Linear Algebra

October 31, 2016

Contents

1 Vector spaces

1.1 Vector spaces

Notation. We will use $\mathbb{F}$ to denote an arbitrary field.

Definition. An $\mathbb{F}$-vector space is an abelian group $(V,+)$ equipped with a function

$$\mathbb{F} \times V \to V$$
$$(\lambda, v) \to \lambda v$$

which is called scalar multiplication such that

$$\begin{aligned} &\bullet \lambda(\mu v) = (\lambda\mu) v && \forall \lambda, \mu \in \mathbb{F}, v \in V \\ &\bullet (\lambda + \mu)(v) = \lambda v + \mu v && \\ &\bullet \lambda(v+u) = \lambda v + \lambda u && \forall \lambda \in \mathbb{F}, u, v \in V \\ &\bullet 1 \cdot v = v && \forall v \in V \end{aligned}$$

For convention, we will always write 0 for the identity in a vector space and by 'abuse' of notation, write 0 for the vector space $\{0\}$.

Definition. Suppose VS is a vector space on $\mathbb{F}$. If $U \subset V$, then U is a *($\mathbb{F}$-linear) subspace* if
• $\forall u_1, u_2 \in U,\ u_1 + u_2 \in U$;
• $\forall \lambda \in F, u \in U, \lambda u \in U$;
• $U \neq \phi$.

Remark. A subspace of a vector space is itself a vector space. $U \subset V$ is a subspace iff $0 \in U$, and $\lambda u_1 + \mu u_2 \in U\ \forall \lambda, \mu \in \mathbb{F}, u_1, u_2 \in U$.

Example. Let $V = \mathbb{R}^3$, $U = \left\{(x_1, x_2, x_3)^T | x_1 + x_2 + x_3 = t\right\}$, then U is a subspace if and only if $t = 0$.

Example. If X is a set and $f : X \to \mathbb{R}$, the *support* of f:

$$\operatorname{supp}(f) = \{x \in X | f(x) \neq 0\}$$

Then

$$\left\{f \in \mathbb{R}^X || \operatorname{supp}(f)| < \infty\right\} \subset \mathbb{R}^X$$

is a subspace, since

$$\begin{aligned} &\operatorname{supp}(f+g) \subseteq \operatorname{supp}(f) \cup \operatorname{supp}(g) \\ &\operatorname{supp}(\lambda f) = \operatorname{supp}(f) \text{ if } \lambda \neq 0 \\ &\operatorname{supp}(0) = \phi \end{aligned}$$

Proposition. If U and W are subspaces of a vector space V, then the *sum* of U and W,i.e. $U + W = \{u + w | u \in U, w \in W\}$ and the intersection $U \cap W$ are subspaces of V.
If X is a subspace of V containing U and W, then X contains $U + W$,i.e. $U + W$ is the smallest subspace containing both U and W.
If Y is a subspace of V contained in U and W, then Y is contained in $U \cap W$, i.e. $U \cap W$ is the largest subspace contained in both U and W.

Proof. Certainly $U + W$ and $U \cap W$ both contain 0.
Now suppose $v_1, v_2 \in U \cap W, u_1, u_2 \in U, w_1, w_2 \in W, \lambda, \mu \in \mathbb{F}$. Then

$$\lambda v_1 + \mu v_2 \in U \cap W$$
$$\lambda (u_1 + w_1) + \mu (u_2 + w_2) = (\lambda u_1 + \mu u_2) + (\lambda w_1 + \mu w_2) \in U + W$$

So $U \cap W$ and $U + W$ are subspaces.
Suppose X is as in statement, then if $u \in U, w \in W$,then $u, w \in X$,so $u + w \in X$, so $U + W \subset X$. □

Definition. Suppose V is a vector space and U is a subspace. Then the *quotient space* V/U is the abelian group V/U equipped with

$$\mathbb{F} \times V/U \to V/U$$
$$(\lambda, v + U) \to \lambda v + u$$

Proposition. V/U with this structure is a vector space.

Proof. To see scalar multiplication is well-defined:
Suppose $v_1 + U = v_2 + U \in V/U$. Then $(v_1 - v_2) \in U$. So $\lambda (v_1 - v_2) \in U$ for all $\lambda \in \mathbb{F}$. Thus $\lambda v_1 + U = \lambda v_2 + U$.
Now the four axioms follow easily. □

1.2 Linear independence, bases, and the Steinitz exchange lemma

Definition. Suppose V is a vector field and $S \subset V$. Then the *span* of S in V is

$$\langle S \rangle = \left\{ \sum_{i=1}^{n} \lambda_i s_i | \lambda_i \in \mathbb{F}, s_i \in S \right\}$$

Remark. Several points:
• $\langle S \rangle$ consists only of *finite* linear combination of elements of S.
• For any subset S of V, $\langle S \rangle$ is the smallest subspace of V that contains S.

Example. Suppose $V = \mathbb{R}^3$, $S = \left\{ (1,0,0)^T, (0,1,1)^T, (1,2,2)^T \right\}$. Then

$$\langle S \rangle = \left\{ (a,b,b)^T | a, b \in \mathbb{R} \right\} = \left\{ (x_1, x_2, x_3)^T \in \mathbb{R}^3 | x_2 - x_3 = 0 \right\}$$

Note every subset of S of size 2 has the same span as S.

Example. Let X be a set, and

$$\begin{aligned} \delta_x : X &\to \mathbb{F} \\ x &\to 1 \\ y &\to 0 \quad (y \neq x) \end{aligned}$$

Then $\langle \delta_x \rangle = \left\{ f \in \mathbb{F}^X | \, |\mathrm{supp}\,(f)| < \infty \right\}$.

Definition. Let V be a vector space on $\mathbb{F}$ and $S \subset V$.
(1) We say S spans V if $\langle S\rangle = V$.
(2) We say S is *linearly independent(LI)* if, whenever $\sum_{i=1}^n \lambda_i s_i = 0$ with $s_i \in S$ distinct and $\lambda_i \in \mathbb{F}$, we must have $\lambda_i = 0$ for all i.
If S is not LI, we say is *linearly dependent(LD)*.
(3) S is a *basis* for V if it spans V and is LI.
(4) If V has a finite basis, we say V is *finite-dimensional(f.d.)*.

Note it is not yet clear that every basis of a f.d. vector space has the same size.

Example. Let V and S be the same in the previous example. Then S is LD. Moreover S does not span V. But every subset of S of size 2 is LI and forms a basis for $\langle S\rangle$.

Remark. (1) 0=0 so no LI subset can contain 0. By convention, $\langle \phi\rangle = 0$.

Lemma. A subset S of a vector space V is LD if and only if there exists $s_0, ..., s_n \in S$ distinct such that $s_0 = \sum_{i=1}^n \lambda_i s_i$ for some $\lambda_i \in \mathbb{F}$.

Proof. Suppose S is LD. Then there exists $s_1, ...s_n \in S$, $\lambda_1, ..., \lambda_n \in \mathbb{F}$ with s_i distinct, $\sum \lambda_i s_i = 0$, and suppose $\lambda_j \neq 0$. then

$$s_j = \sum_{i\neq j} -\frac{\lambda_i}{\lambda_j} s_i$$

The converse is trivial. □

Proposition. If S is a basis for V, then every element of V can be written uniquely as $\sum_{s\in S} \lambda_s s$ with $\lambda_s \in \mathbb{F}$ such that all but finitely many λ_s are 0.

Proof. By definition, S spans V if and only if every element of V can be written as at least one way like $\sum_{s\in S} \lambda_s s$ with all but finitely many $\lambda_s = 0$. So we want to show that S is LI if and only if every element $v \in V$ can be written in at most one way of such.
Suppose $v = \sum_{s\in S} \lambda_s s = \sum_{s\in S} \mu_s s$. Then

$$0 = \sum_{s\in S} (\lambda_s - \mu_s)\, s = \sum_{s\in S} 0s$$

So if some $\lambda_s \neq \mu_s$, then S is LD. Conversely, if S is LD, we can write

$$0 = \sum 0s = \sum \lambda_s s$$

for some $\lambda_s \in \mathbb{F}$. □

Theorem. (Steinitz Exchange Lemma) Let V be a $\mathbb{F}$-vector space, and $S = \{e_1, ..., e_n\} \subset V$ is LI, and $T \subset V$ spans V. Then there exists $T' \subset T$ of size n such that $(T\backslash T') \cup S$ spans V. In particular, $|T| \geq n$.

Proof. Idea: replace elements of T one at a time.
Suppose we have found a set D_r of order r for some $0 \leq r < n$ s.t. $(T \backslash D_r) \cup \{e_1, ..., e_r\}$ spans V. $r = 0$ is trivial and $r = n$ is the statement of the theorem. We know

$$e_{r+1} \in \langle (T \backslash D_r) \cup \{e_1, ..., e_r\} \rangle$$

Call RHS (inside $\langle\rangle$) T_r. So we can find $t_1, ..., t_k \in T_r$ and $\lambda_1, ..., \lambda_k \in \mathbb{F}$ s.t.

$$e_{r+1} = \sum_{i=1}^{k} \lambda_i t_i$$

Since $\{e_1, ..., e_{r+1}\}$ is LI, $\exists\ j$ s.t. $\lambda_j \neq 0$ and $t_j \notin \{e_1, ..., e_r\}$. Now

$$t_j \in \langle (T_r \backslash \{t_j\}) \cup \{e_{r+1}\} \rangle$$

Now let $D_{r+1} = D_r \cup \{t_j\}$ so that

$$(T \backslash D_{r+1}) \cup \{e_1, ..., e_{r+1}\} = (T_r \backslash \{t_j\}) \cup \{e_{r+1}\}$$

Then

$$\langle (T \backslash D_{r+1}) \cup \{e_1, ..., e_{r+1}\} \rangle = \langle T_r \cup \{e_{r+1}\} \rangle$$

(t_j is in LHS); while RHS contains $\langle T_r \rangle = V$. So we're done by induction. □

Corollary. If $\{e_1, .., e_n\} \subset V$ is LI, and $\{f_1, ..., f_m\}$ spans V, then $n \leq m$. Also, after re-ordering the f_i, $\{e_1, ..., e_n, f_{n+1}, ..., f_m\}$ spans V.

