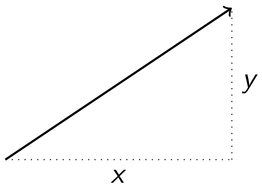
Linear Algebra

Jonny Evans

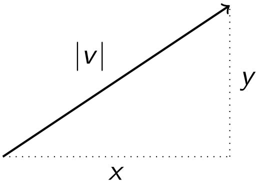
MATH105

Matrices and vectors

A vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ represents an arrow:

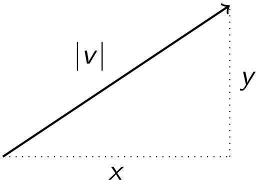


A vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ represents an arrow:



Length of v?

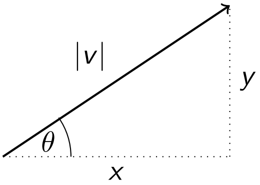
A vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ represents an arrow:



Length of v?

$$|v| = \sqrt{x^2 + y^2}$$

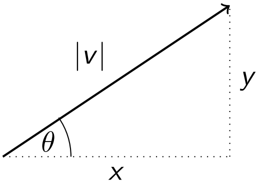
A vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ represents an arrow:



Length of v? $|v| = \sqrt{x^2 + y^2}$

Angle θ?

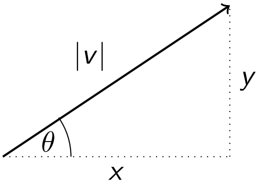
A vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ represents an arrow:



Length of v? $\qquad |v| = \sqrt{x^2 + y^2}$

Angle θ? $\qquad \theta = \arctan(y/x)$

A vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$ in $\mathbb{R}^2$ represents an arrow:



Length of v? $\qquad |v| = \sqrt{x^2 + y^2}$

Angle θ? $\qquad \theta = \arctan(y/x)$

$$v = \begin{pmatrix} |v| \cos \theta \\ |v| \sin \theta \end{pmatrix}.$$

Theorem

If w is obtained by rotating $v = \begin{pmatrix} x \\ y \end{pmatrix}$ by an angle ϕ anticlockwise around its basepoint then

$$w = \begin{pmatrix} x\cos\phi - y\sin\phi \\ x\sin\phi + y\cos\phi \end{pmatrix}.$$

Theorem

If w is obtained by rotating $v = \begin{pmatrix} x \\ y \end{pmatrix}$ by an angle ϕ anticlockwise around its basepoint then

$$w = \begin{pmatrix} x\cos\phi - y\sin\phi \\ x\sin\phi + y\cos\phi \end{pmatrix}.$$

Definition

A *2-by-2 matrix* is a grid of numbers $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Given $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $v = \begin{pmatrix} x \\ y \end{pmatrix}$, define the *linear map*

$$\mathbb{R}^2 \to \mathbb{R}^2, \qquad v \mapsto Mv := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

Similarly, a vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ represents an arrow in $\mathbb{R}^3$.

Similarly, a vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ represents an arrow in $\mathbb{R}^3$.

A 3-by-3 matrix $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ defines a linear map $\mathbb{R}^3 \to \mathbb{R}^3$,

Similarly, a vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ represents an arrow in $\mathbb{R}^3$.

A 3-by-3 matrix $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ defines a linear map $\mathbb{R}^3 \to \mathbb{R}^3$,

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{pmatrix}.$$

Definition

An m-by-n matrix is a grid with m rows and n columns

$$\begin{array}{cc} & n \\ m & \begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & \cdots & A_{2n} \\ \vdots & \vdots & & & \vdots \\ A_{m1} & A_{m2} & \cdots & \cdots & A_{mn} \end{pmatrix} \end{array}$$

Definition

An m-by-n matrix is a grid with m rows and n columns

$$\begin{array}{cc} & \overset{n}{\longleftrightarrow} \\ m \updownarrow & \begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & \cdots & A_{2n} \\ \vdots & \vdots & & & \vdots \\ A_{m1} & A_{m2} & \cdots & \cdots & A_{mn} \end{pmatrix} \end{array}$$

This defines a linear map $\mathbb{R}^n \to \mathbb{R}^m$

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & \cdots & A_{2n} \\ \vdots & \vdots & & & \vdots \\ A_{m1} & A_{m2} & \cdots & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} A_{11}x_1 + \cdots + A_{1n}x_n \\ A_{21}x_1 + \cdots + A_{2n}x_n \\ \vdots \\ A_{m1}x_1 + \cdots + A_{mn}x_n \end{pmatrix}$$

Matrix multiplication

Back to 2-by-2

Given $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, and $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ we get

Back to 2-by-2

Given $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, and $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ we get

$$A(B(v)) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Back to 2-by-2

Given $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, and $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ we get

$$\begin{aligned} A(B(v)) &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &\vdots \end{aligned}$$

