Matrix Multiplication

Last modified: 2026-11-24 09:19:45 (PST)

Matrix multiplication

Definition 0.1 (Matrix multiplication)**.** The **product** [1] of an $m \times k$ matrix $\mathbf{A}$ and a $k \times n$ matrix $\mathbf{B}$ is the $m \times n$ matrix $\mathbf{C} = \mathbf{AB}$ with entries:

$$c_{ij} = \sum_{s=1}^{k} a_{is}\, b_{sj}$$

Remark (Matrix multiplication is not commutative)

Remark 0.1 (Matrix multiplication is not commutative). When $\mathbf{AB}$ and $\mathbf{BA}$ are both defined, they are usually different. For example, with

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},$$

$$\mathbf{AB} = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$$

$$\mathbf{BA} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix},$$

so $\mathbf{AB} \neq \mathbf{BA}$.

Definition 0.2 (Shape rule)**.** The **shape rule** of matrix multiplication says that an $m \times k$ matrix times a $k \times n$ matrix is an $m \times n$ matrix, and that the multiplication is defined only when the inner pair of sizes matches:

$$(m \times k)(k \times n) \longrightarrow m \times n \tag{1}$$

Example 0.1 (Shape of a product of three matrices)**.**
Let $\mathbf{A}$ be 2×3, let $\mathbf{B}$ be 3×4, and let $\underset{\sim}{x} \in \mathbb{R}^4$, which is a 4×1 matrix. Apply the shape rule one product at a time:

- $\mathbf{AB}$ is $(2 \times 3)(3 \times 4)$. The inner pair $3, 3$ matches, so the product is 2×4.
- $(\mathbf{AB})\underset{\sim}{x}$ is $(2 \times 4)(4 \times 1)$. The inner pair $4, 4$ matches, so the product is 2×1, a vector in $\mathbb{R}^2$.

Example 0.2 (A product that the shape rule rules out)**.**
Let $\mathbf{A}$ be 2×3, and let $\mathbf{B}$ be 3×4. The product $\mathbf{BA}$ is $(3 \times 4)(2 \times 3)$. Its inner pair $4, 2$ does not match, so the shape rule says that $\mathbf{BA}$ is not defined, although $\mathbf{AB}$ is.

Example 0.3 (Dot product as matrix multiplication)**.** The dot product of two column vectors $\underset{\sim}{x}$ and $\underset{\sim}{\beta}$ can be written as a matrix product of the row vector $\underset{\sim}{x}^\top$ with the column vector $\underset{\sim}{\beta}$:

$$\begin{aligned} \underset{\sim}{x} \cdot \underset{\sim}{\beta} &= \underset{\sim}{x}^\top \underset{\sim}{\beta} \\ &= [x_1,\ x_2,\ \ldots,\ x_p] \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \\ &= x_1\beta_1 + x_2\beta_2 + \cdots + x_p\beta_p \end{aligned}$$

Exercise 0.1 (Computing the product of a 2×3 and a 3×2 matrix)**.** The product of two matrices is defined only when the first has as many columns as the second has rows. Each entry of the product combines one row of the first factor with one column of the second. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 3 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & -1 \\ 1 & 4 \end{bmatrix}.$$

Find $\mathbf{AB}$.

Hint

$\mathbf{A}$ is 2×3, and $\mathbf{B}$ is 3×2. By Definition 0.1, $\mathbf{AB}$ is therefore 2×2. Its entry c_{ij} is $a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j}$.

Solution

Solution 0.1. **A** has three columns, and **B** has three rows. So the product is defined and is 2×2. Its entries are

$$\mathbf{AB} = \begin{bmatrix} 1\cdot 2 + 0\cdot 0 + 2\cdot 1 & 1\cdot 1 + 0\cdot(-1) + 2\cdot 4 \\ (-1)\cdot 2 + 3\cdot 0 + 1\cdot 1 & (-1)\cdot 1 + 3\cdot(-1) + 1\cdot 4 \end{bmatrix} \quad \text{(Definition 0.1)}$$

$$= \begin{bmatrix} 4 & 9 \\ -1 & 0 \end{bmatrix}. \quad \text{(multiply and add)}$$

Theorem 0.1 (Matrix multiplication is associative)**.** *For an $m \times k$ matrix* **A***, a $k \times l$ matrix* **B***, and an $l \times n$ matrix* **C***:*

$$\underbrace{(\mathbf{AB})\mathbf{C}}_{m\times n} = \underbrace{\mathbf{A}(\mathbf{BC})}_{m\times n}$$

Proof

Proof. Entry (i, j) of each side, from Definition 0.1:

$$\begin{aligned}
[(\mathbf{AB})\mathbf{C}]_{ij} &= \sum_{t=1}^{l} (\mathbf{AB})_{it}\, c_{tj} && \text{(definition of } (\mathbf{AB})\mathbf{C}) \\
&= \sum_{t=1}^{l} \left(\sum_{s=1}^{k} a_{is}\, b_{st} \right) c_{tj} && \text{(definition of } \mathbf{AB}) \\
&= \sum_{s=1}^{k} a_{is} \left(\sum_{t=1}^{l} b_{st}\, c_{tj} \right) && \text{(distribute and swap the finite sums)} \\
&= \sum_{s=1}^{k} a_{is}\, (\mathbf{BC})_{sj} && \text{(definition of } \mathbf{BC}) \\
&= [\mathbf{A}(\mathbf{BC})]_{ij} && \text{(definition of } \mathbf{A}(\mathbf{BC}))
\end{aligned}$$

