ES.1803 Topic 13 Notes

Jeremy Orloff

13 Linear algebra: vector spaces, matrices and linearity

13.1 Goals

1. Know the definition of a vector space and how to show that a given set is a vector space.
2. Know the meaning of the phrase closed under addition and scalar multiplication.
3. Know how to convert a higher order DE into the companion system of first-order DEs.
4. Know how to organize matrix multiplication as a linear combination of the columns of the matrix.
5. Know how to organize matrix multiplication in block form and recognize when block multiplication is valid.

13.2 Introduction

Up to now we have spent most of our time in 18.03 considering linear differential equations. For these, one of our main tools was linearity, or, equivalently, the superposition principle. There are many other domains where linearity is important. For example, systems of linear algebraic equations and matrices. In this next unit on linear algebra we will study the common features of linear systems.

To do this we will introduce the somewhat abstract language of vector spaces. This will allow us to view the plane and space vectors you encountered in 18.02 and the general solutions to a differential equation through the same lens. In 18.02 vectors had both an algebraic and a geometric interpretation. In 18.03 we will focus primarily on the algebraic side of vectors, though we will sometimes use our geometric intuition as a guide.

13.3 Matlab (and alternatives)

We will use Matlab for computation and visualization. It will allow us to work with larger matrices where we wouldn't want to do computations by hand. We will only use a tiny subset of Matlab's enormous set of functions. I'll post some simple (and short) tutorials on its use.

Matlab is available for free to MIT students.

A free substitute for Matlab is Octave. It has the advantage that it loads much faster and doesn't spread digital rights management files all around your computer. The disadvantage is that it can be a little harder to install, especially on the Mac. Look at `https://www.gnu.org/software/octave/download.html`. I can help you get it installed if you want to try.

Another excellent and free substitute is Julia. The syntax is similar, but not identical, to Matlab. Downloads and documentation are available at https://julialang.org.

13.4 Linearity and vector spaces

We've seen before the importance of linearity when solving differential equations $P(D)x = f(t)$. To remind you: the operator $P(D)$ is linear means that

$$P(D)(c_1 f + c_2 g) = c_1 P(D) f + c_2 P(D) g$$

for all functions f, g and constants c_1, c_2.

Matrix multiplication is also linear. If A is a matrix and $\mathbf{v_1}$, $\mathbf{v_2}$ are vectors, then

$$A \cdot (c_1 \mathbf{v_1} + c_2 \mathbf{v_2}) = c_1 A \cdot \mathbf{v_1} + c_2 A \cdot \mathbf{v_2}$$

Example 13.1. In this example we will write an matrix multiplication in a way that emphasizes the linearity.

$$\begin{aligned}\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3+4 \\ 7+8 \end{bmatrix} &= \begin{bmatrix} 6(3+4) + 5(7+8) \\ (3+4) + 2(7+8) \end{bmatrix} \\ &= \begin{bmatrix} 6 \cdot 3 + 5 \cdot 7 + 6 \cdot 4 + 5 \cdot 8 \\ 1 \cdot 3 + 2 \cdot 7 + 1 \cdot 4 + 2 \cdot 8 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} + \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix}\end{aligned}$$

Linearity/Superposition
Exactly like solving linear differential equations, solving linear systems of algebraic equations involves finding a particular solution and superpositioning with the homogeneous solution.

Example 13.2. Solve $\begin{bmatrix} 1 & 3 \\ 4 & 12 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 6 \end{bmatrix}$

Solution: For this example we'll use ad hoc methods to find particular and homogeneous solutions. Later, we will learn systematic methods. The main point here is that the solutions can be superpositioned.

By inspection we can see one solution is $\mathbf{x_p} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. Just as valid would be to take $\mathbf{x_p} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ or $\mathbf{x_p} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$.

Next we have to solve the associated homogeneous equation:

$$\begin{bmatrix} 1 & 3 \\ 4 & 12 \\ 3 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This expands to three equations in two unknowns. You can easily check that the general solution is $\mathbf{x_h} = c \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

By superposition, the solution to the original equation is

$$\mathbf{x} = \mathbf{x_p} + \mathbf{x_h} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

If this is unclear, you should check the solution by substitution.

13.5 Vector spaces

The word space is used in mathematics to describe a set with extra properities. Math has all kinds of spaces. Here we will be concerned with vector spaces.

In order to have the notions of linearity and superposition, we need to have the notions of adding and scaling. This leads to the definition of vectors, whose key property is that they can be added and scaled.

Definition. A vector space is any set V with the following properties.

