Honours Algebra - Week 6 - The Determinant of a Matrix

Antonio León Villares

February 2022

Contents

Based on the notes by Iain Gordon, Sections 4.1 - 4.4

1 The Sign of a Permutation

1.1 The Symmetric Group

- **What is the nth symmetric group?**
  - the group of **permutations** of n elements S_n
  - group under **composition**
  - has $n!$ elements
- **What is a transposition?**
  - a **permutation** which **only** swaps to elements:
  - for example, $(3\ 4) \in S_5$ represents the permutation which swaps 3 and 4, and leaves 1,2,5 unchanged

1.2 Theorem: Permutations as Products of Tranpositions

Any permutation:

$$(a_1\ a_2\ \dots\ a_n)$$

can be written as a ***product of transpositions****.*
In particular, 2 methods are:

$$(a_1\ a_2\ \dots\ a_n) = \prod_{i=2}^{n} (a_1\ a_i)$$

$$(a_1\ a_2\ \dots\ a_n) = \prod_{i=1}^{n-1} (a_i\ a_{i+1})$$

Proof. We prove by induction.

1 **Base Case**

Trivial for $(a_1\ a_2)$

2 **Inductive Hypothesis**

Assume true for $n = k$. In other words, any permutation of k elements can be written as a product of transpositions.

3 **Inductive Step**

Consider a permutation of $n = k + 1$ elements. We can use a single transposition to "place" a_{k+1}. Then, we have k elements left to place in the permutation, but by the inductive hypothesis, these can be written as a product of transpositions. Hence, a permutation of $k+1$ elements can be written as a product of transpositions.

Hence, by induction, any permutation can be expressed as a product of transpositions.

The specific examples provided can be easily proven by using an inductive argument.

$\square$

1.3 The Sign of a Permutation: Original Definition

- **What is the sign of a permutation?**
  - the **parity** of the number of transpositions required to express a permutation
  - symbolically, if $n(\sigma)$ is the number of transpositions used to build σ:

$$sgn(\sigma) = (-1)^{n(\sigma)}$$

- **What is an even permutation?**
  - a **permutation** with $sgn(\sigma) = 1$
  - in other words, a permutation which can be expressed as a product of **evenly** many transpositions
- **What is an odd permutation?**
  - a **permutation** with $sgn(\sigma) = -1$

1.4 The Sign of a Permutation: HAlg Definition

- **What is an inversion of a permutation?**
  - say $\sigma \in S_n$
  - an **inversion** is a tuple:

$$(i, j)$$

    such that:
    1. $1 \leq 1 < j \leq n$
    2. $\sigma(i) > \sigma(j)$

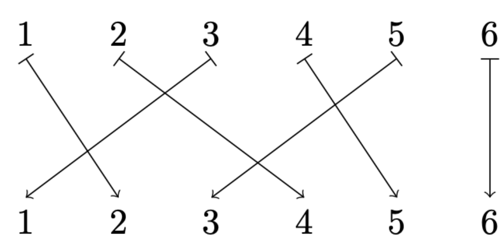


Figure 1: We can visualise the number of inversions by drawing the mappings. In particular, the number of inversions is given by the **number of crossings**. Intuitively this makes sense: if there is a cross, we have an arrow going from left to right (so $i < \sigma(i)$) and from right to left (so $\sigma(j) < j$) such that also $i < j$ and $\sigma(i) > \sigma(j)$, which is precisely the condition for an inversion.
In this diagram, we have that for example $(1, 3)$ is an inversion, since $1 \to 2$ and $3 \to 1$.

- **How do we define the length of a permutation?**

- the length of a permutation is the **number of inversions** of the permutation:

$$l(\sigma) = |\{(i,j) \mid i < j \ \wedge \ \sigma(i) > \sigma(j)\}|$$

- **What is an alternative way of defining the sign of a permutation?**
  - the sign can be defined as the **parity** of the number of inversions (**length of a permutation**):

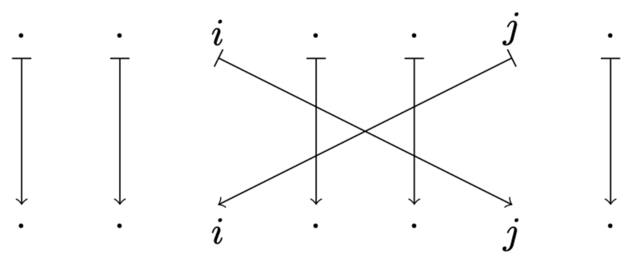
$$sgn(\sigma) = (-1)^{l(\sigma)}$$

1.4.1 Examples

- the **identity** is the only permutation with length 0
- a transposition swapping i, j has length:

$$2|i-j| - 1$$

  This is because i forms an inversion with each of $i+1, i+2, \ldots, j$. Similarly, j forms an inversion with each of $j-1, j-2, \ldots, i$. If we remove the duplicate inversion (i,j), we get the desired figure. This can be easily seen diagrammatically:



  Notice, this says that **transpositions** are **odd** permutations, which coincides with the original idea of sign.

1.5 Lemma: Multiplicativity of the Sign of a Permutation

> *For each $n \in \mathbb{N}$, the* ***sign*** *of a* ***permutation*** *produces a* ***group homomophism****:*
>
> $$sgn : S_n \to \{1, -1\}$$
>
> *In particular, it follows that:*
>
> $$sgn(\sigma\tau) = sgn(\sigma)sgn(\tau), \qquad \forall \sigma, \tau \in S_n$$

Proof. The proof in the notes is not nice or intuitive. I much prfer this one. We can decompose σ, τ into transpositions. Then, it is clear that $\sigma\tau$ can be decomposed into $n(\sigma) + n(\tau)$ transpositions, so:

$$sgn(\sigma\tau) = -1^{n(\sigma)+n(\tau)} = (-1)^{n(\sigma)}(-1)^{n(\tau)} = sgn(\sigma)sgn(\tau)$$

as required.

