Chapter 3. Linear Algebra

Section 3.1. Bases

Definition 1 Let $X$ be a vector space over a field $F$. A linear combination of $x_1, \ldots, x_n \in X$ is a vector of the form

$$y = \sum_{i=1}^{n} \alpha_i x_i$$

where $\alpha_1, \ldots, \alpha_n \in F$

$\alpha_i$ is the coefficient of $x_i$ in the linear combination.

If $V \subseteq X$, the span of $V$, denoted $\text{span} V$, is the set of all linear combinations of elements of $V$. The set $V \subseteq X$ spans $X$ if $\text{span} V = X$.

Definition 2 A set $V \subseteq X$ is linearly dependent if there exist $v_1, \ldots, v_n \in V$ and $\alpha_1, \ldots, \alpha_n \in F$ not all zero such that

$$\sum_{i=1}^{n} \alpha_i v_i = 0$$

A set $V \subseteq X$ is linearly independent if it is not linearly dependent.

Thus $V \subseteq X$ is linearly independent if and only if

$$\sum_{i=1}^{n} \alpha_i v_i = 0, \; v_i \in V \; \forall i \Rightarrow \alpha_i = 0 \; \forall i$$

Definition 3 A Hamel basis (often just called a basis) of a vector space $X$ is a linearly independent set of vectors in $X$ that spans $X$.

Example: $\{(1,0),(0,1)\}$ is a basis for $\mathbb{R}^2$ (this is the standard basis).

$\{(1,1),(-1,1)\}$ is another basis for $\mathbb{R}^2$: Suppose

$$(x, y) = \alpha(1,1) + \beta(-1,1) \text{ for some } \alpha, \beta \in \mathbb{R}$$

$$x = \alpha - \beta$$

$$y = \alpha + \beta$$

$$x + y = 2\alpha$$

$$\Rightarrow \alpha = \frac{x + y}{2}$$
\[ y - x = 2\beta \]
\[ \Rightarrow \beta = \frac{y - x}{2} \]
\[ \Rightarrow (x, y) = \frac{x + y}{2}(1, 1) + \frac{y - x}{2}(-1, 1) \]

Since \((x, y)\) is an arbitrary element of \(\mathbb{R}^2\), \{\(1, 1\), \((-1, 1)\}\} spans \(\mathbb{R}^2\). If \((x, y) = (0, 0)\),
\[ \alpha = \frac{0 + 0}{2} = 0, \quad \beta = \frac{0 - 0}{2} = 0 \]

so the coefficients are all zero, so \{\(1, 1\), \((-1, 1)\}\} is linearly independent. Since it is linearly independent and spans \(\mathbb{R}^2\), it is a basis.

**Example:** \{\((1, 0, 0), (0, 1, 0)\)\} is not a basis of \(\mathbb{R}^3\), because it does not span \(\mathbb{R}^3\).

**Example:** \{\((1, 0), (0, 1), (1, 1)\)\} is not a basis for \(\mathbb{R}^2\).
\[ 1(1, 0) + 1(0, 1) + (-1)(1, 1) = (0, 0) \]

so the set is not linearly independent.

**Theorem 4 (Thm. 1.2')** ¹ Let \(V\) be a Hamel basis for \(X\). Then every vector \(x \in X\) has a unique representation as a linear combination of a finite number of elements of \(V\) (with all coefficients nonzero).²

**Proof:** Let \(x \in X\). Since \(V\) spans \(X\), we can write
\[ x = \sum_{s \in S_1} \alpha_s v_s \]
where \(S_1\) is finite, \(\alpha_s \in F\), \(\alpha_s \neq 0\), and \(v_s \in V\) for each \(s \in S_1\). Now, suppose
\[ x = \sum_{s \in S_1} \alpha_s v_s = \sum_{s \in S_2} \beta_s v_s \]
where \(S_2\) is finite, \(\beta_s \in F\), \(\beta_s \neq 0\), and \(v_s \in V\) for each \(s \in S_2\).

