AFFINE TERM-STRUCTURE MODELS

We study in the present chapter a class of interest-rate models which have the property that the forward rate curve can be represented as an **affine function** of Markov state variables. These are the most important models for practical applications, due to their computational tractability.

The forward rate curve for affine models has the form

$$f(t;T) \;=\; \sum_{i=1}^{N} a_i(t;T)\,X_i(t) \;+\; b(t;T) \tag{1}$$

where $\mathbf{X}(t) \;=\; (\,X_1(t),\, X_2(t),\, ...X_N(t)\,)$ is a vector of **state variables** satisfying a system of stochastic differential equations

$$dX_i \;=\; \sigma_{ik}(\mathbf{X},\,t)\,dZ_k \;+\; \mu_i(\mathbf{X},\,t)\,dt \quad 1 \;\le\; i \;\le\; N, \qquad 1 \;\le\; k \;\le\; \nu\;. \tag{2}$$

(Here, the Z_k 's are independent Brownian motions).[1] Notice that the discount factors P_t^T are **exponential functions** of the state variables:

$$P_t^T \;=\; e^{-\sum_{i=1}^{N}\left(\int_t^T a_i(t;s)\,ds\right)X_i(t) \;-\int_t^T b(t;s)\,ds}$$

$$=\quad e^{-\sum_{i=1}^{N} p_i(t;T)\,X_i(t) \;-\; q_i(t;T)}\;. \tag{3}$$

Converserly, a term-structure model with the property that the discount factors are exponential functions of a (multidimensional) diffusion process is an affine model.

A major advantage of affine models is that they can be implemented on the computer using recombining lattices or finite-difference schemes with relatively low dimensionality (in practice $N \;=\; 1,\,2$ or 3). This is a consequence of the fact that the state variables are

[1] We restrict this discussion to state variables which are Markov Ito processes, or diffusions. The study of Markovian state-variables with jumps, which is relevant for the study of default risk, is beyond the scope of these lectures.

a Markov process. Moreover, the affine form of the forward rate curve or, equivalently, the exponential form of the discount factors, allows us to express any cash-flow which is a function of forward rates as an elementary function – basically a sum or combination of exponential functions of the state variables.[2] Affine models are thus well-suited for pricing and hedging American bond options, American swaptions, callable bonds, exotic interest rate options, etc., as well as European-style derivatives.

As we shall see, the requirement that the forward-rate curve is an affine function implies strong constraints on the process $\mathbf{X}$. In other words, we cannot assign arbitrary probability distributions to the state variables $\mathbf{X}$ and expect to generate an affine model. Roughly speaking, there are three "classes" of affine interest-rate models, namely: Gaussian models, Cox-Ingersoll-Ross (CIR) square-root models and the so-called Li-Ritchken-Sankarasubhramanyam (LRS) models. We will present each of these models in this lecture.

Another advantage of affine models is that the volatilities and correlations of forward rates are easy to compute. Using Itôs Formula, we obtain

$$df(t;T) \;=\; \sum_{k=1}^{\nu} \dot{\sigma}_k(t;T)\, dZ_k \;+\; m(t;T)\, dt$$

where $\dot{\sigma}_k(t;T)$ and $m(t;T)$ are given respectively by

$$\dot{\sigma}_k(t;T) \;=\; \sum_{i=1}^{N} \sigma_{ik}\, a_i(t;T)$$

and

$$m(t;T) \;=\; \sum_{i=1}^{N} \left(\frac{\partial a_i(t;T)}{\partial t} X_i(t) \;+\; a_i(t;T)\, \mu_i \right) \;+\; \frac{\partial b(t;T)}{\partial t}\,. \tag{4}$$

In particular, the standard deviation of the forward rate $f(t;T)$ is given by

$$\sigma_f(t;T) \;=\; \sqrt{\sum_{k=1}^{\nu} \left(\sum_{i=1}^{N} \sigma_{ik}\, a_i(t;T) \right)^2}$$

and the correlation factors are

[2] Notice that, in constrast, the use of a "general" state variable model requires solving partial differential equations to compute the values of the discount factors and forward rates in terms of the state-variables. Hence, the valuation of interest rate derivatives, such as bond options, would require solving several PDEs instead of only one!

$$\phi_k(t;\,T) \;=\; \frac{\sum\limits_{i=1}^{N} \sigma_{ik}\, a_i(t;\,T)}{\sqrt{\sum\limits_{k=1}^{\nu}\left(\sum\limits_{i=1}^{N} \sigma_{ik}\, a_i(t;\,T)\right)^2}} \qquad 1 \;\leq\; k \;\leq\; \nu \;.$$

Hence the specification of the correlation factors and volatilites (for instance, using a principal component analysis) is relatively easy and consists in selecting the parameters of the diffusion equations (2) in such a way that the corresponding coefficients $a_i(t;\,T)$ lead to the desired correlation structure (see previous chapter).

In the following sections, we charaterize the distributions for the state-variables that give rise to affine models and compute the corresponding functions $a_i(t;\,T)$, $b_i(t;\,T,$ $p_i(t;\,T)$ and $q_i(t;\,T)$.

1. A characterization of affine models

Let us study in more detail the conditions on the distribution of the state-variables $\mathbf{X}$ which give rise to an affine model. The main point here is that we want equations (1) or (3) to hold under a *risk-neutral* measure. In the interest rate context, a risk-neutral measure is such that zero-coupon bond prices have a drift equal to the short-term interest rate. This implies certain restrictions on the coefficients μ_i and σ_{ik}, as we now show.

