if f has a local extremum at a point $p \in \mathbb{R}^n$, then p is a critical point of f. For, if f has a local extremum at $p = (p_1, p_2, \ldots, p_n)$, then for any $i = 1, 2, \ldots, n$ the function ϕ_i defined by $\phi_i(t) = f(p_1, p_2, \ldots, p_{i-1}, t, p_{i+1}, \ldots, p_n)$ has a local extremum at $t = p_i$. So, by an elementary single variable argument

$$\frac{\partial f(p)}{\partial t_i} = \frac{d \phi_i(p_i)}{dt} = 0$$

Thus p is a critical point of f. But critical points are not necessarily local extrema.

The second order partial derivatives of f at a critical point p can be used to test a local extremum at p. These partial derivatives determine a matrix $A(p)$ in which the row i column j entry is

$$\frac{\partial^2 f(p)}{(\partial t_i)(\partial t_j)}$$

This matrix is called the Hessian matrix of f at p. Note that if the third order partial derivatives of f are continuous, then the mixed second order partial derivatives of f at p are independent of the order in which they are taken, and hence $A(p)$ is a symmetric matrix. In this all the eigenvalues of $A(p)$ are real.

Theorem 4.9.1 (The Second derivative test). Let $f(t_1, t_2, \ldots, t_n)$ be a real valued function in n real variables for which all third order partial order derivatives exists and are continuous. Let $p = (p_1, p_2, \ldots, p_n)$ be a critical point of f, and let $A(p)$ be the Hessian matrix of f at p.

(a) If all eigenvalues of $A(p)$ are positive, then f has a local minimum at p.
(b) If all eigenvalues of $A(p)$ are negative, then f has a local maximum at p.
(c) If $A(p)$ has at least one positive and at least one negative eigenvalue then f has no local extremum at p (p is called a saddle point of f)
(d) If $\operatorname{rank}(A(p)) < n$ and $A(p)$ does not have both positive and negative eigenvalues, then the second derivative test is inconclusive.

Proof: If $p \neq \theta = (0,0,\ldots,0)$, we may define a function $g: \mathbb{R}^n \to \mathbb{R}$ by $g(t_1, t_2, \ldots, t_n) = f(t_1+p_1, t_2+p_2, \ldots, t_n+p_n) - f(p)$

The following facts are easily verified (verify it):

1. The function f has a local maximum (minimum) at p if and only if g has a local maximum (minimum) at $\theta = (0,0,\ldots,0)$
2. The partial derivatives at θ are equal to the corresponding partial derivatives of f at p.
3. θ is a critical point of g

~~If $A(p) [a_{ij}]_{n\times n}$ then a_{ij}~~

4. If $A(p) = \left[a_{ij}(p)\right]_{n\times n}$ then $a_{ij}(p) = \dfrac{\partial^2 g(\theta)}{(\partial t_i)(\partial t_j)}$ for all i and j.

In view of these ~~for~~ facts, we may assume without loss of generality that $p = \theta$ and $f(p) = 0$

Now we apply Taylor's theorem to $f(\theta)$ to obtain the first order approximation of f around θ. We have

$$f(t_1, t_2, \ldots, t_n) = f(\theta) + \sum_{i=1}^{n} \frac{\partial f(\theta)}{\partial t_i} t_i + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f(\theta)}{(\partial t_i)(\partial t_j)} t_i t_j + S(t_1, t_2, \ldots, t_n)$$

$$= \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f(\theta)}{(\partial t_i)(\partial t_j)} t_i t_j + S(t_1, t_2, \ldots t_n) \quad \ldots (1)$$

where S is a real valued function on $\mathbb{R}^n$ such that

$$\lim_{x \to \theta} \frac{S(x)}{\|x\|^2} = \lim_{(t_1, t_2, \dots, t_n) \to \theta} \frac{S(t_1, t_2, \dots, t_n)}{t_1^2 + t_2^2 + \dots + t_n^2} = 0 \quad \dots (2)$$

Let $K : \mathbb{R}^n \to \mathbb{R}$ be the quadratic form defined by

$$K\begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f(\theta)}{(\partial t_i)(\partial t_j)} t_i t_j \quad \dots \quad (3)$$

Let H be the symmetric bilinear form corresponding to K, and β be the standard ordered basis for $\mathbb{R}^n$. It is easy to verify (verify it) that $\psi_\beta(H) = \frac{1}{2} A(p)$. Since $A(p)$ is symmetric, our previous result says that there exists an orthogonal matrix Q such that

$$Q^t A(p) Q = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

is a diagonal matrix whose diagonal entries are the eigenvalues of $A(p)$. Let $\gamma = \{v_1, v_2, \dots, v_n\}$ be the orthogonal basis for $\mathbb{R}^n$ whose ith vector is the ith column of Q. Then Q is the change of coordinates matrix changing γ-coordinates into β-coordinates and by Theorem 4.4.9

$$\psi_\gamma(H) = Q^t \psi_\beta(H) Q = \frac{1}{2} Q^t A(p) Q = \begin{pmatrix} \frac{\lambda_1}{2} & 0 & \cdots & 0 \\ 0 & \frac{\lambda_2}{2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{\lambda_n}{2} \end{pmatrix}$$

Suppose that $A(p)$ is not the zero matrix. Then $A(p)$ has non-zero eigenvalues. Choose $\varepsilon > 0$ such that $\varepsilon < |\lambda_i| / 2$

for all $\lambda_i \neq 0$. By (2), $\exists\ \delta > 0$ such that for any $x \in \mathbb{R}^n$ satisfying $0 < \|x\| < \delta$, we have $|S(x)| < \varepsilon \|x\|^2$. Consider any $x \in \mathbb{R}^n$ such that $0 < \|x\| < \delta$. Then by (1) and (3)

$$|f(x) - K(x)| = |S(x)| < \varepsilon \|x\|^2$$

and hence

$$K(x) - \varepsilon \|x\|^2 < f(x) < K(x) + \varepsilon \|x\|^2 \quad \cdots (4)$$

Suppose that $x = \sum_{i=1}^{n} s_i v_i$. Then

$$\|x\|^2 = \sum_{i=1}^{n} s_i^2 \quad \text{and} \quad K(x) = \frac{1}{2} \sum_{i=1}^{n} \lambda_i s_i^2$$

Combining these equations with (4), we obtain

$$\sum_{i=1}^{n} \left(\frac{1}{2}\lambda_i - \varepsilon\right) s_i^2 < f(x) < \sum_{i=1}^{n} \left(\frac{1}{2}\lambda_i + \varepsilon\right) s_i^2 \quad \cdots (5)$$

Now suppose all eigenvalues of $A(p)$ are positive. Then $\frac{1}{2}\lambda_i - \varepsilon > 0$ for all i and hence by the inequality of (5),

$$f(\theta) = 0 \leq \sum_{i=1}^{n} \left(\frac{1}{2}\lambda_i - \varepsilon\right) s_i^2 < f(x)$$

Thus $f(\theta) \leq f(x)$ for $\|x\| < \delta$. So, f has a local minimum at θ. By a similar argument using the right inequality in (5), we have that if all the eigenvalues ~~are negative~~ of $A(p)$ are negative, then f has a local maximum ~~extremum~~ at θ. This establishes (a) and (b) of the Theorem.

Next, suppose that $A(p)$ has both a positive and a negative eigenvalue, say $\lambda_i > 0$ and $\lambda_j < 0$. Then $\frac{1}{2}\lambda_i - \varepsilon > 0$