1 Isotropic Covariance Functions

Let \{Z(s)\} be a Gaussian process on \(\mathbb{R}^n\), i.e., a collection of jointly normal random variables \(Z(s)\) associated with \(n\)-dimensional locations \(s \in \mathbb{R}^n\). The joint distribution of \(\{Z(s)\}\) depends only on the means \(\mu(s) = \mathbb{E}Z(s)\) and the covariances \(C(s, t) = \mathbb{E}(Z(s) - \mu(s))(Z(t) - \mu(t))\).

The process is called stationary or translation invariant if the distribution wouldn’t change under a rigid translation of the entire collection of locations, i.e., if \(\mu(s) = \mu(s + h)\) and \(C(s + h, t + h) = C(s, t)\) for all \(h\); in this case \(\mu(s) \equiv \mu\) is constant and \(C(s, t) = C(s - t, 0)\) can only depend on the difference \(h = (s - t)\) between the two locations, so must be of the form \(C(s, t) = C_0(s - t)\) for some function \(C_0(h) = C(h, 0)\) on \(\mathbb{R}^n\). Not just any function \(C_0(h)\) can be a covariance function; let’s see what the choices are.

It’s easy to see that the function \(C_0\) must be even, i.e., must satisfy \(C_0(h) = C_0(-h)\), since \(C(s - t) = \mathbb{E}(Z(s) - \mu(s))(Z(t) - \mu(t)) = C(t - s)\). But more is true: if \(\{s_j\}\) any collection of locations, then complex linear combinations \(a^T(Z - \mu) = \sum a_j(Z_j - \mu_j)\) of the centered random variables \(Z_j = Z(s_j)\) (with means \(\mu_j = \mu(s_j)\)) must have nonnegative squared module \(E|\sum a_j(Z_j - \mu_j)|^2 = \sum a_j^2C(s_j - s_k)\bar{a}_k \geq 0\) for every set of complex numbers \(\{a_j\} \subset \mathbb{C}\). A function \(C_0(h)\) is called positive semi-definite if it always satisfies the inequality \(\sum a_jC(s_j - s_k)\bar{a}_k \geq 0\) for any locations \(s_j\) and complex numbers \(a_j\); this is equivalent to asking that \(C(h) = C(-h)\) for every \(h \in \mathbb{R}^n\) and that \(\sum a_jC(s_j - s_k)a_k \geq 0\) for all real numbers \(a_j \in \mathbb{R}\). One way to get a symmetric positive semi-definite function \(C_0(h)\) is by taking the Fourier transform

\[
C_0(h) = \int_{\mathbb{R}^n} e^{ih \cdot \omega} G(\omega) d^n \omega = \int_{\mathbb{R}^n} e^{ih \cdot \omega} G(d\omega) \tag{1}
\]

of any positive function \(G \in L_1(\mathbb{R}^n)\) or, more generally, of any finite positive Borel measure \(G(d\omega)\), because then

\[
\sum a_jC(s_j - s_k)\bar{a}_k = \int_{\mathbb{R}^n} \sum a_j e^{i_j \cdot \omega}(\bar{a}_k e^{i_k \cdot \omega})G(d\omega)
\]

\[
= \int_{\mathbb{R}^n} \left|\sum a_j e^{i \cdot \omega}\right|^2 G(d\omega) \geq 0.
\]

It turns out that this is the only way to get one—that every positive semi-definite function can be written in this form for some finite positive measure.
$G(d\omega)$, called the \textit{spectral measure} (if $G(d\omega) = G(\omega) \, d\omega$ is absolutely continuous, $G(\omega)$ is called the \textit{spectral density}). Known as 'Bochner’s Theorem,’ this result is really just the Fourier inversion formula in an unfamiliar setting:

$$G(\omega) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-i\omega \cdot h} C_0(h) \, d^n h.$$  

Since the process $\{Z(s)\}$ is real-valued, the spectral density $G(\omega) = G(-\omega)$ must be an even function and so we can write

$$C_0(h) = \int_{\mathbb{R}^n} \cos(h \cdot \omega) G(\omega) \, d^n \omega$$

$$G(\omega) = (2\pi)^{-n} \int_{\mathbb{R}^n} \cos(h \cdot \omega) C_0(h) \, d^n \omega$$

If the Gaussian process is also \textit{isotropic}, or invariant under rotations, then $G(\omega) = g(|\omega|)$ must also be invariant under rotations and depend only on the length $r = |\omega|$ of the vector $\omega \in \mathbb{R}^n$. In this case we can simplify these integrals by transforming to polar coordinates.

\subsection*{1.1 Polar Coordinates for Probabilists}

Polar coordinates are a familiar tool in two-dimensional integrals, where the change of variables from $x \in \mathbb{R}^2$ to $r = \sqrt{x_1^2 + x_2^2}$ and $\theta = \arctan x_2/x_1$ (so $x_1 = r \cos \theta$, $x_2 = r \sin \theta$) and a change from $d^2x$ to $r \, dr \, d\theta$ lead to simple expressions for the integrals of radial functions. Equivalently, we can let $\sigma$ have a uniform probability distribution (denoted by $d\sigma$) over the unit circle $S^1 = \{x : x_1^2 + x_2^2 = 1\}$, and change variables from $x = (x_1, x_2)$ to $(r, \sigma)$, with $d^2x = dx_1 \, dx_2$ replaced by $2\pi r \, dr \, d\sigma$.

