Cambridge University Press
978-1-107-03356-6 - Vectors, Pure and Applied: A General Introduction to Linear Algebra
T. W. Körner
Excerpt
More information

Part I

Familiar vector spaces

1
Gaussian elimination

1.1 Two hundred years of algebra

In this section we recapitulate two hundred or so years of mathematical thought.

Let us start with a familiar type of brain teaser.

Example 1.1.1 *Sally and Martin go to* The Olde Tea Shoppe. *Sally buys three cream buns and two bottles of pop for thirteen shillings, whilst Martin buys two cream buns and four bottles of pop for fourteen shillings. How much does a cream bun cost and how much does a bottle of pop cost?*

Solution. If Sally had bought six cream buns and four bottles of pop, then she would have bought twice as much and it would have cost her twenty six shillings. Similarly, if Martin had bought six cream buns and twelve bottles of pop, then he would have bought three times as much and it would have cost him forty two shillings. In this new situation, Sally and Martin would have bought the same number of cream buns, but Martin would have bought eight more bottles of pop than Sally. Since Martin would have paid sixteen shillings more, it follows that eight bottles of pop cost sixteen shillings and one bottle costs two shillings.

In our original problem, Sally bought three cream buns and two bottles of pop, which, we now know, must have cost her four shillings, for thirteen shillings. Thus her three cream buns cost nine shillings and each cream bun cost three shillings. □

As the reader well knows, the reasoning may be shortened by writing x for the cost of one bun and y for the cost of one bottle of pop. The information given may then be summarised in two equations

$$3x + 2y = 13$$
$$2x + 4y = 14.$$

In the solution just given, we multiplied the first equation by 2 and the second by 3 to obtain

$$6x + 4y = 26$$
$$6x + 12y = 42.$$

Subtracting the first equation from the second yields

$$8y = 16,$$

so $y = 2$ and substitution in either of the original equations gives $x = 3$.

We can shorten the working still further. Starting, as before, with

$$\begin{aligned} 3x + 2y &= 13 \\ 2x + 4y &= 14, \end{aligned}$$

we retain the first equation and replace the second equation by the result of subtracting $2/3$ times the first equation from the second to obtain

$$\begin{aligned} 3x + 2y &= 13 \\ \frac{8}{3}y &= \frac{16}{3}. \end{aligned}$$

The second equation yields $y = 2$ and substitution in the first equation gives $x = 3$.

It is clear that we can now solve any number of problems involving Sally and Martin buying sheep and goats or yaks and xylophones. The general problem involves solving

$$\begin{aligned} ax + by &= \alpha \\ cx + dy &= \beta. \end{aligned}$$

Provided that $a \neq 0$, we retain the first equation and replace the second equation by the result of subtracting c/a times the first equation from the second to obtain

$$\begin{aligned} ax + by &= \alpha \\ \left(d - \frac{cb}{a}\right) y &= \beta - \frac{c\alpha}{a}. \end{aligned}$$

Provided that $d - (cb)/a \neq 0$, we can compute y from the second equation and obtain x by substituting the known value of y in the first equation.

If $d - (cb)/a = 0$, then our equations become

$$\begin{aligned} ax + by &= \alpha \\ 0 &= \beta - \frac{c\alpha}{a}. \end{aligned}$$

There are two possibilities. Either $\beta - (c\alpha)/a \neq 0$, our second equation is inconsistent and the initial problem is insoluble, or $\beta - (c\alpha)/a = 0$, in which case the second equation says that $0 = 0$, and all we know is that

$$ax + by = \alpha$$

so, whatever value of y we choose, setting $x = (\alpha - by)/a$ will give us a possible solution.

There is a second way of looking at this case. If $d - (cb)/a = 0$, then our original equations were

$$\begin{aligned} ax + by &= \alpha \\ cx + \frac{cb}{a}y &= \beta, \end{aligned}$$

that is to say

$$\begin{aligned} ax + by &= \alpha \\ c(ax + by) &= a\beta \end{aligned}$$

so, unless $c\alpha = a\beta$, our equations are inconsistent and, if $c\alpha = a\beta$, the second equation gives no information which is not already in the first.

