Dynamic model:
Local level model

Hedibert F. Lopes
Professor of Statistics and Econometrics

Insper Institute of Education and Research

`https://www.hedibert.org`

March 2026

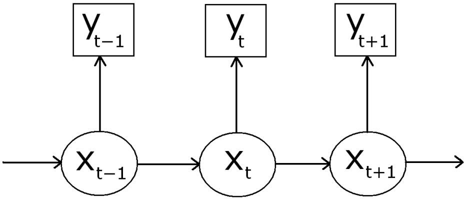
Outline

Dynamic models

• As opposed to observation-driven models, dynamic models or state-space models are characterized by the hierarchical structure that links observables and non-observables.

• For convenience, they have separate names for static non-observables (parameters) and the dynamic non-observables (state, hidden or latent variables).

• Below y_ts are the observed time series and x_t are the state or latent variables:



• In general y_t and y_{t-1} are independent given x_t, with temporal dependence driven the Markovian structure of the latent variables x_ts:

$$p(x_0, x_1, \ldots, x_n) = p(x_0) \prod_{t=1}^{n} p(x_t | x_{0:(t-1)}) = p(x_0) \prod_{t=1}^{n} p(x_t | x_{t-1}).$$

local level model

Let us start with the simples normal dynamic linear model, or normal DLM, the local level model of West and Harrison (1997).

Observation equation:

$$y_{t+1}|x_{t+1}, \theta \sim N(x_{t+1}, \sigma^2)$$

System equation:

$$x_{t+1}|x_t, \theta \sim N(x_t, \tau^2)$$

where

$$x_0 \sim N(m_0, C_0)$$

and

$$\theta = (\sigma^2, \tau^2)$$

fixed (for now).

n-variate normal

It is worth noticing that the normal local level model can be rewritten as a multivariate normal model with multivariate normal prior for the states:

$$
\begin{aligned}
y|x,\theta &\sim N(x,\sigma_2 I_n)\\
x|\theta &\sim N(m_0,C_0)N(x_0 1_n,\tau^2\Omega)
\end{aligned}
$$

where $x=(x_0,x_1,\ldots,x_n)'$, $y=(y_1,\ldots,y_n)'$, and

$$
\Omega = \begin{pmatrix}
1 & 1 & 1 & 1 & \ldots & 1 & 1 & 1\\
1 & 2 & 2 & 2 & \ldots & 2 & 2 & 2\\
1 & 2 & 3 & 3 & \ldots & 3 & 3 & 3\\
1 & 2 & 3 & 4 & \ldots & 4 & 4 & 4\\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\
1 & 2 & 3 & 4 & \ldots & n-2 & n-2 & n-2\\
1 & 2 & 3 & 4 & \ldots & n-1 & n-1 & n-1\\
1 & 2 & 3 & 4 & \ldots & n-2 & n-1 & n
\end{pmatrix}
$$

Ω and Ω^{-1}

The covariance matrix Ω has a tridiagonal inverse:

$$\Omega^{-1} = \begin{pmatrix} 2 & -1 & 0 & \ldots & 0 \\ -1 & 2 & -1 & \ldots & 0 \\ 0 & -1 & 2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

This sparsity is a key computational advantage in Bayesian state-space modeling, as we illustrate in what follows.

States posterior

The joint prior of x given θ is

$$x|\theta \sim N(m_0 1_n; C_0 1_n 1_n' + \tau^2\Omega),$$

while its full conditional posterior distribution is

$$x|y,\theta \sim N(m_1, C_1)$$

where

$$\begin{aligned}
C_1^{-1} &= (C_0 1_n 1_n' + \tau^2\Omega)^{-1} + \sigma^{-2} I_n \\
C_1^{-1} m_1 &= (C_0 1_n 1_n' + \tau^2\Omega)^{-1} m_0 1_n + \sigma^{-2} y
\end{aligned}$$

Sherman-Morrison-Woodbury

Sherman-Morrison-Woodbury: We can rewrite C_1^{-1} in terms of Ω^{-1} . Similarly, we can obtain m_1 without the need to invert C_1^{-1}. Also, sampling form $N(mu_1, C_1)$ can be down via a smart computation of the Cholesky decomposition of C_1^{-1}.

