Optimization

Fourth Class

2020 - 2021

By

Dr. Jawad Mahmoud Jassim

Dept. of Math.

Education College for Pure Sciences

University of Basrah

Iraq

Chapter Two

Line Search

Lecture 1

1: Introduction:

Line search, also called one-dimensional search, refers to an optimization procedure for univariate functions. It is the base of multivariate optimization.

The iterative scheme is

$$X_{k+1} = X_k + \alpha_k d_k \qquad \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots .. (1)$$

Where X_k is given.

The our aim is to find the direction vector d_k and a suitable step size α_k .

Let $\Phi(\alpha) = f(X_k + \alpha_k d_k) \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \; (2)$

<u>Definition (1): (Line search)</u>

The process of finding α_k by estimating a minimizer of Φ is called the line search. **So, the problem that departs from X_k and find a step size in the direction d_k such that $\Phi(\alpha_k) < \Phi(0)$ is just line search about α .**

<u>Definition (2): (Exact Line Search)</u>

If we find α_k such that the objective function in the direction d_k is minimized, i.e.

$$f(X_k + \alpha_k d_k) = \min_{\alpha>0} f(X_k + \alpha d_k) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (3)$$

Or $$\Phi(\alpha_k) = \min_{\alpha>0} \Phi(\alpha) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Such a line search is called exact line search.

<u>Note (1):</u>

Exact line search is also called optimal line search and α_k is called optimal step size.

Definition (3): (Inexact Line Search)

If we choose α_k such that the objective function has acceptable descent amount, i.e. such that the descent

$$f(X_k + \alpha_k d_k) < f(X_k) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

is acceptable by users, such a line search is called inexact line search.

Note (2):

Inexact line search is also called approximation line search or acceptable line search.

The framework of the line search is as follows:

1: determine or give an initial search interval which contains the minimizer.

2: Employ some section techniques or interpolations to reduce the interval iteratively until the length of interval is less than some given tolerance.

Definition (4): (Search Interval)

Let $\Phi: R \rightarrow R$, $\alpha^* \epsilon [0, \infty)$ and $\Phi(\alpha^*) = \min_{\alpha>0} \Phi(\alpha)$. If there exists a closed interval $[a, b] \subset [0, \infty)$ such that $\alpha^* \epsilon [a, b]$, then $[a, b]$ is called a search interval for one – dimensional minimization problem $\min_{\alpha>0} \Phi(\alpha)$.

Note (3):

Since the exact location of minimum of Φ over $[a, b]$ is not known, the search interval is also called the interval of uncertainly.

2: Forward – Backward Method

A simple method to determine an initial interval is called the forward – backward method.

The basic idea of this method is as follows:

Give an initial point and an initial step length, we attempt to determine three points at which three function values show "high – low – high" geometry. If it is not successful to go forward, we will go backward. Concretely, give an initial point α_0 and a step $h_0 > 0$.

$If\ \Phi(\alpha_0 + h_0) < \Phi(\alpha_0)\ then$, **next step, depart from α_0 and go backward until the objective function increases. So, we will obtain an initial interval which contains $the\ minimum\ \alpha^*$.**

Algorithm 1: (Forward – Backward Method)

Step 1:

Give $\alpha_0 \epsilon [0, \infty), h_0 > 0$, the multiple coefficient $t > 1$ ($usually\ t = 2$). Evaluate $\Phi(\alpha_0)$.

Step 2:

$Set\ \alpha_{k+1} = \alpha_k + h_k$, and evaluate $\Phi_{k+1} = \Phi(\alpha_{k+1})$. If $\Phi_{k+1} < \Phi_k$, go to step 3, otherwise go to step 4.

Step 3: (Forward Step)

Set $h_{k+1} = t\, h_k$, $\alpha = \alpha_k$, $\alpha_k = \alpha_{k+1}$, $\Phi_k = \Phi_{k+1}$, $k = k + 1$, go to step 2.

Step 4: (Backward Step)

If $k = 0$, invert the search direction. Set $h_k = -h_k$, $\alpha_k = \alpha_{k+1}$, go to step2, otherwise set $a = min\{\alpha, \alpha_{k+1}\}$, $b = max\{\alpha, \alpha_{k+1}\}$, output $[a, b]$ and stop.

Definition (5): (Unimodal Function)

Let $\Phi : R \rightarrow R$, $[a, b] \subset R$. Let $\alpha^* \epsilon [a, b]$ is a minimizer of $\Phi(\alpha)$ .If $\Phi(\alpha)$ is strictly decreasing on $[a, \alpha^*]$ and strictly increasing on $[\alpha^*, b]$, then $\Phi(\alpha)$ is called a *unimodal function on* $[a, b]$. Such an interval $[a, b]$ is called a unimodal interval related to $\Phi(\alpha)$.

