Applications of Stochastic Calculus in Finance
Chapter 2: Short-rate models

Gechun Liang

1 Arbitrage-free family of zero-coupon bond prices

Fix a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbf{P})$, which satisfies the usual conditions, and supports a one-dimensional Brownian motion W.

In short-rate models, we mainly model the dynamics of short rates. However, as we shall see later, the market under such models would never be complete.

Assumption 1 *(1) The short rate follows SDE*

$$dr_t = b_t dt + \sigma_t dW_t$$

which determines the bank account $B_t = e^{\int_0^t r_s ds}$. Moreover, the drift b and the volatility σ are both progressively measurable processes such that $\int_0^\cdot (b_s ds + \sigma_s dW_s)$ is a semimartingale.

(2) (No arbitrage): There exists an EMM $\mathbf{Q}$ whose Radon-Nikodym density of the form

$$\frac{d\mathbf{Q}}{d\mathbf{P}} = \mathscr{E}(-\int_0^\cdot \Theta_s dW_s)$$

such that the discounted zero-coupon bond price process $P(t,T)/B_t$ for $t \in [0,T]$ is a martingale under $\mathbf{Q}$, and $P(T,T) = 1$.

We make some comments on the above *no arbitrage* assumption. In short-rate models, the only tradeable asset is the bank account. Zero-coupon bonds or more general contingent claims are treated as derivatives as in the Black-Scholes theory, and the short rate (or the corresponding bank account) plays the role of underlying asset. Hence it is not possible to form portfolios which can replicate interesting contingent claims, not even zero-coupon bonds. Such a market is not complete, i.e. ELMM $\mathbf{Q}$ is not unique.

Gechun Liang
Department of Statistics, University of Warwick, U.K. e-mail: g.liang@warwick.ac.uk

Notwithstanding the non-uniqueness of ELMM, if the *no arbitrage* assumption in Assumption 1 holds, $\mathbf{Q}$ is not only an ELMM, but also an EMM. Hence, we have

$$P(t,T) = \mathbf{E}^{\mathbf{Q}}[e^{-\int_t^T r_s ds} \times 1 | \mathscr{F}_t]$$

which coincides the risk-neutral pricing formula.

Definition 1. A family $P(t,T)$ for $0 \leq t \leq T < \infty$ of adapted processes is called an arbitrage-free family of zero-coupon bonds if the *no arbitrage* assumption in Assumption 1 holds.

Proposition 1. *Under Assumption 1, the short rate $r = (r_t)_{t\geq 0}$ follows SDE*

$$dr_t = (b_t - \sigma_t \Theta_t)dt + \sigma_t dW_t^{\mathbf{Q}}$$

under the EMM $\mathbf{Q}$. Moreover, if the filtration $\{\mathscr{F}_t\}_{t\geq 0}$ is the Brownian filtration, then there exists a process $h \in \mathscr{L}^2(\mathbb{R})$ such that

$$\begin{aligned} \frac{dP(t,T)}{P(t,T)} &= r_t dt + h_t dW_t^{\mathbf{Q}} \\ &= (r_t + h_t \Theta_t)dt + h_t dW_t. \end{aligned}$$

Proof. We only show that

$$\frac{dP(t,T)}{P(t,T)} = r_t dt + h_t dW_t^{\mathbf{Q}}.$$

The other two equations follow from the Girsanov's theorem.

Indeed, since $P(t,T)e^{-\int_0^t r_s ds}$ is a $\mathbf{Q}$-martingale, by the martingale representation, there exists a process $\tilde{h} \in \mathscr{L}^2(\mathbb{R})$ such that

$$P(t,T)e^{-\int_0^t r_s ds} = P(0,T) + \int_0^t \tilde{h}_s dW_s^{\mathbf{Q}}.$$

Itô's formula yields

$$\frac{dP(t,T)}{P(t,T)} = r_t dt + \frac{\tilde{h}_t B_t}{P(t,T)} dW_t^{\mathbf{Q}}.$$

