Inference for MS-DSGE Model

Francesco Bianchi

Johns Hopkins University

Estimating a MS-DSGE model

General form:

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{D}(\theta) + \mathbf{Z}\mathbf{s}_t + \mathbf{U}\mathbf{v}_t \\
\mathbf{s}_t &= \mathbf{c}\left(\xi_t^a, \theta, \mathbf{H}_p\right) + \mathbf{T}\left(\xi_t^p, \theta, \mathbf{H}_p\right)\mathbf{s}_{t-1} + \mathbf{R}\left(\xi_t^p, \theta, \mathbf{H}_p\right)\mathbf{Q}\left(\xi_t^v\right) ”_t \\
\mathbf{Q}\left(\xi_t^v\right) &= diag\left(\sigma_{R,\xi_t^v}, \sigma_{g,\xi_t^v}, \sigma_{a,\xi_t^v}\right), \; ”_t \sim N(0, I) \\
\mathbf{U} &= diag\left(\sigma_{GDP}, \sigma_{Infl}, \sigma_{FFR}\right), \; \mathbf{v}_t \sim N(0, I) \\
\xi_t^x &= 1 \ldots m^x, \; \mathbf{H}_{x,i,j} = p\left(\xi_t^x = i \middle| \xi_{t-1}^x = j\right) \text{ for } x = p, a, v.
\end{aligned}
$$

where v_t is a vector of observation errors and Y_t includes a set of observable variables.
In our examples:

$$
\mathbf{y}_t = \begin{bmatrix} GDP_t \\ INFL_t \\ FFR_t \end{bmatrix}, \mathbf{D}(\theta) = \begin{bmatrix} 0 \\ 4\pi \\ 4(\pi + r) \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

Estimating a MS-DSGE model

Under fixed coefficient, a Metropolis-Hastings algorithm can be used to make draws from the posterior, where the **Kalman filter** is used to evaluate the likelihood $\ell\left(\theta, M, \sigma_\varsigma | \mathbf{Y}^T\right)$.

In presence of regime changes:

1. The standard Kalman filter cannot be used: We need to infer the paths of the Markov chain $\left(\xi^T\right)$ and of the DSGE state vector $\left(\mathbf{S}^T\right)$
2. Possible methods:
    - Trimming approximation: Keep track of a limited number of paths (Schorfheide, 2005, Bianchi, 2009)
    - Kim's approximation: Number of paths is kept finite approximating the distribution of S^T (Bianchi 2013)
    - Gibbs sampling algorithm

Kim's approximation of the likelihood

Combine the MS states of the structural parameters and of the heteroskedastic shocks in a unique chain, ξ_t. ξ_t can assume m different values, with $m = m^p * m^a * m^v$, and evolves according to the transition matrix $\mathbf{H} = \mathbf{H}_p \otimes \mathbf{H}_a \otimes \mathbf{H}_v$. I will use the following notation:

$$\mathbf{c}_j = \mathbf{c}(\xi_t = j), \; \mathbf{T}_j = \mathbf{T}(\xi_t = j), \; \mathbf{Q}_j = \mathbf{Q}(\xi_t = j), \; \mathbf{R}_j = \mathbf{R}(\xi_t = j)$$

For a given set of parameters, and some assumptions about the initial DSGE state variables and MS latent variables, we can recursively run the following filter.

Step 1: Build forecasts for Y_t conditional on each possible regime path (ξ_{t-1}, ξ_t):

$$\begin{aligned} \mathbf{s}_{t|t-1}^{(i,j)} &= \mathbf{c}_j + \mathbf{T}_j^i \mathbf{s}_{t-1|t-1} \\ \mathbf{P}_{t|t-1}^{(i,j)} &= \mathbf{T}_j \mathbf{P}_{t-1|t-1}^i \mathbf{T}_j' + \mathbf{R}_j \mathbf{Q}_j^2 \mathbf{R}_j' \end{aligned}$$

$$\mathbf{e}_{t|t-1}^{(i,j)} = \mathbf{y}_t - \mathbf{D} - \mathbf{Z}\mathbf{s}_{t|t-1}^{(i,j)} \text{ and } \mathbf{f}_{t|t-1}^{(i,j)} = \mathbf{Z}\mathbf{P}_{t|t-1}^{(i,j)}\mathbf{Z}' + \mathbf{U}^2$$

