Warmup:

1. Let $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for some $a, b, c, d \in F$.
   (a) Row reduce $\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right)$ to compute X^{-1} for a general 2×2 matrix. Keep track of when you might be accidentally dividing by 0.
   (b) What does it mean if you can't row reduce $(X|I)$ without dividing by 0? Illustrate with an example.
2. Let $X, Y \in M_\ell(F)$. We say Y is **similar** or **conjugate** to X, written $Y \sim X$, if $Y = PXP^{-1}$ for some invertible $P \in M_\ell(F)$. (i.e. Y represents the same linear function as X, just with respect to a different ordered basis—see below for review of context for this problem.)
   (a) Prove that $\sim$ defined an equivalence relation on $M_\ell(F)$.
   (b) What is the equivalence class of I_ℓ (with respect to this relation)?
   (c) Show that if $Y \sim X$, then $Y^k \sim X^k$. [More specifically, that the *same* change of basis that moves you from X to Y will also move you from X^k to Y^k.]
   (d) For $Y \sim X$ and $Y' \sim X'$, is it true that $XX' \sim YY'$?

Context for Problem 2:

Last time, we saw that, in $M_\ell(F)$,

{ invertible matrices } is the same as { change of basis matrices }.

Specifically, let V be a vector space over F, and let $\mathcal{A}$ and $\mathcal{B}$ be ordered bases of V. Then $\mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(\mathrm{id}) = (\mathrm{Rep}_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}))^{-1}$, so that change of basis matrices are all invertible. Conversely, the columns of any invertible matrix is a basis of the corresponding space, so can be viewed as a change of basis matrix. And since

$$\mathrm{Rep}_{\mathcal{B}}^{\mathcal{B}}(f) = \mathrm{Rep}_{\mathcal{A}}^{\mathcal{B}}(\mathrm{id})\mathrm{Rep}_{\mathcal{A}}^{\mathcal{A}}(f)(\mathrm{Rep}_{\mathcal{B}}^{\mathcal{A}}(\mathrm{id}))^{-1},$$

matrix **conjugation** is the same as changing basis (two matrices represent the same function w.r.t. different bases exactly when they are conjugate).

1. Let $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for some $a, b, c, d \in F$.

(a) Compute X^{-1}:

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right)$$

(b) What does it mean if you can't row reduce $(X|I)$ without dividing by 0?

2. (a) "Similarity" defines an equivalence relation on the set $M_\ell(F)$:

$$Y \sim X \qquad \text{whenever} \qquad Y = PXP^{-1}$$

for some invertible $P \in M_\ell(F)$:

- Reflexive: If $X \in M_\ell(F)$, then. . .
- Symmetric: If $X, Y \in M_\ell(F)$ satisfy $X \sim Y$, then. . .
- Transitive: If $X, Y, Z \in M_\ell(F)$ satisfy $X \sim Y$ and $Y \sim Z$, then. . .

(b) If $Y \sim I_\ell$, then for some invertible $P \in M_\ell(F)$, we have

$$Y = PI_\ell P^{-1} =$$

We'll be studying statistics about matrices that are invariant under change of basis, meaning that they're constant on similar matrices—these statistics are important because they pertain to the underlying functions independent of your choice of basis. **Homework:** Trace.

Determinants

The **determinant**, $\det : M_n(F) \to F$, is one *extremely* important statistic about square matrices.

- *Geometrically*, it will measure the volume of a polygon generated by the rows of a matrix.
- *Algebraically*, it will satisfy lots of nice algebraic properties, and will determine whether a matrix is invertible or not for us.

We will also see that it is independent of change of basis, and is therefore *actually* a statistic pertaining to the underlying linear function. We begin by defining it by its desirable properties, and then show that these properties uniquely identify the function.

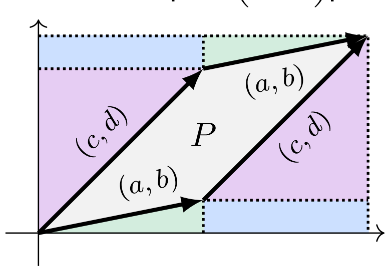
Example: Define

$$\det : M_2(\mathbb{R}) \to \mathbb{R} \quad \text{by} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc.$$

We saw that X is invertible if and only if $\det(X) \neq 0$.

Some other properties:

- The parallelogram P with corners $(0,0)$, (a,b), (c,d), and $(a,b)+(c,d)$ has area $\left|\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right|$:



$$\begin{aligned} \text{Area}(P) &= (a+c)(b+d) - 2(bc) - (cd) - (ab) \\ &= ad - bc. \end{aligned}$$

- **Normalized:** $\det(I_2) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot 1 - 0 \cdot 0 = 1.$
- **Multiplicative:**

$$\begin{aligned} \det(XY) = \det\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}\right) &= \det \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix} \\ &= (aa' + bc')(cb' + dd') - (ab' + bd')(ca' + dc') \\ &= \cdots = (ad - bc)(a'd' - b'c') = \det(X)\det(Y). \end{aligned}$$

Example: Define

$$\det : M_2(\mathbb{R}) \to \mathbb{R} \quad \text{by} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ad - bc.$$

We saw that X is invertible if and only if $\det(X) \neq 0$.

