Lagrange Multiplier Theorem

Thomas Cohn

10/17/18

Given $f \in C^1(\Omega^{\text{osso}\mathbb{R}^{k+n}}, \mathbb{R})$, $\vec{p} \in E = f^{-1}(\vec{0})$, rank $Df(\vec{p}) = n$, $h \in C^1(\Omega, \mathbb{R})$, and $h|_E$ has a local max or min at $\vec{p}$.
Then $\exists \lambda_1, \ldots, \lambda_n$ s.t. $Dh(\vec{p}) = \lambda_1 Df_1(\vec{p}) + \cdots + \lambda_n Df_n(\vec{p})$. λ_j are called the lagrange multipliers.

Ex: What points of $xyz = 1$ lie closest to $\vec{0}$?
Let $f(x,y,z) = xyz - 1$ and $h(x,y,z) = x^2 + y^2 + z^2$. Minimize h over $f^{-1}(\vec{0}) = E$.

Is the existance of the minimum guaranteed? In this case, yes. Pick R s.t. $R > h(x_0, y_0, z_0)$ for some $(x_0, y_0, z_0) \in E \neq \emptyset$. Let $K = \left\{(x,y,z) : x^2 + y^2 + z^2 \leq R\right\}$. Then $\inf K = \inf E \cap K$, and $E \cap K$ is compact. By the extreme value theorem, $\inf E \cap K = \min E \cap K$.

$Dh = \begin{bmatrix} 2x & 2y & 2z \end{bmatrix}$ and $Df = \begin{bmatrix} yz & xz & xy \end{bmatrix}$. So we have the following system of equations:

$$\begin{cases} 2x = \lambda yz \\ 2y = \lambda xz \\ 2z = \lambda xy \\ xyz = 1 \end{cases}$$

Solving this system of equations gives us $(1,1,1)$; $(-1,-1,1)$; $(-1,1,-1)$, and $(1,-1,-1)$.

For extra practice, try $x^a + y^b + c^z = 1$

Ex: $B \in \text{Mat}\,(n,n,\mathbb{R})$ symmetric (that is, $B = B^T$).
Let $h(\vec{x}) = \vec{x}^T B \vec{x}$. Goal: maximize h on $||\vec{x}||^2 = 1$. Use $f(\vec{x}) = ||\vec{x}||^2 - 1$.

Check that Df has rank 1 when $||\vec{x}||^2 = 1$. $Df(\vec{x}) = 2\vec{x}^T$. Then the max exists, and it occurs at a solution of $Dh = \lambda Df$.

Claim: $Dh(\vec{x}) = 2\vec{x}^T B$.
Proof: $h(\vec{x}) = \sum_{j,k} b_{jk} x_j x_k$. So $D_m h(\vec{x}) = \sum_k b_{mk} x_k + \sum_j b_{jm} x_j$.
By symmetry, $D_m h(\vec{x}) = 2 \sum_j b_{jm} x_j = (2\vec{x}^T B)_m$. So $Dh(\vec{x}) = 2\vec{x}^T B$
Proof 2: $Dh(\vec{x}) \cdot \vec{u} = h'(\vec{x}; \vec{u}) = \vec{u}^T B \vec{x} + \vec{x}^T B \vec{u} = 2\vec{x}^T B \vec{u}$.

We need $Dh(\vec{x}) = \lambda Df(\vec{x})$. $Dh(\vec{x}) = 2\vec{x}^T B$ and $Df(\vec{x}) = 2\lambda \vec{x}^T$. So we have $B\vec{x} = \lambda \vec{x}$. So λ is an eigenvalue and $\vec{x}$ is an eigenvector (call it $\vec{x_1}$).

Note that $h(\vec{x}) = \vec{x}^T B \vec{x} = \lambda$, i.e., $\lambda = \max h$ over the sphere. Rename μ_1 as the eigenvalue.

We have previously proved that every symmetric matrix has a real eigenvalue.

Ex: Followup: Now maximize h over the sphere intersected with $\{\vec{x_1}\}^T$, which is just $f^{-1}(\vec{0})$, with
$f(\vec{x}) = ||\vec{x}||^2 - 1 = \begin{bmatrix} \vec{x}^T \cdot \vec{x} - 1 \\ \vec{x_1}^T \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{bmatrix}$

We need $Dh(\vec{x}) = \lambda_1 Df_1(\vec{x}) + \lambda_2 Df_2(\vec{x})$, i.e., we need

$$\begin{cases} \vec{x}^T \cdot \vec{x} = 1 \\ \vec{x_1}^T \cdot \vec{x} = 0 \\ 2\vec{x}^T B = 2\lambda_1 \vec{x}^T + \lambda_2 \vec{x_1}^T \end{cases} \quad \to \text{(right-multiply by } \vec{x_1}) \to 2\vec{x}^T B \vec{x_1} = 0 + \lambda_2$$

So $\lambda_2 = 0$, so $2\vec{x}^T B = 2\lambda_1 \vec{x}^T$, so $B\vec{x} = \lambda_1 \vec{x}$. We get a second real eigenvalue $\mu_2 = \lambda_1$, with eigenvector $\vec{x_2} \in \{\vec{x_1}\}^\perp$.

Ex: Use induction to prove the spectral theorem:
B symmetric real matrix $\to$ B admits an orthonormal basis of eigenvectors $\vec{x_1}, \ldots, \vec{x_n}$ with real eignevalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$.

Ex: $h(c_1\vec{x_1} + \cdots + c_n\vec{x_n}) = c_1^2\mu_1 + \cdots + c_n^2\mu_n$.
All $\mu_i \geq 0 \Leftrightarrow \vec{x}^T B\vec{x} \geq 0$ for all $\vec{x} \in \mathbb{R}^n \setminus \left\{\vec{0}\right\} \overset{\text{def}}{\Leftrightarrow}$ "$B \geq 0$". We say that B is positive semi-definite.
All $\mu_i > 0 \Leftrightarrow \vec{x}^T B\vec{x} > 0$ for all $\vec{x} \in \mathbb{R}^n \setminus \left\{\vec{0}\right\} \overset{\text{def}}{\Leftrightarrow}$ "$B > 0$". We say that B is positive definite.
$B \leq 0 \Leftrightarrow (-B) \geq 0$
$B < 0 \Leftrightarrow (-B) > 0$.

Thm: Given $\Omega \subset \mathbb{R}^n$ convex and open, $f \in C^2(\Omega, \mathbb{R})$.
$Hf(\vec{x}) \overset{\text{def}}{=} (D_j D_k f(\vec{x}))_{j;k}$. This is called the Hessian of f at $\vec{x}$.
$Hf(\vec{x}) \in \text{Symm}(n) \overset{\text{def}}{=} \left\{M \in \text{Mat}\,(n,n) : M^T = M\right\}$
$Hf(\vec{x}) \geq 0$, $\forall \vec{x} \in \Omega$, $Df(\vec{x_0}) = \vec{0}$.

Then $f(\vec{x}) \geq f(\vec{x_0})$, $\forall \vec{x} \in \Omega$.