Definition. (Subspace) A *subspace* of a $\mathbb{F}$-vector space V is a subset $U \subseteq V$ which is a $\mathbb{F}$-vector space itself under the same operations as V. Equivalently, $(U, +)$ is a subgroup of $(V, +)$ and $\forall \lambda \in \mathbb{F}$, $\forall u \in U$ we have that $\lambda u \in U$.

Remark. Axioms (i)-(iv) are always automatically inherited into all subspaces.

Proposition. (Subspace test) Let V be a $\mathbb{F}$-vector space and $U \subseteq V$ then U is a subspace of V if and only if,
(i) U is nonempty.
(ii) $\forall \lambda \in \mathbb{F}$ and $\forall u, w \in U$ we have that $u + \lambda w \in U$.

Proof. If U is a subspace then U satisfies (i) and (ii) since it contains 0 and is closed. Conversely suppose that $U \subseteq V$ satisfies (i) and (ii). Taking $\lambda = -1$ so $\forall u, w \in V$, $u - w \in U$ hence $(U, +)$ is a subgroup of $(V, +)$ by the subgroup test. Finally taking $u = \mathbf{0}$ so we have that $\forall w \in U, \forall \lambda \in \mathbb{F}$ we have that $\lambda w \in U$. So U is a subspace of V. □

We notate U by $U \leq V$.

For some examples

(i)
$$\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x + y + z = t \right\} \subseteq \mathbb{R}^3,$$
for fixed $t \in \mathbb{R}$ is a subspace of $\mathbb{R}^3$ iff $t = 0$.

(ii) Take $\mathbb{R}^\mathbb{R}$ as all the functions from $\mathbb{R}$ to $\mathbb{R}$ then the set of continuous functions is a subspace.

(iii) Also we have that $C^\infty(\mathbb{R})$, the set of infintely differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ is a subspace of $\mathbb{R}^\mathbb{R}$ and the subspace of continuous functions.

(iv) A further subspace of all of those subspaces is the set of polynomial functions.

Lemma. For $U, W \leq V$ we have that $U \cap W \leq V$.

Proof. We'll use the subspace test. Both U, W are subspaces so they contain $\mathbf{0}$ hence $\mathbf{0} \in U \cap W$ so $U \cap W$ is nonempty. Secondly take $x, y \in U \cap W$ with $\lambda \in \mathbb{F}$. Then $U \leq V$ and $x, y \in U$ so $x + \lambda y \in U$. Similarly with W so $x + \lambda y \in W$ hence we have that $x + \lambda y \in U \cap W$ hence $U \cap W \leq V$ □

Remark. This does not apply for subspaces, in fact from IA Groups, we know it doesn't even hold for the underlying abelian group.

Definition. (Subspace sum) For $U, W \leq V$, the *subspace sum* of U, W is
$$U + W = \{u + w : u \in U, w \in W\}.$$

Lemma. If $U, W \leq V$ then $U + W \leq V$.

Proof. Simple application of the subspace test.

Remark. $U + W$ is the smallest subgroup of U, W in terms of inclusion, i.e. if K is such that $U \subseteq K$ and $W \subseteq K$ then $U + W \subseteq K$.

1.2 Linear maps, isomorphisms, and quotients

Definition. (Linear map) For V, W $\mathbb{F}$-vector spaces. A *linear map* from V to W is a group homomorphism, φ, from $(V, +)$ to $(W, +)$ such that $\forall v \in V$
$$\varphi(\lambda v) = \lambda \varphi(v)$$

Equivalently to show any function $\alpha : V \to W$ is a linear map we just need to show that $\forall u, w \in V$, $\forall \lambda \in \mathbb{F}$ we have

$$\alpha(u + \lambda w) = \alpha(u) + \lambda\alpha(w).$$

For some examples of linear maps

(i) $V = \mathbb{F}^n, W = \mathbb{F}^m$ $A \in M_{m\times n}(\mathbb{F})$. Then let $\alpha : V \to W$ be given by $\alpha(v) = Av$. Then α is linear.

(ii) $\alpha : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ defined by taking the derivative.

(iii) $\alpha : C(\mathbb{R}) \to \mathbb{R}$ defined by taking the integral from 0 to 1.

(iv) X any nonempty set, $x_0 \in X$,

$$\begin{aligned} \alpha : \mathbb{F}^X &\to \mathbb{F} \\ f &\to f(x_0) \end{aligned}$$

(v) For any V, W the identity mapping from V to V is linear and so is the zero map from V to W.

(vi) The composition of two linear maps is linear.

(vii) For a non-example squaring in $\mathbb{R}$ is not linear. Similiarly adding constants is not linear, since linear maps preserve the zero vector.

Definition. (Isomorphism) A linear map $\alpha : V \to W$ is an *isomorphism* if it is bijective. We say that V and W are isomorphic, if there exists an isomorphism from $V \to W$ and denote this by $V \cong W$.

An example is the vector space $V = \mathbb{F}^4$ and $W = M_{2\times 2}(\mathbb{F})$ we can define the map

$$\begin{aligned} \alpha : V &\to W \\ \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} &\to \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{aligned}$$

Then α is an isomorphism.

Proposition. If $\alpha : V \to W$ is an isomorphism then $\alpha^{-1} : W \to V$ is also an isomorphism.

Proof. Clearly α^{-1} is a bijection. We need to prove that α^{-1} is linear. Take $w_1, w_2 \in W$ and $\lambda \in \mathbb{F}$. So we can write $w_i = \alpha(v_i)$ for $i = 1, 2$. Then

$$\alpha^{-1}(w_1 + \lambda w_2) = \alpha^{-1}(\alpha(v_1) + \lambda\alpha(v_2)) = \alpha^{-1}(\alpha(v_1 + \lambda v_2)) = v_1 + \lambda v_2 = \alpha^{-1}(w_1) + \lambda\alpha^{-1}(w_2)$$

. Hence α^{-1} is linear, so α^{-1} is an isomorphism. $\square$

Definition. (Kernal) Let V, W be $\mathbb{F}$-vector spaces. Then the *kernel* of the linear map $\alpha : V \to W$ is
$$\ker(\alpha) = \{v \in V : \alpha(v) = \mathbf{0}_W\} \subseteq V$$

Definition. (Image) Let V, W be $\mathbb{F}$-vector spaces. Then the *image* of a linear map $\alpha : V \to W$ is
$$\operatorname{im}(\alpha) = \{\alpha(v) : v \in V\} \subseteq W$$

Lemma. For a linear map $\alpha : V \to W$ the following hold.

