

# R401: Statistical and Mathematical Foundations

## Lecture 14: Unconstrained Optimization. Static Optimization with Equality Constraints. Lagrange Multipliers

Andrey Vassilev

2016/2017

# General Principles and Caveats for the Optimization Module

- Emphasis on practicality over rigour
- Consequently, algorithmic approach and “recipes” rather than proofs
- Also, existence and relevant properties of various objects are often implicitly assumed
- Pathologies and mathematical peculiarities discussed only in special cases

# Lecture Contents

1 Warm-up: Basic Unconstrained Optimization in  $\mathbb{R}^1$

2 Unconstrained Optimization in  $\mathbb{R}^n$

# Warm-up: Basic Unconstrained Optimization in $\mathbb{R}^1$

## Fact 1

For a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  differentiable at a point  $x$ , a necessary condition for a local extreme point (i.e. a maximum or a minimum) at  $x$  is

$$f'(x) = 0.$$

## Example 1

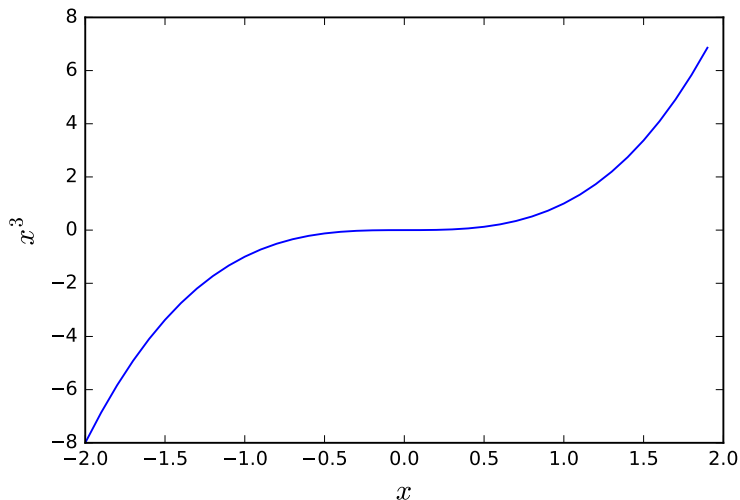
If  $f(x) = ax^2 + bx + c$ , then  $f'(x) = 2ax + b$  and the condition  $f'(x) = 0$  yields the familiar  $x = -\frac{b}{2a}$  (recall your high-school days). Depending on the sign of  $a$ , this is a maximum or a minimum (What is the relationship?).

## Example 2

If  $f(x) = x^3$ , then  $f'(x) = 3x^2$  and  $f'(x) = 0 \Rightarrow x = 0$ .

Does the function attain a maximum or a minimum at  $x = 0$ ?

# Warm-up: Basic Unconstrained Optimization in $\mathbb{R}^1$



# Warm-up: Basic Unconstrained Optimization in $\mathbb{R}^1$

## Example 2 (cont.)

The answer is “neither”! The point  $x = 0$  is not a local extreme point of  $f(x) = x^3$ .

This illustrates the pitfalls of using necessary conditions – they supply only candidates that need to be checked further.

The above examples generalize in the following manner:

## Fact 2

Let a function  $f$  be  $n$  times differentiable at a point  $x$  and

$$f'(x) = f''(x) = \dots = f^{(n-1)}(x) = 0, \quad f^{(n)} \neq 0.$$

- ① If  $n$  is odd, the point  $x$  is not an extreme point of  $f(x)$ .
- ② If  $n$  is even and  $f^{(n)}(x) > 0$ , the point  $x$  is a minimum.
- ③ If  $n$  is even and  $f^{(n)}(x) < 0$ , the point  $x$  is a maximum.

# Unconstrained Optimization in $\mathbb{R}^n$

## Necessary conditions

### Fact 3

For a function  $f : S \rightarrow \mathbb{R}$ ,  $S \subseteq \mathbb{R}^n$ , differentiable at an interior point  $\mathbf{x}$ , a necessary condition for  $\mathbf{x}$  to be a local extreme point is

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

where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad f'(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(x_1, \dots, x_n)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x_1, \dots, x_n)}{\partial x_n} \end{pmatrix}$$

**Note:** A point where all partial derivatives vanish is called a *critical point* or a *stationary point* of the function  $f$ . This also applies to functions on  $\mathbb{R}^1$ .

# Unconstrained Optimization in $\mathbb{R}^n$

## Example 3

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

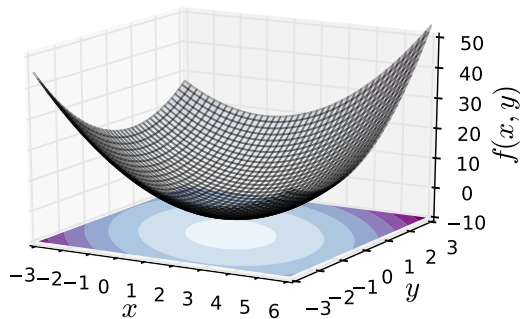
$$\frac{\partial f}{\partial x} = 2x - 3 + y = 0 \quad \Rightarrow \quad x = \frac{3 - y}{2}$$

$$\frac{\partial f}{\partial y} = 4y + x = 0 \quad \Rightarrow \quad y = -\frac{x}{4}$$

$$x = \frac{12}{7}, \quad y = -\frac{3}{7}$$



# Unconstrained Optimization in $\mathbb{R}^n$



# Unconstrained Optimization in $\mathbb{R}^n$

The necessity of the condition  $f'(\mathbf{x}) = \mathbf{0}$  has implications that are similar to the univariate case:

## Example 4

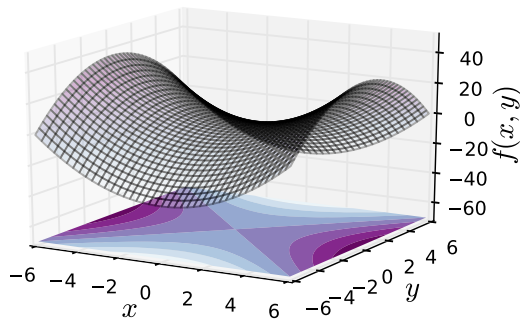
Consider the function  $f(x, y) = x^2 - y^2$ . The NCs yield the following candidate:

$$\frac{\partial f}{\partial x} = 2x = 0 \quad \Rightarrow \quad x = 0,$$

$$\frac{\partial f}{\partial y} = -2y = 0 \quad \Rightarrow \quad y = 0.$$

Let's look at the graph of the function in a neighbourhood of the point  $(0,0)'$ .

# Unconstrained Optimization in $\mathbb{R}^n$



# Unconstrained Optimization in $\mathbb{R}^n$

## Example 4 (cont.)

The critical point  $\mathbf{x} = (0,0)'$  is an example of a *saddle point*. The function  $f$  (obviously) does not attain an extremum at  $\mathbf{x}$ .

Example 4 illustrates the need for a counterpart of ...