□

Exercise 0.2 (Grouping a product of three matrices)**.** By Theorem 0.1, a product of three matrices can be grouped either way, so $\mathbf{AB}\underset{\sim}{x}$ can be computed as $\mathbf{A}(\mathbf{B}\underset{\sim}{x})$, without ever forming **AB**. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \qquad \underset{\sim}{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

where $\underset{\sim}{x}$ is a 2×1 matrix.

Find $\mathbf{AB}\underset{\sim}{x}$ by computing $\mathbf{B}\underset{\sim}{x}$ first.

Hint

By Definition 0.1, $\mathbf{B}\underset{\sim}{x}$ is a 2×1 matrix; multiply $\mathbf{A}$ by it.

Solution

Solution 0.2. First,

$$\begin{aligned} \mathbf{B}\underset{\sim}{x} &= \begin{bmatrix} 2 \cdot 1 + 0 \cdot 2 \\ 1 \cdot 1 + 1 \cdot 2 \end{bmatrix} && \text{(Definition 0.1)} \\ &= \begin{bmatrix} 2 \\ 3 \end{bmatrix}, && \text{(multiply and add)} \end{aligned}$$

so

$$\begin{aligned} \mathbf{AB}\underset{\sim}{x} &= \mathbf{A}(\mathbf{B}\underset{\sim}{x}) && \text{(Theorem 0.1)} \\ &= \begin{bmatrix} 1 \cdot 2 + 1 \cdot 3 \\ 0 \cdot 2 + 1 \cdot 3 \end{bmatrix} && \text{(Definition 0.1)} \\ &= \begin{bmatrix} 5 \\ 3 \end{bmatrix}. && \text{(multiply and add)} \end{aligned}$$

Grouping the other way gives the same answer: $\mathbf{AB} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$, and $(\mathbf{AB})\underset{\sim}{x} = \begin{bmatrix} 3+2 \\ 1+2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$.

Theorem 0.2 (Matrix multiplication is distributive over addition)**.** *Matrix multiplication distributes over addition from the left and from the right. For an $m \times k$ matrix $\mathbf{A}$ and $k \times n$ matrices $\mathbf{B}$ and $\mathbf{C}$:*

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

For $m \times k$ matrices $\mathbf{A}$ and $\mathbf{B}$ and a $k \times n$ matrix $\mathbf{C}$:

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$

Exercise 0.3 (Proving that matrix multiplication distributes from the left)**.** Theorem 0.2 lets a product with a matrix sum be expanded term by term, just as multiplication of numbers distributes over addition. Let $\mathbf{A}$ be an $m \times k$ matrix, and let $\mathbf{B}$ and $\mathbf{C}$ be $k \times n$ matrices.

Prove that $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.

Hint

Express the (i,j) entry of the left side with Definition 0.1 and Definition 8 in Matrices, then split the result with the rule for a sum of sums.

Solution

Solution 0.3. Both sides are $m \times n$ matrices, so they are equal when their (i,j) entries agree for every i and j. For each i and j,

$$\begin{aligned}
[\mathbf{A}(\mathbf{B}+\mathbf{C})]_{ij} &= \sum_{s=1}^{k} a_{is}(\mathbf{B}+\mathbf{C})_{sj} && \text{(Definition 0.1)} \\
&= \sum_{s=1}^{k} a_{is}(b_{sj}+c_{sj}) && \text{(Definition 8 in Matrices)} \\
&= \sum_{s=1}^{k} (a_{is}b_{sj} + a_{is}c_{sj}) && \text{(multiplication of numbers distributes over addition)} \\
&= \sum_{s=1}^{k} a_{is}b_{sj} + \sum_{s=1}^{k} a_{is}c_{sj} && \text{(a sum of sums splits)} \\
&= (\mathbf{AB})_{ij} + (\mathbf{AC})_{ij} && \text{(Definition 0.1)} \\
&= (\mathbf{AB}+\mathbf{AC})_{ij}. && \text{(Definition 8 in Matrices)}
\end{aligned}$$

The third step uses that multiplication of numbers distributes over addition, and the fourth uses the rule for a sum of sums. So $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.

Theorem 0.3 (Product of two sums of matrices)**.** *For $m \times k$ matrices $\mathbf{A}$ and $\mathbf{B}$ and $k \times n$ matrices $\mathbf{C}$ and $\mathbf{D}$,*

$$(\mathbf{A}+\mathbf{B})(\mathbf{C}+\mathbf{D}) = \mathbf{AC}+\mathbf{AD}+\mathbf{BC}+\mathbf{BD}.$$

Each product keeps its factors in the order they appear on the left, because matrix multiplication is not commutative. So $\mathbf{AD}$ and $\mathbf{BC}$ are not collected into $2\mathbf{AD}$.

