Review: Multivariable Calculus used in MATH 331

Partial Derivatives

Consider first a real-valued function $f(x,y)$ depending on two variables $\begin{pmatrix} x \\ y \end{pmatrix}$.

- The partial derivative of f with respect to x at a given point $(x_0, y_0)^T$ is defined as
$$\frac{\partial f(x_0,y_0)}{\partial x} = \lim_{h\to 0} \frac{1}{h}\big[f(x_0+h,y_0) - f(x_0,y_0)\big].$$
$\frac{\partial f(x,y)}{\partial x}$ at a generic point $(x,y)^T$ can be calculated using standard differentiation rules from single-variable calculus, with y treated as a constant. Subbing for (x,y) the coordinates of a special point $(x_0,y_0)^T$ gives $\frac{\partial f(x_0,y_0)}{\partial x}$.
- Analogously one defines $\frac{\partial f(x_0,y_0)}{\partial y}$ and $\frac{\partial f(x,y)}{\partial y}$

Example: $$f(x,y) = x^4 + 6x^2y^2 + y^4 - 6x^2 - 12y^2$$
$$\frac{\partial f(x,y)}{\partial x} = 4x^3 + 12xy^2 - 12x, \quad \frac{\partial f(1,1)}{\partial x} = 4$$
$$\frac{\partial f(x,y)}{\partial y} = 12x^2y + 4y^3 - 24y, \quad \frac{\partial f(1,1)}{\partial y} = -8$$

General case

- If f depends on n variables we write $x = (x_1, \ldots, x_n)^T$ and $f(x) = f(x_1, \ldots, x_n)$.
- For $1 \le i \le n$, the partial derivative $\frac{\partial f(x)}{\partial x_i}$ is calculated by taking the derivative with respect to x_i with all x_j, $j \ne i$, treated as constants.

Example: $$f(x_1,x_2,x_3,x_4) = x_1^2x_2 + x_3x_4^2$$
$$\frac{\partial f(x)}{\partial x_1} = 2x_1x_2, \quad \frac{\partial f(x)}{\partial x_2} = x_1^2, \quad \frac{\partial f(x)}{\partial x_3} = x_4^2, \quad \frac{\partial f(x)}{\partial x_4} = 2x_3x_4.$$

Second Order Derivatives [(∗): Under suitable assumptions on f]

- $\frac{\partial^2 f(x,y)}{\partial x^2}$ is the partial derivative of $\frac{\partial f(x,y)}{\partial x}$ with respect to x.
- $\frac{\partial^2 f(x,y)}{\partial y^2}$ is the partial derivative of $\frac{\partial f(x,y)}{\partial y}$ with respect to y.
- $\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial}{\partial x}\frac{\partial f(x,y)}{\partial y} \overset{(*)}{=} \frac{\partial}{\partial y}\frac{\partial f(x,y)}{\partial x} = \frac{\partial^2 f(x,y)}{\partial y \partial x}$.

Generally:

- $\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{\partial}{\partial x_i}\frac{\partial f(x)}{\partial x_i}$
- $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\frac{\partial f(x)}{\partial x_j} \overset{(*)}{=} \frac{\partial}{\partial x_j}\frac{\partial f(x)}{\partial x_i} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$.

For first Example above:
$$\frac{\partial^2 f(x,y)}{\partial x^2} = 12x^2 + 12y^2 - 12, \quad \frac{\partial^2 f(x,y)}{\partial x \partial y} = 24xy, \quad \frac{\partial^2 f(x,y)}{\partial y^2} = 12x^2 + 12y^2 - 24.$$

Higher Order Derivatives

Continuing this way one defines higher order derivatives (if they exist)

$$\frac{\partial^k f(x,y)}{\partial x^m \partial y^n}\ (m+n=k) \quad \text{or} \quad \frac{\partial^k f(x)}{\partial x_1^{k_1}\partial x_2^{k_2}\cdots\partial x_n^{k_n}}\ (k_1+k_2+\cdots+k_n=k).$$

Critical Points and Minima/Maxima of Functions of Two Variables

Definition A critical point of a function $f(x,y)$ is a point $(x_0,y_0)^T$ for which

$$\frac{\partial f(x_0,y_0)}{\partial x} = \frac{\partial f(x_0,y_0)}{\partial y} = 0.$$

Definition $f(x,y)$ has a strict local minimum (maximum) at $(x_0,y_0)^T$ if there is $\epsilon > 0$ such that $f(x,y) > f(x_0,y_0)$ $(f(x,y) < f(x_0,y_0))$ for all $(x,y)^T$ with $0 < \sqrt{(x-x_0)^2+(y-y_0)^2} < \epsilon$.

