Lecture 7

Optimality Conditions

7.1 Lecture Objectives

- Define local and global minimizers
- Understand the first order conditions for local minimizers in both unconstrained and constrained optimization problems
- Understand the second order conditions for local minimizers

7.2 Local and Global Minimizers

7.2.1 Local minimizers

Consider an optimization problem

$$\min_{\mathbf{x}} f(\mathbf{x}), \text{ s.t. } \mathbf{x} \in \Omega \tag{7.1}$$

where Ω defines the constraint set for $\mathbf{x}$ (ie: either the entire $\mathbb{R}^N$ for an unconstrained problem, or some subset determined by equality or inequality constraints otherwise). A point $\mathbf{x}^*$ is a local minimizer if there exists an $\epsilon > 0$ such that:

$$f(\mathbf{x}^*) \leq f(\mathbf{x}^*), \text{ for all } \mathbf{x} \in \Omega \text{ such that } \|\mathbf{x} - \mathbf{x}^*\| < \epsilon \tag{7.2}$$

Further, if the inequality above is a strict inequality, then $\mathbf{x}^*$ is a *strict* local minimizer, ie:

$$f(\mathbf{x}^*) < f(\mathbf{x}^*), \text{ for all } \mathbf{x} \in \Omega \text{ such that } \|\mathbf{x} - \mathbf{x}^*\| < \epsilon \tag{7.3}$$

7.2.2 Global minimizers

A point $\mathbf{x}^*$ is a global minimizer if:

$$f(\mathbf{x}^*) \leq f(\mathbf{x}^*), \text{ for all } \mathbf{x} \in \Omega \tag{7.4}$$

Further, if the inequality above is a strict inequality, then $\mathbf{x}^*$ is a *strict* global minimizer, ie:

$$f(\mathbf{x}^*) < f(\mathbf{x}^*), \text{ for all } \mathbf{x} \in \Omega \tag{7.5}$$

7.3 Optimality Conditions for Various Problems

7.3.1 Unconstrained Optimization

The gradient $\nabla f(\mathbf{x})$ is defined as:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_N} \end{bmatrix} \tag{7.6}$$

A necessary condition for a minimizer $\mathbf{x}^*$ of an unconstrained optimization problem ($\min_{\mathbf{x}} f(\mathbf{x})$) with a defined gradient $\nabla f(\mathbf{x})$ is as follows

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \tag{7.7}$$

Note that this is just a necessary condition, but it is not sufficient. For instance, a point that satisfies this zero gradient constraint may be a saddle point, or a local maximizer.

The Hessian $\nabla^2 f(\mathbf{x})$ is defined as:

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_N} \\ \vdots & & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_N \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_N^2} \end{bmatrix} \tag{7.8}$$

A sufficient set of conditions condition for a minimizer (given a twice differentiable f) is that: i) the gradient is zero (as described above), and ii) the Hessian is positive definite matrix, ie: a matrix such that:

$$\mathbf{d}^T \left[\nabla^2 f(\mathbf{x})\right] \mathbf{d} > 0 \tag{7.9}$$

for any non-zero $\mathbf{d} \in \mathbb{R}^N$.

In the one-dimensional case, these conditions reduce to: i) derivative equal to zero, and ii) second derivative greater than zero. These concepts are summarized for one-dimensional optimization (ie: where $\mathbf{x}$ is a scalar) in figure 7.1.

