We recall that the \( n \)th order scalar ordinary difference equations and differential equations can be rewritten in the form of systems of first-order difference and differential equations, i.e.

\[
x(k + 1) = f(x(k), k), \\
dx(t)/dt = f(x(t), t),
\]

where \( x(k) \) and \( x(t) \) can be regarded as points in a \( n \)-dimensional real coordinate space (the state space). The evolution of the solutions of these equations with \( k \) or \( t \) gives rise to sequences of points or curves in the state space.

In order to make the notion of state space useful, it is necessary to introduce certain structure on the state space so that it permits performing certain operations among the elements of the set with consistency. It is at this point, algebra enters the picture. Depending on what structure is imposed on the sets, we have various algebraic systems.

The objective here is to develop certain notions and results in linear algebra that are useful in the solution of systems of first-order difference and differential equations.

1. Vector (Linear) Spaces:

First, we introduce the notion of a field.

**Definition 4.1.** A field is a triple \( \{F, +, \cdot\} \), where \( F \) is a nonempty set whose elements are called scalars. The operations “+” and “·” correspond to addition and multiplication between scalars that satisfy the following properties:

(a) To every pair of scalars \( \alpha \) and \( \beta \), there corresponds a scalar \( \alpha + \beta \) (the sum of \( \alpha \) and \( \beta \)) such that

(i) \( \alpha + \beta = \beta + \alpha \) (commutativity) ;

(ii) \( \alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma \) (associativity) ;

(iii) there exists a unique scalar 0 (called the zero scalar) such that \( \alpha + 0 = \alpha \) for every scalar \( \alpha \);

(iv) to every scalar \( \alpha \), there corresponds a unique scalar \(-\alpha\) such that \( \alpha + (-\alpha) = 0 \).

(b) To every pair of scalars \( \alpha \) and \( \beta \), there corresponds a scalar \( \alpha \cdot \beta \) in \( F \) called the product of \( \alpha \) and \( \beta \) such that

(i) \( \alpha \cdot \beta = \beta \cdot \alpha \) (commutativity);

(ii) \( \alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma \) (associativity);

(iii) there exists a unique nonzero scalar “1” (multiplicative identity) such that \( \alpha \cdot 1 = \alpha \) for every scalar \( \alpha \);

(iv) to every nonzero scalar \( \alpha \), there corresponds a unique scalar \( \alpha^{-1} \) (multiplicative inverse) such that \( \alpha \cdot \alpha^{-1} = 1 \).

(c) For any scalars \( \alpha, \beta \) and \( \gamma \), \( \alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma \) (distributivity with respect to addition).
The most important scalar fields which we will encounter in this course are \( \mathcal{R} \) (the field of real numbers) and \( \mathcal{C} \) (the field of complex numbers). Here, “+” and “·” correspond to the usual addition and multiplication of two real or complex numbers.

**Example:** Consider the set of integers \( I_N = \{0, 1, 2, \ldots, N-1\} \). If we define “+” and “·” by the usual addition and multiplication of real numbers respectively, then \( I_N \) is not a scalar field, since additive and multiplicative inverses do not exist (i.e. \(-\alpha\) and \(\alpha^{-1}\) do not belong to \( I_N \)).

**Question:** Can we refine “+” and “·” so that \( I_N \) becomes a scalar field?

Let “+” and “·” be defined by

\[
\alpha + \beta = \text{remainder of } (\alpha + \beta)/N, \text{ and } \alpha \cdot \beta = \text{remainder of } (\alpha \cdot \beta)/N.
\]

For \( N = 3 \), we have
- \( 2 + 2 = \text{remainder of } (2 + 2)/3 = 1; \)
- \( 1 + 2 = \text{remainder of } (1 + 2)/3 = 0; \)
- \( 1 + 1 = \text{remainder of } (1 + 1)/3 = 2, \text{ etc.} \)

Thus, we can construct the following table for addition:

\[
\begin{array}{ccc}
+ & 0 & 1 & 2 \\
0 & 0 & 1 & 2 \\
1 & 1 & 2 & 0 \\
2 & 2 & 0 & 1 \\
\end{array}
\]

- Additive identity (zero scalar “0”) is 0, i.e. \( 2 + 0 = \text{remainder of } (2 + 0)/3 = 2. \)
- Additive inverse of 2 is 1 (remainder of \( (2 + 1)/3 = 0 \));
  additive inverse of 1 is 2 (remainder of \( (1 + 2)/3 = 0 \)).