Let \(S = S_1 \cup S_2\), and define
\[ \alpha_s = 0 \quad \text{for} \quad s \in S_2 \setminus S_1 \]
\[ \beta_s = 0 \quad \text{for} \quad s \in S_1 \setminus S_2 \]

Then
\[ 0 = x - x = \sum_{s \in S_1} \alpha_s v_s - \sum_{s \in S_2} \beta_s v_s = \sum_{s \in S} (\alpha_s - \beta_s) v_s \]

¹See Corrections handout.
²The unique representation of 0 is \(0 = \sum_{i \in \emptyset} \alpha_i b_i\).
Since $V$ is linearly independent, we must have $\alpha_s - \beta_s = 0$, so $\alpha_s = \beta_s$, for all $s \in S$.

$$s \in S_1 \iff \alpha_s \neq 0 \iff \beta_s \neq 0 \iff s \in S_2$$

so $S_1 = S_2$ and $\alpha_s = \beta_s$ for $s \in S_1 = S_2$, so the representation is unique. ■

**Theorem 5** Every vector space has a Hamel basis.

**Proof:** The proof uses the Axiom of Choice. Indeed, the theorem is equivalent to the Axiom of Choice. ■

A closely related result, from which you can derive the previous result, shows that any linearly independent set $V$ in a vector space $X$ can be extended to a basis of $X$.

**Theorem 6** If $X$ is a vector space and $V \subseteq X$ is linearly independent, then there exists a linearly independent set $W \subseteq X$ such that

$$V \subseteq W \subseteq \text{span } W = X$$

**Theorem 7** Any two Hamel bases of a vector space $X$ have the same cardinality (are numerically equivalent).

**Proof:** The proof depends on the so-called Exchange Lemma, whose idea we sketch. Suppose that $V = \{v_\lambda : \lambda \in \Lambda\}$ and $W = \{w_\gamma : \gamma \in \Gamma\}$ are Hamel bases of $X$. Remove one vector $v_{\lambda_0}$ from $V$, so that it no longer spans (if it did still span, then $v_{\lambda_0}$ would be a linear combination of other elements of $V$, and $V$ would not be linearly independent). If $w_\gamma \in \text{span } (V \setminus \{v_{\lambda_0}\})$ for every $\gamma \in \Gamma$, then since $W$ spans, $V \setminus \{v_{\lambda_0}\}$ would also span, contradiction. Thus, we can choose $\gamma_0 \in \Gamma$ such that

$$w_{\gamma_0} \notin \text{span } (V \setminus \{v_{\lambda_0}\})$$

Because $w_{\gamma_0} \in \text{span } V$, we can write

$$w_{\gamma_0} = \sum_{i=0}^{n} \alpha_i v_{\lambda_i}$$

where $\alpha_0$, the coefficient of $v_{\lambda_0}$, is not zero (if it were, then we would have $w_{\gamma_0} \in \text{span } (V \setminus \{v_{\lambda_0}\})$). Since $\alpha_0 \neq 0$, we can solve for $v_{\lambda_0}$ as a linear combination of $w_{\gamma_0}$ and $v_{\lambda_1}, \ldots, v_{\lambda_n}$, so

$$\text{span } ((V \setminus \{v_{\lambda_0}\}) \cup \{w_{\gamma_0}\}) \supseteq \text{span } V = X$$

so

$$((V \setminus \{v_{\lambda_0}\}) \cup \{w_{\gamma_0}\})$$
spans \( X \). From the fact that \( w_{\gamma_0} \not\in \text{span} \left( V \setminus \{v_{\lambda_0}\} \right) \) one can show that

\[
\left( \left( V \setminus \{v_{\lambda_0}\} \right) \cup \{w_{\gamma_0}\} \right)
\]

is linearly independent, so it is a basis of \( X \). Repeat this process to exchange every element of \( V \) with an element of \( W \) (when \( V \) is infinite, this is done by a process called transfinite induction). At the end, we obtain a bijection from \( V \) to \( W \), so that \( V \) and \( W \) are numerically equivalent. ■

**Definition 8** The *dimension* of a vector space \( X \), denoted \( \text{dim} \ X \), is the cardinality of any basis of \( X \).

**Definition 9** Let \( X \) be a vector space. If \( \text{dim} \ X = n \) for some \( n \in \mathbb{N} \), then \( X \) is *finite-dimensional*. Otherwise, \( X \) is *infinite-dimensional*.

Recall that for \( V \subseteq X \), \( |V| \) denotes the cardinality of the set \( V \).³

**Example:** The set of all \( m \times n \) real-valued matrices is a vector space over \( \mathbb{R} \). A basis is given by

\[
\{ E_{ij} : 1 \leq i \leq m, 1 \leq j \leq n \}
\]

where

\[
(E_{ij})_{k\ell} = \begin{cases} 1 & \text{if } k = i \text{ and } \ell = j \\ 0 & \text{otherwise} \end{cases}
\]

The dimension of the vector space of \( m \times n \) matrices is \( mn \).

**Theorem 10 (Thm. 1.4)** Suppose \( \text{dim} \ X = n \in \mathbb{N} \). If \( V \subseteq X \) and \( |V| > n \), then \( V \) is linearly dependent.

