Linear Algebra 2 Notes

by Tyler Wright

github.com/Fluxanoia fluxanoia.co.uk

These notes are not necessarily correct, consistent, representative of the course as it stands today, or rigorous. Any result of the above is not the author's fault.

These notes are marked as unsupported, they were supported up until January 2020.

0 Notation

We commonly deal with the following concepts in Linear Algebra 2 which I will abbreviate as follows for brevity:

Term	Notation
Additive identity of set X	0_X
Multiplicative identity of a set X	1_X
Set of linear maps from V to W	$\mathcal{L}(V, W)$
$\mathcal{L}(V, V)$	$\text{End}(V)$

Contents

1 Groups, Rings, and Fields

1.1 Groups

A group is a set G combined with a group operation $\circ : G \times G \to G$ such that:

- **Associativity**, for all g, h, j in G, $g(hj) = (gh)j$,
- **Identity**, there exists e in G such that $eg = ge = g$ for all g in G
- **Inverses**, for all g in G, there exists g^{-1} in G such that $gg^{-1} = g^{-1}g = e$ where e is the identity of G.

Note that here we have implicitly used the group operation $\circ$.

1.1.1 Subgroups

For a group $\mathcal{G} = (G, \circ)$, we have that $\mathcal{G}' = (G', \circ)$ is a subgroup of $\mathcal{G}$ if and only if $G' \subseteq G$ and $\mathcal{G}'$ is a group.

1.1.2 Group Homomorphisms

A homomorphism between two groups G, H is a function $f : G \to H$ such that $f(gh) = f(g)f(h)$ for all g, h in G.

1.1.3 Properties of Group Homomorphisms

We can derive some properties of homomorphisms, for G, H groups, and $f : G \to H$ a homomorphism:

- The image of the identity in G is the identity in H,
- $\mathrm{Ker}(f)$ is a subgroup of G,
- $\mathrm{Im}(f)$ is a subgroup of H.

1.2 Rings

A ring with unity is a set R along with an addition map $+$, and a multiplication map $\circ$ where $+, \circ : R \times R \to R$ such that:

- $(R, +)$ is an abelian group (of which the identity is called zero),
- The multiplication operation is associative,
- The multiplication operation has a two-sided identity not equal to the zero identity (called one),
- For all a, b, c in R, $a(b + c) = ab + ac$ and $(a + b)c = ac + bc$.

A ring is commutative if the multiplication operation is commutative.

1.2.1 Subrings

For a ring $\mathcal{R} = (R, +, \circ)$, we have that $\mathcal{R}' = (R', +, \circ)$, is a subring of $\mathcal{R}$ if and only if $R' \subseteq R$ and $\mathcal{R}'$ is a ring.

1.2.2 Ring Homomorphisms

For rings with unity R and S, $f : R \to S$ is a ring homomorphism if for all a, b in R:

$$\begin{aligned} f(a + b) &= f(a) + f(b) \\ f(ab) &= f(a)f(b) \\ f(1_R) &= 1_S. \end{aligned}$$

1.3 Fields

A field K is a ring with unity where $(K \setminus \{0\}, \circ)$ is an abelian group.

1.3.1 Characteristic of a Field

For a field K, the field characteristic $\mathrm{char}(K)$ is the smallest positive integer n such that:

$$n \cdot 1 = \sum_{i=1}^{n} 1 = 1 + 1 + \ldots + 1 = 0,$$

or zero if no such value n exists.

Field Characteristics being Prime The characteristic of a field K must be prime (or zero) because if for some a, b integers $\mathrm{char}(K) = ab$ then:

$$0 = \mathrm{char}(K) \cdot 1 = (a \cdot 1)(b \cdot 1),$$

which means $a \cdot 1$ or $b \cdot 1$ is zero so a or b is the characteristic of K.

1.3.2 Algebraic Closure of Fields

A field K is called algebraically closed if all non-constant polynomials with coefficients in K also has a root in K.

2 Vector Spaces

A vector space over a field K is a set V with an addition operation $+ : V \times V \to V$ and a scalar multiplication operation $\circ : K \times V \to V$ such that for all a, b in K and v, w in V:

- $(V, +)$ is an abelian group,
- $1_K \circ v = v$,
- $(ab) \circ v = a \circ (b \circ v)$
- $(a + b) \circ v = a \circ v + b \circ v$
- $a \circ (v + w) = a \circ v + a \circ w$.

2.1 Subspaces

For V a vector space over the field K and W a set, W is a subspace of V if and only if it is a subset of V and is a vector space with respect to the addition and scalar multiplication defined by V. It is sufficient to verify that W is closed under addition and multiplication.

2.2 Linear Combinations of Vectors

For a set V with addition operation $+$, a field K and n in $\mathbb{N}$, a linear combination of $v_1, \ldots, v_n$ in V is:

$$\sum_{i=1}^{n} a_i \cdot v_i = a_1 \cdot v_1 + \cdots + a_n \cdot v_n,$$

for some $a_1, \ldots, a_n$ in K. Such a combination is trivial if each of $a_1, \ldots, a_n$ are zero and non-trivial otherwise.

2.3 Linear Independence

For a vector space V and $W \subseteq V$, we say W is linearly independent if there does not exists a non-trivial linear combination of all the vectors in W equal to zero (and linearly dependent otherwise).

2.3.1 Properties of Linear Independence

For a vector space V with $W \subseteq V$:

- W is linearly dependent if it contains 0_V,
- If W linearly independent, any subset of it is also linearly independent,
- If there's a linearly dependent subset of W, then W is linearly dependent.

2.4 The Span of a Set

For a set V with addition operation $+$ and a field K, the span of $W \subseteq V$ is the set of all the linear combinations of the values in W denoted by $\text{span}(W)$.

2.5 Bases

For a vector space V with $W \subseteq V$, if W is linearly independent and $\text{span}(W) = V$, we say that W is a basis of V. It is a minimal spanning set.

2.5.1 Properties of Bases

We have that if a basis is finite, all other bases have the same size. Additionally, saying W is a basis is equivalent to saying that each vector in V can be **uniquely** written as a linear combination of vectors in W.

2.6 Dimension

For a vector space V with a finite basis, we say that the size of the basis is the dimension of V denoted by $\dim(V)$. By convention, $\dim(\{0_V\}) = 0$. Vector spaces with identical dimension are isomorphic.

2.6.1 Dimension and Subsets

For $V \neq \{0\}$ a vector space, if there is a finite spanning set S of V then:

- V is finite dimensional, particularly, there is a basis B of V where $B \subseteq S$,
- For $X \subseteq V$ such that X is linearly independent, X can be extended to a basis of V,
- All subspaces of V are finite-dimensional.

3 Linear Maps

Let V, W be vector spaces over a field K, we have that $f : V \to W$ is a linear map if for all a, b in K and u, v in V:

$$f(au + bv) = af(u) + bf(v).$$

A bijective linear map is called an isomorphism. If $f : V \to W$ is an isomorphism, we say that V and W are isomorphic, denoted by $V \cong W$.

3.1 Properties of Linear Maps

For a bijective linear map $f : V \to W$, the inverse of f is also linear and if $V = W$, (f is a linear operator) then injectivity or surjectivity imply f is an isomorphism.

3.2 Nilpotence

For a field K, a finite dimensional vector space V over K, with $f : V \to V$ a linear map, we have that f is nilpotent if there exists r in $\mathbb{Z}_{\geq 0}$ such that f^r is the zero map.

3.3 The Rank-Nullity Theorem

For V, W finite-dimensional vector spaces and $f : V \to W$ a linear map, we define the rank and nullity:

$$\begin{aligned} \mathrm{rank}(f) &:= \dim(\mathrm{Im}(f)) \\ \mathrm{nullity}(f) &:= \dim(\mathrm{Ker}(f)), \end{aligned}$$

and we have that:

$$\dim(V) = \mathrm{rank}(f) + \mathrm{nullity}(f).$$

Proof. We have that $\mathrm{Ker}(f)$ is a subspace of V and by the finite-dimensionality of V we have that $\mathrm{Ker}(f)$ is also finite-dimensional, so we take a basis $B_K = \{v_1, \ldots, v_k\}$ of $\mathrm{Ker}(f)$ where $k = \dim(\mathrm{Ker}(f)) = \mathrm{nullity}(f)$. We extend B_K with the linearly independent set $B_I = \{v_{k+1}, \ldots, v_{k+i}\}$ to a basis of V where $i = \dim(V) - k$. Thus, $B = B_K \cup B_I$ is a basis of V (partitioned by B_K, B_I). So, $\mathrm{Im}(f) = \mathrm{span}(f(B))$ as B is a basis but:

$$\begin{aligned} f(B) &= \{f(v_1), \ldots, f(v_k), f(v_{k+1}), \ldots, f(v_{k+i})\} \\ &= \{0_W, \ldots, 0_W, f(v_{k+1}), \ldots, f(v_{k+i})\} \\ &= f(B_I), \end{aligned}$$

as $B_K \subseteq \mathrm{Ker}(f)$. So, $\mathrm{Im}(f) = \mathrm{span}(f(B_I))$. We have that B_I must be linearly independent as it's part of our basis B so $f(B_I)$ must also be linearly independent. Thus, $f(B_I)$ is a basis for $\mathrm{Im}(f)$, so $\mathrm{rank}(f) = \dim(\mathrm{Im}(f)) = |f(B_I)| = |B_I| = i$, thus:

$$\mathrm{rank}(f) + \mathrm{nullity}(f) = i + k = |B| = \dim(V).$$

$\square$

4 Matrices

Let m, n be in $\mathbb{Z}_{>0}$ and let K be a field. An $m \times n$ matrix with entries in K is a map $M : [m] \times [n] \to K$, more commonly written as $M = (a_{ij})$ representing the rectangular array of values held by M.

