КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ
И МОДЕЛИРОВАНИЕ 2016 Т. 8 № 2 С. 323–332

Ки&М

ИНДИВИДУАЛЬНО-ОРИЕНТИРОВАННЫЕ МОДЕЛИ БИОСООБЩЕСТВ И ЭКОСИСТЕМ

УДК: 519.8

Методы оценивания параметров случайных точечных полей с локальным взаимодействием

П. Я. Грабарник

Институт физико-химических и биологических проблем почвоведения РАН,
Россия, 142290, г. Пущино, Московская обл., ул. Институтская, д. 2

E-mail: pavel.grabarnik@gmail.com

Получено 18 февраля 2016 г.,
после доработки 25 марта 2016 г.

В работе дается краткий обзор методов оценивания параметров случайных точечных процессов с локальным взаимодействием между точками. Показано, что общепринятый метод максимального псевдоправдоподобия является частным случаем методов оценивания, основанных на использовании вспомогательного марковского процесса, инвариантная мера которого является гиббсовским точечным полем с параметрами, подлежащими оцениванию. Предложено обобщение данного метода, приводящее к такому виду уравнений для получения оценок неизвестных параметров, который не может быть получен с помощью универсального метода Такача–Фикселя. Компьютерные эксперименты показывают, что новый метод позволяет получать оценки, качество которых выше, чем качество оценок широко используемого метода максимального правдоподобия.

Ключевые слова: гиббсовское точечное поле, оценивающая функция, псевдоправдоподобие, оценивание параметров

Parameter estimation methods for random point fields with local interactions

P. Ya. Grabarnik

Institute of Physicochemical and Biological Problems in Soil Science of RAS, 2 Institutskaya st., Pushchino, Moscow Region, 142290, Russia

Abstract. — The paper gives an overview of methods for estimating the parameters of random point fields with local interaction between points. It is shown that the conventional method of the maximum pseudo-likelihood is a special case of the family of estimation methods based on the use of the auxiliary Markov process, invariant measure of which is the Gibbs point field with parameters to be estimated. A generalization of this method, resulting in estimating equation that can not be obtained by the the universal Takacs–Fiksel method, is proposed. It is shown by computer simulations that the new method enables to obtain estimates which have better quality than those by a widely used method of the maximum pseudolikelihood.

Keywords: Gibbs point process, estimating function, pseudo-likelihood, parametric inference

Citation: *Computer Research and Modeling*, 2016, vol. 8, no. 2, pp. 323–332 (Russian).

Работа выполнена при поддержке РФФИ, проект № 16-04-01348.

Введение

Оценивание параметров — одна из важнейших задач статистических выводов и является обязательной частью учебников по математической статистике [Боровков, 1984]. Однако для задач пространственной статистики, где используются модели случайных точечных полей, получение оценок параметров до сих пор является областью интенсивных исследований [Møller, Waagepetersen, 2004]. Методы вычисления оценок параметров моделей пространственных точечных процессов или полей востребованы в различных областях естествознания, и многочисленные примеры можно найти в биологии, экологии, геофизике, почвоведении, археологии, географии [Baddeley et al., 2006] С такими задачами сталкивается, например, лесная экология, где пространственные точные процессы используются для описания горизонтальной структуры растительных сообществ [Grabarnik, Särkkä, 2009].

Часто системы распределенных в пространстве взаимодействующих объектов удобно описывать с помощью модели случайного точечного поля, вероятностное распределение которого определяется плотностью по отношению к распределению пуассоновского точечного поля. Преимущество так выбранной модели заключается в том, что плотность может быть выражена через *функции взаимодействия* в так называемой форме Гиббса, и такие модели называются *гиббсовскими точечными полями* или процессами [Малышев, Минлос, 1985]. Гиббсовские точечные поля интенсивно изучались в физике, так как хорошо описывают различные физические системы, например газообразное вещество [Минлос, 1967].

Несколько неожиданно эти модели привлекли к себе внимание специалистов по статистике и теории вероятностей, кто интересовался проблемами описания случайных геометрических систем и стохастическими процессами, в первую очередь Бартлета [Bartlett, 1974]. Его ученик, Бисаг, опубликовал работу [Besag, 1974], в которой обозначил контуры нового направления — статистики случайных полей. В этой и нескольких последующих работах Бисагом был развит метод оценивания параметров случайных точечных полей на решетке, названный им методом *максимального псевдоправдоподобия* (МПП). Свойства этого метода для полей на решетке изучались многими авторами; в частности, была доказана состоятельность (см., например, [Winkler, 2012]) и, при некоторых условиях, асимптотическая нормальность МПП-оценок [Jensen, Künsch, 1994].

