Session 10
Mar 9, 2020

Recall: If $f(\vec{x})$ has a max or min at $\vec{a}$, then $(\vec{\nabla} f)(\vec{a}) = \vec{0}$

Generally: if $\vec{\nabla} f(\vec{a}) = \vec{0}$, we call $\vec{a}$ a critical point.

A critical point can be a local max, local min, a saddle point, or neither; see examples from last time.

As in $\mathbb{R}$, we need to look at second derivatives.

Taylor expansion: $f(\vec{a}+\vec{h}) = f(\vec{a}) + (\vec{\nabla} f)(\vec{a})\cdot\vec{h} + \underbrace{\frac{1}{2}(\vec{h}\cdot\vec{\nabla})^2 f(\vec{a})} + R_{2,\vec{a}}(\vec{h})$

$$= \frac{1}{2}\left(\sum_{i=1}^{n} h_i \partial_{x_i}\right)^2 f(\vec{a}) = \frac{1}{2}\sum_{i,j=1}^{n}\left(\partial_{x_i}\partial_{x_j} f\right)(\vec{a})\, h_i h_j$$

<u>Definition:</u>

$$H = H_f(\vec{a}) := \begin{pmatrix} \partial_{x_1}^2 f(\vec{a}) & \partial_{x_1}\partial_{x_2} f(\vec{a}) & \cdots & \partial_{x_1}\partial_{x_n} f(\vec{a}) \\ \partial_{x_2}\partial_{x_1} f(\vec{a}) & \partial_{x_2}^2 f(\vec{a}) & \cdots & \partial_{x_2}\partial_{x_n} f(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_n}\partial_{x_1} f(\vec{a}) & \partial_{x_n}\partial_{x_2} f(\vec{a}) & & \partial_{x_n}^2 f(\vec{a}) \end{pmatrix}$$ is called Hessian of f at $\vec{a}$

Remarks:
- The ij entry H_{ij} of the Hessian matrix is $H_{ij} = (\partial_{x_i}\partial_{x_j} f)(\vec{a})$
- $H_{ij} = H_{ji}$, i.e., H is a symmetric matrix

recall matrix multiplication: A $n\times n$ matrix, $\vec{y}\in\mathbb{R}^n$

$$\Rightarrow A\vec{y} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11}y_1 + \cdots + a_{1n}y_n \\ \vdots \\ a_{n1}y_1 + \cdots + a_{nn}y_n \end{pmatrix} \quad \text{or} \quad (A\vec{y})_i = \sum_{j=1}^{n} a_{ij} y_j$$

if also $\vec{x} \in \mathbb{R}^n \Rightarrow \vec{x} \cdot (A\vec{y}) = \sum_{i=1}^n x_i (A\vec{y})_i = \sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} y_j = \sum_{i,j=1}^n x_i a_{ij} y_j$

Back to Taylor expansion. We can write it as:

$$f(\vec{a}+\vec{h}) = f(\vec{a}) + \underbrace{(\vec{\nabla} f)(\vec{a}) \cdot \vec{h}}_{\substack{=0 \text{ if } \vec{a} \\ \text{critical point}}} + \frac{1}{2} \vec{h} \cdot \underbrace{(\overbrace{H_f}^{\text{matrix}}(\vec{a}) \vec{h})}_{\text{vector}} + R_{2,\vec{a}}(\vec{h})$$

↳ if $\vec{h} \cdot (H\vec{h}) > 0$ for all $\vec{h}$, $\vec{a}$ is a local min

↳ if $\vec{h} \cdot (H\vec{h}) < 0$ for all $\vec{h}$, $\vec{a}$ is a local max

↳ if $\vec{h} \cdot (H\vec{h}) > 0$ for some $\vec{h}$ and <0 for all others, $\vec{a}$ is a saddle point

↳ if $\vec{h} \cdot (H\vec{h}) = 0$ for some $\vec{h}$, we call $\vec{a}$ degenerate

We will discuss good criteria to determine whether $\vec{h} \cdot (H\vec{h}) \gtreqless 0$ later.

Now, consider just $f(x,y)$.

