Linear Algebra ①

Chapter 1 : Vector Spaces and Linear Maps

1.1 Vector Spaces

A vector space over a field F is a set V of vectors equipped with :

a) The structure of an Abelian group, which gives
- addition $V \times V \to V$, $(\underline{u}, \underline{v}) \mapsto \underline{u} + \underline{v}$
- zero $\underline{0} \in V$ (additive identity)
- inverse $V \to V$, $\underline{u} \mapsto -\underline{u}$

b) Scalar Multiplication $F \times V \to V$, $(\lambda, \underline{u}) \mapsto \lambda \underline{u}$
satisfying $1\underline{u} = \underline{u}$, $\lambda(\mu \underline{u}) = (\lambda\mu)\underline{u}$ (action)
$(\lambda + \mu)\underline{u} = \lambda \underline{u} + \mu \underline{u}$
$\lambda(\underline{u} + \underline{v}) = \lambda \underline{u} + \lambda \underline{v}$ (distributivity / bilinearity)

(F will be specified when necessary. We will usually use $\mathbb{R}$ or $\mathbb{C}$ and sometimes $\mathbb{Q}$ or F_p).

Indicative Consequences

$0\underline{x} = \underline{0}$ (for $0\underline{x} + 0\underline{x} = (0+0)\underline{x} = 0\underline{x}$)
$(-1)\underline{x} = -\underline{x}$ (for $\underline{x} + (-1)\underline{x} = (1-1)\underline{x} = 0\underline{x} = \underline{0}$)
Finite linear combinations $\sum_{i=1}^{n} \lambda_i \underline{x}_i$ make sense and can be manipulated as we would expect.
(N.B. $\sum_{i=1}^{0} \lambda_i \underline{x}_i = 0$ and expressions make sense in that case)

Canonical Example : The space F^n of column vectors

$\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ under coordinate wise operations. Note that there are n vectors $\underline{u}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \underline{u}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ such that any $\underline{x}$ as above

can be written uniquely as a finite linear combination $\underline{x} = \sum_{i=1}^{n} \lambda_i \underline{u_i}$.
So F^n has dimension n.

The dimension $\dim V$ of a vector space V is an invariant, and if $\dim V = n$, finite, then $V \cong F^n$ (We shall prove this).

More generally, for any set I, the functions $F^I = \{f : I \to F\}$ form a vector space under pointwise operations
$(f+g)(i) = f(i) + g(i)$, $(\lambda f)(i) = \lambda f(i)$
Note that the cardinality of I is not the dimension; $F^{\mathbb{N}}$ has uncountable dimension.

Observations

1. $\mathbb{R}$ is a subfield of $\mathbb{C}$, and so $\mathbb{C}$ is a vector space over $\mathbb{R}$ (of dimension 2). Similarly, $\mathbb{C}^n$ is a vector space of $\mathbb{R}$ of dimension $2n$.

2. $\mathbb{Q}$ is a subfield of $\mathbb{R}$ and so $\mathbb{R}$ is a vector space over $\mathbb{Q}$. As such, its dimension is uncountable.

3. Not that it is not possible to give an explicit basis for $F^{\mathbb{N}}$ over F or $\mathbb{R}$ over $\mathbb{Q}$.

1·2 Subspaces

A subset U of a vector space V is a subspace $U \leq V$ just when $\underline{0} \in U$, $\underline{x}, \underline{y} \in U \Rightarrow \underline{x} + \underline{y} \in U$, and $\underline{x} \in U \Rightarrow \lambda \underline{x} \in U \ \forall \lambda$.
In other words, U is closed under the basic operations and so under finite linear combinations.

Fact: U is a subspace of V if and only if U is non-empty and $\underline{x}, \underline{y} \in U \Rightarrow \lambda \underline{x} + \mu \underline{y} \in U \ \forall \lambda, \mu$.

This is often the definition as it is quicker to check. We need $U \neq \emptyset$, then we pick $\underline{x} \in U$ and find that $\underline{0} = 0\underline{x} \in U$

Examples

Let X be a space like $[0,1]$, $\mathbb{R}$, $\mathbb{C}$. Then the collection of functions $X \to \mathbb{R}$ (or $\mathbb{C}$) which are:

a) integrable b) continuous c) analytic
d) differentiable e) polynomials

are vector subspaces of $\mathbb{R}^X$ or $\mathbb{C}^X$.

Remark: For $\mathbb{R}$, $\mathbb{C}$, the polynomial functions have a basis $1, x, x^2, \ldots$ and the space is countable dimensional. What about polynomial functions $F_p \to F_p$? (!)

1.3 Linear Maps

Let U, V be vector spaces. A map $\alpha : U \to V$ is linear if and only if $\alpha(\underline{x}+\underline{y}) = \alpha(\underline{x}) + \alpha(\underline{y})$, $\alpha(\lambda\underline{x}) = \lambda\alpha(\underline{x})$ i.e. α is an abelian group homomorphism preserving scalar multiplication

Equivalently, α is linear just when it preserves finite linear combinations: $\alpha(\sum_i \lambda_i \underline{x}_i) = \sum_i \lambda_i \alpha(\underline{x}_i)$
Equivalently, just when $\alpha(\lambda\underline{u} + \mu\underline{v}) = \lambda\alpha(\underline{u}) + \mu\alpha(\underline{v})$

Canonical Example

The linear maps $\alpha : F^n \to F^m$ are given by matrix multiplication

1. If (a_{ij}) is an $m \times n$ matrix then $\underline{x} \mapsto A\underline{x}$ is linear as a map $F^n \to F^m$.

$A\underline{x} = \begin{pmatrix} \sum a_{1j}x_j \\ \vdots \\ \sum a_{mj}x_j \end{pmatrix}$ and we just check that
$$A(\lambda\underline{x} + \mu\underline{y}) = \lambda A\underline{x} + \mu A\underline{y}$$

2. If α is linear, note the following :
Any $\underline{x} \in F^n$ is written $\underline{x} = \sum x_j u_j$ and so by linearity
$\alpha(\underline{x}) = \sum x_j \alpha(\underline{u}_j)$. Let $A = (\alpha(\underline{u}_1) \mid \ldots \mid \alpha(\underline{u}_n))$
with columns $\alpha(\underline{u}_j) = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$

and then $\alpha(\underline{x}) = A\underline{x} = \begin{pmatrix} \sum a_{1j} x_j \\ \vdots \\ \sum a_{nj} x_j \end{pmatrix}$

A takes the standard basis vectors to the columns of A.

Linear Algebra ②

Examples of Linear Maps from Calculus

1. Differentiation, $D = \frac{d}{dt} : C^1(\mathbb{R}) \to C^0(\mathbb{R})$ or $C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ where $C^1(\mathbb{R})$ is the set of continuously differentiable functions.

2. Integration $\int_0^x \square \, dt : C^0(\mathbb{R}) \to C^1(\mathbb{R})$
   or $\int_0^1 \square \, dt : C^0(\mathbb{R}) \to \mathbb{R}$

Closure Properties

Proposition

i) If V is a vector space then the identity map $I = I_V : V \to V$ is linear.

ii) If $\alpha : U \to V$, $\beta : V \to W$ are linear, then so is $\beta\alpha : U \to W$

Proof (of (ii))

Take $\underline{x}, \underline{y} \in U$.
$\beta\alpha(\lambda \underline{x} + \mu \underline{y}) = \beta(\alpha(\lambda \underline{x} + \mu \underline{x}) = \beta(\alpha(\lambda \underline{x}) + \alpha(\mu \underline{y})) = \lambda \beta\alpha(\underline{x}) + \mu\beta\alpha$
In the canonical example, let $\alpha : F^n \to F^m$ be $\alpha(\underline{x}) = A\underline{x}$ and $\beta : F^m \to F^p$ be $\beta(\underline{y}) = B\underline{y}$.

Then $\beta\alpha : F^n \to F^p$ is $\beta\alpha(\underline{x}) = (BA)\underline{x}$, for
$$\beta\alpha(\underline{u}_j) = \beta \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} \sum_1^m b_{1k} a_{kj} \\ \cdots \\ \sum_1^m b_{mk} a_{kj} \end{pmatrix},$$ the j^{th} column of BA.

Proposition

Let U, V be vector spaces. Then $0 : U \to V$, $\underline{x} \mapsto \underline{0}$ is linear, and if $\alpha, \beta : U \to V$ are linear, so is $\lambda\alpha + \mu\beta$, with proof as before

Write $\mathcal{L}(U, V)$ for the set of linear maps from U to V.

Corollary

$\mathcal{L}(U, V) \leq V^U$, the space of all maps U to V, and is in particular a vector space.

In the canonical example, $\mathcal{L}(F^n, F^m)$ is isomorphic to the vector space $M_{m \times n}(F)$ of all $m \times n$ matrices with entries in F.

1.4 Kernels and Images

Definition

Let $\alpha : U \to V$ be linear. We define:
- the kernel, $\ker \alpha = \{\underline{u} \in U \mid \alpha(\underline{u}) = 0\}$
- the image of α, $\operatorname{im} \alpha = \{\underline{v} \in V \mid \underline{v} = \alpha(\underline{u}) \text{ for some } \underline{u} \in U\}$

Proposition

If $\alpha : U \to V$ is linear, then $\ker \alpha \leq U$, $\operatorname{im} \alpha \leq V$

Proof

Let $\underline{x}, \underline{y} \in \ker \alpha$. Then $\alpha(\lambda \underline{x} + \mu \underline{y}) = \lambda \alpha(\underline{x}) + \mu \alpha(\underline{y})$ $= \lambda \underline{0} + \mu \underline{0} = \underline{0}$, so $\lambda \underline{x} + \mu \underline{y} \in \ker \alpha$. Moreover, $\alpha(\underline{0}) = \underline{0}$, so $\underline{0} \in \ker \alpha$, thus $\ker \alpha \leq U$.

Let $\underline{v}, \underline{w} \in \operatorname{Im} \alpha$. Take $\underline{x}, \underline{y} \in U$ with $\alpha(\underline{x}) = \underline{v}$, $\alpha(\underline{y}) = \underline{w}$. Then $\alpha(\lambda \underline{x} + \mu \underline{y}) = \lambda \alpha(\underline{x}) + \mu \alpha(\underline{y}) = \lambda \underline{v} + \mu \underline{w}$ which is in $\operatorname{im} \alpha$. Also $\alpha(\underline{0}) = \underline{0}$, so $0 \in \operatorname{im} \alpha$ and $\operatorname{im} \alpha \leq V$.

Linear Algebra ②

Examples of kernels

1. The set of solutions $(x_1, \dots, x_n)$ of the linear equations

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ \vdots \qquad\qquad \vdots \quad &\ \ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{aligned}$$

(m linear equations in n unknowns)

is the kernel of the map $\underline{x} \mapsto A\underline{x}$, where A is the expected matrix, so it is a subspace of F^n.

2. Consider $D = \frac{d}{dt} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$
By closure properties $\alpha_n D^n + \alpha_{n-1} D^{n-1} + \dots + \alpha_0 I : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ is linear.
In fact, this is still so if the α_i are $C^\infty(\mathbb{R})$ functions of t.
Then the solutions to the differential equation
$\alpha_n \frac{d^n x}{dt^n} + \dots + \alpha_0 x = 0$ form the kernel of the linear map.
In the constant coefficient case, this kernel has dimension n.

Why is this true for $\ddot{x} - 5\dot{x} + 6 = 0$, $\ddot{x} + x = 0$?
Hint : Why is this true for $\dot{x} - x = 0$?
Extension : Recurrence relations

Clearly $\alpha : U \to V$ is surjective $\Leftrightarrow$ im $\alpha = V$

Proposition

$\alpha : U \to V$ is injective $\Leftrightarrow$ $\ker \alpha = \{\underline{0}\}$

Proof

($\Rightarrow$) If $\underline{x} \in \ker \alpha$, $\alpha(\underline{x}) = \underline{0} = \alpha(\underline{0})$, then by injectivity $\underline{x} = \underline{0}$, so $\ker \alpha = \{\underline{0}\}$

($\Leftarrow$) If $\alpha(\underline{x}) = \alpha(\underline{y})$, then $\alpha(\underline{x} - \underline{y}) = \underline{0}$, $\underline{x} - \underline{y} \in \ker \alpha$
then $\underline{x} - \underline{y} = \underline{0}$, $\underline{x} = \underline{y}$, so α is injective.

Isomorphisms

An isomorphism $\alpha : U \to V$ of vector spaces is a linear map with inverse $\hat{\alpha} : V \to U$ (also linear)
(Then $\alpha\hat{\alpha} = I_U$, $\hat{\alpha}\alpha = I_V$)

Proposition

Suppose $\alpha : U \to V$ is a bijective linear map. Then the inverse $\hat{\alpha} : V \to U$ is linear (and so α is an isomorphism).

Proof

Take $\underline{v}, \underline{w} \in V$ and let $\underline{x} = \hat{\alpha}(\underline{v})$, $\underline{y} = \hat{\alpha}(\underline{w})$ so that $\alpha(\underline{x}) = \underline{v}$, $\alpha(\underline{y}) = \underline{w}$.
Then $\hat{\alpha}(\lambda\underline{v} + \mu\underline{w}) = \hat{\alpha}(\lambda\alpha(\underline{x}) + \mu\alpha(\underline{y})) = \lambda\hat{\alpha}(\underline{v}) + \mu\hat{\alpha}(\underline{w})$

$\square$

Corollary

$\alpha : U \to V$ is an isomorphism $\Leftrightarrow$ $\operatorname{im}\alpha = V$, $\ker\alpha = \{\underline{0}\}$

Linear Algebra ③

The Isomorphism Theorem

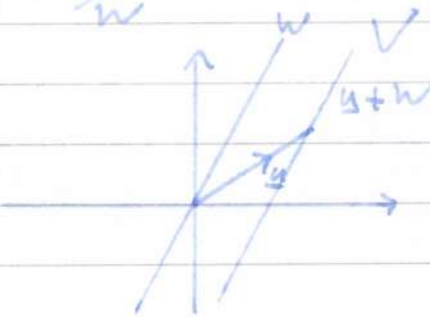
Quotients

Let $W \leq V$. The cosets of W in V are the affine subspaces $\underline{v} + W = \{\underline{v} + \underline{w} \mid \underline{w} \in W\}$. As with groups, the collection V/W is a vector space with operations $(\underline{u} + W) + (\underline{v} + W) = (\underline{u} + \underline{v}) + W$, $\lambda(\underline{u} + W) = \lambda\underline{u} + W$ similar to 'arithmetic modulo W'.

The quotient map $q: V \to V/W$ is linear. $q: \underline{v} \mapsto \underline{v} + W$

$\ker q = \{\underline{v} \mid \underline{v} + W = W\} = W$

$\operatorname{im} q = V/W$



Isomorphism Theorem

Let $\alpha: U \to V$ be linear. Then α factors as:

$$\begin{tikzcd} U \arrow[r, "\alpha"] \arrow[d, "q"'] & V \\ U/\ker\alpha \arrow[r, "\bar{\alpha}"] & \operatorname{Im}\alpha \arrow[u, "\text{inclusion}"'] \end{tikzcd}$$

where $\bar{\alpha}$ is an isomorphism.

What is $\bar{\alpha}$? $\bar{\alpha}(\underline{u} + \ker\alpha) = \alpha(\underline{u})$

$\underline{u} + \ker\alpha = \underline{u}' + \ker\alpha \Leftrightarrow \underline{u} - \underline{u}' \in \ker\alpha \Leftrightarrow \alpha(\underline{u} - \underline{u}') = \underline{0}$
$\Leftrightarrow \alpha(\underline{u}) = \alpha(\underline{u}') \Leftrightarrow \bar{\alpha}(\underline{u} + \ker\alpha) = \bar{\alpha}(\underline{u}' + \ker\alpha)$

$\Rightarrow$ says that $\bar{\alpha}$ is well defined. $\Leftarrow$ says that $\bar{\alpha}$ is injective.

This is evidently surjective and linear, so $\bar{\alpha}$ is an isomorphism.

1.6 Operations on Spaces

Proposition

Let $U, V \leq W$. Then: i) $U \cap V \leq W$ ii) $U + V \leq W$
where $U + V = \{\underline{u} + \underline{v} \mid \underline{u} \in U, \underline{v} \in V\}$

Proof

i) $\underline{0} \in U$, $\underline{0} \in V$, so $\underline{0} \in U \cap V$. Take $\underline{x}, \underline{y} \in U \cap V$. Then, $\underline{x}, \underline{y} \in U$ so $\lambda \underline{x} + \mu \underline{y} \in U$, and similarly for V. Therefore $\lambda \underline{x} + \mu \underline{y} \in U \cap V$.

ii) $\underline{0} = \underbrace{\underline{0}}^{\in U} + \underbrace{\underline{0}}^{\in V} \in U + V$. Let $\underline{x} = \underline{u} + \underline{v}$, $\underline{y} = \underline{u}' + \underline{v}'$ with $\underline{u}, \underline{u}' \in U$, $\underline{v}, \underline{v}' \in V$. Then $\lambda \underline{x} + \mu \underline{y} = (\lambda \underline{u} + \mu \underline{u}') + (\lambda \underline{v} + \mu \underline{v}')$ and $\lambda \underline{x} + \mu \underline{y} \in U + V$

Remark

$U \cap V$ is the greatest subspace $\leq U, V$. $U + V$ is the ~~greatest~~ least subspace $\geq U, V$.

Proposition

Let $U, V \leq W$. Then $U \cap V = \{\underline{0}\} \iff$ every $\underline{w} \in U + V$ has a unique representation as $\underline{w} = \underline{u} + \underline{v}$, $\underline{u} \in U$, $\underline{v} \in V$.

Proof

($\Rightarrow$) Suppose $\underline{w} = \underline{u} + \underline{v} = \underline{u}' + \underline{v}'$. Then $\underline{u} - \underline{u}' = \underline{v} - \underline{v}' \in U \cap V$. So $\underline{u} - \underline{u}' = \underline{v} - \underline{v}' = 0 \Rightarrow \underline{u} = \underline{u}'$, $\underline{v} = \underline{v}'$

Linear Algebra ③

($\Leftarrow$) Take $\underline{w} \in U \cap V$. $\underline{w} = \underline{w} + \underline{0} = \underline{0} + \underline{w}$ are two expressions. So by uniqueness, $\underline{w} = \underline{0}$.

When the above holds, we say that $U + V$ is the (internal) direct sum of U and V, $U + V = U \oplus V$. When $U \oplus V = W$, we say that V is a complement of U in W.

When we have $U_1, \ldots, U_n \leq W$, we clearly have $\bigcap_{i=1}^{n} U_i \leq W$ and $U_1 + \ldots + U_n \leq W$.

$U_1 + \ldots + U_n$ is a direct sum $\Leftrightarrow$ every $\underline{w} \in U_1 + \ldots + U_n$ has a unique representation as $\underline{w} = \underline{u}_1 + \ldots + \underline{u}_n$, $\underline{u}_i \in U_i$.
N.B. $U_1 \oplus \ldots + U_n = (\ldots((U_1 \oplus U_2) \oplus U_3) \ldots \oplus U_n)$

External Direct Sum

Let U, V be vector spaces. We define $U \oplus V$ by $\{ (\underline{u}, \underline{v}) \mid \underline{u} \in U, \underline{v} \in V \}$ under pointwise operations:

$(\underline{u}, \underline{v}) + (\underline{u}', \underline{v}') = (\underline{u} + \underline{u}', \underline{v} + \underline{v}')$, $\lambda(\underline{u}, \underline{v}) = (\lambda \underline{u}, \lambda \underline{v})$
Below are all linear maps:

$$\begin{tikzcd}
U \arrow[dr] & & V \arrow[dl] \\
& U \oplus V \arrow[dl] \arrow[dr] & \\
U & & V
\end{tikzcd}$$

$\underline{u} \mapsto (\underline{u}, \underline{0})$; $\underline{v} \mapsto (\underline{0}, \underline{v})$ $\leftarrow$ (Sum)

$(\underline{u}, \underline{v}) \mapsto \underline{u}$; $(\underline{u}, \underline{v}) \mapsto \underline{v}$ $\leftarrow$ (Product)

Suppose $U, V \leq W$. Then there is a map:

$U \oplus_{\text{external}} V \to U + V \leq W$, $(\underline{u}, \underline{v}) \mapsto \underline{u} + \underline{v}$
This is evidently surjective.

What is the kernel?

The kernel is $\{(\underline{u}, \underline{v}) \mid \underline{u} + \underline{v} = \underline{0}\} = \{(\underline{w}, -\underline{w}) \mid \underline{w} \in U \cap V\}$

The kernel is $\{(\underline{0}, \underline{0})\} \iff U \cap V = \{\underline{0}\}$, and is an isomorphism in this case, i.e. when the internal direct sum exists, it is the same as the external direct sum.

Linear Algebra ④

Dimension

Linear Independence and Spanning

Definition

Let $\underline{x}_1, \dots, \underline{x}_n \in V$, a vector space. The (linear) span or subspace spanned by the $\underline{x}_i$ is

$\langle \underline{x}_1, \dots, \underline{x}_n \rangle = \{ \sum_i \lambda_i \underline{x}_i \mid \lambda_i \in F \}$, the set of linear combinations of the $\underline{x}_i$. We say that $\underline{x}_1, \dots, \underline{x}_n$ span V just when $\langle \underline{x}_1, \dots, \underline{x}_n \rangle = V$.

Observe that $\langle \underline{x}_1, \dots, \underline{x}_n \rangle \leq V$. For $\underline{0} = \sum_i 0 \underline{x}_i$ is a linear combination, and $\lambda \sum_i \lambda_i \underline{x}_i + \mu \sum_i \mu_i \underline{x}_i = \sum_i (\lambda \lambda_i + \mu \mu_i) \underline{x}_i$, so linear combinations are closed under addition and scalar multiplication.

We can extend to arbitrary subsets. If $X \subseteq V$, then

$\langle X \rangle = \{ \sum_i \lambda_i \underline{x}_i \mid \underline{x}_i \in X, \lambda_i \in F \} \leq V$.

Note that $\langle \emptyset \rangle = \{ \underline{0} \}$

Suppose that $X, Y \subseteq V$ and $X \subseteq \langle Y \rangle$. Then $\langle X \rangle \leq \langle Y \rangle$.
For, suppose $\sum_i \lambda_i \underline{x}_i \in \langle X \rangle$, with $\underline{x}_i \in X$. Then we can write $\underline{x}_i = \sum_j \mu_{ij} \underline{y}_j$ with the $\underline{y}_j \in Y$.

