MATH5636 Stochastic Calculus for Finance II Spring 2026

Chapter 6: Interest Rate Models

Lecturer: Kenneth Ng

Preview

Interest rates are not constant over time; rather, they are inherently stochastic and evolve dynamically, much like risky assets. In this chapter, we introduce a range of interest rate models, including one-factor models, multi-factor models, and the Heath–Jarrow–Morton framework. We also discuss the associated bond-pricing equations under these different modeling approaches.

Key topics in this chapter:

1. One-factor interest rate models;
2. Multi-factor interest rate models;
3. Bond-pricing equations;
4. Forward rates and the Heath-Jarrow-Merton model.

1 One-Factor Models

In this chapter, we let $[0, \bar{T}]$ be a fixed horizon, where $\bar{T} > 0$. We denote by $\{r_t\}_{t\in[0,\bar{T}]}$ an interest rate (a.k.a. the *short rate*) process, which is an adapted stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\bar{T}]}, \mathbb{P})$.

A ***zero-coupon bond*** with maturity $T \in (0, \bar{T}]$ is a contract that pays 1 at time T. Let $P(t, T)$, $t \in [0, T]$, denote the price of such a bond at time t. To introduce the interest rate model, we first define the (short-rate) risk-neutral measure.

Definition 1.1 A probability measure $\widetilde{\mathbb{P}}$ is called a (short-rate) risk-neutral measure if, for every maturity $T \in (0, \bar{T}]$, the discounted zero-coupon bond price process

$$\left\{e^{-\int_0^t r_s\,ds}\, P(t, T)\right\}_{t\in[0,T]}$$

is a $\widetilde{\mathbb{P}}$-martingale.

One-factor models are driven by a single source of randomness. In the diffusion framework, a one-factor interest rate $\{r_t\}_{t\in[0,T]}$ is modeled by the following SDE under the risk-neutral measure:

$$dr_t = \beta(t, r_t)\, dt + \gamma(t, r_t)\, d\widetilde{B}_t, \tag{1}$$

where $\beta, \gamma : [0, \bar{T}] \times \mathbb{R} \to \mathbb{R}$, and $\{\widetilde{B}_t\}_{t\in[0,\bar{T}]}$ is a Brownian motion under the risk-neutral probability measure $\widetilde{\mathbb{P}}$.

By the definition of the risk-neutral measure, we have

$$D_t P(t, T) = \widetilde{\mathbb{E}}[D_T P(T, T)|\mathcal{F}_t],$$

where $\{D_t\}_{t\in[0,\bar{T}]}$ is the discount process given by

$$D_t = e^{-\int_0^t r_s\, ds}.$$

Since the payoff of the bond at maturity is given by $P(T, T) = 1$, we deduce that

$$P(t, T) = \widetilde{\mathbb{E}}\left[\frac{D_T}{D_t}|\mathcal{F}_t\right] = \widetilde{\mathbb{E}}\left[e^{-\int_t^T r_s\, ds}|\mathcal{F}_t\right]. \tag{2}$$

We apply the Feynman-Kac formula to characterize the price of the bond given in (2). Recall that the formula provides a PDE for the function

$$v(t, x) = \mathbb{E}^{t,x}\left[e^{-\int_t^T c(s,X_s)\, ds} f(X_T)\right], \tag{3}$$

where $\{X_t\}_{t\in[0,T]}$ satisfies the SDE

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t.$$

By the Markov property, we also have

$$v(t, X_t) = \mathbb{E}\left[e^{-\int_t^T c(s,X_s)\, ds} f(X_T)|\mathcal{F}_t\right].$$

Therefore, by identifying $f \equiv 1$, $X_t = r_t$, $c(t, r) = r$, we see that $P(t, T) = p(t, r_t)$, where

$$p(t, r) := \widetilde{\mathbb{E}}^{t,r}\left[e^{-\int_t^T r_s\, ds}\right].$$

By the Feynman-Kac formula, we see that $p(t, r)$ satisfies the following PDE:

$$\boxed{\begin{cases} p_t(t, r) + \beta(t, r) p_r(t, r) + \dfrac{1}{2}\gamma^2(t, r) p_{rr}(t, r) = r p(t, r), \\ p(T, r) = 1. \end{cases}} \tag{4}$$

The PDE (4) is called the ***bond-pricing equation*** that characterizes the price of a zero-coupon bond under the one-factor model (1). The domain of r depends on the underlying model, i.e., the choice of β and γ. Below, we present different one-factor models and the corresponding solution to the PDE (4).

We can compute the yield of the zero-coupon bond from its price, defined as follow:

Definition 1.2 The ***yield*** of a zero-coupon bond between time t and T is given by

$$Y(t,T) := -\frac{\ln P(t,T)}{T-t} = -\frac{\ln p(t,r_t)}{T-t}.$$

1.1 Hull–White Model

The ***Hull–White model*** was introduced by John C. Hull and Alan White in 1990. In the Hull–White model, the interest rate $\{r_t\}_{t\in[0,\bar{T}]}$ satisfies the following SDE:

$$dr_t = (a(t) - b(t)r_t)\ dt + \sigma(t)\, d\widetilde{B}_t, \tag{5}$$

where $a, b, \sigma : [0,\bar{T}] \to [0,\infty)$ are deterministic functions. When a, b, σ are constants, (5) is reduced to the ***Vasicek model***. In other words, we have

$$\beta(t,r) = a(t) - b(t)r \quad \text{and} \quad \gamma(t,r) = \sigma(t).$$

Using the method of integrating factor, we can deduce the following:

$$\boxed{\begin{aligned} r_t &= r_0\, e^{-\int_0^t b(s)\,ds} + e^{-\int_0^t b(s)\,ds} \int_0^t e^{\int_0^s b(u)\,du}\, a(s)\,ds + e^{-\int_0^t b(s)\,ds} \int_0^t e^{\int_0^s b(u)\,du}\, \sigma(s)\, d\widetilde{B}_s, \\ \mathbb{E}[r_t] &= r_0\, e^{-\int_0^t b(s)\,ds} + e^{-\int_0^t b(s)\,ds} \int_0^t e^{\int_0^s b(u)\,du}\, a(s)\,ds, \\ \mathrm{Var}[r_t] &= e^{-2\int_0^t b(s)\,ds} \int_0^t \sigma^2(s)\, e^{2\int_0^s b(u)\,du}\,ds. \end{aligned}}$$

The Hull–White model is ***mean-reverting***, with the short rate pulled toward a time-dependent long-term level determined by $a(t)/b(t)$. The function $b(t)$ determines the speed of mean-reversion. However, the Hull–White model allows the short rate to be negative due to its normality. Figure 1 depicts the average short-rate paths (computed from 1,000 simulated trajectories) under different constant parameter specifications (a, b, σ) in the Hull–White model.

