Оценка вариограммы гауссовского случайного процесса

Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство.

Случайным процессом $x(t) = x(\omega, t)$ называется семейство случайных величин, заданных на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, где $\omega \in \Omega, t \in \mathbb{T}$, $\mathbb{T}$ — некоторое параметрическое множество.

Функцией распределения случайного процесса $x(t), t \in \mathbb{T}$, называется функция вида

$$F_n(x_1, \ldots, x_n; t_1, \ldots, t_n) = P\{x(t_1) < x_1, \ldots, x(t_n) < x_n\},$$

где $x_j \in \mathbb{R}, t_j \in \mathbb{T}, j = \overline{1, n}$.

Математическим ожиданием случайного процесса $x(t), t \in \mathbb{T}$ называется функция вида

$$m(t) = Ex(t) = \int\limits_{\mathbb{R}} x \, dF(x; t), t \in \mathbb{T}.$$

Дисперсией случайного процесса $x(t), t \in \mathbb{T}$ называется функция вида:

$$V(t) = V(x(t)) = E(x(t) - m(t))^2 = \int\limits_{\mathbb{R}} (x - m(t))^2 \, dF(x; t).$$

Корреляционной функцией случайного процесса $x(t), t \in \mathbb{T}$ называется функция вида:

$$r(t_1, t_2) = E(x(t_1)x(t_2)) = \iint\limits_{\mathbb{R}^2} x(t_1)x(t_2) \, dF(x(t_1), x(t_2); t_1, t_2)$$

Ковариационной функцией случайного процесса $x(t), t \in \mathbb{T}$ называется функция вида:

$$R(t_1, t_2) = E\{(x(t_1) - m(t_1))(x(t_2) - m(t_2))\} = \iint\limits_{\mathbb{R}^2} (x(t_1) - m(t_1))(x(t_2) - m(t_2)) \, dF(x(t_1), x(t_2); t_1, t_2)$$

Случайный процесс $X(s), s \in \mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$, называется внутренне стационарным, если справедливы следующие равенства:

$$E\{X(s_1) - X(s_2)\} = 0, \quad V\{X(s_1) - X(s_2)\} = 2\gamma(s_1 - s_2),$$

где $2\gamma(s_1 - s_2)$ — вариограмма рассматриваемого процесса, $s_1, s_2 \in \mathbb{Z}$.

Пусть $X(s), s \in \mathbb{Z}$ — внутренне стационарный гауссовский случайный процесс с нулевым математическим ожиданием, дисперсией σ^2 и неизвестной вариограммой.

$$2\gamma(h) = V\{X(s + h) - X(s)\}, \ s, h \in \mathbb{Z}.$$

В качестве оценки вариограммы рассмотрим статистику вида:

$$2\tilde{\gamma}(h) = \frac{1}{n - h} \sum_{s=1}^{n-h} (X(s + h) - X(s))^2, \quad h = \overline{0, n - 1}. \tag{1}$$

Вычислим математическое ожидание введённой оценки

$$\begin{gathered} E\{2\tilde{\gamma}(h)\} = \tfrac{1}{n-h} \textstyle\sum_{s=1}^{n-h} E(X(s + h) - X(s))^2 = \\ = [\text{так как процесс является внутренне стационарным}] = \\ = \tfrac{1}{n-h} \textstyle\sum_{s=1}^{n-h} 2\gamma(h) = 2\gamma(h). \end{gathered}$$

Таким образом оценка является несмещённой.

Далее, найдём второй момент:

