Math 313 Lecture #14
§3.3: Cramer's Rule, Volume, and Linear Transformations

We make use of the multiplicative property $\det(AB) = \det(A)\det(B)$ to derive Cramer's Rule for solving $A\vec{x} = \vec{b}$ when A is invertible for its unique solution $\vec{x} = A^{-1}\vec{b}$ without finding A^{-1}.

Cramer's Rule is also useful for theoretical purposes (such as in Math 334).

To set the stage for Cramer's Rule, we let $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$ be an $n \times n$ matrix.

For $\vec{b}$ in $\mathbb{R}^n$ we form a new matrix $A_i(\vec{b})$ determined by $n - 1$ columns of A defined by

$$A_i(\vec{b}) = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{b} & \cdots & \vec{a}_n \end{bmatrix}$$

where $\vec{b}$ replaces the i^{th} column of A.

Theorem 7 (Cramer's Rule). Let A be an invertible $n \times n$ matrix. For any $\vec{b}$ in $\mathbb{R}^n$, the unique solution of $A\vec{x} = \vec{b}$ has entries

$$x_i = \frac{\det(A_i(\vec{b}))}{\det(A)}, i = 1, 2, \ldots, n.$$

Proof. For $I = \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \cdots & \vec{e}_n \end{bmatrix}$ we have for any $i = 1, 2, \ldots, n$ that

$$\begin{aligned} AI_i(\vec{x}) &= A \begin{bmatrix} \vec{e}_1 & \cdots & \vec{x} & \cdots & \vec{e}_n \end{bmatrix} \\ &= \begin{bmatrix} \vec{a}_1 & \cdots & A\vec{x} & \cdots & \vec{a}_n \end{bmatrix} \\ &= \begin{bmatrix} \vec{a}_1 & \cdots & \vec{b} & \cdots & \vec{a}_n \end{bmatrix} \\ &= A_i(\vec{b}). \end{aligned}$$

By the multiplicative property of determinants applied to $AI_i(\vec{x}) = A_i(\vec{b})$ we have

$$\det(A)\det(I_i(\vec{x})) = \det(A_i(\vec{b})).$$

The value of $\det(I_i(\vec{x}))$ is x_i by cofactor expansion along the i^{th} row.

Because $\det(A) \neq 0$, we can solve for x_i, giving the desired formula. $\square$

Example. Use Cramer's Rule to find the solution of

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}.$$

First we compute the determinant of the matrix:

$$\begin{vmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{vmatrix} = (1) \begin{vmatrix} 2 & 1 \\ 2 & 0 \end{vmatrix} + (2) \begin{vmatrix} 0 & 2 \\ 3 & 2 \end{vmatrix} = -14.$$

Then we compute the determinants of $A_1(\vec{b})$, $A_2(\vec{b})$, and $A_3(\vec{b})$:

$$\det(A_1(\vec{b})) = \begin{vmatrix} -3 & 0 & 2 \\ -1 & 2 & 1 \\ 1 & 2 & 0 \end{vmatrix} = (1)\begin{vmatrix} 0 & 2 \\ 2 & 1 \end{vmatrix} - (2)\begin{vmatrix} -3 & 2 \\ -1 & 1 \end{vmatrix} = -4 - 2(-1) = -2,$$

$$\det(A_2(\vec{b})) = \begin{vmatrix} 1 & -3 & 2 \\ 0 & -1 & 1 \\ 3 & 1 & 0 \end{vmatrix} = (-1)\begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} - (1)\begin{vmatrix} 1 & -3 \\ 3 & 1 \end{vmatrix} = -(-6) - 10 = -4,$$

$$\det(A_3(\vec{b})) = \begin{vmatrix} 1 & 0 & -3 \\ 0 & 2 & -1 \\ 3 & 2 & 1 \end{vmatrix} = (1)\begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix} + (-3)\begin{vmatrix} 0 & 2 \\ 3 & 2 \end{vmatrix} = 4 - 3(-6) = 22.$$

Thus the entries of the unique solution $\vec{x} = A^{-1}\vec{b}$ are

$$x_1 = \frac{-2}{-14} = \frac{1}{7}, \quad x_2 = \frac{-4}{-14} = \frac{2}{7}, \quad x_3 = \frac{22}{-14} = -\frac{11}{7}.$$

You should check that this is indeed the solution. ✓

A Formula for the Inverse. We apply Cramer's Rule to find a formula for the inverse of an invertible matrix.

