LS. LINEAR SYSTEMS

LS.1 Review of Linear Algebra

In these notes, we will investigate a way of handling a linear system of ODE's directly, instead of using elimination to reduce it to a single higher-order equation. This gives important new insights into such systems, and it is usually a more convenient and faster way of solving them.

The method makes use of some elementary ideas about linear algebra and matrices, which we will assume you know from your work in multivariable calculus. Your textbook contains a section (5.3) reviewing most of these facts, with numerical examples. Another source is the 18.02 Supplementary Notes, which contains a beginning section on linear algebra covering approximately the right material.

For your convenience, what you need to know is summarized briefly in this section. Consult the above references for more details and for numerical examples.

1. Vectors.

A **vector** (or *n-vector*) is an n-tuple of numbers; they are usually real numbers, but we will sometimes allow them to be complex numbers, and all the rules and operations below apply just as well to n-tuples of complex numbers. (In the context of vectors, a single real or complex number, i.e., a constant, is called a **scalar**.)

The n-tuple can be written horizontally as a **row vector** or vertically as a **column vector**. In these notes it will almost always be a column. To save space, we will sometimes write the column vector as shown below; the small T stands for **transpose**, and means: change the row to a column.

$\mathbf{a} = (a_1, \ldots, a_n)$ *row vector* $\qquad$ $\mathbf{a} = (a_1, \ldots, a_n)^T$ *column vector*

These notes use boldface for vectors; in handwriting, place an arrow $\vec{a}$ over the letter.

Vector operations. The three standard operations on n-vectors are:

addition: $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$.

multiplication by a scalar: $c\,(a_1, \ldots, a_n) = (ca_1, \ldots, ca_n)$

scalar product: $(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = a_1 b_1 + \ldots + a_n b_n$.

2. Matrices and Determinants.

An $m \times n$ **matrix** A is a rectangular array of numbers (real or complex) having m rows and n columns. The element in the i-th row and j-th column is called the *ij-th entry* and written a_{ij}. The matrix itself is sometimes written (a_{ij}), i.e., by giving its generic entry, inside the matrix parentheses.

A $1 \times n$ matrix is a row vector; an $n \times 1$ matrix is a column vector.

Matrix operations. These are

addition: if A and B are both $m \times n$ matrices, they are added by adding the corresponding entries; i.e., if $A = (a_{ij})$ and $B = (b_{ij})$, then $A + B = (a_{ij} + b_{ij})$.

multiplication by a scalar: to get cA, multiply every entry of A by the scalar c; i.e., if $A = (a_{ij})$, then $cA = (ca_{ij})$.

matrix multiplication: if A is an $m \times n$ matrix and B is an $n \times k$ matrix, their product AB is an $m \times k$ matrix, defined by using the scalar product operation:

$$ij\text{-}th\ entry\ of\ AB = (i\text{-}th\ row\ of\ A) \cdot (j\text{-}th\ column\ of\ B)^T .$$

The definition makes sense since both vectors on the right are n-vectors. In what follows, the most important cases of matrix multiplication will be

(i) *A and B are square matrices of the same size*, i.e., both A and B are $n \times n$ matrices. In this case, multiplication is always possible, and the product AB is again an $n \times n$ matrix.

(ii) *A is an $n \times n$ matrix and $B = \mathbf{b}$, a column n-vector.* In this case, the matrix product $A\mathbf{b}$ is again a column n-vector.

Laws satisfied by the matrix operations.

For any matrices for which the products and sums below are defined, we have

$$(A\,B)\,C = A\,(B\,C) \qquad \textit{(associative law)}$$

$$A\,(B+C) = A\,B + A\,C, \quad (A+B)\,C = A\,B + A\,C \qquad \textit{(distributive laws)}$$

$$A\,B \neq B\,A \qquad \textit{(commutative law fails in general)}$$

Identity matrix. We denote by I_n the $n \times n$ matrix with 1's on the main diagonal (upper left to bottom right), and 0's elsewhere. If A is an arbitrary $n \times n$ matrix, it is easy to check from the definition of matrix multiplication that

$$A\,I_n = A \qquad \text{and} \qquad I_n A = A .$$

I_n is called the **identity matrix** of order n; the subscript n is often omitted.

Determinants. Associated with every *square* matrix A is a number, written $|A|$ or $\det A$, and called the **determinant** of A. For these notes, it will be enough if you can calculate the determinant of 2×2 and 3×3 matrices, by any method you like.

Theoretically, the determinant should not be confused with the matrix itself; the determinant is a *number*, the matrix is the *square array*. But everyone puts vertical lines on either side of the matrix to indicate its determinant, and then uses phrases like "the first row of the determinant", meaning the first row of the corresponding matrix.

