The Heath-Jarrow-Morton model for forward rates

Math 622

April 22, 2015

1 Set up of the HJM model

We start with a probability space, $(\Omega, \mathcal{F}, \mathbf{P})$, which supports a Brownian motion $W(t) = (W_1(t), \ldots, W_d(t))$. Let $\{\mathcal{F}(t);\, t \geq 0\}$ be the filtration generated by W.

The HJM model for the forward rate is

$$df(t,T) = \alpha(t,T)\,dt + \sum_{j=1}^{d} \sigma_j(t,T)\,dW_j(t), \qquad 0 \leq t \leq T \leq \bar{T}, \tag{1}$$

where for each T, $\{\alpha(t,T);\, 0 \leq t \leq T\}$ and $\{\gamma_1(t,T);\,;\, 0 \leq t \leq T\}, \ldots, \{\gamma_d(t,T);\, 0 \leq t \leq T\}$ are $\{\mathcal{F}(t);\, t \geq 0\}$-adapted process for each T, $0 < T \leq \bar{T}$.

The interpretation is that we fix a time horizon $\bar{T}$ and we investigate all bonds with expiry T, $T \leq \bar{T}$.

At time $t = 0$, we know the forward rate curve, $T \to f(0,T)$ from market price quotes. This provides an initial condition for equation (1) for each T, and by integrating (1) forward in time,

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T)\,du + \int_0^t \sum_{j=1}^{d} \sigma_j(u,T)\,dW_j(u). \tag{2}$$

In particular, the short rate is

$$R(t) = f(t,t) = f(0,t) + \int_0^t \alpha(u,t)\,du + \int_0^t \sum_{j=1}^{d} \sigma_j(u,t)\,dW_j(u).$$

From now on, for simplicity, we study the case in which $d = 1$:

$$df(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dW(t), \qquad 0 \leq t \leq T \leq \bar{T}. \tag{3}$$

2 Equations for zero-coupon bond prices

What stochastic differential equation for $B(t,T)$ is implied by the HJM model? To answer this, set

$$Y(t,T) = -\int_t^T f(t,v)\,dv.$$

Since

$$B(t,T) = e^{Y(t,T)},$$

Itô's rule implies

$$dB(t,T) = B(t,T)\,dY(t,T) + \frac{1}{2}B(t,T)[dY(t,T)]^2,$$

so to derive an equation for $dB(t,T)$ we need only to compute $dY(t,T)$.

Remark: Note that by $dY(t,T)$ and $dB(t,T)$ here we mean the differentials of Y and B in t, not T. That is for δ small, we mean

$$\begin{aligned} dY(t,T) &\approx Y(t+\delta,T) - Y(t,T); \\ dB(t,T) &\approx B(t+\delta,T) - B(t,T). \end{aligned}$$

Now to calculate $dY(t,T)$, since this is the differential in t, note that t appears in 2 places in the formula

$$Y(t,T) = -\int_t^T f(t,v)dv.$$

If $f(t,v)$ is differentiable in t, then the computation falls under the Newton-Leibniz formula. Here we have a differential form of f, but a similar version of that formula still holds, and thus

$$dY(t,T) = f(t,t)dt - \int_t^T [df(t,v)]dv.$$

Define

$$\alpha^*(t,T) = \int_t^T \alpha(t,v)\,dv \quad \text{and} \quad \sigma^*(t,T) = \int_t^T \sigma(t,v)\,dv.$$

Substituting $df(t,v)$ by (3) and by switching the order of integration if follows that

$$\begin{aligned} dY(t,T) &= f(t,t)dt - \int_t^T \Big[\alpha(t,v)\,dt + \sigma(t,v)\,dW(t)\Big]dv \\ &= f(t,t)dt - \Big[\int_t^T \alpha(t,v)\,dv\Big]dt - \Big[\int_t^T \sigma(t,v)\,dv\Big]dW(t) \\ &= f(t,t)dt - \alpha^*(t,T)dt - \sigma^*(t,T)dW(t). \end{aligned}$$

