Term Structure Models
Workshop at AFIR-ERM Colloquium, Panama, 2017

Michael Sherris
CEPAR and School of Risk and Actuarial Studies
UNSW Business School
UNSW Sydney
m.sherris@unsw.edu.au

UNSW

August 2017

Workshop Aims

- Overview of commonly used term structure models, both continuous time and discrete time and methods use to fit models.
- Provide participants with an understanding of models and financial and actuarial applications rather than full mathematical details.
- Key topics covered (some more detailed than others) will be:
  - Yield curves, Market instruments, Static Term Structure
  - Concepts of no-arbitrage and risk neutral versus real world probabilities,
  - Main continuous time single factor short rate models,
  - Affine models, Dynamic Nelson-Siegel model, Arbitrage-free Nelson Siegel model,
  - HJM (Heath, Jarrow, Morton) framework,
  - Market models - LMM (Libor Market Models), BGM (Brace, Gatarek, Musiela), SABR (stochastic alpha, beta, rho),
  - Discrete time models and binomial implementation including forward induction.

References:

- Cairns, A. J. (2004), Interest Rate Models: An Introduction, Princeton University Press.
- Diebold, F. X. and Rudebusch, G. D., (2013), Yield Curve Modeling and Forecasting: the Dynamic Nelson-Seigel Approach, Princeton University Press.
- Filipovi´c, Damir (2009), Term-Structure Models A Graduate Course, Springer.
- van Deventer, D. R. and Imai, K. (1997), Financial Risk Analytics: A Term Structure Approach for Banking, Insurance and Investment Management, Irwin Professional Publishing.
- Svoboda, S. (2004), Interest Rate Modelling, Palgrave Macmillan.

My Background

- Late 70's and early 80's: Industry experience in banking, life insurance and pensions - bond markets (RBA open market operations), corporate finance (tax based financing), swaps, bill futures arbitrage.
- Mid 80's through late 90's: Professional actuarial education - finance and investment courses for Actuaries Institute in Australia, co-author of text for SOA "Financial Economics", author of text "Money and Capital Markets: Pricing, Yields and Analysis".
- Since 1985: Academic teaching and research - quantitative finance, financial economics, insurance, longevity risk.
  - Sherris, M. (1994), A One Factor Interest Rate Model and the Valuation of Loans with Prepayment Provisions, Transactions of The Society of Actuaries, Vol 46, 251-320
  - Sherris, M. and A. Ang, (1997), Interest Rate Risk Management: Developments in Interest Rate Term Structure Modeling for Risk Management and Valuation of Interest-Rate-Dependent Cash Flows, North American Actuarial Journal, Vol 1, No 2, 1-26.

Assumed knowledge

- Actuarial students studying actuarial professional courses with financial mathematics, probability and statistics background.
- Practitioners with foundation background in interest rate models seeking a deeper understanding of different model approaches.
- Practitioners seeking an overview of interest rate models.
- Practitioners seeking a refresher of interest rate models.

Actuarial Applications

Actuarial Applications

Term structure models are used in many actuarial applications.

- Valuation of interest dependent cash flows - expected claims at different future dates using a spot rate yield curve.
- Interest rate risk management using the models to quantify cash flow duration and convexity using a spot rate yield curve
- Valuation of option and guarantee features in insurance products that depend on uncertain future yield curves.
- Determination of solvency requirements for economic and regulatory capital purposes.

Yield Curves and Market Instruments

Term Structure Models

- Yields for differing maturities for: coupon paying government bonds, par coupon bonds, zero coupon bonds, FRA's, interest rate futures, interest rate swaps.
- Default-free interest rates (yields to maturity, spot interest rates, forward interest rates) and bond prices for different maturities.
- Market instruments and methods to construct the interest rate term structure to be modelled.
- Model for random changes in (risk-free) interest rates and prices of (default-free) discount bonds.
- Historical yield curve time series (P measure) and market pricing (Q measure).

Spot Rates

- Zero coupon bond prices

$$P(t, T) = \exp\left[-(T-t)\,R(t, T)\right] = \frac{1}{1+(T-t)\,L(t, T)}$$

where $P(t, T)$ is the price at time t of \$1 payable at time T (zero coupon bond), $R(t, T)$ is the continuous compounding yield to maturity and $L(t, T)$ simple interest yield to maturity (similar to LIBOR rate).

- Spot interest rates (spot rates - continuous compounding)

$$R(t, T) = -\frac{\log P(t, T)}{T-t}$$

Forward Rates

- Forward rate agreement (FRA) - agreement to fix an interest rate for a future time period (time T to S) at time t

$$\begin{aligned} P(t,T) &= P(t,S)\exp\left[(S-T)F(t,T,S)\right] \\ &= P(t,S)\left[1+(S-T)L(t,T,S)\right] \end{aligned}$$

- Forward rate at time t (continuous compounding) for time T to S $(S > T)$

$$F(t,T,S) = \frac{1}{S-T}\log\frac{P(t,T)}{P(t,S)}$$

and simple forward rate (similar to LIBOR forward rate)

$$L(t,T,S) = \frac{1}{S-T}\left[\frac{P(t,T)}{P(t,S)} - 1\right]$$

Coupon Paying Bonds

- Most bonds traded are coupon paying bonds including government issued bonds (regarded often as default free)
- Price of coupon paying bond paying N coupons (in arrears) with C_i paid at time T_i and nominal (face value) N paid at maturity T_N

$$\begin{aligned} P_{T_N}(t) &= \sum_{i=1}^{N} P(t, T_i) C_i + P(t, T_N) N \\ &= \sum_{i=1}^{N} \frac{1}{[1 + Y(T_N)]^{t_i}} C_i + \frac{1}{[1 + Y(T_N)]^{T_{Ni}}} N \end{aligned}$$

where $Y(T_N)$ is the yield to maturity for the T_N maturity coupon bond.

- Market conventions - most government coupon bonds pay semi-annual coupons and are quoted as semi-annual compounding yields. Times for price quotes are usually between coupon dates (broken periods) and sometimes quoted with accrued interest and sometimes without (dirty and clean prices).

Par Coupon Bonds

- Yields to maturity for coupon paying bonds often used to construct default free term structure using "bootstrapping".
- Quoted market yields are impacted by many factors such as different size of coupons, tax, liquidity.
- Smoothing is usually used to extract a "par yield curve" - the yield curve for coupons bonds with coupons equal to the yield to maturity payable on the same coupon dates and all priced on a coupon date, hence priced at par (face value usually taken as $N = 100$). Par yield to maturity (equal to par coupon rate) given by

$$
\begin{aligned}
N &= c_{T_N} N \sum_{i=1}^{N} P(t, T_i) + P(t, T_N) N \\
c_{T_N} &= \frac{1 - P(t, T_N)}{\sum_{i=1}^{N} P(t, T_i)}
\end{aligned}
$$

Futures and Swaps

- Interest rate futures (Eurodollar futures, futures on bank bills and bonds) are similar to forwards and have mark to market and deposit requirements
- Futures rates require a convexity adjustment

$$\text{Forward rate} = \text{Futures rate } - \frac{1}{2}\sigma^2 (T-t)^2$$

- Interest rate swaps - swap fixed for floating coupons, value of payer swap is:

$$\begin{aligned} \Pi_p(t) &= N\left(P(t,T_0) - \left[\delta K \sum_{i=1}^{n} P(t,T_i) + P(t,T_n)\right]\right) \\ &= N\left(\delta \sum_{i=1}^{n} P(t,T_i)\left[F(t,T_{i-1},T_i) - K\right]\right) \end{aligned}$$

where N is nominal value, K is fixed rate, δ is time between swap payments.

- Par swap rates have $\Pi_p(t) = 0$, similar to par coupon bond rates.

Bootstrapping Par Yields

- Given the par yield to maturity curve for coupon paying bonds then spot interest rates and forward interest rates can be determined.
- All of these are equivalent but spot rates and forward rates can be used more flexibly to value more general cash flows.
- Spot rates are determined working forward by maturity using bootstrapping.
- The first spot rate (continuous compounding) is given using the first par coupon rate:

$$\begin{aligned} 1 + c_{T_1} &= \exp\left[-R\left(t, T_1\right)\right] \\ R\left(t, T_1\right) &= -\log\left[1 + c_{T_1}\right] \end{aligned}$$

- Then having solved for $R\left(t, T_j\right)$ for $j = 1$ to $i - 1$, solve period by period, for $R\left(t, T_i\right)$ using

$$1 = c_{T_i} \sum_{j=1}^{i} \exp\left[-R\left(t, T_j\right)\right] + 1 \exp\left[-R\left(t, T_i\right)\right]$$

Bootstrapping Par Yields

- Rearranging gives spot rates (continuous compounding)

$$R\left(t, T_i\right) = -\log\left[\frac{1 + c_{T_i}}{1 - c_{T_i}\sum_{j=1}^{i-1}\exp\left[-R\left(t, T_j\right)\right]}\right] \quad i = 2, \ldots, N$$

- Then forward rates (continuous compounding) are derived from the spot rates

$$F\left(t, T_{i-1}, T_i\right) = \frac{1}{T_i - T_{i-1}}\left[R\left(t, T_i\right) - R\left(t, T_{i-1}\right)\right]$$

Yield Curve Smoothing

- Methods used to smooth market bond price/yield curve data include:
  - Weighted Least Squares or Regression, MLE or Bayesian
- Fit discount factors, spot rates or forward rates with:
  - Splines,
  - Nelson-Siegel or Svensson Parametric Curve, and similar exponential-polynomial class (see Cairns (2004) Chapter 12),
  - Smith-Wilson Kernel method

Yield Curve Smoothing

- General approach

$$\mathbf{P} = \mathbf{Cd} + \epsilon$$

where

  - $\mathbf{P}$ is (column) vector of market prices,
  - $\mathbf{C}$ is the cash flow matrix with cash flows for the market instruments used to fit term structure,
  - $\mathbf{d}$ discount factors for the term structure (determined from smoothed yield curve in terms of yields to maturity, spot rates or forward rates),
  - ϵ vector of errors to be minimised in the smoothing of market data

- Issues:

  - smoothing spot rates results in saw tooth forward rates so better to smooth forward rates,
  - market instruments include coupon bonds (short maturities, illiquidity) or futures and swaps (longer maturities, more active trading),
  - need for interpolation and extrapolation.

Term Structure Models - R packages

- R packages for Term Structure
  - **termstrc** cubic splines approach of McCulloch (1971, 1975), the Nelson and Siegel (1987) method with extensions by Svensson (1994),Diebold and Li(2006) and DePooter(2007). Weighted constrained optimization procedure with analytical gradients and a globally optimal start parameter search algorithm.
  - **YieldCurve** fits Nelson-Siegel, Diebold-Li and Svensson.

