EE 441: General vector spaces over a field

I. AXIOMS FOR A VECTOR SPACE \( V \) OVER A FIELD \( \mathbb{F} \)

Let \( \mathbb{F} \) be a general field. Let \( V \) be a set of objects called “vectors.” We say that \( V \) is a vector space over the field \( \mathbb{F} \) if there are rules for vector addition and scalar multiplication such that \( V \) is closed with respect to these operations, that is:

- \( v_1 + v_2 \in V \) for any two vectors \( v_1, v_2 \in V \).
- \( \alpha v \in V \) for any vector \( v \in V \) and any scalar \( \alpha \in \mathbb{F} \).

and such that the following additional six properties hold:

1) (Commutativity)

\[
v_1 + v_2 = v_2 + v_1 \quad \text{(for all } v_1, v_2 \in V)\]

2) (Associativity)

\[
(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3) \quad \text{(for all } v_1, v_2, v_3 \in V)\]

\[
(\alpha \beta) v = \alpha (\beta v) \quad \text{(for all } \alpha, \beta, \in \mathbb{F} \text{ and } v \in V)\]

3) (Distributive Properties)

\[
\alpha (v_1 + v_2) = \alpha v_1 + \alpha v_2 \quad \text{(for all } v_1, v_2 \in V, \text{ and } \alpha \in \mathbb{F})\]

\[
(\alpha + \beta) v = \alpha v + \beta v \quad \text{(for all } v \in V, \text{ and } \alpha, \beta \in \mathbb{F})\]

4) (Additive Identity) There exists a vector \( 0 \in V \) such that:

\[
v + 0 = v \quad \text{for all } v \in V\]

5) (Additive Inverse) For every \( v \in V \), there exists a vector \( -v \in V \) such that:

\[
v + (-v) = 0\]

6) (Multiplicative Identity) \( 1v = v \) for every vector \( v \in V \).

Note that the vector space \( \mathbb{R}^n \) over the field \( \mathbb{R} \) satisfies all these properties. Similarly, the set \( \mathbb{F}^n \) is a vector space over \( \mathbb{F} \), where vector addition and scalar multiplication are defined entrywise via the arithmetic of \( \mathbb{F} \). Other examples of vector spaces:

- The vector space \( V \) over the field \( \mathbb{F} \) (where \( \mathbb{F} \) is any general field), consisting of all countably infinite tuples \( (x_1, x_2, x_3, \ldots) \), where \( x_i \in \mathbb{F} \) for all entries \( i \in \{1, 2, \ldots\} \), and where arithmetic is defined entrywise using arithmetic in \( \mathbb{F} \).

- The vector space \( V \) over the field \( \mathbb{R} \), where \( V \) is the set of all continuous functions of time \( t \) for \( t \in (-\infty, \infty) \).

- The vector space \( V \) over the field \( \mathbb{R} \), where \( V \) is the set of all polynomial functions \( f(t) \) of degree less than or equal to \( n \). That is, \( V = \{ f(t) \mid f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_n t^n, \alpha_i \in \mathbb{R} \} \).

II. SIMPLE LEMMAS FOR VECTOR SPACES

Let \( V \) be a vector space over a field \( \mathbb{F} \).

**Lemma 1:** (Uniqueness of 0) The vector \( 0 \in V \) is the unique additive identity.

**Proof:** Suppose that \( w \) satisfies \( v + w = v \) for any \( v \in V \). Then adding \( -v \) to both sides of the equation \( v + w = v \) yields:

\[
-v + v + w = -v + v
\]

and hence: \( 0 + w = 0 \). Therefore, \( w = 0 \). \( \square \)

**Lemma 2:** (Uniqueness of \(-v\)) For any vector \( v \in V \), if there is a vector \( w \in V \) such that \( v + w = 0 \), then \( w = -v \).

