Spatio-Temporal Covariance Functions

Jack Ruder

October 20, 2025

Intro, Matern covariance

Why use the Matern covariance function?

$$C_\nu(d) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{d}{\rho} \right)^\nu K_\nu \left(\sqrt{2\nu} \frac{d}{\rho} \right) \tag{1}$$

Let's break down each of these components.

Bessel's Equation

$$C_\nu(d) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{d}{\rho}\right)^\nu K_\nu\left(\sqrt{2\nu}\frac{d}{\rho}\right) \tag{2}$$

Extremely important PDEs include the Laplace equation, the wave equation, the heat equation, and the Poisson equation. These all are solved by separation of variables, and in polar coordinates, the separation into two ODEs yields the ODE

$$w''(z)z^2 + w'(z)z + (z^2 - v^2)w(z) = 0.$$

Solutions

Solutions to the ODE are called Bessel functions. A series solution via Bessel (1816) has the form

$$w(z) = z^{\upsilon} \sum_{j=0}^{\infty} a_j z^j + z^{-\upsilon} \sum_{j=0}^{\infty} b_j z^j.$$

This yields two orthogonal solutions, one of which is

$$J_{-\upsilon}(z) = z^{-\upsilon} \sum_{k=0}^{\infty} B_k z^{2k}$$

for Fourier-Bessel coefficients

$$B_k = \frac{(-1)^k}{2^{\nu-2k}\Gamma(k-\upsilon+1)k!}.$$

The boundedness of the coefficients depends on ν, so ν is effectively controlling smoothness of the solution. Note: The other solution is given by replacing $-\nu$ with ν.

Modified Bessel Functions

When ν is an integer, the two solutions are not linearly independent. To resolve this, the solution set is extended to include

$$Y_\nu(z) = \lim_{\mu \to \nu} \frac{\cos(\mu\pi) J_\mu(z) - J_{-\mu}(z)}{\sin(\mu\pi)}.$$

where $J_\nu(z)$ and $Y_\nu(z)$ form a basis for solutions to the Bessel equation, for any ν.

Modified Bessel Functions, continued

Now, replacing z with iz gives the modified Bessel equation, which has solutions $w(z) = C_1 I_\nu(z) + C_2 K_\nu(z)$
The corresponding $K_\nu(z) = \frac{\pi}{2}\frac{I_{-\nu}(z) - I_\nu(z)}{\sin(\nu\pi)}$ where
$I_\nu(z) = i^{-\nu} J_\nu(iz)$.
We *can* write down the formula for $K_\nu(z)$, though it is unwieldy and varies depending on the value of ν. When ν is real, $K_\nu(z)$ is real as well, and for any ν, $K_\nu(z) = K_{-\nu}(z)$. The half integers give particularly simple forms:

$$K_{1/2}(z) = \sqrt{\frac{\pi}{2z}}e^{-z}, \quad K_{3/2}(z) = \sqrt{\frac{\pi}{2z}}e^{-z}\left(1 + \frac{1}{z}\right).$$

Modified Bessel Function of the Second Kind in Matern

$$C_\nu(d) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{d}{\rho}\right)^\nu K_\nu\left(\sqrt{2\nu}\frac{d}{\rho}\right) \tag{3}$$

We have $K_\nu\left(\sqrt{2\nu}\frac{d}{\rho}\right)$ as a solution to the modified Bessel equation.

- Useful, exponentially decaying function
- Decay depends on ν, but also impacts shape/smoothness
- d/ρ is the distance, scaled by
- $\sqrt{2\nu}$ standardizes the distance relative to d/ρ, accounting for $2 - D$ space.
- As $\nu \to \infty$, we recover the Gaussian covariance function.

Matern Covariance, other components

$$C_\nu(d) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{d}{\rho}\right)^\nu K_\nu\left(\sqrt{2\nu}\frac{d}{\rho}\right) \tag{4}$$

- σ^2 is overall variance, giving vertical spread
- $\frac{2^{1-\nu}}{\Gamma(\nu)}$ is a normalizing constant ensuring $C_\nu(0) = \sigma^2$
- $\left(\sqrt{2\nu}\frac{d}{\rho}\right)^\nu$ scales the covariance (the bessel function decays more with ν, so this term compensates for that). This term dominates at small d.

