Multivariate spatial nonparametric modelling via kernel processes mixing

Montserrat Fuentes and Brian Reich

Supplementary material

A.1

Properly defined process prior

Kolmogorov existence theorem

We need to prove that the collection of finite-dimensional distributions introduced in (1) define a stochastic process $Y(s)$. We use the two Kolmogorov consistency conditions (symmetry under permutation, and dimensional consistency) to show that (1) defines a proper random process for Y.

Proposition 1. The collection of finite-dimensional distributions introduced in (1) properly define a stochastic process $Y(s)$. We use the two Kolmogorov consistency conditions: symmetry under permutation, and dimensional consistency, to define properly a random process for Y.

Proof of Proposition 1:

Symmetry under permutation.

Let $p_{i_j} = p(Y(s_j) = X(\phi_{i(s_j)}))$, where $\phi_{i(s_j)}$ is the centering knot of the kernel $i(s_j)$ in the representation of $F_{s_j}(Y)$ in (1), and let ϕ_{i_j} be an abbreviation for $\phi_{(i(s_j))}$. Then, $p_{i_1,\ldots,i_n}$ determine the site-specific joint selection probabilities. If $\pi(1),\ldots,\pi(n)$ is any permutation of $\{1,\ldots,n\}$, then we have

$$
\begin{aligned}
& p_{i_{\pi(1)},\ldots,i_{\pi(n)}} = p(Y(s_{\pi(1)}) = X(\phi_{\pi(i_1)}),\ldots,Y(s_{\pi(n)}) = X(\phi_{\pi(i_n)})) \\
= \; & p(Y(s_1) = X(\phi_{i_1}),\ldots,Y(s_n) = X(\phi_{i_n})) = p_{i_1,\ldots,i_n},
\end{aligned}
\tag{12}
$$

since the observations are conditionally independent. Then,

$$
\begin{aligned}
& p(Y(s_1) \in A_1,\ldots,Y(s_n) \in A_n) \\
= & \sum_{i_1,\cdots,i_n} p(Y(s_1) = X(\phi_{i_1}),\ldots,Y(s_n) = X(\phi_{i_n}))\delta_{X(\phi(i_1))}(A_1)\ldots\delta_{X(\phi(i_n))}(A_n) \\
= & \sum_{i_1,\cdots,i_n} p_{i_1,\ldots,i_n}\delta_{X(\phi(i_1))}(A_1)\ldots\delta_{X(\phi(i_n))}(A_n) \\
= & \sum_{i_1,\cdots,i_n} p_{i_{\pi(1)},\ldots,i_{\pi(n)}}\delta_{X(\phi(i_{\pi(1)}))}(A_{\pi(1)})\ldots\delta_{X(\phi(i_{\pi(n)}))}(A_{\pi(n)}) \\
= & \; p(Y(s_{\pi(1)}) \in A_{\pi(1)},\ldots,Y(s_{\pi(n)}) \in A_{\pi(n)}),
\end{aligned}
$$

and, the *symmetry under permutation* condition holds.

Dimensional consistency.

$$
\begin{aligned}
& p(Y(s_1) \in (A_1),\ldots,Y(s_k) \in \mathcal{R},\ldots,Y(s_n) \in (A_n)) \\
= & \sum_{(i_1,\ldots,i_n)\in\{1,2,\ldots\}^n} p_{i_1,\ldots,i_n}\delta_{X(\phi_{i_1})}(A_1)\cdots\delta_{X(\phi_{i_k})}(\mathcal{R})\cdots\delta_{X(\phi_{i_n})}(A_n) \\
= & \sum_{(i_1,\ldots,i_{k-1},i_{k+1},\ldots,i_n)\in\{1,2,\ldots\}^{n-1}} \delta_{X(\phi_{i_1})}(A_1)\cdots\delta_{X(\phi_{i_{k-1}})}(A_{k-1})\delta_{X(\phi_{i_{k+1}})}(A_{k+1}) \\
& \cdots\delta_{X(\phi_{i_n})}(A_n)\sum_{j=1}^{\infty} p_{i_1,\ldots,i_{k-1},j,i_{k+1},\ldots,i_n} \\
= & \; p(Y(s_1) \in (A_1),\ldots,Y(s_{k-1}) \in A_{k-1},Y(s_{k+1}) \in A_{k+1},\ldots,Y(s_n) \in (A_n)).
\end{aligned}
\tag{13}
$$

In (13), we need

$$
p_{i_1,\ldots,i_{k-1},i_{k+1},\ldots,i_n} = \sum_{j=1}^{\infty} p_{i_1,\ldots,i_{k-1},j,i_{k+1},\ldots,i_n}
$$

which holds by Fubini Theorem and the fact that X is a properly defined Gaussian process.