Trick: $$(\partial_x^2 f) \underbrace{\left(h_x + \frac{h_y (\partial_x \partial_y f)}{(\partial_x^2 f)} \right)^2}_{= h_x^2 + \frac{2 h_x h_y (\partial_x \partial_y f)}{(\partial_x^2 f)} + \frac{h_y^2 (\partial_x \partial_y f)^2}{(\partial_x^2 f)^2}} + h_y^2 \left((\partial_y^2 f) - \frac{(\partial_x \partial_y f)^2}{(\partial_x^2 f)} \right)$$

$$= h_x^2 (\partial_x^2 f) + 2 h_x h_y (\partial_x \partial_y f) + h_y^2 (\partial_y^2 f) = \vec{h} \cdot (H\vec{h})$$

since $\partial_x^2 f > 0$

$$\Rightarrow \vec{h} \cdot (H\vec{h}) > 0 \text{ if } \partial_x^2 f > 0 \text{ and } (\partial_y^2 f) - \frac{(\partial_x \partial_y f)^2}{(\partial_x^2 f)} > 0 \quad (\text{i.e., } (\partial_x^2 f) \cdot (\partial_y^2 f) - (\partial_x \partial_y f)^2 > 0)$$

Assuming that remainder $R_{2,\vec{a}}(\vec{h})$ is small enough (it is if $f \in C^2$), we have shown:

Theorem: Let $f: \mathbb{R}^2 \to \mathbb{R}$ be of class C^2, and suppose $(\vec{\nabla} f)(\vec{a}) = \vec{0}$ for some $\vec{a} \in \mathbb{R}^2$. Define $\alpha = (\partial_x^2 f)(\vec{a})$, $\beta = (\partial_x \partial_y f)(\vec{a})$, $\gamma = (\partial_y^2 f)(\vec{a})$ (i.e., $H = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$). Then

- if $\alpha\gamma - \beta^2 > 0$ and $\alpha > 0$, $\vec{a}$ is a local min
- if $\alpha\gamma - \beta^2 > 0$ and $\alpha < 0$, $\vec{a}$ is a local max
- if $\alpha\gamma - \beta^2 < 0$, $\vec{a}$ is a saddle point
- if $\alpha\gamma - \beta^2 = 0$, inconclusive

Ex.:

- $f(x,y) = -x^2 - y^2 \Rightarrow \vec{\nabla} f = \begin{pmatrix} -2x \\ -2y \end{pmatrix} \Rightarrow \vec{0}$ is critical point

  $$H_f(\vec{0}) = \begin{pmatrix} (\partial_x^2 f)(\vec{0}) & (\partial_x \partial_y f)(\vec{0}) \\ (\partial_x \partial_y f)(\vec{0}) & (\partial_y^2 f)(\vec{0}) \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

  $\Rightarrow \alpha = -2, \beta = 0, \gamma = -2 \Rightarrow \alpha < 0$ and $\alpha\gamma - \beta^2 = 4 > 0 \Rightarrow$ local max

- $f(x,y) = x^2 - y^2$

  $\Rightarrow \alpha = 2, \beta = 0, \gamma = -2 \Rightarrow \alpha\gamma - \beta^2 = -4 < 0 \Rightarrow$ saddle point

- $f(x,y) = y^3 - 3x^2 y$

  $(\vec{\nabla} f) = \begin{pmatrix} -6xy \\ 3y^2 - 3x^2 \end{pmatrix} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is the only critical point

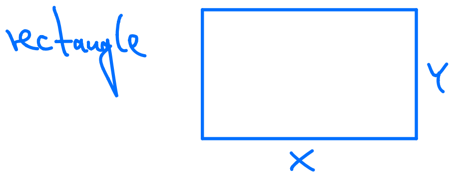
  $H_f = \begin{pmatrix} -6y & -6x \\ -6x & 6y \end{pmatrix} \Rightarrow H_f(\vec{0}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow$ inconclusive

  visualization with geogebra shows that $\vec{0}$ is neither max nor mini; it is a "monkey saddle"

More examples in HW and moodle exercises

2.4 Lagrange Multipliers

Recall the isoperimetric problem (Chapter "Optimization") for rectangles from Calculus and Linear Algebra I:



Question: maximize $f(x,y)=xy$ under the constraint that $2x+2y=L$, for given perimeter L

(Recall the solution: $f(x,y(x))=x\left(\frac{L}{2}-x\right) \Rightarrow 0=\partial_x f=\frac{L}{2}-2x \Rightarrow x_{max}=\frac{L}{4}$

$\Rightarrow y_{max}=\frac{L}{2}-x_{max}=\frac{L}{4} \Rightarrow$ square!)

