Matrix Algebra, Quick and Dirty

W. O. Williams

Fall 2001

1 Apology

This set of notes is not a course in Linear Algebra. It is not a course in Matrix Algebra. It is not even a systematic compendium of results from these areas. It is a casual list of results, intended to be as spare as possible, which can lead a person with no knowledge of either area to those ideas needed for this course (21-260). I refrain from exploring the ideas, or generalizing them. If you find your curiosity piqued or need fuller explanations, grab the nearest text on Linear Algebra or Matrix Algebra (the ones on reserve in the Library or the manuscripts listed on the web page would do). The theory is beautiful and complete and the reading will amply justify whatever time you choose to invest – but it just is not essential to our needs.

2 Linear Equations

We follow Gauss in starting from the theory of linear equations. A **system of linear equations** is something like

$$\left.\begin{aligned} 5x+4y-z &= 2 \\ 3x-y+z &= 8 \end{aligned}\right\} \tag{1}$$

It is linear because the unknowns appear only to the first power. The object of the game is to find all possible x, y, z which simultaneously satisfy all of the equations; this is called the **general solution** of the system of equations. For example (1) has the general solution

$$\left.\begin{aligned} x &= 2-3*\alpha \\ y &= -2+8*\alpha \\ z &= 17*\alpha \end{aligned}\right\} \tag{2}$$

where α is any number at all. In words, putting any value of α in (2) gives a solution of (1), and any set of numbers x, y, z which satisfy (1) can be obtained by proper choice of α in (2). You should choose a few α and verify that the corresponding x, y, z are solutions of the equations.

NB. The *form* of the general solution is not unique. For example, another form of the general solution for these equations is

$$\left.\begin{array}{rcl} x &=& 5/4 - 3/8 * \beta \\ y &=& \beta \\ z &=& 17/4 + 17/8 * \beta \end{array}\right\} \tag{3}$$

It will take you a few minutes of algebra to establish that the two sets of solutions are the same!

2.1 Systematic Solution (Gauss-Jordan)

First, we should look at possibilities.

One way to solve

$$\left.\begin{array}{rcl} 3x + 5y &=& 2 \\ x + y &=& 0 \end{array}\right\} \tag{4}$$

is to use the second equation to find $y = -x$ and then to use this in the first to get

$$-2x = 2 \text{ or } x = -1 \tag{5}$$

so that $x = -1, y = 1$ is the **unique solution**.

By looking carefully at

$$\left.\begin{array}{rcl} 3x + 5y &=& 2 \\ 6x + 10y &=& 6 \end{array}\right\} \tag{6}$$

we see that there can be no numbers x and y which satisfy both equations. **No solution exists.**

Finally, the second equation of the pair

$$\left.\begin{array}{rcl} 3x + 5y &=& 2 \\ 6x + 10y &=& 4 \end{array}\right\} \tag{7}$$

clearly is redundant; it adds no information to the first. Solving the first, we can say that the general solution is

$$\left.\begin{array}{rcl} x &=& 2/3 - 5/3 * \alpha \\ y &=& \alpha \end{array}\right\}. \tag{8}$$

(Here we have taken the point of view that since we can assign one variable arbitrarily and solve for the other, we will formally call the value of that variable (y) the number α. This is a formalism which will prove useful later.) Because of the arbitrariness, we say there are **many solutions**.

The point of doing these simple examples is that the conclusions are absolutely typical: each set of linear equations has either a unique solution, no solution at all, or many solutions, typified by a general solution with one or more arbitrary constants. We now will examine a systematic way of determining solutions, the Gauss-Jordan process. Along the way, the process allows us to see which of these three situations obtains. We proceed by examples: you can deduce the general algorithm.

