1 Last time: determinants

Let n be a positive integer.

If A is an $n \times n$ matrix, then its *determinant* is the number

$$\det A = \sum_{X \in S_n} \Pi(X, A)(-1)^{\mathrm{inv}(X)}$$

where

- S_n is the set of $n \times n$ permutation matrices.
- $\Pi(X, A)$ is the product of the entries of A in the nonzero positions of X.
- $\mathrm{inv}(X)$ is the number of 2×2 submatrices of X equal to $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Theorem. The determinant is the unique function $\det : \{ \ n \times n \text{ matrices } \} \to \mathbb{R}$ such that

(1) $\det I_n = 1$.

(2) Switching two columns reverses the sign of the determinant.

(3) $\det A$ is linear as a function of a single column A if all other columns are fixed.

We have the following formulas for the determinants of small matrices:

$$\det \begin{bmatrix} a \end{bmatrix} = a.$$

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

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a(ei - fh) - b(di - fg) + c(dh - ef).$$

The *diagonal (positions)* of an $n \times n$ matrix are the positions $(1, 1)$, $(2, 2)$, $\dots$, (n, n). The *diagonal entries* of a matrix are the entries in these positions.

A matrix is *upper triangular* if all of its nonzero entries are in positions on or above the diagonal, and *lower triangular* if all of its nonzero entries are in positions on or below the diagonal. A *triangular matrix* is a square matrix which is either upper or lower triangular.

A *diagonal matrix* is a matrix which is both upper and lower triangular: in other words, all of its nonzero entries appear in diagonal positions.

Fact. If A is triangular then $\det A$ is the product of the diagonal entries of A.

Two important theorems:

Theorem. A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem. If A and B are $n \times n$ matrices then $\det(AB) = (\det A)(\det B)$ and $\det(A^T) = \det A$.

Algorithm to compute $\det A$.

Input: an $n \times n$ matrix A.

1. Start by setting $K = 1$.
2. Row reduce A to an echelon form E, while doing the following:
   (a) When you switch two rows, multiply K by -1.
   (b) When you rescale a row by a nonzero factor λ, multiply K by λ.
   (c) When you add a multiple of a row to another row, don't do anything to K.

The determinant of A is then given by $\det A = (\det E)/K =$ the product of the diagonal entries of the echelon form E divided by K.

Example. Suppose $a \neq 0$ Then

$$\begin{aligned} A &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} && K = 1 \\ &\sim \begin{bmatrix} 1 & b/a \\ c & d \end{bmatrix} && \text{(we multiplied the first row by } 1/a) \quad K = 1/a \\ &\sim \begin{bmatrix} 1 & b/a \\ 0 & d - bc/a \end{bmatrix} = E && \text{(we added } -c \text{ times the first row to the second)} \quad K = 1/a \end{aligned}$$

so $\det A = (\det E)/K = (d - bc/a)/(1/a) = ad - bc$, which agrees with our earlier formula.

<u>A recursive formula for $\det A$.</u>

Given an $n \times n$ matrix A, define $A^{(i,j)}$ as the $(n-1) \times (n-1)$ submatrix formed by deleting row i and column j from A.

Example. If $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ then $A^{(1,2)} = \begin{bmatrix} d & f \\ g & i \end{bmatrix}$.

Theorem. If A is the $n \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

then

$$\boxed{\det A = a_{11} \det A^{(1,1)} - a_{12} \det A^{(1,2)} + a_{13} \det A^{(1,3)} - \dots - (-1)^n a_{1n} \det A^{(1,n)}}$$

and also

$$\boxed{\det A = a_{11} \det A^{(1,1)} - a_{21} \det A^{(2,1)} + a_{31} \det A^{(3,1)} - \dots - (-1)^n a_{n1} \det A^{(n,1)}.}$$

Note that each $A^{(1,j)}$ or $A^{(j,1)}$ is a square matrix smaller than A, so $\det A^{(1,j)}$ or $\det A^{(j,1)}$ can be computed by the same formula on a smaller scale.

This formula is most useful if A has many entries which are zero.