We conclude by letting $h_t = \tilde{h}_t B_t / P(t,T)$. $\square$

The summary of short-rate models: the dynamics of the short rate r are

$$\begin{aligned} dr_t &= b_t dt + \sigma_t dW_t \quad \text{under } \mathbf{P} \\ &= (b_t - \sigma_t \Theta_t)dt + \sigma_t dW_t^{\mathbf{Q}} \quad \text{under } \mathbf{Q}. \end{aligned}$$

The dynamics of the zero-coupon bond price $P(t,T)$ are

$$\begin{aligned} \frac{dP(t,T)}{P(t,T)} &= r_t dt + h_t dW_t^{\mathbf{Q}} \quad \text{under } \mathbf{Q} \\ &= (r_t + h_t \Theta_t)dt + h_t dW_t \quad \text{under } \mathbf{P}. \end{aligned}$$

A short-rate model is not fully determined without the exogenous specification of the market price of risk Θ. Hence, it is custom to postulate the $\mathbf{Q}$*-dynamics of the short rate r directly in the context of derivative pricing.*

2 Affine term structure of short-rate models

In the rest of this chapter, suppose that the short rate r follows

$$dr_t = b(t,r_t)dt + \sigma(t,r_t)dW_t^{\mathbf{Q}},$$

where $b(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$ are deterministic functions, and the initial data r_0 is in an open set $\mathscr{O} \subset \mathbb{R}$. Typical choices of $\mathscr{O}$ are $\mathbb{R}$ and $(0,\infty)$.

By the Markov property:

$$P(t,T) = \mathbf{E}^{\mathbf{Q}}[e^{-\int_t^T r_s ds}|\mathscr{F}_t] = \mathbf{E}^{\mathbf{Q}}[e^{-\int_t^T r_s ds}|r_t] = F(t,r_t)$$

for some function $F(\cdot,\cdot)$.

If $F(t,r) \in C^{1,2}([0,T) \times \mathscr{O})$, then by the Feynman-Kac formula, $F(t,r)$ solves the following term-structure equation on $[0,T) \times \mathscr{O}$:

$$\begin{cases} \partial_t F(t,r) + \frac{1}{2}\sigma^2(t,r)\partial_{rr}F(t,r) + b(t,t)\partial_r F(t,r) - rF(t,r) = 0, \\ F(T,r) = 1. \end{cases}$$

For the case with the state space $(0,\infty)$, a parameter condition on the coefficients need to be imposed to guarantee the above term-structure PDE is well posed even without a boundary condition on $r = 0$.

Definition 2. (Affine term structure)

A short-rate model is said to provide an affine term structure (ATS) if the corresponding zero-coupon price $P(t,T) = F(t,r)$ is of the form

$$F(t,r) = e^{-A(t)-B(t)r}$$

for some functions $A(\cdot)$ and $B(\cdot)$, where $A(T)=B(T)=0$.

Theorem 1. *A short-rate model provides an ATS iff the volatility and drift terms are of the form:*

$$\sigma^2(t,r)=a(t)+\alpha(t)r, \quad b(t,r)=b(t)+\beta(t)r$$

for some continuous functions $a(\cdot)$, $b(\cdot)$, $\alpha(\cdot)$ and $\beta(\cdot)$, and moreover, the functions $A(\cdot)$ and $B(\cdot)$ in $F(t,r)=e^{-A(t)-B(t)r}$ solve the following ODEs:

$$\begin{cases} \dfrac{dA(t)}{dt}=\dfrac{1}{2}a(t)B^2(t)-b(t)B(t), \quad A(T)=0; \\ \dfrac{dB(t)}{dt}=\dfrac{1}{2}\alpha(t)B^2(t)-\beta(t)B(t)-1, \quad B(T)=0. \end{cases}$$

Proof. Inserting $F(t,r)=e^{-A(t)-B(t)r}$ into the term-structure equation, we obtain that ATS iff

$$\frac{1}{2}\sigma^2(t,r)B^2(t)-b(t,r)B(t)=\frac{dA(t)}{dt}+(\frac{dB(t)}{dt}+1)r \tag{1}$$

for any $t\in[0,T)$ and $r\in\mathscr{O}\subset\mathbb{R}$.

