MM3001

Lecture 11: Functions of several variables and optimization

Sofia Tirabassi

Review & Clarifications

Stockholm University

Opening the Floor for Clarifications

Before we dive into multi-variable optimization algorithms today, what questions do you have regarding last lecture's material on **Functions of Several Variables, Partial Derivatives, or the Chain Rule**?

60-Second Flash Recall: Warming Up Your Mathematical Brain

To help recall core mechanics while you look through your notes for questions:

- What is the geometric interpretation of a first-order partial derivative $\frac{\partial f}{\partial x}$?
- If $f(x, y) = 3x^2y^4$, what is the expression for the mixed second-order derivative f''_{xy}?

Questions?

Lecture Goal and Outcome

Stockholm University

Goals:

- Find stationary points for a function.
- Use the second-order partial derivatives to classify the stationary points of a function of two variables.

Learning Outcome: After today you will be able to solve problems like the following:

Problem

Find all the stationary points of the function $f(x, y) = (4 - x - y^2)(x + 1)$ and classify them.

Why You Should Care: Multi-Product Monopolies

The Economic Reality of Optimization

Most businesses do not sell just one product. Consider an electronics firm selling smartphones (x) and tablets (y). Their total profit function depends on both quantities simultaneously: $\Pi(x, y)$.

The Mathematical Framework of Multi-Variable Decisions:

- To find the optimal price or quantity mix, the firm must satisfy the **Necessary First-Order Conditions**: $\Pi'_x = 0$ and $\Pi'_y = 0$.
- But how do we know if a stationary point is a profit-maximizing peak, a cost-minimizing valley, or a dead-end trap?
- Today we introduce the **Hessian Matrix** (H) and its determinant—the ultimate tool for multi-variable risk classification!

Lecture Plan

- §17.1 Two Variables: Necessary Conditions
- Hessian of a function.
- Classification of stationary points.
- Sufficient conditions for extrema

Section 1
Classification of Stationary points

The Hessian

Given a function $f : D \to \mathbb{R}$ we define its Hessian as the table (matrix)

$$H(x, y) := \begin{bmatrix} f''_{xx}(x, y) & f''_{xy}(x, y) \\ f''_{yx}(x, y) & f''_{yy}(x, y) \end{bmatrix}$$

and we define its determinant as

$$\det H(x, y) = f''_{xx}(x, y) f''_{yy}(x, y) - f''_{xy}(x, y) f''_{yx}(x, y)$$

For regular functions one has that $f''_{xy} = f''_{yx}$, so we could write

$$\det H(x, y) = f''_{xx}(x, y) f''_{yy}(x, y) - (f''_{xy}(x, y))^2$$

Examples: Calculate the Hessian and its determinant for $f(x, y) = -(x^2 + y^2)$, $f(x, y) = x^2 + y^2$ and $f(x, y) = x^2 - y^2$.

Second derivative test

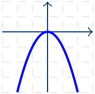
Second derivative test

Let (x_0, y_0) be a stationary point of f. Recall that

$$\det H(x_0, y_0) = f''_{xx}(x_0, y_0) f''_{yy}(x_0, y_0) - f''_{xy}(x_0, y_0) f''_{yx}(x_0, y_0)$$

1. If $\det H(x_0, y_0) > 0$ and $f''_{xx}(x_0, y_0) < 0$, then f has a local maximum at (x_0, y_0).
2. If $\det H(x_0, y_0) > 0$ and $f''_{xx}(x_0, y_0) > 0$, then f has a local minimum at (x_0, y_0).
3. If $\det H(x_0, y_0) < 0$ then f has a saddle point at (x_0, y_0).

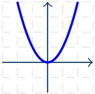
$f(x) = -x^2$



$f''(0) = -2 < 0$

Local Maximum

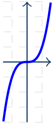
$f(x) = x^2$



$f''(0) = 2 > 0$

Local Minimum

$f(x) = x^3$



$f''(0) = 0$

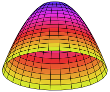
$f''(x) < 0$ if $x < 0$

$f''(x) > 0$ if $x > 0$

$f'(0) = 0$

Stockholm University

$-x^2 - y^2$

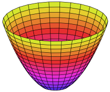


$$H = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

$\det H = 4 > 0$

Local Maximum

$x^2 + y^2$

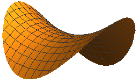


$$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$\det H = 4 > 0$

Local Minimum

$x^2 - y^2$



$$H = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

$\det H = -4 < 0$

Saddle Point

$f'_x(0,0) = 0$

$f'_y(0,0) = 0$

Factoring the System Without Losing Roots

Stockholm University

Think-Pair-Share: Navigating Nonlinear Stationary Points (5 Minutes)

Let us solve our lecture goal from Slide 3 : $f(x, y) = (4 - x - y^2)(x + 1)$. Differentiating yields the first-order system:

$$1) \quad f'_x = 3 - 2x - y^2 = 0 \qquad 2) \quad f'_y = -2y(x + 1) = 0$$

Your Task with Your Bänkkamrat:

1. **Think (2 mins):** Look at equation (2). What two independent mathematical scenarios make $-2y(x + 1) = 0$ true?
2. **Pair (2 mins):** Substitute those scenarios into equation (1) to isolate all matching x and y coordinate values. How many unique stationary points did you locate?
3. **Share (1 min):** Let's verify why losing the boundary root $x = -1$ ruins the classification step.

