Linear Algebra & Geometry

LECTURE 12

- Inverse matrix
- Linear mappings

Theorem. (determinant versus not-quite-matrix-addition)

Suppose $s \in \{1,2,\dots,n\}$ and $A[i,j] = B[i,j] = C[i,j]$ for every i,j such that $j \neq s$ and $C[i,s] = A[i,s] + B[i,s]$. Then $\det(C) = \det(A) + \det(B)$.

Proof..

$$\det\begin{bmatrix} c_{1,1} & \cdots & a_{1,s}+b_{1,s} & \cdots & c_{1,n} \\ c_{2,1} & \cdots & a_{2,s}+b_{2,s} & \cdots & c_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{n,1} & \cdots & a_{m,s}+b_{n,s} & \cdots & c_{n,n} \end{bmatrix} =$$

By Laplace expansion on column s

$$\sum_{i=1}^{n}(-1)^{i+s}(a_{i,s}+b_{i,s})\det(C_{i,s}) = \sum_{i=1}^{n}(-1)^{i+s}a_{i,s}\det(C_{i,s}) + \sum_{i=1}^{n}(-1)^{i+s}b_{i,s}\det(C_{i,s}) = \det(A) + \det(B).$$

Warning. This is NOT about determinant of the sum of two matrices being equal to the sum of their determinants; **that is not true.** This is about determinant of a matrix whose ONE column is the sum of two vectors.

Theorem. (other properties of *det*)

For every $n \times n$ matrices A and B

1. $\det(A) \neq 0$ iff $r(A) = n$, in other words, rows of A are linearly independent
2. If for every i,j such that $i > j$ $a_{i,j} = 0$ (only 0's below the main diagonal, triangular matrix) then $\det(A) = a_{1,1} a_{2,2} \ldots a_{n,n}$
3. In particular, $det(I_{n,n}) = 1$
4. $\det(AB) = \det(A)\det(B)$

Proof. Omitted.

Part 2 suggests a strategy for calculation of determinants of large matrices: row-reduce the matrix to a triangular form.

Determinant and systems of linear equations

Theorem. (Uniqueness theorem)

A system of *n* linear equations with *n* unknowns

$$(*)\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \ \dots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \ \dots + a_{2,n}x_n = b_2 \\ \qquad \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \ \dots + a_{n,n}x_n = b_n \end{cases}$$

has a unique solution iff $\det(A) \neq 0$

Proof.

It follows from the fact that the corresponding homogeneous system has unique solution Θ iff rank(A)=n which, in turn is equivalent to $det(A) \neq 0$. Then, if (and that's a big IF) v_0 is a solution then all solutions v of $(*)$ look like $v = \Theta + v_0 = v_0$.

Warning.

The uniqueness theorem is a "both ways" implication but is often misunderstood. The conclusion should be understood as "the set of solutions of $(*)$ has exactly one element". Hence the negation of this is (contrary to what many people believe) not

"if $\det(A) = 0$ then $(*)$ has no solutions"

but rather (remember de Morgan's Law!)

"if $\det(A) = 0$ then $(*)$ the set of solutions of $(*)$ does not have exactly one element"

which means either zero or more than one element. Look at this:

$\begin{cases} x + y = 2 \\ 2x + 2y = 4 \end{cases}$ $\det\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = 4-4 = 0$ but the system has infinitely many solutions of the form $(t, 2 - t)$ where t is any real number.

Theorem. (Cramer's Rule)
Let A be an $n\times n$ matrix with $\det(A) \neq 0$ and let B be any $n\times 1$ matrix. Then the system of equations $AX = B$ has unique solution $X = [x_1, x_2, \ldots, x_n]^T$ and for each $i = 1, 2, \ldots, n$

$$x_i = \frac{detA_i}{detA},$$

where A_i is obtained by replacing the i-th column of A with B.
Proof (skipped).

