Lecture 33: Lagrange multipliers

We begin with some examples.

Ex 1/ Find the max/min values of $f(x,y) = y^2 - x^2$ on the ellipse $\frac{x^2}{4} + y^2 = 1$.

Following the steps from the end of Lecture 32,

① $g(x,y) = \frac{x^2}{4} + y^2 - 1$

② $\vec{\nabla} f = \lambda \vec{\nabla} g \Rightarrow (-2x, 2y) = \lambda(\frac{1}{2}x, 2y)$

$\Rightarrow$ (A) $-2x = \frac{\lambda x}{2}$, (B) $2y = 2\lambda y$, and

(C) $\frac{x^2}{4} + y^2 = 1$ (or $x^2 + 4y^2 = 4$)

③ (A) $\Rightarrow -4x = \lambda x \Rightarrow$ (i) $\lambda = -4$ or (ii) $x = 0$

(i): (B) becomes $2y = -8y \Rightarrow y = 0$, whence

(C) yields $x = \pm 2 \rightsquigarrow (\pm 2, 0)$

(ii): (C) becomes $y^2 = 1 \Rightarrow y = \pm 1 \rightsquigarrow (0, \pm 1)$

④ Plug in to f:

$f(2,0) = -4 = f(-2,0)$ MIN

$f(0,1) = 1 = f(0,-1)$ MAX

//

Ex 2 / Optimize $f(x,y) = e^{-xy}$ on the elliptical disk $x^2 + 4y^2 \le 1$.

First, let's find the critical (stationary) points in the interior:

$\vec{0} = \vec{\nabla} f = (-ye^{-xy}, -xe^{-xy}) \Rightarrow (x,y) = (0,0)$ at which $f(0,0) = 1$.

Turning to the boundary, to find max/min there we use Lagrange with $g(x,y) = x^2 + 4y^2 - 1$. $\rightsquigarrow$

$\vec{\nabla} f = \lambda \vec{\nabla} g$ is $(-ye^{-xy}, -xe^{-xy}) = \lambda(2x, 8y)$

- $-ye^{-xy} = 2\lambda x \xrightarrow{\cdot x} -xye^{-xy} = 2\lambda x^2$
- $-xe^{-xy} = 8\lambda y \xrightarrow{\cdot y} -xye^{-xy} = 8\lambda y^2$

$\Rightarrow 2\lambda x^2 = 8\lambda y^2 \Rightarrow \cancel{\lambda = 0}$ or $2x^2 = 8y^2$

implies $(x,y) = (0,0)$, not on boundary

$\Rightarrow x = \pm 2y \Rightarrow 8y^2 = 1 \Rightarrow y = \pm 1/2\sqrt{2} \Rightarrow x = \pm 1/\sqrt{2}$

use $x^2 + 4y^2 = 1$

Now compare $f(1/2\sqrt{2}, 1/\sqrt{2}) = f(-1/2\sqrt{2}, -1/\sqrt{2}) = e^{-1/4} < 1$ MIN

$f(-1/2\sqrt{2}, 1/\sqrt{2}) = f(1/2\sqrt{2}, -1/\sqrt{2}) = e^{1/4} > 1$ MAX

GENERAL "ANSATZ" FOR LAGRANGE MULTIPLIERS

To find all candidates for extrema of $f(x_1, \dots, x_n)$ subject to m constraints

$$g_1(\vec{x}) = 0, \; g_2(\vec{x}) = 0, \; \dots, \; g_m(\vec{x}) = 0,$$

we need to solve the system

$$(*) \begin{cases} g_1(\vec{x}) = 0 \\ \vdots \\ g_m(\vec{x}) = 0 \\ \vec{\nabla} f(\vec{x}) = \lambda_1 \vec{\nabla} g_1(\vec{x}) + \dots + \lambda_m \vec{\nabla} g_m(\vec{x}) \end{cases}$$

this last equation is really n equations

$$\begin{cases} \partial f / \partial x_1 = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_1} \\ \vdots \\ \partial f / \partial x_n = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_n} \end{cases}$$