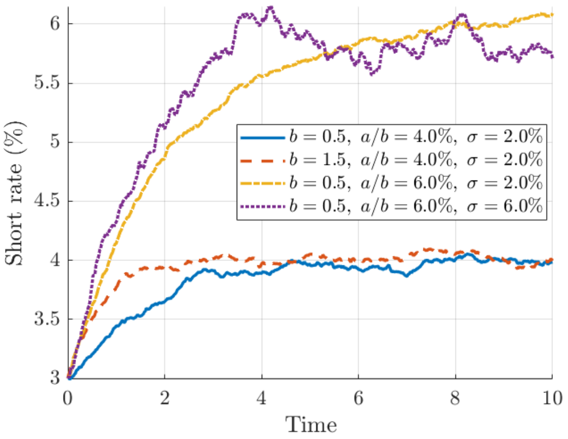

Figure 1: Average short-rate paths (1,000 simulations) under the Hull–White model for different (a, b, σ).

Under the Hull–White model, the bond-pricing equation is reduced to

$$\begin{cases} p_t(t,r) + (a(t) - b(t)r)\, p_r(t,r) + \dfrac{1}{2}\sigma^2(t) p_{rr}(t,r) = rp(t,r), \\ p(T,r) = 1. \end{cases} \tag{6}$$

To solve the PDE (6), we consider the *ansatz*

$$p(t,r) = e^{-rC(t,T) - A(t,T)}, \tag{7}$$

where for fixed $T > 0$, $A(t,T), C(t,T)$ are deterministic functions in time that are to be determined.

We shall derive the expressions of $A(t,T), C(t,T)$ using the PDE (6) and the *ansatz*. By (7), we have

$$\begin{aligned} &p_t(t,r) = \left(-rC_t(t,T) - A_t(t,T)\right) p(t,r), \\ &p_r(t,r) = -C(t,T)p(t,r),\ \ p_{rr}(t,r) = C^2(t,T)p(t,r). \end{aligned} \tag{8}$$

Substituting these into the PDE (6), we have

$$\begin{aligned} 0 &= p_t(t,r) + (a(t) - b(t)r)\, p_r(t,r) + \frac{1}{2}\sigma^2(t) p_{rr}(t,r) - rp(t,r) \\ &= \left[(-rC_t(t,T) - A_t(t,T)) - (a(t) - b(t)r)\, C(t,T) + \frac{1}{2}\sigma^2(t) C^2(t,T) - r \right] p(t,r) \\ &= \left[(-C_t(t,T) + b(t)C(t,T) - 1)\, r - A_t(t,T) - a(t)C(t,T) + \frac{1}{2}\sigma^2(t) C^2(t,T) \right] p(t,r). \end{aligned}$$

Using the fact that $p(t,r) > 0$, the above is satisfied if and only if

$$-C_t(t,T) + b(t)C(t,T) - 1 = 0 \quad \text{and} \quad - A_t(t,T) - a(t)C(t,T) + \frac{1}{2}\sigma^2(t)C^2(t,T) = 0.$$

In addition, using the terminal condition $P(T,T) = 1$, we have

$$1 = P(T,T) = e^{-rC(T,T)-A(T,T)},$$

which yields $A(T,T) = 0 = C(T,T)$. Therefore, we arrive at the following system of ODE:

$$\begin{cases} C_t(t,T) - b(t)C(t,T) + 1 = 0, \\ C(T,T) = 0, \end{cases} \quad \text{and} \quad \begin{cases} A_t(t,T) + a(t)C(t,T) - \dfrac{1}{2}\sigma^2(t)C^2(t,T) = 0, \\ A(T,T) = 0. \end{cases} \tag{9}$$

The solution of the ODE for C can be obtained using the method of integrating factor, which yields:

$$C(t,T) = \int_t^T e^{-\int_t^s b(u)\,du}\,ds.$$

Using this, we can obtain the expression for A by a direct integration:

$$A(t,T) = \int_t^T \left(a(s)C(s,T) - \frac{1}{2}\sigma^2(s)C^2(s,T) \right) ds.$$

Therefore, the price of the bond at time t is given by

$$P(t,T) = p(t,r_t) = e^{-r_tC(t,T)-A(t,T)}.$$

The yield of the bond is given by

$$Y(t,T) = -\frac{\ln P(t,T)}{T-t} = \frac{A(t,T) - r_tC(t,T)}{T-t},$$

which is *affine* in r_t. Hence, the Hull–White model is considered as an example of ***affine yield models***.

1.2 Cox–Ingersoll–Ross Model

The ***Cox–Ingersoll–Ross (CIR) model*** was introduced by John Cox, Jonathan Ingersoll and Stephen Ross in 1985. In the CIR model, the interest rate $\{r_t\}_{t\in[0,\bar{T}]}$ satisfies the following SDE:

$$dr_t = (a - br_t)\;dt + \sigma\sqrt{r_t}\,d\widetilde{B}_t, \tag{10}$$

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

The CIR model is a mean-reverting interest rate model, with $b > 0$ is the speed of mean reversion, and a/b is the long-term mean of the model. Unlike the Hull–White model, the CIR model produces non-negative interest rates whenever $a > 0$; and it satisfies $r_t > 0$ if the ***Feller condition*** is met, i.e., $2a \geq \sigma^2$.