Instantaneous Forward Rates and the Short Rate

- Instantaneous forward rate at time t

$$f(t, T) = \lim_{S \to T} F(t, T, S) = -\frac{\partial}{\partial T} \log P(t, T) = -\frac{\partial P(t, T) / \partial T}{P(t, T)}$$

which gives

$$P(t, T) = \exp\left[-\int_t^T f(t, u)\, du\right] = \exp\left[-R(t, T)(T - t)\right]$$

and

$$f(\tau) = R(\tau) + \tau \frac{\partial}{\partial \tau} R(\tau)$$

where $\tau = T - t$.

- Short rate at time t is

$$r(t) = \lim_{T \to t} R(t, T) = R(t, t) = f(t, t)$$

:

Static Nelson-Siegel

- Parametric formula for instantaneous forward yield curve (cross sectional)

$$f(\tau) = \beta_1 + \beta_2 e^{-\lambda\tau} + \beta_3 \lambda\tau e^{-\lambda\tau}$$

- Nelson-Siegel spot rate yield curve (integrate forward yield curve)

$$y(\tau) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

- Desirable features:
  - $P(0) = 0$ and $\lim_{\tau\to\infty} P(\tau) \to 0$
  - instantaneous short rate $\lim_{\tau\to 0} y(\tau) = f(0) = r$ (equals $\beta_1 + \beta_2$)
  - $\lim_{\tau\to\infty} y(\tau) = \beta_1$ (the long term interest rate)
  - flexible shapes - flat, increasing, decreasing, humped, U-shaped (β_3 determines size and shape of hump)

Dynamic Nelson-Siegel

- Static Nelson Siegel fits constant parameters $\beta_1, \beta_2, \beta_3, \lambda$ to produce a best fit term structure
- Dynamic Nelson Siegel treats parameters $\beta_1, \beta_2, \beta_3$ as variables, or latent factors, and the coefficients as factor loadings

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$$

- Factor loadings
  - 1, the loading on β_{1t}, impacts on all yields, but impacts long term yields relatively more compared to other factors
  - $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)$, the loading on β_{2t}, starts at 1 and decreases quickly to zero - impacts short term yields most, a short term factor,
  - $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$, the loading on β_{3t}, starts at 0, increases and then decreases to zero - impacts medium term yields most, a medium term factor

Svensson Yield Curve

- Parametric formula for instantaneous forward yield curve (cross sectional)

$$f(\tau) = \beta_1 + (\beta_2 + \beta_3 \lambda_1 \tau) e^{-\lambda_1 \tau} + \beta_4 \lambda_2 \tau e^{-\lambda_2 \tau}$$

- Svensson spot rate yield curve (integrate forward yield curve)

$$\begin{aligned} y(\tau) &= \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} - e^{-\lambda_1 \tau} \right) \\ &\quad + \beta_4 \left(\frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} - e^{-\lambda_2 \tau} \right) \end{aligned}$$

  - allows for second hump in the yield curve.

Discussion - Static Term Structure

- Term structure from a static perspsective (P and Q measures).
- Different ways of representing the yield curve (level, slope, curvature).
- Different market instruments used to fit the yield curve (bonds, futures, forwards, swaps).
- Methods of fitting term structure yield curves (yield curve smoothing).
- Parametric form of cross sectional yield curve and links to dynamic models (Nelson-Siegerl).

Concepts of no-arbitrage and risk neutral versus real world probabilities

PDE Approach - Short Rate Models

- Start from stochastic process for interest rates (short interest rate),
- Use Ito formula to derive stochastic process for bond prices (bonds are contingent claims or derivatives on interest rates),
- Require no-arbitrage and derive partial differential equation (PDE) for (zero coupon) bond prices,
- Solve PDE subject to boundary conditions.

Brownian Motion

Sometimes called a "Wiener process"
A stochastic process $\{W(t),\ t \geq 0\}$ is said to be a standard Brownian motion process, or simply *Brownian motion*, if:
$W(0) = W_0 = 0$;
$W(t)$ has stationary and independent increments;
and for every $t > 0$,

$$W(t) \sim N(0, t).$$

A stochastic process $\{X(t)\ \ t \geq 0\}$ is said to be a Brownian motion process with drift coefficient μ and variance parameter σ^2 if:
$X(0) = X_0 = 0$;
$\{X(t)\ \ t \geq 0\}$ has stationary and independent increments; and
for every $t > 0$, $X(t) \sim N\left(\mu t, \sigma^2 t\right)$.

Standard Brownian motion

Integral from

$$W(t) = W_0 + \int_0^t dW(u)$$

Properties

- $W(t)$ is continuous
- $W(t)$ is nowhere differentiable
- $W(t)$ is process of unbounded variation
- $W(t)$ is process of bounded quadratic variation
- Conditional distribution of $W(u) \,|\, W(t)$ for $u > t$ is normal with mean $W(t)$ and variance $(u - t)$
- Variance of a forecast of $W(u)$ increases indefinitely as $u \to \infty$

Generalized univariate Wiener process

Drift and volatility depend on $X(t)$ and t

$$dX(t) = \alpha(X(t), t)\, dt + \sigma(X(t), t)\, dW(t)$$

Special cases:
Arithmetic Brownian motion
Geometric Brownian motion
Mean reverting process

Dynamics for function of Brownian motion $f(X(t), t)$.

Arithmetic Brownian Motion

$$\begin{aligned} dX &= \alpha dt + \sigma dW \\ \alpha\left(X\left(t\right), t\right) &= \alpha \\ \sigma\left(X\left(t\right), t\right) &= \sigma \end{aligned}$$

The process defined by

$$W\left(t\right) = \frac{X\left(t\right) - \alpha t}{\sigma},$$

is a standard Brownian motion.
If $\left\{W\left(t\right),\ t \geq 0\right\}$ is a standard Brownian motion, then by defining

$$X\left(t\right) = \alpha t + \sigma W\left(t\right),$$

the process $\left\{X\left(t\right),\ t \geq 0\right\}$ is a Brownian motion with a drift α and volatility σ.

Arithmetic Brownian Motion

- $X(t)$ may be positive or negative
- distribution of $X(u)$ given $X(t)$ for $u > t$ is normal with mean $X(t) + \alpha(u - t)$ and standard deviation $\sigma\sqrt{(u - t)}$
- variance tends to infinity as u goes to infinity (variance grows linearly with time)

Geometric Brownian Motion

- If $X(t)$ starts at positive value then it remains positive
- $X(t)$ has an absorbing barrier at 0
- Conditional distribution of $X(u)$ given $X(t)$ for $u > t$ is lognormal. Conditional mean of $\ln X(u)$ given by $\ln X(t) + \alpha(u-t) - \frac{1}{2}\sigma^2(u-t)$ and conditional standard deviation of $\ln X(u)$ is $\sigma\sqrt{(u-t)}$.
- Often used to model values - positive and increases at exponential rate.

Geometric Brownian Motion

- $E\left(X\left(t\right)|X\left(u\right),\ 0\leq u\leq s\right)$

$$\begin{aligned}
&= E\left(e^{Y(t)}|X\left(u\right),\ 0\leq u\leq s\right)\\
&= E\left(e^{Y(s)+Y(t)-X(s)}|X\left(u\right),\ 0\leq u\leq s\right)\\
&= e^{Y(s)}E\left(e^{Y(t)-Y(s)}|X\left(u\right),\ 0\leq u\leq s\right)\\
&= Y\left(s\right)\cdot\exp\left[\alpha\left(t-s\right)\right].
\end{aligned}$$

since $\left(Y\left(t\right)-Y\left(s\right)\right)\sim N\left(\left(\alpha-\frac{1}{2}\sigma^2\right)\left(t-s\right),\sigma^2\left(t-s\right)\right)$, then

$$\begin{aligned}
E\left(e^{Y(t)-Y(s)}\right) &= e^{\left[\left(\alpha-\frac{1}{2}\sigma^2\right)(t-s)+\frac{1}{2}\sigma^2(t-s)\right]}\\
&= e^{[\alpha(t-s)]}
\end{aligned}$$

Mean Reverting Process

$$
\begin{aligned}
dX &= \kappa(\mu - X)\,dt + \sigma X^{\gamma} dW \\
\alpha(X(t), t) &= \kappa(\mu - X) \quad \kappa > 0 \\
\sigma(X(t), t) &= \sigma X^{\gamma}
\end{aligned}
$$

κ is speed of adjustment parameter, μ long run mean and σ is volatility parameter

- $X(t)$ is positive as long as $X(t)$ starts positive
- As $X(t)$ approaches zero, the drift is positive and volatility vanishes
- As $u \to \infty$ the variance of $X(u)$ is finite

Mean Reverting Process

- If $\gamma = \frac{1}{2}$ the conditional distribution of $X(u)$ given $X(t)$ for $u > t$ is non-central chi-squared with mean

$$(X(t) - \mu)\exp[-\kappa(u-t)] + \mu$$

and variance

$$X(t)\left(\frac{\sigma^2}{\kappa}\right)\left(\begin{array}{c}\exp[-\kappa(u-t)] - \\ \exp[-2\kappa(u-t)]\end{array}\right)$$
$$+\mu\left(\frac{\sigma^2}{2\kappa}\right)(1-\exp[-\kappa(u-t)])^2$$

- CIR or square root process

One-Dimensional Ito's Formula

Consider the SDE

$$dX = \alpha(X(t), t)dt + \sigma(X(t), t)dW(t) = \alpha dt + \sigma dW$$

and suppose that we have another stochastic process $f\left(X\left(t\right), t\right) = f_t$ which depends on $X\left(t\right)$ and t. *Itô's formula* tells us this process satisfies the following:

$$df_t = \frac{\partial f}{\partial X}dX + \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}\sigma_t^2 dt.$$

This formula is also known as Itô's formula (stochastic Taylor series).

$$\begin{aligned} df &= \frac{\partial f}{\partial X}dX + \frac{\partial f}{\partial t}dt \\ &\quad + \frac{1}{2}\left[\frac{\partial^2 f}{\partial X^2}dX^2 + 2\frac{\partial^2 f}{\partial X \partial t}dXdt + \frac{\partial^2 f}{\partial t^2}dt^2\right] + o\left(dt\right) \\ &= \left[\alpha\frac{\partial f}{\partial X} + \frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial X^2}\right]dt + \sigma\frac{\partial f}{\partial X}dW \end{aligned}$$

Bivariate Ito's Formula

Consider the SDE's

$$\begin{aligned} dX &= \alpha(X, Y, t)\, dt + \sigma(X, Y, t)\, dW_1 \\ dY &= \beta(X, Y, t)\, dt + \nu(X, Y, t)\, dW_2 \\ dW_1 dW_2 &= E[dW_1 dW_2] = \rho dt \end{aligned}$$

