Journal of Computational and Applied Mathematics 233 (2009) 1128–1138

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Application of gradient descent method to the sedimentary grain-size distribution fitting

Haibo Chen [a], Xianqing Lv [a,*], Yansong Qiao [b]

[a] Laboratory of Physical Oceanography, Ocean University of China, Qingdao 266100, China

[b] Institute of Geomechanics, Chinese Academy of Geological Sciences, Beijing 100081, China

ARTICLE INFO

Article history:
Received 2 April 2009
Received in revised form 4 September 2009

MSC:
65D10
62J02
90C31
62P12
93E24

Keywords:
Nonlinear least squares data fitting
Gradient descent
Mixture distribution of three lognormal components
Laser grain-size
Existence theorem

ABSTRACT

Existence of a least squares solution for a sum of several weighted normal functions is proved. The gradient descent (GD) method is used to fit the measured data (i.e. the laser grain-size distribution of the sediments) with a sum of three weighted lognormal functions. The numerical results indicate that the GD method is not only easy to operate but also could effectively optimize the parameters of the fitting function with the error decreasing steadily. Meanwhile the overall fitting results are satisfactory. As a new way of data fitting, the GD method could also be used to solve other optimization problems.

1. Introduction

Sediments such as loess, ice cores, deep-sea soil, etc. which could mirror the changes of the global climate since the quaternary, are considered as good information carriers of the changes of the global environment. The grain-size, as a main feature of the sedimentary deposition and a mature paleoenvironmental index of sediments, is an important parameter for the analysis of the depositional environments, the process and the mechanism of transportation and deposition, and has been widely applied to the study of various depositional environments over the last several years due to its convenient and expedite measurement, definite physical meaning, sensitivity to the changes of the climate, etc. Each component of the sedimentary grain-size distribution has different causes of formation since it is subject to the provenance and the changing climate. As a result, the changes of the proportion of different grain-size fractions, which constitute the components of the sedimentary grain-size distribution, have different paleoclimatic meanings. By now, the analysis of grain-size has been applied to the study of various depositional environments such as loess, paleosol, glacial deposits, ocean, river, lake, etc. Thus how to make good use of the information on the measured data of each grain-size fraction is significant for accurate identification of the types of the depositional environments of different sediments [1–6].

As most sediments in the nature are controlled by one or more different modes of transportation and different dynamics, the grain-size distribution will characterize multi-component and multi-modal with a multi-peak smooth frequency curve

* Corresponding author.
E-mail address: xqinglv@ouc.edu.cn (X. Lv).

0377-0427/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2009.09.005

(see [1–6]). The morphological characteristics of the grain-size distribution curve with several peaks show that every component is actually a lognormal distribution, which could rationally describe the characteristics of the distribution curve and the causes of the formation of each component of the sample and its parameters also have definite physical meanings. Therefore, in this paper the mixture distribution of several lognormal components is chosen to fit the sedimentary grain-size distribution while every component is partitioned mathematically at the same time.

Sun et al. have ever studied a lot about grain-size distribution fitting with a mixture distribution of several components. However, maybe they emphasized particularly on the paleoenvironmental meanings of the fitting results, so in their studies only part of fitting results were shown and the procedure of fitting was also rarely mentioned. Meanwhile we believe that the accuracy of their fitting is still expected to be improved.

The probability density function for a mixture distribution of several weighted lognormal components is given by

$$f(t; a_i, \delta_i, c_i) = \sum_{i=1}^{I} c_i \frac{1}{\sqrt{2\pi}\,\delta_i} \exp\left[-\frac{(\ln t - a_i)^2}{2\delta_i^2}\right],$$

where x denotes the grain-size value, parameters a_i and δ_i represent, respectively, the location and shape parameters of the component i with a total number I, and its percentage in total distribution is given by c_i. Through the natural logarithmic transformation of independent variable $x = \ln t$, we can express the lognormal distributions as the normal distributions for convenience:

$$f(x; a_i, \delta_i, c_i) = \sum_{i=1}^{I} c_i \frac{1}{\sqrt{2\pi}\,\delta_i} \exp\left[-\frac{(x - a_i)^2}{2\delta_i^2}\right].$$

