MTH 656

Uncertainty Quantification

Homework 1

---

1. The covariance function for a Brownian Bridge process on $0 \le t \le 1$ (Brownian motion conditioned on $B(0) = B(1) = 0$) is

$$C_B(s,t) = \min(s,t) - st.$$

   (a) Verify that the eigenfunctions and eigenvalues are as follows

$$\lambda_n = (n\pi)^{-2},$$
$$\phi_n(t) = \sqrt{2}\sin\left(t/\sqrt{\lambda_n}\right)$$

   for $n \ge 1$. (Bonus: Derive this from the intergral equation, by taking derivatives with respect to t twice.)

   (b) Write out and simplify the KL expansion for the Brownian Bridge $B(t), t \in [0,1]$ in terms of i.i.d. normal random variables $\{\xi_n\}_{n=1}^{\infty}$ with zero mean and unit variance.

   (c) Simulate and plot realizations from the truncated KL expansion using 20, 100, and 500 terms. Compare to realizations from `bridge.m` from the course website.

2. Download `KLCodes.zip` from the course website. Given a set of predictions of stream inflows (resulting, say, from equally probable weather forecast realizations), represent the uncertainty in the random field in terms of a finite number of continuous random variables via a Karhunen-Loeve expansion. This expansion can be used to propagate uncertainty through a system, but can also be sampled to produce a large ensemble of realizations with statistics conforming to the original data (data extrapolation). See `KLdriver.m`.

   (a) How many terms are needed to retain 70% of the variance (not energy), as measured in the Frobenius norm (i.e., Frobenius norm squared)? How many for 99%?

   (b) Modify the code to display realizations sampled from the truncated KL expansion which retains only 70% of the variance. Comment qualitatively on these realizations compared to those sampled from the full 5-term expansion. In particular, what undesirable thing is happening and why? Does including one more term fix the problem?

(c) Change the correlation length used to generate the simulated data (indirectly by changing the third `ConfidenceInterval` in `GenerateEnsembles.m`) via multiplication by 0.5 and 2. Which corresponds to the smaller correlation length? Comment on the number of terms *necessary to retain* (i.e., to retain 70% of the variance) in the resulting KL expansion for each case.

(d) In practice it is necessary to preserve positivity of inflows (there is a non-zero probability of negative inflows when using Gaussian random variables in the KL expansion). This can be accomplished by using the logarithm of the data in the construction of the KL expansion. Modify `KLdriver.m` by replacing `ProdData` with its `log`. (Be sure to modify the plots of the sampled realizations at the very end in order to be able to compare these results with the original code.) Comment on the number of terms necessary to retain in this KL expansion (i.e., to retain 70% of the variance). Also comment on the quality of the sampled realizations. Be sure to consider a characteristic set of realizations (e.g., you may have to look carefully at the tails and/or increase `ConfidenceLevel` in `GenerateEnsembles.m` and/or change `rng('default')` to `rng('shuffle')` in `KLdriver.m` to see differences). Feel free to comment on differences between lognormal realizations and data AND/OR differences between log normal realization and normal realizations.

It would be great if you could include some plots or numbers to support your answers.