4.1 Types of Matrices

For m, n in $\mathbb{Z}_{>0}$ and K a field, let M be in $M_{m\times n}(K)$. We have the following types of matrix:

- **Square**: where $m = n$
- **Upper Triangular**: if $a_{ij} = 0$ for $i > j$
- **Lower Triangular**: if $a_{ij} = 0$ for $i < j$
- **Diagonal**: if $a_{ij} = 0$ for $i \neq j$
- **Symmetric**: if $a_{ij} = a_{ji}$
- **Anti-symmetric**: if $a_{ij} = -a_{ji}$.

4.2 The Space of Matrices

For m, n in $\mathbb{Z}_{>0}$ and K a field, we define the set of all $m \times n$ matrices over K by $M_{m\times n}(K)$. We have that $M_{m\times n}(K)$ is a vector space over K where matrices are added and multiplied by scalars component-wise. So, for $M_1 = (a_{ij}), M_2 = (b_{ij})$ in $M_{m\times n}(K)$ and c in K we have:

$$\begin{aligned} cM_1 &= (ca_{ij}) \\ M_1 + M_2 &= (a_{ij} + b_{ij}). \end{aligned}$$

Additionally, the zero vector is $M_0 = (0_K)$, the multiplicative identity is the diagonal matrix of all 1_K's and, the vector space has a basis consisting of M_{ij} where all entries are zero except the $(i, j)^{\text{th}}$ entry. This leads to the conclusion that the dimension is mn and thus that $M_{m\times n} \cong K^{mn}$.

4.3 Matrix Multiplication

For a, b, c in $\mathbb{Z}_{>0}$ and a field K, we can define the multiplication of the two matrices $X = (x_{ij})$ in $M_{a\times b}(K)$ and $Y = (y_{ij})$ in $M_{b\times c}(k)$ as follows:

$$XY := (\sum_{k=1}^{b} x_{ik}y_{kj}).$$

This operation is not commutative in general but is associative.

For A, B in $M_n(K)$, we have that AB is also in $M_n(K)$. This, along with matrix addition, makes M_n a ring with unity with multiplicative identity $I_n := (\delta_{ij})$. However, there exists non-zero A, B in M_n such that $AB = 0$ so, M_n is not a field.

4.4 Matrices of Linear Maps

For V, W vector spaces over a field K, we have $A = \{v_1, \dots, v_n\}$, $B = \{w_1, \dots, w_n\}$ ordered bases for V and W respectively. Given f in $\mathcal{L}(V, W)$, the matrix associated to f (with respect to the bases A and B) is the $m \times n$ matrix:

$$M_{BA}(f) = (a_{ij}),$$

where we define a_{ij} by:

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

for each j in $[n]$.

4.4.1 Matrices of Composed Linear Maps

For U, V, W vector spaces over a field K, for some l, m, n in $\mathbb{Z}_{>0}$ we have $A = \{u_1, \dots, u_l\}$, $B = \{v_1, \dots, v_m\}$, $C = \{w_1, \dots, w_n\}$ bases for U, V, W respectively. Given f in $\mathcal{L}(U, V)$, g in $\mathcal{L}(V, W)$, we have:

$$M_{CA}(g \circ f) = M_{CB}(g)M_{BA}(f).$$

4.5 Transition Matrices

For a finite n-dimensional vector space V with bases A, A' and the n-dimensional identity I, we call $M_{A'A}(I) = C_{A'A}$ the transition matrix from A to A'. We have that $C_{A'A}$ is invertible and $C_{A'A}^{-1} = C_{AA'}$.

4.6 Matrix Transitions

For a finite-dimensional vector space V with bases A, B and $f : V \to V$ a linear operator:

$$\begin{aligned} M_{BB}(f) &= C_{AB}^{-1} M_{AA}(f) C_{AB} \\ &= C_{BA} M_{AA}(f) C_{AB}. \end{aligned}$$

4.7 Similar Matrices

For matrices A', A, we say that A' and A are similar if there exists an invertible matrix C such that:

$$A' = C^{-1} A C.$$

This is denoted by $A' \sim A$. Similarity forms an equivalence relation on the space of square matrices. If we have $A \sim A'$ and A represents some linear operator f for some basis B, then we have that for some basis B', f has matrix A'.

5 Eigenspaces and Root Spaces

5.1 Root Vectors

For a finite dimensional vector space V over K, with $f : V \to V$, and λ in K, we have that v in V is a λ-root vector of f if there exists n in $\mathbb{N}$ such that:

$$(f - \lambda(\mathrm{id}))^n(v) = 0_V.$$

The smallest such n is the height of v denoted by $h(v)$.

5.1.1 Root Spaces

The set $V(\lambda)$ is the set of all root vectors corresponding to λ called the root space of λ under f in V.

5.1.2 Properties of the Root Space

We have the following properties:

1. $V(\lambda)$ is a subspace,
2. $V(\lambda) \neq \{0_V\}$ if and only if λ is an eigenvalue of f,
3. $V(\lambda)$ is f-invariant.

Proof. (1) By the linearity of f, a multiple of a vector in $V(\lambda)$ is also in $V(\lambda)$. If we have the two vectors v_1, v_2 in $V(\lambda)$ with heights n_1, n_2 respectively, set $n = \max(n_1, n_2)$:

$$(f - \lambda(\mathrm{id}))^n(v_1 + v_2) = 0_V.$$

Thus, $V(\lambda)$ is a subspace. □

Proof. (2) Suppose λ is an eigenvalue of f then $V(\lambda)$ contains the corresponding eigenvector(s) and thus, is non-zero. Supposing $V(\lambda)$ is non-zero, we choose $v \neq 0_V$ in $V(\lambda)$, it has some height n, $(f - \lambda(\mathrm{id}))^{n-1}(v)$ is an eigenvector. □

Proof. (3) For some vector v in $V(\lambda)$ with height n, $(f - \lambda(\mathrm{id}))^n(v) = 0_V$ so:

$$\begin{aligned}(f - \lambda(\mathrm{id}))^n(f(v)) &= (f - \lambda(\mathrm{id}))^n(f(v) - \lambda v + \lambda v) \\ &= (f - \lambda(\mathrm{id}))^{n+1}(v) + \lambda(f - \lambda(\mathrm{id}))^n(v) \\ &= 0_V.\end{aligned}$$

Thus, $V(\lambda)$ is f-invariant. □

5.1.3 Primary Decomposition Theorem

For a finite dimensional vector space V over K (algebraically closed) we have that:

$$V = \bigoplus_{i \in [k]} V(\lambda_i),$$

the internal direct sum where $\{\lambda_1, \dots, \lambda_k\}$ is the set of distinct eigenvalues of a linear operator f in $L = \mathcal{L}(V, V)$.

Proof. Let $p_f = \prod_{i=1}^{k} (\lambda_i - x)^{m_i}$ where m_i is the algebraic multiplicity of λ_i. Take F_i, f_i, V_i defined as follows:

$$\begin{aligned} F_i(x) &= p_f(x)(x - \lambda_i)^{-m_i} \\ f_i(x) &= F_i(f)(x) \\ V_i &= \operatorname{Im}(f_i). \end{aligned}$$

We show that $V_i \subseteq \operatorname{Ker}(f - \lambda_i(\mathrm{id}))^{m_i}$ by the Cayley-Hamilton theorem:

$$0_L = p_f(f) = (f - \lambda_i(\mathrm{id}))^{m_i} F_i(f) = (f - \lambda_i(\mathrm{id}))^{m_i} f_i,$$

so $(f - \lambda_i(\mathrm{id}))^{m_i}$ applied to anything in the image of f_i must be zero, so $V_i \subseteq \operatorname{Ker}(f - \lambda_i(\mathrm{id}))^{m_i}$. Consequently, $V_i \subseteq V(\lambda_i)$.

