Linear Algebra – Lesson 3

Determinant, trace and the inverse matrix

In the previous lesson, we talked about matrices, and how they can be thought of as representing linear transformation. We also talked about two important numbers that we can calculate for every matrix: its rank and nullity. We said that the **rank** is the dimension of the matrix image (or column space), and the **nullity** is the dimension of the matrix kernel (or null space).
Today we will start by learning two additional important numbers that characterize a matrix: its **determinant** and its **trace**.
Then we will see how we can use matrices to solve systems of linear equations.
Finally, we will talk about how we can "undo" the transformation induced by a matrix, by calculating the **inverse matrix**.

1. The determinant

1.1 Geometrical explanation

So far, when we looked at matrices, we thought of them as representing some linear transformation, and studied how they transform one vector into another vector. We could also think about how a matrix transform one **shape** to another.

Definition: The determinant of a matrix A (denoted $\det(A)$ or $|A|$) is the signed scaling factor that the matrix induces. Similarly, the determinant of a matrix A is the signed volume of the unit parallelogram after applying A.

$$Area = |\det(A)|$$

As you saw in the videos, the determinant can also be negative. This happens when the orientation of the coordinate system changes, for example when you flip one axis, or when you permute two axes.

Example: Effect of rotation and scaling on the unit square:

$A = \begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix}$

This transformation scales all the basis vectors by $2\sqrt{2}$, and also rotates them by 90 degrees clockwise[1].
What would happen to the unit square after applying A? It will be rotated by 90 degrees clockwise, and its area will increase by $\left(2\sqrt{2}\right)^2 = 8$.

Question: Let A be an $n \times n$ matrix. What does it mean if $\det(A) = 0$?

---

[1] You can see it like that: $\begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. The first matrix just scales both axes by 2. The second matrix rotates by 45 degrees, and also scales vectors by $\sqrt{2}$ (think about the vector $\hat{\jmath}$, for example: it moved from $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and now its length is $\sqrt{2}$.

Answer: A determinant of 0 means that the volume of the unit parallelogram after applying A is 0. This will happen whenever the "effective" dimension of the parallelogram is lower than the original dimension. For example, look at $A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$. This transformation takes any vector, and throws it to the horizontal line spanned by $\hat{\imath}$. The area of the unit square will obviously become 0.
Note that $\det(A) = 0$ means that the columns of A are linearly dependent.

1.2 Determinant properties

Knowing how the determinant changes if we change the matrix is helpful for different calculations:

- The determinant is only defined for square matrices
- $\det(I) = 1$
- $\det(A) = \det(A^T)$ (a proof by induction can be found here)
- If the columns (or rows) of A are linearly dependent, then $\det(A) = 0$
    - If A has a row (or column) of zeros, then $\det(A) = 0$
- $\det(kA) = k^n \det(A)$
- A matrix is called upper triangular if it has non-zero elements only on the diagonal and above it (and it is called lower diagonal if it has non-zero elements only on the diagonal and below it).
  If A is triangular (it has zeros above or below the diagonal), then[2]:

$$\det(A) = \prod_{i=1}^{n} a_{ii} \qquad for\ A = \begin{pmatrix} a_{11} & * & * \\ 0 & \dots & * \\ 0 & 0 & a_{nn} \end{pmatrix}$$

    - If A is diagonal $n \times n$ matrix, then:

$$\det(A) = \prod_{i=1}^{n} a_{ii} \qquad for\ A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & a_{nn} \end{pmatrix}$$

- The determinant of a product of two square matrixes A, B is the product of their determinants:
  $$\det(AB) = \det(A)\det(B)$$
  Remember, that in general it's not true that $AB = BA$. Nevertheless, from this final property we see that $\det(AB) = \det(BA)$.

Two more properties will make more sense when we talk about linear systems of equations:

- If you interchange two rows or columns, the sign of the determinant changes (i.e.,$\det(A) \rightarrow -\det(A)$).
- If you add a multiple of one row to another, the determinant remains unchanged.

1.3 Determinant calculation

The determinant of a 1×1 matrix (a scalar) is the scalar itself:

$$\det(k) = k$$

The determinant of an order 2 matrix is:

[2] You can see why this is true if you know how to calculate the determinant (see next sub-section), and use the first column to do it.

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

In words, the determinant is the product of the elements along the main diagonal, minus the product of the off-diagonal elements.

