Remarks:

(i) $f'(\vec{a}) = \nabla f(\vec{a})$ is not a scalar but a vector in V_n (space of all the vectors in $\mathbb{R}^n$). From the geometric view point, it is a linear transformation $T | \mathbb{R}^n \to \mathbb{R}^1$ mapping the vector change $\vec{h} = \langle h_1, \cdots, h_n \rangle$ in $\vec{x}$ into $\mathbb{R}^1$.

i.e. $T\vec{h} = \nabla f(\vec{a}) \cdot \vec{h} = \frac{\partial f}{\partial x_1}(\vec{a}) h_1 + \cdots + \frac{\partial f}{\partial x_n} h_n$ (now that we use T instead of $\vec{T}$ to emphasize that T is a linear transformation).

(ii) In vector notation, $\overbrace{f(\vec{a}+\vec{h}) - f(\vec{a})}^{\Delta y} = \underbrace{\nabla f(\vec{a}) \cdot \vec{h}}_{dy} + o(|\vec{h}|)$. Now that for $\vec{x}$ nearby $\vec{a}$, if we set $\vec{x} = \vec{a} + \vec{h}$ so that $\vec{h} = \vec{x} - \vec{a}$, we have

$$f(\vec{x}) - f(\vec{a}) = \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + o(|\vec{x} - \vec{a}|)$$

$$\Rightarrow \quad f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + o(|\vec{x} - \vec{a}|)$$

$$\Rightarrow \quad \boxed{f(\vec{x}) = L(\vec{x}) + o(|\vec{x} - \vec{a}|)}$$

where $L(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})$ is the linear approximation of f at the point $\vec{a}$ (or the equation of tangent plane to the surface $w = f(\vec{x})$ at the point $(\vec{a}, f(\vec{a}))$).

The error term when we approximate $f(\vec{x})$ by $L(\vec{x})$ at the pt. a is of order $o(|\vec{x} - \vec{a}|)$ showing $L(\vec{x})$ is a good approximation because the error term $\to 0$ much more rapidly than $\vec{x} \to \vec{a}$ (or $|\vec{x} - \vec{a}| \to 0$)

Ex. $f(x,y) = x^2 + y^2$, show and find $f'(1,3)$.

Solution:

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y$$

$f'(1,3)$ if exists must be given by $\nabla f(1,3) = \langle 2, 6 \rangle$.

It remains to verify the "o" condition i.e.

$$f(\vec{x}) = L(\vec{x}) + o(|\vec{x} - \vec{a}|)$$

where $L(\vec{x}) = f(1,3) + \nabla f(1,3) \cdot \langle x-1, y-3 \rangle = 10 + \langle 2, 6 \rangle \cdot \langle x-1, y-3 \rangle$

$= 10 + 2(x-1) + 6(y-3) = 2x + 6y - 10$

Set $\delta(\vec{x}) = f(\vec{x}) - L(\vec{x}) = x^2 + y^2 - 2x - 6y + 10$

$$= (x^2 - 2x + 1) + (y^2 - 6y + 9)$$

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

$$\lim_{(x,y)\to(1,3)} \delta(x,y) = \lim_{(x,y)\to(1,3)} \frac{(x-1)^2 + (y-3)^2}{\sqrt{(x-1)^2 + (y-3)^2}} = \lim_{(x,y)\to(1,3)} \sqrt{(x-1)^2 + (y-3)^2} = 0$$

Hence, $\vec{f}'(1,3) = \langle 2, 6 \rangle$ //

Thm. (Fundamental Theorem of Differential Calculus)

Given $f(\vec{x})$, $\vec{x} \in D$, D being the domain of f in $\mathbb{R}^n$. Suppose $\frac{\partial f}{\partial x_1}(\vec{a}), \dots, \frac{\partial f}{\partial x_n}(\vec{a})$ exist and are continuous $\forall a \in D$; then we have,

$$f(\vec{a} + \vec{h}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \vec{h} + \vec{\varepsilon}(\vec{h}) \cdot \vec{h}$$

where $\vec{\varepsilon}(\vec{h}) = \langle \varepsilon_1(\vec{h}), \dots, \varepsilon_n(\vec{h}) \rangle$ satisfies

$$\lim_{|\vec{h}|\to 0} \varepsilon_1(\vec{h}) = \dots = \lim_{|\vec{h}|\to 0} \varepsilon_n(\vec{h}) = 0$$

