Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-10-30 (updated version)

Darij Grinberg

December 9, 2019

1. Determinants

1.1. Some determinant identities (recall)

Recall how we defined determinants:

Definition 1.1.1. Let A be an $n \times n$-matrix. Then, the **determinant** $\det A$ of A is defined to be the sum

$$\sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}.$$

Last time we proved the following properties of determinants:

Theorem 1.1.2. If an $n \times n$-matrix A is triangular (i.e., upper-triangular or lower-triangular), then its determinant is the product of its diagonal elements:

$$\det A = A_{1,1} A_{2,2} \cdots A_{n,n}.$$

Theorem 1.1.3. If A is any $n \times n$-matrix, then $\det(A^T) = \det A$.

Theorem 1.1.4. If an $n \times n$-matrix A has two equal rows, then $\det A = 0$.

Theorem 1.1.5. If we scale a row of an $n \times n$-matrix A by a number λ, then $\det A$ gets multiplied by λ.

1.2. More determinant identities

Corollary 1.2.1. If an $n \times n$-matrix A has a zero row (i.e., a row full of zeroes), then $\det A = 0$.

Proof. Assume that the $n \times n$-matrix A has a zero row. Then, scaling this row by 0 does not change A, but (by Theorem 1.1.5) it multiplies $\det A$ by 0. Hence,

$$\det A = 0 \cdot \det A = 0.$$

□

Example 1.2.2. Here is how this argument works if $n = 3$ and if the second row of A is zero:

$$\det \begin{pmatrix} a & b & c \\ 0 & 0 & 0 \\ g & h & i \end{pmatrix} = 0 \cdot \det \begin{pmatrix} a & b & c \\ 0 & 0 & 0 \\ g & h & i \end{pmatrix} = 0.$$

The next theorem goes under the name "multilinearity of the determinant with respect to the rows" or "the determinant is linear in each row".

Theorem 1.2.3. Let n be a nonnegative integer and $k \in [n]$. (Recall that $[n]$ stands for the set $\{1, 2, \ldots, n\}$.)

Let A, B and C be three $n \times n$-matrices that differ from each other only in their k-th rows: i.e., that satisfy

$$\operatorname{row}_i A = \operatorname{row}_i B = \operatorname{row}_i C \qquad \text{for all } i \neq k.$$

Further assume that

$$\operatorname{row}_k C = \operatorname{row}_k A + \operatorname{row}_k B.$$

Then,

$$\det C = \det A + \det B.$$

Example 1.2.4. (a) If $n = 3$ and $k = 2$, then three $n \times n$-matrices A, B, C that differ from each other only in their k-th rows look as follows:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}, \qquad B = \begin{pmatrix} a & b & c \\ d' & e' & f' \\ g & h & i \end{pmatrix}, \qquad C = \begin{pmatrix} a & b & c \\ d'' & e'' & f'' \\ g & h & i \end{pmatrix}$$

for some reals $a, b, c, d, e, f, d', e', f', d'', e'', f'', g, h, i$. If they furthermore satisfy $\operatorname{row}_k C = \operatorname{row}_k A + \operatorname{row}_k B$, then we must have $d'' = d + d'$ and $e'' = e + e'$ and $f'' = f + f'$.

Hence, if $n = 3$ and $k = 2$, then Theorem 1.2.3 is saying that

$$\det \underbrace{\begin{pmatrix} a & b & c \\ d + d' & e + e' & f + f' \\ g & h & i \end{pmatrix}}_{\text{this is } C} = \det \underbrace{\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}}_{\text{this is } A} + \det \underbrace{\begin{pmatrix} a & b & c \\ d' & e' & f' \\ g & h & i \end{pmatrix}}_{\text{this is } B}.$$

(b) If $n = 3$ and $k = 1$, then Theorem 1.2.3 is saying that

$$\det \begin{pmatrix} a+a' & b+b' & c+c' \\ d & e & f \\ g & h & i \end{pmatrix} = \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} + \det \begin{pmatrix} a' & b' & c' \\ d & e & f \\ g & h & i \end{pmatrix}.$$

Warning: Theorem 1.2.3 does **not** say that $\det (A+B) = \det A + \det B$. Indeed, if A and B are two $n \times n$-matrices with $n > 1$, then $\det A$ and $\det B$ do not determine $\det (A+B)$. For example, for $n = 2$, if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$, then

$$\begin{aligned}
\det (A+B) &= \det \begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix} \\
&= \det \begin{pmatrix} a & b \\ c+c' & d+d' \end{pmatrix} + \det \begin{pmatrix} a' & b' \\ c+c' & d+d' \end{pmatrix} \\
&\qquad \text{(by Theorem 1.2.3, applied to } n = 2 \text{ and } k = 2\text{)} \\
&= \underbrace{\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{=A} + \det \begin{pmatrix} a & b \\ c' & d' \end{pmatrix} + \det \begin{pmatrix} a' & b' \\ c & d \end{pmatrix} + \underbrace{\det \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}}_{=B} \\
&\qquad \text{(by Theorem 1.2.3, applied to } n = 2 \text{ and } k = 1\text{)} \\
&\neq \det A + \det B \qquad \text{(in general)}.
\end{aligned}$$

