

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 | Ways to multiply $\mathbf{AB} = \mathbf{C}$ | 7 |
| 1.4 | Solutions to $\mathbf{Ax} = \mathbf{b}$ | 8 |
| 1.5 | Elimination and Back Substitution | 9 |

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

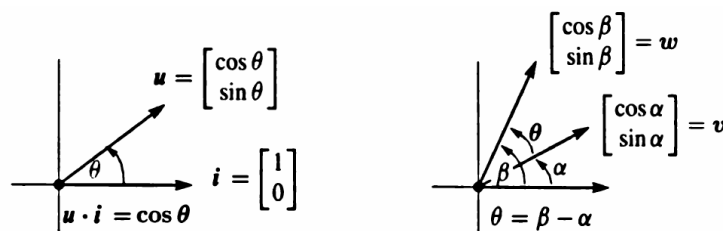


Figure 1.5: Unit vectors: $\mathbf{u} \cdot \mathbf{i} = \cos \theta$. The angle between the vectors is θ .

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 $< \pi/2$):

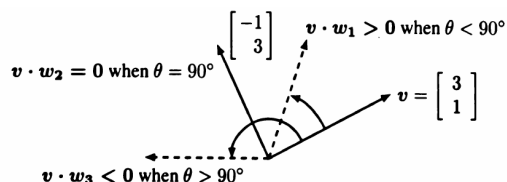


Figure 1.6: Small angle $\mathbf{v} \cdot \mathbf{w}_1 > 0$. Right angle $\mathbf{v} \cdot \mathbf{w}_2 = 0$. Large angle $\mathbf{v} \cdot \mathbf{w}_3 < 0$.

(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 Ways to multiply $AB = C$

Multiplication by columns of A and rows of B

A lesser known way to multiply AB is through considering the columns of A and the rows of B (contrary to the usual ideas where each entry of the result is a dot product of a row of A and column of B):

$$AB = \left[\begin{array}{c|c|c} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{array} \right] \left[\begin{array}{c} -b_1^* - \\ \vdots \\ -b_n^* - \end{array} \right] = a_1 b_1^* + a_2 b_2^* + \cdots + a_n b_n^*.$$

columns a_k rows b_k^* Add columns a_k times rows b_k^*

We multiply each column of A by each row of B ; this gives us n rank 1 matrices, which we then sum together; these matrices are called *outer products*

(usually we see the i th column of the result as a linear combination of the columns of A specified by the i th column of B . However in this case the i th outer product is a matrix of the same size as the result, that contains all the contributions of the i th column of A to the final product. By summing this over all n columns of A we get the result.)

Summary of methods

| (3 by 2)(2 by 4) = (3 by 4) | | Four Ways to Multiply $AB = C$ | |
|--|---|--|--|
| | $\left[\begin{array}{c} \text{grey row } i \\ x \ x \\ x \ x \end{array} \right] \left[\begin{array}{c} \text{grey column } k \\ x \ x \ x \\ x \ x \ x \end{array} \right]$ | (Row i of A) \cdot (Column k of B) = Number C_{ik} $i = 1 \text{ to } 3 \quad k = 1 \text{ to } 4 \quad 12 \text{ numbers}$ | |
| | $\left[\begin{array}{c} \text{grey column } k \\ x \ x \\ x \ x \end{array} \right] \left[\begin{array}{c} x \ x \ x \\ x \ x \ x \end{array} \right]$ | A times (Column k of B) = Column k of C $k = 1 \text{ to } 4 \quad 4 \text{ columns}$ | |
| | $\left[\begin{array}{c} \text{grey row } i \\ x \ x \\ x \ x \end{array} \right] \left[\begin{array}{c} \text{grey row } j \\ x \ x \ x \ x \\ x \ x \ x \ x \end{array} \right]$ | (Row i of A) times B = Row i of C $i = 1 \text{ to } 3 \quad 3 \text{ rows}$ | |
| | $\left[\begin{array}{c} \text{grey column } j \\ x \\ x \\ x \end{array} \right] \left[\begin{array}{c} \text{grey row } j \\ x \ x \ x \ x \end{array} \right]$ | (Column j of A) (Row j of B) = Rank 1 Matrix $j = 1 \text{ to } 2 \quad 2 \text{ matrices}$ | |
| Dot product way, Column way, Row way, Columns times rows | | | |

(A nice way to intuit the the third method is to consider $(AB)^T = B^T A^T$; where the columns of $(AB)^T$ are the rows of AB)

1.4 Solutions to $A\mathbf{x} = \mathbf{b}$

Given a $n \times n$ matrix A and an $n \times 1$ column vector \mathbf{b} , there are three outcomes for the vector \mathbf{x} that solves $A\mathbf{x} = \mathbf{b}$.

First there may be *no vector* \mathbf{x} that solves $A\mathbf{x} = \mathbf{b}$, or there may be exactly *one* solution, or there may be *infinitely many* solution vectors \mathbf{x} . Here are the possibilities:

1. **Exactly one solution** to $A\mathbf{x} = \mathbf{b}$ means that A has independent columns (only one particular linear combination of the columns of A leads to \mathbf{b} . That combination is specified by \mathbf{x}). A is full rank and the only solution to $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$. A has an inverse matrix A^{-1} (given \mathbf{b} , we can work backward to get \mathbf{x} since only one \mathbf{x} leads to \mathbf{b}).

