THE UNIVERSITY OF NEW SOUTH WALES

SCHOOL OF MATHEMATICS AND STATISTICS

November 2015

MATH3311/MATH5335

Mathematical Computing for Finance

Computational Methods for Finance

Solutions

(1) TIME ALLOWED – 2 HOURS

(2) TOTAL NUMBER OF QUESTIONS – 4

(3) ANSWER **QUESTIONS 1 AND 2.**

ANSWER **EITHER** QUESTION 3 **OR** QUESTION 4.
(ONLY ONE OF QUESTIONS 3 AND 4 CAN COUNT.)

(4) THE QUESTIONS ARE OF EQUAL VALUE

(5) THIS PAPER MAY BE RETAINED BY THE CANDIDATE

(6) **ONLY** CALCULATORS WITH AN AFFIXED "UNSW APPROVED" STICKER MAY BE USED

All answers must be written in ink. Except where they are expressly required pencils may only be used for drawing, sketching or graphical work.

1. A deterministic interest rate model $r(t)$ at time $t \in [0, T]$ interpolates the $m = 9$ data values in Table 1. The data is stored in MATLAB in column vectors `td` and `rd`.

i	1	2	3	4	5	6	7	8	9
t_i (years)	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$r(t_i)$ (%)	3.5	3.25	4.1	4.5	4.2	3.3	3.2	3.05	4.12

Table 1: Interest rate data

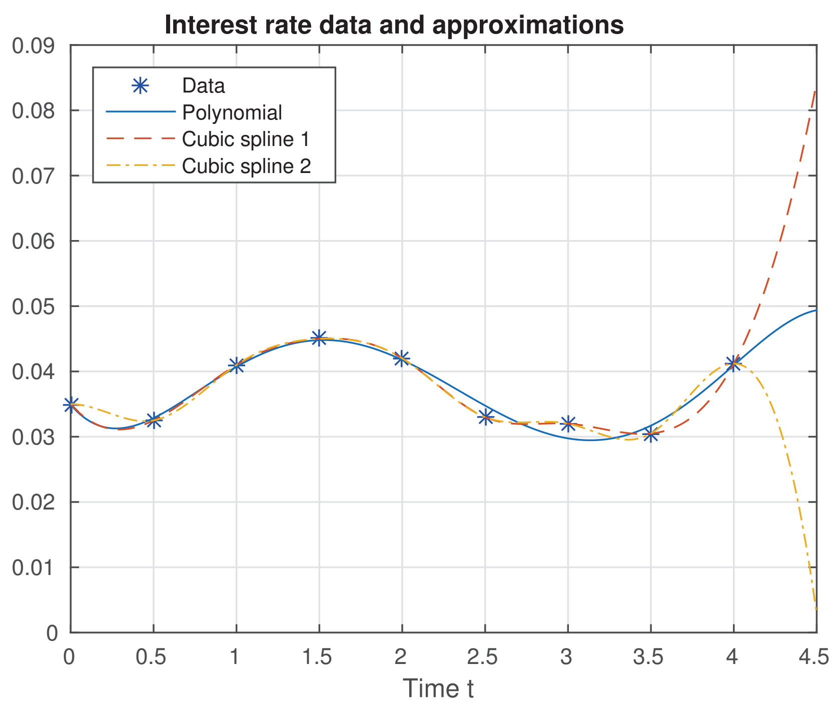

Figure 1: Interest rate data and approximations

i) An analyst decides to try to fit a polynomial of degree $n = 5$ with coefficients in the vector `a` to the interest rate data. They define the Vandermonde matrix $A_{ij} = t_i^{j-1}$ for $i = 1, \ldots, m$, $j = 1, \ldots, n + 1$. Some of the MATLAB commands they use are

```
a = A \ rd;
sv = svd(A);
svmax = sv(1)
      1.2235e+03
svmin = sv(end)
      6.4320e-02
```

They then plot the data and the polynomial as in Figure 1.

a) Explain why the polynomial approximation does not agree with the interest rate data, for example at $t_6 = 2.5, t_7 = 3$ or $t_8 = 3.5$ years.

