

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET LE
RECHERCHE SCIENTIFIQUE

=====

UNIVERSITÉ MENTOURI -CONSTANTINE
FACULTE DES SCIENCES EXACTES

=====

DÉPARTEMENT DE MATHÉMATIQUES

N° d'ordre:.....

N° de série:.....

MÉMOIRE PRÉSENTÉ POUR L'OBTENTION
DU
DIPLOME DE MAGISTÈRE
EN
MATHÉMATIQUES

THÈME:

*Le filtrage linéaire et non linéaire
à temps discret et à temps
continu et applications*

OPTION

Probabilités et Statistique

Par

El-Hadj Ali Thouria

Devant le jury:

Président	S.Boughaba	M.C (A)	Université Mentouri
Rapporteur	M.Bousseboua	Prof	Université Mentouri
Examineur	F.Messaci	Prof	Université Mentouri
Examineur	A.Bennia	Prof	Université Mentouri
Examineur	A.Bibi	Prof	Université Mentouri

Soutenu le: 14/ 06 /2012.

Remerciements

Je tiens tout d'abord à remercier Monsieur Bousseboua.M maître de conférence à l'Université Mentouri de Constantine pour sa disponibilité, sa rigueur et ses conseils et pour tout le temps qu'il m'avait consacré pendant la rédaction de ce mémoire .

Mes vifs remerciements vont à Melle Boughaba.S maître de conférence à l'Université Mentouri de Constantine pour l'honneur qu'elle me fait en présidant le jury de soutenance de ce mémoire.

J'exprime ma gratitude envers Mr Bibi.A Professeur à l'Université Mentouri de Constantine, pour avoir mis à ma disposition toute la documentation dont j'avais besoin pour accomplir ce travail.

Je remercie très sincèrement Mr Bennia.A Professeur à l'Université Mentouri de Constantine qui a été toujours présent pour répondre à mes questions concernant l'application.

J'exprime mes plus sincères remerciements à Mme Messaci.F Professeur à l'Université Mentouri de Constantine qui m'a fait l'honneur d'être examinatrice de cette thèse .

Je remercie énormément Mme Kebbabi.K qui ma beaucoup aidé à l'utilisation du logiciel Latex.

Mes remerciements vont également à ma famille, mes amis et mes collègues qui m'ont soutenue et encouragé.

Je garde le meilleur pour la fin, je remercie tout particulièrement mes parents de m'avoir soutenu, encouragé et d'avoir toujours été présents quand il fallait. Je leurs dédie ce mémoire.

Table des matières

Introduction	4
0 Rappels et définitions	7
0.1 Définitions et propriétés	8
0.1.1 Généralités sur les processus stochastiques	8
0.1.2 Processus à accroissements orthogonaux	11
0.1.3 Le mouvement brownien	13
0.1.4 Martingales	14
0.2 Calcul stochastique	17
0.2.1 L'intégrale stochastique de Wiener	17
0.2.2 L'intégrale stochastique de ITÔ	21
0.2.3 La formule de ITÔ et le théorème de représentation des mar- tingales	24
0.3 Equations différentielles stochastiques (E.D.S)	27
0.3.1 Les équations différentielles stochastiques au sens de Wiener	27
0.3.2 Equations différentielles stochastiques au sens de Itô	29
0.4 Processus de Markov	30
I Le filtrage	32
1 Le prédicteur linéaire optimal	36

1.1	Définition et propriétés du prédicteur linéaire	36
2	Le filtrage linéaire	40
2.1	Le filtrage linéaire à temps discret	41
2.1.1	Le filtre de Kalman	41
2.2	Le filtrage linéaire à temps continu	47
2.2.1	L'approche innovation	47
2.2.2	Le filtre de Kalman	51
2.2.3	Synoptique d'un filtre de Kalman	57
2.2.4	Le filtre de Kalman-Bucy,1961	57
3	Le filtrage non linéaire	61
3.1	Le filtrage non linéaire à temps discret	63
3.1.1	Le filtre de Kalman étendu (FKE)	65
3.2	Le filtrage non linéaire à temps continu	68
3.2.1	Méthode de changement de mesure	68
3.2.2	L'approche innovation	75
3.2.3	Exemple	80
II	Applications	83
4	Application du filtrage linéaire à temps continu au traitement de signal	85
4.0.4	Position du Problème	86
4.0.5	Application du filtre de Kalman-Bucy	87
5	Application du filtrage linéaire à temps discret à un problème de navigation et simulation	90
5.1	Description du problème	91
5.2	Application du filtre de Kalman	92

5.3	Simulation avec Matlab	92
6	L'application du filtrage à la finance	97
6.1	Position du problème	98
6.2	Application du filtrage non linéaire au modèle à volatilité stochastique	99
A	Théorème de Fubini stochastique	101

Introduction

La plupart des systèmes dynamiques physiques ou économiques sont le plus souvent sujets à des perturbations aléatoires. La modélisation de l'évolution dans le temps des états du système considéré s'écrit alors sous la forme d'un terme d'évolution déterministe et d'un terme aléatoire modélisant les perturbations aléatoires de la dynamique ainsi que l'incertitude sur le choix du terme déterministe. L'équation de la dynamique traduit en fait une connaissance à priori que l'on a du système. Afin de déterminer l'état précis du système, on conçoit une procédure d'observation permettant de recueillir des mesures des états du système. Ces observations sont en général entachées d'erreurs dues à l'imperfection de l'instrument de mesures. La suite des observations est modélisée par une équation reliant l'observation à l'état courant du système et bruitée par un processus stochastique connu modélisant l'imperfection des observations.

Le problème du filtrage peut être abordé suivant soit l'approche fréquentielle (filtre de Wiener), soit l'approche temporelle (filtre de Kalman). Ces deux approches ont la même finalité et consistent à élaborer une procédure (filtre) d'estimation optimale de l'état du système dynamique.

Le filtre est un schéma récursif consistant à chaque itération d'une prédiction (estimation à priori) de l'état du système suivie d'une mise à jour de l'estimation de l'état (estimation à posteriori). Cette procédure est reproduite chaque fois que l'on introduit une nouvelle observation.

L'approche fréquentielle initiée par Wiener (1946) pour le filtrage linéaire à temps continu et ensuite par Kolmogorov pour le filtrage linéaire à temps discret, est confinée comme on peut s'y attendre, aux systèmes dynamiques stationnaires. Ce filtre peut être regardé comme un cas particulier du filtre de Kalman qui s'étend au cas de systèmes dynamiques non stationnaires multivariés.

Et en partie, c'est à partir des travaux de Kalman (1960) avec la conception du filtre qui porte son nom, que la théorie du filtrage a connu un développement

important. Cet essor est dû en grande partie à ses innombrables applications dans divers domaines tels le traitement de signal [8], la finance [35], la navigation et le guidage de véhicules spatiaux [6], etc...

L'objet de ce mémoire, porte sur l'estimation dans les espaces d'états et privilégie l'approche temporelle du problème du filtrage. On étudiera successivement le problème du filtrage linéaire et non linéaire à temps discret et à temps continu.

Une solution au problème du filtrage linéaire à temps discret est fournie par le filtre de Kalman et le problème du filtrage linéaire à temps continu ne fut résolu qu'une année plus tard par Kalman et Bucy (1961) avec des bruits gaussiens. Le problème du filtrage étendu aux systèmes dynamiques non linéaires n'a été résolu que vers la fin des années 60 par Stratonovich et Kushner [1967]. Dans ce cas, le filtre est une solution d'une équation non linéaire aux dérivées partielles stochastiques.

L'estimation dans les systèmes dynamiques linéaires et non linéaires à temps continu est basée sur le calcul stochastique qui fait l'objet d'un chapitre introductif. Nous rappelons dans ce chapitre quelques notions fondamentales de la théorie des probabilités et une introduction succincte sur le calcul stochastique et les équations différentielles stochastiques qui s'avèrent un outil nécessaire et incontournable pour la description du modèle d'espace d'état à temps continu.

Notre travail comporte deux parties consacrées respectivement aux aspects théoriques de l'estimation dans les systèmes dynamiques (problème de filtrages) et ses applications. Dans la première partie on commence par rappeler dans un premier chapitre la notion de prédicteur linéaire optimal, suivi de deux chapitres dévolus chacun au problème du filtrage linéaire et non linéaire à temps discret (le filtre de Kalman) et à temps continu. Dans le cas linéaire discret, nous considérons le cas particulier de bruits décorrelés et comment nous ramener à ce cas dans le cas de bruits corrélés. Dans le cas de système linéaire à temps continu, moyennant l'intégrale de Wiener on déduit un estimateur de l'état en intégrant par rapport aux processus des innovations. L'erreur du filtrage dans ce cas satisfait l'équation de

Riccati. En guise d'exemple et pour conclure ce second chapitre, nous reproduisons le filtre de Kalman-Bucy en considérant des bruits gaussiens.

Dans le cas non linéaire à temps discret, une première approche consiste en utilisant la formule de Bayes, à déterminer la densité conditionnelle du filtrage laquelle satisfait une équation intégrale non linéaire. Il s'avère que cette dernière est en générale difficile à résoudre, aussi on adopte une approche plus naturelle appelée le filtre de Kalman étendu et qui consiste à linéariser d'abord le système d'espace d'états et d'appliquer ensuite le filtre de Kalman.

Le problème du filtrage non linéaire à temps continu a été abordé selon deux approches : l'approche changement de mesure et l'approche innovation. L'approche changement de mesure basée sur le théorème de Girsanov, permet de trouver une solution qui satisfait une équation aux dérivées partielles stochastique linéaire (équation de Zakai). L'approche innovation, comme pour le cas linéaire, s'appuyant sur l'intégrale stochastique de ITÔ par rapport au processus des innovations, nous permet de définir une solution qui satisfait une équation aux dérivées partielles stochastique non linéaire appelée équation de Kushner. On remarque que l'approche innovation appliquée aux modèles linéaires permet de retrouver le filtre de Kalman-Bucy.

La seconde partie est consacrée aux applications du filtrage linéaire et non linéaire et comporte trois applications. La première est une application du filtrage linéaire au traitement de signal. On utilise le filtre de Kalman-Bucy pour résoudre le problème du circuit RL et trouver l'intensité du courant électrique. La seconde porte sur une simulation d'un problème de navigation en appliquant le filtre de Kalman à temps discret et dans la dernière nous esquissons brièvement une application du filtrage à la finance (volatilité stochastique).

Chapitre 0

Rappels et définitions

Nous rappelons dans ce chapitre les notions de mathématiques de base pour un développement exhaustif et rigoureux de la théorie du filtrage. Ces résultats ont été puisés dans différentes références de base sur la théorie du processus et du calcul stochastique telles que [4], [14] et [36].

0.1 Définitions et propriétés

0.1.1 Généralités sur les processus stochastiques

Définition 1 Soit \mathbb{T} une partie de \mathbb{R} .

On appelle processus stochastique réel (resp. complexe) une famille de variables aléatoires $(X_t : t \in \mathbb{T})$ définies sur un espace de probabilité (Ω, \mathcal{F}, P) et à valeurs dans \mathbb{R} (resp. \mathbb{C}).

Une réalisation d'un processus aléatoire est dite une trajectoire.

Définition 2 Un processus aléatoire $(X_t : t \in \mathbb{T})$ à valeurs réelles ou complexes, est dit d'ordre k ($k \geq 1$) si pour tout $t \in \mathbb{T}$, $\mathbb{E}(|X_t|^k) < +\infty$, où $|x|$ désigne la valeur absolue (resp. le module) de x .

On note $L^k((\Omega, \mathcal{F}, P))$ ¹ l'espace de toutes les variables aléatoires de puissance $k^{\text{ième}}$ intégrable.

Définition 3 Soit $(X_t : t \in \mathbb{T})$ un processus réel d'ordre 1.

On appelle fonction moyenne du processus (X_t) , la fonction m de \mathbb{T} dans \mathbb{R} définie par

$$\forall t \mapsto m(t) = \mathbb{E}(X_t).$$

Le processus $(X_t : t \in \mathbb{T})$ est dit centré si sa fonction moyenne est la fonction nulle.

Définition 4 Soit $(X_t : t \in \mathbb{T})$ un processus réel d'ordre 2.

On appelle fonction de covariance (ou d'auto-covariance) du processus (X_t) , l'application Γ de \mathbb{T}^2 dans \mathbb{R} et qui associe à chaque couple (s, t) de \mathbb{T}^2 , la covariance des variables X_s et X_t :

$$\Gamma(s, t) = \mathbb{E}(X_t \cdot X_s) - m(t) \cdot m(s)$$

1. En fait, $L^k((\Omega, \mathcal{F}, P))$ est l'espace des classes d'équivalence des variables de puissance $k^{\text{ième}}$ intégrables : $\hat{X} = \{Y : \int_{\Omega} |X(\omega) - Y(\omega)|^k dP(\omega) = 0\}$.

Si le processus est complexe, la fonction de covariance Γ s'écrit

$$\Gamma(s, t) = \mathbb{E} \left\{ (X_s - m(s)) \overline{(X_t - m(t))} \right\}.$$

Processus stationnaires au second ordre

Définition 5 Soit $\mathbf{X} = (X_t : t \in \mathbb{T})$ un processus du 2nd ordre.

On dit que le processus \mathbf{X} est stationnaire au second ordre, si sa fonction moyenne m est constante et sa fonction de covariance Γ est invariante par translation dans le temps :

$$\Gamma(s, t) = \Gamma(s + h, t + h)$$

quels que soient t, s et $h \in \mathbb{T}$.

La fonction d'autocovariance Γ d'un processus stationnaire à l'ordre 2 ne dépend que de l'écart temporel $t - s$, c'est-à-dire fonction de $t - s$, aussi on note γ la fonction

$$\gamma(t - s) = \Gamma(s, t)$$

qu'on identifiera avec la fonction de covariance Γ du processus.

On vérifie qu'une fonction de covariance γ d'un processus stationnaire au second ordre est une fonction paire (resp. hermitienne) selon que le processus est réel ou complexe, bornée et est du type défini non négatif au sens que

$$\sum_{j=1}^p \sum_{l=1}^p \alpha_j \gamma(t_j - t_l) \overline{\alpha_l} \geq 0$$

pour tout $p \geq 1$ et pour tous $t_1, \dots, t_p \in \mathbb{T}$ et $\alpha_1, \dots, \alpha_p \in \mathbb{C}$.

Définition 6 Un processus $(X_t : t \geq 1)$ à valeurs dans \mathbb{R} , est dit un processus gaussien, si toutes les combinaisons linéaires finies du processus $(X_t : t \geq 1)$ suivent une loi normale.

Ainsi, pour tout $n \geq 1$ et $t_1, t_2, \dots, t_n \in \mathbb{N}^*$, la loi de la variable aléatoire $\sum_{j=1}^n \alpha_j X_{t_j}$ est gaussienne et ce pour tous $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Autrement dit, tous les vecteurs finis

dimensionnelles $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ sont gaussiens. Or, nous savons que la loi jointe $P_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}$ est entièrement caractérisée par sa fonction caractéristique $\varphi_{t_1, \dots, t_n}$ définie par :

$$\varphi_{t_1, \dots, t_n}(\theta) = \exp\left(i \langle \theta, m_{\mathbf{t}} \rangle - \frac{1}{2} \theta' \Lambda_{\mathbf{t}} \theta\right)$$

où $\Lambda_{\mathbf{t}}$ est la matrice de covariance du vecteur $(X_{t_1}, X_{t_2}, \dots, X_{t_n})'$, $m_{\mathbf{t}}$ sa fonction moyenne et $\mathbf{t} = (t_1, \dots, t_n)$.

