Vector Calculus

II: Higher Order Derivatives - Maxima and Minima

Oliver Zhao

Contents

1 Iterated Partial Derivatives **14**
1.1 Definition of Iterated Partial Derivatives 14
1.2 Mixed Partial Derivatives 14
1.3 Applications with Partial Differential Equations 14

2 Taylor's Theorem **15**
2.1 Single-Variable Taylor Theorem 15
2.2 First-Order Multi-Variable Taylor Theorem 15
2.3 Second-Order Multi-Variable Taylor Theorem 15
2.4 Examples of Taylor Theorem 16

3 Extrema of Real-Valued Functions **17**
3.1 Definitions of Extreme Points 17
3.2 Hessian Matrix 18
3.3 First Derivative Test for Local Extrema 18
3.4 Second Derivative Test for Local Extrema 18

4 Constrained Extrema and Lagrange Multipliers **19**
4.1 The Lagrange Multiplier Method 19
4.2 Multiple Constraints 20
4.3 Global Maxima and Minima 20
4.4 Second-Derivative Test for Constrained Extrema 21

1 Iterated Partial Derivatives

1.1 Definition of Iterated Partial Derivatives

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be of class C^1. This means that $\frac{\delta f}{\delta x}$, $\frac{\delta f}{\delta y}$, and $\frac{\delta f}{\delta z}$ exist and are continuous. If these derivatives, in turn, have continuous partial derivaatives, we say that f is of class C^2, or is twice continuously differentiable, and so forth. Examples are shown below

$$\frac{\delta^2 f}{\delta x^2} = \frac{\delta}{\delta x}\left(\frac{\delta f}{\delta x}\right), \quad \frac{\delta^2 f}{\delta x \delta y} = \frac{\delta}{\delta x}\left(\frac{\delta f}{\delta y}\right), \quad \frac{\delta^2 f}{\delta z \delta y} = \frac{\delta}{\delta z}\left(\frac{\delta f}{\delta y}\right). \tag{1}$$

1.2 Mixed Partial Derivatives

Consider a function $f(x, y)$ of two variables. If the function is of class C^2, then the mixed partial derivatives are equal, or

$$\frac{\delta^2 f}{\delta x \delta y} = \frac{\delta^2 f}{\delta y \delta x}. \tag{2}$$

This can be easily expanded to higher order derivatives and functions with more than two variables.

1.3 Applications with Partial Differential Equations

Consider the following famous partial differential equations:

Heat Equation:	$\frac{\delta f}{\delta t} = \frac{\delta^2 f}{\delta x_1^2} + \frac{\delta^2 f}{\delta x_2^2} + ... + \frac{\delta^2 f}{\delta x_n^2}$
Laplace Equation:	$\frac{\delta^2 f}{\delta x_1^2} + \cdots + \frac{\delta^2 f}{\delta x_n^2} = 0$
Wave Equation:	$\frac{\delta^2 f}{\delta t^2} = \frac{\delta^2 f}{\delta x_1^2} + \cdots + \frac{\delta^2 f}{\delta x_n^2}$

While deriving the solutions to these partial differential equations can be complex, it is a simple matter to use partial derivatives to show why the given solution is valid. For example, consider the Laplace equation below

$$\frac{\delta^2 f}{\delta x^2} + \frac{\delta^2 f}{\delta y^2} = 0 \tag{3}$$

We can show that $f(x, y) = sin(x)sinh(y)$ is a solution by finding the partial derivatives and substituting it into the partial differential equation. The partial derivatives are:

$$\frac{\delta^2 f}{\delta x^2} = -sin(x)sinh(y), \quad \frac{\delta^2 f}{\delta y^2} = sin(x)sinh(y). \tag{4}$$

Consequently, it becomes apparent that $f(x, y) = sin(x)sinh(y)$ is indeed a valid solution for the 2D Laplace Equation.