Then, we show that $V = V_1 + \cdots + V_k$. The highest common factor of $F_1, \dots, F_k$ must be $1_{K[x]}$ so we have there are polynomials $X_1, \dots, X_k$ in $K[x]$ such that for any v in V:

$$\begin{aligned} &\sum_{i=1}^{k} F_i(x) X_i(x) = 1_{K[x]} \\ \Rightarrow &\sum_{i=1}^{k} F_i(f) X_i(f) = \mathrm{id} \\ \Rightarrow &\sum_{i=1}^{k} [F_i(f) X_i(f)](v) = \sum_{i=1}^{k} f_i(X_i(f)(v)) = v, \end{aligned}$$

writing v as the sum of elements in $\operatorname{Im}(f_i) = V_i$ for each i in $[k]$. Thus, $V = V_1 + \cdots + V_k$.

Now, we show that $V = V_1 \oplus \cdots \oplus V_k$ by showing that for each i in $[k]$:

$$I_i = V_i \bigcap \left[\sum_{i \neq j \in [k]} V_j \right] = \{0_V\}.$$

We consider v in I_i. We have that v is in V_i so $(f - \lambda_i(\mathrm{id}))^{m_i}(v) = 0_V$ and:

$$F_i(f)(v) = \pm \prod_{i \neq j \in [k]} (f - \lambda_j(\mathrm{id}))^{m_j}(v) = 0_V, \tag{1}$$

as v is in $\sum_{i \neq j \in [k]} V_j$. We can see that $(x - \lambda_i)^{m_i}$ and $F_i(x)$ are relatively prime by definition so there are polynomials X and Y in $K[x]$ such that:

$$X(x)(x - \lambda_i)^{m_i} + Y(x)F_i(x) = 1_{K[x]}. \tag{2}$$

Putting this all together, we apply f and v to the above we get:

$$\begin{aligned} v &= [X(f)(f - \lambda_i(\mathrm{id}))^{m_i} + Y(f)F_i(f)](v) && \text{(by (2))} \\ &= X(f)(0_V) + Y(f)(0_V) && \text{(by (1))} \\ &= 0_V. \end{aligned}$$

So, $V = V_1 \oplus \cdots \oplus V_k$.

Finally, we show $V_i = V(\lambda_i)$. We already know that $V_i \subseteq V(\lambda_i)$, so we just have to show that $V(\lambda_i) \subseteq V_i$. Take v in $V(\lambda_i)$ and write it as $v = v_1 + v_2$ where v_1 is in V_1 and v_2 is in $V \setminus V_1$. There is some m such that $(f - \lambda_i)^m v_2 = 0_V$ as:

$$\begin{aligned} v &= v_1 + v_2 \in V(\lambda_i) \\ \Rightarrow v_2 &= v - v_1 \in V(\lambda_i) \end{aligned}$$

as v is in $V(\lambda_i)$ and v_1 is in $V_i \subseteq V(\lambda_i)$. We also have that $F_i(f)(v_2) = 0_V$ as v_2 is in $V(\lambda_i)$. Since $(x - \lambda_i)^m$ and $F_i(x)$ are relatively prime, there are polynomials p and q such that:

$$p(x)(x - \lambda_i)^m + q(x)F_i(x) = 1_{K[x]} \Rightarrow \qquad p(f)(f - \lambda_i(\mathrm{id}))^m + q(x)F_i(f) = \mathrm{id},$$

so, applying this to v_2 gives:

$$\begin{aligned} v_2 &= [p(f)(f - \lambda_i(\mathrm{id}))^m + q(x)F_i(f)](v_2) \\ &= p(f)(0_V) + q(x)(0_V) \\ &= 0_V, \end{aligned}$$

so $v = v_1$ (in $V_i \subseteq V(\lambda_i)$). Thus, $V_i = V(\lambda_i)$ as required. □

5.2 Eigenvectors

For a vector space V over K with $f : V \to V$ a linear operator, a non-zero vector v in V is an eigenvector if $f(v) = \lambda v$ for some λ in K which is called the eigenvalue corresponding to v.

In particular, v is a root vector of height 1.

5.2.1 Eigenspaces

For a vector space V over K with $f : V \to V$ a linear operator and some eigenvalue λ, we define the eigenspace of λ as the set of eigenvectors with eigenvalue λ.

This is denoted by $E(\lambda)$ and $E(\lambda) \cup \{0_V\}$ forms a subspace of V. The dimension of $E(\lambda)$ is the geometric multiplicity of λ.

5.2.2 Nilpotent Maps on Eigenspaces

Let K be a field, V be a finite dimensional vector space over K, f be in $\mathcal{L}(V, V)$, and λ be an eigenvalue of f. Let $g : E(\lambda) \to E(\lambda)$ be $g(v) = f(v)$ for v in $V(\lambda)$. We have that $(g - \lambda \mathrm{id})$ is nilpotent.

5.2.3 Eigenvalues on Nilpotent Maps

Let K be a field, V be a finite dimensional vector space over K, f be in $\mathcal{L}(V, V)$, and λ be an eigenvalue of f. We have that if f is nilpotent, 0_K is the only eigenvalue of f. Thus, for some m, m' in $\mathbb{Z}_{>0}$ with $m' \leq m$:

- $p_f(x) = x^m$,
- $m_f(x) = x^{m'}$,
- $V = V(0)$, the zero root space of f.

5.2.4 Multiplicity

For V a vector space over K with $f : V \to V$ a linear map and λ an eigenvalue in K of f. We have that the algebraic multiplicity of λ is the multiplicity of λ in p_f. The geometric multiplicity of λ is $\dim(E(\lambda))$.

6 Direct Sums and Projections

6.1 Direct Sums

For V, W vector spaces where, we define the external direct sum of V and W as:

$$V \oplus W := \{(v, w) : v \in V, w \in W\},$$

with zero vector $(0_V, 0_W)$. We define addition and scalar multiplication coordinate-wise.

For S, T subspaces of V, V is the internal direct sum of S and T if $S + T = V$ and $S \cap T = \{0_V\}$.

6.1.1 Bases of External Direct Sums

We have that for $B_V = \{v_1, \ldots, v_k\}$, $B_W = \{w_1, \ldots, w_l\}$ bases for V and W respectively:

$$B = \{(v_1, 0_W), \ldots, (v_k, 0_W), (0_V, w_1), \ldots, (0_V, w_l)\},$$

is a basis for $V \oplus W$. Thus, $\dim(V \oplus W) = \dim(V) + \dim(W)$.

6.1.2 The Addition Map for Direct Sums

For V, W subspaces of a vector space U over K, and $f : V \oplus W \to U$ defined by:

$$f((v, w)) = v + w,$$

we have that:

1. f is linear,
2. f is injective if and only if $V \cap W = \{0_U\}$,
3. f is surjective if and only if $V \cup W$ spans U.

Proof. (1) Immediate from the definition. □

Proof. (2) For v in V, w in W and u in $V \cap W$, suppose f is injecive:

$$f((u, -u)) = u + (-u) = 0_U,$$

thus $u = 0_U$ by injectivity. Suppose $V \cap W = \{0_U\}$, if $f((v, w)) = 0_U$:

$$0_U = f((v, w)) = v + w,$$

so $v = -w$ and thus, they are both zero as the intersection of V and W is just 0_U. So, $\mathrm{Ker}(f) = \{0_U\}$. □

Proof. (3) The image of f is just $V + W$, if $V \cup W$ spans U then $V + W$ must equal U. If f is surjective, $\mathrm{Im}(f) = V + W = U$ so $V \cup W$ spans U. □

6.1.3 Consequences of Internal Direct Sums

For $V, W \subseteq U$, where $V \oplus W$ is an internal direct sum, we have that each element in U can be written uniquely as the sum of elements in V and W and the addition map in (6.1.2) is an isomorphism.

6.2 Projections

For $V, W \subseteq U$, where $V \oplus W$ is an internal direct sum, the projection onto V along W is the linear operator $P_{V,W} : U \to U$ where:

$$P_{V,W}(u) = v,$$

where $u = v + w$ for some unique v in V and w in W. We can see that $P_{V,W}^2 = P_{V,W}$ (it is idempotent).

6.2.1 Idempotence and Projections

For a linear operator $E : U \to U$, if E is idempotent ($E^2 = E$) then E is a projection.

Proof. Take u in U, we have that:

$$u = E(u) + (u - E(u)).$$

We see $E(u)$ is in $\mathrm{Im}(E)$ and $u - E(u)$ is in $\mathrm{Ker}(E)$ as:

$$E(u - E(u)) = E(u) - E^2(u) = E(u) - E(u) = 0.$$

Thus, $E = P_{\mathrm{Im}(E),\mathrm{Ker}(E)}$ because if u is in $\mathrm{Ker}(E)$ then $E(u) = 0$ and if u is in $\mathrm{Im}(E)$ then:

$$E(u) = E^2(v) = E(v) = u.$$

Thus, U is the internal direct sum of $\mathrm{Im}(E)$ and $\mathrm{Ker}(E)$. $\square$

6.3 f-invariance

For a vector space V with $U \subseteq V$ a subspace and $f : V \to V$ a linear map, we have that U is f-invariant if for all u in U we have $f(u)$ in U.