Proof

Proof.

$$\begin{aligned}(\mathbf{A}+\mathbf{B})(\mathbf{C}+\mathbf{D}) &= (\mathbf{A}+\mathbf{B})\mathbf{C}+(\mathbf{A}+\mathbf{B})\mathbf{D} && \text{(Theorem 0.2, from the left)}\\ &= \mathbf{AC}+\mathbf{BC}+\mathbf{AD}+\mathbf{BD} && \text{(Theorem 0.2, from the right, twice)}\\ &= \mathbf{AC}+\mathbf{AD}+\mathbf{BC}+\mathbf{BD} && \text{(matrix addition is commutative)}\end{aligned}$$

□

Example 0.4 (Why the middle terms do not combine)**.** Take $\mathbf{A}=\mathbf{C}=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$ and $\mathbf{B}=\mathbf{D}=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}$. Then $\mathbf{A}+\mathbf{B}=\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}$, so

$$\begin{aligned}(\mathbf{A}+\mathbf{B})^2 &= \begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix},\\ \mathbf{A}^2+\mathbf{AB}+\mathbf{BA}+\mathbf{B}^2 &= \begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}+\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix}+\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}+\begin{bmatrix}1 & 0\\ 2 & 1\end{bmatrix}\\ &= \begin{bmatrix}5 & 4\\ 4 & 5\end{bmatrix}.\end{aligned}$$

Here $\mathbf{AB}\neq\mathbf{BA}$, so the middle terms do not combine into $2\mathbf{AB}$. The squared norm of a sum of vectors, $\|\underset{\sim}{x}+\underset{\sim}{y}\|^2$, is a case where they do combine (Theorem 2 in Inner Products and the Cauchy-Schwarz Inequality), because $\underset{\sim}{x}\cdot\underset{\sim}{y}=\underset{\sim}{y}\cdot\underset{\sim}{x}$. Matrices that commute, such as $\mathbf{A}$ and $\mathbf{A}^2$, combine the same way.

Exercise 0.4 (Comparing $(\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B})$ with the rule for numbers)**.** For numbers, $(a+b)^2=a^2+2ab+b^2$. Theorem 0.3 expands a product of two matrix sums into four terms instead, because $\mathbf{AB}$ and $\mathbf{BA}$ can differ. Let

$$\mathbf{A}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix},\qquad \mathbf{B}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}.$$

Is $(\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B})$ equal to $\mathbf{AA}+2\mathbf{AB}+\mathbf{BB}$?

Hint

Compute $\mathbf{A}+\mathbf{B}$ and its product with itself, then compute $\mathbf{AA}$, $\mathbf{AB}$ and $\mathbf{BB}$ with Definition 0.1.

Solution

Solution 0.4. No. On the left, $\mathbf{A}+\mathbf{B}=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$, so

$$(\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B}) = \begin{bmatrix}1\cdot 1+1\cdot 0 & 1\cdot 1+1\cdot 0\\ 0\cdot 1+0\cdot 0 & 0\cdot 1+0\cdot 0\end{bmatrix} \quad \text{(Definition 0.1)}$$
$$= \begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}. \quad \text{(multiply and add)}$$

On the right, the same rule gives $\mathbf{AA}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, $\mathbf{AB}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$ and $\mathbf{BB}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$, so

$$\mathbf{AA}+2\mathbf{AB}+\mathbf{BB}=\begin{bmatrix}1 & 2\\ 0 & 0\end{bmatrix}.$$

The entries in position $(1,2)$ differ, $1\neq 2$, so the two matrices are not equal. The rule for numbers fails because $\mathbf{BA}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$ differs from $\mathbf{AB}$: Theorem 0.3 gives $\mathbf{AA}+\mathbf{AB}+\mathbf{BA}+\mathbf{BB}=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$, which does match the left side.

Exercise 0.5 (Which of these products are defined?)**.** Let $\mathbf{A}$ be 4×3, let $\mathbf{B}$ be 3×2, and let $\underset{\sim}{x}\in\mathbb{R}^3$. For each expression, say whether it is defined, and if so give the shape of the result:

1. $\mathbf{AB}$
2. $\mathbf{BA}$
3. $\mathbf{A}\underset{\sim}{x}$
4. $\underset{\sim}{x}^{\top}\underset{\sim}{x}$
5. $\underset{\sim}{x}\underset{\sim}{x}^{\top}$

Solution

Solution 0.5. A vector in $\mathbb{R}^3$ is a 3×1 matrix, so the same rule settles every case.

1. $\mathbf{AB}$: $(4\times 3)(3\times 2)$. The inner pair is 3 and 3, so it is defined, and the result is 4×2.
2. $\mathbf{BA}$: $(3\times 2)(4\times 3)$. The inner pair is 2 and 4, which do not match, so it is **not defined**. Matrix multiplication is not commutative, and this is the blunt form of that: swapping the order can leave an expression that means nothing.
3. $\mathbf{A}\underset{\sim}{x}$: $(4\times 3)(3\times 1)$, defined, result 4×1 — a vector in $\mathbb{R}^4$.
4. $\underset{\sim}{x}^{\top}\underset{\sim}{x}$: $(1\times 3)(3\times 1)$, defined, result 1×1 — a single number. The product is the inner product of Equation 1 in Vectors.