Probabilistically, $dW_2 = \rho dW_1 + \sqrt{(1-\rho^2)}dZ$

$$\begin{aligned} & df(X, Y, t) = \frac{\partial f}{\partial X} dX + \frac{\partial f}{\partial Y} dY + \frac{\partial f}{\partial t} dt \\ & + \frac{1}{2}\left[\frac{\partial^2 f}{\partial X^2} dX^2 + 2\frac{\partial^2 f}{\partial X \partial Y} dX dY + \frac{\partial^2 f}{\partial Y^2} dY^2\right] + o(dt) \\ = & \left[\alpha \frac{\partial f}{\partial X} + \beta \frac{\partial f}{\partial Y} + \frac{\partial f}{\partial t} + \frac{1}{2}\left[\sigma^2 \frac{\partial^2 f}{\partial X^2} + 2\rho\sigma\nu \frac{\partial^2 f}{\partial X \partial Y} + \nu^2 \frac{\partial^2 f}{\partial Y^2}\right]\right] dt \\ & + \sigma \frac{\partial f}{\partial X} dW_1 + \nu \frac{\partial f}{\partial Y} dW_2 \end{aligned}$$

Covariation Process and Quadratic Variation

Consider

$$dX_i(t) = a_i(t)\,dt + b_i(t)\,dW_i(t)$$

then quadratic covariation process for $X_i(t)$ and $X_j(t)$ is

$$\begin{aligned} \langle X_i, X_j \rangle (t) &= \int_0^t b_i(u)\, b_j(u)\, du \\ d\langle X_i, X_j \rangle (t) &= b_i(t)\, b_j(t)\, dt \text{ sometimes written as } dX_i(t)\, dX_j(t) \end{aligned}$$

Quadratic variation is $\langle X, X \rangle (t) = \int_0^t b(u)^2\, du$.
Product rule

$$d(X(t)\,Y(t)) = X(t)\,dY(t) + Y(t)\,dX(t) + d\langle X, Y \rangle (t)$$

Integration by parts formula

$$X(t)\,Y(t) = X(0)\,Y(0) + \int_0^t X(s)\,dY(s) + \int_0^t Y(s)\,dX(s) + \langle X, Y \rangle (t)$$

Feynman-Kac formula

Consider the pde

$$\left[\alpha\left(X,t\right)\frac{\partial f}{\partial X}+\frac{\partial f}{\partial t}+\frac{1}{2}\sigma^{2}\left(X,t\right)\frac{\partial^{2}f}{\partial X^{2}}\right]-rf+c\left(X,t\right)=0$$

defined on interaval $[0,T]$ with terminal condition

$$f\left(X,T\right)=\Psi\left(X\right)$$

Then Feynman-Kac formula is solution in terms of conditional expectation

$$f\left(X,t\right)=E\left[\begin{array}{c}\int_{t}^{T}\exp\left(-\int_{t}^{s}rd\tau\right)c\left(X,s\right)ds\\+\exp\left(-\int_{t}^{T}rd\tau\right)\Psi\left(X\right)|X_{t}=x\end{array}\right]$$

where X is a diffusion process

$$dX=\alpha\left(X,t\right)dt+\sigma\left(X,t\right)dW$$

PDE for Bond Pricing

Model short interest rate with a diffusion process (single factor)

$$dr(t) = \alpha(r,t)\,dt + \sigma(r,t)\,dZ$$

A bond has a value $P(r,t,T) = P(r,\tau)$ with dynamics (from Itô's formula)

$$\begin{aligned}
dP(r,t,T) &= \frac{\partial P(r,t,T)}{\partial r}dr + \frac{1}{2}\frac{\partial^2 P(r,t,T)}{\partial r^2}dr^2 + \frac{\partial P(r,t,T)}{\partial t}dt \\
&= \frac{\partial P}{\partial r}(\alpha dt + \sigma dZ) + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2 dt + \frac{\partial P}{\partial t}dt \\
&= \left[\alpha\frac{\partial P}{\partial r} + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2 + \frac{\partial P}{\partial t}\right]dt + \frac{\partial P}{\partial r}\sigma dZ \\
&= P\left[mdt + sdZ\right]
\end{aligned}$$

where

$$m = \frac{1}{P}\left(\alpha\frac{\partial P}{\partial r} + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2 + \frac{\partial P}{\partial t}\right) \text{ and } s = \frac{1}{P}\frac{\partial P}{\partial r}\sigma$$

PDE for Bond Pricing

Construct a portfolio, V, of two bonds, V_1 and V_2, with different maturities $\tau_1 > \tau_2$

$$\begin{aligned} V &= V_1(r, \tau_1) + V_2(r, \tau_2) \\ dV &= dV_1 + dV_2 = V_1[m_1 dt + s_1 dZ] + V_2[m_2 dt + s_2 dZ] \\ &= [V_1 m_1 + V_2 m_2]\, dt + [V_1 s_1 + V_2 s_2]\, dZ \end{aligned}$$

Now if we choose V_1 and V_2 so that the volatility of the portfolio is zero

$$\begin{aligned} [V_1 s_1 + V_2 s_2] &= 0 \\ \frac{V_1}{V_2} &= -\frac{s_2}{s_1} \end{aligned}$$

and

$$\begin{aligned} (V - V_2)\, s_1 + V_2 s_2 &= 0 \\ V_2 &= \frac{V s_1}{s_1 - s_2} \text{ and } V_1 = -\frac{V s_2}{s_1 - s_2} \end{aligned}$$

PDE for Bond Pricing

We then have, since the portfolio is now riskless,

$$[V_1 m_1 + V_2 m_2] = rV = r\left(V_1 + V_2\right)$$

$$-\frac{Vs_2}{s_1 - s_2} m_1 + \frac{Vs_1}{s_1 - s_2} m_2 = rV$$

or, rearranging,

$$\frac{(m_1 - r)}{s_1} = \frac{(m_2 - r)}{s_2} = \lambda$$

where λ is the market price of (interest rate) risk. The excess return per unit of risk is the same for all maturity bonds.
We vary V_1 so the portfolio is self financing

$$dV = V\left(\frac{m_2 s_1 - m_1 s_2}{s_1 - s_2}\right) dt$$

PDE for Bond Pricing

Substituting expressions for m and s we obtain the PDE for bond prices

$$\begin{aligned}
\frac{1}{P}\left(\alpha\frac{\partial P}{\partial r}+\frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2+\frac{\partial P}{\partial t}\right)-r &= \frac{1}{P}\frac{\partial P}{\partial r}\sigma\lambda \\
\frac{\partial P}{\partial t}+(\alpha-\sigma\lambda)\frac{\partial P}{\partial r}+\frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2-rP &= 0
\end{aligned}$$

Then solve subject to boundary conditions.
Solving PDEs using known solutions for particular cases, guesses or Feynman-Kac.

PDE for Bond Pricing - Example

Consider random walk for yield $(y(0) = r)$

$$P(\tau) = e^{-y\tau} \text{ and } dy = \sigma dZ$$

$$\frac{\partial P}{\partial t} = yP \quad \frac{\partial P}{\partial y} = -\tau P \quad \frac{\partial^2 P}{\partial y^2} = \tau^2 P$$

From Ito formula

$$dP = \left(y + \frac{1}{2}\sigma^2\tau^2\right) Pdt - \sigma\tau PdZ$$

Hence

$$m = y + \frac{1}{2}\sigma^2\tau^2 \text{ and } s = \sigma\tau$$

and risk premium for interest rate risk gives

$$\frac{\left(y + \frac{1}{2}\sigma^2\tau^2\right) - r}{\sigma\tau} = \lambda$$

PDE for Bond Pricing - Example

In this case

$$\begin{aligned} y(\tau) &= r + \lambda\sigma\tau - \frac{1}{2}\sigma^2\tau^2 \\ &= \text{ risk free rate } + \text{ price of risk } \times \text{ risk} - \text{convexity adjustment} \end{aligned}$$

Note that

- yield curve is quadratic in maturity
- maximum at $\frac{\lambda}{\sigma}$, so humped shaped
- can have negative yields
- instantaneous forward rates are quadratic

$$\begin{aligned} f(\tau) &= -\frac{\partial}{\partial\tau}\log P(\tau) = -\frac{\partial}{\partial\tau}\left(-y(\tau)\tau\right) = y(\tau) + \tau\frac{\partial}{\partial\tau}\left(y(\tau)\right) \\ &= r + \lambda\sigma\tau - \frac{1}{2}\sigma^2\tau^2 - \tau\left(\lambda\sigma - \sigma^2\tau\right) = r + 2\lambda\sigma\tau - \frac{3}{2}\sigma^2\tau^2 \end{aligned}$$

Martingale Approach for Bond Pricing

- Background on Martingales
- Change of Measure
- Fundamental Theorem of Asset Pricing
- Bond Pricing as an Expectation under Risk Neutral Measure

Continuous-Time Martingales

We denote the filtration generated by the process at time t by $\mathcal{F}_t$ and is interpreted as the information (or history) of the process up until time t. We shall assume the process $\{W(t)\}$ is $\mathcal{F}_t$-adapted, which means that given this history $\mathcal{F}_t$, the value of the process at t, $W(t)$, is known. The stochastic process $\{W(t),\ t \geq 0\}$ is said to be a *martingale* if: $E[|W(t)|] < \infty$, i.e. finite expectation of the absolute value, and for all $s < t$, we have

$$E(W(t)|\mathcal{F}_s) = W(s).$$

The best forecast of unobserved future values is the most recently available observed value of the process.

Example *(Brownian Motion)*
Consider a Brownian motion $\{X_t,\ t \geq 0\}$ with drift parameter α and variance parameter σ^2. Then, for $s < t$, we have

$$\begin{aligned} E\left(X_t \left| \mathcal{F}_s\right.\right) &= E\left[X_s + \left(X_t - X_s\right) \left| \mathcal{F}_s\right.\right] \\ &= E\left[X_s \left| \mathcal{F}_s\right.\right] + E\left[X_t - X_s \left| \mathcal{F}_s\right.\right] \\ &= X_s + \alpha\left(t - s\right). \end{aligned}$$

since

$$\left(X_t - X_s\right) \sim N\left(\alpha\left(t - s\right), \sigma^2\left(t - s\right)\right).$$

Thus, we see that $E\left(X_t \left| \mathcal{F}_s\right.\right) = X_s$ only if $\alpha = 0$.
A B.M. with a zero drift is a martingale.

Continuous-Time Martingales

Example *(Geometric Brownian Motion)*
Consider a geometric Brownian motion $\{Y_t,\ t \geq 0\}$ defined as

$$Y_t = \exp\left(X_t\right),$$

where X_t is a B.M. with drift parameter μ and variance parameter σ^2. Then, for $s < t$, we have

$$\begin{aligned}
E\left(Y_t \,|\mathcal{F}_s\right) &= E\left[\exp\left(X_t\right)|\mathcal{F}_s\right] \\
&= E\left[\exp\left(X_s + \left(X_t - X_s\right)\right)|\mathcal{F}_s\right] \\
&= E\left[\exp\left(X_s\right)|\mathcal{F}_s\right] \times E\left[\exp\left(X_t - X_s\right)|\mathcal{F}_s\right] \\
&= Y_s \times E\left[\exp\left(X_t - X_s\right)|\mathcal{F}_s\right] \\
&= Y_s \times \exp\left[\left(\mu + \frac{1}{2}\sigma^2\right)(t-s)\right]
\end{aligned}$$

Continuous-Time Martingales

Recalling the moment generating function of a normal distribution, we have

$$\begin{aligned} E\left(Y_t \,|\mathcal{F}_s\right) &= Y_s \times \exp\left[\alpha\left(t-s\right)+\frac{1}{2}\sigma^2\left(t-s\right)\right] \\ &= Y_s \times \exp\left[\left(\alpha+\frac{1}{2}\sigma^2\right)\left(t-s\right)\right]. \end{aligned}$$

Thus, we see that the geometric Brownian motion is not a martingale unless of course if $\alpha = -\frac{1}{2}\sigma^2$.

Continuous-Time Martingales

We can transform the geometric B.M. to form a martingale by defining the process

$$Z_t = \exp\left(W(t) - \frac{1}{2}\sigma^2 t\right).$$

By following the same argument as above, it is left as an exercise to prove that the process $\{Z_t,\ t \geq 0\}$ is a martingale if the Brownian motion has zero drift.

Stochastic Differential Equations

Definition: A stochastic differential equation, or simply SDE, has the following form:

$$dX_t = \mu\left(X_t, t\right) dt + \sigma\left(X_t, t\right) dB_t,$$

where B_t is the standard Brownian motion process.
The drift $\mu\left(X_t, t\right)$ and $\sigma\left(X_t, t\right)$ can be functions of the process X_t and time t.
A stochastic process X_t that satisfy the above SDE is sometimes called a *diffusion process.*

Stochastic Integration

$\{W(t),\ t \geq 0\}$ is a Brownian motion with zero drift and variance parameter σ^2

Let f be a function with continuous derivative on the interval $[a, b]$.

$$\int_a^b f(t)\, dW(t) = \lim_{\substack{n\to\infty \\ \max(t_k - t_{k-1})\to 0}} \sum_{k=1}^{n} f(t_{k-1})\left(W(k) - W(k-1)\right)$$

where $a = t_0 < t_1 < \cdots < t_n = b$ is a partition of $[a, b]$.

This definition of a stochastic integral is sometimes called the "Ito integral".

Stochastic Integration

To simplify notations, $W(k)$ is the value of the stochastic process at time t_k.
The integral is itself another stochastic process which we can label as

$$Y_t = \int_a^b f(t)\, dW(t).$$

The integration by parts formula applied to sums is given by

$$\sum_{k=1}^{n} f(t_{k-1})\left(W(k) - W(k-1)\right) = \left[f(b)\, W(b) - f(a)\, W(a)\right]$$
$$-\sum_{k=1}^{n}\left[f(t_k) - f(t_{k-1})\right] W(k).$$

Stochastic Integration

Now, we take appropriate limits on both sides, we see that

$$\begin{aligned}\int_a^b f(t)\,dW(t) &= [f(b)\,W(b) - f(a)\,W(a)] - \int_a^b W(t)\,df(t) \\ &= [f(b)\,W(b) - f(a)\,W(a)] - \int_a^b W(t)\,f'(t)\,dt.\end{aligned}$$

Take the expectation of both sides,

$$\begin{aligned}E(Y_t) &= \underbrace{[f(b)\,E(W(b)) - f(a)\,E(W(a))]}_{0} \\ &\quad - \underbrace{\int_a^b E(W(t))\,df(t)}_{0} = 0.\end{aligned}$$

No matter what the form of the function is, the mean of the "Ito integral" is always zero.

Stochastic Integration

Its variance can also be computed as follows.

$$Var\left[\sum_{k=1}^{n} f\left(t_{k-1}\right)\left(W\left(k\right)-W\left(k-1\right)\right)\right]$$

$$\begin{aligned} &= \sum_{k=1}^{n} f^{2}\left(t_{k-1}\right) Var\left(W\left(k\right)-W\left(k-1\right)\right) \\ &= \sigma^{2}\sum_{k=1}^{n} f^{2}\left(t_{k-1}\right)\left(t_{k}-t_{k-1}\right), \end{aligned}$$

$$\begin{aligned} Var\left(Y_{t}\right) &= \lim_{\substack{n\to\infty \\ \max(t_k-t_{k-1})\to 0}} \sigma^{2}\sum_{k=1}^{n}\left[\begin{array}{c} f^{2}\left(t_{k-1}\right) \\ \left(t_{k}-t_{k-1}\right) \end{array}\right] \\ &= \sigma^{2}\int_{a}^{b} f^{2}\left(t\right) dt. \end{aligned}$$

Stochastic Integration - Example

Let the process $\{Y_t,\ t \geq 0\}$ be defined by

$$Y_t = \int_0^t e^{\alpha(t-u)} dW(u), \qquad \text{for } t \geq 0,$$

where α is some constant parameter. Find the mean and variance of this process.

Solution: The process has zero mean (Ito integral).

Variance is

$$\begin{aligned} Var[Y_t] &= \sigma^2 \times \int_0^t \left[e^{\alpha(t-u)}\right]^2 du \\ &= \sigma^2 \times \int_0^t e^{2\alpha(t-u)} du \\ &= \frac{\sigma^2}{2\alpha}\left(e^{2\alpha t} - 1\right). \end{aligned}$$

Cameron-Martin-Girsanov Theorem - Change of Measure

If $W(t)$ is a Brownian motion under measure P and $\lambda(t)$ is a (previsible) process satisfying $E_P \exp\left(\frac{1}{2}\int_0^T \lambda(t)^2\,dt < \infty\right)$, the Novikov condition, then there exists a measure Q such that

- Q is equivalent to P
- $\frac{dQ}{dP} = \exp\left(-\int_0^T \lambda(t)\,dW(t) - \frac{1}{2}\int_0^T \lambda(t)^2\,dt\right)$
- $\widetilde{W}(t) = W(t) + \int_0^t \lambda(t)\,dt$ is a Q Brownian motion

This can be applied to change measure as follows

$$\begin{aligned} E_Q[X] &= E_P\left[\frac{dQ}{dP}X\right] \\ &= E_P\left[\exp\left(-\int_0^T \lambda(t)\,dW(t) - \frac{1}{2}\int_0^T \lambda(t)^2\,dt\right)X\right] \end{aligned}$$

Fundamental Theorem of Asset Pricing

Define the money market account (or cash account) as

$$B(t) = B(0) \exp\left(\int_0^t r(s)\, ds\right)$$

The Fundamental Theorem of Asset Pricing as applied to Bond Prices states that

- Bond prices are arbitrage free if and only if there exists a measure Q, equivalent to P, under which, for each T, the discounted bond price process $P(t, T) / B(t)$ is a martingale for all $t : 0 < t < T$.
- The market is complete if and only if Q is the unique measure for which $P(t, T) / B(t)$ are martingales.

Q is referred to as the risk-neutral measure with the cash account as numeraire since the expected return on a bond under this measure is the risk-free rate.

Martingale Approach for Bond Pricing

Assume SDEs for the risk free short rate and the bond price are (note change in notation from earlier)

$$\begin{aligned} dr(t) &= a(t)\,dt + b(t)\,dW(t) \\ dP(t,T) &= P(t,T)\left[m(t,T)\,dt + s(t,T)\,dW(t)\right] \end{aligned}$$

Money market account

$$dB(t) = r(t)\,B(t)\,dt$$

Market price of risk

$$\lambda(t) = \frac{m(t,T) - r(t)}{s(t,T)}$$

Discounted bond price process

$$Z(t,T) = \frac{P(t,T)}{B(t)} = P(t,T)\exp\left(-\int_0^t r(s)\,ds\right)$$

Martingale Approach for Bond Pricing

We have (Ito product rule)

$$dZ(t,T) = \frac{1}{B(t)} dP(t,T) + P(t,T)\, d\left(\frac{1}{B(t)}\right) + d\left\langle \frac{1}{B}, P \right\rangle (t)$$

and using Ito formula

$$d\left(\frac{1}{B(t)}\right) = -\frac{1}{B(t)^2} dB(t) + \frac{1}{2}\frac{2}{B(t)^3} d\langle B\rangle (t) = -\frac{r(t)\, dt}{B(t)}$$

Hence

$$\begin{aligned} dZ(t,T) &= \frac{P(t,T)\left(m(t,T)\, dt + s(t,T)\, dW(t)\right)}{B(t)} - \frac{r(t)\, P(t,T)\, dt}{B(t)} \\ &= Z(t,T)\left[\left(m(t,T) - r(t)\right) dt + s(t,T)\, dW(t)\right] \\ &= Z(t,T)\left[\begin{array}{c} \left(m(t,T) - r(t) - \lambda(t)\, s(t,T)\right) dt \\ +s(t,T)\left(dW(t) + \lambda(t)\right) \end{array}\right] \\ &= Z(t,T)\, s(t,T)\, d\widetilde{W}(t) \end{aligned}$$

and $Z(t,T)$ is a martingale under measure Q.

Martingale Approach for Bond Pricing

Discounted bond price $\frac{P(t,T)}{B(t)}$ is a martingale under risk neutral measure

$$\begin{aligned} \frac{P(t,T)}{B(t)} &= E_Q\left[\frac{P(T,T)}{B(T)}|\mathcal{F}_s\right] \\ P(t,T) &= E_Q\left[\frac{B(t)}{B(T)}|\mathcal{F}_s\right] \end{aligned}$$

and for $0 < S < T$

$$P(t,S) = E_Q\left[\exp\left(-\int_t^S r(u)\,du\right)|\mathcal{F}_s\right]$$

Discussion - Bond Pricing

- Use of PDE approach versus Martingale Approach.
- Difference between P and Q measure for interest rate dynamics versus bond prices.
- Importance of market price of risk.
- Single factor versus multiple factor models.

Main Continuous Time Single Factor Short Rate Models

Main continuous time single factor short rate models

- One factor models with constant interest rate volatility (affine models)
  - Vasicek (1977) $dr(t) = \alpha(\mu - r(t))\,dt + \sigma d\widetilde{W}(t)$
  - Ho and Lee (1986) $dr(t) = \theta(t)\,dt + \sigma d\widetilde{W}(t)$
  - Extended Vasicek or Hull and White Model (1990, 1993) $dr(t) = \alpha(\mu(t) - r(t))\,dt + \sigma d\widetilde{W}(t)$
- One factor models with rate-dependent interest rate volatility;
  - Cox, Ingersoll and Ross (1985) - $dr(t) = \alpha(\mu - r(t))\,dt + \sigma\sqrt{r(t)}d\widetilde{W}(t)$
  - Black, Derman and Toy (1990) - log-normal short rate - $d\log r(t) = \left(\theta(t) - \frac{\sigma'(t)}{\sigma(t)}\log r(t)\right)dt + \sigma(t)\,d\widetilde{W}(t)$ or with constant volatility $d\log r(t) = \theta(t)\,dt + \sigma d\widetilde{W}(t)$
  - Black and Karasinski (1991) - log-normal short rate - $d\log r(t) = \alpha(t)(\log\mu(t) - \log r(t))\,dt + \sigma(t)\,d\widetilde{W}(t)$

Vasicek model

Risk neutral dynamics

$$dr(t) = \alpha(\mu - r(t))\,dt + \sigma d\widetilde{W}(t)$$

- mean reverting process
- μ long term mean risk free rate under risk neutral measure, α rate $r(t)$ reverts to long term mean
- σ local volatility of short term rate
- $r(t+s)$ given $r(t)$ is normally distributed under Q with mean

$$\mu + (r(t) - \mu)\,e^{-\alpha s}$$

and variance

$$\frac{\sigma^2\left[1 - e^{-2\alpha s}\right]}{2\alpha}$$

Vasicek model

Bond prices determined by:

- using PDE for bond price of zero coupon bonds, guess form of solution and derive form of bond price by solving PDE, or
- using martingale approach and the dynamics for $r(t)$ to evaluate dscounted expected value under risk neutral measure (often solved using Laplace transforms).

Model does not give a perfect fit to an initial yield curve.

Vasicek model

Bond price takes the form

$$P(t, T) = \exp\left[A(t, T) - B(t, T)\, r(t)\right]$$

where

$$B(t, T) = \frac{1 - e^{-\alpha(T-t)}}{\alpha}$$

$$A(t, T) = (B(t, T) - (T - t))\left(\mu - \frac{\sigma^2}{2\alpha^2}\right) - \frac{\sigma^2}{4\alpha} B(t, T)^2$$

Use this to construct the zero coupon bond yield curve and derive spot rates and forward rates.

Ho-Lee model

Gives a perfect fit to an initial yield curve.
Risk neutral dynamics

$$dr(t) = \theta(t)\,dt + \sigma d\widetilde{W}(t)$$

- time varying drift allowing to fit initial yield curve
- short rate normally distributed.

Bond price for zero coupon bonds

$$P(t, T) = \exp\left[A(t, T) - B(t, T)\,r(t)\right]$$

where

$$B(t, T) = (T - t)$$

$$A(t, T) = \frac{\sigma^2}{6}(T - t)^3 - \int_t^T \theta(s)(T - s)\,ds$$

Ho-Lee model

- Setting

$$\theta(t) = \frac{\partial}{\partial T} f(0, T) + \sigma^2 T$$

where

$$f(0, T) = -\frac{\partial}{\partial T} \log P(0, T)$$

- We have

$$\begin{aligned} r(t) &= r(0) + \int_0^t \theta(s)\, ds + \sigma \widetilde{W}(t) \\ &= f(0, t) + \frac{\sigma^2 t^2}{2} + \sigma \widetilde{W}(t) \end{aligned}$$

which allows a perfect fit to the initial yield curve.

- We then have

$$f(t, T) = -\frac{\partial}{\partial T} \log P(t, T) = r(t) + f(0, T) - f(0, t) + \sigma^2 t (T - t)$$

Extended Vasicek or Hull and White Model

Risk neutral dynamics

$$dr(t) = \alpha(\mu(t) - r(t))\,dt + \sigma d\widetilde{W}(t)$$

Bond price for zero coupon bonds, matching initial term structure,

$$\mu(t) = \frac{1}{\alpha}\frac{\partial}{\partial t}f(0,t) + f(0,t) + \frac{\sigma^2}{2\alpha^2}\left(1 - e^{-2\alpha t}\right)$$

$$P(t,T) = \exp\left[A(t,T) - B(t,T)\,r(t)\right]$$

where

$$B(t,T) = \frac{1 - e^{-\alpha(T-t)}}{\alpha}$$

$$\begin{aligned} A(t,T) &= \log\frac{P(0,T)}{P(0,t)} + B(t,T)\,f(0,t) \\ &\quad - \frac{\sigma^2}{4\alpha^3}\left(1 - e^{-\alpha(T-t)}\right)^2\left(1 - e^{-2\alpha t}\right) \end{aligned}$$

Extended Vasicek or Hull and White Model

We also have

$$\begin{aligned} r(t) &= e^{-\alpha t} r(0) + \alpha \int_0^t e^{-\alpha(t-s)} \mu(s)\, ds + \sigma \int_0^t e^{-\alpha(t-s)} d\widetilde{W}(s) \\ &= f(0,t) + \frac{\sigma^2}{2\alpha^2}\left(1 - e^{-\alpha t}\right)^2 + \sigma \int_0^t e^{-\alpha(t-s)} d\widetilde{W}(s) \end{aligned}$$

Special cases

- $\mu(t) = \mu$, a constant, gives Vasicek
- if $\alpha \to 0$ and $\alpha\mu(t) \to \theta(t)$ as $\alpha \to 0$ then gives Ho-Lee.

Cox-Ingersoll-Ross Model

Instantaneous interest rate $r(t)$ dynamics under the risk neutral $\mathbb{Q}$ measure given by:

$$dr(t) = \alpha(\mu - r(t))dt + \sigma\sqrt{r(t)}d\widetilde{W}(s)$$

where $\alpha > 0$ is the speed of mean reversion of $r(t)$, $\mu > 0$ is the long-run mean of $r(t)$, $\sigma\sqrt{r(t)} \geq 0$ is the volatility of the short rate process. Require $2\alpha\mu \geq \sigma^2$ to ensure the process is positive.

Cox-Ingersoll-Ross Model

Zero coupon bond price is given by:

$$P(t,T) = \mathbb{E}_t^Q[e^{-\int_t^T r(u)du}] = \exp\left[A\left(t,T\right) - B\left(t,T\right)r\left(t\right)\right]$$

where $A(t,T)$ and $B(t,T)$ are given by:

$$A(t,T) = \frac{2\alpha\mu}{\sigma^2}\log\left(\frac{2\gamma e^{(\gamma+\alpha)(T-t)/2}}{(\gamma+\alpha)(e^{\gamma(T-t)}-1)+2\gamma}\right)$$

$$B(t,T) = \frac{2(e^{\gamma(T-t)}-1)}{(\gamma+\alpha)(e^{\gamma(T-t)}-1)+2\gamma}$$

with

$$\gamma = \sqrt{\alpha^2+2\sigma^2}$$

with boundary conditions $A(T,T) = 0$ and $B(T,T) = 0$.

Other Short Rate Models

Many other models with differing assumptions, for example:

- Brennan and Schwartz (1979) - two-factor model with short rate and long term rate on a consol bond.
- Longstaff and Schwartz (1992) - two-factor version of CIR with short rate and instantaneous volatility of short rate as factors

Affine Term Structure Models - Short Rate Models

Affine short rate term structure models have bond prices of affine form in the short rate

$$P(t, T) = \exp\left[A(t, T) - B(t, T)\, r(t)\right]$$

and for bond prices to have this form the risk neutral drift and volatility of the short rate take the form

$$\begin{aligned} m(t, r(t)) &= a(t) + b(t)\, r(t) \\ s(t, r(t)) &= \sqrt{\delta(t) + \gamma(t)\, r(t)} \end{aligned}$$

where $a(t)$, $b(t)$, $\delta(t)$, and $\gamma(t)$ are deterministic functions.

Affine Term Structure Models - Short Rate Models

Consider the risk neutral dynamics for the short rate

$$dr(t) = m(t, r(t))\,dt + s(t, r(t))\,d\widetilde{W}(t)$$

Then an application of Ito's formula gives

$$dP(t,T) = P(t,T)\left[\left(\frac{\partial A}{\partial t} - \frac{\partial A}{\partial t}r(t) - Bm + \frac{1}{2}(Bs)^2\right)dt - Bsd\widetilde{W}(t)\right]$$

Under the risk neutral dynamics all bonds have instantaneous expected returns equal to the short rate so that

$$dP(t,T) = P(t,T)\left[r(t)\,dt + S(t,T,r(t))\,d\widetilde{W}(t)\right]$$

where $S(t,T,r(t))$ is the volatility of the bond price $P(t,T)$.

Affine Term Structure Models - Short Rate Models

Define

$$g(t,r) = \left(\frac{\partial A}{\partial t} - \frac{\partial A}{\partial t} r(t) - Bm(t, r(t)) + \frac{1}{2} \left[Bs(t, r(t)) \right]^2 \right) - r$$

then this must be zero for all t and r for the bond prices to be arbitrage-free.
Differentiating twice with respect to r gives

$$\frac{\partial^2 g(t,r)}{\partial r^2} = -B \frac{\partial^2 m(t,r)}{\partial r^2} + \frac{1}{2} B^2 \frac{\partial^2 s(t, r(t))^2}{\partial r^2} = 0$$

and for this to hold

$$\frac{\partial^2 m(t,r)}{\partial r^2} = 0 \text{ and } \frac{\partial^2 s(t, r(t))^2}{\partial r^2} = 0$$

which gives the affine form for the short rate dynamics.

Affine Term Structure Models - Short Rate Models

Special cases include a number of the spot rate models already covered.
Vasicek

$$\begin{aligned} a &= \alpha\mu, b = -\alpha, \delta = \sigma^2, \gamma = 0 \\ dr(t) &= \alpha(\mu - r(t))\,dt + \sigma d\widetilde{W}(t) \end{aligned}$$

Cox, Ingersoll and Ross

$$\begin{aligned} a &= \alpha\mu, b = -\alpha, \delta = 0, \gamma = \sigma^2 \\ dr(t) &= \alpha(\mu - r(t))\,dt + \sigma\sqrt{r(t)}d\widetilde{W}(t) \end{aligned}$$

Discussion - Short Rate Models

- Some models were developed as equilibrium models (Vasicek and CIR) and others as arbitrage-free or relative valuation models.
- Models with constant parameters, drift and volatility, do not provide an exact fit to the current yield curve, so models with time varying drift and volatility were developed to do this.
- Model fitting requires both the current yield curve and a volatility structure for time varying volatility models.
- Gaussian models have geneally small chance of negative interest rates.
- Gaussian models have closed form expressions for bond prices (and options on bonds) because of log-normal distrubution of bond prices.

