6. Time varying Bayesian VARs

6.1. A model with time-varying coefficients

In its simplest form, a time-varying VAR is similar to a traditional VAR, except that it admits the dynamic coefficients to become period-specific. Precisely, the model can expressed as:

$$y_t = A_{1,t} y_{t-1} + A_{2,t} y_{t-2} + ... + A_{p,t} y_{t-p} + C x_t + \varepsilon_t \tag{6.1.1}$$

And the residuals are distributed according to:

$$\varepsilon_t \sim \mathcal{N}(0, \Sigma) \tag{6.1.2}$$

Note that under this simple form, the model only allows for the VAR coefficients to be time-varying, while the residual covariance matrix Σ remains time-invariant. This assumption will be relaxed in the next section. The model can be reformulated for each period in compact form as:

$$y_t = \bar{X}_t \beta_t + \varepsilon_t \tag{6.1.3}$$

with:

$$\bar{X}_t = \underbrace{I_n \otimes X_t}_{n \times q} \quad , \quad X_t = \underbrace{\begin{pmatrix} y_{t-1}^{'} & y_{t-2}^{'} & \cdots & y_{t-p}^{'} & x_t^{'} \end{pmatrix}}_{1 \times k} \tag{6.1.4}$$

and

$$\beta_t^{'} = \underbrace{vec(B_t)}_{q \times 1} \quad , \quad B_t = \underbrace{\begin{pmatrix} A_{1,t}^{'} \\ A_{2,t}^{'} \\ \vdots \\ A_{p,t}^{'} \\ C_t^{'} \end{pmatrix}}_{k \times n} \tag{6.1.5}$$

The VAR coefficients are assumed to follow the following autoregressive process:

$$\beta_t = \beta_{t-1} + \nu_t \qquad , \qquad \nu_t \sim N(0, \Omega) \tag{6.1.6}$$

The covariance matrix Ω is assumed to be a random variable endogenously determined by the model.

This concludes the description of the model. The parameters of interest to be estimated are then: the VAR coefficients $\beta = \{\beta_1, \cdots, \beta_T\}$, the covariance matrix Ω for the shocks on the dynamic process, and the residual covariance matrix Σ.

Bayes rule for the model is given by:

$$\pi(\beta, \Omega, \Sigma \,|y) \propto f(y\,|\beta, \Sigma)\pi(\beta\,|\Omega)\pi(\Omega)\pi(\Sigma) \tag{6.1.7}$$

Start with the likelihood function. Independence between the residuals ε_t implies that the likelihood can be written as:

$$f(y\,|\beta, \Sigma) \propto \prod_{t=1}^{T} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(y_t - \bar{X}_t\beta_t)'\Sigma^{-1}(y_t - \bar{X}_t\beta_t)\right) \tag{6.1.8}$$

It is also possible to obtain a compact form for the likelihood by setting:

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix}}_{nT\times 1} = \underbrace{\begin{pmatrix} \bar{X}_1 & 0 & \cdots & 0 \\ 0 & \bar{X}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \bar{X}_T \end{pmatrix}}_{nT\times qT} \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_T \end{pmatrix}}_{qT\times 1} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix}}_{nT\times 1} \tag{6.1.9}$$

or:

$$y = \bar{X}B + \varepsilon \qquad , \qquad \varepsilon \sim N(0, \bar{\Sigma}) \qquad , \qquad \bar{\Sigma} = I_T \otimes \Sigma \tag{6.1.10}$$

Following:

$$f(y\,|\beta, \Sigma) \propto \left|\bar{\Sigma}\right|^{-1/2} \exp\left(-\frac{1}{2}(y - \bar{X}B)'\bar{\Sigma}^{-1}(y - \bar{X}B)\right) \tag{6.1.11}$$