$\square$

1.6 The Alternating Group

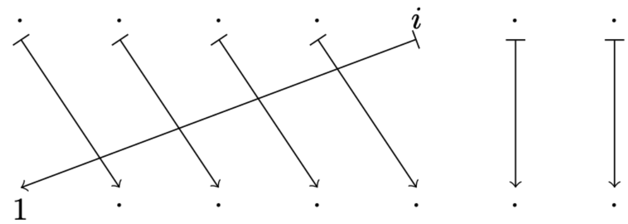
- **What is the alternating group?**
  - a **subgroup** of S_n
  - contains all **even** permutations of S_n, and is denoted A_n
  - its a **subgroup**, since A_n is the **kernel** of the **group homomorphism**:

    $$sgn : S_n \to \{1, -1\}$$

    (since 1 is the identity of $\{1, -1\}$, and only even permutations get mapped there)

1.6.1 Exercises (TODO)

1. **Show that the permutation mapping a_i to a_1, and with $a_j \to a_{j+1}, j \in [1, i-1]$ has $i-1$ inversions:**



2 Defining the Determinant

2.1 Leibniz Formula

- **What is the Leibniz formula for the determinant of a matrix?**
  - the **determinant** is a mapping:

    $$det : Mat(n; R) \to R$$

    where R is a **ring**
  - the **determinant** is computed using the **Leibniz Formula**:

    $$\sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{1\sigma(i)}$$

    In other words, it sums over all possible products of permutations of the diagonal elements of the matrix
  - for an "empty matrix" ($n = 0$), the determinant is 0
- **What does the determinant tell us about its corresponding linear transformation?**
  - if we have a region L which gets mapped to U under a linear transformation A, then:

    $$area(U) = det(A)area(L)$$

    That is, the **determinant** is an **area scaling factor**
  - the **sign** of the **determinant** indicates whether the linear transformation **preserves** or **inverts** orientation
  - you can better understand this by playing with this applet

2.1.1 Examples

- if $n = 1$:

$$A = \begin{pmatrix} a \end{pmatrix} \implies det(A) = a$$

- if $n = 2$:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies det(A) = ab - cd$$

  (there are only 2 permutations: the identity and a transposition)

- for $n = 3$ there are 6 terms: 3 positive and 3 negative, corresponding to the 3 even and 3 odd permutations of S_3.

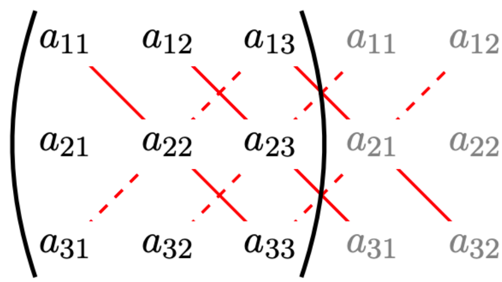


Figure 2: We can use this “trick” to compute the determinant: we multiply along the lines, and add the products; bold lines are positive, dashed lines are negative

- the determinant of **diagonal**, **upper triangular** and **bottom triangular** matrices is the product of the diagonal entries.
  - for upper triangular matrices, notice that:

$$a_{ij} = \begin{cases} 0, & i > j \\ *, & j \geq i \end{cases}$$

  - notice, for the determinant, each summand considers:

$$\prod_{i=1} a_{i\sigma(i)}$$

  - this is non-zero **if and only if**:

$$\sigma(i) \geq i, \qquad \forall i \in [1, n]$$

  - the only permutation which ensures this is the identity permutation; otherwise, we will always have at least one term which leads to $\sigma(i) < i$, in which case the product becomes 0
  - hence,

$$det(A) = \prod_{i=1}^{n} a_{ii}$$

    as required

2.1.2 Exercises (TODO)

1. **Show that the determinant of a block-upper triangular matrix with square blocks along the diagonal is the product of the determinants of the blocks along the diagonal:**

$$\mathsf{det}\left(\begin{array}{c|c|c|c} A_1 & * & * & * \\ \hline 0 & A_2 & * & * \\ \hline 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & A_t \end{array}\right) = \mathsf{det}(A_1)\mathsf{det}(A_2)\cdots\mathsf{det}(A_t)$$

A proof can be found here. It employs induction to prove a simple case, and then shows the general case.

3 Determinants as Multilinear Forms

We now discuss multilinear forms. They are rather abstract, and seem unrelated to determinants, but they provide an alternative way of **characterising** determinants and their properties, beyond the standard definitions.

3.1 Bilinear Forms

- **What is a bilinear form?**
  - a **mapping**:
  $$H : U \times V \to W$$
  where U, V, W are **F-Vector Spaces** (formally, a **bilinear form on** $U \times V$ **with values in** W)
  - it is **bilinear** because it is a **linear mapping** in both entries:
  $$H(u_1 + u_2, v) = H(u_1, v) + H(u_2, v)$$
  $$H(\lambda u, v) = \lambda H(u, v)$$
  $$H(u, v_1 + 1_2) = H(u, v_1) + H(u, v_2)$$
  $$H(u, \lambda v) = \lambda H(u, v)$$
- **When is a bilinear form symmetric?**
  - when $U = V$ and:
  $$H(u, v) = H(v, u), \qquad \forall u, v \in U$$
- **When is a bilinear form antisymmetric/alternating?**
  - when $U = V$ and:
  $$H(u, u) = 0$$

3.2 Remark: Alternating Bilinear Forms

*If H is an **alternating bilinear form**, then:*

$$H(u, v) = -H(v, u)$$

*If H is a **bilinear form** and*

$$H(u, v) = -H(v, u)$$

then:

$$H(u, u) = 0 \iff 1_F + 1_F \neq 0_F$$

*In other words, such a **bilinear form** is **alternating** if and only if $1_F + 1_F \neq 0_F$. [Remark 4.3.2]*

