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 \( V \times V \rightarrow V \) \( \cdot (u, v) \mapsto u + v \)
   - zero \( 0 \in V \) (additive identity)
   - inverse \( V \rightarrow V \) \( \cdot u \mapsto -u \)

b) Scalar Multiplication \( F \times V \rightarrow V \) \( \cdot (\lambda, u) \mapsto \lambda u \)
   satisfying
   - 1 \( \cdot u = u \)
   - \( \lambda (\mu u) = (\lambda \mu) u \) (action)
   - \( (\lambda + \mu) u = \lambda u + \mu u \)
   - \( \lambda (u + v) = \lambda u + \lambda v \) (distributivity / bilinearity)

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

Indicative Consequences

0 \( \cdot u = 0 \) (for \( 0 \cdot u + 0 \cdot u = (0 + 0) \cdot u = 0 \cdot u \))
(-1) \( \cdot u = -u \) (for \( u + (\lambda - 1) \cdot u = (\lambda - 1) \cdot u = 0 \cdot u = 0 \))

Finite linear combinations \( \sum_{i=1}^{n} \lambda_i x_i \) make sense and can be manipulated as we would expect.

(N.B. \( \sum_{i=1}^{n} \lambda_i x_i = 0 \) and expressions make sense in that case.)

Canonical Example: The space \( F^n \) of column vectors

\( x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \) under coordinate-wise operations. Note that there are \( n \) vectors \( u_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \) such that any \( x \) as above
can be written uniquely as a finite linear combination \( x = \sum_{i=1}^{n} \lambda_i v_i \). 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 \( F^I = \{ f : I \rightarrow F \} \) form a vector space under pointwise operations:

\[
(f + g)(i) = f(i) + g(i), \quad (\lambda f)(i) = \lambda f(i)
\]

Note that the cardinality of \( I \) is not the dimension; \( F^I \) 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. Note that it is not possible to give an explicit basis for \( F^n \) over \( F \) or \( \mathbb{R} \) over \( \mathbb{Q} \).

1.2 Subspaces

A subset \( U \) of a vector space \( V \) is a subspace if \( U \subseteq V \) just when \( 0 \in U \), \( x, y \in U \Rightarrow x + y \in U \), and \( \lambda x \in U \Rightarrow \lambda x \in U \).

In other words, \( U \) is closed under the basic operations and \( x \) under finite linear combinations.

Fact: \( U \) is a subspace of \( V \) if and only if \( U \) is non-empty and \( x, y \in U \Rightarrow \lambda x + \mu y \in U \) for all \( \lambda, \mu \).
This is often the definition as it is quicker to check. We need $U \neq \emptyset$, then we pick $x \in U$ and find that $0 = 0x \in U$.

**Examples**

Let $X$ be a space like $[0,1]$, $\mathbb{R}$, $\mathbb{C}$. Then the collection of functions $X \rightarrow \mathbb{R}$ (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 $F_\mathbb{F} \rightarrow F_\mathbb{F}$? (1)

### 1.3 Linear Maps

Let $U, V$ be vector spaces. A map $\alpha : U \rightarrow V$ is linear if and only if

$\alpha(x + y) = \alpha(x) + \alpha(y), \quad \alpha(\lambda x) = \lambda \alpha(x) \quad \text{i.e.} \ \alpha \text{ is an abelian group homomorphism preserving scalar multiplication}$

Equivalently, $\alpha$ is linear just when it preserves finite linear combinations:

$\alpha \left( \sum \lambda_i x_i \right) = \sum \lambda_i \alpha(x_i) \quad \text{Equivalently, just when } \alpha(\lambda x + \mu y) = \lambda \alpha(x) + \mu \alpha(y)$

**Canonical Example**

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

If $\mathbf{A}$ is an $m \times n$ matrix then $x \rightarrow Ax$ is linear as a map $F^n \rightarrow F^m$

$Ax = \begin{pmatrix} a_1^1 & \cdots & a_1^n \\ \vdots & \ddots & \vdots \\ a_m^1 & \cdots & a_m^n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and we just check that

$A(\lambda x + \mu y) = \lambda Ax + \mu Ay$
2. If \( \alpha \) is linear, note the following:

Any \( \mathbf{x} \in \mathbb{F}^n \) is written \( \mathbf{x} = \sum x_i \mathbf{u}_i \) and so by linearity

\[
\alpha(\mathbf{x}) = \sum x_i \alpha(\mathbf{u}_i)
\]

Let \( \mathbf{A} = (\alpha(\mathbf{u}_1) \ldots \alpha(\mathbf{u}_n)) \) with columns

\[
\alpha(\mathbf{u}_i) = \begin{pmatrix}
\mathbf{v}_i \\
\cdot \\
\cdot \\
\mathbf{w}_i
\end{pmatrix}
\]

and then

\[
\alpha(\mathbf{x}) = \mathbf{A} \mathbf{x} = \begin{pmatrix}
\sum v_i x_i \\
\cdot \\
\cdot \\
\sum w_i x_i
\end{pmatrix}
\]

\( \mathbf{A} \) takes the standard basis vectors to the columns of \( \mathbf{A} \).
Linear Algebra

Examples of Linear Maps from Calculus

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

2. Integration
   \[ \int_a^x \frac{d}{dt} dt : C^0(\mathbb{R}) \to C'(\mathbb{R}) \]
   or \( \int_a^x f(t) dt : C^0(\mathbb{R}) \to \mathbb{R} \)

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 \( x, y \in U \).
\[ \beta \alpha (\lambda x + \mu y) = \beta (\alpha (\lambda x + \mu y)) = \beta (\alpha (\lambda x) + \alpha (\mu y)) = \lambda \beta (\alpha (x)) + \mu \beta (\alpha (y)) \]

In the canonical example, let \( \alpha : F^n \to F^m \) be \( \alpha (x) = Ax \)
and \( \beta : F^m \to F^p \) be \( \beta (y) = By \).

Then \( \beta \alpha : F^n \to F^p \) is \( \beta \alpha (x) = (BA)x \), for \( \beta \alpha (yi) = \beta \left( \begin{array}{c} a_{1i} \\ \vdots \\ a_{mi} \end{array} \right) = (\begin{array}{c} b_{11} & b_{12} & \cdots & b_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{11} & b_{12} & \cdots & b_{1m} \end{array}) \left( \begin{array}{c} \alpha (xi) \\ \vdots \\ \alpha (mi) \end{array} \right) \), the \( j \)th column of \( BA \).

Proposition

Let \( U, V \) be vector spaces. Then \( 0 : U \to V, x \mapsto 0 \) is linear,
and if \( \alpha, \beta : U \to V \) are linear, \( \alpha + \mu \beta \), with proof as before
Write \( L(U, V) \) for the set of linear maps from \( U \) to \( V \).

**Corollary**

\[ L(U, V) \leq V^U, \] the space of all maps \( U \) to \( V \), and is in particular a vector space.

In the canonical example, \( L(F^n, F^m) \) 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**

- The kernel, \( \ker \alpha = \{ u \in U \mid \alpha(u) = 0 \} \)
- The image of \( \alpha \), \( \text{im} \alpha = \{ v \in V \mid v = \alpha(u) \text{ for some } u \in U \} \)

**Proposition**

If \( \alpha : U \to V \) is linear, then \( \ker \alpha \leq U \), \( \text{im} \alpha \leq V \)

**Proof**

Let \( x, y \in \ker \alpha \), then \( \alpha(\lambda x + \mu y) = \lambda \alpha(x) + \mu \alpha(y) = \lambda 0 + \mu 0 = 0 \), so \( \lambda x + \mu y \in \ker \alpha \). Moreover, \( \alpha(0) = 0 \), \( 0 \in \ker \alpha \), thus \( \ker \alpha \leq U \).

Let \( v, w \in \text{im} \alpha \). Take \( x, y \in U \) with \( \alpha(x) = v \), \( \alpha(y) = w \). Then \( \alpha(\lambda x + \mu y) = \lambda \alpha(x) + \mu \alpha(y) = \lambda v + \mu w \) which is in \( \text{im} \alpha \). Also \( \alpha(0) = 0 \), so \( 0 \in \text{im} \alpha \) and \( \text{im} \alpha \leq V \).
Examples of Kernels

1. The set of solutions \((x_1, \ldots, x_n)\) of the linear equations
   \[
   a_{11}x_1 + \ldots + a_{1n}x_n = 0
   \]
   \[
   \vdots
   \]
   \[
   a_{m1}x_1 + \ldots + a_{mn}x_n = 0
   \]
   is the kernel of the map \(x \mapsto Ax\), 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 chain properties, \(a_n D^n + a_{n-1} D^{n-1} + \ldots + a_1 D + a_0 I : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})\) is linear.
   In fact, this is still so if the \(a_i\) are \(C^\infty(\mathbb{R})\) functions of \(t\).
   Then the solutions to the differential equation
   \[
   a_n \frac{d^2x}{dt^2} + \ldots + a_0 x = 0
   \]
   form the kernel of the linear map.
   In the constant coefficient case, this kernel has dimension \(n\).

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

   Clearly \(\alpha : U \to V\) is injective \(\iff\) \(\ker \alpha = \{0\}\)

   Proposition

   \(\alpha : U \to V\) is injective \(\iff\) \(\ker \alpha = \{0\}\)

   Proof:

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

   (\(\Leftarrow\)) If \(\alpha(x) = \alpha(y)\), then \(\alpha(x-y) = 0\), \(x-y \in \ker \alpha\)
   \(\Rightarrow\) then \(x - y = 0\), \(x = y\), so \(\alpha\) is injective.
Isomorphisms

An isomorphism $\alpha : U \to V$ of vector spaces is a linear map with
inverse $\alpha^{-1} : V \to U$ (also linear).
(Then $\alpha \circ \alpha^{-1} = I_V$, $\alpha^{-1} \circ \alpha = I_U$.)

Proposition

Suppose $\alpha : U \to V$ is a bijective linear map. Then the inverse $\alpha^{-1} : V \to U$
is linear (and so $\alpha$ is an isomorphism).

Proof

Take $u, v \in V$ and let $x = \alpha(u)$, $y = \alpha(v)$ so that $\alpha(x) = u$, $\alpha(y) = v$.
Then $\alpha^{-1}(\lambda x + \mu y) = \alpha^{-1}(\lambda \alpha(x) + \mu \alpha(y)) = \lambda \alpha^{-1}(x) + \mu \alpha^{-1}(y)$

Corollary

$\alpha : U \to V$ is an isomorphism $\iff$ in $\alpha = V$, ker $\alpha = \{0\}$.
**Linear Algebra**

**The Isomorphism Theorem**

**Quotients**

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

The quotient map $q : V \rightarrow V/W$ is linear. $q : V \rightarrow V + W$

$\ker q = \{ v \in V \mid v + W = W \} = W$

$\text{im } q = \frac{V}{W}$

The quotient map $q : V \rightarrow V/W$ is linear. $q : V \rightarrow V + W$

$\ker q = \{ v \in V \mid v + W = W \} = W$

$\text{im } q = \frac{V}{W}$

**Isomorphism Theorem**

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

\[ \begin{array}{ccc}
U & \xrightarrow{\alpha} & V \\
\downarrow \alpha & & \uparrow \text{inclusion} \\
U/\ker \alpha & \rightarrow & \text{Im } \alpha \\
\end{array} \]

What is $\overline{\alpha}$? $\overline{\alpha} \left( u + \ker \alpha \right) = \alpha(u)$

$u + \ker \alpha = u' + \ker \alpha \iff u - u' \in \ker \alpha \iff \alpha(u - u') = 0$

$\iff \alpha(u) = \alpha(u') \iff \overline{\alpha}(u + \ker \alpha) = \overline{\alpha}(u' + \ker \alpha)$

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

This is evidently injective and linear, so $\overline{\alpha}$ is an isomorphism.
1.6 Operations on Spaces

**Proposition**

Let $U, V \subseteq W$. Then:

1) $U + V \subseteq W$
2) $U \cap V \subseteq W$

where $U + V = \{ u + v \mid u \in U, v \in V \}$

**Proof**

1) $0 \in U, 0 \in V$, so $0 \in U + V$. Take $x, y \in U + V$. Then, $x \in U, y \in V$. Let $x = u + v$ with $u \in U, v \in V$. Then $x = (u + v) + (0 + 0)$. So $x \in U + V$.

2) $0 = 0 + 0 \in U + V$. Let $x = u + v$, $y = u' + v'$ with $u, u' \in U$, $v, v' \in V$. Then $x + y = (u + u') + (v + v')$. So $x + y \in U + V$.

**Remark**

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

**Proposition**

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

**Proof**

(\vDash) Suppose $w = u + v = u' + v'$. Then $u - u' = v' - v \in U \cap V$.

So $u - u = v - v = 0 \Rightarrow u = u', v = v'$.
Linear Algebra 3

(\leftarrow) \text{Take } w \in U \cap V, \ w = w + 0 = 0 + w \text{ are two expressions. So by uniqueness, } w = 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 \( N = u_1 \leq W \) and \( u_1 + \ldots + u_n \leq W \).

\( u_1 + \ldots + u_n \) is a direct sum \( \iff \) every \( w \in u_1 + \ldots + u_n \) has a unique representation as \( w = u_1 + \ldots + u_n \), \( u_i \in U_i \).

N.B. \( u_1 + \ldots + u_n = (u_1 \oplus (u_2 \oplus (u_3 \oplus \ldots \oplus u_n))) \).

**External Direct Sum**

Let \( U, V \) be vector spaces. We define \( U \oplus V \) by

\[
\{(u, v) \mid u \in U, v \in V\}
\]

under pointwise operations:

\[
(u, v) + (u', v') = (u + u', v + v'), \quad \lambda(u, v) = (\lambda u, \lambda v)
\]

Below are all linear maps:

\[
y \mapsto (y, 0) \quad U \rightarrow V \quad \text{and} \quad (0, y) \mapsto y \quad \text{and} \quad (u, v) \mapsto u + v
\]

\[
\text{Suppose } U, V \leq W, \text{ Then there is a map:}
\]

\[
U \oplus V \rightarrow U + V \leq W, \quad (u, v) \mapsto u + v
\]

\[
\text{This is evidently surjective.}
\]

\[
\text{What is the kernel?}
\]
The kernel is \( \{(u, v) \mid u + v = 0\} = \{ (w, -w) \mid w \in U \cup V \} \).

The kernel is \( \{(0, 0)\} \Leftrightarrow U \cup V = \{ 0 \} \), and is an isomorphism.

In this case, i.e., when the internal direct sum exists, it is the same as the external direct sum.
Dimension

Linear Independence and Spanning

Definition

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

\[ \langle \mathbf{x}_1, \ldots, \mathbf{x}_n \rangle = \{ \sum \lambda_i \mathbf{x}_i \mid \lambda_i \in F \} \]

the set of linear combinations of the \( \mathbf{x}_i \). We say that \( \mathbf{x}_1, \ldots, \mathbf{x}_n \) span \( V \) just when \( \langle \mathbf{x}_1, \ldots, \mathbf{x}_n \rangle = V \).

Observe that \( \langle \mathbf{x}_1, \ldots, \mathbf{x}_n \rangle \subseteq V \). For \( \mathbf{0} = \sum \lambda_i \mathbf{x}_i \) is a linear combination, and \( \lambda_1 \mathbf{x}_1 + \mathbf{0} = \sum \lambda_i \mathbf{x}_i \), so linear combinations are closed under addition and scalar multiplication.

We can extend to arbitrary subsets. If \( \mathbf{X} \subseteq V \), then

\[ \langle \mathbf{X} \rangle = \{ \sum \lambda_i \mathbf{x}_i \mid \mathbf{x}_i \in \mathbf{X}, \lambda_i \in F \} \subseteq V \]

Note that \( \langle \emptyset \rangle = \{ \mathbf{0} \} \).

Suppose that \( \mathbf{X}, \mathbf{Y} \subseteq V \) and \( \mathbf{X} \subseteq \langle \mathbf{Y} \rangle \). Then \( \langle \mathbf{X} \rangle \subseteq \langle \mathbf{Y} \rangle \).