Some other properties:

- **Alternating:**

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xmapsto{\mathbf{row}_1 \leftrightarrow \mathbf{row}_2} \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

$$\det(X) = ad - bc \qquad \det(P_{1,2}X) = bc - ad = -\det(X)$$

[Check via multiplicativity: $\det(P_{1,2}) = \det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1$. ✓]

- **Multilinear**, i.e. linear row-by-row:

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

[Check via multiplicativity: $\det(S_1(\lambda)) = \det\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} = \lambda$. ✓]

$$\det\begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c & d \end{pmatrix} = (a_1 + a_2)d - (b_1 + b_2)c$$
$$= (a_1 d - b_1 c) + (a_2 d - b_2 c).$$ (and similar in row 2).

Generalizing: Think of a function $\det : M_n(F) \to F$ as a multivariable function of the row vectors of a matrix, meaning:

If the matrix X has row vectors $\mathbf{r}_1, \ldots, \mathbf{r}_n$, write $\det(X) = \det(\mathbf{r}_1, \ldots, \mathbf{r}_n)$.

A "determinant" is any function $\det : M_n(F) \to F$ that satisfies the following:

(1) **Normalized.** $\det(I_n) = \det(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$.

(2) **Alternating.** Swapping any two rows toggles the sign of the function:

$$\det(\ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots) = -\det(\ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots).$$

(3) **Multilinear.** The determinant is a linear function with respect to every row (individually): for each $i = 1, \ldots, n$, we have

$$\det(\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_i + \lambda\,\mathbf{s}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_n) = \det(\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_n) + \lambda \det(\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{s}_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_n).$$

Consequences:

- If $\mathbf{r}_i = \mathbf{r}_j$, then by (2),

$$\det(\ldots, \mathbf{r}_i, \ldots, \mathbf{r}_i, \ldots) = -\det(\ldots, \mathbf{r}_i, \ldots, \mathbf{r}_i, \ldots),$$

so $\det(\ldots, \mathbf{r}_i, \ldots, \mathbf{r}_i, \ldots) = 0$.

- If X & Y differ by a row combination, then their determinants are equal:

$$\det(\ldots, \mathbf{r}_i + \lambda\mathbf{r}_j, \ldots, \mathbf{r}_j, \ldots) = \det(\ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots) + \lambda \underbrace{\det(\ldots, \mathbf{r}_j, \ldots, \mathbf{r}_j, \ldots)}_{0}.$$

- So if X and Y differ by a sequence of row operations, then $\det(X) = \mu \det(Y)$ for some $0 \neq \mu \in F$.

Let $\det : M_n(F) \to F$ be a determinant, i.e. a function that, as a multivariable function in its row vectors, is normalized, alternating, and multilinear. Then if X and Y differ by a sequence of row operations, we have

$$\det(X) = \mu \det(Y) \text{ for some } 0 \neq \mu \in F.$$

Example.

$$A = \begin{pmatrix} 1 & 3 & 9 & 27 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\mathbf{r}_1 \leftrightarrow \mathbf{r}_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \xrightarrow{\mathbf{r}_3 \mapsto \mathbf{r}_3 - \mathbf{r}_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix}$$

$$\xrightarrow{\mathbf{r}_4 \mapsto \mathbf{r}_4 - \mathbf{r}_1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 4 & 8 \\ 0 & 3 & 9 & 27 \end{pmatrix} \xrightarrow{\mathbf{r}_3 \mapsto \mathbf{r}_3 - 2\mathbf{r}_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 6 \\ 0 & 3 & 9 & 27 \end{pmatrix} \xrightarrow{\mathbf{r}_4 \mapsto \mathbf{r}_4 - 3\mathbf{r}_2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 6 & 24 \end{pmatrix}$$

$$\xrightarrow{\mathbf{r}_3 \mapsto \frac{1}{2}\mathbf{r}_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 6 & 24 \end{pmatrix} \xrightarrow{\mathbf{r}_4 \mapsto \frac{1}{6}\mathbf{r}_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow{\mathbf{r}_2 \mapsto \mathbf{r}_2 - \mathbf{r}_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

$$\xrightarrow{\mathbf{r}_4 \mapsto \mathbf{r}_4 - \mathbf{r}_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\mathbf{r}_2 \mapsto \mathbf{r}_2 + 2\mathbf{r}_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\mathbf{r}_3 \mapsto \mathbf{r}_3 - 3\mathbf{r}_4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\det : 1 \cdot 1 \cdot 1 \cdot 1 \cdot \tfrac{1}{6} \cdot \tfrac{1}{2} \cdot 1 \cdot 1 \cdot 1 \cdot 1 \cdot (-1) \det(A) = \det(I_4) = 1. \quad \text{So } \boxed{\det(A) = -12}.$$

Examples. Each of the elementary row reduction matrices is one row reduction away from I_ℓ, so computing their determinant is simple.