Работа Бесага привлекла внимание А. С. Комарова. В это время, на рубеже 70–80-х годов, он интересовался количественными оценками конкурентного взаимодействия между растениями. Исследования, выполненные под его руководством [Комаров, Щербаков, 1979; Грабарник, Комаров, 1981], доказывали перспективность использования современных статистических методов в лесной экологии для решения практических задач. В теоретическом плане все еще оставалась актуальной проблема математического обоснования метода максимального псевдоправдоподобия, который долгое время рассматривался скорее как эвристическая процедура, чем как строгий математический метод. В настоящей работе предлагается теоретическая конструкция, которая позволяет построить новый класс методов оценивания параметров моделей сложных систем взаимодействующих объектов. Новый класс методов оценивания основан на введении вспомогательного марковского процесса, генератор которого представляет собой оценивающую функцию. Новый класс включает указанный метод максимального псевдоправдоподобия, который ведет далеко не всегда к оптимальным оценкам в данном классе процедур оценивания.

Часто задачей исследователя является нахождение оценок параметров моделей пространственных точечных полей, когда данные представляют собой множество или конфигурацию точек (например, расположение деревьев на участке леса), заданных координатами в некоторой ограниченной области евклидова пространства. Для простоты мы предполагаем, что пространство двумерно, т. е. точки конфигураци расположены на плоскости. Метод максимального правдоподобия для таких моделей не может быть использован из-за трудностей, связанных с отсутствием аналитического представления нормирующей константы. В этом случае оценивание методом максимального псевдоправдоподобия является альтернативой, позволяющей избежать вычислительных проблем. Однако МПП-оценки являются неэффективными и имеют значительное смещение в случае моделей, в которых взаимодействие между точками достаточно сильно. Следовательно, имеет смысл поиск оценивающих процедур, обладающих вычислительной простотой МПП-оценок, но имеющих лучшие статистические свойства. Мы отталкиваемся от нескольких отправных моментов. Во-первых, как было замечено нами ранее [Grabarnik, 1996], между методами оценивания и алгоритмами генерации гиббсовских точечных полей имеется глубокая связь. Так, псевдоправдоподобие для точечного поля X имеет близкое отношение к одному из методов статистического моделирования с помощью марковских цепей (Markov Chain Monte Carlo или MCMC). Этот метод основан на пространственном процессе рождения и гибели с непрерывным временем (Y_t) с интенсивностью гибели 1 и интенсивностью рождения, равной условной интенсивности процесса X. Псевдо-вклад (производная логарифма функции псевдоправдоподобия) для случайного точечного поля X является генератором марковского процесса (Y_t). Оценки, получаемые с помощью генератора марковского процесса, были введены в [Baddeley, 2000] и называются оценками, *инвариантными во времени* (ИВ-оценки). Эта связь позволяет предположить, что относительная эффективность ИВ-оценок имеет отношение к скорости сходимости соответствующей марковской цепи к стационарному распределению. Низкое качество оценок максимального псевдоправдобододобия может быть объяснено тем, что цепь, на которой они основаны, медленно сходится к стационарному распределению в случае моделей с сильно взаимодействующими точками. Чтобы ускорить сходимость, были предложены некоторые альтернативные MCMC-алгоритмы [Geyer, Møller, 1994], адаптирующие алгоритмы Метрополиса и Хастингса [Hastings, 1970] к случаю пространственных точечных полей. Поэтому задача рассмотрения альтернативных МПП-оцениванию методов, основанных на MCMC-алгоритмах, отличных от пространственных процессов рождения и гибели, является важной в теоретичском плане и представляет определенный практический интерес. Во-вторых, основными критериями выбора той или иной оценки служат качество и удобство вычислительной реализации. В работе [Baddeley, Turner, 2000] рассматривались практические аспекты вычислительной реализации оценок максимального псевдоправдоподобия. На основе вычислительного эксперимента показывается, что среди ИВ-оценок существует такая, которая имеет значительно лучшие статистические свойства, чем МПП-оценка, и которая в то же время допускает реализацию с помощью стандартных статистических пакетов, так как относится к классу оценок сложного правдоподобия (composite likelihood).

Гиббсовское точечное поле

В этой работе мы концентрируемся на параметрической модели конечных гиббсовских полей, реализации которых являются конечными множествами точек $\mathbf{x} = \{x_1, \ldots, x_n\}$ в ограниченном борелевском множестве $E \subset R^d$. Эта модель задается вероятностной плотностью $f : \Omega \times \Theta \to R_+$ (где Ω — выборочное пространство, состоящее из всех возможных конечных точечных конфигураций $\mathbf{x}$; Θ — произвольное параметрическое пространство) относительно пуассоновского поля на E с единичной интенсивностью. Мы предполагаем, что вероятностная плотность удовлетворяет условию положительности, то есть из $f(\mathbf{x}) > 0$ следует $f(\mathbf{y}) > 0$ для всех $\mathbf{y} \subset \mathbf{x}$. θ — вектор неизвестных параметров, которые необходимо оценить по имеющейся единственной наблюдаемой конфигурации точек $\mathbf{x}$ в выборочном окне W.