For, suppose \( \sum \lambda_i \mathbf{x}_i \in \langle \mathbf{X} \rangle \), with \( \mathbf{x}_i \in \mathbf{X} \). Then we can write \( \mathbf{x}_i = \sum \mu_i \mathbf{y}_i \) with \( \mathbf{y}_i \in \mathbf{Y} \).

Then \( \sum \lambda_i \mathbf{x}_i = \sum (\lambda_i \mu_i) \mathbf{y}_i \in \langle \mathbf{Y} \rangle \).

Note the special case: if \( \mathbf{X} \subseteq \langle \mathbf{x}_1, \ldots, \mathbf{x}_n \rangle \)

then

\[ \langle \mathbf{x}_1, \ldots, \mathbf{x}_n \rangle = \langle \mathbf{x}_1, \ldots, \mathbf{x}_n \rangle \]

Definition

Let \( \mathbf{x}_1, \ldots, \mathbf{x}_n \in V \), a vector space. We say \( \mathbf{x}_1, \ldots, \mathbf{x}_n \) are linearly independent just when

\[ \sum \lambda_i \mathbf{x}_i = \mathbf{0} \Rightarrow \lambda_i = 0 \quad \forall \lambda_i \]

i.e., no non-trivial linear combination of the \( \mathbf{x}_i \) is equal to \( \mathbf{0} \).
If \( x_1, \ldots, x_n \) are not independent, we say that they are dependent.

**Proposition**

Take \( x_1, \ldots, x_n \in V \). Then \( x_1, \ldots, x_n \) is linearly independent if
\[
\Rightarrow \quad x_k \neq \sum_{i \neq k} \alpha_i x_i \quad \forall k.
\]
i.e. No \( x_k \) is in a linear combination of the other \( x_i \).

**Proof**

(\( \Rightarrow \)) Suppose, for contradiction, that \( x_k \in \langle x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n \rangle \). Then
\[
1 \cdot x_k + \sum_{i \neq k} (-\alpha_i) x_i = 0,
\]
and hence \( x_k, \ldots, x_n \) are linearly dependent.

(\( \Leftarrow \)) Suppose, \( \sum_{i \neq k} \alpha_i x_i = 0 \). If \( \lambda_i \neq 0 \), then
\[
x_k = \sum_{i \neq k} \lambda_i (-\alpha_i) x_i \in \langle x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n \rangle.
\]
So all \( \lambda_i = 0 \).

**Remark**

The definition makes sense as stated for an indexed family \((x_1, x_2, \ldots)\). So if \( x_i = x_j \) for \( i \neq j \), then \( x_1, \ldots, x_n \) is dependent because
\[
1 \cdot x_i + (-1) x_j = 0.
\]

The definition varies for sets thought of as indexed families of distinct vectors. If \( X \subseteq V \) then \( X \) is linearly independent if whenever \( x_1, \ldots, x_n \) are distinct elements of \( X \), then
\[
\sum_{i=1}^n \alpha_i x_i = 0 \Rightarrow \alpha_i = 0 \quad \forall i.
\]
Linear Algebra

Remarks
- \( \emptyset \) is independent.
- Any set \( X \) containing \( \emptyset \) is dependent because \( 1 \cdot 0 = 0 \).
- \( \{x, 2x, 3x\} \) is dependent.
- These notions are finitary, i.e., \( X \) is linearly independent \( \iff \) every finite subset of it is.

Lemma
Suppose \( X \subseteq V \) is independent, and \( y \notin \langle x \rangle \). Then \( X \cup \{y\} \) is independent.

Proofs
Suppose \( \sum_{i=1}^{n} a_i x_i = 0 \), \( a_i \in \mathbb{F} \) distinct. If \( \mu \neq 0 \), then \( y = \frac{1}{\mu} (-\frac{a_i}{\mu}) x_i \in \langle x \rangle \). So \( \mu = 0 \), and \( \mu \sum_{i=1}^{n} a_i x_i = 0 \), and as \( X \) is independent, \( x_i = 0 \) \( \forall i \).

Corollary
\( x_1, \ldots, x_n \) is independent \( \iff \) \( x_{j+1} \notin \langle x_1, \ldots, x_j \rangle \) for any \( 1 \leq j \leq n \).

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 \( A \in \mathbb{N} \), as follows.
Set \( X_0 = \emptyset \). Suppose \( X_r \) is defined. Then either 
\[ \langle x_r \rangle = V, \text{ and we terminate, setting } \lambda = r, \text{ or } \langle x_r \rangle \neq V, \]
and we pick \( x_{r+1} \in \not\langle x_r \rangle \), and set 
\[ x_{r+1} = x_r \cup \{ x_{r+1} \}, \text{ which is independent.} \]

Then, either we stop with \( \langle x_n \rangle = V, \ x_n = \{ x_1, \ldots, x_n \} \)
linearly independent, or \( \bigcup_{i=1}^{\infty} x_i = \{ x_1, \ldots, x_i, \ldots \} \), an infinite independent set.
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 $E \subseteq V$. $E$ is a basis just when it is independent and span.

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 if 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 \( e_1, \ldots, e_n \) is a maximal independent subset of \( V \). Then it is a basis.

Proof:

If it does not span, take \( e_{n+1} \in \langle e_1, \ldots, e_n \rangle \), and \( e_1, \ldots, e_n \) is a basis, contradicting maximality.

Proposition

Suppose that \( e_1, \ldots, e_n \) are linearly independent in a finite dimensional vector space \( V \). Then we can extend \( e_1, \ldots, e_n \) to a basis \( e_1, \ldots, e_n \) for \( V \).

Proof:

Let \( \langle e_1, \ldots, e_n \rangle \) span \( V \). Take \( e_1, \ldots, e_n \) to be a maximal independent subset of \( e_1, \ldots, e_n, e_{n+1} \ldots, e_{n+m} \), including the \( e_1, \ldots, e_n \). If some \( x \in \langle e_1, \ldots, e_n \rangle \), we could add it to get an independent net, contradicting maximality. So \( \langle e_1, \ldots, e_n \rangle = V \leq \langle e_1, \ldots, e_n \rangle \) and \( e_1, \ldots, e_n \) is a basis.
Alternatively, define \( e_1, \ldots, e_n \) inductively for \( n \geq k \) by either
\[ \langle e_1, \ldots, e_k \rangle = V, \text{ and we have a basis } n = r \]
or
\[ \langle e_1, \ldots, e_r \rangle \neq V, \text{ and so } x \in \left[ e_1, \ldots, e_r \right] \text{ and we let } e_{r+1} = x - \sum_{i=1}^r \alpha_i e_i, \text{ and } e_1, \ldots, e_{r+1} \text{ are still linearly independent.} \]

**Proposition**

Let \( e_1, \ldots, e_n \in V \), a vector space. Then \( e_1, \ldots, e_n \) is a basis \( \iff \) every \( x \in V \) has a unique expression \( x = \sum_{i=1}^n x_i e_i \) as a linear combination of the \( e_i \).

**Proof**

\[ (\Rightarrow) \] \( e_1, \ldots, e_n \) span, so we can write \( x = \sum_{i=1}^n x_i e_i \). If also \( x = \sum_{i=1}^n x'_i e_i \), then \( \sum (x_i - x'_i) e_i = 0 \), so \( x_i = x'_i \), by independence.

\[ (\Leftarrow) \] By definition, \( e_1, \ldots, e_n \) span. If \( x = \sum_{i=1}^n x_i e_i = 0 \), then \( x = \sum_{i=1}^n 0 e_i \) and, then \( x_i = 0 \) by uniqueness.

**Consequence**

\( e_1, \ldots, e_2 \) is a basis for \( V \iff (e_1 \neq 0), V = \langle e_1 \rangle \oplus \ldots \oplus \langle e_k \rangle \)

**Observation**

Suppose \( e_1, \ldots, e_n \) is a basis for \( V \) and \( w_1, \ldots, w_k \in W \). Then there is a unique linear map \( \alpha: V \to W \) such that \( \alpha(e_i) = w_i \). For we must define \( \alpha(x) = \alpha(\sum x_i e_i) = \sum_{i=1}^n x_i w_i \). This is well defined by above, and linear, because
\[
\alpha(x_1 e_1 + \mu y) = \alpha(\sum (x_1 e_i + \mu y_i) e_i) = \sum (x_1 e_i + \mu y_i) w_i = \alpha(x_1 e_1) + \mu \alpha(y).
\]
Application

Take the standard basis and some \( x_1, \ldots, x_n \in V \). Then there is a unique linear map \( \hat{E} : F^n \rightarrow V \), \( (a_1, \ldots, a_n) \mapsto \sum a_i x_i \)
are independent, \( \iff \hat{E} \) is injective
are a basis, \( \iff \hat{E} \) is an isomorphism.

2.3 Exchange Lemma

Let \( e_1, \ldots, e_n \) be linearly independent in \( V \), and \( x_1, \ldots, x_m \) span \( V \), a finite-dimensional vector space. Then \( n \leq m \), and, reordering if necessary, we have \( e_1, \ldots, e_n \), \( x_1, x_{n+1}, \ldots, x_m \) spanning.

Proof:

Let \( e_1, \ldots, e_n \) be such that \( n \) is maximized with \( n \leq m \), and \( e_1, \ldots, e_n \), \( x_{n+1}, \ldots, x_m \) spanning after reordering if necessary. Suppose \( n < m \). Then \( e_{n+1} \in \langle e_1, \ldots, e_n, x_{n+1}, \ldots, x_m \rangle \), and so we can write
\[
e_{n+1} = \sum \lambda_i e_i + \sum \mu_i x_i.
\]
If all \( \mu_i = 0 \), we contradict linear independence of the \( e_i \). So some \( \mu_i \neq 0 \), and reordering, we can assume \( \mu_1 \neq 0 \).

Then \( x_{n+1} = \sum (-\lambda_i) e_i + \frac{1}{\mu_1} e_{n+1} + \sum (-\frac{\lambda_i}{\mu_1}) x_i \),
\( x_{n+1} \in \langle e_1, \ldots, e_n, x_{n+1}, \ldots, x_m \rangle \)
But all the other elements of the spanning set \( e_1, \ldots, e_n, x_{n+1}, \ldots, x_m \) are also in this space. So the space spanned is \( V \), and \( e_1, \ldots, e_n, x_{n+1}, \ldots, x_m \) span, contradicting maximality of \( m \). So the maximal \( n \) is \( n \), and we have \( n \leq m \) and \( e_1, \ldots, e_n, x_{n+1}, \ldots, x_m \) spanning.

Corollary

If \( e_1, \ldots, e_n \) and \( x_1, \ldots, x_n \) are bases for a vector space, then \( n \leq m \).
Corollary to the exchange lemma

If \( e_1, \ldots, e_n \) and \( f_1, \ldots, f_m \) for \( V \) then \( n = m \).

Definition

A vector space \( V \) has dimension \( n \) if it has a basis \( e_1, \ldots, e_n \) of \( n \) elements. A finite dimensional vector space has a unique dimension written \( \dim V \).

Another Consequence

If \( \dim V = m \) and \( e_1, \ldots, e_n \) are independent, then \( n \leq m \), either by the exchange lemma, or by extending the independent set to a basis using invariance of dimension.

Deposition

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: \( \phi : F^r \to V \), \((a_1, \ldots, a_r) \mapsto \sum a_i x_i \) is an isomorphism just when \( x_1, \ldots, x_r \) 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 $e_1, \ldots, e_n \in V$ with dimension $n$. Then:

1) $e_1, \ldots, e_n$ independent $\Rightarrow e_1, \ldots, e_n$ form a basis.

2) $e_1, \ldots, e_n$ span $\Rightarrow e_1, \ldots, e_n$ form a basis.

Proof

1) We can extend $e_1, \ldots, e_n$ to a basis, but it will be of size $n$, so $e_1, \ldots, e_n$ is already a basis.

2) Take a minimal spanning subset to get a basis. This will be of size $n$ and so all of $e_1, \ldots, e_n$.

Rank–Nullity Theorem

Definition

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

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

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

Proof:

Take \( e_1, \ldots, e_k \) a basis for \( \ker(\alpha) \), and extend to a basis \( e_1, \ldots, e_n \) for \( U \). Then \( n(\alpha) = k \) and \( \dim U = n \). We claim that \( \alpha(e_1), \ldots, \alpha(e_n) \) is a basis for \( \text{Im}(\alpha) \).

Spanning: Take \( y = \frac{1}{\alpha} \alpha(e_i) \in \text{Im}(\alpha) \) with \( \alpha = \frac{1}{\alpha} \alpha(e_i) \).

Then \( \frac{1}{\alpha} \alpha(e_i) = \frac{\alpha}{\alpha} \alpha(e_i) = \frac{1}{\alpha} \alpha(e_i) \). Thus \( \alpha(e_1), \ldots, \alpha(e_n) \) span \( \text{Im}(\alpha) \).

Independence: Suppose \( \sum_{i=1}^{n} \mu_i \alpha(e_i) = 0 \). Then \( \alpha \left( \sum_{i=1}^{n} \mu_i e_i \right) = 0 \) as \( \mu_i e_i \in \ker \alpha \) and \( \alpha \) is linear. Thus \( \sum_{i=1}^{n} \mu_i e_i = 0 \).

Then \( \sum_{i=1}^{n} \mu_i e_i + \sum_{i=1}^{n} (-\mu_i) e_i = 0 \) and \( \alpha \left( \sum_{i=1}^{n} \mu_i e_i \right) = 0 \) by independence of \( e_1, \ldots, e_n \). This shows independence of \( \alpha(e_1), \ldots, \alpha(e_n) \).

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

Another Version

Suppose \( W \subseteq U \), finite dimensional. Take \( e_1, \ldots, e_k \) a basis for \( W \) and extend to \( e_1, \ldots, e_n \) a basis for \( U \). Then \( (e_1 + W), \ldots, (e_k + W) \) is a basis for \( U/W \). This is essentially the rank-nullity theorem for \( U \rightarrow U/W \).
Remark

1. \(0: U \rightarrow V\) is the only map with \(\text{rank } 0\).

2. If \(x \neq 0\), then \(r(ax) = r(x)\), and \(n(ax) = n(x)\).

3. \(r(x + y)\) is not determined by \(r(x), r(y)\). For \(r(x + y) = r(x) + r(y)\), but \(\text{Im}(x + y) \leq \text{Im}(x) + \text{Im}(y)\) so it is too, but \(r(x + y) = r(x) + r(y)\).

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 \(r: U + V \rightarrow W\), \((u, v) \mapsto u + v\).

\(\text{Im}(r) = U + V\) and \(\text{ker}(r) = \dim(U + V)\).

\(\text{ker}(r) = \{(u, -u) | u \in U\} \leq U + V\), so \(\dim(U \cap V) = n(r)\).

\(\dim(U + V) = \dim(U) + \dim(V)\).

(because for bases \(e_1, \ldots, e_n, f_1, \ldots, f_m\) of \(U, V\), \((e_1, 0), \ldots, (e_n, 0), (0, f_1), \ldots, (0, f_m)\) is a basis for \(U + V\).

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

Proof of Theorem

Take \(e_1, \ldots, e_k\) a basis for \(U \cap V\). Extend it to \(e_1, \ldots, e_n, f_1, \ldots, f_m\) for \(U\) and \(e_1, \ldots, e_n, g_1, \ldots, g_m\) for \(V\).

