



Théorie de système et séries temporelles

Zaka Ratsimalahelo

► To cite this version:

Zaka Ratsimalahelo. Théorie de système et séries temporelles. [Rapport de recherche] Laboratoire d'analyse et de techniques économiques(LATEC). 1994, 43 p., ref. bib.: 1 p. 1/2. hal-01527144

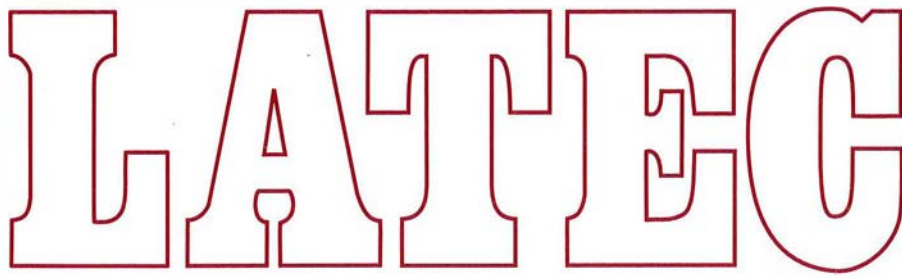
HAL Id: hal-01527144

<https://hal.science/hal-01527144>

Submitted on 24 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**LABORATOIRE D'ANALYSE
ET DE TECHNIQUES ÉCONOMIQUES**

UMR 5601 CNRS

DOCUMENT DE TRAVAIL



Pôle d'Economie et de Gestion

2, boulevard Gabriel - 21000 DIJON - Tél. 03 80 39 54 30 - Fax 03 80 39 54 43

ISSN : 1260-8556

n° 9401



THEORIE DE SYSTEME ET SERIES TEMPORELLES

Zaka RATSIMALAHELO*

juin 1994

** LATEC, Université de Bourgogne et Université de Franche-Comté*

Théorie de système et séries temporelles.

Zaka Ratsimalahelo

Résumé.

L'objectif principal de cet article est de présenter une représentation alternative du modèle espace état appelé modèle innovation, ce modèle est relativement nouveau et est encore inconnu de la littérature économique et économétrique. Nous présentons également ses propriétés.

La représentation état présente un intérêt considérable pour le modèle de séries temporelles, l'approche que nous développons dans cet article couvre une large classe des modèles dont l'avantage est qu'il n'impose aucune décomposition a priori des séries macroéconomiques en tendance et cycle.

Mots-clés.

Séries temporelles, innovation, espace état, stabilité stochastique.

Abstract.

The aim of this paper is to present a different representation of state space models, (innovation state space representation) which is relatively new and apparently unknown in the economics and econometrics literature and to describe some of its properties.

A state space representation is a very flexible form for time series and the approach taken in this paper therefore allows a broad class of models it does not impose a priori the decomposition of data series into trend and cycle.

Keyword.

Time series, innovation, state space, stability stochastic.

Théorie de système et séries temporelles.

1. Introduction.
2. Notion de variables d'états et représentation d'état-mesure.
 - 2.1. Discussion d'un exemple.
 - 2.2. Quelques modélisations de séries macroéconomiques.
3. Formulation générale.
 - 3.1. Propriétés des solutions explicites.
 - 3.2. Estimation du vecteur d'état.
 - 3.3. Hypothèses non gaussiennes.
4. Forme filtre ou modèle innovation.
 - 4.1. Propriétés des solutions explicites.
 - 4.2. Comparaison de la propriété statistique du modèle innovation et du modèle général.
 - 4.3. Relation entre la réponse impulsionnelle et la matrice de covariance.
5. Analyse de la stabilité stochastique.
 - 5.1. Notions de commandabilité et d'observabilité stochastiques.
 - 5.2. Théorème.
6. Les propriétés du modèle innovation.
 - 6.1. Détermination des pôles et zéros de la fonction de transfert.
 - 6.2. Représentation spectrale du modèle innovation.
7. Transformation des processus quelconques en représentation d'état.
 - 7.1. Cas modèle AR(p).
 - 7.2. Cas modèle autorégressif à paramètres aléatoires.
 - 7.3. Cas modèle ARMA(p,q).
 - a) Modèle innovation.
 - b) Forme canonique commandable minimale.
 - c) Forme canonique observable minimale.

1. Introduction.

Ces dernières années, nous avons assisté à un renouvellement profond des analyses descriptives du mouvement macroéconomique sous l'influence du développement des méthodes économétriques des séries temporelles.

Dans les travaux de macroéconomie appliquée, la décomposition des principales séries d'activité comme le PNB en une composante tendancielle et un écart conjoncturel représentait une pratique générale. Le concept de tendance stochastique remet en cause la distinction traditionnelle entre cycle et croissance.

Il est largement reconnu qu'une des limites essentielles de la décomposition de Beveridge et Nelson (1981) est d'être définie pour des modèles ARIMA ce qui suppose une parfaite corrélation des chocs transitoires et permanents. Pour lever cette hypothèse nombreux auteurs ont utilisé les modèles à composantes non observées ou la représentation d'état.

En effet des controverses se sont établies sur l'utilisation de ces modèles de représentation d'état, ces controverses sont dues essentiellement à des divergences de résultats entre Nelson (1988) d'une part et Clark (1987), Campbell et Mankiw (1987) Watson (1986) d'autre part. Il est important de souligner que pour un modèle de séries temporelles quelconque il existe plusieurs représentations du modèle espace état.

Notre objectif est donc d'apporter des éléments des réponses à ces controverses c'est-à-dire de présenter une représentation alternative de modèle espace état appelé modèle innovation ou forme filtre. Le modèle innovation est encore inconnu de la littérature économique et économétrique et pourtant il présente un intérêt considérable dans la pratique grâce à ses propriétés statistiques.

2. Notion de variables d'états et représentation d'état-mesure.

La modélisation d'espace d'état offre une structure et une interprétation simple d'un processus quelconque. Il est utile, au départ, de l'illustrer par un exemple simple.

2.1. Discussion d'un exemple.

Considérons, d'abord, le processus déterministe y_t $t \in \mathbb{Z}$ suivant :

$$(1 - L)^2 y_t = y_t - 2y_{t-1} + y_{t-2} = 0 \quad (2.1)$$

$$\Leftrightarrow y_t - y_{t-1} = y_{t-1} - y_{t-2}$$

$$\Leftrightarrow \Delta y_t = \Delta y_{t-1} \quad (\text{accroissement constant})$$

Il est facile de voir que ce processus est équivalent au système suivant:

$$x_{1,t} = x_{1,t-1} + x_{2,t} \quad (2.2.a)$$

$$x_{2,t} = x_{2,t-1} \quad (2.2.b)$$

$$y_t = x_{1,t} \quad (2.2.c)$$

Quel est l'intérêt de cette représentation ?

Elle permet de spécifier le niveau du processus y_t par (2.2.c) et sa pente dans l'intervalle $(t-1, t)$ par (2.2.b) :

$(y_t - y_{t-1} = x_{1,t} - x_{1,t-1} = x_{2,t})$ deux grandeurs qui ne sont pas directement visibles dans (2.1).

D'après l'équation (2.2.a), le niveau du processus de la période courante est égal à la somme du niveau de la période précédente et de la pente de la période courante.

La représentation matricielle du système (2.2) est

$$X_t = AX_{t-1} \quad (2.3.a)$$

$$y_t = CX_t \quad (2.3.b)$$

où :

$$X'_t = (x_{1,t}, x_{2,t})$$

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} ; \quad C = (1, 0)$$

La première équation (2.3.a) est appelée équation de transition ou d'évolution du système, le vecteur X_t est le vecteur d'état et la matrice A est la matrice d'évolution ou matrice d'état. La deuxième équation (2.3.b) est appelée équation d'observations (ou de sortie), C étant la matrice de sortie.

Considérons maintenant un cadre stochastique. La représentation (2.2) devient :

$$x_{1,t} = x_{1,t-1} + x_{2,t} + v_{1,t} \quad (2.4.a)$$

$$x_{2,t} = x_{2,t-1} + v_{2,t} \quad (2.4.b)$$

$$y_t = x_{1,t} + u_t \quad (2.4.c)$$

Les variables d'état $x_{1,t}$ et $x_{2,t}$ représentent respectivement le niveau et la pente stochastiques du processus y_t , le vecteur

$X'_t = (x_{1,t}, x_{2,t})$ est non stationnaire.

$V'_t = [v_{1t}, v_{2t}]$ est un bruit blanc de variance Q

avec
$$Q = \begin{pmatrix} \sigma_{v1}^2 & \sigma_{v1v2} \\ \sigma_{v2v1} & \sigma_{v2}^2 \end{pmatrix}$$

u_t est l'erreur de mesure (ou d'observations) est un bruit blanc (de variance σ_u^2).

Sa représentation sous forme matricielle est :

$$X_t = AX_{t-1} + BV_t \quad (2.5.a)$$

$$Y_t = CX_t + u_t \quad (2.5.b)$$

où :

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ est la matrice d'entrée aléatoire,}$$

$V'_t = (v_{1,t}, v_{2,t})$ est le vecteur de bruit de l'état,

u_t est le bruit de mesure.

Cet ensemble d'équations (2.5) est appelé représentation d'état d'un système dynamique linéaire appelé aussi représentation markovienne.

On peut montrer que le système (2.5) correspond à un processus ARIMA (0,2,2) ou IMA (2,2) que l'on appelle également selon la terminologie de Granger intégré d'ordre deux, noté I(2). Utilisons (2.4.c):

$$\begin{aligned}(1-L)y_t &= (1-L)x_{1t} + (1-L)u_t \\ &= x_{2t} + v_{1t} + (1-L)u_t\end{aligned}$$

d'où :

$$\begin{aligned}(1-L)^2 y_t &= (1-L)x_{2t} + (1-L)v_{1t} + (1-L)^2 u_t \\ &= v_{2t} + (1-L)v_{1t} + (1-L)^2 u_t\end{aligned}\quad (2.6)$$

En posant $z_t = (1-L)^2 y_t$, le calcul de la fonction d'autocovariance de $(1-L)^2 y_t$ donne :

$$E(z_t z'_t) = \begin{cases} \sigma_{v2}^2 + 2\sigma_{v1}^2 + 2\sigma_{v1v2}^2 + 6\sigma_u^2 & \text{si } s = t \\ -\sigma_{v1}^2 - \sigma_{v1v2} - 4\sigma_u^2 & \text{si } |s - t| = 1 \\ \sigma_u^2 & \text{si } |s - t| = 2 \\ 0 & \text{si } |s - t| > 2 \end{cases} \quad (2.7)$$

Le processus y_t peut, alors, être exprimé par

$$(1-L)^2 y_t = (1 - a_1 L - a_2 L^2) \varepsilon_t \quad (2.8)$$

avec ε_t est i.i.d. $(0, \sigma^2)$

où σ^2 , a_1 et a_2 sont fonction de σ_{v1}^2 , σ_{v2}^2 , σ_{v1v2} et σ_u^2

A partir du modèle (2.4) on peut retrouver la plupart des modélisations des séries macroéconomiques suivant l'hypothèse faite sur les bruits de l'état de et de l'observation.

