1 Methodology

1.1 Factor VAR

We model the factors as following a vector-autoregressive (VAR) process, i.e.,

$$\underbrace{\begin{bmatrix} \mathbf{f1}_t \\ \mathbf{f2}_t \\ \vdots \\ \mathbf{fR}_t \end{bmatrix}}_{z_t} = B \underbrace{\begin{bmatrix} \mathbf{f1}_{t-1} \\ \mathbf{f2}_{t-1} \\ \vdots \\ \mathbf{fR}_{t-1} \end{bmatrix}}_{z_{t-1}} + C + \underbrace{\begin{bmatrix} v1_t \\ v2_t \\ \vdots \\ vR_t \end{bmatrix}}_{v_t},$$

where z_t is the $R \times 1$ matrix of time t factors,
B is the $R \times R$ coefficient matrix,
C is the $R \times 1$ constant matrix,
and v_t is the $R \times 1$ matrix of errors for time t.

We wish to estimate the coefficient matrices B and C. This can be done via OLS estimation. We first rewrite the data as the standard linear equation,

$$\underbrace{\begin{bmatrix} \mathbf{f1}_2 & \mathbf{f2}_2 & \dots & \mathbf{f6}_2 \\ \mathbf{f1}_3 & \mathbf{f2}_3 & \dots & \mathbf{f6}_3 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{f1}_T & \mathbf{f2}_T & \dots & \mathbf{f6}_T \end{bmatrix}}_{\Gamma} = \underbrace{\begin{bmatrix} 1 & \mathbf{f1}_1 & \mathbf{f2}_1 & \dots & \mathbf{f6}_1 \\ 1 & \mathbf{f1}_2 & \mathbf{f2}_2 & \dots & \mathbf{f6}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \mathbf{f1}_{T-1} & \mathbf{f2}_{T-1} & \dots & \mathbf{f6}_{T-1} \end{bmatrix}}_{\Psi} \underbrace{\begin{bmatrix} C' \\ B' \end{bmatrix}}_{\Lambda} + \underbrace{\begin{bmatrix} v1_2 & v2_2 & \dots & vR_2 \\ v1_3 & v2_3 & \dots & vR_3 \\ \vdots & & & \\ v1_T & v2_T & \dots & vR_T \end{bmatrix}}_{V},$$

where Γ is the $T-1 \times R$ dependent data matrix,
Ψ is the $T-1 \times R+1$ independent data matrix,
Λ is the $R+1 \times R$ matrix of coefficient weightings,
and V is the $T-1 \times R$ matrix of residuals.

The coefficient matrix Λ can be estimated by the standard OLS estimator.

$$\widehat{\Lambda} = (\Psi'\Psi)^{-1}(\Psi'\Gamma)$$

It can then be partitioned to calculate $\widehat{B}'$ and $\widehat{C}'$, which can then be transposed to derive our estimates of the original coefficient matrices B and C, $\widehat{B}$ and $\widehat{C}$.

1.2 Dynamic Factor Models

Now let us consider again the monthly covariates which were include in the principal components analysis. We will model these as dynamic factor models (DFMs), i.e. - they are regressed on the factor variables derived from earlier.

The factor models take the following form:

$$\underbrace{\begin{bmatrix} \mathbf{ue}_t \\ \mathbf{pce}_t \\ \vdots \\ \mathbf{claims}_t \end{bmatrix}}_{y_t} = A \underbrace{\begin{bmatrix} \mathbf{f1}_t \\ \mathbf{f2}_t \\ \vdots \\ \mathbf{fR}_t \end{bmatrix}}_{z_t} + \underbrace{\begin{bmatrix} w1_t \\ w2_t \\ \vdots \\ wN_t \end{bmatrix}}_{w_t},$$

where y_t is the $N \times 1$ vector of monthly variables at time t,
A is the $N \times R$ coefficient matrix,
z_t is the $R \times 1$ vector of factors at time t,
and w_t is the $N \times 1$ vector of errors at time t.

