Multivariable Calculus

Aadvik Mohta

aadimohta00@gmail.com

Contents

1 Functions of Two Variables

Definition 1.1. A **function of two variables** is a rule that assigns to each ordered pair of real numbers (x, y) in a set D a unique real number denoted by $f(x, y)$. The set D is the domain of f and its range is the set of values that f takes on, that is, $\{f(x, y)|(x, y) \in D\}$.

Definition 1.2. If f is a function of two variables with domain D, then the graph of f is the set of all points (x, y, z) in $\mathbb{R}^3$ such that $z = f(x, y)$ and $(x, y) \in D$.

Definition 1.3. The level curves of a function f of two variables are the curves with equations $f(x, y) = k$, where $k \in R_f$ is a constant.

Remark. A level curve $f(x, y) = k$ is the set of all points in the domain of f at which f takes on a given value k. Loosely speaking, it shows where the graph has height k.

Remark. The level curves $f(x, y) = k$ are just the traces of the graph of f in the horizontal plane $z = k$ projected down to the xy-plane.

2 Partial Derivatives

Definition 2.1. If f is a function of two variables x and y, its partial derivatives are the functions f_x and f_y defined by

$$f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y)-f(x,y)}{h},$$
$$f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h)-f(x,y)}{h}.$$

Proof. Suppose f is a function of two variables x and y. If we let only x vary while keeping y fixed, such as $y = b$ where b is a constant, then we are effectively considering a function of a single variable in $g(x) = f(x,b)$. If g has a derivative at $x = a$, then we call this derivative the partial derivative of f with respect to x at (a,b), denoted by $f_x(a,b)$. Thus $f_x(a,b) = g'(a)$ where $g(x) = f(x,b)$. By the definition of the derivative, we have

$$g'(a) = \lim_{h\to 0} \frac{g(a+h)-g(a)}{h}.$$

Allowing us to conclude that

$$f_x(a,b) = \lim_{h\to 0} \frac{f(a+h,b)-f(a,b)}{h}.$$

A similar argument ensues for the definition of $f_y(x,y)$.

Notation. If $z = f(x,y)$, we may also write

$$f_x(x,y) = f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x,y) = \frac{\partial z}{\partial x}$$
$$f_y(x,y) = f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} f(x,y) = \frac{\partial z}{\partial y}.$$

2.1 Interpretation of Partial Derivatives

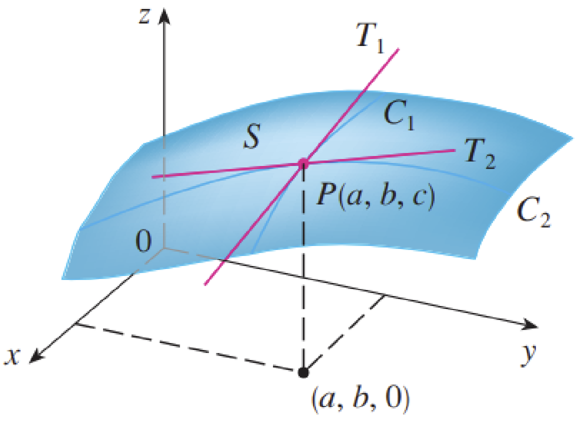

Figure 1: Geometric Interpretation of Partial Derivatives

The equation $z = f(x,y)$ represents a surface S. If $f(a,b) = c$, then the point $P(a,b,c)$ lies on S. By fixing $y = b$, we are restricted to the curve C_1 in which the vertical plane $y = b$ intersects S. Likewise, the vertical plane $x = a$ intersects S in a curve C_2. Both the curves C_1 and C_2 pass through the point

P. The curve C_1 is the graph of the function $g(x) = f(x, b)$, so the gradient of its tangent T_1 at P is thus $g'(a) = f_x(a, b)$. Also, the curve C_2 is the graph of the function $G(y) = f(a, y)$, so the gradient of its tangent T_2 at P is given by $G'(y) = f_y(a, b)$. Therefore, the partial derivatives $f_x(a, b)$ and $f_y(a, b)$ can be interpreted geometrically as the gradients of the tangent lines at $P(a, b, c)$ to the traces C_1 and C_2 of S in the planes $y = b$ and $x = a$ respectively.

Alternatively, partial derivatives may also be interpreted as rates of change. If $z = f(x, y)$, then $\frac{\partial z}{\partial x}$ denotes the rate of change of z with respect to x when y is fixed. A similar definition follows for $\frac{\partial z}{\partial y}$.