$$P_{i,j} \xrightarrow{\mathbf{r}_i \leftrightarrow \mathbf{r}_j} I_\ell, \quad \text{so}$$
$$\det(P_{i,j}) = -\det(I_\ell) = -1 \qquad \text{(because } \det \text{ is alternating)}$$

$$S_i(\lambda) \xrightarrow{\mathbf{r}_i \mapsto \frac{1}{\lambda}\mathbf{r}_i} I_\ell, \quad \text{so}$$
$$\det(S_i(\lambda)) = \lambda \det(I_\ell) = \lambda \qquad \text{(because } \det \text{ is multilinear)}$$

$$C_{i,j}(\lambda) \xrightarrow{\mathbf{r}_j \mapsto \mathbf{r}_j - \lambda \mathbf{r}_i} I_\ell, \quad \text{so}$$
$$\det(C_{i,j}(\lambda)) = \det(I_\ell) = 1 \qquad \text{(see "consequences")}$$

Example: If

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ 0 & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{n,n} \end{pmatrix},$$

then either...

(i) $x_{i,i} = 0$ for some i: then row reducing produces a matrix with at least one row of 0's at the end,

$$X \mapsto \cdots \mapsto E = \begin{pmatrix} \text{---} & \mathbf{r}_1 & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{r}_{n-1} & \text{---} \\ \text{---} & \mathbf{0} & \text{---} \end{pmatrix}.$$

Then $\det(E) = \det(\mathbf{r}_1, \ldots, \mathbf{r}_{n-1}, \mathbf{0}) = \det(\mathbf{r}_1, \ldots, \mathbf{r}_{n-1}, 0 \cdot \mathbf{0})$

$= 0 \det(\mathbf{r}_1, \ldots, \mathbf{r}_{n-1}, \mathbf{0}) = 0$.

So $\det(X) = \mu \det(E) = \mu \cdot 0 = 0$. Or,

(ii) $x_{i,i} \neq 0$ for all i: Then

$$\begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ 0 & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{n,n} \end{pmatrix} \xmapsto{\mathbf{r}_1 \mapsto (1/x_{1,1})\mathbf{r}_1} \begin{pmatrix} 1 & x_{1,2}/x_{1,1} & \cdots & x_{1,n}/x_{1,1} \\ 0 & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{n,n} \end{pmatrix},$$

$$\xmapsto{\mathbf{r}_2 \mapsto (1/x_{2,2})\mathbf{r}_2} \cdots \xmapsto{\mathbf{r}_n \mapsto (1/x_{n,n})\mathbf{r}_n} \begin{pmatrix} 1 & x_{1,2}/x_{1,1} & \cdots & x_{1,n}/x_{1,1} \\ 0 & 1 & \cdots & x_{2,n}/x_{2,2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

$$\xmapsto{\text{sequence of row combinations}} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n.$$

Either way, $\boxed{\det(X) = x_{1,1} x_{2,2} \cdots x_{n,n}}$.

Let $\det : M_n(F) \to F$ be a determinant, i.e. a function that, as a multivariable function in its row vectors, is normalized, alternating, and multilinear. Then if X and Y differ by a sequence of row operations, we have

$$\det(X) = \mu \det(Y) \text{ for some } 0 \neq \mu \in F.$$

Lemma. If X has reduced row echelon form E, then

$$\det(X) = \mu \det(E) \text{ for some } 0 \neq \mu \in F,$$

where μ is determined by the row operations moving from X to E:

$$\mu = (-1)^{\#\text{row swaps}} \left(\prod_{\substack{\text{scaling operations} \\ S_i(\lambda)}} (1/\lambda) \right).$$

Lemma. If $E \in M_n(F)$ has a row of 0's, then $\det(E) = 0$.

Pf. $\det(\mathbf{r}_1, \ldots, \mathbf{r}_{n-1}, \mathbf{0}) = \det(\mathbf{r}_1, \ldots, \mathbf{r}_{n-1}, 0 \cdot \mathbf{0}) = 0 \det(\mathbf{r}_1, \ldots, \mathbf{r}_{n-1}, \mathbf{0}) = 0.$

Proposition. If $\det$ exists, it's unique.

Given existence. . .

AMAZING Theorem. For $X \in M_n(F)$, we have

$$\det(X) \neq 0 \quad \text{if and only if} \quad X \text{ is invertible.}$$

Ok, great. But **does such a function even exist?**

For instance, what if there's more than one sequence of row operations that gets me from X to E? Do I get different answers? (Spoiler: no)

Next time/homework:

- More properties of determinants (if they exists), like what happens to products, inverses, and transposes, and what they mean geometrically.

  Caution: For now, **do not assume** $\det$ **is multiplicative for** $n \geqslant 3$**.** (This has to be proven.)

- The existence of determinant for all n.

- The *symmetric group* and a formula for determinant in terms of *permutations*.

For now, if we take for granted that determinant is well-defined, we can already calculate the determinant of a matrix by row reducing.

Practice exercises:

Chapter Four, Section I: 2.8, 2.11, 2.12, 2.15, 2.18, 2.22

Note: The book uses the notation $\det(A) = |A|$, and shorthands

$$\det \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{pmatrix} = \begin{vmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{vmatrix}.$$

This is common (but not uniformly standard), but it's generally better to use notation that describes what it is.