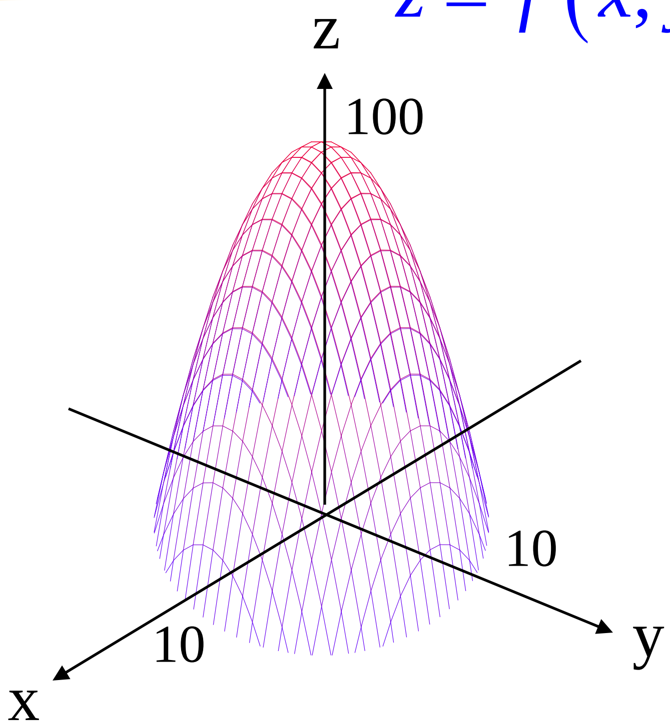
Review of elements of Calculus (functions in more than one variable)

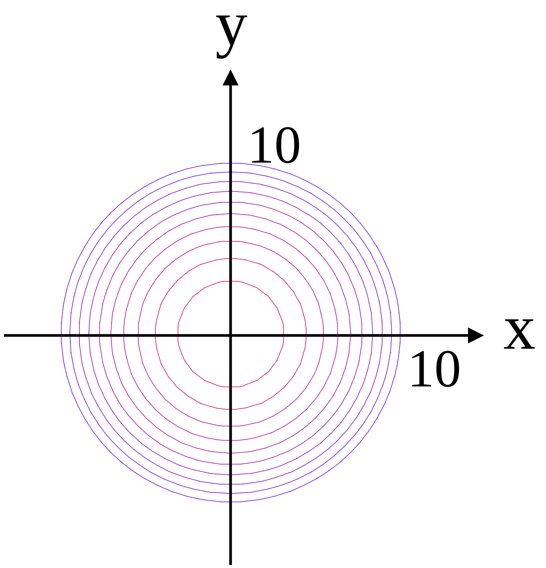
Partly adapted from the lectures of prof Piero Fariselli (University of Padova)

Functions of two variables

$$z = f(x, y) = 100 - x^2 - y^2$$



sketch of graph



level curves

Level curves are drawn by holding the z value constant (similar to contour lines on a topographic map.)

Partial Derivatives

Partial derivatives are defined as derivatives of a function of multiple variables when all but the variable of interest are held fixed during the differentiation.

so the curve you get is the slope of the curve if you keep all the other variables fixed

Definition of Partial Derivatives of a Function of Two Variables
If z = f(x,y), the the first partial derivatives of f with respect to x and y are the functions f_x and f_y defined by

$$f_x(x,y) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x}$$

$$f_y(x,y) = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y}$$

Provided the limits exist.

number of partial derivates is the same of the number of variable in the equation, you can do it i respect to to each of the variables

To find the partial derivatives, hold one variable constant and differentiate with respect to the other.

Example 1: Find the partial derivatives f_x and f_y for the function

$$f(x,y) = 5x^4 - x^2y^2 + 2x^3y$$

To find the partial derivatives, hold one variable constant and differentiate with respect to the other.

Example 1: Find the partial derivatives f_x and f_y for the function

$$f(x,y) = 5x^4 - x^2y^2 + 2x^3y$$

Solution:

$$f(x,y) = 5x^4 - x^2y^2 + 2x^3y$$

$$f_x(x,y) = 20x^3 - 2y^2x + 6yx^2$$

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

5x^4 is considered constant so it disappear

Notation for First Partial Derivative

For z = f(x,y), the partial derivatives fx and fy are denoted by

$$\frac{\partial}{\partial x} f(x,y) = f_x\left(x,y\right) = z_x = \frac{\partial z}{\partial x}$$

and

$$\frac{\partial}{\partial y} f(x,y) = f_y\left(x,y\right) = z_y = \frac{\partial z}{\partial y}$$

The first partials evaluated at the point (a,b) are denoted by

$$\left.\frac{\partial z}{\partial x}\right|_{(a,b)} = f_x\left(a,b\right) \quad \text{and} \quad \left.\frac{\partial z}{\partial y}\right|_{(a,b)} = f_y\left(a,b\right)$$

Example 2: Find the partials f_x and f_y and evaluate them at the indicated point for the function

$$f(x,y)=\frac{xy}{x-y}\ at\ (2,-2)$$

solve the two partial derivates and then substitute the values

Example 2: Find the partials f_x and f_y and evaluate them at the indicated point for the function

$$f(x,y)=\frac{xy}{x-y}\ \ at\ \ (2,-2)$$

Solution:

$$f(x,y)=\frac{xy}{x-y}\ \ at\ \ (2,-2)$$

$$f_x\left(x,y\right)=\frac{\left(x-y\right)y-xy}{(x-y)^2}=\frac{xy-y^2-xy}{(x-y)^2}=\frac{-y^2}{(x-y)^2}$$

$$f_x\left(2,-2\right)=\frac{-\left(-2\right)^2}{(2-\left(-2\right))^2}=\frac{-4}{16}=\frac{-1}{4}$$

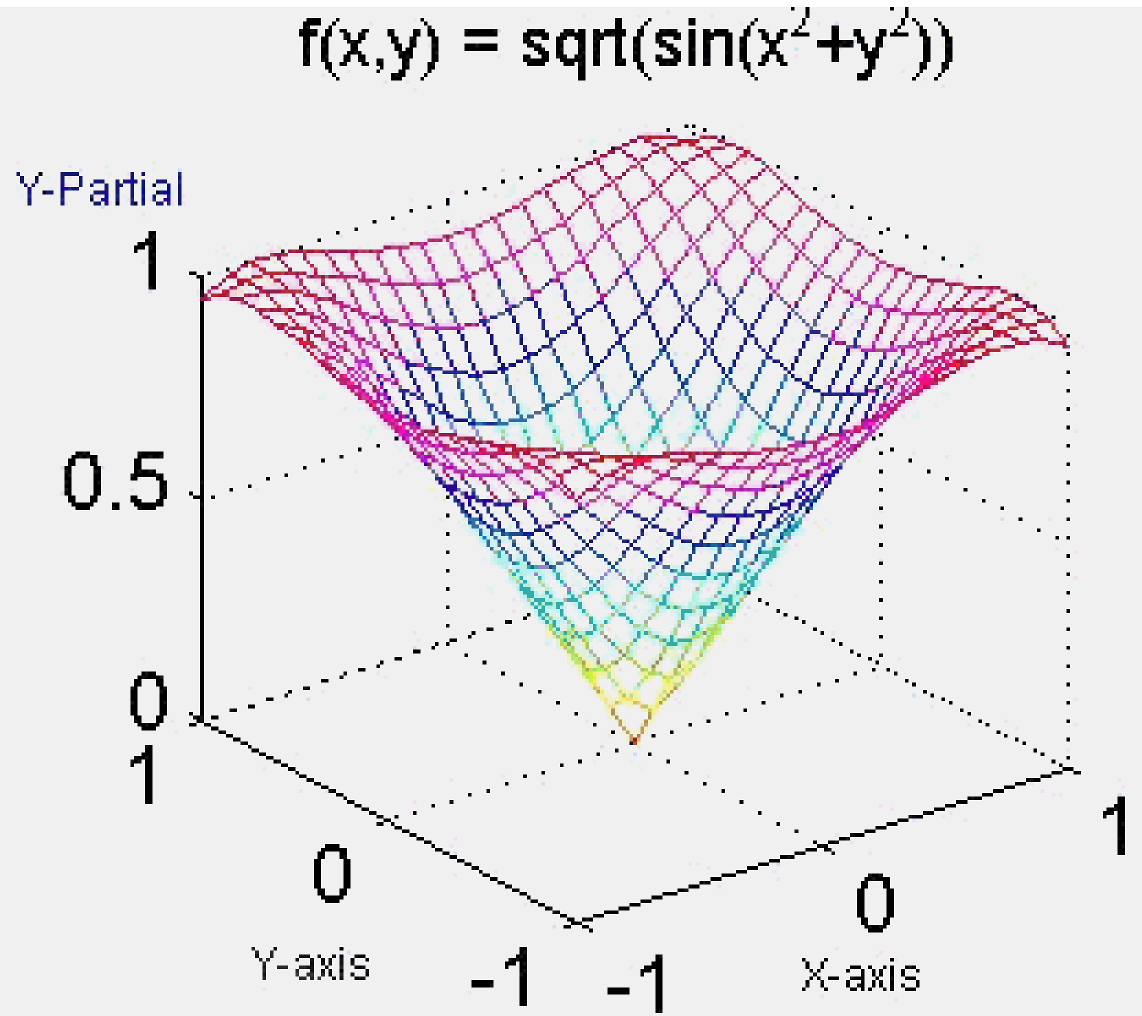
$$f_y\left(x,y\right)=\frac{\left(x-y\right)x+xy}{(x-y)^2}=\frac{x^2-xy+xy}{(x-y)^2}=\frac{x^2}{(x-y)^2}$$

$$f_y\left(2,-2\right)=\frac{x^2}{(x-y)^2}=\frac{4}{16}=\frac{1}{4}$$

The following slide shows the geometric interpretation of the partial derivative.
For a fixed x, z = $f(x_0,y)$ represents the curve formed by intersecting the surface z = f(x,y) with the plane $x = x_0$.

