Empirical Macro-Based Asset Pricing

Part 0: Tools

UCLA | Fall 2016

by Mikhail Chernov

This version: September 22, 2016

1. Stochastic processes as building blocks

Tools

- General probability laws
- The key object is conditional distribution: compute prices and estimate models
- A highly tractable class of affine models: show up both in empirical and equilibrium models

State Space Models

- States, X_t, and data Y_t:

$$\begin{aligned} X_{t+1} &= G(X_t, \varepsilon_{t+1}) \\ Y_t &= F(X_t, e_t) \end{aligned}$$

- We normally think of ε as an economic shock, and e as a measurement error
- Example:

$$\begin{aligned} X_{t+1} &= \bar{X}(I - \Phi) + \Phi X_t + \Sigma \varepsilon_{t+1} \\ Y_t &= A + BX_t + Ce_t \end{aligned}$$

- We need to know $p_X(X_{t+1}|X_t)$ to compute prices
- We need to know $p_Y(Y_{t+1}|Y_t) = p_X(F^{-1}(Y_{t+1})|Y_t) \left| \frac{dF^{-1}(y = Y_{t+1})}{dy} \right|$ to estimate models

MGF, CGF, and CF

- These are extremely useful tools to characterize distributions even if they are not known in closed form
- Moment-generating function: $h_t(X_{t+1}, s) = E_t e^{sX_{t+1}}$
- Property: $\partial^n h_t(X_{t+1}, 0)/\partial s^n = E_t(X_{t+1}^n) \equiv m_{nt}$
- Cumulant-generating function: $k_t(X_{t+1}, s) = \log h_t(X_{t+1}, s)$
- Property: $k_t(X_{t+1}, s) = \sum_{j=1}^{\infty} \kappa_{jt}(X_{t+1}) s^j / j!$
- Cumulants are almost moments

$$
\begin{aligned}
\text{mean} &= \kappa_{1t}(X_{t+1}) \\
\text{variance} &= \kappa_{2t}(X_{t+1}) \\
\text{skewness} &= \kappa_{3t}(X_{t+1}) / \kappa_{2t}^{3/2}(X_{t+1}) \\
\text{(excess) kurtosis} &= \kappa_{4t}(X_{t+1}) / \kappa_{2t}^{2}(X_{t+1})
\end{aligned}
$$

- Characteristic function: $f_t(X_{t+1}, s) = h_t(X_{t+1}, i \cdot s)$

AR(1) process

- Suppose: $X_{t+1} = \bar{X}(1-\phi) + \phi X_t + \sigma\varepsilon_{t+1}$
- The CGF:

$$\begin{aligned} & k_t(X_{t+1}, s) = \log E_t e^{s(\bar{X}(1-\phi)+\phi X_t+\sigma\varepsilon_{t+1})} \\ = \quad & s\left(\bar{X}(1-\phi)+\phi X_t\right) + \log E_t e^{s\sigma\varepsilon_{t+1}} = s(\bar{X}(1-\phi)+\phi X_t) + s^2\sigma^2/2 \end{aligned}$$

- Derivatives:

$$\begin{aligned} & \partial k_t(X_{t+1}, s)/\partial s = \bar{X}(1-\phi) + \phi X_t + s\sigma^2 = \bar{X}(1-\phi) + \phi X_t|_{s=0} \\ & \partial^2 k_t(X_{t+1}, s)/\partial s^2 = \sigma^2 \\ & \partial^n k_t(X_{t+1}, s)/\partial s^n = 0 \end{aligned}$$

- Therefore, it is a conditionally normal variable.
- Of course, we knew this already

MA(∞) process

- Suppose: $X_{t+1} = \bar{X} + \chi(B)\varepsilon_{t+1} = \bar{X} + \chi_0\varepsilon_{t+1} + \chi_1\varepsilon_t + \ldots$
- The CGF:

$$\begin{aligned} & k_t(X_{t+1}, s) = \log E_t e^{s(\bar{X}+\chi_0\varepsilon_{t+1}+\chi_1\varepsilon_t+\ldots)} \\ = \quad & s(\bar{X} + \chi_1\varepsilon_t + \ldots) + \log E_t e^{s\chi_0\varepsilon_{t+1}} = s(\bar{X} + [\chi(B)/B]_+\varepsilon_t) + s^2\chi_0^2/2 \end{aligned}$$

- Examples:

$$\begin{array}{rlll} AR(1) & \chi_0 = \sigma, & \chi_j = \phi\chi_{j-1} & \\ ARMA(1,1) & \chi_0 = \sigma, & \chi_1 = (\theta+\phi)\sigma, & \chi_j = \phi\chi_{j-1} \end{array}$$

- For ARMA(1,1), $k_t(X_{t+1}, s) = s(\bar{X}(1-\phi) + \phi X_t + \theta\sigma\varepsilon_t) + s^2\sigma^2/2$

A jump process: Poisson mixture of normals

- Suppose j is a number of jumps per period, $p(j) = e^{-\omega}\omega^j/j!$

$$E(j) = \sum_{j=0}^{\infty} j \cdot p(j) = \omega \sum_{j=1}^{\infty} e^{-\omega}\omega^{j-1}/(j-1)! = \omega$$

- Next component is a jump size: $Z_{t+1} \sim \mathcal{N}(\theta, \delta^2)$
- The CGF:

$$\begin{aligned} k(Z,s) &= \log E e^{sZ} = (s\theta + s^2\delta^2/2) \\ k(jZ,s) &= \log \sum_{j=0}^{\infty} e^{k(jZ,s)} p(j) = \log \sum_{j=0}^{\infty} e^{-\omega}(\omega e^{s\theta+s^2\delta^2/2})^j/j! \\ &= -\omega + \log \sum_{j=0}^{\infty} (\omega e^{s\theta+s^2\delta^2/2})^j/j! = \omega(e^{s\theta+s^2\delta^2/2} - 1) \end{aligned}$$

- Cumulants: $\kappa_1 = \omega\theta$, $\kappa_2 = \omega(\theta^2 + \delta^2)$, $\kappa_3 = \omega\theta(\theta^2 + 3\delta^2)$, ...
- See Backus, Chernov, and Martin (2011) for details

PDF

- We can recover conditional distributions if we know CF

$$\begin{aligned} p_X(X_{t+1}|X_t) &= (2\pi)^{-1} \int_{-\infty}^{+\infty} e^{-isX_t} f_t(X_{t+1}, s) ds \\ &= \pi^{-1} \int_0^{+\infty} e^{-isX_t + k_t(X_{t+1}, is)} ds \end{aligned}$$