6.3.1 Matrices of Linear Maps (using f-invariance)

For $U, W \subseteq V$ subspaces such that $V = U \oplus W$, let B_U, B_W be finite bases of U and W respectively. For a linear operator $f : V \to V$ such that U and W are f-invariant, we have that the matrix with respect to the basis $B = B_U \cup B_W$ of f has the following block form:

$$M_{BB}(f) = \begin{pmatrix} M_{B_V B_V}(f) & 0 \\ 0 & M_{B_W B_W}(f) \end{pmatrix}.$$

7 Quotient Spaces

For a vector space V over K with $W \subseteq V$ a subspace, we define an equivalence relation on V by declaring:

$$v_1 \sim v_2 \text{ if } v_1 - v_2 \in W.$$

The set of equivalence classes is called the quotient of V by W and is denoted by V/W. For some v in V, we denote the class containing v by $v + W$ (similarly to cosets in Introduction to Group Theory). So, we have:

$$\begin{aligned} v + W &= \{v' \in V : v \sim v'\} = \{v' \in V : v - v' \in W\} \\ V/W &= \{v + W : v \in V\}, \end{aligned}$$

with addition and multiplication defined for v_1, v_2 in V and a in the field:

$$\begin{aligned} (v_1 + W) + (v_2 + W) &= (v_1 + v_2) + W \\ a(v_1 + W) &= av_1 + W. \end{aligned}$$

7.1 Understanding the Quotient Space

For a vector space V over K with W a subspace, consider w in W:

$$\begin{aligned} w + W &= \{v \in V : w - v \in W\} \\ &= W \\ &= \{v \in V : 0_V - v \in W\} \\ &= 0_V + W. \end{aligned} \tag{3}$$

So, $W = w + W$ is $0_{V/W}$. Consequently, for some v in V, we can see that:

$$\begin{aligned} (v + W) + (w + W) &= (v + w) + W \\ &= \{v' \in V : (v + w) - v' \in W\} \\ &= \{v' \in V : v - v' \in W\} \\ &= v + W. \end{aligned}$$

Finally, we can see that $v+W$ is the set of vectors in V such that v and each element in $v + W$ differ by some element in W. So, we are effectively mapping the span of W with the origin at v to v, 'collapsing' W.

7.2 Linear Map to the Quotient Space

For a vector space V over K with W a subspace, we can define $\pi : V \to V/W$ for some v in V by $f(v) = v + W$. We have that:

1. π is linear and surjective,
2. $\operatorname{Ker}(\pi) = W$.

Proof. (1) For each v, v' in V and k, k' in K:

$$\begin{aligned}\pi(kv + k'v')) &= (kv + k'v') + W \\ &= (kv + W) + (k'v' + W) \\ &= k(v + W) + k'(v' + W) \\ &= k\pi(v) + k'\pi(v'),\end{aligned}$$

so π is linear. Also, $v + W = \pi(v)$ so π is surjective. □

Proof. (2) By (3) in (7.1), we can see that for each w in W:

$$\pi(w) = w + W = 0_V + W = 0_{V/W}.$$

So, $W \subseteq \operatorname{Ker}(\pi)$. For each v in $\operatorname{Ker}(\pi)$:

$$\begin{aligned}\pi(v) &= 0_{V/W} \\ &= \{v' \in V : 0_V - v' \in W\} \\ &= W.\end{aligned}$$

So, $\operatorname{Ker}(\pi) \subseteq W$. Thus, $\operatorname{Ker}(\pi) = W$. □

7.3 Isomorphisms formed by Linear Maps

For V, W vector spaces and $f : V \to W$ a linear map, we have an isomorphism $\operatorname{Im}(f) \cong V/\operatorname{Ker}(f)$.

Proof. We define $\tilde{f} : V/\operatorname{Ker}(f) \to \operatorname{Im}(f)$ by $(v + \operatorname{Ker}(f)) \mapsto f(v)$. We first check that for some v, v' in V such that $v \sim v'$, $\tilde{f}(v + \operatorname{Ker}(f)) = \tilde{f}(v' + \operatorname{Ker}(f))$ as $v + \operatorname{Ker}(f) = v' + \operatorname{Ker}(f)$:

$$\begin{aligned}\tilde{f}(v + \operatorname{Ker}(f)) - \tilde{f}(v' + \operatorname{Ker}(f)) &= f(v) - f(v') \\ &= f(v - v') \\ &= 0_V,\end{aligned}$$

as $v \sim v'$ so $v - v'$ is in $\operatorname{Ker}(f)$. We have for each v in $\operatorname{Im}(V)$ there is a w in W such that $f(w) = v$, so $\tilde{f}(w + \operatorname{Ker}(f)) = f(w) = v$, thus $\tilde{f}$ is surjective. Taking v in V, suppose $\tilde{f}(v + \operatorname{Ker}(f)) = 0_V$, so $f(v) = 0_V$, thus $v \in \operatorname{Ker}(f)$ and $v + \operatorname{Ker}(f) = 0_{V/\operatorname{Ker}(f)}$ so $\tilde{f}$ is injective. Thus, $\tilde{f}$ is an isomorphism. □

7.4 Linear Operators on the Quotient Space

For a vector space V with W a subspace and a linear operator $f : V \to V$, there exists a well-defined operator $\bar{f} : V/W \to V/W$; $v + W \mapsto f(v) + W$ if and only if W if f-invariant. We call this the induced operator on V/W.

Proof. $\bar{f}$ is well-defined if and only if for all v, v' in V such that $v + W = v' + W$:

$$f(v) + W = f(v') + W. \tag{4}$$

We have that v' is in $v + W$ as $v + W = v' + W$ and $v' - v' = 0_V$ which is in W so v' is in $v' + W$. Considering (4):

$$\begin{aligned} f(v) + W - f(v') + W &= (f(v) - f(v')) + W \\ &= f(v - v') + W. \end{aligned}$$

We know that $\bar{f}$ is well-defined if and only if $f(v - v') + W$ is zero (so $f(v - v')$ is in W). We know that $v - v'$ is in W as v' is in $v + W$, thus $\bar{f}$ is well-defined if and only if W is f-invariant. Linearity is immediate. □

7.5 Matrices formed using Quotient Spaces

For a finite-dimensional vector space V and $f : V \to V$ a linear operator with W an f-invariant subspace of V, suppose we have B_W a basis for W, that we extend to a basis B of V. Take the set Q:

$$Q = \{v + W : v \in B \setminus B_W\},$$

a basis of V/W and we can form a matrix in block form:

$$M_{BB}(f) = \begin{pmatrix} M_{B_W B_W}(f) & * \\ 0 & M_{QQ}(\bar{f}) \end{pmatrix},$$

where $\bar{f}$ is the induced operator on V/W and $*$ marks the area of the matrix which we cannot determine.

8 Dual Spaces

For V a vector space over K, we have that the dual space V^* is $\mathcal{L}(V, K)$, the set of linear maps from V to K. We have that addition and scalar multiplication are defined for some v in V, f, g in V^*, and a in K:

$$\begin{aligned}(f+g)(v) &:= f(v) + g(v),\\ (af)(v) &:= af(v).\end{aligned}$$

8.1 Dual Bases

For V a finite-dimensional vector space over K, with $\dim(v) = n$ and a basis $B = \{v_1, \dots, v_n\}$. We define the dual basis $B^* = \{v_1^*, \dots, v_n^*\}$ by defining $v_i^* : V \to K$ as the unique linear map such that:

$$v_i^*(v_j) := \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Equivalently, where we decompose v in V into the vectors in B with unique coefficients $a_1, \dots, a_n$ in K, we have that:

$$v_i^*(v) = v_i^*\left(\sum_{j=1}^{n} a_j v_j\right) = \sum_{j=1}^{n} a_j v_i^*(v_j) = a_i.$$

B^* is a basis for V^* and V and V^* are isomorphic by the map sending v_i to v_i^* for each i in $[n]$.

Proof. Suppose we have $a_1, \dots, a_n$ in K such that:

$$\sum_{i=1}^{n} a_i v_i^* = 0_{V^*},$$

thus for each i in $[n]$:

$$0_K = \left(\sum_{i=1}^{n} a_i v_i^*\right)(v_i) = a_i.$$

So, each a_i is zero and thus B^* is linearly independent. Taking some f in V^* and v in V decompose into vectors in B with coefficients $a_1, \dots, a_n$:

$$\begin{aligned} f(v) &= a_1 f(v_1) + \cdots + a_n f(v_n) \\ &= v_1^*(v) f(v_1) + \cdots + v_n^*(v) f(v_n) \\ &= [f(v_1) v_1^* + \cdots + f(v_n) v_n^*](v), \end{aligned}$$

thus B^* spans V^*. So, B^* is a basis. $\square$

8.2 The Annihilator

For V a vector space over K with $S \subseteq V$, the annihilator of S is the subspace S^0 of V^* where for f in S^0, $S \subseteq \mathrm{Ker}(f)$ (or rather, for all s in S, $f(s) = 0$).

