Modèle état mesure

On considère le modèle suivant

$$\begin{cases} \underset{n\times 1}{Y_t} = \underset{n\times r}{H}\ \underset{r\times 1}{\xi_t} + \underset{n\times 1}{w_t} & \text{[équation de mesure]} \\ \underset{r\times 1}{\xi_{t+1}} = \underset{r\times r}{F}\ \underset{r\times 1}{\xi_t} + \underset{r\times 1}{v_{t+1}} & \text{[équation d'état]} \end{cases}$$

avec

$w_t \sim BB(0, R)$ erreurs de mesure

$v_t \sim BB(0, Q)$ innovations des états

$w_t \perp\!\!\!\perp v_s$

et on suppose de plus que l'état est non corrélé avec les erreurs ou les innovations

$\mathbb{E}[\xi_\tau v_t] = 0$ pour $\tau = t-1, t-2, \ldots, 1$

$\mathbb{E}[\xi_\tau w_t] = 0$ pour tout τ

On a aussi

$$\mathbb{E}[w_t y_\tau'] = \mathbb{E}[w_t (H \xi_\tau + w_\tau)'] = 0$$
$$\forall \tau = t-1, t-2, \dots$$

et $$\mathbb{E}[v_t y_\tau'] = \mathbb{E}[v_t (H \xi_\tau + w_\tau)'] = 0$$
$$\forall \tau = t-1, t-2, \dots$$

On pourrait considérer un modèle plus général avec
- des variables exogènes dans l'équation de mesure ou l'équation d'état.
- une corrélation non nulle entre les innovations des états et les erreurs de mesure
- des matrices H, F, R et Q variables dans le temps

<u>Mais</u> cette classe de processus stochastique englobe tout ce que nous avons déjà vu, il y a donc assez de généralité ici -

Exemple 1 Soit le processus AR(p)

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$$

avec $\varepsilon_t \sim iid(0, \sigma^2_\varepsilon)$. On peut réécrire ce modèle sous une forme état mesure

Équation d'état $(r = p)$

$$\underbrace{\begin{pmatrix} Y_{t+1} \\ Y_t \\ \vdots \\ Y_{t-p+2} \end{pmatrix}}_{\xi_{t+1}} = \underbrace{\begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_p \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix}}_{\xi_t} + \underbrace{\begin{pmatrix} \varepsilon_{t+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{v_{t+1}}$$

condition de stabilité
⇓
valeurs propres de F inférieures à 1 en module

$$\text{Lo } V[v_{t+1}] \equiv Q = \begin{pmatrix} \sigma^2_\varepsilon & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & 0 \end{pmatrix}$$

une matrice de rang $1 < r$

Équation de mesure $(n = 1)$

$$Y_t = \underbrace{(1, 0, \cdots, 0)}_{H} \underbrace{\begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix}}_{\xi_t} + \underbrace{0}_{w_t}$$

$\text{Lo } V[w_t] = R = 0$ une matrice nulle.

Exemple 2 Soit le processus MA(1)

$$Y_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

Équation d'état $(r=2)$

$$\underbrace{\begin{pmatrix} \varepsilon_{t+1} \\ \varepsilon_t \end{pmatrix}}_{\xi_{t+1}} = \underbrace{\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}}_{\xi_t} + \underbrace{\begin{pmatrix} \varepsilon_{t+1} \\ 0 \end{pmatrix}}_{v_{t+1}}$$

Équation de mesure $(n=1)$

$$Y_t = \underbrace{(1, \theta)}_{H} \xi_t + \underbrace{0}_{w_t}$$

Notons que la représentation état-mesure n'est pas nécessairement unique. Dans le cas du MA(1) nous pourrions poser :

$$\underbrace{\begin{pmatrix} \varepsilon_{t+1} + \theta \varepsilon_t \\ \theta \varepsilon_{t+1} \end{pmatrix}}_{\xi_{t+1}} = \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{F} \underbrace{\begin{pmatrix} \varepsilon_t + \theta \varepsilon_{t-1} \\ \theta \varepsilon_t \end{pmatrix}}_{\xi_t} + \underbrace{\begin{pmatrix} \varepsilon_{t+1} \\ \theta \varepsilon_{t+1} \end{pmatrix}}_{v_{t+1}}$$

et

$$Y_t = (1, 0) \, \xi_t$$

__Exemple 3__ Plus généralement, considérons le processus ARMA(p,q) :

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

où $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc de variance σ_ε^2.