Remarks:

(i) $\vec{\varepsilon}(\vec{h}) \cdot \vec{h}$ is the error term when we approximate $\Delta f = f(\vec{a} + \vec{h}) - f(\vec{a})$ by $df = \nabla f(\vec{a}) \cdot \vec{h}$.

(ii) Since $\dfrac{|\vec{\varepsilon}(\vec{h}) \cdot \vec{h}|}{|\vec{h}|} \le |\vec{\varepsilon}(\vec{h})| = \sqrt{\varepsilon_1(\vec{h})^2 + \dots + \varepsilon_n(\vec{h})^2}$.

$$\lim_{|\vec{h}|\to 0} \frac{|\vec{\varepsilon}(\vec{h}) \cdot \vec{h}|}{|\vec{h}|} = 0 \Rightarrow f(\vec{a} + \vec{h}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \vec{h} + o(|\vec{h}|)$$

implying f is differentiable at $\vec{a}$ with $f'(\vec{a}) = \nabla f(\vec{a})$.

Hence, if f has all its 1st order partial derivatives exist and are continuous, then f is also differentiable with its total derivative $f' = \nabla f$.

11° Lagrange Multipliers — Optimization with constraints

Motivating Example: Find the point on the surface $\frac{x^2}{4}+\frac{y^2}{9}+\frac{z^2}{16}=1$ that is farest away from the origin. Let (x,y,z) be any point on the surface, its distance from the origin is given by $\sqrt{x^2+y^2+z^2}$. On the other hand, since (x,y,z) is on the ellipsoid, (x,y,z) also satisfies $\frac{x^2}{4}+\frac{y^2}{9}+\frac{z^2}{16}=1$

The whole problem would be like minimizing the function

$$f(x,y,z)=x^2+y^2+z^2 \quad \text{(our objective function)}.$$

subject to the constraint or side condition,

$$\frac{x^2}{4}+\frac{y^2}{9}+\frac{z^2}{16}=1$$

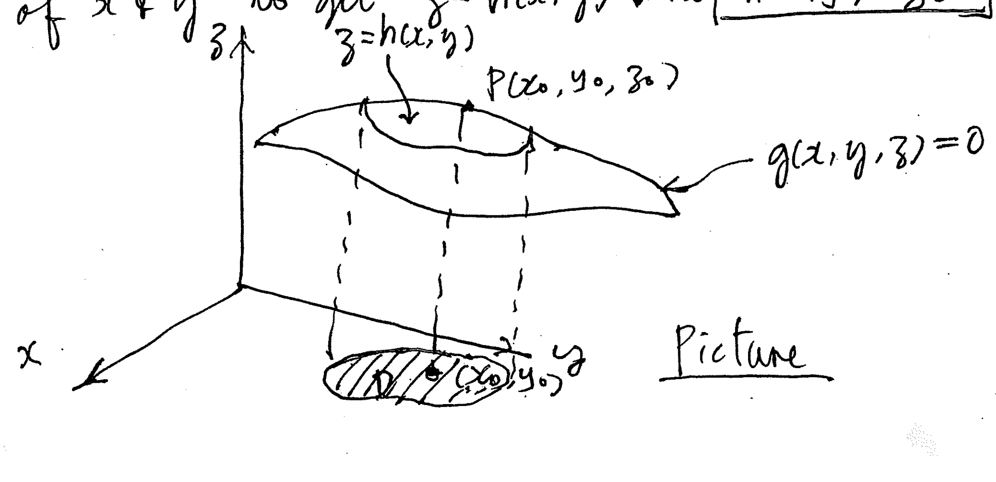
Alternatively, we could look upon it as minimizing

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

on the surface $\frac{x^2}{4}+\frac{y^2}{9}+\frac{z^2}{16}=1$ as its domain.