Corollary. Suppose V is a finite dimension vector space with basis $S = \{e_1, ..., e_n\}$. Then
(a) Every basis of V has order n;
(b) Every spanning set of V of size n is a basis;
(c) Every LI subset of V of size n is a basis;
(d) Every LI subset of V is contained in a basis;
(e) Every finite spanning set in V has a subset that is a basis.

Proof. (a) Suppose S is a finite basis and T is another basis, then any finite subset T' of T is LI. So $|T'| \leq |S|$ by the theorem. The other way is similar.
(b) Suppose T spans VS and has size n. If T is LD then $\exists t_0, ..., t_k \in T$ distinct and $\lambda_1, ..., \lambda_n \in \mathbb{F}$ s.t. $t_0 = \sum_{i=1}^{k} \lambda_i t_i$. Then $\langle T \backslash \{t_0\} \rangle = \langle T \rangle = V$. So $T \backslash \{t_0\}$ spans V and has order $n - 1$. Contradiction with the theorem.
(c) Suppose T is LI and has order n. If $\langle T \rangle \neq V$, then $\exists v \in V \backslash \langle T \rangle$ and $T \cup \{v\}$ is LI (and has order $n + 1$). Contradiction with the theorem.
(d) Suppose $T \subset V$ is LI. Since S spans V, $\exists D \subset S$ with order $|T|$ s.t. $(S \backslash D) \cup T$ spans V and has size $\leq n$ (in fact, $= n$ by the theorem). By (b) $(S \backslash D) \cup T$ is a basis containing T.
(e) Suppose T is a finite spanning set. Let $T' \subset T$ be of minimal size such that $\langle T' \rangle = V$. If $|T'| < n$ we get a contradiction. If $|T'| > n$ then T' is LD by the theorem, so we can reduce the number of vectors in T' by similar methods as above, contradiction. So $|T'| = n$, and by (b) we are done. □

Definition. If V is a finite dimensional vector space on $\mathbb{F}$, the *dimension* of V is

$$\dim_{\mathbb{F}} V = \dim V = |S|$$

where S is any basis f of V.

Remark. • By the above corollary (a), this definition does not depend on S (so is well defined), but does depend on $\mathbb{F}$. For example, $\dim_{\mathbb{C}} \mathbb{C} = 1$ since $\{1\}$ is a basis, but $\dim_{\mathbb{R}} \mathbb{C} = 2$ since $\{1, i\}$ is a basis.
• we could also define the dimension of a non-finite dimensional vector space as the size of a basis, but we haven't proved that such a vector space has a basis, nor that all are ? bijection.

Proposition. If V is a finite dimensional $\mathbb{F}$-vector space and $U \leq V$. Then

$$\dim V = \dim U + \dim V/U$$

Proof. First prove a lemma:

Lemma. If V is a finite dimensional vector space and $U \leq V$, then U is finite dimensional. Indeed $\dim U \leq \dim V$.

Proof. Every LI subset of U is finite. So we choose one *as large as possible.* $|S| \leq \dim V$ by Steinitz.
If $\langle S \rangle \not\leq U$ then $\exists u \leq U \setminus \langle S \rangle$, and $S \cup \{u\}$ is still LI. Contradiction. So S is a basis for U. So $\dim U = |S| \leq \dim V$. □

We must prove that by choosing bases.
Let $\{u_1, ..., u_n\}$ be a basis for U, and extend (by some previous corollary) it to a basis $\{u_1, ..., u_m, v_{m+1}, ..., v_n\}$ for V.
We claim that $\{v_{m+1} + U, ..., v_n + U\} = S$ is a basis for V/U. This claim gives the result by counting.
First we prove that S spans V:
If $v + U \in V/U$, $\exists \lambda_1, ..., \lambda_n \in \mathbb{F}$ s.t.

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

Now

$$v + U = \sum_{i=1}^{m} \lambda_i (u_i + U) + \sum_{i=m+1}^{n} \lambda_i (v_i + U)$$

while the first term is zero.
Then we prove that S is LI:
If

$$\sum_{i=m+1}^{n} \lambda_i (v_i + U) = 0 + U$$

then

$$\sum_{i=m+1}^{n} \lambda_i v_i \in U$$

So $\exists \lambda_1, ..., \lambda_m \in \mathbb{F}$ s.t.

$$\sum_{i=m+1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \lambda_i u_i$$

But since $\{u_1, ..., u_m, v_{m+1}, ..., v_n\}$ is LI, we must have $\lambda_1 = ... = \lambda_m = 0$. □

Corollary. If $U \leq V$ is a proper subspace, then $\dim U < \dim V$.

Proof. We know that $\dim V = \dim U + \dim V/U$. Since U is a proper subspace of V, $V/U \neq 0$. So ϕ does not span V/U. So $\dim V/U \neq 0$. So $\dim U < \dim V$. □

1.3 Direct Sums

Definition. Suppose V is a $\mathbb{F}$-vector space, and $U, W \leq V$. Recall $U + W = \{u + w | u \in U, w \in W\}$. We say V is the *(internal) direct sum* of U and W if $V = U + W$ and $U \cap W = 0$; equivalently, if every element $v \in V$ can be written uniquely as $u + w$ with $u \in U, w \in W$. We write $V = U \oplus W$.
We say U and W are *complementary subspaces* of V.

Example. Let $V = \mathbb{R}^2$ and $U = \left\langle (0,1)^T \right\rangle$, and $W_1 = \left\langle (1,0)^T \right\rangle$ and $W_2 = \left\langle (1,1)^T \right\rangle$. Then W_1 and W_2 are both complementary to U in V. So complementary subspaces need *not* be unique.

Definition. If U, W are $\mathbb{F}$-vector space. The *(external) direct sum of* U and W,

$$U \oplus W = \{(u, w) \, | u \in U, w \in W\}$$

with addition

$$(u_1, w_1) + (u_2, w_2) = (u_1 + u_2, w_1 + w_2)$$

and scalar multiplication

$$\lambda_1 (u, w) = (\lambda u, \lambda w)$$

For $\lambda \in \mathbb{F}, u_1, u_2, u \in U, w_1, w_2, w \in W$.

This defines a vector space.

Problem. Show $U \oplus W$ is a vector space and is the internal direct sum of $\{(u, 0) \, | u \in U\}$ and $\{(0, w) \, | w \in W\}$.

Definition. If $U_1, ..., U_n \leq V$ are subspaces of an $\mathbb{F}$-fector space V. Then V is the (internal) direct sum of $U_1, ..., U_n$ written

$$V = U_1 \oplus ... \oplus U_n = \oplus_{i=1}^n U_i$$

if every $v \in V$ can be written uniquely as

$$v = \sum_{i=1}^n u_i$$

with $u_i \in U_i$.

See Example sheet 1 Q9.

Definition. If $U_1, ..., U_n$ be $\mathbb{F}$-vector spaces, then the external direct sum

$$\oplus_{i=1}^n U_i = \{(u_1, ..., u_n) \, u_i \in U_i\}$$

and coordinate-wise operations.

2 Linear maps

2.1 Definitions and examples

Definition. Suppose U and W are $\mathbb{F}$-vector spaces. Then $\alpha : U \to W$ is a *linear map* if:
(i) $\alpha(u_1 + u_2) = \alpha(u_1) + \alpha(u_2)$ for all $u_1, u_2 \in U$;
(ii) $\alpha(\lambda u) = \lambda\alpha(u)$ for all $\lambda \in \mathbb{F}$, $u \in U$;

Notation. We'll write $\mathcal{L}(U, V) = \{\alpha : U \to V | \alpha \text{ is linear}\}$.

Remark. (1) If α is linear, then $\alpha(0) = 0$.
(2) α is linear $\Leftrightarrow \alpha(\lambda u_1 + \mu u_2) = \lambda\alpha(u_1) + \mu\alpha(u_2)$ for all $\lambda, \mu \in \mathbb{F}$, $u_1, u_2 \in U$.
(3) If we want to stress the $\mathbb{F}$ we say $\mathbb{F}$-linear. For example, complex conjugation $\mathbb{C} \to \mathbb{C}$ is $\mathbb{R}$-linear but not $\mathbb{C}$-linear.

Example. Let A be an $m \times n$ matrix with coefficients in $\mathbb{F}$ (write $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$). Then $\alpha : \mathbb{F}^n \to \mathbb{F}^m$, $\alpha(v) = Av$ defines a linear map.
To see this, let $\lambda, \mu \in \mathbb{F}$, $u, v \in \mathbb{F}^n$, and A_{ij} the i, j^{th} entry of A, and u_j (resp v_j) for the j^{th} coordinate of u (resp v) etc,. Then for $1 \leq i \leq m$,

$$\begin{aligned} (\alpha(\lambda u + \mu v))_i &= \sum_{j=1}^{n} A_{ij}(\lambda u_j + \mu v_j) \\ &= \lambda \sum_{j=1}^{n} A_{ij} u_j + \mu \sum_{j=1}^{n} A_{ij} v_j \\ &= \lambda\alpha(u)_i + \mu\alpha(v)_i \end{aligned}$$

and α is linear as required.