En posant $r = \max\{p, q+1\}$, $\varphi_j = 0$ $\forall\, j > p$ et $\theta_j = 0$ $\forall\, j > q$, nous pouvons réécrire ce processus ARMA sous la forme :

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_r Y_{t-r} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_{r-1} \varepsilon_{t-r+1}$$

__Équation d'état__ $(r = \max\{p, q+1\})$

$$\xi_{t+1} = \underbrace{\begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_{r-1} & \varphi_r \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}}_{F} \xi_t + \underbrace{\begin{pmatrix} \varepsilon_{t+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{v_{t+1}}$$

Equation de mesure (n=1)

$$Y_t = \underbrace{(1, \theta_1, \ldots, \theta_{2-1})}_{H} \xi_t$$

Pour vérifier qu'il s'agit bien du même processus stochastique, notons que l'équation d'état nous dit que :

$$\xi_{2,t+1} = \xi_{1,t}$$

deuxième élément de ξ_{t+1}

premier élément de ξ_t

d'après la deuxième ligne, et plus généralement en observant les lignes suivantes :

$$\xi_{j,t+1} = \xi_{j-1,t} = \xi_{j-2,t-1} \ldots$$

Par récurrence, on a donc :

$$\xi_{j,t+1} = L^{j-1} \xi_{1,t+1}$$

La première ligne de l'équation d'état s'interprète donc comme :

$$\xi_{1,t+1} = \varphi_1 \xi_{1,t} + \varphi_2 \xi_{2,t} + \dots + \varphi_r \xi_{r,t} + \varepsilon_{t+1}$$

$$\Leftrightarrow \xi_{1,t+1} = \left(\varphi_1 + \varphi_2 L + \dots + \varphi_r L^{r-1}\right) \xi_{1,t} + \varepsilon_{t+1}$$

$$\Leftrightarrow \left(1 - \varphi_1 L - \dots - \varphi_r L^r\right) \xi_{1,t+1} = \varepsilon_{t+1}$$

nous retrouvons ici la partie autoregressive du processus ARMA. L'équation de mesure nous dit que :

$$Y_t = \left(1 + \theta_1 L + \dots + \theta_{r-1} L^{r-1}\right) \xi_{1t}$$

en multipliant les deux membres par le polynôme retard associé à la partie autoregressive, il vient

$$\left(1 - \varphi_1 L - \dots \varphi_r L^r\right) Y_t = \left(1 + \theta_1 L + \dots + \theta_{r-1} L^{r-1}\right)\left(1 - \varphi_1 L - \dots - \varphi_r L^r\right) \xi_{1t}$$

$$\Leftrightarrow \left(1 - \varphi_1 L - \dots - \varphi_r L^r\right) Y_t = \left(1 + \theta_1 L + \dots + \theta_{r-1} L^{r-1}\right) \varepsilon_t$$

$$\Leftrightarrow Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_r Y_{t-r} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_{r-1} \varepsilon_{t-r+1}$$

il s'agit bien de notre processus ARMA(p,q).

■

Au delà des processus ARMA, la représentation état mesure peut être employée pour caractériser d'autres processus stochastiques. Cette représentation est par exemple souvent employée pour les décompositions Tendance/Cycle

$$Y_t = X_t + Z_t$$

$$X_t = X_{t-1} + \varepsilon_t \qquad \text{[Tendance]}$$

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \eta_t \quad \text{[Cycle]}$$

$$\xi_{t+1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \varphi_1 & \varphi_2 \\ 0 & 1 & 0 \end{pmatrix} \xi_t + \begin{pmatrix} \varepsilon_t \\ \eta_t \\ 0 \end{pmatrix}$$

$$Y_t = (1,0,0)\, \xi_t$$

▶ Dans un modèle état-mesure la (les) variable(s) est (sont) déterminé(e)(s) par une (des) combinaison(s) linéaire(s) de variables cachées (latentes) : les variables d'état.

