McGill University

MATH 251

Class Notes

Alexandre St-Aubin

Taught by
Mikaël Pichot

March 23, 2023

Contents

1 Fields and Rings

Definition 1.1. (Ring) A set R is a ring if:

1. $a + b = b + a \ \forall a, b \in R$
2. $(a + b) + c = a + (b + c) \ \forall a, b, c \in R$
3. $\exists 0 \in R \ s.t. \ a + 0 = a \ \forall a \in R$
4. $\forall a \in R, \ \exists (-a) \ s.t. \ a + (-a) = 0$
5. $(ab)c = a(bc) \ \forall a, b, c \in R$
6. $a(b + c) = ab + ac \ \forall a, b, c \in R$

Definition 1.2. (Ring with unity)

$$1 \in R$$

Definition 1.3. (Commutative Ring)

$$ab = ba \ \forall a, b \in R$$

Definition 1.4. (Integral domain) If the following hold:

1. $1 \in R$
2. $ab = ba \ \forall a, b \in R$
3. $\forall a, b \in R, \ ab = 0 \implies a = 0 \ or \ b = 0$

i.e. commutative ring with unity and no zero divisors.

Definition 1.5. (Division Ring) Ring with unity such that $\forall R \ni a \neq 0$, a is a unit.

Definition 1.6. (**Field**) A field $\mathbb{F}$ is a commutative ring with unity such that $\forall x \neq 0 \in \mathbb{F}$, x is a unit

Theorem 1.1. *For any prime p,* $\mathbb{F}_p := (\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ *is a field, i.e.* $\mathbb{Z}_p$ *is a field.*

Definition 1.7. (**Subfield**) Let $(S, *, \cdot)$ be be an algebraic structure with 2 operations. Let T be a subset of S such that $(T, *, \cdot)$ is a field. Then, $(T, *, \cdot)$ is a **subfield** of S.

2 Vector spaces

Definition 2.1. (Vector Space) A vector space V over a field $\mathbb{F}$ (or $\mathbb{F}$-vector space) is a set on which two operations (addition and scalar multiplication) are defined so that for each pair of elements x, y in V, there is a unique element $x + y$ in V, and for each element a in $\mathbb{F}$ and each element x in V there is a unique element ax in V, such that the following conditions hold.

- Additive axioms:
  1. For all x, y in V, $x + y = y + x$ (commutativity of addition)
  2. For all x, y, z in V, $x + (y + z) = (x + y) + z$ (associativity of addition)

3. There exists an element 0 in V such that $x + 0 = 0 + x = x, \forall x \in V$
4. For each $x \in V, \exists y \in V$ such that $x + y = 0$

- Multiplicative axioms:

1. For each $x \in V$, $1x = x$
2. For each $x \in V$, $0x = 0$
3. For each pair of elements a, b in $\mathbb{F}$, and $x \in V$, $(ab)x = a(bx)$

- Distributive axioms:

1. For each element $a \in \mathbb{F}$ and pair of elements $x, y \in V$, $a(x + y) = ax + ay$
2. For each pair of elements $a, b \in \mathbb{F}$ and $x \in V$, $(a + b)x = ax + bx$

Remark. The elements of the field $\mathbb{F}$ are called scalars and the elements of the vector space $\mathbb{V}$ are called vectors.

Definition 2.2. (n-tuple) An object of the form $(a_1, a_2, ..., a_n)$ with entries in $\mathbb{F}$ is called an n-tuple. Two n-tuples are equal if they are equal component-wise.

Definition 2.3. ($\mathbb{F}^n$) The set of all n-tuples with entries from a field $\mathbb{F}$ is denoted $\mathbb{F}^n$. This set is a vector space over $\mathbb{F}$ with operations addition and scalar multiplication, and we have that

$$\mathbb{F}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_i \in \mathbb{F} \right\}$$

Theorem 2.1. *(Cancellation Law for Vector Addition) If x, y, z are vectors in a vector space V such that $x + y = x + z$, then $y = z$.*

Theorem 2.2. *In any vector space V, the following are true:*

1. $0x = 0, \forall x \in V$
2. $(-a)x = -(ax) = a(-x), \forall a \in \mathbb{F}, x \in V$
3. $a0 = 0, \forall a \in \mathbb{F}$

3 Subspaces

Definition 3.1. (Subspace) A subset W of a vector space V over a field $\mathbb{F}$ is called a subspace of V if W is a vector space over $\mathbb{F}$ with the operations of addition and scalar multiplication defined on V.

Remark. In any vector space, V and 0 are subspaces. The latter is called the zero subspace of V.

Theorem 3.1. *A subset W of a vector space V is a **subspace** of V if and only if the following properties hold.*

1. *$x + z \in W$ whenever $x, y \in W$. (W is closed under addition)*

2. *$cx \in W$ whenever $c \in \mathbb{F}$ and $x \in W$. (W is closed under scalar multiplication)*

3. *W has a zero vector.*

Theorem 3.2. *Let $U \subseteq V$, with V a vector space over $\mathbb{F}$, TFAE:*

1. *U is a vector space over $\mathbb{F}$*

2. *$\forall n, v \in U, \lambda \in \mathbb{F}$, we have $n + \lambda v \in U$*

3. *$\exists S \subseteq U$ such that $U = span(S)$*

Definition 3.2. (Transpose of a matrix) The transpose A' of an $m \times n$ matrix A is the $n \times m$ matrix obtained from A by interchanging the rows with the columns.

Definition 3.3. (Symmetric matrix) A symmetric matrix is a matrix A' such that $A' = A$.

Definition 3.4. (Trace of a matrix) The trace of an $n \times n$ matrix M, denoted $tr(M)$, is the sum of the diagonal entries of M.

Theorem 3.3. *Any intersection of subspaces of a vector space V is a subspace of V.*

4 Linear Combinations and Systems of linear equations

Definition 4.1. (Linear Combination) Let V be a vector space and S a nonempty subset of V. A vector $v \in V$ is called a **linear combination** of vectors of S if there exist a finite number of vectors $u_1, u_2, ..., u_n \in S$ and scalars $a_1, a_2, ..., a_n \in \mathbb{F}$ such that $v = a_1u_1 + a_2u_2 + ... + a_nu_n$. In this case we also say that v is a linear combination of $u_1, u_2, ..., u_n$ and call $a_1, a_2, ..., a_n$ the **coefficients** of the linear combination.

Definition 4.2. (Span) Let S be a nonempty subset of a vector space V. The **span** of S, denoted $span(S)$, is the set consisting of all linear combinations of the vectors in S. For convenience, we define $span(0) = (\emptyset)$.

Example 4.1. In $\mathbb{R}^3$, the span of the set $\{(1, 0, 0), (0, 1, 0)\}$ consists of all vectors in $\mathbb{R}^3$ that have the form $a(1, 0, 0) + b(0, 1, 0) = (a, b, 0)$ for some scalars a and b. Thus the span of $\{(1, 0, 0), (0, 1, 0)\}$ contains all the points in the xy-plane. In this case, the span of the set is a subspace of $\mathbb{R}^3$. This fact is true in general.

Theorem 4.1. *The span of any subset S of a vector space V is a subspace of V. Moreover, any subspace of V that contains S must also contain the span of S.*

Definition 4.3. A subset S of a vector space V **generates** (or **spans**) V if $span(S) = V$. In this case, we also say that the vectors of S generate (or span) V.

5 Linear Dependence and Linear Independence

Remark. Suppose that V is a vector space over an infinite field and that W is a subspace of V. Unless W is the zero subspace, W is an infinite set. It is desirable to find a "small" finite subset S that generates W because we can then describe each vector in W as a linear combination of the finite number of vectors in S. Indeed, the smaller that S is, the fewer computations that are required to represent vectors in W.

Definition 5.1. (Linear Dependence) A subset S of a vector space V is called **linearly dependent** if there exist a finite number of distinct vectors $u_1, u_2, ..., u_n \in S$ and scalars $a_1, a_2, ..., a_n \in \mathbb{F}$ not all zero, such that

$$a_1u_1 + a_2u_2 + ... + a_nu_n = 0$$

Definition 5.2. (Linear Independence) A subset S of a vector space that is not linearly dependent is called **linearly independent**. As before, we also say that the vectors of S are linearly independent.

Theorem 5.1. *A set is linearly independent $\iff$ the only representations of 0 as linear combinations of its vectors are trivial representations.*

Theorem 5.2. *Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. If S_1 is linearly dependent, then S_2 is linearly dependent.*

Corollary 5.2.1. Let V be a vector space, and let $S_1 \subseteq S_2 \subseteq V$. If S_2 is linearly independent, then S_1 is linearly independent.