2.2. Quelques modélisations de séries macroéconomiques.

La plupart des travaux de macroéconomie pratiquent la décomposition des séries d'activités, comme le PNB, en une composante tendancielle et une composante cyclique, Harvey (1985) Clark (1987, 1989) Aoki (1987, 1988) Cochrane J. H (1988) Nelson (1988) Stock et Watson (1988). On considère, par exemple, le trend comme une fonction déterministe du temps et la composante cyclique comme un processus stationnaire. Plusieurs cas peuvent être envisagé.

1er cas: le bruit d'état v_{2t} est nul.

Supposons que $\sigma_{v2}^2 = 0$ i.e. $v_{2t} = 0 \forall t$. La matrice de variance covariance du bruit de mesure v_t devient:

$$Q = \begin{pmatrix} \sigma_{v1}^2 & 0 \\ 0 & 0 \end{pmatrix}$$

Dans ce cas la pente du processus est constante tandis que le niveau est stochastique. Nous avons alors le système suivant :

$$x_{1t} = x_{1t-1} + x_{2t} + v_{1t} \quad (2.9.a)$$

$$x_{2t} = x_{2t-1} \quad (2.9.b)$$

$$y_t = x_{1t} + u_t \quad (2.9.c)$$

Posons $x_{2t} = \beta \forall t$, la solution par récurrence de (2.9.a), pour une condition initiale $x_{10} = \alpha$ donnée, est:

$$x_{1t} = \alpha + \beta t + \sum_{i=0}^t v_{1t-i} \quad (2.10)$$

Le processus y_t est, alors, défini par

$$y_t = \alpha + \beta t + w_t + u_t \quad (2.11)$$

avec $w_t = \sum_{i=0}^t v_{1t-i}$ tel que
$$\begin{cases} E(w_t) = 0 \quad \forall t \\ E(w_t w_s') = \sigma_{v1}^2 \min(t,s) \end{cases}$$

La composante stochastique w_t représente l'accumulation d'accroissements aléatoires v_{1t} . C'est une composante aléatoire non stationnaire qui s'interprète comme la tendance stochastique.

Le terme βt est la tendance déterministe, ainsi le terme $\beta t + w_t$ représente la composante permanente.

Le terme u_t est la composante transitoire pour les séries d'observations supposée être un bruit blanc.

L'expression en différence première (ou taux de croissance) du processus y_t est alors :

$$(1-L)y_t = \beta + v_{1t} + (1-L)u_t \quad (2.12)$$

ou la somme des accroissements déterministe et stochastique: $\beta + v_{1t}$ représente le taux de croissance permanent.

Puisque le membre droit de (2.12) est stationnaire, la série y_t est alors

stationnaire en différence première (ou D.S, pour difference stationary).

2e cas: Les bruits d'état v_{1t} et v_{2t} sont nuls.

Quand $Q = 0$ i.e. $v'_t = [v_{1t}, v_{2t}] = 0$, la pente et le niveau du processus y_t sont déterministes et la série y_t , en remplaçant w_t par zéro dans (2.11), devient :

$$y_t = \alpha + \beta t + u_t \quad (2.13)$$

ou βt est la tendance déterministe.

Comme le terme aléatoire u_t a été supposé un bruit blanc l'équation (2.13) peut être estimée par les moindres carrés.

Soient $\hat{\alpha}$ et $\hat{\beta}$ les paramètres estimés. Le terme d'écart à la tendance

$$\tilde{y}_t = y_t - \hat{\alpha} - \hat{\beta}t = (\beta - \hat{\beta})(t - \bar{t}) + (u_t - \bar{u}) \quad (2.14)$$

est stationnaire et on qualifiera la variable considérée de stationnaire en écart à une tendance (ou T.S, pour trend stationary).

3e cas: le bruit d'observation u_t est nul.

Quand la série est observée sans erreur le système devient :

$$x_{1t} = x_{1t-1} + x_{2t} + v_{1t} \quad (2.15.a)$$

$$x_{2t} = x_{2t-1} + v_{2t} \quad (2.15.b)$$

$$y_t = x_{1t} \quad (2.15.c)$$

Ce système correspond pour y_t à un processus ARIMA (0,2,1) ou IMA (2,1). Cela se voit immédiatement en remplaçant u_t par zéro dans (2.6). il vient:

$$(1-L)^2 y_t = v_{2t} + (1-L)v_{1t} \quad (2.16)$$

dont la fonction d'autocovariance est

$$E(z_t z'_t) = \begin{cases} \sigma_{v2}^2 + 2\sigma_{v1}^2 + 2\sigma_{v1v2} & \text{si } s = t \\ -\sigma_{v1}^2 - \sigma_{v1v2} & \text{si } |s-t| = 1 \\ 0 & \text{si } |s-t| > 1 \end{cases} \quad (2.17)$$

Si en plus $v_{2t} = 0$, on obtient par référence à (2.11) et (2.12)

$$y_t = \alpha + \beta t + w_t \quad \text{avec } w_t = \sum_{i=0}^t v_{1t-i}$$

d'où $(1-L)y_t = \beta + v_{1t}$

Il s'agit d'un modèle stationnaire en différence première, ce modèle est équivalent au modèle de Nelson et Plosser (1982). On peut considéré

un cadre beaucoup général en supposant que les bruits sont des processus ARIMA ou ARMA.

4e cas: le bruit d'observation u_t est un processus ARIMA.

Comme Beveridge et Nelson (1982), on adopte ici le processus ARIMA suivant:

$$(1-L)A(L)u_t = B(L)\varepsilon_t,$$

où ε_t est un bruit blanc (de variance σ^2) et $A(L)$ et $B(L)$ sont deux polynômes de retards inversibles, stationnaires et n'ayant aucune racine en commun. Puisque $A(L)$ est inversible, on peut écrire

$$(1-L)u_t = C(L)\varepsilon_t$$

avec $C(L) = B(L)/A(L)$. Son introduction dans (2.12) donne

$$(1-L)y_t = \beta + v_{1t} + C(L)\varepsilon_t \quad (2.18)$$

où $\beta + v_{1t}$ est le taux de croissance permanent.

Il s'agit d'un modèle stationnaire en différence première, la formulation nous permet d'avoir l'hypothèse de non corrélation des chocs temporaire et permanent. Nous avons, ainsi, une généralisation du modèle de Beveridge et Nelson. De plus si $v_{1t} = 0$ l'expression du taux de croissance devient

$$(1-L)y_t = \beta + C(L)\varepsilon_t \quad (2.19)$$

Dans ce cas le taux de croissance permanent devient déterministe. Néanmoins la factorisation du polynôme $C(L)$ autour de la racine unité sous la forme

$$C(L) = C(1) + (1-L)C^*(L) \quad (2.20)$$

permet d'avoir le taux de croissance permanent stochastique. Introduisons (2.20) dans (2.19) nous avons:

$$(1-L)y_t = \beta + C(1)\varepsilon_t + C_1(L)\varepsilon_t \quad (2.21)$$

avec $C_1(L) = (1-L)C^*(L)$

$C(1)$ est la somme des coefficients du polynôme, il correspond donc à la réponse à long terme.

Le taux de croissance permanent est la somme d'un accroissement tendanciel déterministe β et d'un accroissement tendanciel stochastique $C(1)\varepsilon_t$.

$C_1(L)\varepsilon_t$ correspond à la composante transitoire stationnaire.

Ce modèle correspond au modèle de Beveridge et Nelson où la tendance stochastique est un multiple de l'aléa ε_t , ce qui traduit l'hypothèse de parfaite corrélation des chocs temporaire et permanent.

5e cas le bruit d'état v_{1t} est un processus M.A.

Dans ce qui suit nous supposons que les bruits d'observation u_t et d'état v_{2t} sont nuls. Le bruit d'état v_{1t} est processus de moyenne mobile (M.A) définit par

$$v_{1t} = C(L)\varepsilon_t$$

où $\varepsilon_t \sim \text{i.i.d } (0, \sigma_\varepsilon^2)$

$C(L)$ est un polynôme de retard infini: $C(z) = \sum_{i=0}^{\infty} c_i z^i$ avec $c_0 = 1$.

Nous avons alors le système suivant:

$$\begin{cases} x_{1t} = x_{1t-1} + x_{2t} + C(L)\varepsilon_t \\ x_{2t} = x_{2t-1} \\ y_t = x_{1t} \end{cases}$$

Admettons maintenant que $x_{2t} = \beta \forall t$, le système est alors équivalent à un modèle IMA(∞).

$$(1-L)y_t = \beta + C(L)\varepsilon_t$$

Ce modèle est identique au modèle (2.19).

6e cas: le trend cyclique.

Harvey (1985) a montré que "le cycle est une part intrinsèque de la tendance plutôt qu'une composante séparée qui peut lui être ajoutée ultérieurement". Il a introduit un processus stochastique cyclique stationnaire w_t dans l'équation de niveau conduisant à l'équation suivante

$$\begin{cases} x_{1t} = x_{1t-1} + x_{2t-1} + w_{t-1} + v_{1t} & (2.22.a) \\ x_{2t} = x_{2t-1} + v_{2t} & (2.22.b) \\ y_t = x_{1t} + u_t & (2.22.c) \end{cases}$$

Soulignons la différence entre le modèle de Harvey et le modèle (2.4), dans (2.24.a) mis à part le terme w_{t-1} , le niveau x_{1t} est fonction de la pente de la période précédente. Si $\sigma_{v2}^2 = 0$ alors on obtient un modèle D.S. correspond au modèle (2.11).