$f_x(x_0, y_0)$ represents the slope of this curve at the point $(x_0,y_0,f(x_0,y_0))$

Thanks to http://astro.temple.edu/~dhill001/partial-demo/
For the animation.



The surface.

Definition of Partial Derivatives of a Function of Three or More Variables

If w = f(x,y,z), then there are three partial derivatives each of which is formed by holding two of the variables

$$\frac{\partial w}{\partial x} = f_x\left(x,y,z\right) = \lim_{\Delta x \to 0} \frac{f\left(x+\Delta x, y, z\right) - f(x,y,z)}{\Delta x}$$

$$\frac{\partial w}{\partial y} = f_y\left(x,y,z\right) = \lim_{\Delta y \to 0} \frac{f\left(x, y+\Delta y, z\right) - f(x,y,z)}{\Delta y}$$

$$\frac{\partial w}{\partial z} = f_z\left(x,y,z\right) = \lim_{\Delta z \to 0} \frac{f\left(x, y, z+\Delta z\right) - f(x,y,z)}{\Delta z}$$

In general, if

$w = f(x_1, x_2, \ldots x_n)$ *there are* n *partial derivatives*

$\dfrac{\partial w}{\partial x_k} = f_{x_k}\left(x_1, x_2, \ldots x_n\right), k = 1, 2, \ldots n$ where all but the kth variable is held constant

Notation for Higher Order Partial Derivatives

Below are the different 2nd order partial derivatives:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)=\frac{\partial^2 f}{\partial x^2}=f_{xx}$$

Differentiate twice with respect to x

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)=\frac{\partial^2 f}{\partial y^2}=f_{yy}$$

Differentiate twice with respect to y

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)=\frac{\partial^2 f}{\partial y\partial x}=f_{xy}$$

Differentiate first with respect to x and then with respect to y

$$\frac{\partial}{\partial \boxed{x}}\left(\frac{\partial f}{\partial y}\right)=\frac{\partial^2 f}{\partial x\partial y}=f_{yx}$$

Differentiate first with respect to y and then with respect to x

Theorem

If f is a function of x and y such that f_{xy} and f_{yx} are continuous on an open disk R, then, for every (x,y) in R,

$$f_{xy}(x,y) = f_{yx}(x,y)$$

the order of derivation do not matter!!

Example 3:

Find all of the second partial derivatives of $f(x,y) = 3xy^2 - 2y + 5x^2y$

Work the problem first then check.

Example 3:

Find all of the second partial derivatives of $f(x,y) = 3xy^2 - 2y + 5x^2y$

$$f(x,y) = 3xy^2 - 2y + 5x^2y$$
$$f_x(x,y) = 3y^2 + 10xy$$
$$f_{xx}(x,y) = 10y$$

$$f(x,y) = 3xy^2 - 2y + 5x^2y$$
$$f_y(x,y) = 6xy - 2 + 5x^2$$
$$f_{yy}(x,y) = 6x$$

$$f(x,y) = 3xy^2 - 2y + 5x^2y$$
$$f_x(x,y) = 3y^2 + 10xy$$
$$f_{xy}(x,y) = 6y + 10x$$

Notice that $f_{xy} = f_{yx}$

$$f(x,y) = 3xy^2 - 2y + 5x^2y$$
$$f_y(x,y) = 6xy - 2 + 5x^2$$
$$f_{yx}(x,y) = 6y + 10x$$

Example 4: Find the following partial derivatives for the function $f(x,y,z) = ye^x + x\ln z$

a. f_{xz}

b. f_{zx}

c. f_{xzz}

d. f_{zxz}

e. f_{zzx}

Work it out then go to the next slide.

Example 4: Find the following partial derivatives for the function $f(x,y,z) = ye^x + x\ln z$

a. f_{xz}

$$f(x,y,z) = ye^x + x\ln z$$

$$f_x(x,y,z) = ye^x + \ln z$$

$$f_{xz}(x,y,z) = \frac{1}{z}$$

Again, notice that the 2nd partials $f_{xz} = f_{zx}$

b. f_{zx}

$$f(x,y,z) = ye^x + x\ln z$$

$$f_z(x,y,z) = \frac{x}{z}$$

$$f_{zx}(x,y,z) = \frac{1}{z}$$

c. f_{xzz}

$$f(x,y,z) = ye^x + x\ln z$$
$$f_x(x,y,z) = ye^x + \ln z$$
$$f_{xz}(x,y,z) = \frac{1}{z}$$
$$f_{xzz}(x,y,z) = \frac{-1}{z^2}$$

e. f_{zzx}

$$f(x,y,z) = ye^x + x\ln z$$
$$f_z(x,y,z) = \frac{x}{z}$$
$$f_{zz}(x,y,z) = \frac{-x}{z^2}$$
$$f_{zzx}(x,y,z) = \frac{-1}{z^2}$$

Notice All Are Equal

d. f_{zxz}

$$f(x,y,z) = ye^x + x\ln z$$
$$f_z(x,y,z) = \frac{x}{z}$$
$$f_{zx}(x,y,z) = \frac{1}{z}$$
$$f_{zxz}(x,y,z) = \frac{-1}{z^2}$$

Gradient

f: $R^n \rightarrow R$. If *f*(x) is of class C^2, ***objective function***

- Gradient of *f*

Is a vector containing all the partial derivatives of the first order

$$\nabla f(\boldsymbol{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right)$$

for a function in N variables the gradient is N long of all the partial derivates of order 1

gradient is always perpendicular to the contour line of a funcion and rappresent the LOCAL direction in wich the function increase the most

Hessian

f: $R^n \rightarrow R$. If *f*(x) is of class C^2, *objective function*

- Hessian of *f*

Is a square matrix of order n containing all the partial derivatives of the second order

$$H=\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

It is a rotation so we can diagonalize the equation and looking lambda we can uderstand if it is a MAX or a MIN

Since the mixed partial derivatives are equal irrespectively of the order of derivation, the matrix is symmetric

Taylor expansion

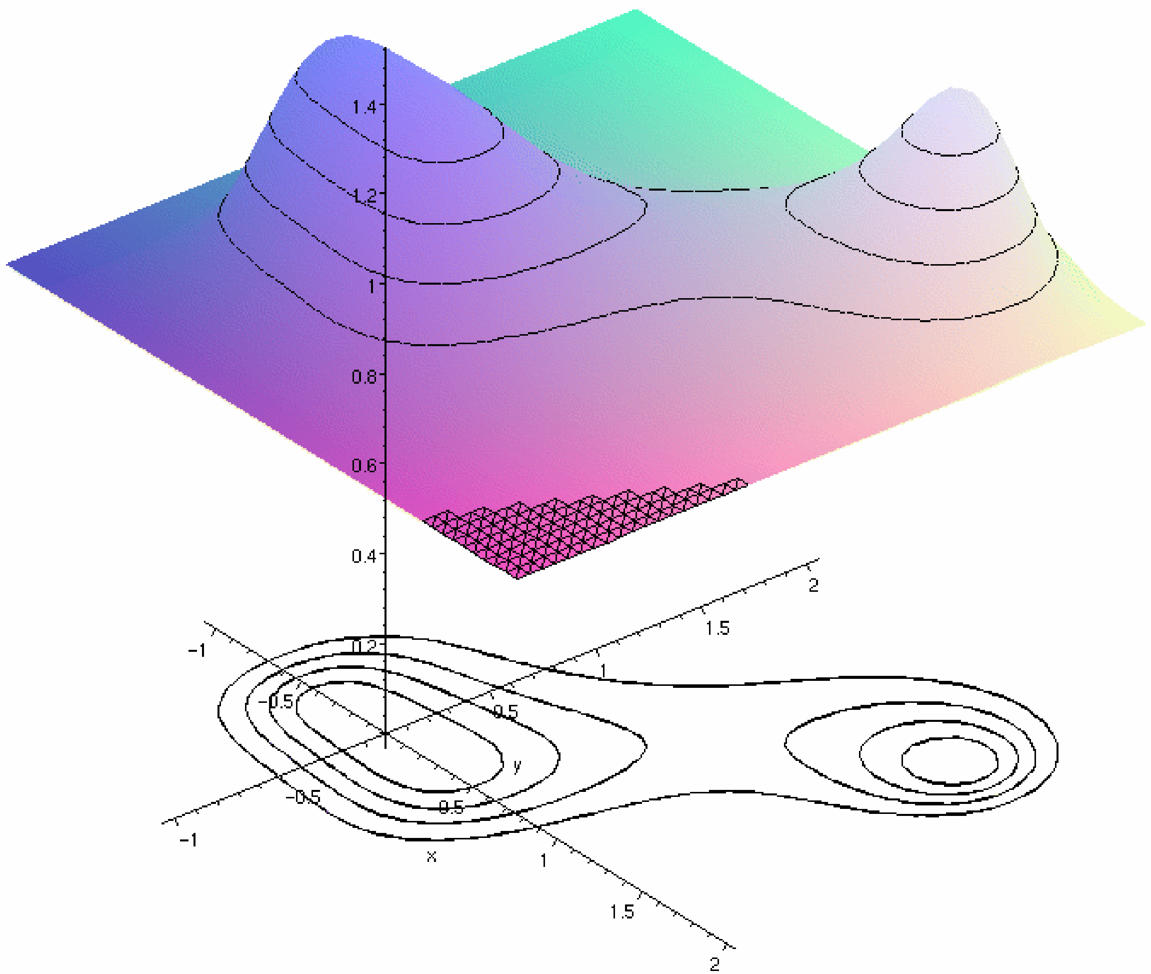
- If the variables are represented with a n-valued vector **x**, the Taylor expansion around a point **x*** is