Example. If X is any set, $g \in F^X$, then $Mg : \mathbb{F}^X \to \mathbb{F}^X$ by $Mg(f)(x) = g(x) f(x)$ for all $x \in X$ is linear.

Example. For all $x \in [a, b]$, $C([a, b], \mathbb{R}) \to \mathbb{R}, f \to f(x)$ is linear.

Example. $D : C^\infty([a, b], \mathbb{R}) \to C^\infty([a, b], \mathbb{R})$, $f \to \frac{df}{dx}$ is linear.

Example. $I : C([a, b], \mathbb{R}) \to \mathbb{R}$, $f \to \int_a^b f dx$ is linear.

Example. If $\alpha, \beta : U \to V$ are linear, then $\alpha + \beta$ is linear (example sheet 1 Q4), and $\lambda\alpha$ is linear for all $\lambda \in \mathbb{F}$. In this way, $\mathcal{L}(U, V)$ is a $\mathbb{F}$-vector space.
Also, If $\gamma : V \to W$ is linear, then $\gamma\beta : U \to W$ is linear.

Definition. We say a linear map $\alpha : U \to W$ is an *isomorphism* if $\exists \beta \in \mathcal{L}(W, U)$ s.t. $\alpha\beta = \iota_W$ and $\beta\alpha = \iota_V$, where ι is the identity map in the respective space. U and W are *isomorphic* if there is such an isomorphism between them.

Lemma. If $\alpha \in \mathcal{L}(U, V)$, then α is an isomorphism if and only if α is a bijection.

Proof. $\Longrightarrow$ is clear. If α is an isomorphism then it has an inverse as a function, so is a bijection.

$\Longleftarrow$ Suppose α is a bijection, and $\beta : V \to U$ is its inverse. We must show β is linear.

Suppose $\lambda, \mu \in \mathbb{F}$, $v_1, v_2 \in V$. Then

$$\begin{aligned}\alpha\beta\left(\lambda v_1 + \mu v_2\right) &= \lambda\alpha\beta\left(v_1\right) + \mu\alpha\beta\left(v_2\right) \\ &= \alpha\left(\lambda\beta\left(v_1\right) + \mu\beta\left(v_2\right)\right)\end{aligned}$$

So

$$\beta\left(\lambda v_1 + \mu v_2\right) = \lambda\beta\left(v_1\right) + \mu\beta\left(v_2\right)$$

i.e. β is linear as required. $\square$

Proposition. Suppose $\alpha \in \mathcal{L}\left(U, V\right)$.
(a) If α is injective and $S \subset U$ is LI, then $\alpha\left(S\right)$ is LI.
(b) If α is surjective and $S \subset U$ s.t. S spans U then $\alpha\left(S\right)$ spans V.
(c) If α is an isomorphism and $S \subset U$ is a basis, then $\alpha\left(S\right)$ is a basis.

Proof. (c) follows immediately form (a) and (b).

(a) Suppose $\alpha\left(S\right)$ is LD, so $\exists s_0, ..., s_n \in S$ distinct and $\lambda_1, ..., \lambda_n \in \mathbb{F}$ s.t.

$$\begin{aligned}\alpha\left(s_0\right) &= \sum_{i=1}^{n} \lambda_i \alpha\left(s_i\right) \\ &= \alpha\left(\sum_{i=1}^{n} \lambda_i s_i\right)\end{aligned}$$

So

$$S_0 = \sum_{i=1}^{n} \lambda_i s_i$$

Since α is injective. Contradiction.

(b) Since α is surjective, $\forall v \in V\ \exists u \in U$ s.t. $\alpha\left(u\right) = v$.
Since S spans U, $\exists s_1, ..., s_n \in S$ and $\lambda_i \in \mathbb{F}$ s.t. $u = \sum \lambda_i s_i$.

Then $v = \alpha\left(u\right) = \sum \lambda_i \alpha\left(s_i\right)$, and $v \in \langle \alpha\left(S\right) \rangle$ as required. $\square$

Corollary. Any two finite dimensional vector spaces that are isomorphic must have the same dimension.

Proof. Suppose $U \cong v$ and both of them are finite dimensional. Let S be a basis for U and $\alpha \in \mathcal{L}\left(U, V\right)$ is an isomorphism then $\alpha\left(S\right)$ is a basis for V.
Since α is injective, $|S| = |\alpha\left(S\right)|$ and $\dim U = \dim V$. $\square$

Proposition. Suppose V is an $\mathbb{F}$-vector space of dimension n, then there is a bijection from $\Phi : \{$ isomorphisms $\alpha : \mathbb{F} \to V\}$ to the set of (ordered) bases for V. by $\alpha \to \left(\alpha\left(e_1\right), ..., \alpha\left(e_n\right)\right)$, where $e_1, ..., e_n$ is the standard basis for $\mathbb{F}^n$.

Proof. $\square$

That Φ is a function follows from above proposition (c).
If $\Phi(\alpha) = \Phi(\beta)$ then

$$\begin{aligned} \alpha (x_1, ..., x_n)^T &= \sum_{i=1}^{n} x_i \alpha (e_i) \\ &= \sum_{i=1}^{n} x_i \beta (e_i) \\ &= \beta (x_1, ..., x_n)^T \end{aligned}$$

So $\alpha = \beta$.
Suppose $(v_1, ..., v_n)$ is a basis for V. Define $\alpha : \mathbb{F}^n \to V$ by $(x_1, ..., x_n)^T \to \sum_{i=1}^{n} x_i v_i$. Then α is linear, injective and surjective. So α is an isomorphism, and $\Phi(\alpha) = (\alpha(e_1), ..., \alpha(e_n)) = (v_1, ..., v_n)$.

2.2 Linear maps and matrices

Proposition. Suppose U and V are vector spaces on $\mathbb{F}$ and $S = \{e_1, ..., e_n\}$ is a basis for U. Then every function $f : S \to V$ extends uniquely to a linear map $\alpha : U \to V$.

Proof. • Uniqueness: suppose $\alpha, \beta \in \mathcal{L}(U, V)$ that agree with f on S. Let $u \in U$. Then there exists $\lambda_1, ..., \lambda_n \in \mathbb{F}$ s.t. $u = \sum \lambda_i e_i$.
Then $\alpha(u) = \sum \lambda_i f(e_i) = \beta(u)$. So $\alpha = \beta$.

• Existence: Every $u \in U$ can be written as $u = \sum \lambda_i e_i$ with $\lambda_i \in \mathbb{F}$ without ambiguity. So we can define $\alpha : U \to V$ by $u \to \sum \lambda_i f(e_i)$.
Suppose $\lambda, \mu \in \mathbb{F}$ and $u_1, u_2 \in U$ s.t. $u_1 = \sum \lambda_i e_i$, $u_2 = \sum \mu_u e_i$. Then

$$\begin{aligned} \alpha(\lambda u_1 + \mu u_2) &= \alpha\left(\sum (\lambda\lambda_i + \mu\mu_i) e_i\right) \\ &= \sum (\lambda\lambda_i + \mu\mu_i) f(e_i) \\ &= \lambda \sum \lambda_i f(e_i) + \mu \sum \mu_i f(e_i) \\ &= \lambda\alpha(u_i) + \mu\alpha(u_2) \end{aligned}$$

So $\alpha \in \mathcal{L}(U, V)$.

Finally $\alpha(e_i) = f(e_i)$ as required. □

Remark. • With a little care, we can remove condition U is finite dimensional (and so S finite).
• It isn't too hard to see that $S \subset U$ satisfies the conclusion of the proposition only if S is a basis. This is a key motivation for the definition of basis.

Corollary. If U and V are finite dimensional vector spaces on $\mathbb{F}$ with (ordered) bases $(e_1, ..., e_n)$ and $(f_1, ..., f_n)$ respectively, then there is a bijection

$$\begin{aligned} \mathcal{L}(U, V) &\to \text{ Mat }_{m,n}(\mathbb{F}) \\ \alpha &\to A \end{aligned}$$

s.t.

$$\alpha\left(e_i\right)=\sum_{j=1}^{n} A_{ji} f_j$$

We interpret this as: the i^{th} column of A tells where i^{th} bases vector of U gets sent by α (as a linear combination of basis vectors in V).

Proof. If $\alpha \in \mathcal{L}(U, V)$, then we can write

$$\alpha\left(e_i\right)=\sum_{j=1}^{n} A_{ji} f_j$$

for $1 \leq i \leq n$ for unique $A_{ji} \in \mathbb{F}$. So function is injective and surjective as required. $\square$

Proposition. Show that this bijection is an isomorphism, and deduce that if U and V are finite dimensional, then $\dim \mathcal{L}(U, V)=(\dim U)(\dim V)$.
Show also that if $U_1, \ldots, U_n$ are vector spaces, then

$$\mathcal{L}\left(\oplus_{i=1}^{n} U_i, V\right) \cong \oplus_{i=1}^{n} \mathcal{L}\left(U_i, V\right)$$

and

$$\mathcal{L}\left(V, \oplus_{i=1}^{n} U_i\right) \cong \oplus_{i=1}^{n} \mathcal{L}\left(V, U_i\right)$$

We leave this as an exercise.