- Inverse Fourier transform

Affine processes

- The processes are defined by the tractability of their CGF

$$k_t(X_{t+1}, s) = \alpha(s) + \beta(s) X_t$$

- The most famous examples are the Vasicek and CIR term structure models

$$\begin{aligned} X_{t+1} &= \bar{X}(1-\phi) + \phi X_t + \sigma \varepsilon_{t+1} && \text{(Vasicek)} \\ X_{t+1} &= \bar{X}(1-\phi) + \phi X_t + \sigma X_t^{1/2} \varepsilon_{t+1} && \text{(CIR)} \end{aligned}$$

- Vasicek: $\alpha(s) = s\bar{X}(1-\phi) + s^2\sigma^2/2$, $\beta(s) = s\phi$
- The CGF for CIR: $\alpha(s) = s\bar{X}(1-\phi)$, $\beta(s) = s\phi + s^2\sigma^2/2$

$$\begin{aligned} & k_t(X_{t+1}, s) = \log E_t e^{s(\bar{X}(1-\phi) + \phi X_t + \sigma X_t^{1/2} \varepsilon_{t+1})} \\ = \quad & s(\bar{X}(1-\phi) + \phi X_t) + s^2 \sigma^2 X_t / 2 \end{aligned}$$

More of affine processes in discrete time

- Poisson mixture of Gammas, or Autoregressive Gamma – ARG(1)
- The CIR process can become negative in discrete time, even if $\sigma^2 < 2\bar{X}(1-\phi)$

$$\begin{aligned} X_{t+1}|j_{t+1} &\sim \mathcal{G}amma((1-\phi)\bar{X}/(\sigma^2/2)+j_{t+1},\sigma^2/2) \\ j_{t+1}|X_t &\sim \mathcal{P}oisson(\phi X_t/(\sigma^2/2)) \end{aligned}$$

- The resulting process is:

$$X_{t+1} = \bar{X}(1-\phi)+\phi X_t+\sigma((1-\phi)\bar{X}/2+\phi X_t)^{1/2}\eta_{t+1}, \eta \text{ is not } \mathcal{N}$$

- cgf is $k_t(X_{t+1},s) = \alpha(s)+\beta(s)X_t$ with

$$\beta(s) = s\phi/(1-s\sigma^2/2),\ \alpha(s) = -(1-\phi)\bar{X}/(\sigma^2/2)\log(1-s\sigma^2/2)$$

- Bertholon, Monfort, and Pegoraro (2008) have many more examples

Entropy

- Definition of entropy: for $x > 0$

$$L(x) \equiv \log E(x) - E(\log x) \geq 0$$

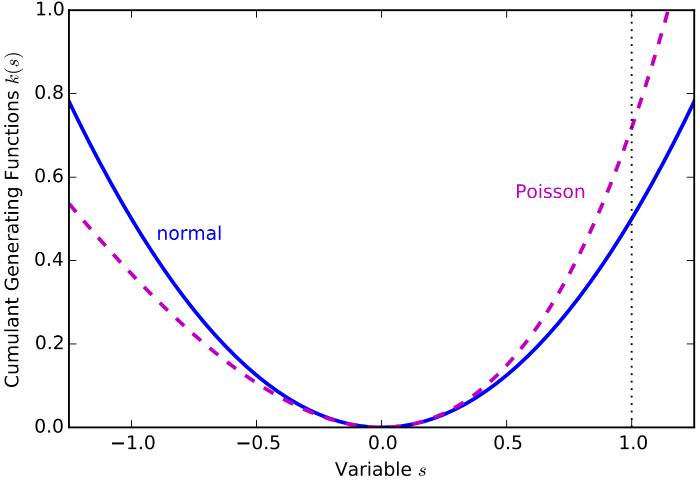
- Entropy and cumulants

$$\begin{aligned} L(x) &= \log E(e^{\log x}) - E(\log x) = k(\log x; 1) - \kappa_1 \\ &= \underbrace{\kappa_2/2!}_{\text{(log)normal term}} + \underbrace{\kappa_3/3! + \kappa_4/4! + \cdots}_{\text{high-order cumulants}} \end{aligned}$$

- Entropy is a measure of dispersion

- Properties:
  - $L(ax) = L(x)$ (a is a constant)
  - $L(xy) = L(x) + L(y)$ (x and y are indepdent)

Normal and Poisson CGF



Vasicek model

- Pricing kernel

$$\log x_{t+1} \quad = \quad x + \theta y_t + \lambda w_{t+1}$$

  with $\{w_t\}$ iid, mean zero, variance one, and cgf $k(s)$ (normality is not required)

- Conditional entropy

$$\begin{aligned} L_t(x_{t+1}) &= k(\lambda) \\ &= \lambda^2/2! + \lambda^3\kappa_3/3! + \lambda^4\kappa_4/4! + \cdots \end{aligned}$$

Coentropy

- Coentropy is a measure of dependence: for $x_1, x_2 > 0$

$$C(x_1, x_2) \equiv L(x_1 x_2) - L(x_1) - L(x_2)$$

- Features
  - Invariant to scaling
  - Equals zero if x_1 and x_2 are independent

- Related to (joint) cgf $k(s_1, s_2) = \log E(e^{s_1 \log x_1 + s_2 \log x_2})$

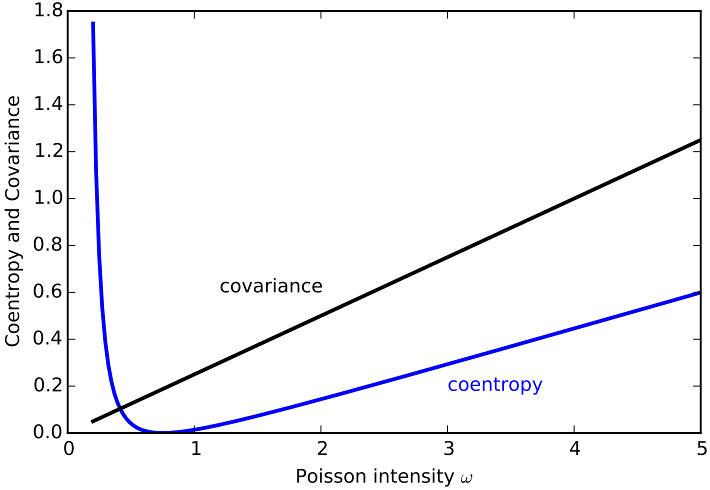
$$C(x_1, x_2) = \underbrace{k(1,1)}_{x_1 x_2} - \underbrace{k(1,0)}_{x_1} - \underbrace{k(0,1)}_{x_2}$$

Coentropy (continued)

- If $\log x = (\log x_1, \log x_2)$ is normal, coentropy = covariance
- Can also be much different
- Example: Poisson mixture ("jump process")
  - Poisson jumps: probability $e^{-\omega}\omega^j/j!$ of $j = 0, 1, 2, \ldots$.
  - Conditional on j, $\log x \sim \mathcal{N}(j\theta, j\Delta)$
- Properties

$$\begin{aligned}
\mathrm{Cov}(\log x_1, \log x_2) &= \omega(\theta_1\theta_2 + \delta_{12}) \\
C(x_1, x_2) &= \omega(e^{(\theta_1+\theta_2)+(\delta_{11}+\delta_{22}+2\delta_{12})/2} - e^{\theta_1+\delta_{11}/2} - e^{\theta_2+\delta_{22}/2})
\end{aligned}$$

Coentropy and covariance



2. GMM

Estimation

- Likelihood is king, because it reflects full information about the hypothesised model
- We like using Markov processes because the likelihood is particularly tractable in this case
- If conditional distributions are not available in the closed form, we have two principle approaches:
  1. Simulation-based methods (here Bayesian MCMC is the most successful tool)
  2. Moment-based methods, such as GMM (here MGF comes in handy)

Problem

- In the standard Consumption CAPM the stochastic discount factor is the ratio of marginal utilities at t and $t+1$.

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-(1-\alpha)}$$

- The Law of One Price (LOOP) for N assets implies the "moment conditions"

$$E_t h_{t+1} = 0$$

where

$$h_{t+1} = \begin{bmatrix} M_{t+1}\left(P_{1t+1} + D_{1t+1}\right) - P_{1t} \\ M_{t+1}\left(P_{2t+1} + D_{2t+1}\right) - P_{2t} \\ \vdots \\ M_{t+1}\left(P_{Nt+1} + D_{Nt+1}\right) - P_{Nt} \end{bmatrix}$$

- How do we use moment conditions to estimate β and α?

Conditioning down

- The Law of Iterated Expectations states that

$$E_{t-1}\left[E_t x_{t+1}\right] = E_{t-1}\left[x_{t+1}\right]$$

- Taking unconditional expectation of the LOOP and using the Law of Iterated Expectation, we get

$$E\left[P_t\right] = E\left[M_{t+1}\left(P_{t+1} + D_{t+1}\right)\right]$$

- Can test the model, but this is a test whether the model is right on average and not whether it holds period by period.
- Use managed portfolios or instruments:

$$P_t z_t = E_t\left[M_{t+1}\left(P_{t+1} + D_{t+1}\right) z_t\right],$$

  where the instrument z_t is any variable known at time t.
- Therefore,

$$E\left[P_t z_t\right] = E\left[M_{t+1}\left(P_{t+1} + D_{t+1}\right) z_t\right],$$

  which is an additional implication of the conditional model.

Moment conditions / GMM

- So, we replace the conditional moments $E_t h_{t+1} = 0$ with the unconditional ones $Ef_{t+1} = 0$, $f_{t+1} = h_{t+1} \otimes z_t$
- Let θ be a $K \times 1$ vector of parameters and define the $(N \times Q) \times 1$ vector of sample moment conditions as:

$$g_T(\theta) \equiv \frac{1}{T} \sum_{t=1}^{T} f_t(\theta)$$

- Estimate the parameters by minimizing a ”squared sum of errors” of the form

$$\hat{\theta} = \arg\min_{\theta} g_T(\theta)' W g_T(\theta) \tag{1}$$

where W is a $(N \times Q) \times (N \times Q)$ positive definite, symmetric weighting matrix.
- If $K < (N \times Q)$, the model is overidentified.

Comments

- Valuation-based moments are not the only ones that can be used
  - Statistical properties of the underlying objects: consumption, returns, etc.
  - MGF is a great way to obtain such moments
- Instruments: something that's helpful in forecasting returns or consumption
- W assigns relative importance to the moments. The statistically optimal one is S^{-1}, where

$$S = \sum_{j=-\infty}^{\infty} E\left[f_t(\theta) f_{t-j}(\theta)^\top\right]$$

  is the covariance matrix of moment conditions.

3. Likelihood

Likelihood of a Markov process

- $F_o(x) \triangleq P[X_0 < x]$

$$
\begin{aligned}
& P[X_0 \in A_0, X_1 \in A_1, ..., X_n \in A_n] \\
= \quad & P[X_n \in A_n | X_0 \in A_0, ..., X_{n-1} \in A_{n-1}] P[X_0 \in A_0, ..., X_{n-1} \in A_{n-1}] \\
\underset{\text{MrkP}}{=} \quad & P[X_n \in A_n | X_{n-1} \in A_{n-1}] P[X_0 \in A_0, ..., X_{n-1} \in A_{n-1}] \\
= \quad & p(t_{n-1}, X_{n-1}, t_n, A_n) P[X_{n-1} \in A_{n-1} | X_0 \in A_0, ..., X_{n-2} \in A_{n-2}] \\
\times \quad & P[X_0 \in A_0, X_1 \in A_1, ..., X_{n-2} \in A_{n-2}] \\
= \quad & \Pi_{i=1}^{n} p(t_{i-1}, X_{i-1}, t_i, A_i) P[X_0 \in A_0] \\
= \quad & \Pi_{i=1}^{n} p(t_{i-1}, X_{i-1}, t_i, A_i) \int_{A_0} dF_0(x)
\end{aligned}
$$

Likelihood of an AR(1) process

- Recall: $X_{t+1} = \bar{X}(1-\phi) + \phi X_t + \sigma\varepsilon_{t+1}$

- The conditional distribution is:

$$p(t_{i-1}, X_{i-1}, t_i, A_i) \equiv p_X(X_{t+1}|X_t) = n(\bar{X}(1-\phi) + \phi X_t, \sigma^2)$$
$$p_X(X_0) = n(\bar{X}, \sigma^2/(1-\phi^2))$$

- Therefore, the (log) likelihood is

$$\begin{aligned} \mathcal{L} &= -(T+1)\log\sigma - 1/(2\sigma^2)\sum_{t=0}^{T-1}(X_{t+1} - \bar{X}(1-\phi) - \phi X_t)^2 \\ &- (1-\phi^2)/(2\sigma^2)(X_0 - \bar{X})^2 + \log(1-\phi^2)/2 \end{aligned}$$

Likelihood of a state-space model

- Suppose we do not observe X_t, but we observe Y_t
- Case 1: $dim(X) = dim(Y)$, and $e \equiv 0$
- Then $X_t = B^{-1}(Y_t - A)$, and

$$p_Y(Y_{t+1}|Y_t) = p_X(B^{-1}(Y_{t+1} - A)|Y_t)\left|B^{-1}\right|$$

- Case 2: $dim(X) < dim(Y)$, and $e \equiv 0$ for the first $dim(X)$ Y's
- Split $Y_t = (Y_t^1, Y_t^2)$, $dim(Y_t^1) = dim(X)$, then $X_t = B_1^{-1}(Y_t^1 - A_1)$,

$$\begin{aligned} & p_Y(Y_{t+1}|Y_t) = p_{Y2}(Y_{t+1}^2|Y_{t+1}^1, Y_t)p_{Y1}(Y_{t+1}^1|Y_t) \\ = \quad & p_e(Y_{t+1}^2 - A_2 - B_2 B_1^{-1}(Y_{t+1}^1 - A_1)|Y_{t+1}^1) \\ \times \quad & p_X(B_1^{-1}(Y_{t+1} - A_1)|Y_t)\left|B_1^{-1}\right| \end{aligned}$$

- Case 3: none of e's is equal to zero, then use Kalman filter (see, e.g., Hamilton 1994)

The Kalman filter

- Goal: construct $\hat{X}_{t|t} = E(X_t | Y^t)$.