In three dimensions the first approach has its analogue in the Euler angles, while the second is simpler with uniform measure for $\sigma$ on the unit sphere $S^2 \subset \mathbb{R}^3$, with $d^3x = dx_1 \, dx_2 \, dx_3$ replaced by $4\pi r^2 \, dr \, d\sigma$. Notice that $2\pi r$ and $4\pi r^2$ are the circumference of the circle and the area of the sphere of radius $r$, respectively. In any number $n$ of dimensions the sphere $S^{n-1}$ has area $2\pi^{n/2} r^{n-1}/\Gamma(n/2)$, and we can again evaluate integrals in polar coordinates with the uniform probability distribution $d\sigma$ for $\sigma \in S^{n-1} \subset \mathbb{R}^n$, and $d^n x = 2\pi^{n/2} r^{n-1} dr \, d\sigma$. This makes it easy to compute integrals of radial functions; for functions that also depend on one or more of the components $x_j$, it is sometimes helpful to note that the squares $\{\sigma_j^2\}$ have a Dirichlet $\text{Di}(\frac{1}{2}, \ldots, \frac{1}{2})$ joint distribution, so each $\sigma_j$ is distributed as the square root of a $\text{Be}(\frac{1}{2}, \frac{n-1}{2})$ random variable.
1.2 Evaluating $C_0(h)$

Switching to polar coordinates $r = |\omega| \geq 0$ and $\sigma = \omega/|\omega| \in S^{n-1}$ (where $d\sigma$ denotes the uniform probability measure on the unit sphere $S^{n-1}$ in $\mathbb{R}^n$), and noting that the component $\sigma_h = \sigma \cdot h/|h|$ of $\sigma \in S^{n-1}$ in the direction $h$ again has the same distribution as the square root of a $\text{Be}(\frac{1}{2}, \frac{n-1}{2})$ random variable, writing $\rho$ for $|h|$, 

$$C_0(h) = \int_{\mathbb{R}^n} \cos(h \cdot \omega) g(|\omega|) d^n \omega$$

$$= \int_{\mathbb{R}_+ \times S^{n-1}} \cos(r \rho \sigma_h) g(r) \frac{2\pi^{n/2} r^{n-1}}{\Gamma(n/2)} dr \, d\sigma$$

$$= \int_{\mathbb{R}_+} \int_0^1 \cos(r \rho \sqrt{u}) g(r) \frac{2\pi^{n/2} r^{n-1}}{\Gamma(n/2)} \frac{\Gamma(n/2)}{\Gamma(\frac{n}{2})\Gamma(\frac{n-1}{2})} u^{1/2} (1-u)^{(n-1)/2} dr \, du$$

$$= \int_0^\infty \rho (2\pi r/\rho)^{\nu+1} J_\nu(r \rho) g(r) dr, \quad \nu = \frac{n}{2} - 1$$

$$= \int_0^\infty (r \rho/2)^{-\nu} \Gamma(\nu + 1) J_\nu (r \rho) \gamma(dr)$$

where

$$J_\nu(z) = \frac{(z/2)^\nu}{\sqrt{\pi} \Gamma(\nu + 1/2)} \int_0^\pi \cos(z \cos \theta) \sin(\theta)^{2\nu} d\theta$$

is the Bessel function of the first kind of order $\nu$ (see Watson, 1944). Bessel functions aren’t as familiar as sines and cosines, but they’re common in engineering and physics and are in the standard C library, the GNU Scientific library (GSL), R, Maple and Mathematica, MatLab, Python’s SciPy library, etc.; for more details, see Abramowitz and Stegun (1964, Chapter 9). Here’s a plot of $J_0(z)$:
The plot of \( J_0(z) \) looks a little like a sine or cosine, but falls off like \( 1/\sqrt{z} \) as \( z \to \infty \).

The most general isotropic covariance is given in \( \Box \), with the absolutely continuous measure \( g(r) \frac{2\pi^{n/2}}{\Gamma(n/2)} r^{n-1} dr \) replaced by an arbitrary positive finite measure \( \gamma(dr) \) on \([0, \infty)\). Any isotropic covariance function may be approximated by one with a discrete spectral measure \( \gamma(dr) = \sum \gamma_j \delta_{r_j}(dr) \) assigning mass \( \gamma_j \) to finitely many points \( r_j \):

\[
C_0(h) = C(\rho) \approx \sum_j \frac{(2/\tau_j \rho)}{T(\nu + 1)J_\nu(\tau_j \rho)} \gamma_j
\]

\[
= \begin{cases}
\sum_j \gamma_j \cos(\tau_j \rho) & \text{if } n = 1 \\
\sum_j \gamma_j \tau_j J_0(\tau_j \rho) & \text{if } n = 2 \\
\sum_j \gamma_j \sqrt{\pi/2} r_j \rho J_1(\tau_j \rho) & \text{if } n = 3
\end{cases}
\]

but a more common approach is to choose small parametric families of densities \( g^\mu(r) \) or measures \( \gamma^\mu(dr) \).