Definition. If $\alpha \in \mathcal{L}(U, V)$ and $\left(u_1, \ldots, u_r\right)$ is a basis for U and $\left(v_1, \ldots, v_s\right)$ is a basis for V, and A is the matrix s.t.

$$\alpha\left(u_i\right)=\sum_{j=1}^{s} A_{ji} v_j$$

then we call A the matrix associated to α with respect to the basis $\left(u_1, \ldots, u_r\right)$ and $\left(v_1, \ldots, v_s\right)$.

Lemma. Suppose U, V, W are vector spaces on $\mathbb{F}$ and $R:\left\{u_1, \ldots, u_r\right\}$ is a basis for U, $S=\left\{v_1, \ldots, v_s\right\}$ is a basis for V, and $T=\left\{w_1, \ldots, w_t\right\}$ is a basis for W. Let $\alpha \in \mathcal{L}(U, V)$ and $\beta \in \mathcal{L}(V, W)$. Then if α is represented by A with respect to R and S and β is represented by B with respect to S and T, then $\beta\alpha$ is represented by $=BA$ with respect to R and T.

Proof.

$$\begin{aligned}
\beta \alpha\left(u_i\right) &= \beta\left(\sum_{j=1}^{s} A_{ji} v_j\right) \\
&= \sum_{j=1}^{s} A_{ji} \sum_{k=1}^{t} B_{kj} w_k \\
&= \sum_{k=1}^{t}[BA]_{ki} w_k
\end{aligned}$$

as required. $\square$

2.3 The first isomorphism theorem, and the rank-nullity theorem

Definition. Suppose $\alpha \in \mathcal{L}(U, V)$ (U,V are vector spaces on $\mathbb{F}$). Then the kernel of α,

$$\ker \alpha = \{u \in U | \alpha(u) = 0\}$$

the image of α,

$$\operatorname{im} \alpha = \{\alpha(u) | u \in U\}$$

Note: α is injective if and only if $\ker \alpha = 0$, α is surjective if and only if $\operatorname{im} \alpha = V$.

Example. If $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$ and $\alpha : \mathbb{F}^n \to \mathbb{F}^m$ is the linear map $x \to Ax$. Then the system of equations

$$\sum_{j=1}^{n} A_{ij} x_j = b_i$$

for $1 \leq i \leq m$ has a solution if and only if $(b_1, ..., b_m)^T \in \operatorname{im} \alpha$. Also, $\ker \alpha$ is the set of solutions to the set of homogeneous equations

$$\sum_{j=1}^{n} A_{ij} x_j = 0$$

for $1 \leq i \leq m$.

Example. Let $\beta : C^\infty(\mathbb{R}, \mathbb{R}) \to C^\infty(\mathbb{R}, \mathbb{R})$ given by

$$\beta(f(t)) = f''(t) + p(t) f'(t) + q(t) f(t)$$

for some $p, q \in C^\infty(\mathbb{R}, \mathbb{R})$. Then $g(t) \in \operatorname{im} \beta$ if and only if $\beta(f(t)) = g(t)$ has a solution in $C^\infty(\mathbb{R}, \mathbb{R})$, $\ker \beta$ is the set solutions to the homogeneous equation $\beta(f(t)) = 0$.

Theorem. (First Isomorphism Theorem) Suppose U, V are vector spaces on $\mathbb{F}$, and $\alpha \in \mathcal{L}(U, V)$, then $\ker \alpha$ is a subspace of U, $\operatorname{im} \alpha$ is a subspace of V, and α induces isomorphism $\bar{\alpha}$ by

$$\bar{\alpha} : U / \ker \alpha \to \operatorname{im} \alpha$$
$$\bar{\alpha}(u + \ker \alpha) \to \alpha(u)$$

Proof. $\alpha(0) = 0$, so $0 \in \ker \alpha$.
If $\lambda, \mu \in \mathbb{F}$ and $u_1, u_2 \in \ker \alpha$, then

$$\alpha(\lambda u_1 + \mu u_2) = \lambda \alpha(u_1) + \mu \alpha(u_2) = 0 + 0 = 0$$

Similarly, if $\lambda, \mu \in \mathbb{F}$ and $u_1, u_2 \in u$, then

$$\lambda \alpha(u_1) + \mu \alpha(u_2) = \alpha(\lambda u_1 + \mu u_2) \in \operatorname{im} \alpha$$

(and $0 \in \operatorname{im} \alpha$) so $\operatorname{im} \alpha \leq V$ and $\bar{\alpha}$ is linear if it's well defined.
To show that α is well defined, suppose $u + \ker \alpha = u' + \ker \alpha \in U / \ker \alpha$. Then $u - u' \in \ker \alpha$, so $\alpha(u - u') = 0$. So $\bar{\alpha}(u + \ker \alpha) = \bar{\alpha}(u' + \ker \alpha)$ as required.
$\bar{\alpha}$ surjective is clear.
If $\bar{\alpha}(u + \ker \alpha) = 0$ then $\alpha(u) = 0$. So $u \in \ker \alpha$. Thus $\ker \bar{\alpha} = 0$ as required. $\square$

Definition. If $\alpha \in \mathcal{L}(U,V)$, the *rank* of α is $r(\alpha) = \dim \operatorname{im} \alpha$, and the *nullity* of α is $n(\alpha) = \dim \ker \alpha$.

Corollary. if U and V are finite dimensional vector spaces on $\mathbb{F}$ and $\alpha \in \mathcal{L}(U,V)$, then

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

This is called the *rank-nullity theorem.*

Proof. $U/\ker \alpha \cong \operatorname{im} \alpha$, so $\dim U/\ker \alpha = \dim \operatorname{im} \alpha$. But we've seen $\dim U = \dim \ker \alpha + \dim U/\ker \alpha$, so $\dim U = n(\alpha) + r(\alpha)$.

We'll give other less slick proofs but we've used the fact $\dim V = \dim U + \dim V/U$ which required some work. $\square$

Exercise: deduce the above equation from the rank-nullity theorem.

Alternative proof:

Proposition. If U and V are finite dimensional vector spaces on $\mathbb{F}$ and $\alpha \in \mathcal{L}(U,V)$, then there are bases $(e_1, ..., e_n)$ for U and $(f_1, ..., f_m)$ for V s.t. α is represented by

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

(where I_r is a $r \times r$ identity matrix) with respect to $(e_1, ..., e_n)$ and $(f_1, ..., f_m)$ where $r = r(\alpha)$. In particular, $\dim U = r(\alpha) + n(\alpha)$.

Proof. Let $(e_{k+1}, ..., e_n)$ be bases for $\ker \alpha$, and extend this to a basis $(e_1, ..., e_n)$ for U. Define $f_i = \alpha(e_i)$ for $1 \leq i \leq k$.

Claim. $(f_1, ..., f_k)$ is a basis for $\operatorname{im} \alpha$.
Suppose $\sum_{i=1}^{k} \lambda_i f_i = 0$, then $\sum_{i=1}^{k} \lambda_i \alpha(e_i) = 0$ and $\alpha\left(\sum_{i=1}^{k} \lambda_i e_i\right) = 0$.
So $\sum_{i=1}^{k} \lambda_i e_i \in \ker \alpha$, but $\ker \alpha \cap \langle \{e_1, ..., e_k\} \rangle = 0$.
Thus $\sum_{i=1}^{k} \lambda_i e_i = 0$. Since $\{e_1, ..., e_k\}$ are LI, each $\lambda_i = 0$.
Now $\alpha\left(\sum_{i=1}^{n} \mu_i e_i\right) = \sum_{i=1}^{n} \mu_i \alpha(e_i) = \sum_{i=1}^{k} \mu_i \alpha(e_i) \in \langle \{f_1, ..., f_k\} \rangle$.

Now extend $(f_1, ..., f_k)$ to a basis $(f_1, ..., f_m)$ for V. So

$$\alpha(e_i) = \begin{cases} f_i & 1 \leq i \leq k \\ 0 & k+1 \leq i \leq n \end{cases}$$

So matrix representing α is as required, note $k = \dim \operatorname{im} \alpha = r$, and $n(\alpha) = n - k = n - r$. $\square$

Note it follows from the statement that the only basis independent numerical invariants of α are $\dim U$, $\dim V$, $r(\alpha)$, or deducible from these.

Example. Suppose $W = \left\{x \in \mathbb{R}^5 | x_1 + x_2 + x_3 = 0, x_3 - x_4 - x_5 = 0\right\} \leq \mathbb{R}^5$. We want to find the dimension of W. Consider

$$\begin{aligned} \alpha : \mathbb{R}^5 &\to \mathbb{R}^2 \\ x &\to \begin{pmatrix} x_1 + x_2 + x_5 \\ x_3 - x_4 - x_5 \end{pmatrix} \end{aligned}$$

which is a linear map. Then $\dim W = n(\alpha) = 5 - r(\alpha)$. But α is surjecitve, $\alpha(1,0,0,0,0)^T = (1,0)^T$, $\alpha(0,0,1,0,0)^T = (0,1)^T$, so $r(\alpha) = 2$, and $\dim W = 3$.

More generally, the space of solutions of n linear equations in m unknowns has dimension at least $m - n$.

