Part IB — Linear Algebra

Based on lectures by Prof. P. Raphael and notes by thirdsgames.co.uk

Michaelmas 2022

Contents

1 Vector spaces and linear dependence **3**
1.1 Vector spaces . . . 3
1.2 Subspaces . . . 4
1.3 Sum of subspaces . . . 5
1.4 Quotients . . . 6
1.5 Span . . . 6
1.6 Dimensionality . . . 7
1.7 Linear independence . . . 8
1.8 Bases . . . 8
1.9 Steinitz exchange lemma . . . 10
1.10 Consequences of Steinitz exchange lemma . . . 10
1.11 Dimensionality of sums . . . 11
1.12 Direct sums . . . 12

2 Linear maps **15**
2.1 Linear maps . . . 15
2.2 Isomorphism . . . 16
2.3 Kernel and image . . . 17
2.4 Rank and nullity . . . 18
2.5 Space of linear maps . . . 20
2.6 Matrices . . . 20
2.7 Linear maps as matrices . . . 21
2.8 Change of basis . . . 24
2.9 Equivalent matrices . . . 26
2.10 Column rank and row rank . . . 28
2.11 Conjugation and similarity . . . 30
2.12 Elementary operations . . . 30
2.13 Gauss' pivot algorithm . . . 32
2.14 Representation of square invertible matrices . . . 32

3 Dual spaces 34
3.1 Dual spaces . . . 34
3.2 Annihilators . . . 36
3.3 Dual maps . . . 37
3.4 Properties of the dual map . . . 39
3.5 Double duals . . . 40

4 Bilinear Forms 43
4.1 Introduction . . . 43
4.2 Change of basis for bilinear forms . . . 46

5 Determinant and Traces 48
5.1 Trace . . . 48
5.2 Permutations and transpositions . . . 49
5.3 Determinant . . . 49
5.4 Volume forms . . . 50
5.5 Multiplicative property of determinant . . . 52
5.6 Singular and non-singular matrices . . . 53
5.7 Determinants of linear maps . . . 54
5.8 Determinant of block-triangular matrices . . . 55

6 Adjugate Matrices 57
6.1 Column and row expansions . . . 57
6.2 Adjugates . . . 58
6.3 Cramer's rule . . . 59

§1 Vector spaces and linear dependence

§1.1 Vector spaces

Definition 1.1 (F-vector space)
Let F be an arbitrary field. A **F-vector space** is an abelian group $(V, +)$ equipped with a function

$$F \times V \to V; \quad (\lambda, v) \mapsto \lambda v$$

such that

1. $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$
2. $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$
3. $\lambda(\mu v) = (\lambda\mu)v$
4. $1v = v$

Such a vector space may also be called a *vector space over F*.

Example 1.1
Let $n \in \mathbb{N}$. F^n is the space of column vectors of length n with entries in F.

$$v \in F^n, v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, x_i \in F, 1 \leq i \leq n.$$

$$v + w = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}, \quad \lambda v = \begin{bmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{bmatrix}.$$

F^n is a F-vector space.

Example 1.2
Let X be a set, and define $\mathbb{R}^X = \{f \colon X \to \mathbb{R}\}$ (set of real valued functions on X). Then $\mathbb{R}^X$ is an $\mathbb{R}$-vector space:

- $(f_1 + f_2)(x) = f_1(x) + f_2(x)$.
- $(\lambda f)(x) = \lambda f(x), \lambda \in \mathbb{R}$.

Example 1.3

Define $M_{n,m}(F)$ to be the set of $n \times m$ F-valued matrices. This is an F-vector space, where the sum of matrices is computed elementwise.

Remark 1. The axioms of scalar multiplication imply that $\forall v \in V,\ 0_F \cdot v = 0_V$.

§1.2 Subspaces

Definition 1.2 (Subspace)

Let V be an F-vector space. The subset $U \subseteq V$ is a vector subspace of V, denoted $U \leq V$, if

1. $0_V \in U$
2. $u_1, u_2 \in U \implies u_1 + u_2 \in U$
3. $(\lambda, u) \in F \times U \implies \lambda u \in U$

Conditions (ii) and (iii) are equivalent to

$$\forall \lambda_1, \lambda_2 \in F, \forall u_1, u_2 \in U, \lambda_1 u_1 + \lambda_2 u_2 \in U$$

This means that U is *stable* by vector addition and scalar multiplication.

Proposition 1.1

If V is an F-vector space, and $U \leq V$, then U is an F-vector space.

Example 1.4

Let $V = \mathbb{R}^{\mathbb{R}}$ be the space of functions $\mathbb{R} \to \mathbb{R}$. The set $C(\mathbb{R})$ of continuous real functions is a subspace of V. The set $\mathbb{P}(\mathbb{R})$ of real polynomials is a subspace of $C(\mathbb{R})$ so $\mathbb{P}(\mathbb{R}) \leq V$.

Example 1.5

Consider the subset of $\mathbb{R}^3$ such that $x_1 + x_2 + x_3 = t$ for some real t. This is a subspace for $t = 0$ only, since no other t values yields the origin as a member of the subset.

Proposition 1.2 (Intersection of two subspaces is a subspace)

Let V be an F-vector space. Let $U, W \leq V$. Then $U \cap W$ is a subspace of V.

Proof. First, note $0_V \in U, 0_V \in W \implies 0_V \in U \cap W$. Now, consider stability:

$$\lambda_1, \lambda_2 \in F, v_1, v_2 \in U \cap W \implies \lambda_1 v_1 + \lambda_2 v_2 \in U, \lambda_1 v_1 + \lambda_2 v_2 \in W$$

Hence stability holds. □

§1.3 Sum of subspaces

Warning 1.1

The union of two subspaces is not, in general, a subspace. For instance, consider $\mathbb{R}, i\mathbb{R} \subset \mathbb{C}$. Their union does not span the space; for example, $1 + i \notin \mathbb{R} \cup i\mathbb{R}$.

Definition 1.3 (Subspace Sum)

Let V be an F-vector space. Let $U, W \leq V$. The sum $U + W$ is defined to be the set

$$U + W = \{u + w \colon u \in U, w \in W\}$$

Proposition 1.3

$U + W$ is a subspace of V.

Proof. First, note $0_{U+W} = 0_U + 0_W = 0_V$. Then, for $\lambda_1, \lambda_2 \in F$ and $f, g \in U + W$ we have

$$\begin{aligned} f &= f_1 + f_2 \\ g &= g_1 + g_2 \end{aligned}$$

with $f_1, g_1 \in U$ and $f_2, g_2 \in W$. Hence

$$\begin{aligned} \lambda_1 f + \lambda_2 g &= \lambda_1(f_1 + f_2) + \lambda_2(g_1 + g_2) \\ &= \underbrace{(\lambda_1 f_1 + \lambda_2 g_1)}_{\in U} + \underbrace{(\lambda_1 f_2 + \lambda g_2)}_{\in W} \in U + W. \end{aligned}$$

□

Proposition 1.4

The sum $U + W$ is the smallest subspace of V that contains both U and W.

Proof. Left as an exercise. □

§1.4 Quotients

Definition 1.4 (Quotient)

Let V be an F-vector space. Let $U \leq V$. The **quotient space** V/U is the abelian group V/U equipped with the scalar multiplication function

$$F \times V/U \to V/U; \quad (\lambda, v+U) \mapsto \lambda v + U$$

Note. We must check that the multiplication operation is well-defined. Indeed, suppose $v_1 + U = v_2 + U$. Then,

$$v_1 - v_2 \in U \implies \lambda(v_1 - v_2) \in U \implies \lambda v_1 + U = \lambda v_2 + U \in V/U$$

Proposition 1.5

V/U is an F-vector space.

Proof. Left as an exercise □

§1.5 Span

Definition 1.5 (Span of a family of vectors)

Let V be an F-vector space. Let $S \subset V$ be a subset (so S is a set of vectors). We define the **span** of S, written $\langle S \rangle$, as the set of finite linear combinations of elements of S. In particular,

$$\langle S \rangle = \left\{ \sum_{s \in S} \lambda_s v_s \colon \lambda_s \in F, v_s \in S, \text{only finitely many nonzero } \lambda_s \right\}$$

By convention, we specify

$$\langle \varnothing \rangle = \{0\}$$

so that all spans are subspaces.

Remark 2. $\langle S \rangle$ is the smallest vector subspace of V containing S.

Example 1.6

Let $V = \mathbb{R}^3$, and

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \\ -4 \end{pmatrix} \right\}$$

Then we can check that

$$\langle S \rangle = \left\{ \begin{pmatrix} a \\ b \\ 2b \end{pmatrix} : (a, b) \in \mathbb{R} \right\}$$

Example 1.7

Let $V = \mathbb{R}^n$. We define

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where the 1 is in the ith position. Then $V = \langle (e_i)_{1 \le i \le n} \rangle$.

Example 1.8

Let X be a set, and $\mathbb{R}^X = \{f \colon X \to \mathbb{R}\}$. Then let $S_x \colon X \to \mathbb{R}$ be defined by

$$S_x(y) = \begin{cases} 1 & y = x \\ 0 & \text{otherwise} \end{cases}$$

Then, $\langle (S_x)_{x \in X} \rangle = \left\{ f \in \mathbb{R}^X : f \text{ has finite support} \right\}$, where the support of f is defined to be $\{x \colon f(x) \neq 0\}$.

§1.6 Dimensionality

Definition 1.6

Let V be an F-vector space. Let $S \subset V$. We say that S spans V if $\langle S \rangle = V$. If S spans V, we say that S is a generating family of V.

Definition 1.7 (Finite dimensional)

Let V be an F-vector space. V is **finite dimensional** if it is spanned by a finite set.

Definition 1.8 (Infinite dimensional)

Let V be an F-vector space. V is **infinite dimensional** if there is no family S with finitely many elements which span V.

Example 1.9

Consider the set $V = \mathbb{P}[x]$ which is the set of polynomials on $\mathbb{R}$. Further, consider $V_n = \mathbb{P}_n[x]$ which is the subspace with degree less than or equal to n. Then V_n is spanned by $\{1, x, x^2, \dots, x^n\}$, so V_n is finite-dimensional.

Conversely, V is infinite-dimensional; there is no finite set S such that $\langle S \rangle = V$. The proof is left as an exercise.

§1.7 Linear independence

Definition 1.9 (Linear independence)

We say that $v_1, \dots, v_n \in V$ are **linearly independent** or **free**, if, for $\lambda_i \in F$,

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \implies \forall i, \lambda_i = 0.$$

Remark 3. Linear dependence implies $\exists\ \lambda_i \in F$ and $j \in [1, n]$ s.t. $\sum_{i=1}^{n} \lambda_i v_i = 0$ and $\lambda_j \neq 0$. This implies $v_j = -\frac{1}{\lambda_j} \sum_{i \neq j}^{n} \lambda_i v_i$, i.e. one of the vectors can be written as a linear combination of the remaining ones.

Remark 4. If $(v_i)_{1 \le i \le n}$ are linearly independent, then

$$\forall i \in \{1, \dots, n\}, v_i \neq 0$$

§1.8 Bases

Definition 1.10 (Basis)

$S \subset V$ is a basis of V if

1. $\langle S \rangle = V$
2. S is a linearly independent set

So, a basis is a linearly independent/free generating family.

Example 1.10

Let $V = \mathbb{R}^n$. The *canonical basis* (e_i) is a basis since we can show that they are free and span V. Proof is left as an exercise.

Example 1.11

Let $V = \mathbb{C}$, considered as a $\mathbb{C}$-vector space. Then $\{1\}$ is a basis. If V is a $\mathbb{R}$-vector space, $\{1, i\}$ is a basis.

Example 1.12

Consider again $\mathbb{P}[x]$, polys on $\mathbb{R}$. Then $S = \{x^n : n \geq 0\}$ is a basis of $\mathbb{P}$.