We can recover a spectral density \( g(r) = G(\omega) \) (for \( r = |\omega| \)) through the Fourier inversion formula, using polar coordinates with \( \rho = |h| \in \mathbb{R}_+ \) and \( \sigma = h/|h| \in S^{n-1} \):

\[
g(r) = G(\omega) = \frac{1}{(2\pi)^n} \int \cos(-h \cdot \omega) C_0(h) \, d^n h \\
= \frac{1}{(2\pi)^n} \int_{\mathbb{R}_+ \times S^{n-1}} \cos(-\rho \sigma \omega) C(\rho) \frac{2\pi^{n/2}}{\Gamma(n/2)} \rho^{n-1} \, d\rho \, d\sigma \\
= \int_0^\infty r(\rho/2\pi \tau)^{n/2} J_\nu(\rho \tau) C(\rho) \, d\rho, \quad \nu \equiv \frac{n}{2} - 1
\]

\[
= \begin{cases}
\int_0^\infty \frac{2}{\pi} \cos(\rho \tau) C(\rho) \, d\rho & \text{if } n = 1 \\
\int_0^\infty (\rho/2\pi \tau)^n J_0(\rho \tau) C(\rho) \, d\rho & \text{if } n = 2 \\
\int_0^\infty (\rho/2\pi \tau)^{3/2} J_1(\rho \tau) C(\rho) \, d\rho & \text{if } n = 3
\end{cases}
\]

It is hard to imagine what \( C_0(h) \) would look like for different choices of \( g(r) \); a simple approach is to take whatever symmetric functions \( G(\omega) \) whose Fourier transforms we can find, and see what we get. Here are some commonly used covariance families, in \( n = 2 \) dimensions; in each case \( \theta_1 = C_0(0) \) is an overall level parameter and \( \theta_2 \) is a distance scale parameter:

- **Power Exponential family**
  
  \[
  C(\rho \mid \theta_1, \theta_2) = \theta_1 \exp\{-|\rho/\theta_2|^p\}, \quad 0 < p \leq 2
  \]
Two notable covariograms in this family are the exponential ($p = 1$, solid below) and the Gaussian ($p = 2$, dashed below):

Notice that the exponential has a negative derivative at $z = 0$, so it falls off quickly at first, then slowly levels off, while the Gaussian has zero derivative near $z = 0$ then falls off very quickly. From (31) it follows that the exponential has spectral density function 

$$g(r) = \frac{\theta_1 \theta_2^2 / 2\pi}{(1 + r^2 \theta_2^2)^{3/2}},$$

proportional to a bivariate Cauchy density function, while the Gaussian has spectral density 

$$g(r) = \frac{\theta_1 \theta_2^2}{4\pi} \exp \left(-\frac{r^2 \theta_2^2}{4}\right),$$

proportional to a normal density.

- Matérn

$$C(\rho \mid \theta) = \frac{2\theta_1}{\Gamma(\theta_3)} \left(\frac{\rho}{2\theta_2}\right)^{\theta_3} K_{\theta_3}(\rho/\theta_2),$$

where $K_{\nu}(z)$ is the modified Bessel function of the third kind of order $\nu$ [Abramowitz and Stegun 1964, §9.6.2):

The displayed plot has shape parameter $\theta_3 = 2$. The Matérn class is quite flexible and includes the exponential family (with $\theta_3 = \frac{1}{2}$), the Gaussian family (in the limit as $\theta_3 \to \infty$), and many others. In $n$ dimensions its spectral density function is

$$g(r) = \frac{\theta_1 \theta_2^n}{\Gamma(\theta_3)n^{3/2}} (1 + \theta_2^2 r^2)^{-\theta_3 - n/2}.$$
proportional to the familiar \( n \)-variate Student’s \( t \) density function with 
\( 2\theta_3 \) degrees of freedom and variance scale \( \sigma^2 = 1/2\theta_2^2\theta_3 \). This lends
more insight into how the Matérn reduces to the exponential when
\( \theta_3 = 1/2 \) and to the Gaussian when \( \theta_3 \to \infty \).

- Spherical

\[
C(\rho \mid \theta) = \begin{cases} 
\theta_1 \left[ 1 - \frac{2}{\pi} \left( \frac{\rho}{\theta_2} \sqrt{1 - \left( \frac{\rho}{\theta_2} \right)^2} + \sin^{-1} \left( \frac{\rho}{\theta_2} \right) \right) \right] & \text{for } \rho < \theta_2 \\
0 & \text{for } \rho \geq \theta_2
\end{cases}
\]

The spherical covariance function is proportional to the area of in-
tersection for two discs of diameter \( \theta_2 \) with centers separated by dis-
tance \( \rho \). In this model the Gaussian quantities \( Z_j \) and \( Z_k \) at loci \( s_j \) 
and \( s_k \) separated by a distance greater than \( \theta_2 \) will be independent.

\[
\text{This is not quite linear. Like the exponential, it has a negative slope at}
\text{ } z = 0 \text{ and falls off rapidly at first; like the Gaussian, it falls off rapidly}
\text{ later and in fact reaches zero. The spectral density, while available}
\text{ in closed form, isn’t illuminating; it’s best to think of the spherical}
\text{ process as a convolution or moving average of Gaussian white noise,}
\text{ integrated at each locus over the surrounding ball of diameter } \theta_2.
\]

