

Strang

Malcolm

Started 5th June 2025

Contents

1	Vectors and Matrices	2
1.1	Intuition for Dot product, Cosine formula, Schwarz and Triangle inequalities	2
1.2	Intuition for column rank being equal to row rank	4
1.3	7

Chapter 1

Vectors and Matrices

1.1 Intuition for Dot product, Cosine formula, Schwarz and Triangle inequalities

Intuition for dot product

The unit vectors $\mathbf{v} = (\cos \alpha, \sin \alpha)$ and $\mathbf{w} = (\cos \beta, \sin \beta)$ are plotted as follows

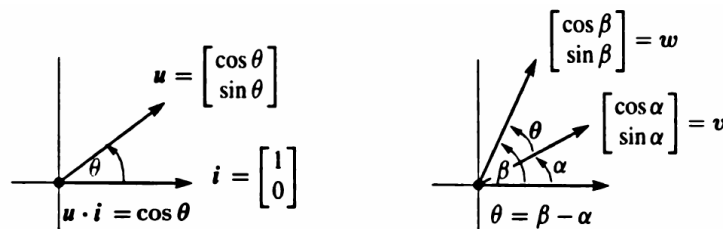


Figure 1.5: Unit vectors: $\mathbf{u} \cdot \mathbf{i} = \cos \theta$. The angle between the vectors is θ .

See first that when fixed in this form, the magnitude of both vectors is 1, with an angle $\beta - \alpha$ between them. These unit vectors have dot product

$$\mathbf{v} \cdot \mathbf{w} = \cos \alpha \cos \beta + \sin \alpha \sin \beta = \cos(\beta - \alpha)$$

We have θ as the angle between the two vectors; see that the sign of $\mathbf{v} \cdot \mathbf{w}$ tells us whether θ is below or above a right angle (due to the cosine function being negative for its argument $> \pi/2$ and positive for $< \pi/2$):

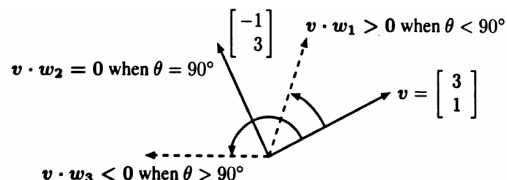


Figure 1.6: Small angle $\mathbf{v} \cdot \mathbf{w}_1 > 0$. Right angle $\mathbf{v} \cdot \mathbf{w}_2 = 0$. Large angle $\mathbf{v} \cdot \mathbf{w}_3 < 0$.

(next page)

Cont.

The idea here is that the dot product reveals the exact angle θ ; for unit vectors \mathbf{u} and \mathbf{U} , the dot product $\mathbf{u} \cdot \mathbf{U}$ is the cosine of θ . This remains true in n dimensions (not shown).

See that any \mathbf{u} and \mathbf{v} can be fixed in the above form by normalising their lengths to get $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$ and $\mathbf{U} = \mathbf{w}/\|\mathbf{w}\|$. After which their dot product would give $\cos \theta$. This leads us to the *cosine formula*:

$$\text{Cosine formula: } \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|} = \cos \theta \quad \text{if } \mathbf{v} \text{ and } \mathbf{w} \text{ are nonzero vectors}$$

Perpendicular vectors

See that when the angle between \mathbf{v} and \mathbf{w} is 90° , its cosine is 0; this gives us a way to test this. Also see that for perpendicular vectors:

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$$

because

$$\|\mathbf{v} + \mathbf{w}\|^2 = (\mathbf{v} + \mathbf{w}) \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{v} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w}$$

where $\mathbf{v} \cdot \mathbf{w} = 0$.

Schwarz and Triangle inequalities

First, see from the cosine formula that the dot product of $\mathbf{v}/\|\mathbf{v}\|$ and $\mathbf{w}/\|\mathbf{w}\|$ never exceeds one (since $\cos \theta$ never exceeds one). This is the *Schwarz inequality*:

$$\text{Schwarz inequality: } |\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|$$

The *Triangle inequality* comes directly from the Schwarz inequality:

$$\text{Triangle inequality: } \|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$$

This can be seen from

$$\|\mathbf{v} + \mathbf{w}\|^2 = \mathbf{v} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} \leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\| \|\mathbf{w}\| + \|\mathbf{w}\|^2$$

The square root gives us the triangle equality (side 3 cannot exceed side 1 + side 2).

1.2 Intuition for column rank being equal to row rank

If all columns are in the same direction, why does it happen that all the rows are the same direction?

