Chapter 2 Matrix Algebra

Math 39, San Jose State University

Prof. Guangliang Chen

Fall 2022

Sections 2.1-2.3 Matrix operations

- Matrix addition/subtraction
- Matrix multiplication
- Matrix powers
- Matrix transpose
- Matrix inverse
- The Invertible Matrix Theorem

Section 2.4 Partitioned matrices

Section 2.5 LU decomposition

Introduction

Matrices are **two dimensional arrays** of real numbers that are arranged along rows (first dimension) and columns (second dimension):

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [\mathbf{a}_1\, \mathbf{a}_2\, \ldots\, \mathbf{a}_n].$$

We denote matrices that have m rows and n columns by $\mathbf{A} \in \mathbb{R}^{m\times n}$, and say that the **size** of the matrix is $m \times n$.

Vectors can be regarded as matrices with size $n \times 1$ (column) or $1 \times n$ (row).

Sometimes, we also use notation like $\mathbf{A} = (a_{ij})_{1\le i\le m, 1\le j\le n}$, or even $\mathbf{A} = (a_{ij})$.

Special matrices

We say that $\mathbf{A}$ is a **square** matrix if $m = n$ (i.e., equally many rows and columns).

Diagonal matrices are square matrices whose only nonzero entries are in the main diagonal of the matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \qquad \longleftarrow \text{empty spaces indicate zero}$$

An **identity matrix** is a diagonal matrix with constant value 1 along the diagonal:

$$\mathbf{I}_n = \mathrm{diag}(1, \ldots, 1) \in \mathbb{R}^{n \times n}.$$

Lastly, a **zero matrix** is a matrix with all entries being 0, and denoted as $\mathbf{O}$.

Matrix operations

- Scalar multiple of a matrix
- Matrix-vector product
- Adding two matrices of the same size (also letting them subtract)
- Multiplying two matrices of "matching" sizes
- Transpose of a matrix
- Inverse of a square matrix

Def 0.1 (Scalar multiple). Let r be a real number and $\mathbf{A} \in \mathbb{R}^{m\times n}$. Then $\mathbf{B} = r\mathbf{A}$ is defined as a matrix of the same size with entries $b_{ij} = ra_{ij}$.

In matrix form, this is

$$r\mathbf{A} = \begin{bmatrix} ra_{11} & ra_{12} & \cdots & ra_{1n} \\ ra_{21} & ra_{22} & \cdots & ra_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ra_{m1} & ra_{m2} & \cdots & ra_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

Def 0.2 (Matrix sum/difference). Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$. Then the matrix sum $\mathbf{C} = \mathbf{A} + \mathbf{B}$ is defined as a matrix of the same size with the following entries

$$\mathbf{C} = (c_{ij}), \qquad c_{ij} = a_{ij} + b_{ij}$$

In matrix form, the above definition becomes

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

Remark. The difference of two matrices, $\mathbf{A} - \mathbf{B}$, is defined similarly (with every $+$ sign being changed to - sign).

Example 0.1. Let

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Find $\mathbf{A}+\mathbf{B}, \mathbf{A}-\mathbf{B}, 3\mathbf{B}$ and $\mathbf{A}+3\mathbf{B}$.

The scalar multiple of a matrix and matrix sum satisfy the following commutative, associative and distributive laws.

Theorem 0.1. Let $\mathbf{A}, \mathbf{B}, \mathbf{C}$ be three matrices of the same size and r, s be scalars. Then

- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- $\mathbf{A} + \mathbf{O} = \mathbf{O} + \mathbf{A} = \mathbf{A}$ ($\mathbf{O}$ is the zero matrix of same size)
- $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
- $r(s\mathbf{A}) = (rs)\mathbf{A}$
- $r(\mathbf{A} + \mathbf{B}) = r\mathbf{A} + r\mathbf{B}$
- $(r + s)\mathbf{A} = r\mathbf{A} + s\mathbf{A}$

Matrix-vector product

Def 0.3. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$. Their product is defined as a vector $\mathbf{y} \in \mathbb{R}^m$ of the following form

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

In compact notation,

$$\mathbf{y} = (y_i) \in \mathbb{R}^m, \qquad \text{with} \quad y_i = \sum_{j=1}^{n} a_{ij}x_j, \; 1 \le i \le m$$

Alternatively (as we have already seen previously), we can multiply a matrix and a vector in a columnwise fashion.

Theorem 0.2. Let $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n] \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$. Then

$$\mathbf{A}\mathbf{x} = [\mathbf{a}_1 \ldots \mathbf{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \cdot \mathbf{a}_1 + \cdots + x_n \cdot \mathbf{a}_n.$$

Proof. By definition,

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

Two properties about matrix-vector multiplication

Theorem 0.3. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $r \in \mathbb{R}$. Then

- $\mathbf{A}(\mathbf{x}+\mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}$
- $\mathbf{A}(r\mathbf{x}) = r(\mathbf{A}\mathbf{x})$

Remark. They were needed for showing that transformations of the form $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ must be linear.