(i) $\ker \alpha \leq V$ and $\operatorname{im} \alpha \leq W$

(ii) α is surjective if and only if $\operatorname{im} \alpha = W$

(iii) α is injective if and only if $\ker \alpha = \{\mathbf{0}_V\}$

Proof. $\mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V$, so applying α to both sides any using the fact that α is linear gives that $\alpha(\mathbf{0}_V) = \mathbf{0}_W$. So $\ker \alpha$ is nonempty. The rest of the proof is a simple application of the subspace test.
The second statement is immediate from the definition.
For the final statement suppose α injective. Suppose $v \in \ker \alpha$. Then $\alpha(v) = \mathbf{0}_W = \alpha(\mathbf{0}_w)$ so $v = \mathbf{0}_V$ by injectivity. Hence $\ker \alpha$ is trivial. Conversely suppose that $\ker \alpha = \{0_V\}$ Let $u, v \in V$ and suppose that $\alpha(u) = \alpha(v)$. The $\alpha(u - v) = \mathbf{0}_W$, so $u - v \in \ker \alpha$, so $u = v$. $\square$

For V a $\mathbb{F}$-vector space, $W \leq V$ write
$$\frac{V}{W} = \{v + W : v \in V\}$$
as the left cosets of W in V. Recall that two cosets $v + V$ and $u + W$ are the same coset if and only if $v - u \in W$.

Proposition. V/W is an $\mathbb{F}$-vector space under operations
$$(u + W) + (v + W) = (u + v) + W$$
$$\lambda(v + W) = (\lambda v) + W$$
We call V/W the quotient space of V by W.

Proof. The proof is long and requires a lot of vector space axioms so we'll just sketch out the proof.
We check that operations are well-defined, so for $u, \overline{u}, v, \overline{v} \in V$ and $\lambda \in \mathbb{F}$ if
$$u + W = \overline{u} + W, \quad v + W = \overline{v} + W$$
then
$$(u + v) + W = (\overline{u} + \overline{w}) + W$$
and
$$(\lambda u) + W = (\lambda \overline{u}) + W$$
The vector space axioms are inherited from V. $\square$

Proposition. (Quotient map) The function $\pi_W : V \to \frac{V}{W}$ called a *quotient map* is given by

$$\pi_W(v) = v + W$$

is a well-defined, surjective, linear map with $\ker \pi_W = W$.

Proof. Surjectivity is clear. For linearity let $u, v \in V$ and $\lambda \in \mathbb{F}$. Then

$$\begin{aligned}\pi_W(u + \lambda v) &= (u + \lambda v) + W \\ &= (u + W) + (\lambda v + W) \\ &= (u + W) + \lambda(v + W) \\ &= \pi_W(u) + \lambda \pi_W(v)\end{aligned}$$

For $v \in V$, we have that $v \in \ker \pi_W \iff \pi_W(v) = \mathbf{0}_{V/W}$. So $v + W = \mathbf{0}_V + W$ so finally $v = v - \mathbf{0}_V \in W$. $\square$

Theorem. (First isomorphism theorem) Let V, W be $\mathbb{F}$-vector spaces and $\alpha : V \to W$ linear. Then there is an isomorphism

$$\overline{\alpha} : \frac{V}{\ker \alpha} \to \operatorname{im} \alpha$$

given by $\overline{\alpha}(v + \ker \alpha) = \alpha(v)$

Proof. For $u, v \in V$,

$$u + K = v = K \iff u - v \in K \quad \iff \alpha(u - v) = \mathbf{0}_W \iff \alpha(u) = \alpha(v) \quad \iff \overline{\alpha}(u + \ker \alpha) = \overline{\alpha}(v + \ker \alpha)$$

The forward direction shows that $\overline{\alpha}$ is well-defined, and the converse shows that $\overline{\alpha}$ is injective. For surjectivity given $w \in \operatorname{im} \alpha$, there exists some $v \in V$ s.t. $w = \alpha(v)$. Then $w = \overline{\alpha}(v + \ker \alpha)$. Finally for linearity given $u, v \in V$, $\lambda \in \mathbb{F}$,

$$\begin{aligned}\overline{\alpha}((u + \ker \alpha) + \lambda(v + \ker \alpha)) &= \overline{\alpha}((u + \lambda v) + \ker \alpha) \\ &= \alpha(u + \lambda v) \\ &= \alpha(u) + \lambda \alpha(v) \\ &= \overline{\alpha}(u + \ker \alpha) + \lambda \overline{\alpha}(v + \ker \alpha)\end{aligned}$$

So $\overline{\alpha}$ is linear hence is an isomorphism $\square$

1.3 Basis

Definition. (Span) Let V be a $\mathbb{F}$-vector space. Then the *span* of some subset $S \subseteq V$ is

$$\langle S \rangle = \left\{ \sum_{s \in S} \lambda_s \cdot s : \lambda_s \in \mathbb{F} \right\}$$

where $\sum$ denotes finite sums. An expression the form above is called a *linear combination* of S.

We say that S *spans* V if $\langle S \rangle = V$

Definition. (Finite-dimensional) For a vector space V we say that it is *finite-dimensional* if there exists a finite spanning set.

We'll give some simple remarks without proof.

(i) $\langle S \rangle \leq V$ and conversely if $W \leq V$ and $S \subseteq W$ then $\langle S \rangle \leq W$.

(ii) If $S, T \subseteq W$ and S spans V and $S \subseteq \langle V \rangle$ then T spans V.

(iii) By convention $\langle \emptyset \rangle = \{\mathbf{0}_V\}$.

(iv) $\langle S \cup T \rangle = \langle S \rangle + \langle T \rangle$

For an example consider $V = \mathbb{R}^3$ and consider the sets

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\}$$

$$T = \left\{ \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} \right\}$$

Then $\langle S \rangle = \langle T \rangle = \left\{ \begin{pmatrix} x \\ y \\ 2y \end{pmatrix} : x, y \in \mathbb{R} \right\} \leq \mathbb{R}^3$.