Then claim that \(e_1, \ldots, e_k, f_1, \ldots, f_m, g_1, \ldots, g_m\) is a basis for \(U + V\).
2.5 Rank and Nullity of a Composite

**Proposition** Suppose \( \alpha : U \rightarrow V \), \( \beta : V \rightarrow W \) are linear maps of finite dimensional vector spaces. Then:

a) \( \text{r}(\beta \alpha) \leq \text{r}(\beta) \) and \( \text{r}(\beta \alpha) \leq r(\alpha) \)

b) \( n(\beta \alpha) \leq n(\alpha) + n(\beta) \) and \( n(\beta \alpha) \geq n(\beta) + \dim U - \dim V \)

**Proof**

a) \( \text{Im} \beta \alpha \subseteq \text{Im} \beta \iff \text{r}(\beta \alpha) \leq \text{r}(\beta) \)

\( \text{Ker} \beta \alpha \supseteq \text{Ker} \alpha \iff n(\beta \alpha) \geq n(\alpha) \)

From this we deduce \( \text{r}(\beta \alpha) = \dim U - n(\beta \alpha) \leq \dim U - n(\alpha) = \text{r}(\alpha) \)

And further \( \text{r}(\beta \alpha) \leq \text{r}(\beta) \)

\( \implies \dim U - n(\beta \alpha) \leq \dim V - n(\beta) \) and so \( n(\beta \alpha) \geq n(\beta) + \dim U - \dim V \)

b) To show \( n(\beta \alpha) \leq n(\alpha) + n(\beta) \) it suffices to show that \( \dim U - \text{r}(\beta \alpha) \leq \dim U - \text{r}(\alpha) + \text{r}(\beta) \), that is:

\[
\text{r}(\alpha) - \text{r}(\beta \alpha) \leq \text{r}(\beta) \tag{†}
\]

Note that \( \text{Im} \beta \alpha = \text{Im}(\beta | \text{Im} \alpha) \). Now \( \text{r}(\alpha) = \dim \text{Im} \alpha \)

\( \text{r}(\beta \alpha) = \text{r}(\beta | \text{Im} \alpha) \), so it suffices to show that \( \text{r}(\beta | \text{Im} \alpha) \leq n(\beta) \)

\( \text{Ker}(\beta | \text{Im} \alpha) = \text{Ker}(\beta \alpha | \text{Im} \alpha) \leq \text{Ker}(\beta) \) and this follows

Also, from † we get \( \text{r}(\alpha) - \text{r}(\beta \alpha) \leq \dim V - \text{r}(\beta) \)

\[
\text{r}(\beta \alpha) \geq \text{r}(\alpha) + \text{r}(\beta) - \dim V
\]

Chapter 3: Matrices

3.1 Coordinates

Let \( e_1, \ldots, e_n \) be a basis for \( U \). We saw that give an isomorphism \( \hat{E} : F^n \rightarrow U \), \( (a_i) \mapsto \sum a_i e_i \)

Let \( E : U \rightarrow F^n \) be the inverse. So if \( x = \sum a_i e_i \), then \( E(x) = (a_1, \ldots, a_n) \)

We say the \( (x_i) \) are the coordinates of \( x \) with respect to \( e_1, \ldots, e_n \)

If we set \( E(x) = (E_1(x), \ldots, E_n(x)) \) with \( E_i : U \rightarrow F \) the coordinate functions,

then later we will identify the \( (E_i) \) as the dual basis to \( (e_i) \) in \( U^* \).

Note \( x = \sum E_j(x) e_i \)
3.2 Matrices

Let \( e_1, \ldots, e_n \) be a basis for \( U \), giving \( \varepsilon : U \to F^n \) and \( f_1, \ldots, f_m \) be a basis for \( V \), giving \( \phi : V \to F^m \). Suppose \( \alpha : U \to V \) is linear. Then we have maps \( \varepsilon, \phi, \phi \circ \alpha \) in the direction:

\[
\begin{array}{ccc}
U & \xrightarrow{\varepsilon} & V \\
\downarrow \alpha & & \downarrow \phi \\
F^n & \xrightarrow{\phi} & F^m
\end{array}
\]

and we let \( A \) be the matrix making the diagram commute, i.e. multiplication by \( A \) in the map \( \phi \circ \alpha \).

What does \( A \varepsilon = \phi \alpha \) mean?

Take \( x \in U \) with \( x = x^i e_i \). Then \( \varepsilon(x) = \left( x^i \right) \).

\[
A \varepsilon(x) = \begin{bmatrix} \sum a_{ij} x^j \end{bmatrix}
\]

So \( A \varepsilon(x) = \phi \alpha(x) \) says that

\[
\begin{bmatrix} a_{ij} \end{bmatrix} \alpha(x) = \sum \left( \sum a_{ij} x^j \right) f_i
\]

That is, if \( (x^i) \) are the coordinates of \( x \), then

\[
(\sum a_{ij} x^j) \] are the coordinates of \( \alpha(x) \).

Set \( x = e_k \) and we get

\[
\alpha(e_k) = \sum a_{ik} f_i
\]

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

3.3 Change of basis

Suppose that \( e_1, \ldots, e_n \) and \( e'_1, \ldots, e'_n \) are bases for \( U \). Let \( P \) be the matrix for the identity with respect to these two bases, i.e.

\[
\begin{array}{ccc}
U & \xrightarrow{P} & U \\
\downarrow \varepsilon & & \downarrow \varepsilon \\
F^n & \xrightarrow{\varepsilon} & F^n
\end{array}
\]

If \( x = x^i e_i \), then

\[
\varepsilon(x) = \sum (\sum a_{ij} x^j) e'_i
\]

So \( P \) gives the new coordinates of \( x \) in terms of the old.

On the other hand, \( e'_i = \sum p_{ij} e_i \) so \( P \) gives the old basis vector in terms of the new. Note that \( P \) is invertible. Write \( P^{-1} = \beta \).

Suppose \( e_1, \ldots, e_n, e'_1, \ldots, e'_n \) are bases for \( U \) with \( \varepsilon, \varepsilon' : U \to F^n \) and \( f_1, \ldots, f_m, f'_1, \ldots, f'_m \) are bases for \( V \) with \( \phi, \phi' : V \to F^m \).

Suppose \( \alpha : U \to V \) has matrix \( A \) with respect to \( (e_i) \) and \( (f_i) \). What is the relation between \( A \) and \( A' \)?
We see that \( A' = QAP^{-1} \)
Recall that we had bases \( e_i, \ldots, e_n \) and \( e'_i, \ldots, e'_n \) for \( U \).
Change of basis matrix \( P = (p_{ik}) \) with \( e_k = \sum p_{ik} e'_i \) and so \( e'_k = \sum p_{ik} e_i \) where \( P = (p_{ik}) = P^{-1} \).
Also bases \( f_1, \ldots, f_m \), and \( f'_1, \ldots, f'_m \) for \( V \). Change of basis matrix \( Q = (q_{is}) \) with \( f_i = \sum q_{is} f'_s \).

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

\[
\alpha(e'_i) = \alpha \left( \sum_P p_{ij} e_j \right) = \sum_P p_{ij} \left( \sum_s a_{is} f_s \right)
\]

\[
= \sum_P a_{is} \sum_P p_{ij} f_s = \sum_P a_{is} p_{ij} f_s = \sum_P a_{is} p_{ij} f_s
\]

\[
\rightarrow a'_{ij} = \sum_P q_{is} a_{sr} p_{ij} \quad \Rightarrow A' = QAP = QAP^{-1}
\]

3.4 Row and column rank

Let \( A \) be an \( m \times n \) matrix. Then we regard \( A \) as a map

\( A : F^n \rightarrow F^m \)

\( r(A) = \dim \text{Im} A \)

\( r(A) = \dim \text{space spanned by columns of } A \)

\( = \text{max number of linearly independent columns} \)

\( = \text{colrank}(A) \), the column rank of \( A \)

\( \text{rowrank}(A) = \dim \text{space spanned by row vectors of } A \)

\( = \text{max number of independent rows} \)

Note that \( \text{rowrank}(A) = r(A^T) \), \( A^T : F^m \rightarrow F^n \)

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

\( F^n \rightarrow F^r \rightarrow F^m \)

Proof:

First we show that in any such factorization we have \( r(A) \leq r \), \( r(A) \leq r(CB) \leq r(B) \), \( r(C) \) by 2.5

\( r(B) = \dim \text{Im} B \leq \dim \text{Im} F^r = r \)

\( r(C) = \dim F^r - n(C) \leq r \)
Second we show that we can realise \( r = r(A) \). For we have a factorisation of \( A \) as \( F^n \to \text{Im} A \to F^m \) and \( \text{Im}(A) \subseteq F^{r(A)} \) and so a factorisation \( F^n \to F^{r(A)} \to F^m \)

**Theorem** For any \( m \times n \) matrix \( A \), \( \text{col} \text{rank}(A) = \text{row} \text{rank}(A) \)

**Proof**

\[
A = CB \iff A^T = B^T C^T
\]

\[
F^n \to F^m \iff F^{r(A)} \to F^m
\]

So the minimal \( r \) on the left hand side = minimal \( r \) on the right hand side

i.e. \( r(A) = r(A^T) \), \( \text{col} \text{rank}(A) = \text{row} \text{rank}(A) \)

### 3.5 Row Echelon Form

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

1. **Transposing Rows**
   
   \( A \to TA \) where \( T = \left( \begin{array}{ccc} 1 & & \vspace{10pt} \cr & \ddots & \vspace{10pt} \cr \vspace{10pt} & & 1 \end{array} \right) \)

   (i.e. interchanging rows \( \left( \right) \))

2. **Scaling**
   
   \( A \to MA \) where \( M = \left( \begin{array}{ccc} & & \vspace{10pt} \\
   \vspace{10pt} & \ddots & \\
   \vspace{10pt} & & \end{array} \right) \text{ k} \neq 0 \)

   (multiply the \( k \)-th row by \( \lambda \))

3. **Adding a scalar multiple of one row to another**
   
   \( A \to LA \)

   (adding \( \lambda \times \text{row } j \) to \( \text{row } i \))

- The row operations and their corresponding matrices are invertible
- A combination of row operations is of the form \( A \to QA \) with \( Q \) invertible and so amounts to choosing a new basis for the image space.
Linear Algebra

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.
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 $\text{col-rank}(B) = \text{row-rank}(B)$. Also, the process leaves the row space the same. So $\text{row-rank}(A) = \text{row-rank}(B)$. The process doesn't affect independence of columns so $\text{col-rank}(A) = \text{col-rank}(B)$. Clearly $\text{row-rank}(A) = \text{row-rank}(B)$.

(Abtractly it follows from $B = QA$, $Q$ invertible
\[ \Rightarrow r(A) = r(B) \]
\[ B^T = A^TQ^T, \text{Q invertible } \Rightarrow r(A^T) = r(B^T) \])
Linear Algebra

Last time we took an mxn matrix A and put it in row echelon form by $A \rightarrow QA$

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 $e_1, \ldots, e_n$ for $\mathbb{F}^n$ and $E, \ldots, F_m$ for $\mathbb{F}^m$. We go through the list $\alpha(e_1), \ldots, \alpha(e_n)$ and when we can remove an $e_i$ and replace it with $\alpha(e_i)$, we do so to get a basis $(\alpha(e_{i_1}), \ldots, \alpha(e_{i_n}),$ remaining $e_i)$, s.t. $j(1) < j(2) < \ldots < j(n)$.

Chapter 4: Determinants

4.1 Endomorphism

An endomorphism of a vector space $V$ is a linear map $\alpha: V \rightarrow V$. Given an EM $\alpha$ of $V$, a finite dimensional vector space, its matrix with respect to a basis $e_1, \ldots, e_n$ is its matrix as a linear map with respect to $(e_i)$ in both domain and range i.e. it is the matrix $A = (a_{ij})$ such that $\alpha(e_i) = \sum_{j=1}^n a_{ij} e_j$.

So we have $V \xrightarrow{E} \xrightarrow{F}$ commuting.

If $e_i'$ is another basis with change of basis matrix $P = (p_{ij})$ so that $e_i = \sum_{j=1}^n p_{ij} e_j'$ then the matrix of $A'$ of $\alpha$ with respect to $(e_i')$ 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 $a_1, \ldots, a_n$ of vectors determines a volume element

$V(a_1, \ldots, a_n) = a_1 \wedge \ldots \wedge a_n \in \Lambda^n(V)$

In 2D:

```
\[ a + b \]
```
Volume is multilinear \( V(-, \lambda a + \mu b, -) = \lambda V(-, a, -) + \mu V(-, b, -) \)
and alternating \( V(-, a, -, a, -) = 0 \)

**Lemma** \( V(-, a, -, b, -) = -V(-, b, -, a, -) \)

**Proof**

\[
V(-, a + b, -, a + b, -) = 0
\]
\[
V(-, a, -, b, -) + V(-, a, -, b, -) + V(-, b, -, a, -) + V(-, b, -, b, -) = 0
\]
Hence the result.

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

\[
a_{ij} = \sum a_{ij} e_i
\]

Then \( V(a_1, \ldots, a_n) = \sum \prod_{j=1}^n a_{ij} e_j \) \( V(e_{i_1}, \ldots, e_{i_n}) \) = 0 unless for a bijection

\[
= \sum \prod_{i \in S_n} a_{i \sigma(i)} e_i(e_{i(1)}, \ldots, e_{i(n)})
\]

So any \( V(a_1, \ldots, a_n) \) is a scalar multiple of \( V(e_1, \ldots, e_n) \) and so \( \wedge (V) \) is of dimension \( \leq 1 \).

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

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

\[4.3\] Volume Forms

Let \( V \) be of dimension \( n \). A volume form \( w \) on \( V \) is a map

\( \omega : V^n \to F \) which is:

multilinear \( \omega(-, \lambda a + \mu b, -) = \lambda \omega(-, a, -) + \mu \omega(-, b, -) \)
alternating \( \omega(-, a, -a, -) = 0 \)

Note: a volume form \( w \) induces a linear map \( \wedge_n(V) \to F \)

\( \alpha \cdot \wedge a_1 \wedge a_2 \wedge \ldots \wedge a_n \) \( \Rightarrow \) \( \omega(a_1, \ldots, a_n) \)
As before, $e_1, \ldots, e_n$ are a basis for $V$ and $a_i$ defined by $a_i = \sum_{j=1}^{n} a_{ij} e_i$ then

$$\omega(a_1, \ldots, a_n) = \left( \sum_{\sigma \in S_n} \prod_{j} a_{\sigma(j) i} e(\sigma) \right) \omega(e_1, \ldots, e_n)$$

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

$$(a_1, \ldots, a_n) \mapsto \sum_{\sigma \in S_n} e(\sigma) \prod_{j} a_{\sigma(j) i}$$

- This is manifestly multilinear
- If $a_i = a_j$, $i \neq j$, then let $T(i \leftrightarrow j)$, and for any $\sigma$

$$\prod_k a_{\sigma(k) i} = \prod_k a_{\sigma(k) j}$$

so in the sum, the terms cancel in pairs and $\omega = 0$. So we have a volume form.

The image of $V(e_1, \ldots, e_n)$ under this is 1, so $\Lambda^n V$ does have dimension 1.
Last time we saw:
- Any volume form \( \omega \) satisfies
  \[
  \omega (a_1, \ldots, a_n) = \left( \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^n a_{\sigma(i)} \right) \omega(e_1, \ldots, e_n)
  \]
  where \( a_\sigma = \sum_{i=1}^n a_{\sigma(i)} e_i \)
  \[a_1, \ldots, a_n \mapsto \left( \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^n a_{\sigma(i)} \right) = \det(a_1, \ldots, a_n) = \det A = \det \alpha\]
  with the property \( \det(e_1, \ldots, e_n) = 1 \) (so it is non-trivial).
  This shows that any volume form \( \omega \) is a scalar multiple
  \[\omega(x_1, \ldots, x_n) = \omega(e_1, \ldots, e_n) \det(x_1, \ldots, x_n)\]
  of the normalized form \( \det \).

Given \( \alpha \) an endomorphism of \( V \), set \( \alpha(e_i) = C_i \).
Then \( \det(\alpha(x_1), \ldots, \alpha(x_n)) \) is a volume form and we get
\[
\det(\alpha(x_1), \ldots, \alpha(x_n)) = \det(\alpha(e_1), \ldots, \alpha(e_n)) \det(x_1, \ldots, x_n)
\]
So
\[
\beta(e_i) = b_i, \quad B = (b_1, \ldots, b_n), \quad C = (c_1, \ldots, c_n)
\]
\[
\det(B) = \det < \det B
\]
In particular
\[
\det(\alpha(x_1), \ldots, \alpha(x_n)) = \det \alpha \det(x_1, \ldots, x_n)
\]
and so \( \omega(\alpha(x_1), \ldots, \alpha(x_n)) = \det \alpha \omega(x_1, \ldots, x_n) \) for any volume form.
It follows that the definition of \( \det \alpha \) is independent of the choice of \( e_1, \ldots, e_n \).