Proof. By the columnwise way of multiplying a matrix and a vector,

$$\begin{aligned}\mathbf{A}(\mathbf{x}+\mathbf{y}) &= [\mathbf{a}_1 \ldots \mathbf{a}_n] \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} \\ &= (x_1+y_1)\mathbf{a}_1 + \cdots + (x_n+y_n)\mathbf{a}_n \\ &= (x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n) + (y_1\mathbf{a}_1 + \cdots + y_n\mathbf{a}_n) \\ &= \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}.\end{aligned}$$

Similarly,

$$\mathbf{A}(r\mathbf{x}) = [\mathbf{a}_1 \ldots \mathbf{a}_n] \begin{bmatrix} rx_1 \\ \vdots \\ rx_n \end{bmatrix} = (rx_1)\mathbf{a}_1 + \cdots + (rx_n)\mathbf{a}_n = r\underbrace{(x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n)}_{\mathbf{A}\mathbf{x}}.$$

A third property about matrix-vector multiplication

Theorem 0.4. Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ and $\mathbf{x} \in \mathbb{R}^n$. Then

$$(\mathbf{A}+\mathbf{B})\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}.$$

Proof. Let $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$ and $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$. Then

$$\mathbf{A}+\mathbf{B} = [\mathbf{a}_1+\mathbf{b}_1, \ldots, \mathbf{a}_n+\mathbf{b}_n].$$

It follows that

$$\begin{aligned}(\mathbf{A}+\mathbf{B})\mathbf{x} &= x_1(\mathbf{a}_1+\mathbf{b}_1) + \cdots + x_n(\mathbf{a}_n+\mathbf{b}_n) \\ &= (x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n) + (x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) \\ &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}.\end{aligned}$$

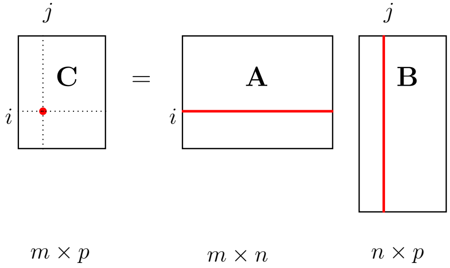
Matrix-matrix multiplications

Def 0.4. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Their product is defined as a matrix $\mathbf{C} \in \mathbb{R}^{m\times p}$ with entries

$$c_{ij} = [a_{i1} \ \dots \ a_{in}] \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix}$$

$$= a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$$

$$= \sum_{k=1}^{n} a_{ik}b_{kj}.$$



Remark. The matrix-vector product is just the special case of $p = 1$.

Example 0.2. Let

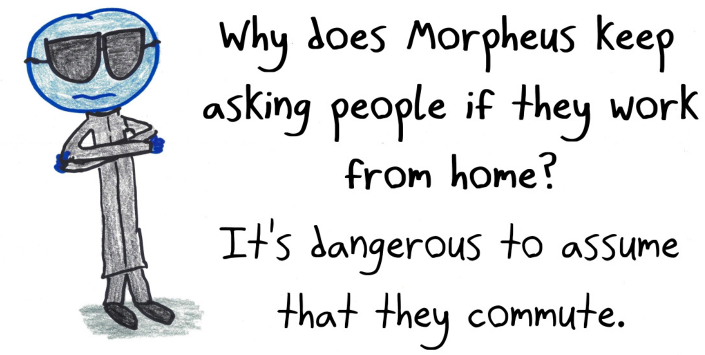
$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

Find $\mathbf{AB}$ and $\mathbf{BA}$. Are they the same?

Example 0.3. Let

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

Find $\mathbf{AB}$. Is $\mathbf{BA}$ defined?



(Taken from https://mathwithbaddrawings.com/2018/03/07/matrix-jokes/)

WARNINGS

- There is no commutative law between matrices: $\mathbf{AB} \neq \mathbf{BA}$. In fact, not both of them need to be defined at the same time.
- If $\mathbf{AB} = \mathbf{O}$, then we cannot conclude that $\mathbf{A} = \mathbf{O}$ or $\mathbf{B} = \mathbf{O}$.
- There is no cancellation law, i.e., $\mathbf{AB} = \mathbf{AC}$ does not necessarily imply $\mathbf{B} = \mathbf{C}$.

Can you give an example for the last statement?