Theorem 5.3. *Let S be a linearly independent subset of a vector space V, and let v be a vector in V that is not in S. Then $S \cup \{v\}$ is linearly dependent $\iff$ $v \in span(S)$.*

6 Bases and Dimensions

Definition 6.1. (Basis) A **basis** β for a vector space V is a linearly independent subset of V that generates V. If β is a basis for V, we also say that the vectors of β form a basis for V.

Example 6.1. In $\mathbb{F}^n$, let $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), ..., e_n = (0, 0, \ldots, 1)$. Then, $\{e_1, e_2, \ldots, e_n\}$ is easily seen to be the basis for $\mathbb{F}^n$ and is said to be the standard basis for $\mathbb{F}^n$.

Theorem 6.1. *Let V be a vector space over $\mathbb{F}$ and $\beta = u_1, u_2, ..., u_n$ be a subset of V. Then β is a basis for V $\iff$ each $v \in V$ can be uniquely expressed as a linear combination of vectors of β, that is, can be expressed in the form*

$$v = a_1u_1 + a_2u_2 + \ldots + a_nu_n$$

for unique scalars $a_1, a_2, ..., a_n \in \mathbb{F}$

Remark. Theorem 5.1 shows that if the vectors $u_1, u_2, ..., u_n$ form a basis for a vector space V, then every vector in V can be uniquely expressed in the form

$$v = a_1u_1 + a_2u_2 + \ldots + a_nu_n$$

for unique scalars $a_1, a_2, ..., a_n \in \mathbb{F}$. Thus v determines a unique n-tuple of scalars $a_1, a_2, ..., a_n$ and, conversely, each n-tuple of scalars determines a unique vector $v \in V$ by using the entries of the n-tuple as the coefficients of a linear combination of $u_1, u_2, ..., u_n$.

Theorem 6.2. *If a vector space V is generated by a finite set S, then some subset of S is a basis for V. Hence V has a finite basis.*

Theorem 6.3. *(Replacement Theorem) Let V be a vector space that is generated by a set G containing exactly n vectors, and let L be a linearly independent subset of V containing exactly m vectors. Then $m < n$ and there exists a subset H of G containing exactly $n - m$ vectors such that $L \cup H$ generates V.*

Corollary 6.3.1. Let V be a vector space having a finite basis. Then every basis for V contains the same number of vectors.

Definition 6.2. (Dimension) A vector space is called **finite-dimensional** if it has a basis consisting of a finite number of vectors. The unique number of vectors in each basis for V is called the **dimension** of V and is denoted by $dim(V)$. A vector space that is not finite-dimensional is called **infinite-dimensional**.

Example 6.2. The vector space $\{0\}$ has dimension zero.

Example 6.3. The vector space $\mathbb{F}^n$ has dimension n.

Example 6.4. The vector space $M_{m\times n}(\mathbb{F})$ has dimension mn.

Example 6.5. The vector space $P_n(\mathbb{F})$ has dimension $n + 1$.

Corollary 6.3.2. Let V be a vector space with dimension n, then

1. Any finite generating set for V contains at least n vectors, and a generating set for V that contains exactly n vectors is a basis for V.
2. Any linearly independent subset of V that contains exactly n vectors is a basis for V.
3. Every linearly independent subset of V can be extended to a basis for V.

Definition 6.3. (Line) A **line** is a set of the form $span(v)$, $v \in V$ such that $v \neq 0$, i.e. $dim_K(span(v)) = 1$. Note that $0 \in span(v)$, simply by choosing $\lambda = 0$. In $\mathbb{F}_2^n$, there are $2^n - 1$ lines.

Definition 6.4. (Plane) A **plane** is a set of the form $span(u, v)$, with $u, v \in V$ such that $u, v \neq 0$, u, v L.I., i.e. $dim_K(span(u, v)) = 2$. Note that $0 \in span(u, v)$.

Theorem 6.4. *Let V be a vector space with dimension n.*

1. *Any finite generating set for V contains at least n vectors, and a gener- ating set for V that contains exactly n vectors is a basis for V.*
2. *Any linearly independent subset ofV that contains exactly n vectors is a basis for V.*
3. *Every linearly independent subset of V can be extended to a basis for V.*

Theorem 6.5. ***(Major Theorem on Basis)*** *Let V a finite dimension vector space (V f.d. v.s.) over $\mathbb{F}$, where $v_1, \ldots, v_n$ are basis vectors in V. Then, TFAE:*

1. *$v_1, \ldots, v_n$ is a minimal spanning set, i.e. $\nexists$ a spanning set of $< n$ elements.*
2. *$v_1, \ldots, v_n$ is a maximal L.I. (linearly independent) set. (see definition 6.3)*
3. *$v_1, \ldots, v_n$ = spanning + L.I.*
4. *Any $v \in V$ can be written as a linear combination $v = \sum \lambda_i v_i$ where the λ_i's (called the **coordinates** of v in basis $v_1, \ldots, v_n$) are unique and $1 < i < n$.*

Corollary 6.5.1. Any two bases are the same size, i.e. if both $u_1, u_2, ..., u_n$ and $u_1, u_2, ..., u_m$ satisfy (3), then $m = n$.

Lemma 6.6. Suppose $V = span(v_1, \ldots, v_n)$, Then, any $u_1, u_2, ..., u_{n+1}$ are L.D.

Theorem 6.7. *(Basis Extension Theorem) Every linearly independent list of vectors in a finite-dimensional vector space V can be extended to a basis of V. That is, suppose you have a v.s. V,* $(u_1, ..., u_n)$ *L.I. vectors and* $(v_1, ..., v_m)$ *span(V). Then,* $\exists(u_{n+1}, ..., u_k) \in \{v_1, ..., v_m\}$ *such that* $(u_1, ..., u_k)$ *is a basis of V.*

Proof. Suppose V is finite-dimensional and $(u_1, ..., u_n)$ is linearly independent. Since V is finite-dimensional, there exists a list $\{v_1, ..., v_m\}$ of vectors that spans V . We wish to adjoin some of the v_m to $(u_1, ..., u_n)$ to create a basis of V .
Step 1. If $v_1 \in span(u_1, ..., u_n)$, let $S = (u_1, ..., u_n)$, otherwise, $S = (u_1, ..., u_n, v_1)$.
Step k. If $v_k \in span(S)$, leave S unchanged, otherwise, adjoin v_k to S.
After each step the list S is still linearly independent since we only adjoined v_k if v_k was not in the span of the previous vectors. After n steps $v_k \in span(S) \forall k = 1, 2, ..., m$. Since $(v_1, ..., v_m)$ was a spanning list, S spans V , so that S is indeed a basis of V . □

Corollary 6.7.1. $U \subseteq V \implies dim(U) \leq dim(V)$

Proof. Simply extend a basis of U to a basis of V (basis ext. thm). □

Proposition 6.1. (Finding a basis) Let $v_1, ..., v_m \in R^n$. We want to find a basis of the span of $v_1, ..., v_m$.

1. Write the vectors as rows on a matrix.
2. Use Gaussian algorithm to reduce to REF.
3. Then, $span(v_1, ..., v_m) = span$(rows of the REF), i.e. discard the rows of zeros as they are L.D.

7 Maximal Linearly Independent Subsets

Definition 7.1. (Maximal) Let τ be a family of sets. A member M of τ is called **maximal** (with respect to set inclusion) if M is contained in no member of τ other than M itself.

Definition 7.2. (Chain) A collection of sets C is called a **chain** (or nest or tower) if for each pair of sets A and B in C, either $A \subseteq B$ or $B \subseteq A$.

Lemma 7.1. (**The Maximal Principle**) Let T be a family of sets. If, for each chain $C \subseteq T$, there exists a member of T that contains each member of C, then T contains a maximal member.

Definition 7.3. (Maximal Linearly Independent Subset) Let S be a subset of a vector space V. A **maximal linearly independent subset** of S is a subset B of S satisfying both of the following conditions.

1. B is linearly independent.
2. The only linearly independent subset of S that contains B is B itself.

Theorem 7.2. *Let V be a vector space and S a subset that generates V. If* β *is a maximal linearly independent subset of S, then* β *is a basis for V.*

Theorem 7.3. *Let S be a linearly independent subset of a vector space V. There exists a maximal linearly independent subset of V that contains S.*

Corollary 7.3.1. Every vector space has a basis.

8 Sums

Let W a $v.s.$, with $U, V \subseteq W$ subspaces.

