Numerical Methods II
M. Giles

Problem sheet 4: solutions

1. (a) If a and b have zero expectation, then

$$\mathbb{V}[a+b] \;=\; \mathbb{E}[(a+b)^2] \;=\; \mathbb{E}[a^2] + \mathbb{E}[b^2] + 2\,\mathbb{E}[a\,b].$$

Since the correlation between a and b

$$\mathrm{corr}(a,b) = \frac{\mathbb{E}[a\,b]}{\sqrt{\mathbb{E}[a^2]\;\mathbb{E}[b^2]}}$$

is less than or equal to 1, it follows that

$$\mathbb{E}[a\,b] \leq \sqrt{\mathbb{E}[a^2]\;\mathbb{E}[b^2]}$$

and so

$$\mathbb{V}[a+b] \;\leq\; \mathbb{E}[a^2] + \mathbb{E}[b^2] + 2\sqrt{\mathbb{E}[a^2]\,\mathbb{E}[b^2]} \;=\; \left(\sqrt{\mathbb{E}[a^2]} + \sqrt{\mathbb{E}[b^2]}\right)^2$$

$$\Longrightarrow \quad \sqrt{\mathbb{V}[a+b]} \;\leq\; \sqrt{\mathbb{V}[b]} + \sqrt{\mathbb{V}[b]}$$

When a and b have non-zero expectation the result remains true because a has the same variance as $a - \mathbb{E}[a]$ which has zero expectation.
Equality requires that either $b\!=\!0$, or a is perfectly correlated with b, i.e.

$$a - \mathbb{E}[a] = \lambda\,(b - \mathbb{E}[b])$$

for some $\lambda \geq 0$.

<u>Extra bit</u>
Here is a proof from first principles that the correlation between two non-zero random variables a and b, each with zero expectation, is less than or equal to 1.
For any $\lambda > 0$,

$$\mathbb{E}[(a - \lambda b)^2] \;\geq\; 0$$

and hence

$$\mathbb{E}[a\,b] \;\leq\; \frac{1}{2\lambda}\,\mathbb{E}[a^2] + \frac{\lambda}{2}\,\mathbb{E}[b^2].$$

The expression on the right is minimised by taking

$$\lambda = \sqrt{\mathbb{E}[a^2]/\mathbb{E}[b^2]}$$

which then gives

$$\mathbb{E}[a\,b] \leq \sqrt{\mathbb{E}[a^2]\;\mathbb{E}[b^2]}.$$

Equality is achieved when there is a $\lambda > 0$ such that $\mathbb{E}[(a - \lambda b)^2] = 0$ and hence $a = \lambda\,b$.
Starting from $\mathbb{E}[(a + \lambda b)^2] \;\geq\; 0$ one can prove that $\mathbb{E}[a\,b] \geq -\sqrt{\mathbb{E}[a^2]\;\mathbb{E}[b^2]}$.

(b) Writing

$$a = (a+b) - b$$

we have

$$\sqrt{\mathbb{V}[a]} \;\leq\; \sqrt{\mathbb{V}[a+b]} + \sqrt{\mathbb{V}[b]}$$

and re-arranging gives the desired result.
Equality requires that either $b\!=\!0$, or

$$a\!+\!b - \mathbb{E}[a\!+\!b] = -\lambda\,(b - \mathbb{E}[b])$$

for some $\lambda \geq 0$, which is true if and only if

$$a - \mathbb{E}[a] = \mu\,(b - \mathbb{E}[b])$$

for some $\mu \leq -1$.

(c) Trivial proof by induction since

$$\sqrt{\mathbb{V}\left[\sum_{n=1}^{N} a_n\right]} \leq \sqrt{\mathbb{V}[a_N]} + \sqrt{\mathbb{V}\left[\sum_{n=1}^{N-1} a_n\right]}.$$

