Lecture Notes on Financial Markets, Part 2

Claudio Albanese

Contents

CHAPTER 1

The Fundamental Theorem of Finance

1. The Fundamental Theorem of Linear Inequalities

In [Farkas, Julius (Gyula) (1902), "Über die Theorie der Einfachen Ungleichungen", Journal fr die Reine und Angewandte Mathematik 124: 127], Farkas established a result which became ubiquitous in Mathematical Economics and is known as Farkas Lemma or as the Fundamental Theorem of Linear Inequalities. This result also lies at the basis of de Finetti's Fundamental Theorem of Probability and Mathematical Finance in the next section. Here, we give a constructive proof of Farkas Lemma based on an explicit solution algorithm introduced by [Avis-Kaluzni].

Theorem 1. **(Farkas)** *Let* $\mathbf{A}$ *be a* $n \times m$ *matrix and let* $\mathbf{b}$ *be a real non-zero* n*-vector. Then either the primal system:*

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}, \qquad \text{and} \qquad \mathbf{x} \geq 0 \tag{1.1}$$

has a solution or the dual system

$$\mathbf{A}^T\mathbf{y} \geq 0, \qquad \text{and} \qquad y \cdot \mathbf{b} < 0 \tag{1.2}$$

has a solution but never both.

Proof. We begin by noticing that there cannot exist both a vector $\mathbf{x}$ and a vector $\mathbf{y}$ satisfying the conditions of the theorem. For otherwise, $0 > y^T b \geq y^T A x \geq 0$. What remains to be shown is that in case the primal system does not admit a solution, then the dual system does.

The proof is based on a iterative algorithm introduced in [Avis-Kaluzni] that solves the system:

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}, \qquad \text{and} \qquad x \geq 0. \tag{1.3}$$

The algorithm is designed in such a way to halt after a finite number of iterations and to either give rise to a solution or to allow one to conclude that a solution does not exist and that instead there is a solution for the dual problem.

It is useful to explain the algorithm by means of examples. Consider the linear system of inequalities

$$
\left\{
\begin{array}{rcl}
-x_1 - 2x_2 + x_3 & \leq & -1 \\
x_1 - 3x_2 - x_3 & \leq & 2 \\
-x_1 - 2x_2 + 2x_3 & \leq & -2 \\
x_1, x_2, x_3 & \geq & 0.
\end{array}
\right.
\tag{1.4}
$$

This can be turned into an equivalent linear system of equalities by introducing three additional variables x_4, x_5, x_6 called *slacks*:

$$
\left\{
\begin{array}{lcc}
x_4 & = & -1 + x_1 + 2x_2 - x_3 \\
x_5 & = & 2 - x_1 + 3x_2 + x_3 \\
x_6 & = & -2 + x_1 + 2x_2 - 2x_3.
\end{array}
\right.
\tag{1.5}
$$

with the positivity constraint $x_i \geq 0$ for $i = 1, ..6$. This equation provides an example of a **dictionary**, i.e. a system of linear equations where

- variables are constrained to be non-negative,
- variables are divided into left-hand-side (l.h.s.) variables and right-hand-side (r.h.s.) variables,
- each l.h.s. variable appears only in one equation and is on the left hand side with a coefficient 1,
- each r.h.s. appears only on the right hand side of the linear equations and may appear on any equation with any coefficient,
- known numeric terms appear on the right hand side of the equations.

Dictionaries can be transformed into each other by means of **pivoting operations** whereby

- one selects a r.h.s. variable,
- one uses one particular linear equation to express it in terms of the r.h.s. variables and the l.h.s. variable appearing in that equation,
- one inserts this expression into the other equations, thus obtaining a new dictionary whereby the selected r.h.s. variable becomes a l.h.s. variable, while the l.h.s. variable in the selected equation becomes a r.h.s. variable.

The solution algorithm consists in an iteration of pivoting transformations aimed at obtaining a dictionary whereby the numeric terms on the right hand side are all non-negative. If this turns out to be possible, then a solution is found by setting all r.h.s. variables to zero.

To continue with the example above, we find the smallest-indexed l.h.s. variable with a positive coefficient appearing in an equation with a negative known term on the right hand side. In this case, we are looking at x_4. In the equation for x_4, let us pivot around the r.h.s. variable with the smallest index that appears with a positive coefficient (in this case x_1. We find

$$\begin{cases} x_1 &= & 1 - 2x_2 + x_3 + x_4 \\ x_5 &= & 1 + 5x_2 - x_4 \\ x_6 &= & -1 - x_3 + x_4. \end{cases} \tag{1.6}$$

Now there is only one equation left with a negative known term on the right hand side. We repeat the procedure and choose the smallest indexed right hand variable appearing with a positive coefficient in this equation, namely x_4 and carry out a pivot operation around it. We find the new dictionary

$$\begin{cases} x_1 &= & 2 - 2x_2 + 2x_3 + x_6 \\ x_4 &= & 1 + x_3 + x_6 \\ x_5 &= & 0 + 5x_2 - x_3 - x_6. \end{cases} \tag{1.7}$$

Since the known terms on the right hand side are all positive, we conclude that a solution is given by $x_1 = 2, x_2 = 0, x_3 = 0$.

Although in this example we arrived to a solution, in general the algorithm may halt otherwise. This happens when there are equations with a negative known term and in these equations all right hand variables appear with negative coefficients. If this happens, then the above rules do not apply and the algorithm halts. In this case, one can show that a solution does not exist. To explain, consider a second example

$$\begin{cases} x_4 &= & 3 + x_1 - 2x_2 - x_3 \\ x_5 &= & -17 - 3x_1 + 2x_2 - x_3 \\ x_6 &= & 19 + x_1 + 6x_2 + 23x_3. \end{cases} \tag{1.8}$$

with the positivity constraint $x_1, x_2, x_3, x_4, x_5, x_6 \geq 0$. The algorithm proceeds as before by choosing the equation for x_5 and solving for x_2:

$$
\begin{cases}
x_2 & = & \frac{17}{2} + \frac{3}{2}x_1 + \frac{1}{2}x_3 + \frac{1}{2}x_5 \\
x_4 & = & -14 - 2x_1 - 2x_3 - x_5 \\
x_6 & = & 70 + 10x_1 + 26x_3 + 3x_5.
\end{cases}
\tag{1.9}
$$

At this point the algorithm halts as in the second equation all right hand variables appear with a negative sign and the known term is also negative. At closer inspection one realizes that this equation does not possibly have a non-negative solution. Since the problem is equivalent to the original one, that one also does not admit a solution.

Let's now discuss the general case of a system with m equations and n unknowns. We introduce m slack variables and re-write the system in dictionary form:

$$
x_{l(i)} = d_{l(i)} + \sum_{j=1}^{n} C_{ij} x_{r(j)}. \tag{1.10}
$$

Here, $x_{r(j)}, j = 1, ...n$ is the collection of right hand variables and $x_{l(i)}, i = 1, ...m$ is the collection of left hand variables. We then carry out the algorithm: if all the known terms $d_{l(i)}$ are non-negative then the algorithm halts and a solution is found by setting all the r.h.s variables to zero. Otherwise, from among the equations with a negative $d_{l(i)}$ select the one with the least index $l(i)$. In this equation, check if there are r.h.s. variables which appear with a positive coefficient. If there aren't such variables the algorithm halts and the conclusion is that no solution exists. Otherwise, from among the r.h.s. variables with positive coefficients, select the one with the least index $r(j)$ and pivot around it. Then iterate.

If the variable x_{n+m} moves from the right to the left during the cycle, based on the previous rules it ought to do so as a consequence of pivoting about it in an equation of the form:

$$
x_{l(i)} = d_{l(i)} + \sum_{j=1}^{n-1} C_{ij} x_{r(j)} + C_{i,n} x_{n+m}. \tag{1.11}
$$

We need to have $d_{l(i)} < 0$ or else the equation would not be selected. We also have $C_{i,n} > 0$ or else we would not consider the variable x_{n+m} as a candidate to pivot around. Finally, we need to have $C_{ij} \leq 0$ for all $j = 1, ..n - 1$ or else we would pick a variable of lesser index to pivot around. This equation shows that, if $x_i \geq 0$ for all $i = 1, ...n + m - 1$, then $x_{n+m} > 0$.

