Part 1: Geostatistical processes

Friday January 20 10.15-12 in room 656 Sentralbygg 2

Last week:

- How to model spatial continuous processes.
  - Hierarchical models
  - Gaussian Random Field (GRF)
  - Covariance functions (and variograms)

This week:

- Predictors for $Y|Z$ (assume Z known) (Tuesday)
- Estimation of covariance function (parameters θ) / variogram.
- Simulations + more models.

Hierachical statistical models (HM) for Mountain temperature

Temperature:

Data model: $[Z|Y,\theta]$ Observations given true temperature:
Independent $Z(s_i) = N(Y(s_i), \sigma_\epsilon^2)$

Process model: $[Y|\theta]$ Distribution for temperature.
$Y(s) = \beta_0 + \beta_1 h(s) + \delta(\theta)$ where $\delta \sim GRF$ with $E(\delta) = 0$ and covariance function $C_Y(s_1, s_2)$

Parameter model: $[\theta]$ Prior for parameters. $[\theta] = [\sigma_\epsilon^2][\beta][C_Y()]$.

- $h(s)$ is elevation (meters above see level)
- Can write as matrices: $Y(s) = X\beta$ with $X = [1, h(s)]^T$ and $\beta = [\beta_0, \beta_1]^T$

Modeling challenge: Daily precipitation

SeNorge

1. Suggest a spatial HM for daily precipitation?
2. What model/method/covariates do you think is used at SeNorge? Why?
3. How would you fit your model (make inference)?
4. How would you evaluate your model?

Inference about θ

Maximum likelihood: $\max_\theta L(\theta|Z) = \max_\theta f(Z|\theta)$

Bayesian: Find posterior $[\theta|Z]$

Computational cost?

Example: Temperature difference (pg 134-135)

$$Y(\mathbf{s}) = \mathbf{x}(\mathbf{s})'\boldsymbol{\beta} + \delta(\mathbf{s}), \qquad \mathbf{s} \in D_s\,,$$

where for $\mathbf{s} = (s_1, s_2)'$, $\mathbf{x}(\mathbf{s}) = (1, s_1, s_2, s_2^2, s_2^3)'$, and $\boldsymbol{\beta} \equiv (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)'$

GEOSTATISTICAL PROCESSES 135

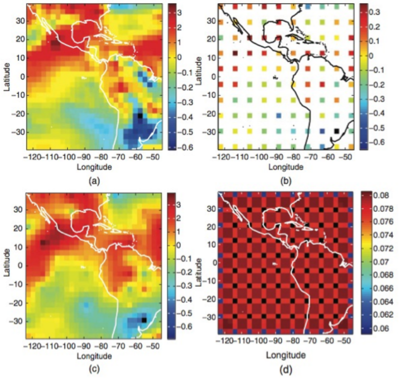

Figure 4.6 (a) Same plot as Figure 4.5a: Temperature change (1990s minus 1980s); the 28 × 28 values represent "the truth." (b) The 10 × 10 observations are obtained by subsampling "the truth" and adding mean-zero Gaussian noise. (c) Simple kriging predictor obtained from the missing and noisy data in (b). (d) Kriging standard error corresponding to simple kriging; the pattern is expected due to regular sampling in (b).

Multivariate Normal distribution

Multivariate Normal(MVN) density

$Y = (Y_1, Y_2, \ldots, Y_n)$ is MVN with expected value μ and covarance Σ, $Y \sim MVN(\mu, \Sigma)$ if

$$f(y) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp(-\frac{1}{2}(y-\mu)^T\Sigma^{-1}(t-\mu))$$

Conditional MVN

Let $Y = (Y_1, Y_2)^T$, $\mu = (\mu_1, \mu_2)^T$ and

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Then the $[Y_1|Y_2 = a] \sim MVN(\mu_{1|2}, \Sigma_{1|2})$ with

- $\mu_{1|2} = \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(a - \mu_2)$ and
- $\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$