Linear Equations Matrices and Determinants

Introduction

The primary purpose of this topic is to develop the mathematical discipline that is necessary to solve a complex problem that requires many numerical and algebraic computations and which it is expected that one would make many errors, which have to be identified and corrected. This seems to be an impossible task unless one can effectively layout the problem and identify many patterns.

This problem will acquaint one with the approach that a mathematician would take to develop tools necessary to accomplish the task. One will also see how these tools can be used to develop a new mathematical topic.

It is easy for some one to solve a set of linear equation with two unknowns and a bit more difficult to solve one with three unknowns -- probably about 3 times harder. For four unknowns, it is 12 times and for five, 60 times more difficult – near impossible.

We begin by developing a new notation.

(a11 a12)	=	(1	2)	
(a21 a22)		(3	4)	

 a_{rc} represents a number in an array where subscript r is the row and c is the column. The value of the subscript is as follows: a11=1, a12=2, a21=3, and a22=4. We can multiply these matrices as follows:

(a11 a12)	2) $(b11 b12) = (a11b11+a12b21 a11b12+a21)$	b22)
(a21 a22)	2) x (b21 b22) (a21b11+a22b21 a21b12+a22	2b22)
To see this more clearly let loo	ok how each element is calculated:	
(a11 a12	2) $(b11 b12) = (a11b11 + a12b21 a11b12 + a21)$	1b22)
(a21 a22)	2) x (b21 b22) (a21b11+a22b21 a21b12+a22	2b22)
(a11 a12)	2) $(b11 b12) = (a11b11 + a12b21 a11b12 + a21)$	b22)
(a21 a22	2) x (b21 b22) (a21b11+a22b21 a21b12+a2	2b22)
(-11-12	A (111110) (-11111) -10101 -11110 -01	(1-22)
	2) (b11 b12) = (a11b11 + a12b21 a11b12 + a21)	
(a21 a22)	2) x (b21 b22) (a21b11+a22b21 a21b12+a22) (a21b11+a22b21 a21b12+a22) (a21b11+a22b21 a21b12+a22) (a21b12+a22) (a21b12+a22b21 a21b12+a22) (a21b12+a22) (a21a+a22)	2b22)
(all al2)	2) $(b11 b12) = (a11b11 + a12b21 a11b12 + a21)$	b22)
(a21 a22	2) x (b21 b22) (a21b11+a22b21 a21b12+a2	2b22)
Consider the notation		
2		

 $\Sigma a_{ij}b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} = a_{ij}b_{jk}$ (Einstein notation) j=1

You might notice that the outer subscripts of the product match the row/column indices of where the product is located:

(a11 a12) (b11 b12) = (a1jbj1 a1jbj2) (a21 a22) x (b21 b22) (a2jbj1 a2jbj2) (using Einstein notation) Matrixes satisfy the associative and commutative properties for addition and the associative properties of multiplication, but not the commutative properties. They also follow the distributive property. We can evaluate the matrix (determinat) as follows:

|a11 a12|

$$a21 a22| = a11a22-a21a12$$

Graphically, we are multiplying on the downward diagonal and subtracting the multiplication on the upper diagonal. To determine the sign of the product we use a concept called permutation. The following numbers are assigned permutation directions as follows:

Positive	Negative
123 312 231	321 213 132

We in essence rotate the forward sequence on backward sequence until we get the original positions. If both the row and column indices are in the same direction, the sign is positive. Consider a three dimensional determinant:

 $\begin{aligned} |a11 \ a12 \ a13| \\ |a21 \ a22 \ a23| &= a11a22a33 + a12a23a31 + a13a32a21 \\ |a31 \ a32 \ a33| & -(a22a13a31 + a11a23a32 + a33a21a12) \end{aligned}$

You will notice that each product has two of each index—two 1s, two 2s, and two threes. If we took allal2al3, it has 4=four ones and only one 2 and 3. It cannot be combined with other elements in its row or column. This suggest that we could rewrite the determinant as follows:

a11 x | a22 a23| - a12 x | a21 a23| + a13 x |a21 a22| |a32 a33| |a31 a33| |a31 a32|

To keep the signs correct we had to alternate signs in the decomposition. Note the pattern for say a12: |a11 a12 a13|

| a21 a22 a23| | a31 a32 a33|

We eliminate those elements in the row (1) and the column (2). You can see that the rules are built upon the observation of patterns.

What is the incentive for developing the concept and rules for matrices and determinants?