If the variable x_{n+m} moves from the left to the right instead, it does so while pivoting about it in an equation of the form:

$$x_{n+m} = d_{n+m} + \sum_{j=1}^{n-1} C_{ij} x_{r(j)}. \tag{1.12}$$

We need to have $d_{n+m} < 0$ or else the equation would not be selected. We also need to have $d_{l(i)} \geq 0$ for all indices i such that $l(i) \neq n+m$ or another equation would be selected. We conclude that, there is a solution of the system of linear inequalities for which $x_i \geq 0$ for all $i = 1, ...n+m-1$ and $x_{n+m} < 0$. To obtain this solution one just has to set to zero all right hand variables. Since this eventuality contradicts the conclusion we reached by assuming that x_{n+m} moves from the right to the left during the algorithm, we arrive to the conclusion that just one of the following three cases must hold:

- The variable x_{n+m} never moves from one side to the other of the equation during the course of the algorithm;
- The variable x_{n+m} moves from the left to the right just once and then never moves again to the left;
- The variable x_{n+m} moves from the right to the left just once and then never moves again to the right.

Since only a finite number of combinations of left and right hand variables is possible, either the algorithm stops after a finite number of steps or it develops a cycle. We can show that the latter case is not possible by contradiction.

As we have seen, it is not possible that over the cycle the variable x_{n+m} moves from the left to the right and then from the right to the left. One possibility is that the variable x_{n+m} sits on the right hand side throughout the cycle; in this case, one can set this variable to zero and pass to a system with one fewer variable and the same number of equations which would repeat the same cycle. The alternative possibility is that x_{n+m} sits constantly on the left hand side, then one can eliminate the corresponding equation and still obtain a cycling dictionary. Either way, we reduce to a system with either one equation fewer or one variable less. The inductive argument can be repeated until one either eliminates all equations or one eliminates all right hand variables. Either way, such a system cannot cycle. So we conclude that our algorithm halts in a finite number of steps. In the worst case scenario, the algorithm halts after having explored all possible rearrangements of left and right and variables, i.e. after at most $\begin{pmatrix} n \\ n+m \end{pmatrix}$ steps.

Cycles being impossible, there are two ways the algorithm can halt. Either it finds a solution. In this case the dual system has no solution.

A second possibility is that one arrives to an incosistent equation of the form

$$x_{l(i)} = d_{l(i)} + \sum_{j=1}^{n} C_{ij} x_{r(j)}. \tag{1.13}$$

with $d_{l(i)} < 0$ and $C_{ij} \leq 0$ for all j. In this case, we define the vector $y_k, k = 1, ...m$ so that

- $y_k = C_{ij}$ if there is a j for which $r(j) = n + k$,
- $y_k = -1$ if there is a j for which $l(i) = n + k$,
- $y_k = 0$ otherwise.

The starting system has the form

$$x_{n+k} = b_{n+k} - \sum_{j=1}^{n} A_{kj} x_j. \tag{1.14}$$

The incosistent equation can be obtained by multiplying each equation in this original system by y_k and summing over k, i.e. it is equivalent to the equation

$$\sum_{k=1}^{m} x_{n+k} y_k = \sum_{k=1}^{m} y_k b_{n+k} - \sum_{k=1}^{m} \sum_{j=1}^{n} y_k A_{kj} x_j. \tag{1.15}$$

Comparing terms, we find that

$$\sum_{k=1}^{m} y_k b_{n+k} = d(l(i)) < 0. \tag{1.16}$$

Moreover

- $-\sum_{k=1}^{m} \sum_{j=1}^{n} y_k A_{k,j} = C_{kl} <= 0$, for each l such that $r(l) = j$
- $-\sum_{k=1}^{m} \sum_{j=1}^{n} y_k A_{k,j} = -1 < 0$, if there is an l such that $l(i) = j$,
- $-\sum_{k=1}^{m} \sum_{j=1}^{n} y_k A_{k,j} = 0$, otherwise.

In conclusion, the vector $\mathbf{y}$ solves the dual problem

$$\mathbf{y} \cdot \mathbf{b} < 0, \quad \text{and} \quad A^T \mathbf{y} \geq 0. \tag{1.17}$$

$\square$

2. The First Fundamental Theorem of Finance

Consider a family of time points $t_i = t_0 + i\delta t$ where t_0 is today's date, δt is a constant step and $i = 0, 1, 2, ...$ is an integer. Consider also a discrete state space $\Lambda = \{0, ...d-1\}$ where $d \geq 1$. Let $\mathcal{P}(\Lambda)$ denote the set of all paths $\gamma = (\gamma_i)_{i=0,1,...}$ with $\gamma_i \in \Lambda$. γ_i is the state variable of path γ at time t_i.

Definition 1. *If $j \geq 0$ is a non-negative integer, a function $\phi(\gamma, t_i)$ with $\gamma \in \mathcal{P}(\Lambda), i = 0, 1, ...$ is called step-j non-anticipatory if*

$$\phi(\gamma, t_i) = \phi(\gamma', t_i) \tag{2.1}$$

for all i and whenever $\gamma_k = \gamma'_k$ for all $k \leq i + j$.

Definition 2. *A pathspace $\mathcal{P}(\Lambda, \kappa)$ is characterized by a sequence of incidence matrices taking only values 0 and 1 and given by step-1 non-anticipatory functions $\kappa(\gamma, t_i) \in \{0, 1\}$ such that if $\kappa(\gamma, t_i) = 0$ for some $i \geq 0$ then $\kappa(\gamma, t_k) = 0$ for all $k \geq i$. A path γ belongs to the set $\mathcal{P}(\Lambda, k)$ if $\kappa(\gamma, t_i) = 1$ for all $i \geq 0$.*

Definition 3. *A real valued process adapted to the pathspace $\mathcal{P}(\Lambda, \kappa)$ is given by a real valued step-0 non-anticipatory function $A(\gamma, t_i)$.*

Pricing is carried out relative to a valuation benchmark, also called numeraire.

Definition 4. *A numeraire is a positive valued adapted process $g(\gamma, t_i) > 0$.*

An example of a numeraire is given by the price of a commodity with negligible carry costs and negligible convenience yield such as, for instance, gold. A second example of a numeraire is defined through a positive valued process $r(\gamma, t_i)$ interpreted as a short rate, i.e. the money market account process given by

$$B(\gamma, t_i) = (1 + \delta t r(\gamma_0, t_0)) \quad ... \quad (1 + \delta t r(\gamma_{i-1}, t_{i-1})). \tag{2.2}$$

Definition 5. *Let $\mathcal{P}(\Lambda, k)$ be a pathspace characterized by the incidence matrices $\kappa(\gamma, t_i)$ and let $g(\gamma, t_i)$ be a numeraire process. Let $A^s(\gamma, t_i)$ be a family of non-anticipatory path functionals indexed by $s = 1, ...M$ with $M > 0$ on the time interval $t_i \in [t_0, t_j]$ where $j > 0$. The family of processes $A^s(\gamma, t_i)$ is called g-coherent if for any $t_i \in [t_0, t_j]$ and any set of coefficients $\zeta^s, s = 1, ...M$, the* *following property holds: if γ is a possible path for which $\kappa(\gamma, t_i) = 1$ and*

$$\frac{1}{g^s(\gamma, t_{i+1})} \sum_s \zeta^s A^s(\gamma, t_{i+1}) - \frac{1}{g^s(\gamma, t_i)} \sum_s \zeta^s A^s(\gamma, t_i) > 0 \tag{2.3}$$

then there is a second possible path γ' such that $\gamma'_k = \gamma_k$ for all $k \le i$ and $\kappa(\gamma', t_i) = 1$ and

$$\frac{1}{g^s(\gamma\prime, t_{i+1})} \sum_s \zeta^s A^s(\gamma\prime, t_{i+1}) - \frac{1}{g^s(\gamma, t_i)} \sum_s \zeta^s A^s(\gamma, t_i) < 0. \tag{2.4}$$

Definition 6. *An elementary transition probability kernel on $\mathcal{P}(\Lambda, k)$ is a family of 1-step non-anticipatory path functionals $q(\gamma, t_i)$ defined for $i = 0, 1, .., \gamma \in \mathcal{P}(\Lambda, k)$ satisfying the following properties:*

(i) $q(\gamma, t_i) \ge 0$,

(ii) $\sum_{\gamma' : \gamma'_k = \gamma_k, k \le i} q(\gamma', t_i) = 1$,

(iii) $q(\gamma, t_i) > 0$ *if and only if* $\kappa(\gamma, t_i) = 1$*, while otherwise* $q(\gamma, t_i) = 0$.