Proof. The first part is clear. If H is **alternating**:

$$\begin{aligned} & H(u+v, u+v) = 0 \\ \implies & H(u, u+v) + H(v, u+v) = 0 \\ \implies & H(u, v) + H(u, u) + H(v, u) + H(v, v) = 0 \\ \implies & H(u, v) + H(v, u) = 0 \\ \implies & H(u, v) = -H(v, u) \end{aligned}$$

If H is a **bilinear form** and $H(u, v) = -H(v, u)$, in particular:

$$H(u, u) = -H(u, u) \implies H(u, u) + H(u, u) = 0$$

We will have $H(u, u) = 0$ if and only if $1_F + 1_F \neq 0$. This can happen, for example, with $\mathbb{F} = \mathbb{F}_2 = \mathbb{Z}_2$ □

3.3 Multilinear Forms

- **How are multilinear forms defined?**
  - **multilinear forms** generalise **bilinear forms**
  - given **F-vector spaces** $V_1, \ldots, V_n, W$, a **multilinear form** is a mapping:

    $$H : V_1 \times \ldots \times V_n \to W$$

  - it is a **linear mapping** in each entry; in other words:

    $$V_j \to W$$

    $$v_j \to H(v_1, \ldots, v_j, \ldots, v_n)$$

    is a linear mapping (here the $v_i, i \neq j$ are fixed)
- **When is a multilinear form alternating?**
  - whenever we have $v_i = v_j, \qquad i \neq j$ and:

    $$H(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = 0$$

  - in other words, the mapping **vanishes** if it has (at least) 2 equal entries

3.4 Remark: Alternating Multilinear Forms

*If H is an **alternating multilinear form**, then:*

$$H(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -H(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

*In other words, if we swap 2 entries in an **alternating multilinear form**, we **negate** the value of the mapping.*
Conversely it H is a multilinear map, and

$$H(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -H(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

*then H is **alternating** if and only if:*

$$1_F + 1_F \neq 0_F$$

*More generally, if σ is a **permutation**:*

$$H(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = sgn(\sigma)H(v_1, \ldots, v_n)$$

[Remark 4.3.5]

Proof. The first one is similar as in the case for bilinear forms.

The second one follows from the fact that every permutation can be written as a **product of transpositions**. Hence, applying σ can be viewed as applying many consecutive transpositions ($n(\sigma)$ of them), from which we see the result.

□

3.5 Theorem: Characterisation of the Determinant

*Let F be a **field**. The mapping:*

$$det : Mat(n; F) \to F$$

*is the **unique alternating multilinear form** on n-tuples of **column vectors** with values in F, and which takes value 1_F on the **identity matrix**.*
*Notice, we treat elements in $Mat(n; F)$ as both **matrices** over F, and as an **ordered list** of **column vectors** (namely the **matrix columns)**, such that:*

$$det : F^n \times \times \ldots \times F^n \to F$$

$$(\underline{v}_1, \ldots, \underline{v}_n) \to det(Mat(\underline{v}_1, \ldots, \underline{v}_n))$$

[Theorem 4.3.6]

Proof. 1. **The Determinant is Multilinear** This is pretty intuitive if we use the Leibniz formula, but here is an example for the 2×2 case

2. **The Determinant Evaluates to 1_F on the Identity Matrix** The identity matrix is a diagonal matrix with diagonal entries 1_F, so its determinant is the product of these entreis, which is 1_F.

3. **The Determinant is Alternating** Assume $\underline{v}_i = \underline{v}_j$. In particular, we must have that:

$$a_{ki} = a_{kj}$$

for any row k.

Now, let $\tau \in S_n$ be the transposition which switches $\underline{v}_i$ and $\underline{v}_j$. Then:

$$a_{ki} = a_{kj} \ \wedge \ a_{kj} = a_{k\tau(i)} \implies a_{ki} = a_{k\tau(i)}$$

But then, for any $\sigma \in S_n$, we must have that:

$$\prod_{i=1}^{n} a_{i\sigma(i)} = \prod_{i=1}^{n} a_{i\tau\sigma(i)}$$

By multiplicity of the sign:

$$sgn(\tau\sigma) = sgn(\tau)sgn(\sigma) = -sgn(\sigma)$$

since $sgn(\tau)$ is a transposition, and so $sgn(\tau) = -1$.

Furthermore, the subgroup of S_n generated by τ is:

$$H = \{id_{S_n}, \tau\}$$

and since cosets of subgroups partition a group (since they define equivalence classes; see here for more), we must have that, if X is the set of right coset representatives of H:

$$\bigcup_{\sigma \in X} H\sigma = S_n$$

where each $H\sigma$ is disjoint. In other words, each $x \in X$ generates 2 (unique) elements in H, namely x and τx. We can now put this together. By Leibniz:

$$det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{1\sigma(i)}$$

Instead of iterating through S_n, we can iterate through the set of representatives X, and then include the elements in S_n generated by each representative:

$$det(A) = \sum_{x \in X} \left(sgn(x) \prod_{i=1}^{n} a_{1x(i)} + sgn(\tau x) \prod_{i=1}^{n} a_{1\tau x(i)} \right)$$

But recall from above that $sgn(\tau x) = -sgn(x)$, and

$$\prod_{i=1}^{n} a_{ix(i)} = \prod_{i=1}^{n} a_{i\tau x(i)}$$

so it follows that:

$$det(A) = \sum_{x \in X} \left(sgn(x) \prod_{i=1}^{n} a_{1x(i)} - sgn(x) \prod_{i=1}^{n} a_{1x(i)} \right) = 0$$

Hence, det is **alternating**.

Notice, this can be extended to show that a square matrix with coefficients in a **commutative ring** has $det(A) = 0$ whenever 2 columns are equal.

