# Section 3.3

## Extrema of real valued functions.

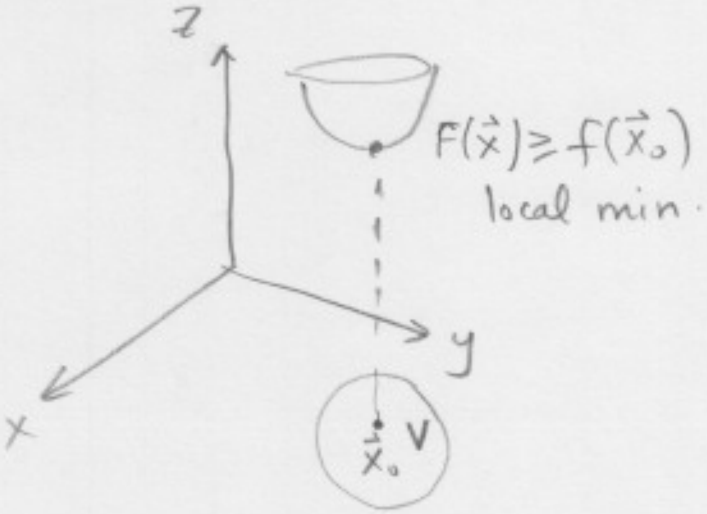
Let $f: U \subset \mathbb{R}^n \to \mathbb{R}$, and $\vec{x}_0 \in U$.

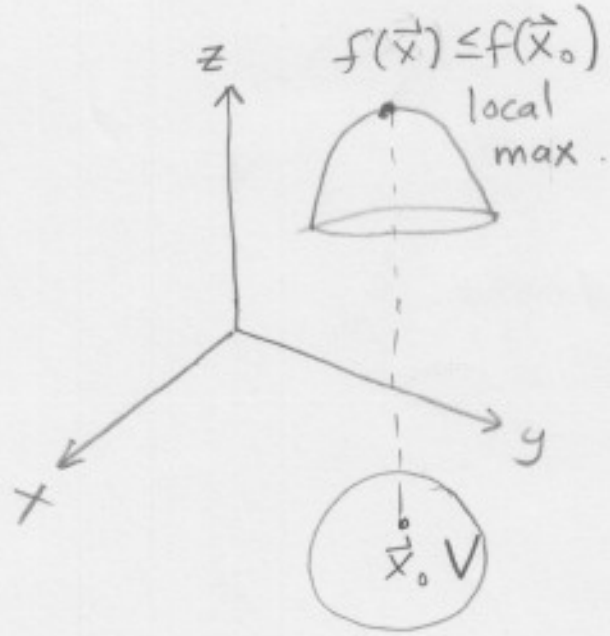
<u>Definition</u>: We say that $\vec{x}_0$ is a local minimum of $f$ if there is a neighborhood $V$ of $\vec{x}_0$ such that:

$$f(\vec{x}) \geq f(\vec{x}_0), \text{ for all } \vec{x} \in V.$$

We say that $\vec{x}_0$ is a local maximum if

$$f(\vec{x}) \leq f(\vec{x}_0), \text{ for all } \vec{x} \in V.$$




These are called local (or relative) extrema.

<u>Definition</u>: The point $\vec{x}_0$ is a critical point for $f$, if either $f$ is not differentiable at $\vec{x}_0$, or, if it is, $\nabla f(\vec{x}_0) = \vec{0}$.

We have the following:

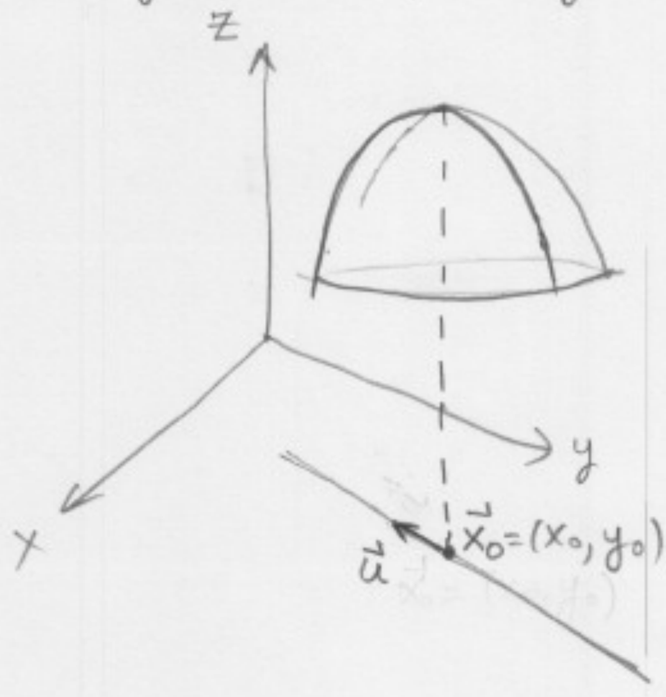
<u>Theorem</u>: If $f$ has a local extrema at $\vec{x}_0$, then $\vec{x}_0$ is a critical point; that is, $\nabla f(\vec{x}_0)=\vec{0}$.

