Extrema of Functions of Several Variables

Extrema of Single Variable Functions

We let $\mathbb{R}$ denote the set of real numbers.
For $a, b \in \mathbb{R}$ with $a \leq b$, we let

$$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}, \qquad \text{closed interval}$$
$$(a, b) := \{x \in \mathbb{R} : a < x < b\}, \qquad \text{open interval}$$

Let $f : (a, b) \to \mathbb{R}$ be a differentiable function. We say that $c \in (a, b)$ is a *critical point* of f if $f'(c) = 0$.

Extrema of Single Variable Functions

Theorem (Second Derivative Test)

Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function and $c \in (a, b)$ be a critical point of f (that is, $f'(c) = 0$).

(i) *If $f''(c) > 0$, then f has a relative minimum at c.*

(ii) *If $f''(c) < 0$, then f has a relative maximum at c.*

(iii) *If $f''(c) = 0$ AND f'' changes sign when passing through c then c is an inflection point for f*

Partial derivatives

Let U be an open subset of $\mathbb{R}^n$ and $f\colon U \to \mathbb{R}$.
We introduce the idea of partial derivatives of f.
For $n = 2$, $f(x, y)$ has two partial derivatives:

$$f_x = \frac{\partial f}{\partial x}$$

is the derivative with respect to x, obtained by differentiating f with respect to x treating y as a constant

$$f_y = \frac{\partial f}{\partial y}$$

is the derivative with respect to y, obtained by differentiating f with respect to y treating x as a constant

Partial derivatives

For functions of n variables, $f(x_1, x_2, \ldots, x_n)$ we have n first order partial derivatives

$$f_{x_1} = \frac{\partial f}{\partial x_1}, \qquad f_{x_2} = \frac{\partial f}{\partial x_2}, \ldots, f_{x_n} = \frac{\partial f}{\partial x_n}$$

Example:

$$f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$$

$$f_x(x, y, z) = 10x + 4y + e^z, \qquad f_y(x, y, z) = 4x + 3y^2 - 2,$$

$$f_z(x, y, z) = 5z^4 + xe^z.$$

Critical points

As with functions of one variable an extremum (maximum or minimum) can only occur at a vector where all first order partial derivatives are 0.
We write this as

$$\nabla f(\mathbf{x}) = 0,$$

where

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{x})\right).$$

Critical points

For $n = 3$ we have

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x}(\mathbf{x}), \frac{\partial f}{\partial y}(\mathbf{x}), \frac{\partial f}{\partial z}(\mathbf{x}) \right)$$

Example:

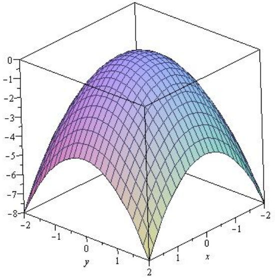
$$f(x, y) = -2x^2 - y^2 + 8x + 10y - 5xy$$

Nature of critical points

At a critical point, a function can have

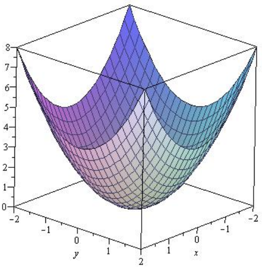
- a (local) maximum,
- a (local) minimum,
- a saddle point,
- non of the above.

An example of a maximum point:

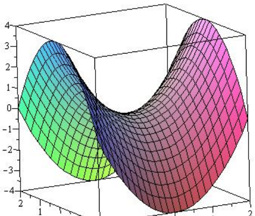


Nature of critical points

An example of a minimum point:



An example of a saddle point:



Higher Order Partial Derivatives

To classify critical points of functions of several variables we need higher order partial derivatives.
For functions of two variables, there are four second order partial derivatives:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}, \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x},$$
$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}, \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}.$$

We also use the notation $f_{xy} = \frac{\partial^2 f}{\partial y \partial x}$. Note the order of the indices!

Higher Order Partial Derivatives

For functions of n variables there are n^2 second order partial derivatives

$$f_{x_i x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).$$

If $f : \mathbb{R}^3 \to \mathbb{R}$ is give by

$$f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$$

$$\begin{array}{ccc} f_{xx} = 10 & f_{xy} = 4 & f_{xz} = e^z \\ f_{yx} = 4 & f_{yy} = 6y & f_{yz} = 0 \\ f_{zx} = e^z & f_{zy} = 0 & f_{zz} = 20z^3 + xe^z \end{array}$$

Hessian matrix

The *Hessian matrix* H of the function f is the matrix of all second order partial derivatives of f.

$$H(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & f_{x_1x_3}(\mathbf{x}) & \cdots & f_{x_1x_n}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & f_{x_2x_3}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ f_{x_3x_1}(\mathbf{x}) & f_{x_3x_2}(\mathbf{x}) & f_{x_3x_3}(\mathbf{x}) & \cdots & f_{x_2x_n}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{x}) & f_{x_nx_2}(\mathbf{x}) & f_{x_nx_3}(\mathbf{x}) & \cdots & f_{x_nx_n}(\mathbf{x}) \end{pmatrix},$$

For most functions used in practical applications, we have

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i},$$

hence the Hessian matrix is *symmetric*.

