Russian Physics Journal, Vol. 40, No. 4, 1997

SPLINE APPROXIMATION FOR CORRELATION FUNCTION WITH MEASUREMENTS AT RANDOM TIMES

F. F. Idrisov

UDC 519.2

We use first-order splines to approximate correlation functions for measurement times forming Poisson or recurrent event flows.

The correlation function of a stationary random process is its most important characteristic, and research on a process usually begins by approximating this function.

Measurements of process values are usually made over equal time intervals. Sometimes, however, there are any of a number of reasons this is impossible, and measurements must be taken at random times. Under these conditions, the problem of approximating the correlation function is substantially more difficult.

In this paper we consider using first-order splines to approximate correlation functions when the measurement times form Poisson or recurrent event flows.

CONSTRUCTION OF APPROXIMATIONS FOR POISSON FLOWS OF MEASUREMENT TIMES

Let x(t) be a stationary Gaussian random process with mathematical expectation $M\{x(t)\} = 0$ and correlation function $R(\tau) = M\{x(t)x(t + \tau)\}$.

In order to approximate the correlation function, we consider the case in which the correlation function is of the form

$$R(\tau) = R_{\kappa-1}\frac{\kappa\tau_0 - \tau}{\tau_0} + R_\kappa\frac{\tau - (\kappa-1)\tau_0}{\tau_0},\ \kappa = \overline{1, n}, \tag{1}$$

over the interval $(\kappa-1)\tau_0 \le \tau \le \kappa\tau_0$, i.e., the correlation function has the form of a first-order spline. Of course, such a correlation function is hardly to be expected to occur in reality, but we need it to construct an approximation. In this case, approximating of the correlation function reduces to approximating the parameters $R_0, R_1, \ldots, R_n$.

Suppose that the times t_i at which measurements of the random process are made form a Poisson flow of events with constant intensity λ, and assume that the measurements of the process x(t) are conducted in the interval [0,T]. We denote the measured values by $x_i = x(t_i)$, $i = \overline{1, N}$. We should note that in this case, the number N of measurements is random.

Consider a recursive approximation for the parameters R_κ. The parameter R_0 can is usually approximated as

$$\hat{R}_0 = \frac{1}{N}\sum_{i=1}^{N} x_i^2. \tag{2}$$

Since $M\{x_i^2\} = R(0) = R_0$, we have $M\{\hat{R}_0\} = R_0$, i.e., this approximation is unbiased.

Suppose, now, we have constructed approximations $\hat{R}_0, \hat{R}_1, \ldots, \hat{R}_{\kappa-1}$, for the parameters $R_0, R_1, \ldots, R_{\kappa-1}$. We can now construct an approximation $\hat{R}_\kappa$ for the parameter R_κ.

Anzhero-Sudzhensk Affiliate of Tomsk Pedagogical University. Translated from Izvestiya Vysshikh Uchebnykh Zavedenii, Fizika, No. 4, pp. 32-38, April, 1997. Original article submitted July 2, 1996.

1064-8887/97/4004-0338$18.00

Let M_κ denote the set of indices (i, j) that satisfy the condition

$$M_\kappa=\{(i,\ j):(\kappa-1)\tau_0\le t_j-t_i\le\kappa\tau_0\}, \tag{3}$$

and consider a statistic of the form

$$S_\kappa=\sum_{i,j\in M_\kappa}x_i x_j\left[\gamma\frac{\kappa\tau_0-(t_j-t_i)}{\tau_0}+\frac{(t_j-t_i)-(\kappa-1)\tau_0}{\tau_0}\right] \tag{4}$$

with some still undetermined coefficient γ.

We will find the mathematical expectation for this statistic. Averaging over realizations of the process x(t) at fixed measurement times $\{t_i\}$, we obtain

$$M\{S_\kappa|\{t_i\}\}=\sum_{i,j\in M_\kappa}R(t_j-t_i)\left[\gamma\frac{\kappa\tau_0-(t_j-t_i)}{\tau_0}+\right.$$
$$\left.+\frac{(t_j-t_i)-(\kappa-1)\tau_0}{\tau_0}\right].$$

If we now average over the times t_i as in [1], we obtain

$$M\{S_\kappa\}=\lambda^2 T\int\limits_{(\kappa-1)\tau_0}^{\kappa\tau_0}R(\tau)\left[\gamma\frac{\kappa\tau_0-\tau}{\tau_0}+\frac{\tau-(\kappa-1)\tau_0}{\tau_0}\right]d\tau.$$