Lemma 1.1 (Unique decomposition for everything equivalent to being a basis)

Let V be an F-vector space. Then, $(v_1, \ldots, v_n)$ is a basis of V if and only if any vector $v \in V$ has a *unique* decomposition

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

Remark 5. In the above definition, we call $(\lambda_1, \ldots, \lambda_n)$ the *coordinates* of v in the basis $(v_1, \ldots, v_n)$.

Proof. Suppose $(v_1, \ldots, v_n)$ is a basis of V. Then $\forall v \in V$ there exists $\lambda_1, \ldots, \lambda_n \in F$ such that

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

So there exists a tuple of λ values. Suppose two such λ tuples exist. Then

$$v = \sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \lambda_i' v_i \implies \sum_{i=1}^{n} (\lambda_i - \lambda_i') v_i = 0 \implies \lambda_i = \lambda_i'$$

since v_i linearly independent. The converse is left as an exercise. □

Lemma 1.2 (Some subset of a spanning set is a basis)

If $\langle\{v_1, \dots, v_n\}\rangle = V$, then some subset of this set is a basis of V.

Proof. If $(v_1, \dots, v_n)$ are linearly independent, this is a basis. Otherwise, one of the vectors can be written as a linear combination of the others. So, up to reordering,

$$\begin{aligned} v_n \in \langle\{v_1, \dots, v_{n-1}\}\rangle &\implies \langle\{v_1, \dots, v_n\}\rangle = \langle\{v_1, \dots, v_{n-1}\}\rangle \\ &\implies \langle\{v_1, \dots, v_{n-1}\}\rangle = V \end{aligned}$$

So we have removed a vector from this set and preserved the span. By induction, we will eventually reach a basis. □

§1.9 Steinitz exchange lemma

Theorem 1.1 (Steinitz exchange lemma)

Let V be a finite dimensional F-vector space. Let $(v_1, \dots, v_m)$ be linearly independent, and $(w_1, \dots, w_n)$ span V. Then,

1. $m \leq n$; and
2. up to reordering, $(v_1, \dots, v_m, w_{m+1}, \dots w_n)$ spans V.

Proof. Suppose that we have replaced $\ell \geq 0$ of the w_i.

$$\langle v_1, \dots, v_\ell, w_{\ell+1}, \dots w_n \rangle = V$$

If $m = \ell$, we are done. Otherwise, $\ell < m$. Then, $v_{\ell+1} \in V = \langle v_1, \dots, v_\ell, w_{\ell+1}, \dots w_n \rangle$ Hence $v_{\ell+1}$ can be expressed as a linear combination of the generating set. Since the $(v_i)_{1 \leq i \leq m}$ are linearly independent (free), one of the coefficients on the w_i are nonzero. In particular, up to reordering we can express $w_{\ell+1}$ as a linear combination of $v_1, \dots, v_{\ell+1}, w_{\ell+2}, \dots, w_n$. Inductively, we may replace m of the w terms with v terms. Since we have replaced m vectors, necessarily $m \leq n$. □

§1.10 Consequences of Steinitz exchange lemma

Corollary 1.1

Let V be a finite-dimensional F-vector space. Then, any two bases of V have the same number of vectors. This number is called the dimension of V, $\dim_F V$.

Proof. Suppose the two bases are $(v_1, \dots, v_n)$ and $(w_1, \dots, w_m)$. Then, $(v_1, \dots, v_n)$ is free and $(w_1, \dots, w_m)$ is generating, so the Steinitz exchange lemma shows that $n \leq m$. Vice versa, $m \leq n$. Hence $m = n$. □

Corollary 1.2

Let V be an F-vector space with finite dimension n. Then,

1. Any independent set of vectors has at most n elements, with equality if and only if it is a basis.
2. Any spanning set of vectors has at least n elements, with equality if and only if it is a basis.

Proof. Exercise. □

§1.11 Dimensionality of sums

Proposition 1.6

Let V be an F-vector space. Let U, W be subspaces of V. If U, W are finite-dimensional, then so is $U + W$, with

$$\dim_F(U + W) = \dim_F U + \dim_F W - \dim_F(U \cap W)$$

Proof. Consider a basis $(v_1, \dots, v_n)$ of the intersection. Extend this basis to a basis $(v_1, \dots, v_n, u_1, \dots, u_m)$ of U and $(v_1, \dots, v_n, w_1, \dots, w_k)$ of W. Then, we will show that $(v_1, \dots, v_n, u_1, \dots, u_m, w_1, \dots, w_k)$ is a basis of $\dim_F(U + W)$, which will conclude the proof. Indeed, since any component of $U + W$ can be decomposed as a sum of some element of U and some element of W, we can add their decompositions together. Now we must show that this new basis is free.

$$\sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^m \beta_i u_i + \sum_{i=1}^k \gamma_i w_i = 0$$
$$\underbrace{\sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^m \beta_i u_i}_{\in U} = \underbrace{\sum_{i=1}^k \gamma_i w_i}_{\in W}$$

$$\sum_{i=1}^{k} \gamma_i w_i \in U \cap W$$
$$\sum_{i=1}^{k} \gamma_i w_i = \sum_{i=1}^{n} \delta_i v_i$$
$$\sum_{i=1}^{n} (\alpha_i + \delta_i) v_i + \sum_{i=1}^{m} \beta_i u_i = 0$$
$$\beta_i = 0, \alpha_i = -\delta_i$$
$$\sum_{i=1}^{n} \alpha_i v_i + \sum_{i=1}^{k} \gamma_i w_i = 0$$
$$\alpha_i = 0, \gamma_i = 0$$

□

Proposition 1.7

If V is a finite-dimensional F-vector space, and $U \leq V$, then U and V/U are also finite-dimensional. In particular, $\dim_F V = \dim_F U + \dim_F (V/U)$.

Proof. Let $(u_1, \dots, u_\ell)$ be a basis of U. We extend this basis to a basis of V: $(u_1, \dots, u_\ell, w_{\ell+1}, \dots, w_n)$. We claim that $(w_{\ell+1} + U, \dots, w_n + U)$ is a basis of the vector space V/U. □

Remark 6. If V is an F-vector space, and $U \leq V$, then we say U is a proper subspace if $U \neq V$. Then if U is proper, then $\dim_F U < \dim_F V$ and $\dim_F (V/U) > 0$ because $(V/U) \neq \varnothing$.

§1.12 Direct sums

Definition 1.11

Let V be an F-vector space and U, W be subspaces of V. We say that $V = U \oplus V$, read as the direct sum of U and V, if $\forall v \in V, \exists! u \in U, \exists! w \in W, u + w = v$. We say that W is *a* direct complement of U in V; there is no uniqueness of such a complement.

Lemma 1.3

Let V be an F-vector space, and $U, W \leq V$. Then the following statements are equivalent.

1. $V = U \oplus W$
2. $V = U + W$ and $U \cap W = \{0\}$
3. For any basis B_1 of U and B_2 of W, $B_1 \cup B_2$ is a basis of V

Proof. First, we show that (ii) implies (i). If $V = U + W$, then certainly $\forall v \in V, \exists u \in U, \exists w \in W, v = u + w$, so it suffices to show uniqueness. Note, $u_1 + w_1 = u_2 + w_2 \implies u_1 - u_2 = w_2 - w_1$. The left hand side is an element of U and the right hand side is an element of W, so they must be the zero vector; $u_1 = u_2, w_1 = w_2$.

Now, we show (i) implies (iii). Suppose B_1 is a basis of U and B_2 is a basis of W. Let $B = B_1 \cup B_2$. First, note that B is a generating family of $U + W$. Now we must show that B is free.

$$\underbrace{\sum_{u \in B_1} \lambda_u u}_{\in U} + \underbrace{\sum_{w \in B_2} \lambda_w w}_{\in W} = 0$$

Hence both sums must be zero. Since B_1, B_2 are bases, all λ are zero, so B is free and hence a basis.

Now it remains to show that (iii) implies (ii). We must show that $V = U + W$ and $U \cap W = \{0\}$. Now, suppose $v \in V$. Then, $v = \sum_{u \in B_1} \lambda_u u + \sum w \in B_2 \lambda_w w$. In particular, $V = U + W$, since the λ_u, λ_w are arbitrary. Now, let $v \in U \cap W$. Then

$$v = \sum_{u \in B_1} \lambda_u u = \sum_{w \in B_2} \lambda_w w \implies \lambda_u = \lambda_w = 0$$

□

Definition 1.12

Let V be an F-vector space, with subspaces $V_1, \dots, V_p \leq V$. Then

$$\sum_{i=1}^{p} V_i = \{v_1, \dots, v_\ell, v_i \in V_i, 1 \leq i \leq \ell\}$$

We say the sum is direct, written

$$\bigoplus_{i=1}^{p} V_i$$

if the decomposition is unique. Equivalently,

$$V = \bigoplus_{i=1}^{p} V_i \iff \exists! v_1 \in V_1, \ldots, v_n \in V_n, v = \sum_{i=1}^{n} v_i$$

Lemma 1.4

The following are equivalent:

1. $\sum_{i=1}^{p} V_i = \bigoplus_{i=1}^{p} V_i$
2. $\forall 1 \leq i \leq l,\ V_i \cap \left(\sum_{j \neq i} V_j\right) = \{0\}$
3. For any basis B_i of V_i, $B = \bigcup_{i=1}^{n} B_i$ is a basis of $\sum_{i=1}^{n} V_i$.

Proof. Exercise. □

§2 Linear maps

§2.1 Linear maps

Definition 2.1

If V, W are F-vector spaces, a map $\alpha\colon V \to W$ is *linear* if

$$\forall \lambda_1, \lambda_2 \in F, \forall v_1, v_2 \in V, \alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2)$$

Example 2.1

Let M be a matrix with n rows and m columns. Then the map $\alpha\colon \mathbb{R}^m \to \mathbb{R}^n$ defined by $x \mapsto Mx$ is a linear map.

Example 2.2

Let $\alpha\colon \mathcal{C}([0,1], \mathbb{R}) \to \mathcal{C}([0,1], \mathbb{R})$ defined by $f \mapsto a(f)(x) = \int_0^x f(t)\,\mathrm{d}t$. This is linear.

Example 2.3

Let $x \in [a, b]$. Then $\alpha\colon \mathcal{C}([a,b], \mathbb{R}) \to \mathbb{R}$ defined by $f \mapsto f(x)$ is a linear map.

Remark 7. Let U, V, W be F-vector spaces. Then,

1. The identity function $i_V\colon V \to V$ defined by $x \mapsto x$ is linear.
2. If $\alpha\colon U \to V$ and $\beta\colon V \to W$ are linear, then $\beta \circ \alpha$ is linear.

Lemma 2.1

Let V, W be F-vector spaces. Let B be a basis for V. If $\alpha_0\colon B \to V$ is *any* map (not necessarily linear), then there exists a unique linear map $\alpha\colon V \to W$ extending α_0: $\forall v \in B, \alpha_0(v) = \alpha(v)$.

Proof. Let $v \in V$. Then, given $B = (v_1, \dots, v_n)$.

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

By linearity,

$$\alpha(v) = \alpha\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \alpha(\lambda_i v_i) = \sum_{i=1}^{n} \alpha_0(\lambda_i v_i)$$

□

Remark 8. This lemma is also true in infinite-dimensional vector spaces. Often, to define a linear map, we instead define its action on the basis vectors, and then we 'extend by linearity' to construct the entire map.

Remark 9. If $\alpha_1, \alpha_2 \colon V \to W$ are linear maps, then if they agree on any basis of V then they are equal.

§2.2 Isomorphism

Definition 2.2

Let V, W be F-vector spaces. A map $\alpha \colon V \to W$ is an *isomorphism* if and only if

1. α is linear
2. α is bijective

If such an α exists, we say that V and W are isomorphic, written $V \cong W$.

Remark 10. If α in the above definition is an isomorphism, then $\alpha^{-1} \colon W \to V$ is linear. Indeed, if $w_1, w_2 \in W$ with $w_1 = \alpha(v_1)$ and $w_2 = \alpha(v_2)$,

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

Similarly, for $\lambda \in F, w \in W$,

$$\alpha^{-1}(\lambda w) = \lambda \alpha^{-1}(w)$$

Lemma 2.2

Isomorphism is an equivalence relation on the class of all vector spaces over F.