F(x)=F(x0)+Dx*(x-x0)+Dy*(y-y0)

$$f(\vec{x}) = f(\vec{x}^*) + \nabla f\Big|_{x^*}^{T} \cdot (\vec{x} - \vec{x}^*) + \frac{1}{2}(\vec{x} - \vec{x}^*)^T H (\vec{x} - \vec{x}^*) + ...$$

hessian

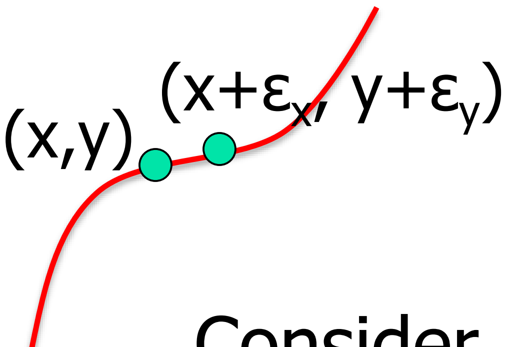
Contour lines of a function



The Gradient is locally perpendicular to level curves

Given a function f(xy) and a level curve f(x,y) = c

the gradient of f is: $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$



Consider 2 points of the curve: (x,y); $(x+\varepsilon_x, x+\varepsilon_y)$, for small ε

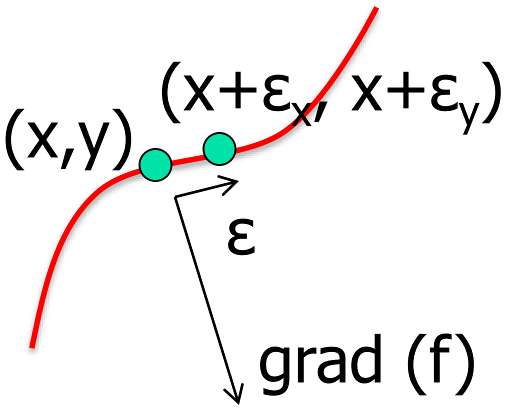
Taylor series in two variables

$$f\left(x+\varepsilon_x, y+\varepsilon_y\right) \approx \boxed{f\left(x,y\right)+\varepsilon_x \left.\frac{\partial f}{\partial x}\right|_{(x,y)} + \varepsilon_y \left.\frac{\partial f}{\partial y}\right|_{(x,y)}} =$$

$$= f\left(x,y\right)+\boldsymbol{\varepsilon}^T \left.\nabla g\right|_{(x,y)}$$

The local perpendicular to a curve: Gradient

Since both : (x,y); $(x+\varepsilon_x, x+\varepsilon_y)$, points satisfy the curve equation:

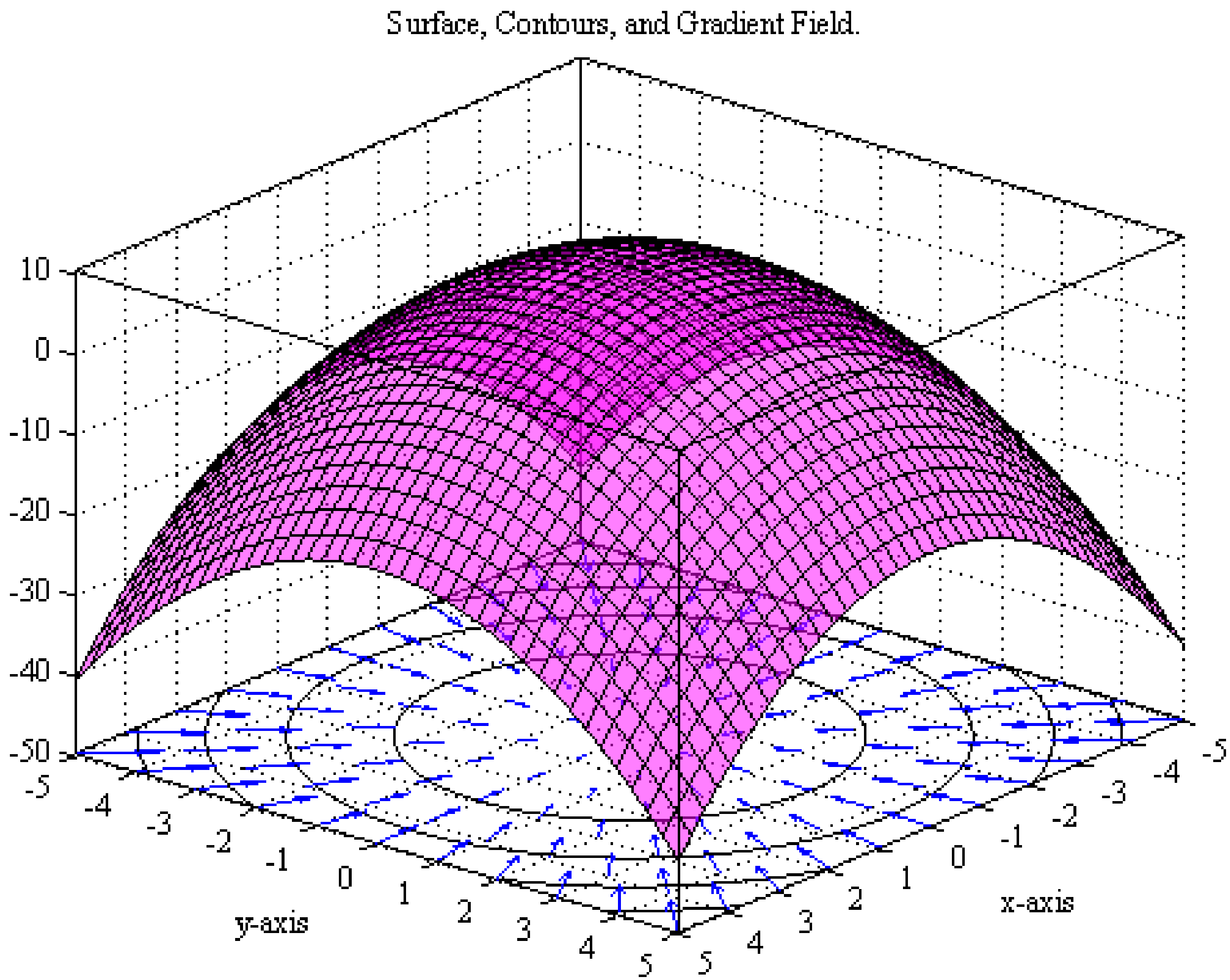


$$c \approx c + \boldsymbol{\varepsilon}^T \nabla f \big|_{(x,y)}$$

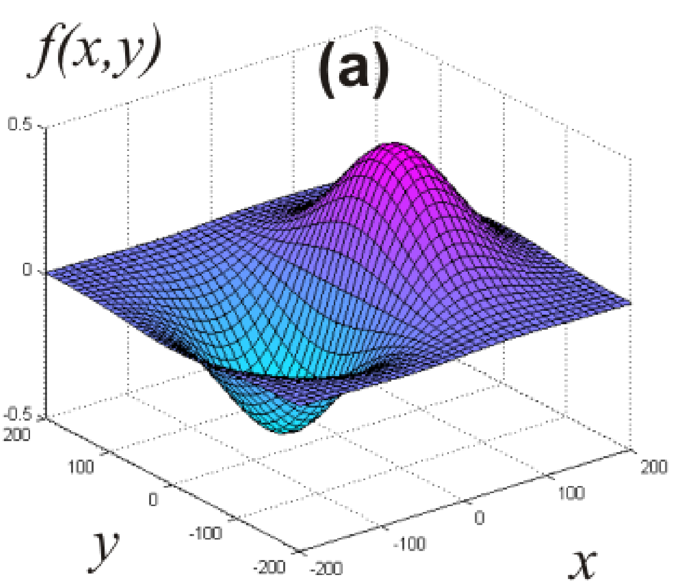
$$\boldsymbol{\varepsilon}^T \nabla f \big|_{(x,y)} = 0$$

The gradient is perpendicular to ε.
For small ε, ε is parallel to the curve and,by consequence, the gradient is perpendicular to the curve.