$$\begin{aligned} \mathbf{s}_{t|t}^{(i,j)} &= \mathbf{s}_{t|t-1}^{(i,j)} + \mathbf{P}_{t|t-1}^{(i,j)} \mathbf{Z}' \left(\mathbf{f}_{t|t-1}^{(i,j)}\right)^{-1} \mathbf{e}_{t|t-1}^{(i,j)} \\ \mathbf{P}_{t|t}^{(i,j)} &= \mathbf{P}_{t|t-1}^{(i,j)} - \mathbf{P}_{t|t-1}^{(i,j)} \mathbf{Z}' \left(\mathbf{f}_{t|t-1}^{(i,j)}\right)^{-1} \mathbf{Z}\mathbf{P}_{t|t-1}^{(i,j)} \end{aligned}$$

Kim's approximation of the likelihood

Step 2: Compute the the likelihood density of observation Y_t is given by:

$$\ell\left(\mathbf{y}_t|\mathbf{Y}^{t-1}\right) = \sum_{j=1}^{m}\sum_{i=1}^{m} f\left(\mathbf{Y}_t|\xi_{t-1}=i,\xi_t=j,\mathbf{Y}^{t-1}\right) * \Pr\left[\xi_{t-1}=i,\xi_t=j|\mathbf{Y}^{t-1}\right]$$

$$f\left(\mathbf{y}_t|\xi_{t-1}=i,\xi_t=j,\mathbf{Y}^{t-1}\right) = \frac{\left|\mathbf{f}_{t|t-1}^{(i,j)}\right|^{-1/2}}{(2\pi)^{N/2}}\exp\left\{-\frac{1}{2}\mathbf{e}_{t|t-1}^{(i,j)\prime}\mathbf{f}_{t|t-1}^{(i,j)}\mathbf{e}_{t|t-1}^{(i,j)}\right\}$$

where $\mathbf{Y}^{t-1} \equiv \{\mathbf{y}_1, ..., \mathbf{y}_{t-1}\}$ and

$$\Pr\left[\xi_{t-1}=i,\xi_t=j|\mathbf{Y}^{t-1}\right] = \mathbf{H}_{ji}\Pr\left[\xi_{t-1}=i|\mathbf{Y}^{t-1}\right]$$

where $\mathbf{H}_{ji} = \Pr\left(\xi_t=j|\xi_{t-1}=i\right)$.

Kim's approximation of the likelihood

Step 3: Update probabilities:

$$\begin{aligned}\Pr\left(\xi_t,\xi_{t-1}|\mathbf{Y}^t\right) &= \frac{\ell\left(\mathbf{y}_t,\xi_t,\xi_{t-1}|\mathbf{Y}^{t-1}\right)}{\ell\left(\mathbf{y}_t|\mathbf{Y}^{t-1}\right)} \\ &= \frac{\ell\left(\mathbf{y}_t,\xi_t,\xi_{t-1}|\mathbf{Y}^{t-1}\right)\Pr\left(\xi_t,\xi_{t-1}|\mathbf{Y}^{t-1}\right)}{\ell\left(\mathbf{y}_t|\mathbf{Y}^{t-1}\right)}\end{aligned}$$

$$\Pr\left(\xi_t|\mathbf{Y}^t\right)=\sum_{i=1}^{m}\Pr\left(\xi_t,\xi_{t-1}=i|\mathbf{Y}^t\right)$$

Kim's approximation of the likelihood

Step 4: Collapse the $m \times m$ elements of $\mathbf{s}_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ into m elements which are represented by $\mathbf{s}_{t|t}^{j}$ and $P_{t|t}^{j}$:

$$\mathbf{s}_{t|t}^{j} = \frac{\sum_{i=1}^{m} \Pr\left[\xi_{t-1} = i, \xi_t = j | Y^t\right] \mathbf{s}_{t|t}^{(i,j)}}{\Pr\left[\xi_t = j | \mathbf{Y}^t\right]}$$

$$\mathbf{P}_{t|t}^{j} = \frac{\sum_{i=1}^{m} \Pr\left[\xi_{t-1} = i, \xi_t = j | \mathbf{Y}^t\right] \left(\mathbf{P}_{t|t}^{(i,j)} + \left(\mathbf{s}_{t|t}^{j} - \mathbf{s}_{t|t}^{(i,j)}\right)\left(\mathbf{s}_{t|t}^{j} - \mathbf{s}_{t|t}^{(i,j)}\right)'\right)}{\Pr\left[\xi_t = j | \mathbf{Y}^t\right]}$$

Step 5: If $t = T$, stop. Otherwise, go back to step 1.