4. **The Determinant is a Unique Such Mapping** As we have seen before (Lemma 1.7.8), linear mappings are completely determined by the values they take on a basis, so we only need to check the values of mappings on the basis elements.

   Assume there exists some other mapping:

   $$d : Mat(n; F) \to F$$

   with the properties of the theorem (multilinear form, alternating, maps identity to 1_F).

   We consider the value of:

   $$d(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)}))$$

   where $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ (since we don't care how each of the basis vectors are organised within the matrix).

   If $\sigma(i) = \sigma(j)$, since d is alternating, we must have that:

   $$d(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)})) = 0 = det(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)}))$$

   Thus, if σ is **not** bijective (in other words, $\sigma \notin S_n$), $d(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)})) = 0$. Otherwise, if $\sigma \in S_n$, then:

   $$d(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)})) = sgn(\sigma) d(Mat(e_1, \dots, e_n))$$

   since d is a multilinear form. Now notice, by assumption, we must have that:

   $$d(Mat(e_1, \dots, e_n)) = 1$$

   so if $\sigma \in S_n$, then:

   $$d(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)})) = sgn(\sigma)$$

   But notice, again if $\sigma \in S_n$ and using the multilinearity of the determinant:

   $$det(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)})) = sgn(\sigma) d(Mat(e_1, \dots, e_n)) = sgn(\sigma)$$

   So it follows that:

   $$d(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)})) = det(Mat(e_{\sigma(1)}, \dots, e_{\sigma(n)}))$$

   as required.

□

3.5.1 Exercises (TODO)

1. **Adapt the argument above to show that if:**

   $$d : Mat(n; F) \to F$$

   is an alternating multilinear form on n-tuples of column vectors with values in F, then:

   $$d(A) = d(Mat(e_1, \dots, e_n)) det(A), \qquad \forall A \in Mat(n; F)$$

4 Calculating With Determinants

4.1 Theorem: Multiplicativity of the Determinant

Let R be a ***commutative ring****, and let $A, B \in R$. Then:*

$$det(AB) = det(A)det(B)$$

[Theorem 4.4.1]

Proof. Recall, when multiplying 2 matrices together, entry $(AB)_{ik}$ is given by:

$$(AB)_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}$$

Let I_n be the set of all mappings from $\{1, \ldots, n\}$ to itself.

From definition:

$$\begin{aligned}
det(AB) &= \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} (AB)_{i\sigma(i)} \\
&= \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} \sum_{j=1}^{n} a_{ij}b_{j\sigma(i)} \\
&= \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} (a_{i1}b_{1\sigma(i)} + a_{i2}b_{2\sigma(i)} + \ldots + a_{in}b_{n\sigma(i)})
\end{aligned}$$

Now, think about the expression above. For example, with $n = 2$:

$$\begin{aligned}
\prod_{i=1}^{2} \sum_{j=1}^{2} a_{ij}b_{j\sigma(i)} &= \prod_{i=1}^{n} (a_{i1}b_{1\sigma(i)} + a_{i2}b_{2\sigma(i)}) \\
&= (a_{11}b_{1\sigma(1)} + a_{12}b_{2\sigma(1)}) \times (a_{21}b_{1\sigma(2)} + a_{22}b_{2\sigma(2)}) \\
&= a_{11}b_{1\sigma(1)}a_{21}b_{1\sigma(2)} + a_{11}b_{1\sigma(1)}a_{22}b_{2\sigma(2)} + a_{12}b_{2\sigma(1)}a_{21}b_{1\sigma(2)} + a_{12}b_{2\sigma(1)}a_{22}b_{2\sigma(2)}
\end{aligned}$$

But notice, each term can be characterised by an element of I_n. For example:

$$\kappa_1(x) = \begin{cases} 1, & x = 1 \\ 1, & x = 2 \end{cases} \implies a_{11}b_{1\sigma(1)}a_{21}b_{1\sigma(2)} = a_{1\kappa_1(1)}b_{\kappa_1(1)\sigma(1)}a_{2\kappa_1(2)}b_{\kappa_1(2)\sigma(2)}$$

$$\kappa_2(x) = \begin{cases} 1, & x = 1 \\ 2, & x = 2 \end{cases} \implies a_{11}b_{1\sigma(1)}a_{22}b_{2\sigma(2)} = a_{1\kappa_2(1)}b_{\kappa_2(1)\sigma(1)}a_{2\kappa_2(2)}b_{\kappa_2(2)\sigma(2)}$$

$$\kappa_3(x) = \begin{cases} 2, & x = 1 \\ 1, & x = 2 \end{cases} \implies a_{12}b_{2\sigma(1)}a_{21}b_{1\sigma(2)} = a_{1\kappa_3(1)}b_{\kappa_3(1)\sigma(1)}a_{2\kappa_3(2)}b_{\kappa_3(2)\sigma(2)}$$

$$\kappa_4(x) = \begin{cases} 2, & x = 1 \\ 2, & x = 2 \end{cases} \implies a_{12}b_{2\sigma(1)}a_{22}b_{2\sigma(2)} = a_{1\kappa_4(1)}b_{\kappa_4(1)\sigma(1)}a_{2\kappa_4(2)}b_{\kappa_4(2)\sigma(2)}$$

Hence, we can succintly write:

$$\prod_{i=1}^{2}\sum_{j=1}^{2} a_{ij}b_{j\sigma(i)} = \sum_{\kappa\in I_2}\prod_{i=1}^{2} a_{i\kappa(i)}b_{\kappa(i)\sigma(i)}$$