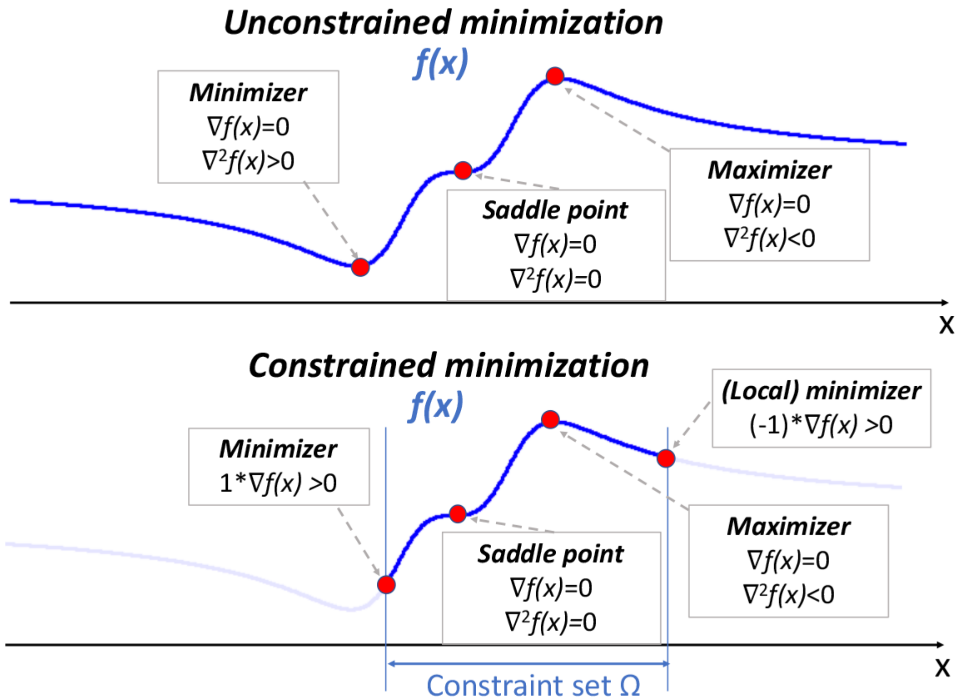

Figure 7.1: One-dimensional examples of unconstrained and constrained optimization, with various minimizers, a saddle point, and a maximizer.

7.3.2 Constrained Optimization

In the presence of constraints, the conditions above are not necessary for a local minimizer.

In this case, a necessary condition for a local minimizer is that, for any feasible direction $\mathbf{d}$ (ie: a direction in which we do not exit the feasible set Ω):

$$\mathbf{d}^T \nabla f(\mathbf{x}^*) \geq 0 \tag{7.10}$$

In other words, at a local minimizer $\mathbf{x}^*$ there is "no way but up" (or constant) within the feasible set. This condition is depicted graphically for a 1D example in figure 7.1, and for a 2D example in figure 7.2.

Similarly to the unconstrained case, a second order condition is that, for any feasible direction $\mathbf{d}$, if $\mathbf{d}^T \nabla \mathbf{x}^* = 0$, then:

$$\mathbf{d}^T \left[\nabla^2 f(\mathbf{x}^*)\right] \mathbf{d} > 0 \tag{7.11}$$

7.3.3 Brief Historical Perspective

The idea that maxima and minima can be found at locations where the gradient is zero (ie: where the tangent is flat) is quite old, and was formulated by Pierre de Fermat in his treatise entitled "Methodus ad Disquirendam Maximam et Minimam" (Fermat, 1637)[1].

[1]See "Method for the Study of Maxima and Minima", https://science.larouchepac.com/fermat/fermat-maxmin.pdf for an English translation from the original Latin

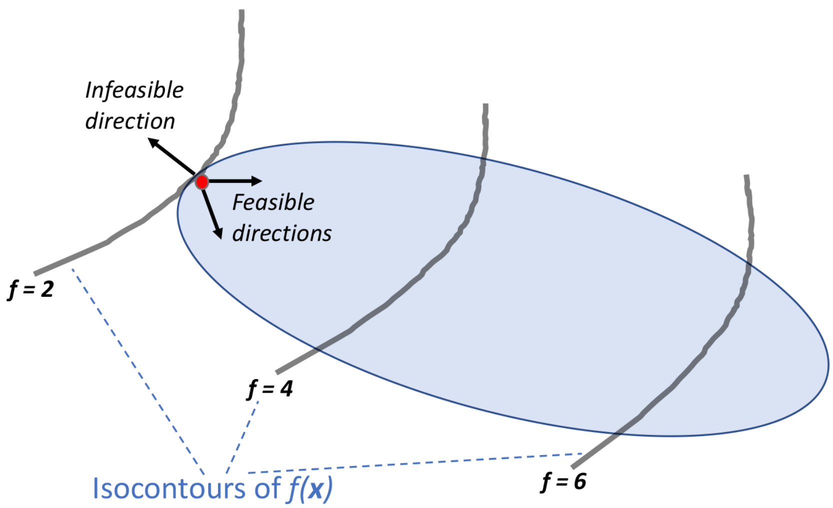

Figure 7.2: Geometric interpretation of the solution to a constrained optimization problem. In this case, the minimizer $\mathbf{x}^*$ is marked by the red dot, with a value $f(\mathbf{x}^*) = 2$.