Woodbury matrix identity

Given a square invertible $n \times n$ matrix A, an $n \times k$ matrix U, and a $k \times n$ matrix V, let B be an $n \times n$ matrix such that $B = A + UV$. Then,

$$B^{-1} = A^{-1} - A^{-1}U(I_k + VA^{-1}U)^{-1}VA^{-1},$$

as long as $(I_k + VA^{-1}U)$ is invertible.

Kalman filter

Let $y^t = (y_1, \ldots, y_t)$. The previous joint posterior posterior for x given y (omitting θ for now) can be constructed as

$$p(x|y^n) = p(x_1|y^n, x_2) \prod_{t=1}^{n} p(x_t|y^n, x_{t+1}),$$

which is obtained from

$$p(x^n|y^n)$$

and noticing that given y^t and x_{t+1},

- x_t and x_{t+h} are independent, and
- x_t and y_t are independent,

for all integer $h > 1$.

Therefore, we first need to derive the above joint and this is done forward via the well-known Kalman filter recursions.

$$p(x_t|y^t) \Longrightarrow p(x_{t+1}|y^t) \Longrightarrow p(y_{t+1}|x_t) \Longrightarrow p(x_{t+1}|y^{t+1})$$

- Posterior at t: $(x_t|y^t) \sim N(m_t, C_t)$
- Prior at $t+1$: $(x_{t+1}|y^t) \sim N(m_t, R_{t+1})$

$$R_{t+1} = C_t + \tau^2$$

- Marginal likelihood: $(y_{t+1}|y^t) \sim N(m_t, Q_{t+1})$

$$Q_{t+1} = R_{t+1} + \sigma^2$$

- Posterior at $t+1$: $(x_{t+1}|y^{t+1}) \sim N(m_{t+1}, C_{t+1})$

$$\begin{aligned} m_{t+1} &= (1 - A_{t+1})m_t + A_{t+1}y_{t+1} \\ C_{t+1} &= A_{t+1}\sigma^2 \end{aligned}$$

where $A_{t+1} = R_{t+1}/Q_{t+1}$.

Kalman smoother

For $t = n$, $x_n | y^n \sim N(m_n^n, C_n^n)$, where $m_n^n = m_n$ and $C_n^n = C_n$.

For $t < n$,

$$
\begin{aligned}
x_t | y^n &\sim N(m_t^n, C_t^n) \\
x_t | x_{t+1}, y^n &\sim N(a_t^n, R_t^n)
\end{aligned}
$$

where

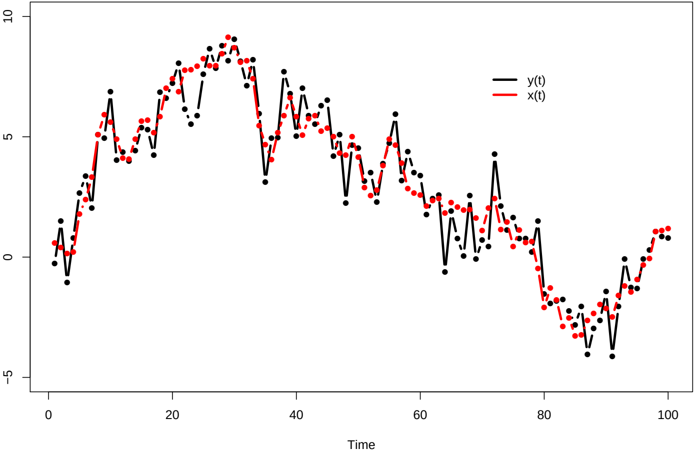
$$
\begin{aligned}
m_t^n &= (1 - B_t) m_t + B_t m_{t+1}^n \\
C_t^n &= (1 - B_t) C_t + B_t^2 C_{t+1}^n \\
a_t^n &= (1 - B_t) m_t + B_t x_{t+1} \\
R_t^n &= B_t \tau^2
\end{aligned}
$$

