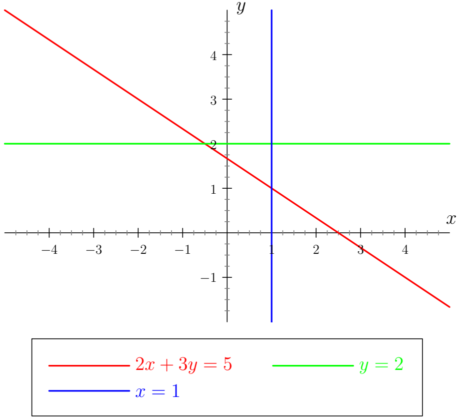
Geometric interpretation of linear equations: 2D

The set of solutions to

$$\alpha x + \beta y = \gamma$$

is a line in the plane. Conversely, any line in the plane is the solution set of some linear equation.



Geometric interpretation of linear equations: 3D

The set of solutions to

$$\alpha x + \beta y + \gamma z = \delta$$

is a plane in $\mathbf{R}^3$. And so forth in higher dimensions.

$$2x + 2y + 3z = 3$$

Systems of linear equations: 3D

System of linear equations with one solution.

$$\begin{array}{ccccccc} x & & & + & z & = & 1 \\ & & y & + & z & = & 1 \\ x & + & y & + & z & = & 3 \end{array}$$

Systems of linear equations: 3D

System of linear equations with infinitely many solutions.

$$\begin{array}{ccccccc} x & & & + & z & = & 1 \\ & & y & + & z & = & 1 \\ x & + & y & + & 2z & = & 2 \end{array}$$

Systems of linear equations: 3D

System of linear equations with no solution.

$$\begin{array}{rcrcrcr} x & & & + & z & = & 1 \\ & & y & + & z & = & 1 \\ x & + & y & + & 2z & = & 3 \end{array}$$

Computational complexity of Gaussian elimination

1. $r := 1$, $c := 1$
2. If $A_{i,j} = 0$ for all $i \geq r$ and $j \geq c$, then end.
3. Let $c := \min\{j \geq c : A_{i,j} \neq 0 \text{ for some } i \geq r\}$.
4. Choose arbitrary $i \geq r$ such that $A_{c,i} \neq 0$, and swap i-th and r-th row.
5. For every $i > r$, subtract $\frac{A_{i,c}}{A_{r,c}}$-times the r-th row from the i-th row.
6. Let $r := r + 1$, $c := c + 1$ and repeat from step 2.

Step 3: each column searched once: mn operations in total.

Step 5: repeated at most $\min(m, n)\times$, at most $2mn$ operations.

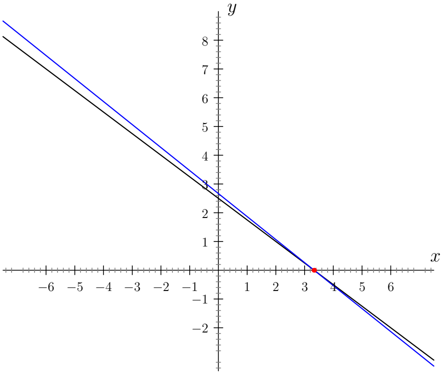
In total, $\approx mn \min(m, n)$ operations; $\approx n^3$ when $m = n$.

Real-world problems with linear equations

Changing the coefficients a little may change the solution a lot:

$$\begin{array}{rcrcl} [3, 3.1, 3.2, 3.3, 3.4]x & + & 4y & = & 10 \\ 4x & + & 5y & = & 13.3 \end{array} \rightarrow$$

$$(x, y) = [(3.3, 0), (6.7, -2.7), \text{no solution}, (-6.7, 8), (-3.3, 5.3)]$$

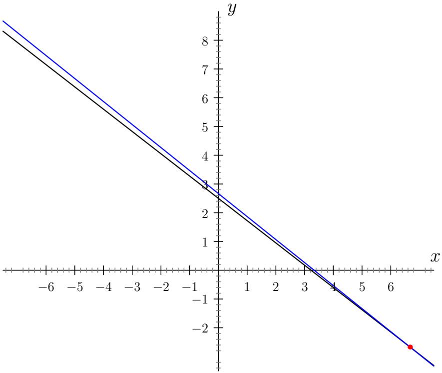


Real-world problems with linear equations

Changing the coefficients a little may change the solution a lot:

$$\begin{array}{lcccl} [3, \color{red}{3.1}, 3.2, 3.3, 3.4]x & + & 4y & = & 10 \\ \color{blue}{4x} & \color{blue}{+} & \color{blue}{5y} & \color{blue}{=} & \color{blue}{13.3} \end{array} \rightarrow$$

$$(x, y) = [(3.3, 0), \color{red}{(6.7, -2.7)}, \text{no solution}, (-6.7, 8), (-3.3, 5.3)]$$