The first observation is that, since $r_t \;=\; f(t;\,t)$, the short rate satisfies

$$r_t \;=\; \sum_{i=1}^{N} \overline{a}_i(t)\, X_i(t) \;+\; \overline{b}(t) \;, \tag{5}$$

where, for simplicity, we introduced the notation[3]

$$\overline{a}_i(t) \;=\; a_i(t;t) \quad , \qquad \overline{b}(t) \;=\; b(t;\,t) \;.$$

We shall make use of the fact that each discount factor P_t^T satisfies, under a risk-neutral measure,

[3]It is often possible to "normalize" the functions $a_i(t;\,T)$ by imposing the condition $\overline{a}_i(t) \;=\; 1$. In fact, whenever $a_i(t;\,t) \;\neq\; 0$, this entails no loss of generality, since we can always redefine the state-variables X_i using the trasformation $X_i \;\longrightarrow\; a_i(t;\,t)\, X_i$. Nevertheless, there are situations in which $a_i(t\,t) \;=\; 0$ for some indices i. This corresponds to models in which the short rate depends on a smaller number of state variables than the entire forward rate curve. One important example is the LRS model.

$$dP_t^T \;=\; P_t^T \left[\sum_{k=1}^{\nu} \sigma_k(t;T)\, dZ_k \;+\; r_t\; dt \right] . \tag{6}$$

Applying Itô's Formula to equation (3), we find that the (lognormal) drift of P_t^T is given by

$$\delta \;=\; \frac{1}{P_t^T} \left[\frac{\partial P_t^T}{\partial t} \;+\; \frac{1}{2} \sum_{i,j\,=\,1}^{N} A_{ij}\, \frac{\partial^2 P_t^T}{\partial X_i\, \partial X_j} \;+\; \sum_{i=1}^{N} \mu_i\, \frac{\partial P_t^T}{\partial X_i} \right] ,$$

where

$$A_{ij} \;=\; \sum_{k=1}^{\nu} \sigma_{ik}\, \sigma_{jk}$$

is the *diffusion matrix* associated with the process $\mathbf{X}$. Due to the exponential form of P_t^T, we can compute all partial derivatives in the last equation explicitly. The resulting expression for the drift of the discount factor is

$$\delta \;=-\left(\sum_{i=1}^{N} \dot{p}_i\, X_i \;+\; \dot{q} \right) \;+\; \frac{1}{2} \sum_{i,j=1}^{N} A_{ij}\, p_i\, p_j \;-\; \sum_{i=1}^{N} \mu_i\, p_i \;,$$

where dots represent derivatives with respect to "calendar time" t. Using equation (6), which states that $\delta = \;r_t\;$, and the formula for the short-term interest rate in (5), we conclude that

$$-\left(\sum_{i=1}^{N} \dot{p}_i\, X_i \;+\; \dot{q} \right) \;+\; \frac{1}{2} \sum_{i,j=1}^{N} A_{ij}\, p_i\, p_j$$

$$-\sum_{i=1}^{N} \mu_i\, p_i \;\;=\; \sum_{i=1}^{N} \bar{a}_i\, X_i \;+\; \bar{b}\;,$$

or

$$\frac{1}{2} \sum_{i,j=1}^{N} A_{ij}\, p_i\, p_j \;-\; \sum_{i=1}^{N} \mu_i\, p_i \;=\; \sum_{i=1}^{N} (\dot{p}_i \;+\; \bar{a}_i\;)\; X_i \;+\; \dot{q} \;+\; \bar{b} \tag{7}$$

Notice that equation (7) implies that the combination

$$\frac{1}{2} \sum_{i,j=1}^{N} A_{ij}\, p_i\, p_j \; - \; \sum_{i=1}^{N} \mu_i\, p_i \tag{8}$$

is an *affine function* of the variables X_i. This leads is to consider the following cases:

Case 1: A_{ij} and μ_i are affine functions of the state variables.

Case 2: A_{ij} and μ_i are not affine but the combination (8) is an affine function.

We shall analyze first Case 1 in detail and leave the study of Case 2 to the end of the chapter. We set, accordingly,

$$A_{ij} \;=\; A_{ij}^{(0)} \;+\; \sum_k A_{ij,\,k}^{(1)}\, X_k \tag{9a}$$

and

$$\mu_i \;=\; \mu_i^{(0)} \;+\; \sum_k \mu_{i,\,k}^{(1)}\, X_k \;, \tag{9b}$$

Equating the coefficients of X_i and the constant terms in the resulting equation, we obtain a system of ordinary differential equations for the coefficients p_i and q, namely,

$$\dot{p}_i \;+\; \sum_k \mu_{ki}^{(1)}\, p_k \;+\; \overline{a}_i \;=\; \frac{1}{2} \sum_{k,\,l} A_{k\,l,\,i}^{(1)}\, p_k\, p_l \qquad 1 \;\leq\; i \;\leq\; N \;, \tag{10a}$$

$$\dot{q} \;+\; \sum_i \mu_i^{(0)}\, p_i \;+\; \overline{b} \;=\; \sum_{i,\,j} \frac{1}{2}\, A_{i\,j}^{(0)}\, p_i\, p_j \;. \tag{10b}$$

In view of the fact that $P_T^T \;=\; 1$ and equation (3), the functions p_i and q must also satisfy the conditions

$$p_i(T;\,T) \;=\; 0 \;, \quad \text{and} \quad q(T;\,T) \;=\; 0 \;\;. \tag{11}$$

Equations (10a) and the first boundary conditions in (11) determine the coefficients p_i. The last equation is used to obtain q by integating with respect to t and using (11). In the case of time-independent coefficients $\mu^{(1)}$ and $A_{ij}^{(1)}$, these equations can be solved in closed form.

In order to characterize the class of affine models corresponding to Case 1, we must look for conditions on A_{ij} and μ_i which guarantee that the stochastic differential equations (2) admit a solution for all times (so that the state-variables are well-defined quantities) and , in addition to this, we must solve the ordinary differential equations satisfied by the coefficients p_i and q. For simplicity, we begin with the simplest instance where equations (2) can be solved, which is the Gaussian case.