In order to see that this theorem is true, we assume, without loss of generality, that $f$ has a local maximum at $\vec{x}_0$.

We let $\vec{u}$ be unit vector. We form the composition:

$$h(t)=f(\vec{x}_0+t\vec{u})$$

A picture in 3 dimensions is as follows (in this case $\vec{u}=(u_1,u_2)$ and $\vec{x}_0+t\vec{u}$ is a line in the $xy$-plane passing through $\vec{x}_0=(x_0,y_0)$ )



Notice that $h: \mathbb{R} \to \mathbb{R}$ has a local maximum at $t=0$. Therefore:

$$h'(0) = 0.$$

From chain rule:

$$h'(t) = \nabla f(\vec{x}_0 + t\vec{u}) \cdot \vec{u}$$

$$\Rightarrow h'(0) = \nabla f(\vec{x}_0) \cdot \vec{u} = 0$$

Since $\nabla f(\vec{x}_0)$ is orthogonal to $\vec{u}$, and $\vec{u}$ is arbitrary (i.e, the same argument holds if we choose a different $\vec{u}$), it follows that $\nabla f(\vec{x}_0)$ is orthogonal to all unit vectors $\vec{u}$, but this can only be true if

$$\nabla f(\vec{x}_0) = \vec{0}. \quad \blacksquare$$

We have shown that, at a local max or min, if $f$ has a tangent plane, it must be horizontal.

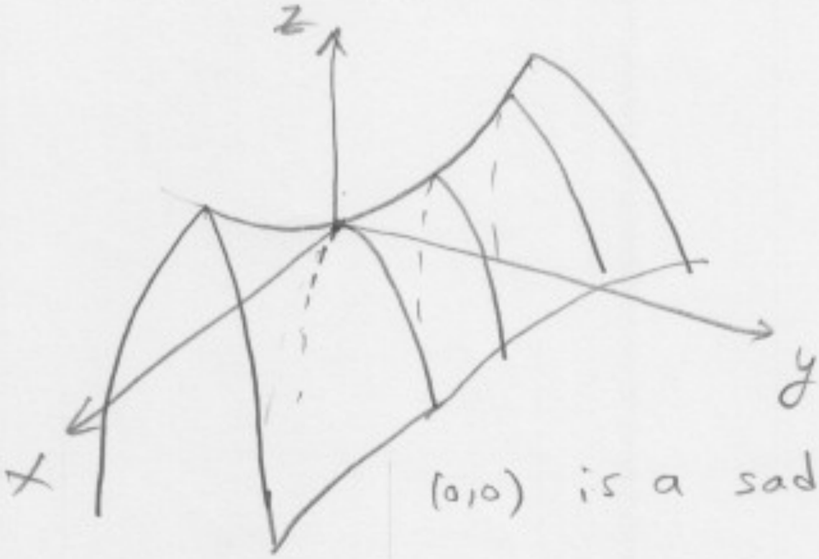
The converse of the previous Theorem is not true; that if, if $\nabla f(\vec{x}_0) = \vec{0}$, for some $\vec{x}_0$, this does not imply that $\vec{x}_0$ has to be a local extrema. For example, consider the function:

$$f(x,y) = x^2 - y^2$$

$$\Rightarrow \nabla f = (2x, -2y)$$

$$\nabla f(0,0) = (0,0) = \vec{0},$$

but $(0,0)$ is not a local extrema:




$(0,0)$ is a saddle point.