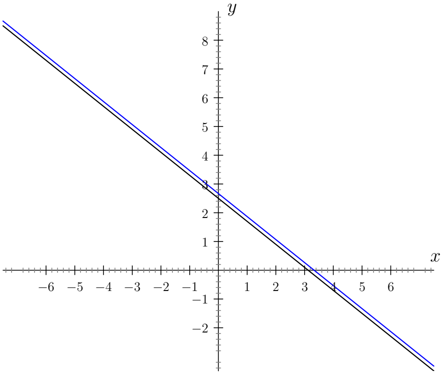


Real-world problems with linear equations

Changing the coefficients a little may change the solution a lot:

$$\begin{array}{lcccc} [3, 3.1, \color{red}{3.2}, 3.3, 3.4]x & + & 4y & = & 10 \\ \color{blue}{4x} & \color{blue}{+} & \color{blue}{5y} & \color{blue}{=} & \color{blue}{13.3} \end{array} \rightarrow$$

$$(x, y) = [(3.3, 0), (6.7, -2.7), \color{red}{\text{no solution}}, (-6.7, 8), (-3.3, 5.3)]$$

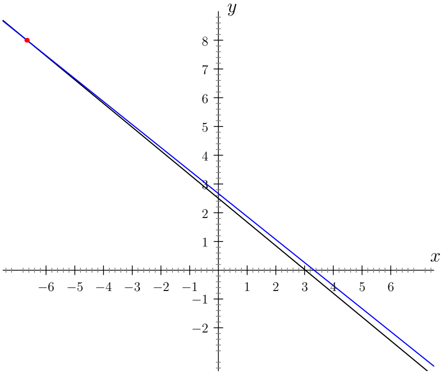


Real-world problems with linear equations

Changing the coefficients a little may change the solution a lot:

$$\begin{array}{lcccc} [3, 3.1, 3.2, \color{red}{3.3}, 3.4]x & + & 4y & = & 10 \\ \color{blue}{4x} & \color{blue}{+} & \color{blue}{5y} & \color{blue}{=} & \color{blue}{13.3} \end{array} \rightarrow$$

$$(x, y) = [(3.3, 0), (6.7, -2.7), \text{no solution}, \color{red}{(-6.7, 8)}, (-3.3, 5.3)]$$

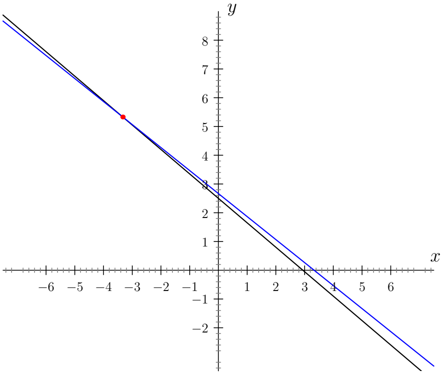


Real-world problems with linear equations

Changing the coefficients a little may change the solution a lot:

$$\begin{array}{lcccl} [3, 3.1, 3.2, 3.3, {\color{red}3.4}]x & + & 4y & = & 10 \\ {\color{blue}4x} & {\color{blue}+} & {\color{blue}5y} & {\color{blue}=} & {\color{blue}13.3} \end{array} \rightarrow$$

$$(x, y) = [(3.3, 0), (6.7, -2.7), \text{no solution}, (-6.7, 8), {\color{red}(-3.3, 5.3)}]$$



Consequences for Gaussian elimination

Doing the arithmetics with limited precision introduces rounding errors.

- This may change the solution a lot.
- Heuristic: in step 4, when choosing the row to swap, select the one with $|A_{c,i}|$ maximum.
    - In step 5, dividing by a large number decreases the absolute error.
    - Not guaranteed to help.

Doing the arithmetics with full precision (rational numbers):

- The denominators may become too large (doubly exponential).