A small, useful result on matrix-matrix-vector product

Theorem 0.5. Let $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{n\times p}$ and $\mathbf{x} \in \mathbb{R}^p$. Then

$$(\mathbf{AB})\mathbf{x} = \mathbf{A}(\mathbf{Bx}).$$

Proof. We compare the entries of both sides. For any $1 \le i \le m$,

$$\begin{aligned} ((\mathbf{AB})\mathbf{x})_i &= \sum_j (\mathbf{AB})_{ij} x_j = \sum_j \sum_k a_{ik} b_{kj} x_j \\ &= \sum_k a_{ik} \sum_j b_{kj} x_j = \sum_k a_{ik} (\mathbf{Bx})_k = (\mathbf{A}(\mathbf{Bx}))_i \,. \end{aligned}$$

Remark. The right hand side is much more efficient to compute, especially when having large matrices $\mathbf{A}, \mathbf{B}$.

Matrix computing in Matlab (optional)

See the following lecture:

`https://www.sjsu.edu/faculty/guangliang.chen/Math250/lec2matrixcomp.pdf`

Matlab scripts available on the Math 250 course page:

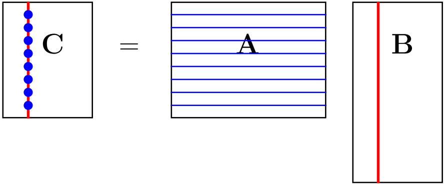
`https://www.sjsu.edu/faculty/guangliang.chen/Math250.html`

The columnwise matrix multiplication (very important)

Theorem 0.6. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Then

$$\mathbf{C} = \mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \dots \mathbf{b}_p] = [\mathbf{Ab}_1 \dots \mathbf{Ab}_p]$$

This shows that for each $j = 1, \dots, p$, the jth column of $\mathbf{AB}$ is equal to $\mathbf{A}$ times the jth column of $\mathbf{B}$.



Properties of matrix multiplication

Theorem 0.7. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Then

- $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (for $\mathbf{B} \in \mathbb{R}^{n\times p}, \mathbf{C} \in \mathbb{R}^{p\times q}$)
- $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ (for $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n\times p}$)
- $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$ (for $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{\ell\times m}$)
- $r(\mathbf{AB}) = (r\mathbf{A})\mathbf{B} = \mathbf{A}(r\mathbf{B})$ (for $\mathbf{B} \in \mathbb{R}^{n\times p}$)
- $\mathbf{I}_m\mathbf{A} = \mathbf{AI}_n = \mathbf{A}$.

Proof. Enough to compare columns. □

Example 0.4. Compute the following product

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Matrix powers

Def 0.5. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a square matrix and k a positive integer. Then the kth **power** of $\mathbf{A}$ is defined as

$$\mathbf{A}^k = \underbrace{\mathbf{A}\cdot\mathbf{A}\cdots\mathbf{A}}_{k\text{ copies}}.$$

Example 0.5. Let

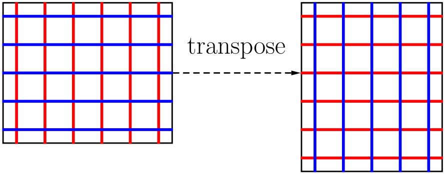
$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Find $\mathbf{A}^3$ and $\mathbf{B}^3$. What are $\mathbf{A}^k$ and $\mathbf{B}^k$ for $k > 3$?

Transpose of a matrix

Def 0.6. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be any matrix. Its transpose, denoted as $\mathbf{A}^T$ is defined to the $n \times m$ matrix $\mathbf{B}$ with entries $b_{ij} = a_{ji}$.

Remark. During the transpose operation, rows (of $\mathbf{A}$) become columns (of $\mathbf{B}$), and columns become rows.



Example 0.6. Find the transpose of the following matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

Properties of the matrix transpose

Theorem 0.8. Let $\mathbf{A}, \mathbf{B}$ be matrices with appropriate sizes for each statement.

- $(\mathbf{A}^T)^T = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- For any scalar r, $(r\mathbf{A})^T = r\mathbf{A}^T$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ (not the other product $\mathbf{A}^T\mathbf{B}^T$, which may not even be defined)

Proof. The first three are obvious. To prove the last one, check the ij-entry of each side. We show the work in class. □

Matrix inverse

Just like nonzero real numbers ($a \in \mathbb{R}$) have their reciprocals ($\frac{1}{a}$), certain (not all) square matrices have matrix inverses.

Def 0.7. A square matrix $\mathbf{A} \in \mathbb{R}^n$ is said to be invertible if there exists another matrix of the same size $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n.$$

In this case, $\mathbf{B}$ is called the inverse of $\mathbf{A}$ and we write $\mathbf{B} = \mathbf{A}^{-1}$ ($\mathbf{A}$ is also called the inverse of $\mathbf{B}$).

Example 0.7. Verify that $\mathbf{A} = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$ are inverses of each other and then use this fact to solve the matrix equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ for $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

From the previous example, we can formulate the following theorem.