The same idea works in any dimension \( d \) — a constant \( \theta_1 \), times the
volume of intersection of two balls of radius \( \theta_2 \) whose centers are sep-
arated by a distance \( \rho \), is a valid covariance function in dimensions
\( n \leq d \) (for any real \( d \geq 1 \), not just \( d \in \mathbb{N} \)):

\[
C(\rho \mid \theta) = 2\theta_1 I_{(1-\rho/\theta_2)/2} \left( \frac{d+1}{2}, \frac{d+1}{2} \right), \quad 0 \leq \rho \leq \theta_2
\]

where \( I_z(\alpha, \beta) \) is the normalized incomplete Beta function (Abramowitz
and Stegun, 1964, §6.6.2), the CDF at \( z \) of the Be(\( \alpha, \beta \)) distribution.
For \( d = 2 \) this reduces to the equation above, and for \( d = 3 \)
to \( C(\rho \mid \theta) = \theta_1 (\theta_2 - \rho)^2 (2\theta_2 + \rho)/2\theta_2^2 \) for \( 0 \leq \rho \leq \theta_2 \).
2 Constructing GPs from their Spectral Measures

The integral representations of isotropic covariance functions can be exploited to construct random fields with specified covariance structure, in at least two different ways.

2.1 White-Noise Convolutions

Many processes may be constructed similarly as kernel integrals of standard Gaussian white noise,

\[ Z(h) = \int_{\mathbb{R}^n} k(h - s) \zeta(ds); \]

where “standard” means that \( E[\zeta(ds)] = 0 \) and \( E[\zeta(ds)^2] = ds \) (more formally, that \( \zeta \) is countably additive with \( \zeta(A) \sim \text{No}(0, |A|) \) for Borel \( A \subset \mathbb{R}^n \) of finite Lebesgue measure \(|A|\)). The covariance is

\[ C_0(h) = E[Z(0)Z(h)] = \int_{\mathbb{R}^n} k(h - s) \overline{k(-s)} ds \]

with spectral density

\[ G(\omega) = (2\pi)^{-n} \int e^{-i\omega \cdot h} C_0(h) dh \]

\[ = (2\pi)^{-n} \int \int e^{-i\omega \cdot h} k(h - s) \overline{k(-s)} ds dh \]

\[ = (2\pi)^{-n} \left| \int e^{-i\omega \cdot x} k(x) dx \right|^2. \]

Thus an isotropic kernel may be computed from the spectral density as

\[ k(x) = (2\pi)^{-n/2} \int e^{i\omega \cdot x} G(\omega)^{1/2} d^n \omega \]

or, in polar coordinates,

\[ k(\rho) = \int_{0}^{\infty} r^{\nu+1} \rho^{-\nu} J_\nu(r \rho) g(r)^{1/2} dr \]

\[ = \begin{cases} \int_{0}^{\infty} \sqrt{\frac{2}{\pi}} \cos(r \rho) \sqrt{g(r)} dr & \text{if } n = 1 \\ \int_{0}^{\infty} J_0(r \rho) r \sqrt{g(r)} dr & \text{if } n = 2 \\ \int_{0}^{\infty} J_{1/2}(r \rho) r^{3/2} \rho^{-1/2} \sqrt{g(r)} d \rho & \text{if } n = 3 \end{cases} \]
provided that the square root of the spectral density is the Fourier transform of a finite positive function, i.e., is itself positive semidefinite. For the Matérn class, the root spectral density \( \sqrt{g(r)} \propto (1 + \theta_3^2 r^2)^{-\frac{\theta_3 n}{2}} \) will be another \( n \)-variate \( t \) density provided \( \theta_3 > n/2 \) and in this case, setting \( \epsilon = (2\theta_3 - n)/4 > 0 \), we find

\[
k(\rho) = \frac{2\theta_1^{1/2}(2\theta_2)^{-\epsilon-n/2}}{\Gamma(\epsilon + n/2)\sqrt{\Gamma(2\epsilon + n/2)}} K_{\epsilon}(\rho/\theta_2)
\]

leads to a moving-average kernel representation for the Matérn covariance class. In any number \( n \geq 1 \) of dimensions the restriction \( \epsilon > 0 \) entails \( \theta_3 > n/2 \geq 1/2 \), ruling out the exponential covariance, but the Gaussian covariance (the limiting case as \( \theta_3 \to \infty \)) is available in any number of dimensions, with

\[
k(\rho) = \theta_1^{1/2}(\pi \theta_2^2/4)^{-n/2} e^{-\rho^2/\theta_2^2}.
\]

### 2.2 Spectral Representation

We generalize Hida and Hitsuda (1993, §III.2, pp. 42 ff). Let \( \Theta \subseteq \mathbb{R}^n \) be the spectrum, the support of the spectral measure \( G(d\omega) \) (see (I)), and let \( \zeta(d\omega) \) be a Gaussian random measure on the Borel sets of \( \Omega \) with control measure \( G(d\omega) \) (so \( \zeta(A) \sim \text{N}(0, G(A)) \) for Borel sets \( A \subset \Omega \) of finite \( G \)-measure). For \( t \in \mathbb{R} \) set

\[
Z_t := \int_{\Omega} e^{i\omega \cdot \xi} \zeta(d\omega).
\]

Then \( Z_t \) is a complex-valued Gaussian stochastic process with mean zero and autocovariance

\[
E Z_s \bar{Z}_t = \int_{\Omega} e^{i\omega \cdot (s-t)} G(d\omega) = C_0(s - t)
\]

I think we can arrange for \( Z \) to be real-valued, but probably not \( \zeta \) since if \( \zeta \) were real then we would have \( Z_{-t} \equiv \bar{Z}_t \).