2 Taylor's Theorem

2.1 Single-Variable Taylor Theorem

For a smooth function $f : \mathbb{R} \to \mathbb{R}$, Taylor's theorem asserts that

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{f"(x_0)}{2}h^2 + \cdots + \frac{f^{(k)}(x_0)}{k!}h^k + R_k(x_0, h), \quad (5)$$

Where

$$R_k(x_0, h) = \int_{x_0}^{x_0+h} \frac{(x_0 + h - \tau)^k}{h^k} f^{k+1}(\tau)d\tau. \quad (6)$$

2.2 First-Order Multi-Variable Taylor Theorem

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\mathbf{x_0} \ \epsilon \ U$. Then

$$f(\mathbf{x_0} + \mathbf{h}) = f(\mathbf{x_0}) + \sum_{i=1}^{n} h_i \frac{\delta f}{\delta x_i}(\mathbf{x_0}) + R_1(\mathbf{x_0}, \mathbf{h}), \quad (7)$$

Where,

$$R_1(\mathbf{x_0}, \mathbf{h}) = \sum_{i,j=1}^{n} \int_0^1 (1 - t)\frac{\delta^2 f}{\delta x_i \delta x_j}(\mathbf{x_0} + t\mathbf{h})h_i h_j dt = \sum_{i,j=1}^{n} \frac{1}{2} \frac{\delta^2 f}{\delta x_i \delta x_j}(\mathbf{c}_{ij})h_i h_j, \quad (8)$$

Where $\mathbf{c}_{ij}$ lies somewhere on the line joining $\mathbf{x_0}$ to $\mathbf{x_0} + \mathbf{h}$, and $R_1(\mathbf{x_0}, \mathbf{h})/||\mathbf{h}|| \to 0$ as $\mathbf{h} \to \mathbf{0}$ in $\mathbb{R}^n$.

2.3 Second-Order Multi-Variable Taylor Theorem

Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\mathbf{x_0} \ \epsilon \ U$. Then

$$f(\mathbf{x_0} + \mathbf{h}) = f(\mathbf{x_0}) + \sum_{i=1}^{n} h_i \frac{\delta f}{\delta x_i}(\mathbf{x_0}) + \frac{1}{2} \sum_{i,j=1}^{n} h_i h_j \frac{\delta^2 f}{\delta x_i \delta x_j}(\mathbf{x_0}) + R_2(\mathbf{x_0}, \mathbf{h}), \quad (9)$$

Where,

$$\begin{aligned} R_2(\mathbf{x_0}, \mathbf{h}) &= \sum_{i,j,k=1}^{n} \int_0^1 \frac{(1-t)^2}{2} \frac{\delta^3 f}{\delta x_i \delta x_j \delta x_k}(\mathbf{x_0} + t\mathbf{h}) h_i h_j h_k dt \\ &= \sum_{i,j,k=1}^{n} \frac{1}{3!} \frac{\delta^3 f}{\delta x_i \delta x_j \delta_k}(\mathbf{c}_{ijk}) h_i h_j h_k, \end{aligned} \tag{10}$$

Where $\mathbf{c}_{ijk}$ lies somewhere on the line joining $\mathbf{x_0}$ to $\mathbf{x_0}+\mathbf{h}$, and $R_2(\mathbf{x_0}, \mathbf{h})/||\mathbf{h}|| \to 0$ as $\mathbf{h} \to \mathbf{0}$ in $\mathbb{R}^n$.

2.4 Examples of Taylor Theorem

Example 1

Find the quadratic approximation of $f(x, y) = cos(x-2y)+sin(x-y)$ at $P(0, 0)$.

$$\begin{array}{lll} f & = cos(x-2y) + sin(x-y) & = 1 \\ f_x & = -sin(x-2y) + cos(x-y) & = 1 \\ f_y & = 2sin(x-2y) - cos(x-y) & = -1 \\ f_{xx} & = -cos(x-2y) - sin(x-y) & = -1 \\ f_{yy} & = -4cos(x-2y) - sin(x-y) & = -4 \\ f_{xy} & = 2cos(x-2y) + sin(x-y) & = 2 \end{array}$$

Thus,

$$\begin{aligned} Q(x, y) &= f + f_x x + f_y y + \frac{f_{xx}}{2} + \frac{f_{yy}}{2} + f_{xy} \\ &= 1 + x - y - \frac{x^2}{2} - 2y^2 + 2xy \end{aligned}$$

Example 2

Suppose $f(x, y, z) = x^2 + y^2 + z^2$. Using the first-order Taylor approximation, find the approximation of $1.01^2 + 2.99^2 + 0.98^2$, where $\mathbf{x_0} = (1, 3, 1)$. Also found the upper boundary of the error. The value of the partial derivatives at $\mathbf{x_0}$ are

$$\begin{array}{lllll} \frac{\delta f}{\delta x} & = 2x & = 2.02 & \frac{\delta^2 f}{\delta x^2} & = 2 \\ \frac{\delta f}{\delta y} & = 2y & = 5.98 & \frac{\delta^2 f}{\delta y^2} & = 2 \\ \frac{\delta f}{\delta z} & = 2z & = 1.96 & \frac{\delta^2 f}{\delta z^2} & = 2 \end{array}$$

