MATH 217- LINEAR ALGEBRA
AND DIFFERENTIAL EQUATIONS
SPRING 2022-2023

***Instructor*:** Emil Novruz, ***e-mail*:** enovruz@gtu.edu.tr
***Office*:** Faculty of Management – Z44,
***Office Hours*:** Thursday 10:30-12:30
Textbooks:
1) D. Kreider, R. Kuller, D. Ostberg, F. Perkins, An Introduction to Linear Analysis;
2) Shepley L. Ross, Differential Equations;
3) B.Kolman, D.R.Hill, Elementary Linear Algebra with Applications.

Exams and grading: There will be one midterm exam and one final exam. The weights are as follows:

Final Grading	Percentage
Midterm Exam	%40
Final exam	%60

Attendance: At least 70%

Week	Topics
1	Vector Spaces, Dimension.
2	Linear Transformations.
3	Matrices.
4	Determinants.
5	Eigenvalues and Eigenvectors. Orthogonality.
6	General Theory of Linear System.
7	Exact Differential Equations, Separable Equations, Homogeneous Equations, Linear Equations.
8	**Midterm Exam,** Bernoulli and Ricatti equations.
9	Method of Integrating Factors.
10	Special Integrating Factors Transformations.
11	General Theory of Linear Differential Equations.
12	Linear Differential Equations with Constant Coefficients.
13	The method of undetermined coefficients. Variation of parameters.
14	The Cauchy-Euler equation. Laplace Transform

Vector Spaces

Definition. A *real vector space* $\mathcal{V}$ is a collection of objects called *vectors*, together with operations of addition and multiplication by real numbers which satisfy the following axioms.

Axioms for addition. Given any pair of vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{V}$ there exists a (unique) vector $\mathbf{x} + \mathbf{y}$ in $\mathcal{V}$ called the *sum* of $\mathbf{x}$ and $\mathbf{y}$. It is required that

(i) addition be *associative*,

$$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z},$$

(ii) addition be *commutative*,

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x},$$

(iii) there exist a vector $\mathbf{0}$ in $\mathcal{V}$ (called the *zero vector*) such that

$$\mathbf{x} + \mathbf{0} = \mathbf{x}$$

for all $\mathbf{x}$ in $\mathcal{V}$, and

(iv) for each $\mathbf{x}$ in $\mathcal{V}$ there exist a vector $-\mathbf{x}$ in $\mathcal{V}$ such that

$$\mathbf{x} + (-\mathbf{x}) = \mathbf{0}.$$

Axioms for scalar multiplication. Given any vector $\mathbf{x}$ in $\mathcal{V}$ and any real number α there exists a (unique) vector $\alpha\mathbf{x}$ in $\mathcal{V}$ called the *product*, or *scalar product*, of α and $\mathbf{x}$. It is required that

(v) $$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y},$$

(vi) $$(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x},$$

(vii) $$(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x}),$$

(viii) $$1\mathbf{x} = \mathbf{x}.$$

EXAMPLE. Let n be a fixed positive integer, and let $\mathcal{R}^n$ denote the totality of ordered n-tuples $(x_1, \ldots, x_n)$ of real numbers. If $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ are two such n-tuples, and α is a real number, set

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$$

and

$$\alpha\mathbf{x} = (\alpha x_1, \ldots, \alpha x_n).$$

Then $\mathcal{R}^n$ becomes a real vector space.

EXAMPLE. The set l_2 of all infinite sequences

$$x = (x_1, x_2, \ldots, x_k, \ldots) \tag{1}$$

of real or complex numbers $x_1, x_2, \ldots, x_k, \ldots$ satisfying the convergence condition

$$\sum_{k=1}^{\infty} |x_k|^2 < \infty,$$

equipped with operations

$$\begin{aligned}(x_1, x_2, \ldots, x_k, \ldots) + (y_1, y_2, \ldots, y_k, \ldots) \\ = (x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k, \ldots), \\ \alpha(x_1, x_2, \ldots, x_k, \ldots) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_k, \ldots),\end{aligned} \tag{2}$$

is a linear space. The fact that

$$\sum_{k=1}^{\infty} |x_k|^2 < \infty, \qquad \sum_{k=1}^{\infty} |y_k|^2 < \infty$$

implies

$$\sum_{k=1}^{\infty} |x_k + y_k|^2 < \infty$$

is an immediate consequence of the elementary inequality

$$(x_k + y_k)^2 \leqslant 2(x_k^2 + y_k^2).$$

EXAMPLE

Let Map($\mathbb{R}, \mathbb{R}$) be the set of all mappings $f : \mathbb{R} \to \mathbb{R}$. For two such mappings f, g define $f + g : \mathbb{R} \to \mathbb{R}$ to be the mapping given by the prescription

$$(f+g)(x) = f(x) + g(x),$$

and for every scalar $\lambda \in \mathbb{R}$ define $\lambda f : \mathbb{R} \to \mathbb{R}$ to be the mapping given by the prescription

$$(\lambda f)(x) = \lambda f(x).$$

the vector 0 the mapping ϑ such that $\vartheta(x) = 0$ for every $x \in \mathbb{R}$