Si de plus $\Lambda_{\mathbf{t}}$ est inversible, alors le vecteur $(X_{t_1}, X_{t_2}, \dots, X_{t_p})'$ admet une densité de probabilité f donnée par

$$f(x_1, \dots, x_p) = \frac{1}{\sqrt{(2\pi)^p} \sqrt{|\Lambda_{\mathbf{t}}|}} \exp\left\{-\frac{1}{2} \mathbf{x}' \Lambda_{\mathbf{t}}^{-1} \mathbf{x}\right\} \quad (1)$$

où $\mathbf{x} = (x_1, x_2, \dots, x_p)'$ et $|\Lambda_{\mathbf{t}}|$ le jacobien de la matrice de covariance $\Lambda_{\mathbf{t}}$

Définition 7 On appelle processus de bruit blanc, un processus $(\varepsilon(t) : t \in \mathbb{Z})$ du second ordre, centré et deux à deux non corrélés.

La fonction de covariance γ d'un bruit blanc $(\varepsilon(t) : t \in \mathbb{Z})$ s'écrit

$$\gamma(t) = \begin{cases} \sigma^2 & \text{si } t = 0 \\ 0 & \text{si } t \neq 0 \end{cases}$$

Et si $n \geq 1$ et t_1, \dots, t_n une suite croissante d'entiers, alors la matrice de covariances $\Lambda(n)$ du vecteur $(\varepsilon(t_1), \dots, \varepsilon(t_n))'$ est diagonale : $\Lambda(n) = \sigma^2 \mathbf{I}_n$ où \mathbf{I}_n est la matrice unité d'ordre n . Ceci nous permet de déduire le résultat presque évident suivant :

Proposition 1 Un bruit blanc $(\varepsilon(t) : t \in \mathbb{Z})$ est gaussien si et seulement si, ses marginales unidimensionnelles sont gaussiennes, auquel cas tout vecteur $(\varepsilon(t_1), \dots, \varepsilon(t_n))'$ n dimensionnel admet une densité de probabilité f donnée par

$$f(x_1, \dots, x_n) = \frac{1}{\sqrt{2\pi}^n \sigma^n} \exp\left\{-\frac{1}{\sigma^2} \sum_{i=1}^n x_i^2\right\}.$$

Enfin nous rappelons le résultat qui d'ailleurs est aisé à démontrer :

Proposition 2 Un processus gaussien est stationnaire au second ordre si et seulement si, sa fonction moyenne est constante et sa fonction de covariance est invariante par translation dans le temps.

Il est tout aussi facile de voir que tout bruit blanc de variance constante est stationnaire et il en est encore de tout bruit blanc gaussien de variance constante. Un tel bruit blanc gaussien forme une suite de variables aléatoires indépendantes et équidistribuées selon une loi normale centrée et de variance σ^2 .

0.1.2 Processus à accroissements orthogonaux

On désigne par \mathcal{H}_0 le sous espace de Hilbert de $L^2(\Omega, \mathcal{F}, P)$ ² formé de variables centrées.

Définition 8 *Un processus $\{X_t, t \geq 0\} \subset \mathcal{H}_0$, est un processus à accroissements orthogonaux (p.a.o) si pour tous instants s, s', t, t' tels que $s < t \leq s' < t'$ les variables $X_t - X_s$ et $X_{t'} - X_{s'}$ sont orthogonales :*

$$\langle X_t - X_s, X_{t'} - X_{s'} \rangle = \int_{\Omega} (X_t - X_s) \cdot (\overline{X_{t'} - X_{s'}}) dP = 0.$$

Remarque 1 *Si $X \in L^2(\Omega, \mathcal{F}, P)$ et si (X_t) est un p.a.o, alors il en est aussi du processus (Y_t) défini par $Y_t = X_t + X$.*

Proposition 3 *Si $\{X_t, t \geq 0\}$ est un p.a.o de fonction de covariance r . La fonction réelle Ψ définie sur \mathbb{R}_+ par $\Psi(t) = \|X_t - X_0\|_2^2$ est une fonction non décroissante et on a*

$$r(t, s) = \Psi(t \wedge s)$$

où $\Psi(1) = \sigma^2$, $t \wedge s = \min(t, s)$ et $\|\cdot\|_2$ la norme dans $L^2(\Omega, \mathcal{F}, P)$.

Démonstration

Cette proposition découle directement de l'orthogonalité des accroissements et de ce que toute variable X_t , peut s'écrire

$$X_t - X_0 = (X_t - X_s) + (X_s - X_0)$$

pour tout $s : 0 < s < t$. ■

2. On rappelle que la forme $\langle X, Y \rangle = \int_{\Omega} X \overline{Y} dP$ est un produit scalaire sur $L^2(\Omega, \mathcal{F}, P)$.

Définition 9 *Un processus $(X_t, t \geq 0)$ est un processus à accroissements orthogonaux stationnaires si c'est un p.a.o et si le processus des accroissements $(X_t - X_s : s, t \geq 0)$ est stationnaire au second ordre.*

Définition 10 *Un processus $(X_t, t \geq 0)$ est un processus à accroissement indépendantes si les accroissements $(X_t - X_s)$ et $(X_{t'} - X_{s'})$ sont indépendantes quelques soient s, s', t, t' tels que $s < t \leq s' < t'$.*

Un processus à accroissements indépendants est un p.a.o mais la réciproque est fautive en général, sauf si le p.a.o est gaussien.

Proposition 4 *Un processus $(X_t, t \geq 0)$ continue en moyenne quadratique (m.q) est à accroissements orthogonaux stationnaires si et seulement si, sa fonction de covariance est de la forme :*

$$r(t, s) = \sigma^2 t \wedge s \quad (2)$$

Démonstration

L'implication directe résulte de la proposition (3) ci-dessus. Réciproquement, si la fonction de covariance r est de la forme (2), alors celle ci est continue en tout point $(s, t) \in \mathbb{R}_+^2$, ce qui assure selon le lemme (1) ci-dessous que le processus (X_t) est continu en m.q. D'autre part, la bilinéarité du produit scalaire et la forme de la fonction de covariance permettent de conclure que le processus (X_t) est un p.a.o stationnaire. ■

Lemme 1 *(Davis M.H. [15])*

Pour qu'un processus (X_t) soit continu en m.q en tout point $t \in \mathbb{R}_+$, il faut et il suffit que sa fonction de covariance soit continue sur la diagonale $\Delta = \{(t, t) : t \in \mathbb{R}_+\}$.

Processus à accroissements orthogonaux multidimensionnels

Définition 11 *Un processus vectoriel $(X(t) : t \geq 0)$ à valeurs dans \mathbb{R}^n , est un processus à accroissements orthogonaux (resp. p.a.o.s) si pour tout $j : 0 \leq j \leq n$, les*

processus $(X_j(t) : t \geq 0)$, $j = \overline{1, n}$ où $X_j(t)$ est la $j^{\text{ième}}$ composante du processus $(X(t) : t \geq 0)$ sont des p.a.o (resp. p.a.o.s).

La fonction de covariance $R(t, s)$ d'un p.a.o vectoriel $(X(t) : t \geq 0)$, satisfait

$$R(t, s) = \Gamma(t \wedge s)$$

où $\Gamma(t)$ est la matrice de variance covariance du vecteur $X(t)$.

$(X(t) : t \geq 0)$ est un p.a.o.s si cette fonction de covariance est de la forme

$$R(t, s) = \Gamma t \wedge s$$

où Γ est une matrice définie non négative. Si de plus, Γ est diagonale, alors les processus coordonnées $(X_j(t) : t \geq 0, j = \overline{1, n})$ engendrent des espaces orthogonaux.

Ci-dessous nous étudions un cas particulier de processus à accroissements indépendants et stationnaires.

0.1.3 Le mouvement brownien

Définition 12 On appelle mouvement brownien (m.b, ou processus de Wiener) un processus gaussien $\{W_t, t \geq 0\}$ centré, à accroissements indépendants et stationnaires.

Le mouvement brownien $\{W_t, t \geq 0\}$ issu de 0 : $W_0 = 0$ et de variance $E(W_1)^2 = 1$ est dit le mouvement brownien standard.

Remarque 2 1. Les trajectoires du m.b sont presque sûrement continues.

Réciproquement, on montre que tout processus réel continu $(W(t) : t \geq 0)$ à accroissements indépendants est un mouvement brownien.

2. le mouvement brownien n'est différentiable presque sûrement en aucun point $t \geq 0$.

Sa dérivée au sens des distributions est un bruit blanc gaussien.

Formellement, le mouvement brownien standard peut être regardé comme l'intégrale d'un b.b gaussien de variance infinie :

$$W(t) = \int_0^t \varepsilon_t ds,$$

pour tout t positif.

Définition 13 (*Mouvement brownien à valeurs dans \mathbb{R}^m*)

Un processus stochastique $(W(t) : t \geq 0)$ à valeurs dans \mathbb{R}^m est un mouvement brownien si pour tout $i : i = \overline{1, m}$, le processus $(W_i(t) : t \geq 0)$ où $W_i(t) = \pi_i(W(t))$ et π_i la $i^{\text{ème}}$ projection, est un mouvement brownien et les processus $(W_i(t))$, $i = \overline{1, m}$ sont mutuellement indépendants.

On montre facilement le résultat suivant :

Proposition 5 *Un mouvement brownien vectoriel $\{W_t\}$ est un processus vectoriel gaussien de covariance $I_m(t \wedge s)$ si ses composantes $\{W_i(t)\}$ sont des mouvements browniens indépendants.*

0.1.4 Martingales

Définition 14 *On appelle filtration sur (Ω, \mathcal{F}) , toute famille croissante $(\mathcal{F}_t : t \geq 0)$ de sous tribus de \mathcal{F} .*

Soit $(X_t : t \geq 0)$ un processus aléatoire, alors la famille des sous σ -algèbres $(\mathcal{F}_t^X : t \geq 0)$ où $\mathcal{F}_t^X = \sigma(X_s : 0 \leq s \leq t)$ (la σ -algèbre générée par les variables $X_s : 0 \leq s \leq t$) est une filtration de (Ω, \mathcal{F}) , appelée filtration naturelle associée au processus $(X_t : t \geq 0)$.

Définition 15 *Un processus $(X_t)_{t \geq 0}$ est dit adapté à la filtration $(\mathcal{F}_t)_{t \geq 0}$, si pour tout $t \geq 0$, la variable X_t est \mathcal{F}_t -mesurable.*

Définition 16 Un processus $(X_t, t \geq 0)$ \mathcal{F} -adapté est dit une \mathcal{F} -martingale si pour tout $t \geq 0$, X_t est P - intégrable et

$$E(X_t/\mathcal{F}_s) = X_s \text{ p.s.} \quad (3)$$

pour tout t et pour tout $s : s \leq t$.

Définition 17 Un processus $(X_t, t \geq 0)$ \mathcal{F} -adapté est dit une \mathcal{F} -sous martingale (resp. \mathcal{F} -sur martingale) si pour tout $t \geq 0$, X_t^+ (resp. X_t^-) est P - intégrable et

$$E(X_t/\mathcal{F}_s) \geq X_s \text{ p.s.} \quad (4)$$

resp.

$$E(X_t/\mathcal{F}_s) \leq X_s \text{ p.s.} \quad (5)$$

pour tout t et pour tout $s : s \leq t$.

Il est aisé de vérifier les assertions suivantes :

1. Un mouvement brownien $(X_t)_{t \geq 0}$ est une martingale par rapport à sa filtration naturelle \mathcal{F}
2. (Inégalité de Jensen)
L'image $(\varphi(X_t) : t \geq 0)$ d'une \mathcal{F} - martingale $(X_t : t \geq 0)$ par une fonction convexe mesurable φ , est une \mathcal{F} - sous martingale.
3. Si $(X_t)_{t \geq 0}$ est un mouvement brownien, alors le processus $X_t^2 - t$ est une \mathcal{F} - martingale par rapport à la filtration naturelle \mathcal{F} associée au mouvement brownien (X_t) .

Définition 18 Soit $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}_+})$ un espace de probabilité filtré. On appelle σ -algèbre prévisible associée à la filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$, la σ -algèbre sur $\mathbb{R}_+ \times \Omega$ générée par toutes les parties de la forme $\{0\} \times A$, $A \in \mathcal{F}_0$ et $]a, b] \times A$ et $0 \leq a < b < \infty$, $A \in \mathcal{F}_a$.

Définition 19 *Un processus à valeurs réelles $(X_t)_{t \in \mathbb{R}_+}$ est dit **prévisible** par rapport à la filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ (ou \mathcal{F}_t - prévisible) s'il est mesurable par rapport à la σ -algèbre prévisible généré par cette filtration.*

On a le

Lemme 2 *(Vincenzo (2005) [36]) Un processus est prévisible si et seulement si, il est mesurable par rapport à la plus petite σ -algèbre sur $\mathbb{R}_+ \times \Omega$ généré par les processus $(\mathcal{F}_{t^-} : t > 0)$ adaptés et continues à gauche.*

On termine cette section par le rappel de mesures absolument continues.

Définition 20 *Soit (Ω, \mathcal{F}, P) un espace de probabilité. Une probabilité Q sur (Ω, \mathcal{F}) est dite absolument continue par rapport à P et on note $Q \ll P$, si :*

$$\forall A \in \mathcal{F} : P(A) = 0 \Rightarrow Q(A) = 0.$$

On dit que P et Q sont équivalentes si $Q \ll P$ et $P \ll Q$.

0.2 Calcul stochastique

Dans ce paragraphe on cherche à définir l'intégrale du type :

$$\int_0^t f(s, \omega) dW(s) \quad (6)$$

où $W(\cdot)$ est un processus à accroissements orthogonaux et $f(s, \omega)$ une fonction déterministe ou aléatoire.

Lorsque f est déterministe, l'intégrale (6) est dite **intégrale de Wiener** et lorsque f est un processus aléatoire et W un mouvement brownien, cette intégrale est dite **intégrale stochastique de Itô** .

Ci-dessous on essaie de décrire brièvement chacune de ces deux intégrales.

0.2.1 L'intégrale stochastique de Wiener

Dans ce paragraphe, on suppose que les variables aléatoires sont centrées et de carré intégrables.

Soit $\{X_t, t \geq 0\}$ un processus à accroissements orthogonaux stationnaires et tel que $X_0 = 0$ et $E(X_1^2) = 1$. Il s'agit de définir l'intégrale $\int_0^t g(s) dX_s$ où g est une fonction de carré intégrable sur $[0, t]$.

Par analogie avec la construction de l'intégrale de Riemann - Stieltjes, l'intégrale de Wiener par rapport au p.a.o (X_t) , est d'abord définie pour les fonctions simples $g(t) = \sum_{i=0}^{n-1} c_i 1_{[t_i, t_{i+1}[}$ en posant

$$\int_0^\infty g(s) dX_s = \sum_{i=0}^{n-1} c_i (X_{i+1} - X_i)$$

et ensuite par passage à la limite, on définit l'intégrale de Wiener pour toute fonction g de carré intégrable sur $[0, t]$, pour tout $t \geq 0$.

On note $\int_0^t g(t) dX(t)$ l'intégrale de Wiener de g par rapport au p.a.o.s (X_t) .

Cette intrégrale définit une variable aléatoire centrée appartenant au sous espace de Hilbert \mathcal{H}_t^X généré par les variables $\{X_s : 0 \leq s \leq t\}$. De plus, il est facile de vérifier

que la transformation M définie sur $L^2[0, t]$ par

$$M(g) = \int_0^{+t} g(s) dX_s$$

est une *isométrie linéaire*.

Cette intégrale se prolonge de façon naturelle aux fonctions $g \in L^2[0, +\infty[$. A cette fin, si g_n est la restriction de g à l'intervalle $[0, t_n[$ où (t_n) est une suite de nombres réelles positifs convergeant vers $+\infty$ et si la suite $(M(g_n))$ est convergente, alors sa limite notée $\int_0^{+\infty} g(s) dX_s$ est dite aussi l'intégrale de Wiener de g par rapport au p.a.o.s (X_t) sur l'intervalle $[0, +\infty[$.