and

$$B_t = C_t / (C_t + \tau^2).$$

Example

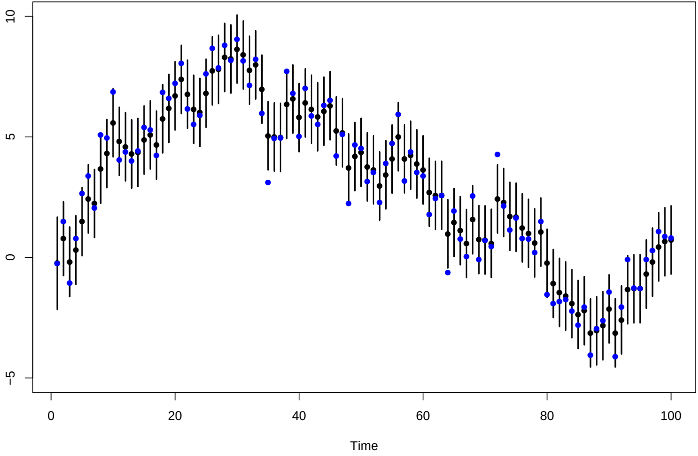
$n = 100$, $\sigma^2 = 1.0$
$\tau^2 = 0.5$ and $x_0 = 0$.



$p(x_t|y^t)$ via Kalman filter

$m_0 = 0.0$ and $C_0 = 10.0$

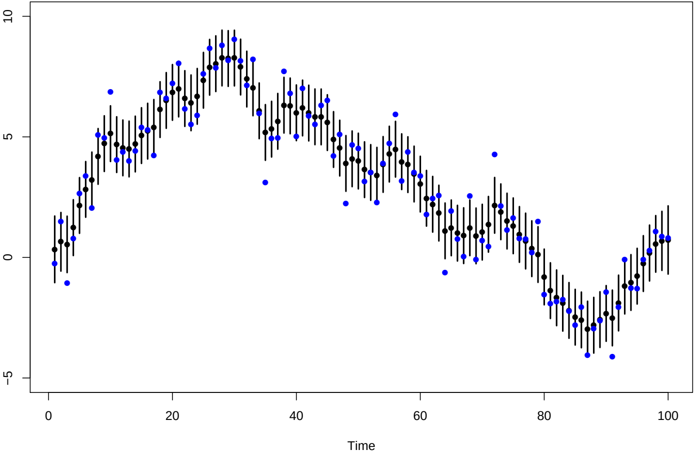
given τ^2 and σ^2



$p(x_t|y^n)$ via Kalman smoother

$m_0 = 0.0$ and $C_0 = 10.0$

given τ^2 and σ^2



Integrating out x^n

We showed earlier that

$$(y_t|y^{t-1}) \sim N(m_{t-1}, Q_t)$$

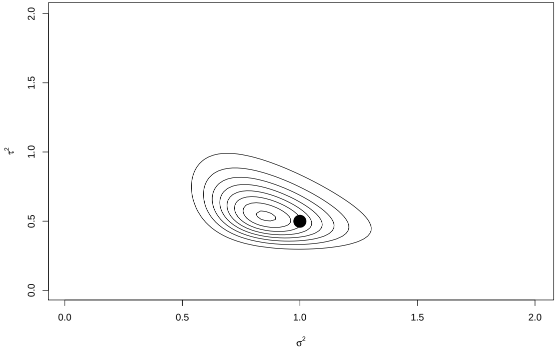
where both m_{t-1} and Q_t were presented before and are functions of $\theta = (\sigma^2, \tau^2)$, y^{t-1}, m_0 and C_0.

Therefore, by Bayes' rule,

$$\begin{aligned} p(\theta|y^n) &\propto p(\theta)p(y^n|\theta) \\ &= p(\theta)\prod_{t=1}^{n} f_N(y_t; m_{t-1}, Q_t). \end{aligned}$$

Example: $p(y|\sigma^2, \tau^2)p(\sigma^2)p(\tau^2)$

$\sigma^2 \sim IG(\nu_0/2, \nu_0\sigma_0^2/2)$, where $\nu_0 = 5$ and $\sigma_0^2 = 1$.
$\tau^2 \sim IG(n_0/2, n_0\tau_0^2/2)$, where $n_0 = 5$ and $\tau_0^2 = 0.5$



In this particular case, we can sample from the posterior of (σ^2, τ^2) via, say, SIR.

MCMC alternative

Alternatively, we could iteratively sample from $p(x|\theta, data)$ and then from $p(\theta|x, data)$.