Example.

$$\begin{cases} 2x + 4y - z = 11 \\ -4x - 3y + 3z = -20 \\ 2x + 4y + 2z = 2 \end{cases} \quad |A| = \begin{vmatrix} 2 & 4 & -1 \\ -4 & -3 & 3 \\ 2 & 4 & 2 \end{vmatrix} = -12+16+24-6-24+32=30,$$

$$|A_1| = \begin{vmatrix} 11 & 4 & -1 \\ -20 & -3 & 3 \\ 2 & 4 & 2 \end{vmatrix} = -66+80+24-6-132+160=264-204=60, \; x = 2$$

$$|A_2| = \begin{vmatrix} 2 & 11 & -1 \\ -4 & -20 & 3 \\ 2 & 2 & 2 \end{vmatrix} = -80+8+66-40-12+88 = 162-132 = 30, \; y = 1$$

$$|A_3| = \begin{vmatrix} 2 & 4 & 11 \\ -4 & -3 & -20 \\ 2 & 4 & 2 \end{vmatrix} = -12-160-176+66+160+32 = 258-348 = -90, \; z = -3$$

Definition. (Inverse matrix)
Let A be an $n \times n$ matrix. If there exists a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$ then A^{-1} is called *the inverse* (*matrix*) of A.

Fact. The inverse matrix for A, if it exists, is unique.

This follows from the very general fact the in every associative algebra the inverse element, if there is one, is unique.

Theorem.
A matrix is A invertible iff $\det(A) \neq 0$.

Proof.(⇒)
If A^{-1} exists, then $\det(AA^{-1}) = \det(I) = 1 = \det(A)\det(A^{-1})$ hence both $\det(A)$ and $\det(A^{-1})$ are different from zero.

(⇐)

If $\det(A) \neq 0$ then, from the uniqueness theorem for $n \times n$ systems of equations, for every $n \times 1$ matrix B there exists a (unique) solution of the system $AX = B$. Replacing B with consecutive columns of the identity matrix I we get the existence of the corresponding columns of the inverse matrix which in turn proves the existence of the inverse matrix itself.

To be more specific:

$$A^{-1} = \mathrm{X} = \begin{bmatrix} \mathrm{x}_{1,1} & \mathrm{x}_{1,2} & \cdots & \mathrm{x}_{1,n} \\ \mathrm{x}_{2,1} & \mathrm{x}_{2,2} & \cdots & \mathrm{x}_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ \mathrm{x}_{n,1} & \mathrm{x}_{n,2} & \cdots & \mathrm{x}_{n,n} \end{bmatrix}$$

$$\mathrm{A} = \begin{bmatrix} \mathrm{a}_{1,1} & \mathrm{a}_{1,2} & \cdots & \mathrm{a}_{1,n} \\ \mathrm{a}_{2,1} & \mathrm{a}_{2,2} & \cdots & \mathrm{a}_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ \mathrm{a}_{n,1} & \mathrm{a}_{n,2} & \cdots & \mathrm{a}_{n,n} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathrm{I}$$

Consider $\begin{bmatrix} \mathrm{a}_{1,1} & \mathrm{a}_{1,2} & \cdots & \mathrm{a}_{1,n} \\ \mathrm{a}_{2,1} & \mathrm{a}_{2,2} & \cdots & \mathrm{a}_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ \mathrm{a}_{n,1} & \mathrm{a}_{n,2} & \cdots & \mathrm{a}_{n,n} \end{bmatrix} \begin{bmatrix} \mathrm{x}_{1,1} \\ \mathrm{x}_{2,1} \\ \vdots \\ \mathrm{x}_{n,1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. The system is uniquely solvable and the solution, X_1 is the first column of A^{-1}. The same can be said about the second, third and each next column of X and I. QED

The proof provides a method (two methods, really) for calculating A^{-1} (that's one reason I insist on doing proofs):

Method 1.