The prior distribution $\pi(\beta\,|\Omega)$ for the VAR coefficients can take two forms. The first is a compact form proposed by Chan and Jeliazkov (2009). To obtain it, note that it is possible to express the dynamic process 6.1.6 in compact form as:

$$\underbrace{\begin{pmatrix} I_q & 0 & 0 & \cdots & 0 \\ -I_q & I_q & 0 & \cdots & 0 \\ 0 & -I_q & I_q & & \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & \cdots & I_q \end{pmatrix}}_{Tq \times Tq} \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_T \end{pmatrix}}_{Tq \times 1} = \underbrace{\begin{pmatrix} \beta_0 + \nu_1 \\ \nu_2 \\ \nu_3 \\ \vdots \\ \nu_T \end{pmatrix}}_{Tq \times 1} \tag{6.1.12}$$

or:

$$HB = \nu \tag{6.1.13}$$

The initial condition β_0 is assumed to follow the following normal distribution: $\beta_0 \sim N(0, (\tau - 1)\Omega)$, where τ denotes some positive scalar. Then independence implies that:

$$\text{var}(\beta_0 + \nu_1) = \text{var}(\beta_0) + \text{var}(\nu_1) = (\tau - 1)\Omega + \Omega = \tau\Omega \tag{6.1.14}$$

so that:

$$\text{var}(\nu) = \underbrace{\begin{pmatrix} \tau\Omega & 0 & 0 & \cdots & 0 \\ 0 & \Omega & 0 & \cdots & 0 \\ 0 & 0 & \Omega & & \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & \cdots & \Omega \end{pmatrix}}_{Tq \times Tq} = I_\tau \otimes \Omega \qquad \text{with} \qquad I_\tau = \underbrace{\begin{pmatrix} \tau & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & \cdots & 1 \end{pmatrix}}_{T \times T} \tag{6.1.15}$$

and hence $\nu \sim N(0, I_\tau \otimes \Omega)$.

Note that the law of motion 6.1.13 implies:

$$B = H^{-1}\nu \tag{6.1.16}$$

And from the affine property of the normal distribution, one then obtains that the prior distribution for β is given by:

$$\pi(\beta \,|\Omega\,) \sim N(0, H^{-1}(I_\tau \otimes \Omega)(H^{-1})') \tag{6.1.17}$$

or:

$$\pi(\beta\,|\Omega\,) \sim N(0,\Omega_0) \quad , \quad \Omega_0 = \underbrace{(H^{,}(I_{-\tau}\otimes\Omega^{-1})H)}_{Tq\times Tq}^{-1} \quad , \quad I_{-\tau} = \underbrace{\begin{pmatrix} \tau^{-1} & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & \cdots & 1 \end{pmatrix}}_{T\times T} \tag{6.1.18}$$

Then:

$$\pi(\beta\,|\Omega\,) \propto |\Omega_0| \exp\left(-\frac{1}{2}B^{,}\Omega_0^{-1}B\right) \tag{6.1.19}$$

The second form relies on a joint conditional formulation, using the fact that:

$$\pi(\beta\,|\Omega\,) = \pi(\beta_1\,|\Omega\,)\prod_{t=2}^{T}\pi(\beta_t\,|\Omega,\beta_{t-1}\,) \tag{6.1.20}$$

This yields:

$$\pi(\beta\,|\Omega\,) \propto |\Omega|^{-T/2} \exp\left(-\frac{1}{2}\left\{\beta_1^{,}(\tau\Omega)^{-1}\beta_1 + \sum_{t=2}^{T}(\beta_t-\beta_{t-1})^{,}\Omega^{-1}(\beta_t-\beta_{t-1})\right\}\right) \tag{6.1.21}$$

Consider now the prior for Ω. For simplicity, Ω is assumed to be a diagonal matrix. Each of the q diagonal entries ω_i, $i = 1,\ldots,q$ is assumed to follow an inverse gamma prior with shape $\frac{\chi_0}{2}$ and scale $\frac{\psi_0}{2}$.

$$\pi(\omega_i) \propto \omega_i^{-\frac{\chi_0}{2}-1} exp\left(-\frac{\psi_0}{2\omega_i}\right) \tag{6.1.22}$$

In order to implement a loose prior, the hyperparameter values are set at $\chi_0 = \psi_0 = 0.001$.

