Interest Rate Models

I. Classical Models

A. Pricing Interest Rate Derivatives

1. The Spot Interest Rate as a Non-Traded Factor

   a. Dynamics of the factor $F=r$, where $r$ is the short term (or spot) interest rate:

   \[ dr = m(r, t) r dt + s(r, t) r dZ \]

   b. Risk-neutral dynamics:

   \[ dr = \{m - \lambda s\} r dt + s r dZ^Q_t \]

   c. Price of a derivative is the solution of the PDE:

   \[ \frac{\partial f}{\partial t} + \frac{\partial f}{\partial r} \{m - \lambda s\} r + \frac{1}{2} \frac{\partial^2 f}{\partial r^2} s^2 r^2 + h - rf = 0 \]

   - At $t=T$: $f(r, T) = g(r)$, and boundary conditions at $r = 0$ and $r = \infty$.

   d. Feynman-Kac Solution:

   \[ f(r, t) = E^Q g(r_T) \exp\left\{-\int_t^T r_u du\right\} + \int_t^T \exp\left\{-\int_t^\tau r_u du\right\} h(r_\tau, \tau) d\tau \bigg| r_t = r \]
2. Examples

a. Bonds

- A zero-coupon bond has $g(r)=1$ and $h=0$.
- A coupon bond has $g(r)=1$ and $h =$ coupon payment.

b. European call option maturing at $T_1$ on a zero-coupon bond maturing at $T_2$, $T_2>T_1$:

$g(r)=\max(0,B(T_1,T_2,r)-K)$ and $h=0$.

c. Structured notes (obligations whose payments are contingent upon stock prices, interest rates, commodities or currencies), floaters (fixed-income instrument with varying coupon rate), LYONs (liquid yield option notes: a zero-coupon, convertible, callable and putable bond).

d. A cap is a loan at variable rate that is capped at some level $r^\text{high}$. Can be seen as: $h=\min(r,r^\text{high})$ and $g(r)=1$.

e. A floor: $h=\max(r,r^\text{low})$ and $g(r)=1$.

f. A yield curve call option has $h=0$ and $g(r)=\max(0,Y(T,T+n,r)-K)$ where $Y(T,T+n,r)$ is the yield-to-maturity at $T$ of a $n$-year bond.

g. A yield curve slope call option has $h=0$ and $g(r)=\max(0,[Y(T,T+n,r)-Y(T,T+m,r)]/[n-m]-K)$, for some $n>m$. 
B. The Vasicek Model

1. Specification

a. Actual spot rate dynamics:

\[ dr = m(r, t) r dt + s(r, t) r dZ = \beta(\alpha - r) dt + \sigma dZ \]

- \( \alpha, \beta \) and \( \sigma \) are constant parameters
- This is known as an Ornstein-Uhlenbeck process
- Conditional moments:

\[
\begin{align*}
E[r_T | r_t] &= \alpha + e^{-\beta(T-t)}(r_t - \alpha) \\
V[r_T | r_t] &= \frac{\sigma^2}{2\beta} \left(1 - e^{-2\beta(T-t)}\right)
\end{align*}
\]

b. Risk-neutral dynamics:

- Market price of risk: \( \lambda(r, t) = \lambda \) is constant.
- The risk-neutral dynamics are the same as the actual dynamics, with \( \alpha \) replaced by

\[
\alpha^* = \alpha - (\sigma \lambda / \beta)
\]

\[
\begin{align*}
\{m(r, t) - \lambda(r, t)s(r, t)\} \ r dt + s(r, t) r dZ^Q \\
= \{\beta(\alpha - r) - \sigma \lambda\} dt + \sigma dZ^Q \\
= \{\beta(\alpha^* - r)\} dt + \sigma dZ^Q
\end{align*}
\]
• So:

\[
\begin{align*}
E^Q[r_T | r_t] &= \alpha^* + e^{-\beta(T-t)}(r_t - \alpha^*) \\
V^Q[r_T | r_t] &= \frac{\sigma^2}{2\beta} \{1 - e^{-2\beta(T-t)}\}
\end{align*}
\]

c. Resulting risk-neutral density is Gaussian:

\[
\rho^Q(r_T | r_t) = \frac{1}{\sqrt{2\pi V^Q[r_T | r_t]}} \exp\left\{-\frac{(r_T - E^Q[r_T | r_t])^2}{2V^Q[r_T | r_t]}\right\}
\]

d. Problems

• The spot rate can become negative with non-zero (albeit small) probability.

• The volatility of the interest rate is constant, which contradicts the data.