2.2 Higher Derivatives

Notation. If f is a function of two variables, then its partial derivatives are also functions of two variables. We write:

$$(f_x)_x = f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 z}{\partial x^2}$$

$$(f_x)_y = f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 z}{\partial y \partial x}$$

$$(f_y)_x = f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 z}{\partial x \partial y}$$

$$(f_y)_x = f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 z}{\partial y^2}.$$

Theorem 2.2 (Clairaut's Theorem). Suppose f is defined on a disk D that contains the point (a, b). If the functions f_{xy} and f_{yx} are both continuous on D, then

$$f_{xy}(a, b) = f_{yx}(a, b).$$

3 The Chain Rule

Theorem 3.1 (Chain Rule Case 1). Suppose that $z = f(x, y)$ is a differentiable function of two variables x and y, where $x = g(t)$ and $y = h(t)$ are both differentiable functions of t. Then z is a differentiable function of t and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Theorem 3.2 (Chain Rule Case 2). Suppose that $z = f(x, y)$ is a differentiable function of two variables x and y, where $x = g(s, t)$ and $y = h(s, t)$ are differentiable functions of s and t. Then

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$$

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

The first case of the the chain rule can be used to derive an important result in the study of implicit functions, known as the Implicit Function Theorem.

Theorem 3.3 (Implicit Function Theorem). Suppose that an equation of the form $F(x, y) = 0$ defines y implicitly as a differentiable function of x. If F is differentiable, then

$$\frac{dy}{dx} = -\frac{F_x}{F_y}.$$

Proof. Suppose $F(x, y)$ is differentiable and the equation $F(x, y) = 0$ defines y implicitly as a differentiable function of x. Let $z = F(x, y) = 0$. Using the Chain Rule, we have

$$F_x \frac{dx}{dx} + F_y \frac{dy}{dx} = \frac{dz}{dx} = 0.$$

$$F_x + F_y \frac{dy}{dx} = 0$$

$$\therefore \frac{dy}{dx} = -\frac{F_x}{F_y}.$$

□

4 Directional Derivatives and the Gradient Vector

Definition 4.1. If f is a differentiable function of x and y, then f has a directional derivative in the direction of any unit vector $\boldsymbol{u} = \begin{pmatrix} a \\ b \end{pmatrix}$ and the directional derivative is defined to be

$$D_{\boldsymbol{u}} f(x, y) = \nabla f \cdot \boldsymbol{u}.$$

If the unit vector $\boldsymbol{u}$ makes an angle θ with the positive x-axis, then $\boldsymbol{u} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ and hence

$$D_{\boldsymbol{u}} f(x, y) = \nabla f \cdot \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$

The directional derivative gives the rate of change of f in any direction $\boldsymbol{u}$, not just the directions $\hat{\boldsymbol{\imath}}$ and $\hat{\boldsymbol{\jmath}}$ for f_x and f_y respectively. The directional derivative thus generalises the notion of a partial derivative.

Note. $\boldsymbol{u}$ should not have a magnitude lest we scale our rate of change by another quantity.

Definition 4.2. If f is a function of two variables x and y, then the gradient of f is the vector function given by

$$\nabla f(x, y) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} = \frac{\partial f}{\partial x}\hat{\boldsymbol{\imath}} + \frac{\partial f}{\partial y}\hat{\boldsymbol{\jmath}}.$$

Theorem 4.3. Suppose f is a differentiable function of two variables. The maximum value of the directional derivative $D_{\boldsymbol{u}} f(\mathbf{x})$ is $|\nabla f(\mathbf{x})|$ and it occurs when $\boldsymbol{u}$ has the same direction as $\nabla f(\mathbf{x})$.

Proof. Note that

$$D_{\boldsymbol{u}} f = \nabla f \cdot \boldsymbol{u} = |\nabla f||\boldsymbol{u}| \cos\theta$$

where θ is the angle between ∇f and $\boldsymbol{u}$. The maximum value of $\cos\theta$ is 1, which occurs when $\theta = 0$. Therefore, the maximum value of $D_{\boldsymbol{u}} f$ is $|\nabla f|$ and it occurs when $\theta = 0$ i.e. $\boldsymbol{u}$ has the same direction as ∇f. □

Note. The function f decreases most rapidly in the direction of $-\nabla f$. The directional derivative in this derivative is $D_{\boldsymbol{u}} f = |\nabla f| \cos\pi = -|\nabla f|$.