**Proof:** Suppose that \( v + w = 0 \). Adding \(-v \) to both sides yields \( 0 + w = -v \), and hence \( w = -v \). \( \square \)

**Lemma 3:** \( 0v = 0 \) for any \( v \in V \).

**Proof:** Take any vector \( v \in V \). Then:

\[
v = 1v = (1 + 0)v = 1v + 0v
\]

and hence \( v = v + 0v \). Adding \(-v \) to both sides yields \( 0 = 0v \), proving the result. \( \square \)
Lemma 4: $\alpha 0 = 0$ for any $\alpha \in \mathbb{F}$.

Proof: Take any $\alpha \in \mathbb{F}$ and any vector $v \in \mathcal{V}$. Then:

$$\alpha v = \alpha(v + 0) = \alpha v + \alpha 0$$

Thus, we have $\alpha v = \alpha v + \alpha 0$. Adding $-\alpha v$ to both sides yields $0 = \alpha 0$, proving the result.

Lemma 5: $-v = (-1)v$ for any $v \in \mathcal{V}$.

Proof: Take any vector $v \in \mathcal{V}$. Then:

$$0 = 0v = (1 + (-1))v = 1v + (-1)v = v + (-1)v$$

Thus, $0 = v + (-1)v$. Adding $-v$ to both sides yields $-v = (-1)v$, proving the result.

III. SUBSPACES

Definition 1: Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. Let $S \subset \mathcal{V}$ be a subset of $\mathcal{V}$. We say that $S$ is a subspace if:

$$v_1 + v_2 \in S \quad \text{for all } v_1, v_2 \in S$$

$$\alpha v \in S \quad \text{for all } v \in \mathcal{V}, \alpha \in \mathbb{F}$$

where addition and scalar multiplication are the same in $S$ as they are in $\mathcal{V}$.

It is easy to prove that if $S$ is a subspace of vector space $\mathcal{V}$ over field $\mathbb{F}$, then $S$ is itself a vector space over field $\mathbb{F}$. (It is important to note that $-v = (-1)v$ in proving this...why?).

IV. LINEAR COMBINATIONS AND LINEAR INDEPENDENCE

Definition 2: Let $\{x_1, \ldots, x_k\}$ be a collection of vectors in a vector space $\mathcal{V}$ over a field $\mathbb{F}$. We say that a vector $v \in \mathcal{V}$ is a linear combination of $\{x_1, \ldots, x_k\}$ if it can be written: $v = \alpha_1 x_1 + \ldots + \alpha_k x_k$ for some scalars $\alpha_i \in \mathbb{F}$ for $i \in \{1, \ldots, k\}$.

Definition 3: Let $\{x_1, \ldots, x_k\}$ be a collection of vectors in a vector space $\mathcal{V}$ over a field $\mathbb{F}$. We define $\text{Span}\{x_1, \ldots, x_k\}$ as the set of all linear combinations of $\{x_1, \ldots, x_k\}$. Note that $\text{Span}\{x_1, \ldots, x_k\} \subset \mathcal{V}$.

Definition 4: Let $S$ be a subspace of a vector space $\mathcal{V}$ over a field $\mathbb{F}$. We say that a collection of vectors $\{x_1, \ldots, x_k\}$ span $S$ if $\text{Span}\{x_1, \ldots, x_k\} = S$.

Definition 5: We say that a collection of vectors $\{x_1, \ldots, x_k\}$ in a vector space $\mathcal{V}$ (over a field $\mathbb{F}$) are linearly independent if the equation $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = 0$ can only be true if $\alpha_i = 0$ for all $i \in \{1, \ldots, k\}$.

Definition 6: A collection of vectors $\{x_1, \ldots, x_k\}$ is a basis for a subspace $S$ if the collection $\{x_1, \ldots, x_k\}$ is linearly independent in $S$ and spans $S$.

The following lemmas have proofs that are identical (or nearly identical) to the corresponding lemmas proven in class for the vector space $\mathbb{R}^n$. The proofs are left as an exercise.

Lemma 6: A collection of vectors $\{x_1, \ldots, x_k\}$ are linearly independent if and only if none of the vectors can be written as a linear combination of the others.

