MATH 146 Linear Algebra 1, Lecture Notes by Stephen New

Chapter 1. Concrete Vector Spaces and Affine Spaces

Rings and Fields

1.1 Definition: For a set S we write $S \times S = \{(a, b) | a \in S, b \in S\}$. A **binary operation** on S is a map $* : S \times S \to S$, where for $a, b \in S$ we usually write $*(a, b)$ as $a * b$.

1.2 Definition: A **ring** (with identity) is a set R together with two binary operations $+$ and $\cdot$ (called addition and multiplication), where for $a, b \in R$ we usually write $a \cdot b$ as ab, and two distinct elements 0 and 1, such that

(1) $+$ is associative: $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$,
(2) $+$ is commutative: $a + b = b + a$ for all $a, b \in R$,
(3) 0 is an additive identity: $0 + a = a$ for all $a \in R$,
(4) every element has an additive inverse: for every $a \in R$ there exists $b \in R$ with $a + b = 0$,
(5) $\cdot$ is associative: $(ab)c = a(bc)$ for all $a, b, c \in R$,
(6) 1 is a multiplicative identity: $1 \cdot a = a$ for all $a \in R$, and
(7) $\cdot$ is distributive over $+$: $a(b + c) = ab + ac$ for all $a, b, c \in R$,

A ring R is called **commutative** when
(8) $\cdot$ is commutative: $ab = ba$ for all $a, b \in R$.

For $a \in R$, we say that a is **invertible** (or that a has an **inverse**) when there exists an element $b \in R$ such that $ab = 1 = ba$. A **field** is a commutative ring F such that
(9) every non-zero element has a multiplicative inverse: for every $a \in F$ with $a \neq 0$ there exists $b \in F$ such that $ab = 1$.

An element in a field F is called a **number** or a **scalar**.

1.3 Example: The set of **integers Z** is a commutative ring, but it is not a field because it does not satisfy Property (9). The set of **positive integers** $\mathbf{Z}^+ = \{1, 2, 3, \cdots\}$ is not a ring because $0 \notin \mathbf{Z}^+$ and $\mathbf{Z}^+$ does not satisfy Properties (3) and (4). The set of **natural numbers** $\mathbf{N} = \{0, 1, 2, \cdots\}$ is not a ring because it does not satisfy Property (4). The set of **rational numbers Q**, the set of **real numbers R** and the set of **complex numbers C** are all fields. For $2 \leq n \in \mathbf{Z}$, the set $\mathbf{Z}_n = \{0, 1, \cdots, n - 1\}$ of **integers modulo** n is a commutative ring, and $\mathbf{Z}_n$ is a field if and only if n is prime (in $\mathbf{Z}_1 = \{0\}$ we have $0 = 1$, so $\mathbf{Z}_1$ is not a ring).

1.4 Remark: In a field, we can perform all of the usual arithmetical operations. The next few theorems illustrate this.

1.5 Theorem: *(Uniqueness of Inverse) Let R be a field. Let $a \in R$. Then*

(1) the additive inverse of a is unique: if $a + b = 0 = a + c$ then $b = c$,
(2) if a has an inverse then it is unique: if $ab = 1 = ac$ then $b = c$.

Proof: To prove (1), suppose that $a + b = 0 = a + c$. Then

$$b = 0 + b = (a + c) + b = b + (a + c) = (b + a) + c = (a + b) + c = 0 + c = c\,.$$

To prove (2), suppose that $a \neq 0$ and that $ab = 1 = ac$. Then

$$b = 1 \cdot b = (ac)b = b(ac) = (ba)c = (ab)c = 1 \cdot c = c\,.$$

1.6 Definition: Let R be a ring and let $a, b \in R$. We write the (unique) additive inverse of a as $-a$, and we write $b - a = b + (-a)$. If a has a multiplicative inverse, we write the (unique) multiplicative inverse of a as a^{-1}. When R is commutative, we also write a^{-1} as $\frac{1}{a}$, and we write $\frac{b}{a} = b \cdot \frac{1}{a}$.

