Lecture 32: Optimization

In the last lecture, we were mainly concerned with local extrema of interior points of a set, showing that — for a differentiable function — these must be stationary points (where the gradient $= \vec{0}$). What about global extrema?

Theorem 1: (i) If $\mathcal{S} \subseteq \mathbb{R}^n$ is a closed, bounded set, and f is continuous on $\mathcal{S}$, then it attains (absolute/global) max + min on $\mathcal{S}$.

(ii) These extrema can occur at

- $\partial \mathcal{S}$ = boundary of $\mathcal{S}$
- singular points of $f(\vec{x})$ (i.e. f is not differentiable on a neighborhood of the point)
- stationary points of $f(\vec{x})$ (in $\mathrm{int}(\mathcal{S})$ by definition).

I defer the proof of (i) for now; (ii) is immediate from yesterday's result: let $\mathcal{S}_0 \subseteq \mathcal{S}$ be the interior of the subset on which f is differentiable, & let $\vec{a}$ be an extremum of f. If $\vec{a} \in \mathcal{S}_0$, we know $\vec{\nabla} f(\vec{a}) = \vec{0}$. Otherwise, either $\vec{a} \in \partial \mathcal{S}$ or $\vec{a} \in \mathrm{int}(\mathcal{S}) \setminus \mathcal{S}_0$ = singular points.

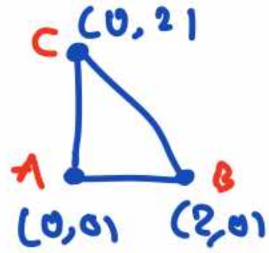
<u>Upshot</u>: we get an algorithm for locating f's extrema, known as <u>optimizing</u> f:

$\boxed{\text{Step 1}}$ minimize/maximize f on $\partial\mathcal{S}$

$\boxed{\text{Step 2}}$ evaluate f at stationary points (& any obvious bad pts.)

$\boxed{\text{Step 3}}$ pick the largest/smallest from Steps 1 & 2.

Ex 1/ Optimize $f(x,y) = 8xy - x - y$ on the closed triangular region shown:



- $\overline{AC}$: $y=0$. $f(x,0) = -x$ $\nearrow$ max at $A: 0$, $\searrow$ min at $C: -2$
- $\overline{AB}$: $x=0$. $f(0,y) = -y$ $\nearrow$ max at $A: 0$, $\searrow$ min at $B: -2$
- $\overline{BC}$: $f(2-t,t) = 16t - 8t^2 - 2$

  $\Rightarrow 0 = \frac{d}{dt} f(2-t,t) = 16 - 16t$ gives $t=1$, and $f(1,1) = 6$
- interior: $\vec{0} = \vec{\nabla} f = (8y-1, 8x-1)$

  $\Rightarrow (x,y) = (\frac{1}{8}, \frac{1}{8})$, at which $f = -\frac{1}{8}$ (it's a saddle point, but we don't care)

Conclude that overall $\diagup$ max $= 6$, at $(1,1)$; $\diagdown$ min $= -2$, at B & C. //

Ex 2/ Optimize $f(x,y) = 2x^2 + y^2 - 4x - 2y + 5$ on the closed set $\mathcal{S} = \{(x,y) \mid x^2 + \frac{y^2}{2} \leq 1\}$.

- $\partial\mathcal{S}$ = ellipse, parametrized by $(x,y) = (\cos t, \sqrt{2}\sin t)$

Write $F(t) = f(x(t), y(t))$, so that Chain Rule $\Rightarrow$

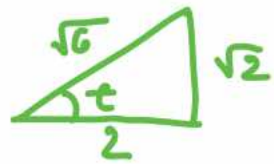
$$F'(t) = f_x(x(t),y(t))\,x'(t) + f_y(x(t),y(t))\,y'(t)$$

$$= (4x(t)-4)(-\sin t) + (2y(t)-2)(\sqrt{2}\cos t)$$

$$= (\cancel{4\cos t}-4)(-\sin t) + (\cancel{2\sqrt{2}\sin t}-2)(\sqrt{2}\cos t)$$

$$= 4\sin t - 2\sqrt{2}\cos t.$$

Setting $0 = F'(t)$ gives $\tan(t) = \frac{\sqrt{2}}{2}$



$\Rightarrow x(t) = \cos\left(\underset{+\pi}{\arctan}\left(\frac{\sqrt{2}}{2}\right)\right) = -\frac{2}{\sqrt{6}}$

$y(t) = \sqrt{2}\sin\left(\underset{+\pi}{\arctan}\left(\frac{\sqrt{2}}{2}\right)\right) = -\frac{2}{\sqrt{6}}$

$f\left(\frac{2}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right) = 7 - 2\sqrt{6}$, $f\left(-\frac{2}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right) = 7 + 2\sqrt{6}$

- interior : $\vec{0} = \vec{\nabla} f = (4x-4, 2y-2) \Longrightarrow (x,y) = (1,1)$

... and $f(1,1) = 2$ $(< 7 - 2\sqrt{6})$ is the minimum

NOOOO! $(1,1)$ is not in S !