1. The set V has the arithmetic operations of addition and scalar multiplication.

2. Closure under addition: The sum of any two elements in V is another member V. That is, if $\mathbf{v}$, $\mathbf{w} \in V$, then $\mathbf{v} + \mathbf{w} \in V$.

3. Closure under scalar multiplication: Scaling an element of V results in another member of V. That is, if $\mathbf{v} \in V$ and c is a scalar, then $c\mathbf{v} \in V$.

4. Distributive law: If $\mathbf{v}$, $\mathbf{w} \in V$ and c is a scalar, then $c(\mathbf{v} + \mathbf{w}) = c\mathbf{v} + c\mathbf{w}$.

- An element $\mathbf{v}$ in V is called a vector.
- The formal definition of a vector space requires some more technical properties, but this definition will suffice for 18.03.
- If the scalars are required to be real numbers, we say we have a real vector space. If we allow them to be complex numbers, then we have a complex vector space.

The next few examples will introduce some important vector spaces and show how to check whether or not a given set is a vector space.

Key point. Checking whether or not a given set is a vector space is always easy. This is similar to checking whether a given operator is linear.

Example 13.3. Show that the set of ordered pairs (x, y), under the usual rules of addition and scalar multiplication, is a vector space.

Solution: We have to show the set satisfies the four properties in the definition of vector space. As we said, this is straightforward.

1. Multiplication and scalar multiplication: By definition we have these operations.

2. Closure under addition: Take any two ordered pairs (x_1, y_1) and (x_2, y_2). Their sum, $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ is also an ordered pair.

3. Closure under scalar multiplication: Take any ordered pair (x, y) and scalar c, then $c(x, y) = (cx, cy)$ is also an ordered pair.

4. Distributive law: We show this without any commentary:

$$c\left((x_1, y_1) + (x_2, y_2)\right) = c(x_1 + x_2, y_2 + y_2) = ... = c(x_1, y_1) + c(x_2, y_2).$$

Since the set satisfies the four properties, it is a vector space.

Example 13.4. (Vector spaces.) The following are all examples of vector spaces. You should be able to check this exactly as we did in the previous example.

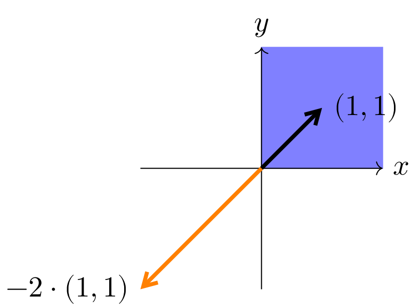
- We denote the plane by $\mathbf{R}^2$. It is the set of all ordered pairs (x, y).
- We denote space by $\mathbf{R}^3$. It is the set of all ordered triples (x, y, z).
- The powers indicate the dimension of each space. Likewise we can work with high dimensional vector spaces like $\mathbf{R}^{1000}$ which consists of all lists of 1000 numbers.
- In 18.03 we have used the fact that functions can be added and scaled. The set of solutions to the homogeneous DE

$$x'' + 8x' + 7x = 0$$

is a vector space. That is, the set $\{c_1 e^{-t} + c_2 e^{-7t}\}$ satisfies the above requirements 1-4 for a vector space.

Example 13.5. (Non-vector spaces.) The following are not vector spaces.

1. The set of plane vectors in the first quadrant. This fails to be closed under scalar multiplication. For example, $(1, 1)$ is in the first quadrant, but $-2 \cdot (1, 1) = (-2, -2)$ is not.



2. The set of functions of the form $\cos(6t) + c_1 e^{-t} + c_2 e^{-7t}$. This fails to be closed under addition. For example,

$$(\cos(6t) + 2e^{-t} + 3e^{-7t}) + (\cos(6t) + e^{-t} + 4e^{-7t}) = 2\cos(6t) + 3e^{-t} + 7e^{-7t}$$

The sum is not in the set because of the factor of 2 in front of $\cos(6t)$.

13.6 Connection to DEs

We will give two examples showing directly how matrices arise in differential equations.

Example 13.6. The companion matrix -converting a higher order DE to a first-order system. Here we are going to convert a higher order differential equation into a system of first-order equations. Later we will see how this technique allows us to understand DEs in a new way and also how it allows us to use numerical techniques on higher order equations.

Consider the second-order linear differential equation

$$\ddot{x} + 8\dot{x} + 7x = 0.$$

We've worked this example many times. The general solution is $x = c_1 e^{-t} + c_2 e^{-7t}$.