For multiplication, we have
- \( 2 \cdot 2 = \text{remainder of } (2 \times 2)/3 = 1; \)
- \( 1 \cdot 2 = \text{remainder of } (1 \times 2)/3 = 2, \text{ etc.} \)

Thus, we can construct the following table for multiplication:

\[
\begin{array}{ccc}
\cdot & 0 & 1 & 2 \\
0 & 0 & 0 & 0 \\
1 & 0 & 1 & 2 \\
2 & 0 & 2 & 1 \\
\end{array}
\]

- The multiplicative identity is 1 (i.e. \( \alpha \cdot 1 = \alpha \)).
- The multiplicative inverse of 1\((1^{-1})\) is 1 (i.e \( 1 \cdot 1^{-1} = \text{remainder of } (1 \times 1)/3 = 1 \)).
  - The multiplicative inverse of 2\((2^{-1})\) is 2 (i.e. \( 2 \cdot 2^{-1} = \text{remainder of } (2 \times 2)/3 = 1(\text{identity}) \)).

Thus, with the foregoing definitions of “+” and “·”, \( I_3 \) is a scalar field.

**Question:** Is \( I_4 \) with the foregoing definition of “+” and “·” a scalar field? What about \( N > 4 \)?
Definition 4.2. A vector space $V$ (over a scalar field $F$) is a set of elements called vectors satisfying the following conditions:

(i) For every pair of vectors $x$ and $y$ in $V$, there corresponds a unique vector $x + y$ in $V$ (addition of $x$ and $y$) (the closure property). The operation “$+$” (vector addition) has the following properties:
   (a) $x + y = y + x$ (commutativity);
   (b) $x + (y + z) = (x + y) + z$ (associativity) for all $x, y$ and $z \in V$;
   (c) there is a unique vector $0$ (the zero vector or identity element) such that $x + 0 = x$ for all $x \in V$;
   (d) for every vector $x$ in $V$, there is a unique vector $-x$ in $V$ (the additive inverse of $x$) such that $x + (-x) = 0$.

(ii) Scalar multiplication is defined, i.e. to every scalar $\alpha$ in a scalar field $F$ and a vector $x \in V$, there is a vector $\alpha x \in V$ called the product of $\alpha$ and $x$. It has the following properties:
   (a) $\alpha(\beta x) = (\alpha \cdot \beta)x$;
   (b) $1x = x$ for every $x \in V$ (“1” is the identity element in $F$);
   (c) $\alpha(x + y) = \alpha x + \alpha y$;
   (d) $(\alpha + \beta)x = \alpha x + \beta x$, where $\alpha \cdot \beta$ denotes multiplication between scalars $\alpha$ and $\beta$.

Remark: Note the difference in meaning between the “$+$” signs in $x + y$ and $\alpha + \beta$.

Examples of vectors spaces:

- $R^n$ — $n$-dimensional real coordinate space. The vectors are ordered $n$-tuple of real numbers (i.e. $x = (x_1, \ldots, x_n), x_i \in R$), and the scalar field $F$ is the set of all real numbers. The addition of two vectors $x = (x_1, \ldots, x_n)$ and $x' = (x'_1, \ldots, x'_n)$ is defined by
  \[ x + x' = (x_1 + x'_1, \ldots, x_n + x'_n), \]
  where $x_i + x'_i$ corresponds to ordinary addition of two real numbers. The multiplication of $x$ by a scalar $\alpha$ is defined by $\alpha x = (\alpha x_1, \ldots, \alpha x_n)$, where $\alpha x_i$ corresponds to ordinary multiplication of two real numbers.