Thus,

$$\begin{aligned} f(1.01, 2.99, 0.98) &\approx f(1, 3, 1) + [2(0.01) + 2(-0.01) + 2(-0.02)] \\ &\approx 11 - 0.08 = 10.92 \end{aligned}$$

Where

$$|R_1(\mathbf{x_0}, \mathbf{h})| \leq \frac{1}{2}\left(2(0.01)^2 + 2(-0.01)^2 + 2(-0.02)^2\right) = \frac{6}{10000}.$$

Example 3

Suppose $f(x, y) = tan(\frac{x}{y})$. Using the second-order Taylor approximation, find the approximation of $tan\left(\frac{\pi+0.01}{3.97}\right)$, where $\mathbf{x_0} = (\pi, 4)$. The value of the partial derivatives at $\mathbf{x_0}$ are

$$\begin{aligned}
\frac{\delta f}{\delta x} &= \frac{1}{y} sec^2\left(\frac{x}{y}\right) && = \frac{1}{2} \\
\frac{\delta f}{\delta y} &= -\frac{x}{y^2} sec^2\left(\frac{x}{y}\right) && = \frac{-\pi}{8} \\
\frac{\delta^2 f}{\delta x^2} &= 2\frac{1^2}{y^2} sec\left(\frac{x}{y}\right) sec\left(\frac{x}{y}\right) tan\left(\frac{x}{y}\right) && = \frac{1}{4} \\
\frac{\delta^2 f}{\delta y^2} &= 2\left(\frac{-x}{y^2}\right)^2 sec^2\left(\frac{x}{y}\right) tan\left(\frac{x}{y}\right) + \frac{2x}{y^3} sec^2\left(\frac{x}{y}\right) && = \frac{\pi}{64} + \frac{\pi}{16} \\
\frac{\delta^2 f}{\delta x^2} &= 2\frac{1}{y}\frac{-x}{y^2} 2ec\left(\frac{x}{y}\right) sec\left(\frac{x}{y}\right) tan\left(\frac{x}{y}\right) + \frac{-1}{y^2} sec^2\left(\frac{x}{y}\right) && = -\frac{\pi}{16} - \frac{1}{8}
\end{aligned}$$

Thus,

$$\begin{aligned}
tan\left(\frac{\pi + 0.01}{3.97}\right) &\approx 1 + (0.1)\left(\frac{1}{2}\right) + (-0.03)\left(\frac{-\pi}{8}\right) + \frac{1}{2}\frac{1}{4}(0.01)^2 \\
&+ (0.01)(-0.03)\left(-\frac{\pi}{16} - \frac{1}{8}\right) + \frac{1}{2}(-0.03)^2\left(\frac{\pi^2}{64} + \frac{\pi}{16}\right) \\
&\approx 1.062
\end{aligned}$$

3 Extrema of Real-Valued Functions

3.1 Definitions of Extreme Points

Local Minimum
For a function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, $\mathbf{x_0} \; \epsilon \; U$ is called a local minimum if there is a ball $B = B_\epsilon(\mathbf{x_0}) = \{\mathbf{x} \; \epsilon \; U : ||\mathbf{x} - \mathbf{x_0}\}|| < \epsilon$, with $f(\mathbf{x_0}) \leq f(\mathbf{x})$ for all $\mathbf{x} \; \epsilon \; B$. If $f(\mathbf{x_0}) < f(\mathbf{x})$ for all $\mathbf{x} \; \epsilon \; B$, then $\mathbf{x_0}$ is a strict local minimum.

Local Maximum
For a function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, $\mathbf{x_0} \; \epsilon \; U$ is called a local maximum if there is a ball $B = B_\epsilon(\mathbf{x_0}) = \{\mathbf{x} \; \epsilon \; U : ||\mathbf{x} - \mathbf{x_0}\}|| < \epsilon$, with $f(\mathbf{x_0}) \geq f(\mathbf{x})$ for all $\mathbf{x} \; \epsilon \; B$. If $f(\mathbf{x_0}) > f(\mathbf{x})$ for all $\mathbf{x} \; \epsilon \; B$, then $\mathbf{x_0}$ is a strict local maximum.