Theorem 1 (*Necessary condition for a strict local minimum or maximum*) If f has a strict local minimum or maximum at $(x_0,y_0)^T$ then $(x_0,y_0)^T$ is a critical point of f.

Theorem 2 (*Sufficient condition for a strict local minimum or maximum*) If $(x_0,y_0)^T$ is a critical point of f and

$$(1) \quad \frac{\partial^2 f(x_0,y_0)}{\partial x^2} > 0 \quad \Big(\frac{\partial^2 f(x_0,y_0)}{\partial x^2} < 0\Big)$$

$$(2) \quad \frac{\partial^2 f(x_0,y_0)}{\partial x^2}\cdot\frac{\partial^2 f(x_0,y_0)}{\partial y^2} - \Big(\frac{\partial^2 f(x_0,y_0)}{\partial x \partial y}\Big)^2 > 0$$

then f has a strict local minimum (maximum) at $(x_0,y_0)^T$.

Meaning of (1) and (2): Let $Hf(x_0,y_0)$ be the symmetric 2×2-matrix defined by

$$Hf(x_0,y_0) = \begin{pmatrix} \frac{\partial^2 f(x_0,y_0)}{\partial x^2} & \frac{\partial^2 f(x_0,y_0)}{\partial x \partial y} \\ \frac{\partial^2 f(x_0,y_0)}{\partial x \partial y} & \frac{\partial^2 f(x_0,y_0)}{\partial y^2} \end{pmatrix}.$$

This matrix is called the "Hesse-matrix" or "Hessian" of f at $(x_0,y_0)^T$. The conditions (1) and (2) imply that $Hf(x_0,y_0)$ is positive (negative) definite, that is,

$$(\xi,\eta)Hf(x_0,y_0)\begin{pmatrix}\xi\\ \eta\end{pmatrix} = \frac{\partial^2 f(x_0,y_0)}{\partial x^2}\xi^2 + 2\frac{\partial^2 f(x_0,y_0)}{\partial x\partial y}\xi\eta + \frac{\partial^2 f(x_0,y_0)}{\partial y^2}\eta^2 > 0 \ \ (<0)$$

for all (ξ,η) with $\xi^2+\eta^2>0$ (lecture Ch. 4.2-2).

Note: in 2d there are three types of generic critical points:

- a point at which f has a strict local minimum
- a point at which f has a strict local maximum
- a saddle point: $\det Hf(x_0,y_0) = \frac{\partial^2 f(x_0,y_0)}{\partial x^2}\cdot\frac{\partial^2 f(x_0,y_0)}{\partial y^2} - \Big(\frac{\partial^2 f(x_0,y_0)}{\partial x\partial y}\Big)^2 < 0.$

Example: $$f(x,y) = x^4 + 6x^2y^2 + y^4 - 6x^2 - 12y^2$$

Critical point equations: $\frac{\partial f(x,y)}{\partial x} = 4x^3 + 12xy^2 - 12x = 4x(x^2+3y^2-3) = 0$

$\frac{\partial f(x,y)}{\partial y} = 12x^2y + 4y^3 - 24y = 4y(3x^2+y^2-6) = 0$

Solutions:

(1) $(0,0)^T$
(2) $x=0$ and $y^2=6 \Rightarrow (0,\pm\sqrt{6})^T$
(3) $y=0$ and $x^2=3 \Rightarrow (\pm\sqrt{3},0)^T$

(4) If $xy \neq 0 \Rightarrow \left\{ \begin{array}{rcl} x^2+3y^2 & = & 3 \\ 3x^2+y^2 & = & 6 \end{array} \right\} \Rightarrow x^2 = \frac{15}{8},\, y^2 = \frac{3}{8} \Rightarrow (\pm\sqrt{\frac{15}{8}}, \pm\sqrt{\frac{3}{8}})^T$.