An important formula which everyone uses and no one can prove is

$$\det(A\,B) = \det A \cdot \det B . \tag{1}$$

Inverse matrix. A square matrix A is called **nonsingular** or **invertible** if $\det A \neq 0$.

If A is nonsingular, there is a unique square matrix B of the same size, called the **inverse** to A, having the property that

$$B\,A = I, \qquad \text{and} \qquad A\,B = I .$$

This matrix B is denoted by A^{-1}. To confirm that a given matrix B is the inverse to A, you only have to check one of the above equations; the other is then automatically true.

Different ways of calculating A^{-1} are given in the references. However, if A is a 2×2 matrix, as it usually will be in the notes, it is easiest simply to use the formula for it:

(2) $$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad A^{-1} = \frac{1}{|A|}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} .$$

Remember this as a procedure, rather than as a formula: switch the entries on the main diagonal, change the sign of the other two entries, and divide every entry by the determinant. (Often it is better for subsequent calculations to leave the determinant factor outside, rather than to divide all the terms in the matrix by det$A$.) As an example of (2),

$$\begin{pmatrix} 1 & -2 \\ -1 & 4 \end{pmatrix}^{-1} = \tfrac{1}{2}\begin{pmatrix} 4 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} .$$

To calculate the inverse of a nonsingular 3×3 matrix, see for example the 18.02 notes.

3. Square systems of linear equations. Matrices and determinants were originally invented to handle in an efficient way the solution of a system of simultaneous linear equations. This is still one of their most important uses. We give a brief account of what you need to know. This is not in your textbook, but can be found in the 18.02 Notes. We will restrict ourselves to *square* systems — those having as many equations as they have variables (or "unknowns", as they are frequently called). Our notation will be:

$$\begin{aligned} A &= (a_{ij}), \quad \text{a square } n \times n \text{ matrix of constants,} \\ \mathbf{x} &= (x_1, \ldots, x_n)^T, \quad \text{a column vector of unknowns,} \\ \mathbf{b} &= (b_1, \ldots, b_n)^T, \quad \text{a column vector of constants;} \end{aligned}$$

then the square system

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n &= b_2 \\ \cdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

can be abbreviated by the matrix equation

(3) $$A\mathbf{x} = \mathbf{b} .$$

If $\mathbf{b} = \mathbf{0} = (0, \ldots, 0)^T$, the system (3) is called **homogeneous**; if this is not assumed, it is called **inhomogeneous**. The distinction between the two kinds of system is significant. There are two important theorems about solving square systems: an easy one about inhomogeneous systems, and a more subtle one about homogeneous systems.

Theorem about square inhomogeneous systems.

If A is nonsingular, the system (3) has a unique solution, given by

(4) $$\mathbf{x} = A^{-1}\mathbf{b} .$$

Proof. Suppose $\mathbf{x}$ represents a solution to (3). We have

$$\begin{aligned} A\mathbf{x} = \mathbf{b} \Rightarrow A^{-1}(A\mathbf{x}) &= A^{-1}\mathbf{b}, \\ \Rightarrow (A^{-1}A)\mathbf{x} &= A^{-1}\mathbf{b}, \quad \text{by associativity;} \\ \Rightarrow \quad I\mathbf{x} &= A^{-1}\mathbf{b}, \quad \text{definition of inverse;} \\ \Rightarrow \quad \mathbf{x} &= A^{-1}\mathbf{b}, \quad \text{definition of } I. \end{aligned}$$

This gives a formula for the solution, and therefore shows it is unique if it exists. It does exist, since it is easy to check that $A^{-1}\mathbf{b}$ is a solution to (3). □

The situation with respect to a homogeneous square system $A\mathbf{x} = \mathbf{0}$ is different. This always has the solution $\mathbf{x} = \mathbf{0}$, which we call the *trivial* solution; the question is: when does it have a nontrivial solution?

Theorem about square homogeneous systems. *Let A be a square matrix.*

(5) $\quad A\mathbf{x} = \mathbf{0}$ *has a nontrivial solution* $\Leftrightarrow$ $\det A = 0$ *(i.e., A is singular).*

Proof. The direction $\Rightarrow$ follows from (4), since if A is nonsingular, (4) tells us that $A\mathbf{x} = \mathbf{0}$ can have only the trivial solution $\mathbf{x} = \mathbf{0}$.

The direction $\Leftarrow$ follows from the criterion for linear independence below, which we are not going to prove. But in 18.03, you will always be able to show by calculation that the system has a nontrivial solution if A is singular.