Answer: We have 9 data points and use a polynomial of degree 5. The polynomial has 6 coefficients which we can choose. Hence in general it is not possible to fit the polynomial such that it goes through all 9 data points (in general it can only go through at most 6 data points, which is the case in this example).

b) What is the rank of the matrix A? Give reasons for your answer.

Answer: All singular values are positive and hence the matrix has full rank, which is 6.

c) Estimate the condition number $\kappa_2(A)$ of A.

Answer:

$$\kappa_2(A) = \frac{\text{svmax}}{\text{svmin}} = \frac{1.2235 * 10^3}{6.4320 * 10^{-2}} = 1.9022 * 10^4$$

d) Estimate the relative change in the polynomial coefficients `a` if there is a one basis point (0.01 of a percent) decrease in the interest rate at some time t_i. What confidence do you have in the values of `a`?

Answer: The relative change is

$$0.01\kappa_2(A) = 10^{-4} * 1.9022 * 10^4 = 1.9022.$$

Since the relative change is larger than 1, I don't have any confidence in the solution.

ii) The analyst then decides to interpolate the data with cubic splines using

```
tp = linspace(0, 4.5, 5001);
rsp1 = spline(td, rd, tp);
rsp2 = spline(td, [0 rd 0], tp);
```

and adds these two splines to the plot in Figure 1.

a) Explain what a "cubic spline interpolant" is.

Answer: A cubic spline interpolant is a piecewise polynomial of degree 3 which goes through all the data points such that the resulting function is twice continuously differentiable.

b) Why do the two cubic spline interpolants predict very different estimates for the interest rate at $t = 4.5$ years?

Answer: The cubic spline requires two additional conditions to be defined uniquely. These two conditions at the end points are different for those two splines, yielding different behaviour outside the range. (Spline 2 has derivative 0 at the end points, whereas Spline 1 is a not-a-knot spline.)

iii) Suppose that $m \geq n$ and that the rectangular matrix $A \in \mathbb{R}^{m\times n}$ has the singular value decomposition

$$A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T = Y\Sigma V^T = \sum_{j=1}^{n} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^T,$$

where

$$\begin{aligned} U &= [Y \quad Z] \in \mathbb{R}^{m\times m}, \\ Y &= [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_n] \in \mathbb{R}^{m\times n}, \\ Z &= [\boldsymbol{u}_{n+1} \quad \boldsymbol{u}_{n+2} \quad \cdots \quad \boldsymbol{u}_m] \in \mathbb{R}^{m\times(m-n)}, \\ \Sigma &= \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \in \mathbb{R}^{n\times n}, \\ V &= [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_n] \in \mathbb{R}^{n\times n}, \end{aligned}$$

with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.

iv) State the definition of an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$.

Answer: The matrix Q is orthogonal if $QQ^T = I$, where I is the identity matrix.

v) Show that if Q is orthogonal then $\|Q\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Hint: $(\|\boldsymbol{x}\|_2)^2 = \boldsymbol{x}^T\boldsymbol{x}$.

Answer: We have

$$\|Q\boldsymbol{x}\|_2^2 = (Q\boldsymbol{x})^T \cdot (Q\boldsymbol{x}) = \boldsymbol{x}^T Q^T Q\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|_2^2.$$

The result follows now by taking square roots on both sides.

vi) For $\boldsymbol{b} \in \mathbb{R}^m$ and $\boldsymbol{y} = V^T\boldsymbol{x}$, show that $\|A\boldsymbol{x} - \boldsymbol{b}\|_2$ is minimized when $\|\Sigma\boldsymbol{y} - Y^T\boldsymbol{b}\|_2$ is minimized.

Answer: We have

$$\|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 = \|U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T\boldsymbol{x} - \boldsymbol{b}\|_2^2 = \|U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \boldsymbol{y} - \boldsymbol{b}\|_2^2.$$

Please see over ...

Since U and also U^T is an orthogonal matrix, we have by the answer to the previous question that

$$\|A\boldsymbol{x}-\boldsymbol{b}\|_2^2 = \|U^T A\boldsymbol{x}-U^T\boldsymbol{b}\|_2^2 = \|U^T U \begin{bmatrix}\Sigma \\ 0\end{bmatrix}\boldsymbol{y}-U^T\boldsymbol{b}\|_2^2 = \|\begin{bmatrix}\Sigma \\ 0\end{bmatrix}\boldsymbol{y}-U^T\boldsymbol{b}\|_2^2.$$

Using Pythagoras Theorem we obtain

$$\|A\boldsymbol{x}-\boldsymbol{b}\|_2^2 = \|\begin{bmatrix}\Sigma \\ 0\end{bmatrix}\boldsymbol{y}-U^T\boldsymbol{b}\|_2^2 = \|\Sigma\boldsymbol{y}-Y^T\boldsymbol{b}\|_2^2 + \|Z^T\boldsymbol{b}\|_2^2.$$

The term $\|Z^T\boldsymbol{b}\|_2^2$ does not depend on $\boldsymbol{x}$ and $\boldsymbol{y}$, hence $\|A\boldsymbol{x}-\boldsymbol{b}\|_2^2$ is minimized if and only if $\|\Sigma\boldsymbol{y}-Y^T\boldsymbol{b}\|_2^2$ is minimized.

2. A consultant wishes to model the weekly price data shown in the first plot of Figure 2.

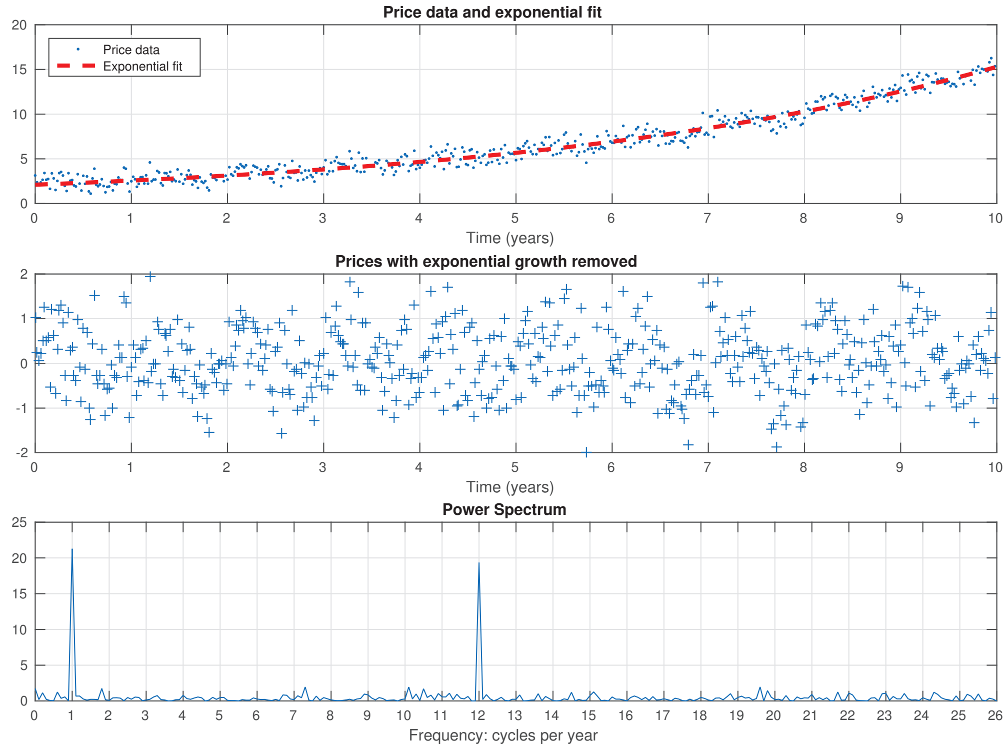

Figure 2: Price data, exponential fit and power spectrum.

Let

$$N = 520, \qquad T = 10, \qquad \Delta t = \frac{T}{N},$$

and denote the jth price by P_j and the corresponding time (in years) by

$$t_j = \frac{j}{52} = j\,\Delta t \quad \text{for } j = 0,\,1,\,2,\,\ldots,\,N-1.$$

Assume that the data values are already stored in two $N \times 1$ MATLAB arrays `P` $= [P_0, P_1, \ldots, P_{N-1}]^T$ and `t` $= [t_0, t_1, \ldots, t_{N-1}]^T$. (Thus, $P_j =$ `P(j+1)` and $t_j =$ `t(j+1)`.)

