Matrices and Gaussian Elimination : Part 1

P. Sam Johnson

December 30, 2019

A Motivating Example

A shopkeeper offers two standard packets because he is convinced that north indians each more wheat than rice and south indians each more rice than wheat.

Packet one P_1 : 5 kg wheat and 2 kg rice ;

Packet two P_2 : 2 kg wheat and 5 kg rice.

Notation :

(m, n) for m kg wheat and n kg rice.

Suppose I need 19 kg of wheat and 16 kg of rice. Then I need to buy x packets of P_1 and y packets of P_2 so that $x(5,2) + y(2,5) = (10,16)$.

Suppose I need 34 kg of wheat and 1 kg of rice. Then I must buy 8 packets of P_1 and -3 packets of P_2. What does this mean? I buy 8 packets of P_1 and from these I make three packets of P_2 and give them back to the shopkeeper.

Motivating Example for Linear Map

Let the packet $P_1 = (5, 2)$ be priced at Rs.R_1 and $P_2 = (2, 5)$ be priced at Rs.R_2.

Then if f is the price of m packets of P_1 and n packets of P_2, we see that $f(mP_1 + nP_2) = mf(P_1) + nf(P_2)$.

Linear algebra is about linear spaces, also called vector spaces, and linear maps between them.

The central problem of linear algebra is the solution of linear equations. The way to understand this subject is by example. **We consider first n equations in n unknowns**. Two well-established ways to solve linear equations.

Method of elimination (**Gaussian elimination**). Multiples of the first equations are subtracted from the other equations, so as to remore the first unknowns from those equations. This leaves a smaller system of $n-1$ equations and in $n-1$ unknowns. The process is repeated until there is only one equation and one unknown, which can be solved immediately.

Sophisticated way of using the idea of determinants: Cramer's rule. The solution is as a ratio of two $n \times n$ determinants.

Geometry of Linear Equations

Suppose we have n equations with n unknowns. Each equation represents a $(n-1)$-dimensional plane in n-dimensional space. The first two equations intersect (we hope) in a smaller set of "dimension $n-2$".

Assuming all goes well, every new plane (every new equation) reduces the dimension by one. At the end, when all n planes are accounted for, the intersection has dimension zero. It is a point, it lies on all the planes, and its coordinates satisfy all n equations. It is the solution.

1. **Row picture.** Intersection (solution) of n planes (each plane is of $n-1$ dimension). With n equations in n unknowns, there are n-planes in the row picture.
2. **Column picture.** With n equations in n unknowns, there are n vectors in the column picture, plus a vector b on the right side. The right side b is a linear combination of the column vectors. Solution is the coefficients in the linear combination of columns.

Vector Equation

The n separate equations are really one "vector equation".

$x_1[\vdots] + \cdots + x_n[\vdots] = b$. The problem is to find the combination of the column vectors on the left side which produces the vector on the right side.

The geometry exactly breaks down, in what is called the "**singular case**".

1. **Row picture.** All planes are parallel - no solution ; two planes parallel - no solution ; no common intersection - no solution ; intersection of three plane is a line - infintely many solutions.

2. **Column picture.** $u[\vdots] + v[\vdots] + w[\vdots] = b$. Suppose "three column vectors" span a plane. Suppose if the vector b is not in that plane, then "no solution" case. Suppose b lies in the plane of the columns, there are too many solutions. In that case the **three columns can be combined in infinitely many ways to produce b. How do we know that the three columns lie in the same plane? We will check whether the three column vectors are linearly independent or not?**

An example of Gaussian elimination

Consider the system

$$\begin{aligned} 2u + v + w &= 5 \\ 4u - 6v &= -2 \\ -2u + 7v + 2w &= 9. \end{aligned}$$

The method starts by substracting multiples of the first equation from the others, so as to eliminate u from the last two equations. This requires that we

1. subtract 2 times the first equation from the second;
2. substract -1 times the first equation from the third.

The numbers 2 and -1 are called **multipliers**.

The result is an equivalent system of equations

$$\begin{aligned} 2u + v + w &= 5 \\ -8v - 2w &= -12 \\ 8v + 3w &= 14. \end{aligned}$$

The coefficient 2, which multiplied the first unknown u in the first equation, is known as the **first pivot**. Elimination is constantly dividing the pivot into the numbers underneath it, to find out the right multipliers. At the second stage of elimination, we ignore the first equation. We add the second equation to the third or, in other words, we "substract -1 times the second equation from the third". The elimination process is now complete, at least in the "forward" direction.

$$\begin{aligned} 2u + v + w &= 5 \\ -8v - 2w &= -12 \\ w &= 2. \end{aligned}$$

There is an obvious order in which to solve this system. The last equation gives $w = 2$. Substituting into the second equation, we find $v = 1$. Then the first equation gives $u = 1$. The process is called **back-substitution**.

Forward elimination produced the pivots $2, -8, 1$. It subtracted multiples of each row from the rows beneath. It reached the "triangular" system. Then this system was solved in reverse order, from bottom to top, by substituting each newly computed value into the equation above. By definition, **pivots cannot be zero**. We need to divide by them.

The breakdown of elimination. Under what cicumstances could the process break down? Something must go wrong in the singular case, and something might go wrong in the nonsingular case. The question is not geometric but algebraic.

If the algorithm produces n pivots, then there is only one solution to the equations. The system is nonsingular, and it is solved by forward elimination and back-substitution. But if a zero appears in a pivot position, elimination has to stop - either temporarily or permanently. The sytem might or might not be singular.

Cost of elimination.

For n equations in n unknowns, how many separate arithmetical operations does elimination require? For the moment, we ignore the right-hand sides of the equations, and count only operations on the left.