For a second example consider $V = \mathbb{R}^{\mathbb{N}}$ and set $T = \{\delta_n : n \in \mathbb{N}\}$. This is not a spanning set, since we require infinitely many elements from T to make an element in V. In fact we can write that

$$\langle T \rangle = \{f \in \mathbb{R}^{\mathbb{N}} : f(n) = 0 \text{ for all but finitely many terms}\}.$$

Definition. (Linear Independence) A subset $S \subseteq V$ is called *linearly independent* if, for all finite linear combinations

$$\sum_{s \in S} \lambda_s s \quad \text{of S}$$

if the sum is the zero vector in V the $\lambda_s = 0$ for all $s \in S$.

If S is not linearly indepedent we say that S is linearly dependent.

We'll make some more remarks

(i) If $\mathbf{0} \in S$ then S is not linearly independent.

(ii) If we have a finite set, then to show linearly independent, we only need to consider the linear combination of all elements, not all finite lienar combinations.

(iii) However is S is infinite, then we have to consider every possible finite subset of S and show it's linearly independent.

(iv) Every subset of a linearly independent set is itself linearly indepedent.

Definition. (Basis) A subset $S \subseteq V$ is a *basis* for V if S is linearly independent and a spanning set.

For an example consider $e_i \in \mathbb{F}^n$ be given by

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{with the 1 in the } i\text{th entry}$$

then the set $\{e_i : 1 \leq i \leq n\}$ is the standard basis for $\mathbb{F}^n$.

For $P(\mathbb{R})$ the set of real polynomial functions and let $p_n \in P(\mathbb{R})$ be given by $p_n(x) = x^n$, then $\{p_n : n \in \mathbb{Z}_{\geq 0}\}$ is a basis for $P(\mathbb{R})$.

Proposition. If $S \subseteq V$ is a finite spanning set, then there exists a subset $S' \subseteq S$ such that S' is a basis.

Proof. If S is linearly independent then we're done. Otherwise write $S = \{v_1, \ldots, v_n\}$. Then there exists $\lambda_1, \ldots, \lambda_n$ such that $\lambda_1 v_1 + \cdots \lambda_n v_n = \mathbf{0}$ wlog suppose that λ_n is nonzero. Then

$$v_n = -\frac{1}{\lambda_n} \sum_{i=1}^{n-1} \lambda_i v_i$$

so v_n is in the span of the other vectors. Hence $S \setminus \{v_n\}$ is still a spanning set. Repeat which the set is linearly independent, must terminate since the set is finite and the empty set is not a spanning set. $\square$

Corollary. Every finite-dimensional vector space has a finite basis.

Proof. Trivial application of the proposition $\square$

Theorem. (Steinitz Exchange Lemma) Let $S, T \subseteq V$ finite with S linearly independent and T a spanning set of V. Then
(i) $|S| \leq |T|$,
(ii) and there exists $T' \subseteq T$ which has size $|T'| = |T| - |S|$ and $S \cup T'$ spans V.

Proof. To come later...

Let's look at some consequences of the lemma first.

Corollary. For a finite-dimensional vector space V,
(i) Every basis for V is finite.
(ii) All finite basis have the same size.

Proof. V has a finite basis B, suppose we have some other basis B' infinite. Let $B'' \subseteq B'$ with $|B''| = |B| + 1$ then $|B''|$ is linearly independent, so applying (i) of the Steinitz exchange lemma with $S = B''$ and $T = B$ we get a contradiction.

For the second part, let B_1, B_2 be finite basis for V then apply Steinitz symmetrically since both are spanning set and linearly independent, so we get that $|B_1| \geq |B_2|$ and $|B_1| \geq |B_2|$ so $|B_1| = |B_2|$. □

Definition. (Dimension) For a vector space V the *dimension* of V is the size of any basis. We write this as $\dim V$.

This definition is well-defined by the previous corollary.

For an example $\dim \mathbb{F}^n = n$ since we've shown the standard basis has size n. As a complex vector space $\mathbb{C}$ is one-dimensional as a complex vector space and two-dimension as a real vector space, with basis $\{1\}$ and $\{1, i\}$ repectively.

Corollary. For a vector space V let $S, T \subseteq V$ finite, with S linearly independent and T a spanning set, then

$$|S| \leq \dim V \leq |T|$$

with equality if and only if S spans or V is linearly independent respectively.

Proof. The inequalities are immediate from Steinitz. If S is a basis then $|S| = \dim V$ from the previous corollary. Conversely if $|S| = \dim V$ and let B be a basis for V so we have that $|B| = |S|$ so B is a spanning set. So we can apply Steinitz (ii) to B so there exists $B' \subseteq B$ with $|B'| = |B| - |S| = 0$ and $S \cup B' = S \cup \emptyset$ spans V. So S is a basis. Similiar we have a very similar proof for equality in V. □

We will not prove that every vector space has a basis, however some non-finitely dimensional vector spaces have an infinite basis, for example $P(\mathbb{R})$.

Proposition. If V is a finite-dimensional vector space, then if $U \leq V$ then U is finite-dimensional, namely, $\dim U \leq \dim V$ with equality if and only if $U = V$.

Proof. If $U = \{\mathbf{0}\}$, we're done. Otherwise let $\mathbf{0} \neq u_1 \in U$. Then $\{u_1\} \subseteq U$ is linearly indepedent. Repeating, after repeating k times suppose we have $\{u_1, \ldots, u_k\}$ linearly indepedent with $k \leq \dim(V)$ by the previously corollary. If the set spans U we're done, if not we'll add another vector, u_{k+1} outside of the span of our space. If $\{u_1, \ldots, u_{k+1}\}$ is not linearly indepedent, we can write $\mathbf{0}$ non-trivially, so

$$\sum_{i=1}^{k+1} \lambda_i u_i = \mathbf{0}$$

with $\lambda_{k+1} \neq 0$ since $\{u_1, \ldots, u_k\}$ linearly indepedent. Thus we have that

$$u_{k+1} = -\frac{1}{\lambda_{k+1}} \left(\sum_{i=1}^{k} \lambda_i u_i \right)$$

this process must terminate after at most $\dim V$ many steps, by the previous corollary. If $\dim U = \dim V$ apply the previous corollary with S being any basis for U. $\square$

Proposition. (Extending a basis) Let $U \leq V$. For any basis B_U of U there exists a basis B_V of V such that $B_U \subseteq B_V$.