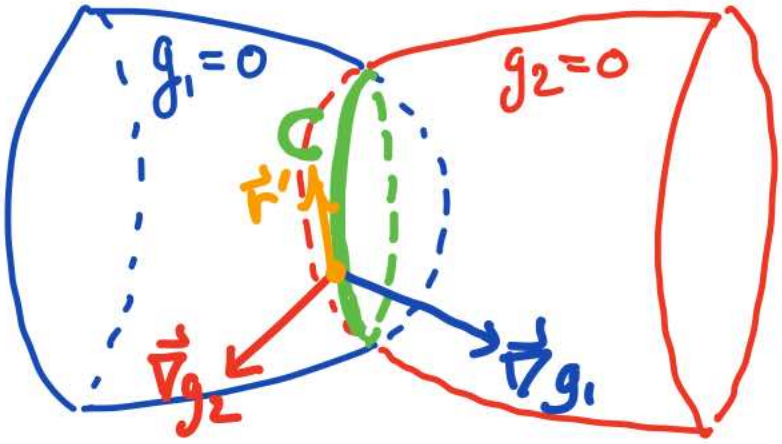
Altogether we have $n+m$ equations in $n+m$ variables.

You would expect the solution set to be a finite set of points, which will include our desired min/max.

WHY DOES IT WORK?

Say $n=3$, $m=2$, and we are trying to max/minimize $f(x,y,z)$ on $C = \{g_1 = 0\} \cap \{g_2 = 0\}$.

If $\vec{r}(t)$ parametrizes C, then $\vec{r}\,'(t)$ is tangent to C, hence tangent to $g_1=0$ & $g_2=0$. Thus $\vec{r}\,' \perp \vec{\nabla} g_1$ & $\vec{\nabla} g_2$.



At an extremum of f on C (at $\vec{r}_0 = \vec{r}(t_0)$),

$$0 = \frac{d}{dt} f(\vec{r}(t))\Big|_{t=t_0} = \vec{\nabla} f(\vec{r}_0) \cdot \vec{r}\,'(t_0) \Rightarrow \vec{r}\,' \perp \vec{\nabla} f.$$

Therefore, if $\vec{\nabla} g_1(\vec{r}_0)$ & $\vec{\nabla} g_2(\vec{r}_0)$ span the subspace $\perp$ to $\vec{r}\,'(t_0)$, we have

$$(\#) \qquad \vec{\nabla} f(\vec{r}_0) = \lambda_1 \vec{\nabla} g_1(\vec{r}_0) + \lambda_2 \vec{\nabla} g_2(\vec{r}_0)$$

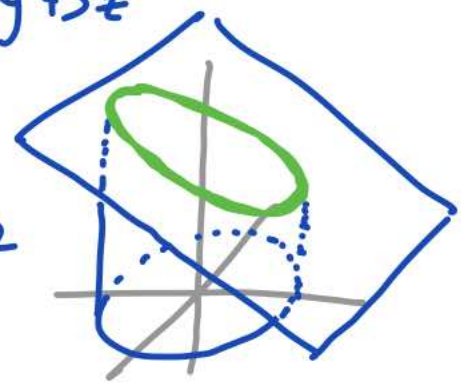
for some $\lambda_1, \lambda_2 \in \mathbb{R}$.

<u>Remark</u>: The $\vec{\nabla} g_1, \vec{\nabla} g_2$ will span as long as they are independent, i.e. $\vec{\nabla} g_1 \times \vec{\nabla} g_2 \neq \vec{0}$.

If $\vec{\nabla} g_1 \times \vec{\nabla} g_2 = \vec{0}$ at some point $\vec{r}_0$, the analysis doesn't work and $\vec{r}_0$ could be an extremum without $(\#)$ holding.

So in general, we really need to throw into the critical set, along with solutions to $(*)$, all points of $g_1 = \cdots = g_m = 0$ where $\vec{\nabla} g_1, \ldots, \vec{\nabla} g_m$ are <u>dependent</u>.

Ex 3/ Optimize $f(x,y,z) = x+2y+3z$ on the ellipse that is the intersection of the cylinder $x^2+y^2=2$ with the plane $y+z=1$.