Полезной характеристикой гиббсовских точечных полей является функция условной интенсивности, определяемая как отношение плотностей вида

$$\lambda(u, \mathbf{x}; \theta) = \frac{f(\mathbf{x} \cup \{u\}; \theta)}{f(\mathbf{x}; \theta)}.$$

Основной моделью в данной работе является так называемое поле Штраусса [Strauss, 1975] с твердой сердцевиной, которое имеет плотность

$$f(\mathbf{x};\theta) = \frac{1}{z(\theta)}\exp(\alpha\, n(\mathbf{x}) + \beta\, s(\mathbf{x}))\, \mathbf{1}(h(\mathbf{x}) = 0),$$

где $\theta = (\alpha, \beta)$ — параметры, которые управляют плотностью точек (параметр α) и взаимодействием между точками (параметр β); $n(\mathbf{x})$ обозначает число точек в точечном множестве $\mathbf{x}$; статистика

$$s(\mathbf{x}) = \#\{(x_i, x_j) : i < j, r_0 < \|x_i - x_j\| \le r, x_i, x_j \in \mathbf{x}\}$$

соответствует числу пар точек, которые удалены далее, чем r_0, и ближе, чем r единиц друг от друга; $z(\theta)$ — нормирующий множитель. Статистика

$$h(\mathbf{x}) = \#\{(x_i, x_j) : i < j, \|x_i - x_j\| \le r_0, x_i, x_j \in \mathbf{x}\}$$

определяет условие твердой сердцевины, которое проверяет отсутствие пар точек, расположенных ближе друг к другу, чем r_0 единиц. r_0 называется *радиусом твердой сердцевины*, а r — *радиусом взаимодействия*, они могут быть выбраны заранее или оцениваться по данным, $\mathbf{1}(\cdot)$ — индикаторная функция, равная 1, если условие в скобках выполнено, и 0 в противном случае. Мы включили условие твердой сердцевины в модель Штраусса, чтобы позволить параметру взаимодействия β быть положительным. В этом случае реализациями модели являются точечные конфигурации, имеющие характер группового (кластерного) размещения точек.

Методы псевдоправдоподобия и Такача–Фикселя

Метод максимального правдоподобия не может быть использован для гиббсовских точечных полей, так как нормирующий множитель $z(\theta)$ в функции плотности не имеет аналитического представления. Бисаг [Besag, 1975] предложил максимизировать *псевдоправдоподобие*, которое в случае дискретных марковских полей определяется как произведение условных вероятностей. В нашей работе [Diggle et al., 1994] была введена функция псевдоправдоподобия для случая конечных гиббсовских полей, которая в общей форме имеет вид

$$PL(\theta, \mathbf{x}) = \prod_{x_i \in \mathbf{x}} \lambda(x_i, \mathbf{x} \setminus x_i; \theta) \exp\left(-\int_W \lambda(u, \mathbf{x}; \theta)\, \mathrm{d}u\right). \tag{1}$$

Другой подход к проблеме оценивания параметров гиббсовских точечных полей (заданных на всем пространстве R^d) был предложен Такачем и развит Фикселем [Fiksel, 1984]. Они предложили оценивать параметр θ, решая оптимизационную задачу

$$\sum_k \left[\sum_{x_i \in \mathbf{x}} g_k(x_i, \mathbf{x} \setminus x_i) - \int_W g_k(u, \mathbf{x})\lambda(u, \mathbf{x}; \theta)\, \mathrm{d}u\right]^2 \to \min_\theta$$

для множества различных пробных функций $g_k(u, \mathbf{x})$. Обоснование метода Такача–Фикселя основано на применении формулы Нгуена–Цезина [Nguyen, Zessin, 1979]

$$\mathbb{E}\left[\sum_{x_i \in \mathbf{x}} g(x_i, \mathbf{x} \setminus x_i)\right] = \mathbb{E}\left[\int_W g(u, \mathbf{x})\lambda(u, \mathbf{x})\, \mathrm{d}u\right],$$

левая сторона которой благодаря уточненной теореме Кэмпбелла может быть выражена через ожидание относительно распределения Пальма [Stoyan, Kendall, Mecke, 1995]. Различный выбор функции g ведет к различным оценивающим уравнениям, и, в частности, выбор $g_k = \frac{\partial \lambda}{\partial \theta_k}$ дает уравнение псевдоправдоподобия [Diggle et al., 1994].