5. $\underset{\sim}{x}\underset{\sim}{x}^\top$: $(3 \times 1)(1 \times 3)$, defined, result 3×3 — a matrix.

The last two use the same two vectors and differ only in order, and they return objects of different kinds. Reading a transpose as decoration rather than as a shape change is the commonest way to lose track of an expression.

Square and identity matrices

Definition 0.1 (Square matrix)**.** A matrix is **square** [2] if it has the same number of rows as columns.

Example 0.1 (Square and not square)**.** $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is square (2 rows, 2 columns); $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ is not (2 rows, 3 columns).

Exercise 0.1 (Whether a data matrix times its transpose is square)**.** Many statistical methods start from an $n \times p$ data matrix $\mathbf{X}$, with one row per observation, and form the product $\mathbf{X}^\top\mathbf{X}$ of its transpose with itself. Whether that product is square matters, because some operations, such as taking the inverse of a matrix, apply only to square matrices.
Is $\mathbf{X}^\top\mathbf{X}$ square?

> **Hint**
>
> Find the shape of $\mathbf{X}^\top$ from Definition 6 in Matrices, then the shape of the product from Definition 0.1.

Solution

Solution 0.1. Yes. The transpose $\mathbf{X}^\top$ is $p \times n$, so $\mathbf{X}^\top$ has n columns. $\mathbf{X}$ has n rows, so the product $\mathbf{X}^\top\mathbf{X}$ is defined and is $p \times p$. It has p rows and p columns, so it is square for every n and p.

Definition 0.2 (Main diagonal)**.** The **main diagonal** [3] of an $m \times n$ matrix $\mathbf{A}$ is the list of its entries $a_{11}, a_{22}, \ldots, a_{kk}$, where k is the smaller of m and n.

Definition 0.3 (Diagonal entry)**.** The entries on the main diagonal (Definition 0.2) of a matrix $\mathbf{A}$ are the **diagonal entries** of $\mathbf{A}$.

Definition 0.4 (Off-diagonal entry)**.** In an $m \times n$ matrix $\mathbf{A}$ with main diagonal (Definition 0.2), every entry a_{ij} with $i \neq j$ is an **off-diagonal entry**.

Example 0.2 (Diagonal entries of a square and a non-square matrix)**.**

- For $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, the diagonal entries are $a_{11} = 1$ and $a_{22} = 4$, and the off-diagonal entries are $a_{12} = 2$ and $a_{21} = 3$.
- For $\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, $m = 2$ and $n = 3$, so $k = 2$, and the diagonal entries are $b_{11} = 1$ and $b_{22} = 5$; the other four entries, 2, 3, 4 and 6, are off-diagonal.

Exercise 0.2 (The main diagonal of a tall matrix)**.** The main diagonal of a matrix that is not square stops when it runs out of rows or columns, whichever comes first. A tall matrix, with more rows than columns, runs out of columns first. Let

$$\mathbf{E} = \begin{bmatrix} 7 & -2 \\ 1 & 3 \\ 0 & 5 \\ 4 & -6 \end{bmatrix}.$$

Find the main diagonal of $\mathbf{E}$.

Hint

Find m, n and k in Definition 0.2.

Solution

Solution 0.2. Write e_{ij} for the entry of $\mathbf{E}$ in row i and column j. $\mathbf{E}$ is 4×2, so $m = 4$, $n = 2$ and $k = 2$. The main diagonal of $\mathbf{E}$ is therefore the list e_{11}, e_{22}, that is, $7, 3$. The list stops at e_{22} because $\mathbf{E}$ has no third column. It therefore has no entry e_{33}.

Exercise 0.3 (Adding the diagonal entries of a square matrix)**.** The sum of the diagonal entries of a square matrix summarizes that matrix in one number, called its trace. Let

$$\mathbf{F} = \begin{bmatrix} 2 & 5 & -1 \\ 0 & -3 & 8 \\ 4 & 1 & 6 \end{bmatrix}.$$

Find the sum of the diagonal entries of $\mathbf{F}$.

Hint

Find the main diagonal of $\mathbf{F}$ first.

Solution

Solution 0.3. Write f_{ij} for the entry of $\mathbf{F}$ in row i and column j. $\mathbf{F}$ is 3×3, so its diagonal entries are $f_{11} = 2$, $f_{22} = -3$ and $f_{33} = 6$, and adding them gives

$$f_{11} + f_{22} + f_{33} = 2 + (-3) + 6 = 5.$$

Exercise 0.4 (Counting the off-diagonal entries of a square matrix)**.** In a table of pairwise comparisons among several items, the off-diagonal entries are the comparisons between two different items, so counting them counts the ordered pairs of different items. Let $\mathbf{A}$ be a 5×5 matrix, so $\mathbf{A}$ is square.
How many off-diagonal entries does $\mathbf{A}$ have?

