3-31-1994

Determinants.

$M_n(F)$ denotes the $n \times n$ matrices with entries in F.

Definition. A **determinant function** is a function $D : M_n(F) \to F$ such that:

1. D is a linear function in each row. That is, if $a \in F$ and $x, y \in F^n$,

$$D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & ax + y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = a \cdot D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} + D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

2. D is an **alternating function** in the rows. That is,

$$D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = -D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

3. $D(I) = 1$, where I is the $n \times n$ identity matrix.

Example. If two rows of a matrix A are equal, then $D(A) = 0$. If two rows are equal to $x \in F^n$, then by property 2

$$D(A) = D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = -D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = -D(A).$$

Therefore, $D(A) = 0$. $\square$

Example. Define $|\cdot| : M_2(\mathbb{R}) \to \mathbb{R}$ by

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Then $|\cdot|$ is a determinant function.
First,

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = bc - ad = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Therefore, property 2 holds.

Next,

$$\left|\begin{matrix} ka+a' & kb+b' \\ c & d \end{matrix}\right| = kad + a'd - kbc - b'c = k(ad-bc) + (a'd - b'c) = k\left|\begin{matrix} a & b \\ c & d \end{matrix}\right| + \left|\begin{matrix} a' & b' \\ c & d \end{matrix}\right|.$$

This proves linearity in the first row. To prove linearity in the second row, use linearity in the first row and the alternating property (proved earlier):

$$\left|\begin{matrix} c & d \\ ka+a' & kb+b' \end{matrix}\right| = -\left|\begin{matrix} ka+a' & kb+b' \\ c & d \end{matrix}\right| = -k\left|\begin{matrix} a & b \\ c & d \end{matrix}\right| - \left|\begin{matrix} a' & b' \\ c & d \end{matrix}\right| = k\left|\begin{matrix} c & d \\ a & b \end{matrix}\right| + \left|\begin{matrix} c & d \\ a' & b' \end{matrix}\right|.$$

Finally,

$$\left|\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\right| = 1. \square$$

Example. (Elementary row operations) Property 2 implies that swapping two rows multiplies the determinant by -1. By linearity, multiplying a row by a multiplies the determinant by a.

Consider the operation of adding a multiple of a row to another row. Suppose, for example, I'm performing the operation $r_i \to r_i + a \cdot r_j$. Let

$$A = \begin{bmatrix} & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \end{bmatrix}.$$

Then

$$D\begin{bmatrix} & \vdots & \\ \leftarrow & r_i + a \cdot r_j & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \end{bmatrix} = D\begin{bmatrix} & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \end{bmatrix} + a \cdot D\begin{bmatrix} & \vdots & \\ \leftarrow & r_j & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \end{bmatrix} = D\begin{bmatrix} & \vdots & \\ \leftarrow & r_i & \rightarrow \\ & \vdots & \\ \leftarrow & r_j & \rightarrow \end{bmatrix} = D(A).$$

Therefore, this kind of row operation leaves the determinant unchanged. $\square$

There are two problems here. First, I've defined a determinant function on 2×2 matrices. This solves the *existence* problem for determinant functions on $M_2(F)$, but I to show that there is a determinant function on $M_n(F)$. I'll do this inductively, using **expansion by cofactors**.

Next, I'll show that the determinant function defined in this way is the only determinant function on $M_n(F)$. This solves the *uniqueness* problem for determinants on $M_n(F)$.

Mathematicians often proceed in this way: Define an object by identifying properties which characterize it, rather than simply writing down a formula. I could have defined the determinant by cofactor expansion, but it says more about the structure of the determinant to *characterize* it by the axioms given above.

Moreover, the axiomatic approach will provide a clean proof of properties such as multiplicativity. Try proving that $D(AB) = D(A)D(B)$ from the cofactor expansion!

Lemma. Let $f : M_n(F) \to F$ be a function which is linear in each row and satisfies $f(A) = 0$ whenever two adjacent rows are equal. Then A is alternating.

