Optimization

Thomas Cohn

10/12/18

Situation 2: Constraints
Given $f \in C^1(\Omega^{\text{osso}\mathbb{R}^{k+n}}, \mathbb{R}^n)$, $\vec{p} \in E = f^{-1}(\vec{0})$, $h \in C^1(\Omega, \mathbb{R})$, $h|_E$ has a local min/max at $\vec{p}$.
Consider $\gamma \in C^1(\text{osso}\mathbb{R}, E)$. From Wednesday, $0 = Dh(\vec{p}) \cdot \gamma'(0)$. What do we know about $\gamma'(0)$?
Note that if we define $f \circ \gamma = 0$, $Df(\gamma(t)) \cdot \gamma'(t) = 0$. When $t = 0$, $Df(\vec{p}) \cdot \gamma'(0) = 0$.
Thus, $\gamma'(0) \in \ker Df(\vec{p})$.

Lemma: If $Df(\vec{p})$ has maximal rank n, then there are no other constarints on $\gamma'(0)$.
Proof: Homework 6 problem 1.

Altogether, we have $Dh(\vec{p}) \in (\ker Df(\vec{p}))^\perp = ((\text{row space } Df(\vec{p}))^\perp)^\perp = \text{row space } Df(\vec{p})$.
This is Span $\{Df_1(\vec{p}), Df_2(\vec{p}), \ldots, Df_n(\vec{p})\}$. I.e. $Dh(\vec{p}) \sum_{i=1}^n \lambda_i Df_i(\vec{p})$. The λ_i's are called Lagrange multipliers.

So we have the equations $\begin{cases} f(\vec{p}) = \vec{0} \\ Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p}) \end{cases}$

This gives us $k + 2n$ unknowns: $\vec{p} = (p_1, \ldots, p_{k+n})$ and $\lambda_1, \ldots, \lambda_n$.
$f(\vec{p}) = \vec{0}$ gives us n "scalar equations".
$Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p})$ gives us $n + k$ "scalar equations".

Global aspects: $K^{\text{cpt}} \subset \mathbb{R}^m$, $h : K \to \mathbb{R}$, the extreme value theorem implies that h has a global max and min on K. Points we need to check:

1. $\vec{p} \in \text{Int}\,(K)$ if $Dh(\vec{p}) = \vec{0}$.
2. $\vec{p} \in \text{Int}\,(K)$ if h is not differentiable at $\vec{p}$.
3. $\vec{p} \in \text{Bd}\,(K)$.

Ex: Maximize and minimize $h(x, y) = x^4 + y^6$ on $K = \big\{(x, y) : x^2 + y^2 \leq 1\big\}$.

$$Dh\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 4x^3 & 6x^5 \end{bmatrix}$$

$$Df\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 2x & 2y \end{bmatrix}$$

The minimum occurs at $(0, 0)$ with $h(0, 0) = 0$. The maximum occurs on the boundary of K.
$\text{Bd}\,(K) = E = \big\{(x, y) : x^2 + y^2 = 1\big\}$. So we have the system of equations

$$\begin{cases} x^2 + y^2 = 1 \\ 4x^3 = \lambda \cdot 2x \\ 6y^5 = \lambda \cdot 2y \end{cases}$$

$x = 0 \to y = \pm 1 \to h = 1$.
$y = 0 \to x = \pm 1 \to h = 1$.

$$x, y \neq 0 \to \begin{cases} x^2 + y^2 = 1 \\ x^2 + \frac{1}{2}\lambda \\ y^4 = \frac{1}{3}\lambda \end{cases} \to \frac{\lambda}{2} + \frac{\sqrt{\lambda}}{\sqrt{3}} - 1 = 0 \to \lambda = \frac{2}{3}(4 - \sqrt{7}) \;\ldots\; \to h = 0.368$$

A variant: Replace $x^2 + y^2 \leq 1$ by $x^8 + y^8 \leq 1$. Then you get a "non-trivial maximum".