Simultaneous equations

Consider the following equations:

3x+5y=98 7x+2y=103 We are going to write them in a matrix notation: a11=3, a12=5 x1=x a21=7, a22=2, x2=y (3 5) (x1) = (98) (7 2) (x2) = (103)

We are now going to solve for x and y watching the algebraic and matrix approaches:

3x+5y=98 $(3 \ 5)(x) = (98)$ $|3 \ 5| = 3x2-7x5 = -29$ 7x + 2y = 103 $(7 \ 2)(y) = (103) |7 2|$ 7x 7*3x+7*5y=7x98 3 98 $3x \quad 3*7x+3x2y=3x103$ |7 103| 3x103-7x98 309-686 -377 (3*2-7*5)y=3x103-7x98y = (3*103-7*98)/(3*2-7*5)|3 5| $3x^2 - 7x^5$ 6 - 35 -29 |7 2 $2x \quad 2*3x+2*5y=2x98$ |98 5| 98x2-103x5 196-515 $5x \quad 5*7x+5*2y=5x103$ 103 2 -319 (2*3-5*7)x=2x98-5x103x = (2x108 - 5x103)/(3x2 - 5*7)3 5 3x2-7x5 6 – 35 -29 |7 2|

We have shown by using numbers how the terms of the determinants and how we solve the problem algebraically relate. You will note that we solve for each variable by replacing the contents of the column number for the variable of interest. In the case of two unknowns, the use of the matrix determinant notation does not make it easier to solve the problem.

Simultaneous equations of three unknowns

We are going to be more rigorous as we show a one to one correspondence between the Coefficients in the algebraic expressions as to the terms in the determinants.

(a33a11-a13a31)x1+(a33a12-a13a32)x2=a33b1-a13b3(a33a21-a23a31)x1+(a33a22-a23a32)x2=a33b2-a23b3 solution for x1 ((a33a22-a23a32)(a33a11-a13a31)-(a33a12-a13a32)(a33a21-a23a31))x1=(a33a22-a23a32)(a33b1-a13b3)-(a33a12-a13a32)(a33b2-a23b3)

Coefficients of x1 match determinant and constant term matches determinant after dividing by a33

a33(a33a22a11-a23a32a11-a13a31a22-a33a12a21+a12a23a31+a13a32a21)x1=a33((a22a33-a23a32)b1-(a33a12-a13a32)b2+(-a22a13+a12a23)b3)

```
| b1 a12 a13 |
| b2 a22 a23 | =(a22a33-a32a23)b1+(a13a32-a12a33)b2+(a12a23-a22a13)b3
| b3 a32 a33 |
```

Back to where x3 was eliminated (a33a11-a13a31)x1+(a33a12-a13a32)x2=a33b1-a13b3 (a33a21-a23a31)x1+(a33a22-a23a32)x2=a33b2-a23b3

solution for x2 ((a33a21-a23a31)(a33a12-a13a32)-(a33a11-a13a31)(a33a22-a23a32))x2= (a33a21-a23a31)(a33b1-a13b3)-(a33a11-a13a31)(a33b2-a23b3)

Coefficients of x2 match determinant and constant term matches determinant after dividing by a33

a33(a33a21a12-a23a31a12-a21a13a32-a33a11a22+a13a31a22+a11a23a32)x2=a33((a21a33-a23a31)b1+(a13a31-a11a33)b2+(-a21a13+a11a23)b3

multiply both sides by -1 to correct signs

| a11 b1 a13 | | a21 b2 a23 | =(a23a31-a21a33)b1+(a11a33-a31a13)b2+(a21a13-a23a11)b3 | a31 b3 a33 |

Let us check with numbers 2x+3y+4z=20|234| 5x+6y+7z=38D= | 5 6 7 | = 108+168+180-192-126-135=3 8x + 9y + 9z = 538991 2034 | 38 6 7 | / D = (1080+1113+1368-1272-1260-1026)/D=3/D=3/3=1 5399 |2204| | 5 38 7 | / D = (684+1120+1060-1216-742-900)/D=6/D=6/3=2 8539 |2320| | 2 | 1 20 | |5638|/D = 3x|5238|/D = (3x(212+304+300-320-228-265))/D=3x3/D=9/3=38353 8953

Therefore x=1, y=2, and z=3

Now we can see the power of organization and patterns in making it very easy to solve a set of equations with three unknowns.

Inverse matrices

The inverse matrix is defined such that when we multiply it times the original matrix we will get the unit matrix: (1 0)

Applying the inverse matrix to the numbers to the right of the equal signs, we can get the answer with fewer calculations if we exclude the calculations it took to determine the inverse matrix.