8.2.1 Properties of the Annihilator

For V a vector space over K with $U, W \subseteq V$ subspaces, we have that:

- $(U + W)^0 = U^0 \cap W^0$
- $U \subseteq W \Rightarrow W^0 \subseteq U^0$,

and for V finite-dimensional,

- $(U \cap W)^0 = W^0 + U^0$
- $\dim(W) + \dim(W^0) = \dim(V)$.

Proof of $\dim(W) + \dim(W^0) = \dim(V)$.
Let $B_W = \{v_1, \dots, v_m\}$ be a basis for W, we extend it with $v_{m+1}, \dots, v_n$ to a basis B for V and take the dual basis $B^* = \{v_1^*, \dots, v_n^*\}$. Each w in W can be decomposed into the vectors in B_W with coefficients $a_1, \dots, a_m$ in K. Thus, for i in $[n]$ with $i > m$ we have that:

$$v_i^*(w) = v_i^*(a_1 v_1 + \cdots a_m v_m) = 0_V.$$

So, v_i^* is in W^0 for each i as defined above. They are linearly independent, and we want to show they span W^0. If f is in W^0 then, f can be decomposed into the vectors in B^* with coefficients $b_1, \dots, b_n$ in K. However, for $i \leq m$, v_i is in W so:

$$0_K = f(v_i) = b_i,$$

which means that:

$$f = b_{m+1} v_{m+1}^* + \cdots + b_n v_n^*.$$

Thus, $(B \setminus B_W)^*$ is a basis for W^0 which means:

$$\dim(W^0) = n - m = \dim(V) - \dim(W)$$

as required. □

8.3 Isomorphism to the Double Dual

For V a finite-dimensional vector space over K, we have $F : V \to V^{**}$ defined for some v in V and f in V^* as follows:

$$F(v)(f) = f(v).$$

We have that F is an isomorphism.

Proof. Omitted. $\square$

8.4 Transposing Linear Maps

For V, W vector spaces with $f : V \to W$ a linear map. We define the transpose as $f^t : W^* \to V^*$ where for g in W^*, v in V:

$$f^t(g) := (g \circ f).$$

So, for some v in V:

$$f^t(g)(v) = (g \circ f)(v) = g(f(v)).$$

8.5 Transposed Linear Maps and Matrices

If we have V, W finite-dimensional vector spaces over K with bases $A = \{v_1, \dots, v_n\}$, $B = \{w_1, \dots, w_m\}$ respectively. We have that for some linear map $f : V \to W$, and $f^t : W^* \to V^*$ the transpose map with respect to f:

$$M_{BA}(f) = \left(M_{A^*B^*}(f^t)\right)^t.$$

9 Rank and Determinants

9.1 Elementary Row Operations

For a field K, take A in $M_{m,n}(K)$. For some c in K, the elementary row operations are:

- Swapping,
- Multiplying a row by $c \neq 0$,
- Adding c multiples of one row to another.

9.1.1 Elementary Matrices

The $n \times n$ elementary matrices are:

- $E_1(i,j)$: obtained by swapping the i^{th} and j^{th} rows of the identity
- $E_2(c,i)$: obtained by scaling the i^{th} row of the identity by c non-zero
- $E_3(c,i,j)$: obtained by adding c times row i to row j where $i \neq j$.

We have that any elementary row operation can be realised as left-multiplication by a corresponding elementary matrix. As a consequence of the definition, we have that elementary matrices are invertible and have elementary inverses.

9.1.2 Echelon Form

A matrix A is in echelon form if each row has the form:

$$(0, \dots, 0, 1, *, \dots, *),$$

where each row has more leading zeroes than the one above and the first row has any amount of leading zeroes. Every matrix can be put in this form via Gaussian elimination.

9.1.3 Decomposition via Elementary Matrices

For an $n \times n$ matrix A, there exists elementary matrices $E_1, \dots, E_k$ such that $E_1 \cdots E_k A = B$ where:

$$B = \begin{cases} \text{the identity} & \text{if } A \text{ is invertible} \\ \text{a matrix with a final row consisting of all zeroes} & \text{otherwise.} \end{cases}$$

9.2 Rank

For $A = (a_{ij})$ a matrix in $M_{m,n}(K)$, we denote its rows by $A_{(1)}, \ldots, A_{(m)}$ and columns by $A^{(1)}, \ldots, A^{(n)}$. We say:

- The row rank of A is the dimension of the subspace of spanned by $A^t_{(1)}, \ldots, A^t_{(m)}$ in K^m

- The column rank of A is the dimension of the subspace of spanned by $A^{(1)}, \ldots, A^{(n)}$ in K^n.

We have these are equal, so can generally refer to the rank of a matrix.

If $E_1, \ldots, E_k$ are elementary matrices, we have that the rank of A is equal to the rank of $E_1 \cdots E_k A$. Similarly, matrices that are similar have the same rank.

9.2.1 Rank of Matrices from Linear Maps

For A an $m \times n$ matrix on K, we can define a map $f : K^n \to K^m$ by $v \mapsto Av$. We have that the rank of A is the dimension of the image of f. Thus, invertible $n \times n$ matrices have rank n.

9.2.2 Partially Diagonalising Matrices

For an $m \times n$ matrix A, there exists some:

- $p \times m$ matrix P,

- $n \times q$ matrix Q,

such that $PAQ = D = (d_{ij})$ where:

$$d_{ij} = \begin{cases} \delta_{ij} & \text{for } i \leq \operatorname{rank}(A) \\ 0 & \text{otherwise,} \end{cases}$$

the matrix with $\operatorname{rank}(A)$ units on the diagonal.

9.3 Permutations

For some n in $\mathbb{Z}_{>0}$, a permutation of $[n]$ is a bijection $\sigma : [n] \to [n]$. We define the set of all permutations on $[n]$ as S_n.

9.4 Properties of S_n

For some n in $\mathbb{Z}_{>0}$, S_n is:

- A group under function composition,
- Of order $n!$,
- Non-abelian for $n > 2$.

9.5 Decomposition of Permutations

All permutations can be written as a product of disjoint cycles. Thus, all permutations can be written as a product of transpositions.

9.6 Parity of Permutations

Even permutations are permutations that can be expressed as the product of an even number of transpositions. Otherwise, a permutation is odd.

9.6.1 The Signature

We define the sgn function for a given permutation σ:

$$\operatorname{sgn}(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{otherwise.} \end{cases}$$

We have that for another permutation τ:

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau).$$

In other words, sgn is a homomorphism from S_n to $\{1, -1\}$.

9.6.2 The Alternating Group

We have that A_n the set of even permutations in S_n is a subgroup as it is the kernel of sgn.

9.7 Determinants

For $A = (a_{ij})$ a $n \times n$ matrix over K, we have the determinant is a scalar defined by:

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

A more pratical but equivalent definition would be:

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

where i is in $[n]$ and $\tilde{A}^{ij}$ is A with the i^{th} row and j^{th} column removed. We may represent the k^{th} column vector of A by $A^{(k)}$ and then write:

$$\det(A) = \det(A^{(1)}, \ldots, A^{(n)}),$$

as a function on the columns of A.

9.7.1 Multi-linearity of the Determinant

For $A = (a_{ij})$ a $n \times n$ matrix over K with $A^{(k)} = c_1 v_1 + c_2 v_2$ with c_1, c_2 in K and v_1, v_2 in K^n for some k in $[n]$, we have that:

$$\begin{aligned} \det(A) &= \det(A^{(1)}, \ldots, A^{(n)}) \\ &= c_1 \cdot \det(A^{(1)}, \ldots, v_1, \ldots, A^{(n)}) \\ &\qquad + c_2 \cdot \det(A^{(1)}, \ldots, v_2, \ldots, A^{(n)}). \end{aligned}$$

From this we can show that any square matrix with a column of all zeroes has zero determinant.

9.7.2 Alternativity of the Determinant

For $A = (a_{ij})$ a $n \times n$ matrix over K with $i \neq j$, we have that:

$$\begin{gathered} \det(A^{(1)}, \ldots, A^{(i)}, \ldots, A^{(j)} \ldots, A^{(n)}) \\ = \\ -\det(A^{(1)}, \ldots, A^{(j)}, \ldots, A^{(i)} \ldots, A^{(n)}). \end{gathered}$$

From this we can show that a matrix with a pair of identical columns must have zero determinant.

9.7.3 Normality of the Determinant

For a square upper (or lower) triangular matrix A, we have that the determinant is the product of the diagonal entries.

From this we can show that the determinant of the identity is 1.

9.7.4 The Determinants of Elementary Matrices

We have the following:

$$\begin{aligned}\det\left[E_1(i,j)\right] &= -1\\ \det\left[E_2(c,i)\right] &= c\\ \det\left[E_3(c,i,j)\right] &= 1.\end{aligned}$$

9.7.5 The Determinant of the Transpose

For a matrix A, we have that: $\det(A) = \det(A^t)$.