Using the method of integrating factor, we can deduce the following:

$$\boxed{\begin{aligned} \mathbb{E}[r_t] &= r_0 e^{-bt} + \frac{a}{b}\left(1 - e^{-bt}\right), \\ \mathrm{Var}[r_t] &= \frac{\sigma^2}{2b}\left(1 - e^{-bt}\right)\left(r_0 e^{-bt} + \frac{a}{b}\right). \end{aligned}}$$

Figure 2 depicts the average short-rate paths (computed from 1,000 simulated trajectories) under different constant parameter specifications (a, b, σ) in the CIR model. Note that the paths appear smoother than in the Hull-White model because the diffusion scales with the square root of the rate, reducing volatility when rates are low and ensuring positivity.

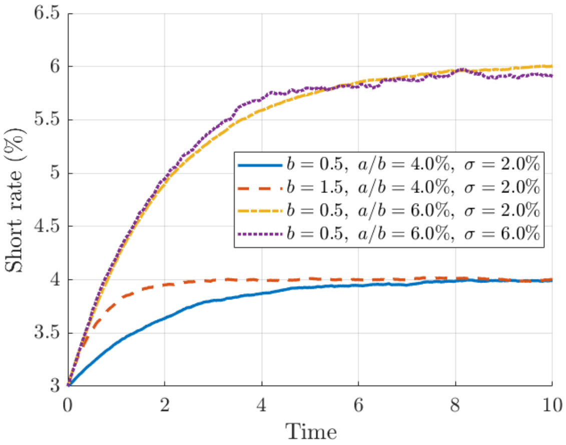

Figure 2: Average short-rate paths (1,000 simulations) under the CIR model for different (a, b, σ).

Under the CIR model, the bond-pricing equation is reduced to

$$\begin{cases} p_t(t,r) + (a - br)\, p_r(t,r) + \dfrac{1}{2}\sigma^2 r p_{rr}(t,r) = r p(t,r), \\ p(T,r) = 1. \end{cases} \tag{11}$$

To solve the PDE (11), we consider the *ansatz*

$$p(t,r) = e^{-rC(t,T) - A(t,T)},$$

which is again affine in the short rate r. Since the form of the *ansatz* is the same as the Hull–White model, we can substitute (8) into (11), which yields

$$\begin{aligned}
0 &= p_t(t,r) + (a-br)\,p_r(t,r) + \frac{1}{2}\sigma^2 r p_{rr}(t,r) - rp(t,r) \\
&= \left[(-rC_t(t,T) - A_t(t,T)) - (a-br)\,C(t,T) + \frac{1}{2}\sigma^2 r C^2(t,T) - r\right] p(t,r) \\
&= \left[\left(-C_t(t,T) + bC(t,T) + \frac{1}{2}\sigma^2 C^2(t,T) - 1\right) r - A_t(t,T) - aC(t,T)\right] p(t,r).
\end{aligned}$$

Therefore, we arrive at the following ODEs:

$$\begin{cases} C_t(t,T) - bC(t,T) - \frac{1}{2}\sigma^2 C^2(t,T) + 1 = 0, \\ C(T,T) = 0, \end{cases} \quad \text{and} \quad \begin{cases} A_t(t,T) + a(t)C(t,T) = 0, \\ A(T,T) = 0. \end{cases} \tag{12}$$

The ODE satisfied by $C(t,T)$ depends on the square term $C^2(t,T)$, which is known as a ***Riccati equation***, whose solution is given by

$$\begin{aligned}
C(t,T) &= \frac{\sinh(\gamma(T-t))}{\gamma\cosh(\gamma(T-t)) + \frac{b}{2}\sinh(\gamma(T-t))}, \\
A(t,T) &= -\frac{2a}{\sigma^2}\ln\left(\frac{\gamma e^{\frac{b}{2}(T-t)}}{\gamma\cosh(\gamma(T-t)) + \frac{b}{2}\sinh(\gamma(T-t))}\right),
\end{aligned}$$

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

1.3 Limitations of One-Factor Models

One-factor short-rate models provide a tractable description of the term structure. In many classical specifications, the yield admits an affine representation of the form

$$Y(t,T) = \frac{A(t,T) - r_t C(t,T)}{T-t} =: \tilde{A}(t,T) - \tilde{C}(t,T) r_t,$$

where the short rate is driven by a single Brownian motion $\tilde{B}$. Despite their analytical convenience, such models impose strong structural restrictions on the joint dynamics of yields across maturities.

1. **Perfect instantaneous correlation across maturities.**

   In practice, zero-coupon bonds of different maturities are traded separately in the market, and their yields are computed from distinct bond prices. Empirically, yield changes across maturities are highly, but not perfectly, correlated.

However, under the affine structure above, the yields for two maturities, $Y(t,T_1)$ and $Y(t,T_2)$, are perfectly correlated via r_t: for $i=1,2$,

$$dY(t,T_i) = \left(\tilde{A}_t(t,T_i) - r_t\tilde{C}_t(t,T_i) - \tilde{C}(t,T_i)\beta(t,r_t)\right)\,dt - \tilde{C}(t,T_i)\gamma(t,r_t)\,d\widetilde{B}_t,$$

which is driven by a common Brownian motion $\widetilde{B}$. This is inconsistent with observed market data.

2. **Restriction on slope and curvature movements.** Empirically, yield curve changes can be decomposed into three factors:
    (a) *Level*: yields across all maturities move up or down together.
    (b) *Slope*: short-term and long-term yields move in opposite directions, causing the curve to steepen or flatten.
    (c) *Curvature*: intermediate-term yields move differently from short- and long-term yields, creating a hump or dip.

   Under a one-factor short-rate model, all stochastic yield changes are driven by $\widetilde{B}$ and lie along one fixed direction. Consequently, the model can produce a parallel shift (level) but cannot generate independent slope or curvature movements observed in the market.

2 Multi-Factor Models

To address the limitations of one-factor models, we consider two-factor extensions of the Hull–White and CIR frameworks, in which the short-rate dynamics are driven by two Brownian motions that may be correlated.