Back to 2-by-2

Given $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, and $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ we get

$$\begin{aligned} A(B(v)) &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &\vdots \\ &= \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Definition (Matrix multiplication: 2-by-2)

Define

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

Definition (Matrix multiplication: 2-by-2)

Define

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

Equivalently,

$$(AB)_{ij} = \sum_{k=1}^{2} A_{ik}B_{kj}.$$

Definition (Matrix multiplication: 2-by-2)

Define

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

Equivalently,

$$(AB)_{ij} = \sum_{k=1}^{2} A_{ik}B_{kj}.$$

Definition (Matrix multiplication)

Given an m-by-n matrix A and an n-by-p matrix B, define AB to be the m-by-p matrix with entries

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}.$$

Definition (Matrix powers)

If A is a square matrix and k is a positive integer then A^k denotes the product $AA \cdots A$ (k times, where $k = 0$ means $A = I$).

Definition (Matrix powers)

If A is a square matrix and k is a positive integer then A^k denotes the product $AA \cdots A$ (k times, where $k = 0$ means $A = I$).

Definition (Matrix sum)

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

Definition (Matrix powers)

If A is a square matrix and k is a positive integer then A^k denotes the product $AA\cdots A$ (k times, where $k = 0$ means $A = I$).

Definition (Matrix sum)

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

Definition (Matrix scaling)

$$(\lambda A)_{ij} = \lambda A_{ij}.$$

Matrix exponential

Matrix exponential

Definition (Matrix exponential)

Given an n-by-n matrix A, define

$$\exp(A) = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{n \geq 0} \frac{1}{n!}A^n.$$

Dot products & orthogonal matrices

Definition

Given vectors $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ and $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ define the *dot product*

$$v \cdot w := v_1 w_1 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i.$$

Definition

Given vectors $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ and $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ define the *dot product*

$$v \cdot w := v_1 w_1 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i.$$

Theorem (Proof later!)

If v and w make an angle ϕ then

$$v \cdot w = |v||w| \cos \phi.$$

Note that

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}.$$

Note that

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}.$$

Definition

Define the *transpose* A^T of an m-by-n matrix A to be the n-by-m matrix with entries $(A^T)_{ij} = A_{ji}$.

So $v \cdot w = v^T w$.

Lemma

$$(AB)^T = B^T A^T.$$

Lemma

$$(AB)^T = B^T A^T.$$

Definition (Orthogonal matrix)

A square matrix A is *orthogonal* if $A^T A = I$.

Lemma

$$(AB)^T = B^T A^T.$$

Definition (Orthogonal matrix)

A square matrix A is *orthogonal* if $A^T A = I$.

Lemma

If A is orthogonal then $(Av) \cdot (Aw) = v \cdot w$.

i.e. orthogonal matrices preserve lengths and angles.

3-d rotations

Example

$$A = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

Example

$$D = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

- Find axis u: solve $Du = u$.
- Find angle: pick $v \perp u$ and compute angle between v and Dv.

Example

$$E = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} + \frac{1}{\sqrt{3}} & -\frac{1}{3} + \frac{1}{\sqrt{3}} \\ \frac{1}{3} - \frac{1}{\sqrt{3}} & \frac{1}{3} & -\frac{1}{3} - \frac{1}{\sqrt{3}} \\ -\frac{1}{3} - \frac{1}{\sqrt{3}} & -\frac{1}{3} + \frac{1}{\sqrt{3}} & \frac{1}{3} \end{pmatrix}.$$

- Find axis u: solve $Eu = u$.
- Find angle: pick $v \perp u$ and compute angle between v and Ev.

We saw earlier that $\exp\begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. In fact...

Theorem

$\exp(tA)$ *is orthogonal for all* $t \in \mathbb{R}$ *if and only if* A *is* antisymmetric, *i.e.* $A^T = -A$.

Proof.

If $A^T = -A$ then

$$(\exp(tA))^T = \exp(tA^T) = \exp(-tA).$$

$$\exp(-tA)\exp(tA) = I,$$

so $\exp(tA)$ is orthogonal. Conversely...

□

Theorem

$\exp(tA)$ *is orthogonal for all* $t \in \mathbb{R}$ *if and only if* A *is* antisymmetric, *i.e.* $A^T = -A$.

Proof.

If $\exp(tA)$ is orthogonal, then $\exp(tA)^T \exp(tA) = I$ for all t.
Differentiate with respect to t:

$$A^T \exp(tA^T) \exp(tA) + \exp(tA^T) A \exp(tA) = 0,$$

and set $t = 0$:

$$A^T + A = 0, \qquad \text{as } \exp(0A) = I.$$

$\square$

Need to show:

$$\begin{aligned}
\exp(B)^T &= \exp(B^T) \\
\exp(-B)\exp(B) &= I \\
\frac{d}{dt}\exp(tA) &= A\exp(tA) \\
\frac{d}{dt}(A(t)B(t)) &= \frac{dA(t)}{dt}B(t) + A(t)\frac{dB(t)}{dt}.
\end{aligned}$$