2. Gaussian models: general case

If we assume that $A_{ij} \; = \; A_{ij}^{(0)}$ in equation (9a), the process $\mathbf{X}$ satisfies a linear system of stochastic differential equations

$$dX_i \;=\; \sigma_{i\,k}\, dZ_k \;+\; \mu_i^{(0)}\, dt \;+\; \mu_{i\,j}^{(i)}\, X_j\; dt\;. \tag{12}$$

Here $\sigma_{i\,k}$ is the *square root* of A_{ij}. To solve equation (12), we introduce the auxillary matrix-valued function $\Psi(t;\,T)$ (or *transfer function*) which solves the (matrix-valued) differential equation

$$\frac{d}{dT}\,\Psi(t;\,T) \;=\; \mu^{(1)}(T)\,\Psi(t;T)\;, \qquad \Psi(t;\,t) \;=\; \mathbf{I}\;,$$

where $\mathbf{I}$ represents the identity matrix ($\mathbf{I}_{ij} \;=\; 1$ if $i=j$ and $\mathbf{I}_{ij} \;=\; 0$ if $i \neq j$). It is easy to verfy, using the method of variation of constants, that the solution of the SDE (12) satisfies

$$\mathbf{X}(t') \;=\; \Psi(t;\,t')\cdot X(t) \;+\; \cdot \int_t^{t'} \Psi(s;\,t')\;\cdot\; \mu^{(0)}(s)\,ds \;+$$

$$\int_t^{t'} \Psi(s;\,t')\;\cdot\;\sigma(s)\cdot dZ(s)\;, \tag{13}$$

for all $0 \;\leq\; t \;<\; t'$. This formula shows that $\mathbf{X}$ has Gaussian distribution.[4]

Let us compute the coefficients $p_i(t;\,T)$. Notice that equations (10a) reduce to a linear system of ordinary differential equations

$$\frac{d\,p_i}{d\,t} \;+\; \sum_j p_j\;\mu_{j\,i}^{(1)} \;+\; \overline{a}_i \;=\; 0\;,$$

[4] Solutions of linear stochastic differential equations such as (12) are called a **Gauss-Markov process**.

with boundary conditions $[\, p_i(t;T)\,]_{t=T} \;=\; 0$. The solution of this system can be computed as follows: let $\Phi(t;\,T)$ be the solution of the matrix-valued differential equation

$$\frac{d}{dt}\,\Phi(t;\,T) \;=\; -\,\Phi(t;\,T)\,\mu^{(1)}(t)$$

$$\Phi(T;\,T) \;=\; \mathbf{I}\;.$$

Using again the method of variation of constants and condition (11), we find that $p(t;\,T) \;=\; (p_1(t;\,T),\; ...p_N(t;\,T)\,)$ is given by

$$p(t;\,T) \;=\; \int_t^T \overline{a}(s)\;\cdot\;\Phi(t;s)\;ds\;, \tag{14}$$

where $\overline{a}(s) \;=\; (\overline{a}_1(s),\;...\,\overline{a}_N(s)\,)$.

Given the expressions obtained for X_i and p_i, we conclude that the forward rates satisfy the SDE

$$df(t;\,T) \;=\; \sum_i \dot{p}_i(t;\,T)\,dX_i(t)$$

$$=\; \sum_{i,k} \dot{p}_i(t;\,T)\,\sigma_{i\,k}\,dZ_k(t) \;+\; \text{drift terms}\;.$$

Recall that, from the HJM theorem, the instantaneous covariance structure of the forward rates determines completey their dynamics under the risk-neutral measure. We conclude, in particular, that the risk-neutral dynamics are independent of $\mu^{(0)}$ and of the initial value of the state-variables, $X_i(0)$. This allows us to simplify the calculations by introducing the vector of "reduced state variables" $\mathbf{Y} \;=\; (Y_i(t),\;...,Y_N(t))$ such that

$$\mathbf{Y}(t) \;=\; \int_0^t \Psi(s;\,t)\;\cdot\;\sigma(s)\cdot\;dZ(s)\;.$$

(Notice that these state variables satisfy the reduced linear SDE

$$dY_i \;=\; \sigma_{i\,k}\,dZ_k \;+\; \mu^{(i)}_{i\,k}\,dt \quad,\quad Y_i(0) \;=\; 0.\;)$$

The forward rate curve and discount factors can be expressed in the form

$$f(t;\, T) \;=\; \sum_i a_i(t;\, T)\, Y_i(t) \;+\; b(t;\, T) \tag{15}$$

and

$$P_t^T \;=\; \exp\left[-\, p_i(t;\, T)\, Y_i(t) \;-\; q_i(t;\, T) \,\right] \;, \tag{16}$$

where $b(t;\, T)$ and $q_i(t;\, T)$ depend only on $\overline{a}_i$, $\sigma_{i\,k}$, $\mu_{i\,k}^{(1)}$ and the current forward rate curve $f(0;\, T)$, but not on $\mu^{(0)}$.[5]

It remains to compute the coefficient $q(t;\, T)$. This is done by integrating both sides of equation (10b) from t to T, assuming $\mu^{(0)} \;=\; 0$. The result is

$$q(t;\, T) \;=\; \int_t^T \overline{b}(s)\, ds \;-\; \frac{1}{2} \int_t^T \sum_{i,\, j} A_{ij}^{(0)}(s)\, p_i(s;\, T)\, p_j(s;\, T)\; ds \;. \tag{17}$$

Notice that we have not yet specified the function $\overline{b}$. We claim that (consistently with the HJM theorem) this function is determined by the condition that the model prices correctly all zero-coupon bonds P_0^T, $T \;>\; 0$. In fact, equating the forward rate curve at time $t = 0$ to the affine function in (15) defining the forward rate curve, we have

$$f(0;\, T) \;=\; b(0;\, T) \;=\; \frac{\partial q(0;\, T)}{\partial T} \;.$$

We shall use this relation and equation (17) to determine $\overline{b}$. Differentiating equation (17) with respect to T and setting $t = 0$, we have

$$\frac{\partial q(0;\, T)}{\partial T} \;=\; \overline{b}(T) \;-\; \frac{\partial}{\partial T}\, \frac{1}{2} \int_0^T \sum_{i,\, j} A_{ij}^{(0)}(s)\, p_i(s;\, T)\, p_j(s;\, T)\; ds$$

$$=\; \overline{b}(T) \;-\; \int_0^T \sum_{i,\, j} A_{ij}^{(0)}(s)\, \frac{\partial p_i(s;\, T)}{\partial T}\; p_j(s;\, T)\;\; ds \;,$$

[5] Notice that we have kept the same notation for the coefficients $b(t;\, T)$ and $q(t;\, T)$ in the reduced representation. This is a harmless abuse of notation, since the latter functions have not yet been specified.

where we used the boundary conditions (11) and the symmetry of the matrix A_{ij}.

