Chapter 0

Mathematical preliminaries

0.1 Even and odd functions

Definition 0.1 *The function $f : \mathbb{R} \to \mathbb{R}$ is an even function if and only if*

$$f(-x) = f(x), \quad \forall x \in \mathbb{R}. \tag{1}$$

The graph of any even function is symmetric with respect to the y-axis.

Lemma 0.1 *Let $f(x)$ be an integrable even function. Then,*

$$\int_{-a}^{0} f(x)dx = \int_{0}^{a} f(x)dx, \quad \forall a \in \mathbb{R}, \tag{2}$$

and therefore

$$\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx, \quad \forall a \in \mathbb{R}. \tag{3}$$

Moreover, if $\int_0^\infty f(x)dx$ exists, then

$$\int_{-\infty}^{0} f(x)dx = \int_{0}^{\infty} f(x)dx, \tag{4}$$

and

$$\int_{-\infty}^{\infty} f(x)dx = 2\int_{0}^{\infty} f(x)dx. \tag{5}$$

Definition 0.2 *The function $f : \mathbb{R} \to \mathbb{R}$ is an odd function if and only if*

$$f(-x) = -f(x), \quad \forall x \in \mathbb{R}. \tag{6}$$

If we let $x = 0$ in (6), we find that $f(0) = 0$ for any odd function $f(x)$. Also, the graph of any odd function is symmetric with respect to the point $(0, 0)$.

Lemma 0.2 *Let $f(x)$ be an integrable odd function. Then,*

$$\int_{-a}^{a} f(x)dx = 0, \quad \forall a \in \mathbb{R}. \tag{7}$$

Moreover, if $\int_0^\infty f(x)dx$ exists, then

$$\int_{-\infty}^{\infty} f(x)dx = 0. \tag{8}$$

0.2 Useful sums with interesting proofs

The following sums occur frequently when estimating operation counts of numerical algorithms:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}; \tag{9}$$

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}; \tag{10}$$

$$\sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2. \tag{11}$$

0.3 Sequences satisfying linear recursions

Definition 0.3 *A sequence $(x_n)_{n\geq 0}$ satisfies a linear recursion of order k if and only if there exist constants a_i, $i = 0 : k$ with $a_k \neq 0$, such that*

$$\sum_{i=0}^{k} a_i x_{n+i} = 0, \quad \forall n \geq 0. \tag{12}$$

The recursion (12) is called a linear recursion because of the following linearity properties:
(i) If the sequence $(x_n)_{n\geq 0}$ satisfies the linear recursion (12), then the sequence $(z_n)_{n\geq 0}$ given by

$$z_n = Cx_n, \quad \forall n \geq 0, \tag{13}$$

where C is an arbitrary constant, also satisfies the linear recursion (12).
(ii) If the sequences $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ satisfies the linear recursion (12), then the sequence $(z_n)_{n\geq 0}$ given by

$$z_n = x_n + y_n, \quad \forall n \geq 0, \tag{14}$$

also satisfies the linear recursion (12).

Definition 0.4 *The characteristic polynomial $P(z)$ corresponding to the linear recursion $\sum_{i=0}^{k} a_i x_{n+i} = 0$, for all $n \geq 0$, is defined as*

$$P(z) = \sum_{i=0}^{k} a_i z^i. \tag{15}$$

$P(z)$ is a polynomial of degree k, i.e., $\deg(P(z)) = k$. If $P(z)$ has p different roots, λ_j, $j = 1 : p$, with $p \leq k$, and if $m(\lambda_j)$ denotes the multiplicity of the root λ_j, then $\sum_{j=1}^{p} m(\lambda_j) = k$ where λ_j can be a complex number.

Theorem 0.1 *Let $(x_n)_{x\geq 0}$ be a sequence satisfying the linear recursion*

$$\sum_{i=0}^{k} a_i x_{n+i} = 0, \quad \forall n \geq 0, \tag{16}$$

with $a_k \neq 0$, and let $P(z) = \sum_{i=0}^{k-1} a_i z^i$ be the characteristic polynomial associated with recursion (16). Let λ_j, $j = 1 : p$, where $p \leq k$, be the roots of $P(z)$, and let $m(\lambda_j)$ be the multiplicity of λ_j. The general form of the sequence $(x_n)_{n \geq 0}$ satisfying the linear recursion (16) is

$$x_n = \sum_{j=1}^{p} \left(\sum_{i=0}^{m(\lambda_j)-1} C_{i,j} n^i \right) \lambda_j^n, \quad \forall n \geq 0, \tag{17}$$

where $C_{i,j}$ are constant numbers.

0.4 The "Big O" and "little o" notations

Definition 0.5 *Let $f, g : \mathbb{R} \to \mathbb{R}$. We write that $f(x) = O(g(x)$, as $x \to \infty$, if and only if there exist constants $C > 0$ and $M > 0$ such that $\left| \frac{f(x)}{g(x)} \leq C \right|$, for any $x \geq M$. This can be written equivalently as*

$$f(x) = O(g(x)), \quad \textit{as} \quad x \to \infty, \quad \textit{iff} \quad \limsup_{x \to \infty} \left| \frac{f(x)}{g(x)} \right| < \infty. \tag{18}$$

The "little o" notation refers to functions whose ratios tend to 0 at certain points, and can be defined for $x \to \infty$, $x \to a$, and $x \to -\infty$ as follows:

Definition 0.6 *Let $f, g : \mathbb{R} \to \mathbb{R}$. Then*

$$f(x) = o(g(x)), \quad \textit{as} \quad x \to \infty, \quad \textit{iff} \quad \lim_{x \to \infty} \left| \frac{f(x)}{g(x)} \right| = 0; \tag{19}$$

$$f(x) = o(g(x)), \quad \textit{as} \quad x \to -\infty, \quad \textit{iff} \quad \lim_{x \to -\infty} \left| \frac{f(x)}{g(x)} \right| = 0; \tag{20}$$

$$f(x) = o(g(x)), \quad \textit{as} \quad x \to a, \quad \textit{iff} \quad \lim_{x \to a} \left| \frac{f(x)}{g(x)} \right| = 0. \tag{21}$$

0.5 Exercises

0.5.1 Question 1

Let $f : \mathbb{R} \to \mathbb{R}$ be an odd function.

Part (i)

Show that $xf(x)$ is an even function and $x^2 f(x)$ is an odd function.

Answer Let $g(x) = xf(x)$, using definition (6), we have

$$\begin{aligned} g(-x) &= -xf(-x) \\ &= -x(-f(x)) \\ &= xf(x) \\ &= g(x). \end{aligned}$$

Therefore, $xf(x)$ is an even function.

Let $g(x) = x^2 f(x)$, using definition (6), we have

$$\begin{aligned} g(-x) &= (-x)^2 f(-x) \\ &= x^2(-f(x)) \\ &= -x^2 f(x) \\ &= -g(x) \end{aligned}$$

Therefore, $x^2 f(x)$ is an odd function.

Part (ii)

Show that the function $g_1 : \mathbb{R} \to \mathbb{R}$ given by $g_1(x) = f(x^2)$ is an even function and that the function $g_2 : \mathbb{R} \to \mathbb{R}$ given by $g_2(x) = f(x^3)$ is an odd function.

Answer

$$\begin{aligned} g_1(-x) &= f((-x)^2) \\ &= f(x^2) \\ &= g_1(x) \end{aligned}$$

Therefore g_1 is an even function.
Let $y = x^3$:

$$\begin{aligned} g_2(-x) &= f((-x)^3) \\ &= f(-x^3) \\ &= f(-y) \\ &= -f(y) \\ &= -f(x^3) \\ &= -g_2(x) \end{aligned}$$

Therefore g_2 is an odd function.

Part (iii)

Let i be even, j be odd, and $y = x^j$:

$$\begin{aligned} h(-x) &= (-x)^i f((-x)^j) \\ &= x^i f(-x^j) \\ &= x^i f(-y) \\ &= -x^i f(y) \\ &= -x^i f(x^j) \\ &= -h(x) \end{aligned}$$

Let i be odd, j be even:

$$\begin{aligned} h(-x) &= (-x)^i f((-x)^j) \\ &= -x^i f(x^j) \\ &= -h(x) \end{aligned}$$

When $i + j$ is odd, $h(x)$ is an odd function.

0.5.2 Question 2

Let $S(n,2) = \sum_{k=1}^{n} k^2$ and $S(n,3) = \sum_{k=1}^{n} k^3$.

Part (i)

Let $T(n,2,x) = \sum_{k=1}^{n} k^2 x^k$. Use formulas,

$$T(n,2,x) = x\frac{d}{dx}(T(n,1,x)),$$

and

$$T(n,1,x) = \frac{x-(n+1)x^{n+1}+nx^{n+2}}{(1-x)^2},$$

to show that

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

Answer Using quotient rule,

$$\begin{aligned}
T(n,2,x) &= x\frac{d}{dx}(T(n,1,x)) \\
&= x\frac{d}{dx}\left(\frac{x-(n+1)x^{n+1}+nx^{n+2}}{(1-x)^2}\right) \\
&= x\left(\frac{\begin{array}{l}\frac{d}{dx}(x-(n+1)x^{n+1}+nx^{n+2})(1-x)^2 \\ \qquad -(x-(n+1)x^{n+1}+nx^{n+2})\frac{d}{dx}(1-x)^2\end{array}}{(1-x)^4}\right) \\
&= x\left(\frac{\begin{array}{l}(1-(n+1)^2x^n+(n^2+2n)x^{n+1})(1-x)^2 \\ \qquad -(x-(n+1)x^{n+1}+nx^{n+2})(-2(1-x))\end{array}}{(1-x)^4}\right) \\
&= x\left(\frac{\begin{array}{l}(1-(n+1)^2x^n+(n^2+2n)x^{n+1})(1-x)^2 \\ \qquad\qquad +2(x-(n+1)x^{n+1}+nx^{n+2})(1-x)\end{array}}{(1-x)^4}\right) \\
&= x\left(\frac{(1-(n+1)^2x^n+(n^2+2n)x^{n+1})(1-x)+2(x-(n+1)x^{n+1}+nx^{n+2})}{(1-x)^3}\right) \\
&= x\left(\frac{\begin{array}{l}(1-(n+1)^2x^n+(n^2+2n)x^{n+1}) \\ \qquad -(x-(n+1)^2x^{n+1}+(n^2+2n)x^{n+2})+\cdots\end{array}}{(1-x)^3}\right) \\
&= x\left(\frac{1-x-(n+1)^2x^n+(2n^2+4n+1)x^{n+1}-(n^2+2n)x^{n+2}+\cdots}{(1-x)^3}\right) \\
&= x\left(\frac{1+x-(n+1)^2x^n+(2n^2+2n-1)x^{n+1}-n^2x^{n+2}}{(1-x)^3}\right) \\
&= \frac{x+x^2-(n+1)^2x^{n+1}+(2n^2+2n-1)x^{n+2}-n^2x^{n+3}}{(1-x)^3}
\end{aligned}$$

Part (ii)

Note that $S(n,2) = T(n,2,1)$. Use l'Hôpitals's rule to evaluate $T(n,2,1)$, and conclude that $S(n,2) = \frac{n(n+1)(2n+1)}{6}$.

Answer

$$
\begin{aligned}
\lim_{x\to 1} T(n,2,x) &= \lim_{x\to 1} \frac{x + x^2 - (n+1)^2 x^{n+1} + (2n^2+2n-1)x^{n+2} - n^2 x^{n+3}}{(1-x)^3} \\
&= \lim_{x\to 1} \frac{\begin{array}{r}(n-1)n(n+1)^3 x^{n-2} - n(n+1)(n+2)(2n^2+2n-1)x^{n-1} \\ + (n+1)(n+2)(n+3)n^2 x^n\end{array}}{6} \\
&= \frac{\begin{array}{r}(n-1)n(n+1)^3 - n(n+1)(n+2)(2n^2+2n-1) \\ + (n+1)(n+2)(n+3)n^2\end{array}}{6} \\
&= \frac{n(n+1)(2n+1)}{6} \\
S(n,2) &= \frac{n(n+1)(2n+1)}{6}
\end{aligned}
$$

Part (iii)

Compute $T(n,3,x) = \sum_{k=1}^{n} k^3 x^k$ using the formula

$$
T(n,3,x) = x\frac{d}{dx}(T(n,2,x)).
$$

Answer

$$
\begin{aligned}
T(n,3,x) &= x\frac{d}{dx}(T(n,2,x)) \\
&= x\frac{d}{dx}\frac{x + x^2 - (n+1)^2 x^{n+1} + (2n^2+2n-1)x^{n+2} - n^2 x^{n+3}}{(1-x)^3} \\
&= \frac{\begin{array}{r}x((n^3(x-1)^3 - 3n^2(x-1)^2 + 3n(x^2-1) \\ - x(x+4) - 1)x^n + (x+4)x + 1)\end{array}}{(1-x)^4}
\end{aligned}
$$

Part (iv)

Note that $S(n,3) = T(n,3,1)$. Use l'Hôpital's rule to evaluate $T(n,3,1)$, and conclude that $S(n,3) = \left(\frac{n(n+1)}{2}\right)^2$.

$$
\lim_{x\to 1} T(n,3,x) = \lim_{x\to 1} \frac{\begin{array}{r}x((n^3(x-1)^3 - 3n^2(x-1)^2 + 3n(x^2-1) \\ - x(x+4) - 1)x^n + (x+4)x + 1)\end{array}}{(1-x)^4}
$$

0.5.3 Question 3

Compute $S(n,4) = \sum_{k=1}^{n} k^4$ using the recursion formula for $i = 4$, the fact that $S(n,0) = n$, and formulas for $S(n,1)$, $S(n,2)$, and $S(n,3)$.

Answer

$$\begin{aligned} S(n,4) &= \frac{1}{5}\left((n+1)^5 - 1 - \sum_{j=0}^{3}\binom{5}{j} xS(n,j)\right) \\ &= \frac{1}{5}\big((n+1)^5 - 1 - S(n,0) - 5S(n,1) - 10S(n,2) - 10S(n,3)\big) \\ &= \frac{1}{5}\left((n+1)^5 - 1 - n - 5\frac{n(n+1)}{2} - 10\frac{n(n+1)(2n+1)}{6} - 10\left(\frac{n(n+1)}{2}\right)^2\right) \\ &= \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30} \end{aligned}$$

0.5.4 Question 4

It is easy to see that the sequence $(x_n)_{n\geq 1}$ given by $x_n = \sum_{k=1}^{n} k^2$ satisfies the recursion

$$x_{n+1} = x_n + (n+1)^2, \quad \forall n \geq 1, \tag{22}$$

with $x_1 = 1$.