Hessian matrix

Example: The Hessian matrix of
$f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$ is

$$H(\mathbf{x}) = \begin{pmatrix} 10 & 4 & e^z \\ 4 & 6y & 0 \\ e^z & 0 & 20z^3 + xe^z \end{pmatrix}.$$

Principal minors

For an $n \times n$ Hessian matrix we identify n submatrices. a set of n *submatrices* The first of these, denoted by $H_1(\mathbf{x})$, is the 1×1 submatrix consisting of the element located in row 1 and column 1,

$$H_1(\mathbf{x}) = \left(f_{x_1x_1}(\mathbf{x})\right).$$

The second submatrix is the 2×2 matrix

$$H_2(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) \end{pmatrix}.$$

The third submatrix is the 3×3 matrix

$$H_3(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & f_{x_1x_2}(\mathbf{x}) & f_{x_1x_3}(\mathbf{x}) \\ f_{x_2x_1}(\mathbf{x}) & f_{x_2x_2}(\mathbf{x}) & f_{x_2x_3}(\mathbf{x}) \\ f_{x_3x_1}(\mathbf{x}) & f_{x_3x_2}(\mathbf{x}) & f_{x_3x_3}(\mathbf{x}) \end{pmatrix},$$

Principal minors

The j^{th} submatrix is the $j \times j$ matrix

$$H_j(\mathbf{x}) = \begin{pmatrix} f_{x_1x_1}(\mathbf{x}) & \cdots & f_{x_1x_j}(\mathbf{x}) \\ \vdots & & \vdots \\ f_{x_jx_1}(\mathbf{x}) & \cdots & f_{x_jx_j}(\mathbf{x}) \end{pmatrix}$$

and $H_n(\mathbf{x}) = H(\mathbf{x})$.

Principal minors (continued)

The determinants of these submatrices are called *principal minors* and are denoted by $\Delta_k(\mathbf{x})$

$$\Delta_k(\mathbf{x}) = \det(H_k(\mathbf{x})) \qquad k = 1, \ldots, n.$$

Example: The principal minors of $f(x, y, z) = 5x^2 + 4xy + y^3 - 2y + z^5 + xe^z$ are

$$\Delta_1(\mathbf{x}) = H_1(\mathbf{x}) = 10$$

$$\Delta_2(\mathbf{x}) = \det H_2(\mathbf{x}) = 60y - 16$$

$$\Delta_3(\mathbf{x}) = \det H_3(\mathbf{x}) = e^z(-6ye^z) + (20z^3 + xe^z)(60y - 16)$$

Second Derivative Test

Theorem (Second Derivative Test)

Let $\mathbf{p}$ *be a critical point of a function* $f : \mathbb{R}^n \to \mathbb{R}$.

- *If* $(-1)^k \Delta_k(\mathbf{p}) > 0$ *for all* $k = 1, \ldots, n$, *then* f *has a local maximum at* $\mathbf{p}$.
- *If* $\Delta_k(\mathbf{p}) > 0$ *for all* $k = 1, \ldots, n$, *then* f *has a local minimum at* $\mathbf{p}$.
- *If* $\Delta_n(\mathbf{p}) = \det(H(\mathbf{p})) \neq 0$, *but none of the previous conditions holds, then* f *has a saddle point at* $\mathbf{p}$.
- *If* $\Delta_n(\mathbf{p}) = \det(H(\mathbf{p})) = 0$ *the test in inconclusive,*

Second Derivative Test

Find and classify all critical points of the function

$$f(x, y, z) = 2x^2 + xy + 4y^2 + xz + z^2 + 2.$$

Find and classify all critical points of the function

$$f(x, y, z) = x^2 + y^3 + z^2 - 2xy - 2yz + 1.$$

Second Derivative Test

Exercise:

For $a \in \mathbb{R}$ and consider the function

$$f(x, y, z) = x^2 + ay^2 + z^2 - 4xy.$$

Determine how the critical points of f and their nature depends on a.