$$cov(2\tilde{\gamma}(h_1), 2\tilde{\gamma}(h_2)) = E((2\tilde{\gamma}(h_1) - E(2\tilde{\gamma}(h_1)))(2\tilde{\gamma}(h_2) - E(2\tilde{\gamma}(h_2)))) =$$
$$= E(\tfrac{1}{n-h_1}\textstyle\sum_{s=1}^{n-h_1}((x(s+h_1) - x(s))^2 - E((x(s+h_1) - x(s))^2)) \times$$
$$\times \tfrac{1}{n-h_2}\textstyle\sum_{t=1}^{n-h_2}((x(t+h_2) - x(t))^2 - E((x(t+h_2) - x(t))^2))) =$$
$$= \tfrac{1}{(n-h_1)(n-h_2)}\textstyle\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2} cov((x(s+h_1) - x(s))^2, (x(t+h_2) - x(t))^2) =$$
$$= [\text{по определению } cov(a,b) = corr(a,b)\sqrt{V(a)V(b)}] =$$
$$= \tfrac{1}{(n-h_1)(n-h_2)}\textstyle\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2} corr((x(s+h_1) - x(s))^2, (x(t+h_2) - x(t))^2) \times$$
$$\times\sqrt{(V((x(s+h_1) - x(s))^2)V((x(t+h_2) - x(t))^2))} =$$
$$= [V\{X(s+h) - X(s)\}^2 = 2\{2\gamma(h)\}^2, \text{TODO: добавить ссылку на источник}] =$$
$$= \tfrac{2(2\gamma(h_1))(2\gamma(h_2))}{(n-h_1)(n-h_2)}\textstyle\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2} corr((x(s+h_1) - x(s))^2, (x(t+h_2) - x(t))^2) =$$
$$= [\text{по лемме 1 [\#Брест2005]}] =$$
$$= \tfrac{2(2\gamma(h_1))(2\gamma(h_2))}{(n-h_1)(n-h_2)}\textstyle\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2}\{corr((x(s+h_1) - x(s)), (x(t+h_2) - x(t)))\}^2 =$$
$$= \tfrac{2(2\gamma(h_1))(2\gamma(h_2))}{(n-h_1)(n-h_2)}\textstyle\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2}\{\tfrac{cov((x(s+h_1)-x(s)),(x(t+h_2)-x(t)))}{\sqrt{(V((x(s+h_1)-x(s)))V((x(t+h_2)-x(t))))}}\}^2 =$$
$$= [\text{по лемме 3 [\#Брест2005]}] =$$
$$= \tfrac{2(2\gamma(h_1))(2\gamma(h_2))}{(n-h_1)(n-h_2)}\textstyle\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2}\{\tfrac{\gamma(s+h_1-t)+\gamma(s-t-h_2)-\gamma(s+h_1-t-h_2)-\gamma(s-t)}{\sqrt{2\gamma(h_1)}\sqrt{2\gamma(h_2)}}\}^2$$

Таким образом получены соотношения математического ожидания и ковариации для статистики (1) $2\tilde{\gamma}(h)$:

$$E\{2\tilde{\gamma}(h)\} = 2\gamma(h), \tag{2}$$

$$cov(2\tilde{\gamma}(h_1), 2\tilde{\gamma}(h_2)) = \frac{2(2\gamma(h_1))(2\gamma(h_2))}{(n-h_1)(n-h_2)}\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2}\{\frac{\gamma(s+h_1-t) + \gamma(s-t-h_2) - \gamma(s+h_1-t-h_2) - \gamma(s-t)}{\sqrt{2\gamma(h_1)}\sqrt{2\gamma(h_2)}}\}^2. \tag{3}$$

Тогда, из (3) нетрудно получить соотношение для дисперсии оценки вариограммы $2\tilde{\gamma}(h)$, если положить $h_1 = h_2 = h$:

$$V\{2\tilde{\gamma}(h)\} = \tfrac{2(2\gamma(h))^2}{(n-h)^2}\textstyle\sum_{s=1}^{n-h}\sum_{t=1}^{n-h}\{\tfrac{\gamma(s+h-t)+\gamma(s-t-h)-\gamma(s+h-t-h)-\gamma(s-t)}{\sqrt{2\gamma(h)}\sqrt{2\gamma(h)}}\}^2 =$$
$$= \tfrac{2}{(n-h)^2}\textstyle\sum_{s,t=1}^{n-h}\{\gamma(s+h-t) + \gamma(s-t-h) - 2\gamma(s-t)\}^2.$$

Тем самым доказана следующая теорема.

Теорема 1. *Для оценки $2\tilde{\gamma}(h)$, представленной равенством (1), имеют место следующие соотношения:*

$$E\{2\tilde{\gamma}(h)\} = 2\gamma(h),$$

$$cov(2\tilde{\gamma}(h_1), 2\tilde{\gamma}(h_2)) = \frac{2(2\gamma(h_1))(2\gamma(h_2))}{(n-h_1)(n-h_2)}\sum_{s=1}^{n-h_1}\sum_{t=1}^{n-h_2}\{\frac{\gamma(s+h_1-t) + \gamma(s-t-h_2) - \gamma(s+h_1-t-h_2) - \gamma(s-t)}{\sqrt{2\gamma(h_1)}\sqrt{2\gamma(h_2)}}\}^2,$$

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

где $\gamma(h), h \in \mathbb{R},$ — семивариограмма процесса $X(s), s \in \mathbb{R},\ h, h_1, h_2 = \overline{0, n-1}$.