Lemma 7: If $\{x_1, \ldots, x_k\}$ are linearly independent in a vector space $\mathcal{V}$, and if $w \in \mathcal{V}$ and $w \notin \text{Span}\{x_1, \ldots, x_k\}$, then $\{x_1, x_2, \ldots, x_k, w\}$ are linearly independent.
Lemma 8: \((k \leq m)\) Let \(\{x_1, \ldots, x_k\}\) be a collection of vectors that are linearly independent in a subspace \(S\). Let \(\{y_1, \ldots, y_m\}\) be a collection of vectors that span \(S\). Then \(k \leq m\).

Lemma 9: Any two bases of a subspace \(S\) have the same size, defined as the dimension of the subspace.

Lemma 10: The dimension of \(F^n\) is \(n\).

Lemma 11: Let \(S\) be a subspace of a vector space \(V\), where \(V\) has dimension \(n\). Then \(S\) has a finite basis, and the dimension of \(S\) is less than or equal to \(n\).

Note that the standard basis for \(F^n\) is given by \(\{e_1, \ldots, e_n\}\), where \(e_i\) is a \(n\)-tuple with all entries equal to 0 except for entry \(i\), which is equal to 1.

Note that the collection \(\{1, t, t^2, \ldots, t^n\}\) is a basis for the vector space \(V\) over the field \(R\), where \(V\) is the space of all polynomial functions of degree less than or equal to \(n\) (why is this true?). Thus, this vector space has dimension \(n + 1\). Note also that, for any \(n\), this vector space is a subspace of the vector space over \(R\) defined by all continuous functions. Thus, the dimension of the vector space of all continuous functions is infinite (as it contains subspaces of dimension \(n\) for arbitrarily large \(n\)).

V. Matrices

Let \(A\) be a \(m \times n\) matrix with elements in \(F\). Note that the equation \(Ax = 0\) (where \(0 \in F^m\) and \(x \in F^n\)) has only the trivial solution \(x = 0 \in F^n\) if and only if the columns of \(A\) are linearly independent.

Lemma 12: A square \(n \times n\) matrix \(A\) (with elements in \(F\)) is non-singular if and only if its columns are linearly independent, if and only if \(Ax = 0\) has only the trivial solution.

The above lemma follows from the fact that if \(A\) is non-singular, it has a single unique solution to \(Ax = b\) for all \(b \in F^n\) (which is true by Gaussian Elimination), and if it is singular it does not have a solution for some vectors \(b \in F^n\) and it has multiple solutions for the remaining \(b \in F^n\).

Lemma 13: Let \(A, B\) be square \(n \times n\) matrices. If \(AB = I\), then both \(A\) and \(B\) are invertible, and \(A^{-1} = B\) and \(B^{-1} = A\). The above lemma follows from the fact that \(AB = I\) implies \(A\) is non-singular (why?) and hence invertible.

Lemma 14: A square \(n \times n\) matrix \(A\) (with elements in \(F\)) has linearly independent columns (and hence is invertible) if and only if its transpose \(A^T\) has linearly independent columns (and hence is invertible). Thus, a square invertible matrix \(A\) has both linearly independent rows and linearly independent columns.

The above lemma follows from the fact that \(AA^{-1} = I\), and hence \((A^{-1})^T A^T = I\).

VI. Basic Probability

The next several lectures on erasure coding will use the following simple but important probability facts:

- If a probability experiment has \(K\) equally likely outcomes, then the probability of each individual outcome is \(1/K\).
- Let \(E_1, E_2, \ldots, E_m\) be a set of independent events (say, from \(m\) independent probability experiments). Then:
  \[ Pr[E_1 \cap E_2 \cap \cdots \cap E_m] = Pr[E_1] Pr[E_2] \cdots Pr[E_m] \]

That is, the probability that all independent events occur is equal to the product of the individual event probabilities.

- (Union Bound) Let \(E_1, E_2, \ldots, E_m\) be a collection of \(m\) events (possibly not independent). Then:
  \[ Pr[E_1 \cup E_2 \cup \cdots \cup E_m] \leq \sum_{i=1}^{m} Pr[E_i] \]

That is, the probability that at least one of the events occurs is less than or equal to the sum of the individual event probabilities.