

E600 Mathematics

Chapter 2: Matrix Algebra

Martin Reinhard

August 25/26, 2022

1. Introduction

Outline

In this chapter, we discuss

- The formal matrix concept and key definitions/types of matrices
- The matrix-based linear independence test
- Matrix inversion and its usefulness for solving equation systems
 - understand why matrix inversion works (elementary operations)
 - invert matrices using the Gauss-Jordan algorithm
- Key concepts related to matrices: rank, determinant, eigenvalues, definiteness, and how they are related

1. Introduction

Why Matrices?

- Linear equation systems are a fundamental building block of advanced mathematics
- Optimization: first and second derivatives of more general functions typically have matrix form
- Matrices are the default format of empirical data, and thus a central object type in Econometrics
- Matrices facilitate otherwise complex problems (recall: LI test)

... have ever wondered why economists use MATLAB so much?

1. Introduction

Matrices – an Equation System Motivation

Consider the system

$$\begin{array}{rccccrcr} x_1 & + & x_2 & + & x_3 & = & 6 \\ & & & & x_2 & - & x_3 & = & 0 \\ 5 \cdot x_1 & & & & & + & x_3 & = & 1 \end{array}$$

- Important questions of general interest:
 - 1 Does a solution exist?
 - 2 If so, is the solution unique? Otherwise, how many solutions are there?
 - 3 How can we (or our computer) efficiently derive the (set of) solutions?
- Matrix representation “ $Ax = b$ ”:

$$\begin{pmatrix} 1 \cdot x_1 + 1 \cdot x_2 + 1 \cdot x_3 \\ 0 \cdot x_1 + 1 \cdot x_2 + (-1) \cdot x_3 \\ 5 \cdot x_1 + 0 \cdot x_2 + 1 \cdot x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 1 \end{pmatrix} \Leftrightarrow \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 5 & 0 & 1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{=x} = \underbrace{\begin{pmatrix} 6 \\ 0 \\ 1 \end{pmatrix}}_{=b}$$

\Rightarrow we can use A (and b) to answer 1. – 3.!

2. The Vector Space of $n \times m$ -Matrices

Definition: Matrix and Basis Operations

Definition (Matrix of dimension $n \times m$)

Let $(a_{ij} : i \in \{1, \dots, n\}, j \in \{1, \dots, m\})$ be a collection of elements from a basis set X , i.e. $\forall i \in \{1, \dots, n\}, j \in \{1, \dots, m\} : a_{ij} \in X$. Then, the matrix A of these elements is the ordered two-dimensional collection

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}.$$

We call n the *row dimension* and m the *column dimension* of A . We write $A \in X^{n \times m}$, and $A = (a_{ij})_{i \in \{1, \dots, n\}, j \in \{1, \dots, m\}}$, or, if $n = m$, $A = (a_{ij})_{i, j \in \{1, \dots, n\}}$.

2. The Vector Space of $n \times m$ -Matrices

Representation of Matrices

- Real Matrix ($X = \mathbb{R}$) as a **vector of vectors**:

- $a_i^r = (a_{i1}, \dots, a_{im})$ i -th row of A (row vector)

- $a_j^c = (a_{1j}, \dots, a_{mj})' = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ j -th column of A (column vector)

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} a_1^r \\ \vdots \\ a_n^r \end{pmatrix} = (a_1^c, \dots, a_m^c)$$

- Typically: leave superscript r/c away if clear from context
- Vector space of real matrices: not vectors of real *numbers* (as \mathbb{R}^n), but vectors of real *vectors*!

2. The Vector Space of $n \times m$ -Matrices

Matrix Basis Operations and Vector Space of Matrices

Set of real $m \times n$ -Matrices: $\mathcal{M}_{n \times m} = \{A \in \mathbb{R}^{n \times m}\}$

Definition (Addition of Matrices)

For matrices A, B of *identical dimension*, i.e. $\exists n, m \in \mathbb{N} : A, B \in \mathcal{M}_{n \times m}$, their sum is obtained from addition of their elements, that is

$$A + B = (a_{i,j} + b_{i,j})_{i \in \{1, \dots, n\}, j \in \{1, \dots, m\}}.$$

Definition (Scalar Multiplication of Matrices)

Let $A = (a_{i,j})_{i \in \{1, \dots, n\}, j \in \{1, \dots, m\}}$ and $\lambda \in \mathbb{R}$. Then, scalar multiplication of A with λ is defined element-wise, that is

$$\lambda A := (\lambda a_{i,j})_{i=1, \dots, n, j=1, \dots, m}.$$

2. The Vector Space of $n \times m$ -Matrices

Norms in the Vector Space of Matrices

- Maximum Norm of $(\mathcal{M}_{n \times m}, +, \cdot)$:

$$\|A\|_{\infty} = \max\{|a_{ij}| : i \in \{1, \dots, n\}, j \in \{1, \dots, m\}\}$$

- 1 $\|A\|_{\infty} \geq 0$ and $(\|A\|_{\infty} = 0 \Leftrightarrow \forall i, j : a_{ij} = 0, \text{ i.e. } A = \mathbf{0}_{n \times m})$
 - 2 Triangle inequality transfers from $|\cdot|$
 - 3 Absolute homogeneity easily checked as well \Rightarrow norm ✓
- Can define norm-induced metrics, convergence, etc. like for \mathbb{R}^n !

2. The Vector Space of $n \times m$ -Matrices

Special Matrices 1/3

- $A \in \mathbb{R}^{n \times m}$ is **equal** to the matrix B if (and only if) $B \in \mathbb{R}^{n \times m}$ and $\forall i, j : a_{ij} = b_{ij}$ (e.g. $\mathbf{0}_{2 \times 2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{0}_{2 \times 3}$!)
- **Transposed** matrix $A'/A^T/A^t$: swap of row and column index, e.g.

$$\begin{pmatrix} 1 & 0 & 4 \\ 3 & 1 & 2 \end{pmatrix}' = \begin{pmatrix} 1 & 3 \\ 0 & 1 \\ 4 & 2 \end{pmatrix}$$

- **Vector** of length n : (column) vector $a \in \mathbb{R}^{n \times 1}$, row vector $a \in \mathbb{R}^{1 \times n}$
- **Zero matrix**: $\mathbf{0}_{n \times m} \in \mathbb{R}^{n \times m}$: $\mathbf{0}_{n \times m} = (a_{ij})_{i \in \{1, \dots, n\}, j \in \{1, \dots, m\}}$ where $\forall i, j : a_{ij} = 0$ (neutral element of addition!)

2. The Vector Space of $n \times m$ -Matrices

Special Matrices 2/3

- Square matrix: $A \in \mathbb{R}^{n \times m}$ with $n = m$, i.e. $A \in \mathbb{R}^{n \times n}$
- Symmetric matrix: *square* matrix $A \in \mathbb{R}^{n \times n}$ with $a_{ij} = a_{ji} \forall i, j$, e.g.

$$\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix} \quad \text{but not} \quad \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$$

- Diagonal matrix: $A = (a_{ij})_{i,j \in \{1, \dots, n\}}$ with $(i \neq j \Rightarrow a_{ij} = 0)$, e.g.

$$\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

We write $A = \text{diag}\{a_{11}, a_{22}, \dots, a_{nn}\}$

- Identity matrix: $\mathbf{I}_n = \text{diag}\{1, 1, \dots, 1\}$ (write down \mathbf{I}_3)

2. The Vector Space of $n \times m$ -Matrices

Special Matrices 3/3

Definition (Upper and Lower Triangular Matrix)

A **square** matrix $A = (a_{ij})_{i,j \in \{1, \dots, n\}} \in \mathbb{R}^{n \times n}$ is said to be *upper triangular* if $(i > j \Rightarrow a_{ij} = 0)$, i.e. when the entry a_{ij} equals zero whenever it lies below the diagonal. Conversely, A is said to be *lower triangular* if A' is upper triangular, i.e. $(i < j \Rightarrow a_{ij} = 0)$.

Example:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -4 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{for which} \quad A' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 1 & -4 & 0 \\ 4 & 0 & 3 & 2 \end{pmatrix}.$$

3. Matrix Calculus

Matrix Product Definition

Definition (Matrix Product)

Consider $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times k}$ where the *column dimension of A is equal to the row dimension of B*. Then, the matrix $C \in \mathbb{R}^{n \times k}$ of *column dimension equal to the one of A and row dimension equal to the one of B* is called the product of A and B, denoted $C = A \cdot B$, if

$$\forall i \in \{1, \dots, n\}, j \in \{1, \dots, k\} : c_{ij} = \sum_{l=1}^m a_{il} b_{lj}.$$

What is $A \cdot B$ if

- $A = \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & 2 & 4 \\ 0 & 1 & 3 \end{pmatrix}$?
- $A = \begin{pmatrix} -1 & 2 & 4 \\ 0 & 1 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}$? Commutativity?

3. Matrix Calculus

Product as Scalar Product of Rows and Columns

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, B = (b_1 \ \cdots \ b_k) \text{ (row/column notation). Then,}$$

$$AB = \begin{pmatrix} a'_1 \cdot b_1 & a'_1 \cdot b_2 & \cdots & a'_1 \cdot b_k \\ a'_2 \cdot b_1 & a'_2 \cdot b_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a'_{n-1} \cdot b_k \\ a'_n \cdot b_1 & \cdots & a'_n \cdot b_{k-1} & a'_n \cdot b_k \end{pmatrix}$$

Example:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 4 & 4 & 0 \\ -2 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 0 + 3 \cdot 5 & 1 \cdot 0 + 2 \cdot 1 + 3 \cdot (-3) \\ 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 5 & 0 \cdot 0 + 1 \cdot 1 + 0 \cdot (-3) \\ 4 \cdot 1 + 4 \cdot 0 + 0 \cdot 5 & 4 \cdot 0 + 4 \cdot 1 + 0 \cdot (-3) \\ (-2) \cdot 1 + 4 \cdot 0 + 1 \cdot 5 & (-2) \cdot 0 + 4 \cdot 1 + 1 \cdot (-3) \end{pmatrix} = \begin{pmatrix} 16 & -7 \\ 0 & 1 \\ 4 & 4 \\ 3 & 1 \end{pmatrix}$$

3. Matrix Calculus

Matrix Product: Key Properties 1/2

Theorem (Associativity and Distributivity of the Product)

Assuming that A, B, C are matrices of appropriate dimension, the product for matrices is

- (i) *Associative: $(AB)C = A(BC)$ (order of multiplication is irrelevant!)*
- (ii) *Distributive over matrix addition: $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$*

Theorem (Zero and Identity matrix)

Let $A \in \mathbb{R}^{n \times m}$. Then,

- (i) $A + \mathbf{0}_{n \times m} = A$.
- (ii) For any $k \in \mathbb{N}$, $A \cdot \mathbf{0}_{m \times k} = \mathbf{0}_{n \times k}$ and $\mathbf{0}_{k \times n} \cdot A = \mathbf{0}_{k \times m}$.
- (iii) For any $k \in \mathbb{N}$, $A \cdot \mathbf{I}_m = A$ and $\mathbf{I}_n \cdot A = A$.

Take-away: product (almost) behaves naturally! (exception: commutativity)

3. Matrix Calculus

Matrix Product: Key Properties 2/2

Theorem (Transposition, Sum, and Product)

- (i) Let $A, B \in \mathbb{R}^{n \times m}$. Then, $(A + B)' = A' + B'$
- (ii) Let $A \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{m \times k}$. Then, $(AB)' = B' A'$
- (iii) If $A \in \mathbb{R}^{1 \times 1}$, then A is actually a scalar and $A' = A$. (?!)

- Recall: scalar product compactly displays sum of products:

$$x \cdot y = x' y = \sum_{i=1}^n x_i y_i$$

- More generally: when $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ with $a_i, b_i \in \mathbb{R}^{1 \times m}$, then

$$A' B = \sum_{i=1}^n a'_i b_i \text{ where } a'_i b_i \in \mathbb{R}^{m \times m}$$

- Recall intuition: vectors of vectors vs. vectors of real numbers

4. Matrices and Linear Equation Systems

Introduction 1/2

- Let's verify our matrix representation $Ax = b$:

$$\begin{array}{rclcl} 1 \cdot x_1 & + & 1 \cdot x_2 & + & 1 \cdot x_3 & = & 6 \\ 0 \cdot x_1 & + & 1 \cdot x_2 & + & (-1) \cdot x_3 & = & 0 \\ 5 \cdot x_1 & + & 0 \cdot x_2 & + & 1 \cdot x_3 & = & 1 \end{array} \Leftrightarrow \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 5 & 0 & 1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{=x} = \underbrace{\begin{pmatrix} 6 \\ 0 \\ 1 \end{pmatrix}}_{=b}$$

- Interpretation: search of arguments x that solve $f_A(x) = b$ for $f_A(x) = Ax$

→ how to use A to characterize the solutions and determine them?

- Easiest scenarios: “square system” – as many equations as unknowns
- Simplest of all: one equation in one unknown: $f_a(x) = ax = b$, $a, b \in \mathbb{R}$
 - If $a \neq 0$, then unique solution $x = a^{-1}b = b/a$
 - If $a = 0$ and $b = 0$: infinitely many solutions $x \in \mathbb{R}$
 - If $a = 0$ and $b \neq 0$: no solution (“ b never reached by $a \cdot x$ ”)

4. Matrices and Linear Equation Systems

Introduction 2/2

- 1×1 system: generalization step
 - Unique solution if and only if a is **invertible**, i.e.

$$\exists a^{-1} \in \mathbb{R} : a^{-1}a = aa^{-1} = 1$$

$$\rightarrow x^* = a^{-1}b$$

- Otherwise: infinitely many/no solutions; $b \in \text{im}(f_a)$, $f_a(x) = ax$?
- For $n \times n$ systems ($Ax = b$):
 - Unique solution if and only if A is **invertible**, i.e.

$$\exists A^{-1} \in \mathbb{R}^{n \times n} : A^{-1}A = AA^{-1} = \mathbf{I}_n$$

$$\rightarrow x^* = A^{-1}b$$

- Otherwise: infinitely many/no solutions; $b \in \text{im}(f_A)$, $f_A(x) = Ax$?
- Now: (i) when $\exists A^{-1}$? (ii) when $b \in \text{im}(f_A)$?

4. Matrices and Linear Equation Systems

Invertability: Definition and Uniqueness

- Inverse matrix of **square** matrix $A \in \mathbb{R}^{n \times n}$: $A^{-1} \in \mathbb{R}^{n \times n}$ so that

$$A^{-1}A = AA^{-1} = \mathbf{I}_n$$

- The inverse matrix is unique!

Proof. Let $B, C \in \mathbb{R}^{n \times n}$ and assume that $BA = AB = \mathbf{I}_n$ and $CA = AC = \mathbf{I}_n$. Then, it holds that

$$C = C\mathbf{I}_n = C(AB) = \underbrace{(CA)}_{=\mathbf{I}_n} B = \mathbf{I}_n B = B,$$

i.e. $C = B$.

- (1) How to (dis)prove invertability? (2) How to find the inverse?

4. Matrices and Linear Equation Systems

Invertability: A Helpful Theorem

Proposition (Invertability: Transpose and Product)

Suppose that $A, B \in \mathbb{R}^{n \times n}$ are invertible. Then,

- A' is invertible and $(A')^{-1} = (A^{-1})'$,
- AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$,
- $\forall \lambda \in \mathbb{R}, \lambda \neq 0, \lambda A$ is invertible and $(\lambda A)^{-1} = 1/\lambda A^{-1}$.

Corollary: if A_1, \dots, A_k are invertible matrices, then

$$A_k^{-1} A_{k-1}^{-1} \dots A_2^{-1} \underbrace{A_1^{-1} A_1}_{=I_n} A_2 \dots A_k = I_n \Rightarrow \left(\prod_{i=1}^k A_i \right)^{-1} = \prod_{i=k}^1 A_i^{-1}$$

... but how to proceed more fundamentally?

4. Matrices and Linear Equation Systems

Elementary Matrix Operations 1/2

Definition (Elementary Matrix Operations)

$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ with rows $a'_1, \dots, a'_n \in \mathbb{R}^n$, $\tilde{A} = \begin{pmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_n \end{pmatrix}$. The three elementary matrix operations are mappings $A \rightarrow \tilde{A}$ given by

- (E1) Interchange of two rows i, j : $\tilde{a}_i = a_j$, $\tilde{a}_j = a_i$ and $\tilde{a}_k = a_k \forall k \notin \{i, j\}$,
- (E2) Scalar multiplication of a row i with $\lambda \neq 0$: $\tilde{a}_i = \lambda a_i$ and $\tilde{a}_j = a_j \forall j \neq i$,
- (E3) Addition of a multiple of row j to row i : $\tilde{a}_i = a_i + \lambda a_j$, $\lambda \in \mathbb{R}$, and $\tilde{a}_j = a_j$ for all $j \neq i$.

- Crucial tool for **finding the inverse** using Gauß-Jordan's algorithm
- Do not confuse with *basis* operations + and !

4. Matrices and Linear Equation Systems

Elementary Operations 2/2

- Elementary Operations (EO) can be represented by **left-multiplication** of operator matrix E : $\tilde{A} = EA$ (formal definition: see script)
- Generally: $e_i' A$ **extracts** the i -th row of A , and Ae_j the j -th column!
→ E stacks **row extraction** operators
- Examples for 2×2 system?
- Can perform multiple operations at once using the same logic!
- Example: 4×4 system, what do these EO matrices do?

$$E^1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad E^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad E^3 = \begin{pmatrix} -3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- How do EOs help in finding the inverse?

4. Matrices and Linear Equation Systems

Triangularization and Identity Transformation 1/3

- Suppose we can bring A to an identity matrix using only EOs E_1, \dots, E_k , i.e. $E_k E_{k-1} \dots E_1 A = \mathbf{I}_n$
- Then, $E = E_k E_{k-1} \dots E_1$ **fits the definition of A^{-1} !**
- $A^{-1} = E = E \cdot \mathbf{I}_n \rightarrow$ Gauß-Jordan algorithm
- General systems ($n \neq m$): EOs help characterize the number of solutions! Key concept: (generalized) upper triangular form

Definition (Generalized Upper Triangular Matrix)

Let $A \in \mathbb{R}^{n \times m}$. If $n \geq m$, A has generalized upper triangular form if $A = \begin{pmatrix} A_T \\ \mathbf{0}_{n-m \times m} \end{pmatrix}$ where $A_T \in \mathbb{R}^{m \times m}$ is upper triangular. If $n \leq m$, A has generalized upper triangular form if $A = (A_T, X)$ where $A_T \in \mathbb{R}^{m \times m}$ is upper triangular and X is an arbitrary $n \times m - n$ matrix.

4. Matrices and Linear Equation Systems

Triangularization and Identity Transformation 2/3

Examples Generalized Upper Triangular Form: A, B, D , but not C

$$A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 4 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 1 & 4 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Theorem (Triangulizability of a Matrix)

Consider a matrix $A \in \mathbb{R}^{n \times m}$. Then, if $n = m$, A can be brought to upper triangular form applying *only elementary operations* to A . Generally, A can be brought to generalized upper triangular form.

(Quite extensive) proof can be found in script

4. Matrices and Linear Equation Systems

Triangularization and Identity Transformation 3/3

- Recall: If we can bring A to identity form, then we know how to invert it
- From the upper triangular form, the identity form is usually not far away: take for instance

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{pmatrix}$$

- Condition (first step): non-zero diagonal
- Let's find more general (and direct) invertability conditions; concepts will also help characterize general equation systems
- Afterwards: solving (we actually already know how to)

5. Determinant, Ranks, Definiteness and Eigenvalues

Intro Determinant 1/4: Intuition

- Consider $f_A(x) = Ax$ for a **square** matrix $A \in \mathbb{R}^{n \times n}$
 - Ax is proportional to x : $A(\lambda x) = \lambda(Ax)$ for all $\lambda \in \mathbb{R}$

- $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ implies a **transformation of axes** of \mathbb{R}^2 :

$$f_A \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f_A \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 3 \end{pmatrix} = 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- A implies *expansion* of axes, summarized by $\det(A) = 6 = 2 \cdot 3$
- $\det(A) \neq 0$: expansion can be reverted, therefore A is invertible!

- $A = \begin{pmatrix} 2 & 3 \\ 0 & 0 \end{pmatrix}$ implies the transformation of axes

$$f_A \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f_A \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 3 \\ 0 \end{pmatrix} = 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- A implies an *infinite compression* of axes (losing the second axis), summarized by $\det(A) = 0 = 2 \cdot 0 \Rightarrow$ not invertible!

5. Determinant, Ranks, Definiteness and Eigenvalues

Intro Determinant 2/4: Definition

- A_{-ij} : A w/o column i and row j , e.g.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 4 \end{pmatrix} \rightarrow A_{-23} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

Definition (Determinant)

Let $A \in \mathbb{R}^{n \times n}$. Then, we define the determinant of A , denoted $\det(A)$ or $|A|$, as

(i) if $n = 1$ and $A = (a)$ is a scalar, $\det(A) = \det(a) := a$.

(ii) for all $n \in \mathbb{N} : n \geq 2$, when $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$,

$$\det(A) := \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{-ij}) \text{ with } i = 1.$$

5. Determinant, Ranks, Definiteness and Eigenvalues

Intro Determinant 3/4

- **Laplace Expansion Theorem:** It holds that
 - $\det(A) := \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{-ij})$ for arbitrary $i \in \{1, \dots, n\}$
(expansion by the i -th row)
 - $\det(A) := \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{-ij})$ for arbitrary $j \in \{1, \dots, n\}$
(expansion by the j -th column)
- Look for rows/columns with many zeros (e.g. A on slide before)!
- Determinant of “small” matrices:
 - (i) If $n = 2$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $\det(A) = ad - bc$.
 - (ii) If $n = 3$ and $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, then
$$\det(A) = aei + bfg + cdh - (ceg + bdi + afh).$$

5. Determinant, Ranks, Definiteness and Eigenvalues

Intro Determinant 4/4

- Usually: do Laplace expansions until one may use the 3x3 rule
- We typically consider at most 4x4 matrices analytically
- 3x3 graphically:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

- Example: $\det \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ -1 & 0 & 4 \end{pmatrix}$

5. Determinant, Ranks, Definiteness and Eigenvalues

Invertability: Determinant and Triangular Matrices

- The determinant of triangular matrices is equal to the **trace** (product of diagonal elements)

$$\begin{aligned}\det(A) &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{32} & \cdots & a_{3n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} = a_{11} a_{22} \det \begin{pmatrix} a_{33} & a_{34} & \cdots & a_{3n} \\ 0 & a_{43} & \cdots & a_{4n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} \\ &= \dots = \prod_{i=1}^n a_{ii} = \text{tr}(A)\end{aligned}$$

- $\det(A) \neq 0 \Leftrightarrow (\forall i \in \{1, \dots, n\} : a_{ii} \neq 0)$ (does this look familiar?)
- Triangular A is invertible if and only if $\det(A) \neq 0$!
- This holds also for non-triangular A , let's see why

5. Determinant, Ranks, Definiteness and Eigenvalues

Invertability: Determinant and Elementary Operations

Theorem (Determinant and Elementary Operations)

Let $A \in \mathbb{R}^{n \times n}$ and \tilde{A} the resulting matrix for the respective EO. Then,

- (i) for (E1) (interchange of two rows), we have $\det(\tilde{A}) = -\det(A)$,
- (ii) for (E2) (row multiplication with a scalar $\lambda \neq 0$), $\det(\tilde{A}) = \lambda \det(A)$,
- (iii) for (E3) (addition of multiple of row to another row), $\det(\tilde{A}) = \det(A)$, i.e. (E3) does not change the determinant.

- EOs do not affect the property “ $\det \neq 0$ ”!
- So, for any square matrix A :
 - We can use EOs to bring A to a triangular matrix \tilde{A}
 - We can use EOs to bring \tilde{A} to \mathbf{I}_n if $\det(\tilde{A}) \neq 0$
 - By the theorem, $\det(\tilde{A}) \neq 0$ if and only if $\det(A) \neq 0$

⇒ **we can invert A if $\det(A) \neq 0$!**

5. Determinant, Ranks, Definiteness and Eigenvalues

Determinant: Equivalence and Central Result

- So far: “ $\det(A) \neq 0 \Rightarrow A$ invertible”
 - Product rule allows to show “ A invertible $\Rightarrow \det(A) \neq 0$ ”
 - Product rule: $\det(AB) = \det(A) \det(B)$ (if A, B are square!)

\Rightarrow **we can invert A if and only if $\det(A) \neq 0$!**

\Rightarrow Square systems $Ax = b$: **unique solution $x^* = A^{-1}b \Leftrightarrow \det(A) \neq 0$**

Conclusion: A square matrix is invertible if and only if it has non-zero determinant! In a Square System, we (almost) always use the determinant invertability criterion!