Put its in the most general setting, suppose we are to optimize a objective function $f(x,y,z)$ subject to a side condition $g(x,y,z)=0$.

Let $P(x_0,y_0,z_0)$ be the optimal point (say. a local maximum) we look for, suppose further that $\boxed{\nabla g(P)\neq\vec{0}}$. Without loss of generality, let us assume $\frac{\partial g}{\partial z}(P)\neq 0$, then locally at P, we could solve z explicitly as a function of x & y to get $z=h(x,y)$ with $\boxed{h(x_0,y_0)=z_0}$



Picture

Locally at $P(x_0, y_0, z_0)$, i.e. on the small patch near P and containing P, we could reduce $f(x, y, z)$ into

$$F(x, y) = f(x, y, h(x, y)), \quad (x, y) \in D \text{ for some domain } D \text{ in } \mathbb{R}^2$$

At P where the optimal point is, we have $\frac{\partial F}{\partial x}(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0) = 0$.

Hence,

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0) + \frac{\partial f}{\partial z}(x_0, y_0, z_0) \cdot \frac{\partial h}{\partial x}(x_0, y_0) = 0 \quad \text{------ (i)}$$

$$\frac{\partial f}{\partial y}(x_0, y_0, z_0) + \frac{\partial f}{\partial z}(x_0, y_0, z_0) \frac{\partial h}{\partial y}(x_0, y_0) = 0 \quad \text{------ (ii)}$$

This is equivalent to having.

$$\nabla f(P) \cdot \left\langle 1, 0, \frac{\partial h}{\partial x}(x_0, y_0) \right\rangle = 0 \text{ and } \nabla f(P) \cdot \left\langle 0, 1, \frac{\partial h}{\partial y}(x_0, y_0) \right\rangle = 0$$

Similarly, locally at P (see above figure), we also have

$$g(x, y, h(x, y)) \equiv 0 \quad \forall x, y \in D$$

implying (on differentiating both sides of the equation),

$$\nabla g(P) \cdot \left\langle 1, 0, \frac{\partial h}{\partial x}(x_0, y_0) \right\rangle = 0 \text{ and } \nabla g(P) \cdot \left\langle 0, 1, \frac{\partial h}{\partial y}(x_0, y_0) \right\rangle = 0$$

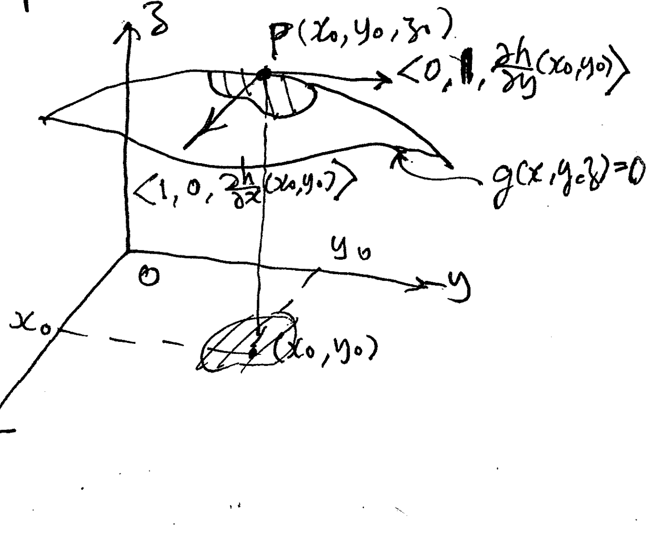
But $\left\langle 1, 0, \frac{\partial h}{\partial x}(x_0, y_0) \right\rangle$ and $\left\langle 0, 1, \frac{\partial h}{\partial y}(x_0, y_0) \right\rangle$ are a pair of non-parallel tangential vectors to the surface $g(x, y, z) = 0$ (or $z = h(x, y)$) near P.

Therefore both of $\nabla f(P)$ and $\nabla g(P)$ are normal to the surface $g(x, y, z) = 0$ at the point P.

Since, $\nabla g(P) \neq \vec{0}$, $\exists \lambda \in \mathbb{R}$ s.t.

$$\boxed{\nabla f(P) = \lambda \nabla g(P)}$$

λ is known as the lagrange multiplier.



Back to the example of minimizing $f(x,y,z) = x^2+y^2+z^2$ subject to the side condition $\frac{x^2}{4}+\frac{y^2}{9}+\frac{z^2}{16}=1$. We have at an optimal point,

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

implying

$$\langle 2x, 2y, 2z \rangle = \lambda \langle \frac{x}{2}, \frac{2y}{9}, \frac{z}{8} \rangle$$

$$\Rightarrow \begin{cases} x = \frac{\lambda}{4}x & \text{--- (i)} \\ y = \frac{\lambda}{9}y & \text{--- (ii)} \\ z = \frac{\lambda}{16}z & \text{--- (iii)} \end{cases}$$

along with the side condition

$$\frac{x^2}{4}+\frac{y^2}{9}+\frac{z^2}{16}-1=0$$

$$(g(x,y,z)=0).$$

Rewriting the system in the standard forms

$$\begin{cases} (1-\lambda/4)x = 0 & \text{--- (i)} \\ (1-\frac{\lambda}{9})y = 0 & \text{--- (ii)} \\ (1-\frac{\lambda}{16})z = 0 & \text{--- (iii)} \end{cases}$$

Since the trivial solution $(x,y,z)=(0,0,0)$ is obviously not a feasible point. we must have

$$\begin{vmatrix} (1-\lambda/4) & 0 & 0 \\ 0 & (1-\frac{\lambda}{9}) & 0 \\ 0 & 0 & (1-\frac{\lambda}{16}) \end{vmatrix} = 0 \quad \text{(in order to have non-trivial solution)}$$

Hence, we must have

$$(1-\lambda/4)(1-\lambda/9)(1-\lambda/16)=0 \Rightarrow \lambda = 4, 9 \text{ or } 16$$

$\lambda=4 \Rightarrow y=z=0$ & $x=\pm 2$, $f(\pm 2,0,0)=4$

$\lambda=9 \Rightarrow x=z=0$ & $y=\pm 3$, $f(0,\pm 3,0)=9$

$\lambda=16 \Rightarrow x=y=0$ & $z=\pm 4$, $f(0,0,\pm 4)=16$

$\therefore$ minimal pt. is at $(\pm 2,0,0)$ & maximal pt. at $(0,0,\pm 4)$, Max distance $=4$//.

Ex. Let x, y and z be the three sides of a triangle, suppose further that the perimeter of the triangle is fixed, prove the triangle attains its maximum area when the triangle is equilateral (i.e. with $x=y=z$).

Pf:

Set $s = \dfrac{x+y+z}{2}$, by Heron's formula, the area A of the $\triangle$ is given by,

$$A = \sqrt{s(s-x)(s-y)(s-z)}$$

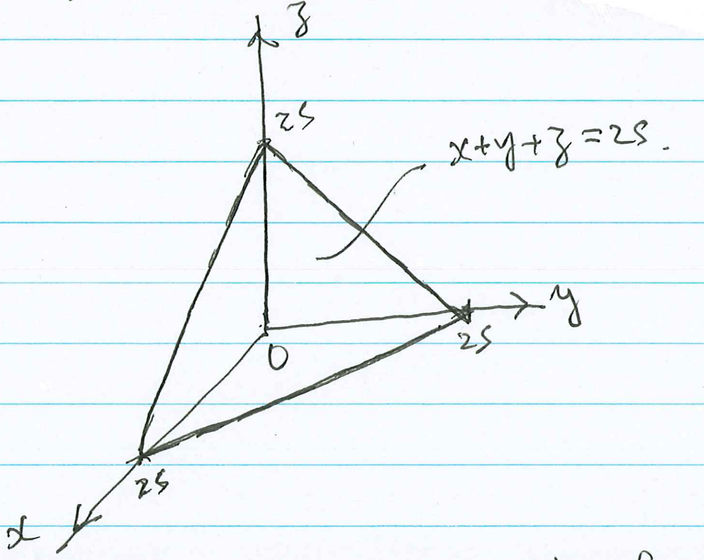
This problem is a natural Larange Multiplier Problem, it boils down to maximizing

