

7.2 Matrices and linear systems

Note: 1.5 lectures, first part of §5.1 in [EP], §7.2 and §7.3 in [BD], see also [\appendixref{linalg:appendix}](#)

7.2.1 Matrices and vectors

Before we start talking about linear systems of ODEs, we need to talk about matrices, so let us review these briefly. A *matrix* is an $m \times n$ array of numbers (m rows and n columns). For example, we denote a 3×5 matrix as follows

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{bmatrix}.$$

The numbers a_{ij} are called *elements* or *entries*.

By a *vector* we usually mean a *column vector*, that is an $m \times 1$ matrix. If we mean a *row vector*, we will explicitly say so (a row vector is a $1 \times n$ matrix). We usually denote matrices by upper case letters and vectors by lower case letters with an arrow such as \vec{x} or \vec{b} . By $\vec{0}$ we mean the vector of all zeros.

We define some operations on matrices. We want 1×1 matrices to really act like numbers, so our operations have to be compatible with this viewpoint.

First, we can multiply a matrix by a *scalar* (a number). We simply multiply each entry in the matrix by the scalar. For example,

$$2 \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix}.$$

Matrix addition is also easy. We add matrices element by element. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 2 \\ 4 & 7 & 10 \end{bmatrix}.$$

If the sizes do not match, then addition is not defined.

If we denote by 0 the matrix with all zero entries, by c, d scalars, and by A, B, C matrices, we have the following familiar rules:

$$\begin{aligned}
 A + 0 &= A = 0 + A, \\
 A + B &= B + A, \\
 (A + B) + C &= A + (B + C), \\
 c(A + B) &= cA + cB, \\
 (c + d)A &= cA + dA.
 \end{aligned}$$

- Another useful operation for matrices is the so-called *transpose*. This operation just swaps rows and columns of a matrix. The transpose of A is denoted by A^T . Example:

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

7.2.2 Matrix multiplication

- Let us now define matrix multiplication. First we define the so-called *dot product* (or *inner product*) of two vectors. Usually this will be a row vector multiplied with a column vector of the same size. For the dot product we multiply each pair of entries from the first and the second vector and we sum these products. The result is a single number. For example,

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1b_1 + a_2b_2 + a_3b_3.$$

- Similarly for larger (or smaller) vectors.

- Armed with the dot product we define the *product of matrices*. First let us denote by $\text{row}_i(A)$ the i^{th} row of A and by $\text{column}_j(A)$ the j^{th} column of A . For an $m \times n$ matrix A and an $n \times p$ matrix B , we can define the product AB . We let AB be an $m \times p$ matrix whose ij^{th} entry is the dot product

$$\text{row}_i(A) \cdot \text{column}_j(B).$$

- Do note how the sizes match up: $m \times n$ multiplied by $n \times p$ is $m \times p$. Example:

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

$$= \begin{bmatrix} 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 & 1 \cdot 0 + 2 \cdot 1 + 3 \cdot 0 & 1 \cdot (-1) + 2 \cdot 1 + 3 \cdot 0 \\ 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1 & 4 \cdot 0 + 5 \cdot 1 + 6 \cdot 0 & 4 \cdot (-1) + 5 \cdot 1 + 6 \cdot 0 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 \\ 15 & 5 & 1 \end{bmatrix}$$

- For multiplication we want an analogue of a 1. This analogue is the so-called *identity matrix*. The identity matrix is a square matrix with 1s on the diagonal and zeros everywhere else. It is usually denoted by I . For each size we have a different identity matrix and so sometimes we may denote the size as a subscript. For example, the I_3 would be the 3×3 identity matrix

$$I = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

- We have the following rules for matrix multiplication. Suppose that A, B, C are matrices of the correct sizes so that the following make sense. Let α denote a scalar (number).

$$\begin{aligned} A(BC) &= (AB)C, \\ A(B + C) &= AB + AC, \\ (B + C)A &= BA + CA, \\ \alpha(AB) &= (\alpha A)B = A(\alpha B), \\ IA &= A = AI. \end{aligned}$$

- A few warnings are in order.