To solve

$$\left.\begin{array}{rcl} x+2y-2z &=& 2 \\ 2x+2y-3z &=& 0 \\ x-2y+z &=& 0 \end{array}\right\} \tag{9}$$

we first subtract 2*equation 1 from equation 2 to get a new equation 2, and subtract 1*equation 1 from equation 3 to get a new equation 3. We have a revised set of equations – *which have exactly the same set of solutions as the original set*.

$$\left.\begin{array}{rcll} x+2y-2z &=& 2 & (\mathsf{stet}) \\ -2y+z &=& -4 & (\mathsf{eqn2}-2*\mathsf{eqn1}) \\ -4y+3z &=& -2 & (\mathsf{eqn3}-\mathsf{eqn1}) \end{array}\right\} \tag{10}$$

The object was to obtain this form, with x appearing only in the first equation. We next do this for the variable y, leaving it only in the second equation.

$$\left.\begin{array}{rcll} x+2y-2z &=& 2 & (\mathsf{stet}) \\ -2y+z &=& -4 & (\mathsf{stet}) \\ z &=& 6 & (\mathsf{eqn3}-2*\mathsf{eqn2}) \end{array}\right\} \tag{11}$$

So $z=6$. We use this in equation 2 to find

$$-2y+6=-4 \quad \text{or} \quad y=5, \tag{12}$$

and substitute both into equation 1 to find

$$x+10-12=2 \quad \text{or} \quad x=4. \tag{13}$$

Thus there is a unique solution, and we have it.

Note. This is the **Gauss-Jordan procedure**. The goal is to wind up with a set of equations in which the ith variable appears in no equation past the ith one. The only elaboration which you need apply is occasionally to rearrange the order in which the equations are written. The systematic way of substituting the answer from one equation into in the previous equations to enable their solution is called **back-substitution**.

Another example: pretend that the G-J procedure has given us

$$\left.\begin{array}{rcl} x+2y-z &=& 2 \\ y-z &=& 6 \\ 0 &=& 0 \end{array}\right\} \tag{14}$$

(which tells us that the original third equation was redundant: it was an algebraic combination of the two previous equations). Then we cannot solve uniquely for z; we let it be undetermined and solve equation 2 for y, as

$$y = 6 + z, \tag{15}$$

and substitute this into equation 1:

$$x + 2(6+z) - z = 2 \quad \text{or} \quad x = -10 - z. \tag{16}$$

There are many solutions, according to our choice of z. We dress this up by saying

$$\left.\begin{array}{rcl} x &=& -10-\alpha \\ y &=& 6+\alpha \\ z &=& \alpha \end{array}\right\} \alpha \text{ arbitrary.} \tag{17}$$

A final example: if G-J leads us to

$$\left.\begin{array}{rcl} x+2y-z &=& 2 \\ y-z &=& 6 \\ 0 &=& 5 \end{array}\right\} \tag{18}$$

then we see there can be no solution (you can easily find a candidate for the previous step which led to this). This means that the original set of equations was inconsistent, even though it might not have been obvious at that stage. G-J makes it obvious.

2.2 Matrix Notation

Doing the G-J procedure by hand is tedious. A way of making it easier (and making it easier to avoid errors) is to replace the equations by arrays of coefficients. In this notation, the first example of the last section becomes

$$\left.\begin{array}{rcl} x+2y-2z &=& 2 \\ 2x+2y-3z &=& 0 \\ x-2y+z &=& 0 \end{array}\right\} \iff \begin{bmatrix} 1 & 2 & -2 \\ 2 & 2 & -3 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \tag{19}$$

$$\implies \begin{bmatrix} 1 & 2 & -2 \\ 0 & -2 & 1 \\ 0 & -4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ -2 \end{bmatrix} \tag{20}$$

$$\implies \begin{bmatrix} 1 & 2 & -2 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix} \tag{21}$$

The array on the left we call the **coefficient matrix** and note that it has size 3×3 (3 down and 3 across). The array with the unknowns is 3×1, and is called a **3-vector**, and the right-hand-side is also a 3-vector. In these terms, G-J is a method of reducing the coefficient matrix to one to an equivalent one which is **upper-triangular**, as in (21)

Exercise: Use this notation to carry out G-J on each of the 2×2 systems which we started with.

