Assignment 4, due October 7

Corrections: [none yet]

1. (*Prediction, linear algebra, and conditional Gaussians*) A multi-component random variable is multivariate normal if its PDF has the form

$$p(x) = \frac{1}{Z} e^{-\frac{1}{2}(x-\mu)^t H(x-\mu)} \; .$$

The *precision* matrix H should be symmetric and positive definite. The mean is $\mu \in \mathbb{R}^n$. Let A be any symmetric positive definite matrix. Then it is possible to find a matrix L with $A = LL^t$. The *Cholesky factorization* is one way to do this, but not the only way. The covariance matrix of X is the $n \times n$ matrix $C = \text{cov}(X)$ with entries $C_{ii} = \text{var}(X_i)$ and $C_{ij} = \text{cov}(X_i, X_j)$ if $j \neq i$. This PDF is written $\mathcal{N}(\mu, C)$.

(a) Show that the covariance matrix is

$$C = \mathrm{E}\big[\,(X-\mu)(X-\mu)^t\big] = H^{-1} \; .$$

Hint: Use the substitution $y = \left(L^t\right)^{-1}(x - \mu)$ in the integral that represents the covariance matrix.

(b) Suppose $n = n_1 + n_2$ and we consider the first n_1 and the last n_2 components of X separately. Block matrix/block vector notation for this can be

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \; , \;\; X_1 \in \mathbb{R}^{n_1} \; , \;\; X_1 \in \mathbb{R}^{n_1} \; , \;\; X_2 \in \mathbb{R}^{n_2}$$

$$C = \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{12}^t & C_{22} \end{array}\right) \; , \;\; C_{11} = \text{cov}(X_1) \; , \; \text{etc.}$$

Show that X_1 and X_2 are independent if and only if $C_{12} = \text{cov}(X_1, X_2) = 0$. *Hint*: independence means the PDF is a product. A matrix is block diagonal if and only if its inverse is block diagonal.

(c) Let M be an $n \times n$ non-singular matrix. Suppose $Y = MX$. Show that Y is multivariate normal if and only if X is multivariate normal. *Hint*: find the precision matrix H_Y for Y.

(d) Suppose we know X_2 but not X_1. Let $\widehat{X}_1$ be a predictor of X_1 from X_2. A *linear predictor* has the form

$$\widehat{X}_1 = KX_2 + b \; .$$

Here K is an $n_1 \times n_2$ prediction matrix (also called *regression* matrix) and $b \in \mathbb{R}^{n_1}$ is to get the mean of X_1 right. The prediction *residual* is $R = X_1 - \widehat{X}_1$. Write a formula for K so that the conditional mean of X_1 is correct, which means

$$\widehat{X}_1 = \mathrm{E}[\, X_1 \mid X_2] \ .$$

Show that the prediction residual is independent of X_2. *Hint for the last part*: R and X_2 are jointly Gaussian because there is an M (why?) so that

$$\begin{pmatrix} R \\ X_2 \end{pmatrix} = M \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \ .$$

(e) Find a formula for $\mathrm{cov}(R)$, the covariance matrix of the prediction residual.

(f) Let $p(x_1|x_2)$ be the conditional PDF of X_1, conditional on knowing $X_2 = x_2$. Show that this conditional density is $\mathcal{N}(\mu(x_2), C(x_2))$. Find formulas for the conditional mean $\mu(x_2)$ and the conditional covariance $C(x_2)$.

2. (*Brownian bridge*) Let W_t be a standard Brownian motion path (this has $\mathrm{var}(\Delta W) = \Delta t$, and no drift). Suppose there are three times $0 \leq t_1 < t_2 < t_3$. Denote the values at these times by $W_{t_j} = W_j$, which is an abuse of notation. The *Brownian bridge* construction is a stochastic interpolation procedure where we create a Brownian motion path in a "top down" way. First we create the large scale structure (see a later assignment), then we fill in middle values in a way suggested by this exercise.

   (a) Find the mean vector and covariance for the three component random variable $W = (W_1, W_21, .W_3)$. To do this, multiply Gaussian transition densities with the unconditional density of W_1 to write a formula for the joint density. Use this to identify H and μ. Find the inverse of H to get the covariance matrix. This is an example of the general calculations in Exercise 1.

   (b) Use the method of Exercise 1 and the covariance calculations from part (a) to find the conditional distribution (mean and 2×2 covariance matrix) of (W_2, W_3) conditional on $W_1 = w_1$.

   (c) Use the method of Exercise 1 to find the conditional distribution of W_2 conditional on $W_1 = w_1$ and $W_3 = w_3$ known. Interpret the conditional mean $\overline{W}_2(w_1, w_3, t_1, t_3)$ using linear interpolation. Why is the conditional variance given both W_1 and W_3 lower than the conditional variance given only W_1?

3. Suppose W_t is a Brownian motion path and T is a random hitting time. The *stopped process* is

$$X_t = \begin{cases} W_t & \text{if } t < T \\ W_T & \text{if } t \geq T \ . \end{cases}$$

There are tricks involving martingales to calculate things about hitting times.