Proof.

1. $i_V \colon V \to V$ is an isomorphism
2. If $\alpha \colon V \to W$ is an isomorphism, $\alpha^{-1} \colon W \to V$ is an isomorphism.
3. If $\beta \colon U \to V, \alpha \colon V \to W$ are isomorphisms, then $\alpha \circ \beta \colon U \to W$ is an isomorphism.

The proofs of each part are left as an exercise. $\square$

Theorem 2.1

If V is an F-vector space of dimension n, then $V \cong F^n$.

Proof. Let $B = (v_1, \dots, v_n)$ be a basis for V. Then, consider $\alpha \colon V \to F^n$ defined by

$$v = \sum_{i=1}^{n} \lambda_i v_i \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

We claim that this is an isomorphism. This is left as an exercise. $\square$

Remark 11. Choosing a basis for V is analogous to choosing an isomorphism from V to F^n.

Theorem 2.2

Let V, W be F-vector spaces with finite dimensions n, m. Then,

$$V \cong W \iff n = m$$

Proof. If $\dim V = \dim W = n$, then there exist isomorphisms from both V and W to F^n. By transitivity, therefore, there exists an isomorphism between V and W.

Conversely, if $V \cong W$ then let $\alpha \colon V \to W$ be an isomorphism. Let B be a basis of V, then we claim that $\alpha(B)$ is a basis of W. Indeed, $\alpha(B)$ spans W from the surjectivity of α, and $\alpha(B)$ is free due to injectivity. $\square$

§2.3 Kernel and image

Definition 2.3

Let V, W be F-vector spaces. Let $\alpha \colon V \to W$ be a linear map. We define the kernel and image as follows.

$$N(\alpha) = \ker \alpha = \{v \in V : \alpha(v) = 0\}$$

$$\operatorname{Im}(\alpha) = \{w \in W : \exists v \in V, w = \alpha(v)\}$$

Lemma 2.3

$\ker \alpha$ is a subspace of V, and $\operatorname{Im} \alpha$ is a subspace of W.

Proof. Let $\lambda_1, \lambda_2 \in F$ and $v_1, v_2 \in \ker \alpha$. Then

$$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) = 0$$

Hence $\lambda_1 v_1 + \lambda_2 v_2 \in \ker \alpha$.

Now, let $\lambda_1, \lambda_2 \in F$, $v_1, v_2 \in V$, and $w_1 = \alpha(v_1), w_2 = \alpha(v_2)$. Then

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) = \alpha(\lambda_1 v_1 + \lambda_2 v_2) \in \operatorname{Im} \alpha$$

$\square$

Remark 12. $\alpha \colon V \to W$ is injective if and only if $\ker \alpha = \{0\}$. Further, $\alpha \colon V \to W$ is surjective if and only if $\operatorname{Im} \alpha = W$.

Theorem 2.3

Let V, W be F-vector spaces. Let $\alpha \colon V \to W$ be a linear map. Then $\overline{\alpha} \colon V / \ker \alpha \to \operatorname{Im} \alpha$ defined by

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

is an isomorphism. *This is the isomorphism theorem from IA Groups.*

Proof. First, note that $\overline{\alpha}$ is well defined. Suppose $v + \ker \alpha = v' + \ker \alpha$. Then $v - v' \in \ker \alpha$, hence

$$\alpha(v - v') = 0 \implies \alpha(v) - \alpha(v') = 0$$

so $\overline{\alpha}$ is indeed well defined.

Linearity of $\overline{\alpha}$ follows from linearity of α.

Now, we show $\overline{\alpha}$ is injective.

$$\overline{\alpha}(v + \ker \alpha) = 0 \implies \alpha(v) = 0 \implies v \in \ker \alpha$$

Hence, $v + \ker \alpha = 0 + \ker \alpha$.

Further, $\overline{\alpha}$ is surjective as if $w \in \operatorname{Im} \alpha$, $\exists\, v \in V$ s.t. $w = \alpha(v) = \overline{\alpha}(v + \ker \alpha)$. $\square$

§2.4 Rank and nullity

Definition 2.4

Rank and nullity The *rank* of α is

$$r(\alpha) = \dim \operatorname{Im} \alpha.$$

The *nullity* of α is

$$n(\alpha) = \dim \ker \alpha.$$

Theorem 2.4 (Rank-nullity theorem)

Let U, V be F-vector spaces such that the dimension of U is finite. Let $\alpha\colon U \to V$ be a linear map. Then,

$$\dim U = r(\alpha) + n(\alpha)$$

Proof. We have proven that $U/\ker \alpha \cong \operatorname{Im} \alpha$. Hence, the dimensions on the left and right match: $\dim(U/\ker \alpha) = \dim \operatorname{Im} \alpha$.

$$\dim U - \dim \ker \alpha^{a} = \dim \operatorname{Im} \alpha$$

and the result follows. □

[a] by proposition 1.7

Lemma 2.4 (Characterisation of isomorphisms)

Let V, W be F-vector spaces with equal, finite dimension. Let $\alpha\colon V \to W$ be a linear map. Then, the following are equivalent.

1. α is injective.
2. α is surjective.
3. α is an isomorphism.

Proof. Clearly, (iii) follows from (i) and (ii) and vice versa. The rest of the proof is left as an exercise, which follows from the rank-nullity theorem. □

Example 2.4

$$
\begin{aligned}
V &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x + y + z = 0 \right\} \\
\alpha &: \mathbb{R}^3 \to \mathbb{R} \\
&\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto x + y + z \\
\implies &\ker \alpha = V \\
&\operatorname{Im} \alpha = \mathbb{R}.
\end{aligned}
$$

So by rank nullity

$$
3 = n(\alpha) + 1 \implies \dim V = 2
$$

§2.5 Space of linear maps

Let V and W be F-vector spaces. Consider the space of linear maps from V to W. Then $L(V, W) = \{\alpha \colon V \to W \text{ linear}\}$.

Proposition 2.1 (Linear maps form a vector space)
$L(V, W)$ is an F-vector space under the operation

$$
\begin{aligned}
(\alpha_1 + \alpha_2)(v) &= \alpha_1(v) + \alpha_2(v) \\
(\lambda\alpha)(v) &= \lambda(\alpha(v))
\end{aligned}
$$

Further, if V and W are finite-dimensional, then so is $L(V, W)$ with

$$
\dim_F L(V, W) = \dim_F V \dim_F W
$$

Proof. Proving that $L(V, W)$ is a vector space is left as an exercise. The dimensionality part is proven later, proposition 2.4. $\square$

§2.6 Matrices

Definition 2.5 (Matrix)
An $m \times n$ matrix over F is an array with m rows and n columns, with entries in F.

Notation. We write $M_{m\times n}(F)$ for the set of $m \times n$ matrices over F.

Proposition 2.2

$M_{m\times n}(F)$ is an F-vector space under

$$((a_{ij}) + (b_{ij})) = (a_{ij} + b_{ij});$$

$$\lambda(a_{ij}) = (\lambda a_{ij})$$

Proof. Left as an exercise □

Proposition 2.3

$\dim_F M_{m,n}(F) = mn.$

Proof. Consider the basis defined by, the 'elementary matrix' for all i, j:

$$e_{pq} = \delta_{ip}\delta_{jq}$$

Then (e_{ij}) is a basis of $M_{m\times n}(F)$, since it spans $M_{m\times n}(F)$[a] and we can show that it is free. □

[a] given $A = (a_{ij}) \in M_{n\times n}(F)$, $A = a_{ij}e_{ij}$

§2.7 Linear maps as matrices

Let V, W be F-vector spaces and $\alpha : V \to W$ be a linear map. Consider bases B of V and C of W:

$$B = (v_1, \ldots, v_n);\ C = (w_1, \ldots, w_m)$$

Then let $v \in V$. We have

$$v = \sum_{j=1}^{n} \lambda_j v_j \equiv [v]_B = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in F^n$$

where the vector given is the coordinates in basis B.

Notation. $[v]_B$ is the coordinates of v in basis B.

We can equivalently find $[w]_C$, the coordinates of w in basis C. We can now define a matrix of some linear map α in the B, C basis.

Definition 2.6 (Matrix of linear map)
The matrix representing α wrt B, C basis is

$$[\alpha]_{B,C} = \Big([\alpha(v_1)]_C, \dots, [\alpha(v_n)]_C\Big) \in M_{m\times n}(F)$$

Note. Let $[\alpha]_{B,C} = (a_{ij})$, then by definition

$$\alpha(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

Lemma 2.5
For all $v \in V$,

$$[\alpha(v)]_C = [\alpha]_{B,C} \cdot [v]_B$$

Proof. We have

$$v = \sum_{i=1}^{n} \lambda_j v_j$$

Hence

$$\alpha\left(\sum_{i=1}^{n} \lambda_j v_j\right) = \sum_{j=1}^{n} \lambda_j \alpha(v_j) = \sum_{j=1}^{n} \lambda_i \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\lambda_j\right) w_i$$

□

Lemma 2.6
Let $\beta\colon U \to V$ and $\alpha\colon V \to W$ be linear maps. Then, if A, B, C are bases of U, V, W respectively, then

$$[\alpha \circ \beta]_{A,C} = [\alpha]_{B,C} \cdot [\beta]_{A,B}$$

Proof. Let $A = [\alpha]_{B,C}$ and $B = [\beta]_{A,B}$. Consider $u_l \in A$ (basis of U). Then

$$(\alpha \circ \beta)(u_l) = \alpha(\beta(u_l))$$

giving

$$\alpha\left(\sum_j b_{jl}v_j\right) = \sum_j b_{jl}\alpha(v_j) = \sum_j b_{jl}\sum_i a_{ij}w_i = \sum_i\left(\sum_j a_{ij}b_{jl}\right)w_i$$

where $a_{ij}b_{jl}$ is the (i,l) element of AB by the definition of the product of matrices. □

Proposition 2.4

If V, W are F-vector spaces, and $\dim_F V = n, \dim_F W = m$, then

$$L(V,W) \cong M_{m\times n}(F)$$

which implies the dimensionality of $L(V,W)$ in F is $m\times n$.

Proof. Consider two bases B, C of V, W. We claim that

$$\begin{aligned}\theta\colon L(V,W) &\to M_{m\times n}(F)\\ \alpha &\mapsto [\alpha]_{B,C}\end{aligned}$$

is an isomorphism.

First, note that θ is linear.

$$[\lambda_1\alpha_1 + \lambda_2\alpha_2] = \lambda_1[\alpha_1]_{B,C} + \lambda_2[\alpha_2]_{B,C}.$$

Also, θ is surjective; consider any matrix $A = (a_{ij})$ and consider $\alpha\colon v_j \mapsto \sum_{i=1}^m a_{ij}w_i$ defined on B. Then this is certainly a linear map which extends uniquely by linearity to A, giving $[\alpha]_{B,C} = (a_{ij}) = A$[a].

Now, θ is injective since $[\alpha]_{B,C} = 0 \implies \alpha = 0$. □

[a] Proving this left as an exercise

Remark 13. If B, C are bases of V, W respectively, and $\varepsilon_B\colon V \to F^n$ is defined by $v \mapsto [v]_B$, and analogously for ε_C, then the following diagram commutes

$$\begin{tikzcd} V \arrow[r, "\alpha"] \arrow[d, "\varepsilon_B"'] & W \arrow[d, "\varepsilon_C"] \\ \mathbb{F}^n \arrow[r, "{[\alpha]_{B,C}}"'] & \mathbb{F}^m \end{tikzcd}$$

We can see that

$$[\alpha]_{B,C} \circ \varepsilon_B = \varepsilon_C \circ \alpha$$

so the operations commute.