In matrix form, the possible outcomes of a G-J reduction on a 3×3 system are

$$\begin{bmatrix} \bowtie & \bowtie & \bowtie \\ 0 & \bowtie & \bowtie \\ 0 & 0 & \bowtie \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \bowtie \\ \bowtie \\ \bowtie \end{bmatrix} \tag{22}$$

$$\begin{bmatrix} \bowtie & \bowtie & \bowtie \\ 0 & \bowtie & \bowtie \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \bowtie \\ \bowtie \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \bowtie & \bowtie & \bowtie \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \bowtie \\ 0 \\ 0 \end{bmatrix} \tag{23}$$

$$\begin{bmatrix} \bowtie & \bowtie & \bowtie \\ 0 & \bowtie & \bowtie \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \bowtie \\ \bowtie \\ \bowtie \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \bowtie & \bowtie & \bowtie \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \bowtie \\ a \\ b \end{bmatrix} \tag{24}$$

where one of a or b is not zero. These yield the three outcomes mentioned before: respectively, a unique solution, many solutions with one or two parameters, and no solution.

Remark: Our examples so far have involved only square coefficient matrices. The same process is used for any number of equations in any number of unknowns, that is , for non-square coefficient matrices.

2.3 Determinants

The case in which there is a unique answer for the system, as in equation (22), when we get exactly one solution solving n equations for n unknowns, is, in several senses, the most desirable. If this is the case, we say that the (original or at any stage) coefficient matrix is **invertible**. Sometimes it is useful to determine *a priori* when the coefficient matrix is invertible. The classical way of doing this is to use determinants (although it is arguable that it is easier to do the G-J reduction to see if it becomes upper-triangular).

An $n \times n$ matrix is invertible if and only if its determinant is not zero.

You probably have learned to find a determinant of a 2×2 matrix:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc. \tag{25}$$

A 3×3 can be computed by the **cofactor expansion**:

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix} = a \det \begin{bmatrix} e & f \\ h & j \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & j \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix}. \tag{26}$$

Look carefully to see how the 2×2 matrices are chosen and how the sign alternates. In fact, the expansion can use any **row** (set of entries across) or **column** (set of entries down) from the big matrix, as long as a sign convention is followed (start with plus or minus according as you are on the even or odd numbered row or column). Thus:

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix} = -b \det \begin{bmatrix} d & f \\ g & j \end{bmatrix} + e \det \begin{bmatrix} a & c \\ g & j \end{bmatrix} - h \det \begin{bmatrix} a & c \\ d & f \end{bmatrix}. \tag{27}$$

Larger matrices follow a similar rule, if you have the patience. For example, a 4×4 requires evaluation of four 3×3 determinants. In these cases, if you do it by hand, it may be easier to use the G-J reduction-to-upper-triangular test.

2.4 Homogeneous Equations

We end this section with a couple of useful results. Because it is most important to us, we consider only the $n \times n$ case. For homogeneous equations (right-hand side all zeros) you can immediately write down one solution: all of the variables equal zero. If the coefficient matrix is invertible, then this is the only solution.

But if the matrix is not invertible, there always are solutions (think about the end of the G-J process; the other possibilities do not obtain); moreover, if, for example, x_o, y_o, z_o is a solution, then so is $\gamma x_o, \gamma y_o, \gamma z_o$, for any γ.

3 Algebra of Vectors and Matrices

3.1 Multiplication

We have implicitly used multiplication of matrices and vectors above: think of

$$\begin{bmatrix} 1 & 2 & -2 \\ 2 & 2 & -3 \\ 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \tag{28}$$

or

$$\begin{bmatrix} 5 & 4 & -1 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \tag{29}$$

formally as

$$A\mathbf{x} = \mathbf{b} \tag{30}$$

so that the matrix A multiplies the vector of unknowns $\mathbf{x}$ to get the right-hand-side vector $\mathbf{b}$. You deduce the formal process of multiplication from looking at the equations these represent. The rule is

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + jz \end{bmatrix} \tag{31}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \end{bmatrix} \tag{32}$$

Remark: It is easy to extend to multiplication of matrices of various sizes. The rule is $m \times n$ times $n \times q$ yields $m \times q$. See our text for the straightforward extension. It is interesting that the multiplication of square matrices, which can occur in either order, generally is not the same if the order of multiplication is reversed: $AB \neq BA$.

