Session 11: Optimality conditions

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

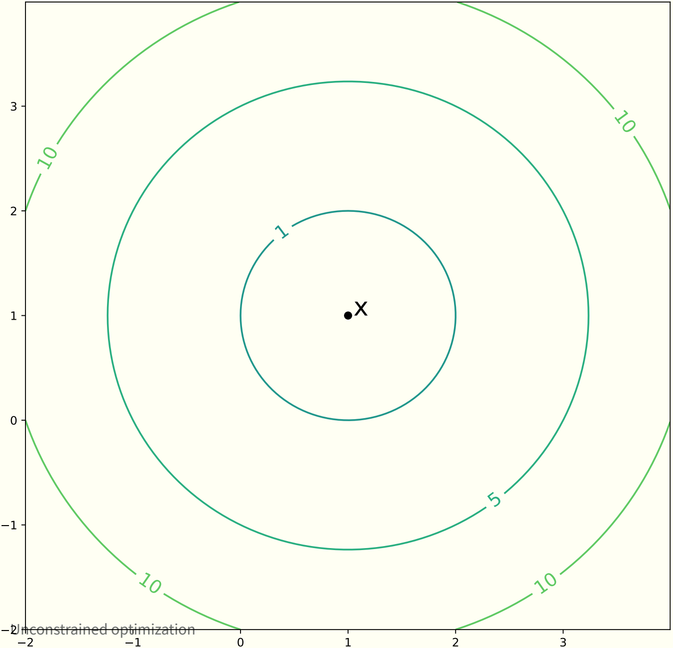
Contents

Unconstrained optimization

Questions about the video?

- Global minimizer $\implies$ local minimizer $\implies$ critical point.
- Critical point + positive definite Hessian $\implies$ local minimizer.

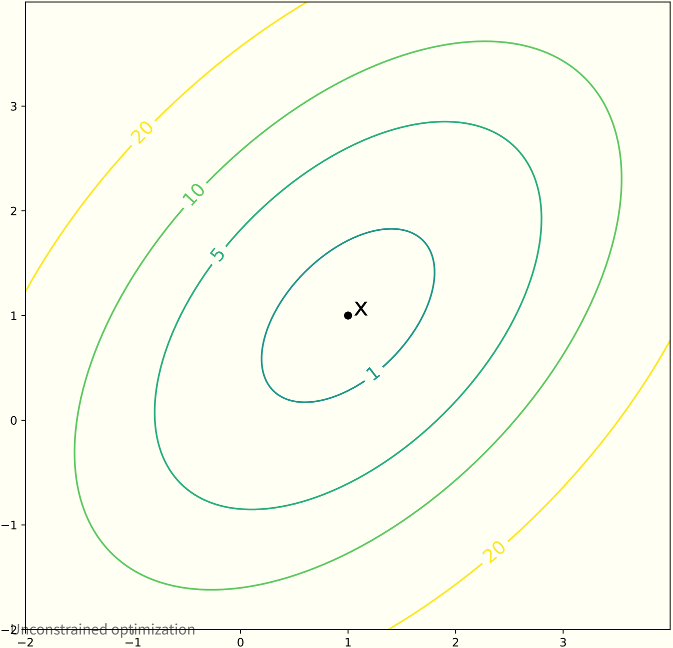
Hessian at a critical point



The eigenvalues of the Hessian at x are

1. $\begin{cases} \lambda_1 = 1 \\ \lambda_2 = -1 \end{cases}$
2. $\begin{cases} \lambda_1 = 1 \\ \lambda_2 = 1 \end{cases}$
3. $\begin{cases} \lambda_1 = 2 \\ \lambda_2 = 1 \end{cases}$
4. $\begin{cases} \lambda_1 = -1 \\ \lambda_2 = -1 \end{cases}$