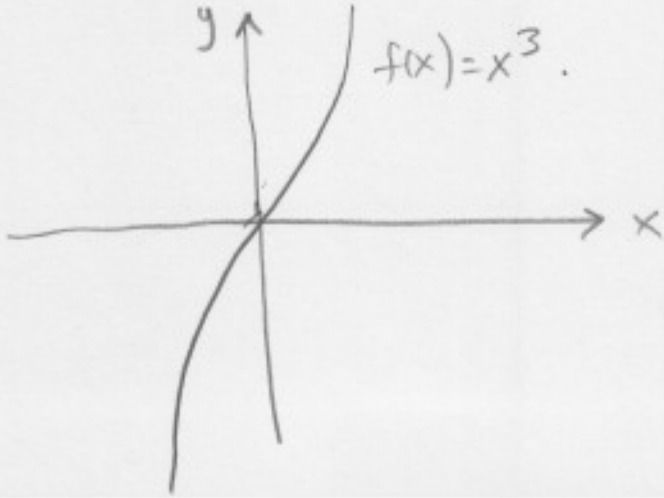
In one-dimension, consider:

$$f(x) = x^3$$

$$f'(x) = 3x^2$$

$$f'(0) = 0,$$

but $0$ is not a local max or min:



What the Theorem tell us is that, in order to find local max. and min. we must find all the critical points that solve:

$$\nabla f(\vec{x}) = 0.$$

Then, the points where there are local extrema (if they exist) must be found among the critical points.

Ex: Find the critical points of

$$z = (x^2 - y^2)\, e^{\frac{-x^2-y^2}{2}}.$$

$$\frac{\partial z}{\partial x} = 2x\, e^{\frac{-x^2-y^2}{2}} - x\, e^{\frac{-x^2-y^2}{2}}(x^2-y^2)$$

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

$$\frac{\partial z}{\partial y} = -2y\, e^{\frac{-x^2-y^2}{2}} - y\, e^{\frac{-x^2-y^2}{2}}(x^2-y^2)$$

$$= e^{\frac{-x^2-y^2}{2}} y\,(-2 - x^2 + y^2) = 0.$$

We need to solve (since the exponential is never zero) the following 2 equations:

$$x\,(2-x^2+y^2)=0 \rightarrow (1)$$
$$y\,(-2-x^2+y^2)=0 \rightarrow (2)$$

Clearly:

$\boxed{(0,0) \text{ is a critical point}}$

If $x=0$, $y\neq 0$, equation (1) is satisfied and, in order to satisfy equation (2) we request:

$$-2-x^2+y^2=0$$
$$\Rightarrow -2+y^2=0 \;;\ \text{since } x=0$$
$$\Rightarrow y=\pm\sqrt{2}$$

$\boxed{\text{Hence, } (0,\sqrt{2}) \text{ and } (0,-\sqrt{2}) \text{ are critical points}}$

If $x\neq 0$, $y=0$, equation (2) is satisfied. In order to satisfy equation (1) we impose:

$$2-x^2+y^2=0$$
$$\Rightarrow 2-x^2=0 \;;\ \text{since } y=0$$
$$\Rightarrow x=\pm\sqrt{2}$$

$\boxed{\text{Hence, } (\sqrt{2},0),\ (-\sqrt{2},0) \text{ are critical points}}$

There are all the points that solve $\nabla f(x,y)=\vec{0}$.

Second derivative test to find local max/min for functions of one variable $f:\mathbb{R}\to\mathbb{R}$, $f$ of class $C^2$.

Let $f:\mathbb{R}\to\mathbb{R}$ and $x_0\in\mathbb{R}$ a critical point for $f$, that is, $f'(x_0)=0$. Suppose that $f''(x_0)>0$. We will show next that $f$ has a local min. at $x=x_0$.

Since $f$ is $C^2$, we can write the Taylor's formula for $f$:

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + R_2(x)$$

where:

$$\lim_{x\to x_0} \frac{R_2(x)}{(x-x_0)^2} = 0 \quad \to (*)$$

From $(*)$ it follows, using the definition of limit, that for every $\varepsilon>0$, there exists a $\delta>0$ such that:

If $|x-x_0|<\delta$ then $\dfrac{|R_2(x)|}{(x-x_0)^2}<\varepsilon$

Hence, if $x$ is close enough to $x_0$, we have:

$$|R_2(x)| < \varepsilon (x-x_0)^2 \quad \to (**)$$

Since $f'(x_0)=0$, the Taylor's formula simplifies to:

$$f(x) = f(x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + R_2(x)$$

or

$$(A) \quad \boxed{f(x) - f(x_0) = \frac{1}{2} f''(x_0)(x-x_0)^2 + R_2(x)}$$

Choosing, for example, $\varepsilon = \frac{1}{4} f''(x_0)$, (recall that $f''(x_0) > 0$) we have from (**):

$$(B) \quad \boxed{|R_2(x)| < \frac{1}{4} f''(x_0)(x-x_0)^2} \; ; \text{ if } x \text{ is close enough to } x_0$$