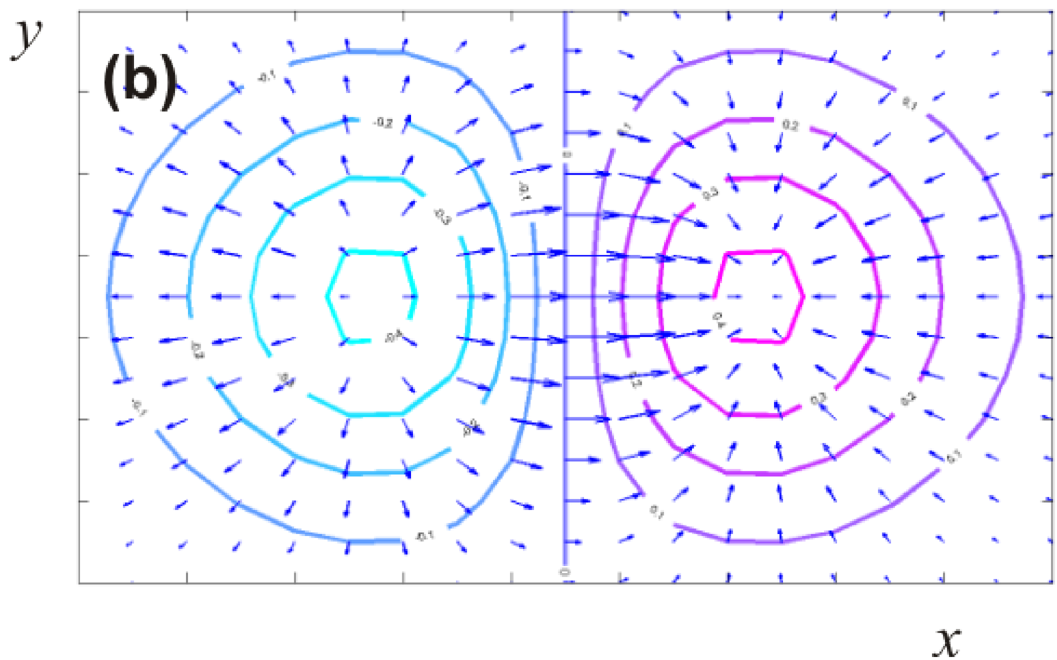
The gradient points towards the direction of maximum increase of f



Surface, Contours, and Gradient Field.
10
0
-10
-20
-30
-40
-50
y-axis
x-axis
-5
-4
-3
-2
-1
0
1
2
3
4
5



f(x,y)
(a)
0.5
0
-0.5
200
100
0
-100
-200
y
-200
-100
0
100
200
x



y
(b)
x

Local optima (without constraints)

The condition to have a local optimum in **x*** is

$$\nabla f\Big|_{x^*} = \vec{0}^{(n)}$$

i.e.: all the partial derivatives must be null

if the gardient is = 0 it can be a local MAX or MIN

In order to study whether the point is a minimum or a maximum, the matrix H must be considered

hassian

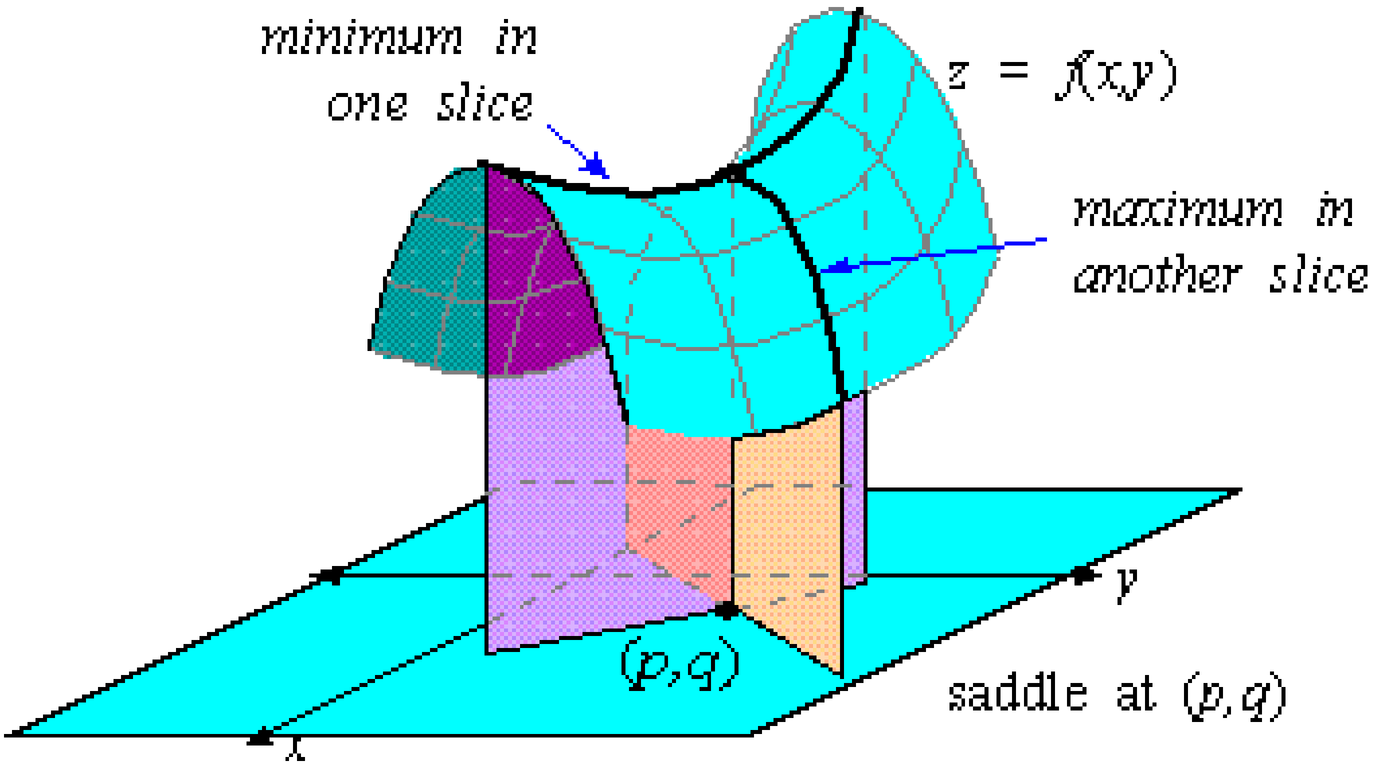
H is a nxn symmetric real matrix: it has n real eigenvalues

The point **x*** is
a maximum if all eigenvectors are positive,
a minimum if all the eigenvectors are negative

Saddle points

you have to diagonalize the hessian matrix to be able to get a system of reference in wich the eigenvectors are orthogonal and sowe can get the MAX,MIN of the function

When eigenvalues of the Hessian matrix have different sign , the critical point is a Saddle point



Example 1

$$f(x,y) = x^2 - 4y^2$$ gradient =0 (the two partial derivate must be=0)