We wish to estimate the coefficient matrix A. As before, we can do this by estimating this as an OLS equation, writing the data matrices as follows

$$\underbrace{\begin{bmatrix} \mathbf{ue}_2 & \mathbf{pce}_2 & \dots & \mathbf{claims}_2 \\ \mathbf{ue}_3 & \mathbf{pce}_3 & \dots & \mathbf{claims}_3 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{ue}_T & \mathbf{pce}_T & \dots & \mathbf{claims}_T \end{bmatrix}}_{\Phi} = \underbrace{\begin{bmatrix} \mathbf{f1}_2 & \mathbf{f2}_2 & \dots & \mathbf{f6}_2 \\ \mathbf{f1}_3 & \mathbf{f2}_3 & \dots & \mathbf{f6}_3 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{f1}_T & \mathbf{f2}_T & \dots & \mathbf{f6}_T \end{bmatrix}}_{\Omega} A' + \underbrace{\begin{bmatrix} w1_2 & w2_2 & \dots & wR_2 \\ w1_3 & w2_3 & \dots & wR_3 \\ \vdots & & & \\ w1_T & w2_T & \dots & wR_T \end{bmatrix}}_{W} \tag{1}$$

As before we can estimate A with the standard OLS estimator.

$$\widehat{A'} = (\Omega'\Omega)^{-1}(\Omega'\Phi)$$

1.3 Kalman Filter

Now, combining our equations for the DFM and the VAR, we have the below system.

$$\begin{aligned} z_t &= Bz_{t-1} + Cx + v_t \\ y_t &= Az_t + w_t \end{aligned}$$

Since we have estimated our values B, C, and A in our previous two steps, this system is now fully specified and in state-space form. The first equation is our state (or transition) equation. The second equation is our measurement equation.

For Kalman filtration, we also require an assumed distribution on v_t and w_t. We assume that v_t is distributed normally with mean 0 and constant covariance calculated by taking the average covariance is.

We begin with the unconditional mean of $\widehat{Z}_{0|-1} = 0$ and unconditional variance of VAR $\Sigma_{0|-1} = 0$.

2 Test

$$z_t = Bz_{t-1} + v_t. \tag{state equation}$$
$$y_t = Hz_t + w_t. \tag{measurement equation}$$

Let v_t be distributed with variance Q, w_t be distributed with variance R. Let R be time-dependent with R_t set as 0 for vintages where data is unavailable.

$$\begin{aligned}\mathbf{z}_{1|0} &= B\mathbf{z}_{0|0}\\ \mathbf{CovZ}_{1|0} &= B\mathbf{CovZ}_{0|0}B' + Q\\ \mathbf{y}_{1|0} &= H\mathbf{z}_{1|0}\\ \mathbf{CovY}_{1|0} &= H\mathbf{CovZ}_{1|0}H' + R\end{aligned}$$

$$\begin{aligned}P_1 &= \mathbf{CovZ}_{1|0}H'\mathbf{CovY}_{1|0}\\ \mathbf{z}_{1|1} &= \mathbf{z}_{1|0} + P_1(\mathbf{y}_1 - \mathbf{y}_{1|0})\\ \mathbf{Covz}_{1|1} &= \mathbf{z}_{1|0} - P_1\mathbf{CovY}_{1|0}P_1'\end{aligned}$$

Forecasting step:

$$\begin{aligned}P_1 &= \mathbf{CovZ}_{1|0}H'\mathbf{CovY}_{1|0}^{-1}\\ \mathbf{z}_{1|1} &= \mathbf{z}_{1|0} + P_1(\mathbf{y}_1 - \mathbf{y}_{1|0})\\ \mathbf{Covz}_{1|1} &= \mathbf{z}_{1|0} - P_1\mathbf{CovY}_{1|0}P_1'\end{aligned}$$

diag(c(5, 10, 20, .5, .3, Inf, Inf))