Theorem 0.9. Consider a matrix equation $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a square matrix. If $\mathbf{A}$ is invertible, then for any vector $\mathbf{b} \in \mathbb{R}^n$, the system has a unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Proof. Since $\mathbf{A}$ is invertible, its inverse $\mathbf{A}^{-1}$ exists and we can use it to multiply both sides of the equation

$$\mathbf{A}^{-1}(\mathbf{Ax}) = \mathbf{A}^{-1}\mathbf{b}$$

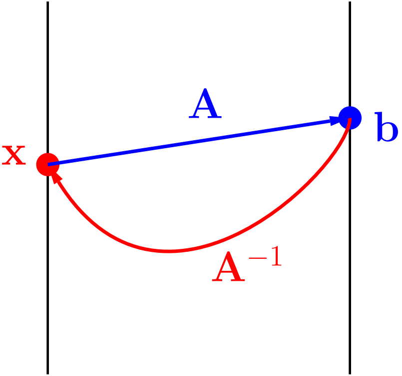
By the associative law,

$$\underbrace{(\mathbf{A}^{-1}\mathbf{A})}_{\mathbf{I}}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

which yields that

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

Illustration of A^{-1} as a transformation



Properties of matrix inverse

Theorem 0.10. Let $\mathbf{A}, \mathbf{B}$ be two invertible matrices of the same size. Then

- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$
- For any nonzero scalar r, $(r\mathbf{A})^{-1} = \frac{1}{r}\mathbf{A}^{-1}$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (not the other product $\mathbf{A}^{-1}\mathbf{B}^{-1}$)

Proof. We verify them in class. □

The Invertible Matrix Theorem (part 1)

"For a square matrix, lots of things are the same."

Theorem 0.11. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a square matrix. Then the following statements are all equivalent:

(1) $\mathbf{A}$ is invertible.

(2) There is an $n \times n$ matrix $\mathbf{C}$ such that $\mathbf{CA} = \mathbf{I}$.

(3) The equation $\mathbf{Ax} = \mathbf{0}$ only has the trivial solution.

(4) $\mathbf{A}$ has n pivot positions.

(5) $\mathbf{A}$ is row equivalent to $\mathbf{I}_n$.

The Invertible Matrix Theorem (part 2)

Theorem 0.12. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a square matrix. Then the following statements are all equivalent:

(1) $\mathbf{A}$ is invertible.

(6) There is an $n \times n$ matrix $\mathbf{D}$ such that $\mathbf{AD} = \mathbf{I}$.

(7) The equation $\mathbf{Ax} = \mathbf{b}$ (for any $\mathbf{b}$) is always consistent.

(8) The columns of $\mathbf{A}$ span $\mathbb{R}^n$.

(9) The linear transformation $f(\mathbf{x}) = \mathbf{Ax}$ (from $\mathbb{R}^n$ to $\mathbb{R}^n$) is onto.

The Invertible Matrix Theorem (part 3)

Theorem 0.13. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a square matrix. Then the following statements are all equivalent:

(1) $\mathbf{A}$ is invertible.

(10) $\mathbf{A}^T$ is invertible.

(3) The equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ only has the trivial solution.

(11) The columns of $\mathbf{A}$ form a linearly independent set.

(12) The linear transformation $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ is one-to-one.

Summary

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a square matrix.

If $\mathbf{A}$ is invertible, then all of the following statements are true.

Conversely, if any of the following statement is true, then $\mathbf{A}$ must be invertible.

(2) There is an $n \times n$ matrix $\mathbf{C}$ such that $\mathbf{CA} = \mathbf{I}$.

(6) There is an $n \times n$ matrix $\mathbf{D}$ such that $\mathbf{AD} = \mathbf{I}$.

(3) The equation $\mathbf{Ax} = \mathbf{0}$ only has the trivial solution.

(7) The equation $\mathbf{Ax} = \mathbf{b}$ (for any $\mathbf{b}$) has at least one solution.

(8) The columns of $\mathbf{A}$ span $\mathbb{R}^n$.

(11) The columns of $\mathbf{A}$ form a linearly independent set.

(9) The linear transformation $f(\mathbf{x}) = \mathbf{Ax}$ (from $\mathbb{R}^n$ to $\mathbb{R}^n$) is onto.

(12) The linear transformation $f(\mathbf{x}) = \mathbf{Ax}$ is one-to-one.

