Lecture 4-22

Last time we saw that the determinant $\det A$ of a square matrix A is nonzero if and only if A is invertible. We can actually give an explicit formula for the inverse A^{-1} of A whenever $\det A \neq 0$. To do this, let B be the *transpose* of the cofactor matrix of A, so that its ji-th entry b_{ji} is the ij-cofactor $C^{ij} = (-1)^{i+j} \det A^{ij}$ of A. Computing the product AB, we see that its ii-th entry for any i is given by a sum matching the expansion of $\det A$ about its i-th row, whence this entry is $\det A$. On the other hand, if $i \neq j$, then the ij-th entry of AB is again given by the expansion of $\det A$ about the j-th row of A, but with the the k-th entry a_{jk} of that row replaced by a_{ik}. Thus it is the expansion about the j-th row of the determinant of the matrix A' obtained from A by replacing its j-th row with another copy of its i-th row, whence this entry is 0. Hence $AB = (\det A)I$, whence $A^{-1} = (1/\det A)B$. In words, *the inverse matrix is the transpose of the cofactor matrix divided by the determinant.* This formula is not generally useful for computations, except for 2×2 matrices; in that case it says that the inverse of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $(1/h)\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, where $h = \det A = ad - bc$. Thus the diagonal entries a, d switch places in the inverse, while the off-diagonal entries b, c switch signs. Returning to the general $n \times n$ case, we get something called *Cramer's Rule* as a consequence of the formula for A^{-1}. Given a linear system $AX = B$ such that A is invertible, so that the unique solution is $A^{-1}B$, we can write down an explicit formula for the value of the i-th variable x_i in the solution, namely $\det A'_i / \det A$, where A'_i is the matrix obtained form A by replacing its i-th column with the vector B on the right side. In particular, the value of $x_i$is given by a ratio of determinants. The formula follows at once by a direct computation of $A^{-1}B$, using the above formula for A^{-1}.

We also have the extremely useful *product rule for determinants*: we have $\det AB = (\det A)(\det B)$ if A, B are $n \times n$ matrices. The sneaky way to prove this formula is to note first that if B is singular, so that the rows of B fails to span $\mathbb{R}^n$, then the same is true of the rows of AB, these being combinations of the rows of B. Thus $\det AB = \det B = (\det A)(\det B) = 0$ in this case. Otherwise, if B is nonsingular (and fixed), then the function mapping the rows of A to the ratio $(\det AB)/(\det B)$ is easily shown to be alternating and multilinear in these rows and moreover $(\det IB)/(\det B) = 1$. This forces $(\det AB)/(\det B) = \det A$, as claimed, by the uniqueness of the determinant as an alternating multilinear functions of the rows equalling 1 on the rows of the identity matrix. In particular, for an invertible matrix A, we have $\det A^{-1} = (\det A)^{-1}$, since $\det I = 1 = (\det A)(\det A^{-1})$.

Finally let me mention the standard geometric interpretation of the determinant: given n independent vectors $\vec{v}_1, \ldots, \vec{v}_n$ in $\mathbb{R}^n$, let A be the matrix whose i-th column is $\vec{v}_i$. Let P be the parallelepiped spanned by the $\vec{v}_i$ consisting by definition of all linear combinations $\sum_{i=1}^n r_i \vec{v}_i$ where $r_i \in [0, 1]$. Then the n-dimensional volume of P is the absolute value $|\det A|$ of the determinant of A. We will not attempt to prove this formula, as we have not even defined n-dimensional volumes yet, but we mention the formula now because it will appear in the change of variable formula for multiple integrals, which we will see toward the end of the quarter.

In the remaining time today we start to review the course so far. We began with a piece of calculus which led directly to linear algebra. Given a real-valued function f defined on an an open subset of $\mathbb{R}^n$ a necessary condition for f to have a local maximum or minimum at $\vec{a}$ is that $\nabla f(\vec{a}) = \vec{0}$ or $\nabla f(\vec{a})$ is undefined; in practice we ignore the latter possibility. If this holds we need a way to decide whether $\vec{a}$ is a local maximum, a local minimum, or neither for f. We do this by setting up the Hessian matrix $H = H(\vec{a})$ whose ij-th entry is $\partial^2 f/\partial x_i \partial x_j(\vec{a})$, a symmetric matrix, and then working out whether H is positive definite (in the sense that $\vec{v}^T H\vec{v} > 0$ for all column vectors $\vec{v} \neq \vec{0}$, or negative definite, or neither. In the 2×2 case, setting $A = f_{xx}(\vec{a}), B = f_{xy}(\vec{a})$, and $C = f_{yy}(\vec{a}), H$ is positive definite and $\vec{a}$ is a local minimum exactly when $A > 0, AC - B^2 > 0$, while H is negative definite exactly when $A < 0, AC - B^2 > 0$. If $AC - B^2 < 0$, then $\vec{a}$ is a saddle point for f.

The last bit of calculus that we did before starting linear algebra had to do with maximizing or minimizing real-valued functions restricted to level sets of other functions. The criterion for $\vec{a}$ to be a critical point for a function f restricted to a level set for the function g is that $\nabla f(\vec{a}) = \lambda \nabla g(\vec{a})$ (or $\nabla g(\vec{a}) = \vec{0}$, but we will ignore this possibility). There is no second-derivative test for the nature of a critical point in this setting; we are forced to look at the values of f at all critical points and compare these values to decide where the global maximum and minimum of f occur.