Mickey Warner

1. Prove the results about the smoothness of the members of the Matérn family.

We use the theorem that if

$$\frac{d^{2\nu}}{d\tau^{2\nu}}\rho(\tau)$$

exists and is finite at $\tau = 0$, then the random field having $\rho(\tau)$ as its correlation function is ν times differentiable at 0.

Without loss of generality, let $\phi = 1$. The Matérn correlation function is given by

$$\rho(\tau) = \frac{\tau^\nu}{2^{\nu-1}\Gamma(\nu)}K_\nu(\tau), \qquad \tau \geq 0, \nu > 0.$$

For small τ and $\nu > 0$, $K_\nu(\tau) \approx \Gamma(\nu)2^{\nu-1}\tau^{-\nu}$. Also, $\frac{d}{d\tau}\tau^\nu K_\nu(\tau) = -\tau^\nu K_{\nu-1}(\tau)$ and $K_\nu(\tau) = K_{-\nu}(\tau)$. We will be taking $\tau \to 0$, so we use the approximation for $K_\nu(\tau)$. This leads to the derivative,

$$\begin{aligned}
\frac{d}{d\tau}\rho(\tau) &= -\frac{\tau^\nu}{2^{\nu-1}\Gamma(\nu)}K_{\nu-1}(\tau) \\
&= \begin{cases} -\frac{\tau^\nu}{2^{\nu-1}\Gamma(\nu)}K_{\nu-1}(\tau), & \nu - 1 \geq 0 \\ -\frac{\tau^\nu}{2^{\nu-1}\Gamma(\nu)}K_{1-\nu}(\tau), & \nu - 1 < 0 \end{cases} \\
&\approx \begin{cases} -\frac{\tau^\nu}{2^{\nu-1}\Gamma(\nu)}\Gamma(\nu-1)2^{\nu-2}\tau^{-\nu+1}, & \nu - 1 \geq 0 \\ -\frac{\tau^\nu}{2^{\nu-1}\Gamma(\nu)}\Gamma(1-\nu)2^{-\nu}\tau^{\nu-1}, & \nu - 1 < 0 \end{cases} \\
&\approx \begin{cases} -\tau G_1(\nu), & \nu - 1 \geq 0 \\ -\tau^{2\nu-1}G_2(\nu), & \nu - 1 < 0 \end{cases}.
\end{aligned}$$

Therefore,

$$\rho'(0) \begin{cases} = 0, & \nu \geq 1 \\ \in (-\infty, 0), & 1/2 \leq \nu < 1 \\ = -\infty, & 0 < \nu < 1/2 \end{cases}.$$

The second derivative is given by

$$\rho''(\tau) = \begin{cases} \frac{-\tau^{\nu-1}K_{\nu-1}(\tau) + \tau^\nu K_{\nu-2}(\tau)}{2^{\nu-1}\Gamma(\nu)}, & \nu \geq 2 \\ \frac{-\tau^{\nu-1}K_{\nu-1}(\tau) + \tau^\nu K_{2-\nu}(\tau)}{2^{\nu-1}\Gamma(\nu)}, & 1 \leq \nu < 2 \\ \frac{-\tau^{\nu-1}K_{1-\nu}(\tau) + \tau^\nu K_{2-\nu}(\tau)}{2^{\nu-1}\Gamma(\nu)}, & 0 < \nu < 1 \end{cases},$$

and is evaluated at $\tau = 0$ to

$$\rho''(0) \begin{cases} = 0, & \nu \geq 2 \\ \in (-\infty, 0), & 1 \leq \nu < 2 \\ = -\infty, & 0 < \nu < 1 \end{cases} .$$

We have that the second derivative is finite when $\nu \geq 1$, leading to a random field that is one time mean square differentiable. To show this generalizes to $\nu \geq d$, we need to keep taking derivatives of $\rho(\tau)$. I suspect that on each even derivative, the orders of certain Bessel functions are negated when those orders are less than ν. When this happens, the approximation will lead to a term having τ raised to a negative exponent causing the derivative to evaluate to $-\infty$.