$-f$ the mapping given by $(-f)(x) = -f(x)$ for every $x \in \mathbb{R}$.

These operations make Map($\mathbb{R}, \mathbb{R}$) into a real vector space.

EXAMPLE

Let $\mathbb{R}_n[X]$ be the set of polynomials of degree at most n with real coefficients The reader will recognise this as the set of objects of the form

$$a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$$

where each $a_i \in \mathbb{R}$ and X is an 'indeterminate', the largest suffix i for which $a_i \neq 0$ being the *degree* of the polynomial.

We can define an addition on $\mathbb{R}_n[X]$ by setting

$$\begin{aligned}&(a_0 + a_1X + \cdots + a_nX^n) + (b_0 + b_1X + \cdots + b_nX^n)\\ &= (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_n + b_n)X^n\end{aligned}$$

and a multiplication by scalars by

$$\lambda(a_0 + a_1X + \cdots + a_nX^n) = \lambda a_0 + \lambda a_1X + \cdots + \lambda a_nX^n.$$

In this way $\mathbb{R}_n[X]$ has the structure of a real vector space.

Now we note a number of immediate consequences of Definition. The first of these concerns the zero vector and asserts that this vector behaves very much as one might expect. Specifically,

$$0\mathbf{x} = \mathbf{0} \qquad \text{for every } \mathbf{x}, \tag{1}$$

and

$$\alpha\mathbf{0} = \mathbf{0} \qquad \text{for every } \alpha. \tag{2}$$

To prove the first of these assertions set $\alpha = \beta = 0$ in $\alpha\mathbf{x} + \beta\mathbf{x} = (\alpha + \beta)\mathbf{x}$. This gives

$$0\mathbf{x} + 0\mathbf{x} = (0 + 0)\mathbf{x} = 0\mathbf{x}.$$

Now subtract $0\mathbf{x}$ from both sides of this equation, and then use the fact that $0\mathbf{x} - 0\mathbf{x} = \mathbf{0}$ to obtain

$$0\mathbf{x} + (0\mathbf{x} - 0\mathbf{x}) = 0\mathbf{x} - 0\mathbf{x}$$

and

$$0\mathbf{x} + \mathbf{0} = \mathbf{0}.$$

Hence $0\mathbf{x} = \mathbf{0}$.

The proof of (2) is similar; this time set $\mathbf{x} = \mathbf{y} = \mathbf{0}$ in $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.

Next, an equally elementary proof establishes the fact that the vector $-\mathbf{x}$ and $(-1)\mathbf{x}$ are one and the same. Indeed, since $1\mathbf{x} = \mathbf{x}$ and $0\mathbf{x} = \mathbf{0}$, we have

$$\mathbf{x} + (-1)\mathbf{x} = 1\mathbf{x} + (-1)\mathbf{x} = (1 - 1)\mathbf{x} = 0\mathbf{x} = \mathbf{0}.$$

Now subtract $\mathbf{x}$ from both sides of this equation to obtain $(-1)\mathbf{x} = -\mathbf{x}$, as asserted.

Lemma 1. *If $\mathbf{0}'$ is a vector in $\mathcal{V}$ such that $\mathbf{x} + \mathbf{0}' = \mathbf{x}$ for every $\mathbf{x}$ in $\mathcal{V}$, then $\mathbf{0}' = \mathbf{0}$. Similarly, if $\mathbf{x}'$ is any vector in $\mathcal{V}$ such that $\mathbf{x} + \mathbf{x}' = \mathbf{0}$, then $\mathbf{x}' = -\mathbf{x}$.*