To fit the grain-size distribution with a sum of several normal functions is indeed to estimate the parameters a_i, δ_i and c_i. One canonical method for parameter estimation is the least squares (LS) approximation, and numerical methods for solving the nonlinear LS problem are described in [7–9]. Before the iterative minimization of the sum of squares it is still necessary to answer the two questions below beforehand:

(1) Does the solution (i.e. the LSE described in the following section) to a certain LS problem exist?
(2) How to determine a possible good initial approximation $(a_{i0}, \delta_{i0}, c_{i0})$?

In the case of nonlinear LS problems it is still extremely difficult to answer the first question. Some theorems about the existence of the LSE for some special function can be found, e.g., in [10–17]. In [10] Jukić and Scitovski have given the existence theorem of LS fitting Gaussian type curve. Actually, the existence theorem of the LSE for the sum of several normal functions, which are indeed a series of special Gaussian functions, could be readily proved with the given condition in this paper in the similar way as Jukić and Scitovski did in [10], and the details will be given in Section 3.

There are many methods developed for the initial approximation and the choice is dependent on both the object of study and the main optimization method. To answer the second question, some authors have given valuable proposals (see [10–15,18,19]). In [18] Hölmstrom and Petersson used a modified Proney method for the initial approximation for a sum of exponential functions; Jukić and Marković et al. [10–15] have proposed several practical initial value techniques for some special functions; Hasdorff [19] derived an expression for initial approximation from gradient. In this paper we will determine the initial approximation by linearizing the nonlinear LS problem.

There are various numerical methods for the nonlinear LS problem (see, e.g., [7–9]) and the favorable choice still depends. One essential suggestion is the Gauss–Newton (GN) method, which has been applied to nonlinear data fitting in many studies (see, e.g., [10,13,18,20]). In addition Hölmstrom and Petersson also presented some other methods such as the quasi-Newton method, Levenberg–Marquardt (LM) method, etc in [18]. Atieg and Watson [20] made a summary and comparison of some methods which mainly include Newton method and GN method. Böckmann [21] proposed a modification of the trust-region Gauss–Newton method with a two-parameter approximation for a sum of weighted functions. Nyarko and Scitovski [22] gave a method for solving the parameter identification problem for ordinary second order differential equations using genetic algorithms. Ahn, Rauh and Warnecke [23] proposed some simple and robust nonparametric algorithms for the geometric fitting of circle/sphere/ellipse/hyperbola/parabola. Furthermore, in [18,20] the fitting function is a sum of weighted functions, which is very useful to our study here. The GD method, which takes the negative gradient direction as the search direction of the minimization algorithm, is the simplest one of the unconstrained optimization algorithms that require calculation of derivatives, and the discussion on the application of GD method to nonlinear data fitting is of great benefit.

In this paper, we present the proof of the existence of the LS solution for a sum of I weighted normal functions and our effort to fit the laser grain-size analytical data of Ganzi loess as an example with the mixture distribution of three lognormal components, and the principal objective is to give an introduction of the application of GD method to the grain-size distribution fitting and to show its effectiveness and feasibility. In Section 2 the problem of data fitting is described as the one of nonlinear LS for a sum of I weighted normal functions. In Section 3 we present the existence theorem (Theorem 3.1) for a sum of I weighted normal functions as well as their proofs. In the remainder of this paper we take $I = 3$ to show our fitting methods. In Section 4 we use the linearization method for the initial approximation of the parameters. Note that the choice is arbitrary and it is possible to take another method. Here we present the linearization method with the following two purposes: the first principal purpose is to provide a possible good initial approximation for GD method, and the other is