Proof of Theorem 1.2.3. Let us first prove Theorem 1.2.3 on an example: Set $n = 3$ and $k = 2$. Then, the definition of a determinant of a 3×3-matrix yields

$$\begin{aligned}
&\det \begin{pmatrix} a & b & c \\ d+d' & e+e' & f+f' \\ g & h & i \end{pmatrix} \\
&= a\left(e+e'\right)i + b\left(f+f'\right)g + c\left(d+d'\right)h - a\left(f+f'\right)h - b\left(d+d'\right)i - c\left(e+e'\right)g \\
&= aei + ae'i + bfg + bf'g + cdh + cd'h - afh - af'h - bdi - bd'i - ceg - ce'g \\
&= \underbrace{(aei + bfg + cdh - afh - bdi - ceg)}_{=\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}} + \underbrace{(ae'i + bf'g + cd'h - af'h - bd'i - ce'g)}_{=\det \begin{pmatrix} a & b & c \\ d' & e' & f' \\ g & h & i \end{pmatrix}} \\
&\qquad \left(\begin{array}{c} \text{here, we have sorted all products that involve one of } d, e, f \\ \text{into one pair of parentheses, and all products that involve} \\ \text{one of } d', e', f' \text{ into another} \end{array} \right) \\
&= \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} + \det \begin{pmatrix} a & b & c \\ d' & e' & f' \\ g & h & i \end{pmatrix}.
\end{aligned}$$

This proves Theorem 1.2.3 when $n = 3$ and $k = 2$.

A similar argument applies in general (but of course, much more bookkeeping is required in the general case): The definition of a determinant yields

$$\det A = \sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot \underbrace{A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}}_{\text{let me call this product } A_\sigma}$$
$$= \sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot A_\sigma, \tag{1}$$

where we set $A_\sigma := A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$ for any permutation σ. Similarly,

$$\det B = \sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot B_\sigma, \tag{2}$$

where we set $B_\sigma = B_{1,\sigma(1)} B_{2,\sigma(2)} \cdots B_{n,\sigma(n)}$. Similarly,

$$\det C = \sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot C_\sigma, \tag{3}$$

where we set $C_\sigma = C_{1,\sigma(1)} C_{2,\sigma(2)} \cdots C_{n,\sigma(n)}$.

Now, fix a permutation σ of $[n]$. Consider the product $A_\sigma = A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$. One of its factors is $A_{k,\sigma(k)}$, which is an entry of the k-th row of A. All its other factors come from the other $n-1$ rows of A. Thus, we can write it as $A_\sigma = A_{k,\sigma(k)} \widetilde{A}_\sigma$, where

$$\widetilde{A}_\sigma = A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{k-1,\sigma(k-1)} A_{k+1,\sigma(k+1)} \cdots A_{n,\sigma(n)} \tag{4}$$

is the product of all factors of $A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{n,\sigma(n)}$ other than $A_{k,\sigma(k)}$. Similarly, we can write B_σ as $B_\sigma = B_{k,\sigma(k)} \widetilde{B}_\sigma$, where

$$\widetilde{B}_\sigma = B_{1,\sigma(1)} B_{2,\sigma(2)} \cdots B_{k-1,\sigma(k-1)} B_{k+1,\sigma(k+1)} \cdots B_{n,\sigma(n)} \tag{5}$$

is a product similar to $\widetilde{A}_\sigma$ but using the entries of B instead of A. Similarly, we can write C_σ as $C_\sigma = C_{k,\sigma(k)} \widetilde{C}_\sigma$, where

$$\widetilde{C}_\sigma = C_{1,\sigma(1)} C_{2,\sigma(2)} \cdots C_{k-1,\sigma(k-1)} C_{k+1,\sigma(k+1)} \cdots C_{n,\sigma(n)} \tag{6}$$

is a product similar to $\widetilde{A}_\sigma$ but using the entries of C instead of A.