Example 2.5

Let $\alpha\colon V \to W$ be a linear map and $Y \leq V$, where V, W are finite-dimensional. Then let $\alpha(Y) = Z \leq W$. Consider a basis B of V, such that $B' = (v_1, \dots, v_k)$ is a basis of Y completed by $B'' = (v_{k+1}, \dots, v_n)$ into $B = B' \cup B''$. Then let C be a basis of W, such that $C' = (w_1, \dots, w_\ell)$ is a basis of Z completed by $C'' = (w_{\ell+1}, \dots, w_m)$ into $C = C' \cup C''$. Then

$$[\alpha]_{B,C} = \begin{pmatrix} \alpha(v_1) & \dots & \alpha(v_k) & \alpha(v_{k+1}) & \dots & \alpha(v_n) \end{pmatrix}$$

For $1 \leq i \leq k$, $\alpha(v_i) \in Z$ since $v_i \in Y, \alpha(Y) = Z$. So the matrix has an upper-left $\ell \times k$ block A which is $\alpha\colon Y \to Z$ on the basis B', C'. We can show further that α induces a map $\overline{\alpha}\colon V/Y \to W/Z$ by $v + Y \mapsto \alpha(v) + Z$. This is well-defined; $v_1 + Y = v_2 + Y$ implies $v_1 - v_2 \in Y$ hence $\alpha(v_1 - v_2) \in Z$ as required. The bottom-right block is $[\overline{\alpha}]_{B'',C''}$.

§2.8 Change of basis

Suppose we have two bases $B = \{v_1, \dots, v_n\}, B' = \{v'_1, \dots, v'_n\}$ of V and corresponding C, C' for W. If we have a linear map $[\alpha]_{B,C}$, we are interested in finding the components of this linear map in another basis, that is,

$$[\alpha]_{B,C} \mapsto [\alpha]_{B',C'}$$

Definition 2.7 (Change of basis matrix)

The **change of basis** matrix P from B' to B is

$$P = \begin{pmatrix} [v'_1]_B & \cdots & [v'_n]_B \end{pmatrix}$$

which is the identity map in B', written

$$P = [I]_{B',B}$$

Lemma 2.7

For a vector v,

$$[v]_B = P[v]_{B'}$$

Proof. We have

$$[\alpha(v)]_C = [\alpha]_{B,C} \cdot [v]_C$$

Since $P = [I]_{B',B}$,

$$[I(v)]_B = [I]_{B',B} \cdot [v]_{B'} \implies [v]_B = P[v]_{B'}$$

as required. □

Remark 14. P is an invertible $n \times n$ square matrix. In particular,

$$P^{-1} = [I]_{B,B'}$$

Indeed,

$$\begin{aligned} [\alpha \circ \beta]_{A,C} &= [\alpha]_{B,C}[\beta]_{A,B} \\ \implies I_n &= [I \cdot I]_{B,B} = [I]_{B',B} \cdot [I]_{B,B'} \end{aligned}$$

where I_n is the $n \times n$ identity matrix.

Warning 2.1

$$\begin{aligned} P &= ([v_1']_B, \dots, [v_n']_B) \\ \implies [v]_B &= P[v]_{B'} \\ \implies [v]_{B'} &= P^{-1}[v]_B \end{aligned}$$

Proposition 2.5

If α is a linear map from V to W, and $P = [I]_{B',B}, Q = [I]_{C',C}$[a], we have

$$A' = [\alpha]_{B',C'} = [I]_{C,C'}[\alpha]_{B,C}[I]_{B,'B} = Q^{-1}AP$$

where $A = [\alpha]_{B,C}, A' = [\alpha]_{B',C'}$.

[a] P, Q invertible.

Proof.

$$\begin{aligned}
[\alpha(v)]_C &= Q[\alpha(v)]_{C'} \\
&= Q[\alpha]_{B',C'}[v]_{B'} \\
[\alpha(v)]_C &= [\alpha]_{B,C}[v]_B \\
&= AP[v]_{B'} \\
\therefore \forall v,\ QA'[v]_{B'} &= AP[v]_{B'} \\
\therefore QA' &= AP
\end{aligned}$$

as required. □

§2.9 Equivalent matrices

Definition 2.8 (Equivalent matrices)
Matrices $A, A' \in M_{m,n}(F)$ are called **equivalent** if

$$A' = Q^{-1}AP$$

for some invertible $m \times m, n \times n$ matrices Q, P.

Remark 15. This defines an equivalence relation on $M_{m,n}(F)$.

- $A = I_m^{-1}AI_n$;
- $A' = Q^{-1}AP \implies A = QA'P^{-1}$;
- $A' = Q^{-1}AP, A'' = (Q')^{-1}A'P' \implies A'' = (QQ')^{-1}A(PP')$.

Proposition 2.6
Let V, W be vector spaces over F with $\dim_F V = n$, $\dim_F W = m$. Let $\alpha\colon V \to W$ be a linear map. Then there exists a basis B of V and a basis C of W such that

$$[\alpha]_{B,C} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

so the components of the matrix are exactly the identity matrix of size r in the top-left corner, and zeroes everywhere else.

Proof. We first fix $r \in \mathbb{N}$ such that $\dim\ker\alpha = n - r$. Then we will construct a basis $\{v_{r+1}, \dots, v_n\}$ of the kernel. We extend this to a basis of the entirety of V, that is, $\{v_1, \dots, v_n\}$. Then, we want to show that

$$\{\alpha(v_1), \dots, \alpha(v_r)\}$$

is a basis of $\operatorname{Im}\alpha$. Indeed, it is a generating family:

$$\begin{aligned} v &= \sum_{i=1}^{n} \lambda_i v_i \\ \alpha(v) &= \sum_{i=1}^{n} \lambda_i \alpha(v_i) \\ &= \sum_{i=1}^{r} \lambda_i \alpha(v_i) \text{ as } v_{r+i} \in \ker\alpha \end{aligned}$$

Then if $y \in \operatorname{Im}\alpha$, there exists v such that $\alpha(v) = y$. So

$$y = \sum_{i=1}^{r} \lambda_i \alpha(v_i) \in \langle \alpha(v_1), \dots, \alpha(v_r) \rangle.$$

Further, it is a free family:

$$\begin{aligned} \sum_{i=1}^{r} \lambda_i \alpha(v_i) &= 0 \\ \alpha\left(\sum_{i=1}^{r} \lambda_i v_i\right) &= 0 \\ \sum_{i=1}^{r} \lambda_i v_i &\in \ker\alpha \\ \sum_{i=1}^{r} \lambda_i v_i &= \sum_{i=r+1}^{n} \lambda_i v_i \text{ as } v_{r+i} \text{ is a basis of } \ker\alpha. \\ \sum_{i=1}^{r} \lambda_i v_i - \sum_{i=r+1}^{n} \lambda_i v_i &= 0 \end{aligned}$$

But since $\{v_1, \dots, v_n\}$ is a basis, $\lambda_i = 0$ for all i.

Hence $\{\alpha(v_1), \dots, \alpha(v_r)\}$ is a basis of $\operatorname{Im}\alpha$. Now, we extend this basis to the whole of W to form

$$\{\alpha(v_1), \dots, \alpha(v_r), w_{r+1}, \dots, w_n\}$$

Now,

$$[\alpha]_{BC} = \begin{pmatrix} \alpha(v_1) & \cdots & \alpha(v_r) & \alpha(v_{r+1}) & \cdots & \alpha(v_n) \end{pmatrix}$$

$$= \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

□

Remark 16. This also proves the rank-nullity theorem:

$$\operatorname{rank} \alpha + \operatorname{null} \alpha = n$$

Corollary 2.1

Any $m \times n$ matrix A is equivalent to a matrix of the form

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

where $r = \operatorname{rank} A$.

§2.10 Column rank and row rank

Definition 2.9 (Column rank)

Let A[a] $\in M_{m,n}(F)$. Then, the **column rank** of A, here denoted $r_c(A)$, is the dimension of the subspace of F^n spanned by the column vectors.

$$r_c(A) = \dim \operatorname{span} \{c_1, \ldots, c_n\}$$

[a] $A = (c_1 \mid \cdots \mid c_n),\ c_n \in F^m$.

Definition 2.10 (Row rank)

The **row rank** is the column rank of $A^\intercal$.

Remark 17. If α is a linear map, represented by A with respect to some basis, then:

$$\operatorname{rank} \alpha = r_c(A) = \dim \operatorname{Im} \alpha$$

Proof. Proof of $\operatorname{rank} \alpha = r_c(A)$ is left as an exercise. □

Proposition 2.7

Two matrices are equivalent if they have the same column rank:

$$r_c(A) = r_c(A').$$

Proof. ($\implies$) If the matrices are equivalent, then they correspond to the same linear map α in two different basis

$$\begin{aligned} r_c(A) &= \operatorname{rank}\alpha \\ r_c(A') &= \operatorname{rank}\alpha \\ \implies r_c(A) &= r_c(A') \end{aligned}$$

($\impliedby$) Conversely, if $r_c(A) = r_c(A') = r$, then A, A' are equivalent to

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

By transitivity, A, A' are equivalent. $\square$

Theorem 2.5

Column rank $r_c(A)$ and row rank $r_c(A^\intercal)$ are equivalent.

Proof. Let $r = r_c(A)$. Then,

$$Q^{-1}AP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m \times n}$$

Then take the transpose:

$$\begin{aligned} (Q^{-1}AP)^\intercal &= P^\intercal A^\intercal \left(Q^{-1}\right)^\intercal \\ &= P^\intercal A^\intercal (Q^\intercal)^{-1} \\ &= \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}^\intercal_{m \times n} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{n \times m} \end{aligned}$$

Then $r_c(A^\intercal) = r = r_c(A)$.

Note. We can swap the transpose and inverse on Q because

$$\begin{aligned} (AB)^\intercal &= B^\intercal A^\intercal \\ \left(QQ^{-1}\right)^\intercal &= \left(Q^{-1}\right)^\intercal Q^\intercal \end{aligned}$$

$$
\begin{aligned}
I &= \left(Q^{-1}\right)^{\intercal} Q^{\intercal} \\
(Q^{\intercal})^{-1} &= \left(Q^{-1}\right)^{\intercal}
\end{aligned}
$$

□

So we can drop the concepts of column and row rank, and just talk about rank as a whole.

§2.11 Conjugation and similarity

Consider the following special case of changing basis.

Definition 2.11

If $\alpha \colon V \to V$ is linear, α is called an **endomorphism**.

If $B = C, B' = C'$ then the special case of the change of basis formula is

$$[\alpha]_{B',B'} = P^{-1}[\alpha]_{B,B}P$$

Definition 2.12 (Similar matrices)

Let A, A' be $n \times n$ (square) matrices. We say that A and A' are **similar** or **conjugate** iff there exists P ($n \times n$ square invertible matrix) such that $A' = P^{-1}AP$.

This is a central concept when we will study diagonalisation of matrices, Spectral theory.

§2.12 Elementary operations

Definition 2.13 (Elementary column operation)

An **elementary column operation** is

1. swap columns i, j $(i \neq j)$
2. replace column i by λ multiplied by the column $(\lambda \neq 0, \lambda \in F)$
3. add λ multiplied by column i to column j $(i \neq j)$

We define analogously the elementary row operations. Note that these elementary operations are invertible (for $\lambda \neq 0$). These operations can be realised through the action of elementary matrices. For instance, the column swap operation can be realised using

$$T_{ij} = \begin{pmatrix} I_{i-1} & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & I \end{pmatrix}; \quad A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & I & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

To multiply a column by λ,

$$n_{i,\lambda} = \begin{pmatrix} I_{i-1} & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & I \end{pmatrix}$$

To add a multiple of a column,

$$c_{ij,\lambda} = I + \lambda E_{ij}$$

where E_{ij} is the matrix defined by elements $(e_{ij})_{pq} = \delta_{ip}\delta_{jq}$.

An elementary column (or row) operation can be performed by multiplying A by the corresponding elementary matrix from the right (on the left for row operations).