To convert the DE to a matrix system, we introduce a new variable: $y = \dot{x}$. Now, substituting y for $\dot{x}$ in the original DE we get the equation $\dot{y} + 8y + 7x = 0$. Altogether we have the system of two first-order linear DEs:

$$\begin{array}{rcrcr} \dot{x} & = & & & y \\ \dot{y} & = & -7x & - & 8y \end{array}$$

This can be rewritten in matrix form:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -7 & -8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Notice two things:

1. If we write this abstractly with $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $A = \begin{bmatrix} 0 & 1 \\ -7 & -8 \end{bmatrix}$, it looks like $\dot{\mathbf{x}} = A\mathbf{x}$. Ignoring the fact that $\mathbf{x}$ is a vector and A is a matrix, this looks like our most important DE: $\dot{x} = ax$.

2. Solving the system is equivalent to solving the original equation. That is, if we solve the original equation, we'll have found x and hence $y = \dot{x}$. Conversely, if we solve the matrix system, we'll have found x (the solution to the orginal DE) and $y = \dot{x}$.

In this case we already know the solution to the DE, so the solution to the system is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} c_1 e^{-t} + c_2 e^{-7t} \\ -c_1 e^{-t} + -7c_2 e^{-7t} \end{bmatrix} = c_1 e^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 e^{-7t} \begin{bmatrix} 1 \\ -7 \end{bmatrix}$$

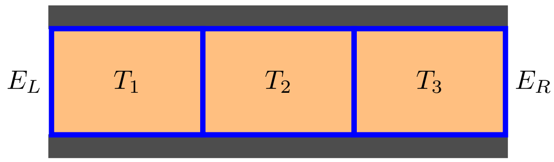
Notice that the basic solutions are of the form $e^{rt}\mathbf{v}$, where $\mathbf{v}$ is a constant vector. Later, we will use the method of optimism to guess solutions of this form.

Definition. The matrix A of coefficients that arises from this technique will be called the companion matrix to the original DE.

Example 13.7. Heat Flow. In this example we will set up a model for heat flow. We won't solve it for a few days.

Suppose we have a metal rod where different parts are at different temperatures. We divide it into 3 regions and imagine that each region exchanges heat with the adjacent regions. The regions on either end also exchange heat with the environment. We assume that the top and bottom of the rod are insulated, so that heat can only flow out of the bar at the ends. We also assume that the heat transfer follows Newton's law and the rate constant is k at each interface.

The figure below shows the metal bar divided into 3 regions and insulated above and below. The temperature of each region and the temperature of the environment on the left and right ends are indicated in the figure.



Using Newton's law we can write a DE for the temperature of each region.

$$\begin{array}{lclclclclclcl}
\dot{T}_1 & = & -k(T_1 - E_L) & - & k(T_1 - T_2) & = & -2kT_1 & + & kT_2 & + & & & kE_L \\
\dot{T}_2 & = & -k(T_2 - T_1) & - & k(T_2 - T_3) & = & kT_1 & - & 2kT_2 & + & kT_3 & & \\
\dot{T}_3 & = & -k(T_3 - T_2) & - & k(T_3 - E_R) & = & & & kT_2 & - & 2kT_3 & + & kE_R
\end{array}$$

We can write this in matrix form

$$\begin{bmatrix} \dot{T}_1 \\ \dot{T}_2 \\ \dot{T}_3 \end{bmatrix} = \begin{bmatrix} -2k & k & 0 \\ k & -2k & k \\ 0 & k & -2k \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} + \begin{bmatrix} kE_L \\ 0 \\ kE_R \end{bmatrix}$$

Remark: This particular coefficient matrix occurs quite often in applications. You should make sure you know how to modify the equation if we use n divisions of the rod instead of 3.

13.7 Matrix Multiplication

Here we will assume that you are comfortable with matrices and matrix multiplication. For completeness, we've added a quick review of some of the basics below.

Combination of columns

We can view the result of multiplying a matrix times a vector as a linear combination of the columns of the matrix. **We will use this again and again,** so you should internalize it now! We illustrate with an example:

Example 13.8. Consider the following product

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \cdot 3 + 5 \cdot 4 \\ 1 \cdot 3 + 2 \cdot 4 \end{bmatrix} = 3 \begin{bmatrix} 6 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Notice that the result is a linear combination of the columns of the matrix.

To express this abstractly we write a matrix as

$$A = \begin{bmatrix} \\ \mathbf{v_1} & \mathbf{v_2} & \mathbf{v_3} & \mathbf{v_4} & \mathbf{v_5} \\ \\ \end{bmatrix}$$

Here each $\mathbf{v_j}$ is a vector representing a column of A. We then have

$$\begin{bmatrix} \\ \mathbf{v_1} & \mathbf{v_2} & \mathbf{v_3} & \mathbf{v_4} & \mathbf{v_5} \\ \\ \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3} + c_4\mathbf{v_4} + c_5\mathbf{v_5}$$

That is, the product is a linear combination of the columns of A.