2.1 Two-Factor Vasicek Model

The two-factor Vasicek model is driven by the factors (X_t^1, X_t^2), which satisfies the following dynamics:

$$\begin{cases} dX_t^1 = \left(a_1 - b_{11}X_t^1 - b_{12}X_t^2\right)\,dt + \sigma_1\,d\widetilde{W}_t^1, \\ dX_t^2 = \left(a_2 - b_{21}X_t^1 - b_{22}X_t^2\right)\,dt + \sigma_2\,d\widetilde{W}_t^2, \end{cases} \tag{13}$$

where $\widetilde{W}^1, \widetilde{W}^2$ are standard Brownian motions under $\widetilde{\mathbb{P}}$ with a correlation $\rho \in (-1,1)$, i.e., $d\langle \widetilde{B}^1, \widetilde{B}^2\rangle_t = \rho\,dt$. We assume that $\sigma_1, \sigma_2 > 0$, and the matrix

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

has strictly positive eigenvalues λ_1, λ_2. Consequently, the factor (X^1, X^2) is mean-reverting. The short rate under the two-factor Vasicek model is then given by

$$r_t = \varepsilon_0 + \varepsilon_1 X_t^1 + \varepsilon X_t^2, \tag{14}$$

where $\varepsilon_0, \varepsilon_1, \varepsilon_2 \in \mathbb{R}$.

The representation in (13)–(14) is overparametrized, i.e., there exists other choices of parameters $a_i, b_{ij}, \varepsilon_k$, $i, j = 1, 2$, $k = 0, 1, 2$, that will lead to the same model. In fact, the two-factor Vasicek model admits a canonical representation involving only six parameters $\lambda_1, \lambda_2 > 0$ and $\lambda_{21}, \delta_0, \delta_1, \delta_2 \in \mathbb{R}$, given by

$$\begin{cases} d\phi_t^1 = -\lambda_1 \phi_t^1 \, dt + d\widetilde{B}_t^1, \\ d\phi_t^2 = -\lambda_{21} \phi_t^1 \, dt - \lambda_2 \phi_t^2 \, dt + d\widetilde{B}_t^2, \\ r_t = \delta_0 + \delta_1 \phi_t^1 + \delta_2 \phi_t^2, \end{cases} \tag{15}$$

and $\widetilde{B}^1, \widetilde{B}^2$ are independent standard Brownian motions.

The derivation of (15) is omitted, and we shall directly work with this canonical representation in the sequel. Note that by the joint normality of (ϕ^1, ϕ^2), the resulting interest rate r_t could take negative values.

2.1.1 Bond Prices

By the risk-neutral pricing formula, the price of the zero-coupon bond that matures at T is given by, for $t \in [0, T]$,

$$P(t, T) = \widetilde{\mathbb{E}} \left[e^{-\int_t^T r_s \, ds} \middle| \mathcal{F}_t \right].$$

By the Markov property, we can represent $P(t, T)$ as a function of r_t. Under the two-factor model, r_t is in turn an affine function of ϕ_t^1, ϕ_t^2. Therefore, we can represent the bond price as a function of the factors:

$$P(t, T) = p(t, \phi_t^1, \phi_t^2).$$

Using the Feynman–Kac formula, $p(t, y_1, y_2)$ satisfies the following PDE:

$$\boxed{\begin{cases} p_t(t, y_1, y_2) - \lambda_1 y_1 p_{y_1}(t, y_1, y_2) - (\lambda_{21} y_1 + \lambda_2 y_2) p_{y_2}(t, y_1, y_2) + \dfrac{1}{2} p_{y_1 y_1}(t, y_1, y_2) \\ \qquad\qquad + \dfrac{1}{2} p_{y_2 y_2}(t, y_1, y_2) = (\delta_0 + \delta_1 y_1 + \delta_2 y_2) p(t, y_1, y_2), \\ p(T, y_1, y_2) = 1. \end{cases}} \tag{16}$$

We propose an *ansatz* for the PDE (16):

$$p(t, y_1, y_2) = e^{-C_1(t,T) y_1 - C_2(t,T) y_2 - A(t,T)}, t \in [0, T].$$

This *ansatz* implies that the yield of the bond is affine in the factors (instead of simply in r_t):

$$Y(t, T) = -\frac{\ln P(t, \phi_t^1, \phi_t^2)}{T - t} = \frac{C_1(t, T)\phi_t^1 + C_2(t, T)\phi_t^2 + A(t, T)}{T - t}.$$

For notational convenience, we shall let $\tau := T - t$, $C_1(\tau) = C_1(t,T)$, $C_2(\tau) = C_2(t,T)$, and $A(\tau) = A(t,T)$, as these function shall depend on t via $T - t = \tau$.

Using and *ansatz*, we have

$$\begin{aligned} p_t(t,y_1,y_2) &= \left(C_1'(\tau)y_1 + C_2'(\tau)y_2 + A'(\tau)\right) p(t,y_1,y_2), \\ p_{y_1}(t,y_1,y_2) &= -C_1(\tau)p(t,y_1,y_2),\ p_{y_1y_1}(t,y_1,y_2) = C_1^2(\tau)p(t,y_1,y_2), \\ p_{y_2}(t,y_1,y_2) &= -C_2(\tau)p(t,y_1,y_2),\ p_{y_2y_2}(t,y_1,y_2) = C_2^2(\tau)p(t,y_1,y_2). \end{aligned} \tag{17}$$

Substituting these into (16), we obtain

$$\begin{aligned} 0 &= p(t,y_1,y_2)\Big\{ C_1'(\tau)y_1 + C_2'(\tau)y_2 + A'(\tau) - \lambda_1 y_1(-C_1(\tau)) - (\lambda_{21}y_1 + \lambda_2 y_2)(-C_2(\tau)) \\ &\qquad + \frac{1}{2}C_1^2(\tau) + \frac{1}{2}C_2^2(\tau) - (\delta_0 + \delta_1 y_1 + \delta_2 y_2)\Big\} \\ &= p(t,y_1,y_2)\Big\{ \left[C_1'(\tau) + \lambda_1 C_1(\tau) + \lambda_{21}C_2(\tau) - \delta_1\right] y_1 + \left[C_2'(\tau) + \lambda_2 C_2(\tau) - \delta_2\right] y_2 \\ &\qquad + A'(\tau) + \frac{1}{2}C_1^2(\tau) + \frac{1}{2}C_2^2(\tau) - \delta_0\Big\} \end{aligned}$$