Proof. Apply the second result from Steinitz with $S = B_U$ and T is any basis for V. We obtain that $T' \subseteq T$ s.t.

$$|T'| = |T| - |S| = \dim V - \dim U$$

and $B_V = B_U \cup T'$ spans V. But we have that

$$|B_V| \leq |B_U| + |T'| = \dim V$$

so by the previous corollary, B_V is a basis for V. $\square$

Now we'll finally prove the Steinitz exchange lemma.

Proof. Let $S = \{u_1, \ldots, u_m\}$, $T = \{v_1, \ldots, v_n\}$ with $|T| = m$ and $|T| = n$. If S is empty then we're done. Otherwise there exists $\lambda_i \in \mathbb{F}$ such that

$$u_1 = \sum_{i=1}^{n} \lambda_i v_i$$

so by renumbering we can say that $\lambda_1 \neq 0$. Then

$$v_1 = \frac{1}{\lambda_1}\left(u_1 - \sum_{i=2}^{n} \lambda_i v_i\right)$$

So $\{u_1, v_2, \ldots, v_n\}$ spans V. After repeating k times with $k < m$ suppose $\{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\}$ spans V, then there exists $\lambda_i, \mu_j \in \mathbb{F}$ such that

$$u_{k+1} = \sum_{j=1}^{k} \mu_j u_j + \sum_{i=k+1}^{n} \lambda_i v_i$$

If for all $\lambda_i = 0$ then

$$\left(\sum_{j=1}^{k} \mu_j u_j\right) - u_{k+1} = \mathbf{0}$$

which is a contradiction since S is linearly independent. So by relabeling we have that $\lambda_{k+1} \neq 0$ such that

$$v_{k+1} = \frac{1}{\lambda_{k+1}}\left(u_{k+1} - \sum_{j=1}^{k} \mu_j u_j - \sum_{i=k+1}^{n} \lambda_i v_i\right)$$

so $(u_1, \ldots, u_{k+1}, v_{k+2}, \ldots, v_n\}$ spans V. So we can conclude that $m \neq n$ and $\{u_1, \ldots, u_m, v_{m+1}, \ldots, v_n\}$ spans V hence the set $T' = \{v_{m+1}, \ldots, v_n\}$ exists as claimed. $\square$

Definition. (Nullity) For a linear map $\alpha : V \to W$ we define the *nullity* of α as

$$\mathrm{n}(\alpha) = \dim \ker \alpha.$$

Definition. (Rank) For a linear map $\alpha : V \to W$ we define the *rank* of α as

$$\mathrm{rk}(\alpha) = \dim \operatorname{im} \alpha.$$

Theorem. (Rank-nullity theorem) If V is a finite dimensional $\mathbb{F}$-vector space and W is a $\mathbb{F}$-vector space. Then if $\alpha : V \to W$ is linear then $\operatorname{im}\alpha$ is finite dimensional and

$$\dim V = \mathrm{n}(\alpha) + \mathrm{rk}(\alpha).$$

Proof. Recall the first isomorphism theorem so

$$\frac{V}{\ker \alpha} \cong \operatorname{im} \alpha$$

It is sufficient to prove the lemma

Lemma. For $U \leq V$,

$$\dim(V/U) = \dim V - \dim U$$

Proof. Let $B_U = \{u_1, \dots, u_m\}$ be a basis of U. Extend to a basis $B_V = \{u_1, \dots, u_m, v_{m+1}, \dots, v_n\}$ of V where $m = \dim U$ and $n = \dim V$.
Set $B_{V/U} = \{v_i + U : m+1 \leq i \leq n$. The we claim that $B_{V/U}$ is a basis for V/U of size $n - m$.
To show spanning, for $v \in V$ write

$$v = \sum_i \lambda_i v_i + \sum_j \mu_j v_j$$

Then $v + U = \sum_i \lambda_i (v_i + U) \in \langle B_{V/U} \rangle$. For linear independence, suppose

$$\sum_i \lambda_i (v_i + U) = \mathbf{0} + U$$

hence

$$= \left(\sum_i \lambda_i v_i \right) + U$$

$$\sum_i \lambda_i v_i \in U$$

$$\sum_i \lambda_i v_i = \sum_j \mu_j u_j$$

since B_V is linearly independent, we have that all λ_i and μ_j are zero. Similiarly if $v_i + U = v_j + U$ with $i \neq j$ then we can write $v_i - v_j = \sum_j \mu_j u_j$ which is a contradiction. $\square$

Remark. We can maek a direct proof without quotient spaces by rearranging some of the arguments of the proof.

Corollary. (Linear Pigeonhole principle) If $\dim V = \dim W = n$ and $\alpha : V \to W$ then the following conditions are equivalent.

(i) α is injective,
(ii) α is surjective,
(iii) α is an isomorphism.

Proof. If α injective then $\mathrm{n}(\alpha) = 0$ so by rank nullity we have that $\mathrm{rk}(\alpha) = n$ so α is surjective. If α is surjective then $\mathrm{rk}(\alpha) = n$ so by rank nullity, the dimension of the kernel is 0 hence the kernel is trivial, so α injective, hence α is an isomorphism. If α is an isomorphism, clearly it's injective, so all equivalent. $\square$

Proposition. Suppose V is a vector space with a basis B. For any vector space W and any function $f : B \to W$ there is a unique linear map $F : V \to W$ such that $F(B) = W$.

Proof. First we'll show existance. For $v \in V$ write $v = \sum_b \lambda_b b$ for a finite sum. Then define

$$F(v) = \sum_b \lambda_b f(b).$$

This is well-defined, since B is a basis the λ_b are uniquely determined by v. For $u, v \in V$ and $\lambda \in \mathbb{F}$ we write

$$u = \sum_b \mu_b b, \quad \sum_b \lambda_b b.$$

Then

$$\begin{aligned} F(u + \lambda v) &= F(\sum_b (\mu_b + \lambda\lambda_b) f(b) \\ &= \sum_b \mu_b f(b) + \lambda \sum_b \lambda_b f(b) \\ &= F(u) + \lambda F(v). \end{aligned}$$

So F is linear. To show uniqueness $\overline{F} : V \to W$ is another linear map extending f then,

$$\overline{F}\left(\sum_b \lambda b b\right) = \sum_b \lambda_b \overline{F}(b)$$

which is the same as our definition for F hence they are the same function.