- $C^n$ — $n$-dimensional complex coordinate space. The vectors are ordered $n$-tuple of complex numbers, and the scalar field $F$ is the set of all complex numbers. Vector addition corresponds to the componentwise addition of two complex numbers, and multiplication of a vector by a scalar corresponds to componentwise multiplication by a scalar.

- $P^n$-space of real or complex polynomials in $t$ with degree $\leq n - 1$. The vectors are polynomials with real or complex coefficients in a variable $t$. An element of $P^n$ has the form:
  \[ x = \sum_{i=0}^{n-1} \eta_i t^i, \quad \eta_i \text{ is a real or complex number}. \]
  Vector addition: $x = \sum_{i=0}^{n-1} \eta_i t^i, y = \sum_{i=0}^{n-1} \eta'_i t^i$;
\[ x + y = \sum_{i=0}^{n-1} (\eta_i + \eta'_i) t^i. \]

Zero vector: \( x = 0 \) (a polynomial with all coefficients \( \eta_i = 0 \)).

- The unit interval \( I = \{ x \in R : 0 \leq x \leq 1 \} \) over the real scalar field is not a vector space. (The closure property is not satisfied).

**Definition 4.3.** A set of vectors \( \{x_1, \ldots, x_n\} \) is said to be **linearly independent**, if

\[ \sum_{i=1}^{n} \alpha_i x_i = 0 \implies \alpha_i = 0 \text{ for all } i = 1, \ldots, n. \]

It is **linearly dependent**, if there exists a set of scalars \( \alpha_1, \ldots, \alpha_n \), not all zero, such that

\[ \sum_{i=1}^{n} \alpha_i x_i = 0. \]

**Example:** In \( \mathbb{R}^3 \), the vectors \( x_1 = (0, 0, 1), x_2 = (0, 1, 0) \) and \( x_3 = (2, 0, 0) \) are linearly independent. If \( x_4 = (1, 1, 0) \) is added to the set, the set is linearly dependent, since \( x_4 = x_2 + x_3/2 \).

**Assertion:** The set of vectors \( \{x_1, \ldots, x_m\} \) is linearly dependent if and only if one of them, say \( x_j \), \( 1 \leq j \leq m \), can be expressed as a linear combination of others.

**Proof:** Linear dependence \( \iff \) there exists a set of scalars \( \alpha_i, i = 1, \ldots, m \), not all zero, such that \( \sum_{i=1}^{m} \alpha_i x_i = 0 \). Solve for \( x_j \), whose coefficient \( \alpha_j \neq 0 \), i.e. \( x_j = \sum_{i=1, i \neq j}^{m} (\alpha_i \cdot \alpha_j^{-1}) x_i \). Conversely, if \( x_j = \sum_{i=1, i \neq j}^{m} \alpha_i x_i \), then \( x_j - \sum_{i=1, i \neq j}^{m} \alpha_i x_i = 0 \iff \text{linear dependence.} \|

**Theorem 4.1.** Let \( x_i = (x_{i1}, \ldots, x_{in}), i = 1, \ldots, n \) be vectors in \( \mathbb{R}^n \) or \( \mathbb{C}^n \). The set of vectors \( \{x_1, \ldots, x_m\} \) is linearly dependent if and only if \( \det[A] = 0 \), where \( [A] \) is the \( n \times n \) matrix defined by

\[
[A] = \begin{bmatrix}
x_{11} & x_{21} & \cdots & x_{n1} \\
\vdots & \vdots & \ddots & \vdots \\
x_{1n} & x_{2n} & \cdots & x_{nn}
\end{bmatrix}.
\]

**Proof:** Necessity: Linear dependence \( \iff \) there exist \( \alpha_i, i = 1, \ldots, n \), not all zero, such that \( \sum_{i=1}^{n} \alpha_i x_i = 0 \). This can be rewritten as \( [A] \alpha = 0 \), where \( \alpha = (\alpha_1, \ldots, \alpha_n)^T \) and \( (\cdot)^T \) denotes transpose. For nontrivial solution, \( \det[A] = 0 \).