The determinant of a 3×3 matrix can be calculated using determinants of 2×2 submatrices:
It is a linear combination of three determinants of order 2, whose coefficients (with alternating signs) are the first row the matrix:

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Note that we can find each 2×2 submatrix by deleting the row and column of the corresponding coefficient in the original matrix:

$$a_{11}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

In fact, you can use any row/column and use it to calculate the determinant as we did here for the first row. But be careful, you have to know how the signs change in every line/column:

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

2. The trace

Another useful number you can extract from a matrix is its trace.

<u>2.1 Trace definition</u>

The trace of a matrix A is the sum of its diagonal elements:

$$tr(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

<u>2.2 Trace properties</u>

- $tr(A + B) = tr(A) + tr(B)$
- $tr(kA) = k\, tr(A)$
- $tr(A^T) = tr(A)$
- $tr(AB) = tr(BA)$ (notice that the trace is equal although in general $AB \neq BA$)

3. The inverse matrix as the reverse transformation

Now we know a lot about matrices. Given a matrix A, we can tell what it does to the basis vectors, what its rank is (the dimension of its image), what its nullity is (the dimension of its kernel), and also its determinant and trace. Now we'll see how we can use some of them to find the inverse matrix.

3.1 Inverse matrix – intuition and examples

The inverse matrix A^{-1} is the matrix which "undoes" what the matrix A does.
Let's see two examples.

Example 1 – an inverse matrix exists

Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. What is the geometrical interpretation of what A does? It's a rotation of 90 degrees counter clockwise. For example:

$$A\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

How can we "undo" its effect? We can apply a rotation of 90 degrees counter clockwise:

$$A^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Indeed:

$$A^{-1}\begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Or in other words:

$$A^{-1}Av = v \quad \rightarrow \quad A^{-1}A = I$$

Example 2 – an inverse matrix does not exist

Now let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. What is the geometrical interpretation of what A does? We've seen it before: it's a projection of 3D space to the XY plane. For example:

$$A\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$$

How can we "undo" its effect? In other words, how can we go back from $\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$ to the original $\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$? Well, we can't. There is an infinite number of vectors of the form $\begin{pmatrix} 1 \\ 3 \\ z \end{pmatrix}$ that are all projected to $\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$.

In general, let's imagine that A did have an inverse A^{-1}, and we tried to solve this equations:

$$A\bar{x} = 0 \quad \rightarrow \quad A^{-1}A\bar{x} = A^{-1}0 \quad \rightarrow \quad \bar{x} = 0$$

We found that $\bar{x} = 0$. But that's obviously not true, because we know that our $\bar{x}$ could have been non-zero, like $\begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix}$.

3.2 Inverse matrix – definition

A square matrix A is invertible if there exists a matrix B such that:

$$AB = BA = I$$

The inverse matrix B is denoted A^{-1}.

An invertible matrix is also called "regular".
A non-invertible matrix is called "singular".

3.3 Conditions for invertibility
What can we learn from the second example above? What is "wrong" with the matrix A, that makes it non-invertible? To answer this, let's examine some of its properties. Together we will fill this table, for the equivalent conditions that determine that A is not-invertible:

For $A_{n\times n}$:

Kernel	$\ker(A) \neq \{0\}$
Nullity	$null(A) \neq 0$
Image	$Im(A) \neq \mathbb{R}^n$
Rank	$rank(A) < n$
Determinant	$\det(A) = 0$
Columns linear dependence	Columns of A are linearly dependent

We can see that $ker(A) = \left\{\begin{pmatrix} 0 \\ 0 \\ v_3 \end{pmatrix} \middle| v_3 \in \mathbb{R}\right\}$. In other words, the kernel of A includes more than the zero vector.
This means that the dimension of the kernel (the nullity of A) is not 0.
What about its image and rank? The image of A (the span of its columns) is the XY plane:
$Im = \left\{\begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} \middle| v_1, v_2 \in \mathbb{R}\right\}$. In other words, $rank(A) = 2 < 3$.
Its determinant is 0 (remember, it has a row of zeros).
Finally, since one of the columns is the zero vector, we get that the columns of A are linearly dependent.

As an exercise, study the rotation matrix A from our first example, and see that none of these conditions is satisfied.

Note: Another way to think about singular matrices[3] – Let's take a matrix with linearly dependent columns, like $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$, and look for its inverse B (meaning, we look for $AB = BA = I$). Each column in the result of AB will be some linear combination of the columns of A (because this is always the case for matrix multiplication: if $AB = C$, then the columns of C are linear combinations of the columns of A). But this means they will all be linearly dependent as well, so they definitely can't give us the two linearly independent columns we are looking for, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

3.2 Calculation of the inverse matrix
Given an invertible matrix A, how can we calculate its inverse A^{-1}?

[3] Based on Strang's video 3, around 26:00.