If part: Substitute the ODEs for $A(\cdot)$ and $B(\cdot)$ into the RHS of the above equation:

$$RHS=\frac{1}{2}a(t)B^2(t)-b(t)B(t)+(\frac{1}{2}\alpha(t)B^2(t)-\beta(t)B(t))r.$$

Substitute $\sigma^2(t,r)=a(t)+\alpha(t)r$ and $b(t,r)=b(t)+\beta(t)r$ into its LHS:

$$\begin{aligned} LHS&=\frac{1}{2}(a(t)+\alpha(t)r)B^2(t)-(b(t)+\beta(t)r)B(t) \\ &=\frac{1}{2}a(t)B^2(t)-b(t)B(t)+(\frac{1}{2}\alpha(t)B^2(t)-\beta(t)B(t))r. \end{aligned}$$

Only if part: We only consider the case that $B_T(t)$ and $B_T^2(t)$ are linearly independent for any fixed $t\geq 0$, where we use sub T to emphasize the dependence on the maturity T. The linear dependent case is left as an exercise (see Filipovic [1] Chapter 5).

For any $T_1>T_2>t$,

$$\begin{pmatrix} B_{T_1}^2(t), & -B_{T_1}(t) \\ B_{T_2}^2(t), & -B_{T_2}(t) \end{pmatrix}\begin{pmatrix} \frac{1}{2}\sigma^2(t,r) \\ b(t,r) \end{pmatrix}=\begin{pmatrix} \frac{dA_{T_1}(t)}{dt} \\ \frac{dA_{T_2}(t)}{dt} \end{pmatrix}+\begin{pmatrix} \frac{dB_{T_1}(t)}{dt}+1 \\ \frac{dB_{T_2}(t)}{dt}+1 \end{pmatrix}r$$

Since

$$\begin{pmatrix} B_{T_1}^2(t) \\ B_{T_2}^2(t) \end{pmatrix} \text{ and } \begin{pmatrix} B_{T_1}(t) \\ B_{T_2}(t) \end{pmatrix}$$

are linearly independent by the assumption, we obtain that

$$\begin{pmatrix} \frac{1}{2}\sigma^2(t,r) \\ b(t,r) \end{pmatrix} = \begin{pmatrix} B_{T_1}^2(t), & -B_{T_1}(t) \\ B_{T_2}^2(t), & -B_{T_2}(t) \end{pmatrix}^{-1} \left(\begin{pmatrix} \frac{dA_{T_1}(t)}{dt} \\ \frac{dA_{T_2}(t)}{dt} \end{pmatrix} + \begin{pmatrix} \frac{dB_{T_1}(t)}{dt} + 1 \\ \frac{dB_{T_2}(t)}{dt} + 1 \end{pmatrix} r \right).$$

Hence, $\sigma^2(t,r)$ and $b(t,r)$ are affine functions of r. Plugging this in, LHS of (1) reads

$$\frac{1}{2}a(t)B_T^2(t) - b(t)B_T(t) + (\frac{1}{2}\alpha(t)B_T^2(t) - \beta(t)B_T(t))r.$$

Terms containing t must match. This implies the two ODE. □

3 Some standard short-rate models

1. Vasicek Model

$$dr_t = (a - br_t)dt + \sigma dW_t^{\mathbf{Q}}$$

with $a, b, \sigma > 0$.