Proof. Left as an exercise. □

Example 2.6

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}.$$

We can prove corollary 2.1 constructively:

Proof. This will essentially provide a constructive proof that any $m \times n$ matrix is equivalent to

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

We will start with a matrix A. If all entries are zero, we are done.
So we will pick $a_{ij} = \lambda \neq 0$, and swap rows $i, 1$ and columns $j, 1$. This ensures that $a_{11} = \lambda \neq 0$.
Now we multiply column 1 by $\frac{1}{\lambda}$ so $a_{11} = 1$ now.
Finally, we can clear out row 1 and column 1 by subtracting multiples of rows or columns (3rd elementary operation). Then we can perform similar operations on the $(m-1) \times (n-1)$ matrix in the bottom right block and inductively finish this

process. We end up with:

$$\begin{aligned}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} &= \underbrace{E'_p \dots E'_1}_{\text{row operations}} \; A \; \underbrace{E_1 \dots E_c}_{\text{column operations}} \\ &= Q^{-1}AP\end{aligned}$$

□

§2.13 Gauss' pivot algorithm

If only row operations are used, we can reach the **row echelon form** of the matrix, a specific case of an upper triangular matrix.

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

On each row, there are a number of zeroes until there is a one, called the pivot.

First, we assume that $a_{ij} \neq 0$.
We swap rows $i, 1$.
Then divide the first row by $\lambda = a_{i1}$ to get a one in the top left.
We can use this one to clear the rest of the first column.
Then, we can repeat on the next column, and iterate.

This is a technique for solving a linear system of equations.

§2.14 Representation of square invertible matrices

Lemma 2.8

If A is an $n \times n$ square invertible matrix, then we can obtain I_n using only row elementary operations, or only column elementary operations.

Proof. We show an algorithm that constructs this I_n. This is exactly going to invert the matrix, since the resultant operations can be combined to get the inverse matrix. We will show here the proof for column operations.

We argue by induction on the number of rows.

Suppose we can make the form

$$\begin{pmatrix} I_k & 0 \\ A & B \end{pmatrix}$$

We want to obtain the same structure with $k+1$ rows.

We claim that there exists $j > k$ such that $a_{k+1,j} \neq 0$. Indeed, otherwise we can show that the vector

$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \delta_{k+1,i}$$

is not in the span of the column vectors of A.[a] This contradicts the invertibility of the matrix.

Now, we will swap columns $k+1, j$ and divide this column by λ. We can now use this 1 to clear the rest of the $k+1$ row using elementary operations of type 3.

The desired results follows from induction. □

[a] Left as an exercise to check this.

Remark 18. Inductively, we have found $AE_1 \dots E_c = I_n$ where E_c are elementary. Thus, $A^{-1} = E_1 \dots E_c$ and so this is an algorithm for computing A^{-1} and so solving linear systems of equations.

Proposition 2.8

Any invertible square matrix is a product of elementary matrices.

Proof. The proof is exactly the proof of the lemma above. □

§3 Dual spaces

§3.1 Dual spaces

Definition 3.1 (Dual Space)

Let V be an F-vector space. Then $V^\star$ is the **dual** of V, defined by

$$V^\star = L(V, F) = \{\alpha \colon V \to F\}$$

where the α are linear.
If $\alpha \colon V \to F$ is linear, then we say α is a **linear form**. So the dual of V is the set of linear forms on V.

Example 3.1

For instance, the trace $\operatorname{tr} \colon M_{n,n}(F) \to F$ is a linear form on $M_{n,n}(F)$. So $\operatorname{tr} \in M^*_{n,n}(F)$

Example 3.2

Consider functions $f : [0, 1] \to \mathbb{R}$. We can define $T_f \colon \mathcal{C}^\infty([0, 1], \mathbb{R}) \to \mathbb{R}$ such that $\varphi \mapsto \int_0^1 f(x)\varphi(x)\,\mathrm{d}x$. I.e. $T_f(\varphi) = \int_0^1 f(x)\varphi(x)\,\mathrm{d}x$.
Then T_f is a linear form on $\mathcal{C}^\infty([0, 1], \mathbb{R})$ ($\mathbb{R}$ vector space).
The function defines a linear form. We can then reconstruct f given T_f. This mathematical formulation is called distribution (which is about the generalisation of the notion of functions).

Remark 19. Duality is not that useful in finite dimensions but it is in infinite.

Lemma 3.1 (Dual Basis)

Let V be an F-vector space with a finite basis $B = \{e_1, \dots, e_n\}$. Then there exists a basis $B^\star$ for $V^\star$ given by

$$B^\star = \{\varepsilon_1, \dots, \varepsilon_n\}; \quad \varepsilon_j{}^{a}\left(\sum_{i=1}^n a_i e_i\right) = a_j$$

We call $B^\star$ the **dual basis** for B.

[a] Recall ε_j is a linear form

Remark 20. Kronecker symbol, δ_{ij}.

$$\varepsilon_j\left(\sum_{i=1}^{n} a_i e_i\right) = a_j \iff \varepsilon_j(e_i) = \delta_{ij}$$

Proof. Let

$$\varepsilon_j(e_i) = \delta_{ij}$$

First, we will show that the set of linear forms as defined is free. For all i,

$$\begin{aligned}
\sum_{j=1}^{n} \lambda_j \varepsilon_j &= 0 \\
\therefore \left(\sum_{j=1}^{n} \lambda_j \varepsilon_j\right) e_i &= 0 \\
\sum_{j=1}^{n} \lambda_j \underbrace{\varepsilon_j(e_i)}_{\delta_{ij}} &= 0 \\
\lambda_i &= 0
\end{aligned}$$

Now we show that the set spans $V^\star$. Suppose $\alpha \in V^\star$, $x \in V$.

$$\begin{aligned}
\alpha(x) &= \alpha\left(\sum_{j=1}^{n} \lambda_j e_j\right) \\
&= \sum_{i=1}^{n} \lambda_j \alpha(e_j)
\end{aligned}$$

Conversely, we can write

$$\sum_{i=1}^{n} \underbrace{\alpha(e_j)}_{\in F} \varepsilon_j \in V^\star$$

Thus,

$$\begin{aligned}
\left(\sum_{i=1}^{n} \alpha(e_j)\varepsilon_j\right)(x) &= \sum_{j=1}^{n} \alpha(e_j)\varepsilon_j\left(\sum_{k=1}^{n} \lambda_k e_k\right) \\
&= \sum_{j=1}^{n} \alpha(e_j) \sum_{k=1}^{n} \lambda_k \varepsilon_j(e_k) \\
&= \sum_{j=1}^{n} \alpha(e_j) \sum_{k=1}^{n} \lambda_k \delta_{jk}
\end{aligned}$$

$$= \sum_{j=1}^{n} \alpha(e_j)\lambda_j$$
$$= \alpha(x)$$

We have then shown that

$$\alpha = \sum_{j=1}^{n} \alpha(e_j)\varepsilon_j$$

as required. □

Corollary 3.1

If V is finite-dimensional, $V^\star$ has the same dimension.[a]

[a] Very different in infinite dimension.

Remark 21. It is sometimes convenient to think of $V^\star$ as the spaces of row vectors of length $\dim V$ over F. For instance, consider the basis $B = (e_1, \dots, e_n)$, so $x = \sum_{i=1}^n x_i e_i$. Then we can pick $(\varepsilon_1, \dots, \varepsilon_n)$ a basis of $V^\star$, so $\alpha = \sum_{i=1}^n \alpha_i \varepsilon_i$. Then

$$\alpha(x) = \sum_{i=1}^{n} \alpha_i \varepsilon_i(x) = \sum_{i=1}^{n} \alpha_i \varepsilon \left(\sum_{j=1}^{n} x_j e_j \right) = \sum_{i=1}^{n} \alpha_i x_i$$

This is exactly

$$\begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

which essentially defines a scalar product between the two spaces.

§3.2 Annihilators

Definition 3.2 (Annihilator)

Let $U \le V$. Then the **annihilator** of U is

$$U^0 = \{\alpha \in V^\star \colon \forall u \in U, \alpha(u) = 0\}$$

Lemma 3.2 1. $U^0 \le V^\star$;

2. If $U \leq V$ and $\dim V < \infty$, then $\dim V = \dim U + \dim U^0$.

Proof. 1. First, note that $0 \in U^0$. If $\alpha, \alpha' \in U^0$, then for all $u \in U$,

$$(\alpha + \alpha')(u) = \alpha(u) + \alpha'(u) = 0$$

Further, for all $\lambda \in F$,

$$(\lambda\alpha)(u) = \lambda\alpha(u) = 0$$

Hence $U^0 \leq V^\star$.

2. Let $U \leq V$ and $\dim V = n$. Let $(e_1, \ldots, e_k)$ be a basis of U, completed into a basis $B = (e_1, \ldots, e_k, e_{k+1}, \ldots, e_n)$ of V. Let $(\varepsilon_1, \ldots, \varepsilon_n)$ be the dual basis $B^\star$. We then will prove that

$$U^0 = \langle \varepsilon_{k+1}, \ldots, \varepsilon_n \rangle$$

Pick $i > k$, then $\varepsilon_i(e_j) = \delta_{ij} = 0$ for $1 \leq j \leq k$. Hence $\varepsilon_i \in U^0$. Thus $\langle \varepsilon_{k+1}, \ldots, \varepsilon_n \rangle \subset U^0$.
Conversely, let $\alpha \in U^0$. Then $\alpha = \sum_{i=1}^n \alpha_i \varepsilon_i$. For $i \leq k$, $\alpha \in U^0$ hence $\alpha(e_i) = 0$ for $1 \leq i \leq k$. Hence,

$$\alpha = \sum_{i=k+1}^{n} \alpha_i \varepsilon_i$$

Thus

$$\alpha \in \langle \varepsilon_{k+1}, \ldots, \varepsilon_n \rangle$$

so $U^0 \subset \langle \varepsilon_{k+1}, \ldots, \varepsilon_n \rangle$ as required.

□

§3.3 Dual maps

Lemma 3.3 (Dual Map)
Let V, W be F-vector spaces. Let $\alpha \in L(V, W)$. Then there exists a unique $\alpha^\star \in L(W^\star, V^\star)$

$$\begin{aligned} \alpha^\star : W^\star &\to V^\star \\ \varepsilon &\mapsto \varepsilon \circ \alpha \end{aligned}$$

called the **dual map**.

Proof. First, note $\varepsilon(\alpha)\colon V \to F$ is a linear map. Hence, $\varepsilon \circ \alpha \in V^\star$. Now we must show $\alpha^\star$ is linear.

$$\alpha^\star(\theta_1 + \theta_2) = (\theta_1 + \theta_2)(\alpha) = \theta_1 \circ \alpha + \theta_2 \circ \alpha = \alpha^\star(\theta_1) + \alpha^\star(\theta_2)$$

Similarly, we can show

$$\alpha^\star(\lambda\theta) = \lambda\alpha^\star(\theta)$$

as required. Hence $\alpha^\star \in L(W^\star, V^\star)$. □

Proposition 3.1

Let V, W be finite-dimensional F-vector spaces with bases B, C respectively. Let $B^\star, C^\star$ be the dual basis of $V^\star, W^\star$. Then

$$[\alpha^\star]_{C^\star, B^\star} = [\alpha]^\intercal_{B,C}$$

Thus, we can think of the dual map as the *adjoint* of α.