Finding matrix inverse

First consider 2×2 matrices

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

If $ad - bc \neq 0$, then $\mathbf{A}$ is invertible and its inverse is given by the following empirical rule

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Example 0.8. Use the above rule to find the inverse of

$$\mathbf{A} = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$$

In general, given an invertible matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ (for any n), finding its inverse is equivalent to solving the matrix equation

$$\mathbf{A}\mathbf{X} = \mathbf{I}_n, \qquad \text{or equivalently} \quad \mathbf{A}[\mathbf{x}_1, \ldots, \mathbf{x}_n] = [\mathbf{e}_1, \ldots, \mathbf{e}_n]$$

This leads to n separate systems of linear equations:

$$\mathbf{A}\mathbf{x}_1 = \mathbf{e}_1 \text{ (i.e. } [\mathbf{A} \mid \mathbf{e}_1]), \quad \ldots, \quad \mathbf{A}\mathbf{x}_n = \mathbf{e}_n \text{ (i.e. } [\mathbf{A} \mid \mathbf{e}_n]).$$

which may be solved simultaneously:

$$[\mathbf{A} \mid [\mathbf{e}_1, \ldots, \mathbf{e}_n]] = [\mathbf{A} \mid \mathbf{I}_n] \longrightarrow [\mathbf{I}_n \mid \mathbf{A}^{-1}].$$

Example 0.9. Find the inverse of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix},$$

if its exists.

Partitioned matrices

A **partitioned matrix**, also called a **block matrix**, is a matrix whose elements have been divided into blocks (called **submatrices**).

For example,

$$\mathbf{A} = \left[\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 & 0 \\ \hline 0 & 0 & 0 & 7 & 8 \\ \hline 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix}$$

Partitioned matrices are very useful because they reduce large matrices into a collection of smaller matrices (which are easier to deal with).

Addition and scalar multiplication

If two matrices $\mathbf{A}, \mathbf{B}$ have the same size and have been partitioned in exactly the same way, then we can just add the corresponding blocks to get their sum (with the same partition):

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix} = \begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \\ A_{31}+B_{31} & A_{32}+B_{32} \end{bmatrix}$$

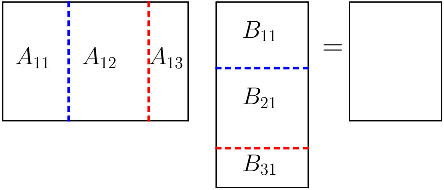
The scalar multiple of a partitioned matrix is

$$r\mathbf{A} = \begin{bmatrix} rA_{11} & rA_{12} \\ rA_{21} & rA_{22} \\ rA_{31} & rA_{32} \end{bmatrix}$$

Multiplication of partitioned matrices: simple cases

Let $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{n\times p}$ be two matrices that may be multiplied together.

When the columns of $\mathbf{A}$ and rows of $\mathbf{B}$ are divided in a conformable way, we can carry out block multiplication:



$$\mathbf{AB} = A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31}$$

Remark.

- All terms $\mathbf{AB}, A_{11}B_{11}, A_{12}B_{21}, A_{13}B_{31}$ are $m \times p$ matrices.
- Such partitions do not show up in the product matrix.

Example 0.10. Let

$$\mathbf{A} = \left[\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 & 0 \\ 7 & 8 & 9 & 0 & 0 \end{array}\right], \quad \mathbf{B} = \left[\begin{array}{cc} 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ \hline 1 & -1 \\ 1 & -1 \end{array}\right]$$

Find $\mathbf{AB}$ using two ways: (a) direct multiplication (b) block multiplication.

Answer.

$$\mathbf{AB} = \underbrace{\begin{bmatrix} 6 & -6 \\ 15 & -15 \\ 24 & -24 \end{bmatrix}}_{3\times 2} = \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix}}_{3\times 2} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}}_{3\times 2}$$

A joke

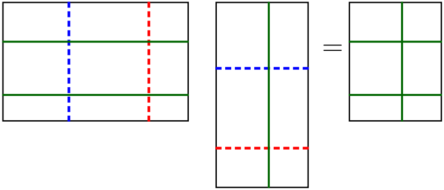
How does a mathematician change three light bulbs at the same time?

He gives them to three engineers and ask them to do it in parallel.

Multiplication of partitioned matrices: more general cases

Let $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{n\times p}$ be two matrices that are partitioned in a conformable way (i.e., column partition of $\mathbf{A}$ matches row partition of $\mathbf{B}$).

Regardless of the row partition of $\mathbf{A}$ and column partition of $\mathbf{B}$, we can carry out block multiplications by treating the blocks as numbers.



Remark. Row partition of $\mathbf{A}$ + column partition of $\mathbf{B}$ = partition of $\mathbf{AB}$ (such two partitions do not need to match).

In terms of math symbols, that is

$$\mathbf{AB} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \cdot \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \\ A_{31}B_{11} + A_{32}B_{21} + A_{33}B_{31} & A_{31}B_{12} + A_{32}B_{22} + A_{33}B_{32} \end{bmatrix}$$

In the above, we can think of $\mathbf{A}$ as a 3×3 partitioned matrix and $\mathbf{B}$ as a 3×2 partitioned matrix, so that we must obtain a 3×2 partitioned matrix.