Proof. If $\mathbf{x} + \mathbf{0}' = \mathbf{x}$ for every $\mathbf{x}$ in $\mathcal{V}$, we have, in particular,

$$\mathbf{0} + \mathbf{0}' = \mathbf{0}.$$

On the other hand, the zero vector has the property that $\mathbf{0} + \mathbf{x} = \mathbf{x}$ for every $\mathbf{x}$. Hence

$$\mathbf{0} + \mathbf{0}' = \mathbf{0}',$$

and it follows that

$$\mathbf{0} = \mathbf{0}'.$$

The second statement of the lemma follows from the sequence of equalities

$$\mathbf{x}' = \mathbf{0} + \mathbf{x}' = (-\mathbf{x} + \mathbf{x}) + \mathbf{x}' = -\mathbf{x} + (\mathbf{x} + \mathbf{x}') = -\mathbf{x} + \mathbf{0} = -\mathbf{x}. \blacksquare$$

Definition. A subset $\mathcal{W}$ of a vector space $\mathcal{V}$ is said to be a *subspace* of $\mathcal{V}$ if $\mathcal{W}$ itself is a vector space under the operations of addition and scalar multiplication defined in $\mathcal{V}$.

Example.

In the real vector space $\mathbb{R}^2$ the set $X = \{(x,0)\ ;\ x \in \mathbb{R}\}$ is a subspace; for we have

$$(x_1,0) + (x_2,0) = (x_1 + x_2, 0);$$
$$\lambda(x,0) = (\lambda x, 0),$$

and so X is closed under addition and multiplication by scalars. This subspace is simply the 'x-axis' in the cartesian plane $\mathbb{R}^2$. Similarly, the 'y-axis'

$$Y = \{(0,y)\ ;\ y \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^2$.

Subspace Criterion. *If every vector of the form*

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$$

belongs to $\mathcal{W}$ whenever $\mathbf{x}_1$ and $\mathbf{x}_2$ belong to $\mathcal{W}$, and α_1 and α_2 are arbitrary scalars, then $\mathcal{W}$ is a subspace of $\mathcal{V}$.

To prove this assertion, we must show that $\mathcal{W}$ satisfies first Definition . But by first setting $\alpha_1 = \alpha_2 = 1$, and then setting $\alpha_2 = 0$, we deduce in turn that
(a) the sum of any two vectors in $\mathcal{W}$ again belongs to $\mathcal{W}$, and

(b) $\alpha\mathbf{x}$ belongs to $\mathcal{W}$ for every real number α and every $\mathbf{x}$ in $\mathcal{W}$.

From (b) it follows in particular that $-\mathbf{x}$ belongs to $\mathcal{W}$ whenever $\mathbf{x}$ does, and that $\mathcal{W}$ also contains the zero vector. Thus $\mathcal{W}$ satisfies Axioms (iii) and (iv) of Definition . Finally, we observe that the remaining axioms certainly hold in $\mathcal{W}$, since they are valid everywhere in $\mathcal{V}$. Hence $\mathcal{W}$ is a subspace of $\mathcal{V}$. ▌

EXAMPLE 1. Every vector space has two subspaces: (a) the whole space, and (b) the subspace consisting of the zero vector by itself, called the *trivial subspace*. A subspace of $\mathcal{V}$ which is distinct from $\mathcal{V}$ is called a *proper subspace*.

EXAMPLE 2. If $\mathcal{W}$ is the subset of $\mathcal{R}^3$ consisting of all those vectors whose third component is zero, then the above criterion implies at once that $\mathcal{W}$ is a subspace of $\mathcal{R}^3$. When the components of each vector in $\mathcal{R}^3$ are viewed as its ordinary x, y, z-components, then $\mathcal{W}$ is just the (x, y)-plane in 3-space.

EXAMPLE 3. Let $\mathcal{C}^1[a, b]$ denote the set of all functions which possess a continuous derivative at every point of the interval $[a, b]$; i.e., the so-called *continuously differentiable functions* on $[a, b]$. Since a differentiable function is continuous, each function in $\mathcal{C}^1[a, b]$ also belongs to $\mathcal{C}[a, b]$. But both the scalar multiple of a continuously differentiable function and the sum of two such functions are continuously differentiable. Hence $\mathcal{C}^1[a, b]$ is closed under addition and scalar multiplication and thus is a subspace of $\mathcal{C}[a, b]$. More generally, if $\mathcal{C}^n[a, b]$ denotes the set of all *n times continuously differentiable functions* on $[a, b]$, then $\mathcal{C}^m[a, b]$ is a subspace of $\mathcal{C}^n[a, b]$ whenever $m \geq n$.