L'intégrale stochastique de Wiener permet une caractérisation des espaces de Hilbert \mathcal{H}_t^X et \mathcal{H}^X le sous espace de Hilbert généré par le processus (X_t) . On montre ([21]) que

1. $\mathcal{H}^X = \{\int_0^{+\infty} g(s) dX_s : g \in L^2[0, +\infty[]\}$
2. $\mathcal{H}_t^X = \{\int_0^t g(s) dX_s : g \in L^2[0, t]\}$ pour tout $t > 0$.

Proposition 6 *Pour qu'un processus (Z_t) du 2nd ordre continue en m.q puisse s'écrire à tout instant t sous la forme d'une intégrale de Wiener par rapport à un p.a.o.s (X_t) (i.e. $Z_t = \int_0^t g(s) dX_s$ où $g \in L^2([0, t]) \forall t > 0$), il faut et il suffit que pour tout s et $t \in \mathbb{R}^+ : s \leq t$, on ait*

$$Z_t \in \mathcal{H}_t^X \quad \text{et} \quad (Z_t - Z_s) \perp \mathcal{H}_s^X.$$

Démonstration

S'il existe $g \in L^2([0, +\infty[)$ telle que pour tout $t : Z_t = \int_0^t g(s) dX_s$, alors $Z_t \in \mathcal{H}_t^X$ et de plus l'identité

$$\begin{aligned} Z_t - Z_s &= \int_0^t g(u) dX_u - \int_0^s g(u) dX_u \\ &= \int_s^t g(u) dX_u \end{aligned}$$

montre que $Z_t - Z_s$ est orthogonale à \mathcal{H}_s^X .

Réciproquement, si $Z_t \in \mathcal{H}_t^X$ pour tout t et $(Z_t - Z_s) \perp \mathcal{H}_s^X$ pour $s \leq t$, ceci exprime

d'une part, que le processus Z_t est un p.a.o et d'autre part, si $\mathcal{P}_{\mathcal{H}_s^X}$ est le projecteur dans \mathcal{H}_s^X , alors

$$\mathcal{P}_{\mathcal{H}_s^X} \left(\int_0^t g(u) dX_u \right) = \int_0^s g(u) dX_u \quad (7)$$

Comme $Z_t \in \mathcal{H}_t^X$, alors pour tout $t > 0$, il existe une fonction $g(t, \cdot) \in L^2([0, t])$ telle que

$$Z_t = \int_0^t g(t, u) dX_u$$

Il suffit de montrer que la fonction $g(t, \cdot)$ est indépendante de t .

Or, l'orthogonalité de $(Z_t - Z_s)$ avec \mathcal{H}_s^X pour tous $s \leq t$ fait que $\mathcal{P}_s Z_t = Z_s$, ce qui assure compte tenu de ce que l'intégrale de Wiener préserve le produit scalaire, que

$$\int_0^s (g(t, u) - g(s, u))^2 du = 0$$

Ceci exprime que les fonctions $g(t, u)$ et $g(s, u)$ sont équivalentes dans $L^2([0, s])$.

Par suite, si on définit par exemple $g(u) = g(n, u), \forall u \in]n-1, n], n = 1, 2, \dots$, alors pour tout t , $g(\cdot)$ et $g(t, \cdot)$ sont équivalentes dans $L^2([0, t])$, d'où le résultat. ■

Ainsi l'intégrale de Wiener est un p.a.o et si de plus le p.a.o est un mouvement brownien, alors le processus de Wiener est aussi un mouvement brownien.

L'intégrale de Wiener multidimensionnelle

Soit $(W_t, t \geq 0)$ un processus vectoriel à accroissements orthogonaux de dimension m de fonction de covariance $I_m(t \wedge s)$. Pour une matrice B de dimension $n \times m$ dont les éléments sont fonctions du temps t , l'intégrale de Wiener par rapport à (W_t) est le processus vectoriel d'ordre n , dont la $i^{\text{ème}}$ composante est définie par

$$X_t^i = \sum_j \int_0^t b_{ij} dW_s^j$$

où b_{ij} sont les coefficients de la matrice B , cette intégrale existe ssi les coefficients b_{ij} sont de carrée intégrables. On note $X_t = \int_0^t B(s) dW_s$ cette intégrale.

Proposition 7 Soit (W_t) un p.a.o vectoriel de fonction de covariance $(\Gamma \cdot t \wedge s)$ où Γ est une matrice définie non négative et soit (X_t) et (Y_t) deux processus définis par

$X_t = \int_0^t B(s)dW_s$ et $Y_t = \int_0^t C(s)dW_s$ où B et C sont deux matrices d'ordre $n \times m$ et $r \times m$ respectivement dont les coefficients sont de carrée intégrables. On a

$$\begin{cases} E(X_t) &= E(Y_t) = 0 \\ Cov(X_t, Y_s) &= E(X_t Y_s') = \int_0^{t \wedge s} B(u)\Gamma(u)C'(u)du \end{cases}$$

Démonstration

Il est évident que (X_t) et (Y_t) sont centrés (l'intégrale de Wiener est centrée).

Pour calculer la covariance on pose $\Gamma = \mathbb{I}_m$. On a $(X_s), (Y_s) \in \mathcal{H}_s^W$ et $\int_s^t B(u)dW_u \perp \mathcal{H}_s^W$, alors pour tout $s < t$ on obtient

$$\begin{aligned} Cov(X_t, Y_s) &= E \left[\left(\int_0^s B(u)dW_u + \int_s^t B(u)dW_u \right) \left(\int_0^s C(u)dW_u \right) \right] \\ &= Cov(X_s, Y_s) \end{aligned}$$

et la préservation du produit scalaire permet d'écrire

$$\begin{aligned} E(X_s^i Y_s^j) &= E \left(\sum_{p=1}^n \int_0^s b_{ip}(u)dW_u^p \sum_{p=1}^n \int_0^s c_{jp}(u)dW_u^p \right) \\ &= \sum_p \int_0^s b_{ip}(u)c_{jp}(u)du \\ &= \int_0^s (B(u)C'(u))_{ij}du. \end{aligned}$$

Si $\Gamma \neq \mathbb{I}_m$, il est commode d'écrire $\Gamma = DD'$ car Γ est définie non négative et si on pose $W_t'(u) = D^{-1}(u)W_t(u)$, alors

$$Cov(W_t', W_s') = (D^{-1})DD'(D^{-1})' = \mathbb{I}_m$$

ainsi

$$\begin{aligned} \left\langle \int_0^s B(u)dW_u, \int_0^s C(u)dW_u \right\rangle &= \left\langle \int_0^s B(u)D(u)dW_u', \int_0^s C(u)D(u)dW_u' \right\rangle \\ &= \int_0^s B(u)D(u)D'(u)C'(u)du \\ &= \int_0^s B(u)\Gamma(u)C'(u)du. \end{aligned}$$

d'où le résultat.

■

Proposition 8 $Z \in \mathcal{H}^W$ si et seulement si

$$Z = \int_0^{+\infty} b(s) dW_s \quad (8)$$

pour $b(s) = (b_1(s), \dots, b_m(s)) \in (L^2[0, +\infty])^m$.

Pour la démonstration de ce résultat on renvoie à [21].

Lemme 3 Soit $\{W_t, t \geq 0\}$ un p.a.o.s et $g : [0, 1] \times [0, 1] \mapsto \mathbb{R}$ une fonction continue et $f, h : [0, 1] \rightarrow \mathbb{R}$ deux fonctions mesurables telles que :

$$\int_0^1 |f(t)| dt < +\infty \quad \int_0^1 h^2(t) dt < +\infty$$

alors

1. le processus $\{X_t\}$ défini par $X_t = \int_0^1 g(t, s) h(s) dW_s$ est continue en moyenne quadratique.
2. $A_1 = \int_0^1 f(t) (\int_0^1 g(t, s) h(s) dW_s) dt$ et $A_2 = \int_0^1 (\int_0^1 f(t) g(t, s) dt) h(s) dW_s$ sont deux variables aléatoires équivalentes dans \mathcal{H}_1^w

0.2.2 L'intégrale stochastique de ITÔ

Soit (Ω, \mathcal{F}, P) un espace de probabilité et \mathcal{F} une filtration sur cet espace. On introduit la classe \mathcal{C} des processus stochastiques f définis sur $([0, T], \mathcal{B}_{[0, T]}, dt) \otimes (\Omega, \mathcal{F}, P)$ et satisfaisant :

1. f est $\mathcal{B}_{[0, T]} \otimes \mathcal{F}$ - mesurable,
2. le processus f est \mathcal{F} - adapté,
3. $\forall t \in [0, T], f(t, \cdot) \in L^2(\Omega, \mathcal{F}, P)$ et $\int_0^T E(|f(t)|^2) dt < +\infty$.

Il s'agit de définir l'intégrale stochastique du processus f par rapport au mouvement brownien $(W(t) : t \geq 0)$:

$$\int_0^t f(s, \omega) dW(s)$$

appelée intégrale stochastique de ITÔ.

La construction de cette intégrale à une nuance près, est analogue à celle de l'intégrale

de Lebesgue. Il s'agit de remplacer les fonctions étagées par des processus étagés et d'adapter ensuite la même approche.

Définition 21 *Un processus \mathcal{F} -adapté f est dit étagé, s'il existe une subdivision $\{t_0, \dots, t_n\}$ de l'intervalle $[0, T]$ et une suite finie $\{f_0, \dots, f_{n-1}\}$ de variables aléatoires réelles telles que :*

$$f(t, \omega) = \sum_{i=0}^{n-1} f_i(\omega) I_{[t_i, t_{i+1}[}(t)$$

et où pour tout i : f_i est \mathcal{F}_{t_i} - mesurable

On fait remarquer pour qu'un processus étagé f \mathcal{F} -adapté appartenant à \mathcal{C} , il faut et il suffit que les variables aléatoires f_i soient de carré intégrables.

Intégrale de ITÔ d'un processus étagé

Pour tout processus étagé f \mathcal{F} -adapté appartenant à \mathcal{C} , on définit son intégrale stochastique par rapport au mouvement brownien ($W(t)$) en posant :

$$\int_0^t f(t) dW(t) = \sum_{i=0}^{n-1} f_i(\omega) (W_{t_{i+1}}(\omega) - W_{t_i}(\omega)) \quad (9)$$

Si maintenant $f \in \mathcal{C}$, alors il existe une suite de processus étagés $(f_n : n \geq 1) \subset \mathcal{C}$ et tel que :

$$f_n \xrightarrow[n \nearrow +\infty]{} f \text{ dans } \mathbf{L}^2(dP \otimes dt)$$

Considérons maintenant la suite de processus $(\xi_n(t) : n \geq 1)$ de terme général défini par

$$\xi_n(t) = \int_0^t f_n(s) dW(s)$$

On montre alors qu'il existe un processus stochastique $(\xi(t))$ mesurable, \mathcal{F} -adapté et continu et tel que presque sûrement la suite $(\xi_n(t) : n \geq 1)$ converge vers $\xi(t)$ dans $L^2(dP \otimes dt)$. Le processus $(\xi(t))$ est unique presque sûrement.

Définition 22 (L'intégrale de ITÔ) On appelle intégrale stochastique d'un processus $f \in \mathcal{C}$, le processus stochastique $(\xi(t) : t \geq 0)$ défini par :

$$\xi(t) = \int_0^t f(s)dW(s) = \lim_{n \rightarrow \infty} \int_0^t f_n(s)dW(s) \quad (10)$$

dans $L^2(dP \otimes dt)$ et où (f_n) est une suite de processus étagés \mathcal{F} -adapté dans \mathcal{C} et telle que $\|f_n - f\| \rightarrow 0$ quand n tend vers $+\infty$.

La norme $\|\cdot\|$ étant celle de l'espace $L^2(dP \otimes dt)$.

Propriétés de l'intégrale stochastique d'ITÔ

L'intégrale stochastique d'un processus $f \in \mathcal{C}$ possède les propriétés suivantes :

1. L'intégrale stochastique sur \mathcal{C} est linéaire et \mathcal{F}_t -mesurable.
2. l'intégrale stochastique est centrée : $E\{\int_0^T f(t)dW(t)\} = 0$ et satisfait **l'isométrie de ITÔ** :

$$E \left\{ \left(\int_0^T f(t)dW(t) \right)^2 \right\} = E \left\{ \int_0^T (f(t))^2 dt \right\} \quad (11)$$

3. L'intégrale de ITÔ d'un processus $f \in \mathcal{C}$ est une martingale continue et satisfait

$$P \left(\sup_{0 \leq t \leq T} |\xi(t)| > \varepsilon \right) \leq \frac{1}{\varepsilon^2} E \left\{ \int_0^T [\varphi(t)]^2 dt \right\} \quad (12)$$

pour tout $\varepsilon > 0$.

L'intégrale de ITÔ peut être prolongée à la classe \mathcal{C}_2 des processus f tels que f est $\mathcal{B}_{[0,T]} \otimes \mathcal{F}$ mesurable, \mathcal{F} -adapté et $\int_0^T |f(t, \omega)|^2 dt < +\infty$ presque sûrement.

Cette classe contient évidemment la classe \mathcal{C} et son intégrale de ITÔ possède les propriétés précédentes à l'exception de la propriété 3 ci-dessus.

L'intégrale de ITÔ multidimensionnelle

Notons par $\mathbb{R}^{m \times n}$ l'espace des matrices d'ordre (m, n) à coefficients réels et par $W(t) = (W_1(t), \dots, W_n(t))'$; $t \geq 0$ le mouvement brownien de dimension n .

Pour tout réel $T \geq 0$, on définit les classes de processus vectoriels suivants :

$$\mathcal{C}_{1W}([0, T]) = \{f : [0, T] \times \Omega \rightarrow \mathbb{R}^{n \times m} \text{ tel que } f_{ij} \in \mathcal{C}_{1W_j}([0, T]) : i = \overline{1, m}, j = \overline{1, n}\},$$

et

$$\mathcal{C}_{2W}([0, T]) = \{f : [0, T] \times \Omega \rightarrow \mathbb{R}^{n \times m} \text{ tel que } f_{ij} \in \mathcal{C}_{2W_j}([0, T]) : i = \overline{1, m}, j = \overline{1, n}\},$$

où $\mathcal{C}_{1W_j}([0, T])$ et $\mathcal{C}_{2W_j}([0, T])$ correspondent respectivement à la classe $\mathcal{C}_1([0, T])$ et $\mathcal{C}_2([0, T])$ et où \mathcal{C}_1 est formé des processus f tels que f est $\mathcal{B}_{[0, T]} \otimes \mathcal{F}$ mesurable, \mathcal{F} -adapté et $\int_0^T |f(t, \omega)| dt < +\infty$ presque sûrement.

Définition 23 *L'intégrale stochastique d'un processus $f \in \mathcal{C}_{1W}[0, T]$ par rapport au brownien n -dimensionnel $(W(t))$ est un processus m -dimensionnel défini par :*

$$\int_0^T f(s) dW(s) = \left(\sum_{i=1}^n \int_0^T f_{ij}(s) dW_j(s) \right)'_{1 \leq j \leq m} \quad (13)$$

où chaque intégrale du second membre est défini au sens de l'intégrale de ITÔ.

0.2.3 La formule de ITÔ et le théorème de représentation des martingales

Définition 24 *On dit qu'un processus stochastique $(X(t) : t \geq 0)$ possède la **différentielle stochastique***

$$dX(t) = a(t) dt + b(t) dW(t) \quad (14)$$

si pour tout t

$$X(t) = X(0) + \int_0^t a(s) ds + \int_0^t b(s) dW(s) \quad (15)$$

où a et b sont deux processus tels que $a \in \mathcal{C}_1$ et $b \in \mathcal{C}_2$.

