WARNING: REFERENCES TO OTHER PARTS OF THE DOCUMENT WON'T WORK HERE

1 Model

The general dynamic linear model (DLM) is a linear, gaussian, state space model and can be written as

$$y_t = F_t\theta_t + v_t \qquad v_t \overset{ind}{\sim} N_k(0, V_t) \tag{1}$$

$$\theta_t = G_t\theta_{t-1} + w_t \qquad w_t \overset{ind}{\sim} N_p(0, W_t) \tag{2}$$

for $t = 1, 2, \cdots, T$, and $v_{1:T}$, $w_{1:T}$ independent. Equation (1) is called the *observation equation* and equation (2) is called the *system equation*. Similarly, $v_{1:T}$ are called the observation errors, $V_{1:T}$ are called the observation variances, $w_{1:T}$ are called the system disturbances and $W_{1:T}$ are called the system variances. The observed data is $y_{1:T}$ while $\theta_{0:T}$ are called the latent states. For each $t = 1, 2, \cdots, T$, F_t is a $k\times p$ matrix and G_t is a $p \times p$ matrix. Let ϕ denote the vector of unknown parameters in the model. Then possibly $F_{1:T}$, $G_{1:T}$, $V_{1:T}$, and $W_{1:T}$ are all functions of ϕ.

We will focus our attention on a simpler version of the DLM. Typically additional model structure is used to learn about $V_{1:T}$ and $W_{1:T}$ if time dependence is enforced – e.g. a stochastic volatility prior which would require a statespace model describing the $V_{1:T}$'s and $W_{1:T}$'s as data. Because of this additional complexity, we focus on the time-constant variances model, though many of our results may be useful in more complicated time-varying variance models. So we set $V_t = V$ and $W_t = W$ for $t = 1, 2, \cdots, T$. We will also suppose that F_t and G_t are known matrices for $t = 1, 2, \cdots, T$, though this constraint is immaterial since relaxing it will simply add one or more Gibbs steps to the algorithms we explore so long as no parameter that enters any F_t or G_t also enters V or W. Note, however, that in one of the data augmentations that we discuss, the scaled error data augmentation, there is a bit more housekeeping associated with $F_{1:T}$ depending on an unknown parameter (Section **??**).

When $\phi = (V, W)$ is our unkown parameter vector and we can write the model as

$$y_t|\theta_{0:T} \overset{ind}{\sim} N(F_t\theta_t, V) \tag{3}$$

$$\theta_t|\theta_{0:t-1} \sim N(G_t\theta_{t-1}, W) \tag{4}$$

To complete the model specification in a Bayesian context, we need priors on θ_0, V, and W. We'll use the standard approach and assume that they're mutually independent a priori and that $\theta_0 \sim N(m_0, C_0)$, $V \sim IW(\Lambda_V, \lambda_V)$ and $W \sim IW(\Lambda_W, \lambda_W)$ where m_0, C_0, Λ_V, λ_V, Λ_W, and λ_W are known hyperparameters and $IW(\Lambda, \lambda)$ denotes the inverse Wishart distribution with degrees of freedom λ and positive definite scale matrix Λ. This allows us to write the full joint distribution of $(V, W, \theta_{0:T}, y_{1:T})$ as

$$\begin{aligned}
p(V, W, \theta_{0:T}, y_{1:T}) \propto{}& \exp\left[-\frac{1}{2}(\theta_0 - m_0)'C_0^{-1}(\theta_0 - m_0)\right]\\
&\times |V|^{-(\lambda_V+k+T+2)/2}\exp\left[-\frac{1}{2}\,\mathrm{tr}\left(\Lambda_V V^{-1}\right)\right]\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(y_t - F_t\theta_t)'V^{-1}(y_t - F_t\theta_t)\right]\\
&\times |W|^{-(\lambda_W+p+T+2)/2}\exp\left[-\frac{1}{2}\,\mathrm{tr}\left(\Lambda_W W^{-1}\right)\right]\exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\theta_t - G_t\theta_{t-1})'W^{-1}(\theta_t - G_t\theta_{t-1})\right]
\end{aligned} \tag{5}$$

where $p = dim(\theta_t)$, $k = dim(y_t)$, and tr(.) is the matrix trace operator.