Lemma. *If $\mathcal{W}_1$ and $\mathcal{W}_2$ are subspaces of $\mathcal{V}$, then the set consisting of all vectors belonging to both $\mathcal{W}_1$ and $\mathcal{W}_2$ is a subspace of $\mathcal{V}$.*

Proof. Let $\mathcal{W}$ be the set in question, and note that $\mathcal{W}$ contains the zero vector since this vector belongs to both $\mathcal{W}_1$ and $\mathcal{W}_2$. Now let $\mathbf{x}_1$ and $\mathbf{x}_2$ be any two vectors in $\mathcal{W}$. Then $\mathbf{x}_1$ and $\mathbf{x}_2$ belong to $\mathcal{W}_1$ and to $\mathcal{W}_2$, and hence so does $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$ for any pair of real numbers α_1 and α_2. This implies that $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$ belongs to $\mathcal{W}$, and the assertion that $\mathcal{W}$ is a subspace of $\mathcal{V}$ now follows from the subspace criterion. ▌

The subspace $\mathcal{W}$ of this lemma is known as the *intersection* of $\mathcal{W}_1$ and $\mathcal{W}_2$, and is denoted $\mathcal{W}_1 \cap \mathcal{W}_2$ (read "$\mathcal{W}_1$ intersect $\mathcal{W}_2$").

We now return to the problem of finding all subspaces of an arbitrary vector space $\mathcal{V}$. Rather than attempt a frontal assault on this problem, it turns out to be much more profitable to proceed as follows: Let $\mathfrak{X}$ be *any* (nonempty) subset of $\mathcal{V}$. Then, as was noted above, there is at least one *subspace* of $\mathcal{V}$ containing $\mathfrak{X}$, namely $\mathcal{V}$ itself. This being so, we attempt to find the "smallest" subspace of $\mathcal{V}$ containing $\mathfrak{X}$, where by this we mean *that* subspace of $\mathcal{V}$ which contains $\mathfrak{X}$, and which in turn is contained in every subspace of $\mathcal{V}$ containing $\mathfrak{X}$. To show that such a subspace actually exists, consider the totality of *all* subspaces of $\mathcal{V}$ which contain $\mathfrak{X}$, and let $\mathcal{S}(\mathfrak{X})$ denote the set of vectors belonging to *every one* of these subspaces; i.e., $\mathcal{S}(\mathfrak{X})$ is the intersection of these subspaces. Reasoning as in the proof of Lemma, we see that $\mathcal{S}(\mathfrak{X})$ is a subspace of $\mathcal{V}$, and from its very definition it is clear that there is no subspace of $\mathcal{V}$ which contains $\mathfrak{X}$ and is *properly* contained in $\mathcal{S}(\mathfrak{X})$. Thus $\mathcal{S}(\mathfrak{X})$ is the desired subspace. It is called the subspace of $\mathcal{V}$ *spanned* by $\mathfrak{X}$ and, as we shall see, is uniquely determined by the set $\mathfrak{X}$.

All this is well and good, but unless we can discover an easy method for finding $\mathcal{S}(\mathfrak{X})$ in terms of the vectors belonging to $\mathfrak{X}$, we will have made little progress on the problem of surveying the subspaces of $\mathcal{V}$. Fortunately (and this is the reason for introducing $\mathcal{S}(\mathfrak{X})$ in the first place) such a method is easy to derive. To do so, we introduce the following definition.

Definition. An expression of the form

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n, \qquad (*)$$

where $\alpha_1, \ldots, \alpha_n$ are real numbers, is called a *linear combination* of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$.

And now we can describe $\mathcal{S}(\mathfrak{X})$: *it is the set of all linear combinations of the vectors in* $\mathfrak{X}$.

Theorem. *Let $\mathfrak{X}$ be a (nonempty) subset of a vector space $\mathcal{V}$. Then the subspace of $\mathcal{V}$ spanned by $\mathfrak{X}$ consists of all linear combinations of the vectors in $\mathfrak{X}$.*

Proof. In the first place, the set of all linear combinations of vectors in $\mathfrak{X}$ is closed under addition and scalar multiplication, and hence is a subspace $\mathcal{W}$ of $\mathcal{V}$. Moreover, the equation $\mathbf{x} = 1\mathbf{x}$ shows that each $\mathbf{x}$ in $\mathfrak{X}$ is a linear combination of vectors in $\mathfrak{X}$, thus proving that $\mathfrak{X}$ is contained in $\mathcal{W}$. Finally, every subspace of $\mathcal{V}$ which contains $\mathfrak{X}$ must contain all vectors of the form $(*)$ by virtue of the fact that a subspace is closed under addition and scalar multiplication. In other words, $\mathcal{W}$ is contained in every subspace of $\mathcal{V}$ containing $\mathfrak{X}$, and it follows that $\mathcal{W} = \mathcal{S}(\mathfrak{X})$. ▌

LINEAR DEPENDENCE AND INDEPENDENCE; BA$ES

Definition. A vector $\mathbf{x}$ is said to be *linearly dependent* on $\mathbf{x}_1, \ldots, \mathbf{x}_n$ if $\mathbf{x}$ can be written in the form

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n,$$

where the α_i are scalars. If, on the other hand, no such relation exists, $\mathbf{x}$ is said to be *linearly independent* of $\mathbf{x}_1, \ldots, \mathbf{x}_n$.