Notons que les intégrales de l'équation (15) sont bien définies (la première par définition de l'espace \mathcal{C}_1 et la seconde est une intégrale stochastique au sens de ITÔ).

Théorème 1 (Formule de différenciation)

$X_1(t)$ et $X_2(t)$ sont deux processus admettant chacun une différentielle stochastique donnée par

$$dX_1(t) = a_1(t)dt + b_1(t)dW(t) \text{ et } dX_2(t) = a_2(t)dt + b_2(t)dW(t)$$

où a_1 et $a_2 \in \mathcal{C}_1$ et b_1 et $b_2 \in \mathcal{C}_2$, alors le processus produit $(X_1(t) \cdot X_2(t))$ satisfait la différentielle stochastique suivante :

$$d(X_1(t) \cdot X_2(t)) = X_1(t)dX_2(t) + X_2(t)dX_1(t) + b_1(t)b_2(t)dt$$

Théorème 2 (La formule de Itô)

Si le processus $(X(t))$ possède la différentielle stochastique (14) et si f est une fonction réelle de la variable $(t, x) \in [0, T] \times \mathbb{R}$ telle que f est de classe \mathcal{C}^1 par rapport à la variable t et de classe \mathcal{C}^2 par rapport à la variable x , alors le processus $f(t, X(t))$ possède une différentielle stochastique donnée par :

$$\begin{aligned} df(t, X(t)) &= \left[\frac{\partial}{\partial t} f(t, X(t)) + \frac{\partial}{\partial x} f(t, X(t))a(t) \right] dt \\ &+ \frac{\partial}{\partial x} f(t, X(t))b(t)dW(t) + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, X(t))b^2(t)dt \end{aligned} \quad (16)$$

Pour une démonstration nous renvoyons au livre de Karatzas and Shreve [14].

Un moyen mnémotechnique pour retrouver la formule de Itô est de rajouter à la différentielle ordinaire de $f(t, \xi(t))$ le terme $\frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, X(t)) \cdot (dX(t))^2$, ce qui permet de réécrire (16) sous la forme

$$\begin{aligned} df(t, X(t)) &= \frac{\partial}{\partial t} f(t, X(t)) \cdot dt + \frac{\partial}{\partial x} f(t, X(t)) \cdot dX(t) \\ &+ \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, X(t)) \cdot (dX(t))^2 \end{aligned} \quad (17)$$

et où $[dX(t)]^2$ est calculé selon les règles suivantes :

$$dt \cdot dt = dW(t) \cdot dt = dt \cdot dW(t) = 0 \text{ et } dW(t) \cdot dW(t) = dt$$

Exemple 1 Soit α un processus de la classe $\mathcal{C}_2[0, T]$ et tel que $\alpha^2 \in \mathcal{C}_2[0, T]$. On considère le processus $(X(t))$ qui possède la différentielle stochastique

$$dX(t) = -\frac{1}{2}\alpha^2(t)dt + \alpha(t)dW(t)$$

et soit $\eta(t, X(t)) = f(t, X(t))$ où $f(t, x) = e^x$.

En appliquant la formule de ITÔ à $\eta(t, X(t))$, on trouve

$$d\eta(t, X(t)) = \alpha(t) \exp\left\{\int_0^t -\frac{1}{2}\alpha^2(s)ds + \int_0^t \alpha(s)dW(s)\right\}dW(t)$$

Ce résultat sera appliqué au §2 du chapitre 2 portant sur le filtrage non linéaire à temps continu.

Le théorème de représentation des martingales

Le résultat suivant représente une réciproque de la propriété 3 de l'intégrale stochastique et dit en substance que toute martingale est l'intégrale stochastique de ITÔ d'un processus unique de la classe \mathcal{C} . Ce résultat est connu sous le nom de théorème de représentation des martingales. De façon précise on a

Théorème 3 Soit $(X_t)_{t \in [0, T]} \in \mathbb{L}^2(\Omega, \mathcal{F}, P)$ une martingale et $W(t)$ un mouvement brownien. Alors il existe un unique processus $f \in \mathcal{C}[0, T]$ tel que

$$X(t) = E(X_0) + \int_0^t f(s)dW(s).$$

pour tout $t \in [0, T]$.

0.3 Equations différentielles stochastiques (E.D.S)

Définition 25 On appelle équation différentielle stochastique du premier ordre, une équation de la forme

$$dX(t) = a(t, X(t)) dt + b(t, X(t)) dW(t) \quad (18)$$

où les coefficients $a(t, x)$ et $b(t, x)$ sont deux fonctions définies et mesurables sur $[0, T] \times]-\infty, +\infty[$, $(W(t) : t \geq 0)$ un p.a.o.s et $(X(t) : t \geq 0)$ un processus stochastique.

Si les accroissements de $(W(t))$ sont gaussiens alors cette intégrale est appelée l'E.D.S au sens de ITÔ. Si la fonction a s'écrit $a(t, X(t)) = A(t).X(t)$ et b ne dépend que de t , cette équation est dite une E.D.S au sens de Wiener.

0.3.1 Les équations différentielles stochastiques au sens de Wiener

Considérons l'équation différentielle stochastique suivante :

$$\begin{cases} dX(t) &= A(t)X(t)dt + C(t)dW(t) \\ X(0) &= X_0 \end{cases} \quad (19)$$

où $A(\cdot)$, $C(\cdot)$ sont deux matrices de dimensions $n \times n$ et $n \times m$ respectivement dont les coefficients sont des fonctions du temps continues par morceaux, W_t est un p.a.o.s de dimension m et X_0 un vecteur aléatoire de dimension n .

Théorème de l'existence et de l'unicité

Soit $\{W_t\}$ un p.a.o.s et X_0 un vecteur aléatoire de dimension n orthogonale à \mathcal{H}^W de moyenne m_0 et de covariance Q_0 . L'équation stochastique (19) admet une unique solution X_t donnée par

$$X(t) = \Phi(t, 0)X_0 + \int_0^t \Phi(t, s)C(s)dW(s) \quad (20)$$

où $\Phi(t, s)$ est une application linéaire qui à X_0 associe l'unique solution X_t de l'équation homogène :

$$\begin{cases} dX(t) = A(t)X(t)dt \\ X(0) = X_0 \end{cases}$$

Φ est appelée *la matrice de transition* correspondante à $A(t)$ ou encore *la résolvante de l'équation homogène*.

On rappelle que la solution du système homogène vérifie :

$$X(t) = \Phi(t, s)X(s).$$

En outre, la moyenne $m(t) = E(X(t))$ et la covariance $Q_t = cov(X(t))$ sont les uniques solutions des équations :

$$\dot{m}(t) = A(t)m(t), \quad m(0) = m_0 \quad (21)$$

$$\dot{Q}(t) = A(t)Q(t) + Q(t)A'(t) + C(t)C'(t), \quad Q(0) = Q_0. \quad (22)$$

Remarque 3 *La résolvante Φ satisfait les propriétés suivantes :*

1. $\Phi(t, t) = I$ (I est la matrice d'identité)
2. $\forall r \leq s \leq t, \Phi(t, s)\Phi(s, r) = \Phi(t, r)$
3. $\Phi(t, s)$ est non singulière (invertible) et $\Phi^{-1}(t, s) = \Phi(s, t)$
4. $\frac{\partial}{\partial t}\Phi(t, s) = A(t)\Phi(t, s)$, $\frac{\partial}{\partial s}\Phi(t, s) = -\Phi(t, s)A(s)$.

Corollaire 1 *Si en plus des hypothèses ci-dessus, le p.a.o.s (W_t) est un mouvement brownien, alors la solution (20) de l'équation (19) est un processus gaussien de moyenne $m(t)$ et de covariance $Q(t)$ tels que :*

$$m(t) = \Phi(t, 0)X$$

et

$$Q(t) = \int_0^t \Phi(t, s)C(s)C'(s)\Phi'(t, s)ds + \Phi(t, 0)Q_0\Phi'(t, 0)$$

où Φ' désigne la transposée de Φ .

Exemple 2 On considère l'équation différentielle stochastique linéaire

$$dX(t) = -A.X(t)dt + dW(t)$$

où A est une constante réelle, W_t un mouvement brownien et $X_0 \perp \mathcal{H}^W$.

On vérifie facilement que la résolvante du système homogène $dX(t) = -AX(t)dt$ est donnée par $\Phi(t, s) = e^{-A(t-s)}$. Alors, d'après le théorème de l'existence et l'unicité il existe une solution unique de cette équation donnée par

$$X(t) = X(0).e^{-At} + \int_0^t e^{-A(t-s)}dW(s).$$

Ce processus est connu sous le nom de processus d'Orstein-Uhlenbek.

0.3.2 Equations différentielles stochastiques au sens de Itô

Considérons l'équation différentielle stochastique suivante :

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t) \quad (23)$$

où les coefficients $a(t, x)$ et $b(t, x)$ sont deux fonctions définies et mesurables sur $[0, T] \times]-\infty, +\infty[$, $(W(t) : t \geq 0)$ le mouvement brownien standard.

Le théorème de l'existence et l'unicité

Sous les conditions suivantes :

1. $\forall t \in [0, T], \forall (x, y) \in \mathbb{R} \times \mathbb{R} : |a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K^*|x - y|$
(on dit que les fonctions a et b sont globalement Lipschitziennes),
2. $\forall t \in [0, T], \forall x \in \mathbb{R} : |a(t, x)|^2 + |b(t, x)|^2 \leq K^2(1 + |x|^2)$ où K, K^* deux constantes positive (on dit que a et b sont linéairement croissantes),
3. $E(|X_0|^2) < \infty$,
4. X_0 est indépendante de $W(t)$,

il existe une unique solution presque sûrement continue $(X(t) : t \in [0, T]) \in \mathcal{C}[0, t]$ appelée solution forte de l'E.D.S (23). Pour la démonstration nous suggérons les références [14] et [33].

Exemple 3 *Considérons l'équation différentielle stochastique donnée dans l'exemple (2).*

On vérifie facilement que les conditions assurant l'existence de la solution sont satisfaites et donc cette équation admet une unique solution. En utilisant la formule de ITO, on trouve que cette solution est un processus d'Orstein-Uhlenbeck.

Remarque 4 *Si pour tous $t > 0$, les fonctions $a(t, x)$ et $b(t, x)$ sont bornées et continues avec des dérivées partielles d'ordre 2 par rapport à x bornées et continues, alors l'équation (23) admet une unique solution appelée solution faible.*

0.4 Processus de Markov

Définition 26 *Un processus stochastique défini sur $[0, T]$ est dit un processus de Markov si :*

$$P(X(t) \in A/\mathcal{F}_s) = P(X(t) \in A/X(s))p.s$$

où \mathcal{F}_t est la filtration naturelle associée au processus $X(t)$, A un borélien, $x \in \mathbb{R}$ et $0 < s < t < T$.

La probabilité $P(X(t) \in A/X(s))$ notée $p(s, X(s), t, A)$ est appelée **la probabilité de transition** du processus de Markov ($X(t)$). Elle satisfait les propriétés suivantes :

1. $p(s, x, t, A)$ est une fonction réelle mesurable de x .
2. $p(s, x, t, A)$ est une mesure par rapport à A pour s, t et x fixés.
3. Pour x et $0 \leq s < \tau < t \leq T$, la probabilité de transition $p(s, x, t, A)$ satisfait l'équation de Chapman-Kolmogorov

$$p(s, x, t, A) = \int_{-\infty}^{+\infty} p(\tau, y, t, A)p(s, x, \tau, dy)$$

Cette équation sera utilisée au chapitre 2 afin de résoudre le problème du filtrage non linéaire à temps discret.

Théorème 4 *La solution $X(t)$ de l'E.D.S*

$$dX(t) = a(t, X(t))dt + b(t, X(t))dW(t)$$

dont les coefficients $a(t, x)$ et $b(t, x)$ satisfont les conditions du théorème de l'existence et l'unicité, est un processus de Markov de noyau de transition défini par

$$p(s, x, t, A) = P(X(t) \in A / X(s) = x) = P(X_{x,s}(t) \in A)$$

où $X_{x,t}$ est solution de l'équation :

$$X_{x,s}(t) = x + \int_s^t (a(u, X_{x,s}(u)))du + \int_s^t b(u, X_{x,s}(u))dW(u)$$

sur l'intervalle $[0, T]$.

Le générateur infinitésimal de cette solution, suivant la définition donnée par Byörk (2003) [34], est donnée par

$$\mathcal{L}h(t, x) = a(t, x)\frac{\partial h(t, x)}{\partial x} + \frac{1}{2}b^2(t, x)\frac{\partial^2}{\partial x^2}h(t, x) \quad (24)$$

où $h(., .)$ est une fonction telle que $h \in \mathcal{C}^2(\mathbb{R})$.

Et l'opérateur adjoint \mathcal{L}^* est défini par :

$$\mathcal{L}^*h(t, x) = -\frac{\partial}{\partial x}[a(t, x)h(t, x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[b(t, x)^2h(t, x)].$$

Pour plus de détail sur les propriétés des opérateurs différentiels (voir [34] page 67).

Première partie

Le filtrage

Position du problème

Le problème du filtrage consiste à estimer l'état X_t d'un système dynamique (X_t) non observable directement, à partir des observations bruitées ($Y_t : 0 \leq t \leq T$). Le processus d'état et le processus des observations sont décrit par un système d'équations différentielles stochastiques

$$\begin{cases} dX_t &= b(X_t)dt + f(t)dV_t \\ dY_t &= h(X_t)dt + g(t)dW_t \end{cases} \quad (25)$$

où $V(t)$ et (W_t) sont deux mouvements browniens, b, f, h et g des fonctions supposées continues et de sorte que les intégrales stochastiques de ITÔ par rapport aux browniens V et W respectivement, soient bien définies.

Le problème est d'estimer l'état présent X_t du système en se basant sur l'information fournie par les observations passées, c'est à dire par l'histoire antérieure à l'instant t :

$$\mathcal{F}_t^Y = \sigma(Y_s : 0 \leq s \leq t).$$

Plus généralement, étant donnée l'histoire $\mathcal{F}_{t_1}^Y$, cette estimation est dite un problème de

1. prédiction si $t > t_1$
2. filtrage si $t = t_1$
3. lissage si $t < t_1$.

Notre travail porte sur le problème du filtrage qui consiste à trouver le meilleur estimateur \mathcal{F}_t^Y – mesurable de l'état X_t du système. Le choix d'un tel estimateur n'est pas unique. Si l'on choisit le critère classique des moindres carrés, l'on sait (lemme 4 ci-dessous) que l'estimateur optimal \widehat{X}_t de X_t au vu des observations Y_0, Y_1, \dots, Y_n est l'espérance conditionnelle $E(X_t/Y_0, \dots, Y_n)$ laquelle satisfait

$$E(|X_t - \widehat{X}_t|^2) = \min_{U \in \mathcal{V}_t} \{E(|X_t - U|^2)\}$$

où \mathcal{V}_t est le sous espace de Hilbert générée par les variables aléatoires de carré intégrables U qui sont \mathcal{F}_t^Y – mesurables et nous rappellerons ci-dessous que l’estimateur optimal est déterminé par l’espérance conditionnelle $E^{\mathcal{F}_t^Y}(X_t)$ laquelle coïncide avec la projection orthogonale de la variable X_t sur le sous espace de Hilbert \mathcal{V}_t . Cette formulation du problème du filtrage puise sa légitimité dans le cas de bruits gaussiens. Mais en général, l’information la plus riche disponible au vu des observations Y_1, \dots, Y_n est contenu dans la loi conditionnelle de X_t sachant les observations passées. Aussi, on préfère une estimation du maximum à posteriori (maximisant la loi conditionnelle $\mathcal{L}(X_t/Y_1, \dots, Y_t)$). Car la donnée d’une telle loi permet de choisir une estimation de l’état du système. Dans ce cas, le problème du filtrage consiste à calculer cette loi de probabilité conditionnelle.