4.4 The determinant of a matrix
Let \( A \) be an \( n \times n \) matrix (and consider \( A : F^n \to F^n \)).
Write \( a_i = A e_i = \sum \epsilon_{i,j} e_j \).
Set \( \det A = \sum \epsilon(\sigma) \prod_{i=1}^n a_{\sigma(i)} \) and regard it as a map
\[a_1, \ldots, a_n \mapsto \prod_{i=1}^n a_{\sigma(i)} \] of the \( n \) column vectors.

1. \( \det (a_1, \ldots, a_n) \) is alternating and multilinear.
2. \( \det (e_1, \ldots, e_n) = \det I = 1 \).
3. If \( \omega \) is a volume form in \( F^n \) then
\[
\omega(a_1, \ldots, a_n) = \omega(e_1, \ldots, e_n) \det (a_1, \ldots, a_n)
\]
If \( a_r = a_s, \ r \neq s \), the \( a_{o(i)} \) and \( a_{o(i)} \) are equal.

\[
\mathbf{c} = (r, s) \quad \prod a_{o(i)} = \prod a_{o(i)}
\]

\[
\omega (a_1, \ldots, a_n) = \sum_{f: \{1, \ldots, n\} \rightarrow \{r, s\}} \prod a_{f(i)} \omega (e_{f(1)}, \ldots, e_{f(n)})
\]

\[
= \sum_{o \in G} \prod a_{o(i)} \omega (e_{o(1)}, \ldots, e_{o(n)})
\]

\[
= \sum_{o \in G} \prod a_{o(i)} \varepsilon(o) \omega (e_1, \ldots, e_n)
\]

**Aside:** \((a_1 e_1 + a_{12} e_2 + \ldots) \wedge (a_{13} e_1 + \ldots)\)

For this, multiplying out means choosing an entry from each bracket; that is, given by the choice \( f(1) \mapsto f(1) \), \( \wedge 1 \mapsto f(n)\).

Given a matrix \( C \), the function

\[(b_1, \ldots, b_n) \mapsto \det \begin{pmatrix} C_{b_1} & \cdots & C_{b_n} \end{pmatrix} = 6 \det CB \]

is alternating. Therefore \( \det \begin{pmatrix} C_{b_1} & \cdots & C_{b_n} \end{pmatrix} = \det \begin{pmatrix} C_{e_1} & \cdots & C_{e_n} \end{pmatrix} \det CB \).

Now suppose \( \alpha \) is an endomorphism of \( V \) with matrix \( A \)
with respect to \( e_1, \ldots, e_n \) and \( A' \) with respect to \( e_1', \ldots, e_n' \).

Then \( \det A = \det A' \).

For \( \det A' = \det P \det A P^{-1} = \det P \det A \det P^{-1} = \det A \)
because \( \det P \det P^{-1} = \det I = 1 \).

**Definition**

The determinant of an endomorphism \( \alpha \) of a finite-dimensional vector space is \( \det \alpha \) where \( \alpha \) is any matrix for \( \alpha \).

\( \det A \) is alternating multilinear in the columns, so the elementary column operations have the effect:

- **Transposition:** Multiply \( \det \) by \(-1\)
- **Scaling a column by \( \lambda \):** \( \lambda \)
- **Adding a multiple of:** no effect on \( \det \)

"Analyze the columns. If we add a multiple of column \( j \) to column \( i \neq j \),"
Note that
\[ \det(A) = \sum_{\sigma \in S_n} E(\sigma) \prod_{j=1}^{n} a_{\sigma(j)j} = \sum_{\rho \in S_n} E(\rho) \prod_{k=1}^{n} \rho_k a_{1k} = \det(A^T) \]

So \( \det A \) is alternating multilinear in the rows of \( A \) with the same effect as for the row operations.

4.5 The adjugate matrix

Expand \( \det A = \det (a_{ij}, \sigma_{ij} e_i, \ldots, a_{nj}) \) in the 5th column,
\[ \det A = \sum_i a_{ij} \det (a_{ij}, e_i, \ldots, a_{nj}) \]

\[ \det \begin{pmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix} \]

by column operations we can clear the rest of the 5th row
\[ \det \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix} \] It takes \( i+j-2 \) transpositions to move the 1 to the top left.

So \( \det (a_{1j}, \ldots, e_i, \ldots, a_{nj}) = (-1)^{i+j} \det A_{ij} \) is a \( (n-1) \times (n-1) \) matrix in \( \mathbb{F}^n \) with \( i,j \) th entry gone.

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

We have \( \det A = \sum_i (\text{adj}(A))_{ii} a_{jj} \) (for each \( j \))

Also \( \sum_i (\text{adj}(A))_{ij} a_{ik} = \det \text{of } A \) with \( j \)th column replaced by \( k \)th
\[ = 0 \]

Theorem:

\[ \text{adj}(A) \cdot A = (\det A) I \]

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

1. Suppose an $n \times n$ matrix is of the form
$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} \text{ with } B = m \times m, \quad D = (n-m) \times (n-m)$$
Then $\det A = \det B \det D$.

\text{EITHER}

Consider that $\prod_{i=1}^{m} a_{\sigma(i)} = 0$ unless $\sigma$ permutes $\{1, \ldots, m\}$ and $\{m+1, \ldots, n\}$ amongst themselves. Then all possible choices

$$\left( \sum_{\sigma} \prod_{i=1}^{m} b_{\sigma(i)} \right) \left( \sum_{\tau} \prod_{i=1}^{n-m} d_{\tau(i)} \right)$$

\text{OR}

$$A = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} B & C \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$
$$\Rightarrow \det A = \det D \cdot \det B \cdot \det I = \det D \cdot \det B = \det A$$

2. Suppose $\alpha$ is an endomorphism of $V$ (finite dimensional).

Suppose $A$ is an $n \times n$ matrix. Then:

$A$ is invertible iff $\det A \neq 0$.

($\Rightarrow$) If $\hat{A}$ exists, then $\hat{A}A = I$, and so $\det \hat{A} \det A = 1$, and $\det A \neq 0$.

($\Leftarrow$) We don't need to use the formula for $A^{-1} = \left( \frac{\det A \ adj(A)}{\det A} \right)$.

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, $a_1 = \sum_{i=2}^{n} a_{i1} \cdot a_i$. Then

$$\det A = \det (a_1, \ldots, a_n) = \det (a_1 - \sum_{i=2}^{n} a_{i1} \cdot a_i, \ldots)$$
using column operation $3$.

3. The trace of a matrix $A = (a_{ij})$ is the sum $tr(A) = \sum_{i=1}^{n} a_{ii}$ of the diagonal elements. Suppose $A$ is $m \times n$ and $B$ is $n \times m$, so $BA$ is $m \times m$ and $AB$ is $n \times n$.

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

\text{Consequences}

If $A$ and $A'$ are matrices of an endomorphism $\alpha$,
then $tr(A') = tr(PAP^{-1}) = tr(P^{-1}PA) = tr(A)$

So we can define the trace $tr_A$ of an endomorphism $\alpha$ of a finite-dimensional vector space to be $tr_A$, where $A$ is any matrix for $\alpha$. 


Chapter 5 Theory of an endomorphism

5.1 Invariant subspaces

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

Suppose $\alpha$ is an endomorphism of $V$, finite dimensional, and $V = U_1 \oplus \ldots \oplus U_k$ where the $U_i$ are invariant.

Take successively bases for the $U_i$ to give a basis for $V$. With respect to this basis, $\alpha$ has a matrix of the form

$$
\begin{pmatrix}
B_1 & & \\
& \ddots & \\
& & B_k
\end{pmatrix}
$$

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

This is a handle on $\alpha$.

Suppose $U \leq V$ is an invariant for $\alpha$ and endomorphism of $V$ (finite dimensional).

Take a basis $e_1, \ldots, e_m$ for $U$ ($m = \dim U$) and extend to a basis $e_1, \ldots, e_n$ for $V$ ($\dim V = n$). With respect to this basis $\alpha$ has a matrix of the form

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

where $B$ is the matrix for $\alpha|_U$ with respect to $e_1, \ldots, e_m$.

What is $D$? $\alpha$ induces a map $\bar{\alpha} : V/U \to V/U$.

$\bar{\alpha}(x + U) = \alpha(x) + U$. $D$ is the matrix for $\bar{\alpha}$ with respect to the basis $e_m+U, \ldots, e_n+U$ for $V/U$.

Cyclic Subspaces. Let $\alpha$ be an endomorphism of $V$. For any $x \in V$ there is a least invariant subspace $\langle x \rangle$ containing $x$.

Why? If $x \in U$ invariant, then $\alpha(x) \in U$, $\alpha^2(x) \in U$, etc.

So $\langle x, \alpha(x), \alpha^2(x), \ldots \rangle \leq U$. But this is also invariant. So $\langle x \rangle = \langle x, \alpha(x), \alpha^2(x), \ldots \rangle$. Now suppose that $V$ is finite dimensional. Then $\{x, \alpha(x), \ldots \}$ is dependent and we can take $r$ least such that $\alpha^r(x) \in \langle x, \alpha(x), \ldots, \alpha^{r-1}(x) \rangle$. By choice $x, \ldots, \alpha^{r-1}(x)$ are independent, and we have

$$\alpha^r(x) = \sum_{i=0}^{r-1} m_i \alpha^i(x), \text{ so this space is invariant.}$$

So $\langle x \rangle = \langle x, \alpha(x), \ldots, \alpha^{r-1}(x) \rangle$. 


Moreover, with respect to $x_1, \ldots, x_n$, $\alpha^{-1}(x)$, $x \in \mathbb{R}^n$ has the matrix

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

(Compute $\det(B - AI)$)
5.2 Eigenvector and Eigenvalues

**Definition**
Let \( \lambda \) be an endomorphism of \( V \). Then \( x \in V \) is an eigenvector of \( \lambda \) with eigenvalue \( \lambda \) just when \( \lambda(x) = \lambda x \) and \( x \neq 0 \).

**Remarks**
1. \( x \) is an eigenvector just when \( \langle x, x \rangle \lambda = 1 \) 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 \( \mathbb{R}^2 \), \( \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2 \) 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., \( V \neq \{0\} \)).

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

**Observation**
\( \lambda \) is an eigenvalue iff \( \ker(\lambda - \lambda I) \neq \{0\} \)
iff \( \lambda - \lambda I \) is not invertible (a singular)
iff \( \det(\lambda - \lambda I) = 0 \).

We want to set \( \chi_\lambda(t) = \det(\lambda - tI) \), the characteristic polynomial of \( \lambda \). Then \( \lambda \) is an eigenvalue iff \( \chi_\lambda(\lambda) = 0 \).

Then the existence of eigenvalues follows from the Fundamental Theorem of Algebra, because \( \chi_\lambda(t) \) is of degree \( n = \dim V \), and so if \( n \geq 1 \), \( \chi_\lambda(t) \) has root over \( \mathbb{C} \).

We define the characteristic polynomial \( \chi_\lambda(t) \) of an \( n \times n \) matrix by \( \chi_\lambda(t) = \det(\lambda - tI) \). Note that \( \lambda - tI = \begin{pmatrix} \lambda - t & -a_1 \\ \vdots & \ddots & \ddots \\ -a_{n-1} & \cdots & \lambda - t \end{pmatrix} \) has entries in \( \mathbb{F}[t] \), the ring of polynomials in \( t \).

Using the formula we get \( \chi_\lambda(t) \in \mathbb{F}[t] \). It's clear that the degree of \( \chi_\lambda(t) \) is \( n \).
If $\alpha$ is an endomorphism of $V$ with matrices $A, A'$ with respect to bases $(e_i)$ and $(e_i')$, then $A = P A P^{-1}$ and so $(A' - t I) = P (A - t I) P^{-1}$. So $X_{A'}(t) = \det(A' - t I) = \det(A - t I) = X_A(t)$. It follows that we can define $X_\alpha(t)$ to be $X_A(t)$ where $A$ is any matrix for $\alpha$. Substituting $t$ for $\alpha$, we see that $X_\alpha(\alpha) = \det(\alpha - \alpha I)$. So the roots of the characteristic polynomial are the eigenvalues.

5.3 The Cayley-Hamilton Theorem

If $\alpha$ is an endomorphism of $V$ then we can define $\alpha^n$ 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 $p(\alpha) \cdot \alpha^n + \ldots + a_1 \alpha + a_0 I$. (N.B. If $p(t)$ and $q(t) \in F[t]$, we do have $p(q) = q(p(q))$).

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

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

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

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

So $X_A(A) = \det\begin{pmatrix} a_{11} - a_{11} I & \cdots & a_{1n} - a_{1n} I \\ \vdots & \ddots & \vdots \\ a_{n1} - a_{n1} I & \cdots & a_{nn} - a_{nn} I \end{pmatrix}$ and it is this that must be 0.
This shows that the following is nonsense:
\[ X_\alpha(t) = \det(A - tI), \text{ so setting } t = \alpha \]
\[ X_\alpha(\alpha) = \det(A - \alpha I) = \det 0 = 0 \]

\section*{SECRET}
Take a matrix \( A \) with respect to \( e_1, ..., e_n \).
Take the transpose of \( (t) \) and apply it to the column vector of
basis vectors.
\[
\begin{pmatrix}
  a_{11} - t & \cdots & a_{1n} \\
  \vdots & \ddots & \vdots \\
  a_{n1} & \cdots & a_{nn} - t
\end{pmatrix}
\begin{pmatrix}
  e_1 \\
  e_2 \\
  \vdots \\
  e_n
\end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\]
So \((t)\) is singular and \( \det(t) = 0 \).

\section*{Proof by cyclic subspaces}
Let \( \alpha \in \mathcal{V} \) be \( \neq 0 \) and consider \( \langle \alpha \rangle_\alpha = \langle \alpha, ..., \alpha^\ell(\alpha) \rangle \),
where \( \alpha^\ell(\alpha) = \sum \lambda_i \alpha^i(\alpha) \).
Let \( \beta \) be the restriction of \( \alpha \) to \( \langle \alpha \rangle_\alpha \). With respect to the basis
\( e_1, ..., \alpha^{\ell-1}(\alpha) \),
\( \beta \) has matrix
\[
\begin{pmatrix}
  0 & \lambda_1 \\
  0 & 0 & \lambda_2 \\
  \vdots & \ddots & \ddots \\
  0 & \cdots & \cdots & 0 & \lambda_\ell
\end{pmatrix}
\]
\[ X_\beta(t) = \det(B - tI) = (-1)^{\ell+1}(\lambda_0 + \lambda_1 t + ... + \lambda_\ell t^{\ell-1} - t^\ell) \]
Expanding the last column.
\[ = (-1)^{\ell+1}(t^\ell + a_{\ell-1} t^{\ell-1} + ... + a_0) \text{ where } a_i = -\lambda_i \]
By \((\alpha)\), \( X_\beta(\alpha)(\alpha) = 0 \) and so \( X_\beta(\alpha)(\alpha^i(\alpha)) = \alpha^i X_\beta(\alpha)(\alpha) = 0 \).
So \( X_\beta(\alpha) = 0 \) on \( \langle \alpha \rangle_\alpha \).

