Chapter 3 Higher order derivatives, maxima + minima

Theorem If $f(x,y)$ is C^2 then $f_{xy} = f_{yx}$. In particular the hessian matrix

$$H(f) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

is symmetric.

Example Let $f(x,y,z)$

$$= \frac{1}{\sqrt{x^2+y^2+z^2}} \quad \text{for } (x,y,z) \neq (0,0,0)$$

Then $\Delta f := f_{xx} + f_{yy} + f_{zz} = 0$

The operator Δ (Laplace operator) is the most famous operator of mathematical physics and appears in the basic pdes describing

electrostatics potential eqn (or gravitational)

$$\Delta u(x,y,z) = 0$$

electrostatic potential in absence of external charge

(Poisson eqn $\Delta u(x,y,z) = f(x,y,z)$)

heat eqn $$\frac{\partial u}{\partial t} = \Delta u$$

wave eqn $$c^2 \frac{\partial^2 u}{\partial t^2} = \Delta u$$

wave equation

$n=1$ $\quad u_{xx} = c^2 u_{tt}$

$$u(x,t) = f(x+c\cdot t) + g(x-ct)$$

$$u_{xx} = f''(x+ct) + g''(x-ct)$$

$$u_{tt} = c^2 (f''(x+ct) + g''(x-ct)$$

Example $u(x,t) = \frac{1}{\sqrt{t}} e^{-\frac{x^2}{4t}}$.

$0 \le x \le l$

(describes heat conduction in a thin rod)

$n=1$

Verify u satisfies heat eqn with $c=1$ (normalized)

$$u_t = e^{-\frac{x^2}{4t}}\left(-\frac{1}{2}t^{-3/2} + \frac{1}{\sqrt{t}} \cdot \frac{x^2}{4t^2}\right)$$

$$u_{xx} = \frac{1}{\sqrt{t}}\left(e^{-\frac{x^2}{4t}}\left(-\frac{2x}{4t}\right)\right)_x$$

$$= \frac{-1}{2t^{3/2}} e^{-\frac{x^2}{4t}}\left(1 - \frac{2x^2}{4t}\right) = u_t$$

Since f_{xy} is assumed continuous in a nbhd of (x_0, y_0)

$$f_{yx}(x_0,y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = f_{xy}(x_0, y_0)$$

Example

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

$$f_x = e^y + 2yx, \quad f_y = xe^y + x^2$$

$$f_{xy} = e^y + 2x, \quad f_{yx} = e^y + 2x$$

(3.2) Taylor approximation

For a "nice function" $g : \mathbb{R} \to \mathbb{R}$ of one variable we have the Taylor approximation

$$g(x_0+h) = g(x_0) + g'(x_0)\,h + \frac{g''(x_0)}{2!}h^2 + \dots + \frac{g^{(k)}(x_0)}{k!}h^k + R_k(x_0,h)$$

$$R_k(x_0,h) = \frac{g^{(k+1)}(x_0^*)\,h^{k+1}}{(k+1)!}$$

"lower order error term"

of order k (a polynomial approximation)

where $\lim_{h\to 0} \frac{R_k(x_0,h)}{h^k} = 0$

explicitly: $R_k(x_0,h) = \int_{x_0}^{x_0+h} \frac{(x_0+h-\tau)^k}{k!} g^{(k+1)}(\tau)\,d\tau$

$$= g^{(k+1)}(x_0^*)$$

$x_0^* \in (x_0, x_0+h)$

in application $x_0 = 0$ $h = 1$

We can use this to give a similar approximation for $f(\vec{x})$

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

with $\vec{x}_0, \vec{h}$ fixed

We take $k = 2$: $0 \le t \le 1$

$g(0) = f(\vec{x}_0)$, $g(1) = f(\vec{x}_0 + \vec{h})$

$$g'(t) = \nabla f(\vec{x}_0 + t\vec{h}) \cdot \vec{h} = \sum_{k=1}^{n} h_k f_k(\vec{x}_0 + t\vec{h})$$