Note. Any direction $\boldsymbol{u}$ orthogonal to a gradient $\nabla f \neq \mathbf{0}$ is a direction of zero change in f because then $\theta = 0.5\pi$ and

$$D_{\boldsymbol{u}} f = |\nabla f| \cos\theta = 0.$$

5 Tangent Planes and Normal Lines

5.1 Tangent Plane to a Level Surface of The Form $F(x, y, z) = k$.

Suppose S is a surface with equation $F(x, y, z) = k$, so it is a level surface of a function F of 3 variables. Let $P(x_0, y_0, z_0)$ be a point on S. Then let C be any curve that lies on the surface and passes through the point P. The curve C can be described by the continuous vector function $\boldsymbol{r}(t) = \left\langle x(t), y(t), z(t) \right\rangle$. Let t_0 be the parameter value corresponding to P. Therefore $\boldsymbol{r}(t_0) = \left\langle x(t_0), y(t_0), z(t_0) \right\rangle$. Since C lies on S, any point $(x(t), y(t), z(t))$ must satisfy the equation of S. Therefore,

$$F(x(t), y(t), z(t)) = k.$$

Differentiating with respect to t,

$$\frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt} = 0$$
$$\Rightarrow \nabla F \cdot \boldsymbol{r}'(t) = 0.$$

when $t = t_0$, $\boldsymbol{r}(t_0) = \left\langle x_0, y_0, z_0 \right\rangle$ and thus

$$\nabla F(x_0, y_0, z_0) \cdot \boldsymbol{r}'(t_0) = 0.$$

This allows us to define the tangent vector to the level surface precisely, assuming that

Definition 5.1. Suppose $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$.The tangent plane to the level surface $F(x, y, z) = k$ at $P(x_0, y_0, z_0)$ is the plane that passes through P and has normal vector $\nabla F(x_0, y_0, z_0)$.

The equation of the tangent plane to the level surface is thus

$$\boldsymbol{r} \cdot \boldsymbol{n} = \boldsymbol{a} \cdot \boldsymbol{n}$$
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} F_x(x_0, y_0, z_0) \\ F_y(x_0, y_0, z_0) \\ F_z(x_0, y_0, z_0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} \cdot \begin{pmatrix} F_x(x_0, y_0, z_0) \\ F_y(x_0, y_0, z_0) \\ F_z(x_0, y_0, z_0) \end{pmatrix}$$
$$\Rightarrow (x - x_0)F_x(x_0, y_0, z_0) + (y - y_0)F_y(x_0, y_0, z_0) + (z - z_0)F_z(x_0, y_0, z_0) = 0.$$

5.2 Normal Line to a Level Surface of The Form $F(x, y, z) = k$.

Definition 5.2. The normal line to the level surface $F(x, y, z) = k$ at the point $P(x_0, y_0, z_0)$ is the line that passes through P and is parallel to the gradient vector $\nabla F(x_0, y_0, z_0)$.

The equation of the normal line to the level surface is thus

$$\boldsymbol{r} = \boldsymbol{a} + \lambda \boldsymbol{b}, \quad \lambda \in \mathbb{R}$$
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \lambda \begin{pmatrix} F_x(x_0, y_0, z_0) \\ F_y(x_0, y_0, z_0) \\ F_z(x_0, y_0, z_0) \end{pmatrix}$$
$$x - x_0 = \lambda F_x(x_0, y_0, z_0)$$
$$y - y_0 = \lambda F_y(x_0, y_0, z_0)$$
$$z - z_0 = \lambda F_z(x_0, y_0, z_0)$$
$$\Rightarrow \frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)}.$$

5.3 Tangent Plane and Normal Lineto a Surface $z = f(x, y)$.

Consider the case where $z = f(x, y)$. Then $F(x, y, z) = f(x, y) - z = 0$. Therefore

$$F_x(x_0, y_0, z_0) = f_x(x_0, y_0)$$
$$F_y(x_0, y_0, z_0) = f_y(x_0, y_0)$$
$$F_z(x_0, y_0, z_0) = -1$$

The equation of the tangent plane thus becomes

$$(x - x_0)f_x(x_0, y_0) + (y - y_0)f_y(x_0, y_0) = z - z_0$$

The equation of the normal line also becomes

$$\frac{x - x_0}{f_x(x_0, y_0)} = \frac{y - y_0}{f_y(x_0, y_0)} = z_0 - z.$$

5.4 Tangent Line to a Level Curve

Consider the level curve $f(x, y) = k$. Then $F(x, y) = f(x, y) - k = 0$. By the Implicit Function Theorem, we have