1.7 Theorem: *(Cancellation) Let R be a field. Then for all $a, b, c \in R$, we have*

(1) if $a + b = a + c$ then $b = c$,
(2) if $a + b = a$ then $b = 0$, and
(3) if $a + b = 0$ then $b = -a$.

Let F be a field. Then for all $a, b, c \in F$ we have

(4) if $ab = ac$ then either $a = 0$ or $b = c$.
(5) if $ab = a$ then either $a = 0$ or $b = 1$,
(6) if $ab = 1$ then $b = a^{-1}$, and
(7) if $ab = 0$ then either $a = 0$ or $b = 0$.

Proof: To prove (1), suppose that $a + b = a + c$. Then we have

$$b = 0 + b = -a + a + b = -a + a + c = 0 + c = c\,.$$

Part (2) follows from part (1) since if $a + b = a$ then $a + b = a + 0$, and part (3) follows from part (1) since if $a + b = 0$ then $a + b = a + (-a)$. To prove part (4), suppose that $ab = ac$ and $a \neq 0$. Then we have

$$b = 1 \cdot b = a^{-1}ab = a^{-1}ac = 1 \cdot c = c\,.$$

Note that parts (5), (6) and (7) all follow from part (4).

1.8 Remark: In the above proof, we used associativity and commutativity implicitly. If we wished to be explicit then the proof of part (1) would be as follows. Suppose that $a + b = a + c$. Then we have

$$b = 0+b = (a-a)+b = (-a+a)+b = -a+(a+b) = -a+(a+c) = (-a+a)+c = 0+c = c.$$

In the future, we shall often use associativity and commutativity implicitly in our calculations.

1.9 Theorem: *(Multiplication by 0 and -1) Let R be a ring and let $a \in R$. Then*

(1) $0 \cdot a = 0$, and
(2) $(-1)a = -a$.

Proof: We have

$$0a = (0 + 0)a = 0a + 0a.$$

Subtracting $0a$ from both sides (using part 2 of the Cancellation Theorem) gives $0 = 0a$. Also, we have

$$a + (-1)a = (1)a + (-1)a = (1 + (-1))a = 0a = 0,$$

and subtracting a from both sides (part 3 of the Cancellation Theorem) gives $(-1)a = -a$.

The Standard Vector Space

1.10 Definition: Let S be a set. An **n-tuple** on S is a function $a : \{1, 2, \cdots, n\} \to S$. Given an n-tuple a on S, for $k \in \{1, 2, \cdots, n\}$ we write $a_k = a(k)$. The set $\{1, 2, \cdots, n\}$ is called the **index set**, an element $k \in \{1, 2, \cdots, n\}$ is called an **index**, and the element $a_k \in S$ is called the k^{th} **entry** of a. We sometimes write $a = (a_1, a_2, \cdots, a_n)$ but we more often write

$$a = (a_1, a_2, \cdots, a_n)^T = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

to indicate that a is the n-tuple with entries $a_1, a_2, \cdots, a_n$. The set of all n-tuples on S is denoted by S^n, so we have

$$S^n = \Big\{ a = (a_1, a_2, \cdots, a_n)^T \Big| \text{ each } a_i \in S \Big\}.$$

1.11 Definition: For a ring R, we define the **zero element** $0 \in R^n$ to be

$$0 = (0, 0, 0, \cdots, 0)^T$$

or equivalently we define $0 \in R^n$ to be the element with entries $0_i = 0$ for all i. We define the **standard basis elements** $e_1, e_2, \cdots, e_n \in R^n$ to be

$$\begin{aligned} e_1 &= (1, 0, 0, 0 \cdots, 0)^T, \\ e_2 &= (0, 1, 0, 0, \cdots, 0^T, \\ e_3 &= (0, 0, 1, 0, \cdots, 0)^T, \\ &\vdots \\ e_n &= (0, 0, 0, \cdots, 0, 1)^T. \end{aligned}$$