Proof. First, I'll show that swapping two adjacent rows multiplies the value of f by -1. I'll show the required manipulations in schematic form:

$$f\begin{bmatrix} & \vdots & \\ \leftarrow & x & \rightarrow \\ \leftarrow & y & \rightarrow \\ & \vdots & \end{bmatrix} = f\begin{bmatrix} & \vdots & \\ \leftarrow & y+(x-y) & \rightarrow \\ \leftarrow & y & \rightarrow \\ & \vdots & \end{bmatrix} = f\begin{bmatrix} & \vdots & \\ \leftarrow & y & \rightarrow \\ \leftarrow & y & \rightarrow \\ & \vdots & \end{bmatrix} + f\begin{bmatrix} & \vdots & \\ \leftarrow & x-y & \rightarrow \\ \leftarrow & y & \rightarrow \\ & \vdots & \end{bmatrix} =$$

$$f\begin{bmatrix} & \vdots & \\ \leftarrow & x-y & \rightarrow \\ \leftarrow & y & \rightarrow \\ & \vdots & \end{bmatrix} = f\begin{bmatrix} & \vdots & \\ \leftarrow & x-y & \rightarrow \\ \leftarrow & x+(y-x) & \rightarrow \\ & \vdots & \end{bmatrix} = f\begin{bmatrix} & \vdots & \\ \leftarrow & x-y & \rightarrow \\ \leftarrow & x & \rightarrow \\ & \vdots & \end{bmatrix} + f\begin{bmatrix} & \vdots & \\ \leftarrow & x-y & \rightarrow \\ \leftarrow & y-x & \rightarrow \\ & \vdots & \end{bmatrix} =$$

$$f\begin{bmatrix} & \vdots & \\ \leftarrow & x & \rightarrow \\ \leftarrow & x & \rightarrow \\ & \vdots & \end{bmatrix} - f\begin{bmatrix} & \vdots & \\ \leftarrow & y & \rightarrow \\ \leftarrow & x & \rightarrow \\ & \vdots & \end{bmatrix} - f\begin{bmatrix} & \vdots & \\ \leftarrow & x-y & \rightarrow \\ \leftarrow & x-y & \rightarrow \\ & \vdots & \end{bmatrix} = -f\begin{bmatrix} & \vdots & \\ \leftarrow & y & \rightarrow \\ \leftarrow & x & \rightarrow \\ & \vdots & \end{bmatrix}.$$

To complete the proof that f is alternating, I must show that swapping two non-adjacent rows multiplies the value of f by -1. Since swapping adjacent rows multiplies the determinant by -1, it will suffice to show that swapping any two rows can be accomplished by an odd number of adjacent row swaps.

Without loss of generality, then, I may suppose the rows to be swapped are rows 1 and n. I'll indicate how to do the swaps by just displaying the row numbers. First, I do $n-1$ adjacent row swaps to move row 1 into the n-th position:

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & \ldots & n-2 & n-1 & n \\ 2 & 1 & 3 & 4 & \ldots & n-2 & n-1 & n \\ & & \vdots & & \ldots & & \vdots & \\ 2 & 3 & 4 & 5 & \ldots & n-1 & 1 & n \\ 2 & 3 & 4 & 5 & \ldots & n-1 & n & 1 \end{array}$$

Now I do $n-2$ more swaps to move (the old) row n into the first position:

$$\begin{array}{cccccccc} 2 & 3 & 4 & 5 & \ldots & n-1 & n & 1 \\ 2 & 3 & 4 & 5 & \ldots & n & n-1 & 1 \\ & & \vdots & & \ldots & & \vdots & \\ 2 & 1 & 3 & 4 & \ldots & n-2 & n-1 & 1 \\ 1 & 2 & 3 & 4 & \ldots & n-2 & n-1 & 1 \end{array}$$

I've done a total of $(n-1)+(n-2) = 2n-3$ swaps, an odd number. This proves the result. $\square$

Definition. Let $A \in M_n(F)$. $A(i \mid j)$ is the $(n-1)\times(n-1)$ matrix obtained by deleting the i-th row and j-th column of A.

Theorem. If C is a determinant function on $M_{n-1}(F)$, then for any $j \in \{1, \ldots, n\}$, the function

$$D(A) = \sum_i (-1)^{i+j} A_{ij} C\left(A(i \mid j)\right)$$

is a determinant function on $M_n(F)$.

Proof. I need to show that D is linear in each row, D is alternating, and $D(I) = 1$.