One useful tool: the **identity matrix** I does not alter a vector:

$$I\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{x}. \tag{33}$$

3.2 Addition and Scalar Multiplication

When we solved

$$\begin{bmatrix} 5 & 4 & -1 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \tag{34}$$

we found the solutions (now put in vector format)

$$\mathbf{x} = \begin{bmatrix} 2 - 3 * \alpha \\ -2 + 8 * \alpha \\ 17 * \alpha \end{bmatrix} \tag{35}$$

Let us write this as

$$\mathbf{x} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -3 \\ 8 \\ 17 \end{bmatrix}, \tag{36}$$

introducing two straightforward ideas: adding vectors and multiplying them by a number (referred to as a **scalar**, to distinguish it from a vector).

These operations follow the standard rules of arithmetic. Note in particular that

$$-\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} -a \\ -b \\ -c \\ -d \end{bmatrix} = (-1) \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}. \tag{37}$$

For example

$$\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ -3 \\ -6 \end{bmatrix} \tag{38}$$

In the next section we will have occasion to deal with linear combinations of vectors. Given vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, we call a combination like

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} \tag{39}$$

a **linear combination** of the vectors. Here α, β, γ are scalars.

3.3 Linear Independence

Suppose we are given vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and we find it useful to write another vector as a linear combination of them, as

$$\mathbf{x} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}. \tag{40}$$

This is most convenient, presumably, if the number of vectors on the right is minimal. Suppose, say, we knew that

$$\mathbf{b} = -\mathbf{a} + 3\mathbf{c}. \tag{41}$$

Then we could replace (40) by

$$\begin{aligned} \mathbf{x} &= \alpha\mathbf{a} + \beta(-\mathbf{a} + 3\mathbf{c}) + \gamma\mathbf{c} & (42)\\ &= (\alpha - \beta)\mathbf{a} + (\gamma + 3\beta)\mathbf{c}, & (43) \end{aligned}$$

and the representation of $\mathbf{x}$ becomes more efficient, since only two vectors are involved.

Generally, we say that the set of vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is **linearly dependent** if we can find scalars ζ, η, θ, *not all zero*, so that

$$\zeta\mathbf{a} + \eta\mathbf{b} + \theta\mathbf{c} = \mathbf{0}. \tag{44}$$

Here $\mathbf{0}$ is the zero vector. If the set of vectors is linearly dependent, then one of the vectors can be written as a linear combination of the other two, and always can be eliminated from a representation like (40). If there is no such set of scalars, then the set of vectors is said to be **linearly independent** and representations like (40) are irreducible.

It is a trivial calculation to show that the standard basis for the 3-vectors,

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{45}$$

is linearly independent.

Testing for linear dependence is a knee-jerk operation: to find whether the set

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \tag{46}$$

is linearly dependent, set up the equation

$$\zeta \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \eta \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \theta \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{47}$$

to solve for ζ, η, θ. This is a system of linear equations:

$$\left.\begin{aligned} \eta + 2\theta &= 0 \\ \zeta + \theta &= 0 \\ \zeta + \eta + 3\theta &= 0 \end{aligned}\right\} \tag{48}$$

or

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} \zeta \\ \eta \\ \theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{49}$$

Now we can be clever: notice that the columns of this matrix are the three vectors, so we could have leapt directly to this point. Moreover, the system of equations is homogeneous, so that we have two options. We may choose to calculate the determinant of this matrix or we can choose to solve for the unknowns. If we chose the former, and find that the determinant is non-zero, the matrix is invertible and the only solution is the zero one: $\zeta = 0, \eta = 0, \theta = 0$, so the set of vectors is linearly independent. If the determinant is zero, then there must be non-zero solutions and we conclude that the set of vectors is linearly dependent.

Let's take the second method, solving the system (49). We rearrange the equations (why?) to

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} \zeta \\ \eta \\ \theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{50}$$

and reduce:

$$\Longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} \zeta \\ \eta \\ \theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{51}$$

$$\Longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \zeta \\ \eta \\ \theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{52}$$

Clearly this has non-zero solutions, $\zeta = -\theta, \eta = -2\theta$, so that the vectors are linearly dependent. For example, if $\theta = -1, \zeta = 1, \eta = 2$ and we find

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{53}$$

and we can write any one of the three vectors as a linear combination of the other two.