Trimming approximation

- The idea is to keep track of a limited number of alternative paths for the Markov-switching states.
- Paths with low probability are trimmed or approximated using Kim's algorithm.
- Combine ξ_t^p, ξ_t^a, and ξ_t^v to obtain ξ_t. ξ_t can assume values from 1 to m, where $m = m_p * m_a * m_v$, and it evolves according to the transition matrix $\mathbf{H} = \mathbf{H}_p \otimes \mathbf{H}_a \otimes \mathbf{H}_v$.
- Suppose the algorithm has reached time t. From previous steps, we have a $((t-1) \times l_{t-1})$ matrix L containing the l_{t-1} retained paths, a vector L_p collecting the probabilities assigned to the different paths, and a $(n \times l_{t-1})$ matrix $L_{\mathbf{s}}$ and a $(n \times n \times l_{t-1})$ matrix L_P containing respectively means and covariance matrices of the DSGE state vector corresponding to each of the l_{t-1} paths.

Trimming approximation

The goal is to approximate the likelihood for time t, $\ell\left(\mathbf{y}_t|\mathbf{Y}^{t-1}\right)$ for a given a set of parameters:

1. $\forall\ i = 1...l_{t-1}$, $\forall\ j = 1...L\left(t-1, i\right) m$, compute a one-step-ahaed Kalman filter with $\mathbf{s}^i_{t-1|t-1} = L_{\mathbf{s}}\left(:, i\right)$ and $\mathbf{P}^i_{t-1|t-1} = L_{\mathbf{P}}\left(:, :, i\right)$. This will return $f\left(\mathbf{y}_t|\xi^{t-1} = i, \xi_t = j, Y_{t-1}\right)$, i.e. the probability of observing $\mathbf{y}_t$ given *history* i and $\xi_t = j$. At the end of this step we will have a total of $l_{t-1} * m$ possible histories that are stored in L'. $\forall\ i$ and $\forall\ j$ save $\widetilde{\mathbf{s}}^{(i,j)}_{t|t}$ and $\widetilde{\mathbf{P}}^{(i,j)}_{t|t}$ and store them in $L'_{\mathbf{s}}$ and $L'_{\mathbf{P}}$.
2. Compute the ex-ante probabilities for each of the $l_{t-1} * m$ possible paths using the transition matrix $\mathbf{H}$:

$$
\begin{aligned}
p_{t|t-1}\left(j, i\right) &= p_{t-1|t-1}\left(i\right) * \mathbf{H}\left(j, L\left(t-1, i\right)\right) \\
p_{t-1|t-1}\left(i\right) &= L_p\left(i\right)
\end{aligned}
$$

where $L\left(t-1, i\right)$ is the regime in place at time $t-1$ based on the *ith* history.

Trimming approximation

3. Compute the likelihood density of observation $\mathbf{y}_t$ as a weighted average of the conditional likelihoods:

$$f\left(\mathbf{y_t}|\mathbf{Y}^{t-1}\right) = \sum_{j=1}^{m} \sum_{i=1}^{l_t} p_{t|t-1}(j,i)\, f\left(\mathbf{y_t}|\xi^{t-1} = i, \xi_t = j, \mathbf{Y}^{t-1}\right)$$

4. Update the probabilities for the different paths:

$$\begin{aligned} \widetilde{p}_{t|t}\left(i'\right) &= \frac{p_{t|t-1}(j,i)\, f\left(\mathbf{y_t}|\xi^{t-1} = i, \xi_t = j, \mathbf{Y}^{t-1}\right)}{f\left(\mathbf{y_t}|\mathbf{Y}^{t-1}\right)} \\ i' &= 1 ... l_{t-1} * m \end{aligned}$$

and store them in L_p'.

Trimming approximation

5. Reorder L'_p in decreasing order and rearrange $L'_{\mathbf{s}}$, L'_P and L' accordingly. Retain l_t of the possible paths where $l_t = \min\{B, l\}$, where B is an arbitrary integer and $l > 0$ is such that

$$\sum_{i'=1}^{l} \widetilde{p}_{t|t}\left(i'\right) \geq tr$$

where $tr > 0$ is an arbitrary threshold (for example: $B = 100$, $tr = 0.99$). Update the matrices L_P, $L_{\mathbf{s}}$, and L:

$$L_{\mathbf{P}} = L'_{\mathbf{P}}\left(:,:,1:l_t\right),\ L_{\mathbf{s}} = L'_{\mathbf{s}}\left(:,1:l_t\right),\ L = L'\left(:,1:l_t\right)$$

6. Rescale the probabilities of the retained paths and update L_p:

$$L_p\left(i\right) = p_{t|t}\left(i\right) = \frac{\widetilde{p}_{t|t}\left(i\right)}{\sum_{j=1}^{l_t} \widetilde{p}_{t|t}\left(j\right)}, \quad i = 1...l_t$$

Note: Kim's approximation can be applied to the trimmed paths.