Linearity: I'll prove linearity in row k. Let $a \in F$, $x, y \in F^n$. I want to prove that

$$D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & ax + y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = a \cdot D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} + D \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

(All the action is taking place in the k-th row, and the other rows are the same in the three matrices). Label the three matrices above:

$$P = \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & ax + y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}, \quad Q = \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}, \quad R = \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & y & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

Expand each of the D's in the equation using the cofactor summation. Since C is a determinant function, C is linear. For $i \neq k$, then, the terms of the summation on the two sides clearly agree.

Consider the terms generated on both sides when $i = k$. On the left, I have

$$(-1)^{k+j}(ax_j + y_j)C\left(P(k \mid j)\right).$$

On the right, I have

$$(-1)^{k+j}ax_j C\left(Q(k \mid j)\right) + (-1)^{k+j}y_j C\left(R(k \mid j)\right).$$

However,

$$C\left(P(k \mid j)\right) = C\left(Q(k \mid j)\right) = C\left(R(k \mid j)\right),$$

because P, Q, and R only differ in row k, which is being deleted.

Therefore, the terms on the left and right are the same for all i, and D is linear.

Alternating: Since D is now known to be linear, the previous lemma applies: To show D is alternating, it suffices to show that D vanishes when two adjacent rows are equal.

Suppose without loss of generality that rows 1 and 2 are equal. All the terms in the cofactor expansion vanish except the first and second, since in all other cases the determinant function C will be applied to a matrix with two equal rows.

The first and second terms are

$$(-1)^{1+j}A_{1j}C\left(A(1 \mid j)\right) + (-1)^{2+j}A_{2j}C\left(A(2 \mid j)\right).$$

Now $A_{1j} = A_{2j}$, since the first and second rows are equal. Likewise, since the rows are equal, I get the same matrix by deleting one or the other. Therefore, $C\left(A(1 \mid j)\right) = C\left(A(2 \mid j)\right)$. The only things left are the signs, and $(-1)^{1+j}$ and $(-1)^{2+j}A_{2j}$ have opposite parity. Therefore, the cofactor expansion vanishes, and $D(A) = 0$, as I wished to prove.

$D(I) = 1$: Evidently $A_{ij} = 0$ unless $i = j$. Therefore, the cofactor expansion of $D(I)$ reduces to

$$(-1)^{j+j}A_{jj}C\left(A(j \mid j)\right) = 1 \cdot 1 \cdot C(I) = 1. \square$$

Theorem. If $D : M_n(F) \to F$ is an alternating function that is linear in each row, then

$$D(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)} D(I).$$

Proof. Write

$$A = \begin{bmatrix} \leftarrow & r_1 & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

Observe that

$$r_i = \sum_j A_{ij} e_j,$$

where e_j is the j-th standard basis vector (equivalently, the j-th row of the identity matrix).

By linearity,

$$D(A) = D \begin{bmatrix} \leftarrow & \sum_j A_{1j} e_j & \rightarrow \\ \leftarrow & r_2 & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix} = \sum_j A_{1j} D \begin{bmatrix} \leftarrow & e_j & \rightarrow \\ \leftarrow & r_2 & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

Now do the same with row 2:

$$D(A) = \sum_j \sum_k A_{1j} A_{2k} D \begin{bmatrix} \leftarrow & e_j & \rightarrow \\ \leftarrow & e_k & \rightarrow \\ & \vdots & \\ \leftarrow & r_n & \rightarrow \end{bmatrix}.$$

Do this for all n rows, using $j_1, \ldots, j_n$ as the summation indices:

$$D(A) = \sum_{j_1, \ldots, j_n} A_{1j_1} A_{2j_2} \cdots A_{nj_n} D \begin{bmatrix} \leftarrow & e_{j_1} & \rightarrow \\ \leftarrow & e_{j_2} & \rightarrow \\ & \vdots & \\ \leftarrow & e_{j_n} & \rightarrow \end{bmatrix}.$$