Hint

Count all the entries, then subtract the diagonal entries.

Solution

Solution 0.4. $\mathbf{A}$ has $5 \cdot 5 = 25$ entries. Its diagonal entries are the 5 entries a_{ii}, $i = 1, \ldots, 5$, and every other entry has $i \neq j$, so $\mathbf{A}$ has $25 - 5 = 20$ off-diagonal entries.

Definition 0.5 (Identity matrix)**.** The $p \times p$ **identity matrix** [4] $\mathbf{I}_p$ (or $\mathbf{I}$ when the size is clear from context) has ones on the main diagonal and zeros elsewhere:

$$(\mathbf{I}_p)_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \qquad \mathbf{I}_p = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Example 0.3 (The 3×3 identity)**.**

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

so, for instance, $(\mathbf{I}_3)_{22} = 1$ and $(\mathbf{I}_3)_{23} = 0$.

Exercise 0.5 (Whether a matrix with the identity pattern is an identity matrix)**.** An identity matrix has ones on its main diagonal and zeros elsewhere. The matrix below has the same pattern of ones and zeros. Let

$$\mathbf{J} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Is $\mathbf{J}$ an identity matrix?

Hint

Compare the shape of $\mathbf{J}$ with the shape of $\mathbf{I}_p$ in Definition 0.5.

Solution

Solution 0.5. No. Every identity matrix $\mathbf{I}_p$ is $p \times p$, so it is square. $\mathbf{J}$ is 2×3, so it is not square. Therefore $\mathbf{J}$ is not an identity matrix, even though its main diagonal holds ones and its other entries are zeros.

Remark (The identity matrix and the Kronecker delta)

Remark 0.1 (The identity matrix and the Kronecker delta). The entries of the identity matrix are given by the Kronecker delta: $(\mathbf{I}_p)_{ij} = \delta_{ij}$.

Theorem 0.1 (Identity matrix is a multiplicative identity)**.** *For any $m \times p$ matrix $\mathbf{A}$:*

$$\mathbf{A}\,\mathbf{I}_p = \mathbf{A}$$

$$\mathbf{I}_m\,\mathbf{A} = \mathbf{A}$$

Exercise 0.6 (Multiplying by an identity matrix)**.** Theorem 0.1 says that an identity matrix of the right size leaves a matrix unchanged under matrix multiplication, just as

multiplying a number by 1 leaves it unchanged. Let

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 \\ 5 & 3 & 4 \end{bmatrix}.$$

Compute $\mathbf{I}_2\mathbf{A}$ entry by entry, from the definition of matrix multiplication.

Hint

Entry (i, j) of $\mathbf{I}_2\mathbf{A}$ is the dot product of the ith row of $\mathbf{I}_2$ with the jth column of $\mathbf{A}$, as in Example 0.3.

Solution

Solution 0.6. $\mathbf{I}_2$ is 2×2 and $\mathbf{A}$ is 2×3, so the product is defined and is 2×3. The first row of $\mathbf{I}_2$ is $\begin{bmatrix} 1 & 0 \end{bmatrix}$ and the second is $\begin{bmatrix} 0 & 1 \end{bmatrix}$, so by Definition 0.1

$$\begin{aligned} \mathbf{I}_2\mathbf{A} &= \begin{bmatrix} 1\cdot 2 + 0\cdot 5 & 1\cdot 0 + 0\cdot 3 & 1\cdot(-1) + 0\cdot 4 \\ 0\cdot 2 + 1\cdot 5 & 0\cdot 0 + 1\cdot 3 & 0\cdot(-1) + 1\cdot 4 \end{bmatrix} && \text{(Definition 0.1)} \\ &= \begin{bmatrix} 2 & 0 & -1 \\ 5 & 3 & 4 \end{bmatrix} && \text{(multiply and add)} \\ &= \mathbf{A}. \end{aligned}$$

Each row of $\mathbf{I}_2$ has a single 1, so entry (i, j) of the product picks out entry (i, j) of $\mathbf{A}$, as Theorem 0.1 says.

Matrix-vector multiplication

Definition 0.1 (Matrix-vector multiplication)**.** The **matrix-vector product** of an $m \times p$ matrix $\mathbf{A}$ and a $p \times 1$ column vector $\underset{\sim}{x}$ is the $m \times 1$ column vector $\mathbf{A}\underset{\sim}{x}$ with entries:

$$(\mathbf{A}\underset{\sim}{x})_i = \sum_{j=1}^{p} a_{ij}\, x_j$$

Remark (Each entry is a dot product)

Remark 0.1 (Each entry is a dot product). Matrix-vector multiplication extends the dot product (Definition 10 in Vectors): entry i of $\mathbf{A}\underset{\sim}{x}$ is the dot product of row i of $\mathbf{A}$ with $\underset{\sim}{x}$. For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 2 \cdot (-1) \\ 3 \cdot 1 + 4 \cdot (-1) \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}.$$

Exercise 0.1 (A matrix-vector product with more columns than rows)**.** A matrix-vector product $\mathbf{A}\underset{\sim}{x}$ computes one weighted sum of the entries of $\underset{\sim}{x}$ for each row of $\mathbf{A}$, so one multiplication computes several weighted sums at once. The matrix $\mathbf{A}$ below has more columns than rows, so $\mathbf{A}\underset{\sim}{x}$ is shorter than $\underset{\sim}{x}$. Let

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 \\ 5 & 3 & 4 \end{bmatrix}, \qquad \underset{\sim}{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Compute $\mathbf{A}\underset{\sim}{x}$.