9.7.6 The Determinant under Matrix Multiplication

For A, B two $n \times n$ matrices, we have that: $\det(AB) = \det(A) \cdot \det(B)$.

9.7.7 The Determinant and Invertibility

A square matrix is invertible if and only if it has non-zero determinant.

10 Polynomials

For R a ring with unity, a polynomial over R is of the form:

$$p(x) = \sum_{i=0}^{n} a_i x^i,$$

for some sequence $(a_i)_{i\in[n]}$ in R called the coefficients of the polynomial. In this case, x is the interdeterminate.

10.1 The Set of Polynomials

For a ring R, $R[x]$ is the set of all polynomials on R:

1. $R[x]$ is a ring with unity that is commutative if and only if R is commutative.

2. $R[x]$ has a multiplicative identity if and only if R has a multiplicative identity.

Proof. (1) Immediate from the definition of polynomial multiplication. □

Proof. (2) If R has a multiplicative identity 1_R, the multiplicative identity in $R[x]$ is 1_R. If $R[x]$ has multiplicative identity $1_{R[x]}$ then for each r in R, $1_{R[x]}(r) = 1_{R[x]} \cdot r = r$ so $1_{R[x]} = 1_R$. □

10.2 Polynomial Degree

For a polynomial p with coefficients (a_i) the degree is the greatest i such that $a_i \neq 0$ and if no such a_i exists we call this the zero polynomial and the degree is zero. The degree is denoted as $\deg(p)$. The leading coefficient is $a_{\deg(p)}$.

10.3 Degree and Composition in $R[x]$

For a ring with unity R, p, q non-zero elements of $R[x]$, we have that:

- $\deg(p+q) \leq \max(\deg(p), \deg(q))$

- $\deg(pq) \leq \deg(p) + \deg(q)$

- $\deg(pq) = \deg(p) + \deg(q)$ if the leading coefficient of p or q is an invertible element of R (or R is a field).

10.4 Evalutation of Polynomials

For $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ in $R[x]$ and c in R, we have the value of p at c is:

$$p(c) = a_0 + a_1 c + \cdots + a_n c^n.$$

If $p(c) = 0$, then we call c a root of p.

10.5 The Division Algorithm of Polynomials

For a ring with unity R, f, g in $R[x]$ with the leading coefficient of g being a unit (invertible element) in R, we have that there exists q, r in $R[x]$ such that:

$$f(x) = q(x)g(x) + r(x),$$

where r is the zero polynomial or $\deg(r) < \deg(g)$.

Proof. If $\deg(f) < \deg(g)$, q is the zero polynomial and $r = f$. If $\deg(f) \geq \deg(g)$, suppose:

$$\begin{aligned} f(x) &= a_0 + \cdots + a_n x^n \\ g(x) &= b_0 + \cdots + b_m x^m \\ q(x) &= c_0 + \cdots + c_{n-m} x^{n-m}, \end{aligned}$$

where each set of coefficients is in R and $c_0, \ldots, c_{n-m}$ is unknown. We break down f by considering:

$$\begin{aligned} f(x) = a_0 + \cdots + a_n x^n &= (c_0 + \cdots + c_{n-m} x^{n-m})(b_0 + \cdots + b_m x^m) + r(x) \\ &= q(x)g(x) + r(x), \end{aligned}$$

at the greatest power of x. This gives us that $c_{n-m} = a_n b_m^{-1}$. We notice that $f(x) - c_{n-m} x^{n-m} g(x)$ is a polynomial of degree strictly less than the degree of f. If we repeat this process on $f(x) - c_{n-m} x^{n-m} g(x)$ until $\deg(f) < \deg(g)$ we have the result. $\square$

10.5.1 Factorisation by Roots

For p a polynomial in $R[x]$ where $\deg(p) > 0$ and c in R, c is a root of p if and only if we can write $p(x) = (x - c)q(x)$ for some q in $R[x]$.

Proof. Suppose c is a root of p. By (10.5), we have that:

$$p(x) = (x - c)q(x) + r(x),$$

for some q, r in $R[x]$. But, $p(c) = 0_R$ so:

$$\begin{aligned} p(c) &= (c - c)q(x) + r(x) = 0_R \\ &= 0_R \cdot q(x) + r(x) \\ &= r(x), \end{aligned}$$

thus, r is the zero polynomial. So, $p(x) = (x - c)q(x)$. Supposing $p(x) = (x - c)q(x)$ for some q in $R[x]$:

$$\begin{aligned} p(c) &= (c - c)q(x) \\ &= 0_R \cdot q(x) \\ &= 0_R, \end{aligned}$$

so c is a root of p. Thus, we have the result. □

10.6 The Divisibility of Polynomials

For a field K, we have that for p, q in $K[x]$, if q divides p (written as $q|p$) then there exists r in $K[x]$ such that:

$$p(x) = q(x)r(x).$$

10.6.1 Highest Common Factors of Polynomials

For a field K, we have that for p, q in $K[x]$, the highest common factor of p and q is a polynomial h with **maximal** degree such that h divides both p and q.

We also have that there exists a, b in $K[x]$ such that $h = ap + bq$.

10.7 Irreducible Polynomials

An irreducible polynomial over a field K is a non-constant (degree greater than zero) polynomial in $K[x]$ such that it cannot be written as the product of two polynomials (both with smaller degree).

10.7.1 Consequences of Irreducible Divisibility

For a field K, suppose we have f, p, q in $K[x]$ such that f is irreducible. If $f|pq$ then either $f|p$ or $f|q$ or both.

Proof. The highest common factor of f and p is either λ or $\lambda \cdot f$ for some λ in $K \setminus \{0_K\}$ as f is irreducible. If the highest common factor is $\lambda \cdot f$ then we have that $f|p$. Otherwise, f and p are relatively prime so there exists polynomials a, b in $K[x]$ such that:

$$1_{K[x]} = af + bp,$$

Thus, multiplying through by q:

$$q = qaf + qbp.$$

Clearly $f|qaf$ and we can see that $f|qbp$ as $f|pq$. Thus, $f|h$. $\square$

10.7.2 Decomposition into Irreducible Polynomials

For a field K, we have that for every f in $K[x]$ where $\deg(f) \geq 1$ we have that f can be written as the product of irreducible polynomials uniquely up to order and multiplication by constants. If f is monic (leading coefficient equal to one), it is a product of monic irreducible polynomials, unique up to order.

10.8 Definition of the Minimal Polynomial

For a field K and V a finite n-dimensional vector space let $f : V \to V$. The minimal polynomial $m_f(x)$ in $K[x]$ is the polynomial such that:

- $m_f(f) = 0_L$ where $L = \mathcal{L}(V, V)$,
- $\deg(m_f)$ is minimal,
- m_f is monic (leading coefficient equal to one).

We have that this polynomial always exists and is unique.

Proof. Suppose p and m_f are two distinct monic polynomials of the same degree so that $p(f) = 0_L = m_f(f)$. As p and m_f are distinct, $(p - m_f)$ is a non-zero polynomial for which $(p - m_f)(f) = 0_L$. Taking λ to be the leading coefficient of $(p - m_f)$, we have that $\lambda^{-1}(p - m_f)$ annihilates f, is monic but, has degree less than m_f which is impossible. So, the minimal polynomial is distinct.

L is n^2 dimensional as it's isomorphic to $M_n(K)$. Thus, $f^0, f, f^2, \dots, f^{n^2}$ must be linearly dependent on L. Thus, there is $a_0, \dots, a_{n^2}$ not all zero such that:

$$a_0 f^0 + \dots + a_{n^2} f^{n^2} = 0_L.$$

Take k to be maximal such that $a_k \neq 0_K$. Thus:

$$p(x) = a_k^{-1}[a_0 + a_1 x + \dots + a_k x^k],$$

is monic and annihilates f. Thus, the minimal polynomial exists. □

10.8.1 Properties of the Minimal Polynomial

For a field K and V a finite n-dimensional vector space, let f be in $L = \mathcal{L}(V, V)$ and m_f be the corresponding minimal polynomial in $K[x]$. We have that:

1. If p in $K[x]$ satisfies $p(f) = 0$ then $m_f | p$,
2. For λ in K, $m_f(\lambda) = 0$ if and only if λ is an eigenvalue of f.