But we assumed that the matrices A, B and C differ from each other only in their k-th rows. Thus, the entries $A_{i,\sigma(i)}$ that appear on the right hand side of (4) are the same as the entries $B_{i,\sigma(i)}$ that appear on the right hand side of (5) and also the same as the entries $C_{i,\sigma(i)}$ that appear on the right hand side of (6) (since no entries from the k-th row of any matrix appear on any of these right hand sides). Therefore, the right hand sides of the three equalities (4), (5) and (6) are equal. Thus, the left hand sides are equal as well. In other words,

$$\widetilde{A}_\sigma = \widetilde{B}_\sigma = \widetilde{C}_\sigma. \tag{7}$$

Moreover, we assumed that $\operatorname{row}_k C = \operatorname{row}_k A + \operatorname{row}_k B$. Thus, $C_{k,\sigma(k)} = A_{k,\sigma(k)} + B_{k,\sigma(k)}$ (since $A_{k,\sigma(k)}$, $B_{k,\sigma(k)}$ and $C_{k,\sigma(k)}$ are corresponding entries in $\operatorname{row}_k A$, $\operatorname{row}_k B$ and $\operatorname{row}_k C$). Hence,

$$C_\sigma = \underbrace{C_{k,\sigma(k)}}_{=A_{k,\sigma(k)}+B_{k,\sigma(k)}} \widetilde{C}_\sigma = \left(A_{k,\sigma(k)} + B_{k,\sigma(k)}\right) \widetilde{C}_\sigma = A_{k,\sigma(k)} \underbrace{\widetilde{C}_\sigma}_{\substack{=\widetilde{A}_\sigma \\ \text{(by (7))}}} + B_{k,\sigma(k)} \underbrace{\widetilde{C}_\sigma}_{\substack{=\widetilde{B}_\sigma \\ \text{(by (7))}}}$$
$$= \underbrace{A_{k,\sigma(k)} \widetilde{A}_\sigma}_{=A_\sigma} + \underbrace{B_{k,\sigma(k)} \widetilde{B}_\sigma}_{=B_\sigma} = A_\sigma + B_\sigma.$$

Now, forget that we fixed σ. We thus have proved that $C_\sigma = A_\sigma + B_\sigma$ for every permutation σ. Hence, (3) becomes

$$\begin{aligned}
\det C &= \sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot \underbrace{C_\sigma}_{=A_\sigma + B_\sigma} \\
&= \sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot (A_\sigma + B_\sigma) \\
&= \underbrace{\sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot A_\sigma}_{\substack{=\det A \\ \text{(by (1))}}} + \underbrace{\sum_{\sigma \text{ is a permutation of } [n]} \operatorname{sign}(\sigma) \cdot B_\sigma}_{\substack{=\det B \\ \text{(by (2))}}} \\
&= \det A + \det B.
\end{aligned}$$

This proves Theorem 1.2.3. $\square$

Corollary 1.2.5. Let A be an $n \times n$-matrix, and let p and q be two distinct elements of $[n]$. If we add $\lambda \cdot \operatorname{row}_p A$ to the q-th row of A, then $\det A$ does not change.

Proof of Corollary 1.2.5. This is another "proof by example": We only consider the case when $n = 3$, $p = 3$ and $q = 2$, and we write our 3×3-matrix A as $\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Then, we need to prove

$$\det \begin{pmatrix} a & b & c \\ d + \lambda g & e + \lambda h & f + \lambda i \\ g & h & i \end{pmatrix} = \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

(because the matrix on the left hand side is what we obtain if we add $\lambda \cdot \operatorname{row}_p A$ to the q-th row of A).

By Theorem 1.2.3 (applied to $n = 3$ and $k = 2$), we have

$$\det \begin{pmatrix} a & b & c \\ d + \lambda g & e + \lambda h & f + \lambda i \\ g & h & i \end{pmatrix}$$

$$= \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} + \underbrace{\det \begin{pmatrix} a & b & c \\ \lambda g & \lambda h & \lambda i \\ g & h & i \end{pmatrix}}_{\substack{= \lambda \det \begin{pmatrix} a & b & c \\ g & h & i \\ g & h & i \end{pmatrix} \\ \text{(by Theorem 1.1.5,} \\ \text{because the matrix has had its second row} \\ \text{scaled by } \lambda)}}$$

$$= \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} + \lambda \underbrace{\det \begin{pmatrix} a & b & c \\ g & h & i \\ g & h & i \end{pmatrix}}_{\substack{=0 \\ \text{(by Theorem 1.1.4,} \\ \text{since the matrix here has two equal rows)}}}$$

$$= \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix},$$

which is exactly what we need to prove.

The general case (with n, p and q arbitrary) is proved in the same way. □

Theorem 1.2.6. If we swap two rows of an $n \times n$-matrix, then its determinant gets multiplied by -1 (that is, it flips its sign but preserves its magnitude).

Proof of Theorem 1.2.6. You can swap rows i and j of a matrix by performing the following sequence of row operations:[1]

- add row i to row j;
- then subtract row j from row i;
- then add row i to row j;
- then scale row i by -1.

