Chapter 6

Functions of several variables

6.1 Limits and continuity

Definition 6.1 (Euclidean distance). Given two points $P(x_1, y_1)$ and $Q(x_2, y_2)$ on the plane, we define their distance by the formula

$$||P - Q|| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

Lemma 6.2 (Properties of distance). Each of the following statements is true.

(a) Distance is symmetric: one has $||P - Q|| = ||Q - P||$ for all $P, Q \in \mathbb{R}^2$;

(b) Distance is non-negative: one has $||P - Q|| \geq 0$ with equality if and only if $P = Q$;

(c) Triangle inequality: one has $||P - Q|| \leq ||P - R|| + ||R - Q||$ for all $P, Q, R \in \mathbb{R}^2$.

Definition 6.3 (Limits). Let $f(x, y)$ be a function of two variables and $(x_*, y_*) \in \mathbb{R}^2$ be fixed. If there exists a number L that the values $f(x, y)$ approach as (x, y) approaches (x_*, y_*), then one expresses this fact by writing

$$\lim_{(x,y)\to(x_*,y_*)} f(x, y) = L.$$

More precisely, this equation means that given any $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$0 \neq ||(x, y) - (x_*, y_*)|| < \delta \quad \Longrightarrow \quad |f(x, y) - L| < \varepsilon.$$

If there exists no number L with this property, then we say that the limit does not exist.

Example 6.4. We show that the limit

$$L = \lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x^2 + y^2}$$

does not exist. First of all, let us use polar coordinates to express the given fraction as

$$\frac{x^2 - y^2}{x^2 + y^2} = \frac{r^2 \cos^2\theta - r^2 \sin^2\theta}{r^2 \cos^2\theta + r^2 \sin^2\theta} = \cos^2\theta - \sin^2\theta.$$

If the point (x, y) approaches the origin at an angle of $\theta = 0$, then the last equation gives

$$\frac{x^2 - y^2}{x^2 + y^2} = \cos^2 0 - \sin^2 0 = 1.$$

On the other hand, if (x, y) approaches the origin at an angle of $\theta = \pi/2$, then we get

$$\frac{x^2 - y^2}{x^2 + y^2} = \cos^2(\pi/2) - \sin^2(\pi/2) = -1.$$

Thus, the given function does not really approach any particular value as (x, y) approaches the origin, and this means that the given limit does not really exist.

Example 6.5. We show that the limit

$$M = \lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2 + y^2}$$

is equal to zero. Once again, we use polar coordinates to express the given fraction as

$$f(x, y) = \frac{x^2y^2}{x^2 + y^2} = \frac{r^2\cos^2\theta \cdot r^2\sin^2\theta}{r^2\cos^2\theta + r^2\sin^2\theta} = r^2\cos^2\theta\sin^2\theta.$$

Since (x, y) approaches the origin, we have $r = \sqrt{x^2 + y^2} \to 0$ and so the given function must approach zero as well. More precisely, we have

$$0 \leq f(x, y) = r^2\cos^2\theta\sin^2\theta \leq r^2$$

and the fact that $r^2 \to 0$ implies that $f(x, y) \to 0$ because of the Squeeze Law.

Proposition 6.6 (Properties of limits). Each of the following statements is true.

(a) The limit of a sum/product is equal to the sum/product of the limits, respectively.

(b) When defined, the limit of a quotient is equal to the quotient of the limits.

Definition 6.7 (Special functions). A linear function is one that has the form

$$f(x, y) = Ax + By + C$$

for some constants $A, B, C \in \mathbb{R}$. A polynomial function is one that has the form

$$f(x, y) = \sum_{i,j=0}^{n} a_{ij}\, x^i y^j$$

for some coefficients $a_{ij} \in \mathbb{R}$. Finally, a rational function is the quotient of two polynomials.

Definition 6.8 (Continuity). Let $f(x, y)$ be a function of two variables and $(x_*, y_*) \in \mathbb{R}^2$ be fixed. We say that f is continuous at the point (x_*, y_*), if

$$\lim_{(x,y)\to(x_*,y_*)} f(x, y) = f(x_*, y_*).$$

This means that limits of continuous functions can be computed by simple substitution.

Proposition 6.9 (Continuous functions). Each of the following statements is true.

(a) The sum/product/quotient of two continuous functions is continuous wherever defined.

(b) All linear/polynomial/rational functions are continuous wherever defined.

(c) Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables and let $g\colon \mathbb{R} \to \mathbb{R}$ be a function of a single variable. If f, g are both continuous, then so is their composition $g \circ f\colon \mathbb{R}^2 \to \mathbb{R}$.

