Unit 2D: Vector spaces and Dot Product

Prof.Dr.P.M.Bajracharya

School of Mathematical Sciences
T.U., Kirtipur

March 2, 2025

Summary

1. **Vector spaces**
   - Subspaces of a vector space
   - Examples of subspaces and not-subspaces in two dimensions
   - Subspaces in $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, and beyond
2. **Dot product**

Vector spaces

Vector spaces

A nonempty set V is a **vector space** over $\mathbb{R}$ if for every $x, y, z \in V$ and $a, b \in \mathbb{R}$

I. $x + y \in V$ with the properties:

1. $x + y = y + x$
2. $(x + y) + z = y + (x + z)$
3. $\exists\ 0 \in V$ such that $x + 0 = x$
4. $\exists\ -x \in V$ such that $x + (-x) = 0$

II $ax \in V$ with the properties:

1. $a(x + y) = ax + ay$
2. $(a + b)x = ax + bx$
3. $a(bx) = (ab)x$
4. $1x = x$.

Elements of a vector space are called **vectors**.

Note that

- The element 0 in Axiom 1 (c) is called the **zero vector**.
- The element $-x$ in Axiom 1 (d) is called the **negative** of x.

Theorem

1. Let x be a vector in $\mathbb{R}^n$. Then the vector $-x$ which satisfies property $x + (-x) = 0$ is unique.
2. Let x be a vector in $\mathbb{R}^n$. Then $(-1)x = -x$.

In many problems, a vector space consists of an appropriate subset of vectors from some larger vector space. In this case, only two of the ten vector space axioms need to be checked:

For every $x, y \in V$ and $a \in \mathbb{R}$

$$x + y \in V, \;\; ax \in V.$$

The rest are automatically satisfied.

Subspaces

Let S be a non-empty subset of a vector space V. The set S is called a **subspace** of V if S is closed under the same vector addition and scalar multiplication as V, that is, for all $x, y \in S$ and for all $a \in \mathbb{R}$

$$x + y \in S \text{ and } ax \in S.$$

Example

The L_2 ball $B(0;1) = \{x \in \mathbb{R}^2 : \ \|x\|_2 \leq 1\}$ is not a subspace.

Solution.

- It is not closed under addition: for example,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in B(0; 1),$$

  but their sum

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin B(0; 1).$$

Solution.

- It is not closed under addition: for example,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in B(0;1),$$

but their sum

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin B(0;1).$$

- Also, it is not closed under scalar multiplication: for example,

$$2\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \notin B(0;1).$$

The same is true for the L_2 sphere. Similarly, L_1 and L_∞ balls and spheres are not subspaces of $\mathbb{R}^2$.

Example

The **positive orthant** is not a subspace.

Example

The **positive orthant** is not a subspace.

Solution: It is not closed under scalar multiplication, for example: the vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is in positive orthant (first quadrant), but

$$-2\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

has negative entry and so, it is not in prositive orthant (first quadrant). $\square$

Remark

The positive orthant is, however, closed under multiplication of nonnegative scalars, a fact that is sometimes of interest.

Example

The positive simplex is not a subspace.

Example

The positive simplex is not a subspace.

Solution: It is not closed under the addition of two vectors, and it is nor closed under scalar multiplication. $\square$

Example

The positive simplex is not a subspace.

Solution: It is not closed under the addition of two vectors, and it is nor closed under scalar multiplication. $\square$

It is, however, closed under the special class of vector additions of the form $z = ax + by$, where x, y, z are vectors, and a, b are numbers such that $a, b \geq 0$ and $a + b = 1$. We will see later why this is of interest.

Example

The line L given by the equation $x_1 + x_2 = 1$ is not a subspace.

Solution. For example,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in L,$$

Example

The line L given by the equation $x_1 + x_2 = 1$ is not a subspace.

Solution. For example,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in L,$$

but

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin L.$$

$\square$

Examples of subspaces.

The set $\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$ is a “trivial” subspace of $\mathbb{R}^2$

since if we multiply it by any scalar or add it to itself, then the output is still the 0 vector.

We can see that $\mathbb{R}^2$ itself is a "trivial" subspace of $\mathbb{R}^2$ since $\mathbb{R}^2$ is a vector space and a subset of itself.

Problem

The line $x_2 = ax_1 + b$ (perhaps more familiar as $y = ax + b$) is not a subspace $\mathbb{R}^2$ for $b \neq 0$.

Solution

Let u, v be points on the line $x_2 = ax_1 + b$ and

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Solution

Let u, v be points on the line $x_2 = ax_1 + b$ and

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Then

$$u_2 = au_1 + b$$
$$v_2 = av_1 + b.$$

Solution ...

