Dynamic GLMs

Hedibert F. Lopes
Professor of Statistics and Econometrics

Insper Institute of Education and Research

https://www.hedibert.org

March 2026

Outline

Dynamic generalized linear model

DGLMs were introduced by West, Harrison and Migon (1985).

The model is based on the exponential family of models:

$$
\begin{aligned}
f(y_t|\theta_t) &= a(y_t)\exp\{y_t\theta_t + b(\theta_t)\} \\
E(y_t|\theta_t) &= \mu_t \\
g(\mu_t) &= F_t'\beta_t \\
\beta_t &= G_t\beta_{t+1} + w_t
\end{aligned}
$$

with $w_t \sim N(0, W_t)$ and link function g.

The model is completed with a prior $\beta_1 \sim N(a, R)$.

It combines the prior specification of normal dynamic models with the observational structure of generalized linear models.

Dynamic binomial regression

Dynamic logistic regression with a series of binomial observations y_t with respective success probabilities π_t dynamically related to explanatory variables $x = (x_1, \ldots, x_d)'$ through the logistic link $\text{logit}(\pi_t) = x_t'\beta_t$.

$$\begin{aligned} y_t &\sim Binomial(n_t, \pi_t) \\ \text{logit}(\pi_t) &= x_t'\beta_t \\ \beta_t &= G_t\beta_{t+1} + w_t, \end{aligned}$$

with $w_t \sim N(0, W_t)$.

Dynamic Poisson regression

Poisson counts with means λ_t dynamically related through multiplicative perturbations $\lambda_t = \lambda_{t-1} w_t^*$. After a logarithmic transformation, one obtains $\log\,\lambda_t = \log\,\lambda_{t-1} + w_t$ with $w_t = \log\,w_t^*$.

$$
\begin{aligned}
y_t &\sim Poi(\lambda_t) \\
\log(\lambda_t) &= x_t'\beta_t \\
\beta_t &= G_t\beta_{t+1} + w_t,
\end{aligned}
$$

with $w_t \sim N(0, W_t)$.

Posterior inference via MCMC

Assuming that the variances of the system disturbances are constant, the model parameters are given by the state parameters $\beta = (\beta_1, \ldots, \beta_n)'$ and the system variance $W = \Phi^{-1}$.

The model is specified with the observation and system equations and completed with the independent prior distributions $\beta_1 \sim N(a, R)$ and $\Phi \sim W(n_W/2, n_W S_W/2)$.

The posterior distribution is given by

$$\pi(\beta, \Phi) \propto \prod_{t=1}^{n} f(y_t|\beta_t) \prod_{i=2}^{n} p(\beta_t|\beta_{t-1}, \Phi)\, p(\beta_1)p(\Phi)\ .$$

Full conditional for Φ

$$
\begin{aligned}
\pi_\Phi(\Phi) &\propto \prod_{t=2}^{n} p(\beta_t|\beta_{t-1},\Phi)\, p(\Phi) \\
&\propto \prod_{t=2}^{n} |\Phi|^{1/2} \exp\left\{-\frac{1}{2}\mathrm{tr}[(\beta_t - G_t\beta_{t-1})(\beta_t - G_t\beta_{t-1})'\Phi]\right\} \\
&\times |\Phi|^{[n_W-(p+1)]/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(n_W S_W \Phi)\right\} \\
&\propto |\Phi|^{[n_W^*-(d+1)]/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[(n_W^* S_W^*)\,\Phi\right]\right\} .
\end{aligned}
$$

that is the density of the $W(n_W^*/2, n_W^* S_W^*/2)$ distribution with

$$
\begin{aligned}
n_W^* &= n_W + n - 1 \\
n_W^* S_W^* &= n_W S_W + \sum_{t=2}^{n} (\beta_t - G_t\beta_{t-1})(\beta_t - G_t\beta_{t-1})'
\end{aligned}
$$

Full conditionals for β

For block β

$$
\begin{aligned}
\pi_\beta(\beta) &\propto \prod_{t=1}^n f(y_t|\beta_t) \prod_{t=2}^n p(\beta_t|\beta_{t-1},\Phi)\, p(\beta_1) \\
&\propto \exp\left\{\sum_{t=1}^n [y_t\theta_t + b(\theta_t)] - \frac{1}{2}\sum_{t=1}^n (\beta_t - G_t\beta_{t-1})'\Phi(\beta_t - G_t\beta_{t-1})\right\}.
\end{aligned}
$$

