Multivariate DLMs

Dr. Jarad Niemi

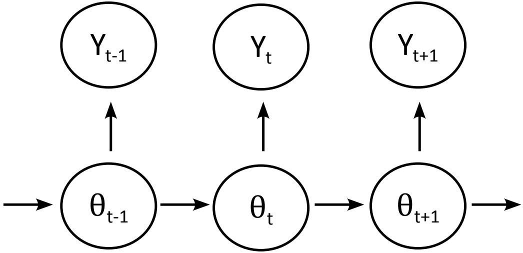
STAT 6150 - Iowa State University

October 14, 2025

Dynamic Linear Models

$$
\begin{aligned}
Y_t &= F_t\theta_t + v_t & v_t &\stackrel{ind}{\sim} N_m(0, V_t) \\
\theta_t &= G_t\theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_p(0, W_t) \\
& & \theta_0 &\sim N_p(m_0, C_0)
\end{aligned}
$$

where v_t and w_t are independent across time and all are independent of θ_0.



Longitudinal data

Suppose at each time point, you have observations from multiple units, e.g. people, countries, stocks, etc. The data can be represented as a matrix.

	Time			
Item	1	2	$\cdots$	T
1	$Y_{1,1}$	$Y_{1,2}$	$\cdots$	$Y_{1,T}$
2	$Y_{2,1}$	$Y_{2,2}$	$\cdots$	$Y_{2,T}$
$\vdots$		$\vdots$		
m	$Y_{m,1}$	$Y_{m,2}$	$\cdots$	$Y_{m,T}$

Individual univariate DLMs

Suppose for each unit $i = 1, \ldots, m$, we can write down a univariate DLM:

$$\begin{aligned} Y_{i,t} &= F\theta_t^{(i)} + v_{i,t}, & v_{i,t} &\stackrel{ind}{\sim} N(0, V_i) \\ \theta_t^{(i)} &= G\theta_{t-1}^{(i)} + w_t^{(i)}, & w_t^{(i)} &\stackrel{ind}{\sim} N_p(0, W_i) \end{aligned}$$

- F and G are constant for all i and t
- V_i and W_i are constant for all t
- $\theta_t^{(i)}$ are still vectors, thus the notation
- Is there a relationship between V_i and V_j or W_i and W_j for $i \neq j$?
- How about $Cov(v_{i,t}, v_{j,t})$ or $Cov(w_t^{(i)}, w_t^{(j)})$ for $i \neq j$?

Seemingly unrelated time series (SUTSE)

Linear trend models

Suppose you have a linear trend model for each unit i:

$$\begin{aligned} Y_{i,t} &= F_{i,t}\theta_t^{(i)} + v_{i,t} & v_{i,t} &\stackrel{ind}{\sim} N_m(0, V_{i,t}) \\ \theta_{i,t} &= G_{i,t}\theta_{t-1}^{(i)} + w_t^{(i)} & w_t^{(i)} &\stackrel{ind}{\sim} N_p(0, W_{i,t}) \end{aligned}$$

- What is $F_{i,t}$? $F_{i,t} = (1, 0)^\top$
- What is $G_{i,t}$?

$$G_{i,t} = \left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}\right]$$

- What is $V_{i,t}$? $V_{i,t} = V_i$
- What is $W_{i,t}$? $W_{i,t} = W_i$
- What is $\theta_t^{(i)}$? $\theta_t^{(i)} = (\mu_{i,t}, \beta_{i,t})$

Combine these into a DLM with $m = 2$:

$$\begin{aligned} Y_t &= F_t\theta_t + v_t & v_t &\stackrel{ind}{\sim} N_m(0, V_t) \\ \theta_t &= G_t\theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_p(0, W_t) \end{aligned}$$

- $\theta_t = (\mu_{1,t}, \mu_{2,t}, \beta_{1,t}, \beta_{2,t})$
- What is F_t?

$$F_t = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$$

- What is G_t?

$$G_t = \left[\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

- What is V_t?

$$V_t = V = \left[\begin{array}{cc} V_1 & ? \\ ? & V_2 \end{array}\right]$$

- What is W_t?

$$W_t = W = \left[\begin{array}{cc} W_\mu & 0 \\ 0 & W_\beta \end{array}\right]$$

- What are W_μ and W_β? If diagonal?