- i. $AB \neq BA$ in general (it may be true by fluke sometimes). That is, matrices do not commute. For example, take $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.
- ii. $AB = AC$ does not necessarily imply $B = C$, even if A is not 0.
- iii. $AB = 0$ does not necessarily mean that $A = 0$ or $B = 0$. Try, for example, $A = B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

- For the last two items to hold we would need to “divide” by a matrix. This is where the *matrix inverse* comes in. Suppose that A and B are $n \times n$ matrices such that

$$AB = I = BA.$$

- Then we call B the inverse of A and we denote B by A^{-1} . If the inverse of A

exists, then we call A *invertible*. If A is not invertible, we sometimes say A is *singular*.

- ✎ If A is invertible, then $AB = AC$ does imply that $B = C$ (in particular, the inverse of A is unique). We just multiply both sides by A^{-1} (on the left) to get $A^{-1}AB = A^{-1}AC$ or $IB = IC$ or $B = C$. It is also not hard to see that $(A^{-1})^{-1} = A$.

✎ 7.2.3 The determinant

- ✎ For square matrices we define a useful quantity called the *determinant*. We define the determinant of a 1×1 matrix as the value of its only entry. For a 2×2 matrix we define

$$\det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) \stackrel{\text{def}}{=} ad - bc.$$

- ✎ Before trying to define the determinant for larger matrices, let us note the meaning of the determinant. Consider an $n \times n$ matrix as a mapping of the n -dimensional euclidean space \mathbb{R}^n to itself, where \vec{x} gets sent to $A\vec{x}$. In particular, a 2×2 matrix A is a mapping of the plane to itself. The determinant of A is the factor by which the area of objects changes. If we take the unit square (square of side 1) in the plane, then A takes the square to a parallelogram of area $|\det(A)|$. The sign of $\det(A)$ denotes changing of orientation (negative if the axes get flipped). For example, let

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

- ✎ Then $\det(A) = 1 + 1 = 2$. Let us see where the (unit) square with vertices $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$ gets sent. Clearly $(0, 0)$ gets sent to $(0, 0)$.

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

- ✎ The image of the square is another square with vertices $(0, 0)$, $(1, -1)$, $(1, 1)$, and $(2, 0)$. The image square has a side of length $\sqrt{2}$ and is therefore of area 2.
- ✎ If you think back to high school geometry, you may have seen a formula for computing the area of a parallelogram with vertices $(0, 0)$, (a, c) , (b, d) and $(a + b, c + d)$. And it is precisely

$$\det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) .$$

🔗 The vertical lines above mean absolute value. The matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ carries the unit square to the given parallelogram.

🔗 Let us look at the determinant for larger matrices. We define A_{ij} as the matrix A with the i^{th} row and the j^{th} column deleted. To compute the determinant of a matrix, pick one row, say the i^{th} row and compute:

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

🔗 For the first row we get

$$\det(A) = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + a_{13} \det(A_{13}) - \dots \begin{cases} +a_{1n} \det(A_{1n}) & \text{if } n \text{ is odd} \\ -a_{1n} \det(A_{1n}) & \text{if } n \text{ is even} \end{cases}$$

🔗 We alternately add and subtract the determinants of the submatrices A_{ij} multiplied by a_{ij} for a fixed i and all j . For a 3×3 matrix, picking the first row, we get $\det(A) = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + a_{13} \det(A_{13})$. For example,

$$\begin{aligned} \det \left(\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \right) &= 1 \cdot \det \left(\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} \right) - 2 \cdot \det \left(\begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} \right) + 3 \cdot \det \left(\begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} \right) \\ &= 1(5 \cdot 9 - 6 \cdot 8) - 2(4 \cdot 9 - 6 \cdot 7) + 3(4 \cdot 8 - 5 \cdot 7) = 0. \end{aligned}$$

🔗 The numbers $(-1)^{i+j} \det(A_{ij})$ are called *cofactors* of the matrix and this way of computing the determinant is called the *cofactor expansion*. No matter which row you pick, you always get the same number. It is also possible to compute the determinant by expanding along columns (picking a column instead of a row above). It is true that $\det(A) = \det(A^T)$.

🔗 A common notation for the determinant is a pair of vertical lines:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right).$$

🔗 I personally find this notation confusing as vertical lines usually mean a positive quantity, while determinants can be negative. Also think about how to write the absolute value of a determinant. I will not use this notation in this book.

- Think of the determinants telling you the scaling of a mapping. If B doubles the sizes of geometric objects and A triples them, then AB (which applies B to an object and then A) should make size go up by a factor of 6. This is true in general:

$$\det(AB) = \det(A) \det(B).$$

- This property is one of the most useful, and it is employed often to actually compute determinants. A particularly interesting consequence is to note what it means for existence of inverses. Take A and B to be inverses of each other, that is $AB = I$. Then

$$\det(A) \det(B) = \det(AB) = \det(I) = 1.$$

- Neither $\det(A)$ nor $\det(B)$ can be zero. Let us state this as a theorem as it will be very important in the context of this course.