5. Determinant, Ranks, Definiteness and Eigenvalues

Determinant: Helpful Properties

- Product: $\det(AB) = \det(A) \det(B)$ (if A, B are square!)
- Inverse Matrix: $\det(A^{-1}) = 1/\det(A)$
- Scalar Product: If $A \in \mathbb{R}^{n \times n}$, then $\det(\lambda A) = \lambda^n \det(A)$
 - follows from the product rule with $\lambda A = \text{diag}\{\lambda, \dots, \lambda\}A$

5. Determinant, Ranks, Definiteness and Eigenvalues

Intro Rank 1/3

- Why another criterion for characterizing the solution set? Not all systems are square!
- Recall: a solution exists if we can “reach” b using $f_A(x) = Ax$
- Example:

$$x_1 + 2x_2 = 2$$

$$x_1 - x_2 = 0$$

Matrix form

$$Ax = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 2 \\ -1 \end{pmatrix} x_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = b$$

Column space criterion: A solution exists if and only if

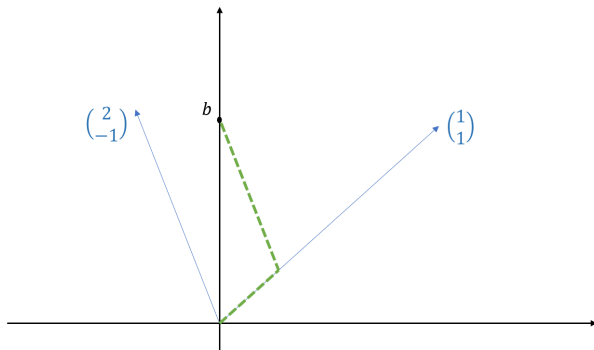
$$b \in \text{Span} \left(\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\} \right) =: \text{Co}(A)$$

Column space: spanned by the columns of A ; x 's as LC multipliers!

5. Determinant, Ranks, Definiteness and Eigenvalues

Intro Rank 2/3

Column space criterion: “reaching” the point b



Note: condition for *existence*, not *uniqueness*!

5. Determinant, Ranks, Definiteness and Eigenvalues

Intro Rank 3/3

- $b \notin f_A[\mathbb{R}^n]$: no solution = **information conflict**
 - Why “information conflict”? Rows provide contradicting info, e.g.

$$\begin{array}{rcl} x_1 & + & x_2 = 2 \\ 2x_1 & + & 2x_2 = 5 \end{array}$$

where

$$Co(A) = Span \left(\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\} \right) = \left\{ \begin{pmatrix} x \\ 2x \end{pmatrix} : x \in \mathbb{R} \right\} \not\ni \begin{pmatrix} 2 \\ 5 \end{pmatrix} = b$$

- Can occur in any (potentially non-square) system with ≥ 2 equations
- Can be more complex/non-obvious (especially for larger systems)
- Number of LI columns of A = dimension of column space
 - LI columns of A are basis of $Co(A)$
 - The more dimensions $Co(A)$ has, the more “likely” $b \in Co(A)$

5. Determinant, Ranks, Definiteness and Eigenvalues

Rank: Definition and Equivalence

- **Column rank**: Number of LI columns of A
- **Row rank**: Number of LI rows of A
- Column rank is crucial for solution existence
- Theorem: Column rank = row rank =: **Rank**, denoted $\text{rk}(A)$ or $\text{rk } A$
- Corollary: rank bound $\text{rk } A \leq \min\{m, n\}$ for $n \times m$ matrix A
- $n \times m$ matrix A : **full row (column) rank** if $\text{rk } A = n$ ($\text{rk } A = m$)
- If $m = n = \text{rk } A$, then A has **full rank**
- **The rank is unchanged by elementary operations (E1) to (E3)!**

5. Determinant, Ranks, Definiteness and Eigenvalues

Rank and General Equation Systems

Proposition (Elementary Operations and the Set of Solutions)

Consider a system $Ax = b$ of linear equations. Then, for an elementary operation characterized $\tilde{A} = EA$ with operation matrix E , the system $\tilde{A}x = \tilde{b}$ with $\tilde{b} = Eb$ is equivalent to $Ax = b$ in terms of the solutions x .

In other words, **EOs do not change the set of solutions!**

Theorem (Rank Condition)

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Then, the system $Ax = b$ has a unique solution if and only if $b \in \text{Co}(A)$ and $\text{rk}(A) = n$.

- m equations in n unknowns (!! $m \times n$ -matrices from now on !!)
- Intuition $\text{rk } A = n$: one LI column for every direction into which b expands \rightarrow unique LC of columns

5. Determinant, Ranks, Definiteness and Eigenvalues

Non-square Systems 1/2

- Definitions:
 - Under-identified system: $m < n$, strictly less equations than unknowns
 - Over-identified system: $m > n$, strictly more equations than unknowns
- Under-identified system: $\text{rk } A \leq m < n \Rightarrow$ no unique solution
 - Intuition: Column space interpretation; e.g.

$$\lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + \lambda_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

- More LC coefficients than dimensions in $b \rightarrow$ no unique LC!
- “Information deficit” in pinning down n -dim. LC

5. Determinant, Ranks, Definiteness and Eigenvalues

Non-square Systems 2/2

- Over-identified system: $m > n$: $A = \begin{pmatrix} A_{n \times n} \\ A_{(m-n) \times n} \end{pmatrix}$
 - Recall: we can bring A to gen. upper triangular form with EOs E :

$$\tilde{A} = \begin{pmatrix} A_T \\ \mathbf{0}_{(m-n) \times n} \end{pmatrix} = EA$$

- $\tilde{A}x = Eb = \tilde{b}$ features the **same solutions!** When are there any? E.g.