Part (i)

By substituting $n+1$ for n in (22), obtain that

$$x_{n+2} = x_{n+1} + (n+2)^2. \tag{23}$$

Subtract (22) from (23) to find that

$$x_{n+2} = 2x_{n+1} - x_n + 2n + 3, \quad \forall n \geq 1, \tag{24}$$

with $x_1 = 1$ and $x_2 = 5$.

Answer

$$\begin{aligned} x_{(n+1)+1} &= x_{n+1}((n+1)+1)^2 \\ x_{n+2} &= x_{n+1}(n+2)^2 \end{aligned}$$

Subtract (22) from (23):

$$\begin{aligned} x_{n+2} - x_{n+1} &= x_{n+1} + (n+2)^2 - x_n - (n+1)^2 \\ x_{n+2} &= 2x_{n+1} - x_n + n^2 + 4n + 4 - n^2 - 2n - 1 \\ &= 2x_{n+1} - x_n + 2n + 3 \end{aligned}$$

Part (ii)

Similarly, substitute $n+1$ for n in (24) and obtain that

$$x_{n+3} = 2x_{n+2} - x_{n+1} + 2(n+1) + 3. \tag{25}$$

Subtract (24) from (25) to find that

$$x_{n+3} = 3x_{n+2} - 3x_{n+1} + x_n + 2, \quad \forall n \geq 1, \tag{26}$$

with $x_1 = 1$, $x_2 = 5$, and $x_3 = 14$.

Answer

$$\begin{aligned} x_{(n+1)+2} &= 2x_{(n+1)+1} - x_{n+1} + 2(n+1) + 3 \\ x_{n+3} &= 2x_{n+2} - x_{n+1} + 2(n+1) + 3 \end{aligned}$$

Subtract (24) from (25)

$$\begin{aligned} x_{n+3} - x_{n+2} &= 2x_{n+2} - x_{n+1} + 2(n+1) + 3 - 2x_{n+1} + x_n - 2n - 3 \\ x_{n+3} &= 3x_{n+2} - 3x_{n+1} + x_n + 2 \end{aligned}$$

Part (iii)

Use a similar method to prove that the sequence $(x_n)_{n\geq 0}$ satisfies the linear recursion

$$x_{n+4} - 4x_{n+3} + 6x_{n+2} - 4x_{n+1} + x_n = 0, \quad \forall n \geq 1. \tag{27}$$

The characteristic polynomial associated to the recursion (27) is

$$P(z) = z^4 - 4z^3 + 6z^2 - 4z + 1 = (z-1)^4.$$

Use the fact that $x_1 = 1$, $x_2 = 5$, $x_3 = 14$, and $x_4 = 30$ to show that

$$x_n = \frac{n(n+1)(2n+1)}{6}, \quad \forall n \geq 1,$$

and conclude that

$$S(n,2) = \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}, \quad \forall n \geq 1.$$

Answer Substitute $n+1$ for n in (26) to obtain

$$x_{n+4} = 3x_{n+3} - 3x_{n+2} + x_{n+1} + 2. \tag{28}$$

Subtract (26) from (28) to obtain that

$$\begin{aligned} x_{n+4} - x_{n+3} &= 3x_{n+3} - 3x_{n+2} + x_{n+1} + 2 \\ &\quad - (3x_{n+2} - 3x_{n+1} + x_n + 2) \\ x_{n+4} - 4x_{n+3} + 6x_{n+2} - 4x_{n+1} - x_n &= 0 \end{aligned}$$

The characteristic polynomial has root $\lambda = 1$ with multiplicity 4. The linear recursion can be expressed as

$$\begin{aligned} x_n &= \sum_{j=1}^{p} \left(\sum_{i=0}^{3} C_{i,j} n^i \right) \lambda_j^n \\ &= \sum_{j=1}^{p} \left(C_{0,j} + C_{1,j} n + C_{2,j} n^2 + C_{3,j} n^3 \right) \lambda_j^n \\ &= C_1 + C_2 n + C_3 n^2 + C_4 n^3 \end{aligned}$$

Since $x_1 = 1$, $x_2 = 5$, $x_3 = 14$, and $x_4 = 30$, C_1, C_2, C_3, and C_4 must solve the linear system

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 14 \\ 30 \end{pmatrix}.$$

We obtain that $C_1 = 0$, $C_2 = \frac{1}{6}$, $C_3 = \frac{1}{2}$, and $C_4 = \frac{1}{3}$ and therefore

$$x_n = \frac{n}{6} + \frac{n^2}{2} + \frac{n^3}{3} = \frac{n(n+1)(n+2)}{6}$$

0.5.5 Question 5

Find the general form of the sequence $(x_n)_{n \geq 0}$ satisfying the linear recursion

$$x_{n+3} = 2x_{n+1} + x_n, \quad \forall n \geq 0,$$

with $x_0 = 1$, $x_1 = -1$, and $x_2 = 1$.

Answer Rewrite the recursion in the form (12) as

$$x_{n+3} - 2x_{n+1} - x_n = 0, \quad \forall n \geq 0.$$

The characteristic polynomial associated to the linear recursion is

$$\begin{aligned} P(z) &= z^3 - 2z - 1 \\ &= (z+1)(z^2 - z - 1) \end{aligned}$$

and the roots of $P(z)$ are

$$\lambda_1 = -1, \quad \lambda_2 = \frac{1+\sqrt{5}}{2}, \quad \lambda_3 = \frac{1-\sqrt{5}}{2}.$$

From Theorem 0.1, we find that

$$x_n = C_1 \lambda_1^n + C_2 \lambda_2^n + C_3 \lambda_3^n, \quad \forall n \geq 0.$$

Given $x_0 = 1$, $x_1 = -1$, and $x_2 = 1$, we obtain the linear system

$$\begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

By solving the linear system, we find that $C_1 = 1$, $C_2 = 0$, and $C_3 = 0$. The general formula for is

$$x_n = (-1)^n, \quad \forall n \geq 0.$$

0.5.6 Question 6

The sequence $(x_n)_{n\geq 0}$ satisfies the recursion

$$x_{n+1} = 3x_n + 2, \quad \forall n \geq 0,$$

with $x_0 = 1$.

Part (i)

Show that the sequence $(x_n)_{n\geq 0}$ satisfies the linear recursion

$$x_{n+2} = 4x_{n+1} - 3x_n, \quad \forall n \geq 0,$$

with $x_0 = 1$ and $x_1 = 5$.

Answer Substitute $n+1$ for n to obtain

$$x_{n+2} = 3x_{n+1} + 2$$

Subtract the original recursion to get

$$\begin{aligned} x_{n+2} - x_{n+1} &= 3x_{n+1} + 2 - 3x_n - 2 \\ x_{n+2} &= 4x_{n+1} - 3x_n \end{aligned}$$

Part (ii)

Find the general formula for x_n, $n \geq 0$.

Answer The characteristic polynomial has the form

$$P(z) = z^2 - 4z + 3 = (z-1)(z-3)$$

which has roots $\lambda_1 = 1$ and $\lambda_2 = 3$. We obtain the linear system

$$\begin{cases} C_1 + C_2 = 1; \\ C_1\lambda_1 + C_2\lambda_2 = 5. \end{cases}$$

The solution to the linear system is $C_1 = -1$ and $C_2 = 2$. Therefore, the general form is

$$x_n = 2(3)^n - 1$$

0.5.7 Question 7

The sequence $(x_n)_{n\geq 0}$ satisfies the recursion

$$x_{n+1} = 3x_n + n + 2, \quad \forall n \geq 0,$$

with $x_0 = 1$.

Part (i)

Show that the sequence $(x_n)_{n\geq 0}$ satisfies the linear recursion

$$x_{n+3} = 5x_{n+2} - 7x_{n+1} + 3x_n, \quad \forall n \geq 0,$$

with $x_0 = 1$, $x_1 = 5$, and $x_2 = 18$.

Answer Substitute $n+1$ for n, we obtain

$$x_{n+2} = 3x_{n+1} + n + 3$$

Subtract

$$x_{n+2} = 4x_{n+1} - 3x_n + 1$$

Substitute $n+1$ for n, we obtain

$$x_{n+3} = 4x_{n+2} - 3x_{n+1} + 1$$

Subtract

$$x_{n+3} = 5x_{n+2} - 7x_{n+1} + 3x_n$$

Part (ii)

Find the general formula for x_n, $n \geq 0$.

Answer The characteristic polynomial is given by

$$P(z) = z^3 - 5z^2 + 7z - 3 = (z-1)^2(z-3),$$

with roots $\lambda_1 = 1$ and $\lambda_2 = 3$. The general form is

$$\begin{aligned} x_n &= \sum_{j=1}^{2}\left(\sum_{i=0}^{m(\lambda_j)-1} C_{i,j}n^i\right)\lambda_j^n \\ &= \lambda_1^n \sum_{i=0}^{1} C_{i,1}n^i + \lambda_2^n C_2 \\ &= \lambda_1^n C_{0,1} + \lambda_1^n C_{1,1}n + \lambda_2^n C_2 \\ &= C_1 + C_2 n + C_3 3^n \end{aligned}$$

Since $x_0 = 1$, $x_1 = 5$, and $x_2 = 18$, we find $C_1 = -\frac{1}{2}$, $C_2 = -\frac{5}{4}$, and $C_3 = \frac{9}{4}$. We conclude that

$$x_n = \frac{3^{n+2} - 2n - 5}{4}$$

0.5.8 Question 8

Let $P(z) = \sum_{i=0}^{k} a_i z^i$ be the characteristic polynomial corresponding to the linear recursion

$$\sum_{i=0}^{k} a_i x_{n+i} = 0, \quad \forall n \geq 0.$$

Assume that λ is a root of multiplicity 2 of $P(z)$. Show that the sequence $(y_n)_{n\geq 0}$ given by

$$y_n = Cn\lambda^n, \quad n \geq 0,$$

where C is an arbitrary constant, satisfies the recursion.

Answer

$$\begin{aligned}
\sum_{i=0}^{k} a_i y_{n+i} &= \sum_{i=0}^{k} a_i C(n+i)\lambda^{n+i} \\
&= Cn\sum_{i=0}^{k} a_i\lambda^{n+i} + C\sum_{i=0}^{k} a_i i\lambda^{n+i} \\
&= Cn\lambda^n\sum_{i=0}^{k} a_i\lambda^{i} + C\lambda^{n+1}\sum_{i=0}^{k} i a_i\lambda^{i-1} \\
&= Cn\lambda^n P(\lambda) + C\lambda^{n+1}P'(\lambda) \\
&= 0.
\end{aligned}$$

0.5.9 Question 9

Let $n > 0$. Show that

$$\begin{aligned}
O(x^n) + O(x^n) = O(x^n), &\quad \text{as} \quad x \to 0; \\
o(x^n) + o(x^n) = o(x^n), &\quad \text{as} \quad x \to 0.
\end{aligned}$$

Answer Let $f(x) = O(x^n)$ and $g(x) = O(x^n)$, then

$$\limsup_{x\to 0}\left|\frac{f(x)}{x^n}\right| < \infty \quad \text{and} \quad \limsup_{x\to 0}\left|\frac{g(x)}{x^n}\right| < \infty.$$

We see that

$$\limsup_{x\to 0}\left|\frac{f(x)+g(x)}{x^n}\right| \le \limsup_{x\to 0}\left|\frac{f(x)}{x^n}\right| + \limsup_{x\to 0}\left|\frac{g(x)}{x^n}\right| < \infty,$$

and therefore $O(x^n) + O(x^n) = O(x^n)$.
Let $f(x) = o(x^n)$ and $g(x) = o(x^n)$, then

$$\lim_{x\to 0}\left|\frac{f(x)}{x^n}\right| = 0 \quad \text{and} \quad \lim_{x\to 0}\left|\frac{g(x)}{x^n}\right| = 0.$$

We see that

$$\lim_{x\to 0}\left|\frac{f(x)+g(x)}{x^n}\right| \le \lim_{x\to 0}\left|\frac{f(x)}{x^n}\right| + \lim_{x\to 0}\left|\frac{g(x)}{x^n}\right| = 0,$$

and therefore $o(x^n) + o(x^n) = o(x^n)$.

0.5.10 Question 10

Prove that

$$\begin{aligned}
\sum_{k=1}^{n} k^2 = O(n^3), &\quad \text{as} \quad n \to \infty; \\
\sum_{k=1}^{n} k^2 = \frac{n^3}{3} + O(n^2), &\quad \text{as} \quad n \to \infty,
\end{aligned}$$

i.e., show that

$$\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} k^2}{n^3} < \infty$$

and that

$$\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} k^2 - \frac{n^3}{3}}{n^2} < \infty.$$

Similarly, prove that

$$\sum_{k=1}^{n} k^3 = O(n^4), \quad \text{as} \quad n \to \infty;$$

$$\sum_{k=1}^{n} k^3 = \frac{n^4}{4} + O(n^3), \quad \text{as} \quad n \to \infty,$$

Answer Using (10)

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} k^2}{n^3} &= \limsup_{n\to\infty} \frac{\frac{n(n+1)(2n+1)}{6}}{n^3} \\
&= \limsup_{n\to\infty} \frac{2n^3 + 3n^2 + n}{6n^3} \\
&= \frac{1}{3} < \infty \\
\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} k^2 - \frac{n^3}{3}}{n^2} &= \limsup_{n\to\infty} \frac{\frac{2n^3+3n^2+n}{6} - \frac{n^3}{3}}{n^2} \\
&= \limsup_{n\to\infty} \frac{3n^2 + n}{6n^2} \\
&= \frac{1}{2} \\
&< \infty
\end{aligned}$$

Using (11)

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} k^3}{n^4} &= \limsup_{n\to\infty} \frac{\left(\frac{n(n+1)}{2}\right)^2}{n^4} \\
&= \limsup_{n\to\infty} \frac{n^2(n^2 + 2n + 1)}{4n^4} \\
&= \frac{1}{4} \\
&< \infty \\
\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} k^3 - \frac{n^4}{4}}{n^3} &= \limsup_{n\to\infty} \frac{\frac{(n^4+2n^3+n^2)}{4} - \frac{n^4}{4}}{n^3} \\
&= \limsup_{n\to\infty} \frac{2n^3 + n^2}{4n^3} \\
&= \frac{1}{2} \\
&< \infty
\end{aligned}$$

Chapter 1

Calculus review. Plain vanilla options.