We will observe that to obtain the inverse matrix we interchange the top to bottom diagonal, change the signs of the variables on the diagonal from the bottom to top, and divide by the value of the determinant. In the next section we will prove this. Inverse for $2x^2$ $(35)^{-1} = (2-5) / |35| = (2-5) / -29$

(72) (-73) |72| (-73)

$$\frac{1}{-29} \times (2-5) (98) = (2 \times 98-5 \times 103) / -29 = (196-515) / -29 = (-319) / -29 = (11) (-7 - 3) (103) (-7 \times 98+3 \times 103) (-686+309) (377) = (13)$$

Inverse for the 3x3 matrix

Let us review some of the basic principles.

Expanding determinant

```
all al2 al3
|a21 a22 a23|
                 = a11 x | a22 a23 | - a12 x | a21 a23 | + a13 x | a21 a22 |
|a31 a32 a33|
                         | a32 a33|
                                            | a31 a32|
                                                               |a31 a32|
```

Matrix Multiplication

```
(c11 c12 c13) (a11 a12 a13) (c11a11+c12a21+c13a31 c11a12+c12a22+c13a32 c11a13+c12a23+c13a33)
(c21 c22 c23)x(a21 a22 a23)=(c21a11+c22a21+c23a31 c21a12+c22a22+c23a32 c21a13+c32a23+c13a32)
(c31 c32 c33) (a31 a32 a33) (c31a11+c32a21+c33a31 c31a12+c32a22+c33a32 c31a13+c32a23+c33a33)
```

 $d_{ij} = \sum_{k=1}^{\infty} c_{ik} a_{ki} = c_{ik} a_{ki}$ notice summation of center indices.

Permutations

sgn(a11a22a3) = + because both sequences in order sgn(a11a32a23))=- because right index sequential right and left is sequential left

123 231 312 are right sequential 321 213 132 are left sequential

(c11 c12 c13) (a11 a12 a13) (100) $(c21 c22 c23) \times (a21 a22 a23) = (010)$ the unit matrix (c31 c32 c33) (a31 a32 a33) (001)

Finding inverse c_{ii} let x=c11 y=c12 z=c13 where we look at matrix multiplications of the top row of the left matrix by the three columns of the right matrix.

a11 x + a21y + a31 z = 1a12 x + a22y + a32 z = 0a13 x + a23y + a33 z = 0

eliminate y then x:

a23a12 x + a23 a22 y + a23 a32 z = 0a22a13 x + a22 a23 y + a22 a33 z = 0(a23a12-a22a13)x+(a23a32-a22a33)z=0x = ((a22a33-a23a32)/(a12a23-a22a13)) z

```
a13a12 x + a13a22 y + a13a32 z = 0
a12a13 x + a12a23 y + a12a33 z = 0
           (a13a22-a12a23)y+(a13a32-a12a33)z=0
                           y= ((a13a32-a12a33/(a12a23-a13a22)z
Substituting:
z(a11(a22a33-a23a32)/(a12a23-a22a13)+a21(a13a32-a12a33)/(a12a23-a13a22)+a31) = 1
z=
 (a12a23-a22a13)/((a11(a22a33-a23a32)+a21(a13a32-a12a33)+a31(a12a23-a22a13)))
 z= | a12 a13| / D
                          all al2 al3
                      D = |a21 \ a22 \ a23|
     | a22 a23|
                         |a31 a32 a33|
y = - |a12 a13| /D
                         all al2 al3
    | a32 a33 |
                      D=|a21 a22 a23|
                         |a31 a32 a33|
x = |a22 \ a23|/D
                         |a11 a12 a13|
    |a32 a33|
                      D=|a21 a22 a23|
                         |a31 a32 a33|
```

Finding inverse c_{ji} let x=c21 y=c22 z=c23 where we look at matrix multiplications of the second row of the left matrix by the three columns of the right matrix.