Equivalently, for each index k we define $e_k \in R^n$ to be given by $(e_k)_i = \delta_{ki} = \begin{cases} 1 \text{ if } k = i, \\ 0 \text{ if } k \neq i. \end{cases}$

1.12 Definition: Given $t \in R$, $x = (x_1, x_2, \cdots, x_n)^T \in R^n$ and $y = (y_1, y_2, \cdots, y_n)^T \in R^n$, where R is a ring, we define the product tx and the sum $x + y$ by

$$tx = t \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} t\,x_1 \\ t\,x_2 \\ \vdots \\ t\,x_n \end{pmatrix},$$

$$x + y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Equivalently, we can define tx to be the element with entries $(tx)_i = t\,x_i$ for all i, and we can define $x + y$ to be the element with entries $(x + y)_i = x_i + y_i$ for all i.

1.13 Note: For $x_1, x_2, \cdots, x_n \in R$, notice that

$$(x_1, x_2, \cdots, x_n)^T = \sum_{i=1}^{n} x_i e_i = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n.$$

1.14 Theorem: *(Basic Properties of R^n) Let R be a ring. Then*

(1) $+$ is associative: $(x + y) + z = x + (y + z)$ for all $x, y, z \in R^n$,
(2) $+$ is commutative: $x + y = y + x$ for all $x, y \in R^n$,
(3) $0 \in R^n$ is an additive identity $0 + x = x$ for all $x \in R^n$,
(4) every $x \in R^n$ has an additive inverse: for all $x \in R^n$ there exists $y \in R^n$ with $x + y = 0$,
(5) $\cdot$ is associative: $(st)x = s(tx)$ for all $s, t \in R$ and all $x \in R^n$,
(7) $\cdot$ distributes over addition in R: $(s + t)x = sx + tx$, for all $s, t \in R$ and $x \in R^n$,
(8) $\cdot$ distributes over addition in R^n: $t(x + y) = tx + ty$ for all $t \in R$ and $x, y \in R^n$, and
(9) $1 \in R$ acts a multiplicative identity: $1x = x$ for all $x \in R^n$.

Proof: To prove part (4), let $x \in R^n$ and choose $y = (-1)x$. Then for all indices i we have $y_i = ((-1)x)_i = -x_i$ and so $(x + y)_i = x_i - x_i = 0$. Since $(x + y)_i = 0$ for all i, we have $x + y = 0$, as required. To prove part (8), let $t \in R$ and let $x, y \in R^n$. Then for all i we have $\big(t(x + y)\big)_i = t(x + y)_i = t(x_i + y_i) = t\,x_i + t\,y_i = (tx + ty)_i$. Since $\big(t(x + y)\big)_i = (tx + ty)_i$ for all i, we have $t(x + y) = tx + ty$. The other parts can be proven similarly.

1.15 Definition: When F is a field, F^n is called the **standard n-dimensional vector space** over F, and an element of F^n is called a **point** or a **vector**.

1.16 Example: Let $n = 2$ or 3 and let $u, v \in \mathbf{R}^n$. If $u \neq 0$ then the set $\{tu | t \in \mathbf{R}\}$ is the line in $\mathbf{R}^n$ through the points 0 and u, and the set $\{tu | 0 \leq t \leq 1\}$ is the line segment in $\mathbf{R}^2$ between 0 and u. If $u \neq 0$ and v does not lie on the line through 0 and u, then the points 0, u, v and $u + v$ are the veritices of a parallelogram P in $\mathbf{R}^n$, the set $\{su + tv | s \in \mathbf{R}, t \in \mathbf{R}\}$ is the plane which contains P (in that case that $n = 2$, this plane is the entire set $\mathbf{R}^2$), the set $\{su + tv | 0 \leq s \leq 1, 0 \leq t \leq 1\}$ is the set of points inside (and on the edges of) P. As an exercise, describe the sets $\{su + tv | s + t = 1\}$ and $\{su + tv | s \geq 0, t \geq 0, s + t \leq 1\}$.

