Analysis in $\mathbb{R}^n$
Math 204, Section 30
Winter Quarter 2008
Paul Sally, e-mail: `sally@math.uchicago.edu`
John Boller, e-mail: `boller@math.uchicago.edu`
website: `http://www.math.uchicago.edu/~boller/M203`

Differentiation

0.1 Review of Differentiation in One Variable

Definition 0.1.1 Let $[a,b]$ be an interval in $\mathbb{R}$ and consider $f:[a,b]\to\mathbb{R}$. We say that f is *differentiable* at a point $x_0\in(a,b)$ if there exists $L\in R$ such that

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=L.$$

Observe that this definition can be phrased in the following way. The function f is differentiable at $x\in(a,b)$ if there exists $L\in\mathbb{R}$ such that

$$\lim_{h\to 0}\frac{f(x+h)-f(x)-Lh}{h}=0.$$

The number L is called the *derivative* of f at x, and is denoted by $f'(x)$ or $Df(x)$.

Exercise 0.1.2 If L exists, show that it is unique.

Exercise 0.1.3 Show that f is differentiable at $x\in(a,b)$ iff there exists a constant L such that

$$\lim_{h\to 0}\frac{|f(x+h)-f(x)-Lh|}{|h|}=0.$$

Theorem 0.1.4 Suppose $f:[a,b]\to\mathbb{R}$ is differentiable at a point $x\in(a,b)$. Then f is continuous at x.

Exercise 0.1.5 Let $f:\mathbb{R}\to\mathbb{R}$ be defined by $f(x)=x^2$ if x is rational and 0 if x is irrational. Show that this function is discontinuous at every nonzero x, and f is differentiable at $x=0$ and nowhere else.

Exercise 0.1.6 Consider the function f_r defined as follows. Let $r\geq 1$ and set

$$f_r(x)=\begin{cases}\frac{1}{q^r} & \text{if } x=\frac{p}{q}\neq 0 \text{ and}\\ 0 & \text{if } x=0 \text{ or } x \text{ is irrational.}\end{cases}$$

i. Show that for any $r \geq 1$, $f_r(x)$ is continuous at 0 and the irrational numbers and is discontinuous at the nonzero rationals.

ii. If $1 \leq r \leq 2$, show that f_r is not differentiable at any irrational point.

iii. For which r is f_r differentiable at $x = 0$?

Exercise 0.1.7 Define

$$f_1(x) = \begin{cases} x & \text{if } x \leq 1/2 \\ 1 - x & \text{if } 1/2 \leq x \leq 1, \end{cases}$$

and extend periodically to $[0, \infty)$ by $f_1(x+1) = f_1(x)$. Then for all $n \geq 2$, define $f_n(x) = \frac{1}{2} f_{n-1}(2x)$. Let $S_m(x) = \sum_{n=1}^{m} f_n(x)$.

i. Show that S_m is a continuous function on $[0, \infty)$ for each $m \in \mathbb{N}$.

ii. Show that the sequence $(S_m)_{m \in \mathbb{N}}$ converges uniformly to a continuous function S.

iii. Show that S is not differentiable at any point in $[0, \infty)$.

Theorem 0.1.8 Suppose $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ are both differentiable at $x \in (a, b)$. Then for any $\alpha \in \mathbb{R}$, $\alpha f + g$ is differentiable at x. Also, the product fg and the quotient $\frac{f}{g}$ are differentiable at x (for $\frac{f}{g}$ we must have $g(x) \neq 0$). We have:

i. $(\alpha f + g)'(x) = \alpha f'(x) + g'(x)$;

ii. $(fg)'(x) = f(x)g'(x) + f'(x)g(x)$;

iii. $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$.

Theorem 0.1.9 (Chain Rule) Let f be differentiable at a point a and let g be differentiable at $f(a)$. Then $g \circ f$ is differentiable at a and $D(g \circ f)(a) = (Dg)(f(a))Df(a)$.

