Dimension reduction in a model of random set given by union of interacting discs

Jakub Staněk [1]

joint work with Kateřina Staňková Helisová [2]

[1] Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 18675 Prague 8, Czech Republic
Email: stanekj@karlin.mff.cuni.cz

[2] Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Mathematics, Technická 2, 16627 Prague 6, Czech Republic
Email: helisova@math.feld.cvut.cz

Abstract

Consider a planar random set observed in a bounded window $W \subset \mathbf{R}^2$. The set is given by a union of randomly scattered interacting discs with random radii and described by a density $f_\theta(\mathbf{x})$ of any finite configuration $\mathbf{x} = (x_1, ..., x_n)$ of the discs $x_1, ..., x_n$ with respect to the probability measure of a stationary random-disc Boolean model. The density is of the form

$$f_\theta(\mathbf{x}) = \frac{\exp\{\theta \cdot T(U_\mathbf{x})\}}{c_\theta},$$

where $T(U_\mathbf{x})$ is a k-dimensional vector of geometrical characteristics of the union $U_\mathbf{x}$ composed of the discs from the configuration $\mathbf{x}$, $\theta = (\theta_1, \ldots, \theta_k)$ is a vector of parameters, $\cdot$ denotes the inner product and c_θ is a normalizing constant.

This process with $\theta = (\theta_1, \ldots, \theta_6)$ and $T = (A, L, \chi, N_h, N_{id}, N_{bv})$, where A denotes the area, L the perimeter, χ the Euler-Poincare characteristic, N_h the number of holes, N_{id} the number of isolated discs and N_{bv} the number of boundary vertices was studied in [1] and consequently fitted to real data in [2], where the parameters were estimated by maximum likelihood method using MCMC simulations (MCMC MLE). However, this method appeared to be very time-consuming, mainly because of looking for the best estimate $\hat{\theta}$ of the parameter θ in the space of high dimension k.

This contribution concerns different methods for reduction of dimension of the vector T (and accordingly of the vector θ) and consequent statistical inference of the reduced model. More precisely, it will be described

how to use main components method and sliced inverse regression method with different ways of slicing data to obtain a new vector $\tilde{T}$, which supports the same information about the set as the vector T, but its dimension is lower. Then, estimation of new parameters by MCMC MLE and checking the accordance of the model to data will be shown.

References

[1] Møller J., Helisová K. (2008): *Power diagrams and interaction processes for unions of discs.* Advances in Applied Probability **40**(2), 321–347.

[2] Møller J., Helisová K. (2010): *Likelihood inference for unions of interacting discs.* Scandinavian Journal of Statistics **37**(3), 365–381.

[3] Li K.-C. (1991): *Sliced inverse regression for dimension reduction.* JASA **86**(3), 316–327.

[4] Rencher A.C. (2002): Methods of Multivariate Analysis, 2nd edn. Wiley & Sons, New York.