Linear Gaussian Affine Term Structure Models with Unobservable Factors: Calibration and Yield Forecasting

Paresh Date, I. C. Wang

The Centre for the Analysis of Risk and Optimisation Modelling Applications (CARISMA), School of Information Systems, Computing and Mathematics, Brunel University.

Outline

- Linear Kalman filter
- Linear Gaussian interest rate models
- Numerical experiments
- Nonlinear filter for term structure matching
- Summary

CARISMA

Brunel UNIVERSITY WEST LONDON

Linear Kalman filter: basic theory

- Continuous time SDEs are often used to model dynamics of interest rates. Discretization of these equations may lead to variables which do not have a physical value.
- Two approaches:
  - Use "proxies" for unobserved variables (`e.g.` overnight interest rate/ 3-month LIBOR for short rate).
  - Estimate them from observed variables.
- Kalman filter is the main tool for the latter approach.
- Used in both model calibration and forecasting.

CARISMA

Brunel UNIVERSITY WEST LONDON

Linear Kalman filter: the set-up

- Consider a discrete time, linear state space system

$$x_{n+1} = Ax_n + B + \epsilon_{n+1} \quad \textit{Evolution Equation}$$
$$y_n = Cx_n + D + w_n \quad \textit{Measurement Equation}$$

- ϵ_n, w_n zero mean, Gaussian and uncorrelated.
- $A, B, C, D, \mathbb{E}(\epsilon_n \epsilon_n^T) = Q > 0, \mathbb{E}(w_n w_n^T) = R \geq 0$ are constants.
- Only y_n is measured; x_n is of interest (not directly observable).

Kalman filter: the aims

- To estimate the conditional mean $\hat{x}_{n+1|n}$ (and possibly, $\hat{x}_{n+1|n+1}$);
- To estimate the variance of error; $P_{n+1|n} = \mathbb{E}(x_{n+1} - \hat{x}_{n+1|n})^2$;
- To estimate the parameters.

Kalman filter equations

$$
\begin{aligned}
\textit{innovations} \qquad & v_n = y_n - \underbrace{(C\hat{x}_{n|n-1} + D)}_{\hat{y}_n} \\
\textit{variance of innovations} \qquad & \Sigma_n = CP_{n|n-1}C^T + R \\
\textit{Kalman gain} \qquad & K_n = AP_{n|n-1}C^T\Sigma_n^{-1} \\
\textit{conditional mean} \quad & \hat{x}_{n+1|n} = A\hat{x}_{n|n-1} + B + K_n v_n \\
\textit{conditional variance} \quad & P_{n+1|n} = AP_{n|n-1}A^T + Q \\
& \qquad - AP_{n|n-1}C^T\Sigma_n^{-1}CP_{n|n-1}A^T
\end{aligned}
$$

Kalman filter equations (cont'd)

- Equations are recursive. $\hat{x}_{n+1|n}$ is conditional mean of unobserved variable; $P_{n+1|n}$ is its conditional variance.
- We need x_0, P_0 to start with.
- Even if noise terms are not Gaussian, K_n as above yields smallest error variance amongst all linear filters.
- Under certain conditions on A, $K_n \to K$ as $n \to \infty$ where K is a constant matrix.

Kalman Filter equations: applications

- **Calibration:** Maximise, over parameter vector θ, joint (Gaussian) probability function of observations:

$$L(y_n, \theta) = \sum_{i=n}^{T} \log p(y_n | \mathcal{F}_{n-1}, \theta)$$
$$= -\sum_{n=0}^{T} \left(\log \det(\Sigma_n) + v_n \Sigma_n^{-1} v_n \right) + \text{a constant}$$

  where T is number of samples.

- **Forecasting:** Given parameters, find K_n at t_n and hence find conditional mean of x_{n+1} at "next" time-step; $\hat{x}_{n+1|n}$ using the above equations. Also yields a predicted variance $P_{n+1|n}$.

Basic term structure models

- With continuous compounding at constant rate r, the price of 1 Euro payable at time T is

$$\lim_{N\to\infty}\left(1+\frac{r}{N}\right)^{-NT}=e^{-rT}.$$

- If, r_t is driven by a stochastic process, the price at time t is $\mathbb{E}\left(e^{-\int_t^T r_s ds}\right)$ (under correct measure).
- Dynamics of the continuously compounded interest rate r_t can be inferred from observed bond/ securities prices or *spot rates*.

Basic term structure models (cont'd)

- *Spot rates* R_{t,τ_k} are interest rates at time t for time to maturity τ_k.
- *Yield curve* (or spot rate curve) gives (static) relationship between different R_{t,τ_k} and τ_k, at a given t.
- *Short rate* at time t, r_t, is the instantaneous spot rate. This is a theoretical construct.
- In a linear Gaussian model (and certain other models), R_{t,τ_k} is an affine function of r_t.