- Consider two vectors of normal variables z_1 and z_2 :

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} \right)$$

- Then $z_2 | z_1 \sim N(m, \Sigma)$, where

$$\begin{aligned} m &= E(z_2 | z_1) = \mu_2 + \Omega_{21}\Omega_{11}^{-1}(z_1 - \mu_1), \\ \Sigma &= E((z_2 - m)(z_2 - m)^\top | z_1) = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12} \end{aligned}$$

- We will show that $X_t | Y^{t-1} \sim N(\hat{X}_{t|t-1}, P_{t|t-1})$

- Then we can derive $\hat{X}_{t|t}$ using the result above

The Kalman filter steps

- Initialize: $\hat{X}_{1|0} = 0$, $P_{1|0} = E(X_t X_t^\top)$:

$$vec(P_{1|0}) = [I_{dim(X)^2} - \Phi \otimes \Phi]^{-1} vec(\Sigma\Sigma^\top)$$

(A and $\bar{X}$ are not jointly identified, so set $\bar{X} = 0$.)

- Recursion: assume that $\hat{X}_{t|t-1}$, and $P_{t|t-1}$ are known. Then:
- Forecast of Y_t is

$$E(Y_t|Y^{t-1}) = A + B\hat{X}_{t|t-1} \equiv M_{t|t-1}$$

- The forecast error is

$$Y_t - E(Y_t|Y^{t-1}) = B(X_t - \hat{X}_{t|t-1}) + Ce_t$$

- Thus (e_t is independent of X_t and $\hat{X}_{t|t-1}$):

$$E[(Y_t - E(Y_t|Y^{t-1}))(Y_t - E(Y_t|Y^{t-1}))^\top|Y^{t-1}] = BP_{t|t-1}B^\top + CC^\top \equiv V_{t|t-1}$$

$$E[(Y_t - E(Y_t|Y^{t-1}))(X_t - \hat{X}_{t|t-1})^\top|Y^{t-1}] = BP_{t|t-1}$$

The Kalman filter steps

- Therefore,

$$\left[\begin{array}{c} Y_t|Y^{t-1} \\ X_t|Y^{t-1} \end{array}\right] \sim N\left(\left[\begin{array}{c} A+B\hat{X}_{t|t-1} \\ \hat{X}_{t|t-1} \end{array}\right], \left[\begin{array}{cc} BP_{t|t-1}B^\top + CC^\top & BP_{t|t-1} \\ P_{t|t-1}B^\top & P_{t|t-1} \end{array}\right]\right)$$

- Therefore (using the result), $X_t|Y^t \sim N(\hat{X}_{t|t}, P_{t|t})$, where

$$\begin{aligned} \hat{X}_{t|t} &= \hat{X}_{t|t-1} + P_{t|t-1}B^\top(BP_{t|t-1}B^\top + CC^\top)^{-1}(Y_t - A - B\hat{X}_{t|t-1}) \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}B^\top(BP_{t|t-1}B^\top + CC^\top)^{-1}BP_{t|t-1} \end{aligned}$$

- Final step

$$\begin{aligned} \hat{X}_{t+1|t} &= \Phi\hat{X}_{t|t} \\ P_{t+1|t} &= \Phi P_{t|t}\Phi^\top + \Sigma\Sigma^\top \end{aligned}$$

Comments

- The update of P is deterministic
- One may also smooth the state, that is, compute $\hat{X}_{t|T} = E(X_t|Y^T)$:

$$\hat{X}_{t|T} = \hat{X}_{t|t} + P_{t|t}\Phi^\top P_{t+1|t}^{-1}(\hat{X}_{t+1|T} - \hat{X}_{t+1|t})$$

- Likelihood:

$$\begin{aligned} p_Y(Y_t|Y^{t-1}) &= (2\pi)^{-dim(Y)/2}|V_{t|t-1}|^{-1/2} \\ &\times \exp[-1/2(Y_t - M_{t|t-1})^\top V_{t|t-1}^{-1}(Y_t - M_{t|t-1})]. \end{aligned}$$

Other approaches

- All of this works well for linear observation equations and Gaussian states
- Non-linear observation equations can be approximated using Taylor formula
- Non-Gaussian states can be approximated by the Gaussian ones with the same mean/variance
- A more general approach is to use simulation-based methods

4. Bayesian MCMC

Bayesian MCMC

- Bayesian Markov Chain Monte Carlo as applied to finance problems (e.g., Johannes and Polson, 2009)
- A model connects observable prices Y to state variables X (some unobservable) via parameters Θ.
- Econometrician's task is to estimate the unknown paramteers and state variables.
- Bayesians view all the unknowns as random variables that have some distribution (frequentists view parameters as constants whose estimates have an asymptotically normal distribution).
- In other words, the task is to construct $p(\Theta, X|Y)$

Example: SVOL

- The model

$$
\begin{aligned}
r_t &= \mu + \sigma_{t-1}\varepsilon_t \\
\log \sigma_t^2 &= \omega + \beta \log \sigma_{t-1}^2 + \sigma_v \nu_t
\end{aligned}
$$

  - $Y = \{r_t\}_{t=1}^T$
  - $X = \{\sigma_t\}_{t=1}^T$
  - $\Theta = \{\mu, \omega, \beta, \sigma_v\}$

- How do we construct densities in practice? A histogram!
  - Simply simulate draws $\left\{\Theta^{(g)}, X^{(g)}\right\}_{g=1}^G$ from the target density $p(\Theta, X|Y)$
  - Then for each Θ_i construct a histogram using the available $\Theta_i^{(g)}$
  - Obtain a finite-sample distribution of Θ_i

- How do we simulate from $p(\Theta, X|Y)$?