6.2 Partial derivatives

Definition 6.10 (Partial derivatives). Given a function $f(x, y)$ of two variables, we define its partial derivative f_x as the derivative of f with respect to x when y is treated as a constant. Its partial derivative f_y is defined similarly by interchanging the roles of x and y.

Lemma 6.11 (Rules of differentiation). The usual rules of differentiation for functions of one variable may still be used to compute partial derivatives for functions of two variables.

Example 6.12. The partial derivatives of $f(x, y) = \sin(x^2 y)$ are given by

$$\begin{aligned} f_x(x, y) &= \cos(x^2 y) \cdot (x^2 y)_x = \cos(x^2 y) \cdot 2xy, \\ f_y(x, y) &= \cos(x^2 y) \cdot (x^2 y)_y = \cos(x^2 y) \cdot x^2. \end{aligned}$$

Theorem 6.13 (Mixed partials). If the mixed partial derivatives f_{xy} and f_{yx} happen to be continuous, then they must also be equal to one another.

Definition 6.14 (Directional derivative). Let $f(x, y)$ be a function of two variables and let $(x_*, y_*) \in \mathbb{R}^2$ be fixed. Given a unit vector $\mathbf{u} = \langle a, b \rangle$, we define the directional derivative of f in the direction of $\mathbf{u}$ as the rate at which f changes in that direction, namely

$$D_{\mathbf{u}} f(x_*, y_*) = a f_x(x_*, y_*) + b f_y(x_*, y_*).$$

Definition 6.15 (Gradient). Given a function $f(x, y)$ of two variables, we define its gradient as the vector $\nabla f(x, y) = \langle f_x, f_y \rangle$. Using this notation, one can then write

$$D_{\mathbf{u}} f(x_*, y_*) = \nabla f(x_*, y_*) \cdot \mathbf{u}.$$

Example 6.16. Let $f(x,y) = 3x^2 - 4xy^2$. When it comes to the point $(1,1)$, we have

$$\nabla f(x,y) = \langle f_x, f_y \rangle = \langle 6x - 4y^2, -8xy \rangle \quad \Longrightarrow \quad \nabla f(1,1) = \langle 2, -8 \rangle.$$

Thus, the directional derivative of f in the direction of the unit vector $\mathbf{u} = \langle 3/5, 4/5 \rangle$ is

$$D_{\mathbf{u}} f(1,1) = \nabla f(1,1) \cdot \mathbf{u} = \frac{3}{5} \cdot 2 - \frac{4}{5} \cdot 8 = -\frac{26}{5}.$$

Remark. For a function $f(x,y,z)$ of three variables, our last two definitions take the form

$$\nabla f = \langle f_x, f_y, f_z \rangle, \qquad D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u}$$

and the vector $\mathbf{u}$ is supposed to be a unit vector as before. To deal with an arbitrary vector, one may simply divide it by its length to turn it into a unit vector.

Theorem 6.17 (Interpretation of gradient). The gradient vector ∇f gives the direction in which f increases most rapidly. Similarly, $-\nabla f$ gives the direction of most rapid decrease.

Theorem 6.18 (Chain rule, version 1). Suppose $f(x,y)$ depends on two variables, each of which depends on a third variable t. Then the derivative of f with respect to t is given by

$$f_t = f_x x_t + f_y y_t.$$

Theorem 6.19 (Chain rule, version 2). Suppose $f(x,y)$ depends on two variables, each of which depends on the variables s,t. Then the partial derivatives f_s and f_t are given by

$$f_s = f_x x_s + f_y y_s, \qquad f_t = f_x x_t + f_y y_t.$$

Remark. Similar versions of the chain rule apply for functions $f(x,y,z)$ of three variables. In that case, the derivative f_s with respect to a variable on which f depends indirectly (a variable other than x,y,z) can be expressed in terms of derivatives with respect to variables on which it depends directly. When $f = f(x,y,z)$, for instance, we have $f_s = f_x x_s + f_y y_s + f_z z_s$.

Example 6.20. Suppose that $u = x^2 y$, where $x = r\cos\theta$ and $y = r\sin\theta$. Then

$$\begin{aligned} u_r &= u_x x_r + u_y y_r = 2xy\cos\theta + x^2\sin\theta, \\ u_\theta &= u_x x_\theta + u_y y_\theta = -2xyr\sin\theta + x^2 r\cos\theta. \end{aligned}$$

Example 6.21. Suppose that $z = z(r,s,t)$, where $r = u + 2v$, $s = 3u$ and $t = 4v$. Then

$$\begin{aligned} z_u &= z_r r_u + z_s s_u + z_t t_u = 1z_r + 3z_s + 0z_t = z_r + 3z_s, \\ z_v &= z_r r_v + z_s s_v + z_t t_v = 2z_r + 0z_s + 4z_t = 2z_r + 4z_t. \end{aligned}$$

6.3 Applications of partial derivatives

Definition 6.22 (Convergence in $\mathbb{R}^2$). We say that a sequence $\{(x_n, y_n)\}$ of points in $\mathbb{R}^2$ is convergent, if the sequences $\{x_n\}$ and $\{y_n\}$ are both convergent.

