Scalar Product

Daniel H. Luecking

15 March 2024

Definition

If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ then the *scalar product of* $\mathbf{x}$ *and* $\mathbf{y}$ is $\mathbf{x}^T\mathbf{y}$

Definition

If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ then the *scalar product of* $\mathbf{x}$ *and* $\mathbf{y}$ is $\mathbf{x}^T\mathbf{y}$

In terms of the coordinates, if $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, then

$\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{y}^T\mathbf{x}$.

Definition

If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ then the *scalar product of* $\mathbf{x}$ *and* $\mathbf{y}$ is $\mathbf{x}^T\mathbf{y}$

In terms of the coordinates, if $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, then $\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{y}^T\mathbf{x}$.

Definition

We write $\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2}$ and call this the *norm of* $\mathbf{x}$.

Definition

If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ then the *scalar product of* $\mathbf{x}$ *and* $\mathbf{y}$ is $\mathbf{x}^T\mathbf{y}$

In terms of the coordinates, if $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, then $\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{y}^T\mathbf{x}$.

Definition

We write $\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2}$ and call this the *norm of* $\mathbf{x}$.

In terms of coordinates $\|\mathbf{x}\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$.

Definition

If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ then the *scalar product of* $\mathbf{x}$ *and* $\mathbf{y}$ is $\mathbf{x}^T\mathbf{y}$

In terms of the coordinates, if $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, then $\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{y}^T\mathbf{x}$.

Definition

We write $\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2}$ and call this the *norm of* $\mathbf{x}$.

In terms of coordinates $\|\mathbf{x}\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$.

In $\mathbb{R}^2$ we have $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$,

Definition

If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ then the *scalar product of* $\mathbf{x}$ *and* $\mathbf{y}$ is $\mathbf{x}^T\mathbf{y}$

In terms of the coordinates, if $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, then
$\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{y}^T\mathbf{x}$.

Definition

We write $\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2}$ and call this the *norm of* $\mathbf{x}$.

In terms of coordinates $\|\mathbf{x}\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$.
In $\mathbb{R}^2$ we have $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$, and in $\mathbb{R}^3$ we have
$\|\mathbf{x}\| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$.

Definition

If $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^n$ then the *scalar product of $\mathbf{x}$ and $\mathbf{y}$* is $\mathbf{x}^T\mathbf{y}$

In terms of the coordinates, if $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, then

$\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{y}^T\mathbf{x}$.

Definition

We write $\|\mathbf{x}\| = (\mathbf{x}^T\mathbf{x})^{1/2}$ and call this the *norm of $\mathbf{x}$*.

In terms of coordinates $\|\mathbf{x}\| = (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2}$.

In $\mathbb{R}^2$ we have $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$, and in $\mathbb{R}^3$ we have $\|\mathbf{x}\| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$. In either case this is the length of the arrow that we use to visualize $\mathbf{x}$.

A couple of simple properties of scalar products.

1. The scalar product of $\alpha\mathbf{x}$ and $\beta\mathbf{y}$ is $(\alpha\mathbf{x})^T(\beta\mathbf{y}) = \alpha\beta(\mathbf{x}^T\mathbf{y})$

A couple of simple properties of scalar products.

1. The scalar product of $\alpha\mathbf{x}$ and $\beta\mathbf{y}$ is $(\alpha\mathbf{x})^T(\beta\mathbf{y}) = \alpha\beta(\mathbf{x}^T\mathbf{y})$
2. The norm of $\alpha\mathbf{x}$ is $\|\alpha\mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|$. Note that if $\alpha = 1/\|\mathbf{x}\|$ the $\|\alpha\mathbf{x}\| = 1$.

A couple of simple properties of scalar products.

1. The scalar product of $\alpha \mathbf{x}$ and $\beta \mathbf{y}$ is $(\alpha \mathbf{x})^T (\beta \mathbf{y}) = \alpha\beta(\mathbf{x}^T \mathbf{y})$
2. The norm of $\alpha \mathbf{x}$ is $\|\alpha \mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|$. Note that if $\alpha = 1/\|\mathbf{x}\|$ the $\|\alpha \mathbf{x}\| = 1$.

In dimension 2 or 3 we have a geometric intrerpretation of the scalar product.

A couple of simple properties of scalar products.

1. The scalar product of $\alpha\mathbf{x}$ and $\beta\mathbf{y}$ is $(\alpha\mathbf{x})^T(\beta\mathbf{y}) = \alpha\beta(\mathbf{x}^T\mathbf{y})$
2. The norm of $\alpha\mathbf{x}$ is $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$. Note that if $\alpha = 1/\|\mathbf{x}\|$ the $\|\alpha\mathbf{x}\| = 1$.

In dimension 2 or 3 we have a geometric intrerpretation of the scalar product.

Theorem

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ or $\mathbb{R}^3$. Suppose $\|\mathbf{x}\| \neq 0$ and $\|\mathbf{y}\| \neq 0$. Let θ be the angle between $\mathbf{x}$ and $\mathbf{y}$ with $0 \leq \theta \leq 180^\circ$. Then

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

A couple of simple properties of scalar products.

1. The scalar product of $\alpha\mathbf{x}$ and $\beta\mathbf{y}$ is $(\alpha\mathbf{x})^T(\beta\mathbf{y}) = \alpha\beta(\mathbf{x}^T\mathbf{y})$
2. The norm of $\alpha\mathbf{x}$ is $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$. Note that if $\alpha = 1/\|\mathbf{x}\|$ the $\|\alpha\mathbf{x}\| = 1$.

In dimension 2 or 3 we have a geometric intrerpretation of the scalar product.

Theorem

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ or $\mathbb{R}^3$. Suppose $\|\mathbf{x}\| \neq 0$ and $\|\mathbf{y}\| \neq 0$. Let θ be the angle between $\mathbf{x}$ and $\mathbf{y}$ with $0 \leq \theta \leq 180°$. Then

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

We're actually going to prove this in the case $\mathbb{R}^2$.

A couple of simple properties of scalar products.

1. The scalar product of $\alpha\mathbf{x}$ and $\beta\mathbf{y}$ is $(\alpha\mathbf{x})^T(\beta\mathbf{y}) = \alpha\beta(\mathbf{x}^T\mathbf{y})$
2. The norm of $\alpha\mathbf{x}$ is $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$. Note that if $\alpha = 1/\|\mathbf{x}\|$ the $\|\alpha\mathbf{x}\| = 1$.