Substituting for $R(\tau)$ with (1) and evaluating the resultant integrals, we obtain

$$M\{S_\kappa\}=\lambda^2 T\tau_0\left[\left(\frac{\gamma}{3}+\frac{1}{6}\right)R_{\kappa-1}+\left(\frac{\gamma}{6}+\frac{1}{3}\right)R_\kappa\right].$$

In view of this, if we have an approximation $\hat{R}_{\kappa-1}$ for the parameter $R_{\kappa-1}$, we can obtain the approximation $\hat{R}_\kappa$ for the parameter R_κ with the equation

$$S_\kappa=\lambda^2 T\tau_0\left[\left(\frac{\gamma}{3}+\frac{1}{6}\right)\hat{R}_{\kappa-1}+\left(\frac{\gamma}{6}+\frac{1}{3}\right)\hat{R}_\kappa\right], \tag{5}$$

from which we obtain a recurrence relation for the approximation $\hat{R}_\kappa$ of the parameter R_κ:

$$\hat{R}_\kappa=-\frac{1+2\gamma}{2+\gamma}\hat{R}_{\kappa-1}+$$
$$+\frac{6}{2+\gamma}\frac{1}{\lambda^2 T\tau_0}\sum_{i,j\in M_\kappa}x_j x_i\left[\gamma\frac{\kappa\tau_0-(t_j-t_i)}{\tau_0}+\frac{(t_j-t_i)-(\kappa-1)\tau_0}{\tau_0}\right]. \tag{6}$$

We now consider the problem of choosing the parameter γ. Equation (6) is a first-order finite difference equation in $\hat{R}_\kappa$. Its solution is stable only if $|(1+2\gamma)/(2+\gamma)|<1$, which occurs when $-1<\gamma<1$.

It appears that the best value for γ will be one for which $(1+2)/(2+\gamma)=0$; which occurs when $\gamma=-1/2$. Now the approximation for the parameter R_κ takes the form

$$\hat{R}_\kappa=\frac{4}{\lambda^2 T\tau_0}\sum_{i,j\in M_\kappa}x_j x_i\left[\frac{t_j-t_i-(\kappa-1)\tau_0}{\tau_0}-\frac{\kappa\tau_0-(t_j-t_i)}{2\tau_0}\right], \tag{7}$$

and it generally does not contain $\hat{R}_{\kappa-1}$. Computation of its mathematical expectation shows that $M\{\hat{R}_\kappa\}=R_\kappa$, i.e., the approximation is unbiased. It can be shown that its dispersion decreases with $1/\lambda T$.

CONSTRUCTION OF APPROXIMATIONS FOR RECURRENT FLOWS OF MEASUREMENT TIMES

Suppose, now, that the measurement times t_i form a recurrent flow of events. This means that the $\tau_i = t_i - t_{i-1}$ are independent, identically distributed random variables with probability density $p(\tau)$.

Let $\bar{\tau} = M(\tau)$ and $\lambda = 1/\bar{\tau}$, and suppose g(s) is the Laplace transform of $p(\tau)$ and $\pi(\tau)$ is the inverse Laplace transform of the function $g(s)/(1 - g(s))$.

By analogy with (4), consider statistics of the form

$$S_\kappa = \sum_{l,j \in M_\kappa} x_j x_l \left[\gamma_\kappa \frac{\kappa\tau_0 - (t_j - t_l)}{\tau_0} + \frac{(t_j - t_l) - (\kappa - 1)\tau_0}{\tau_0} \right], \tag{8}$$

which differ from statistics of the form (4) in that the parameter γ now depends on κ. Its mathematical expectation for realizations of the process x(t) is of a similar form, and only the results of averaging over the measurement times t_i changes. As in [2], we can show that

$$M\{S_\kappa\} = \lambda T \int_{(\kappa-1)\tau_0}^{\kappa\tau_0} R(\tau) \left[\gamma_\kappa \frac{\kappa\tau_0 - \tau}{\tau_0} + \frac{\tau - (\kappa - 1)\tau_0}{\tau_0} \right] \pi(\tau)\, d\tau. \tag{9}$$

