Theory sheet 3

Motivation for matrix inversion

When we are solving linear equation $ax = y$ for x, we simply divide by a and obtain $x = y/a$. As it turns out, a much more complicated problem of finding solutions to square systems of linear equations

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nn}x_n & = & b_n \end{cases}$$

may be performed in a very similar way. This system of equations may be rewritten in the matrix form as

$$A\vec{x} = \vec{b}, \tag{1}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

It turns out, the Equation (1) may be solved for $\vec{x}$ by

$$\vec{x} = A^{-1}\vec{b}$$

with some new matrix A^{-1} called the inverse of A. Below we discuss how to define and how to find this matrix.

Definition and properties of the matrix inverse

Definition 1. *If $A \in M_{n,n}(\mathbb{R})$ is a square matrix and there exists another square matrix $B \in M_{n,n}(\mathbb{R})$ such that*

$$AB = BA = I,$$

then we say that A is invertible and that B is the inverse of A. We denote this by

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

For example, if $A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$, then we can check that $AB = BA = I$:

$$AB = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 3 \cdot 2 + 5 \cdot (-1) & 3 \cdot (-5) + 5 \cdot 3 \\ 1 \cdot 2 + 2 \cdot (-1) & 1 \cdot (-5) + 2 \cdot 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

and

$$BA = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot 3 + (-5) \cdot 1 & 2 \cdot 5 + (-5) \cdot 2 \\ (-1) \cdot 3 + 3 \cdot 1 & (-1) \cdot 5 + 3 \cdot 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Here are a few remarks.

- Matrix inverse is <u>not</u> defined componentwise: $(A^{-1})_{ij} \neq 1/(A)_{ij}$.
- If A contains a zero column or a zero row, it is <u>not</u> invertible, because

$$\begin{pmatrix} b_{11} & b_{12} & & \\ b_{21} & b_{22} & & \\ & & \ddots & \\ & & & b_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & a_{nn} \end{pmatrix} = \begin{pmatrix} * & * & & & \\ * & * & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & * \end{pmatrix}$$

and

$$\begin{pmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & & \\ b_{21} & b_{22} & & \\ & & \ddots & \\ & & & b_{nn} \end{pmatrix} = \begin{pmatrix} * & * & & & \\ * & * & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & * \end{pmatrix},$$

so the matrix on the right <u>cannot</u> be I because it has zero column or row.

- I is clearly the inverse of itself: $I = I^{-1}$, but I is not the only matrix with this property:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- <u>Not every matrix is invertible!</u> In fact, the following theorem is true (without proof):

Theorem 1. *Let $A \in M_{n,n}(\mathbb{R})$. Then*

$$A \textit{ is invertible} \iff \operatorname{Rank} A = n \iff \boxed{\det A \neq 0}.$$

We will only prove the " $\implies$ " direction:

Proof. If A is invertible, then there exists A^{-1} such that $I = AA^{-1}$. Taking $\det$ of both sides, we obtain $\det I = \det(AA^{-1})$. Since $\det(AB) = \det(A) \cdot \det(B)$, on the right we have $\det(A) \cdot \det(A^{-1})$, whereas on the left $\det I = 1$. Hence,

$$\det(A) \cdot \det(A^{-1}) = 1,$$

and therefore $\det A \neq 0$. $\square$

Properties of the matrix inverse

Here's a list of the properties of matrix inverse:

1. If $AB = I$, then $BA = I$ automatically, so A and B are invertible. *In other words, we don't need to check two equalities* $AB = I$ *and* $BA = I$, *checking one is enough.*

Proof.
- $AB = I \implies \det(AB) = \det A \cdot \det B = \det I = 1$
- $\implies \det A, \det B \neq 0$
- By theorem above B is invertible
- Multiplying $AB = I$ by B on the right we obtain $(AB)B^{-1} = IB^{-1}$ or $A = B^{-1}$.
- Multiplying this by B on the left we obtain $BA = BB^{-1} = I$. □

2. Note that there exist non-square matrices $A \in M_{n,p}$ and $B \in M_{p,n}$ such that $AB = I$. This does not mean that A or B are invertible. Only square matrices may be invertible.

3. If A is invertible, the inverse is unique. In other words, the equalities $AB = I$ and $AC = I$ cannot be satisfied with $B \neq C$.

Proof. Let B, C be two inverses of A. Multiplying the equality $AB = I$ by C on the left we obtain $C(AB) = C$ or $(CA)B = C$ or $B = C$. □

4. $AA^{-1} = I$ and $A^{-1}A = I$.

