Affine Processes

Econometric specifications

Eduardo Rossi

University of Pavia

December, 2010

Introduction

The *affine* processes are among the most widely studied time series processes in the empirical finance literature.

- Accommodation of stochastic volatility
- Volatility
- Jumps
- Correlations

among risk factors.

Affine Processes

Affine Process

An affine process Y is one for which the conditional mean and variance are affine functions of Y.

Characterization of affine processes in terms of

- *Exponential-affine Fourier* transform for continuous time
- *Laplace* transform for discrete-time.

Markov Process

Probability Space $(\Omega, \mathcal{F}, P)$, information set $\mathcal{F}_t$.
First-order Markov Process (MP) Y taking on values in a state space $D \subset \mathbb{R}^N$.

Markov Process

A process is Markov if, for any measurable function, $g : D \to \mathbb{R}$ and for any fixed times t and s, $s > t$

$$E_t[g(Y_s)] = h(Y_t)$$

for some function $h : D \to \mathbb{R}$.

Conditional Characteristic Fucntion

CCF of a Markov Process

The conditional characteristic function (CCF) of a MP, Y_T, conditioned on current and lagged information about Y at date t, is given by the *Fourier* transform of its conditional density function:

$$\begin{aligned} CCF_t(\tau, u) &\equiv E_t\left(e^{iu'Y_T}|Y_t\right) \qquad u \in \mathbb{R}^N \\ &= \int_{\mathbb{R}^N} e^{iu'Y_T} f_Y(Y_T|Y_t;\gamma)\mathrm{d}Y_T \end{aligned}$$

where $\tau = (T - t)$, $i = \sqrt{-1}$, and f_Y is the conditional density of Y.

Conditional Moment-generating function

CMGF of a Markov Process

The conditional Moment-generating function (CMGF) of a MP, Y_T, conditioned on current and lagged information about Y at date t, is given by the *Laplace* transform of its conditional density:

$$\begin{aligned} CMGF_t(\tau, u) &\equiv E_t\left(e^{u' Y_T} | Y_t\right) \qquad u \in \mathbb{R}^N \\ &= \int_{\mathbb{R}^N} e^{u' Y_T} f_Y(Y_T | Y_t; \gamma) \mathrm{d} Y_T \end{aligned}$$

where $\tau = (T - t)$, $i = \sqrt{-1}$, and f_Y is the conditional density of Y.

Affine Processes

Affine Process

A MP Y is said to be an *affine process* if either its CCF or CMGF has the exponential affine form

$$CCF_t = e^{\phi_{0t} + \phi'_{Y_t} Y_t}$$

or

$$CMGF_t = e^{\phi_{0t} + \phi'_{Y_t} Y_t}$$

where ϕ_{0t} and ϕ_{Y_t} are complex (real) coefficients in the case of the CCF (CMGF). They are indexed by t to allow for the possibility of time dependence of the moments of Y.

Continuous-time Affine Processes

Jump-diffusion process

A jump-diffusion process is a MP solving the SDE

$$\mathrm{d}Y_t = \mu(Y_t, \gamma_0)\mathrm{d}t + \sigma(Y_t, \gamma_0)\mathrm{d}W_t + \mathrm{d}Z_t$$

where

- W_t is an $(\mathcal{F})$-standard Brownian Motion in $\mathbb{R}^N$
- $\mu : D \rightarrow \mathbb{R}^N$;
- $\sigma : D \rightarrow \mathbb{R}^{N \times N}$;
- Z pure-jump process whose jump amplitudes have a fixed probability distribution ν on $\mathbb{R}^N$ and arrive with intensity $\{\lambda(Y_t) : t \geq 0\}$, for some $\lambda : D \rightarrow [0, \infty)$;
- $\gamma \in \mathbb{R}^K$ is the vector of unknown parameters.

