Linear Algebra Notes

Samuel Lindskog

August 15, 2024

Vector Spaces

Definition 1.1 (Fields)**.** A field F is a set on which operations $+$ and $\cdot$ (addition and multiplication, respectively), defined so that for each pair of elements $x, y \in F$, there exists unique elements $x + y$ and $x \cdot y$ in F for which the following conditions hold for all $a, b, c, 0, 1 \in F$:

$a + b = b + a$ and $a \cdot b = b \cdot a$ (commutativity)
$(a + b) + c = a + (b + c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (associativity)
$\exists 0 \exists 1 (0 + a = a$ and $1 \cdot a = a$ and $0 \neq 1)$ (identity elements)
$\forall a \forall b (b \neq 0) \exists c \exists d (a + c = 0$ and $b \cdot d = 1)$ (inverse elements)
$a \cdot (b + c) = a \cdot b + a \cdot c$ (distributivity)

Definition 1.2 (Vector Space)**.** A vector space V over a field F consists of a set on which two operations (vector 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 $a \in F$ and for each element $x \in V$ there is a unique element $ax \in V$ such that the following conditions hold:

$\forall x, y \in V (x + y = y + x)$ (commutativity)
$\forall x, y, z \in V ((x + y) + z = x + (y + z))$ (associativity)
$\exists 0 \in V \forall x \in V (x + 0 = x)$ (zero vector)
$\forall x \in V \exists y \in V (x + y = 0)$ (additive inverse)
$\forall x \in V (1x = x)$ (identity element)
$\forall a, b \in F \forall x \in V ((ab)x = a(bx))$ (associativity)
$\forall x, y \in V \forall a \in F (a(x + y) = ax + ay)$ (distributivity)
$\forall a, b \in F \forall x \in V ((a + b)x = ax + bx)$ (distributivity)

Definition 1.3 (N-Tuple)**.** An object of the form $(a_1, a_2, \ldots, a_n)$, with $a_1, \ldots, a_n \in F$ is called an n-tuple with entries from F. The elements $a_1, \ldots, a_n$ are called entries or components of the n-tuple.

Notation. The set of all n-tuples with entries from a field F is denoted F^n. Incidentally this set is a vector space over F.

Definition 1.4 (Matrix)**.** An $m \times n$ matrix with entries from a field F

is a rectangular array of the form:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & \dots & a_{mn} \end{pmatrix}$$

where each entry a_{ij} $(1 \leq i \leq m,\ 1 \leq j \leq n)$ is an element of F.

Matrix addition and scalar multiplication for $M_{m\times n}(F)$

The set of all $m \times n$ matrices with entries from a field F is a vector space, which we denote by $M_{m\times n}(F)$, with the following operations of matrix addition and scalar multiplicaiton defined as follows:

For $A, B \in M_{m\times n}(F)$ and $c \in F$:

$$(A+B)_{ij} = A_{ij} + B_{ij}$$
$$(cA)_{ij} = cA_{ij}$$

Theorem 1.1. Vector Addition Cancallation Law If x, y, z are vectors in a vector space V such that $x + z = y + z$, then $x = y$.

Theorem 1.2. In any vector space V, the following statements are true:

$0x = 0$ for all $x \in V$

$(-a)x = -(ax) = a(-x)$ for each $a \in F$ and each $x \in V$

$a0 = 0$ for all $a \in F$

Definition 1.5 (Subspace). A subset W of a vector space V over a field F is called a subspace of V if W is a vector space over F with the operations of addition and scalar multiplication defined on V.

Properties of a subspace

A subset W of a vector space V is a subspace of V iff:

$x \in W \wedge y \in W \Rightarrow x + y \in W$

$c \in F \wedge x \in W \Rightarrow cx \in W$

$0 \in W$

each vector in W has an additive inverse in W (redundant)

In any vector space V, V and $\{0\}$ are subspaces of V.

Theorem 1.3. Let V be a vector space and $W \subseteq V$. Then W is a subspace of V iff:

$$x \in W \wedge y \in W \Rightarrow x + y \in W$$
$$c \in F \wedge x \in W \Rightarrow cx \in W$$
$$0 \in W$$

Definition 1.6 (Transpose Matrix)**.** The transpose A^t of a matrix A is defined by $(A^t)_{ij} = A_{ji}$.