1.1 Brief review of differentiation

The function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at the point $x \in \mathbb{R}$ if the limit

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

exists, in which case the derivative $f'(x)$ is defined as

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}. \tag{1.1}$$

Theorem 1.1 (Product Rule.) *The product $f(x)g(x)$ of two differentiable functions $f(x)$ and $g(x)$ is differentiable, and*

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x). \tag{1.2}$$

Theorem 1.2 (Quotient Rule.) *The quotient $\frac{f(x)}{g(x)}$ of two differentiable functions $f(x)$ and $g(x)$ is differentiable at every point x where the function $\frac{f(x)}{g(x)}$ is well defined, and*

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}. \tag{1.3}$$

Theorem 1.3 (Chain Rule.) *The composite function $(g \circ f)(x) = g(f(x))$ of two differentiable functions $f(x)$ and $g(x)$ is differentiable at every point x where $g(f(x))$ is well defined, and*

$$(g(f(x)))' = g'(f(x))f'(x). \tag{1.4}$$

The Chain Rule formula (1.4) can also be written as

$$\frac{dg}{dx} = \frac{dg}{du}\frac{du}{dx},$$

where $u = f(x)$ is a function of x and $g = g(u) = g(f(x))$.

Chain Rule is often used for power functions, exponential functions, and logarithmic function:

$$\frac{d}{dx}((f(x))^n) = n(f(x))^{n-1}f'(x); \tag{1.5}$$

$$\frac{d}{dx}(e^{f(x)}) = e^{f(x)}f'(x); \tag{1.6}$$

$$\frac{d}{dx}(\ln f(x)) = \frac{f'(x)}{f(x)}. \tag{1.7}$$

Lemma 1.1 *Let $f : [a, b] \to [c, d]$ be a differentiable function, and assume that $f(x)$ has an inverse function denoted by $f^{-1}(x)$ with $f^{-1} : [c, d] \to [a, b]$. The function $f^{-1}(x)$ is differentiable at every point $x \in [c, d]$ where $f'(f^{-1}(x)) \neq 0$ and*

$$(f^{-1}(x))' = \frac{1}{f'(f^{-1}(x))}. \tag{1.8}$$

1.2 Brief review of integration

Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function. Recall that $F(x)$ is the antiderivative of $f(x)$ iff $F'(x) = f$, i.e.,

$$F(x) = \int f(x)dx \quad \Longleftrightarrow \quad F'(x) = f(x).$$

Theorem 1.4 (Fundamental Theorem of Calculus.) *Let $f(x)$ be a continuous function on the interval $[a, b]$, and let $F(x)$ be the antiderivative of $f(x)$. Then*

$$\int_a^b f(x)dx = F(x)|_a^b = F(b) - F(a).$$

Integration by parts is the counterpart for integration of the product rule.

Theorem 1.5 (Integration by parts.) *Let $f(x)$ and $g(x)$ be continuous function. Then*

$$\int f(x)g(x)dx = F(x)g(x) - \int F(x)g'(x)dx, \tag{1.9}$$

where $F(x) = \int f(x)dx$ is the antiderivative of $f(x)$. For definite integrals,

$$\int_a^b f(x)g(x)dx = F(b)g(b) - F(a)g(a) - \int_a^b F(x)g'(x)dx. \tag{1.10}$$

Integration by substitution if the counterpart for integration of the chain rule.

Theorem 1.6 (Integration by substitution) *Let $f(x)$ be an integrable function. Assume that $g(u)$ is an invertible and continuously differentiable function. The substitution $x = g(u)$ changes the integration variable from x to u as follows:*

$$\int f(x)dx = \int f(g(u))g'(u)du. \tag{1.11}$$

For definite integrals,

$$\int_a^b f(x)dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(u))g'(u)du. \tag{1.12}$$

1.3 Differentiating definite integrals

If a definite integral has functions as limits of integration, e.g.,

$$\int_{a(t)}^{b(t)} f(x)dx,$$

or if the function to be integrated is a function of the integrating variable and of another variable, e.g.,

$$\int_a^b f(x,t)dx$$

then the result of the integration is a function (of the variable t in both cases above).

Lemma 1.2 *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then,*

$$\frac{d}{dt}\left(\int_{a(t)}^{b(t)} f(x)dx\right) = f(b(t))b'(t) - f(a(t))a'(t), \tag{1.13}$$

where $a(t)$ and $b(t)$ are differentiable functions.

Lemma 1.3 *Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a continuous function such that the partial derivative $\frac{\partial f}{\partial t}(x,t)$ exists and is continuous in both variables x and t.*

$$\frac{d}{dt}\left(\int_a^b f(x,t)dx\right) = \int_a^b \frac{\partial f}{\partial t}(x,t)dx. \tag{1.14}$$

Lemma 1.4 *Let $f(x,t)$ be a continuous function such that the partial derivative $\frac{\partial f}{\partial t}(x,t)$ exists and is continuous. Then,*

$$\frac{d}{dt}\left(\int_{a(t)}^{b(t)} f(x,t)dx\right) = \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(x,t) + f(b(t),t)b'(t) - f(a(t),t)a'(t).$$

Note that Lemma 1.2 and Lemma 1.3 are special cases of Lemma 1.4.

1.4 Limits

Definition 1.1 *Let $g : \mathbb{R} \to \mathbb{R}$. The limit of $g(x)$ as $x \to x_0$ exists and is finite and equal to l iff for any $\epsilon > 0$ there exists $\delta > 0$ such that $|g(x) - l| < \epsilon$ for all $x \in (x_0 - \delta, x_0 + \delta)$, i.e.,*

$$\lim_{x\to x_0} g(x) = l, \quad \textit{iff} \quad \forall \epsilon > 0 \exists \delta > 0 \quad \textit{such that} \quad |g(x) - l| < \epsilon, \ \forall |x - x_0| < \delta.$$

Similarly,

$$\lim_{x\to x_0} g(x) = \infty, \quad \textit{iff} \quad \forall C > 0 \exists \delta > 0 \quad \textit{such that} \quad g(x) > C, \ \forall |x - x_0| < \delta.$$

$$\lim_{x\to x_0} g(x) = -\infty, \quad \textit{iff} \quad \forall C < 0 \exists \delta > 0 \quad \textit{such that} \quad g(x) < C, \ \forall |x - x_0| < \delta.$$

Theorem 1.7 *If $P(x)$ and $Q(x)$ are polynomials and $c > 1$ is a fixed constant, then*

$$\lim_{x\to\infty} \frac{P(x)}{c^x} = 0, \quad \forall c > 1; \tag{1.15}$$

$$\lim_{x\to\infty} \frac{\ln|Q(x)|}{P(x)} = 0. \tag{1.16}$$

Lemma 1.5 *Let $c > 0$ be a positive constant. Then,*

$$\lim_{x\to\infty} x^{\frac{1}{x}} = 1; \tag{1.17}$$

$$\lim_{x\to\infty} c^{\frac{1}{x}} = 1; \tag{1.18}$$

$$\lim_{x\searrow 0} x^x = 1, \tag{1.19}$$

$$\tag{1.20}$$

where the notation $x \searrow 0$ means that x goes to 0 while always being positive, i.e., $x \to 0$ with $x > 0$.

Lemma 1.6 *If k is a positive integer number, and if $c > 0$ is a positive fixed constant, then*

$$\lim_{k\to\infty} k^{\frac{1}{k}} = 1; \tag{1.21}$$

$$\lim_{k\to\infty} c^{\frac{1}{k}} = 1; \tag{1.22}$$

$$\lim_{k\to\infty} \frac{c^k}{k!} = 0, \tag{1.23}$$

$$\tag{1.24}$$

where $k! = 1 \cdot 2 \cdot \ldots \cdot k$.

1.5 L'Hôpital's rule and connections to Taylor expansions

Theorem 1.8 (L'Hôpital's Rule) *Let x_0 be a real number; allow $x_0 = \infty$ and $x_0 = -\infty$ as well. Let $f(x)$ and $g(x)$ be two differentiable functions.*

(i) Assume that $\lim_{x\to x_0} f(x) = 0$ and $\lim_{x\to x_0} g(x) = 0$. If $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ exists, and if there exists an interval (a, b) around x_0 such that $g'(x) \neq 0$ for all $x \in (a, b) \setminus 0$, then the limit $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ also exists and

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}.$$

(ii) Assume that $\lim_{x\to x_0} f(x)$ is either $-\infty$ or ∞, and that $\lim_{x\to x_0} g(x)$ is either $-\infty$ or ∞. If the limit $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ exists, and if there exists an interval (a, b) around x_0 such that $g'(x) \neq 0$ for all $x \in (a, b) \setminus 0$, then the limit $\lim_{x\to x_0} \frac{f(x)}{g(x)}$ also exists and

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}.$$

1.6 Multivariable functions

Scalar Valued Functions Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n variables denoted by $x_1, x_2, ..., x_n$, and let $x = (x_1, x_2, ..., x_n)$.

Definition 1.2 *Let $f : \mathbb{R}^n \to \mathbb{R}$. The partial derivative of the function $f(x)$ with respect to the variable x_i is denoted by $\frac{\partial f}{\partial x_i}(x)$ and is defined as*

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \to 0} \frac{f(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n) - f(x_1, x_2, ..., x_n)}{h}, \tag{1.25}$$

if the limit from (1.25) *exists and is finite.*

A compact formula for (1.25) can be given as follows: Let e_i be the vector with all entries equal to 0 with the exception of the i-th entry, which is equal for 1, i.e., $e_i(j) = 0$, for $j \neq i$, $1 \leq j \leq n$, and $e_i(j) = 1$. Then,

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h}.$$

Theorem 1.9 *If all the partial derivatives of order k of the function $f(x)$ exist and are continuous, then the order in which partial derivatives of $f(x)$ of order at most k is computed does not matter.*

Definition 1.3 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n variables and assume that $f(x)$ is differentiable with respect to all variables x_i, $i = 1 : n$. The gradient $Df(x)$ of the function $f(x)$ is the following row vector of size n:*

$$Df(x) = \left(\frac{\partial f}{\partial x_1}(x) \frac{\partial f}{\partial x_2}(x) \cdots \frac{\partial f}{\partial x_n}(x) \right). \tag{1.26}$$

Definition 1.4 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of n variables. The Hessian of $f(x)$ is denoted by $D^2 f(x)$ and is defined as the following $n \times n$ matrix:*

$$D^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}. \tag{1.27}$$

Another commonly used notations for the gradient and Hessian of $f(x)$ are $\nabla f(x)$ and $Hf(x)$, respectively.

Vector Valued Functions Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a vector valued function given by

$$F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix},$$

where $x = (x_1, x_2, ..., x_n)$.

Definition 1.5 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ given by $F(x) = (f_j(x))_{j=1:m)}$, and assume that the functions $f_j(x)$, $j = 1 : m$, are differentiable with respect to all variables $x_i = 1 : n$. The gradient $DF(x)$ of the function $F(x)$ is the following matrix of size $m \times n$:*

$$DF(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}. \tag{1.28}$$

1.6.1 Functions of two variables

Scalar Valued Functions Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables denoted by x and y. The partial derivatives of the function $f(x, y)$ with respect to the variables x and y are denoted by $\frac{\partial f}{\partial x}(x, y)$ and $\frac{\partial f}{\partial y}(x, y)$, respectively, and defined as follows:

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h}; \\ \frac{\partial f}{\partial y}(x, y) &= \lim_{h \to 0} \frac{f(x, y+h) - f(x, y)}{h}. \end{aligned}$$

The gradient of $f(x, y)$ is

$$Df(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{pmatrix}. \tag{1.29}$$

The Hessian of $f(x, y)$ is

$$D^2 f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial y \partial x}(x, y) \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{pmatrix}. \tag{1.30}$$

Vector Valued Functions Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$F(x, y) = \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \end{pmatrix}.$$

The gradient of $F(x, y)$ is

$$DF(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x, y) & \frac{\partial f_1}{\partial y}(x, y) \\ \frac{\partial f_2}{\partial x}(x, y) & \frac{\partial f_2}{\partial y}(x, y) \end{pmatrix}. \tag{1.31}$$

1.7 Plain vanilla European Call and Put options

Call Option A Call Option on an underlying asset is a contract between two parties which gives the buyer of the option the right, but not the obligation, to **buy** from the seller of the option one unit of the asset at a predetermined time T in the future, call the maturity of the option, for a predetermined price K, call the strike of the option. For this right, the buyer of the option pays $C(t)$ at time $t < T$ to the seller of the option.

Put Option A Put Option to an underlying asset is a contract between two parties which gives the buyer of the option the right, but not the obligation, to **sell** to the seller of the option one unit of the asset at a predetermined time T in the future, called the maturity of the option, for a predetermined price K, called the strike of the option. For this right, the buyer of the option pays $P(t)$ at time $t < T$ to the seller of the option.
The options described above are plain European options. An American option can be exercised at any time prior to maturity.
In an option contract, two parties exist: the buyer of the option and the seller of the option. The buyer of the option is long the option (or has a long position in the option) and the seller of the option is short the option (or has a short position in the option).
Let $S(t)$ and $S(T)$ be the price of the underlying asset at time t and at maturity T, respectively.
At time t, a call option is

$$\begin{cases} ITM & \text{if} \quad S(t) > K; \\ ATM & \text{if} \quad S(t) = K; \\ OTM & \text{if} \quad S(t) < K, \end{cases}$$

where ITM is in the money, ATM is at the money, and OTM is out of the money (OTM). A put option is

$$\begin{cases} ITM & \text{if} \quad S(t) < K; \\ ATM & \text{if} \quad S(t) = K; \\ OTM & \text{if} \quad S(t) > K. \end{cases}$$

At maturity T, a call option expires

$$\begin{cases} ITM & \text{if} \quad S(T) > K; \\ ATM & \text{if} \quad S(T) = K; \\ OTM & \text{if} \quad X(T) < K. \end{cases}$$

A put option expires

$$\begin{cases} ITM & \text{if} \quad S(T) < K; \\ ATM & \text{if} \quad S(T) = K; \\ OTM & \text{if} \quad X(T) > K. \end{cases}$$

The payoff of a call option at maturity is

$$C(T) = \max(S(T) - K, 0) = \begin{cases} S(T) - K, & \text{if} \quad S(T) > K; \\ 0, & \text{if} \quad S(T) \leq K. \end{cases}$$

The payoff of a put option at maturity is

$$P(T) = \max(K - S(T), 0) = \begin{cases} 0, & \text{if} \quad S(T) \geq K; \\ K - S(T), & \text{if} \quad S(T) < K. \end{cases}$$

1.8 Arbitrage-free pricing

An arbitrage opportunity is an investment opportunity that is guaranteed to earn money without any risk involved.
In an arbitrage-free market, we can infer relationships between prices of various securities, based on the following principle:

Theorem 1.10 (The (Generalized) Law of One Price.) *If two portfolios are guaranteed to have the same value at a future time $\tau > t$ regardless of the state of the market at time τ, then they must have the same value at time t. If one portfolio is guaranteed to be more valuable (or less valuable) than another portfolio at a future time $\tau > t$ regardless of the state of the market at time τ, then that portfolio is more valuable (or less valuable, respectively) than the other one at time $t < \tau$ as well:*
If there exists $\tau > t$ such that $V_1(\tau) = V_2(\tau)$ (or $V_1(\tau) > V_2(\tau)$, or $V_1(\tau) < V_2(\tau)$, respectively for any state of the market at time τ, then $V_1(t) = V_2(t)$ (or $V_1(t) > V_2(t)$, or $V_1(t) < V_2(t)$, respectively).