For block β_t, $t = 2, \ldots, n-1$

$$
\begin{aligned}
\pi_t(\beta_t) &\propto f(y_t|\beta_t)\, p(\beta_t|\beta_{t-1},\Phi) p(\beta_{t+1}|\beta_t,\Phi) \\
&\propto \exp\{y_t\theta_t + b(\theta_t)\} \exp\left\{-\frac{1}{2}\left[(\beta_t - G_t\beta_{t-1})'\Phi(\beta_t - G_t\beta_{t-1})\right.\right. \\
&+ \left.\left.(\beta_{t+1} - G_{t+1}\beta_t)'\Phi(\beta_{t+1} - G_{t+1}\beta_t)\ \right]\right\}.
\end{aligned}
$$

Similar results follow for blocks β_1 and β_n.

Sampling schemes

Knorr-Held and Besag (1998) and Knorr-Held (2000) suggested the use of independence chains with prior proposals.

Shephard and Pitt (1997) used independence chains with proposals based on both prior and a normal approximation to the likelihood.

Migon *et al.* (2013) used independence normal proposals for the block β with moments given by the approximation of West, Harrison and Migon (1985).

Singh and Roberts (1982) and Fahrmeir and Wagenpfeil (1997) extended to the dynamic setting the method of mode evaluation for static regression.

An alternative previously discussed is the reparametrization in terms of the system disturbances w_t (Gamerman, 1998)

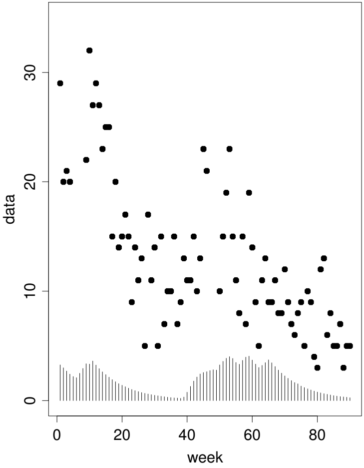
Example: Advertising awareness

Samples of $n_t = 66$ people were selected at random every week for an opinion poll and asked whether they remembered having seen a given advertising campaign on TV. A weekly cumulative measure of campaign expenditure was constructed.

Following Migon and Harrison (1985), the model used for this problem was a dynamic logistic regression

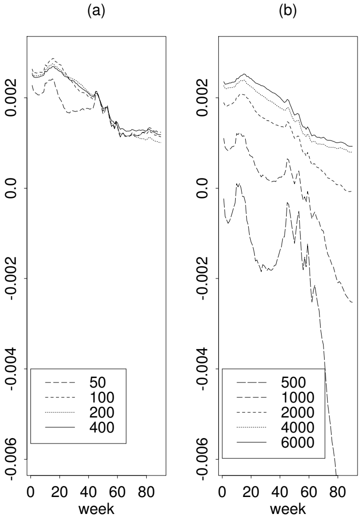
$$
\begin{aligned}
y_t &\sim bin(n_t, \pi_t) \\
\mu_t &= n_t \pi_t \\
\text{logit}(\pi_t) &= \beta_{1t} + \beta_{2t} x_t \\
\beta_t | \beta_{t-1} &\sim N(\beta_{t-1}, W)
\end{aligned}
$$

Data

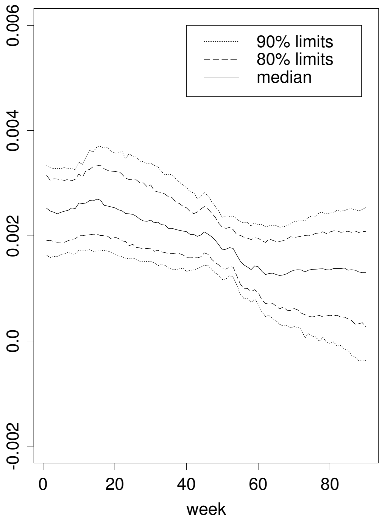


MCMC schemes

Average trajectory of β_{2t} in 500 parallel chains with number of iterations for sampling from: (a) system disturbances; (b) state parameters.