For the special case of \( n = 1 \), we can start with independent real-valued Gaussian random measures \( \zeta_0(d\omega) \) and \( \zeta_1(d\omega) \) on \( \Omega_+ = \Omega \cap \mathbb{R}_+ \), each with control measure \( \frac{1}{2} G(d\omega) \) restricted to \( \Omega_+ \), and set

\[
\zeta(A) = [\zeta_0(A_+) + \zeta_0(A_-)] + i \left[ \zeta_1(A_+) - \zeta_1(A_-) \right]
\]

8
(where of course $A_+ = A \cap \mathbb{R}_+$ and $A_- = A \cap \mathbb{R}_-$). Then $\zeta$ is a complex Gaussian measure on $\mathbb{R}$ with mean zero and covariance

$$E\zeta(ds)\bar{\zeta}(dt) = \delta_s(dt)G(ds)$$

Set

$$Z_s = \int_{\Omega} e^{i\omega s} \zeta(d\omega) = \int_{\Omega} \cos(\omega s) \zeta_0(d\omega) - \int_{\Omega} \sin(\omega s) \zeta_1(d\omega)$$ \hspace{1cm} (7)

I’m not sure how to think of this in Polar coordinates, or how to extend it to $\geq 2$ dimensions, but I expect that’s possible useful interesting and, probably, old. The Free Euclidean Field should be an example.

3 Compact Ranges

Fourier transforms and harmonic analysis are most commonly studied on Euclidean space $\mathbb{R}^d$, but many of the ideas extend to bounded spaces like the closed unit ball $B^d \subset \mathbb{R}^d$ or its boundary the unit sphere $S^{d-1}$ (the disc and circle, for $d = 2$), or even to locally compact groups.

In this section we’ll consider functions on the unit ball $B^d$. Because it is compact, functions in $L_2(B^d)$ will have Fourier series with a discrete index set, rather than the Fourier transforms of Section I.

3.1 Fourier-Bessel Expansions in $d = 2$ Dimensions

The key for $d = 2$ is the orthogonality relationship for Bessel functions of the first kind ([Abramowitz and Stegun, 1964, §9.1]):

$$\int_0^1 J_\alpha(r,j_\alpha,n) J_\alpha(r,j_\alpha,n') r\,dr = \frac{1}{2} \delta_{nn'} [J_{\alpha + 1}(j_\alpha,n)]^2,$$

where $j_\alpha,n$ is the $n^{th}$ positive zero of $J_\alpha(x)$, for any $n, n' \in \mathbb{Z}$ and any $\alpha \in \mathbb{R}$. For integers $\alpha$, $J_\alpha(-x) = (-1)^\alpha J_\alpha(x)$, so $[J_{\alpha + 1}(j_\alpha,n)]^2 = [J_{\alpha + 1}(j_\alpha,n)]^2$ and $j_\alpha,n = j_{-\alpha},n$. It follows that for $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ the functions

$$\phi_{mn}(r, \theta) \equiv \frac{1}{\sqrt{\pi} J_{|m|+1}(j_{m,n})} J_m(r,j_{m,n}) e^{im\theta}$$ \hspace{1cm} (8)

are orthonormal in $L_2(B^2)$. In fact they form a CONS, so any function $f \in L_2(B^2)$ has a convergent expansion

$$f(r, \theta) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{N}} a_{mn} \phi_{mn}(r, \theta) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{N}} \frac{a_{mn}}{c_{mn}} J_m(r,j_{m,n}) e^{im\theta}$$
\[
    c_{mn} = \sqrt{\pi} J_{\nu+1}(j_{m,n}), \quad a_{mn} = \int_{B^d} f(r, \theta) \phi_{mn}(r, \theta) r \, dr \, d\theta. \tag{9}
\]

Note \(a_{mn}\) from (2) can also be evaluated in Cartesian coordinates if, for example, \(f\) is known on a rectangular grid containing \(B^2\) — just multiply the integrand by \(1_{B^2}\) and replace \(r \, dr \, d\theta\) with \(dx \, dy\). Both the Bessel functions \(J_m\) and the zeros \(j_{m,n}\) are available in \(\mathbb{R}\). From Bessel’s differential equation

\[
    z^2 \frac{d^2}{dz^2} J_m(z) + \frac{d}{dz} J_m(z) + (z^2 - m^2) J_m(z) = 0
\]