Reminder: Gaussian elimination

1. $r := 1$, $c := 1$
2. If $A_{i,j} = 0$ for all $i \geq r$ and $j \geq c$, then end.
3. Let $c := \min\{j \geq c : A_{i,j} \neq 0 \text{ for some } i \geq r\}$.
4. Choose arbitrary $i \geq r$ such that $A_{c,i} \neq 0$, and swap i-th and r-th row.
5. For every $i > r$, subtract $\frac{A_{i,c}}{A_{r,c}}$-times the r-th row from the i-th row.
6. Let $r := r + 1$, $c := c + 1$ and repeat from step 2.

More elimination

Suppose that A is in REF, with basis column indices $p_1 < \ldots < p_r$. We can

- Make $A_{i,p_i} = 1$ (by dividing the i-th row by A_{i,p_i}).
- Make A_{i,p_i} the only non-zero entry in the p_i-th column (by subtracting a multiple of i-th row from the preceding ones).

Matrix satisfying these additional properties is in reduced row echelon form (RREF).

$$\begin{pmatrix} 1 & 0 & \star & 0 & \star & 0 \\ 0 & 1 & \star & 0 & \star & 0 \\ 0 & 0 & 0 & 1 & \star & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

RREF: example

$$\begin{pmatrix} 3 & 1 & 1 & 1 & 1 & 5 \\ 0 & 1 & 3 & 2 & 0 & -1 \\ 0 & 0 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1/3 & 1/3 & 1/3 & 1/3 & 5/3 \\ 0 & 1 & 3 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \sim$$

$$\begin{pmatrix} 1 & 1/3 & 1/3 & 1/3 & 1/3 & 0 \\ 0 & 1 & 3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1/3 & 1/3 & 0 & 1/3 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \sim$$

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

Gauss-Jordan elimination

- Perform Gaussian elimination on A, let $p_1 < \ldots < p_r$ be basis column indices.
- For $k = r, r-1, \ldots, 1$:
  - Divide k-th row by A_{k,p_k}.
  - For $i = 1, \ldots, k-1$, subtract $A_{i,p_k} \times$ the k-th row from the i-th row.

Theorem

Gauss-Jordan elimination applied to matrix B returns a row-equivalent matrix A in RREF.

Uniqueness of RREF

Theorem (for now without proof)

For every matrix B, there exists exactly one row-equivalent matrix A in RREF.

Definition

Let $\mathrm{RREF}(B)$ denote the matrix in RREF that is row-equivalent to B.

Corollary

If A and A' are any matrices in REF and $A \sim A'$, then A and A' have the same basis column indices.

Uniqueness of RREF: example

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

$$\begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

RREF and systems of equations

In RREF, we can directly read the solution.

$$\begin{aligned} x_1 + x_2 + x_3 + 2x_4 + x_5 &= 6 \\ x_1 + x_2 + 3x_3 + 4x_4 + 3x_5 &= 12 \\ 2x_1 + 2x_2 + 3x_3 + 5x_4 + 4x_5 &= 16 \end{aligned} \to \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 2 & 1 & 6 \\ 1 & 1 & 3 & 4 & 3 & 12 \\ 2 & 2 & 3 & 5 & 4 & 16 \end{array}\right) \sim$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 2 & 1 & 6 \\ 0 & 0 & 2 & 2 & 2 & 6 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right) \sim \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right) \to$$

$$\begin{aligned} x_1 + x_2 \quad + x_4 \quad &= 3 \\ x_3 + x_4 \quad &= 2 \\ x_5 &= 1 \end{aligned}$$

Basis column indices 1, 3, 5, rest will be parameters:

$$\begin{aligned} x_2 &= s & x_4 &= t \\ x_1 &= 3 - s - t & x_3 &= 2 - t \\ x_5 &= 1 \end{aligned}$$

Operations with matrices

Let $A = \begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,n} \\ & \dots & \\ \alpha_{m,1} & \dots & \alpha_{m,n} \end{pmatrix}$ and $B = \begin{pmatrix} \beta_{1,1} & \dots & \beta_{1,n} \\ & \dots & \\ \beta_{m,1} & \dots & \beta_{m,n} \end{pmatrix}$
be $m \times n$ matrices, and γ a real number.
Then

$$A + B = \begin{pmatrix} \alpha_{1,1} + \beta_{1,1} & \dots & \alpha_{1,n} + \beta_{1,n} \\ & \dots & \\ \alpha_{m,1} + \beta_{m,1} & \dots & \alpha_{m,n} + \beta_{m,n} \end{pmatrix}$$

$$A - B = \begin{pmatrix} \alpha_{1,1} - \beta_{1,1} & \dots & \alpha_{1,n} - \beta_{1,n} \\ & \dots & \\ \alpha_{m,1} - \beta_{m,1} & \dots & \alpha_{m,n} - \beta_{m,n} \end{pmatrix}$$

$$\gamma A = A\gamma = \begin{pmatrix} \gamma\alpha_{1,1} & \dots & \gamma\alpha_{1,n} \\ & \dots & \\ \gamma\alpha_{m,1} & \dots & \gamma\alpha_{m,n} \end{pmatrix}$$

