NOTES ON EXTREMA AND CRITICAL POINTS

JAKE MARCINEK
HARVARD UNIVERSITY
CAMBRIDGE, MA
MARCINEK@MATH.HARVARD.EDU

1. Notions

Let $f(x, y)$ be a function of 2 variables. We are now primarily interested in finding *local / global extrema* of f.

Definition 1.1. The point (a, b) is a

- *local minimum* if $f(a, b) \leq f(x, y)$ for all (x, y) near (a, b).
- *global minimum* if $f(a, b) \leq f(x, y)$ for all (x, y).
- *local minimum* if $f(a, b) \geq f(x, y)$ for all (x, y) near (a, b).
- *global minimum* if $f(a, b) \geq f(x, y)$ for all (x, y).

Note 1.2. The only parts of the definition that change are

(1) the direction of the inequality depending on max/min and
(2) the quantifier "near (a, b)" on the points (x, y) depending on local/global.

Strategy 1.3. To find them, we follow the same strategy as in single variable calculus:

(1) Local extrema occur when $f'(x) = 0$ (now $\nabla f = 0$) and are classified by the sign of $f''(x)$ (now $\nabla\nabla f$).
(2) Global extrema are tested case by case. The possible candidates for global extrema are the local extrema and the boundary. We will talk about checking the boundary next time using Lagrange multipliers.

Definition 1.4. The point (a, b) is a critical point of f if $\nabla f(a, b) = 0$.

Theorem 1.5 (First derivative test)**.** *Local extrema are critical points.*

Proof. ∇f points in the direction of greatest increase. At a local extremum, there is no such direction so $\nabla f = 0$. $\square$

Definition 1.6. The *Hessian* of f is $H(a, b) = \nabla\nabla f(a, b) = \begin{pmatrix} f_{xx} & f_{yx} \\ f_{xy} & f_{yy} \end{pmatrix}$. We set $D = \det H = f_{xx} f_{xy} - f_{xy}^2$ to be the determinant of the Hessian.

Theorem 1.7 (Second derivative test)**.** *Suppose (a, b) is a critical point.*

(1) *If $D(a, b) < 0$, then (a, b) is a saddle point.*
(2) *If $D(a, b) > 0$ and $f_{xx} < 0$, then (a, b) is a maximum.*
(3) *If $D(a, b) > 0$ and $f_{yy} > 0$, then (a, b) is a minimum.*
(4) *If $D(a, b) = 0$, then the test is inconclusive.*

Date: October 18, 2016.

2. Proof of second derivative test

I did not want to discuss any linear algebra in class since not everyone is familiar. However, the logic behind the second derivative is very mysterious without using linear algebra, so I want to give some intuition on it to those of you with the sufficient background.

Basically, we want to generalize the 1-variable case:

(1) Concave up in every direction means minimum.
(2) Concave down in every direction means maximum.
(3) Concave up in some directions and concave down in other directions means saddle.

Just as in one variable, concavity in the direction $\vec{v}$ is given by the second derivative in the direction $\vec{v}$, i.e.

$$D_{\vec{v}}(D_{\vec{v}}f) = \vec{v} \cdot \nabla(\nabla f \cdot \vec{v}) = \vec{v} \cdot H \cdot \vec{v}.$$

From linear algebra, we know that there are two eigenvalues $m \leq M$ with corresponding eigenvectors $\vec{v}_m$ and $\vec{v}_M$ corresponding to the directions that have concavity m and M, respectively. In any other direction, $\vec{w}$,

$$m = D_{\vec{v}_m} D_{\vec{v}_m} f \leq D_{\vec{w}} D_{\vec{w}} f \leq D_{\vec{v}_M} D_{\vec{v}_M} f = M$$

or as matrices

$$m = \vec{v}_m \cdot H \cdot \vec{v}_m \leq \vec{w} \cdot H \cdot \vec{w} \leq \vec{v}_M \cdot H \cdot \vec{v}_M = M.$$

Furthermore, the determinant of a matrix is always equal to the product of its eigenvalues so $D = mM$.

Corresponding to the list above, we break into cases:

(1) minima have $m, M < 0$, so $D > 0$.
(2) maxima have $m, M > 0$, so $D > 0$.
(3) saddles have $m < 0 < M$, so $D < 0$.

To distinguish between the first two cases, we just have to check the second derivative in any direction (f_{xx} works, for example).

3. Lagrange multipliers: Relative extremal problems

Now we want to solve extrema problems subject to a constraint. In other words, maximize/minimize f on a level set of g.

Example 3.1 (Economics)**.** You are running a business and are trying to decide how to split your monthly spending between costs for: $e =$ employees, $s =$ supplies, and $a =$ advertisements. You want to spend \$1000 this month and maximize your profit $\pi(e, s, a)$ which is a function of the three. This is a Lagrange multipliers problem: Maximize pi subject to the constraint $e + s + a = 1000$. Lagrange multipliers will tell us to solve $\nabla \pi(e, s, a) =< 1, 1, 1 >$.

Let f and g_i be differentiable functions on $\mathbb{R}^n$, and let $S = \cap_i \{g_i = 0\}$ be a level set of a collection of functions.

Problem 3.2. Maximize f on S.

Definition 3.3. The point $P \in S$ is a

- *local minimum of* $f|_S$ if $f(P) \leq f(Q)$ for all Q near P.
- *global minimum of* $f|_S$ if $f(P) \leq f(Q)$ for all Q.
- *local maximum of* $f|_S$ if $f(P) \leq f(Q)$ for all Q near P.
- *global maximum of* $f|_S$ if $f(P) \leq f(Q)$ for all Q.

Definition 3.4. The gradient of f splits into two components, $\nabla f = \nabla^T f + \nabla^N f$, the component of the gradient tangent to S and the component normal to S.

Definition 3.5. A point $P \in S$ is a *critical point of* $f|_S$ if $\nabla^T f = 0$.

Theorem 3.6 (Lagrange Multipliers)**.** *All local extrema of* $f|_S$ *are critical points.*

Proof. While ∇f points in the direction of greatest increase by the gradient theorem, we may not be able to move in that direction while remaining on S. The projection onto the tangent space, $\nabla^T f$, points in the direction of greatest increase which we are allowed to move in while remaining on S. At a local extremum, there is no such direction so $\nabla^T f = 0$. □

Remark 3.7. By the gradient theorem, the normal direction is spanned by ∇g_i, so $\nabla^N f = \sum_i \lambda_i \nabla g_i$ for some scalars λ_i. Therefore, a restatement of the Lagrange multiplier theorem is: at a local extremum of $f|_S$,

$$\nabla f = \nabla^N f = \sum \lambda \nabla g_i. \qquad (1)$$