Definition 8.1. $U + V := \{u + v : u \in U, v \in V\}$

Definition 8.2. (Direct Sum) We say that $W = U \oplus V$ if

1. $W = U + V$
2. $U \cap V = \{0\}$

Proposition 8.1. $W = U \oplus V \iff$ any $w \in W$ can be written as $w = u + v$ in a unique way ($u \in U, v \in V$).

Remark. Let $v_1, ..., v_n$ a basis of V, then $V = \{Kv_1 \oplus ... \oplus Kv_n\}$, where $Kv_1 = span(v_1)$ = a line generated by v_1.

Theorem 8.1. *(Direct Sums and Dimension) Suppose U and V are subspaces of W, with $U \oplus V = W$, then $dim(W) = dim(U) + dim(V)$.*

Theorem 8.2. ***(Gluing Basis)*** *Let $E_i = (e_{1,i}, ..., e_{k,i})$ be a basis of V_i. Then $E_1 \sqcup ... \sqcup E_n$ is a basis of $V_1 \oplus ... \oplus V_n$.*

Definition 8.3. (Complement) Let $U \subseteq W$, we say $V \subseteq W$ is a complement of U if $U \oplus V = W$.

Proposition 8.2. The complement always exists.

Proof. Choose a basis of U $(u_1, ..., u_n)$, extend it to a basis of W $(u_1, ..., u_n, u_{n+1}, ..., u_m)$. Then, let $V = span(u_{n+1}, ..., u_m)$, so we have that V =complement of U. □

Proposition 8.3. If U_1 and U_2 are subspaces of a finite dimensional vector space then:

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

Proof. Let $u_1, ..., u_m$ be a basis of $U_1 \cap U_2$; thus $\dim(U_1 \cap U_2) = m$. Because $u_1, ..., u_m$ is a basis in $U_1 \cap U_2$, it is linearly independent in U_1. Hence this list can be extended to a basis $u_1, ..., u_m, v_1, ..., v_j$ of U_1. Thus, $\dim U_1 = m + j$. Also extend $u_1, .., u_m$ to a basis $u_1, ..., u_m, w_1, ..., w_k$ of U_2. $\dim U_2 = m + k$. We will show that $u_1, ..., u_m, v_1, ..., v_j, w_1, ..., w_k$ is a basis of $U_1 + U_2$. This will complete the proof because then we will have

$$\dim(U_1 + U_2) = m + j + k = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

We just need to show that the list $u_1, ..., u_m, v_1, ..., v_j, w_1, ..., w_k$ is linearly independent. To prove this, suppose:

$$a_1u_1 + ... + au_m + b_1v_1 + ... + b_jv_j + c_1w_1 + ... + c_kw_k = 0$$

where all a, b, c's are scalars. We need to show that all the a, b and c's are 0.
The equation can be rewritten as

$$c_1w_1 + ... + c_kw_k = -a_1u_1 - ... - a_mu_m - b_1v_1 - ... - b_jv_j$$

Which shows that $c_1w_1 + ... + c_kw_k \in U_1$. But actually all w's are in U_2. So the LHS must be an element of $U_1 \cap U_2$.

$c_1w_1+...+c_kw_k = d_1u_1+...+d_mu_m$ for some choice of scalars $d_1, d_2, ..., d_m$. But $u_1, ..., u_m, w_1, ..., w_k$ is linearly independent. So our last equation implies that all the c's equal 0.

Thus our original equation involving a, b, c becomes

$$a_1u_1 + ... + a_mu_m + b_1v_1 + ... + b_jv_j = 0$$

But we already knew that the list $u_1, ..., u_m, v_1, ..., v_j$ is linearly independent. This equation implies that all the a's and b's are 0. We now know that all a, b and c's are 0, hence proving our original claim. □

Second Proof

Proof. let

$$U = (U \cap V) \oplus U'$$
$$V = (U \cap V) \oplus V'$$

where U' is the complement of $(U \cap V)$ inside U, and V' the complement of $(U \cap V)$ inside V, which both exist by Prop 8.2. Hence, we have that

$$U + V = (U \cap V) \oplus U' \oplus V'$$
$$dim(U + V) = dim(U \cap V) \oplus dim(U') \oplus dim(V')$$
$$dim(U + V) = dim(U \cap V) + dim(U - (U \cap V)) + dim(V - (U \cap V))$$
$$dim(U + V) = dim(U) + dim(V) - dim(U \cap V)$$

Must also show that

$$(U \cap V) \cap (U' \cap V') = 0$$

Let $u \in (U \cap V) \cap (U' + V')$. Then, $u \in (U' + V') \implies u = u' + v' \implies v' \in V' \implies v' = u - u' \in (U \cap V) \implies v' = 0$

And, show

$$U' \cap (U \cap V + V') = 0$$

Finally,

$$V' \cap (U \cap V + U') = 0$$

□

9 Linear Maps

Definition 9.1. (Linear Map) Let U, V be vector spaces. Then, $f : U \to V$ is a linear map if:

1. $f(U + V) = f(U) + f(V)$
2. $f(\lambda U) = \lambda f(U)$

$\iff f(\sum \lambda_i u_i) = \sum \lambda_i f(u_i)$

Example 9.1. Let $f : K^n \to V$, $v_1, ..., v_n \in V$ then,

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \to \sum_i \lambda_i v_i$$

is a linear map.

Theorem 9.1. *Let V and W be vector spaces over F, and suppose that $v_i, v_2, ..., v_n$ is a basis for V. For $w_1, w_2, ..., w_n$ in W ,there exists exactly one linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for $i = 1, 2, ..., n$*

Definition 9.2. (Isomorphism) An isomorphism is a map which is bijective and linear.

Proposition 9.1. The following are equivalent:

1. f is an isomorphism.
2. $u_1, ..., u_n$ is a basis of V.

Proof. Suppose (1), $v_1, ..., v_n$ span, so f is surjective.

$$\forall v \in V, \exists \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{R}^n \; s.t. \; f \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = V \implies V = \sum \lambda_i v_i$$

L.I.: Assume $\sum \lambda_i v_i = 0$, then

$$f \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0 = f \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \implies \lambda_i = 0 \; \forall i$$

So, $v_1, ..., v_n$ is a basis.
Now, suppose (2), we show (1). Clearly, f is linear. Also, f is injective: suppose

$$f \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = f \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

Then, $\sum \lambda_i v_i = \sum \mu_i v_i$, so $\lambda_i = \mu_i$.
f is surjective: take $v \in V$, $\exists \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in K^n \; s.t. f \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = v$ since $v_1, ..., v_n$ spans V. □

Corollary 9.1.1. Every finite dimension vector space is isomorphic to K^n for some n.

Example 9.2. Let $P_n(K)$ be the set of polynomials of degree $\leq n$. Then,

$$P_n(K) \cong K^{n+1}$$

$$a_n x^n + ... + a_1 x + a_0 \iff \begin{pmatrix} a_n \\ \vdots \\ a_0 \end{pmatrix}$$

$$\phi : K^{n+1} \to P_n(K)$$

$$\begin{pmatrix} a_n \\ \vdots \\ a_0 \end{pmatrix} \to \sum a_k X^K \iff (1, x, x^2, ..., x^n) \; is \; basis$$

Definition 9.3. (Kernel) $ker(f) = \{u \in U : f(u) = 0\}$ is a subspace of U.

Theorem 9.2. *Let $f : U \to V$ be linear. Then,*

$$f \text{ is surjective} \iff Im(f) = V$$

$$f \text{ is injective} \iff ker(f) = 0$$

Proof. (injective $\implies ker(f) = 0$):

$$f(u) = 0 = f(0) \implies u = 0$$

($ker(f) = 0 \implies$ injective):

$$\begin{aligned} ker(f) = 0,\ f(u) = f(v) &\implies f(u) - f(v) = 0 \\ &\implies f(u - v) = 0 \quad [\text{since } f \text{ is linear}] \\ &\implies (u - v) \in ker(f) \implies u = v \end{aligned}$$

□

Lemma 9.3. Let $f : U \to V$ be linear, let $U = ker(f) \oplus U'$, where U' is the complement of $ker(f)$ Then, $f|_{U'}$ is injective.

Proof. Suppose $f(u) = f(v)$, with $u, v \in U'$. Then, $u - v \in (U' \cap ker(f)) \implies u = v$ □

Corollary 9.3.1. $f : U' \to f(U)$ is an isomorphism.

