Spatial prediction (Kriging):

We seek to predict the data y_u another set of locations.

Use the posterior predictive distn:

$$[y_u \mid y_o] = \int\int\int \underbrace{[y_u \mid \beta, \sigma^2, \phi, y_o]}_{\text{predictive full-conditional}} \underbrace{[\beta, \sigma^2, \phi \mid y_o]}_{\text{posterior}} \, d\beta \, d\sigma^2 \, d\phi$$

y_u: unobs. data; y_o: obs. data

Important: if we could observe y_u, then we'd have the joint model:

$$\begin{pmatrix} y_o \\ y_u \end{pmatrix} \sim N\left(\begin{pmatrix} X_o\beta \\ X_u\beta \end{pmatrix}, \begin{pmatrix} \Sigma_{oo} & \Sigma_{ou} \\ \Sigma_{uo} & \Sigma_{uu} \end{pmatrix} \right)$$

We can use the MVN properties to find the pred. full-cond. distn as:

$$[y_u \mid \beta, \sigma^2, \phi, y_o] = N\left(X_u\beta + \Sigma_{uo}\Sigma_{oo}^{-1}(y_o - X_o\beta), \; \Sigma_{uu}^{-1} - \Sigma_{uo}\Sigma_{oo}^{-1}\Sigma_{ou} \right)$$

Use composition sampling to obtain $y_u^{(k)}$ for $k = 1, \ldots, K$

1.) obtain $\beta^{(k)}, \sigma^{2(k)}, \phi^{(k)}$ from posterior

2.) sample $y_u^{(k)} \sim [y_u \mid \beta^{(k)}, \sigma^{2(k)}, \phi^{(k)}, y_o]$

Point prediction is posterior mean: $E(y_u \mid y_o) = \dfrac{\sum_{k=1}^{K} y_u^{(k)}}{K}$

Hierarchical spatial models:

Treat spatially correlated random effects as latent process:

$$\underline{y} \sim N(X\underline{\beta} + \underline{\eta} \; , \; \sigma_y^2 I) \quad \Big] \text{ Data Model}$$

$$\underline{\eta} \sim N(\underline{0} \; , \; \Sigma) \; , \; \Sigma = \sigma_\eta^2 R(\phi) \quad \Big] \text{ process model}$$

$$\left.\begin{aligned} \underline{\beta} &\sim N(\underline{\mu}_\beta, \Sigma_\beta) \\ \sigma_\eta^2 &\sim IG(q_\eta, r_\eta) \\ \phi &\sim Gamma(a, b) \\ \sigma_y^2 &\sim IG(q_y, r_y) \end{aligned}\right] \text{ parameter model}$$

Notes:

1.) $\underline{\eta}$ can be viewed as a missing covariate or simply as a way to account for leftover dependence in $\underline{y}$ beyond that explained by $X\underline{\beta}$.

2.) σ_y^2 accounts for small-scale variation and/or measurement error.

Notes cont.

3.) Latent process η can be <u>integrated out</u>:

$$[y \mid \beta, \sigma_y^2, \sigma_\eta^2, \phi] = \int [y \mid \beta, \eta, \sigma_y^2][\eta \mid \sigma_\eta^2, \phi]\, d\eta$$
$$= N(X\beta, \sigma_y^2 I + \sigma_\eta^2 R(\phi))$$

which yields a <u>non-hierarchical model</u>!

we can fit either the hier. or non-hier. model and get the same inference about the parameters.