Hint

Entry i of $\mathbf{A}\underset{\sim}{x}$ is the dot product of the ith row of $\mathbf{A}$ with $\underset{\sim}{x}$, as in Remark 0.1.

Solution

Solution 0.1. $\mathbf{A}$ is 2×3 and $\underset{\sim}{x}$ is 3×1, so the matrix-vector product is defined and is 2×1:

$$\begin{aligned} \mathbf{A}\underset{\sim}{x} &= \begin{bmatrix} 2 \cdot 1 + 0 \cdot 2 + (-1) \cdot 3 \\ 5 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 \end{bmatrix} && \text{(Definition 0.1)} \\ &= \begin{bmatrix} 2 + 0 - 3 \\ 5 + 6 + 12 \end{bmatrix} && \text{(multiply)} \\ &= \begin{bmatrix} -1 \\ 23 \end{bmatrix}. && \text{(add)} \end{aligned}$$

Definition 0.2 (Linear system, system of linear equations)**.** A **linear system** [5] (or **system of linear equations**) with m equations in n unknowns is an equation $\mathbf{A}\underset{\sim}{x} = \underset{\sim}{b}$ in which the $m \times n$ matrix $\mathbf{A}$ and the vector $\underset{\sim}{b} \in \mathbb{R}^m$ are given. A **solution** of the system is a vector $\underset{\sim}{x} \in \mathbb{R}^n$ for which $\mathbf{A}\underset{\sim}{x} = \underset{\sim}{b}$ holds.

Example 0.1 (Two equations in two unknowns)**.** By Definition 0.1, row i of $\mathbf{A}\underset{\sim}{x} = \underset{\sim}{b}$ is the equation $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$. So the two equations $x_1 + 2x_2 = 5$ and $3x_1 + 4x_2 = 6$

form the linear system $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} \underset{\sim}{x} = \begin{bmatrix}5\\ 6\end{bmatrix}$. The vector $\underset{\sim}{x} = (-4, 4.5)$ is a solution:

$$\begin{aligned}\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}-4\\ 4.5\end{bmatrix} &= \begin{bmatrix}1\cdot(-4) + 2\cdot 4.5\\ 3\cdot(-4) + 4\cdot 4.5\end{bmatrix} && \text{(Definition 0.1)}\\ &= \begin{bmatrix}-4+9\\ -12+18\end{bmatrix} && \text{(multiply)}\\ &= \begin{bmatrix}5\\ 6\end{bmatrix} && \text{(add)}\end{aligned}$$

The vector $(1, 2)$ is not a solution: its first equation holds, since $1 + 2 \cdot 2 = 5$, but its second gives $3 \cdot 1 + 4 \cdot 2 = 11 \neq 6$.

Exercise 0.2 (Checking whether a vector solves a linear system)**.** Checking whether a vector solves a linear system takes only one matrix-vector product, which is much less work than solving the system. Let

$$\mathbf{A} = \begin{bmatrix}3 & 1\\ 1 & -2\end{bmatrix}, \qquad \underset{\sim}{b} = \begin{bmatrix}5\\ 4\end{bmatrix}, \qquad \underset{\sim}{x} = \begin{bmatrix}2\\ -1\end{bmatrix}.$$

Is $\underset{\sim}{x}$ a solution of $\mathbf{A}\underset{\sim}{x} = \underset{\sim}{b}$?

Hint

Compute $\mathbf{A}\underset{\sim}{x}$ with Definition 0.1. Then compare it with $\underset{\sim}{b}$, as in Example 0.1.

Solution

Solution 0.2. Yes. The matrix-vector product is

$$\begin{aligned}\mathbf{A}\underset{\sim}{x} &= \begin{bmatrix}3\cdot 2 + 1\cdot(-1)\\ 1\cdot 2 + (-2)\cdot(-1)\end{bmatrix} && \text{(Definition 0.1)}\\ &= \begin{bmatrix}5\\ 4\end{bmatrix}, && \text{(multiply and add)}\end{aligned}$$

which matches $\underset{\sim}{b}$ entry by entry, so $\underset{\sim}{x}$ is a solution of the system.