Example. Suppose U and W are subspaces of V.
Consider $\alpha : U \oplus W \to V : (u, w) \to u + w$, a linear map. Then $\operatorname{im}\alpha = U + W$, $\ker\alpha = \{(u, -u) | u \in U \cap W\} \cong U \cap W$.
So $\dim U + \dim W = \dim U \oplus W = \dim(U + W) + \dim(U \cap W)$ (rank-nullity).

Corollary. (of rank-nullity) If $\alpha \in \mathcal{L}(U, V)$, then the following are equivalent if $\dim U = \dim V = n < \infty$:
(a) α is injective;
(b) α is surjecitve;
(c) α is an isomorphism.

Proof. We've already seen (a)+(b) $\iff$ (c). So we just need to prove (a) $\iff$ (b).
α ijective $\iff n(\alpha) = 0$
$\iff r(\alpha) = \dim U = n = \dim V$ (by rank-nullity)
$\iff \alpha$ is surjective. $\square$

Lemma. Suppose $A \in \mathrm{Mat}_n(\mathbb{F}) := \mathrm{Mat}_{n,n}(\mathbb{F})$. Then the following are equivalent:
(a) $\exists B \in \mathrm{Mat}_n(\mathbb{F})$ s.t. $BA = I_n$;
(b) $\exists C \in \mathrm{Mat}_n(\mathbb{F})$ s.t. $AC = I_n$.
If (a) and (b) hold, then $B = C$, and we write $A^{-1} = B (= C)$ and say A *is invertible.*

Proof. Let α, β, γ and ι be the linear maps represented by A, B, C and I_n respectively with respect to standard bases.
Now (a) holds $\implies \exists \beta : \mathbb{F}^n \to \mathbb{F}^n$ s.t. $\beta\alpha = \iota$
$\implies \alpha$ injective
$\implies \alpha$ is an isomorphism
$\implies \exists \beta : \mathbb{F}^n \to \mathbb{F}^n$ s.t. $\beta\alpha = \iota$
$\implies$ (a) holds $\iff \alpha$ is an isomorphism.

(b) holds $\implies \exists \gamma : \mathbb{F}^n \to \mathbb{F}^n$ s.t. $\alpha\gamma = \iota$
$\implies \alpha$ is surjective
$\implies \alpha$ is an isomorphism
$\implies \exists \gamma : \mathbb{F}^n \to \mathbb{F}^n$ s.t. $\alpha\gamma = \iota$
$\implies$ (b) holds.

So (b) holds $\iff$ α is an isomorphism $\iff$ (a) holds.

Note if α is an isomorphism, then β and γ are the (set-the??) inverse to α. So $\beta = \gamma$ and $B = C$. $\square$

2.4 Change of basis

Theorem. Suppose $(e_1, ..., e_m)$ and $(u-1, ..., u_m)$ be bases for an $\mathbb{F}$-vector space U and $(f_1, ..., f_n)$ and $(v_1, ..., v_n)$ are bases for another $\mathbb{F}$-vector space V. Let $\alpha \in \mathcal{L}(U, V)$, A be the matrix representing α with respect to $(e_1, ..., e_m)$ and $(f_1, ..., f_n)$ and B be the matrix representing α with respect to $(u_1, ..., u_m)$ and $(v_1, ..., v_n)$. Then

$$B = Q^{-1}AP$$

where

$$u_i = \sum_k P_{ki} e_k,$$
$$v_j = \sum_l Q_{lj} f_l$$

$$\begin{array}{cccc} U & \overrightarrow{\iota} U & \overrightarrow{\alpha} V & \overleftarrow{\iota} V \\ \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbb{F}^m & \overrightarrow{P} \mathbb{F}^m & \overrightarrow{A} \mathbb{F}^n & \overleftarrow{Q} \mathbb{F}^n \end{array}$$

Proof. Both AP and QB represent linear map $\alpha : U \to V$ with respect to $(u_1, ..., u_n)$ and $(f_1, ..., f_n)$, so $AP = QB$. But P and Q are invertible, as ι is an isomorphism and $Q^{-1}AP = B$ as required. $\square$

Definition. We say $A, B \in \mathrm{Mat}_{m,n}(\mathbb{F})$ are *equivalent* if $\exists P \in \mathrm{Mat}_n(\mathbb{F})$ and $Q \in \mathrm{Mat}_n(\mathbb{F})$ both invertible s.t. $B = Q^{-1}AP$.

Corollary. (of 2nd proof of rank-nullity theorem) If $A \in \mathrm{Mat}\, m, n\,(\mathbb{F})$, there are invertible $P \in \mathrm{Mat}_n(\mathbb{F})$ and $Q \in \mathrm{Mat}_m(\mathbb{F})$ s.t.

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

Moreover, r is uniquely determined by A, i.e. every equivalence class contains precisely one matrix of the above block form.

Definition. If $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$, then the *column rank* of A is the dimension of span of the column vectors of A as a subspace of $\mathbb{F}^n$. The *row rank* of A is the column rank of A^T.
Note the column rank of A, $r(A)$ is just the rank of the linear map $\mathbb{F}^m \to \mathbb{F}^n$ represented by A with respect to the standard bases. So equivalent matrices have the same column rank.

Corollary. (row-rank = column-rank) If $A \in \mathrm{Mat}\, n, m\,(F)$, then

$$r(A) = r\left(A^T\right)$$

Proof. $\exists P, R$ invertible s.t.

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

Then

$$P^T A^T \left(Q^{-1}\right)^T = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$P^T \left(P^{-1}\right)^T = \left(\left(P^{-1}\right) P\right)^T = I^T = I$$

So $\left(P^T\right)^{-1}$ and Q^T are invertible. So $r\left(A^T\right) = r\left(A\right)$. □

2.5 Elementary matrix and operations

Definition. We call the following types of invertible $n \times n$ matrices *elementary*:

$$S_{ij}^n = \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & ... & & & & \\ & & & 0 \; ... & 1 & & \\ & & & 1 \; ... & 0 & & \\ & & & & & ... & \\ & & & & & & 1 \end{pmatrix}$$

$$E_{ij}^n = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & ... & & & \\ & & & ... & \lambda & \\ & & & & ... & \\ & & & & & 1 \end{pmatrix}$$

for $i \neq j$,

$$T_{ij}^n = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & ... & & & \\ & & & \lambda & & \\ & & & & ... & \\ & & & & & 1 \end{pmatrix}$$

for $\lambda \in \mathbb{F} \setminus \{0\}$.

Observation: If $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$, then AS_{ij}^n ($S_{ij}^n A$) is obtained from A by swapping column (row) i and j.

$AE_{ij}^n(\lambda)$ ($E_{ij}^m(\lambda) A$) is obtained from A by adding λ times of column (row) i to column (row) j.

$AT_i^n(\lambda)$ $(T_i^n(\lambda)A)$ is obtained from A by multiplying column (row) i by λ.

Recall:

Proposition. If $A \in \mathrm{Mat}\, m, n(\mathbb{F})$, then there exists an invertible $P \in \mathrm{Mat}_n(\mathbb{F})$ and $Q \in \mathrm{Mat}_m(\mathbb{F})$ s.t.

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

for some $r \geq 0$. Now we prove this purely by matrices:

Proof. We will prove that there exists elementary matrices $E_1^n, ..., E_k^n$ and $F_1^m, ..., F_l^m$ s.t.

$$F_l^m ... F_1^m A E_1^n ... E_k^n = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

for some $r \geq 0$ which is sufficient.
In fact, equivalently we will prove that there exists elementary row and column operations transforming A into the desired block matrix form.

If $A = 0$ then we are done. Otherwise, $\exists i, j$ s.t. $A_{ij} \neq 0$. By swapping columns 1 and j, and then rows 1 and i, we may assume $A_{11} \neq 0$. By multiplying row 1 by $1/A_{11}$ we may assume $A_{11} = 1$.

By adding $-A_{1j}$ times column 1 to column j for each $j > 1$, and $-A_{i1}$ times row 1 to row i for each $i > 1$, we can assume A is of the form

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

where $B \in \mathrm{Mat}_{m-1,n-1}(\mathbb{F})$. We can complete by induction on $\min(m, n)$. □

The elementary row and column operations preserve the rank of A and A^T. We leave this as an exercise.

3 Duality

3.1 Dual spaces

To specify a subspace of $\mathbb{F}^n$, we can write down a set of suitable linear equations that each vector in the subspace satisfies.

Example. Let $U = \left\langle (1,2,1)^T \right\rangle \subset \mathbb{F}^3$, then

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid 2x_1 - x_2 = 0,\ x_1 - x_3 = 0 \right\}$$

The choice of equations is not caconical (i.e. no best choice).

Each equation can be (conceived?) as defining the kernel of a linear map $g : \mathbb{F}^n \to \mathbb{F}$ e.g. $x_1 - x_3 = 0 \iff x \in \ker \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \to x_1 - x_3 \right)$.

If $g_1, g_2 \in \mathcal{L}(\mathbb{F}^n, \mathbb{F})$ s.t. $g_1(U) = g_2(U) = 0$, then $(\lambda g_1 + \mu g_2)(U) = 0$ for all $\lambda, \mu \in \mathbb{F}$.

More over, $0 \in \mathcal{L}(\mathbb{F}^n, \mathbb{F})$ and $0(U) = 0$. So $\{\theta \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}) \mid \theta(U) = 0\} \leq \mathcal{L}(\mathbb{F}^n, \mathbb{F})$.