There is another definition for unimodal function, given below.

Definition (6): (Unimodal Function)

Let $\Phi : R \rightarrow R$, $[a, b] \subset R$. Let $\alpha^* \epsilon [a, b]$ is a unique minimizer of $\Phi(\alpha)$.

Then $\Phi(\alpha)$is called a *unimodal function on* $[a, b]$ *if and only if any two points* α_1 *and* α_2

$$\alpha^* < \alpha_1 < \alpha_2 \text{ , then } \Phi(\alpha^*) < \Phi(\alpha_1) < \Phi(\alpha_2), \text{ and}$$

$$\alpha^* > \alpha_1 > \alpha_2, \text{ then } \Phi(\alpha^*) < \Phi(\alpha_1) < \Phi(\alpha_2).$$

Note (4):

In some textbooks a function which is strictly increasing on $[a, \alpha^*]$and strictly decreasing on $[\alpha^*, b]$ is also called a unimodal function. Of course in this case α^* *is a maximizer*.

<u>Note (5):</u>

A sufficient condition for a function $\boldsymbol{\Phi}$ to be unimodal on the interval $[\boldsymbol{a}, \boldsymbol{b}]$ is that $\boldsymbol{\Phi}$ strictly convex function on $[\boldsymbol{a}, \boldsymbol{b}]$.

$\boldsymbol{But\ a\ unimodal\ function\ \Phi\ is\ NOT\ necessarily\ a\ strictly\ convex}$.

<u>Note (6):</u>

The unimodal function does not require the continuity and differentiability of the function.

<u>Note (7):</u>

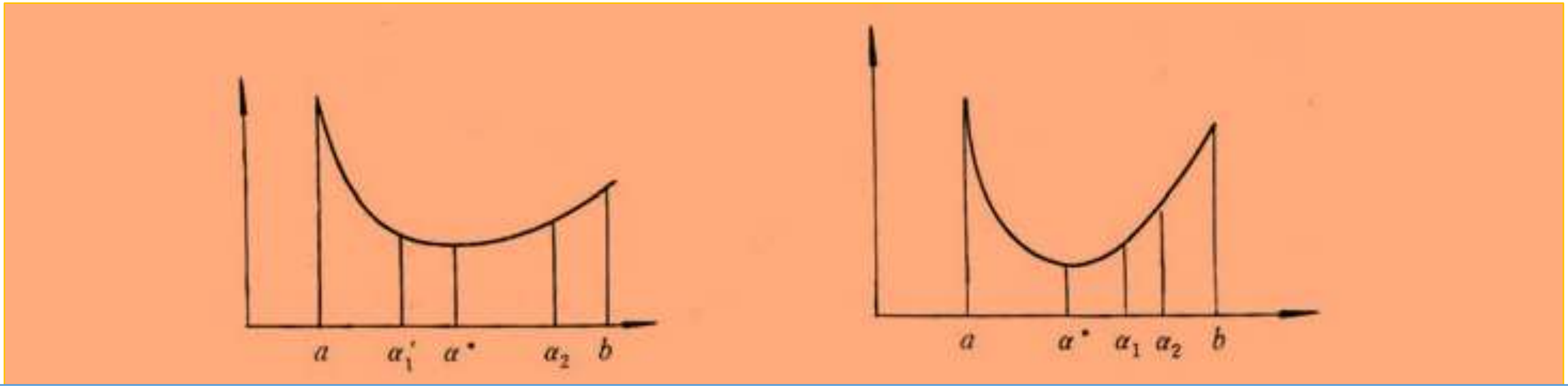
Using the property of the unimodal function, we can exclude portions of the interval of uncertainly that do not contain the minimum such that the interval of uncertainly is reduced.

Theorem (1):

Let $\Phi : R \to R$ be unimodal function on $[a, b]$. Let $\alpha_1, \alpha_2 \epsilon [a, b], \alpha_1 < \alpha_2$. Then:

1: $if\ \Phi(\alpha_1) \leq \Phi(\alpha_2)$, then $[a, \alpha_2]$ is a unimodal interval related to Φ.

2: $if\ \Phi(\alpha_1) \geq \Phi(\alpha_2)$, then $[\alpha_1, b]$ is a unimodal interval related to Φ.



Properties of unimodal interval and unimodal function