Thus

$$\begin{aligned} Y(t,T) &= Y(0,T) + \int_0^t f(u,u)\,du - \int_0^t \alpha^*(u,T)\,du - \int_0^t \sigma^*(u,T)\,dW(u) \\ &= -\int_0^T f(0,u)\,du + \int_0^t f(u,u)\,du - \int_0^t \alpha^*(u,T)\,du - \int_0^t \sigma^*(u,T)\,dW(u). \end{aligned} \tag{4}$$

Since $f(t,t) = R(t)$, it follows that

$$dY(t,T) = \Big[R(t) - \alpha^*(t,T)\Big]\,dt - \sigma^*(t,T)\,dW(t), \tag{5}$$

and hence that

$$dB(t,T) = B(t,T)\Big[R(t) - \alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2\Big]\,dt - B(t,T)\sigma^*(t,T)\,dW(t). \tag{6}$$

It is interesting to note that $\sigma^*(T,T) = 0$. This makes sense: the volatility of $B(T,T)$ is zero because $B(T,T) = 1$.

Remark: Recall the dynamics for $f(t,T)$ is

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t). \tag{7}$$

It is important to observe that equations (5), (6), (7) are equivalent. That is if we have the dynamics of $B(t,T)$ following (6) then the dynamics of $Y(t,T)$ and $f(t,T)$ have to follow (5) and (7) respectively. The other cases are similar. The reason is because $B(t,T)$, $f(t,T)$ and $Y(t,T)$ are connected via the relations

$$\begin{aligned} B(t,T) &= e^{Y(t,T)} \\ \frac{\partial Y}{\partial T} &= -f(t,T). \end{aligned}$$

The implication of this will be seen in Section (4), where we conclude that the drift of $f(t,T)$ have to be $\sigma(t,T)\sigma^*(t,T)$ under the risk neutral measure, even when we derive f from a short rate model, for example the Hull-White model.

3 Existence of a risk-neutral model

In order that there be no arbitrage in the HJM model it is necessary that there exists an equivalent probability measure $\widetilde{\mathbf{P}}$ with respect to which $\{D(t)B(t,T);\, t \leq T\}$ is a martingale for all $0 \leq T \leq \bar{T}$. Note that this is a subtle condition since for each T, we have a different financial product $B(t,T)$. So this statement requires that the discounted price of *infinitely many* financial products must be $\widetilde{\mathbf{P}}$ martingale. But in our model, we only have finitely many sources of noise. This has implication about the existence of solution to the market price of risk equation. (Recall from what we have learned that if the number of assets m is less than the number of random sources d then the risk neutral measure many not exist).

From the previous section, we have the dynamics of the bond $B(t,T)$ under the HJM model under the physical measure $\mathbf{P}$:

$$dB(t,T) = B(t,T)\left[R(t) - \alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2\right] dt - B(t,T)\sigma^*(t,T)\, dW(t).$$

This is a geometric Brownian motion model with drift term

$$\gamma(t,T) := R(t) - \alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2$$

and volatility $-\sigma^*(t,T)$.

Thus, the market price of risk equation for this *particular bond with expiry* T is: to find an adapted process $\theta(t)$ such that

$$\begin{aligned} -\sigma^*(t,T)\theta(t) &= \gamma(t,T) - R(t) \\ &= \left[R(t) - \alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2\right] - R(t) \\ &= -\alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2. \end{aligned}$$

Or equivalently

$$\sigma^*(t,T)\theta(t) = \alpha^*(t,T) - \frac{1}{2}(\sigma^*(t,T))^2.$$

Note that $\theta(t)$ does not depend on T (the risk neutral measure, which is defined from θ should not depend on an expiry). On the other hand, for each T, we have such an equation and thus this gives an infinite system of equations for $\theta(t)$.