Since D is alternating,

$$D \begin{bmatrix} \leftarrow & e_{j_1} & \rightarrow \\ \leftarrow & e_{j_2} & \rightarrow \\ & \vdots & \\ \leftarrow & e_{j_n} & \rightarrow \end{bmatrix} = 0$$

if two rows are equal. Therefore, I only need to consider terms where all the rows are different — i.e. terms where $j_1, \ldots, j_n$ are distinct. This means that $\{j_1, \ldots, j_n\}$ is a permutation of $\{1, \ldots, n\}$, and I'm summing over $\sigma \in S_n$:

$$D(A) = \sum_{\sigma \in S_n} A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)} D \begin{bmatrix} \leftarrow & e_{\sigma(1)} & \rightarrow \\ \leftarrow & e_{\sigma(2)} & \rightarrow \\ & \vdots & \\ \leftarrow & e_{\sigma(n)} & \rightarrow \end{bmatrix}.$$

Now

$$\begin{bmatrix} \leftarrow & e_{\sigma(1)} & \rightarrow \\ \leftarrow & e_{\sigma(2)} & \rightarrow \\ & \vdots & \\ \leftarrow & e_{\sigma(n)} & \rightarrow \end{bmatrix}$$

is just the identity matrix with its rows permuted by σ. Every permutation is a product of transpositions, and a transposition (i.e. a row swap) multiplies the determinant by -1. Hence,

$$D \begin{bmatrix} \leftarrow & e_{\sigma(1)} & \rightarrow \\ \leftarrow & e_{\sigma(2)} & \rightarrow \\ & \vdots & \\ \leftarrow & e_{\sigma(n)} & \rightarrow \end{bmatrix} = \operatorname{sgn}(\sigma) D(I),$$

where $\operatorname{sgn}(\sigma)$ is the **sign** of the permutation σ: $+1$ is σ is even, -1 if σ is odd.

Therefore,

$$D(A) = \sum_{\sigma \in S_n} A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)} \operatorname{sgn}(\sigma) D(I). \quad \square$$

Corollary. There is a unique determinant function $|\cdot| : M_n(F) \to F$.

Proof. A determinant function is alternating and linear in each row, and satisfies $D(I) = 1$. The theorem then shows that $D(A)$ is completely determined by A. $\square$

Remark. Another way to express the proof of the Corollary is: If $D : M_n(F) \to F$ is alternating and linear in each row, then

$$D(A) = |A| D(I),$$

where $|\cdot|$ denotes the determinant function on $M_n(F)$. $\square$

Corollary. If $A \in M_n(F)$, then $|A| = |A^t|$.

Proof.

$$|A^t| = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A^t_{i\sigma(i)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{\sigma(i)i} =$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} A_{j\sigma^{-1}(j)} = \sum_{\sigma \in S_n} \operatorname{sgn}\left(\sigma^{-1}\right) \prod_{j=1}^{n} A_{j\sigma^{-1}(j)} =$$

$$\sum_{\sigma^{-1} \in S_n} \operatorname{sgn}\left(\sigma^{-1}\right) \prod_{j=1}^{n} A_{j\sigma^{-1}(j)} = \sum_{\tau \in S_n} \operatorname{sgn}(\tau) \prod_{j=1}^{n} A_{j\tau(j)} = |A|. \quad \square$$

Remark. By combining this result with the *construction* of the determinant using expansion by cofactors along columns, you can show that determinants may also be computed by by cofactor expansion along rows. $\square$

The uniqueness of determinant functions is useful in deriving properties of determinants.

Theorem. Let $A, B \in M_n(F)$. Then $|AB| = |A||B|$.

Proof. Fix B, and define

$$D(A) = |AB|.$$

Let r_i denote the i-th row of A. Then

$$D(A) = \begin{vmatrix} \leftarrow & r_1 B & \rightarrow \\ \leftarrow & r_2 B & \rightarrow \\ & \vdots & \\ \leftarrow & r_n B & \rightarrow \end{vmatrix}.$$

Now $|\cdot|$ is alternating, so interchanging two rows in the determinant above multiplies $D(A)$ by -1. Hence, D is alternating.

Next, I'll show that D is linear:

$$D \begin{bmatrix} & \vdots & \\ \leftarrow & kx + y & \rightarrow \\ & \vdots & \end{bmatrix} = \begin{vmatrix} & \vdots & \\ \leftarrow & (kx+y)B & \rightarrow \\ & \vdots & \end{vmatrix} =$$

$$k \cdot \begin{vmatrix} & \vdots & \\ \leftarrow & xB & \rightarrow \\ & \vdots & \end{vmatrix} + \begin{vmatrix} & \vdots & \\ \leftarrow & yB & \rightarrow \\ & \vdots & \end{vmatrix} = k \cdot D \begin{bmatrix} & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \end{bmatrix} + D \begin{bmatrix} & \vdots & \\ \leftarrow & y & \rightarrow \\ & \vdots & \end{bmatrix}.$$

This proves that D is linear in each row.