Pour faire des prévisions sur une variable observée, il faut faire des prédictions sur les variables cachées. En principe, étant donnée la forme de l'équation d'état (autoregressive d'ordre 1) cela ne devrait pas être compliqué... Sauf que nous n'observons pas les variables d'état (ξ_t). Il faut donc révéler ces variables. Le filtre de Kalman est un algorithme récursif d'inference dont le but est de dévoiler les variables latentes à partir d'un échantillon.

Filtre de Kalman

- On observe un échantillon $\mathcal{Y}_T = \{y_1, y_2, \dots, y_T\}$
- On suppose que les matrices H, F, R et Q du modèle état-mesure sont connues (nous pourrions les estimer par MV).
- À partir de ces informations on souhaite révéler les variables latentes ξ_t.

Le filtre de Kalman est un algorithme récursif pour calculer des prévisions des variables latentes en $t+1$ sachant l'information disponible en t :

l'échantillon jusqu'à la date t

$$\hat{\xi}_{t+1|t} = \mathbb{E}\left[\xi_{t+1} \mid \mathcal{Y}_t\right]$$

L'algorithme génère une chronique de prévisions :

$$\hat{\xi}_{1|0}, \hat{\xi}_{2|1}, \dots, \hat{\xi}_{T|T-1}$$

ainsi qu'une suite associée de matrices mesurant l'incertitude attachée à ces prévisions

$$P_{1|0}, P_{2|1}, \dots, P_{T|T-1}$$

où

$$P_{t+1|t} = \mathbb{E}\left[\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)' \right]$$

⊛ Initialisation du filtre

La prévision $\hat{\xi}_{1|0}$ ne repose sur aucune observation. On pose :

$$\hat{\xi}_{1|0} = \mathbb{E}[\xi_1]$$

l'espérance inconditionnelle du vecteur d'état.

L'erreur quadratique moyenne associée est

$$P_{1|0} = \mathbb{E}\left[\left(\xi_1 - \mathbb{E}[\xi_1]\right)\left(\xi_1 - \mathbb{E}[\xi_1]\right)' \right]$$

Si toutes les valeurs propres de F sont inférieures à 1 en module, l'espérance inconditionnelle de ξ_t est solution de

$$\mathbb{E}[\xi_{t+1}] = F\,\mathbb{E}[\xi_t]$$

En supposant que $\{\xi_t, t \in \mathbb{Z}\}$ est stationnaire au second ordre, on doit donc avoir :

$$(I_r - F)\,\mathbb{E}[\xi_t] = 0$$

Puisque 1 n'est pas valeur propre de F, la matrice $I_r - F$ est de plein rang et $\mathbb{E}[\xi_t]$ est déterminé de façon unique :

$$\mathbb{E}[\xi_t] = \underset{r\times 1}{0}$$

La variance inconditionnelle du vecteur des variables latentes vérifie

$$\mathbb{E}[\xi_{t+1}\xi_{t+1}'] = \mathbb{E}\left[(F\xi_t + v_{t+1})(F\xi_t + v_{t+1})'\right]$$

$$\Leftrightarrow \mathbb{E}[\xi_{t+1}\xi_{t+1}'] = F\,\mathbb{E}[\xi_t\xi_t']F' + \mathbb{E}[v_{t+1}v_{t+1}']$$

$$\Leftrightarrow \Sigma = F\Sigma F' + Q$$

admet une unique solution finie si les valeurs propres de F sont < 1 en module.

$$\text{Lo} \quad \boxed{\operatorname{vec}(\Sigma) = (I_{r^2} - F\otimes F)^{-1}\operatorname{vec}(Q)}$$

Si les valeurs propres de F sont à l'intérieur du cercle unité, on pose donc :