These operations are of two kinds.

1. One is a **division** by the pivot, to find out what multiple (say ℓ) of the pivot equation is to be subtracted.
2. Second is **multiplication-subtraction**: the terms in the pivot equation are multiplied by ℓ, and then subtracted from the equation beneath it.

Suppose we call each division, and each multiplication-subtraction, a single operation. There are $n-1$ rows underneath the first one, so the first stage of elimination needs $n(n-1) = n^2 - n$ operations.

(Another approach to $n^2 - n$ is this: All n^2 entries need to be changed, except the n in the first row). When the elimination is down to k equations, only $k^2 - k$ operations are needed to clear out the column below the pivot- by the same reasoning that applied to the first stage, when k equaled n. Altogether, the total number of operations on the left side of the equations is $\sum_{k=1}^{n} k(k-1) = (n^3 - n)/3$. Forward elimination is about a third of a million steps, a good code on a PC would take 41 seconds. If n is at all large, a good estimate for the number of operators is $n^3/3$.

Back substitution is considerably faster. The last unknown is found in only one operation (a division by the last pivot). The second to last unknown requires two operations, and so on. Then the total for back-substitution is $\sum_{k=1}^{n} = n(n+1)/2 \approx n^2/2$.

If we have 3×3 example, we could list the elimination steps which subtracts a multiplie of one equation from another and reach a triangular form.

Matrix Multiplication

Given a system with n equations with n unknowns. $AX = B$, A is **the coefficient matrix, square matrix; if m equations and n unknowns, rectangular matrix.** The first component of B is the the first component of the product AX must come from "multiplying" the first row of A into the column vector X.

It starts with a row and a column vector of matching lengths, and it produces a single number. This single quantity is called **the inner product of two vectors.**

If we look at the whole computation, multiplying a matrix by a vector, there are two ways to do it. One is to combine **a row at a time.** Each row of the matrix combines with the vector to give a component of the product. **There are three inner products when there are three rows.**

By rows: $$Ax = \begin{pmatrix} 1 & 1 & 6 \\ 3 & 0 & 3 \\ 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 1.2 + 1.5 + 6.0 \\ 3.2 + 0.5 + 3.0 \\ 1.2 + 1.5 + 4.0 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \\ 7 \end{pmatrix}$$

The second way is equally important. In fact it is more important. It does the multiplication **a column at a time.** The product Ax is found at all once, and it is combination of the three columns of A.

By columns: $$Ax = 2\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} + 5\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 6 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \\ 7 \end{pmatrix}.$$

The answer is twice column 1 plus 5 times column 2. **It corresponds to "the column picture" of the linear system $Ax = b$. The column rule will be used over and over throughout the course.**

Matrix Notation for the individual entries in A: The entry in the ith row and jth column is always denoted by a_{ij}. The first subscript gives the row number and the second subscript indicates the column. If A is an $m \times n$ matrix, then the index i goes 1 to m - there are m rows ; and the index j goes from 1 to n. Altogether, there matrix has mn entries, forming a rectangular array, and a_{mn} is the lower right corner.

$\sum_{j=1}^{n} a_{ij}x_j$ is the ith component of Ax, formed the inner product of ith row of A with x. This sum takes us along the ith row of A, forming its inner product with x. The length of the rows (the number of columns in A) must match the length of x. An $m \times n$ matrix multiplies an n-dimensional vector (and produces an m-dimensional vector). Summations are simpler to work with than writing everything out in full, but they are not as good as matrix notation itself.

Why matrix notation is preferred: We want to get on with the connection between matrix multiplication and Guassian elimination.

Matrix Multiplication.

Three different ways to look at matrix multiplication:

1. **Inner product.** Each entry AB is the product of a row and a column: $(AB)_{ij}$ = row i of A times column j of B (inner product of the ith row of A and the jth column of B).
2. **Column picture.** Each column of AB is the product of a matrix and a column: column j of $AB = A$ times column j of B.
3. **Row picture.** Each row of AB is the product of a row and a matirx: row i of $AB =$ row i of A times B.

The i, j entry of AB is the inner product of the ith row of A and the jth column of B.

Columns of AB: $\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 2a+3c & 2b+3d \end{pmatrix}$.

By columns, B consists of two columns side by side, and A multiplies each of them separately. Therefore each column of AB is the combination of the columns of A.

$$\begin{pmatrix} a \\ 2a+3c \end{pmatrix} = a \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

$$\begin{pmatrix} b \\ 2b+3d \end{pmatrix} = b \begin{pmatrix} 1 \\ 2 \end{pmatrix} + d \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

The first columns of AB is "a" times column 1 plus "c" times column 2.

Rows of AB: $AB = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 2a+3c & 2b+3d \end{pmatrix}$.

The first row of AB is $1[a\ b] + 0[c\ d] = [a\ b]$. The second row of AB is $2[a\ b] + 3[c\ d] = [2a+3c,\ 2b+3d]$. Each row of AB is a combination of the rows of B.

When the matrix B contains several columns, say x_1, x_2, x_3. We hope that the columns of AB are just Ax_1, Ax_2, Ax_3. The first column of AB equals A times the first column of B, and the same is true for the other columns.

1. Matrix multiplication is associative: $(AB)C = A(BC)$.
2. Matrix operations are distributive: $A(B + C) = AB + AC$ and $(B + C)D = BD + CD$.
3. Matrix multiplication is not commutative: Usually $FE \neq EF$.

References

- S. Kumaresan, "*Linear Algebra - A Geometric Approach*", PHI Learning Pvt. Ltd., 2011.
- Gilbert Strang, "*Linear Algebra and its Applications*", Cengage Learning, New Delhi, 2006.