\section*{N.B}
\( X_\beta(\beta) = 0 \)

Extend the basis \( \alpha, ..., \alpha^{\ell-1}(\alpha) \) for \( \langle \alpha \rangle_\alpha \) to a basis for \( \mathcal{V} \)
With respect to that matrix \( \alpha \) has the matrix
\[
\begin{pmatrix}
  B & C \\
  0 & D
\end{pmatrix}
\]
and it follows that \( X_\alpha(t) = X_\alpha(t) = \det(A - tI) \)
\[ = \det(B - tI) \det(D - tI) = X_\beta(t) \det(D - tI) \]
Hence as \( X_\beta(\alpha) = 0 \) on \( \langle \alpha \rangle_\alpha \) we have \( X_\alpha(\alpha) = 0 \) on \( \mathcal{V} \).
But $x \neq 0$ is arbitrary so in particular we have

$$X_a(x)(x) = 0 \quad \forall x \in V$$

Thus $X_a(x) = 0$
Linear Algebra

S.4 Triangular Form

Let $A$ be the matrix for an endomorphism $\alpha$ of $V$ with respect to the basis $e_1, \ldots, e_n$. $A$ is in upper triangular form if $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix}$ i.e. $a_{ij} = 0$ for $i > j$.

 iff $\alpha(e_i) \in \langle e_1, \ldots, e_s \rangle = V_s \text{ } \forall \text{ } j$

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

$(\alpha - a_{11} I) : V = V \Rightarrow V_{n-1}$

$(\alpha - a_{n-1,n-1} I) : V_{n-1} \Rightarrow V_{n-2}$

$(\alpha - a_{ii} I) : V_i \Rightarrow V_0 = \{0\}$

So $\text{Tr}(\alpha - a_{ii} I) = 0$ but also $\det(A - \lambda I) = (-1)^{n} \text{Tr}(\alpha - \lambda I)$

So the Cayley-Hamilton Theorem is clear for upper-triangular matrices.

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

Proof Induction on $\dim V$ (assume $> 1$)

Over $\mathbb{C}$ eigenvalues and so eigenvectors exist. So choose $e_1$ to be an eigenvector for $\alpha$ and complete it to a basis $e_1, e_2, \ldots, e_n$ for $V$. The matrix of $\alpha$ is

$$
\begin{pmatrix}
\lambda_{11} & \cdots & 0 \\
0 & \ddots & \vdots \\
0 & \cdots & \lambda_{nn}
\end{pmatrix}
$$

where $D$ is the matrix for $\delta : \langle e_2, \ldots, e_n \rangle \to \mathbb{C}$

$C$ is the matrix for $\phi : \langle e_2, \ldots, e_n \rangle \to \langle e_1 \rangle$

with respect to the given bases on $\langle e_2, \ldots, e_n \rangle$, $\alpha(e_i) = \lambda_i e_i + \delta_i e_1$

By the induction hypothesis there is a basis $e_2, \ldots, e_n$ for $\langle e_2, \ldots, e_n \rangle$ with respect to which $\delta$ is upper triangular i.e.

$\delta(e_j) \in \langle e_1, \ldots, e_i \rangle$, $j \geq 2$

Then $\alpha(e_i) = \lambda_i e_i + \delta_i e_1 e_j \in \langle e_1, \ldots, e_i \rangle$, $j \geq 2$

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


5.5 Diagonal form

Let \( A \) be the matrix for an endomorphism \( \alpha \) with respect to \( e_1, \ldots, e_n \). \( A \) is diagonal if \( A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \), i.e., \( a_{ij} = 0 \) for \( i \neq j \).

If \( A(e_i) = \lambda_i e_i \), that is, \( \lambda_i \) is an eigenvalue, usually we collect the vectors for a given eigenvalue together in the list, so that the diagonal matrix looks like

\[
\begin{pmatrix}
A_1 & 0 & \cdots & 0 \\
0 & A_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A_k
\end{pmatrix}
\]

with \( \lambda_1, \ldots, \lambda_k \) the distinct eigenvalues and
\[\lambda_1 + \cdots + \lambda_k = n = \dim V.\]

The first \( n \) vectors span \( k(A - \lambda_i I) \) and so on.

\( V = V_1 \oplus \cdots \oplus V_r \), if \( x_i \in V_i \), \( \alpha(x_i) = \lambda_i x_i \).

Note

1. An eigenspace \( \ker (A - \lambda I) \) is \( \alpha \) invariant. For if \( (A - \lambda I)x = 0 \), then \( (A - \lambda I)(\alpha x) = \alpha(A - \lambda I)x = \alpha(0) = 0 \).

2. If \( x \in \ker (A - \lambda I) \) then \( \alpha(x) = \lambda x \) and so \( \alpha^n(x) = \lambda^n x \) and
so for any polynomial \( p(A)(x) = p(A)x \).

Proposition

Let \( \lambda_1, \ldots, \lambda_k \) be the distinct eigenvalues for an endomorphism \( \alpha \) of \( V \).

Then the direct sum of the corresponding eigenspaces is a direct sum:
\[\ker (A - \lambda_1 I) \oplus \cdots \oplus \ker (A - \lambda_k I)\]

Proof

It suffices to show that no non-trivial sum \( x_1 + \cdots + x_k \) for \( x_i \in \ker (A - \lambda_i I) \) can be 0. So assume
\[x_1 + \cdots + x_k = 0\]
with \( x_i \in \ker (A - \lambda_i I) \).

Apply \( \prod_{i=1}^k (\lambda_i - \lambda_i) \). We get:
\[\prod_{i=1}^k (\lambda_i - \lambda_i) x_i = 0,\] Thus \( x_1 = 0 \).

Similarly, all \( x_i \) are 0.
Corollary

An endomorphism $\alpha$ of a finite dimensional vector space $V$ is diagonalisable iff $\bigoplus \ker (\alpha - \lambda_i I) = V$ with $\lambda_i$ distinct eigenvalues.

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

Corollary

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

For there are $n$ eigenspaces $\ker (\alpha - \lambda_i I) = \ker (\alpha - \lambda_i I)$ each of dimension at least 1 and so exactly 1.
5.6 The minimal polynomial

Let \( \alpha \) be an EM of a finite dimensional \( V \). For if \( \dim V = n \), then \( \dim L(V, V) = n^2 \), and then \( I = \alpha \cdot \alpha \cdot \alpha \cdots \), \( \alpha \) is \( n^2 + 1 \) elements and are dependent, i.e. we have \( \sum_{i=1}^{n^2+1} \alpha^i = 0 \) with \( \lambda_i \) 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 \( q \) 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 \( \alpha \), \( m(t) \) is this unique \( m(t) \), monic

of least degree.

Cayley Hamilton \( M_\alpha \mid \lambda \alpha \)

Corollary Suppose \( U \subseteq V \) is \( \lambda \) invariant. Then \( M_{\lambda U} \mid \lambda M_\alpha \)

Suppose \( \alpha \) is an endomorphism of \( V \), \( \alpha \in \text{End}(V) \).

that we have a basis \( \{ \alpha \} \), \( \alpha^2 \), \( \alpha^3 \), \( \alpha^n \) for \( V \), and with

\( \alpha^{n+i}(\alpha) = \sum_{i=1}^{n^2+1} \alpha^i \alpha^j(\alpha). \)

Then \( m(t) = t^n + a_1t^{n-1} + \ldots + a_0 \) with \( a_i = -\langle \alpha \rangle \) satisfies

\( m(\alpha)(\alpha) = 0 \) and \( \alpha m(\alpha)(\alpha) = 0 \).

But suppose \( p(t) = t^n + b_1t^{n-1} + \ldots + b_0 \) is a polynomial \( r < n \). Then \( p(\alpha) = \alpha^r(\alpha) + b_1\alpha^{r-1}(\alpha) + \ldots + b_0 \alpha \)

is a linear combination of basis vectors and \( \alpha \) is non-zero. That shows

that \( m(t) \) is the minimal polynomial \( M_\alpha(t). \)
Recall that we showed $X_\alpha(t) = (-1)^n m_\alpha(t)$.

We could argue that $m_\alpha(t) X_\alpha$ and both are of degree $n = \dim V$ and $X_\alpha$ is a scalar multiple of $m_\alpha$. This shows that any monic polynomial is the minimum polynomial and $(t - \lambda)$ a characteristic polynomial for some $\lambda$.

$\lambda$ is such for $\lambda : F^n \rightarrow F^n$.

**Proposition**

The roots of $m_\alpha(t)$ are exactly the eigenvalues of $\alpha$.

*(Aside: Since $m_\alpha(t)$ is any root of $m_\alpha(t)$ is a root of $X_\alpha$ and so an eigenvalue.)*

**Proof**

Let $\lambda$ be a root of $m_\alpha(t)$ and write $m_\alpha(t) = (t - \lambda) q(t)$.

By minimality $q(\lambda) \neq 0$ so we can choose $\mathbf{x} \in V$ such that

$\mathbf{y} = q(\lambda) \mathbf{x} \neq 0$. But $(\lambda - \lambda I) \mathbf{y} = (\lambda - \lambda I) q(\lambda) \mathbf{x} = m_\alpha(\lambda) \mathbf{x}$

So $\mathbf{y}$ is an eigenvector with eigenvalue $\lambda$.

Now suppose $\lambda$ is an eigenvalue with eigenvector $\mathbf{x} \neq 0$.

$0 = m_\alpha(\lambda) \mathbf{x} = m_\alpha(\lambda) \mathbf{x}$ with $\mathbf{x} \neq 0$, so $m_\alpha(\lambda) = 0$ and $\lambda$ is a root of $m_\alpha$.

(Or the minimum polynomial of $\alpha$, $\alpha(t)$ is $(t - \lambda)$ and so $(t - \lambda) m_\alpha(t)$)

Suppose $\alpha$ is diagonalizable so that $V = \ker (\alpha - \lambda I) \oplus \cdots \oplus \ker (\alpha - \lambda_k I)$

with $\lambda_i$ the distinct eigenvalues. Clearly the minimum polynomial of $\alpha$ is $\ker (\alpha - \lambda_i I) \oplus (t - \lambda_i) \ldots (t - \lambda_k) | m_\alpha(t)$.

But $(\alpha - \lambda_i I) \ldots (\alpha - \lambda_k I)$ is $0$ on each $\ker (\alpha - \lambda_i I)$ and so $m_\alpha(t) = (t - \lambda_i) \ldots (t - \lambda_k)$. Here $X_\alpha(t) = (-1)^n (t - \lambda_i) \ldots (t - \lambda_k)^n$ with $n_i = \deg (\alpha - \lambda_i I)$

**Proposition** Suppose $\alpha$ is an endomorphism of finite dimensional $V$ whose minimal polynomial splits as a product of $(t - \lambda_1) \ldots (t - \lambda_k)$ of distinct linear factors. Then $\alpha$ is diagonalizable.
Proof

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

Either by dimension: The dimension of the direct sum

\[
\dim V = \dim \ker (\alpha - \lambda, I) + \ldots + \dim \ker (\alpha - \lambda_k, I) 
\]

Then \( V = \ker (\alpha - \lambda, I) \oplus \ldots \oplus \ker (\alpha - \lambda_k, I) \) = \dim V = n \)

Or: By the Chinese Remainder Theorem. Let

Let \( q_i(t) = \frac{\text{gcd}(t)}{t - \lambda_i} \). Then the highest common factor of the

\( q_i(t) \) is 1.

And so we can write \( 1 = \sum q_i(t) q_i(t) \)

For any \( x \in V \) we have

\[
x = \sum a_i(\alpha) q_i(\alpha)(x)
\]

but \( (\alpha - \lambda_i, I) q_i(\alpha)(x) = m_i(\alpha)(x) = 0 \) 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 \( x \in \ker (\alpha - \lambda_i, I) \oplus \ldots \oplus \ker (\alpha - \lambda_k, I) \).
Suppose that $\alpha$ and $\beta$ are endomorphisms of $V$, both of which have diagonal matrices $D_\alpha$ and $D_\beta$ with respect to a basis $e_1, \ldots, e_n$.

Then $\alpha \beta(e_i) = \alpha(\beta(e_i)) = \lambda_i \alpha(e_i) = \lambda_i \beta(e_i) = \beta(e_i)$.

And so $\alpha$ and $\beta$ commute. (OR, evidently the two matrices commute!)

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

Proof:
Write $V = V_1 \oplus \cdots \oplus V_k$ where the $V_i$ are the eigenspaces of $\alpha$.

Claim: the $V_i$ are $\beta$-invariant.

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

Write $\beta_i \beta = \beta_i V_i$. $M_\beta$ $\perp$ $M_\beta$. But as $\beta$ is diagonalizable, $M_\beta$ splits into linear factors, hence so does $M_\beta$.

So $\beta_i \beta = \beta_i V_i$ is diagonalizable. In each $V_i$, choose a basis of eigenvectors of $\beta$. Put these together to give a basis for $V$ of vectors which are eigenvectors of both $\alpha$, and $\beta$.

5.7 Jordan Normal Forms

Let $\alpha$ be an endomorphism of a finite dimensional complex $V$. Take $x_k(t) = (-1)^k (t - \lambda_i)_d \cdots (t - \lambda_k)_d$ with $a_i$ the algebraic multiplicities of the eigenvalues. $n = \dim V = \sum \alpha_i$ then $x_k(t) = (t - \lambda_i)^{n_i} \cdots (t - \lambda_k)^{n_k}$ where $1 \leq n_i \leq a_i$.

FACT: $\alpha$ has a matrix $A$ in Jordan Normal Form. This will be described in stages.
$A = \begin{pmatrix} B_1 & \mathbf{0} \\ \mathbf{0} & B_2 \end{pmatrix}$ where the $B_i$ are blocks corresponding to the eigenvalues $\lambda_i$. We shall see that $B_i$ has $\lambda_i$ down the diagonal. The $B_i$ are of size $q_i \times q_i$ matrices.

This comes from $V = \ker (\alpha - \lambda_i I)^m + \ldots + \ker (\alpha - \lambda_i I)^m$ where the $\ker (\alpha - \lambda_i I)^m$ are the generalised eigenspaces.

Note $\ker (\alpha - \lambda_i I)^m = \ker (\alpha - \lambda_i I)^n$ for any $m \geq n$.

Why? Suppose $x_1 + \ldots + x_k = 0$ with $x_i \in \ker (\alpha - \lambda_i I)^m$.

Set $q_i(t) = \frac{m(t)}{t - \lambda_i}$. Apply $q_i(\alpha)$ to get $q_i(\alpha) x = 0$. So $x_1 = 0$. So we have a direct sum.

Then this is the whole of $V$, by the Chinese Remainder Theorem.

For $\lambda_i$: of the form $\alpha$

$B_i = \begin{pmatrix} \lambda_i^{n_i} & \mathbf{0} \\ \mathbf{0} & \lambda_i^{n_i} \end{pmatrix}$ where $n_i = \dim (\ker (\alpha - \lambda_i I)) = \dim (\alpha - \lambda_i I)$.

Each $C_i$ corresponds to a cyclic subspace.

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

(N.B. A cyclic subspace for $\alpha$ is cyclic for $\alpha - \lambda_i I$ and vice versa)

For there must be a vector $x$ such that $x$, $(\alpha - \lambda_i I)x$, $(\alpha - \lambda_i I)^2x$, $\ldots$, $(\alpha - \lambda_i I)^nx$ is a basis and with $(\alpha - \lambda_i I)^nx = 0$

$\alpha(x) = \lambda_i x + (\alpha - \lambda_i I)x$ (look at form of $C_i$).

Each $C_i$ contains a 1D eigenspace. Finally, the maximal size of a $C_i$ in $B_i$ is $\lambda_i$. 
Linear Algebra

Example \( \frac{d^{n-1}x}{dt^{n-1}} + c_{n-1} \frac{d^{n-2}x}{dt^{n-2}} + \ldots + c_0 x = 0 \), an ODE

Solutions = \( \left\{ f \mid f(D)(t) \leq C^\infty(C) \right\} \) where \( p \) is the auxiliary polynomial and \( D = \frac{d}{dt} \)

Rewrite as \( D \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -c_0 & -c_1 \\ \vdots & \vdots \\ -c_{n-2} & -c_{n-1} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \)

Analysis \( \frac{d}{dt} Z = A Z \) has solution \( e^{At} Z(0) \) and so a solution is uniquely determined by \( x(0), x'(0), \ldots, x^{(n-1)}(0) \).