$$f(x,y,z) = s(s-x)(s-y)(s-z) \quad (=A^2)$$

subjected to the constraint

$$x+y+z = 2s \quad \text{or} \quad \underbrace{x+y+z-2s}_{g(x,y,z)} = 0 \quad \text{where } s \text{ is fixed.}$$

Geometrically speaking, it would be like maximizing $f(x,y,z)$ over the triangular plane $x+y+z=2s$ in the first octant where $x \geq 0, y \geq 0$ & $z \geq 0$.



When f attains its maximum, by Larange Multiplier

$$\nabla f = \lambda \nabla g \quad (\text{Note that } \nabla g = \langle 1,1,1 \rangle \neq \vec{0}).$$

$$\Rightarrow \langle -s(s-y)(s-z), -s(s-x)(s-z), -s(s-x)(s-y) \rangle = \lambda \langle 1,1,1 \rangle$$

Hence,

$$\begin{cases} -s(s-y)(s-z) = \lambda & \text{------- (i)} \\ -s(s-x)(s-z) = \lambda & \text{------- (ii)} \\ -s(s-x)(s-y) = \lambda & \text{------- (iii)} \end{cases}$$

$$\Rightarrow \quad (s-y)(s-z) = (s-x)(s-z) = (s-x)(s-y) = -\lambda/s$$

Consider $(s-y)(s-z) = (s-x)(s-z)$, we have

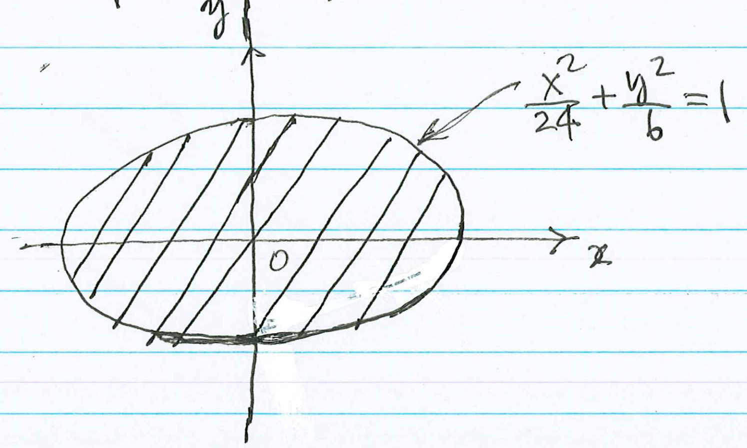
either $z = s$ or $y = x$. But $z = s \Rightarrow x+y = s$ which yields a degenerated triangle whose area is zero, therefore we must have $x = y$ when the Δ attains its maximum area.

Similarly, we also have $z = y$, whence the result.

Ex. Suppose the temperature of a metal plate is given by the temperature function

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

where the metal plate occupies the elliptical region $x^2 + 4y^2 \le 24$



Find the points on the plate where the minimum and the maximum temperatures are being attained and the corresponding minimum and maximum temperatures.

Solution: It boils down to optimize $T(x,y)$ on the elliptical region $x^2+4y^2 \leq 24$. We first look for critical points inside the domain where

$$\frac{\partial T}{\partial x}(x,y) = 2x+2=0 \quad \text{and} \quad 2y=0$$

$\Rightarrow$ $(-1,0)$ is the only critical point.