Linear Gaussian term structure model with unobservable factors

We consider a two factor, linear Gaussian short rate model (Babbs-Nowman model)

$$r_t = \mu - \sum_{i=1}^{2} X_{i,t}$$

$$dX_{i,t} = -\alpha_i X_{i,t} + \sum_{j=1}^{i} \sigma_{ij} dz_{j,t}. \tag{1}$$

Here, r_t is the (unobservable) short rate, α_i, σ_{ij}, μ *etc* are (unknown) constants. $z_{i,t}$ are uncorrelated standard Wiener processes.

- Theoretical spot rates $\left(= \frac{-1}{\tau_k} \log \mathbb{E}\left[\exp\left(-\int_t^{t+\tau_k} r_s\, ds\right)\right]\right)$:

$$r_t(\tau_k|\theta) = A_0(\tau_k|\theta) + \sum_{i=1}^{2} A_i(\tau_k|\theta) X_{i,t}$$

where $A_i(\cdot|\cdot)$ are known functions, θ is the vector of parameters.

- Actual spot rates are "observed in noise":

$$R_{t,\tau_k} = r_t(\tau_k, \theta) + \epsilon_{t,k} \tag{2}$$

with $\mathbb{E}(\epsilon_{t,\tau_k}) = 0$, $\mathbb{E}(\epsilon_{t,k}^2) = h_k^2$.

- Given observed R_{t,τ_k} for various maturities at each t and (discretised) short rate equation given earlier, we can use Kalman filtering to calibrate this model and forecast short rate/ spot rates.

Numerical experiments I: data

- Weekly data on 7 UK gilt yields (= government bond spot rates, published daily) from January 2001 to January 2006 from *Datastream*.
- 261 observations; 150 used for calibration and 100 for validation.
- 15 parameter, 2 factor linear model calibrated using Kalman filter +QML.

Numerical experiments II: parameter values

Parameter values			
α_1	0.7036	α_2	0.2807
σ_{11}	0.0080	σ_{22}	0.1806
λ_1	0.4249	λ_2	-0.0414
μ	0.2768	h_1	0.0006
h_2	0.0007	h_3	0.0014
h_4	0.0017	h_5	0.0011
h_6	0.0007	h_7	0.0010
σ_{12}	-0.0003		

$h_1^2, \cdots, h_7^2$ are measurement noise variances; λ_i are prices of risk.

Numerical experiments III: relative errors in yield prediction

- Measure of error: sample mean of the relative absolute error

$$\frac{|\text{observed yield-predicted yield}|}{\text{observed yield}}$$

  over the relevant set of observations (either in-sample or out-of sample).

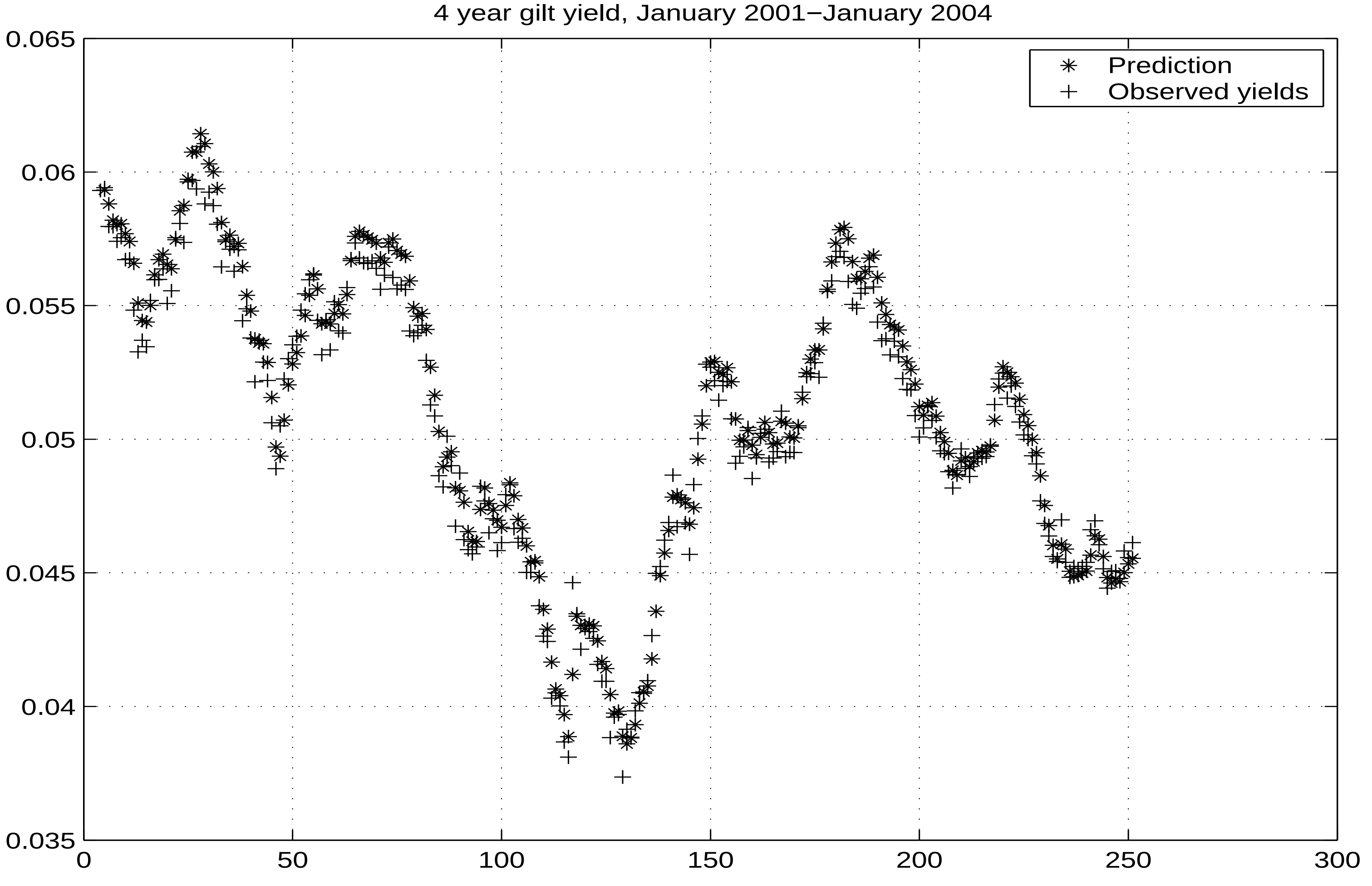
- Computed for one step ahead prediction of yields.