Definition 7. *Let us fix a time $t_j > t_0$ and let $A(\gamma, t_i), i = 0, ...j$ be a process. The expectation of this process at time t_i prior to t_j with respect to the kernel $q(\gamma, t_i)$ is a process denoted by $E^q_{t_i}[A(\gamma, t_j) \mid \{\gamma_k\}_{k \le i}]$ and defined as follows:*

$$E^q_{t_i}[A(\gamma, t_j) | \{\gamma_k\}_{k \le i}] = \sum_{\gamma' : \gamma'_k = \gamma_k \forall k \le i} q(\gamma', t_i) \; .... \; q(\gamma', t_{j-1}) A(\gamma', t_j). \tag{2.5}$$

This formula can be recast in terms of conditional probabilities of a path given by the functional

$$p(\gamma, t_i, t_j) \equiv q(\gamma, t_i) \; .... \; q(\gamma, t_{j-1}). \tag{2.6}$$

In terms of this functional, we can express expectations in (2.5) as follows:

$$E^q_{t_i}[A(\gamma, t_j) | \{\gamma_k\}_{k \le i}] = \sum_{\gamma' : \gamma'_k = \gamma_k \forall k \le i} p(\gamma', t_i, t_j) A(\gamma', t_j). \tag{2.7}$$

The interpretation of $p(\gamma, t_i, t_j)$ is that this is the probability for a path to be equal to γ in the time interval $[t_{i+1}, t_j]$ conditioned to knowing that it coincides with γ on the preceding time interval $[t_0, t_i]$.

Definition 8. *Let $g(\gamma, t_i)$ be a numeraire asset, the adapted process $A(\gamma, t_i)$ is called g-discounted martingale with respect to the elementary kernel $q(\gamma, t_i)$ if*

$$A(\gamma, t_i) = E^q_{t_i} \left[\frac{g(\gamma, t_i)}{g(\gamma, t_j)} A(\gamma, t_j) \middle| \{\gamma_k\}_{k \le i} \right]. \tag{2.8}$$

where E^q denotes the expectation with respect to the elementary transition probability kernels $q(\gamma, t_i)$.

As an example, consider the case where the numeraire is given by the money market account $B(\gamma, t_i)$ in (2.2). Fix a time $t_j > t_0$ and let $A(\gamma, t_j)$ be a process. The discounted expectation of this process at time t_i prior to t_j with respect to the money market account $B(\gamma, t_i)$ can be expressed as follows:

$$E^q_{t_i}\left[\frac{B(\gamma,t_i)}{B(\gamma,t_j)}A(\gamma,t_j)\;\middle|\;\{\gamma_k\}_{k\le i}\right] = \sum_{\gamma':\gamma'_k=\gamma_k\ \forall k\le i} q(\gamma',t_i)\quad ....\quad q(\gamma',t_{j-1})\ \frac{B(\gamma',t_i)}{B(\gamma',t_j)}A(\gamma',t_j). \tag{2.9}$$

Elementary discounted transition probability kernels defined as

$$q^D(\gamma,t_i) = \frac{1}{1+\delta tr\,(\gamma,t_i)}q(\gamma,t_i) \tag{2.10}$$

can be used to implicitly account for the numeraire asset in the path expansion for discounted expectations by recasting it as follows:

$$E^q_{t_i}\left[\frac{B(\gamma,t_i)}{B(\gamma,t_j)}A(\gamma,t_j)\middle|\{\gamma_k\}_{k\le i}\right] = \sum_{\gamma':\gamma'_k=\gamma_k\ \forall k\le i} q^D(\gamma',t_i)\quad ....\quad q^D(\gamma',t_{j-1})\ A(\gamma',t_j). \tag{2.11}$$

Definition 9. *The family of adapted processes $A^s(\gamma, t_i), s = 1, ..M$ is called a family of g-discounted equivalent martingales if there exists an elementary kernel $q(\gamma, t_i)$ on the path-space $\mathcal{P}(\Lambda, k)$ with respect to which the processes $A^s(\gamma, t_i)$, $s = 1, ..M$, are all g-discounted martingales.*

Theorem 2. ***(First Fundamental Theorem of Finance)*** *Let $g(\gamma, t_i)$ be a numeraire process and let $A^s(\gamma, t_i), s = 1, ...M$ be family of processes adapted to $\mathcal{P}(\Lambda, k)$ and defined on the time interval $t_i \in [t_0, t_j]$. Then $A^s, s = 1, ...M$ is a family of equivalent g-discounted martingales if and only if they are a g-coherent family.*

Proof. This theorem was first stated and proved by Bruno de Finetti in (**?**). The proof depends on Farkas Lemma in the previous section and goes as follows.

Assume that $A^s(\gamma, t_i)$ is a g-coherent family of processes. Let $i \geq 0$, $t_i \in [t_0, t_j]$. We need to show that for all $i \geq 0$, there exists a transition probability kernel $q(\gamma, t_i)$ such that

$$q(\gamma,t_i) > 0 \text{ if and only if }\ k(\gamma,t_i) = 1, \tag{2.12}$$

with respect to which the processes $A^s(\gamma, t_i)$ are discounted martingales, i.e. are such that

$$\sum_{\gamma':\gamma'_k=\gamma_k\ \forall k\le i} q(\gamma', t_i) A^s(\gamma', t_{i+1}) \frac{g(\gamma', t_i)}{g(\gamma', t_{i+1})} = A^s(\gamma, t_i) \tag{2.13}$$

for all $s = 1, ...M$, all $i = 0, ..j$ and all $\gamma \in \mathcal{P}(\Lambda, k)$. This is sufficient as, by iterating this equation one arrives to the discounted martingale condition over arbitrarily long time intervals. It is convenient to recast this last equation (2.13) as follows:

$$\sum_{\gamma':\gamma'_k=\gamma_k\ \forall k\le i} q(\gamma', t_i) \left(A^s(\gamma', t_{i+1}) \frac{g(\gamma', t_i)}{g(\gamma', t_{i+1})} - A^s(\gamma, t_i) \right) = 0. \tag{2.14}$$

Let γ^abe a family of paths where $a = 1, \dots, n$ with the following properties:

(i) they coincide for $t \le t_i$, i.e. $\gamma^a_k = \gamma^b_k$ for all $a, b = 1, \dots, n$ and all $k \le i$;
(ii) they differ at time t_{i+1}, i.e. $\gamma^a_{i+1} \neq \gamma^b_{i+1}$ if $a \neq b$.
(iii) For all $a = 1, \dots, n$ we have that $k(\gamma, t_i) = 1$;
(iv) If y satisfies $k(\gamma, t_i) = 1$ then there is a path γ^a in the family such that $y = \gamma^a_{i+1}$.

Let

$$w^{as} = A^s(\gamma^a, t_{i+1}) \frac{g(\gamma^a, t_i)}{g(\gamma^a, t_{i+1})} - A^s(\gamma^a, t_i). \tag{2.15}$$

Let $\mathcal{V}$ be a n-dimensional vector space with basis vectors $\mathbf{v}^1, ..., \mathbf{v}^n$. Let $\mathbf{w}^s$ be the vector in $\mathcal{V}$ of components

$$\mathbf{w}^s = \sum_{a=1}^{n} w^{as} \mathbf{v}^a. \tag{2.16}$$

It suffices to show that there is a vector $\mathbf{q} = (q^a)$ such that $q^a \ge 0$, $\forall a = 1, \dots, n$, $\sum_{a=1}^n q^a = 1$ and

$$\mathbf{q} \cdot \mathbf{w}^s \equiv \sum_{a=1}^{n} q^a w^{as} = 0 \qquad \forall s = 1, \dots, M. \tag{2.17}$$

The coherence hypothesis can be recast in this language as the statement according to which a vector (ζ^s) satisfies

$$\sum_{s=1}^{M} \zeta^s w^{as} \ge 0, \quad \forall a = 1, ..., n \tag{2.18}$$

only if it is the zero vector, i.e. $\zeta^s = 0$.