Инвариантный во времени метод оценивания параметров

Бэддлей [Baddeley, 2000] предложил общий подход к конструированию оценок для параметрических выводов в случае, когда вероятностная мера статистической модели может быть определена как инвариантное распределение некоторой марковской цепи или процесса. Предположим, что $(Y_n, n \geq 0)$ — дискретный марковский процесс для вероятностной модели $f(\mathbf{x};\theta)$. Другими словами, (Y_n) — марковский процесс такой, что для любого θ инвариантное распределение процесса (Y_n) имеет плотность вероятности $f(\mathbf{x};\theta)$ относительно пуассоновской меры. Определим генератор $\mathcal{A}_\theta$ процесса (Y_n) как оператор вида

$$(\mathcal{A}_\theta S)(\mathbf{x}) = \mathbb{E}_\theta\left[S(Y_{i+1}) - S(Y_i) \mid Y_i = \mathbf{x}\right], \tag{2}$$

где S — произвольная статистика, определенная для точечной конфигурации $\mathbf{x}$. *Инвариантная во времени* оценка $\hat{\theta}$ есть решение уравнения $(\mathcal{A}_\theta S)(\mathbf{x}) = 0$. Поскольку $\mathbb{E}(\mathcal{A}_\theta S)(\mathbf{x}) = 0$, уравнение (2) является *несмещенным оценивающим уравнением*.

В работе [Baddeley, 2000] установлено, что уравнения псевдоправдоподобия могут быть получены из ИВ-оценивающих уравнений для экспоненциального семейства моделей, основанных на марковском процессе рождения и гибели с непрерывным временем (РГНВ), таком, что интенсивности рождения и гибели имеют вид

$$\begin{aligned} b_\theta(\mathbf{x}, u) &= \frac{f(\mathbf{x} \cup u;\theta)}{f(\mathbf{x};\theta)} = \lambda(u, \mathbf{x};\theta), \\ d_\theta(\mathbf{x}, x_i) &\equiv 1. \end{aligned}$$

Отметим, что можно построить целый класс ИВ-оценивающих уравнений, основанных на РГНВ-марковских процессах с помощью генератора

$$(\mathcal{A}_\theta S)(\mathbf{x}) = \sum_{x_i}\left[S(\mathbf{x} \setminus x_i) - S(\mathbf{x})\right]\lambda^{(1-p)}(x_i, \mathbf{x};\theta) + \int_W \left[S(\mathbf{x} \cup u) - S(\mathbf{x})\right]\lambda^{p}(u, \mathbf{x};\theta)\,\mathrm{d}u. \tag{3}$$

Можно показать, что в случае $p = 1$ уравнение $(\mathcal{A}_\theta S)(\mathbf{x}) = 0$ совпадает с уравнением псевдоправдоподобия.

Далее будет установлено, что новый вид оценивающих уравнений может быть получен с помощью дискретного марковского процесса (алгоритма) Метрополиса–Хастингса.

Пусть (Y_n) — дискретный марковский процесс Метрополиса–Хастингса [Hastings, 1970; Geyer, Moller, 1994]. Определим вспомогательный процесс (proposal process) Q с помощью функции $q(\mathbf{x} \to \mathbf{x} \cup u)$ такой, что $\mathbf{x} \to \mathbf{x} \cup u$ соответствует рождению с вероятностью p, где u равномерно распределено в W, и $\mathbf{x} \to \mathbf{x} \setminus x_i$ соответствует гибели с вероятностью $1 - p$, где x_i — одна из точек, выбранная случайно с равной вероятностью. Тогда вероятности принятия перехода для марковского процесса Метрополиса–Хастинга имеют вид

$$A(\mathbf{x}, \mathbf{x} \cup u) = c(\mathbf{x}, \mathbf{x} \cup u)\Big/\left(1 + \frac{f(\mathbf{x})q(\mathbf{x} \to \mathbf{x} \cup u)}{f(\mathbf{x} \cup u)q(\mathbf{x} \cup u \to \mathbf{x})}\right),$$

где $c(\cdot)$ — симметричная функция, выбранная так, чтобы $0 \leq A(\cdot) \leq 1$.

В работе [Geyer, Moller, 1994] предложена конструкция марковского процесса, максимизирующая вероятности принятия перехода так, что плотность ядра переходных вероятностей процесса (Y_n) имеет вид

$$\begin{aligned} r_\theta(\mathbf{x} \to \mathbf{x} \cup u) &= \frac{p}{|W|}A(\mathbf{x}, \mathbf{x} \cup u) = \frac{p}{|W|}\min\left(1, \lambda(u, \mathbf{x};\theta)\frac{(1-p)}{p}\frac{|W|}{n(\mathbf{x}) + 1}\right), \\ r_\theta(\mathbf{x} \to \mathbf{x} \setminus x_i) &= \frac{1-p}{n(\mathbf{x})}A(\mathbf{x}, \mathbf{x} \setminus x_i) = \frac{1-p}{n(\mathbf{x})}\min\left(1, \frac{1}{\lambda(x_i, \mathbf{x};\theta)}\frac{p}{(1-p)}\frac{n(\mathbf{x})}{|W|}\right). \end{aligned}$$