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

$$g''(t) = \sum_{k,l=1}^{n} h_k h_l f_{kl}(\vec{x}_0 + t\vec{h})$$

$$g''(0) = \sum_{k,l=1}^{n} h_k h_l f_{kl}(\vec{x}_0)$$

$$g'''(t) = \sum_{k,l,j=1}^{n} h_k h_l h_j f_{klj}(\vec{x}_0 + t\vec{h})$$

5.1

$n=1$

$$f(x_0+h) = f(x_0) + \int_{x_0}^{x_0+h} f'(\tau)\,d\tau$$

$$= f(x_0) - \underbrace{(x_0+h-\tau)\,f'(\tau)\Big|_{x_0}^{x_0+h}}_{h f'(x_0)} + \int_{x_0}^{x_0+h} (x_0+h-\tau)\,f''(\tau)\,d\tau$$

$$= f(x_0) + h f'(x_0) + \frac{h^2}{2} f''(x_0) + \cdots + \frac{h^k}{k!} f^k(x_0) + \int_{x_0}^{x_0+h} \frac{(x_0+h-\tau)^k}{k!} f^{k+1}(\tau)$$

$$R_k(x_0,h) = \frac{h^{k+1}}{(k+1)!} f^{k+1}(x_0^*)$$

$$x_0^* \in (x_0, x_0+h)$$

$$g(t) = f(\vec{x}_0 + th)$$
$$g'(t) = (h \cdot D) f(\vec{x}_0 + th)$$
$$g^{(j)}(t) = (h \cdot D)^j f(\vec{x}_0 + th)$$

where $h \cdot D = h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n}$

Multinomial theorem

For any $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$
and positive integer J

$$(x_1 + \dots + x_n)^k = \sum_{|\alpha| = k} \frac{k!}{\alpha!} x^\alpha$$

$\alpha = (\alpha_1, \dots, \alpha_n)$ multiindex $\alpha_i \geq 0$

$|\alpha| = \alpha_1 + \dots + \alpha_n$ $\alpha! = \alpha_1! \dots \alpha_n!$

$x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$

$n = 2$ binomial theorem

$$(x_1 + x_2)^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} x_1^j x_2^{k-j} = \sum_{\alpha_1 + \alpha_2 = k} \frac{k!}{\alpha_1! \alpha_2!} x_1^{\alpha_1} x_2^{\alpha_2}$$
$$= \sum_{|\alpha| = k} \frac{k!}{\alpha!} \vec{x}^\alpha$$

(5.4)

$$f(\vec{x}_0+\vec{h}) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x_0)}{\alpha!} h^\alpha + R_{\alpha,k}(h)$$

$$R_{\alpha,k}(h) = \sum_{|\alpha|=k+1} \partial^\alpha f(\vec{x}_0+c\vec{h}) \frac{h^\alpha}{\alpha!}$$

some $c \in (0,1)$

$$\partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}$$

scalar
Then (take $t=1$) $h = \Delta t = 1$

$$f(\vec{x}_0+\vec{h}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot\vec{h} + \frac{1}{2}\sum_{i,j=1}^{n} f_{ij}(\vec{x}_0)\, h_i h_j + R_2(\vec{x}_0,\vec{h})$$

where $R_2(\vec{x}_0,\vec{h}) = \int_0^1 \frac{(1-\tau)^2}{2} f_{ijk}(\vec{x}_0+\tau\vec{h})\, h_i h_j h_k \, d\tau$

<u>Example 1</u> Compute the second order Taylor formula for $f(x,y) = \sin(x+2y)$ about $\vec{x}_0 = (0,0)$

$f(0,0) = 0$ $\quad f_x(0,0) = 1$ $\quad f_y(0,0) = 2$

$f_{xx}(0,0) = 0 = f_{yy}(0,0) = f_{xy}(0,0)$

So, $f(\vec{h}) = f(h_1,h_2) = h_1 + 2h_2 + R_2(\vec{0},$

(here all second order derivatives at $(0,0)$ vanish)

Example 2

$f(x,y) = \sqrt{1+4x^2+y^2}$ about

$(x_0, y_0) = (1, 2)$ $\quad f(1,2) = 3$

Approximate $f(1.1, 2.05)$

$f_x = \dfrac{4x}{\sqrt{1+4x^2+y^2}}$ $\quad f_x(1,2) = 4/3$

$f_y = \dfrac{y}{\sqrt{1+4x^2+y^2}}$ $\quad f_y(1,2) = 2/3$

$f_{xx} = \dfrac{4}{\sqrt{1+4x^2+y^2}} - \dfrac{16x^2}{(1+4x^2+y^2)^{3/2}}$ $\quad f_{xx}(1,2) = 4/3 - \frac{16}{27} = \frac{20}{27}$