Jump process

Cox process construction of *jump arrivals* in which, conditional on the path

$$\{Y_s := 0 \leq s \leq t\}$$

to time t, the times of jumps arriving during the interval $[0, t]$ are assumed to be the jump times of a Poisson process with time-varying intensity

$$\{\lambda(Y_s) : 0 \leq s < T\}$$

Affine Processes

The special case of an affine-jump diffusion is obtained by requiring that μ, $\sigma\sigma'$ and λ all be affine functions on $\mathcal{D}$. Y follows a jump-affine diffusion if

$$\mathrm{d}Y_t = \mathcal{K}(\Theta - Y_t)\mathrm{d}t + \Sigma\sqrt{S_t}\mathrm{d}W_t + \mathrm{d}Z_t$$

- W_t N-dimensional independent standard Brownian Motion
- $\mathcal{K}$ and Σ are $N \times N$ matrices, which be nondiagonal and asymmetric;
- $S_{ii,t} = \alpha_i + \beta_i' Y_t$

Both drifts and the instantaneous conditional variances are affine in Y_t. The jump intensity is assumed to be a positive, affine function of the state Y_t

$$\lambda_t = l_0 + l_Y' Y_t$$

The jump-size distribution f_J is assumed to be determined by its characteristic function

$$\mathcal{J}(u) = \int \exp\{ius\} f_J(s) ds$$

Affine Term Structure Models

Zero-coupon price

The time-t price of a zero-coupon bond maturing at time $t+\tau$:

$$P_t(\tau) = E_t^Q\left[\exp\left(-\int_t^{t+\tau} r_s ds\right)\right]$$

Instantaneous short rate

The instantaneous short rate r_t is an affine function of a $(N \times 1)$ vector of unobservable state variables $x_t = (x_{1t}, \ldots, x_{Nt})'$:

$$r_t = \delta_0 + \sum_{i=1}^{N} \delta_{1i} x_{it} = \delta_0 + \delta_1' x_t$$

Affine Term Structure Models

State Variable

Under the risk neutral probability measure $\mathcal{Q}$ the state variables x_t follow an *affine diffusion*:

$$dx_t = K^Q(\theta^Q - x_t)dt + \Sigma S_t^{1/2}\mathrm{d}W_t^Q \tag{1}$$

where

- W_t^Q is a $(N \times 1)$ vector of independent Brownian motions under $\mathcal{Q}$
- K^Q and Σ are $(N \times N)$ general matrices of parameters (Σ may be asymmetric) θ^Q is a $(N \times 1)$ vector of parameters
- S_t is a $(N \times N)$ diagonal matrix with (i, i) element given by:

  $$[S_t]_{ii} = \alpha_i + \beta_i' x_t$$

  and for each $i = 1, \ldots, N$, α_i and β_i respectively are a scalar and a $(N \times 1)$ vector of parameters.

Affine Term Structure Models

The drift and the conditional variance of x_t are affine functions of $\mathbf{x}_t$. Let us denote with:

$$\mu_t^Q = K^Q(\theta^Q - x_t)$$

$$\Sigma_t = \Sigma S_t^{1/2}$$

the drift vector under $\mathcal{Q}$ and the diffusion matrix of x_t, respectively.

Affine Term Structure Models

If the parametrization is admissible, Duffie and Kan (1996) have shown that:

$$P_t(\tau) = exp\left[a(\tau) - b(\tau)'\mathbf{x}_t\right]$$

where the coefficients $a(\tau)$ and $b(\tau)$ satisfy the following system of *ordinary differential equation* (ODEs):

$$\frac{da(\tau)}{d\tau} = -\theta^{Q'} K^{Q'} b(\tau) + \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma' b(\tau)\right]_i^2 \alpha_i - \delta_0$$

$$\frac{db(\tau)}{d\tau} = -K^{Q'} b(\tau) - \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma' b(\tau)\right]_i^2 \beta_i + \delta_1$$

subject to the initial conditions $a(0) = 0$ and $b(0) = 0$.

Affine Term Structure Models

The dynamics of x_t under the actual probability measure $\mathcal{P}$ can be derived with an application for the Girsanov's Theorem:

$$\mathrm{d}x_t = (\mu_t^Q + \Sigma_t \lambda_t)\mathrm{d}t + \Sigma S_t^{1/2} \mathrm{d}W_t$$

where λ_t is a $(N \times 1)$ vector of risk premium functions.