The system

$$\begin{cases} \sum_{a=1}^{n} q^a = 1 \\ \sum_{a=1}^{n} q^a w^{as} = 0 \\ q^a \geq 0, \end{cases} \tag{2.19}$$

can be recast as a primal system of linear inequalities in the standard form $Aq = c$, $q \geq 0$, where

$$\mathbf{A} = \begin{pmatrix} 1 \ 1 \ .... \ 1 \\ \\ W^T \end{pmatrix} \quad \text{and} \quad c = \begin{pmatrix} 1 \\ 0 \\ ... \\ 0 \end{pmatrix} \tag{2.20}$$

The corresponding dual system is

$$\xi_{-1} + \sum_{s=1}^{M} w^{as} \xi^s \geq 0, \quad \xi_{-1} < 0, \tag{2.21}$$

i.e. $\sum_{s=1}^{M} w^{as} \xi^s > 0$. Assuming coherence, the dual system does not have a solution. Hence the primal equation does and this establishes one direction of the theorem. Vice versa, thanks again to Farkas Lemma, if the dual has a solution, i.e. there is no g-coherence, then the processes are not g-discounted martingales. □

3. Cash Flow Streams and Dynamic Trading Strategies

Consider a finite lattice Λ and a finite time-step δt_n. Let $\mathcal{P}(\Lambda, k)$ be a pathspace characterized by the incidence matrices $\kappa(\gamma, t_i)$ and let $g(\gamma, t_i)$ be a numeraire asset price process. Let $A^s(\gamma, t_i)$ be a family of g-coherent non-anticipatory path functionals indexed by $s = 1, ...M$ with $M > 0$ on the time interval $t_i \in [t_0, t_j]$ where $j > 0$. Let Q be any probability measure with respect to which the processes A_n^s are g-discounted martingales. The existence of such a measure is assured by the Fundamental Theorem of Finance in the previous section.

Definition 10. *A* **contingent cash flow stream** *is given by a non-anticipatory path functional $c(\gamma, t)$ and describes unit of accounts payable by one party to another at time t in case the path γ is realized up to time t.*

A special case of a contingent cash flow stream is given by a single payment equal to one unit of account at a fixed maturity date T. A contract that entitles to a payment of amount $c(\gamma, T)$ at time T and zero at all other times is worth $c(\gamma, T)$ units at time T. Hence, applying the Fundamental Theorem, the worth of this contract at times prior to T is given by

$$A(\gamma, t; T) = E_t^q \left[\frac{g(\gamma, t_i)}{g(\gamma, T)} c(\gamma, T) \;\middle|\; \{\gamma_k\}_{k \le i} \right] = \sum_{\gamma' : \gamma'_k = \gamma_k \ \forall k \le i} q(\gamma', t_i) \quad .... \quad q(\gamma', T - \delta t) \, \frac{g(\gamma', t_i)}{g(\gamma', T)} c(\gamma', T). \tag{3.1}$$

A contract which entitles to a generic cash flow stream $c(\gamma, t')$ at all times $t' \ge t$ is equivalent to the portfolio of all elementary contracts which entitle to a cash flow of the same amount $c(\gamma, t')$ at one single date in the future. In fact, the difference between the cash flow stream contract and the corresponding portfolio entitles to no cash flows at any time and is thus worth zero. Hence, the price of a cash flow stream $c(\gamma, t')$ for $t' > t$ is given by

$$E_t^q \left[\sum_{t' > t} \frac{g(\gamma, t)}{g(\gamma, t')} c(\gamma, t') \;\middle|\; \{\gamma_k\}_{k \le i} \right]. \tag{3.2}$$

Definition 11. *A* **dynamic trading strategy** *is given by a family of M non-anticipatory path functionals $\zeta^s(\gamma, t)$ expressing positions in the assets $A^s(\gamma, t_i)$.*

Notice that in general a trading strategy will be costly to implement as transactions are required to rebalance a position when needed. The rebalancing cost at time t is given by

$$\sum_{s=1}^{M} (\zeta^s(\gamma, t) - \zeta^s(\gamma, t - \delta t)) A^s(\gamma, t) \tag{3.3}$$

Hence, the cumulative cost at time t to implement a trading strategy $\zeta^s(\gamma, t)$, including all past incurred costs and the worth of future ones, is given by

$$\begin{aligned} D(\gamma, t; \zeta) \quad &= E_t^q \left[\textstyle\sum_{t' > t} \frac{g(\gamma, t)}{g(\gamma, t')} (\zeta^s(\gamma, t') - \zeta^s(\gamma, t' - \delta t)) A^s(\gamma, t') \;\middle|\; \{\gamma_{t''}\}_{t'' \le t} \right] \\ &\quad + \textstyle\sum_{t' \le t} \frac{g(\gamma, t)}{g(\gamma, t')} (\zeta^s(\gamma, t') - \zeta^s(\gamma, t' - \delta t)) A^s(\gamma, t'). \end{aligned} \tag{3.4}$$

(3.5)

Notice that the cost process of a trading strategy $D(\gamma, t; \zeta)$ is itself a non-anticipatory functional which, in virtue of the Fundamental Theorem, is coherent with respect to the base asset price processes $A^s(\gamma, t)$.

DEFINITION 12. *A dynamic trading strategy is called* **self-financing** *if*

$$\sum_{s=1}^{M} (\zeta^s(\gamma, t + \delta t) - \zeta^s(\gamma, t)) A^s(\gamma, t + \delta t) = 0 \tag{3.6}$$

for all paths γ and all times t. The financial meaning of this condition is that the cost for updating a position according to the given strategy from a generic time t to the following instant $t + \delta t$ is zero.

THEOREM 3. *If $\zeta^s(\gamma, t)$ is a self-financing trading strategy, then if one adds the portfolio value process*

$$\Pi(\gamma, t) \equiv \sum_{s=1}^{M} \zeta^s(\gamma, t) A^s(\gamma, t) \tag{3.7}$$

to the family of asset price processes $A^s(\gamma, t)$, the extended family is still g-coherent.

PROOF. Consider the extended family of $M + 1$ assets where $A^{M+1}(\gamma, t) = \Pi(\gamma, t)$ and consider the elementary time step from time t to time $t + \delta t$. Let $\xi_s, s = 1, ...M + 1$ be a position vector and assume that it is possible to have

$$\frac{1}{g^s(\gamma, t_{i+1})} \sum_{s=1}^{M+1} \xi^s A^s(\gamma, t_{i+1}) - \frac{1}{g^s(\gamma, t_i)} \sum_{s=1}^{M+1} \xi^s A^s(\gamma, t_i) > 0. \tag{3.8}$$

Since we are assuming self-financing, we have that

$$A^{M+1}(\gamma, t + \delta t) = \Pi(\gamma, t + \delta t) = \sum_{s=1}^{M} \zeta^s(\gamma, t + \delta t) A^s(\gamma, t + \delta t) = \sum_{s=1}^{M} \zeta^s(\gamma, t) A^s(\gamma, t + \delta t). \tag{3.9}$$

We also have that

$$A^{M+1}(\gamma, t) = \sum_{s=1}^{M} \zeta^s(\gamma, t) A^s(\gamma, t). \tag{3.10}$$

Inserting these two equations into (3.8) we find that this reduces to the following condition:

$$\frac{1}{g^s(\gamma, t_{i+1})} \sum_{s=1}^{M} (\xi^s + \zeta^s(\gamma, t)) A^s(\gamma, t_{i+1}) - \frac{1}{g^s(\gamma, t_i)} \sum_{s=1}^{M} (\xi^s + \zeta^s(\gamma, t)) A^s(\gamma, t_i) > 0. \tag{3.11}$$

Due to the assumed coherence of the asset price processes $A^s(\gamma, t_i)$, we conclude that also the extended set of asset price processes is coherent. □

Definition 13. *A family of asset price processes $A^s(\gamma, t_i)$ is called complete if it contains also cost processes of all trading strategies and the value process of all self-financing trading strategies.*

4. Futures and Forward Price Processes

A futures price process is an non-anticipatory functional $f(\gamma, t)$ such that, at any time t, the value at time t of the cash flow at time $t + \delta t$ of amount

$$f(\gamma, t + \delta t) - f(\gamma, t) \tag{4.1}$$

is zero.

Examples of futures' price processes are given by contracts with a set final future's price at a maturity T given by the value of an asset price $A(\gamma, t)$ at time T.