Definition. If V is an $\mathbb{F}$-vector space, then the *dual* of V is

$$V^* = \mathcal{L}(V, \mathbb{F}) = \{\alpha : V \to \mathbb{F} \mid \alpha \text{ is linear }\}$$

Elements of V^* are often called *linear-forms* or *linear functions.*

Example.

$$\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \to x_1 - x_3 \right) \in \left(\mathbb{R}^3\right)^*$$

Example. If X is a set, $x \in X$, then

$$(f \to f(x)) \in \left(\mathbb{F}^X\right)^*$$

Example.

$$f \to \int_0^1 \sin(2n\pi x) f(x)\, dx \in (C[0,1], \mathbb{R})^*$$

Example.

$$\begin{aligned} \operatorname{tr} : \operatorname{Mat}_n(\mathbb{F}) &\to \mathbb{F} \\ A &\to \sum_{i=1}^n A_{ii} \end{aligned}$$

is in $\operatorname{Mat}_n(\mathbb{F})^*$.

Lemma. If V is a finite dimensional $\mathbb{F}$-vector space with basis $(e_1, ..., e_n)$, there is a basis $(\varepsilon_1, ..., \varepsilon_n)$ of V^* s.t. $\varepsilon_i(e_j) = \delta_{ij}$ for $1 \leq i, j \leq n$. We say $(\varepsilon_1, ..., \varepsilon_n)$ is the basis *dual* to $(e_1, ..., e_n)$.

Proof. Since a linear map is determined by its value on a basis, $\varepsilon_1, ..., \varepsilon_n$ are uniquely determined by $\varepsilon_i(e_j) = \delta_{ij}$.

Suppose $\theta \in V^*$. Let $\lambda_i = \theta(e_i)$, then $\theta(e_i) = \left(\sum_{v=1}^n \lambda_j \varepsilon_j\right)(e_i)$ for $1 \leq i \leq n$.

Since $(e_1, ..., e_n)$ is a basis, $g = \sum \lambda_j \varepsilon_j$ and $\theta \in \langle \varepsilon_1, ..., \varepsilon_n \rangle$.

Next, suppose $\sum_{i=1}^n \lambda_i \varepsilon_i = 0$. Then

$$\left(\sum_{i=1}^n \lambda_i \varepsilon_i\right)(e_j) = \lambda_j = 0$$

for $1 \leq j \leq n$. So $\varepsilon_1, ..., \varepsilon_n$ is LI as required. $\square$

Remark. If $(a_1, ..., a_n)$ is a row vector and $(x_1, ..., x_n)^T$ is a column vector, then $(a_1, ..., a_n)(x_1, ..., x_n)^T = \sum_{i=1}^n a_i x_i = \left(\sum a_i \varepsilon_i\right)\left(\sum x_j e_j\right)$ where $e_1, ..., e_j$ is the standard basis for $\mathbb{F}^n$ and $\varepsilon_1, ..., \varepsilon_n$ is the dual basis for $(\mathbb{F}^n)^*$. So ε_i is the row vector with 1 in entry i and 0 elsewhere.

Corollary. If V is finite dimensional, then $\dim V = \dim V^*$.

Definition. If $U \subset V$, the *annihilator* of U is $U^\circ = \{\theta \in V^* \mid \theta(u) = 0 \forall u \in U\}$.

Example. If $U = \left\langle (1,2,1)^T \right\rangle \subset \mathbb{F}^3$, then $U^\circ = \langle (1,0,-1), (2,-1,0) \rangle \in \left(\mathbb{F}^3\right)^*$.

Proposition. Suppose V is a finite dimensional $\mathbb{F}$-vector space, and $U \leq V$. Then

$$\dim U + \dim U^\circ = \dim V$$

Proof. Let $(e_1, ..., e_k)$ be a basis for U and extend to a basis $(e_1, ..., e_n)$ for V. Let $(\varepsilon_1, ..., \varepsilon_n)$ be dual basis to $(e_1, ..., e_n)$ in V^*. We claim that $(\varepsilon_{k+1}, ..., \varepsilon_n)$ is a basis for U°.

To show this, since $\varepsilon_i(e_j) = 0$ for $1 \leq j \leq k$ and $k+1 \leq i \leq n$, so $\varepsilon_{k+1}, ..., \varepsilon_n \in U^\circ$. So it's enough to show that they span U°.

If $\theta \in U^\circ$, then $\theta(e_i) = 0$ for $1 \leq i \leq k$. But $\theta = \sum_{j=1}^n \lambda_j \varepsilon_j$ for some $\lambda_j \in \mathbb{F}$, and $\lambda_1, ..., \lambda_k = 0$. So $\theta \in \langle e_{k+1}, ..., e_n \rangle$.

So $\dim U^\circ + \dim U = (n-k) + k = n = \dim V$. $\square$

Another proof:

Proof. Let $r : V^* \to U^*$ by $\theta \to \theta|_U$. This is a linear surjection since every linear map $U \to \mathbb{F}$ can be extended to a linear map $V \to \mathbb{F}$. Moreover, $\ker r = U^\circ$. So by rank-nullity theorem,

$$\begin{aligned} \dim V^* &= \dim U^* + \dim U^\circ, \\ \dim V^* &= \dim V, \\ \dim U^* &= \dim U \end{aligned}$$

So done. ☐

Another proof:

Proof. There is a linear isomorphism $U^\circ \to (V/U)^*$ by $\theta \to \bar{\theta}$ where $\bar{\theta}(v+U) = \theta(v)$.

So $\dim U^\circ = \dim (V/U)^* = \dim V/U = \dim V - \dim U$. ☐

Proposition. (Change of dual basis) If V is finite dimensional $\mathbb{F}$-vector space with bases $(e_1, ..., e_n)$ and $(f_1, ..., f_n)$ and change of basis matrix from (e_k) to (f_k) given by P, i.e. $f_i = \sum_{k=1}^n P_{ki} e_k$ for $1 \le i \le n$, then if $(\varepsilon_1, ..., \varepsilon_n)$ and $(\eta_1, ..., \eta_n)$ are the corresponding dual bases, then the change of basis matrix from $(\varepsilon_1, ..., \varepsilon_n)$ to $(\eta_1, ..., \eta_n)$ is $\left(P^{-1}\right)^T$, i.e.

$$\varepsilon_i = \sum_{k=1}^n \left(p_{ki}^T\right) \eta_k$$

Proof. Let $Q = P^{-1}$ so $e_i = \sum_k Q_{ki} f_k$.

$$\begin{aligned} \left(\sum_{k=1}^n (P_{ki})^T \eta_k\right)(e_j) &= \sum_{k,l} (P_{ik}\eta_k)(Q_{lj} f_l) \\ &= \sum_{k,l} P_{ik}\delta_{kl} Q_{lk} \\ &= [PQ]_{ij} = \delta_{ij} = \varepsilon_i(e_j) \end{aligned}$$

for $1 \le j \le n$. So

$$\sum_{k=1}^n (P_{ki})^T \eta_k = \varepsilon_i$$

as required. ☐

3.2 Dual maps

Definition. If $\alpha \in \mathcal{L}(V, W)$ where V, W are $\mathbb{F}$-vector spaces, the *dual map*

$$\alpha^* : W^* \to V^*$$

is given by

$$\alpha^*(\theta) = \theta \circ \alpha$$

Note $\alpha^*(\theta) \in V^*$ for all $\theta \in W^*$ since composite of linear map is linear. Moreover, $\alpha^* \in \mathcal{L}(W^*, V^*)$ since $f: \lambda, \mu \in \mathbb{F}, \theta_1, \theta_2 \in W^*$ and $v \in V$, then

$$\begin{aligned}\alpha^*(\lambda\theta_1 + \mu\theta_2)(v) &= (\lambda\theta_1 + \mu\theta_2)(\alpha(v)) \\ &= \lambda\theta_1(\alpha(v)) + \mu\theta_2(\alpha(v)) \\ &= \lambda\alpha^*(\theta_1)(v) + \mu\alpha^*(\theta_2)(v)\end{aligned}$$

as required.

Proposition. Suppose V and W are vector spaces with bases $(e_1, ..., e_n)$ and $(f_1, ..., f_n)$ and $\alpha \in \mathcal{L}(V, W)$ represented by A with respect to these bases. Then $\alpha^* \in \mathcal{L}(W^*, V^*)$ is represented by A^T with the dual bases $(\varepsilon_1, ..., \varepsilon_n)$ and $(\eta_1, ..., \eta_n)$.

Proof. We have $\alpha(e_i) = \sum_{k=1}^n A_{ki} f_k$ and $\varepsilon_i(e_j) = \delta_{ij} = \eta_i(f_j)$. We want to show

$$\alpha^*(\eta_i) = \sum_{k=1}^n \left(A^T\right)_{ki} \varepsilon_k$$

Then

$$\begin{aligned}\alpha^*(\eta_i)(e_j) &= \eta_i(\alpha(e_j)) \\ &= \eta_i\left(\sum_{k=1}^n A_{kj} f_k\right) \\ &= \sum_k A_{kj}\delta_{ik} \\ &= A_{ij} \\ &= \sum_{k=1}^n \left(A^T\right)_{ki} \varepsilon_k(e_j)\end{aligned}$$

So $\alpha^*(\eta_i)$ and $\sum \left(A^T\right)_{ki} \varepsilon_k$ agree on a basis, so are equal. □

Remark. • If $\alpha \in \mathcal{L}(U, V)$ and $\beta \in \mathcal{L}(V, W)$, then $(\beta\alpha)^* = \alpha^*\beta^*$.
• If $\alpha, \beta \in \mathcal{L}(U, V)$, $\lambda, \mu \in \mathbb{F}$, then $(\lambda\alpha + \mu\beta)^* = \lambda\alpha^* + \mu\beta^*$.
• If $B = Q^{-1}AP$ with P, Q invertible, then

$$B^T = P^T A^T \left(Q^{-1}\right)^T = \left(\left(P^{-1}\right)^T\right)^{-1} A^T \left(Q^{-1}\right)^T$$

as we should expect.