In doing this we will encounter a matrix $\text{adj}(A)$, called the **adjugate** of A, whose entries are the cofactors $C_{ij} = (-1)^{i+j}\det(A_{ij})$ in *transposed* position:

$$\text{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \ldots & C_{nn} \end{bmatrix}.$$

Theorem 8. If A is an invertible $n \times n$ matrix, then

$$A^{-1} = \frac{1}{\det(A)}\,\text{adj}(A).$$

Proof. For an invertible A we have $AA^{-1} = I = \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \cdots & \vec{e}_n \end{bmatrix}$.

If $\vec{x}$ is the j^{th} column of A^{-1}, then $\vec{x}$ satisfies $A\vec{x} = \vec{e}_j$.

By Cramer's Rule, the i^{th} entry of $\vec{x}$, which is the (i, j) entry of A^{-1}, is given by

$$x_i = \frac{\det(A_i(\vec{e}_j))}{\det(A)}.$$

Recall that A_{ji} is the submatrix of A obtained by deleting the j^{th} row and the i^{th} column of A.

Observe that the non-zero entry of the i^{th} column of $A_i(\vec{e}_j)$ occurs on the j^{th} row.

Cofactor expansion down the i^{th} column of $A_i(\vec{e}_j)$ gives

$$\det(A_i(\vec{e}_j)) = (-1)^{j+i}\det(A_{ji}) = C_{ji}.$$

Thus the (i, j) entry of A^{-1} is C_{ji} divided by $\det(A)$. □

Example. Use this new formula to find the inverse of

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{bmatrix}.$$

Recall that the determinant of the matrix A is $\det(A) = -14$.

We compute the entries of the adjugate of A:

$$\text{adj } A = \begin{bmatrix} \begin{vmatrix} 2 & 1 \\ 2 & 0 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 3 & 0 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ 3 & 2 \end{vmatrix} \\ -\begin{vmatrix} 0 & 2 \\ 2 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 3 & 2 \end{vmatrix} \\ \begin{vmatrix} 0 & 2 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} \end{bmatrix}^T = \begin{bmatrix} -2 & 3 & -6 \\ 4 & -6 & -2 \\ -4 & -1 & 2 \end{bmatrix}^T = \begin{bmatrix} -2 & 4 & -4 \\ 3 & -6 & -1 \\ -6 & -2 & 2 \end{bmatrix}.$$

We then compute the inverse:

$$A^{-1} = \frac{1}{\det(A)} \text{ adj}(A) = \frac{1}{-14} \begin{bmatrix} -2 & 4 & -4 \\ 3 & -6 & -1 \\ -6 & -2 & 2 \end{bmatrix}.$$

We can check this:

$$A\left(\frac{1}{\det(A)} \text{ adj}(A)\right) = \frac{-1}{14} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} -2 & 4 & -4 \\ 3 & -6 & -1 \\ -6 & -2 & 2 \end{bmatrix} = \frac{-1}{14} \begin{bmatrix} -14 & 0 & 0 \\ 0 & -14 & 0 \\ 0 & 0 & -14 \end{bmatrix} \checkmark.$$

Determinants as Area or Volume. When A is a 2×2 matrix, it columns determine a parallelogram in the plane, and when A is 3×3, its columns determine a parallelepiped in $\mathbb{R}^3$.

We give a geometric interpretation of the absolute value of the determinant of A as the area of the parallelogram or the volume of the parallelepiped.

Theorem 9. For a square matrix A, the value $|\det(A)|$ is (a) the area of the parallelogram determined by its columns when A is 2×2, and (b) the volume of the parallelepiped determined by its columns when A is 3×3.

Proof. (a) We present an argument here that differs from the book.

Let $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix}$ with $\vec{a}_1 \neq 0$ (for if $\vec{a}_1 = 0$, then the parallelogram with sides $\vec{a}_1$ and $\vec{a}_2$ has area 0 and $\det(A) = 0$.

Now $\vec{a}_1$ makes an angle θ with the positive horizontal axis (as measured in the counterclockwise sense).

Rotate both $\vec{a}_1$ and $\vec{a}_2$ by the same angle so that $\vec{a}_1$ now points in the positive horizontal direction.

That is we apply the matrix

$$R(\theta) = \begin{bmatrix} \cos(2\pi - \theta) & -\sin(2\pi - \theta) \\ \sin(2\pi - \theta) & \cos(2\pi - \theta) \end{bmatrix}$$

to each column of A to get

$$\begin{bmatrix} R(\theta)\vec{a}_1 & R(\theta)\vec{a}_2 \end{bmatrix} = R(\theta)A.$$

Since $\det(R(\theta)A) = \det(R(\theta))\det(A)$ and $\det(R(\theta) = \cos^2(2\pi - \theta) + \sin^2(2\pi - \theta) = 1$, then $\det(R(\theta)A) = \det(A)$, which is to say that rotating the sides $\vec{a}_1$ and $\vec{a}_2$ of the parallelogram by the same angle does not change the area of the parallelogram.