Corollary 1.1 *If the value of a portfolio of securities is guaranteed to be equal to 0 at a future time $\tau > t$ regardless of the state of the market at time τ, then the value of the portfolio at time t must have been 0 as well:*
If there exists $\tau > t$ such that $V(\tau) = 0$ for any state of the market at time τ, then $V(t) = 0$.

A consequence of Theorem 1.10 is the fact that, if the value of a portfolio at time T in the future is independent of the state of the market at that time, then the value of the portfolio in the present is the risk-neutral discounted present value of the portfolio at time T.
"Risk-neutral discounted present value" refers to the time value of money: cash can be deposited at time t_1 to be returned at time t_2 ($t_2 > t_1$), with interest. The interest rate depends on several factors, one of them being the probability of default of the party receiving the cash deposit. If this probability is zero, or close to zero, then the return is considered risk-free.
For continuously compounded interest, the value $B(t_2)$ at time $t_2 > t_1$ of $B(t_1)$ cash at time t_1 is

$$B(t_2) = e^{r(t_2-t_1)}B(t_1),$$

where r is the risk-free rate between t_1 and t_2. The value $B(t_1)$ at time $t_1 < t_2$ of $B(t_2)$ cash at time t_2 is

$$B(t_1) = e^{-r(t_2-t_1)}B(t_2).$$

Lemma 1.7 *If the value $V(T)$ of a portfolio at time T in the future is independent of the state of the market at time T, then*

$$V(t) = V(T)e^{-r(T-t)}, \qquad (1.32)$$

where $t < T$ and r is the constant risk-free rate.

1.9 The Put-Call parity for European options

The Put-Call parity states that

$$P(t) + S(t) - C(t) = Ke^{-r(T-t)}, \qquad (1.33)$$

where $C(t)$ and $P(t)$ are the values at time t of a European call and put option, respectively, with maturity T and strike K, on the same non-dividend paying asset with spot price $S(t)$. If the underlying asset pays dividends continuously at the rate q, the Put-Call parity has the form

$$P(t) + S(t)e^{-q(T-t)} - C(t) = Ke^{-r(T-t)}. \qquad (1.34)$$

1.10 Forward and Futures contracts

Forward contract A forward contract is an agreement between two parties: one party (the long position) agrees to buy the underlying asset from the other party (the short position) at a specified time in the future and for a specified price, called the forward price. The forward price is chosen such that the forward contract has value zero at the time when the forward contract is entered into.
The contractual forward price F of a forward contract with maturity T and struck at time 0 on a non-dividend paying underlying asset with spot price $S(0)$ is

$$F = S(0)e^{rT},$$

where the interest rate r is assumed to be constant over the life of the forward contract, i.e., between times 0 and T.
If the underlying asset pays dividends continuously at the rate q, the forward price is

$$F = S(0)e^{(r-q)T}.$$

Futures contract A futures contract has a similar structure as a forward contract, but it requires the delivery of the underlying asset for the futures price. The forward and futures prices are, in theory, the same, if the risk-free interest rates are constant or deterministic, i.e., just functions of time. Major differences exist between the ways forward and futures contracts are structures, settles, and traded:

- Futures contracts trade on an exchange and have standard features, while forward contracts are over-the-counter instruments;
- Futures are marked to market and settled in a margin account on a daily basis, while forward contracts are settled in cash at maturity;
- Futures have a range of delivery dates, while forward contracts have a specified delivery date;
- Futures carry almost no credit risk, since they are settled daily, while entering into a forward contract carries some credit risk.

1.11 Exercises

1.11.1 Question 1

Use the integration by parts to compute $\int \ln(x)dx$.

Answer Let $f(x) = 1$ and $g(x) = \ln(x)$,

$$\begin{aligned}\int \ln(x)dx &= \int 1 \cdot \ln(x)dx \\ &= x\ln(x) - \int x\frac{1}{x}dx \\ &= x\ln(x) - x + C\end{aligned}$$

1.11.2 Question 2

Compute $\int \frac{1}{x\ln(x)}dx$ by using the substitution $u = \ln(x)$.

Answer Let $u = \ln(x)$, then $du = \frac{dx}{x}$

$$\begin{aligned}\int \frac{1}{x\ln(x)}dx &= \int \frac{1}{u}du \\ &= \ln(|u|) + C \\ &= \ln(|\ln(x)|) + C\end{aligned}$$

1.11.3 Question 3

Show that $(\tan x)' = 1/(\cos x)^2$ and

$$\int \frac{1}{1+x^2}dx = \arctan(x) + C.$$

Note: The antiderivative of a rational function is often computed using the substitution $x = \tan(\frac{x}{2})$.

Answer Using Quotient Rule,

$$\begin{aligned}(\tan(x))' &= \left(\frac{\sin(x)}{\cos(x)}\right)' \\ &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} \\ &= \frac{1}{\cos^2(x)}\end{aligned}$$

Let $x = \tan(u)$, then $dx = \frac{1}{\cos^2(u)}du$.

$$\begin{aligned}\int \frac{1}{1+x^2}dx &= \int \frac{1}{1+\tan^2(u)}\frac{1}{\cos^2(u)}du \\ &= \int \frac{1}{\frac{\cos^2(u)+\sin^2(u)}{\cos^2(u)}}\frac{1}{\cos^2(u)}du \\ &= \int du \\ &= u + C \\ &= \arctan(x) + C\end{aligned}$$

1.11.4 Question 4

Use l'Hôpital's rule to show that the following two Taylor approximations hold when x is close to 0:

$$\sqrt{1+x} \approx 1 + \frac{x}{2};$$

$$e^x \approx 1 + x + \frac{x^2}{2}.$$

In other words, show that the following limits exist and are constant:

$$\lim_{x\to 0}\frac{\sqrt{1+x}-(1+\frac{x}{2})}{x^2} \quad \text{and} \quad \lim_{x\to 0}\frac{e^x-(1+x+\frac{x^2}{2})}{x^3}.$$

Answer

$$\begin{aligned}
\lim_{x\to 0}\frac{\sqrt{1+x}-(1+\frac{x}{2})}{x^2} &= \lim_{x\to 0}\frac{\frac{1}{2}(1+x)^{-\frac{1}{2}}-\frac{1}{2}}{2x}\\
&= \lim_{x\to 0}\frac{-\frac{1}{4}(1+x)^{-\frac{3}{2}}}{2}\\
&= -\frac{1}{8}\\
\lim_{x\to 0}\frac{e^x-(1+x+\frac{x^2}{2})}{x^3} &= \lim_{x\to 0}\frac{e^x-1-x}{3x^2}\\
&= \lim_{x\to 0}\frac{e^x-1}{6x}\\
&= \lim_{x\to 0}\frac{e^x}{6}\\
&= \frac{1}{6}
\end{aligned}$$

1.11.5 Question 5

Use the definition of e, i.e., $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$, to show that

$$\frac{1}{e} = \lim_{x\to\infty}\left(1-\frac{1}{x}\right)^x.$$

Hint: Use the fact that

$$\frac{1}{1+\frac{1}{x}} = \frac{x}{x+1} = 1-\frac{1}{x+1}.$$

Answer

$$\begin{aligned}
e &= \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x\\
\frac{1}{e} &= \lim_{x\to\infty}\left(\frac{1}{1+\frac{1}{x}}\right)^x\\
&= \lim_{x\to\infty}\left(1-\frac{1}{x+1}\right)^x\\
&= \lim_{x\to\infty}\left(1-\frac{1}{x+1}\right)^{x+1}\\
&= \lim_{y\to\infty}\left(1-\frac{1}{y}\right)^y
\end{aligned}$$

where $y = x+1$, since

$$\lim_{x\to\infty}\left(1-\frac{1}{x+1}\right) = 1$$

1.11.6 Question 6

Let K, T, σ, and r be positive constants, and define the function $g : \mathbb{R} \to \mathbb{R}$ as

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_0^{b(x)} e^{-\frac{y^2}{2}} dy,$$

where $b(x) = (\ln(\frac{x}{K}) + (r + \frac{\sigma^2}{2})T)/(\sigma\sqrt{T})$. Compute $g'(x)$. Note: This function is related to the Delta of a plain vanilla Call option.

Answer

$$\begin{aligned}
g(x) &= \frac{1}{\sqrt{2\pi}} \int_0^{b(x)} e^{-\frac{y^2}{2}} dy \\
g'(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(b(x))^2}{2}} b'(x) \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{(b(x))^2}{2}} \frac{1}{\sigma\sqrt{T}x} \\
&= \frac{1}{x\sigma\sqrt{2\pi T}} \exp\left(- \frac{\left(\ln\left(\frac{x}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right)^2}{2\sigma^2 T} \right)
\end{aligned}$$

1.11.7 Question 7

Let $f(x)$ be a continuous function. Show that

$$\lim_{h\to 0} \frac{1}{2h} \int_{a-h}^{a+h} f(x)dx = f(a), \quad \forall a \in \mathbb{R}.$$

Note: Let $F(x) = \int f(x)dx$. The central finite difference approximation of $F'(a)$ is

$$F'(a) = \frac{F(a+h) - F(a-h)}{2h} + O(h^2), \tag{1.35}$$

as $h \to 0$ (if $F^{(3)}(x) = f''(x)$ is continuous). Since $F'(a) = f(a)$, formula (1.35) can be written as

$$f(a) = \frac{1}{2h} \int_{a-h}^{a+h} f(x)dx + O(h^2).$$

Answer

$$\begin{aligned}
\lim_{h\to 0} \frac{1}{2h} \int_{a-h}^{a+h} f(x)dx &= \lim_{h\to 0} \frac{f(a+h) - f(a-h)(-1)}{2} \\
&= \frac{f(a) + f(a)}{2} \\
&= \frac{f(a)}{2}
\end{aligned}$$

1.11.8 Question 8

Let $f : \mathbb{R} \to \mathbb{R}$ given by $f(y) = \sum_{i=1}^{n} c_i e^{-yt_i}$, where c_i and t_i, $i = 1 : n$, are positive constants. Compute $f'(y)$ and $f''(y)$.
Note: The function $f(y)$ represents the price of a bond with cash flows c_i paid at time t_i as a function of the yield y of the bond. When scaled appropriately, the derivative of $f(y)$ with respect to y give the duration and convexity of the bond.

Answer

$$\begin{aligned} f(y) &= \sum_{i=1}^{n} c_i e^{-yt_i} \\ f'(y) &= -\sum_{i=1}^{n} c_i t_i e^{-yt_i} \\ f''(y) &= \sum_{i=1}^{n} c_i t_i^2 e^{-yt_i} \end{aligned}$$

1.11.9 Question 9

Let $f : \mathbb{R}^3 \to \mathbb{R}$ given by $f(x) = 2x_1^2 - x_1x_2 + 3x_2x_3 - x_3^2$, where $x = (x_1, x_2, x_3)$.

Part (i)

Compute the gradient and Hessian of the function $f(x)$ at the point $a = (1, -1, 0$, i.e., compute $Df(1, -1, 0)$ and $D^2 f(1, -1, 0)$.

Answer

$$\begin{aligned} f(x) &= 2x_1^2 - x_1x_2 + 3x_2x_3 - x_3^2 \\ Df(x) &= \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) & \frac{\partial f}{\partial x_3}(x) \end{pmatrix} \\ &= \begin{pmatrix} 4x_1 - x_2 & -x_1 + 3x_3 & 3x_2 - 2x_3 \end{pmatrix} \\ &= \begin{pmatrix} 5 & -1 & -3 \end{pmatrix} \\ D^2 f(x) &= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_3 \partial x_1}(x) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \frac{\partial^2 f}{\partial x_3 \partial x_2}(x) \\ \frac{\partial^2 f}{\partial x_1 \partial x_3}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_3}(x) & \frac{\partial^2 f}{\partial x_3^2}(x) \end{pmatrix} \\ &= \begin{pmatrix} 4 & -1 & 0 \\ -1 & 0 & 3 \\ 0 & 3 & -2 \end{pmatrix} \end{aligned}$$

Part (ii)

Show that

$$f(x) = f(a) + Df(a)(x - a) + \frac{1}{2}(x - a)^t D^2 f(a)(x - a). \tag{1.36}$$

Answer

$$
\begin{aligned}
&f(a) + Df(a)(x-a) + \frac{1}{2}(x-a)^t D^2 f(a)(x-a) \\
&= 3 + \begin{pmatrix} 5 & -1 & -3 \end{pmatrix} \begin{pmatrix} x_1 - 1 \\ x_2 + 1 \\ x_3 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x_1 - 1 \\ x_2 + 1 \\ x_3 \end{pmatrix}^t \begin{pmatrix} 4 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 3 & -2 \end{pmatrix} \begin{pmatrix} x_1 - 1 \\ x_2 + 1 \\ x_3 \end{pmatrix} \\
&= 3 + (5x_1 - x_2 - 3x_3 - 6) + (2x_1^2 - 5x_1 - x_1x_2 + x_2 + 3x_2x_3 + 3x_3 - x_3^2 + 3) \\
&= 2x_1^2 - x_1x_2 + 3x_2x_3 - x_3^2 \\
&= f(x)
\end{aligned}
$$

1.11.10 Question 10

Let

$$
u(x,t) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}, \quad \text{for} \quad t > 0, x \in \mathbb{R}.
$$

Compute $\frac{\partial u}{\partial t}$ and $\frac{\partial^2 u}{\partial x^2}$, and show that

$$
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.
$$

Answer

$$
\begin{aligned}
\frac{\partial u}{\partial t} &= -\frac{1}{2\sqrt{4\pi}} t^{\frac{3}{2}} e^{-\frac{x^2}{4t}} + \frac{1}{\sqrt{4\pi t}} \frac{x^2}{4t^2} e^{-\frac{x^2}{4t}} \\
&= -\frac{1}{2t\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} + \frac{x^2}{4t^2} \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \\
\frac{\partial u}{\partial x} &= -\frac{x}{2t} \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \\
\frac{\partial^2 u}{\partial x^2} &= \frac{1}{2t\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} + \frac{x^2}{4t^2} \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}} \\
&= \frac{\partial u}{\partial t}
\end{aligned}
$$

1.11.11 Question 11

Consider a portfolio with the following positions:

- long one call option with strike $K_1 = 30$;
- short two call options with strike $K_2 = 35$;
- long one call option with strike $K_3 = 40$.

