

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ РОЗПОВСЮДЖЕННЯ ВІРУСНОЇ ХВОРОБИ В ПОПУЛЯЦІЯХ З ПРИХОВАНИМИ НОСІЯМИ

О.С. САМОСЬОНОК,
Інститут кібернетики ім. В.М. Глушкова НАН
України, Київ, Україна
samosyonok@gmail.com

***Анотація.** Приводиться алгоритм оцінювання невідомих параметрів моделі розповсюдження вірусної хвороби, що побудована на основі інструментарію марковських полів за допомогою методу максимальної вірогідності. Сама модель побудована на основі припущення, що стан популяції є певною конфігурацією скінченного марковського випадкового поля.*

***Ключові слова:** марковські випадкові поля, гіббсовський розподіл, метод максимальної вірогідності.*

Нехай $X = (X_1, \dots, X_{|S|})$ – багатовимірний випадковий величина, де кожна компонента X_i є одновимірною випадковою величиною, що приймає значення зі скінченної множини станів A певної особини в популяції, S – множина індексів. Конкретна реалізація $x = (x_1, \dots, x_{|S|})$ випадкової величини X представляє собою спільну появу всіх подій $(X_1 = x_1, \dots, X_{|S|} = x_{|S|})$ і фактично характеризує загальну епідеміологічну ситуацію X над певною популяцією. Спільний розподіл станів елементів поля може бути виражено безпосередньо через умовні розподіли:

$$P(x) = P(x_n | x_1, \dots, x_{n-1}) P(x_1, \dots, x_{n-1}).$$

У випадку одновимірного випадкового поля – ланцюга Маркова останній множник можна обчислити по умовним розподілам:

$$P(x) = P(x_0) \prod_{i=1}^{n-1} P(x_i | x_{i-1}),$$

за умови, що відомо розподіл ймовірностей станів деякого крайнього елемента X_0 і всі умовні розподіли. Властивість марковості дозволяє спростити проблему дослідження взаємодії великої (або навіть нескінченної) кількості елементів в моделі завдяки цілком природному припущенню про те, що стан деякого елемента системи залежить лише від станів скінченної кількості сусідніх елементів.

Вводячи поняття сусідства запишемо спільний розподіл ймовірностей станів елементів марковського випадкового поля X по простору його конфігурацій x як

$$P(x) = \frac{\prod_{\chi \in S} F_{\chi}(x_{\chi})}{\sum_{x' \in X} \prod_{\chi \in S} F_{\chi}(x'_{\chi})}$$

де $F_{\chi}(x_{\chi}, v)$ - додатньовизначена функція, відмінна від константи. \mathcal{X} - кліка, множина елементів випадкового поля, яка відповідає повнзв'язному підграфу графа сусідства.

Записавши згадані функції в експоненціальній формі

$$F_{\chi}(x_{\chi}, v) = e^{-H(x_{\chi}, v)}$$

спільний розподіл ймовірностей прийме вигляд:

$$P(x) = Z(v)^{-1} \exp\left(-\sum_{j=1}^{|\mathcal{C}|} \sum_{\chi \subset S} H_{\chi}(x_{\chi}, v^j)\right), \quad v = (v^j)_{j=1}^l,$$

де Z - нормуючий множник.

Розглядаючи послідовність незалежних спостережень $\{x^1, \dots, x^n\}$ за станами особин популяції виникає задача оцінювання істинного значення параметра v^* за результатами спостережень $x \in \Lambda$.

Для розв'язку задачі використовується метод максимальної вірогідності. В контексті нашої задачі функція максимальної вірогідності приймає вигляд

$$L_n(x, v) = \sum_{i=1}^n \left(\sum_{j=1}^{|\mathcal{C}|} \sum_{\chi \subset S} H_{\chi}(x_{\chi}^i, v^j) \right) + n \ln Z(v)$$

Для чисельного пошуку оцінки гіббсовського розподілу методом максимальної вірогідності розглянемо алгоритм, побудований на

основі стохастичного квазіградієнтного методу розв'язку задач опуклого стохастичного програмування [2]. Цей метод є розвитком процедур випадкового пошуку і може розглядатися як узагальнення алгоритмів стохастичної апроксимації на багатовимірні негладкі задачі.

Згідно з [1] практичний алгоритм пошуку оцінки невідомих параметрів можна подати в наступному вигляді:

1. З ймовірністю $\frac{1}{n}$ на k -ому кроці ($v = v_k$) обирається одне з чисел $i = 1, \dots, n$, де n – кількість спостережень. Позначимо обране число через i_{fix} .
2. Для поточного значення параметру v_k за допомогою алгоритму породження марковських ланцюгів методом Метрополіса або Гіббса генерується скінченна вибірка $\{\bar{x}(v_k)\}$ конфігурацій гіббсовського поля довжини p .
3. Розраховується стохастичний квазіградієнт γ^k в точці v_k :

$$\gamma^k \approx \tilde{\nabla} L_n(x, v_k) = \left(\sum_{\chi \in S} \frac{\partial H_\chi(x_{\chi}^{i_{fix}}, v_k^j)}{\partial v_k^j} - \frac{1}{p} \sum_{q=1}^p \left(\sum_{\chi \in S} \frac{\partial H_\chi(\bar{x}_\chi^q, v_k^j)}{\partial v_k^j} \right) \right)_{j=1}^l$$

якщо функції $H_\chi(x, v)$ є диференційованими.

4. Знаходимо точку
$$v_{k+1} = \pi_\Theta \left(v_k - \frac{1}{k} \gamma^k \right)$$
. В загальному вигляді для цього необхідно розв'язати оптимізаційну задачу

$$v_{k+1} = \arg \min_{v \in \Theta} \left\| v_k - \rho_k \gamma^k - v \right\|^2,$$

проте якщо параметрична множина \cup компактна, то

$$v_{k+1} = v_k - \frac{1}{k} \gamma^k$$

5. Перевіряємо один вибраний критерій зупинки: $|v_{k+1} - v_k| < \varepsilon$, $|L_n(x, v_{k+1}) - L_n(x, v_k)| < \varepsilon$ або $|\nabla L_n(x, v_{k+1})| < \varepsilon$. Якщо він виконується то вважаємо, що істинне значення параметра $v^* = v_{k+1}$, інакше $k = k + 1$ і повертаємося до кроку 1.

Суттєвим недоліком запропонованого алгоритму є висока обчислювальна складність, тому що для кожної ітерації необхідно здійснювати процес породження марковського ланцюга, тобто генерувати повноцінну вибірку тестових конфігурацій гіббсовського поля. У випадку наявності обмежень на обчислювальні ресурси для вирішення цієї проблеми пропонується модифікувати алгоритм, здійснюючи апроксимацію математичного сподівання

$$E\left(\sum_{x \in S} \frac{\partial H_\chi(x, v)}{\partial v}\right)$$

на кожній ітерації лише за однією згенерованою вибіркою або ж породжувати допоміжні конфігурації через певну кількість кроків. Звичайно, отримана таким чином оцінка найшвидшого напрямку спуску до глобального мінімуму не буде оптимальною і, відповідно, алгоритм буде потребувати додаткових ітерацій, та цей недолік буде компенсуватися значним зменшенням обчислювального часу на кожній ітерації.

Литература.

1. Кнопов П.С., Самосьюнок О.С., Біла Г.Д. Модель розповсюдження інфекційних захворювань з прихованими носіями. – Кібернетика та системний аналіз. 2021. Том 57, №4. С. 166-176.
2. Ермольев Ю.М. Методы стохастического программирования / Ермольев Ю.М. — М.: Наука, 1976. — 240 с.