Hence, we conclude that the function $\overline{b}$ is given by

$$\overline{b}(T) = f(0;T) - \int_0^T \sum_{i,j} A_{ij}^{(0)}(s) \frac{\partial p_i(s;T)}{\partial T} p_j(s;T)\, ds .$$

The coefficient $q(t;T)$ is recovered by substituting this expression into formula (17). After some computation, we obtain the following expression for $q(t;T)$:

$$q(t;T) = \int_t^T f(0;s)\, ds - \frac{1}{2} \int_0^t \sum_{i,j} A_{ij}^{(0)} \left(p_i(s;T)\, p_j(s;T) - p_i(s;t)\, p_j(s;t) \right) ds$$

$$= \int_t^T f(0;s)\, ds + q_0(t;T) \tag{18}$$

where we set

$$q_0(t;T) = -\frac{1}{2} \int_0^t \sum_{i,j} A_{ij}^{(0)} \left(p_i(s;T)\, p_j(s;T) - p_i(s;t)\, p_j(s;t) \right) ds . \tag{19}$$

This result shows that the term-structure of interest rates can be "fitted" to the Gaussian term structure model with specified volatility structure by choosing $\overline{b}$ (or, equivalently, $q(t;T)$, or $b(t;T)$) as a function of $\overline{a}_i, \sigma_{ik}, \mu^{(1)}$ and the current forward rate curve. Some practitioners call $\overline{b}(t)$ the "fudge factor" – it is the term that needs to added to the linear combination of state variables in order to fit the current term-structure.

Gaussian models enjoy an interesting "factorization property" with regards to the discount factor. In fact, recalling that

$$e^{-\int_0^t f(0;s)\, ds} = \frac{P_0^T}{P_0^t} ,$$

we conclude from (18) that the value of the discount factor at time t is given by

$$P_t^T = \frac{P_0^T}{P_0^t} \cdot \exp\left[\sum_i - p_i(t;T)\, X_i(t) - q_0(t;T) \right] \tag{20}$$

This formula shows that the discount factor can be factorized into the product of a term that depends on the current term-structure of interest rates P_0^T/P_0^t (this is the "forward price of a loan of \$1 at the future time t for the period of time $T-t$) and an term that depends only on the volatility of forward rates. This factorization will prove useful later on for computing the values of caps.

3. Gaussian models: explicit formulas

In the previous section, we derived formulas for the discount factors of Gaussian models for general parameters $\mu_{ij}^{(1)}$ and $A_{ij}^{(0)}$, using the formalism of trasfer matrices. In this section, we obtain more explicit expressions by making two simplifying assumptions: first, that the matrix $\mu^{(1)}$ is diagonal and second, that the parameters are constasnt in time.[6] Accordingly, we set

$$\mu_{ij}^{(1)} \;=\; -\,\kappa_i\,\delta_{ij}\;,$$

and

$$A_{ij}^{(0)} \;=\; \sigma_i\,\sigma_j\,\rho_{ij}\;.$$

Under these assumptions, it is easy to check that the resolvents $\Phi(t;T)$ and $\Psi(t;T)$ are diagonal, with

$$\Phi_{ii}(t;T) \;=\; \Psi_{ii}(t;T) \;=\; e^{-\,\kappa_i(T-t)}\;, \quad i \;=\; 1,2,..,N\;,$$

and that the functions $p_i(t;T)$ and $a_i(t;T)$ are given by

$$p_i(t;T) \;=\; \frac{\overline{a}_i}{\kappa_i}\left(1 \;-\; e^{-\kappa_i(T\,-\,t)}\right)\;, \qquad a_i(t;T) \;=\; \overline{a}_i\,e^{-\,\kappa_i\,(T-t)}\;.$$

Notice that the SDE for the (reduced) state-variables is

$$dY_i \;=\; -\kappa_i\,Y_i\,dt \;+\; \sigma_i\,dZ_i\;,$$

[6]In practice, the latter assumption may not be appropriate if we wish to calibrate the model to a "term structure" of option prices. Nevetheless, we discuss the constant coefficients case because it leads to simple mathematical expressions. The assumption that $\mu^{(1)}$ is diagonal is beneficial, in our opinion, because it makes the specification of the correlation structure more "transparent", as we shall see.

with $\mathbf{E}(\, dZ_i \; dZ_j \,) \;=\; \rho_{ij}\, dt$. Each state variable behaves like a *mean-reverting* or *mean-repelling* Gaussian process, according to the sign of κ_i.

Using equation (18), we conclude after a straightforward (but tedious) calculation that

$$q(t;\, T) \;=\; \int_t^T f(0;\, s)\, ds \;-$$

$$\sum_{i\,j} \frac{\overline{a}_i\, \overline{a}_j\, \sigma_i\, \sigma_j\, \rho_{ij}}{2\, \kappa_i\, \kappa_j} \left[\frac{1 \;-\; e^{-\kappa_i\,(T-t)}}{\kappa_i} \;+\; \frac{1 \;-\; e^{-\kappa_j\,(T-t)}}{\kappa_j} \;-\; \frac{1 \;-\; e^{-(\kappa_i+\kappa_j)\,(T-t)}}{\kappa_i \;+\; \kappa_j} \right] .$$

The coefficient $b(t;\, T)$ is obtained by differentiating with respect to T. Accordingly,

$$b(t;\, T) \;=\; f(0;\, T) \;-$$

$$\sum_{i\,j} \frac{\overline{a}_i\, \overline{a}_j\, \sigma_i\, \sigma_j\, \rho_{ij}}{2\, \kappa_i\, \kappa_j} \left[e^{-\kappa_i\,(T-t)} \;+\; e^{-\kappa_j\,(T-t)} \;-\, e^{-(\kappa_i+\kappa_j)\,(T-t)} \right] .$$

The expression for the "fudge factor" for the short-rate process, $\overline{b}$, follows from setting $T \,=\, t$ in this last expression. We have

$$\overline{b}(t) \;=\; f(0;\, t) \;-$$

$$\sum_{i\,j} \frac{\overline{a}_i\, \overline{a}_j\, \sigma_i\, \sigma_j\, \rho_{ij}}{2\, \kappa_i\, \kappa_j} \left[e^{-\kappa_i\, t} \;+\; e^{-\kappa_j\, t} \;-\, e^{-(\kappa_i+\kappa_j)\, t} \right] .$$

We conclude that the short rate process for the Gaussian model with constant coefficients and diagona $\mu^{(1)}$ has the form

$$r_t \;=\; f(0;\, t) \;+\; \sum_i \overline{a}_i\, \sigma_i \int_0^t e^{-\kappa_i\,(t\,-\,s)}\, d\, Z_i \;-$$

$$\sum_{i,\,j} \frac{\overline{a}_i\, \overline{a}_j\, \sigma_i\, \sigma_j\, \rho_{ij}}{2\, \kappa_i\, \kappa_j} \left[e^{-\kappa_i\, t} \;+\; e^{-\kappa_j\, t} \;-\, e^{-(\kappa_i+\kappa_j)\, t} \right] .$$

This formula is analogous to the one obtained in the previous chapter when we discussed the Modified Vasicek model.

Notice that the "mean-reversion parameters" κ_1 determine the shape of the correlation factors, as explained in the previous chapter.

4. Square-root models: probability distribution of the state variables

Another imortant class of stochastic process that give rise to affine models are the **square-root processes**. These processes are discussed in classical Probability textbooks (*e.g* Feller, 195?) and were introduced in term-structure modelling in Finance by Cox, Ingersoll and Ross ( ) (See Longstaff and Schwartz ( ), Scott ( ) and Duffie and Kan ( ) for in-depth studies and extensiones of these models). In this section, we concentrate on one-factor square-root models, with emphasis in the probability distribution of the state variable.

Historically, the use of square-root processes was motivated by the fact that the state-variables are positive. This is an important advantage over Gaussian models, which lead unavoidably to negative interest rates. In addition to the issue of positive rates, the square-root processes offer a greater variety of distributions for the state variables and hence for the forward rates.

The material from this section draws from the classical discussion of the CIR model (see, for instance, ????).Nevertheless, our presentation differs from classical discussions of the CIR model because we treat the square-root process as a *state-variable* of an affine term-structure model, rather than as the short-term interest rate.

Following the general classification of affine models in §1, we consider the case of one-factor models ($N = 1$) for which the diffusion coefficient $a = a_{1\,1}$ is an affine function of the state variable, i.e.,

$$a \;=\; a^{(0)} \;+\; a^{(1)}\,X\;.$$

According to this equation, X satisfies the SDE

$$dX \;=\; \sqrt{a^{(0)} \;+\; a^{(1)}\,X}\;dZ \;+\; \left(\mu^{(0)} \;+\; \mu^{(1)}\,X\;\right)\,dt\;.$$

For simplicity, we shall assume that all the coefficients appearing in the latter equation are constant (independent of t). This equation defines a stochastic process $X(t)$ for all times provided that $a^{(0)} + a^{(1)} X(t)$ remains positive for all times. We will investigate this issue in detail now.

To fix ideas, assume that

$$a^{(1)} = \sigma^2 > 0$$

and set

$$Y = \frac{a^{(0)}}{a^{(1)}} + X .$$

From ( ), we see that the stochastic process Y satisfies formally the SDE

$$dY = \sigma \sqrt{Y}\, dZ + \kappa\, (\theta - Y)\, dt$$

where

$$\kappa = -\mu^{(1)} \quad \text{and} \qquad \theta = -\frac{\mu^{(0)}}{\mu^{(1)}} + \frac{a^{(0)}}{a^{(1)}} .$$

Equation ( ) can thus be viewed as the "standard form" of a one-dimensional square-root process. The main question of interest is to determine conditions on the coefficients σ, κ and θ which ensure that $Y(t)$ is well-defined for all $t > 0$. Unlike the Gaussian case, the solution of equation ( ) cannot be expressed in a simple form using Ito integrals. Nevertheless, the distribution of $Y(t)$ can be studied using the PDE satisfied by its probability density. The main result is sumarized in the following proposition.

Proposition 1: *Suppose that $Y(0) > 0$. Then,*

(i) the SDE ( ) admits a solution $Y(t)$ which is strictly positive for all $t > 0$ if and only if

$$\kappa\, \theta > \frac{1}{2} \sigma^2 .$$

(ii) If

$$0 < \kappa\, \theta \le \frac{1}{2} \sigma^2 ,$$

the process $Y(t)$ vanishes with probability 1. Nevertheless, there exists a unique solution of equation ( ) which remains non-negative for all times $t > 0$.

Assuming that $\kappa\,\theta \;>\; 0$, there are two subcases: (a): *If $\kappa \;> 0$ and $\theta \;>\; 0$, the process $Y(t)$ has a long-term equilibrium Gamma distribution with density function*

$$p(Y) \;=\; \frac{\left(\frac{2\kappa}{\sigma^2}\right)}{\Gamma\left(\frac{2\kappa\,\theta}{\sigma^2} \;-\; 1\right)} \;\cdot\; Y^{\frac{2\kappa\,\theta}{\sigma^2} \;-\; 1} \exp\left(-\; \frac{2\kappa}{\sigma^2}\; Y \right) \;,$$

where $\Gamma(p) \;=\; \int_0^\infty x^p\, e^{-\,x}\, dx$ is the Gamma function. (b) If $\kappa \;< 0$ and $\theta <\; 0$, then the process $Y(t)$ converges to $+\infty$ as $t \;\to\; \infty$ with probability 1.

(iii) If $\kappa\;\theta \;\le\; 0$, the process $Y(t)$ vanishes at a finite time and is absorbed at zero with probability 1.

We give a proof of these statements in the Appendix. This proposition can be interpreted intuitively as follows: for $Y(t) \;\approx\; 0$, the contribution to the dynamics of $Y(t)$ which comes from the Brownian motion becomes negligible. The dynamics are therefore controlled by the drift term, which is proportional to $\kappa\,\theta$. A positive drift has the effect of "pushing" the state variable into the half-line $\{Y \;>\; 0\}$. In contrast, a negative or vanishing drift will not drive the state variable to positive values after $Y = 0$ is touched for the first time. This means that the solution of the SDE cannot exist beyond the first time that that $Y(t) \;=\; 0$. (Unless a restoring mechanism is specified exogenously, for instance using a jump process at $Y \;=\; 0$.) The distinction between cases (i) and (ii) is more subtle but can also be uderstood heuristically. In fact, assume that the state variable is very near the boundary, at a distance $\epsilon \ll\; 1$. The time required for diffusing towards zero by Brownian motion is proportional to the distance to zero(ϵ) divided by the diffusion coefficient. Therefore, it is of order (at least)

$$\tau_{diff} \;\propto\; \frac{\epsilon^2}{\sigma^2\,\epsilon} \;=\; \frac{\epsilon}{\sigma^2}$$

On the other hand, the time rquired to drift upward (in the positive direction) by an amount ϵ is

$$\tau_{drift} \;\propto\; \frac{\epsilon}{\kappa\;\theta} \;.$$

Therefore, the probability of hitting zero is controlled by the ratio $\frac{\tau_{diff}}{\tau_{drift}} \;\approx\; \frac{\kappa\,\theta}{\sigma^2}$. This argument is made more precise in the Appendix using PDE methods.