Affine Term Structure Models

Some (more or less) common choices of λ_t are:

Dai and Singleton (2000)

$$\lambda_t = S_t^{1/2} \lambda_1$$

where λ_1 is a$(N \times 1)$ vector of parameters. Under this choice:

$$\mu_t^Q + \Sigma_t \lambda_t = K(\theta - x_t)$$

$$K = K^Q - \Sigma \Phi$$

$$\theta = K^{-1}(K^Q \theta^Q + \Sigma \phi)$$

where Φ is a $(N \times N)$ matrix whose *i*-th row is given by $\lambda_{1i} \beta_i'$, and ϕ is a $(N \times 1)$ vector whose *i*-th element is given by $\lambda_{1i} \alpha_i$. This choice gives rise to the so called *completely affine class* of term structure models.

Affine Term Structure Models

Duffee (2002)

$$\lambda_t = S_t^{1/2}\lambda_1 + (S_t^-)^{1/2}\lambda_2 x_t$$

where λ_1 is as above, λ_2 is a $(N \times N)$ matrix of parameters, and S_t^- is diagonal $(N \times N)$ matrix with:

$$\left[S_t^-\right]_{ii} = \begin{cases} (\alpha_i + \beta_i' x_t)^{-1} & \text{if } \inf(\alpha_i + \beta_i' x_t) > 0 \\ 0 & \text{otherwise} \end{cases}$$

This choice gives rise to the so called *essentially affine class* of term structure models.

Affine Term Structure Models

Duffee and Stanton (2001)

$$\lambda_t = \Sigma^{-1}\lambda_0 + S_t\Sigma^{-1}\lambda_1$$

where λ_0 is a $(N \times 1)$ vector of parameters, and λ_1 is defined as above.

Duarte (2002)

$$\lambda_t = \Sigma^{-1}\lambda_0 + S_t^{1/2}\lambda_1 + (S_t^-)^{1/2}\Lambda_2 x_t$$

with λ_0, λ_1, λ_2 and S_t^- defined as above.

Affine Term Structure Models

In a completely affine model, the relevant equations are given by:

- The dynamics of the state variables under $\mathcal{P}$:

$$\mathrm{d}x_t = K(\theta - x_t)\mathrm{d}t + \Sigma S_t^{1/2}\mathrm{d}W_t$$

- The risk premium functions:

$$\lambda_t = S_t^{1/2}\lambda$$

- the dynamics of x_t under $\mathcal{Q}$:

$$\mathrm{d}x_t = K^Q(\theta^Q - x_t)dt + \Sigma S_t^{1/2}\mathrm{d}W_t$$

where:

$$K^Q = K + \Sigma\Phi$$

$$\theta^Q = K^{Q-1}(K\theta - \Sigma\phi)$$

where Φ is a $(N \times N)$ matrix whose i-th row vis given by $\lambda_i\beta_i'$, and ϕ is a $(N \times 1)$ vector whose i-th element is given by $\lambda_i\alpha_i$

Affine Term Structure Models

- Zero coupon prices:

$$P_t(\tau) = \exp\left[a(\tau) - b(\tau)' x_t\right]$$

and zero coupon yields:

$$y_t(\tau) = -\frac{1}{\tau}\ln P_t(\tau) = -\frac{a(\tau)}{\tau} + \frac{b(\tau)'}{\tau} b_t$$

where $a(\tau)$ and $b(\tau)$ satisfy the ODE:

$$\frac{da(\tau)}{d\tau} = -\theta^{Q'} K^{Q'} b(\tau) + \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma' b(\tau)\right]_i^2 \alpha_i - \delta_0$$

$$\frac{db(\tau)}{d\tau} = -K^{Q'} b(\tau) - \frac{1}{2}\sum_{i=1}^{N}\left[\Sigma' b(\tau)\right]_i^2 \beta_i + \delta_1$$

subject to the initial conditions $a(0) = 0$ and $b(0) = 0$.

Affine Term Structure Models

Let us denote with ψ the vector of parameters:

$$\psi = (K, \theta, \Sigma, B, \alpha, \delta_0, \delta_1, \lambda)$$

where $B = (\beta_1, \ldots, \beta_N)$ is a $(N \times N)$ matrix. To be *admissible*, a parametrization must guarantee that $[S_t]_{ii}$ is strictly positive, for all i.