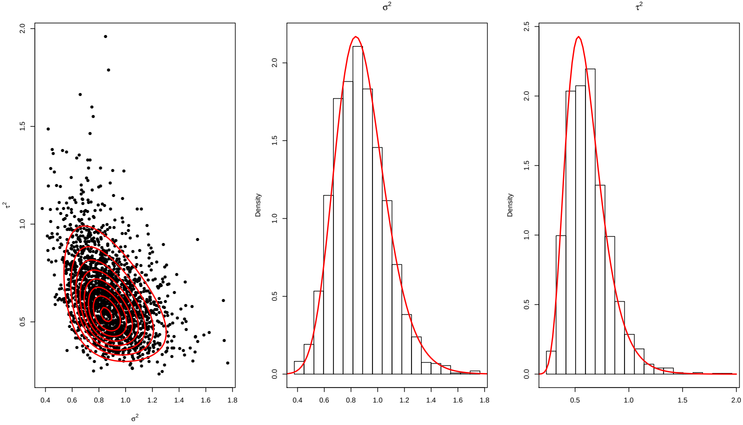
- Sample θ from $p(\theta|y^n, x^n)$

$$p(\theta|y^n, x^n) \propto p(\theta) \prod_{t=1}^{n} p(y_t|x_t, \theta) p(x_t|x_{t-1}, \theta).$$

- Sample x^n from $p(x^n|y^n, \theta)$

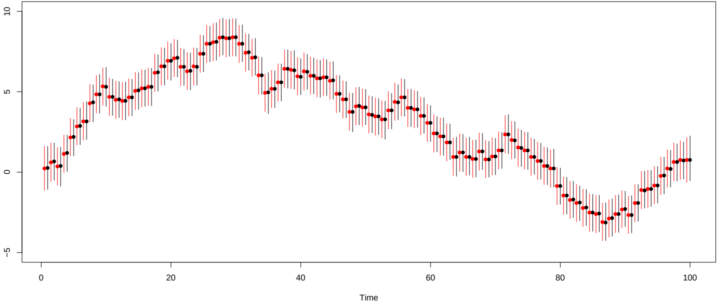
$$p(x^n|y^n, \theta) = \prod_{t=1}^{n} f_N(x_t|a_t^n, R_t^n)$$

Example: $p(x_t|y^n)$



Example: Comparison

$p(x_t|y^n)$ versus $p(x_t|y^n, \tilde{\sigma}^2 = 0.87, \tilde{\tau}^2 = 0.63)$.



Lessons from the 1st order DLM

Sequential learning in non-normal and nonlinear dynamic models $p(y_{t+1}|x_{t+1})$ and $p(x_{t+1}|x_t)$ in general rather difficult since

$$
\begin{aligned}
p(x_{t+1}|y^t) &= \int p(x_{t+1}|x_t)p(x_t|y^t)dx_t \\
p(x_{t+1}|y^{t+1}) &\propto p(y_{t+1}|x_{t+1})p(x_{t+1}|y^t)
\end{aligned}
$$

are usually unavailable in closed form.

Over the last 35+ years:

- FFBS for conditionally Gaussian DLMs;
- Gamerman (1998) for generalized DLMs;
- Carlin, Polson and Stoffer (2002) for more general DMs.

References

West and Harrison (1997) *Bayesian Forecasting and Dynamic Models (2nd edition)*. New York: Springer-Verlag.

Migon, Gamerman, Lopes and Ferreira (2005) Dynamic models, In *Handbook of Statistics, Volume 25: Bayesian Thinking, Modeling and Computation* (Eds. D. Dey and C. R. Rao), Amsterdam: Elsevier, 553-588.

Gamerman and Lopes (2006) *MCMC: Stochastic Simulation for Bayesian Inference*. Baton Rouge: Chapman & Hall/CRC.

Petris, Petrone and Campagnoli (2009) *Dynamic Linear Models with R*. New York: Springer.

Schmidt and Lopes (2019) Dynamic models. In Gelfand, Fuentes, Hoeting and Smith (Eds.), *Handbook of Environmental and Ecological Statistics*, pages 57-80. Chapman & Hall.

Prado, Ferreira and West (2021) *Time Series: Modeling, Computation and Inference (2nd edition)*. Baton Rouge: Chapman & Hall/CRC.