Then $\sum_i \lambda_i \underline{x}_i = \sum_j (\sum_i \lambda_i \mu_{ij}) \underline{y}_j \in \langle Y \rangle$.

Note the special case: if $\underline{y} \in \langle \underline{x}_1, \dots, \underline{x}_n \rangle$ then $\langle \underline{x}_1, \dots, \underline{x}_n, \underline{y} \rangle = \langle \underline{x}_1, \dots, \underline{x}_n \rangle$

Definition

Let $\underline{x}_1, \dots, \underline{x}_n \in V$, a vector space. We say $\underline{x}_1, \dots, \underline{x}_n$ are (linearly) independent just when $\sum_{i=1}^{n} \lambda_i \underline{x}_i = \underline{0} \Rightarrow \lambda_i = 0 \ \forall i$.
i.e. no non-trivial linear combination of the $\underline{x}_i$ is equal to $\underline{0}$.

If $\underline{x_1}, \dots, \underline{x_n}$ are not independent, we say that they are dependent.

Proposition

Take $\underline{x_1}, \dots, \underline{x_n} \in V$. Then $\underline{x_1}, \dots, \underline{x_n}$ is linearly independent $\Leftrightarrow \underline{x_k} \notin \langle \underline{x_1}, \dots, \underline{x_{k-1}}, \underline{x_{k+1}}, \dots, \underline{x_n} \rangle \; \forall k$.
i.e. No $\underline{x_k}$ is a linear combination of the other $\underline{x_i}$.

Proof

($\Rightarrow$) Suppose, for contradiction, that $\underline{x_k} \in \langle \underline{x_1}, \dots, \underline{x_{k-1}}, \underline{x_{k+1}}, \dots, \underline{x_n} \rangle$ so that $\underline{x_k} = \sum_{i \neq k} \lambda_i \underline{x_i}$. Then $1 \cdot \underline{x_k} + \sum_{i \neq k} (-\lambda_i) \underline{x_i} = \underline{0}$, and so $\underline{x_1}, \dots, \underline{x_n}$ is linearly dependent.

($\Leftarrow$) Suppose $\sum_{i=1}^{n} \lambda_i \underline{x_i} = \underline{0}$. If $\lambda_k \neq 0$, then
$\underline{x_k} = \sum_{i \neq k} (-\lambda_i/\lambda_k) \underline{x_i} \in \langle \underline{x_1}, \dots, \underline{x_{k-1}}, \underline{x_{k+1}}, \dots, \underline{x_n} \rangle$ ※.
So all $\lambda_k = 0$.

Remark

The definition makes sense as stated for an indexed family (or list) $\underline{x_1}, \dots, \underline{x_n}$. So if $\underline{x_i} = \underline{x_j}$ for $i \neq j$, then $\underline{x_1}, \dots, \underline{x_n}$ is dependent because $1 \cdot \underline{x_i} + (-1) \underline{x_j} = \underline{0}$.

The definition varies for sets thought of as indexed families of distinct vectors. If $X \subseteq V$, then X is linearly independent $\Leftrightarrow$ whenever $\underline{x_1}, \dots, \underline{x_n}$ are distinct elements of X, then $\sum_i \lambda_i \underline{x_i} = \underline{0}$ $\Rightarrow \lambda_i = 0 \; \forall i$.

Linear Algebra ④

Remarks

- $\emptyset$ is independent.

- Any set X containing $\underline{0}$ is dependent because $1 \cdot \underline{0} = \underline{0}$.

- $\{\underline{a}, 2\underline{a}\}$ is dependent.

- These notions are finitary, i.e. X is linearly independent $\Leftrightarrow$ every finite subset of it is.

Lemma

Suppose $X \subseteq V$ is independent, and $\underline{y} \notin \langle X \rangle$. Then $X \cup \{\underline{y}\}$ independent.

Proof

Suppose $\sum \lambda_i \underline{x_i} + \mu \underline{y} = \underline{0}$, $\underline{x_i} \in X$ distinct. If $\mu \neq 0$, then $\underline{y} = \sum_i (-\lambda_i/\mu) \underline{x_i} \in \langle X \rangle$ #
So $\mu = 0$, and so $\sum \lambda_i \underline{x_i} = \underline{0}$, and as X is independent, $\lambda_i = 0 \ \forall i$.

Corollary

$\underline{x_1}, \ldots, \underline{x_n}$ is independent $\Leftrightarrow \underline{x_k} \notin \langle \underline{x_1}, \ldots, \underline{x_{k-1}} \rangle$ for any $1 \leq k \leq n$.

This suggests the following process. Given a vector space V, we construct a sequence X_r of linearly independent subsets with $|X_r| = r$, either terminating in X_n or for $\forall r \in \mathbb{N}$, as follows.

Set $X_0 = \emptyset$. Suppose X_r is defined. Then <u>either</u> $\langle X_r \rangle = V$, and we terminate, setting $n = r$, or $\langle X_r \rangle \neq V$, and we pick ~~pick~~ $\underline{x}_{r+1} \notin \langle X_{r+1} \rangle$, and set $X_{r+1} = X_r \cup \{\underline{x}_{r+1}\}$, which is independent.

Then, either we stop with $\langle X_n \rangle = V$, $X_n = \{\underline{x}_1, \ldots, \underline{x}_n\}$, linearly independent, <u>or</u> $\bigcup_{i=1}^{\infty} X_i = \{\underline{x}_1, \ldots, \underline{x}_i, \ldots\}$, an infinite independent set.

Linear Algebra ⑤

Bases

Definition

Let $e_1, \ldots, e_n \in V$, a vector space. $e_1, \ldots, e_n$ form a basis just when they are independent and span.

Note

We usually consider vectors indexed over $1, \ldots, n$. The e_i must then be distinct. But the definition makes sense for arbitrary $E \subset V$. E is a basis just when it is independent and spanning.

Examples

In $\mathbb{R}^3$:
- $(1,1,1), (1,1,0)$ is independent, but not spanning
- $(1,1,1), (1,1,0), (1,0,0)$ is a basis
- $(1,1,1), (1,1,0), (1,0,0), (1,0,1)$ spans, but is not independent
- $(1,1,1), (1,1,0), (1,0,1), (0,1,1)$ is not independent

Proposition

Suppose $e_1, \ldots, e_n$ is a minimal spanning set in V. Then, it is a basis for V.

Proof:

If it were not independent then $e_k \in \langle e_1, e_2, \ldots, e_{k-1}, e_{k+1}, \ldots, e_n \rangle$ and then $e_1, \ldots, e_{k-1}, e_{k+1}, \ldots, e_n$ spans, contradicting minimality.

Definition

A vector space is finite dimensional $\Leftrightarrow$ it has a finite spanning set.

Corollary

If V is a finite dimensional vector space then it has a finite basis.

Proof:

Let X be a spanning set, and take E, a minimal spanning subset.

Proposition

Suppose $\underline{e}_1, \ldots, \underline{e}_n$ is a maximal independent subset of V. Then it is a basis.

Proof:

If it does not span, take $\underline{e}_{n+1} \notin \langle \underline{e}_1, \ldots, \underline{e}_n \rangle$, and $\underline{e}_1, \ldots, \underline{e}_{n+1}$ is a basis, contradicting maximality.

Proposition

Suppose that $\underline{e}_1, \ldots, \underline{e}_k$ are linearly independent in a finite dimensional vector space V. Then, we can extend $\underline{e}_1, \ldots, \underline{e}_k$ to a basis $\underline{e}_1, \ldots, \underline{e}_n$ for V.

Proof:

Let $\underline{x}_1, \ldots, \underline{x}_m$ span V. Take $\underline{e}_1, \ldots, \underline{e}_n$ to be a maximal independent subset of $\underline{e}_1, \ldots, \underline{e}_k, \underline{x}_1, \ldots, \underline{x}_m$ including the $\underline{e}_1, \ldots, \underline{e}_k$. If some $\underline{x}_j \notin \langle \underline{e}_1, \ldots, \underline{e}_n \rangle$, we could add it to get an independent set, contradicting maximality. So $\langle \underline{x}_1, \ldots, \underline{x}_m \rangle = V \leq \langle \underline{e}_1, \ldots, \underline{e}_n \rangle$ and $\underline{e}_1, \ldots, \underline{e}_n$ is a basis.

Linear Algebra (5)

Alternatively, define $\underline{e_1}, \dots, \underline{e_r}$ inductively for $r \geq k$.
by <u>either</u> $\langle \underline{e_1}, \dots, \underline{e_r} \rangle = V$, and we have a basis $n = r$
<u>or</u> $\langle \underline{e_1}, \dots, \underline{e_r} \rangle \neq V$, and so $\underline{x_j} \notin \langle \underline{e_1}, \dots, \underline{e_r} \rangle$ and we let $\underline{e_{r+1}} = \underline{x_r}$, and $\underline{e_1}, \dots, \underline{e_{r+1}}$ are still linearly independent.

Proposition

Let $\underline{e_1}, \dots, \underline{e_n} \in V$, a vector space. Then $\underline{e_1}, \dots, \underline{e_n}$ is a basis $\Leftrightarrow$ every $\underline{x} \in V$ has a <u>unique</u> expression $\underline{x} = \sum x_i \underline{e_i}$ as a linear combination of the $\underline{e_i}$.

Proof:

($\Rightarrow$) $\underline{e_1}, \dots, \underline{e_n}$ span, so we can write $\underline{x} = \sum x_i \underline{e_i}$. If also $\underline{x} = \sum x_i' \underline{e_i}$ then $\sum (x_i - x_i') \underline{e_i} = 0$, so $x_i = x_i'$, by independence.

($\Leftarrow$) By definition, the $\underline{e_i}$ span. If $\sum \lambda_i \underline{e_i} = \underline{0}$, then $\sum \lambda_i \underline{e_i} = \sum 0 \underline{e_i}$ and then $\lambda_i = 0$ by uniqueness.

Consequence

$\underline{e_1}, \dots, \underline{e_n}$ is a basis for $V \Leftrightarrow (\underline{e_i} \neq 0)$, $V = \langle \underline{e_1} \rangle \oplus \dots \oplus \langle \underline{e_n} \rangle$

Observation

Suppose $\underline{e_1}, \dots, \underline{e_n}$ is a basis for V and $\underline{w_1}, \dots, \underline{w_n} \in W$. Then there is a unique linear map $\alpha: V \to W$ such that $\alpha(\underline{e_i}) = \underline{w_i}$. For we must define $\alpha(\underline{x}) = \alpha(\sum x_i \underline{e_i}) = \sum x_i \underline{w_i}$. This is all well defined by above, and linear, because

$$\alpha(\lambda \underline{x} + \mu \underline{y}) = \alpha(\sum (\lambda x_i + \mu y_i) \underline{e_i}) = \sum (\lambda x_i + \mu y_i) \alpha(\underline{e_i})$$
$$= \lambda \sum x_i \underline{w_i} + \mu \sum y_i \underline{w_i} = \lambda \alpha(\underline{x}) + \mu \alpha(\underline{y}).$$

Application

Take the standard basis and some $\underline{x}_1, \dots, \underline{x}_n \in V$. Then there is a unique linear map $\xi : F^n \to V$, $(a_1, \dots, a_n) \mapsto \sum a_i \underline{x}_i$

$\underline{x}_1, \dots, \underline{x}_n$ span $\iff \xi$ surjective

are independent $\iff \xi$ injective

are a basis $\iff \xi$ is an isomorphism.

2·3 Exchange Lemma

Let $\underline{e}_1, \dots, \underline{e}_n$ be linearly independent in V, and $\underline{x}_1, \dots, \underline{x}_m$ span V, a finite dimensional vector space. Then $n \leq m$, and reordering the $\underline{x}_i$ if necessary, we have $\underline{e}_1, \dots, \underline{e}_n, \underline{x}_{n+1}, \dots, \underline{x}_m$ spanning.

Proof:

Let $\underline{e}_1, \dots, \underline{e}_r$ be such that r is maximised with $r \leq m$, and $\underline{e}_1, \dots, \underline{e}_r, \underline{x}_{r+1}, \dots, \underline{x}_m$ spanning after reordering if necessary. Suppose $r < n$. Then $\underline{e}_{r+1} \in \langle \underline{e}_1, \dots, \underline{e}_r, \underline{x}_{r+1}, \dots, \underline{x}_m \rangle$, and so we can write $\underline{e}_{r+1} = \sum_{i=1}^{r} \lambda_i \underline{e}_i + \sum_{i=r+1}^{m} \mu_i \underline{x}_i$. If all $\mu_i = 0$, we contradict linear independence of the $\underline{e}_i$. So some $\mu_i \neq 0$, and reordering, we can assume $\mu_{r+1} \neq 0$.

Then $\underline{x}_{r+1} = \sum_{i=1}^{r} \left(-\frac{\lambda_i}{\mu_{r+1}}\right) \underline{e}_i + \frac{1}{\mu_{r+1}} \underline{e}_{r+1} + \sum_{i=r+2}^{m} \left(-\frac{\mu_i}{\mu_{r+1}}\right) \underline{x}_i$.

$\underline{x}_{r+1} \in \langle \underline{e}_1, \dots, \underline{e}_r, \underline{x}_{r+2}, \dots, \underline{x}_m \rangle$

But all the other elements of the spanning set $\underline{e}_1, \dots, \underline{e}_r, \underline{x}_{r+1}, \dots, \underline{x}_m$ are also in this space. So the space spanned is V, and $\underline{e}_1, \dots, \underline{e}_{r+1}, \underline{x}_{r+2}, \dots, \underline{x}_m$ span, contradicting maximality of r. So the maximal r is n, and we have $n \leq m$ and $\underline{e}_1, \dots, \underline{e}_n, \underline{x}_{n+1}, \dots, \underline{x}_m$ spanning.

Corollary

If $\underline{e}_1, \dots, \underline{e}_n$ and $\underline{f}_1, \dots, \underline{f}_m$ are bases for a vector space, then $m = n$.

Linear Algebra 6

Corollary to the exchange lemma

If $\underline{e_1}, \dots, \underline{e_n}$ and $\underline{f_1}, \dots, \underline{f_m}$ for V then $n = m$.

Definition

A vector space V has dimension $n \iff$ it was a basis $\underline{e_1}, \dots, \underline{e_n}$ of n elements. A finite dimensional vector space has a unique dimension written $\dim V$.

Another Consequence

If $\dim V = m$ and $\underline{e_1}, \dots, \underline{e_n}$ are independent, then $n \leq m$, either by the exchange lemma, or by extending the independent set to a basis using invariance of dimension.

Proposition

Suppose $U \leq V$, V finite dimensional. Then U is finite dimensional, and $\dim U \leq \dim V$. Moreover, if $\dim U = \dim V$, then $U = V$.

Proof

Any independent set in U is independent in V and so of size $\leq V$. Take the maximal such set to give a basis of U so that U is finite dimensional and $\dim U \leq \dim V$. Given a basis for U, we can extend it to a basis for V. If $\dim U = \dim V$, the extension is trivial, it is also a basis for V, and $U = V$.

Aside : $\xi : F^n \to V$, $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto \sum_{i=1}^{n} a_i \underline{x_i}$ is an isomorphism just when $\underline{x_1}, \dots, \underline{x_n}$ is a basis. So if $\dim V = n$, then $V \cong F^n$.

~~Proposition~~

~~Suppose $U \leq V$, finite dimensional. Then U is finite dimensional, and $\dim U \leq \dim V$~~

Proposition

Suppose $\underline{e}_1, \dots, \underline{e}_n \in V$ with dimension n. Then :

i) $\underline{e}_1, \dots, \underline{e}_n$ independent $\Rightarrow$ $\underline{e}_1, \dots, \underline{e}_n$ form a basis.

ii) $\underline{e}_1, \dots, \underline{e}_n$ span $\Rightarrow$ $\underline{e}_1, \dots, \underline{e}_n$ form a basis.

Proof:

i) We can extend $\underline{e}_1, \dots, \underline{e}_n$ to a basis, but it will be of size n, so $\underline{e}_1, \dots, \underline{e}_n$ is already a basis.

ii) Take a minimal spanning subset to get a basis. This will be of size n and so all of $\underline{e}_1, \dots, \underline{e}_n$.

Rank - Nullity Theorem

Definition

Let $\alpha : U \to V$ be linear between finite dimensional U, V.
We define the rank : $r(\alpha) = \dim \operatorname{Im}(\alpha)$
the nullity : $n(\alpha) = \dim \operatorname{Ker}(\alpha)$

(Aside : It is enough here, and henceforth, to take U finite dimensional).

Linear Algebra (6)

Theorem

~~Theorem~~ Let $\alpha : U \to V$ be linear with U (and V) finite dimensional. Then $r(\alpha) + n(\alpha) = \dim U$.

Proof

Take $\underline{e_1}, \dots, \underline{e_k}$, a basis for $\ker(\alpha)$, and extend to a basis $\underline{e_1}, \dots, \underline{e_n}$ for U. Then $n(\alpha) = k$ and $\dim U = n$. We claim that $\alpha(\underline{e_{k+1}}), \dots, \alpha(\underline{e_n})$ is a basis for $\mathrm{Im}(\alpha)$.

Spanning: Take $\underline{y} = \alpha(\underline{x}) \in \mathrm{Im}(\alpha)$ with $\underline{x} = \sum_{i=1}^{n} x_i \underline{e_i}$. Then $\underline{y} = \alpha(\underline{x}) = \sum_{i=1}^{n} x_i \alpha(\underline{e_i}) = \sum_{k+1}^{n} x_i \alpha(\underline{e_i})$. Thus $\alpha(\underline{e_{k+1}}), \dots, \alpha(\underline{e_n})$ span $\mathrm{Im}\,\alpha$.

Independence: Suppose $\sum_{k+1}^{n} \mu_i \alpha(\underline{e_i}) = \underline{0}$. Then $\alpha\left(\sum_{k+1}^{n} \mu_i \underline{e_i}\right) = \underline{0}$ so $\sum_{k+1}^{n} \mu_i \underline{e_i} \in \ker \alpha$ and we can write $\sum_{k+1}^{n} \mu_i \underline{e_i} = \sum_{1}^{k} \lambda_i \underline{e_i}$.

Then $\sum_{1}^{k} \lambda_i \underline{e_i} + \sum_{k+1}^{n} (-\mu_i) \underline{e_i} = \underline{0}$ and so $(\lambda_i = 0$ and$)$ $\mu_i = 0$ by independence of $\underline{e_1}, \dots, \underline{e_n}$. This shows independence of $\alpha(\underline{e_{k+1}}), \dots, \alpha(\underline{e_n})$.

It follows that $r(\alpha) = n - k$, and so $r(\alpha) + n(\alpha) = \dim U$.

Another Version

Suppose $W \leq U$, finite dimensional. Take $\underline{e_1}, \dots, \underline{e_k}$, a basis for W, and extend to $\underline{e_1}, \dots, \underline{e_n}$, a basis for U. Then $(\underline{e_{k+1}} + W), \dots, (\underline{e_n} + W)$ is a basis for U/W. This is essentially the rank-nullity theorem for $U \to U/W$.

Remark

1. $0: U \to V$ is the only map with rank 0.

2. If $\lambda \neq 0$, then $r(\lambda\alpha) = r(\alpha)$, and $n(\lambda\alpha) = n(\alpha)$.

3. $r(\alpha+\beta)$ is not determined by $r(\alpha)$, $r(\beta)$. For $r(\alpha + -\alpha) = r(0)$, $r(\alpha+\alpha) = r(\alpha)$. But $\mathrm{Im}(\alpha+\beta) \leq \mathrm{Im}(\alpha) + \mathrm{Im}(\beta)$ so it is true, that $r(\alpha+\beta) \leq r(\alpha) + r(\beta)$.

Theorem (Corollary)

Let $U, V \leq W$ finite dimensional. Then $\dim(U+V) + \dim(U \cap V) = \dim(U) + \dim(V)$.

Proof:

Consider the map $\gamma: U \oplus V \to W$, $(\underline{u}, \underline{v}) \mapsto \underline{u} + \underline{v}$. (external)
$\mathrm{Im}(\gamma) = U+V$ and so $r(\gamma) = \dim(U+V)$.
$\mathrm{Ker}(\gamma) = \{(\underline{w}, -\underline{w}) \mid \underline{w} \in U \cap V\} \cong U \cap V$, so $\dim(U \cap V) = n(\gamma)$
$\dim(U \oplus V) = \dim U + \dim V$
(because for bases $\underline{e_1}, \dots, \underline{e_n}$, $\underline{f_1}, \dots, \underline{f_m}$ of U, V, $(\underline{e_1}, \underline{0}), \dots, (\underline{e_n}, \underline{0})$, $(\underline{0}, \underline{f_1}), \dots, (\underline{0}, \underline{f_m})$ is a basis for $U \oplus V$.

So by the Rank-Nullity Theorem, $\dim U + \dim V = \dim(U+V) + \dim(U \cap V)$

Usual Proof

Take $\underline{e_1}, \dots, \underline{e_k}$, a basis for $U \cap V$. Extend it to $\underline{e_1}, \dots, \underline{e_k}, \underline{f_1}, \dots, \underline{f_n}$ for U and $\underline{e_1}, \dots, \underline{e_k}, \underline{g_1}, \dots, \underline{g_m}$ for V.
Then claim that $\underline{e_1}, \dots, \underline{e_k}, \underline{f_1}, \dots, \underline{f_n}, \underline{g_1}, \dots, \underline{g_m}$ is a basis for $U+V$.

21/10/11

Linear Algebra (7)

2.5 Rank and Nullity of a Composite

<u>Proposition</u> Suppose $\alpha : U \to V$, $\beta : V \to W$ are linear maps of finite-dimensional vector spaces. Then:

a) $r(\beta\alpha) \leq r(\beta)$ and $r(\beta\alpha) \leq r(\alpha)$
$n(\beta\alpha) \geq n(\alpha)$ and $n(\beta\alpha) \geq n(\beta) + \dim U - \dim V$

b) $n(\beta\alpha) \leq n(\alpha) + n(\beta)$
$r(\beta\alpha) \geq r(\alpha) + r(\beta) - \dim V$

<u>Proofs</u>

a) $\mathrm{Im}\,\beta\alpha \leq \mathrm{Im}\,\beta \Rightarrow r(\beta\alpha) \leq r(\beta)$
$\ker \beta\alpha \geq \ker \alpha \Rightarrow n(\beta\alpha) \geq n(\alpha)$
From this we deduce $r(\beta\alpha) = \dim U - n(\beta\alpha) \leq \dim U - n(\alpha) = r(\alpha)$
And further $r(\beta\alpha) \leq r(\beta)$
$\Rightarrow \dim U - n(\beta\alpha) \leq \dim V - n(\beta)$ and so $n(\beta\alpha) \geq n(\beta) + \dim U - \dim V$

$\beta|_{\mathrm{Im}\alpha}$ b) To show $n(\beta\alpha) \leq n(\alpha) + n(\beta)$ it suffices to show that
$\dim U - r(\beta\alpha) \leq \dim U - r(\alpha) + n(\beta)$, that is:
$r(\alpha) - r(\beta\alpha) \leq n(\beta)$ $\quad (\dagger)$
Note that $\mathrm{Im}\,\beta\alpha = \mathrm{Im}(\beta|_{\mathrm{Im}\alpha})$. Now $r(\alpha) = \dim \mathrm{Im}(\alpha)$
$r(\beta\alpha) = r(\beta|_{\mathrm{Im}\alpha})$. So it suffices to show that $n(\beta|_{\mathrm{Im}\alpha}) \leq n(\beta)$
$\ker(\beta|_{\mathrm{Im}\alpha}) = \ker(\beta \cap \mathrm{Im}\,\alpha) \leq \ker(\beta)$ and this follows
Also, from $(\dagger)$ we get $r(\alpha) - r(\beta\alpha) \leq \dim V - r(\beta)$
$r(\beta\alpha) \geq r(\alpha) + r(\beta) - \dim V$

Chapter 3: Matrices

3.1 Coordinates

Let $\underline{e}_1, \ldots, \underline{e}_n$ be a basis for U. We saw that gives an isomorphism
$\hat{\varepsilon} : \mathbb{F}^n \to U$, $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto \sum_{i=1}^{n} a_i \underline{e}_i$
Let $\varepsilon : U \to \mathbb{F}^n$ be the inverse. So if $\underline{x} = \sum_i x_i \underline{e}_i$, then $\varepsilon(\underline{x}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
We say the $(x_i)_i$ are the coordinates of $\underline{x}$ with respect to $\underline{e}_1, \ldots, \underline{e}_n$
If we set $\varepsilon(\underline{x}) = \begin{pmatrix} \varepsilon_1(\underline{x}) \\ \vdots \\ \varepsilon_n(\underline{x}) \end{pmatrix}$ with $\varepsilon_j : U \to \mathbb{F}$ the coordinate functions,
then later we will identify the (ε_j) as the dual basis to $(\underline{e}_j)_j$ in U^*.
<u>Note</u> $\underline{x} = \sum_j \varepsilon_j(\underline{x}) \underline{e}_j$

3.2 Matrices

Let $\underline{e_1}, \dots, \underline{e_n}$ be a basis for U, giving $\varepsilon : U \xrightarrow{\cong} \mathbb{F}^n$ and $\underline{f_1}, \dots, \underline{f_m}$ be a basis for V, giving $\phi : V \xrightarrow{\cong} \mathbb{F}^m$. Suppose $\alpha : U \to V$ is linear. Then we have maps $\alpha, \varepsilon, \phi$ in the direction:

$$\begin{array}{ccc} U & \xrightarrow{\alpha} & V \\ \varepsilon \downarrow & & \downarrow \phi \\ \mathbb{F}^n & \xrightarrow{A} & \mathbb{F}^m \end{array}$$

and we let A be the matrix making the diagram commute i.e. multiplication by A is the map $\phi\alpha\hat{\varepsilon}$.

What does $A\varepsilon = \phi\alpha$ mean?

Take $\underline{x} \in U$ with $\underline{x} = \sum^n x_j \underline{e_j}$. Then $\varepsilon(\underline{x}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$A\varepsilon(\underline{x}) = \begin{pmatrix} \sum a_{1j} x_j \\ \vdots \\ \sum a_{mj} x_j \end{pmatrix}$ So $A\varepsilon(\underline{x}) = \phi\alpha(\underline{x})$ says that

$$\alpha(\underline{x}) = \sum_i \Big(\sum_j a_{ij} x_j \Big) \underline{f_i}$$

That is, if $(x_j)_j$ are the coordinates of $\underline{x}$, then $(\sum a_{ij} x_j)_i$ are the coordinates of $\alpha(\underline{x})$.

Set $\underline{x} = \underline{e_k}$ and we get $\alpha(\underline{e_k}) = \sum_i a_{ik} \underline{f_i}$

(Note: we know this in some sense as $\alpha(\underline{e_k})$ has coordinates the k^{th} column vector of A).

3.3 Change of basis

Suppose that $\underline{e_1}, \dots, \underline{e_n}$ and $\underline{e_1}', \dots, \underline{e_n}'$ are bases for U.

Let P be the matrix for the identity with respect to these two bases, i.e.

$$\begin{array}{ccc} U & \xrightarrow{I} & U \\ \varepsilon \downarrow & & \downarrow \varepsilon' \\ \mathbb{F}^n & \xrightarrow{P} & \mathbb{F}^n \end{array}$$

If $\underline{x} = \sum x_j \underline{e_j}$, then $\underline{x} = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} p_{ij} x_j \Big) \underline{e_i}'$. So P gives the new coordinates of $\underline{x}$ in terms of the old

On the other hand, $\underline{e_j} = \sum_{i=1}^{n} p_{ij} \underline{e_i}'$ so P gives the old basis vectors in terms of the new. Note that P is invertible. Write $P^{-1} = \hat{P}$.

Suppose $\underline{e_1}, \dots, \underline{e_n}, \underline{e_1}', \dots, \underline{e_n}'$ are bases for U with $\varepsilon, \varepsilon' : U \to \mathbb{F}^n$ and $\underline{f_1}, \dots \underline{f_m}, \underline{f_1}', \dots, \underline{f_m}'$ are bases for V with $\phi, \phi' : V \to \mathbb{F}^m$.

Suppose $\alpha : U \to V$ has matrix A with respect to $(\underline{e_j})_j$ and $(\underline{f_i})_i$.

What is the relation between A and A'?

21/10/11

Linear Algebra (7)

$$\begin{tikzcd}
U \arrow[rr, "\alpha"] \arrow[dr, "I"] \arrow[dd, "\varepsilon"'] & & V \arrow[dr, "I"] \arrow[dd] & \\
& U \arrow[rr, "\alpha"] \arrow[dd, "\varepsilon'"] & & V \arrow[dd] \\
\mathbb{F}^n \arrow[rr, "A"] \arrow[dr, "P"'] & & \mathbb{F}^m \arrow[dr, "Q"] & \\
& \mathbb{F}^n \arrow[rr, "A'"'] & & \mathbb{F}^m
\end{tikzcd}$$

We see that $\underline{A' = QAP^{-1}}$

24/10/11

Linear Algebra (8)

Recall that we had bases $\underline{e_1}, \dots, \underline{e_n}$ and $\underline{e_1}', \dots \underline{e_n}'$ for U.
Change of basis matrix $P = (p_{ij})$ with $\underline{e_k} = \sum_i p_{ik} \underline{e_i}'$ and so $\underline{e_k}' = \sum_i \hat{p}_{ik} \underline{e_i}$ where $\hat{P} = (\hat{p}_{ij}) = P^{-1}$
Also bases $\underline{f_1}, \dots, \underline{f_m}$, and $\underline{f_1}', \dots, \underline{f_m}'$ for V. Change of basis matrix $Q = (q_{ij})$ with $\underline{f_k} = \sum_i q_{ik} \underline{f_i}'$.

Suppose $\alpha : U \to V$ has matrix $A = (a_{ij})$ for the old basis and $A' = (a_{ij}')$ for the new basis.

$$\alpha(\underline{e_j}') = \alpha\left(\sum_r \hat{p}_{rj} \underline{e_r}\right) = \sum_r \hat{p}_{rj} \left(\sum_s a_{sr} \underline{f_s}\right)$$

$$= \sum_{r,s} a_{sr} \hat{p}_{rj} \left(\sum_i q_{is} \underline{f_i}'\right) = \sum_i \left(\sum_{r,s} q_{is} a_{sr} \hat{p}_{rj}\right) \underline{f_i}'$$

$$\Rightarrow a_{ij}' = \sum_{r,s} q_{is} a_{sr} \hat{p}_{rj} \Rightarrow A' = QA\hat{P} = QAP^{-1}$$

3.4 Row and column rank

Let A be an $m \times n$ matrix. Then we regard A as a map $A : \mathbb{F}^n \to \mathbb{F}^m$ $\quad r(A) = \dim \operatorname{Im} A$

$\therefore r(A) = \dim$ (space spanned by columns of A)
$\qquad = $ max number of linearly independant columns
$\qquad = \operatorname{colrnk}(A)$, the column rank of A
$\operatorname{rowrnk}(A) = \dim$ (space spanned by row vectors of A) $\leftarrow \dim \mathbb{F}^{m*}$
$\qquad = $ max number of independent rows

Note that $\operatorname{rowrank}(A) = r(A^T)$, $A^T : \mathbb{F}^m \to \mathbb{F}^n$

Proposition Let A be an $m \times n$ matrix. Then $r(A)$ is the least r such that there is a factorisation $A = CB$ with C an $n \times r$ matrix and B an $r \times n$ matrix

$$\begin{tikzcd} \mathbb{F}^n \arrow[r, "B"'] \arrow[rr, bend right, "A"'] & \mathbb{F}^r \arrow[r, "C"'] & \mathbb{F}^m \end{tikzcd}$$

Proof

First we show that in any such factorisation we have $r(A) \leq r$,
$r(A) = r(CB) \leq r(B), r(C)$ by 2.5
$r(B) = \dim \operatorname{Im} B \leq \dim \mathbb{F}^r = r$
$r(C) = \dim \mathbb{F}^r - n(C) \leq r$

Second we show that we can realise $r = r(A)$. For we have a factorisation of A as $\mathbb{F}^n \to \operatorname{Im} A \to \mathbb{F}^m$ and $\operatorname{Im}(A) \cong \mathbb{F}^{r(A)}$ and so a factorisation $\mathbb{F}^n \to \mathbb{F}^{r(A)} \to \mathbb{F}^m$

<u>Theorem</u> For any $m \times n$ matrix A, $\operatorname{colrnk}(A) = \operatorname{row\,rnk}(A)$

<u>Proof</u>

$A = CB \iff A^T = B^T C^T$

$$\begin{tikzcd} \mathbb{F}^n \arrow[rr, "A"] \arrow[dr, "B"'] & & \mathbb{F}^m \\ & \mathbb{F}^r \arrow[ur, "C"'] & \end{tikzcd} \iff \begin{tikzcd} \mathbb{F}^m \arrow[rr, "A^T"] \arrow[dr, "C^T"'] & & \mathbb{F}^n \\ & \mathbb{F}^r \arrow[ur, "B^T"'] & \end{tikzcd}$$

So the minimal r on the left hand side = minimal r on the right hand side

i.e. $r(A) = r(A^T)$, $\operatorname{colrnk}(A) = \operatorname{rowrnk}(A)$

3.5 Row Echelon Form

Given an $m \times n$ matrix A, the elementary row operations are the following:

(1) Transposing Rows

$A \mapsto TA$ where $T = \begin{pmatrix} 1 & & & & & \\ & 0 & & 1 & & \\ & & 1 & & & \\ & 1 & & 0 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}$

(i.e. interchanging rows !!)

(2) Scaling

$A \mapsto MA$ where $M = \begin{array}{cc} \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ k \\ \\ \\ \end{matrix} \end{array}$ $\lambda \neq 0$

(multiply the k^{th} row by λ)

(3) Adding a scalar multiple of one row to another

$A \mapsto LA$ $\quad L = \begin{pmatrix} 1 & & \lambda \\ & \ddots & \\ & & 1 \end{pmatrix}$ ij^{th} place

(adding $\lambda \times$ row j to row i)

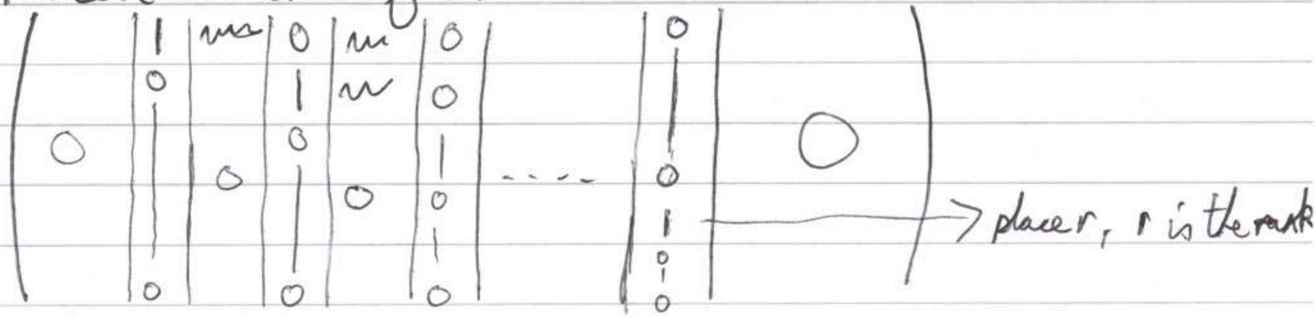
(1) The row operations and their corresponding matrices are invertible

(2) A combination of row operations is of the form $A \mapsto QA$ with Q invertible and so amounts to choosing a new basis for the image space.

24/10/11

Linear Algebra (7)

Using these transformations we transform a matrix A into a matrix B in reduced row echelon form.



1. Find the first non-zero column. Take a non 0 and transpose it to the top row. Scale that entry to 1; clear the rest of the column using (3)
2. Find the first column (following) with a non-zero entry not in the first row. Transpose that entry to the second row, scale the column and clear as before.
3. Repeat.

If B is in this form, then the non zero rows are independent and so span the row space, and the chosen column vectors are independent, and span the column space. So $\text{colrnk}(B) = \text{rowrnk}(B)$. Also, the process leaves the row space the same [So $\text{row rnk}(A) = \text{row rnk}(B)$. The process doesn't affect independence of columns so $\text{col rnk}(A) = \text{col rnk}(B)$]. $\dagger$
Clearly $\text{colrnk}(A) = \text{rowrnk}(A)$

(Abstractly $\dagger$ follows from $B = QA$, Q invertible
$\Rightarrow r(A) = r(B)$ $\quad B^T = A^T Q^T$, Q^T invertible $\Rightarrow r(A^T) = r(B^T)$)

26/10/11

Linear Algebra (9)

Last time we took an $m \times n$ matrix A and put it in row echelon form by $A \mapsto QA$

Remarks

1. If A is invertible then $QA = I$ and we get a method for finding the inverse of a matrix.
2. Abstractly, we have bases $\underline{e}_1, \dots, \underline{e}_n$ for $\mathbb{F}^n$ and $\underline{f}_1, \dots, \underline{f}_m$ for $\mathbb{F}^m$. We go through the list $\alpha(\underline{e}_1), \dots, \alpha(\underline{e}_n)$ and when we can remove an $\underline{f}_i$ and replace it with $\alpha(\underline{e}_j)$, we do so to get a basis $(\alpha(\underline{e}_{j(1)}), \dots, \alpha(\underline{e}_{j(r)}),$ remaining $\underline{f}_i)$ $\quad j(1) < j(2) < \dots < j(r)$

Chapter 4 Determinants

4.1 Endomorphisms $\to$ EM

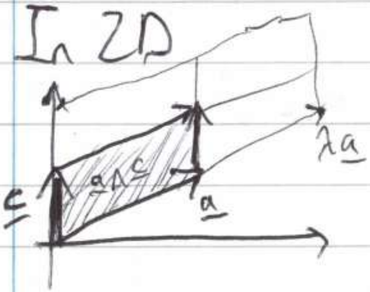
An endomorphism of a vector space V is a linear map $\alpha : V \to V$.
Given an EM α, of V, a finite dimensional vector space, its matrix with respect to a basis $\underline{e}_1, \dots, \underline{e}_n$ is its matrix as a linear map with respect to $(\underline{e}_j)$ in both domain and range i.e. it is the matrix $A = (a_{ij})$ such that
$$\alpha(\underline{e}_j) = \sum_{i=1}^{n} a_{ij} \underline{e}_i$$
So we have
$$\begin{array}{ccc} V & \xrightarrow{\alpha} & V \\ \varepsilon \downarrow & & \downarrow \varepsilon \\ \mathbb{F}^n & \xrightarrow{A} & \mathbb{F}^n \end{array}$$
commuting.

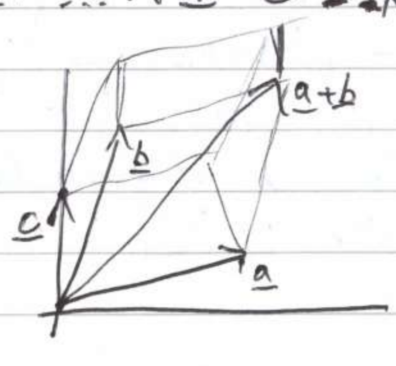
If $\underline{e}_j'$ is another basis with change of basis matrix $P = (p_{ij})$ so that $\underline{e}_j = \sum_{i=1}^{n} p_{ij} \underline{e}_i'$ then the matrix of A' of α with respect to $(\underline{e}_j')$ is given by $A' = PAP^{-1}$, a conjugate of A.

4.2 The determinant line

Suppose V is a vector space of dimension n. Then an n-tuple $\underline{a}_1, \dots, \underline{a}_n$ of vectors determines a volume element
$$V(\underline{a}_1, \dots, \underline{a}_n) = \underline{a}_1 \wedge \dots \wedge \underline{a}_n \in \Lambda_n(V)$$

In 2D





Volume is multilinear $V(-, \lambda \underline{a} + \mu \underline{b}, -) = \lambda V(-, \underline{a}, -) + \mu V(-, \underline{b}, -)$

and alternating $V(-, \underline{a}, -, \underline{a}, -) = 0$

<u>Lemma</u> $V(-, \underline{a}, -, \underline{b}, -) = -V(-, \underline{b}, -, \underline{a}, -)$

<u>Proof</u> $V(-, \underline{a}+\underline{b}, -, \underline{a}+\underline{b}, -) = 0$

$V(-, \underline{a}, -, \underline{a}, -) + V(-, \underline{a}, -, \underline{b}, -) + V(-, \underline{b}, -, \underline{a}, -) + V(-, \underline{b}, -, \underline{b}, -) = 0$

Hence the result.

Let $\underline{e_1}, \dots, \underline{e_n}$ be a basis for V and let $\underline{a_1}, \dots, \underline{a_n}$ be the image under some linear map α with matrix $A = (a_{ij}) \in M$ so that

$\underline{a_j} = \sum_{i=1}^{n} a_{ij} \underline{e_i}$

Then $V(\underline{a_1}, \dots, \underline{a_n}) = \sum_{f: \{1, \dots, n\} \circlearrowleft} \prod_j a_{f(j)j} V(\underline{e_{f(1)}}, \dots, \underline{e_{f(n)}})$ $= 0$ unless f is a bijection

$= \sum_{\sigma \in S_n} \prod_j a_{\sigma(j)j} V(\underline{e_{\sigma(1)}}, \dots, \underline{e_{\sigma(n)}})$

$= \left(\sum_{\sigma \in S_n} \prod_j a_{\sigma(j)j} \varepsilon(\sigma) \right) V(\underline{e_1}, \dots, \underline{e_n})$

So any $V(\underline{a_1}, \dots, \underline{a_n})$ is a scalar multiple of $V(\underline{e_1}, \dots, \underline{e_n})$ and so $\Lambda_n(V)$ is of dimension ≤ 1.

If $\dim \Lambda_n(V) = 1$ then any $\alpha: V \to V$ induces a linear map $\Lambda_n(\alpha): \Lambda_n(V) \to \Lambda_n(V)$ which is multiplication by a scalar, the determinant of α by definition. All properties follow easily.

But to show it has dimension 1 we must do something with the determinant.

<u>4.3 Volume Forms</u>

Let V be of dimension n. A volume form ω on V is a map $\omega: V^n \to \mathbb{F}$ which is:

multilinear $\omega(-, \lambda \underline{a} + \mu \underline{b}, -) = \lambda \omega(-, \underline{a}, -) + \mu \omega(-, \underline{b}, -)$

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

Note: a volume form ω induces a linear map $\Lambda_n(V) \to \mathbb{F}$

$\underline{a_1} \wedge \underline{a_2} \wedge \dots \wedge \underline{a_n} \mapsto \omega(\underline{a_1}, \dots, \underline{a_n})$

26/10/11

Linear Algebra (8)

As before, $\underline{e_1}, \ldots, \underline{e_n}$ are a basis for V and $\underline{a_j}$ defined by $\underline{a_j} = \sum_{i=1}^{n} a_{ij} \underline{e_i}$ then

$$\omega(\underline{a_1}, \ldots, \underline{a_n}) = \left[\sum_{\sigma \in S_n} \left(\prod_j a_{\sigma(j)j} \right) \varepsilon(\sigma) \right] \omega(\underline{e_1}, \ldots, \underline{e_n})$$

Now we set $\omega(\underline{e_1}, \ldots, \underline{e_n}) = 1$ and consider

$$(\underline{a_1}, \ldots, \underline{a_n}) \mapsto \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_j a_{\sigma(j)j}$$

- This is manifestly multilinear
- If $\underline{a_i} = \underline{a_j}$, $i \neq j$, then let $\tau = (i\ j)$, and for any σ

$$\prod_k a_{\sigma(k)k} = \prod_k a_{\tau\sigma(k)k}, \quad \varepsilon(\sigma) = -\varepsilon(\tau\sigma)$$ so in the sum, the terms cancel in pairs and $\omega = 0$. So we have a volume form.

The image of $v(\underline{e_1}, \ldots, \underline{e_n})$ under this is 1, so $\Lambda_n(V)$ does have dimension 1.