So far, we have not dealt with the case $a = 0$. If $b \neq 0$, we can interchange the roles of x and y. If $c \neq 0$, we can interchange the roles of the two equations. If $d \neq 0$, we can interchange the roles of x and y and the roles of the two equations. Thus we only have a problem if $a = b = c = d = 0$ and our equations take the simple form

$$\begin{aligned} 0 &= \alpha \\ 0 &= \beta. \end{aligned}$$

These equations are inconsistent unless $\alpha = \beta = 0$. If $\alpha = \beta = 0$, the equations impose no constraints on x and y which can take any value we want.

Now suppose that Sally, Betty and Martin buy cream buns, sausage rolls and bottles of pop. Our new problem requires us to find x, y and z when

$$\begin{aligned} ax + by + cz &= \alpha \\ dx + ey + fz &= \beta \\ gx + hy + kz &= \gamma. \end{aligned}$$

It is clear that we are rapidly running out of alphabet. A little thought suggests that it may be better to try and find x_1, x_2, x_3 when

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= y_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= y_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= y_3. \end{aligned}$$

Provided that $a_{11} \neq 0$, we can subtract a_{21}/a_{11} times the first equation from the second and a_{31}/a_{11} times the first equation from the third to obtain

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= y_1 \\ b_{22}x_2 + b_{23}x_3 &= z_2 \\ b_{32}x_2 + b_{33}x_3 &= z_3, \end{aligned}$$

Cambridge University Press
978-1-107-03356-6 - Vectors, Pure and Applied: A General Introduction to Linear Algebra
T. W. Körner
Excerpt
More information

where

$$b_{22} = a_{22} - \frac{a_{21}a_{12}}{a_{11}} = \frac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}}$$

and, similarly,

$$b_{23} = \frac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}}, \quad b_{32} = \frac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} \quad \text{and} \quad b_{33} = \frac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}}$$

whilst

$$z_2 = \frac{a_{11}y_2 - a_{21}y_1}{a_{11}} \quad \text{and} \quad z_3 = \frac{a_{11}y_3 - a_{31}y_1}{a_{11}}.$$

If we can solve the smaller system of equations

$$\begin{aligned} b_{22}x_2 + b_{23}x_3 &= z_2 \\ b_{32}x_2 + b_{33}x_3 &= z_3, \end{aligned}$$

then, knowing x_2 and x_3, we can use the equation

$$x_1 = \frac{y_1 - a_{12}x_2 - a_{13}x_3}{a_{11}}$$

to find x_1. In effect, we have reduced the problem of solving '3 linear equations in 3 unknowns' to the problem of solving '2 linear equations in 2 unknowns'. Since we know how to solve the smaller problem, we know how to solve the larger.

Exercise 1.1.2 *Use the method just suggested to solve the system*

$$\begin{aligned} x + y + z &= 1 \\ x + 2y + 3z &= 2 \\ x + 4y + 9z &= 6. \end{aligned}$$

So far, we have assumed that $a_{11} \neq 0$. A little thought shows that, if $a_{ij} \neq 0$ for some $1 \leq i, j \leq 3$, then all we need to do is reorder our equations so that the ith equation becomes the first equation and reorder our variables so that x_j becomes our first variable. We can then reduce the problem to one involving fewer variables as before.

If it is not true that $a_{ij} \neq 0$ for some $1 \leq i, j \leq 3$, then it must be true that $a_{ij} = 0$ for all $1 \leq i, j \leq 3$ and our equations take the peculiar form

$$\begin{aligned} 0 &= y_1 \\ 0 &= y_2 \\ 0 &= y_3. \end{aligned}$$

These equations are inconsistent unless $y_1 = y_2 = y_3 = 0$. If $y_1 = y_2 = y_3 = 0$, the equations impose no constraints on x_1, x_2 and x_3 which can take any value we want.

Cambridge University Press
978-1-107-03356-6 - Vectors, Pure and Applied: A General Introduction to Linear Algebra
T. W. Körner
Excerpt
More information

We can now write down the general problem when n people choose from a menu with n items. Our problem is to find $x_1, x_2, \dots, x_n$ when

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= y_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= y_n. \end{aligned}$$

We can condense our notation further by using the summation sign and writing our system of equations as

$$\sum_{j=1}^{n} a_{ij}x_j = y_i \qquad [1 \le i \le n]. \qquad \bigstar$$

We say that we have 'n linear equations in n unknowns' and talk about the '$n \times n$ problem'.