$$\nabla f = (2x, -8y)$$

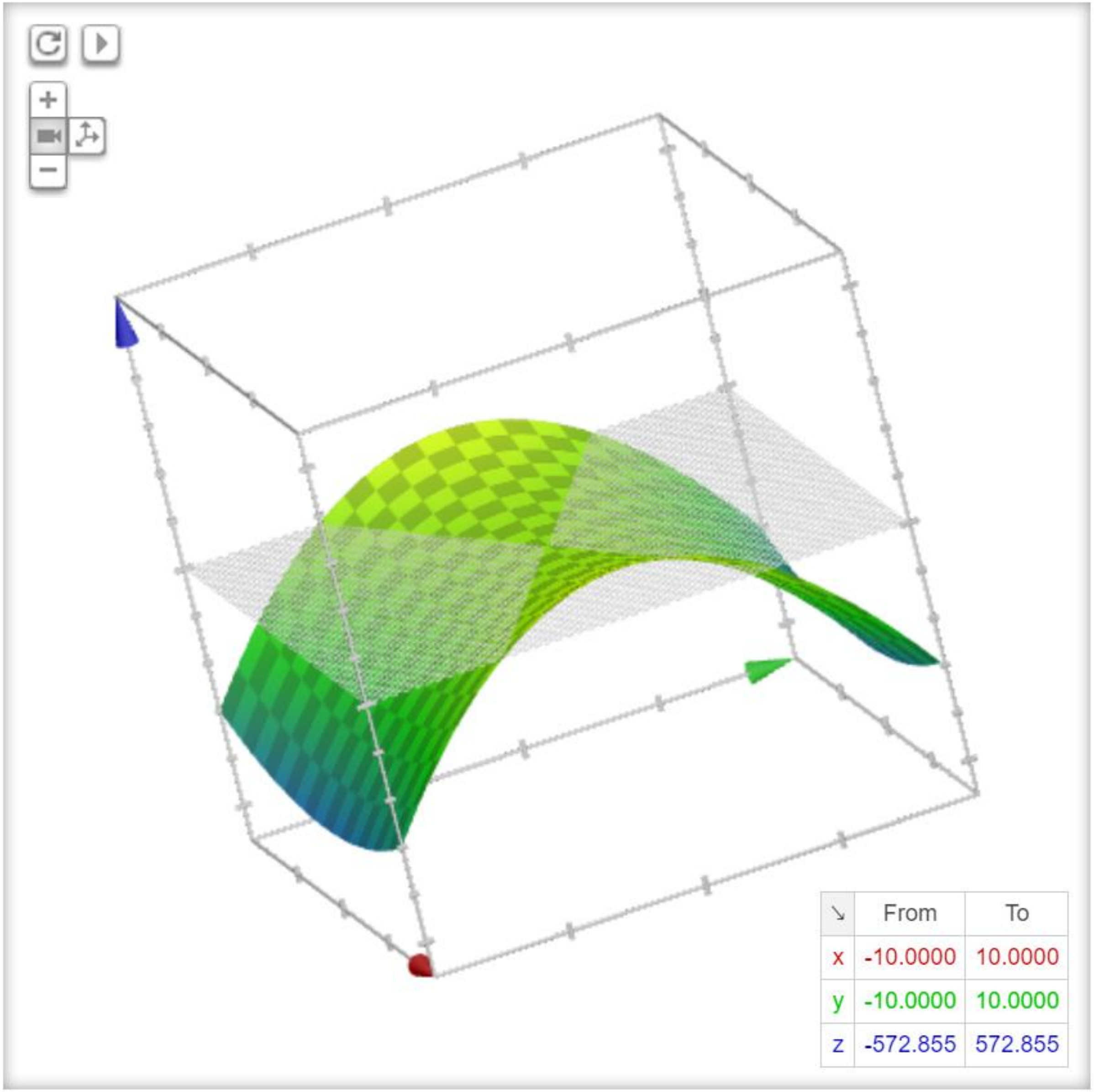
Critical point in (0,0)

hessian(second grade derivates)

$$H = \begin{pmatrix} 2 & 0 \\ 0 & -8 \end{pmatrix}$$ For any point eigenvalues= 2,-8 and so a saddle

The critical point is a saddle point

Example 1



Example 2

$$f(x,y) = (x+y)\exp(-x^2 - y^2)$$

$$\nabla f = \begin{pmatrix} \exp(-x^2-y^2) - 2x(x+y)\exp(-x^2-y^2) \\ \exp(-x^2-y^2) - 2y(x+y)\exp(-x^2-y^2) \end{pmatrix}$$

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

$$\begin{cases} \exp(-x^2-y^2)\left[1 - 2x^2 - 2xy\right] = 0 \\ \exp(-x^2-y^2)\left[1 - 2y^2 - 2xy\right] = 0 \end{cases}$$

$$\begin{cases} \left[1 - 2x^2 - 2xy\right] = 0 \\ \left[2x^2 - 2y^2\right] = 0 \end{cases} \begin{cases} \left[1 - 2x^2 - 2xy\right] = 0 \\ x = \pm y \end{cases}$$

Example 2

$$f(x,y) = (x+y)\exp(-x^2 - y^2)$$

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

$$1)x = y \Rightarrow 1 - 2x^2 - 2x^2 = 0$$

$$(1/2;1/2);(-1/2;-1/2)$$

$$2)x = -y \Rightarrow 1 - 2x^2 + 2x^2 = 0 \Rightarrow NoSolution$$

Critical points in (1/2,1/2) and (-1/2,-1/2)

Example 2

$$f(x,y) = (x+y)\exp(-x^2-y^2)$$

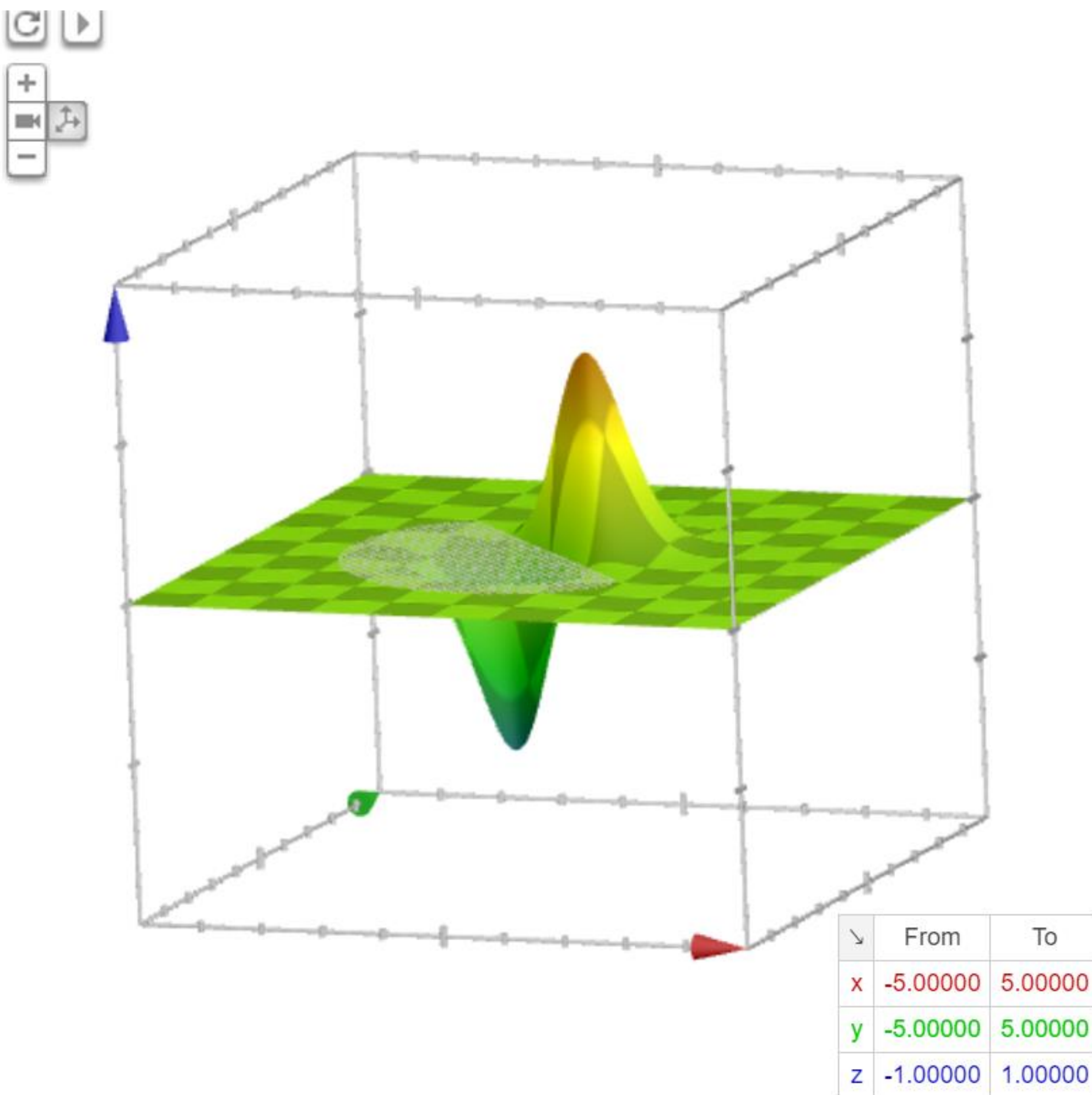
$$\nabla f = \begin{pmatrix} \exp(-x^2-y^2)\left[1-2x^2-2xy\right] \\ \exp(-x^2-y^2)\left[1-2y^2-2xy\right] \end{pmatrix}$$

$$H = \begin{pmatrix} \exp(-x^2-y^2)\left[-4x-2y-2x(1-2x^2-2xy)\right] & \exp(-x^2-y^2)\left[-2x-2y(1-2x^2-2xy)\right] \\ \exp(-x^2-y^2)\left[-2y-2x(1-2x^2-2xy)\right] & \exp(-x^2-y^2)\left[-4y-2x-2y(1-2x^2-2xy)\right] \end{pmatrix}$$

$$H\big|_{(1/2;1/2)} = \exp(-1/2)\begin{pmatrix} -3 & -1 \\ -1 & -3 \end{pmatrix} \qquad \lambda_1 = -2; \lambda_2 = -4; MAXIMUM$$

$$H\big|_{(-1/2;-1/2)} = \exp(-1/2)\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \qquad \lambda_1 = +2; \lambda_2 = +4; MINIMUM$$

