

### III. Linear state-space models

A. State-space representation of a dynamic system

---

---

---

---

---

---

---

---

Consider following model

State equation:

$$\underset{r \times 1}{\xi_{t+1}} = \underset{r \times r}{\mathbf{F}} \underset{r \times 1}{\xi_t} + \underset{r \times 1}{\mathbf{v}_{t+1}}$$

Observation equation:

$$\underset{n \times 1}{\mathbf{y}_t} = \underset{n \times k}{\mathbf{A}'} \underset{k \times 1}{\mathbf{x}_t} + \underset{n \times r}{\mathbf{H}'} \underset{r \times 1}{\xi_t} + \underset{n \times 1}{\mathbf{w}_t}$$

Observed variables:  $\mathbf{y}_t, \mathbf{x}_t$

Unobserved variables:  $\xi_t, \mathbf{v}_t, \mathbf{w}_t$

Matrices of parameters:  $\mathbf{F}, \mathbf{A}, \mathbf{H}$

---

---

---

---

---

---

---

---

$$\begin{bmatrix} \mathbf{v}_t \\ \mathbf{w}_t \end{bmatrix} \sim \text{i.i.d. } N \left( \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \right)$$

$$\mathbf{Q} = r \times r$$

$$\mathbf{R} = n \times n$$

---

---

---

---

---

---

---

---

Example 1:

$$\xi_{t+1} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{r-1} & \phi_r \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \xi_t + \begin{bmatrix} \varepsilon_{t+1} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

---

---

---

---

---

---

---

---

$$\begin{aligned} \xi_{j,t+1} &= L^{j-1} \xi_{1,t+1} \quad \text{for } j = 2, 3, \dots, r \\ \xi_{1,t+1} &= \phi_1 \xi_{1t} + \phi_2 L^1 \xi_{1t} + \phi_3 L^2 \xi_{1t} \\ &\quad + \cdots + \phi_p L^{p-1} \xi_{1t} + \varepsilon_{t+1} \\ \phi(L) \xi_{1,t+1} &= \varepsilon_{t+1} \end{aligned}$$

---

---

---

---

---

---

---

---

Observation equation:

$$y_t = \mu + \begin{bmatrix} 1 & \theta_1 & \theta_2 & \cdots & \theta_{r-1} \end{bmatrix} \xi_t$$

$$y_t - \mu = \theta(L) \xi_{1t}$$

put together with state equation:

$$\begin{aligned} \phi(L) \xi_{1t} &= \varepsilon_t \\ \phi(L)(y_t - \mu) &= \theta(L) \varepsilon_t \end{aligned}$$

Conclusion: any ARMA process can be written as a state-space model.

---

---

---

---

---

---

---

---

Example 2:

$C_t$  = state of business cycle

$\chi_{it}$  = idiosyncratic component for sector  $i$

$C_t, \chi_{it}$  unobserved

$y_{it}$  = growth in sector  $i$  (observed)

---

---

---

---

---

---

---

---

$$\xi_t = (C_t, \chi_{1t}, \chi_{2t}, \dots, \chi_{nt})'$$

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1}$$

$$\mathbf{F} = \begin{bmatrix} \phi_C & 0 & 0 & \dots & 0 \\ 0 & \phi_1 & 0 & \dots & 0 \\ 0 & 0 & \phi_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \phi_r \end{bmatrix}$$

---

---

---

---

---

---

---

---

Observation equation:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} + \begin{bmatrix} \gamma_1 & 1 & 0 & \dots & 0 \\ \gamma_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \gamma_n & 0 & 0 & \dots & 1 \end{bmatrix} \xi_t$$

---

---

---

---

---

---

---

---

Purpose of state-space representation:  
state vector  $\xi_t$  contains all information about  
system dynamics and forecasting.

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1}$$

$$\mathbf{y}_t = \mathbf{A}'\mathbf{x}_t + \mathbf{H}'\xi_t + \mathbf{w}_t$$

$$E(\mathbf{y}_{t+j} | \xi_t, \xi_{t-1}, \dots, \xi_1, \mathbf{y}_t, \mathbf{y}_{t-1},$$

$$\dots, \mathbf{y}_1, \mathbf{x}_{t+j}, \mathbf{x}_{t+j-1}, \dots, \mathbf{x}_1)$$

$$= \mathbf{A}'\mathbf{x}_{t+j} + \mathbf{H}'\mathbf{F}^j\xi_t$$

---

---

---

---

---

---

---

---

### III. Linear state-space models

- A. State-space representation of a dynamic system
- B. Kalman filter

---

---

---

---

---

---

---

---

Purpose of Kalman filter: calculate  
distribution of  $\xi_t$  conditional on

$$\Omega_t = \{\mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_1, \mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_1\}$$

$$\xi_t | \Omega_t \sim N(\hat{\xi}_{t|t}, \mathbf{P}_{t|t})$$

---

---

---

---

---

---

---

---

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1}$$

$$\mathbf{y}_t = \mathbf{A}'\mathbf{x}_t + \mathbf{H}'\xi_t + \mathbf{w}_t$$

$$\begin{bmatrix} \mathbf{v}_t \\ \mathbf{w}_t \end{bmatrix} \sim \text{i.i.d. } N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}\right)$$

---

---

---

---

---

---

---

---

Begin with the prior:

$$\xi_0 \sim N(\hat{\xi}_{0|0}, \mathbf{P}_{0|0})$$

$\hat{\xi}_{0|0}$  = prior best guess as to value of  $\xi_0$

$\mathbf{P}_{0|0}$  = uncertainty about this guess

(much uncertainty = large diagonal elements of  $\mathbf{P}_{0|0}$ )