Substituting (1) for $R(\tau)$, we find that

$$M\{S_\kappa\} = \lambda T [R_{\kappa-1}(\gamma_\kappa A_\kappa + C_\kappa) + R_\kappa(\gamma_\kappa C_\kappa + B_\kappa)], \tag{10}$$

where

$$\begin{aligned} A_\kappa &= \int_{(\kappa-1)\tau_0}^{\kappa\tau_0} \left(\kappa - \frac{\tau}{\tau_0}\right)^2 \pi(\tau)\, d\tau, \\ B_\kappa &= \int_{(\kappa-1)\tau_0}^{\kappa\tau^0} \left(\frac{\tau}{\tau_0} - (\kappa - 1)\right)^2 \pi(\tau)\, d\tau, \\ C_\kappa &= \int_{(\kappa-1)\tau_0}^{\kappa\tau_0} \left(\kappa - \frac{\tau}{\tau_0}\right)\left(\frac{\tau}{\tau_0} - (\kappa - 1)\right) \pi(\tau)\, d\tau. \end{aligned} \tag{11}$$

As a result, we propose the following recursive approximation $\hat{R}_\kappa$ for the parameter R_κ:

$$\begin{aligned} \hat{R}_\kappa = &- \frac{\gamma_\kappa A_\kappa + C_\kappa}{\gamma_\kappa C_\kappa + B_\kappa} \hat{R}_{\kappa-1} + \\ &+ \frac{1}{\gamma_\kappa C_\kappa + B_\kappa} \frac{1}{\lambda T} \sum_{l,j \in M_\kappa} x_j x_l \left[\gamma_\kappa \frac{\kappa\tau_0 - (t_j - t_l)}{\tau_0} + \frac{(t_j - t_l) - (\kappa - 1)\tau_0}{\tau_0} \right]. \end{aligned} \tag{12}$$

This approximation is stable only if $|(\gamma_\kappa A_\kappa + C_\kappa)/(\gamma_\kappa C_\kappa + B_\kappa)| < 1$, which is possible when γ_κ belongs to the interval

$$-(B_\kappa + C_\kappa)/(A_\kappa + C_\kappa) < \gamma_\kappa < (B_\kappa - C_\kappa)/(A_\kappa - C_\kappa).$$

It is clear that the best value for γ_κ is $\gamma_\kappa = -C_\kappa/A_\kappa$, where the approximation $\hat{R}_\kappa$ contains no terms including $\hat{R}_{\kappa-1}$ and takes the form

$$\hat{R}_\kappa = \frac{A_\kappa}{A_\kappa B_\kappa - C_\kappa^2} \frac{1}{\lambda T} \sum_{i,j \in M_\kappa} x_j x_i \left[\frac{t_j - t_i - (\kappa - 1)\tau_0}{\tau_0} - \right.$$
$$\left. - \frac{C_\kappa}{A_\kappa} \frac{\kappa\tau_0 - (t_j - t_i)}{\tau_0} \right]. \tag{13}$$

It is an unbiased approximation for the parameter R_κ, i.e., its dispersion decreases in proportion to $1/\lambda T$.

CONSTRUCTION OF APPROXIMATIONS IN THE PRESENCE OF ERROR IN MEASUREMENT TIMES

We now consider the situation in which the times t_i at which the random process x(t) is measured are known only with some error.

Suppose that $\{t_i\}$ is a Poisson flow of events with constant intensity λ. We assume that we know not the times t_i themselves, but the quantities $\tau_i = t_i + \xi_i$, where the ξ_i are independent normal random variables with mathematical expectation $M\{\xi_i\} = 0$ and dispersion $D\{\xi_i\} = \sigma^2$. From now on, we assume that $\sigma << \tau_0$.

The fundamental difficulty presented by construction of approximations is that $t_j - t_i$ may lie outside the interval $(\kappa - 1)\tau_0 < t_i - t_i < \kappa\tau_0$, even when the times $\tau_j - \tau_i$ don't, i.e., when $(\kappa - 1)\tau_0 < \tau_j - \tau_i < \kappa\tau_0$. In order to assure that the quantities $t_j - t_i$ lie in the interval $[(\kappa - 1)\tau_0, \kappa\tau_0$, we reduce the size of the differences $\tau_j - \tau_i$ and consider the set

$$M_{\sigma\kappa} = \{j, i : (\kappa - 1)\tau_0 + g\sigma\sqrt{2} \leq \tau_j - \tau_i \leq \tau_0 - g\sigma\sqrt{2}\}. \tag{14}$$

If we take g = 1.96, we can assert that with probability 0.95, $t_j - t_i \in [(\kappa - 1)\tau_0, \kappa\tau_0]$; when g = 3.29, this probability rises to 0.999. Thus, we can chose g in the interval [2, 3].

