

Лабораторная работа 4

Исследование статистических свойств оценок семивариограммы

Исследовать статистические свойства классической и робастной оценок семивариограммы гауссовского случайного процесса.

Пусть $X(1), \dots, X(n)$ – n последовательных, полученных через равные промежутки времени наблюдений за случайным процессом. В качестве классической оценки семивариограммы рассмотрим статистику вида:

$$\tilde{\gamma}(h) = \frac{1}{2(n-h)} \sum_{s=1}^{n-h} (X(s+h) - X(s))^2,$$

$h = \overline{0, n-1}$. Положим $\tilde{\gamma}(-h) = \tilde{\gamma}(h)$, $h = \overline{0, n-1}$, и $\tilde{\gamma}(h) = 0$ для $|h| \geq n$.

В качестве робастной оценки семивариограммы рассмотрим статистику:

$$\hat{\gamma}(h) = \frac{\left(\frac{1}{(n-h)} \sum_{s=1}^{n-h} |X(s+h) - X(s)|^{1/2} \right)^4}{2 \left(0.457 + \frac{0.494}{n-h} + \frac{0.045}{(n-h)^2} \right)},$$

$h = \overline{0, n-1}$. Положим $\hat{\gamma}(-h) = \hat{\gamma}(h)$, $h = \overline{0, n-1}$, и $\hat{\gamma}(h) = 0$ для $|h| \geq n$.

- 1) Вычислить оценки семивариограммы и представить их графически.

Необходимо:

- Смоделировать n , $n = 100$, наблюдений за гауссовским стационарным случайным процессом с непрерывным временем и известной ковариационной функцией (можно воспользоваться результатами лабораторной работы 1).
- Найти вид семивариограммы процесса. Построить графики наблюдений, ковариационной функции и семивариограммы.
- Вычислить классическую и робастную оценки семивариограммы. На одном графике представить классическую, робастную оценки и истинную семивариограмму.
- Вычислить отклонения $\gamma(h) - \tilde{\gamma}(h)$ и $\gamma(h) - \hat{\gamma}(h)$ для $h = 0, \dots, n-1$. Сделать вывод. Отобразить графически.

- 2) Исследовать оценки семивариограммы на состоятельность в среднеквадратическом смысле.

Необходимо:

- Смоделировать 10 наборов наблюдений за гауссовским стационарным случайным процессом с непрерывным временем и известной ковариационной функцией. Число наблюдений n , соответствующих каждому набору, выбрать из таблицы 5 согласно варианту.
- Вычислить значения дисперсий классической и робастной оценок для требуемых n и h (см. таблицу вариантов).

$$D\{\tilde{\gamma}(h)\} = \frac{1}{2(n-h)^2} \sum_{s,t=1}^{n-h} (\gamma(s-t+h) + \gamma(s-t-h) - 2\gamma(s-t))^2,$$

$$\begin{aligned}
D\{\hat{\gamma}(h)\} &\approx \frac{10}{(n-h)^2} \left(\pi^{-\frac{1}{2}} - \frac{\Gamma^2\left(\frac{3}{4}\right)}{\pi} \right) \frac{\Gamma^6\left(\frac{3}{4}\right)}{\pi^3} \sum_{s,t=1}^{n-h} \frac{(\gamma(s-t+h) + \gamma(s-t-h) - 2\gamma(s-t))^2}{\left(0.457 + \frac{0.494}{n-h} + \frac{0.045}{(n-h)^2}\right)^2} = \\
&= \frac{10}{(n-h)^2} \frac{1}{\left(0.457 + \frac{0.494}{n-h} + \frac{0.045}{(n-h)^2}\right)^2} \left(\pi^{-\frac{1}{2}} - \frac{\Gamma^2\left(\frac{3}{4}\right)}{\pi} \right) \frac{\Gamma^6\left(\frac{3}{4}\right)}{\pi^3} \times \\
&\quad \times \sum_{s,t=1}^{n-h} (\gamma(s-t+h) + \gamma(s-t-h) - 2\gamma(s-t))^2,
\end{aligned}$$

где $\Gamma(\cdot)$ – гамма функция, т.е. $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$, $\alpha > 0$.

- Для каждого фиксированного h из таблицы вариантов построить графики зависимостей дисперсий $D\{\tilde{\gamma}(h)\}$ и $D\{\hat{\gamma}(h)\}$ от количества наблюдений n , представить таблицы соответствующих зависимостей. Сделать сравнительный анализ и вывод о состоятельности оценок.

Таблица 5. Исходные данные

Вариант 1	$n = 10, 25, 50, 75, 100, 150, 200, 250, 300, 350$; $h = 2, 5, 7$.
Вариант 2	$n = 15, 25, 50, 75, 100, 150, 200, 250, 300, 350$; $h = 1, 4, 8$.
Вариант 3	$n = 25, 50, 75, 100, 125, 150, 200, 250, 300, 350$; $h = 1, 7, 15$.
Вариант 4	$n = 30, 50, 75, 100, 125, 150, 200, 250, 300, 350$; $h = 2, 9, 18$.
Вариант 5	$n = 35, 60, 100, 125, 150, 200, 250, 300, 350, 400$; $h = 2, 10, 20$.
Вариант 6	$n = 40, 70, 100, 125, 150, 200, 250, 300, 350, 400$; $h = 1, 17, 31$.
Вариант 7	$n = 45, 80, 100, 125, 150, 200, 250, 300, 350, 400$; $h = 3, 8, 35$.
Вариант 8	$n = 50, 75, 100, 125, 150, 200, 250, 300, 350, 400$; $h = 2, 25, 40$.
Вариант 9	$n = 55, 75, 100, 125, 150, 200, 250, 300, 350, 450$; $h = 4, 30, 50$.
Вариант 10	$n = 15, 35, 70, 105, 130, 165, 200, 300, 350, 450$; $h = 2, 6, 12$.
Вариант 11	$n = 20, 40, 75, 100, 125, 150, 200, 250, 300, 350$; $h = 3, 11, 17$.
Вариант 12	$n = 40, 70, 100, 130, 160, 200, 250, 300, 380, 450$; $h = 5, 22, 35$.

- 3*)** Исследовать статистические свойства классической и робастной оценок семивариограммы гауссовского случайного процесса в случае нерегулярной сети наблюдения. Для этого выполнить пункты 1) и 2) задания для нерегулярного случая. Вид нерегулярности выбрать самостоятельно. Выполнить сравнительный анализ результатов.

Вспомогательная литература:

Cressie N. Statistics for Spatial Data / N. Cressie – New York: Wiley Classics Library, 2015. – 928 p.