The Hammerlsey-Clifford theorem

- Consider two simpler problems
  - simulate X from $p(X|Y,\Theta)$
  - simulate Θ from $p(\Theta|Y,X)$
  - HC: the combined (X,Θ) will form a draw from the orginal target $p(\Theta,X|Y)$
- How?
  - Given $\Theta^{(0)}$, draw $X^{(1)} \sim p\left(X|Y,\Theta^{(0)}\right)$
  - Given $X^{(1)}$, draw $\Theta^{(1)} \sim p\left(\Theta|Y,X^{(1)}\right)$
  - ... and so forth
  - The sequence $\left\{\Theta^{(g)},X^{(g)}\right\}_{g=1}^{G}$ forms a Markov chain that converges to draws from $p(\Theta,X|Y)$
- The devil is in the detail: how do we simulate from the marginal densities?
  - Sometimes these densities are known, then one can use standard simulation methods – Gibbs sampler
  - When the densities are unknown, we use the Metropolis-Hastings algorithm

What happens afterwards?

- We can get point estimates of parameters as
  $E(\Theta_i|Y) = \frac{1}{G}\sum_{g=1}^{G}\Theta_i^{(g)}$
- Inference? Construct *finite-sample* confidence bounds by picking pth and $(1-p)$th percentiles from the distribution
- We can estimate the unobservable state variables in a number of ways
  1. Smoothing,
     $p(X_t|Y^T) = \int p(\Theta, X|Y)\, d\Theta dX_{-t} = \frac{1}{G}\sum_{g=1}^{G} p(X_t|\Theta_i^{(g)}, X_{-t}^{(g)}, Y)$
  2. Filtering $p(X_t|Y^t)$
  3. Forecasting $p(X_{t+1}|Y^t)$

General representation of asset pricing models

- Bayes rule:

$$p(A|B) = \frac{P(A,B)}{p(B)} = \frac{P(B|A)P(A)}{p(B)}$$

- Therefore,

$$\begin{aligned} p(\Theta, X|Y) &= \frac{P(Y|\Theta, X)P(\Theta, X)}{p(Y)} \propto P(Y|\Theta, X)P(\Theta, X) \\ &= \underbrace{P(Y|\Theta, X)}_{\text{full info likelihood}} \; \underbrace{P(X|\Theta)}_{\text{dist'n of state}} \; \underbrace{p(\Theta)}_{\text{prior}} \end{aligned}$$

- This representaion will be used over and over while implementing the HC theorem

Gibbs sampler: Black-Scholes

- Stock returns follow

$$r_t = \mu + \sigma\varepsilon_t, \;\; Y = \{r_t\}_{t=1}^{T}, \; \Theta = \left\{\mu, \sigma^2\right\}$$

- The objective is $p\left(\Theta|\,Y\right) = p(\mu, \sigma^2|\,Y)$

- HC suggests the following approach
  - Draw $\mu^{(g+1)} \sim p(\mu|\left(\sigma^2\right)^{(g)}, Y)$; then $\left(\sigma^2\right)^{(g+1)} \sim p(\sigma^2|\mu^{(g+1)}, Y)$

- What are these marginal densities? Use the AP decomposition:

$$\begin{aligned}
p(\mu|\sigma^2, Y) &\propto p(Y|\mu, \sigma^2)p(\mu) \\
p(\sigma^2|\mu, Y) &\propto p(Y|\mu, \sigma^2)p(\sigma^2) \\
p(Y|\mu, \sigma^2) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^T \exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{Y_t - \mu}{\sigma}\right)^2\right)
\end{aligned}$$

Priors

- In order to figure out the posterior distributions, we need to make assumptions about priors.

- Conjugate priors are very attractive because they are tractable and lead to correct posteriors
  - Inverse gamma distribution is often used as a prior for σ^2

$$f(\sigma^2|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\sigma^2\right)^{-\alpha-1} \exp\left(-\beta/\sigma^2\right)$$

  - Normal distribution is often used as a prior for μ

$$f(\mu|\theta,\delta) = \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{1}{2}\left(\frac{\mu-\theta}{\delta}\right)^2\right)$$

Posterior of σ^2

$$
\begin{aligned}
p(\sigma^2|\mu, Y) &\propto p(Y|\mu,\sigma^2) \times p(\sigma^2) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^T \exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{Y_t-\mu}{\sigma}\right)^2\right) \\
&\times \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\sigma^2\right)^{-\alpha-1} \exp\left(-\beta/\sigma^2\right) \\
&\propto \left(\sigma^2\right)^{-T/2-\alpha-1} \exp\left(-\left[\frac{1}{2}\sum_{t=1}^{T}(Y_t-\mu)^2+\beta\right]/\sigma^2\right) \\
&\propto \mathcal{IG}\left(\alpha+\frac{T}{2}, \beta+\frac{1}{2}\sum_{t=1}^{T}(Y_t-\mu)^2\right)
\end{aligned}
$$

Posterior of μ

- Start with

$$\begin{aligned}
p(\mu|\sigma^2, Y) &\propto p(Y|\mu, \sigma^2) \times p(\mu) \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^T \exp\left(-\frac{1}{2}\sum_{t=1}^{T}\left(\frac{Y_t - \mu}{\sigma}\right)^2\right) \\
&\times \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{1}{2}\left(\frac{\mu - \theta}{\delta}\right)^2\right)
\end{aligned}$$

- Next, need a trick

Completing the square

- Denote $\widehat{\mu} = \left(\sum_{t=1}^{T} Y_t \right) / T$ and

$$\begin{aligned}\sum_{t=1}^{T} (Y_t - \mu)^2 &= \sum_{t=1}^{T} (Y_t - \widehat{\mu} + \widehat{\mu} - \mu)^2 \\ &= \sum_{t=1}^{T} (Y_t - \widehat{\mu})^2 + 2(\widehat{\mu} - \mu) \sum_{t=1}^{T} (Y_t - \widehat{\mu}) + \sum_{t=1}^{T} (\widehat{\mu} - \mu)^2 \\ &= \sum_{t=1}^{T} (Y_t - \widehat{\mu})^2 + T(\widehat{\mu} - \mu)^2\end{aligned}$$

- Note that the first member of the sum does not depend on the parameter of interest.

Back to posterior of μ

$$
\begin{aligned}
p(\mu|\sigma^2, Y) &\propto \exp\left(-\frac{T}{2\sigma^2}(\widehat{\mu}-\mu)^2 - \frac{1}{2\delta^2}(\mu-\theta)^2\right) \\
&\propto \exp\left(-\frac{T}{2\sigma^2}\left(-2\widehat{\mu}\mu+\mu^2\right) - \frac{1}{2\delta^2}\left(\mu^2-2\mu\theta\right)\right) \\
&\propto \exp\left(-\frac{1}{2\delta^{*2}}\left(\mu-\left(\frac{T\widehat{\mu}}{\sigma^2}+\frac{\theta}{\delta^2}\right)\delta^{*2}\right)^2\right) \\
&\propto \mathcal{N}\left(\left(\sum_{t=1}^{T} Y_t/\sigma^2+\theta/\delta^2\right)\delta^{*2}, \delta^{*2}\right),
\end{aligned}
$$

where $\delta^{*2} = (T/\sigma^2 + 1/\delta^2)^{-1}$.