Test for linear independence. *The vectors* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *are linearly independent if and only if the equation*

$$\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n = \mathbf{0}$$

implies that $\alpha_1 = \cdots = \alpha_n = 0.$

For instance, $\mathbf{x}_1 = (1, 3, -1, 2)$, $\mathbf{x}_2 = (2, 0, 1, 3)$, $\mathbf{x}_3 = (-1, 1, 0, 0)$ are linearly independent in $\mathcal{R}^4$ since the equation

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 = \mathbf{0}$$

implies that

$$\begin{aligned} \alpha_1 + 2\alpha_2 - \alpha_3 &= 0, \\ 3\alpha_1 \qquad\quad + \alpha_3 &= 0, \\ -\alpha_1 + \alpha_2 \qquad\quad &= 0, \\ 2\alpha_1 + 3\alpha_2 \qquad\quad &= 0, \end{aligned}$$

from which it easily follows that $\alpha_1 = \alpha_2 = \alpha_3 = 0.$

And now we are ready to show how one can weed the extraneous vectors from any finite set $\mathbf{x}_1, \ldots, \mathbf{x}_n$ without disturbing $\mathcal{S}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. The basic idea is obvious; just get rid of as many linearly dependent vectors as possible.

To accomplish this we begin with the vector $\mathbf{x}_n$. If $\mathbf{x}_n$ is linearly dependent on $\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}$, then

$$\mathbf{x}_n = \alpha_1\mathbf{x}_1 + \cdots + \alpha_{n-1}\mathbf{x}_{n-1},$$

and we can rewrite the expression

$$\mathbf{x} = \beta_1\mathbf{x}_1 + \cdots + \beta_n\mathbf{x}_n$$

for an arbitrary vector in $\mathcal{S}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ in the form

$$\mathbf{x} = (\beta_1 + \alpha_1\beta_n)\mathbf{x}_1 + \cdots + (\beta_{n-1} + \alpha_{n-1}\beta_n)\mathbf{x}_{n-1}.$$

This proves that $\mathbf{x}$ is already a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}$, and hence that $\mathcal{S}(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}) = \mathcal{S}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. In this case we drop the vector $\mathbf{x}_n$ from the set $\mathbf{x}_1, \ldots, \mathbf{x}_n$. If, on the other hand, $\mathbf{x}_n$ is *not* linearly dependent on $\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}$, we keep it.

If we repeat this procedure with each of the $\mathbf{x}_i$ in turn, dropping $\mathbf{x}_i$ if it is linearly dependent on the remaining vectors in the (possibly modified) set, keeping it otherwise, it is clear that we obtain a linearly independent subset of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ which spans the subspace $\mathcal{S}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. This, of course, is what we started out to show, and we have proved

Theorem. *Every finite set of vectors $\mathfrak{X}$ contains a linearly independent subset which spans the subspace $\mathcal{S}(\mathfrak{X})$.*

Definition. A finite linearly independent subset $\mathcal{B}$ of a vector space $\mathcal{V}$ is said to be a *basis* for $\mathcal{V}$ if $\mathcal{S}(\mathcal{B}) = \mathcal{V}$.

As an example, we cite the vectors $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, $\mathbf{k} = (0, 0, 1)$, which form a basis for $\mathcal{R}^3$. We shall prove this assertion in Example 1 below, and now merely wish to observe that every vector $\mathbf{x} = (x_1, x_2, x_3)$ in $\mathcal{R}^3$ can be written in *one and only one* way as a linear combination of these basis vectors, namely, $\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$. This last property actually serves to characterize a basis in a vector space, as we now show.

Theorem. *A set of vectors* $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *is a basis for a vector space* $\mathcal{V}$ *if and only if every vector in* $\mathcal{V}$ *can be written uniquely as a linear combination of* $\mathbf{e}_1, \ldots, \mathbf{e}_n$.