(1) The solution is

$$r_t = r_0 e^{-bt} + \frac{a}{b}(1 - e^{-bt}) + \sigma \int_0^t e^{b(s-t)} dW_s^{\mathbf{Q}}.$$

(2) The expectation of r_t is

$$\mathbf{E}^{\mathbf{Q}}[r_t] = r_0 e^{-bt} + \frac{a}{b}(1 - e^{-bt}) \rightarrow \frac{a}{b}, \quad \text{as } t \uparrow \infty.$$

(3) The variance of r_t is

$$Var[r_t] = \frac{\sigma^2}{2b}(1 - e^{-2bt}) \rightarrow \frac{\sigma^2}{2b}, \quad \text{as } t \uparrow \infty.$$

(4) The price of the zero-coupon bond is $P(t,T) = e^{-A(t)-B(t)r}$ where

$$\begin{cases} \frac{dA(t)}{dt} = \frac{1}{2}\sigma^2 B^2(t) - aB(t), \quad A(T) = 0; \\ \frac{dB(t)}{dt} = bB(t) - 1, \quad B(T) = 0. \end{cases}$$

The solution is

$$\begin{cases} B(t) = \frac{-1}{b}(e^{-b(T-t)} - 1) \\ A(t) = \int_t^T \left[aB(s) - \frac{1}{2}\sigma^2 B^2(s) \right] ds. \end{cases}$$

(5) The drawback is that the short rate could be negative : $\mathbf{Q}(r_t < 0) > 0$.

2. Cox-Ingersoll-Ross (CIR) Model

$$dr_t = (a - br_t)dt + \sigma\sqrt{r_t}dW_t^{\mathbf{Q}}$$

with $a, b, \sigma > 0$.

(1) No explicit solution since r_t is non Gaussian. However, the short rate is always nonnegative: $r_t \geq 0$. Moreover, by using Feller's test, one can show that $r_t > 0$ if the parameters satisfy $\sigma^2 \leq 2a$ and $r_0 > 0$ (so no need to impose a boundary condition on $r = 0$ for the term-structure PDE). See Jeanblanc et al [2] Chapter 6 for the proof.

(2) The expectation of r_t. Applying Itô's formula to $e^{bt}r_t$ yields

$$d(e^{bt}r_t) = ae^{bt}dt + \sigma e^{bt}\sqrt{r_t}dW_t^{\mathbf{Q}}.$$

Hence,

$$\begin{aligned} e^{bt}r_t &= r_0 + a\int_0^t e^{bs}ds + \sigma\int_0^t e^{bs}\sqrt{r_s}dW_s^{\mathbf{Q}} \\ &= r_0 + \frac{a}{b}(e^{bt} - 1) + \sigma\int_0^t e^{bs}\sqrt{r_s}dW_s^{\mathbf{Q}}. \end{aligned}$$

Taking expectation gives us

$$\mathbf{E}^{\mathbf{Q}}[r_t] = r_0 e^{-bt} + \frac{a}{b}(1 - e^{-bt}) \to \frac{a}{b}, \quad \text{as } t \uparrow \infty.$$

(3) The variance of r_t. Introduce $X_t = e^{bt}r_t$. Then

$$dX_t = ae^{bt}dt + \sigma e^{\frac{1}{2}bt}\sqrt{X_t}dW_t^{\mathbf{Q}}.$$

Applying Itô's formula to X_t^2 yields

$$dX_t^2 = 2ae^{bt}X_t dt + 2\sigma e^{\frac{1}{2}bt}X_t^{\frac{3}{2}}dW_t^{\mathbf{Q}} + \sigma^2 e^{bt}X_t dt.$$

Hence,

$$X_t^2 = X_0^2 + (2a + \sigma^2)\int_0^t e^{bs}X_s ds + 2\sigma\int_0^t e^{\frac{1}{2}bs}X_s^{\frac{3}{2}}dW_s^{\mathbf{Q}}.$$

Taking expectation gives us

$$\begin{aligned} \mathbf{E}^{\mathbf{Q}}[X_t^2] &= X_0^2 + (2a + \sigma^2)\int_0^t e^{bs}\mathbf{E}^{\mathbf{Q}}[X_s]ds \\ &= r_0^2 + (2a + \sigma^2)\int_0^t e^{bs}(r_0 + \frac{a}{b}(e^{bs} - 1))ds. \end{aligned}$$