---

---

---

---

---

---

---

---

$$\xi_1 = \mathbf{F}\xi_0 + \mathbf{v}_1$$

$$\xi_1 \sim N(\hat{\xi}_{1|0}, \mathbf{P}_{1|0})$$

$$\hat{\xi}_{1|0} = \mathbf{F}\hat{\xi}_{0|0}$$

$$\mathbf{P}_{1|0} = \mathbf{F}\mathbf{P}_{0|0}\mathbf{F}' + \mathbf{Q}$$

---

---

---

---

---

---

---

---

Useful result: suppose that

$$\begin{bmatrix} \mathbf{y}_1 | \mathbf{x} \\ \mathbf{y}_2 | \mathbf{x} \end{bmatrix} \sim N \left( \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right)$$

where  $\boldsymbol{\mu}_i$  and  $\boldsymbol{\Sigma}_{ij}$  may depend on  $\mathbf{x}$ . Then

$$\mathbf{y}_2 | \mathbf{y}_1, \mathbf{x} \sim N(\mathbf{m}^*, \mathbf{M}^*)$$

$$\mathbf{m}^* = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{y}_1 - \boldsymbol{\mu}_1)$$

$$\mathbf{M}^* = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}$$

---

---

---

---

---

---

---

---

Here

$$\begin{bmatrix} \mathbf{y}_1 | \mathbf{x}_1, \Omega_0 \\ \boldsymbol{\xi}_1 | \mathbf{x}_1, \Omega_0 \end{bmatrix} \sim N \left( \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right)$$

$$\boldsymbol{\mu}_2 = \hat{\boldsymbol{\xi}}_{1|0} \quad \boldsymbol{\Sigma}_{22} = \mathbf{P}_{1|0}$$

$$\boldsymbol{\mu}_1 = \mathbf{A}' \mathbf{x}_1 + \mathbf{H}' \hat{\boldsymbol{\xi}}_{1|0} \quad \boldsymbol{\Sigma}_{11} = \mathbf{H}' \mathbf{P}_{1|0} \mathbf{H} + \mathbf{R}$$

$$\boldsymbol{\Sigma}_{21} = \mathbf{P}_{1|0} \mathbf{H}$$

---

---

---

---

---

---

---

---

Hence

$$\boldsymbol{\xi}_1 | \mathbf{y}_1, \mathbf{x}_1, \Omega_0 = \boldsymbol{\xi}_1 | \Omega_1 \sim N(\hat{\boldsymbol{\xi}}_{1|1}, \mathbf{P}_{1|1})$$

$$\hat{\boldsymbol{\xi}}_{1|1} = \hat{\boldsymbol{\xi}}_{1|0} + \mathbf{P}_{1|0} \mathbf{H} (\mathbf{H}' \mathbf{P}_{1|0} \mathbf{H} + \mathbf{R})^{-1} \times (\mathbf{y}_1 - \mathbf{A}' \mathbf{x}_1 - \mathbf{H}' \hat{\boldsymbol{\xi}}_{1|0})$$

$$\mathbf{P}_{1|1} = \mathbf{P}_{1|0} -$$

$$\mathbf{P}_{1|0} \mathbf{H} (\mathbf{H}' \mathbf{P}_{1|0} \mathbf{H} + \mathbf{R})^{-1} \mathbf{H}' \mathbf{P}_{1|0}$$

---

---

---

---

---

---

---

---

Identical calculations: if  $\xi_t | \Omega_t \sim N(\hat{\xi}_{t|t}, \mathbf{P}_{t|t})$ ,

then  $\xi_{t+1} | \Omega_{t+1} \sim N(\hat{\xi}_{t+1|t+1}, \mathbf{P}_{t+1|t+1})$

$$\mathbf{P}_{t+1|t} = \mathbf{F} \mathbf{P}_{t|t} \mathbf{F}' + \mathbf{Q}$$

$$\mathbf{P}_{t+1|t+1} = \mathbf{P}_{t+1|t} -$$

$$\mathbf{P}_{t+1|t} \mathbf{H} (\mathbf{H}' \mathbf{P}_{t+1|t} \mathbf{H} + \mathbf{R})^{-1} \mathbf{H}' \mathbf{P}_{t+1|t}$$

$$\hat{\xi}_{t+1|t} = \mathbf{F} \hat{\xi}_{t|t}$$

$$\hat{\mathbf{e}}_{t+1|t} = \mathbf{y}_{t+1} - \mathbf{A}' \mathbf{x}_{t+1} - \mathbf{H}' \hat{\xi}_{t+1|t}$$

$$\hat{\xi}_{t+1|t+1} = \hat{\xi}_{t+1|t} +$$

$$\mathbf{P}_{t+1|t} \mathbf{H} (\mathbf{H}' \mathbf{P}_{t+1|t} \mathbf{H} + \mathbf{R})^{-1} \hat{\mathbf{e}}_{t+1|t}$$

---

---

---

---

---

---

---

---

Iterating on these calculations for  $t = 1, 2, \dots, T$  to produce the sequences

$\{\mathbf{P}_{t|t}\}_{t=1}^T$  and  $\{\hat{\xi}_{t|t}\}_{t=1}^T$  is called the

Kalman filter.

---

---

---

---

---

---

---

---

$\hat{\xi}_{t|t}$  is the posterior Bayesian expectation of  $\xi_t$  given observation of  $\Omega_t = \{\mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_1, \mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_1\}$ .

$$\mathbf{P}_{t|t} = E\left(\hat{\xi}_{t|t} - \xi_t\right)\left(\hat{\xi}_{t|t} - \xi_t\right)'$$

where these expectations condition on the values of  $\mathbf{F}, \mathbf{Q}, \mathbf{A}, \mathbf{H}, \mathbf{R}$ .