Theorem 7.2.1. *An $n \times n$ matrix A is invertible if and only if $\det(A) \neq 0$.*

- In fact, $\det(A^{-1}) \det(A) = 1$ says that $\det(A^{-1}) = \frac{1}{\det(A)}$. So we even know what the determinant of A^{-1} is before we know how to compute A^{-1} .

- There is a simple formula for the inverse of a 2×2 matrix

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

- Notice the determinant of the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in the denominator of the fraction. The formula only works if the determinant is nonzero, otherwise we are dividing by zero.

7.2.4 Solving linear systems

- One application of matrices we will need is to solve systems of linear equations. This is best shown by example. Suppose that we have the following system of linear equations

$$\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 2, \\ x_1 + x_2 + 3x_3 &= 5, \\ x_1 + 4x_2 + x_3 &= 10. \end{aligned}$$

- Without changing the solution, we could swap equations in this system, we

could multiply any of the equations by a nonzero number, and we could add a multiple of one equation to another equation. It turns out these operations always suffice to find a solution.

- It is easier to write the system as a matrix equation. The system above can be written as

$$\begin{bmatrix} 2 & 2 & 2 \\ 1 & 1 & 3 \\ 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 10 \end{bmatrix}.$$

- To solve the system we put the coefficient matrix (the matrix on the left-hand side of the equation) together with the vector on the right and side and get the so-called *augmented matrix*

$$\left[\begin{array}{ccc|c} 2 & 2 & 2 & 2 \\ 1 & 1 & 3 & 5 \\ 1 & 4 & 1 & 10 \end{array} \right].$$

- We apply the following three elementary operations.

- Swap two rows.
- Multiply a row by a nonzero number.
- Add a multiple of one row to another row.

- We keep doing these operations until we get into a state where it is easy to read off the answer, or until we get into a contradiction indicating no solution, for example if we come up with an equation such as $0 = 1$.

- Let us work through the example. First multiply the first row by $\frac{1}{2}$ to obtain

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 1 & 3 & 5 \\ 1 & 4 & 1 & 10 \end{array} \right].$$

- Now subtract the first row from the second and third row.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 4 \\ 0 & 3 & 0 & 9 \end{array} \right]$$

Multiply the last row by $\frac{1}{3}$ and the second row by $\frac{1}{2}$.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 3 \end{array} \right]$$

Swap rows 2 and 3.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array} \right]$$

Subtract the last row from the first, then subtract the second row from the first.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array} \right]$$

If we think about what equations this augmented matrix represents, we see that $x_1 = -4$, $x_2 = 3$, and $x_3 = 2$. We try this solution in the original system and, voilà, it works!

Exercise 7.2.1. Check that the solution above really solves the given equations.

We write this equation in matrix notation as

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

where A is the matrix $\begin{bmatrix} 2 & 2 & 2 \\ 1 & 1 & 3 \\ 1 & 4 & 1 \end{bmatrix}$ and \vec{b} is the vector $\begin{bmatrix} 2 \\ 5 \\ 10 \end{bmatrix}$. The solution can also be computed via the inverse,

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

It is possible that the solution is not unique, or that no solution exists. It is easy to tell if a solution does not exist. If during the row reduction you come up with a row where all the entries except the last one are zero (the last entry in a row corresponds to the right-hand side of the equation), then the system is *inconsistent* and has no solution. For example, for a system of 3 equations and 3 unknowns, if you find a row such as $[0 \ 0 \ 0 \mid 1]$ in the augmented matrix, you know the system is inconsistent. That row corresponds to $0 = 1$.

You generally try to use row operations until the following conditions are

satisfied. The first (from the left) nonzero entry in each row is called the *leading entry*.

- i. The leading entry in any row is strictly to the right of the leading entry of the row above.
- ii. Any zero rows are below all the nonzero rows.
- iii. All leading entries are 1.
- iv. All the entries above and below a leading entry are zero.

Such a matrix is said to be in *reduced row echelon form*. The variables corresponding to columns with no leading entries are said to be *free variables*. Free variables mean that we can pick those variables to be anything we want and then solve for the rest of the unknowns.