3.4 Homogeneous and Particular Solutions

This topic is very important for us, as it extends to linear differential equations as well.

We proceed by example: solving

$$\begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 2 & -1 & -1 & 1 \\ 2 & 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \\ 6 \end{bmatrix} \tag{54}$$

$$\Longrightarrow \begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & -2 & 0 & -2 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ -4 \\ 2 \end{bmatrix} \tag{55}$$

$$\Longrightarrow \begin{bmatrix} 2 & 1 & -1 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \\ 0 \end{bmatrix} \tag{56}$$

We find by back-substitution that

$$w = 1 + \frac{1}{2} y - z, \ x = 2 - z, \tag{57}$$

with y and z arbitrary. Thus the solution can be written as

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}. \tag{58}$$

This is more than a convenient formalism. One can verify (please do) that the first vector in (58) is a solution of the system (54). But we also find (verify this as well) that each of the other two vectors satisfies the homogeneous version (zero right-hand-side) of (54)! Moreover it is easy to see that these two vectors are linearly independent.

This is a general result. If the G-J procedure and back-substitution are carried out, the form of the solution is always a **particular solution** of the equation plus a linear combination of linearly independent solutions of the homogeneous equation (the **general solution of the homogeneous equation**). If the solution is unique, then the general solution of the homogeneous equation is the zero vector.

The arbitrariness in form of the general solution referred to earlier can be explained in these terms. The particular solution can be anything that satisfies the equation: here $\begin{bmatrix}1\\2\\0\\0\end{bmatrix}$ can be replaced by $\begin{bmatrix}0\\1\\0\\1\end{bmatrix}$, for example. Also the two (two in this case) linearly independent solutions of the homogeneous equation could be replaced by any two other linearly independent solutions – and there are many.

4 Geometry

So far, we have ignored vectors as geometric objects. Undoubtedly you are familiar with the interpretation of 2-vectors and 3-vectors as arrows. Using the standard basis vectors (45) as unit length arrows in the axis directions, we draw, or, more likely, imagine that we draw

$$\begin{bmatrix}3\\-1\\2\end{bmatrix} = 3\begin{bmatrix}1\\0\\0\end{bmatrix} - \begin{bmatrix}0\\1\\0\end{bmatrix} + 2\begin{bmatrix}0\\0\\1\end{bmatrix} \tag{59}$$

as an arrow of three units in the x-direction, one unit in the negative y-direction and two units in the z-direction.

The most important notions for us are orthogonality and parallelness. We say that two vectors are **orthogonal** if their **inner product** is zero:

$$\mathbf{a}\cdot\mathbf{b} = \begin{bmatrix}a\\b\\c\end{bmatrix}\cdot\begin{bmatrix}d\\e\\f\end{bmatrix} = ad + be + cf = 0. \tag{60}$$

This is the usual dot product you should be familiar with. We say that two vectors $\mathbf{c}, \mathbf{d}$ are **parallel** if one is a scalar multiple of the other

$$\mathbf{c} = \alpha\mathbf{d}, \tag{61}$$

ie, if they are linearly dependent.

5 Eigenvalue Problems

Sometimes it is useful to find whether a matrix A has the property that there is one or more vectors $\mathbf{x}$ with $A\mathbf{x}$ parallel to $\mathbf{x}$. For example we see that

$$\begin{bmatrix}1 & 2\\2 & 1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}3\\3\end{bmatrix} = 3\begin{bmatrix}1\\1\end{bmatrix}. \tag{62}$$

The vector, $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ here, is called an **eigenvector** of the matrix; the multiplier, 3 here, is called an **eigenvalue**.