From (A) and (B) we have:

$$\frac{1}{2} f''(x_0)(x-x_0)^2 + R_2(x) > \frac{1}{2} f''(x_0)(x-x_0)^2 - \frac{1}{4} f''(x_0)(x-x_0)^2$$

$$= \frac{3}{4} f''(x_0)(x-x_0)^2$$

$$> 0.$$

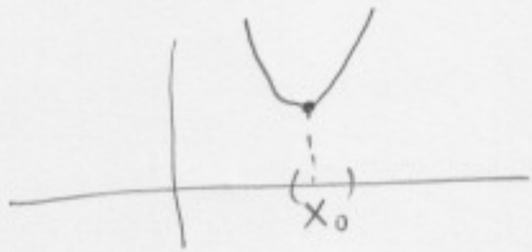
That is, even if $R_2(x)$ is negative, this term is absorbed by the term $\frac{1}{2} f''(x_0)(x-x_0)^2$ and the sum is still positive, if $x$ is close enough to $x_0$. Therefore; from (A):

$$f(x) - f(x_0) > 0 \quad \text{for } x \text{ close enough to } x_0$$

and hence $f$ has a local min at $x_0$.

Again, we have shown that if $f'(x_0)=0$ and $f''(x_0)>0$, $f$ has a local min at $x_0$ since

$$f(x) > f(x_0) \text{ for } x \text{ close to } x_0$$




We can proceed with an analogous argument to show that:

If $f'(x_0)=0$ and $f''(x_0)<0$, then $f$ has a local max at $x=x_0$. Indeed we would have:

$$f(x)-f(x_0) = \frac{1}{2}f''(x_0)(x-x_0)^2 + R_2(x)$$

If $f''(x_0)<0$, since $|R_2(x)| < \varepsilon(x-x_0)^2$, where $\varepsilon$ can be made arbitrarily small, the term $R_2(x)$ is "absorbed" in $\frac{1}{2}f''(x_0)(x-x_0)^2$ and the sum remains negative, that is:

$$f(x)-f(x_0) < 0$$

or $f(x) < f(x_0)$, for $x$ close enough to $x_0$

$\Rightarrow$ $f$ has a local max. at $x_0$.

Second derivative test to find local max/min for functions of two variables $f: \mathbb{R}^2 \to \mathbb{R}$.

Suppose that $f$ is of class $C^2$ and that $(x_0, y_0)$ is a critical point; that is, $\nabla f(x_0, y_0) = 0$. From Taylor's theorem:

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)\,(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)\,(y-y_0)$$

$$+ \frac{1}{2}\Big(f_{xx}(x_0,y_0)\,(x-x_0)^2 + 2f_{xy}(x_0,y_0)\,(x-x_0)(y-y_0)$$

$$+ f_{yy}(x_0,y_0)\,(y-y_0)^2\Big) + R_2(x,y),$$

where $$\lim_{(x,y)\to(x_0,y_0)} \frac{R_2(x,y)}{\|(x,y)-(x_0,y_0)\|^2} = 0.$$

Note: The hypothesis $C^2$ is enough for our analysis. $f \in C^3$ would be necessary if we had to use a formula for $R_2(x,y)$.

Since $\nabla f(x_0, y_0) = 0$, Taylor's Theorem reduces to:

$$f(x,y) - f(x_0,y_0) = \underbrace{Q(x,y)}_{\text{quadratic term}} + R_2(x,y)$$

To find out if $f(x,y)$ is above or below $f(x_0,y_0)$ we must investigate the quadratic term (if the quadratic term is 0 we need to move to the third order approximation, etc).

We will se that, using the same type of argument as for $f:\mathbb{R}\to\mathbb{R}$, if $Q(x,y)$ is always positive, then $(x_0,y_0)$ is a local min. If $Q(x,y)$ is always negative, $(x_0,y_0)$ is a local max, and if $Q(x,y)$ can change sign, $(x_0,y_0)$ is a saddle point.

