



On the scale free laws of urban facilities

Liang Wu, Yang Li, Xuezheng Chen*

School of Economics, Sichuan University, 29 Wangjiang Rd., Chengdu, Sichuan 610065, PR China



ARTICLE INFO

Article history:

Received 11 August 2015
 Received in revised form 12 October 2015
 Accepted 13 October 2015
 Available online 30 October 2015
 Communicated by C.R. Doering

Keywords:

Urban facilities
 Double stochastic model
 Power law
 Covariance function
 Model check

ABSTRACT

We implement a double stochastic process as the mathematical model for the spatial point patterns of urban facilities. We find that the model with power covariance function can produce the best fit not only to K function (whose derivative gives the radial distribution $\rho(t) = K'(t)/2\pi t$) but also to additional facts of spatial point patterns. These facts include the mean-variance relationship of number of events in a series of expanding bins, and other statistics beyond the first two orders, such as inter-event distribution function $H(t)$ and nearest neighbor distribution functions $G(t)$ and $F(t)$.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The rapid urbanization becomes one of the predominant processes in the human history. It arouses much interest to study cities and urban lives in the scientific communities [1,2]. Like many other physical systems, despite the complex underlying structure, macro statistical regularities such as power laws emerge [1,3]. With the increasing ability to collect the spatial coordinate data of facilities and buildings from electronic map providers, researchers begin to study the spatial substructure of cities. It is reported in [4] that K -function (whose derivative gives the radial distribution function $\rho(t) = K'(t)/2\pi t$) is a power function. Besides, the mean-variance relationship of number of events in a series of expanding bins is also a power function. These empirical results can be embedded in a double stochastic process model when the covariance function is a power function.

Scientific interest in power law relations stems partly from the possibility that the power function might point to a deep origin in the dynamical process that generates the power law relation. Unfortunately, the detection and characterization of power laws are complicated by the large fluctuations that occur in the tail of the distribution and by the difficulty of identifying the range over which power law behavior holds [5]. The fact that actual physical systems are finite also hinders the test of power laws which characterizes long range interactions. As [6] points out, one should not only consider a detailed mechanism in driving dynamics, but also

the extent of statistical support for a reported power law. The linear relationship on a log–log plot is clearly established for urban facilities in [4] over more than 2 orders of magnitude, however the considered statistics K -function and mean-variance relationship only capture the statistical properties of the first two orders. They do not give a complete picture. As an illustration of the insufficiency of second-order statistics, Ref. [7] describes a class of non-Poisson processes however for which $K(t) = \pi t^2$ coincides with the Poisson random process. Besides, the empirical evidence from K -function is not strong enough to support the power laws. Except for the power function, there are many other functions that can give good fit to $K(t)$. The power law rule would be on a more solid basis if the model estimated from the first and the second order statistics fits other statistics including higher order statistics.

Higher order statistics can be defined in terms of the joint intensity functions for the occurrence of specified configurations of three, four, etc. events. Interpretation would be difficult in practice since, for example, the third-order intensity functions of a stationary, isotropic process requires three arguments, the fourth-order function, five, and so on. There are several distribution functions which are easy to interpret. They can serve as additional statistics summary of spatial point process. These are $H(t)$ the distribution function of inter-event distance, $G(t)$ the distribution function of the distance from an arbitrary event to its nearest other event, and $F(t)$, the distribution function of the distance from an arbitrary point to its nearest event [8].

Except these statistics, mean-variance relationship can also serve as a statistical method to compare models. Denote M and V as the average and variance of number of events in a series of expanding bins, respectively. As shown in [4], the power law rule of

* Corresponding author.

E-mail addresses: liangwu@scu.edu.cn (L. Wu), xzchen@scu.edu.cn (X. Chen).

mean-variance relationship $V = aM^b$ is a natural result of a double stochastic model if the K -function is given as $K(t) = \pi t^2 + K_0 t^f$. Besides, their exponents are related by $b = 1 + f/2$.

