COM4509/6509

Lecture 2a: Review of Vector Notation & End-to-End ML

Mike Smith* and Matt Ellis

*m.t.smith@sheffield.ac.uk

A scalar is just a numeric value. E.g. 3423.3, -41, etc.

Denoted with non-bold text: λ, x_2, N.

A scalar is just a numeric value. E.g. 3423.3, -41, etc.

Denoted with non-bold text: λ, x_2, N.

A vector is a list of scalar values, we refer to these as elements or entries. Usually use a lowercase **bold** letter. They can be column vectors:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{or} \quad \boldsymbol{\lambda} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Or row vectors:

$$\mathbf{z} = \begin{pmatrix} a & a^2 & a^3 & a^4 & \dots & a^N \end{pmatrix}$$

$$\boldsymbol{\theta} = \begin{pmatrix} 2 & 4 & 6 \end{pmatrix}$$

Or row vectors:

$$\mathbf{z} = \begin{pmatrix} a & a^2 & a^3 & a^4 & \dots & a^N \end{pmatrix}$$

$$\boldsymbol{\theta} = \begin{pmatrix} 2 & 4 & 6 \end{pmatrix}$$

We often want to say what dimensions (or shape) a vector or matrix is.

Or row vectors:

$$\mathbf{z} = \begin{pmatrix} a & a^2 & a^3 & a^4 & \dots & a^N \end{pmatrix}$$

$$\boldsymbol{\theta} = \begin{pmatrix} 2 & 4 & 6 \end{pmatrix}$$

We often want to say what dimensions (or shape) a vector or matrix is.

We report this as number of rows by number of columns. So $\boldsymbol{\theta}$ is "1×3".

(The vectors in this module will all be column vectors).

A **matrix** is a rectangular array of scalars. Usually written as **bold uppercase**:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \quad \mathbf{\Pi} = \begin{pmatrix} 2 & 5 & 9 & 12 \\ 1 & 3 & 6 & 10 \end{pmatrix}$$

A **matrix** is a rectangular array of scalars. Usually written as **bold uppercase**:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \quad \boldsymbol{\Pi} = \begin{pmatrix} 2 & 5 & 9 & 12 \\ 1 & 3 & 6 & 10 \end{pmatrix}$$

$\boldsymbol{A}$ has dimensions 3×2 (three rows, two columns).

A **matrix** is a rectangular array of scalars. Usually written as **bold uppercase**:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \quad \boldsymbol{\Pi} = \begin{pmatrix} 2 & 5 & 9 & 12 \\ 1 & 3 & 6 & 10 \end{pmatrix}$$

$\boldsymbol{A}$ has dimensions 3×2 (three rows, two columns).

Notice how we can write the entries as a_{ij}, the first index is the row, the second the column.

The transpose of a matrix $\boldsymbol{X}$ is simply swapping the rows for the columns, and is written $\boldsymbol{X}^\top$.

The transpose of a matrix $\boldsymbol{X}$ is simply swapping the rows for the columns, and is written $\boldsymbol{X}^\top$.

For example, if

$$\boldsymbol{X} = \begin{pmatrix} 2 & 3 \\ 4 & 7 \\ 0 & 0 \end{pmatrix}$$

Then its transpose is

$$\boldsymbol{X}^\top = \begin{pmatrix} 2 & 4 & 0 \\ 3 & 7 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$AB = \begin{pmatrix} \quad & \quad \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