28/10/11 Linear Algebra 10

Last time we saw:

- Any volume form ω satisfies
$$\omega(\underline{a_1}, \ldots, \underline{a_n}) = \left(\sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_j a_{\sigma(j)j}\right) \omega(\underline{e_1}, \ldots, \underline{e_n})$$
where $\underline{a_j} = \sum a_{ij} \underline{e_i}$
- $\underline{a_1}, \ldots, \underline{a_n} \mapsto \left(\sum_{\sigma \in S_n} \varepsilon(\sigma) \prod a_{\sigma(j)j}\right)$ is a volume form
$$= \det(\underline{a_1}, \ldots, \underline{a_n}) = \det A = \det \alpha$$
with the property $\det(\underline{e_1}, \ldots, \underline{e_n}) = 1$ (so it is non trivial)

This shows that any volume form ω is a scalar multiple
$\omega(\underline{x_1}, \ldots, \underline{x_n}) = \omega(\underline{e_1}, \ldots, \underline{e_n}) \det(\underline{x_1}, \ldots, \underline{x_n})$ of the normalised form det.

Given γ, an endomorphism of V, set $\gamma(\underline{e_i}) = \underline{c_i}$

Then $\det(\gamma(\underline{x_1}), \ldots, \gamma(\underline{x_n}))$ is a volume form and we get
$$\underbrace{\det(\gamma(\underline{b_1}), \ldots, \gamma(\underline{b_n}))}_{\det \gamma\beta} = \underbrace{\det(\gamma(\underline{e_1}), \ldots, \gamma(\underline{e_n}))}_{= \det \gamma} \underbrace{\det(\underline{b_1}, \ldots, \underline{b_n})}_{\det \beta}$$
So

$\beta(\underline{e_i}) = \underline{b_i}$, $B = (\underline{b_1}, \ldots, \underline{b_n})$, $C = (\underline{c_1}, \ldots, \underline{c_n})$
$$\det CB = \det C \det B$$
In particular
$$\det(\alpha(\underline{d_1}), \ldots, \alpha(\underline{d_n})) = \det \alpha \det(\underline{d_1}, \ldots, \underline{d_n})$$
and so $\omega(\alpha(\underline{d_1}), \ldots, \alpha(\underline{d_n})) = \det \alpha \; \omega(\underline{d_1}, \ldots, \underline{d_n})$ for any volume form.

It follows that the definition of $\det \alpha$ is independant of the choice of $\underline{e_1}, \ldots, \underline{e_n}$.

4.4 The determinant of a matrix

Let A be an $n \times n$ matrix (and consider $A : \mathbb{F}^n \to \mathbb{F}^n$)

Write $\underline{a_j} = A\underline{e_j} = \sum_i a_{ij} \underline{e_i}$

Set $\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_j a_{\sigma(j)j}$ and regard it as a map
$\underline{a_1}, \ldots, \underline{a_n} \mapsto \det A = \det(\underline{a_1}, \ldots, \underline{a_n})$. of the n column vectors

1. $\det(\underline{a_1}, \ldots, \underline{a_n})$ is alternating and multilinear
2. $\det(\underline{e_1}, \ldots, \underline{e_n}) = \det \underline{I} = 1$
3. If ω is a volume form in $\mathbb{F}^n$ then
$$\omega(\underline{a_1}, \ldots, \underline{a_n}) = \omega(\underline{e_1}, \ldots, \underline{e_n}) \det(\underline{a_1}, \ldots, \underline{a_n})$$

If $\underline{a_r} = \underline{a_s}$, $r \neq S$, the $a_{\sigma(j)j}$ and $a_{\sigma(j)\tau(j)}$ are equal $\tau = (r\ S)$ $\prod_j a_{\sigma(j)j} = \prod_j a_{\tau\sigma(j)j}$

$$\omega(\underline{a_1}, \dots, \underline{a_n}) = \sum_{f:\{1,\dots,n\}\circlearrowleft} \prod_j a_{f(j)j}\, \omega(\underline{e_{f(1)}}, \dots, \underline{e_{f(n)}})$$

$$= \sum_{\sigma \in S_n} \prod a_{\sigma(j)j}\, \omega(\underline{e_{\sigma(1)}}, \dots, \underline{e_{\sigma(n)}})$$

$$= \sum_{\sigma \in S_n} \prod_j a_{\sigma(j)j}\, \varepsilon(\sigma)\, \omega(\underline{e_1}, \dots, \underline{e_n})$$

ASIDE: $(a_{11}\underline{e_1} + a_{21}\underline{e_2} + \dots) \wedge (a_{12}\underline{e_1} + a_{22}\underline{e_2} + \dots) \wedge (a_{13}\underline{e_1} + \dots\ )$
For this, multiplying out means choosing an entry from each bracket; that is given by the choice $1 \mapsto f(1)$, $n \mapsto f(n)$

Given a matrix C, the function
$(\underline{b_1}, \dots, \underline{b_n}) \mapsto \det(C\underline{b_1}, \dots, C\underline{b_n})$ $(= \det CB)$ is alternating multilinear. Therefore $\det(C\underline{b_1}, \dots, C\underline{b_n}) = \det(C\underline{e_1}, \dots, C\underline{e_n}) \det(\underline{b_1}, \dots, \underline{b_n})$
$\det CB = \det C \det B$

Now suppose α is an endomorphism of V with matrix A with respect to $\underline{e_1}, \dots, \underline{e_n}$ and A' with respect to $\underline{e_1}', \dots, \underline{e_n}'$
Then $\det A = \det A'$
For $\det A' = \det PAP^{-1} = \det P \det A \det P^{-1} = \det A$
because $\det P \det P^{-1} = \det I = 1$

<u>Definition</u>
The determinant of an endomorphism α of a finite dimensional vector space is $\det A$ where A is any matrix for α.

$\det A$ is alternating multilinear in the columns, so the elementary column operations have the effect:
Transposition : Multiply det by -1
Scaling a column by λ : " " " λ
Adding a multiple of column j to column $i \neq j$: no effect on det

28/10/11

Linear Algebra

Note that

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_j a_{\sigma(j) j} \overset{\text{setting } \rho = \sigma^{-1}}{=} \sum_{\rho \in S_n} \varepsilon(\rho) \prod_k a_{k \rho(k)} = \det(A^T)$$

So ~~the~~ $\det A$ is alternating multilinear in the rows of A with the same effects for the row operations.

4.5 The adjugate matrix

Expand $\det A = \det(\underline{a_1}, \sum_i a_{ij} \underline{e_i}, \dots, \underline{a_n})$ in the j^{th} column.

$$\det A = \sum_i a_{ij} \det(\underline{a_1}, \dots, \overset{i}{\underline{e_i}}, \dots, \underline{a_n})$$

$\leftarrow j^{th}$ column

$$\det \begin{pmatrix} a_{11} & 0 & a_{1n} \\ \vdots & 1 & \vdots \\ a_{m1} & 0 & a_{mn} \end{pmatrix}$$

by column operations we can clear the rest of the i^{th} row

$$\det \begin{pmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ 0 & \cdots & 1 & \cdots & 0 \\ a_{m1} & \cdots & 0 & \cdots & a_{nn} \end{pmatrix}$$

It takes $i+j-2$ transpositions to move the 1 to the top left.

So $\det(\underline{a_1}, \dots, \underline{e_i}, \dots, \underline{a_n}) = (-1)^{i+j} \det A_{\hat{i}\hat{j}} \Leftarrow A$ with i, j^{th} row, j^{th} column gone

Set $\operatorname{adj}(A)$ be the matrix $(\operatorname{adj}(A))_{ij} = (-1)^{i+j} \det A_{\hat{j}\hat{i}}$

We have $\det A = \sum_i (\operatorname{adj}(A))_{ji} a_{ij}$ (for each j)

Also $\sum_i (\operatorname{adj}(A)_{ji}) a_{ik}$, $k \neq j$

$= \det$ of A with j^{th} column replaced by k^{th}

$= 0$

Theorem

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

In particular if A is invertible, then $A^{-1} = \frac{1}{\det A} \operatorname{adj}(A)$

Linear Algebra (11)

A few more things

1. Suppose an $n\times n$ matrix is of the form
$A = \left(\begin{array}{c|c} B & C \\ \hline O & D \end{array}\right)$ with $B = m\times m$, $D = (n-m)\times(n-m)$, $C = m\times(n-m)$.
Then $\det A = \det B \det D$.

EITHER
Consider that $\prod_i a_{\sigma(i)i} = 0$ unless σ permutes $\{1,\dots,m\}$ and $\{m+1,\dots,n\}$ amongst themselves. Then all possible choices
$\left[\sum_\tau \prod_i b_{\tau(i)i}\right]\left(\sum_\rho \prod_i d_{\rho(i)i}\right)$

OR $A = \left(\begin{array}{c|c} I & O \\ \hline O & D \end{array}\right)\left(\begin{array}{c|c} I & C \\ \hline O & I \end{array}\right)\left(\begin{array}{c|c} B & O \\ \hline O & I \end{array}\right)$ so $\det A = \det D \cdot 1 \cdot \det$

2. Suppose α is an endomorphism of V (finite dimensional) or Suppose A is an $n\times n$ matrix. Then:
$\overset{\alpha}{A}$ is invertible iff $\overset{\det\alpha}{\det A} \neq 0$
($\Rightarrow$) If $\hat{A} = A^{-1}$ exists, then $\hat{A}A = I$, and so $\det\hat{A}\det A = 1$, and $\det A \neq 0$
($\Leftarrow$) We don't <u>need</u> to use the formula for A^{-1} $\left(= \frac{1}{\det A}\operatorname{adj}(A)\right)$
For if A is not invertible, then the columns of A are dependent. Suppose e.g. that the first column is dependent on the others, $\underline{a_1} = \sum_{i=2}^{n} \mu_i \underline{a_i}$. Then $\det A = \det(\underline{a_1},\dots,\underline{a_n}) = \det(\underline{a_1} - \sum_i \mu_i a_i, \dots\dots)$ using column operation 3 $= \det(\underline{0},\dots)$ which is 0 by linearity in the first factor.

3. The trace of a matrix $\overset{\nearrow n\times n}{A} = (a_{ij})$ is the sum $tr(A) = \sum_i a_{ii}$ of the diagonal elements. Suppose A is $m\times n$ and B is $n\times m$, so BA is $n\times n$ and AB is $m\times m$.
Then $tr(AB) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}b_{ji}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} b_{ji}a_{ij}\right) = tr(BA)$
<u>Consequences</u> If A and A' are matrices of an endomorphism α, then $tr(A') = tr((PA)P^{-1}) = tr(P^{-1}PA) = tr A$
So we can define the trace $tr\,\alpha$ of an endomorphism α of a finite dimensional vector space to be $tr A$, where A is any matrix for α

Chapter 5 Theory of an endomorphism

5.1 Invariant subspaces

Definition Suppose α is an endomorphism of V. A subspace $U \leq V$ is invariant (with respect to α) just when $\alpha|_U : U \to U$ i.e. $\mathrm{Im}(\alpha|_U) \leq U$.

Suppose α an endomorphism of V, finite dimensional, and $V = U_1 \oplus \dots \oplus U_k$ where the U_i are invariant.
Take successively bases for the U_i to give a basis for V.
With respect to this basis, α has a matrix of the form

$$\begin{pmatrix} \boxed{B_1} & & & 0 \\ & \boxed{B_2} & & \\ & & \ddots & \\ 0 & & & \boxed{B_k} \end{pmatrix}$$

where B_i is the matrix for $\alpha|_{U_i}$ with respect to the chosen basis for U_i.
This is a handle on α.

Suppose $U \leq V$ is α invariant for α an endomorphism of V (finite dimensional). Take a basis $\underline{e}_1, \dots, \underline{e}_m$ for U ($m = \dim U$) and extend to a basis $\underline{e}_1, \dots, \underline{e}_n$ for V ($\dim V = n$). With respect to this basis α has a matrix of the form $A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$ where B is the matrix for $\alpha|_U$ with respect to $\underline{e}_1, \dots, \underline{e}_m$

[What is D? α induces a map $\bar{\alpha} : V/U \to V/U$
$\bar{\alpha}(\underline{x} + U) = \alpha(\underline{x}) + U$. D is the matrix for $\bar{\alpha}$ with respect to the basis $\underline{e}_{m+1} + U, \dots, \underline{e}_n + U$ for V/U]

Cyclic Subspaces Let α be an endomorphism of V. For any $\underline{x} \in V$ there is a least invariant subspace $\langle \underline{x} \rangle_\alpha$ containing $\underline{x}$.
Why? If $\underline{x} \in U$ invariant, then $\alpha(\underline{x}) \in U$, $\alpha^2(\underline{x}) \in U$, etc.
So $\langle \underline{x}, \alpha(\underline{x}), \alpha^2(\underline{x}), \dots \rangle \leq U$. But this is also invariant. So $\langle \underline{x} \rangle_\alpha = \langle \underline{x}, \alpha(\underline{x}), \alpha^2(\underline{x}), \dots \rangle$. Now suppose that V is finite dimensional. Then $\underline{x}, \alpha(\underline{x}), \dots$ is dependent and we can take r least such that $\alpha^r(\underline{x}) \in \langle \underline{x}, \alpha(\underline{x}), \dots, \alpha^{r-1}(\underline{x}) \rangle$. By choice $\underline{x}, \dots, \alpha^{r-1}(\underline{x})$ are independent, and we have

$$\alpha^r(\underline{x}) = \sum_{i=0}^{r-1} \mu_i \alpha^i(\underline{x})$$

, so this space is invariant.

So $\langle \underline{x} \rangle_\alpha = \langle \underline{x}, \alpha(\underline{x}), \dots, \alpha^{r-1}(\underline{x}) \rangle$

Linear Algebra 11

Moreover, with respect to $\underline{x}, \ldots, \alpha^{r-1}(\underline{x})$, $\alpha|_{\langle \underline{x} \rangle_\alpha}$ has the matrix

$$B = \begin{pmatrix} 0 & 0 & & & \mu_0 \\ 1 & 0 & & & \vdots \\ 0 & 1 & 0 & & \vdots \\ \vdots & \vdots & 1 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \mu_{r-1} \end{pmatrix}$$

$\boxed{\text{Compute } \det(B - \lambda I)}$

02/11/11

Linear Algebra (12)

5.2 Eigenvectors and Eigenvalues

Definition

Let α be an endomorphism of V. Then $\underline{x} \in V$ is an *eigenvector* of α with *eigenvalue* λ just when $\alpha(\underline{x}) = \lambda \underline{x}$ and $\underline{x} \neq \underline{0}$

Remarks

1. $\underline{x}$ is an eigenvector just when $\langle \underline{x} \rangle_\alpha$ is 1 dimensional.
2. The eigenvalue for an eigenvector is uniquely determined. However $\lambda I : V \to V$ has all non zero vectors as eigenvectors with eigenvalue 1.
3. Over $\mathbb{R}$: $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$ rotates through $90°$ and there are no eigenvectors or eigenvalues.

However, over $\mathbb{C}$, we shall see that eigenvalues exist for any non trivial V (i.e. $V \neq \{\underline{0}\}$)

If λ is an eigenvalue, then the eigenspace for λ is $\{\underline{x} \mid \alpha(\underline{x}) = \lambda \underline{x}\} = \ker(\alpha - \lambda I) \leq V$. It consists of all eigenvectors together with $\underline{0}$. Assume now V is finite dimensional.

Observation

λ is an eigenvalue iff $\ker(\alpha - \lambda I) \neq \{\underline{0}\}$

iff $(\alpha - \lambda I)$ is not invertible (is singular)

iff $\det(\alpha - \lambda I) = 0$

We want to set $\chi_\alpha(t) = \det(\alpha - tI)$, the characteristic polynomial of α. Then λ is an eigenvalue iff $\chi_\alpha(\lambda) = 0$. Then the existence of eigenvalues follows from the Fundamental Theorem of Algebra, because $\chi_\alpha(t)$ is of degree $n = \dim V$, and so if $n \geq 1$, $\chi_\alpha(t)$ has roots over $\mathbb{C}$.

We define the characteristic polynomial $\chi_A(t)$ of an $n \times n$ matrix by $\chi_A(t) = \det(A - tI)$. Note that

$$A - tI = \begin{pmatrix} a_{11}-t & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn}-t \end{pmatrix}$$

has entries in $F[t]$, the ring of polynomials in t.

Using the formula we get $\chi_A(t) \in F[t]$. It's clear that the degree of $\chi_A(t)$ is n.

If α ~~and~~ is an endomorphism of V with matrices A, A' with respect to bases $(\underline{e_i})_i$ and $(\underline{e_i'})_i$ then $A = PAP^{-1}$ and so $(A' - tI) = P(A - tI)P^{-1}$

So $\chi_{A'}(t) = \det(A' - tI) = \det(A - tI) = \chi_A(t)$. It follows that we can define $\chi_\alpha(t)$ to be $\chi_A(t)$ where A is any matrix for α. Substituting λ for t we see that $\chi_\alpha(\lambda) = \det(\alpha - \lambda I)$. So the roots of the characteristic polynomial are the eigenvalues.

5.3 The Cayley - Hamilton Theorem

If α is an endomorphism of V then we can define α^n e.g. by setting $\alpha^0 = I$ and $\alpha^{n+1} = \alpha \cdot \alpha^n$. Then if $p(t) = a_0 + a_1 t + \ldots + a_n t^n \in F[t]$ is a polynomial, then we can define $p(\alpha) = a_0 I + a_1 \alpha + \ldots + a_n \alpha^n$

(N.B. If $p(t)$ and $q(t) \in F[t]$, we do have $p(\alpha)\, q(\alpha) = q(\alpha)\, p(\alpha)$)

Theorem

If α is an endomorphism of a finite dimensional V, then $\chi_\alpha(\alpha) = 0$.

Meaning in terms of matrices Suppose $A = (a_{ij})$ is an $n \times n$ matrix. Then

$$A - tI = \begin{pmatrix} a_{11} - t & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - t \end{pmatrix}$$ ~~and so~~

and so $\chi_A(t)$ is the determinant of this.

So $$\chi_A(A) = \det\begin{pmatrix} a_{11}I - A & \cdots & a_{1n} I \\ \vdots & \ddots & \vdots \\ a_{n1} I & \cdots & a_{nn} I - A \end{pmatrix}$$ and it is this that must be 0.

(†)

2/11/11

Linear Algebra (12)

This shows that the following is nonsense:

$\chi_\alpha(t) = \det(\alpha - tI)$, so setting $t = \alpha$

$\chi_\alpha(\alpha) = \det(\alpha - \alpha) = \det 0 = 0$

SECRET

Take a matrix for α with respect to $\underline{e}_1, \ldots, \underline{e}_n$.

Take the transpose of $(\dagger)$ and apply it to the column vector of basis vectors.

$$\begin{pmatrix} a_{11}I - A & \cdots & a_{n1}I \\ \vdots & \ddots & \vdots \\ a_{1n}I & \cdots & a_{nn}I - A \end{pmatrix}\begin{pmatrix} \underline{e}_1 \\ \underline{e}_2 \\ \vdots \\ \underline{e}_n \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \vdots \\ \underline{0} \end{pmatrix}$$

So $(\dagger)$ is singular and $\det(\dagger) = 0$

Proof by cyclic subspaces

Let $\underline{x} \in V$ be $\neq \underline{0}$ and consider $\langle \underline{x} \rangle_\alpha = \langle \underline{x}, \ldots, \alpha^{r-1}(\underline{x}) \rangle$

where $\alpha^r(\underline{x}) = \sum_0^{r-1} \mu_i \alpha^i(\underline{x})$

Let β be the restriction of α to $\langle \underline{x} \rangle_\alpha$. With respect to the basis $\underline{x}, \ldots, \alpha^{r-1}(\underline{x})$,

β has matrix
$$\begin{pmatrix} 0 & & 0 & \mu_1 \\ 1 & & & \mu_2 \\ 0 & \ddots & & \vdots \\ \vdots & & & \vdots \\ 0 & \cdots & 0\ 1 & \mu_{r-1} \end{pmatrix}$$

$\chi_\beta(t) = \det(\beta - tI) = (-1)^{r+1}(\mu_0 + \mu_1 t + \ldots + \mu_{r-1}t^{r-1} - t^r)$

expanding the last column.

$= (-1)^r(t^r + a_{r-1}t^{r-1} + \ldots + a_0)$ where $a_i = -\mu_i$ $(*)$

By $(*)$, $\chi_\beta(\alpha)(\underline{x}) = \underline{0}$ and so $\chi_\beta(\alpha)(\alpha^i(\underline{x})) = \alpha^i \chi_\beta(\alpha)(\underline{x}) =$

So $\chi_\beta(\alpha)$ is 0 on $\langle \underline{x} \rangle_\alpha$.

[N.B $\chi_\beta(\beta) = 0$]

Extend the basis $\underline{x}, \ldots, \alpha^{r-1}(\underline{x})$ for $\langle \underline{x} \rangle_\alpha$ to a basis for V

With respect to that matrix α has the matrix

$\left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$ and it follows that $\chi_\alpha(t) = \chi_A(t) = \det(A - tI)$

$= \det(B - tI)\det(D - tI)$

$= \chi_\beta(t)\det(D - tI)$

Hence as $\chi_\beta(\alpha) = 0$ on $\langle \underline{x} \rangle_\alpha$ we have $\chi_\alpha(\alpha) = 0$ on $\langle \underline{x}$

But $\underline{x} \neq \underline{0}$ is arbitrary so in particular we have

$$\chi_a(a)(\underline{x}) = 0 \quad \forall \underline{x} \in V$$

Thus $\chi_a(a) = 0$

04/11/11 Linear Algebra 13

5.4 Triangular Form

Let A be the matrix for an endomorphism α of V with respect to the basis $\underline{e_1}, \ldots, \underline{e_n}$. A is in upper triangular form

i.e. $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix}$ i.e. $a_{ij} = 0$ for $i > j$.

iff $\alpha(\underline{e_j}) \in \langle \underline{e_1}, \ldots, \underline{e_j} \rangle = V_j \ \forall j$

Suppose A, in upper triangular form, is the matrix for α with respect to $\underline{e_1}, \ldots, \underline{e_n}$.

$(\alpha - a_{nn}I) : V = V_n \to V_{n-1}$

$(\alpha - a_{n-1n-1}I) \ V_{n-1} \to V_{n-2}$

$\vdots$

$(\alpha - \alpha_{11}I) : V_1 \to V_0 = \{\underline{0}\}$

So $\prod (\alpha - a_{ii}I) = 0$. But also $\det(A - tI) = (-1)^n \prod_i (t - a_{ii})$

So the Cayley-Hamilton Theorem is clear for upper-triangular matrix

__Proposition__ Let α be an endomorphism of a (non-trivial) finite dimensional vector space V over $\mathbb{C}$. Then there is a basis $\underline{e_1}, \ldots, \underline{e_n}$ with respect to which the matrix of α is triangular.