(c11 c12 c13) (a11 a12 a13) (100)(c21 c22 c23) x (a21 a22 a23)= (010)(c31 c32 c33) (a31 a32 a33) (001)a11 x + a21 y a31 z = 0a12 x + a22y + a32 z = 1a13 x + a23y + a33 z = 0eliminate y then x:a23a11 x + a23 a21 y + a23 a31 z = 0a21a13 x + a21 a23 y + a21 a33 z = 0(a23a11-a21a13)x+(a23a31-a21a33)z=0x = -((a23a31-a21a33)/(a11a23-a21a13)) za13a11 x + a13a21 y + a13a31 z = 0a11a13 x + a11a23 y + a11a33 z = 0(a13a21-a11a23)y+(a13a31-a11a33)z = 0y = -(a11a33-a13a31)/(a11a23-a21a13) Substituting:

z(-a12(a23a31-a21a33)/(a11a23-a21a13)-a22(a11a33-a13a31)/(a11a23-a21a13)+a32)=1

Z=

-(a11a23-a21a13)/((a12(a23a31-a21a33))+a22(a11a33-a13a31)-a32(a11a23-a21a13))

z= - a11 a13 / D a21 a23	a11 a12 a13 D= a21 a22 a23 a31 a32 a33
y= 11 13 /D 31 33	a11 a12 a13 D= a21 a22 a23 a31 a32 a33
x= - 21 23 / D 31 33	a11 a12 a13 D= a21 a22 a23 a31 a32 a33

Finding inverse c_{ji} let x=c31 y=c32 z=c33 where we look at matrix multiplications of the bottom row of the left matrix by the three columns of the right matrix.

Substituting:

z(a13(a21a32-a22a31)/(a22a11-a21a12)-a23(a11a32-a12a31)/a22a32-a12a31)+a33)=1

(a11a22-a21a12)/(a13(a21a32-a22a31)-a23(a11a32-a12a31)+a33(a22a11-a21a12))

z=	a11 a a12 a	a12 / D a22	a11 a12 a13 D= a21 a22 a23 a31 a32 a33
y=	- a11 a31	a12 / D a32	a11 a12 a13 D= a21 a22 a23 a31 a32 a33
x=	a21 ; a31 ;	a22 / D a32	a11 a12 a13 D= a21 a22 a23 a31 a32 a33

The inverse matrix is:

(a22 a23	- a12 a13	a12 a13)	
(a32 a33	a32 a33	a22 a23)	
()	a11 a12 a13
(- a21 a23	a11 a13	- a11 a13)	/ a21 a22 a23
(a31 a33	a31 a33	a21 a23)	a31 a32 a33
()	
(a21 a22	- a11 a12	a11 a12)	
(a31 a32	a31 a32	a21 a22)	

Sample

```
(2 3 4)(x) (20)
                |2 3 4|
                            |2 3 4|
(3 6 9)(y) = (42) |3 6 9| = 3x|1 2 3| = 3x(36+36+28-32-42-27)=3x(100-101)=-3
(479)(z) (45)
               |4 7 9| |4 7 9|
( |69| -|34| |34|)
( |79| |79| |69|)
                    (-9 1 3)
(
               )
(-|39| |24| -|24|) = (92-6)/-3
( |4 9| |4 9| |3 9| )
                   (-3 -2 3)
(
               )
( |36| -|23| |23|)
( |47| |47| |36|)
```

z=

Check:

It is always important to use examples so that you can develop techniques for solving the problem and minimizing errors.

It we let $a_{13}=a_{31}=a_{32}=0$ and $a_{33}=1$ we get the following inverse matrix:

|a12 0|) (|a22 0| -|a120| (|0 1 |0 1| |a22 0|) | a11 a12 0| () (- |a21 0 | |a11 0 | -|a11 0|) / | a21 a22 0| (|0 1 |0 1| $|a21 \ 0|$) | 0 0 1| () (|a21 a22| -|a11 a12| |a11 a12|) (|0 0| | 0 0| |a21 a22|

(a2	2 -a12	0)		
= (-a2	l a11	0)	/	a11 a12
(() 0	a11 a1	12)		a21 a22
		a21 a	12)		

We get an answer consistent with that for the two by two inverse.

The 4x4 matrix

We are going to find the matrix and its inverse for the 4x4 matrix by using patterns that we observed from the 3x3 matrix. The determinant from the four by four matrix has to be viewed differently because we can only generate 8 terms rather that the 24 that we need if we use the diagonal approach. Thus we use the expanded format testing the patterns that we saw from the 3x3.

```
\begin{array}{c} x+2y+3z+4w=30\\ 2x+3y+4z+ w=24\\ 3x+4y+ z+2w=22\\ 4x+ y+2z+3w=24\\ \end{array}
```