Due to the Fundamental Theorem, there exists a pricing measure Q for which

$$E_{T-\delta t}[A(\gamma, T) - f(\gamma, t - \delta t)] = 0. \tag{4.2}$$

Hence

$$f(\gamma, t - \delta t)] = E_{T-\delta t}[A(\gamma, T)]. \tag{4.3}$$

We also have

$$f(\gamma, t - 2\delta t)] = E_{T-2\delta t}[f(\gamma, T - \delta t)] = E_{T-2\delta t}[A(\gamma, T)]. \tag{4.4}$$

By iterating this equation backward in time, we find that

$$F(\gamma, t)] = E_t[A(\gamma, T). \tag{4.5}$$

Futures price processes are used in futures contracts. A futures contract is written with respect to a futures price process. Two parties, a long and a short party, and is stipulated in such a way that at all times after inception t, the short party in the contract pays the amount $F(\gamma, t) - F(\gamma, t - \delta t)$ to the long party of the contract. With this arrangement, the contract is always worth zero.

If $A(\gamma, t)$ is an asset price process and T a fixed maturity, the corresponding **forward price process** is defined as follows:

$$F(\gamma, t) = \frac{A(\gamma, t)}{Z_t(\gamma, T)}, \tag{4.6}$$

where

$$Z_t(\gamma, T) = E_t\left[\frac{g(\gamma,t)}{g(\gamma,T)}\right] \tag{4.7}$$

is the price process for a zero coupon bond, i.e. the asset paying one unit of account at maturity T.

5. The Stop-Loss-Start-Gain Strategy

In this section, we introduce a limit notion of arbitrage freedom in the continuous space and continuous time limit.

Let $h_n, n = 0, 1, ...$ be a sequence of lattice discretization mesh parameters and let us consider a lattice of the form $\Lambda_n = h_n\mathbb{Z}^D \cap (0,1)$ where $D \geq 1$ is a dimension. We assume that $h_n \to 0$ as $n \to \infty$. We assume that also the corresponding time steps $\delta t_n \to 0$ as $n \to \infty$.

For each element n in the sequence of lattices, let $\mathcal{P}(\Lambda_n, k)$ be a path-space characterized by the incidence matrices $\kappa_n(\gamma, t_i)$ and let $g_n(\gamma, t_i)$ be a numeraire process. Let $A^s_n(\gamma, t_i)$ be a family of g-coherent non-anticipatory path functionals indexed by $s = 1, ...M$ with $M > 0$ on the time interval $t_i \in [t_0, t_j]$ where $j > 0$. Let Q_n is any probability measure with respect to which the processes A^s_n are g_n-discounted martingales. The existence of such a measure is assured by the Fundamental Theorem of Finance in the previous section.

Consider a situation with an asset price process given by the non-anticipatory functional $S_n(\gamma, t)$. Let us fix a strike price K and a maturity T. Let $F_n(\gamma, t; T)$ be the forward price process corresponding to $S_n(\gamma, t)$ and of maturity T.

Assume completeness and let $C_n(\gamma, t; K, T)$ be the price process for the call option contract with payoff $(S_n(\gamma, T) - K)_+$ at maturity T. According to the fundamental theorem, we have

$$C_n(\gamma, t; K, T) = E^{Q_n}\left[\frac{g(\gamma', t)}{g(\gamma', T)}(S_n(\gamma', T) - K)_+ \ \ |\gamma'_{t'} = \gamma_t \ \ \forall\ t' \leq t\right] \tag{5.1}$$

Assume that

$$\liminf_{n\to\infty} C_n(\gamma, t; K, T) > 0; \tag{5.2}$$

Let's specialize further to the case where possible paths are continuous in the sense that incidence matrices satisfy the following constraint:

$$\kappa_n(\gamma, t_i) = 0 \quad if \quad |\gamma_{t_i} - \gamma_{t_{i+1}}| \geq 2; \tag{5.3}$$

Namely, a path can only hop from one value to the nearest neighbor at any given time. In this case, a call payoff can also be replicated by means of the trading strategy according to which whenever $F_n(\gamma, t; T) > K$ one holds a long position in the stock at time t and a short position $KZ_t(T)$ in a zero coupon bond maturing at time T . In case $F_n(\gamma, t; T) <= K$ instead, according to this strategy one holds nothing. The strategy obviously replicates the call payoff and it is costly, i.e. it is not self-financing. The cost of the strategy is

$$E_t\left[\sum_{t'=t+\delta t_n,\dots T} |S_n(\gamma,t') - S_n(\gamma,t'-\delta t)| 1(F_n(\gamma,t;T) > K) 1(F_n(\gamma,t-\delta t;T) \le K)\right]. \tag{5.4}$$

Assuming absence of arbitrage, the cost to implement this strategy should match the call price as given in (5.1). As we shall see below, this consideration places severe limitations on the underlying process.

Theorem 4. *Assume that*

- *for all times $t' > t$ we have that*

$$\begin{aligned} 0 &< \liminf_{n\to\infty} \tfrac{1}{h_n} E_t^{Q_n}\Big[\delta(F_n(\gamma,t;T) > K)\delta(F_n(\gamma,t-\delta t;T) \le K)\Big] \\ &< \limsup_{n\to\infty} \tfrac{1}{h_n} E_t^{Q_n}\Big[\delta(F_n(\gamma,t;T) > K)\delta(F_n(\gamma,t-\delta t;T) \le K)\Big] < \infty \end{aligned} \tag{5.5}$$

- *for all possible paths γ and all times $t' > t$ we have that:*

$$0 < \liminf_{n\to\infty} \frac{1}{h_n}(\bar S_n(\gamma,t') - \bar S_n(\gamma,t'-\delta t)) < \limsup_{n\to\infty} \frac{1}{h_n}(\bar S_n(\gamma,t') - \bar S_n(\gamma,t'-\delta t)) < \infty \tag{5.6}$$

for some constant $c > 0$.

Then

$$0 < \liminf_{n\to\infty} \frac{\delta t_n}{T h_n^2} < \limsup_{n\to\infty} \frac{\delta t_n}{T h_n^2} < \infty. \tag{5.7}$$

Proof. Due to the hypothesis above, the cost of the trading strategy is

$$E_t\left[\sum_{t'=t+\delta t_n,\dots T} (S_n(\gamma,t') - S_n(\gamma,t'-\delta t))\delta(F_n(\gamma,t;T) > K)\delta(F_n(\gamma,t-\delta t;T) \le K)\right] = O\left(\frac{h_n^2 T}{\delta t_n}\right) \tag{5.8}$$

However, due to the hypothesis of absence of arbitrage, this limit should converge in the limit as $n \to \infty$ to the call price $C_n(\gamma, t; K, T)$ which, by assumption, is finite and positive. This is not possible if the ratio $\frac{\delta t_n}{Th_n^2}$ can either be arbitrarily small or arbitrarily large in the limit as $n \to \infty$. Hence the conclusion. $\square$

The condition in the Theorem can be rephrased in terms of the so called Hurst exponent H defined so that

$$E_t\left[|S_n(\gamma, t') - S_n(\gamma, t' - \delta t))|\right] = O\left(\delta t_n^H\right). \tag{5.9}$$

The conclusion is that the Hurst exponent needs to be equal to 1/2 in order to achieve arbitrage freedom in a scalable fashion, i.e. also along a sequence $h_n \to 0$ as $n \to \infty$.

In the next sections, we discuss Markovian diffusion processes with Hurst exponent 1/2 and show that they can indeed be taken as the basis for arbitrage free pricing models with continuous paths. We then discuss fractional Brownian motion and apply the theorem in this section to show that they allow for arbitrage trading strategies. Fractional Brownian motions are appealing empirically as they admit fat tailed and auto-correlated return distributions. However, lack of arbitrage freedom makes it impossible to use them for valuation theory. This raises a problem as empirically one observes fat tailed returns which indicate that historical asset price processes are not diffusions. In the concluding section in this chapter we discuss Markovian jump processes which admit fat tailed distribution. We also discuss stochastic volatility models which allow one to model auto-correlations albeit in a arbitrage-free way.

6. Problems

Problem 1. Consider a situation with two time points $t_0 = 0 < t_1 = 1$ and three assets: a money market account $B(\gamma, t)$, a stock of price $S(\gamma, t)$ and a call option on the stock of strike K, maturity t_1 and price process $C(\gamma, t)$.