Corollary. For a vector space, V, with $\dim V = n$ with a basis $B = \{v_1, \dots, v_n\}$ for V then there is a unique isomorphism

$$\begin{aligned} F_B &: V \to \mathbb{F}^n \\ \sum_{i=1}^n \lambda_i v_i &\to \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \end{aligned}$$

Proof. Let $E = \{e_1, \dots, e_n\}$ be the standard basis for $\mathbb{F}^n$. Define

$$\begin{aligned} f : B &\to W \\ v_i &\to e_i \end{aligned}$$

and let F_B be the unique linear extension of f to V. We see that f defines a bijection from $B \to E$. Let $\bar{F}_B$ be the unique linear extension of $f^{-1} : E \to B$. Then $\bar{F}_B \cdot F_B$ is the composition of two linear maps, hence it's linear, moreover it is id_B. But also id_V is also a linear extension of id_B, by the proposition, they are the same map so $\bar{F}_B \cdot F_B = F_B \cdot \bar{F}_B = \mathrm{id}_B$. Hence F_B is bijective, so it is an isomorphism. □

Corollary. If V, W are finite dimensional $\mathbb{F}$-vector spaces. Then

$$V \cong W \iff \dim V = \dim W$$

Proof. Trivial from the corollary using the transitivity of the isomorphism relation. □

Definition. (Coordinate vector) $F_B(v) = [v]_B$ is the *coordinate vector* of v with respect to the basis B

For an example if $V \cong \mathbb{F}^n$ and $U \leq V$ with $U \cong \mathbb{F}^m$ then $\dim(V/U) = n - m$, so $\frac{V}{U} \cong \mathbb{F}^{n-m}$.

1.4 Direct sums

Definition. (External direct sum) For $\mathbb{F}$-vector spaces, V and W, we dnote the *external direct sum* of V and W as $V \oplus W$ with underlying set $V \times W$ with addition and scalar multiplication given in the obvious sense.

We can similarly define

$$V_1 \oplus \cdots \oplus V_n = \bigoplus_{i=1}^{n} V_i.$$

Lemma. For V, W finite dimensional vector spaces,

$$\dim(V \oplus W) = \dim V + \dim W$$

Proof.
(First Proof) Let B, C be basis for V, W respectively. Set

$$D = (B \times \{\mathbf{0}_W\}) \cup (\{\mathbf{0}_V\} \times C)$$

it is straightfoward to check that D is basis of $V \oplus W$ of the size $\dim V + \dim W$. □

(Second Proof) Suppose $V \cong \mathbb{F}^n$ and $W \cong \mathbb{F}^m$ construct an isomorphism $V \oplus W \cong \mathbb{F}^{n+m}$. □

Proposition. Let V be a vector space with $U, W \leq V$. There is a surjective linear map

$$\begin{aligned} \varphi : U \oplus W &\to U + W \\ (u, w) &\to u + w \end{aligned}$$

with $\ker \varphi \cong U \cap W$.

Proof. Surjectively and linearity are clear. Note for $(u, w) \in U \oplus W$ then $(u, w) \in \ker \varphi$ if and only if $w = -u$. Hence

$$\ker \varphi = \{(x, -x) : x \in U \cap W\}$$

the map $\psi : U \cap W \to \ker \varphi$ sending $x \to (x, -x)$ is an isomorphism.

Corollary. (Sum-Intersection Formula) If V is finite dimensional and $U, W \leq V$ then

$$\dim(U + W) = \dim U + \dim V - \dim(U \cup V)$$

Applying the rank-nullity theorem to the linear map φ in the proposition we get that

$$\begin{aligned} \dim U + \dim W &= \dim(U \oplus V) \\ &= \dim(\ker \varphi) + \dim(\operatorname{im} \varphi) \\ &= \dim(U + W) + \dim(U \cap W) \quad \square \end{aligned}$$

We can also give an explicit basis. Given a basis B for $U \cap W$, extend B to a basis B_U for U, and a basis B_W for W. Then $B_U \cap B_W$ spans $U + W$ and

$$|B_U \cup B_W| \leq |B_U| + |B_W| - |B| = \dim(U + V)$$

hence $B_U \cup B_W$ is linearly independent so it's a basis for $U + W$.

Remark. We could also check directly that $B_U \cup B_W$ is linearly independent of the size $\dim(U+V)$ without assuming the sum-intersection formula, so this alos servers as an alternative proof of the sum-intersection formula.

Definition. (Internal direct sum) Suppose $U, W \leq V$ satisify
(i) $U + W = V$,
(ii) $U \cap W = \{\mathbf{0}_V\}$.
Then

$$\varphi : U \oplus W \to V$$

is an isomorphism, and we say that V is the *internal direct sum* of U and W, and we write that $V = U \oplus W$.

Alternatively, every element $v \in V$ can be written *uniquely* as $v = u + w$ for $u \in U, w \in W$.

Definition. (Direct complement) For $U \leq V$ a *direct complement* to U in V is a subspace $W \leq V$ satisfying $V = U \oplus W$.

Proposition. If V is finite dimensional then every subspace has a direct complement.

Proof. Let $U \leq V$ and let B_U be a basis for U. Extend to a basis B_V for V. Set $W =_V \setminus B_U\rangle$. Then

$$\begin{aligned} V = \langle B_V \rangle &= \langle B_U \cup (B_V \setminus B_U) \rangle \\ &= \langle B_U \rangle + \langle B_V \setminus B_U \rangle \\ &= U + W. \end{aligned}$$

Moreover using the sum-intersection formula

$$\dim(U \cap W) = |B_V| + |B_U| - |B_V \setminus B_U| = 0.$$

Hence $U \oplus W = V$. $\square$

More generally for $U_1, \ldots, U_n \leq V$ we say that V is the direct sum of the U_i and write that

$$V = U_1 \oplus + \cdots + \oplus V_n = \bigoplus_{i=1}^{n} V_i$$

if the map

$$\begin{aligned} \varphi : U_1 \oplus \cdots \oplus U_n &\to V \\ (u_1, \ldots, u_n) &\to u_1, \ldots, u_n \end{aligned}$$

is an isomorphism. Equivalently every $v \in V$ can be uniquely written as $v = u_1 + \cdots + u_n$ for $u_i \in U_i$.

2 Matrices and Linear Maps

2.1 Vector spaces of linear maps

Definition. For V, W $\mathbb{F}$-vector spaces we define

$$\mathcal{L}(V, W) = \{\alpha : V \to W : \alpha \text{ is linear}\}$$

which forms a $\mathbb{F}$-vector space under pointwise addition and obvious scalar multiplication.

Recall that $M_{m\times n}$ is the space of matrices over $\mathbb{F}$ with m rows and n columns. For $A \in M_{m\times n}(\mathbb{F})$ we write $A = (a_{ij})$ where $a_{ij} \in \mathbb{F}$ is the entry in the ith row and the jth column.

Let $B = \{v_1, \ldots, v_n\}, C = \{w_1, \ldots, w_m\}$ are *ordered* basis for V, W.

Let $\alpha \in \mathcal{L}(V, W)$. We can write

$$\begin{aligned} \alpha(v_1) &= a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m \\ \alpha(v_2) &= a_{12}w_1 + a_{22}w_2 + \cdots + a_{m1}w_m \\ &\vdots \\ \alpha(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m \end{aligned}$$

Definition. (Matrix) The *matrix* of α with respect to the ordered basis B, C is

$$[\alpha]_C^B = (a_{ij}) \in M_{m\times n}(\mathbb{F})$$

Recall we have a linear isomorphism

$$\begin{aligned} \varepsilon_B : V &\to \mathbb{F}^n \\ v = \sum_{i=1}^n \lambda_i v_i &\to (\lambda_i)_i = [v]_B \end{aligned}$$

where $[v]_B$ is the coordinate vector of v with respect to B.

Proof. Let $v \in V$ write $v = \sum_{j=1}^n \lambda_j v_j$. Then

$$\begin{aligned} \alpha(v) &= \sum_{j=1}^n \lambda_j \alpha(v_j) \\ &= \sum_{j=1}^n \lambda_j \sum_{i=1}^m a_{ij} w_i \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n \lambda_j a_{ij} \right) w_i. \end{aligned}$$

So

$$\begin{aligned} [\alpha(v)]_C &= \left(\sum_{j=1}^n a_{ij} \lambda_j \right)_i \\ &= (a_{ij}) \cdot (\lambda_j) \\ &= [\alpha]_C^B [v]_B. \end{aligned}$$

Hence (i) is proved. For (ii), take $1 \le j \le n$, so $[v_j]_B = e_j$. Hence for $A \in M_{m\times n}(\mathbb{F})$, $A[v_j]_B$ is the jth column of A. But if $A[v_j]_B = [\alpha(v_j)]_C = [\alpha]_C^B [v_j]_B = [\alpha]_C^B e_j$, then $A[v_j]_B$ is also the jth column of $[\alpha]_C^B$. Since this holds for all j in our range, they are the same matrix.

Now for part (iii), let $\alpha, \beta \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$. Then

$$\begin{aligned} [\alpha + \lambda\beta]_C^B [v]_B &= [(\alpha + \lambda\beta)(v)]_C \\ &= [\alpha(v) + \lambda\beta(v)]_C \\ &= [\alpha(v)]_C + \lambda[\beta(v)]_C \\ &= ([\alpha]_C^B + \lambda[\beta]_C^B)[v]_B \end{aligned}$$

for all $v \in V$. Hence by (ii) we get that $[\alpha + \lambda\beta]_c^b = [\alpha]_C^B + \lambda[\beta]_C^B$ so the map is linear. Let $\alpha \in \ker(\varepsilon_C^B)$ so that $[\alpha]_C^B = 0 \in M_{m\times n}(\mathbb{F})$. Then by (i) we have that $[\alpha(v)]_C = 0$ for all $v \in V$. But $\varepsilon : w \to [w]_C$ is an isomorphism so $\alpha(v) = 0$ for all $v \in V$ hence $\alpha = 0$ and α is injective. For surjectivity let $A \in M_{m\times n}(\mathbb{F})$ and define $f : B \to W$ by $f(v_j) = \sum_{i=1}^n a_{ij} w_I$ and extend f to a linear map $F : V \to W$. Then $[F]_C^B = A$. So ε_C^B is an isomorphism. $\square$

Proposition. Let V, W, X be finite-dimensional $\mathbb{F}$-vector spaces with basis B, C, D and $\alpha \in \mathcal{L}(V, W)$ and $\beta \in \mathcal{L}(W, X)$. Then

$$[\beta \circ \alpha]_D^B = [\beta]_D^C [\alpha]_C^B.$$

Proof. By the theorem $[\beta \circ \alpha]_D^B$ is the unique matrix A satisfying

$$A[v]_B = [\beta(\alpha(v))]_D, \quad \forall v \in V.$$

But $[\beta]_D^C [\alpha]_C^B [v]_B = [\beta]_D^C [\alpha(v)]_C = [\beta(\alpha(v))]_D$. So by (ii) of theorem they are equal. □

Remark. For any basis B of V,

$$[\mathrm{id}_V]_B^B = I_{\dim V}.$$

Definition. (Change of basis matrix) Let B, B' be basis for V and $\dim V = n$. The *change of basis matrix* from B to B' is given by

$$P = [\mathrm{id}_V]_{B'}^B \in M_{m \times n}(\mathbb{F})$$

Equivalently letting $B = \{v_i\}_{i=1}^n$ and $B' = \{v_i'\}_{i=1}^n$, then

$$P = (p_{ij}) \quad \text{where} \quad v_j = \sum_{i=1}^n p_{ij} v_i'$$

so the jth column of P is $[v_j]_{B'}$.

Proposition. For V, W finite-dimensional vector spaces,
(i) $[\mathrm{id}_V]_{B'}^B \in GL_n(\mathbb{F})$ with inverse $[\mathrm{id}_V]_B^{B'}$.
(ii) If $\alpha \in \mathcal{L}(V, W)$ and B, B' basis for V and C, C' basis for W, then

$$[\alpha]_{C'}^{B'} = [\mathrm{id}_W]_{C'}^C [\alpha]_C^B [\mathrm{id}_V]_B^{B'}.$$

Proof. By the remark,

$$I_n = [\mathrm{id}_V]_B^B = [\mathrm{id}_V]_B^{B'} [\mathrm{id}_V]_{B'}^B$$

and symmetrically swapping B and B'. For the second part the result is immediate from the proposition.

Definition. (Equivalent matrices) Let $A, A' \in M_{m \times n}(\mathbb{F})$. We say that A and A' are *equivalent* if $\exists P \in GL_m(\mathbb{F})$, $Q \in GL_n(\mathbb{F})$ such that $A' = PAQ$.

Remark. Certianly A is equivalent to itself by $P = I_m$ and $Q = I_n$.
If $A' = PAQ$ then $A = P^{-1} A' Q^{-1}$.
If $A'' = RA'S$ too, then $A'' = (RP)A(QS)$, so the equivalence of matrices is an equivance relation on $M_{m \times n}(\mathbb{F})$.

Theorem. Let V, W be finite-dimensional $\mathbb{F}$-vector spaces. Let $\dim V = n$, $\dim W = m$ and let $\alpha \in \mathcal{L}(V, W)$. Let $r = \mathrm{rk}(\alpha)$. Then,

(i) There exists basis B, C for V, W respectively such that

$$[\alpha]_C^B = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m \times n}(\mathbb{F})$$

where I_r is the identity matrix of size r, and the zeros are block zero matrices.

(ii) If

$$[\alpha]_{C'}^{B'} = \begin{pmatrix} I_{r'} & 0 \\ 0 & 0 \end{pmatrix} \in M_{m \times n}(\mathbb{F})$$

for some basis B', C' of V, W respectively, then $r' = r$

Proof. By rank-nullity $\mathrm{n}(\alpha) = n - r$. Let $\{v_{r+1}, \ldots, v_n\}$ be a basis for $\ker \alpha$. Extend to a basis $B = \{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$. Then $\{\alpha(v_1), \ldots, \alpha(v_r)\}$ spans the image, and has size at most $\dim(\mathrm{im}(\alpha))$, so it's linearly independent, hence we can extend it to form a basis of W.

$$C = \{w_1 = \alpha(v_1), \ldots, w_r = \alpha(v_r), w_{r+1}, \cdots, w_m\}$$

Then

$$\alpha(v_j) = \begin{cases} w_j & 1 \leq j \leq r \\ \mathbf{0} & \text{otherwise} \end{cases}$$

hence we have that $[\alpha]_C^B = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

For the second part, if $[\alpha]_{C'}^{B'} = \begin{pmatrix} I_{r'} & 0 \\ 0 & 0 \end{pmatrix}$ then

$$\alpha(v'_j) = \begin{cases} w'_j & 1 \leq j \leq r' \\ \mathbf{0} & \text{otherwise} \end{cases}.$$

Hence $w'_1, \ldots, w'_{r'}$ span $\mathrm{im}(\alpha)$ and are linearly independent. Hence $\mathrm{rk}(\alpha) = r'$. □

Definition. (Column-space) For $A \in M_{m \times n}(\mathbb{F})$ the *column-space* $\mathrm{Col}(A)$ is the subspace of $\mathbb{F}^m$ spanned by the columns of A. The dimension of the column-space is called the *column-rank* of A.

Definition. (Row-space) For $A \in M_{m \times n}(\mathbb{F})$ the *row-space* $\mathrm{Row}(A)$ is the subspace of $\mathbb{F}^m$ spanned by the rows of A (when transposed as column vectors). The dimension of the row-space is called the *row-rank* of A.

Remark.

$$\mathrm{Row}(A) = \mathrm{Col}(A^T)$$

hence the row-rank of A is the same as the column-rank of A^T.

Remark. Given a matrix $A \in M_{m\times n}(F)$ we can define a linear map $\alpha : \mathbb{F}^n \to \mathbb{F}^m$ by $\alpha(v) = Av$. Then $\mathrm{im}(\alpha) = \mathrm{Col}(A)$, so the rank of α is the same as the column-rank of A. Moreover, $A = [\alpha]_{E_m}^{E_n}$ where E_k are the standard basis for $\mathbb{F}^k$.

We may write $\mathrm{im}\, A, \ker A, \mathrm{rk}(A), \mathrm{n}(A)$ to refer to the corresponding concepts for α.

Theorem. Let $A, A' \in M_{m\times n}(\mathbb{F})$, then
(i) A is equivalent to

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \text{ where } r \text{ is the column-rank of } A$$

(ii) A and A' are equivalent if and only if the have the same column-rank.

Proof. We'll first prove a lemma.

Lemma. For $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{n\times p}(\mathbb{F})$ then $\mathrm{rk}(A \cdot B) \leq \min(\mathrm{rk}(A), \mathrm{rk}(B))$.

Proof. We have that $\mathrm{im}(AB) \leq \mathrm{im}(A)$ so $\mathrm{rk}(AB) \leq \mathrm{rk}(A)$. If $Bv = \mathbf{0}$ for $v \in \mathbb{F}^p$, then $ABv = \mathbf{0}$, so $\mathrm{n}(B) \geq \mathrm{n}(AB)$, so applying rank-nullity, we get that

$$p - \mathrm{rk}(B) \leq p - \mathrm{rk}(AB) \implies \mathrm{rk}(AB) \leq \mathrm{rk}(B) \quad \square$$

Now we'll prove the first part of the theorem. Let α the natural linear map corresponding to A, so $A = [\alpha]_{E_m}^{E_n}$. By the previous theorem, there exists matrices B, C of $\mathbb{F}^n, \mathbb{F}^m$ such that

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = [\alpha]_C^B = [\mathrm{id}_{\mathbb{F}^m}]_C^{E_m} [\alpha]_{E_m}^{E_n} [\mathrm{id}_{\mathbb{F}^n}]_{E_n}^B = PAQ$$

where $r = \mathrm{rk}(\alpha)$ which we know is equal to the column-rank of A.

