Chapter 3, Linear Algebra Section 3.1, Bases

Definition 1 Let $X$ be a vector space over a field $F$. A linear combination of $x_1, \ldots, x_n$ 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 $V \subseteq X$, span$V$ denotes the set of all linear combinations of $V$.

A set $V \subseteq X$ is linearly dependent if there exist $v_1, \ldots, v_n \in X$ 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.

A set $V \subseteq X$ spans $X$ if span$V = X$.

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 $\mathbb{R}^2$.

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

$$\begin{align*}
(x, y) &= \alpha(1,1) + \beta(-1,1) \\
x &= \alpha - \beta \\
y &= \alpha + \beta \\
x + y &= 2\alpha \\
\alpha &= \frac{x + y}{2}
\end{align*}$$
\[ y - x = 2\beta \]
\[ \beta = \frac{y - x}{2} \]
\[ (x, y) = \frac{x + y}{2}(1, 1) + \frac{y - x}{2}(-1, 1) \]

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.

**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 2 (1.2’, see Corrections handout)** Let \(V\) be a Hamel basis for \(X\). Then every vector \(x \in X\) has a unique representation as a linear combination (with all coefficients nonzero) of a finite number of elements of \(V\).

(Aside: the unique representation of 0 is \(0 = \sum_{i \in \emptyset} \alpha_i b_i\).)

**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, \(\alpha_s \in F, \alpha_s \neq 0, v_s \in V\) for \(s \in S_1\). 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 $s \in S_2$.

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

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

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 3** 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. ■

**Theorem 4** Any two Hamel bases of a vector space $X$ 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 $V$, \[ 0 = x - x = \sum_{s \in S_1} \alpha_s v_s - \sum_{s \in S_2} \beta_s v_s = \sum_{s \in S} \alpha_s v_s - \sum_{s \in S} \beta_s v_s = \sum_{s \in S} (\alpha_s - \beta_s) v_s \]

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 3** 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. ■

**Theorem 4** Any two Hamel bases of a vector space $X$ 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 $V$, \[ 0 = x - x = \sum_{s \in S_1} \alpha_s v_s - \sum_{s \in S_2} \beta_s v_s = \sum_{s \in S} \alpha_s v_s - \sum_{s \in S} \beta_s v_s = \sum_{s \in S} (\alpha_s - \beta_s) v_s \]

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 3** 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. ■

**Theorem 4** Any two Hamel bases of a vector space $X$ 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 $V$, \[ 0 = x - x = \sum_{s \in S_1} \alpha_s v_s - \sum_{s \in S_2} \beta_s v_s = \sum_{s \in S} \alpha_s v_s - \sum_{s \in S} \beta_s v_s = \sum_{s \in S} (\alpha_s - \beta_s) v_s \]

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.
and $V$ would not be linearly independent). If $w_{\gamma} \in \text{span} \left( V \setminus \{v_{\lambda_0}\} \right)$ 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

$$w_{\gamma_0} \not\in \text{span} \left( V \setminus \{v_{\lambda_0}\} \right)$$

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 $w_{\gamma_0} \in \text{span} \left( V \setminus \{v_{\lambda_0}\} \right)$). 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

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

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

$$((V \setminus \{v_{\lambda_0}\}) \cup \{w_{\gamma_0}\})$$

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 5** Let $\text{dim} \ X$ (read “the dimension of $X$”) denote the cardinal number of any basis of $X$. 


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

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

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

**Theorem 6 (1.4)** Suppose $\dim X = n \in \mathbb{N}$. If $V \subseteq X$ and $|V| > n$ (recall $|V|$ denotes the number of elements in the set $V$), then $V$ is linearly dependent.

**Theorem 7 (1.5')** Suppose $\dim X = n \in \mathbb{N}$, $V \subseteq X$, $|V| = n$.

- 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.

Read the material on Affine Spaces on your own.

**Section 3.2, Linear Transformations**

**Definition 8** Let $X, Y$ be two vector spaces over the field $F$. We say $T : X \rightarrow Y$ is a linear transformation if

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

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

**Theorem 9** $L(X, Y)$ is a vector space over $F$. 
Proof: The hard part 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.

We define

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

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 is too tedious to reproduce here.

Composition of Linear Transformations

Given \(R \in L(X,Y)\) and \(S \in L(Y,Z)\), \(S \circ R : X \rightarrow Z\). We will show that \(S \circ R \in L(X,Z)\).

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

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

Definition 10

\[\text{Im } T = T(X) \text{ (image of } T)\]
\[ \ker T = \{ x : T(x) = 0 \} \text{ (kernel of } T) \]
\[ \text{Rank } T = \dim(\text{Im } T) \]

**Theorem 11 (2.9, 2.7, 2.6)** Let \( X \) be a finite-dimensional vector space, \( T \in L(X, Y) \). Then \( \text{Im } T \) and \( \ker T \) are vector subspaces of \( Y \) and \( X \) respectively, and

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

**Theorem 12 (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

\[ T(x_1 - x_2) = T(x_1) - T(x_2) = 0 \]

so \( x_1 - x_2 \in \ker T \), so \( x_1 - x_2 = 0, x_1 = x_2 \). Thus, \( T \) is one-to-one.\[ \blacksquare \]

**Definition 13** \( T \in L(X, Y) \) is invertible if there is a function \( S : Y \to X \) such that

\[ \forall x \in X S(T(x)) = x \]
\[ \forall y \in Y T(S(y)) = y \]

In other words \( S \circ T = \text{id}_X \) and \( T \circ S = \text{id}_Y \), where \( \text{id} \) denotes the identity map. 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 14 (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,

$$
\exists! v', w' \in X \quad \begin{cases} T(v') = v & T^{-1}(v) = v' \\ T(w') = w & T^{-1}(w) = w' \end{cases}
$$

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 $T^{-1} \in L(Y, X)$.

Although the next theorem is in Section 3.3, it really belongs here:

Theorem 15 (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 $V$, i.e.

1. Given any set of values $\{y_\lambda : \lambda \in \Lambda\} \subseteq Y$,

$$
\exists T \in L(X, Y) \forall \lambda \in \Lambda \; T(v_\lambda) = y_\lambda
$$

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

$$
x = \sum_{i=1}^{n} \alpha_i v_{\lambda_i} \quad \alpha_i \neq 0 (i = 1, \ldots, n)
$$
(Aside: for \( x = 0 \), we have \( n = 0 \).) 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 \( S(\lambda) = T(\lambda) \) for all \( \lambda \in \Lambda \). Given \( x \in X \),

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

so \( S = T \).

Section 3.3, Isomorphisms

**Definition 16** Two vector spaces \( X, Y \) over a field \( F \) are *isomorphic* if there is an invertible (recall this means one-to-one and onto) \( T \in L(X, Y) \). \( T \) is called an *isomorphism*.

Isomorphic vector spaces are essentially indistinguishable as vector spaces.

**Theorem 17 (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, via the isomorphism $T$. Let

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

be a basis of $X$, and let

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

Since $T$ is one-to-one, $U$ and $V$ are numerically equivalent. If $y \in Y$, then there exists $x \in X$ such that

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

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

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

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

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

\( \Rightarrow \) \( \alpha_1 = \cdots = \alpha_n = 0 \) since \( V \) is a basis

\( \Rightarrow \) \( x = 0 \)

\( \Rightarrow \) \( \ker T = \{0\} \)

\( \Rightarrow \) \( T \) 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, so \( T \) is an isomorphism and \( X, Y \) are isomorphic. \( \blacksquare \)