March 1, 2021

2. UNCONSTRAINED OPTIMIZATION

All throughout this section, D denotes an **open subset** of $\mathbb{R}^n$.

1. First order necessary condition

Proposition 1.1. Let $f : D \to \mathbb{R}$ be differentiable. If $x_0 \in D$ is a local maximum or a local minimum of f on D, then

$$\nabla f(x_0) = 0$$

Proof Fix $i = 1 \ldots, n$ and consider the curve

$$g(t) = f(x_0 + te_i)$$

where $\{e_1 \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$. Note that g is a 1-variable differentiable function that attains a local maximum at $t_0 = 0$. Hence,

$$g'(0) = 0$$

But,

$$g'(0) = \frac{d}{dt}\bigg|_{t=0} f(x_0 + te_i) = \lim_{x \to 0} \frac{f(x_0 + te_i - f(p)}{t} = \frac{\partial f}{\partial x_i}(x_0)$$

Definition 1.2. Let $f : D \to \mathbb{R}$ we say that $x_0 \in D$ is a **critical point** if either f is not differentiable at x_0 or if

$$\nabla f(x_0) = 0.$$

Remark 1.3. If x_0 is a local extremum of f, then x_0 is a critical point of f.

Definition 1.4. If $\nabla f(x_0) = 0$, but x_0 is not a local extremum of f, then x_0 is a **saddle point**.

2. Second order necessary conditions

Proposition 2.1. Let $f : D \to \mathbb{R}$ be of class $C^2(D)$. Fix a point $x_0 \in D$.

(1) If x_0 is a local maximum of f on D, then the Hessian matrix $\mathrm{H}\, f(x_0)$ is negative semidefinite or negative definite.
(2) If x_0 is a local minimum of f on D, then the Hessian matrix $\mathrm{H}\, f(x_0)$ is positive semidefinite or positive definite.

3. Second order sufficient condition

Proposition 3.1. Let $f : D \to \mathbb{R}$ be of class $C^2(D)$. Fix a point $x_0 \in D$ and suppose

$$\nabla f(x_0) = 0.$$

We have,

(1) If $\mathrm{H}\, f(x_0)$ is negative definite, then x_0 is a (strict) local maximum of f.
(2) If $\mathrm{H}\, f(x_0)$ is positive definite, then x_0 is a (strict) local minimum of f.
(3) If $\mathrm{H}\, f(x_0)$ is indefinite, then x_0 is a saddle point.

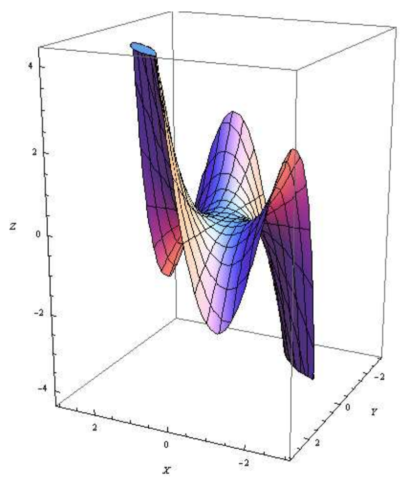
Example 3.2. Consider the function,

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

Then, $\nabla f(x, y) = (2xy + y^2, 2xy + x^2)$ so the only critical point is $(0, 0)$. To determine if it is a maximum, minimum or a saddle point, we compute the Hessian matrix,

$$\mathrm{H}\, f(0, 0) = \begin{pmatrix} 2y & 2x + 2y \\ 2x + 2y & 2x \end{pmatrix}\Bigg|_{x=y=0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

We see that the second order conditions are not informative. But, note that $f(x, x) = 2x^3$. So, $(0, 0)$ is a saddle point. The graph of f is the following one