Il est entendu que dans le cas de bruits gaussiens, ces deux formulations de la théorie du filtrage sont les mêmes.

Il est utile de rappeler que si la probabilité conditionnelle de X_t par rapport à la σ –algèbre \mathcal{F}_t^Y admet une version régulière π_t de la probabilité conditionnelle³ alors on a

$$E^{\mathcal{F}_t^Y}(X_t) = \int_{\Omega} X_t \pi_t(d\omega)$$

Dans ce cas, le problème du filtrage revient à estimer la loi conditionnelle du processus d’états non observable relativement au passé \mathcal{F}_t^Y .

Dans cette partie, nous développons le problème du filtrage linéaire à temps discret et à temps continu dans le chapitre 2 et le problème du filtrage non linéaire dans le chapitre 3. Un premier chapitre est consacré aux problèmes d’estimation linéaire

3. Une fonction π_t de $\mathcal{B}(\mathbb{R}^n) \times \Omega$ dans $[0, 1]$ telle que :

1. pour tout ω fixé $\pi_t(\cdot, \omega)$ est une mesure de probabilité sur $(\Omega, \mathcal{F}_t^Y)$,
2. pour tout $A \in \mathcal{B}(\mathbb{R}^n)$, $\pi_t(A, \cdot)$ est \mathcal{F}_t^Y – mesurable,
3. pour tout $A \in \mathcal{F}_t^Y$ et pour presque tout ω , on a

$$\pi_t(A, \omega) = P(X_t \in A / \mathcal{F}_t^Y)(\omega)$$

optimale au sens de la variance minimum.

Chapitre 1

Le prédicteur linéaire optimal

Nous rappellerons dans ce chapitre certains résultats sur les problèmes de prédiction.

1.1 Définition et propriétés du prédicteur linéaire

Lemme 4 *Soit X une variable de carré intégrable définie sur un espace de probabilité (Ω, \mathcal{F}, P) et \mathcal{G} une sous σ -algèbre de \mathcal{F} . L'espérance conditionnelle $E(X/\mathcal{G})$ est le meilleur estimateur de $E(X)$ dans $L^2(\Omega, \mathcal{F}, P)$ parmi les variables aléatoires de carré intégrables et qui sont \mathcal{G} -mesurables.*

Démonstration

Il suffit de montrer que

$$E(X - \nu)^2 \geq E(X - E(X/\mathcal{G}))^2$$

pour toute variable aléatoire ν \mathcal{G} -mesurable et de carré intégrable.

A cette fin, écrivons le terme $E(X - \nu)^2$ sous la forme

$$\begin{aligned} E(X - \nu)^2 &= E(X - E(X/\mathcal{G}))^2 + E(E(X/\mathcal{G}) - \nu)^2 \\ &\quad - 2E\{E[(X - E(X/\mathcal{G}))(E(X/\mathcal{G}) - \nu)]/\mathcal{G}\} \end{aligned}$$

La \mathcal{G} mesurabilité des variables $E(X/\mathcal{G})$ et $E(X/\mathcal{G}) - \nu$ fait que le troisième terme de cette somme est nul et donc

$$E(X - \nu)^2 = E(X - E(X/\mathcal{G}))^2 + E(E(X/\mathcal{G}) - \nu)^2$$

d'où le résultat. ■

Lemme 5 *Soit \mathcal{G} une sous σ -algèbre de \mathcal{F} et \mathcal{V} le sous espace de Hilbert de $L^2(\Omega, \mathcal{F}, P)$ des variables aléatoires \mathcal{G} -mesurables. On désigne par $\mathcal{P}_{\mathcal{V}}$ le projecteur orthogonal de $L^2(\Omega, \mathcal{F}, P)$ dans \mathcal{V} . Pour toute variable aléatoire X de carré intégrable, on a*

$$\mathcal{P}_{\mathcal{V}}(X) = E^{\mathcal{G}}(X) \tag{1.1}$$

où $E^{\mathcal{G}}(X)$ est l'espérance conditionnelle de X par rapport à la σ -algèbre \mathcal{G} .

Démonstration

On rappelle que $E^{\mathcal{G}}(X)$ est \mathcal{G} mesurable, unique presque sûrement et vérifie

$$\int_A X dP = \int_A E^{\mathcal{G}}(X) dP$$

pour tout événement $A \in \mathcal{G}$. D'autre part, la variable aléatoire $\mathcal{P}_{\mathcal{V}}(X)$ est \mathcal{G} mesurable et est orthogonale à $X - \mathcal{P}_{\mathcal{V}}(X)$. On déduit alors que

$$\int_{\Omega} \mathcal{P}_{\mathcal{V}}(X) dP = \int_{\Omega} X dP$$

et par l'unicité de l'espérance conditionnelle, la formule (1.1) en résulte. ■

Définition 27 *Soient X une variable aléatoire de carré intégrable et $(Y_t : t \in \{0, 1, \dots, T\})$ une suite finie de variables aléatoires de carré intégrable.*

On appelle meilleur prédicteur linéaire de X en terme de $\{Y_t, t \in \{0, 1, \dots, T\}\}$, la variable $\widehat{X} \in \mathcal{H}^Y$ (\mathcal{H}^Y étant le sous espace de Hilbert généré par les variables $Y_t : t = 0, 1, \dots, T$) telle que

$$\| X - \widehat{X} \| = \min_{U \in \mathcal{H}^Y} \| X - U \| .$$

Par le théorème de projection, il s'avère que le prédicteur linéaire optimal \widehat{X} n'est rien d'autre que la projection orthogonale de X sur \mathcal{H}^Y : $\widehat{X} = \mathcal{P}_{\mathcal{H}^Y}(X)$.

On déduit du lemme (4) ci-dessus :

Corollaire 2 *Le prédicteur linéaire de la variable X en terme de $Y_s : 0 \leq s \leq t$ pour tout $t \in \{0, 1, \dots, T\}$ satisfait*

$$\widehat{X} = \mathcal{P}_{\mathcal{H}_t^Y}(X) = E^{\mathcal{F}_t^Y}(X)$$

où \mathcal{F}_t^Y est la σ -algèbre engendrée par les variables $Y_s : s \leq t$.

Corollaire 3 *Si en plus des hypothèses du corollaire 2, le processus $\{Y_t : 0 \leq t \leq T\}$ est un p.a.o.s, alors le prédicteur linéaire de la variable X en terme de $Y_s : 0 \leq s \leq t$ pour tout $t \in \{0, 1, \dots, T\}$ est donné par*

$$\widehat{X} = \int_0^t \left(\frac{d}{ds} E[XY_s] \right) dY_s$$

Démonstration

Ceci vient de ce que $\widehat{X} \in \mathcal{H}_t^Y$ et de ce que

$$\widehat{X} = \int_0^t \Phi(s) dY_s$$

où $\Phi \in L^2([0, t])$. Comme $(X - \widehat{X}_t) \perp \mathcal{H}_t^Y$, alors il résulte de ce que l'intégrale de Wiener conserve le produit scalaire que

$$E(X \int_0^t \Psi(s) dY_s) = \int_0^t \Phi(s) \Psi(s) ds$$

et ce pour tout $\Psi \in L^2([0, t])$. On déduit alors en prenant $\Psi(s) = 1_{[0, r]}(s) \forall r \leq t$ que

$$E(XY_r) = \int_0^r \Phi(s) ds$$

d'où l'on tire $\Phi(r) = \frac{d}{dr} E(XY_r)$ ce qui conduit au résultat attendu. ■

Définition 28 *Soit $X_t = (X_{t_1}, \dots, X_{t_n})' \in (\mathbb{L}^2)^n$ et soit $\{Y_s, 0 \leq s \leq k\}$ une famille de vecteurs aléatoires tels que $Y_s \in (\mathbb{L}^2)^m$. On appelle meilleur prédicteur*

linéaire à un pas \widehat{X}_t de X_t , le vecteur aléatoire dont la i ème composante est le meilleur prédicteur linéaire de X_{t_i} en terme de toutes les composantes des vecteurs Y_0, \dots, Y_{t-1} .

On note $\widehat{X}_t = X_{t/t-1}$ ce prédicteur linéaire. L'erreur de prédiction à un pas notée $P_{t/t-1}$, est la matrice des erreurs en moyenne quadratiques définie par

$$P_{t/t-1} = E[(X_t - \widehat{X}_t)(X_t - \widehat{X}_t)']$$

Plus généralement, le meilleur estimateur de X_t basé sur Y_0, Y_1, \dots, Y_k est le vecteur $X_{t/k}$ dont la i ème composante est le meilleur estimateur linéaire de X_{t_i} en terme de toutes les composantes des vecteurs Y_0, Y_1, \dots, Y_k . Et la matrice de l'erreur quadratique moyenne de cette estimation est donnée par

$$P_{t/k} = E[(X_t - X_{t/k})(X_t - X_{t/k})'].$$

Propriétés

1. La projection orthogonale $\mathcal{P}_t(X)$ est linéaire, de plus par définition de $\mathcal{P}(X_i/\mathcal{H}^Y)$, $\mathcal{P}_t(X)$ est le seul vecteur aléatoire dont les composantes sont dans \mathcal{H}^Y , tel que :

$$[X - \mathcal{P}_t(X)] \perp Y_s, \quad s = 0, 1, \dots, t.$$

2. Si $X \in (L_2)^n$ et $Y \in (L_2)^m$, on a :

$$\mathcal{P}(X/Y) = MY$$

où $M = E(XY')[E(YY')]^{-1}$ est une matrice de dimension $(n \times m)$ et $[E(YY')]^{-1}$ est l'inverse généralisé de $E(YY')$.

3. Si Y et Z sont deux vecteurs orthogonaux, alors :

$$\mathcal{P}(X/Y, Z) = \mathcal{P}(X/Y) + \mathcal{P}(X/Z)$$

Chapitre 2

Le filtrage linéaire

Le problème du filtrage linéaire résolu par Kalman (1960) pour le filtrage à temps discret et par Kalman - Bucy (1961) pour le filtrage à temps continu, est fondé sur l'approche innovation. Le filtrage linéaire suppose que les fonctions b et h des équations différentielles stochastiques (25) ci-dessus sont linéaires, autrement dit ce système d'équations est linéaire :

$$\begin{cases} dX_t &= A(t)X_t dt + C(t)dV_t \\ dY_t &= H(t)X_t dt + G(t)dW_t \end{cases} \quad (2.1)$$

où (V_t) et W_t sont des mouvements browniens et $A(t)$, $H(t)$, $C(t)$ et $G(t)$ sont des fonctions matricielles du temps. La première équation du système (2.1) est dite **équation d'état** ou **équation de transition** et la deuxième équation est dite **équation de mesure** ou **équation d'observation**. Ce système d'équations (d'état et d'observation) est dit un **modèle d'espace d'états**.

Ce chapitre comprend deux sections dévolues chacune au filtrage linéaire : la première au filtrage à temps discret et la seconde au filtrage à temps continu.

2.1 Le filtrage linéaire à temps discret

Le **modèle d'espace d'états** (2.1) s'écrit dans cette situation sous la forme

$$\begin{cases} X_{t+1} &= F(t)X_t + V_t \\ Y_t &= G(t)X_t + W_t, \end{cases} \quad (2.2)$$

pour tout $t = 1, 2, \dots$ et où V_t et W_t sont deux bruits blancs indépendants de matrice de covariance respective Q_t et R_t indépendants de X_1 , et $F(t)$ et $G(t)$ deux fonctions matricielles du temps t .

Le modèle d'espace d'état est dit **stationnaire** si les matrices $F(t)$, $G(t)$, Q_t et R_t sont indépendantes du temps t et c'est ce que l'on rencontre dans bon nombre d'applications.

Le filtre est un schéma récursif consistant à chaque itération d'une prédiction (estimation à priori) de l'état du système suivie d'une mise à jour de l'estimation de l'état (estimation à postérieur). Cette procédure est reproduite chaque fois que l'on introduit une nouvelle observation.

2.1.1 Le filtre de Kalman

Le filtre de Kalman est un algorithme récursif dont le but est de construire des estimateurs linéaires optimaux (au sens qu'ils minimisent l'erreur quadratique moyenne) du vecteur d'état X_t en fonction des observations Y_0, Y_1, \dots, Y_t . Il s'agit en fait de calculer à chaque instant t la valeur de $\mathcal{P}_t(X_t)$ ainsi que la matrice de l'erreur $P_{t|t}$ commise.

Considérons, le modèle dynamique défini par (2.2) et on suppose que X_1 est une variable aléatoire centrée de carré intégrable et orthogonale au sous espace fermé $\mathcal{H}_{V,W}$. Il est facile de vérifier que

$$\begin{aligned} V_t \perp X_t & & V_t \perp Y_s, & & 1 \leq s \leq t, \\ & \text{et} & & & \\ W_t \perp X_s, 1 \leq s \leq t & & W_t \perp Y_s, & & 1 \leq s < t. \end{aligned}$$

Description du filtre de Kalman

Le filtre de Kalman est un schéma récursif alternant à chaque itération t une étape prédiction et une étape correction ou encore une mise à jour de l'estimation de l'état ainsi que de l'erreur d'estimation.

L'étape prédiction consiste à prédire l'état $\mathcal{P}_{t-1}(X_t)$ et à évaluer la matrice de l'erreur de prédiction $P_{t|t-1}$.

L'étape correction consiste à

1. prédire l'observation $\mathcal{P}_{t-1}(Y_t)$,
2. déterminer l'innovation ν_t et sa covariance S_t ,
3. évaluer le gain du filtrage.
4. calculer $\mathcal{P}_t(X_t)$ ainsi que le calcul de la matrice de covariance $P_{t/t}$ de l'erreur d'estimation.

Pour initialiser le filtre, on suppose que Y_0 est orthogonale au sous espace fermé $\mathcal{H}_{V,W}$ et on pose $\hat{X}_1 = \mathcal{P}(X_1/Y_0)$.

L'hypothèse sur Y_0 est naturellement satisfaite lorsque ce dernier est un vecteur aléatoire dégénéré. En pratique, on choisit $Y_0 = (1, 1, \dots, 1)'$.

On résume ces différentes étapes dans le théorème suivant.

Théorème 5 (*Le filtre de Kalman*)

Sous les hypothèses précédentes, l'estimateur $X_{t/t} = \mathcal{P}_t(X_t)$ et la matrice de covariance $P_{t/t}$ de l'erreur d'estimation, sont déterminés par les récursions suivantes :

$$\mathcal{P}_t(X_t) = \mathcal{P}_{t-1}(X_t) + P_{t|t-1}G'_tS_t^{-1}(Y_t - G_t\hat{X}_t) \quad (2.3)$$

et

$$P_{t/t} = P_{t|t-1} - P_{t|t-1}G'_tS_t^{-1}G_tP'_{t|t-1} \quad (2.4)$$

où $S_t = G_tP_{t|t-1}G'_t + R_t$ est la matrice de covariance de l'innovation ν_t et $P_{t|t-1}$ est la matrice de covariance de l'erreur de prédiction.

Pour la démonstration, on procède en suivant les étapes du filtre de Kalman telles que décrites ci-dessus.

Lemme 6 (*Étape prédiction*)

La prédiction \widehat{X}_t de l'état X_t sur la base des observations Y_0, Y_1, \dots, Y_{t-1} et l'erreur de prédiction $P_{t|t-1}$ sont données par

$$\begin{cases} \widehat{X}_t = F_{t-1}X_{t-1|t-1} \\ P_{t|t-1} = F_{t-1}P_{t-1|t-1}F'_{t-1} + Q_{t-1} \end{cases}$$

Ce résultat dérive directement de l'équation d'état et de l'orthogonalité de V_{t-1} avec X_{t-1} .