Kronecker products

Given two matrices A and B, the Kronecker product $A \otimes B$ is defined as

$$\begin{bmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \vdots & \vdots \\ a_{m,1}B & \cdots & a_{m,n}B \end{bmatrix}$$

- If A is 2x2, what is $A \otimes I_2$?

SUTSE model

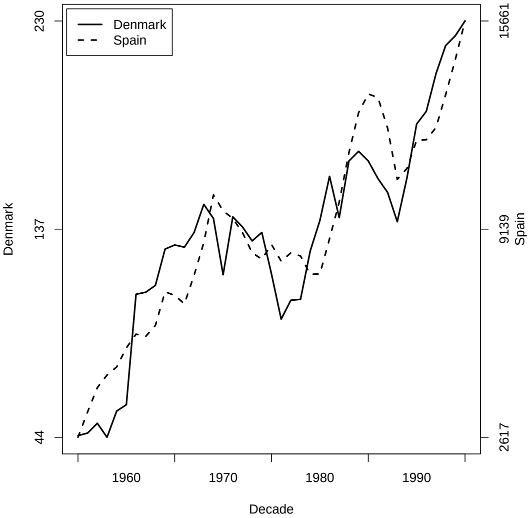
$$
\begin{aligned}
Y_t &= (F \otimes I_m)\theta_t + v_t & v_t &\stackrel{ind}{\sim} N_m(0, V) \\
\theta_t &= (G \otimes I_m)\theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_p(0, W)
\end{aligned}
$$

where

The covariance matrices V and W are typically somewhat sparse. Often

- off-diagonal elements of V are zero
- off-diagonal elements of W are zero
- some diagonal elements of W are zero

The data


Denmark
Spain
230
137
44
15661
9139
2617
Denmark
Spain
1960
1970
1980
1990
Decade

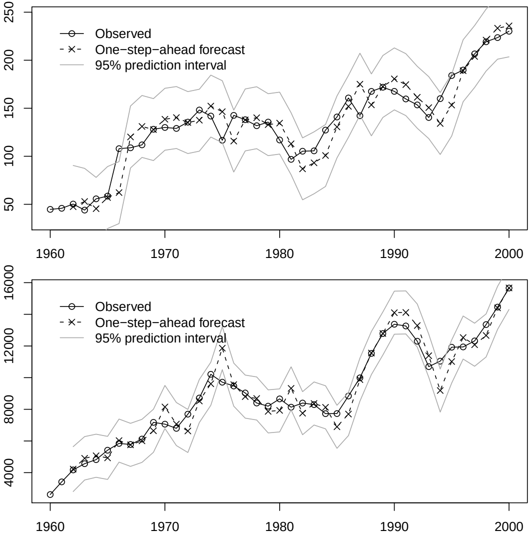
Model setup

```
> mod <- dlmModPoly(2)
> mod$FF <- mod$FF %x% diag(2)
> mod$GG <- mod$GG %x% diag(2)
>
> W1 <- matrix(0, 2, 2)
> W2 <- diag(c(49, 437266))
> W2[1, 2] <- W2[2, 1] <- 155
> mod$W <- bdiag(W1, W2)
>
> V <- diag(c(72, 14353))
> Var[1, 2] <- Var[2, 1] <- 1018
> mod$V <- V
>
> mod$m0 <- rep(0, 4)
> mod$C0 <- diag(4) * 1e7
> mod <- as.dlm(mod)
```

Filtering and one-step ahead prediction errors

```
> investFilt <- dlmFilter(invest, mod)
>
> sdev <- residuals(investFilt)$sd
> lwr <- investFilt$f + qnorm(0.025) * sdev
> upr <- investFilt$f - qnorm(0.025) * sdev
```

Filtering


Observed
One−step−ahead forecast
95% prediction interval
50
100
150
200
250
1960
1970
1980
1990
2000
Observed
One−step−ahead forecast
95% prediction interval
4000
8000
12000
16000
1960
1970
1980
1990
2000