(Abramowitz and Stegun 1964, §9.1.1), \(\phi_{mn}\) is an eigenfunction of the Laplacian with eigenvalue \(-j_{m,n}^2\) satisfying Dirichlet boundary conditions \(\phi_{mn}(1, \theta) \equiv 0\) on \(S^1 = \partial B^2\), so the negative Laplacian of \(f\) is given in polar coordinates by

\[
    -\Delta f(r, \theta) \equiv - \left\{ \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} \right\} = \sum (j_{m,n})^2 \left( \frac{a_{mn}}{c_{mn}} \right) J_m(r_j m, n) e^{im\theta}.
\]

The squared \(L_2\) and Sobolev \(H_1\) norms of \(f\) are

\[
    \langle f, f \rangle = \sum |a_{mn}|^2, \quad \|f\|_1^2 = \langle (-\Delta + I)f, f \rangle = \sum (1 + j_{m,n}^2) |a_{mn}|^2,
\]

so one reasonable scale-invariant measure of roughness would be

\[
    \frac{\langle \nabla f, \nabla f \rangle}{\langle f, f \rangle} = \frac{\langle (-\Delta + I)f, f \rangle}{\langle f, f \rangle} = \frac{\sum j_{m,n}^2 |a_{mn}|^2}{\sum |a_{mn}|^2}
\]

or, similarly but more generally, the ratio of squared Sobolev norms

\[
    \frac{\|f\|_s^2}{\|f\|_0^2} = \frac{\langle (-\Delta + I)^s f, f \rangle}{\langle f, f \rangle} = \frac{\sum (1 + j_{m,n}^2)^s |a_{mn}|^2}{\sum |a_{mn}|^2}
\]

for any \(s > 0\). Another measure depending only on angular roughness would be

\[
    \frac{\sum |m|^{2s} |a_{mn}|^2}{\sum |a_{mn}|^2}
\]

for any \(s > 0\). Similar results hold for \(d = 3\) (using the spherical Bessel function \(j_0\)) or for other boundary conditions. Each of these measures is the moment of some “roughness” aspect, for the discrete probability distribution with masses proportional to the squared Fourier coefficients \(|a_{mn}|^2\).
3.2 Fourier Expansions in $d = 1$ Dimension

This section has a brief review of one-dimensional Sobolev spaces on an interval, for context and intuition.

Let $\mathcal{T} = [0, T]$ be a closed interval in $\mathbb{R}^1$. The functions

$$\phi_n(x) = \sin(n\pi x / T) \sqrt{2 / T}$$

are a CONS in $L_2(\mathcal{T}, dx)$ of eigenfunctions of $-\Delta$ (with eigenvalues $n^2 \pi^2 / T^2$), with Dirichlet boundary conditions at $\partial \mathcal{T}$, so every function $f \in L_2(\mathcal{T}, dx)$ has a convergent expansion of the form

$$f(x) = \sum_{n \in \mathbb{N}} a_n \phi_n(x),$$

(10)

with coefficients

$$a_n := \int_{\mathcal{T}} f(x) \phi_n(x) dx.$$

For $s \in \mathbb{R}$ the closed linear span $H_s$ of $\{\phi_n\}$ in the norm

$$\|f\|_s := \left\{ \sum_{n \in \mathbb{N}} |a_n|^2 \left(1 + n^2 \pi^2 / T^2\right)^s \right\}^{1/2}$$

(11)

is a Banach space (complete separable normed metric space) whose dual space is $H_{-s}$. The self-dual Hilbert space $H_0$ is just $L_2(\mathcal{T}, dx)$.

Differentiating (10) gives

$$\|f\|_0^2 = \left\| \sum_{n \in \mathbb{N}} a_n \frac{n \pi}{T} \cos\left(\frac{n \pi}{T}\right) x \sqrt{2 / T} \right\|_0^2$$

$$= \sum_{n,m \in \mathbb{N}} a_n a_m \frac{nm \pi^2}{T^2} \frac{2}{T} \int_0^T \cos\left(\frac{n \pi x}{T}\right) \cos\left(\frac{m \pi x}{T}\right) dx$$

$$= \sum_{n \in \mathbb{N}} a_n^2 \frac{n^2 \pi^2}{T^2} = \sum_{n \in \mathbb{N}} |a_n|^2 n^2 \pi^2 / T^2$$

$$= \|f\|_1^2 - \|f\|_0^2,$$

so $f \in H_1$ if and only if $f \in H_0$ and $f' \in H_0$. Similarly one can show that $f \in H_s \Rightarrow f' \in H_{s-1}$ for any $s \in \mathbb{R}$. For $s > 0$, $H_s$ consists of those
functions with \( s \) derivatives in \( L_2 \), while \( H_{-s} \) consists of \( s^{th} \) derivatives of \( L_2 \) functions.

In particular, for \( t \in \mathcal{T} \) the Dirac delta function \( \delta_t \in H_{-1} \) but \( \delta_t \notin H_0 \).
For example, for \( T = \pi \) and \( t = \pi/2 \), \( f \equiv \delta_t \) has coefficients

\[
a_n = \int_0^\pi f(x) \sin(nx) \sqrt{2/\pi} \, dx = \begin{cases} \sqrt{2/\pi} & n \equiv 1 \pmod{4} \\ -\sqrt{2/\pi} & n \equiv 3 \pmod{4} \end{cases},
\]

otherwise zero, with eigenvalues \( n^2 \), so

\[
\|\delta_t\|_2^2 = \frac{2}{\pi} \sum_{k=1}^{\infty} \left(1 + (2k + 1)^2\right)^s.
\]

This is only finite for \( s < -\frac{1}{2} \), so \( \delta_t \in H_s \) only for \( s < -\frac{1}{2} \).