Тогда генератор для такого процесса (Y_n) имеет вид

$$\begin{aligned}(\mathcal{A}_\theta S)(\mathbf{x}) &= \sum_{x_i} r_\theta(\mathbf{x} \to \mathbf{x} \setminus x_i)\,[S(\mathbf{x} \setminus x_i) - S(\mathbf{x})] + \\ &+ \int_W r_\theta(\mathbf{x} \to \mathbf{x} \cup u)\,[S(\mathbf{x} \cup u) - S(\mathbf{x})]\ \mathrm{d}u = \\ &= \sum_{x_i} \frac{1-p}{n(\mathbf{x})} \min\left(1, \frac{1}{\lambda(x_i, \mathbf{x}; \theta)} \frac{p}{(1-p)} \frac{n(\mathbf{x})}{|W|}\right)[S(\mathbf{x} \setminus x_i) - S(\mathbf{x})] + \\ &+ \int_W \frac{p}{|W|} \min\left(1, \lambda(u, \mathbf{x}; \theta) \frac{(1-p)}{p} \frac{|W|}{n(\mathbf{x})+1}\right)[S(\mathbf{x} \cup u) - S(\mathbf{x})]\ \mathrm{d}u.\end{aligned}$$

Более компактно это может быть записано как

$$\begin{aligned}(\mathcal{A}_\theta S)(\mathbf{x}) = \sum_{x_i} m_\theta(x_i, \mathbf{x} \setminus x_i)\ &[S(\mathbf{x} \setminus x_i) - S(\mathbf{x})] + \\ + \int_W m_\theta(u, \mathbf{x})\ &[S(\mathbf{x} \cup u) - S(\mathbf{x})]\,\lambda(u, \mathbf{x}; \theta)\ \mathrm{d}u,\end{aligned} \tag{4}$$

где

$$m_\theta(u, \mathbf{x}) = \min\left(\frac{1-p}{n(\mathbf{x} \cup u)}, \frac{1}{\lambda(u, \mathbf{x}; \theta)} \frac{p}{|W|}\right).$$

Таким образом, полученная оценивающая функция аналогична функции вклада псевдоправдоподобия, но имеет дополнительный весовой сомножитель $m_\theta(\cdot)$. Иной выбор функции $c(\cdot) \equiv 1$ ведет к форме марковского процесса, который для вспомогательной цепи Q с симметричными переходными вероятностями является марковским процессом Баркера–Хастингса [Hastings, 1970)], для которого соответствующая оценивающая функция принимает вид

$$\begin{aligned}(\mathcal{A}_\theta S)(\mathbf{x}) = \sum_{x_i} &\frac{1}{p\,n(\mathbf{x}) + (1-p)\lambda(x_i, \mathbf{x}; \theta)|W|}[S(\mathbf{x} \setminus x_i) - S(\mathbf{x})] + \\ + \int_W &\frac{\lambda(u, \mathbf{x}; \theta)}{p\,n(\mathbf{x} \cup u) + (1-p)\lambda(u, \mathbf{x}; \theta)|W|}\ [S(\mathbf{x} \cup u) - S(\mathbf{x})]\ \mathrm{d}u,\end{aligned} \tag{5}$$

что можно переписать в виде, аналогичным виду функции (4), но с другим весовым множителем

$$m_\theta(u, \mathbf{x}) = \frac{1}{p\,n(\mathbf{x} \cup u) + (1-p)\lambda(u, \mathbf{x}; \theta)|W|}.$$

Преимущество последней формы оценивающей функции состоит в том, что можно определить такую функцию, производная от логарифма которой дает оценивающую функцию, соответствующую процессу Баркера–Хастингса. Определим новую функцию QL_p, которая может рассматриваться как модификация функции псевдоправдоподобия, в виде

$$\begin{aligned}QL_p(\theta; \mathbf{x}) = \exp\Bigg(&\frac{1}{p} \sum_{x_i} \log \frac{\lambda(x_i, \mathbf{x}; \theta)}{p\,n(\mathbf{x}) + (1-p)\lambda(x_i, \mathbf{x}; \theta)|W|} - \\ &- \frac{n(\mathbf{x})}{(1-p)|W|} \int_W \log(p\,n(\mathbf{x} \cup u) + (1-p)\lambda(u, \mathbf{x}; \theta)|W|)\ \mathrm{d}u\Bigg),\end{aligned}$$

и тогда частные производные $\log QL$ дают оценивающие уравнения Баркера–Хастингса (5) [Grabarnik, Baddeley, 2005].

Заметим, что эта идея в несколько ином виде была реализована в работе [Baddeley et al., 2014].

Вычислительный эксперимент (см. следующие секции) демонстрирует, что статистические свойства оценки, основанной на QL-функции, лучше, чем у оценки максимального псевдоправдоподобия, в то время как получение QL-оценки может быть выполнено на базе любого пакета статистических программ, который имеет модуль расчета оценок обобщенной линейной модели (GLM).

Обобщение инвариантного во времени метода оценивания

Инвариантный во времени подход к задаче оценивания позволяет конструировать широкий класс оценивающих уравнений, но он тем не менее уже, чем класс уравнений Такача–Фикселя. Чтобы расширить класс инвариантных во времени оценок, предлагается следующая модификация этого метода.

Пусть (Y_n) — обратимый однородный во времени дискретный марковский процесс. Рассмотрим оператор $\mathcal{B}_\theta$, действующий на функции $V : \Omega \times \Omega \to R^k$ (где Ω — выборочное конфигурационное пространство, k — размерность статистики V), определенный как

$$(\mathcal{B}_\theta V)(\mathbf{x}) = \mathbb{E}_\theta \left[V(Y_n, Y_{n+1}) | Y_n = \mathbf{x} \right].$$

Тогда может быть показано, что для антисимметричных функций $V(\mathbf{x}, \mathbf{y})$, т. е. $V(\mathbf{x}, \mathbf{y}) = -V(\mathbf{y}, \mathbf{x})$,

$$\mathbb{E}_\theta(\mathcal{B}_\theta V)(\mathbf{x}) = 0.$$

Таким образом, $(\mathcal{B}_\theta V)(\mathbf{x}) = 0$ является несмещенным оценивающим уравнением.

Можно показать, что класс уравнений Такача–Фикселя содержится в классе обобщенных ИВ-оценивающих уравнений. Для этого следует выбрать в качестве вспомогательного марковского процесса пространственный процесс рождения и гибели функции V в виде

$$V(\mathbf{x}, \mathbf{y}) = \begin{cases} -w(\mathbf{x}, \mathbf{y}), & \text{если } y = \mathbf{x} \setminus u, \ u \in \mathbf{x}, \\ w(\mathbf{x}, \mathbf{y}), & \text{если } y = \mathbf{x} \cup u, \ u \notin \mathbf{x}, \\ 0 & \text{в остальных случаях,} \end{cases}$$

где $w(\mathbf{x}, \mathbf{y})$ — произвольная симметричная функция.

Еще один выбор функции V в виде

$$V(\mathbf{x}, \mathbf{y}) = \operatorname{sign}\left(\log \frac{f(\mathbf{y}; \theta)}{f(\mathbf{x}; \theta)} \right) w(\mathbf{x}, \mathbf{y})$$

приводит к оценивающим уравнениям, которые не могут быть получены с помощью уравнений Такача–Фикселя. В частном случае, если взять $w(\mathbf{x}, \mathbf{y}) = 1$ и процесс с динамикой Метрополиса, мы получим следующие оценивающие уравнения:

$$(B_\theta V)(\mathbf{x}) = \sum_{x_i \in \mathbf{x}} \int_{W \setminus \mathbf{x}} \mathbf{1}(\Delta E < 0) - \mathbf{1}(\Delta E > 0) \exp(-\Delta E) \, \mathrm{d}u, \tag{6}$$

где $\Delta E = E(\mathbf{x} \setminus x_i \cup u; \theta) - E(\mathbf{x}; \theta)$ и $E(\cdot) = -\log(f(\cdot)/f(\emptyset))$ называется *энергией системы*.

Сравнение статистистических свойств ИВ-оценок

Мы включили следующие оценки в список сравниваемых оценок:

- оценки максимального псевдоправдоподобия (1), МПП;
- оценки, основанные на марковской цепи рождения и гибели с параметром $p = 0.5$ (3), РГНВ;
- оценки, основанные на марковской цепи с динамикой Гейера и Мёллера с параметром $p = 0.5$ (4), Г–М;
- оценки, основанные на марковской цепи с динамикой Беркера–Хастингса с параметром $p = 0.5$ (5), Б–Х;
- оценки, основанные на марковской цепи с динамикой Метрополиса (6), МТ.

Для изучения статистических свойств выбранных оценок было проведено два вычислительных эксперимента. В первом случае исследовались оценки для конечных точечных множеств, взаимодействия между точками которых носят характер отталкивания. В качестве статистической модели было использовано точечное поле Штраусса со случайным числом точек. Для каждого из пяти значений параметра β (-0.8, -1.2, -1.6, -2.0, -2.4) были сгенерированы 1000 реализаций в единичном квадрате. Чтобы избежать граничных эффектов, был использован метод периодических границ. Радиус взаимодействия был выбран равным 0.1, а параметр α — так, чтобы число точек в реализации насчитывало в среднем 100 точек для каждого значения параметра взаимодействия β.