In dimension 2 or 3 we have a geometric intrerpretation of the scalar product.

Theorem

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ or $\mathbb{R}^3$. Suppose $\|\mathbf{x}\| \neq 0$ and $\|\mathbf{y}\| \neq 0$. Let θ be the angle between $\mathbf{x}$ and $\mathbf{y}$ with $0 \leq \theta \leq 180°$. Then

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

We're actually going to prove this in the case $\mathbb{R}^2$. But first recall the linear transformations on $\mathbb{R}^2$ given by rotations.

A couple of simple properties of scalar products.

1. The scalar product of $\alpha\mathbf{x}$ and $\beta\mathbf{y}$ is $(\alpha\mathbf{x})^T(\beta\mathbf{y}) = \alpha\beta(\mathbf{x}^T\mathbf{y})$
2. The norm of $\alpha\mathbf{x}$ is $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$. Note that if $\alpha = 1/\|\mathbf{x}\|$ the $\|\alpha\mathbf{x}\| = 1$.

In dimension 2 or 3 we have a geometric intrerpretation of the scalar product.

Theorem

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ or $\mathbb{R}^3$. Suppose $\|\mathbf{x}\| \neq 0$ and $\|\mathbf{y}\| \neq 0$. Let θ be the angle between $\mathbf{x}$ and $\mathbf{y}$ with $0 \leq \theta \leq 180°$. Then

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

We're actually going to prove this in the case $\mathbb{R}^2$. But first recall the linear transformations on $\mathbb{R}^2$ given by rotations. That is, if γ is any angle let $R_\gamma\mathbf{x}$ be the arrow $\mathbf{x}$ rotated counterclockwise by the angle γ.

Since R_γ goes from $\mathbb{R}^2$ to $\mathbb{R}^2$ it is a matrix transformation.

Since R_γ goes from $\mathbb{R}^2$ to $\mathbb{R}^2$ it is a matrix transformation. We get that matrix by

$$R_\gamma \mathbf{e}_1 = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad \text{and} \quad R_\gamma \mathbf{e}_2 = \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}$$

Since R_γ goes from $\mathbb{R}^2$ to $\mathbb{R}^2$ it is a matrix transformation. We get that matrix by

$$R_\gamma \mathbf{e}_1 = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad \text{and} \quad R_\gamma \mathbf{e}_2 = \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}$$

and therefore the matrix is $A = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$

Since R_γ goes from $\mathbb{R}^2$ to $\mathbb{R}^2$ it is a matrix transformation. We get that matrix by

$$R_\gamma \mathbf{e}_1 = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad \text{and} \quad R_\gamma \mathbf{e}_2 = \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}$$

and therefore the matrix is $A = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$

An easy computation show that $A^T A = I$, the identity matrix.

Since R_γ goes from $\mathbb{R}^2$ to $\mathbb{R}^2$ it is a matrix transformation. We get that matrix by

$$R_\gamma \mathbf{e}_1 = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad \text{and} \quad R_\gamma \mathbf{e}_2 = \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}$$

and therefore the matrix is $A = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$

An easy computation show that $A^T A = I$, the identity matrix. Therefore, the scalar product of $R_\gamma \mathbf{x}$ and $R_\gamma \mathbf{y}$ is

$$(R_\gamma \mathbf{x})^T (R_\gamma \mathbf{y}) = (A\mathbf{x})^T (A\mathbf{y}) = \mathbf{x}^T A^T A \mathbf{y} = \mathbf{x}^T I \mathbf{y} = \mathbf{x}^T \mathbf{y}$$

Since R_γ goes from $\mathbb{R}^2$ to $\mathbb{R}^2$ it is a matrix transformation. We get that matrix by

$$R_\gamma \mathbf{e}_1 = \begin{pmatrix} \cos\gamma \\ \sin\gamma \end{pmatrix} \quad \text{and} \quad R_\gamma \mathbf{e}_2 = \begin{pmatrix} -\sin\gamma \\ \cos\gamma \end{pmatrix}$$

and therefore the matrix is $A = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$

An easy computation show that $A^T A = I$, the identity matrix. Therefore, the scalar product of $R_\gamma \mathbf{x}$ and $R_\gamma \mathbf{y}$ is

$$(R_\gamma \mathbf{x})^T (R_\gamma \mathbf{y}) = (A\mathbf{x})^T (A\mathbf{y}) = \mathbf{x}^T A^T A \mathbf{y} = \mathbf{x}^T I \mathbf{y} = \mathbf{x}^T \mathbf{y}$$

Similarly

$$\|R_\gamma \mathbf{x}\| = \big((R_\gamma \mathbf{x})^T (R_\gamma \mathbf{x})\big)^{1/2} = (\mathbf{x}^T \mathbf{x})^{1/2} = \|\mathbf{x}\|$$

We now make several observations.

- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change under a rotation R_γ

We now make several observations.

- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change under a rotation R_γ
- The angle between $\mathbf{x}$ and $\mathbf{y}$ does not change when both are rotated by the same angle.

We now make several observations.

- The expression $\dfrac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}$ does not change under a rotation R_γ
- The angle between $\mathbf{x}$ and $\mathbf{y}$ does not change when both are rotated by the same angle.
- The expression $\dfrac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.

We now make several observations.

- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change under a rotation R_γ
- The angle between $\mathbf{x}$ and $\mathbf{y}$ does not change when both are rotated by the same angle.
- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.
- The angle beween $\mathbf{x}$ and $\mathbf{y}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.

We now make several observations.

- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change under a rotation R_γ
- The angle between $\mathbf{x}$ and $\mathbf{y}$ does not change when both are rotated by the same angle.
- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.
- The angle beween $\mathbf{x}$ and $\mathbf{y}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.

So, in $\mathbb{R}^2$, to prove that $\cos\theta = \dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$, we may change $\mathbf{x}$ and $\mathbf{y}$ by scalar multiplication so that $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$

We now make several observations.

- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change under a rotation R_γ
- The angle between $\mathbf{x}$ and $\mathbf{y}$ does not change when both are rotated by the same angle.
- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.
- The angle beween $\mathbf{x}$ and $\mathbf{y}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.

