Lecture 9 : Monday April 20th

jacques@ucsd.edu

Key concepts : Local extrema, Hessian, positive definite, determinant
Know how to find local extrema using first and second derivatives

9.1 Local extremes

For a function $f : \mathbb{R}^n \to \mathbb{R}$ defined on an open set U, a point $x \in U$ is called a local minimum (respectively, maximum) of f if there exists an open ball B containing x such that $f(y) \geq f(x)$ (respectively, $f(y) \leq f(x)$ for all $y \in B$. A local extremum of f is either a local minimum or maximum of f. The point x is a critical point of f if $\nabla f(x) = 0$ or if f is not differentiable at x. As in the case of single variable functions, every local extremum of f is a critical point.

Proposition.

> If $U \subset \mathbb{R}^n$ is an open set, and $f : U \to \mathbb{R}^n$ is differentiable, then if $x \in U$ is a local extremum of f, $\nabla f(x) = 0$.

9.2 Hessian Matrix

In the case of functions of one variable, we test whether a critical point x is a local minimum or maximum or a saddle point of a function f using the second derivative of f. A similar test holds for functions of several variables, but involves matrices. The crucial matrix for the second derivative test on a function $f : U \to \mathbb{R}$ where $f \in C^3(U)$ is the textcolorDarkGreenHessian matrix of f. The Hessian matrix for $f = f(x_1, x_2, \dots, x_n)$ is defined by the following matrix of second order partial derivatives:

$$H_f = \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ f_{x_3x_1} & f_{x_3x_2} & \cdots & f_{x_3x_n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_n} \end{pmatrix}.$$

In other words, the ijth entry of H_f is $\frac{\partial^2 f}{\partial x_i \partial x_j}$ or $f_{x_ix_j}$. Remember that since $f \in C^3(U)$, all corresponding mixed second order partial derivatives are equal, so $f_{x_ix_j} = f_{x_jx_i}$. In particular, this means that H_f is a symmetric matrix. We write $H_f(a)$ for the evaluation of the Hessian of f at a point a.

Recall the definition of the determinant of a square matrix H. The principal submatrices of an $n \times n$ matrix H are the matrices obtained by deleting the last i rows and last i columns of H, for $i = 0, 1, \ldots, n-1$. For example, the principal submatrices of the matrix

$$\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \quad \text{are} \quad \begin{pmatrix} 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Let H_i denote the $i \times i$ principal submatrix of H. The matrix H is positive definite if the determinants of all H_i are positive, and negative definite if $\det(H_1) < 0$, $\det(H_2) > 0$, $\det(H_3) < 0$, and so on (in other words the subdeterminants alternate in sign starting with a $-$ sign). These notions are key to the second derivative test for critical points:

Second Derivative Test.

> Let $U \subset \mathbb{R}^n$ be an open set, and $f : U \to \mathbb{R}^n$ in $C^3(U)$. At a critical point $x \in U$ of f, if $H_f(x)$ is positive definite, then x is a local minimum of f, and if $H_f(x)$ is negative definite, then x is a local maximum of f. If $H_f(x)$ is neither positive nor negative definite, then x is called a saddle point of f.

9.3 Examples.

Example 1. Find all local extremes of the function $f(x, y) = \log(1+x^2+y^2)$ and classify them. First note that

$$\nabla f = (f_x, f_y) = \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2} \right).$$

Therefore $\nabla f = 0$ only at the origin, $(0, 0)$. Therefore $(0, 0)$ is the only critical point of f. Since $f(0, 0) = 0$ and $f(a, b) > 0$ for all $(a, b) \neq (0, 0)$, we conclude that $(0, 0)$ is a local minimum of f, by definition. Alternatively, we could use the second derivative test:

$$\begin{aligned} f_{xx} &= \frac{2(1+x^2+y^2) - 4x^2}{(1+x^2+y^2)^2} \\ f_{yy} &= \frac{2(1+x^2+y^2) - 4y^2}{(1+x^2+y^2)^2} \\ f_{xy} &= -\frac{4xy}{(1+x^2+y^2)^2} = f_{yx} \end{aligned}$$

and at $(0, 0)$ these derivatives are 2, 2 and zero respectively. So the Hessian is

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Since $\det(H_1) = 2$ and $\det(H_2) = 4$, $H_f(0, 0)$ is positive definite and so $(0, 0)$ represents a local minimum.