Discussion - Short Rate Models

- Log-normal short rate models require numerical implementation to compute bond prices usually with a lattice.
- Many models developed as discrete time lattice models have continuous time limits (Ho-Lee, Black-Derman-Toy).
- Many are single factor models, with perfect correlation of instantaneous bond returns across maturities, but can be extended to multiple factors.
- Matching volatility and correlations between different maturity bond returns and option prices requires time varying volatility and multiple factor models.

Affine Term Structure Models - Multi-Factor Models

More generally affine term structure (ATS) models are arbitrage-free multifactor model of interest rates in which the yield on any risk-free zero-coupon bond is an affine function of a set of unobserved latent factors.

Affine term structure (ATS) models have bond prices of affine form

$$\begin{aligned} P(t,T) &= \exp\left[A(t,T) + \sum_{j=1}^{n} B_j(t,T) X_j(t)\right] \\ &= \exp\left[A(t,T) + B(t,T)' X(t)\right] \end{aligned}$$

where $X(t) = (X_1(t), X_2(t), \ldots, X_n(t))'$ is a vector of state variables and $B(t,T) = (B_1(t,T), B_2(t,T), \ldots, B_n(t,T))'$.

We will only consider time homogeneous models where $A(t,T)$ and $B(t,T)$ are functions of $(T-t)$ only and the state variables $X(t)$ are time homogeneous.

Affine Term Structure Models - Multi-Factor Models

If bond prices have the form

$$P(t, t+\tau) = \exp\left[A(\tau) + B(\tau)' X(t)\right]$$

then $X(t)$ must have SDE's $dX(t) = (\alpha + \beta) X(t) + SD(X(t)) d\widetilde{W}(t)$ where $\widetilde{W}(t)$ is an n-dimensional Brownian motion under Q, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)'$ is a constant vector, $\beta = \left(\beta_{ij}\right)_{i.j=1}^{n}$ is a constant matrix, $S = (\sigma_{ij})_{i.j=1}^{n}$ is a constant matrix and $D(X(t))$ a diagonal matrix

$$\begin{bmatrix} \sqrt{\gamma_1' X(t) + \delta_1} & 0 & \cdots & \cdots & 0 \\ 0 & \sqrt{\gamma_2' X(t) + \delta_2} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \cdots & \vdots \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \sqrt{\gamma_n' X(t) + \delta_n} \end{bmatrix}$$

where $\delta_1, \delta_2, \ldots, \delta_n$ are constants and $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ is a constant

Affine Term Structure Models - Gaussian Model

As an example, a 3-factor Gaussian ATS model has

$$r(t) = \mu + \mathbf{1}' X(t)$$

where $\mathbf{1} = (1, 1, 1)'$ and $X(t) = (X_1(t), X_2(t), X_3(t))$ follows a zero-mean Ornstein-Uhlenbeck process under the real world probability measure

$$dX(t) = -KX(t)\,dt + \Sigma dW(t)$$

with K is lower triangular and Σ is diagonal, both 3×3 matrices. Market prices of risk for the latent factors are given by

$$\lambda(t) = \lambda(0) + \Lambda X(t)$$

where $\lambda(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t))'$ and Λ a 3×3 matrix.

Affine Term Structure Models - Gaussian Model

Bond prices for a $\tau-$maturity bond given by

$$P(t,\tau) = E_Q\left[\exp\left(-\int_t^{t+\tau} r(u)\,du\right)|\mathcal{F}_s\right]$$

and the risk neutral dynamics are gven by

$$\begin{aligned} dX(t) &= (-K-\lambda(t)\Sigma)X(t)\,dt + \Sigma d\widetilde{W}(t) \\ &= -((K+\Sigma\Lambda)X(t)+\Sigma\lambda(0))\,dt + \Sigma d\widetilde{W}(t) \end{aligned}$$

Zero coupon bond prices can be written as

$$P(t,\tau) = \exp\left[A(\tau)+B(\tau)X(t)\right]$$

where expressions for $A(\tau)$ and $B(\tau)$ are factor loadings that are functions of the parameters of the model solved from a set of differential equations.

Dynamic Nelson-Siegel Model

The Dynamic Nelson Siegel (DNS) model is effectively a 3-factor Gaussian model, although not arbitrage-free. Recognising the 3 factors as level, slope and curvature of the yield curve, we can write it as

$$y_t(\tau) = l_t + s_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + c_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$$

and can be written in state-space form as a measurement equation

$$y_t = \Lambda f_t + \varepsilon_t$$

where

$$y_t = \begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix}, \quad f_t = \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_N) \end{pmatrix}$$

Dynamic Nelson-Siegel Model

The parameter matrix is

$$\Lambda = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix}$$

where we have observable yields at times $t = 1, \ldots, T$, and N maturities for zero coupon bond yields at each time.
The l_t, s_t, and c_t are common factors with dynamics given by the transition equation.

Dynamic Nelson-Siegel Model

The transition equation assumes a first-order vector autoregression

$$(f_t - \mu) = A\,(f_{t-1} - \mu) + \eta_t$$

where

$$\mu = \begin{pmatrix} \mu^l \\ \mu^s \\ \mu^c \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \eta_t^l \\ \eta_t^s \\ \eta_t^c \end{pmatrix}$$

Dynamic Nelson-Siegel Model

Finally the state space formulation requires assumptions for the covariance structure of the measurement and transition errors. Here we assume white noise and orthogonal errors so that

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim WN \begin{pmatrix} 0 & Q & 0 \\ 0 & 0 & H \end{pmatrix}$$

This formulation allows the use of Kalman filter for optimal filtering and smoothing as well as prediction of yield factors and observed yields. Estimation is then by maximum likelihood with the state space representation and the Kalman filter.

Dynamic Nelson-Siegel Model

Kalman filter provides one-step-ahead prediction errors for which the Gaussian pseudo likelihood can be evaluated for any set of parameters.
The parameter configuration that maximises the likleihood is found numerically using some form of either gradient based methods or analytic score functions.
Other model estimation procedures are possible including an EM based optimization or a Bayesian approach using Markov-chain Monte Carlo.
There are a large number of parameters so that traditonal gradient based numerical optimization methods may be intractable.
If there are N maturities at each time point to fit then the measurement equation requires one-parameter λ, unless a fixed value is used, the transition equation has 12 parameters (3 means and 9 dynamic parameters), the measurement error covariance matrix has $\left(N^2 + N\right)/2$ parameters and the transition disturbance covariance marix has 6 parameters.
So if there were 15 yield maturities to be fitted then there would be 139 parameters.

Arbitrage-Free Nelson-Siegel Model

The Dynamic Nelson-Seigel Model is flexible.
It can fit cross section and time series of yields well.
However it does not impose restrictions to make the model arbitrage-free.
The DNS model can be made arbitrage-free and formulated as an affine term structure model.
The arbitrage free Nelson Seigel (AFNS) maintains the DNS factor-loading structure and is an affine arbitrage-free term structure model.

Arbitrage-Free Nelson-Siegel Model

Consider a three-factor affine model with $X_t = \left(X_t^1, X_t^2, X_t^3\right)$ then we require a yield function of the form

$$\begin{aligned} y(t,T) &= -\frac{A(t,T)}{T-t} + X_t^1 + \left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}\right) X_t^2 \\ &\quad + \left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)}\right) X_t^3 \end{aligned}$$

so the factor loadings in the affine model need to be

$$\begin{aligned} B^1(t,T) &= -(T-t), \quad B^2(t,T) = -\frac{1-e^{-\lambda(T-t)}}{\lambda}, \\ \text{and } B^3(t,T) &= -\frac{1-e^{-\lambda(T-t)}}{\lambda} + (T-t)\, e^{-\lambda(T-t)} \end{aligned}$$

Arbitrage-Free Nelson-Siegel Model

The required model has

$$r(t) = X_t^1 + X_t^2$$

and risk neutral dynamics for the state variables

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\left(\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right) dt \right]$$
$$+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} d\widetilde{W}_t^1 \\ d\widetilde{W}_t^2 \\ d\widetilde{W}_t^3 \end{pmatrix}$$

Arbitrage-Free Nelson-Siegel Model

Zero coupon bond prices are

$$P(t,T) = \exp\left(C(t,T) + B^1(t,T)X_t^1 + B^2(t,T)X_t^2 + B^3(t,T)X_t^3\right)$$

and the ODEs for the factor loadings are

$$\begin{pmatrix} \frac{dB^1(t,T)}{dt} \\ \frac{dB^2(t,T)}{dt} \\ \frac{dB^3(t,T)}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \begin{pmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{pmatrix}$$

and

$$\frac{dC(t,T)}{dt} = -B(t,T)'\theta^Q - \frac{1}{2}\sum_{j=1}^{3}\left(\Sigma' B(t,T) B(t,T)'' \Sigma\right)_{j.,j}$$

with boundary conditions
$C(T,T) = B^1(T,T) = B^2(T,T) = B^3(T,T) = 0.$

Arbitrage-Free Nelson-Siegel Model

The relationship between the real-world dynamics and the risk-neutral dynamics is

$$d\widetilde{W}_t = dW_t + \Gamma_t dt$$

and use the essentially affine risk premium specification in which Γ_t is affine in the state variables so that

$$\Gamma_t = \begin{pmatrix} \gamma_1^0 \\ \gamma_2^0 \\ \gamma_3^0 \end{pmatrix} + \begin{pmatrix} \gamma_{11}^1 & \gamma_{12}^1 & \gamma_{13}^1 \\ \gamma_{21}^1 & \gamma_{22}^1 & \gamma_{23}^1 \\ \gamma_{31}^1 & \gamma_{32}^1 & \gamma_{33}^1 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}$$

The P measure dynamics are

$$dX_t = K^P \left[\theta^P - X_t\right] dt + \Sigma dW_t$$

and we are free to choose the mean vector θ^P and the mean reversion matrix K^P under the P measure and preserve the Q dynamics.

Arbitrage-Free Nelson-Siegel Model

The independent-factor AFNS model has independent state variables under the P measure

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ 0 & \kappa_{22}^P & 0 \\ 0 & 0 & \kappa_{33}^P \end{pmatrix} \left[\left(\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right) dt \right]$$
$$+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \\ dW_t^3 \end{pmatrix}$$

Arbitrage-Free Nelson-Siegel Model

The correlated-factor AFNS model has independent state variables under the P measure

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \left[\left(\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right) dt \right]$$
$$+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \\ dW_t^3 \end{pmatrix}$$

Arbitrage-Free Nelson-Siegel Model

The measurement equation is

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}$$

$$- \begin{pmatrix} \frac{C(\tau_1)}{\tau_1} \\ \frac{C(\tau_2)}{\tau_2} \\ \vdots \\ \frac{C(\tau_N)}{\tau_N} \end{pmatrix} + \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_N) \end{pmatrix}$$

Arbitrage-Free Nelson-Siegel Model

Estimation of the AFNS uses a Kalman filter maximum likelihood approach.