- $\hat{\xi}_{1|0} = 0$
- $P_{1|0}$ tq $\operatorname{vec}(P_{1|0}) = (I_{r^2} - F\otimes F)^{-1}\operatorname{vec}(Q)$

⊛ Prédire y_t

Notons $\hat{y}_{t|t-1} = \mathbb{E}[y_t | \mathcal{Y}_{t-1}]$ la prévision sur les observables en t sachant l'information disponible en $t-1$. On a :

$$\hat{y}_{t|t-1} = \mathbb{E}[y_t | \mathcal{Y}_{t-1}]$$

$$= \mathbb{E}[H \xi_t + v_t | \mathcal{Y}_{t-1}]$$

$$= H \mathbb{E}[\xi_t | \mathcal{Y}_{t-1}] + \mathbb{E}[v_t | \mathcal{Y}_{t-1}]$$

$\hookrightarrow$ $$\boxed{\hat{y}_{t|t-1} = H \hat{\xi}_{t|t-1}}$$

Pour prédire la variable observable il faut prédire les variables latentes. Plus généralement on peut montrer que

$$\hat{y}_{t+p|t} = H \hat{\xi}_{t+p|t}$$

L'erreur de prévision :

$$Y_t - \hat{Y}_{t|t-1} = H\xi_t + w_t - H\hat{\xi}_{t|t-1}$$

$$\Leftrightarrow Y_t - \hat{Y}_{t|t-1} = H\underbrace{\left(\xi_t - \hat{\xi}_{t|t-1}\right)}_{\text{erreur de prévision sur les variables latentes.}} + w_t$$

erreur de mesure.

L'erreur quadratique moyenne associée à la prévision est (la variance de l'erreur de prévision) :

$$\mathbb{E}\left[\left(Y_t - \hat{Y}_{t|t-1}\right)\left(Y_t - \hat{Y}_{t|t-1}\right)'\right] = H\,\mathbb{E}\left[\left(\xi_t - \hat{\xi}_{t|t-1}\right)\left(\xi_t - \hat{\xi}_{t|t-1}\right)'\right]H' + \mathbb{E}\left[w_t w_t'\right]$$

où les termes croisés n'apparaissent pas car

$$\mathbb{E}\left[w_t\left(\xi_t - \hat{\xi}_{t|t-1}\right)'\right] = 0$$

La variance de l'erreur de prévision est donc :

$$\mathbb{E}\left[(Y_t - \hat{Y}_{t|t-1})(Y_t - \hat{Y}_{t|t-1})'\right] = H P_{t|t-1} H' + R$$

où $P_{t|t-1}$ est l'erreur quadratique moyenne associée à la prévisions des variables latentes.

⊛ <u>Comment prédire ξ_{t+1} ?</u>

La prévision des variables latentes est

$$\hat{\xi}_{t+1|t} = \mathbb{E}\left[\xi_{t+1} \mid \mathcal{Y}_t\right]$$

$$\Leftrightarrow \hat{\xi}_{t+1|t} = F\,\mathbb{E}\left[\xi_t \mid \mathcal{Y}_t\right] + \mathbb{E}\left[v_{t+1} \mid \mathcal{Y}_t\right]$$

$$\Leftrightarrow \boxed{\hat{\xi}_{t+1|t} = F\,\hat{\xi}_{t|t}}$$

Pour calculer l'espérance de ξ_{t+1}

sachant l'échantillon jusqu'à la date t, il faut déjà connaître l'espérance de ξ_t sachant l'échantillon jusqu'à la date t... Supposons que l'on connaisse $\hat{\xi}_{t|t-1}$ (cf. l'initialisat° de l'algorithme), on veut mettre à jour cette information sur ξ_t quand une nouvelle observation (y_t) apparaît $\to$ Calculer $\hat{\xi}_{t|t}$ en partant de $\hat{\xi}_{t|t-1}$ avec l'information apportée par la nouvelle observation y_t.