Proof.

$$f : U' \to f(U') = f(U)$$

□

Theorem 9.4. *Let $f : U \to V$ be a linear map, U, V finite dimension. Then,*

$$dim(U) = dim(ker(f)) + dim(Im(f))$$

Proof. Let $U = ker(f) \oplus U'$. Since $f : U' \to f(U)$ is an isomorphism,

$$dim(U') = dim(f(U)) = dim(Im(f))$$

$$\implies dim(U) = dim(ker(f)) + dim(U') = dim(ker(f)) + dim(Im(f))$$

□

Corollary 9.4.1. Suppose $f : U \to U$ is a linear map (U f.d.), TFAE:

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

Remark. A linear map from a vector space to itself is called an **Operator**.

Remark. Every linear transformation can be represented by a matrix

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \to A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

9.1 Null space (nullity) and Range (rank)

Definition 9.4. Let V and W be vector spaces, and let $T : V \to W$ be linear. We define the **null space** (or **kernel**) $N(T)$ of T to be the set of all vectors x in V such that $T(x) = 0$; that is, $N(T) = x \in V : T(x) = 0$. We define the **range** (or **image**) $R(T)$ of T to be the subset of W consisting of all images (under T) of vectors in V; that is, $R(T) = T(x) : x \in V$.

Theorem 9.5. *Let V and W be vector spaces, and let $T : V \to W$ be linear. If $\beta = \{v_i, v_2, ..., v_n\}$ is a basis for V, then*

$$Range(T) = span(T(\beta))$$

Definition 9.5. Let V and W be vector spaces, and let $T : V \to W$ be linear. If $N(T)$ and $R(T)$ are finite-dimensional, then we define the nullity of T, denoted $nullity(T)$, and the rank of T, denoted $rank(T)$, to be the dimensions of $N(T)$ and $R(T)$, respectively.

Theorem 9.6. *Let V and W be vector spaces and $T : V \to W$ be linear. Then $N(T)$ and $R(T)$ are subspaces of V and W, respectively.*

Theorem 9.7. *(Dimension Theorem) Let V and W be vector spaces, and let $T : V \to W$ be linear. If V is finite-dimensional, then*

$$nullity(T) + rank(T) = dim(V)$$

9.2 Matrix Representation

Definition 9.6. (**Matrix Representation**) We call the $m \times n$ matrix A defined by $A_{ij} = a_{ij}$ the matrix representation of T in the ordered bases β and γ and write $A = [T]^{\gamma}_{\beta}$.If $V = W$ and $\beta = \gamma$, then we write $A = [T]_{\beta}$.

Notice that the jth column of A is simply $[T(v_j)]_{\gamma}$. Also observe that if $U : V \to W$ is a linear transformation such that $[U]^{\gamma}_{\beta} = [T]^{\gamma}_{\beta}$, then $U = T$.

Definition 9.7. (Kronecker Delta) We define the Kronecker delta δ_{ij} by $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The $n \times n$ identity matrix I_n is defined by $(I_n)_{ij} = \delta_{ij}$.

Theorem 9.8. *Let V and W be finite-dimensional vector spaces with ordered bases β and γ, respectively, and let $T, U : V \to W$ be linear transformations. Then*

1. $[T + U]^{\gamma}_{\beta} = [T]^{\gamma}_{\beta} + [U]^{\gamma}_{\beta}$.
2. $[aT]^{\gamma}_{\beta} = a[T]^{\gamma}_{\beta}$.

Definition 9.8. Let A be an $n \times n$ matrix with entries from a field F. We denote by L_A the mapping $L_A : F^n \to F^m$ defined by $L_A(x) = Ax$ (i.e. the matrix product of x and A) for each column vector $x \in F$. We call L_A a **Left-multiplication Transformation**.

9.3 Scaling transformation

Definition 9.9. (Scaling transformation) $T : U \to U$ is a scaling transformation if $\exists(v_1, ..., v_n)$ basis of U and $\lambda_1, ...\lambda_n \in K$ s.t. $T(v_i) = \lambda_i v_i$.

9.4 Projection

Definition 9.10. (Projection) Let V a v.s., with $V = V_1 \oplus V_2$ be a direct sum decomposition, $\forall v \in V, \exists! v_1 \in V_1, v_2 \in V_2 \; s.t. \; v = v_1 + v_2$. The projection of v onto V_1 is defined as taking only the v_1 element of v. It is a linear transformation.

Proof. Let P be a projection, let $u = u_1 + u_2$, $v = v_1 + v_2$, then

$$u + v = (u_1 + u_2) + (v_1 + v_2)$$

$$P(u + v) = P[(u_1 + u_2) + (v_1 + v_2)]$$

$$P(u + v) = P[(u_1 + v_1) + (u_2 + v_2)]$$

$$P(u + v) = u_1 + v_1$$

$$P(u + v) = P(u) + P(v)$$

And,

$$P(\lambda u) = P(\lambda(u_1 + u_2))$$

$$P(\lambda u) = P(\lambda u_1 + \lambda u_2)$$

$$P(\lambda u) = \lambda u_1 = \lambda P(u_1 + u_2) = \lambda P(u)$$

□

Proposition 9.2. P is a projection $\iff P = P^2$

Proof. For the forward implication, suppose $P : V \to V$ is a projection on the subspace W. Then, let $v \in V$, we have $P(v) = w_1 \in W$, so $P^2(v) = P(P(v)) = P(w_1) = w_1 = P(v)$ since $P(w) = w, \; \forall w \in W$

Now, for the other implication, suppose we have a linear operator $P : V \to V$ such that $P^2 = P$. Since P is a linear operator, we have a decomposition of V into a direct sum

$$V = \ker P \oplus \operatorname{im} P$$

because

1. Let $v \in ker(P) \cap Im(P)$, then $v = P(u)$ and $P(v) = 0$ (because it's in the kernel), but then, since $P = P^2$, we get $0 = P(v) = P^2(u) = P(u) = 0$, so $ker(P) \cap Im(P) = \{0\}$.

2. We show $ker(P) + Im(P)$ spans V. Let $v \in V$, $P(v) \in Im(P)$. Let's show $V = (V - P(V)) + P(v)$, then $(V - P(V))$ must be in $ker(P)$.

   $$P(V - P(V)) = P(V) - P^2(U) = P(V) - P(V) = 0$$

   So, $(V - P(V)) \in ker(P)$ and we have shown that $V = \ker P \oplus \operatorname{im} P$.

We show that P is a projection onto the subspace $\operatorname{im} P$. By definition, $P(V) = \operatorname{im} P$. Secondly, for any w in the image of P we have $P(w) = P(P(v))$ for some $v \in V$. But by assumption $P^2 = P$, so $P(w) = P(v) = w$. Hence P restricts to the identity on $\operatorname{im} P$ and so P is a projection operator. □

Theorem 9.9. *(Complementarity of image and kernel)*

Let W be a finite-dimensional vector space and P be a projection on W. Suppose the subspaces U and V are the image and kernel of P respectively. Then P has the following properties:

1. *P is the identity operator I on U:* $\forall \mathbf{x} \in U : P\mathbf{x} = \mathbf{x}$.

2. *we have a direct sum* $W = U \oplus V$. *Every vector* $\mathbf{x} \in W$ *may be decomposed uniquely as* $\mathbf{x} = \mathbf{u} + \mathbf{v}$ *with* $\mathbf{u} = P\mathbf{x}$ *and* $\mathbf{v} = \mathbf{x} - P\mathbf{x} = (I - P)\,\mathbf{x}$, *and where* $\mathbf{u} \in U, \mathbf{v} \in V$.

The image and kernel of a projection are "complementary", as are P and $Q = I - P$. *The operator Q is also a projection as the image and kernel of P become the kernel and image of Q and vice versa. We say P is a projection along V onto U (kernel/image) and Q is a projection along U onto V.*

9.5 Rotation

Definition 9.11. A rotation is a linear transformation $R : \mathbb{R}^2 \to \mathbb{R}^2$ such that $R(v)$ rotates the vector v by θ degrees. The transformation matrix of a rotation is:

$$R_\theta = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix}$$

where $T(e_1) = \begin{pmatrix} cos\theta \\ sin\theta \end{pmatrix}$ and $T(e_2) = \begin{pmatrix} -sin\theta \\ cos\theta \end{pmatrix}$, with (e_1, e_2) forming a basis of $\mathbb{R}^2$

Remark. $\{\mathbb{R}^2\}$ forms a group.

$$R_\theta^{-1} = R_{-\theta}$$

$$R_{\theta'} R_\theta = R_{\theta' + \theta}$$

Definition 9.12. You can define the group $SO(2)$ as 2×2 matrices:

$$SO(2) = \left\{ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} : \theta \in \mathbb{R} \right\}.$$

So the 2 comes from the fat that you have 2×2 matrices. In general

$$O(n) = \{A \in GL(n) : A^T A = A^T A = I\}. \tag{1}$$

$$SO(n) = \{A \in O(n) : det(A) = 1\}. \tag{2}$$

will be a group of $n \times n$ matrices. Note that all the matrices above are orthogonal and have determinant 1.