All options are on the same underlying asset and have maturity T. Draw the payoff diagram at maturity of the portfolio, i.e., plot the value of the portfolio $V(T)$ at maturity as a function of $S(T)$, the price of the underlying asset at time T.

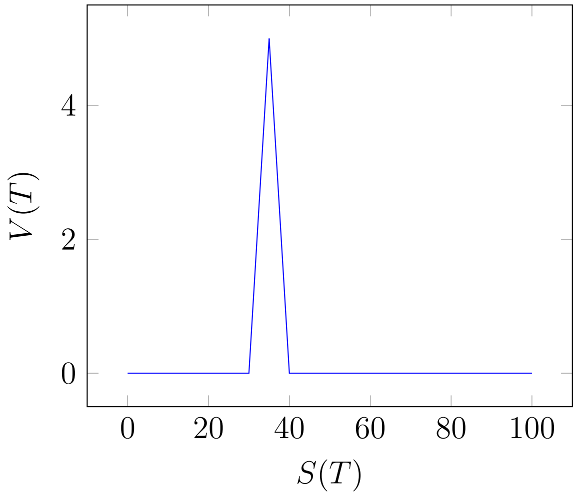
Answer Payoffs of the call options at maturity are:

$$C_1(T) = \max(S(T) - K_1, 0) = \max(S(T) - 30, 0);$$
$$C_2(T) = \max(S(T) - K_2, 0) = \max(S(T) - 35, 0);$$
$$C_3(T) = \max(S(T) - K_3, 0) = \max(S(T) - 40, 0).$$

Value of the portfolio at maturity is

$$V(T) = C_1(T) - 2C_2(T) + C_3(T).$$

Payoff diagram of butterfly spread



1.11.12 Question 12

Draw the payoff diagram at maturity of a bull spread with a long position in a call with strike 30 and short a call with strike 35, and of a bear spread with long a put of strike 20 and short a put of strike 15.

Answer Payoffs of the call options in bull spread at maturity are:

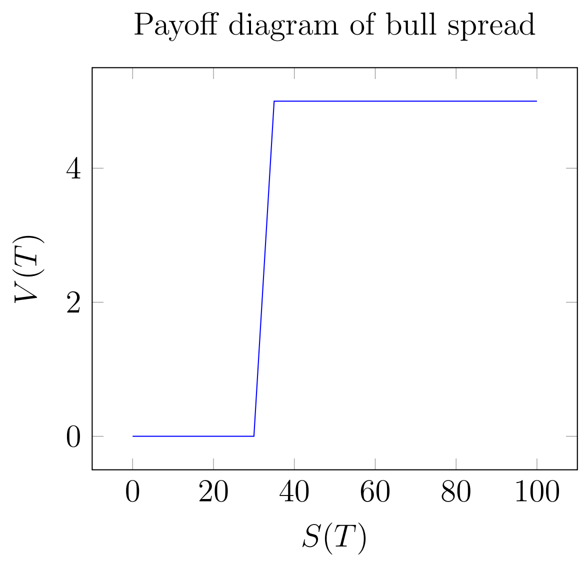
$$C_1(T) = \max(S(T) - K_{C_1}, 0) = \max(S(T) - 30, 0);$$
$$C_2(T) = \max(S(T) - K_{C_2}, 0) = \max(S(T) - 35, 0).$$

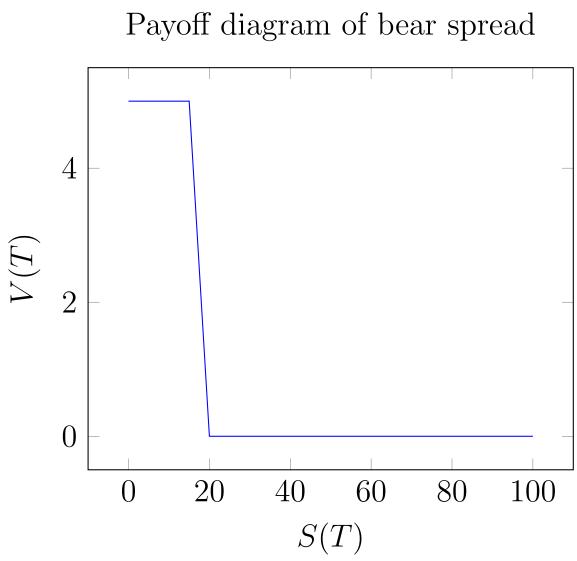
Payoffs of the put options in bear spread at maturity are:

$$P_1(T) = \max(K_{P_1} - S(T), 0) = \max(20 - S(T), 0);$$
$$P_2(T) = \max(K_{P_2} - S(T), 0) = \max(15 - S(T), 0).$$

Value of the portfolios at maturity are

$$V_{bull}(T) = C_1(T) - C_2(T);$$
$$V_{bear}(T) = P_1(T) - P_2(T).$$





1.11.13 Question 13

Which of the following two portfolios would you rather hold:

- Portfolio 1: Long one call option with strike $K = X - 5$ and long one call option with strike $K = X + 5$;
- Portfolio 2: Long two call options with strike $K = X$?

(All options are on the same asset and have the same maturity.)

Answer Payoffs of the call options in Portfolio 1 at maturity are:

$$C_1(T) = \max(S(T) - (X - 5), 0);$$
$$C_2(T) = \max(S(T) - (X + 5), 0).$$

Payoffs of the call options in Portfolio 2 at maturity are:

$$C_3(T) = \max(S(T) - X, 0).$$

Value of the portfolios at maturity are

$$V_1(T) = C_1(T) + C_2(T) = \max(S(T) - (X - 5), 0) + \max(S(T) - (X + 5), 0);$$
$$V_2(T) = 2C_3(T) = 2\max(S(T) - X, 0).$$

Long Portfolio 1 and short Portfolio 2 is equivalent to long a butterfly spread, which is always non-negative at maturity. Therefore, it is better owning Portfolio 1 since its payoff at maturity will always be at least as big as the payoff of Portfolio 2.

1.11.14 Question 14

Call options with strikes 100, 120, and 130 on the same underlying asset and with the same maturity are trading for 8, 5, and 3, respectively (there is no bid-ask spread). Is there an arbitrage opportunity present? If yes, how can you make a riskless profit?

S(T)	V(T)
$S(T) \le K_1$	0
$K_1 < S(T) \le K_2$	$x_1 S(T) - x_1 K_1)$
$K_2 < S(T) \le K_3$	$(x_1 + x_2)S(T) - x_1 K_1 - x_2 K_2$
$K_3 < S(T)$	$(x_1 + x_2 + x_3)S(T) - x_1 K_1 - x_2 K_2 - x_3 K_3$

Answer For arbitrage opportunity to be present, the portfolio must have non-negative payoff at maturity and with a negative cost setting up.
Let $K_1 = 100 < K_2 = 120 < K_3 = 130$ be the strike of the options. Let x_1, x_2, and x_3 be options positions at time 0. The portfolio value at time 0 is

$$V(0) = x_1 C_1(0) + x_2 C_2(0) + x_3 C_3(0).$$

The portfolio value at maturity T is

$$\begin{aligned} V(T) &= x_1 C_1(T) + x_2 C_2(T) + x_3 C_3(T) \\ &= x_1 \max(S(T) - K_1, 0) + x_2 \max(S(T) - K_2, 0) + x_3 \max(S(T) - K_3, 0) \end{aligned}$$

The payoff at maturity is as follows:
$V(T)$ is non-negative when $S(T) \le K_2$ only if a long position is taken for the option with strike K_1, i.e., $x_1 \ge 0$. $V(T)$ is non-negative for any value of $S(T)$ if and only if $x_1 \ge 0$, if $V(T)$ when $S(T) = K_3$ is non-negative, and if $x_1 + x_2 + x_3 \ge 0$.
An arbitrage exists if and only if the values $C_1(0)$, $C_2(0)$, $C_3(0)$ are such that the we find x_1, x_2, and x_3 with the following properties:

$$\begin{aligned} x_1 C_1(0) + x_2 C_2(0) + x_3 C_3(0) &< 0; \\ x_1 &\ge 0; \\ (x_1 + x_2)K_3 - x_1 K_1 - x_2 K_2 &\ge 0; \\ x_1 + x_2 + x_3 &\ge 0. \end{aligned}$$

For $C_1(0) = 8$, $C_2(0) = 5$, $C_3(0) = 3$ and $K_1 = 100$, $K_2 = 120$, $K_3 = 130$, the problem becomes finding $x_1 \ge 0$, x_2, and x_3 such that

$$\begin{aligned} 8x_1 + 5x_2 + 3x_3 &< 0; \\ 30x_1 + 10x_2 &\ge 0; \\ x_1 + x_2 + x_3 &\ge 0. \end{aligned}$$

Arbitrage can occur for a portfolio with long positions in the options with lowest and highest strikes and a short position in the option with the middle strike. Choosing $x_3 = -x_1 - x_2$, the constraints become

$$\begin{aligned} 5x_1 + 2x_2 &< 0; \\ 3x_1 + x_2 &\ge 0. \end{aligned}$$

The constraints are satisfied when

$$x_1 = 1; \quad x_2 = -3; \quad x_3 = 2.$$

Buying one option with strike 100, selling three options with strike 120, and buying two options with strike 130 will generate a positive cash flow of \$1, and will result in a portfolio that will not lose money, regardless of the value of the underlying asset at maturity of the options.

1.11.15 Question 15

A stock with spot price 40 pays dividends continuously at a rate of 3%. The four months at-the-money put and call options on this asset are trading at \$2 and \$4, respectively. The risk-free rate is constant and equal to 5% for all times. Show that the Put-Call parity is not satisfied and explain you would take advantage of this arbitrage opportunity.

Answer Given the following values: $S = 40$, $K = 40$, $q = 0.03$, $C = 4$, $P = 2$, $r = 0.05$, $T = 1/3$, we have

$$P + Se^{-qT} - C = \approx 37.6020 < 39.3389 \approx Ke^{-rT}$$

To take advantage of the arbitrage opportunity, we:

1. Short one call (+ \$4)
2. Long one put (- \$2)
3. Short one forward-like position
   - Short one share (- \$39.6020)
   - Invest in risk-free bond (+ \$39.3389)

The riskless profit at maturity will be approximately \$1.74.

1.11.16 Question 16

The bid and ask prices for a six months European call option with strike 40 on a non-dividend-paying stock with spot price 42 are \$5 and \$5.5, respectively. The bid and ask prices for a six months European put option with strike 40 on the same underlying asset are \$2.75 and \$3.25, respectively. Assume that the risk free rate is equal to 0. Is there an arbitrage opportunity present?

Answer The Put-Call parity is

$$\begin{aligned} C - P &= S - K \\ C - P &= 42 - 40 = 2 \end{aligned}$$

An arbitrage opportunity exists if $C - P$ can be "bought" for less than \$2, of "sold" for more than \$2.
The bid-ask spreads are

$$\begin{aligned} C &\in [5, 5.5] \\ P &\in [2.75, 3.25] \end{aligned}$$

Thus, the possible range of $C - P$ is $C - P \in [1.75, 2.75]$.
Since $C - P$ can be "bought" for more than \$2 and "sold" for less than \$2, there is no arbitrage opportunity.

1.11.17 Question 17

You expect that an asset with spot price \$35 will trade in the \$40 - \$45 range in one year. One year at-the-money (ATM) calls on the asset can be bought for \$4. To act on the expected stock price appreciation, you decide to either buy the asset, or to buy ATM calls. Which strategy is better, depending on where the asset price will be in a year?

Answer Payoff for the strategy of buying the asset is

$$V_1(T) = \frac{x}{35}S(T),$$

where x is the amount invested and $S(T)$ is the spot price of the asset in one year. Payoff for the strategy of buying call option is

$$V_2(T) = \frac{x}{4}\max(S(T) - 35)$$

If the asset price is less than \$35 in a year, the calls will expire worthless while the first strategy of buying the asset will not lose all its value. However, the strategy of buying call options can be more profitable if the asset appreciates in value, i.e., the spot price is larger than \$35. The break even point of the two strategies is approximately \$39.5161, since

$$\frac{x}{35}S(T) = \frac{x}{4}\max(S(T) - 35) \Leftrightarrow S(T) = 39.5161$$

If the price of the asset is expected to be \$40 - \$45, buying call options will be the more profitable strategy.

1.11.18 Question 18

The risk free rate is 8% compounded continuously and the dividend yield of a stock index is 3%. The index is 12,000 and the futures prices of a contract deliverable in three months is 12,100. Is there an arbitrage opportunity, and how would you take advantage of it?

Answer Futures fair price is given by

$$\begin{aligned} Se^{(r-q)T} &= 12000e^{(0.08-0.03)/4} \\ &= 12150.94 \\ &> 12100 \end{aligned}$$

The futures contract is underpriced. To take advantage of the arbitrage opportunity, we long the future and short the index (spot).
At maturity, the asset is bought for 12100 and the short is closed. The realized gain is the interest on the cash from the short position minus 12,100, i.e.,

$$e^{0.08/4}(12000e^{-0.03/4}) - 12100 = 50.94$$

Chapter 2

Improper integrals. Numerical integration. Interest rates. Bonds.