Vector Spaces and Affine Spaces in F^n

1.17 Definition: Let F be a field. Given a point $p \in F^n$ and a non-zero vector $u \in F^n$, we define the **line** in F^n through p in the direction of u to be the set

$$L = \big\{p + tu \big| t \in F\big\}.$$

Given a point $p \in \mathbf{F}^n$ and two vectors $u, v \in \mathbf{R}^n$ with $u \neq 0$ and $v \neq tu$ for any $t \in F$, we define the **plane** in F^n through p in the direction of u and v to be the set

$$P = \big\{p + su + tv \,\big|\, s, t \in F\big\}.$$

1.18 Remark: We wish to generalize the above definitions by defining higher dimensional versions of lines and planes.

1.19 Note: For a finite set S, the **cardinality** of S, denoted by $|S|$, is the number of elements in S. When we write $S = \{a_1, a_2, \cdots, a_m\}$, we shall always tacitly assume that the elements $a_i \in S$ are all distinct so that $|S| = m$ unless we explicitly indicate otherwise.

1.20 Definition: Let R be a ring, let $\mathcal{A} = \{u_1, u_2, \cdots, u_m\} \subseteq R^n$. A **linear combination** on $\mathcal{A}$ (over R) is an element $x \in R^n$ of the form

$$x = \sum_{i=1}^{m} t_i u_i = t_1 u_1 + t_2 u_2 + \cdots + t_m u_m \quad \text{with each } t_i \in R.$$

The **span** of $\mathcal{A}$ (over R) (also called the **submodule** of R^n **spanned** by $\mathcal{A}$ over R), is the set of all linear combinations on $\mathcal{A}$. We denote the span of $\mathcal{A}$ by $\operatorname{Span}\mathcal{A}$ (or by $\operatorname{Span}_R\mathcal{A}$) so we have

$$\operatorname{Span}\mathcal{A} = \operatorname{Span}_R\mathcal{A} = \Big\{ \sum_{i=1}^{m} t_i u_i \Big| \text{ each } t_i \in R\Big\}.$$

For convenience, we also define $\operatorname{Span}\emptyset = \{0\}$, where $\emptyset$ is the empty set. Given an element $p \in R^n$, we write

$$p + \operatorname{Span}\mathcal{A} = \Big\{p + u \Big| u \in \operatorname{Span}\mathcal{A}\Big\} = \Big\{p + \sum_{i=1}^{m} t_i u_i \Big| \text{each } t_i \in R\Big\}.$$

1.21 Definition: Let F be a field. For a finite set $\mathcal{A} \subseteq F^n$ the set $U = \operatorname{Span}\mathcal{A}$ is called the **vector space** in F^n (or the **subspace** of F^n) spanned by $\mathcal{A}$ (over F). A **vector space** in F^n (or a **subspace** of F^n) is a subset $U \subseteq F^n$ of the form $U = \operatorname{Span}\mathcal{A}$ for some finite subset $\mathcal{A} \subseteq F^n$. Given a point $p \in F^n$ and a finite set $\mathcal{A} \subseteq F^n$, the set $P = p + \operatorname{Span}\mathcal{A}$ is called the **affine space** in F^n (or the **affine subspace** of F^n) through p in the direction of the vectors in $\mathcal{A}$. An **affine space** in F^n (or an **affine subspace** of F^n) is a subset $P \subseteq F^n$ of the form $P = p + U$ for some point $p \in F^n$ and some vector space U in F^n. An element in a subspace of F^n can be called a **point** or a **vector**. An element in an affine subspace of F^n is usually called a **point**.