Proposition 9.3. The map

$$\pi : (\mathbb{R}, +) \to (SO(2), \times) \subseteq M_2(\mathbb{R})$$

defined by

$$\pi : \theta \to R_\theta$$

is a surjective homomorphism with $ker(\pi) = 2\pi\mathbb{Z}$. It follows, from the first isomorphism theorem, that

$$\mathbb{R}/2\pi\mathbb{Z} \cong SO(2)$$

9.6 Differential operators

Let $V = K[x]$ = the set of all polynomials with coefficients in K. Let $f \in V$ be of the form $f = \sum a_n x^n$.

Definition 9.13.

$$D(f) := \sum n a_n x^{n-1} := f'$$

D defines a linear map $P_n \to P_n$. $X^n \notin Im(D)$ because we lose a degree when taking the derivative. $\implies$ D is not surjective

Proposition 9.4. $D : P_n \to P_n$ is **nilpoltent**, that is, $\exists N \in \mathbb{N}\ s.t.\ D^n = 0$

10 Linear forms

Let V a vector space over K.

Definition 10.1. A **linear form** on V is a linear map $f : V \to K$ (K being the vector space on which V is defined)

Definition 10.2. (**Coordinate vector**) Let $\beta = \{u_1, u_2, ..., u_n\}$ be an ordered basis for a finite-dimensional vector space V. For $x \in V$, let $(a_1, a_2, ..., a_n)$ be the unique scalars such that

$$x = \sum_{i=1}^{n} a_i u_i$$

We define the coordinate vector of x relative to β, denoted $[x]_\beta$, by

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

Example 10.1. Let $v_1, ..., v_n$ a basis of V $\implies$ any $v \in V$ can be uniquely written as $\sum \lambda_i v_i,\ i \in \{1, ..., n\}$

Proposition 10.1. $f : V \to K$ defined by $f : v_i \to \lambda_i$ is a linear form.

Proof. let $v = \sum \lambda_i v_i,\ v' = \sum \lambda'_i v_i$

$$\implies v + v' = \sum (\lambda_i + \lambda'_i) v_i$$

$$f(v + v') = coordinate\ i\ of\ v + v' = \lambda_i + \lambda'_i = f(v) + f(v')$$

And,

$$f(\lambda v) = \lambda f(v)$$

□

Proposition 10.2. Let V a vector space over K. Then, the set of all linear forms on V is a vector space over K. It is called the **Dual Space** and is denoted V^*

$$V^* = L(V, (F)) = \text{space of linear transformations } f : V \to \mathbb{F}$$

If $Dim(v) \leq \infty$ then $V \cong V^*$.

Proof.

$$Dim(V^*) = Dim(L(V, \mathbb{F}))$$

and since

$$Dim(L(V, W)) = Dim(V)Dim(W),$$

we have

$$Dim(V^*) = Dim(V)Dim(\mathbb{F}) = Dim(V) \times 1$$

So, V & V^* have the same dimension hence are isomorphic. □

Proposition 10.3. Let $v_1, ..., v_n$ be a basis of V, let

$$\begin{aligned} f_i : & V \to K \\ & v \to \lambda_i \end{aligned}$$

with $v = \sum \lambda_i v_i$, $1 \leq i \leq n$. Then, $f_1, ..., f_n$ is a basis of V^*.

Corollary 10.0.1. If V is finite dimension, then $dim(V) = dim(V^*)$ ($V \cong V^*$). This isomorphism depends on the basis for V.

$$V \cong V^* \cong V^{**}$$

Proposition 10.4. There exists a natural homomorphism

$$V \to V^{**}$$

that is, the isomorphism is constructed from V itself, without a basis.

Definition 10.3. (**Dual Basis**) We call the ordered basis $\beta^* = f_1, f_2, \ldots, f_n$ of V^* that satisfies $f_i(x_j) = \delta_{ij} (1 < i, j < n)$ the dual basis of β, where δ_{ij} is the Kronecker delta.

Theorem 10.1. *Let $T : V \to W$ a linear transformation, $\beta = \{v_1, ..., v_n\}$ a basis of V and $\gamma = \{w_1, ..., w_m\}$ a basis of W. Then we have a matrix*

$$A = [T]_\beta^\gamma$$

Let $\beta^* = \{f_1, ..., f_n\}$ be the dual basis of V (i.e. basis of V^*) and $\gamma^* = \{g_1, ..., g_m\}$ the basis of W^*. Then,

$$[T^t]_{\gamma^*}^{\beta^*} = A^t$$

Proof. Consider

$$[T^t]_{\gamma^*}^{\beta^*} = \begin{array}{c} \\ f_1 \\ \vdots \\ f_n \end{array} \begin{array}{c} \begin{array}{cccc} T^t(g_1) & \ldots T^t(g_j) \ldots & T^t(g_m) \end{array} \\ \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \end{array}$$

And

$$[T^t]_\beta^\gamma = \begin{array}{c} \\ w_1 \\ \vdots \\ w_k \\ \vdots \\ w_m \end{array} \begin{array}{c} \begin{array}{c} T(v_i) \end{array} \\ \begin{pmatrix} \\ A_{ki} \\ \\ \end{pmatrix} \end{array}$$

Recall that if $f \in V^*$,

$$f = \sum_{i=1}^{n} f(v_i) f_i$$

And

$$T^t(g_j) \in V^* = \sum_{i=1}^{n} T^t(g_j)(v_i) f_i = \sum_{i=1}^{n} g_j(T(v_i)) f_i$$

Now,

$$T(v_i) = \sum_{k=1}^{m} A_{ki} w_k \implies g_j(T(v_i)) = g_j(\sum_{k=1}^{m} A_{ki} W_k) = \sum_{k=1}^{m} A_{ki} g_j(w_k)$$
$$A_{1i} g_j(w_1) + A_{2i} g_j(w_2) + \ldots + A_{ji} g_j(w_j) + \ldots + A_{mi} g_j(w_m)$$

And since the g_j's are dual basis vectors, they are equal to 1 precicely at w_j and 0 elsewhere, we get

$$T(v_i) = A_{ji}$$

To summarize,

$$T^t(g_j) = \sum_{i=1}^{n} A_{ji} f_i = \sum_{i=1}^{n} (A^t)_{ij} f_i$$

so

$$[T^t]_{\gamma^*}^{\beta^*} = A^t$$

□

11 Change of Basis

Theorem 11.1. *Let β and β' be two ordered bases for a finite-dimensional vector space V, and let $Q = [I_v]_{\beta'}^{\beta}$. Then*

1. *Q is invertible.*
2. *For any $v \in V$, $[v]_\beta = Q[v]_{\beta'}$.*

The matrix $Q = [I_v]_{\beta'}^{\beta}$ defined in the above Theorem is called a change of coordinate matrix. Because of part (2) of the theorem, we say that Q changes β'-coordinates into β-coordinates. Observe that if $\beta = \{x_1, x_2, ..., x_n\}$ and $\beta' = \{x'_1, x'_2, ..., x'_n\}$, then

$$x'_j = \sum_{i=1}^{n} Q_{ij} x_i \quad \text{for } j = 1, 2, ..., n$$

That is, the jth column of the matrix Q is $[x'_j]_\beta$, i.e. the coordinate vector of x' in basis β. Moreover, notice that if Q changes β'-coordinates into β-coordinates, then Q^{-1} changes β-coordinates into β'-coordinates.