Expenditure coefficient β_{2t}



Example: Nonlinear dynamic model

Let y_t, for $t = 1, \ldots, n$, be generated by the following nonlinear dynamic model

$$
\begin{aligned}
(y_t | x_t, \sigma^2) &\sim N(x_t^2/20, \sigma^2) \\
(x_t | x_{t-1}, \theta, \tau^2) &\sim N(G'_{x_{t-1}}\theta, \tau^2) \\
x_0 &\sim N(m_0, C_0)
\end{aligned}
$$

where $G'_{x_t} = \left(x_t, x_t/(1 + x_t^2), \cos(1.2t)\right)$ and $\theta = (\alpha, \beta, \gamma)'$.

Prior distribution

$$
\begin{aligned}
\sigma^2 &\sim IG(n_0/2, n_0\sigma_0^2/2) \\
\theta|\tau^2 &\sim N(\theta_0, \tau^2 V_0) \\
\tau^2 &\sim IG(\nu_0/2, \nu_0\tau_0^2/2)
\end{aligned}
$$

Sampling $(\theta, \sigma^2, \tau^2 | x_{0:n}, y^n)$

Let $y^n = (y_1, \ldots, y_n)$ and $x_{0:n} = (x_0, \ldots, x_n)'$.

It follows that

$$
\begin{aligned}
(\theta, \tau^2 | x_{0:n}) &\sim N(\theta_1, \tau^2 V_1) IG(\nu_1/2, \nu_1 \tau_1^2/2) \\
(\sigma^2 | y^n, x^n) &\sim IG(n_1/2, n_1 \sigma_1^2/2)
\end{aligned}
$$

where $\nu_1 = \nu_0 + n$, $n_1 = n_0 + n$,

$$
\begin{aligned}
Z &= (G_{x_0}, \ldots, G_{x_{n-1}})' \\
V_1^{-1} &= V_0^{-1} + Z'Z \\
V_1^{-1}\theta_1 &= V_0^{-1}\theta_0 + Z' x_{1:n} \\
\nu_1 \tau_1^2 &= \nu_0 \tau_0^2 + (y - Z\theta_1)'(y - Z\theta_1) + (\theta_1 - \theta_0)' V_0^{-1} (\theta_1 - \theta_0) \\
n_1 \sigma_1^2 &= n_0 \sigma_0^2 + \sum_{t=1}^{n} (y_t - x_t^2/20)^2
\end{aligned}
$$

Sampling $(x_t|x_{-t},\theta,\sigma^2,\tau^2,y^n)$

Let $x_{-t}=(x_0,\ldots,x_{t-1},x_{t+1},\ldots,x_n)$, for $t=1,\ldots,n-1$, $x_{-0}=x^n$, $x_{-n}=x_{0:(n-1)}$, $y_0=\emptyset$, and $\psi=(\theta,\sigma^2,\tau^2)$.

For $t=0$

$$p(x_0|x_{-0},y_0,\psi)\propto f_N(x_0;m_0,C_0)f_N(x_1;G'_{x_0}\theta,\tau^2)$$

For $t=1,\ldots,n-1$

$$\begin{aligned}p(x_t|x_{-t},y_t,\psi) &\propto f_N(y_t;x_t^2/20,\sigma^2)\\ &\times f_N(x_t;G'_{x_{t-1}}\theta,\tau^2)\\ &\times f_N(x_{t+1};G'_{x_t}\theta,\tau^2)\end{aligned}$$

For $t=n$

$$p(x_n|x_{-n},y_n,\psi)\propto f_N(y_n;x_n^2/20,\sigma^2)f_N(x_n;G'_{x_{n-1}}\theta,\tau^2)$$

Metropolis-Hastings algorithm

A simple random walk Metropolis algorithm with tuning variance v_x^2 would work as follows. For $t = 0, \ldots, n$

