Lecture Topics:

a) Vector space, subspace
b) Matrices: sum, product, invertible matrix
c) Column and row spaces, right and left nullspaces

Vector space: A vector space is constituted of a pair \( (V, S) \) where \( V \) is a set of vectors and \( S \) a field of scalars. Throughout this course \( S \) will be either \( \mathbb{R} \) (the field of real numbers) or \( \mathbb{C} \) (the field of complex numbers). The pair \( (V, S) \) is endowed with two operations: the addition of two vectors, and the multiplication of a vector by a scalar. In the following a vector \( \mathbf{x} \in V \) will be denoted with in bold font and a scalar \( a \in S \) in regular font.

Vector addition: For \( \mathbf{x}, \mathbf{y} \in V \), the addition \( \mathbf{x} + \mathbf{y} \in V \) must satisfy the following axioms.

(i) Associativity:

\[ \mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z} \]

for all \( \mathbf{x}, \mathbf{y}, \mathbf{z} \in V \).

(ii) Commutativity:

\[ \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} . \]

(iii) There exists a zero vector \( \mathbf{0} \) such that

\[ \mathbf{x} + \mathbf{0} = \mathbf{x} . \]

(iv) For every vector \( \mathbf{x} \in V \), there exists an additive inverse \( -\mathbf{x} \in V \) such that \( \mathbf{x} - \mathbf{x} = \mathbf{0} \).

Scalar multiplication: For \( a \in S \) and \( \mathbf{x} \in V \), the multiplication \( a\mathbf{x} \in V \) must satisfy the following axioms.

(i) Associativity: \( a(b\mathbf{x}) = (ab)\mathbf{x} \) for all \( a, b \in S \) and \( \mathbf{x} \in V \).

(ii) Distributivity of the scalar multiplication with respect to the vector addition:

\[ a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y} \]

for all \( a \in S \) and \( \mathbf{x}, \mathbf{y} \in V \).
(iii) Distributivity of the multiplication with respect to the scalar addition:

$$(a + b)x = ax + bx$$

for all $a, b \in S$ and $x \in V$.

(iv) There exists a unit scalar $1 \in S$ such that $1x = x$ for all $x \in V$.

**Example 1:** $V = \mathbb{R}^n$, $S = \mathbb{R}$. In this case, an arbitrary vector

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}$$

is an $n$-tuple of real entries $x_i$ with $1 \leq i \leq n$. Then if

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$$

and $a \in \mathbb{R}$, we have

$$x + y = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_i + y_i \\ \vdots \\ x_n + y_n \end{bmatrix} \quad \text{and} \quad ax = \begin{bmatrix} ax_1 \\ \vdots \\ ax_i \\ \vdots \\ ax_n \end{bmatrix}.$$ 

The zero-vector and additive inverse $-x$ are given respectively by

$$0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad -x = \begin{bmatrix} -x_1 \\ \vdots \\ -x_i \\ \vdots \\ -x_n \end{bmatrix}.$$  

When $n = 3$, $\mathbb{R}^3$ is the set of real vectors in 3 dimensions, and the addition $x + y$ and scalar multiplication $ax$ can be represented geometrically as shown in Fig. 1 below. Note that the vector $ax$ is colinear with $x$. 

2
Figure 1: Vector addition and multiplication of a vector by a scalar in $\mathbb{R}^3$.

Also, an arbitrary vector of $\mathbb{R}^n$ can be represented as

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$$  \hspace{1cm} (1)

where for $1 \leq i \leq n$, the entries of basis vector are all zero, except for its $i$-th entry which equals one. The basis $\{\mathbf{e}_i, 1 \leq i \leq n\}$ is orthonormal since

$$\mathbf{e}_i^T \mathbf{e}_j = \delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}.$$

**Example 2:** Let $S = \mathbb{R}$, and let $V$ be the space of piecewise continuous real functions over interval $[0, T]$. A function $f(t), 0 \leq t \leq T$ belonging to this space is shown in Fig. 2.

Figure 2: Piecewise continuous function over interval $[0, T]$.

The sum of two such functions is given by

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

for $0 \leq t \leq T$, which is also piecewise continuous. The scalar multiplication $a f(t)$ corresponds to scaling the function $f(t)$ by $a$.  

3
The inner product of two functions is given by

\[ \langle f, g \rangle = \int_0^T f(t) g(t) \, dt. \]

Then, if we consider

\[ c_k(t) = \sqrt{\frac{2}{T}} \cos(k \omega_0 t), \quad s_k(t) = \sqrt{\frac{2}{T}} \sin(k \omega_0 t), \]

with \( k \) integer, where \( \omega_0 = 2\pi/T \) is the fundamental frequency associated to the interval \([0, T]\), the functions \( \{c_k(t), k \geq 0\} \) and \( \{s_k(t), k \geq 1\} \) are orthonormal since

\[
\langle c_k, c_\ell \rangle = \langle s_k, s_\ell \rangle = \delta_{k\ell} \quad \langle c_k, s_\ell \rangle = 0
\]

for all \( k, \ell \). Furthermore, an arbitrary piecewise continuous function \( f(t) \) over \([0, T]\) can be expressed in Fourier series as

\[ f(t) = \sum_{k=0}^{\infty} a_k c_k(t) + \sum_{k=1}^{\infty} b_k s_k(t) \quad (2) \]

where the Fourier coefficients \( \{a_k, k \geq 0\} \) and \( \{b_k, k \geq 1\} \) are given by

\[ a_k = \int_0^T f(t) c_k(t) \, dt, \quad b_k = \int_0^T f(t) s_k(t) \, dt. \]

**Subspace:** If \( V \) is a vector space and \( W \) is a set of vectors from \( V \), \( W \) is a subspace of \( V \) if it is closed under the operation of vector addition and scalar multiplication, i.e., if \( \mathbf{x} \) and \( \mathbf{y} \) are arbitrary vectors of \( W \) and \( a, b \in S \), we have \( a\mathbf{x} + b\mathbf{y} \in W \).

**Example 1:** Let \( V = \mathbb{R}^3 \), and let \( W \) be the set of vectors belonging to the \( x_1-x_2 \) plane, i.e.,

\[ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}. \]

This space is depicted in Fig. 3 below.

\( W \) is a subspace of \( \mathbb{R}^3 \) since for

\[ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ 0 \end{bmatrix} \in W \]

we have

\[ a\mathbf{x} + b\mathbf{y} = \begin{bmatrix} ax_1 + by_1 \\ ax_2 + by_2 \\ 0 \end{bmatrix}, \]
where the vector $a\mathbf{x} + b\mathbf{y}$ belongs to the $x_1$-$x_2$ plane, since its last entry is zero.

**Example 2:** Let $V = \mathbb{R}^2$, and let $W$ be the set of vectors $\mathbf{x}$ colinear with the two axes, i.e.,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

with either $x_1$ or $x_2$ zero. $W$ is shown in Fig. 4 below. Then $W$ is not a subspace since for

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in W,$$

the vector

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \not\in W.$$

**Example 3:** Let $V = \mathbb{R}^2$, and let $W$ be the set of vectors $\mathbf{x}$ in the first quadrant, i.e.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
with \( x_1 \geq 0, x_2 \geq 0 \). This set is shown in Fig. 5 below. \( W \) is not a subspace since
\[
\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in W \quad \text{but} \quad -\mathbf{x} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \not\in W.
\]

![Figure 5: First quadrant of \( \mathbb{R}^2 \).](image)

**Linear dependence/independence:** A set of vectors \( \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \) is said to be linearly dependent if we can find a set of scalars \( a_i, 1 \leq i \leq k \), which not all zero, such that
\[
\sum_{i=1}^{k} a_i \mathbf{v}_i = 0. \tag{3}
\]
If no such scalars exist, i.e., if
\[
\sum_{i=1}^{k} a_i \mathbf{v}_i = 0 \Rightarrow a_1 = a_2 \ldots = a_k = 0,
\]
the vectors \( \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \) are linearly independent.