Proof. This follows from the definition of the dual map. Let $B = (b_1, \dots, b_n)$, $C = (c_1, \dots, c_m)$, $B^\star = (\beta_1, \dots, \beta_n)$, $C^\star = (\gamma_1, \dots, \gamma_m)$. Let $[\alpha]_{B,C} = (a_{ij})$. Recall $\alpha^\star : W^\star \to V^\star$. Then, we compute

$$\begin{aligned}
\alpha^\star(\gamma_r)(b_s) &= \underbrace{\gamma_r}_{\in W^\star} \circ \underbrace{\alpha(b_s)}_{\in W} \\
&= \gamma_r \underbrace{\left(\sum_t a_{ts} c_t\right)}_{s\text{th column vector}} \\
&= \sum_t a_{ts} \gamma_r(c_t) \\
&= \sum_t a_{ts} \delta_{tr} \\
&= a_{rs}
\end{aligned}$$

We can conversely write $[\alpha^\star]_{C^\star, B^\star} = (m_{ij})$ and

$$\begin{aligned}
\alpha^\star(\gamma_r) &= \sum_{i=1}^n m_{ir} \beta_i \\
\alpha^\star(\gamma_r)(b_s) &= \sum_{i=1}^n m_{ir} \beta_i(b_s)
\end{aligned}$$

$$= \sum_{i=1}^{n} m_{ir}\delta_{is}$$

$$= m_{sr}$$

Thus,

$$a_{rs} = m_{sr}$$

as required. □

§3.4 Properties of the dual map

Let $\alpha \in L(V, W)$, and $\alpha^\star \in L(W^\star, V^\star)$. Let B and C be bases of V, W respectively, and $B^\star, C^\star$ be their duals. We have proven that

$$[\alpha]_{B,C} = [\alpha^\star]^\intercal_{B,C}$$

Lemma 3.4

Suppose that $E = (e_1, \dots, e_n)$ and $F = (f_1, \dots, f_n)$ are bases of V. Let $P = [I]_{F,E}$ be a change of basis matrix from F to E. The bases $E^\star = (\varepsilon_1, \dots, \varepsilon_n)$, $F^\star = (\eta_1, \dots, \eta_n)$ are the corresponding dual bases.
Then, the change of basis matrix from $F^\star$ to $E^\star$ is

$$\left(P^{-1}\right)^\intercal$$

Proof. Consider

$$[I]_{F^\star,E^\star} = [I]^\intercal_{E,F} = \left([I]^{-1}_{F,E}\right)^\intercal = \left(P^{-1}\right)^\intercal$$

□

Lemma 3.5

Let V, W be F-vector spaces. Let $\alpha \in L(V, W)$. Let $\alpha^\star \in L(W^\star, V^\star)$ be the corresponding dual map. Then, denoting $N(\alpha)$ for the kernel of α,

1. $N(\alpha^\star) = (\operatorname{Im}\alpha)^0$[a], so $\alpha^\star$ is injective if and only if α is surjective.
2. $\operatorname{Im}\alpha^\star \le (N(\alpha))^0$, with equality if V, W are finite-dimensional. In this finite-dimensional case, $\alpha^\star$ is surjective if and only if α is injective.

[a]The annihilator of $\operatorname{Im}\alpha$

Remark 22. This is a fundamental property.
In many applications (especially in infinite dimensions) e.g. controllability, it is often simpler to understand the dual map $\alpha^\star$ than it is to understand α.

Proof. First, we prove (i). Let $\varepsilon \in W^\star$. Then, $\varepsilon \in N(\alpha^\star) \iff \alpha^\star(\varepsilon) = 0$. Hence, $\alpha^\star(\varepsilon) = \varepsilon \circ \alpha = 0$. So for any $v \in V$, $\varepsilon(\alpha(v)) = 0$. Equivalently, ε is an element of the annihilator of $\operatorname{Im}\alpha$.

Now, we will show (ii). Let $\varepsilon \in \operatorname{Im}\alpha^\star$. Then $\alpha^\star(\varphi) = \varepsilon$ for some $\varphi \in W^\star$. Then, for all $u \in N(\alpha)$, $\varepsilon(u) = (\alpha^\star(\varphi))(u) = \varphi \circ \alpha(u) = \varphi(\alpha(u)) = 0$. Certainly then $\varepsilon \in (N(\alpha))^0$. Then, $\operatorname{Im}\alpha^\star \leq (N(\alpha))^0$.

In the finite-dimensional case, we can compare the dimension of these two spaces.

$$\dim \operatorname{Im}\alpha^\star = r(\alpha^\star) = r([\alpha^\star]_{C^\star,B^\star}) = r\left([\alpha]_{B,C}^\intercal\right) = r([\alpha]_{B,C}) = r(\alpha) = \dim \operatorname{Im}\alpha$$

Due to the rank-nullity theorem, $\dim \operatorname{Im}\alpha^\star = \dim \operatorname{Im}\alpha = \dim V - \dim N(\alpha) = \dim\left[(N(\alpha))^0\right]$ by lemma 3.2. Hence,

$$\operatorname{Im}\alpha^\star \leq (N(\alpha))^0; \quad \dim \operatorname{Im}\alpha^\star = \dim(N(\alpha))^0$$

The dimensions are equal, and one is a subspace of the other, hence the spaces are equal. □

§3.5 Double duals

Definition 3.3 (Double Dual)

Let V be an F-vector space. Let $V^\star$ be the dual of V. The **double dual** or **bidual** of V is

$$V^{\star\star} = L(V^\star, F) = (V^\star)^\star$$

Remark 23. This is a very important space in infinite dimensions.
In general, there is no obvious relation between V and $V^\star$ (unless Hilbertian structure). However, the following useful facts hold about V and $V^{\star\star}$.

1. There is a large class of function spaces where $V \cong V^{\star\star}$. These are called **reflexive spaces**.

Example 3.3

$p > r$, $L^p(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : \int_{\mathbb{R}} |f(x)|^p \, \mathrm{d}x < +\infty\}$. This is a reflexive space (this uses the Lebesgue integral as this space is not complete using Riemann integral.)

Such spaces are investigated in the study of Banach spaces.

2. There is a **canonical embedding** from V to $V^{\star\star}$. In particular, there exists i in $L(V, V^{\star\star})$ which is injective.

Theorem 3.1

V embeds into $V^{\star\star}$.

Proof. Choose a vector $v \in V$ and define the linear form $\hat{v} \in L(V^\star, F)$ such that

$$\hat{v}(\varepsilon) = \varepsilon(v)$$

We want to show $\hat{v} \in V^{\star\star}$. If $\varepsilon \in V^\star, \varepsilon(v) \in F$. Further, $\lambda_1, \lambda_2 \in F$ and $\varepsilon_1, \varepsilon_2 \in V^\star$ give

$$\hat{v}(\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2) = (\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2)(v) = \lambda_1\varepsilon_1(v) + \lambda_2\varepsilon_2(v) = \lambda_1\hat{v}(\varepsilon_1) + \lambda_2\hat{v}(\varepsilon_2)$$

$\square$

Theorem 3.2

If V is a finite-dimensional vector space over F, then $i \colon V \to V^{\star\star}$ given by $i(v) = \hat{v}$ is an isomorphism[a].

[a]In infinite dimension, we can show under canonical assumptions (Banach space) that this is an injection.

Proof. We will show i is linear. If $v_1, v_2 \in V, \lambda_1, \lambda_2 \in F, \epsilon \in V^\star$, then

$$i(\lambda_1 v_1 + \lambda_2 v_2)(\varepsilon) = \varepsilon(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1\varepsilon(v_1) + \lambda_2\varepsilon(v_2) = \lambda_1\hat{v}_1(\varepsilon) + \lambda_2\hat{v}_2(\varepsilon).$$

Now, we will show that i is injective for finite-dimensional V. Let $e \in V \setminus \{0\}$. We will show that $e \notin \ker i$. We extend e into a basis $(e, e_2, \dots, e_n)$ of V. Now, let $(\varepsilon, \varepsilon_2, \dots, \varepsilon_n)$ be the dual basis. Then $\hat{e}(\varepsilon) = \varepsilon(e) = 1$. In particular, $\hat{e} \neq 0$. Hence $\ker i = \{0\}$, so it is injective.

We now show that i is an isomorphism. We need to simply compute the dimension of the image under i. Certainly, $\dim V = \dim V^\star = \dim (V^\star)^\star = \dim V^{\star\star}$. Since i is injective, $\dim V = \dim V^{\star\star}$. So i is surjective as required. $\square$

Lemma 3.6

Let V be a finite-dimensional F-vector space. Let $U \leq V$. Then,

$$\hat{U}^{a} = U^{00}$$

After identifying V and $V^{\star\star}$, we typically say

$$U = U^{00}$$

although this is is incorrect notation and not an equality (but an isomorphism).

[a] Image of U under i map

Proof. We will show that $\hat{U} \leq U^{00}$. Indeed, let $u \in U$, then by definition

$$\forall \varepsilon \in U^0, \varepsilon(u) = 0 \implies \hat{u}(\varepsilon) = 0$$

Hence $\hat{u} \in U^{00}$ and so $\hat{U} \leq U^{00}$.

Now, we will compute dimension: $\dim U^{00} = \dim V - \dim U^0 = \dim U$. Since $\hat{U} \cong U$, their dimensions are the same, so $U^{00} = \hat{U}$. □

Remark 24. Due to this identification of $V^{\star\star}$ and V, we can define

$$T \leq V^\star, T^0 = \{v \in V : \forall \theta \in T, \theta(v) = 0\}$$

Lemma 3.7

Let V be a finite-dimensional F-vector space. Let U_1, U_2 be subspaces of V. Then

1. $(U_1 + U_2)^0 = U_1^0 \cap U_2^0$;
2. $(U_1 \cap U_2)^0 = U_1^0 + U_2^0$

Proof. Let $\theta \in V^\star$. Then $\theta \in (U_1 + U_2)^0 \iff \forall u_1 \in U_1, u_2 \in U_2, \theta(u_1 + u_2) = 0$. Iff $\theta(u) = 0$ for all $u \in U_1 \cup U_2$ by linearity. Iff $\theta \in U_1^0 \cap U_2^0$.

Now, take the annihilator of (i) and $U^{00} = U$ to complete part (ii). □

§4 Bilinear Forms

§4.1 Introduction

Definition 4.1 (Bilinear Forms)
Let U, V be F-vector spaces. Then $\varphi\colon U \times V \to F$ is a **bilinear form** if it is 'linear in both components'. For example, φ at a fixed $u \in U$ is a linear form $V \to F$ and an element of $V^\star$; and φ at a fixed $v \in V$ is a linear form $U \to F$ and an element of $U^\star$

Example 4.1
Consider the map $V \times V^\star \to F$ given by

$$(v, \theta) \mapsto \theta(v).$$

You can check this is a bilinear map.

Example 4.2 (Scalar Product)
The scalar product on $U = V = \mathbb{R}^n$ is given by

$$\begin{aligned} \psi : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (x, y) &\mapsto \sum_{i=1}^{n} x_i y_i \end{aligned}$$

You can check this is a bilinear map.

Example 4.3
Let $U = V = C([0, 1], \mathbb{R})$ and consider

$$\varphi(f, g) = \int_0^1 f(t) g(t) \,\mathrm{d}t$$

You can check this is a bilinear map.

Definition 4.2 (Matrix of a bilinear form in a basis)
If $B = (e_1, \dots, e_m)$ is a basis of U and $C = (f_1, \dots, f_n)$ is a basis of V, and $\varphi\colon U \times V \to F$ is a bilinear form, then the **matrix of the bilinear form in this**

basis is

$$[\varphi]_{B,C} = \left(\underbrace{\varphi(e_i, f_j)}_{\in F} \right)_{1 \le i \le m, 1 \le j \le n}$$

Lemma 4.1

We can link φ with its matrix in a given basis as follows.

$$\varphi(u,v) = [u]_B^\intercal [\varphi]_{B,C} [v]_C$$

Proof. Let $u = \sum_{i=1}^m \lambda_i e_i$ and $v = \sum_{j=1}^n \mu_j f_j$. Then by linearity:

$$\varphi(u,v) = \varphi\left(\sum_{i=1}^m \lambda_i e_i, \sum_{j=1}^n \mu_j f_j \right) = \sum_{i=1}^m \sum_{j=1}^n \lambda_i \mu_j \varphi(e_i, f_j) = [u]_B^\intercal [\varphi]_{B,C} [v]_C.$$

Check these equality signs are correct. □

Remark 25. Note that $[\varphi]_{B,C}$ is the only matrix such that $\varphi(u,v) = [u]_B^\intercal [\varphi]_{B,C} [v]_C$.