Exercise 0.1.10 Give a critique of the following supposed proof of the chain rule:

$$\begin{aligned} \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x - a} &= \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \frac{f(x) - f(a)}{x - a} \\ &= Dg(f(a))Df(a). \end{aligned}$$

Theorem 0.1.11 Suppose that f satisfies the hypotheses above and f assumes a local maximum or minimum at a point $c \in (a, b)$. Then $f'(c) = 0$.

Theorem 0.1.12 Let $f : [a, b] \to \mathbb{R}$ be continuous and suppose that f is differentiable on (a, b).

i. (Rolle's Theorem) If $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.

ii. (Mean Value Theorem) In any case, there exists $c \in (a, b)$ such that $f'(c)(b - a) = f(b) - f(a)$.

iii. (Generalized Mean Value Theorem) If g satisfies the same hypotheses as f, then there exists $c \in (a, b)$ such that $(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c)$.

Corollary 0.1.13 Suppose f is continuous on $[a, b]$ and differentiable on (a, b). If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant.

Corollary 0.1.14 Suppose f is continuous on $[a, b]$ and differentiable on (a, b). If $f'(c) > 0$ for all $c \in (a, b)$, then f is monotonic increasing.

Corollary 0.1.15 Suppose f is continuous on $[a, b]$ and differentiable on (a, b). If $f'(c) < 0$ for all $c \in (a, b)$, then f is monotonic decreasing.

Corollary 0.1.16 (L'Hôpital's Rule) Let $I = (a, b)$ be any open interval in $\mathbb{R}$ and suppose f and g are differentiable on I. Let $c \in I$, and suppose $\lim_{x \to c} f(x) = \lim_{x \to c} g(x) = 0$ and $\lim_{x \to c} \frac{f'(x)}{g'(x)}$ exists. Then $\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)}$.

Exercise 0.1.17 Prove a similar result for limits as x approaches ∞.

Exercise 0.1.18 Let

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{when } x \neq 0 \\ 0 & \text{when } x = 0. \end{cases}$$

Show that $f'(0)$ exists, but f' is not continuous at 0.

Theorem 0.1.19 (Intermediate value theorem for Derivatives) Let $f : [a, b] \to \mathbb{R}$ be continuous and let f be differentiable on (a, b). Suppose that $(c, d) \subset (a, b)$, $f'(c) < 0$, and $f'(d) > 0$. Then there exists a point $x \in (c, d)$ so that $f'(x) = 0$.

Definition 0.1.20 Let f be a continuous function from (a, b) to $\mathbb{R}$. If k is an integer greater than or equal to 1, we say that $f \in C^k(a, b)$ if f has k derivatives at each point in (a, b) and each of these derivatives is continuous on (a, b). We denote the k-th derivative of f by $f^{(k)}$. In particular, $f^{(0)} = f$. We say that $f \in C^\infty(a, b)$ if f has derivatives of all orders in (a, b). If U is any open set in $\mathbb{R}$, the expressions $C^k(U)$ and $C^\infty(U)$ are defined similarly.

Exercise 0.1.21 Let

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{for } x \neq 0 \\ 0 & \text{when } x = 0 \end{cases}.$$

i. Show that $f \in C^\infty(\mathbb{R})$.

ii. Using L'Hôpital's rule, or anything you wish, show that $f^{(k)}(0) = 0$ for all $k \geq 0$.

Corollary 0.1.22 (Taylor's Theorem) Suppose $f \in C^{k+1}(a, b)$ and $x_0 \in (a, b)$. Then, for any $x \in (a, b)$, we can write

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots \\ &+ \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{1}{(k+1)!} f^{(k+1)}(c)(x - x_0)^{k+1} \end{aligned}$$

where c is a point between x and x_0.

Definition 0.1.23 What Taylor's theorem allows us to do is to approximate a C^{k+1} function in the neighborhood of a point by a polynomial of degree k. Usually the remaining term is reasonably small because of the $(k + 1)!$ in the denominator. The expansion in Taylor's theorem is called the *Taylor expansion of f about x_0*.

Exercise 0.1.24 Find the Taylor expansion of the following functions about the indicated point, to at least 6 terms.

i. $f(x) = \sin(x)$ about $x_0 = \pi$;

ii. $f(x) = \frac{1}{x-1}$ about $x_0 = -1$;

iii. $f(x) = e^{-1/x^2}$ about $x_0 = 0$;

iv. $f(x) = \sqrt{x^2 + 1}$ about $x_0 = 2$.