So, in $\mathbb{R}^2$, to prove that $\cos\theta = \dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$, we may change $\mathbf{x}$ and $\mathbf{y}$ by scalar multiplication so that $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$ and then change both by rotation so that $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$.

We now make several observations.

- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change under a rotation R_γ
- The angle between $\mathbf{x}$ and $\mathbf{y}$ does not change when both are rotated by the same angle.
- The expression $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.
- The angle beween $\mathbf{x}$ and $\mathbf{y}$ does not change if we multiply $\mathbf{x}$ and $\mathbf{y}$ by positive scalars.

So, in $\mathbb{R}^2$, to prove that $\cos\theta = \dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$, we may change $\mathbf{x}$ and $\mathbf{y}$ by scalar multiplication so that $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$ and then change both by rotation so that $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$. And then $\dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} = \dfrac{\cos\theta}{1}$.

In $\mathbb{R}^3$ we can do something similar. Roughly speaking we can perform a few rotations so that both vectors have the third component 0 and then apply the $\mathbb{R}^2$ case.

In $\mathbb{R}^3$ we can do something similar. Roughly speaking we can perform a few rotations so that both vectors have the third component 0 and then apply the $\mathbb{R}^2$ case.

There is a useful special case: If $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 1$ and θ is the angle between them, then $\cos\theta = \mathbf{u}^T\mathbf{v}$.

In $\mathbb{R}^3$ we can do something similar. Roughly speaking we can perform a few rotations so that both vectors have the third component 0 and then apply the $\mathbb{R}^2$ case.

There is a useful special case: If $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 1$ and θ is the angle between them, then $\cos\theta = \mathbf{u}^T\mathbf{v}$. We can always convert to this case by letting $\mathbf{u} = \dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$ and $\mathbf{v} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$.

In $\mathbb{R}^3$ we can do something similar. Roughly speaking we can perform a few rotations so that both vectors have the third component 0 and then apply the $\mathbb{R}^2$ case.

There is a useful special case: If $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 1$ and θ is the angle between them, then $\cos\theta = \mathbf{u}^T\mathbf{v}$. We can always convert to this case by letting $\mathbf{u} = \dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$ and $\mathbf{v} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$.

Example: if $\mathbf{x} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ then $\|\mathbf{x}\| = 5$ and $\|\mathbf{y}\| = \sqrt{2}$, $\mathbf{x}^T\mathbf{y} = 1$.

In $\mathbb{R}^3$ we can do something similar. Roughly speaking we can perform a few rotations so that both vectors have the third component 0 and then apply the $\mathbb{R}^2$ case.

There is a useful special case: If $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 1$ and θ is the angle between them, then $\cos\theta = \mathbf{u}^T\mathbf{v}$. We can always convert to this case by letting $\mathbf{u} = \dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$ and $\mathbf{v} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$.

Example: if $\mathbf{x} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ then $\|\mathbf{x}\| = 5$ and $\|\mathbf{y}\| = \sqrt{2}$, $\mathbf{x}^T\mathbf{y} = 1$. So, if θ is the angle between $\mathbf{x}$ and $\mathbf{y}$ we can compute $\cos\theta$ by $\frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{1}{5\sqrt{2}}$

In $\mathbb{R}^3$ we can do something similar. Roughly speaking we can perform a few rotations so that both vectors have the third component 0 and then apply the $\mathbb{R}^2$ case.

There is a useful special case: If $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 1$ and θ is the angle between them, then $\cos\theta = \mathbf{u}^T\mathbf{v}$. We can always convert to this case by letting $\mathbf{u} = \dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$ and $\mathbf{v} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$.

Example: if $\mathbf{x} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ then $\|\mathbf{x}\| = 5$ and $\|\mathbf{y}\| = \sqrt{2}$, $\mathbf{x}^T\mathbf{y} = 1$. So, if θ is the angle between $\mathbf{x}$ and $\mathbf{y}$ we can compute $\cos\theta$ by $\frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{1}{5\sqrt{2}}$ or by

$$\cos\theta = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}^T \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \frac{-3}{5\sqrt{2}} + \frac{4}{5\sqrt{2}} = \frac{1}{5\sqrt{2}}$$

The Cauchy-Schwarz Inequality

Theorem

If $\mathbf{x}$ *and* $\mathbf{y}$ *are vectors in* $\mathbb{R}^n$*, then* $|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$*.*

The Cauchy-Schwarz Inequality

Theorem

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, then $|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$.

In $\mathbb{R}^2$ or $\mathbb{R}^3$ this follows from $\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta$ and the fact that $|\cos\theta| \leq 1$.

The Cauchy-Schwarz Inequality

Theorem

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, then $|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$.

In $\mathbb{R}^2$ or $\mathbb{R}^3$ this follows from $\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$ and the fact that $|\cos\theta| \leq 1$. In $\mathbb{R}^n$ for $n > 3$ we can rotate 2 variables at a time (without changing the scalar product or the norms) until coordinates x_j and y_j are 0 for all $j > 2$. Then apply the $\mathbb{R}^2$ case.

The Cauchy-Schwarz Inequality

Theorem

If $\mathbf{x}$ *and* $\mathbf{y}$ *are vectors in* $\mathbb{R}^n$*, then* $|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$.

In $\mathbb{R}^2$ or $\mathbb{R}^3$ this follows from $\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$ and the fact that $|\cos\theta| \leq 1$. In $\mathbb{R}^n$ for $n > 3$ we can rotate 2 variables at a time (without changing the scalar product or the norms) until coordinates x_j and y_j are 0 for all $j > 2$. Then apply the $\mathbb{R}^2$ case.

Definition

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$, we say that $\mathbf{x}$ *is orthogonal to* $\mathbf{y}$ if $\mathbf{x}^T\mathbf{y} = 0$. We denote this by writing $\mathbf{x} \perp \mathbf{y}$.

The Cauchy-Schwarz Inequality

Theorem

If $\mathbf{x}$ *and* $\mathbf{y}$ *are vectors in* $\mathbb{R}^n$*, then* $|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$.

In $\mathbb{R}^2$ or $\mathbb{R}^3$ this follows from $\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$ and the fact that $|\cos\theta| \leq 1$. In $\mathbb{R}^n$ for $n > 3$ we can rotate 2 variables at a time (without changing the scalar product or the norms) until coordinates x_j and y_j are 0 for all $j > 2$. Then apply the $\mathbb{R}^2$ case.