Gibbs sampling for a MS-DSGE: Heteroskedasticity

General form:

$$
\begin{aligned}
Y_t &= \mathbf{D}(\theta) + \mathbf{Z}\mathbf{s}_t + \mathbf{U}\mathbf{v}_t \\
\mathbf{s}_t &= \mathbf{T}(\theta)\,\mathbf{s}_{t-1} + \mathbf{R}(\theta)\,\mathbf{Q}\left(\xi_t^v\right)\text{”}_t \\
\mathbf{Q}_{\xi_t^v} &= diag\left(\sigma_{R,\xi_t^v}, \sigma_{g,\xi_t^v}, \sigma_{a,\xi_t^v}\right),\ \text{”}_t \sim N(0, \mathbf{I}) \\
\mathbf{U} &= diag\left(\sigma_{GDP}, \sigma_{Infl}, \sigma_{FFR}\right),\ \mathbf{v}_t \sim N(0, \mathbf{I}) \\
\xi_t^v &= 1...m^v,\ \mathbf{H}_{i,j}^v = p\left(\xi_t^v = i | \xi_{t-1}^v = j\right)
\end{aligned}
$$

where v_t is a vector of observation errors and Y_t includes a set of observable variables.
In our examples:

$$
\mathbf{y}_t = \begin{bmatrix} GDP_t \\ INFL_t \\ FFR_t \end{bmatrix}, \mathbf{D}(\theta) = \begin{bmatrix} 0 \\ 4\pi \\ 4(\pi + r) \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

- Assume a Dirichlet distribution for the columns of $\mathbf{H}^v$ and an inverse gamma prior for the elements of $\mathbf{Q}$ and $\mathbf{U}$.

Gibbs sampling for a MS-DSGE: Heteroskedasticity

While $n < n_{sim}$:

1. Given $\mathbf{S}_{n-1}^{T}$, $\mathbf{Q}_{\xi^{v},n-1}$, and $\mathbf{H}_{v,n-1}$, use Bayesian updating to get a filtered estimate of $\widetilde{\xi}_{n}^{v,T}$ and then draw a sequence for $\xi_{n}^{v,T}$ (see Kim and Nelson and Hamilton)
2. Given $\xi_{n}^{v,T}$, $\mathbf{H}_{v,n}$ can be drawn according to a Dirichlet distribution
3. Conditional on $\xi_{n}^{v,T}$, evaluate the likelihood of the state space form model using the Kalman filter. Draw θ_{n} by using a Metropolis-Hastings algorithm. This step also returns filtered estimates for $\widetilde{\mathbf{S}}_{n}^{T}$
4. Use a backward procedure to draw $\mathbf{S}_{n}^{T}$
5. Conditional on $\mathbf{S}_{n}^{T}$, model innovations are observable:
    - Conditional on $\xi_{n}^{v,T}$, draw each of the elements of $\mathbf{Q}_{\xi^{v},n}$ with an inverse gamma.
    - Draw each of the elements of $\mathbf{U}$ with an inverse gamma.

Gibbs sampling for a MS-DSGE: General case

General form:

$$\begin{array}{rcl} Y_t & = & \mathbf{D}(\theta) + \mathbf{Z}\mathbf{s}_t + \mathbf{U}v_t \\ \mathbf{s}_t & = & C\left(\xi_t^a, \theta, \mathbf{H}^a\right) + \mathbf{T}\left(\xi_t^p, \theta, \mathbf{H}^p\right)\mathbf{s}_{t-1} + \mathbf{R}\left(\xi_t^p, \theta, \mathbf{H}^p\right)\mathbf{Q}\left(\xi_t^v\right)\varepsilon_t \\ \mathbf{Q}_{\xi_t^v} & = & diag\left(\sigma_{R,\xi_t^v}, \sigma_{g,\xi_t^v}, \sigma_{a,\xi_t^v}\right), \ \varepsilon_t \sim N\left(0, \mathbf{I}\right) \\ \mathbf{U} & = & diag\left(\sigma_{GDP}, \sigma_{Infl}, \sigma_{FFR}\right), \ v_t \sim N\left(0, \mathbf{I}\right) \\ \xi_t^x & = & 1...m^{xx}, \ \mathbf{H}_{x,i,j} = p\left(\xi_t^x = i | \xi_{t-1}^x = j\right) \text{ for } x = p, a, v. \end{array}$$