Theorem 0.1.25 (Liouville's Theorem) Let $\alpha \in \mathbb{R}$ be algebraic of degree $n \geq 2$. Then there exists $C = C(\alpha)$ depending on α such that $|\alpha - p/q| > C/q^n$ for all $p/q \in \mathbb{Q}$.

Exercise 0.1.26 Suppose $n \geq 2$ and α is real algebraic of degree n. Show that if $r > n$, then f_r is differentiable at α.

0.2 Differential Calculus in $\mathbb{R}^n$

Definition 0.2.1 Suppose $U \subset \mathbb{R}^n$ is an open set. A function $f : U \to \mathbb{R}^m$ is *differentiable at* $x \in U$ if there is a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\substack{h \to 0 \\ h \in \mathbb{R}^n}} \frac{|f(x+h) - f(x) - Th|}{|h|} = 0.$$

Notice that the absolute value sign in the numerator denotes the standard Euclidean norm in $\mathbb{R}^m$ while the absolute value sign in the denominator denotes the standard Euclidean norm in $\mathbb{R}^n$. This notation will be used throughout this chapter. We write $T = Df(x)$ and we call this the *derivative* of f at x. We say that f is *differentiable on* U if f is differentiable at each point in U.

Theorem 0.2.2 Suppose U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is differentiable at a point $x_0 \in U$. Then f is continuous at x_0.

Definition 0.2.3 Let $U \subset \mathbb{R}^n$ be an open set, and let $f : U \to \mathbb{R}$ be differentiable on U. Let $x \in U$ and $v \in \mathbb{R}^n$ be a unit vector. The *directional derivative* of f at x in the direction v is defined as

$$D_v f(x) = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}.$$

In the particular case when $v = e_j$, a standard basis vector, we obtain the *partial derivative in the* j*th direction*

$$D_j f(x) = D_{e_j} f(x) = \lim_{t \to 0} \frac{f(x+te_j) - f(x)}{t}.$$

Exercise 0.2.4

i. Let $U \subset \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}^m$ be differentiable. Write $f = (f_1, f_2, \ldots, f_m)$, where $f_k : U \to \mathbb{R}$ is the k^{th} coordinate function of f. Let $v \in \mathbb{R}^n$ be a unit vector. If $x \in U$, define

$$D_v f(x) = \lim_{t \to\ 0} \frac{f(x+tv) - f(x)}{t}.$$

Show that $D_v f(x)$ exists if and only if $D_v f_k$ exists for all k, $1 \le k \le m$, and, in this case, $D_v f(x) = (D_v f_1(x), D_v f_2(x), \ldots, D_v f_m(x))$.

ii. Explain why it is useful for v to be a unit vector.

Remark 0.2.5 Partial derivatives play a special role, as we shall see below. It is worth observing that the classical notation for $D_j f(x)$ is $\frac{\partial f}{\partial x_j}(x)$. All sorts of theorems and properties can be stated much more easily with the notation $D_j f(x)$.

Exercise 0.2.6

i. For a fixed k with $1 \le k \le n$, let $p_k : \mathbb{R}^n \to \mathbb{R}$ be defined by $p_k(x_1, x_2, \ldots, x_n) = x_k$. Show that p_k is differentiable at any point and that $Dp_k = p_k$.

ii. Find $D_v p_k(x)$ for any unit vector v in $\mathbb{R}^n$.

Definition 0.2.7 The map p_k in the above exercise is called the *projection onto the* k^{th} *coordinate*. More generally, if $m \le n$, we can pick indices $1 \le i_1 < i_2 < \cdots < i_m \le n$ and define a projection $p : \mathbb{R}^n \to \mathbb{R}^m$ by $p(x_1, x_2, \ldots, x_n) = (x_{i_1}, x_{i_2}, \ldots, x_{i_m})$.

Exercise 0.2.8 Show that p is differentiable and find its derivative.