---

---

---

---

---

---

---

---

### III. Linear state-space models

- A. State-space representation of a dynamic system
- B. Kalman filter
- C. Using the Kalman filter
  - 1. Estimating the unknown parameters

---

---

---

---

---

---

---

---

Classical perspective

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1}$$

If eigenvalues of  $\mathbf{F}$  are all inside unit circle, set

$$\hat{\xi}_{0|0} = E(\xi_0) = \mathbf{0}$$

$$\mathbf{P}_{0|0} = E(\xi_0 \xi_0')$$

$$\text{vec}(\mathbf{P}_{0|0}) = [\mathbf{I}_{r^2} - (\mathbf{F} \otimes \mathbf{F})]^{-1} \text{vec}(\mathbf{Q})$$

---

---

---

---

---

---

---

---

Let  $\theta$  be vector containing unknown elements of  $\mathbf{F}, \mathbf{Q}, \mathbf{A}, \mathbf{H}, \mathbf{R}$

$$\mathbf{y}_t | \Omega_{t-1}, \mathbf{x}_t; \theta \sim N(\hat{\mathbf{y}}_{t|t-1}, \mathbf{C}_{t|t-1})$$

$$\hat{\mathbf{y}}_{t|t-1} = \mathbf{A}' \mathbf{x}_t + \mathbf{H}' \hat{\xi}_{t|t-1}$$

$$\mathbf{C}_{t|t-1} = \mathbf{H}' \mathbf{P}_{t|t-1} \mathbf{H} + \mathbf{R}$$

---

---

---

---

---

---

---

---



Classical econometrician: choose

$\hat{\theta}$  so as to maximize log likelihood:

$$-\frac{Tn}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |\mathbf{C}_{t|t-1}|$$
$$-\frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})' \mathbf{C}_{t|t-1}^{-1} (\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1})$$

---

---

---

---

---

---

---

---

### III. Linear state-space models

- A. State-space representation of a dynamic system
- B. Kalman filter
- C. Using the Kalman filter
  - 1. Estimating the unknown parameters
  - 2. Forecasting

---

---

---

---

---

---

---

---

Forecasting:

$$\mathbf{y}_t = \mathbf{A}' \mathbf{x}_t + \mathbf{H}' \xi_t + \mathbf{w}_t$$
$$E(\mathbf{y}_{t+j} | \Omega_t, \mathbf{x}_{t+j}, \mathbf{F}, \mathbf{Q}, \mathbf{A}, \mathbf{H}, \mathbf{R})$$
$$= \mathbf{A}' \mathbf{x}_{t+j} + \mathbf{H}' \mathbf{F}^j \hat{\xi}_{t|t}$$

MSE for  $j = 1$ :

$$E(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})(\mathbf{y}_{t+1} - \hat{\mathbf{y}}_{t+1|t})'$$
$$= \mathbf{H}' \mathbf{P}_{t+1|t} \mathbf{H} + \mathbf{R}$$

---

---

---

---

---

---

---

---

### III. Linear state-space models

- A. State-space representation of a dynamic system
- B. Kalman filter
- C. Using the Kalman filter
  - 1. Estimating the unknown parameters
  - 2. Forecasting
  - 3. Smoothed inference

---

---

---

---

---

---

---

---

Smoothed inference: might also want to form inference about  $\xi_t$  using all the data  $\Omega_T$ :

$$\xi_t | \Omega_T \sim N(\hat{\xi}_{t|T}, \mathbf{P}_{t|T})$$

---

---

---

---

---

---

---

---

To derive formula, consider instead

$$\xi_t | \xi_{t+1}, \Omega_t \sim N(\xi_{t|t}^*, \mathbf{P}_{t|t}^*)$$

Same kind of derivation as for Kalman filter establishes that

$$\xi_{t|t}^* = \hat{\xi}_{t|t} + \mathbf{J}_t (\xi_{t+1} - \hat{\xi}_{t+1|t})$$

$$\mathbf{J}_t = \mathbf{P}_{t|t} \mathbf{F}' \mathbf{P}_{t+1|t}^{-1}$$

$$\mathbf{P}_{t|t}^* = \mathbf{P}_{t|t} - \mathbf{J}_t \mathbf{F} \mathbf{P}_{t|t}$$

Generalization: what if  $\mathbf{P}_{t+1|t}$  is singular?

---

---

---

---

---

---

---

---

If  $\mathbf{P}_{t+1|t}$  is singular, then some linear combinations of  $\xi_{t+1}$  can be forecast perfectly from  $\Omega_t$ , implying inference about  $\xi_t$  given  $\Omega_t$  and these linear combinations of  $\xi_{t+1}$  is identical to inference about  $\xi_t$  given  $\Omega_t$  alone.

---

---

---

---

---

---

---

---

Let  $\xi_t$  be  $(r \times 1)$  and let the rank of  $\mathbf{P}_{t+1|t}$  be  $s \leq r$ . Define the  $(s \times 1)$  vector  $\xi_t^{**} = \mathbf{H}^{**} \xi_t$  for an arbitrary  $(s \times r)$  matrix  $\mathbf{H}^{**}$  such that  $\mathbf{P}_{t+1|t}^{**} \equiv \mathbf{H}^{**} \mathbf{P}_{t+1|t} \mathbf{H}^{**'}$  has rank  $s$ . For example,  $\xi_t^{**}$  might be the first  $s$  elements of  $\xi_t$  in which case  $\mathbf{P}_{t+1|t}^{**}$  would be first  $s$  rows and columns of  $\mathbf{P}_{t+1|t}$ . Then  $\xi_t | \xi_{t+1}, \Omega_t$  has same distribution as  $\xi_t | \xi_{t+1}^{**}, \Omega_t$ .

---

---

---

---

---

---

---

---

Generalization of previous results for singular  $\mathbf{P}_{t+1|t}$ :

$$\xi_{t|t}^* = \hat{\xi}_{t|t} + \mathbf{J}_t^{**} (\xi_{t+1}^{**} - \hat{\xi}_{t+1|t}^{**})$$

$$\mathbf{J}_t^{**} = \mathbf{P}_{t|t} (\mathbf{H}^{**} \mathbf{F})' \mathbf{P}_{t+1|t}^{** - 1}$$

$$\mathbf{P}_{t|t}^* = \mathbf{P}_{t|t} - \mathbf{J}_t^{**} \mathbf{H}^{**} \mathbf{F} \mathbf{P}_{t|t}$$

---

---

---

---

---

---

---

---

Next suppose that, in addition to  $\xi_{t+1}$ , we had also observed  $\mathbf{y}_{t+1}, \mathbf{y}_{t+2}, \dots, \mathbf{y}_T$ . This would contain no more information about  $\xi_t$  than was provided by  $\xi_{t+1}$  and  $\Omega_t$  alone:

$$\xi_t | \xi_{t+1}, \Omega_T \sim N(\xi_{t|t}^*, \mathbf{P}_{t|t}^*)$$

for the same  $\xi_{t|t}^*, \mathbf{P}_{t|t}^*$ .