Then we go to the boundary $\frac{x^2}{24}+\frac{y^2}{6}=1$ and look for the optimal points on the boundary. It becomes a Lagrange Multiplier Problem of optimizing $T(x,y)=x^2+2x+y^2$ subjected to the constraint

$$g(x,y)=\frac{x^2}{24}+\frac{y^2}{6}-1=0$$

At an optimal pt. on the boundary, we have

$$\nabla T(x,y) = \lambda \nabla g(x,y), \quad g(x,y)=0$$

$$\Rightarrow \begin{cases} 2x+2 = 2\lambda x & \text{------ (i)} \\ 2y = 8\lambda y & \text{------ (ii)} \\ x^2+4y^2 = 24 & \text{------ (iii)} \end{cases}$$

$\leftarrow$ easiest one to begin with.

(ii) $\Rightarrow 2y(1-4\lambda)=0 \quad \therefore \lambda = 1/4$ or $y=0$

Case 1° substituting $\lambda = 1/4$ back into (i) $\Rightarrow x = -4/3$

substituting $x=-4/3$ into (iii) $\Rightarrow \left(-\frac{4}{3}, \frac{\sqrt{50}}{3}\right), \left(-\frac{4}{3}, -\frac{\sqrt{50}}{3}\right)$ are critical pts.

Case 2° substituting $y=0$ into (iii) $\Rightarrow (-\sqrt{24},0)$ and $(\sqrt{24},0)$ are critical pts.

Thus we end up with 5 critical pts., on testing out the values of T at those pts, we have

$$T(-1,0)=-1, \quad T\left(-\frac{4}{3}, \frac{\sqrt{50}}{3}\right)=\frac{14}{3}, \quad T\left(-\frac{4}{3}, -\frac{\sqrt{50}}{3}\right)=\frac{14}{3},$$

$$T(\sqrt{24},0)=24+2\sqrt{24}, \quad T(-\sqrt{24},0)=24-2\sqrt{24}$$

Hence, the minimum temperature is given by $T(-1,0)=-1$ // and the maximum temperature is given by $T(\sqrt{24},0)=24+2\sqrt{24}$ //. #.

Ex Prove the inequality A.M. $\geq$ G.M.

Th^m. Let $a_1, \cdots, a_n$ be n positive numbers, then we have

$$\sqrt[n]{a_1 \cdots a_n} \leq \frac{a_1 + \cdots + a_n}{n}$$

(The case when some of the $a_1, \cdots, a_n$ are zero is trivial).

Out-line of proof:

Define $x_1, \cdots, x_n \geq 0$ by

$x_1^2 = \alpha a_1, \cdots, x_n^2 = \alpha a_n$ where $\alpha = \dfrac{1}{a_1 + \cdots + a_n}$ (assuming $\alpha_1, \cdots, \alpha_n$ not all zero).

As a result, we have $0 \leq x_1 \leq 1, \cdots, 0 \leq x_n \leq 1$, so that the domain for $(x_1, \cdots, x_n)$ is a unit super-cube in $\mathbb{R}^n$. Now that proving the original inequality is equivalent to proving

$$(x_1^2 \cdots x_n^2)^{1/n} \leq \frac{1}{n}$$

where $x_1^2 + \cdots + x_n^2 = \alpha(a_1 + \cdots + a_n) = 1$

In the Larange Multiplier's setting, the original problem is reduced or transformed to maximizing

$f(x_1, \cdots, x_n) = x_1^2 \cdots x_n^2$ subjected to the constraint

$g(x_1, \cdots, x_n) = x_1^2 + \cdots + x_n^2 - 1 = 0$

Now that

$$\nabla f = \lambda \nabla g \Rightarrow \begin{cases} 2x_1 x_2^2 \cdots x_n^2 = 2\lambda x_1 & \text{------ (1)} \\ 2x_1^2 x_2 \cdots x_n^2 = 2\lambda x_2 & \text{------ (2)} \\ \vdots & \\ 2x_1^2 \cdots x_{n-1}^2 x_n = 2\lambda x_n & \text{------ (n)} \end{cases}$$

(once more, we may assume $x_1 > 0, \cdots, x_n > 0$ since any one of them being zero would corresponds to the minimum & not the maximum).