Proposition 0.2.9 If U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is differentiable, then the derivative of f is unique.

Proposition 0.2.10 If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map and $x \in \mathbb{R}^n$, then f is differentiable at x and $Df(x) = f$.

Proposition 0.2.11 Let U be an open set in $\mathbb{R}^n$ and $f, g : U \to \mathbb{R}^m$ be differentiable on U. Then $f + g$ is differentiable on U, and $D(f + g)(x) = Df(x) + Dg(x)$ for $x \in U$.

Proposition 0.2.12 If U is an open set in $\mathbb{R}^n$, and $f : U \to \mathbb{R}$ and $g : U \to R$ are real-valued functions which are differentiable on U, then fg is differentiable on U. For $x \in U$, $D(fg)(x) = f(x)Dg(x) + g(x)Df(x)$.

Theorem 0.2.13 (Chain Rule) Suppose that U is an open set in $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ is differentiable on U. Take a point $a \in U$ and let V be an open set in $\mathbb{R}^m$ such that $f(a) \in V$. Suppose $g : V \to \mathbb{R}^p$ is differentiable on V. Then $g \circ f : U \to \mathbb{R}^p$ is differentiable at a, and $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$.

Definition 0.2.14 If $f : \mathbb{R}^n \to \mathbb{R}^m$, we can write $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$. Then, for each k, $1 \leq k \leq m$, the map $f_k : \mathbb{R}^n \to \mathbb{R}$ is called the k-th *component function* of f.

Corollary 0.2.15 If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable, then the component function f_k is differentiable for each k.

Exercise 0.2.16 Show that if each f_k is differentiable on an open set $U \subset \mathbb{R}^n$, then $f = (f_1, f_2, \ldots, f_m)$ is differentiable on U and $Df = (Df_1, Df_2, \ldots, Df_m)$.

Proposition 0.2.17 If $U \subset \mathbb{R}^n$ and $f : U \to \mathbb{R}$ is differentiable, then

$$Df(x) = (D_1 f(x), D_2 f(x), \ldots, D_n f(x)).$$

For $h \in \mathbb{R}^n$,

$$Df(x)h = D_1 f(x)h_1 + D_2 f(x)h_2 + \cdots + D_n f(x)h_n.$$

Exercise 0.2.18 Let $U \subset \mathbb{R}^n$ and $f = (f_1, f_2 \ldots, f_m) : U \to \mathbb{R}^m$ be differentiable. If $x \in U$ and (a_{ij}) is the matrix of $Df(x)$ with respect to the standard basis, then $a_{ij} = D_j f_i(x)$.

Definition 0.2.19 The vector-valued function $(D_1 f(x), D_2 f(x), \ldots, D_n f(x))$ is called the *gradient vector* and is denoted $\nabla f(x)$. Thus $Df(x)(h) = \nabla f(x) \cdot h$.

Exercise 0.2.20 Let $U \subset \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}$ be differentiable. If $x \in U$ and v is a unit vector in $\mathbb{R}^n$, then $D_v f(x) = \nabla f(x) \cdot v$.

Exercise 0.2.21 Let

$$f(x, y) = \begin{cases} 0 & \text{when } (x, y) = (0, 0) \\ \frac{xy}{x^2+y^2} & \text{otherwise.} \end{cases}$$

Show that $D_1 f(0, 0) = D_2 f(0, 0) = 0$, but that f is not continuous at the origin, and hence not differentiable at $(0, 0)$.

Exercise 0.2.22

i. Let

$$f(x, y) = \begin{cases} 0 & \text{when } (x, y) = (0, 0) \\ \frac{x^3}{x^2+y^2} & \text{otherwise.} \end{cases}$$

Show that f has a directional derivative in every direction at every point, but that these directional derivatives are not continuous at the origin.

ii. Let

$$f(x, y) = \begin{cases} 0 & \text{when } (x, y) = (0, 0) \\ (x^2 + y^2) \sin(\frac{1}{\sqrt{x^2+y^2}}) & \text{otherwise.} \end{cases}$$

Show that $D_1 f$ and $D_2 f$ exist at the origin but are not continuous.