Thus, generalising in the above:

$$\begin{aligned}
det(AB) &= \sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^{n}(AB)_{i\sigma(i)} \\
&= \sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{j\sigma(i)} \\
&= \sum_{\sigma\in S_n} sgn(\sigma)\sum_{\kappa\in I_n}\prod_{i=1}^{n} a_{i\kappa(i)}b_{\kappa(i)\sigma(i)} \\
&= \sum_{\sigma\in S_n} sgn(\sigma)\sum_{\kappa\in I_n}\left(\prod_{i=1}^{n} a_{i\kappa(i)}\right)\left(\prod_{i=1}^{n} b_{\kappa(i)\sigma(i)}\right) \\
&= \sum_{\kappa\in I_n}\sum_{\sigma\in S_n} sgn(\sigma)\left(\prod_{i=1}^{n} a_{i\kappa(i)}\right)\left(\prod_{i=1}^{n} b_{\kappa(i)\sigma(i)}\right) \\
&= \sum_{\kappa\in I_n}\left(\prod_{i=1}^{n} a_{i\kappa(i)}\right)\sum_{\sigma\in S_n} sgn(\sigma)\left(\prod_{i=1}^{n} b_{\kappa(i)\sigma(i)}\right)
\end{aligned}$$

Let B_κ be the matrix obtained from shuffling its rows by using κ (so $b_{\kappa(i)}$ is its ith row). Furthermore, notice that:

$$det(B_\kappa) = \sum_{\sigma\in S_n} sgn(\sigma)\left(\prod_{i=1}^{n} b_{\kappa(i)\sigma(i)}\right)$$

If $\kappa \notin S_n$, we will have the $det(B_\kappa) = 0$ (where B_κ is the matrix resulting from applying κ to each of the rows of B), since we will have at least 2 identical rows. Furthermore, if $\kappa \in S_n$, we know from the multilinearity of the determinant that:

$$det(B_\kappa) = sgn(\kappa)det(B)$$

Thus:

$$\begin{aligned}
det(AB) &= \sum_{\kappa\in I_n}\left(\prod_{i=1}^{n} a_{i\kappa(i)}\right)\sum_{\sigma\in S_n} sgn(\sigma)\left(\prod_{i=1}^{n} b_{\kappa(i)\sigma(i)}\right) \\
&= \sum_{\kappa\in I_n}\left(\prod_{i=1}^{n} a_{i\kappa(i)}\right)det(B_k) \\
&= \sum_{\kappa\in S_n}\left(\prod_{i=1}^{n} a_{i\kappa(i)}\right)sgn(\kappa)det(B), \qquad \textit{(since if } \kappa\notin S_n \textit{ we have } det(B_k)\textit{, so terms in sum vanish)} \\
&= \left(\sum_{\kappa\in S_n} sgn(\kappa)\prod_{i=1}^{n} a_{i\kappa(i)}\right)det(B) \\
&= det(A)det(B)
\end{aligned}$$

as required.

□

4.2 Theorem: Determinantal Criterion for Invertibility

> *The **determinant** of a **square matrix** with entries in a field F is non-zero **if and only if** the matrix is **invertible**. [Theorem 4.4.2]*

Proof. 1. **Matrix is Invertible**

If A is invertible, then:

$$\exists B : AB = I_n$$

By multiplicativity of determinant:

$$det(A)det(B) = 1$$

Since $det(A), det(B) \in F$, this is only possible if $det(A) \neq 0$, since fields are **integral domains**

2. **Matrix is not Invertible**

A non-invertible matrix in particular won't have full rank, so, without loss of generality, we can write the first column vector of A as a **linear combination** of the other column vectors. That is:

$$a_{*1} = \sum_{i=2}^{n} \lambda_i a_{*i}, \lambda_i \in F$$

Then, we can exploit the multilinearity and alternating properties of the determinant:

$$\begin{aligned} det(A) &= det(Mat(\sum_{i=2}^{n} \lambda_i a_{*i}, a_{*2}, \ldots, a_{*n}) \\ &= \sum_{i=2}^{n} \lambda_i det(Mat(a_{*i}, a_{*2}, \ldots, a_{*n}) \\ &= \sum_{i=2}^{n} \lambda_i 0 \\ &= 0 \end{aligned}$$

Where we use the fact that det is alternating, and so 0 whenever there is a repeated entry.

□

4.3 Remark: Determinant and Similar Matrices

> *From the Theorem above, it is clear that:*
>
> $$det(A^{-1}) = det(A)^{-1}$$
>
> *By multiplicativity of determinants, and since we are working over **commutative rings**, it thus follows that:*
>
> $$det(A^{-1}BA) = det(A^{-1})det(B)det(A) = det(B)$$
>
> *[Remark 4.4.3]*

4.4 Lemma: Determinant of the Transpose

*If $A \in Mat(n; R)$, and R is a **commutative ring**, then:*

$$det(A^T) = det(A)$$

[Lemma 4.4.4]

Proof. From definition:

$$det(A^T) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$$

Now, if $\tau = \sigma^{-1}$, then:

$$sgn(\tau) = sgn(\sigma)$$

(the inverse of a transposition is itself, so the inverse of σ will be composed of the same number of transpositions, just "reflected" in their order)

Moreover, since we operate over a **commutative ring**, we must have that:

$$\prod_{i=1}^{n} a_{\sigma(i)i} = \prod_{i=1}^{n} a_{i\tau(i)}$$

Thus:

$$det(A^T) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} = \sum_{\tau \in S_n} sgn(\tau) \prod_{i=1}^{n} a_{i\tau(i)} = det(A)$$

□

4.4.1 Exercises (TODO)

(1) Let

$$V = \left(\lambda_j^{i-1}\right) = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & & \lambda_n^2 \\ \vdots & & & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_3^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix}$$

where $\lambda_i \neq \lambda_j$. Calculate $|V|$.

(2) Let

$$C = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{pmatrix}.$$

Calculate $a(\lambda) = |\lambda 1_n - C|$.