Example 0.11. Verify that

$$\left[\begin{array}{ccc|cc} 1 & 2 & 3 & 0 & 0 \\ \hline 4 & 5 & 6 & 0 & 0 \\ 7 & 8 & 9 & 0 & 0 \end{array}\right] \cdot \left[\begin{array}{cc} 1 & -1 \\ 1 & -1 \\ 1 & -1 \\ \hline 1 & -1 \\ 1 & -1 \end{array}\right] = \left[\begin{array}{cc} 6 & -6 \\ \hline 15 & -15 \\ 24 & -24 \end{array}\right]$$

Example 0.12. Show that

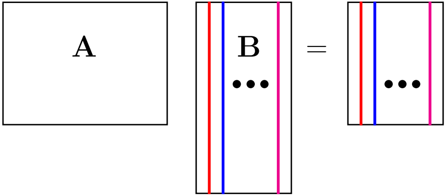
$$\begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = U_1 \, \Sigma \, V_1$$

(assuming all submatrices are compatible with each other)

Matrix multiplication again

The columnwise multiplication of two compatible matrices $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{n\times p}$ actually has already used simple partitions of matrices:

$$\mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \dots \mathbf{b}_p] = [\mathbf{Ab}_1 \dots \mathbf{Ab}_p]$$

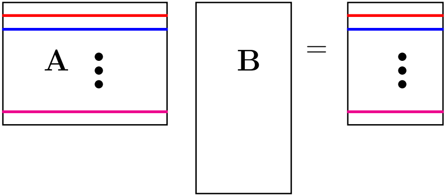


We present two new ways of performing matrix multiplication:

- Rowwise multiplication

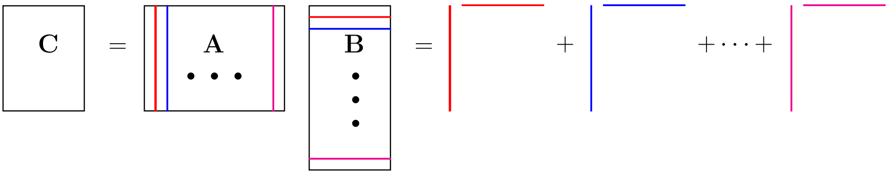
$$\mathbf{AB} = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \mathbf{B} = \begin{bmatrix} A_1\mathbf{B} \\ \vdots \\ A_m\mathbf{B} \end{bmatrix}$$

where $A_1, \ldots, A_m$ are the rows of $\mathbf{A}$.



- Column-row expansion

$$\mathbf{AB} = [\mathbf{a}_1 \ldots \mathbf{a}_n] \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix} = \mathbf{a}_1 B_1 + \cdots + \mathbf{a}_n B_n$$



Example 0.13. Find the product of $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}$ by using three different ways:

(a) Columnwise multiplication

(b) Rowwise multiplication and

(c) Column-row multiplication

Block diagonal matrices

Def 0.8. A matrix is said to be **block diagonal** if it is of the form

$$\mathbf{A} = \begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix}$$

Example 0.14.

$$\left[\begin{array}{ccc|cc} 1 & 2 & 3 & & \\ 4 & 5 & 6 & & \\ 7 & 8 & 9 & & \\ \hline & & & 1 & 1 \\ & & & 2 & 2 \end{array}\right]$$

Theorem 0.14. Let $\mathbf{A}, \mathbf{B}$ be two block diagonal matrices with conformable partitions:

$$\mathbf{A} = \begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} B_{11} & \\ & B_{22} \end{bmatrix}$$

Then we have

$$\mathbf{AB} = \begin{bmatrix} A_{11}B_{11} & \\ & A_{22}B_{22} \end{bmatrix}.$$

Proof. By direct verification. □

Remark. This formula also generalizes to three or more blocks.

The previous result immediately implies the following.

Theorem 0.15. For a block diagonal matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & \\ & A_{22} \end{bmatrix},$$

if the two blocks are both square and invertible, then $\mathbf{A}$ is also invertible. Moreover,

$$\mathbf{A}^{-1} = \begin{bmatrix} A_{11}^{-1} & \\ & A_{22}^{-1} \end{bmatrix}$$

Proof. By direct verification. □

Example 0.15. Find the inverse of

$$\left[\begin{array}{cc|c} 1 & 2 & \\ 1 & 3 & \\ \hline & & 4 \end{array}\right]$$

Block upper triangular matrices

Def 0.9. A matrix is said to be **block upper triangular** if it is of the form

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ & A_{22} \end{bmatrix}$$

Example 0.16.