By differentiating both sides of the above equation with respect to T, this reduces to the equivalent and simpler sufficient condition:

$$\sigma(t,T)\theta(t) = \alpha(t,T) - \sigma^*(t,T)\sigma(t,T), \qquad \text{for all } 0 \leq t \leq T \leq \bar{T}. \tag{8}$$

Indeed, if $\theta(t)$ satisfies (8) then by taking the anti-derivative (in T) we get

$$\sigma^*(t,T)\theta(t) = \alpha^*(t,T) - \frac{1}{2}(\sigma^*(t,T))^2 + C(T),$$

where $C(T)$ is a constant depending on T. Plug in $t = T$ we conclude $C(T) = 0$ and thus $\theta(t)$ solves the market price of risk equation.

Therefore, if equation (8) has a solution and if

$$E[\exp\{-\int_0^{\bar{T}} \theta(u)\, dW(u) - \int_0^{\bar{T}} \theta^2(u)\, du\}] = 1,$$

(this is a technical condition to make sure the change of measure is well-defined) then there is an equivalent risk-neutral measure.

Again, (8) places heavy restrictions on the relation of $\alpha(t,T)$ to $\sigma(t,T)$. If $\sigma(t,T)$ is non-zero always then $\theta(t) = \sigma^{-1}(t,T)\alpha(t,T) - \sigma^*(t,T)$ and the right-hand side *must not actually depend on* T. Thus $\alpha(t,T)$ and $\sigma(t,T)$ must be related in a very special way.

You are asked to derive a version of (8) when W is multidimensional in Shreve, Exercise 10.9.

4 The risk-neutral form of HJM

Assume that the market price of risk equations of (8) have a solution θ. Then under the risk-neutral measure, $\widetilde{W}(t) = W(t) + \int_0^t \theta(u)\, du$ is a Brownian motion, and

$$\begin{aligned} df(t,T) &= \alpha(t,T)\, dt + \sigma(t,T)\, [d\widetilde{W}(t) - \theta(t)\, dt] \\ &= [\alpha(t,T) - \sigma(t,T)\theta(t)]\, dt + \sigma(t,T)\, d\widetilde{W}(t) \\ &= \sigma^*(t,T)\sigma(t,T)\, dt + \sigma(t,T)\, d\widetilde{W}(t) \end{aligned} \tag{9}$$

Here is an important observation for the risk neutral form of the forward rate: *there is no freedom in choosing the drift of the forward rate under the risk-neutral model.* This, in some sense is similar to the situation we are used to with modeling the asset in Black-Scholes framework, where the drift term of the asset must be $R(t)S(t)dt$ under the risk neutral measure.

For the forward rate, the drift must be $\sigma^*(t,T)\sigma(t,T)$, which is completely determined by the volatility $\sigma(t,T)$. This fact limits the nature of the models that can be proposed for the forward rate.

For example, suppose we wanted to have a lognormal model for $f(t,T)$ uder the risk-neutral measure (So that we can have explicit computation for the distribution of $f(t,T)$ and possibly $B(t,T)$). A reasonable way to approach this would be to let $\sigma(t,T) = \sigma f(t,T)$, where σ is a constant. (This is possibly the simplest way to propose the volatility of a log-normal model).

However, if we do this the drift term will become

$$\sigma^*(t,T)\sigma(t,T) = \sigma^2(t)f(t,T)\int_t^T f(t,v)\,dv$$

which is *not a linear function* of $f(t,T)$. (So even if we want, a log-normal model under the risk neutral measure is not possible for the forward rate).

Even worse, with this drift (9) will not even have a solution defined for all sample paths (near the expiry T the forward rate may get very large - due to the drift being almost like quadratic in $f(t,T)$. See Shreve's Section 10.4.1 for the more details. This is also a motivating example for us to switch our attention to the forward LIBOR model in the 2nd part of this lecture.