Along with the terminal condition $p(T,y_1,y_2) = 0$, we deduce that $A(0) = C_1(0) = C_2(0) = 0$. Therefore, we arrive at the following ODEs:

$$\begin{cases} C_1'(\tau) = -\lambda_1 C_1(\tau) - \lambda_{21}C_2(\tau) + \delta_1, \\ C_1(0) = 0, \end{cases}$$
$$\begin{cases} C_2'(\tau) = -\lambda_2 C_2(\tau) + \delta_2, \\ C_2(0) = 0, \end{cases}$$
$$\begin{cases} A'(\tau) = -\frac{1}{2}C_1^2(\tau) - \frac{1}{2}C_2^2(\tau) + \delta_0 \\ A(0) = 0. \end{cases}$$

Using the method of integrating factor, it is easy to see that

$$C_2(\tau) = \int_0^\tau \delta_2 e^{\lambda_2(s-\tau)}\, ds = \frac{\delta_2}{\lambda_2}\left(1 - e^{-\lambda_2\tau}\right).$$

Substituting the solution of $C_2(\tau)$ into the ODE of $C_1(\tau)$, using the method of integrating factor, we obtain

$$\begin{aligned} C_1(\tau) &= \int_0^\tau e^{-\lambda_1(\tau-s)}\left(\delta_1 - \lambda_{21}C_2(s)\right) ds \\ &= \int_0^\tau e^{-\lambda_1(\tau-s)}\left(\delta_1 - \frac{\delta_2\lambda_{21}}{\lambda_2}\left(1 - e^{-\lambda_2 s}\right)\right) ds \end{aligned}$$

$$
\begin{aligned}
&= \int_0^\tau \left[\left(\delta_1 - \frac{\delta_2 \lambda_{21}}{\lambda_2} \right) e^{-\lambda_1(\tau - s)} + \frac{\delta_2 \lambda_{21}}{\lambda_2} e^{-\lambda_1 \tau + (\lambda_1 - \lambda_2) s} \right] ds \\
&= \begin{cases} \dfrac{1}{\lambda_1} \left(\delta_1 - \dfrac{\delta_2 \lambda_{21}}{\lambda_2} \right) (1 - e^{-\lambda_1 \tau}) + \dfrac{\delta_2 \lambda_{21}}{\lambda_2 (\lambda_1 - \lambda_2)} (e^{-\lambda_2 \tau} - e^{-\lambda_1 \tau}), & \text{if } \lambda_1 \neq \lambda_2, \\ \dfrac{1}{\lambda_1} \left(\delta_1 - \dfrac{\delta_2 \lambda_{21}}{\lambda_2} \right) (1 - e^{-\lambda_1 \tau}) + \dfrac{\delta_2 \lambda_{21}}{\lambda_2} \tau e^{-\lambda_1 \tau}. & \text{if } \lambda_1 = \lambda_2. \end{cases}
\end{aligned}
$$

Finally, substituting the solution of $C_1(\tau)$ and $C_2(\tau)$ to the ODE of $A(\tau)$, followed by a direct integration, we obtain

$$
A(\tau) = \int_0^\tau \left[-\frac{1}{2} C_1^2(s) - \frac{1}{2} C_2^2(s) + \delta_0 \right] ds.
$$

2.2 Two-Factor CIR Model

The two-factor CIR model is driven by the factors (ϕ^1, ϕ^2), which satisfies the following system of SDEs:

$$
\begin{cases} d\phi_t^1 = (\mu_1 - \lambda_{11} \phi_t^1 - \lambda_{12} \phi_t^2)\, dt + \sqrt{\phi_t^1}\, d\widetilde{B}_t^1, \\ d\phi_t^2 = (\mu_2 - \lambda_{21} \phi_t^1 - \lambda_{22} \phi_t^2)\, dt + \sqrt{\phi_t^2}\, d\widetilde{B}_t^2, \end{cases} \tag{18}
$$

where $\widetilde{B}^1, \widetilde{B}^2$ are independent standard Brownian motions under $\widetilde{\mathbb{P}}$, and the parameters satisfy the following conditions:

$$
\mu_1, \mu_2 \geq 0, \ \ \lambda_{11}, \lambda_{22} > 0, \ \ \lambda_{12}, \lambda_{21} \leq 0. \tag{19}
$$

The interest rate under the two-factor CIR model is then given by

$$
r_t = \delta_0 + \delta_1 \phi_t^1 + \delta_2 \phi_t^2,
$$

where $\delta_0 \geq 0$, $\delta_1, \delta_2 > 0$.

The condition (19) ensures that $\phi_t^1, \phi_t^2 \geq 0$ whenever $\phi_0^1, \phi_0^2 \geq 0$. Heuristically, if $\phi_t^1 = 0$ and $\phi_t^2 \geq 0$, the diffusion term in the dynamics of ϕ_t^1 is zero, whereas drift term of ϕ_t^1 is positive:

$$
\mu_1 - \lambda_{11} \phi_t^1 - \lambda_{12} \phi_t^2 = \mu_1 - \lambda_{12} \phi_t^2 > \mu_1 \geq 0.
$$

Hence, the drift term of ϕ_t^1 will drive up its value and thus ϕ_t^1 is always non-negative. A similar argument applies to ϕ_t^2. With $\delta_0 \geq 0$ and $\delta_1, \delta_2 > 0$, we see that $r_t \geq 0$.

2.2.1 Bond Prices

By the risk-neutral pricing formula, the price of the zero coupon bond at $t \in [0, T]$ is given by

$$
P(t, T) = \widetilde{\mathbb{E}} \left[e^{-\int_t^T r_s\, ds} \middle| \mathcal{F}_t \right].
$$

By the Markov property, the price is Markovian in the factors (ϕ_t^1, ϕ_t^2), i.e., $P(t,T) = P(t, \phi_t^1, \phi_t^2)$. By the Feynman–Kac formula, $p(t, y_1, y_2)$ satisfies the following PDE:

$$\boxed{\begin{cases} p_t(t,y_1,y_2) + (\mu_1 - \lambda_{11}y_1 - \lambda_{12}y_2)p_{y_1}(t,y_1,y_2) + (\mu_2 - \lambda_{21}y_1 - \lambda_{22}y_2)p_{y_2}(t,y_1,y_2) \\ \qquad\qquad + \dfrac{1}{2}y_1 p_{y_1y_1}(t,y_1,y_2) + \dfrac{1}{2}y_2 p_{y_2y_2}(t,y_1,y_2) = (\delta_0 + \delta_1 y_1 + \delta_2 y_2)p(t,y_1,y_2), \\ p(T,y_1,y_2) = 1. \end{cases}} \tag{20}$$