La composante cyclique w_t est modélisée par

$$\begin{bmatrix} w_t \\ * \\ w_t \end{bmatrix} = r \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} w_{t-1} \\ * \\ w_{t-1} \end{bmatrix} + \begin{bmatrix} e_t \\ * \\ e_t \end{bmatrix} \quad (2.23)$$

où e_t et e_t^* sont des bruits blancs non corrélés de moyenne nulle et de variances respectives σ_e^2 et $\sigma_{e^*}^2$,

le paramètre θ ($0 \leq \theta \leq \pi$) est la fréquence angulaire du cycle ($\theta = \frac{2\pi}{T}$

si T est la période),

le paramètre r ($0 \leq r \leq 1$) est l'amplitude du cycle.

On peut écrire (2.23) après élimination de w_t^* par

$$(1 - 2r\cos\theta.L + r^2L^2)w_t = (1 - r\cos\theta.L)e_t + r\sin\theta.L.e_t^*$$

qui est donc un processus cyclique ARMA(2,1).

On peut considérer quelques cas particuliers suivant les valeurs de θ .

cas 1 si $\theta = 0$ alors $w_t = rw_{t-1} + e_t$

cas 2 si $\theta = \pi$ alors $w_t = -rw_{t-1} + e_t$

Dans les deux cas le processus w_t devient donc un processus autorégressif de 1er ordre. Pour que le processus soit cyclique il faut donc restreindre la valeur de θ dans l'intervalle $]0, \pi[$.

En choisissant les variables x_{1t} , x_{2t} , w_t et w_t^* comme des variables d'état nous pouvons représenter le modèle état-mesure par le système suivant:

Equation d'évolution.

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ w_t \\ w_t^* \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r\cos\theta & r\sin\theta \\ 0 & 0 & -r\sin\theta & r\cos\theta \end{bmatrix} \begin{bmatrix} x_{1t-1} \\ x_{2t-1} \\ w_{t-1} \\ w_{t-1}^* \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \\ e_t \\ e_t^* \end{bmatrix}$$

Equation de mesure

$$y_t = [1 \ 0 \ 0 \ 0] \begin{bmatrix} x_{1t} \\ x_{2t} \\ w_t \\ w_t^* \end{bmatrix} + u_t$$

Nous venons de montrer la formalisation d'un processus intégré dans l'espace d'états et les modélisations des séries macroéconomiques qui y sont associées. Considérons maintenant un cadre plus général.

3. Formulation générale.

D'une manière générale, des variables exogènes ou variables de commandes peuvent apparaître dans les équations de transition et des variables prédéterminées dans les équations d'observations.

Un modèle état-mesure comporte deux ensembles d'équations :

1) équation de transition ou équation d'état.

$$x_{t+1} = f(x_t, z_t^1, v_t, t) \quad (3.1.a)$$

où :

x_t : le vecteur d'état

z_t^1 : le vecteur des variables exogènes ou variables de commandes

v_t : le vecteur de perturbation du processus dynamique

2) équation de mesure ou d'observations.

$$y_t = g(x_t, z_t^2, u_t, t) \quad (3.1.b)$$

où :

y_t : le vecteur de mesure ou vecteur d'observations

z_t^2 : le vecteur de variables prédéterminées

u_t : le vecteur de perturbation d'observations

Définition.

On appelle réalisation, la transformation en représentation d'état ou modèle markovien qui est, dans ce qui suit, supposé être un système dynamique linéaire

$$X_t = A_t X_{t-1} + D_t^1 Z_t^1 + B_t V_t \quad (3.2.a)$$

$$Y_t = C_t X_t + D_t^2 Z_t^2 + U_t \quad (3.2.b)$$

ou: A_t est la matrice d'état ou d'évolution de dimension (n,n)

B_t est la matrice d'entrée aléatoire de dimension (n,n)

C_t est la matrice d'observations ou de mesure de dimension (m,n)

D_t^1 est la matrice d'application de la commande de dimension (n,1)

D_t^2 est la matrice de dimension (m,1)

X_t est le vecteur d'état de dimension (n,1)

Y_t est le vecteur d'observations de dimension (m,1)

La représentation d'état d'un système dynamique a été développée pour les systèmes physiques par les automaticiens, le vecteur d'état joue un rôle fondamental, il se présente comme facteur d'énergie d'accumulation c'est-à-dire comme un système pouvant accumuler de l'énergie potentielle, électrique, ... Une analogie de ce système en économie est alors d'interpréter le vecteur d'état comme le facteur d'accumulation de l'information.

Les hypothèses.

H1 l'état initial X_{t_0} du système est un vecteur aléatoire gaussien tel que :

$$E(X_{t_0}) = \bar{X}_0$$

$$E[(X_{t_0} - \bar{X}_0)(X_{t_0} - \bar{X}_0)'] = P_0$$

où P_0 est matrice (n,n) symétrique de type non négatif

H2 l'état initial X_0 est non corrélé avec les vecteurs de bruits

U_t et V_t pour tout $t \in \mathbb{Z}$

$$E(U_t X_0') = 0$$

$$E(V_t X_0') = 0$$

H3 les bruits dynamiques V_t et d'observations U_t sont supposés gaussiens temporellement indépendants, indépendants entre eux de moyennes nulles et de matrice variances-covariances connues

$$E(V_t) = E(U_t) = 0 \quad \forall t$$

$$E(V_s V_t') = \begin{cases} Q_t & \text{si } s = t \\ 0 & \text{sinon} \end{cases}$$

$$E(U_s U_t') = \begin{cases} R_t & \text{si } s = t \\ 0 & \text{sinon} \end{cases}$$

$$E(U_s V_t') = S_t \quad \forall t$$

les matrices Q_t , R_t , et S_t sont respectivement de dimension (n,n), (m,m) et (m,n)

$$\text{en résumé } \begin{bmatrix} V_t \\ U_t \end{bmatrix} \text{ N.I.D. } \begin{bmatrix} 0, \begin{pmatrix} Q_t & S_t \\ S'_t & R_t \end{pmatrix} \end{bmatrix} \quad (3.3)$$

H4 les matrices A_t , B_t , D_t^1 , D_t^2 , et C_t sont supposées connues et non aléatoires.

Cadre général.

Considérons d'une part un système dynamique en temps discret à n dimension décrit par les équations d'état suivantes

$$X_t = AX_{t-1} + BV_t \quad (3.4.a)$$

et d'autre part un vecteur d'observations à m dimension

$$Y_t = CX_t + U_t \quad (3.4.b)$$

Ce modèle obéit aux hypothèses H1 à H4 précédentes. Notons toutefois que la matrice S est supposée nulle dans la suite (la généralisation au cas $S \neq 0$ s'obtient facilement).

Il nous paraît nécessaire de donner une solution explicite à l'équation d'état.

3.1. Propriétés des solutions explicites.

a) Résolution de l'équation d'état.

Pour une condition initiale donnée X_{t_0} , on obtient par récurrence

$$X_t = A^t X_{t_0} + \sum_{j=0}^{t-1} A^{t-1-j} B V_j \quad (3.5)$$

où A^t est la matrice de transition.

Cette relation met en évidence le fait que X_t est une combinaison linéaire des vecteurs aléatoires conjointement gaussiens $(X_{t_0}, V_0, V_1, \dots, V_{t-1})$, ces vecteurs aléatoires sont conjointement gaussiens du fait qu'ils sont individuellement gaussiens et indépendants. Par suite X_t est une variable aléatoire gaussienne. Il est évident que si X_{t_0} est centré alors X_t est un processus centré.

Remarque.

Toute solution stationnaire X_t est purement non déterministe. La matrice A est stable (valeurs propres de module inférieur à l'unité noté $|\lambda_i(A)| < 1$).

quand t tend vers l'infini alors $\lim_{t \rightarrow \infty} A^t X_0 = 0$

$$\text{d'où } X_t = \sum_{j=1}^{\infty} A^{t-j} B V_j$$

On montre que X_t est un processus gaussien markovien et on dit que les équations (3.4) constituent une représentation gaussienne markovienne de y_t .

b) Résolution de l'équation d'observations.

En reportant l'expression développée (3.5) de X_t dans l'équation (3.4.b) nous pouvons écrire

$$Y_t = C A^t X_0 + \sum_{j=0}^{t-1} C A^{t-1-j} B V_j + U_t$$

Cette expression nous montre que Y_t est une combinaison linéaire des vecteurs aléatoires conjointement gaussiens ($X_0, V_0, V_1, \dots, V_{t-1}, U_t$). Par conséquent Y_t est lui-même un processus gaussien. Cependant à la différence de X_t , Y_t n'est pas un processus markovien.

Propriétés des processus.

Nous avons montré que les processus X_t et Y_t sont gaussiens, leurs propriétés de probabilités sont entièrement définies par les deux premiers moments.

$$E(X_t) = E[A^t X_0 + \sum A^{t-1-j} B V_j] = A^t \bar{X}_0$$

$$E(Y_t) = E[C A^t X_0 + \sum C A^{t-1-j} B V_j + U_t] = C A^t \bar{X}_0$$

L'hypothèse $|\lambda_i(A)| < 1$ implique que $\lim_{t \rightarrow \infty} A^t \bar{X}_0 = 0$

La matrice de variance-covariance du processus X est définie par

$$E(X_t X_t') = A E(X_{t-1} X_{t-1}') A' + B Q B'$$

$$P_0 = A^{t-t_0} P_0 (A')^{t-t_0} + \sum_{i=0}^{t-t_0-1} A^i B Q B' A'^i$$

$$E(X_{t+j} X_t') = P_j = A^j P_0 \quad j > 0$$

$$= P_0 A^{-j} \quad j < 0$$

où P_0 est solution d'une équation de Lyapunov discrète

$$P_0 - AP_0A' = BQB'$$

La transformation en z de la covariance est

$$G_{XX}(z) = (ZI - A)^{-1}BQB'(Z^{-1}I - A')^{-1}$$

La matrice de variance-covariance de l'observation.

$$E(Y_{t+j}Y_t') = C E(X_{t+j}X_t')C' + E(V_{t+j}X_t')C' + CE(X_tV_{t+j}') + E(V_{t+j}V_t')$$

$$\Gamma(j) = CA^jP_0C' \quad \text{pour } j > 0$$

$$= CP_0C' + R \quad \text{pour } j = 0$$

$$= CP_0A'^{-j}C' \quad \text{pour } j < 0$$

La transformation en z de la covariance de l'observation est :

$$G_{YY}(z) = C \left[(ZI - A)^{-1}BQB'(Z^{-1}I - A')^{-1} \right] C' + R.$$

Considérons maintenant le problème de l'estimation de l'état X du système.