Operations with matrices: example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + 3 \begin{pmatrix} 1 & 1 & 1 \\ -3 & 2 & -1 \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 3 & 3 & 3 \\ -9 & 6 & -3 \end{pmatrix} =$$

$$\begin{pmatrix} 4 & 5 & 6 \\ -5 & 11 & 3 \end{pmatrix}$$

Properties of matrix operations

Let A, B, C be $m \times n$ matrices, α and β real numbers. Let $O = \begin{pmatrix} 0 & \dots & 0 \\ & \dots & \\ 0 & \dots & 0 \end{pmatrix}$ denote the $m \times n$ matrix with all 0 entries.

- $A + B = B + A$ commutativity
- $A + (B + C) = (A + B) + C$ associativity
- $A + O = O + A = A$ neutral element
- $A - A = O$
- $0A = O$
- $1A = A$
- $A - B = A + (-1)B$
- $(\alpha + \beta)A = \alpha A + \beta A$, $\alpha(A + B) = \alpha A + \alpha B$ distributivity
- $(\alpha\beta)A = \alpha(\beta A)$

Matrix transposition

If A is an $m \times n$ matrix, then let A^T denote the $n \times m$ matrix obtained by exchanging rows and columns,

$$(A^T)_{r,c} = A_{c,r} \text{ for all } r = 1, \ldots, n \text{ and } c = 1, \ldots, m.$$

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$$

A square matrix A is symmetric if $A = A^T$.
Example:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \text{ is symmetric, } \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \text{ is not symmetric.}$$

Properties of matrix transposition

Let A and B be $m \times n$ matrices and α a real number.

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(\alpha A)^T = \alpha A^T$

Matrix multiplication: motivation

We want

$$\begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,n} \\ \alpha_{2,1} & \dots & \alpha_{2,n} \\ & \dots & \\ \alpha_{m,1} & \dots & \alpha_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}x_1 + \alpha_{1,2}x_2 + \dots + \alpha_{1,n}x_n \\ \alpha_{2,1}x_1 + \alpha_{2,2}x_2 + \dots + \alpha_{2,n}x_n \\ \dots \\ \alpha_{m,1}x_1 + \alpha_{m,2}x_2 + \dots + \alpha_{m,n}x_n \end{pmatrix}$$

so that

$$\begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,n} \\ & \dots & \\ \alpha_{m,1} & \dots & \alpha_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \dots \\ \beta_m \end{pmatrix}$$

is the same as

$$\begin{aligned} \alpha_{1,1}x_1 + \alpha_{1,2}x_2 + \dots + \alpha_{1,n}x_n &= \beta_1 \\ &\dots \\ \alpha_{m,1}x_1 + \alpha_{m,2}x_2 + \dots + \alpha_{m,n}x_n &= \beta_m \end{aligned}$$

Matrix multiplication

Definition

If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then AB or $A \cdot B$ is the $m \times p$ matrix such that

$$(AB)_{r,c} = A_{r,1}B_{1,c} + A_{r,2}B_{2,c} + \ldots + A_{r,n}B_{n,c}.$$

$$\begin{pmatrix} & & \ldots & \\ \alpha_{r,1} & \alpha_{r,2} & \cdots & \alpha_{r,n} \\ & & \ldots & \end{pmatrix} \cdot \begin{pmatrix} \ldots & \beta_{1,c} & \ldots \\ \ldots & \vdots & \ldots \\ \ldots & \beta_{n,c} & \ldots \end{pmatrix} =$$

$$\begin{array}{cc} & \\ r\text{-th row} & \begin{pmatrix} \ldots & \ldots & \ldots \\ \ldots & \alpha_{r,1}\beta_{1,c} + \alpha_{r,2}\beta_{2,c} + \ldots + \alpha_{r,n}\beta_{n,c} & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix} \\ & c\text{-th column} \end{array}$$