Theorem 0.2.23 Let U be an open set in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a function with the property that $D_j f_i$ is continuous on U for $1 \leq i \leq m$, $1 \leq j \leq n$. Then f is differentiable on U and, as we might expect, $Df(x) = (D_j f_i(x))_{i=1,2,\ldots m, j=1,2,\ldots,n}$.

Exercise 0.2.24 Let $f : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$f(x, y, z) = (x^2 y + e^{xz}, \sin(xyz)).$$

Show that f is differentiable on all of $\mathbb{R}^3$ and compute $Df(x, y, z)$.

Theorem 0.2.25 Let U be a open set in $\mathbb{R}^n$, and let $f : U \to \mathbb{R}$ be differentiable. Let x, y be two distinct points in U such that the line segment joining x to y lies entirely in U. Then, there exists $\xi \in (0, 1)$ such that

$$f(y) - f(x) = Df(z)(y - x),$$

where $z = (1 - \xi)x + \xi y$.

Corollary 0.2.26 Let U be a connected open set in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be differentiable. If $Df(x) = 0$ for all $x \in U$, then f is constant on U.

Theorem 0.2.27 Let U be an open subset if $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be differentiable on U. For any two distinct points $x, y \in U$ such that the line segment joining x to y lies entirely in U, and any vector $v \in \mathbb{R}^m$, there exists $\xi \in (0, 1)$ such that

$$v \cdot (f(y) - f(x)) = v \cdot (Df(z)(y - x)),$$

where $z = (1 - \xi)x + \xi y$.

Definition 0.2.28 Consider a function $f = (f_1, f_2, \ldots, f_m)$ from an open set in $\mathbb{R}^n$ to $\mathbb{R}^m$. Then, we define the *mixed partial derivative*

$$D_{ij} f_k = D_i(D_j f_k)$$

assuming that the $D_j f_k$ has partial derivatives for all j, k.

Theorem 0.2.29 Let $U \subset \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}$ be differentiable. Suppose that D_{ij} and D_{ji} exist and are continuous on U. Then $D_{ij} f(x) = D_{ji} f(x)$ for all $x \in U$.

Exercise 0.2.30 Let $f(x, y) = \begin{cases} 0 & \text{when } (x, y) = (0, 0), \\ \frac{x^3 y - x y^3}{x^2 + y^2} & \text{otherwise.} \end{cases}$

Show that f is differentiable. Show that $D_{12} f(0, 0)$, $D_{21} f(0, 0)$ exist, but $D_{12} f(0, 0) \neq D_{21} f(0, 0)$.

In one variable calculus, higher derivatives are defined by differentiating the derivative considered as a function on some set. In the present situation, where we have function $f : \mathbb{R}^n \to \mathbb{R}^m$, the derivative is map from $\mathbb{R}^n$ to the space $L(\mathbb{R}^n, \mathbb{R}^m)$ of linear maps $\mathbb{R}^n \to \mathbb{R}^m$. From our discussion of linear algebra, we know that $L(\mathbb{R}^n, \mathbb{R}^m)$ can be identified with the space $\mathbb{R}^{mn}$. We can then intepret the second derivative $D^2 f$ at a point as an element of

$$L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m)) \cong L(\mathbb{R}^n, \mathbb{R}^{mn}).$$

This space has mn^2 dimensions, and the entries of this matrix are the partial derivatives $D_{ij} f_k$. We can continue this process and define the k^{th} derivative for $k \geq 3$.

On the other hand, it will be useful to talk about all orders of differentiation for the partial dervatives. There is an interesting combinatorial problem that arises here. That is, when we consider the k^{th} partial derivatives, we are considering partitions of k into indices $i_1, i_2, \ldots, i_h$ such that $i_1 + i_2 + \cdots + i_h = k$. Here there are many possibilities for the mixed partials. For example, we might consider whether $D_5(D_2(D_3 f)) = D_2(D_3(D_5 f))$. One might expect, of course, that if one set of indices in a permuation of the other and the mixed partials satisfy an appropriate condition, that his equality holds.

Exercise 0.2.31

i. How may ways can one partition k into an ordered sum of h nonnegative summands?

ii. How may ways can one partition k into h nonnegative summands if the order is ignored?