2. Term Structure

a. Closed-form bond prices

- Price at $t$ of a zero-coupon bond with face value $\$1$, maturing at $T$, when the spot rate is $r_t=r$:

$$B(r, t, T) = e^{a(t,T) - r b(t,T)}$$

- Define the yield of a bond of maturity $\tau+T-t$ to be $y(r,t,T)$ such that: $B(r, t, T) = e^{-y(r,t,T)(T-t)}$, that is:

$$y(r, t, T) = -\frac{1}{T-t} \ln(B(r, t, T))$$

And so:

$$y(r, t, T) = -\frac{\ln B(r, t, T)}{(T-t)} = -\frac{a(t, T)}{(T-t)} + \frac{b(t, T)}{(T-t)} r$$

That is $y(r,t,T)$ is an affine function of the short-term rate $r$.

where:

- $b(t, T) \equiv \frac{1-e^{-\beta(T-t)}}{\beta}$

- $a(t, T) \equiv \left\{ \frac{b(t, T) - (T-t)}{\beta^2 \alpha - \beta \lambda \sigma - \sigma^2 / 2} - \frac{\sigma b(t, T)^2}{4 \beta} \right\}$
b. Possible shapes of the term structure

- Term structures under the Vasicek model have an asymptote: \( \lim_{T \to \infty} y(r, t, T) = y_\infty = \alpha - \frac{\lambda\sigma}{\beta} - \frac{\sigma^2}{2\beta^2} \).

- If \( r \leq y_\infty - \frac{\sigma^2}{4\beta^2} \) the TS is monotonically increasing.

- If \( y_\infty - \frac{\sigma^2}{4\beta^2} < r \leq y_\infty + \frac{\sigma^2}{2\beta^2} \), the TS is a humped curve.

- If \( r > y_\infty + \frac{\sigma^2}{2\beta^2} \) the TS is monotonically decreasing.

(Hull 1993)

C. The Cox-Ingersoll-Ross Model

1. Specification

a. Actual spot rate dynamics:

\[ dr = m(r, t) r dt + s(r, t) r dZ = \beta (\alpha - r) dt + \sigma \sqrt{r} dZ \]

• \( \alpha, \beta \) and \( \sigma \) are constant parameters

• This is known as Feller's square-root process (Feller, Annals of Mathematics, 1951).

• Conditional moments:

\[
\begin{align*}
E[r_T | r_t] &= \alpha + e^{-\beta(T-t)} (r_t - \alpha) \\
V[r_T | r_t] &= r_t \beta \left( e^{-\beta(T-t)} - e^{-2\beta(T-t)} \right) + \alpha \beta \left( 1 - e^{-2\beta(T-t)} \right) 
\end{align*}
\]

Resulting density is non-central chi-squared:

\[
\rho(r_T | r_t) = c e^{-u-v} \left( \frac{v}{u} \right)^{q/2} I_q \left( 2\sqrt{uv} \right)
\]

where:

\[
\begin{align*}
c &= 2\beta / \left( \sigma^2 \left( 1 - e^{-\beta(T-t)} \right) \right) \\
u &= cr_t e^{-\beta(T-t)} \\
v &= cr_T \\
q &= \left( 2\beta \alpha / \sigma^2 \right) - 1
\end{align*}
\]
and $I_q(2\sqrt{uv})$ is the modified Bessel function of the first kind of order $q$ (see Feller --also in Handbook of Mathematical Functions).

b. Risk-neutral dynamics:

- Market price of risk: $\lambda(r,t) = \frac{\lambda}{\sigma} \sqrt{r}$ where $\lambda$ is a constant parameter.

- The risk-neutral dynamics are the same as the actual dynamics, with $\alpha$ replaced by $\alpha^*$ and $\beta$ by $\beta^*$.

\[
\begin{align*}
\text{dr} &= \{m(r,t) - \lambda(r,t)s(r,t)\}r\,dt + s(r,t)r\,dZ_t^Q \\
&= \left\{\beta(\alpha - r) - \left(\frac{\lambda}{\sigma} \sqrt{r}\right)\left(\sigma \sqrt{r}\right)\right\}dt + \sigma \sqrt{r} \,dZ_t^Q \\
&= \left\{\beta(\alpha - r) - \lambda r\right\}dt + \sigma \sqrt{r} \,dZ_t^Q \\
&= \left\{(\beta + \lambda)\left(\frac{\beta \alpha}{\beta + \lambda} - r\right)\right\}dt + \sigma \sqrt{r} \,dZ_t^Q \\
&\equiv \left\{\beta^*(\alpha^* - r)\right\}dt + \sigma \sqrt{r} \,dZ_t^Q \\
\end{align*}
\]

- So: $E^Q[r_t|t]$ and $V^Q[r_t|t]$ are obtained as above with $\alpha^*$ (resp. $\beta^*$) replacing $\alpha$ (resp. $\beta$).

d. Features

- The spot rate is restricted to $(0, +\infty)$.

- The volatility of the interest rate increases with the level of the interest rate.

2. Term Structure

a. Closed-form bond prices

- Price at t of a zero-coupon bond with face value $1, maturing at T, when the spot rate is \( r_t = r \):

\[
B(r, t, T) = a(t, T) e^{-r b(t, T)}
\]

where:

- \( b(t, T) = \frac{2 \left\{ e^{\gamma (T-t)} - 1 \right\}}{(\gamma + \beta + \lambda) \left\{ e^{\gamma (T-t)} - 1 \right\} + 2\gamma}
\]

- \( a(t, T) = \frac{2 \gamma e^{[(\gamma + \beta + \lambda) (T-t)]/2}}{\left\{ \gamma + \beta + \lambda \right\} \left\{ e^{\gamma (T-t)} - 1 \right\} + 2\gamma} \)

- \( \gamma = \left\{ (\beta + \lambda)^2 + 2\sigma^2 \right\}^{1/2} \)

- Notice, that yields of zero coupon bonds are affine functions of the short-term interest rate in the CIR model.
b. Possible shapes of the term structure

- Term structures under the CIR model also have an asymptote: \( \lim_{T \to \infty} y(r, t, T) = y_{\infty} = \frac{2\beta \alpha}{\gamma + \beta + \lambda} \).