Active Learning 2

Stockholm University

Peer Instruction: Testing the Global Boundary Criteria

Suppose a production cost function yields the following generalized Hessian matrix everywhere across its convex domain D:

$$H(x, y) = \begin{bmatrix} 6 & -2 \\ -2 & 4 \end{bmatrix}$$

What classification type does this cost matrix reflect?

A) $\det H < 0 \implies$ A saddle point landscape.

B) $\det H > 0$ and $f''_{xx} > 0 \implies$ Globally convex (Global Min).

C) $\det H > 0$ and $f''_{xx} < 0 \implies$ Globally concave (Global Max).

D) The matrix is inconclusive without knowing (x_0, y_0).

Action: On my count of three, show 1, 2, 3, or 4 fingers. Let's apply the sufficient conditions from before to verify why a constant positive determinant guarantees a global result!

Examples

- (2015) Let $F(x,y) = x^4 + y^4 - 36xy$. Find all the stationary points for this function and determine whether they are local maximum, minimum or saddle points.
- (2024) Let $F(x,y) = \ln(1 + x^2 + y^2)$. Find all the stationary points for this function and determine their type.
- (2024)Let $F(x,y) = \sqrt{x^2 + y^2} - y^2 - 1$. Find all the stationary points for this function and determine their type.
- (2024) $F(x,y) = xye^{-xy}$. Find all the stationary points for this function and determine their type.

Exam Workshop 1:

Stockholm University

Think-Pair-Share: Finding Stationary Points

Consider the 2015 exam function:

$$F(x, y) = x^4 + y^4 - 36xy$$

Taking first derivatives yields: $F'_x = 4x^3 - 36y = 0$ and $F'_y = 4y^3 - 36x = 0$.

Your Task with Your Neighbor:

1. **Think (2 mins):** Simplify the equations to $y = \frac{x^3}{9}$ and substitute this into the second equation. Factor out x to avoid losing the origin root!
2. **Pair (2 mins):** Solve for all real values of x. Remember that $x^8 - 9^4 = 0$ splits into real and imaginary roots. How many real pairs (x, y) did you locate?
3. **Share (1 min):** Why is $(0, 0)$ a common trap when calculating the Hessian determinant later?

Exam Workshop 2: Peer Instruction on the 2024 Log Model

Peer Instruction: Testing Chain Rule Proportionality (3 Minutes)

Look at the first 2024 exam problem on Slide 10:

$$F(x, y) = \ln(1 + x^2 + y^2)$$

Using the Chain Rule, the partial derivatives are:

$$F'_x = \frac{2x}{1 + x^2 + y^2} = 0 \quad \text{and} \quad F'_y = \frac{2y}{1 + x^2 + y^2} = 0$$

Concept Check: What is the single stationary point, and its type?

A) $(0, 0)$ is a saddle point.

B) $(0, 0)$ is a local minimum.

C) No stationary points exist because of the log.

D) Any point on the unit circle

Exam Workshop 3: Group Challenge on the Product Rule

Stockholm University

Group Challenge: Factoring the Exponential Boundary (4 Minutes)

Look at the final 2024 exam function on Slide 10:

$$F(x, y) = xye^{-xy}$$

Let $u = xy$. Then the function simplifies to a univariate form $f(u) = ue^{-u}$.

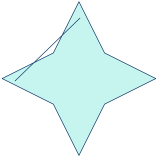
Work with your team to decode the stationary condition:

1. Compute the derivative with respect to u: $f'(u) = (1 - u)e^{-u}$.
2. Set $f'(u) = 0$. Since $e^{-u} \neq 0$, what specific value must the product xy equal?
3. **The Critical Question:** Does this function have isolated stationary points, or an entire **stationary curve** of points? What does this mean for the standard Hessian test?

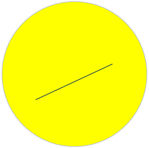
Section 2
Sufficient conditions for extrema

Convexity

We say that a set $D \subset \mathbb{R}^2$ is convex, if **for every** pair of points $X, Y \in D$ the segment $[X, Y]$ joining them in contained in D.



