FORECASTING WITH DYNAMIC FACTOR MODELS

Victor Pontines, PhD

Senior Economist

https://sites.google.com/site/victorpontineshomepage/

SEACEN Online Course on Forecasting for Monetary and Financial Stability

1 August – 5 August 2022

The South East Asian Central Banks Research and Training Centre

The SEACEN Centre

Outline

- Motivation

- The Set-Up of a Dynamic Factor Model

- Estimation of a Dynamic Factor Model

    - Kalman Filter and Kalman Smoother

- Dynamic Factor Model Forecasting

    - Application: Using the Stock-Watson (2016) Real Activity Dataset to Forecast using Dynamic Factor Modelling (Handbook of Macroeconomics Chapter on Dynamic Factor Model)

Motivation (1)

- The data structure that macroeconomists often encounter is peculiar:
  - The number of years for which there is a reliable and relevant data is limited and cannot readily be increased other than by the passage of time
  - On the other hand, in several countries, for much of the postwar period statistical agencies have collected monthly or quarterly data on a great number of macroeconomic, financial and sectoral variables
- Thus, macroeconomists encounter datasets that have hundreds or even thousand of series, but the number of observations per series is relatively short (e.g., 10, 20 to 30 years of quarterly data)

Motivation (2)

- The premise of dynamic factor models (DFMs):
  - The common dynamics of a large number of time series variables stem from a relatively small number of unobserved (or latent) factors, which in turn evolve over time
  - And, idiosyncratic terms that represent measurement errors or a specific feature relevant to an individual series
- Examples: GDP which are driven by a few factors representing the business cycle plus some measurement errors; equity returns which are driven by a few factors representing the market effect plus some idiosyncratic features specific to a company or sector
- DFMs stand out because the complex co-movements of a potentially large number of observable series are summarized by a small number of common factors which drive the common fluctuations of all the series

The Set-Up (1)

L: Log operator

- There are two ways to write DFMs:
  - Dynamic form: Represents the dependence of Y_t on lags of the factors explicitly
  - Static form: represents those dynamics implicitly

Dynamic Form of the DFM

- Expresses a $N \times 1$ vector Y_t of observed time series variables as depending on a reduced number q of unobserved or latent factors f_t and a mean-zero idiosyncratic component e_t, where both the latent factors and idiosyncratic terms are in general serially correlated:

$$Y_t = \lambda(\mathrm{L}) f_t + e_t \qquad (1)$$

$$F_t = \Psi(\mathrm{L}) f_{t-1} + \eta_t \qquad (2)$$

Error term

where the lag polynomial matrices $\lambda(\mathrm{L})$ and $\Psi(\mathrm{L})$ are $N \times q$ and $q \times q$, respectively, and η_t is the $q \times 1$ vector of (serially uncorrelated) mean-zero innovations to the factors

$\lambda_1 f_t + \lambda_2 f_{t-1} + \lambda_3 f_{t-2} + \dots + \lambda$

The Set-Up (2)

Dynamic Form of the DFM (cont.)

- The idiosyncratic disturbances are assumed to be uncorrelated with the factors
- The jth row of $\lambda(L)$ ($\lambda_j(L)$) is called the dynamic factor loading for the jth series, Y_{jt}
- The term $\lambda_j(L)f_t$ in equation (1) is the common component of the jth series

— $\lambda_j(L)$ is treated as one sided, which means that the common component of each series is a distributed lag of current and past values of f_t

- For some purposes, it is desirable to specify a parametric model for the idiosyncratic dynamics. A simple model is to suppose that the ith idiosyncratic disturbance, e_{it}, follows the univariate regression: $e_{it} = \delta_i(L)e_{it-1} + \nu_{it}$

where ν_{it} is serially uncorrelated

The SEACEN Centre

The Set-Up (3)

Stacked Form of the DFM

- The static, or stacked, form of the DFM rewrites the dynamic form (1) and (2) to depend on *r static factors* F_t instead of the *q* dynamic factors f_t

  — This rewriting makes the model amenable to principal components analysis and to other least squares methods

- Let *p* be the degree of the lag polynomial matrix λ(L) and let $F_t = (f_t', f_{t-1}', \ldots, f_{t-p}')'$ denote an $r \times 1$ vector of so-called "static" factors

- Also let $\Lambda = (\lambda_0, \lambda_1, \ldots, \lambda_p)$, where λ_l is the $N \times q$ matrix of coefficients on the hth lag in λ(L)

The SEACEN Centre

The Set-Up (4)

Stacked Form of the DFM (cont.)

