Lecture 17

New question: Suppose we are studying a function $f(x,y)$, but are only interested in the values of f along a curve in the domain of f (the xy-plane). Or on some closed domain where we are assured an extremum can occur.

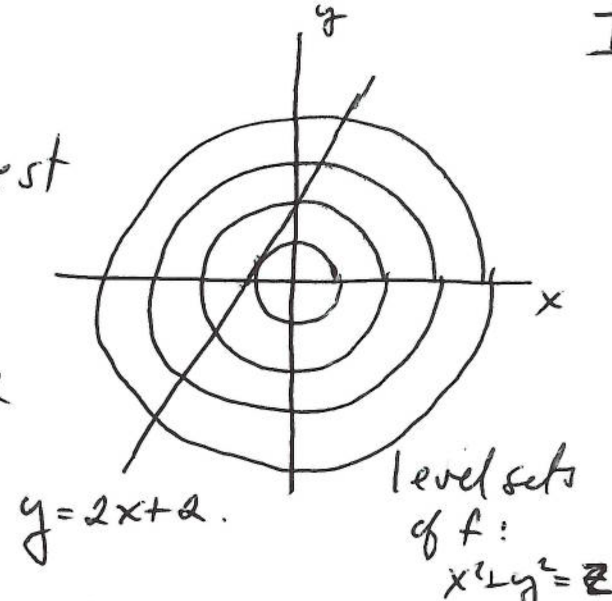
Then we are <u>constraining</u> f in its domain and possibly looking for <u>constrained</u> <u>extrema</u>:

ex. Constrain $f(x,y) = x^2 + y^2$ to the points on the curve $y = 2x + 2$. Then, f varies along the curve as $\frac{d}{dt}(f \circ \vec{c})(t) = Df(\vec{c}(t)) \cdot \vec{c}\,'(t)$,

where $\vec{c}(t) = \begin{bmatrix} t \\ 2t+2 \end{bmatrix}$

Q: Does f, along $\vec{c}$, have extrema?

A: Yes, f will attain its lowest value along $\vec{c}(t)$ precisely where $\vec{c}$ is closest to the origin (do you see this?)



f, along $\vec{c}$, will also be critical, where $\vec{c}$ is tangent to a level set of f. (why is this?)

<u>Method of Lagrange Multipliers</u>

Let $f: U \subset \mathbb{R}^n \to \mathbb{R}$, and $g: U \subset \mathbb{R}^n \to \mathbb{R}$ be C^1, with $\vec{x}_0 \in U$ a pt on the c-level set of g,

Thm 3 p. 186

$$S_c = \{ \vec{x} \in \mathbb{R}^n \mid g(\vec{x}) = g(\vec{x}_0) = c \}$$

Then, if f, restricted to S_c, denoted $f|_{S_c}$, has an extremum at $\vec{x}_0$, then there exists $\lambda \in \mathbb{R}$, where

$$\nabla f(\vec{x}_0) = \lambda g(\vec{x}_0).$$

Notes ① The real number, λ, which may be 0, is not important. The fact that $\nabla f(\bar{x}_0)$ is a multiple of (is parallel to) $\nabla g(\bar{x}_0)$ is !!!

② We call $\bar{x}_0$, in this case, a critical pt of $f|_S$ (though it usually is not a critical pt of f).

ex. For $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + y^2$, define
$g: \mathbb{R}^2 \to \mathbb{R}$, $g(x,y) = y - (2x+2)$, so that our curve $y = 2x+2$ is the 0-level set of $g(x,y)$.

Is there a place $\vec{x}_0 = (x_0, y_0)$ where along $y = 2x+2$, $f(x,y)$ has a local extremum?