We let:

$$a_{11} := f_{xx}(x_0,y_0),\ a_{22} := f_{yy}(x_0,y_0)$$
$$a_{12} := f_{xy}(x_0,y_0) = f_{yx}(x_0,y_0)$$
$$h_1 := x-x_0,\quad h_2 := y-y_0.$$

Hence, $Q(x,y)$ can be rewritten as a function of $h_1, h_2$, say $g(h_1,h_2)$:

$$g(h_1,h_2) = \frac{1}{2}\left(a_{11}h_1^2 + 2a_{12}h_1h_2 + a_{22}h_2^2\right).$$

We need to investigate conditions on $a_{11}$, $a_{12}$ and $a_{22}$

<u>Definition</u>: A quadratic function is a function $g(h_1, \dots, h_n) = \sum_{i,j=1}^{n} a_{ij} h_i h_j$.

(a) $g$ is positive definite if $g > 0$ for any $(h_1, \dots, h_n) \neq 0$

(b) $g$ is negative definite if $g < 0$ for any $(h_1, \dots, h_n) \neq 0$

(c) $g$ is indefinite if it can change sign.

<u>Ex</u>: $g(h_1, h_2) = h_1^2 + h_2^2$, positive definite
$g(h_1, h_2) = -h_1^2 - h_2^2$, negative definite
$g(h_1, h_2) = h_1^2 - h_2^2$, indefinite

In a more advance class in linear algebra, one can derive many tests on quadratic functions to determine if they are positive definite, negative definite, or indefinite.

For $n=2$ (our situation) the test is easier. Indeed, notice that $g(h_1, h_2)$ can be rewritten as:

$$g(h_1, h_2) = \frac{1}{2}\left(a_{11}h_1^2 + 2a_{12}h_1h_2 + a_{22}h_2^2\right)$$
$$= \frac{1}{2}a_{11}\left(h_1 + \frac{a_{12}}{a_{11}}h_2\right)^2 + \frac{1}{2}\left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)h_2^2$$

From this formula we can easily check that:

$g$ is pos. def. $\Longleftrightarrow a_{11} > 0$ and $a_{11}a_{22} - a_{12}^2 > 0$

$g$ is neg. def. $\Longleftrightarrow a_{11} < 0$ and $a_{11}a_{22} - a_{12}^2 > 0$

indefinite $\Longleftrightarrow a_{11}a_{22} - a_{12}^2 < 0$

Therefore, we have:

## Second derivative test

(i) $f$ has a local max at $(x_0, y_0)$ if $f_{xx}(x_0, y_0) < 0$ and $D(x_0, y_0) > 0$, where:

$$D(x_0, y_0) = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}$$

(ii) $f$ has a local min at $(x_0, y_0)$ if

$$f_{xx}(x_0, y_0) > 0 \quad \& \quad D(x_0, y_0) > 0$$

(iii) $f$ has a saddle point at $(x_0, y_0)$ if $D(x_0, y_0) < 0$

(iv) If $D(x_0, y_0) = 0$ we do not have any information.

We go back to our example:

<u>Ex</u>: $z = (x^2 - y^2)\, e^{-\frac{x^2 - y^2}{2}}$

We computed earlier the critical point of $f$ which are:

$(0,0)$, $(\sqrt{2}, 0)$, $(-\sqrt{2}, 0)$, $(0, \sqrt{2})$, $(0, -\sqrt{2})$

We need:

$$\frac{\partial^2 f}{\partial x^2} = \left[2 - 5x^2 + x^2(x^2 - y^2) + y^2\right] e^{-\frac{x^2 - y^2}{2}}$$

$$\frac{\partial^2 f}{\partial x \partial y} = xy\,(x^2 - y^2)\, e^{-\frac{x^2 - y^2}{2}}$$

$$\frac{\partial^2 f}{\partial y^2} = \left[5y^2 - 2 + y^2(x^2 - y^2) - x^2\right] e^{-\frac{x^2 - y^2}{2}}$$

| Point | $f_{xx}$ | $f_{xy}$ | $f_{yy}$ | D | Type |
|---|---|---|---|---|---|
| $(0,0)$ | $2$ | $0$ | $-2$ | $-4$ | Saddle point |
| $(\sqrt{2}, 0)$ | $-4/e$ | $0$ | $-4/e$ | $16/e^2$ | local max |
| $(-\sqrt{2}, 0)$ | $-4/e$ | $0$ | $-4/e$ | $16/e^2$ | local max |
| $(0, \sqrt{2})$ | $4/e$ | $0$ | $4/e$ | $16/e^2$ | local min |
| $(0, -\sqrt{2})$ | $4/e$ | $0$ | $4/e$ | $16/e^3$ | local min |