Not convex



Convex

A function is *convex/concave* if the region above/below its graph is a convex set.

Sufficient conditions

Second derivative test

Assume that:

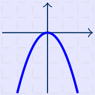
- D is a convex domain;
- (x_0, y_0) is an interior stationary point of f
- and that **for all** $(x, y) \in D$

$$\det H(x,y) = f''_{xx}(x,y)f''_{yy}(x,y) - f''_{xy}(x,y)f''_{yx}(x,y) \geq 0$$

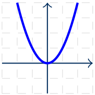
1. If $f''_{xx}(x,y) \leq 0$ and $f''_{yy}(x,y) \leq 0$ **for all** $(x,y) \in D$, then f is concave on D and (x_0, y_0) is a maximum point for f in D.
2. If $f''_{xx}(x,y) \geq 0$ and $f''_{yy}(x,y) \geq 0$ **for all** $(x,y) \in D$, then f is convex on D and (x_0, y_0) is a minimum point for f in D.

How can we recall the theorem?

$f(x) = -x^2$



$f(x) = x^2$

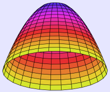


$x = 0$ is a stationary point: $f'(0) = 0$

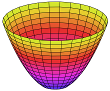
$f''(x) = -2 < 0$ (Concave)
Maximum

$f''(x) = 2 > 0$ (Convex)
Minimum

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



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



$(0, 0)$ is a stationary point $f'_x(0, 0) = f'_y(0, 0) = 0$

$H = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$

$\det H = f''_{xx} f''_{yy} - (f''_{xy})^2 \geq 0$

$f''_{xx}(x, y) \leq 0 \quad f''_{yy}(x, y) \leq 0$

(Concave) Maximum

$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$\det H = f''_{xx} f''_{yy} - (f''_{xy})^2 > 0$

$f''_{xx}(x, y) \geq 0 \quad f''_{yy}(x, y) \geq 0$

(Convex) Minimum

Example

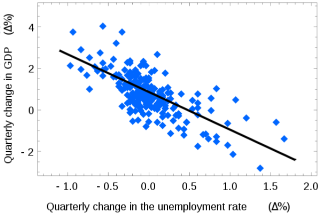
Show that the function

$$f(x,y) = 2x^2 + 4xy + 3y^2 - x - 2y$$

is convex and has a global minimum on $\mathbb{R}^2$.

Application:Linear regression

Given some data (x_t, y_y) $t = 1, \dots, T$, we want to find a line $y = \alpha + \beta x$ such that, if $e_t := \alpha + \beta x_t - y_t$ (error term at time t), it minimises the function



$$L(\alpha, \beta) = \frac{1}{T} \sum_{t=1}^{T} e_t^2$$

Define $\mu_x = \frac{1}{T} \sum_{t=1}^{T} x_t \quad \mu_y = \frac{1}{T} \sum_{t=1}^{T} y_t$ and

$$\sigma_{xx} = \frac{1}{T} \sum_{t=1}^{T} (x_t - \mu_x)^2, \qquad \sigma_{yy} = \frac{1}{T} \sum_{t=1}^{T} (y_t - \mu_y)^2,$$

$$\sigma_{xy} = \frac{1}{T} \sum_{t=1}^{T} (x_t - \mu_x)(y_t - \mu_y)$$

Assume that $\sigma_{xx} > 0$ (otherwise, all the points lie on a vertical line)

One can show (after some calculations) that $L(\alpha, \beta) =$
$= \alpha^2 + \mu_y^2 + \beta^2\mu_x^2 - 2\alpha\mu_y - 2\beta\mu_x\mu_y + 2\alpha\beta\mu_x + \beta^2\sigma_{xx} - 2\beta\sigma_{xy} + \sigma_{yy}$
Also that the unique stationary point is

$$\hat{\alpha} = \mu_y - \frac{\sigma_{xy}}{\sigma_{xx}}\mu_x, \qquad \hat{\beta} = \frac{\sigma_{xy}}{\sigma_{xx}}$$

Note that

$$H(\alpha, \beta) = \begin{bmatrix} 2 & 2\mu_x \\ 2\mu_x & 2\mu_x^2 + 2\sigma_{xx} \end{bmatrix} \Rightarrow \det H(\alpha, \beta) = 4\sigma_{xx} > 0$$

$$2 = f''_{\alpha\alpha} > 0 \qquad 2\mu_x^2 + 2\sigma_{xx} = L''_{\beta\beta} > 0$$

Since $\mathbb{R}^2$ is convex, by the previous theorem, L is convex and the stationary point is a global minimum for $L(\alpha, \beta)$.

$y = \hat{\alpha} + \hat{\beta}x$ is called the regression line.

Thank you for your attention!