En utilisant l'équation d'observation du système (2.2) et l'orthogonalité de W_t et de Y_s , pour tous s et $t : s < t$, nous déduisons aussi le

Lemme 7 (*Prédiction de l'observation*)

Le meilleur estimateur linéaire de Y_t basé sur les observations Y_0, Y_1, \dots, Y_{t-1} est la projection orthogonale de Y_t sur le sous espace fermé \mathcal{H}_{t-1}^Y engendré par Y_0, Y_1, \dots, Y_{t-1} :

$$\mathcal{P}_{t-1}(Y_t) = G_t X_{t|t-1}$$

Définition 29 On appelle processus des innovations le processus $(\nu_t : t \in \mathbb{N})$ défini par

$$\nu_t = \begin{cases} Y_0 & \text{si } t = 0 \\ Y_t - G_t \widehat{X}_t & \text{si } t = 1, 2, \dots \end{cases} \quad (2.5)$$

où $G_t \widehat{X}_t$ est le prédicteur de Y_t sur le sous espace de Hilbert \mathcal{H}_{t-1}^Y .

L'innovation ν_t s'écrit encore

$$\nu_t = G_t(X_t - \widehat{X}_t) + W_t$$

pour tout $t \geq 1$.

Sous cette forme, la matrice de covariance S_t du processus des innovations se déduit facilement. On trouve

$$S_t = G_t P_{t|t-1} G'_t + R_t \quad (2.6)$$

Lemme 8 (*Calcul du gain de Kalman*)

Le meilleur estimateur linéaire de X_t basé sur les observations Y_0, Y_1, \dots, Y_t (i.e. $\mathcal{P}_t(X_t)$) est donné par

$$X_{t|t} = \hat{X}_t + M\nu_t \quad (2.7)$$

où

$$M = P_{t|t-1}G'_tS_t^{-1}. \quad (2.8)$$

est appelé le **gain de Kalman**.

Démonstration

L'innovation (ν_t) étant orthogonale à $\{Y_s, \forall s < t\}$, fait que

$$\mathcal{P}_t = \mathcal{P}_{t-1} + \mathcal{P}(\cdot/\nu_t)$$

où $\mathcal{P}(\cdot/\nu_t)$ est le projecteur sur la droite engendré de par ν_t . Et par suite,

$$X_{t/t} = \mathcal{P}_t(X_t) = \mathcal{P}_{t-1}(X_t) + M\nu_t. \quad (2.9)$$

où M selon la propriété (2) du prédicteur linéaire donnée au chapitre 1, est telle que énoncé par la formule 2.8

$$M = P_{t/t-1}G'_tS_t^{-1}.$$

■

Revenons à la démonstration du théorème

L'estimation du filtre de Kalman est fournie par l'expression 2.9 ci-dessus et pour ce qui est de la matrice de covariance de l'erreur de filtrage, celle ci s'écrit

$$X_t - \hat{X}_t = X_t - X_{t|t} + M\nu_t$$

et compte tenu de l'orthogonalité de $X_t - \mathcal{P}_t X_t$ et ν_t , on obtient

$$P_{t/t} = P_{t/t-1} - P_{t/t-1}G'_tS_t^{-1}G_tP'_{t/t-1}.$$

Remarques

1. L'estimateur de Kalman $X_{t/t}$ a utilisé l'information d'innovation pour corriger la prédiction de l'état en tenant compte de l'estimation antérieure. C'est ce qu'on appelle le schéma de type prédiction-correction.
2. Le gain de Kalman varie selon la confiance que l'on peut accorder aux observations. En effet, d'après le lemme d'inversion matriciel, il est facile de vérifier que le gain peut s'écrire :

$$M_t = P_{t/t} G_t' R_t^{-1}$$

cette équation montre que le gain de Kalman fluctue :

- (a) Le gain M_t diminue si l'estimation par le modèle devient plus précise.
 - (b) Le gain M_t augmente si les observations deviennent plus précises.
3. Dans le cas où les bruits du modèle d'espace d'état définie par (2.2) sont corrélés c-à-d $E(V_t W_t') = H_t$, on se ramène dans ce cas à un modèle à bruit non corrélé développé ci-dessus, par modification de l'équation d'état. On définit par (2.10) le modèle équivalent avec des bruits non corrélés :

$$\begin{cases} X_{t+1} &= \tilde{F}_t X_t + \tilde{G}_t Y_t + \tilde{V}_t \\ Y_t &= G_t X_t + W_t \end{cases} \quad (2.10)$$

où $\tilde{V}_t = V_t - H_t R_t^{-1} W_t$, $\tilde{F}_t = F_t - H_t R_t^{-1} G_t$, $\tilde{G}_t = H_t R_t^{-1}$.

Les bruits de ce nouveau modèle satisfont les propriétés suivantes :

$$E(\tilde{V}_t) = 0, E(W_t) = 0, E(W_t W_t') = R_t, E(\tilde{V}_t \tilde{V}_t') = \tilde{Q}_t = Q_t - H_t R_t^{-1} H_t'.$$

En appliquant le filtre de Kalman discret au système (2.10), on trouve que les étapes de correction sont conservées, la seule étape qui va être modifiée est celle de prédiction dans ce cas le prédicteur devient :

$$X_{t+1/t} = \underbrace{F_t X_{t/t}}_{\text{partie pour } H_t = 0} + H_t S_t^{-1} (Y_t - G_t X_{t/t-1})$$

avec une matrice de covariance de l'erreur

$$P_{t+1/t} = \underbrace{F_t P_{t/t} F_t' + Q_t + F_t M_t S_t M_t' F_t'}_{\text{partie pour } H_t = 0} - L_t S_t L_t'$$

où $L_t = H_t S_t^{-1} + F_t M_t$.

4. Dans le cas où le bruit d'état à l'instant t est corrélé avec le bruit de mesure à l'instant $t - 1$, on estime l'état du système en utilisant l'algorithme de Kalman avec la même étape de prédiction mais contrairement à la remarque précédente, c'est l'étape de correction qui va être modifier (voir [19]).
5. Selon que l'étape prédiction suit ou précède l'étape de correction, le filtre correspondant est un filtre estimateur (i.e. $(\mathcal{P}_t(X_t), P_{t/t}) \rightarrow (\mathcal{P}_{t+1}(X_{t+1}), P_{t+1/t+1})$) ou un filtre prédictif à un pas (i.e. $(\mathcal{P}_{t-1}(X_t), P_{t-1/t}) \rightarrow (\mathcal{P}_t(X_{t+1}), P_{t+1/t})$). On notera que les expressions permettant de calculer le gain M et les matrices de covariance $P_{t/t}, P_{t+1/t}$ ne dépendent pas des observations y_t et des valeurs de $\mathcal{P}_t(X_t)$ et $\mathcal{P}_t(X_{t+1})$. Elles peuvent donc être calculées à l'avance. On trouve à partir des expressions ci-dessus, l'équation de récurrence fournissant $P_{t+1|t}$:

$$P_{t+1|t} = F_t P_{t|t-1} F_t^T + Q_t - F_t P_{t|t-1} G_t^T (R_t + G_t P_{t|t-1} G_t^T)^{-1} G_t P_{t|t-1} F_t^T$$

qui est une équation algébrique de Riccati à temps discret permettant de calculer $P_{t|t}$ et M_t par les expressions des étapes corrections et l'expression du filtre. Une fois que ces matrices sont connues, les filtres peuvent être réalisés suivant que l'on dispose ou non de sortie à l'instant $t+1$.

2.2 Le filtrage linéaire à temps continu

Le processus d'états $\{X_t; t \geq 0\}$ et le processus des observations $\{Y_t; t \geq 0\}$ sont solutions du système d'équations différentielles stochastiques :

$$\begin{cases} dX_t &= A(t)X_t dt + C(t)dV_t \\ dY_t &= H(t)X_t dt + G(t)dW_t \end{cases} \quad (2.11)$$

avec les conditions initiales X_0 et $Y_0 = 0$ et où (V_t) et (W_t) sont deux processus supposés ici à accroissements orthogonaux stationnaires non nécessairement gaussiens tels que \mathcal{H}^V est orthogonale à \mathcal{H}^W , X_0 un vecteur aléatoire orthogonale à $\mathcal{H}^{V,W}$ la fermeture de l'enveloppe linéaire générée par les p.a.o.s $((V_t)$ et $(W_t))$ et A, C, H et G sont des matrices de dimensions convenables et dont les coefficients sont des fonctions du temps continues par morceaux. Rappelons que dans le cas de bruits gaussiens et non corrélés, le problème du filtrage linéaire à temps continu a été résolu par Kalman et Bucy (1961). Dans ce cas la loi conditionnelle de X_t est gaussienne de moyenne \hat{X}_t et de covariance $P(t)$. Cette loi est calculée de façon récursive.

Le filtre de Kalman consiste ici à rechercher le meilleur estimateur linéaire de X_t conditionnellement aux observations $Y_s : 0 \leq s \leq t$ au sens que l'erreur d'estimation est de variance minimum. On décrit ci-dessous l'approche innovation.

2.2.1 L'approche innovation

Considérons le modèle d'espace d'état défini par le système (2.11) et posons pour des commodités d'écriture $Z_t = H(t)X_t$. On considère les hypothèses suivantes :

1. $\{X_t\}$ est un processus du second ordre, continu en moyenne quadratique.
2. Pour tous s et t tels que $s < t$, $(W_t - W_s) \perp \mathcal{H}_s^{W,X} = \overline{\mathcal{L}\{W_u, X_u, u \leq s\}}$
3. $\forall t \geq 0, G(t)G'(t)$ est une matrice définie positive.

Définition 30 On appelle processus d'innovation le processus $(\nu_t, t > 0)$ défini par

l'équation

$$d\nu_t = dY_t - \widehat{Z}_t dt \quad (2.12)$$

où $\widehat{Z}_t = \mathcal{P}_{\mathcal{H}_t^Y} Z_t = H(t)\widehat{X}_t$.

En remplaçant l'observation Y_t par son expression (équation d'observation), l'innovation ν_t devient

$$d\nu_t = \widetilde{Z}_t dt + G(t)dW_t \quad (2.13)$$

où $\widetilde{Z}_t = Z_t - \widehat{Z}_t = H(t)\widetilde{X}_t$ est l'erreur d'estimation.

Remarques

1. L'existence du processus des innovations est subordonnée à celle de l'intégrale $\int_0^t \widehat{Z}_s ds$, en d'autres termes, il faut que \widehat{Z}_t soit une fonction mesurable de t . Or, le processus (X_t) étant continu en m.q (hypothèse 1 ci-dessus) ceci assure la continuité en m.q de \widehat{Z}_t , d'où l'on déduit ensuite en utilisant un résultat de Doob ([16]) qu'il existe une version mesurable en t du processus \widehat{Z}_t , ce qui donne un sens à l'intégrale $\int_0^t \widehat{Z}_s ds$ et au processus des innovations.
2. Si les p.a.o (V_t) et (W_t) sont gaussiens, alors le processus des innovations est gaussiens.

Théorème 6 *Sous les hypothèses (1),(2) et (3) ci-dessus, le processus des innovations $(\nu_t : t \geq 0)$ est un p.a.o et de matrice de covariance*

$$\text{cov}(\nu_t) = \int_0^t G(s)G'(s)ds$$

Démonstration

Notons d'abord que $\nu_s \in \mathcal{H}_s^Y$ et par suite montrer que le processus des accroissements ν est un p.a.o, revient à montrer que $(\nu_t - \nu_s) \perp \mathcal{H}_s^Y$. Il suffit alors de montrer que

toutes les composantes du vecteur $\nu_t - \nu_s$ sont orthogonales à \mathcal{H}_s^Y . Soit $X \in \mathcal{H}_s^Y$ et notons ν_t^i la i^{ieme} composante de ν_t et g_{ij} les coefficients de la matrice G . On a

$$E[X(\nu_t^i - \nu_s^i)] = E \left[X \left\{ \left(Y_t^i - \int_0^t \widehat{Z}_u^i du \right) - \left(Y_s^i - \int_0^s \widehat{Z}_u^i du \right) \right\} \right]$$

qui s'écrit encore en remplaçant Y_t^i par son expression de l'équation d'observation

$$E[(X(\nu_t^i - \nu_s^i))] = E \left[X \left(\int_s^t \widetilde{Z}_u^i du \right) \right] + \sum_j E \left(X \int_s^t g^{ij}(u) dW_u^j \right)$$

Il s'agit de montrer que cette somme est nulle. Or, la nullité du premier terme résulte en appliquant Fubini et ensuite l'orthogonalité de \widetilde{Z}_u^i avec \mathcal{H}_s^Y pour tout $s \leq u$ et la nullité du second terme résulte directement de l'hypothèse (2). Ainsi, $(\nu_t^i - \nu_s^i) \perp \mathcal{H}_s^Y$ d'où le résultat.

Concernant la détermination de la matrice de covariance de ν_t , on peut écrire compte tenu de ce que ce processus est centré que

$$\begin{aligned} E(\nu_t \nu_t') &= E \left\{ \left[\int_0^t \widetilde{Z}_s + \int_0^t G(s) dW(s) \right] \cdot \left[\left(\int_0^t \widetilde{Z}_s ds \right)' + \left(\int_0^t G(s) dW(s) \right)' \right] \right\} \\ &= E \left\{ \left(\int_0^t \widetilde{Z}_s ds \right) \cdot \left(\int_0^t \widetilde{Z}_s ds \right)' \right\} + E \left\{ \left(\int_0^t G(s) dW(s) \right) \left(\int_0^t G(s) dW(s) \right)' \right\} \\ &+ E \left\{ \left(\int_0^t \widetilde{Z}_s ds \right) \cdot \left(\int_0^t G(s) dW(s) \right)' \right\} + E \left\{ \left(\int_0^t G(s) dW(s) \right) \cdot \left(\int_0^t \widetilde{Z}_s ds \right)' \right\} \end{aligned}$$

Le second terme de cette somme est d'après la proposition 7 chap.0, égale à

$$E \left\{ \left(\int_0^t G(s) dW(s) \right) \left(\int_0^t G(s) dW(s) \right)' \right\} = \int_0^t G(s) G'(s) ds$$

Et au vu de ce résultat, nous devons prouver que les trois autres termes de cette somme sont nuls. Ce qui revient à montrer que les coefficients de chacune de ces matrices sont nuls.

1. Étude du premier terme

Soit $\{0 = t_0, t_1, \dots, t_n = t\}$ une subdivision de $[0, t]$ et écrivons

$$E \left\{ \left(\int_0^t \widetilde{Z}_i(s) ds \right) \cdot \left(\int_0^t \widetilde{Z}_j(s) ds \right) \right\} = \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} E \left(\int_{t_k}^{t_{k+1}} \int_{t_l}^{t_{l+1}} \widetilde{Z}_i(u) \widetilde{Z}_j(v) du dv \right)$$

En utilisant le théorème de Fubini et ensuite l'inégalité de Schwartz, on voit que

$$\begin{aligned}
E \left(\int_{t_k}^{t_{k+1}} \int_{t_l}^{t_{l+1}} |\tilde{Z}_i(u) \tilde{Z}_j(v)| dudv \right) &= \int_{t_k}^{t_{k+1}} \int_{t_l}^{t_{l+1}} |E(\tilde{Z}_i(u) \tilde{Z}_j(v))| dudv \\
&\leq \int_{t_k}^{t_{k+1}} \int_{t_l}^{t_{l+1}} \|\tilde{Z}_i(u)\|_2 \|\tilde{Z}_j(v)\|_2 dudv \\
&= \int_{t_k}^{t_{k+1}} \int_{t_l}^{t_{l+1}} \sqrt{\tilde{\Gamma}_i(u, u)} \sqrt{\tilde{\Gamma}_j(v, v)} dudv \\
&\leq \left[\int_{t_k}^{t_{k+1}} \sqrt{\tilde{\Gamma}_i(u, u)} du \right] \cdot \left[\int_{t_k}^{t_{k+1}} \sqrt{\tilde{\Gamma}_j(u, u)} du \right]
\end{aligned}$$

où $\tilde{\Gamma}_i(u, v) = cov(\tilde{Z}_i(u), \tilde{Z}_i(v))$ et où $\|\cdot\|_2$ désigne la norme dans L^2 .