Calculate $\nabla f(\vec{x}) = \lambda \nabla g(\vec{x})$ and look for a solution.

$$\nabla f(\vec{x}) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \overset{?}{=} \lambda \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \lambda \nabla g(\vec{x})$$

This gives us 2 equations: $2x = \lambda(-2)$,
$2y = \lambda(1)$

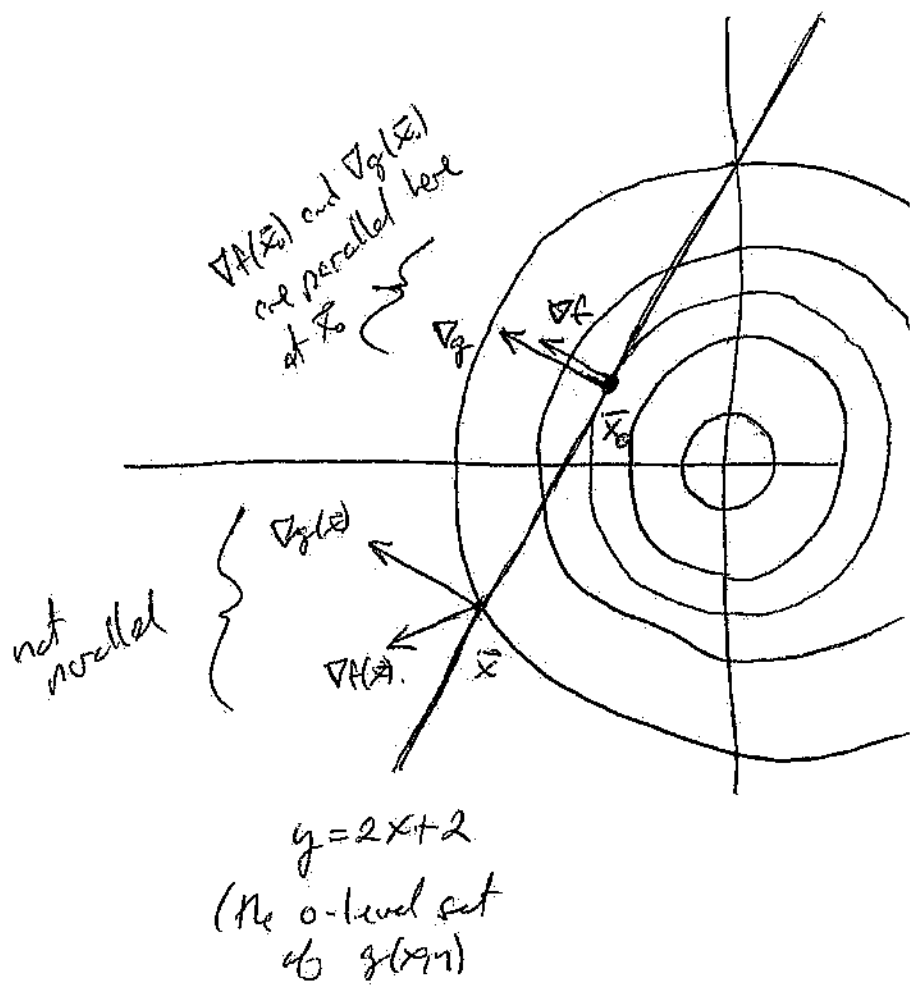
ex cont'd

The first makes sense when $x = -\lambda$ and the second when $y = \lambda/2$. And since this point must be on the line $g(x,y) = y-(2x+2)=0$ we get

$$y-(2x+2)=0 = \frac{\lambda}{2} - (2(-\lambda)+2) = 0$$
$$= \frac{5}{2}\lambda - 2 = 0$$

or $\lambda = \frac{4}{5}$.

Thus $\bar{X}_0 = \begin{bmatrix} -4/5 \\ 2/5 \end{bmatrix}$.



<u>example cont'd.</u>

Why is this point special? Because f, evaluated along the curve $\vec{c}(t) = \begin{bmatrix} t \\ 2t+2 \end{bmatrix}$ stops falling and starts rising at this point (visually one can see this: $\vec{c}(t)$ stops cutting through smaller and smaller level sets of f and starts cutting through larger and larger ones after $\vec{x}_0$.

At a pt where $\nabla f = \lambda \nabla g$, $\vec{c}\,'(t)$ is tangent to the level set. $\frac{d}{dt}(f \circ \vec{c})(t) = 0$ there.