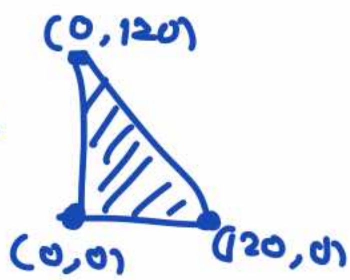
So $7 \pm 2\sqrt{6}$ are the max/min values. //

Ex 3/ Find 3 positive numbers that add up to 120 & such that their product is a maximum:

- What are the set S & function f ?

Want to maximize xyz on $\begin{cases} x+y+z = 120 \\ x,y,z \geq 0 \end{cases}$

i.e. $f(x,y) = xy(120-x-y)$ on $S =$



- f is $\underline{\underline{0}}$ on ∂S

- so consider $\vec{0} = \vec{\nabla} f = \vec{\nabla}(120xy - x^2y - xy^2)$

$\Rightarrow (0,0) = (120y - 2xy - y^2,\ 120x - 2xy - x^2)$

$\Rightarrow y(120 - 2x - y) = 0 = x(120 - 2y - x)$

$\Rightarrow 120 - 2x - y = 0 = 120 - 2y - x$

don't want x or y = 0 $\qquad \Rightarrow y = x$

$\Rightarrow 120 - 3x = 0 \Rightarrow x = 40 \Rightarrow y = 40 \Rightarrow z = 40$

$\Rightarrow \max = 40^3 = 64{,}000.$ //

Ex 4/ Find the minimum distance between the origin and the surface $z^2 = x^2y + 4$:

- let $P = (x,y,z)$ be any point on the surface.

$(\text{dist}(P,0))^2 = x^2 + y^2 + z^2$ is easier to work with and optimizing it is equivalent to optimizing $\text{dist}(P,0)$

So consider $f(x,y) = x^2 + y^2 + (x^2y + 4)$ on $S_0 = \{(x,y) \mid x^2y + 4 \geq 0\}$.
(why?)

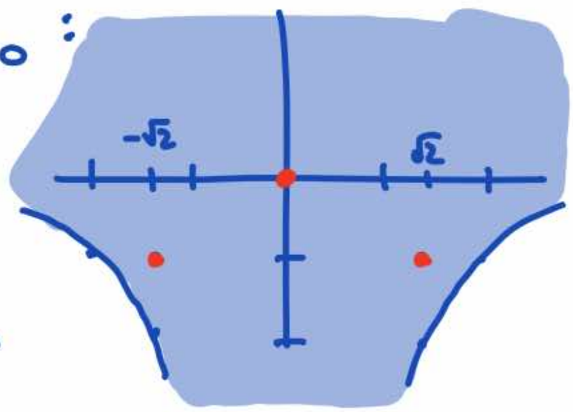
- start by finding the critical points:

$\vec{0} = \nabla f = (2x + 2xy,\ 2y + x^2) \Rightarrow y = \frac{-x^2}{2}$

$\Rightarrow 0 = 2x + 2xy = 2x - x^3 \Rightarrow x = 0, \pm\sqrt{2}$

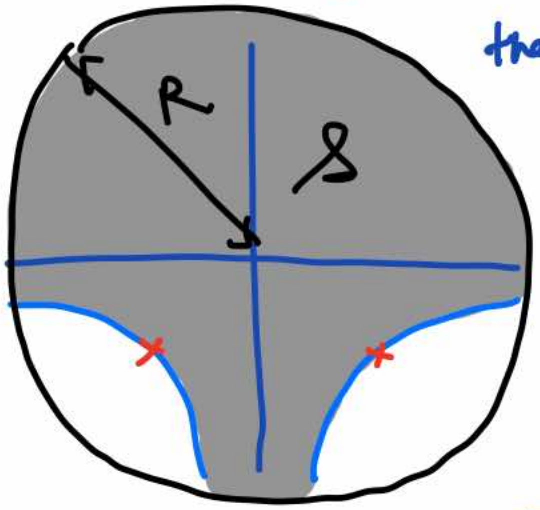
$\Rightarrow (x,y) = (0,0), (\sqrt{2}, -1), (-\sqrt{2}, -1) \Rightarrow f = 4, 5, 5$

- now look at S_0:



It is defined by $y \geq -4/x^2$

This isn't bounded, so let S be its intersection with a big disk of radius R: on the part of the boundary



that's circular, $f \geq R^2$ which we can take arbitrarily large. On the curve $t \mapsto (t, -4/t^2)$ $(t \neq 0)$, we have $f = x^2 + y^2 = t^2 + \frac{16}{t^4}$ which is minimised when $0 = f' = 2t - \frac{64}{t^5}$