i) Explain how to use MATLAB to estimate the exponential growth underlying the price of this asset (shown by the dashed curve)

$$P_j \approx \alpha e^{\beta t_j} \quad \text{for } j = 0, \ldots, N-1.$$

Answer: Taking logarithms we obtain $\log P_j \approx \log\alpha + \beta t_j$. We can now use least-squares to estimate $\gamma = \log\alpha$ and β. Define the matrix A by $A = [\mathrm{ones}(N,1), \mathtt{t}]$ and solve the normal equations $A^T A \begin{pmatrix}\gamma\\ \beta\end{pmatrix} = A^T \log \mathtt{P}$ using $[\gamma,\beta]^T = A\backslash \log \mathtt{P}$. Then set $\alpha = \exp(\gamma)$.

ii) It is suspected there is a periodic component y_j in the data. Let

$$P_j = \alpha e^{\beta t_j} + y_j, \quad \text{for } j = 0, \ldots, N-1,$$

using the values of α and β found in part i). The discrete Fourier transform is

$$\hat{y}_k = \sum_{j=0}^{N-1} y_j e^{-i2\pi jk/N} \quad \text{for } k = 0,\, 1,\, \ldots,\, N-1.$$

where $i = \sqrt{-1}$.

a) Show that

$$\hat{y}_{N-k} = \overline{\hat{y}}_k = \text{complex conjugate of } \hat{y}_k \quad \text{for } k = 1,\, 2,\, \ldots,\, N-1.$$

Answer: We have

$$\begin{aligned}
\hat{y}_{N-k} &= \sum_{j=0}^{N-1} y_j e^{-i2\pi j(N-k)/N} = \sum_{j=0}^{N-1} y_j e^{i2\pi jk/N} e^{-i2\pi j} \\
&= \sum_{j=0}^{N-1} y_j e^{i2\pi jk/N} = \overline{\sum_{j=0}^{N-1} y_j e^{-i2\pi jk/N}} = \overline{\hat{y}}_k,
\end{aligned}$$

where we used that $\overline{y_j} = y_j$ (i.e., y_j is real).

b) Let $n = 260$ so that $N = 2n$. Show that

$$y_j = a_0 + \sum_{k=1}^{n} \big[a_k \cos(2\pi k t_j/T) + b_k \sin(2\pi k t_j/T)\big]$$

for $j = 0,\, 1,\, 2,\, \ldots,\, N-1$ and express the real coefficients a_k and b_k in terms of the (complex) transform $\hat{y}_k$. You may assume the inversion formula

$$y_j = \frac{1}{N} \sum_{k=0}^{N-1} \hat{y}_k e^{i2\pi jk/N}.$$

Answer: Using Euler's formual $e^{i2\pi x} = \cos(2\pi x) + i\sin(2\pi x)$ we obtain from the inversion formula that

$$y_j = \frac{1}{N}\sum_{k=0}^{N-1} \hat{y}_k e^{i2\pi jk/N} = \frac{\hat{y}_0}{N} + \sum_{k=1}^{N-1}\left[\frac{\hat{y}_k}{N}\cos(2\pi jk/N) + i\frac{\hat{y}_k}{N}\sin(2\pi jk/N)\right].$$

Since the y_j are real numbers the imaginary part in the above expression is 0 and we have

$$y_j = \frac{\mathrm{Re}(\hat{y}_0)}{N} + \sum_{k=1}^{N-1}\left[\frac{\mathrm{Re}(\hat{y}_k)}{N}\cos(2\pi jk/N) - \frac{\mathrm{Im}(\hat{y}_k)}{N}\sin(2\pi jk/N)\right].$$

Since $N = 2n$ is an even integer, we have for $k = 1, 2, \ldots, n-1$ that

$$\begin{aligned}\cos(2\pi jk/N) + \cos(2\pi j(N-k)/N) &= \cos(2\pi jk/N) + \cos(2\pi j - 2\pi jk/N) \\ &= 2\cos(2\pi jk/N)\end{aligned}$$

and

$$\begin{aligned}\sin(2\pi jk/N) - \sin(2\pi j(N-k)/N) &= \sin(2\pi jk/N) - \sin(2\pi j - 2\pi jk/N) \\ &= \sin(2\pi jk/N) + \sin(2\pi jk/N) \\ &= 2\sin(2\pi jk/N).\end{aligned}$$