Finally, the prior distribution for Σ is inverse Wishart with scale S_0 and degrees of freedom κ_0. Hence:

$$\pi(\Sigma) \propto |\Sigma|^{-(\kappa_0+n+1)/2} \exp\left(-\frac{1}{2}tr\left\{\Sigma^{-1}S_0\right\}\right) \tag{6.1.23}$$

Given Bayes rule 6.1.7, the likelihood function 6.1.8, the prior distributions 6.1.19, 6.1.22 and 6.1.23, the joint posterior distribution is given by:

$$
\begin{aligned}
&\pi(\beta,\Omega,\Sigma\,|y\,) \\
&\propto \prod_{t=1}^{T}(2\pi)^{-n/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(y_t-\bar{X}_t\beta_t)'\Sigma^{-1}(y_t-\bar{X}_t\beta_t)\right) \\
&\times|\Omega_0|\exp\left(-\frac{1}{2}B'\Omega_0^{-1}B\right) \\
&\times\prod_{i=1}^{q}\omega_i^{-\frac{\chi_0}{2}-1}exp\left(-\frac{\psi_0}{2\omega_i}\right) \\
&\times|\Sigma|^{-(\kappa_0+n+1)/2}\exp\left(-\frac{1}{2}tr\left\{\Sigma^{-1}S_0\right\}\right)
\end{aligned}
$$

This function does not admit analytical solutions, so that the Gibbs sampling algorithm has to be used. Start with the conditional posterior for β. Consider Bayes rule 6.1.7 and relegate to the proportionality constant any term not involving β to obtain:

$$\pi(\beta\,|y,\Omega,\Sigma\,)\propto f(y\,|\beta,\Sigma\,)\pi(\beta\,|\Omega\,) \tag{6.1.24}$$

Substituting for 6.1.11 and 6.1.19 to obtain:

$$\pi(\beta\,|y,\Omega,\Sigma\,)\propto\left|\bar{\Sigma}\right|^{-1/2}\exp\left(-\frac{1}{2}(y-\bar{X}B)'\bar{\Sigma}^{-1}(y-\bar{X}B)\right)\times|\Omega_0|\exp\left(-\frac{1}{2}B'\Omega_0^{-1}B\right) \tag{6.1.25}$$

Rearranging then yields:

$$\pi(\beta\,|y\;,\Omega,\Sigma)\propto\exp\left(-\frac{1}{2}(B-\bar{B})'\bar{\Omega}^{-1}(B-\bar{B})\right) \tag{6.1.26}$$

with:

$$\bar{\Omega}^{-1}=\left(\Omega_0^{-1}+\bar{X}'\bar{\Sigma}^{-1}\bar{X}\right) \tag{6.1.27}$$

and

$$\bar{B}=\bar{\Omega}\left(\bar{X}'\bar{\Sigma}^{-1}y\right) \tag{6.1.28}$$

This is the kernel of a multivariate normal distribution with mean $\bar{B}$ and variance $\bar{\Omega}$:

$$\pi(\beta\,|y\;,\Omega,\Sigma)\sim N(\bar{B},\bar{\Omega}) \tag{6.1.29}$$

Consider then the conditional posterior distribution of Ω. More precisely, consider the conditional posterior for each ω_i, $i=1,\ldots,q$. Start from Bayes rule 6.1.7 and relegate to the proportionality constant any term not involving Ω to obtain:

$$\pi(\Omega\,|y,\beta,\Sigma\,) \propto \pi(\beta\,|\Omega\,)\pi(\Omega) \tag{6.1.30}$$

Substituting for the reformulated prior $\pi(\beta\,|\Omega\,)$ given by 6.1.21 and for the prior $\pi(\omega_i)$ given by 6.1.22, then rearranging, one obtains:

$$\begin{aligned} &\pi(\omega_i\,|y\;,\beta,\Sigma) \\ &\propto |\Omega|^{-T/2}\exp\left(-\frac{1}{2}\left\{\beta_1^{,}(\tau\Omega)^{-1}\beta_1+\sum_{t=2}^{T}(\beta_t-\beta_{t-1})^{,}\Omega^{-1}(\beta_t-\beta_{t-1})\right\}\right) \\ &\times\,\omega_i^{-\frac{\chi_0}{2}-1}exp\left(-\frac{\psi_0}{2\omega_i}\right) \end{aligned} \tag{6.1.31}$$