Lemma. Suppose $\alpha = \mathcal{V}, \mathcal{W}$ and V, W are finite dimensional vector spaces. Then
(a) $\ker \alpha^* = (\operatorname{im}\alpha)^\circ$;
(b) $r(\alpha) = r(\alpha^*)$; and
(c) $\operatorname{im}\alpha^* = (\ker\alpha)^\circ$.

Proof. (a)Suppose $\theta \in W^*$. Then $\theta \in \ker\alpha^* \iff \theta\alpha = 0 \iff \theta\alpha(v) = 0 \forall v \in V \iff \theta \in (\operatorname{im}\alpha)^\circ$.

(b) We've seen $\dim \operatorname{im} \alpha + \dim (\operatorname{im} \alpha)^\circ = \dim W$ since $\operatorname{im} \alpha \leq W$. Using (a), we deduce $r(\alpha) + n(\alpha^*) = \dim W = \dim W^*$. But rank-nullity theorem implies $r(\alpha^*) + n(\alpha^*) = \dim W^*$.

(c) If $\theta \in \operatorname{im} \alpha^*$, then there exists $\phi \in W^*$ s.t. $\alpha^*(\phi) = \theta$ i.e. $\phi\alpha = \theta$.
So if $v \in \ker \theta$, then $\theta(v) = \phi\alpha(v) = 0$. Thus $\theta \in (\ker \alpha)^\circ$, i.e. $\operatorname{im} \alpha^* \leq (\ker \alpha)^\circ$.
But $\dim (\ker \alpha)^\circ + n(\alpha) = \dim V$, and $r(\alpha) + n(\alpha) = \dim V$. So $r(\alpha) = \dim (\ker \alpha)^\circ = r(\alpha^*)$ by (b).
So $\dim \operatorname{im} \alpha^* = \dim (\ker \alpha)^\circ$ and the inclusion is an equality. □

Lemma. Let V be a $\mathbb{F}$-vector space. There is a canonical linear map $ev : V \to (V^*)^*$ given by $ev(v)(\theta) = \theta(v)$ for all $\theta \in V^*, v \in V$.

Proof. We need to prove: (a) $ev(v) \in (V^*)^*$ for all $v \in V$, and (b) ev is actually a linear map.

$ev(v)$ is a function $V^* \to \mathbb{F}$, we want $ev(v)(\lambda\theta_1 + \mu\theta_2) = \lambda ev(v)(\theta_1) + \mu ev(v)(\theta_2)$ for all $\lambda, \mu \in \mathbb{F}$, $\theta_1, \theta_2 \in V^*$.

$$\begin{aligned} \text{LHS} &= (\lambda\theta_1 + \mu\theta_2)(v) \\ &= \lambda\theta_1(v) + \mu\theta_2(v) = \text{RHS}. \end{aligned}$$

Moreover, if $\lambda, \mu \in \mathbb{R}$ and $v_1, v_2 \in V$ and $\theta \in V^*$, then

$$\begin{aligned} ev(\lambda v_1 + \mu v_2)(\theta) &= \theta(\lambda v_1 + \mu v_2) \\ &= \lambda\theta(v_1) + \mu\theta(v_2) \\ &= \lambda ev(v_1)(\theta) + \mu ev(v_2)(\theta) \\ &= (\lambda ev(v_1) + \mu ev(v_2))(\theta) \end{aligned}$$

So $ev(\lambda v_1 + \mu v_2) = \lambda ev(v_1) + \mu ev(v_2)$ as required. □

Lemma. If V is finite dimensional on $\mathbb{F}$, then $ev : V \to V^{**}$ is an isomorphism.

Proof. We know $\dim V^{**} = \dim V^* = \dim V$, so it suffices to show ev is an injection, i.e. $ev(v) = 0 \implies v = 0$.

Suppose $ev(v) = 0$. Then $\theta(v) = ev(v)(\theta) = 0$ for all $\theta \in V^*$, i.e. $V^* = \langle v \rangle^\circ$.
But $\dim \langle v \rangle + \dim \langle v \rangle^\circ = \dim V = \dim V^*$. So $\dim \langle v \rangle = 0$, i.e. $v = 0$. □

Remark. If V is any vector space with a basis, then $ev(v)$ can be shown to be injective but it will not be surjective unless V is finite dimensional.

Lemma. Suppose V, W are vector spaces on $\mathbb{F}$ and $\alpha \in \mathcal{L}(V, W)$. Then

$$\alpha^{**} \circ ev = ev \circ \alpha$$

where α^{**} is the corresponding map from V^{**} to W^{**}.

Proof. Suppose $v \in V$, $\theta \in W^*$. Then

$$\begin{aligned}\alpha^{**} \circ ev\left(v\right)\left(\theta\right) &= ev\left(v\right) \circ \alpha^*\left(\theta\right)\\ &= ev\left(v\right)\left(\theta \circ \alpha\right)\\ &= \theta\left(\alpha\left(v\right)\right)\\ &= ev\left(\alpha\left(v\right)\right)\left(\theta\right)\end{aligned}$$

i.e. $\alpha^{**} \circ ev\left(v\right) = ev\left(\alpha\left(v\right)\right)$ for all $v \in V$. $\square$

Proposition. Suppose V is finite dimensional vector space over $\mathbb{F}$, and U, U_1, $U_2 \leq V$.
(a) $U^{\circ\circ} = ev\left(U\right)$;
(b) $ev\left(U^\circ\right) = ev\left(U^\circ\right)$;
(c) $\left(U_1 + U_2\right)^\circ = U_1^\circ \cap U_2^\circ$;
(d) $\left(U_1 \cap U_2\right)^\circ = U_1^\circ + U_2^\circ$.

Proof. (a) Let $u \in U$ and $\theta \in V^*$. Then $\theta\left(u\right) = 0 \iff ev\left(u\right)\left(\theta\right) = 0$.
So $\theta \in U^\circ \implies ev\left(u\right)\left(\theta\right) = 0$, and $ev\left(u\right) \in \left(U^\circ\right)^\circ$, i.e. $ev\left(U\right) \subset U^{\circ\circ}$.
But $\dim\left(ev\left(U\right)\right)\dim U = \dim V - \dim U^\circ = \dim V - \left(\dim V^* - \dim U^{\circ\circ}\right) = \dim U^{\circ\circ}$ and $ev\left(U\right) = U^{\circ\circ}$ as desired.

(b) $ev\left(U^\circ\right) = \left(U^\circ\right)^{\circ\circ} = \left(U^{\circ\circ}\right)^\circ = ev\left(U\right)^\circ$.

(c) Suppose $\theta \in V^*$, then $\theta \in \left(U_1 + U_2\right)^\circ \iff \theta\left(u_1 + u_2\right) = 0$ for all $u_1 \in U_1$, $u_2 \in U_2$, $\iff \theta\left(u_1\right) = 0$ for all $u_1 \in U_1$ and $\theta\left(u_2\right) = 0$ for all $u_2 \in U_2 \iff \theta \in U_1^\circ \cap U_2^\circ$.

(d) By (c), $U_1^{\circ\circ} \cap U_2^{\circ\circ} = \left(U_1^\circ + U_2^\circ\right)^\circ$. Thus $\left(ev\left(U_1\right) \cap ev\left(U_2\right)\right)^\circ = ev\left(U_1^\circ + U_2^\circ\right) = ev\left(U_1^\circ\right) + ev\left(U_2^\circ\right)$, $ev\left(U_1 \cap U_2\right)^\circ = ev\left(U_1^\circ\right) + ev\left(U_2^\circ\right)$, and $\left(ev\left(U_1 \cap U_2\right)\right)^\circ = ev\left(U_1\right)^\circ + ev\left(U_2\right)^\circ$ by part (b). So $\left(U_1 \cap U_2\right)^\circ = U_1^\circ + U_2^\circ$ since ev is an isomorphism and commute with $^\circ$. $\square$

4 Bilinear forms (I)

Suppose U, V are vector spaces on $\mathbb{F}$.

Definition. A bilinear form on $U \times V$ is a function $\phi : U \times V \to \mathbb{F}$ that is linear in each variable, i.e.

$$\phi(u, -) : V \to \mathbb{F} \in V^* \ \forall u \in U,$$
$$\phi(-, v) : U \to \mathbb{F} \in U^* \ \forall v \in V$$

Example.

$$V \times V^* \to \mathbb{F}$$
$$(v, \theta) \to \theta(v)$$

is a bilinear form.

Example. If $V = W = \mathbb{R}^n$, then

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
$$(x, y) \to \sum_{i=1}^{n} x_i y_i$$

is a bilinear form.

Example. If $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$, then

$$\mathbb{F}^m \times \mathbb{F}^n \to \mathbb{F}$$
$$(v, w) \to v^T A w$$

is a bilinear form.

Example. If $V = W = C([0,1], \mathbb{R})$, then

$$(f, g) \to \int_0^1 f(t) g(t) \, dt$$

is a bilinear form.