Individual univariate dynamic regressions

Suppose for each unit i, we have a univariate dynamic regression model:

$$\begin{aligned} Y_{i,t} &= F_t^{(i)}\theta_t^{(i)} + v_{i,t}, & v_{i,t} &\stackrel{ind}{\sim} N(0, V_i) \\ \theta_t^{(i)} &= G_{i,t}\theta_{t-1}^{(i)} + w_t^{(i)}, & w_t^{(i)} &\stackrel{ind}{\sim} N_p(0, W_i) \end{aligned}$$

- What is $F_t^{(i)}$? $F_t^{(i)} = (x_{1,t}, \ldots, x_{p,t})$ if covariates are common to all series, as they will be in the example to follow
- What is $G_{i,t}$? $G_{i,t} = I_p$
- What is $V_{i,t}$? $V_{i,t} = V_i$
- What is $W_{i,t}$? $W_{i,t} = I_p$
- What is $\theta_t^{(i)}$? $\theta_t^{(i)} = (\beta_{1,t}^{(i)}, \ldots, \beta_{p,t}^{(i)})^\top$

Simple multivariate dynamic regression

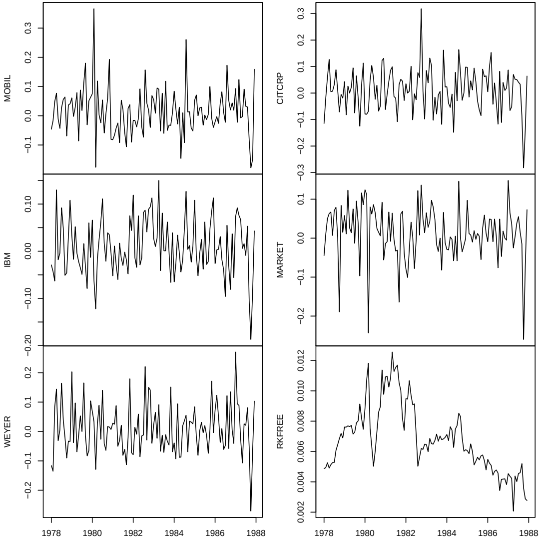
$$
\begin{aligned}
y_t &= (F_t \otimes I_m)\theta_t + v_t & v_t &\stackrel{ind}{\sim} N(0, V) \\
\theta_t &= (G \otimes I_m)\theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N(0, W)
\end{aligned}
$$

where

$$
y_t = \begin{bmatrix} y_{1,t} \\ \vdots \\ y_{m,t} \end{bmatrix}, \quad \theta_t = \begin{bmatrix} \alpha_{1,t} \\ \vdots \\ \alpha_{m,t} \\ \beta_{1,t} \\ \vdots \\ \beta_{m,t} \end{bmatrix}, \quad v_t = \begin{bmatrix} v_{1,t} \\ \vdots \\ v_{m,t} \end{bmatrix}, \quad w_t = \begin{bmatrix} w_{1,t} \\ \vdots \\ w_{m,t} \\ w_{m+1,t} \\ \vdots \\ w_{2m,t} \end{bmatrix}
$$

with $F_t = (1, x_t)^\top, G = I_2$, and $W = \text{blockdiag}(W_\alpha, W_\beta)$.