In fact, we can fit non-hier. model and then obtain samples of η post hoc by sampling from its full-conditional distn afterward

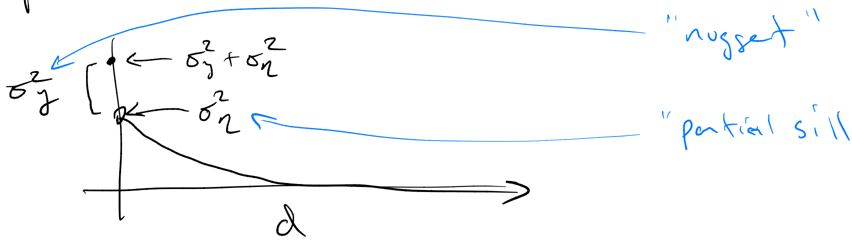
$$[\eta \mid \cdot] \propto [y \mid \beta, \eta, \sigma_y^2][\eta \mid \sigma_\eta^2, \phi]$$

$$\propto \exp\{-\tfrac{1}{2}(y - X\beta - \eta)'(\sigma_y^2 I)^{-1}(y - X\beta - \eta)\}\exp\{-\tfrac{1}{2}\eta'\Sigma^{-1}\eta\}$$
$$\propto \exp\{-\tfrac{1}{2}(-2\underbrace{(y - X\beta)'(\sigma_y^2 I)^{-1}}_{b'}\eta + \eta'\underbrace{((\sigma_y^2 I)^{-1} + \Sigma^{-1})}_{A}\eta)\}$$

$$= N(A^{-1}b, A^{-1})$$

Notes cont.

4.) Implied covariance fcn:



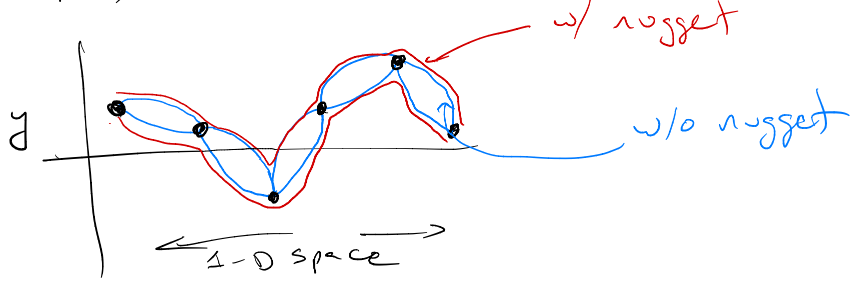
5.) Including nugget $\sigma_y^2 I$ adds a "ridge" to the spatial covariance matrix:

$$\sigma_y^2 I + \sigma_\eta^2 R(\phi)$$

which increases the diagonal elements and stabilizes the covariance matrix, making it easier to invert.

6.) If $\sigma_y^2 = 0$ the Bayesian geostat model is a <u>perfect interpolator</u> (i.e., predictions are forced to go through the data).

For example, in 1-D:



Notes cont.

7.) Could allow parameter to vary spatially:

For example, in time series:

$$y_{it} \sim N(\alpha_i y_{it-1}, \sigma_y^2), \quad i=1,\dots,n \quad t=1,\dots,T$$

spatially correlated dynamics

$$\left[\begin{array}{l} \underline{\alpha} = (\alpha_1, \dots, \alpha_n)' \\ \underline{\alpha} \sim N(\underline{0}, \sigma_\alpha^2 R(\phi)) \end{array}\right.$$

8.) Could change the data model. For counts,

$$\underline{y} \sim \text{Pois}(\underline{\lambda})$$

$$\log(\underline{\lambda}) \sim N(X\beta, \Sigma) \qquad \Sigma = \sigma^2 R(\phi)$$

This is a form of overdispersed Poisson model.

9.) Could use as a surrogate for a mechanistic model (i.e., statistical emulator).

10.) For discrete spatial process, change covariance function:

$$\Sigma = \sigma^2(\text{diag}(W\underline{1}) - \rho W)^{-1}$$

spatial correlation param.

adjacency matrix

$$W = \begin{bmatrix} 0 & 1 & 1 & 0 & \cdots \\ 1 & & & & \vdots \\ 1 & & \ddots & & \vdots \\ 0 & & & & \\ \vdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}$$

neighbors

not neighbors