Linear and Nonlinear Optimization

Islam S. M. Khalil

German University in Cairo

September 5, 2016

Outline

- Motivation
- Agenda
- Static optimization
- Minimum and maximum
- First order optimality conditions
- Second order sufficient conditions
- Constrained optimization
- Lagrange multiplier method
- Summary and lessons

Motivation

Definition

The goal in optimization problems is to find the minimum or maximum of a cost function $f(x_1, \ldots, x_n)$ that depends on n variables $x_1, \ldots, x_n$.

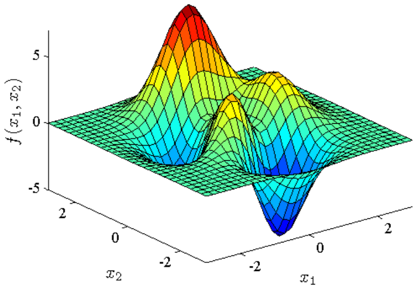

Figure: Function $f(x_1, x_2)$ with multiple extrema.

Agenda

Week	Subject
1	Static optimization
2	Algorithms for continuous optimization without constrains
3	Gradient-based techniques
4	Newton and quasi-Newton techniques
5	Algorithms for continuous optimization with constrains
6	Linear problems: simplex and primal-dual interior point
7	Quadratic problems: active-set and interior point
8	Convex optimization: the concept duality and algorithms
9	General non-linear optimization
10	A number of special optimization problems
11	Design of an electromagnetic system
12	Ongoing research examples and problems

Minimum and Maximum

The function $f(x_1, \ldots, x_n)$ is said to have a *global minimum* at the point $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}$ if

$$f(x_1, \ldots, x_n) \leq f(x_1 + h_1, \ldots, x_n + h_n), \tag{1}$$

for all $\mathbf{h}^{\mathrm{T}} = \begin{bmatrix} h_1 & h_2 & \ldots & h_n \end{bmatrix} \neq 0$.

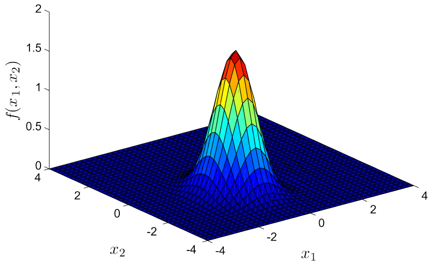

Figure: Function $f(x_1, x_2)$ with a *global maximum*.

Minimum and Maximum

The point $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$ is said to be *local minimum* if the condition (1) holds for all $\mathbf{h}$ such that

$$\|\mathbf{h}\|_2 = (h_1^2 + h_2^2 + \ldots + h_n^2)^{1/2} < \rho, \qquad (2)$$

where $\rho > 0$.

The points at which the function $f(x_1, \ldots, x_n)$ becomes minimum or maximum are the extrema of the function. Consider the Taylor series expansion of $f(x_1 + h_1, \ldots, x_n + h_n)$ about $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}$

$$f(x_1 + h_1, \ldots, x_n + h_n) = f(x_1, \ldots, x_n) + \partial f(x_1, \ldots, x_n)$$

$$+\partial^2 f(x_1, \ldots, x_n) + \text{higher-order terms},$$

where,

$$\partial f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \frac{\partial f(x_1, \ldots, x_n)}{\partial x_i} h_i,$$

$$\partial^2 f(x_1, \ldots, x_n) = \frac{1}{2} \begin{bmatrix} h_1 & \ldots & h_n \end{bmatrix} \mathbf{H}(x_1, \ldots, x_n) \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix},$$

where $\mathbf{H}(x_1, \ldots, x_n)$ is the Hessian matrix.

The Hessian matrix is given by

$$\mathbf{H}(x_1,\ldots,x_n)=\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{bmatrix}. \tag{3}$$

The first order optimality conditions for $x_1,\ldots,x_n$ to be a minimum of $f(x_1,\ldots,x_n)$

$$\frac{\partial f(x_1,\ldots,x_n)}{\partial x_i}=0 \text{ for } i=1,\ldots,n. \tag{4}$$

Example 1
Find the extremum of the function $f_1(x_1, x_2, x_3) = x_1 + 2x_3 + x_2x_3 - x_1^2 - x_2^2 - x_3^2$.