- Similarly, let Φ (L) be the companion matrix consisting of 1s, 0s, and the elements of Ψ(L) such that the vector autoregression in (2) can then be written:

$$Y_t = \Lambda F_t + e_t \qquad (3)$$

$$F_t = \Phi(L)F_{t-1} + G\eta_t \qquad (4)$$

- For example, suppose that there is a single dynamic factor f_t (so q = 1), that all Y_{it} depend only on the current and first lagged values of f_t, and that f_t in (2) has two lags, so $f_t = \Psi_1 f_{t-1} + \Psi_2 f_{t-2} + \eta_t$

The SEACEN Centre

The Set-Up (5)

Stacked Form of the DFM (cont.)

- Then the correspondence between the dynamic and stacked forms for Y_{it} is:

$$Y_{it} = \lambda_{i0} f_t + \lambda_{i1} f_{t-1} + e_{it} = [\lambda_{i0} \quad \lambda_{i1}] \begin{bmatrix} f_t \\ f_{t-1} \end{bmatrix} + e_{it} = \Lambda_i F_t + e_{it} \qquad (5)$$

- where first expression in (5) writes out the equation for Y_{it} in dynamic form (1), $\Lambda = [\lambda_{i0}, \lambda_{i1}]$ is the ith row of Λ, and the final expression in (5) is the equation for Y_{it} in stacked form (3)

- In the stacked form of the DFM, the common component of the ith variable is $\Lambda_i F_t$ and the idiosyncratic component is e_{it}

The SEACEN Centre

The Set-Up (6)

Stacked Form of the DFM (cont.)

$$F_t = \begin{bmatrix} f_t \\ f_{t-1} \end{bmatrix} = \begin{bmatrix} \Psi_1 & \Psi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_{t-1} \\ f_{t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \eta_t = \Phi F_{t-1} + G\eta_t \qquad (6)$$

- The first row in the above is the evolution equation of the dynamic factor in equation (2) and the second row is the identity used to express (2) in first-order form (so called companion form)

The SEACEN Centre

The Set-Up (7)

Stacked Form of the DFM (cont.)

- Another example (q = 2 with same number of lags):

$$F_t = \begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} = \begin{bmatrix} \psi_{1,1}^{1} & \psi_{1,2}^{1} & \psi_{1,1}^{2} & \psi_{1,2}^{2} \\ \psi_{2,1}^{1} & \psi_{2,2}^{1} & \psi_{2,1}^{2} & \psi_{2,2}^{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_{1,t-1} \\ f_{2,t-1} \\ f_{1,t-2} \\ f_{2,t-2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \eta_t = \begin{bmatrix} \Psi_1 & \Psi_2 \\ 1 & 0 \end{bmatrix} F_{t-1} + G\eta_t \qquad (7)$$

- Most applications set r = q in equation (6). The reason being that when the numbers of static and dynamic factors are estimated using macroeconomic data, the difference between the estimated values of r and q is often small

Determining the Number of Factors: Two Formal Tests (1)

- Arguably, the most popular are the information criteria proposed by Bai and Ng (2002)

- Intuitively, the approach consists of estimating the factor model for an increasing number of factors beginning with $r = 1$ and comparing the goodness-of-fit of each factor model adjusted for the number of estimated factors in the model relative to the total number of observations, $N \times T$

- Similar to the use of information criteria in determining the optimal number of lags in a vector autoregression, among the set of candidates of r, the optimal number of factors is the value of r that minimises the computed Bai and Ng information criteria

Determining the Number of Factors: Two Formal Tests (2)

- There are further tests that modify and extend the information criteria proposed by Bai and Ng (2002)

- One of this tests is the eigenvalue ratio estimator of Ahn and Horenstein (2009) which proposes in estimating r as the maximiser of the ratio of two adjoining eigenvalues (ratio of eigenvalue k to eigenvalue $k + 1$)

DFM: Estimation

- Equations (1) and (2) of the dynamic form of the DFM and equations (3) and (4) of the stacked form of the DFM constitute a complete linear state-space system

- State-space estimation can then be used to obtain the parameters

  – The Kalman filter can be used to compute the likelihood and can be maximised to obtain maximum likelihood estimates of the parameters

  – The Kalman smoother can be used to compute estimates of f_t given the full-sample data on Y_t

How Does the Kalman Filter Work? (1)

- To illustrate the Kalman filter algorithm, the following simple linear state-space system is specified consisting of a single observable variable y_t and a single latent factor f_t

$$y_t = \lambda f_t + u_t$$

$$f_t = \phi f_{t-1} + v_t$$

where $u_t \sim N(0, \sigma^2)$ and $v_t \sim N(0, \sigma_v^2)$. $\lambda, \phi, \sigma, \sigma_v$ are unknown parameters