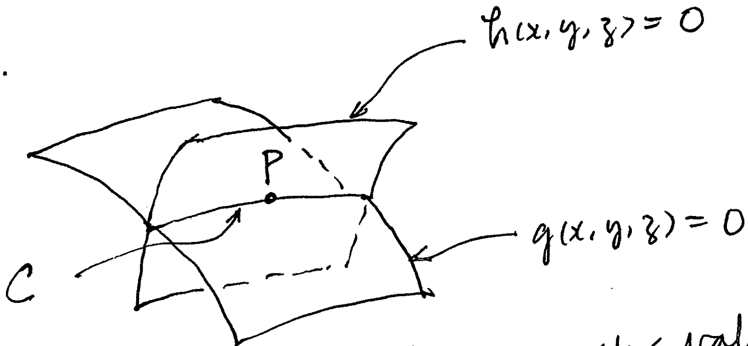
Then we could show that the maximum is attained when $x_1 = \cdots = x_n = \frac{1}{\sqrt{n}}$ with the corresponding value given by $1/n \Rightarrow (x_1^2 \cdots x_n^2)^{1/n} \leq 1/n$ //.

Generalization to the case of having two constraints

Suppose we are to optimize $f(x,y,z)$ subject to the constraints, $g(x,y,z)=0$ and $h(x,y,z)=0$ where f, g & h are all first continuously differentiable (i.e. $f, g, h \in C^1$).

Suppose further that P is an optimal point where $\nabla g(P)$ and $\nabla h(P)$ are both non-zero and non-parallel.

In this case, the problem is equivalent to optimizing $f(x,y,z)$ along the curve C of intersection between the two surfaces $g(x,y,z)=0$ and $h(x,y,z)=0$. P being an optimal point must therefore also be on the curve C.



We could parametrize C by a vector value function $\vec{r}(t)$ with $\vec{r}(t_0) = P$ for some $t_0 \in \mathbb{R}$ s.t. $\vec{r}\,'(t_0) \neq \vec{0}$.

Since P is an optimal point when f is restricted to C, by setting $F(t) = f(\vec{r}(t))$, we must have $F'(t_0)=0$ implying

$$\nabla f(P) \cdot \vec{r}\,'(t_0) = 0 \quad \text{------} \quad (i)$$

Similarly, we also have $g(\vec{r}(t)) \equiv 0$ and $h(\vec{r}(t)) \equiv 0 \ \forall t$ implying in particular,

$$\nabla g(P) \cdot \vec{r}\,'(t_0) = 0 \quad \text{----} \quad (ii), \qquad \nabla h(P) \cdot \vec{r}\,'(t_0) = 0 \quad \text{----} \quad (iii)$$

(i) — (iii) $\Rightarrow$ $\nabla f(P)$, $\nabla g(P)$ and $\nabla h(P)$ are all normal to the non-zero vector $\vec{r}\,'(t_0)$ & must therefore be co-planar.

But by assumption, $\nabla g(P)$ & $\nabla h(P)$ are non-zero & non-parallel, they must form a basis for the plane. Thus, $\exists\, \lambda, \mu \in \mathbb{R}$ s.t.

$$\boxed{\nabla f(P) = \lambda \nabla g(P) + \mu \nabla h(P)}$$

Ex. The plane $x+y+z=12$ intersects the paraboloid $z=x^2+y^2$ in an ellipse. Find the point P that (on the ellipse) is nearest to the origin.

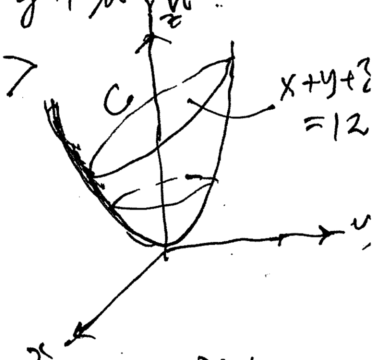
Solution: It boils down to minimize $f(x,y,z)=x^2+y^2+z^2$ subject to the side conditions

$$g(x,y,z)=x+y+z-12=0 \quad \text{and} \quad h(x,y,z)=x^2+y^2-z=0$$

At a optimal point, $\exists \lambda, \mu \in \mathbb{R}$ s.t. $\nabla f = \lambda \nabla g + \mu \nabla h$.