2.1 Double integrals

Let $D \subset \mathbb{R}^2$ be a bounded region and let $f : D \to \mathbb{R}$ be a continuous function. The double integral of f over D, denoted by

$$\int\int_D f,$$

represents the volume of the three dimensional body between the domain D in the two dimensional plane and the graph of the function $f(x, y)$.
For simplicity, assume that the domain D is bounded and convex, i.e., for any two points x_1 and x_2 in D, all the points on the segment joining x_1 and x_2 are in D as well. Also, assume that there exist two continuous functions $f_1(x)$ and $f_2(x)$ such that D can be described as follow:

$$D = (x, y)|a \leq x \leq b \quad \text{and} \quad f_1(x) \leq y \leq f_2(x). \tag{2.1}$$

The functions $f_1(x)$ and $f_2(x)$ are well defined by (2.1) since the domain D is bounded and convex. Then, by definition,

$$\int\int_d f(x, y)dydx = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} f(x, y)dy \right) dx. \tag{2.2}$$

If there exists two continuous functions $g_1(x)$ and $g_2(x)$ such that
$D = (x, y)|c \leq y \leq d$ and $g_1(y) \leq x \leq g_2(x)$, then, by definition,

$$\int\int_D f(x, y)dxdy = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} f(x, y)dx \right) dy. \tag{2.3}$$

Theorem 2.1 (Fubini's Theorem) *With the notations above, if the function $f(x, y)$ is continuous, then the integrals (2.2) and (2.3) are equal to each other and to the double integral of $f(x, y)$ over D, i.e., the order of integration does not matter:*

$$\int\int_D f = \int\int_D f(x, y)dxdy = \int\int_D f(x, y)dydx. \tag{2.4}$$

2.2 Improper integrals

We consider three types of improper integrals:

Type 1 Integrate the function $f(x)$ over an infinite interval of the form $[a, \infty)$ or $(-\infty, b]$. The integral $\int_a^\infty f(x)dx$ exists if and only if the limit as $t \to \infty$ of the definite integral of $f(x)$ between a and t exists and is finite. The integral $\int_{-\infty}^b f(x)dx$ exists if and only if the limit as $t \to -\infty$ of the definite integral of $f(x)$ between t and b exists and is finite. Then

$$\int_a^\infty f(x)dx = \lim_{t\to\infty} \int_a^t f(x)dx;$$
$$\int_{-\infty}^b f(x)dx = \lim_{t\to-\infty} \int_t^b f(x)dx.$$

Adding and subtracting improper integrals of this type follows rules similar to those for definite integrals:

Lemma 2.1 *Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function over the interval $[a, \infty)$. If $b > a$, then $f(x)$ is also integrable over the interval $[b, \infty)$ and*

$$\int_a^\infty f(x)dx - \int_b^\infty f(x)dx = \int_a^b f(x)dx.$$

Let $f(x)$ be an integrable function over the interval $(-\infty, b]$. If $a < b$, then $f(x)$ is also integrable over the interval $(-\infty, a]$ and

$$\int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx = \int_a^b f(x)dx.$$

Type 2 Integrate the function $f(x)$ over an interval $[a, b]$ where $f(x)$ is unbounded as x approaches the end points a and/or b. For example, if the limit as $x \searrow a$ of $f(x)$ is infinite, then $\int_a^b f(x)dx$ exists if and only if the limit as *tsearrowa* of the definite integral of $f(x)$ between t and b exists and is finite, i.e.,

$$\int_a^b f(x)dx = \lim_{x\searrow a} \int_t^b f(x)dx.$$

Type 3 Integrate the function $f(x)$ on the entire real axis, i.e., on $(-\infty, \infty)$. The integral $\int_{-\infty}^\infty f(x)dx$ exists if and only if a real number a such that both $\int_{-\infty}^a f(x)dx$ and $\int_a^\infty f(x)dx$ exist. Then,

$$\int_{-\infty}^\infty f(x)dx = \int_{-\infty}^a f(x)dx + \int_a^\infty f(x)dx \tag{2.5}$$
$$= \lim_{t_1\to-\infty} \int_{t_1}^a f(x)dx + \lim_{t_2\to\infty} \int_a^{t_2} f(x)dx. \tag{2.6}$$

It is incorrect to use, instead of (2.6), the following definition for the integral $f(x)$ over the real axis $(-\infty, \infty)$:

$$\int_{-\infty}^\infty f(x)dx = \lim_{t\to\infty} \int_t^t f(x)dx. \tag{2.7}$$

However, if we know that the function $f(x)$ is integrable over the entire real axis, then we can use formula (2.7) to evaluate it:

Lemma 2.2 *If the improper integral $\int_{-\infty}^{\infty} f(x)dx$ exists, then*

$$\int_{-\infty}^{\infty} f(x)dx = \lim_{t\to\infty} \int_{-t}^{t} f(x)dx. \tag{2.8}$$

2.3 Differentiating improper integrals with respect to the integration limits

Lemma 2.3 *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function such that the improper integral $\int_{-\infty}^{\infty} f(x)dx$ exists. Let $g, h : \mathbb{R} \to \mathbb{R}$ be given by*

$$g(t) = \int_{-\infty}^{b(t)} f(x)dx; \quad h(t) = \int_{a(t)}^{\infty} f(x)dx,$$

where $a(t)$ and $b(t)$ are differentiable functions. Then $g(t)$ and $h(t)$ are differentiable, and

$$g'(t) = f(b(t))b'(t);$$
$$h'(t) = -f(a(t))a'(t).$$

2.4 Numerical methods for computing definite integrals: Midpoint rule, Trapezoidal rule, and Simpson's rule

Computing the value of a definite integral using the Fundamental Theorem of Calculus is not always possible. The approximate values of the definite integral are computed using numerical integration methods in these cases. We present three of the most common such methods.

Let $f : [a, b] \to \mathbb{R}$ be an integrable function. To compute an approximate value of the integral

$$I = \int_a^b f(x)dx,$$

we partition the interval $[a, b]$ into n intervals of equal size $h = \frac{b-a}{h}$ by using the nodes $a_i = a + ih$, for $i = 0 : n$, i.e.,

$$a = a_0 < a_1 < a_2 < ... < a_{n-1} < a_n = b.$$

Note that $a_i - a_{i-1} = h$, $i = 1 : n$. Let x_i be the midpoint of the interval $[a_{i-1}, a_i]$, i.e.,

$$x_i = \frac{a_{i-1} + a_i}{2}, \quad \forall i = 1 : n.$$

The integral I can be written as

$$I = \sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} f(x)dx. \tag{2.9}$$

On each interval $[a_{i-1}, a_i]$, $i = 1 : n$, the function $f(x)$ is approximated by a simpler function whose integral on $[a_{i-1}, a_i]$ can be computed exactly. The resulting values are summed up to obtain an approximate value of I. Depending on whether constant functions, linear functions, or quadratic functions are used to approximate $f(x)$, the resulting numerical integration methods are called the Midpoint rule, the Trapezoidal rule, and the Simpson's rule, respectively.

Midpoint rule Approximate $f(x)$ on the interval $[a_{i-1}, a_i]$ by the constant function $c_i(x)$ equal to the value of the function f at the midpoint x_i of the interval $[a_{i-1}, a_i]$, i.e.,

$$c_i(x) = f(x_i), \quad \forall x \in [a_{i-1}, a_i]. \tag{2.10}$$

Then,

$$\int_{a_{i-1}}^{a_i} f(x)dx \approx \int_{a_{i-1}}^{a_i} c_i(x)dx = (a_i - a_{i-1})f(x_i) = hf(x_i). \tag{2.11}$$

From (2.9) and (2.11), we obtain that the Midpoint Rule approximation I_n^M of I corresponding to n partition intervals is

$$\begin{aligned} I_n^M &= \sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} c_i(x)dx \\ &= h\sum_{i=1}^{n} f(x_i). \end{aligned} \tag{2.12}$$

Trapezoidal Rule Approximate $f(x)$ on the interval $[a_{i-1}, a_i]$ by the linear function $l_i(x)$ equal to $f(x)$ at the end points a_{i-1} and a_i, i.e.,

$$l_i(a_{i-1}) = f(a_{i-1}) \quad \text{and} \quad l_i(a_i) = f(a_i).$$

By linear interpolation, it is easy to see that

$$l_i(x) = \frac{x - a_{i-1}}{a_i - a_{i-1}} f(a_i) + \frac{a_i - x}{a_i - a_{i-1}} f(a_{i-1}), \quad \forall x \in [a_{i-1}, a_i]. \tag{2.13}$$

Then,

$$\int_{a_{i-1}}^{a_i} f(x)dx \approx \int_{a_{i_1}}^{a_i} l_i(x)dx = \frac{h}{2}(f(a_{i-1}) + f(a_i)). \tag{2.14}$$

From (2.9) and (2.14), we obtain that the Trapezoidal Rule approximation I_n^T of I corresponding to n partition intervals is

$$\begin{aligned} I_n^T &= \sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} l_i(x)dx \\ &= \frac{h}{2}\left(f(a_0) + 2\sum_{i=1}^{n-1} f(a_i) + f(a_n)\right) \end{aligned} \tag{2.15}$$

Simpson's Rule Approximate $f(x)$ on the interval $[a_{i-1}, a_i]$ by the quadratic function $q_i(x)$ equal to $f(x)$ at a_{i-1}, a_i, and at the midpoint $x_i = \frac{a_{i-1}-a_i}{2}$, i.e.,

$$q_i(a_{i-1}) = f(a_{i-1}); q_i(x_i) = f(x_i) \quad \text{and} \quad q_i(a_i) = f(a_i).$$

By quadratic interpolation, we find that

$$\begin{aligned} q_i(x) =& \frac{(x-a_{i-1})(x-x_i)}{(a_i-a_{i-1})(a_i-x_i)} f(a_i) + \frac{(a_i-x)(x-a_{i-1}}{(a_i-x_i)(x_i-a_{i-1})} f(x_i) \\ &+ \frac{(a_i-x)(x_i-x)}{(a_i-a_{i-1})} f(a_{i-1}), \quad \forall x \in [a_{i-1}, a_i]. \end{aligned} \tag{2.16}$$

Then,

$$\int_{a_{i-1}}^{a_i} f(x)dx \approx \int_{a_{i-1}}^{a_i} q_i(x)dx = \frac{h}{6}(f(a_{i-1}) + 4f(x_i) + f(a_i)). \tag{2.17}$$

From (2.9) and (2.17), we obtain that the Simpson's Rule approximation I_n^S of I corresponding to n partition intervals is

$$\begin{aligned} I_n^S &= \sum_{i=1}^{n} \int_{a_{i-1}}^{a_i} q_i(x)dx \\ &= \frac{h}{6}\left(f(a_0) + 2\sum_{i=1}^{n-1} f(a_i) + f(a_n) + 4\sum_{i=1}^{n} f(x_i) \right). \end{aligned} \tag{2.18}$$

2.5 Convergence of the Midpoint, Trapezoidal, and Simpson's rules

We derived the formulas (2.12), (2.15), and (2.18) for computing approximate values I_n^M, I_n^T, and I_n^S of the integral

$$I = \int_a^b f(x)dx$$

corresponding to the Midpoint, Trapezoidal, and Simpson's rules, respectively. In this section, we discuss the convergence of these methods.

Definition 2.1 *Denote by I_n the approximation of I obtained using a numerical integration method with n partition integrals. The method is convergent if and only if the approximation I_n converge to I as the number of intervals n goes to infinity (and therefore as $h = \frac{b-a}{h}$ goes to 0), i.e.,*

$$\lim_{n\to\infty} |I - I_n| = 0.$$

The order of convergence of the numerical integration method is $k > 0$ if and only if

$$|I - I_n| = O(h^k) = O\left(\frac{1}{n^k}\right).$$

Theorem 2.2 *Let $I = \int_a^b f(x)dx$, and let I_n^M, I_n^T, or I_n^S be the approximations of I given by the Midpoint, Trapezoidal, and Simpson's rules corresponding to n partition intervals of size $h = \frac{b-a}{n}$.*

(i) If $f''(x)$ exists and is continuous on $[a,b]$, then the approximation errors of the Midpoint and Trapezoidal rules can be bounded from above as follows:

$$|I - I_n^M| \le \frac{h^2}{24}(b-a) \max_{a \le x \le b} |f''(x)|; \tag{2.19}$$

$$|I - I_n^T| \le \frac{h^2}{12}(b-a) \max_{a \le x \le b} |f''(x)|. \tag{2.20}$$

Thus, the Midpoint and Trapezoidal rules are quadratically convergent, i.e.,

$$|I - I_n^M| = O\left(\frac{1}{n^2}\right), \quad as \quad n \to \infty; \tag{2.21}$$

$$|I - I_n^T| = O\left(\frac{1}{n^2}\right), \quad as \quad n \to \infty. \tag{2.22}$$

$$\tag{2.23}$$

(ii) If $f^{(4)}(x)$ exists and is continuous on $[a,b]$, then

$$|I - I_n^S| \le \frac{h^4}{2880}(b-a) \max_{a \le x \le b} |f^{(4)}(x)|, \tag{2.24}$$

and Simpson's rule is fourth order convergent, i.e.,

$$|I - I_n^S| O\left(\frac{1}{n^4}\right), \quad as \quad n \to \infty. \tag{2.25}$$

The upper bounds (2.19), (2.20), and (2.24) can be established using the following approximation error results: For any $i = 1 : n$, there exist points $\xi_{i,T}$, $\xi_{i,M}$, and $\xi_{i,S}$ in the interval (a_{i-1}, a_i) such that

$$\int_{a_{i-1}}^{a_i} f(x)dx - \int_{a_{i-1}}^{a_i} c_i(x)dx = \frac{h^3}{24} f''(\xi_{i,T}); \tag{2.26}$$

$$\int_{a_{i-1}}^{a_i} f(x)dx - \int_{a_{i-1}}^{a_i} l_i(x)dx = -\frac{h^3}{12} f''(\xi_{i,M}); \tag{2.27}$$

$$\int_{a_{i-1}}^{a_i} f(x)dx - \int_{a_{i-1}}^{a_i} q_i(x)dx = -\frac{h^5}{2880} f^{(4)}(\xi_{i,S}), \tag{2.28}$$

where $c_i(x)$, $l_i(x)$, and $q_i(x)$ are given by (2.10), (2.13), and (2.16), respectively.

2.5.1 Implementation of numerical integration methods

Computing approximate values of the definite integral of a given function $f(x)$ on the interval $[a,b]$ using the Midpoint, Trapezoidal, or Simpson's rules requires implementation of formulas (2.12), (2.15), and (2.18), i.e.,

$$I_n^M = h \sum_{i=1}^{n} f(x_i); \tag{2.29}$$

$$I_n^T = h\left(\frac{f(a_0)}{2} + \frac{f(a_n)}{2}\right) + h \sum_{i=1}^{n-1} f(a_i); \tag{2.30}$$

$$I_n^S = h\left(\frac{f(a_0)}{6} + \frac{f(a_n)}{6}\right) + \frac{h}{3} \sum_{i=1}^{n-1} f(a_i) + \frac{2h}{3} \sum_{i=1}^{n} f(x_i). \tag{2.31}$$

$$\tag{2.32}$$

Here, $h = \frac{b-a}{n}$, $a_i = a + ih$, $i = 0 : n$, and $x_i = a + \left(i - \frac{1}{2}\right) h$, $i = 1 : n$.
A routine f_int(x) evaluating the function to be integrated at the point x is required. The end points a and b of the integration interval and the number of intervals n must also be specified.