- If \( r \leq y_{\infty} \) the TS is monotonically increasing.

- If \( y_{\infty} < r \leq \frac{\beta \alpha}{\beta + \lambda} \), the TS is a humped curve.

- If \( r > \frac{\beta \alpha}{\beta + \lambda} \) the TS is monotonically decreasing.

c. Closed-form European bond option prices also available: CIR page 396.

II. The Estimation of the Spot Rate Process

A. Method-of-Moments Estimation

1. Need both the drift and the diffusion function of the spot rate process in order to price interest rates derivatives

   a. Formulas require estimates of \( \alpha, \beta, \sigma \)

   b. More general models require more parameters.
2. Discretization and GMM

a. Approximate the continuous-time SDE by a discrete model:

\[ dr_t = \beta (\alpha - r_t) \, dt + \sigma r_t^\gamma \, dZ_t \]

\[ r_{t+\Delta t} - r_t = \beta \alpha - \beta r_t + \varepsilon_{t+\Delta t} \]

\[ E[\varepsilon_{t+\Delta t} | r_t] = 0, \quad E[\varepsilon_{t+\Delta t}^2 | r_t] = \sigma^2 r_t^{2\gamma} \]

b. GMM Estimation

\[ \hat{\theta} = \arg \min_{\theta} \left[ \frac{1}{T} \sum_{t=1}^{T} u_t(\theta) \right] W_T \left[ \frac{1}{T} \sum_{t=1}^{T} u_t(\theta) \right] \]

\[ u_t(\theta) = \begin{bmatrix} \varepsilon_{t+\Delta t} \\ \varepsilon_{t+\Delta t} r_t \\ \varepsilon_{t+\Delta t}^2 - \sigma^2 r_t^{2\gamma} \\ \{\varepsilon_{t+\Delta t}^2 - \sigma^2 r_t^{2\gamma}\} r_t \end{bmatrix} \]

\[ \varepsilon_{t+\Delta t} = \{ r_{t+\Delta t} - r_t \} - \{ \beta \alpha - \beta r_t \} \]

\[ \theta = [\alpha \quad \beta \quad \sigma \quad \gamma] \]

- The weighting matrix \( W_T \) is chosen optimally.
- Estimated \( \gamma \) is 1.5
c. **Problems**

- The discrete approximation will work only if $\Delta t$ goes to 0. The data available to construct the GMM estimator do not satisfy this (e.g., weekly or daily = fixed sampling interval).

- This is very different from constructing a tree or solving numerically a PDE, where $\Delta t$ is under your control. Here $\Delta t$ is imposed by the data at hand.

3. **MLE (Maximum Likelihood Estimation)**

a. When conditional densities are known in closed-form

b. Can estimate the parameters using data over any intervals of time.


**Maximum Likelihood Estimation Primer:**

Suppose that $x$ is a continuous random variable with transition density function

$$f (x; \theta_1, \theta_2, \ldots, \theta_k)$$

where $\theta_1, \theta_2, \ldots, \theta_k$ are $k$ unknown constant parameters that need to be estimated.
Collect \( n \) (independent) observations for \( x \), hence get: 
\[ x_1, x_2, \ldots, x_n \quad \text{(a time series of interest rates r, for instance)} \]

As long as the \( n \) observations’ “randomness components” are independent (ex: the various “dZ” are independent in the case of a Brownian motion), the likelihood of observing all of the \( x \)’s is the product of the likelihoods of observing each \( x \).

Let us thus write this likelihood as:
\[
L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k) = L = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \ldots, \theta_k)
\]

The Maximum Likelihood Estimators are obtained by maximizing the function \( L \). However, it is actually easier to maximize the logarithm of \( L \). Note that maximizing \( L \) or \( \ln(L) \) produces the same estimates: if \( \ln(L) \) is maximized \( L \) must also be maximized since the logarithm is a strictly increasing function. Also note that you can also \textbf{minimize} \( -\ln(L) \) instead.

The function to maximize is thus:
\[
\ln(L) = \sum_{i=1}^{n} \ln f(x_i; \theta_1, \theta_2, \ldots, \theta_k)
\]

And the Maximum Likelihood Estimates satisfy:
\[
(\theta_1, \theta_2, \ldots, \theta_k) = \text{Argmax}_\theta \sum_{i=1}^{n} \ln f(x_i; \theta_1, \theta_2, \ldots, \theta_k)
\]

Closed-form solutions can in some cases be obtained by taking derivatives of \( \ln(L) \) with respect to the various \( \theta \), but in general numerical optimization must be used.