To solve the PDE (20), we again consider the affine-yield *ansatz*:

$$p(t,y_1,y_2) = e^{-C_1(T-t)y_1 - C_2(T-t)y_2 - A(T-t)}, \ t \in [0,T].$$

Since the expression of the *ansatz* is the same as the two-factor Vasicek model, the derivatives of P are also given by (17). Substituting (17) into the PDE and using the notation $\tau := T-t$, we obtain

$$\begin{aligned} 0 = p(t,y_1,y_2)\Bigg\{ & (C_1'(\tau)y_1 + C_2'(\tau)y_2 + A'(\tau)) + (\mu_1 - \lambda_{11}y_1 - \lambda_{12}y_2)(-C_1(\tau)) \\ & + (\mu_2 - \lambda_{21}y_1 - \lambda_{22}y_2)(-C_2(\tau)) + \frac{C_1^2(\tau)y_1 + C_2^2(\tau)y_2}{2} - (\delta_0 + \delta_1 y_1 + \delta_2 y_2)\Bigg\} \\ = p(t,y_1,y_2)\Bigg\{ & \left[C_1'(\tau) + \lambda_{11}C_1(\tau) + \lambda_{21}C_2(\tau) + \frac{1}{2}C_1^2(\tau) - \delta_1\right] y_1 \\ & + \left[C_2'(\tau) + \lambda_{12}C_1(\tau) + \lambda_{22}C_2(\tau) + \frac{1}{2}C_2^2(\tau) - \delta_2\right] y_2 \\ & + [A'(\tau) - \mu_1 C_1(\tau) - \mu_2 C_2(\tau) - \delta_0]\Bigg\}. \end{aligned}$$

By the terminal condition $P(T,T) = 1$, we have $A(0) = C_1(0) = C_2(0) = 0$. Therefore, we arrive at the following system of ODEs:

$$\begin{cases} C_1'(\tau) = -\lambda_{11}C_1(\tau) - \lambda_{21}C_2(\tau) - \dfrac{1}{2}C_1(\tau)^2 + \delta_1, \\ C_2'(\tau) = -\lambda_{12}C_1(\tau) - \lambda_{22}C_2(\tau) - \dfrac{1}{2}C_2(\tau)^2 + \delta_2, \\ C_1(0) = 0, \\ C_2(0) = 0, \end{cases} \tag{21}$$
$$\begin{cases} A'(\tau) = \mu_1 C_1(\tau) + \mu_2 C_2(\tau) + \delta_0, \\ A(0) = 0. \end{cases}$$

The ODE satisfied by the pair $(C_1(\tau), C_2(\tau))$ in (21) is a ***non-symmetric Riccati equation***, whose solution can be derived using Radon's lemma. To this end, we define the following 2×2 matrices:

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix}, \quad \Delta = \operatorname{diag}(\delta_1, \delta_2),$$

and a 4×4 matrix $\mathbf{M}$ by

$$\mathbf{M} = \begin{pmatrix} \mathbf{0}_{2\times 2} & \frac{1}{2} I_2 \\ \Delta & -\mathbf{\Lambda}^\top \end{pmatrix}. \tag{22}$$

Proposition 2.1 (Radon's lemma on Riccati differential equations) Denote the matrix exponential as

$$\exp(\mathbf{M}\tau) = \begin{pmatrix} \mathcal{X}_{11}(\tau) & \mathcal{X}_{12}(\tau) \\ \mathcal{X}_{21}(\tau) & \mathcal{X}_{22}(\tau) \end{pmatrix}. \tag{23}$$

The solution to the Riccati equations for $\mathbf{C}(\tau) = (C_1(\tau), C_2(\tau))^\top$ is given by the diagonal entries of $\mathcal{X}_{21}(\tau)\mathcal{X}_{11}(\tau)^{-1}$, i.e.,

$$\mathbf{C}(\tau) = \operatorname{diag}\left(\mathcal{X}_{21}(\tau)\mathcal{X}_{11}(\tau)^{-1}\right). \tag{24}$$

Using the solution $\mathbf{C}(\tau)$, we can compute $A(\tau)$ by

$$A(\tau) = \int_0^\tau \left(\mu_1 C_1(s) + \mu_2 C_2(s) + \delta_0\right) ds.$$

3 Heath—Jarrow—Morton Model

In this section, we introduce the ***Heath–Jarrow–Morton (HJM) model***. Rather than modeling the short rate directly, the HJM framework describes the evolution of the entire yield curve by modeling the instantaneous forward rate. This approach aligns with market practice, where the term structure is inferred from the prices of zero-coupon bonds; from these prices, yields and forward rates are derived, and the short rate arises naturally as the limit of the forward rate as the maturity approaches the current time.

3.1 Forward Rates

Definition 3.1 Let $P(t, T)$ be the price of a zero-coupon bond at time t that matures at $T > t$. The ***forward rate*** $f(t, T)$ at time t for investing at time T is defined by

$$f(t, T) = -\lim_{\delta \downarrow 0} \frac{\ln P(t, T + \delta) - \log P(t, T)}{\delta} = -\frac{\partial}{\partial T} \ln P(t, T).$$

Remark 3.1.

1. If the interest rate is positive, then $P(t, T+\delta) < P(t,T)$ (i.e., bonds with longer maturity tend to be cheaper), then $f(t,T) > 0$.
2. By integrating the forward rate over the maturity, the bond price can be computed using the forward rate via

$$P(t,T) = \exp\left(-\int_t^T f(t,s)\,ds\right). \tag{25}$$

Using the forward rate, the interest rate r_t is defined as

$$r_t = f(t,t),$$

i.e., the interest rate is the instantaneous forward rate at time t.