Definition

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$, we say that $\mathbf{x}$ *is orthogonal to* $\mathbf{y}$ if $\mathbf{x}^T\mathbf{y} = 0$. We denote this by writing $\mathbf{x} \perp \mathbf{y}$.

Note that if $\mathbf{x} \perp \mathbf{y}$ then also $\mathbf{y} \perp \mathbf{x}$.

The Cauchy-Schwarz Inequality

Theorem

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, then $|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$.

In $\mathbb{R}^2$ or $\mathbb{R}^3$ this follows from $\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$ and the fact that $|\cos\theta| \leq 1$. In $\mathbb{R}^n$ for $n > 3$ we can rotate 2 variables at a time (without changing the scalar product or the norms) until coordinates x_j and y_j are 0 for all $j > 2$. Then apply the $\mathbb{R}^2$ case.

Definition

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$, we say that $\mathbf{x}$ *is orthogonal to* $\mathbf{y}$ if $\mathbf{x}^T\mathbf{y} = 0$. We denote this by writing $\mathbf{x} \perp \mathbf{y}$.

Note that if $\mathbf{x} \perp \mathbf{y}$ then also $\mathbf{y} \perp \mathbf{x}$. Also $\mathbf{0}$ is orthogonal to any vector.

The Cauchy-Schwarz Inequality

Theorem

If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, then $|\mathbf{x}^T\mathbf{y}| \le \|\mathbf{x}\|\,\|\mathbf{y}\|$.

In $\mathbb{R}^2$ or $\mathbb{R}^3$ this follows from $\mathbf{x}^T\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\| \cos\theta$ and the fact that $|\cos\theta| \le 1$. In $\mathbb{R}^n$ for $n > 3$ we can rotate 2 variables at a time (without changing the scalar product or the norms) until coordinates x_j and y_j are 0 for all $j > 2$. Then apply the $\mathbb{R}^2$ case.

Definition

If $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$, we say that $\mathbf{x}$ *is orthogonal to* $\mathbf{y}$ if $\mathbf{x}^T\mathbf{y} = 0$. We denote this by writing $\mathbf{x} \perp \mathbf{y}$.

Note that if $\mathbf{x} \perp \mathbf{y}$ then also $\mathbf{y} \perp \mathbf{x}$. Also $\mathbf{0}$ is orthogonal to any vector. In $\mathbb{R}^2$ or $\mathbb{R}^3$, if neither $\mathbf{x}$ nor $\mathbf{y}$ is $\mathbf{0}$ then $\mathbf{x} \perp \mathbf{y}$ means that the angle between them is $90°$.

Some properties of orthogonality:

1. If $\mathbf{x} \perp \mathbf{y}_1, \mathbf{x} \perp \mathbf{y}_2, \ldots, \mathbf{x} \perp \mathbf{y}_n$ then $\mathbf{x}$ is orthogonal to every vector in $\mathrm{Span}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Proof:

$$\mathbf{x}^T(c_1\mathbf{y}_1 + \cdots + c_n\mathbf{y}_n) = c_1\mathbf{x}^T\mathbf{y}_1 + \cdots + c_n\mathbf{x}^T\mathbf{y}_n = 0 + \cdots + 0 = 0$$

Some properties of orthogonality:

1. If $\mathbf{x} \perp \mathbf{y}_1, \mathbf{x} \perp \mathbf{y}_2, \ldots, \mathbf{x} \perp \mathbf{y}_n$ then $\mathbf{x}$ is orthogonal to every vector in $\mathrm{Span}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Proof:

$$\mathbf{x}^T(c_1\mathbf{y}_1 + \cdots + c_n\mathbf{y}_n) = c_1\mathbf{x}^T\mathbf{y}_1 + \cdots + c_n\mathbf{x}^T\mathbf{y}_n = 0 + \cdots + 0 = 0$$

2. If $\mathbf{x} \in \mathbb{R}^n$ then the set of vectors orthogonal to $\mathbf{x}$ is a subspace of $\mathbb{R}^n$.

Some properties of orthogonality:

1. If $\mathbf{x} \perp \mathbf{y}_1, \mathbf{x} \perp \mathbf{y}_2, \ldots, \mathbf{x} \perp \mathbf{y}_n$ then $\mathbf{x}$ is orthogonal to every vector in $\mathrm{Span}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Proof:

$$\mathbf{x}^T(c_1\mathbf{y}_1 + \cdots + c_n\mathbf{y}_n) = c_1\mathbf{x}^T\mathbf{y}_1 + \cdots + c_n\mathbf{x}^T\mathbf{y}_n = 0 + \cdots + 0 = 0$$

2. If $\mathbf{x} \in \mathbb{R}^n$ then the set of vectors orthogonal to $\mathbf{x}$ is a subspace of $\mathbb{R}^n$.
3. $\mathbf{x} \perp \mathbf{x}$ if and only if $\mathbf{x} = \mathbf{0}$. Proof: $0 = \mathbf{x}^T\mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2$ can only be true if all $x_j = 0$.

Some properties of orthogonality:

1. If $\mathbf{x} \perp \mathbf{y}_1, \mathbf{x} \perp \mathbf{y}_2, \ldots, \mathbf{x} \perp \mathbf{y}_n$ then $\mathbf{x}$ is orthogonal to every vector in $\mathrm{Span}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Proof:
$$\mathbf{x}^T(c_1\mathbf{y}_1 + \cdots + c_n\mathbf{y}_n) = c_1\mathbf{x}^T\mathbf{y}_1 + \cdots + c_n\mathbf{x}^T\mathbf{y}_n = 0 + \cdots + 0 = 0$$
2. If $\mathbf{x} \in \mathbb{R}^n$ then the set of vectors orthogonal to $\mathbf{x}$ is a subspace of $\mathbb{R}^n$.
3. $\mathbf{x} \perp \mathbf{x}$ if and only if $\mathbf{x} = \mathbf{0}$. Proof: $0 = \mathbf{x}^T\mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2$ can only be true if all $x_j = 0$.

Example: find a vector orthogonal to $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$.

