Linear Algebra

Lecture Notes

Rostyslav Hryniv

Ukrainian Catholic University
Computer Science Programme

3rd term
Autumn 2017

Lecture 2. Matrix algebra

Outline

Terminology

A matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

is a rectangular array of *numbers*, or *scalars*

- A has m rows and n columns; its size is $m \times n$
- $m \times 1$ matrix is a column vector; $1 \times n$ matrix is a row vector
- a_{ij} is the entry of A in i^{th} row and j^{th} column; $a_{ij} = (A)_{ij}$
- a shorthand notation: $A = (a_{ij})_{m \times n}$ or $A = (a_{ij})$
- entries $a_{11}, a_{22}, \dots$ form the main diagonal of A
- special matrices:
  - zero matrix: $a_{ij} = 0$;
  - square matrix if $m = n$
  - diagonal matrix $D = \operatorname{diag}\{d_1, d_2, \dots, d_n\}$
  - identity matrix: $I_n = \operatorname{diag}\{1, 1, \dots, 1\}$

Basic operations

- equality:

$$B = C \iff \text{the same size and } \forall i,j\colon\ (B)_{ij} = (C)_{ij}$$

- multiplication by scalars:

$$k \in \mathbb{R},\ A = (a_{ij})_{m\times n} \implies (kA) = (ka_{ij})_{m\times n} \quad \text{i.e.} \quad (kA)_{ij} = ka_{ij}$$

- addition/subtraction:

B, C of the same size $\implies (B \pm C)_{ij} = (B)_{ij} \pm (C)_{ij}$

- transposition $A^\top$ of $A = (a_{ij})_{m\times n}$:

$$A^\top \quad \text{is of size } n\times m \quad \text{and} \quad (A^\top)_{ij} = a_{ji}$$

i.e. rows of A become columns of $A^\top$

- trace of a square matrix $A = (a_{ij})_{n\times n}$:

$$\operatorname{tr} A := a_{11} + a_{22} + \cdots + a_{nn}$$

Examples

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

$$(-2)A = \begin{pmatrix} -4 & 2 & 0 \\ 2 & -6 & -2 \end{pmatrix} \qquad A + B = \begin{pmatrix} 2 & 1 & 1 \\ -2 & 3 & 4 \end{pmatrix}$$

$$A^\top = \begin{pmatrix} 2 & -1 \\ -1 & 3 \\ 0 & 1 \end{pmatrix} \qquad A^\top + B = \quad ?$$

Properties of matrix operations

Theorem

Let A, B, and C be matrices of the same size and r and s any scalars. Then the following relations hold:

1. $A + B = B + A$ *(commutative)*
2. $(A + B) + C = A + (B + C)$ *(associative)*
3. $A + \mathbf{0} = \mathbf{0} + A$ *(matrix zero)*
4. $r(A + B) = rA + rB$ *(distributive)*
5. $(r + s)A = rA + sA$ *(distributive)*
6. $r(sA) = (rs)A$ *(associative)*

Inner product of vectors

$A\mathbf{x} = \mathbf{b}$ is not just a symbolic shorthand notation for the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Inner product of row and column vectors:

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{r}_2(A)\mathbf{x}$$

Analogously,

$$b_i = \mathbf{r}_i(A) \cdot \mathbf{x} = \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_j a_{ij}x_j$$

Definition (Row-column multiplication)

The inner product of a $1 \times n$ (row) vector $\mathbf{u} = (u_1\ u_2\ \dots\ u_n)$ and a $n \times 1$ (column) vector $\mathbf{v} = (v_1\ v_2\ \dots\ v_n)^\top$ is

$$\mathbf{uv} = (u_1\ u_2\ \dots\ u_n)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n$$

Example

$$\begin{pmatrix} -2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} = (-2) \cdot 1 + 1 \cdot (-2) + 0 \cdot 2 = -4$$

Matrix-vector multiplication

- $A = (a_{ij})_{m\times n}$ a matrix; $\mathbf{x} = (x_1 \; x_2 \; \dots \; x_n)^\top$ a column vector