Matrix multiplication: example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix} =$$

$$\begin{pmatrix} 1\cdot 7 + 2\cdot 9 + 3\cdot 11 & 1\cdot 8 + 2\cdot 10 + 3\cdot 12 \\ 4\cdot 7 + 5\cdot 9 + 6\cdot 11 & 4\cdot 8 + 5\cdot 10 + 6\cdot 12 \end{pmatrix} = \begin{pmatrix} 58 & 64 \\ 139 & 154 \end{pmatrix}$$

$$\begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} =$$

$$\begin{pmatrix} 7\cdot 1 + 8\cdot 4 & 7\cdot 2 + 8\cdot 5 & 7\cdot 3 + 8\cdot 6 \\ 9\cdot 1 + 10\cdot 4 & 9\cdot 2 + 10\cdot 5 & 9\cdot 3 + 10\cdot 6 \\ 11\cdot 1 + 12\cdot 4 & 11\cdot 2 + 12\cdot 5 & 11\cdot 3 + 12\cdot 6 \end{pmatrix} = \begin{pmatrix} 39 & 54 & 69 \\ 49 & 68 & 87 \\ 59 & 82 & 105 \end{pmatrix}$$

Matrix multiplication: computational complexity

Multiplying two $n \times n$ matrices:

- trivial algorithm: $\approx n^3$ operations
- Strassen algorithm: $\approx n^{2.81}$ operations
- the best known algorithm (for extremely large matrices): $\approx n^{2.373}$ operations
- do $\approx n^2$ operations suffice?

Matrix multiplication: non-commutativity

Let A be an $m \times n$ matrix, and B an $n \times p$ matrix.

- If $m \neq p$, then BA is undefined.
- If $m = p \neq n$, then AB and BA have different sizes ($m \times m$ and $n \times n$).
- If $m = p = n$, then AB and BA may be different.

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

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

Matrix multiplication: associativity

Let A be an $n \times m$ matrix, B an $m \times p$ matrix, and C and $p \times q$ matrix. Then

$$(AB)C = A(BC).$$

Proof.

AB is an $n \times p$ matrix, BC is an $m \times q$ matrix, so the products $(AB)C$ and $A(BC)$ are defined and are $n \times q$ matrices.

$$[(AB)C]_{r,c} = \sum_{i=1}^{p} (AB)_{r,i} C_{i,c} = \sum_{i=1}^{p} \left(\sum_{j=1}^{m} A_{r,j} B_{j,i} \right) C_{i,c} =$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{m} A_{r,j} B_{j,i} C_{i,c} = \sum_{j=1}^{m} \sum_{i=1}^{p} A_{r,j} B_{j,i} C_{i,c} =$$

$$= \sum_{j=1}^{m} A_{r,j} \sum_{i=1}^{p} B_{j,i} C_{i,c} = \sum_{j=1}^{m} A_{r,j} (BC)_{j,c} = [A(BC)]_{r,c}$$

Matrix multiplication: properties

Let A and A' be $n \times m$ matrices, let B and B' be $m \times p$ matrices, let C be an $p \times q$ matrix, let D be an $m \times s$ matrix, let α and β be real numbers.

- $(AB)C = A(BC)$
- $A(B + B') = AB + AB'$
- $(A + A')B = AB + A'B$
- $(\alpha A)B = A(\alpha B) = \alpha(AB)$
- $A(B|D) = (AB|AD)$
- $(AB)^T = B^T A^T$

Identity matrix, standard basis vectors

Let $I^{(n)} = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ & & \dots & \end{pmatrix}$ denote the $n \times n$ matrix with 1's on diagonal and 0's elsewhere, $I^{(n)}_{r,c} = \begin{cases} 1 & \text{if } r = c \\ 0 & \text{if } r \neq c \end{cases}$

Let $e^{(n)}_k$ denote the k-th column of $I^{(n)}$,

$$e^{(n)}_k = \begin{pmatrix} 0 \\ 0 \\ \dots \\ \color{red}{1} \\ 0 \\ \dots \end{pmatrix} \begin{array}{l} \\ \\ \\ \color{red}{k\text{-th row}} \\ \\ \\ \end{array}$$

When n is clear from the context, we write only I and e_k.

Identity matrix, standard basis and multiplication

Let A be an $m \times n$ matrix.

- $A \cdot e_k^{(n)} = A_{\star,k}$ is the k-th column of A.
- $(e_k^{(m)})^T \cdot A = A_{k,\star}$ is the k-th row of A.
- $AI^{(n)} = I^{(m)}A = A$.