---

---

---

---

---

---

---

---

And since

$$E(\xi_t | \xi_{t+1}, \Omega_T) = \hat{\xi}_{t|t} + \mathbf{J}_t^{**} (\xi_{t+1}^{**} - \hat{\xi}_{t+1|t}^{**}),$$

it follows from law of iterated expectations that

$$E(\xi_t | \Omega_T) = \hat{\xi}_{t|t} + \mathbf{J}_t^{**} (\hat{\xi}_{t+1|T}^{**} - \hat{\xi}_{t+1|t}^{**})$$

which we can calculate by iterating backwards for  $t = T-1, T-2, \dots$

---

---

---

---

---

---

---

---

Procedure to calculate smoothed inferences  $\{\hat{\xi}_{t|T}\}_{t=1}^T$ .

(1) Perform Kalman filter recursion and save the values of

$$\{\hat{\xi}_{t|t}, \hat{\xi}_{t+1|t}, \mathbf{P}_{t|t}, \mathbf{P}_{t+1|t}\}_{t=1}^T.$$

---

---

---

---

---

---

---

---

(2) Calculate

$$\mathbf{J}_t^{**} = \mathbf{P}_{t|t}(\mathbf{H}^{**}\mathbf{F})'(\mathbf{H}^{**}\mathbf{P}_{t+1|t}\mathbf{H}^{**'})^{-1}$$

for  $t = 1, 2, \dots, T-1$ , where  $\mathbf{H}^{**}$  is an  $(s \times r)$  matrix selecting the nonredundant elements of  $\xi_t$ .

---

---

---

---

---

---

---

---

(3) Calculate

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} + \mathbf{J}_t^{**}\mathbf{H}^{**}(\hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t})$$

for  $t = T-1$  where  $\hat{\xi}_{T-1|T-1}$ ,  $\hat{\xi}_{T|T}$ , and

$\hat{\xi}_{T|T-1}$  are all known from step (1).

---

---

---

---

---

---

---

---

(4) Evaluate

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} + \mathbf{J}_t^{**}\mathbf{H}^{**}(\hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t})$$

for  $t = T-2$  where right-hand variables are all known from step (3). Iterate

for  $t = T-3, T-4, \dots$

---

---

---

---

---

---

---

---

The MSE's of these smoothed inferences are given by

$$E(\xi_t - \hat{\xi}_{t|T})(\xi_t - \hat{\xi}_{t|T})' = \mathbf{P}_{t|T}$$

where  $\mathbf{P}_{t|T}$  can be found by iterating on

$$\mathbf{P}_{t|T} = \mathbf{P}_{t|t} + \mathbf{J}_t^{**} \mathbf{H}^{**} (\mathbf{P}_{t+1|T} - \mathbf{P}_{t+1|t}) \mathbf{H}^{**'} \mathbf{J}_t^{**'}$$

backward starting from  $t = T - 1$ .

---

---

---

---

---

---

---

---

### III. Linear state-space models

C. Using the Kalman filter

1. Estimating the unknown parameters
2. Forecasting
3. Smoothed inference
4. Time-varying parameters and missing observations

---

---

---

---

---

---

---

---

Suppose that  $\mathbf{F}, \mathbf{Q}, \mathbf{A}, \mathbf{H}, \mathbf{R}$  are known functions of  $t$  (or more generally, known functions of  $\mathbf{x}_t$ ):

$$\xi_{t+1} = \mathbf{F}_t \xi_t + \mathbf{v}_{t+1} \quad E(\mathbf{v}_{t+1} \mathbf{v}_{t+1}') = \mathbf{Q}_t$$

$$\mathbf{y}_t = \mathbf{A}_t' \mathbf{x}_t + \mathbf{H}_t' \xi_t + \mathbf{w}_t \quad E(\mathbf{w}_t \mathbf{w}_t') = \mathbf{R}_t$$

Then Kalman filter recursion immediately generalizes to:

---

---

---

---

---

---

---

---

$$\begin{aligned}
\mathbf{P}_{t+1|t} &= \mathbf{F}_t \mathbf{P}_{t|t} \mathbf{F}_t' + \mathbf{Q}_t \\
\mathbf{P}_{t+1|t+1} &= \mathbf{P}_{t+1|t} - \\
\mathbf{P}_{t+1|t} \mathbf{H}_{t+1} (\mathbf{H}_{t+1}' \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} + \mathbf{R}_{t+1})^{-1} \mathbf{H}_{t+1}' \mathbf{P}_{t+1|t} \\
\hat{\boldsymbol{\xi}}_{t+1|t} &= \mathbf{F}_t \hat{\boldsymbol{\xi}}_{t|t} \\
\hat{\boldsymbol{\epsilon}}_{t+1|t} &= \mathbf{y}_{t+1} - \mathbf{A}_{t+1}' \mathbf{x}_{t+1} - \mathbf{H}_{t+1}' \hat{\boldsymbol{\xi}}_{t+1|t} \\
\hat{\boldsymbol{\xi}}_{t+1|t+1} &= \hat{\boldsymbol{\xi}}_{t+1|t} + \\
\mathbf{P}_{t+1|t} \mathbf{H}_{t+1} (\mathbf{H}_{t+1}' \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} + \mathbf{R}_{t+1})^{-1} \hat{\boldsymbol{\epsilon}}_{t+1|t}
\end{aligned}$$

---

---

---

---

---

---

---

---

One simple trick for handling missing observations: if observation  $y_{it}$  is missing for date  $t$ , set  $i$ th rows of  $\mathbf{A}'_t$  and  $\mathbf{H}'_t$  to zero, take  $y_{it} = 0$ , set row  $i$ , col  $i$  of  $\mathbf{R}_t$  to 1 and all other elements of row  $i$  or col  $i$  of  $\mathbf{R}_t$  to zero.

