Methods of Mathematics

Kenneth A. Ribet

UC Berkeley

Math 10B
April 14, 2016

I'm really glad to be back!

The next breakfast will be one week from today, April 21 at 8:30AM. Send me email to sign up.

There's a "pop-up" lunch tomorrow, April 15 at the Faculty Club at 12:30PM. All are welcome.

A reminder: there's office hour in 885 Evans on Friday, April 15: 11AM–12:20PM.

Recap of Tuesday

A system of m linear equations in n unknowns can be recast as the single matrix equation

$$Ax = b,$$

where b is a column matrix of length m, x is a column matrix of length n (the column of “unknowns”) and A is a rectangular matrix of size $m \times n$. This means that A has n columns and m rows.

We are multplying $m \times n$ by $n \times 1$ to get $m \times 1$.

In the general matrix multiplcation $AB = C$, the dimensions might be $m \times n$, $n \times \ell$, $m \times \ell$. The rule is that the middle numbers have to match; the outer numbers survive in the product.

We can add two matrices of the same size. We can multiply a matrix by a number. We can take the transpose of a matrix.

I hope that Jason told you also about the inner product of two vectors of length n and about the norm ("length") of a single vector. The norm of $(a_1, \ldots, a_n)$ is $\sqrt{(a_1^2 + \cdots + a_n^2)}$.

The famous Cauchy–Schwarz inequality states that the absolute value of the dot product of two vectors is at most the product of the norms of the two vectors.

A proof is sketched in the online "textbook."

For $n \geq 1$, the $n \times n$ identity matrix $I = I_n$ is the matrix of the indicated size that has 1s on the diagonal and 0s off the diagonal. A computation shows that

$$IA = A, \quad BI = B$$

whenever A has n rows and B has n columns. Multiplying by I is a no-op: nothing happens.

If A is a square matrix of size n (i.e., an $n \times n$ matrix), an *inverse* for A is a matrix A^{-1} with the property that

$$A^{-1}A = AA^{-1} = I_n.$$

It's easy to see that an inverse for A is unique if it exists—in other words, a matrix can't have two inverses.

It's also obvious that some matrices do not have inverses. For example the $n \times n$ matrix with all 0s has no inverses.

Only sqaure matrices are invertible. This is a consequence of the definition of *inverse*.

If A is invertible, multiplication by A^{-1} "undoes" multiplication by A. For example, if B has n rows, then

$$A^{-1}(AB) = (A^{-1}A)B = I_nB = B$$

and a similar computation can be made for $(CA)A^{-1}$ when C has n columns.

If you have a system

$$Ax = b$$

of n equations in n unknowns, and if A happens to be invertible, then you can multiply both sides of $Ax = b$ by A^{-1} to get the equivalent equation $x = A^{-1}b$. This latter equation just says that $Ax = b$ has a unique solution, namely $A^{-1}b$.

If someone hands you an $n \times n$ matrix A, you might want to answer questions like these:

- Does A have an inverse?
- Can I find an inverse for A?

These are serious questions if n is very large. You might find it amusing to do a `google` search on something like "inverting a large matrix."

Actually a typical hit was the post Don't invert that matrix. In many situations, it's more efficient for calculations to avoid explicit inverses.

By the way, if you want to read a rant about the teaching of differential equations in math courses, I suggest this essay by Gian-Carlo Rota.

I don't know the date of this essay. As you can see from Wikipedia, Rota died in 1999.

My friend's $e^{2i\pi}$ license plate makes him number ONE.

Facts about invertibility

- If A is an $n \times n$ matrix, A is invertible if and only if $Ax = b$ has at least one solution x for every column matrix b of length n.
- If A is an $n \times n$ matrix, A is invertible if and only if $Ax = b$ has at most one solution x for every column matrix b of length n.
- If A is an $n \times n$ matrix, A is invertible if and only if $Ax = b$ has exactly one solution x for every column matrix b of length n.
- An $n \times n$ matrix A is invertible if and only if the equation $Ax = 0$ (for x a column of length n) has only the solution $x = 0$.

We now come to a discussion of determinants–and their connection with invertibility.

Down With Determinants!

Sheldon Axler

1. INTRODUCTION. Ask anyone why a square matrix of complex numbers has an eigenvalue, and you'll probably get the wrong answer, which goes something like this: The characteristic polynomial of the matrix—which is defined via determinants—has a root (by the fundamental theorem of algebra); this root is an eigenvalue of the matrix.

What's wrong with that answer? It depends upon determinants, that's what. Determinants are difficult, non-intuitive, and often defined without motivation. As we'll see, there is a better proof—one that is simpler, clearer, provides more insight, and avoids determinants.

To every square matrix A, we associate a number $\det A$. The main properties of "det" are:

1. For $n \times n$ matrices A and B, $\det(AB) = \det A \cdot \det B$;
2. a square matrix is invertible if and only if its determinant is non-zero;
3. If A is invertible, we can calculate A^{-1} fairly easily if we happen to know $\det A$;
4. Sheldon Axler is right when he says that determinants are not very useful for matrices of very large size.

A common notation for the determinant is a set of vertical bars. If A is a matrix, $|A|$ usually denotes the determinant of A (i.e., $\det A$).

As we'll see in the next slide,

$$\begin{vmatrix} 8 & -2 \\ 3 & 4 \end{vmatrix} = 38.$$

That's how the notation is used.

If $A = (a)$ is a 1×1 matrix, then $\det A = a$.

If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 matrix, then $\det A = ad - bc$.

A useful formula is

$$\det A \cdot I = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

An immediate consequence is that

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

whenever $\det A$ is non-zero.

If $A = (a)$ is a 1×1 matrix, then $\det A = a$.

If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 matrix, then $\det A = ad - bc$.

A useful formula is

$$\det A \cdot I = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

An immediate consequence is that

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

whenever $\det A$ is non-zero.

For A an $n \times n$ matrix, one can write down a companion matrix B such that

$$\det A \cdot I = AB;$$

this generalizes one of the formulas on the previous slide. Up to signs and a transpose, the entries in B are determinants of $(n-1) \times (n-1)$ matrices that are derived from A by the following silly rule: to get the (i, j)th $(n-1) \times (n-1)$ matrix, you remove from A its ith row and jth column.

That's a mouthful, but I intend to illustrate the procedure for $n = 3$ on the document camera.

Speaking of 3×3 matrices, you might guess that it would be hard to remember the formula for the determinant of a 3×3 matrix, namely

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + dhc - (gec + hfa + idb).$$

As I'll explain on the board, there's a good mnemonic that enables me to remember the formula without fail.

An alternative equivalent formula for the determinant above is

$$a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

This formula expresses a 3×3 determinant as an alternating sum of three 2×2 determinants.

A natural and fairly obvious generalization of this formula expresses an $n \times n$ determinant as an alternating sum of n determinants of matrices of size $n - 1$. I'll illustrate on the board (or document camera) how to compute a 4×4 determinant by this method.

In Math 110 (our upper-division linear algebra course), the two most common textbooks are by Axler ("Linear algebra done right") and by Friedberg, Insel and Spence ("Linear Algebra"). The first book barely mentions determinants, except to say that they are evil. The second book defines determinants recursively by the formula that I'm about to illustrate.

A matrix is called *diagonal* if the off-diagonal entries are 0. If A is such a matrix, then its determinant in the product of its diagonal entries.

In particular, $\det I = 1$ if I is an identity matrix.

Because the determinant of a product of matrices is the product of their determinants, it follows that $\det A^{-1} = 1/(\det A)$ if A is invertible.

A final fact is that a matrix A and its transpose have the same determinant.