1. Current state: $x_t^{(j)}$
2. Sample x_t^* from $N(x_t^{(j)}, v_x^2)$
3. Compute the acceptance probability

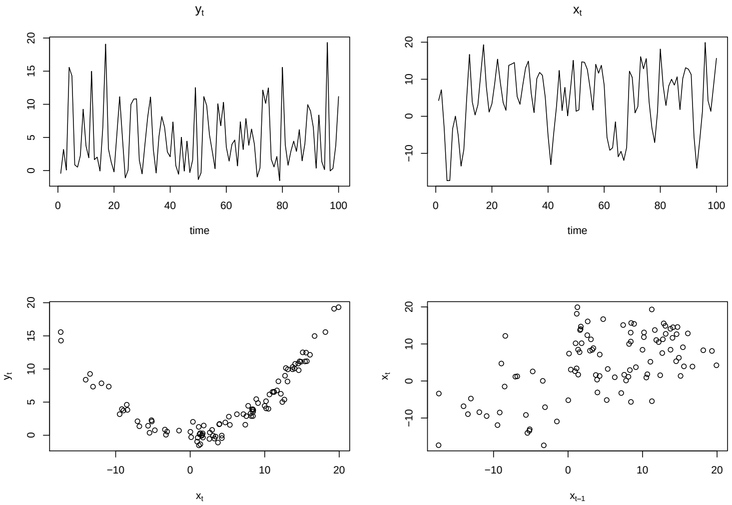
$$\alpha = \min\left\{1, \frac{p(x_t^*|x_{-t}, y_t, \psi)}{p(x_t^{(j)}|x_{-t}, y_t, \psi)}\right\}$$

4. New state:

$$x_t^{(j+1)} = \begin{cases} x_t^* & \text{w. p. } \alpha \\ x_t^{(j)} & \text{w. p. } 1-\alpha \end{cases}$$

Simulated exercise

We simulated $n = 100$ observations based on $\theta = (0.5, 25, 8)'$, $\sigma^2 = 1$, $\tau^2 = 10$ and $x_0 = 0.1$.



Prior hyperparameters

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

$$m_0 = 0.0 \quad \text{and} \quad C_0 = 10$$

- $\theta|\tau^2 \sim N(\theta_0, \tau^2 V_0)$

$$\theta_0 = (0.5, 25, 8)' \quad \text{and} \quad V_0 = \text{diag}(0.0025, 0.1, 0.04)$$