3.2. Estimation du vecteur d'état.

Le problème est de trouver le meilleur estimateur $\hat{X}$ de l'état X à l'instant s à partir d'observations effectuées jusqu' à l'instant t . Suivant que $s < t$, $s = t$, ou $s > t$ on dit que $\hat{X}$ est une valeur lissée, filtrée ou prédite de X et le problème d'estimation correspondant est appelé respectivement problème de lissage, de filtrage ou de prédiction. Plusieurs méthodes d'obtention d'un tel estimateur sont possibles conduisant au même système représenté par un algorithme appelé filtrage de Kalman.

Dans notre cas sous l'hypothèse gaussienne du vecteur d'état X_t et d'observations Y_t , on recherche l'estimateur selon le critère de la variance conditionnelle minimum.

Notons par $y_t = \{ Y_0, Y_1, \dots, Y_t \}$ les vecteurs d'observations jusqu'à un temps t inclus.

$$E \left[\|X_s - \hat{X}\|^2 / y_t^- \right] \leq E \left[\|X_s - Z\|^2 / y_t^- \right]$$

pour tout vecteur Z fonction des observations $\{Y_0, Y_1, \dots, Y_t\}$ où $||.||$ représente la norme euclidienne.

En effet, $\hat{X}$ est égal à la moyenne conditionnelle de X_s

$$\hat{X} = E \left[X_s / y_t^- \right] = \hat{X}_{s/t}.$$

Proposition 1. (Filtre de Kalman).

Le meilleur estimateur $\hat{X}$ de l'état X_t en termes de vecteurs d'observations y_t peut être exprimé par une relation de récurrence utilisant la prévision $\hat{X}_{t/t-1}$ et l'observation courante y_t , donné par

$$\hat{X}_{t/t} = A\hat{X}_{t-1/t-1} + K_t e_{t/t-1} \quad (3.6.1)$$

$$y_t = C\hat{X}_{t/t-1} + e_{t/t-1} \quad (3.6.2)$$

avec la condition initiale

$$\bar{X}_0 = E(X_{t_0})$$

où $e_{t/t-1} = y_t - \hat{y}_{t/t-1}$ est le processus d'innovation gaussien.

Le processus d'innovation est blanc, centré de matrice de covariance

$$H_t = E(e_t e_t') = C P_{t/t-1} C' + R. \quad (3.6.3)$$

La matrice de gain est donnée par

$$K_t = P_{t/t-1} C' [C P_{t/t-1} C' + R]^{-1} \quad (3.6.4)$$

La matrice $P_{t/t-1}$ définie non négative de dimension (n,n) est la matrice de covariance de l'erreur d'estimation

$$\begin{aligned} \varepsilon_{t/t-1} &= X_t - \hat{X}_{t/t-1} \\ \text{i.e. } P_{t/t-1} &= E(\varepsilon_{t/t-1} \varepsilon_{t/t-1}') \\ &= A P_{t-1} A' + B Q B' \end{aligned} \quad (3.6.5)$$

et peut être précalculée en utilisant la relation itérative suivante appelée équation de Riccati

$$P_t = A P_{t/t-1} A' - P_{t/t-1} C' [C P_{t/t-1} C' + R]^{-1} C P_{t/t-1} + B Q B' \quad (3.6.6)$$

avec la condition initiale

$$P_0 = E[(X_{t_0} - \bar{X}_0)(X_{t_0} - \bar{X}_0)'].$$

Preuve (voir Annexe).

Examinons maintenant la structure de l'estimateur ainsi définie.

Le filtre est composé de trois parties :

- 1) la prévision à un pas de l'état X_t
- 2) une matrice de gain K_t variant avec le temps
- 3) une boucle de rétroaction de gain unitaire et négatif.

A chaque instant $t = 0, 1, 2, \dots$ la boucle de rétroaction engendre le processus d'innovation

$$e_{t/t-1} = Y_t - C\hat{X}_{t/t-1} \text{ (erreur de la projection orthogonale de } Y_t \text{ sur l'espace } \mathcal{Y}_{t-1} \text{ engendré par } \{Y_0, Y_1, \dots, Y_{t-1}\})$$

qui est non corrélé avec la séquence des observations passées $\mathcal{Y}_{t-1}$ (l'erreur associée à un estimateur par projection est orthogonale à l'observation) et représente l'information supplémentaire obtenue de par la nouvelle observation Y_t . Ceci agit sur le système par une matrice de gain K_t qui indique comment utiliser, au mieux, au sens des moindres carrés l'innovation $e_{t/t-1}$ et génère alors le meilleur estimateur linéaire

$$E(X_t / e_{t/t-1}) = K_t (Y_t - C\hat{X}_{t/t-1})$$

de l'état X_t du système.

En même temps la prévision à un pas de l'état X_t engendre le meilleur estimateur linéaire

$$E(X_t / \mathcal{Y}_{t-1}) = \hat{X}_{t/t-1} = A\hat{X}_{t-1/t-1}$$

lequel peut être simplement ajouté à $E(X_t / e_{t/t-1})$ pour donner le meilleur estimateur linéaire de l'état

$$\hat{X}_{t/t} = E(X_t / \mathcal{Y}_t) = E(X_t / \mathcal{Y}_{t-1}, e_{t/t-1})$$

puisque l'innovation $e_{t/t-1}$ est non corrélée avec les observations $\mathcal{Y}_{t-1}$.

Nous avons mis en évidence le rôle du processus d'innovation dans l'estimation récursive de l'état X_t . L'innovation est par définition la part non prévisible de l'information.

On pourrait maintenant s'intéresser à l'hypothèse où la loi de distribution de la condition initiale X_0 et des bruits U_t et V_t n'est pas connue.

3.3. Hypothèses non gaussiennes.

En l'absence de l'hypothèse gaussienne, on peut faire une hypothèse de connaissance de second ordre de l'état et d'observations.

L'état initial X_0 du système est un vecteur aléatoire de second ordre.

$$X_0 \sim \text{I.I.D} (0, P_0)$$

Les bruits dynamiques U_t et V_t sont supposés de second ordre

$$\begin{bmatrix} U_t \\ V_t \end{bmatrix} \sim \text{I.I.D.} \left[0, \begin{pmatrix} Q & S \\ S' & R \end{pmatrix} \right]$$

Sous ces hypothèses le problème qui consiste à calculer l'estimation de l'état X_t revient à rechercher l'estimateur affine selon le critère du minimum de l'erreur quadratique moyenne

$$\min_Z E || X - Z ||^2$$

Cet estimateur a pour composantes, dans l'espace de Hilbert, des variables aléatoires du second ordre. Ce sont les projections orthogonales des composantes de variables aléatoires X_t sur le sous-espace de Hilbert $H(t-1) = \{Y_0, Y_1, \dots, Y_{t-1}\}$ (combinaisons linéaires des composantes des variables aléatoires d'observations). L'estimateur de l'état X_t est défini par

$$\hat{X}_t = E(X_t / H(t-1))$$

Nous savons que X_t est une fonction affine des variables aléatoires d'observations $\{Y_0, Y_1, \dots, Y_{t-1}\}$, les deux estimateurs sont donc identiques i.e. le meilleur estimateur linéaire coïncide avec l'estimateur conditionnel.

Quelques remarques sur le filtre de Kalman.

- Le filtre est linéaire, récursif, le système est de dimension finie.

- Si l'hypothèse de bruits gaussiens n'est plus satisfaite alors le filtre de Kalman devient l'estimateur linéaire de variance minimum.

- L'estimateur est non biaisé

$$E(X_t - E(X_t / Y_t)) = 0$$

- La condition d'orthogonalité est satisfaite

$$E\left[(X_t - \hat{X}_{t/t-1})Y_{t-1}\right] = 0$$

- Le calcul du gain K_t et des matrices de covariances $P_{t/t-1}$ et H_t n'est pas fonction des observations $y_{t-1} = \{Y_0, Y_1, \dots, Y_{t-1}\}$; ce calcul peut donc être fait au préalable si on dispose de l'ensemble des paramètres du modèle (A, C, Q, R). Il ne reste alors plus que le calcul de l'innovation $e_{t/t-1}$ et la valeur prédite $\hat{X}_t$ à effectuer au fur et à mesure que les observations Y_t sont fournies.

- On peut également établir les équations du filtre par calcul matriciel en minimisant la matrice de covariance de l'erreur d'estimation de l'état.

- Cas : observation non bruitée.

Si l'erreur associée à l'observation est nulle i.e. $R = 0$ ou d'une manière générale la matrice de l'erreur d'observations R est singulière alors la matrice de variance-covariance de l'innovation

$$H_{t/t-1} = CP_{t/t-1}C' + R$$

n'est pas inversible.

Dans ce cas le calcul de la matrice de gain

$$K_t = P_{t/t-1}C'H_{t/t-1}^{-1}$$

est devenu impossible.

- Nous avons supposé les matrices A, B, et C fixes, la généralisation de l'étude s'effectue sans difficulté en considérant ces matrices en fonction du temps c'est-à-dire cadre non stationnaire.

4. Forme filtre ou modèle innovation.

Un modèle équivalent permettant de représenter le processus Y_t est obtenu par les équations (3.6.1) et (3.6.2) du filtre de Kalman.

$$S'1 \begin{cases} \hat{X}_t = A\hat{X}_{t-1} + K_t e_{t/t-1} & (3.7.a) \\ Y_t = C\hat{X}_{t-1} + e_{t/t-1} & (3.7.b) \end{cases}$$

avec $e_{t/t-1} = Y_t - \hat{Y}_{t/t-1}$ le processus d'innovation.

Ce modèle est appelé modèle innovation ou forme filtre, les mêmes observations Y_t , $t \in \mathbb{N}$ sont décrites soit par le système S1 soit par le système S'1. L'état X_t du système explicatif de Y_t est alors remplacé par l'estimé $\hat{X}_{t/t-1}$ calculable à partir des informations passées y_{t-1} .