The necessary conditions for extremum are:

$$\begin{aligned} \frac{\partial f_1}{\partial x_1} &= 1 - 2x_1 = 0 \\ \frac{\partial f_1}{\partial x_2} &= x_3 - 2x_2 = 0 \\ \frac{\partial f_1}{\partial x_3} &= 2 + x_2 - 2x_3 = 0 \end{aligned}$$

on simultaneously solving these three equations, we can obtain the $x_1 = 1/2$, $x_2 = 2/3$, and $x_3 = 4/3$.

First Order Optimality Conditions

Example 2

Find the extremum of the function $f_2(x_1, x_2, x_3) = 8x_1x_2 + 3x_2^2$.

The necessary conditions for extremum are:

$$
\begin{aligned}
\frac{\partial f_1}{\partial x_1} &= 8x_2 = 0 \\
\frac{\partial f_1}{\partial x_2} &= 8x_1 + 6x_2 = 0
\end{aligned}
$$

on simultaneously solving these three equations, we can obtain the $x_1 = 0$ and $x_2 = 0$.

Second Order Sufficient Conditions

Assume that $x_1, \ldots, x_n$ is a minimum, then using (1) and the first order necessary condition (4) yields

$$\partial^2 f(x_1, \ldots, x_n) = \frac{1}{2} \begin{bmatrix} h_1 & \ldots & h_n \end{bmatrix} \mathbf{H}(x_1, \ldots, x_n) \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} > 0. \quad (5)$$

Any point $x_1, \ldots, x_n$ that satisfies both (4) and (5) must be at least a local minimum of the function $f(x_1, \ldots, x_n)$.

Second Order Sufficient Conditions

- $\mathbf{h}^{\mathrm{T}}\mathbf{Hh}$ (or $\mathbf{H}$) is positive definite iff: $\mathbf{h} \cdot \mathbf{Hh} > 0$ for all $\mathbf{h} \neq 0$,
  or, all Re(λ_i)> 0.
- $\mathbf{h}^{\mathrm{T}}\mathbf{Hh}$ (or $\mathbf{H}$) is semipositive definite iff: $\mathbf{h} \cdot \mathbf{Hh} \geq 0$ for all $\mathbf{h} \neq 0$,
  or, Re(λ_i)≥ 0.
- $\mathbf{h}^{\mathrm{T}}\mathbf{Hh}$ (or $\mathbf{H}$) is negative definite iff: $\mathbf{h} \cdot \mathbf{Hh} < 0$ for all $\mathbf{h} \neq 0$,
  or, all Re(λ_i)< 0.
- $\mathbf{h}^{\mathrm{T}}\mathbf{Hh}$ (or $\mathbf{H}$) is seminegative definite iff: $\mathbf{h} \cdot \mathbf{Hh} \leq 0$ for all $\mathbf{h} \neq 0$,
  or, Re(λ_i)≤ 0.

Second Order Sufficient Conditions

- $\mathbf{h}^{\mathrm{T}}\mathbf{H}\mathbf{h}$ (or $\mathbf{H}$) is indefinite iff: $\mathbf{h} \cdot \mathbf{H}\mathbf{h} \geq 0$ for some $\mathbf{h} \neq 0$, and $\mathbf{h} \cdot \mathbf{H}\mathbf{h} \leq 0$ for other $\mathbf{h} \neq 0$, or some Re$(\lambda_i)> 0$ and other Re$(\lambda_i)< 0$.

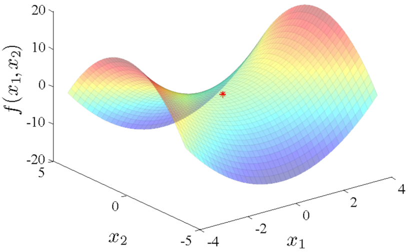

Figure: Point (0,0) is a saddle point, and is the optimal solution of the following problem: $\min_{x_1 \in \mathbb{R}} \max_{x_2 \in \mathbb{R}} \{x_1^2 - x_2^2\}$.

Example 3
Determine if the extremum in Example 1 is a maximum, a minimum, or a saddle point.

The Hessian matric is given by

$$\mathbf{H}(x_1, x_2, x_3) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix},$$

the eigenvalues of the Hessian matrix are -1, -2, and -3. Therefore, the extremum is a maximum.

Thank You

Thank You!
Questions please