Definition 0.2.32 A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be C^k if f has all continuous partial derivatives up to and including order k. f is said to be C^∞ or *smooth* if f has all partial derivatives of all orders.

Exercise 0.2.33 If f is C^k and $i_1, i_2, \dots, i_h$ is a set of indices such that $i_1 + i_2 + \cdots + i_h = k$ then for any $\sigma \in S_h$,

$$D_{i_1,i_2,\dots,i_h} f = D_{\sigma(i_1),\sigma(i_2),\dots,\sigma(i_h)} f.$$

Exercise 0.2.34 This is to make sure you understand the notation. If $f : \mathbb{R}^n \to \mathbb{R}$, and $v = (v_1, v_2, \dots, v_n)$, $w = (w_1, w_2, \dots, w_n)$ are vectors in $\mathbb{R}^n$, we know

$$Df(x)(v) = \sum_{j=1}^{n} D_j f(x) v_j.$$

Prove that

$$(D^2 f(x)(v))(w) = \sum_{i,j=1}^{n} D_{ij} f(x) v_i w_j,$$

and similarly in higher dimensions. If $f : \mathbb{R}^n \to \mathbb{R}^m$, we can write the same equations but $D_j f(x)$ is now a vector in R^m being multiplied by a scalar v_j.

Exercise 0.2.35 If $\gamma : [-\delta, 1+\delta] \to \mathbb{R}^n$ is a straight line path such that $\gamma(0) = 0$ and $\gamma(1) = v$, and $f : \mathbb{R}^n \to \mathbb{R}$ is a C^k function, then $D^k f(0)(v) = D^k(f \circ \gamma)(0)(1)$. (Remember that $D^k f(0)(v)$ is an abbreviation for $D^k f(0)(v, \dots, v)$.)

Theorem 0.2.36 Let $U \subset \mathbb{R}^n$ be convex and open, let $f : U \to \mathbb{R}$ be C^{k+1}, and let $x, y \in U$. Then

$$\begin{aligned} f(y) &= f(x) + \sum_{j=1}^{k} \frac{1}{j!} \sum_{i_1,\dots,i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \cdots \partial x_{i_j}}(x)(y_{i_1} - x_{i_1}) \cdots (y_{i_j} - x_{i_j}) \\ &+ \frac{1}{(k+1)!} \sum_{i_1,\dots,i_{k+1}=1}^{n} \frac{\partial^{k+1} f}{\partial x_{i_1} \cdots \partial x_{i_{k+1}}}(z)(y_{i_1} - x_{i_1}) \cdots (y_{i_{k+1}} - x_{i_{k+1}}), \end{aligned}$$

where $z = (1-t_0)x + t_0 y$ for some $t_0 \in (0,1)$.

0.3 Max-Min problems

We now consider the problem of finding maximum and/or minimum values of a function from $\mathbb{R}^n$ to $\mathbb{R}$ using properties of the derivative. Suppose that $f : U \to \mathbb{R}$ is differentiable, where U is an open subset of $\mathbb{R}^n$. Then, for any unit vector v we can consider the directional derivative $D_v f(x)$ for $x \in U$. The directional derivative $D_v f$ measures the rate of change of the function f in the direction v. So it is in fact a derivative of a function of one variable $g(t) = f(x + tv)$, for which we can ask "What is the maximum or minimum value of g on some interval $a \le t \le b$ such that $\{x + tv \mid t \in [a,b]\} \subset U$?" Since g is differentiable, we consider the points where $g'(t) = 0$.

Definition 0.3.1 As in the one-variable case, a point where Df vanishes is called a *critical point* of f, and f need not have either a maximum or minimum at a critical point.

Example 0.3.2 Show that the function $f(x,y) = xy$ has vanishing derivative at the origin, but has neither maximum nor minimum there.

Example 0.3.3 If $f(x,y) = 2xy^2/(x^2+y^4)$, show that f has all partial derivatives equal to zero at the origin, yet is not even continuous there.