Proof. First suppose that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis for $\mathcal{V}$. Then the $\mathbf{e}_i$ span $\mathcal{V}$, and hence every vector in $\mathcal{V}$ can be written in *at least one* way as

$$\mathbf{x} = \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n. \tag{1}$$

To show that this is the only such expression possible, let

$$\mathbf{x} = \beta_1\mathbf{e}_1 + \cdots + \beta_n\mathbf{e}_n \tag{2}$$

be another. Then, subtracting (2) from (1), we obtain

$$\mathbf{0} = (\alpha_1 - \beta_1)\mathbf{e}_1 + \cdots + (\alpha_n - \beta_n)\mathbf{e}_n. \tag{3}$$

But since $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis for $\mathcal{V}$, these vectors are linearly independent. Hence, by our test for linear independence, each of the coefficients in (3) is zero, and it follows that $\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$, as desired.

Conversely, suppose every vector in $\mathcal{V}$ can be written *uniquely* as a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Then these vectors certainly span $\mathcal{V}$, and we need only prove their linear independence in order to show that they are a basis for $\mathcal{V}$. To accomplish this, we observe that $\mathbf{0} = 0\mathbf{e}_1 + \cdots + 0\mathbf{e}_n$ and that our assumption concerning the uniqueness of such expressions implies that this is the *only* representation of $\mathbf{0}$ as a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Thus if $\alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0}$, we must have $\alpha_1 = \cdots = \alpha_n = 0$, and the test for linear independence now applies. ▌

EXAMPLE 1. The vectors

$$\begin{aligned}
\mathbf{e}_1 &= (1, 0, \ldots, 0), \\
\mathbf{e}_2 &= (0, 1, \ldots, 0), \\
&\vdots \\
\mathbf{e}_n &= (0, 0, \ldots, 1)
\end{aligned}$$

are a basis for $\mathcal{R}^n$, since $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ is the only way of expressing the vector $\mathbf{x} = (x_1, \ldots, x_n)$ as a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$. This particular basis is called the *standard basis* for $\mathcal{R}^n$.

EXAMPLE 2. Again in $\mathcal{R}^n$, let

$$\begin{aligned}
\mathbf{e}'_1 &= (1, 0, \ldots, 0), \\
\mathbf{e}'_2 &= (1, 1, \ldots, 0), \\
&\vdots \\
\mathbf{e}'_n &= (1, 1, \ldots, 1),
\end{aligned}$$

where, in general, $\mathbf{e}'_i$ is the n-tuple having 1's in the first i places and 0's thereafter. Then $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$ is a basis for $\mathcal{R}^n$. To prove this let $\mathbf{x} = (x_1, \ldots, x_n)$ be given, and let us attempt to find real numbers $\alpha_1, \ldots, \alpha_n$ such that $\mathbf{x} = \alpha_1\mathbf{e}'_1 + \cdots + \alpha_n\mathbf{e}'_n$. In order that such an equality hold we must have

$$\begin{aligned}
(x_1, \ldots, x_n) &= \alpha_1(1, 0, \ldots, 0) + \alpha_2(1, 1, \ldots, 0) + \cdots + \alpha_n(1, 1, \ldots, 1) \\
&= (\alpha_1, 0, \ldots, 0) + (\alpha_2, \alpha_2, \ldots, 0) + \cdots + (\alpha_n, \alpha_n, \ldots, \alpha_n) \\
&= (\alpha_1 + \alpha_2 + \cdots + \alpha_n, \alpha_2 + \cdots + \alpha_n, \ \ldots \ , \alpha_n),
\end{aligned}$$

which leads to the system of equations

$$\begin{aligned} \alpha_1 + \alpha_2 + \cdots + \alpha_n &= x_1, \\ \alpha_2 + \cdots + \alpha_n &= x_2, \\ &\vdots \\ \alpha_n &= x_n. \end{aligned}$$

Hence

$$\begin{aligned} \alpha_1 &= x_1 - x_2, \\ \alpha_2 &= x_2 - x_3, \\ &\vdots \\ \alpha_{n-1} &= x_{n-1} - x_n, \\ \alpha_n &= x_n, \end{aligned}$$

which simultaneously shows that $\mathbf{x}$ can be written as a linear combination of $\mathbf{e}_1', \ldots, \mathbf{e}_n'$, and that the coefficients of this relation are uniquely determined. Thus the $\mathbf{e}_i'$ are a basis for $\mathcal{R}^n$, as asserted.