4. Linear independence of vectors.

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, x_k$ be a set of n-vectors. We say they are **linearly dependent** (or simply, *dependent*) if there is a non-zero relation connecting them:

(6) $$c_1\mathbf{x}_1 + \ldots + c_k\mathbf{x}_k = \mathbf{0}, \qquad (c_i \text{ constants, not all } 0).$$

If there is no such relation, they are called **linearly independent** (or simply, *independent*). This is usually phrased in a positive way: the vectors are *linearly independent* if the only relation among them is the zero relation, i.e.,

(7) $$c_1\mathbf{x}_1 + \ldots + c_k\mathbf{x}_k = \mathbf{0} \quad \Rightarrow \quad c_i = 0 \quad \text{for all } i.$$

We will use this definition mostly for just two or three vectors, so it is useful to see what it says in these low-dimensional cases. For $k = 2$, it says

(8) $\quad \mathbf{x}_1$ *and* $\mathbf{x}_2$ *are dependent* $\Leftrightarrow$ *one is a constant multiple of the other.*

For if say $\mathbf{x}_2 = c\mathbf{x}_1$, then $c\mathbf{x}_1 - x_2 = \mathbf{0}$ is a non-zero relation; while conversely, if we have non-zero relation $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 = \mathbf{0}$, with say $c_2 \neq 0$, then $\mathbf{x}_2 = -(c_1/c_2)\,\mathbf{x}_1$.

By similar reasoning, one can show that

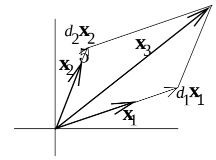
(9) $\quad \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ *are dependent* $\Leftrightarrow$ *one of them is a linear combination of the other two.*

Here by a **linear combination** of vectors we mean a sum of scalar multiples of them, i.e., an expression like that on the left side of (6). If we think of the three vectors as origin vectors in three space, the geometric interpretation of (9) is

(10) $\quad$ *three origin vectors in 3-space are dependent* $\Leftrightarrow$ *they lie in the same plane.*

For if they are dependent, say $\mathbf{x}_3 = d_1\mathbf{x}_1 + d_2\mathbf{x}_2$, then (thinking of them as origin vectors) the parallelogram law for vector addition shows that $\mathbf{x}_3$ lies in the plane of $\mathbf{x}_1$ and $\mathbf{x}_2$ — see the figure.

Conversely, the same figure shows that if the vectors lie in the same plane and say $\mathbf{x}_1$ and $\mathbf{x}_2$ span the plane (i.e., don't lie on a line), then by completing the



parallelogram, $\mathbf{x}_3$ can be expressed as a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$. (If they all lie on a line, they are scalar multiples of each other and therefore dependent.)

Linear independence and determinants. We can use (10) to see that

(11) $\quad$ *the rows of a* 3×3 *matrix* A *are dependent* $\Leftrightarrow$ *det* $A = \mathbf{0}$.

Proof. If we denote the rows by $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$, then from 18.02,

$$\begin{array}{c}\text{volume of the parallelepiped} \\ \text{spanned by } \mathbf{a}_1,\ \mathbf{a}_2,\ \mathbf{a}_3\end{array} = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = \det A,$$

so that

$$\mathbf{a}_1,\ \mathbf{a}_2,\ \mathbf{a}_3 \quad \text{lie in a plane} \quad \Leftrightarrow \quad \det A = 0.$$

The above statement (11) generalizes to an $n \times n$ matrix A ; we rephrase it in the statement below by changing both sides to their negatives. (We will not prove it, however.)

Determinantal criterion for linear independence

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *be* n*-vectors, and* A *the square matrix having these vectors for its rows (or columns). Then*

(12) $\quad \mathbf{a}_1, \ldots, \mathbf{a}_n$ *are linearly independent* $\Leftrightarrow$ *det*$A \neq 0$.

Remark. The theorem on square homogeneous systems (5) follows from the criterion (12), for if we let $\mathbf{x}$ be the column vector of n variables, and A the matrix whose columns are $\mathbf{a}_1, \ldots, \mathbf{a}_n$, then

(13) $$A\mathbf{x} = (\mathbf{a}_1 \ldots \mathbf{a}_n)\begin{pmatrix} x_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} = \mathbf{a}_1 x_1 + \ldots + \mathbf{a}_n x_n$$

and therefore

$A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$

$\Leftrightarrow$ $\mathbf{a}_1 x_1 + \ldots + \mathbf{a}_n x_n = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$, by (13);

$\Leftrightarrow$ $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent, by (7);

$\Leftrightarrow$ $\det A \neq 0$, by the criterion (12).

Exercises: Section 4A