The model needs to be converted from continuous to discrete time.

The conditional mean vector is

$$E^P\left[X_T\left|\mathcal{F}_s\right.\right] = \left(I - \exp\left(-K^P\left(T-t\right)\right)\right)\theta^P + \exp\left(-K^P\left(T-t\right)\right)X_t$$

and the conditional variance is

$$V^P\left[X_T\left|\mathcal{F}_s\right.\right] = \int_0^{(T-t)} e^{-K^P s}\Sigma\Sigma' - \left(K^P\right)' s ds$$

The state transition equation is

$$X_t = \left(I - \exp\left(-K^P\Delta t\right)\right)\theta^P + \exp\left(-K^P\Delta t\right)X_{t-1} + \eta_t$$

where Δt is the time between observations.

The variance of η_t is

$$Q = \int_0^{\Delta t} e^{-K^P s}\Sigma\Sigma' - \left(K^P\right)' s ds$$

Arbitrage-Free Nelson-Siegel Model

The AFNS measurement equation is

$$y_t = BX_t + C + \varepsilon_t$$

and the stochastic error term is

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N \begin{pmatrix} 0 & Q & 0 \\ 0 & 0 & H \end{pmatrix}$$

For the Kalman filter we start with unconditional mean and variance of state variables under the P measure

$$X_0 = \theta^P, \text{ and } \Sigma_0 = \int_0^\infty e^{-K^P s} \Sigma \Sigma' - \left(K^P\right)' s ds$$

Arbitrage-Free Nelson-Siegel Model

Denote model parameters by ψ and the information at time t by $Y_t = (y_1, y_2, \ldots, y_T)$, then at time $t-1$ assume we have the state update X_{t-1} and the mean square error matrix Σ_{t-1}.
The prediction step is

$$\begin{aligned} X_{t|t-1} &= E^P\left[X_t | Y_{t-1}\right] = \Phi_t^{X,0}(\psi) + \Phi_t^{X,1}(\psi) X_{t-1} \\ \Sigma_{t|t-1} &= \Phi_t^{X,1}(\psi) \Sigma_{t-1} \Phi_t^{X,1}(\psi)' + Q_t(\psi) \end{aligned}$$

where

$$\begin{aligned} \Phi_t^{X,0}(\psi) &= \left(I - \exp\left(-K^P \Delta t\right)\right) \theta^P \\ \Phi_t^{X,1}(\psi) &= \exp\left(-K^P \Delta t\right) \\ Q_t(\psi) &= \int_0^{\Delta t} e^{-K^P s} \Sigma \Sigma' - \left(K^P\right)' s \, ds \end{aligned}$$

Arbitrage-Free Nelson-Siegel Model

The time t update step improves $X_{t|t-1}$ using the additional information contained in Y_t.

$$\begin{aligned} X_t &= E\left[X_t | Y_t\right] = X_{t|t-1} + \Sigma_{t|t-1} B\left(\psi\right)' F_t^{-1} \nu_t \\ \Sigma_t &= \Sigma_{t|t-1} - \Sigma_{t|t-1} B\left(\psi\right)' F_t^{-1} B\left(\psi\right) \Sigma_{t|t-1} \end{aligned}$$

where

$$\begin{aligned} \nu_t &= y_t - E\left[y_t | Y_{t-1}\right] = y_t - B\left(\psi\right) X_{t|t-1} - C\left(\psi\right) \\ F_t &= cov\left(\nu_t\right) = B\left(\psi\right) \Sigma_{t|t-1} B\left(\psi\right)' + H\left(\psi\right) \\ H\left(\psi\right) &= diag\left(\sigma_\varepsilon^2\left(\tau_1\right), \ldots, \sigma_\varepsilon^2\left(\tau_N\right)\right) \end{aligned}$$

Arbitrage-Free Nelson-Siegel Model

A single pass of the Kalman filter allows us to evaluate the Gaussian log likelihood

$$\log l\left(y_1, y_2, \ldots, y_T | \psi\right) = \sum_{t=1}^{N} \left(-\frac{1}{2} N \log\left(2\pi\right) - \frac{1}{2} \log\left|F_t\right| - \frac{1}{2} \nu_t^{'} F_t^{-1} \nu_t \right)$$

where N is the number of observed yields.
This is then maximised with respect to ψ.

Discussion - Affine Arbitrage-Free and Nelson-Siegel Model

- Gaussian versus other affine models.
- Fitting current yield curve.
- Number of factors.
- Benefits of closed form solutions.

HJM (Heath, Jarrow, Morton)

HJM Framework

- General framework for arbitrage-free models.
- Starts with initial forward-rate curve as input - guaranteed to fit initial term structure.
- Imposes no-arbitrage conditions to derive relationship between drift and volatility of forward curve.

HJM Framework

- Consider a one factor model. We will not cover technical conditions in details.
- Forward curve $f(t, T)$ $0 \leq t < T$ with initial forward curve $f(0, T)$
- For fixed maturity T dynamics of $f(t, T)$ given by

$$df(t, T) = \alpha(t, T)\, dt + \sigma(t, T)\, dW(t)$$

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

- For the risk free interest rate

$$\begin{aligned} r(T) &= \lim_{t \to T^-} f(t, T) = f(0, T) + \int_0^T \alpha(s, T)\, ds \\ &\quad + \int_0^T \sigma(s, T)\, dW(s) \end{aligned}$$

HJM Framework

- Cash account

$$
\begin{aligned}
dB(t) &= r(t)\,B(t)\,dt \\
B(t) &= B(0)\exp\left[\int_0^t r(u)\,du\right] \\
&= B(0)\exp\left[\begin{array}{c}\int_0^t f(0,u)\,du + \int_0^t \int_s^t \alpha(s,u)\,duds \\ + \int_0^t \left(\int_s^t \sigma(s,u)\,du\right) dW(s)\end{array}\right]
\end{aligned}
$$

- Uses technical condition to change order of integration in third term

$$
\int_0^t \left(\int_0^u \sigma(s,u)\,dW(s)\right) du = \int_0^t \left(\int_s^t \sigma(s,u)\,du\right) dW(s)
$$

HJM Framework

- Zero coupon bond market prices

$$
\begin{aligned}
P(t,T) &= \exp\left[-\int_t^T f(t,u)\,du\right] \\
&= \exp\left[\begin{array}{c} -\int_0^t f(0,u)\,du - \int_0^t \int_t^T \alpha(s,u)\,duds \\ -\int_0^t \left(\int_t^T \sigma(s,u)\,du\right) dW(s) \end{array}\right]
\end{aligned}
$$

HJM, Framework

- Discounted bond price

$$
\begin{aligned}
Z(t,T) &= \frac{P(t,T)}{B(t)} \\
&= \exp\left[\begin{array}{c} -\int_0^T f(0,u)\,du - \int_0^t \int_s^T \alpha(s,u)\,duds \\ +\int_0^t S(s,T)\,dW(s) \end{array} \right]
\end{aligned}
$$

where

$$
S(s,T) = -\left(\int_s^T \sigma(s,u)\,du \right)
$$

and $S(t,T)$ is interpreted as the volatility of $P(t,T)$ - zero coupon bond price volatility for maturity T.

HJM Framework

- Using Ito formula we have

$$dZ(t,T) = Z(t,T)\left[\begin{array}{c}\left(\frac{1}{2}S^2(t,T) - \int_t^T \alpha(t,u)\,du\right)dt \\ +S(t,T)\,dW(t)\end{array}\right]$$

- Change of measure to make the discounted bond price a martingale

$$\gamma(t) = \frac{1}{2}S(t,T) - \frac{1}{S(t,T)}\int_t^T \alpha(t,u)\,du$$

HJM Framework

- Rearranging gives the no arbitrage condition

$$\int_t^T \alpha(t,u)\,du = \frac{1}{2}^2 S(t,T) - \gamma(t)\,S(t,T)$$

- Differentiate with respect to T

$$\alpha(t,T) = \sigma(t,T)\left(\gamma(t) - S(t,T)\right)$$

- The original model was

$$\begin{aligned} df(t,T) &= \alpha(t,T)\,dt + \sigma(t,T)\,dW(t) \\ &= \alpha(t,T)\,dt + \sigma(t,T)\left(d\widetilde{W}(t) - \gamma(t)\,dt\right) \\ &= -\sigma(t,T)\,S(t,T)\,dt + \sigma(t,T)\,d\widetilde{W}(t) \end{aligned}$$

- Forward rate process is completely specified by the volatility functions $\sigma(s,t)$.

HJM Framework

- We have for HJM (single factor)
  $r(t) = f(0,t) - \int_0^t \sigma(s,t) S(s,t)\, ds + \int_0^t \sigma(s,t)\, d\widetilde{W}(s)$.
- In general $r(t)$ is not Markov but special cases for volatility specifications correspond to cases considered previously.
- Set

$$\sigma(s,t) = \sigma$$

  for all s and t then we get the Ho-Lee model

$$r(t) = f(0,t) - \frac{1}{2}\sigma^2 t^2 + \sigma \widetilde{W}(t)$$

HJM Framework

- Set

$$\sigma(s,t) = \sigma e^{-\alpha(t-s)}$$

- Then

$$\begin{aligned} S(s,t) &= -\frac{\sigma}{\alpha}\left(1-e^{-\alpha(t-s)}\right) \\ -\int_0^t \sigma(s,t)\, S(s,t)\, ds &= \frac{\sigma^2}{2\alpha^2}\left(1-e^{-\alpha t}\right)^2 \end{aligned}$$

and

$$r(t) = f(0,t) + \frac{\sigma^2}{2\alpha^2}\left(1-e^{-\alpha t}\right)^2 + \sigma\int_0^t e^{-\alpha(t-s)}\widetilde{W}(t)$$

which is a form of Hull-White model.

HJM Framework

- For log-normal forward rates set

$$\sigma(t, T) = \sigma f(t, T)$$

- Under this volatility structure we have

$$df(t, T) = \sigma^2 f(t, T) \int_t^T f(t, s)\, ds dt + \sigma f(t, T)\, d\widetilde{W}(t)$$

- Here the drift grows as the square of the forward rate and the forward rate explodes in finite time. This is a characteristic of all log-normal models of instantaneous forward rates (but not discrete froward rates).

Market Models

LMM (Libor Market Models),
BGM (Brace, Gatarek, Musiela),
SABR (stochastic alpha, beta, rho)

Market Models

- Models aim to fit market pricing formulae such as Black's formulae for caps and swaptions.
- They generally use discrete time interest rates such as LIBOR (market observed rates).
- Black model is used for pricing interest rate options with an assumption of log-normality of interest rates or log-normality of bond or swap prices.
- Pricing models often involve approximations and require changes of measure.
- We don't cover details here where our focus is on term structures of interest rates but pricing caps, swaptions is important in practice.