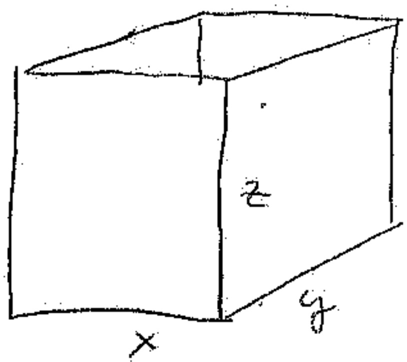
Another way to see this?

$$\frac{d}{dt}(f \circ \vec{c})(t) = \frac{d}{dt} f(t, 2t+2) = \frac{d}{dt}(t^2 + (2t+2)^2)$$
$$= \frac{d}{dt}(5t^2 + 8t + 4)$$
$$= 10t + 8$$

Here $\frac{d}{dt}(f \circ \vec{c})(t) = 0$ when $t = \frac{-4}{5}$. But this is the pt $\vec{c}(\frac{-4}{5}) = \begin{bmatrix} -4/5 \\ 2/5 \end{bmatrix} = \vec{x}_0$

And since $\frac{d}{dt}(f \circ \vec{c})(t) < 0$ for $t < -\frac{4}{5}$, and $\frac{d}{dt}(f \circ \vec{c})(t) > 0$ for $t > \frac{4}{5}$, by the 1st derivative test, $t = -4/5$ is a local min for $(f \circ \vec{c})(t)$.

ex. Given a box of dimensions x, y, z (sides and bottom, no top), which size box that can contain $V = 2048\,cm^3$ minimizes surface area?



Strategy: This is a standard optimization problem in Calculus I where the base of the box is square of size x with z the second variable. With also a top, the optimal box size was a cube ($z = x$). With no top, the optimal size was $z = \frac{1}{2}x$. Here, we use Lagrange Multipliers to solve.

Solution: Here, we seek to minimize

$$f(x,y,z) = 2xz + 2yz + xy$$

subject to

$$\underbrace{g(x,y,z) = xyz = 2048}_{\text{Call this } S.}$$

ex. Solution (cont'd).

According to the method, any critical pts of f along the level set of g must satisfy the system

$$\nabla f(\vec{x}) = \lambda \nabla g(\vec{x}), \quad 2048 = xyz$$

(we assume $x, y, z > 0$ since they are lengths)

$$\nabla f(\vec{x}) = \begin{bmatrix} 2z+y \\ 2z+x \\ 2x+2y \end{bmatrix} = \lambda \begin{bmatrix} yz \\ xz \\ xy \end{bmatrix} = \lambda \nabla g \quad , \text{ or}$$

I name these for bookkeeping purposes.

$$\left\{ \begin{array}{ll} ① & 2z+y = \lambda yz \\ ② & 2z+x = \lambda xz \\ ③ & 2x+2y = \lambda xy \\ ④ & 2048 = xyz \end{array} \right\} \text{ 4 eqns in 4 unknowns.}$$

Note: This system is nonlinear, so writing this as a matrix system will not work.

First, try whatever works to start eliminating variables from the equations, just like for linear systems.

ex Solution (cont'd)

First, use ④ to eliminate z from ① and ②
since ① and ② are similar ~~in~~

Since by ④ $z = \frac{2048}{xy}$, sub into ① and ②:

① $\frac{4096}{xy} + y = \lambda\left(\frac{2048}{x}\right)$, or $\frac{4096}{y} + xy = 2048\lambda$

② $\frac{4096}{xy} + x = \lambda\left(\frac{2048}{y}\right)$, or $\frac{4096}{x} + xy = 2048\lambda$

These last 2 eqns can only both be true if $x = y$ why?

Hence by ③ $\left.\begin{array}{r} \underbrace{2x + 2y} = \lambda xy \\ 4x = \lambda x^2 \end{array}\right\} \Rightarrow x = \frac{4}{\lambda} = y$

Back to ① $2z + \left(\frac{4}{\lambda}\right) = \lambda\left(\frac{4}{\lambda}\right) z \implies \frac{4}{\lambda} = 2z$, or $z = \frac{2}{\lambda}$

And by ④, $xyz = \left(\frac{4}{\lambda}\right)\left(\frac{4}{\lambda}\right)\left(\frac{2}{\lambda}\right) = 2048 = \frac{32}{\lambda^3} \Rightarrow \lambda = \frac{1}{4}$

Hence only critical pt of f|S is $\vec{x} = (16, 16, 8)$

Q: To actually show this is a local min (and hence a global one since it is alone) can be tricky.
Do you have any ideas?