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{f_x(x, y)}{f_y(x, y)}.$$

At the point (x_0, y_0), the equation of the tangent line to the level curve is given by

$$y - y_0 = m(x - x_0)$$
$$y - y_0 = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}(x - x_0)$$
$$(y - y_0)f_y(x_0, y_0) = -(x - x_0)f_x(x_0, y_0)$$
$$\therefore (y - y_0)f_y(x_0, y_0) + (x - x_0)f_x(x_0, y_0) = 0.$$

6 Maximum and Minimum Values

Definition 6.1. A function of two variables has a local maximum at (a, b) if $f(x, y) \leq f(a, b)$ when (x, y) is near (a, b). The number $f(a, b)$ is called a local maximum value. If $f(x, y) \geq f(a, b)$ when (x, y) is near (a, b), then f has a local minimum at (a, b) and $f(a, b)$ is called a local minimum value.

Definition 6.2. If $f(x, y) \leq f(a, b)$ for all points $(x, y) \in D_f$, then f has an absolute maximum at (a, b). If $f(x, y) \geq f(a, b)$ for all points $(x, y) \in D_f$, then f has an absolute minimum at (a, b).

Theorem 6.3. If f has an local maximum or minimum at (a, b) and the first-order partial derivatives of f exist there, then $\nabla f(a, b) = \mathbf{0}$.

Proof. Let $g(x) = f(x, b)$. If f has a local maximum or minimum at (a, b), then g has a local maximum or minimum at $x = a$, so $g'(a) = 0$ by the Interior Extremum Theorem. But $g'(a) = f_x(a, b)$ and hence $f_x(a, b) = 0$. Similarly, let $G(y) = f(a, y)$. If f has a local maximum or minimum at (a, b), then G has a local maximum or minimum at $y = b$, so $G'(b) = 0$ by the Interior Extremum Theorem. But $G'(b) = f_y(a, b)$ and hence $f_y(a, b) = 0$. This allows us to conclude that $\nabla f(a, b) = \begin{pmatrix} f_x(a, b) \\ f_y(a, b) \end{pmatrix} = \mathbf{0}$. □

Theorem 6.3 can be interpreted in a geometrical sense too. Consider the equation of the tangent plane to $z = f(x, y)$ at the point (a, b), that is,

$$z = z_0 + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Substituting $f_x(a, b) = 0$ and $f_y(a, b) = 0$, we get $z = z_0$. This means that if the graph of f has a tangent plane at a local maximum or minimum, then that tangent plane must be horizontal.

Definition 6.4 (Critical Point). A point (a, b) is called a critical point if $f_x(a, b) = 0$ and $f_y(a, b) = 0$, or if one of these partial derivatives does not exist. If f has a local maximum or minimum at (a, b), then (a, b) is a critical point of f. At a critical point, a function could have a local maximum, a local minimum, or neither.

6.1 The Second Partial Derivative Test

Suppose the second partial derivatives of f are continuous on a disk with centre (a, b), and suppose that $\nabla f(a, b) = \mathbf{0}$. We define the discriminant of f at the point (a, b) to be the determinant of the Hessian at (a, b) i.e.

$$D(a, b) = \det \begin{pmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{yx}(a, b) & f_{yy}(a, b) \end{pmatrix} = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2.$$

- If $D > 0$ and $f_{xx}(a, b) > 0$, then $f(a, b)$ is a local maximum.
- If $D > 0$ and $f_{xx}(a, b) < 0$, then $f(a, b)$ is a local minimum.
- If $D < 0$, then $f(a, b)$ is a saddle point i.e it is neither a local maximum nor a local minimum.
- If $D = 0$, then the Second Partial Derivative Test is inconclusive.

6.2 Absolute Maximum and Minimum Values

Definition 6.5 (Closed Set and Bounded Set). A closed set in $\mathbb{R}^2$ is one that contains all its boundary points. A boundary point of D is a point (a, b) such that every disk with centre (a, b) contains points in D and also points not in D. A bounded set in $\mathbb{R}^2$ is one that is contained within some disk and is finite in extent.

Theorem 6.6 (Extreme Value Theorem). If f is continuous on a closed, bounded set D in $\mathbb{R}^2$, then f attains an absolute maximum value $f(x_1, y_1)$ and an absolute minimum value $f(x_2, y_2)$ at some points (x_1, y_1) and (x_2, y_2) in D.

Therefore, we can come up with a "brute force" approach to finding maxima and minima.

1. Find values of f at critical point of f in D.
2. Find extreme values on the boundary,
3. The largest of these values found is the absolute maximum, while the smallest is the absolute minimum.

