Elements of Calculus

Prof. Sören Petrat, Constructor University

Lecture notes from Spring 2025

Session 23
(Week 12A)

6. Multivariable Calculus

6.2 Optimization in Many Variables

Topic for Week 12 A: Lagrange Multipliers

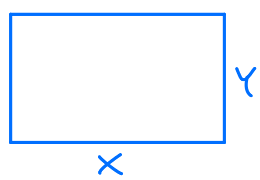
We continue our discussion on optimizing functions of several variables.

Problem for today: Optimize $f(x,y)$, but under the constraint that $G(x,y)=0$.

Now it does not help to simply find the max./min. of f (using $\nabla f = 0$) because these might not satisfy the constraint $G(x,y)=0$.

Let us start with a simple example: the "isoperimetric problem".

Consider a rectangle with side lengths x and y:



Question: Maximize the area $f(x,y)=xy$ given that the perimeter $2x+2y$ is equal L, i.e., under the constraint $G(x,y)=2x+2y-L=0$.

Here we know how to solve this: From $G(x,y)=0$ we know $y=\frac{1}{2}(L-2x)$

$\Rightarrow f(x,y(x)) = x\,\frac{1}{2}(L-2x) = \frac{L}{2}x - x^2$

Extrema: $0 = \frac{d}{dx} f(x,y(x)) = \frac{L}{2} - 2x \Rightarrow x = \frac{L}{4} \Rightarrow y = \frac{1}{2}\left(L - 2\left(\frac{L}{4}\right)\right) = \frac{L}{4}$ (max. since $\frac{d^2 f}{dx^2} < 0$).

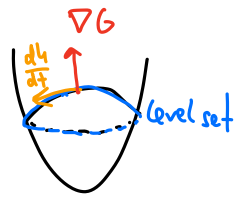
$\Rightarrow$ The area is maximal for a square!

However, often it is impractical or impossible to explicitly solve $G(x,y)=0$ for x (or y), e.g., $G(x,y)=e^{\sin(xy)}-\cos(x^2y^2)=0$.

Hence we want to find a better method.

Let us first consider the constraint more carefully. We assume $G:\mathbb{R}^n\to\mathbb{R}$ is once continuously differentiable. Now consider the $\boxed{\text{level sets } S_c=\{x\in\mathbb{R}^n : G(x)=c\},}$ for any $c\in\mathbb{R}$ (all x's where G has the constant value c; geometrically for $G:\mathbb{R}^2\to\mathbb{R}$ this means all x where $G(x)$ has the same height w.r.t. the x-y-plane, see also the picture at the end of the notes).

Now let $h(t)$ be any curve in S_c ($h(t)\in S_c \ \forall t\in\mathbb{R}$).



Then $G(h(t))=c \ \forall t\in\mathbb{R}$ by definition (h is a curve in S_c).

$$\Rightarrow 0=\frac{d}{dt}G(h(t))\overset{\text{chain rule}}{=}\nabla G(h(t))\cdot\frac{dh}{dt}$$

Hence we have proven that $\nabla G(h(t))\cdot\frac{dh}{dt}=0$, i.e., $\boxed{\nabla G \text{ is orthogonal to } S_c}$ (at any point).

Example: see geogebra picture on the last page

↳ $G(x,y)=x^2+y^3+2$, $c=3$; i.e., the level set is the intersection of G with the $z=3$-plane

We find $\nabla G=\begin{pmatrix}2x\\3y^2\end{pmatrix}$

↳ E.g., the points $\overbrace{(1,0),\ (0,1),\text{ and } \left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt[3]{2}}\right) \text{ are in the level set}}^{G=3 \text{ at all three points}}$,

there, we have $\nabla G(1,0)=\begin{pmatrix}2\\0\end{pmatrix}$, $\nabla G(0,1)=\begin{pmatrix}0\\3\end{pmatrix}$, $\nabla G\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt[3]{2}}\right)=\begin{pmatrix}2/\sqrt{2}\\3/\sqrt[3]{4}\end{pmatrix}$

In the picture one can clearly see how ∇G is orthogonal to the level set $G=3$ at these points.

Let's apply this to our problem of optimizing $f: \mathbb{R}^n \to \mathbb{R}$ given $G(x)=0$.

Suppose f restricted to $S=\{x \in \mathbb{R}^n : G(x)=0\}$ has an extremum at $a \in \mathbb{R}^n$.
Consider any curve $h(t)$ in S such that $h(0)=a$.

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

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

$\Rightarrow$ $(\nabla f)(a)$ is orthogonal to tangent vector of any curve in S, i.e., $(\nabla f)(a)$ is orthogonal to S.

But from above we also know that $(\nabla G)(a)$ is orthogonal to S!