1.22 Theorem: *(Closure under Addition and Multiplication) Let R be a ring, let $\mathcal{A}$ be a finite subset of R, and let $U = \operatorname{Span}\mathcal{A}$. Then*

(1) U is closed under addition: for all $x, y \in U$ we have $x + y \in U$, and
(2) U is closed under multiplication: for all $t \in R$ and all $x \in U$ we have $tx \in U$.

Proof: Let $\mathcal{A} = \{u_1, u_2, \cdots, u_m\}$, let $t \in R$ and let $x, y \in U = \operatorname{Span}\mathcal{A}$, say $x = \sum_{i=1}^{m} s_i u_i$ and $y = \sum_{i=1}^{m} t_i u_i$. Then $x + y = \sum_{i=1}^{m} (s_i + t_i) u_i \in U$ and $tx = \sum_{i=1}^{m} (t s_i) u_i \in U$.

1.23 Theorem: *Let R be a ring, let $p, q \in R^n$, let $\mathcal{A}$ and $\mathcal{B}$ be finite subsets of R^n, and let $U = \operatorname{Span}\mathcal{A}$ and $V = \operatorname{Span}\mathcal{B}$. Then*

(1) $p + U \subseteq q + V$ if and only if $U \subseteq V$ and $p - q \in V$, and
(2) $p + U = q + V$ if and only if $U = V$ and $p - q \in U$.

Proof: Suppose that $p + U \subseteq q + V$. Since $p = p + 0 \in p + U$, we also have $p \in q + V$, say $p = q + v$ where $v \in V$. Then $p - q = v \in V$. Let $u \in U$. Then we have $p + u \in p + U$ and so $p + u \in q + V$, say $p + u = q + w$ where $w \in V$. Then $u = w - (p - q) = w - v = w + (-1)v \in V$ by closure. Conversely, suppose that $U \subseteq V$ and $p - q \in V$, say $p - q = v \in V$. Let $a \in p + U$, say $a = p + u$ where $u \in U$. Then we have $a = p + u = (q + v) + u = q + (u + v) \in q + V$ by closure, since $u, v \in V$. This proves Part (1), from which Part (2) immediately follows.

1.24 Theorem: *Let $\mathcal{A} = \{u_1, u_2, \cdots, u_l\} \subseteq R^n$ and let $\mathcal{B} = \{v_1, v_2, \cdots, v_m\} \subseteq R^n$, where R is a ring. Then*

(1) $\operatorname{Span}\mathcal{A} \subseteq \operatorname{Span}\mathcal{B}$ *if and only if each* $u_j \in \operatorname{Span}\mathcal{B}$, *and*
(2) $\operatorname{Span}\mathcal{A} = \operatorname{Span}\mathcal{B}$ *if and only if each* $u_j \in \operatorname{Span}\mathcal{B}$ *and each* $v_j \in \operatorname{Span}\mathcal{A}$.

Proof: Note that each $u_j \in \operatorname{Span}\mathcal{A}$ because we can write u_j as a linear combination on $\mathcal{A}$, indeed we have

$$u_j = 0u_1 + 0u_2 + \cdots + 0u_{j-1} + 1u_j + 0u_{j+1} + \cdots + 0u_l = \sum_{i=1}^{l} t_i u_i \text{ with } t_i = \delta_{ij}.$$

It follows that if $\operatorname{Span}\mathcal{A} \subseteq \operatorname{Span}\mathcal{B}$ then we have each $u_j \in \operatorname{Span}\mathcal{B}$. Suppose, conversely, that each $u_j \in \operatorname{Span}\mathcal{B}$, say $u_j = \sum_{i=1}^{m} s_{ji} v_i$. Let $x \in \operatorname{Span}\mathcal{A}$, say $x = \sum_{j=1}^{l} t_j u_j$. Then

$$x = \sum_{j=1}^{l} t_j u_j = \sum_{j=1}^{l} t_j \sum_{i=1}^{m} s_{ji} v_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{l} t_j s_{ji} \Big) v_i \in \operatorname{Span}\mathcal{B}.$$

This Proves part (1), and Part (2) follows immediately from part (1).