Definition. A vector space is said to be of *dimension n* if it has a basis consisting of n vectors, and is said to be *infinite dimensional* otherwise. We denote the fact that $\mathcal{V}$ is n-dimensional by writing $\dim \mathcal{V} = n$.

DIMENSION

Theorem*. *If $\mathcal{V}$ has a basis containing n vectors, then any $n + 1$ or more vectors in $\mathcal{V}$ are linearly dependent.*

Proof. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be a basis for $\mathcal{V}$, and suppose, contrary to the assertion of the theorem, that $\mathcal{V}$ contains a linearly independent set $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$ in which $m > n$. Express each of the $\mathbf{e}'_j$ as a linear combination of the $\mathbf{e}_i$, thereby obtaining the system of equations

$$\begin{aligned}
\mathbf{e}'_1 &= \alpha_{11}\mathbf{e}_1 + \alpha_{21}\mathbf{e}_2 + \cdots + \alpha_{n1}\mathbf{e}_n,\\
\mathbf{e}'_2 &= \alpha_{12}\mathbf{e}_1 + \alpha_{22}\mathbf{e}_2 + \cdots + \alpha_{n2}\mathbf{e}_n,\\
&\vdots\\
\mathbf{e}'_m &= \alpha_{1m}\mathbf{e}_1 + \alpha_{2m}\mathbf{e}_2 + \cdots + \alpha_{nm}\mathbf{e}_n,
\end{aligned} \tag{1}$$

in which the α_{ij} are scalars. Since none of the $\mathbf{e}'_j$ is the zero vector, at least one of the α_{ij} is different from zero in each of these equations. (Recall that the zero vector is linearly dependent on every vector in $\mathcal{V}$.) Thus, by relabeling the $\mathbf{e}_i$ if necessary, we may assume that $\alpha_{11} \neq 0$. This done, solve the first equation for $\mathbf{e}_1$, and substitute the value obtained in the remaining $m - 1$ equations. This eliminates $\mathbf{e}_1$ from (1), and yields a system of equations of the form

$$\begin{aligned}
\mathbf{e}'_2 &= \beta_{22}\mathbf{e}_2 + \beta_{32}\mathbf{e}_3 + \cdots + \beta_{n2}\mathbf{e}_n + \beta_{12}\mathbf{e}'_1,\\
\mathbf{e}'_3 &= \beta_{23}\mathbf{e}_2 + \beta_{33}\mathbf{e}_3 + \cdots + \beta_{n3}\mathbf{e}_n + \beta_{13}\mathbf{e}'_1,\\
&\vdots\\
\mathbf{e}'_m &= \beta_{2m}\mathbf{e}_2 + \beta_{3m}\mathbf{e}_3 + \cdots + \beta_{nm}\mathbf{e}_n + \beta_{1m}\mathbf{e}'_1.
\end{aligned} \tag{2}$$

Focusing our attention on the first of these equations we note that the linear independence of $\mathbf{e}'_1$ and $\mathbf{e}'_2$ implies that at least one of the coefficients β_{22}, $\beta_{32}, \ldots, \beta_{n2}$ is different from zero. Assume that the $\mathbf{e}_i$ are labeled so that $\beta_{22} \neq 0$. Then a repetition of the above argument, now applied to $\mathbf{e}_2$, reduces (2) to the system

$$
\begin{aligned}
\mathbf{e}'_3 &= \gamma_{33}\mathbf{e}_3 + \cdots + \gamma_{n3}\mathbf{e}_n + \gamma_{13}\mathbf{e}'_1 + \gamma_{23}\mathbf{e}'_2, \\
&\vdots \\
\mathbf{e}'_m &= \gamma_{3m}\mathbf{e}_3 + \cdots + \gamma_{nm}\mathbf{e}_n + \gamma_{1m}\mathbf{e}'_1 + \gamma_{2m}\mathbf{e}'_2.
\end{aligned}
$$

Let us now speculate on the effect of our assumption that m is greater than n. A moment's thought will reveal that by continuing the above process of elimination we will eventually find ourselves confronted with a system of $m - n$ equations expressing each of the vectors $\mathbf{e}'_{n+1}, \ldots, \mathbf{e}'_m$ as a linear combination of $\mathbf{e}'_1, \ldots, \mathbf{e}'_n$. But this cannot be. Hence $m \leq n$ after all. ▌

Corollary. *If $\mathcal{V}$ has a basis containing n vectors, then every basis for $\mathcal{V}$ contains n vectors.*

Proof. If $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and $\mathbf{e}'_1, \ldots, \mathbf{e}'_m$ are bases for $\mathcal{V}$, then the above theorem implies that $m \leq n$, and $n \leq m$. Hence $m = n$. ▌

Theorem. *If $\mathcal{W}$ is a subspace of an n-dimensional vector space $\mathcal{V}$, then* $\dim \mathcal{W} \leq n$.