__Proof__ Induction on $\dim V$ (assume ≥ 1)

Over $\mathbb{C}$ eigenvalues and so eigenvectors exist. So choose $\underline{e_1}$ to be an eigenvector for α and complete it to a basis $\underline{e_1}, \underline{f_2}, \ldots, \underline{f_n}$ for V.

The matrix of α is

$$\left(\begin{array}{c|c} a_{11} & C \\ \hline 0 & D \end{array}\right)$$

where D is the matrix for $\delta : \langle \underline{f_2}, \ldots, \underline{f_n} \rangle \circlearrowleft$

C is the matrix for $\gamma : \langle \underline{f_2}, \ldots, \underline{f_n} \rangle \mapsto \langle \underline{e_1} \rangle$

with respect to the given bases

On $\langle \underline{f_2}, \ldots, \underline{f_n} \rangle$, $\alpha(\underline{x}) = \gamma(\underline{x}) + \delta(\underline{x})$

By the induction hypothesis there is a basis $\underline{e_2}, \ldots, \underline{e_n}$ for $\langle \underline{f_2}, \ldots, \underline{f_n} \rangle$ with respect to which δ is upper triangular i.e.

$\delta(\underline{e_j}) \in \langle \underline{e_2}, \ldots, \underline{e_j} \rangle$, $j \geq 2$

Then $\alpha(\underline{e_j}) = \gamma(\underline{e_j}) + \delta(\underline{e_j}) \in \langle \underline{e_1}, \ldots, \underline{e_j} \rangle \quad j \geq 2$

$\alpha(\underline{e_1}) = a_{11}\underline{e_1} \in \langle \underline{e_1} \rangle$. So the matrix for α is upper triangular with respect to the basis $\underline{e_1}, \ldots, \underline{e_n}$

5.5 Diagonal form

Let A be the matrix for an endomorphism α with respect to $\underline{e_1}, \dots, \underline{e_n}$

A is diagonal $\begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$ i.e. $a_{ij} = 0$ for $i \neq j$

iff $\alpha(\underline{e_i}) = a_{ii}\underline{e_i}$, that is iff $\underline{e_1}, \dots, \underline{e_n}$ is a basis of eigenvectors.

Usually we collect the vectors for a given eigenvalue together in the list, so that the diagonal matrix looks like

$$\begin{pmatrix} \boxed{\overbrace{\lambda_1 \ddots \lambda_1}^{n_1}} & & & \\ & \boxed{\overbrace{\lambda_2 \ddots \lambda_2}^{n_2}} & & \\ & & \ddots & \\ & & & \boxed{\overbrace{\lambda_k \ddots \lambda_k}^{n_k}} \end{pmatrix}$$

with $\lambda_1, \dots, \lambda_k$ the distinct eigenvalues and $n_1 + \dots + n_k = n = \dim V$

The first n_1 vectors span $k(\alpha - \lambda_1 I)$ and so on.

$(V = V_1 \oplus \dots \oplus V_r$, if $\underline{x_i} \in V_i$, $\alpha(\underline{x_i}) = \lambda_i \underline{x_i})$

Note

1. An eigenspace $\ker(\alpha - \lambda I)$ is α invariant. For if $(\alpha - \lambda I)(\underline{x}) = \underline{0}$ then $(\alpha - \lambda I)(\alpha\underline{x}) = \alpha(\alpha - \lambda I)\underline{x} = \alpha(\underline{0}) = \underline{0}$
2. If $\underline{x} \in \ker(\alpha - \lambda I)$ then $\alpha(\underline{x}) = \lambda\underline{x}$ and so $\alpha^n(\underline{x}) = \lambda^n(\underline{x})$ and so for any polynomial $p(\alpha)(\underline{x}) = p(\lambda)\underline{x}$

Proposition

Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues for an endomorphism α of $\mathbf{V}$. Then the ~~form~~ sum of the corresponding eigenspaces is a direct sum:

$$\ker(\alpha - \lambda_1 I) \oplus \dots \oplus \ker(\alpha - \lambda_k I)$$

Proof

It suffices to show that no non-trivial sum $\underline{x_1} + \dots + \underline{x_k}$ for $\underline{x_i} \in \ker(\alpha - \lambda_i I)$, can be $\underline{0}$. So assume $\underline{x_1} + \dots + \underline{x_k} = \underline{0}$ with $x_i \in \ker(\alpha - \lambda_i I)$

Apply $\prod_{i=2}^{k}(\alpha - \lambda_i I)$. We get:

$$\prod_{i=2}^{k}(\lambda_1 - \lambda_i)\underline{x_1} = \underline{0}$$

. Thus $\underline{x_1} = \underline{0}$

Similarly, all $\underline{x_i}$ are $\underline{0}$

04/11/11 Linear Algebra 13

Corollary

An endomorphism α of a finite dimensional vector space V is diagonalisable iff $\ker(\alpha - \lambda_1 I) \oplus \dots \oplus \ker(\alpha - \lambda_k I) = V$

with λ_i distinct eigenvalues

iff $n_1 + \dots + n_k = n$, where $n_i = n(\alpha - \lambda_i I)$, $n = \dim V$

Corollary

Suppose α an endomorphism of V such that $\chi_\alpha(t)$ splits into distinct linear factors. Then α is diagonalisable.

For there are n eigenspaces $\ker(\alpha - \lambda_1 I) \dots \ker(\alpha - \lambda_n I)$ eac of dimension at least 1 and so exactly 1.

Linear Algebra (14)

5.6 The minimal polynomial

Let α be an EM of a finite dimensional V. For if $\dim V = n$, then $\dim L(V,V) = n^2$, and then $I = \alpha^0, \alpha, \ldots, \alpha^{n^2}$ is n^2+1 elements and so dependent, i.e. we have $\sum_0^{n^2} \lambda_i \alpha^i = 0$ with λ_i not all 0.

So that gives a polynomial p of degree $\leq n^2$ with $p(\alpha) = 0$

(Cayley - Hamilton : there is such a polynomial of degree n)

Take a monic polynomial m of least degree such that $m(\alpha) = 0$. (Don't allow 0). Assume that V is non-trivial, and $\deg m \geq 1$.

Proposition If p is a polynomial with $p(\alpha) = 0$ then $m(t) \mid p(t)$

Proof Write $p(t) = q(t)m(t) + r(t)$ where r is either 0 or of degree less than m.

Then $0 = p(\alpha) = q(\alpha)m(\alpha) + r(\alpha) = r(\alpha)$. So if $r \neq 0$ we can scale to make it monic contradicting minimality.

So $r = 0$ and $m \mid p$.

Corollary The m given above is unique. For if m_1, m_2 were two such, then $m_1 \mid m_2$, $m_2 \mid m_1$ and they are both monic and so equal

The minimal polynomial for α, $m_\alpha(t)$ is this unique $m(t)$, monic of least degree.

Cayley Hamilton $m_\alpha \mid \chi_\alpha$

Corollary Suppose $U \leq V$ is α invariant. Then $m_{\alpha|U} \mid m_\alpha$

Suppose α is an endomorphism of V, $\underline{x} \in V$ with $\langle \underline{x} \rangle_\alpha = V$, so that we have a basis $\underline{x}, \alpha(\underline{x}), \ldots, \alpha^{n-1}(\underline{x})$ for V, and with $\alpha^n(\underline{x}) = \sum_0^{n-1} \mu_i \alpha^i(\underline{x})$.

Then $m(t) \overset{\circ}{=} t^n + a_{n-1}t^{n-1} + \ldots + a_0$ with $a_i = -\mu_i$ satisfies $m(\alpha)(\underline{x}) = \underline{0}$ and so $m(\alpha)(\alpha^i(\underline{x})) = \alpha^i m(\alpha)(\underline{x}) = \underline{0}$ and so $m(\alpha) = 0$. But suppose $p(t) = t^r + b_{r-1}t^{r-1} + \ldots + b_0$ is a polynomial $r < n$. Then $p(\alpha)\underline{x} = \alpha^r(\underline{x}) + b_{r-1}\alpha^{r-1}(\underline{x}) + \ldots + b_0\underline{x}$ is a linear combination of basis vectors and so is non zero. That shows that $m(t)$ is the minimum polynomial $m_\alpha(t)$.

Recall that we showed $\chi_\alpha(t) = (-1)^n m_\alpha(t)$
We could argue that $m_\alpha \mid \chi_\alpha$ and both are of degree $n = \dim V$
χ_α is a scalar multiple of m_α. This shows that any monic polynomial is the minimum polynomial and $(-1)^n \times$ characteristic polynomial for some α.
(†) is such for $\begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & 0 & \vdots \\ & \ddots & & \\ 0 & & 1 \quad 0 & -a_{n-1} \end{pmatrix} : \mathbb{F}^n \to \mathbb{F}^n$

<u>Proposition</u>
The roots of $m_\alpha(t)$ are exactly the eigenvalues of α.
(<u>Aside</u> Since $m_\alpha \mid \chi_\alpha$ any root of m_α is a root of χ_α and so an eigenvalue)

<u>Proof</u>
Let λ be a root of $m_\alpha(t)$ and write $m_\alpha(t) = (t-\lambda)q(t)$
By minimality $q(\alpha) \neq 0$ so we can choose $\underline{x} \in V$ such that $\underline{y} = q(\alpha)(\underline{x}) \neq \underline{0}$. But $(\alpha - \lambda I)\underline{y} = (\alpha - \lambda I)q(\alpha)\underline{x} = m_\alpha(\alpha)\underline{x} = \underline{0}$
So $\underline{y}$ is an eigenvector with eigenvalue λ.
Now suppose λ is an eigenvalue with eigenvector $\underline{x} \neq \underline{0}$.
$\underline{0} = m_\alpha(\alpha)(\underline{x}) = m_\alpha(\lambda)\underline{x}$ with $\underline{x} \neq 0$, so $m_\alpha(\lambda) = 0$ and λ is a root of m_α.
(Or the minimum polynomial of $\alpha|_{\langle \underline{x} \rangle}$ is $(t-\lambda)$ and so $(t-\alpha) \mid m_\alpha(t)$)

Suppose α is diagonalisable so that $V = \ker(\alpha - \lambda_1 I) \oplus \dots \oplus \ker(\alpha - \lambda_k I)$ with λ_i the distinct eigenvalues. Clearly the minimal polynomial of $\alpha|_{\ker(\alpha-\lambda_i I)}$ is $(t-\lambda_i)$ so $(t-\lambda_1)\dots(t-\lambda_k) \mid m_\alpha(t)$
But $(\alpha - \lambda_1 I) \dots (\alpha - \lambda_k I)$ is 0 on each $\ker(\alpha - \lambda_i I)$ and so $m_\alpha(t) = (t-\lambda_1)\dots(t-\lambda_k)$. Here $\chi_\alpha(t) = (-1)^n (t-\lambda_1)^{n_1} \dots (t-\lambda_k)^{n_k}$ with $n_i = n((\alpha - \lambda I))$

<u>Proposition</u> Suppose α is an endomorphism of finite dimensional V whose minimal polynomial splits as a product
$m_\alpha(t) = (t-\lambda_1)\dots(t-\lambda_k)$ of distinct linear factors. Then α is ~~diag~~ diagonalisable.

Linear Algebra (4)

Proof

It suffices to show that the direct sum $\ker(\alpha-\lambda_1 I)\oplus\ldots\oplus\ker(\alpha-\lambda_k I)$ is all of V.

Either by dimension: The dimension of the direct sum

!! $n_1+\ldots+n_k = n(\alpha-\lambda_1 I)+\ldots+n(\alpha-\lambda_k I) \geq n\left(\prod(\alpha-\lambda_i I)\right)$
$= n(m_\alpha(\alpha)) = n(0) = \dim V = n$

Then $V=\ker(\alpha-\lambda_1 I)\oplus\ldots\oplus\ker(\alpha-\lambda_k I)$

Or: By the Chinese Remainder Theorem. ~~Let~~

Let $q_i(t) = \frac{m_\alpha(t)}{t-\lambda_i}$. Then the highest common factor of the $q_i(t)$ is 1.

And so we can write $1 = \sum_1^k a_i(t)\, q_i(t)$

For any $x \in V$ we have

$$x = \sum_1^k a_i(\alpha)\, q_i(\alpha)(x)$$

but $(\alpha-\lambda_i I)\, q_i(\alpha)(x) = m_\alpha(\alpha)(x) = 0$ so $q_i(\alpha)(x) \in \ker(\alpha-\lambda_i I)$

So $a_i(\alpha)\, q_i(\alpha)(x) \in \ker(\alpha-\lambda_i I)$. So $x \in \ker(\alpha-\lambda_1 I)\oplus\ldots\oplus\ker(\alpha-\lambda_k I)$

Linear Algebra (15)

~~Suppose that~~

Application to simultaneous diagonalisability

Suppose that α, β are endomorphisms, both of which have diagonal matrix $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ and $\begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$ with respect to a basis, $\underline{e_1}, \dots, \underline{e_n}$.

Then $\alpha\beta(\underline{e_i}) = \alpha(\cancel{\beta}\underline{e_i}) = \mu_i\, \alpha(\underline{e_i}) = \lambda_i \mu_i \underline{e_i} = \lambda_i \beta(\underline{e_i}) = \beta(\lambda_i \underline{e_i})$
$= \beta\alpha(\underline{e_i})$

And so α, β commute. (OR, evidently the two matrices commute!)

Proposition Suppose α, β are endomorphisms of a finite dimensional V and that both α, β are diagonalisable. If $\alpha\beta = \beta\alpha$ then α, β are simultaneously diagonalisable.

Proof

Write $V = V_1 \oplus \dots \oplus V_k$ where the V_i are the eigenspaces of α. $(\ker(\alpha - \lambda_i I) = V_i)$

Claim: the V_i are β invariant.

For, suppose $\underline{x} \in \ker(\alpha - \lambda I)$ so that $\alpha(\underline{x}) = \lambda \underline{x}$. Then, $\alpha(\beta(\underline{x})) = \beta(\alpha(\underline{x})) = \beta(\lambda \underline{x}) = \lambda \beta(\underline{x})$ and so $\beta(\underline{x}) \in \ker(\alpha - \lambda I)$

Write $\beta_i = \beta|_{V_i}$. $m_{\beta_i} \mid m_\beta$. But as β is diagonalisable, m_β splits into linear factors, hence so does m_{β_i}

So $\beta_i = \beta|_{V_i}$ is diagonalisable. In each V_i, choose a basis of eigenvectors of β. Put these together ~~with~~ to give a basis for V of vectors which are eigenvectors of both α, and β.

5.7 Jordan Normal Forms

Let α be an endomorphism of a finite dimensional complex V.

Take $\chi_\alpha(t) = (-1)^n (t - \lambda_1)^{a_1} \dots (t - \lambda_k)^{a_k}$ with a_i the algebraic multiplicities of the eigenvalues. $n = \dim V = \sum a_i$

Then $m_\alpha(t) = (t - \lambda_1)^{m_1} \dots (t - \lambda_k)^{m_k}$ where $1 \le m_i \le a_i$

FACT α has a matrix A in Jordan Normal Form

This will be described in stages.

① $A = \begin{pmatrix} \boxed{B_1} & & & 0 \\ & \boxed{B_2} & & \\ & & \ddots & \\ 0 & & & \boxed{B_k} \end{pmatrix}$ where the B_i are blocks corresponding to the eigenvalues λ_i. We shall see that B_i has λ_i down the diagonal. The B_i are of size a_i ($a_i \times a_i$ matrices).

This comes from $V = \ker(\alpha - \lambda_1 I)^{m_1} \oplus \dots \oplus \ker(\alpha - \lambda_k I)^{m_k}$ where the $\ker(\alpha - \lambda I)^m$ are the generalised eigenspaces.
(N.B $\ker(\alpha - \lambda I)^m = \ker(\alpha - \lambda I)^M$ for any $M \geq m$)
Why? Suppose $\underline{x_1} + \dots + \underline{x_k} = \underline{0}$ with $\underline{x_i} \in \ker(\alpha - \lambda_i I)^{m_i}$

Set $q_j(t) = \dfrac{m(t)}{(t-\lambda_j)^{m_j}}$. Apply $q_j(\alpha)$ to get
$q_j(\lambda_j)\, \underline{x_j} = \underline{0}$, So $\underline{x_j} = \underline{0}$. So we have a direct sum.
Then this is the whole of V, by the Chinese Remainder Theorem.

② B_i (for λ_i) is of the form:

$B_i = \begin{pmatrix} \boxed{C_1} & & & \\ & \boxed{} & & \\ & & \ddots & \\ & & & \boxed{C_{n_i}} \end{pmatrix}$ where $n_i = \dim(\ker(\alpha - \lambda_i I)) = n(\alpha - \lambda_i I)$
Each C corresponds to a cyclic subspace.

C can therefore be put in the form $\begin{pmatrix} \lambda_i & & 0 \\ 1 & \ddots & \\ 0 & 1 & \lambda_i \end{pmatrix}$

(N.B. A cyclic subspace for α is cyclic for $\alpha - \lambda$ and vice versa)
For there must be a vector $\underline{x}$ such that
$\underline{x}, (\alpha - \lambda_i I)\underline{x}, (\alpha - \lambda_i I)^2\underline{x}, \dots, (\alpha - \lambda_i I)^{n-1}\underline{x}$ is a basis
and with $(\alpha - \lambda_i I)^n \underline{x} = 0$
$\alpha(\underline{x}) = \lambda_i \underline{x} + (\alpha - \lambda_i I)\underline{x}$ (Look at form of C_i)

Each C_i contains a 1D eigenspace. Finally, the maximal size of a C in B_i is m_i.

Linear Algebra (15)

<u>Example</u>

$$\frac{d^n x}{dt^n} + c_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \dots + c_0 x = 0 \quad \text{, an ODE}$$

Solutions $= \{ f \mid p(D)(f) \} \subseteq C^\infty(\mathbb{C})$ where p is the auxiliary polynomial and $D = \frac{d}{dt}$

Rewrite as
$$D\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ & & & 1 \\ -c_0 & - & - & -c_{n-1} \end{pmatrix}}_{A} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}$$

<u>Analysis</u> $\frac{d}{dt}\underline{Z} = A\underline{Z}$ has solution $e^{At}\underline{Z}(0)$ and so a solution is uniquely determined by $x(0), x'(0), \dots, x^{(n-1)}(0)$.
So $V \cong \mathbb{C}^n$. And $D: V \to V$ has matrix A with respect to some basis. $(-1)\chi_D = m_D = p$

$$\chi_D(t) = (-1)^n (t-\lambda_1)^{a_1} \dots (t-\lambda_k)^{a_k}$$
$$m_D(t) = (t-\lambda_1)^{a_1} \dots (t-\lambda_k)^{a_k}$$

So the Jordan Normal Form is
$$\begin{pmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{bmatrix} & & & \\ & \begin{bmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{bmatrix} & & \\ & & \ddots & \\ & & & \begin{bmatrix} \lambda & & 0 \\ & \ddots & 1 \\ 0 & & \lambda \end{bmatrix} \end{pmatrix}$$

Find an explicit basis for the generalised eigenspace $\ker(D-\lambda I)^a$

$$\frac{d}{dt} e^{\lambda t} \mapsto \lambda e^{\lambda t}$$
$$\frac{d}{dt} te^{\lambda t} \mapsto \lambda te^{\lambda t} + e^{\lambda t}$$
$$\frac{d}{dt} \frac{t^2}{2} e^{\lambda t} \mapsto \lambda \frac{t^2}{2} e^{\lambda t} + te^{\lambda t}$$
$$\dots \quad \frac{d}{dt}\frac{t^{a-1}}{(a-1)!}e^{\lambda t} \mapsto \lambda \frac{t^{a-1}}{(a-1)!} e^{\lambda t} + \frac{t^{a-2}}{(a-2)!} e^{\lambda t}$$

$$(D-\lambda I)\ \frac{t^{a-1}}{(a-1)!}e^{\lambda t} \to \frac{t^r}{r!}e^{\lambda t} \dots te^{\lambda t} \to e^{\lambda t} \to 0$$

It is standard linear algebra that these are independent and give a basis for the solution space.

14/11/11

Linear Algebra (16)

Correction

Remark: If $\underline{x} \in V$ finite dimensional, and α is an endomorphism of V, then there is a minimum polynomial $m_{\underline{x}}(t)$ such that $m_{\underline{x}}(\alpha)(\underline{x}) = 0$. Then $m_{\underline{x}}$ divides any polynomial $p(t)$ such that $p(\alpha)(\underline{x}) = \underline{0}$

In fact : $m_{\underline{x}} = m_{\alpha|_{\langle \underline{x}\rangle_\alpha}}$

To show $\ker(\alpha - \lambda_1 I)^{m_1} + \dots + \ker(\alpha - \lambda_k I)^{m_k}$ is a direct sum.

We took $\underline{x}_1 + \dots + \underline{x}_k = \underline{0}$, $\underline{x}_i \in \ker(\alpha - \lambda_i I)^{m_i}$ (†)

and took $q_j(t) = \frac{m(t)}{(t-\lambda_j)^{m_j}} = \prod_{i \neq j}(t-\lambda_i)^{m_i}$ and applied $q_j(\alpha)$ to (†)

We deduce that $q_j(\alpha)(\underline{x}_j) = \underline{0}$

But, $\underline{x}_j$ is not an eigenvector necessarily. However, if $\underline{x}_j \neq \underline{0}$, its minimum polynomial divides $(t-\lambda_j)^{m_j}$ and is of the form $(t-\lambda_j)^{m_j'}$, $m_j' \geq$

But $(t-\lambda_j)^{m_j'} \nmid q_j(t)$ ※ $\Rightarrow \underline{x}_j = \underline{0}$.

6 Duality

6.1 Dual Spaces and dual bases

Definition If V is a vector space then its dual V^* is the vector space $L(V, F)$ of linear functionals on V.

Suppose $\underline{e}_1, \dots, \underline{e}_n$ is a basis for V. We define the "dual basis" for V^* $\varepsilon_1, \dots, \varepsilon_n$ by setting $\varepsilon_i(\underline{e}_j) = \delta_{ij}$

Note that $\varepsilon_i(\sum x_j \underline{e}_j) = x_i$. So the ε_i are the old coordinate functions $V \xrightarrow[\varepsilon]{\underline{x}} \mathbb{F}^n$; $\underline{x} \mapsto \begin{pmatrix} \varepsilon_1(\underline{x}) \\ \\ \varepsilon_n(\underline{x}) \end{pmatrix}$

Proposition ~~Suppose~~ $\varepsilon_1, \dots, \varepsilon_n$ is a basis for V^*.