Gibbs sampler: Jumps

- Stock returns follow

$$
\begin{aligned}
r_{t+1} &= \mu + \sigma\varepsilon_{t+1} + Z_{t+1}\xi \\
Z_t &= \begin{cases} 1 & \text{with prob } \lambda \\ 0 & \text{with prob } 1-\lambda \end{cases} \\
\xi &\sim \mathcal{N}(\mu_s, \sigma_s^2) \\
Y &= \{r_t\}_{t=1}^T \\
\Theta &= \{\mu, \sigma^2, \lambda, \mu_s, \sigma_s^2\} \\
X &= \{z_t, \xi_t\}_{t=1}^T
\end{aligned}
$$

- The objective is $p(\Theta, X|Y) = p(\Theta, Z, \xi|Y)$
- HC suggests the following approach
  - Draw $\Theta_i^{(g+1)} \sim p(\Theta_i|\Theta_{-i}^{(g)}, Z^{(g)}, \xi^{(g)}, Y)$; then $Z^{(g+1)} \sim p(Z|\Theta^{(g+1)}, \xi^{(g)}, Y)$;then $\xi^{(g+1)} \sim p(\xi|\Theta^{(g+1)}, Z^{(g+1)}, Y)$

Implementing the decomposition

- Use the decomposition above to obtain marginal densities

$$p(\Theta, X|Y) \propto P(Y|\Theta, X)P(X|\Theta)p(\Theta)$$

- Use the model to obtain

$$\begin{aligned} P(Y|\Theta, Z, \xi) &= \prod_{t=1}^{T} p(Y_t|\Theta, Z_t, \xi_t) \\ p(Y_t|\Theta, Z_t, \xi_t) &= n\left(Y_t, \mu + \xi_t Z_t, \sigma^2\right) \end{aligned}$$

- Use the following priors

$$\begin{aligned} \mu &\sim \mathcal{N}\left(\theta, \delta^2\right) \\ \sigma^2 &\sim \mathcal{IG}(\alpha, \beta) \\ \mu_s &\sim \mathcal{N}\left(\theta_s, \delta_s^2\right) \\ \sigma_s^2 &\sim \mathcal{IG}(\alpha_s, \beta_s) \\ \lambda &\sim \mathcal{B}(\gamma, \eta) \end{aligned}$$

where $\mathcal{B}(\gamma, \eta)$ is a Beta distribution $\frac{\Gamma(\gamma+\eta)}{\Gamma(\gamma)\Gamma(\eta)}\lambda^{\gamma-1}(1-\lambda)^{\eta-1}$

Posterior of Θ

- Posteriors of μ and σ^2 are the same as in the BS case
- Posteriors of μ_s and σ_s^2 are by analogy
- Posterior of λ

$$
\begin{aligned}
p(\lambda|\Theta/\lambda, Z, \xi, Y) &= p(\lambda|Z) = p(Z|\lambda)p(\lambda) \propto \prod_{t=1}^{T} p(Z_t|\lambda)p(\lambda) \\
&= \prod_{t=1}^{T} \lambda^{Z_t}(1-\lambda)^{1-Z_t} p(\lambda) \\
&\propto \lambda^{\sum_t Z_t}(1-\lambda)^{T-\sum_t Z_t}\lambda^{\gamma-1}(1-\lambda)^{\eta-1} \\
&\propto \mathcal{B}\left(\sum\nolimits_t Z_t + \gamma, T - \sum\nolimits_t Z_t + \eta\right)
\end{aligned}
$$

Posterior of X_t

- Posterior of ξ_t

$$
\begin{aligned}
p(\xi_t|\Theta, Z_t, Y_t) &\propto p(Y_t|\Theta, Z_t, \xi_t)\, p(\xi_t|\Theta) \\
&\propto \exp\left(-\frac{1}{2}\left(\frac{Y_t-\mu-\xi_t Z_t}{\sigma}\right)^2 - \frac{1}{2}\left(\frac{\xi_t-\mu_s}{\sigma_s}\right)^2\right) \\
&\propto \mathcal{N}\left(\left((Y_t-\mu)Z_t/\sigma^2+\mu_s/\sigma_s^2\right)\sigma_t^{*2}, \sigma_t^{*2}\right)
\end{aligned}
$$

where $\sigma_t^{*2} = \left(Z_t/\sigma^2 + 1/\sigma_s^2\right)^{-1}$

- Posterior of Z_t

$$
\begin{aligned}
p(Z_t &= 1|\Theta, \xi_t, Y_t) \propto p(Y_t|\Theta, Z_t = 1, \xi_t)\, p(Z_t = 1|\Theta) \\
&\propto \exp\left(-\frac{1}{2}\left(\frac{Y_t-\mu-\xi_t}{\sigma}\right)^2\right)\lambda
\end{aligned}
$$

- Similalry, compute $p(Z_t = 0|\Theta, \xi_t, Y_t)$.The "integrating" constant is determined by insuring that the two probs add up to one.

Metropolis – Hastings sampler: SVOL

- Stock returns follow (Jacquier, Polson, and Rossi, 1994)

$$
\begin{aligned}
r_t &= \sqrt{V_t}\varepsilon_t \\
\log V_t &= \omega + \beta \log V_{t-1} + \sigma_v \nu_t \\
Y &= \{r_t\}_{t=1}^T \\
X &= \{V_t\}_{t=1}^T \\
\Theta &= \left\{\omega, \beta, \sigma_v^2\right\}
\end{aligned}
$$

- HC suggests the following approach
  - Draw $(\omega,\beta)^{(g+1)} \sim p(\omega,\beta|\left(\sigma_v^2\right)^{(g)}, X^{(g)}, Y)$; then $\left(\sigma_v^2\right)^{(g+1)} \sim p(\sigma_v^2|(\omega,\beta)^{(g+1)}, X^{(g)}, Y)$; then $X^{(g+1)} \sim p(X|\Theta, Y)$

Posterior of ω and β

- Prior for $m = (\omega, \beta)'$: $\mathcal{N}(\theta, \Delta)$ (bivariate normal, can be independent)
- Posterior

$$
\begin{aligned}
& p(\omega, \beta | \sigma_v^2, X, Y) \propto p(Y | \omega, \beta, \sigma_v^2, X) p(X | \omega, \beta, \sigma_v^2) p(\omega, \beta) \\
\propto \quad & \prod_t p(X_t | X_{t-1}, \Theta)\, p(m) \\
= \quad & \prod_t \frac{1}{X_t \sqrt{2\pi\sigma_v^2}} \exp\left(-\frac{1}{2\sigma_v^2} (\log X_t - \omega - \beta \log X_{t-1})^2 \right) \\
& \times \exp\left(-\frac{1}{2} (m - \theta)' \Delta^{-1} (m - \theta) \right) \\
\propto \quad & \exp\left(-\frac{1}{2\sigma_v^2} \sum_{t=1}^{T} (\log X_t - \omega - \beta \log X_{t-1})^2 \right) \\
& \times \exp\left(-\frac{1}{2} (m - \theta)' \Delta^{-1} (m - \theta) \right)
\end{aligned}
$$