Row-reduce the following matrix to a row-canonical one

$$[A|I] = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} & 1 & 0 & \dots & 0 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} & 0 & 0 & \dots & 1 \end{bmatrix} \sim \dots \sim \dots \sim \dots \sim$$

$$\sim \dots \sim \begin{bmatrix} 1 & 0 & \dots & 0 & x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ 0 & 1 & \dots & 0 & x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 & x_{n,1} & x_{n,2} & \dots & x_{n,n} \end{bmatrix} = [I|A^{-1}]$$

This is always possible if A is invertible. It proves that A is invertible iff it may be row-reduced to the identity matrix I.

Method 2.

Using Cramer's rule to calculate each $x_{i,j}$ of A^{-1}.

This method involves calculation of det(A) and n^2 determinants of the size $(n-1) \times (n-1)$. For large matrices it takes forever.

$x_{i,j}$ appears in j-th column of A^{-1} which means must consider

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_{1,j} \\ \vdots \\ x_{i,j} \\ \vdots \\ x_{n,j} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$ $= I_j$ where the solitary 1 in I_j is in the j-th position. So, in order to find the i-th unknown we need divide the determinant of $A^*_{i,j}$ (A with i-th column replaced by I_j) by det(A).

$$detA_{i,j}^{*} = det\begin{bmatrix} a_{1,1} & \dots & 0 & \dots & a_{1,n} \\ \vdots & & \vdots & & \vdots \\ a_{j,1} & \dots & 1 & \dots & a_{j,n} \\ \vdots & & \vdots & \dots & \vdots \\ a_{n,1} & & 0 & \dots & a_{n,n} \end{bmatrix}$$

in j-th row and i-th column of $A_{i,j}^{*}$

If you do this determinant by *i*-th column, the only nonzero term in the Laplace expansion will be $(-1)^{i+j}$ times the determinant obtained by the removal of *j*-th row and *i*-th column from $A_{i,j}^{*}$. Here is the funny thing: *A* and $A_{i,j}^{*}$ only differ on the *i*-th column, which is being removed. Hence $detA_{i,j}^{*} = (-1)^{i+j}detA_{j,i}$ and, finally, $x_{i,j} = \frac{(-1)^{i+j}detA_{j,i}}{detA}$. In other words

$$A^{-1} = \frac{1}{detA}\left[(-1)^{i+j}detA_{i,j}\right]^{T}$$

Example.

$$A=\begin{bmatrix}0&2&1&1\\1&1&2&0\\2&1&0&0\\1&3&4&1\end{bmatrix}.\ \text{Find } A^{-1}.$$

Method 1. (Gauss elimination). Notice and remember the strategy used:

Step one: get number 1 in the upper left corner

$$\begin{bmatrix}0&2&1&1&1&0&0&0\\1&1&2&0&0&1&0&0\\2&1&0&0&0&0&1&0\\1&3&4&1&0&0&0&1\end{bmatrix}\sim \boldsymbol{r_4-r_2,-r_3+2r_2}\begin{bmatrix}0&2&1&1&1&0&0&0\\1&1&2&0&0&1&0&0\\0&1&4&0&0&2&-1&0\\0&2&2&1&0&-1&0&1\end{bmatrix}\sim$$

$$\boldsymbol{r_1-2r_3,r_2-r_3,r_4-2r_3}\begin{bmatrix}0&0&-7&1&1&-4&2&0\\1&0&-2&0&0&-1&1&0\\0&1&4&0&0&2&-1&0\\0&0&-6&1&0&-5&2&1\end{bmatrix}\boldsymbol{-r_1+r_4}$$

$$\begin{bmatrix}0&0&1&0&-1&-1&0&0\\1&0&-2&0&0&-1&1&0\\0&1&4&0&0&2&-1&0\\0&0&-6&1&0&-5&2&1\end{bmatrix}$$

A^{-1}

Method 2. (Cramer's Rule, cofactors)

$\det A = 1$. We are cheating here; this is based on method 1. Only two transformations in the previous slide affected the determinant, in both cases they were like $-r_s + cr_t$ which really means two operations: *scale* r_s *by* (-1) and *add to the new* r_s *another row (perhaps scaled by some factor)*. Scaling a row by (-1) changes the sign of the determinant and we did it twice.