Linear Independence, Bases and Dimension

1.25 Definition: Let R be a ring. For $\mathcal{A} = \{u_1, u_2, \cdots, u_m\} \subseteq R^n$, we say that $\mathcal{A}$ is **linearly independent** (over R) when for all $t_1, t_2, \cdots, t_m \in R$, if $\sum_{i=1}^{m} t_i u_i = 0$ then each $t_i = 0$, and otherwise we say that $\mathcal{A}$ is **linearly dependent**. For convenience, we also say that the empty set $\emptyset$ is linearly independent. For a finite set $\mathcal{A} \subseteq F^n$, when $\mathcal{A}$ is linearly independent and $U = \operatorname{Span} \mathcal{A}$, we say that $\mathcal{A}$ is a **basis** for U.

1.26 Example: Let F be a field. The empty set $\emptyset$ is linearly independent and $\operatorname{Span} \emptyset = \{0\}$ and so $\emptyset$ is a basis for the vector space $\{0\}$ in F^n. If $0 \neq u \in F^n$ then $\{u\}$ is linearly independent and so $\{u\}$ is a basis for $\operatorname{Span}\{u\}$. As an exercise, verify that for $u, v \in F^n$, the set $\{u, v\}$ is linearly independent if and only if $u \neq 0$ and for all $t \in F$ we have $v \neq tu$.

1.27 Example: Verify that the set $\{e_1, e_2, \cdots, e_n\}$ is a basis for F^n. We call it the **standard basis** for F^n.

1.28 Theorem: *Let F be a field and let $\mathcal{A} = \{u_1, u_2, \cdots, u_m\} \subseteq F^n$. Then*
(1) for $1 \leq k \leq m$, we have $u_k \in \operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big)$ if and only if $\operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big) = \operatorname{Span} \mathcal{A}$,
(2) $\mathcal{A}$ is linearly dependent if and only if $u_k \in \operatorname{Span}\big(\mathcal{A} \setminus \{u_m\}\big)$ for some index k.

Proof: Note that if $\operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big) = \operatorname{Span} \mathcal{A}$ then $u_k \in \operatorname{Span} \mathcal{A} = \operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big)$. Suppose, conversely, that $u_k \in \operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big)$, say $u_k = \sum_{i \neq k} s_i u_i$ where each $s_i \in F$. Since $\mathcal{A} \setminus \{u_k\} \subseteq \mathcal{A}$ it is clear that $\operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big) \subseteq \operatorname{Span} \mathcal{A}$. Let $x \in \operatorname{Span} \mathcal{A}$, say $x = \sum_{i=1}^{m} t_i u_i$. Then we have $x = t_k u_k + \sum_{i \neq k} t_i u_i = t_k \sum_{i \neq k} s_i u_i + \sum_{i \neq k} t_i u_i \in \operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big)$. This proves Part (1).

Note that since $\mathcal{A} \setminus \{u_k\} \subseteq \mathcal{A}$, we have $\operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big) \subseteq \operatorname{Span} \mathcal{A}$. Suppose that $\mathcal{A}$ is linearly dependent. Choose coefficients $s_i \in F$, not all equal to zero, so that $\sum_{i=1}^{m} s_i u_i = 0$. Choose an index k so that $s_k \neq 0$. Since $0 = s_k u_k + \sum_{i \neq k} s_i u_i$ we have $u_k = -\sum_{i \neq k} \frac{s_i}{s_k} u_i$. For $x = \sum_{i=1}^{m} t_i u_i \in \operatorname{Span} \mathcal{A}$ we have $x = t_k u_k + \sum_{i \neq k} t_i u_i = -t_k \sum_{i \neq k} \frac{s_i}{s_k} u_i + \sum_{i \neq k} t_i u_i \in \operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big)$. This shows that if $\mathcal{A}$ is linearly dependent then $\operatorname{Span}\big(\mathcal{A} \setminus \{u_k\}\big) = \operatorname{Span} \mathcal{A}$.