1×3

$$AB = \begin{pmatrix} & \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$1 \times 3 + 2 \times 0$$

$$AB = \begin{pmatrix} \\ \\ \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$1 \times 3 + 2 \times 0 = 3$$

$$AB = \begin{pmatrix} 3 & \\ & \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$AB = \begin{pmatrix} 3 & \\ & \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$1 \times 1$$

$$AB = \begin{pmatrix} 3 & \\ & \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$1 \times 1 + 2 \times 5 = 11$$

$$AB = \begin{pmatrix} 3 & 11 \\ & \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$3 \times 3 + 4 \times 0$$

$$AB = \begin{pmatrix} 3 & 11 \\ 9 & \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$3 \times 1 + 4 \times 5$

$$AB = \begin{pmatrix} 3 & 11 \\ 9 & 23 \\ & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$5 \times 3 + 6 \times 0$$

$$AB = \begin{pmatrix} 3 & 11 \\ 9 & 23 \\ 15 & \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$5 \times 1 + 6 \times 5$$

$$AB = \begin{pmatrix} 3 & 11 \\ 9 & 23 \\ 15 & 35 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

$$5 \times 1 + 6 \times 5$$

$$AB = \begin{pmatrix} 3 & 11 \\ 9 & 23 \\ 15 & 35 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

3×2 $\qquad$ 2×2

$$AB = \begin{pmatrix} 3 & 11 \\ 9 & 23 \\ 15 & 35 \end{pmatrix}$$

3×2.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & 1 \\ 0 & 5 \end{pmatrix}$$

3×2 $\qquad$ 2×2

$$AB = \begin{pmatrix} 3 & 11 \\ 9 & 23 \\ 15 & 35 \end{pmatrix}$$

3×2

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 5 & 6 & 2 \\ 0 & 1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

3×3 $\qquad$ 3×1

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 5 & 6 & 2 \\ 0 & 1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

3×3 $\qquad$ 3×1

$$AB = \begin{pmatrix} 2 \times 1 + 1 \times 2 + 4 \times 3 \\ 5 \times 1 + 6 \times 2 + 2 \times 3 \\ 0 \times 1 + 1 \times 2 + 2 \times 3 \end{pmatrix}$$

$$\underset{3\times 3}{A} = \begin{pmatrix} 2 & 1 & 4 \\ 5 & 6 & 2 \\ 0 & 1 & 2 \end{pmatrix} \qquad \underset{3\times 1}{B} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$AB = \begin{pmatrix} 2\times 1 + 1\times 2 + 4\times 3 \\ 5\times 1 + 6\times 2 + 2\times 3 \\ 0\times 1 + 1\times 2 + 2\times 3 \end{pmatrix}$$

$$= \underset{3\times 1}{\begin{pmatrix} 16 \\ 23 \\ 8 \end{pmatrix}}$$

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 5 & 6 & 2 \\ 0 & 1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

3×3 Have to match 3×1

$$AB = \begin{pmatrix} 2\times1 + 1\times2 + 4\times3 \\ 5\times1 + 6\times2 + 2\times3 \\ 0\times1 + 1\times2 + 2\times3 \end{pmatrix}$$

$$= \begin{pmatrix} 16 \\ 23 \\ 8 \end{pmatrix}$$

3×1

Number of columns in A has to match number of rows in B.

Let $\boldsymbol{A}$ be a matrix with entries a_{ik} of dimension $p \times q$.

Let $\boldsymbol{A}$ be a matrix with entries a_{ik} of dimension $p \times q$.

Let $\boldsymbol{B}$ be a matrix with entries b_{kj} of dimension $t \times s$.

Let $\boldsymbol{A}$ be a matrix with entries a_{ik} of dimension $p \times q$.

Let $\boldsymbol{B}$ be a matrix with entries b_{kj} of dimension $t \times s$.

Matrix multiplication $\boldsymbol{AB}$ is only possible if $q = t$.

Let $\boldsymbol{A}$ be a matrix with entries a_{ik} of dimension $p \times q$.

Let $\boldsymbol{B}$ be a matrix with entries b_{kj} of dimension $t \times s$.

Matrix multiplication $\boldsymbol{AB}$ is only possible if $q = t$.

The matrix $\boldsymbol{C} = \boldsymbol{AB}$ has dimensions $p \times s$ and entries:

$$c_{ij} = \sum_{k} a_{ik} b_{kj}$$

There are lots of useful identities that let us quickly work with matrices. Lots of these are summarised in the matrix cookbook.