In turns,

$$\begin{aligned} \mathbf{E}^{\mathbf{Q}}[r_t^2] &= e^{-2bt}\mathbf{E}^{\mathbf{Q}}[X_t^2] \\ &= e^{-2bt}r_0^2 + \frac{2a + \sigma^2}{b}(r_0 - \frac{a}{b})(e^{-bt} - e^{-2bt}) + \frac{a(2a + \sigma^2)}{2b^2}(1 - e^{-2bt}). \end{aligned}$$

Finally, we obtain

$$\begin{aligned} Var[r_t] &= \mathbf{E}^{\mathbf{Q}}[r_t^2] - (\mathbf{E}^{\mathbf{Q}}[r_t])^2 \\ &= \frac{\sigma^2}{b} r_0 (e^{-bt} - e^{-2bt}) + \frac{a\sigma^2}{2b^2}(1 - 2e^{-bt} + e^{-2bt}) \to \frac{a\sigma^2}{2b^2}, \quad \text{as } t \uparrow \infty. \end{aligned}$$

(4) The price of the zero-coupon bond is $P(t,T) = e^{-A(t)-B(t)r}$ where

$$\begin{cases} \dfrac{dA(t)}{dt} = -aB(t), \quad A(T) = 0; \\ \dfrac{dB(t)}{dt} = \dfrac{1}{2}\sigma^2 B^2(t) + bB(t) - 1, \quad B(T) = 0. \end{cases}$$

The equation for $B(\cdot)$ is a Riccati ODE. The solution is

$$\begin{cases} B(t) = \dfrac{\sinh(\gamma(T-t))}{\gamma \cosh(\gamma(T-t)) + \frac{b}{2}\sinh(\gamma(T-t))} \\ A(t) = -\dfrac{2a}{\sigma^2} \ln\left[\dfrac{\gamma e^{\frac{1}{2}b(T-t)}}{\gamma \cosh(\gamma(T-t)) + \frac{b}{2}\sinh(\gamma(T-t))}\right] \end{cases}$$

where $\gamma = \frac{1}{2}\sqrt{b^2 + 2\sigma^2}$ and

$$\sinh(\gamma(T-t)) = \frac{1}{2}(e^{\gamma(T-t)} - e^{-\gamma(T-t)}); \quad \cosh(\gamma(T-t)) = \frac{1}{2}(e^{\gamma(T-t)} + e^{-\gamma(T-t)}).$$

See Shreve [3] Chapter 6 for the proof.
3. Extended Vasicek Model

$$dr_t = (a(t) - b(t)r_t)dt + \sigma(t)dW_t^{\mathbf{Q}}$$

with deterministic functions $a(t), b(t), \sigma(t) > 0$.

4.Extended CIR Model

$$dr_t = (a(t) - b(t)r_t)dt + \sigma(t)\sqrt{r_t}dW_t^{\mathbf{Q}}$$

with deterministic functions $a(t), b(t), \sigma(t) > 0$.

4 Exercises

Exercise 1. (CIR model)
Let $W = (W^1, \cdots, W^d)$ be a d-dimensional Brownian motion. For $1 \le j \le d$, let X^j be the solution of the following Ornstein-Uhlenbeck SDE:

$$dX_t^j = -\frac{b}{2}X_t^j dt + \frac{\sigma}{2}dW_t^j, \quad X_0^j = x^j$$

where $b > 0$, $\sigma > 0$.