Data


MOBIL
CITCRP
IBM
MARKET
WEYER
RKFREE
1978
1980
1982
1984
1986
1988

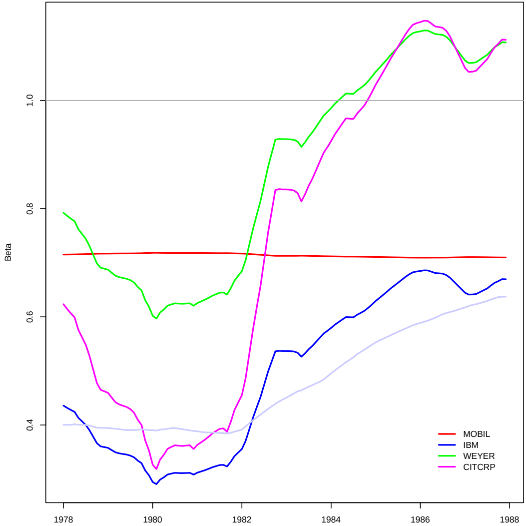
Model setup

```
> CAPM <- dlmModReg(market)
> CAPM$FF <- CAPM$FF %x% diag(m)
> CAPM$GG <- CAPM$GG %x% diag(m)
> CAPM$JFF <- CAPM$JFF %x% diag(m)
> CAPM$W <- CAPM$W %x% matrix(0, m, m)
> CAPM$W[-(1 : m), -(1 : m)] <-
+     c(8.153e-07,  -3.172e-05, -4.267e-05, -6.649e-05,
+       -3.172e-05, 0.001377,   0.001852,   0.002884,
+       -4.267e-05, 0.001852,   0.002498,   0.003884,
+       -6.649e-05, 0.002884,   0.003884,   0.006057)
> CAPM$V <- CAPM$V %x% matrix(0, m, m)
> CAPM$Var[] <- c(41.06,     0.01571, -0.9504, -2.328,
+                  0.01571, 24.23,     5.783,   3.376,
+                 -0.9504,   5.783,   39.2,     8.145,
+                 -2.328,    3.376,    8.145,  39.29)
> CAPM$m0 <- rep(0, 2 * m)
> CAPM$C0 <- diag(1e7, nr = 2 * m)
```

Smoothing inference

```
> CAPMsmooth <- dlmSmooth(y, CAPM)
```

Smoothed estimates



Longitudinal data

Suppose at each time point, you have observations from multiple units, e.g. people, countries, stocks, etc. The data can be represented as a matrix.

	Time			
Item	1	2	$\cdots$	T
1	$Y_{1,1}$	$Y_{1,2}$	$\cdots$	$Y_{1,T}$
2	$Y_{2,1}$	$Y_{2,2}$	$\cdots$	$Y_{2,T}$
$\vdots$		$\vdots$		
m	$Y_{m,1}$	$Y_{m,2}$	$\cdots$	$Y_{m,T}$

Individual DLMs

$$Y_{i,t} = F_{i,t}\theta_{i,t} + v_{i,t}, \quad v_{i,t} \stackrel{ind}{\sim} N(0, \sigma^2_{i,t})$$

SUTSE: Necessary to estimate blocks of $m \times m$ matrices in the covariance matrix W. Instead:

$$\begin{aligned} \theta_{i,t} &= \lambda_t + \epsilon_{i,t}, & \epsilon_{i,t} &\stackrel{ind}{\sim} N_k(0, \Sigma_t) \\ \lambda_t &= G\lambda_{t-1} + w_t, & w_t &\stackrel{ind}{\sim} N_k(0, W_t) \end{aligned}$$

Hierarchical DLM:

$$\begin{aligned} Y_t &= F_{y,t}\theta_t + v_t, & v_t &\stackrel{ind}{\sim} N_m(0, V_{y,t}) \\ \theta_t &= F_{\theta,t}\lambda_t + \epsilon_t, & \epsilon_t &\stackrel{ind}{\sim} N_P(0, V_{\theta,t}) \\ \lambda_t &= G_t\lambda_{t-1} + w_t, & w_t &\stackrel{ind}{\sim} N_k(0, W_t) \end{aligned}$$

Hierarchical DLMs

Hierarchical DLM:

$$\begin{aligned}
Y_t &= F_{y,t}\theta_t + v_t, & v_t &\stackrel{ind}{\sim} N_m(0, V_{y,t}) \\
\theta_t &= F_{\theta,t}\lambda_t + \epsilon_t, & \epsilon_t &\stackrel{ind}{\sim} N_P(0, V_{\theta,t}) \\
\lambda_t &= G_t\lambda_{t-1} + w_t, & w_t &\stackrel{ind}{\sim} N_k(0, W_t)
\end{aligned}$$

where

$$\begin{aligned}
\theta_t &= (\theta_{1,t}^\top, \ldots, \theta_{m,t}^\top)^\top \\
F_{y,t} &= \text{blockdiag}(F_{i,t}) \\
F_{\theta,t} &= [I_p|\cdots|I_p]^\top \\
V_{y,t} &= \text{diag}(\sigma_{i,t}^2) \\
V_{\theta,t} &= \text{blockdiag}(\Sigma_t) \\
G_t &= G
\end{aligned}$$