The boundary \( \partial \mathcal{T} \) and boundary conditions play a subtle rôle in \( H_s(\mathcal{T}) \).
For example, in the interior of \( \mathcal{T} = [0, T] \) for \( T > 0 \) the function

\[f(x) = x(T - x)\]
satisfies \( f''(x) \equiv -2 \) and hence

\[(-\Delta + I)f(x) = x(T - x) + 2\]

and so, at least for non-negative integers \( s \),

\[(-\Delta + I)^s f(x) = x(T - x) + 2s.\]

One might expect \( f \in H_s \) for all \( s > 0 \), with

\[
\|f\|_2^2 \equiv \langle (-\Delta + I)^s f, f \rangle = \int_0^T [x(T - x) + 2s][x(T - x)] \, dx = T^5/30 + sT^3/3,
\]
suggesting (wrongly) that \( f \in H_s \) for all \( s > 0 \). In fact the coefficients are

\[
a_n := \int_T f(x) \phi_n(x) \, dx = 4\sqrt{2T^{5/2}}/(n\pi)^3
\]
for odd $n$, and zero for even $n$, so

$$
\| f \|_s^2 \equiv \sum_{n=1}^{\infty} |a_n|^2 \left( n^2 \pi^2 / T^2 + 1 \right)^s
$$

$$
= \frac{32 T^5}{\pi^6} \sum_{k=1}^{\infty} \frac{(2k+1)^2 \pi^2 / T^2 + 1)^s}{(2k+1)^6}
$$

which is only finite for $s < 5/2$. The problem is that even the first derivative $f'$ fails to satisfy Dirichlet boundary conditions $f(\partial T) = 0$. Integrating by parts,

$$
\int_T f'(x) f'(x) \, dx = f'(x) f(x) \bigg|_0^T - \int_T f''(x) \, f(x) \, dx
$$

$$
= - \int_T f''(x) \, f(x) \, dx = \langle (-\Delta) f, f \rangle
$$

$$
\int_T f''(x) f''(x) \, dx = f''(x) f''(x) \bigg|_0^T - \int_T f'''(x) \, f'(x) \, dx
$$

$$
= 4T - f'''(x) f(x) \bigg|_0^T + \int_T f'''(x) \, f'(x) \, dx
$$

$$
= \langle (-\Delta)^2 f, f \rangle + 4T
$$

so for $s \geq 2$ a non-zero boundary term enters the integration by parts and the naive calculation fails. Intuitively, what happens is that $f(x)$ takes the constant value of zero outside $T$ and so derivatives beyond the first pick up something like a Dirac delta function at the boundary.

Something similar happens in $d = 2$ dimensions for functions $f$ that satisfy Dirichlet boundary conditions at $\partial B^2$ but whose derivatives do not, such as

$$
f(x, y) = (1 - x^2 - y^2) \vee 0
$$

for which $-\Delta f \equiv 4$ and hence $(-\Delta + I)^s f = f + 4s$ in the interior. Still, $f \in H_s$ only for $s < 5/2$ because by circular symmetry its Fourier-Bessel coefficients are $a_{mn} = 0$ for $m \neq 0$, while

$$
a_{0n} = \frac{2\pi}{\sqrt{\pi} J_1(j_{0,n})} \int_0^1 (r - r^3) J_0(r j_{0,n}) \, dr = \frac{8\sqrt{\pi}}{j_{0,n}^3},
$$

so

$$
\| f \|_s^2 = 64\pi \sum_{n=1}^{\infty} (1 + \lambda_{0n}^2)^2 j_{0,n}^{-6},
$$

which is only finite for $s < 5/2$ since $j_{0,n} > n\pi$. 

13
4 Stochastic Expansions

4.1 Brownian Bridge on One Dimensional Interval

As in Sec. 3.2 let

\[ \phi_n(x) = \sin(n\pi x / T) \sqrt{2/T} \]

be a CONS in \( L_2(T, dx) \) for \( T = [0, T] \subset \mathbb{R}^1 \), and consider the Brownian Bridge Gaussian stochastic process \( X_t \) with mean and covariance

\[ \mathbb{E}X_t = 0 \quad \text{Cov}(X_s, X_t) = \mathbb{E}X_s X_t = (s \wedge t)(1 - (s \vee t)/T), \]

identical to unit-rate Brownian motion conditioned to reach \( X_T = 0 \). The Fourier coefficients for a random path \( \{X_t\} \) are Gaussian random variables

\[ A_n := \int_T Z_t \phi_n(s) \, dt \]

with mean zero and covariance

\[
\mathbb{E}A_m A_n = \frac{2}{T^2} \int_{T^2} (s \wedge t)(T - (s \vee t)) \sin(m\pi s / T) \sin(n\pi t / T) \, ds \, dt \\
= \begin{cases} 
0 & m \neq n \\
\frac{T^2}{n^2\pi^2} & m = n
\end{cases}
\]

so the \( \{A_n\} \stackrel{\text{ind}}{\sim} \text{No} \left(0, \left(\frac{T}{n\pi}\right)^2\right) \) are independent and the expected squared \( H_s \) norm of the sample path \( s \sim Z_s \) is

\[
\mathbb{E}\|Z_s\|_2^2 = \sum_{n \in \mathbb{N}} \frac{T^2 (1 + n^2\pi^2 / T^2)^t}{n^2\pi^2} < \infty \text{ if and only if } s < \frac{T}{2},
\]

with \( \mathbb{E}\|Z_s\|_2^2 = T^2/6 \). Roughly speaking, Brownian motion paths have just under one-half of a derivative.

