Computational Tutorial of Steepest Descent Method and Its Implementation in Digital Image Processing

Vorapoj Patanavijit
Department of Electrical and Electronic Engineering, Faculty of Engineering
Assumption University, Bangkok, Thailand
E-mail: <Patanavijit@yahoo.com>

Abstract

In the last decade, optimization techniques have extensively come up as one of principal signal processing techniques, which are used for solving many previous intractable problems in both digital signal processing (DSP) problems and digital image processing (DIP) problem. Due to its low computational complexity and uncomplicated implementation, the Gradient Descent (GD) method [1] is one of the most popular optimization methods for problems, which can be formulated as a differentiable multivariable functions. The GD method is ubiquitously used from basic to advanced researches. First, this paper presents the concept of GD method and its implementations for general mathematical problems. Next, the computation of GD processes is shown step by step with the aim to understand the effect of important parameters (such as its initial value and step size) to the performance of GD. Later, the computational concept of GD method for DIP problems [2-5] is formulated and the computation of GD is demonstrated step by step. The effect of the initial value and the step size to the performance of GD method in DIP is also presented.

Keywords: Gradient Descent (GD) method, Digital Image Processing and Digital Signal Processing.

1. General Introduction of Gradient Descent (GD) Method [1]

The estimation of the multivariable $\mathbf{x} = \left[x_1, x_2, \ldots, x_m\right]^{\mathrm{T}}$ that minimize the error or the cost function, $F(\mathbf{x})$, is often found in both DSP and DIP problems. When $F(\mathbf{x})$ is Differentiable, $\mathbf{x}$ can be estimated simply by

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \lambda_n \cdot \nabla F(\mathbf{x}_n) \tag{1.1}$$

where

- $\mathbf{x}_n$ is the estimated multivariable at n^{th} iteration.
- $\nabla F(\mathbf{x})$ is the gradient of function $F(\mathbf{x})$ that can be determined as

$$\nabla F(\mathbf{x}) = \nabla F\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}\right) = \begin{bmatrix} \dfrac{\partial F(\mathbf{x})}{\partial x_1} & \dfrac{\partial F(\mathbf{x})}{\partial x_2} & \cdots & \dfrac{\partial F(\mathbf{x})}{\partial x_m} \end{bmatrix}^{\mathrm{T}} \tag{1.2}$$

- λ_n is the step size of GD method at the n^{th} iterative.

The parameter $\mathbf{x}_0$ (so call the initial value) and λ_n in Eq. (1.1) directly influence the performance of GD as shown in the following section.