Nous avons donc un processus de Markov avec comme entrée le bruit blanc $e_{t/t-1}$.

Le modèle innovation présente un intérêt considérable dans la pratique grâce à ses propriétés statistiques, nous pouvons constater que le modèle innovation est déterminé par les mêmes paramètres dynamiques (A, C) que le modèle initial S1, en revanche, ses paramètres statistiques sont (K_t, H_t) .

4.1. Propriétés des solutions explicites.

La solution récursive de l'équation de l'état, étant donné la condition initiale, s'écrit :

$$\hat{X}_t = A^t X_0 + \sum_{j=0}^{t-1} A^{t-j-1} K_j e_j$$

ce qui nous permet d'écrire l'équation d'observations

$$\begin{aligned} Y_t &= CA^t X_0 + C \sum_{j=0}^{t-1} A^{t-j-1} K_j e_j + e_t \\ &= CA^t X_0 + \sum_{j=0}^t CA^{t-j} K_j e_j. \end{aligned}$$

Posons par $r(t-j) = CA^{t-j} K_j$

$$\text{avec } r(t-j) = \begin{cases} CA^{t-j} K_j & 0 \leq j < t \\ I & j = t \\ 0 & j > t \end{cases}$$

Le terme $r(t-j)$ représente la réponse impulsionnelle du filtre, l'équation d'observations devient alors:

$$Y_t = CA^t X_0 + \sum_{j=0}^t r(t-j) e_j$$

Puisque $\|A\| < 1$ et pour une condition initiale donnée, lorsque t tend vers l'infini $\lim_{t \rightarrow \infty} CA^t X_0 = 0$, l'équation d'observations devient :

$$Y_t = \sum_{j=1}^{\infty} r_j e_j$$

Cette dernière relation nous montre que l'équation d'observations est fonction de l'innovation.

4.2. Comparaison de la propriété statistique du modèle innovation et du modèle initial.

Une comparaison de la propriété statistique du modèle innovation avec le modèle initial permet d'apprécier celui-ci.

Rappelons, d'abord, la matrice de covariance du modèle initial avec bruits corrélés.

$$E \left[\begin{pmatrix} V \\ U \end{pmatrix} \begin{pmatrix} V \\ U \end{pmatrix}' \right] = \begin{bmatrix} Q & S \\ S' & R \end{bmatrix}$$

Pour le modèle innovation, le processus d'innovation e_t est gaussien de moyenne nulle et de variance H_t , par suite $K_t e_t$ est gaussien de moyenne nulle et de covariance $K_t' H_t K_t$.

$$E \left[\begin{pmatrix} K_t e_t \\ e_t \end{pmatrix} \begin{pmatrix} K_t e_t \\ e_t \end{pmatrix}' \right] = \begin{bmatrix} K_t' H_t K_t & K_t' H_t \\ H_t K_t & H_t \end{bmatrix}$$

En comparant ces deux matrices de covariances, nous pouvons dire que le triplet (Q, R, S) du modèle initial est statistiquement équivalent au doublet (K_t, H_t) du modèle innovation.

D'après les principaux résultats ci-dessus, nous pouvons dresser un tableau de comparaison entre le modèle général et le modèle innovation

Type de modèle	Paramètre dynamique	Paramètres statistiques	Nombre
Modèle général	A	Q, R, S	$n(n+1) + n + 1$
Modèle innovation	A	K, H	$n + 1$

Au vu de cette comparaison le modèle innovation présente un intérêt considérable.

Le multiplicateur dynamique de long terme ou la réponse impulsionnelle d'une série en différence stationnaire est un indicateur de persistance Campbell et Mankiw (1987). Nous nous intéressons ici à la relation entre la réponse impulsionnelle et la matrice de covariance.

4.3. Relation entre la réponse impulsionnelle et la matrice de covariance.

La covariance de y_t peut être calculée en fonction de la réponse impulsionnelle $r(t-j)$; pour une condition initiale nulle, nous avons :

$$\tilde{r}(t-s) = \sum_{j=0}^t \sum_{i=0}^s r(t-j) E(e_j e_i') r'(t-i)$$

$$= \sum_{i=0}^{\min(t,s)} r(t-i) H_1 r'(s-i)$$

Considérons la matrice de covariance

$$E(y_t y_t') = \tilde{r}_t$$

où $y_t' = (Y_0, Y_1, \dots, Y_t)'$

$$\text{avec } \tilde{r}_t = \begin{bmatrix} \Gamma_{t,t} & \Gamma_{t,t-1} & \Gamma_{t,0} \\ & \Gamma_{t-1,t-1} & \\ & & \Gamma_{0,0} \end{bmatrix}$$

Si le processus est stationnaire, cette matrice de covariance devient

$$\tilde{r}_t = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \Gamma_t \\ & \Gamma_0 & \\ & & \Gamma_0 \end{bmatrix}$$

c'est une matrice de Toeplitz.

Reprenons la solution de l'équation d'observations

$$Y_t = \sum_{j=1}^{\infty} r_j e_j$$

l'écriture sous forme matricielle est

$$Y_t = \tilde{R}_t \tilde{E}_t$$

$$\text{avec } \tilde{R}_t = \begin{bmatrix} I & r_{t,t-1} & r_{t,0} \\ & I & \\ 0 & & I \end{bmatrix} ; \tilde{E}_t = (e_0, e_1, \dots, e_t)$$

d'où la matrice de covariance est $E(Y_t Y_t') = \tilde{\Gamma}_t = \tilde{R}_t \tilde{H}_t \tilde{R}_t'$

$$\text{avec } E(\tilde{E}_t \tilde{E}_t') = \tilde{H}_t = \begin{bmatrix} H_t & & \\ & H_{t-1} & \\ & & H_0 \end{bmatrix}$$

Nous avons donc une décomposition de la matrice de covariance de l'observation en fonction d'une matrice triangulaire $\tilde{R}_t$ et d'une matrice diagonale $\tilde{H}_t$ (matrice de variance-covariance de l'innovation), ce résultat nous montre qu'on peut obtenir la réponse impulsionnelle à partir de la matrice de covariance.

5. Analyse de la stabilité stochastique.

La notion de stabilité est indispensable pour le système mais avant de la définir rappelons d'abord les notions de commandabilité, d'observabilité et d'atteignabilité des états d'un système, ce sont des notions peu familières aux économètres, statisticiens et économistes.

5.1. Notions de commandabilité et d'observabilité stochastiques.

Ces deux notions ne sont pas habituelles aux économètres et aux statisticiens et pourtant elles possèdent des propriétés statistiques.

a) Commandabilité.

Reprenons l'équation d'état

$$X_t = A X_{t-1} + B V_t$$

Pour une condition initiale donnée $X_{t_0} = X_0$, la solution récursive est définie par

$$X_t = A^t X_0 + \sum_{i=0}^{t-1} A^{t-1-i} B V_i.$$

Posons $d_t = X_t - A^t X_0$ qui satisfait à
 $E(d_t) = 0$

et

$$\begin{aligned}\text{Var}(d_t) &= \sum_{i=0}^{t-1} A^{t-1-i} B Q B' (A^{t-1-i})' \\ &= G_c\end{aligned}$$

$\|G_c\|$ demeure bornée pour tout t , où $\|\cdot\|$ est une norme euclidienne, alors $\|d_t\|$ reste bornée pour tout t . G_c est appelée matrice de commandabilité stochastique du système dynamique.

Etant donné que la matrice de variance covariance de l'erreur Q est définie positive, la matrice G_c est alors définie positive si et seulement si le rang de $[B, AB, A^2B, \dots, A^{t-1}B]$ vaut n où n désigne la dimension de l'état du système.

Cette définition implique que l'effet de la perturbation du système reste borné.

Pour $B = I$ la condition devient :

$$\text{La matrice } G_c = \sum_{i=0}^{t-1} A^{t-1-i} Q (A^{t-1-i})' \text{ est inversible si et seulement}$$

si le rang $[Q^{1/2}, A Q^{1/2}, \dots, A^{t-1} Q^{1/2}]$ vaut n .

b) Observabilité.

Un système est dit observable si, pour un instant initial t_0 , il est possible de trouver un instant t_1 tel que l'observation de Y_t et la connaissance d'une commande sur l'intervalle de temps $[t_0, t_1]$ permettent de trouver l'état initial X_{t_0} . La notion d'observabilité est souvent confondue dans les cadres déterministe et stochastique et pourtant elle n'a pas la même propriété statistique dans les deux cas.

Observabilité déterministe.

$$Y_t = C X_t$$

Observabilité stochastique.

$$Y_t = C X_t + U_t$$

A cause de la présence du bruit, la définition de l'observabilité déterministe n'a plus de sens dans le cas stochastique.

Théorème (FAURRE.P et ROBINS. M (1984)).

Un système linéaire est observable si et seulement si le rang de la matrice d'observabilité $O = [C', A'C', \dots, A^{m-1}C']'$ de dimension $(n \times m, n)$ vaut n .

Lemme (AOKI. M (1987)) : observabilité gramienne.

Le système linéaire est observable si la matrice

$$G_O = \sum_{i=0}^{m-1} (A')^i C' C A^i > 0.$$

On dit que le couple $(A, C) = [C', A'C', \dots, A^{m-1}C']$ est complètement observable.

Ainsi le rang $O = n$ si et seulement si $OO' = G_O > 0$.

Soit $\hat{X}_t$ l'estimateur de X et P_{t/t_0} la matrice de variance-covariance des erreurs, on peut se demander ce qui se passe lorsqu'on augmente les observations, on peut espérer améliorer l'estimation de l'état ou tout au moins ne pas le détériorer.

Le système est stochastiquement observable si la matrice de variance-covariance P_{t/t_0} demeure bornée en norme lorsque t tend vers l'infini.

Théorème (KUSNER (1971)).

Si le couple (A, C) est observable, la matrice de variance-covariance P_{t/t_0} est uniformément bornée quand t tend vers l'infini.

Le concept d'observabilité a une interprétation statistique, c'est la condition sur le comportement de l'erreur d'estimation des paramètres ou vecteur d'état quand le nombre d'observations augmente.