7 Lagrange Multipliers

Theorem 7.1 (Method of Lagrange Multipliers). To find the extreme values of $f(x, y, z)$ subject to the constraint $g(x, y, z) = k$, assuming such extreme values exist, and that $\nabla g \neq \mathbf{0}$ on the surface $g(x, y, z) = k$, first find all x, y, z, λ such that $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$, before evaluating f at all points (x, y, z) found earlier. The largest of these values is the maximum value of f and the smallest is the minimum value of f.

Proof. Suppose a function $f(x, y, z)$ has an extreme value at the point $P(x_0, y_0, z_0)$ on the surface S and let C be a curve with vector equation $\boldsymbol{r}(t) = \Big\langle x(t), y(t), z(t) \Big\rangle$ that lies on S and passes through P. If t_0 is the parameter value corresponding to the point P, then $\boldsymbol{r}(t_0) = \Big\langle x_0, y_0, z_0 \Big\rangle$. The function $h(t) = f(x(t), y(t), z(t))$ represents the values f takes on the curve C. Since f has an extreme value at $P(x_0, y_0, z_0)$, $h'(t_0) = 0$.

$$\Rightarrow 0 = h'(t_0) = x'(t_0)f_x(x_0, y_0, z_0) + y'(t_0)f_y(x_0, y_0, z_0) + z'(t_0)f_z(x_0, y_0, z_0)$$

$$\Rightarrow \nabla f(x_0, y_0, z_0) \cdot \boldsymbol{r}'(t_0) = 0$$

Hence $\nabla f(x_0, y_0, z_0)$ is orthogonal to $\boldsymbol{r}'(t_0)$. But it is also known that $\nabla g(x_0, y_0, z_0)$ is also orthogonal to $\boldsymbol{r}'(t_0)$ for every such curve. Hence $\nabla f(x_0, y_0, z_0)$ and $\nabla g(x_0, y_0, z_0)$ are parallel. Therefore, if $\nabla g(x_0, y_0, z_0) \neq \mathbf{0}$, then there exists some real λ such that

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0).$$

□

In the case where f and g are functions of two variables, the two equations to solve by the method of Lagrange Multipliers are

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$
$$g(x, y) = k.$$

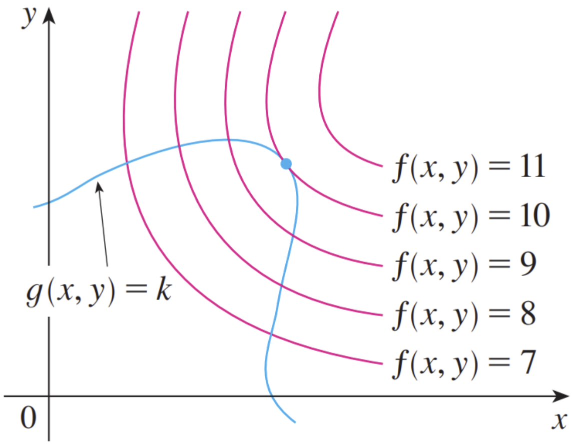

Figure 2: Geometric Intuition for Lagrange Multipliers in the Case of 2 Variables

Intuition. We seek to find the extreme values of $f(x, y)$ when the point (x, y) is restricted to lie on the curve $g(x, y) = k$. To maximise $f(x, y)$ subject to $g(x, y) = k$ is to find the largest value of c such that the level curve $f(x, y) = c$ intersects $g(x, y) = k$. From the diagram, it appears that this happens when the curves have a common tangent line(otherwise the value of c would be increased further). This means the normal lines at the point (x_0, y_0), where the curves touch, are identical i.e. for some scalar λ,

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0).$$

8 Approximations for Functions of 2 Variables

8.1 Linear Approximations

The equation of a tangent plane to a function f of two variables at the point (a, b) is given by

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

The linear function L whose graph is this tangent plane

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is the **linearisation** of f at (a, b) and the approximation

$$f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is the linear approximation of f at (a, b).

8.2 Quadratic Approximations

Theorem 8.1 (Taylor Series in Two Variables). If $f(x,y)$ is an analytic function of two variables, then

$$f(x,y) = \sum_{n,m=0}^{\infty} \left[\frac{1}{n!m!} \frac{\partial^{m+n} f}{\partial x^n \partial y^n}(a,b) \right] (x-a)^n (y-b)^m.$$

Expanding up to the second terms, we get the quadratic approximation to a function $f(x,y)$ at a point (a,b), which is given by

$$Q(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) + 0.5 f_{xx}(a,b)(x-a)^2 + f_{xy}(a,b)(x-a)(y-b) + 0.5 f_{yy}(y-b)^2.$$