(You can easily convince yourself of this claim by tracking what happens to the entries in rows i and j when these four row operations are performed.[2])

What happens to the determinant when we perform these four row operations? The first three operations do not change the determinant, because they all have

[1] This is a version of what is known as the "XOR swapping technique".

[2] Proof by example: Let us assume that $n = 1$ and $i = 1$ and $j = 2$. Thus, we want to swap rows 1 and 2 of the 2×2-matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Our sequence of row operations then does the following:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{\text{add row 1 to row 2}} \begin{pmatrix} a & b \\ a+c & b+d \end{pmatrix} \xrightarrow{\text{subtract row 2 from row 1}} \begin{pmatrix} -c & -d \\ a+c & b+d \end{pmatrix}$$

the form "add $\lambda \cdot \operatorname{row}_p A$ to the q-th row" for some values of p, q and λ (and as we know from Corollary 1.2.5, the determinant does not change when we perform such operations). The fourth operation multiplies the determinant by -1, because of Theorem 1.1.5. Thus, altogether, the determinant gets multiplied by -1. $\square$

1.3. Computing determinants by row operations

We recall our old definition of row operations (from the 2019-09-30 classwork):

Definition 1.3.1. The following operations on a matrix A are called **elementary row operations** (short **EROs**):

- **ERO1:** Exchange two rows.
- **ERO2:** Scale a row by a nonzero constant.
- **ERO3:** Add a multiple of one row to another row. (That is, add $\lambda \operatorname{row}_i A$ to $\operatorname{row}_j A$ for $\lambda \in \mathbb{R}$ and $i \neq j$.)

Let us give these operations some more descriptive names:

- ERO1 will be called a **row swap**.
- ERO2 will be called a **row scaling** with scaling factor λ, where λ is the nonzero constant by which the row is being scaled.
- ERO3 will be called a **row addition**.

We can now tell what happens to the determinant of an $n \times n$-matrix when the matrix undergoes any of these row operations:

- If the matrix undergoes a row swap (i.e., ERO1), then its determinant gets multiplied by -1 (by Theorem 1.2.6).).
- If the matrix undergoes a row scaling (i.e., ERO2) with scaling factor λ, then its determinant gets multiplied by λ (by Theorem 1.1.5).
- If the matrix undergoes a row addition (i.e., ERO3), then its determinant does not change (by Corollary 1.2.5).

$$\xrightarrow{\text{add row 1 to row 2}} \begin{pmatrix} -c & -d \\ a & b \end{pmatrix} \xrightarrow{\text{scale row 1 by } -1} \begin{pmatrix} c & d \\ a & b \end{pmatrix}.$$

And indeed, the result is what you would get from the original matrix by swapping rows 1 and 2.

Corollary 1.3.2. Let A be an $n \times n$-matrix, and let B be a matrix obtained from A by a sequence of EROs. Assume that this sequence contains exactly k row swaps and exactly ℓ row scalings with scaling factors $\lambda_1, \lambda_2, \ldots, \lambda_\ell$. (We don't care how many row additions it contains.) Then,

$$\det B = (-1)^k \underbrace{\lambda_1 \lambda_2 \cdots \lambda_\ell}_{\substack{\text{this is understood} \\ \text{to mean 1 if } \ell = 0}} \det A.$$

Proof. The previous results show how each row operation affects the determinant:

- A row addition doesn't change it at all.
- A row swap multiplies it by -1.
- A row scaling with scaling factor λ multiplies it by λ.

Thus, our sequence of EROs that took A to B multiplies the determinant by $\underbrace{-1, -1, \ldots, -1}_{k \text{ times}}, \lambda_1, \lambda_2, \ldots, \lambda_\ell$. Thus, in total, it multiplies the determinant by $(-1)^k \lambda_1 \lambda_2 \cdots \lambda_\ell$. This proves Corollary 1.3.2. $\square$

Corollary 1.3.2 gives us a quick way of computing a determinant of a square matrix A: We perform row-reduction to bring A into RREF. The resulting RREF matrix is upper-triangular (indeed, you can easily check that any square matrix in RREF is upper-triangular[3]), so we can compute its determinant easily (by Theorem 1.1.2). Then, by Corollary 1.3.2, we get $\det A$. (We don't even need to go all the way to an RREF; usually we will obtain an upper-triangular matrix in the row-reduction process long before we obtain a RREF matrix.)