L'observabilité stochastique peut donc être définie comme une condition de convergence en probabilité de l'estimateur.

Le théorème d'observabilité ou le lemme associé permet de choisir un modèle d'état adéquat pour représenter l'évolution d'un processus réel.

Remarques.

1) Il y a équivalence entre la notion d'observabilité en système et la notion d'identification en économétrie, l'étude de la condition de rang de la matrice d'observabilité permet de déterminer le vecteur X_0 par les n observations, une telle condition correspond en économétrie à identifier le modèle.

2) les notions d'observabilité et de commandabilité sont duales.

3) Ces deux notions peuvent avoir des interprétations économiques. AOKI.M (1976) a donné une interprétation économique de la notion d'observabilité en étudiant l'équilibre d'un marché soumis à des perturbations.

c) Définition de la stabilité stochastique.

On dit qu'un système dynamique est stable relativement à une trajectoire lorsque des faibles perturbations appliquées au système entraînent de faibles écarts par rapport à cette trajectoire.

L'équation d'évolution de l'état est asymptotiquement stable si et seulement si toutes les valeurs propres de la matrice A sont à l'intérieur du cercle unité $|\lambda_i(A)| < 1$.

Nous avons vu à la section (3.1) que la condition $|\lambda_i(A)| < 1$ assure également la stationnarité du processus X_t .

Par conséquent, pour une équation d'évolution de l'état invariante, la condition $|\lambda_i(A)|$ permet de déterminer le filtre optimal invariant. Rappelons d'abord les équations du prédictor optimal et la matrice de covariance des erreurs.

$$\hat{X}_t = (A - K_t C) \hat{X}_{t-1} + K_t Y_t$$

$$P_t = A[P_{t/t-1} - P_{t/t-1} C(C' P_{t/t-1} C + R)^{-1} C' P_{t/t-1}] A + B Q B'$$

ou également

$$P_t = (A - K_t C') P_{t/t-1} (A - K_t C')' + K_t R K_t' + B Q B'$$

L'analyse de stabilité de ce système consiste à montrer la convergence des matrices de covariances P_t , H_t et la matrice de gain K_t .

5.2. Théorème.

Si l'équation de l'évolution de l'état est invariante, asymptotiquement stable i.e. $|\lambda_1(A)| < 1$, la paire (A, B) complètement commandable et la paire (A, C) complètement observable alors

a) pour une condition initiale donnée P_{t_0} , matrice définie non négative, on obtient

$$\lim_{t \rightarrow \infty} P_{t/t+1} = P$$

où P est indépendant de P_{t_0} et vérifie l'état stationnaire

$$P = A[P - PC(C'PC + R)^{-1}C'P]A' + BQB'$$

appelé équation stationnaire de Riccati que l'on écrit également

$$P = (A - KC)P(A - KC)' + KRK' + BQB'$$

dans le cas où les bruits d'état et d'observations sont non corrélés. soit encore

$$P = A[P - (PC + BS)(C'PC + R)^{-1}(C'P + BS)']A + BQB'$$

dans le cas où les bruits d'état et d'observations sont corrélés.

b) la limite de la matrice de gain de Kalman

$$\lim_{t \rightarrow \infty} K_t = K$$

$$\text{avec } K = APC(C'PC + R)^{-1} \quad (\text{cas bruits non corrélés})$$

$$\text{ou } K = (APC + BS)(C'PC + R)^{-1} \quad (\text{cas bruits corrélés})$$

c) les valeurs propres de la matrice $|A - KC'|$ sont à l'intérieur du cercle unité i.e. $|\lambda_1(A - KC)| < 1$ alors le prédicteur optimal

$$\hat{X}_t = (A - KC')\hat{X}_{t-1} + KY_t$$

est asymptotiquement stable uniformément.

Lemme.

La matrice de covariance de l'innovation

$$H_t = CP_{t/t-1}C' + R$$

est uniformément bornée.

$$\lim_{t \rightarrow \infty} H_t = H = CPC' + R$$

Preuve.

Sachant que $P_{t/t-1}$ décroît vers P , il vient

$H_t = CP_{t/t-1}C' + R$ décroît vers $H = CPC' + R$ or R est une matrice strictement positive donc la matrice H est aussi strictement positive.

Quelques remarques.

1) la stabilité de la matrice $(A - KC)$ assure la stationnarité de l'équation de prédictor optimal.

2) On peut exprimer l'estimé de la composante permanente $\hat{X}_t$ en fonction de l'observation Y_t

$$\hat{X}_t = \sum_{i=0}^t (A - KC)^{t-i} KY_{t+1}$$

la stabilité assure également que la puissance matricielle décroît géométriquement (i.e. le multiplicateur de l'observation est convergent).

Ces résultats nous permettent d'étudier les propriétés du modèle innovation.

6. Les propriétés du modèle innovation.

6.1. Détermination des pôles et zéros de la fonction de transfert.

Considérons le modèle innovation suivant :

$$X_t = AX_{t-1} + Ke_t$$

$$Y_t = CX_t + e_t$$

où la matrice de variance de l'innovation est définie par

$$E(e_t e_t') = H = C'PC + R.$$

La matrice de covariance du modèle innovation est

$$E \begin{bmatrix} \begin{pmatrix} Ke_t \\ e_t \end{pmatrix} \begin{pmatrix} Ke_t \\ e_t \end{pmatrix}' \end{bmatrix} = \begin{bmatrix} KHK' & KH \\ H K' & H \end{bmatrix} = \begin{pmatrix} K \\ I \end{pmatrix} H \begin{pmatrix} K \\ I \end{pmatrix}'$$

l'équation d'évolution s'écrit

$$X_t = (ZI - A)^{-1} K e_t$$

et l'équation d'observations

$$Y_t = [C'(ZI - A)^{-1}K + I].e_t.$$

Soient $e(z)$ et $Y(z)$ les transformations respectives de e_t et Y_t en z , la fonction de transfert du modèle

$$G(z) = Y(z) / e(z) = C(ZI - A)^{-1}K + I$$

est une fraction rationnelle en Z stable et inverse stable.

La fonction de transfert $G(z)$ s'écrit aussi

$$\begin{aligned} C(ZI - A)^{-1}K + I &= \frac{C \operatorname{adj}(ZI - A) K}{\det(ZI - A)} + I \\ &= \frac{C \operatorname{adj}(ZI - A) K + \det(ZI - A)}{\det(ZI - A)} \end{aligned}$$

($\operatorname{adj}(X)$ signifie matrice adjointe de X).

En effet, il est évident que les pôles de $G(z) = (C(ZI - A)^{-1}K + I)$ sont les zéros de $\det(ZI - A)$ c'est-à-dire les valeurs propres de A parce que

$$\det(ZI - A) = Z^p + a_1 Z^{p-1} + \dots + a_p$$

or, la matrice A est supposée stable, ses valeurs propres sont de module inférieur à l'unité et par suite les pôles de $G(z)$ sont eux-mêmes de module inférieur à 1.

Considérons le numérateur de la fonction de transfert $G(z)$, en utilisant la propriété suivante (BALESTRA P. 1978)

$$\det(L + MN) = \det(L) + N \operatorname{adj}(L) M$$

où L est une matrice d'ordre (n,n) et M, N sont deux vecteurs d'ordre respectif $(n,1), (1,n)$.

Posons $L = ZI - A$; $M = K$ et $N = C$, il vient

$$C \operatorname{adj}(ZI - A)K + \det(ZI - A) = \det [ZI - (A - KC)]$$

La détermination des racines du numérateur c'est-à-dire des zéros de $G(z)$ consiste alors à rechercher les valeurs propres de la matrice $(A - KC)$.

Or, la matrice $(A - KC)$ est stable, ses valeurs propres sont de module inférieur à l'unité c'est-à-dire les zéros de $G(z)$ sont de module inférieur à 1.

En résumé, la fonction de transfert $G(z)$ a pour pôles les valeurs propres de la matrice A et pour zéros les valeurs propres de $(A - KC)$.

6.2. Représentation spectrale du modèle innovation.

Le spectre $G_{YY}(z)$ s'écrit

$$G_{YY}(z) = [C'(ZI - A)^{-1}K + I] E(e_t e_t') [K(Z^{-1}I - A)^{-1}C + I]$$

$$\begin{aligned} G_{YY}(z) &= [C'(ZI - A)^{-1}K + I] [C'PC + R] [K'(Z^{-1}I - A)^{-1}C + I] \\ &= B(z) B(z^{-1}) \end{aligned}$$

$$\begin{aligned} \text{avec } B(z) &= [C'(ZI - A)^{-1}K + I] [C'PC + R]^{1/2} \\ &= G(z)H^{1/2} \end{aligned}$$

où $G(z) = C'(ZI - A)^{-1}K + I$ est la fonction de transfert.

Le spectre $G_{YY}(z)$ s'écrit aussi

$$G_{YY}(z) = G(z) G_{ee}(z) G(z^{-1})$$

En remplaçant z par e^{iw} on obtient la densité spectrale du processus Y_t :

$$\rho(w) = \frac{1}{2\pi} B(e^{iw})B(e^{-iw}) = \frac{1}{2\pi} G(e^{iw}) H G(e^{-iw})$$

Ainsi nous venons de démontrer la proposition suivante :

Proposition 2.

La fonction de transfert $G_{yy}(z)$ est factorisable

$$G_{yy}(z) = B(z).B(z^{-1})$$

où $B(z)$ est la transformée en Z du processus Y_t

avec $B(z) = [I + C(ZI - A)^{-1}K] [C'PC + R]^{1/2}$

$B(z)$ est analytique à l'extérieur du cercle unité

tel que $\det Y(z) \neq 0$

Elle peut également être exprimée par

$$G_{yy}(z) = G(z) G_{ee}(z) G(z^{-1})$$

où $G(z)$ est la fonction de transfert

La densité spectrale de Y_t vaut

$$\rho(\omega) = \frac{1}{\pi} G(e^{i\omega}) H G(e^{-i\omega})$$

7. Transformation des processus ARMA en représentation d'état.**7.1. cas modèle AR(p).**

considérons le processus autorégressif d'ordre p suivant

$$A(L) y_t = \varepsilon_t$$

avec $A(L) = 1 - \sum_{i=1}^p a_i L^i$

Soient $Y(z)$ et $\varepsilon(z)$, les transformations respectives de y_t et ε_t en z , la fonction de transfert du processus vaut pour tout complexe z

$$g(z) = Y(z) / \varepsilon(z) = 1 / A(z).$$

Choix des variables d'états.