On the other hand, we can use (9) to define *a risk-neutral model directly.* We suppose the Brownian motion $\widetilde{W}$ on $(\Omega, \mathcal{F}, \widetilde{\mathbf{P}})$ to be given, we assume we have a $\sigma(t,T)$ such that (9) *has a solution* (in the case when $\sigma(t,T)$ depends on $f(t,T)$), and this gives us a risk-neutral model with

$$R(t) = f(t,t) = f(0,t) + \int_0^t \sigma^*(u,t)\sigma(u,t)\,du + \int_0^t \sigma(u,t)\,d\widetilde{W}(u),$$

and with

$$B(t,T) = \exp\{-\int_t^T f(t,v)\,dv\}$$

satisfying

$$dB(t,T) = R(t)B(t,T)\,dt - \sigma^*(t,T)B(t,T)\,d\widetilde{W}(t) \qquad (10)$$

(We know this has to be the right equation for $B(t,T)$ from equation (6) and from the fact that $D(t)B(t,T)$ is a martingale under $\widetilde{\mathbf{P}}$.)

5 Example of HJM models

The fact that the market price of risk equations

$$\sigma(t,T)\theta(t) = \alpha(t,T) - \sigma^*(t,T)\sigma(t,T), \qquad \text{for all } 0 \le t \le T \le \bar{T}$$

is an infinite systems for $\theta(t)$ might give us the impression that it is not easy to have an example of HJM model, except perhaps for trivial cases when α and σ do not depend on T. But this is not the case. All previous short rate models we discussed (Vasicek, CIR and also the one factor model Hull-White) are examples of HJM model as we will see. The key is again, if we *start directly under the risk neutral measure* (which is the case for all short rate models) then we do not have the market price of risk equation to deal with.

More generally, we can always obtain a risk-neutral HJM model as described in the previous section, if we choose an exogenously determined volatility $\sigma(t,T)$ (that is, a volatility that *is not a function of* $f(t,T)$ but is defined *a priori.*

Indeed this will be so for any affine yield model based on a model for the short rate, under the risk-neutral measure.

$$B(t,T) = \exp\{-\sum_{j=1}^{m} C_j(t,T)X_j(t) - A(t,T)\},$$

where $X_1(t), \ldots, X_m(t)$ solves a system of stochastic differential equations driven by $\widetilde{W}$ (possibly multi-dimensional), and $C_1(t,T), \ldots, C_m(t,T)$, and $A(t,T)$ are differentiable in T. Then

$$f(t,T) = \sum_{j=1}^{m} X_j(t)\frac{\partial C_j(t,T)}{\partial T} + \frac{\partial A(t,T)}{\partial T}$$

and

$$df(t,T) = \sum_{j=1}^{m} d[\frac{\partial C_j(t,T)}{\partial T} X_j(t)] + d[\frac{\partial A(t,T)}{\partial T}]$$

This enables one to recover an HJM type model for the forward rate.

5.1 Example with one-factor model

For simpliciy, suppose $m = 1$ (that is we have a one-factor model), from which $R(t)$ has dynamics under $\widetilde{\mathbf{P}}$

$$dR(t) = \beta(t)dt + \gamma(t)d\widetilde{W}(t),$$

where we do allow β, γ to depend on $R(t)$, we just do not explicitly write it out that way.

The form of the bond is

$$B(t,T) = \exp\{-C(t,T)R(t) - A(t,T)\},$$

where, $C(t,T)$ and $A(t,T)$ are functions of β, γ (that is obtained from the risk neutral pricing formula).

The forward rate $f(t,T)$ is derived from $B(t,T)$ via the relation

$$f(t,T) = -\frac{\partial}{\partial T}\log B(t,T) = R(t)\frac{\partial}{\partial T}C(t,T) + \frac{\partial}{\partial T}A(t,T).$$

Therefore, $f(t,T)$ has the dynamics

$$\begin{aligned} df(t,T) &= d[\frac{\partial C(t,T)}{\partial T}R(t)] + d[\frac{\partial A(t,T)}{\partial T}] \\ &= \Big[\frac{\partial C(t,T)}{\partial T}\beta(t) + R(t)\frac{\partial C'(t,T)}{\partial T} + \frac{\partial A'(t,T)}{\partial T}\Big]dt + \frac{\partial C(t,T)}{\partial T}\gamma(t)d\widetilde{W}(t), \end{aligned}$$

where $C'(t,T), A'(t,T)$ denotes the derivative of these functions with respect to t.