The general method for calculating the inverse matrix is called “Gaussian Elimination”. We will not learn it now. For now, let’s find a formula for the inverse of a 2×2 matrix.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We are looking for a matrix B that satisfies: $AB = I$.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} ax_1 + by_1 & ax_2 + by_2 \\ cx_1 + dy_1 & cx_2 + dy_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

These are four equations:
$ax_1 + by_1 = 1$
$cx_1 + dy_1 = 0$
$ax_2 + by_2 = 0$
$cx_2 + dy_2 = 1$

Taking the first two equations, and multiplying by d and b respectively:
$adx_1 + dby_1 = d$
$bcx_1 + dby_1 = 0$

Subtracting the second equation from the first:
$(ad - bc)x_1 = d$
$$x_1 = \frac{d}{ad - bc}$$
But what is the denominator? That’s just the determinant of A:
$$x_1 = \frac{d}{\det(A)}$$
Similarly, we can get the full solution:
$$A^{-1} = \frac{1}{\det(A)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$
Notice that we just have to switch the elements on the main diagonal, add a minus sign to the elements off the diagonal, and divide everything by $\det(A)$.
Also notice that we can easily see that if $\det(A) = 0$, then A^{-1} is not defined.

3.3 The inverse of a composite matrix

Now, given two square and invertible matrices A, B, what is the inverse of the composite matrix AB?

Before I tell you the answer, let’s try to think about it intuitively.
Remember what happens when we apply a composite matrix on a vector:
$$(AB)\bar{v} = AB\bar{v}$$
We read it from right to left: first we apply B and get a new vector $B\bar{v}$. Then we apply A and get the result. To “undo” this, we just have to apply the reverse transformations in the correct order – we always undo the most recent transformation. So first, we have to undo A:
$$A^{-1}(AB\bar{v})$$
Now, we have to undo B:

$$B^{-1}(A^{-1}AB\bar{v})$$

But this means that:

$$(B^{-1}A^{-1})(AB)\bar{v} = \bar{v}$$

By definition, $B^{-1}A^{-1}$ is exactly the inverse matrix we were looking for. In general:

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

A formal proof is based on associativity – if we "guess" that the inverse matrix is $B^{-1}A^{-1}$, all we have to do is test this:

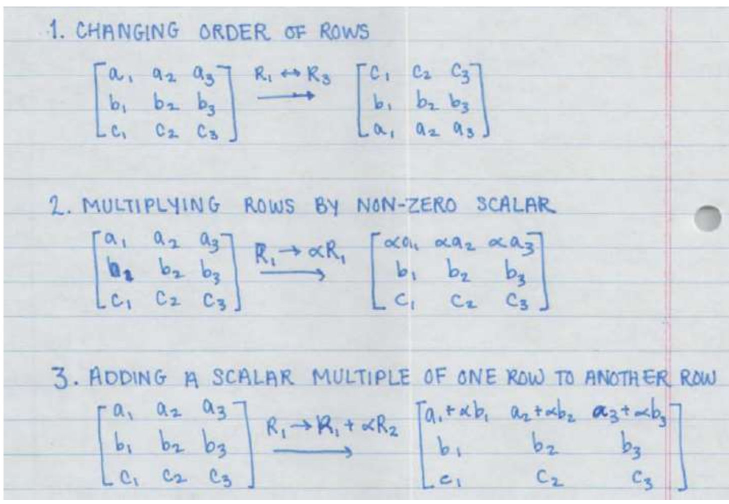
$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

We can prove the other side in exactly the same way, by multiplying from the right.

3.4 A general algorithm for finding the inverse matrix

See lesson presentation on Moodle for an explanation of this section.

There is a general algorithm for finding the inverse of a matrix (or figuring out that it does not exist, if the matrix is singular). This algorithm is based on three "elementary row operations":



It turns out that every such operation has a matrix associated with it, which we call an elementary matrix. The elementary matrix is simply the identity matrix after applying the elementary row operation to it.

For example, for a 3x3 matrix the elementary matrix of changing rows 1 and 2 is:

$$R1 \leftrightarrow R2\text{:} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The elementary matrix that adds a multiple of 3 of the first row to the third row is:

$$R3 \rightarrow R3 + 3R1\text{:} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

For an example on how to use row operation to find the inverse matrix, watch this short video: https://www.youtube.com/watch?v=HwRRdG_E4Yo.