Example 1.3.3. Let us compute $\det A$ for $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$. Indeed, we row-reduce A as follows:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$
$$\xrightarrow{\text{add } -1 \cdot \text{row } 1 \text{ to row } 2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} \xrightarrow{\text{add } -1 \cdot \text{row } 1 \text{ to row } 3} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

[3] Actually, any square matrix in RREF is diagonal, which is better than upper-triangular.

$$\xrightarrow{\text{add } -1 \cdot \text{row } 1 \text{ to row } 4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} \xrightarrow{\text{add } -1 \cdot \text{row } 2 \text{ to row } 3} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix}$$
$$\xrightarrow{\text{add } -1 \cdot \text{row } 2 \text{ to row } 4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \xrightarrow{\text{add } -1 \cdot \text{row } 3 \text{ to row } 4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We have obtained a triangular matrix with $1, 1, 1, 1$ on the diagonal and thus determinant $= 1$. Since our row operations have included no row scalings and no row swaps, we thus conclude that $\det A = 1$.

Example 1.3.4. Let us compute $\det A$, where $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$. We row-reduce A as follows:

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix} \xrightarrow{\text{swap rows 1 and 3}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix we have obtained is upper-triangular with diagonal $1, 1, 1$, so it has determinant 1. Thus,

$$1 = (-1)^1 \cdot \det A,$$

where the $(-1)^1$ comes from the fact that we have used 1 row swap operation. Thus, solving this for $\det A$, we obtain $\det A = -1$.

Example 1.3.5. Let us compute $\det A$, where $A = \begin{pmatrix} 0 & 2 \\ 3 & 6 \end{pmatrix}$. Assume that you have forgotten the formula $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$. We row-reduce A as follows:

$$A = \begin{pmatrix} 0 & 2 \\ 3 & 6 \end{pmatrix} \xrightarrow{\text{swap rows 1 and 2}} \begin{pmatrix} 3 & 6 \\ 0 & 2 \end{pmatrix} \xrightarrow{\text{scale row 1 by } 1/3} \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$
$$\xrightarrow{\text{scale row 2 by } 1/2} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

We have obtained an upper-triangular matrix with diagonal $1, 1$, which therefore has determinant 1. Thus, Corollary 1.3.2 yields

$$1 = (-1)^1 \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \det A.$$

Solving this for $\det A$, we obtain

$$\det A = -6.$$

1.4. The determinant determines invertibility

As a consequence of this method for computing determinants using RREF, we can get a new criterion for invertibility of a square matrix (which we can add to our Inverse Matrix Theorem):

Theorem 1.4.1. Let A be an $n \times n$-matrix. Then, A is invertible if and only if $\det A \neq 0$.

Proof. $\Longrightarrow$: Assume that A is invertible. We must prove that $\det A \neq 0$.

We have assumed that A is invertible. Hence, the Invertible Matrix Theorem (more precisely, the implication **(k)** $\Longrightarrow$ **(a)** from Theorem 1.2.1 in the notes from 2019-10-16) shows that A can be row-reduced to I_n. In other words, we can obtain I_n from A by a sequence of row operations. As we have seen in the previous section, each row operation multiplies the determinant by a nonzero number (either 1 in the case of a row addition, or -1 in the case of a row swap, or λ if you are scaling a row by λ). Hence, $\det(I_n)$ is obtained from $\det A$ by multiplying with a bunch of nonzero numbers. Therefore, $\det(I_n) \neq 0$ if and only if $\det A \neq 0$. But I_n is an upper-triangular matrix with diagonal entries $1, 1, \ldots, 1$; therefore, Theorem 1.1.2 shows that $\det(I_n) = 1 \neq 0$. Hence, $\det A \neq 0$ (since $\det(I_n) \neq 0$ if and only if $\det A \neq 0$). This completes the proof of the "$\Longrightarrow$" direction of Theorem 1.4.1.

$\Longleftarrow$: Assume that $\det A \neq 0$. We must prove that A is invertible.

Let B be the RREF of A. Then, we can obtain B from A by a sequence of row operations. As we have seen in the previous section, each row operation multiplies the determinant by a nonzero number (either 1 in the case of a row addition, or -1 in the case of a row swap, or λ if you are scaling a row by λ). Hence, $\det B$ is obtained from $\det A$ by multiplying with a bunch of nonzero numbers. Therefore, $\det B \neq 0$ (since $\det A \neq 0$). From this, we can easily see that the RREF matrix B has a pivot in each column[4]. Hence, $B = I_n$ (by Lemma 1.1.14 **(b)** in the notes from 2019-10-07), since B is a square matrix. But A row-reduces to B; in other words, A row-reduces to I_n (since $B = I_n$). Therefore, the Invertible Matrix Theorem (more precisely, the implication **(a)** $\Longrightarrow$ **(k)** from Theorem 1.2.1 in the notes from 2019-10-16) shows that A is invertible. This proves the "$\Longleftarrow$" direction of Theorem 1.4.1. $\square$

Example 1.4.2. Given any numbers a, b, c, d, e, f, g, we claim that the matrix $\begin{pmatrix} 0 & 0 & 0 & g \\ 0 & 0 & 0 & f \\ 0 & 0 & 0 & e \\ a & b & c & d \end{pmatrix}$ is not invertible.