Transposes of sums and products

Theorem 0.1 (Transpose of a sum)**.** *For $m \times n$ matrices* $\mathbf{A}$ *and* $\mathbf{B}$*:*

$$(\mathbf{A}+\mathbf{B})^{\top} = \mathbf{A}^{\top} + \mathbf{B}^{\top}$$

In particular, for column vectors $\underset{\sim}{x}$ *and* $\underset{\sim}{y}$ *of the same length:*

$$(\underset{\sim}{x}+\underset{\sim}{y})^{\top} = \underset{\sim}{x}^{\top} + \underset{\sim}{y}^{\top}$$

Proof

Proof. Entry (i, j) of each side:

$$\begin{aligned}
\left[(\mathbf{A}+\mathbf{B})^{\top}\right]_{ij} &= (\mathbf{A}+\mathbf{B})_{ji} && \text{(definition of the transpose)} \\
&= a_{ji} + b_{ji} && \text{(definition of matrix addition)} \\
&= (\mathbf{A}^{\top})_{ij} + (\mathbf{B}^{\top})_{ij} && \text{(definition of the transpose)} \\
&= \left[\mathbf{A}^{\top} + \mathbf{B}^{\top}\right]_{ij} && \text{(definition of matrix addition)}
\end{aligned}$$

□

Exercise 0.1 (Transposing a sum of two matrices)**.** Theorem 0.1 lets a transpose be taken before or after adding, whichever is less work. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & -3 \end{bmatrix}.$$

Find $(\mathbf{A}+\mathbf{B})^{\top}$.

Hint

Add first with Definition 8 in Matrices, then take the transpose.

Solution

Solution 0.1. Adding entry by entry,

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix} 1 & 3 & 1 \\ 5 & 1 & 1 \end{bmatrix},$$

so

$$(\mathbf{A}+\mathbf{B})^\top = \begin{bmatrix} 1 & 5 \\ 3 & 1 \\ 1 & 1 \end{bmatrix}.$$

By Theorem 0.1, adding $\mathbf{A}^\top = \begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 4 \end{bmatrix}$ and $\mathbf{B}^\top = \begin{bmatrix} 0 & 2 \\ 1 & 2 \\ 1 & -3 \end{bmatrix}$ gives the same matrix.

Theorem 0.2 (Transpose of a product)**.** *For an $m \times k$ matrix $\mathbf{A}$ and a $k \times n$ matrix $\mathbf{B}$:*

$$\underbrace{(\mathbf{AB})^\top}_{n\times m} = \underbrace{\mathbf{B}^\top}_{n\times k}\ \underbrace{\mathbf{A}^\top}_{k\times m}$$

Proof

Proof. Entry (i, j) of each side:

$$\begin{aligned}
\left[(\mathbf{AB})^\top\right]_{ij} &= (\mathbf{AB})_{ji} && \text{(definition of the transpose)} \\
&= \sum_{s=1}^{k} a_{js}\, b_{si} && \text{(definition of matrix multiplication)} \\
&= \sum_{s=1}^{k} (\mathbf{B}^\top)_{is}\, (\mathbf{A}^\top)_{sj} && \text{(definition of the transpose; reorder each product)} \\
&= \left[\mathbf{B}^\top \mathbf{A}^\top\right]_{ij} && \text{(definition of matrix multiplication)}
\end{aligned}$$

□

Remark (The order of the factors reverses)

Remark 0.1 (The order of the factors reverses). Transposing a product reverses the order of the factors. Keeping the original order usually gives an undefined product or a different matrix. For example, with $\mathbf{A} = [1,\ 2]$ (1×2) and $\mathbf{B} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ (2×1),

$$\begin{aligned}
\mathbf{AB} &= 1 \cdot 3 + 2 \cdot 4 \\
&= 11,
\end{aligned}$$

so $(\mathbf{AB})^\top = 11$, and

$$\mathbf{B}^\top \mathbf{A}^\top = [3,\ 4] \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 11, \qquad \mathbf{A}^\top \mathbf{B}^\top = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [3,\ 4] = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}.$$

If instead $\mathbf{A}$ is 2×3 and $\mathbf{B}$ is 3×1, then $\mathbf{B}^\top \mathbf{A}^\top$ is a $(1 \times 3)(3 \times 2)$ product, but $\mathbf{A}^\top \mathbf{B}^\top$ is a $(3 \times 2)(1 \times 3)$ product, which is undefined.

Exercise 0.2 (The order of the factors in a transposed square product)**.** Remark 0.1 shows that, for factors that are not square, keeping the original order of the factors in the transpose of a product can give an undefined product. For two square matrices of the same size, both $\mathbf{A}^\top\mathbf{B}^\top$ and $\mathbf{B}^\top\mathbf{A}^\top$ are defined, so the shapes alone cannot settle which order gives $(\mathbf{AB})^\top$. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Does $\mathbf{A}^\top\mathbf{B}^\top$ equal $(\mathbf{AB})^\top$ for these $\mathbf{A}$ and $\mathbf{B}$?

Hint

Compute $\mathbf{AB}$ with Definition 0.1 and take its transpose, then compute $\mathbf{A}^\top\mathbf{B}^\top$ and compare the two entry by entry.

Solution

Solution 0.2. No: the two matrices differ, so keeping the order of the factors gives the wrong transpose.
The product $\mathbf{AB}$ and its transpose are

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} 1\cdot 1 + 2\cdot 1 & 1\cdot 0 + 2\cdot 1 \\ 0\cdot 1 + 1\cdot 1 & 0\cdot 0 + 1\cdot 1 \end{bmatrix} && \text{(Definition 0.1)} \\ &= \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}, && \text{(multiply and add)} \end{aligned}$$

so $(\mathbf{AB})^\top = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$.