Suppose that the yearly interest rate using simple compounding with yearly period is $r = 1\%$, that the stock price at time 0 is \$100 and that $K = \$100$.

- Consider a situation where the stock price can take the values \$95 and \$105 at time t_1.
- Under what condition on the price of the call option at time 0 is there no arbitrage?
- Suppose that the call is worth \$1. Find a pricing measure satisfying the fundamental theorem.
- Answer the same two questions above in case the stock price at time t_1 can possibly be \$95, \$100 and \$105.

- Same question as above in case the stock price can be at time t_1 can possibly be \$94, \$98, \$102 and \$106.

PROBLEM 2. Consider a stock worth S_0 at time 0 and let T be a maturity in the future. By using the Fundamental Theorem of Finance, assuming that interest rates are deterministic and selecting the money market account as the numeraire asset, what is the expected value of the stock price at time T, i.e. $E_0[S_T]$?

How does the answer change if interest rates are stochastic?

PROBLEM 3. Consider again the example in Chapter 1 of a perpetual double barrier option and assume zero interest rates. Let L and U be the lower and upper barriers, respectively and let S_0 be the spot price, assuming it is in the interval $[L, U]$.

By using the Fundamental Theorem of Finance and selecting the money market account as the numeraire asset, infer the probability that either the upper or the lower barrier is the one hit first.

PROBLEM 4. Consider two contracts, a forward and a futures contract on a zero coupon bond starting in the future.

- For what choice of numeraire is the forward price equal to the expected price of the zero coupon bond?
- For what choice of numeraire is the futures price equal to the expected price of the zero coupon bond?
- How do the forward and futures prices differ from each other?

PROBLEM 5. Consider a stock of spot price S_0 and suppose one can trade all call and put options on this stock of maturity T and of strikes comprised in the interval $[K_1, K_2]$. Suppose that interest rates are deterministic and choose the money market account as numeraire. How would you infer the probability that the stock price fall in the interval $[K_1, K_2]$ based on market prices for calls and puts?

PROBLEM 6. Consider a stock of spot price S_0 and suppose one can trade all call and put options on this stock of maturity T. Suppose that interest rates are deterministic and choose the money market account as numeraire.

How would you infer the expectation of realized variance of the stock in the time interval $[0, T]$?

PROBLEM 7. How would you replicate a futures contract on a foreign currency?

PROBLEM 8. How can one use a trading strategy in futures to replicate a variance swap?

PROBLEM 9. On December 3rd 2008, at the Comex in New York, December gold futures were in backwardation: December 31 deliveries were quoted at 2% discount to spot, while gold futures with delivery February 27 were quoted at 0.29% discount to spot. (All percentages annualized.) How to interpret these quotes?

PROBLEM 10. A bank sells forward contracts on oil with delivery in one year to two different counter-parties. The contracts are otherwise identical but the forward price to one counter-party is \$45 per barrel, while the other counterparty pays \$45.5 per barrel. What reasons can motivate the difference? Of how much would the futures prices on the same underlying differ among the two counter-parties?

PROBLEM 11. Metallgesellshaft used to be a major firm that would sell forward contracts for oil delivery to clients hedging them with futures contracts. In December 1993 the strategy gave rise to mark-to-market losses exceeding 1.5 billions. How could this hedging strategy go so wrong?

PROBLEM 12. Formulate a stop-loss-start-gain strategy to dynamically hedge put options.

PROBLEM 13. Formulate a stop-loss-start-gain strategy to dynamically hedge a variance swap instead of replicating the final log payoff with a static position in call and put options.

CHAPTER 2

Pricing Models

1. Markov Processes

Markov processes are a particularly important special class of processes characterized by the fact that transition probability kernels are independent of the past values attained by the path γ. More precisely, elementary transition probability kernels for a Markov process have the special form

$$q(\gamma, t_i) = u_{\delta t}(\gamma_i, \gamma_{i+1}; t_i) \tag{1.1}$$

where $u_{\delta t}(y_1, y_2; t_i)$ is a function of $y_1, y_2 \in \Lambda$ and time t_i. The Markov generator or Markovian is given by the matrix $\mathcal{L}(y_1, y_2; t_i)$ such that

$$u_{\delta t}(y_1, y_2; t_i) = \delta_{y_1,y_2} + \delta t \mathcal{L}(y_1, y_2; t_i) \tag{1.2}$$

for all $y_1, y_2 \in \Lambda$ and all t_i, $i = 0, 1, \ldots$. Here

$$\delta_{y_1,y_2} = \begin{cases} 1 & \text{if } y_1 = y_2 \\ 0 & \text{if } y_1 \neq y_2 \end{cases} \tag{1.3}$$

is the so called Kronecker Delta. The constraints of positivity and probability conservation imply the following two conditions on the matrix $\mathcal{L}(y_1, y_2; t_i)$:

(i) $\mathcal{L}(y_1, y_2; t_i) \geq 0$ for all $y_1 \neq y_2$ and all t_i;

(ii) $\sum_{y_2} \mathcal{L}(y_1, y_2; t_i) = 0$ for all y_1 and all t_i.

These two conditions are necessary but not sufficient as, due to condition (i) and (ii), the diagonal matrix elements

$$\mathcal{L}(y_1, y_1; t_i) = -\sum_{y_2 \neq y_1} \mathcal{L}(y_1, y_2; t_i) \leq 0 \tag{1.4}$$

are non-positive. To ensure positivity of the diagonal elements of the elementary kernel, we thus need to postulate the following third property called the Courant condition:

(iii) $\delta t \le \frac{1}{\max_y |\mathcal{L}(y,y;t_i)|}$.

Definition 14. *A matrix $\mathcal{L}(y_1, y_2; t_i)$ satisfying the properties (i), (ii) and (iii) above is called Markov matrix or Markovian.*

In practice, one builds elementary transition probability kernels starting from a Markovian, given which one finds the elementary time interval $\delta t > 0$ (typically one day or a fraction of a day) satisfying the Courant condition.

As a matter of terminology, we distinguish between operators and kernels. A kernel is a matrix $A(x, y)$ with indices x, y taking up a finite set of values. In the previous example $x, y = 0, ...d - 1$. A vector is instead represented by an array $v(x)$. The matrix $A(x, y)$ can also be put in relation with an operator A that transforms vectors linearly, so that

$$(Av)(x) = \sum_y A(x, y)v(y). \tag{1.5}$$

Vice versa, to each operator that transforms vectors linearly there corresponds a matrix. In fact, if one considers the vector $\delta_y(x) = \delta(x - y)$, one finds

$$(A\delta_y)(x) = A(x, y). \tag{1.6}$$

Given an operator A, the matrix $A(x, y)$ is called the *kernel of A*.

A *transition probability kernel* over finite time intervals is given by a two-parameter family of matrices $u(y_1, t_1; y_2, t_2)$ dependent on the time coordinates $t_1 \le t_2$ and representing the transition probabilities from state y_1 at time t_1 to state y_2 at time t_2. For fixed t_1 and t_2, these are the transition probability kernels. The operator corresponding to a transition probability kernel is called *propagator*.

The incidence matrix κ is characterized by the set of all admissible set $\mathcal{P}(\Lambda, \kappa)$. In turn, admissible paths are characterized by the condition

$$u(\gamma_i, t_i; \gamma_{i+1}, t_{i+1}) > 0. \tag{1.7}$$

being satisfied for all $i \ge 0$.

Given the family of elementary transition probability kernels $u_{\delta t}(y_1, y_2; t_i)$ for a Markov process, one can compute more general transition probability kernels $u(y_1, t_i; y_2, t_j)$ over arbitrary time intervals $[t_i, t_j]$ with $i < j$. In the particular case of a time step twice the size of δt from the equation (2.6), we find

$$u(\gamma_i, t_i; \gamma_{i+2}, t_{i+2}) = \sum_{\gamma_{i+1}} u_{\delta t}(\gamma_i, \gamma_{i+1}; t_i)\, u_{\delta t}(\gamma_{i+1}, \gamma_{i+2}; t_{i+1}) \tag{1.8}$$

Remarkably, this law is the same as the standard rule for multiplying matrices rows by columns. This is perhaps the single most noteworthy property of Markov processes which allows one to reduce problems in probability theory to linear algebra.