$g_1(x,y,z) = x^2+y^2-2 \rightarrow x^2+y^2=2$

$g_2(x,y,z) = y+z-1$

$\vec{\nabla} f = \lambda_1 \vec{\nabla} g_1 + \lambda_2 \vec{\nabla} g_2 \rightsquigarrow (1,2,3) = \lambda_1(2x,2y,0) + \lambda_2(0,1,1)$

$$\rightsquigarrow \begin{cases} 1 = 2\lambda_1 x & \rightarrow x = \frac{1}{2\lambda_1} \\ 2 = 2\lambda_1 y + \lambda_2 & \\ 3 = \lambda_2 & \end{cases} \rightarrow y = \frac{-1}{2\lambda_1}$$

$\lambda_1 = \pm\frac{1}{2} \Leftarrow 2\frac{1}{4\lambda_1^2} = 2 \Leftarrow \left(\frac{1}{2\lambda_1}\right)^2 + \left(\frac{-1}{2\lambda_1}\right)^2 = 2$

2 cases:

- $\lambda_1 = \frac{1}{2} \Rightarrow (x,y,z) = (1,-1,2) \Rightarrow f = 5$ MAX
- $\lambda_1 = -\frac{1}{2} \Rightarrow (x,y,z) = (-1,1,0) \Rightarrow f = 1$ MIN

//

Ex 4/ Prove that the arithmetic mean of n nonnegative numbers is always larger than their geometric mean.

To do so, we shall maximize $\sqrt[n]{x_1 \cdots x_n}$ subject to $x_1 + \cdots + x_n = c$ $\left(g(\vec{x}) = x_1 + \cdots + x_n - c\right)$.

Use $f(\vec{x}) = x_1 \cdots x_n$. Then $\vec{\nabla} f = \lambda \vec{\nabla} g \Rightarrow$

$$(x_2 \cdots x_n, x_1 x_3 \cdots x_n, \ldots, x_1 \cdots x_{n-1}) = \lambda (1, 1, \ldots, 1).$$

$$\left.\begin{array}{lcl} x_2 \cdots x_n = \lambda & \xrightarrow[\cdot x_1]{} & x_1 \cdots x_n = \lambda x_1 \\ x_1 x_3 \cdots x_n = \lambda & \xrightarrow[\cdot x_2]{} & x_1 \cdots x_n = \lambda x_2 \\ \vdots & \vdots & \vdots \\ x_1 \cdots x_{n-1} = \lambda & \xrightarrow[\cdot x_n]{} & x_1 \cdots x_n = \lambda x_n \end{array}\right\} \rightarrow \lambda x_1 = \cdots = \lambda x_n$$

If can't have $\lambda = 0$ if maximizing the product!

$\Rightarrow x_1 = \cdots = x_n$. Since $x_1 + \cdots + x_n = c$, we get $x_j = \frac{c}{n}$ for each j.

The maximum value of $x_1 \cdots x_n$ is therefore $\frac{c^n}{n^n}$

$\Rightarrow$ maximum value of $\sqrt[n]{x_1 \cdots x_n}$ is $\frac{c}{n}$

$$\Rightarrow \sqrt[n]{x_1 \cdots x_n} \underset{\text{always}}{\leq} \frac{c}{n} = \frac{x_1 + \cdots + x_n}{n}.$$

//

We finish by proving Theorem 1(i) from Lecture 32 in the special case where $S = [\vec{a}, \vec{b}] := [a_1, b_1] \times \cdots \times [a_n, b_n]$.

(The following 3 results are still true on an arbitrary closed & bounded subset of $\mathbb{R}^n$, but we won't prove that.)

For the following, let $f : S \to \mathbb{R}$ be a continuous function.

<u>Theorem A</u>: f is bounded.

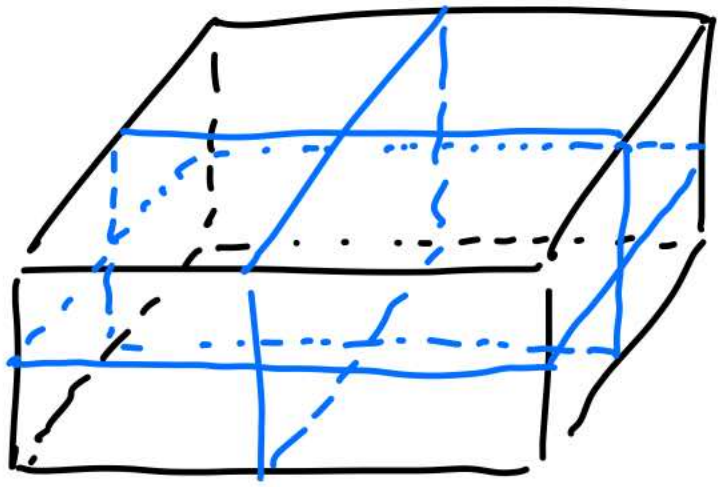
<u>Theorem B</u>: f attains a max & min.