Beginning with iid \( \zeta_n \stackrel{iid}{\sim} \text{No}(0, 1) \) one can construct a Brownian Bridge process as the sum

\[ Z_t := \sum_{n=1}^{\infty} \frac{T}{n\pi} \zeta_n \phi_n(t) \]

or fractional derivatives of it by

\[ Z_t^\alpha := \sum_{n=1}^{\infty} \left[ \frac{T}{n\pi} \right]^{1-\alpha} \zeta_n \phi_n(t), \]
which will be centered Gaussian processes with about $\frac{1}{2} - \sigma$ derivatives—smooth paths, for $\sigma \ll 0$ or, for $\sigma > 0$, generalized stochastic processes that only make sense as convolutions $Z[\psi] = \int Z_t \psi(t) \, dt$ for $\psi \in H_s$ with $s > \sigma - \frac{1}{2}$.

### 4.2 Stochastic Expansions in Two Dimensions

As in Section 3.1, let

$$\phi_{mn}(r, \theta) = \frac{1}{\sqrt{\pi}} \frac{J_{|m|+1}(j_{m,n})}{J_m(r \, j_{m,n})} e^{im\theta}$$

be a CONS in $L_2(B^2)$, and let $\zeta_{mn} \overset{iid}{\sim} \mathcal{N}(0, 1)$ for $m \in \mathbb{Z}, n \in \mathbb{N}$. Fix $\sigma \in \mathbb{R}$ and set

$$Z^\sigma(r, \theta) := \sum [j_{m,n}]^\sigma \zeta_{mn} \phi_{mn}(r, \theta)$$

for $(r, \theta) \in B^2$. The expected squared $H_s$ norm of $Z^\sigma$ is

$$\mathbb{E}\langle Z^\sigma, Z^\sigma \rangle_s = \mathbb{E}\langle (-\Delta + I)^s Z^\sigma, Z^\sigma \rangle
= \sum (j_{m,n}^2 + 1)^s j_{m,n}$$

Since $j_{m,n} \sim \pi (n + |m|/2 - 1/4)$ (Abramowitz and Stegun, 1964, §9.5.12), this is finite by the integral test if and only if

$$\infty > \int_1^\infty (r^2 + 1)^s r^\sigma \, dr,$$

so $Z^\sigma \in H_s$ almost-surely for $\sigma < -2(s + 1)$—or $\sigma < -2$ for $s = 0$.

### 4.3 Reproducing Kernel Hilbert Space (RKHS)

Now let $X_s$ be an arbitrary mean-zero Gaussian process with (positive definite) covariance kernel

$$k(s, t) = \mathbb{E}X_sX_t$$

for $s, t \in \mathcal{T}$, and consider the Hilbert space $\mathcal{H}$ consisting of limits of linear combinations of the $\{X_s\}$ for $\{s_j\} \subset \mathcal{T}$. For tame enough functions $f, g : \mathcal{T} \to \mathbb{R}$ this space will include elements $X_f \equiv \int_\mathcal{T} f(s)Z_s \, ds$ and $X_g \equiv \int_\mathcal{T} g(t)Z_t \, dt$, each Gaussian with mean zero and covariance

$$\langle X_f, X_g \rangle \equiv \int_{\mathcal{T}^2} f(s)k(s, t)g(t) \, ds \, dt,$$
the inner-product in $L_2(T)$ of $f$ and $Kg$ for the operator $K : L_2(T) \to L_2(T)$ with kernel $k$,

$$Kg(s) \equiv \int_T k(s, t) g(t) \, dt.$$ 

If $k \in L_2(T^2)$ that operator is Hilbert-Schmidt, with a complete orthonormal set $\{\phi_n\}$ of eigenfunctions

$$K \phi_n(s) \equiv \int_T k(s, t) \phi_n(t) \, dt = \lambda_n \phi_n(s)$$

whose eigenvalues are square-summable, with

$$\int_{T^2} k(s, t)^2 \, ds \, dt = \sum \lambda_n^2 < \infty.$$ 

In this case for iid $\{\zeta_n\} \sim \text{No}(0, 1)$ the series

$$Z_s \equiv \sum \lambda_n \zeta_n \phi_n(s)$$

converges almost-surely in $L_2$ to a Gaussian process with mean zero and covariance $k$, the so-called Karhunen-Loève expansion. The example of Sec. 4.1 was an example of this. As in that example, often $k$ is the Greens function for a differential operator (here $-\partial^2 / \partial x^2$ with Dirichlet boundary conditions), and the eigenfunctions may be found as solutions to differential equations.

References


By: Robert L Wolpert Last revised: August 14, 2013

16