- $\tau^2 \sim IG(\nu_0/2, \nu_0\tau_0^2/2)$

$$\nu_0 = 6 \quad \text{and} \quad \tau_0^2 = 20/3$$

  such that $E(\tau^2) = \sqrt{V(\tau^2)} = 10.$

- $\sigma^2 \sim IG(n_0/2, n_0\sigma_0^2)$

$$n_0 = 6 \quad \text{and} \quad \sigma_0^2 = 2/3$$

  such that $E(\sigma^2) = \sqrt{V(\sigma^2)} = 1.$

MCMC setup

- Metropolis-Hastings tuning parameter

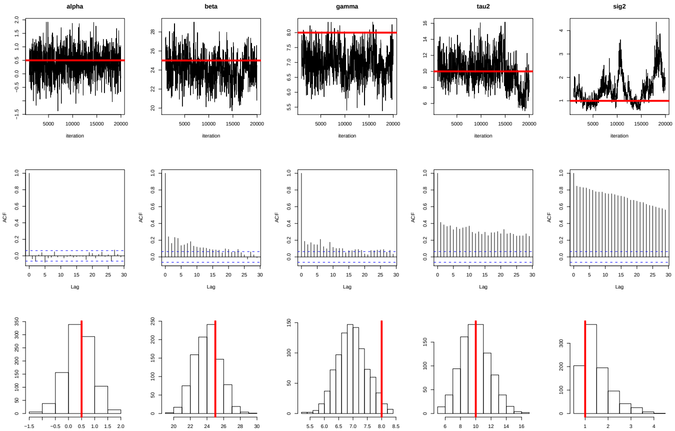
$$v_x^2 = (0.1)^2$$

- Burn-in period, step and MCMC sample size

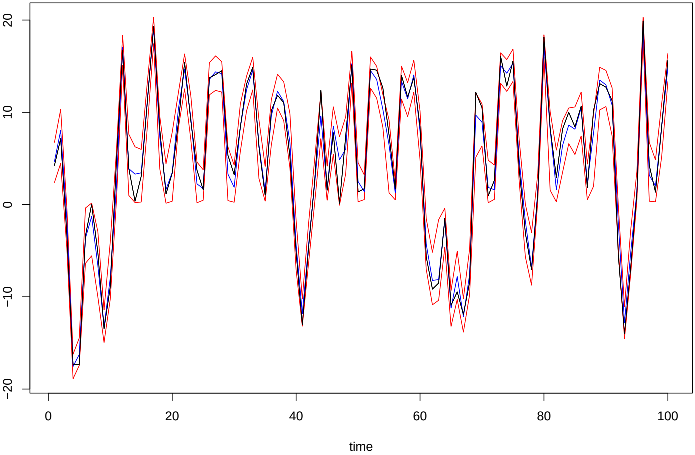
$$M_0 = 1,000 \quad L = 20 \quad M = 950 \;\Rightarrow\; 20,000 \text{ draws}$$

- Initial values
  - $\theta = (0.5, 25, 8)'$
  - $\tau^2 = 10$
  - $\sigma^2 = 1$
  - $x_{0:n} = x_{0:n}^{\text{true}}$

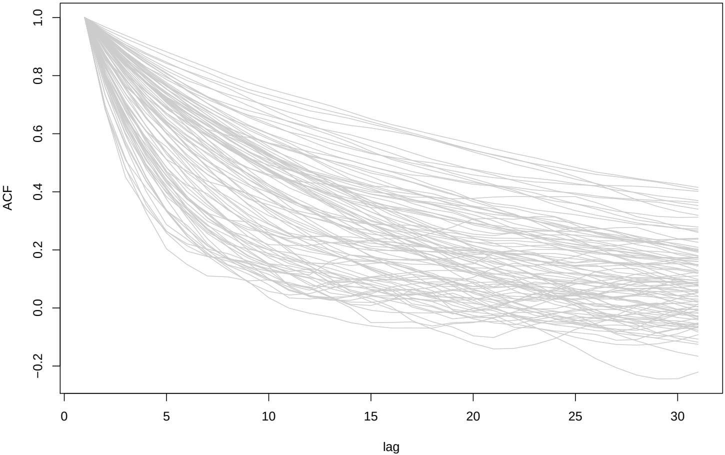
Parameters



States



States



References

Carlin, Polson and Stoffer (1992) A Monte Carlo approach to nonnormal and nonlinear state space modeling. *Journal of the American Statistical Association*, 87, 493-500.

Carter and Kohn (1994) On Gibbs Sampling for State Space Models. *Biometrika*, 81(3), 541-553.

Frühwirth-Schnatter (1994) Data augmentation and dynamic linear models,, *Journal of time series analysis*, 15(2), 183-202.

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.

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.

References (continuing)

Fahrmeir and Wagenpfeil (1997) Penalized likelihood estimation and iterative Kalman smoothing for non-Gaussian dynamic regression models. *Computational Statistics & Data Analysis*, 24(3), 295-320.

Gamerman (1998) Markov chain Monte Carlo for dynamic generalised linear models. *Biometrika*, 85(1), 215-227.

Knorr-Held and Besag (1998). Modelling risk from a disease in time and space. Journal of the Royal Statistical Society: Series C (Applied Statistics), 47(3), 415-431.

Knorr-Held (2000). Bayesian modelling of inseparable space-time variation in disease risk. Statistics in Medicine, 19(17-18), 2555-2567.

Migon, Schmidt, Ravines and Pereira (2013) An efficient sampling scheme for dynamic generalized models. *Computational Statistics*, 28, 2267-2293.

Shephard and Pitt (1997) Likelihood analysis of non-Gaussian measurement time series. Biometrika, 84(3), 653-667.

Singh and Roberts (1982) Analysis of discrete data from a survey sampling perspective. Technical Report, Statistics Canada. (Also related to their work in the Proceedings of the Section on Survey Research Methods).

West, Harrison and Migon (1985) Dynamic Generalized Linear Models and Bayesian Forecasting. *Journal of the American Statistical Association*, 80(389), 73-83.