Definition 1.7 (Symmetric Matrix)**.** A symmetric matrix is a matrix A such that $A^t = A$

Theorem 1.4. Any intersection of subspaces of a vector space V is a subspace of V.

The Sum of Subsets of a Vector Space

If S_1 and S_2 are nonempty subsets of a vector space V, then the sum of S_1 and S_2, denoted $S_1 + S_2$ is the set $\{x + y \mid x \in S_1 \wedge y \in S_2\}$.

Definition 1.8 (Direct Sum)**.** A vector space V is called the direct sum of W_1 and W_2 if W_1 and W_2 are subspaces of V such that $W_1 \cap W_2 = \{0\}$ and $W_1 + W_2 = V$. We denote this with $V = W_1 \bigoplus W_2$.

Definition 1.9 (Skew-Symmetric Matrix)**.** A matrix M is called skew-symmetric if $M^t = -M$.

Definition 1.10 (Linear Combination)**.** Let V be a vector space over F, and S be a nonempty subset of V. A vector $v \in V$ is called a linear combination of vectors in S if there exist a finite number of vectors $u_1, u_2, \ldots, u_n$ in S and scalars $a_1, a_2, \ldots, a_n$ in F such that $v = a_1 u_1, a_2 u_2, \ldots, a_n u_n$.

Definition 1.11 (Span)**.** Let S be a nonempty subset of a vector space V. The span of S, is the set consisting of all linear combinations of the vectors in S. For convenience, we define span$(\emptyset) = \{0\}$.

Theorem 1.5. 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 1.12 (Generate)**.** A subset S of a vector space V generates (or spans) V if span$(S) = V$. We can also say that the vectors of S generate (or span) V.

Definition 1.13 (Linear Dependence)**.** A subset S of a vector space V is called linearly dependent if there exist a finite number of distinct vectors $u_1, u_2, \ldots, u_n$ in S and scalars $a_1, a_2, \ldots, a_n$, not all zero, such that:

$$a_1 u_1 + a_2 u_2 + \ldots + a_n u_n = 0$$

Definition 1.14 (Linear Independence)**.** A subset S of a vector space that is not linearly dependent is called linearly independent.

Theorem 1.6. 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 1.1. If S_2 is linearly independent, then S_1 is linearly independent.

Theorem 1.7. 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 if and only if $v \in \text{span}(S)$.

Proof: Suppose S is a linearly independent subset of a vector space V over field F, and v a vector in V such that $v \notin S$. Suppose the set $S \cup \{v\}$ of vectors in V is linearly dependent. It follows from the definition of linear dependence that we have n coefficients $a_i \in F$ and n distinct vectors $x_i \in S \cup \{v\}$ such that $0 = \sum_{i=1}^{n} a_i x_i$ with not all $a_i = 0$. Following the fact that S is linearly independent, it must be so that if $S \cup \{v\}$ is linearly dependent at least one of the distinct vectors from $S \cup \{v\}$ satisfying linear dependence must not be from S, and must therefore be v. By assigning v to a_1, we see:

$$\begin{aligned}
0 &= \sum_{i=1}^{n} a_i x_i \\
-a_1 v &= \sum_{i=2}^{n} a_i x_i \\
v &= (-a_1)^{-1} \sum_{i=2}^{n} a_i x_i \\
v &= \sum_{i=2}^{n} (-a_1)^{-1} a_i x_i
\end{aligned}$$

Thus v must be a linear combination of vectors in S. Therefore if $S \cup \{v\}$ is linearly dependent, $v \in \text{span}(S)$.

To prove the converse, suppose that $v \in \text{span}(S)$. It follows that v is a linear combination of a finite number of vectors in S. By assigning scalars $a_i \in F$ and distinct vectors $x_i \in S$ for n vectors in S, we can form a linear combination of vectors in S with $-v$ such that:

$$0 = -v + \sum_{i=1}^{n} a_i x_i$$

Note the scalar -1 multiplying v. Because v is a linear combination of vectors in S, there always exists scalars $a_i \in F$ such that $v = \sum_{i=1}^{n} a_i x_i$, and thus in the above equation can be satisfied for any $v \in F$. Because the v coefficient in the above equation is -1, it follows that there always exists a non-zero coefficient such that a linear combination of v and distinct vectors from S equals zero. Thus, $v \in \text{span}(S)$ implies $S \cup \{v\}$ is linearly dependent. □

Definition 1.15 (Basis)**.** A basis β for a vector V is a linearly independet subset of V that generates V. If β is a basis of V, we also say that the vectors of β form a basis of V.