Example 2

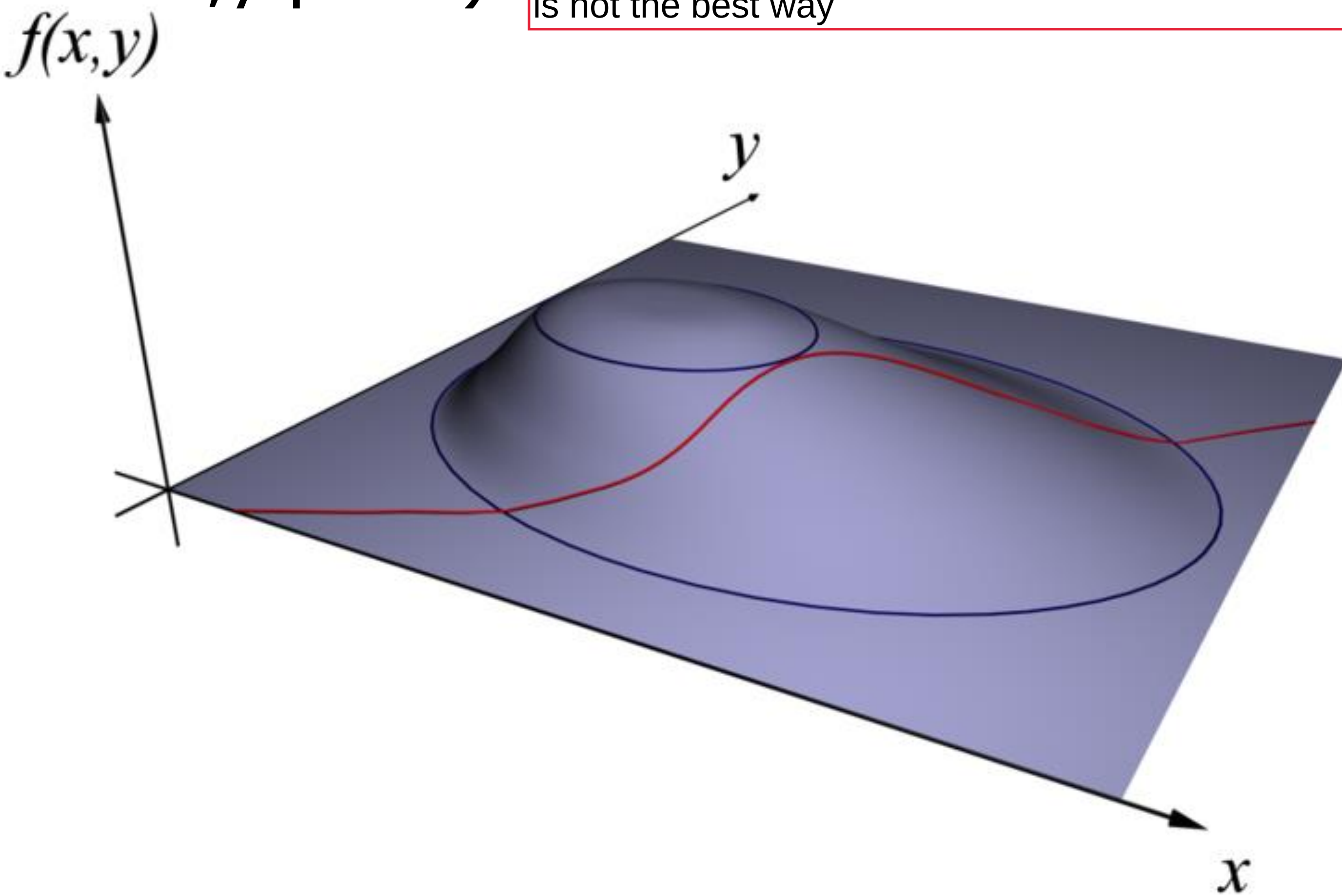


Constrained optimization problems: Lagrange Multipliers

Aim

- We want to maximise the function z = f(x,y) subject to the constraint g(x,y) = c (curve in the x,y plane)

if you can express X in Y or viceversa you can solve it, but it is not the best way



Simple solution

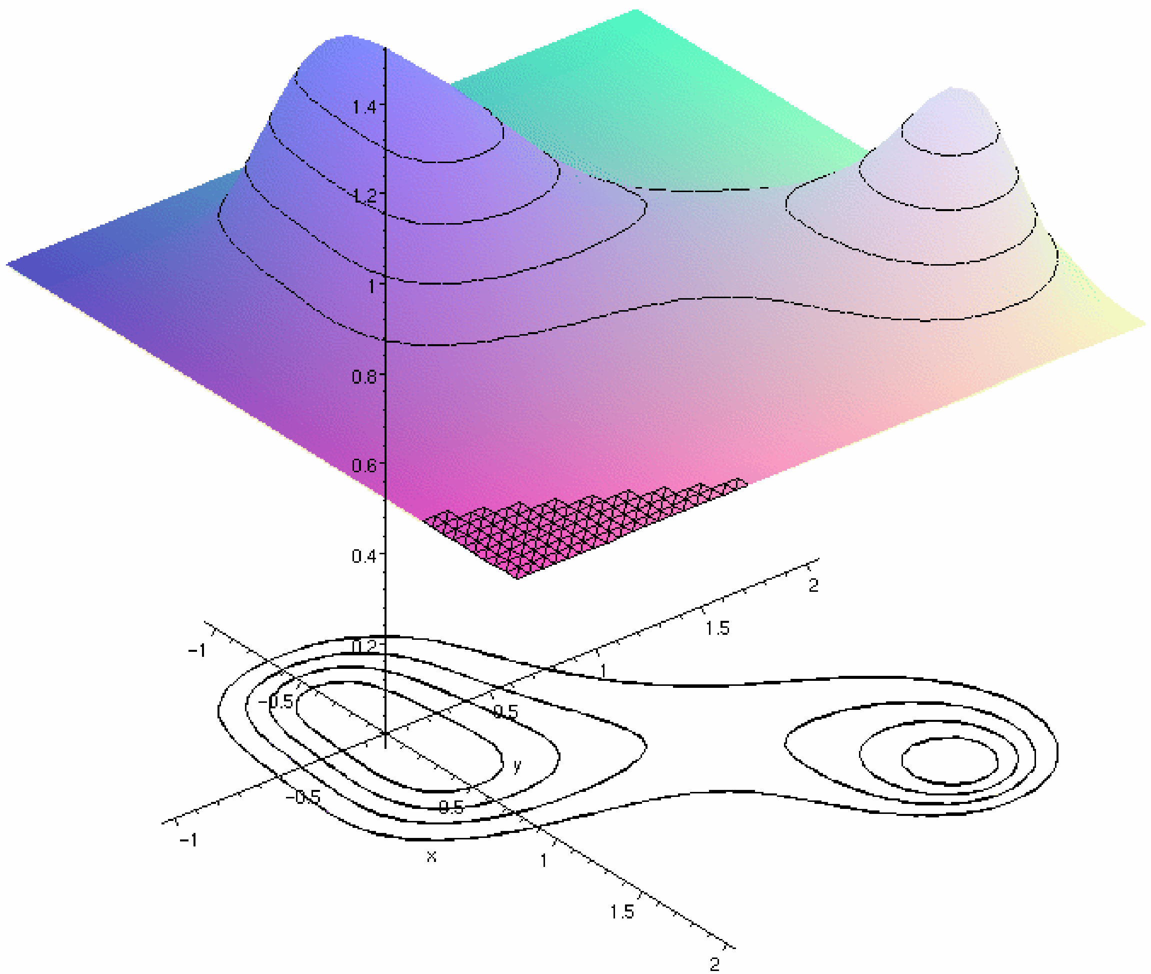
Solve the constraint g(x,y) = c and express, for example, y=h(x)