Example 11.1. Let $e_1 = \begin{pmatrix}1\\0\end{pmatrix}$, $e_2 = \begin{pmatrix}0\\1\end{pmatrix}$ be a canonical basis in $\mathbb{R}^2$. Let $v = \begin{pmatrix}1\\1\end{pmatrix}$ = coord. of v in e-basis. Take $f_1 = \begin{pmatrix}1\\0\end{pmatrix}$, $f_2 = \begin{pmatrix}1\\1\end{pmatrix}$ be a basis of $\mathbb{R}^2$. Then, the coordinates of v in f-basis are $\begin{pmatrix}0\\1\end{pmatrix}$ because $v = 0 \times (f_1) + 1 \times (f_2)$

Example 11.2. Let $f_1 = \begin{pmatrix}1\\2\\3\end{pmatrix}$, $f_2 = \begin{pmatrix}0\\1\\2\end{pmatrix}$ in $\mathbb{R}^3$. Let $v = \begin{pmatrix}1\\1\end{pmatrix}$ in the f-basis. Then, the coordinates of v in $\mathbb{R}^3$ is $v = 1 \cdot f_1 + 1 \cdot f_2 = \begin{pmatrix}1\\3\\5\end{pmatrix}$

Example 11.3. (Change of basis matrix) Let $\begin{pmatrix}1 & 0\\2 & 1\\3 & 2\end{pmatrix}$ be a matrix which changes f-coord to e-coord.

$$\begin{pmatrix}1 & 0\\2 & 1\\3 & 2\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\3\\5\end{pmatrix}$$

Definition 11.1. (**Change of basis matrix**) Let $e_1, ...e_n$ be a basis of V. Let $f_1, ...f_n$ be vectors in V written in e-coordinates. Then, $Q = (f_1, ...f_n)$ is called the change of basis matrix $f - coord \to e - coord$. Where

$$Q\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = f_1$$

Remark. Both $e = e_1, ...e_n$ and $f = e_1, ...f_n$ are bases of V.

$$Q = (f_1, ..., f_n) \ : \ f - coord \to e - coord$$

$$P = (f_1, ..., f_n) \ : \ e - coord \to f - coord$$

$$Q = P^{-1}$$

Corollary 11.1.1. Let $f_1...f_n$ be a basis of V, written in e-coord. Then, $(Q = f_1...f_n)$ is invertible. That is, if the vectors form a basis, the matrix is invertible.

Theorem 11.2. *Let $T : V \to V$ a linear transformation, e_i a basis of V, f_i a basis of V, both of finite dimension. Let A =(matrix of T in e-basis),B =(matrix of T in f-basis),*
Q = (change of basis matrix f-coord → e-coord). Then,

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

12 Application: Coding Theory

Let p a prime number, $\mathbb{F}_p$ = Field with p elements. Let k an integer, and the elements of $\mathbb{F}_p^k$, a vector space, are called words of length k.

Example 12.1. $p = 2, 0101$ a word of length 4 in $\mathbb{F}_2^4$.

Goal: transmit words through a channel.

$$(word) \to_{encoding} (longer\ word) \to_{transmit} (modified\ word) \to_{decode} (original\ word)$$

Definition 12.1. (Linear codes) A linear encoding is an injective map

$$\phi \ : \ \mathbb{F}_p^k \to \mathbb{F}_p^n$$

with $im(\phi)$ = code words

Example 12.2. Let $\phi = \frac{Id}{Id}$

$$\phi : \mathbb{F}_p^k \to \mathbb{F}_p^{2k}$$

e.g. with $p = 2$, $\phi(0101) = 01010101$

$$0101 \to_E 01010101 \to_T 01010100$$

12.1 Problems

1. Make the code words as small as possible.
2. Make sure to recover the original words from the transmitted words, knowing the transmition error rate.

Definition 12.2. (parity code) Let $p = 2$

$$0101 \to 01010$$

The last 0 in the image is a parity ($\sum$ of digits modulo 2)

Example 12.3. Let $\phi = \mathbb{F}_2^k \to \mathbb{F}_2^{k+1}$.

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

(adds parity bit to word)

12.2 Hamming Code

Acts on words of length 4, makes them into words of length 7.

$$\phi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

$$\phi(0101) = 0101010$$

Definition 12.3. (Hamming distance on $\mathbb{F}_p^n$) Let $u, v \in \mathbb{F}_p^n$, $d(u,v) = \#\{i : u_i \neq v_i\}$. e.g.

$$d(0000, 1111) = 4$$

$$d(0000, 0001) = 1$$

Example 12.4. Show the triangle inequality for the hamming distance.

Definition 12.4. (Hamming norm) $||u|| = d(u, 0)$

Example 12.5. Let $\phi = parity\ bit$, then $im(\phi) = \{v \in \mathbb{F}_2^{n+1} \ : \ \sum_1^{n+1} v_i = 0\}$

Proof. Let $v \in im(\phi)$, then $v_{n+1} = \sum_1^n v_i$, so

$$\sum_1^{n+1} v_i = 2\left(\sum_1^n v_i\right) = 0$$

$$im(\phi) \subseteq \left\{v : \sum_1^{n+1} v_i = 0\right\}$$

Other inclusion is left as an exercise. □

Definition 12.5. (Separation between code words) $Sep(\phi) = \min\{d(u,v) : u, v \in \ im(\phi),\ u \neq v\}$, with d the Hamming distance.

Example 12.6. Show $Sep(\phi) = \min\{||u|| \ : \ u \in im(\phi), u \neq 0\}$

13 Revisiting linear systems (chpt 3 in textbook)

End of material for midterm.

13.1 Hyperplanes

Definition 13.1. Let V a vector space (f.d.). A hyperplane is a subspace of $dim(V) - 1$. Equivalently, a hyperplane V in a vector space W is any subspace such that W/V is one-dimensional. Equivalently, a hyperplane is the linear transformation kernel of any nonzero linear map from the vector space to the underlying field.

Example 13.1. $x + y + z = 0$ a hyperplane in $\mathbb{R}^3$.
Implicit definition

$$H = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \ x + y + z = 0 \right\}$$

Parametric definition
$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in H$, then

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y - z \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$H = span\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}\right\}$$

Definition 13.2. (System of k equations)

$$\begin{Bmatrix} a_{11}x_1 + \ldots + a_{1n}x_n = 0 \\ \vdots \\ a_{k1}x_1 + \ldots + a_{kn}x_n = 0 \end{Bmatrix}$$

Remark. The intersection of subspaces is a subspace.

Example 13.2. To find the solution of a system: Gaussian elimination.
Elementary operations:

1. swap two rows $R_i \leftrightarrow R_j$
2. Mutltiply by a non zero constant $R_i \rightarrow \lambda R_i$
3. Add to a row a multiple of a different row $R_i \rightarrow R_i + \lambda R_j$

13.2 Elementary Matrices

Definition 13.3. An elementary matrix is a matrix which differs from the identity matrix by one single elementary row operation. The elementary matrices generate the general linear group $GL_n(F)$ when F is a field. Left multiplication (pre-multiplication) by an elementary matrix represents elementary row operations, while right multiplication (post-multiplication) represents elementary column operations.

1. Swap $R_i \leftrightarrow R_j$

   Remark. To act on the rows, multiply by an elementary matrix on the left, to act on the columns, multiply by a elementary matrix on the right.

   Example 13.3.

$$\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} = \begin{pmatrix}1&2&3\\7&8&9\\4&5&6\end{pmatrix}$$

2. $R_i \rightarrow \alpha R_i$, $\alpha \neq 0$

$$\begin{pmatrix}1&0&0\\0&2&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&2&3\\1&2&3\\1&2&3\end{pmatrix} = \begin{pmatrix}1&2&3\\2&4&8\\1&2&3\end{pmatrix}$$

3. $R_i \rightarrow R_i + \alpha R_j$, $i \neq j$

Theorem 13.1. *Elementary matrices are invertible, and the inverse of an elementary matrix is an elementary matrix of the same type.*

Definition 13.4. (REF) Let A be a matrix, then $\exists(E_1, \ldots, E_n)$ elementary matrices such that

$$E_n \circ \ldots \circ E_2 \circ E_1 \circ A = REF(A)$$

Theorem 13.2. *Let $A \in M_{m\times n}(F)$, and suppose that B is obtained from A by performing an elementary row [column] operation. Then there exists an $m \times m$ $[n \times n]$ elementary matrix E such that $B = EA$ $[B = AE]$. In fact, E is obtained from I_m $[I_n]$ by performing the same elementary row [column] operation as that which was performed on A to obtain B. Conversely, if E is an elementary $m \times m$ $[n \times n]$ matrix, then EA $[AE]$ is the matrix obtained from A by performing the same elementary row [column] operation as that which produces E from I_m $[I_n]$.*

13.3 Rank theorem for systems

Let $V = K^n$,

$$X = \begin{Bmatrix} eq_1 = 0 \\ \vdots \\ eq_k = 0 \end{Bmatrix} \iff AX = 0$$

We have k equations, n variables. Let S =(subspace of solutions). The dimension of S can be found with the theorem below. We view A as a linear transformation matrix:

$$A : K^n \to K^k$$

A is a matrix with k rows and n columns.

$$S = ker(A) = \{x \in K^n : \ AX = 0\}$$

Theorem 13.3.

$$dim(S) = (num\ of\ variables\ (n)) - (\text{"}True\text{"}\ num\ of\ equations),$$

where S is the set of solutions. To find the true number of equations, start with A, apply gaussian elimination, and the non zero rows in the REF is the number of true equations.