– The first equation above is also known as the measurement/observation equation, while the second equation is also referred to as the state/transition equation

How Does the Kalman Filter Work? (2)

- The Kalman filter is an iterative algorithm that starts from an initial estimate of the factor, obtained from the state equation for the first observation, without using any information on the observed variable at time $t = 1$

- This estimate of the factor is then used to compute an estimate of the observed variable at time $t = 1$ based on the measurement equation

- However, as the observed variable is indeed observed at $t = 1$, an updated estimate of the latent factor is then obtained by using information on y_1

- This recursion or sequential procedure is then applied to successive observations until observation T

- In rolling through the sample, it is assumed that the parameters of the state-space system are known

How Does the Kalman Filter Work? (3)

- For the linear state-space system, the Kalman filter equations are:

- **Prediction**: $f_{t|t-1} = \phi f_{t-1|t-1}$

  $P_{t|t-1} = \phi^2 P_{t-1|t-1} + \sigma_v^2$

  where P is the variance of the factor f_t

- **Observation**: $y_{t|t-1} = \lambda f_{t|t-1}$

  $V_{t|t-1} = \lambda^2 P_{t|t-1} + \sigma^2$

  where V is the variance of the observed variable y

How Does the Kalman Filter Work? (4)

- **Updating**: $$f_{t|t} = f_{t|t-1} + \frac{\lambda P_{t|t-1}}{V_{t|t-1}}(y_t - y_{t|t-1})$$

$$P_{t|t} = P_{t|t-1} - \frac{\lambda^2 P^2{}_{t|t-1}}{V_{t|t-1}}$$

The term $\frac{\lambda P_{t|t-1}}{V_{t|t-1}}$ is commonly referred to as the Kalman gain

- At t = 1, starting values for $f_{1|0}$ and $P_{1|0}$ are commonly chosen as:

$$f_{1|0} = 0 \qquad P_{1|0} = \sigma_v^2/(1 - \phi^2)$$

The starting value for $f_{1|0} = 0$ is its unconditional mean and the starting value for $P_{1|0}$ is its unconditional variance

How Does the Kalman Filter Work? (5)

- Numerical illustration of the recursion of the Kalman filter

- The parameters are: $\lambda = 1.0$, $\phi = 0.8$, $\sigma = 0.5$, $\sigma_v = 1$

- The data (y_t):

$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$
–0.680	0.670	0.012	–0.390	–1.477

- Prediction: $t = 1$ (Initialisation)

$$f_{1|0} = 0.0 \qquad P_{1|0} = \frac{\sigma_v^2}{1-\phi^2} = \frac{1}{1-0.8^2} = 2.778$$

The SEACEN Centre

How Does the Kalman Filter Work? (6)

- Numerical illustration of the recursion of the Kalman filter (cont.)

- Observation: $t = 1$

$$y_{1|0} = \lambda f_{1|0} = 1 \times 0.0 = 0.0$$
$$V_{1|0} = \lambda^2 P_{1|0} + \sigma^2 = 1.0^2 \times 2.778 + 0.5^2 = 3.028$$

- Updating: $t = 1$

$$f_{1|1} = f_{1|0} + \frac{\lambda P_{1|0}}{V_{1|0}}\left(y_1 - y_{1|0}\right) = 0.0 + \frac{1.0 \times 2.778}{3.028}(-0.680 - 0) = -0.624$$

$$P_{1|1} = P_{1|0} - \frac{\lambda^2 P^2{}_{1|0}}{V_{1|0}} = 2.778 - \frac{1.0^2 \times 2.778^2}{3.028} = 0.229$$

How Does the Kalman Filter Work? (7)

- Numerical illustration of the recursion of the Kalman filter (cont.)

- Prediction: $t = 2$

$$f_{2|1} = \phi f_{1|1} = 0.8 \times (-0.624) = -0.499$$

$$P_{2|1} = \phi^2 P_{1|1} + \sigma_v^2 = 0.8^2 \times 0.229 + 1 = 1.147$$

- Observation: $t = 2$

$$y_{2|1} = \lambda f_{2|1} = 1 \times -0.499 = -0.499$$

$$V_{2|1} = \lambda^2 P_{2|1} + \sigma^2 = 1.0^2 \times 1.147 + 0.5^2 = 1.397$$

How Does the Kalman Filter Work? (8)

- Numerical illustration of the recursion of the Kalman filter (cont.)