$f_{yy} = \dfrac{1}{\sqrt{1+4x^2+y^2}} - \dfrac{y^2}{(1+4x^2+y^2)^{3/2}}$ $\quad f_{yy}(1,2) = 1/3 - \frac{4}{27} = 5/27$

$f_{xy} = -\dfrac{4xy}{(1+4x^2+y^2)^{3/2}}$ $\quad f_{xy}(1,2) = -\frac{8}{27}$

6.2

$$f(x_0+h_1, y_0+h_2) \simeq$$

$$f(x_0,y_0) + f_x(x_0,y_0)\,h_1 + f_y(x_0,y_0)\,h_2$$

$$+ \frac{1}{2}\left(f_{xx}(x_0,y_0)\,h_1^2 + 2 f_{xy}(x_0,y_0)\,h_1 h_2 + f_{yy}(x_0,y_0)\,h_2^2 \right)$$

$$h_1 = 0.1 \qquad h_2 = 0.05$$

$$f(1.1,\ 2.05) \approx 3 + 4/3\,(0.1) + 2/3\,(0.05)$$

$$+ 1/2\left(\frac{20}{27}(0.1)^2 - 2\,\frac{8}{27}(0.1)(0.05) + \frac{5}{27}(0.05)^2 \right)$$

$$= \quad 3.169\cancel{0}$$

$$\Big(\text{actual } f(1.1,\ 2.05) = 3.1690\ldots$$

<u>maxima</u> and <u>minima</u>

" extreme pts "

<u>Def'n</u> $f : U \subset \mathbb{R}^n \to \mathbb{R}$

$\vec{x}_0 \in U$ is a local min (max) of f if $\exists$ nbhd V of $\vec{x}_0$ s. that $\forall \vec{x} \in V$

$$f(\vec{x}) \geq f(\vec{x}_0)$$

$$\left(f(\vec{x}) \leq f(\vec{x}_0) \right)$$

$\vec{x}_0$ is a critical pt of f if $\nabla f(\vec{x}_0) = 0$

(book allows f to be non-diff. at crit. pt)

a critical pt which is not a local max or min is called a "saddle pt"

First derivative test

If f is differentiable at $\vec{x}_0$ which is a local max or min, then $\vec{x}_0$ is a crit. pt

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

Pf | Fix $\vec{h} \in \mathbb{R}^n$; then $g(t) = f(\vec{x}_0 + t\vec{h})$ has a local extremum at $t = 0$ so

$g'(0) = 0$ i.e (chain rule)

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

since $\vec{h}$ is arbitrary,

$$\nabla f(x_0) = \vec{0}$$

So the first derivative test says we should look among the critical pts of f to find local extremum.

<u>Example</u> Let $f(x,y) = x^2y + y^2x$

Then $f_x = 2xy + y^2 = y(2x+y) = 0$

$f_y = x^2 + 2xy = x(x+2y) = 0$

So we find $x = y = 0$

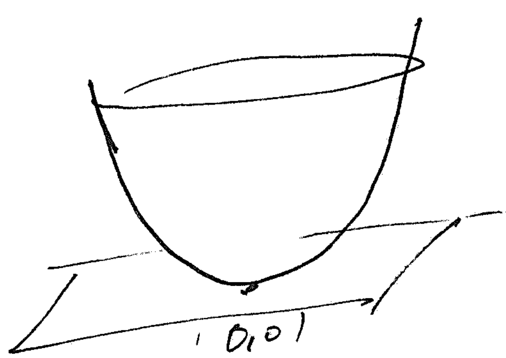
i.e $(0,0)$ only crit. pt

$f(x,y) = xy(x+y)$

We see f vanishes along x + y axes. Along the diagonals $y = \pm x$ we find $f = 2x^3, 0$ respect.

So $(0,0)$ is a saddle pt

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



$(0,0)$ local (and absolute) min.

