Session 11
March 11, 2020

Lagrange multipliers:

Problem: Find extrema of $f(\vec{x})$ under some constraint $G(\vec{x})=0$

Approach via differentials:

- for f to have extremum we need $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = 0$
- but dx, dy not independent, since they are such that $0 \overset{\text{G does not change along the constraint}}{=} dG = \frac{\partial G}{\partial x}dx + \frac{\partial G}{\partial y}dy$

$\Rightarrow 0 = d(f+\lambda g) = \left(\frac{\partial f}{\partial x} + \lambda\frac{\partial G}{\partial x}\right)dx + \left(\frac{\partial f}{\partial y} + \lambda\frac{\partial G}{\partial y}\right)dy$ for any $\lambda \in \mathbb{R}$

regarded as independent now, since constraint taken into account

$\Rightarrow$ need to solve $\frac{\partial f}{\partial x} + \lambda\frac{\partial G}{\partial x} = 0$ and $\frac{\partial f}{\partial y} + \lambda\frac{\partial G}{\partial y} = 0$

More geometric approach:

- consider f restricted to $S = \{\vec{x} \in \mathbb{R}^n : G(\vec{x}) = 0\}$ with an extremum at $\vec{a}$,

  and any curve $\vec{h}(t)$ in S such that $\vec{h}(0) = \vec{a}$.

  $\Rightarrow \varphi(t) = f(\vec{h}(t))$ has an extremum at $t=0$

  $\Rightarrow 0 = \varphi'(0) = (\vec{\nabla} f)(\underbrace{\vec{a}}_{\vec{h}(0)}) \cdot \vec{h}'(0)$

  $\Rightarrow (\vec{\nabla} f)(\vec{a})$ orthogonal to tangent vector of any curve in S

  but $(\vec{\nabla} G)(\vec{a})$ is also orthogonal to S

  $\Rightarrow (\vec{\nabla} f)(\vec{a}) = \lambda(\vec{\nabla} G)(\vec{a})$ for some $\lambda \in \mathbb{R}$ (both vectors are linearly dependent)

Conclusion (Lagrange's method):

As necessary condition to find extrema of $f(\vec{x})$ under constraint $G(\vec{x})=0$, we need to solve

$$\underbrace{\vec{\nabla}(f-\lambda G)=\vec{0}}_{n \text{ equations}} \quad \text{and} \quad \underbrace{G(\vec{x})=0}_{1 \text{ equation}}$$

, i.e., $n+1$ equations for $n+1$ variables $x_1,\dots,x_n,\lambda$.

λ is called $\boxed{\text{Lagrange multiplier}}$.

In short: if $\boxed{L(\vec{x},\lambda)=f(\vec{x})-\lambda G(\vec{x}) \text{, we need } \underbrace{\vec{\nabla}}_{=\begin{pmatrix}\partial_{x_1}\\ \vdots \\ \partial_{x_n} \\ \partial_\lambda\end{pmatrix}} L=0}$

Ex.: • rectangle: $f(x,y)=xy$, $G(x,y)=2x+2y-L=0$

$$\Rightarrow L(x,y,\lambda)=xy-\lambda(2x+2y-L)$$

$$\Rightarrow \vec{\nabla}L=\begin{pmatrix}\partial_x L\\ \partial_y L\\ \partial_\lambda L\end{pmatrix}=\begin{pmatrix}y-2\lambda\\ x-2\lambda\\ 2x+2y-L\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

$$\Rightarrow 0=2x+2y-L=2(2\lambda)+2(2\lambda)-L=0 \Rightarrow \underbrace{\lambda=\frac{L}{8}}_{\substack{\text{usually not interested}\\ \text{in the value of } \lambda}} \Rightarrow x=\frac{L}{4}, y=\frac{L}{4}$$

• $f(x,y)=x^2+y^2+y$, $G(x,y)=x^2+y^2-1=0$ (circle)

$$\Rightarrow L=x^2+y^2+y-\lambda(x^2+y^2-1) \Rightarrow \begin{pmatrix}\partial_x L\\ \partial_y L\\ \partial_\lambda L\end{pmatrix}=\begin{pmatrix}2x-2\lambda x\\ 2y+1-2\lambda y\\ x^2+y^2-1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

$\Rightarrow x=0$ or $\lambda=1$ $\Rightarrow$ only $x=0 \Rightarrow y=\pm\sqrt{x^2-1}=\pm 1$

$\hookrightarrow 2y+1-2y=1$ cannot be zero

$\Rightarrow (0,-1)$ and $(0,1)$ are critical points

$\downarrow$ $\downarrow$

$f(0,-1)=0$, $f(0,1)=2$

$\Rightarrow$ indeed a minimum $\Rightarrow$ indeed a maximum

Note: There are at least two advantages of using Lagrange's method:

- We can use it even if we cannot solve the constraint for one of the variables, or generally when substituting it into the function does not work

- Even if we could substitute the constraint into the function, and then find the critical points as usual, the computation with Lagrange's method is usually easier and faster, see examples in the homework/moodle/quiz exercises

2.5 Vector-valued Functions

We know what derivatives of $f:\mathbb{R}\to\mathbb{R}$ and $f:\mathbb{R}^n\to\mathbb{R}$ are.

What about $f:\mathbb{R}\to\mathbb{R}^n$?

$$\vec{f}(t)=\begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t)\end{pmatrix} \Rightarrow \frac{d\vec{f}}{dt} := \lim_{h\to 0} \frac{\vec{f}(t+h)-\vec{f}(t)}{h} = \lim_{h\to 0}\begin{pmatrix} \frac{1}{h}(f_1(t+h)-f(h)) \\ \vdots \\ \frac{1}{h}(f_n(t+h)-f(h))\end{pmatrix},$$

so $\dfrac{d\vec{f}(t)}{dt} = \begin{pmatrix} \frac{df_1(t)}{dt} \\ \vdots \\ \frac{df_n(t)}{dt}\end{pmatrix}$ i.e., derivative is taken component-wise

$\Rightarrow$ the usual product rules hold: for $\vec{f},\vec{g}:\mathbb{R}\to\mathbb{R}^n$, $\phi:\mathbb{R}\to\mathbb{R}$, then

- $\dfrac{d}{dt}(\phi\vec{f}) = \dfrac{d\phi}{dt}\vec{f} + \phi\dfrac{d\vec{f}}{dt}$
- $\dfrac{d}{dt}(\vec{f}\cdot\vec{g}) = \dfrac{d\vec{f}}{dt}\cdot\vec{g} + \vec{f}\cdot\dfrac{d\vec{g}}{dt}$
- $\dfrac{d}{dt}(\vec{f}\times\vec{g}) = \dfrac{d\vec{f}}{dt}\times\vec{g} + \vec{f}\times\dfrac{d\vec{g}}{dt}$

recall: cross product $\vec{f}\times\vec{g} = \begin{pmatrix} f_2g_3-f_3g_2 \\ -f_1g_3+f_3g_1 \\ f_1g_2-f_2g_1\end{pmatrix}$