where v_t is a vector of observation errors and Y_t includes a set of observable variables.
In our examples:

$$\mathbf{y}_t = \begin{bmatrix} GDP_t \\ INFL_t \\ FFR_t \end{bmatrix}, \mathbf{D}(\theta) = \begin{bmatrix} 0 \\ 4\pi \\ 4(\pi + r) \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- Assume a Dirichlet distribution for the columns of $\mathbf{H}_v$ and an inverse gamma prior for the elements of $\mathbf{Q}$ and $\mathbf{U}$.

Gibbs sampling for a MS-DSGE: General case

While $n < n_{sim}$:

1. Given $\mathbf{S}_{n-1}^{T}$, $\mathbf{Q}_{\xi^{v},n-1}$, and $\mathbf{H}_{v,n-1}$, use Bayesian updating to get a filtered estimate of $\widetilde{\xi}_{n}^{v,T}$ and then draw a sequence for $\xi_{n}^{v,T}$ (see Kim and Nelson and Hamilton)
2. Given $\xi_{n}^{v,T}$, $\mathbf{H}_{v,n}$ can be drawn according to a Dirichlet distribution
3. Conditional on $\xi_{n}^{v,T}$, draw $\mathbf{H}_{p,n}$, $\mathbf{H}_{a,n}$, and θ_n by using a Metropolis-Hastings algorithm. Evaluate the likelihood of the state space form model using a modified Kalman filter. This step also returns filtered estimates for the joint distribution of $\widetilde{\xi}_{n}^{p,T}$, $\widetilde{\xi}_{n}^{a,T}$, and $\widetilde{\mathbf{S}}_{n}^{T}$.
4. Use a backward procedure to draw $\xi_{n}^{p,T}$ and $\xi_{n}^{a,T}$.
5. Conditional on $\xi_{n}^{p,T}$, $\xi_{n}^{a,T}$ and θ_n, use a backward procedure to draw $\mathbf{S}_{n}^{T}$
6. Conditional on $\xi_{n}^{p,T}$, $\xi_{n}^{a,T}$, and $\mathbf{S}_{n}^{T}$, model innovations are observable:
   - Conditional on $\xi_{n}^{v,T}$, draw each of the elements of $\mathbf{Q}_{\xi^{v},n}$ with an inverse gamma.
   - Draw each of the elements of $\mathbf{U}$ with an inverse gamma.

Gibbs sampling for a MS-DSGE: Alternatives

A bit more work would allow us to get rid of the bottle-neck represented by the likelihood approximation

For example, we could draw a sequence for the regime sequences $\xi_n^{p,T}$ and $\xi_n^{a,T}$:

1. Likelihood can be computed by using the basic Kalman filter
2. A Metropolis-Hastings algorithm needs to be used to draw the regime sequences and the transition matrix.
3. There could be problems with convergence of the algorithm, especially if the chain is not initialized close to the posterior mode.

Counterfactual simulations

For each draw we can compute a sequence of shocks: ”^{T}

- Therefore, for each draw we can reconstruct a counterfactual path for the macroeconomic variables changing
    1. Regime sequence (for example, always hawkish)
    2. Policy makers' behavior (for example, more hawkish when hawkish)
    3. Agents' beliefs (agents can be more optimistic/pessimistic, other regimes can be introduced)

$$\begin{aligned} \mathbf{s}_t &= \mathbf{c}\left(\widehat{\xi}_t^a, \widehat{\theta}, \widehat{\mathbf{H}}^a\right) + \mathbf{T}\left(\widehat{\xi}_t^p, \widehat{\theta}, \widehat{\mathbf{H}}^p\right)\mathbf{s}_{t-1} \\ &\quad + \mathbf{R}\left(\widehat{\xi}_t^p, \widehat{\theta}, \widehat{\mathbf{H}}^p\right)\mathbf{Q}\left(\xi_t^v\right)\text{”}_t \end{aligned}$$

- Counterfactual simulations are robust to the Lucas' critique because model is re-solved taking into account the change in agents' information set