Formally, given a matrix A, we seek a number λ and a vector $\mathbf{x}$ such that

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{63}$$

There is an elaborate theory of when a matrix has eigenvectors and how many. We will not look into that, but consider how to find them. We note that the equation can be rewritten

$$A\mathbf{x} - \lambda\mathbf{x} = A\mathbf{x} - \lambda I\mathbf{x} = \mathbf{0} \tag{64}$$

$$(A - \lambda I)\mathbf{x} = \mathbf{0}. \tag{65}$$

The last is a homogeneous equation. We can choose to solve it directly by G-J reduction, finding that we must make particular choices of λ so as to ensure there is a non-zero solution, or we can first ensure that there is a non-zero solution by choosing λ so as to make the determinant of the coefficient matrix zero and then proceed to find $\mathbf{x}$. It is traditional to do the latter, *ie*, to set

$$\det(A - \lambda I) = 0, \tag{66}$$

which yields an equation to solve for λ, then to substitute λ in (65) and solve for $\mathbf{x}$.

For the 2×2 matrix we just introduced equation (66) becomes

$$\det\left(\begin{bmatrix}1 & 2\\ 2 & 1\end{bmatrix} - \lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right) = 0 \tag{67}$$

$$\det\begin{bmatrix}1-\lambda & 2\\ 2 & 1-\lambda\end{bmatrix} = 0 \tag{68}$$

$$(1-\lambda)^2 - 4 = 0. \tag{69}$$

This is easy to solve: the solutions are $\lambda = 3$ and $\lambda = -1$. Thus there are two eigenvalues.

Taking $\lambda = -1$, and putting it into equation (65), we seek to solve

$$(A + I)\mathbf{x} = \mathbf{0} \quad \text{or} \quad \begin{bmatrix}2 & 2\\ 2 & 2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}, \tag{70}$$

which, of course, reduces to

$$\begin{bmatrix}2 & 2\\ 0 & 0\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}, \tag{71}$$

so all solutions are multiples of $\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]$.

If we do the same thing for $\lambda = 3$ we find that all eigenvectors are multiples of $\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$, which we have already confirmed is an eigenvector associated to this eigenvalue.

It is a useful and easily demonstrated result that if the matrix is **symmetric**, *ie*, the entries are the same reflected across the main diagonal, the upper-left to lower-right diagonal, then the eigenvalues are all real numbers. The case in which the eigenvalues are complex, however, is no problem: the computations to find eigenvalues and eigenvectors are the same. The eigenvectors, of course, may involve complex numbers in this case.

6 Some Exercises

6.1 Solving Linear Equations

1. Solve the following

   (a) $$\left.\begin{array}{rcl} x + 2y &=& 1 \\ x - y &=& 2 \end{array}\right\}$$

   (b) $$\left.\begin{array}{rcl} 2x - y &=& 1 \\ x + y &=& -1 \end{array}\right\}$$

   (c) $$\left.\begin{array}{rcl} x + y + z &=& 1 \\ x - y + z &=& -1 \\ 4x + 2y + z &=& 1 \end{array}\right\}$$

2. Solve the following

   (a) $$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 4 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

   (b) $$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

   (c) $$\begin{bmatrix} 1 & -2 & 1 \\ 2 & -1 & -1 \\ 3 & 0 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

3. For what values of a, b, c is there a solution to

$$\begin{bmatrix} 1 & -2 & 1 \\ 2 & -1 & -1 \\ 3 & 0 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

6.2 Algebra, Linear Dependence

1. Compute the following ($i = \sqrt{-1}$ is the complex unit.)

   (a) $\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

   (b) $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - 4 \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

   (c) $\begin{bmatrix} 2 \\ -2 \\ -3 \end{bmatrix} + i \begin{bmatrix} 1+i \\ -1 \\ i \end{bmatrix}$

2. Find $\mathbf{a}$ if

   (a) $\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \mathbf{a} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$

   (b) $\begin{bmatrix} 2 \\ 1 \end{bmatrix} - 2\mathbf{a} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$

   (c) $\begin{bmatrix} 2 \\ 1 \end{bmatrix} + (2+i)\mathbf{a} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$

3. Determine linear dependence or independence of

   (a) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

   (c) $\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$

   (d) $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$

6.3 Orthogonality

1. Find a vector orthogonal to both $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

2. Show that if the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are mutually orthogonal, then they are linearly independent.

6.4 Eigenvalue Problems

Solve for all eigenvalues and eigenvectors of the matrix.

1. $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

2. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & -1 & 3 \end{bmatrix}$