- Fuentes, Chen, Davis (2007). "A class of nonseparable and nonstationary spatial temporal covariance functions"
- Motivation: Ozone data modeling around early 2000s used stationary models, or considered location-independent models.
- The most prevelant/useful models made use of separable covariance functions, but these fail to capture space-time interactions when using simpler models which do not encapsulate these interactions.
- Problem: Ozone data is generally non-separable (ex. poles vs equator have different dynamics).
- Goal: handle ST interactions via the covariance function.

- The process $Z(\mathbf{x}, t)$ is separable if $\text{cov}(Z(\mathbf{x}, t), Z(\mathbf{x}', t')) = C_S(\mathbf{x}, \mathbf{x}')C_T(t, t')$ where C_S is a spatial covariance function and C_T is a temporal covariance function.
- If $Z(\mathbf{x}, t)$ is separable, then model each component independently (build a model for space, and a model for time). Advantage: computationally easier, and can use a plethora of existing spatial and temporal models.

Let $\{Z(\mathbf{x}, t), x \in D \subset \mathbb{R}^D, t \in T \subset \mathbb{R}\}$ be a *stationary* random field with spatial-temporal covariance function.
The typical representation for the joint process is

$$Z(\mathbf{x}, t) = \int_{\mathbb{R}_d} \int_{\mathbb{R}} \exp(i\omega^T \mathbf{x} + i\tau t) dY(\omega, \tau),$$

where the (Fourier-Stieltjes) integral is taken over the random measure $Y(\omega, \tau)$ with uncorrelated increments and complex symmetry, other than the constraint $dY(\omega, \tau) = dY^c(-\omega, \tau)$.

$Y(\omega, \tau)$

- ω is the spatial frequency, and τ is the temporal frequency.
- Y is a complex-valued random measure, representing the amplitude and phase of each frequency component.
- Y^c is the complex conjugate of Y
- Uncorrelated increments means that for any two disjoint sets $A, B \subset \mathbb{R}^d \times \mathbb{R}$,

$$E\left[dY(\omega,\tau)dY^c(\omega',\tau')\right] = 0, \quad (\omega,\tau) \in A, (\omega',\tau') \in B.$$

- Complex symmetry means that

$$E\left[dY(\omega,\tau)dY^c(\omega',\tau')\right] = E\left[dY(-\omega,-\tau)dY^c(-\omega',-\tau')\right].$$

$C(\mathbf{x}, t)$

Using the spectral representation of $Z(\mathbf{x}, t)$, the covariance function is

$$C(\mathbf{x}, t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\omega^T \mathbf{x} + i\tau t) F(d\omega d\tau),$$

where F is known as the spectral measure,

$$F(\omega, \tau) = \mathbb{E}[Y(\omega, \tau) Y^c(\omega, \tau)] = E[|Y(\omega, \tau)|^2].$$

If F has a density $f(\omega, \tau)$ with respect to Lebesgue measure, then f is the Fourier transform of $C(\mathbf{x}, t)$, and

$$C(\mathbf{x}, t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} \exp(i\omega^T \mathbf{x} + i\tau t) f(\omega, \tau) d\omega d\tau.$$

Separability of spectral density

Note: if $f(\omega, \tau) = f_1(\omega) f_2(\tau)$, then the covariance function is separable! Separability in the covariance domain corresponds to factorization in the spectral domain.

Ex: Matern spectral density

The spectral density of the (non-reparameterized) Matern covariance function is

$$f(\omega, \tau) = f(\omega) = \gamma(\alpha^2 + |\omega|^2)^{-\nu - \frac{d}{2}}$$

- This is MUCH nicer than working with a Bessel function
- But, we have to FFT to get the covariance function back.
- Low frequencies (small $|\omega|$) correspond to large-scale, smooth spatial features
- High frequencies (large $|\omega|$) indicate lots of small-scale spatial variation.
- All frequencies decay as $|w|^{-2\nu-d}$, so high frequencies are more heavily penalized for larger ν.