Example 7.2.1. The following augmented matrix is in reduced row echelon form.

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Suppose the variables are x_1 , x_2 , and x_3 . Then x_2 is the free variable, $x_1 = 3 - 2x_2$, and $x_3 = 1$.

On the other hand if during the row reduction process you come up with the matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 13 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 3 \end{array} \right],$$

there is no need to go further. The last row corresponds to the equation $0x_1 + 0x_2 + 0x_3 = 3$, which is preposterous. Hence, no solution exists.

7.2.5 Computing the inverse

If the matrix A is square and there exists a unique solution \vec{x} to $A\vec{x} = \vec{b}$ for any \vec{b} (there are no free variables), then A is invertible. Multiplying both sides by A^{-1} , you can see that $\vec{x} = A^{-1}\vec{b}$. So it is useful to compute the inverse if you want to solve the equation for many different right-hand sides \vec{b} .

We have a formula for the 2×2 inverse, but it is also not hard to compute inverses of larger matrices. While we will not have too much occasion to compute inverses for larger matrices than 2×2 by hand, let us touch on how to do it. Finding the inverse of A is actually just solving a bunch of linear equations. If we can solve $A\vec{x}_k = \vec{e}_k$ where \vec{e}_k is the vector with all zeros except a 1 at the k^{th} position, then the inverse is the matrix with the columns \vec{x}_k for $k = 1, 2, \dots, n$ (exercise: why?). Therefore, to find the inverse we write a larger $n \times 2n$ augmented matrix $[A \mid I]$, where I is the identity matrix. We then perform row reduction. The reduced row echelon form of $[A \mid I]$ will be of the form $[I \mid A^{-1}]$ if and only if A is invertible. We then just read off the inverse A^{-1} .

7.2.6 Exercises

Exercise 7.2.2. Solve $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \vec{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ by using matrix inverse.

Exercise 7.2.3. Compute determinant of $\begin{bmatrix} 9 & -2 & -6 \\ -8 & 3 & 6 \\ 10 & -2 & -6 \end{bmatrix}$.

Exercise 7.2.4. Compute determinant of $\begin{bmatrix} 1 & 2 & 3 & 1 \\ 4 & 0 & 5 & 0 \\ 6 & 0 & 7 & 0 \\ 8 & 0 & 10 & 1 \end{bmatrix}$. Hint: Expand along the proper row or column to make the calculations simpler.


Exercise 7.2.5. Compute inverse of $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.


Exercise 7.2.6. For which h is $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & h \end{bmatrix}$ not invertible? Is there only one such h ? Are there several? Infinitely many?

Exercise 7.2.7. For which h is $\begin{bmatrix} h & 1 & 1 \\ 0 & h & 0 \\ 1 & 1 & h \end{bmatrix}$ not invertible? Find all such h .

Exercise 7.2.8. Solve $\begin{bmatrix} 9 & -2 & -6 \\ -8 & 3 & 6 \\ 10 & -2 & -6 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.


Exercise 7.2.9. Solve $\begin{bmatrix} 5 & 3 & 7 \\ 8 & 4 & 4 \\ 6 & 3 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$.

 **Exercise 7.2.10.** Solve $\begin{bmatrix} 3 & 2 & 3 & 0 \\ 3 & 3 & 3 & 3 \\ 0 & 2 & 4 & 2 \\ 2 & 3 & 4 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 0 \\ 4 \\ 1 \end{bmatrix}$.


 **Exercise 7.2.11.** Find 3 nonzero 2×2 matrices A , B , and C such that $AB = AC$ but $B \neq C$.

 **Exercise 7.2.101.** Compute determinant of $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & -5 \\ 1 & -1 & 0 \end{bmatrix}$


► Answer.

 **Exercise 7.2.102.** Find t such that $\begin{bmatrix} 1 & t \\ -1 & 2 \end{bmatrix}$ is not invertible.

► Answer.

 **Exercise 7.2.103.** Solve $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \vec{x} = \begin{bmatrix} 10 \\ 20 \end{bmatrix}$.

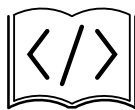
► Answer.

 **Exercise 7.2.104.** Suppose a, b, c are nonzero numbers. Let $M = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$,
 $N = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$.

a. Compute M^{-1} .

b. Compute N^{-1} .

► Answer.



For a higher quality printout use the PDF version:

<https://www.jirka.org/diffyqs/diffyqs.pdf>