Nous avons déjà mentionné que le vecteur d'état joue un rôle fondamental parce qu'il spécifie complètement les informations transmises du passé vers le futur. Le choix des variables d'états permet alors de définir différents modèles.

$$\text{Posons } x_1 = y_t ; x_2 = y_{t-1} ; \dots ; x_p = y_{t-p+1}$$

$$\text{et } x'_t = (x_1, x_2, \dots, x_p)'$$
 le vecteur d'état.

La représentation du modèle d'état correspondante est déterminée

par le système d'équations suivantes :

équation d'évolution

$$x_t = Ax_{t-1} + B\varepsilon_t$$

équation d'observation.

$$y_t = Cx_t$$

où le vecteur d'état $x'_t = (y_t, y_{t-1}, \dots, y_{t-p+1})'$
(p,1)

$$\text{la matrice de transition } A = \begin{bmatrix} a_1 & a_2 & \dots & a_p \\ & & & 0 \\ & & I_{p-1} & \\ & & & 0 \end{bmatrix}$$

$$B' = (1, 0, 0 \dots 0)',$$

(p,1)

la matrice d'observations $C = (1, 0, 0, \dots 0)$ de dim(1,p).

L'équation de transition est une équation de différence de premier ordre de p équations obtenues par une simple transformation.

La transformation en z de l'équation de transition est

$$X(z) = (ZI - A)^{-1}B \varepsilon(z)$$

et la fonction de transfert du modèle d'état est

$$G(z) = Y(z) / \varepsilon(z) = C(ZI - A)^{-1}B.$$

On peut constater que $G(z) = g(z)$

$$\text{soit } C(ZI - A)^{-1}B = 1 / A(z).$$

La représentation d'état obtenue est appelée forme canonique commandable minimale, la dimension du vecteur d'état vaut p.

Ce modèle peut être généralisé en supposant que les paramètres du système sont aléatoires.

7.2. Cas modèle autorégressif à paramètres aléatoires.

Supposons que les paramètres $(a_i; i = 1, p)$ sont aléatoires, définis par

$$a_{it} = a_{it-1} + u_t$$

où u_t est une variable aléatoire de moyenne nulle et de variance constante σ_u^2 , dans ce cas les variables d'états sont définies par

$$x_1 = a_1; x_2 = a_2; \dots; x_p = a_p.$$

La représentation du modèle d'état est devenue :

l'équation de transition.

$$x_t = x_{t-1} + u_t^*$$

l'équation d'observation.

$$y_t = Cx_t + \varepsilon_t$$

où :

$$\text{le vecteur d'état } x_t' = (a_{1t}, a_{2t}, \dots, a_{pt})' \\ (p, 1)$$

$$\text{le bruit d'état } u_t^* = (u_{1t}, u_{2t}, \dots, u_{pt})' \\ (p, 1)$$

$$\text{la matrice d'observations } C = (y_{t-1}, y_{t-2}, \dots, y_{t-p}). \\ (1, p)$$

Cette représentation a été suggérée par Harrison et Stevens (1976) afin d'avoir la révision des estimations des paramètres.

De ces deux modèles, le choix du vecteur d'état dépend du contexte à étudier d'où l'intérêt de l'utilisation de la représentation d'état.

7.3. Cas modèle ARMA(p,q).

Soit le modèle ARMA(p,q) suivant :

$$a(L) y_t = b(L) \varepsilon_t$$

$$\text{où } a(L) = 1 - \sum_{i=1}^p a_i L^i$$

$$b(L) = \sum_{j=0}^q b_j L^j \quad \text{avec } b_0 = 1.$$

La transformation d'un modèle ARMA en représentation d'état n'est pas unique, nous considérons dans ce qui suit trois modèles équivalents.

a) Modèle innovation.

Sa représentation est définie par le système suivant :

équation d'évolution

$$x_t = Ax_{t-1} + B\varepsilon_t$$

équation d'observations.

$$y_t = Cx_t + \varepsilon_t$$

où

$$\text{le vecteur d'état } x_t = (x_{1t}, x_{2t}, \dots, x_{pt})'$$

la matrice de transition $A = \begin{bmatrix} a_1 & & & \\ a_2 & & & \\ \vdots & & & \\ a_p & & I_{p-1} & \end{bmatrix}$
 (p, p)

la matrice $B = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_p \end{bmatrix}$ telle que $d_j = a_j + b_j$ pour $j = 1, q$

la matrice d'observations $C = (1, 0, \dots, 0)$.

Preuve.

Supposons que $q \leq p$, le vecteur d'état est défini par

$x_t = (x_{1t}, x_{2t}, \dots, x_{pt})$ tel que

$$x_{1t} = y_t - \varepsilon_t$$

et

$$\begin{aligned} x_{2t-1} &= x_{1t} - a_1 y_{t-1} - b_1 \varepsilon_{t-1} \\ &= x_{1t} - a_1 (x_{1t-1} + \varepsilon_{t-1}) - b_1 \varepsilon_{t-1} \\ &= x_{1t} - a_1 x_{1t-1} - d_1 \varepsilon_{t-1} \end{aligned}$$

$$\text{avec } d_1 = a_1 + b_1$$

$$\text{d'où } x_{1t} = a_1 x_{1t-1} + x_{2t-1} + d_1 \varepsilon_{t-1}$$

En procédant de la même manière, on définit x_{j+1} par

$$\begin{aligned} x_{j+1,t-1} &= x_{jt} - a_j y_{t-1} - b_j \varepsilon_{t-1} \\ &= x_{jt} - a_j (x_{1t-1} + \varepsilon_{t-1}) - b_j \varepsilon_{t-1} \\ &= x_{jt} - a_j x_{1t-1} - d_j \varepsilon_{t-1} \end{aligned}$$

$$\text{avec } d_j = a_j + b_j$$

$$\text{d'où } x_{jt} = a_j x_{1t-1} + x_{j+1,t-1} + d_j \varepsilon_{t-1}$$

finalement nous avons donc

$$x_{pt} = a_p x_{1t-1} + x_{p+1,t-1} + d_p \varepsilon_{t-1} \quad \blacksquare \text{ C.Q.F.D.}$$

Nous allons montrer dans un cadre général qu'à partir d'une fonction de transfert on obtient des modèles de représentation d'état.

Effectuons la transformation en z du processus ARMA(p,q)

$$Y(z) - a_1 Z Y(z) - \dots - a_p Z^p Y(z) = \varepsilon(z) + b_1 Z \varepsilon(z) + \dots + b_q Z^q \varepsilon(z)$$

$$[1 - a_1 Z - \dots - a_p Z^p] Y(z) = [1 + b_1 Z + \dots + b_q Z^q] \varepsilon(z)$$

la fonction de transfert est définie par

$$g(z) = \frac{Y(z)}{\varepsilon(z)} = \frac{1 + b_1 Z + \dots + b_q Z^q}{1 - a_1 Z - \dots - a_p Z^p}.$$

Pour avoir la représentation d'état correspondante il faut déterminer les matrices A, B, et C vérifiant

$$x_t = A x_{t-1} + B u_t$$

$$y_t = C x_t$$

il suffit alors de vérifier l'égalité des fonctions de transfert des deux modèles.

La fonction de transfert du modèle d'état

$$G(z) = \frac{Y(z)}{u(z)} = C(ZI - A)^{-1}B$$

d'où $G(z) = g(z)$ implique

$$C(ZI - A)^{-1}B = \frac{1 + b_1 Z + \dots + b_q Z^q}{1 - a_1 Z - \dots - a_p Z^p}$$

b) Proposition 3.

Le processus linéaire stationnaire décrit par la fonction de transfert

$$g(z) = \frac{1 + b_1 Z + \dots + b_q Z^q}{1 - a_1 Z - \dots - a_p Z^p}$$

avec $p > q$, est équivalent au modèle d'état dit forme canonique

commandable minimale, suivant :

$$\begin{aligned} X_t &= A X_{t-1} + B_c V_t \\ Y_t &= C_c X_t \end{aligned}$$

où la dimension de l'état X_t vaut p

$$\text{la matrice de transition } A_{(p,p)} = \begin{bmatrix} & & & I_{p-1} \\ & & & \\ & & & \\ a_1 & a_2 & & a_p \end{bmatrix}$$

la matrice $B_c = (0, 0, \dots, 1)$

la matrice de sortie $C_c = (1, b_1, b_2, \dots, b_q, 0 \dots)$.

Preuve.

On doit vérifier l'égalité des fonctions de transfert des deux systèmes

$$\text{la matrice } ZI - A = \begin{bmatrix} Z-1 & 0 & & 0 \\ 0 & Z-1 & & 0 \\ & & & -1 \\ -a_1 & & & Z-a_p \end{bmatrix}$$

$$\det (ZI - A) = 1 - a_1 Z - \dots - a_p Z^p$$

$$\begin{aligned} \text{d'où } (ZI - A)^{-1} &= \frac{\text{adj}(ZI - A)}{\det (ZI - A)} \\ &= \frac{\text{adj}(ZI - A)}{1 - a_1 Z - \dots - a_p Z^p} \end{aligned}$$

Posons $\text{adj}(ZI - A) = E$ matrice quelconque de dimension (p,p)

Pour établir l'égalité avec la fonction de transfert

$$G(z) = C_c (ZI - A)^{-1} B_c$$

il suffit donc de vérifier que

$$1 + b_1 Z + \dots + b_q Z^q = C_c E B_c$$

Soit encore en effectuant le produit matriciel $E B$

$$1 + b_1 Z + \dots + b_q Z^q = (1, b_1, \dots, b_q, 0 \dots) (E_{1,p}, E_{2,p}, \dots, E_{p,p})'$$

or les mineurs $m_{p,j}$ de la dernière ligne de la matrice $ZI - A$ se calculent facilement, nous avons

$$E_{j,p} = (-1)^{p+j} m_{p,j} = (-1)^{p+j} Z^{j-1} (-1)^{p-j} = Z^{j-1}$$

donc, puisque $p > q$

$$C_c E B_c = (1, b_1, b_2, \dots, b_q, 0 \dots) \begin{bmatrix} 1 \\ Z \\ \vdots \\ Z^q \\ Z^{p-1} \end{bmatrix} = 1 + b_1 Z + \dots + b_q Z^q$$

C.Q.F.D.

c) Proposition 4.