Proof

Suppose $\sum \lambda_i \varepsilon_i = 0$. Then for any j, $\sum_i \lambda_i \varepsilon_i(\underline{e}_j) = 0$ and so $\lambda_j = 0$. This shows independence.

Suppose that $\theta \in V^*$ and let $t_i = \theta(\underline{e}_i)$

$\sum_i t_i \varepsilon_i(\underline{e}_j) = t_j = \theta(\underline{e}_j)$ for all $\underline{e}_j$, and so

$\theta = \sum_i t_i \varepsilon_i$ and the ε_i span V^*

Corollary $\dim V = \dim V^*$

 WARNING Let V be the space of sequences from $\mathbb{F}$.

generated by the obvious unit vectors $(1,0,\dots)$, $(0,1,0,\dots)$ etc.
It is countable dimensional but $V^* \cong \mathbb{F}^{\mathbb{N}}$ is <u>not</u>

<u>Aside</u>

If $\theta = \sum t_i \varepsilon_i$ and $\underline{x} = \sum x_i \underline{e_i}$, then $\theta(\underline{x}) = \sum_k t_k x_k$
$= (t_1, \dots, t_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

<u>Also Aside</u> The map $V^* \times V \to \mathbb{F}$, $\theta, \underline{x} \mapsto \theta(\underline{x})$ is bilinear.

- $\theta(\lambda \underline{x} + \mu \underline{y}) = \lambda\theta(\underline{x}) + \mu\theta(\underline{y})$, θ linear
- $(\lambda\phi + \mu\psi)(\underline{x}) = \lambda\phi(\underline{x}) + \mu\psi(\underline{x})$ by definition of $L(V, F)$

"Think $\theta(\underline{x}) = \langle \theta, \underline{x} \rangle$"

6.2 The dual of a linear map

<u>Definition</u> Let $\alpha: U \to V$ be linear. Then the dual map $\alpha^*: V^* \to U^*$ is defined by $\alpha^*(\theta)(\underline{u}) = \theta(\alpha(\underline{u}))$ for $\theta \in V^*$, $\underline{u} \in U$

<u>Remark</u>

$\alpha^*(\theta)$ is the composite $\theta \circ \alpha$ and so is automatically linear. $\alpha^*(\theta) \in U^*$ for $\theta \in V^*$. "Think $\langle \alpha^*(\theta), u \rangle = \langle \theta, \alpha(u) \rangle$"

<u>Proposition</u>

$\alpha^*: V^* \to U^*$ is linear.

<u>Proof</u> $\alpha^*(\lambda\phi + \mu\psi)(\underline{u}) = (\lambda\phi + \mu\psi)(\alpha(\underline{u})) = \lambda\phi(\alpha(\underline{u})) + \mu\psi(\alpha(\underline{u}))$
(for $\phi, \psi \in V^*$, $\underline{u} \in U$). $\quad = \lambda\alpha^*(\phi)(\underline{u}) + \mu\alpha^*(\psi)(\underline{u})$
$\quad = (\lambda\alpha^*\phi + \mu\alpha^*\psi)(\underline{u})$

And so $\alpha^*(\lambda\phi + \mu\psi) = \lambda\alpha^*\phi + \mu\alpha^*\psi$ i.e. α^k is linear.

<u>Note</u> $\begin{array}{c} U \xrightarrow{I} U \\ U^* \xleftarrow{I} U^* \end{array}$ $\qquad \begin{array}{c} U \xrightarrow{\alpha} V \xrightarrow{\beta} W \\ U^* \xleftarrow{\alpha^*} V^* \xleftarrow{\beta^*} W^* \end{array}$ $\qquad (\beta\alpha)^* = \alpha^*\beta^*$

<u>Proposition</u> Suppose $\alpha: U \to V$ has matrix A with respect to bases $\underline{e_1}, \dots, \underline{e_n}$ for U and $\underline{f_1}, \dots, \underline{f_m}$ for V. Then, α^* has the matrix A^T with respect to the dual basis $\varepsilon_1, \dots, \varepsilon_n$ and $\phi_1, \dots, \phi_m$

<u>Proof</u>

$$\alpha^*(\phi_j)(\underline{e_k}) = \phi_j(\alpha(\underline{e_k})) = \phi_j\left(\sum_i a_{ik} \underline{f_i}\right) = \sum_i a_{ik}\, \phi_j(\underline{f_i})$$
$$= a_{jk}.$$

So $\alpha^*(\phi_j) = \sum_k a_{jk}\varepsilon_k = \sum_k [A^T]_{kj}\varepsilon_k$. Then α^* has matrix A^T.

Note $r(\alpha) = \text{colrk}(A) = \text{rowrk}(A) = \text{colrk}(A^T) = r(\alpha^*)$

14/11/11

Linear Algebra (17)

5.3 Annihilators

Definition Let $W \leq V$. The annihilator W° of W is
$$W^\circ = \{\theta \in V^* \mid \theta(\underline{w}) = 0 \ \forall \underline{w} \in W\}$$

Note Let $\imath : W \to V$ be the inclusion map.
Then $W^\circ = \ker(\imath^*)$, $\imath^* : V^* \to W^*$
For $\imath^*(\theta) = 0$ iff $\theta \circ \imath = 0$ iff θ is 0 on W.

Proposition $W^\circ \leq V^*$

Proof either by the above Note
OR: $0(\underline{w}) = 0 \ \forall \underline{w} \in W$ so $0 \in W^\circ$
If $\phi, \psi \in W^\circ$ then $\forall \underline{w} \in W$, $(\lambda\phi + \mu\psi)(\underline{w}) = \lambda\phi(\underline{w}) + \mu\psi(\underline{w}) = 0$
then $\lambda\phi + \mu\psi \in W^\circ$.

Facts $\{\underline{0}\}^\circ = V^*$, $V^\circ = \{0\}$
If $W_1 \leq W_2$ then $W_1^\circ \geq W_2^\circ$

Proposition Let $W \leq V$ be finite dimensional. Then $\dim W + \dim W^\circ = \dim V$

Proof
Let $\underline{e}_1, \dots, \underline{e}_r$ be a basis for W and extend to a basis $\underline{e}_1, \dots, \underline{e}_n$ for V.
Let $\varepsilon_1, \dots, \varepsilon_n$ be the dual basis for V^*.
CLAIM $\varepsilon_{r+1}, \dots, \varepsilon_n$ is a basis for W°.
- If $r+1 \leq i \leq n$ then $\forall j \leq r$ we have $\underline{e}_i(\varepsilon_j) = 0$, so ε_i is 0 on a basis for W, and so on W. Thus $\varepsilon_i \in W^\circ$.
- $\varepsilon_{r+1}, \dots, \varepsilon_n$ is independent because $\varepsilon_1, \dots, \varepsilon_n$ is a basis.
- Let $\theta \in W^\circ$. Write $\theta = \sum_{i=1}^n t_i \varepsilon_i$. If $1 \leq j \leq r$, then
$0 = \theta(\underline{e}_j) = \sum_{i=1}^n t_i \varepsilon_i(\underline{e}_j) = t_j$
Thus $\theta = \sum_{i=r+1}^n t_i \varepsilon_i$. So $\varepsilon_{r+1}, \dots, \varepsilon_n$ span W°. That proves the claim. Now, $\dim W = r$, $\dim W^\circ = n - r$ so $\dim W + \dim W^\circ = \dim V$

Remark This says that $r(\imath) = r(\imath^*)$ for $\imath : W \to V$ the inclusion.

Observation If $U, W \leq V$ then:
i) $(U + W)^\circ = U^\circ \cap W^\circ$
ii) If V is finite dimensional then $U^\circ + W^\circ = (U \cap W)^\circ$

Why? i) $\theta \in (U+W)^\circ$ iff $U + W \leq \ker \theta$

$\theta \in U^\circ \cap W^\circ$ iff $U \leq \ker\theta$, $W \leq \ker\theta$

Now the result follows as $U + W$ is the least subspace $\geq U, W$

~~ii) Clearly $U^\circ \geq (U+W)^\circ$, $W^\circ \geq (U+W)^\circ$. So $(U+W)^\circ \leq U^\circ \cap W^\circ$~~

~~$\dim (U+W)^\circ = \dim V - \dim(U+W) = n - (\dim U + \dim W - \dim(U\cap W))$~~

ii) $(U\cap W)^\circ \geq U^\circ$, $(U \cap W)^\circ \geq W^\circ$, so $(U\cap W)^\circ \geq U^\circ + W^\circ$

$$\begin{aligned}\dim(U^\circ + W^\circ) &= \dim U^\circ + \dim W^\circ - \dim(U^\circ \cap W^\circ)\\ &= n - \dim U + n - \dim W - (n - \dim(U+W))\\ &= n - (\dim U + \dim W - \dim(U+W))\\ &= n - \dim(U\cap W) = \dim(U\cap W)^\circ\end{aligned}$$

6.4 The rank of the Dual

Theorem Suppose $\alpha : U \to V$ is linear with dual $\alpha^* : V^* \to U^*$

Then i) $\ker \alpha^* = (\mathrm{Im}\,\alpha)^\circ$

ii) $\mathrm{Im}\,\alpha^* = (\ker\alpha)^\circ$ and $r(\alpha) = r(\alpha^*)$ ((ii) only in the finite dimensional case)

Proof

i) $\theta \in \ker\alpha^*$ iff $\alpha^*(\theta) = 0$ iff $\theta \circ \alpha = 0$

iff $\theta = 0$ on any $w \in \mathrm{Im}\,\alpha$ iff $\theta \in (\mathrm{Im}\,\alpha)^\circ$

ii) In the finite dimensional case we have:

$r(\alpha^*) = n - n(\alpha^*) = n - \dim(\mathrm{Im}\,\alpha)^\circ = \dim(\mathrm{Im}\,\alpha) = r(\alpha)$

Finally, we always have $\mathrm{Im}\,\alpha^* \leq (\ker\alpha)^\circ$. For if $\phi = \alpha^*(\theta) \in \mathrm{Im}\,\alpha^*$ then for $\underline{x} \in \ker\alpha$, we have $\phi(\underline{x}) = \alpha^*(\theta)(\underline{x}) = \theta(\alpha(\underline{x})) = \theta(0) = 0$

Thus $\phi \in \ker\alpha$.

But $\dim(\ker\alpha)^\circ = n - n(\alpha) = r(\alpha) = r(\alpha^*) = \dim(\mathrm{Im}\,\alpha^*)$

6.5 The Double Dual

If $\underline{x} \in V$ then we get $\hat{\underline{x}} : V^* \to F$ by $\hat{\underline{x}}(\theta) = \theta(\underline{x})$

(i.e. $\hat{\underline{x}}$ is "evaluated at $\underline{x}$")

Consequence of bilinearity:

1. $\hat{\underline{x}}$ is a linear map so $\hat{\underline{x}} \in L(V^*, F) = V^{**}$
2. The map $K: V \to V^*$, $\underline{x} \mapsto \hat{\underline{x}}$ is itself linear.

14/11/11

Linear Algebra (17)

This is natural in the sense that given $\alpha : U \to V$,
the diagram

$$\begin{tikzcd} U \arrow[r, "\alpha"] \arrow[d, "K"'] & V \arrow[d, "K"] \\ U^{**} \arrow[r, "\alpha^{**}"] & V^{**} \end{tikzcd}$$

commutes.

Why? $K_V(\alpha(u))$ acts on $\theta \in V^*$ to give $\theta(\alpha(u))$
and $\alpha^{**}(K_U(u))$ acts on $\theta \in V$ to give $\hat{u} \circ \alpha^*(\theta) = \theta \circ \alpha(\underline{u})$

Proposition If V is finite dimensional, then $K : V \to V^{**}$ is an isomorphism
Proof
We show that K is injective. Suppose $\underline{x} \in V$ with $\underline{x} \neq \underline{0}$. Set $\underline{e}_1 = \underline{x}$ and extend to a basis $\underline{e}_1, \dots, \underline{e}_n$ for V. Take the dual basis $\varepsilon_1, \dots, \varepsilon_n$. Then $\hat{x}(\varepsilon_1) = \varepsilon_1(\underline{e}_1) = 1$ so $\hat{x} \neq 0$.
Thus, the kernel of K is $\{\underline{0}\}$.
Thus K is injective, but $V^{**} = V^* = V$. So K is an isomorphism.

Observe If $U \leq V$ then $K(U) \leq V^{**}$ is equal to U^{oo}
We can therefore identify V with V^{**} and get a real duality $(\ )^o$ on subspaces.

16/11/11

Linear Algebra (18)

Chapter 7 Bilinear Forms

7.1 Bilinear Maps

$\beta : U \times V \to \mathbb{F}$ is bilinear just when

$\beta(\lambda \underline{u} + \mu \underline{u}', \underline{v}) = \lambda \beta(\underline{u}, \underline{v}) + \mu \beta(\underline{u}', \underline{v})$

$\beta(\underline{u}, \lambda \underline{v} + \mu \underline{v}') = \lambda \beta(\underline{u}, \underline{v}) + \mu \beta(\underline{u}, \underline{v}')$

In this situation we have

$\beta_L : U \to V^* \ : \ \beta_L(\underline{u})(\underline{v}) = \beta(\underline{u}, \underline{v})$

$\beta_R : V \to U^* \ : \ \beta_R(\underline{v})(\underline{u}) = \beta(\underline{u}, \underline{v})$ both linear maps.

Note The composite $V \xrightarrow{k} V^{**} \xrightarrow{\beta_L^*} U^*$ is equal to β_R

This "duality" is reflected in terms of matrices.

Suppose that $e_1, \dots, e_n$ and $f_1, \dots, f_m$ are bases for U and V, and let B be the matrix $B = (b_{ij})$ with $b_{ij} = \beta(e_i, f_j)$

Then $\beta_L(e_i)(f_j) = b_{ij}$ and so $\beta_L(e_i) = \sum b_{ij} \phi_j$ where $\phi_1, \dots, \phi_m$ is the basis dual to $f_1, \dots, f_m$.

Similarly, β_R has matrix B with respect to $f_1, \dots, f_m$ and $\varepsilon_1, \dots, \varepsilon_n$

β is left non-degenerate iff β_L is injective.

β is right non-degenerate iff β_R is injective.

$\Rightarrow$ If $\underline{u} \in U$ is such that $\beta(\underline{u}, \underline{v}) = 0 \ \forall \underline{v} \in V$ then $\underline{u} = \underline{0}$

In the finite dimensional case we are interested in β_L being an isomorphism.

This happens iff the matrix B above is non-singular and so iff β_R is an isomorphism (In this case $\dim U = \dim V$).

If $X \leq U$ we have $X^{\perp} = \{ \underline{v} \in V \mid \beta(\underline{x}, \underline{v}) = 0 \ \forall \underline{x} \in X \} \leq V$

If $Y \leq V$ we have ${}^{\perp}Y = \{ \underline{u} \in U \mid \beta(\underline{u}, \underline{y}) = 0 \ \forall \underline{y} \in Y \} \leq U$

$\underline{v} \in X^{\perp}$ iff $\beta_L(\underline{x})(\underline{v}) = 0 \ \forall \underline{x} \in X$

iff $\hat{v} \in (\beta_L(X))^{\circ}$

So in the finite dimensional case, we have (for β non-singular)

$\dim X + \dim X^{\perp} = \dim \beta_L(X) + \dim (\beta_L(X))^{\circ} = \dim V$

Suppose again in the finite dimensional case that β is non singular.
Then for any $\theta \in V^*$ there is a unique $\underline{u}_\theta \in U$ such that $\beta(\underline{u}_\theta, \underline{v}) = \theta(\underline{v})$ $\forall \underline{v} \in V$.

Proposition Let $\beta: U \times V \to \mathbb{F}$ be non singular bilinear with U, V finite dimensional. Then for any endomorphism α of V there is a unique adjoint endomorphism α^+ of U such that $\beta(\alpha^+(\underline{u}), \underline{v}) = \beta(\underline{u}, \alpha(\underline{v}))$

Proof
For fixed $\underline{u} \in U$, $\underline{v} \mapsto \beta(\underline{u}, \alpha(\underline{v}))$ is in V^* so we have a unique $\alpha^+(\underline{u}) \in U$ such that $\beta(\alpha^+(\underline{u}), \underline{v}) = \beta(\underline{u}, \alpha(\underline{v}))$.
It remains to show that α^+ is linear.

$$\beta(\alpha^+(\lambda\underline{x}+\mu\underline{y}), \underline{v}) = \beta(\lambda\underline{x}+\mu\underline{y}, \alpha(\underline{v})) = \lambda\beta(\underline{x}, \alpha(\underline{v})) + \mu\beta(\underline{y}, \alpha\underline{v})$$
$$= \lambda\beta(\alpha^+(\underline{x}), \underline{v}) + \mu\beta(\alpha^+(\underline{y}), \underline{v}) = \beta(\lambda\alpha^+(\underline{x}) + \mu\alpha^+(\underline{y}), \underline{v})$$
and this holds for all $\underline{v}$, hence α^+ is linear.

7.2 Bilinear Forms
A bilinear form β on V is a bilinear map $\beta: V \times V \to \mathbb{F}$
Given a basis $\underline{e}_1, \dots, \underline{e}_n$ for V, the matrix $B = (b_{ij})$ for β with respect to $\underline{e}_1, \dots, \underline{e}_n$ is $b_{ij} = \beta(\underline{e}_i, \underline{e}_j)$
If $\underline{x} = \sum x_i \underline{e}_i$ and $\underline{y} = \sum y_j \underline{e}_j$ then $\beta(\underline{x}, \underline{y}) = \sum_{ij} x_i b_{ij} y_j$
That is $(x_1 \dots x_n)\begin{pmatrix} & & \\ & B & \\ & & \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \mathcal{E}(\underline{x})^T B \mathcal{E}(\underline{y})$

Suppose β has the matrix $B' = (b_{ij}')$ with respect to $\underline{e}_1', \dots, \underline{e}_n'$
EITHER $\underline{x} = \sum x_i' \underline{e}_i'$ and $\underline{y} = \sum y_j' \underline{e}_j'$ and
$\beta(\underline{x}, \underline{y}) = \sum x_i' b_{ij}' y_j' = \sum_{i,r,j,s} p_{ir} x_r b_{ij}' p_{js} y_s = \sum_{r,s} x_r (P^T B' P)_{rs} y_s$
and so $B = P^T B' P$
OR $\beta(\underline{e}_i', \underline{e}_j') = \beta(\sum \hat{p}_{ri} \underline{e}_r, \sum \hat{p}_{sj} \underline{e}_s) = \sum \hat{p}_{ri} b_{rs} \hat{p}_{sj}$
$= (\hat{P}^T B \hat{P})_{ij}$

And so $B' = \hat{P}^T B \hat{P}$, $\hat{P} = P^{-1}$
We are interested in finding good bases with respect to which a form has a simple matrix: e.g. if we have $V = U \oplus W$ with $\beta(\underline{u}, \underline{w}) = 0 = \beta(\underline{w}, \underline{u})$ $\forall \underline{u} \in U, \underline{w} \in W$, then we have a matrix

Linear Algebra (18)

of the form $\left(\begin{array}{c|c} B_U & 0 \\ \hline 0 & B_W \end{array}\right)$

WARNING! It is natural to try "$U \oplus U^\perp$"
BUT in general this is not a direct sum even though $\dim U + \dim U^\perp = \dim$ (assuming β is non-singular).
For example, take $\begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}$

There are null vectors e.g. $(1, 1, 0, 0)$ such that

$(1\ 1\ 0\ 0)\begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ i.e. self-orthogonal

So if $(1\ 1\ 0\ 0) \in U$ then $U \cap U^\perp \neq \{\underline{0}\}$

16/11/11

Linear Algebra (19)

Proposition Suppose β is a bilinear form which is non-singular on a subspace $U \leq V$. Then $V = U \oplus U^\perp$

Proof

Suppose $\underline{u} \in U \cap U^\perp$. Then $\forall \underline{x} \in U$, $\beta(\underline{x}, \underline{u}) = 0$. But as $\beta_R(\underline{u})|_U = 0$ and β is non singular on U, that implies $\underline{u} = \underline{0}$.

Take $\underline{v} \in V$. Consider the map $U \to \mathbb{F}$, $\underline{y} \mapsto \beta(\underline{y}, \underline{v})$. It is in U^* and so as β on U is non-singular there exists a (unique) $\underline{w} \in U$ with $\beta(\underline{y}, \underline{v}) = \beta(\underline{y}, \underline{w})$ $\forall \underline{y} \in U$.

Then $\underline{v} - \underline{w} \in U^\perp$ and $\underline{v} = \underbrace{\underline{w}}_{\in U} + \underbrace{(\underline{v} - \underline{w})}_{\in U^\perp}$

7.3 Symmetric Bilinear Forms

A bilinear form $\beta: V \times V \to \mathbb{F}$ is

symmetric just when $\beta(\underline{x}, \underline{y}) = \beta(\underline{y}, \underline{x})$ $\forall \underline{x}, \underline{y} \in V$

skew-symmetric $\beta(\underline{x}, \underline{y}) = -\beta(\underline{y}, \underline{x})$

ASSUME for this section that char $\mathbb{F} \neq 2$ (i.e. $2 \neq 0$ in $\mathbb{F}$)

Any bilinear form can be written:

$$\beta(\underline{x}, \underline{y}) = \tfrac{1}{2}[\beta(\underline{x}, \underline{y}) + \beta(\underline{y}, \underline{x})] + \tfrac{1}{2}[\beta(\underline{x}, \underline{y}) - \beta(\underline{y}, \underline{x})]$$

as the sum of a symmetric and skew-symmetric form.

A quadratic form $q: V \to \mathbb{F}$ is a map such that $q(\underline{x}) = \beta(\underline{x}, \underline{x})$ for some bilinear β. We might as well take β to be symmetric and then β is determined by q: $\beta(\underline{x}, \underline{y}) = \frac{1}{2}(q(x+y) - q(x) - q(y))$

Proposition Suppose β is a symmetric bilinear form on a finite dimensional V. Then there is a basis $\underline{e}_1, \ldots, \underline{e}_n$ with respect to which β has a diagonal matrix (i.e. $\beta(\underline{e}_i, \underline{e}_j) = 0$ for $i \neq j$)

Lemma If β is a symmetric bilinear form such that the corresponding quadratic form $q(\underline{x}) = \beta(\underline{x}, \underline{x})$ is 0, then β is 0.