2. Following the reasoning in lecture 8, the conditional p.d.f. for the value of $W(t)$ is

$$p\Big(W(t) = w \mid W(t_n), W(t_{n+1})\Big)$$

$$= \frac{\dfrac{1}{\sqrt{2\pi(t\!-\!t_n)}}\ \exp\left(-\dfrac{(w\!-\!W(t_n))^2}{2\,(t\!-\!t_n)}\right)\ \dfrac{1}{\sqrt{2\pi(t_{n+1}\!-\!t)}}\ \exp\left(-\dfrac{(W(t_{n+1})\!-\!w)^2}{2\,(t_{n+1}\!-\!t)}\right)}{\dfrac{1}{\sqrt{2\pi(t_{n+1}\!-\!t_n)}}\ \exp\left(-\dfrac{(W(t_{n+1})\!-\!W(t_n))^2}{2\,(t_{n+1}\!-\!t_n)}\right)}$$

$$= \sqrt{\frac{(t_{n+1}\!-\!t_n)}{2\pi(t_{n+1}\!-\!t)(t\!-\!t_n)}}\ \exp\left(-\frac{\Big((t_{n+1}\!-\!t_n)\,w - (t\!-\!t_n)\,W(t_{n+1}) - (t_{n+1}\!-\!t)\,W(t_n)\Big)^2}{2\ (t_{n+1}\!-\!t_n)\ (t_{n+1}\!-\!t)\ (t\!-\!t_n)}\right)$$

and thus $W(t)$ is Normally distributed with mean

$$W(t_n) + \frac{t\!-\!t_n}{t_{n+1}\!-\!t_n}\ \Big(W(t_{n+1})\!-\!W(t_n)\Big)$$

and variance

$$\frac{(t_{n+1}\!-\!t)\ (t\!-\!t_n)}{(t_{n+1}\!-\!t_n)}.$$

This can be used in a Brownian Bridge construction to generate the value of intermediate points on the Brownian path, using a recursive approximate bisection approach. For example, suppose we wish to use 13 unit Normals Z_n to generate a path with 13 equally spaced timesteps. This would be done in the following way:

- Z_1 would be used to define W_{13}
- Z_2 would be used to define W_6
- Z_3, Z_4 would be used to define W_3, W_9
- Z_5, Z_6, Z_7, Z_8 would be used to define W_1, W_4, W_7, W_{11}
- $Z_9, Z_{10}, Z_{11}, Z_{12}, Z_{13}$ would be used to define $W_2, W_5, W_8, W_{10}, W_{12}$

At each level, we bisect each interval as closely as possible, rounding down where necessary. To explain some of the steps above in more detail, we would define

$$\begin{aligned}
W_6 &= \frac{6}{13} W_{13} + \sqrt{\frac{7\times 6}{13\times 13}\, T}\; Z_2 \\
W_9 &= \frac{4}{7} W_6 + \frac{3}{7} W_{13} + \sqrt{\frac{3\times 4}{7\times 13}\, T}\; Z_4 \\
W_7 &= \frac{2}{3} W_6 + \frac{1}{3} W_9 + \sqrt{\frac{1\times 2}{3\times 13}\, T}\; Z_7
\end{aligned}$$

In a QMC simulation, the Z_n would each equal $\Phi^{-1}(U_n)$ where $\Phi^{-1}(\cdot)$ is the inverse cumulative Normal function and U_n is a quasi-uniform value on the unit interval.

3. Given that $W(0) = W(1) = 0$, then for $0 < t_j < t_k < 1$ we know from the previous question that W_j is Normally distributed with zero mean and variance $t_j\,(1-t_j)$. Conditional on both this and $W(1)$, W_k is Normally distributed with mean

$$\frac{1-t_k}{1-t_j}\, W_j$$

and hence

$$\mathbb{E}[W_j\, W_k] \;=\; \frac{1-t_k}{1-t_j}\, \mathbb{E}[W_j^2] \;=\; t_j\,(1-t_k) \;=\; t_j - t_j\, t_k$$

If $t_j \geq t_k$ then $\mathbb{E}[W_j\, W_k] = t_k - t_j\, t_k$ and thus $\Omega_{jk} = \min(t_j, t_k) - t_j\, t_k$.

Ω has the structure

$$\Omega = \frac{1}{N^2} \begin{pmatrix}
N-1 & N-2 & N-3 & \ldots & 3 & 2 & 1 \\
N-2 & 2(N-2) & 2(N-3) & \ldots & 6 & 4 & 2 \\
N-3 & 2(N-3) & 3(N-3) & \ldots & 9 & 6 & 3 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
3 & 6 & 9 & \ldots & 3(N-3) & 2(N-3) & N-3 \\
2 & 4 & 6 & \ldots & 2(N-3) & 2(N-2) & N-2 \\
1 & 2 & 3 & \ldots & N-3 & N-2 & N-1
\end{pmatrix}.$$

It can be verified, slightly tediously, that multiplying Ω by the Ω^{-1} given in the question does indeed produce the identity matrix. It is then easily checked that the specified eigenvectors V_m and inverse eigenvalues λ_m^{-1} are the eigenvectors and eigenvalues of Ω^{-1} and hence that V_m and λ_m are the eigenvectors and eigenvalues of Ω.