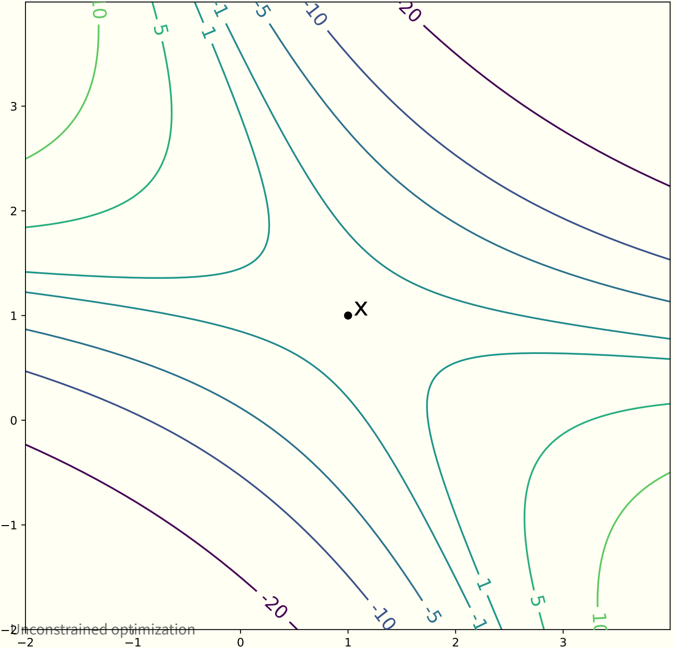
Hessian at a critical point



The eigenvalues of the Hessian at x are

1. $\begin{cases} \lambda_1 = 1 \\ \lambda_2 = -3 \end{cases}$
2. $\begin{cases} \lambda_1 = 2 \\ \lambda_2 = 2 \end{cases}$
3. $\begin{cases} \lambda_1 = 1 \\ \lambda_2 = 3 \end{cases}$
4. $\begin{cases} \lambda_1 = -1 \\ \lambda_2 = -1 \end{cases}$

Hessian at a critical point



The eigenvalues of the Hessian at x are

1. $\begin{cases} \lambda_1 = 1 \\ \lambda_2 = -3 \end{cases}$
2. $\begin{cases} \lambda_1 = 2 \\ \lambda_2 = 2 \end{cases}$
3. $\begin{cases} \lambda_1 = 1 \\ \lambda_2 = 3 \end{cases}$
4. $\begin{cases} \lambda_1 = -1 \\ \lambda_2 = -1 \end{cases}$

Constrained optimization

General formulation

Constrained optimization problems take the form:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p, \end{array}$$

with variable $x \in \mathbb{R}^n$.

Question

If x is a solution to

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p, \end{array}$$

do we have $\nabla f(x) = 0$?

Feasible points

Definition

A point $x \in \mathbb{R}^n$ is *feasible* if it satisfies all the constraints:
$g_1(x) \leq 0, \ldots, g_m(x) \leq 0$ and $h_1(x) = 0, \ldots, h_p(x) = 0$.

First order optimality condition

First order optimality condition

First order optimality condition

First order optimality condition

First order optimality condition

Theorem

If x is a solution and if $\nabla h_1(x), \ldots, \nabla h_p(x), \{\nabla g_i(x) \mid g_i(x) = 0\}$ are linearly independent, then there exists $\lambda_1, \ldots, \lambda_m \geq 0$ and $\nu_1, \ldots, \nu_p \in \mathbb{R}$ such that:

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0. \tag{1}$$

Moreover, for all $i \in \{1, \ldots, m\}$, if $g_i(x) < 0$ then $\lambda_i = 0$.

Example

Let $u \in \mathbb{R}^n$ be a non-zero vector.

$$\begin{array}{ll} \text{Minimize} & \langle x, u \rangle \\ \text{subject to} & \|x\|^2 = 1. \end{array}$$

Example

Let $u \in \mathbb{R}^n$ be a non-zero vector.

$$\begin{array}{ll} \text{Minimize} & \langle x, u \rangle \\ \text{subject to} & \|x\|^2 = 1. \end{array}$$

Convex constrained optimization

General formulation

We say that the constrained optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p, \end{array} \tag{2}$$

is convex when $f, g_1, \ldots, g_m$ are convex and $h_1, \ldots, h_p$ are affine.

First order optimality condition

Theorem

Assume that the problem is convex and that there exists a feasible point x_0 such that $g_i(x_0) < 0$ for all i.
Then x is a solution if and only if x is feasible and there exists $\lambda_1, \dots, \lambda_m \geq 0$, $\nu_1, \dots, \nu_p \in \mathbb{R}$ such that:

$$\begin{cases} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0. \\ \lambda_i g_i(x) = 0, \ \text{for all } i \in \{1, \dots, p\}. \end{cases}$$

Example: Ridge regression

$$
\begin{array}{ll}
\text{minimize} & \|Ax - y\|^2 \\
\text{subject to} & \|x\|^2 \le r^2.
\end{array}
$$

Example

Let $u, v \in \mathbb{R}^n$ such that $\|v\| = 1$. Solve:

$$\begin{array}{ll} \text{minimize} & \|x - u\|^2 \\ \text{subject to} & x \perp v. \end{array}$$

Example

Questions?

Questions?