Definition 13.5. (Row space of A) Take only the rows of the matrix A.

$$Row(A) = span\{r_1, ..., r_k\} \subseteq K^n,$$

where

$$A = \begin{pmatrix} r_1 \\ \vdots \\ r_k \end{pmatrix}$$

Definition 13.6. (Column space of A) Take only the columns of A.

$$col(A) = span\{c_1, ..., c_k\}$$

Remark. True number of equations: $dim(Row(A))$

Theorem 13.4. *(old rank theorem)*

$$dim(V) = dim(ker(A)) + dim(Col(A))$$

Definition 13.7. If $A \in M_{n\times m}(K)$, we define the rank of A to be the rank of the linear transformation $L_A : K^n \to K^m$.

Theorem 13.5. *Elementary row and column operations are rank preserving.*

Theorem 13.6. *The rank of any matrix equals the maximum number of its linearly independent columns; that is, the rank of a matrix is the dimension of the subspace generated by its columns.*

$$Rank(A) = dim(Im(A)),\ where\ Im(A) = Col(A)$$

Theorem 13.7. *(Rank transpose theorem)*

$$rank(A) = rank(A^t)$$

$$Col(A) = Row(A^t)$$

Proof. Will come later. □

Corollary 13.7.1. $dim(Row(A) = dim(Col(A)))$

Corollary 13.7.2. Let A be an $n \times m$ matrix of rank r. Then, there exist invertible matrices $B_{m\times m}$ and $C_{n\times n}$ such that $D = BAC$

Theorem 13.8. *The rank of any matrix equals the maximum number of its linearly independent columns; that is, the rank of a matrix is the dimension of the subspace generated by its columns. Moreover, the rows and columns of any matrix generate subspaces of the same dimension, numerically equal to the rank of the matrix.*

Theorem 13.9. *Let A be an $m \times n$ matrix, if P and Q are invertible $m \times m$ and $n \times n$ matrices, respectively, then*

(a) $rank(AQ) = rank(A)$,
(b) $rank(PA) = rank(A)$, *and therefore,*
(c) $rank(PAQ) = rank(A)$.

Theorem 13.10. *(Reduction to a Projection) Let A be an $(m \times n)$ matrix of rank r. Then $r \leq m$, $r \leq n$, and $\exists E_1, \ldots, E_n$ and $F_1, \ldots, F_m$ elementary matrices such that*

$$E_n \circ \ldots \circ E_1 \circ A \circ F_1 \circ \ldots \circ F_m = \begin{pmatrix} Id_r & 0_1 \\ 0_2 & 0_3 \end{pmatrix} = D$$

where $0_1, 0_2, 0_3$ are zero matrices. Thus, $D_{ii} = 1$ for $i \leq r$ and $D_{ij} = 0$ otherwise. That is, the elementary matrices on the left act on the columns and the elementary matrices on the right act on the rows to reduce A to a projection matrix.

Corollary 13.10.1. Let A be an m x n matrix of rank r. Then there exist invertible matrices B and C of sizes $m \times m$ and $n \times n$, respectively, such that $D = BAC$, where

$$D = \begin{pmatrix} Id_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

is the $m \times n$ matrix in which O_i are zero matrices.

Theorem 13.11. $rank(A) = rank(A^t)$

Proposition 13.1. $rank(A) = r$ if $A \to \begin{pmatrix} Id_r & 0 \\ 0 & 0 \end{pmatrix}$

Example 13.4. Consider the matrix

$$A = \begin{pmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix} \xrightarrow{\text{elementary operations}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = D$$

So, since we know $rk(A) = rk(D)$, we get that $rank(A) = 3$

Lemma 13.12. If E, F are invertible, then $rank(A) = rank(EAF)$, which implies the proposition above by induction on the theorem.

Corollary 13.12.1. $rk(A) = rk(A^t)$

Proof.

$$rk(A) = E_n \circ \ldots \circ E_1 \circ A \circ F_1 \circ \ldots \circ F_m = \begin{pmatrix} Id_r & 0 \\ 0 & 0 \end{pmatrix} = r$$

$$rk(A^t) = (E_n \circ \ldots \circ E_1 \circ A \circ F_1 \circ \ldots \circ F_m)^t = \begin{pmatrix} Id_r & 0 \\ 0 & 0 \end{pmatrix}^t = r$$

□

13.4 Invertible matrices

Theorem 13.13. *A is an $n \times n$ invertible matrix $\iff$ $rank(A) = n$*

Theorem 13.14. *Every invertible matrix is the product of elementary matrices.*

Example 13.5. To find the inverse of a matrix; let A be an $n \times n$ invertible matrix, and consider the $n \times 2n$ augmented matrix $C = (A|I_n)$. Then,

$$A^{-1}C = (A^{-1}A|A^{-1}I_n) = (I_n|A^{-1})$$

That is, it is possible to transform the augmented matrix $C = (A|I_n)$ into $(I_n|A^{-1})$ with a finite number of elementary matrices, finding the inverse matrix.

Theorem 13.15. *Let $Ax = b$ be a system of n linear equations in n unknowns. If A is invertible, then the system has exactly one solution, namely, $A^{-1}b$. Conversely, if the system has exactly one solution, then A is invertible.*

Definition 13.8. Let V and W be vector spaces. We say that V is isomorphic to W if there exists a linear transformation $T : V \to W$ that is invertible. Such a linear transformation is called an isomorphism from V onto W.

Theorem 13.16. *Let V and W be finite-dimensional vector spaces (over the same field). Then V is isomorphic to W if and only if $dim(V) = dim(W)$.*

Theorem 13.17. *Let $Ax = b$ be a system of linear equations. Then the system is consistent if and only if $rank(A) = rank(A|b)$.*

13.5 The General Linear Group

Theorem 13.18. *Let $A_{n\times n}$ an invertible matrix. Then, $\exists E_1, \ldots, E_k$ such that*

$$A = E_1 \ldots E_k$$

viz. any invertible matrix is a product of elementary matrices.

Proof. $A \to E_n \ldots E_1 A F_1 \ldots F_m = Id_r \implies rk(A) = r$ □

Definition 13.9. $GL_n(K)$ is the set of all invertible matrices $(n \times n)$, this is a group.

Corollary 13.18.1. $GL_n(K)$ is generated by the elementary matrices. It is a symmetric generating set S because $S = S^{-1}$, as inverses of elementary matrices are also elementary matrices.

Definition 13.10. (Generating set) Let G a group, then S is a generating set if every $s \in G$ can be written as $s = \pi s_i$, where $s_i \in S \cup S^{-1}$. A generating set is a set of vectors that spans a vector space. All the vectors in the space can be written as a linear combination of the vectors of the generating set.

13.6 The symmetric group

Definition 13.11. Let X a set. $Sym(X)$ = the set of all bijections $X \to X$. $X = \{1, ..., n\}$, $S_n = Sym(X)$

Definition 13.12. (Cycle decomposition of a permutation) Let $\sigma \in S_n \implies \sigma$ is a bijection. Any bijection can be written as a product of disjoint cycles in a unique way. e.g.

$$\sigma = (x_1 x_2 x_3 x_4)(x_5 x_6 x_7)$$

$$\sigma = (1\ 2\ 3\ 4)(5\ 6\ 7) = (5\ 6\ 7)(1\ 2\ 3\ 4)$$

Lemma 13.19. Every cycle is a product of transpositions which is a transformation which permutes two elements, while fixing every other element, e.g. (xy).

Example 13.6. $(123) = (13)(12) = (32)(31)$

Theorem 13.20. *S_n is generated by transpositions.*

13.7 Permutation matrices

$$S_n \leftrightarrow GL_n(K)$$

Let $\sigma \in S_n$. Permutation matrix:

$$A_\sigma = \begin{cases} a_{ij} = 1 & if \quad \sigma(j) = i \\ a_{ij} = 0 & if \quad \sigma(j) \neq i \end{cases}$$

Example 13.7. For example, the permutation matrix P_π corresponding to the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 2 & 5 & 3 \end{pmatrix}$$

is

$$P_\pi = \begin{bmatrix} \mathbf{e}_{\pi(1)} \\ \mathbf{e}_{\pi(2)} \\ \mathbf{e}_{\pi(3)} \\ \mathbf{e}_{\pi(4)} \\ \mathbf{e}_{\pi(5)} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_4 \\ \mathbf{e}_2 \\ \mathbf{e}_5 \\ \mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Observe that the j^{th} column of the I_5 identity matrix now appears as the $\pi(j)$th column of P_π.

Proposition 13.2. The map

$$S_n \to GL_n$$
$$\sigma \to A_\sigma$$

is an injective homomorphism.

14 The Determinant

14.1 Determinant of order 2

Definition 14.1. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is a 2×2 matrix with entries from a field F, then we define the determinant of A, denoted $\det(A)$ or $|A|$, to be the scalar $ad - bc$.