(3) Suppose that $a(\lambda) = \prod_{i=1}^n (\lambda - \lambda_i)$ for distinct roots λ_i. Calculate $V^{-1}CV$. Deduce that

$$(VV^T)^{-1}C(VV^T) = C^T.$$

(4) Let B be an $n \times n$ matrix with eigenvalues $\lambda_1, \dots, \lambda_n$ and define a $n \times n$ matrix $\hat{B} = (\operatorname{Tr} B^{i+j-2})$ (with rows i and columns j). Verify that

$$VV^T = \hat{B} = \begin{pmatrix} n & \operatorname{Tr} B & \cdots & \operatorname{Tr} B^{n-1} \\ \operatorname{Tr} B & \operatorname{Tr} B^2 & \cdots & \operatorname{Tr} B^n \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{Tr} B^{n-1} & \operatorname{Tr} B^n & \cdots & \operatorname{Tr} B^{2n-2} \end{pmatrix}.$$

and hence deduce that $|\hat{B}| = \prod_{i<j} (\lambda_j - \lambda_i)^2$.

4.5 ILA Definition of Determinants: The Cofactor

- **What is the cofactor of a matrix?**
  - let $A \in Mat(n; R)$, where R is a **commutative ring**
  - the (i, j) **cofactor** of A is:

    $$C_{ij} = (-1)^{i+j} det(A\langle i, j\rangle$$

    where $A\langle i, j\rangle$ is the matrix obtained by removing row i and column j of A

$$C_{23} = (-1)^{2+3}\text{det}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = -a_{11}a_{32} + a_{31}a_{12}$$

4.6 Theorem: Laplace's Expansion of the Determinant

*Let $A = (a_{ij}) \in Mat(n; R)$, where R is a **commutative ring**.*
*For a **fixed** i the i**th row expansion of the determinant** is:*

$$det(A) = \sum_{j=1}^{n} a_{ij} C_{ij}$$

*For a **fixed** j the j**th column expansion of the determinant** is:*

$$det(A) = \sum_{i=1}^{n} a_{ij} C_{ij}$$

[Theorem 4.4.7]

Proof. Since $det(A) = det(A^T)$, it is sufficient to only prove the column expansion. Moreover, moving the jth column to the first position (as in (1.6.1)) is the same as applying the permutation:

$$\sigma = (1\ j)(12)(23)\dots(j-1\ j)^1$$

so it will change the determinant by a factor of $sgn(\sigma) = (-1)^{j-1}$.

Thus, it is sufficient to show that $det(A) = \sum_{i=1}^{n} a_{ij} C_{ij}$ for expansion along the first column, $j = 1$.

Say we have:

$$A = Mat(a_{*1}, \dots, a_{*n})$$

We write the first column as a linear combination of basis vectors:

$$a_{*1} = \sum_{i=1}^{n} a_{i1} e_i$$

We can then apply multilinearity of the determinant:

$$det(A) = det(Mat(a_{*1}, \dots, a_{*n})) = \sum_{i=1}^{n} a_{i1} det(Mat(e_i, \dots, a_{*n}))$$

Notice, if we move the ith row of $Mat(e_i, \dots, a_{*n})$ to the first row, we will obtain the matrix:

$$\left(\begin{array}{c|c} 1 & * \\ \hline 0 & A\langle i, j\rangle \end{array}\right)$$

($Mat(a_{*1}, \dots, a_{*n})$ is A without the $j = 1$ column, and moving the ith row is equivalent to removing the ith row of A) In doing this, we will change the value of the determinant by a factor of $(-1)^{i-1}$

Now recall the exercise in which we show that the determinant of a block-upper triangular matrix is the product of the determinants of the matrices in the main diagonal. In other words:

$$det(Mat(e_i, \dots, a_{*n})) = (-1)^{i-1} det(A\langle i, j\rangle) = C_{i1}$$

[1] When writing this I cam up with this permutation on the spot, and I'm pretty proud of that yeet

Thus, as required, if we expand along $j = 1$:

$$det(A) = \sum_{i=1}^{n} a_{i1} C_{i1}$$

If we do this for an arbitrary j, we first need to move the jth column to the first column, so we would get:

$$\begin{aligned} det(Mat(e_i, \ldots, a_{*n})) &= (-1)^{j-1}(-1)^{i-1} det(A\langle i, j\rangle) \\ &= (-1)^{i+j-2} det(A\langle i, j\rangle) \\ &= (-1)^{i+j}(-1)^{-2} det(A\langle i, j\rangle) \\ &= (-1)^{i+j} det(A\langle i, j\rangle) \\ &= (-1)^{i+j} C_{ij} \end{aligned}$$

□

4.7 Defining the Adjugate Matrix

- **What is an adjugate matrix?**
  - let $A \in Mat(n; R)$, where R is a comuutative ring
  - the **adjugate matrix** is:

$$adj(A) \in Mat(n; R) \qquad adj(A)_{ij} = C_{ji}$$

4.8 Theorem: Cramer's Rule

Let $A \in Mat(n; R)$, where R is a ***commutative ring****.*
Then:

$$A \cdot adj(A) = (det(A)) I_n$$

Proof. From the matrix product formula, the ik entry of $A \cdot adj(A)$ is:

$$\sum_{j=1}^{n} a_{ij} adj(A)_{jk}$$

Hence, we need to show that:

$$\sum_{j=1}^{n} a_{ij} adj(A)_{jk} = \delta_{ik} det(A)$$

But $adj(A)_{jk} = C_{kj}$ so we require:

$$\sum_{j=1}^{n} a_{ij} C_{kj} = \delta_{ik} det(A)$$

Thereare 2 cases to consider:

1. $\boldsymbol{i = k}$ Then, $\delta_{ik} = 1$, so we require:

$$\sum_{j=1}^{n} a_{ij} C_{ij} = det(A)$$

which is nothing but the i**th row expansion of the determinant**, so it is correct.

2. $\boldsymbol{i \neq k}$ Now define the matrix $\hat{A}$, which is identical to A, except for the fact that the kth row of $\hat{A}$ is the same as the ith row of A. In other words, each entry $\hat{a}_{kj}$ is given by a_{ij}.