(a) Find a representation of X_t as an Ito integral

$$X_t = \int_0^t f_s dW_s \; .$$

Make sure the f_s you use is non-anticipating. You can do that by explaining the strategy function $F(w,t)$ so that $f_t = F(W_{[0,t]}, t)$. Conclude that the stopped process is a martingale and that

$$\mathrm{E}[X_t] = W_0 \; .$$

(b) Suppose $W_0 = 0$, and $x_l < 0 < x_r$, and that T is the first hitting time, which is

$$T = \min\left\{ t \mid W_t = x_l \text{ or } W_t = x_r \right\} .$$

Use the fact that this stopped process is a martingale to find a formula for $\Pr(W_T = x_l)$. There is some tricky reasoning here, beyond the simple martingale calculation. You may assume that $\mathrm{E}[T] < \infty$.

(c) In general, suppose Y_t is a martingale, T is a hitting time related to Y, and X_t is the stopped process (stopped at $t = T$). then (we will show in a later class) that X_t is a martingale too. Here, we apply this to hitting time for Brownian motion with drift $W_t - at$ with $a > 0$ and $W_0 = w_0 > 0$ as in Assignment 2. The martingale is $Y_t = (W_t - at) + at$ and T is the first time $W_t - at = 0$. Show (using methods from class 4, not the future) that Y_t is a martingale and find the Ito integral representation for the stopped process (not quite the same as part (a) above). Assume that $E[T] < \infty$ and find a simple formula for $E[T]$ in terms of a and w_0.

(d) Find an integral representation for the martingale $Y_t = W_t^2 - t$. You may use Ito's lemma in the integral form. Suppose $w_0 = 0$, and $x_l < 0 < x_r$. Let T be the first time either $W_t = a$ or $W_t = -a$. Find a formula for $E[T]$. Does it "scale" with a in a way that makes sense?

4. Suppose S_t is geometric Brownian motion with parameters r (risk free rate) and σ (vol). Let $v(s) = (K - s)_+$, which is the "positive part" of $(K - s)$. This is $(K - s)$ if $K - s > 0$ and 0 if $K - s < 0$. Find a formula for

$$f(s_0, t) = \mathrm{E}[\,(K - S_t)_+ \mid S_0 = s_0] \;\; .$$

Do this by representing S_t in terms of a Brownian motion W_t as in the class 4 notes. Write the expectation as an integral over w with the PDF of W_t, then calculate the integral. The answer has two terms, both involving the cumulative normal $N(a) = \Pr(Z < a)$, with $Z \sim \mathcal{N}(0,1)$. *Warning:* This is not the Black Scholes formula for the price of a European put because it isn't properly discounted. *Hint:* The integral over w runs from $-\infty$ to some number w_*, which depends on the parameters. In this range $(K - S_t)_+ = K - S_t$, so you get two integrals. The integrals involve exponentials with quadratics. You do those by completing the square, as we have done with Gaussian integrals before.

Computing exercise

This is an exercise in simulating a diffusion process using the SDE and using the approximate sample paths to estimate an expected value. There is no demo code to download for this. Most of what you need is in the demo codes from the first three assignments. You can build the code for this exercise mostly using cut-and-paste from earlier codes. Part of the exercise is to maintain coding quality and style. This applies particularly to output and graphics.

The *Euler method* (also called *Euler Mayurama*) for $dX_t = a(X_t)dt + b(X_t)dW_t$ is

$$X_{k+1} = X_k + a(X_k)\Delta t + b(X_k)\Delta W_k \ .$$

This uses the abuse of notation X_k for the approximation of X_{t_k}, where $t_k = k\Delta t$. The random variables ΔW_k are independent normals with mean zero and variance (you know this) Δt. For this problem, the SDE with be geometric Brownian motion $dS_t = rS_t dt + \sigma S_t dW_t$. You may use the formula $\mathrm{E}[S_t] = s_0 e^{rt}$.

1. Create a Python code to generate N sample paths up to time t with time step Δt [Adjust Δt down a little to make an integer number of time steps as was done in Assignment 3.] Use this to generate N independent approximate (approximate because $\Delta t > 0$) samples of the random variable S_t. Print out the sample mean and standard deviation to see that it converges to what you know is the right answer. Choose parameters that make the problem not too easy (parameters not too close to zero) and not too hard (very large). Give computational evidence that that the estimates converge to the right answer as $\Delta t \to 0$ and $N \to \infty$. The numbers you print should be formatted, well presented, and easy to read.

2. Now take $r = 0$ and $s_0 = 1$ and $\sigma = 1$. Show that the work needed to get an accurate estimate (code from part (a)) increases as t increases.

3. Write code to make a histogram of S_t. Look at the distribution for various values of t to see that it is more skew when t is larger. Explain the difficulty in estimating $\mathrm{E}[S_t]$ for large t.