Given a finite horizon $[0,\bar{T}]$, $\bar{T} > 0$, the mapping $T \in [0,\bar{T}] \mapsto f(0,T)$ is called the *initial forward rate curve*, which represents the rate of borrowing at time 0 with different maturity. The HJM framework models the forward rate $f(t,T)$, $t \leq T$, via the following SDE:

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)\,ds + \int_0^t \sigma(s,T)\,dB_t,$$

where $\{B_t\}_{t\in[0,\bar{T}]}$ is a standard Brownian motion under the real-world measure $\mathbb{P}$, and $\alpha(\cdot,T)$, $\sigma(\cdot,T)$ represents the drift and volatility of the forward rate under $\mathbb{P}$, respectively, which themselves can be random processes. The equivalent differential form of the forward rate is given by

$$df(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dB_t. \tag{26}$$

Using (26) and (25), we can derive the dynamics of the bond price $\{P(t,T)\}_{t\in[0,T]}$ as follows. Let

$$Y_t := -\int_t^T f(t,s)\,ds.$$

By the Leibniz integral rule,

$$\begin{aligned} dY_t &= f(t,t)\,dt - \int_t^T df(t,s)\,ds \\ &= r_t\,dt - \int_t^T \left[\alpha(t,s)\,dt + \sigma(t,s)\,dB_t\right]\,ds \\ &= (r_t - \alpha^*(t,T))\,dt - \sigma^*(t,T)\,dB_t, \end{aligned}$$

where

$$\alpha^*(t,T) := \int_t^T \alpha(t,s)\,ds \quad \text{and} \quad \sigma^*(t,T) := \int_t^T \sigma(t,s)\,ds. \tag{27}$$

Therefore, by Itô's lemma,

$$\begin{aligned}
dP(t,T) &= de^{Y_t} \\
&= e^{Y_t} dY_t + \frac{1}{2} e^{Y_t} d\langle Y \rangle_t \\
&= P(t,T)\left((r_t - \alpha^*(t,T))\,dt - \sigma^*(t,T)\,dB_t\right) + \frac{1}{2} P(t,T)\left(\sigma^*(t,T)\right)^2\,dt \\
&= P(t,T)\left(r_t - \alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2\right)dt - \sigma^*(t,T)P(t,T)\,dB_t. \qquad (28)
\end{aligned}$$

3.2 Risk-Neutral Measure

Under the HJM framework, we model the forward rate $f(t,T)$, and thus the associated bond prices $P(t,T)$, for any maturity $T \in (0,\bar{T}]$. However, one needs to ensure such construction would not lead to any market arbitrage, i.e., the bond prices with different maturities have to be "consistent".

Invoking the first fundamental theorem of asset pricing, the bond market under the HJM model is arbitrage-free if one can construct a probability measure (i.e., the risk-neutral measure) $\widetilde{\mathbb{P}}$, under which the discounted bond price

$$\{D_t P(t,T)\}_{t\in[0,T]} = \left\{ e^{-\int_0^t r_s\,ds} P(t,T) \right\}_{t\in[0,T]},$$

is a $\widetilde{\mathbb{P}}$-martingales simultaneously for all maturity $T \in (0,\bar{T}]$.

To construct such a measure $\widetilde{\mathbb{P}}$, by Itô's lemma and (28), we have

$$\begin{aligned}
d\left(D_t P(t,T)\right) &= P(t,T)\,dD_t + D_t\,dP(t,T) \\
&= D_t P(t,T)\left[\left(\alpha^*(t,T) + \frac{1}{2}\left(\sigma^*(t,T)\right)^2\right)dt - \sigma^*(t,T)\,dB_t\right]. \qquad (29)
\end{aligned}$$

Suppose that there exists a process $\{\theta_t\}_{t\in[0,\bar{T}]}$ such that, for any $T \in (0,\bar{T}]$ and $t \in [0,T]$, it holds that

$$\alpha^*(t,T) + \frac{1}{2}\left(\sigma^*(t,T)\right)^2 = -\sigma^*(t,T)\theta_t. \qquad (30)$$

Then, we can rewrite (29) as

$$d\left(D_t P(t,T)\right) = -\sigma^*(t,T) D_t P(t,T)\left[\theta_t\,dt + dB_t\right] = -\sigma^*(t,T) D_t P(t,T)\,d\widetilde{B}_t, \qquad (31)$$

where

$$\widetilde{B}_t := \int_0^t \theta_s\,ds + B_t.$$

As such, we can define $\widetilde{\mathbb{P}}$ by, for any $A \in \mathcal{F}_{\bar{T}}$, $\widetilde{\mathbb{P}}(A) = \mathbb{E}[Z_{\bar{T}} \mathbb{1}_A]$, where

$$Z_{\bar{T}} = \exp\left(-\int_0^{\bar{T}} \theta_t \, dt - \frac{1}{2} \int_0^{\bar{T}} \theta_t^2 \, dt \right).$$

By Girsanov's theorem, $\widetilde{\mathbb{P}}$ is a probability measure, under which $\{\widetilde{B}_t\}_{t\in[0,\bar{T}]}$ is a standard Brownian motion. By (31), the discounted bond price $\{D_t P(t,T)\}_{t\in[0,T]}$ would be a $\widetilde{\mathbb{P}}$-martingale for any maturity $T \in (0, \bar{T}]$.

Therefore, the bond market is arbitrage-free if there exists a process $\{\theta_t\}_{t\in[0,\bar{T}]}$, known as the ***market price of risk***, such that it solves the ***market equations*** (30). Note that the market equations indeed consist of infinitely many equations, since we require a single θ_t, independent of the maturity, that solves (30) for all $T \in (0, \bar{T}]$.

Theorem 3.2 The bond market is arbitrage-free if there exists an adapted process $\{\theta_t\}_{t\in[0,\bar{T}]}$ that solves the market equations (30) for all maturity $T \in (0, \bar{T}]$. In particular, (30) holds if $\{\theta_t\}_{t\in[0,\bar{T}]}$ solves the following equation:

$$-\alpha(t,T) + \sigma^*(t,T)\sigma(t,T) = -\sigma(t,T)\theta_t, \ \ t \in [0,T], \ \ T \in (0,\bar{T}]. \tag{32}$$

Proof. The first part of the theorem has been proven based on the discussion preceding it. Hence, it suffices to prove that, if $\{\theta_t\}_{t\in[0,\bar{T}]}$ solves (32), it also solves (30).