In order to construct an approximation, we consider the function

$$\varphi_G(x) = \begin{cases} x, & \text{if} \quad G \leq x \leq 1 - G, \\ 0, & \text{if} \quad x < G \text{ or } x > 1 - G, \end{cases}$$

where $G = g\sigma\sqrt{2}/\tau_0$.

We now consider a statistic of the form

$$S_\kappa = \frac{1}{\lambda T} \sum_{i,j \in M_{\sigma\kappa}} x_j x_i \left[\gamma\varphi_G\left(\frac{\kappa\tau_0 - (\tau_j - \tau_i)}{\tau_0}\right) + \varphi_G\left(\frac{\tau_j - \tau_i - (\kappa - 1)\tau_0}{\tau_0}\right) \right]. \tag{15}$$

Averaging S_κ over realizations of the process x(t) with fixed t_i and τ_i, we obtain

$$M\{S_\kappa | \{t_i, \tau_i\}\} = \frac{1}{\lambda T} \sum_{i,j \in M_{\sigma\kappa}} \left[R_{\kappa-1} \frac{\kappa\tau_0 - (t_j - t_i)}{\tau_0} + \right.$$
$$\left. + R_\kappa \frac{t_j - t_i - (\kappa - 1)\tau_0}{\tau_0} \right] \left[\gamma\varphi_G\left(\frac{\kappa\tau_0 - (\tau_j - \tau_i)}{\tau_0}\right) + \right. \tag{16}$$
$$\left. + \varphi_G\left(\frac{\tau_j - \tau_i - (\kappa - 1)\tau_0}{\tau_0}\right) \right].$$

We now average this expression over the measurement times t_i and the errors ξ_i. For example, consider the quantity

$$A = \frac{1}{\lambda^2 T\tau_0} M\left\{ \frac{\kappa\tau_0 - (t_j - t_i)}{\tau_0} \varphi_G\left(\frac{\kappa\tau_0 - (\tau_j - \tau_i)}{\tau_0}\right) \right\}. \tag{17}$$

Since $\tau_j - t_i = t_j - t_i + (\xi_j - \xi_i)$, while ξ_i and ξ_j are independent, $\xi = \xi_j - \xi_i$ is a normal random variable with $M\{\xi\} = 0$ and $D\{\xi\} = 2\sigma^2$. As a result, averaging over $\{t_i\}$ and then over ξ yields

$$A = \frac{1}{\tau_0} \int_{-\infty}^{\infty} p(\xi)\, d\xi \int_{(\kappa-1)\tau_0}^{\kappa\tau_0} \frac{\kappa\tau_0 - \tau}{\tau_0}\, \varphi_G\left(\frac{\kappa\tau_0 - \tau - \xi}{\tau_0}\right) d\tau.$$

Making the substitution $(\kappa\tau_0 - \tau)/\tau_0 = x$, we obtain

$$A = \int_{-\infty}^{\infty} p(\xi)\, d\xi \int_0^1 x\varphi_G\left(x - \frac{\xi}{\tau_0}\right) dx.$$

In view of the form of φ_G, the substitution $x - \xi/\tau_0 = z$ yields

$$A = \int_{-\infty}^{\infty} p(\xi)\, d\xi \int_G^{1-G} \left(z + \frac{\xi}{\tau_0}\right) z dz =$$

$$= \int_G^{1-G} z^2 dz = \frac{1}{3}\left[(1-G)^3 - G^3\right],$$

since $M\{\xi\} = 0$.

