Supplement: Lagrange Multipliers

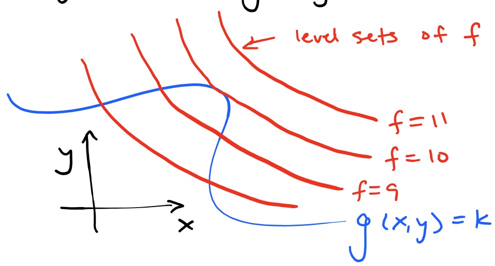
Monday, February 15, 2021 1:55 PM

We'll do a general / powerful duality theory which includes Lagrange Multipliers

Here, let's connect w/ what you saw in calc (Ref: §11.8 Stewart's "Essential Calculus")

To solve:

$$\min_{x,y \in \mathbb{R}} f(x,y) \quad \text{s.t.} \quad g(x,y) = k$$



idea is that the optimal pt (x^*, y^*) will be where f, g have same tangent line,

$$\nabla f(x^*, y^*) = \lambda \nabla g(x^*, y^*)$$

No convexity needed

Thm $f, g \in C^1(\mathbb{R}^2)$, then if a min or max of f (subject to $g(x,y)=0$) occurs at a point $\vec{p} = (x^*, y^*)$ where $\nabla g(\vec{p}) \neq 0$, then $\exists\, \lambda \in \mathbb{R}$ s.t.

$$\nabla f(\vec{p}) = \lambda \nabla g(\vec{p})$$

Lagrange multiplier

NOTATION: Note: we'll use "ν" for such, saving "λ" for inequality constraints

proof

Since $\nabla g(\vec{p}) \neq 0$, use implicit function theorem to represent $g(x,y)=0$ locally by a parametric curve $r(t)$ w/ tangent vector r'. So $\exists\, t_0$ s.t. $r(t_0) = \vec{p}$ since $\vec{p}$ is feasible, and wlog let $t_0 = 0$ so $r(0) = \vec{p}$, and f obtains its min/max at $\vec{p}$

Then, locally, we have unconstrained problem, $f(r(t))$, and it obtains its min/max at $t=0$.

For unconstrained problem, at min/max, derivative must vanish.

$$\text{So} \quad \underbrace{\frac{d}{dt} f(r(t)) \Big|_{t=0}}_{= \nabla f(\vec{p}) \cdot \vec{r}\,'(0)} = 0 \qquad r(0) = \vec{p}$$

Also, $g(r(t)) = 0$ $\forall t$ by design

i.e., it's constant,

i.e., $\underbrace{\frac{d}{dt} g(r(t))}$ $= 0$ $(\forall t)$

at $t=0$: $\nabla g(\vec{p}) \cdot \vec{r}\,'(0) = 0$ $\quad r(0) = \vec{p}$

i.e.

$$\nabla f(\vec{p}) \perp \vec{r}\,'(0)$$

and $\nabla g(\vec{p}) \perp \underbrace{\vec{r}\,'(0)}_{\neq 0 \text{ by assumption}}$

in $\mathbb{R}^2$, the set of vectors orthogonal to $\vec{r}\,'(0)\ (\neq \vec{0})$ is one-dimensional

i.e., $\nabla f(\vec{p})$, $\nabla g(\vec{p})$ in same 1D subspace (linearly dependent)

i.e., $\exists\ \lambda$ s.t. $\nabla f(\vec{p}) = \lambda \nabla g(\vec{p})$ (specifically, assumed $\nabla g(\vec{p}) \neq 0$, but $\nabla f(\vec{p}) = 0$ is ok w/ $\lambda = 0$)