Consider the point

$$w = u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

Solution ...

Consider the point

$$w = u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

Then we have

$$u_2 + v_2 = a(u_1 + v_1) + 2b.$$

Solution ...

Consider the point

$$w = u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

Then we have

$$u_2 + v_2 = a(u_1 + v_1) + 2b.$$

This line is not on the line $x_2 = ax_1 + b$, the intercepts being different. $\square$

Problem

Every line through the origin in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$. That is, the line $y = ax$ is a subspace of $\mathbb{R}^2$.

Solution

Let u, v be two points on the line $y = ax$ and

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Then

$$u_2 = au_1$$
$$v_2 = av_1.$$

Solution ...

Consider the point

$$w = u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

Solution ...

Consider the point

$$w = u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

Then we have

$$u_2 + v_2 = a(u_1 + v_1).$$

This line is on the line $y = ax$. $\square$

Problem

The set of points on two lines through the origin in $\mathbb{R}^2$ is not a subspace of $\mathbb{R}^2$.

Proof.

Let $a, b \in \mathbb{R}$ such that $a \neq b$, and

$$\Omega_x = \{(x_1, x_2) : \; x_1 = ax_2)\},$$
$$\Omega_y = \{(y_1, y_2) : \; y_1 = by_2)\}.$$

Then $\Omega = \Omega_x \cup \Omega_y$ is not a subspace of $\mathbb{R}^2$.

Proof.

Let $a, b \in \mathbb{R}$ such that $a \neq b$, and

$$\Omega_x = \{(x_1, x_2) : \; x_1 = ax_2)\},$$
$$\Omega_y = \{(y_1, y_2) : \; y_1 = by_2)\}.$$

Then $\Omega = \Omega_x \cup \Omega_y$ is not a subspace of $\mathbb{R}^2$. In fact, the set Ω is not closed under addition of two vectors, since adding two vectors lying on two different lines results a vector that is not on either of those lines (except in the degenerate case when the two lines are the same). $\square$

Remark

The last two examples indicate that the union of two subspaces may not be a subspace.

Subspaces in $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, and beyond

There are two types of subspaces of $\mathbb{R}$:

$$\mathbb{R}, \quad \{0\}.$$

$\mathbb{R}$ is a one-dimensional subspace and $\{0\}$ a zero dimensional subspace.

There are two types of subspaces of $\mathbb{R}$:

$$\mathbb{R}, \quad \{0\}.$$

$\mathbb{R}$ is a one-dimensional subspace and $\{0\}$ a zero dimensional subspace.

Both of these are "trivial," in the sense that there is not too much interesting going on, and so one typically does not spend much time discussing the subspace aspects of $\mathbb{R}$, but it's good to understand such "extreme cases" in the definitions of vector spaces and subspaces.

The examples of the previous section suggest (correctly, as we discussed above) that there are three kinds of subspaces for $\mathbb{R}^2$:

The examples of the previous section suggest (correctly, as we discussed above) that there are three kinds of subspaces for $\mathbb{R}^2$:

- $\mathbb{R}^2$ itself is a two-dimensional subspace of $\mathbb{R}^2$.

The examples of the previous section suggest (correctly, as we discussed above) that there are three kinds of subspaces for $\mathbb{R}^2$:

- $\mathbb{R}^2$ itself is a two-dimensional subspace of $\mathbb{R}^2$.
- The singleton set $\{0\}$ is a subspace of dimension 0.

The examples of the previous section suggest (correctly, as we discussed above) that there are three kinds of subspaces for $\mathbb{R}^2$:

- $\mathbb{R}^2$ itself is a two-dimensional subspace of $\mathbb{R}^2$.
- The singleton set $\{0\}$ is a subspace of dimension 0.
- A line through the origin is a subspace of dimension 1, and it takes 1 number to specify a point on a line.

We see that there are four kinds of subspaces for $\mathbb{R}^3$:

We see that there are four kinds of subspaces for $\mathbb{R}^3$:

- $\mathbb{R}^3$ is a three-dimensional subspace of $\mathbb{R}^3$.

We see that there are four kinds of subspaces for $\mathbb{R}^3$:

- $\mathbb{R}^3$ is a three-dimensional subspace of $\mathbb{R}^3$.
- A plane through the origin is a two-dimensional subspace of $\mathbb{R}^3$.