Example.

$$\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$
$$(a, b)^T, (c, d)^T \to ad - bc$$

is a bilinear form.

Definition. Suppose U has basis $(e_1, ..., e_n)$ and V has basis $(f_1, ..., f_m)$ and $\phi : U \times V \to \mathbb{F}$ is a bilinear form. The matrix A representing ϕ with respect to $(e_1, ..., e_n)$ and $(f_1, ..., f_m)$ is given by

$$A_{ij} = \phi(e_i, f_j) \ \forall 1 \le i \le n, 1 \le j \le m$$

Remark. If $u = \sum \lambda_i e_i$ and $v = \sum \mu_j f_j$, then

$$\begin{aligned} \phi(u,v) &= \phi\left(\sum \lambda_i e_i, \sum_j \mu_j f_j\right) \\ &= \sum_{i=1}^n \lambda_i \phi\left(e_i, \sum_j \mu_j f_j\right) \\ &= \sum_{i=1}^n \sum_{j=1^m} \lambda_i \mu_j \phi(e_i, f_j) \\ &= \sum_{i,j} \lambda_i A_{ij} \mu_j \\ &= \lambda^T A \mu \end{aligned}$$

So ϕ is uniquely determined by A.

Definition. A bilinear form $\phi : U \times V \to \mathbb{F}$ determines two linear maps

$$\begin{aligned} \phi_L &: U \to V^* \\ \phi_R &: V \to U^* \end{aligned}$$

given by

$$\phi_L(u)(v) = \phi(u,v) = \phi_R(v)(u)$$

Example. If

$$\begin{aligned} \phi : V \times V^* &\to \mathbb{F} \\ (v,\theta) &\to \theta(v) \end{aligned}$$

then $\phi_L : V \to V^{**}$ is ev.
Moreover, $\phi_R : V^* \to V^*$ is ι_{V^*} (identity).

In fact, $\phi_R = \phi_L^* \circ ev$ and $\phi_L = \phi_R^* \circ ev$. That motives the following lemma:

Lemma. Suppose $(e_1, ..., e_n)$ is a basis for U, and $(f_1, ..., f_n)$ is a basis for V, and $\phi : U \times V \to \mathbb{F}$ is a bilinear form. Then if A represents ϕ with respect to these bases then A also represents ϕ_R with respect to $(f_1, ..., f_m)$, and the basis $(\varepsilon_1, ..., \varepsilon_n)$ that is dual to $(e_1, ..., e_n)$ and A^T represents ϕ_L with respect to $(e_1, ..., e_n)$, and the basis $(\eta_1, ..., \eta_m)$ that is dual to $(f_1, ..., f_m)$.

Proof.

$$\phi_R(f_i)(e_j) = \phi(e_j, f_i) = A_{ji}$$

So

$$\phi_R(f_i) = \sum_k A_{ki} \varepsilon_k$$

as required.

Similarly,

$$\phi_L(e_i)(f_j) = \phi(e_i, f_j) = A_{ij} = A^T_{ji}$$

So
$$\phi_L(e_i) = \sum_k \left(A^T\right)_{ki} \eta_k$$
as required. □

We call $\ker \phi_L$ the *left kernel* of ϕ, and $\ker \phi_R$ the *right kernel* of ϕ.

Note
$$\ker \phi_L = \{u \in U \mid \phi(u,v) = 0 \forall v \in V\},$$
$$\ker \phi_R = \{v \in V \mid \phi(u,v) = 0 \forall u \in U\}$$

More generally, if $T \subset U$, then define
$$T^\perp = \{v \in V \mid \phi(t,v) = 0 \forall t \in T\} \leq V$$
and
$$^\perp S = \{u \in U \mid \phi(u,s) = 0 \forall s \in S\}$$

Definition. We say ϕ is *non-degenerate* if $\ker \phi_L = 0$ and $\ker \phi_R = 0$. Otherwise we say ϕ is *degenerate*.

Lemma. If $\phi : U \times V \to \mathbb{F}$ is a bilinear form, $(e_1, ..., e_n)$ is a basis for U, $(f_1, ..., f_n)$ is a basis for V, and ϕ is represented by A with respect to the bases. Then ϕ is non-degenerate $\iff$ A is invertible. In particular, A has to be a square matrix $(n = m)$.

Proof. ϕ is non-degenerate $\iff$ $\ker \phi_L = 0$ and $\ker \phi_R = 0$
$\iff$ $r\left(A^T\right) = n$ and $r(A) = n$ by rank-nullity
$\iff$ $n = m = r(A)$ since row rank = column rank
$\iff$ A is invertible. □

So to define a bilinear form $\phi : U \times V \to \mathbb{F}$ that is non-degenerate is to define an isomorphism $\phi_L : U \to V^*$ (equivalently an isomorphism ϕ_R from V to U^*) when U and V are finite dimensional.

Proposition. (Change of basis) Suppose $(e_1, ..., e_n)$ and $(u_1, ..., u_n)$ are bases for U and $(f_1, ..., f_m)$ and $(v_1, ..., v_m)$ are bases for V s.t.
$$u_i = \sum_{k=1}^{n} P_{ki} e_k$$
for $i = 1, ..., n$,
$$v_j = \sum_{i=1}^{n} Q_{lj} f_l$$
for $j = 1, ..., m$. Let ϕ be a bilinear form $u \times V \to \mathbb{F}$ represented by A wrt $(e_1, ..., e_n)$ and $(f_1, ..., f_m)$ and by B wrt $(u_1, ..., u_n)$ and $(v_1, ..., v_m)$. Then $B = P^T A Q$.

Proof.

$$\begin{aligned}
B_{ij} &= \phi(u_i, v_j) \\
&= \phi\left(\sum_k P_{ki} e_k, \sum_l Q_{lj} f_l\right) \\
&= \sum_{k,l} P_{ki} Q_{lj} \phi(e_k, f_l) \\
&= \sum_{k,l} P^T_{ik} A_{kl} Q_{lj} \\
&= \left[P^T A Q\right]_{ij}
\end{aligned}$$

□

So now we can define

Definition. Let $\phi : U \times V \to \mathbb{F}$ be a bilinear form with U, V finite dimensional. The *rank* of ϕ, $r(\phi)$ is the rank of any matrix representing ϕ. By the above lemma, this does not depend on the choice of basis.

Remark. $r(\phi) = r(\phi_R) = r(\phi_L)$.

5 Determinants of matrices

Recall that S_n is the group of permutations of $\{1, ..., n\}$ and there is a group homomorphism $\varepsilon : S_n \to (\{\pm 1\}, \cdot)$ s.t. $\varepsilon(\sigma) = 1$.
If σ is a product of an even number of transposition then $\varepsilon(\sigma) = -1$; otherwise $\varepsilon(\sigma) = 1$.

Definition. Let $A \in \mathrm{Mat}_n(\mathbb{F})$. The *determinant* of A is

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) \left(\cap_{i=1}^n A_i \sigma(i)\right)$$

Example. Consider $n = 2$. We have

$$\det A = A_{11} A_{22} - A_{12} A_{21}$$

Lemma. $\det A = \det A^T$.

Proof.

$$\begin{aligned}
\det A^T &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \left(\cap_{i=1}^n A_{\sigma(i)i}\right) \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma) \left(\cap_{j=1}^n A_{j\varepsilon^{-1}(j)}\right) \\
&= \sum_{\tau \in S_n} \varepsilon\left(\tau^{-1}\right) \left(\cap_{j=1}^n A_{j\tau(j)}\right) \\
&= \det A
\end{aligned}$$

□

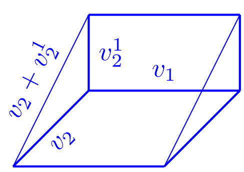
Definition. A *volume form* on $\mathbb{F}^n$ is a function

$$d : \mathbb{F}^n \times ... \times \mathbb{F}^n \to \mathbb{F}$$

that is
(a) multilinear, i.e. if $v_1, ..., v_{i-1}, v_{i+1}, ..., v_n \in \mathbb{F}^n$, then

$$d(v_1, ..., v_{i-1}, -, v_{i+1}, ..., v_n) \in (\mathbb{F}^n)^*$$

For each $1 \leq i \leq n$;
(b) alternating, i.e. whenever $v_i = v_j$ for some $i \neq j$, then $d(v_1, ..., v_n) = 0$.



One may view a matrix $A \in \mathrm{Mat}_n(\mathbb{F})$ as an n-tuple of elements of $\mathbb{F}^n$ (its columns):

$$A = \left(A^{(1)}, ..., A^{(n)}\right)$$

Lemma. det is a volume form.

Proof. To see det is multi-linear, it suffices to see that

$$\cap_{i=1}^{n} A_{i\sigma(1)}$$

is multilinear for each $\sigma \in S_n$, since any linear combination of (multi)-linear functions is also (multi)-linear.

But $\cap_{i=1}^{n} A_{i\sigma(1)}$ contains one entry from each column, so is clearly multi-linear.

Suppose $A^{(k)} = A^{(l)}$ for some $k \neq l$. Let $\tau = (kl)$. Then $A_{ij} = A_{i\tau(j)}$ for all $1 \leq i, j \leq n$.

But S_n is the disjoint union of A_n and τA_n, and

$$\sum_{\sigma \in A_n} \cap_{i=1}^{n} A_{i\sigma(i)} = \sum_{\sigma \in \tau A_n} \cap_{i=1}^{n} A_{i\sigma(i)}$$

so

$$\det A = LHS - RHS = 0$$

□