Specification of C via f

Let $f(\omega, \tau) = \gamma(\alpha^2\beta^2 + \beta^2|\omega|^2 + \alpha^2\tau^2 + \varepsilon|\omega|^2 t^2)^{-\nu}$, where γ, α, β are positive constants, $\nu > \frac{d+1}{2}$, $\varepsilon \in [0, 1]$.

- α^{-1} is spatial range, explaining spatial decay.
- β^{-1} is temporal range, explaining temporal decay.
- γ is a scaling constant, proportional to the sill σ^2.
- ν controls smoothness of the process.

$$f(\omega,\tau) = \gamma(\alpha^2\beta^2 + \beta^2|\omega|^2 + \alpha^2\tau^2 + \varepsilon|\omega|^2\tau^2)^{-\nu}$$

- When $\varepsilon = 1$, the covariance function is separable:

$$f(\omega,\tau) = \gamma(\alpha^2 + |\omega|^2)^{-\nu}(\beta^2 + \tau^2)^{-\nu}$$

- Notice that the spatial frequency component looks like the Matern spectral density.
- When $\varepsilon = 0$, there is no space-time interaction,

$$f(\omega,\tau) = \gamma(\alpha^2\beta^2 + \beta^2|\omega|^2 + \alpha^2\tau^2)^{-\nu}$$

- When $\varepsilon \in (0,1)$, there is space-time interaction, and the covariance function is non-separable.

Taylor's Hypothesis

- A Covariance function satisfies Taylor's hypothesis if there exists a velocity vector $\mathbf{v} \in \mathbb{R}^d$ such that $C(\mathbf{0}, t) = C(\mathbf{v}t, 0)$ for all $t \in \mathbb{R}$.

Stationary Nonseparable ST Covariance Function

Cressie, Huang (1999) derive, for integrable and continuous C:

$$f(\omega, \tau) = (2\pi)^{d-1} \int \int \exp(-i\mathbf{x}^T\omega - i\tau t))C(\mathbf{x}, t)d\mathbf{x}dt$$

$$= (2\pi)^{-1} \int h(\omega, t)dt$$

$$h(\omega, t) = (2\pi)^{-d} \int \exp(-i\mathbf{x}^T\omega)C(\mathbf{x}, t)d\mathbf{x} \qquad = \int \exp(-i$$

Then, specify $h(\omega, t)$ to get $C(\mathbf{x}, t)$. Assume $h(\omega, t) = \rho(\omega, t)k(\omega)$, where ρ is an autocorrelation function in t, and $k(\omega)$ is a spectral density in $\mathbf{x}$. Then,

$$C(\mathbf{x}, t) = \int \exp(i\mathbf{x}^T\omega)\rho(\omega, t)k(\omega)d\omega.$$

- Separability is obtained when $\rho(\omega, t) = \rho(t)$, i.e spatial autocorrelation depends only on time lag.
- Only need to specify $\rho(\omega, t)$ and $k(\omega)$ to be integrable and allow $C(\mathbf{x}, t)$ to be computed as an integral.

$$\rho(\omega, t) = \exp\{-||\omega||^2 t^2/4\} \exp\{-\delta t^2\} k(\omega) \quad = \exp\{-\alpha||\omega||^2\}$$

Then,

$$C(\mathbf{x}, t) \propto \frac{1}{(t^2 + c_0)^{d/2}} \exp\left\{-\frac{\mathbf{x}^2}{t^2 + c_0}\right\} \exp\{-\delta t^2\}$$

References

- Fuentes, Chen, Davis (2007). "A class of nonseparable and nonstationary spatial temporal covariance functions".
- Cressie, Huang (1999). "Classes of nonseparable, spatio-temporal stationary covariance functions".