Critical Points and Saddle Points
A critical point $\mathbf{x_0}$ of f satisfies $\mathbf{D}f(\mathbf{x_0}) = \mathbf{0}$ or is not differentiable at $\mathbf{x_0}$. If a point $\mathbf{x_0}$ is a critical point but is not a local extremum, it is called a saddle point.

3.2 Hessian Matrix

Suppose that $f : U \subset \mathbb{R}^n \to \mathbb{R}$ has second-order continuous derivatives $(\delta^2 f/\delta x_i \delta x_j)(\mathbf{x_0})$, for $i, j = 1, \ldots, n$ at a point $\mathbf{x_0} \, \epsilon \, U$. The Hessian of f at $\mathbf{x_0}$ is the quadratic function defined by:

$$\begin{aligned} Hf(\mathbf{x_0})(\mathbf{h}) &= \frac{1}{2} \sum_{i,j=1}^{n} \frac{\delta^2 f}{\delta x_i \delta x_j}(\mathbf{x_0}) h_i h_j \\ &= \frac{1}{2}[h_1, \ldots, h_n] \begin{bmatrix} \frac{\delta^2 f}{\delta x_1 \delta x_1} & \cdots & \frac{\delta^2 f}{\delta x_1 \delta x_n} \\ \vdots & & \vdots \\ \frac{\delta^2 f}{\delta x_n \delta x_1} & \cdots & \frac{\delta^2 f}{\delta x_n \delta x_n} \end{bmatrix} \end{aligned} \tag{11}$$

3.3 First Derivative Test for Local Extrema

If $U \subset \mathbb{R}^2$ is open, the function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable, and $\mathbf{x_0} \, \epsilon \, U$ is a local extremum, then $\mathbf{D}f(\mathbf{x_0}) = \mathbf{0}$. That is, $\mathbf{x_0}$ is a critical point of f.

3.4 Second Derivative Test for Local Extrema

Two-Variable Case for Second Derivative Test
Let us define D as the determinant of the Hessian matrix $\mathbf{H}$, or

$$D = f_{xx}(x,y) f_{yy}(x,y) - (f_{xy}(x,y))^2 . \tag{12}$$

Now suppose that (x_0, y_0) is a critical point, that is,

$$\frac{\delta f}{\delta x}(x_0, y_0) = \frac{\delta f}{\delta y}(x_0, y_0) = 0 \tag{13}$$

The second partial derivative test asserts the following:

1. If $D(x_0, y_0) > 0$ and $f_{xx}(x_0, y_0) > 0$, then (x_0, y_0) is a local minimum of f.

2. If $D(x_0, y_0) > 0$ and $f_{xx}(x_0, y_0) < 0$, then (x_0, y_0) is a local maximum of f.

3. If $D(x_0, y_0) < 0$ then (x_0, y_0) is a saddle point of f.

4. If $D(x_0, y_0) = 0$ then the second derivative test is inconclusive. Point (x_0, y_0) can be a minimum, maximum, or saddle point.

Note that in order for there to be a saddle point, either $\left(\frac{\delta^2 f}{\delta x^2}\right) = 0$ or $\left(\frac{\delta^2 f}{\delta y^2}\right) = 0$, with $\left(\frac{\delta^2 f}{\delta x \delta y}\right)^2 \neq 0$.

Many-Variable Case for Second Derivative Test
For a function of f of two or more variables, we examine the eigenvalues of the Hessian matrix instead of the determinant of the Hessian matrix. The following test can be applied at any critical point a for which the Hessian matrix is invertible:

1. If the Hessian has all positive eigenvalues, f has a local minimum at a.

2. If the Hessian has all negative eigenvalues, f has a local maximum at a.

3. If the Hessian has all mixed eigenvalues, f has a saddle point at a.

4 Constrained Extrema and Lagrange Multipliers

Often we are required to maximize or minimize a function within certain constraints or side conditions. This section will develop some methods for handling these types of problems.

4.1 The Lagrange Multiplier Method

Suppose that $f : U \subset \mathbb{R}^n \to \mathbb{R}$ and $g : U \subset \mathbb{R}^n \to \mathbb{R}$ are given C^1 real-valued functions. Let $\mathbf{x_0} \,\epsilon\, U$ and $g(\mathbf{x_0}) = c$, and let S be the level set for g with value c. This means S is the set of points $\mathbf{x} \,\epsilon\, \mathbb{R}^n$ satisfying $g(\mathbf{x})$. Assume $\nabla g(\mathbf{x_0}) \neq \mathbf{0}$.