We see that there are four kinds of subspaces for $\mathbb{R}^3$:

- $\mathbb{R}^3$ is a three-dimensional subspace of $\mathbb{R}^3$.
- A plane through the origin is a two-dimensional subspace of $\mathbb{R}^3$.
- A line through the origin is a one-dimensional subspace of $\mathbb{R}^3$.

We see that there are four kinds of subspaces for $\mathbb{R}^3$:

- $\mathbb{R}^3$ is a three-dimensional subspace of $\mathbb{R}^3$.
- A plane through the origin is a two-dimensional subspace of $\mathbb{R}^3$.
- A line through the origin is a one-dimensional subspace of $\mathbb{R}^3$.
- The set $\{0\}$ is a zero-dimensional subspace of $\mathbb{R}^3$.

Dot product

Dot product on $\mathbb{R}^2$

If $x = (x_1, x_2)$ and $y = (y_1, y_2))$ are points on the plane $\mathbb{R}^2$, then the **dot product** or **inner product** between those two vectors is

$$x \cdot y = \sum_{1}^{2} x_i y_i = x_1 y_1 + x_2 y_2.$$

The following relation establishes the relationship between the dot poduct and the L_2 norm:

$$\|x\|_2 = \left(\sum_1^2 x_i^2\right)^{1/2} = \sqrt{x \cdot x}.$$

We can easily generalize the dot product on $\mathbb{R}^2$ to that on $\mathbb{R}^n$.

Dot product on $\mathbb{R}^n$

If $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n))$ are points in $\mathbb{R}^n$, then the dot product or inner product between those two vectors is

$$x \cdot y = x_1 y_1 + ... + x_n y_n = \sum_{1}^{n} x_i y_i.$$

Key points about the dot product in data science:

- **Vector Representation:** To use the dot product, data points are often converted into vectors, allowing for calculations based on their direction and magnitude.

Key points about the dot product in data science:

- **Vector Representation:** To use the dot product, data points are often converted into vectors, allowing for calculations based on their direction and magnitude.
- **Interpretation of Result:** A high dot product value indicates a high degree of similarity between two vectors, while a low value suggests a low similarity.

Key points about the dot product in data science:

- **Vector Representation:** To use the dot product, data points are often converted into vectors, allowing for calculations based on their direction and magnitude.
- **Interpretation of Result:** A high dot product value indicates a high degree of similarity between two vectors, while a low value suggests a low similarity.
- **Normalization:** When using the dot product for similarity calculations, it is often necessary to normalize the vectors to ensure that the length of the vectors does not influence the result.

Applications of dot products in data science

One common application of dot products in data science is in calculating the **similarity between vectors**.

Applications of dot products in data science

One common application of dot products in data science is in calculating the **similarity between vectors**. In machine learning, it is often useful to compare the similarity of two vectors in order to classify or cluster data.

Applications of dot products in data science

One common application of dot products in data science is in calculating the **similarity between vectors**. In machine learning, it is often useful to compare the similarity of two vectors in order to classify or cluster data. The dot product can be used to measure the similarity between vectors by calculating the angle between them.

Dot products are also used to optimize linear regression models.

Dot products are also used to optimize linear regression models. Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables.

Dot products are also used to optimize linear regression models. Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables. By minimizing the sum of the squared errors between the predicted values and the actual values, linear regression models can be optimized using the dot product.

Dot products are also used to optimize linear regression models. Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables. By minimizing the sum of the squared errors between the predicted values and the actual values, linear regression models can be optimized using the dot product.

There are many other applications of dot products in data science.

Key applications of dot product in data science:

- **Cosine Similarity:** The most common application of the dot product is to calculate cosine similarity, which measures the angle between two vectors, providing a way to assess how similar two data points are, especially when dealing with high-dimensional data.

Key applications of dot product in data science:

- **Cosine Similarity:** The most common application of the dot product is to calculate cosine similarity, which measures the angle between two vectors, providing a way to assess how similar two data points are, especially when dealing with high-dimensional data.
  For example, if we have two vectors, A and B, the similarity between them is calculated as:

$$\text{Similarity}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$

Key applications of dot product in data science:

- **Cosine Similarity:** The most common application of the dot product is to calculate cosine similarity, which measures the angle between two vectors, providing a way to assess how similar two data points are, especially when dealing with high-dimensional data.
  For example, if we have two vectors, A and B, the similarity between them is calculated as:
  $$\text{Similarity}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$

- **Neural Networks:** In neural networks, the dot product is used to compute the weighted sum of inputs to a neuron during forward propagation,

- **Text Analysis:** By representing words as vectors (word embeddings), the dot product can be used to calculate the similarity between words, enabling tasks like sentiment analysis and topic modeling.