|30 2 3 4|

= 160

1 30 3 4 $|2\ 24\ 4\ 1|=-30x|\ 2\ 4\ 1|+24x|1\ 3\ 4|-22x|1\ 3\ 4|+24x|1\ 3\ 4|=-30x-4+24x4+-22x-44+24x-36$ 3 22 1 2 312 3 1 2 |2 4 1| |2 4 1| | 4 2 3| |4 24 2 3| |4 2 3| |4 2 3| |3 1 2| =320 |1 2 30 4| $|2\ 3\ 24\ 1|=30x|\ 2\ 3\ 1|\ -24x|1\ 2\ 4|+22x|1\ 2\ 4|-24x|1\ 2\ 4|=30x4+-24x-44+22x-36+-24x-44$ 3 4 22 2 342 342 |2 3 1| |2 3 1| |4 1 24 3| | 4 1 3| |4 1 3| 342 4 1 3 =480

You can see that this was a lot of work, but with the organization and layout, we could solve it straightforwardly. The five by five matrix would be five times as much work.

When we analyzed the inverse for the 3x3 matrix, we found that starting with a positive sign in the upper left hand corner, we alternated signs in rows and columns which always assured that the sign of the diagonal element in each row was positive. We found the element for a cell by eliminating the row and column from the 4x4 determinant corresponding to the row and column of that element. However, the result of the operation was placed in the cell in which the row and column numbers of the element were interchanged. There were more patterns, but finding them is a task left to the reader.

Inverse matrix for 4x4

|a22 a23 a24| |a12 a13 a14| |a12 a13 a14| |a12 a13 a14| |a32 a33 a34| -|a32 a33 a34| |a22 a23 a24| -|a22 a23 a24 |a42 a43 a44| |a42 a43 a44| |a42 a43 a44| |a32 a33 a34| |a21 a23 a24| | a11 a13 a14| |a11 a13 a14| |a11 a13 a14| -|a31 a33 a34| | a31 a33 a34| -|a21 a23 a24| |a21 a23 a24| |a41 a43 a44| | a41 a43 a44| |a41 a43 a44| |a31 a33 a34| |a21 a22 a24| |a11 a12 a14| |a11 a12 a14| |a11 a12 a14| |a31 a32 a34| -|a31 a32 a34| |a21 a22 a24| -|a21 a22 a24| |a41 a42 a44| |a31 a32 a34| |a41 a42 a44| |a41 a42 a44| |a21 a22 a23| |a11 a12 a13| | a11 a12 a13| |a11 a12 a13| -|a31 a32 a33| |a31 a32 a33| -|a21 a22 a23| |a21 a22 a23| |a41 a42 a43| |a41 a42 a43| |a41 a42 a43| |a31 a32 a33| _____ | a11 a12 a13 a14| | a21 a22 a23 a24| | a31 a32 a33 a34| | a41 a42 a43 a44| Testing with numbers 3 4 1 2 3 4 2 3 4 2 3 4 |4 1 2| -|4 1 2| |3 4 1| -|3 4 1| |1 2 3| |1 2 3| |1 2 3| |4 1 2| |2 4 1| |1 3 4| |1 3 4| |1 3 4| -312 312 -241 241 |4 2 3| |4 2 3| |4 2 3| |3 1 2| |2 3 1| |1 2 4| |1 2 4| |1 2 4| |3 4 2| -|3 4 2| |2 3 1| -|2 3 1| |4 1 3| |4 1 3| |4 1 3| |3 4 2| |-36 4 4 44| |2 3 4| | 1 2 3| |1 2 3| |1 2 3| 4 4 44 - 36 -341 341 -234 234 | 4 44 -36 4| |44-36 4 4| 4 1 2 4 1 2 4 1 2 3 4 1 _____ -----|1 2 3 4| 160 |2 3 4 1| |3 4 1 2| |4 1 2 3|

(-9 1 11)	1) (30)	(-9x30+24+22+11x24)	(40) (1)
(1111-9	9) (24) =	1/40 x(30+24+11x22-9x24) =	(80) = (2)
1/40 x (1 11 – 9	1) (22)	(30+11x24 - 9x22 + 24)	1/40 x (120) (3)
(11-9 1 1	1) (24)	(11x30 - 9x24 + 22 + 24)	(160) (4)

The use of the inverse matrix made the problem look easy, but there was a lot of work, errors, and correction that were needed to find it. If the same set of equations with different constants on the right of the equal sign were presented, then we could see the value of finding it. We found patterns that would have made it a lot easier to generate the matrix. We could use this technique to generate the inverse for the 5x5 with little work. When a task becomes this big, we can write a program using our knowledge of the patterns to make it sufficiently general to solve for the inverse of a matrix of any size.