Proof. The theorem is obviously true if $n = 0$, or if $\mathcal{W}$ is the trivial subspace of $\mathcal{V}$. Thus we can assume $n > 0$, and $\mathcal{W}$ nontrivial.

By virtue of this last assumption, $\mathcal{W}$ contains linearly independent sets of vectors, since any nonzero vector in $\mathcal{W}$ is, by itself, such a set. Moreover, every linearly independent set in $\mathcal{W}$ is also linearly independent as a set in $\mathcal{V}$. Thus, by the previous theorem, the number of vectors in such a set cannot exceed n. Finally, if $\mathbf{e}_1, \ldots, \mathbf{e}_m$ is a linearly independent set in $\mathcal{W}$ containing a *maximum* number of vectors, then $\mathcal{S}(\mathbf{e}_1, \ldots, \mathbf{e}_m) = \mathcal{W}$. Hence $\dim \mathcal{W} = m \leq n$, as advertised. ▌

This theorem may be read as asserting that every nontrivial subspace $\mathcal{W}$ of an n-dimensional space $\mathcal{V}$ has a basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$ with $m \leq n$. If $m = n$, then $\mathbf{e}_1, \ldots, \mathbf{e}_m$ is also a basis for $\mathcal{V}$, and $\mathcal{W} = \mathcal{V}$. On the other hand, if $m < n$, then $\mathcal{W}$ is a *proper* subspace of $\mathcal{V}$ (i.e., $\mathcal{W} \neq \mathcal{V}$), and there exist vectors in $\mathcal{V}$ which do not belong to $\mathcal{W}$. Choose any such vector, and label it $\mathbf{e}_{m+1}$. Then it is all but obvious that $\mathbf{e}_1, \ldots, \mathbf{e}_{m+1}$ are linearly independent in $\mathcal{V}$.

To prove the truth of this observation, we apply the test for linear independence as follows. Suppose that

$$\alpha_1\mathbf{e}_1 + \cdots + \alpha_m\mathbf{e}_m + \alpha_{m+1}\mathbf{e}_{m+1} = \mathbf{0}. \tag{1}$$

Then $\alpha_{m+1} = 0$, for otherwise

$$\mathbf{e}_{m+1} = -\frac{\alpha_1}{\alpha_{m+1}}\mathbf{e}_1 - \cdots - \frac{\alpha_m}{\alpha_{m+1}}\mathbf{e}_m,$$

and $\mathbf{e}_{m+1}$ is in $\mathcal{W}$. Thus

$$\alpha_1\mathbf{e}_1 + \cdots + \alpha_m\mathbf{e}_m = \mathbf{0},$$

and it follows from the linear independence of $\mathbf{e}_1, \ldots, \mathbf{e}_m$ that $\alpha_1 = \cdots = \alpha_m = 0$.

Hence all of the coefficients in (1) are zero, and $\mathbf{e}_1, \ldots, \mathbf{e}_{m+1}$ are linearly independent.

We now repeat the above argument, this time starting with the subspace $\mathcal{S}(\mathbf{e}_1, \ldots, \mathbf{e}_{m+1})$. If $\mathcal{S}(\mathbf{e}_1, \ldots, \mathbf{e}_{m+1})$ is a proper subspace of $\mathcal{V}$ we can enlarge $\mathbf{e}_1, \ldots, \mathbf{e}_{m+1}$ to a linearly independent set in $\mathcal{V}$ containing $m + 2$ vectors. But

Theorem* implies that this process must come to a halt after $n - m$ steps, at which point we will have a *basis* for $\mathcal{V}$. With this we have proved the following important and useful result.

Theorem. *Let $\mathcal{V}$ be an n-dimensional vector space, and let* $\mathbf{e}_1, \ldots, \mathbf{e}_m$ *be a basis for an m-dimensional subspace of* $\mathcal{V}$. *Then there exist* $n - m$ *vectors* $\mathbf{e}_{m+1}, \ldots, \mathbf{e}_n$ *in* $\mathcal{V}$ *such that* $\mathbf{e}_1, \ldots, \mathbf{e}_m, \mathbf{e}_{m+1}, \ldots, \mathbf{e}_n$ *is a basis for* $\mathcal{V}$.