- **Text Analysis:** By representing words as vectors (word embeddings), the dot product can be used to calculate the similarity between words, enabling tasks like sentiment analysis and topic modeling.
- **Recommendation Systems:** By calculating the similarity between users or items based on their vector representations, the dot product can be used to recommend items that are likely to be preferred by a user.

- **Text Analysis:** By representing words as vectors (word embeddings), the dot product can be used to calculate the similarity between words, enabling tasks like sentiment analysis and topic modeling.
- **Recommendation Systems:** By calculating the similarity between users or items based on their vector representations, the dot product can be used to recommend items that are likely to be preferred by a user.
- **Image Similarity:** Similar to text analysis, images can be represented as vectors, and the dot product can be used to determine the similarity between images.

The following relation establishes the relationship between the dot poduct on $\mathbb{R}^n$ and the L_2 norm on $\mathbb{R}^n$:

$$\|x\|_2 = \left(\sum_1^n x_i^2 \right)^{1/2} = \sqrt{x \cdot x}.$$

The dot product allows us to define the perpendicularity (or orthogonality).

Orthogonality

Let x and y be two vectors in $\mathbb{R}^n$. We say that x is **perpendicular** (or **orthogonal**) to y, if

$$x \cdot y = 0.$$

Orthogonal compliment

Let x be vector in $\mathbb{R}^n$. The set of vectors perpendicular to a vector $x \in \mathbb{R}^n$ is denoted by $x^\perp$ and is defined by

$$x^\perp = \{y \in \mathbb{R}^n : \ x \cdot y = 0\}.$$

The set $x^\perp$ is called the **orthogonal compliment** of $\{x\}$.

Theorem

Given a vector $x \in \mathbb{R}^2$, such that $x \neq 0$, let $x^\perp$ be the set of vectors that are perpendicular to x. Then, $x^\perp$ is a subspace of $\mathbb{R}^2$.

Proof.

If $x = 0$, then $x^{\perp} = \mathbb{R}^2$, which we know is a subspace.

Proof.

If $x = 0$, then $x^{\perp} = \mathbb{R}^2$, which we know is a subspace. Suppose that $x \neq 0$. Let $u, v \in x^{\perp}$.

Proof.

If $x = 0$, then $x^\perp = \mathbb{R}^2$, which we know is a subspace. Suppose that $x \neq 0$. Let $u, v \in x^\perp$. If $x = (x_1, x_2), u = (u_1, u_2)$ and $v = (v_1, v_2)$, then we have

$$x \cdot (u + v) = x_1(u_1 + v_1) + x_2(u_2 + v_2)$$

Proof.

If $x = 0$, then $x^{\perp} = \mathbb{R}^2$, which we know is a subspace. Suppose that $x \neq 0$. Let $u, v \in x^{\perp}$. If $x = (x_1, x_2), u = (u_1, u_2)$ and $v = (v_1, v_2)$, then we have

$$\begin{aligned} x \cdot (u + v) &= x_1(u_1 + v_1) + x_2(u_2 + v_2) \\ &= (x_1u_1 + x_2u_2) + (x_1v_1 + x_2v_2) = 0 \end{aligned}$$

Proof.

If $x = 0$, then $x^\perp = \mathbb{R}^2$, which we know is a subspace. Suppose that $x \neq 0$. Let $u, v \in x^\perp$. If $x = (x_1, x_2), u = (u_1, u_2)$ and $v = (v_1, v_2)$, then we have

$$\begin{aligned} x \cdot (u + v) &= x_1(u_1 + v_1) + x_2(u_2 + v_2) \\ &= (x_1u_1 + x_2u_2) + (x_1v_1 + x_2v_2) = 0. \end{aligned}$$

Therefore,

$$u + v \in x^\perp.$$

Proof ...

Now, if $a \in \mathbb{R}$, then

$$x \cdot (au) = x_1(au_1) + x_2(au_2)$$

Proof ...

Now, if $a \in \mathbb{R}$, then

$$\begin{aligned} x \cdot (au) &= x_1(au_1) + x_2(au_2) \\ &= a(x_1u_1 + x_2u_2) = 0. \end{aligned}$$

Proof ...

Now, if $a \in \mathbb{R}$, then

$$\begin{aligned} x \cdot (au) &= x_1(au_1) + x_2(au_2) \\ &= a(x_1u_1 + x_2u_2) = 0. \end{aligned}$$

Therefore,

$$au \in x^{\perp}.$$

Thus, by definition, $x^{\perp}$ is a subspace of $\mathbb{R}^2$. $\square$