So \( V \cong \mathbb{C}^n \). And \( D : V \to V \) has matrix \( A \) with respect to some basis. \((-1)^n X_D = M_0 = p\)

\[ X_D(t) = \sum_{k=1}^{n} \frac{(t-\lambda_k)^{a_k}}{a_k!} (t - \lambda_k)^{a_k} \]

\[ M_0(t) = (t - \lambda_1)^{a_1} \ldots (t - \lambda_k)^{a_k} \]

So the Jordan Normal Form is

\[ \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \]

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

\[ \frac{dt}{dt} e^{At} \to A e^{At} \]

\[ \frac{dt}{dt} te^{At} \to t e^{At} + e^{At} \]

\[ \frac{dt}{dt} \frac{t^2}{2} e^{At} \to \frac{t^2}{2} e^{At} + t e^{At} \]

\[ \vdots \]

\[ \frac{dt}{dt} \frac{t^{a-1}}{(a-1)!} e^{At} \to \frac{t^{a-1}}{(a-1)!} e^{At} + \frac{t^{a-2}}{(a-2)!} e^{At} \]

\[ (D - \lambda I) \frac{t^{a-1}}{(a-1)!} e^{At} \to \frac{t^a}{a!} e^{At} - t e^{At} + e^{At} \to 0 \]

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

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

In fact, $m_\alpha = m_{x_1^{(1)} x_2^{(1)} x_3^{(1)}}$.

To show $\text{ker}(\alpha - \lambda_1 I)^{m_1} + \cdots + \text{ker}(\alpha - \lambda_k I)^{m_k} = \text{null}(\alpha)$, we took $x_1^{(1)} + x_2^{(1)} = 0$, $x_i \in \text{ker}(\alpha - \lambda_i I)^{m_i}$. Then we took $q(t) = \prod_{\lambda_i}(t - \lambda_i)^{m_i}$ and applied $q(t)(x)$ to $t$.

We deduce that $q_i(\alpha)(x_i) = 0$.

But $x_i$ is not an eigenvector necessarily. However, if $x_i \neq 0$, its minimum polynomial divides $(t - \lambda_i)^m$, and is of the form $(t - \lambda)^m / m'$. But $(t - \lambda)^m / m' x_i(x) \neq x_i = 0$.

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 $e_1, \ldots, e_n$ is a basis for $V$. We define the "dual basis" for $V^*$ $e_1, \ldots, e_n$ by setting $E_i(e_j) = \delta_{ij}$.

Note that $E_i(e_j x ; e_j) = x_i$. So the $E_i$ are the old coordinate functions $V \xrightarrow{\delta} F^n; x \mapsto (E_i(x))$.

Proposition

Suppose $e_1, \ldots, e_n$ is a basis for $V^*$.

Proof.

Suppose $\sum \lambda_i e_i = 0$. Then for any $i$, $\sum \lambda_i E_i(e_i) = 0$ and so $\lambda_i = 0$. This shows independence.

Suppose that $\theta \in V^*$ and let $t_i = \theta(e_i)$, $\sum t_i e_i(e_i) = t_i = \theta(e_i)$ for all $e_i$, and so $\theta = \sum t_i e_i$ and the $e_i$ span $V^*$.

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

WARNING. Let $V$ be a finite dimensional vector space. Then every sequence from $F^n$. 

generated by the obvious unit vectors \((1,0,\ldots), (0,1,0,\ldots)\) etc.

It is countable dimensional but \(V^* \cong \mathbb{F}^N\) is not.

Aside

If \(\theta = \sum t_i e_i\) and \(x = \sum x_k e_k\), then \(\theta(x) = \sum t_k x_k\)

\[
(\begin{array}{c}
\ell_1 \\
\vdots \\
\ell_N
\end{array})
\]

Also Aside The map \(V^* \times V \rightarrow \mathbb{F}, \theta, x \mapsto \theta(x)\) is bilinear.

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

Think \(\theta(x) = \langle \theta, x \rangle\)

6.2 The dual of a linear map

Definition Let \(\alpha : U \rightarrow V\) be linear. Then the dual map \(\alpha^* : V^* \rightarrow U^*\) is defined by \(\alpha^*(\theta)(u) = \theta(\alpha(u))\) for \(\theta \in V^*, u \in U\).

Remark

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

Proposition

\(\alpha^* : V^* \rightarrow U^*\) is linear.

Proof. \(\alpha^* (\lambda \phi + \mu \psi)(x) = (\lambda \phi + \mu \psi)(\alpha(u)) = \lambda \phi(\alpha(u)) + \mu \psi(\alpha(u))\)

for \(\phi, \psi \in V^*, u \in U\).

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

Note \(\overline{U \rightarrow U^*}\) \(\overline{V^* \rightarrow V^*}\)

\(\alpha^* \beta^* = (\beta \alpha)^*\)

Proposition Suppose \(\alpha : U \rightarrow V\) has matrix \(A\) with respect to bases \(e_1, \ldots, e_n\) for \(U\) and \(f_1, \ldots, f_m\) for \(V\). Then, \(\alpha^*\) has the matrix \(A^T\) with respect to the dual basis \(e_1, \ldots, e_n\) and \(f_1, \ldots, f_m\)

Proof

\(\alpha^*(\phi_i)(e_k) = \phi_i(\alpha(e_k)) = \phi_i(\sum \alpha_{ik} f_i) = \sum \alpha_{ik} \phi_i(e_k)\)

Thus \(\phi_i(e_k) = \sum \alpha_{ik} \phi_i(e_k) = \sum [A^T]_{ki} e_k\). Then \(\alpha^*\) has matrix \(A^T\).

Note \(r(A) = \text{colrk}(A) = \text{rank}(A) = \text{colrk}(A^T) = r(\alpha^*)\)
5.3 Annihilator

Definition Let $W \leq V$. The annihilator $W^0$ of $W$ is
\[ W^0 = \{ \theta \in V^* \mid \theta(w) = 0 \text{ for all } w \in W \} \]

Note Let $\iota : W \to V$ be the inclusion map. Then $W^0 = \ker (\iota^*)$, $\iota^* : V^* \to W^*$

For $\iota^*(\theta) = 0$ if $\theta \circ \iota = 0$ if $\iota$ is 0 on $W$.

Proposition $W^0 \leq V^*$

Proof: either by the above note OR $\mathcal{O}(W) = 0 \iff \forall w \in W \implies 0 \in W^0$

If $\phi, \psi \in W^0$ then $\forall w \in W$, $(\lambda \phi + \mu \psi)(w) = \lambda \phi(w) + \mu \psi(w) = 0$

then $\lambda \phi + \mu \psi \in W^0$

Fact $\{0\}^* = V^* \quad V^0 = \{0\}$

If $W_1 \leq W_2$ then $W_1^0 \geq W_2^0$

Proposition Let $W \leq V$ be finite dimensional. Then $\dim W + \dim W^0 = \dim V$

Proof: Let $e_1, \ldots, e_n$ be a basis for $W$ and extend to a basis $e_1, \ldots, e_n$ for $V$.

Let $e_1, \ldots, e_n$ be the dual basis for $V^*$.

CLAIM $e_1, \ldots, e_n$ is a basis for $W^0$.

- If $r + 1 \leq i \leq n$ then $\forall i \leq r$ we have $e_i(e_j) = 0$, so $e_i$ is 0 on a basis for $W$, and so on $W$. Thus $e_i \in W^0$.

- $e_1, \ldots, e_n$ is independent because $e_1, \ldots, e_n$ is a basis.

- Let $\theta \in W^0$. Write $\theta = \sum b_i e_i$. If $1 \leq i \leq r$, then
  \[ 0 = \theta(e_i) = \sum b_i e_i(e_i^{*}) = b_i \]

Thus $\theta = \sum b_i e_i$. So $e_1, \ldots, e_n$ span $W^0$. That proves the claim. Now, $\dim W = r$, $\dim W^0 = n - r$, so $\dim W + \dim W^0 = \dim V$

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

Observation If $U, W \leq V$ then:

i) $(U + W)^0 = U^0 \cap W^0$

ii) If $V$ is finite dimensional then $U + W = (U \cap W)^0$
Why? i) $\theta \in (U + W)^o$ if $U + W \leq \ker \theta$

$\theta \in U^o \cap W^o$ 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 $U^o \supset (U + W)^o$, $W^o \supset (U + W)^o$. So $(U + W)^o \leq U^o \cap W^o$

$\dim(U + W)^o = \dim V - \dim(U + W) = n - (\dim U + \dim W - \dim(U + W))$

$(U \cap W)^o \supset U^o$, $(U \cap W)^o \supset W^o$, so $(U \cap W)^o \supset U^o + W^o$

$\dim(U^o + W^o) = \dim U^o + \dim W^o - \dim(U \cap W)^o$

$= n - \dim U + n - \dim W - (n - \dim(U + W))$

$= n - (\dim U + \dim W - \dim(U + W))$

$= n - \dim(U \cap W) = \dim(U \cap W)^o$

6.4 The rank of the Dual

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

Then i) $\ker \alpha^* = (\text{Im} \alpha)^o$

ii) $\text{Im} \alpha^* = (\ker \alpha)^o$ and $r(\alpha) = r(\alpha^*)$ (ii) only in the finite-dimensional case

Proof:

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

iff $\Theta = 0$ on any $w \in \text{Im} \alpha$ iff $\Theta \in (\text{Im} \alpha)^o$

ii) In the finite dimensional case we have:

$r(\alpha^*) = n - n(\alpha) = n - \dim(\text{Im} \alpha)^o = \dim(\text{Im} \alpha) = r(\alpha)$

Finally, we always have $\text{Im} \alpha^* \leq (\ker \alpha)^o$. For if $\Phi = \alpha^*(\Theta) \in \text{Im} \alpha^*$

then for $x \in \ker \alpha$, we have $\Phi(x) = \alpha^*(\Theta)(x) = \Theta(\alpha(x)) = \Theta(0) = 0$

Thus $\Phi \in \ker \alpha$.

But $\dim(\ker \alpha)^o = n - n(\alpha) = r(\alpha) = r(\alpha^*) = \dim(\text{Im} \alpha^*)$

6.5 The Double Dual

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

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

Consequence of bilinearity:

1. $\hat{x}$ is a linear map $x \in L(V^*, F) = V^{**}$

2. The map $\hat{\cdot}: V \to V^{**}$, $x \mapsto \hat{x}$ is itself linear.
This is natural in the sense that given $\alpha : U \rightarrow V$, the diagram:

```
          K
        U  α  V
          K
```

commutes.

Why? $K\alpha (\alpha(u))$ acts on $\theta \in V^*$ to give $\theta (\alpha(u))$ and $\alpha^{**}(K\alpha(u))$ acts on $\theta \in V$ to give $\hat{\theta} \circ \alpha^*(\theta) = \theta_0 \alpha(u)$.

**Proposition.** If $V$ is finite dimensional, then $K : V \rightarrow V^*$ is an isomorphism.

**Proof.** We show that $K$ is injective. Suppose $x \in V$ with $x \neq 0$. Set $e_1 = x$ and extend to a basis $e_1, \ldots, e_n$ for $V$. Take the dual basis $E_1, \ldots, E_n$. Then $S(x) = E_i (e_i) = 1 \Rightarrow x \neq 0$.

Thus, the kernel of $K$ is $\{0\}$.

Thus, $K$ is injective, but $V^{**} = V^* = V$. So $K$ is an isomorphism.

Observe, if $U \subseteq V$ then $K(U) \subseteq V^{**}$ is equal to $U^{**}$. We can therefore identify $V$ with $V^{**}$ and get a real duality ($\circ$) on subspaces.
Chapter 7 Bilinear Forms

7.1 Bilinear Maps

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

$\beta (u + u', v) = \beta (u, v) + \beta (u', v)$

$\beta (\lambda u, v + v') = \lambda \beta (u, v) + \beta (u, v')$

In this situation we have

$\beta_u : U \to V^* : \beta_u (u) (v) = \beta (u, v)$

$\beta^*_v : V \to U^* : \beta^*_v (v) (u) = \beta (u, v)$

Both linear maps.

Note: The composite $V^* \to V^* \to U^*$ is equal to $\beta^*$.

This "duality" is reflected in terms of matrices.

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

Then $\beta_u (e_i) (f_j) = b_{ij}$ and so $\beta^*_v (e_i) = \sum b_{ij} \phi_j$ where $\phi_1, \ldots, \phi_m$ is the basis dual to $f_1, \ldots, f_m$.

Similarly, $\beta^*_v$ has matrix $B$ with respect to $e_1, \ldots, e_n$ and $E_1, \ldots, E_n$.

$\beta$ is left non-degenerate iff $\beta_u$ is injective.

$\beta$ is right non-degenerate iff $\beta^*_v$ is injective.

If $u \in U$ is such that $\beta (u, x) = 0 \forall v \in V$ then $u = 0$.

In the finite dimensional case we are interested in $\beta$ being an isomorphism.

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

If $x \in U$ we have $x^* = \{ y \in V : \beta (x, y) = 0 \}$.

If $y \in V$ we have $y^* = \{ x \in U : \beta (x, y) = 0 \}$.

$y \in x^*$ iff $\beta_u (x) (y) = 0 \forall x \in x$.

$\beta_u$ is injective iff $\beta (x) (y) = 0 \forall x \in x$.

$\beta^*_v$ is surjective iff $\beta (x) (y) = 0 \forall y \in y^*$.

So in the finite dimensional case, we have (for $\beta$ non-singular)

$\dim x + \dim x^* = \dim \beta_u (x) + \dim (\beta_u (x))^* = \dim V$. 

Suppose again in the finite dimensional case that \( \beta \) is non-singular. Then for any \( \theta \in V^* \) there is a unique \( u_0 \in U \) such that 
\[
\beta(u_0, v) = \theta(v) \quad \forall v \in V.
\]

**Proposition** Let \( \beta : U \times V \to F \) be a non-singular bilinear with \( U, V \) finite dimensional. Then for any endomorphism \( \alpha \) of \( V \) there is a unique adjoint endomorphism \( \alpha^+ \) of \( U \) such that 
\[
\beta(\alpha^+(x), y) = \beta(\alpha(x), y) \quad \forall x, y \in U.
\]

**Proof** For fixed \( u \in U \), \( v \mapsto \beta(\alpha(x), y) \) is in \( V^* \) so we have a unique \( \alpha^+(x) \in U \) such that 
\[
\beta(\alpha^+(x), y) = \beta(u, \alpha(x)).
\]
It remains to show that \( \alpha^+ \) is linear.

\[
\beta(\alpha^+(x + \lambda y), z) = \beta(x + \lambda y, \alpha(x + \lambda z)) = \lambda \beta(x, \alpha^+(y), z) + \mu \beta(x, \alpha(y), z) = \beta(\alpha^+(x), y) + \mu \beta(\alpha^+(y), z) = \beta(\alpha^+(x) + \mu \alpha^+(y), z)
\]
and this holds for all \( z \), hence \( \alpha^+ \) is linear.

### 7.2 Bilinear Forms

A bilinear form \( \beta \) on \( V \) is a bilinear map \( \beta : V \times V \to F \).

Given a basis \( e_1, \ldots, e_n \) for \( V \), the matrix \( B = (b_{ij}) \) for \( \beta \) with respect to \( e_1, \ldots, e_n \) is \( b_{ii} = \beta(e_i, e_i) \).

If \( x = \sum x_i e_i \) and \( y = \sum y_i e_i \) then 
\[
\beta(x, y) = \sum x_i \beta(e_i, y_i)
\]
That is, \( (x_1, \ldots, x_n) \begin{bmatrix} B & (e_i) \end{bmatrix} = e(x) \beta(e) \).

Suppose \( \beta \) has the matrix \( B' = (b_{ij}') \) with respect to \( e_1', \ldots, e_n' \).

EITHER \( x = \sum x_i e_i' \) and \( y = \sum y_i e_i' \) and 
\[
\beta(x, y) = \sum x_i' b_{ii}' y_i'
\]
and \( x \beta = p^T B' p \)

OR \( \beta(e_i', e_i') = \beta(\sum p_{i'i} e_i, \sum p_{k'k} e_k) = \sum p_{i'i} b_{ii'} p_{k'k} \)

And \( x \beta = p^T B' p \), \( p = p' \).