BGM (Brace, Gatarek, Musiela)

- BGM derive processes followed by market quoted rates in HJM framework.
- Uses forward LIBOR rates assumed to evolve as lognormal.
- Consistent with Black's model for pricing interst rate caps.
- Each forward rate has time dependent volatility ands correlations with other forward rate.
- Standard LMM model does not produce market observed caplet volatility smile/skew.
- Local volatility or stchastic volatility models are used to do this.

BGM (Brace, Gatarek, Musiela)

- Select tenor τ which will be required to be log-normal (only one tenor at a time can be log-normal)
- We have
  - $f(t,x)$ the instantaneous forward rate at time t for time $t+x$
  - $L(t,x)$ the market quoted forward rate at time t for time $t+x$ of tenor τ
- We have

$$1+\tau L(t,x)=\exp\left(\int_x^{x+\tau} f(t,u)\,du\right)$$

- Require $L(t,x)$ to have a log-normal process

$$dL(t,x)=\mu(t,x)\,dt+\gamma(t,x)\,L(t,x)\,dW(t)$$

BGM (Brace, Gatarek, Musiela)

- BGM show that $\mu(t,x)$ must have the form

$$\frac{\partial}{\partial x} L(t,x) + L(t,x)\,\gamma(t,x)\,\sigma(t,x) + \frac{\tau L^2(t,x)}{1+\tau L(t,x)} \left|\gamma(t,x)\right|^2$$

and $\sigma(t,x)$ related to $\gamma(t,x)$ by

$$\sigma(t,x) = \begin{array}{ll} 0, & 0 \le x \le \tau \\ \sum_{k=1}^{\lfloor \frac{x}{\tau} \rfloor} \frac{\tau L(t,x-k\tau)}{1+\tau L(t,x-k\tau)} \gamma(t,x-k\tau) & \tau \le x \end{array}$$

SABR model

- This is a stochastic volatility model which attempts to capture the volatility smile.
- Stands for "stochastic alpha, beta, rho".
- Details are relevant for interest rate derivative trading, maybe less relevant for most actuarial applications.
- We only introduce the concept here.

SABR model

- Model for forward rates

$$\begin{aligned} df(t,T) &= \alpha(t) f(t,T)^{\beta} dW_1(t) \\ d\alpha(t) &= \upsilon\alpha(t) dW_2(t) \\ E\left[dW_1(t) dW_2(t)\right] &= \rho dt \end{aligned}$$

- Different models $\beta = 0$ produces a stochastic Gaussian model, $\beta = 1$ produces a stochastic log-normal model and $\beta = \frac{1}{2}$ produces stochastic CIR model.

Discussion - Market Models

- Fitting models used for pricing caps and swaptions - market models (Black model).
- Emphasis on volatility and correlations.
- Ability to calibrate models to interest rate options data.
- Stochastic volatility and fitting market volatility (smiles etc).

Discrete time models

Discrete time models and binomial implementation

- For log-normal models there are no closed form bond price expressions for the term structure of bond prices.
- Many term structure models were developed as discrete time models, especially log-normal models.
- We would like to calibrate the model to both current term structure of interest rates and volatility term structure.
- We do this on a binomial lattice for a single factor model.

Discrete time models and binomial implementation

- Work in discrete time steps with time steps of Δt and nodes on a binomial lattice.
- Denote by $P(i)$ the price of a zero coupon bond maturing at time $i\Delta t$ and $\sigma_R(i)$ the volatility of the yield of this bond.
- $u(i)$ is the median interest rate at time $i\Delta t$.
- $\sigma(i)$ is the volatility of the short term interest rate at time $i\Delta t$.
- $r(i,j)$ is the short term interest rate at time $i\Delta t$. at node j applicable for period $[i\Delta t, (i+1)\Delta t]$.
- $d(i,j)$ the time $i\Delta t$, state j value of a discount bond maturing at time $(i+1)\Delta t$ so that $d(i,j) = 1/(1 + r(i,j)\Delta t)$.

Discrete time models and binomial implementation

- Use Arrow-Debreu securities in calibrating the binomial model of interest rates to market data.
- $Q(i,j)$ is the time 0 value of a security pating 1 if node (i,j) is reached and 0 otherwise with $Q(0,0) = 1$.
- The time 0 price of a discount bond maturing at time $(i+1)\Delta t$ is

$$P(i+1) = \sum_j Q(i,j)\, d(i,j)$$

- Note that

$$P(i) - \exp(-R(i)\, i\Delta t)$$

- Calibration uses forward induction and moves UP and DOWN on the lattice to calibrate to volatility.

Discrete time models and binomial implementation

- Arrow-Debreu prices at time $i\Delta t$ are updated using the time $(i-1)\,\Delta t$ prices using

$$\begin{aligned}
Q(i,i) &= \frac{1}{2}Q(i-1,i-1)\,d(i-1,i-1) \\
Q(i,j) &= \frac{1}{2}Q(i-1,j-1)\,d(i-1,j-1) \\
&\quad + \frac{1}{2}Q(i-1,j+1)\,d(i-1,j+1) \\
Q(i,-i) &= \frac{1}{2}Q(i-1,-i+1)\,d(i-1,-i+1)
\end{aligned}$$

- The short term interest rate at node (i,j) is

$$r(i,j) = u(i)\exp\left(\sigma(i)\,j\sqrt{\Delta t}\right)$$

Discrete time models and binomial implementation

- The lattice is constructed starting at time $i = 0$ with single state $j = 0$.
- At each subsequent time i there are $i + 1$ states indexed as $j = -i, -i + 2, \ldots, i - 2, i$ representing net moves to the node.
- From the initial node $(0, 0)$ we can go to $(1, 1)$ - an UP move, or to $(1, -1)$ - a DOWN move (binomial tree).
- At these nodes $P_U(i)$ and $P_D(i)$are the prices of a discount bond with maturity $i\Delta t$ and $R_U(i)$ and $R_D(i)$ are the discount bond yields at these nodes.
- These values are calibrated to be consistent with observed prices $P(i)$ and volatilities $\sigma_R(i)$.

Discrete time models and binomial implementation

- Require

$$\frac{1}{1+r(0,0)\Delta t}\left(\frac{1}{2}P_U(i)+\frac{1}{2}P_D(i)\right)=P(i) \quad i=2,\ldots,N$$

$$\sigma_R(i)\sqrt{\Delta t}=\frac{1}{2}\ln\frac{\ln P_U(i)}{\ln P_D(i)} \quad i=2,\ldots,N$$

- These are solved numericallyusing Newton-Raphson by rewriting the equations as

$$\begin{aligned} P_D(i)^{\exp\left(2\sigma_R(i)\sqrt{\Delta t}\right)}+P_D(i) &= 2P(i)\,1+r(0,0)\Delta t \\ P_U(i) &= P_D(i)^{\exp\left(2\sigma_R(i)\sqrt{\Delta t}\right)} \end{aligned}$$

Discrete time models and binomial implementation

- Need state prices at the UP and DOWN nodes so define these as $Q_U(i,j)$ and $Q_D(i,j)$.
- By definition $Q_U(1,1) = 1$ and $Q_D(1,-1) = 1$.
- Then

$$\begin{aligned} P_U(i+1) &= \sum_j Q_U(i,j)\, d(i,j) \quad i = 1, \ldots, N-1 \\ P_D(i+1) &= \sum_j Q_D(i,j)\, d(i,j) \quad j \in \{-i, -i-2, \ldots, i-2, i\} \end{aligned}$$

Discrete time models and binomial implementation

State prices are then updated

$$
\begin{aligned}
Q_U(i,i) &= \frac{1}{2} Q_U(i-1,i-1)\, d(i-1,i-1) \\
Q_U(i,j) &= \frac{1}{2} Q_U(i-1,j-1)\, d(i-1,j-1) \\
&\quad + \frac{1}{2} Q_U(i-1,j+1)\, d(i-1,j+1) \\
Q_U(i,-i) &= \frac{1}{2} Q_U(i-1,-i+1)\, d(i-1,-i+1) \\
Q_D(i,-i) &= \frac{1}{2} Q_D(i-1,-i+1)\, d(i-1,-i+1) \\
Q_D(i,j) &= \frac{1}{2} Q_D(i-1,j-1)\, d(i-1,j-1) \\
&\quad + \frac{1}{2} Q_D(i-1,j+1)\, d(i-1,j+1) \\
Q_D(i,i-2) &= \frac{1}{2} Q_D(i-1,i-3)\, d(i-1,i-3)
\end{aligned}
$$

Discrete time models and binomial implementation

- Algorithm for calibration is then as follows.
- Set initial values

$$
\begin{aligned}
r(0,0) &= u(0) = \frac{e^{(R(1)\Delta t)} - 1}{\Delta t} \\
Q_U(1,1) &= 1 \text{ and } Q_D(1,-1) = 1 \\
\sigma(0) &= \sigma_R(1) \\
d(0,0) &= \frac{1}{1 + r(0,0)\Delta t}
\end{aligned}
$$

Discrete time models and binomial implementation

- Then for $i = 1, \ldots, N-1$
  - solve for $P_D(i+1)$ using Newton-Raphson (or other numerical method) then solve for $P_U(i+1)$
  - use these to solve for $u(i)$ and $\sigma(i)$ substituting

$$d(i,j) = \frac{1}{1 + u(i)\exp\left(\sigma(i) j \sqrt{\Delta t}\right)}$$

to get

$$\begin{aligned} P_U(i+1) &= \sum_j \frac{Q_U(i,j)}{1 + u(i)\exp\left(\sigma(i) j \sqrt{\Delta t}\right)} \\ P_D(i+1) &= \sum_j \frac{Q_D(i,j)}{1 + u(i)\exp\left(\sigma(i) j \sqrt{\Delta t}\right)} \end{aligned}$$

for $i = 1, \ldots, N-1$, $j \in \{-i, -i-2, \ldots, i-2, i\}$.

Discrete time models and binomial implementation

- The one period short term interest rates are determined for each node using the calibrated values for $u(i)$ and $\sigma(i)$.
- State prices can then be updated.
- Any set of cash flows on the lattice can now be valued using the state prices including options on interest rates.
- Many other ways of contructing lattices to include mean reversion and other interest rate models.

Summary and Wrap Up

Summary and Wrap Up

- Aim has been to give a broad overview of Term Structure Models focussing on useful models for determining discount rate curves consistent with market data.
- Wide range of models and modelling techniques reviewed and summarised.
- Application to guarantees and options in liabilities not covered in details.
- Participants have a foundation for further study - encourage reading of more advanced texts and research articles.