Let's calculate just a single entry of A^{-1}, say $A^{-1}(2,3)$. According to the cofactor theorem $A^{-1}(2,3) = \frac{1}{detA}(-1)^{2+3}\det(A_{3,2})$

$$\det(A_{3,2}) = \begin{vmatrix} 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 3 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 4 & 1 \end{vmatrix} = 4 - 2 - 1 = 1, \text{ which}$$

means $A^{-1}[2,3]$ should be -1. We move back one slide and … surprise, surprise! it checks. Now you must calculate the remaining 15 entries of A^{-1}!

Linear Mappings

Definition.

Let V and W be vector spaces over a field $\mathbb{F}$. A function $\phi: V \to W$ is called a *linear mapping* iff

(a) $(\forall u, v \in V)\ \phi(u+v) = \phi(u) + \phi(v)$,

(b) $(\forall v \in V)(\forall p \in \mathbb{F})\ \phi(pv) = p\phi(v)$.

Proposition.

A function $\phi: V \to W$ is a linear mapping iff

(c) $(\forall q, p \in \mathbb{F})\ (\forall u, v \in V)\ \phi(qu + pv) = q\phi(u) + p\phi(v)$.

Proof. ($\Rightarrow$) Suppose ϕ is a linear mapping. Then $\phi(qu + pv) = \phi(qu) + \phi(pv) = q\phi(u) + p\phi(v)$, by (a) and (b).

($\Leftarrow$) To prove (a), we put $p = q = 1$ in (c) and to prove (b) we put $q = 0$. Then $\phi(pv) = \phi(\Theta + pv) = \phi(0u + pv) = 0\phi(u) + p\phi(v) = \Theta + p\phi(v) = p\phi(v)$. QED

Example 1. $V = W = \mathbb{R}_n[x]$, $\phi(f(x)) = f\,'(x)$. Differentiation of polynomials is obviously a linear mapping. Clearly, we can extend this observation to any space of differentiable functions.

Example 2. ϕ: $\mathbb{R}^3 \to \mathbb{R}^4$, $\phi(x, y, z) = (x + y, 2x - z, x + y + z, y)$. We have $\phi((x, y, z) + (a, b, c)) = \phi(x + a, y + b, z + c) = ((x + a) + (y + b), 2(x + a) - (z + c), (x + a) + (y + b) + (z + c), (y + b)) = (x + y + a + b, 2x - z + 2a - c, x + y + z + a + b + c, y + b) = \phi(x, y, z) + \phi(a, b, c)$ and $\phi(p(x, y, z)) = \phi(px, py, pz) = (px + py, 2px - pz, px + py + pz, py) == (p(x + y), p(2x - z), p(x + y + z), py) = p\phi(x, y, z)$. Hence ϕ is a linear mapping.

Example 3.

ϕ: $\mathbb{R}^3 \to \mathbb{R}^2$, $\phi(x, y, z) = (x + y - 1, x - z)$. ϕ is NOT a linear mapping as $\phi(\Theta + \Theta) = \phi(\Theta) = \phi(0,0,0) = (0 + 0 - 1, 0 - 0) = (-1,0)$, while $\phi(\Theta) + \phi(\Theta) = (-1,0) + (-1,0) = (-2,0)$. This is enough to show that ϕ is not linear. Let us note that ϕ does not satisfy the second condition either, as $\phi(2\Theta) = \phi(\Theta) = (-1,0)$ and $2\phi(\Theta) = 2(-1,0) = (-2,0)$.

We will now define two important parameters of a linear mapping, rank and nullity.