Consider the matrix, see that column 2 is m times column 1:

$$\mathbf{A} = \begin{bmatrix} a & ma \\ b & mb \end{bmatrix}$$

See that the second row is just b/a times the first row—if the column rank is 1, then the row rank is 1. See that transposing the matrix, we have

$$\mathbf{A} = \begin{bmatrix} a(1) & b(1) \\ a(m) & b(m) \end{bmatrix}$$

which still has one independent column. Now consider the 3x3 case:

$$\mathbf{A} = \begin{bmatrix} a & ma & pa \\ b & mb & pb \\ c & mc & pc \end{bmatrix}$$

See that a similar deduction can also be made in this case, where the row rank of A is equal to its column rank.

(next page)

An informal proof

Consider any matrix \mathbf{A} , suppose we go from left to right, looking for independent columns of \mathbf{A} using the following procedure:

1. If column 1 of \mathbf{A} is not zero, put it in matrix \mathbf{C}
2. If column 2 of \mathbf{A} is not a multiple of column 1, put it in into \mathbf{C}
3. If column 3 of \mathbf{A} is not a combination of columns 1 and 2, put it into \mathbf{C} .
continue

See that at the end \mathbf{C} will have r columns taken from \mathbf{A} , where r is the rank of \mathbf{A} and \mathbf{C} . While the n columns of \mathbf{A} are dependent, the r columns of \mathbf{C} will surely be independent.

For instance consider \mathbf{A} with rank 2

$$\mathbf{A} = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ 1 & 3 & 5 \end{bmatrix} \quad \text{leads to} \quad \mathbf{C} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \\ 1 & 5 \end{bmatrix}$$

Now consider another matrix \mathbf{R} to be multiplied by \mathbf{C} such that $\mathbf{A} = \mathbf{C}\mathbf{R}$. The first and third columns of \mathbf{A} are already in \mathbf{C} , so those respective columns in \mathbf{R} make up a *identity matrix*; the second column of \mathbf{A} is a multiple of the first, so we have

$$\mathbf{A} = \mathbf{C}\mathbf{R} \quad \text{is} \quad \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(See that the i th row of \mathbf{A} can be seen as a linear combination of the rows of \mathbf{R} specified the i th row of \mathbf{C} . (or just consider $\mathbf{A}^T = \mathbf{R}^T \mathbf{C}^T$). We know that

1. \mathbf{C} contains the full set of r independent columns of \mathbf{A} .
2. $\mathbf{R} = [\mathbf{I} \mathbf{F}]$ contains the identity matrix \mathbf{I} in the same r columns that held \mathbf{C} .
3. The dependent columns of \mathbf{A} are combinations of $\mathbf{C}\mathbf{F}$ of the independent columns in \mathbf{C} .

Where the matrix \mathbf{F} goes into the other $n - r$ columns of $\mathbf{R} = [\mathbf{I} \mathbf{F}]$. ($\mathbf{A} = \mathbf{C}\mathbf{R}$ becomes $\mathbf{A} = \mathbf{C}[\mathbf{I}, \mathbf{F}] = [\mathbf{C}, \mathbf{C}\mathbf{F}] = [\text{indep cols of } \mathbf{A}, \text{ dep cols of } \mathbf{A}]$ (in correct order).

See that \mathbf{C} has the same column space as \mathbf{A} , and \mathbf{R} has the same row space as \mathbf{A} (every row of \mathbf{A} is a combination of the rows of \mathbf{R}).

(next page)

Cont.

We had the example

$$\mathbf{A} = \mathbf{C}\mathbf{R} \quad \text{is} \quad \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Here is an informal proof that the row rank of \mathbf{A} equals the column rank of \mathbf{A} (based from facts we already know)

1. The r columns of \mathbf{C} are independent (chosen that way from \mathbf{A})
2. Every column of \mathbf{A} is a combination of those r columns of \mathbf{C} (since $\mathbf{A} = \mathbf{C}\mathbf{R}$)
3. The r rows of \mathbf{R} are independent (they contain the r by r matrix \mathbf{I})
4. Every row of \mathbf{A} is a combination of the r rows of \mathbf{R}

See that for every column of \mathbf{A} that goes into \mathbf{C} , a column of \mathbf{I} goes into \mathbf{R} , where each column of \mathbf{I} in \mathbf{R} adds an independent row.

This means that the column rank of \mathbf{C} (column space of \mathbf{A}) is always equal to the row rank of \mathbf{R} (row space of \mathbf{A})—the column rank of \mathbf{A} is equal to the row rank of \mathbf{A} .

More examples

Rank 2:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

Rank 2:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Rank 1:

$$\begin{bmatrix} 1 & 2 & 10 & 100 \\ 3 & 6 & 30 & 300 \\ 2 & 4 & 20 & 200 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 10 & 100 \end{bmatrix}$$

1.3