Proof

$0 = \beta(\underline{x}+\underline{y}, \underline{x}+\underline{y}) = \beta(\underline{x}, \underline{x}) + 2\beta(\underline{x}, \underline{y}) + \beta(\underline{y}, \underline{y}) = 2\beta(\underline{x}, \underline{y})$

So $\beta(\underline{x}, \underline{y}) = 0$

Proof of Proposition (By induction on $\dim V \geq 1$)
The initial case with $\dim V = 1$ is trivial.
For the induction step, assume true for all vector spaces of $\dim V < n$ and suppose $\dim V = n$.
EITHER $\beta : V \times V \to \mathbb{F}$ is 0 and any basis will do.
OR we can find a vector $\underline{e_1}$ with $\beta(\underline{e_1}, \underline{e_1}) = \lambda \neq 0$

CLAIM
$V = \langle \underline{e_1} \rangle \oplus \langle \underline{e_1} \rangle^{\perp}$

Take any $\underline{v} \in V$. $\underline{v} = \frac{\beta(\underline{v}, \underline{e_1})}{\lambda} \underline{e_1} + \left(v - \frac{\beta(\underline{v}, \underline{e_1})}{\lambda} \underline{e_1} \right)$

$\frac{1}{\lambda} \beta(\underline{v}, \underline{e_1}) \, \underline{e_1} \in \langle \underline{e_1} \rangle$

$\beta\left(\underline{e_1}, v - \frac{\beta(\underline{v}, \underline{e_1})}{\lambda} \underline{e_1}\right) = \beta(\underline{e_1}, \underline{v}) - \frac{\beta(\underline{e_1}, \underline{v})}{\cancel{\lambda}} \cancel{\beta(\underline{e_1}, \underline{e_1})} = 0$

So $\left(\underline{v} - \frac{1}{\lambda} \beta(\underline{v}, \underline{e_1}) \underline{e_1} \right) \in \langle \underline{e_1} \rangle^{\perp}$

This expression is unique because if $\underline{v} = \mu \underline{e_1} + \underline{u}$ with $\underline{u} \in \langle \underline{e_1} \rangle^{\perp}$
then $\beta(\underline{e_1}, \underline{v}) = \mu \beta(\underline{e_1}, \underline{e_1}) + 0$, so $\mu = \frac{1}{\lambda} \beta(\underline{e_1}, \underline{v})$ and
$\underline{u} = (\underline{v} - \mu \underline{e_1})$

By the induction hypothesis, there is a basis $\underline{e_1}, \dots, \underline{e_n}$ of $U = \langle \underline{e_1} \rangle^{\perp}$ with respect to which β has diagonal matrix. And then β has a diagonal matrix with respect to $\underline{e_1}, \dots, \underline{e_n}$.

Special Case $\mathbb{F} = \mathbb{R}$. We have a matrix

$$\left(\begin{array}{cc|c} d_1 & 0 & \\ & \ddots & 0 \\ 0 & d_r & \\ \hline & 0 & 0 \end{array} \right)$$

for β, $d_i \neq 0$ and we can ensure that $d_1, \dots, d_p > 0$ and then $d_{p+1}, \dots, d_r < 0$.
Set a new basis $\underline{e_i}' = \underline{e_i} / \sqrt{|d_i|}$ for $1 \leq i \leq r$ and we get a new matrix of the form

$$\left(\begin{array}{ccc} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0 \end{array} \right) \qquad p + q = r$$

Special Case $\mathbb{F} = \mathbb{C}$. We have a matrix as for $\mathbb{R}$ and set
$\underline{e_i}' = \frac{\underline{e_i}}{\sqrt{d_i}}$ for $1 \leq i \leq r$.
We get a matrix of the form

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right)$$

16/11/11

Linear Algebra (19)

7.4 Sylvester's Law of Inertia

Suppose β is a symmetric bilinear form on a finite dimensional real V.
Suppose that β has matrices

$$\begin{pmatrix} I_p & & 0 \\ & -I_q & \\ 0 & & 0 \end{pmatrix} \text{ and } \begin{pmatrix} I_{p'} & & 0 \\ & -I_{q'} & \\ 0 & & 0 \end{pmatrix}$$ with respect to two bases.

Then $p = p'$, $q = q'$.

β is positive semi-definite on U iff $\beta(\underline{u}, \underline{u}) \geq 0 \ \forall \underline{u} \in U$
positive definite on U iff $\beta(\underline{u}, \underline{u}) > 0 \ \forall \underline{u} \in U$
negative semi-definite ≤ 0
negative definite < 0

Suppose U is a space on which β is +ve definite and W a space on which β is -ve semi definite. Take $\underline{x} \in U \cap W$, $\beta(\underline{x}, \underline{x}) \leq 0$ and $\beta(\underline{x}, \underline{x}) > 0$ unless $\underline{x} = \underline{0}$. So $\underline{x} = \underline{0}$. Thus $U \oplus W$ exist

For the first basis we have a space P, $\dim P = p$, on which β is +ve definite, and a space S, $\dim S = n-p$, on which β is -ve semi-definite. Similarly, for the 2nd basis we have a space P', $\dim P' = p'$, on which β is +ve definite and a space S', $\dim S' = n-p'$ on which β is -ve semi-definite.

The direct sum $P \oplus S'$ exists, so $p + n - p' \leq n$ i.e. $p \leq p'$
Similarly, $p' \leq p$.
Similarly, for Q -ve definite and T positive semi-definite, and Q', T', q, q' then $q \leq q'$, $q' \leq q$.

The rank of the bilinear form is the rank of any matrix is $p+q$ in this case
signature $p-q$
These are invariants of the bilinear forms.

3

17/11/11

Linear Algebra 20

Invariance of Rank

1. Suppose B, B' are matrices for a bilinear form β. We have $B' = Q^T B Q$ where Q is invertible.
Then $r(B') = r(B)$. So we can define the rank of β to be the rank of any matrix for β.
[Aside: Suppose S is invertible. $r(SA) \leq r(A) = r(S^{-1}(SA)) \leq r(SA)$
So $r(SA) = r(A)$ and similarly $r(AS) = r(A)$]

2. The rank of β as a form is the rank of the linear map $\beta_L : V \to V^*$. Then we ~~wrote~~ noted that if β has matrix B as a form, then β_L has matrix B^T and so the rank of a matrix for β is an invariant.

Consequence

When we have two matrices

$$\begin{pmatrix} I_p & & 0 \\ & -I_q & \\ 0 & & 0 \end{pmatrix} \qquad \begin{pmatrix} I_{p'} & & 0 \\ & -I_{q'} & \\ 0 & & 0 \end{pmatrix}$$

a real symmetric form, then we know $p+q = p'+q'$. So it suffices to show that $p = p'$ and deduce $q = q'$.

Examples

Diagonalising is "completing the square".

1. $2x^2 + 2xy + y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$
$= 2(x + \frac{1}{2}y)^2 + \frac{1}{2}y^2$

$$= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$= x^2 + (x+y)^2 \qquad = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

2. $x^2 + y^2$
Take $e_1 = \frac{1}{\sqrt{5}}(2 \ \ 1)$. Find $\underline{e_2} \in \langle \underline{e_1} \rangle^{\perp}$. Suppose $\underline{e_2} = (u, v)$
$\begin{pmatrix} 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$, $2u + v = 0$. So $u = 1$, $v = -2$ for example

$$= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

7.5 Hermitian Forms

For any vector space V over $\mathbb{C}$ there is a vector space $\bar{V}$ over $\mathbb{C}$ which is V with the same addition and with scalar multiplication $(\lambda, v) \mapsto \bar{\lambda} v$

A sesquilinear form is a bilinear map $\tau : \bar{V} \times V \to \mathbb{C}$. What does it mean?

$\tau(\underline{x}, \underline{y})$ is linear in $\underline{y}$ and antilinear in $\underline{x}$:

$\tau(\lambda \underline{x} + \mu \underline{x}', \underline{y}) = \bar{\lambda} \tau(\underline{x}, \underline{y}) + \bar{\mu} \tau(\underline{x}', \underline{y})$

e.g. the complex inner product

$(z_1, \dots z_n) \cdot (w_1 \dots w_n) = \sum_{i=1}^{n} \bar{z}_i w_i$

Note that if $\tau(\underline{x}, \underline{y})$ is sesquilinear then so is $\overline{\tau(\underline{y}, \underline{x})}$. We say τ is Hermitian just when $\tau(\underline{x}, \underline{y}) = \overline{\tau(\underline{y}, \underline{x})}$ and skew Hermitian just when $\tau(\underline{x}, \underline{y}) = -\overline{\tau(\underline{y}, \underline{x})}$.

Any sesquilinear form is the sum of an Hermitian and a skew Hermitian form.

The matrix C for a Hermitian τ satisfies $C = C^{\dagger} = \overline{C}^{T}$

Note that for τ, C Hermitian we have $\tau(\underline{x}, \underline{x}) = \overline{\tau(\underline{x}, \underline{x})}$ real and the diagonal entries of C are real.

As with last time, we can diagonalise a Hermitian form, getting it to the form $\begin{pmatrix} I_p & & 0 \\ & -I_q & \\ 0 & & 0 \end{pmatrix}$. Note that we cannot do more because $\tau(\lambda \underline{x}, \lambda \underline{x}) = |\lambda|^2 \tau(\underline{x}, \underline{x})$

[Aside: we have a change of basis $C' = Q^{\dagger} C Q$, Q invertible]

Appendix I Given $\begin{pmatrix} \overset{P}{I_p} & & 0 \\ & \overset{Q}{-I_q} & \\ 0 & & 0 \end{pmatrix}$ what is the maximal dimension of a space $Z \leq V$ in which β is 0?

Basis $\underline{e_1}, \dots, \underline{e_n}$

β is 0 on $\langle \underline{e}_{p+q+1}, \dots, \underline{e_n}, \underline{e_1} + \underline{e}_{p+1}, \dots \rangle$

which is of dimension $n - (p+q) + \min(p, q) \quad = \min(n-p, n-q)$

But also, if β is 0 on Z then $P \cap Z = Q \cap Z = \{\underline{0}\}$

so $p + \dim Z \leq n$, $q + \dim Z \leq n$, $\dim Z \leq {}^{n-p}_{n-q}$

17/11/11

Linear Algebra 20

Appendix II

An abstract approach to symmetric and Hermitian forms.

Let $p = \max\{\dim U \mid \beta \text{ is +ve definite on } U \leq V\}$

$q = \max\{\dim W \mid \beta \text{ is -ve definite on } W \leq V\}$

Pick U, W realising these maxima.

β on U is +ve definite and we can find an orthonormal basis

$-\beta$ on W is -ve definite, so the same idea applies.

β is non-singular on $U \oplus W$ and so

$$V = (U \oplus W) \oplus (U \oplus W)^{\perp}$$

Let $v \in (U \oplus W)^{\perp}$. Suppose $\beta(\underline{v}, \underline{v}) > 0$. Then β is +ve definite on $U \oplus \langle \underline{v} \rangle$ ※

Similarly if $\beta(\underline{v}, \underline{v}) < 0$ then β is -ve definite on $W \oplus \langle \underline{v} \rangle$ ※.

So β is 0 on $(U \oplus W)^{\perp}$ and taking a basis for it with the orthonormal basis gives

$$\begin{pmatrix} I_p & & 0 \\ & -I_q & \\ 0 & & 0 \end{pmatrix}$$

21/11/11

Linear Algebra (21)

Chapter 8 Inner Product Spaces

8.1 Inner Products

Definition: A real/complex inner product space is a real/complex vector space V equipped with a positive definite symmetric/hermitian form $\langle\ ,\ \rangle$ the inner product.

Remarks

(1) An inner product is automatically non-singular. If $\underline{x} \in V$ is such that $(y \mapsto \langle \underline{y}, \underline{x} \rangle)$ is 0 then in particular $\langle \underline{x}, \underline{x} \rangle = 0$ so $\underline{x} = 0$.
So the map $V \to \bar{V}^*$, $\underline{x} \mapsto (\underline{y} \mapsto \langle \underline{y}, \underline{x} \rangle)$ is an isomorphism.

(2) Write the quadratic form as $\|\underline{x}\|^2 = \langle \underline{x}, \underline{x} \rangle$. We have $\|\underline{x}\|^2 \geq 0$ in $\mathbb{R}$ and so set $\|\underline{x}\| = +\sqrt{\|\underline{x}\|^2}$. This is the norm or length of $\underline{x}$.

(3) Parallelogram Law: $\|\underline{x}+\underline{y}\|^2 + \|\underline{x}-\underline{y}\|^2 + 2\|\underline{x}\|^2 + 2\|\underline{y}\|^2$ holds

(4) The Cauchy-Schwarz Inequality:
$|\langle \underline{x}, \underline{y} \rangle|^2 \leq \|\underline{x}\|^2 \|\underline{y}\|^2$ with equality just when $\underline{x}, \underline{y}$ are independent.
For example, we have $\langle \underline{x} - \lambda \underline{y}, \underline{x} - \lambda \underline{y} \rangle \geq 0$ (equality just when $\underline{x} = \lambda \underline{y}$.
i.e. $\|\underline{x}\|^2 - \lambda \langle \underline{x}, \underline{y} \rangle - \bar{\lambda} \langle \underline{y}, \underline{x} \rangle + |\lambda|^2 \|\underline{y}\|^2 \geq 0$
Assume $\underline{y} \neq \underline{0}$ (otherwise the result is trivial) and set $\lambda = \frac{\langle \underline{x}, \underline{y} \rangle}{\|\underline{y}\|^2} = \frac{\langle \underline{y}, \underline{x} \rangle}{\|\underline{x}\|}$

We get $2\frac{|\langle x, y \rangle|^2}{\|y\|^2} \leq \|\underline{x}\|^2 + \frac{|\langle x, y \rangle|^2 \cdot \|y\|^2}{\|y\|^4}$ giving the result, with equality just when $\underline{x}$ depends on $\underline{y}$.

(5) The inner products are determined by the quadratic forms.
Real $\langle \underline{x}, \underline{y} \rangle = \frac{1}{4}\left(\|\underline{x}+\underline{y}\|^2 - \|\underline{x}-\underline{y}\|^2\right)$
Complex $\langle \underline{x}, \underline{y} \rangle = \frac{1}{4}\big(\|\underline{x}+\underline{y}\|^2 - i\|\underline{x}+i\underline{y}\|^2$
$- \|\underline{x}-\underline{y}\|^2 + i\|\underline{x}-i\underline{y}\|^2\big)$

8.2 Orthogonality

We say $\underline{x}, \underline{y}$ are orthogonal just when $\langle \underline{x}, \underline{y} \rangle = 0$ $\nearrow \underline{x} \perp \underline{y}$
Clearly $\underline{x} \perp \underline{y} \Leftrightarrow \underline{y} \perp \underline{x}$. In particular $W^{\perp}$ and ${}^{\perp}W$ are the same.
Vectors $\underline{x_1}, \ldots, \underline{x_k}$ are <u>orthogonal</u> iff the $\underline{x_i} \neq \underline{0}$ and $\underline{x_i} \perp \underline{x_j}$ for $i \neq j$. Any such $\underline{x_1}, \ldots, \underline{x_k}$ is independent.
For if $\sum_{i=1}^{k} \lambda_i \underline{x_i} = \underline{0}$, then $\langle \underline{x_j}, \sum_{i=1}^{k} \lambda_i \underline{x_i} \rangle = \lambda_j \|\underline{x_j}\|^2 = 0$ and so $\lambda_j = 0$ as $\|\underline{x_j}\|^2 \neq 0$.

$\underline{e_1}, \ldots, e_k$ in an inner product space V are orthonormal if and only if $\langle \underline{e_i}, \underline{e_j} \rangle = \delta_{ij}$ $\forall i, j$. This is an orthogonal set/sequence normalised so its vectors are of unit length.
In Chapter 7 we saw that we can diagonalise symmetric/hermitian forms. If the form is +ve definite, the corresponding matrix is I.
Hence we have a basis $\underline{e_1}, \ldots, \underline{e_n}$, $\langle \underline{e_i}, \underline{e_j} \rangle = \delta_{ij}$
Thus any finite dimensional inner product space has an orthonormal basis.

Suppose we have an orthonormal basis $\underline{e_1}, \ldots, \underline{e_n}$ for V. Let $\underline{x} = \sum_{i=1}^{n} x_i \underline{e_i}$.
Then for any j, $\langle \underline{e_j}, \underline{x} \rangle = \sum_{i=1}^{n} x_i \langle \underline{e_j}, \underline{e_i} \rangle = x_j$

So $\underline{x} = \sum_{i=1}^{n} \langle \underline{e_i}, \underline{x} \rangle \underline{e_i}$

8.3 Orthogonal Projection

Suppose W is a subspace of an inner product space V and $\underline{v} \in V$
We seek "the foot of the perpendicular from $\underline{v}$ to W".

<u>Theorem</u> Let W be a finite dimensional subspace of an inner product space V. Then
1) $V = W \oplus W^{\perp}$
2) The map $\pi : V \to W$ (such that $\underline{v} = \underbrace{\pi(\underline{v})}_{\in W} + \underbrace{(\underline{v} - \pi(\underline{v}))}_{\in W^{\perp}}$) is an orthogonal projection.
3) $\pi(v)$ is vector in W closest to $\underline{v}$.

21/11/11

Linear Algebra (21)

Proof:

<u>Abstract view</u> $\langle , \rangle$ restricted to W is +ve definite and so non-singular. Hence $\underline{w}_0 \longmapsto (\underline{w} \mapsto \langle \underline{w}, \underline{w}_0 \rangle)$ is an isomorphism $W \to \bar{W}^*$. For $\underline{v} \in V$, $(\underline{w} \mapsto \langle \underline{w}, \underline{v} \rangle) \in \bar{W}^*$ and we have a linear map $V \to \bar{W}^*$. So $(\underline{w} \mapsto \langle \underline{w}, \underline{v} \rangle) = (\underline{w} \mapsto \langle \underline{w}, \pi(\underline{v}) \rangle)$ for a unique $\pi(\underline{v}) \in W$ depending linearly on $\underline{v}$. We deduce that $\langle \underline{w}, \underline{v} - \pi(\underline{v}) \rangle = 0 \ \forall \underline{w} \in W$ so $\underline{w} - \pi(\underline{v}) \in W^\perp$ and $\underline{v} = \pi(\underline{v}) + (\underline{v} - \pi(\underline{v}))$, so $V = W + W^\perp$

<u>But</u> if $\underline{w} \in W \cap W^\perp$ then $\langle \underline{w}, \underline{w} \rangle = 0$ so $\underline{w} = \underline{0}$.
Thus $W \cap W^\perp = \{\underline{0}\}$ and $V = W \oplus W^\perp$ (Not needed, see chapter 7)

π is a projection for e.g. $\pi(\underline{v}) - \pi^2(\underline{v}) \in W \cap W^\perp$, so $\pi^2 = \pi$ (projection)
π is an <u>orthogonal</u> projection because $\underline{v} - \pi(\underline{v}) \in W^\perp \ \forall \underline{v}$.

For 3) take $\underline{w} \in W$ and consider $\|\underline{v} - \underline{w}\|$

$$\begin{aligned}\|\underline{v} - \underline{w}\|^2 &= \|(\pi(\underline{v}) - \underline{w}) + (\underline{v} - \pi(\underline{v}))\|^2 \\ &= \|\pi(\underline{v}) - \underline{w}\|^2 + \|\underline{v} - \pi(\underline{v})\|^2\end{aligned}$$

because $(\pi(\underline{v}) - \underline{w}) \perp (\underline{v} - \pi(\underline{v}))$. This takes its minimum when $\underline{w} = \pi(\underline{v})$.

<u>Note</u> We used Pythagoras' Theorem. If $\underline{x}_1, \ldots, \underline{x}_k$ are pairwise orthogonal then $\|\underline{x}_1 + \ldots + \underline{x}_k\|^2 = \|\underline{x}_1\|^2 + \ldots + \|\underline{x}_k\|^2$

22/11/11

Linear Algebra (22)

Orthogonal Projections : Alternative view

Suppose W is a finite dimensional subspace of an inner product space V. Take an orthonormal basis $\underline{e_1}, \dots, \underline{e_m}$ for W.

Define $\pi(\underline{v}) = \sum_{i=1}^{m} \langle \underline{e_i}, \underline{v} \rangle \underline{e_i} \in W$. Clearly π is linear.

Also, for any j, $\langle \underline{e_j}, \underline{v} - \pi(\underline{v}) \rangle = \langle \underline{e_j}, \underline{v} \rangle - \sum_{i=1}^{m} \langle \underline{e_i}, \underline{v} \rangle \langle \underline{e_j}, \underline{e_i} \rangle$
$= \langle \underline{e_j}, \underline{v} \rangle - \langle \underline{e_j}, \underline{v} \rangle = 0$

Thus $\underline{v} - \pi(\underline{v}) \in W^{\perp}$ ($\perp$ to a basis for W and so to all of W)

Note that $\pi(\underline{e_j}) = \sum_{i=1}^{m} \langle \underline{e_i}, \underline{e_j} \rangle \underline{e_i} = \underline{e_j}$. So π is the identity on W and $\pi^2(\underline{v}) = \pi(\underline{v})$

Finally if $\underline{w} \in W$, then $\|\underline{v} - \underline{w}\|^2 = \|(\pi(\underline{v}) - \underline{w}) + (\underline{v} - \pi(\underline{v}))\|^2$
$= \|\pi(\underline{v}) - w\|^2 + \|\underline{v} - \pi(\underline{v})\|^2$ by Pythagoras.