Notice that in case (i) the density function vanishes at $Y = 0$ whereas case (ii) gives rise to an asymptotic distribution that has a singularity at $Y \;=\; 0$ if $\kappa\;\theta \;\;<\; \frac{\sigma^2}{2}$ and to an exponential distribution in the "marginal case" $\kappa\;\theta \;\;=\;\; \frac{\sigma^2}{2}$. It is easy to check that $p(Y)$ converges to a Dirac delta function at $Y = 0$ in the limit $\kappa\,\theta \;\;\to\;\; 0$, which show that for $\kappa\,\theta \;=\; 0$ the process is absorbed with probability 1. The tails of the equilibrium distribution are therefore controlled by the ratio $\frac{\kappa\,\theta}{\sigma^2}$.

Next, we characterize the distribution of $Y(t)$ at finite times. Setting $Y(0) \;=\; Y_0$, the probability density can be represented as a function of three variables:

$$p\,(Y_0,\; t;\; Y)\;\; .$$

This function is most easily characterized by its Laplace transform (or moment-generating function)

$$\hat{p}(Y_0,\, t;\; \lambda) \;\;=\;\; \int_0^\infty e^{-\lambda\, Y}\, p\,(Y_0,\; t;\; Y)\; dY \;\;=\;\; \mathbf{E}\left\{ e^{-\lambda\, Y(t)} \mid Y(0) \;=\; Y_0 \right\}\;.$$

This moment-generating function can be computed in closed form.

Proposition 2: *Assume that* $Y_0 \;>\; 0$ *and that* $\kappa\,\theta \;>\; 0$. *Then*

$$\hat{p}(Y_0,\, t;\; \lambda) \;\;=\;\; \frac{1}{\left[\, + \,\frac{\lambda\,\sigma^2}{2\kappa}\;(1 \;-\; e^{-\kappa\,t}\;)\,\right]^{\frac{2\kappa\,\theta}{\sigma^2}}} \cdot \exp\left[\frac{-\,\lambda\;Y_0\;e^{-\kappa\,t}}{1\;+\;\frac{\lambda\,\sigma^2}{2\kappa}(1\;-\;e^{-\kappa\,t})}\right]\;.$$

Proof: The calculation of $\hat{p}(Y_0,\, t;\; \lambda)$ is done observing that the expectation

$$p(Y,\, t,\, T) \;\;=\;\; \mathbf{E}_t\left\{ e^{-\lambda\, Y(T)} \mid Y(t) \;=\; Y \right\}$$

satisfies the partial differential equation

$$p_t \;+\; \frac{\sigma^2\, Y}{2}\, p_{YY} \;+\; \kappa\;(\theta \;-\; Y)\;\, p_Y \;=\; 0$$

$$p(Y,\, T,\, T) \;\;=\;\; e^{-\lambda\, Y}$$

A solution of this PDE can be sought in the form

$$p(Y,\, T,\, T) \;\;=\;\; \exp\;(\; -r(t;\, T)\, Y \;\;-\;\; r(t;\, T)\;)\;\; .$$

Substitution of this function into the PDE ( ) shows that the functions $r(t;T)$ and $s(t;T)$ satisfy the ordinary differential equations

$$\dot{r} \; - \; \kappa\, r \; = \; \frac{1}{2}\sigma^2\, r^2$$

$$\dot{s} \; + \; \kappa\,\theta\, r \; = \; 0\;,$$

with boundary conditions

$$r(T;T) \; = \; \lambda \quad , \quad s(T;T) \; = \; 0$$

These ordinary differential equations can be solved in closed form (the trasfomation $y \; = \; \frac{1}{r}$ can be used to liearize the first equation). The final result is obtained by setting $t = 0$, $T = t$, and $Y = Y_0$. We leave the details of the calculation to the reader.

The probability distribution corresponding to ( ) is called a *non-central Chi-square with* $\frac{4\,\kappa\,\theta}{\sigma^2}$ *degrees of freedom*. (Footnote: the terminology comes from the fact that if ν is an integer, x_1, x_2, $...x_\nu$ are standard normal random variables (with mean zero and variance 1) and a and b are real numbers, then the random variable

$$\sum_{i=1}^{\nu} (a\,x_i \; + \; b)^2$$

has moment generating function

$$\frac{1}{(1 \; + \; 2\,\lambda\,a^2)^{\frac{\nu}{2}}} \cdot e^{-\frac{\lambda\,\nu\,b^2}{(1 \; + \; 2\,\lambda\,a^2)}}\;.$$

The case $b \; = \; 0$ corresponds to the standard Chi-square distribution (sum of squares of ν independent normals). For $b \neq \; 0$, we have a *non-central* Chi-square with non-centrality parameter b. The concept of a fractional number of degrees of freedom stems from analytic continuation of the dimension ν to arbitrary positive real numbers in the above formula.)

From the Laplace transform of the distribution, we can compute the moments of the distribution, as well as the behavior of the probability density for $Y \; \ll \; 1$.