Definition 4.3

Let $\varphi \colon U \times V \to F$ be a bilinear form. Then φ induces two linear maps given by the partial application of a single parameter to the function.

$$\varphi_L \colon U \to V^\star; \quad \varphi_L(u) \colon V \to F; \quad v \mapsto \varphi(u,v)$$

$$\varphi_R \colon V \to U^\star; \quad \varphi_R(v) \colon U \to F; \quad u \mapsto \varphi(u,v)$$

In particular,

$$\varphi_L(u)(v) = \varphi(u,v) = \varphi_R(v)(u)$$

Lemma 4.2

Let $B = (e_1, \dots, e_m)$ be a basis of U, and let $B^\star = (\varepsilon_1, \dots, \varepsilon_m)$ be its dual; and let $C = (f_1, \dots, f_n)$ be a basis of V, and let $C^\star = (\eta_1, \dots, \eta_n)$ be its dual. Let $A = [\varphi]_{B,C}$. Then

$$[\varphi_R]_{C,B^\star} = A; \quad [\varphi_L]_{B,C^\star} = A^\intercal$$

Proof.

$$\varphi_L(e_i)(f_j) = \varphi(e_i, f_j) = A_{ij}$$

Since η_j is the dual of f_j,

$$\varphi_L(e_i) = \sum_i A_{ij}\eta_j$$

Further,

$$\varphi_R(f_j)(e_i) = \varphi(e_i, f_j) = A_{ij}$$

and then similarly

$$\varphi_R(f_j) = \sum_i A_{ij}\varepsilon_i$$

□

Definition 4.4 (Left/ Right Kernel)

$\ker \varphi_L$ is called the **left kernel** of φ. $\ker \varphi_R$ is the **right kernel** of φ.

Definition 4.5 (Degenerate/ Non-Degenerate Bilinear Form)

We say that φ is **non-degenerate** if $\ker \varphi_L = \ker \varphi_R = \{0\}$. Otherwise, φ is **degenerate**.

Lemma 4.3

Let B be a basis of U, and let C be a basis of V, where U, V are finite-dimensional.
Let $\varphi\colon U \times V \to F$ be a bilinear form. Let $A = [\varphi]_{B,C}$.
Then, φ is non-degenerate if and only if A is invertible.

Corollary 4.1

If φ is non-degenerate, then $\dim U = \dim V$.

Proof. Suppose φ is non-degenerate. Then $\ker \varphi_L = \ker \varphi_R = \{0\}$. This is equivalent to saying that $n(\varphi_L) = n(\varphi_R) = 0$. We can use the rank-nullity theorem to state that $r(A^\intercal) = \dim U$ and $r(A) = \dim V$. This is equivalent to saying that A is invertible. Note that this forces $\dim U = \dim V$ as $r(A^\intercal) = r(A)$. □

Remark 26. The canonical example of a non-degenerate bilinear form is the scalar product $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ represented by the identity matrix in the standard basis[1].

Corollary 4.2

If U and V are finite-dimensional with $\dim U = \dim V$, then choosing a non-degenerate bilinear form $\varphi \colon U \times V \to F$ is equivalent to choosing an isomorphism $\varphi_L \colon U \to V^\star$.

Definition 4.6

If $T \subset U$, then we define

$$T^\perp = \{v \in V \colon \forall t \in T, \varphi(t,v) = 0\}^a$$

Further, if $S \subset V$, we define

$$^\perp S = \{u \in U \colon \forall s \in S, \varphi(u,s) = 0\}$$

These are called the **orthogonals** of T and S.

[a] $\varphi : (U,V) \to F$.

§4.2 Change of basis for bilinear forms

Proposition 4.1 (Change of basis for bilinear forms)

Let B, B' be bases of U and $P = [I]_{B',B}$, let C, C' be bases of V and $Q = [I]_{C',C}$, and finally let $\varphi \colon U \times V \to F$ be a bilinear form. Then

$$[\varphi]_{B',C'} = P^\intercal [\varphi]_{B,C} Q$$

Proof. We have $\varphi(u,v) = [u]_B^\intercal [\varphi]_{B,C} [v]_C$. Changing coordinates, we have

$$\varphi(u,v) = (P[u]_{B'})^\intercal [\varphi]_{B,C} (Q[v]_{C'}) = [u]_{B'}^\intercal (P^\intercal [\varphi]_{B,C} Q)[v]_{C'}{}^a$$

□

[a] There is only one matrix A s.t. $\varphi(u,v) = [u]_{B'}^\intercal A [v]_{C'}$, see earlier remark.

Lemma 4.4

The **rank** of a bilinear form φ, denoted $r(\varphi)$ is the rank of any matrix representing

[1] $[\varphi]_{B,B} = I$ where B the standard bases as $\varphi(e_i, e_j) = \delta_{ij}$

φ. This quantity is well-defined.

Proof. For any invertible matrices P, Q, $r(P^\intercal AQ) = r(A)$. $\square$

Remark 27. $r(\varphi) = r(\varphi_R) = r(\varphi_L)$, since $r(A) = r(A^\intercal)$.

We will see more applications later in the course, especially when we say scalar products.

§5 Determinant and Traces

§5.1 Trace

Definition 5.1 (Trace)

The **trace** of a square matrix $A \in M_{n,n}(F) \equiv M_n(F)$ is defined by

$$\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}$$

Remark 28.

$$\begin{aligned} M_n(F) &\to F \\ A &\mapsto \operatorname{tr} A \end{aligned}$$

The trace is a linear form.

Lemma 5.1

$\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for any matrices $A, B \in M_n(F)$.

Proof. We have

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ji}a_{ij} = \operatorname{tr}(BA)$$

□

Corollary 5.1

Similar matrices have the same trace.

Proof.

$$\operatorname{tr}\left(P^{-1}AP\right) = \operatorname{tr}\left(AP^{-1}P\right) = \operatorname{tr} A$$

□

Definition 5.2

If $\alpha\colon V \to V$ is linear, we can define the trace of α as

$$\operatorname{tr} \alpha = \operatorname{tr}[\alpha]_B$$

for any basis B. This is well-defined by the corollary above.

Lemma 5.2

If $\alpha\colon V \to V$ is linear, $\alpha^\star\colon V^\star \to V^\star$ satisfies

$$\operatorname{tr}\alpha = \operatorname{tr}\alpha^\star$$

Proof.

$$\operatorname{tr}\alpha = \operatorname{tr}[\alpha]_B = \operatorname{tr}[\alpha]_B^\intercal = \operatorname{tr}[\alpha^\star]_{B^\star} = \operatorname{tr}\alpha^\star$$

□

§5.2 Permutations and transpositions

Recall the following facts about permutations and transpositions. S_n is the group of permutations of the set $\{1,\dots,n\}$; the group of bijections $\sigma\colon \{1,\dots,n\} \to \{1,\dots,n\}$. A transposition $\tau_{k\ell} = (k,\ell)$ is defined by $k \mapsto \ell, \ell \mapsto k, x \mapsto x$ for $x \neq k,\ell$. Any permutation σ can be decomposed as a product of transpositions. This decomposition is not necessarily unique, but the parity of the number of transpositions is well-defined. We say that the signature of a permutation, denoted $\varepsilon\colon S_n \to \{-1,1\}$, is 1 if the decomposition has even parity and -1 if it has odd parity. We can then show that ε is a homomorphism.

§5.3 Determinant

Definition 5.3

Let $A \in M_n(F)$. We define

$$\det A = \sum_{\sigma\in S_n} \varepsilon(\sigma) A_{\sigma(1)1} \dots A_{\sigma(n)n}$$

Example 5.1

Let $n = 2$. Then,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \implies \det A = a_{11}a_{22} - a_{12}a_{21}$$

Lemma 5.3

If $A = (a_{ij})$ is an upper (or lower) triangular matrix (with zeroes on the diagonal), then $\det A = 0$.

Proof. Let $(a_{ij}) = 0$ for $i > j$. Then

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}$$

For the summand to be nonzero, $\sigma(j) \leq j$ for all j. Thus,

$$\det A = a_{11} \dots a_{nn} = 0$$

$\square$

Lemma 5.4

Let $A \in M_n(F)$. Then, $\det A = \det A^\intercal$.

Proof.

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} \\
&= \sum_{\sigma^{-1} \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma^{-1}) a_{1\sigma(1)} \dots a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)} \\
&= \det A^\intercal
\end{aligned}$$

$\square$

§5.4 Volume forms

Definition 5.4

A volume form d on F^n is a function $d\colon \underbrace{F^n \times \dots \times F^n}_{n \text{ times}} \to F$ satisfying

1. d is multilinear: for all $i \in \{1, \dots, n\}$ and for all $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n \in F^n$,

the map from F^n to F defined by

$$v \mapsto (v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_n)$$

is linear. In other words, this map is an element of $(F^n)^\star$.

2. d is alternating: for $v_i = v_j$ for some $i \neq j$, $d = 0$.

So an alternating multilinear form is a volume form. We want to show that, up to multiplication by a scalar, the determinant is the only volume form.

Lemma 5.5

The map $(F^n)^n \to F$ defined by $(A^{(1)}, \dots, A^{(n)}) \mapsto \det A$ is a volume form. This map is the determinant of A, but thought of as acting on the column vectors of A.

Proof. We first show that this map is multilinear. Fix $\sigma \in S_n$, and consider $\prod_{i=1}^n a_{\sigma(i)i}$. This product contains exactly one term in each column of A. Thus, the map $(A^{(1)}, \dots, A^{(n)}) \mapsto \prod_{i=1}^n a_{\sigma(i)i}$ is multilinear. This then clearly implies that the determinant, a sum of such multilinear maps, is itself multilinear.

Now, we show that the determinant is alternating. Let $k \neq \ell$, and $A^{(k)} = A^{(\ell)}$. Let $\tau = (k\ell)$ be the transposition exchanging k and ℓ. Then, for all $i, j \in \{1, \dots, n\}$, $a_{ij} = a_{i\tau(j)}$. We can decompose permutations into two disjoint sets: $S_n = A_n \cup \tau A_n$, where A_n is the alternating group of order n. Now, note that $\prod_{i=1}^n a_{\sigma(i)i} + \prod_{i=1}^n a_{(\tau\circ\sigma)(i)i} = 0$. So the sum over all $\sigma \in A_n$ gives zero. So the determinant is alternating, and hence a volume form. □

Lemma 5.6

Let d be a volume form. Then, swapping two entries changes the sign.

Proof. Take the sum of these two results:

$$\begin{aligned}
d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + d(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\
&= d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) \\
&+ d(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\
&+ d(v_1, \dots, v_i, \dots, v_i, \dots, v_n) \\
&+ d(v_1, \dots, v_j, \dots, v_j, v_n) \\
&= 2d(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) \\
&= 0
\end{aligned}$$

as required. □

Corollary 5.2

If $\sigma \in S_n$ and d is a volume form, $d(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma) d(v_1, \dots, v_n)$.

Proof. We can decompose σ as a product of transpositions $\prod_{i=1}^{n_\sigma} e_i$. □

Theorem 5.1

Let d be a volume form on F^n. Let A be a matrix whose columns are $A^{(i)}$. Then

$$d(A^{(1)}, \dots, A^{(n)}) = \det A \cdot d(e_1, \dots, e_n)$$

So there is a unique volume form up to a constant multiple. We can then see that $\det A$ is the only volume form such that $d(e_1, \dots, e_n) = 1$.

Proof.

$$d(A^{(1)}, \dots, A^{(n)}) = d\left(\sum_{i=1}^n a_{i1} e_i, A^{(2)}, \dots, A^{(n)}\right)$$

Since d is multilinear,

$$d(A^{(1)}, \dots, A^{(n)}) = \sum_{i=1}^n a_{i1} d\left(e_i, A^{(2)}, \dots, A^{(n)}\right)$$

Inductively on all columns,

$$d(A^{(1)}, \dots, A^{(n)}) = \sum_{i=1}^n \sum_{j=1}^n a_{i1} a_{j2} d\left(e_i, e_j, A^{(3)}, \dots, A^{(n)}\right) = \dots = \sum_{1 \leq i_1, \leq \dots \leq n} \prod_{k=1}^n a_{i_\ell k} d(e_{i_1}, \dots e_{i_n})$$

Since d is alternating, we know that for $d(e_{i_1}, \dots, e_{i_n})$ to be nonzero, the i_k must be different, so this corresponds to a permutation $\sigma \in S_n$.

$$d(A^{(1)}, \dots, A^{(n)}) = \sum_{\sigma \in S_n} \prod_{k=1}^n a_{\sigma(k)k} \varepsilon(\sigma) d(e_1, \dots, e_n)$$

which is exactly the determinant up to a constant multiple. □

§5.5 Multiplicative property of determinant

Lemma 5.7

Let $A, B \in M_n(F)$. Then $\det(AB) = \det(A)\det(B)$.