---

---

---

---

---

---

---

---

Why it works: suppose for illustration the first  $r$  elements of  $\mathbf{y}_{t+1}$  are missing.

$$\begin{aligned}
\mathbf{A}'_{t+1} &= \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{A}}' \end{bmatrix} & \mathbf{H}'_{t+1} &= \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{H}}' \end{bmatrix} \\
\mathbf{R}_{t+1} &= \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{R}} \end{bmatrix}
\end{aligned}$$


---

---

---

---

---

---

---

---

Then

$$\mathbf{H}_{t+1} = \begin{bmatrix} \mathbf{0} & \tilde{\mathbf{H}} \end{bmatrix}$$

$$\mathbf{P}_{t+1|t} \mathbf{H}_{t+1} = \begin{bmatrix} \mathbf{0} & \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} \end{bmatrix}$$

$$\mathbf{H}_{t+1}' \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{H}}' \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} \end{bmatrix}$$

---

---

---

---

---

---

---

---

$$\begin{aligned} & \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} (\mathbf{H}_{t+1}' \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} + \tilde{\mathbf{R}})^{-1} \\ &= \begin{bmatrix} \mathbf{0} & \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} \end{bmatrix} \times \\ & \quad \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & (\tilde{\mathbf{H}}' \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} + \tilde{\mathbf{R}})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0} & \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} (\tilde{\mathbf{H}}' \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} + \tilde{\mathbf{R}})^{-1} \end{bmatrix} \end{aligned}$$

---

---

---

---

---

---

---

---

$$\begin{aligned} & \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} (\mathbf{H}_{t+1}' \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} + \tilde{\mathbf{R}})^{-1} \\ &= \begin{bmatrix} \mathbf{0} & \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} (\tilde{\mathbf{H}}' \mathbf{P}_{t+1|t} \tilde{\mathbf{H}} + \tilde{\mathbf{R}})^{-1} \end{bmatrix} \end{aligned}$$

$$\hat{\boldsymbol{\xi}}_{t+1|t+1} = \hat{\boldsymbol{\xi}}_{t+1|t} +$$

$$\mathbf{P}_{t+1|t} \mathbf{H}_{t+1} (\mathbf{H}_{t+1}' \mathbf{P}_{t+1|t} \mathbf{H}_{t+1} + \mathbf{R}_{t+1})^{-1} \hat{\boldsymbol{\epsilon}}_{t+1|t}$$

acts as if first  $r$  elements of  $\mathbf{y}_t$  weren't there

---

---

---

---

---

---

---

---



### III. Linear state-space models

- C. Using the Kalman filter
  - 1. Estimating the unknown parameters
  - 2. Forecasting
  - 3. Smoothed inference
  - 4. Time-varying parameters and missing observations
  - 5. Using mixed-frequency data as they arrive in real time

---

---

---

---

---

---

---

---

Practical problem for economic forecasters:

Different data are of different, asynchronous frequencies and are subsequently revised

Example: "Introducing the Euro-Sting: Short Term Indicator of Euro Area Growth", Maximo Camacho and Gabriel Perez-Quiros

---

---

---

---

---

---

---

---

Assumption: there is an unobserved scalar  $f_t$  representing the monthly growth rate of real economic activity.

$\mathbf{z}_t^h = (4 \times 1)$  vector of "hard" indicators of  $f_t$

---

---

---

---

---

---

---

---

$z_{1t}^h$  = industrial production growth  
 $z_{2t}^h$  = retail sales growth  
 $z_{3t}^h$  = new industrial orders growth  
 $z_{4t}^h$  = Euro area export growth  
 $z_{it}^h = k_i^h + \beta_i^h f_t + u_{it}^h$

---

---

---

---

---

---

---

---

$z_{it}^h = k_i^h + \beta_i^h f_t + u_{it}^h$   
 $f_t = a_1 f_{t-1} + a_2 f_{t-2} + \dots + a_6 f_{t-6} + \varepsilon_t^f$   
 $\varepsilon_t^f \sim N(0, 1)$   
 $u_{it}^h = c_{i1}^h u_{i,t-1}^h + c_{i2}^h u_{i,t-2}^h + \dots + c_{i6}^h u_{i,t-6}^h + \varepsilon_{it}^h$   
 $\varepsilon_{it}^h \sim N(0, \sigma_{ii}^2)$   
 $\mathbf{z}_t^h = \mathbf{k}^h + \beta^h f_t + \mathbf{u}_t^h$   
 $\mathbf{u}_t^h = \mathbf{C}_1^h \mathbf{u}_{t-1}^h + \mathbf{C}_2^h \mathbf{u}_{t-2}^h + \dots + \mathbf{C}_6^h \mathbf{u}_{t-6}^h + \boldsymbol{\varepsilon}_t^h$   
 $\boldsymbol{\xi}_t = (f_t, f_{t-1}, \dots, f_{t-5}, \mathbf{u}_t^h, \mathbf{u}_{t-1}^h, \dots, \mathbf{u}_{t-5}^h)'$

---

---

---

---

---

---

---

---

Also have some “soft” survey measures intended to reflect year-over-year growth

$z_{1t}^s$  = Belgium overall business indicator  
 $z_{2t}^s$  = Euro-zone economic sentiment  
 $z_{3t}^s$  = German IFO business climate  
 $z_{4t}^s$  = Euro manufacturing purchasing managers index  
 $z_{5t}^s$  = services PMI