<u>Theorem C</u>: f is uniformly continuous.

will skip this one: see "small span thm." in Apostol (proof very similar to A).

<u>Proof of A</u>: $\underbrace{\text{Assume otherwise.}}_{\text{to get a contradiction}}$ Chop S into half in each direction

S



$= \bigcup_{\alpha \in \mathcal{A}_1} S_\alpha^{(1)}$

Then into 4ths

$\vdots$

and so on.

$$= \bigcup_{\alpha \in \mathcal{A}_2} S_\alpha^{(2)}$$

$$\vdots$$

$$= \bigcup_{\alpha \in \mathcal{A}_m} S_\alpha^{(m)}$$

Side lengths $\dfrac{b_k - a_k}{2^m}$

There must be an $\alpha_m \in \mathcal{A}_m$ for each m on which $f|_{S_{\alpha_m}^{(m)}}$ is unbounded, and we can choose these so that

$$\underset{[\vec{a},\vec{b}]}{S} \supset \underset{[\vec{a}^{(1)},\vec{b}^{(1)}]}{S_{\alpha_1}^{(1)}} \supset \underset{[\vec{a}^{(2)},\vec{b}^{(2)}]}{S_{\alpha_2}^{(2)}} \supset \cdots \supset \underset{[\vec{a}^{(m)},\vec{b}^{(m)}]}{S_{\alpha_m}^{(m)}} \supset \cdots$$

Hence $a_k^{(m)}$ resp. $b_k^{(m)}$ are increasing resp. decreasing sequences with $a_k^{(m)} \le b_k^{(m)}$ $\forall m, k$. ($\Rightarrow$ $a_k^{(m)}$ resp. $b_k^{(m)}$ is bounded above resp. below.) Since

$$0 \le b_k^{(m)} - a_k^{(m)} \le \frac{b_k - a_k}{2^m} \xrightarrow[m \to \infty]{} 0$$

we have $\lim_{m\to\infty} a_k^{(m)} = \lim_{m\to\infty} b_k^{(m)} =: t_k$.

By continuity of f at $\vec{t}$,

$\exists\, \epsilon > 0$ s.t. $|f(\vec{x}) - f(\vec{t})| < 1$ for every $\vec{x} \in B(\vec{t};\epsilon) \cap S$

$\Rightarrow$ $|f(\vec{x})| < 1 + |f(\vec{t})|$ " " "

$\Rightarrow$ f bounded on $B(\vec{t};\epsilon)$ $\left(\supset S_{\alpha_m}^{(m)} \text{ for } m \gg 0\right)$

$\Rightarrow$ f bounded on $S_{\alpha_m}^{(m)}$ for $m \gg 0$ $\times$ $\square$

Proof of B : Let $M := \sup\{f(x) \mid x \in S\}$

(= least upper bound of f)

Then $g(\vec{x}) := M - f(\vec{x}) \geq 0$ on S.

Assume $g(\vec{x}) > 0$ on S, i.e. $f(\vec{x})$ never actually attains the value M on S.

we will get a contradiction below

Then $\frac{1}{g}$ is continuous on S

$$\underset{\text{Thm. A}}{\Longrightarrow} \exists C \text{ s.t. } \frac{1}{g(\vec{x})} \leq C \quad \forall \vec{x} \in S$$

$$\Longrightarrow M - f(\vec{x}) \ (= g) \geq \frac{1}{C} \quad \forall \vec{x} \in S$$

$$\Longrightarrow f(\vec{x}) \leq M - \frac{1}{C} \quad \forall \vec{x} \in S \qquad \times$$

So $\exists \vec{x}_0 \in S$ s.t. $g(\vec{x}_0) = 0$

$$\Rightarrow f(\vec{x}_0) = M \geq f(\vec{x}) \quad \forall \vec{x} \in S$$

$\Rightarrow f$ attains maximum at $\vec{x}_0$.

For min, apply the same argument to $-f$. $\square$