The idea of this paper is to implement a double stochastic model which can generate random samples in resemblance to the actual point pattern of urban facilities. Then we propose some other covariance functions in addition to the power function to fit the K -function. We compare these models in their capability to fit additional second order statistics such as mean-variance relationship, and other statistics beyond the first two orders, such as inter-event distribution function $H(t)$ and nearest neighbor distribution functions $G(t)$ and $F(t)$.

2. Choice of covariance function

A double stochastic process (DSP) model is introduced to model the spatial structure of urban facilities [4]. The DSP model assumes that there are two layers of stochastic process. The first layer stochastic process is a correlated random field of density function $\Lambda(x)$, which models the inhomogeneous concentration of facilities in a city. Conditional on the density, the location of urban facilities is based on a Poisson process. The DSP model simulates the growth process of urban facilities in the sense that urban facilities are evolving stochastically on top of an existing structure of a city, while the city structure itself follows a separate stochastic process.

A relatively flexible and tractable construction to encompass the non-negative constraint for density processes is log-Gaussian processes [9], i.e., the density function is drawn from a log-Gaussian random process $\Lambda(\vec{x}) = \exp(S(\vec{x}))$. S is assumed to be a stationery isotropic Gaussian, $S(\vec{x}) \sim \mathcal{N}(\mu, \sigma^2)$. Its spatial dependence is given by its covariance density $\text{Cov}(S(\vec{x}), S(\vec{x} + \vec{u})) = \sigma^2 r(|\vec{u}|)$. The first and second order statistics of S and Λ fields are related by $m = E[\Lambda(\vec{x})] = \exp(\mu + \sigma^2/2)$, and $\gamma(|\vec{u}|) = \text{Cov}(\Lambda(\vec{x}), \Lambda(\vec{x} + \vec{u})) = \exp(2\mu + \sigma^2)[\exp(\sigma^2 r(\vec{u})) - 1]$.

The radial distribution function $\rho(t) = (2\pi t)^{-1} K'(t)$ are related to the second order statistics of S and $\Lambda(\vec{x})$ by

$$\sigma^2 r(t) = \log(\rho(t)) \quad (1)$$

$$\gamma(t) = m^2(\rho(t) - 1) \quad (2)$$

Now, we propose 3 models of the radial distribution function $\rho(t)$ as

$$\rho_1(t) = 1 + \theta_{11}(1 + t^2)^{-\theta_{12}/2} \quad (3)$$

$$\rho_2(t) = 1 + \exp(\theta_{21} - \theta_{22}t) \quad (4)$$

$$\rho_3(t) = \exp(\theta_{31} \exp(-\theta_{32}t)) \quad (5)$$

These different models of $\rho(t)$ are hereafter referred as m1, m2 and m3 respectively. The radial distribution function $\rho_1(t)$ of m1 is actually a power function $\rho_1(t) \propto 1 + \theta_{11}t^{-\theta_{12}}$ at large values of t . 1 is added to avoid the divergence at $t = 0$. Note that for m1 $K(t) = \pi t^2 + \frac{\theta_{11}}{1-\theta_{12}}t^{-\theta_{12}+1}$, which is the model suggested in [4]. Note that m2 is the first order expansion of m3 $\rho_3(t) \approx 1 + \theta_{31} \exp(-\theta_{32}t) + O(\exp(-2\theta_{32}t))$. Therefore, the covariance of both m2 and m3 are exponential function, which dies quickly as t increases and the interaction between events are short-ranged.

3. Results

Similar to [4], we take banks in Beijing as an illustrative example. We present the original spatial data of banks in Fig. 1(a). The data is constrained to the metropolitan area of Beijing, which covers $2^{14} \times 2^{14} \text{ m}^2$. We map the data to a $2^{10} \times 2^{10}$ lattice by taking a transformation of the coordinate x of each point as $x' = [x/2^4]$, rounded to the nearest integer.