Example $f(x,y) = x^2 - y^2$ "saddle surface"

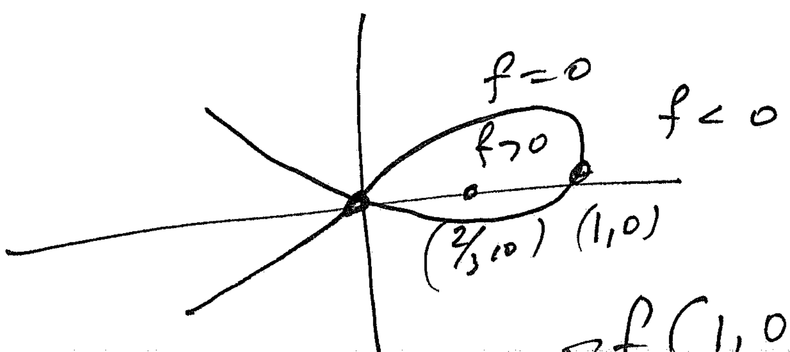
<u>Example</u> $f(x,y) = x^2 - x^3 - y^2$

$f_x = 2x - 3x^2$, $\quad f_y = -2y = 0$

$= x(2-3x)$

So critical pts are $(0,0)$ saddle

$(2/3, 0)$ local max



$\nabla f(1,0) = -i$

<u>Example</u> $f(x,y) = x - x^3 - xy^2$

$f_x = 1 - 3x^2 - y^2$, $\quad f_y = -2xy$

$f = x(1-x^2-y^2)$ (12)

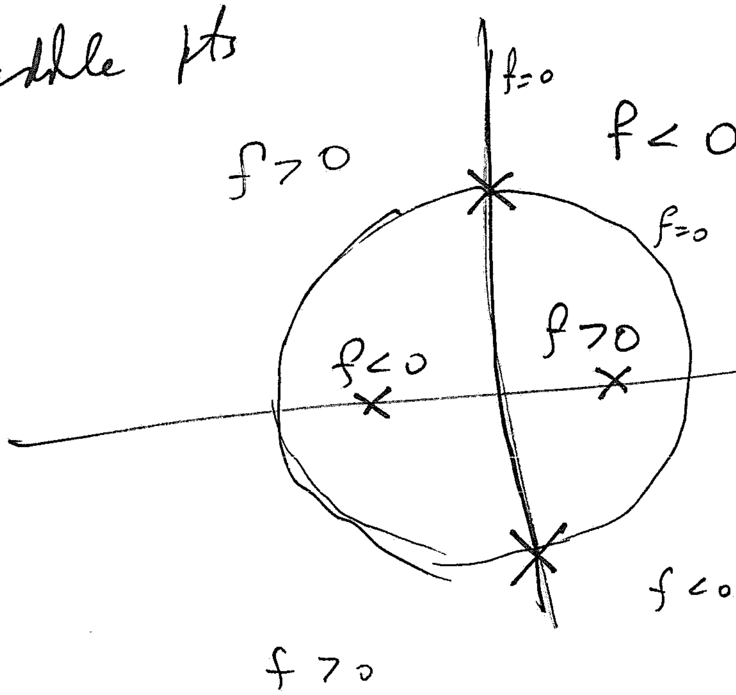
odd in x

critical pts

$1-3x^2-y^2 = 0$

$-2xy = 0 \Rightarrow x=0$ or $y=0$

$x=0$	$y=0$
$y=\pm 1$	$x = \pm \frac{1}{\sqrt{3}} = \pm \frac{\sqrt{3}}{3}$
$(0,1)$ $(0,-1)$ "saddle pts"	$(\frac{\sqrt{3}}{3}, 0)$ local max $(-\frac{\sqrt{3}}{3}, 0)$ local min



The zero set of f has "two pieces"

$x=0$ (y axis)

and unit circle $x^2+y^2=1$

and f odd in x

The second derivative test

<u>Example</u> $f = x^3 + y^3 - 3xy$

$f_x = 3x^2 - 3y$ $\quad f_y = 3y^2 - 3x$

$f_{xx} = 6x$ $\quad f_{xy} = -3$ $\quad f_{yy} = 6y$

critical pts $\quad x^2 - y = 0 \Rightarrow y = x^2$

$-x + y^2 = 0 \Rightarrow$

$$H(f) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$$

$-x + x^4 = 0$

$x(x^3 - 1) = 0$

$x(x-1)\underbrace{(x^2+x+1)}_{\text{no real roots}} = 0$

$x = 0 \quad x = 1$

$y = 0 \quad y = 1$

so $\boxed{(0,0)\ (1,1) \text{ are the critical pts}}$

let's look at the Taylor expansion near $(0,0)$ first

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

$$Hf(0,0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$$

so $f(h_1,h_2) \approx -3h_1h_2 + o(|\vec{h}|^2)$ for $\|\vec{h}\|$ small

Hence $(0,0)$ is a saddle pt.