If $f|S$, which is read at "f is restricted to S", has a local maximum or minimum on S at $\mathbf{x_0}$, then there must be a real number λ (which may be zero) such that the following is true:

$$\nabla f(\mathbf{x_0}) = \lambda \nabla g(\mathbf{x_0}). \tag{14}$$

More generally, if the condition above is held, point $\mathbf{x_0}$ is a critical point. Note that if f, when constrained to a surface S, has a maximum or minimum at $\mathbf{x_0}$, then $\nabla f(\mathbf{x_0})$ is perpendicular to S at $\mathbf{x_0}$.

Based on the condition stated earlier, in order to solve for λ and the other variables, we must solve a systems of equations:

$$\begin{aligned}
\frac{\delta f}{\delta x_1}(x_1,\ldots,x_n) &= \lambda \frac{\delta g}{\delta x_1}(x_1,\ldots,x_n)\\
\frac{\delta f}{\delta x_1}(x_2,\ldots,x_n) &= \lambda \frac{\delta g}{\delta x_2}(x_1,\ldots,x_n)\\
&\cdots\\
\frac{\delta f}{\delta x_n}(x_1,\ldots,x_n) &= \lambda \frac{\delta g}{\delta x_n}(x_1,\ldots,x_n)\\
g(x_1,\ldots,x_n) &= c
\end{aligned} \tag{15}$$

4.2 Multiple Constraints

If a surface S is defined by several constraints, namely

$$\begin{aligned}
g_1(x_1,\ldots,x_n) &= c_1\\
g_2(x_2,\ldots,x_n) &= c_2\\
&\cdots\\
g_k(x_n,\ldots,x_n) &= c_k
\end{aligned} \tag{16}$$

Then the Lagrange multiplier theorem can be generalized. If f has a maximum or minimum at $\mathbf{x_0}$ on S, there must exist constants $\lambda_1,\ldots,\lambda_k$ such that

$$\nabla f(\mathbf{x_0}) = \lambda_1 \nabla g_1(\mathbf{x_0}) + \cdots + \lambda_k \nabla g_k(\mathbf{x_0}). \tag{17}$$

4.3 Global Maxima and Minima

Let U be an open region in $\mathbb{R}^n$ with boundary δU. We say that δU is smooth if δU is the level set of a smooth function g whose gradient ∇g never vanishes (ie. $\nabla g \neq \mathbf{0}$). Then we have the following strategy:

Let f be a differentiable function on a closed and bounded region $D = U \cup \delta U$, U open in $\mathbb{R}^n$, with smooth boundary δU. To find the absolute maximum and minimum of f on D:

(i) Locate all critical points of f in U.

(ii) Use the method of Lagrange multiplier to find the critical points of $f|\delta U$.

(iii) Compute the values of f at all of these critical points.

(iv) Selected the largest and the smallest values.

4.4 Second-Derivative Test for Constrained Extrema

Let $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ and $g : U \subset \mathbb{R}^2 \to \mathbb{R}$ be smooth (at least C^2) functions. Let $g(\mathbf{v_0}) = c$, and S be the level curve for g with value c. Assume that $\nabla g(\mathbf{v_0}) \neq \mathbf{0}$ and that there is a real number λ such that $\nabla f(\mathbf{v_0}) = \lambda \nabla g(\mathbf{v_0})$. Form the auxillary function $h = f - \lambda g$ and the bordered Hessian determinant:

$$|\bar{H}| = \begin{vmatrix} 0 & -\frac{\delta g}{\delta x} & -\frac{\delta g}{\delta y} \\ -\frac{\delta g}{\delta x} & \frac{\delta^2 h}{\delta x^2} & \frac{\delta^2 h}{\delta x \delta y} \\ -\frac{\delta g}{\delta y} & \frac{\delta^2 h}{\delta x \delta y} & \frac{\delta^2 h}{\delta x^2} \end{vmatrix} \text{ evaluated at } \mathbf{v_0}. \tag{18}$$

(i) If $|\bar{H}| > 0$, then $\mathbf{v_0}$ is a local maximum point for $f|S$.

(ii) If $|\bar{H}| < 0$, then $\mathbf{v_0}$ is a local minimum point for $f|S$.

(iii) If $|\bar{H}| = 0$, the test is inconclusive and $\mathbf{v_0}$ may be a minimum, maximum, or neither.