Theorem 6.23 (Bolzano-Weierstrass in $\mathbb{R}^2$). If a sequence of points in $\mathbb{R}^2$ is bounded, then it has a convergent subsequence.

Definition 6.24 (Closed in $\mathbb{R}^2$). We say that a subset $A \subset \mathbb{R}^2$ is closed, if the limit of every sequence of points in A must itself lie in A. Intuitively speaking, this means that the set A contains its boundary. For example, the unit disk $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ is closed.

Theorem 6.25 (EXTREME VALUE THEOREM). Suppose $f(x, y)$ is continuous on a closed, bounded region. Then f attains both its minimum and its maximum value.

Theorem 6.26 (Location of min/max). Suppose $f(x, y)$ is continuous on a closed, bounded region R. Then the only points at which the min/max values of f may occur are

- points where one of the partial derivatives f_x, f_y does not exist;
- points where $f_x = f_y = 0$ (also known as critical points); and
- points on the boundary of R.

Example 6.27. We find the minimum value of $f(x, y) = 2x^3 - 3x^2 + 9y^2$ over the disk

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 9\}.$$

In this case, both f_x and f_y exist at all points, so we need only check the critical points and the points on the boundary. To find the critical points, we need to solve the equations

$$0 = f_x(x, y) = 6x^2 - 6x = 6x(x - 1), \qquad 0 = f_y(x, y) = 18y.$$

The only critical points are then $(0, 0)$ and $(1, 0)$, while the corresponding values are

$$f(0, 0) = 0, \qquad f(1, 0) = -1.$$

To check the points on the boundary, we note that $y^2 = 9 - x^2$ for all such points, hence

$$f(x, y) = 2x^3 - 3x^2 + 9(9 - x^2) = 2x^3 - 12x^2 + 81$$

and we need to find the minimum value of this function on $[-3, 3]$. Noting that

$$g(x) = 2x^3 - 12x^2 + 81 \quad \Longrightarrow \quad g'(x) = 6x^2 - 24x = 6x(x - 4),$$

we see that the minimum value may only occur at $x = -3$, $x = 3$ or $x = 0$. Since

$$g(-3) = -81, \qquad g(3) = 27, \qquad g(0) = 81,$$

the minimum value of f over the whole region is the value $g(-3) = -81$.

Theorem 6.28 (Local extrema). Suppose that (x_*, y_*) is a critical point of f and that the mixed partials f_{xy}, f_{yx} are continuous at (x_*, y_*). Let H denote the Hessian matrix

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

consisting of all second-order partial derivatives evaluated at the given point.

(a) If H has both positive and negative eigenvalues, then f has a saddle point at (x_*, y_*).

(b) If the eigenvalues of H are all positive, then f has a local minimum at (x_*, y_*).

(c) If the eigenvalues of H are all negative, then f has a local maximum at (x_*, y_*).

Theorem 6.29 (Local extrema in $\mathbb{R}^2$). Let (x_*, y_*) and H be as in the previous theorem.

(a) If $\det H < 0$ at the given point, then f has a saddle point there.

(b) If $\det H > 0$ and $f_{xx} > 0$ at the given point, then f has a local minimum there.

(c) If $\det H > 0$ and $f_{xx} < 0$ at the given point, then f has a local maximum there.

Example 6.30. We classify the critical points of

$$f(x, y) = x^2 - xy + y^2 - 2x - 2y.$$

In order to find these points, we have to solve the equations

$$0 = f_x(x, y) = 2x - y - 2, \qquad 0 = f_y(x, y) = -x + 2y - 2.$$

We multiply the first equation by 2 and then add it to the second equation to get

$$0 = 3x - 6 = 3(x - 2) \quad \Longrightarrow \quad x = 2 \quad \Longrightarrow \quad y = 2x - 2 = 2.$$

This makes $(2, 2)$ the only critical point, while the Hessian at that point is

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Since $\det H = 4 - 1 > 0$ and $f_{xx} = 2 > 0$, the critical point $(2, 2)$ is a local minimum.