Now look at $(1,1)$

$$f(1,1) = -1$$

$$H(f)(1,1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

So near $(1,1)$

$$f(1+h_1, 1+h_2) = -1 + \underbrace{3h_1^2 - 3h_1h_2 + 3h_2^2}_{3(h_1^2 - h_1h_2 + h_2^2)} + o(|\vec{h}|^2)$$

quadratic form - complete the square

$$= 3\left(\left(h_1 - \tfrac{1}{2}h_2\right)^2 + \tfrac{3}{4}h_2^2\right)$$ positive definite

Hence $(1,1)$ is a <u>local min</u>

<u>Theorem</u> (Second derivative test)

If (x_0, y_0) is a critical pt of $f(x,y)$ and if

B.5

Quadratic form in 2 variables

$$Q(x,y) = Ax^2 + 2Bxy + Cy^2$$

assume $A \neq 0$

$$Q(x,y) = A\left(\left(x + \frac{B}{A}y\right)^2 + \frac{AC - B^2}{A^2}\, y^2\right)$$

define the discriminant

$$D = AC - B^2$$

So if $D > 0$ (definite case)

$A > 0 \Rightarrow Q$ is positive definite

$A < 0 \Rightarrow Q$ is negative definite

15.6

If $D < 0$ (indefinite case)

$Q(x, y)$ changes sign

The case $D = 0$

is called

$A > 0$ positive semidefinite

$A < 0$ negative semidefinite

$Hf(x_0, y_0)$ is positive ~~(semi)~~ definite

(i.e the quadratic form
$$Q(h_1, h_2) = \frac{1}{2}\left(f_{xx}(x_0, y_0)\, h_1^2 + 2 f_{xy}(x_0, y_0)\, h_1 h_2 + f_{yy}(x_0, y_0)\, h_2^2 \right)$$

is positive definite)

Then (x_0, y_0) is a strict local min

If $Hf(x_0, y_0)$ is negative ~~semi~~ definite then (x_0, y_0) is a (strict) local max

(If $Hf(x_0, y_0)$ is only semi-definite we have either local min or local max but not necessarily strict

If $Hf(x_0, y_0)$ indefinite

then (x_0, y_0) is a saddle pt.

Constrained extrema and Lagrange multipliers

Real world problems are of the form:

Find the min or max of $f(x_1, .., x_n) = f(\vec{x})$

where $\vec{x}$ lies on the level set

(constraint set) $\{x : g(\vec{x}) = c\} = S$

<u>Example</u> maximize $f(x,y) = x+2y$

where (x,y) lies on the unit circle $x^2+y^2=1$

This problem is simple enough that we can do it "by hand"

Introduce polar coordinates

$x = \cos\theta, \quad y = \sin\theta \qquad 0 \le \theta \le 2\pi$

Then $F(\theta) = f(\cos\theta, \sin\theta)$

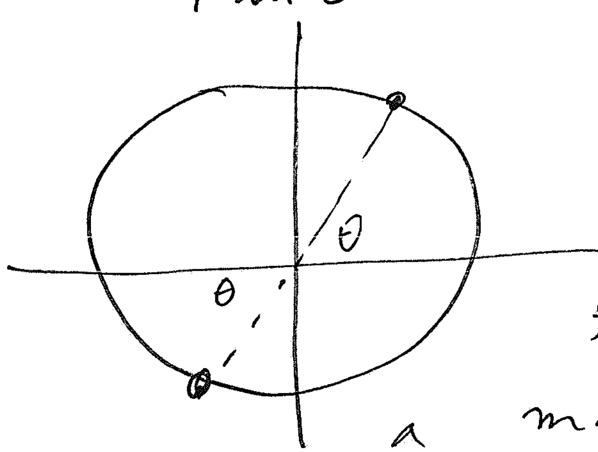
$= \cos\theta + 2\sin\theta \to \max$

This is now an unconstrained calc 1 problem

$F'(\theta) = -\sin\theta + 2\cos\theta = 0$

$\tan\theta = 2 \qquad \theta = \tan^{-1} 2$

(2 critical pts)