Following Dai and Singleton (2000), it turns out that the condition to impose on ψ depend on the rank of B, which we shall denote by m. Formally, for any given N, there are $N+1$ subfamilies $\mathcal{A}_m(N)$ of admissible N factor models, corresponding to $m = 0, 1, \ldots, N$.

Affine Term Structure Models

For each m, the *canonical representation* of $\mathcal{A}_m(N)$ is defined as follows.

- To begin with, partition x_t as:

$$x_{t\,N\times 1} = \left[\begin{array}{c} x_t^B \\ m\times 1 \\ x_t^D \\ (N-m)\times 1 \end{array}\right]$$

- Accordingly, if $m > 0$, partition K as follows:

$$K_{N\times N} = \left[\begin{array}{cc} K^{BB} & 0 \\ m\times m & m\times(N-m) \\ K^{DB} & K^{DD} \\ (N-m)\times m & (N-m)\times(N-m) \end{array}\right]$$

Affine Term Structure Models

If $m = 0$, K is either upper or lower triangular. Furthermore:

$$[K]_{ij} \leq 0 \text{ for } 1 \leq j \leq m, i \neq j$$

To insure stationarity, the (real part of the) eigenvalues of K must be positive.

Affine Term Structure Models

- On θ:

$$\theta_{N\times 1} = \begin{bmatrix} \theta^B \\ m \times 1 \\ 0 \\ (N-m) \times 1 \end{bmatrix}$$

Furthermore:

$$[K\theta]_i = \sum_{j=1}^{m} [K]_{ij}[\theta]_j > 0 \text{ for } 1 \leq i \leq m,$$

This condition must be strengthened to:

$$[K\theta]_i = \sum_{j=1}^{m} [K]_{ij}[\theta]_j > \frac{1}{2} \text{ for } 1 \leq i \leq m$$

to insure that 0 is not an absorbing state for x_t^B. Finally:

$$[\theta]_i \geq 0 \text{ for } 1 \leq i \leq m$$

Affine Term Structure Models

- On Σ:

$$\Sigma_{(N\times N)} = I_N$$

- On α:

$$\alpha = \left[\begin{array}{c} 0_{m\times 1} \\ \iota_{(N-m)\times 1} \end{array} \right] \quad (N \times 1)$$

- On B:

$$B_{N\times N} = \left[\begin{array}{cc} I_m & B^{BD} \\ m \times m & m \times (N-m) \\ 0 & 0 \\ (N-m) \times m & (N-m) \times (N-m) \end{array} \right]$$

Furthermore:

$$[B]_{ij} \geq 0 \text{ for } 1 \leq j \leq m, m+1 \leq j \leq N$$

Affine Term Structure Models

- On δ_1

$$\delta_{1i} \geq 0, \text{ for } m+1 \leq i \leq N$$

Affine Term Structure Models

- To complete the definition, the subfamilies $\mathcal{A}_m(N)$ consist of all the models which are nested special cases of the canonical representation, or of any equivalent model obtained by an *invariant* transformation of it.
- Invariant transformations preserve admissibility and identification.
- These are sufficient (but not necessary) conditions for admissibility;
- A$\ddot{i}$t-Sahalia and Kimmel (2002)provide some examples of admissible two- and three-factors affine models which are not invariant transformations of the corresponding canonical representation.

$\mathcal{A}_1(1)$

- Dynamics under $\mathcal{P}$:

$$\mathrm{d}x_t = k(\theta - x_t)\mathrm{d}t + \sqrt{x_t}\mathrm{d}W_t$$

- Short rate:

$$r_t = \delta_0 + \delta_1 x_t$$

- Risk premium:

$$\lambda_t = \lambda\sqrt{x_t}$$

- Dynamics under $\mathcal{Q}$:

$$\mathrm{d}x_t = k^Q(\theta^Q - x_t)dt + \sqrt{x_t}\mathrm{d}\widetilde{W}_t$$

where:

$$k^Q = k + \lambda$$

and

$$\theta^Q = \frac{k\theta}{k+\lambda}$$

$\mathcal{A}_1(1)$

- Yields:
$$y_t(\tau) = -\frac{a(\tau)}{\tau} + \frac{b(\tau)}{\tau} x_t$$
where $a(\tau)$ and $b(\tau)$ solve:
$$\frac{da(\tau)}{d\tau} = -\theta^{Q'} k^Q b(\tau) - \delta_0$$
$$\frac{db(\tau)}{d\tau} = -k^Q b(\tau) - \frac{1}{2} b(\tau)_i^2 \beta_i + \delta_1$$
- Parameters:
$$\psi = (k, \theta, \delta_0, \delta_1, \lambda)'$$
- Restrictions on ψ:
$$k > 0$$
$$\theta \geq 0$$
$$k\theta > \frac{1}{2}$$