**Example 1:** The vectors
\[
\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}
\]
are linearly dependent since
\[
\mathbf{v}_1 + \mathbf{v}_3 - 2\mathbf{v}_2 = 0.
\]

**Example 2:** The vectors
\[
\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}
\]
are linearly independent since

$$a_1v_1 + a_2v_2 = \begin{bmatrix} a_1 + 2a_2 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has

$$a_2 = a_1 = 0$$

as its unique solution.

**Spanning set:** A set of vectors \( \{v_1, \ldots, v_k\} \) spans a vector space \( V \) if for an arbitrary vector \( x \) of \( V \) we can find some scalars \( a_i \), \( 1 \leq i \leq k \) such that

$$x = \sum_{i=1}^{k} a_i v_i ,$$  

(4)

i.e., \( x \) is a linear combination of the vectors \( v_i \).

**Example:**

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} , \quad v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad v_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

form a spanning set of \( \mathbb{R}^2 \), but it is not a minimum spanning set since \( \{v_1, v_2\} \) or \( \{v_1, v_3\} \) are also spanning.

**Matrix:** A matrix \( A \) is an \( m \text{ by } n \) array of real or complex numbers \( a_{ij} \):

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} ,$$

where \( a_{ij} \) represents the element of row \( i \) and column \( j \) with \( 1 \leq i \leq m \) and \( 1 \leq j \leq n \).

The special case of \( n = 1 \) is a column vector, and \( n = 1 \) is a row vector. Note that \( A \) can be written column-wise as

$$A = \begin{bmatrix} c_1 & \cdots & c_j & \cdots & c_n \end{bmatrix}$$

where

$$c_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{bmatrix}$$
denotes the $j$-th column of $A$. Similarly, $A$ can be written row-wise as

$$A = \begin{bmatrix} r_1^T \\ \vdots \\ r_i^T \\ \vdots \\ r_m^T \end{bmatrix}$$

where

$$r_i^T = [ a_{i1} \ldots a_{ij} \ldots a_{in} ]$$

denotes the $i$-th row of $A$.

If both $A$ and $B$ are $m \times n$ matrices, their sum $C = A + B$ is an $m \times n$ matrix with elements $c_{ij} = a_{ij} + b_{ij}$. It is easy to check that the sum is associative and commutative, i.e.

$$(A + B) + C = A + (B + C) \quad A + B = B + A.$$ 

If $A$ and $B$ are $m \times q$ and $q \times n$ matrices, respectively, the product $C = AB$ is an $m \times n$ matrix with elements

$$c_{ij} = \sum_{k=1}^{q} a_{ik} b_{kj}, \quad (5)$$

e.i., the elements of the $i$-th row of $A$ are multiplied term by term with those of the $j$-th column of $B$ and the terms summed.

**Example:** Let

$$A = \begin{bmatrix} 1 & 3 & 7 \\ -1 & 0 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}.$$ 

Then

$$AB = C = \begin{bmatrix} 4 & 9 \\ -1 & 3 \end{bmatrix},$$

where the entries of $C$ are obtained by using (5). For example

$$c_{11} = 1.1 + 3.1 + 7.0 = 4.$$ 

Note that in order to be able to compute the product of $A$ and $B$, the number of columns of $A$ must be equal to the number of rows of $B$. The product is associative: $(AB)C = A(BC)$, but not necessarily commutative, i.e., in general $AB \neq BA$. To see this, consider

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$ 

Then

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad BA = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$ 

8
so that $AB \neq BA$. Also, the sum and product are distributive, i.e., $(A + B)C = AC + BC$
and $C(A + B) = CA + CB$.
A square matrix is an $m \times n$ matrix with $m = n$. The identity matrix $I_n$ is an $n \times n$ matrix
with ones on the diagonal and zeros elsewhere, i.e.,

$$a_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases}$$

The identity has the property $AI_n = I_nA = A$ for any $n \times n$ matrix $A$.
An $n \times n$ matrix $A$ is invertible if there exists an $n \times n$ matrix $A^{-1}$ such that $AA^{-1} = A^{-1}A = I_n$. We shall see later that if $A$ is invertible then $A^{-1}$ is unique; $A^{-1}$ is called the
inverse of $A$.

