Consider a random field $Z(s) = m(s)\beta + e(s)$, if we observe on $s_1, \cdots, s_n$, and $m(s)$ is of rank $p$ ($p < n$).

We said if we can find a matrix $K$ such that $E(KZ) = 0$ (so that $KZ$ is a contrast) and work with $KZ$ to get maximum likelihood estimates for the covariance parameters, we call these estimates REML estimates.

REML overcomes the problems of ML estimates.

Note though that here we consider $K$ such that its dimension is $(n - p) \times n$ (NOT $n \times n$) and each rows of $K$ are linearly independent.

Therefore, just using $I - m(s)(m(s)^T m(s))^{-1} m(s)^{-1}$ as $K$ would not be the right thing.

We may want to find an $(n - p) \times n$ matrix $K$ such that $KK^T = I$ and $KK^T = I - m(s)(m(s)^T m(s))^{-1} m(s)^{-1}$.
Soil Carbon and Nitrogen Data

- Source: Carbon Dioxide Information Analysis Center (CDIAC, http://cdiac.ornl.gov)
- The dataset available on the website is quite large
- I am going to use subset of the data for examples of this class
Location of samples

There are several measurements measured at the same location.
- The unit for Nitrogen is $g/m^2$ and the unit for Carbon is $Kg/m^2$
- We take log transformation to make things close to Gaussian
It is natural to have a model:

\[ C(s) = \beta_0 + \beta_1 N(s) + e(s) \]

Here \( C \) is log transformed Carbon and \( N \) is log transformed Nitrogen.

Now the question is how to estimate \( \beta_0 \) and \( \beta_1 \) and how to model \( e(\cdot) \).

This type of problem is called a \textit{spatial regression} problem.

Let’s start from a simple thing.
We assume fitting

\[ Z(s) = m(s)\beta + e(s) \]

where \( e(s) \sim (0, \sigma^2 I) \)

We get the estimates:

\[ \hat{\beta} = (m(s)^T m(s))^{-1} m(s)^T Z(s) \]
\[ \hat{\sigma}^2 = \frac{1}{n - \text{rank}(m(s))} (Z(s) - m(s)\hat{\beta})^T (Z(s) - m(s)\hat{\beta}) \]

How did we get the above estimates?

What other ways are there to estimate the parameters \( \beta \) and \( \sigma \)?
Some facts of the estimators

- For linear models, least squares and ML produce equivalent estimators.
- In fact, the ML estimator for $\beta$ and the bias-corrected ML estimator of $\sigma^2$ (the REML estimator) are equivalent to the OLS estimators given in the previous slide.
- For the OLS estimators, if the linear model is correct, we get the following:
  - $E(\hat{\beta}) = \beta$
  - $V(\hat{\beta}) = \sigma^2 (m(s)^T m(s))^{-1}$
  - $E(\hat{e}(s)) = 0$
  - $V(\hat{e}(s)) = \sigma^2 (I - m(s)(m(s)^T m(s))^{-1} m(s)^T)$
  - $1^T \hat{e}(s) = 0$
Estimates for the soil data

- We consider log transformed data
- Note that for our data, $Z(\cdot)$ is the total carbon amount and $m(\cdot)$ consists of two columns (one with one’s and the other with the total nitrogen)

- Let’s assume the residuals are independent across space and have constant unknown variance $\sigma^2$
- Using the formula previously, we get

$$\hat{\beta} = (-2.73, 0.76), \quad \hat{\sigma}^2 = 0.085$$
Fit

\begin{figure}
\begin{center}
\begin{tabular}{cc}
\hspace{1cm} & \hspace{1cm} \\
\hspace{1cm} & \\
\end{tabular}
\end{center}
\end{figure}
Variogram of residuals

Mikyoung Jun (Texas A&M)

stat647 lecture 10

October 6, 2009 11 / 18
Problems of residuals from OLS

- Although we assume $e(\cdot)$ is uncorrelated and has constant variance, the residuals do not have this property
  - rank deficient: if $m(s)$ is of rank $p$, than only $n - p$ residuals carry information and the remainders are redundant
  - correlated: recall the covariance matrix of the residuals
  - heteroscedastic: again recall the covariance matrix. Furthermore, the sill of semivariogram from the residuals does not reflect the variability of the process
- If we fit semivariograms of the residuals, we have to be careful
Some fits with residuals

- We now suppose the error field $e(\cdot)$ has spatial dependence.

- If I fit a Matérn covariance model with nugget effect to the residuals, I get $\hat{\alpha} = 0.046, \hat{\beta} = 97.44, \hat{\nu} = 0.40, \hat{\delta} = 0.029$.

- If I do the regular MLE (not with the residuals), I get $\hat{\alpha} = 0.040, \hat{\beta} = 50.73, \hat{\nu} = 0.52, \hat{\delta} = 0.029, \hat{\beta}_0 = -3.12, \hat{\beta}_1 = 0.80$.

- Note that the variance estimate from the regular MLE is around 0.078.
Fitted variogram

Green curves are from the fit with residuals
Red line is from the MLE estimates
Fitted variogram using MLE

Blue curves are from the MLE estimates
If I use the first $n - 2$ eigenvectors of $I - m(s)(m(s)^T m(s))^{-1} m(s)^T$, they are linearly independent and the corresponding eigenvalues are zero and thus they give contrasts. From that, I get the following covariance estimates:

$$\hat{\alpha} = 0.048, \hat{\beta} = 78.71, \hat{\nu} = 0.47, \hat{\delta} = 0.030$$

Note that the variance estimate from these is around 0.093