In order that this be a valid HJM model, the drift term must be related to the volatility term as in equation (9). That is denoting

$$\sigma(t,T) = \frac{\partial C(t,T)}{\partial T}\gamma(t),$$

then it must follow that

$$\frac{\partial C(t,T)}{\partial T}\beta(t) + R(t)\frac{\partial C'(t,T)}{\partial T} + \frac{\partial A'(t,T)}{\partial T} = \sigma(t,T)\sigma^*(t,T).$$

But you do not really need to check it to know that it must be true, because the Hull-White model was created under the risk-neutral measure so that $D(t)B(t,T)$ was automatically a martingale. And we know from the analysis of the previous section that if $D(t)B(t,T)$ is a martingale under the risk-neutral measure and if $df(t,T)$ has a stochastic differential, it must be of the form in (9).

Indeed, as discussed in the remark at the end of Section (2), once we know the dynamics of the bond $B(t,T)$, then the dynamics of $f(t,T)$ is forced. Recall the dynamics of $B(t,T)$ and $f(t,T)$ as discussed in Section (2) under the physical measure P:

$$\begin{aligned} dB(t,T) &= B(t,T)\Big[R(t) - \alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2\Big]\,dt - B(t,T)\sigma^*(t,T)\,dW(t); \\ df(t,T) &= \alpha(t,T)dt + \sigma(t,T)dW(t). \end{aligned}$$

Since the derivation is via Ito's formula, this relation also holds under the risk neutral measure $\widetilde{\mathbf{P}}$, we just need to change $W(t)$ to $\widetilde{W}(t)$ and modify the drift accordingly. Under risk neutral, the drift of $B(t,T)$ is $B(t,T)R(t)$. Thus one must have

$$\alpha^*(t,T) = \frac{1}{2}(\sigma^*(t,T))^2.$$

But this implies

$$\alpha(t,T) = \frac{\partial \alpha^*(t,T)}{\partial T} = \sigma(t,T)\sigma^*(t,T),$$

which is what the drift of $f(t,T)$ must be under risk neutral. So as long as we derive $B(t,T)$ and $f(t,T)$ from a consistent risk neutral model (which is any factor model for the short rate) then the dynamics of $f(t,T)$ and $B(t,T)$ would be consistent as discussed in the above.

Apply this to the Hull-White example above, we see that since

$$B(t,T) = \exp\{-C(t,T)R(t) - A(t,T)\},$$

the dynamics of $B(t,T)$ under $\widetilde{\mathbf{P}}$ is

$$dB(t,T) = \textit{ something } dt - B(t,T)C(t,T)\gamma(t)d\widetilde{W}(t).$$

(The fact that the *something* becomes $B(t,T)R(t)$ comes from the relation between $C(t,T), A(t,T), \beta(t), \gamma(t)$. The fact that the volatility has this form comes from the fact that only $R(t)$ has Brownian motion component in its dynamics and the Ito's formula applied to $B(t,T)$). This implies that

$$\sigma^*(t,T) = C(t,T)\gamma(t).$$

Note that this is consistent with the fact that we assign

$$\sigma(t,T) = \frac{\partial C(t,T)}{\partial T}\gamma(t)$$

for the dynamics of $f(t,T)$. Then the conclusion that

$$\frac{\partial C(t,T)}{\partial T}\beta(t) + R(t)\frac{\partial C'(t,T)}{\partial T} + \frac{\partial A'(t,T)}{\partial T} = \sigma(t,T)\sigma^*(t,T).$$

can be seen as a consequence of the dynamics of $f(t,T)$ being forced once we known the dynamics of $B(t,T)$ as discussed above.

Because the Hull-White, Vasicek, etc. models are examples of HJM models, all theory we develop later assuming the HJM model is valid for these earlier models.