By the remark following the uniqueness theorem,

$$|AB| = D(A) = |A|D(I) = |A||B|. \square$$

Theorem. Let $A \in M_n(F)$. A is invertible if and only if $|A| \neq 0$.

Proof. If A is invertible, then

$$|A||A^{-1}| = |AA^{-1}| = |I| = 1.$$

Hence, $|A| \neq 0$.

Conversely, suppose $|A| \neq 0$. The discussion of the effect of row operations on determinants shows that if R is the row reduced echelon form for A, then $|R| \neq 0$. Hence, R has no all-zero rows. Therefore, R has n leading coefficients, and $R = I$. Hence, A is invertible. $\square$

The proof shows that if A is invertible, then

$$|A^{-1}| = \frac{1}{|A|}.$$

Here's an easy application of this result.

Corollary. Similar matrices have the same determinant.

Proof.

$$|PAP^{-1}| = |P||A||P^{-1}| = |P||P^{-1}||A| = |A|. \square$$

Next, I'll obtain two results that are of theoretical interest. These formulas are *not* efficient for computation, though they aren't bad for small matrices and systems since they're relatively easy to remember. To invert matrices or solve systems by machine, row reduction (and its variants) is the approach of choice.

Definition. The **classical adjoint** $\operatorname{adj} A$ is the matrix whose i-j-th entry is

$$(\operatorname{adj} A)_{ij} = (-1)^{i+j}|A(j \mid i)|.$$

Theorem. If A is invertible, then

$$A^{-1} = \frac{1}{|A|}\operatorname{adj} A.$$

Proof. Expand $|A|$ by cofactors of row i:

$$|A| = \sum_j (-1)^{i+j} A_{ij} |A(i \mid j)|.$$

Take $k \neq i$, and construct a new matrix by replacing row k of A with row i. Since the new matrix has two equal rows, its determinant is 0. On the other hand, expanding by cofactors of row k,

$$0 = \sum_j (-1)^{k+j} A_{kj} |A(k \mid j)| = \sum_j (-1)^{k+j} A_{ij} |A(k \mid j)|.$$

In words, this means that if you expand by cofactors of a row using the cofactors from *another* row, you get 0. You only get $|A|$ if you use the "right" cofactors — i.e. the cofactors for the given row. In terms of the Kronecker delta function,

$$\sum_j (-1)^{k+j} A_{ij} |A(k \mid j)| = \delta_{ik} |A| \quad \text{for all} \quad i, k.$$

Interpret this equation as a matrix equation, where the two sides represent the i-k-th entries of their respective matrices. What *are* the respective matrices? The right side is obviously the i-k-th entry of $|A| \cdot I$. For the left side,

$$(A \cdot \operatorname{adj} A)_{ik} = \sum_j A_{ij} (\operatorname{adj} A)_{jk} = \sum_j A_{ij} (-1)^{j+k} |A(k \mid j)|.$$

Therefore,

$$|A| \cdot I = A \cdot \operatorname{adj} A,$$

from which the result follows immediately. □

Corollary. (Cramer's rule) If A is an invertible $n \times n$ matrix, the unique solution to $Ax = y$ is given by

$$x_i = \frac{|B_i|}{|A|},$$

where B_i is the matrix obtained from A by replacing its i-th column by y.

Proof.

$$Ax = y,$$

$$(\operatorname{adj} A) A x = (\operatorname{adj} A) y,$$

$$|A| x_i = \sum_j (\operatorname{adj} A)_{ij} y_j = \sum_j (-1)^{i+j} y_j |A(j \mid i)|.$$

What is the last sum? It is a cofactor expansion of A along column i, where instead of the elements of A's column i I'm using the components of y. This is exactly $|B_i|$. □

Example. Let

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

Here's how to compute the inverse using cofactors:

$$A^{-1} = -\frac{1}{6} \begin{bmatrix} -2 & 0 & -4 \\ -3 & 3 & -6 \\ 5 & -3 & 4 \end{bmatrix}.\ \square$$