General question: Find extrema of $f(\vec{x})$ under some constraint $\underbrace{G(\vec{x})=0}_{\text{e.g., above, } G(x,y)=2x+2y-L}$

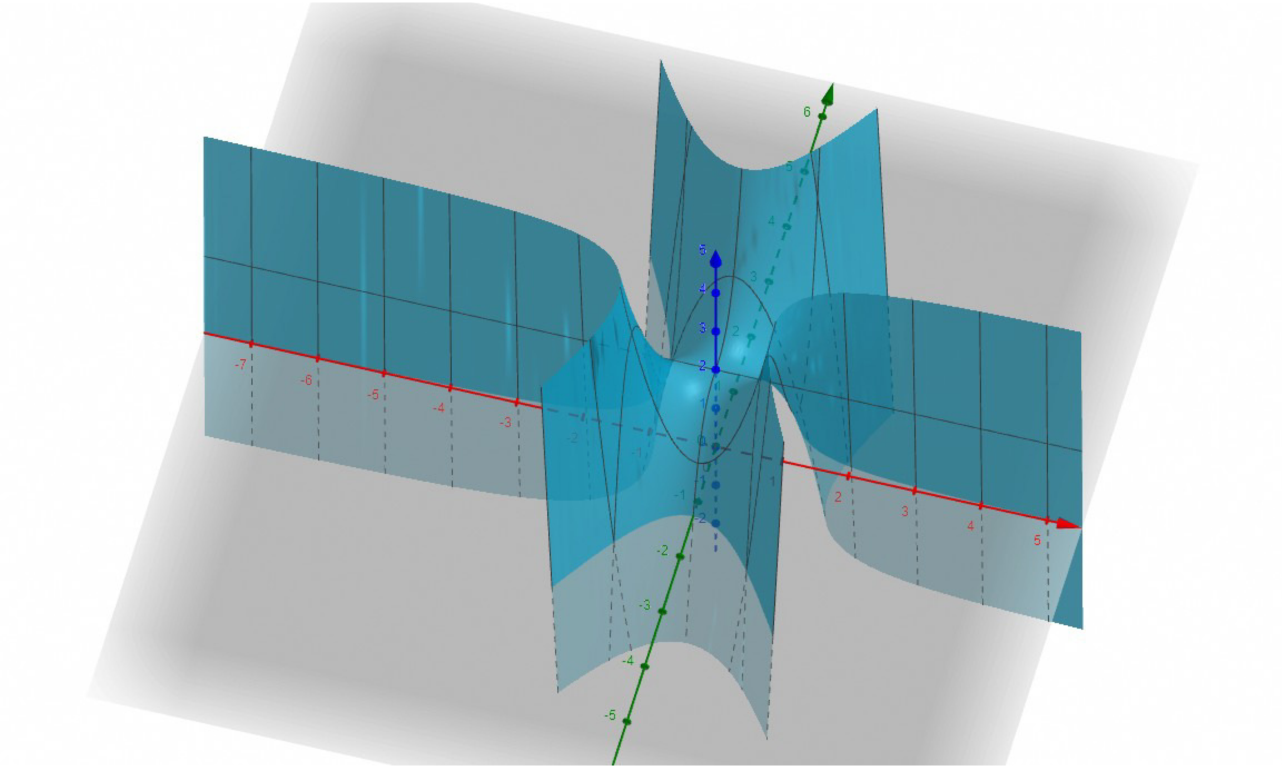
Note: it is often impractical or impossible to explicitly solve $G(x,y)=0$ for x (or y), e.g.,

$$G(x,y)=e^{\sin(xy)}-\cos(x^2y^2)=0$$

Goal next time: find a better method

Visualizations using geogebra.org:

The monkey saddle: $f(x,y) = y^3 - 3x^2y$



The isoperimetric problem for rectangles:
$f(x,y) = xy$ with constraint $G(x,y) = 2x+2y-1$ (perimeter $L=1$ here)