$$\begin{pmatrix} 2 & 3 \\ 0 & 4 \\ 0 & 0 \end{pmatrix} x = \begin{pmatrix} 2 \\ -1 \\ c \end{pmatrix} \rightarrow c \stackrel{!}{=} 0$$

- No solution if \tilde{b} has non-zero entries below index n
- Otherwise: omit “ $0 = 0$ ” rows (“delete redundant information”) and investigate square system

5. Determinant, Ranks, Definiteness and Eigenvalues

Non-square Systems: Summary

- Under-identified systems: insufficient information to pin down a unique solution
 - ⇒ Usually not interesting
- Over-identified systems
 - Include contradicting information/rows, or
 - are actually a square system with added redundant information
 - In latter case: triangularize system and solve the square remainder (example?)

⇒ **Any system with a unique solution is or can be reduced to a square system**
- Square system: $\text{rk } A = n \Leftrightarrow \text{Co}(A) = \mathbb{R}^n$
 - $b \in \text{Co}(A)$ guaranteed \rightarrow no information conflict
 - Unique solution exists if and only if $\text{rk } A = n$ (cf. rank condition)

5. Determinant, Ranks, Definiteness and Eigenvalues

Rank, Determinant and Linear Equation Systems

- As a corollary of what we did so far, we get this powerful result:

Corollary (Invertability, Rank and Determinant of Square Matrices)

$A \in \mathbb{R}^{n \times n}$ (*square matrix*). Then, the following statements are equivalent:

- (i) A is invertible,
- (ii) $\det(A) \neq 0$,
- (iii) $\text{rk } A = n$,
- (iv) for any $b \in \mathbb{R}^n$, $Ax = b$ has a unique solution for x ,
- (v) Any triangular matrix \tilde{A} obtained from applying EOs to A has only non-zero diagonal entries.

5. Determinant, Ranks, Definiteness and Eigenvalues

Testing Linear Independence 1/2

Theorem (Testing Linear Independence)

A equivalent condition for the set of vectors $B = \{b_1, b_2, \dots, b_k\}$, $b_j \in \mathbb{R}^n$, to be linearly independent is that

$$\sum_{j=1}^k \lambda_j b_j = \mathbf{0} \Rightarrow (\forall j \in \{1, \dots, k\} : \lambda_j = 0).$$

- Ch. 1 promised a LI check using matrices
- $\sum_{j=1}^k \lambda_j b_j = \mathbf{0}$ is an equation system (LC of columns)!
- Matrix form: $B_{mat} \lambda = \mathbf{0}$ with $\lambda = (\lambda_1, \dots, \lambda_k)'$,
 $B_{mat} = (b_1, b_2, \dots, b_k)$
- We need (i) a unique solution (ii) equal to $\lambda = \mathbf{0}$

5. Determinant, Ranks, Definiteness and Eigenvalues

Testing Linear Independence 2/2

Theorem (Testing Linear Independence)

A equivalent condition for the set of vectors $B = \{b_1, b_2, \dots, b_k\}$, $b_j \in \mathbb{R}^n$, to be linearly independent is that

$$\sum_{j=1}^k \lambda_j b_j = \mathbf{0} \Rightarrow (\forall j \in \{1, \dots, k\} : \lambda_j = 0).$$

- $k > n$: solution can not be unique (more vectors than dimensions)
 - $k > n$ vectors of length n can not be linearly independent
- $k = n$: B_{mat} is square \rightarrow apply determinant criterion
 - If $\det(B_{mat}) \neq 0$, unique solution $\lambda = B_{mat}^{-1} \mathbf{0} = \mathbf{0}$
- $k < n$: Bring B_{mat} to general upper triangular form
 - Equivalent system: $B_{mat} \lambda = \mathbf{0} \Leftrightarrow EB_{mat} \lambda = E \mathbf{0} = \mathbf{0}$
 - Drop zero rows \rightarrow square system, apply determinant criterion
 - Example?

5. Determinant, Ranks, Definiteness and Eigenvalues

Eigenvectors and Eigenvalues

- Eigenvalues and definiteness are linked to inversion but also important in their own right
- **Only square matrices have eigenvalues and are (in)definite!**
- Def. **Eigenvalue** $\lambda \in \mathbb{R}$ of $A \in \mathbb{R}^{n \times n}$: $\exists x \in \mathbb{R}^n \setminus \{0\} : Ax = \lambda x$
 - $x \in \mathbb{R}^n \setminus \{0\}$ with $Ax = \lambda x$ is called an **eigenvector** of the eigenvalue λ
 - Eigenvectors are not unique: $Ax = \lambda x \Rightarrow A(cx) = \lambda(cx) \forall c \in \mathbb{R}$
 - **Eigenspace** of λ : $\text{Span}(\{x \in \mathbb{R}^n : Ax = \lambda x\})$
- How to find eigenvalues?

$$Ax = \lambda x = \lambda \cdot (\mathbf{I}_n x) \Leftrightarrow \mathbf{0} = Ax - \lambda \mathbf{I}_n x = (A - \lambda \mathbf{I}_n)x.$$

\Rightarrow Eigenvectors solve this equation system!