Thus,

$$\langle 2x, 2y, 2z\rangle = \lambda\langle 1,1,1\rangle + \mu\langle 2x, 2y, -1\rangle$$



Geom. Picture

Together with the constraint equations, we have

$$\begin{cases} 2x = \lambda + 2\mu x & \text{--- (i)} \\ 2y = \lambda + 2\mu y & \text{--- (ii)} \\ 2z = \lambda - \mu & \text{--- (iii)} \\ x+y+z-12=0 & \text{--- (iv)} \\ x^2+y^2-z=0 & \text{--- (v)} \end{cases}$$

(i) and (ii) $\Rightarrow$ $\lambda = 2(1-\mu)x$ and $\lambda = 2(1-\mu)y$

Setting them equal $\Rightarrow$ $(1-\mu)x = (1-\mu)y$

Case 1° $\mu = 1$, then $\lambda = 0$ and from (iii) $z = -\frac{1}{2}$ which contradicts (v)

Case 2° $y = x$, subsituting into (v), $z = 2x^2$

subsituting further into (iv), $x + x + 2x^2 - 12 = 0 \Rightarrow x = -3$ or 2

Therefore we come up with 2 optimal points $(-3,-3,18)$ and $(2,2,8)$.

Evaluating f at these points, we have

$f(2,2,8) = 72$ and $f(-3,-3,18) = 342$

Hence, the point that is nearest to the origin is at $(2,2,8)$ and the coorresponding distance from the origin is $\sqrt{72} = 6\sqrt{2}$ //.

On the other hand, the point that is farest away is at $(-3,-3,18)$ & the farest distance is $\sqrt{342} = 3\sqrt{38}$ //.

A final picture: Generalization of Larange Multipliers to more constraints

Given $f \mid \mathbb{R}^n \to \mathbb{R}^1$ which is C^1 ($\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ all exist and are continuous),

suppose we want to optimize f subject to k constraints where $k < n$

$$\begin{cases} g_1(x_1, \dots, x_n) = 0 \\ g_2(x_1, \dots, x_n) = 0 \\ \vdots \\ g_k(x_1, \dots, x_n) = 0 \end{cases}$$

where $k < n$ and $g_1, \dots, g_n \in C^1$ as well

Let P be an optimal point at which $\{\nabla g_1(P), \dots, \nabla g_k(P)\}$ forms a linearly independent set of $\mathbb{R}^n$ (note that since $k \leq n$, it's possible and for $k > n$, it becomes impossible). Then we have

$$\nabla f(P) = \lambda_1 \nabla g_1(P) + \dots + \lambda_k \nabla g_k(P)$$

for some $\lambda_1, \dots, \lambda_k \in \mathbb{R}^1$ (the Larange Multipliers).

Solution for problem 2 of assignment 5

$f(x, y) = y \sin x$, $\frac{\partial f}{\partial x} = y \cos x$, $\frac{\partial f}{\partial y} = \sin x$.

At a pt. (a, b), $f'(a, b)$ if exists, has to be given by $\nabla f(a, b) = \langle b\cos a, \sin a \rangle$

It remains to verify the "o" condition i.e.

$$f(a+h, b+k) - f(a, b) = \nabla f(a, b) \cdot \langle h, k \rangle + o(\sqrt{h^2+k^2})$$

Set
$$\begin{aligned} \delta(a, b; h, k) &= f(a+h, b+k) - f(a, b) - \nabla f(a, b) \cdot \langle h, k \rangle \\ &= (b+k)\sin(a+h) - b\sin a - (b\cos a)h - (\sin a)k \\ &= b\left[\sin(a+h) - \sin a - (\cos a)h\right] + k\left[\sin(a+h) - \sin a\right] \\ &= b\, o(h) + o(\sqrt{h^2+k^2}) \quad \text{as } \sqrt{h^2+k^2} \to 0 \\ &= o(\sqrt{h^2+k^2}) \quad \text{whence the result} // \end{aligned}$$