- Updating: $t = 2$

$$f_{2|2} = f_{2|1} + \frac{\lambda P_{2|1}}{V_{2|1}}\left(y_2 - y_{2|1}\right) = -0.499 + \frac{1.0 \times 1.147}{1.397}(0.670 - (-0.499)) = 0.461$$

$$P_{2|2} = P_{2|1} - \frac{\lambda^2 P^2{}_{2|1}}{V_{2|1}} = 1.147 - \frac{1.0^2 \times 1.147^2}{1.397} = 0.205$$

How Does the Kalman Filter Work? (9)

- Numerical illustration of the recursion of the Kalman filter (cont.)

- Updating: $t = 3$

$$f_{3|3} = 0.077$$
$$P_{3|3} = 0.205$$

- Updating: $t = 4$

$$f_{4|4} = -0.308$$
$$P_{4|4} = 0.205$$

- Updating: $t = 5$

$$f_{5|5} = -1.255$$
$$P_{5|5} = 0.205$$

- I leave the calculations as an exercise for you

How Does the Kalman Filter Work? (10)

- The maximum likelihood is to choose values of the parameters that generate latent factors that maximise the likelihood of the observed variables

- The log-likelihood function for the t[th] observation is given by:

$$\ln l_t = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|V_{t|t-1}| - \frac{1}{2}(y_t - y_{t|t-1})' V_{t|t-1}^{-1}(y_t - y_{t|t-1})$$

- Then, the log-likehood function at t = 1 is computed as:

$$\ln l_t = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|V_{1|0}| - \frac{1}{2}(y_1 - y_{1|0})' V_{1|0}^{-1}(y_1 - y_{1|0})$$

$$\ln l_t = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|3.028| - \frac{1}{2}(-0.680 - 0.0)' 3.028^{-1}(-0.680 - 0.0) = \mathbf{-1.549}$$

The SEACEN Centre

How Does the Kalman Filter Work? (11)

- The values for $\ln l_2, \ln l_3, \ln l_4, \ln l_5$ are also computed recursively

- For a sample of $t = 1,2,\cdots$, T observations, the log-likelihood function is:

$$\ln L_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} \ln l_t\,(\theta)$$

In which $\theta = [\lambda, \phi, \sigma, \sigma_v]$

- The above log-likelihood function is maximised

How Does the Kalman Smoother Work? (1)

- There is an additional recursion that delivers smooth estimates of the unobserved state variables or latent factors

- The smoothed estimate takes account of all the data in the observed sample and is obtained by conditioning on all T observations

- Formally, the equations for the smoothed conditional mean and variance of the factors are respectively:

$$f_{t|T} = f_{t|t} + J_t(f_{t+1|T} - f_{t+1|t})$$
$$P_{t|T} = P_{t|t} + J_t(P_{t+1|T} - P_{t+1|t})J_t'$$

where $J_t = P_{t|t}\phi' P_{t+1|t}^{-1}$

How Does the Kalman Smoother Work? (2)

- This full-set conditioning has the effect of generating smoother estimates of the factor than the filtered estimates

- The process of computing $f_{t|T}$ and $P_{t|T}$ essentially requires running the filter backward after the parameters have been estimated

- Numerical illustration of Kalman smoothing using the previous observed data

- at t = 5 (since T = 5), the smoothed estimates equal the updated factor estimates

$$f_{5|5} = -1.255$$
$$P_{5|5} = 0.205$$

How Does the Kalman Smoother Work? (3)

- Numerical illustration of Kalman smoothing using the previous observed data (cont.)

- at t = 4, $J_4 = P_{4|4}\phi' P_{5|4}^{-1} = 0.205 \times 0.8 \times 1.131^{-1} = 0.145$

  where: $P_{5|4} = \phi^2 P_{4|4} + \sigma_v^2 = 0.8^2 \times 0.205 + 1 = 1.131$

- Which then yields the respective smoothed estimates:

$$f_{4|5} = f_{4|4} + J_4\left(f_{5|5} - f_{5|4}\right) = -0.308 + 0.145(-1.255 - (-0.246)) = -0.454$$

where: $f_{5|4} = \phi f_{4|4} = 0.8 \times (-0.308) = -0.246$

$$P_{4|5} = P_{4|4} + J_4\left(P_{5|5} - P_{5|4}\right)J_4' = 0.205 + 0.145 \times (0.205 - 1.131) \times 0.145 = 0.185$$

How Does the Kalman Smoother Work? (4)

- <u>Numerical illustration of Kalman smoothing using the previous observed data (cont.)</u>

- Similarly, the remaining smooth-factor estimates conditional on information at time T=5 are:

- At t = 3

$$f_{3|5} = 0.002$$
$$P_{3|5} = 0.185$$

- At t = 2

$$f_{2|5} = 0.408$$
$$P_{2|5} = 0.185$$

- At t = 1

$$f_{1|5} = -0.479$$
$$P_{1|5} = 0.205$$

DFM: Forecasting

- H-step ahead forecasts can be computed in two ways:
  - Direct multistep forecasts are computed by regressing Y_{t+h} on F_t, y_t, lags of y_t
  - Alternatively, iterated multistep forecasts can be computed by first estimating a VAR for F_t then using this VAR to iterate forward h periods

Application: Stock-Watson (2016) Handbook of Macroeconomics Chapter on DFM (1)

- In section 6.2 of the chapter, Stock and Watson (S-W) (2016) constructed a dataset which they referred to as the real activity dataset consisting of 58 disaggregated series

  – A disaggregated series would refer, for instance, to the seven sectoral industrial production (IP) series, whereas, total IP is the higher-level aggregate series

  – To estimate the factors of the real activity dataset, S-W (2016) used the 58 disaggregated series to estimate the factors (of course, incorrect to include the aggregated series in the estimation of the factors)

Application: Stock-Watson (2016) Handbook of Macroeconomics Chapter on DFM (2)

- The 58 disaggregated series (last column of the table below) in the real activity dataset of S-W (2016):

	Category	Number of series	Number of series used for factor estimation
(1)	NIPA	20	12
(2)	Industrial production	11	7
(3)	Employment and unemployment	45	30
(4)	Orders, inventories, and sales	10	9

- Setting a maximum number of 11 factors, S-W (2016) finds 3 factors and 1 factor, respectively, using the Bai-Ng (2002) information criteria and Ahn and Horenstein (2009) test

The SEACEN Centre

Application: Stock-Watson (2016) Handbook of Macroeconomics Chapter on DFM (3)

- The estimated first factor of the real activity dataset of S-W (2016) for the period 1959Q1-2014Q4:

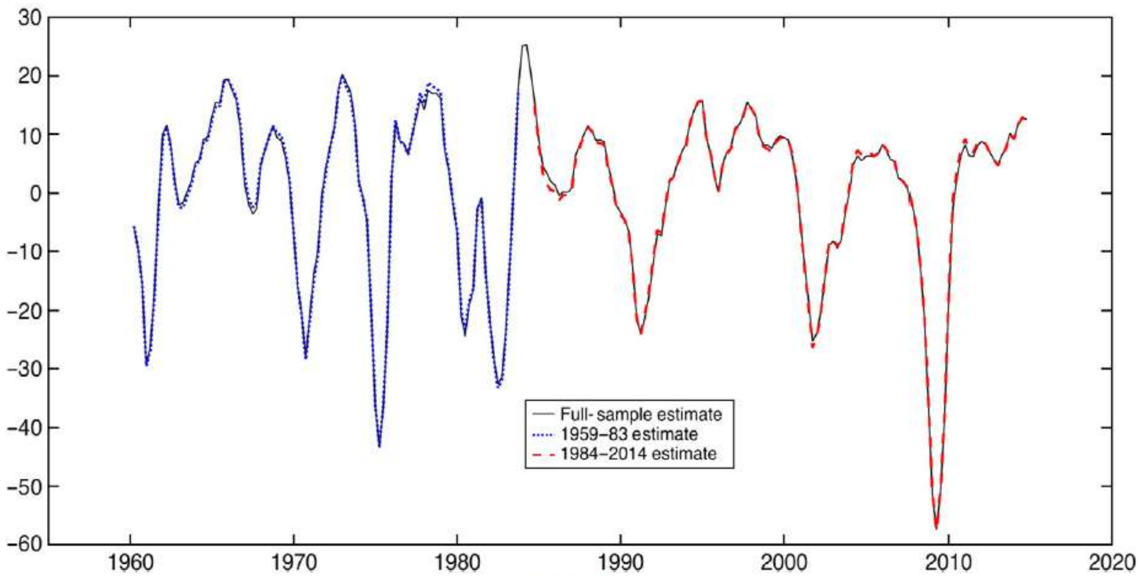

Fig. 5 First factor, real activity dataset: full sample, 1959–84, and 1984–2014.

The SEACEN Centre

Application: Stock-Watson (2016) Handbook of Macroeconomics Chapter on DFM (4)

- In our application, we will forecasts each of the 58 variables in the real activity dataset of S-W (2016) for the out-of-sample period of 2013Q1-2014Q4

Thank you for listening!

If you have any clarifications, comments or suggestions, please feel free to write me at vpontines@seacen.org

The SEACEN Centre