Using $\hat{y}_{N-k} = \overline{\hat{y}_k}$ we obtain

$$\begin{aligned}y_j &= \frac{\mathrm{Re}(\hat{y}_0)}{N} + \sum_{k=1}^{n-1}\left[\frac{2\mathrm{Re}(\hat{y}_k)}{N}\cos(2\pi jk/N) - \frac{2\mathrm{Im}(\hat{y}_k)}{N}\sin(2\pi jk/N)\right] \\ &\quad + \left[\frac{\mathrm{Re}(\hat{y}_n)}{N}\cos(2\pi jn/N) - \frac{\mathrm{Im}(\hat{y}_k)}{N}\sin(2\pi jn/N)\right].\end{aligned}$$

We have $t_j = j\Delta t = j\frac{T}{N}$, hence $\frac{j}{N} = \frac{t_j}{T}$. Therefore we obtain

$$\begin{aligned}y_j &= \frac{\mathrm{Re}(\hat{y}_0)}{N} + \sum_{k=1}^{n-1}\left[\frac{2\mathrm{Re}(\hat{y}_k)}{N}\cos(2\pi kt_j/T) - \frac{2\mathrm{Im}(\hat{y}_k)}{N}\sin(2\pi kt_j/T)\right] \\ &\quad + \left[\frac{\mathrm{Re}(\hat{y}_n)}{N}\cos(2\pi nt_j/T) - \frac{\mathrm{Im}(\hat{y}_k)}{N}\sin(2\pi nt_j/T)\right].\end{aligned}$$

Thus we obtain

$$a_0 = \frac{\mathrm{Re}(\hat{y}_0)}{N}, a_k = \frac{2\mathrm{Re}(\hat{y}_k)}{N}, b_k = -\frac{2\mathrm{Im}(\hat{y}_k)}{N} \text{ for } k = 1, 2, \ldots, n-1$$

and $a_n = \frac{\mathrm{Re}(\hat{y}_n)}{N}$ and $b_n = -\frac{\mathrm{Im}(\hat{y}_n)}{N}$. Note that $\sin(2\pi nt_j/T) = \sin(2\pi jn/N) = \sin(\pi j) = 0$, hence the value of b_n is irrelevant and we can also set $b_n = 0$. Further, from the previous question it also follows that $\hat{y}_n = \overline{\hat{y}_{N-n}} = \overline{\hat{y}_{2n-n}} = \overline{\hat{y}_n}$, thus $\hat{y}_n \in \mathbb{R}$. Thus $a_n = \frac{\hat{y}_n}{N}$.

c) Define the power spectrum in terms of the Discrete Fourier Transform $\hat{y}_k$ for $k = 0, \ldots, N-1$.

Answer: The power spectrum is given by $\frac{|\hat{y}_k|^2}{N}$.

d) Justify or refute the claim that the data in Figure 2 has both yearly and monthly cycles of roughly the same intensity.

Answer: The third graph in Figure 2 shows peaks at 1 and 12, where 1 and 12 denote the cycles per year. Hence the yearly frequency and the frequency with 12 cycles per year, i.e., a monthly cycle have the most intensity (of roughly the same size).

ANSWER EITHER QUESTION 3 OR QUESTION 4, NOT BOTH

3. The random vector $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]^T \in \mathbb{R}^n$ is normally distributed with mean $\boldsymbol{\mu} = [\mu_1, \mu_2, \ldots, \mu_n]^T \in \mathbb{R}^n$ and covariance matrix $C \in \mathbb{R}^{n\times n}$. Thus, $\boldsymbol{X}$ has the probability density function

$$p(\boldsymbol{x}) = \frac{\exp\bigl(-(\boldsymbol{x}-\boldsymbol{\mu})^T C^{-1}(\boldsymbol{x}-\boldsymbol{\mu})/2\bigr)}{\sqrt{(2\pi)^n \det(C)}} \quad \text{for } \boldsymbol{x} \in \mathbb{R}^n.$$

i) What tests should you apply to check that a provided matrix C is a valid covariance matrix for use in calculating $p(\boldsymbol{x})$?

