Available online at www.sciencedirect.com

SCIENCE DIRECT®

International Journal of Rock Mechanics & Mining Sciences 43 (2006) 877–893

International Journal of
Rock Mechanics
and Mining Sciences

www.elsevier.com/locate/ijrmms

Inference of discontinuity trace length distributions using statistical graphical models

R. Jimenez-Rodriguez*, N. Sitar

Department of Civil and Environmental Engineering, University of California at Berkeley, CA-94720-1710 USA

Accepted 16 December 2005
Available online 17 February 2006

Abstract

The characterization of discontinuities within rock masses is often accomplished using stochastic discontinuity network models, in which the stochastic nature of the discontinuity network is represented by means of statistical distributions. We present a flexible methodology for maximum likelihood inference of the distribution of discontinuity trace lengths based on observations at rock outcrops. The inference problem is formulated using statistical graphical models and target distributions with several Gaussian mixture components. We use the Expectation–Maximization algorithm to exploit the relations of conditional independence between variables in the maximum likelihood estimation problem. Initial results using artificially generated discontinuity traces show that the method has good inference capabilities, and inferred trace length distributions closely reproduce those used for generation. In addition, the convergence of the algorithm is shown to be fast.

Keywords: EM-algorithm; Fracture network; Discontinuity size; Maximum likelihood

1. Introduction

Discontinuities have a significant impact on the deformability, strength, and permeability of rock masses [1,2]; consequently, their characterization is an important element of rock mass characterization [1–5]. However, deterministic characterization of individual discontinuities in the rock mass is usually an insurmountable site characterization challenge and, in general, it is only feasible for major features. The characterization of other (i.e., not major) discontinuities can be, on the other hand, accomplished using stochastic discontinuity network models. In these models, the rock mass is represented as an assemblage of discontinuities intersecting a volume of intact rock and the stochastic nature of the discontinuity network is represented using statistical distributions [6,7]; in some cases, additional non-geometrical aspects are included to imitate geological processes leading to the formation of discontinuities within the rock mass [5,8].

To be able to use stochastic discontinuity network models in engineering applications, the problem of calibration of network parameters remains, and we need methods for the characterization of discontinuity networks based on information available at the design stage. In particular, one of the main difficulties in estimating discontinuity sizes is the fact that direct observation of their complete three dimensional extent is not possible. As a result, the distribution of discontinuity dimensions is commonly inferred using information about the distribution of trace lengths at rock exposures by means of, for example, stereological or fractal considerations [9–15].

Hence, proper characterization of the distribution of trace lengths is an essential step in the characterization of the distribution of discontinuity dimensions. There are two additional difficulties in the solution of the problem of estimating the trace length distribution: The first is that the observations of discontinuity traces are biased [15–20]; and the second is due to the complexity of the geological processes leading to the development of rock discontinuities

*Corresponding author. Present address: ETS Ing. de Caminos, Canales y Puertos. Universidad Politécnica de Madrid, Spain.
Tel.: +34 91 336 6710; fax: +34 91 336 6774.
E-mail address: rjimenez@caminos.upm.es (R. Jimenez-Rodriguez).

1365-1609/$ - see front matter
doi:10.1016/j.ijrmms.2005.12.008

[11]. Such complexity is responsible for uncertainties about the most adequate type of distribution to be used in real applications. Exponential and lognormal distributions are most commonly employed (see e.g., [21,22]), but there are cases in which bi-modal types of distributions—which cannot be properly described by the distributions mentioned above—seem to be more adequate, as suggested by case histories or geomechanical modeling results (see e.g., [18,23,24]).