Sufficiency: If \( \det[A] = 0 \), there exists a nonzero \( \alpha \) such that \( [A] \alpha = 0 \iff \text{linear dependence.} \|

**Example:** The set of vectors \( \{(1, 1, 0), (1, 0, 0), (1, 2, 1)\} \) in \( \mathbb{R}^3 \) is linearly independent, since

\[
\det \begin{bmatrix}
1 & 1 & 1 \\
1 & 0 & 2 \\
0 & 0 & 1
\end{bmatrix} = -1 \neq 0.
\]
**Definition 4.4:** A set $\mathcal{B}$ of vectors in a vector space $V$ is a *basis* (coordinate system) of $V$, if (i) $\mathcal{B}$ is linearly independent, (ii) every vector in $V$ can be expressed as a linear combination of vectors in $\mathcal{B}$. The elements of $\mathcal{B}$ are called *basis vectors*.

The set of vectors $\{v_1 = (1,0,\ldots,0), v_2 = (0,1,0,\ldots,0), \ldots, v_n = (0,\ldots,0,1)\}$ forms a *standard* or *natural basis* for $\mathbb{R}^n$.

**Definition 4.5:** A set of vectors $\mathcal{E}$ is said to *span* a vector space $V$, if any vector in $V$ can be expressed as a linear combination of vectors of $\mathcal{E}$.

**Note:** $\mathcal{E}$ needs not be linearly independent, but it must contain a linearly independent set of vectors.

**Example:** The set of vectors $\{(1,0),(0,1),(1,1)\}$ in $\mathbb{R}^2$ spans $\mathbb{R}^2$, but they do not constitute a basis for $\mathbb{R}^2$. However, $\{(1,0),(0,1)\}$ or $\{(1,0),(1,1)\}$ constitutes a basis for $\mathbb{R}^2$.

**Theorem 4.2.** If $\{v_1, \ldots, v_n\}$ forms a basis for the vector space $V$, then every vector $x$ in $V$ has a *unique* representation of the form:

$$x = \sum_{i=1}^{n} \eta_i v_i,$$

where $\eta_i$ the component of $x$ with respect to basis vector $v_i$.

**Proof:** Suppose there exist two representations for $x$, i.e.

$$x = \sum_{i=1}^{n} \eta_i v_i, \quad x = \sum_{i=1}^{n} \eta'_i v_i.$$ 

Thus,

$$\sum_{i=1}^{n} (\eta_i - \eta'_i) v_i = 0.$$ 

Since $\{v_1, \ldots, v_n\}$ is linearly independent, therefore $\eta_i = \eta'_i, i = 1, \ldots, n$. ||

**Definition 4.6.** The *dimension* of a vector space $V$ is the number of elements in any basis of $V$.

**Note:** (i) $V = \{0\}$ has zero-dimension, since the empty set of vectors is a basis for $V$.

(ii) Every set of $(n+1)$ vectors $\{x_1, \ldots, x_{n+1}\}$ in a $n$-dimensional vector space $V$ is *linearly dependent*.

Since $\dim V = \text{number of elements in any basis of } V$. By definition of a basis, every vector $x_i$ in the set of $(n+1)$ vectors can be expressed as a linear combination of $n$ basis vectors $v_1, \ldots, v_n$ in $\mathcal{B}$. Consider

$$\sum_{i=1}^{n+1} \eta_i x_i = \sum_{i=1}^{n+1} \eta_i \left( \sum_{j=1}^{n} \xi_{ij} v_j \right) = \sum_{j=1}^{n} \gamma_j v_j = 0,$$

where $\gamma_j = \sum_{i=1}^{n+1} \eta_i \xi_{ij}$. 

32
By linear independence of \( \{v_1, \ldots, v_n\} \), \( \sum_{j=1}^{n} \gamma_j = 0, j = 1, \ldots, n \) or
\[
\sum_{i=1}^{n+1} \eta_i \xi_{ij} = 0, j = 1, \ldots, n,
\]
which is a set of linear algebraic equations with \( n + 1 \) unknowns \( \eta_1, \ldots, \eta_{n+1} \). Evidently, there exist \( \eta_i, i = 1, \ldots, n+1 \), not all zero such that the above equations are satisfied. Hence the set \( \{x_1, \ldots, x_{n+1}\} \) is linearly dependent.

(iii) Any linearly independent set of \( n \) vectors in a \( n \)-dimensional vector space is a basis for \( V \).

**Definition 4.7.** A nonempty set \( M \) of a vector space \( V \) is a subspace, if for any pair of vectors \( x, x' \in M \), every linear combination
\[
\alpha x + \beta x' \in M.
\]

**Note:** A subspace is a vector space. If \( M \) is a proper subset of \( V \), then \( M \) is a proper subspace. A subspace \( M \) in a \( n \)-dimensional vector space \( V \) is a vector space of dimension \( \leq n \).

**Examples:**
(i) The set \( \{0\} \) is a subspace of \( V \).
(ii) The whole space is a subspace of \( V \).
(iii) In \( \mathbb{R}^n \), lines and planes passing through the origin are subspaces of \( \mathbb{R}^n \).
(iv) A ray \( R = \{x \in \mathbb{R}^n : x = \alpha v, \alpha \geq 0\} \), where \( v \) is a nonzero vector in \( \mathbb{R}^n \), is not a subspace.
(v) The intersection of a collection of subspaces \( M_i, i = 1, \ldots, m \) is a subspace.

**Proof:** Let \( M = \bigcap_i M_i \). Since any \( M_i \) contains the zero vector \( 0 \), so does \( M \). Therefore \( M \) is a nonempty set. If \( x, x' \in M \) (i.e. they belong to all \( M_i \)), then \( \alpha x + \beta x' \in \) all \( M_i \). Hence \( M \) is a subspace. ||

(vi) The union of any collection of subspaces is not necessarily a vector space. e.g. in \( \mathbb{R}^2 \), let \( M_1 = \{(\eta_1, \eta_2) : \eta_1 = \eta_2\} \) and \( M_2 = \{(\eta_1, \eta_2) : \eta_2 = 2\eta_1\} \) as represented by lines in the \((\eta_1, \eta_2)\)-plane:

**Definition 4.8.** Let \( V_1 \) and \( V_2 \) be vectors spaces over the same scalar field \( F \). A linear transformation (or operator) \( A \) on \( V_1 \) into \( V_2 \) (Notation: \( A : V_1 \rightarrow V_2 \)) is a correspondence that assigns to every vector \( x \) in \( V_1 \) a vector \( Ax \) in \( V_2 \) such that
\[
A(\alpha x + \beta y) = \alpha Ax + \beta Ay \text{ for all } x, y \in V_1 \text{ and all scalars } \alpha, \beta.
\]

**Note:** If \( A \) is a linear transformation, then \( A0 = 0 \), i.e. it maps the zero vector in \( V_1 \) to the zero vector in \( V_2 \).

**Remarks:**
(1) \( V_1 \) and \( V_2 \) need not be the same, they may have different dimensions.
(2) \( V_1 \) is called the domain of \( A \). The range of \( A \) is the set \( \mathcal{R}(A) = \{y \in V_2 : y = Ax, x \in V_1\} \). In general, \( \mathcal{R}(A) \subseteq V_2 \) (i.e. \( \mathcal{R}(A) \) is a subset of \( V_2 \)). When \( \mathcal{R}(A) = V_2 \), we say that \( A \) maps \( V_1 \) onto \( V_2 \).
Examples:
(1) Consider the vector space $\mathbb{R}^2$ with vectors $\mathbf{x} = (\eta_1, \eta_2)$. Each of the following $A$’s is a linear transformation on $\mathbb{R}^2$ into $\mathbb{R}^2$:

(a) $A((\eta_1, \eta_2)) = (\alpha \eta_1, \alpha \eta_2)$, $\alpha$ is a real number. The effect of $A$ is to multiply each vector in $\mathbb{R}^2$ by a scalar $\alpha$.