5. $(A^{-1})^{-1} = A$.

Proof. (a) Multiplying $I = AA^{-1}$ by $(A^{-1})^{-1}$ on the right we obtain

$$I(A^{-1})^{-1} = AA^{-1}(A^{-1})^{-1}$$

(b) On the left, we have $(A^{-1})^{-1}$.

(c) The expression in blue is equal to I by definition of $(A^{-1})^{-1}$. □

6. $(A^\top)^{-1} = (A^{-1})^\top$. We say that inversion commutes with transposition.

Proof. Taking the transpose of $I = AA^{-1}$ we obtain

$$I^\top = (AA^{-1})^\top.$$

Since $I^\top = I$ and $(AA^{-1})^\top = (A^{-1})^\top A^\top$ (recall that transposition flips the oder of matrix product), we obtain

$$I = (A^{-1})^\top A^\top.$$

Hence, $(A^{-1})^\top$ is the inverse of $A^\top$. □

7. By previous point, if A is symmetric, then A^{-1} is symmetric.

8. $\boxed{(A + B)^{-1} \neq A^{-1} + B^{-1}}$. Take $A = B = I$ for example.

9. If A and B are invertible, then AB is also invertible and $\boxed{(AB)^{-1} = B^{-1}A^{-1}}$. *Note that matrix inversion also flips the order!*

Proof.
- Invertibility follows from $\det(AB) = \det A \cdot \det B$. Indeed, if $\det A, \det B \neq 0$, then $\det(AB) \neq 0$ and by the theorem above AB is invertible.
- Hence, $I = (AB)(AB)^{-1}$.
- Multiplying the last equality by A^{-1} on the left, we obtain $A^{-1} = B(AB)^{-1}$.
- Multiplying this by B^{-1} on the left, we obtain $B^{-1}A^{-1} = (AB)^{-1}$, which is what we wanted to prove. ☐

10. $\det(A^{-1}) = \frac{1}{\det A}$. *See the proof of Theorem 1.*

11. The inverse of a diagonal matrix is easy to find:

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}^{-1} = \begin{pmatrix} 1/a_{11} & 0 & 0 & \cdots & 0 \\ 0 & 1/a_{22} & 0 & \cdots & 0 \\ 0 & 0 & 1/a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1/a_{nn} \end{pmatrix}.$$

Unfortunately, this is the only type of matrices for which finding inverse is this easy.

Elementary transformations

- To find the inverse A^{-1} of A, we shall use the elementary transformations again.
- Elementary transformations may be interpreted as multiplying A by some special matrices.
- The operation of swapping rows

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix} \rightsquigarrow \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

may be represented as multiplication by

$$E_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

on the left:

$$E_1 \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}.$$

- The operation of multiplying one row by λ

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix} \rightsquigarrow \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

may be represented as multiplication by

$$E_2 = \begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

<u>on the left</u>:

$$E_2 \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

.

- The operation of adding one line to another

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix} \rightsquigarrow \begin{pmatrix} a_{11} + \lambda a_{21} & a_{12} + \lambda a_{22} & \cdots & a_{1n} + \lambda a_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

may be represented as multiplication by

$$E_3 = \begin{pmatrix} 1 & \lambda & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

<u>on the left</u>:

$$E_3 \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix} = \begin{pmatrix} a_{11} + \lambda a_{21} & a_{12} + \lambda a_{22} & \cdots & a_{1n} + \lambda a_{2n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}.$$

- We have seen already that elementary transformations do not change rank and that the determinant has nice properties with respect to these operations. Now we can prove these properties easily:

  (i) $\det(E_1 A) = \det(E_1) \cdot \det(A) = -\det(A)$, because $\det(E_1) = -1$ (prove this!).

  (ii) $\det(E_2 A) = \det(E_2) \cdot \det(A) = \lambda \det(A)$, because $\det(E_2) = \lambda$ (prove this!).

  (iii) $\det(E_3 A) = \det(E_3) \cdot \det(A) = \det(A)$, because $\det(E_3) = 1$ (prove this!).

- Note that elementary transformations <u>of columns</u> are equivalent to multiplication by elementary matrices <u>on the right</u>, not on the left!

- Idea: to find A^{-1} we will multiply A on the left by elementary transformation matrices $F_1, F_2, F_3, \dots$ until
$$F_k F_{k-1} \dots F_3 F_2 F_1 A = I.$$
Then $A^{-1} = F_k F_{k-1} \dots F_3 F_2 F_1$.