Corollary: *For* $\kappa \; > \; 0$, $\theta \; > \; 0$, *we have*

$$\mathbf{E}\left\{Y(t)\,|\,Y(0) \; = \; Y_0\;\right\} \; = \; e^{-\kappa\,t}\,Y_0 \; + \; \theta\,\left(1 \; - \; e^{\kappa\,t}\right) \; + \; Y_0\,e^{-\kappa\,t}$$

and

$$\mathbf{Var}\,\{Y(t)\,|\,Y(0) \;=\; Y_0\,\} \;=\frac{\sigma^2}{2\,\kappa}\,\left(1\,-\,e^{-\kappa\,t}\right)\,\left[\,\frac{\sigma^2}{2\,\kappa}\,\left(1\,-\,e^{-\kappa\,t}\right)\,+\,\frac{4\kappa\,\theta}{\sigma^2}\,Y_0\,e^{-\kappa\,t}\,\right]\,.$$

Moreover,

$$p(Y_0,\,t;\,Y)\;\propto\;Y^{\frac{2\kappa\,\theta}{\sigma^2}\,-\,1}\,,\qquad Y\;\ll\;1\,.$$

It follows from ( ) that θ corresponds to the long-term mean of $Y(t)$ and the long-term variance is $\left(\frac{\sigma^2}{2\kappa}\right)^2$. Notice the difference with the Gauss-Markov case, where the asymptotic variance is $\frac{\sigma^2}{2\kappa}$. Notice also that the probability density of $Y(t)$ behaves like a power of $Y(t)$ near $Y \;=\; 0$, according to ( ). Hence, the noncentral Chi-square with $\frac{2\kappa\theta}{\sigma^2}\;>\;1$ has smaller tails than the Gaussian density with same mean and variance (the latter assigns finite mass to $\{Y\leq\;0\}$) and fatter tails than the lognormal distribution (which vanishes to all orders at $Y=0$).

5. The one-factor square-root model: formulas for discount factors and forward rates

Having characterized the distribution of the state-variable Y, we compute the functions $p(t;\,T)$ and $q(t;\,T)$ which satisfy the equation

$$P_t^T\;=\;\exp\,[\,-\,p(t;\,T)\,Y(t)\;-\;q(t;\,T)\,]$$

and, in particular,

$$P_0^T\;=\;\exp\,[\,-\,p(0;\,T)\,Y(0)\;-\;q(0;\,T)\,]\;.$$

From the general considerations in §1 (equations (10a) and (10b)), the functions $p(t;\,T)$ and $q(t;\,T)$ satisfy the ordinary differential equations for p and q:

$$\dot{p}\;-\;\kappa\,p\;+\;\overline{a}\;=\;\frac{\sigma^2}{2}\,p^2\qquad p(T;\,T)\;=\;0\,,$$

$$\dot{q} \; + \; \kappa\theta\, p \; + \; \overline{b} \; = \; 0 \qquad , \qquad q(T;\, T) \; = \; 0 \; .$$

We observe that the transformations

$$Y \;\leftrightarrow\; \overline{a}\, Y, \qquad \sigma \;\leftrightarrow\; \overline{a}^{1/2}\, \sigma, \qquad \kappa \;\leftrightarrow\; \overline{a}\, \kappa$$

have the effect of reducing the computation to the case $\overline{a} \; = \; 1$. Thus, without loss of generality, we assume assume in the sequel that $\overline{a} \; = \; a(t;\, t) \; = \; 1$.

Equation ( ) which has a quadratic nonlinearity in p, is known as a **Ricatti** differential equation. It is well known that Ricatti equations can be "linearized" via the transformation

$$p \; = \; \frac{x}{y} \; ,$$

where x and y satisfy a linear system of ordinary differential equations of the form

$$\dot{x} \; = \; A_1\, x \; + \; B_1\, y$$

$$\dot{y} \; = \; A_2\, x \; + \; B_2\, y \; .$$

In fact, it follows from these two equations that

$$\left(\frac{x}{y}\right)^{\cdot} \; = \; \frac{\dot{x}}{y} \; - \; \frac{x\, \dot{y}}{y^2}$$

$$= \quad \frac{A_1\, x \; + \; B_1\, y}{y} \; - \; \frac{x \; (A_2\, x \; + \; B_2\, y\,)}{y^2}$$

$$= \; B_1 \; + \; (A_1 \; - \; B_2) \left(\frac{x}{y}\right) \; - \; A_2 \left(\frac{x}{y}\right)^2 \; .$$

In particular, setting

$$A_1 \; = \; -B_2 \; = \; \frac{\kappa}{2} \quad ; \quad B_1 \; = \; -1 \quad ; \quad A_2 \; = \; -\, \frac{\sigma^2}{2} \; ,$$

the function $p \; = \; x/y$ satisfies the Ricatti equation ( ). The solution of this equation can thus be obtained by solving the system

$$\dot{x} \; = \; \frac{\kappa}{2}\, x \; - \; \frac{\sigma^2}{2}\, y$$

$$\dot{y} \;=\; -\,x \;-\; \frac{\kappa}{2}\,y$$

with the boundary conditions

$$x(T;\,T) \;=\; 0 \quad \text{and} \quad \dot{x}(T;\,T) \;+\; y(T;\,T) \;=\; 0\;.$$

(The latter boundary condition arises from the fact that we assume that $\overline{a} \;=\; -\;[\dot{p}]_{t\,=\,T} \;=$ 1.)

A straightforward computation of the solution of the system ( ) gives the result

$$p(t;\,T) \;=\; \frac{2\left(1 \;-\; e^{-\,\nu\,(T-t)}\right)}{\kappa\,\left(1 \;-\; e^{-\,\nu\,(T-t)}\right) \;+\; \nu\,\left(1 \;+\; e^{-\,\nu\,(T-t)}\right)}\;,$$

where

$$\nu \;=\; \sqrt{\kappa^2 \;+\; 2\,\sigma^2}\;.$$

The function $q(t;\,T)$ is obtained, by integrating both sides of equation ( ), which gives

$$q(t;\,T) \;=\; \int_t^T \overline{b}(s)\,ds \;+\; \kappa\theta \int_t^T p(s;\,T)\,ds$$

$$=\; \int_t^T \overline{b}(s)\,ds \;+\; \kappa\theta\,\left(\frac{2\nu}{\sigma^2}\right)\log\left[\;\frac{\kappa\,\left(1 \;-\; e^{-\,\nu\,(T-t)}\right) \;+\; \nu\,\left(1 \;+\; e^{-\,\nu\,(T-t)}\right)}{2\nu\,e^{-\frac{\sigma^2\,(T-t)}{\kappa\;+\nu}}}\;\right]\;.$$