Proof. (1) By (10.5) we can write $p(x) = m_f(x)q(x) + r(x)$. As p annihilates f:

$$\begin{aligned} p(f) &= m_f(f)q(f) + r(f) = 0_L \\ &= 0_L \cdot q(f) + r(f) \\ &= r(f), \end{aligned}$$

thus, m_f divides p. □

Proof. (2) Taking v an eigenvalue of f with eigenvalue λ, $f(v) = \lambda v$ so for each p in $K[x]$, $p(f)(v) = p(\lambda)(v)$. Taking $p = m_f$:

$$0_V = m_f(f)(v) = m_f(\lambda)(v),$$

so $m_f(\lambda) = 0_K$. Conversely, suppose λ is a root of m_f so $m_f(x) = (x - \lambda)p(x)$ for some monic p in $K[x]$. We have that $\deg(p) < \deg(m_f)$ so $p(f) \neq 0$ as otherwise this would contradict the minimality of m_f. We take v in V such that $v' = p(f)(v)$ is non-zero. However:

$$\begin{aligned} 0_K = m_f(f)(v) &= [(x - \lambda)p(x)](f)(v) \\ &= (f - \lambda(\mathrm{id}))p(f)(v) \\ &= (f - \lambda(\mathrm{id}))(v'), \end{aligned}$$

so v' is an eigenvector of f with eigenvalue λ. □

10.9 Characteristic Polynomials

The characteristic polynomial of an operator $f : V \to V$ is the polynomial:

$$p_f(x) = \det(A - xI),$$

where A is the matrix of f relative to some basis. This doesn't change based on the choice of basis as similar matrices have the same determinant. Additionally, this is divisible by m_f by the Cayley-Hamilton theorem.

10.9.1 The Cayley-Hamilton Theorem

For V a finite n-dimensional vector space over a field K where p_f the characteristic polynomial of an operator f in $L = \mathcal{L}(V, V)$, we have that:

$$p_f(f) = 0_V.$$

We also have that:

$$p_f(M_{BB}(f)) = 0_V,$$

for some basis B of V.

Proof over an algebraically closed field. If $\dim(V) = 1$, f simply scales vectors in V by some scalar c in K. Thus, $p_f(x) = c - x$ so $p_f(f) = 0_L$. Suppose $\dim(V) > 1$ and the theorem holds for all spaces of dimension $\dim(V) - 1$. As K is algebraically closed and $\deg(p_f) > 0$, we have that f has an eigenvalue λ with corresponding eigenvector v_1. We expand $\{v_1\}$ to $B = \{v_1, \ldots, v_n\}$ a basis for V. We have that:

$$p_f(x) = (\lambda - x)p(x),$$

for some p in $K[x]$. Taking $V_1 = \operatorname{span}(\{v_1\})$ the f-invariant subspace of V, we consider the induced map $\bar{f} : V/V_1 \to V/V_1$. We have that $p_{\bar{f}} = p$ by (7.5) and by assumption $p_{\bar{f}}(\bar{f}) = 0_L$ meaning for any v in V:

$$p_{\bar{f}}(\bar{f})(v + V_1) = 0_{V/V_1}.$$

As $p_{\bar{f}}(\bar{f})$ maps elements of V/V_1 to $0_{V/V_1}$, $p_{\bar{f}}(f)$ must map elements of V to V_1. Thus, as $p_f(x) = (\lambda - x)p_{\bar{f}}(x)$:

$$\begin{aligned} p_f(f)(v) &= (\lambda - f)p_{\bar{f}}(f)(v) \\ &= (\lambda - f)(cv_1) && \text{(for some } c \text{ in } K\text{)} \\ &= 0_V. \end{aligned}$$

Proving the result by induction. □

11 Jordan

11.1 Jordan Blocks

For a field K, h in $\mathbb{Z}_{>0}$, a Jordan block of size $h \times h$ on λ in K is the matrix of the form:

$$J_h(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix},$$

or alternatively:

$$J_h(\lambda) = (a_{ij}),\ a_{ij} = \begin{cases} \lambda & i = j \\ 1 & j = i + 1 \\ 0 & \text{otherwise.} \end{cases}$$

11.1.1 Jordan Matrices

For a field K, a Jordan matrix consisting of Jordan blocks of sizes $\{h_1, \ldots, h_n\}$ in $\mathbb{Z}_{>0}$ and values $\{\lambda_1, \ldots, \lambda_n\}$ in K has the form:

$$J = \begin{pmatrix} J_{h_1}(\lambda_1) & 0 & \cdots & \cdots & 0 \\ 0 & J_{h_2}(\lambda_2) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & J_{h_n}(\lambda_n) \end{pmatrix}$$

11.1.2 Jordan Normal Form

A Jordan normal form of a matrix A is a Jordan matrix that is similar to A.

11.2 Jordan Bases

For an algebraically closed field K, a finite dimensional vector space V over K, with f in $\mathcal{L}(V, V)$, we have that a basis B of V is a Jordan basis for f if $M_{BB}(f)$ is a Jordan matrix.

11.2.1 Existence of Jordan Bases

Let K be an algebraically closed field, V be a finite dimensional vector space over K, f be in $L = \mathcal{L}(V, V)$. There exists a Jordan basis for f and the Jordan normal form of f is unique up to permutations of the Jordan blocks.

Proof. We first suppose f is nilpotent, so for some t in $\mathbb{Z}_{\geq 0}$, $f^t = 0_L$. We know that $V = V(0)$ as f has a single eigenvalue, zero. Take $V_0 = E(0)$ and V_i for i in $[t]$ to be defined as follows:

$$V_i = \mathrm{Im}(f^i) \cap V_0,$$

so V_i is the eigenvectors of f^i (specifically the ones in $E(0)$). Take v in V_i, there exists some v' such that $f^i(v') = v$ as v is in $\mathrm{Im}(f^i)$. So:

$$\begin{aligned} v &= f^i(v') \\ &= f^{i-1}(f(v')), \end{aligned}$$

meaning that v is in $\mathrm{Im}(f^{i-1})$. But, we also have that v must be in V_0 by the definition of V_i so v is in V_{i-1}. Hence, $V_i \subseteq V_{i-1}$ and we can say that:

$$\{0_V\} \subseteq V_t \subseteq \cdots \subseteq V_1 \subseteq V_0.$$

Let $h : E(0) \to \mathbb{Z}_{\geq 0}$ be the height function. Consider $B_{t-1} = \{e_1, \ldots, e_{b_1}\}$ a basis for V_{t-1}. For each b in $[b_1]$, let $h(e_b) = t-1$ and choose $v_b^{h(e_b)} = v_b^{t-1}$ in V such that:

$$f^{h(e_b)}(v_b^{h(e_b)}) = e_b.$$

We can choose such a value in V as $V_{t-1} = V_{h(e_b)}$ is in the image of $f^{h(e_b)}$. Now, for k in $[h(e_b)]$, we let:

$$v_b^{k-1} = f(v_b^k).$$

so that:

$$\begin{aligned} v_b^{t-2} &= f(v_b^{t-1}) \\ v_b^{t-3} &= f^2(v_b^{t-1}) \\ &\cdots \\ v_b^1 &= f^{t-2}(v_b^{t-1}) \\ e_b = v_b^0 &= f^{t-1}(v_b^{t-1}). \end{aligned}$$

We extend B_{t-1} to B_{t-2} (a basis for V_{t-2}) with new basis elements:

$$B_{t-2} = \{e_1, \ldots, e_{b_1}, e_{b_1+1}, \ldots, e_{b_2}\}.$$

Similarly to the above, take b in $\{b_1+1,\ldots,b_2\}$ and let $h(e_b)=t-2$ and choose v_b^{t-2} such that $f^{t-2}(v_b^{t-2})=e_b$. We then define for k in $[h(e_b)]$:

$$v_b^{k-1}=f(v_b^k).$$

so that:

$$\begin{aligned}
v_b^{t-3}&=f(v_b^{t-2})\\
v_b^{t-4}&=f^2(v_b^{t-2})\\
&\cdots\\
v_b^1&=f^{t-3}(v_b^{t-2})\\
e_b=v_b^0&=f^{t-2}(v_b^{t-2}).
\end{aligned}$$

We continue this for all $V_{t-3},\ldots,V_0$. We take $B_0=V_0=\{e_1,\ldots,e_n\}$ to be our basis extended to V_0 noting that $n=b_1+\cdots+b_t=\dim(V_0)$. We have:

$$B=\{v_i^j:\underbrace{1\le i\le n}_{\text{stack number}},\overbrace{0\le j\le h(e_i)}^{\text{stack height}}\},$$

our Jordan basis.

We want to show linear independence of B, so we take $\{a_i^j\in K:1\le i\le n,0\le j\le h(e_i)\}$ such that:

$$\sum_{i=1}^{n}\sum_{j=0}^{h(e_i)}a_i^jv_i^j=0_V. \tag{5}$$

We first notice that for each v_i^j, $f^j(v_i^j)=e_i$ and so $f^{j+1}(v_i^j)=0_V$. Thus, applying f^{t-1} to both sides of (5) we get a linear combination of basis vectors (specifically in B_{t-1}) meaning each element of $\{a_i^{t-1}:1\le i\le n\}$ must be zero. Thus, taking $0\le j_0<j\le t-1$, where all elements of $\{a_i^j:1\le i\le n\}$ are zero. By applying f^{j_0} to both sides of (5) we see we have a linear combination of vectors in V_{j_0} with coefficients $\{a_i^{j_0}:1\le i\le n\}$ necessarily zero. So, by induction, all the coefficients are zero and B is linearly independent.

