

Wednesday, Sep 24, Lecture 12: Optimization and Lagrange multipliers

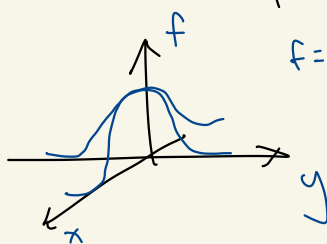
Last time... recall vector calc.

Optimization

Ex: in machine learning, need to minimize an "objective function" $f: \mathbb{R}^n \rightarrow \mathbb{R}$, n is huge.

$f: \mathbb{R}^2 \rightarrow \mathbb{R}$ tells a story

Max:

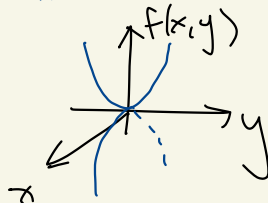


$$f = -x^2 - y^2$$

In all cases,

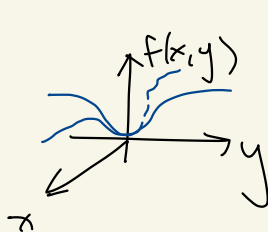
$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = 0.$$

Saddle:



$$f = -x^2 + y^2$$

Min

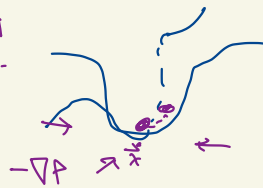


$$f = x^2 + y^2$$

Notes: the "second derivative test" for $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ is much nastier than $f: \mathbb{R} \rightarrow \mathbb{R}$. Often, a better way is with "stability":



check $-\nabla f$ near $\vec{x}^*$!



Ex:

Make a tent of fixed volume V with minimal fabric of following shape:



Soln:

$$V = \frac{1}{2} \cdot 2w \cdot l \cdot w \tan \theta = w^2 l \tan \theta$$

$$A = 2w^2 \tan \theta + \frac{2lw}{\cos \theta}$$

- Minimize A subject to fixed V . At first, $A = A(l, w, \theta)$ but since V is fixed,

(note: $\theta \in (0^\circ, 90^\circ)$, otherwise degenerate case:

$\theta = 0^\circ$ | $\theta = 90^\circ$)

$$l = \frac{V}{w^2 \tan \theta}$$

$$\Rightarrow A = A(w, \theta)$$

$$= 2w^2 \tan \theta + \frac{2w}{\cos \theta} \frac{V}{w^2 \tan \theta}$$

$$= 2w^2 \tan \theta + \frac{2V}{w} \csc \theta$$

set:

$$\nabla A = 0 \quad \text{so}$$

$$0 = \frac{\partial A}{\partial w} = 4w \tan \theta - \frac{2V \csc \theta}{w^2} \quad (*)$$

$$0 = \frac{\partial A}{\partial \theta} = 2w^2 \sec \theta - \frac{2V}{w} \csc \theta \cot \theta$$

$$\Rightarrow \frac{V \csc \theta}{2 \tan \theta} = w^3 = \frac{V \csc \theta \cot \theta}{\sec \theta}$$

$$\Rightarrow \cos^2 \theta = \frac{1}{2}, \text{ so } \theta = 45^\circ.$$

$$\Rightarrow \tan \theta = 1, \text{ so } V = w^2 l \text{ and } (*) \text{ gives } 2w = l\sqrt{2}.$$

To compute these exactly, need to fix V .

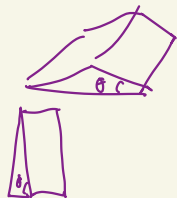
How do we know this is a min?

- Option 1: plug in other values for fixed value V ,
see that $A \geq A_{critical}$

- Option 2: This is the only extremum,
and as we approach extreme cases,

$\theta \rightarrow 0$, Big area, small volume.

$\theta \rightarrow 90^\circ$, Big area, small volume



This must be a happy medium.

Lagrange Multipliers

Another strategy for optimization with constraints.

Ex: find the point $\vec{p} = (x, y, z)$ on the sphere

$S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ which maximizes $f(x, y, z) = z$.

Soln:

pretty obvious that $\vec{p} = (0, 0, 1)$ does it, but let's see how this works.

- Blindly computing $\nabla f = (0, 0, 1)$ and setting $\nabla f = 0$
doesn't work - after all, no real vector $\vec{p} \in \mathbb{R}^3$ maximizes f !