(b) $A((\eta_1, \eta_2)) = (\eta_2, \eta_1)$, $A$ reflects $\mathbb{R}^2$ about the diagonal line $\eta_1 = \eta_2$.

(c) $A((\eta_1, \eta_2)) = (\eta_1, 0)$, $A$ projects $\mathbb{R}^2$ onto the $\eta_1$ – axis.

(2) Let $V_1$ be the space of all integrable real-valued functions defined on $[0, 1]$. Let $A$ be defined by

$$A(x) = \int_0^1 x(t) dt.$$ 

$A$ is a linear transformation on $V_1$ into $\mathbb{R}$. Here, $V_1$ is infinite dimensional.

If we change $V_1$ to the space of real-valued functions spanned by $v_i = \sin(i\pi t), i = 1, \ldots, n, 0 \leq t \leq 1$, then $V_1$ is a $n$-dimensional vector space. In fact, $\{v_i, i = 1, \ldots, n\}$ is a basis for $V_1$.

(3) The set of all linear transformations on a vector space $V_1$ into vector space $V_2$ (with the same scalar field) forms a vector space $V$. Here, each “vector” or element in $V$ is a linear transformation. We define:

(i) vector addition: $A + B$
(A + B)x = Ax + Bx for all x ∈ V₁.

(ii) Additive identity: O (the zero transformation)

Ox = 0 (zero vector in V₁ for all x in V₁).

(iii) additive inverse corresponding to A : −A defined by

(−A)x = −Ax (a vector in V₂) for all x in V₁.

(iv) scalar multiplication: αA defined by

(αA)(x) = α(Ax) for all x ∈ V₁.

With the foregoing definitions, {V, +, ·, F} is a vector space, where F is the same scalar field as that of V₁.

Remark: The vector space of linear transformations is useful in applications. For example, for discrete-time systems described by linear difference equations of the form:

x(k + 1) = Φx(k), where Φ is a linear transformation on V₁. We may consider a perturbed linear system of the form:

x(k + 1) = (Φ + ΔΦ)x(k),

where ΔΦ corresponds to a perturbation of Φ. Here, we may consider Φ and ΔΦ as elements of a vector space consisting of linear transformations on V₁ into itself.

8. Products (composition) of Linear Transformations: Consider two linear transformations A and B on V into V.

Definition 4.8. The product P of A and B (P = AB) is defined by Px = A(Bx).

Remarks: (1) P is a composite of two transformations. x is first acted on by B, and then Bx is acted on by A.
(2) The motion of a discrete-time linear system can be considered as a sequence of linear transformations on a state space $V$:

$$x(k+2) = \Phi(k+1)x(k+1) = \Phi(k+1)\Phi(k)x(k) = \Phi(k+1)\Phi(k)\Phi(k-1)x(k-1) = \ldots$$

(3) In general, $AB \neq BA$. If $AB = BA$, then we say $A$ and $B$ commute.

Example: Consider the following linear transformations $A : \mathcal{P}^n \to \mathcal{P}^{n+1}$, and $B : \mathcal{P}^n \to \mathcal{P}^{n-1}$. Let

$$Ax = tx, \quad Bx = \frac{d}{dt}(x), \quad x \text{ is a polynomial in } t \text{ of degree } < n.$$ 

Thus, $AB \neq BA$.

9. Polynomials of Linear Transformations:

The associative law of multiplication enable us to write three or more factors without any parenthesis, in particular,

$$A \circ \circ \circ A = A^m, \quad \text{m times}$$

Thus,

$$A^{n+m} = A^n A^m, \quad (A^n)^m = A^{nm}, \quad A^1 = A, \quad A^0 = I \text{ (identity)}.$$ 

It is meaningful to form polynomials of $A$:

$$p(A) = \alpha_0 I + \alpha_1 A + \cdots + \alpha_n A^n.$$ 

Note: Due to noncommutativity,

$$(A + B)^2 \neq A^2 + 2AB + B^2,$$

(Equality holds if and only if $AB = BA$).