Главная цель вычичслительного эксперимента состояла в том, чтобы сравнить оценки параметра взаимодействия β, поэтому мы не приводим результаты о качестве оценок параметра α, имеющие второстепенный интерес. В таблице 1 приведены результаты первого эксперимента, проведенного для сравнения оценок, в терминах средних и стандартных ошибок оценок, вычисленных на основе смоделированных реализаций.

Таблица 1. Средние и стандартные ошибки (в скобках) для ИВ-оценок параметра взаимодействия β в модели Штраусса с взаимодействием отталкивающего типа

Оценка	$\beta = -0.8$	$\beta = -1.2$	$\beta = -1.6$	$\beta = -2.0$	$\beta = -2.4$
МПП	−0.842 (0.192)	−1.27 (0.252)	−1.73 (0.353)	−2.23 (0.508)	-2.48^* (0.870)
РГНВ	−0.851 (0.193)	−1.27 (0.241)	−1.70 (0.296)	−2.14 (0.356)	−2.58 (0.423)
Г–М	−0.836 (0.189)	−1.24 (0.230)	−1.66 (0.278)	−2.08 (0.328)	−2.50 (0.373)
Б–Х	−0.842 (0.188)	−1.25 (0.230)	−1.67 (0.275)	−2.09 (0.327)	−2.51 (0.377)
МТ	−0.813 (0.183)	−1.22 (0.222)	−1.63 (0.261)	−2.05 (0.310)	−2.47 (0.355)

* В 5.2 % реализаций конечная оценка не была найдена.

Результаты эксперимента демонстрируют, что оценка, основанная на марковском процессе с динамикой Метрополиса, является наилучшей в терминах как смещения, так и стандартной ошибки. ИВ-оценки, основанные на марковских процессах с динамикой обновления Гейера–Мёллера и Баркера–Хастингса, имеют сходное качество, и их статистические свойства несколько лучше, чем у оценки, основанной на марковском процессе рождения и гибели с параметром $p = 0.5$. Наш эксперимент подтверждает низкое качество МПП-оценок в случае моделей с большими значениями параметра взаимодействия, как это раньше уже было отмечено в литературе [Diggle et al., 1994].

Во втором эксперименте были исследованы свойства оценок для моделей с взаимодействием между точками типа притяжения, которые порождают точечные множества с кластерными свойствами. В качестве модели было использовано случайное точечное поле Штраусса с твердой сердцевиной. В данном случае генерировались реализации модели с фиксированным числом точек, равным 100. Результаты эксперимента суммированы в таблице 2.

Таблица 2. Средние и стандартные ошибки (в скобках) для ИВ-оценок параметра взаимодействия β в кластерной модели Штраусса

Оценка	$\beta = 0.4$	$\beta = 0.8$	$\beta = 1.2$	$\beta = 1.6$
МПП	0.371 (0.187)	0.780 (0.183)	1.22 (0.249)	1.92* (0.604)
РГНВ	0.371 (0.189)	0.770 (0.171)	1.17 (0.169)	1.60 (0.212)
Г–М	0.377 (0.203)	0.772 (0.196)	1.16 (0.196)	1.58 (0.226)
Б–Х	0.374 (0.197)	0.773 (0.187)	1.17 (0.190)	1.58 (0.224)
МТ	0.401 (0.186)	0.796 (0.169)	1.20 (0.175)	1.66 (0.255)

* В 2.9 % реализаций конечная оценка не была найдена.

В отличие от предыдущего случая в модели с притяжением лучшие результаты были показаны оценкой, основанной на марковском процессе с РГНВ-динамикой и параметром $p = 0.5$. ИВ-оценки, основанные на марковском процессе с динамикой обновления Гейера–Мёллера и Баркера–Хастингса, вновь демонстрируют сходное качество. Интересно, что МТ-оценка, которая была лучшей для моделей с отталкиванием, в этом эксперименте уступила РГНВ-оценке.

Заключение

В работе рассмотрены методы нахождения оценок параметров случайных точечных полей, заданных плотностью в гиббсовской форме. Эти методы основаны на использовании вспомогательной марковской цепи, инвариантные меры которой представляют собой распределения статистической модели, подгоняемой к данным. Так как конкретный метод получения оценок зависит от вида динамики цепи, встает вопрос об оптимальности такого выбора. Критерием оптимальности служит наименьшая потеря эффективности оценки по сравнению с оценками максимального правдоподобия, получение которых потребовало бы значительных вычислительных ресурсов. Наши эксперименты показывают, что оценки псевдоправдоподобия могут быть далеки от оптимальных оценок в классе инвариантных во времени оценивающих процедур.