Using the insight obtained by reducing the 3×3 problem to the 2×2 case, we see at once how to reduce the $n \times n$ problem to the $(n-1) \times (n-1)$ problem. (We suppose that $n \ge 2$.)

Step 1. If $a_{ij} = 0$ for all i and j, then our equations have the form

$$0 = y_i \qquad [1 \le i \le n].$$

Our equations are inconsistent unless $y_1 = y_2 = \dots = y_n = 0$. If $y_1 = y_2 = \dots = y_n = 0$, the equations impose no constraints on $x_1, x_2, \dots, x_n$ which can take any value we want.

Step 2. If the condition of Step 1 does not hold, we can arrange, by reordering the equations and the unknowns, if necessary, that $a_{11} \neq 0$. We now subtract a_{i1}/a_{11} times the first equation from the ith equation $[2 \le i \le n]$ to obtain

$$\sum_{j=2}^{n} b_{ij}x_j = z_i \qquad [2 \le i \le n] \qquad \bigstar\bigstar$$

where

$$b_{ij} = \frac{a_{11}a_{ij} - a_{i1}a_{1j}}{a_{11}} \quad \text{and} \quad z_i = \frac{a_{11}y_i - a_{i1}y_1}{a_{11}}.$$

Step 3. If the new set of equations ★★ has no solution, then our old set ★ has no solution. If our new set of equations ★★ has a solution $x_i = x_i'$ for $2 \le i \le n$, then our old set ★ has the solution

$$\begin{aligned} x_1 &= \frac{1}{a_{11}}\left(y_1 - \sum_{j=2}^{n} a_{1j}x_j'\right) \\ x_i &= x_i' \qquad [2 \le i \le n]. \end{aligned}$$

Note that this means that if ★★ has exactly one solution, then ★ has exactly one solution, and if ★★ has infinitely many solutions, then ★ has infinitely many solutions. We have already remarked that if ★★ has no solutions, then ★ has no solutions.

Once we have reduced the problem of solving our $n \times n$ system to that of solving an $(n-1) \times (n-1)$ system, we can repeat the process and reduce the problem of solving the new $(n-1) \times (n-1)$ system to that of solving an $(n-2) \times (n-2)$ system and so on. After $n-1$ steps we will be faced with the problem of solving a 1×1 system, that is to say, solving a system of the form

$$ax = b.$$

If $a \neq 0$, then this equation has exactly one solution. If $a = 0$ and $b \neq 0$, the equation has no solution. If $a = 0$ and $b = 0$, every value of x is a solution and we have infinitely many solutions.

Putting the observations of the two previous paragraphs together, we get the following theorem.

Theorem 1.1.3 *The system of simultaneous linear equations*

$$\sum_{j=1}^{n} a_{ij}x_j = y_i \qquad [1 \leq i \leq n]$$

has 0, 1 *or infinitely many solutions.*

We shall see several different proofs of this result (for example, Theorem 1.4.5), but the proof given here, although long, is instructive.

1.2 Computational matters

The method just described for solving 'simultaneous linear equations' is called *Gaussian elimination*. Those who rate mathematical ideas by difficulty may find the attribution unworthy, but those who rate mathematical ideas by utility are happy to honour Gauss in this way.

In the previous section we showed how to solve $n \times n$ systems of equations, but it is clear that the same idea can be used to solve systems of m equations in n unknowns.

Exercise 1.2.1 *If $m, n \geq 2$, show how to reduce the problem of solving the system of equations*

$$\sum_{j=1}^{n} a_{ij}x_j = y_i \qquad [1 \leq i \leq m] \qquad ★$$

to the problem of solving a system of equations

$$\sum_{j=2}^{n} b_{ij}x_j = z_i \qquad [2 \leq i \leq m]. \qquad ★★$$

Exercise 1.2.2 *By using the ideas of Exercise 1.2.1, show that, if $m, n \geq 2$ and we are given a system of equations*

$$\sum_{j=1}^{n} a_{ij} x_j = y_i \qquad [1 \leq i \leq m], \qquad \bigstar$$

then at least one of the following must be true.