$$\left[\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 4 & 5 & 6 & 0 & 1 \\ 7 & 8 & 9 & 3 & 3 \\ \hline & & & 1 & 1 \\ & & & 2 & 2 \end{array}\right]$$

Theorem 0.16. For a block upper triangular matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ & A_{22} \end{bmatrix},$$

if the two main blocks are both square and invertible, then $\mathbf{A}$ is also invertible, and

$$\mathbf{A}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ & A_{22}^{-1} \end{bmatrix}$$

Proof. By direct verification. □

Example 0.17. Find the inverse of

$$\left[\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 3 & 1 \\ \hline & & 4 \end{array}\right]$$

LU decomposition

In this part, we will derive a factorization scheme to express a given matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ as a product of two matrices of special forms

$$\mathbf{A} = \mathbf{L}\cdot\mathbf{U} = \begin{bmatrix} 1 & & & \\ * & 1 & & \\ \vdots & \vdots & \ddots & \\ * & * & * & 1 \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ & * & * & * & * \\ & & & * & * \\ & & & & * \end{bmatrix}$$

where $\mathbf{L} \in \mathbb{R}^{m\times m}$ is square, lower-triangular with 1's on the diagonal (called unit lower triangular), and $\mathbf{U} \in \mathbb{R}^{m\times n}$ is the REF of $\mathbf{A}$ (which is upper triangular).

Such a factorization is very useful for solving linear systems $\mathbf{A}\mathbf{x} = \mathbf{b}$.

For example, the following is an LU decomposition (verify this):

$$\underbrace{\begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} 1 & & \\ -1 & 1 & \\ 2 & -5 & 1 \end{bmatrix}}_{\mathbf{L}} \cdot \underbrace{\begin{bmatrix} 3 & -7 & -2 \\ & -2 & -1 \\ & & -1 \end{bmatrix}}_{\mathbf{U}}$$

To use it to solve the system of linear equations

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \qquad \text{where} \quad \mathbf{b} = \begin{bmatrix} -7 & 5 & 2 \end{bmatrix}^T$$

we first rewrite the equation as

$$\mathbf{A}\mathbf{x} = (\mathbf{L}\mathbf{U})\mathbf{x} = \mathbf{L}\underbrace{(\mathbf{U}\mathbf{x})}_{\mathbf{y}} = \mathbf{b}$$

and then solve two simper systems in the order

$$\mathbf{L}\mathbf{y} = \mathbf{b} \qquad \xrightarrow{\mathbf{y}} \quad \mathbf{U}\mathbf{x} = \mathbf{y}$$

That is, from the first equation, we obtain that $\mathbf{y} = \begin{bmatrix} -7 & -2 & 6 \end{bmatrix}^T$ and then use it to solve the second equation for $\mathbf{x} = \begin{bmatrix} 3 & 4 & -6 \end{bmatrix}^T$ (work done in class).

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

However, how to find such a decomposition in the first place will require the introduction of the so-called **elementary matrices**.

Elementary matrices

Elementary matrices are (square) matrices that can be obtained from the identity matrix through a single elementary row operation.

$$\mathbf{M}_i(r) = \begin{array}{c|ccccc|} & & & & i & \\ \hline & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ i & & & & r & \\ & & & & & 1 \\ \hline \end{array}$$

$$\mathbf{R}_{i\leftarrow j}(k) = \begin{array}{c|ccccc|} & & j & & & \\ \hline & 1 & & & & \\ j & & 1 & & & \\ & & & 1 & & \\ i & & k & & 1 & \\ & & & & & 1 \\ \hline \end{array}$$

$$\mathbf{P}_{ij} = \begin{array}{c|ccccc|} \hline j & 0 & & & 1 & \\ & & 1 & & & \\ & & & 1 & & \\ i & 1 & & & 0 & \\ & & & & & 1 \\ \hline \end{array}$$

Performing an elementary row operation on a given matrix can now is equivalent to matrix multiplication (the elementary matrix left multiplies the given matrix).

- $\mathbf{M}_i(r)$ - Multiply row i by a nonzero scalar r

$$\mathbf{M}_3(r)\mathbf{A} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & r \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ ra_{31} & ra_{32} & ra_{33} & ra_{34} \end{bmatrix}$$

- $\mathbf{R}_{i\leftarrow j}(k)$ - Add a scalar multiple (k) of one row (j) to another row (i) to replace that row (i):
  - Downward replacement

$$\mathbf{R}_{3\leftarrow 1}(k)\mathbf{A} = \begin{bmatrix} 1 & & \\ & 1 & \\ k & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ ka_{11} + a_{31} & ka_{12} + a_{32} & ka_{13} + a_{33} & ka_{14} + a_{34} \end{bmatrix}$$

– Upward replacement

$$\mathbf{R}_{1\leftarrow 3}(k)\mathbf{A} = \begin{bmatrix} 1 & & k \\ & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} + ka_{31} & a_{12} + ka_{32} & a_{13} + ka_{33} & a_{14} + ka_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