Block matrices and multiplication

Consider the following matrix $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 6 & 5 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix}$. We can divide this into blocks

$$A = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 6 & 5 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} 0 & I \\ \hline B & 0 \end{array}\right]$$

As long as the sizes of the blocks are compatible, block matrices multiply just like matrices:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} E \\ F \end{bmatrix} = \begin{bmatrix} AE + BF \\ CE + DF \end{bmatrix}$$

To convince yourself of this look at the following product and see that the blocks in the first column on the left only touch the top block on the right etc.

$$\left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 6 & 5 & 0 & 0 \\ 1 & 2 & 0 & 0 \end{array}\right] \cdot \left[\begin{array}{cc} a & b \\ c & d \\ \hline e & f \\ g & h \end{array}\right]$$

13.7.1 Review: matrix notation

For a matrix A, we give its size as rows $\times$ columns. So a 2×3 matrix has 2 rows and 3 columns. We write $A_{i,j}$ for the entry in the i^{th} row and j^{th} column.

Example 13.9. For the 2×3 matrix $A = \begin{bmatrix} 1 & 3 & 5 \\ 7 & 9 & 11 \end{bmatrix}$, the $1, 2$ entry is $A_{1,2} = 3$. Likewise, the $2, 3$ entry is $A_{2,3} = 11$.

13.7.2 Review: matrix multiplication

Written out formally the i, j-entry of AB is given by the dot product of the i^{th} row of A dotted with the j^{th} column of B. That is

$$i, j\text{-entry of } AB = \langle i^{\text{th}} \text{ row of } A \rangle \cdot \langle j^{\text{th}} \text{ column of } B \rangle$$

This is illustrated in the following example.

Example 13.10. In the matrix product below, we've put a line through the 3[rd] row of first matrix and the 2[nd] column of the second matrix. The dot product of this row and column is $3, 2$-entry of the product, in this case it's 51.

$$\begin{bmatrix} 3 & 4 \\ 5 & 6 \\ \not{7} & \not{8} \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 & |5 \\ 1 & |2 \end{bmatrix} = \begin{bmatrix} * & * \\ * & * \\ * & 51 \\ * & * \end{bmatrix}$$

Example 13.11.

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \cdot 3 + 5 \cdot 4 \\ 1 \cdot 3 + 2 \cdot 4 \end{bmatrix} = \begin{bmatrix} 38 \\ 11 \end{bmatrix}.$$

Only compatibly sized matrices can be multiplied. For matrices A and B: the product AB only makes sense if the number of columns of A equals the number of rows of B.

That is, the product AB only makes sense if the A is an $m \times n$ matrix and B is an $n \times p$ matrix. The product AB is an $m \times p$ matrix.

Example 13.12. (i) A 4×2 times a 2×3 gives a 4×3 matrix:

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \\ 7 & 8 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 & 5 & 3 \\ 1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 41 & 40 & 38 \\ 8 & 9 & 11 \\ 50 & 51 & 53 \\ 1 & 2 & 4 \end{bmatrix}$$

(ii) A 2×2 times a 2×3 gives a 2×3 matrix:

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 16 & 38 & 60 \\ 5 & 11 & 17 \end{bmatrix}$$

(iii) A 2×3 times a 3×1 gives a 2×1 matrix:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 50 \\ 122 \end{bmatrix}$$

(iv) A 2×2 times a 3×2 is **not okay**.

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}$$ **NOT A VALID EXPRESSION**

Matrix multiplication is **NOT commutative**. That is, except in rare cases, $AB \neq BA$. Indeed, sometimes the matrices are only compatible for one order of multiplication.

Example 13.13. Let $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ then the product AB is legitimate, but the product BA does not make sense.

Even, if the product is legitimate in either order, the products can be different.

Example 13.14. Let $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ then the following multiplications show that $AB \neq BA$

$$\begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 27 \\ 1 & 8 \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 9 \\ 3 & 6 \end{bmatrix}$$

Lesson: You need to be careful and precise when doing matrix algebra. Make sure you multiply in the correct order.

2. Identity: The following matrices are called **identity** matrices:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

They are called the identity for the same reason the scalar 1 is called the multiplicitave identity. That is if you multiply the identity times anything you get back that anything. For example

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$$

Identity matrices are always **square** matrices. That is they have the same number of rows and columns. We use the subscripts in I_2, I_3 etc. to indicate the size of the identity. If the size is clear from the context we drop the subscript and just write I for the identity matrix.

MIT OpenCourseWare
https://ocw.mit.edu

ES.1803 Differential Equations
Spring 2024

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.