Answer: Check that the matrix is symmetric (using a matrix norm and tolerance level) and postive definite (using the MATLAB function `chol`).

ii) Explain how the Cholesky factorization $C = R^T R$ can be used to

a) evaluate $\sqrt{\det(C)}$ efficiently;

Answer: The Cholesky factor R is upper triangular, hence $\det(R) = \prod_i R_{i,i}$, where $R_{i,i}$ are the diagonal elements of R. Then

$$\det(C) = \det(R^T R) = \det(R^T)\det(R) = \det(R)\det(R).$$

Thus

$$\sqrt{\det(C)} = \det(R) = \prod_i R_{i,i},$$

i.e., the product of the diagonal elements of R.

b) evaluate $p(\boldsymbol{x})$ efficiently.

Answer: Since $C = R^T R$, we have $C^{-1} = R^{-1}R^{-T}$. Hence

$$(\boldsymbol{x}-\boldsymbol{\mu})^T C^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = (\boldsymbol{x}-\boldsymbol{\mu})^T R^{-1}R^{-T}(\boldsymbol{x}-\boldsymbol{\mu}) = [R^{-T}(\boldsymbol{x}-\boldsymbol{\mu})]^T[R^{-T}(\boldsymbol{x}-\boldsymbol{\mu})].$$

We first compute the Cholesky factor R of C. We compute $\sqrt{\det(C)}$ by computing the product of the diagonal elements of R. Then we compute $\boldsymbol{y}$ which satisfies $R^T\boldsymbol{y} = \boldsymbol{x}-\boldsymbol{\mu}$ by solving a linear system (using forward substitution). Then we compute

$$\frac{\exp(-\boldsymbol{y}\cdot\boldsymbol{y}/2)}{(2\pi)^{n/2}\prod_i R_{i,i}}$$

to get the value of $p(\boldsymbol{x})$.

iii) From now on, suppose for simplicity that $n = 2$, and consider the problem of evaluating the expected value of $f(X_1, X_2)$,

$$\mathbb{E}[f(X_1, X_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) p(x_1, x_2)\, dx_1\, dx_2, \qquad (1)$$

for some given function $f(x_1, x_2)$. The standard normal pdf and cdf functions are, for $x, y \in \mathbb{R}$,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad \Phi(y) = \int_{-\infty}^{y} \phi(x) dx.$$

a) Show that the substitution $\boldsymbol{x} = \boldsymbol{\mu} + R^T \boldsymbol{y}$, where R is the Cholesky factor of C, gives

$$p(x_1, x_2) = \frac{\phi(y_1)\phi(y_2)}{R_{11} R_{22}}.$$

Answer: Using the answer from the previous question, we have

$$p(x_1, x_2) = \frac{\exp(-(y_1^2 + y_2^2)/2)}{2\pi R_{1,1} R_{2,2}} = \frac{\phi(y_1)\phi(y_2)}{R_{1,1} R_{2,2}}.$$

b) Transform the expected value (1) to

$$\mathbb{E}[f(X_1, X_2)] = \int_0^1 \int_0^1 F(z_1, z_2)\, dz_1\, dz_2. \qquad (2)$$

and find an expression for the function F.

Answer: We have

$$\mathbb{E}[f(X_1, X_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\boldsymbol{\mu} + R^T \boldsymbol{y}) \frac{\phi(y_1)\phi(y_2)}{R_{1,1} R_{2,2}}\, dy_1\, dy_2.$$

Set $z_i = \Phi(y_i)$, or in other words $y_i = \Phi^{-1}(z_i)$, for $i = 1, 2$. Then $dz_i = \phi(y_i) dy_i$ and hence

$$\mathbb{E}[f(X_1, X_2)] = \int_0^1 \int_0^1 \frac{f(\boldsymbol{\mu} + R^T(\Phi^{-1}(z_1), \Phi^{-1}(z_2))^T)}{R_{1,1} R_{2,2}}\, dz_1\, dz_2.$$ |det(R^T)|

Thus

$$F(z_1, z_2) = \frac{f(\boldsymbol{\mu} + R^T(\Phi^{-1}(z_1), \Phi^{-1}(z_2))^T)}{\cancel{R_{1,1} R_{2,2}}}.$$ |det(R^T)|

c) Consider now the Monte Carlo method.