Some properties of orthogonality:

1. If $\mathbf{x} \perp \mathbf{y}_1, \mathbf{x} \perp \mathbf{y}_2, \ldots, \mathbf{x} \perp \mathbf{y}_n$ then $\mathbf{x}$ is orthogonal to every vector in $\mathrm{Span}(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n)$. Proof:
$$\mathbf{x}^T(c_1\mathbf{y}_1 + \cdots + c_n\mathbf{y}_n) = c_1\mathbf{x}^T\mathbf{y}_1 + \cdots + c_n\mathbf{x}^T\mathbf{y}_n = 0 + \cdots + 0 = 0$$
2. If $\mathbf{x} \in \mathbb{R}^n$ then the set of vectors orthogonal to $\mathbf{x}$ is a subspace of $\mathbb{R}^n$.
3. $\mathbf{x} \perp \mathbf{x}$ if and only if $\mathbf{x} = \mathbf{0}$. Proof: $0 = \mathbf{x}^T\mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2$ can only be true if all $x_j = 0$.

Example: find a vector orthogonal to $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$. Solution: solve

$$\mathbf{a}^T\mathbf{x} = 0: \;\; x_1 - 2x_2 + 3x_3 = 0 \quad \text{or} \quad x_1 = 2x_2 - 3x_3$$

Some properties of orthogonality:

1. If $\mathbf{x} \perp \mathbf{y}_1, \mathbf{x} \perp \mathbf{y}_2, \dots, \mathbf{x} \perp \mathbf{y}_n$ then $\mathbf{x}$ is orthogonal to every vector in $\mathrm{Span}(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)$. Proof:
$$\mathbf{x}^T(c_1\mathbf{y}_1 + \cdots + c_n\mathbf{y}_n) = c_1\mathbf{x}^T\mathbf{y}_1 + \cdots + c_n\mathbf{x}^T\mathbf{y}_n = 0 + \cdots + 0 = 0$$
2. If $\mathbf{x} \in \mathbb{R}^n$ then the set of vectors orthogonal to $\mathbf{x}$ is a subspace of $\mathbb{R}^n$.
3. $\mathbf{x} \perp \mathbf{x}$ if and only if $\mathbf{x} = \mathbf{0}$. Proof: $0 = \mathbf{x}^T\mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2$ can only be true if all $x_j = 0$.

Example: find a vector orthogonal to $\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$. Solution: solve

$$\mathbf{a}^T\mathbf{x} = 0\text{:} \ \ x_1 - 2x_2 + 3x_3 = 0 \quad \text{or} \quad x_1 = 2x_2 - 3x_3$$

Here x_1 is leading, x_2 and x_3 are free. The basic solutions

$$\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

are orthogonal to $\mathbf{a}$, as is every vector in their span.

Example: find a vector orthogonal to both $\mathbf{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ and $\mathbf{a}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$.

Example: find a vector orthogonal to both $\mathbf{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ and $\mathbf{a}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$. Solution: solve the system $\mathbf{a}_1^T \mathbf{x} = 0$, $\mathbf{a}_2^T \mathbf{x} = 0$:

Example: find a vector orthogonal to both $\mathbf{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ and $\mathbf{a}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$. Solution: solve the system $\mathbf{a}_1^T \mathbf{x} = 0$, $\mathbf{a}_2^T \mathbf{x} = 0$: This system is equivalent to $A\mathbf{x} = \mathbf{0}$ where $A = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{pmatrix}$. We solve that by row-reducing A:

$$\begin{pmatrix} 1 & -1 & 3 \\ -1 & 2 & -1 \end{pmatrix} \xrightarrow{R_2+R_1} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & 2 \end{pmatrix} \xrightarrow{R_1+R_2} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 2 \end{pmatrix}$$

Example: find a vector orthogonal to both $\mathbf{a}_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$ and $\mathbf{a}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$. Solution: solve the system $\mathbf{a}_1^T \mathbf{x} = 0$, $\mathbf{a}_2^T \mathbf{x} = 0$: This system is equivalent to $A\mathbf{x} = \mathbf{0}$ where $A = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{pmatrix}$. We solve that by row-reducing A:

$$\begin{pmatrix} 1 & -1 & 3 \\ -1 & 2 & -1 \end{pmatrix} \xrightarrow{R_2+R_1} \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & 2 \end{pmatrix} \xrightarrow{R_1+R_2} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 2 \end{pmatrix}$$

This has leading variables x_1 and x_2 with free variable x_3, also $x_1 = -5x_3$ and $x_2 = -2x_3$ and so the vectors $\begin{pmatrix} -5\alpha \\ -2\alpha \\ \alpha \end{pmatrix}$ are orthogonal to $\mathbf{a}_1$ and $\mathbf{a}_2$, for any choice of α.

Orthogonal Projection

If $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$ is the distance from the point $(0, 0)$ to (x_1, x_2).

Orthogonal Projection

If $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$ is the distance from the point $(0,0)$ to (x_1, x_2). If $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, then $\|\mathbf{x} - \mathbf{y}\|$ is the distance between (x_1, x_2) and (y_1, y_2).

Orthogonal Projection

If $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$ is the distance from the point $(0,0)$ to (x_1, x_2). If $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, then $\|\mathbf{x} - \mathbf{y}\|$ is the distance between (x_1, x_2) and (y_1, y_2). Because of this, and our formula for angles between vectors, we can do geometry via linear algebra.

Orthogonal Projection

If $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$ is the distance from the point $(0,0)$ to (x_1, x_2). If $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, then $\|\mathbf{x} - \mathbf{y}\|$ is the distance between (x_1, x_2) and (y_1, y_2). Because of this, and our formula for angles between vectors, we can do geometry via linear algebra.

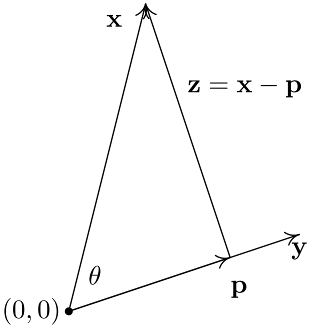
Suppose we want to find the point on a line that is closest to the point (x_1, x_2).

Orthogonal Projection

If $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $\|\mathbf{x}\| = (x_1^2 + x_2^2)^{1/2}$ is the distance from the point $(0,0)$ to (x_1, x_2). If $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, then $\|\mathbf{x} - \mathbf{y}\|$ is the distance between (x_1, x_2) and (y_1, y_2). Because of this, and our formula for angles between vectors, we can do geometry via linear algebra.