   Then, we can compute the determinant of $\hat{A}$ using the kth row expansion:

$$det(\hat{A}) = \sum_{j=1}^{n} \hat{a}_{kj} C_{kj} = \sum_{j=1}^{n} a_{ij} C_{kj}$$

   But notice, $\sum_{j=1}^{n} a_{ij} C_{kj} = \delta_{ik} det(A)$, so we need to show that:

$$det(\hat{A}) = \delta_{ik} det(A) = 0$$

   since $\delta_{ik} = 0$, as $i \neq k$. But this is true, since $\hat{A}$ has rows i and k equal, so by the alternating property of the determinant, $det(\hat{A}) = 0$, as required.

□

4.9 Remark: Cramer's Rule to Solve Linear Equations

> ***Cramer's Rule*** *can also be stated in the context of solving a linear system:*
>
> $$A\underline{x} = \underline{b}$$
>
> *where:*
>
> $$x_i = \frac{det(Mat(a_{*1}, \ldots, \underline{b}, \ldots, a_{*n}))}{det(A)}$$

4.10 Corollary: Cramer's Rule and the Invertibility of Matrices

> *$A \in Mat(n; R)$, where R is a* ***commutative ring*** *is invertible* ***if and only if****:*
>
> $$det(A) \in R^{\times}$$
>
> *That is, $det(A)$ must be a unit in R (so it has a* ***multiplicative inverse*** *in R). For instance, matrices over $\mathbb{Z}$ will be invertible only when $det(A) = 1, -1$, whilst matrices over fields will be invertible whenever $det(A) \neq 0$ (since every element in a field has a multiplicative inverse except 0). [Corollary 4.4.11]*

Proof. 1. A **is Invertible** Then, $\exists B \in Mat(n; R)$ such that:

$$AB = I_n \implies det(A)det(B) = 1_R$$

Hence, $det(A)$ must be a **unit** in R.

2. $det(A)$ **is a Unit in** R Recall, we need to show the existence of 2 matrices $B, C \in Mat(n; R)$ such that:

$$AB = CA = I_n$$

In the first case, if we have $\hat{B} = adj(A)$, then **Cramer's Rule** says:

$$A\hat{B} = (det(A))I_n$$

Since $det(A)$ is a unit, it has an inverse, so:

$$A(det(A)^{-1}\hat{B}) = I_n$$

Thus, setting $B = det(A)^{-1}\hat{B}$ satisfies the first condition.

Since $det(A^T) = det(A)$, then $det(A^T)$ must also be a unit. Again applying Cramer's Rule with $\hat{C} = adj(A^T)$:

$$A^T\hat{C} = (det(A^T))I_n \implies A^T(det(A)^{-1}\hat{C}) = I_n$$

If we then take the transpose:

$$(det(A)^{-1}\hat{C}^T)A = I_n$$

Hence, setting $C = det(A)^{-1}\hat{C}^T$ satisfies the second condition.

$\square$

5 Workshop

1. **True or false. Let R be an integral domain and let $A \in Mat(n, R)$ be a matrix with non-zero determinant. Then A is invertible,**

This is false. By Corollary 4.4.11:

> *$A \in Mat(n; R)$, where R is a **commutative ring** is invertible **if and only if**:*
>
> $$det(A) \in R^\times$$
>
> *That is, $det(A)$ must be a unit in R (so it has a **multiplicative inverse** in R). For instance, matrices over $\mathbb{Z}$ will be invertible only when $det(A) = 1, -1$, whilst matrices over fields will be invertible whenever $det(A) \neq 0$ (since every element in a field has a multiplicative inverse except 0). [Corollary 4.4.11]*

Hence, it is sufficient to find an integral domain R, such that $det(A) \notin R^\times$. Picking $R = \mathbb{Z}$, then $R^\times = \{-1, +1\}$. Consider the matrix:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Then, $det(A) = 2$ so clearly $det(A) \notin R^\times$. We can confirm that $A^{-1} \notin Mat(2, R)$ since:

$$A^{-1} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

2. **Let:**

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 2 & 1 \end{pmatrix}$$

   (a) **Write π as a product of disjoint cycles.**
   We get:
   $$\pi = (1\ 4\ 2\ 5)$$

   (b) **Write each nontrivial disjoint cycle of π as a product of transpositions.**
   We get:
   $$(1\ 5)(1\ 2)(1\ 4)$$

   (c) **Write each transposition in the previous part as a product of transpositions of the form $(i, i+1)$.**

   > *This is definitely not trivial. The key is to exploit the fact that a transposition is its own inverse.*

   We can write:
   $$(1\ 5) = (4\ 5)(3\ 4)(2\ 3)(1\ 2)(2\ 3)(3\ 4)(4\ 5)$$
   This ensures that if a 5 goes in, we "cascade" down the transposition chain, until we reach (1 2), which is the only transposition with a 1, and so returns 1. Alternatively, if 1 goes in, we "cascade" up the transposition chain, and return 5. All other numbers will get mapped to themselves.
   We can write:
   $$(1\ 4) = (3\ 4)(2\ 3)(1\ 2)(2\ 3)(3\ 4)$$
   Hence, we have that:
   $$\pi = (4\ 5)(3\ 4)(2\ 3)(1\ 2)(2\ 3)(3\ 4)(4\ 5)(1\ 2)(3\ 4)(2\ 3)(1\ 2)(2\ 3)(3\ 4)$$

3. (a) **Evaluate the following determinant:**

$$\Delta_n := \begin{vmatrix} 0 & x_1 & x_2 & \dots & x_{n-1} \\ y_1 & 1 & 0 & \dots & 0 \\ y_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{n-1} & 0 & 0 & \dots & 1 \end{vmatrix}_{n\times n}$$