Completing the square

- Introduce additional notation:

$$W_t = (1, \log X_t)', \ U_t = \log X_t \text{ and } \widehat{m} = W'U / W'W$$

- Then,

$$\begin{aligned} & \sum_{t=1}^{T} (\log X_t - \omega - \beta \log X_{t-1})^2 = \sum_{t=1}^{T} \left(U_t - m' W_{t-1}\right)^2 \\ = & \sum_{t=1}^{T} \left(U_t - \widehat{m}' W_{t-1} + \widehat{m}' W_{t-1} - m' W_{t-1}\right)^2 \\ = & \sum_{t=1}^{T} \left(U_t - \widehat{m}' W_{t-1}\right)^2 + (\widehat{m} - m)\, W' W\, (\widehat{m} - m)' \end{aligned}$$

Posterior of ω and β, cont...

- Therefore,

$$\begin{aligned} p(\omega,\beta|\sigma_v^2,X,Y) &\propto \exp\left(-\frac{1}{2\sigma_v^2}(\widehat{m}-m)\,W'W\,(\widehat{m}-m)'\right) \\ &\times \exp\left(-\frac{1}{2}(m-\theta)'\,\Delta^{-1}\,(m-\theta)\right) \\ &\propto \mathcal{N}\left(\left(W'U/\sigma_v^2+\Delta^{-1}\theta\right)\Delta^*,\Delta^*\right) \end{aligned}$$

where $\Delta^* = (W'W/\sigma_v^2+\Delta^{-1})^{-1}$

Posterior of σ_v^2

- Prior of σ_v^2: $IG(\alpha, \beta)$

- Posterior

$$
\begin{aligned}
& p(\sigma_v^2|\omega,\beta,X,Y) \propto p(Y|\omega,\beta,\sigma_v^2,X)p(X|\omega,\beta,\sigma_v^2)p(\sigma_v^2) \\
\propto \quad & \prod\nolimits_t p(X_t|X_{t-1},\Theta)\, p(\sigma_v^2) \\
\propto \quad & \left(\frac{1}{\sqrt{2\pi\sigma_v^2}}\right)^T \exp\left(-\frac{1}{2\sigma_v^2}\sum_{t=1}^{T}(\log X_t - \omega - \beta \log X_{t-1})^2\right) \\
& \times (\sigma_v^2)^{-\alpha-1} \exp\left(-\beta/\sigma_v^2\right) \\
\propto \quad & \mathcal{IG}\left(\alpha + \frac{T}{2}, \beta + \frac{1}{2}\sum_{t=1}^{T}(\log X_t - \omega - \beta \log X_{t-1})^2\right)
\end{aligned}
$$

Posterior of V

Break down $p(X|\Theta, Y)$ into

$$
\begin{aligned}
& p(X_t|X_{-t},\Theta, Y_t) = p(X_t|X_{t-1}, X_{t+1}, \Theta, Y_t) \\
\propto\quad & p(Y_t|X_t,\Theta)p(X_t|X_{t-1},\Theta)\,p(X_{t+1}|X_t,\Theta) \\
\propto\quad & \frac{1}{\sqrt{X_t}}\exp\left(-\frac{Y_t^2}{2X_t}\right)\frac{1}{X_t}\exp\left(-\frac{1}{2\sigma_v^2}\left(\log X_t - \omega - \beta\log X_{t-1}\right)^2\right) \\
& \times\exp\left(-\frac{1}{2\sigma_v^2}\left(\log X_{t+1} - \omega - \beta\log X_t\right)^2\right)
\end{aligned}
$$

Unrecognizable distribution – have to use Metropolis-Hastings

Accept/Reject Method

- This is a general method of drawing samples from distributions that are difficult to simulate from
- Suppose we want to draw from $f(x)$. We have density $g(x)$ such that $f(x) \leq Mg(x)$. This condition implies that f/g is bounded, i.e. g has thicker tails than f.
  1. Generate $X \sim g(x)$ and $U \sim U[0,1]$
  2. $Y = X$ if $U \leq f(X)/(Mg(X)) \Rightarrow Y$ is from the target distribution
  3. Return to step 1 otheriwse

$$P(Y \leq y) = P\left(X \leq y | U \leq \frac{f(X)}{Mg(X)}\right) = \frac{P\left(X \leq y, U \leq \frac{f(X)}{Mg(X)}\right)}{P\left(U \leq \frac{f(X)}{Mg(X)}\right)}$$

$$= \frac{\int\limits_{-\infty}^{y} \int\limits_{0}^{f(x)/(Mg(x))} du g(x) dx}{\int\limits_{-\infty}^{\infty} \int\limits_{0}^{f(x)/(Mg(x))} du g(x) dx} = \frac{\frac{1}{M} \int\limits_{-\infty}^{y} f(x) dx}{\frac{1}{M} \int\limits_{-\infty}^{\infty} f(x) dx} = \int\limits_{-\infty}^{y} f(x) dx$$

Metropolis – Hastings

- We will use a similar approach here
  - Suppose we cannot simulate $\Theta_i^{(g+1)}$ from $\pi(\Theta_i) \equiv p\left(\Theta_i | \Theta_{-i}^{(g+1)}, X, Y\right)$
  - Use proposal density $q\left(\Theta_i^{(g+1)} | \Theta_i^{(g)}\right)$
  - Assume we can compute posterior density ratio $\pi\left(\Theta_i^{(g+1)}\right) / \pi\left(\Theta_i^{(g)}\right)$

- The algorithm
  1. Draw $\Theta_i^{(g+1)}$ from $q\left(\Theta_i^{(g+1)} | \Theta_i^{(g)}\right)$
  2. Accept $\Theta_i^{(g+1)}$ with probability $\alpha\left(\Theta_i^{(g+1)}, \Theta_i^{(g)}\right)$

$$\alpha\left(\Theta_i^{(g+1)}, \Theta_i^{(g)}\right) = \min\left(\frac{\pi\left(\Theta_i^{(g+1)}\right) q\left(\Theta_i^{(g)} | \Theta_i^{(g+1)}\right)}{\pi\left(\Theta_i^{(g)}\right) q\left(\Theta_i^{(g+1)} | \Theta_i^{(g)}\right)}, 1\right)$$