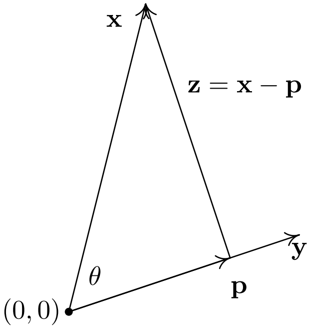
Suppose we want to find the point on a line that is closest to the point (x_1, x_2). Suppose we can express a line as all points that are tips of the vectors $\alpha\mathbf{y}$ for all $\alpha \in \mathbb{R}$.

The following is a picture of this setup:



We seek to find the tip of $\mathbf{p}$.

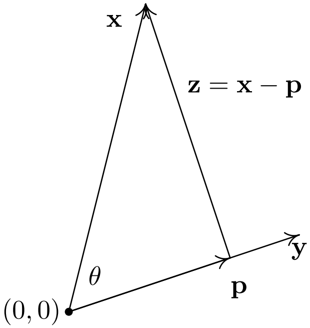
The following is a picture of this setup:



We seek to find the tip of $\mathbf{p}$.

We can find the length of $\mathbf{p}$ via trigonometry: $\|\mathbf{p}\| = \|\mathbf{x}\| \cos\theta$.

The following is a picture of this setup:

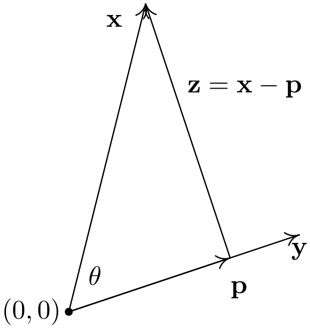


We seek to find the tip of $\mathbf{p}$.

We can find the length of $\mathbf{p}$ via trigonometry: $\|\mathbf{p}\| = \|\mathbf{x}\| \cos\theta$. With our formula for $\cos\theta$, we have

$$\|\mathbf{p}\| = \|\mathbf{x}\| \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}$$

The following is a picture of this setup:



We seek to find the tip of $\mathbf{p}$.

We can find the length of $\mathbf{p}$ via trigonometry: $\|\mathbf{p}\| = \|\mathbf{x}\| \cos\theta$. With our formula for $\cos\theta$, we have

$$\|\mathbf{p}\| = \|\mathbf{x}\| \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}$$

To get $\mathbf{p}$ itself we observe that $\mathbf{u} = (1/\|\mathbf{y}\|)\mathbf{y}$ has the same direction as both $\mathbf{y}$ and $\mathbf{p}$, but has length 1.

Thus $\mathbf{p} = \|\mathbf{p}\| \, \mathbf{u}$.

Thus $\mathbf{p} = \|\mathbf{p}\| \, \mathbf{u}$. Putting these together:

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}\frac{1}{\|\mathbf{y}\|}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y}.$$

Thus $\mathbf{p} = \|\mathbf{p}\| \, \mathbf{u}$. Putting these together:

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}\frac{1}{\|\mathbf{y}\|}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y}.$$

Definition

If $\mathbf{x}$ and $\mathbf{y}$ belong to $\mathbb{R}^2$ or $\mathbb{R}^3$ then:

The number $\alpha = \dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}$ is called the *scalar projection of $\mathbf{x}$ onto $\mathbf{y}$*.

The vector $\mathbf{p} = \dfrac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y}$ is called the *vector projection of $\mathbf{x}$ onto $\mathbf{y}$*

Thus $\mathbf{p} = \|\mathbf{p}\| \, \mathbf{u}$. Putting these together:

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}\frac{1}{\|\mathbf{y}\|}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y}.$$

Definition

If $\mathbf{x}$ and $\mathbf{y}$ belong to $\mathbb{R}^2$ or $\mathbb{R}^3$ then:

The number $\alpha = \dfrac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{y}\|}$ is called the *scalar projection of $\mathbf{x}$ onto $\mathbf{y}$*.

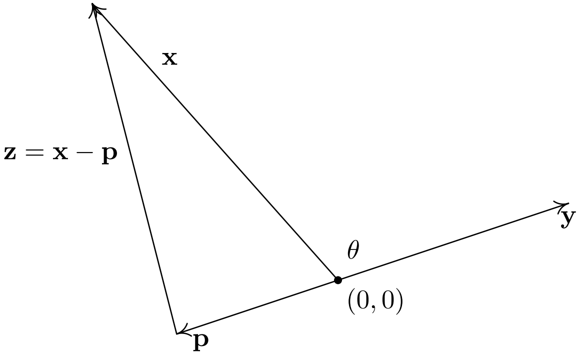
The vector $\mathbf{p} = \dfrac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y}$ is called the *vector projection of $\mathbf{x}$ onto $\mathbf{y}$*

Note that the scalar projection α times the vector $\mathbf{u} = (1/\|\mathbf{y}\|)\mathbf{y}$ is the vector projection.

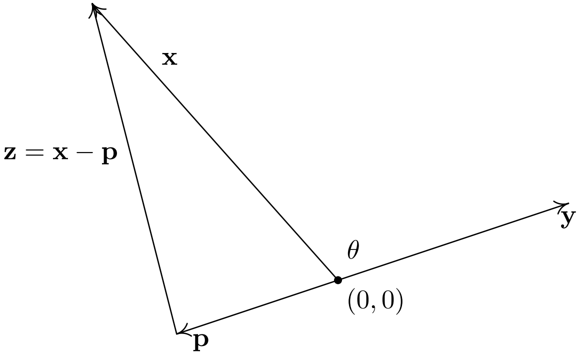
The calculation of α and $\mathbf{p}$ was based on a figure where $\theta < 90^\circ$.

The calculation of α and $\mathbf{p}$ was based on a figure where $\theta < 90^\circ$. The following figure ilustrates what happens when $\theta > 90^\circ$.

The calculation of α and $\mathbf{p}$ was based on a figure where $\theta < 90°$. The following figure ilustrates what happens when $\theta > 90°$.