2. Use the spectral representation to show that the product of two valid correlation functions is a valid correlation function.

A valid correlation function is the characteristic function of some random variable,

$$\rho(\boldsymbol{\tau}) = E\left[e^{i\boldsymbol{\tau}^\top \mathbf{X}}\right].$$

Suppose we have two valid correlation functions $\rho_1(\tau)$ and $\rho_2(\tau)$ associated with independent random variables $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively. Then the product is written

$$\begin{aligned} \rho(\boldsymbol{\tau}) = \rho_1(\boldsymbol{\tau})\rho_2(\boldsymbol{\tau}) &= E\left[e^{i\boldsymbol{\tau}^\top \mathbf{X}_1}\right] E\left[e^{i\boldsymbol{\tau}^\top \mathbf{X}_2}\right] \\ &= E\left[e^{i\boldsymbol{\tau}^\top \mathbf{X}_1} e^{i\boldsymbol{\tau}^\top \mathbf{X}_2}\right] \\ &= E\left[e^{i\boldsymbol{\tau}^\top (\mathbf{X}_1 + \mathbf{X}_2)}\right], \end{aligned}$$

so ρ is the characteristic function of $\mathbf{X}_1 + \mathbf{X}_2$ and thus the product of two valid correlation functions is a valid correlation function. Note, our assumption of independence for $\mathbf{X}_1$ and $\mathbf{X}_2$ presents no issues. $\mathbf{X}_1$ and $\mathbf{X}_2$ may be dependent, but then we could simply define new independent random variables $\mathbf{Y}_1$ and $\mathbf{Y}_2$ with the same marginal distributions as $\mathbf{X}_1$ and $\mathbf{X}_2$, resulting in the same correlation functions in either case.

3. The spectral density of a correlation in the Matérn family has tails whose thickness depends on the smoothness parameter. Conjecture: the smoothness of the corresponding random field depends on the number of moments of the spectral density. What can you say about this conjecture?

For correlation function

$$\rho(\tau) \propto (a\tau)^\nu K_\nu(a\tau), \qquad \tau \geq 0, \nu > 0, a = 1/\phi > 0,$$

we have the corresponding spectral density

$$f(x) \propto \frac{1}{(1+(x/a)^2)^{\nu+n/2}},$$

where n is the dimension τ (and x). This density has a form comparable to the t-distribution. Using integration by parts, we calculate the kth moment as

$$\begin{aligned}
E(X^k) &\propto \int x^k (1+(x/a)^2)^{-(\nu+n/2)} dx \\
&= -\frac{a^2}{2\nu+n-2} \frac{x^{k-1}}{(1+(x/a)^2)^{(2\nu+n-2)/2}}\Bigg|_{-\infty}^{\infty} + \int \frac{(k-1)a^2}{2\nu+n-2} \frac{x^{k-2}}{(1+(x/a)^2)^{-(2\nu+n-2)/2}} dx.
\end{aligned}$$

The first term (and hence the second term also) will be finite when $2\nu + n - 2 \geq k - 1$, or $\nu \geq (k-n+1)/2$. In one dimension, $n = 1$, we see that when $\nu \geq k/2$ the kth moment exists. This may be related to the theorem used in the first problem, that we need to have $2d$-differentiable correlation function to have a d-differentiable random field. Here, we need the $2d$th moment to exist, i.e. $\nu \geq d$, to have smoothness.

4. Use the K-L representation to approximate the exponential correlation for range parameter equal to 1. Plot the approximation for several orders and compare to the actual correlation.

