Geostatistical models – Part III

Applied Spatial Statistics

Spatial correlation models

- We have discussed general ideas including stationarity and isotropy
- Assuming the process has these properties dramatically simplifies the analysis
- However, there are still many correlation functions that are stationary and isotropic (spherical, power, exponential, etc)
- In this lecture we will introduce a few common examples, and discuss their properties

Spatial covariance models

- The spatial covariance models we describe below are isotropic (and thus stationary)
- This is rarely true, but mathematical assumptions are usually over-simplifications
- Often it is better for prediction and estimation to fit a simple stable model than a rich unstable model
- In the estimation section, we will discuss ways to determine if the isotropy assumption is grossly violated

Spatial covariance models

- Just like a correlation has to be between $[-1,1]$, the correlation function ρ has to satisfy constraints
- Of course $\rho(\mathbf{s}_i,\mathbf{s}_j)$ has to be in $[-1,1]$, but it must also be symmetric and positive definite
- Symmetric means that $\rho(\mathbf{s}_i,\mathbf{s}_j)=\rho(\mathbf{s}_j,\mathbf{s}_i)$
- The function ρ is positive definite if for any n locations $\mathbf{s}_1,...,\mathbf{s}_n$ and any values $y_1,...,y_n$ (not all zeros),

$$\sum_{i=1}^n\sum_{j=1}^n\rho(\mathbf{s}_i,\mathbf{s}_j)y_iy_j>0$$

- The correlation functions below are positive definite

Spatial covariance models

- The most common model is exponential,

$$\rho(d) = \exp(-d/\phi)$$

- The parameter $\phi > 0$ is the spatial range

- Solving $\rho(d) = 0.05$ gives $d = \log(20)\phi \approx 3\phi$

- So 3ϕ is often called the effective range, i.e., the distance at which observations are approximately uncorrelated

Spatial covariance models

- Another model is the squared exponential,

$$\rho(d) = \exp\{-(d/\phi)^2\}$$

- The Matern covariance includes both exponential and squared exponential as special cases

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

- This expression is hideous

- The key is that it adds a second parameter $\nu > 0$ to control the smoothness Z_i

Spatial covariance models

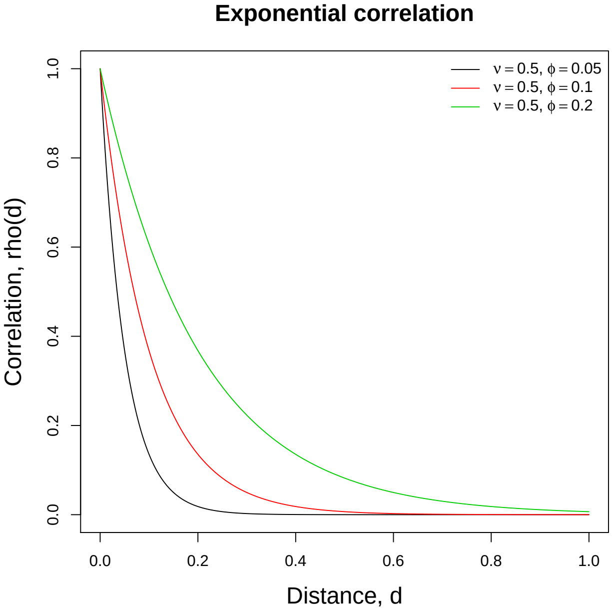
- Imagine taking a sample from Z_i at an infinite number of spatial locations $\mathbf{s}_1, \mathbf{s}_2, ...$
- Under the Matern correlation, Z as a function of $\mathbf{s}$ is $\nu - 1$ times mean-square differentiable
- Small ν and Z is a bumpy function of $\mathbf{s}$, e.g., if $\nu = 1/2$

$$\rho(d) = \exp(-d/\phi)$$

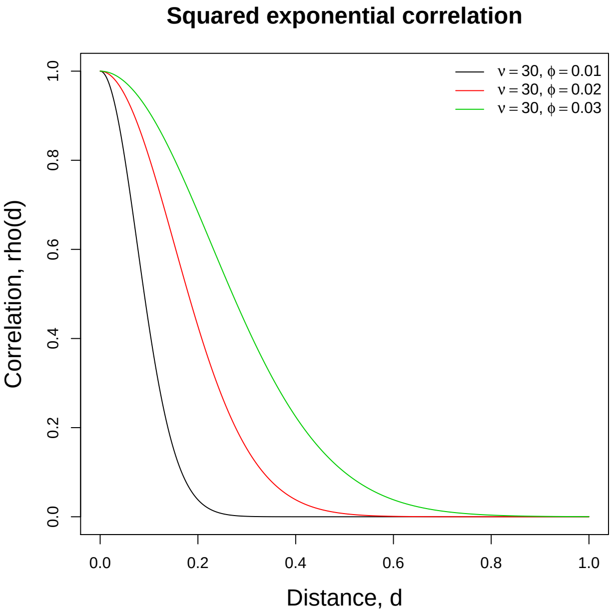
- Large ν and Z is a smooth function of $\mathbf{s}$, e.g., if $\nu = \infty$