[4] *Proof.* Assume the contrary. Thus, the matrix B has a column without pivot. Hence, B has $< n$ pivots (since B has n columns, and at most one pivot per column). Therefore, the matrix B has a row without pivot (since B has n rows). This row must be a zero row (since any nonzero row contains a pivot). But Corollary 1.2.1 shows that any matrix with a zero row has determinant 0. Thus, $\det B = 0$. This contradicts $\det B \neq 0$. This contradiction shows that our assumption was false, qed.

To prove this, it suffices (by Theorem 1.4.1) to show that its determinant is 0. Why is its determinant 0 ?

One way to see this is the following: If we swap row 1 and row 4, then we obtain the matrix $\begin{pmatrix} a & b & c & d \\ 0 & 0 & 0 & f \\ 0 & 0 & 0 & e \\ 0 & 0 & 0 & g \end{pmatrix}$, which is upper-triangular with diagonal $a, 0, 0, g$ and thus (by Theorem 1.1.2) has determinant $a \cdot 0 \cdot 0 \cdot g = 0$. Thus, our original matrix must also have determinant 0 (because the row swap only multiplied the determinant by -1).

Another way to see this is the following: By the definition of the determinant,

$$\det \begin{pmatrix} 0 & 0 & 0 & g \\ 0 & 0 & 0 & f \\ 0 & 0 & 0 & e \\ a & b & c & d \end{pmatrix} = (\text{a sum of 24 products}).$$

Each product has 4 factors, using 1 entry from each row and 1 from each column. Thus, at least two of its 4 factors must be 0 (because only 1 of the 4 factors can come from row 4, and only one can come from column 4, but this means that the remaining 2 or more factors come neither from row 4 nor from column 4 and thus are 0). Hence, the product is 0. Thus,

$$\det \begin{pmatrix} 0 & 0 & 0 & g \\ 0 & 0 & 0 & f \\ 0 & 0 & 0 & e \\ a & b & c & d \end{pmatrix} = (\text{a sum of 24 products, each of which is } 0) = 0.$$

This shows that the matrix is not invertible.

1.5. The determinant of AB

The following theorem is one of the most important properties of determinants:

Theorem 1.5.1. We have $\det(AB) = \det A \cdot \det B$ for any two $n \times n$-matrices A and B.

See [Strickland, Theorem B.17] and [Lay, §3.2, Theorem 6] for two different proofs of this theorem.

Example 1.5.2. Let $C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$ again. Here is another way to find $\det C$.

Recall (from midterm 1 Exercise 1) that

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Hence, by Theorem 1.5.1, we have

$$\det C = \underbrace{\det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}}_{\substack{=1 \\ \text{(since this is a lower-triangular matrix} \\ \text{with diagonal entries } 1,1,\ldots,1)}} \cdot \underbrace{\det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\substack{=1 \\ \text{(since this is a upper-triangular matrix} \\ \text{with diagonal entries } 1,1,\ldots,1)}}$$
$$= 1.$$

Another application of Theorem 1.5.1 is the following fact:

Corollary 1.5.3. Let B be an $n \times n$-matrix. Let λ be a scalar. Then,

$$\det(\lambda B) = \lambda^n \det B.$$

Example 1.5.4. For $n = 3$, this is saying

$$\det \begin{pmatrix} \lambda a & \lambda b & \lambda c \\ \lambda d & \lambda e & \lambda f \\ \lambda g & \lambda h & \lambda i \end{pmatrix} = \lambda^3 \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

First proof of Corollary 1.5.3. We have

$$\begin{aligned} \det(\lambda B) &= \det(\lambda I_n \cdot B) \\ &= \det(\lambda I_n) \cdot \det B \qquad (\text{by Theorem 1.5.1, applied to } A = \lambda I_n). \end{aligned}$$

But $\lambda I_n = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix}$ is an upper-triangular (and lower-triangular, too) matrix with diagonal entries $\lambda, \lambda, \ldots, \lambda$. Hence, Theorem 1.1.2 says that $\det(\lambda I_n) = \underbrace{\lambda\lambda\cdots\lambda}_{n \text{ times}} = \lambda^n$. Thus, $\det(\lambda B) = \underbrace{\det(\lambda I_n)}_{=\lambda^n} \cdot \det B = \lambda^n \det B$. This proves Corollary 1.5.3. $\square$

Alternatively, we can prove Corollary 1.5.3 without using Theorem 1.5.1:

Second proof of Corollary 1.5.3. The matrix λB is obtained from the $n \times n$-matrix B by successively scaling **each** row by λ. Each time a row is scaled by λ, the determinant gets multiplied by λ (by Theorem 1.1.5). Hence, when going from B to λB, the determinant gets multiplied by λ^n (since there are n rows). This proves Corollary 1.5.3 as well. $\square$

1.6. Determinants and columns

Theorem 1.6.1. Theorem 1.1.4, Theorem 1.1.5, Corollary 1.2.1, Theorem 1.2.3, Corollary 1.2.5 and Theorem 1.2.6 remain valid if we replace "row" by "column" throughout them.