   We claim that:
   $$\Delta_n = -\sum_{i=1}^{n-1} x_i y_i$$
   We work by induction.
   1 **Base Case:** $n = 1$

   We see that trivially $\Delta_1 = 0 = -\sum_{i=1}^{0} x_i y_i$.
   2 **Inductive Hypothesis:** $n = k$

Assume true for $n = k$. Then:

$$\Delta_k = -\sum_{i=1}^{k-1} x_i y_i$$

3 **Inductive Step:** $n = k + 1$

We compute Δ_{k+1}:

$$\Delta_{k+1} := \begin{vmatrix} 0 & x_1 & x_2 & \dots & x_k \\ y_1 & 1 & 0 & \dots & 0 \\ y_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_k & 0 & 0 & \dots & 1 \end{vmatrix}_{(k+1)\times(k+1)}$$

If we expand along the last row, we see that:

$$\Delta_{k+1} = (-1)^{k+1+1} y_k \begin{vmatrix} x_1 & x_2 & \dots & x_k \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{vmatrix}_{k\times k} + \Delta_k$$

Furthermore:

$$\begin{vmatrix} x_1 & x_2 & \dots & x_k \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{vmatrix}_{k\times k} = (-1)^{k+1} x_k det(I_k) = (-1)^{k+1} x_k$$

Hence, we have that:

$$\Delta_{k+1} = (-1)^{k+1+1} y_k (-1)^{k+1} x_k - \sum_{i=1}^{k-1} x_i y_i = (-1)^{2k+3} y_k x_k - \sum_{i=1}^{k-1} x_i y_i = -\sum_{i=1}^{k} x_i y_i$$

as required.

(b) **Let $A = (a_1, \dots, a_m) \in Mat(n \times m; F), B = (b_1, \dots, b_m) \in Mat(n \times m; F)$ where $a_i, b_j \in F^n$. If $n > m$, what is $det(AB^T)$?**

Notice,

$$im(AB^T) \subseteq im(A)$$

since $im(AB^T)$ is just the image of A corresponding to vectors of the form $B^T \underline{v}$. This means that:

$$rank(AB^T) \leq rank(A)$$

Moreover, since $n > m$, we must have that:

$$rank(A) \leq m$$

In particular, this means that:

$$rank(AB^T) \leq m$$

But notice, AB^T is a $n \times n$ matrix, so if $rank(AB^T) \leq m < n$, then AB^T has linearly dependent rows. In particular, this means that:

$$det(AB^T) = 0$$

(recall, the determinant is a bilinear form, so rows being equal tells us that the determinant is 0)

(c) **Let $a_i \neq 0 \in \mathbb{R}$ with $i \in [0, n]$. Prove that:**

$$a_n + \cfrac{1}{a_{n-1} + \cfrac{1}{\ldots + \cfrac{1}{a_1 + \cfrac{1}{a_0}}}} = \frac{\Delta_n}{\Delta_{n-1}}$$

where:

$$\Delta_n = \begin{vmatrix} a_0 & 1 & 0 & \ldots & 0 & 0 \\ -1 & a_1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & a_{n-1} & 1 \\ 0 & 0 & 0 & \ldots & -1 & a_n \end{vmatrix}_{(n+1)\times(n+1)}$$

Again, we proceed by induction.

1 **Base Case:** $n = 0$

The result follows trivially.

2 **Inductive Hypothesis:** $n = k$

Assume that:

$$a_k + \cfrac{1}{a_{k-1} + \cfrac{1}{\ldots + \cfrac{1}{a_1 + \cfrac{1}{a_0}}}} = \frac{\Delta_k}{\Delta_{k-1}}$$

3 **Inductive Step:** $n = k + 1$

We compute $\frac{\Delta_{k+1}}{\Delta_k}$. Indeed, we expand along the last row:

$$\Delta_{k+1} = \begin{vmatrix} a_0 & 1 & 0 & \ldots & 0 \\ -1 & a_1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix}_{(k+1)\times(k+1)} + a_{k+1}\Delta_k$$

Again, if we expand along the last row:

$$\begin{vmatrix} a_0 & 1 & 0 & \dots & 0 \\ -1 & a_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}_{(k+1)\times(k+1)} = \Delta_{k-1}$$

so we get that:

$$\Delta_{k+1} = a_{k+1}\Delta_k + \Delta_{k-1}$$

Dividing through by Δ_k:

$$\frac{\Delta_{k+1}}{\Delta_k} = a_{k+1} + \frac{\Delta_{k-1}}{\Delta_k} = a_{k+1} + \frac{1}{\frac{\Delta_k}{\Delta_{k-1}}} = a_{k+1} + \cfrac{1}{a_k + \cfrac{1}{a_{k-1} + \cfrac{1}{\dots + \cfrac{1}{a_1 + \cfrac{1}{a_0}}}}}$$

as required.

4. **Given the linear equation:**

$$A\underline{x} = \underline{b}$$

where:

$$A = (\underline{a}_1, \dots, \underline{v}_n) \in Mat(n;F) \qquad \underline{x} = (x_1, \dots, x_n)^T \qquad \underline{b} = (b_1, \dots, b_n)^T$$

we set:

$$A_i = (\underline{a}_1, \dots, \underline{b}, \dots, \underline{a}_n)$$

as the matrix A but with the ith column changed to $\underline{b}$. Show that:

$$x_i = \frac{|A_i|}{|A|}$$

Define I_i as the matrix obtained by changing the ith column of the identity matrix by $\underline{x}$. Then:

$$AI_i = \begin{pmatrix} A\underline{e}_1 & \dots & A\underline{x} & \dots & A\underline{e}_n \end{pmatrix} = A_i$$

Moreover, I_i is a diagonal matrix, so:

$$det(I_i) = x_i$$

Hence:

$$AI_i = A_i \implies |A|x_i = |A_i|$$

so if $|A| \neq 0$ then:

$$x_i = \frac{|A_i|}{|A|}$$