To decide which of these critical points is a maximum, minimum or a saddle point, we have to calculate the Hessian at these points using

$$Hf(x,y) = 12 \begin{pmatrix} x^2+y^2-1 & 2xy \\ 2xy & x^2+y^2-2 \end{pmatrix}.$$

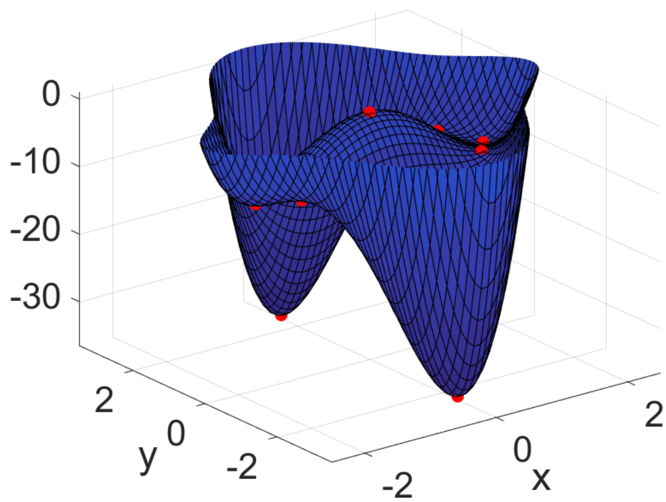
(1) $Hf(0,0) = 12 \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} \Rightarrow (0,0)^T$ is a strict local maximum

(2) $Hf(0,\pm\sqrt{6}) = 12 \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix} \Rightarrow (0,\pm\sqrt{6})^T$ are two strict local minima

(3) $Hf(\pm\sqrt{3},0) = 12 \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow (\pm\sqrt{3},0)^T$ are two strict local minima

(4) $Hf(\pm\sqrt{\frac{15}{8}}, \pm\sqrt{\frac{3}{8}}) = 12 \begin{pmatrix} \frac{5}{4} & \pm\frac{\sqrt{45}}{4} \\ \pm\frac{\sqrt{45}}{4} & \frac{1}{4} \end{pmatrix} \Rightarrow \det Hf(\pm\sqrt{\frac{15}{8}}, \pm\sqrt{\frac{3}{8}}) = 144 \cdot \left(-\frac{40}{16}\right) < 0$

$\Rightarrow (\pm\sqrt{\frac{15}{8}}, \pm\sqrt{\frac{3}{8}})^T$ are four saddle points.



Critical Points and Minima/Maxima in Higher Dimensions

- x_0 is a critical point of $f(x)$ ($x = (x_1, \ldots, x_n)^T$) if $\frac{\partial f(x_0)}{\partial x_i} = 0$ for all $1 \leq i \leq n$.
- $f(x)$ has a strict local minimum (maximum) at x_0 if there is $\epsilon > 0$ such that $f(x) > f(x_0)$ ($f(x) < f(x_0)$) whenever $0 < \|x - x_0\| < \epsilon$ (where $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$).
- If $f(x)$ has a strict local minimum or maximum at x_0 then x_0 is a critical point of f.
- If x_0 is a critical point of $f(x)$ and the (symmetric) Hessian matrix $Hf(x_0)$ of second order partial derivatives, $\big(Hf(x_0)\big)_{ij} = \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}$, is positive (negative) definite, that is, $\xi^T Hf(x_0)\xi > 0$ (< 0) for all $\xi = (\xi_1, \ldots, \xi_n)^T$ with $\|\xi\| > 0$, then f has a strict local minimum (maximum) at x_0.
- A critical point x_0 of $f(x)$ with $\det Hf(x_0) \neq 0$ can be a point at which f has a strict local minimum, a point at which f has a strict local maximum, or an n-dimensional saddle point.