The Karhunen-Loeve representation for a random process is given by

$$X(s) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} \psi_j(s) Z_j$$

where Z_j are mean zero processes with $E(Z_j Z_k) = \delta_{jk}$ and ψ_j is a basis of orthogonal functions such that

$$\int \psi_j(s) \overline{\psi_k(s)} ds = \delta_{jk}$$

and

$$\int C(s,s') \psi_j(s) ds = \lambda_j \psi_j(s').$$

For an exponential correlation function with $\phi = 1$ on an interval $[-L, L]$, we have

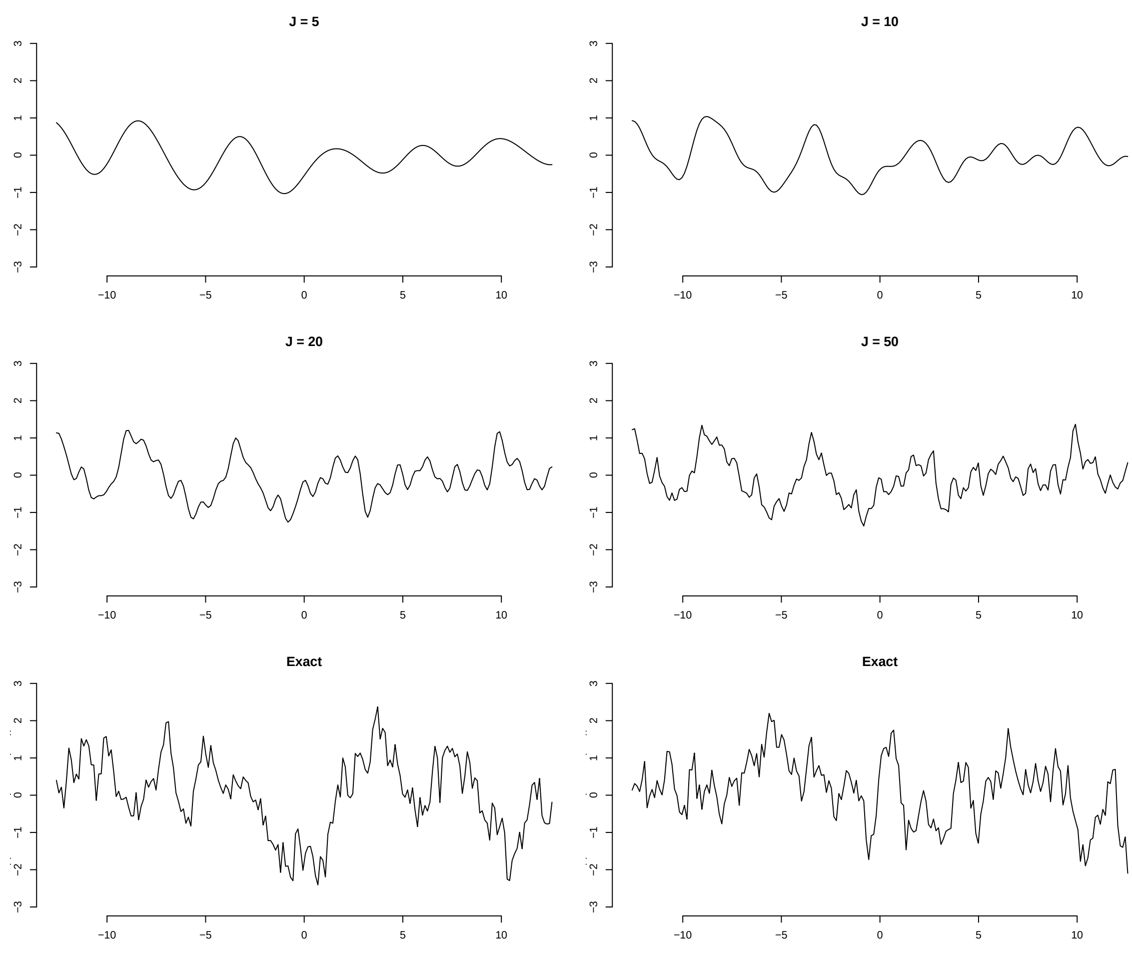
$$\lambda_i = \begin{cases} 2/(1+v_j^2), & i = 2j+1 \\ 2/(1+w_k^2), & i = 2k \end{cases}$$