**Proof:** If not, so \( V \) is linearly independent, then there is a basis \( W \) for \( X \) that contains \( V \). But \( |W| \geq |V| > n = \text{dim} \ X \), a contradiction. ■

**Theorem 11 (Thm. 1.5’)** Suppose \( \text{dim} \ X = n \in \mathbb{N} \) and \( V \subseteq X \), \( |V| = n \).

- If \( V \) is linearly independent, then \( V \) spans \( X \), so \( V \) is a Hamel basis.
- If \( V \) spans \( X \), then \( V \) is linearly independent, so \( V \) is a Hamel basis.

**Proof:** (Sketch)

³See the Appendix to Lecture 2 for some facts about cardinality.
• If $V$ does not span $X$, then there is a basis $W$ for $X$ that contains $V$ as a proper subset. Then $|W| > |V| = n = \dim X$, a contradiction.

• If $V$ is not linearly independent, then there is a proper subset $V'$ of $V$ that is linearly independent and for which $\text{span} V' = \text{span} V = X$. But then $|V'| < |V| = n = \dim X$, a contradiction.

\[ \text{Note: Read the material on Affine Spaces on your own.} \]

\section*{Section 3.2. Linear Transformations}

\textbf{Definition 12} Let $X$ and $Y$ be two vector spaces over the field $F$. We say $T : X \to Y$ is a \emph{linear transformation} if

\[ T(\alpha x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2) \quad \forall x_1, x_2 \in X, \alpha_1, \alpha_2 \in F \]

Let $L(X, Y)$ denote the set of all linear transformations from $X$ to $Y$.

\textbf{Theorem 13} $L(X, Y)$ is a vector space over $F$.

The hard part of proving this theorem is figuring out what you are being asked to prove. Once you figure that out, this is completely trivial, although writing out a complete proof that checks all the vector space axioms is rather tedious. The key is to define scalar multiplication and vector addition, and show that a linear combination of linear transformations is a linear transformation.

\textbf{Proof:} First, define linear combinations in $L(X, Y)$ as follows. For $T_1, T_2 \in L(X, Y)$ and $\alpha, \beta \in F$, define $\alpha T_1 + \beta T_2$ by

\[ (\alpha T_1 + \beta T_2)(x) = \alpha T_1(x) + \beta T_2(x) \]

We need to show that $\alpha T_1 + \beta T_2 \in L(X, Y)$.

\[ (\alpha T_1 + \beta T_2)(\gamma x_1 + \delta x_2) \]
\[ = \alpha T_1(\gamma x_1 + \delta x_2) + \beta T_2(\gamma x_1 + \delta x_2) \]
\[ = \alpha (\gamma T_1(x_1) + \delta T_1(x_2)) + \beta (\gamma T_2(x_1) + \delta T_2(x_2)) \]
\[ = \gamma (\alpha T_1(x_1) + \beta T_2(x_1)) + \delta (\alpha T_1(x_2) + \beta T_2(x_2)) \]
\[ = \gamma (\alpha T_1 + \beta T_2)(x_1) + \delta (\alpha T_1 + \beta T_2)(x_2) \]

so $\alpha T_1 + \beta T_2 \in L(X, Y)$.

The rest of the proof involves straightforward checking of the vector space axioms. \[ \text{\blacksquare} \]
Composition of Linear Transformations

Given \( R \in L(X, Y) \) and \( S \in L(Y, Z) \), \( S \circ R : X \to Z \). We will show that \( S \circ R \in L(X, Z) \), that is, the composition of two linear transformations is also linear.

\[
(S \circ R)(\alpha x_1 + \beta x_2) = S(R(\alpha x_1 + \beta x_2)) = S(\alpha R(x_1) + \beta R(x_2)) = \alpha S(R(x_1)) + \beta S(R(x_2)) = \alpha(S \circ R)(x_1) + \beta(S \circ R)(x_2)
\]

so \( S \circ R \in L(X, Z) \).

Definition 14 Let \( T \in L(X, Y) \).

- The image of \( T \) is \( \text{Im} \, T = T(X) \)
- The kernel of \( T \) is \( \ker \, T = \{ x \in X : T(x) = 0 \} \)
- The rank of \( T \) is \( \text{Rank} \, T = \dim(\text{Im} \, T) \)

Theorem 15 (Thms. 2.9, 2.7, 2.6: The Rank-Nullity Theorem) Let \( X \) be a finite-dimensional vector space and \( T \in L(X, Y) \). Then \( \text{Im} \, T \) and \( \ker \, T \) are vector subspaces of \( Y \) and \( X \) respectively, and

\[
\dim X = \dim \ker T + \text{Rank} \, T
\]

Proof: (Sketch) First show that \( \text{Im} \, T \) is a vector subspace of \( Y \) and \( \ker \, T \) is a vector subspace of \( X \) (exercise).