4. (a) If we define $X_{1,n} \equiv \log S_1(nh)$ and $X_{2,n} \equiv \log S_2(nh)$ where $h = \frac{1}{4}$ then the SDES can be integrated exactly to give

$$\begin{aligned} X_{1,n+1} &= X_{1,n} + (r - \tfrac{1}{2}\sigma_1^2)\, h + \sigma_1\, \Delta W_{1,n} \\ X_{2,n+1} &= X_{2,n} + (r - \tfrac{1}{2}\sigma_2^2)\, h + \sigma_2\, \Delta W_{2,n} \end{aligned}$$

Given the computed values for X_n at each of the timesteps, the payoff evaluation is straightforward.

The Brownian increments $\Delta W_{1,n}$ and $\Delta W_{2,n}$ would each be Normally distributed with the required covariance

$$\Omega = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

This could be achieved by defining

$$\Delta W_n \equiv \begin{pmatrix} \Delta W_{1,n} \\ \Delta W_{2,n} \end{pmatrix} = \sqrt{h} \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} Z$$

where the two components of Z are i.i.d. unit Normals.

(b) The p.d.f. for the transition from X_n to X_{n+1} is

$$p_n = \frac{1}{2\pi\sqrt{|\Omega|}} \exp\Big(-\tfrac{1}{2}(X_{n+1} - X_n - a\, h)^T\ \Omega^{-1}\ (X_{n+1} - X_n - a\, h)\Big)$$

where

$$a \equiv \begin{pmatrix} r - \frac{1}{2}\sigma_1^2 \\ r - \frac{1}{2}\sigma_2^2 \end{pmatrix}$$

and

$$\Omega = \begin{pmatrix} \sigma_1^2 h & \rho\sigma_1\sigma_2 h \\ \rho\sigma_1\sigma_2 h & \sigma_2^2 h \end{pmatrix}.$$

Using the Likelihood Ratio Method, the "score function" for the Deltas is

$$\begin{aligned} \left(\frac{\partial \log p_0}{\partial S(0)}\right)^T &= \left(\frac{\partial \log X(0)}{\partial S(0)}\right)^T \left(\frac{\partial \log p_0}{\partial X(0)}\right)^T \\ &= \begin{pmatrix} S_1(0)^{-1} & 0 \\ 0 & S_2(0)^{-1} \end{pmatrix} \Omega^{-1}\ (X_1 - X_0 - a\, h) \end{aligned}$$

i.e. the Deltas are equal to the expectation of this score function multiplied by the payoff function.

(c) The problem with using the pathwise sensitivity approach is that the payoff is discontinuous; an infinitesimal perturbation to the path can make the difference between zero and unit payoff.

An overestimate for the value of the option is given by defining the smoothed payoff to be

$$H_\varepsilon\left(\sum_n H_\varepsilon\Big(B - \min(S_{1,n}, S_{2,n})\Big) - 2.5\right)$$

where the approximate Heaviside function $H_\varepsilon(\cdot)$ for $\varepsilon > 0$ is defined as

$$H_\varepsilon(x) = \begin{cases} 0, & x \le -\varepsilon \\ 1 + \dfrac{x}{\varepsilon}, & -\varepsilon < x < 0 \\ 1, & x \ge 0 \end{cases}$$

It is clearly an overestimate since $H_\varepsilon(x) \ge H_0(x)$, and $H_0(x) \ge H_0(y)$ for $x \ge y$, and hence

$$\begin{aligned} & H_\varepsilon\left(\sum_n H_\varepsilon\Big(B - \min(S_{1,n}, S_{2,n})\Big) - 2.5\right) \\ & \qquad \ge \; H_0\left(\sum_n H_\varepsilon\Big(B - \min(S_{1,n}, S_{2,n})\Big) - 2.5\right) \\ & \qquad \ge \; H_0\left(\sum_n H_0\Big(B - \min(S_{1,n}, S_{2,n})\Big) - 2.5\right) \end{aligned}$$

An underestimate is obtained by changing the definition of $H_\varepsilon(\cdot)$ to

$$H_\varepsilon(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{x}{\varepsilon}, & 0 < x < \varepsilon \\ 1, & x \ge \varepsilon \end{cases}$$