In matrix language, the equation above can be recast as follows:

$$u(t_i; t_{i+2}) = u_{\delta t}(t_i)\, u_{\delta t}(t_{i+1})\,. \tag{1.9}$$

For two generic time indices $i < j$, we have that

$$u(t_i; t_j) = u_{\delta t}(t_i) \cdots u_{\delta t}(t_{j-1})\,. \tag{1.10}$$

The path-wise representation for the latter formula is

$$u(\gamma_i, t_i; \gamma_j, t_j) = \sum_{\gamma:\gamma_i \to \gamma_j} u_{\delta t}(\gamma_i, \gamma_{i+1}; t_i) \cdots u_{\delta t}(\gamma_{j-1}, \gamma_j; t_{j-1}) \tag{1.11}$$

where the sum is over all paths $\gamma = \{\gamma_i, \ldots, \gamma_j\}$, γ_k is a state variable and $k = i, \ldots j$.

In the important case where the transition probability kernels $u_{\delta t}(x, y; t)$ are time-homogeneous, i.e. independent of time, over a certain time interval, then the numerical evaluation of kernels by matrix products can be greatly sped up by the so called fast exponentiation algorithm. Fast exponentiation proceeds iteratively evaluating first

$$u_{2\delta t} = u_{\delta t} \cdot u_{\delta t}. \tag{1.12}$$

Here we suppress writing the arguments x, y as we are writing this equation in operator notation as opposed to in kernel or matrix notation. We also suppress denoting the time coordinate as we are assuming time homogeneity. As a second step one evaluates

$$u_{4\delta t} = u_{2\delta t} \cdot u_{2\delta t} \tag{1.13}$$

and then iterates until one finds after n steps

$$u_{2^n\delta t} = u_{2^{n-1}\delta t} \cdot u_{2^{n-1}\delta t}. \tag{1.14}$$

2. Brownian Motion

Brownian motion can be described as a limit of the discrete process on the lattice $h\mathbb{Z} \equiv \{0, \pm h, \pm 2h, ...\}$ with Markov generator

$$\mathcal{L}(x, y) = \frac{1}{2}\Delta_h(x, y) \tag{2.1}$$

where $\Delta_h(x, y)$ is the matrix of the discrete Laplace operator given by

$$\Delta_h(x, y) = \frac{\delta_{x+h,y} + \delta_{x-h,y} - 2\delta_{x,y}}{h^2}. \tag{2.2}$$

The elementary transition probability matrix is given by

$$u_{\delta t}(x, y) = \delta_{xy} + \delta t \mathcal{L}(x, y) \tag{2.3}$$

where δt satisfies the Courant condition

$$\delta t \leq \frac{1}{\mathcal{L}(x, x)} \equiv h^2. \tag{2.4}$$

To numerically evaluate the transition probability matrix over a long time step, one can impose boundary conditions, form a finite matrix and then use the fast exponentiation algorithm. An alternative is to use the analytic approximation below. Until recent years, fast exponentiation was not technically viable and analytic approximations were the only route to compute kernels.

We start from observing that a complete set of orthogonal eigenvectors for the matrix $\mathcal{L}(x, y)$ is given by the functions

$$v_k(x) = e^{ikx} \tag{2.5}$$

Here, $k \in \left[-\frac{\pi}{h}, \frac{\pi}{h}\right]$. In fact, the following eigenvalue equation is satisfied:

$$\mathcal{L}v_k = \left(\frac{\cos kh - 1}{h^2}\right)v_k \tag{2.6}$$

Furthermore, according to the theory of Fourier series, all lattice functions $f(x)$ with $\sum_x |f(x)^2| < \infty$ can be written as follows

$$f(x) = \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \hat{f}(k) e^{-ikx} \frac{dk}{2\pi} \tag{2.7}$$

where

$$\hat{f}(k) = h \sum_{-\infty}^{\infty} f(x) e^{ikx} \tag{2.8}$$

Because of the eigenvalue equation in (2.6), we have that

$$u_{\delta t} v_k = \left(1 + \delta t \frac{\cos kh - 1}{h^2}\right) v_k. \tag{2.9}$$

The N-th power of this matrix thus satisfies the equation

$$u_{N\delta t} v_k = u_{\delta t}^N v_k = \left(1 + \delta t \frac{\cos kh - 1}{h^2}\right)^N v_k. \tag{2.10}$$

Notice that the matrix elements $\mathcal{L}(x, y)$ depend only on the difference $x - y$. This property is also shared by $u_{N\delta t}$ whose matrix elements can be written as follows

$$u_{N\delta t}(x, y) = U_{N\delta t}(x - y). \tag{2.11}$$

Hence

$$\sum_y U_{N\delta t}(x - y) e^{iky} = \left(1 + \delta t \frac{\cos kh - 1}{h^2}\right)^N e^{ikx}. \tag{2.12}$$

Otherwise stated

$$\sum_y U_{N\delta t}(x - y) e^{-ik(y-x)} = \left(1 + \delta t \frac{\cos kh - 1}{h^2}\right)^N. \tag{2.13}$$

The Fourier inversion formula yields the density

$$h^{-1} U_{N\delta t}(x - y) = \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \left(1 + \delta t \frac{\cos kh - 1}{h^2}\right)^N e^{ik(y-x)} \frac{dk}{2\pi} \tag{2.14}$$

To fulfill the Courant condition, let us set $\delta t = h^2$. We intend to take a limit as $h \to 0$ while holding $T = \delta t N$ constant. Hence N depends on h so that

$$N = N(h) = T h^{-2}. \tag{2.15}$$

By expanding the integrand in powers of h we find

$$h^{-1} U_T(x - y) = \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \left(1 - \frac{T}{N(h)} \frac{k^2}{2} + O(h^2)\right)^{N(h)} e^{ik(y-x)} \frac{dk}{2\pi} \tag{2.16}$$

Taking the limit as $h \to 0$ and applying Neper's formula we arrive at the following equation for the limit probability density

$$p_T(x,y) = \lim_{h\to 0} h^{-1} U_T(x-y) = \lim_{h\to 0} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-\frac{Tk^2}{2}} e^{ik(y-x)} \frac{dk}{2\pi} = \frac{1}{\sqrt{2\pi T}} e^{-\frac{(x-y)^2}{2T}}. \tag{2.17}$$

This density gives the transition probability for Brownian motion.

Notice that the time coordinate t is dimensionless in this equation. In practice, time is measure in some physical units (typically years in Finance). It is thus necessary to accompany calendar time with a factor that transforms it in a pure number. Typically, one writes this factor as the square of a parameter σ, so that $\sigma^2 T$ is dimensionless. The transition probability for Brownian motion with dimensional time is thus written as follows:

$$p_T(x,y) = \lim_{h\to 0} h^{-1} U_{\sigma^2 T}(x-y) = \lim_{h\to 0} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-\frac{\sigma^2 Tk^2}{2}} e^{ik(y-x)} \frac{dk}{2\pi} = \frac{1}{\sigma\sqrt{2\pi T}} e^{-\frac{(x-y)^2}{2\sigma^2 T}}. \tag{2.18}$$

Let us introduce the notations

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{2.19}$$

for the normal distribution density function and

$$\Phi(x) = \int_{-\infty}^{x} \phi(y)dy \tag{2.20}$$

for the cumulative normal distribution function. Let us notice that $\lim_{x\to-\infty}\Phi(x) = 0$ and $\lim_{x\to+\infty}\Phi(x) = 1$.

We have that

$$E_t[x_T | x_t = x] = \frac{1}{\sigma\sqrt{2\pi T}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2\sigma^2 T}} y dy = x. \tag{2.21}$$

Hence, Brownian motion is a martingale.

We have

$$E_t[\exp(x_T)] = \exp\left(\frac{x + \sigma^2 T}{2}\right). \tag{2.22}$$

Hence, also the process

$$X_t = x_t e^{\frac{-\sigma^2 t}{2}} \tag{2.23}$$

is a martingale. The process following by X_t is called *log-normal*.