- Interchange two rows

$$\mathbf{P}_{12}\mathbf{A} = \begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

$$\mathbf{P}_{13}\mathbf{A} = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix}$$

$$\mathbf{P}_{23}\mathbf{A} = \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix}$$

An important fact

Elementary matrices are all invertible (because elementary row operations are all reversible)

$$\begin{aligned} \mathbf{M}_i(1/r) \cdot \mathbf{M}_i(r) &= \mathbf{I} \\ \mathbf{R}_{i \leftarrow j}(-k) \cdot \mathbf{R}_{i \leftarrow j}(k) &= \mathbf{I} \\ \mathbf{P}_{ij} \cdot \mathbf{P}_{ij} &= \mathbf{I} \end{aligned}$$

and their inverses are the same kind of elementary matrices!

$$\begin{aligned} \mathbf{M}_i(r)^{-1} &= \mathbf{M}_i(1/r) \\ \mathbf{R}_{i \leftarrow j}(k)^{-1} &= \mathbf{R}_{i \leftarrow j}(-k) \\ \mathbf{P}_{ij}^{-1} &= \mathbf{P}_{ij} \end{aligned}$$

Application of elementary matrices in finding matrix inverse

Previously we presented a procedure for finding the inverse of a square, invertible matrix

$$[\mathbf{A} \mid \mathbf{I}_n] \xrightarrow{\text{elementary row operations}} [\mathbf{I}_n \mid \mathbf{A}^{-1}]$$

This is equivalent to using a sequence of elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_\ell$ to left multiply the augmented matrix:

$$\mathbf{E}_\ell \cdots \mathbf{E}_2 \cdot \mathbf{E}_1 \cdot [\mathbf{A} \mid \mathbf{I}_n] = [\mathbf{I}_n \mid \mathbf{A}^{-1}]$$

Through matrix block multiplication, we obtain

$$[\mathbf{E}_\ell \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A} \mid \mathbf{E}_\ell \cdots \mathbf{E}_2\mathbf{E}_1] = [\mathbf{I}_n \mid \mathbf{A}^{-1}]$$

This shows that

$$\mathbf{A}^{-1} = \mathbf{E}_\ell \cdots \mathbf{E}_2\mathbf{E}_1$$

Application of elementary matrices in finding matrix REF

Similarly, give any matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, one can perform a sequence of elementary row operations through corresponding elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_\ell$ to transform the given matrix into its REF

$$\mathbf{E}_\ell \cdots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{U}$$

This yields that

$$\mathbf{A} = (\mathbf{E}_\ell \cdots \mathbf{E}_2 \mathbf{E}_1)^{-1} \mathbf{U} = \underbrace{\mathbf{E}_1^{-1} \mathbf{E}_2^{-1} \cdots \mathbf{E}_\ell^{-1}}_{\text{elementary matrices}} \mathbf{U}$$

Note that $\mathbf{U}$ (as REF) must be upper triangular.

Existence of the LU decomposition

In some cases, one only needs to use a sequence of downward replacement operations (i.e., $\mathbf{R}_{i\leftarrow j}(k)$ for $j < i$) to transform a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ into its REF $\mathbf{U} \in \mathbb{R}^{m\times n}$. That is,

$$\underbrace{\mathbf{E}_\ell \cdots \mathbf{E}_2\mathbf{E}_1}_{\text{all downward replacements}} \quad \mathbf{A} = \mathbf{U}$$

Then

$$\mathbf{A} = \underbrace{\mathbf{E}_1^{-1}\mathbf{E}_2^{-1} \cdots \mathbf{E}_\ell^{-1}}_{\text{also downward replacements}} \mathbf{U} = \underbrace{\mathbf{L}}_{\text{lower triangular}} \underbrace{\mathbf{U}}_{\text{REF}}$$

Remark. In other cases, one can always rearrange the rows of $\mathbf{A}$ in a way such that an LU decomposition exists.

Finding the $\mathbf{L}$ matrix

When a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ has an LU decomposition, we can find it as follows:

$$\mathbf{E}_\ell \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A} = \underbrace{\mathbf{U}}_{\text{REF}}$$

$$\mathbf{E}_\ell \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{L} = \underbrace{\mathbf{I}}_{\text{identity matrix}} \longleftarrow \mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\cdots\mathbf{E}_\ell^{-1}$$

That is, we will try to design a matrix $\mathbf{L}$ (lower triangular with 1's on the diagonal) so that the same row operations performed on $\mathbf{A}$ toward its REF will transform $\mathbf{L}$ into the identity matrix.

Example 0.18. Find the LU decomposition of

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

Example 0.19. Find the LU decomposition of

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