As in the analysis of Gaussian models, $\overline{b}$ is determined by matching the value of the zero-coupon bonds P_0^T to the market prices. This can be expressed by equation ( ), which is equivalent to

$$q(0;\,T) \;+\; p(0;\,T)\,Y(0) \;=\; \int_0^T f(0;\,s)\,ds\;.$$

Setting $t \;=\; 0$ in equation ( ) and solving for $\int_0^T \overline{b}$ gives therefore

$$\int_0^T \bar{b}(s)\,ds \;=\; \int_0^T f(0;\,s)\,ds \;-\; p(0;\,T)\,Y(0) \;-$$

$$\frac{2\nu\kappa\theta}{\sigma^2}\log\left[\frac{\kappa\left(1 \;-\; e^{-\nu T}\right) \;+\; \nu\;\left(1 \;+\; e^{-\nu T}\right)}{2\nu\, e^{-\frac{\sigma^2\,T}{\kappa\;+\nu}}}\right] \quad .$$

Using this identity, we conclude that

$$q(t;\,T) \;=\; \int_t^T f(0;\,s),ds \;-\; (p(0;\,T) \;-\; p(0;\,t))\;Y(0) \;-\; \log A(t;\,T)$$

where

$$A(t;\,T) \;=\; \left[\frac{\kappa\left(1 \;-\; e^{-\nu T}\right) \;+\; \nu\;\left(1 \;+\; e^{-\nu T}\right)}{\left(\kappa\;(1 \;-\; e^{-\nu t}) \;+\; \nu\;(1 \;+\; e^{-\nu t})\right)\cdot\left(\kappa\left(1 \;-\; e^{-\nu(T-t)}\right) \;+\; \nu\;\left(1 \;+\; e^{-\nu(T-t)}\right)\right)}\right]^{\frac{2\nu\kappa\theta}{\sigma^2}}$$

We conclude from this that the value at time t of a zero-coupon bond paying \$1 at time T is

$$P_t^T \;=\; \left(\frac{P_0^T}{P_0^t}\right)\cdot A(t;\,T)\;\cdot\exp\;\left[\,-\,p(t;\,T)\,Y(t)\;+(p(0;\,T)\;-\;p(0;\,t))\;Y(0)\,\right]\;.$$

Therefore, the stochastic discount factor P_t^T can be expressed as the product of the forward price for derlivery at time t of a zero-coupon bond maturing at time T and a model-dependent quantity which depends on the volatility of the forward rate curve.

Unlike in case of Gausssian models, the initial value of the state-variable $Y(0)$ and the long-term mean θ appear in the expression of the discount factor. This is due to the fact that the SDE that governs the dynamics of $Y(t)$ is non-linear and hence the model is not translation-invariant with respect to Y.

Let us derive formulas for the instantaneous forward rates and the short-term rate. We have

$$f(t;\,T) \;=\; a(t;\,T)\,Y(t) \;+\; b(t;\,T)$$

where $a(t;T) = \frac{\partial p(t;T)}{\partial T}$ and and $b(t;T) = \frac{\partial q(t;T)}{\partial T}$. Therefore, differentiating equations ( ) and ( ) with respect to T, we obtain

$$a(t;T) = \frac{4\,\nu^2\, e^{-\nu\,(T-t)}}{\left[\,\kappa\,\left(1 - e^{-\nu\,(T-t)}\right) + \nu\,\left(1 + e^{-\nu\,(T-t)}\right)\right]^2} .$$

and, from ( ),

$$b(t;T) = f(0;T) - \frac{4\,\nu^2\, e^{-\nu\,T}\, Y(0)}{\left[\,\kappa\,(1 - e^{-\nu\,T}) + \nu\,(1 + e^{-\nu\,T})\right]^2} + \left(\frac{2\nu^2\,\kappa\,(\nu-\kappa)\,\theta}{\sigma^2}\right) \times$$

$$\left[\frac{e^{-\nu\,T}}{\kappa\,(1 - e^{-\nu\,T}) + \nu\,(1 + e^{-\nu\,T})} - \frac{e^{-\nu\,(T-t)}}{\kappa\,\left(1 - e^{-\nu\,(T-t)}\right) + \nu\,\left(1 + e^{-\nu\,(T-t)}\right)}\right] .$$

Setting $T = t$ in this last formula, we conclude that the short rate process is given by

$$r_t = Y(t) + f(0\,t) - \frac{4\,\nu^2\, e^{-\nu\,t}\, Y(0)}{\left[\,\kappa\,(1 - e^{-\nu\,t}) + \nu\,(1 + e^{-\nu\,t})\right]^2} -$$

$$\frac{2\nu\,\kappa\,\theta\,\left(1 - e^{-\nu\,t}\right)}{\kappa\,(1 - e^{-\nu\,t}) + \nu\,(1 + e^{-\nu\,t})}$$

Notice that this function satisfies

$$a(t;T) \approx \frac{4\nu^2}{(\kappa + \nu)^2}\, e^{-\nu\,(T-t)} \qquad \text{for } T - t \gg 1 .$$

This is in qualitative agreement with the Gaussian models, in which ν is replaced by κ. Thus, the standard deviation of forward rates decays exponentially with rate ν as the maturity increases. We shall make use of this fact in the study of multifactor square-root models. In the latter case, "square-root" state-variables with different values of ν can be used to model a desired correlation structure.

5. Multi-factor square-root models

The analysis presented in the previous section suggests the following multifactor model:

$$f(t;T) \;=\; \sum_{i=1}^{N} a_i(t;T)\,Y_i(t) \;+\; b(t;T)$$

where the state-variables $Y_1(t)$, ... $Y_N(t)$ satisfy the system of partial differential equations

$$dY_i \;=\; \sigma_i\,\sqrt{Y_i}\,dZ_i \;+\; \kappa_i\,(\theta_i \;-\; Y_i\,)\,dt \qquad 1 \;\leq\; i \;\leq\; N.$$

This is a special case of the general affine structure ( ) for which the diffusion matrix has diagonal form

$$A_{i\,j} \;=\; \delta_{i,\,j}\,Y_i \qquad 1\; leq\; i,\; j \;\leq\; N\;,$$

with coefficients that depend linearly on the state-variables.