We also have that

$$E_t[(x_T - x_t)^2 | x_t = x] = \frac{1}{\sigma\sqrt{2\pi T}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2\sigma^2 T}} (y-x)^2 dy = \sigma^2 T. \tag{2.24}$$

Hence, the standard deviation at time T of the distribution of Brownian motion of volatility σ is $\sigma\sqrt{T}$. A similar formula for the log-normal process is

$$E_t[(X_T - X_t)^2 | X_t = X] = \frac{1}{\sigma\sqrt{2\pi T}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2\sigma^2 T}} (e^{y-\frac{\sigma^2 T}{2}} - e^x)^2 dy = e^{2\sigma^2 T}. \tag{2.25}$$

The equation

$$E_t[(x_T - K)_+ | x_t = x] = (x-K)\Phi\left(\frac{x-K}{x\sigma\sqrt{T}}\right) + x\sigma\sqrt{T}\phi\left(\frac{x-K}{x\sigma\sqrt{T}}\right) \tag{2.26}$$

is called *Bachelier formula*. The equation

$$E_t[(X_T - K)_+ | X_t = X] = X\Phi\left(\frac{\log\frac{X}{K} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - K\Phi\left(\frac{\log\frac{X}{K} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) \tag{2.27}$$

is called the *Black-Scholes formula*.

3. Local Calibration with the Bachelier and Black-Scholes Model

The Bachelier formula was introduced by Bachelier in 1900 in his PhD thesis and is currently used in as a convention for price quoting. The Bachelier model applies to situations where one is interested in evaluating the expectation of a stochastic process which is known to be a martingale. If this process is modeled as Brownian motion, the Bachelier formula provides a closed form expression for call option prices. If instead one postulates that the martingale process is a log-normal martingale, one obtains the Black-Scholes formula for call options.

Applications of the Bachelier and Black-Scholes formula in equity markets apply to various payoffs including call options of payoff $(S_T - K)_+$. Let us assume here that the stock does not pay dividends. Here, S_t denotes the stock price process, T is the maturity, K is the strike. To setup a Bachelier model we need to consider a restricted universe in which only the option, the stock price S_t and a numeraire asset g_t are traded and then apply the Fundamental Theorem to this ideal situation. This approach is called *local calibration*. Regarding the numeraire asset, two possibilities are convenient: the forward measure of numeraire $g_t = Z_t(T)$ and the risk neutral measure where the numeraire is the money-market account.

To use the forward measure, consider the forward price process

$$F_t(T) = \frac{S_t}{Z_t(T)}. \tag{3.1}$$

According to the Fundamental Theorem, we require that

$$F_t(T) = E_t^{Z(T)}[S_T]. \tag{3.2}$$

because $S_T = F_T(T)$. To achieve this condition, we model the forward as a Brownian motion, i.e. set $F_t(T) = x_t$, so that the price of a call option can be evaluated as follows by means of a Bachelier formula

$$C_t(T,K) = Z_t(T)E_t^{Z(T)}\left[(S_T - K)_+ | F_t(T) = F\right] = Z_t(T)(F-K)\Phi\left(\frac{F-K}{F\sigma\sqrt{T}}\right) + Z_t(T)F\sigma\sqrt{T}\phi\left(\frac{F-K}{F\sigma\sqrt{T}}\right) \tag{3.3}$$

We can also set $F_t(T) = X_t$ and in this case we arrive at the Black-Scholes formula

$$C_t(T,K) = Z_t(T)E_t^{Z_t(T)}[(S_T-K)_+ | F_t(T) = F] = Z_t(T)F\Phi\left(\frac{\log\frac{F}{K} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - Z_t(T)K\Phi\left(\frac{\log\frac{F}{K} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right). \tag{3.4}$$

The Bachelier and the Black-Scholes formula are obviously different. However, the differences are numerically minor as long as maturities are short and one uses a Bachelier volatility σ_B equal to $\sigma_{BS}F$, where σ_{BS} is the Black-Scholes volatility.

In alternative, one can choose the money-market account B_t as a numeraire and express the Bachelier formula in terms of the futures price process $f_t(T)$ which is a martingale under this choice of measure. In this case, one can consider the case of options traded on a futures basis whereby the price changes of option contracts are settled on a daily basis, similarly to what happens for futures contracts on the underlying. The future price process for call options is also a martingale under the risk neutral measure. The Bachelier formula for futures price of call options reads

$$c_t(T,K) = E_t^{Z(T)}\left[(S_T - K)_+ | f_t(T) = f\right] = (f-K)\Phi\left(\frac{f-K}{f\sigma\sqrt{T}}\right) + f\sigma\sqrt{T}\phi\left(\frac{f-K}{f\sigma\sqrt{T}}\right) \tag{3.5}$$

We can also set $F_t(T) = X_t$ and in this case we arrive at the Black-Scholes formula for the future price of call options is

$$c_t(T,K) = E_t^{B_t}[(S_T - K)_+ | f_t(T) = f] = f\Phi\left(\frac{\log\frac{f}{K} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right) - K\Phi\left(\frac{\log\frac{f}{K} + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}\right). \tag{3.6}$$

Since the forward and the futures price of a stock are in general different, one arrives to slightly different prices for the same choice of volatility.

The Bachelier and Black-Scholes formula are commonly used also as quoting conventions. The idea is that if one fixes the price, then there is a volatility which reproduces it. That is called the *implied* Bachelier or Black-Scholes volatility.

The Bachelier and Black-Scholes local calibration procedure also extend to interest rate swaptions. In this case, the procedure involves identifying a small ideal market in which the only traded contracts are: a fixed rate annuity and a floating rate annuity with the same start and end date, and a swaption whose maturity is the start date of the annuities and whose tenor is the difference between their end and start date. Assuming that this small market niche is isolated from the rest of the world, one sets up a model which conforms to the Fundamental Theorem. Of course, if one wishes to include other swaptions of different strike, different maturity or different tenor, then the construction will be repeated and one will identify an instrument specific numeraire and implied volatility. Thus local calibration is not globally consistent across all assets as the Fundamental Theorem demands.

To price payer swaptions, consider that the equilibrium swap rate at time t is give by

$$SR_t = \frac{Z_t(T) - Z_t(T_0)}{\sum_{i=0}^{N} \tau Z_t(T_i)} \tag{3.7}$$

where T_0 is the start date of the swap and we assume that $t < T_0$. A payer swaption has a payoff equal to

$$(SR_t - \kappa)_+ \sum_{i=0}^{N} \tau Z_t(T_i) \tag{3.8}$$

at maturity T_0. If we select the numeraire asset $g_t = \sum_{i=0}^{N} \tau Z_t(T_i)$ and consider the asset price process for the swaption SO_t of strike κ, we find

$$SO_t = g_t E_t^g \left[(SR_t - \kappa)_+ | SR_t = r \right]. \tag{3.9}$$

Postulating that the swap rate follows a Brownian motion, we can evaluate the expectation as follows by means of the Bachelier formula

$$SO_t = g_t E_t^g \left[(SR_t - \kappa)_+ | SR_t = r \right] = g_t (r - \kappa) \Phi\left(\frac{r - \kappa}{r\sigma\sqrt{T}} \right) + g_t r\sigma\sqrt{T} \phi\left(\frac{r - \kappa}{r\sigma\sqrt{T}} \right) \tag{3.10}$$

If instead we set $SR_t = X_t$, we arrive at the Black-Scholes formula for swaptions

$$SO_t = g_t E_t^g \left[(SR_t - \kappa)_+ | SR_t = r \right] = g_t r \Phi\left(\frac{\log \frac{r}{\kappa} + \frac{1}{2}\sigma^2 T}{\sigma \sqrt{T}} \right) - g_t \kappa \Phi\left(\frac{\log \frac{r}{\kappa} + \frac{1}{2}\sigma^2 T}{\sigma \sqrt{T}} \right). \tag{3.11}$$

4. Problems

PROBLEM 14.

- Derive the Bachelier formula for equity put options under the forward measure.
- Derive the Bachelier formula for equity put futures options under the risk neutral measure.
- Derive the Black-Scholes formula for equity put options under the forward measure.
- Derive the Black-Scholes formula for equity put future options under the risk neutral measure.
- Derive the Bachelier formula for receiver swaptions of payoff

$$(\kappa - SR_t)_+ \sum_{i=0}^{N} \tau Z_t(T_i) \tag{4.1}$$

- Derive the Black-Scholes formula for receiver swaptions.

PROBLEM 15. How should one modify the Bachelier and Black-Scholes pricing formulas for equity call options in case the underlying pays dividends?

PROBLEM 16. Derive Bachelier and Black-Scholes formula for foreign exchange options.