A) Briefly describe how you would use the Monte Carlo method to compute a numerical approximation to the integral (2).

Answer: Generate N samples $\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_N$ in the square $[0,1]^2$ where $\boldsymbol{z}_n = (z_{n,1}, z_{n,2})$ and then compute

$$\frac{1}{N}\sum_{n=1}^{N} F(\boldsymbol{z}_n)$$

to approximate the integral (2).

B) What can you say about the accuracy of the Monte Carlo method?

Answer: The expectation value of the approximation is $\int_{[0,1]^2} F(Z)\,dZ$, and the standard deviation of $\frac{1}{N}\sum_{n=1}^{N} F(\boldsymbol{z}_n)$ is $\frac{\sigma(F)}{\sqrt{N}}$.

C) How could you improve the accuracy of the Monte Carlo method?

Answer: One could use a control variate G, whose integral can be computed exactly and then approximate the integral over $F - G$ using Monte Carlo. Another option is to use antithetic variables. The third option is to replace Monte Carlo sampling with Quasi-Monte Carlo sampling.

ANSWER EITHER QUESTION 3 OR QUESTION 4, NOT BOTH

4. The value $V(S,t)$ of an exotic option on an underlying asset with price S at time t satisfies the Black–Scholes PDE

$$\frac{\partial V}{\partial t} + rS\,\frac{\partial V}{\partial S} + \tfrac{1}{2}\sigma^2 S^2\,\frac{\partial^2 V}{\partial S^2} = rV \tag{3}$$

for $0 < S < \infty$ and $0 < t < T$. We assume that the prevailing risk-free interest rate r and the volatility σ are known positive constants.

Consider using (3) to price an exotic option with barrier B, strike price $0 < X < B$ and expiry T. If at any time $t \in [0,T]$ the asset prices S reaches B, the option is immediately exercised with a pay-out of $B - X$. Otherwise the option can only be exercised at expiry $t = T$, where the owner has the right to **buy** the asset for a price X.

i) Carefully explain what the "initial" conditions are for this problem.

Answer: At expiry time $t = T$ we have

$$V(S,T) = \begin{cases} 0 & \text{if } 0 \le S < X, \\ S - X & \text{if } X \le S < B, \\ B - X & \text{if } B \le S. \end{cases}$$

ii) Carefully explain what the boundary conditions are for this problem.

Answer: We have $V(0,t) = 0$ since if the asset price drops to 0 at some time t it will stay at 0. The option is then never exercised and is therefore worthless. For the case where $S \to \infty$ we only need to consider $S = B$, in which case the option is exercised and the price of the option is $B - X$.

iii) Set up a grid $(S_j, t_\ell) = (j\Delta S, \ell\Delta t)$ for $j = 0, 1, 2, \ldots, n+1$ and $\ell = 0, 1, 2, \ldots, m$, where

$$\Delta S = \frac{B}{n+1} \quad \text{and} \quad \Delta t = \frac{T}{m}.$$

Let $v_j^\ell \approx V(S_j, t_\ell)$ denote the approximate solution.

a) At the time point t_ℓ and asset value S_j give

A) central difference approximations of $O(\Delta S^2)$ for the partial derivatives

$$\frac{\partial V}{\partial S} \quad \text{and} \quad \frac{\partial^2 V}{\partial S^2};$$

Answer:

$$\frac{\partial V}{\partial S} \approx \frac{v_{j+1}^{\ell} - v_{j-1}^{\ell}}{2\Delta S}$$
$$\frac{\partial^2 V}{\partial S^2} \approx \frac{v_{j+1}^{\ell} - 2v_{j}^{\ell} + v_{j-1}^{\ell}}{\Delta S^2}$$

B) a forward difference approximation of $O(\Delta t)$ for the partial derivative

$$\frac{\partial V}{\partial t}.$$

Answer:

$$\frac{\partial V}{\partial t} \approx \frac{v_{j}^{\ell+1} - v_{j}^{\ell}}{\Delta t}$$

b) We want to use an **implicit** finite difference method for computing an approximate solution v_j^{ℓ}.

A) Show that the finite difference equation has the form

$$\alpha_j v_{j-1}^{\ell} + \beta_j v_j^{\ell} + \gamma_j v_{j+1}^{\ell} = v_j^{\ell+1}, \tag{4}$$

and find the coefficients α_j, β_j and γ_j.