---

---

---

---

---

---

---

---

$$z_{it}^s = k_i^s + \beta_i^s \sum_{j=0}^{11} f_{t-j} + u_{it}^s$$

$$u_{it}^s = c_{i1}^s u_{i,t-1}^s + c_{i2}^s u_{i,t-2}^s + \dots + c_{i,6}^s u_{i,t-6}^s + \varepsilon_{it}^s$$

---

---

---

---

---

---

---

---

$q_t$  = true monthly growth rate  
of real GDP in deviation  
from mean (not observed)

$$q_t = \frac{1}{3} \beta^q f_t + u_t^q$$

$$u_t^q = c_1^q u_{t-1}^q + c_2^q u_{t-2}^q + \dots + c_6^q u_{t-6}^q + \varepsilon_t^q$$

---

---

---

---

---

---

---

---

Every three months we do  
observe a second revision of  
quarterly GDP growth

$$y_t^2 = k^2 + \frac{1}{3} q_t + \frac{2}{3} q_{t-1} + q_{t-2}$$

$$+ \frac{2}{3} q_{t-3} + \frac{1}{3} q_{t-4}$$

---

---

---

---

---

---

---

---

40 days earlier a more preliminary first revision was available

$$y_t^1 = y_t^2 + e_{2t}$$

20 days before that the initial “flash” estimate of GDP was released

$$y_t^0 = y_t^1 + e_{1t}$$

---

---

---

---

---

---

---

---

Model also uses quarterly employment growth  $\ell_t$ .

Potential observation vector:

$$\mathbf{y}_t = (y_t^2, \mathbf{z}_t^h, \mathbf{z}_t^s, \ell_t, y_t^1, y_t^0)'$$

---

---

---

---

---

---

---

---

Potential observation vector:

$$\mathbf{y}_t = (y_t^2, \mathbf{z}_t^h, \mathbf{z}_t^s, \ell_t, y_t^1, y_t^0)'$$

In every month, some of these (e.g.,  $y_t^2$ ,  $\ell_t$ , and  $y_t^0$ ) are treated as missing observations

On any given day before the end of the month, a smaller subset is observed.

---

---

---

---

---

---

---

---

$$\xi_t = (f_t, f_{t-1}, \dots, f_{t-11}, u_t^q, u_{t-1}^q, \dots, u_{t-5}^q, \dots, u_t^{h'}, \dots, u_{t-5}^{h'}, u_t^{s'}, \dots, u_{t-5}^{s'}, u_t^l, \dots, u_{t-5}^l)'$$

Model allows forecast of any variable using all information available as of any day

---

---

---

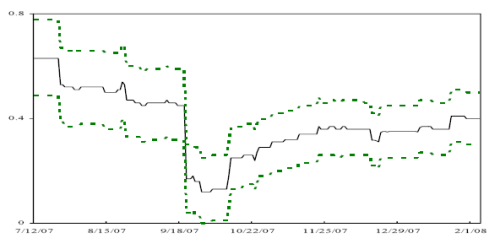
---

---

---

---

---



Real-time forecasts of 2007:Q4 real GDP growth from release of second revision on 2007/07/12 until 2008/02/13

---

---

---

---

---

---

---

---

### III. Linear state-space models

- A. State-space representation of a dynamic system
- B. Kalman filter
- C. Using the Kalman filter
- D. Bayesian analysis of linear state-space models

---

---

---

---

---

---

---

---

How do we obtain  $\theta$ , the vector containing unknown elements of  $F, Q, A, H, R$ ?

Classical approach: choose  $\hat{\theta}$  so as to maximize log likelihood  $\log p(\mathbf{Y}|\theta)$  (done by numerical search methods).

---

---

---

---

---

---

---

---

Asymptotic standard errors from

$$\hat{\theta} \approx N(\theta_0, \hat{\mathbf{C}})$$

$$\hat{\mathbf{C}} = \left[ -\frac{\partial^2 \log p(\mathbf{Y}|\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\hat{\theta}} \right]^{-1}$$

Parametric bootstrap for small-sample standard errors typically infeasible (requires separate numerical optimization for each Monte Carlo draw  $j$ ).

---

---

---

---

---

---

---

---

Analytical Bayesian results also unknown (problem:  $\xi_t$  is unobserved)

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1} \quad E(\mathbf{v}_t \mathbf{v}_t') = \mathbf{Q}$$

$$\mathbf{y}_t = \mathbf{A}'\mathbf{x}_t + \mathbf{H}'\xi_t + \mathbf{w}_t \quad E(\mathbf{w}_t \mathbf{w}_t') = \mathbf{R}$$

Solution: Gibbs sampler

---

---

---

---

---

---

---

---

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1} \quad E(\mathbf{v}_t\mathbf{v}_t') = \mathbf{Q}$$

$$\mathbf{y}_t = \mathbf{A}'\mathbf{x}_t + \mathbf{H}'\xi_t + \mathbf{w}_t \quad E(\mathbf{w}_t\mathbf{w}_t') = \mathbf{R}$$

Solution: Gibbs sampler

- $\theta_1$  = unknown elements of  $\mathbf{Q}, \mathbf{R}$
- $\theta_2$  = unknown elements of  $\mathbf{F}, \mathbf{A}, \mathbf{H}$
- $\theta_3$  = unknown elements of  $\{\xi_0, \xi_1, \dots, \xi_T\}$

---

---

---

---

---

---

---

---

$$\xi_{t+1} = \mathbf{F}(\theta_2)\xi_t + \mathbf{v}_{t+1} \quad E(\mathbf{v}_t\mathbf{v}_t') = \mathbf{Q}(\theta_1)$$

$$\mathbf{y}_t = \mathbf{A}(\theta_2)'\mathbf{x}_t + \mathbf{H}(\theta_2)'\xi_t + \mathbf{w}_t$$

$$E(\mathbf{w}_t\mathbf{w}_t') = \mathbf{R}(\theta_1)$$

$$(1) p(\theta_1, \theta_2, \theta_3, \mathbf{Y}, \mathbf{X})$$

Knowledge of  $\theta_2, \theta_3, \mathbf{Y}, \mathbf{X}$  is equivalent to direct observation of  $\{(\mathbf{v}_t', \mathbf{w}_t')'\}_{t=1}^T$ .