Example 6.31. We classify the critical points of

$$f(x, y) = 3xy - x^3 - y^3.$$

In order to find these points, we have to solve the equations

$$0 = f_x(x, y) = 3y - 3x^2, \qquad 0 = f_y(x, y) = 3x - 3y^2.$$

These give $y = x^2$ and also $x = y^2$, so we easily get

$$x = y^2 = x^4 \quad \Longrightarrow \quad x^4 - x = 0 \quad \Longrightarrow \quad x(x^3 - 1) = 0 \quad \Longrightarrow \quad x = 0, 1.$$

Thus, the only critical points are $(0, 0)$ and $(1, 1)$, while the Hessian is given by

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} -6x & 3 \\ 3 & -6y \end{bmatrix}.$$

When it comes to the critical point $(0, 0)$, we get

$$H = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} \quad \Longrightarrow \quad \det H = -9 < 0$$

so this point is a saddle point. When it comes to the critical point $(1, 1)$, we similarly get

$$H = \begin{bmatrix} -6 & 3 \\ 3 & -6 \end{bmatrix} \quad \Longrightarrow \quad \det H = 36 - 9 > 0.$$

Since $f_{xx}(1, 1) = -6 < 0$, however, the critical point $(1, 1)$ is a local maximum.

6.4 Double integrals

Definition 6.32 (Darboux sums). Suppose f is bounded on the rectangle $R = [a, b] \times [c, d]$. Given partitions $P = \{x_0, x_1, \ldots, x_n\}$ and $Q = \{y_0, y_1, \ldots, y_m\}$ of $[a, b]$ and $[c, d]$, respectively, we may then define the lower Darboux sum as

$$S^-(f, P, Q) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} \inf_{R_{kl}} f(x, y) \cdot (x_{k+1} - x_k)(y_{l+1} - y_l),$$

where $R_{kl} = [x_k, x_{k+1}] \times [y_l, y_{l+1}]$. The upper Darboux sum is defined similarly by setting

$$S^+(f, P, Q) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} \sup_{R_{kl}} f(x, y) \cdot (x_{k+1} - x_k)(y_{l+1} - y_l).$$

Definition 6.33 (Integrability). Suppose f is bounded on the rectangle $R = [a, b] \times [c, d]$. If it happens that $\sup S^- = \inf S^+$, then we say that f is integrable over R and we also write

$$\iint_R f(x, y)\, dA = \sup_{P,Q} S^-(f, P, Q) = \inf_{P,Q} S^+(f, P, Q).$$

Theorem 6.34 (Continuous functions are integrable). If a function is continuous on a rectangle R, then it is also integrable over R.

Definition 6.35 (Integrals over general regions). Suppose f is continuous on a closed, bounded region R and let f^* be the function defined by

$$f^*(x,y) = \left\{ \begin{array}{cl} f(x,y) & \text{if } (x,y) \in R \\ 0 & \text{if } (x,y) \notin R \end{array} \right\}.$$

Then the double integral of f over R is defined by the formula

$$\iint_R f(x,y)\,dA = \iint_{R^*} f^*(x,y)\,dA,$$

where R^* is any rectangle which is large enough to contain R.

Theorem 6.36 (Fubini's theorem). Suppose f is continuous on a closed, bounded region R that can be described in two different ways, say

$$R = \{(x,y) \in \mathbb{R}^2 : a \leq x \leq b, \quad g_1(x) \leq y \leq g_2(x)\},$$
$$R = \{(x,y) \in \mathbb{R}^2 : c \leq y \leq d, \quad h_1(y) \leq x \leq h_2(y)\}.$$

Then the double integral of f over R can be computed in two different ways, namely

$$\iint_R f(x,y)\,dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\,dx = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\,dy.$$

Example 6.37. Switching the order of integration, one easily finds that

$$\int_0^1 \int_y^1 e^{x^2}\,dx\,dy = \int_0^1 \int_0^x e^{x^2}\,dy\,dx = \int_0^1 xe^{x^2}\,dx = \left[\frac{e^{x^2}}{2}\right]_0^1 = \frac{e-1}{2}.$$

Example 6.38. We switch the order of integration in order to compute the integral

$$I = \int_0^4 \int_{y/2}^2 \frac{e^{2y/x}}{x}\,dx\,dy = \int_0^2 \int_0^{2x} \frac{e^{2y/x}}{x}\,dy\,dx.$$

In this case, the inner integral is given by

$$\int_0^{2x} \frac{e^{2y/x}}{x}\,dy = \frac{1}{x}\int_0^{2x} e^{2y/x}\,dy = \frac{1}{x}\left[\frac{xe^{2y/x}}{2}\right]_{y=0}^{y=2x} = \frac{e^4-1}{2}$$

and so the double integral is equal to

$$I = \int_0^2 \int_0^{2x} \frac{e^{2y/x}}{x}\,dy\,dx = \int_0^2 \frac{e^4-1}{2}\,dx = e^4 - 1.$$