1.29 Theorem: *Let F be a field, let $\mathcal{A} = \{u_1, u_2, \cdots, u_m\} \subseteq F^n$, and let $U = \operatorname{Span} \mathcal{A}$. Then $\mathcal{A}$ contains a basis for U.*

Proof: If $\mathcal{A}$ is linearly independent, then $\mathcal{A}$ is a basis for U. Suppose that $\mathcal{A}$ is linearly dependent. Then for some index k we have $u_k \in \operatorname{Span}(\mathcal{A} \setminus \{u_k\})$. Reordering the vectors if necessary, let us assume that $u_m \in \operatorname{Span}(\mathcal{A} \setminus \{u_m\})$. Then we have $\operatorname{Span}\{u_1, u_2, \cdots, u_{m-1}\} = \operatorname{Span}\{u_1, u_2, \cdots, u_m\} = U$. If $\{u_1, u_2, \cdots, u_{m-1}\}$ is linearly independent then it is a basis for U. Otherwise, as above, we can reorder $u_1, u_2, \cdots, u_{m-1}$ if necessary so that $\operatorname{Span}\{u_1, u_2, \cdots, u_{m-2}\} = \operatorname{Span}\{u_1, u_2, \cdots, u_{m-1}\} = U$. Repeating this procedure we will eventually obtain a linearly independent subset $\{u_1, u_2, \cdots, u_k\} \subseteq \mathcal{A}$ with $\operatorname{Span}\{u_1, u_2, \cdots, u_k\} = U$ (if the procedure continues until no vectors are left then we have $k = 0$ and $\{u_1, \cdots, u_k\} = \emptyset$, which is linearly independent).

1.30 Corollary: *For a field F, every subspace of F^n has a basis.*

1.31 Theorem: *Let F be a field, let $\mathcal{A} = \{u_1, u_2, \cdots, u_m\} \subseteq F^n$, let $a_1, a_2, \cdots, a_m \in F$ with $a_k \neq 0$, and let $\mathcal{B} = \{v_1, v_2, \cdots, v_m\}$ where $v_i = u_i$ for $i \neq k$ and $v_k = \sum_{i=1}^{m} a_i u_i$. Then*
(1) $\operatorname{Span}\mathcal{A} = \operatorname{Span}\mathcal{B}$ *and*
(2) $\mathcal{A}$ is linearly independent if and only if $\mathcal{B}$ is linearly independent.

Proof: For $x = \sum_{i=1}^{m} t_i v_i \in \operatorname{Span}\mathcal{B}$, we have $x = t_k v_k + \sum_{i\neq k} t_i v_i = t_k \sum_{i=1}^{m} a_i u_i + \sum_{i\neq k} t_i u_i$ and so $x \in \operatorname{Span}\mathcal{A}$. This shows that $\operatorname{Span}\mathcal{B} \subseteq \operatorname{Span}\mathcal{A}$.

Suppose $\mathcal{A}$ is linearly independent. Suppose $\sum_{i=1}^{m} t_i v_i = 0$ where each $t_i \in F$. Then

$$0 = t_k v_k + \sum_{i\neq k} t_i v_i = t_k \sum_{i=1}^{m} a_i u_i + \sum_{i\neq k} t_i u_i = t_k a_k u_k + \sum_{i\neq k} (t_k a_i + t_i) u_i.$$

Since $\mathcal{A}$ is linearly independent, all of the coefficients in the above linear combination on $\mathcal{A}$ must be equal to zero, so we have $t_k a_k = 0$ and $t_k a_i + t_i = 0$ for $i \neq k$. Since $t_k a_k = 0$ and $a_k \neq 0$ we have $t_k = 0$ and hence $0 = t_k a_i + t_i = t_i$ for all $i \neq k$. This shows that $\mathcal{B}$ is linearly independent.