This is minimum for $\underline{w} = \pi(\underline{v})$

Bessell's Inequality. Given $e_1, \dots, e_n$ an orthonormal sequence, we have $\underline{x} = \left(\sum_{i}^{m} \langle \underline{e_i}, \underline{x} \rangle \underline{e_i} \right) + \left(\underline{x} - \sum \langle \underline{e_i}, \underline{x} \rangle \underline{e_i} \right)$

By Pythagoras, $\|\underline{x}\|^2 = \sum_{i=1}^{m} |\langle \underline{e_i}, \underline{x} \rangle|^2 + \|$ " $\|^2$

and so $\|\underline{x}\|^2 \geq \sum_{i=1}^{m} |\langle \underline{e_i}, \underline{x} \rangle|^2$

Cauchy Schwarz. By orthogonal projection. Suppose $\underline{y} \neq 0$ and write $\underline{x}$ in $\langle \underline{y} \rangle \oplus \langle \underline{y} \rangle^{\perp}$. So

$\underline{x} = \frac{\langle \underline{y}, \underline{x} \rangle}{\|\underline{y}\|^2} \underline{y} + \left(\underline{x} - \frac{\langle \underline{y}, \underline{x} \rangle}{\|\underline{y}\|^2} \underline{y} \right)$

So $\|\underline{x}\|^2 \geq \frac{|\langle \underline{y}, \underline{x} \rangle|^2 \|\underline{y}\|^2}{\|\underline{y}\|^4}$ and the result follows.

(Equality if $\underline{x} = \frac{\langle \underline{y}, \underline{x} \rangle}{\|\underline{y}\|^2} \underline{y}$

8.4 Gram-Schmidt Orthogonalisation

Given an an independent sequence $\underline{x_1}, \dots, \underline{x_m}$ or $\underline{x_1}, \underline{x_2}, \dots$ in an inner product space V, we define an orthonormal sequence $\underline{e_1}, \dots, \underline{e_m}$ or $\underline{e_1}, \underline{e_2}, \dots$ (together with an auxiliary sequence $\underline{e_i}'$) inductively as follows:

1. $\underline{e_1}' = \underline{x_1}$ and $\underline{e_1} = \frac{\underline{e_1}'}{\|\underline{e_1}'\|}$
2. $\underline{e_{r+1}}' = \underline{x_{r+1}} - \sum_{1}^{r} \langle \underline{e_i}, \underline{x_{r+1}} \rangle \underline{e_i}$, the component of $\underline{x_{r+1}}$ in $\langle \underline{e_1}, \dots, \underline{e_r} \rangle^{\perp}$. Then $\underline{e_{r+1}} = \frac{\underline{e_{r+1}}'}{\|\underline{e_{r+1}}'\|}$

Note inductively that $\langle \underline{e_1}, \dots, \underline{e_k} \rangle = \langle \underline{e_1}', \dots, \underline{e_k}' \rangle = \langle \underline{x_1}, \dots, \underline{x_k} \rangle$

(The last equality follows as in 8·3, $\langle W, \underline{v} \rangle = \langle W, \underline{v} - \pi(\underline{v}) \rangle$)
From this it follows that the $\underline{x}_i$ are independent, each $\underline{e}_k' \neq 0$ and the definition makes sense.
This produces an orthonormal sequence $\underline{e}_1, \ldots, \underline{e}_m$ or $\underline{e}_1, \underline{e}_2, \ldots$ with the property that $\langle \underline{x}_1, \ldots, \underline{x}_k \rangle = \langle \underline{e}_1, \ldots, \underline{e}_k \rangle$

8·5 The Adjoint of an Endomorphism

Proposition Let V be a finite dimensional inner product space. Then for every endomorphism α of V, there is a unique endomorphism α^* of V with the property that $\langle \alpha^*(\underline{x}), \underline{y} \rangle = \langle \underline{x}, \alpha(\underline{y}) \rangle \ \forall \underline{x}, \underline{y} \in V$

Proof

For $\underline{x}$ fixed, the map $\underline{y} \mapsto \langle \underline{x}, \alpha(\underline{y}) \rangle \in V^*$. The inner product provides an isomorphism $\bar{V} \to V^*$ and so there is a unique vector $\alpha^*(\underline{x}) \in V$ with $\langle \alpha^*(\underline{x}), \underline{y} \rangle = \langle \underline{x}, \alpha(\underline{y}) \rangle$

It remains to show that $\alpha^*(\underline{x})$ is linear in $\underline{x}$.

(Abstractly, the map $\underline{x} \mapsto (\underline{y} \mapsto \langle \underline{x}, \alpha(\underline{y}) \rangle)$ is linear $\bar{V} \to V^*$ and we compose with a linear isomorphism $V^* \xrightarrow{\cong} \bar{V}$. It is then linear $\bar{V} \to \bar{V}$ and so linear $V \to V$).

OR $\langle \alpha^*(\lambda\underline{x} + \mu\underline{y}), \underline{z} \rangle = \langle \lambda\underline{x} + \mu\underline{y}, \alpha(\underline{z}) \rangle$

$= \lambda \langle \underline{x}, \alpha(\underline{z}) \rangle + \mu \langle \underline{y}, \alpha(\underline{z}) \rangle = \lambda \langle \alpha^*(\underline{x}), \underline{z} \rangle + \mu \langle \alpha^*(\underline{y}), \underline{z} \rangle$

$= \langle \lambda\alpha^*(\underline{x}) + \mu\alpha^*(\underline{y}), \underline{z} \rangle$

As this holds $\forall \underline{z}$, $\alpha^*(\lambda\underline{x} + \mu\underline{y}) = \lambda\alpha^*(\underline{x}) + \mu\alpha^*(\underline{y})$

Define then α^* to be the endomorphism adjoint to α.

Easy consequences of the proposition:

1. $\alpha^{**} = \alpha$ 2. $(\beta\alpha)^* = \alpha^*\beta^*$

Application An endomorphism α is orthogonal/unitary iff $\langle \alpha(\underline{x}), \alpha(\underline{y}) \rangle = \langle \underline{x}, \underline{y} \rangle \ \forall \underline{x}, \underline{y}$ (α preserves the inner product)

Note that $\langle \alpha(\underline{x}), \alpha(\underline{y}) \rangle = \langle \underline{x}, \underline{y} \rangle \ \forall \underline{x}, \underline{y}$

$\iff \langle \alpha^*\alpha(\underline{x}), \underline{y} \rangle = \langle \underline{x}, \underline{y} \rangle \ \forall \underline{x}, \underline{y}$

$\iff \alpha^*\alpha(\underline{x}) = \underline{x} \ \forall \underline{x} \iff \alpha^*\alpha = I$

So an orthogonal or unitary inverse is one whose adjoint is its inverse.

27/11/11

Linear Algebra (22)

With inner product spaces we are interested in the matrices for endomorphisms with respect to orthonormal bases.

Suppose α an endomorphism of V has matrix $A = (a_{ij})$ with respect to an orthonormal basis $\underline{e_1}, \ldots, \underline{e_n}$ and that α^* has matrix $\tilde{A} = (\tilde{a}_{ij})$

Then:

$$\langle \alpha^*(\underline{e_i}), \underline{e_j} \rangle = \langle \underline{e_i}, \alpha(\underline{e_j}) \rangle$$

$$\tilde{a}_{ji} = \langle \sum_s \tilde{a}_{si} \underline{e_s}, \underline{e_j} \rangle \qquad \langle \underline{e_i}, \sum_r a_{rj} \underline{e_r} \rangle = a_{ij}$$

Thus, $\tilde{a}_{ji} = a_{ij}^*$, and $\tilde{A} = A^\dagger$

Consequence

An orthogonal/unitary endomorphism has a matrix A with $A^\dagger A = I$ with respect to any orthonormal basis.

Note that α has this property just where it carries orthonormal bases to orthonormal bases.

Observation Suppose P is the change of basis matrix for $(\underline{e_i}), (\underline{e_i}')$, both orthonormal. Then P is the matrix with respect to $(\underline{e_i}')$ of the map taking $\underline{e_i}' \mapsto \underline{e_i}$. It follows that $P^\dagger P = I$

23/11/11

Linear Algebra (23)

8.6 Symmetric and Hermitian Endomorphisms

Definition Let V be an inner product space, finite dimensional. An endomorphism α of V is symmetric (real), Hermitian (complex) just when $\alpha = \alpha^*$. (Recall: the adjoint endomorphism satisfies $\langle \alpha^*(\underline{x}), \underline{y} \rangle = \langle \underline{x}, \alpha($

Note α is symmetric / Hermitian just when its matrix A with respect to an orthonormal basis is symmetric / Hermitian $A = A^*$.

Proposition

1. Let α be an Hermitian endomorphism. Then the eigenvalues of α are real.
2. Let α be a symmetric endomorphism in a real vector space. Then $\chi_\alpha(t)$ has only real roots (i.e. it factorises into real linear factors over $\mathbb{C}$

Proof

1. Let $\alpha(\underline{e}) = \lambda \underline{e}$ with $\underline{e} \neq 0$. Then:

$\bar{\lambda} \langle \underline{e}, \underline{e} \rangle = \langle \lambda \underline{e}, \underline{e} \rangle = \langle \alpha(\underline{e}), \underline{e} \rangle = \langle \underline{e}, \alpha(\underline{e}) \rangle = \langle \underline{e}, \lambda \underline{e} \rangle = \lambda \langle \underline{e}, \underline{e}$

So $\lambda = \lambda^*$ as $\langle \underline{e}, \underline{e} \rangle \neq 0$

2. α has a symmetric matrix A with respect to an orthonormal basis. Use A to give an endomorphism of $\mathbb{C}^n$. That endomorphism is Hermitian and we apply 1 to it. But $\chi_\alpha(t) = \chi_A(t)$

Theorem

Let α be a symmetric / Hermitian endomorphism of a real/complex finite dimensional inner product space V. Then there is an orthonormal basis $\underline{e}_1, \dots, \underline{e}_n$ for V consisting of eigenvectors for α.

Proof

1 By induction on $\dim V$. Both $\dim V = 0$, 1 are trivial.

Induction step

There is an eigenvalue and so we can pick $\underline{e}$, an eigenvector.

Set $\underline{e_1} = \frac{\underline{e}}{\|\underline{e}\|}$, an eigenvector of length 1. Now $V = \langle \underline{e_1} \rangle \oplus \langle \underline{e_1} \rangle^\perp$

CLAIM: $\alpha|_{\langle \underline{e_1}\rangle^\perp} : \langle \underline{e_1}\rangle^\perp \to \langle \underline{e_1}\rangle^\perp$

Proof Suppose $\langle \underline{e_1}, \underline{x} \rangle = 0$. Then $\langle \underline{e_1}, \alpha(\underline{x}) \rangle = \langle \alpha(\underline{e_1}), \underline{x} \rangle$

$= \langle \lambda \underline{e_1}, \underline{x} \rangle = \bar{\lambda} \langle \underline{e}, \underline{x} \rangle = 0$ $\quad (\lambda = \lambda^* \text{ here})$

Now $\alpha|_{\langle \underline{e_1}\rangle^\perp}$ is itself Hermitian, so by our induction hypothesis, there is an orthonormal basis $\underline{e_2}, \dots, \underline{e_n}$ consisting of eigenvectors.

Then $\underline{e_1}, \dots, \underline{e_n}$ is an orthonormal basis as required.

Matrix Interpretation Let A be a symmetric / Hermitian matrix thought of as a symmetric/Hermitian endomorphism of $\mathbb{R}^n / \mathbb{C}^n$ with the standard inner product. The theorem gives P such that PAP^{-1} is diagon

[Aside: P has the property that $\underline{e_k} = \sum_i p_{ik} \underline{e_i}'$. So the new basis is the columns of $\hat{P} = P^{-1}$]

"So" $\hat{P}$ and so P are orthogonal / unitary $\quad \begin{matrix} \hat{P}\hat{P}^\dagger = I = \hat{P}^\dagger \hat{P} \\ PP^\dagger = I = P^\dagger P \end{matrix}$

Now $PAP^{-1} = \hat{P}^\dagger A \hat{P}$ - the change of basis formula for a form !!

8-7 Simultaneous Diagonalisation of forms

Lemma

Let V be an inner product space. Then for any symmetric/Hermitian form, γ on V, there is a unique symmetric/Hermitian endomorphism α such that $\langle \underline{x}, \alpha(\underline{y}) \rangle = \gamma(\underline{x}, \underline{y})$

23/11/11 Linear Algebra (23)

Proof

Fix $\underline{y}$. The map $\underline{x} \mapsto \gamma(\underline{x}, \underline{y}) \in \bar{V}^*$, So there is a unique $\alpha(\underline{y})$ with $\langle \underline{x}, \alpha(\underline{y}) \rangle = \gamma(\underline{x}, \underline{y})\ \forall \underline{x} \in V$

α is Linear:

EITHER $V \to \bar{V}^* \xrightarrow{\alpha} V$ is a composite of linear maps
$\underline{y} \mapsto (\underline{x} \mapsto \gamma(\underline{x}, \underline{y})) \mapsto \alpha(\underline{y})$

OR $\langle \underline{x}, \alpha(\lambda \underline{y} + \mu \underline{z}) \rangle = \gamma(\underline{x}, \lambda \underline{y} + \mu \underline{z}) = \bar{\lambda} \gamma(\underline{x}, \underline{y}) + \bar{\mu} \gamma(\underline{x}, \underline{z})$
$= \bar{\lambda} \langle \underline{x}, \alpha(\underline{y}) \rangle + \bar{\mu} \langle \underline{x}, \alpha(\underline{z}) \rangle = \langle \underline{x}, \lambda \alpha(\underline{y}) + \mu \alpha(\underline{z}) \rangle$

This holds for all $\underline{x}$. So $\alpha(\lambda \underline{y} + \mu \underline{z}) = \lambda \alpha(\underline{y}) + \mu \alpha(\underline{z})$

Finally $\langle \underline{x}, \alpha(\underline{y}) \rangle = \gamma(\underline{x}, \underline{y}) = \overline{\gamma(\underline{y}, \underline{x})} = \overline{\langle \underline{y}, \alpha(\underline{x}) \rangle} = \langle \alpha(\underline{x}), \underline{y} \rangle$

and so α is symmetric/Hermitian.

Theorem

Suppose β, γ are symmetric/Hermitian forms on a real/complex finite dimensional vector space. Further, suppose β is positive definite. Then there is a basis with respect to which the form β has the matrix I the identity, and γ has a diagonal matrix.

Proof

Regard β as an inner product on the vector space V. Take α to be the symmetric/Hermitian endomorphism representing γ, in the sense that $\gamma(\underline{x}, \underline{y}) = \beta(\underline{x}, \alpha(\underline{y}))$

Take $\underline{e_1}, \dots, \underline{e_n}$ an orthonormal (with respect to β) basis of eigenvectors of α, say with $\alpha(\underline{e_i}) = \lambda_i \underline{e_i}$. Then

- $\beta(\underline{e_i}, \underline{e_j}) = \delta_{ij}$, and β has matrix I
- $\gamma(\underline{e_i}, \underline{e_j}) = \beta(\underline{e_i}, \alpha(\underline{e_j})) = \beta(\underline{e_i}, \lambda_j \underline{e_j}) = \lambda_j \delta_{ij}$ and γ has a diagonal matrix.

3

Appendix

Suppose there is an orthonormal basis of eigenvectors of α. Then α has matrix $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = A$ and α^* has matrix $\begin{pmatrix} \lambda_1^* & & 0 \\ & \ddots & \\ 0 & & \lambda_n^* \end{pmatrix} = A^\dagger$.

Evidently $AA^\dagger = A^\dagger A$ and so $\alpha\alpha^* = \alpha^*\alpha$.

Moreover, observe that if $\alpha\alpha^* = \alpha^*\alpha$ then $\underline{e}$ is an eigenvector for α with eigenvalue λ iff it is also an eigenvector of α^* with eigenvalue λ^*.

Proof $(\alpha - \lambda I)\underline{e} = 0$ iff $\langle (\alpha-\lambda I)\underline{e}, (\alpha-\lambda I)\underline{e}\rangle = 0$

iff $\langle (\alpha^* - \lambda^* I)(\alpha-\lambda I)\underline{e}, \underline{e}\rangle = 0$ ~~iff $\langle (\alpha-\lambda I)\underline{e}, (\alpha^*-\lambda^* I)\underline{e}\rangle = 0$~~

iff $\langle (\alpha-\lambda I)(\alpha^*-\lambda^* I)\underline{e}, \underline{e}\rangle = 0$ iff $\langle (\alpha^*-\lambda^* I)\underline{e}, (\alpha^*-\lambda^* I)\underline{e}\rangle = 0$

iff $(\alpha^* - \lambda^* I)\underline{e} = 0$

Note in the Theorem of 8·6, the key is the claim $\alpha|_{\langle e_1\rangle^\perp} : \langle e_1\rangle^\perp \to \langle e_1\rangle^\perp$

This still goes through. Suppose $\langle \underline{e}_1, \underline{x}\rangle = 0$. Then

$\langle \underline{e}_1, \alpha(\underline{x})\rangle = \langle \alpha^*(\underline{e}_1), \underline{x}\rangle = \langle \lambda^* \underline{e}_1, \underline{x}\rangle = \lambda\langle \underline{e}_1, \underline{x}\rangle = 0$

So if $\alpha^*\alpha = \alpha\alpha^*$ we can diagonalise in the Main Theorem.

Special Case α is orthogonal/unitary, $\alpha^*\alpha = I = \alpha\alpha^*$

This is the case when we diagonalise with entries of all absolute value 1, 25/11/11 Linear Algebra (24)

8.8 Simultaneous Diagonalisation

Matrix Interpretation of 8.7

We have matrices B (positive definite), and C (both are Hermitian) and we seek a (non-singular) matrix Q with:

$$Q^*BQ = I \quad , \quad Q^*CQ = D, \text{ diagonal } \left(= \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}\right)$$

[Recall from 7.2 that here $Q = P^{-1} = \hat{P}$ where P is the change of basis matrix, and that P appears in the new formula]

Observation ① The columns $\underline{q_1}, \ldots, \underline{q_n}$ of Q form an orthonormal (with respect to B) basis of eigenvectors of some representing endomorphism, A. A must satisfy $\underline{x}^* C \underline{y} = \underline{x}^* B(A\underline{y})$, that is $C = BA$ or $A = B^{-1}C$.

Note that $\underline{x}$ is an eigenvector for A with eigenvalue λ iff $A\underline{x} = \lambda\underline{x}$ iff $C\underline{x} = \lambda B\underline{x}$

Note $Q^*BQ = I$ says that $\underline{q_i}^* B \underline{q_j} = \delta_{ij}$

i.e. it says that the $\underline{q_i}$ are orthonormal with respect to B.

$Q^*CQ = D$ says that $\underline{q_i}^* C \underline{q_j} = d_j \delta_{ij} = \underline{q_i}^* d_j B \underline{q_j}$

That holds for all $\underline{q_i}$ iff $C\underline{q_j} = d_j B \underline{q_j}$ iff the $\underline{q_j}$ are eigenvectors for A.

Observation ② We have a method for finding Q.

λ is an eigenvalue of A iff $C - \lambda B$ is singular iff λ is a root of $\det(C - tB) = 0$.

With luck, all the eigenspaces are 1 dimensional and we find the $\underline{q_i}$ by taking eigenvectors with appropriate scaling.

Otherwise, we take an orthonormal (with respect to B) basis of the eigenspaces for A.

The quadratic forms will be $\sum_{i=1}^{n} \left| \sum_{j=1}^{n} p_{ij} x_j \right|^2$ and $\sum_{i=1}^{n} d_i \left| \sum_{j=1}^{n} p_{ij} x_j \right|^2$

Observation (3)

Outline of a purely matrix treatment. Suppose $\underline{u} \in \ker(C - \lambda B)$ and $\underline{v} \in \ker(C - \mu B)$, $\lambda \neq \mu$.

$$\underline{v}^* C \underline{u} = \underline{v}^* \lambda B \underline{u} = \lambda \underline{v}^* B \underline{u}$$

$$\underline{u}^* C \underline{v} = \underline{u}^* \mu B \underline{v} = \mu \underline{u}^* B \underline{v} = \mu \underline{v}^* B \underline{u}, \text{ Thus } \underline{v}^* B \underline{u} = 0$$

So it follows that if $\lambda_1, \ldots, \lambda_k$ are the distinct roots of $\det(C - tB)$ then we have a direct sum $\ker(C - \lambda_1 B) \oplus \ldots \oplus \ker(C - \lambda_k B)$

Why is this all of $\mathbb{C}^n$? We can write:

$\mathbb{C}^n = \left(\bigoplus_{i=1}^{k} \ker(C - \lambda_i B) \right) \oplus W$, where W is the $\perp$ complement with respect to B. Then we see easily that $A = B^{-1} C : W \to W$ and so if $W \neq \{\underline{0}\}$ there would exist eigenvalues and so eigenvectors ※

Worked Example

$$2x^2 + 2y^2 \qquad x^2 + 6xy + y^2$$

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

$$\det(C - tB) = \det \begin{pmatrix} 1-2t & 3 \\ 3 & 1-2t \end{pmatrix} = 1 - 4t + 4t^2 - 9$$

$$= 4(t^2 - t - 2) = 4(t-2)(t+1)$$

We know the final forms will be $X^2 + Y^2$, $2X^2 - Y^2$

$\lambda = 2$, $\begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}$. So $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector, $\| \cdot \|^2 = 4$

$\lambda = -1$, $\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$ so $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector, $\| \cdot \|^2 = 4$

So $Q = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

The forms are $(x+y)^2 + (x-y)^2$ and $2(x+y)^2 - (x-y)^2$

<u>Remark</u>

In Chapter 7, we thought of rank and signature in terms of $\pm$ve definite, semi-definite or zero spaces. Now, we see that to find the rank and signature of C, it suffices to find the eigenvalues with multiplicities. The signature is the difference between the number of positives and negatives.

<u>Appendix</u>

Let V be a finite dimensional inner product space and α a Hermitian (symmetric) endomorphism. Then, for any eigenvalue λ, the dimension of $\ker(\alpha - \lambda I)$ is the algebraic multiplicity of λ (that is, the degree to which it appears in $\chi_\alpha(t)$).

Why? $V = \ker(\alpha - \lambda I) \oplus (\ker(\alpha - \lambda I))^\perp = U \oplus W$

Clearly $\alpha : U \to U$ and it follows that $\alpha : W \to W$.

Hence the matrix of α has the form $\left(\begin{array}{c|c} A_U & 0 \\ \hline 0 & A_W \end{array}\right)$ and

$\chi_\alpha(t) = \chi_{A_U}(t)\,\chi_{A_W}(t) = (\lambda - t)^k \chi_{A_W}(t)$

But $(t-\lambda) \nmid \chi_{A_W}(t)$ otherwise λ would be an eigenvalue for $\alpha|_W$ and there would be an eigenvector for λ in W ※