**Theorem:** If $A$ is invertible, then $Ax = b$ has a unique solution $x$ for each choice of $b$, and $y^T A = c^T$ has a unique solution $y$ for each $c$.

**Proof:** If $x$ satisfies $Ax = b$, pre-multiplying this equation by $A^{-1}$ yields $A^{-1}Ax = A^{-1}b$, or equivalently

$$I_n x = x = A^{-1}b.$$ 

Since $A^{-1}b$ is uniquely specified, the solution is unique. Similarly post-multiplying $y^T A = c^T$ by $A^{-1}$ yields $y^T = c^T A^{-1}$. \quad \Box

**Matrix transpose:** The transpose of an arbitrary $m \times n$ matrix $A = (a_{ij})$ is the $n \times m$
matrix $A^T = (b_{ij})$ with $b_{ij} = a_{ji}$. Under transposition the rows of $A$ become the columns of $A^T$, and vice-versa.

**Example:** If

$$A = \begin{bmatrix} 1 & 0 & 1 \\ -7 & 3 & 2 \end{bmatrix} \quad \text{then} \quad A^T = \begin{bmatrix} 1 & -7 \\ 0 & 3 \\ 1 & 2 \end{bmatrix}.$$ 

To an arbitrary $m \times n$ matrix $A$, we can associate four fundamental spaces.

The **column space** $\mathcal{R}(A)$ is the subspace of $\mathbb{R}^n$ spanned by the columns $c_j$ of $A$, i.e., it is the set of vectors $b \in \mathbb{R}^n$ that can be expressed as $b = \sum_{j=1}^{n} x_j c_j$ for some real numbers $x_j$, $1 \leq j \leq n$. In matrix form this gives $b = Ax$ with

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$ 

so

$$\mathcal{R}(A) = \{ b \in \mathbb{R}^n : b = Ax \text{ for some } x \in \mathbb{R}^n \}.$$ 

The **row space** $R(A^T)$ is the space spanned by the rows $r_i$ of $A$, or equivalently by the columns of $A^T$. It is the set of vectors $c \in \mathbb{R}^n$ such that $c^T = \sum_{i=1}^{m} y_i r_i$, or equivalently
\( c^T = y^T A \) for some \( y \in \mathbb{R}^m \). Thus

\[
\mathcal{R}(A^T) = \{ c \in \mathbb{R}^n : c^T = y^T A \text{ for some } y \in \mathbb{R}^m \} = \{ c \in \mathbb{R}^n : c = A^T y \text{ for some } y \in \mathbb{R}^m \}.
\]

The **right null space** \( N(A) \) is the set of solutions \( x \) of the homogeneous equation \( Ax = 0 \), or equivalently the set of dependence relations existing between the columns of \( A \), i.e.

\[
N(A) = \{ x \in \mathbb{R}^n : Ax = 0 \}.
\]

The **left null space** \( N(A^T) \) is the set of solutions \( y \in \mathbb{R}^m \) of \( y^T A = 0^T \), i.e.,

\[
N(A^T) = \{ y \in \mathbb{R}^m : y^T A = 0^T \} = \{ y \in \mathbb{R}^m : A^T y = 0 \}.
\]

**Example:** If

\[
A = \begin{bmatrix}
1 & 2 & 1 \\
0 & 1 & 1
\end{bmatrix},
\]

the right null space \( N(A) \) is obtained by solving

\[
\begin{bmatrix}
1 & 2 & 1 \\
0 & 1 & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix} = \begin{bmatrix}
0 \\
0
\end{bmatrix}.
\]

This gives

\[
x_1 = x_3 = -x_2,
\]

so that \( N(A) \) is the space spanned by

\[
\begin{bmatrix}
1 \\
-1 \\
1
\end{bmatrix}.
\]