Example. If $A = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 3 & 4 & 5 \\ 1 & 6 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$ then $\det A = \det \begin{bmatrix} 3 & 4 & 5 \\ 6 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} - 0 + 2 \det \begin{bmatrix} 0 & 3 & 5 \\ 1 & 6 & 0 \\ 0 & 1 & 1 \end{bmatrix} - 0.$

We have

$$\det \begin{bmatrix} 3 & 4 & 5 \\ 6 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} 3 & 6 & 1 \\ 4 & 0 & 1 \\ 5 & 0 & 1 \end{bmatrix} = -\det \begin{bmatrix} 6 & 3 & 1 \\ 0 & 4 & 1 \\ 0 & 5 & 1 \end{bmatrix} = -\det \begin{bmatrix} 6 & 0 & 0 \\ 3 & 4 & 5 \\ 1 & 1 & 1 \end{bmatrix} = -6 \det \begin{bmatrix} 4 & 5 \\ 1 & 1 \end{bmatrix} = 6$$

since we taking transposes doesn't change the determinant, and switching columns reverses the sign of the determinant. Similarly, we have

$$\det \begin{bmatrix} 0 & 3 & 5 \\ 1 & 6 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} 0 & 1 & 0 \\ 3 & 6 & 1 \\ 5 & 0 & 1 \end{bmatrix} = -\det \begin{bmatrix} 3 & 1 \\ 5 & 1 \end{bmatrix} = -(3-5) = 2.$$

Therefore $\det A = 6 + 2 \cdot 2 = 10$.

2 Interpreting the determinant geometrically

The last thing we'll mention is how the determinant of a matrix A can be interpreted as measuring the volume of a certain region bounded by the columns of A. This easiest way to explain this is in 2 dimensions:

Proposition. If $u, v \in \mathbb{R}^2$ are two vectors and $A = \begin{bmatrix} u & v \end{bmatrix}$ then the area of the parallelogram with sides u and v is $|\det A|$.

Proof idea. Make things simple by putting u and v both in the first quadrant. Draw a picture of the parallelogram P inside the rectangle R whose diagonal is $u + v$ and whose sides are on the x- and y-axes. Then compute the area of P by subtracting the areas of an appropriate number of rectangular and triangular regions from R. One finds that this area is $ad - bc$ if $u = \begin{bmatrix} a \\ c \end{bmatrix}$ and $v = \begin{bmatrix} b \\ c \end{bmatrix}$. □

This property generalises to 3 and higher dimensions, once we replace "area" by "volume" and then specify what the higher dimensional analogue of a parallelogram is. (In 3 dimensions, it's the *parallelepiped*).

Theorem. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation with standard matrix A. If S is a parallelogram in the $\mathbb{R}^2$ plane then the area of $T(S)$ is the area of S times $|\det A|$.

3 Review topics

We'll spend the rest of today giving a quick summary of the main important topics to review for the midterm. Check out the list of practice problems on the course website to get an idea of the types of problems that might be asked.

Things to know for the exam (from lectures 1–12):

1. **Linear systems.**
   (a) Know definition of a *linear system* and its *solutions*, and when two linear systems are *equivalent*.
   (b) Understand that every linear system has either 0, 1, or infinitely many solutions.
   (c) Know how to extract the *coefficient matrix* and *augmented matrix* from a linear system. Understand the difference between these two matrices.
2. **Row reduction.**

(a) Know the definitions of the elementary row operations: (1) swapping two rows, (2) rescaling a row by a nonzero constant, (3) adding a multiple of one row to another.

(b) Understand how row operations, performed on a matrix A, correspond to multiplying A on the right by an elementary matrix E.

(c) Matrices are *row equivalent* if one is obtained from the other by a sequence of row operations. Each matrix is row equivalent to a unique matrix in reduced echelon form. Linear systems whose augmented matrices are row equivalent have the same solutions.

(d) Know definitions of *echelon form*, *reduced echelon form*.

(e) **Important:** know how to carry out the *row reduction algorithm* to reduce a matrix, quickly and without using a calculator, to its unique reduced echelon form.

(f) Understand the meaning of the *pivot positions* and *pivot columns* of a matrix A. Remember that these cannot be determined without first reducing A to echelon form.

(g) If A is the coefficient matrix of a linear system in variables $x_1, x_2, \dots, x_n$ then x_i is a *basic variable* if i is a pivot column of A, and otherwise is a *free variable*.

(h) A linear system has zero solutions if the last column in its augment matrix is a pivot column. Assume this does not happen. Then the system has infinitely many solutions if it has at least one free variable, and has exactly one solution otherwise.

3. **Vectors and matrix operations.**

(a) A *vector* is just a list of numbers oriented as a column: $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ or $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

(b) $\mathbb{R}^n$ is the set of vectors $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ with n rows where $v_1, v_2, \dots, v_n \in \mathbb{R}$.