Shock decomposiiton

The law of motion for the DSGE state vector $\mathbf{s}_t$ is given by:

$$\underset{(n\times 1)}{\mathbf{s}_t} = \mathbf{c}_{\xi_t^a} + \underset{(n\times n)(n\times 1)}{\mathbf{T}_{\xi_t^p}\mathbf{s}_{t-1}} + \underset{(n\times k)(k\times 1)}{\mathbf{R}_{\xi_t^p}\mathbf{e}_t}$$

where $\mathbf{e}_t = \mathbf{Q}\left(\xi_t^v\right)\boldsymbol{"}_t$.
Let $\mathbf{s}_0$ denote the DSGE state vector at time 0. Then:

$$\begin{aligned}\mathbf{s}_t &= \sum_{x=0}^{t-1}\left[\prod_{v=0}^{x-1}\mathbf{T}_{\xi_{t-v}^p}\right]\mathbf{c}_{\xi_{t-x}^a} + \prod_{x=0}^{t-1}\mathbf{T}_{\xi_{t-x}^p}\mathbf{s}_0 \\ &+ \sum_{x=0}^{t-1}\left[\prod_{v=0}^{x-1}\mathbf{T}_{\xi_{t-v}^p}\right]\mathbf{R}_{\xi_{t-x}^p}\mathbf{e}_{t-x}\end{aligned}$$

where for $x = 0$ I assume $\prod_{v=0}^{-1}\mathbf{T}_{\xi_{t-v}^p} = \mathbf{I}$.

Shock decomposiiton

If we define $\overline{\mathbf{ee}}_t \equiv \left[1_{(1,n)} \otimes \mathbf{e}_t\right]'$, where $1_{(1,n)}$ is a $(1,n)$ vector with all the elements equal to 1, and the operator Γ that given a matrix X returns a column vector containing the sums of the elements of each row, we have:

$$\begin{aligned} \mathbf{s}_t &= \sum_{x=0}^{t-1} \left[\prod_{v=0}^{x-1} \mathbf{T}_{\xi_{t-v}^p}\right] C_{\xi_{t-x}^a} + \prod_{x=0}^{t-1} \mathbf{T}_{\xi_{t-x}} \mathbf{s}_0 \\ &\quad + \Gamma\left(\sum_{x=0}^{t-1} \left[\prod_{v=0}^{x-1} \mathbf{T}_{\xi_{t-v}^p}\right] \mathbf{R}_{\xi_{t-x}^p} \circ \overline{\mathbf{ee}}_{t-x}\right) \end{aligned}$$

where $\circ$ represents the Hadamard product, a binary operation that takes two matrices of the same dimensions and produces another matrix where each element *ij* is the product of elements *ij* of the original two matrices. At each point in time the cumulative contribution of the shock e_i to the *jth* variable in the vector $\mathbf{s}_t$ can then be obtained extracting the *ji* element of the matrix $\sum_{x=0}^{t-1} \left[\prod_{v=0}^{x-1} \mathbf{T}_{\xi_{t-v}^p}\right] \mathbf{R}_{\xi_{t-x}^p} \circ \overline{ee}_{t-x}$. Notice that in general a sequence of regimes will also affect the impact of the initial conditions $\mathbf{s}_0$.

Shock decomposiiton

The shock decomposition can be computed recursively:

1. Contribution of Gaussian shocks. For each shock k and a regime sequence ξ^T, we can compute:

$$\mathbf{s}_t^k = \mathbf{T}_{\xi_t^p} \mathbf{s}_{t-1}^k + \mathbf{R}_{\xi_t^p} \mathbf{e}_t^k$$

for $t = 1, ..., T$ and where $\mathbf{e}_t^k$ sets all shocks to zero except for the *kth* shock and $\mathbf{s}_0^k = 0$.

2. Contribution of the intercept:

$$\mathbf{s}_t^C = \mathbf{c}_{\xi_t^a} + \mathbf{T}_{\xi_t^p} \mathbf{s}_{t-1}^C$$

for $t = 1, ..., T$ and $\mathbf{s}_0^C = 0$.

3. Contribution of initial values interacted with the regime sequence:

$$\mathbf{s}_t^{\mathbf{s}_0} = \mathbf{T}_{\xi_t^p} \mathbf{s}_{t-1}^{\mathbf{s}_0}$$

for $t = 1, ..., T$ and $\mathbf{s}_0^{\mathbf{s}_0} = \mathbf{s}_0$.