2. **No solution** to $A\mathbf{x} = \mathbf{b}$ means that \mathbf{b} is not in the column space of A , so A is not full rank.

3. **Infinitely many solutions.** See that when the columns of A are not independent (not full rank), then there are infinitely many ways to produce the zero vector $\mathbf{b} = \mathbf{0}$ (this is the meaning of dependent columns), and so there are infinitely many solutions to $A\mathbf{x} = \mathbf{0}$.

Also see that if A is not full rank it means that its column space is some subspace, where solutions only exist for \mathbf{b} within that subspace.

As such, if there so happens to be a solution to $A\mathbf{x} = \mathbf{b}$ then we can add any solution to $A\mathbf{x} = \mathbf{0}$:

$$A(\mathbf{x} + \alpha\mathbf{x}) = A\mathbf{x} + \alpha A\mathbf{x} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

For some constant α , which gives us \mathbf{b} again—we have infinitely many solutions.

1.5 Elimination and Back Substitution

Elimination

We want to produce an *upper triangular* matrix \mathbf{U} from a square matrix \mathbf{A} . This is done through elimination; the procedure (for a 3x3 matrix) is as follows (assuming no row exchanges):

1. Use the first equation(row) to produce zeros in column 1 below the first pivot.
2. Use the new second equation(row) to clear out column 2 below pivot 2 in row 2.
3. *Continue to column 3.* The expected result is an upper triangular matrix \mathbf{U} .

These steps can be carried out using elimination matrices \mathbf{E} .

Consider \mathbf{A} and \mathbf{b}

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 11 & 14 \\ 2 & 8 & 17 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 19 \\ 55 \\ 50 \end{bmatrix}$$

\mathbf{E}_{21} multiplies equation 1 by 2 and subtracts that from equation 2 to get a zero in the first column below the first pivot:

$$\mathbf{E}_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \implies \mathbf{E}_{21}\mathbf{A} = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 2 & 8 & 17 \end{bmatrix}, \quad \mathbf{E}_{21}\mathbf{b} = \begin{bmatrix} 19 \\ 17 \\ 50 \end{bmatrix}$$

(For intuition on the elimination matrices, consider the row perspective of matrix multiplication) This produced the desired zero in column 1. It changed equation 2. To make the first pivot column zero, we subtract row 1 from row 3 using \mathbf{E}_{31}

$$\mathbf{E}_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \implies \mathbf{E}_{31}\mathbf{E}_{21}\mathbf{A} = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 0 & 5 & 13 \end{bmatrix}, \quad \mathbf{E}_{31}\mathbf{E}_{21}\mathbf{b} = \begin{bmatrix} 19 \\ 17 \\ 31 \end{bmatrix}$$

This completes elimination in column 1. Moving on to column 2 row 2 (the second pivot row). We use \mathbf{E}_{32} to subtract equation 2 from equation 3

$$\mathbf{E}_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \implies \mathbf{U} = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & 7 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 19 \\ 17 \\ 14 \end{bmatrix}$$

$\mathbf{E}_{32}\mathbf{E}_{31}\mathbf{E}_{21}\mathbf{A} = \mathbf{U}$ is triangular. See that the same steps were applied to the right hand side \mathbf{b} to produce a new right hand side \mathbf{c} .
(next page)

Possible breakdown of elimination

Elimination might fail. This occurs when zero appears in a pivot position—subtracting that zero from lower rows will not clear out the column below that pivot. For example:

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 6 & 14 \\ 2 & 8 & 17 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 3 & 4 \\ 0 & 0 & 6 \\ 0 & 5 & 13 \end{bmatrix} = \mathbf{B}$$

A possible way to get around this would be to *exchange* row 2 (with the zero pivot) for row (with the nonzero in that column), then carry out elimination as per usual. This exchange is carried out using the permutation matrix \mathbf{P} :

$$\mathbf{PB} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \\ 0 & 0 & 6 \\ 0 & 5 & 13 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 0 & 5 & 13 \\ 0 & 0 & 6 \end{bmatrix}$$

In this example the exchange produced \mathbf{U} with nonzero pivots; normally there may be more columns eliminate before \mathbf{U} is reached.

At times this exchange strategy may not work. This occurs when there is no pivot is available. Consider \mathbf{A}' :

$$\mathbf{A}' = \begin{bmatrix} 2 & 3 & 4 \\ 4 & 6 & 14 \\ 2 & 3 & 17 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 3 & 4 \\ 0 & 0 & 6 \\ 0 & 0 & 13 \end{bmatrix} = \mathbf{U}'$$

There is no second pivot in this case. This tells us that the matrix \mathbf{A}' *did not have full rank* (intuitively see that the non-pivot row can be expressed as a linear combination of the other pivot rows, so the row space (which is equal to the row space of the original \mathbf{A}' since the elimination steps are just linear combinations of the existing rows) is not full rank, and so the column space of the original \mathbf{A}' is also not full rank.

In this case there will be nonzero solutions \mathbf{X} to $\mathbf{A}'\mathbf{X} = 0$. The columns of \mathbf{U}' (and \mathbf{A}') are not independent.
(next page)

Back Substitution to solve $U\mathbf{x} = \mathbf{c}$
Elimination (ideally) produces an upper