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 
\[
\beta(u, v) = 0 = \beta(w, x) \quad \forall u \in U, v \in V, \quad w, x \in W,
\]
then we have a matrix.
of the form \[
\begin{pmatrix}
B & 0 \\
0 & B^T \\
\end{pmatrix}
\]

**WARNING!** It is natural to try \(U + U^+\) but in general this is not a direct sum even though \(\dim U + \dim U^+ = \dim V\) (assuming \(B\) is non-singular). For example, take
\[
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
\end{pmatrix}
\]

There are null vectors e.g. \((1, 1, 0, 0)\) such that
\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
1 \\
1 \\
0 \\
0 \\
\end{pmatrix}
\]

i.e. self-orthogonal

So if \((1, 1, 0, 0) \in U\) then \(U \cap U^+ = \{0\}\)
Proposition. Suppose $B$ is a bilinear form which is non-singular on a subspace $U \leq V$. Then $V = U \oplus U^\perp$.

Proof.
Suppose $u \in U \cap U^\perp$. Then $B(x,u) = 0$. But as $B(u,u) = 0$ and $B$ is non-singular on $U$, that implies $u = 0$.
Take $v \in V$. Consider the map $U \to F$, $u \mapsto B(u,v)$. It is in $U^\perp$ and so as $B$ on $U$ is non-singular there exists a (unique) $w \in U$ with $B(u,w) = B(u,v)$ if $u \in U \setminus U^\perp$.

Then $v - w \in U^\perp$ and $v = w + (v - w)$.

### 7.3 Symmetric Bilinear Forms

A bilinear form $B : V \times V \to F$ is *symmetric* if $B(x,y) = B(y,x)$ for all $x, y \in V$.

A *skew-symmetric* bilinear form $B(x,y) = -B(y,x)$.

**Assume** for this section that char $F \neq 2$ (i.e., $2 \neq 0$ in $F$).

Any bilinear form can be written:

$$B(x,y) = \frac{1}{2} \left[ B(x,y) + B(y,x) \right] + \frac{1}{2} \left[ B(x,y) - B(y,x) \right]$$

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

A quadratic form $q : V \to F$ is a map such that

$q(x) = B(x,x)$ for some bilinear $B$. We might as well take $B$ to be symmetric and then $B$ is determined by

$$q : B(x,y) = \frac{1}{2} \left( q(x+y) - q(x) - q(y) \right).$$

**Proposition.** Suppose $B$ is a symmetric bilinear form on a finite-dimensional $V$.

Then there is a basis $e_1, \ldots, e_n$ with respect to which $B$ has a diagonal matrix (i.e., $B(e_i, e_j) = 0$ for $i \neq j$).

**Lemma.** If $B$ is a bilinear form such that the corresponding quadratic form $q(x)^2 = B(x,x)$ is 0, then $B$ is 0.

**Proof.**

$$0 = B(x+y, x+y) = B(x,x) + 2B(x,y) + B(y,y) = 2B(x,y)$$

So $B(x,y) = 0$. 


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 \( \beta : V \times V \to F \) is 0 and any basis will do.

Or we can find a vector \( e_1 \) with \( \beta(e_1, e_1) = 1 \neq 0 \)

Claim

\( V = \langle e_1 \rangle + \langle e_2 \rangle \)

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

\( \beta(v, e_1) e_1 \in \langle e_1 \rangle \)

\( \beta(e_1, v - \frac{\beta(v, e_1)}{\lambda} e_1) = \beta(e_1, v) - \frac{\beta(v, e_1)}{\lambda} \beta(e_1, e_1) = 0 \)

So \( (v - \frac{1}{\lambda} \beta(v, e_1) e_1) \in \langle e_2 \rangle \)

This expression is unique because if \( v = \mu e_1 \) with \( \mu \in \langle e_1 \rangle \)

then \( \beta(e_1, v) = \mu \beta(e_1, e_1) = 0 \), so \( \mu = \frac{1}{\lambda} \beta(e_1, v) \) and

\( v = (v - \mu e_1) \)

By the induction hypothesis, there is a basis \( e_1, \ldots, e_n \) of \( U = \langle e_1 \rangle \)

with respect to which \( \beta \) has diagonal matrix. And then \( \beta \) has a diagonal matrix with respect to \( e_1, \ldots, e_n \).

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

\[
\begin{pmatrix}
1 & 0 & 0 \\
0 & d_2 & 0 \\
0 & 0 & d_n
\end{pmatrix}
\]

for \( \beta \), \( d_i \neq 0 \) and we can ensure that \( d_1, \ldots, d_p > 0 \) and then \( d_{p+1}, \ldots, d_n < 0 \).

Set a new basis \( e_i' = \frac{e_i}{\sqrt{d_i}} \) for \( 1 \leq i \leq n \) and we get a new matrix of the form

\[
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}
\]

\( p + q = V \)

Special Case. \( F = \mathbb{C} \). We have a matrix as for \( \mathbb{R} \) and set

\( e_i' = \frac{1}{\sqrt{d_i}} \) for \( 1 \leq i \leq n \)

we get a matrix of the form

\[
\begin{pmatrix}
I & 0 \\
0 & 0
\end{pmatrix}
\]
7.4 Sylvester's Law of Inertia

Suppose $\beta$ is a symmetric bilinear form on a finite-dimensional real $V$.

Suppose that $\beta$ has matrices

$$
\begin{pmatrix}
I_p & 0 \\
0 & -I_q
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
I_r & 0 \\
0 & -I_t
\end{pmatrix}
$$

with respect to two bases.

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

$\beta$ is positive-definite on $U$ iff $\beta(x, x) > 0 \forall x \in U$.

$\beta$ is positive semi-definite on $U$ iff $\beta(x, x) \geq 0 \forall x \in U$.

$\beta$ is negative semi-definite.

$\beta$ is negative-definite.

Suppose $U$ is a space on which $\beta$ is the definite, and $W$ a space on which $\beta$ is negative semi-definite. Take $x \in U \cap W$, $\beta(x, x) \leq 0$ and $\beta(x, x) > 0$ unless $x = 0$. So $x = 0$. Then $U \cap W = \{0\}$.

For the first basis we have a space $P$, $\dim P = p$, on which $\beta$ is positive definite, and a space $S$, $\dim S = n - p$, on which $\beta$ is negative semi-definite. Similarly, for the second basis we have a space $P'$, $\dim P' = p'$, on which $\beta$ is positive definite and a space $S'$, $\dim S' = n - p'$, on which $\beta$ is negative 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 $Q$ positive 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 $\text{rk}(\beta)$ in this case.

These are invariants of the bilinear form.
Invariance of Rank

1. Suppose $B, B'$ are matrices for a bilinear form $B$. We have $B' = Q^T B Q$ where $Q$ is invertible.

Then $r(B') = r(B)$. So we can define the rank of $B$ to be the rank of any matrix for $B$.

Aside: Suppose $S$ is invertible. $r(SA) \leq r(A) = r(S' SA) \leq r(SA)$

So $r(SA) = r(A)$ and similarly $r(AS) = r(A)$.

2. The rank of $B$ as a form in the rank of the linear map $\beta_B: V \to V^*$. Then we note that if $B$ has matrix $B$ as a form, then $\beta_B$ has matrix $B^*$ and so the rank of a matrix for $B$ is an invariant.

Consequence

When we have two matrices

\[
\begin{pmatrix}
I_p & 0 \\
0 & I_q
\end{pmatrix}
\]

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

Examples: Diagonalizing is "completing the square".

1. $2x^2 + 2xy + y^2 = (x + y)^2$.

\[
= 2(x + \frac{1}{2}y)^2 + \frac{1}{2}y^2
\]

\[
\begin{pmatrix}
x & y \\
y & \frac{1}{2}
\end{pmatrix}
\begin{pmatrix}
\frac{1}{2} & 0 \\
0 & \frac{1}{2}
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
\]

\[
= x^2 + (x + y)^2
\]

2. $x^2 + y^2$

Take $e_1 = \frac{1}{\sqrt{2}} (2, 1)$. Find $e_2 \in \langle e_1 \rangle^\perp$. Suppose $e_2 = (4, 1)$.

\[
\begin{pmatrix}
2 & 1 \\
1 & 0
\end{pmatrix}
\begin{pmatrix}
u \\
v
\end{pmatrix} = 0
\]

So $u = 1$, $v = -2$ for example.

\[
= (x, y)\begin{pmatrix}
\frac{1}{2} & 0 \\
0 & \frac{1}{2}
\end{pmatrix}\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}\begin{pmatrix}
\frac{1}{2} & 0 \\
0 & \frac{1}{2}
\end{pmatrix}\begin{pmatrix}
x \\
y
\end{pmatrix}
\]
7.3 Hermitian Forms

For any vector space \( V \) over \( \mathbb{C} \) there is a vector space \( V \) over \( \mathbb{C} \) which is \( V \) with the same addition and with scalar multiplication:
\[
\lambda, \mu \rightarrow \lambda \mu
\]

A sesquilinear form is a bilinear map \( r : V \times V \rightarrow \mathbb{C} \). What does it mean?
\[
r(x, y) \text{ is linear in } y \text{ and antilinear in } x:
\]
\[
r(\lambda x + \mu x', y) = \lambda r(x, y) + \mu r(x', y)
\]

E.g. the complex inner product
\[
(z_1, \ldots, z_n) \cdot (w_1, \ldots, w_n) = \sum_{i=1}^{n} \overline{z_i} w_i
\]

Note that if \( r(x, y) \) is sesquilinear then \( y \rightarrow r(x, y) \) is \( \overline{r(y, x)} \). We say \( r \) is Hermitian just when \( r(x, y) = \overline{r(y, x)} \) and skew Hermitian just when \( r(x, y) = -\overline{r(y, x)} \).

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

The matrix \( C \) for a Hermitian \( r \) satisfies \( C = C^\dagger = \overline{C} \).

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

As with last time, we can diagonalize a Hermitian form, getting it to the form \( \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} \). Note that we cannot do more because
\[
r(\lambda x, \lambda x) = |\lambda|^2 r(x, x)
\]

Aside: we have a change of basis \( C' = Q^\dagger C Q \), \( Q \) invertible.

Appendix I: Given \( \begin{pmatrix} I_p & \beta \\ 0 & 0 \end{pmatrix} \) what is the maximal dimension \( B \) is 0 on \( \mathbb{C}^{p+q+1}, \ldots, \mathbb{C}^{p+q+1} \) which is of dimension \( n - (p + q) + \min(p, q) = \min(n - p, n - q) \).

But also, if \( B \) is 0 on \( Z \) then \( P \cap Z = \mathbb{Q} \cap Z = \{0\} \).

So \( P + \dim Z \leq n \), \( q + \dim Z \leq n \), \( \dim Z \leq n - p \).
Appendix II

An abstract approach to symmetric and Hermitian forms.

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

Pick $U, W$ realizing these maxima.

$\beta$ on $U$ is +ve definite and we can find an orthonormal basis

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

$\beta$ 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 $\beta(v, v) > 0$. Then $\beta$ is +ve
definite on $U + \langle v \rangle^\perp$.

Similarly if $\beta(v, v) < 0$ then $\beta$ is -ve definite on $W + \langle v \rangle^\perp$.

So $\beta$ is 0 on $U \oplus W$ and taking a basis for it with the
orthonormal basis gives:

\[
\begin{pmatrix}
I_p & 0 \\
0 & -I_q
\end{pmatrix}
\]
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 \( \langle \cdot, \cdot \rangle \) the inner product.

Remarks

1. An inner product is automatically non-singular. If \( x \in V \) such that \( \langle y, x \rangle = 0 \) for all \( y \), then in particular \( \langle x, x \rangle = 0 \). So \( x = 0 \).

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

3. Parallelogram Law: \( \|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \) holds.

4. The Cauchy-Schwarz Inequality:

\[ |\langle x, y \rangle|^2 \leq \|x\|^2 \|y\|^2 \quad \text{with equality just when } x, y \text{ are independent.} \]

For example, we have \( \langle x-\lambda y, x-\lambda y \rangle \geq 0 \) (equivalently just when \( \frac{x}{x} = \frac{y}{y} \)), i.e. \( \|x\|^2 - \lambda \langle x, y \rangle - \lambda \langle y, x \rangle + |\lambda|^2 \|y\|^2 \geq 0 \).

Assume \( y \neq 0 \) (otherwise the result is trivial) and set \( \lambda = \frac{\langle y, x \rangle}{\|y\|^2} = \frac{\langle x, y \rangle}{\|x\|^2} \) giving the result,

\[ \frac{\|x\|^2 \langle x, y \rangle^2}{\|y\|^2} \leq \|x\|^2 + \frac{\|x\|^2 \langle x, y \rangle^2}{\|y\|^2} \quad \text{with equality just when } x \text{ depends on } y. \]

5. The inner products are determined by the quadratic forms.

\[
\begin{align*}
\text{Real} & \quad \langle x, y \rangle = \frac{1}{2} \left( \|x+y\|^2 - \|x-y\|^2 \right) \\
\text{Complex} & \quad \langle x, y \rangle = \frac{1}{4} \left( \|x+y\|^2 - i \|x+i y\|^2 - \|x-y\|^2 + i \|x-i y\|^2 \right)
\end{align*}
\]
8.2 Orthogonality

We say $x, y$ are orthogonal just when $\langle x, y \rangle = 0$

Clearly $x \perp y \iff y \perp x$. In particular $W^\perp$ and $W$ are
the same.

Vectors $x_1, \ldots, x_k$ are orthogonal iff the $x_i \neq 0$ and
$x_i \perp x_j$ for $i \neq j$. Any such $x_1, \ldots, x_k$ is independent.

For if $\sum x_i x_j = 0$, then $\langle x_i, \sum x_j x_k \rangle = \langle x_i, x_k \rangle = 0$
and so $\langle x, x \rangle = 0$ as $\|x\|^2 \neq 0$.

$e_1, \ldots, e_k$ in an inner product space $V$ are orthonormal if and
only if $\langle e_i, e_j \rangle = \delta_{ij} \neq 0$. This is an orthogonal
set/sequence normalized so its vectors are of unit length.

In Chapter 7 we saw that we can diagonalize symmetric hermitian
forms. If the form is the definite, corresponding matrix is $I$.

Here we have a basis $e_1, \ldots, e_k$, $\langle e_i, e_j \rangle = \delta_{ij}$. Thus any finite dimensional inner product space has an orthonormal basis.

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

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

8.3 Orthogonal Projection

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

Theorem 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 $v = \Pi(v) + (v - \Pi(v))$) is an
orthogonal projection.

3) $\Pi(v)$ is vector in $W$ closest to $v$. 
Proof:

Abstract view $\langle , \rangle$ restricted to $V$, the defining and $w$ non-singular. Hence $w \rightarrow (w \rightarrow \langle w, w \rangle)$ is an isomorphism $W \rightarrow W^*$. For $v \in V$, $(w \rightarrow \langle w, v \rangle) \in W^*$ and we have a linear map $v \rightarrow W^*$. So $(w \rightarrow \langle w, v \rangle) = (w \rightarrow \langle w, \pi(w) \rangle)$ for a unique $\pi(w) \in W$ depending linearly on $w$. We deduce that $\langle w, v - \pi(w) \rangle = 0$ for $w \in W$, so $w - \pi(w) \in W^*$ and $v = \pi(w) + (v - \pi(w))$, so $v = w + W^*$

But if $w \in W \cap W^*$ then $\langle w, w \rangle = 0$ so $w = 0$.

Thus $W \cap W^* = \{0\}$ and $V = W \oplus W^*$. (Not needed, see chapter 7)

$\pi$ is a projection for $v \in V$, $\pi(w) \in W \cap W^*$, so $\pi^2 = \pi$ (projection)

$\pi$ is an orthogonal projection because $v - \pi(v) \in W^* \perp V$.

For 3) take $w \in W$ and consider $\|v - w\|^2$

$\|v - w\|^2 = \|\pi(w) - w) + (v - \pi(v))\|^2$ $\|\pi(w) - w\|^2 + \|v - \pi(v)\|^2$

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

Note We used Pythagorean Theorem. If $x_1, \ldots, x_k$ are pairwise orthogonal then $\|x_1 + \ldots + x_k\|^2 = \|x_1\|^2 + \ldots + \|x_k\|^2$
Linear Algebra

Orthogonal Projections: Alternative View

Suppose \( W \) is a finite dimensional subspace of an inner product space \( V \). Take an orthonormal basis \( e_1, \ldots, e_n \) for \( W \).