Row form of matrix-vector multiplication

Denote by $\mathbf{r}_i = (a_{i1} \; a_{i2} \; \dots \; a_{in})$ the i^{th} row of A; then

$$A\mathbf{x} = \begin{pmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_m \end{pmatrix} \mathbf{x} = \begin{pmatrix} \mathbf{r}_1\mathbf{x} \\ \vdots \\ \mathbf{r}_m\mathbf{x} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$$

Column form of matrix-vector multiplication:

Denote by $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$ the columns of A; then the product of A and $\mathbf{x}$ is

$$A\mathbf{x} = (\mathbf{c}_1 \quad \mathbf{c}_2 \quad \dots \quad \mathbf{c}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$$

In particular, $A\mathbf{x}$ is a *linear combination* of columns of A

Linear combinations, span, independence

Definition (Linear combination)

A linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$ is a vector of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k$ with any scalars $c_1, c_2, \dots, c_k$

Definition (Linear independence)

A collection of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ is linearly independent if no nontrivial linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ equals $\mathbf{0}$.

Definition (Linear span)

The set of all linear combinations of given vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ is called their linear span and denoted $\mathrm{ls}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$

Example

- Vectors $\mathbf{v}_1 = (1, 0, 1)^\top$ and $\mathbf{v}_2 = (0, 1, -1)^\top$ are linearly independent and span the plane $s\mathbf{v}_1 + t\mathbf{v}_2$ for $s, t \in \mathbb{R}$;
- the vectors $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3 := (2, 1, 1)^\top$ are linearly dependent as $2\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$

Example

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

- Row form:

$$A\mathbf{x} = \begin{pmatrix} 1 \cdot 2 + 2 \cdot (-1) + 3 \cdot 0 + 4 \cdot 1 \\ 4 \cdot 2 + 3 \cdot (-1) + 2 \cdot 0 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

- Column form:

$$A\mathbf{x} = 2 \cdot \begin{pmatrix} 1 \\ 4 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} + 0 \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

Matrix-matrix multiplication

Definition

If A is an $m \times n$ matrix and B is an $n \times k$ matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_k$, then the product AB of A and B is

$$AB = A\begin{pmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k\end{pmatrix} = \begin{pmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k\end{pmatrix}$$

and is an $m \times k$ matrix.

Corollary (Matrix product AB by column method)

- *The j^{th} column of AB is $A\mathbf{b}_j$*
- *Columns of AB are linear combinations of the columns of A*

Corollary (Row-by-column method)

The ij^{th} entry of AB is the product of i^{th} row of A and j^{th} column of B:

$$(AB)_{ij} = \sum_p (A)_{ip}(B)_{pj} = \mathbf{r}_i(A)\mathbf{c}_j(B)$$

Rows of AB

Vector-matrix multiplication

For a row vector $\mathbf{a} = (a_1 \ a_2 \ \ldots \ a_n)$ and an $n \times k$ matrix B with rows $\mathbf{r}_i(B)$ and columns $\mathbf{c}_j(B)$,

$$\begin{aligned} \mathbf{a}\, B &= \mathbf{a}\left(\mathbf{c}_1(B) \quad \mathbf{c}_2(B) \quad \ldots \quad \mathbf{c}_k(B)\right) \\ &= \left(\mathbf{a}\mathbf{c}_1(B) \quad \mathbf{a}\mathbf{c}_2(B) \quad \ldots \quad \mathbf{a}\mathbf{c}_k(B)\right) \end{aligned}$$

By combining a_i:

$$= a_1\mathbf{r}_1(B) + a_2\mathbf{r}_2(B) + \cdots + a_n\mathbf{r}_n(B)$$

so $\mathbf{a}\, B$ is a linear combination of the rows of B

Corollary (Matrix product AB by row method)

- *The i^{th} row of AB is* $\mathbf{r}_i(A)B$
- *Rows of AB are linear combinations of the rows of B*

Example

$$A = \begin{pmatrix} 0 & 2 & 1 \\ -1 & 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \qquad C = AB = (c_{ij})_{2\times 3}$$

$$c_{21} = \begin{pmatrix} * & * & * \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & * & * \\ 1 & * & * \\ 2 & * & * \end{pmatrix} = -1 \cdot 1 + 0 \cdot 1 + 2 \cdot 2 = 3$$

$$\mathbf{c}_1(C) = \begin{pmatrix} c_{11} \\ c_{21} \end{pmatrix} = A\mathbf{c}_1(B) = 1 \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

$$\begin{aligned} \mathbf{r}_2(C) = \begin{pmatrix} c_{21} & c_{22} & c_{23} \end{pmatrix} = \mathbf{r}_2(A)B &= -1 \cdot \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \\ &+ \ 0 \cdot \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} \\ &+ \ 2 \cdot \begin{pmatrix} 2 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 2 & -1 \end{pmatrix} \end{aligned}$$

Properties of matrix multiplication

Theorem

Let A be an $m \times n$ matrix and matrices B and C are of sizes permitting the multiplications below. Then

(a) $(AB)C = A(BC)$ *(associative law)*

(b) $(A+B)C = AC + BC$ *(right distributive law)*

(c) $A(B+C) = AB + AC$ *(left distributive law)*

(d) $r(AB) = (rA)B = A(rB)$ *(for every scalar r)*

(e) $I_m A = A = A I_n$ *(identity for matrix multiplication)*

(f) $(AB)^\top = B^\top A^\top$ *(multiplication and transposition)*

(g) $AB \neq BA$ *(no commutativity in general)*

(e) $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ *(order independence of trace)*

Proof.

Follow from the respective properties of scalar operations $\square$

Corollaries for linear systems

Solutions to a homogeneous system $A\mathbf{x} = \mathbf{0}$:

- If $\mathbf{u}$ is a solution of $A\mathbf{x} = \mathbf{0}$, so is $t\mathbf{u}$:
$$A(t\mathbf{u}) = t(A\mathbf{u}) = t\mathbf{0} = \mathbf{0}$$
- If $\mathbf{u}$ and $\mathbf{v}$ are solutions, so is $\mathbf{u} + \mathbf{v}$:
$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$
- If $\mathbf{u}$ and $\mathbf{v}$ are solutions, so is $s\mathbf{u} + t\mathbf{v}$ for all scalars s and t

Solutions to a nonhomogeneous system $A\mathbf{x} = \mathbf{b}$:

- If $\mathbf{u}$ and $\mathbf{v}$ are solutions, $\mathbf{u} - \mathbf{v}$ solves $A\mathbf{x} = \mathbf{0}$:
$$A(\mathbf{u} - \mathbf{v}) = A\mathbf{u} - A\mathbf{v} = \mathbf{b} - \mathbf{b} = \mathbf{0}$$
- If $\mathbf{u}$ is a particular solution of $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{v} = \mathbf{0}$, then
$$A(\mathbf{u} + t\mathbf{v}) = A\mathbf{u} + tA\mathbf{v} = \mathbf{b} + t\mathbf{0} = \mathbf{b}$$
- For $A\mathbf{x} = \mathbf{b}$ to be consistent, $\mathbf{b}$ must be a linear combination of the columns of A
- $A\mathbf{x} = \mathbf{b}$ consistent for all $\mathbf{b}$ $\iff$ columns of A span $\mathbb{R}^m$
- unique solution $\implies$ columns of A linearly independent

Invertible matrices

Definition

A square $n \times n$ matrix A is called invertible if there is a matrix B s.t. $AB = BA = I_n$; B is then called the inverse of A and denoted A^{-1}.

Example (Invertible A)

$$A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$$

$$AB = \begin{pmatrix} 2\cdot 3 + 5\cdot(-1) & 2\cdot(-5) + 5\cdot 2 \\ 1\cdot 3 + 3\cdot(-1) & 1\cdot(-5) + 3\cdot 2 \end{pmatrix} = I_2$$

$$BA = \begin{pmatrix} 3\cdot 2 + (-5)\cdot 1 & 3\cdot 5 + (-5)\cdot 3 \\ (-1)\cdot 2 + 2\cdot 1 & (-1)\cdot 5 + 2\cdot 3 \end{pmatrix} = I_2$$

Example (Non-invertible A)

$A = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$ is not invertible: rows of BA are linear combinations of $(2\ 4)$ and $(1\ 2)$ and cannot be $(1\ 0)$ or $(0\ 1)$

First properties of invertible matrices

Right- and left-invertibility

If A, B and C are square matrices s.t. $BA = AC = I$, then

$$B = C = A^{-1}$$

i.e., if A is left-invertible and right-invertible, then it is invertible

Proof.

$$B = B(AC) = (BA)C = C$$

□

Invertibility of a product

If A and B are invertible, then so is AB and $(AB)^{-1} = B^{-1}A^{-1}$

Proof.

$$(B^{-1}A^{-1})(AB) = ((B^{-1}A^{-1})A)B = (B^{-1}(A^{-1}A))B$$

$$= (B^{-1}I)B = B^{-1}B = I$$

$$(AB)(B^{-1}A^{-1}) = \dots$$

Elementary row transformations revisited

A is an $m \times n$ coefficient matrix of a system $A\mathbf{x} = \mathbf{b}$

1. Row multiplication: Multiply i^{th} row by t

Amounts to matrix multiplication EA, with

$$E^{-1} = \operatorname{diag}\{\underbrace{1, \ldots, 1}_{i-1}, t^{-1}, 1, \ldots, 1\}$$

2. Row replacement: Add α times k^{th} row to ℓ^{th} row

Amounts to EA, with $(E)_{ii} = 1$, $(E)_{\ell k} = \alpha$, $(E)_{ij} = 0$ otherwise:

$$E^{-1} = \begin{array}{c c} & \begin{array}{ccccc} & k & & & \\ & \downarrow & & & \end{array} \\ \begin{array}{c} \\ \ell \rightarrow \\ \\ \\ \end{array} & \begin{pmatrix} 1 & & & & \\ & & & & \\ & -\alpha & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \end{array}$$

Elementary row transformations revisited

3. Row interchange: Interchange k^{th} and ℓ^{th} rows

Amounts to matrix multiplication EA, with

$$E = E^{-1} = \begin{array}{c} \begin{pmatrix} 1 & k & & & \\ & \downarrow & 0 & 1 & \\ & & 1 & & \\ & 1 & 0 & \uparrow & \\ & & & \ell & 1 \end{pmatrix} \end{array} \begin{array}{l} \\ \leftarrow k \\ \\ \leftarrow \ell \\ \\ \end{array}$$

Definition

The above matrices performing the elementary row operations are called elementary matrices

Example: ERO via matrix multiplication

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix} \qquad \begin{matrix} \\ -2 \times (1) \\ -1 \times (1) \end{matrix} \qquad E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \\ 0 & -1 & 0 \end{pmatrix} \qquad \begin{matrix} \\ \\ -\frac{1}{3} \times (2) \end{matrix} \qquad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{3} & 1 \end{pmatrix}$$

$$E_2 E_1 A = \underbrace{\begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}}_{U} \implies$$

$$A = \underbrace{E_1^{-1} E_2^{-1}}_{L} U = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & \frac{1}{3} & 1 \end{pmatrix} U = LU$$

L: lower-triangular $\qquad$ U: upper-triangular

Products of lower-triangular matrices

Definition (Lower- and upper-triangular matrices)

A square matrix is called lower-triangular (upper-triangular) if all its entries above (below) the main diagonal are zero

Lemma

Product of two lower-triangular (upper-triangular) matrices is lower-triangular (upper-triangular)

Proof.