$$\rho(d) = \exp\{-(d/\phi)^2\}$$

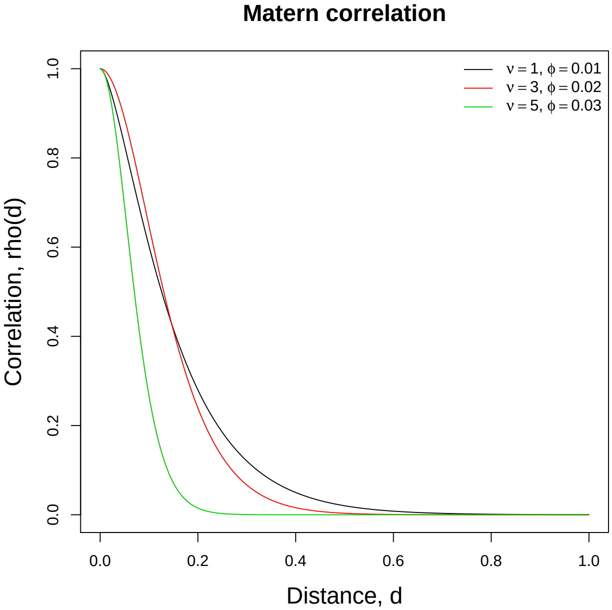
Spatial correlation models



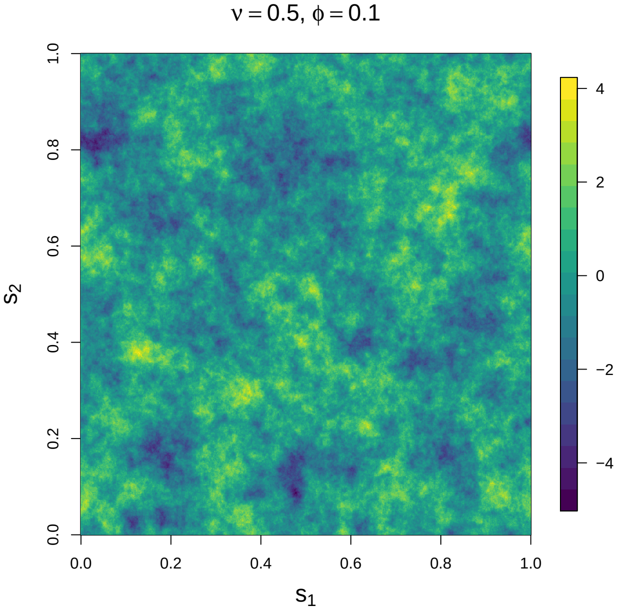
Spatial correlation models



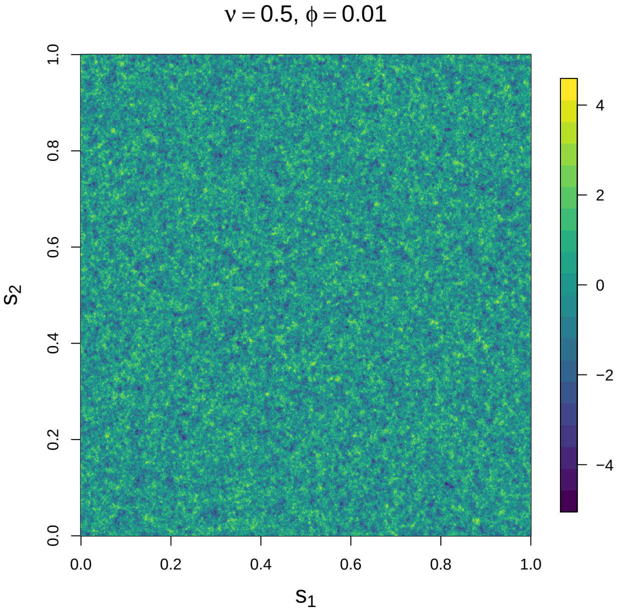
Spatial correlation models



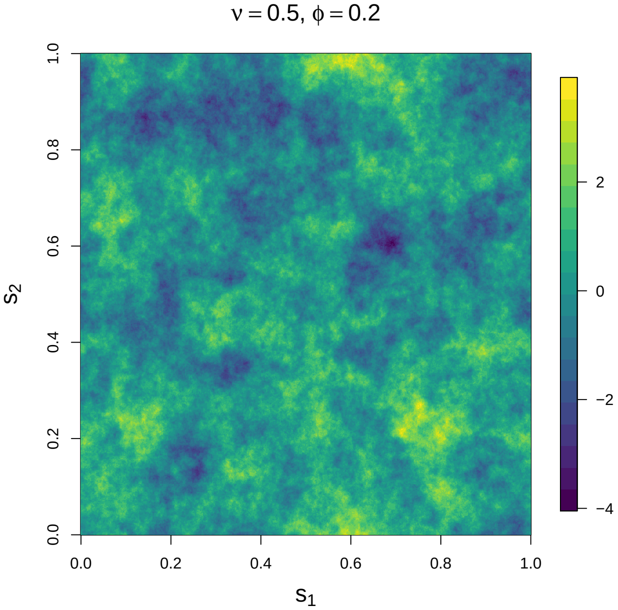
Spatial covariance models - random draws