- But on the sphere is a different story.

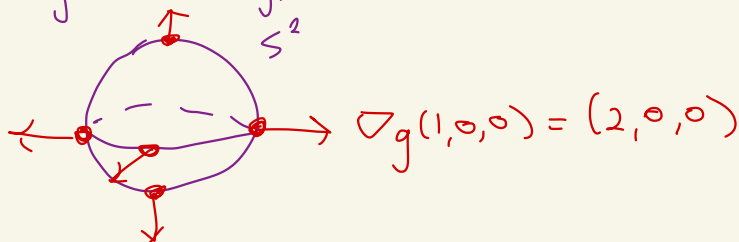
- S^2 is given by a constraint equation $g: \mathbb{R}^3 \rightarrow \mathbb{R}$

$$g(x, y, z) = x^2 + y^2 + z^2 - 1$$

(Notice that S^2 is the level set of $g(x, y, z) = 0$).

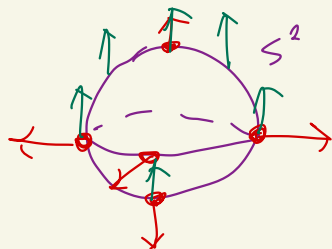
Notice that ∇g is perpendicular to its level sets.

$$\nabla g = (2x, 2y, 2z)$$



• We may increase f by moving in direction of gradient.

$$\nabla f = (0,0,1)$$



• Notice: when ∇f and ∇g point in the same direction ($\nabla f = \lambda \nabla g$), we have a local extremum.

This observation gives the method of Lagrange multipliers:
Find a nonzero λ st.

$$\nabla f - \lambda \nabla g = 0 \quad \text{at some point } (x_0, y_0, z_0).$$

$$\begin{aligned} \text{so } \nabla f &= (0, 0, 1) \\ \nabla g &= (2x, 2y, 2z) \end{aligned}$$

$$\Rightarrow \nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$$

forces $x_0 = y_0 = 0$. This leaves

$$1 = \lambda 2z_0 \Rightarrow \frac{1}{2\lambda} = z_0.$$

But $(x_0, y_0, z_0) \in S^2$, so

$$1 = 0 + 0 + z_0^2.$$

Thus, either $z_0 = -1$ or $z_0 = 1$. check picture,

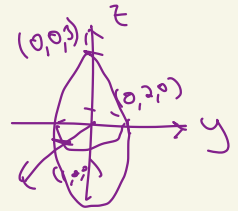
$z_0 = 1$ is max.

$\Rightarrow (0, 0, 1)$ is max.

Ex

Find the volume of largest rectangular box inscribed in the "egg"

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$



Soln:



$$\Rightarrow \text{Volume } f(x, y, z) = 2x \cdot 2y \cdot 2z = 8xyz.$$

Lagrange: call constraint $g(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$

$\Rightarrow$ Find solns to

$$\nabla f = \lambda \nabla g$$

$$\nabla f = (8yz, 8xz, 8xy)$$

$$\nabla g = \left(\frac{2x}{a^2}, \frac{2y}{b^2}, \frac{2z}{c^2} \right)$$

$$8yz = \frac{2x}{a^2}$$

$$8xz = \frac{2y}{b^2}$$

$$8xy = \frac{2z}{c^2}$$

Degenerate cases: if $x=0$ (or any other $x, y, z=0$),
Volume = 0. Suffices to assume all nonzero.

Multiply them all by the missing variable and add them:

$$3 \cdot xyz = 2\lambda \underbrace{\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right)}_{=1}$$

$$\Rightarrow \lambda = \frac{3}{2}xyz.$$

So, plugging into $\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}$

$$3yz = \frac{3}{2}xyz \cdot \frac{(2x)}{a^2}$$

and dividing

$$x^2 = \frac{1}{3}a^2$$

Similarly,

$$y^2 = \frac{1}{3}b^2$$

$$z^2 = \frac{1}{3}c^2$$

$$\Rightarrow (x, y, z) = \frac{1}{\sqrt{3}}(a, b, c)!$$

Physics note:

We are working with a Lagrangian

$$\mathcal{L}(x, \lambda) = f(x) - \lambda g(x) \text{ and checking } \nabla_x \mathcal{L}(x, \lambda) = 0.$$

$$[\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow g(x, y) = 0$$