Proof. Given A, we define the volume form $d_A \colon (F^n)^n \to F$ by

$$d_A(v_1, \dots, v_n) \mapsto \det(Av_1, \dots, Av_n)$$

$v_i \mapsto Av_i$ is linear, and the determinant is multilinear, so d_A is multilinear. If $i \neq j$ and $v_i = v_j$, then $\det(\dots, Av_i, \dots, Av_j, \dots) = 0$ so d_A is alternating. Hence d_A is a volume form. Hence there exists a constant C_A such that $d_A(v_1, \dots, v_n) = C_A \det(v_1, \dots, v_n)$. We can compute C_A by considering the basis vectors; $Ae_i = A_i$ where A_i is the ith column vector of A. Then,

$$C_A = d_A(e_1, \dots, e_n) = \det(Ae_1, \dots, Ae_n) = \det A$$

Hence,

$$\det(AB) = d_A(B) = \det A \det B$$

□

§5.6 Singular and non-singular matrices

Definition 5.5

Let $A \in M_n(F)$. We say that

1. A is *singular* if $\det A = 0$;
2. A is *non-singular* if $\det A \neq 0$.

Lemma 5.8

If A is invertible, it is non-singular.

Proof. If A is invertible, there exists A^{-1}. Then, since the determinant is a homomorphism,

$$\det\left(AA^{-1}\right) = \det I = 1$$

Thus $\det A \det A^{-1} = 1$ and hence neither of these determinants can be zero. □

Theorem 5.2

Let $A \in M_n(F)$. The following are equivalent.

1. A is invertible;

2. A is non-singular;

3. $r(A) = n$.

Proof. We have already shown that (i) implies (ii). We have also shown that (i) and (iii) are equivalent by the rank-nullity theorem. So it suffices to show that (ii) implies (iii).

Suppose $r(A) < n$. Then we will show A is singular. We have $\dim \operatorname{span}(A_1, \dots, A_n) < n$. Therefore, since there are n vectors, $(A_1, \dots, A_n)$ is not free. So there exist scalars λ_i not all zero such that $\sum_i \lambda_i A_i = 0$. Choose j such that $\lambda_j \neq 0$. Then,

$$A_j = -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i A_i$$

So we can compute the determinant of A by

$$\det A = \det\left(A_1, \dots, -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i A_i, \dots, A_n \right)$$

Since the determinant is alternating and linear in the jth entry, its value is zero. So A is singular as required. □

Remark 29. The above theorem gives necessary and sufficient conditions for invertibility of a set of n linear equations with n unknowns.

§5.7 Determinants of linear maps

Lemma 5.9

Similar matrices have the same determinant.

Proof.

$$\det\left(P^{-1}AP\right) = \det\left(P^{-1}\right) \det A \det P = \det A \det\left(P^{-1}P\right) = \det A$$

□

Definition 5.6

If α is an endomorphism, then we define

$$\det \alpha = \det[\alpha]_{B,B}$$

where B is any basis of the vector space. This is well-defined, since this value does not depend on the choice of basis.

Theorem 5.3

$\det\colon L(V,V) \to F$ satisfies the following properties.

1. $\det I = 1$;
2. $\det(\alpha\beta) = \det\alpha\det\beta$;
3. $\det\alpha \neq 0$ if and only if α is invertible, and in this case, $\det(\alpha^{-1})\det\alpha = 1$.

This is simply a reformulation of the previous theorem for matrices. The proof is simple, and relies on the invariance of the determinant under a change of basis.

§5.8 Determinant of block-triangular matrices

Lemma 5.10

Let $A \in M_k(F)$, $B \in M_\ell(F)$, $C \in M_{k,\ell}(F)$. Consider the matrix

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

Then $\det M = \det A \det B$.

Proof. Let $n = k + \ell$, so $M \in M_n(F)$. Let $M = (m_{ij})$. We must compute

$$\det M = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n m_{\sigma(i)i}$$

Observe that $m_{\sigma(i)i} = 0$ if $i \leq k$ and $\sigma(i) > k$. Then, we need only sum over $\sigma \in S_n$ such that for all $j \leq k$, we have $\sigma(j) \leq k$. Thus, for all $j \in \{k+1, \dots, n\}$, we have $\sigma(j) \in \{k+1, \dots, n\}$. We can then uniquely decompose σ into two permutations $\sigma = \sigma_1\sigma_2$, where σ_1 is restricted to $\{1, \dots, k\}$ and σ_2 is restricted to $\{k+1, \dots, n\}$. Hence,

$$\begin{aligned} \det M &= \sum_{\sigma_1 \in S_k} \sum_{\sigma_2 \in S_{n-k}} \varepsilon(\sigma) \prod_{i=1}^n m_{\sigma(i)i} \\ &= \sum_{\sigma_1 \in S_k} \sum_{\sigma_2 \in S_{n-k}} \varepsilon(\sigma_1)\varepsilon(\sigma_2) \prod_{i=1}^k m_{\sigma(i)i} \prod_{i=k+1}^n m_{\sigma(i)i} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\sigma_1 \in S_k} \varepsilon(\sigma_1) \prod_{i=1}^{k} m_{\sigma(i)i} \sum_{\sigma_2 \in S_{n-k}} \varepsilon(\sigma_2) \prod_{i=k+1}^{n} m_{\sigma(i)i} \\
&= \det A \det B
\end{aligned}
$$

□

Corollary 5.3

We need not restrict ourselves to just two blocks, since we can apply the above lemma inductively. In particular, this implies that an upper-triangular matrix with diagonal elements λ_i has determinant $\prod_i \lambda_i$.

§6 Adjugate Matrices

§6.1 Column and row expansions

Let $A \in M_n(F)$ with column vectors $A^{(i)}$. We know that

$$\det\left(A^{(1)}, \dots, A^{(j)}, \dots, A^{(k)}, \dots, A^{(n)}\right) = -\det\left(A^{(1)}, \dots, A^{(k)}, \dots, A^{(j)}, \dots, A^{(n)}\right)$$

Using the fact that $\det A = \det A^\intercal$ we can similarly see that swapping two rows will invert the sign of the determinant.

Remark 30. We could have proven all of the properties of the determinant above by using the decomposition of A into elementary matrices.

Definition 6.1 (Minor)

Let $A \in M_n(F)$. Let $i, j \in \{1, \dots, n\}$. We define the **minor** $A_{\widehat{ij}} \in M_{n-1}(F)$ to be the matrix obtained by removing the ith row and the jth column.

Lemma 6.1

Let $A \in M_n(F)$.

1. Let $j \in \{1, \dots, n\}$. The determinant of A is given by the *column expansion with respect to the jth column*:

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{\widehat{ij}}$$

2. Let $i \in \{1, \dots, n\}$. The same determinant is also given by the *row expansion with respect to the ith row*:

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{\widehat{ij}}$$

This is a process of reducing the computation of $n \times n$ determinants to $(n-1) \times (n-1)$ determinants.

Proof. We will prove case (i), the column expansion with respect to the jth column. Then (ii) will follow from the transpose of the matrix. Let $j \in \{1, \dots, n\}$. We can write $A^{(j)} = \sum_{i=1}^n a_{ij} e_i$ where the e_i are the canonical basis. Then, by swapping rows and columns,

$$\begin{aligned}
\det A &= \det\left(A^{(1)}, \dots, \sum_{i=1}^n a_{ij} e_i, \dots, A^{(n)}\right) \\
&= \sum_{i=1}^n a_{ij} \det\left(A^{(1)}, \dots, e_i, \dots, A^{(n)}\right) \\
&= \sum_{i=1}^n a_{ij} (-1)^{j-1} \det\left(e_i, A^{(1)}, \dots, A^{(n)}\right) \\
&= \sum_{i=1}^n a_{ij} (-1)^{j-1} (-1)^{i-1} \det\left(e_1, \overline{A}^{(1)}, \dots, \overline{A}^{(n)}\right)
\end{aligned}$$

This has brought the matrix into block form, where there is an element of value 1 in the top left, and the matrix $A_{\widehat{ij}}$ in the bottom right. The bottom left block is entirely zeroes. Hence,

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{\widehat{ij}}$$

as required. □

Remark 31. We have proven that

$$\det\left(A^{(1)}, \dots, e_i, \dots, A^{(n)}\right) = (-1)^{i+j} \det A_{\widehat{ij}}$$

§6.2 Adjugates

Definition 6.2

Let $A \in M_n(F)$. The *adjugate matrix* of A, denoted $\operatorname{adj} A$, is the $n \times n$ matrix given by

$$(\operatorname{adj} A)_{ij} = (-1)^{i+j} \det A_{\widehat{ji}}$$

Hence,

$$\det\left(A^{(1)}, \dots, e_i, \dots, A^{(n)}\right) = (\operatorname{adj} A)_{ji}$$

Theorem 6.1

Let $A \in M_n(F)$. Then

$$(\operatorname{adj} A)A = (\det A)I$$

In particular, when A is invertible,

$$A^{-1} = \frac{\operatorname{adj} A}{\det A}$$

Proof. We have

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{\widehat{ij}}$$

Hence,

$$\det A = \sum_{i=1}^{n} (\operatorname{adj} A)_{ji} a_{ij} = ((\operatorname{adj} A)A)_{jj}$$

So the diagonal terms match. Off the diagonal,

$$0 = \det\left(A^{(1)}, \dots, \underbrace{A^{(k)}}_{j\text{th position}}, \dots, A^{(k)}, \dots, A^{(n)} \right)$$

By linearity,

$$\begin{aligned}
0 &= \det\left(A^{(1)}, \dots, \underbrace{\sum_{i=1}^{n} a_{ik} e_i}_{j\text{th position}}, \dots, A^{(k)}, \dots, A^{(n)} \right) \\
&= \sum_{i=1}^{n} a_{ik} \det\left(A^{(1)}, \dots, \underbrace{e_i}_{j\text{th position}}, \dots, A^{(k)}, \dots, A^{(n)} \right) \\
&= \sum_{i=1}^{n} a_{ik} (\operatorname{adj} A)_{ji} \\
&= ((\operatorname{adj} A)A)_{jk}
\end{aligned}$$

□

§6.3 Cramer's rule

Proposition 6.1

Let A be an invertible square matrix of dimension n. Let $b \in F^n$. Then the unique solution to $Ax = b$ is given by

$$x_i = \frac{1}{\det A} \det\left(A_{\widehat{ib}}\right)$$

where $A_{\widehat{ib}}$ is obtained by replacing the ith column of A by b. This is an algorithm to compute x, avoiding the computation of A^{-1}.

Proof. Let A be invertible. Then there exists a unique $x \in F^n$ such that $Ax = b$. Then, since the determinant is alternating,

$$\begin{aligned}
\det\left(A_{\widehat{ib}}\right) &= \det\left(A^{(1)}, \dots, A^{(i-1)}, b, A^{(i+1)}, \dots, A^{(n)}\right) \\
&= \det\left(A^{(1)}, \dots, A^{(i-1)}, \sum_{j=1}^{n} x_j A^{(j)}, A^{(i+1)}, \dots, A^{(n)}\right) \\
&= \det\left(A^{(1)}, \dots, A^{(i-1)}, x_i A^{(i)}, A^{(i+1)}, \dots, A^{(n)}\right) \\
&= x_i \det A
\end{aligned}$$

So the formula works. $\square$