Or, $\tilde{\Gamma}_i(u, u) = \|\tilde{Z}_i(u)\|_2^2$ est bornée sur $[0, t]$ car $(Z_i(u))$ est continue en m.q ce qui assure d'après le lemme (1), chap(0), que la fonction de covariance Γ_i du processus scalaire $Z_i(\cdot)$ est continue sur $[0, t]$ et l'inégalité $\|\tilde{Z}_i(u)\|_2^2 \leq \|Z_i(u)\|_2^2 < +\infty$ sur $[0, t]$ permet de conclure à la bornitude de la fonction $u \mapsto \tilde{\Gamma}_i(u, u)$. Et donc si K_i désigne la borne supérieure de $\tilde{\Gamma}_i(u, u)$, on obtient

$$\begin{aligned}
\left| E \left[\int_0^t \int_0^t \tilde{Z}_i(u) \tilde{Z}_j(v) dudv \right] \right| &\leq \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} E \left(\int_{t_k}^{t_{k+1}} \int_{t_l}^{t_{l+1}} |\tilde{Z}_i(u) \tilde{Z}_j(v)| dudv \right) \\
&\leq \sqrt{K} \sum_{k,l} (t_{k+1} - t_k)(t_{l+1} - t_l) \\
&\leq t\sqrt{K} \max_k (t_{k+1} - t_k) \rightarrow 0
\end{aligned}$$

lorsque $\max_k (t_{k+1} - t_k) \rightarrow 0$ et où $K = K_i K_j$. Et donc

$$E \left\{ \left(\int_0^t \tilde{Z}_s ds \right) \cdot \left(\int_0^t \tilde{Z}_s ds \right)' \right\} = 0.$$

2. S'agissant du troisième terme, on trouve en utilisant l'inégalité de Schwartz d'abord et l'isométrie de Wiener ensuite

$$\begin{aligned}
\sum_k \sum_l \left| E \left(\int_{t_k}^{t_{k+1}} \tilde{Z}_i(u) du \int_{t_l}^{t_{l+1}} g_{ij}(u) dW_j(u) \right) \right| &\leq \sum_k \sum_l \left(E \left[\int_{t_k}^{t_{k+1}} \tilde{Z}_i(u) du \right]^2 \right)^{\frac{1}{2}} \\
&\times \left(\int_{t_l}^{t_{l+1}} (g_{ij}(u))^2 du \right)^{\frac{1}{2}} \\
&\leq K_1 \sum_{k,l} (t_{k+1} - t_k)(t_{l+1} - t_l)^{\frac{1}{2}} \\
&\leq K_1 t \max_k (t_{k+1} - t_k)^{\frac{1}{2}} \rightarrow 0
\end{aligned}$$

qui tend vers 0 quand $\max_k (t_{k+1} - t_k)^{\frac{1}{2}} \rightarrow 0$. D'ici on voit que le quatrième terme est nul. Ainsi

$$cov(\nu_t) = E(\nu_t \nu_t') = \int_0^t G(s) G'(s) ds.$$

■

Le résultat suivant précise que le processus des innovations engendre les mêmes sous espaces que le processus des observations $\{(Y_t) : t > 0\}$ (Davis [21]).

Théorème 7 *Sous les hypothèses 1, 2 et 3 et pour tout t , on a*

$$\mathcal{H}_t^Y = \mathcal{H}_t^\nu.$$

2.2.2 Le filtre de Kalman

On rappelle que le filtre de Kalman est un schéma récursif qui permet de calculer l'estimateur linéaire optimal de l'état X_t en fonction seulement de l'observation (Y_t) et de l'estimation (\hat{X}_s) où $s < t$. Ce résultat est précisé par le théorème suivant :

Théorème 8 (Le filtre de Kalman)

Sous les hypothèses 1, 2 et 3 ci-dessus, l'estimateur \hat{X}_t est solution de l'équation différentielle stochastique suivante :

$$\begin{cases} d\hat{X}_t &= A\hat{X}_t dt + K(t) d\nu_t, \\ \hat{X}_0 &= E(X_0) = m_0. \end{cases} \quad (2.14)$$

où $K(t) = P(t)H'(t)(G(t)G'(t))^{-1}$ est le gain de Kalman et $P(t) = E(\tilde{X}_t\tilde{X}_t')$ est la matrice de covariance de l'erreur du filtrage.

La matrice $P(t)$ satisfait l'équation matricielle de **Riccati**

$$\begin{cases} \frac{dP}{dt} &= C(t)C'(t) - P(t)H'(t)(G(t)G'(t))^{-1}H(t)P(t) + A(t)P(t) + P(t)A'(t) \\ P(0) &= Cov(X_0) = P_0. \end{cases} \quad (2.15)$$

Démonstration

Pour alléger les écritures des expressions ci-dessous, on omettra des fonctions matricielles la variable temps t . La démonstration de ce théorème sera exécutée sur plusieurs étapes.

Et pour commencer on examine le cas $E(X_0) = 0$ ce qui entraîne que $\forall t, E(X_t) = 0$.

Étape 1

Considérons la matrice $D(t)$ définie par $D = (GG')^{-\frac{1}{2}}$ telle que :

$$DD' = D'D = (GG')^{-1}$$

et soit (π_t) un processus défini par $\pi_t = \int_0^t D(s)d\nu_s$, où ν_s est le processus d'innovation, alors de la proposition (7) du chapitre (0), on a

$$cov(\pi_t) = \int_0^t DGG'D' ds = I \cdot t$$

où I est la matrice identité. Ainsi π_t est un p.a.o.s et comme $\mathcal{H}_t^\pi = \mathcal{H}_t^\nu = \mathcal{H}_t^Y$ et $\hat{X}_t \in \mathcal{H}_t^Y$, d'après le corollaire (3) chapitre 1, on obtient

$$\hat{X}_t = \int_0^t k(t, s)d\pi_s \quad (2.16)$$

où $k(t, s) = \frac{d}{ds}E(X_t\pi_s')$

Étape 2

On définit le processus Q_t par

$$Q_t = \hat{X}_t - \hat{X}_0 - \int_0^t A(s)\hat{X}_s ds \quad (2.17)$$

Et on montre que (Q_t) est un p.a.o. En effet, pour tous $\forall s < t$

$$Q_t - Q_s = \widehat{X}_t - \widehat{X}_s - \int_s^t A\widehat{X}_u du$$

et comme

$$\mathcal{P}_{\mathcal{H}_s^Y} \widehat{X}_t = \mathcal{P}_{\mathcal{H}_s^Y} \mathcal{P}_{\mathcal{H}_t^Y} X_t = \mathcal{P}_{\mathcal{H}_s^Y} X_t \quad (2.18)$$

alors

$$\mathcal{P}_{\mathcal{H}_s^Y} \left(\int_s^t A\widehat{X}_u du \right) = \mathcal{P}_{\mathcal{H}_s^Y} \int_0^t AX_u du \quad (2.19)$$

En utilisant les deux formules (2.18) et (2.19) données ci avant et d'après la linéarité de la projection, on obtient

$$\begin{aligned} \mathcal{P}_{\mathcal{H}_s^Y} (Q_t - Q_s) &= \mathcal{P}_{\mathcal{H}_s^Y} \left(X_t - X_s - \int_s^t AX_u du \right) \\ &= \mathcal{P}_{\mathcal{H}_s^Y} \left(\int_s^t C dV_t \right) \\ &= 0. \end{aligned}$$

ce qui montre que (Q_t) est un p.a.o par rapport à $\{\mathcal{H}_t^Y\}$ et donc aussi par rapport à $\{\mathcal{H}_t^\pi\}$, alors on peut appliquer la proposition (6) du chapitre (0) pour conclure qu'il existe une fonction g à valeurs matricielles, telle que pour tous t on a

$$\sum_{i,j} \int_0^t (g^{ij}(s))^2 ds < \infty \quad \text{telle que} \quad Q_t = \int_0^t g(s) d\pi_s.$$

et l'on obtient le système différentiel stochastique suivant

$$\begin{aligned} d\widehat{X}_t &= A\widehat{X}_t dt + g(t) d\pi_t \\ \widehat{X}_0 &= E(X_0) = 0. \end{aligned} \quad (2.20)$$

On aura l'équation du filtre de Kalman quand on aura déterminé la matrice g .

Étape 3

Pour déterminer g , on voit que l'équation (2.20) admet une solution de la forme

$$\begin{aligned} \widehat{X}_t &= \Phi(t, 0) \widehat{X}_0 + \int_0^t \Phi(t, s) g(s) d\pi_s \\ &= \int_0^t \Phi(t, s) g(s) d\pi_s \end{aligned}$$

où $\Phi(t, s)$ est la matrice de transition correspondante à A . Comparons ceci avec (2.16), nous voyons que

$$\Phi(t, s)g(s) = k(t, s)$$

et pour $s = t$ la première propriété de la résolvante Φ montre que

$$g(t) = \Phi(t, t)g(t) = k(t, t)$$

On cherche maintenant à calculer $E(X_t \pi'_s)$, l'expression de g s'en déduira immédiatement.

Étape 4

On a $X_t \perp \mathcal{H}_s^W$ et l'expression (2.13) permet d'écrire

$$\begin{aligned} E(X_t \pi'_s) &= E(X_t \int_0^s (DH \tilde{X}_u)' du) + E(X_t (\int_0^s DG dW_u)') \\ &= E(X_t \int_0^s (DH \tilde{X}_u)' du) \\ &= \int_0^s E(X_t \tilde{X}_u') H' D' du \end{aligned}$$

Pour $t > u$, l'équation d'état possède la solution suivante

$$X_t = \phi(t, u)X_u + \int_u^t \Phi(t, \tau)C dV_\tau$$

le deuxième terme est orthogonal à $\mathcal{H}_u^{V,W}$ et comme $\tilde{X}_u \in \mathcal{H}_u^{V,W}$, on obtient

$$\begin{aligned} E(X_t \tilde{X}_u') &= \Phi(t, u)E(X_u \tilde{X}_u') \\ &= \Phi(t, u)E(\tilde{X}_u \tilde{X}_u') \\ &= \Phi(t, u)P(u) \end{aligned}$$

D'où

$$E(X_t \pi'_s) = \int_0^s \Phi(t, u)P(u)H'(u)D'(u)du$$

donc

$$g(t) = k(t, t) = \frac{d}{dt}E(X_t \pi'_t) = \Phi(t, t)P(t)H'(t)D'(t) = P(t)H'(t)D'(t)$$

on obtient donc l'équation du filtre

$$\begin{aligned}
d\widehat{X}_t &= A\widehat{X}_t dt + PH'D' d\pi_t \\
&= A\widehat{X}_t dt + PH'D' Dd\nu_t \\
&= A\widehat{X}_t dt + PH'(GG')^{-1} d\nu_t
\end{aligned}$$

Ce qui complète la démonstration de la première partie du théorème sous l'hypothèse $E(X_0) = 0$.

Étape 5

On remplace ν_t par son expression dans l'équation précédente, on obtient

$$d\widehat{X}_t = A\widehat{X}_t dt + PH'D'D(H\widetilde{X}_t dt + GdW_t)$$

Soustrayons ceci de l'équation d'état, il vient

$$d\widetilde{X}_t = (A - PH'D'DH)\widetilde{X}_t dt + CdV_t - PH'D'DGdW_t$$

et soit $\Psi(t, s)$ la matrice de transition correspondante à $(A - PH'D'DH)$. Par suite la solution de cette équation est donnée par

$$\widetilde{X}_t = \Psi(t, 0)X_0 + \int_0^t \Psi(t, s)C(s)dV_s - \int_0^t \Psi(t, s)PH'D'DGdW_s$$

Utilisant les hypothèses $X_0 \perp \mathcal{H}_t^V$, $X_0 \perp \mathcal{H}_t^W$ et $\mathcal{H}_t^W \perp \mathcal{H}_t^V$, la proposition (7) chapitre (0) assure que

$$\begin{aligned}
P(t) &= E(\widetilde{X}_t \widetilde{X}_t') \\
&= \Psi(t, 0)P_0\Psi'(t, 0) + \int_0^t \Psi(t, s)C(s)C'(s)\Psi'(t, s)ds \\
&+ \int_0^t \Psi(t, s)PH'D'DGG'DD'HP'\Psi'_t(t, s)ds \\
&= \Psi(t, 0)P_0\Psi'(t, 0) + \int_0^t \Psi CC'\Psi' ds + \int_0^t \Psi PH'D'DHP'\Psi' ds
\end{aligned}$$

Alors, (2.15) est déduite par différenciation de cette équation par rapport à t .

Prenons par exemple le second terme, on utilise la quatrième propriété de la résolvante

et que $\frac{d}{dt} \int_0^t f(t, u) du = \int_0^t \frac{d}{dt} f(t, u) du + f(t, t)$ on obtient

$$\begin{aligned} \frac{d}{dt} \left(\int_0^t \Psi(t, s) C(s) C'(s) \Psi'(t, s) ds \right) &= CC' + \int_0^t \left(\frac{\partial}{\partial t} \Psi \right) CC' \Psi' ds + \int_0^t \left(\Psi CC' \frac{\partial}{\partial t} \Psi' \right) ds \\ &= CC' + (A - PH' D' DH) \int_0^t \Psi CC' \Psi' ds \\ &\quad + \int_0^t \Psi CC' \Psi' ds (A' - H' D' DHP) \end{aligned}$$

et on fait de même pour le troisième terme, on trouve

$$\begin{aligned} \frac{dP}{dt} &= CC' + (PH' D' DHP) + (A - PH' D' DH)P + P(A' - H' D' DHP) \\ &= CC' - PH' D' DHP + AP + PA' \end{aligned}$$

Pour $t = 0$, $P(0) = \Psi(0, 0) P_0 \Psi(0, 0) = P_0$ où $P_0 = P(0)$.

Étape 6

Finalement, on suppose que $m_0 \neq 0$ donc $m(t) = E(X_t)$ satisfait l'équation différentielle

$$\frac{dm}{dt} = Am(t) \text{ où } m(0) = m_0 \quad (2.21)$$

et on définit le processus centré par

$$X_t^c = X_t - m(t)$$

c-à-d \widehat{X}_t^c satisfait l'équation 2.14 avec $\widehat{X}_0^c = E(X_0^c) = 0$

Quant à $\widehat{X}_t = \widehat{X}_t^c + m(t)$, il satisfait la même équation avec $\widehat{X}_0 = E(X_0)$ non nécessairement nul. Ce qui termine la démonstration . ■

Remarques

1. Une des raisons que le théorème du filtre de Kalman fournit une solution effective et calculatoire au problème du filtrage est que l'équation de covariance

(2.15) est non aléatoire, donc on peut calculer $P(t)$ en premier et ensuite on pourra calculer \widehat{X}_t comme solution du système de l'équation d'état.

2. Un résultat dont on admette (voir [21]) affirme que si la matrice des conditions initiales d'un système différentielle de Riccati est définie positive (c'est bien le cas ici avec $P(0) = Cov(X_0)$), la solution $P(t)$ est définie positive, unique et bornée sur tout intervalle $[0, t]$.
3. Si le modèle d'espace d'état stationnaire est scalaire, la solution de l'équation de Riccati est donnée par :

$$P(t) = \frac{\mathcal{N}(t)}{\mathcal{D}(t)}$$

telle que

$$\mathcal{N}(t) = G^2[P(0)(\phi + A) + C^2] + G^2[P(0)(\phi - A) - C^2]e^{-2\phi t}$$

et

$$\mathcal{D}(t) = [H^2P(0) + G^2(\phi - A)] - [H^2P(0) - G^2(A + \phi)]e^{-2\phi t}$$

où $\phi(t) = \sqrt{A^2 + \frac{H^2C^2}{G^2}}$, et pour plus de détail sur la solution des équations de Riccati, voir[24].