We want to show that B spans V, remembering that $V=V(0)$, we take v in V (necessarily a root vector), and has some height $h_v\le t$. Suppose $h_v=1$, B contains a basis for V_0 so v is in the span of B. Suppose $h_v=k+1$ and for each vector v' of height less than or equal to k we have that v' is in the span of B. We have that

$f^k(v)$ is in V_k as it's in $\mathrm{Im}(f^k)$ and $f(f^k(v)) = 0_V$ so is in V_0. Thus, considering $V_k \cap B = \{e_1, \dots, e_j\}$ a basis for V_k, we have $a_1, \dots, a_j$ in K such that:

$$f^k(v) = \sum_{i=1}^{j} a_i e_i.$$

Considering that each $e_i = f^k(v_i^k)$:

$$\begin{aligned} f^k(v) &= \sum_{i=1}^{j} a_i e_i \\ &= \sum_{i=1}^{j} a_i f^k(v_i^k) \\ &= f^k\left(\sum_{i=1}^{j} a_i v_i^k\right). \end{aligned}$$

So, we can deduce that v is in the span of B. Hence, B is a Jordan basis.

To show uniqueness, the number of $m \times m$ Jordan blocks is:

$$\dim(\mathrm{Im}(f^{m-1}) \cap \mathrm{Ker}(f))) - \dim(\mathrm{Im}(f^m) \cap \mathrm{Ker}(f)),$$

which is uniquely defined.

Finally, we take f to be any linear map. We consider $\lambda_1, \dots, \lambda_k$ to be the distinct eigenvalues of f. By the Primary Decomposition Theorem, we have that:

$$V = \bigoplus_{i=1}^{k} V(\lambda_i).$$

For each i in $[k]$, by the properties of the root space, $V(\lambda_i)$ is f-invariant. So, we consider $f_i : V(\lambda_i) \to V(\lambda_i); v \mapsto f(v)$. We have that $(f_i - \lambda_i(\mathrm{id}))$ has one eigenvalue, zero. By using the Cayley-Hamilton theorem we can see that:

$$(f_i - \lambda_i(\mathrm{id}))^{\dim(V(\lambda_i))} = 0_{\mathcal{L}(V,V)},$$

so $(f_i - \lambda_i(\mathrm{id}))$ is nilpotent. By the above, $(f_i - \lambda_i(\mathrm{id}))$ has a Jordan basis B_i. Putting this together, $B = \bigcup_{i=1}^{k} B_i$ is a Jordan basis for f. $\square$

11.2.2 Relation to Eigenvalue Multiplicity

For a given eigenvalue:

- The geometric multiplicity is the number of Jordan blocks with the eigenvalue,
- The algebraic multiplicity is the sum of the sizes of all the Jordan blocks corresponding to the eigenvalue,
- The algebraic multiplicity in m_f is the maximum size of a Jordan block corresponding to the eigenvalue.

11.2.3 Computing Jordan Bases

The process is as follows:

- Compute the characteristic polynomial and factorise it,
- For each eigenvalue, find the eigenspaces,
- If the geometric multiplicity of an eigenvalue is less than the algebraic multiplicity then the difference is the number of root vectors that need to be found.

12 Bilinear and Quadratic Forms

12.1 Bilinear Forms

For V a vector space over a field K, a bilinear form on V is a map $\langle\,,\rangle : V \times V \to K$ such that:

$$\begin{aligned}\langle au + bv, w\rangle &= a \cdot \langle u, w\rangle + b \cdot \langle v, w\rangle \\ \langle u, av + bw\rangle &= a \cdot \langle u, v\rangle + b \cdot \langle u, w\rangle,\end{aligned}$$

for all a, b in K, u, v, w in V. Additionally, $\langle\,,\rangle$ is symmetric if $\langle u, v\rangle = \langle v, u\rangle$.

12.2 Quadratic Forms

For V a vector space over a field K with $\langle\,,\rangle$ a symmetric bilinear form on V. The quadratic form $Q : V \to K$ associated to $\langle\,,\rangle$ is $Q(v) := \langle v, v\rangle$.

12.2.1 Determining Bilinear Forms from Quadratic Forms

We have that if $\mathrm{char}(K) \neq 2$, then $\langle\,,\rangle$ is uniquely defined by Q as:

$$\begin{aligned}2^{-1}[Q(v+w) - Q(v) - Q(w)] &= 2^{-1}[\langle v+w, v+w\rangle - Q(v) - Q(w)] \\ &= 2^{-1}[\langle v, v+w\rangle + \langle w, v+w\rangle - Q(v) - Q(w)] \\ &= 2^{-1}[2 \cdot \langle v, w\rangle] \qquad (\langle\,,\rangle \text{ is symmetric}) \\ &= \langle v, w\rangle.\end{aligned}$$

12.3 Orthogonality

Let $\langle\,,\rangle$ be a bilinear form on V with v, u in V. We say that v in V is orthogonal to u if $\langle u, v\rangle = 0$. Note, be very careful as a bilinear form is not necessarily symmetric.

12.3.1 Orthogonal Spaces

For $W \subseteq V$, we have that $W^\perp$ is defined as:

$$W^\perp = \{v \in V : w \in W, \langle w, v\rangle = 0\},$$

the set of vectors such that for all v in $W^\perp$, v is orthogonal to all of W. This is a subspace of V.

12.3.2 The Kernel for Bilinear maps

The kernel of $\langle\,,\rangle$ is $V^\perp$. If the kernel is $\{0_V\}$, then the form is called non-degenerate and is called degenerate otherwise.

12.3.3 Dimension and Orthogonal Spaces

We have that if V is finite dimensional and $\langle\,,\rangle$ is non-degenerate then:

$$\dim(W^\perp) + \dim(W) = \dim(V).$$

Proof. We note that f, the map corresponding to $\langle\,,\rangle$ is an isomorphism. Thus:

$$\dim(W^\perp) = \dim(f(W^\perp)) = \dim(W^0) = \dim(V) - \dim(W). \qquad \square$$

12.4 Linear Maps from Bilinear Forms

We can form a linear map $f : V \to V^*$ from a bilinear form $\langle\,,\rangle$ as follows:

$$f(v)(u) = \langle u, v\rangle.$$

12.4.1 Isomorphismic Bilinear Maps

If V is finite dimensional, we have that f is an isomorphism if and only if $\langle\,,\rangle$ is non-degenerate. That is, $\mathrm{Ker}(f) = \mathrm{Ker}(\langle\,,\rangle) = V^\perp = \{0_V\}$.

12.5 Matrices from Bilinear Forms

For V a finite n-dimensional vector space over K with $S = \{v_1, \ldots, v_n\}$ an ordered basis for V. Let $B = \langle\,,\rangle$ be a bilinear form. The matrix corresponding to B with respect to S is $M_{SS}(B) = (b_{ij})$ where:

$$b_{ij} = \langle v_i, v_j\rangle.$$

Similarly, taking $S^* = \{v_1^*, \ldots, v_n^*\}$ to be a dual basis for S^*, we have a matrix $M_{SS^*}(f) = M_{SS}(B)$ for the linear map f corresponding to B.

12.5.1 Determining Bilinear Forms from Matrices

Take u, v in V decomposed into vectors in S with coefficients $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ respectively. Thus:

$$\langle u, v\rangle = (x_1, \ldots, x_n) \cdot M_{SS}(B) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

12.5.2 Properties of Bilinear Matrices

We have that:

- $M_{SS}(B)$ is symmetric if and only if B is symmetric,
- B is non-degenerate if and only if $M_{SS}(B)$ is invertible.

12.5.3 Similarity of Matrices of Bilinear Forms

For V a finite n-dimensional vector space over the field K with $\mathrm{char}(K) \neq 2$, let $S = \{v_1, \ldots, v_n\}$, $S' = \{v'_1, \ldots, v'_n\}$ be ordered bases for V. Let $B = \langle\, , \rangle$ be a symmetric bilinear form. Let $C = C_{SS'}$ be the transition matrix. We have that:

$$M_{S'S'}(B) = C^t M_{SS}(B) C.$$

12.5.4 Diagonal Matrices of Bilinear Forms

For V a finite n-dimensional vector space over the field K with $\mathrm{char}(K) \neq 2$, let $B = \langle\, , \rangle$ be a symmetric bilinear form. There exists a basis $S = \{v_1, \ldots, v_n\}$ for V consisting of pairwise orthogonal vectors and thus, the matrix $M_{SS}(B)$ is diagonal.

12.6 Inner Products

For a vector space V over K with symmetric bilinear form $B : V \times V \to K$, we have that B is an inner product if for all v in V:

$$B(v, v) \geq 0,$$

and $B(v, v) = 0$ if and only if $v = 0_V$.