---

---

---

---

---

---

---

---

E.g., if priors for  $\mathbf{Q}^{-1}$  and  $\mathbf{R}^{-1}$  are independent  $W(N_Q, \Lambda_Q)$  and  $W(N_R, \Lambda_R)$ , respectively, then

$$\mathbf{Q}^{-1} | \theta_2, \theta_3, \mathbf{Y}, \mathbf{X} \sim W(N_Q + T, \Lambda_Q + \mathbf{S}_Q)$$

$$\mathbf{S}_Q = \sum_{t=1}^T \mathbf{v}_t\mathbf{v}_t'$$

$$\mathbf{R}^{-1} | \theta_2, \theta_3, \mathbf{Y}, \mathbf{X} \sim W(N_R + T, \Lambda_R + \mathbf{S}_R)$$

$$\mathbf{S}_R = \sum_{t=1}^T \mathbf{w}_t\mathbf{w}_t'$$

---

---

---

---

---

---

---

---

If  $\mathbf{Q}$  or  $\mathbf{R}$  are singular, apply similar idea to the nonsingular subset.  
 E.g., inverse of (1,1) element of  $\mathbf{Q}$  has prior  $\Gamma(N_Q, \lambda_Q)$  and posterior  $\Gamma(N_Q + T, \lambda_Q + S_Q)$ .

---

---

---

---

---

---

---

---

Given particular numerical values for  $\theta_2^{(j)}$  and  $\theta_3^{(j)} = ([\xi_0^{(j)}]', [\xi_1^{(j)}]', \dots, [\xi_T^{(j)}]')'$ , can generate values for  $\mathbf{Q}^{(j+1)}$  and  $\mathbf{R}^{(j+1)}$  from these distributions.

---

---

---

---

---

---

---

---

$$\xi_{t+1} = \mathbf{F}(\theta_2)\xi_t + \mathbf{v}_{t+1} \quad E(\mathbf{v}_t\mathbf{v}_t') = \mathbf{Q}(\theta_1)$$

$$\mathbf{y}_t = \mathbf{A}(\theta_2)'\mathbf{x}_t + \mathbf{H}(\theta_2)'\xi_t + \mathbf{w}_t$$

$$E(\mathbf{w}_t\mathbf{w}_t') = \mathbf{R}(\theta_1)$$

$$(2) p(\theta_2|\theta_1, \theta_3, \mathbf{Y}, \mathbf{X})$$

With knowledge of  $\theta_1, \theta_3, \mathbf{Y}, \mathbf{X}$ , this is a standard regression model.

---

---

---

---

---

---

---

---



E.g., if  $\mathbf{F}$  is unrestricted,

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1} \quad E(\mathbf{v}_t\mathbf{v}_t') = \mathbf{Q}(\theta_1)$$

$$\mathbf{f} = \text{vec}(\mathbf{F}')$$

$$\text{prior: } \mathbf{f}|\mathbf{Q} \sim N(\mathbf{m}_F, \mathbf{Q} \otimes \mathbf{M}_F)$$

$$\text{posterior: } \mathbf{f}|\mathbf{Y}, \mathbf{X}, \theta_1, \theta_3 \sim N(\mathbf{m}_F^*, \mathbf{Q} \otimes \mathbf{M}_F^*)$$

---

---

---

---

---

---

---

---

$$\mathbf{M}_F^* = \left( \mathbf{M}_F^{-1} + \sum_{t=1}^T \xi_{t-1} \xi_{t-1}' \right)^{-1}$$

$$\mathbf{m}_F^* = (\mathbf{I}_r \otimes \mathbf{M}_F^* \mathbf{M}_F^{-1}) \mathbf{m}_F + (\mathbf{I}_r \otimes \mathbf{M}_F^* \sum_{t=1}^T \xi_{t-1} \xi_{t-1}') \hat{\mathbf{f}}$$

$$\hat{\mathbf{f}} = \text{vec}(\hat{\mathbf{F}}')$$

$$\hat{\mathbf{F}}' = \left( \sum_{t=1}^T \xi_{t-1} \xi_{t-1}' \right)^{-1} \left( \sum_{t=1}^T \xi_{t-1} \xi_t' \right)$$

---

---

---

---

---

---

---

---

So, can generate  $\theta_2^{(j+1)}$  from

$$p(\theta_2|\theta_1^{(j+1)}, \theta_3^{(j)}, \mathbf{Y}, \mathbf{X}).$$

---

---

---

---

---

---

---

---

$$\xi_{t+1} = \mathbf{F}(\theta_2)\xi_t + \mathbf{v}_{t+1} \quad E(\mathbf{v}_t\mathbf{v}_t') = \mathbf{Q}(\theta_1)$$

$$\mathbf{y}_t = \mathbf{A}(\theta_2)' \mathbf{x}_t + \mathbf{H}(\theta_2)' \xi_t + \mathbf{w}_t$$

$$E(\mathbf{w}_t\mathbf{w}_t') = \mathbf{R}(\theta_1)$$

$$(3) p(\theta_3|\theta_1, \theta_2, \mathbf{Y}, \mathbf{X})$$

With knowledge of  $\theta_1, \theta_2, \mathbf{Y}, \mathbf{X}$ , this is a Kalman filter problem.