Roughly speaking, this means that everything we said above about determinants and rows holds all the same for determinants and columns. (For example: If you scale a column of an $n \times n$-matrix by a number λ, then the determinant gets multiplied by λ.)

Proof of Theorem 1.6.1 (rough idea). Let us prove the analogue of Theorem 1.1.5 for columns instead of rows:

Let A be an $n \times n$-matrix, and let λ be a number. Let B be the matrix if we scale a column of A (say, the i-th column) by λ. We must prove that $\det B = \lambda \cdot \det A$.

Theorem 1.1.3 yields $\det A = \det\left(A^T\right)$ and $\det B = \det\left(B^T\right)$. Recall that the matrix B is obtained from the matrix A by scaling a column by λ. Hence, its transpose B^T is obtained from A^T by scaling a **row** by λ (because the columns of a transpose of a matrix are the transposes of the rows of the matrix). Therefore, Theorem 1.1.5 yields $\det\left(B^T\right) = \lambda \cdot \det\left(A^T\right)$. In view of $\det A = \det\left(A^T\right)$ and $\det B = \det\left(B^T\right)$, this rewrites as $\det B = \lambda \cdot \det A$. This proves the analogue of Theorem 1.1.5 for columns instead of rows.

Similar arguments can be used to derive the column analogues of all other theorems from their original (row) versions. $\square$

Example 1.6.2. Here is an instance of the analogue of Theorem 1.2.3 for columns (with $n = 3$ and $k = 2$):

$$\det\begin{pmatrix} a & b+b' & c \\ d & e+e' & f \\ g & h+h' & i \end{pmatrix} = \det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} + \det\begin{pmatrix} a & b' & c \\ d & e' & f \\ g & h' & i \end{pmatrix}.$$

1.7. Laplace expansion

The following fact reduces the computation of the determinant of an $n \times n$-matrix to the computation of several smaller determinants:

Theorem 1.7.1 (Laplace expansion along the p-th row)**.** Let A be an $n \times n$-matrix. For each $p, q \in [n]$, we let $M_{p,q}$ be the $(n-1) \times (n-1)$-matrix obtained from A by removing row p and column q. Then, for each $p \in [n]$, we have

$$\det A = \sum_{q=1}^{n} (-1)^{p+q} A_{p,q} \det \left(M_{p,q} \right).$$

Example 1.7.2. Let $n = 3$. Then,

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{pmatrix}.$$

Thus, the $M_{p,q}$ in Theorem 1.7.1 are

$$M_{1,1} = \begin{pmatrix} A_{2,2} & A_{2,3} \\ A_{3,2} & A_{3,3} \end{pmatrix}, \qquad M_{1,2} = \begin{pmatrix} A_{2,1} & A_{2,3} \\ A_{3,1} & A_{3,3} \end{pmatrix}, \qquad M_{1,3} = \begin{pmatrix} A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{pmatrix},$$

$$M_{2,1} = \begin{pmatrix} A_{1,2} & A_{1,3} \\ A_{3,2} & A_{3,3} \end{pmatrix}, \qquad M_{2,2} = \begin{pmatrix} A_{1,1} & A_{1,3} \\ A_{3,1} & A_{3,3} \end{pmatrix}, \qquad M_{2,3} = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{3,1} & A_{3,2} \end{pmatrix},$$

$$M_{3,1} = \begin{pmatrix} A_{1,2} & A_{1,3} \\ A_{2,2} & A_{2,3} \end{pmatrix}, \qquad M_{3,2} = \begin{pmatrix} A_{1,1} & A_{1,3} \\ A_{2,1} & A_{2,3} \end{pmatrix}, \qquad M_{3,3} = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}.$$