(c) We can add vectors with the same number of rows and multiply vectors by scalars:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} d \\ e \\ f \end{bmatrix} = \begin{bmatrix} a+d \\ b+e \\ c+f \end{bmatrix} \qquad \text{and} \qquad x \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} xa \\ xb \\ xc \end{bmatrix}.$$

(d) We draw vectors in $\mathbb{R}^2$ as arrows from the origin in the xy-plane. Know how to interpret the sum of two vectors in $\mathbb{R}^2$ in terms of these pictures.

(e) Know that the *zero vector* in $\mathbb{R}^n$ is the vector whose entries are all zeros: $0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$.

(f) Know definitions of *linear combinations* and *span* of a set of vectors.

(g) Know how to multiply an $m \times n$ matrix A and a vector $v \in \mathbb{R}^n$.

(h) If the columns of A are $a_1, a_2, \dots, a_n \in \mathbb{R}^m$ and $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ then $Ax = b \in \mathbb{R}^m$ has the same solutions as the vector equation $x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b$.

(i) If A is an $m \times n$ matrix then the following are equivalent:

i. $Ax = b$ has a solution for each $b \in \mathbb{R}^m$.

ii. Each $b \in \mathbb{R}^m$ is a linear combination of the columns of A.

iii. The span of the columns of A is $\mathbb{R}^m$

iv. A has a pivot position in every row.

4. **Linear independence**.

(a) Know the definition of a *linearly independent/dependent* set of vectors.

(b) Understand that two vectors are linearly independent if and only if neither is a scalar multiple of the other, while n vectors are linearly independent if and only if no vector is a linear combination of the others.

(c) If $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ and $p > n$, then the vectors are linearly dependent.

5. **Linear transformations**.

(a) Understand what the function notation "$f : X \to Y$" means.

(b) Know definitions of *domain*, *codomain*, *image*, *range* of function.

(c) Know definition of a *linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$.

(d) Every linear function $T : \mathbb{R}^n \to \mathbb{R}^m$ has the formula $f(v) = Av$ for some $m \times n$ matrix A, called the *standard matrix* of T.

(e) Know how to compute A from T via $A = \begin{bmatrix} T(e_1) & T(e_2) & \dots & T(e_n) \end{bmatrix}$.

(f) Understand definitions of *one-to-one*, *onto*, and *invertible* functions.

(g) T is linear and one-to-one $\Leftrightarrow$ columns of A are linearly independent.

(h) T is linear and onto $\Leftrightarrow$ columns of A span $\mathbb{R}^m$.

(i) Understand how to add linear transformations/matrices and multiply these by scalars.

(j) Know definition of the *composition* $f \circ g$ of two functions and how this corresponds to the product AB of two matrices. What must be the sizes of matrices A and B for AB to be defined?

(k) Recall definitions of the identity matrix and matrix transpose; $(AB)^T = B^T A^T$.

6. **Inverses**.

(a) Know the definitions of an invertible matrix, and what are the definitions of the inverses of an invertible function or invertible matrix.

(b) The inverse of a linear, invertible function $\mathbb{R}^n \to \mathbb{R}^n$ is also linear.

(c) If A is an $n \times n$ matrix then the following are equivalent:

i. A is invertible.

ii. There is a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I_n$.

iii. The columns of A span $\mathbb{R}^n$.

iv. The columns of A are linearly independent.

v. $\texttt{RREF}(A) = I_n$.

vi. $\det A \neq 0$.

(d) Know how to compute A^{-1}: reduce $\begin{bmatrix} A & I_n \end{bmatrix}$ to reduced echelon form to obtain $\begin{bmatrix} I_n & A^{-1} \end{bmatrix}$.

(e) Know formula for inverse of 2×2 matrix: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$, in which case

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

7. **Subspaces.**
   (a) Know definitions of *subspace*, *basis*, and *dimension*.
   (b) Know definitions of *column space*, *null space*, and *rank* of a matrix.
   (c) Know how to compute a basis for $\operatorname{Nul} A$ and $\operatorname{Col} A$, and how to compute the dimension of these subspaces.
   (d) Understand that if A has n columns then $n = \operatorname{rank} A + \dim \operatorname{Nul} A$.
   (e) Basis theorem: if H is a subspace of $\mathbb{R}^n$ with dimension p, then any set of p vectors spanning H is a basis for H, and any set of p linearly independent vectors in H is a basis for H.
8. **Determinants.**
   (a) (Know properties of and algorithms to compute $\det A$ summarised at beginning of this lecture.)