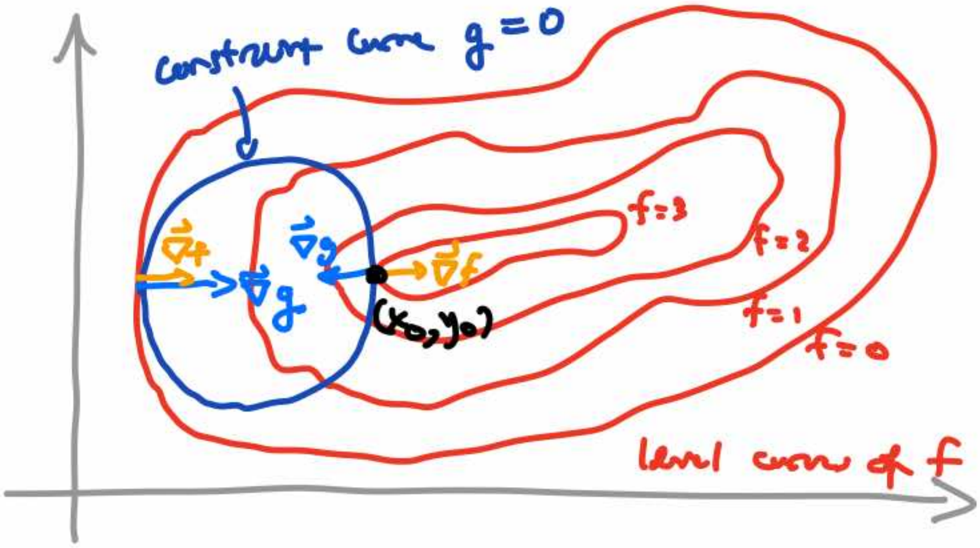
$\Rightarrow t^6 = 32$ $\Rightarrow f = 2^{5/3} + 2^{4/3} = 3 \cdot 2^{2/3}$
$t = 2^{5/6}$

- Since $3 \cdot 2^{2/3} > 4$, the closest points are at $(0,0,\pm 2)$ with distance $\sqrt{4} = 2$.

//

The last 2 examples are constrained extremum problems, which in general can be quite difficult: in particular, sometimes you can't use the constraint to solve for one variable in terms of the others.

Fortunately, there is a better approach: Consider first the case where we want to optimize $\boxed{f(x,y)}$ subject to $\boxed{g(x,y) = 0}$:



We shall take it to be geometrically evident that the level curve $f = k$ with greatest possible k intersecting the constraint curve is tangent to it at the intersection point(s). Hence, their normal vectors are parallel, and

$$\vec{\nabla} f = \lambda \vec{\nabla} g$$

at (x_0, y_0) on $g = 0$ where f is maximized.

A little less heuristically, if $\vec{r}(t) = (x(t), y(t))$ parametrizes $g(x,y) = 0$, then the function f is maximized on $g = 0$ where

$$0 = \frac{d}{dt} f(\vec{r}(t)) = \vec{\nabla} f(\vec{r}(t)) \cdot \vec{r}\,'(t)$$

for $t =$ some t_0; and then $\vec{\nabla} f(\vec{r}(t_0)) \perp \vec{r}\,'(t_0)$.

But $\vec{\nabla} g(\vec{r}(t_0)) \perp \vec{r}\,'(t_0)$ as well (since gradient is normal to $g = 0$, and $\vec{r}\,'(t_0)$ is tangent to it), so $\vec{\nabla} f(\vec{r}(t_0))$ is parallel to $\vec{\nabla} g(\vec{r}(t_0))$.

This approach to solving constrained extremum problems generalizes to 3, 4, etc. variables. For (say) 3 variables, here's the step-by-step:

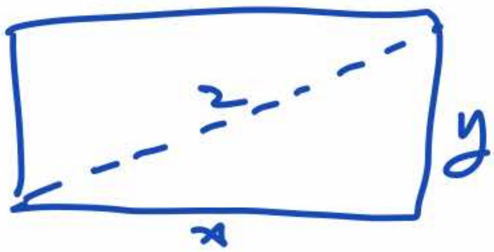
$\boxed{\text{Step 1}}$ Make sure the constraint is in the form $g(x, y, z) = 0$. [e.g. $x^2 + y^2 = z^2$ is not)

$\boxed{\text{Step 2}}$ Set $\vec{\nabla} f = \lambda \vec{\nabla} g$

$\boxed{\text{Step 3}}$ Solve the resulting set of equations (3 from step 2, 1 from Step 1 : $g=0$) to find (x,y,z).

$\boxed{\text{Step 4}}$ Evaluate f at each of the points found, to see where it is really largest/smallest.

Ex 5/ Find the greatest area a rectangle can have with diagonal of length 2.



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

2) $\vec{\nabla} f = \lambda \vec{\nabla} g \rightarrow (y,x) = \lambda(2x, 2y)$

3) $y = 2\lambda x$, $x = 2\lambda y$, $x^2+y^2 = 4$

$\Rightarrow xy = 2\lambda x^2$, $xy = 2\lambda y^2$ $\Longrightarrow$ $\underbrace{\lambda = 0}_{\substack{\text{impossible: would give} \\ (x,y)=(0,0) \Rightarrow g \neq 0}}$ or $x^2 = y^2$

$\Rightarrow \begin{cases} 2x^2 = 4 \\ x = y \end{cases} \Rightarrow x = \sqrt{2} = y$

4) $f(\sqrt{2}, \sqrt{2}) = 2 = A_{max}$. //

The method we've just been describing is that of Lagrange multipliers. We will continue with this on Wednesday.