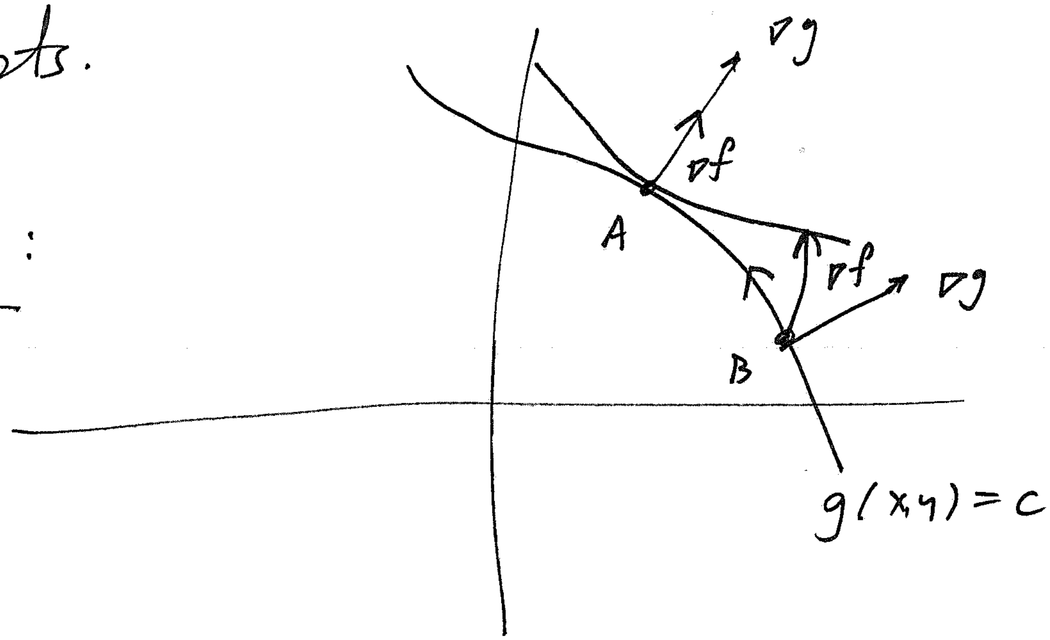


The one in the first quadrant is a max

However in a complicated problem we will probably not be so lucky.

The <u>method of Lagrange</u> tells us where to find the possible critical pts.

<u>idea</u>:



If $\nabla f(B)$ is not perpendicular to the constraint set, then we can move on the constraint set and increase f

So a necessary condition should be that $\nabla f \parallel \nabla g$ at a critical pt.

We can make this precise as follows: Suppose $S = \{\vec{x} : g(\vec{x}) = c\}$ is a "regular surface" i.e $\nabla g \neq 0$ on S and that $\max_{\vec{x} \in S} f(\vec{x}) = f(\vec{P})$.

Then $\nabla f(P) = \lambda \nabla g(P)$.

Lagrange multiplier

<u>Pf</u> Let $\vec{X}(t)$ be a regular curve on S with $\vec{X}(0) = P$. Then $f(\vec{X}(t))$ has a local max at $t=0$ so

$$0 = \frac{d}{dt} f(\vec{X}(t)) \Big|_{t=0} = \nabla f(P) \cdot \vec{X}'(0)$$

On the other hand $g(\vec{X}(t)) = c$

so $\nabla g(P) \cdot \vec{X}'(0) = 0$

~~On the other ha~~ Since S is regular $\vec{X}'(0)$ is arbitrary vector in the tangent plane to S at P, so

we must have $\nabla f(P) = \lambda \nabla g(P)$. //

Note: $\lambda = 0$ is possible

Example: Find the maximum of $f(x,y,z) = x+y+z$ on the sphere $g(x,y,z) := x^2+y^2+z^2 = 1$

$$\nabla g(\vec{x}) = 2\vec{x} = 2(x\vec{i} + y\vec{j} + z\vec{k})$$

$$\nabla f(\vec{x}) = \vec{i} + \vec{j} + \vec{k}$$

So we have to find λ and $\vec{x} \in S$

satisfying

$1 = 2\lambda x$ where $x^2+y^2+z^2=1$

$1 = 2\lambda y$

$1 = 2\lambda z$

square each equation and add:

$3 = 4\lambda^2 \qquad \lambda = \pm\frac{\sqrt{3}}{2}$

$\vec{x}_1 = \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)$ or

$\vec{x}_2 = \left(-\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{3}\right)$

clearly f has a max at $\vec{x}_1$