Use the row or column form of matrix-matrix product □

LU factorization

Theorem

Assume that an $m \times n$ matrix can be reduced to row echelon form U using only row substitution operations. Then $A = LU$ with a lower-triangular $m \times m$ matrix L.

Proof.

$$E_k \cdots E_1 A = U \implies L = (E_k \cdots E_1)^{-1} = E_1^{-1} \cdots E_k^{-1} \qquad \square$$

Definition

The above representation $A = LU$, with an $m \times m$ lower-triangular matrix L and upper-triangular* $m \times n$ matrix U, is the *LU-factorization* of A.

Remark

- *L is unique if all its diagonal entries are* 1
- *If row interchanges are needed, use $PA = LU$, with P encoding all row interchanges*

Why is LU-factorization important?

$$A\mathbf{x} = \mathbf{b} \iff \begin{cases} U\mathbf{x} & = \mathbf{y} \\ L\mathbf{y} & = \mathbf{b} \end{cases}$$

- Typically, $O(n^3)$ flops are needed to solve $A\mathbf{x} = \mathbf{b}$
- To solve $L\mathbf{y} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{y}$, need $O(n^2)$ flops
- E.g., elementary row operations transforming L to I_n only perform on $\mathbf{b}$
- Use to find A^{-1} for nonsingular A

Method of finding A^{-1}:

Computing A^{-1}:

Assume A is nonsingular and can be transformed to its reduced echelon form I_n applying elementary matrices $E_1, E_2, \dots, E_k$. Perform them on I_n; the result is A^{-1}.

Example

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 1 \end{array}\right) \quad \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -3 & 1 & -2 & 1 & 0 \\ 0 & -3 & 0 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & -3 & 0 & -1 & 0 & 1 \end{array}\right) \quad \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & -3 & 0 & -1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{3} & 1 & -\frac{1}{3} \\ 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & -3 & 0 & -1 & 0 & 1 \end{array}\right) \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{3} & 1 & -\frac{1}{3} \\ 0 & 1 & 0 & \frac{1}{3} & 0 & -\frac{1}{3} \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right)$$

Characterization of invertible matrices

Theorem

For an $n \times n$ matrix A, TFAE:

(a) *A is invertible*

(b) *A has n positions in its row echelon form*

(c) *The reduced row echelon form of A is I_n*

(d) *A is expressible as a product of elementary matrices*

(e) *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution*

(f) *$A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b} \in \mathbb{R}^n$*

(g) *There is an $n \times n$ matrix B s.t. $AB = I$*

(h) *There is an $n \times n$ matrix C s.t. $CA = I$*

(i) *$A^\top$ is invertible*

Proof. Already know: $(f) \iff (b) \iff (e)$

$$(b) \implies (c) \implies (d) \implies (a) \implies (h) \implies (e) \iff (b)$$

$(g) \implies (f) \iff (a) \implies (g)$. By transposing, $(i) \iff (g)$

Some corollaries

Corollary:

$$\text{Invertibility} \equiv \text{Non-singularity}$$

Corollary: For an $n \times n$ matrix A, TFAE:

(a) A is invertible
(j) The columns of A span $\mathbb{R}^n$
(k) The columns of A are linearly independent
(l) The rows of A span $\mathbb{R}^n$
(m) The rows of A are linearly independent

Corollary:

If A and B are square matrices and AB is invertible, then both A and B are invertible

Corollary:

A square upper- or lower-triangular matrix A is invertible $\iff$
$\iff$ no zeros on diagonal

Invertibility in a nutshell

To summarise:

- not all matrices are invertible
- an $n \times n$ matrix A is invertible $\iff$ columns of A are linearly independent $\iff$ columns of A span $\mathbb{R}^n$
- the inverse matrix is unique (if exists)
- elementary matrices (*row multiplication, row replacement, row interchange*) are invertible
- $\mathbf{x} := A^{-1}\mathbf{b}$ solves $A\mathbf{x} = \mathbf{b}$
- $(AB)^{-1} = B^{-1}A^{-1}$ for invertible A and B