Algorithm 1 Pseudocode for Midpoint Rule

```
Input:
a = left endpoint of the integration interval
b = right endpoint of the integration interval
n = number of partition interval
f_int(x) = routine evaluating f(x)

Output:
I_midpoint = Midpoint Rule approximation of ∫_a^b f(x)

h = (b − a)/n; I_midpoint = 0
for i = 1 : n do
    I_midpoint = I_midpoint + f_int(a + (i − 1/2)h)
end for
I_midpoint = h · I_midpoint
```

Algorithm 2 Pseudocode for Trapezoidal Rule

```
Input:
a = left endpoint of the integration interval
b = right endpoint of the integration interval
n = number of partition interval
f_int(x) = routine evaluating f(x)

Output:
I_trap = Trapezoidal Rule approximation of ∫_a^b f(x)

h = (b − a)/n; I_midpoint = 0
I_trap = f_int(a) / 2 + f_int(b) / 2
for i = 1 : (n − 1) do
    I_trap = I_trap + f_int(a + ih)
end for
I_trap = h · I_trap
```

In practice, we want to find an approximate value that is within a prescribed tolerance *tol* of the integral I of a given function $f(x)$ over the interval $[a, b]$. Simply using a numerical integration method with n partition intervals cannot work effectively, since we do not know in advance how large n should be chosen to obtain an approximation of I with the desired accuracy.
We choose an integration method and smaller number of intervals, e.g., 4 or 8 intervals, and compute the numerical approximation of the integral. We then double the number of intervals and compute another approximation I. If the absolute value of the difference between the new and old approximations is smaller than the required

Algorithm 3 Pseudocode for Simpson's Rule

```
Input:
a = left endpoint of the integration interval
b = right endpoint of the integration interval
n = number of partition interval
f_int(x) = routine evaluating f(x)

Output:
I_simpson = Simpson's Rule approximation of ∫_a^b f(x)

h = (b − a)/n; I_midpoint = 0
I_simpson = f_int(a) / 6 + f_int(b) / 6
for i = 1 : (n − 1) do
    I_simpson = I_simpson + f_int(a + ih) / 3
end for
for i = 1 : n do
    I_simpson = I_simpson + 2 f_int(a + (i − 1/2)h) / 3
end for
I_simpson = h · I_simpson
```

tolerance *tol*, we declare the last computed approximation of the integral to be the approximate value of I that we are looking for. Otherwise, double the number of intervals again and repeat the process until two consecutive numerical integration approximations are within the desired tolerance *tol* of each other. This condition is called the stopping criterion for the algorithm, and can be written formally as

$$|I_{new} - I_{old}| < tol, \tag{2.33}$$

where I_{old} and I_{new} are the last two approximations of I that were computed. The pseudocode for this method is given in Algorithm 4.

2.5.2 A concrete example

We want to find an approximate value for

$$I = \int_0^2 e^{-x^2} dx$$

which is within $0.5\ 10^{-7}$ of I.
We implement the algorithm from Algorithm 4 for each of the numerical integration methods to compare their convergence properties. We choose $tol = 0.5\ 10^{-7}$. For an initial partition of the interval $[0, 2]$ into $n = 4$ intervals, the following approximate values of I are found using the Midpoint, Trapezoidal, and Simpson's rules, respectively:

$$I_4^M = 0.88278895; \quad I_4^T = 0.88061863; \quad I_4^S = 0.88206551.$$

Then, we double the number of partition intervals and compute the numerical approximates corresponding to each method. We keep doubling the number of partition intervals until the stopping criterion (2.33) is satisfied. The results are recorded below:

Algorithm 4 Pseudocode for computing an approximate value of an integral with given tolerance

Input:
tol = prescribed tolerance
I_numerical(n) = result of the numerical integration rule with n intervals; any integration rule can be used

Output:
I_approx = approximation of $\int_a^b f(x)$ with tolerance tol

```
n = 4; I_old = I_numerical(n)                ▷ 4 intervals initial partition
n = 2n; I_new = I_numerical(n)
while abs(I_new - I_old) > tol do
    I_old = I_new
    n = 2n
    I_new = I_numerical(n)
end while
I_approx = I_new
```

No. Intervals	Midpoint Rule	Trapezoidal Rule	Simpson's Rule
4	0.88278895	0.88061863	0.88206551
8	0.88226870	0.88170379	0.88208040
16	0.88212887	0.88198624	0.88208133
32	0.88209330	0.88205756	0.88208139
64	0.88208437	0.88207543	
128	0.88208214	0.88207990	
256	0.88208158	0.88208102	
512	0.88208144	0.88208130	

2.6 Interest Rate Curves. Zero rates and instantaneous rates

The zero rate $r(0,t)$ between time 0 and time t is the rate of return of a cash deposit made at time 0 and maturing at time t. If specified for all values of t, then $r(0,t)$ is called the zero rate curve and is a continuous function of t.
We assume that interest is compounded continuously. The value at time t of $B(0)$ currency units is

$$B(t) = \exp(tr(0,t))B(0). \tag{2.34}$$

The value at time 0 of $B(t)$ currency units at time t is

$$B(0) = \exp(-tr(0,t))B(t). \tag{2.35}$$

The instantaneous rate $r(t)$ at time t is the rate of return of deposits made at time t and maturing at time $t + dt$, where dt is infinitesimally small, i.e.,

$$r(t) = \lim_{dt \to 0} \frac{1}{dt} \frac{B(t+dt) - B(t)}{B(t)} = \frac{B'(t)}{B(t)}.$$

We conclude that $B(t)$ satisfies the ordinary differential equation (ODE)

$$\frac{B'(\tau)}{B(\tau)} = r(\tau), \quad \forall \tau > 0, \tag{2.36}$$

with the initial condition that $B(\tau)$ at time $\tau = 0$ must be equal to $B(0)$. By integrating (2.36) between 0 and $t > 0$, it follows that

$$\int_0^t r(\tau)d\tau = \int_0^t \frac{B'(\tau)}{B(\tau)} d\tau = \ln(B(\tau))|_{\tau=0}^{\tau=t} = \ln\left(\frac{B(t)}{B(0)}\right).$$

Therefore,

$$B(t) = B(0)\exp\left(\int_0^t r(\tau)d\tau\right), \quad \forall t > 0. \tag{2.37}$$

Formula (2.37) gives the future value at time t of a cash deposit made at time $0 < t$. It can also be used to find the present value at time 0 of a cash deposit $B(t)$ made at time $t > 0$, i.e.,

$$B(0) = B(t)\exp\left(-\int_0^t r(\tau)d\tau\right), \quad \forall t > 0; \tag{2.38}$$

The term $\exp\left(-\int_0^t r(\tau)d\tau\right)$ from (2.38) is called the discount factor.
From (2.34) and (2.37), it follows that

$$r(0,t) = \frac{1}{t}\int_0^t r(\tau)d\tau. \tag{2.39}$$

The zero rate $r(0,t)$ is the average of the instantaneous rate $r(t)$ over the time interval $[0,t]$.
If $r(t)$ is continuous, then it is uniquely determined if the zero rate curve $r(0,t)$ is known. From (2.39), we obtain that

$$\int_0^t r(\tau)d\tau = tr(0,t). \tag{2.40}$$

By differentiating (2.40) with respect to t, we find that

$$r(t) = r(0,t) + t\frac{d}{dt}(r(0,t)). \tag{2.41}$$

Formulas (2.34) and (2.38) can also be written for times $t_1 < t_2$ instead of times 0 and t, as follows:

$$B(t_2) = B(t_1)\exp\left(\int_{t_1}^{t_2} r(\tau)d\tau\right), \quad \forall 0 < t_1 < t_2; \tag{2.42}$$

$$B(t_1) = B(t_2)\exp\left(-\int_{t_1}^{t_2} r(\tau)d\tau\right), \quad \forall 0 < t_1 f, t_2. \tag{2.43}$$

2.6.1 Constant interest rates

When the assumption that interest rates are constant is made, it is never mentioned whether the zero rate or the instantaneous rates are considered to be constant. The reason is that if either one of these rates is constant and equal to r, then the other rate is also constant and equal to r, i.e.,

$$r(0,t) = r, \ \forall 0 \le t \le T \quad \Leftrightarrow \quad r(t) = r, \ \forall 0 \le t \le T.$$

To see this, assume that $r(t) = r, \forall 0 \le t \le T$. From (2.39), we find that

$$r(0,t) = \frac{1}{t}\int_0^t r(\tau)d\tau = \frac{1}{t}\int_0^t r d\tau = r, \quad \forall 0 < t \le T.$$

Since $r(0,t)$ is continuous, we obtain that $r(0,0) = r$.
If $r(0,t) = r, \forall 0 \le t \le T$, recall from (2.41) that

$$r(t) = r(0,t) + t\frac{d}{dt}(r(0,t)) = r, \ \forall 0 < t < T.$$

Then $r(0) = r$ and $r(T) = r$ as well, since $r(t)$ is continuous.
If interest rates are constant and equal to r, the future value and present value formulas (2.37), (2.38), (2.42), and (2.43) become

$$B(t) = e^{rt}B(0), \quad \forall t > 0; \tag{2.44}$$

$$B(t_2) = e^{r(t_2-t_1)}B(t_1), \quad \forall 0 < t_1 < t_2; \tag{2.45}$$

$$B(0) = e^{-rt}B(t), \quad \forall t > 0; \tag{2.46}$$

$$B(t_1) = e^{-r(t_2-t_1)}B(t_2), \quad \forall 0 < t_1 < t_2. \tag{2.47}$$

2.6.2 Forward Rates

The forward rate of return $r(0; t_1, t_2)$ between times t_1 and t_2 is the constant rate of return, *as seen at time 0*, of a deposit that will be made at time $t_1 > 0$ in the future and will mature at time $t_2 > t_1$.
An arbitrage-free value for the forward rate $r(0; t_1, t_2)$ in terms of the zero rate curve $r(0,t)$ can be found using the Law of One Price as follows. Consider two different strategies for investing $B(0)$ currency units at time 0:

First Strategy: At time 0, deposit $B(0)$ currency units until time t_1, with interest rate $r(0, t_1)$. Then, at t_1, deposit the proceeds until time t_2, at the forward rate $r(0; t_1, t_2)$, which was locked in at time 0. The value of the deposit at time t_1 is $V_1(t_1) = B(0) \exp(t_1 r(0, t_1))$. At time t_2, the value is

$$\begin{aligned} V_1(t_2) &= V_1(t_1) \exp((t_2 - t_1) r(0; t_1, t_2)) \\ &= B(0) \exp(t_1 r(0, t_1) + (t_2 - t_1) r(0; t_1, t_2)). \end{aligned} \tag{2.48}$$

Second Strategy: Deposit $B(0)$ at time 0 until time t_2, with interest rate $r(0, t_2)$. At time t_2, the value of the deposit is

$$V_2(t_2) = B(0) \exp(t_2 r(0, t_2)). \tag{2.49}$$

Both investment strategies are risk free and the cash amount invested at time 0 is the same, equal to $B(0)$, for both strategies. From the Law of One Price, it follows that $V_1(t_2) = V_2(t_2)$. From (2.48) and (2.49), we find that

$$t_1 r(0, t_1) + (t_2 - t_1) r(0; t_1, t_2) = t_2 r(0, t_1).$$

By solving for $r(0; t_1, t_2)$, we conclude that

$$r(0; t_1, t_2) = \frac{t_2 r(0, t_2) - t_1 r(0, t_1)}{t_2 - t_1}. \tag{2.50}$$

2.6.3 Discretely compounded interest

Assume that interest is compounded n times every year, and let $r_n(0, t)$, for $t \geq 0$, denote the corresponding zero rate curve. Then,

$$B_n(t) = B(0) \left(1 + \frac{r_n(0, t)}{n}\right)^{nt}, \tag{2.51}$$

where $B_n(t)$ is the amount that accumulates at time t from an amount $B(0)$ at time 0 by compounding interest n times a year between 0 and t.
The most common types of discretely compounded interest are:

- annually compounded, i.e., once a year:
$$B_1(t) = B(0)(1 + r_1(0, t))^t;$$
- semiannually compounded, i.e., every six months:
$$B_2(t) = B(0) \left(1 + \frac{r_2(0, t)}{2}\right)^{2t};$$
- quarterly compounded, i.e., every three months:
$$B_4(t) = B(0) \left(1 + \frac{r_4(0, t)}{4}\right)^{4t};$$
- monthly compounded, i.e., every month:
$$B_{12}(t) = B(0) \left(1 + \frac{r_{12}(0, t)}{12}\right)^{12t}.$$

Continuously compounded interest is the limiting case of discretely compounded interest, i.e., $B_{cont}(t) = \lim_{n \to \infty} B_n(t)$. Here, $B_n(t)$ and $B_{cont}(t)$ are given by (2.51) and (2.34), respectively, i.e., $B_{cont}(t) = e^{tr(0,t)} B(0)$.

2.7 Bond Pricing. Yield of a Bond. Bond Duration and Bond Convexity

A bond is a financial instrument used to issue debt, i.e., to borrow money. The issuer of the bond receives cash when the bond is issued, and must pay the face value (principal) of the bond at a certain time in the future call the maturity (expiry) of the bond. Other payments by the issuer of the bond to the buyer of the bond, called coupon payments, may also be made at predetermined times in the future. The price of the bond is equal to the sum of all future cash flows discounted to the present by using the risk-free zero rate.

Let B be the value of a bond with future cash flows c_i to be paid to the holder of the bond at times t_i, $i = 1 : n$. Let $r(0, t_i)$ be the continuously compounded zero rates corresponding to t_i, $i = 1 : n$. Then,

$$B = \sum_{i=1}^{n} c_i e^{-r(0,t_i)t_i}. \tag{2.52}$$

Note that $e^{-r(0,t_i)t_i}$ is the discount factor corresponding to time t_i, $i = 1 : n$. If the instantaneous interest rate curve $r(t)$ is known, a formula similar to (2.52) can be given, by substituting the discount factor $\exp\left(-\int_0^{t_i} r(\tau)d\tau\right)$ for $e^{-r(0,t_i)t_i}$. Then,

$$B = \sum_{i=1}^{n} c_i \exp\left(-\int_0^{t_i} r(\tau)d\tau\right). \tag{2.53}$$

Definition 2.2 *The yield of a bond is the internal rate of return of the bond, i.e., the constant rate at which the sum of the discounted future cash flows of the bond is equal to the price of the bond. If B is the price of a bond with cash flows c_i at time t_i, $i = 1 : n$, and if y is the yield of the bond, then,*

$$B = \sum_{i=1}^{n} c_i e^{-yt_i}. \tag{2.54}$$

As expressed in (2.54), the price of the bond B can be regarded as a function of the yield. It is easy to see that the price of the bond goes down if the yield goes up, and vice versa.