5. Determinant, Ranks, Definiteness and Eigenvalues

Finding Eigenvalues

- Thus far: interested in unique solutions and invertible matrices
- Here: candidate unique solution $x = (A - \lambda \mathbf{I}_n)^{-1} \mathbf{0} = \mathbf{0}$ is *not* an eigenvector!

→ Interested in scenarios where $A - \lambda \mathbf{I}_n$ is *not* invertible:

$$\mathcal{P}(\lambda) = \det(A - \lambda \mathbf{I}_n) = 0$$

→ Infinitely many solutions/eigenvectors ($\neq \mathbf{0}$), and λ is eigenvalue!

- $\mathcal{P}(\lambda)$: **characteristic polynomial** of A , we want to know its **roots**:

$$\text{Eigenvalues}(A) = \{\lambda \in \mathbb{R} : \mathcal{P}(\lambda) = 0\}$$

- Example: Self-study exercises, online exercises

5. Determinant, Ranks, Definiteness and Eigenvalues

Proposition (Eigenvalues and Invertability)

Let $A \in \mathbb{R}^{n \times n}$. Then, A is invertible if and only if all eigenvalues of A are non-zero.

Proof. A is invertible if and only if $0 \neq \det(A) = \det(A - 0 \cdot \mathbf{I}_n)$, which is the case if and only if 0 is not an eigenvalue of A . \square

- If you know the eigenvalues, you can skip the det-check
- If you know the determinant, you know whether 0 is an eigenvalue

5. Determinant, Ranks, Definiteness and Eigenvalues

Definiteness of Symmetric Matrices

Definition (Definiteness of a Matrix)

A **symmetric** square matrix $A \in \mathbb{R}^{n \times n}$ is called

- *positive semi-definite* if $\forall x \in \mathbb{R}^n : x'Ax \geq 0$
- *negative semi-definite* if $\forall x \in \mathbb{R}^n : x'Ax \leq 0$
- *positive definite* if $\forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\} : x'Ax > 0$
- *negative definite* if $\forall x \in \mathbb{R}^n \setminus \{\mathbf{0}\} : x'Ax < 0$

Otherwise, it is called *indefinite*.

- Why exclude the zero vector for definiteness?
- Intuition: “smaller/greater than zero” equivalent for matrices
- Fundamentally important for optimization: sign of second derivative

5. Determinant, Ranks, Definiteness and Eigenvalues

Definiteness, Eigenvalues and Invertability

Proposition (Definiteness and Eigenvalues)

A symmetric square matrix $A \in \mathbb{R}^{n \times n}$ is

- (i) positive (negative) definite if and only if all eigenvalues of A are strictly positive (negative).
- (ii) positive (negative) semi-definite if and only if all eigenvalues of A are strictly non-negative (non-positive).

Corollary (Definiteness and Invertability)

If $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite or negative definite, it is invertible.

Positive/negative definiteness is **sufficient** for invertability!

Context: OLS (see Chapter 5)

6. Matrix Inversion

Introduction and Gauß-Jordan Validity

- Earlier: if there is an identity transformation of A using EOs, i.e. $EA = \mathbf{I}_n$, then $A^{-1} = E = E\mathbf{I}_n$ (“Gauß-Jordan method”)
 - This means that “ $EA = \mathbf{I}_n$ for an EO matrix $E \Rightarrow A$ invertible”
 - Now: is it also true that “ A invertible $\Rightarrow EA = \mathbf{I}_n$ for an EO matrix E ”?
 - *That is, does Gauß-Jordan always find the inverse?*
- Yes, and you know the arguments already!

6. Matrix Inversion

Gauß-Jordan Validity: Line of Reasoning

- Gauß-Jordan Validity:
 - If $A \in \mathbb{R}^{n \times n}$ is invertible, then $\text{rk } A = n$
 - Because A is square, we can triangularize it to $\tilde{A} = EA$
 - The rank is preserved under EOs, so that $\text{rk } \tilde{A} = n$
 - Thus, \tilde{A} has only non-zero diagonal elements (cf. summary theorem)

$$\tilde{A} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ 0 & \tilde{a}_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tilde{a}_{2,n-1} \\ 0 & \cdots & 0 & \tilde{a}_{nn} \end{pmatrix} \rightarrow \hat{A} = \begin{pmatrix} 1 & \hat{a}_{12} & \cdots & \hat{a}_{1n} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \hat{a}_{2,n-1} \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

→ eliminate the upper triangle using the ones iteratively (right to left) ✓
This gives ...

6. Matrix Inversion

Gauß-Jordan Theorem

Theorem (Gauß-Jordan Algorithm Validity)

Suppose that $A \in \mathbb{R}^{n \times n}$ is an invertible matrix. Then, we can apply elementary operations E_1, \dots, E_k in ascending order of the index to A to arrive at the identity matrix \mathbf{I}_n , and the inverse can be determined as $A^{-1} = E_k \cdot \dots \cdot E_1$.

Example: Define

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

Can you invert A ? If so, what is its inverse?

6. Matrix Inversion

Inverting 2×2 Matrices

Proposition (Inverse of a 2×2 Matrix)

Consider the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $ad \neq bc$. Then,

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Proof: See script (can use Gauß-Jordan, easier: verify $A^{-1}A = \mathbf{I}_n$).