Le processus linéaire stationnaire décrit par la fonction de transfert

$$g(z) = \frac{1 + b_1 Z + \dots + b_q Z^q}{1 - a_1 Z - \dots - a_p Z^p}$$

avec $p > q$ est équivalent à la représentation du modèle d'état dite forme canonique observable minimale

$$X_t = A_o X_{t-1} + B_o V_t$$

$$Y_t = C_o X_t$$

où la dimension de l'état X_t vaut p

$$\text{la matrice de transition } A_o = \begin{bmatrix} 0 & 0 & \dots & 0 & a_1 \\ & I_{p-1} & & & a_2 \\ & & & & \vdots \\ & & & & a_p \end{bmatrix}$$

$$\text{la matrice } B_o = (1, b_1, \dots, b_q, 0 \dots)$$

$$\text{la matrice } C_o = (0, 0, \dots, 1)$$

ces deux formes minimales sont duales, caractérisées par les relations suivantes : $A_c = A_o'$; $B_o = C_c'$ et $C_o = B_c'$.

Annexe.

Démonstration de la proposition 1.

Considérons d'une part un système dynamique en temps discret à n dimension décrit par les équations d'état suivantes

$$X_t = AX_{t-1} + BV_t \quad (1)$$

et d'autre part un vecteur d'observations à m dimension

$$Y_t = CX_t + U_t \quad (2)$$

L'estimateur de X est:

$$\hat{X}_{t/t} = E(X_t / y_t) = E(X_t / y_{t-1}, e_{t/t-1})$$

ou $E(\cdot / \cdot)$ est l'opérateur d'espérance mathématique conditionnelle

Appliquons la propriété de décomposition en somme directe (voir Ratsimalahelo (1992)).

$$\begin{aligned} \hat{X}_t &= E(X_t / y_{t-1}) + E(X_t / e_{t/t-1}) \\ &= E(X_t / y_{t-1}) + K_t (Y_t - \hat{Y}_{t/t-1}) \text{ qui vérifie (3.6.1)} \end{aligned}$$

avec $e_{t/t-1} = Y_t - \hat{Y}_{t/t-1}$ (l'innovation)

$$K_t = E(X_t e'_{t/t-1}) [E(e_{t/t-1} e'_{t/t-1})]^{-1} \text{ (la matrice de gain)}$$

Utilisons l'opérateur de l'espérance mathématique conditionnelle sachant l'ensemble d'informations y_{t-1} à l'équation d'évolution

$$\begin{aligned} E(X_t / y_{t-1}) &= AE(X_{t-1} / y_{t-1}) + B.E(V_t / y_{t-1}) \\ &= A\hat{X}_{t-1/t-1} \end{aligned} \quad (3)$$

on introduit les erreurs d'estimations

$$\varepsilon_{t/t-1} = (X_t - E(X_t / y_{t-1})) = A(X_{t-1} - \hat{X}_{t-1/t-1}) + BV_t$$

qui ne sont pas des processus d'innovation parce que $(X_{t_0}, \dots, X_{t-1}) \notin y_{t-1}$ dont la moyenne est nulle

$$E(\varepsilon_{t/t-1}) = E[X_t - E(X_t / y_{t-1})] = 0 \quad \forall t$$

et la covariance est

$$E(\varepsilon_{t/t-1} \varepsilon'_{t/t-1}) = E \left[[A(X_t - \hat{X}_{t-1}) + BV_t] [A(X_t - \hat{X}_{t-1}) + BV_t]' \right]$$

$$P_{t/t-1} = AP_{t-1}A' + BQB' \text{ qui vérifie (3.6.5)} \quad (4)$$

De la même manière en prenant l'espérance mathématique conditionnelle sachant l'ensemble d'informations $\mathcal{Y}_{t-1}$ de l'équation d'observations

$$E(Y_t / \mathcal{Y}_{t-1}) = CE(X_t / \mathcal{Y}_{t-1}) + E(U_t / \mathcal{Y}_{t-1})$$

$$\hat{Y}_{t/t-1} = C\hat{X}_{t/t-1} \quad (5)$$

introduisons le processus d'innovation

$$e_{t/t-1} = Y_t - C\hat{X}_{t/t-1}$$

$$\text{d'où } Y_t = C\hat{X}_{t/t-1} + e_{t/t-1} \text{ qui vérifie (3.6.2)} \quad (6)$$

le processus d'innovation s'écrit aussi

$$e_{t/t-1} = C(X_t - \hat{X}_{t/t-1}) + u_t = C\varepsilon_{t/t-1} + u_t$$

Sa propriété est qu'il est centré, blanc.

$$E(e_{t/t-1}) = E[C\varepsilon_{t/t-1} + u_t] = 0 \quad \forall t$$

$$E(e_{t/t-1} e'_{s/s-1}) = \begin{cases} H_t = CP_{t/t-1}C' + R & \forall t, s \quad s = t \\ 0 & \text{(d'après la propriété de la condition} \\ & \text{d'orthogonalité) qui vérifie (3.6.3)} \end{cases} \quad (7)$$

Calcul de la covariance croisée

$$E(X_t e'_{t/t-1}) = E(\varepsilon_{t/t-1} e'_{t/t-1}) = E \left[(X_t - \hat{X}_{t/t-1}) [C(X_t - \hat{X}_{t/t-1}) + u_t]' \right]$$

$$= E[(X_t - \hat{X}_{t/t-1})(X_t - \hat{X}_{t/t-1})'C']$$

$$+ E[(X_t - \hat{X}_{t/t-1})u_t']$$

$$= P_{t/t-1}C' \quad (8)$$

En utilisant les deux expressions précédentes on obtient la matrice de gain

$$K_t = P_{t/t-1} C' H_t^{-1} \quad \text{qui vérifie (3.6.4)} \quad (9)$$

Calcul de la matrice de variance de X_t conditionnelle à l'ensemble d'observation Y_t

$$\begin{aligned} P_t &= E(\varepsilon_t \varepsilon_t') = E(\varepsilon_{t/t-1} \varepsilon_{t/t-1}') - E(\varepsilon_{t/t-1} e_{t/t-1}') [E(e_{t/t-1} e_{t/t-1}')]^{-1} \\ &\quad E(\varepsilon_{t/t-1} e_{t/t-1}') \\ &= P_{t/t-1} - K_t C P_{t/t-1} \\ &= [I - K_t C] P_{t/t-1} \end{aligned} \quad (10)$$

compte tenu des expressions (4), (7) et (9) il vient

$$P_t = A P_{t-1} A' + B Q B' - P_{t/t-1} C' [C P_{t/t-1} C' + R]^{-1} C P_{t/t-1} \quad \text{qui vérifie (3.6.6)}$$

Bibliographie.

- AOKI Masanao, (1987a), "*State Space Modeling of Time Series*", Springer Verlag.
- AOKI Masanao, (1987b), "Studies of Economic interdependence by state space modeling of time series", *Annales d'Economie et de Statistique*, n°6-7, pp 225-252.
- AOKI Masanao, (1988) "On alternative state representation of time series models", *Journal of Economic Dynamics and Control*, Vol 12, pp 595-607.
- BALESTRA P., (1978), "Determinant and inverse of a sum of matrices with applications in economics and statistics", *Institut de Mathématiques Economiques*, ronéoté n°24, Dijon, Avril.
- BEVERIDGE S. et NELSON C. R. (1981), "A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle", *Journal of Monetary Economics*, Vol 7, pp 151-174.
- CLARK P. (1987): "The cyclical component of US economic activity" *Quarterly Journal of Economics*, pp 797-814.
- CLARK P. (1989): "Trend revision in real output and unemployment" *Journal of Econometrics*, pp 15-32.
- CAMPBELL J. Y. et MANKIW G. (1987a): "Are output fluctuations transitory", *Quarterly Journal of Economics*, pp 857-880.
- CAMPBELL J. Y. et MANKIW G. (1987b): "Permanent and transitory components in macroeconomic fluctuations" *Journal of Monetary Economics*, Vol 23, pp 319-334.
- COCHRANE J. H. (1988): "How big is the random walk in GNP" *Journal of Political Economy*, pp 893-920.
- DURBIN J., (1984), "Present position and potential developments : Time Series Analysis" *Journal of Royal Statistical Society A*, 147, pp 161-173.
- ENGEL R. et WATSON M., (1983), "Alternative algorithms for estimation of dynamic factor MIMIC and varying coefficient regression models", *Journal of Econometrics*, 23, pp 385-400.
- FAURRE P. et ROBIN M. (1984), "*Eléments d'automatique*", Dunod, Paris.
- GOURIEROUX C. et MONFORT A., (1990), "Séries temporelles et modèles dynamiques", Economica, Paris.

- HANNAN E. J., (1970), *"Multiple Time series"*, Wiley, New-York.
- HARRISON P. J. et STEVENS C. F., (1976), "Bayesian forecasting", *Journal of Royal Statistical Society A*, Vol 38, n° 3, pp 205-247.
- HARVEY A. C., (1984), "A unified view of statistical forecasting procedures" *Journal of Forecasting*, Vol 3, pp 245-275.
- HARVEY A. C., (1985), "Trends and cycles in macroeconomic time series", *Journal of Business and Economic Statistics*, Vol 3, n°3, pp 216-227.
- HENDRY David, (1984), "Time Series Econometrics", *Journal of Royal Statistical Society A*, 147, part 2, pp 327-339.
- KALMAN R. E., FALB P. L., ARIB M. A., (1969), *"Topics in mathematical system theory"*, Mc Graw -Hill.
- KUSHNER Harnold, (1971), *"Introduction to stochastic control"*, Holt Rinehart and Wiston.
- NELSON C. R. (1988): "Spurious trend and cycle in state space decomposition of time series with a unit root" *Journal of Economic dynamics and Control*, pp 475-488.
- RATSIMALAHELO Zaka, (1992), "La demande d'investissement avec coûts d'ajustement" thèse doctorat Université de Bourgogne 552p.
- WATSON M. W., (1986): " Univariate detrending methods with stochastic trends", *Journal of Monetary Economics*, 18, pp 45-75.