In view of Thm 9.1.4, we conclude

Theorem 9.1.7

If $f: \mathbb{R}^n \to \mathbb{R}^m$, all the second order partial derivatives of $f_1, \dots, f_m$ exist and are continuous in an open set G containing a point x, then

$$D_i D_j f_k(x) = D_j D_i f_k(x), \quad \begin{array}{l} 1 \le k \le m \\ 1 \le i, j \le n. \end{array}$$

If the second order partial derivatives are not continuous, then the order of the partial derivatives does matter in general.

—start—

Book practical introduction

§9.2 Classification of Critical Points

We classify the critical points of a function $f: \mathbb{R}^n \to \mathbb{R}^1$.

Definition 9.2.1

Let $A \subset \mathbb{R}^n$ and $f: A \to \mathbb{R}$. A point $x^* \in A$ is a local minimizer of f in A if there is an open neighborhood G of x^* such that for all $x \in A \cap G$

(9.2.1) $$g(x) \geq g(x^*).$$

x^* is a <u>strict</u> <u>minimizer</u> if the inequality (9.2.1) is strict for all $x \in G \cap A$, $x \neq x^*$. If (9.2.1) holds for all $x \in A$, x^* is a <u>global</u> <u>minimizer</u> of f on A.

Of course, we can define analogous terms for a maximizer. But since finding maximizers of f is equivalent to finding minimizers of $-f$, we concentrate on the latter problem.

Critical points arise when we try to find minimizers on open sets.

<u>Definition 9.2.2</u>

Let $G \subset \mathbb{R}^n$ be open and $f: G \to \mathbb{R}$. A point $x^* \in G$ is a <u>critical</u> <u>point</u> of f if f has n partial derivatives at x^* and $D_i f(x^*) = 0$, $1 \leq i \leq n$.

This means that if is differentiable at x^*, then x^* is a critical point if $Df(x^*) = 0$.

Theorem 9.2.1

Let $G \subset \mathbb{R}^n$ be open and $f: G \to \mathbb{R}$. Suppose $x^* \in G$ is a local minimizer of f on G. Then if f has partial derivatives at x^*, x^* is a critical point of f.

In other words, $D_i f(x^*) = 0$, $1 \le i \le n$, if x^* is a local minimizer on an open set.

Proof

Since x^* is a local minimizer,

$$f(x_1^*, x_2^*, \ldots, x_i^* + h, x_{i+1}^*, \ldots, x_n^*) - f(x_1^*, x_2^*, \ldots, x_i^*, \ldots, x_n^*) \ge 0$$

for all h close to zero. So on one hand, we have

$$D_i f(x^*) = \lim_{h \downarrow 0} \frac{f(x_1^*, \ldots, x_i^* + h, \ldots, x_n^*) - f(x_1^*, \ldots, x_n^*)}{h} \ge 0$$

while on the other

$$D_i f(x^*) = \lim_{h \uparrow 0} \frac{f(x_1^*, \ldots, x_i^* + h, \ldots, x_n^*) - f(x_1^*, \ldots, x_n^*)}{h} \le 0.$$

So, $D_i f(x^*) = 0$.

Of course, a local minimizer need not be a critical point.

Example 9.2.1

0 is a minimizer of $f(x)=|x|$.

Also, a critical point need not be a minimizer (or maximizer).

Example 9.2.2

0 is a critical point of $f(x)=x^3$, but is not a minimizer.

Book recall 1-d

The second derivative can help us on the latter issue. We need a multi-dimensional analog of a positive second derivative. In the one dimensional setting, a function that has a zero first derivative and a positive second derivative at a point, looks approximately like a quadratic with a minimum at the point. The higher dimensional analog is a bilinear function with a positive definite matrix.

Definition 9.2.3

An $n \times n$ matrix H is positive semidefinite if

$$(9.2.2) \qquad x^T H x \geq 0 \qquad \text{all } x \in \mathbb{R}^n.$$

It is positive definite if the inequality is strict.

A standard result is

Theorem 9.2.2

Let H be an $n \times n$ matrix.

(1) If H is positive definite (semidefinite), its eigenvalues are real positive (nonnegative).

(2) If H is symmetric and its eigenvalues are positive (nonnegative), it is positive definite (semidefinite).

The first is easy to see, since if $Hv = \lambda v$, $\lambda \in \mathbb{R}$, $v \in \mathbb{R}^n$, then $v^T H v = \lambda \|v\|^2 > 0 \Rightarrow \lambda > 0$.

Theorem 9.2.3

Let $G \subset \mathbb{R}^n$ be open and $f: G \to \mathbb{R}$. Suppose the second derivative $D^2 f(x^*)$ is defined. If x^* is a critical point of f and $H_f(x^*)$ is positive definite, then x^* is a strict local minimizer of f on G. If x^* is a local minimizer, then $H_f(x^*)$ is postive semidefinite.

Example 9.2.3

x^4 has a strict minimizer at 0, but $D^2(x^4) = 12x^2$ is not postive definite at 0.

Example 9.2.4

x^3 has $D^2(x^3) = 6x$, which is positive semidefinite at 0, but x^3 does not have a minimizer at 0.

The proof follows easily from

Theorem 9.2.4

Assume that $f : \mathbb{R}^n \to \mathbb{R}^m$ has a second derivative at $x \in \mathbb{R}^n$. Then

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Df(x)h - \frac{1}{2} D^2 f(x) h h}{\|h\|^2} = 0.$$

We can see that this is close to being a statement about a second order Taylor expansion.

Proof of Thm 9.2.3

Let x^* be a critical point and $H_f(x^*)$ positive definite. For any fixed $h \in \mathbb{R}^n \setminus \{0\}$, Thm 9.2.4 implies

$$\lim_{t \to 0} \frac{1}{t^2} \left(f(x^* + th) - f(x^*) \right) = \frac{1}{2} h^T H_f(x^*) h.$$

Hence, $f(x^* + th) - f(x^*) > 0$ for all sufficiently small t, depending only on $\|h\|$. So, x^* is a strict