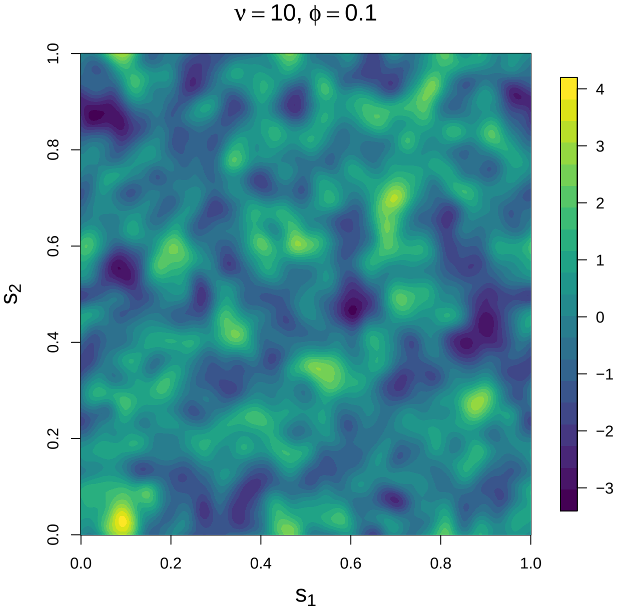
Spatial covariance models - random draws



Spatial covariance models - random draws



Spatial covariance models - random draws

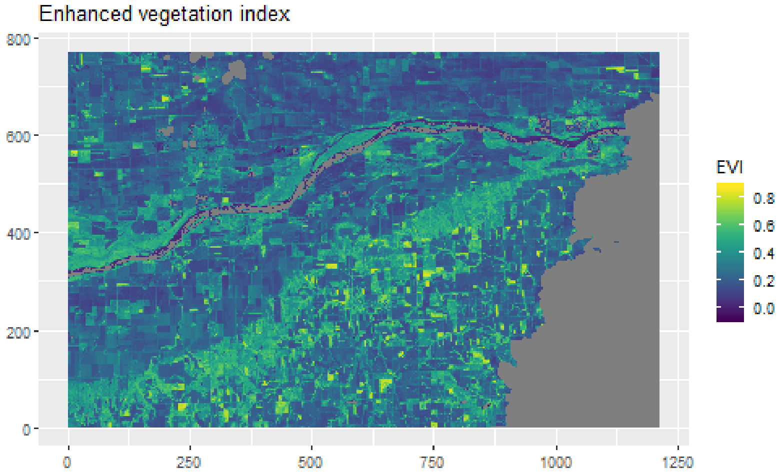


Models for the landsat data

Say we are trying to build a model for the Landsat data to impute the missing values

- What covariates might we include?
- Do you think we need a nugget?
- Does the process look stationary?
- Does the process look isotropic?
- If we fit a Matern correlation, do you think ν would be large or small?

Landsat data



Spatial prediction

- Prediction at a location without data is a fundamental task in spatial statistics
- Say we observe data at locations $\mathbf{s}_1, ..., \mathbf{s}_n$ and we want to make a prediction at location $\mathbf{s}_0$
- How might this be useful for the air pollution example?
-
-

Spatial prediction

- A natural predictive model is to use a linear combination of the observed data
$$\hat{Y}_0 = \sum_{i=1}^{n} w_i Y_i$$
- How to pick the weights, w_i?
- Let d_{i0} be the distance between $\mathbf{s}_i$ and $\mathbf{s}_0$
- In the absence of other information, we might assume the weights decrease with d_{i0}

Spatial prediction

- K nearest neighbors: Set $w_i = 1/K$ for the K sites closest to $\mathbf{s}_0$ and $w_i = 0$ for the rest
- Inverse distance weighting (IDW): $w_i = c/d_i$ where $1/c = \sum_{i=1}^{n} 1/d_i$ makes the weights sum to one
- Kernel smoothing: $w_i = ck(d_i)$ where $1/c = \sum_{i=1}^{n} k(d_i)$ and w_i is a kernel, say $\exp(-d_i^2)$
- Kriging: This is optimal (in some sense) and takes the weights to be a function of the spatial covariance