$\mathcal{A}_1(1)$

- Assume that at each date t we observe P yields:

$$y_t = [y_t(\tau_1), \ldots, y_t(\tau_P)]'$$

- The yield of maturity τ_x is observed without error:

$$y_t(\tau_x) = -\frac{a(\tau_x)}{\tau_x} + \frac{b(\tau_x)}{\tau_x} x_t$$

Hence:

$$\widehat{x}_t = \frac{\tau_x}{b(\tau_x)} \left[y_t(\tau_x) + \frac{a(\tau_x)}{\tau_x} \right]$$

Notice: $\widehat{x}_t$ *must be strictly positive at each date*. We take into account these constraints by introducing a huge penalization in the log-likelihood function.

$\mathcal{A}_1(1)$

- Using the Jacobian formula and Aït-Sahalia (2001) analytic approximation, it is easy to compute the log-likelihood of the yields observed without error(see the following).
- The other $P-1$ yields are observed with error:

  $$y_t(\tau_i) = \widehat{y}_t(\tau_i) + e_i \text{ for } i = 1, \ldots, P-1$$

  where:

  $$\widehat{y}_t(\tau_i) = -\frac{a(\tau_i)}{\tau_i} + \frac{b(\tau_i)}{\tau_i} x_t$$

  In vector terms:

  $$y_{-xt} = \widehat{y}_{-xt} + e_t$$

  where e_t is a $[(P-1) \times 1]$ vector of measurement errors.
- If we make an assumption on the stochastic structure of e_t, than it is straightforward to compute the log-likelihood of the yields observed with error.

Sample log-lik $\mathcal{A}_1(1)$

- Let us denote with ν the vector of parameters entering the distribution of the measurement errors e_t, and define:

$$\gamma = \begin{pmatrix} \psi \\ \nu \end{pmatrix}$$

  the vector of all (structural and auxiliary) parameters in the model.
- We also assume that the measurement errors e_t are independent through time, but may be correlated in cross section.
- The conditional density of the yields at date t given previous observations can be factorized as:

$$\begin{aligned} f_Y(y_t|Y_1^{t-1};\gamma) &= f_Y(y_t|y_{t-1};\gamma) = f_Y(y_t(\tau_x)|y_{t-1};\gamma) f_Y(y_{-xt}|y_t(\tau_x);\gamma \\ &= f_X(\widehat{x}_t|\widehat{x}_{t-1};\psi)\frac{\tau_x}{b(\tau_x)} \times f_E(e_t|\widehat{x}_t;\gamma) \end{aligned}$$

Sample log-lik $\mathcal{A}_1(1)$

- The factor component of the total sample log-likelihood is given by:

$$\ell_Y[\psi; y_{(\tau_x)}] = \sum_{t=2}^{T} \ln f_X(\widehat{x}_t | \widehat{x}_{t-1}; \psi) + (T-1)[\ln(\tau_x) - \ln b(\tau_x)]$$

- The measurement error component of the total sample log-likelihood depend on the assumed stochastic structure of $\mathbf{e}_t$.

Sample log-lik $\mathcal{A}_1(1)$

- If we assume:

$$e_t \sim IIDN(0, \omega^2 I_{P-1})$$

then ν reduces to the scalar ω, and:

$$\ell_Y[\gamma; Y_{(-x)}] = -\frac{T(P-1)}{2}\log(2\pi) - T(P-1)\log(\omega) - \frac{1}{2\omega^2}\sum_{t=1}^{T} e_t' e_t$$