- In other words, to find A^{-1} we need to apply row elementary transformations to both sides of the equation
$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$
For example, if
$$A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix},$$
then
$$\begin{aligned}
\begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad & \xrightarrow[L_1 \leftarrow L_1 + 2L_2]{} \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \\
& \xrightarrow[L_2 \leftarrow -L_2]{} \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \\
& \xrightarrow[L_2 \leftrightarrow L_1]{} \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \\
& \xrightarrow[L_1 \leftarrow L_1 + 3L_2]{} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} \\
& \Longrightarrow A^{-1} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}
\end{aligned}$$

- To simplify the notation, we introduce a notation trick called the augmented matrix $(A \mid I)$ as follows:
$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right).$$
Applying elementary transformations to this augmented matrix is the same as setting up the equation $AA^{-1} = I$ and applying the transformations to both sides. For

example,

$$
\left(\begin{array}{cc|cc} 2 & -5 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{array}\right)
\xrightarrow[L_1 \leftarrow L_1 + 2L_2]{} \left(\begin{array}{cc|cc} 0 & 1 & 1 & 2 \\ -1 & 3 & 0 & 1 \end{array}\right)
$$

$$
\xrightarrow[L_2 \leftarrow -L_2]{} \left(\begin{array}{cc|cc} 0 & 1 & 1 & 2 \\ 1 & -3 & 0 & -1 \end{array}\right)
$$

$$
\xrightarrow[L_2 \leftrightarrow L_1]{} \left(\begin{array}{cc|cc} 1 & -3 & 0 & -1 \\ 0 & 1 & 1 & 2 \end{array}\right)
$$

$$
\xrightarrow[L_1 \leftarrow L_1 + 3L_2]{} \left(\begin{array}{cc|cc} 1 & 0 & 3 & 5 \\ 0 & 1 & 1 & 2 \end{array}\right)
$$

$$
\Longrightarrow A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}.
$$

Recipe for finding the inverse

In order to find a matrix inverse, we play the following game: given an augmented matrix $(A \mid I)$ of the form

$$
\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array}\right),
$$

bring it to the form $(I \mid B)$

$$
\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & 1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{array}\right).
$$

by applying row elementary transformations. Then $A^{-1} = B$.

Is it okay to use both row and column elementary transformations?

No. We need to choose one kind and stick to it.

- Recall that row transformations are equivalent to multiplying by elementary matrices on the left.
- Finding the inverse A^{-1} by row transformations is equivalent to solving $AX = I$ for X.
- Column transformations are equivalent to multiplying by elementary matrices on the right.

- Finding the inverse A^{-1} by column transformations is equivalent to solving $XA = I$ for X.
- Recall that it was okay to mix the transformations when we were finding determinant and rank!

Special case of 2×2 matrices

A matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_{2,2}(\mathbb{R})$ is invertible <u>if and only if</u>

$$\boxed{\det A = ad - bc \neq 0}$$

and its inverse is given by

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

Proof.

$$\begin{aligned} \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \frac{1}{ad-bc}\begin{pmatrix} d\cdot a - b\cdot c & d\cdot b - b\cdot d \\ -c\cdot a + a\cdot c & -c\cdot b + a\cdot d \end{pmatrix} \\ &= \frac{1}{ad-bc}\begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = I. \end{aligned}$$

□

Try finding this formula using elementary transformations!

Another method for matrix inversion: cofactor formula

Recall that the <u>cofactor</u> C_{ij} of the element a_{ij} is defined by

$$C_{ij} = (-1)^{i+j}\det \begin{array}{c} \begin{pmatrix} a_{11} & a_{12} & & & \\ a_{21} & a_{22} & & & \\ & & \ddots & & \\ & & & a_{ij} & \\ & & & & \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i. \\ \\ \end{array} \\ \begin{array}{c} \uparrow \\ j \end{array} \end{array}$$

Definition 2. *The <u>cofactor matrix</u> of A is the matix $\operatorname{cof}(A)$ whose elements are cofactors of A's elements:*

$$(\operatorname{cof}(A))_{ij} = C_{ij}.$$

The <u>adjoint</u> matrix of A is the matrix $\operatorname{adj}(A)$ defined as the transpose of the cofactor matrix:

$$\operatorname{adj}(A) = (\operatorname{cof}(A))^{\top}.$$

The following theorem (without proof) gives another approach to finding the inverse of A:

Theorem 2. *If A is invertible, then* $A^{-1} = \dfrac{1}{\det A}\,\mathrm{adj}(A)\,.$