A.2

<u>*Proof of Theorem 1.*</u>

The covariance function C of the underlying process X has a first order derivative, C'. We introduce a Taylor expansion for C with a Lagrange remainder term,

$$C(|\phi_{i_1} - \phi_{i_2}|) = C(|s - s'|) + C'(\psi_{i_1,i_2})\varepsilon_{i_1,i_2}, \tag{14}$$

where $\varepsilon_{i_1,i_2} = (|\phi_{i_1} - \phi_{i_2}| - |s - s'|)$ and ψ_{i_1,i_2} is in between $|s - s'|$ and $|\phi_{i_1} - \phi_{i_2}|$.

Assuming that s and s' lie on the support of the kernels K_{i_1} and K_{i_2}, respectively, i.e. $|\phi_{i_1} - s| < \epsilon_{i_1}$ and $|\phi_{i_2} - s'| < \epsilon_{i_2}$. We have,

$$\varepsilon_{i_1,i_2} \leq ||\phi_{i_1} - \phi_{i_2}| - |s - s'|| \leq |(\phi_{i_1} - \phi_{i_2}) - (s - s')| \leq \epsilon_{i_1} + \epsilon_{i_2} \leq 2\epsilon,$$

and,

$$\varepsilon_{i_1,i_2} \geq -||\phi_{i_1} - \phi_{i_2}| - |s - s'|| \geq -|(\phi_{i_1} - \phi_{i_2}) - (s - s')| \geq -(\epsilon_{i_1} + \epsilon_{i_2}) \geq -2\epsilon,$$

Thus, $-2\epsilon \leq \varepsilon_{i_1,i_2} \leq 2\epsilon$.

Let $p(s)$ be the potentially infinite vector with all the probabilities masses $p_i(s)$ in $F_s(Y)$. The conditional covariance of the data process Y is written in terms of the covariance C of X,

$$\text{cov}(Y(s), Y(s')|p(s), p(s'), C) = \sum_{i_1 i_2} p_{i_1}(s)p_{i_2}(s')C(|\phi_{i_1} - \phi_{i_2}|),$$

since the kernels all have compact support, the expression above is the same as

$$\sum_{i_1,i_2;|\phi_{i_1}-s|<\epsilon_{i_1},|\phi_{i_2}-s'|<\epsilon_{i_2}} p_{i_1}(s)p_{i_2}(s')C(|\phi_{i_1} - \phi_{i_2}|).$$

Using the Taylor approximation in (14), the $\mathrm{cov}(Y(s), Y(s')|p(s), p(s'), C)$ can be written

$$C(|s-s'|)\left[\sum_{i_1} p_{i_1}(s)\right]\left[\sum_{i_2} p_{i_2}(s')\right] + \sum_{i_1,i_2;|\phi_{i_1}-s|<\epsilon_{i_1},|\phi_{i_2}-s'|<\epsilon_{i_2}} p_{i_1}(s)p_{i_2}(s')C'(\psi_{i_1,i_2})\varepsilon_{i_1,i_2}.$$

Since C' is nonnegative and $\varepsilon_{i_1,i_2} \in (-2\epsilon, 2\epsilon)$, we have that for $J_{i_1,i_2} = \{(i_1, i_2); |\phi_{i_1} - s| < \epsilon_{i_1}, |\phi_{i_2} - s'| < \epsilon_{i_2}\}$,

$$\begin{aligned} -2\epsilon \sum_{i_1,i_2 \in J_{i_1,i_2}} p_{i_1}(s)p_{i_2}(s')C'(\psi_{i_1,i_2}) &\leq \sum_{(i_1,i_2)\in J_{i_1,i_2}} p_{i_1}(s)p_{i_2}(s')C'(\psi_{i_1,i_2})\varepsilon_{i_1,i_2} \\ &\leq 2\epsilon \sum_{(i_1,i_2)\in J_{i_1,i_2}} p_{i_1}(s)p_{i_2}(s')C'(\psi_{i_1,i_2}), \end{aligned} \tag{15}$$

where,

$$2\epsilon \sum_{i_1,i_2 \in J_{i_1,i_2}} p_{i_1}(s)p_{i_2}(s')C'(\psi_{i_1,i_2}) \longrightarrow_{\epsilon\to 0} 0,$$

because C' is bounded and the sum of probability masses is always bounded by 1. The sum of probability masses would be always bounded by 1, because by proposition 1 the prior process is properly defined.

Thus, we obtain

$$\mathrm{cov}(Y(s), Y(s')|p(s), p(s'), C) \longrightarrow_{\epsilon\to 0} C(|s-s'|)\left[\sum_{i_1} p_{i_1}(s)\right]\left[\sum_{i_2} p_{i_2}(s')\right] = C(|s-s'|).$$

Therefore, the conditional covariance of the data process Y approximates the covariance C of the underlying process X as the bandwiths of the kernel functions go to zero.