Example 3.3. Consider the function,

$$f(x, y) = (x - 1)^4 + (y - 1)^2$$

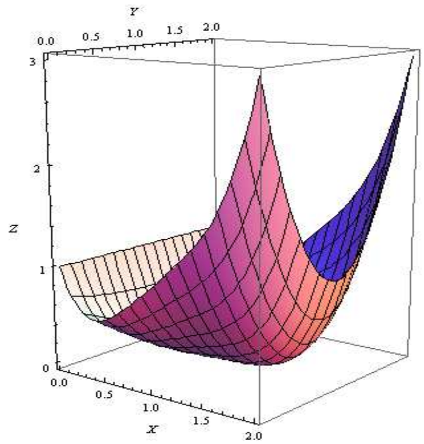
Then,

$$\nabla f(x, y) = (4(x - 1)^3, 2(y - 1))$$

so the only critical point is $(1, 1)$. To determine if it is a maximum, minimum or a saddle point, we compute the Hessian matrix,

$$\mathrm{H}\, f(1, 1) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

Since, it is positive semidefinite, the second order conditions are not informative. But, $f(x, y) \geq 0 = f(1, 1)$. Hence, $(1, 1)$ is a global minimum. The graph of f is the following one



Example 3.4. Consider the function,

$$f(x,y) = (x-1)^3 + y^2$$

The gradient is

$$\nabla f(x,y) = (3(x-1)^2, 2y)$$

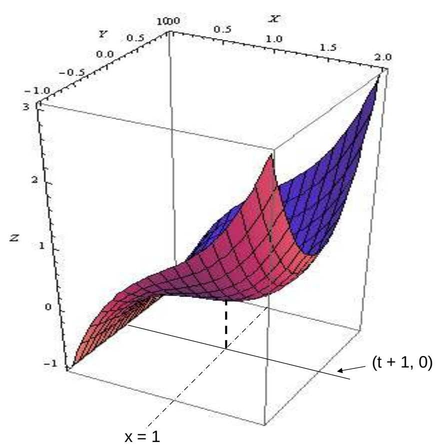
and there is a unique critical point $(1,0)$. To classify it, we compute the Hessian matrix

$$\mathrm{H}\, f(1,0) = \begin{pmatrix} 6(x-1), & 0 \\ 0 & 2 \end{pmatrix}\Big|_{x=1 y=0} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

Since, it is positive semidefinite, the second order conditions are not informative. But,

$$f(1+t,0) = t^3 = \begin{cases} > 0 \text{ if } t > 0 \\ < 0 \text{ if } t < 0 \end{cases}$$

So, $(1,0)$ is a saddle point. The graph of f is the following one



Example 3.5. Consider the function,

$$f(x,y) = x^2 + y^2(x+1)^3$$

The gradient is

$$\nabla f(x,y) = \big(2x + 3y^2(x+1)^2, 2y(x+1)^3\big)$$

and there is unique critical point, $(0,0)$. To classify it we compute the Hessian matrix,

$$\mathrm{H}\, f(0,0) = \begin{pmatrix} 2+6y^2(x+1) & 6y(x+1)^2 \\ 6y(x+1)^2 & 2(x+1)^3 \end{pmatrix}\Bigg|_{x=y=0} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

which is positive definite. Hence, $(0,0)$ is a strict local minimum. But it is not a global minimum, because, $f(-2,y) = 4 - y^2$ can be made arbitrarily small, by taking y very large.

Remark 3.6 (A justification of the second order conditions). Recall that Taylor's polynomial of order 2 of f at the point x_0 is

$$P_2(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)\,\mathrm{H}\, f(x_0)(x - x_0)$$

Recall also that if f is of class C^2 then

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

where

$$R_2(x) = f(x) - P_2(x)$$

is the error produced when we approximate the function f by Taylor's polynomial of order 2. Suppose now x_0 is a critical point of f and, hence $\nabla f(x_0) = 0$. Then,

$$f(x) - f(x_0) = \frac{1}{2}(x - x_0)\,\mathrm{H}\, f(x_0)(x - x_0) + R_2(x_0)$$

and for x near x_0 the term $R_2(x)$ is 'negligible'. Therefore if, for example we know that the term

$$(x - x_0)\,\mathrm{H}\, f(x_0)(x - x_0) > 0$$

then $f(x) - f(x_0) > 0$ for every $x \neq x_0$ 'sufficiently close' to x_0 and the point x_0 would be a local minimum. But, the condition $(x - x_0)\,\mathrm{H}\, f(p)(x - x_0) > 0$ for every $x \neq x_0$ is satisfied if $\mathrm{H}\, f(x_0)$ is positive definite.