1. Show that
$$X_t^j = e^{-\frac{b}{2}t}[x^j + \frac{\sigma}{2}\int_0^t e^{\frac{b}{2}u}dW_u^j]$$
2. Calculate $m(t) = \mathbf{E}[X_t^j]$ and $v(t) = Var(X_t)$ and their limites when $t \to \infty$.
3. Define $r_t = \sum_{j=1}^d |X_t^j|^2$. Show that
$$dr_t = (\frac{d\sigma^2}{4} - br_t)dt + \sigma\sqrt{r_t}dB_t$$
where
$$B_t = \int_0^t \frac{1}{\sqrt{r_s}}\sum_{j=1}^d X_s^j dW_s^j$$
is a Brownian motion.
4. Prove that $X_t^1 \cdots, X_t^d$ are iid normal random variables $N(m(t), v(t))$. Therefore, $r_t = \sum_{j=1}^d |X_t^j|^2$ is the sum of square of iid normal random variables, and hence r_t has χ^2-distribution.
5. Prove that the moment generating function of $|X_t^j|^2$ is given by
$$\mathbf{E}[\exp\{\mu|X_t^j|^2\}] = \frac{1}{\sqrt{1-2v(t)\mu}}\exp\left\{\frac{\mu|m(t)|^2}{1-2v(t)\mu}\right\}$$
for any $\mu < \frac{1}{2v(t)}$.
6. Based on (5), prove that the moment generating function of r_t is given by
$$\mathbf{E}[\exp\{\mu r_t\}] = \frac{1}{(\sqrt{1-2v(t)\mu})^d}\exp\left\{\frac{d\mu|m(t)|^2}{1-2v(t)\mu}\right\}$$
for any $\mu < \frac{1}{2v(t)}$.

Exercise 2. (Ho-Lee model and corresponding forward rate)
The one dimensional Ho-Lee model is given by

$$dr_t = b(t)dt + \sigma dW_t^{\mathbf{Q}}$$

under the EMM $\mathbf{Q}$, where $b(\cdot)$ is some deterministic function, and $\sigma > 0$. The corresponding zero-coupon bond price $P(t,T)$ is calculated as

$$P(t,T) = \mathbf{E}^{\mathbf{Q}}[e^{-\int_t^T r_s ds}|\mathscr{F}_t].$$

1. Since $(r_t)_{t\geq 0}$ admits Markov property, there exists some measurable function $F(t,r)$ on $[0,T]\times\mathbb{R}$ such that $P(t,T) = F(t,r_t)$. Suppose that $F(t,r)$ is in

$C^{1,2}([0,T) \times \mathbb{R})$. Write down the PDE for $F(t,r)$ by using the Feynman-Kac formula.

2. Explain why $F(t,r)$ has the affine form:

$$F(t,r) = e^{-A(t)-B(t)r}.$$

Prove that $A(t)$ and $B(t)$ satisfy the following ODE system:

$$\frac{dA(t)}{dt} = -b(t)B(t) + \frac{1}{2}\sigma^2|B(t)|^2;$$
$$\frac{dB(t)}{dt} = -1,$$

and solve the above ODE system to get the expressions for $A(t)$ and $B(t)$.

3. Recall that the forward rate $f(t,T)$ is defined as

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}.$$

Now suppose $b(t)$ has the form $b(t) = \partial_t f(0,t) + \sigma^2 t$. Prove that the short rate r_t has the dynamic

$$r_t = f(0,t) + \frac{\sigma^2 t^2}{2} + \sigma W_t^{\mathbf{Q}},$$

and the forward rate has $f(t,T)$ has the dynamic

$$f(t,T) = f(0,T) + \sigma^2 t(T - \frac{t}{2}) + \sigma W_t^{\mathbf{Q}}.$$

Therefore the volatility $\sigma(t,T)$ of the forward rate is a constant: $\sigma(t,T) = \sigma$.

4. Show that the drift of the forward rate is nothing but $\sigma(t,T)\int_t^T \sigma(t,s)ds$.

References

1. Filipovic, Damir. *Term-Structure Models. A Graduate Course.* Springer, 2009.
2. Jeanblanc, Monique, Marc Yor, and Marc Chesney. *Mathematical methods for financial markets.* Springer, 2009.
3. Shreve, Steven E. *Stochastic calculus for finance II: Continuous-time models*. Springer, 2004.