So WLOG we may assume that

$$\vec{a}_1 = \begin{bmatrix} a \\ 0 \end{bmatrix}.$$

The line L be the line through $\vec{0}$ and $\vec{a}_1$ is the horizontal axis.

Then $\vec{a}_2 + L$ is the line through $\vec{a}_2$ and parallel to L.

The vector $\vec{a}_2 + c\vec{a}_1$ lies on the line $\vec{a}_2 + L$ for any scalar c.

The parallelogram with sides $\vec{a}_1$ and $\vec{a}_2$ has the same area as the parallelogram with sides $\vec{a}_1$ and $\vec{a}_2 + c\vec{a}_1$ because both have the same base length and because the points $\vec{a}_2$ and $\vec{a}_2 + c\vec{a}_1$ on $\vec{a}_2 + L$ have the same height (or perpendicular distance to L).

In other words, adding a multiple of one column to another column of A does not change the determinant of A.

We choose c so that $\vec{a}_2 + c\vec{a}_1$ has the form

$$\begin{bmatrix} 0 \\ d \end{bmatrix}.$$

Thus we have

$$|\det(A)| = \left| \det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \right| = |ad|$$

which is the area of the rectangle with base length $|a|$ and height $|d|$.

(b) This argument is left for you to read in the text. $\square$

Determinants and Linear Transformations

The unit rectangle U in $\mathbb{R}^2$ has sides $\vec{e}_1$ and $\vec{e}_2$; it is the set

$$U = \{(u_1\vec{e}_1 + u_2\vec{e}_2 : 0 \leq u_1, u_2 \leq 1\},$$

and it has area 1.

If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is linear, what is the area of $T(U)$?

The image of each side $\vec{e}_i$ under T is $T(\vec{e}_i)$, which are the columns of the standard matrix A of T, and so the area of $T(U)$ is precisely $|\det(A)|$.

How does T affect the area of other parallelograms?

Theorem 10. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation with its standard 2×2 matrix A. If S is a parallelogram in $\mathbb{R}^2$ with sides $\vec{b}_1$ and $\vec{b}_2$, then the area of $T(S)$ is $|\det(A)|$ times the area of S. A similar statement holds for a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ for all $n \geq 2$.

Proof. Write $A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) \end{bmatrix} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix}$.

The parallelogram S with sides $\vec{b}_1$ and $\vec{b}_2$ is the set

$$S = \{s_1\vec{b}_1 + s_2\vec{b}_2 : 0 \leq s_1, s_2 \leq 1\}.$$

Its image under T is the set of points

$$\begin{aligned} T(S) &= \{s_1T(\vec{b}_1) + s_2T(\vec{b}_2) : 0 \leq s_1, s_2 \leq 1\} \\ &= \{s_1A\vec{b}_1 + s_2A\vec{b}_2 : 0 \leq s_1, s_2 \leq 1\}. \end{aligned}$$

So $T(S)$ is the parallelogram with sides $A\vec{b}_1$ and $A\vec{b}_2$, and so its area is the absolute value of the determinant of $A\begin{bmatrix} \vec{b}_1 & \vec{b}_2 \end{bmatrix} = AB$.

Thus the area of $T(S)$ is the absolute value of $\det(AB) = \det(A)\det(B)$ which is $|\det(A)|$ times the area of S. □

The conclusion of Theorem 10 also holds when the parallelogram does not have a corner at the origin because translation of any parallelogram to the origin does not change its area.

The conclusion of Theorem 10 further holds when the area is not a parallelogram because any finite area of the plane can be divided into many small parallelograms.

Example. What is the area of the ellipse E determined by

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1?$$

WLOG we assume $a > 0$ and $b > 0$.

We know how to compute the area of the unit circle D determined by $x_1^2 + x_2^2 \leq 1$; the area is $\pi(1)^2 = \pi$.

If we can show that the ellipse E is the image of D under a linear transformation T, then we would obtain the area of the ellipse through the generalization of Theorem 10.

Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ whose standard matrix is

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}.$$

Any point (x_1, x_2) in the ellipse E is the image under T of $(x_1/a, x_2/b)$ in unit circle D, and so $T(D) = E$.

Thus the area of E is $|\det(A)|$ times area of D which is πab.