- Maximum out-of-sample error (over 7 yields) is only 2.31%, over a period of almost two years from calibration.

Mean relative absolute errors

yields	in-sample	out-of-sample
$3m$	0.0111	0.0118
$6m$	0.0178	0.0081
$1y$	0.0331	0.0216
$2y$	0.0337	0.0229
$4y$	0.0262	0.0152
$8y$	0.0209	0.0231
$10y$	0.0161	0.0213

Numerical Experiments IV: sample gilt yield plot



Samples from 150 onwards are post-calibration data.

Other numerical experiments

- Principal component analysis: 95% variation is explained by first two factors.
- One factor model with full or restricted yield spectrum: One factor is adequate if we use only short dated yields or only long dated yields.
- 4-step ahead prediction: Both one factor and two factor models perform similarly.

Nonlinear filter for term structure matching I: assumption

- Apart from yield prediction, another relevant problem is yield curve matching (*i.e.* finding values of $\hat{X}_{i,t_n|t_n}$ which best explain yield vector y_n at time t_n).
- For a calibrated model with parameter vector $\hat{\theta}$, we now assume that measurement noise is unknown-but-bounded:

$$|R_{t,\tau_k} - r_t(\tau_k|\hat{\theta})| < \delta_{t,k} R_{t,\tau_k}$$

  for some bounded $\delta_{t,k} > 0$.
- This does not change yield formulae; measurement noise variances do not affect bond pricing equations.

Nonlinear filter for term structure matching II: LP problem

- Now consider the problem at each time-step t

$$\min_{X_{i,t}} \max_{k} \frac{|R_{t,\tau_k} - r_t(\tau_k|\hat{\theta})|}{R_{t,\tau_k}}$$

- This is a linear programming problem in the state variables $X_{i,t}$; can be solved at each time-step to find $\hat{X}_{i,t|t}$ (instead of using a Kalman state update).

Nonlinear filter for term structure matching III: results

- Worst case mean relative absolute error improves from 2.67% (Kalman state update) to 2.22% (linear prog. update) in term structure matching; deteriorates from 2.31% (Kalman update) to 2.83% (lin. prog. update) in one step ahead prediction.
- We can also update μ at each step simultaneously with the state update ($r_t(\tau_k|\hat{\theta})$ is affine in μ).
- Consistent, *by definition*, to underlying model (unlike Nelsen-Siegel type curves or splines).

Summary

- Extensive numerical evidence presented on short term prediction ability of linear Gaussian term structure models with unobservable states.
- A two state model can perform very well out-of-sample in short term yield prediction.
- A new, simple nonlinear filter to improve term structure matching proposed; provides a term structure curve which is consistent with the assumed dynamics.
- Applications: pricing, scenario generation, bond portfolio optimization.