Rearranging eventually yields:

$$\pi(\omega_i\,|y,\beta,\omega_{-i},\Sigma\,) \propto \omega_i^{-\bar{\chi}-1}\exp\left(-\frac{\bar{\psi}}{\omega_i}\right) \tag{6.1.32}$$

with

$$\bar{\chi} = \frac{\chi_0+T}{2} \tag{6.1.33}$$

and:

$$\bar{\psi} = \frac{\beta_{i,1}^2/\tau+\sum\limits_{t=2}^{T}(\beta_{i,t}-\beta_{i,t-1})^2+\psi_0}{2} \tag{6.1.34}$$

where $\beta_{i,t}$ is element i of β_t, $i = 1,\ldots,q$. This is the kernel of an inverse Gamma distribution with shape $\bar{\chi}$ and degrees of freedom $\bar{\psi}$:

$$\pi(\omega_i\,|y\;,\beta,\omega_{-i},\Sigma) \sim IG(\bar{\chi},\bar{\psi}) \tag{6.1.35}$$

Finally, obtain the conditional posterior for Σ. Start from Bayes rule 6.1.7 and relegate to the proportionality constant any term not involving Σ to obtain:

$$\pi(\Sigma\,|y,\beta,\Omega\,) \propto f(y\,|\beta,\Sigma)\pi(\Sigma) \tag{6.1.36}$$

Substituting for the likelihood function 6.1.8 and the prior distribution 6.1.23, one obtains:

$$\begin{aligned}
&\pi(\Sigma\,|y,\beta,\Omega) \\
&\propto \prod_{t=1}^{T} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}(y_t - \bar{X}_t\beta_t)'\Sigma^{-1}(y_t - \bar{X}_t\beta_t)\right) \\
&\times |\Sigma|^{-(\kappa+n+1)/2} \exp\left(-\frac{1}{2}tr\left\{\Sigma^{-1}S\right\}\right)
\end{aligned} \tag{6.1.37}$$

Rearranging yields:

$$\pi(\Sigma\,|y,\beta,\Omega) \propto \exp\left(-\frac{1}{2}tr\left\{\Sigma^{-1}\bar{S}\right\}\right) \times |\Sigma|^{-(\bar{\kappa}+n+1)/2} \tag{6.1.38}$$

with:

$$\bar{\kappa} = T + \kappa_0 \tag{6.1.39}$$

and:

$$\bar{S} = \sum_{t=1}^{T}(y_t - \bar{X}_t\beta_t)(y_t - \bar{X}_t\beta_t)' + S_0 \tag{6.1.40}$$

Following, it is possible to introduce the general Gibbs sampling algorithm for the model:

Algorithm 6.1.1 (Gibbs sampler for the time-varying Bayesian VAR)

1. Determine the initial values $B^{(0)}, \Omega^{(0)}$ and $\Sigma^{(0)}$. Set $B^{(0)} = 1_T \otimes \hat{\beta}_{OLS}$, $\Omega^{(0)} = diag(\hat{\beta}_{OLS}\hat{\beta}'_{OLS})$, and $\Sigma^{(0)} = \hat{\Sigma}_{OLS}$.

2. At iteration n, draw $B^{(n)}$ from:
$\pi(\beta\,|y\ ,\Omega,\Sigma) \sim N(\bar{B},\bar{\Omega})$
with:
$\bar{\Omega}^{-1} = \left(\Omega_0^{-1} + \bar{X}'\bar{\Sigma}_{(n)}^{-1}\bar{X}\right)$
and
$\bar{B} = \bar{\Omega}\left(\bar{X}'\bar{\Sigma}_{(n)}^{-1}y\right)$

3. At iteration n, draw ${\omega_i}^{(n)}$ from:
$\pi(\omega_i\,|y\ ,\beta,\omega_{-i},\Sigma) \sim IG(\bar{\chi},\bar{\psi})$
with:

$\bar{\chi} = \frac{\chi_0}{2} + T$
and:
$\bar{\psi} = \frac{\beta_{i,1}^2/\tau + \sum\limits_{t=2}^{T}(\beta_{i,t} - \beta_{i,t-1})^2 + \psi_0}{2}$

4. At iteration n, draw $\Sigma^{(n)}$ from:
$\pi(\Sigma\,|y\;,\beta,\Omega) \sim IW(\bar{S},\bar{\alpha})$
with:
$\bar{\kappa} = T + \kappa_0$
and:
$\bar{S} = \sum\limits_{t=1}^{T}(y_t - \bar{X}_t\beta_t)(y_t - \bar{X}_t\beta_t)\text{'} + S_0$