To compute the yield of a bond with a known price B, we must solve (2.54) for y. This can be written as a nonlinear equation in y, i.e.,

$$f(y) = 0, \quad \text{where} \quad f(y) = \sum_{i=1}^{n} c_i e^{-yt_i} - B,$$

which is then solve numerically.

Definition 2.3 *Par yield is the coupon rate that makes the value of the bond equal to its face value.*

Duration and convexity are two of the most important parameters to estimate when investing in a bond, other than its yield. Duration provides the sensitivity of the bond

price with respect to small changes in the yield, while convexity distinguishes between two bond portfolios with the same duration. (The portfolio with higher convexity is more desirable.)
The duration of a bond is the weighted time average of the future cash flows of the bond discounted with respect to the yield of the bond, and normalized by dividing by the price of the bond.

Definition 2.4 *The duration D of a bond with price B and yield y, with cash flows c_i at time t_i, $i = 1 : n$, is*

$$D = \frac{\sum_{i=1}^{n} t_i c_i e^{-yt_i}}{B}. \tag{2.55}$$

From (2.54) and (2.55), it is easy to see that

$$\frac{\partial B}{\partial y} = -\sum_{i=1}^{n} t_i c_i e^{-yt_i} = -BD,$$

and therefore

$$D = -\frac{1}{B}\frac{\partial B}{\partial y}. \tag{2.56}$$

The duration of a bond gives the relative change in the price of a bond for *small* changes Δy in the yield of the bond (parallel shifts of the yield curve). Let ΔB be the corresponding change in the price of the bond, i.e., $\Delta B = B(y + \Delta y) - B(y)$. The discretized version of (2.56) is

$$D \approx -\frac{1}{B}\frac{B(y+\Delta y) - B(y)}{\Delta y} = -\frac{\Delta B}{B \cdot y},$$

which is equivalent to

$$\frac{\Delta B}{B} \approx -\Delta y D. \tag{2.57}$$

For very small parallel shifts in the yield curve, the approximation formula (2.57) is accurate. For larger parallel shifts, convexity is used to better capture the effect of the changes in the yield curve on the price of the bond.

Definition 2.5 *The convexity C of a bond with price B and yield y is*

$$C = \frac{1}{B}\frac{\partial^2 B}{\partial y^2}. \tag{2.58}$$

Using (2.54), it is easy to see that

$$C = \frac{\sum_{i=1}^{n} t_i^2 c_i e^{-yt_i}}{B}.$$

The following approximation of the percentage change in the price of the bond for a given change in the yield of the bond is more accurate than (2.57):

$$\frac{\Delta B}{B} \approx -D\Delta y + \frac{1}{2}C(\Delta y)^2.$$

2.7.1 Zero Coupon Bonds

A zero coupon bond is a bond that pays back the face value of the bond at maturity and has no other payments, i.e., has coupon rate equal to 0. If F is the face value of a zero coupon bond with maturity T, the bond pricing formula (2.52) becomes

$$B = Fe^{-r(0,T)T}, \tag{2.59}$$

where B is the price of the bond at time 0 and $r(0,T)$ is the zero rate corresponding to time T.
If the instantaneous interest rate curve $r(t)$ is given, the bond pricing formula (2.53) becomes

$$B = F\exp\left(-\int_0^T r(\tau)d\tau\right). \tag{2.60}$$

Let y be the yield of the bond. From (2.54), we find that

$$B = Fe^{-yT}. \tag{2.61}$$

From (2.59) and (2.61), we conclude that $y = r(0,T)$. The yield of a zero coupon bond is the same as the zero rate corresponding to the maturity of the bond. This explains why the zero rate curve $r(0,t)$ is also called the yield curve.
As expected, the duration of a zero coupon bond is equal to the maturity of the bond. From (2.56) and (2.61), we obtain that

$$D = -\frac{1}{B}\frac{\partial B}{\partial y} = -\frac{1}{Fe^{-yT}}(-TFe^{-yT}) = T.$$

The convexity of a zero coupon bond can be computed from (2.58) and (2.61):

$$C = \frac{1}{B}\frac{\partial^2 B}{\partial y^2} = \frac{1}{Fe^{-yT}}(T^2Fe^{-yT}) = T^2.$$

2.8 Numerical implementation of bond mathematics

When specifying a bond, the maturity T of the bond, as well as the cash flows c_i and the cash flows dates t_i, $i = 1 : n$ are given. The price of the bond can be obtained from formula (2.52) provided that the zero rate curve $r(0,t)$ is known for any $t > 0$, or at least for the cash flow times, i.e., for $t = t_i$, $i = 1 : n$.
If a routine r_zero(t) for computing the zero rate curve is given, the pseudocode from Algorithm 5 can be used to compute the price of the bond.
If the instantaneous interest rate curve $r(t)$ is known, the price of the bond is given by formula (2.53). If a closed formula for $\int r(\tau)d\tau$ cannot be found, evaluating the discount factors $disc(i) = \exp\left(-\int_0^{t_i} r(\tau)d\tau\right)$, $i = 1 : n$, requires estimating

$$I_i \int_0^{t_i} r(\tau)d\tau, \quad \forall i = 1 : n,$$

using numerical integration methods. This is done by setting a tolerance $tol(i)$ for the numerical approximation of I_i and doubling the number of intervals in the partition of

Algorithm 5 Pseudocode for computing the bond price given the zero rate curve

Input:
n = number of cash flows
t_cash_flow = vector of cash flow dates (of size n)
v_cash_flow = vector of cash flows (of size n)
r_zero(t) = zero rate corresponding to time t

Output:
B = bond price

$B = 0$
for $i = 1 : n$ **do**
  disc(i) = exp(-t_cash_flow(i) r_zero(t_cash_flow(i)))
  B = B = v_cash_flow(i) disc(i)
end for

Algorithm 6 Pseudocode for computing the bond price given the instantaneous interest rate curve

Input:
n = number of cash flows
t_cash_flow = vector of cash flow dates (of size n)
v_cash_flow = vector of cash flows (of size n)
r_inst(t) = instantaneous interest rate at time t
tol = vector of tolerances in the numerical approximation of discount factor integrals (of size n)

Output:
B = bond price

$B = 0$
for $i = 1 : n$ **do**
  I_numerical(i) = result of the numerical integration of r_inst(t) on the
    interval [0, t_cash_flow(i)] with tolerance tol(i)
  disc(i) = exp(-I_numerical(i) r_zero(t_cash_flow(i)))
  B = B = v_cash_flow(i) disc(i)
end for

$[0, t_i]$ until two consecutive approximations of I_i are within $tol(i)$ of each other; see the pseudocode from Algorithm 6 for more details.
From a practical standpoint, we note that the cash flow at maturity $c_n = 100\left(1 + \frac{C}{m}\right)$, is about two orders of magnitude higher than any other cash flow $c_i = 100\frac{C}{m}$, $i < n$, where m is the frequency of annual cash flows, e.g., $m = 2$ for a semiannual coupon bond. Therefore, an optimal vector of tolerances tol has the first $n - 1$ entries equal to each other, and the n-th entry two orders of magnitude smaller than the previous entry, i.e.,

$$tol = \left[\tau \ \tau ... \tau \ \frac{\tau}{100}\right].$$

2.9 Exercise

2.9.1 Question 1

Compute the integral of the function $f(x, y) = x^2 - 2y$ on the region bounded by the parabola $y = (x + 1)^2$ and the line $y = 5x - 1$.

Answer To identify the integration domain D, we note that

$$\begin{aligned}(x+1)^2 &= 5x - 1\\ x^2 - 3x + 2 &= 0\\ (x-1)(x-2) &= 0.\end{aligned}$$

So, $x = 1$ or $x = 2$ and that $(x + 1)^2 \leq 5x - 1$ for $1 < x < 2$. Therefore,

$$D = (x, y)|1 \leq x \leq 2 \text{ and } (x + 1)^2 \leq y \leq 5x - 1.$$

Then,

$$\begin{aligned}\int\int_D f(x,y)dxdy &= \int_1^2 \left(\int_{(x+1)^2}^{5x-1} (x^2 - 2y)dy\right) dx\\ &= \int_1^2 \left(x^2y - y^2\right)\Big|_{(x+1)^2}^{5x-1} dx\\ &= \int_1^2 x^2((5x-1) - (x+1)^2) - ((5x-1)^2 - (x+1)^4)dx\\ &= \int_1^2 (5x - 1 - (x+1)^2)(x^2 - (5x - 1 + (x+1)^2))dx\\ &= \int_1^2 (-x^2 + 3x - 2)(-7x)dx\\ &= \int_1^2 7x^3 - 21x^2 + 14xdx\\ &= \left(\frac{7}{4}x^4 - 7x^3 + 7x^2\right)\Big|_1^2\\ &= -\frac{7}{4}\end{aligned}$$

2.9.2 Question 2

Let $f : (0, \infty) \to \mathbb{R}$ denote the Gamma function, i.e. let

$$f(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

Part (i)

Show that $f(\alpha)$ is well defined for any $\alpha > 0$, i.e., show that both

$$\int_0^1 x^{\alpha-1} e^{-x} dx = \lim_{t \to 0} \int_t^1 x^{\alpha-1} e^{-x} dx$$

and

$$\int_1^\infty x^{\alpha-1} e^{-x} dx = \lim_{t \to \infty} \int_1^t x^{\alpha-1} e^{-x} dx$$

exist and are finite.

Answer Let $\alpha > 0$. Intuitively, as $x \searrow 0$, the function $x^{\alpha-1} e^{-x}$ is on the order of $x^{\alpha-1}$ since $\lim_{x \searrow 0} e^{-x} = 1$. Since

$$\begin{aligned} \lim_{t \searrow 0} \int_t^1 x^{\alpha-1} dx &= \lim_{t \searrow 0} \left. \frac{x^\alpha}{\alpha} \right|_t^1 \\ &= \frac{1}{\alpha} \lim_{t \searrow 0} (1 - t^\alpha) \\ &= \frac{1}{\alpha}, \end{aligned}$$

it follows that

$$\int_0^1 x^{\alpha-1} e^{-x} dx = \lim_{t \searrow 0} \int_t^1 x^{\alpha-1} e^{-x} dx$$

exists and is finite (upper bounded by $\frac{1}{\alpha}$.
Similarly, intuitively, as $x \to \infty$, the function $x^{\alpha-1} e^{-x}$ is on the order of e^{-x}, since the exponential function dominates at infinity. Since

$$\lim_{t \to \infty} \int_1^t e^{-x} dx = \lim_{t \to \infty} (1 - e^{-t}) = 1,$$

it follows that

$$\int_1^\infty x^{\alpha-1} e^{-x} dx = \lim_{t \to \infty} \int_1^t x^{\alpha-1} e^{-x} dx$$

exists and is finite.
The prove the second part rigorously, we need to prove that, for any $\epsilon > 0$, there exists $n(\epsilon) > 0$ such that

$$\int_s^\infty x^{\alpha-1} e^{-x} dx < \epsilon, \quad \forall s > n(\epsilon).$$

There exists $N > 0$ such that

$$x^{\alpha-1} e^{-x} < e^{-\frac{x}{2}}, \quad \forall x > N,$$

since

$$\lim_{x\to\infty} x^{\alpha-1}e^{-\frac{x}{2}} = 0.$$

Since $\lim_{x\to\infty} e^{-\frac{x}{2}} = 0$, it follows that, for any $\epsilon > 0$, there exists $m(\epsilon) > 0$ such that

$$2e^{-\frac{m(\epsilon}{2}} < \epsilon.$$

Choose $n(\epsilon) = \max(m(\epsilon), N)$, we obtain that

$$\begin{aligned} x^{\alpha-1}e^{-x} &< e^{-\frac{x}{2}}, \quad \forall x > n(\epsilon); \\ 2e^{-\frac{n(\epsilon)}{2}} &< \epsilon. \end{aligned}$$

For any $s > n(\epsilon)$,

$$\begin{aligned} \int_s^\infty x^{\alpha-1}e^{-x}dx &= \lim_{t\to\infty}\int_s^t x^{\alpha-1}e^{-x}dx \\ &< \lim_{t\to\infty}\int_s^t e^{-\frac{x}{2}}dx \\ &= \lim_{t\to\infty}\left(-2e^{-\frac{t}{2}} + 2e^{-\frac{s}{2}}\right) \\ &= 2e^{-\frac{s}{2}} \\ &< 2e^{-\frac{n(\epsilon)}{2}} \\ &< \epsilon, \end{aligned}$$

Part (ii)

Prove, using integration by parts, that $f(\alpha) = (\alpha-1)f(\alpha-1)$ for any $\alpha > 1$. Show that $f(1) = 1$ and conclude that, for any $n \geq 1$ positive integer, $f(n) = (n-1)!$.

Answer

$$\begin{aligned} f(\alpha) &= \int_0^\infty x^{\alpha-1}e^{-x}dx \\ &= \left(-x^{\alpha-1}e^{-x}\right)\Big|_0^\infty - \int_0^\infty -(\alpha-1)x^{\alpha-2}e^{-x}dx \\ &= (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}dx, \quad \because \lim_{x\to\infty} e^{-x} = 0 \\ &= (\alpha-1)f(\alpha-1) \\ f(1) &= \int_0^\infty e^{-x}dx \\ &= -e^{-x}\big|_0^\infty \\ &= -(0-1) \\ &= 1 \end{aligned}$$

By induction, $f(n) = (n-1)!$.

No. Intervals	Midpoint Rule	Trapezoidal Rule	Simpson's Rule
4	0.40715731	0.41075744	0.40835735
8	0.40807542	0.40895737	0.40836940
16	0.40829719	0.40851640	0.40837019
32	0.40835199	0.40840674	
64	0.40836569	0.40837937	
128	0.40836911	0.40837253	
256	0.40836996	0.40837082	
512		0.40837039	

2.9.3 Question 3

Compute an approximate value of $\int_1^3 \sqrt{x}e^{-x}dx$ using the Midpoint rule, the Trapezoidal rule, and Simpson's rule. Start with $n = 4$ intervals, and double the number of intervals until two consecutive approximations are within 10^{-6} of each other.

Answer The approximate values can be found in the table below: The approximate value of the integral is 0.408370, and is obtained for a 256 intervals partition using the Midpoint rule, for a 512 intervals partition using the Trapezoidal rule, and for a 16 intervals partition using Simpson's rule.