С практической точки зрения наличие надежных оценок параметров случайных точечных полей имеет важное значение, так как позволяет изучать пространственную структуру многих биосистем при анализе значений и изменений параметров моделей этих систем. Например, в экологии величины параметров могут интерпретироваться как напряженность конкурентного взаимодействия между деревьями, что дает возможность строить правдоподобные биологические гипотезы и ведет к более глубокому пониманию процессов, протекающих в растительных сообществах.

Список литературы

Боровков А. А. Математическая статистика. Оценка параметров. Проверка гипотез. — М.: Наука, 1984. — 472 с.

Грабарник П. Я., Комаров А. С. Статистический анализ горизонтальной структуры древостоя // Моделирование биогеоценотических процессов. — М.: Наука, 1981. — С. 119–135.

Комаров А. С., Щербаков Р. А. Статистический анализ пространственных структур. Двоичная случайная переменная на правильной решетке // Экомодель-2. Материалы по математическому обеспечению ЭВМ. — Пущино: ОНТИ НЦБИ, 1979. — С. 1–48.

Малышев В. А., Минлос Р. А. Гиббсовские случайные поля. Метод кластерных разложений. — М.: Наука, Глав. ред. физ.-мат. лит-ры, 1985. — 356 с.

Минлос Р. А. Предельное распределение Гиббса // Функциональный анализ и его приложения. — 1967. — Т. 1, № 2. — С. 60–73.

Baddeley A. Time-invariance estimating equations // Bernoulli. — 2000. — Vol. 6. — P. 783–808.

Baddeley A., Turner R. Practical maximum pseudolikelihood for spatial point patterns (with discussion) // *Australian and New Zealand Journal of Statistics*. — 2000. — Vol. 42. — P. 283–322.

Baddeley A., Gregori P., Mateu J., Stoica R., Stoyan D. Case Studies in Spatial Point Process Modeling // Lect. Notes Statist. — 2006. — Vol. 185. — N.Y.: Springer, 307 p.

Baddeley A., Coeurjolly J. F., Rubak E., Waagepetersen R. Logistic regression for spatial Gibbs point processes // Biometrika. — 2014. — Vol. 101, No. 2. — P. 377–392.

Bartlett M. S. The statistical analysis of spatial pattern // Advances in Applied Probability. — 1974. — Vol. 6, No. 2. — P. 336–358.

Besag J. Spatial interaction and the statistical analysis of lattice systems // Journal of the Royal Statistical Society. Series B. — 1974. — Vol. 36, No. 2. — P. 192–236.

Besag J. Statistical analysis of non-lattice data // Statistician. — 1975. — Vol. 24, No. 3. — P. 179–195.

Fiksel T. Estimation of parameterized pair potentials of marked and non-marked Gibbsian point processes // Elektron. Informationsverarb. Kybernet. — 1984. — Vol. 20. — P. 270–278.

Geyer C. J., Møller J. Simulation procedures and likelihhod inference for spatial point processes // Scand. J. Statist.. — 1994. — Vol. 21. — P. 359–373.

Grabarnik P. A connection between estimation and simulation methods of spatial point processes// Seminaire European de Statistique on “Stochastic Geometry, Theory and Application”. — Toulouse. — 13-18 May, 1996.

Grabarnik P., Baddeley A. Time-invariance estimators for spatial point processes: performance and implementation // Abstracts of a conference “Stochastic Geometry and its Application”. — Bern. — 3–7 October, 2005.

Grabarnik P., Särkkä A. Modelling the spatial structure of forest stands by multivariate point processes with hierarchical interactions // Ecological Modelling. — 2009. — Vol. 220, No. 9–10. — P. 1232–1240.

Hastings W. K. Monte Carlo sampling methods using Markov chains and their applications // Biometrika. — 1970. — Vol. 57. — P. 97–109.

Jensen J. L., Künsch H. R. On asymptotic normality of pseudo likelihood estimates for pairwise interaction processes // Ann. Inst. Statist. Math. — 1994. — Vol. 46. — P. 475–486.

Møller J., Waagepetersen R. P. Statistical Inference and Simulation for Spatial Point Processes. — Chapman and Hall/CRC, Boca Raton. — 2004. — 301 p.

Nguye X. X. and Zessin H. (1979) Integral and differencial characterizations of the Gibbs process // Math. Nach. — 1979. — Vol. 88. — P. 105–115.

Stoyan D., Kendall W. S., Mecke J. Stochastic Geometry and its Applications, 2nd ed. John Wiley, Chichester. — 1995.

Strauss D. J. A model for clustering // Biometrika. — 1975. — Vol. 62, No. 2. — P. 467–475.

Winkler G. Image analysis, random fields and Markov chain Monte Carlo methods: a mathematical introduction. — Springer Science & Business Media, 2012.