The transposes are $\mathbf{A}^\top = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ and $\mathbf{B}^\top = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, so

$$\begin{aligned} \mathbf{A}^\top\mathbf{B}^\top &= \begin{bmatrix} 1\cdot 1 + 0\cdot 0 & 1\cdot 1 + 0\cdot 1 \\ 2\cdot 1 + 1\cdot 0 & 2\cdot 1 + 1\cdot 1 \end{bmatrix} && \text{(Definition 0.1)} \\ &= \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}. && \text{(multiply and add)} \end{aligned}$$

Entry $(1,1)$ is 3 in $(\mathbf{AB})^\top$ and 1 in $\mathbf{A}^\top\mathbf{B}^\top$. In the reversed order, as Theorem 0.2 says, $\mathbf{B}^\top\mathbf{A}^\top = \begin{bmatrix} 1\cdot 1 + 1\cdot 2 & 1\cdot 0 + 1\cdot 1 \\ 0\cdot 1 + 1\cdot 2 & 0\cdot 0 + 1\cdot 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$, which matches $(\mathbf{AB})^\top$.

Theorem 0.3 (A 1×1 product equals its own transpose)**.** *If* $\underset{\sim}{u}$ *and* $\underset{\sim}{v}$ *are vectors of length* p *and* $\mathbf{M}$ *is a* $p \times p$ *matrix, then* $\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v}$ *is a number, and*

$$\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v} = \underset{\sim}{v}^\top \mathbf{M}^\top \underset{\sim}{u}.$$

In particular, $\underset{\sim}{u}^\top \underset{\sim}{v} = \underset{\sim}{v}^\top \underset{\sim}{u}$.

Proof

Proof. The product $\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v}$ is a $(1 \times p)(p \times p)(p \times 1)$ product, so it is a 1×1 matrix. The transpose of a 1×1 matrix is itself, so

$$\begin{aligned} \underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v} &= \left(\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v}\right)^\top && \text{(a } 1 \times 1 \text{ matrix equals its transpose)} \\ &= \underset{\sim}{v}^\top \mathbf{M}^\top \left(\underset{\sim}{u}^\top\right)^\top && \text{(transpose of a product, twice)} \\ &= \underset{\sim}{v}^\top \mathbf{M}^\top \underset{\sim}{u} && \text{(transposing twice changes nothing)} \end{aligned}$$

The transpose step is Theorem 0.2. Taking $\mathbf{M} = \mathbf{I}$ gives the last statement. □

Example 0.1 (Two ways to write the same number)**.** Take $\underset{\sim}{u} = (1, 2)$, $\underset{\sim}{v} = (3, -1)$ and $\mathbf{M} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. Then $\mathbf{M} \underset{\sim}{v} = (1, -3)$, so

$$\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v} = 1 \cdot 1 + 2 \cdot (-3) = -5.$$

Also $\mathbf{M}^\top = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$ and $\mathbf{M}^\top \underset{\sim}{u} = (1, 8)$, so

$$\underset{\sim}{v}^\top \mathbf{M}^\top \underset{\sim}{u} = 3 \cdot 1 + (-1) \cdot 8 = -5.$$

The two products are equal, as the theorem says. In least squares, this is why the two cross terms of $(\mathbf{X}\underset{\sim}{\beta} - \underset{\sim}{y})^\top (\mathbf{X}\underset{\sim}{\beta} - \underset{\sim}{y})$ are equal and can be combined.

Exercise 0.3 (Evaluating a product of a vector, a matrix and a vector)**.** Theorem 0.3 turns $\underset{\sim}{v}^\top \mathbf{M}^\top \underset{\sim}{u}$ back into $\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v}$, which can be the easier product of a vector, a matrix and a vector to compute. Let

$$\underset{\sim}{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \qquad \underset{\sim}{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \qquad \mathbf{M} = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}.$$

Find $\underset{\sim}{v}^\top \mathbf{M}^\top \underset{\sim}{u}$.

Hint

By Theorem 0.3, it is enough to compute $\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v}$.

Solution

Solution 0.3. By Theorem 0.3, $\underset{\sim}{v}^\top \mathbf{M}^\top \underset{\sim}{u} = \underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v}$. By Definition 0.1,

$$\mathbf{M}\underset{\sim}{v} = \begin{bmatrix} 0 \cdot 2 + 1 \cdot 3 \\ 2 \cdot 2 + 0 \cdot 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix},$$

so

$$\underset{\sim}{u}^\top \mathbf{M} \underset{\sim}{v} = 1 \cdot 3 + (-1) \cdot 4 = -1.$$

The answer is -1.

[1] Wikipedia contributors, "Matrix multiplication — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Matrix_multiplication

[2] Wikipedia contributors, "Square matrix — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Square_matrix

[3] Wikipedia contributors, "Main diagonal — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Main_diagonal

[4] Wikipedia contributors, "Identity matrix — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Identity_matrix

[5] Wikipedia contributors, "System of linear equations — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/System_of_linear_equations