To deal with these problems, the measured (i.e., biased) trace length distribution is commonly estimated first; then, it is assumed that the real (i.e., unbiased) trace length distribution has the same distribution form, so that only the parameters of the distribution (usually the mean and variance) need to be obtained [7,21]. Several methods for the estimation of the mean of the real distribution of trace lengths have been proposed [17,19,25,26]; the variance can be estimated by assuming that the values of the variances of the observed and real trace length distributions are equal [7] or, better yet, by assuming that the coefficients of variation of both distributions are equal [21]. Additional statistical approaches have been used to estimate the distribution of discontinuity trace lengths: Song and Lee [20] estimated trace lengths using areal sampling and probabilistic relations derived for the Poisson disk model; and maximum likelihood estimation methods for scanline or areal sampling have been proposed as well [18,23,27–29].

In this work, we present a novel approach based on the use of statistical graphical models for maximum likelihood inference of the real (i.e., corrected for biases) distribution of discontinuity trace lengths [30,31]. The identification of the type or structure of the trace length distribution is not our main interest; rather, we are interested in working with a model that provides reasonable estimates of the probability distribution without making strong assumptions about its type. Accordingly, we avoid the assumption that the measured and real trace lengths distributions are of the same type, and we approach the inference problem by considering a broad family of target mixture distribution models. That is, we work with a *target* distribution that is flexible enough to mimic the main features of the *real* (and unknown) distribution of trace lengths, allowing the observed data to ''select'' the most adequate distribution for each case.

We believe that this type of model based on a statistical analysis of observed discontinuity trace data will gain even more significance in the years to come, as traditional methods for discontinuity surveying in rock engineering [32] are being replaced by automated techniques, which allow more efficient and detailed acquisition of discontinuity data [33–35].

2. Generation and sampling of discontinuities

We assume that the rock outcrop is a planar surface and that the *sampling domain* (where traces are observed) is a rectangular region within the rock outcrop of dimensions $W_o \times H_o$. We further assume that the sampling domain is contained within the *generation domain*, where we use Monte Carlo simulation to generate populations of discontinuity traces. We consider the size of the generation domain to be ''much larger'' than the size of the sampling domain and ''much larger'' than the length of generated traces, so that the consequences of biases, edge effects, or both are negligible in the generation process. (For a detailed analysis of the influence of stochastic network parameters in the occurrence of edge effects, see [36].) We also assume that discontinuities are parallel and flat circular disks of negligible thickness, with their centers uniformly distributed in space (i.e., the Poisson disk model [37]); accordingly, discontinuity traces are parallel straight lines of negligible width, with centers uniformly located within the generation domain.

The Poisson disk model has been extensively used in rock mechanics applications (see e.g., [5,6,9,20,21,29, 36–41]). Here, we use the Poisson disk model because it has been found to generate fractures and fracture traces that are similar to natural fracture patterns in many cases and because it has been recognized that some fracture systems are best described by this type of models [5,11]. Additional advantage of the Poisson disk model is that it is simple and easy to program [5]; it is also mathematically convenient [14], allowing simple derivations of analytical expressions.

In other cases, however, characteristics of fracture systems in rock masses are best described by power laws and fractal geometry [11,12], and fractals have been widely used to describe fracture geometry (see e.g., [42–46]). (For further arguments in favor of fractals, see [5,11].) Therefore, the method proposed herein should not be applied to rock masses with a demonstrable fractal system without due consideration to the errors that may be introduced.

Fig. 1 illustrates the types trace maps, sampling domains, and generation domains used in this work.

3. Existing biases

3.1. Description of types of bias

Observations of discontinuity traces at rock exposures are subjected to orientation, truncation, censoring and size bias (e.g., [16–21]). The terms *curtailment* and *trimming* have also been used to refer to censoring and truncation [15], but we chose to use the former terminology for overall consistency with the rock mechanics literature.

In this work, we consider the commonly used model of a single set of parallel traces (see e.g., [15,20,25,26,29,47,48]) and, therefore, orientation bias does not affect the results discussed herein. Methods for correction of orientation bias—including the use of circular domains—are quite common [16,21,49–51], and extensions to consider trace data with variable orientation are also available [23,52]. Similarly, truncation bias is not significant in the context of block formation, since the truncation threshold may be