---

---

---

---

---

---

---

---

$$\xi_T|\theta_1, \theta_2, \mathbf{Y}, \mathbf{X} \sim N(\hat{\xi}_{T|T}, \mathbf{P}_{T|T})$$

$$\xi_{T-1}|\xi_T, \theta_1, \theta_2, \mathbf{Y}, \mathbf{X} \sim N(\xi_{T-1|T-1}^*, \mathbf{P}_{T-1|T-1}^*)$$

$$\xi_{T-1|T-1}^* = \hat{\xi}_{T-1|T-1} + \mathbf{J}_{T-1}(\xi_T - \hat{\xi}_{T|T})$$

$$\mathbf{J}_{T-1} = \mathbf{P}_{T-1|T-1} \mathbf{F}' \mathbf{P}_{T|T}^{-1}$$

$$\mathbf{P}_{T-1|T-1}^* = \mathbf{P}_{T-1|T-1} - \mathbf{J}_{T-1} \mathbf{F} \mathbf{P}_{T-1|T-1}$$

---

---

---

---

---

---

---

---

How do we generate an  $(r \times 1)$  vector  $\mathbf{q} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$ ?

If  $\boldsymbol{\Omega}$  is nonsingular, find Cholesky factorization  $\boldsymbol{\Omega} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}'$ . Generate  $\mathbf{u} \sim N(\mathbf{0}, \mathbf{I}_r)$  and  $\mathbf{q} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{u}$ .

---

---

---

---

---

---

---

---

If  $\Omega$  is singular, then some linear combinations of  $\mathbf{q}$  are known deterministically. Set these to their known values and generate rest of  $\mathbf{q}$  from non-redundant elements of distribution.

---

---

---

---

---

---

---

---

Specifically, let  $\mathbf{H}$  be a known nonsingular  $(r \times r)$  matrix such that

$$\mathbf{H}\Omega\mathbf{H}' = \begin{bmatrix} \Omega_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

for  $\Omega_1$  a nonsingular  $(s \times s)$  matrix ( $s < r$ ) with Cholesky factorization  $\Omega_1 = \Lambda_1\Lambda_1'$ .

---

---

---

---

---

---

---

---

Generate

$$\mathbf{u}_1 \sim N(\mathbf{0}, \mathbf{I}_s)$$

$$\mathbf{u} = \begin{bmatrix} \Lambda_1\mathbf{u}_1 \\ \mathbf{0} \end{bmatrix}$$

$$\mathbf{q} = \boldsymbol{\mu} + \mathbf{H}^{-1}\mathbf{u}$$

---

---

---

---

---

---

---

---

Generate  $\xi_t^{(j+1)}$  for  $t = T-1, T-2, \dots$

$$\xi_t^{(j+1)} | \xi_{t+1}^{(j+1)}, \xi_{t+2}^{(j+1)}, \dots, \xi_T^{(j+1)}, \theta_1^{(j+1)}, \theta_2^{(j+1)}, \mathbf{Y}, \mathbf{X} \\ \sim N(\xi_{t|t}^*, \mathbf{P}_{t|t}^*)$$

$$\xi_{t|t}^* = \hat{\xi}_{t|t} + \mathbf{J}_t^{**} \mathbf{H}^{**} (\xi_{t+1} - \hat{\xi}_{t+1|t})$$

$$\mathbf{J}_t^{**} = \mathbf{P}_{t|t} (\mathbf{H}^{**} \mathbf{F})' (\mathbf{H}^{**} \mathbf{P}_{t+1|t} \mathbf{H}^{**})^{-1}$$

$$\mathbf{P}_{t|t}^* = \mathbf{P}_{t|t} - \mathbf{J}_t^{**} \mathbf{H}^{**} \mathbf{F} \mathbf{P}_{t|t}$$

---

---

---

---

---

---

---

---

So, can generate  $\theta_3^{(j+1)}$  from  
 $p(\theta_3 | \theta_1^{(j+1)}, \theta_2^{(j+1)}, \mathbf{Y}, \mathbf{X})$ .

---

---

---

---

---

---

---

---

Gibbs sampler:

(1) Start with arbitrary initial guesses for  $\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}$  (e.g.,  $\theta_3^{(1)} = \mathbf{0}$ ).

(2) Generate:

$$\theta_1^{(j+1)} \text{ from } p(\theta_1 | \mathbf{Y}, \mathbf{X}, \theta_2^{(j)}, \theta_3^{(j)})$$

$$\theta_2^{(j+1)} \text{ from } p(\theta_2 | \mathbf{Y}, \mathbf{X}, \theta_1^{(j+1)}, \theta_3^{(j)})$$

$$\theta_3^{(j+1)} \text{ from } p(\theta_3 | \mathbf{Y}, \mathbf{X}, \theta_1^{(j+1)}, \theta_2^{(j+1)})$$

for  $j = 1, 2, \dots, D$

---

---

---

---

---

---

---

---

From these one can calculate such things as:

(1) Small-sample confidence intervals for elements of  $\mathbf{F}$  (e.g., for 95% of values of  $j$  between  $D_0$  and  $D$  the  $(1, 1)$  element of  $\mathbf{F}(\theta_2^{(j)})$  is between  $a_1$  and  $a_2$ ).

---

---

---

---

---

---

---

---

(2) Best guess as to state of business cycle  $C_t$  at historical date  $t$  (average value of  $(1, 1)$  element of  $\xi_t^{(j)}$  for  $j$  between  $D_0$  and  $D$ ) and uncertainty about this guess (standard deviation of  $\xi_t^{(j)}$ ), where uncertainty incorporates both filter uncertainty,

$$\mathbf{P}_{t|T}(\theta_1, \theta_2) \neq \mathbf{0},$$

and uncertainty about parameter values:

$\theta_1, \theta_2$  unknown.

---

---

---

---

---

---

---

---

(3) Optimal forecast  $\hat{\mathbf{y}}_{t+m|t}$ , incorporating uncertainty about parameter values, or average value of

$$[\mathbf{A}(\theta_2^{(j)})]' \mathbf{x}_{t+m} + [\mathbf{H}(\theta_2^{(j)})]' [\mathbf{F}(\theta_2^{(j)})]^m \hat{\xi}_{jt}^{(j)}$$

---

---

---

---

---

---

---

---