Sample log-lik $\mathcal{A}_1(1)$

- If we assume:

$$e_t \sim IIDN(0, \Omega)$$

with $\Omega = \mathrm{diag}(\omega_1^2, \ldots, \omega_{P-1}^2)$, then $\nu = (\omega_1^2, \ldots, \omega_{P-1}^2)'$, and:

$$\begin{aligned} \ell_Y[\gamma; Y_{(-x)}] &= -\frac{T(P-1)}{2} \log(2\pi) - T \sum_{i=1}^{P-1} \log(\omega_i) \\ &\quad -\frac{1}{2} \sum_{t=1}^{T} \sum_{i=1}^{P-1} \left(\frac{e_{it}'}{\omega} \right)^2 \\ &= -\frac{T(P-1)}{2} \log(2\pi) - T \sum_{i=1}^{P-1} \log(\omega_i) - \frac{1}{2} \sum_{t=1}^{T} v_t' v_t \end{aligned}$$

where:

$$v_t = \left(\frac{e_{it}}{\omega_1}, \ldots, \frac{e_{P-1t}}{\omega_{P-1}} \right)'$$

Sample log-lik $\mathcal{A}_1(1)$

- If we assume:

$$e_t \sim IIDN(0, \Omega)$$

with general symmetric positive definite $\Omega = C'C$, where C is the upper triangular matrix provided by the Cholesky decomposition of Ω, then ν consist of the $P(P-1)/2$ free elements of C, and:

$$\begin{aligned} \ell_Y[\gamma; Y_{(-x)}] &= -\frac{T(P-1)}{2}\log(2\pi) - \frac{T}{2}\log|\Omega| - \frac{1}{2}\sum_{t=1}^{T} e_t'\Omega^{-1}e_t \\ &= -\frac{T(P-1)}{2}\log(2\pi) - \frac{T}{2}\log|C'C| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T} e_t'(C'C)^{-1}e_t \end{aligned}$$

Sample log-lik $\mathcal{A}_1(1)$

$$\begin{aligned}
\ell_Y[\gamma; Y_{(-x)}] &= -\frac{T(P-1)}{2}\log(2\pi) - T\log|C| - \frac{1}{2}\sum_{t=1}^{T} e_t' C^{-1}(C')^{-1}e_t \\
&\quad -\frac{T(P-1)}{2}\log(2\pi) - T\sum_{i=1}(P-1)\log[C]_{ii} \\
&\quad -\frac{1}{2}\sum_{t=1}^{T}[(C')^{-1}e_t]'[(C')^{-1}e_t] \\
&= -\frac{T(P-1)}{2}\log(2\pi) - T\sum_{i=1}(P-1)\log[C]_{ii} - \frac{1}{2}\sum_{t=1}^{T} v_t' v_t
\end{aligned}$$

where $v_t = (C')^{-1}e_t$.

Optimization issues

Things are complicated by the existence of T non linear constraints of the form $\widehat{x}_t > 0$.To solve the problem, we follow Duffee (2002) strategy:

1. Pick a random initial value of γ
2. Check if it is admissible, i.e. if $\widehat{x}_t > 0$ for all t. If the answer is positive, go on; otherwise, go back to step (1).
3. the optimization method; continue iterations until a stable convergence is achieved.
4. In the point thus attained, start a derivative-based optimization routine.

Procedure (1)-(4) is repeated until a given number (say 1000) of estimates are obtained. The final estimate is the one associated to the maximum value of the log-likelihood.

References

1. Aït-Sahalia, Y. (2001),"Closed-Form Likelihood Expansions for Multivariate Diffusions", working paper, Princeton University
2. Aït-Sahalia, Y. and R. Kimmel (2002), "Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansions", working paper, Princeton University
3. Dai, Q. and K.Singleton (2002), "Specification Analysis of Affine Term Structure Models", *Journal of Finance*, vol. LV, 1943-78.
4. Duarte, J (2002), "Evaluating Alternative Risk Preferences in Affine Term Structure Models", working paper, University of Washington.
5. Duffee, G.(2002), "Term Premia and Interest Rate Forecasts in Affine Models", *Journal of Finance*, vol.LVII, 405-443.
6. Duffee, G. and R. Stanton (2001), "Estimation of Dynamic Term Structure Models", working paper, University of California, Berkeley.
7. Duffie, D. and R. Kan (1996), "A Yield-Factor Model of Interest Rates", *Mathematical Finance*, vol. 6, n.4, 379-406.