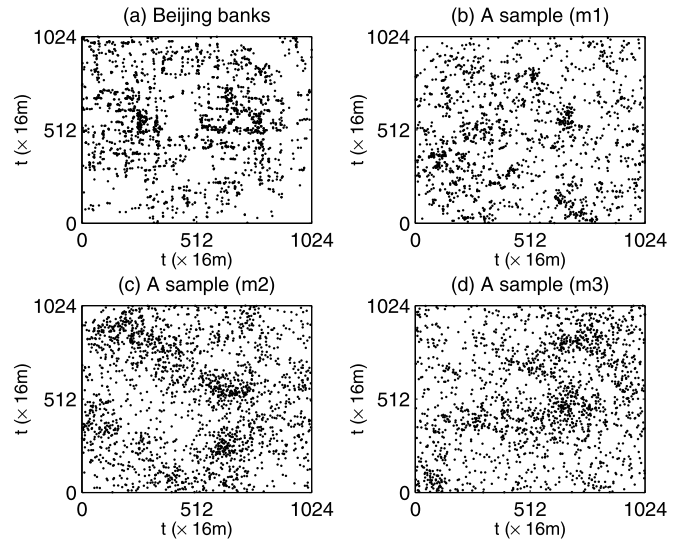


Fig. 1. (a) Original locations of banks in a metropolitan area of Beijing ($2^{14} \times 2^{14} \text{ m}^2$) mapped to a $(2^{10} \times 2^{10})$ lattice; a random sample of the point pattern generated from the DSP when the radial distribution function $\rho(t)$ is given by: (b) Eq. (3); (c) Eq. (4); (d) Eq. (5).

The estimation of $\rho(t)$ is not convenient in the spatial analysis of point patterns. Knowing that

$$K(t) = \int_0^t ds 2\pi s \rho(s) = m^{-1} E[N_0(t)] \quad (6)$$

where $N_0(t)$ is the number of further events within distance t of an arbitrary event, we estimate the models by fitting $K(t)$.

Denote $\hat{K}(t)$ the estimator calculated from the data after edge correction is applied to the square window in our case, and $K(t, \theta)$ the theoretical K -function. θ is estimated to minimize the square error $J = \sum_{i=1}^I (\hat{K}(t_i) - K(t_i, \theta))^2 / I$ between $\hat{K}(t)$ and $K(t, \theta)$. We choose 26 t_i s spanning from 5 to 500 units (each unit is 16 meters) which are separated in such a way that they are somehow uniformly distributed in log space to give more emphasis on small t s as shown in Fig. 2. From the figure, all three models fit the data quite well in the large range from 100 meters to 7000 meters. There is a slight difference when $t < 300$ meters as one can see from the log-log plot in the sub-window. The log-log plot shows that m1 fits slightly better than the other two models.

After we fit the models to $K(t)$, we can use the parameters to run simulations. Random samples can be generated with the Fourier filtering method [10] when the radial distribution function (therefore the covariance function $\sigma^2 r(t)$ of S field) is given. A random sample for each model of $\rho(t)$ is depicted in Fig. 1(b), Fig. 1(c) and Fig. 1(d) respectively. By first checking the graphs with naked eyes, one may conclude that the point pattern generated by m1 matches closer to the actual pattern of banks than the other two models. Rigorous statistical checks are conducted in a later section of this paper.

In Fig. 2, we plot the fitting results of three models in one figure. All three models are nearly indistinguishable in the large range of t from 100 meters to 7000 meters. One cannot really tell the difference in the regular plot. Although the covariance function for m2 and m3 are exponential which would deviate from the power function of m1 at large values of t , we cannot reject them based on the fitting of $K(t)$ given the finite size of actual physical systems. Fortunately, there is slight difference when $r < 300$ meters as one can see from the log-log plot in the sub-window of Fig. 2. The log-log plot shows that m1 fits slightly better than the other two models.

We extend our analysis to include additional statistics to further test the goodness of fit of different models. The first one is

Download English Version:

<https://daneshyari.com/en/article/1863387>

Download Persian Version:

<https://daneshyari.com/article/1863387>

[Daneshyari.com](https://daneshyari.com)