Then let \( V = \{v_1, \ldots, v_k\} \) be a basis for \( \ker T \) (note that \( \ker T \subseteq X \) so \( \dim \ker T \leq \dim X = n \)). If \( \ker T = \{0\} \), take \( k = 0 \) so \( V = \emptyset \). Extend \( V \) to a basis \( W \) for \( X \) with \( W = \{v_1, \ldots, v_k, w_1, \ldots, w_r\} \). Then \( \{T(w_1), \ldots, T(w_r)\} \) is a basis for \( \text{Im} \, T \) (do this as an exercise).

By definition, \( \dim \ker T = k \) and \( \dim \text{Im} \, T = r \). Since \( W \) is a basis for \( X \), \( k + r = |W| = \dim X \), that is,

\[
\dim X = \dim \ker T + \text{Rank} \, T
\]

\( \blacksquare \)

Theorem 16 (Thm. 2.13) \( T \in L(X, Y) \) is one-to-one if and only if \( \ker T = \{0\} \).

Proof: Suppose \( T \) is one-to-one. Suppose \( x \in \ker T \). Then \( T(x) = 0 \). But since \( T \) is linear, \( T(0) = T(0 \cdot 0) = 0 \cdot T(0) = 0 \). Since \( T \) is one-to-one, \( x = 0 \), so \( \ker T = \{0\} \).
Conversely, suppose that \( \ker T = \{0\} \). Suppose \( T(x_1) = T(x_2) \). Then
\[
T(x_1 - x_2) = T(x_1) - T(x_2) = 0
\]
which says \( x_1 - x_2 \in \ker T \), so \( x_1 - x_2 = 0 \), or \( x_1 = x_2 \). Thus, \( T \) is one-to-one. \( \blacksquare \)

**Definition 17** \( T \in L(X,Y) \) is **invertible** if there is a function \( S : Y \to X \) such that
\[
S(T(x)) = x \quad \forall x \in X \quad \quad T(S(y)) = y \quad \forall y \in Y
\]

In other words \( S \circ T = id_X \) and \( T \circ S = id_Y \), where \( id \) denotes the identity map. In this case denote \( S \) by \( T^{-1} \).

Note that \( T \) is invertible if and only if it is one-to-one and onto. This is just the condition for the existence of an inverse function. The linearity of the inverse follows from the linearity of \( T \).

**Theorem 18 (Thm. 2.11)** If \( T \in L(X,Y) \) is invertible, then \( T^{-1} \in L(Y,X) \), i.e. \( T^{-1} \) is linear.

**Proof:** Suppose \( \alpha, \beta \in F \) and \( v, w \in Y \). Since \( T \) is invertible, there exists unique \( v', w' \in X \) such that
\[
T(v') = v \quad T^{-1}(v) = v' \\
T(w') = w \quad T^{-1}(w) = w'.
\]
Then
\[
T^{-1}(\alpha v + \beta w) = T^{-1}(\alpha T(v') + \beta T(w'))
\]
\[
= T^{-1}(T(\alpha v' + \beta w'))
\]
\[
= \alpha v' + \beta w'
\]
\[
= \alpha T^{-1}(v) + \beta T^{-1}(w)
\]
so \( T^{-1} \in L(Y,X) \). \( \blacksquare \)

**Theorem 19 (Thm. 3.2)** Let \( X,Y \) be two vector spaces over the same field \( F \), and let \( V = \{ v_{\lambda} : \lambda \in \Lambda \} \) be a basis for \( X \). Then a linear transformation \( T \in L(X,Y) \) is completely determined by its values on \( V \), that is:

1. Given any set \( \{ y_{\lambda} : \lambda \in \Lambda \} \subseteq Y \), \( \exists T \in L(X,Y) \) s.t.
\[
T(v_{\lambda}) = y_{\lambda} \quad \forall \lambda \in \Lambda
\]
2. If \( S, T \in L(X,Y) \) and \( S(v_{\lambda}) = T(v_{\lambda}) \) for all \( \lambda \in \Lambda \), then \( S = T \).
Proof:

1. If \( x \in X \), \( x \) has a unique representation of the form

\[
x = \sum_{i=1}^{n} \alpha_i v_{\lambda_i}
\]

with \( \alpha_i \neq 0 \ \forall i = 1, \ldots, n \) (Recall that if \( x = 0 \), then \( n = 0 \).) Define

\[
T(x) = \sum_{i=1}^{n} \alpha_i y_{\lambda_i}
\]

Then \( T(x) \in Y \). The verification that \( T \) is linear is left as an exercise.