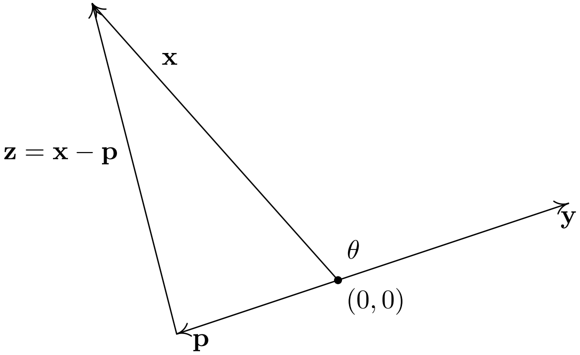


The calculation of α and $\mathbf{p}$ was based on a figure where $\theta < 90°$. The following figure ilustrates what happens when $\theta > 90°$.



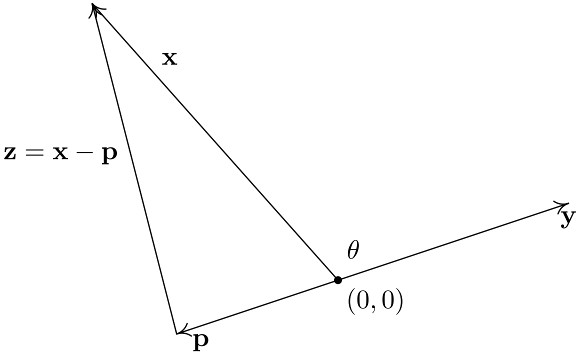
In this case $\cos\theta$ is negative and so is α.

The calculation of α and $\mathbf{p}$ was based on a figure where $\theta < 90°$. The following figure ilustrates what happens when $\theta > 90°$.



In this case $\cos\theta$ is negative and so is α. Then $\alpha = -\|\mathbf{p}\|$, but $\alpha\mathbf{u}$ is still $\mathbf{p}$ because $\mathbf{p}$ points in the direction opposite to $\mathbf{u}$ and $\mathbf{y}$.

The calculation of α and $\mathbf{p}$ was based on a figure where $\theta < 90^\circ$. The following figure ilustrates what happens when $\theta > 90^\circ$.



In this case $\cos\theta$ is negative and so is α. Then $\alpha = -\|\mathbf{p}\|$, but $\alpha\mathbf{u}$ is still $\mathbf{p}$ because $\mathbf{p}$ points in the direction opposite to $\mathbf{u}$ and $\mathbf{y}$. The formulas for α and $\mathbf{p}$ from the definition still hold in this case.

The distance between (x_1, x_2) and the line specified by $\mathbf{y}$ is the length of the vector $\mathbf{x} - \mathbf{p}$: $\|\mathbf{x} - \mathbf{p}\|$.

The distance between (x_1, x_2) and the line specified by $\mathbf{y}$ is the length of the vector $\mathbf{x} - \mathbf{p}$: $\|\mathbf{x} - \mathbf{p}\|$.

Here is an example: Let L be the line passing through $(0, 0)$ and $(3, 1)$.

We can take $\mathbf{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Consider the point $(1, 4)$, whose vector form is

$\mathbf{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$

The distance between (x_1, x_2) and the line specified by $\mathbf{y}$ is the length of the vector $\mathbf{x} - \mathbf{p}$: $\|\mathbf{x} - \mathbf{p}\|$.

Here is an example: Let L be the line passing through $(0, 0)$ and $(3, 1)$. We can take $\mathbf{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Consider the point $(1, 4)$, whose vector form is $\mathbf{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ To get the closest point to $(1, 4)$ on this line we compute

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y} = \frac{7}{10}\mathbf{y} = \begin{pmatrix} 2.1 \\ 0.7 \end{pmatrix}$$

The distance between (x_1, x_2) and the line specified by $\mathbf{y}$ is the length of the vector $\mathbf{x} - \mathbf{p}$: $\|\mathbf{x} - \mathbf{p}\|$.

Here is an example: Let L be the line passing through $(0, 0)$ and $(3, 1)$. We can take $\mathbf{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Consider the point $(1, 4)$, whose vector form is $\mathbf{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ To get the closest point to $(1, 4)$ on this line we compute

$$\mathbf{p} = \frac{\mathbf{x}^T \mathbf{y}}{\mathbf{y}^T \mathbf{y}} \mathbf{y} = \frac{7}{10} \mathbf{y} = \begin{pmatrix} 2.1 \\ 0.7 \end{pmatrix}$$

So the closest point of L to $(1, 4)$ is $(2.1, 0.7)$ and the distance is $\left((1 - 2.1)^2 + (4 - 0.7)^2\right)^{1/2} = \sqrt{12.1} \approx 3.4785$.

The distance between (x_1, x_2) and the line specified by $\mathbf{y}$ is the length of the vector $\mathbf{x} - \mathbf{p}$: $\|\mathbf{x} - \mathbf{p}\|$.

Here is an example: Let L be the line passing through $(0,0)$ and $(3,1)$. We can take $\mathbf{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Consider the point $(1,4)$, whose vector form is $\mathbf{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ To get the closest point to $(1,4)$ on this line we compute

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y} = \frac{7}{10}\mathbf{y} = \begin{pmatrix} 2.1 \\ 0.7 \end{pmatrix}$$

So the closest point of L to $(1,4)$ is $(2.1, 0.7)$ and the distance is $\left((1-2.1)^2 + (4-0.7)^2\right)^{1/2} = \sqrt{12.1} \approx 3.4785$.

If we try this with a line L that does not pass through $(0,0)$ we simply have to pick two points P_0 and P_1 on the line.

The distance between (x_1, x_2) and the line specified by $\mathbf{y}$ is the length of the vector $\mathbf{x} - \mathbf{p}$: $\|\mathbf{x} - \mathbf{p}\|$.

Here is an example: Let L be the line passing through $(0,0)$ and $(3,1)$. We can take $\mathbf{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Consider the point $(1,4)$, whose vector form is $\mathbf{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ To get the closest point to $(1,4)$ on this line we compute

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y} = \frac{7}{10}\mathbf{y} = \begin{pmatrix} 2.1 \\ 0.7 \end{pmatrix}$$

So the closest point of L to $(1,4)$ is $(2.1, 0.7)$ and the distance is $\left((1-2.1)^2 + (4-0.7)^2\right)^{1/2} = \sqrt{12.1} \approx 3.4785$.

If we try this with a line L that does not pass through $(0,0)$ we simply have to pick two points P_0 and P_1 on the line. The displacement vector from P_0 to P_1 will be denoted $\overrightarrow{P_0P_1}$.

