Bandwidth selection for estimating the two-point correlation function of a spatial point pattern using AMSE

Woncheol Jang and Ji Meng Loh

Seoul National University and New Jersey Institute of Technology

Supplementary Material

To show (2.1), it suffices to show

$$\mathsf{E}\left(\frac{DD}{RR}\right) - \frac{\mathsf{E}\,(DD)}{\mathsf{E}\,(RR)} = O\left(\frac{1}{h^2|W_0|}\right).$$

Using a Taylor expansion,

$$\begin{aligned}\frac{DD}{RR} &= \frac{\mathsf{E}\,(DD)}{\mathsf{E}\,(RR)} + \frac{1}{\mathsf{E}\,(RR)}(DD - \mathsf{E}\,(DD)) - \frac{\mathsf{E}\,(DD)}{(\mathsf{E}\,(RR))^2}(RR - \mathsf{E}\,(RR)) \\ &\quad + \frac{2\mathsf{E}\,(DD)}{(\mathsf{E}\,(RR))^3}(RR - \mathsf{E}\,(RR))^2 - \frac{1}{(2\mathsf{E}\,(RR))^2}(DD - \mathsf{E}\,(DD))(RR - \mathsf{E}\,(RR)) + R_n.\end{aligned}$$

where R_n is the remainder term. By taking expectation on both sides,

$$\mathsf{E}\left(\frac{DD}{RR}\right) - \frac{\mathsf{E}\,(DD)}{\mathsf{E}\,(RR)} = \frac{2\mathsf{E}\,(DD)}{\mathsf{E}\,(RR)}\frac{\mathsf{Var}\,(RR)}{(\mathsf{E}\,(RR))^2} - \frac{1}{2}\frac{\mathsf{Cov}\,(DD, RR)}{(\mathsf{E}\,(RR))^2} + \mathsf{E}\,(R_n).$$

The second term on the RHS is zero since DD and RR are independent.

The first term of the RHS can be expressed as follows

$$2\frac{\mathsf{E}\,(DD)}{\mathsf{E}\,(RR)}\frac{\mathsf{Var}\,(RR)}{(\mathsf{E}\,(RR))^2} = 2\frac{\mathsf{Var}\,(RR)}{(\mathsf{E}\,(RR))^2}\left(g(r_0) + O(h^2)\right).$$

Recall

$$(\mathsf{E}\,(RR))^2 = 4h^2\lambda^4\left(|W_0|^2 + \frac{2}{3}|W_0||W_0''|h^2)\right).$$

On the other hand,

$$\begin{aligned}
(RR)^2 &= (2h)^2\left\{\sum_x\sum_{y\neq x} K_h(r_0 - |x-y|)\right\}^2\\
&= 4h^2\left[\sum_x\sum_{y\neq x}\sum_{x'\neq x}\sum_{y'\neq x'} K_h(r_0-|x-y|)K_h(r_0-|x'-y'|)\right]\\
&\quad + 4h^2\left[\sum_x\sum_{y\neq x}\sum_{\substack{y'\neq x\\ y'\neq y}} K_h(r_0-|x-y|)K_h(r_0-|x-y'|)\right]\\
&\quad +4h^2\left[\sum_x\sum_{y\neq x}\sum_{x'\neq y, x'\neq x} K_h(r_0-|x-y|)K_h(r_0-|x'-y|)\right] + 4h^2\left[\sum_x\sum_{y\neq x} K_h^2(r_0-|x-y|)\right]
\end{aligned}$$

Note that the expectation of the first term on the RHS is $(\mathsf{E}\,(RR))^2$. Hence $Var(RR)$ is the sum of the remaining three terms on the RHS.

Recall $K(t) = \frac{1}{2}I\{|t| \leq 1\}$. Hence $K^2(t) = \frac{K(t)}{2}$. The second term can be expressed as follows:

$$\begin{aligned}
&4h^2\lambda^2\int_W\int\int K_h(r_0-r)K_h(r_0-r')C_x(r)C_x(r')\mathrm{d}r\mathrm{d}r'\mathrm{d}x\\
=\ & 4h^2\lambda^2\int_W\int\int K(t)K(s)C_x(r_0+th)C_x(r_0+sh)\mathrm{d}t\mathrm{d}s\mathrm{d}x\\
=\ & 4h^2\lambda^2\int_W\int\int K(t)K(s)\left(C_x(r_0)+\frac{(th)^2}{2}C_x''(r_0)\right)\left(C_x(r_0)+\frac{(sh)^2}{2}C_x''(r_0)\right)\mathrm{d}t\mathrm{d}s\mathrm{d}x\\
=\ & 4h^2\lambda^2|\widetilde{W}_0| + \frac{8}{3}h^4\lambda^2C_x''(r_0)C_x(r_0),
\end{aligned}$$

where $|\widetilde{W_0}| = \int_W C_x^2(r_0)\mathrm{d}x = O(|W_0|)$. Similarly, the third term is of the same order.

Now, we can show the order of the last term on the RHS as follows.

$$\begin{aligned} 4h^2\lambda^2 \int_w \int K_h^2(r_0 - r)(C_x(r))\mathrm{d}r\mathrm{d}x &= 4h\lambda^2 \int_w \int K^2(t) \left(C_x(r_0) + \frac{t^2h^2}{2} C_x''(r_0) \right) \mathrm{d}r\mathrm{d}x \\ &= 2h\lambda^2 \int_w \int K(t) \left(C_x(r_0) + \frac{t^2h^2}{2} C_x''(r_0) \right) \mathrm{d}r\mathrm{d}x \\ &= 2h|W_0|\lambda^2 + \frac{2}{3}h^3|W_0''|\lambda^2. \end{aligned}$$

Therefore the leading term of $Var(RR)$ is of order $O(|W_0|)$. We conclude that

$$\frac{\mathsf{Var}\,(RR)}{(\mathsf{E}\,(RR))^2} = O\left(\frac{1}{h^2|W_0|}\right).$$

Furthermore, using the above result and the Chebyshev inequality,

$$\left|\frac{RR}{\mathsf{E}\,(RR)} - 1\right| = O_p\left(h^{-1}|W_0|^{-1/2}\right).$$

Since DD and RR are independent, it is straightforward to show the reminder term $R_n = \left(h^{-3}|W_0|^{-3/2}\right)$.

Therefore,

$$\mathsf{E}\left(\frac{DD}{RR}\right) - \frac{\mathsf{E}\,(DD)}{\mathsf{E}\,(RR)} = O\left(\frac{1}{h^2|W_0|}\right)\left(g(r_0) + O(h^2)\right) + O\left(h^{-3}|W_0|^{-3/2}\right) = O\left(\frac{1}{h^2|W_0|}\right).$$