2. Suppose \( S(v_{\lambda}) = T(v_{\lambda}) \) for all \( \lambda \in \Lambda \). Given \( x \in X \),

\[
S(x) = S\left( \sum_{i=1}^{n} \alpha_i v_{\lambda_i} \right) = \sum_{i=1}^{n} \alpha_i S(v_{\lambda_i}) = \sum_{i=1}^{n} \alpha_i T(v_{\lambda_i}) = T\left( \sum_{i=1}^{n} \alpha_i v_{\lambda_i} \right) = T(x)
\]

so \( S = T \).

\[ \blacksquare \]

Section 3.3. Isomorphisms

Definition 20 Two vector spaces \( X, Y \) over a field \( F \) are isomorphic if there is an invertible \( T \in L(X, Y) \).

\( T \in L(X, Y) \) is an isomorphism if it is invertible (one-to-one and onto).

Isomorphic vector spaces are essentially indistinguishable as vector spaces.

Theorem 21 (Thm. 3.3) Two vector spaces \( X, Y \) over the same field are isomorphic if and only if \( \dim X = \dim Y \).
**Proof:** Suppose $X, Y$ are isomorphic, and let $T \in L(X, Y)$ be an isomorphism. Let

$$U = \{u_\lambda : \lambda \in \Lambda\}$$

be a basis of $X$, and let

$$v_\lambda = T(u_\lambda), \quad V = \{v_\lambda : \lambda \in \Lambda\}$$

Since $T$ is one-to-one, $U$ and $V$ have the same cardinality. If $y \in Y$, then there exists $x \in X$ such that

$$y = T(x) = T \left( \sum_{i=1}^{n} \alpha_i u_{\lambda_i} \right) = \sum_{i=1}^{n} \alpha_i T(u_{\lambda_i}) = \sum_{i=1}^{n} \alpha_i v_{\lambda_i}$$

which shows that $V$ spans $Y$. To see that $V$ is linearly independent, suppose

$$0 = \sum_{i=1}^{m} \beta_i v_{\lambda_i} = \sum_{i=1}^{m} \beta_i T(u_{\lambda_i}) = T \left( \sum_{i=1}^{m} \beta_i u_{\lambda_i} \right)$$

Since $T$ is one-to-one, $\ker T = \{0\}$, so

$$\sum_{i=1}^{m} \beta_i u_{\lambda_i} = 0$$

Since $U$ is a basis, we have $\beta_1 = \cdots = \beta_m = 0$, so $V$ is linearly independent. Thus, $V$ is a basis of $Y$; since $U$ and $V$ are numerically equivalent, $\dim X = \dim Y$.

Now suppose $\dim X = \dim Y$. Let

$$U = \{u_\lambda : \lambda \in \Lambda\} \text{ and } V = \{v_\lambda : \lambda \in \Lambda\}$$

be bases of $X$ and $Y$; note we can use the same index set $\Lambda$ for both because $\dim X = \dim Y$. By Theorem 3.2, there is a unique $T \in L(X, Y)$ such that $T(u_\lambda) = v_\lambda$ for all $\lambda \in \Lambda$. If $T(x) = 0$, then

$$0 = T(x) = T \left( \sum_{i=1}^{n} \alpha_i u_{\lambda_i} \right)$$
\[
\sum_{i=1}^{n} \alpha_i T(u_{\lambda_i}) = \sum_{i=1}^{n} \alpha_i v_{\lambda_i}
\]

\[\Rightarrow \alpha_1 = \cdots = \alpha_n = 0 \text{ since } V \text{ is a basis}\]

\[\Rightarrow x = 0\]

\[\Rightarrow \ker T = \{0\}\]

\[\Rightarrow T \text{ is one-to-one}\]

If \(y \in Y\), write \(y = \sum_{i=1}^{m} \beta_i v_{\lambda_i}\). Let

\[x = \sum_{i=1}^{m} \beta_i u_{\lambda_i}\]

Then

\[T(x) = T \left( \sum_{i=1}^{m} \beta_i u_{\lambda_i} \right) = \sum_{i=1}^{m} \beta_i T(u_{\lambda_i}) = \sum_{i=1}^{m} \beta_i v_{\lambda_i} = y\]

so \(T\) is onto, hence \(T\) is an isomorphism and \(X, Y\) are isomorphic. ■