The determinant in itself is not a linear transformation. Nevertheless, the determinant does possess an important linearity property, which is explained in the following theorem.

Theorem 14.1. *The function* $\det : M_{2\times 2}(F) \to F$ *is a linear function of each row of a* 2×2 *matrix when the other row is held fixed. That is, if u, v, and w are in* F^2 *and k is a scalar, then*

$$\det \begin{pmatrix} u + kv \\ w \end{pmatrix} = \det \begin{pmatrix} u \\ w \end{pmatrix} + k \det \begin{pmatrix} v \\ w \end{pmatrix}$$

14.2 Determinant of order n

Definition 14.2. Given $A \in M_{n\times n}(F)$, for $n > 2$, denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row i and column j by $\tilde{A}_{ij}$.

Definition 14.3. (**Laplace Cofactor Expansion**) Let $A \in M_{n\times n}(F)$. If $n = 1$ so that $A = (A_{11})$, we define $\det(A) = (A_{11})$. For $n > 2$, we define $\det(A)$ recursively as

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j})$$

The scalar $(-1)^{1+j} \cdot \det(\tilde{A}_{1j})$ is called the **cofactor** of the entry of A in row i, column j.

Letting

$$c_{ij} = (-1)^{1+j} \cdot \det(\tilde{A}_{1j}),$$

we express the formula for the determinant of A as

$$\det(A) = A_{11}c_{11} + A_{12}c_{12} + \ldots + A_{1n}c_{1n}.$$

Thus the determinant of A equals the sum of the products of each entry in row 1 of A multiplied by its cofactor.

Theorem 14.2. *The determinant of an $n \times n$ matrix is a linear function of each row when the remaining rows are held fixed. That is, for $1 < r < n$, we have*

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u + kv \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix} + k \det\begin{pmatrix} a_1 \\ \vdots \\ a_{r-1} \\ v \\ a_{r+1} \\ \vdots \\ a_n \end{pmatrix}$$

Corollary 14.2.1. If $A \in M_{n\times n}(F)$ has a row consisting entirely of zeros, then $det(A) = 0$.

Theorem 14.3. *The determinant of a square matrix can be evaluated by cofactor expansion along any row.*

Theorem 14.4. *The following rules summarize the effect of an elementary row operation on the determinant of a matrix $A \in M_{n\times n}(F)$.*

1. *Let $A \in M_{n\times n}(F)$, and let B be a matrix obtained by adding a multiple of one row of A to another row of A. Then* $\det(B) = \det(A)$.
2. *If $A \in M_{n\times n}(F)$ and B is a matrix obtained from A by interchanging any two rows of A, then* $\det(-B) = -\det(A)$.
3. *If B is a matrix obtained by multiplying a row of A by a nonzero scalar k, then* $\det(B) = k\det(A)$.

These facts can be used to simplify the evaluation of a determinant along with the following theorem.

Theorem 14.5. *The determinant of an upper triangular matrix is the product of its diagonal entries.*

Corollary 14.5.1. If $A \in M_{n\times n}(F)$ has two identical rows, then $det(A) = 0$.

Corollary 14.5.2. If $A \in M_{n\times n}(F)$ has rank less than n, then $det(A) = 0$.

Remark. Since we can transform any square matrix into an upper triangular using elementary row operations, to evaluate its determinant:
(1) reduce to upper triangular
(2) multiply the diagonal entries together
(3) multiply the obtained scalar by -1 if you interchanged rows
(4) the result is the determinant.

14.3 Properties of the determinant

Theorem 14.6. *For any $A, B \in M_{n\times n}(F)$, $det(AB) = det(A) \cdot det(B)$.*

Corollary 14.6.1. A matrix $A \in M_{n\times n}(F)$ is invertible if and only if $\det(A) \neq 0$. Furthermore, if A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$.

Theorem 14.7. *For any $A \in M_{n\times n}(F)$,* $\det(A^t) = \det(A)$.

14.4 Lebnitz formula

Theorem 14.8. *There exists a unique homomorphism*

$$\varepsilon : S_n \to \mathbb{Z}/2\mathbb{Z}$$

and it is surjective. It verifies

$$\varepsilon(ab) = -1$$
$$\varepsilon(abc) = 1$$
$$\varepsilon(abcd) = -1$$

And,

$$(123) = (32)(31)$$
$$\varepsilon(123) = \varepsilon(32)\varepsilon(31)$$
$$(1) = (-1)(-1)$$

Definition 14.4. The kernel of the ε homomorphism is the alternating group, viz. the permutations which can be expanded into an even number of transpositions get sent to 0.

$$A_n = \ker(\varepsilon)$$

Let $A \in M_n(k)$, then

$$\det(A) = \sum_{\sigma \in S_n} \underbrace{\varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n}}_{a_{21}a_{32}a_{13} \text{ if } \sigma = (123)}$$

Example 14.1. Let A_σ = permutation matrix.

$$\begin{aligned}\det(A_\sigma) &= \sum_{z \in S_n} \varepsilon(z) \underbrace{a_{z(1)1} \dots a_{z(n)n}}_{a_{ij}=1 \iff \sigma(j)=1} \\ &= \varepsilon(\sigma) a_{\sigma(1)1} \dots a_{\sigma(n)n} \\ &= \varepsilon(\sigma)\end{aligned}$$

Example 14.2. let A_σ a permutation matrix.

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

$$A_{\sigma(e_i)} = e_{\sigma(i)}$$

If $\sigma = (123)$,

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = A_\sigma$$
$$\det(A_\sigma) = \varepsilon(\sigma)$$

Example 14.3.

$$\det\begin{pmatrix} 0 & 0 & \lambda_1 \\ \lambda_2 & 0 & 0 \\ 0 & \lambda_3 & 0 \end{pmatrix} = \varepsilon(123)\lambda_1\lambda_2\lambda_3 = \lambda_1\lambda_2\lambda_3$$

Example 14.4.

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

It is an upper triangular matrix, hence a permutation σ contributes to the sum if and only if $\sigma(i) \leq i$, because $a_{\sigma(n)n} = 0$ if $\sigma(n) > n$.

Lemma 14.9. Suppose $\sigma(i) \leq i \ \forall i$, i.e. upper triangular matrix. Then $\sigma = id$.

Proof. Let σ such that $\sigma(i) \leq i$

$$\begin{aligned} \sigma(1) \leq 1 &\rightarrow \sigma(1) = 1 \\ \sigma(2) \leq 2 &\rightarrow \sigma(2) = 2 \\ \sigma(3) \leq 3 &\rightarrow \sigma(3) = 3 \\ &\vdots \\ \sigma(n) &= id \end{aligned}$$

□

Definition 14.5. A definition of the determinant:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma_i}$$

In hopefully simpler terms, we range over all possible "patterns" (*ways of picking n entries such that every row and every column is represented exactly once*), computing the product of the entries in the pattern, applying a sign change depending on the number of pairs of entries in the pattern where one is above and to the right of other entries in the pattern (*if odd, change sign. if even, keep sign the same*), and then summing.

15 Diagonalization

15.1 Eigenvalues and Eigenvectors

15.2 Multilinearity

Definition 15.1. A **multilinear map** is a function of several variables that is linear separately in each variable. More precisely, a multilinear map is a function

$$f : V_1 \times \ldots \times V_n \rightarrow W,$$

where $V_1 \times \ldots \times V_n$ and W are vector spaces, with the following property: for each i, if all of the variables but v_i are held constant, then $f(v_1, ..., v_i, ...v_n)$ is a linear function of v_i. A multilinear map of one variable

is a linear map, and of two variables is a bilinear map. More generally, a multilinear map of k variables is called a k-linear map. If the codomain of a multilinear map is the field of scalars, it is called a multilinear form.

Definition 15.2. Let V, W and X be three vector spaces over the same base Field F. A bilinear map is a function

$$B : V \times W \to X$$

such that for all $w \in W$, the map B_w

$$v \mapsto B(v, w)$$

is a linear map from V to X, and for all $v \in V$, the map B_v

$$w \mapsto B(v, w)$$

is a linear map from W to X. In other words, when we hold the first entry of the bilinear map fixed while letting the second entry vary, the result is a linear operator, and similarly for when we hold the second entry fixed.

Such a map B satisfies the following properties.

1. For any $\lambda \in F$, $B(\lambda v, w) = B(v, \lambda w) = \lambda B(v, w)$.
2. The map B is additive in both components: if $v_1, v_2 \in V$ and $w_1, w_2 \in W$, then $B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$ and $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$.