Thus, Theorem 1.7.1 (applied to $p = 1$) yields

$$\begin{aligned}
&\det A \\
&= \sum_{q=1}^{3} (-1)^{1+q} A_{1,q} \det \left(M_{1,q} \right) \\
&= \underbrace{(-1)^{1+1}}_{=1} A_{1,1} \det \left(M_{1,1} \right) + \underbrace{(-1)^{1+2}}_{=-1} A_{1,2} \det \left(M_{1,2} \right) + \underbrace{(-1)^{1+3}}_{=1} A_{1,3} \det \left(M_{1,3} \right) \\
&= A_{1,1} \det \left(M_{1,1} \right) - A_{1,2} \det \left(M_{1,2} \right) + A_{1,3} \det \left(M_{1,3} \right) \\
&= A_{1,1} \det \begin{pmatrix} A_{2,2} & A_{2,3} \\ A_{3,2} & A_{3,3} \end{pmatrix} - A_{1,2} \det \begin{pmatrix} A_{2,1} & A_{2,3} \\ A_{3,1} & A_{3,3} \end{pmatrix} + A_{1,3} \det \begin{pmatrix} A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{pmatrix}.
\end{aligned}$$

Likewise, Theorem 1.7.1 (applied to $p = 2$) yields

$$\begin{aligned}
&\det A \\
&= -A_{2,1} \det \begin{pmatrix} A_{1,2} & A_{1,3} \\ A_{3,2} & A_{3,3} \end{pmatrix} + A_{2,2} \det \begin{pmatrix} A_{1,1} & A_{1,3} \\ A_{3,1} & A_{3,3} \end{pmatrix} - A_{2,3} \det \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{3,1} & A_{3,2} \end{pmatrix}.
\end{aligned}$$

For a proof of Theorem 1.7.1, see [Strickland, Proposition B.24].

Example 1.7.3. Let's compute $\det A$ for $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 3 & 0 & 4 \end{pmatrix}$. Theorem 1.7.1 (for $p = 2$) yields

$$\begin{aligned}
&\det A \\
&= -A_{2,1} \det \begin{pmatrix} A_{1,2} & A_{1,3} \\ A_{3,2} & A_{3,3} \end{pmatrix} + A_{2,2} \det \begin{pmatrix} A_{1,1} & A_{1,3} \\ A_{3,1} & A_{3,3} \end{pmatrix} - A_{2,3} \det \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{3,1} & A_{3,2} \end{pmatrix} \\
&= \underbrace{-0 \det \begin{pmatrix} 0 & 2 \\ 0 & 4 \end{pmatrix}}_{=0} + 1 \det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \underbrace{0 \det \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}}_{=0} \\
&= \det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2.
\end{aligned}$$

Applying Theorem 1.7.1 to some value of p is called "Laplace expansion with respect to the p-th row" or simply "expanding (the determinant) with respect to the p-th row". Generally, a good strategy when computing determinants is the following: If your matrix A has a row with many zeroes, you can try computing $\det A$ by Laplace expansion with respect to this row.

Example 1.7.4. Let us compute $\det A$, where $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$. Theorem 1.7.1 (for $p = 1$) yields

$$\begin{aligned}
\det A &= 1 \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} - \underbrace{0 \det \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}}_{=0} \\
&\qquad + 1 \det \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} - \underbrace{0 \det \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_{=0} \\
&= \underbrace{\det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}}_{\substack{=0 \\ \text{(since this matrix has two equal rows)}}} + \det \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\
&= \det \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}
\end{aligned}$$

$$= \underbrace{0 \det \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}_{=0} \underbrace{-1 \det \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}_{=1} +1 \underbrace{\det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}}_{\substack{=0 \\ \text{(due to two equal rows)}}}$$

(by Theorem 1.7.1 for $p = 1$)

$$= -1.$$

Exercise 1.7.1. Compute

$$\det\left((i+j)_{1\leq i\leq n,\ 1\leq j\leq n}\right) = \det \begin{pmatrix} 1+1 & 1+2 & \cdots & 1+n \\ 2+1 & 2+2 & \cdots & 2+n \\ \vdots & \vdots & \ddots & \vdots \\ n+1 & n+2 & \cdots & n+n \end{pmatrix}$$

$$= \det \begin{pmatrix} 2 & 3 & \cdots & n+1 \\ 3 & 4 & \cdots & n+2 \\ \vdots & \vdots & \ddots & \vdots \\ n+1 & n+2 & \cdots & 2n \end{pmatrix}.$$

You can see that this determinant is 2 if $n = 1$; is -1 if $n = 2$; is 0 if $n = 3$; if 0 if $n = 4$; can you spot and prove the general pattern? E.g., for $n = 5$, this is about $\begin{pmatrix} 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 8 & 9 \\ 6 & 7 & 8 & 9 & 10 \end{pmatrix}$.

References

[Lay] David C. Lay, Steven R. Lay, Judi J. McDonald, *Linear Algebra and its Applications*, 5th edition.

[Strickland] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 28 September 2013.
`http://neil-strickland.staff.shef.ac.uk/courses/MAS201/`