There are lots of useful identities that let us quickly work with matrices. Lots of these are summarised in the matrix cookbook.

For example, how to transpose a product...

$$(\boldsymbol{ABC})^\top = \boldsymbol{C}^\top \boldsymbol{B}^\top \boldsymbol{A}^\top.$$

Rewrite this without brackets: $((\boldsymbol{A}^\top \boldsymbol{B}\boldsymbol{C})^\top \boldsymbol{D}\boldsymbol{E}^\top)^\top$

Rewrite this without brackets: $((\boldsymbol{A}^\top \boldsymbol{B}\boldsymbol{C})^\top \boldsymbol{D}\boldsymbol{E}^\top)^\top$

If, $\boldsymbol{B}$ is 2×2; $\boldsymbol{C}$ is 2×4; and $\boldsymbol{D}$ is 5×3, what shape must $\boldsymbol{A}$ be? What shape is the result? What sort of vector is this?

Rewrite this without brackets: $((\boldsymbol{A}^\top \boldsymbol{B}\boldsymbol{C})^\top \boldsymbol{D}\boldsymbol{E}^\top)^\top$

If, $\boldsymbol{B}$ is 2×2; $\boldsymbol{C}$ is 2×4; and $\boldsymbol{D}$ is 5×3, what shape must $\boldsymbol{A}$ be? What shape is the result? What sort of vector is this?

If $\boldsymbol{X}$ is diagonal matrix, $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\boldsymbol{w}$ is $\begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$, what is $\boldsymbol{X}\boldsymbol{w}$?

An identity element in maths is a thing that doesn't have any effect on any variable for a given operator.

An identity element in maths is a thing that doesn't have any effect on any variable for a given operator.

For example 1 is the multiplicative identity for scalar values.

An identity element in maths is a thing that doesn't have any effect on any variable for a given operator.

For example 1 is the multiplicative identity for scalar values.

For matrix multiplication it is a square matrix with ones down the diagonal: The identity matrix.

An identity element in maths is a thing that doesn't have any effect on any variable for a given operator.

For example 1 is the multiplicative identity for scalar values.

For matrix multiplication it is a square matrix with ones down the diagonal: The identity matrix.

$$\boldsymbol{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The inverse of a matrix, is the matrix that one can multiply it by to get the identity.

The inverse of a matrix, is the matrix that one can multiply it by to get the identity.

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The inverse of a matrix, is the matrix that one can multiply it by to get the identity.

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} ? & \quad & ? \\ ? & \quad & ? \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

If the left matrix is $\boldsymbol{A}$ then the unknown is its inverse, $\boldsymbol{A}^{-1}$.

The inverse of a matrix, is the matrix that one can multiply it by to get the identity.

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} ? & & ? \\ ? & & ? \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

If the left matrix is $\boldsymbol{A}$ then the unknown is its inverse, $\boldsymbol{A}^{-1}$.

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$$

The inverse of a matrix, is the matrix that one can multiply it by to get the identity.

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The inverse of a matrix, is the matrix that one can multiply it by to get the identity.

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Activity: Two minutes for those who aren't used to matrices etc, check this is true!

- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are non-singular [i.e. they can be inverted] then: $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are non-singular [i.e. they can be inverted] then: $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$
- The inverse of a transpose is the transpose of the inverse: $(\boldsymbol{A}^{\top})^{-1} = (\boldsymbol{A}^{-1})^{\top}$

- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are non-singular [i.e. they can be inverted] then: $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$
- The inverse of a transpose is the transpose of the inverse: $(\boldsymbol{A}^{\top})^{-1} = (\boldsymbol{A}^{-1})^{\top}$
- Not all matrices can be inverted: Only square ones can, and they must have a positive determinant (non-singular).

For example, we can't find an inverse of, $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

There's no matrix we could multiply this by to get the identity matrix.