By replacing T with s in (32), and integrating the resulting equation from $s = t$ to $s = T$, we have

$$\begin{aligned} & -\int_t^T \alpha(t,s)\,ds + \int_t^T \sigma^*(t,s)\sigma(t,s)\,ds = -\int_t^T \sigma(t,s)\theta_t\,ds \\ \Rightarrow & -\alpha^*(t,T) + \int_t^T \sigma^*(t,s)\sigma(t,s)\,ds = -\sigma^*(t,T)\theta_t, \end{aligned} \tag{33}$$

where we have used the definition of α^* and σ^*; see (27). In addition,

$$\sigma(t,s) = \frac{d}{ds}\sigma^*(t,s).$$

Hence,

$$\int_t^T \sigma^*(t,s)\sigma(t,s)\,ds = \int_t^T \sigma^*(t,s) d\left(\sigma^*(t,s)\right) = \frac{1}{2}\left(\sigma^*(t,s)\right)^2 \Big|_{s=t}^{s=T} = \frac{1}{2}\left(\sigma^*(t,T)\right)^2,$$

since $\sigma^*(t,t) = 0$. Substituting this into (33), we arrive at (30). $\square$

If (32) is solvable and $\sigma(t,T) > 0$, the solution $\{\theta_t\}_{t\in[0,\bar{T}]}$ must be unique. Indeed, $\{\theta_t\}_{t\in[0,\bar{T}]}$ is uniquely given by

$$\theta_t = \frac{\alpha(t,T)}{\sigma(t,T)} - \sigma^*(t,T).$$

Therefore, the second fundamental theorem of asset pricing implies that the bond market is complete, i.e., all interest rate derivatives are hedgeable by trading zero-coupon bonds with different maturities.

Under the risk-neutral measure $\widetilde{\mathbb{P}}$, using (32), we can express the forward rate dynamics as follows:

$$\begin{aligned} df(t,T) &= \alpha(t,T)\,dt + \sigma(t,T)\,dB_t \\ &= \left(\alpha(t,T) - \sigma(t,T)\theta_t\right)\,dt + \sigma(t,T)\,d\widetilde{B}_t \\ &= \sigma^*(t,T)\sigma(t,T)\,dt + \sigma(t,T)\,d\widetilde{B}_t. \end{aligned} \tag{34}$$

Therefore, under $\widetilde{\mathbb{P}}$, the drift term and the dynamics of $f(t,T)$ is uniquely determined by the volatility structure $\sigma(t,T)$ through the no-arbitrage condition. Consequently, the HJM framework reduces the modeling problem to the specification and calibration of the volatility function $\sigma(t,T)$.

3.3 Connection with the Affine Yield Model

In this subsection, we show an affine-yield term structure model can be identified as a HJM model with the associated coefficients α and σ for the forward rate dynamics.

For simplicity, we consider the one-factor model (Hull–White and CIR). In these models, the interest rate $\{r_t\}_{t\in[0,\bar{T}]}$ and the bond price $\{P(t,T)\}_{t\in[0,T]}$ are given by

$$\begin{aligned} dr_t &= \beta(t,r_t)\,dt + \gamma(t,r_t)\,d\widetilde{B}_t, \\ P(t,T) &= e^{-r_tC(t,T)-A(t,T)}, \end{aligned}$$

for some functions A and C.

Using (25), we have

$$f(t,T) = -\frac{d}{dT}\ln P(t,T) = r_t\frac{\partial}{\partial T}C(t,T) + \frac{\partial}{\partial T}A(t,T).$$

By Itô's lemma,

$$\begin{aligned} df(t,T) &= \frac{\partial}{\partial T}C(t,T)\,dr_t + r_t\frac{\partial}{\partial T}C'(t,T)\,dt + \frac{\partial}{\partial T}A'(t,T)\,dt \\ &= \frac{\partial}{\partial T}C(t,T)\left(\beta(t,r_t)\,dt + \gamma(t,r_t)\,d\widetilde{B}_t\right) + r_t\frac{\partial}{\partial T}C'(t,T)\,dt + \frac{\partial}{\partial T}A'(t,T)\,dt \end{aligned}$$

$$= \left[\beta(t, r_t)\frac{\partial}{\partial T}C(t,T) + r_t\frac{\partial}{\partial T}C'(t,T) + \frac{\partial}{\partial T}A'(t,T)\right] dt + \gamma(t, r_t)\frac{\partial}{\partial T}C(t,T)\, d\widetilde{B}_t.$$

By identifying with the HJM forward rate model, we have

$$\begin{cases} \sigma(t,T) = \gamma(t, r_t)\dfrac{\partial}{\partial T}C(t,T), \\ \sigma^*(t,T)\sigma(t,T) = \beta(t, r_t)\dfrac{\partial}{\partial T}C(t,T) + r_t\dfrac{\partial}{\partial T}C'(t,T) + \dfrac{\partial}{\partial T}A'(t,T) \end{cases}$$

Substituting the first equation to the second, and using the fact that $\sigma^*(t,T) = \int_t^T \sigma(t,s)\, ds$, we have

$$\begin{aligned} &\beta(t, r_t)\frac{\partial}{\partial T}C(t,T) + r_t\frac{\partial}{\partial T}C'(t,T) + \frac{\partial}{\partial T}A'(t,T) \\ &= \gamma(t, r_t)\frac{\partial}{\partial T}C(t,T)\int_t^T \left(\gamma(t, r_t)\frac{\partial}{\partial s}C(t,s)\right) ds \\ &= \gamma^2(t, r_t)\frac{\partial}{\partial T}C(t,T)\,[C(t,s)]_{s=t}^{s=T} \\ &= \gamma^2(t, r_t)\frac{\partial}{\partial T}C(t,T)\,[C(t,T) - C(t,t)] \\ &= \gamma^2(t, r_t)C(t,T)\frac{\partial}{\partial T}C(t,T), \end{aligned} \tag{35}$$

since $C(t,t) = 0$.

It can be verified that both the Hull–White model (including the Vasicek model) and the CIR model satisfy (35). Therefore, these models can be identified as no-arbitrage HJM models.

Further Readings

1. Implementations of the HJM model;
2. London interbank offered rate (LIBOR) model;
3. Forward measures;
4. Caplets and their pricing formula.