Finally note that since $v_k = \sum_{i=1}^{m} a_i u_i = a_k u_k + \sum_{i\neq k} a_i v_i$ with $a_k \neq 0$, it follows that $u_k = \sum_{i=1}^{m} b_i v_i$ where $b_k = \frac{1}{a_k} \neq 0$ and $b_i = -\frac{a_i}{a_k}$ for $i \neq k$. Hence the same arguments used in the previous two paragraphs, with the rôles of $\mathcal{A}$ and $\mathcal{B}$ interchanged, show that $\operatorname{Span}\mathcal{A} \subseteq \operatorname{Span}\mathcal{B}$ and that if $\mathcal{B}$ is linearly independent then so is $\mathcal{A}$.

1.32 Theorem: *Let F be a field, let U be a subspace of F^n and let $\mathcal{A} = \{u_1, u_2, \cdots, u_m\}$ be a basis for U. Let $\mathcal{B} = \{v_1, v_2, \cdots, v_l\} \subseteq U$. Suppose that $\mathcal{B}$ is linearly independent. Then $l \leq m$, if $l = m$ then $\mathcal{B}$ is a basis for U, and if $l < m$ then there exist $m - l$ vectors in $\mathcal{A}$ which, after possibly reordering the vectors u_i we can take to be $u_{l+1}, u_{l+2}, \cdots, u_m$, such that the set $\{v_1, v_2, \cdots, v_l, u_{l+1}, u_{l+2}, \cdots, u_m\}$ is a basis for U.*

Proof: When $l = 0$ so that $\mathcal{B} = \emptyset$, we have $m - l = m$ and we use all m of the vectors in $\mathcal{A}$ to obtain the basis $\{u_1, u_2, \cdots, u_m\}$. Let $l \geq 1$ and suppose, inductively, that for every set $\mathcal{B}_0 = \{v_1, v_2, \cdots, v_{l-1}\} \subseteq U$, if $\mathcal{B}_0$ is linearly independent then we have $l - 1 \leq m$ and we can reorder the vectors u_i so that $\{v_1, v_2, \cdots, v_{l-1}, u_l, u_{l+1}, \cdots, u_m\}$ is a basis for U. Let $\mathcal{B} = \{v_1, v_2, \cdots, v_l\} \subseteq U$. Suppose $\mathcal{B}$ is linearly independent. Let $\mathcal{B}_0 = \{v_1, v_2, \cdots, v_{l-1}\}$. Note that $\mathcal{B}_0$ is linearly independent but $\mathcal{B}_0$ idoes not span U because $v_l \in U$ but $v_l \notin \operatorname{Span}\mathcal{B}_0$. By the induction hypothesis, we have $l - 1 < m$ and we can reorder the vectors u_i so that $\{v_1, v_2, \cdots, v_{l-1}, u_l, u_{l+1}, \cdots, u_m\}$ is a basis for U. Since $v_l \in U$ we can write v_l in the form $v_l = \sum_{i=1}^{l-1} t_i v_i + \sum_{i=l}^{m} s_j u_j$. Note that the coefficients s_j cannot all be equal to zero since $v_l \notin \operatorname{Span}\mathcal{B}_0$. After reordering the vectors u_j we can suppose that $s_l \neq 0$. By Theorem XX, the set $\{v_1, v_2, \cdots, v_{l-1}, v_l, u_{l+1}, \cdots u_m\}$ is a basis fot U (in the case that $l - 1 = m$ this basis is the set $\mathcal{B}$).

1.33 Corollary: *For a vector space U in F^n, any two bases for U have the same number of elements.*

1.34 Definition: For a vector space U in F^n, we define the **dimension** of U, denoted by $\dim U$, to be the number of elements in any basis for U. For an affine space $P = p + U$ in F^n, we define the **dimension** of P to be $\dim P = \dim U$.