Lecture 3-31

Continuing from last time, we studied functions $f : \mathbb{R}^2 \to \mathbb{R}$ of two variables having a critical (also sometimes called *stationary*) point at (a, b) and continuous second-order partials there. We learned that if we set $A = f_{xx}(a, b), B = f_{xy}(a, b), C = f_{yy}(a, b)$, then $A > 0, AC - B^2 > 0$ imply that (a, b) is a local minimum, while $A < 0, AC - B^2 > 0$ imply that (a, b) is a local maximum. The condition $AC - B^2 < 0$ implies that (a, b) is a saddle point. A convenient way to reformulate (and remember) this condition is to set up the 2×2 *Hessian matrix* H of f at (a, b), whose rows and columns are indexed by the variables $x_1 = x, x_2 = y$ and whose ijth entry is $f_{x_i,x_j}(a, b)$. Then the condition to be a local minimum is that $\det H, h_{11} > 0$; the condition to be a local maximum is that $\det H > 0, h_{11} < 0$. The condition to be a saddle point is that $\det H < 0$. We can see at once from simple examples, by the way, that we can say nothing in the case $\det H = 0$: the functions $f(x, y) = x^4 + y^2, g(x, y) = -x^4 - y^2, h(x, y) = x^4 - y^2$ all have $(0, 0)$ as a critical point with $\det H = 0$, with $(0, 0)$ being a local minimum in the first case, a local maximum in the second, and a saddle point in the last case.

What about functions $f(x_1, \ldots, x_n)$ of more than two variables? To analyze such functions, we first of all need the multivariable analogue of the quadratic approximation; in the case of a critical point (a, b), the approximation to $f(\vec{a} + \vec{h}) - f(\vec{a})$ is given by $(1/2) \sum_{i-1}^{n} f_{ii}(\vec{a})h_{ii}^2 + \sum_{i<j\leq n} f_{ij}(\vec{a})h_i h_j$, where $\vec{h} = (h_1, \ldots, h_n)$. Now to understand this last homogeneous quadratic function of n variables, we need more linear algebra (which we will sneakily learn by doing calculus); for now we just note that we can rewrite it as $(\vec{h})^t H\vec{h}$, where we write $\vec{h}$ as a column vector, $(\vec{h})^t$ for the corresponding row vector, and the ijth entry h_{ij} of H is $(1/2)f_{ij}(\vec{a})$ (so that H is the Hessian matrix defined above divided by 2). We will need to study eigenvalues and eigenvectors (to be defined later) of square matrices (having the same number of rows as columns) before we can say more about this situation.

For now we look at a couple of examples. If $f(x, y) = xye^{-(x^2+y^2)/2}$ (Example 6 in the text, p. 833), then we find that the x- and y-partials of f are $y(x^2 - 1)e^{-(x^2+y^2)/2}$, $x(y^2 - 1)e^{-(x^2+y^2)/2}$, respectively, so there are just five critical points, namely $(0, 0)$ and $(\pm 1, \pm 1)$. Defining as usual A, B, C to be the xx, xy, yy-partials of f at these points and setting $D = \det H = AC - B^2$, we find that $B = D = -1, A = C = 0$ at $(0, 0)$, whence this point is a saddle point, while $A, C, D > 0, B = 0$ at $\pm(1, 1)$, whence these points are local minima, and $A, C < 0, D > 0, B = 0$ at $\pm(1, -1)$ and these points are local maxima. This function has $f(x, y) \to 0$ as $x, y- \to \infty$, so it must have an absolute maximum and minimum on the xy-plane (note that this property does not hold generally for differentiable functions f, as the plane is not a bounded set). As there are only two local maxima, the common value e^{-1} of $f(x, y)$ at both of them must be the global maximum (attained by $(1, 1)$); its negative is the global minimum, attained at $(-1, 1)$. If $g(x, y) = xy + (1/x) + (1/y)$, then $g_x = y - (1/x^2), g_y = x - (1/y^2)$, whence $g_x = g_y = 0$ if and only if $x = y$ and $x^3 = 1$, or if and only if $x = y = 1$. Here we don't have to bother with the second-derivative test; since $|g(x, y)| \to \infty$ as $x \to \infty$ or $y \to \infty$ and $g(x, y)$ can have either sign if one or both of x, y are large in absolute value, it follows that $g(x, y)$ has no global maximum or minimum overall, but must have a global minimum when restricted to the first quadrant, which must occur at the unique critical point $(1,1)$.

In practice one often deals with functions defined on a closed rather than an open subset of $\mathbb{R}^n$. As previously noted, such a function f having a local maximum or minimum at a point will typically *not* have partial derivatives equal to 0 at that point. It is still true, however, that if a local maximum or minimum of f occurs at a point not on the boundary, then the partials $\partial f/\partial x_i$ (if defined) must be 0 at such a point. As for the boundary points, they often form one or more parametrized curves; we can maximize or minimize f along such a curve by standard one-variable techniques. Thus for example if $f(x,y) = x^2 + y^2 - 2x - 2y + 4$ and we consider only (x,y) lying on the disk D of radius 3 centered at $(0,0)$ (Example 1 in the text, p. 837), then $\nabla f = (2x-2, 2y-2)$, whence the only critical point is $(1,1)$ (which does lie on the disk). On the boundary of the disk, we have $x = 3\cos t, y = 3\sin t, f(x,y) = 13 - 6\cos t - y\sin t$, whence critical points of f occur only at points where $\sin t = \cos t$. By directly eamining all critical points (as we must to determine global maxima and minima generally) we find that the unique minimum value 2 of f occurs at $(1,1)$ while the maximum value $13 + 6\sqrt{2}$ occurs at $(\frac{-3}{2}\sqrt{2}, \frac{-3}{2}\sqrt{2})$. Note that functions defined on nonclosed sets can fail to have maxima, minima, or both; for example, the function $f(x,y) = x^2 + y^2$ fails to have even a local maximum or minimum on the "punctured disk" defined by the inequalities $0 < x^2 + y^2 < 1$. We will learn a further technique, known as the method of Lagrange multipliers, next time, to deal with functions defined on level sets not admitting explicit parametrizations.