Similarly, we can show that

$$\frac{1}{\lambda^2 T\tau_0} M\left\{\frac{t_j - t_i - (\kappa-1)\tau_0}{\tau_0}\, \varphi_G\left(\frac{\tau_j - \tau_i - (\kappa-1)\tau_0}{\tau_0}\right)\right\} = A,$$

$$\frac{1}{\lambda^2 T\tau_0} M\left\{\frac{\kappa\tau_0 - (t_j - t_i)}{\tau_0}\, \varphi_G\left(\frac{\tau_j - \tau_i - (\kappa-1)\tau_0}{\tau_0}\right)\right\} =$$

$$= \int_G^{1-G} z(1-z)\, dz = \frac{1}{2}\left[(1-G)^2 - G^2\right] - \frac{1}{3}\left[(1-G)^3 - G^3\right] = B.$$

Thus, finally,

$$M\{S_\kappa\} = \lambda^2 T\tau_0\left[R_{\kappa-1}(\gamma A + B) + R_\kappa(A + \gamma B)\right]. \tag{18}$$

As a result, we have the following approximation $\hat{R}_\kappa$ for the parameter R_κ:

$$\hat{R}_\kappa = -\frac{A\gamma + B}{A + \gamma B}\hat{R}_{\kappa-1} + \tag{19}$$

$$+ \frac{1}{A + \gamma B}\frac{1}{\lambda^2 T\tau_0} \sum_{i,j \in M_{\sigma l}} x_j x_i \left[\gamma\frac{\kappa\tau_0 - (\tau_j - \tau_i)}{\tau_0} + \frac{\tau_j - \tau_i - (\kappa-1)\tau_0}{\tau_0}\right].$$

It is clear that once again, it is best to choose γ to satisfy the condition $\gamma A + B = 0$, i.e.,

$$\gamma = -\frac{B}{A} = 1 - \frac{3}{2}\frac{(1-G)^2 - G^2}{(1-G)^3 - G^3}, \tag{20}$$

and then the approximation itself takes the form

$$\hat{R}_\kappa = \frac{A}{A^2 - B^2}\frac{1}{\lambda^2 T\tau_0} \sum_{i,j \in M_{\sigma\kappa}} x_j x_i \left[\frac{\tau_j - \tau_i - (\kappa-1)\tau_0}{\tau_0} - \right.$$

$$\left. - \frac{B}{A}\frac{\kappa\tau_0 - (\tau_j - \tau_i)}{\tau_0}\right]. \tag{21}$$

Similar results can also be obtained for recurrent event flows.

APPROXIMATION OF CORRELATION FUNCTIONS OF ARBITRARY FORM

Now, let the process x(t) be a stationary random process with arbitrary correlation function $R(\tau)$.

We divide the interval of values of τ that is of interest to us into segments of length τ_0, and we approximate the parameters of R_κ by using the same formulas as above. We choose an approximation of the following for the desired correlation function $R(\tau)$:

$$\hat{R}(\tau) = \hat{R}_{\kappa-1} \frac{\kappa\tau_0 - \tau}{\tau_0} + \hat{R}_\kappa \frac{\tau - (\kappa - 1)\tau_0}{\tau_0}. \tag{22}$$

As a result, we obtain a spline approximation for the correlation function in which we are interested. Now, however, even when $\lambda T \to \infty$ there is an error that does not vanish, because the correlation function is not a spline. We will show how to compute the limiting error in the approximation of an arbitrary correlation function by a first-order spline.

Suppose that the measurement times form a Poisson event flow and they are known exactly. We find

$$R_\kappa = \frac{4}{\tau_0} \int\limits_{(\kappa-1)\tau_0}^{\kappa\tau_0} R(\tau) \left[\frac{\tau - (\kappa - 1)\tau_0}{\tau_0} - \frac{\kappa\tau_0 - \tau}{2\tau_0}\right] d\tau.$$

It can be shown that in approximating the parameters $\hat{R}_\kappa$ with algorithm (7), $\hat{R}_\kappa$ converges to R_κ as $\lambda T \to \infty$, at least in the sense of mean square convergence.

The limiting mean square error in approximation of the correlation function can be represented in the form

$$E_n = \frac{1}{n\tau_0} \sum_{\kappa=1}^{n} \int\limits_{(\kappa-1)\tau_0}^{\kappa\tau_0} \left(R(\tau) - \left[R_{\kappa-1} \frac{\kappa\tau_0 - \tau}{\tau_0} + R_\kappa \frac{\tau - (\kappa - 1)\tau_0}{\tau_0}\right]\right)^2 d\tau. \tag{23}$$

The concrete form of $R(\tau)$ can be used to find an explicit expression for the limiting mean square error in approximation of the correlation function.

REFERENCES

1. T. M. Kulikova, Izv. Vuzov. Fiz., No. 4, 23-29 (1996).
2. F. F. Idrisov and V. N. Tkachenko, Izv. Vuzov. Fiz., No. 2, 55-66 (1994).