If A' has column-rank r then both matrices are equivalent to $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, so by transitivity, A and A' are equivalent. Conversely suppose that A and A' are equivalent, so $A' = PAQ$. By the lemma $\mathrm{rk}(A') \geq \mathrm{rk}(AQ) \geq \mathrm{rk}(A)$ and symmetrically we get that $\mathrm{rk}(A) \geq \mathrm{rk}(A')$, hence $\mathrm{rk}(A') = \mathrm{rk}(A)$. $\square$

Theorem. For any $A \in M_{m\times n}(\mathbb{F})$, the row-rank of A is equal to the column-rank of A.

Proof. Note that if P is invertiable, then so it the tranpose with inverse $(P^{-1})^T$. Let r be the column-rank of A. So there exists matrices $P \in GL_m(\mathbb{F})$ and $Q \in GL_n(\mathbb{F})$ such that $PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m\times n}(\mathbb{F})$. Then A^T is equivalent to $Q^T A^T P^T = (PAQ)^T = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{n\times m}(\mathbb{F})$. By the previous theorem, the column-rank of A^T is r which also the row-rank of A. $\square$

Let V be a finite-dimensional vector space and B, B' be basis for V. Now let $\alpha \in \mathrm{End}(V) = \mathcal{L}(V, V)$. Then

$$[\alpha]_{B'}^{B'} = [\mathrm{id}_V]_{B'}^B [\alpha]_B^B [\mathrm{id}_V]_B^{B'}$$

Definition. (Similarity) For matrices $A, A' \in M_{n\times m}(\mathbb{F})$ are *similar* if there exists $P \in GL_n(\mathbb{F})$ such that $A' = P^{-1}AP$.

Remark. We have some remarks showing the similarity and equivalence are not the same thing.

(i) Similarity is an equivalence relation on $M_{n\times n}(\mathbb{F})$.

(ii) Similar matrices are equivalent but equivalent matrices need not be similar.

For example every matrix in $GL_n(\mathbb{F})$ is equivalent to I_n but I_n forms its only single element equivalence class, when we think about similarity.

2.2 Elementary operations on matrices

Definition. (Elementary row operations) Let $r_1, \ldots, r_m$ be the rows of A. We have three types of *elementary row operations* on A

(i) Swap r_i and r_j with $i \neq j$.

(ii) Replace r_i with λr_i with $0 \neq \lambda \in \mathbb{F}$.

(iii) Replace r_i with $r_i + \lambda r_j$ with $\lambda \in \mathbb{F}$ and $i \neq j$.

Similarly there are three types of elementary column operations.

Remark. These are all reversable.

Each elementary operation has a corresponding matrix representation repesentation. All corresponding matrices are invertiable.

Lemma. If E is a matrix of type (i)-(iii) then EA is obtained from A by applying the corresponding ERO to A.

Proof. Direct matrix computation.

Remark. Similarly AE is obtained by applying the corresponding ECO

Remark. EROs preserve $\mathrm{Row}(A)$ (and ECOs preserve $\mathrm{Col}(A)$).

So both EROs and ECOs preserve the row-rank of a matrix, and therefore also the rank of the linear map corresponding to the matrix.

Definition. (Row reduced echelon form) A matrix $A \in M_{m\times n}(\mathbb{F})$ is said to be in *row reduced echelon form* (RRE) if

(i) All non-zero rows of A appear above all zero rows.

(ii) The leftmost non-zero element of a non-zero row is 1 (called the *pivot entry*).

(iii) If row r_i, r_j are non-zero rows with $i < j$ then the index of the pivot entry of i is less than the index of the pivot entry of j.

(iv) In a column containing a pivot entry, every other entry is zero.

For an example consider

$$M = \begin{pmatrix} 1 & a & 0 & 0 & b \\ 0 & 0 & 1 & 0 & c \\ 0 & 0 & 0 & 1 & d \end{pmatrix}$$

which is in row reduced echelon form. Similarly we have column reduced echelon form, which have the exact same rules but transposed.

Lemma. If A is in row reduced echelon form then the row rank of A is the number of non-zero rows of A.

Proof. Let $r_1, \ldots, r_k$ be the non-zero rows of A let $j_i = P(v_i)$ be the pivot entry. Certainly $r_1, \ldots, r_k$ span $\mathrm{Row}(A)$. Suppose that

$$v = \sum_{i=1}^{k} \lambda_i r_i = 0 \quad (\lambda_i \in \mathbb{F}).$$

Then $(v)_{j_i} = \lambda_i = 0$ so the non-zero rows are linearly independent so we're done. □

Proposition. Every matrix $A \in M_{m\times n}(\mathbb{F})$ can be put into row reduced echelon form with elementary row operations.

Proof. Proceed by induction on n. Write that $A = [c_1 \mid \cdots \mid c_n]$. If $c_1 = 0$ apply induction to $[c_2 \mid \cdots \mid c_n]$, so suppose that $c_1 \neq 0$, suppose that element in $(i, 1)$ is non-zero. Applying row operations (i) we can move it to $(1, 1)$. Apply row operation (ii) to rescale it to be 1. Now we can clear the rest of the column by (iii). By induction we can use elementary row operation on rows 2-m to reduce further. This is decreasing the dimension to the process terminates, hence the matrix can be put into row reduced echelon form. □

Remark. Putting a matrix into RRE form preserves the row-space and the RRE of any matrix is unique. Also if A is a square matrix then A either has a zero row or is the identity.

Theorem. For $A \in M_{m\times n}(\mathbb{F})$ the following are equivalent:
(i) $\mathrm{rk}(A) = n$.
(ii) A is a product of elementary matrices.
(iii) A is invertiable.

Proof. Let's prove that (i) $\implies$ (ii). By the proposition there exists elementary matrices E_i such that $E_1 \ldots E_\ell A$ is in RRE form. By the remark this is I_n hence $A = E_\ell^{-1} \cdots E_1^{-1}$ which are also elementary. For (ii) $\implies$ (iii) elementary matrix lie in $GL_n(\mathbb{F})$ which is a group, hence closed. Finially for (iii) $\implies$ (i) suppose there exists $B \in M_{m\times n}(\mathbb{F})$ such that $AB = I_n$. Then for $v \in \mathbb{F}^n$ we have that $v = (AB)v = A(Bv)$, so $v \in \operatorname{im} A$.