Theorem 0.3.4 (Second-derivative test for extrema) If $f : \mathbb{R}^n \to \mathbb{R}$ is C^2 and x is a critical point, let $H(x,v) = D^2 f(x)(v,v)$ ($H(x,\cdot)$ is often called the *Hessian* of f at x). If $H(x,\cdot)$ is positive definite, f has a minimum at x; if $H(x,\cdot)$ is negative definite, f has a maximum at x.

0.4 Lagrange multipliers

There is a method, called the method of *Lagrange multipliers*, for finding relative maxima or minima of a function f, subject to some constraints. The idea is that you have a function $f : \mathbb{R}^n \to \mathbb{R}$, and you have some subset C of $\mathbb{R}^n$ on which you wish to find a minimum of f. This subset is expressed as the set of points where some constraint functions $g_1, \ldots, g_k$ vanish—for instance, the line $y = x$ in the plane has constraint function $g(x,y) = x - y$. Lagrange's method makes use of the observation that the gradient vector to a function is perpendicular to its level surfaces. So if x is a local extremum of f, the level surface of f touches the constraint surface without intersecting it; thus the gradient vector to f must be perpendicular to the constraint surface C. But the space of vectors normal to C is spanned by the gradients of the functions cutting it out, i.e., $g_1, \ldots, g_k$. Hence ∇f is a linear combination of $\nabla g_1, \ldots, \nabla g_k$. This provides a *necessary* (not sufficient!) condition for x to be an extreme point of f.

To make everything work out right, we have to assume that $\nabla g_1, \ldots, \nabla g_k$ are everywhere linearly independent; for the geometric cognoscenti, this assures that C is a *manifold*, that is, doesn't have any sharp corners. Here is the formal statement of the theorem.

Theorem 0.4.1 (Lagrange multipliers) Let $f : \mathbb{R}^m \to \mathbb{R}$ be C^1, and $g_1, \ldots, g_k : \mathbb{R}^m \to \mathbb{R}$ be k C^1 functions, with $k < m$. If x is a local maximum or minimum for f on the common zero set C of $g_1, \ldots, g_k$, and the constraints g_i are independent at x in the sense that $\nabla g_1, \ldots, \nabla g_k$ are everywhere linearly independent, then there exist k numbers $\lambda_1, \ldots, \lambda_k$ satisfying the equation

$$D_x f + \lambda_1 D_x g_1 + \cdots + \lambda_k D_x g_k = 0.$$

0.5 Inverse and implicit function theorems

Suppose that U is an open set in $\mathbb{R}$ and $f : U \to \mathbb{R}$ is C^1. Take a point $x_0 \in U$. We saw earlier in the chapter that if $f'(x_0) \neq 0$, then f is monotonic in an open interval I around x_0. This, of course, implies that f is one to one on I. Moreover, $f(I)$ is an open interval J contained on $\mathbb{R}$ and $f^{-1} : J \to I$ is in C^1 and $(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}$.

Lemma 0.5.1 Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^n$ be C^1. Take $x_0 \in U$ and suppose that $Df(x_0)$ is nonsingular. Then there exists a neighborhood W of x_0 and a constant c such that

$$|f(y) - f(x)| \geq c|y - x| \text{ for all } x, y \in W.$$

Theorem 0.5.2 (Inverse Function Theorem) Let U be an open set in $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^n$ be C^1. Let $x_0 \in U$ such that $Df(x_0)$ is nonsingular. Then there exists a neighborhood W of x_0 such that

i. $f : W \to f(W)$ is a bijection;

ii. $f(W)$ is an open set in $\mathbb{R}^n$;

iii. $f^{-1} : f(W) \to W$ is C^1 and $Df^{-1}(f(x)) = (Df(x))^{-1}$ for $x \in W$.

Theorem 0.5.3 (Implicit Function Theorem) If $f : \mathbb{R}^m \to \mathbb{R}^n$ is C^1, $m \leq n$, and $D_x f$ has maximal rank m, then (after permuting coordinates) the points $(y_1, \ldots, y_n)$ on the graph of f above a neighborhood of x have the form $(y_{m+1}, \ldots, y_n) = g(y_1, \ldots, y_m)$ for a C^1 function g.