LECTURE NOTES

§ 1 | Another Application

Sediment decomposition: Let $S_k =$ (amount of sediment deposited in year k) for $0 \leq K \leq N = 10^4$ and we model:

$$S_{k+1} = \mu + (1-a) \cdot [S_k - \mu] + \eta_{k+1}$$

we could also write this as:

$$[S_{k+1} | S_k \sim N\big((1-a) \cdot [S_k - \mu], \sigma^2\big)]$$

Here $\eta_{k+1} \sim i.i.d N(0, \sigma^2)$ where $\sigma^2 = 10^{-4}$, and $a = 10^{-3}$ So, S_k with $\mu = 0$ would look something like this:

As shown this is a process that stays distributed about a mean.

Suppose we have noisy observations of S_t for a few years Let Y_i be our measurements

$$Y_i = S_{k_i} + \varepsilon_i$$

For $1 \leq i \leq N$ and $\varepsilon_i \sim i.i.d N(0, \sigma_\varepsilon^2)$ Our Goal: Estimate the total amount of sediment deposited

$$S_{\text{total}} = \sum_{k=1}^{N} S_k$$

Let $D_k = S_k - \mu$ we can rewrite this as:

$$D_k = \eta_k + (1-a) D_{k-1} = \eta_k + (1-a)\eta_{k-1} + (1-a)^2 D_{k-2} = \sum_{j \geq 0} (1-a)^j \eta_{k-j}$$

Thus, we know:

$$D_k \sim N\left(0, \sigma^2 \cdot \sum_{j \geq 0} (1-a)^2 j\right) = N\left(0, \sigma^2 \frac{1}{(2-a)a}\right)$$

And since $\sigma^2 = 10^{-4}$ and $a = 10^{-3}$ we can simplify this further, $D_k \sim N(0, \sim 0.05)$. Let's say that $(W_t)_{t \in \mathbb{R}}$ is a Brownian motion and $\eta_k = W_{\frac{k+1}{N}} - W_{\frac{k}{N}} \sim N\big(0, \frac{1}{N} = \sigma^2\big)$ Then,

$$D_k = \sum_{j \geq 0} (1-a)^j \left(W_{\frac{k-j+1}{N}} - W_{\frac{k-j}{N}}\right) \approx \sum_{j \geq 0} e^{-aj} \left(W_{\frac{k-j+1}{N}} - W_{\frac{k-j}{N}}\right) = \int_{-\infty}^{t} e^{-aN(t-s)} dW_s$$

Where, $t = \frac{k}{N}$ and $0 \leq t \leq 1$ Further simplifying we get,

$$D_k = \int_{-\infty}^{t} e^{-aN(t-s)} dW_s \sim N(0, \int_{-\infty}^{t} (e^{-aN(t-s)})^2 ds = \frac{1}{2aN}) \approx \frac{\sigma^2}{2a} + \mathcal{O}(a^2\sigma^2)$$

Let $U_t = \int_{-\infty}^{t} e^{-aN(t-s)} dW_s$ Also, $S_{\text{total}} \simeq N\mu + \sum_{k=1}^{N} U_{\frac{k}{N}} \sim N \cdot (\mu + \int_0^1 U_t dt)$

New question: Let $T = \int_0^1 U_t dt$ and $X_i = U_{t_i} + \varepsilon_i$ where $\varepsilon_i \sim i.i.d N(0, \sigma_\varepsilon^2)$ What is the conditional distribution of $(T | X = X_1, \ldots, X_N = X_N)$?

Date: 01 October 2018.

Now, we are trying to estimate the total integral, or the purple area below.

The points on the blue curve are values of U_t, and the purple area is: $\int_0^1 U_t dt = T$ To do this we need everything inside to be jointly Gaussian. i.e. we need to use the same white noise.

Using the lemma from last lecture, we only need their covariance to calculate this area. Recall the lemma, if (X, Y) is $N\left(0, \begin{bmatrix} \Sigma'_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}\right)$ then, $(X|Y = y) \sim N(\Sigma_{XY}\Sigma_{YY}^{-1}y, \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\intercal})$
To apply this we first need to expand T

$$T = \int_0^1 U_t dt = \int_0^1 \int_{-\infty}^t e^{-aN(t-s)} dW_s dt = \int_{-\infty}^1 \int_{\max(s,0)}^t e^{-aN(t-s)} dt dW_s$$

$$= \int_{-\infty}^1 \frac{1}{aN} e^{aN\min(0,s)} \left(1 - e^{-aN}\right) dW_s := \int_{-\infty}^1 \phi(s) dW_s$$

Now, we know the following:

$$\mathrm{var}[T] = \int_{-\infty}^1 \phi^2(s) ds$$

$$\mathrm{var}[X_i] = \mathrm{var}[U_{t_i}] + \sigma_\varepsilon = \frac{1}{2aN} + \sigma_\varepsilon^2$$

$$\mathrm{cov}[X_i, X_j] = \int_{-\infty}^{t_i} e^{-aN(t_i-s)} \cdot e^{-aN(t_j-s)} ds = \frac{1}{2aN}[e^{-aN(t_j-t_i)}]$$

for $i \neq j$ and $t_i < t_j$ Thus,

$$\mathrm{cov}[X_i, T] = \int_{-\infty}^t e^{-aN(t_i-s)} \phi(s) ds$$

where,

$$\Sigma = \begin{bmatrix} ||\phi^2|| & 0 & \dots & \dots & 0 \\ 0 & \frac{1}{2aN} + \varepsilon^2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \frac{1}{2aN} + \varepsilon^2 \end{bmatrix}$$

Where the columns are associated with T, X_1, X_2, and so forth. And the rows are similarly associated with T, X_1, X_2, and so forth. (This means the diagonal elements are the variances of T, X_1, X_2, $\dots$, X_N).