⊛ <u>Mise à jour des moyennes sur ξ_t</u>

$\hat{\xi}_{t|t-1}$ résume ce que nous pouvons dire du vecteur d'état en t sachant l'échantillon jusqu'à la date t-1 ($\mathcal{Y}_{t-1}$). On veut mettre à jour l'inférence sur ξ_t quand une nouvelle information y_t nous parvient, i.e. calculer $\hat{\xi}_{t|t} = \mathbb{E}[\xi_t | Y_t]$. On peut montrer que :

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + \underbrace{P_{t|t-1} H'}_{\text{covariance entre les latentes et observables}} \underbrace{\left(H P_{t|t-1} H' + R\right)^{-1}}_{\text{inverse de la variance de l'erreur de prévision}} \underbrace{\left(y_t - H \hat{\xi}_{t|t-1}\right)}_{\text{erreur de prévision}}$$

$\hat{\xi}_{t|t} > \hat{\xi}_{t|t-1}$ ssi la prévision sous estime la réalisation

en utilisant les propriétés de la loi normale multivariée, et que l'erreur quadratique moyenne associée est :

$$P_{t|t} = \underset{\gg 0}{P_{t|t-1}} - \underset{\gg 0}{P_{t|t-1}} H' \underset{\gg 0}{\left(H P_{t|t-1} H' + R\right)^{-1}} H \underset{\gg 0}{P_{t|t-1}}$$

la variance de l'estimateur des variables latente diminue d'autant plus que la variance des erreurs de prévision est faible.

* Prévoir ξ_{t+1}

On sait déjà que

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

en substituant la règle de mise à jour des croyances sur ξ, il vient :

$$\hat{\xi}_{t+1|t} = F \hat{\xi}_{t|t-1} + F P_{t|t-1} H' \left(H P_{t|t-1} H' + R \right)^{-1} \left(y_t - H \hat{\xi}_{t|t-1} \right)$$

En notant

$$K_t = F P_{t|t-1} H' \left(H P_{t|t-1} H' + R \right)$$

le gain du filtre de Kalman (qui quantifie ce que nous apprenons des erreurs de prévision), il vient :

$$\boxed{\hat{\xi}_{t+1|t} = F \hat{\xi}_{t|t-1} + K_t \left(y_t - H \hat{\xi}_{t|t-1} \right)}$$

La variance de cette prévision est :

$$P_{t+1|t} = \mathbb{E}\left[\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)'\right]$$

$$\Leftrightarrow \quad P_{t+1|t} = \mathbb{E}\left[\left(F\xi_t + v_t - F\hat{\xi}_{t|t}\right)\left(F\xi_t + v_t - F\hat{\xi}_{t|t}\right)'\right]$$

$$\Leftrightarrow \quad P_{t+1|t} = F\,\mathbb{E}\left[\left(\xi_t - \hat{\xi}_{t|t}\right)\left(\xi_t - \hat{\xi}_{t|t}\right)'\right]F' + Q$$

$$\Leftrightarrow \quad \boxed{P_{t+1|t} = F P_{t|t} F' + Q}$$

Finalement, en substituant l'expression pour $P_{t|t}$, on obtient :

$$\boxed{P_{t+1|t} = F\left(P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1}HP_{t|t-1}\right)F' + Q}$$

l'équation de Ricatti

Résumé du Filtre de Kalman

moments inconditionnels d'ordre 1 & 2

$$\text{Initialisation} \begin{cases} \hat{\xi}_{1|0} = \mathbb{E}[\xi_t] \\ \hat{P}_{1|0} = \mathbb{V}[\xi_t] \end{cases}$$

$$\begin{cases} \bullet\ K_t = F P_{t|t-1} H' \left(H P_{t|t-1} H' + R\right)^{-1} \\ \bullet\ \hat{\xi}_{t+1|t} = F \hat{\xi}_{t|t-1} + K_t \left(y_t - H \hat{\xi}_{t|t-1}\right) \\ \bullet\ P_{t+1|t} = F P_{t|t-1} F' - K_t H P_{t|t-1} + Q \end{cases}$$

boucle pour $t = 1, \ldots, T$

→ Cet algorithme produit aussi des prévisions pour les observables

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

→ À partir des erreurs de prévisions sur les observables, on peut construire la fonct° de vraisemblance [→ estimat°]