The distance between (x_1, x_2) and the line specified by $\mathbf{y}$ is the length of the vector $\mathbf{x} - \mathbf{p}$: $\|\mathbf{x} - \mathbf{p}\|$.

Here is an example: Let L be the line passing through $(0, 0)$ and $(3, 1)$. We can take $\mathbf{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Consider the point $(1, 4)$, whose vector form is $\mathbf{x} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ To get the closest point to $(1, 4)$ on this line we compute

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y} = \frac{7}{10}\mathbf{y} = \begin{pmatrix} 2.1 \\ 0.7 \end{pmatrix}$$

So the closest point of L to $(1, 4)$ is $(2.1, 0.7)$ and the distance is $\left((1 - 2.1)^2 + (4 - 0.7)^2\right)^{1/2} = \sqrt{12.1} \approx 3.4785$.

If we try this with a line L that does not pass through $(0, 0)$ we simply have to pick two points P_0 and P_1 on the line. The displacement vector from P_0 to P_1 will be denoted $\overrightarrow{P_0P_1}$. All the calculations to get the point on L closest to a point P then take place with $\mathbf{x} = \overrightarrow{P_0P}$ and $\mathbf{y} = \overrightarrow{P_0P_1}$. When we find $\mathbf{p}$, the closest point will be the point Q such that $\overrightarrow{P_0Q} = \mathbf{p}$.

Example: Consider the Line with slope $1/3$ that passes through $(0, 1)$ ($y = (1/3)x + 1$).

Example: Consider the Line with slope $1/3$ that passes through $(0,1)$ ($y=(1/3)x+1$). Then take $P_0=(0,1)$ and $P_1=(3,2)$ so $\mathbf{y}=\overrightarrow{P_0P_1}=\begin{pmatrix}3\\1\end{pmatrix}$.

Example: Consider the Line with slope $1/3$ that passes through $(0, 1)$ ($y = (1/3)x + 1$). Then take $P_0 = (0, 1)$ and $P_1 = (3, 2)$ so $\mathbf{y} = \overrightarrow{P_0P_1} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Say we want the closest point from $P = (1, 4)$ to this line. Then $\mathbf{x} = \overrightarrow{P_0P} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

Example: Consider the Line with slope $1/3$ that passes through $(0,1)$ $(y = (1/3)x + 1)$. Then take $P_0 = (0,1)$ and $P_1 = (3,2)$ so $\mathbf{y} = \overrightarrow{P_0P_1} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Say we want the closest point from $P = (1,4)$ to this line. Then $\mathbf{x} = \overrightarrow{P_0P} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$. Then we get the projection of $\mathbf{x}$ onto $\mathbf{y}$

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\mathbf{y} = \frac{6}{10}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1.8 \\ 0.6 \end{pmatrix}$$

To get that closest point, we add the components of $\mathbf{p}$ to P_0 to get $Q = (1.8, 1.6)$ and the distance $((1-1.8)^2 + (4-1.6)^2)^{1/2} = \sqrt{6.4} \approx 2.53$.

Planes in $\mathbb{R}^3$

A plane can be specified in terms of orthogonality.

Planes in $\mathbb{R}^3$

A plane can be specified in terms of orthogonality. If $\mathbf{a}$ is any vector, then the set of vectors orthogonal to $\mathbf{a}$ is a plane passing through $(0, 0, 0)$.

Planes in $\mathbb{R}^3$

A plane can be specified in terms of orthogonality. If $\mathbf{a}$ is any vector, then the set of vectors orthogonal to $\mathbf{a}$ is a plane passing through $(0, 0, 0)$. The expression of orthogonality gives us an equation for the plane:

$$\mathbf{a}^T\mathbf{x} = 0 \text{ gives } a_1x_1 + a_2x_2 + a_3x_3 = 0$$

Planes in $\mathbb{R}^3$

A plane can be specified in terms of orthogonality. If $\mathbf{a}$ is any vector, then the set of vectors orthogonal to $\mathbf{a}$ is a plane passing through $(0,0,0)$. The expression of orthogonality gives us an equation for the plane:

$$\mathbf{a}^T\mathbf{x} = 0 \text{ gives } a_1x_1 + a_2x_2 + a_3x_3 = 0$$

In the usual format (i.e., using variables x, y, z) this says that

$$ax + by + cz = 0 \text{ is the plane through } (0,0,0) \text{ perpendicular to } \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Planes in $\mathbb{R}^3$

A plane can be specified in terms of orthogonality. If $\mathbf{a}$ is any vector, then the set of vectors orthogonal to $\mathbf{a}$ is a plane passing through $(0,0,0)$. The expression of orthogonality gives us an equation for the plane:

$$\mathbf{a}^T\mathbf{x} = 0 \text{ gives } a_1x_1 + a_2x_2 + a_3x_3 = 0$$

In the usual format (i.e., using variables x, y, z) this says that

$$ax + by + cz = 0 \text{ is the plane through } (0,0,0) \text{ perpendicular to } \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

If we want a plane that doesn't pass through $(0,0,0)$ we need to have a point $P_0 = (x_0, y_0, z_0)$ that is on the plane, and then the plane is the set of points $P = (x, y, z)$ such that $\overrightarrow{P_0P} \perp \mathbf{a}$ for some vector $\mathbf{a}$.

Planes in $\mathbb{R}^3$

A plane can be specified in terms of orthogonality. If $\mathbf{a}$ is any vector, then the set of vectors orthogonal to $\mathbf{a}$ is a plane passing through $(0,0,0)$. The expression of orthogonality gives us an equation for the plane:

$$\mathbf{a}^T\mathbf{x} = 0 \;\text{ gives }\; a_1x_1 + a_2x_2 + a_3x_3 = 0$$

In the usual format (i.e., using variables x, y, z) this says that

$$ax + by + cz = 0 \text{ is the plane through } (0,0,0) \text{ perpendicular to } \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

If we want a plane that doesn't pass through $(0,0,0)$ we need to have a point $P_0 = (x_0, y_0, z_0)$ that is on the plane, and then the plane is the set of points $P = (x, y, z)$ such that $\overrightarrow{P_0P} \perp \mathbf{a}$ for some vector $\mathbf{a}$.That is, the equation of the plane through (x_0, y_0, z_0) and perpendicular to $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$