$$\psi_i(s) = \begin{cases} \frac{\cos(v_j s)}{\sqrt{L+\sin(2v_j L)/(2v_j)}}, & i = 2j+1 \\ \frac{\sin(w_k s)}{\sqrt{L-\sin(2w_k L)/(2w_k)}}, & i = 2k \end{cases},$$

where v_j and w_k are the non-negative solutions to

$$1 - v\tan(vL) = 0$$
$$w + \tan(wL) = 0$$

sorted in increasing order. We end the approximation at $J = 5, 10, 20, 50$ summands. We let $Z_j \sim N(0,1)$, $j = 1, \ldots, J$ and $L = 2\pi$ and obtain realizations of $X(s)$ on a grid $s \in [-4\pi, 4\pi]$. We also obtain realizations with `geoR`'s `grf()` function which produces realizations from a Gaussian process. One draw for each J is shown (top four) and two Gaussian process draws are given (bottom two).



5. Repeat for the approximation given on Page 13 of the fifth set of slides.

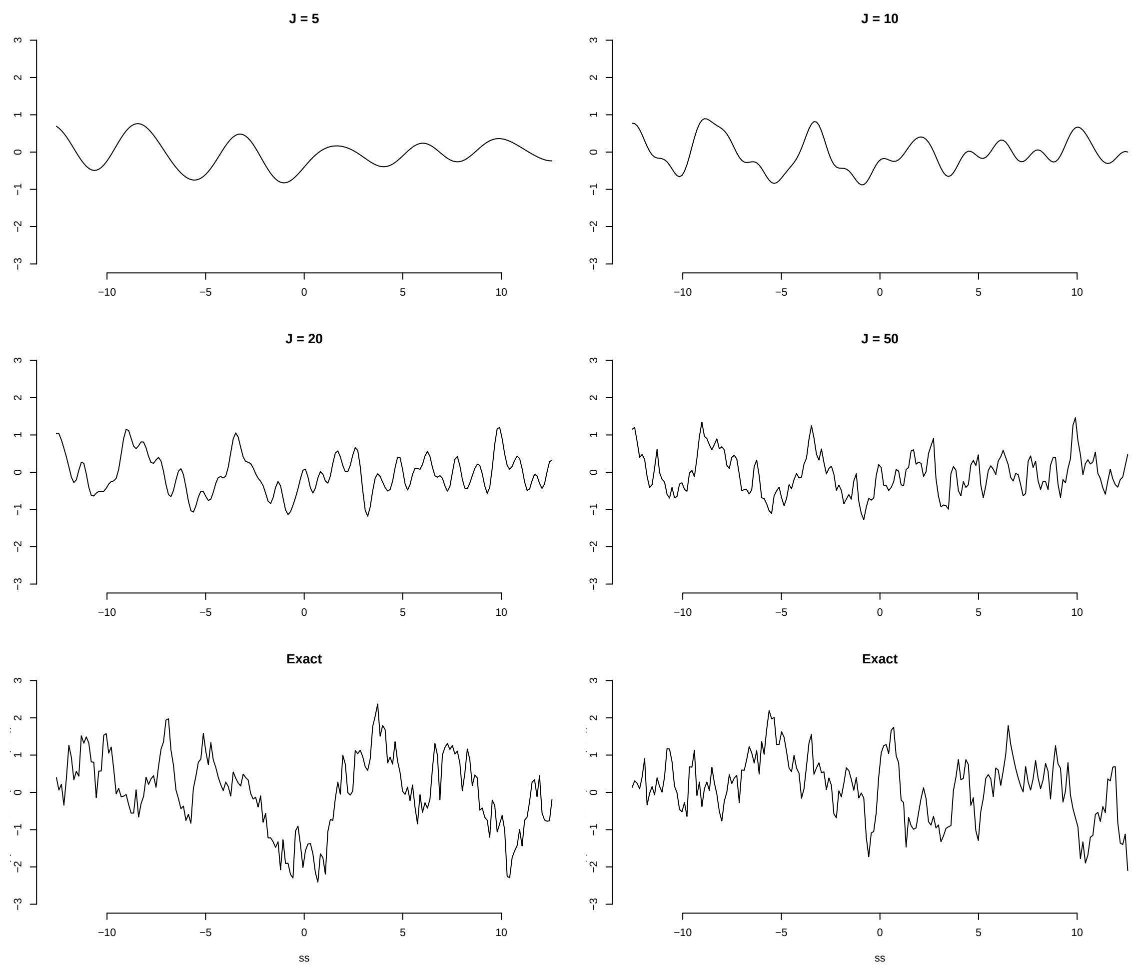
The approximation of interest is given by