Answer: We have

$$\frac{v_j^{\ell+1} - v_j^{\ell}}{\Delta t} + rS_j \frac{v_{j+1}^{\ell} - v_{j-1}^{\ell}}{2\Delta S} + \frac{1}{2}\sigma^2 S_j^2 \frac{v_{j+1}^{\ell} - 2v_j^{\ell} + v_{j-1}^{\ell}}{\Delta S^2} = rv_j^{\ell}.$$

By rearranging the terms we get

$$\begin{aligned} v_j^{\ell+1} &= \left(\frac{rS_j\Delta t}{2\Delta S} - \frac{\Delta t\sigma^2 S_j}{2\Delta S^2}\right) v_{j-1}^{\ell} + \left(1 + r\Delta t + \frac{\Delta t\sigma^2 S_j^2}{\Delta S^2}\right) v_j^{\ell} \\ &\quad + \left(-\frac{\Delta t r S_j}{2\Delta S} - \frac{\Delta t\sigma^2 S_j^2}{2\Delta S^2}\right) v_{j+1}^{\ell}. \end{aligned}$$

Hence we have

$$\alpha_j = \frac{rS_j\Delta t}{2\Delta S} - \frac{\Delta t\sigma^2 S_j^2}{2\Delta S^2}, \beta_j = 1 + r\Delta t + \frac{\Delta t\sigma^2 S_j^2}{\Delta S^2}, \gamma_j = -\frac{\Delta t r S_j}{2\Delta S} - \frac{\Delta t\sigma^2 S_j^2}{2\Delta S^2}.$$

Please see over ...

B) What are the discrete "initial" conditions?

Answer:

$$v_j^m = \begin{cases} 0 & \text{if } 0 \le S_j < X, \\ S_j - X & \text{if } X \le S_j. \end{cases}$$

C) What are the discrete boundary conditions?

Answer:

$$v_0^\ell = 0, \quad v_{n+1}^\ell = B - X.$$

iv) Equation (4) has the form

$$A\boldsymbol{v}^\ell + \boldsymbol{b} = \boldsymbol{v}^{\ell+1}, \quad \text{where } \boldsymbol{v}^\ell = (v_1^\ell, \dots, v_n^\ell)^T \text{ and } \boldsymbol{b} = (0, \dots, 0, \gamma_n(B-X))^T.$$

a) Define the entries $a_{i,j}$ of the matrix $A = (a_{i,j})$. What structure does the matrix A have?

Answer: The matrix A is an $n \times n$ matrix with

$$\begin{aligned} a_{j,j-1} &= \alpha_j, \quad j = 2, 3, \dots, n; \\ a_{j,j} &= \beta_j, \quad j = 1, 2, \dots, n; \\ a_{j,j+1} &= \gamma_j, \quad j = 1, 2, \dots, n-1; \end{aligned}$$

The remaining elements are 0. Hence the matrix is tri-diagonal.

b) Is the matrix A sparse? Calculate its sparsity.

Answer: Yes, the matrix is sparse and its sparsity is

$$\frac{n + 2(n-1)}{n^2} = \frac{3}{n} - \frac{2}{n^2}.$$

c) How can the structure of the matrix A be exploited to make time-stepping as efficient as possible?

Answer: In each timestep we need to solve the linear system $A\boldsymbol{v}^\ell = \boldsymbol{v}^{\ell+1} - \boldsymbol{b}$, where $A, \boldsymbol{v}^{\ell+1}, \boldsymbol{b}$ are known and $\boldsymbol{v}^\ell$ is unknown. The matrix A is tri-diagonal and hence Gaussian elimination can be simplified using the Thomas algorithm. This algorithm requires only $\mathcal{O}(n)$ operations to solve.

d) Provide an estimation of the number of flops needed to compute the vector $\boldsymbol{v}^0$.

Answer: Solving the linear system $A\boldsymbol{v}^\ell = \boldsymbol{v}^{\ell+1} - \boldsymbol{b}$ requires $\mathcal{O}(n)$ flops. This system needs to be solved for each value of $\ell = m-1, m-2, \ldots, 0$. Hence we require $\mathcal{O}(nm)$ flops.