The substitute in function f and find the maximum in x of

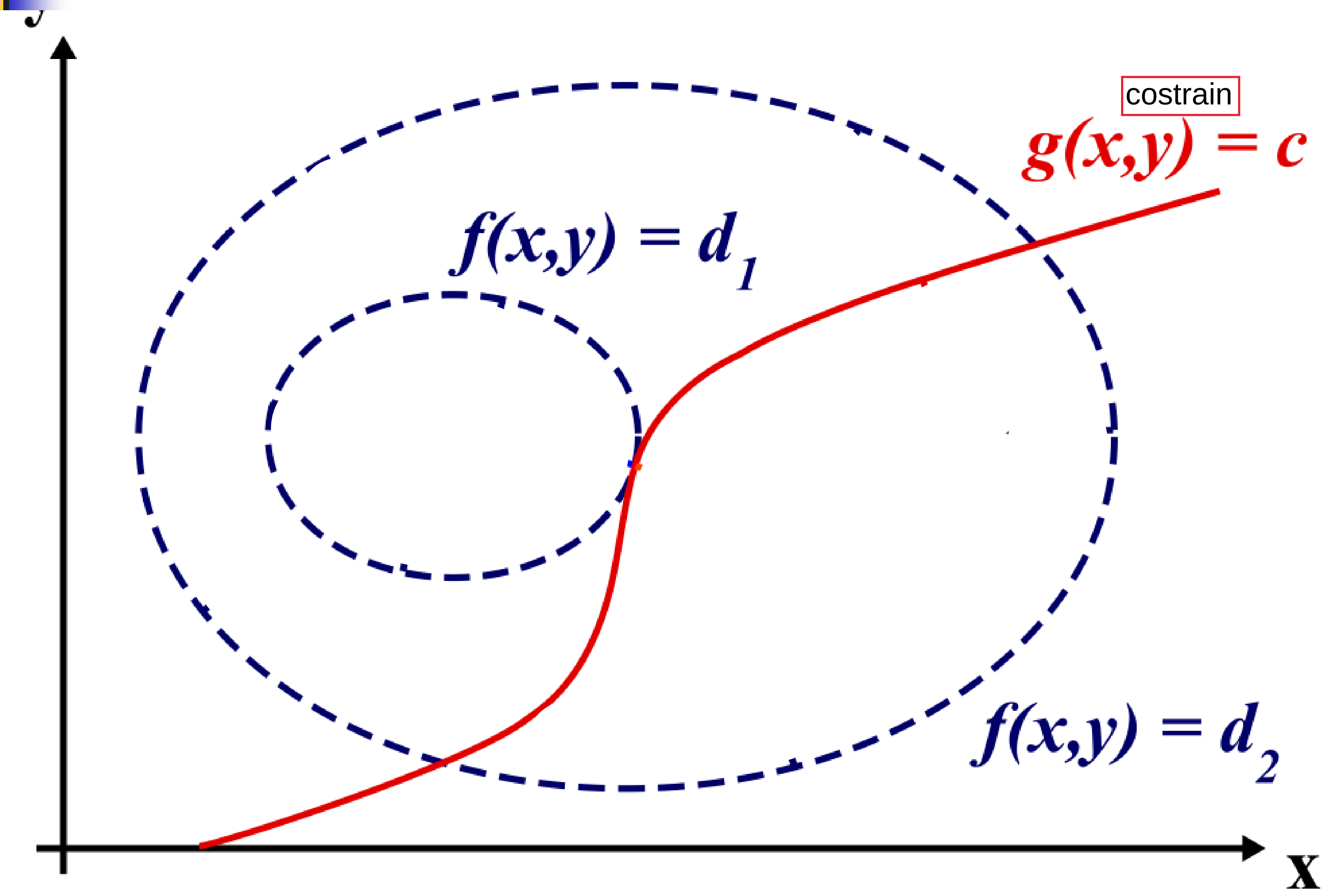
f(x, h(x))

Analytical solution of the constraint can be very difficult

Contour lines of a function



Contour lines of f and constraint



Lagrange Multipliers

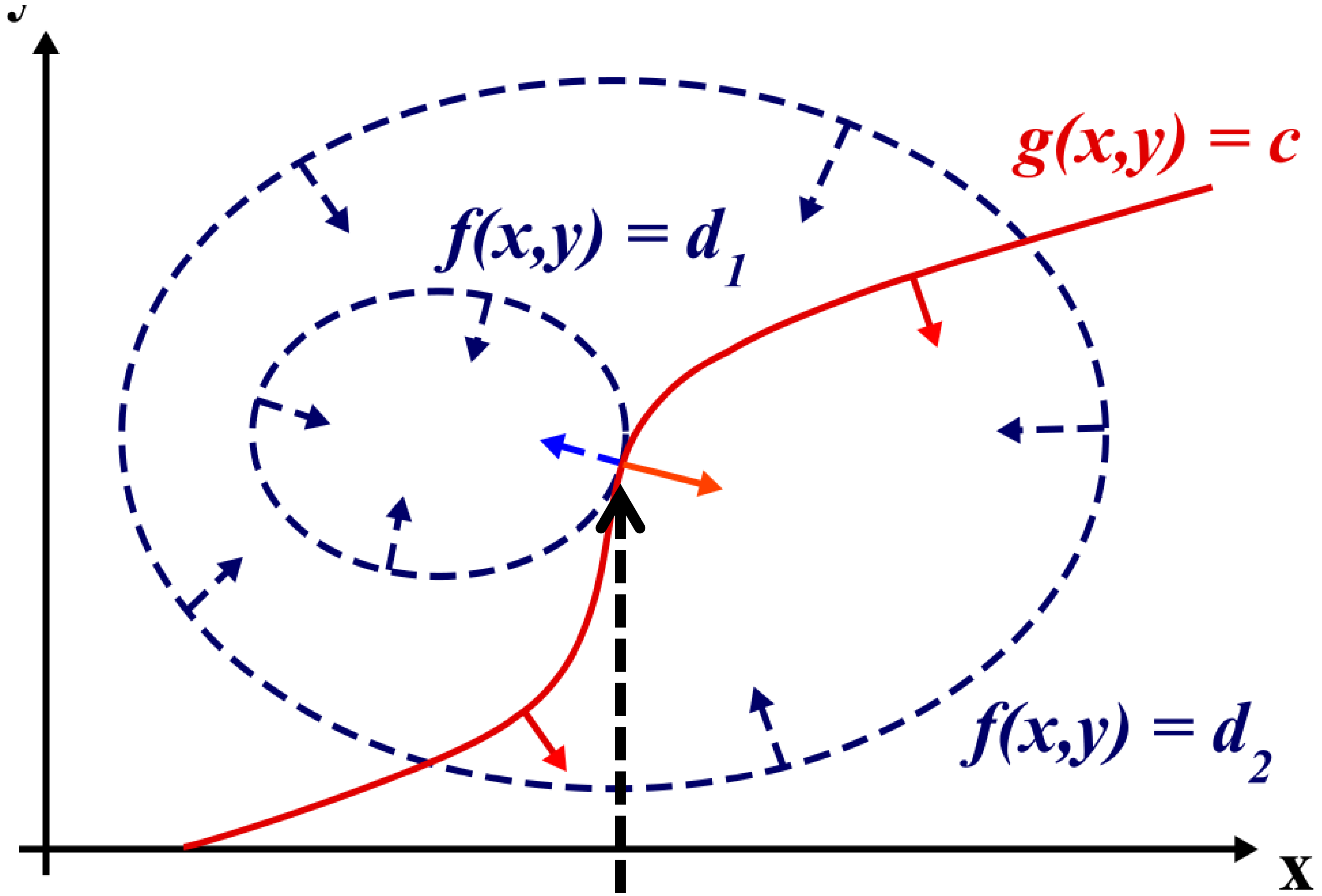
Suppose we walk along the constraint line $g(x,y) = c$.

In general the contour lines of f are distinct from the constraint $g(x,y) = c$.

While moving along the constraint line $g(x,y) = c$ the value of f vary (that is, different contour levels of f are intersected).

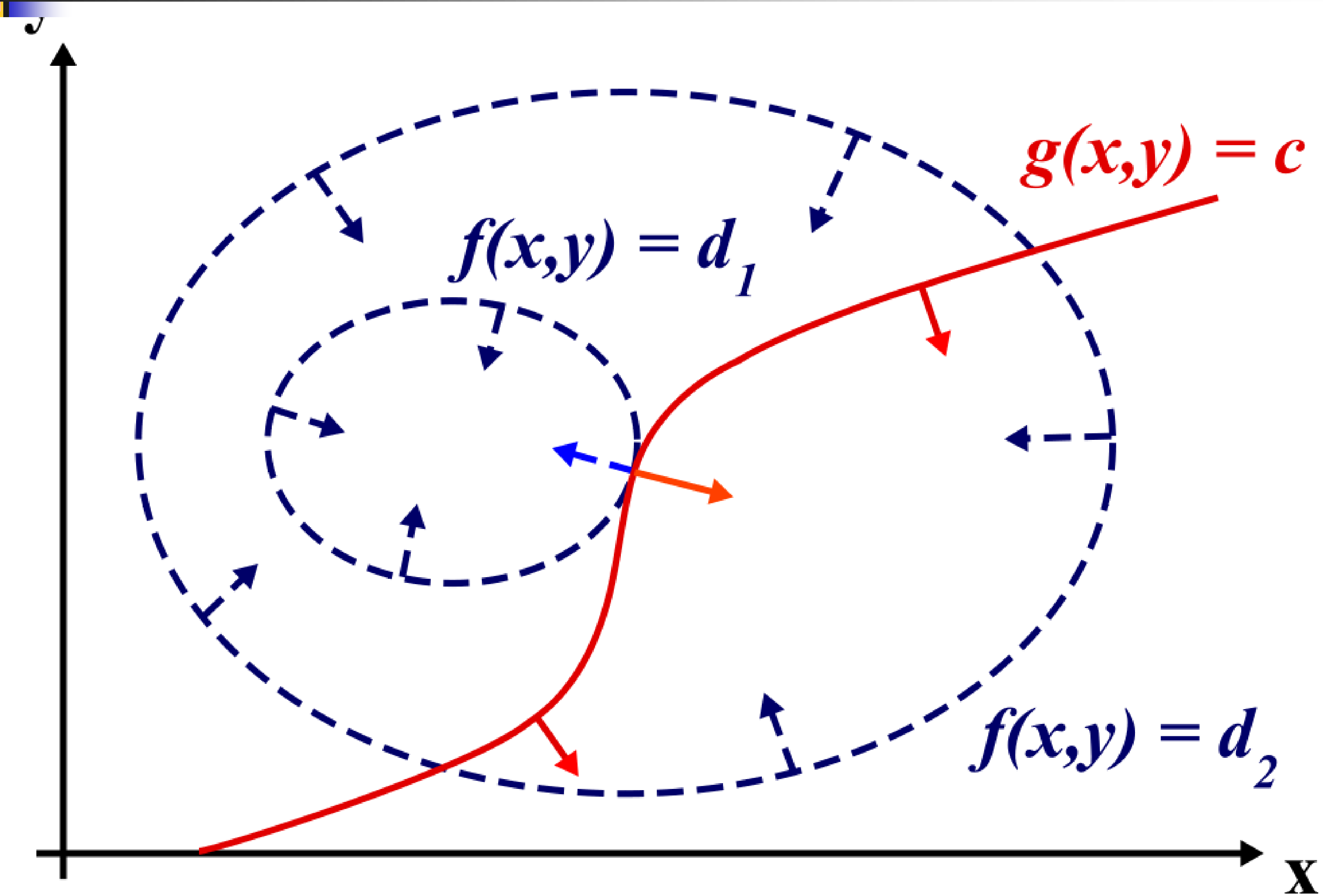
Only when the constraint line $g(x,y) = c$ touches the contour lines of f in a tangential way, we do not increase or decrease the value of f: ***the function f is at its local max or min along the constraint.***