$$\lambda_j \approx f(j\pi/(2L)), \qquad \psi_j(s) \approx ce^{ij\pi s/(2L)}$$

where $f(k)$ is the spectrum at k (and we set $c = 1$). Since the exponential correlation is equivalent to the Matérn with $\nu = 1/2$, the spectral density, for $\phi = 1$, is given by

$$f(k) = \frac{1}{(1+k^2)^{(n+1)/2}}$$

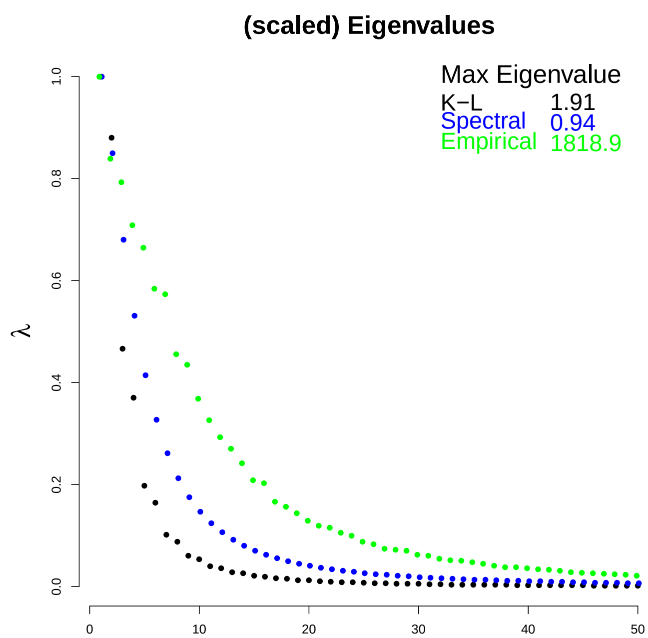
The next figure is comparable to the previous, except with this second approximation.