A.3

Proof of Theorem 2.

The cross-covariance function $C_{1,2}(s, s')$ of the underlying process $X = (X_1, X_2)$ has first order partial derivatives $\delta C_{1,2}(s, s')/\delta s$ and $\delta C_{1,2}(s, s')/\delta s'$. We introduce a Taylor expansion for $C_{1,2}$ with a Lagrange remainder term,

$$C_{1,2}(\phi_{i_1},\phi_{i_2}) = C_{1,2}(s,s')+(\phi_{i_1}-s)\left[\delta C_{1,2}(s,s')/\delta s\right]_{(s,s')=(\psi_{i_1},\psi_{i_2})}+(\phi_{i_2}-s')\left[\delta C_{1,2}(s,s')/\delta s'\right]_{(s,s')=(\psi_{i_1},\psi_{i_2})}$$

where ψ_{i_1} is in between s and ϕ_{i_1}, and ψ_{i_2} is in between s' and ϕ_{i_2}. We follow the same steps as in Theorem 1, to bound the first order term of the Taylor expansion, and obtain that

$$\mathrm{cov}(Y_1(s),Y_2(s')|p_1(s),p_2(s'),C_{1,2}) \longrightarrow_{\epsilon\to 0} C_{1,2}(s,s'),$$

where $p_1(s)$, and $p_2(s')$ are the potentially infinite dimensional vectors with the all probability masses in the spatial stick-breaking prior processes $F_s(Y_1)$, and $F_{s'}(Y_2)$ respectively.

A.4

Proof of Theorem 3.

Let $\psi(t,s)$ be the characteristic function of $Y(s)$. Then,

$$\begin{aligned} &\psi(t,s_1)-\psi(t,s_2) = E_Y[\exp\{itY(s_1)\}] - E_Y[\exp\{itY(s_2)\}] \\ =\ & E_X\left\{\sum_j p_j(s_1)\exp\{itX(\phi_j)\}\right\} - E_X\left\{\sum_j p_j(s_2)\exp\{itX(\phi_j)\}\right\} \\ =\ & E_X\left\{\sum_j (p_j(s_1)-p_j(s_2))\exp\{itX(\phi_j)\}\right\} \longrightarrow_{|s_1-s_2|\to 0} 0. \qquad (16)\end{aligned}$$

Then, $F_{s_1}(Y)$ converges to $F_{s_2}(Y)$ for any locations s_1, s_2, as long as $|s_1-s_2|\to 0$.

A.5

Proof of Theorem 4.

The probability masses $p_i(s)$ in (1) are $p_i(s) = V_iK_i(s)\prod_{j=1}^{i-1}(1-V_jK_j(s))$. Since the bandwiths ϵ_i converge uniformly to zero, then, $p_i(s)\to 1$, as $|\phi_i - s|\to 0$, where ϕ_i is the knot of kernel K_i. This holds because $\sum_j p_j(s) = 1$ a.s. (since the process Y is properly defined).

Assume now $|s_1 - s_2| \to 0$, we need to prove that $Y(s_1)$ converges a.s. to $Y(s_2)$.

Let ϕ_1 and ϕ_2 satisfy,

$$|\phi_1 - s_1| \to 0, \quad \text{and} \quad |\phi_2 - s_2| \to 0. \tag{17}$$

Thus, we obtain that with probability 1, $Y(s_1)$ converges to $X(\phi_1)$, and $Y(s_2)$ to $X(\phi_2)$.

Since $|s_1 - s_2| \to 0$, and $|\phi_1 - \phi_2| \leq |\phi_1 - s_1| + |\phi_2 - s_2| + |s_1 - s_2|$. Then, by (17)

$$|\phi_1 - \phi_2| \to 0. \tag{18}$$

We have,

$$|Y(s_1) - Y(s_2)| \leq |Y(s_1) - X(\phi_1)| + |Y(s_2) - X(\phi_2)| + |X(\phi_1) - X(\phi_2)|,$$

where $|Y(s_1) - X(\phi_1)| \to 0$ a.s., as $|s_1 - \phi_1| \to 0$; $|Y(s_2) - X(\phi_2)| \to 0$ a.s., as $|s_2 - \phi_2| \to 0$; and since X is a.s. continuous, $|X(\phi_1) - X(\phi_2)| \to 0$ a.s., as $|\phi_1 - \phi_2| \to 0$ (which holds by 18).

Therefore, $|Y(s_1) - Y(s_2)| \to 0$ a.s.