Geometrical interpretation

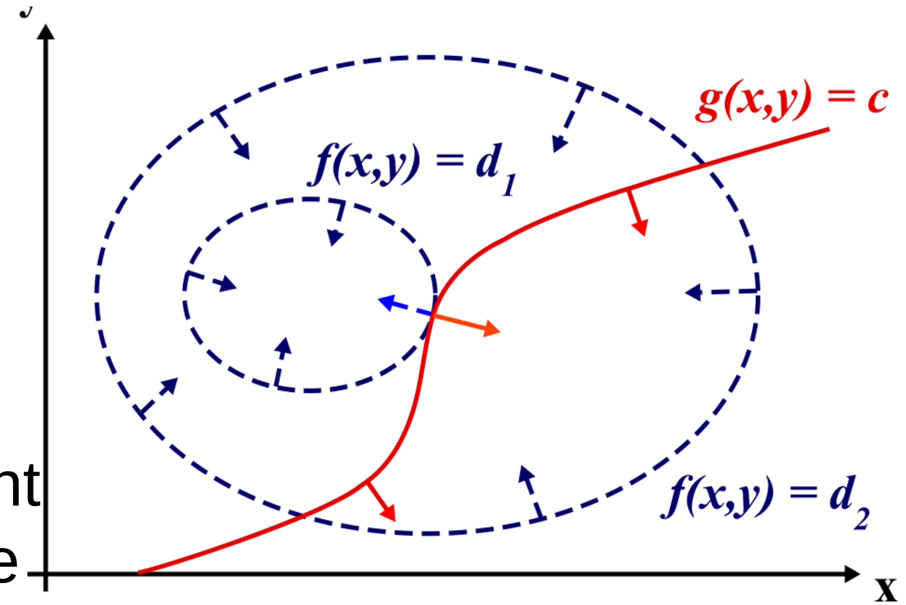


Contour line and constraint are tangential: their local perpendicular to are parallel

Normal to a curve



Lagrange Multipliers



On the point of g(x,y)=c that Max-min-imize f(x,y), the gradient of f is perpendicular to the curve g(x,y) =c, otherwise we should increase or decrease f by moving locally on the curve.

So, the two gradients are parallel

$$\nabla f = \lambda \nabla g$$

gradient of F=lagrange multiplier*gradient of G

for some scalar λ (where ∇ is the gradient).

$$\nabla_{x,y} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right).$$

Lagrange Multipliers

Thus we want points (x,y) where $g(x,y) = c$ and $\nabla_{x,y} f = \lambda \nabla_{x,y} g$

,

To incorporate these conditions into one equation, we introduce an auxiliary function (Lagrangian)

$$F(x,y,\lambda) = f(x,y) - \lambda\big(g(x,y) - c\big)$$

and solve

.

$$\nabla_{x,y,\lambda} F(x,y,\lambda) = 0$$

Recap of Constrained Optimization

- Suppose we want to: minimize/maximize f(**x**) subject to g(**x**) = 0
- A necessary condition for $\mathbf{x}_0$ to be a solution:

$$\begin{cases} \frac{\partial}{\partial \mathbf{x}}\Big(f(\mathbf{x}) - \alpha g(\mathbf{x})\Big)\Big|_{\mathbf{x}=\mathbf{x}_0} = \mathbf{0} \\ g(\mathbf{x}) = 0 \end{cases}$$

- α: the Lagrange multiplier
- For multiple constraints $g_i(\mathbf{x}) = 0$, i=1, …, m, we need a Lagrange multiplier α_i for each of the constraints

$$\begin{cases} \frac{\partial}{\partial \mathbf{x}}\Big(f(\mathbf{x}) - \sum_{i=1}^{n} \alpha_i g_i(\mathbf{x})\Big)\Big|_{\mathbf{x}=\mathbf{x}_0} = \mathbf{0} \\ g_i(\mathbf{x}) = 0 \qquad \text{for } i = 1, \ldots, m \end{cases}$$

Example

- Find the dimensions of the box (rectangular parallelepiped) with largest volume if the total surface area is 64 cm^2

- We need to maximize

$$f(x, y, z) = xyz$$

with the constraint

$$2xy + 2xz + 2yz = 64$$

Example

MAXIMIZATION WITH CONSTRAIN

Write G in the form G(x,y,z)=0

L(lagrangian)=F(x,y,z)-LambdaG(x,y,z)

$$L = f(x,y,z) - \lambda g(x,y,z) = xyz - \lambda\left[xy + xz + yz - 32\right]$$

$$\begin{cases} \dfrac{\partial L}{\partial x} = 0 \\ \dfrac{\partial L}{\partial y} = 0 \\ \dfrac{\partial L}{\partial z} = 0 \\ \dfrac{\partial L}{\partial \lambda} = 0 \end{cases} \Rightarrow \begin{cases} yz - \lambda y - \lambda z = 0 \\ xz - \lambda x - \lambda z = 0 \\ xy - \lambda x - \lambda y = 0 \\ xy + xz + yz - 32 = 0 \end{cases}$$

make a system of the derivates of L

put DL=0 soit means that the two gradient nullify each other

solve the system IN ALL VARIABLE : X,Y,Z and also LAMBDA because in L it is a variable

Example

x*(1° eq)-y*(2° eq)

y*(2° eq)-z*(3° eq)

z*(3° eq)-x*(1° eq)

$$\begin{cases} \lambda z(y-x)=0 \\ \lambda y(z-x)=0 \\ \lambda x(z-y)=0 \\ xy+xz+yz-32=0 \end{cases}$$

$$\begin{cases} x=y=z \\ xy+xz+yz-32=0 \end{cases} \Rightarrow \begin{cases} x=y=z \\ 3x^2=32 \end{cases} \Rightarrow x=y=z=4\sqrt{\frac{2}{3}}$$

Exercises

Find the maximum and minimum of $f(x, y) = 5x - 3y$

subject to the constraint $x^2 + y^2 = 136$

Find the maximum and minimum of $f(x, y, z) = xyz$

subject to the constraint, $x + y + z = 1$

assuming $x, y, z \geq 0$

Exercises

Find the maximum and minimum of $f(x, y, z) = 4y - 2z$

subject to the constraints $2x - y - z = 2$ and $x^2 + y^2 = 1$

you have to introduce one Lagrange multiplier for every constrain

L=F(x,y,z)-sum(lambda*constrain)

Information entropy

$$p_i, i = 1 \ldots n$$

$$\sum p_i = 1$$

$$\boxed{S = \sum (-p_i \ln p_i)}$$

Maximize the function

$$L = \sum (-p_i \ln p_i) - \lambda \left(\sum p_i - 1 \right)$$

$$\frac{\partial L}{\partial p_k} = -\ln p_k - 1 - \lambda = 0, \forall k$$

$$\frac{\partial L}{\partial \lambda} = -\sum p_i + 1 = 0$$

Information entropy

$p_i, i = 1 \ldots n$ $\sum p_i = 1$ $S = \sum(-p_i \ln p_i)$

$$-\ln p_k = 1 + \lambda, \forall k \Rightarrow p_k = \exp(-1-\lambda), \forall k$$

$$\sum p_i = 1$$

All p_k are identical and sum to 1: $p_k=1/n$

$$S_{Max} = \sum(-\frac{1}{n}\ln\frac{1}{n}) = \ln n$$

the information you get from a random event is greater if greater the possible outcome