Theorem 1.8. Let V be a vector space and $\beta = \{a_1u_1 + a_2u_2 + \ldots + a_nu_n\}$ 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 $a_1, \ldots, a_n$.

Theorem 1.9. 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 1.10 (Replacement)**.** 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. Then $m \leq n$ and there exists a subset H of G containing exactly $n - m$ vectors such that $L \cup H$ generates V.

Proof: It can be proven (dont feel like doing it here) that every set of vectors which generates a vector space has a linearly independent subset of vectors which generates that vector space. Because a linearly dependent set will have greater than or equal to as many elements than any of its linearly independent subsets, proving the replacement theorem for linearly independent generating sets proves the theorem for linearly dependent generating sets.[1] Throughout the proof, subscripts on sets denote their cardinality.

We shall prove the replacement theorem with the modification that G is linearly independent through induction on m, the number of elements in L. Suppose $m = 0$. Then L is the empty set, $\emptyset \cup G$ generates V and is linearly independent.[2]

Suppose $m = k$, and let L_k be an arbitrary set. Case $k < n$: By inductive hypothesis, if L_k is a linearly independent subset of V with k elements, then there exists a subset of G with $n - k$ elements, H_{n-k}, such that $L_k \cup H_{n-k}$ generates V and is linearly independent. Let L_{k+1} be a linearly independent subset of V equal to $L_k \cup \{x\}$, with $x \in V \backslash \text{span}(L_k)$.[3] Such an x exists because H_{n-k} (which is L.I. from L_k by definition) has at least one element. Thus, x is a linear combination of vectors: l_i in L_k with scalars a_i, plus a nonzero[4] linear combination of vectors h_i in H_{n-k} with scalars b_i:

$$x = a_1l_1 + \ldots + a_kl_k + b_1h_1 + \ldots + b_{n-k}h_{n-k}$$

Because there exists a vector h_i with nonzero coefficient b_i, rename these terms vector h_y and coefficient b_y. Let $H_{n-(k+1)}$ be equal to $H_{n-k} \setminus \{h_y\}$. We can now show that h_y can be constructed from a linear combination of vectors in $L_{k+1} \cup H_{n-(k+1)}$, and thus the set

[1] I also like proving this using a linearly independent generating set because as a consequence $L \cup H$ is linearly independent, from which we can extrapolate that the cardinality of a basis for V is unique.

[2] The fact that $L \cup H$ is linearly independent is crucial in establishing the existence of vector x below.

[3] It can be proven that any linearly independent subset of V can be constructed in this way. This fact is crucial in proving the requirement that $m \leq n$.

[4] At least one coefficient b_i must be nonzero in order for x not to be in span(L_k)

$L_k \cup H_{n-k}$ is a subset of $\text{span}(L_{k+1} \cup H_{n-(k+1)})$.

$$x = a_1 l_1 + \ldots + a_k l_k + b_1 h_1 + b_{n-(k+1)} h_{n-(k+1)} + b_y h_y$$
$$\Big(x - (a_1 l_1 + \ldots + a_k l_k + b_1 h_1 + \ldots + b_{n-(k+1)} h_{n-(k+1)})\Big) \cdot b_y^{-1} = h_y$$

Trivially, $L_{k+1} \cup H_{n-(k+1)}$ is a subset of $L_k \cup H_{n-k}$, so $\text{span}(L_{k+1} \cup H_{n-(k+1)}) = V$. This new set is also linearly independent because h_y was a vector in the linearly independent set H_{n-k}, and h_y is not in $\text{span}(L_k)$.

Case $k \geq n$: Because L_k is linearly independent and L_k spans V, there does not exist a vector x in V such that $L_k \cup \{x\}$ is linearly independent. Therefore, $|L_{k+1}|$ must equal k violating the fact that $|L_{k+1}| = k + 1$. Thus the inductive hypothesis is vacuously true for $k + 1$. □

Corollary 1.2. Let V be a vector space having a finite basis. Then every basis for V contains the same number of vectors.

Definition 1.16 (dimension)**.** A vector space is called finite-dimensional is 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 V and is denoted by $\dim(V)$. A vector space that is not finite-dimensional is called infinite-dimensional.