Define \( \mathbf{\Pi}(v) = \sum_{i=1}^{n} \langle e_i, v \rangle e_i \in W \). Clearly \( \mathbf{\Pi} \) is linear.

Also, for any \( v \),
\[
\langle e_i, v - \mathbf{\Pi}(v) \rangle = \langle e_i, v \rangle - \langle e_i, \sum_{i=1}^{n} \langle e_i, v \rangle e_i \rangle = \langle e_i, v \rangle - \langle e_i, v \rangle = 0
\]

Thus \( v - \mathbf{\Pi}(v) \in W^\perp \) (plus to a basis for \( W \) and so to all of \( W \))

Note that \( \mathbf{\Pi}(e_i) = \sum_{i=1}^{n} \langle e_i, e_i \rangle e_i = e_i \). So \( \mathbf{\Pi} \) is the identity on \( W \) and \( \mathbf{\Pi}(x) = \mathbf{\Pi}(x) \).

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

Thus so minimum for \( w = \mathbf{\Pi}(v) \)

Bessel's Inequality. Given \( e_1, \ldots, e_n \) an orthonormal sequence, we have
\[
\|x\|_2^2 = \sum_{i=1}^{n} \left( \langle e_i, x \rangle \right)^2
\]

By Pythagoras, \( \|x\|_2^2 = \sum_{i=1}^{n} \left( \langle e_i, x \rangle \right)^2 + \|e_i \|^2 \) and so \( \|x\|_2^2 \geq \sum_{i=1}^{n} \langle e_i, x \rangle^2 \)

Cauchy-Schwarz. By orthogonal projection, suppose \( y \neq 0 \) and write \( x = \sum_{i=1}^{n} \langle e_i, x \rangle e_i \). So
\[
\|x\|_2^2 = \sum_{i=1}^{n} \left( \langle e_i, x \rangle \right)^2 + \sum_{i=1}^{n} \left( \langle e_i, x \rangle \right)^2
\]

So \( \|x\|_2^2 \geq \sum_{i=1}^{n} \langle e_i, x \rangle^2 \) and the result follows.

Equality if \( x = \sum_{i=1}^{n} \langle e_i, x \rangle e_i \)

8.4 Gram-Schmidt Orthogonalization

Given an independent sequence \( x_1, \ldots, x_m \) or \( x_1, x_2, \ldots \) in an inner product space \( V \), we define an orthonormal sequence \( e_1, \ldots, e_m \) or \( e_1, e_2, \ldots \) (together with an auxiliary sequence \( e'_1 \)) inductively as follows:

1. \( e_1 = x_1 \) and \( e_1' = \frac{x_1}{\|x_1\|} \)

2. \( e_m' = x_m - \sum_{i=1}^{m-1} \frac{\langle e_i', x_m \rangle}{\langle e_i', e_i \rangle} e_i \), the component of \( x_m \) in \( e_i' \).

Then \( e_m = \frac{x_m}{\|x_m\|} \).

Note inductively that \( \langle e_1', \ldots, e_k' \rangle = \langle e_1', \ldots, e_k \rangle \).
The last equality follows as in 8.3. \( \langle W, v \rangle = \langle W, v - \Pi(v) \rangle \)
From this it follows that the \( x_i \) are independent, each \( x_i \neq 0 \) and the definition makes sense.
This produces an orthonormal sequence \( e_1, \ldots, e_m \) or \( e_1, e_2, \ldots \) with the property that \( \langle x_1, \ldots, x_k \rangle = \langle e_1, \ldots, e_k \rangle \)

8.5 The Adjoint of an Endomorphism

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

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

It remains to show that \( \alpha^*(x) \) is linear in \( x \).
(Abtractly, the map \( x \mapsto \langle x, \alpha(y) \rangle \) is linear \( \mathbb{V} \to \mathbb{V}^* \) and we compare with a linear isomorphism \( \mathbb{V}^* \to \mathbb{V} \). It is then linear \( \mathbb{V} \to \mathbb{V} \) and so linear \( \mathbb{V} \to \mathbb{V}^* \).)

\( \langle x, \alpha(y) \rangle = \langle x \alpha^*(y), y \rangle \epsilon \mathbb{V}^* \)
\( = \lambda \langle x, \alpha(y) \rangle + \mu \langle x, \alpha(y) \rangle = \lambda \langle x, \alpha^*(y), y \rangle + \mu \langle x^\alpha(y), y \rangle \)
\( = \lambda \langle x, \alpha^*(y), y \rangle + \mu \langle x, \alpha^*(y), y \rangle \)

As this holds for \( \mathbb{V} \), \( \alpha^*(\lambda x + \mu y) = \lambda \alpha^*(x) + \mu \alpha^*(y) \)
Define then \( \alpha^* \) to be the endomorphism adjoint to \( \alpha \).

Easy consequences of the proposition:
\( \alpha^{**} = \alpha \)
\( \alpha^*(\alpha(x))^** = \alpha^*(\alpha(x)) \)

Application An endomorphism \( \alpha \) is orthogonal/unitary iff
\( \langle \alpha(x), \alpha(y) \rangle = \langle x, y \rangle \quad \forall x, y \) (a preserve the inner product)

Note that \( \langle \alpha(x), \alpha(y) \rangle = \langle x, y \rangle \quad \forall x, y \) (a preserve the inner product)
\( \Rightarrow \langle \alpha^*(x), y \rangle = \langle x, y \rangle \quad \forall x, y \) (a preserve the inner product)
\( \Rightarrow \alpha^*(\alpha(x)) = x \quad \forall x \) (a preserve the inner product)
So an orthogonal or unitary endomorphism is one whose adjoint is its inverse.
With inner product spaces we are interested in the matrices for endomorphisms with respect to orthonormal bases. Suppose \( \alpha \) is an endomorphism of \( V \) has matrix \( A = (a_{ij}) \) with respect to an orthonormal basis \( e_1, \ldots, e_n \) and that \( \alpha^* \) has matrix \( A^* = (a_{ij}^*) \).

Then:
\[
\langle \alpha^*(e_i), e_j \rangle = \langle e_i, \alpha(e_j) \rangle
\]

\[
\alpha^*_{ij} = \sum_{k} a_{ik}^* a_{kj} = a_{ij}^*
\]

Thus, \( \alpha^* = a_{ij}^* \) and \( A = A^* \).

**Consequence**

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

Note that \( \alpha^* \) has this property just where it carries orthonormal bases to orthonormal bases.

**Observation** Suppose \( P \) is the change of basis matrix for \( (e_i, e'_j) \), both orthonormal. Then \( P \) is the matrix with respect to \( (e'_i) \) of the map taking \( e'_i \mapsto e_i \). It follows that \( P^* P = I \).
2.6 Symmetric and Hermitian Endomorphisms

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

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

Proposition

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

Proof:

1. Let $\lambda(e) = \lambda e$ with $e \neq 0$. Then:
   $\langle \lambda e, e \rangle = \langle \lambda e, e \rangle = \langle \lambda(e), e \rangle = \langle e, \lambda(e) \rangle = \langle e, \lambda e \rangle = \lambda \langle e, e \rangle$.

So $\lambda = \lambda^*$ as $\langle e, e \rangle \neq 0$.

2. $\alpha$ 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 $\lambda\alpha(t) = \lambda\alpha(t)$.

Theorem

Let $\alpha$ be a symmetric/Hermitian endomorphism of a real/complex finite dimensional inner product space $V$. Then there is an orthonormal basis $e_1, \ldots, e_n$ for $V$ consisting of eigenvectors for $\alpha$.

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 $\mathbf{e}_i$, an eigenvector.

Set $\mathbf{e}_i = \frac{\mathbf{e}_i}{\|\mathbf{e}_i\|}$, an eigenvector of length 1. Now $V = \langle \mathbf{e}_i \rangle + \langle \mathbf{e}_i \rangle^*$

**Claim:** $\langle \lambda \mathbf{e}_i, \mathbf{e}_i^* \rangle : \langle \mathbf{e}_i \rangle^* \rightarrow \langle \mathbf{e}_i \rangle^*$

**Proof:** Suppose $\langle \mathbf{e}_i, \mathbf{v} \rangle = 0$. Then $\langle \mathbf{e}_i, \lambda(\mathbf{v}) \rangle = \langle \lambda \mathbf{e}_i, \mathbf{v} \rangle = \lambda \langle \mathbf{e}_i, \mathbf{v} \rangle = 0 \quad (\lambda = \lambda^*)$

Now $\lambda \mathbf{e}_i$ is itself Hermitian, so by induction hypothesis, there is an orthonormal basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ consisting of eigenvectors.

Then $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 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 diagonal.

**Aside:** $P$ has the property that $\mathbf{e}_k = \sum \lambda_k \mathbf{e}_i$. So the new basis is the columns of $P = P^{-1}$

"So" $P$ and $P^*$ are orthogonal/unitary \[ P^*P = I = PP^* \]

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

8.7 Simultaneous Diagonalization of forms

**Lemma:**

Let $V$ be an inner product space. Then for any symmetric/Hermitian form, $\gamma$ on $V$, there is a unique symmetric/Hermitian endomorphism $\lambda$ such that $\langle \mathbf{v}, \lambda(\mathbf{v}) \rangle = \gamma(\mathbf{v}, \mathbf{v})$.
Proof
Fix $y$. The map $x \mapsto r(x, y) \in V^*$, so there is a unique $\alpha(y)$ with $\langle x, \alpha(y) \rangle = r(x, y)$ $\forall x \in V$.

$\alpha$ is linear:
EITHER $V \rightarrow V^* \xrightarrow{\alpha} V$ is a composite of linear maps $y \mapsto (x \mapsto r(x, y)) \mapsto \alpha(y)$
OR $\langle x, \alpha(y + z) \rangle = \langle x, \alpha(y) + \alpha(z) \rangle = \langle x, \alpha(y) \rangle + \langle x, \alpha(z) \rangle = \langle x, \alpha(y) \rangle + \langle x, \alpha(z) \rangle$.

This holds for all $x$. So $\alpha(y + z) = \alpha(y) + \alpha(z)$.

Finally $\langle x, \alpha(y) \rangle = r(x, y) = r(y, x) = \langle y, \alpha(x) \rangle = \langle x, \alpha(y) \rangle$ and so $\alpha$ is symmetric/Hermitean.

Theorem
Suppose $\beta, \gamma$ are symmetric/Hermitean forms on a real/complex finite dimensional vector space. Further, suppose $\beta$ is positive definite. Then there is a basis with respect to which the form $\beta$ has the matrix $I$, the identity, and $\gamma$ has a diagonal matrix.

Proof
Regard $\beta$ as an inner product on the vector space $V$. Take $\alpha$ to be the symmetric/Hermitean endomorphism representing $\gamma$, in the sense that $r(x, y) = \beta(x, \alpha(y))$.

Take $e_1, ..., e_n$ an orthonormal (with respect to $\beta$) basis of eigenvectors of $\alpha$, say with $\alpha(e_i) = \lambda_i e_i$. Then

$-\beta(e_i, e_i) = \delta_{ii}$, and $\beta$ has matrix $I$.

$-r(e_i, e_i) = b(e_i, a(e_i)) = b(e_i, \lambda_i e_i) = \lambda_i \delta_{ii}$, and $r$ has a diagonal matrix.
Appendix

Suppose there is an orthonormal basis of eigenvectors of $A$. Then $A$ has matrix \( \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \) and $A^*$ has matrix \( \begin{bmatrix} \lambda^* & 0 \\ 0 & \lambda^* \end{bmatrix} \). Evidently $AA^* = A^*A$ and so $\alpha A^* = \alpha^* \alpha$.

Moreover, observe that if $\alpha A = \alpha^* \alpha$ then $e$ is an eigenvector for $A$ with eigenvalue $\lambda$ iff it is also an eigenvector of $A^*$ with eigenvalue $\lambda^*$.

**Proof**: $(\alpha^* - \lambda I)e = 0$ iff $\langle (\alpha^* - \lambda I)e, e \rangle = 0$ and $\langle (\alpha - \lambda I)\alpha^* e, e \rangle = 0$.

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

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

Note in the Theorem of 8.6, the key is the claim $\alpha \langle e_i \rangle^* \rightarrow \langle e_i \rangle^*$.

This still goes through. Suppose $\langle e_i, x \rangle = 0$. Then $\langle e_i, \alpha x \rangle = \langle \alpha^* e_i, x \rangle = \langle A^* e_i, x \rangle = \lambda \langle e_i, x \rangle = 0$.

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

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

This is the case when we diagonalize with entries of all absolute value 1.
8.8 Simultaneous Diagonalization

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:

$$Q^* B Q = I, \quad Q^* C Q = D, \text{ diagonal}$$

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

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

Note that $x$ is an eigenvector for $A$ with eigenvalue $\lambda$ iff $A x = \lambda x$ iff $C x = \lambda B x$.

Note: $Q^* B Q = I$ says that $q_i^* B q_i = \delta_{ij}$, i.e., it says that the $q_i$ are orthonormal with respect to $B$.

$Q^* C Q = D$ says that $q_i^* C q_i = d_i \delta_{ij} = q_i^* d_i B q_i$.

That holds for all $q_i$ iff $C q_i = d_i B q_i$, iff the $q_i$ are eigenvectors for $A$.

Observation 2: We have a method for finding $Q$.

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

With luck, all the eigenspaces are 1-dimensional and we find the $q_i$ 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 $\sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_i x_j$ and $\sum_{i=1}^{n} d_i \sum_{j=1}^{n} p_{ij} x_i x_j$.
Observation 3:
Outline of a purely matrix treatment. Suppose $u \in \ker(C - \lambda B)$
and $v \in \ker(C - \mu B)$, $\lambda \neq \mu$.
$u^* C u = v^* \lambda Bu = \lambda v^* Bu$
$u^* C v = u^* \mu B v = \mu u^* B v = \mu v^* Bu$, Thus $v^* Bu = 0$
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 $C^*$? We can write:
$C^* = (\bigoplus_{i=1}^{k} \ker(C - \lambda_i B)) \oplus W$, where $W$ is the 1 complement with
respect to $B$. Then we see easily that $A = B^{-1} C : W \to W$ and
so if $W \neq \{0\}$ there would exist eigenvalues and $C$ eigenvectors

Worked Example

$2x^2 + 2y^2 - x^2 + 6xy + y^2$
$B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$
$C = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$
$\det(C - tB) = \det \begin{pmatrix} 1 - 2t & 3 \\ 3 & 1 - 2t \end{pmatrix} = 1 - 4t + 4t^2 - 9$
$= 4(t^2 - t - 2) = 4(t-2)(t+1)$

We know the final forms will be $x^2 + y^2$, $2x^2 - y^2$
$\lambda = 2$, $\begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}$. So $(1)$ is an eigenvector, $11 \cdot 11^2 = 4$
$\lambda = -1$, $\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$ so $(1)$ is an eigenvector, $11 \cdot 11^2 = 4$
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 \(2(x+y)^2 - (x-y)^2\).

**Remark**

In Chapter 7, we thought of rank and signature in terms of +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.

**Appendix**

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

\[ \chi_{\alpha}(t) = \text{Ker}(\alpha - \lambda I) + (\text{Ker}(\alpha - \lambda I))^\perp = U + W \]

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

Hence the matrix of \( \alpha \) has the form \( \begin{pmatrix} \lambda I & 0 \\ 0 & A_{\infty} \end{pmatrix} \) and

\[ \chi_{\alpha}(t) = \chi_{A_{\infty}}(t) \chi_{A_{\infty}}(t) = (\lambda - t)^k \chi_{A_{\infty}}(t) \]

But \((\lambda - t)^k \chi_{A_{\infty}}(t)\) otherwise \( \lambda \) would be an eigenvalue for \( \alpha \) in \( W \) and there would be an eigenvector for \( \lambda \) in \( W \).