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

<u>Conclusion (Lagrange's method):</u>

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

$\underbrace{\nabla(f-\lambda G)=0}_{n \text{ equations}}$ and $\underbrace{G(x)=0}_{1 \text{ equation}}$, i.e., $n+1$ equations for $n+1$ variables $x_1,\dots,x_n,\lambda$.

We call λ the **Lagrange multiplier**.

In short: if $\boxed{\underbrace{L(x,\lambda)}_{\text{called "Lagrangian function"}}=f(x)-\lambda G(x), \text{ we need } \nabla L=0.}$

Here, $\nabla=\begin{pmatrix}\partial_{x_1}\\ \vdots \\ \partial_{x_n} \\ \partial_\lambda\end{pmatrix}$

Note: $\nabla L = 0$ is just a necessary condition, one still needs to check that the points are indeed minima or maxima.

Examples:

- The isoperimetric problem from above: $f(x,y) = xy$, $G(x,y) = 2x+2y-L = 0$

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

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

$$\Rightarrow 0 = 2x+2y-L = 2(2\lambda)+2(2\lambda)-L = 0 \Rightarrow \underbrace{\lambda = \frac{L}{8}} \Rightarrow x = \frac{L}{4}, y = \frac{L}{4}.$$

Note: usually we are not interested in the value of λ

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

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

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

$\hookrightarrow$ $2y+1-2y = 1$ and not 0 (as we would need from $\partial_y L = 0$)

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

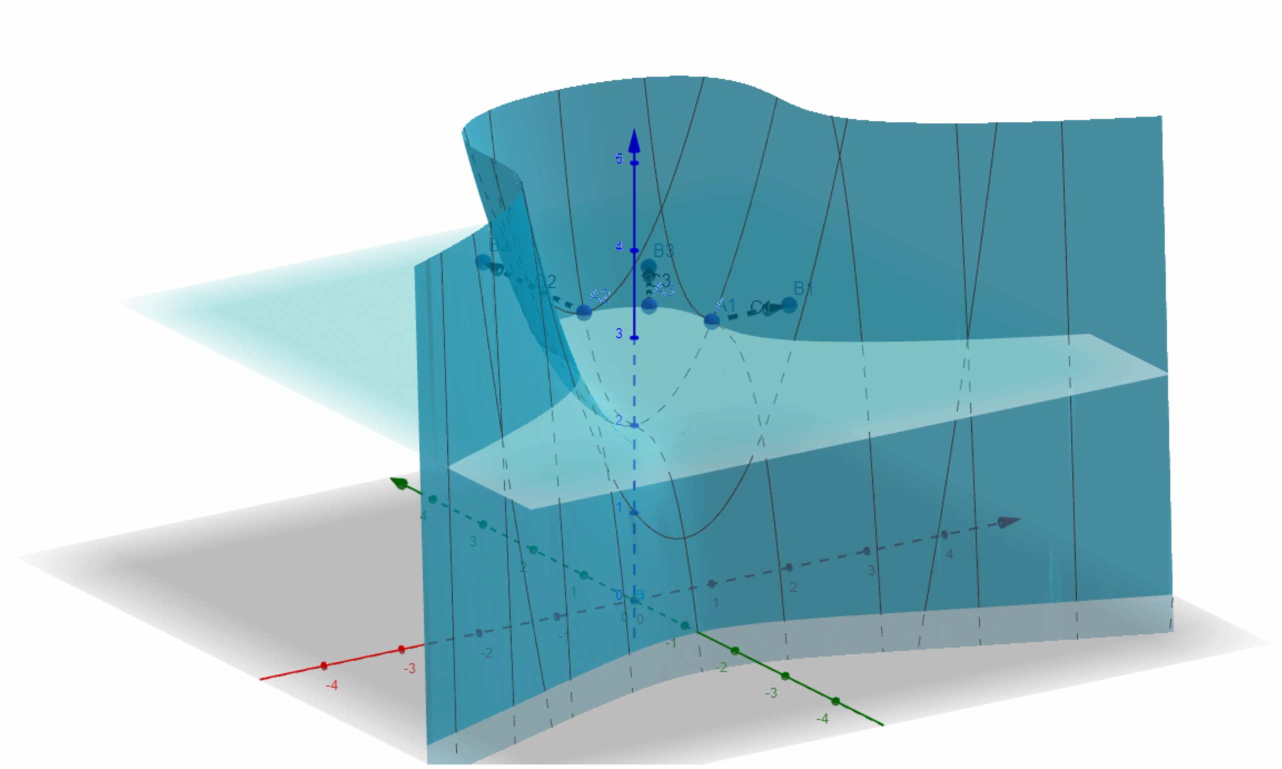
$\downarrow$ $\downarrow$

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

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

Use https://www.geogebra.org/3d for generating the plots.

$G(x,y) = x^2 + y^3 + 2$, level set $G(x,y) = 3$, with three gradient vectors:



View from above to see orthogonality of nabla G and the level set: