Biostats 270: HW 2

Andrew Holbrook

Spring 2022

For the following, please include code and output in a single file. Assignment due by the beginning of class on Tuesday 5/3.

1. We wish to simulate Brownian motions on the interval $[0, 1]$.
   (a) Use Euler-Maruyama with stepsizes of $h \in (0.1, 0.01, 0.001, 0.0001)$ to simulate 20 independent Brownian motions each. For each stepsize, how long does it take to generate all 20? **Plot** the sample paths.
   (b) Use the KL expansion with with $p \in (5, 10, 20, 40)$ eigenfunctions to simulate 20 independent Brownian motions each. For each number of eigenfunctions, how long does it take to generate all 20? **Plot** the sample paths.
   (c) Use the explicit form of the Brownian motion covariance function to obtain covariance matrices on uniform partitions of $[0, 1]$ with 100, 1000, and 10,000 grid points. Use these covariance matrices to simulate 20 independent Brownian motions each. For each partition size, how long does it take to generate all 20? **Plot** the sample paths.

2. We wish to simulate Brownian bridges on the interval $[0, 1]$.
   (a) Use Euler-Maruyama with stepsizes of $h \in (0.1, 0.01, 0.001, 0.0001)$ to simulate 20 independent Brownian bridges each. For each stepsize, how long does it take to generate all 20? **Plot** the sample paths.
   (b) Use the KL expansion with with $p \in (5, 10, 20, 40)$ eigenfunctions to simulate 20 independent Brownian bridges each. For each number of eigenfunctions, how long does it take to generate all 20? **Plot** the sample paths.
   (c) Use the explicit form of the Brownian bridge covariance function to obtain covariance matrices on uniform partitions of $[0, 1]$ with 100, 1000, and 10,000 grid points. Use these covariance matrices to simulate 20 independent Brownian bridges each. For each partition size, how long does it take to generate all 20? **Plot** the sample paths.

3. We wish to draw samples from the 1-dimensional standard normal distribution. Use the 1D Fokker-Plank equation to derive the SDE with solution that leaves the standard normal target invariant. What is the name of this process? How does the parameterization of this process change when we change the variance of the Gaussian target?

4. The scaled Brownian motion $X(t)$ satisies the SDE

$$dX_t = \sigma dB_t \, .$$

(a) Derive its covariance function $K_\sigma(s,t) = E(X_s X_t)$.

(b) Derive its KL expansion over the interval $[0,1]$.

(c) Set seed to 1, simulate a single scaled Brownian motion ($\sigma = 10$) using Euler-Maruyama with stepsize $h = 0.001$ and randomly subsample to obtain 100 finite values $(t_1, y_1), \dots, (t_{100}, y_{100})$.

(d) Use the KL expansion (truncated at $p = 10, 40$) to specify a prior on the scaled Brownian motion $X(t)$ and place a truncated ($\sigma > 0$) standard normal distribution prior on σ. For the joint distribution of data $y_1, \dots, y_{100}$, specify

$$p(y_i|\sigma, t_i, X) \overset{iid}{\sim} \mathrm{N}(X_\sigma(t_i), \tau^2)$$

Finally, specify a truncated ($\tau > 0$) standard normal distribution prior on τ. For $p = 10, 40$, use Metropolis-Hastings to sample from the joint posterior $p(\sigma, \tau|y_1, \dots, y_{100}, t_1, \dots, t_{100})$. **Plot**:

i. The posteriors of σ and τ for both values of p and

ii. the posteriors of the truncated KL expansions for both values of p.

(e) What is the joint distribution $p(y_1, \dots, y_{100}|\sigma, t_1, \dots, t_{100})$? Place a truncated ($\sigma > 0$) standard normal distribution prior on σ and use Metropolis-Hastings to infer the posterior $p(\sigma|y_1, \dots, y_{100}, t_1, \dots, t_{100})$.

i. **Plot** the posterior of σ. How does it compare to the truth and the results from the truncated KL model?

ii. Use a uniform grid of time values $\zeta_1, \dots, \zeta_{500}$ on $[0,1]$ and the posterior distribution of σ to obtain the posterior distribution of the conditional mean

$$E(y(\zeta_1), \dots, y(\zeta_{500})|y_1, \dots, y_{100}, t_1, \dots, t_{100}, \sigma) \, .$$

Hint: use a 600×600 covariance function and the formula for the conditional mean of a multivariate Gaussian. Smooth over values to **plot** a posterior mean over the curve $E(y(t)|y_1, \dots, y_{100}, t_1, \dots, t_{100}, \sigma)$.