  3. The accepted draw is from the target distribution $\pi(\Theta_i)$.

MH Examples

- Independence MH:

$$q\left(\Theta_i^{(g+1)}|\Theta_i^{(g)}\right) = q\left(\Theta_i^{(g+1)}\right),$$

$$\text{then} \quad \alpha\left(\Theta_i^{(g+1)},\Theta_i^{(g)}\right) = \min\left(\frac{\pi\left(\Theta_i^{(g+1)}\right)q\left(\Theta_i^{(g)}\right)}{\pi\left(\Theta_i^{(g)}\right)q\left(\Theta_i^{(g+1)}\right)},1\right)$$

- Random-walk MH

$$\Theta_i^{(g+1)} = \Theta_i^{(g)} + \varepsilon,$$

$$\text{then} \quad q\left(\Theta_i^{(g)}|\Theta_i^{(g+1)}\right) = q\left(\Theta_i^{(g+1)}|\Theta_i^{(g)}\right),$$

$$\text{then} \quad \alpha\left(\Theta_i^{(g+1)},\Theta_i^{(g)}\right) = \min\left(\frac{\pi\left(\Theta_i^{(g+1)}\right)}{\pi\left(\Theta_i^{(g)}\right)},1\right)$$

Back to posterior of V

- We stopped at

$$
\begin{aligned}
p(X_t|X_{-t},\Theta,Y) &= p(X_t|X_{t-1},X_{t+1},\Theta,Y) \\
\propto &\quad \frac{1}{\sqrt{X_t}}\exp\left(-\frac{Y_t^2}{2X_t}\right)\frac{1}{X_t}\exp\left(-\frac{1}{2\sigma_v^2}\left(\log X_t-\omega-\beta\log X_{t-1}\right)^2\right) \\
\times &\quad \exp\left(-\frac{1}{2\sigma_v^2}\left(\log X_{t+1}-\omega-\beta\log X_t\right)^2\right) \\
\propto &\quad \frac{1}{\sqrt{X_t}}\exp\left(-\frac{Y_t^2}{2X_t}\right)\frac{1}{X_t}\exp\left(-\frac{1}{2\sigma_x^2}\left(\log X_t-\mu_t\right)^2\right) \\
\mu_t &= \left(\omega(1-\beta)+\beta\left(\log X_{t+1}+\log X_{t-1}\right)\right)/\left(1+\beta^2\right), \\
\sigma_x^2 &= \sigma_v^2/\left(1+\beta^2\right)
\end{aligned}
$$

- Jacquier, Polson, and Rossi (1994) use an independence MH as one can closely approximate the true conditional distribution, especially in the tails

Proposal density is an $\mathcal{IG}$

- The first term in the posterior is $\mathcal{IG}\left(-1/2, Y_t^2/2\right)$

- The second (log-normal) term can be approximated by a suitably chosen $\mathcal{IG}(\alpha, \beta)$. Specifically, we are matching the first and second moments of the lognormal $\left(\exp\left(\mu_t + 0.5\sigma_x^2\right), \left(\exp\left(\sigma_x^2\right) - 1\right)\exp\left(2\mu_t + \sigma_x^2\right)\right)$ to the moments of $\mathcal{IG}\left(\beta/(\alpha-1), \beta^2/(\alpha-1)^2/(\alpha-2)\right)$. Therefore,

$$\alpha = \frac{2e^{\sigma_x^2} - 1}{e^{\sigma_x^2} - 1}, \quad \beta = e^{\mu_t + 0.5\sigma_x^2}(\alpha - 1)$$

- Thus, the proposal density is $\mathcal{IG}\left(\widetilde{\alpha}, \widetilde{\beta}\right) = \mathcal{IG}\left(\alpha - \frac{1}{2}, e^{\mu_t + 0.5\sigma_x^2}(\alpha - 1) + \frac{Y_t^2}{2}\right)$

Black-Scholes with returns and options

- Stock returns follow

$$
\begin{aligned}
r_t &= \mu + \sigma \varepsilon_t, \\
c_t &= C_t / S_t = BS(\sigma, r_t) + \varepsilon_t^c, \\
Y &= \{r_t, c_t\}_{t=1}^T, \; \Theta = \{\mu, \sigma^2\}
\end{aligned}
$$

- HC suggests the following approach
  - Draw $\mu^{(g+1)} \sim p(\mu | \left(\sigma^2\right)^{(g)}, Y)$ then $\left(\sigma^2\right)^{(g+1)} \sim p(\sigma^2 | \mu^{(g+1)}, Y)$
- What are these marginal densities? Use the AP decomposition:

$$
\begin{aligned}
p(\mu | \sigma^2, Y) &\propto p(\mu | \sigma^2, r) \propto p(r | \mu, \sigma^2) p(\mu) \\
p(\sigma^2 | \mu, Y) &\propto p(c | \sigma^2, r) p(r | \mu, \sigma^2) p(\sigma^2) \\
p(r | \mu, \sigma^2) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^T \exp\left(-\frac{1}{2} \sum_{t=1}^T \left(\frac{r_t - \mu}{\sigma} \right)^2 \right) \\
p(c | \sigma^2, r) &\propto \prod_{t=1}^T p(c_t | \sigma^2, r_t) \propto \prod_{t=1}^T \exp\left(-\frac{1}{2} \left(\frac{c_t - BS(\sigma, r_t)}{\sigma^c} \right)^2 \right)
\end{aligned}
$$

The independence MH step

- Propose from

$$q(\sigma^2) = p(\sigma^2|\mu, r) \propto p(r|\mu, \sigma^2)p(\sigma^2)$$

- If the prior is $\mathcal{IG}$, so is the posterior (see the Gibbs sampler for BS)

- The algorithm:
  1. Draw $(\sigma^2)^{(g+1)}$ from $q(\sigma^2)$
  2. Accept $(\sigma^2)^{(g+1)}$ with probability:

$$\alpha((\sigma^2)^{(g+1)}, (\sigma^2)^{(g)}) = \min\left(\frac{p(c|(\sigma^2)^{(g+1)}, r)}{p(c|(\sigma^2)^{(g)}, r)}, 1\right)$$

Bibliography I

Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters implied by equity index options, *Journal of Finance* 66, 1967–2010.

Bertholon, Henri, Alain Monfort, and Fulvio Pegoraro, 2008, Econometric asset pricing modelling, *Journal of Financial Econometrics* 6, 407–458.

Hamilton, James, 1994, State space models, in R.F. Engle, and D.L. McFadden, ed.: *Handbook of Econometrics, Volume IV* (Elsvier).

Jacquier, Eric, Nicholas G. Polson, and Peter Rossi, 1994, Bayesian analysis of stochastic volatility models, *Journal of Business and Economic Statistics* 12, 69–87.

Johannes, Michael, and Nicholas G. Polson, 2009, MCMC methods in financial econometrics, in Yacine Aït-Sahalia, and Lars Hansen, ed.: *Handbook of Financial Econometrics, 1-72* (Elsevier: Oxford).