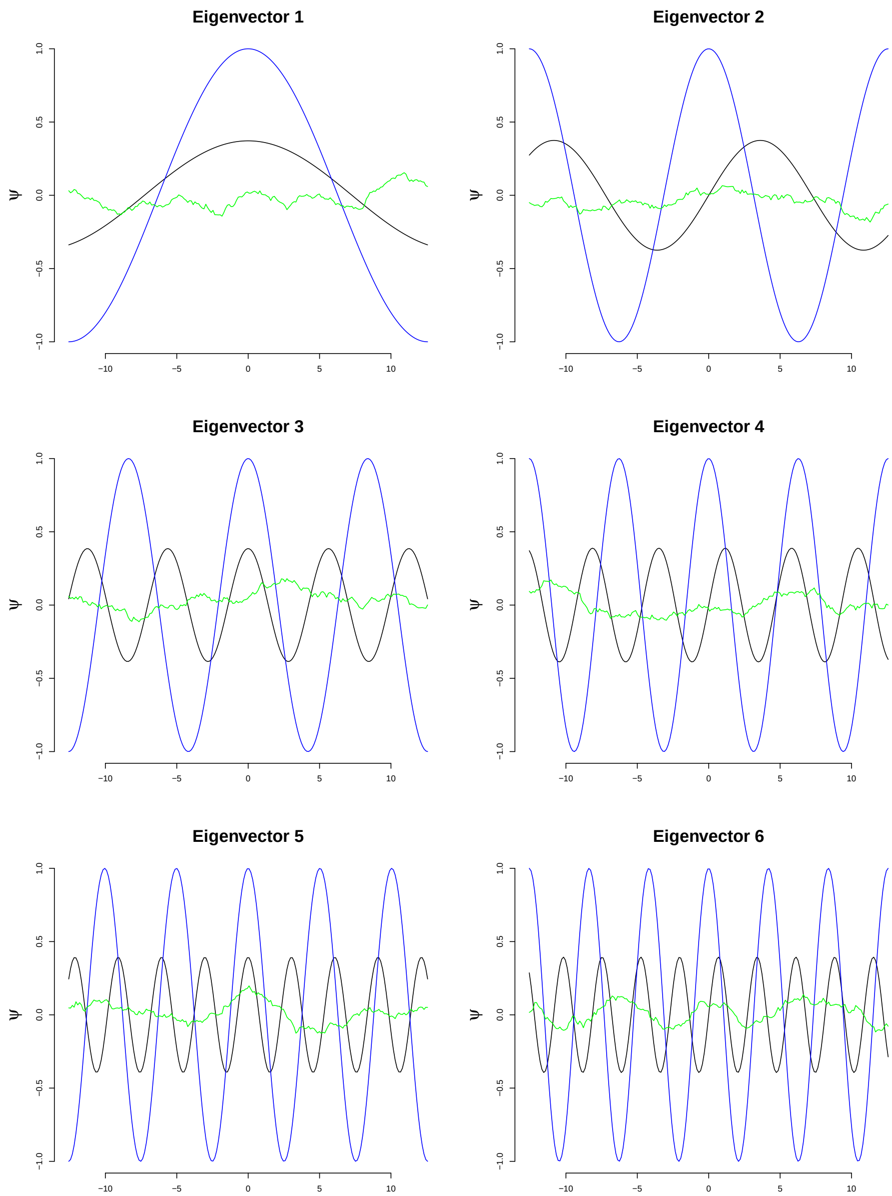
The same seed for generating the random draws was used in both problems 4 and 5. As we can see, the approximations are effectively identical. Both approaches reveal their inadequacies when J is too low, but at just $J = 50$, there is little distinct when obtaining realizations from the definition.

6. Generate 100 realizations of a univariate Gaussian process with exponential correlation with range parameter 1. Compare the empirically estimated eigenvalues and eigenfunctions to the ones given by the K-L and the approximation on Page 12.

Empirical estimates for the eigenvalues and eigenfunctions of 100 realizations were obtained via SVD. Here, we used a cutoff at $J = 50$ for the approximations. We see a substantial difference between all three eigenvalues and eigenfunctions. The eigenvalues are given in the first plot. These are scaled so the first (and largest) eigenvalue is 1 (for plotting purposes).



The last plot shows the eigenvectors associated with the first six eigenvalues. The are plotted against $s \in [-4\pi, 4\pi]$. The periodicity is much more evident in the approximations (perhaps due to construction). As in the first plot, black is K-L representation, blue is the approximation using the spectral density, and green is the empirical estimates from SVD.


Eigenvector 1
Eigenvector 2
Eigenvector 3
Eigenvector 4
Eigenvector 5
Eigenvector 6
ψ
1.0
0.5
0.0
−0.5
−1.0
−10
−5
0
5
10