(*i*) ★ *has no solution.*

(*ii*) ★ *has infinitely many solutions.*

(*iii*) *There exists a system of equations*

$$\sum_{j=2}^{n} b_{ij} x_j = z_i \qquad [2 \leq i \leq m] \qquad \bigstar\bigstar$$

with the property that if ★★ *has exactly one solution, then* ★ *has exactly one solution, if* ★★ *has infinitely many solutions, then* ★ *has infinitely many solutions, and if* ★★ *has no solutions, then* ★ *has no solutions.*

If we repeat Exercise 1.2.1 several times, one of two things will eventually occur. If $n \geq m$, we will arrive at a system of $n - m + 1$ equations in one unknown. If $m > n$, we will arrive at 1 equation in $m - n + 1$ unknowns.

Exercise 1.2.3 (*i*) *If $r \geq 1$, show that the system of equations*

$$a_i x = y_i \qquad [1 \leq i \leq r]$$

has exactly one solution, has no solution or has an infinity of solutions. Explain when each case arises.

(*ii*) *If $s \geq 2$, show that the equation*

$$\sum_{j=1}^{s} a_j x_j = b$$

has no solution or has an infinity of solutions. Explain when each case arises.

Combining the results of Exercises 1.2.2 and 1.2.3, we obtain the following extension of Theorem 1.1.3

Theorem 1.2.4 *The system of equations*

$$\sum_{j=1}^{n} a_{ij} x_j = y_i \qquad [1 \leq i \leq m]$$

has 0, 1 *or infinitely many solutions. If $m > n$, then the system cannot have a unique solution* (*and so will have* 0 *or infinitely many solutions*).

Cambridge University Press
978-1-107-03356-6 - Vectors, Pure and Applied: A General Introduction to Linear Algebra
T. W. Körner
Excerpt
More information

Exercise 1.2.5 *Consider the system of equations*

$$x + y = 2$$
$$ax + by = 4$$
$$cx + dy = 8.$$

(*i*) *Write down non-zero values of a, b, c and d such that the system has no solution.*

(*ii*) *Write down non-zero values of a, b, c and d such that the system has exactly one solution.*

(*iii*) *Write down non-zero values of a, b, c and d such that the system has infinitely many solutions.*

Give reasons in each case.

Exercise 1.2.6 *Consider the system of equations*

$$x + y + z = 2$$
$$x + y + az = 4.$$

For which values of a does the system have no solutions? For which values of a does the system have infinitely many solutions? Give reasons.

How long does it take for a properly programmed computer to solve a system of n linear equations in n unknowns by Gaussian elimination? The exact time depends on the details of the program and the structure of the machine. However, we can get get a pretty good idea of the answer by counting up the number of elementary operations (that is to say, additions, subtractions, multiplications and divisions) involved.

When we reduce the $n \times n$ case to the $(n-1) \times (n-1)$ case, we subtract a multiple of the first row from the jth row and this requires *roughly* $2n$ operations. Since we do this for $j = 2, 3, \dots, n-1$ we need roughly $(n-1) \times (2n) \approx 2n^2$ operations. Similarly, reducing the $(n-1) \times (n-1)$ case to the $(n-2) \times (n-2)$ case requires about $2(n-1)^2$ operations and so on. Thus the reduction from the $n \times n$ case to the 1×1 case requires about

$$2\big(n^2 + (n-1)^2 + \cdots + 2^2\big)$$

operations.

Exercise 1.2.7 (*i*) *Show that there exist A and B with $A \geq B > 0$ such that*

$$An^3 \geq \sum_{r=1}^{n} r^2 \geq Bn^3.$$

(*ii*) (*Not necessary for our argument.*) *By comparing $\sum_{r=1}^{n} r^2$ and $\int_1^{n+1} x^2 dx$, or otherwise, show that*

$$\sum_{r=1}^{n} r^2 \approx \frac{n^3}{3}.$$