Integrating out θ_t

Let's write this as a standard DLM

$$
\begin{aligned}
Y_t &= F_{y,t}F_{\theta,t}\lambda_t + v_t^*, & v_t^* &\stackrel{ind}{\sim} N_m(0, F_{y,t}V_{\theta,t}F_{y,t}^\top + V_{y,t}) \\
\lambda_t &= G_t\lambda_{t-1} + w_t, & w_t &\stackrel{ind}{\sim} N_k(0, W_t)
\end{aligned}
$$

Longitudinal data with common covariates

Suppose at each time point, you have observations from multiple units, e.g. people, countries, stocks, etc., with a common covariate. The data can be represented as a matrix.

	Time				
Item	1	2	$\cdots$	T	$Covariate$
1	$Y_{1,1}$	$Y_{1,2}$	$\cdots$	$Y_{1,T}$	x_1
2	$Y_{2,1}$	$Y_{2,2}$	$\cdots$	$Y_{2,T}$	x_2
$\vdots$		$\vdots$			
m	$Y_{m,1}$	$Y_{m,2}$	$\cdots$	$Y_{m,T}$	x_m

Let's allow flexibility in how Y relates to x:

$$E(Y_t|x) = \sum_{j=1}^{k} \beta_{j,t} h_j(x)$$

where $h_j(x)$ may be, e.g.

powers	$: h_j(x) = x^j$
trig functions	$: h_j(x) = a_j \sin(b_j x)$
cubic splines	$: h_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3$

Then

$$Y_{i,t} = \sum_{j=1}^{k} \beta_{j,t} h_j(x_i) + \epsilon_{i,t}, \qquad \epsilon_{i,t} \stackrel{ind}{\sim} N(0, \sigma^2).$$

In matrix notation,

$$Y_t = F\beta_t + \epsilon_t, \qquad \epsilon_t \stackrel{ind}{\sim} N_m(0, V).$$

What are F and V?

Dynamic nonparametric regression

$$
\begin{aligned}
Y_t &= F\theta_t + \epsilon_t, && \epsilon_t \stackrel{ind}{\sim} N_m(0, V) \\
\theta_t &= G\theta_{t-1} + w_t, && w_t \stackrel{ind}{\sim} N_p(0, W_t)
\end{aligned}
$$

where

$$
\begin{aligned}
\theta_t &= \beta_t \\
F &= \begin{bmatrix} h_1(x_1) & \cdots & h_p(x_1) \\ \vdots & & \vdots \\ h_1(x_m) & \cdots & h_p(x_m) \end{bmatrix} \\
V &= \sigma^2 I_m
\end{aligned}
$$

Dynamic factor model

$$
\begin{aligned}
Y_t &= A\mu_t + v_t, & v_t &\stackrel{ind}{\sim} N_m(0,V) \\
\mu_t &= \mu_{t-1} + w_t, & w_t &\stackrel{ind}{\sim} N_p(0,W)
\end{aligned}
$$

where A is a fixed $m \times p$ matrix of factor loadings with $p < m$.

Identifiability of A. Suppose H is a $p \times p$ invertible matrix. Then

$$
\begin{aligned}
Y_t &= \tilde{A}\tilde{\mu}_t + v_t, & v_t &\stackrel{ind}{\sim} N_m(0,V) \\
\tilde{\mu}_t &= \tilde{\mu}_{t-1} + \tilde{w}_t, & \tilde{w}_t &\stackrel{ind}{\sim} N_p(0,HWH^\top)
\end{aligned}
$$

with $\tilde{\mu}_t = H\mu_t$ and $\tilde{A} = AH^{-1}$.

- How many parameters in A? mp
- How many parameters in W? $\frac{1}{2}p(p+1)$
- Number of free parameters is $mp - \frac{1}{2}p(p-1)$.

Identifiable dynamic factor models

One method of enforcing identifiability is

- Let $W = I_p$,
- Let $A = \left[\begin{array}{c} L \\ B \end{array}\right]$ where L is a lower triangular matrix and B can be anything.
- Total parameters are $mp - \frac{1}{2}p(p-1)$.

Another method is

- Let W be diagonal
- Let $A_{i,i} = 1$ and $A_{i,j} = 0$ for $j > i$.