Corollary 1.3. Let W be a subspace of a finite-dimensional vector space V. Then W is finite-dimensional and $\dim(W) \leq \dim V$. Moreover, if $\dim(W) = \dim(V)$, then $V = W$.

Definition 1.17 (Maximal Family)**.** Let $\mathcal{F}$ be a family of sets. A member M of $\mathcal{F}$ is called maximal (with respect to set inclusion) if M is contained in no member of $\mathcal{F}$ other than M itself.

Definition 1.18 (Chain)**.** A family of sets $\mathcal{C}$ is called a chain if for each pair of sets A and B, in $\mathcal{C}$, either $A \subseteq B$ of $B \subseteq A$.[5]

Definition 1.19 (Maximal Principle)**.** Let $\mathcal{F}$ be a family of sets. If, for each chain $\mathcal{C} \subseteq \mathcal{F}$, there exists a member of $\mathcal{F}$ that contains each member of $\mathcal{C}$, then $\mathcal{F}$ contains a maximal member.

Definition 1.20. 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.

[5] Remember from set theory that no two elements of a set are the same.

Theorem 1.11. 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.

Proof: Because β is maximally linearly independent, it follows that $S \setminus \beta$ must be in $\text{span}(\beta)$, and therefore $\text{span}(\beta) = V$. □

Linear Transformations

Definition 1.21 (Linear Transformation)**.** Let V and W be vector spaces over F. We call a function $T : V \to W$ a linear transformation from V to W if for all $x, y \in V$ and $c \in F$, we have:

1. $T(x + y) = T(x) + T(y)$

2. $cT(x) = T(cx)$

Definition 1.22 (Kernel and Image)**.** Let V and W be vector spaces, and let $T : V \to W$ be linear. We define the kernel $N(T)$ of T to be the set of all vectors x in V such that $T(x) = 0$. The image $R(T)$ of T is the subset of W consisting of all images under T of vectors in V.[6]

Definition 1.23 (Rank and Nullity)**.** Let $T : V \to W$ be a linear transformation from vector spaces V to W. If $N(T)$ and $R(T)$ are finite-dimensional, then we define the *nullity* of T and the *rank* of T to be $\dim(N(T))$ and $\dim(R(T))$ respectively.

Theorem 1.12 (Dimension Theorem)**.** Let V and W be vector spaces and let $T : V \to W$ be linear. If V is finite-dimensional, then:

$$\text{nullity}(T) + \text{rank}(T) = \dim(V)$$

Proof: Let A be a basis of $N(T)$. Let B be a linearly independent subset of V. Suppose C is a basis for V equal to $A \cup B$. Such a basis exists because $A \subseteq C$, A is linearly independent, and any maximally linearly independent subset of V is a basis for V. Because C generates V, the cardinality of C is the dimension of V. It follows that because A is linearly independent, that the cardinality of B is $\dim(V) - \text{nullity}(T)$. Because $T(x) = 0$ for any vector x in $\text{span}(A)$, it follows from definition of a linear transformation that $T(\text{span}(A) + \text{span}(B)) = T(\text{span}(B))$, and thus $T(V) = T(\text{span}(B))$. Because linearly independent vectors in B are linearly independent in $T(B)$, we see $\text{rank}(T) = |B|$, and thus $\text{nullity}(T) + \text{rank}(T) = \dim(V)$. □

Definition 1.24 (T-Invariant)**.** Let V be a vector space and $T : V \to V$ be linear. A subspace W of V is said to be *T-invariant* if $T(x) \in W$ for every $x \in W$.

[6] The nullity of T is the dimension of the kernel, the rank of T is the dimension of the image of T.

Definition 1.25 (Ordered Basis)**.** Let V be a finite-dimensional vector space. An ordered basis for V is a basis for V endowed with a specific order i.e. it is a sequence/tuple.

Definition 1.26 (Standard Ordered Basis)**.** For the vector space F^n, we call $\{e_1, e_2, \ldots, e_n\}$ the *standard ordered basis* for F^n.

Definition 1.27 (Coordinate Vector)**.** Let $\beta = (u_1, u_2, \ldots, u_n)$ be an ordered basis for a finite dimensional vector space V. For $x \in V$, let $a_1, a_2, \ldots, a_n$ be the unique scalars such that

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

The coordinate vector of x relative to β, denoted $[x]_\beta$ is:

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