4. On peut écrire \widehat{X}_t sous la forme suivante :

$$d\widehat{X}_t = (A(t) - K(t) \times H(t))\widehat{X}_t dt + K(t)dY_t$$

cette formule explique bien la récursivité de l'estimateur.

2.2.3 Synoptique d'un filtre de Kalman

La figure (2.1) représente un schéma qui résume le principe du filtre de Kalman pour le système d'espace d'état (2.11) donné ci dessus.

2.2.4 Le filtre de Kalman-Bucy,1961

On a étudié le problème du filtrage linéaire dans le cas général où les bruits d'état et de mesure sont des processus à accroissements orthogonaux. En particulier,

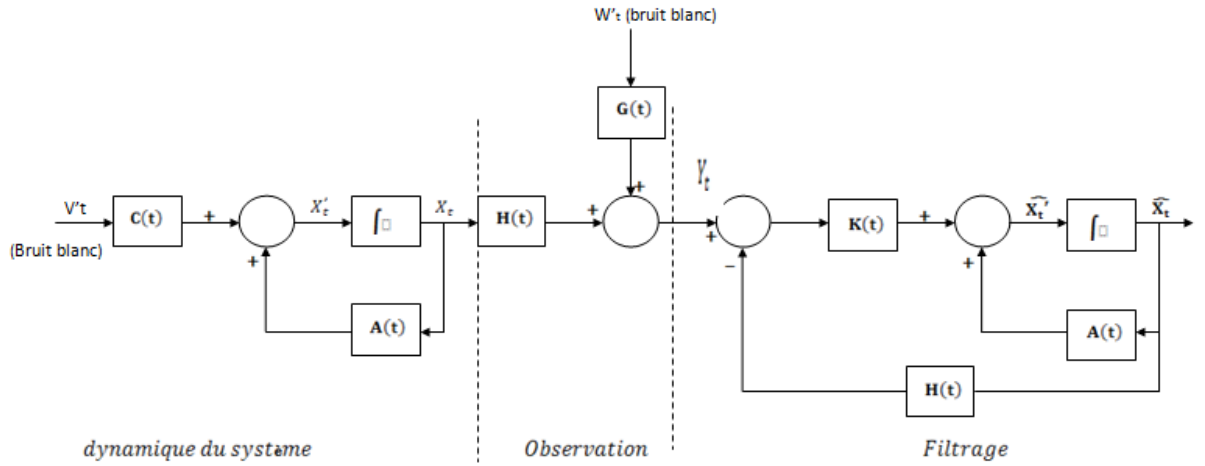


FIGURE 2.1 – Schéma de principe

si on suppose que (V_t) et (W_t) sont deux mouvements browniens indépendants où V_t est indépendant de W_t et X_0 on va obtenir la même solution et dans ce cas le filtre est appelée **le filtre de Kalman-Bucy**.

Théorème 9 (*Le filtre de Kalman-Bucy*)

Considérons le modèle d'espace d'état (2.11) dans le cas unidimensionnel. L'estimateur \hat{X}_t de l'état du système satisfait l'équation différentielle stochastique suivante :

$$d\hat{X}_t = \left(A(t) - \frac{H^2(t)P(t)}{G^2(t)} \right) \hat{X}_t dt + \frac{H(t)P(t)}{G^2(t)} dY_t, \quad \hat{X}_0 = E(X_0)$$

où $P(t) = E(X_t - \hat{X}_t)^2$ satisfait l'équation déterministe de Riccati :

$$\frac{dP}{dt} = 2A(t)P(t) - \frac{H^2(t)P^2(t)}{G^2(t)} + C^2(t), \quad P(0) = E(X_0 - E(X_0))^2.$$

La démonstration de ce théorème est faite avec les mêmes étapes de la démonstration du théorème (8), en remplaçant l'intégrale de Wiener par l'intégrale de Itô, on trouve facilement le résultat, voir [4].

On applique ci-dessus le filtre de kalman-Bucy à deux exemple différents.

Exemple 4 (Des observations bruitées d'un processus constant) Considérons le cas simple :

$$\begin{aligned} dX_t &= 0, i.e X_t = X_0 & E(X_0) &= 0, & E(X_0^2) &= a^2; \\ dY_t &= X_t dt + m dW_t, & Y_0 &= 0. \end{aligned}$$

D'après le théorème précédent l'équation de Riccati est de la forme :

$$\frac{dP(t)}{dt} = -\frac{1}{m^2}P^2(t), \text{ où } P(0) = a^2$$

ainsi

$$\frac{dP(t)}{P^2(t)} = \frac{-1}{m^2} dt$$

la solution de cette équation différentielle non linéaire du premier ordre est :

$$P(t) = \frac{a^2 m^2}{m^2 + a^2 t}$$

donc

$$d\hat{X}_t = -\frac{a^2}{m^2 + a^2 t} \hat{X}_t dt + \frac{a^2}{m^2 + a^2 t} dY_t, \text{ où } \hat{X}_0 = E(X_0) = 0.$$

c'est une équation différentielle stochastique linéaire du premier ordre, la formule de Itô permet d'écrire :

$$\hat{X}_t = \exp \left\{ -\int_0^t \frac{a^2}{m^2 + a^2 s} ds \right\} \left(\hat{X}_0 + \int_0^t \exp \left(\int_0^s \frac{a^2}{m^2 + a^2 u} du \right) \frac{a^2}{m^2 + a^2 s} dY_s \right)$$

Alors

$$d \left(\hat{X}_t \exp \left\{ \int_0^t \frac{a^2}{m^2 + a^2 s} ds \right\} \right) = \exp \left(\int_0^t \frac{a^2}{m^2 + a^2 t} dt \right) \frac{a^2}{m^2 + a^2 t} dY_t$$

donc

$$\hat{X}_t = \frac{a^2}{m^2 + a^2 t} Y_t \Rightarrow \hat{X}_t = \frac{1}{\frac{m^2}{a^2} + t} Y_t, \text{ pour tout } t \geq 0$$

Exemple 5 (les coefficients constants) Considérons le système d'espace d'état stationnaire suivant :

$$\begin{aligned} dX_t &= AX_t dt + CdV_t \\ dY_t &= HX_t + GdW_t \end{aligned}$$

où A, C, H et G sont des constants non nuls. L'équation de Riccati correspondante est :

$$P' = 2AP - \frac{H^2}{G^2}P^2 + C^2, \quad P(0) = a^2$$

admet la solution :

$$P(t) = \frac{\alpha_1 - K\alpha_2 \exp\left(\frac{(\alpha_2 - \alpha_1)H^2 t}{G^2}\right)}{1 - K \exp\left(\frac{(\alpha_2 - \alpha_1)H^2 t}{G^2}\right)}$$

où

$$\alpha_1 = H^{-2}(AG^2 - G\sqrt{A^2G^2 + H^2C^2})$$

$$\alpha_2 = H^{-2}(AG^2 + G\sqrt{A^2G^2 + H^2C^2})$$

et

$$K = \frac{a^2 - \alpha_1}{a^2 - \alpha_2}$$

donc la solution \widehat{X}_t est de la forme suivante :

$$\widehat{X}_t = \exp\left(\int_0^t F(s)ds\right) \widehat{X}_0 + \frac{H}{G^2} \int_0^t \exp\left(\int_0^s F(u)du\right) P(s) dY_s$$

où

$$F(s) = A - \frac{H^2}{G^2}P(s)$$

et pour des grandes valeurs de s , on a $P(s) \approx \alpha_2$, ce qui donne

$$\begin{aligned} \widehat{X}_t &\approx \widehat{X}_0 \exp\left(A - \frac{H^2}{G^2}\alpha_2 t\right) + \frac{H}{G^2}\alpha_2 \int_0^t \exp\left(A - \frac{H^2}{G^2}\alpha_2(t-s)\right) dY_s \\ &\approx \widehat{X}_0 \exp(-\beta t) + H \frac{\alpha_2}{G^2} \exp(-\beta t) \int_0^t \exp(-\beta s) dY_s \end{aligned}$$

où $\beta = G^{-1}\sqrt{A^2G^2 + H^2C^2}$.

Chapitre 3

Le filtrage non linéaire

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ un espace de probabilisé, (V_t) et (W_t) deux mouvements browniens indépendants définis sur cet espace. Considérons le problème du filtrage non linéaire associé au modèle d'espace d'état suivant :

$$\begin{cases} dX_t &= b(X_t)dt + f(X_t)dV_t \\ dY_t &= h(X_t)dt + dW_t \end{cases} \quad (3.1)$$

où les coefficients b , f et h sont des applications continues à valeurs réelles satisfaisant chacune la propriété de Lipschitz :

$$\begin{cases} |b(x_1) - b(x_2)| + |f(x_1) - f(x_2)| &\leq M_1|x_1 - x_2| \\ |h(x_1) - h(x_2)| &\leq M_2|x_1 - x_2| \end{cases}$$

et la condition de croissance linéaire

$$\begin{cases} |b(x)|^2 + |f(x)|^2 &\leq M_1^2(1 + |x|^2) \\ |h(x)|^2 &\leq M_2^2(1 + |x|^2) \end{cases}$$

et où M_1 et M_2 sont des constantes positives.

Pour, $t = 0$ on suppose que X_0 est une variable aléatoire indépendante de (V_t) et (W_t) .

On suppose en outre que pour tout $t > 0$:

$$E\left(\int_0^t |h(X_s)|^2 ds\right) < \infty.$$

Le problème du filtrage non linéaire consiste ici, à déterminer la distribution conditionnelle de X_t sachant \mathcal{F}_t^Y :

$$\Pi_t(\phi) = E_p(\phi(X_t)/\mathcal{F}_t^Y), \phi \in \mathcal{C}_b^2$$

où rappelons que \mathcal{F}_t^Y est la σ -algèbre engendrée par les variables $\{Y_s : s < t\}$, E_p est l'espérance par rapport à la loi de probabilité P et \mathcal{C}_b^2 est la classe des fonctions deux fois continuellement différentiables et bornées .

La nature des équations du système d'état (3.1) impose une autre idée de l'intégrale stochastique. Et contrairement à l'intégrale de Wiener, il s'agit ici de l'intégrale stochastique au sens de ITÔ.

Les techniques du calcul différentiel stochastique développées au cours des années 1960-1970 ont permis à H.Kushner de prouver que le filtre $\Pi(t)$ associé au problème du filtrage non linéaire décrit par le système 3.1 est solution d'une équation aux dérivées partielles stochastique non linéaire appelée équation de Kushner-Stratonovich. Ce résultat a été obtenu par autre approche initiée par M.Fujisaki,G.Kallianpur, H.Kunita [20] en utilisant la notion de processus d'innovation. Toutefois la non linéarité de l'équation de Kushner-Stratonovich rend difficile l'étude des propriétés de sa solution. Ce problème a été résolu par M.Zakai qui a prouvé moyennant un changement de probabilité que le filtre non normalisé est solution d'une équation aux dérivées partielles stochastique linéaire qui porte son nom.

Comme pour le cas du filtrage linéaire, nous examinerons là également, le problème du filtrage non linéaire à temps discret et ensuite le cas du filtrage non linéaire à temps continu.

3.1 Le filtrage non linéaire à temps discret

Le modèle d'espace d'état (3.1) s'écrit dans ce cas sous la forme suivante

$$\begin{cases} X_{k+1} &= b_{k+1}(X_k) + V_{k+1} \\ Y_{k+1} &= h_{k+1}(X_{k+1}) + W_{k+1} \end{cases} \quad (3.2)$$

où la variable d'état $X_k \in \mathbb{R}^n$ avec $X_0 = V_0$ et l'observation $Y_k \in \mathbb{R}^m$, b_k , h_k sont deux fonctions vectorielles mesurables, non linéaires de dimensions n et m respectivement avec h_k continue et bornée. Les processus $\{V_k, k \geq 1\}$ et $\{W_k, k \geq 1\}$ sont deux bruits blancs gaussiens centrés indépendants entre eux et indépendants de X_0 et de matrice de covariance respective Q_k, R_k .

Considérons la suite d'observations $\mathcal{Y}_l = \{Y_1, \dots, Y_l\}$ et rappelons que le filtrage non linéaire à temps discret consiste à déterminer la distribution conditionnelle de X_{k+1} sachant \mathcal{Y}_{k+1} . Il s'agit alors de déterminer à l'instant $k+1$, la densité conditionnelle notée $\varphi_{X_{k+1}}^{\mathcal{Y}_{k+1}=\underline{y}_{k+1}}$ étant donnée la densité initiale φ_0 de l'état initial (X_0) et où $\underline{y}_{k+1} = (y_0, y_1, \dots, y_{k+1})$.

Pour alléger les notations, on convient de noter cette densité de probabilité par $\varphi_{X_{k+1}}^{\underline{y}_{k+1}}$.

Théorème 10 (*La densité conditionnelle du filtrage discret*)

Supposons que la densité $\varphi_{X_{k+1}}^{\underline{y}_{k+1}}$ existe, alors

$$\varphi_{X_{k+1}}^{\underline{y}_{k+1}}(x_{k+1}) = \frac{\varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) \int \varphi_{X_{k+1}}^{x_k}(x_{k+1}) \varphi_{X_k}^{y_k}(x_k) dx_k}{\int \int \varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) \varphi_{X_{k+1}}^{x_k}(x_{k+1}) \varphi_{X_k}^{y_k}(x_k) dx_k dx_{k+1}}$$

où $\varphi_{X_{k+1}}^{x_k}(\cdot)$ est la densité conditionnelle de X_{k+1} sachant $X_k = x_k$ appelée densité de transition de l'équation d'état et $\varphi_{Y_{k+1}}^{x_{k+1}}(\cdot)$ est la densité conditionnelle de Y_{k+1} par rapport à $X_{k+1} = x_{k+1}$ laquelle est donnée par la formule suivante

$$\varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) = \frac{1}{(2\pi)^{\frac{m}{2}} |R_k|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (y_{k+1} - h(x_{k+1}))' R_{k+1}^{-1} (y_{k+1} - h(x_{k+1})) \right\}$$

Démonstration

La démonstration, comme dans le cas du filtrage linéaire, comporte deux étapes. La

première est consacrée à la prédiction et consiste à calculer la densité $\varphi_{X_{k+1}}^{y_k}$ et la seconde concerne la correction qui donne la densité du filtrage $\varphi_{X_{k+1}}^{y_{k+1}}$.

Étape prédiction

D'abord notons que l'équation d'états génère un processus de Markov d'ordre 1 qui fait que $\varphi_{X_{k+1}}^{y_k}$ devient

$$\varphi_{X_{k+1}}^{y_k}(x_{k+1}) = \int \varphi_{X_{k+1}, X_k}^{y_k}(x_{k+1}, x_k) dx_k = \int \varphi_{X_{k+1}}^{x_k}(x_{k+1}) \varphi_{X_k}^{y_k}(x_k) dx_k$$

Étape correction

On a

$$\varphi_{X_{k+1}}^{y_{k+1}}(x_{k+1}) = \varphi_{X_{k+1}}^{y_k, y_{k+1}}(x_{k+1})$$

la formule de Bayes jointe à l'équation précédente font que le filtre devient

$$\varphi_{X_{k+1}}^{y_{k+1}}(x_{k+1}) = \frac{\varphi_{Y_{k+1}}^{y_k, x_{k+1}}(y_{k+1}) \varphi_{X_{k+1}}^{y_k}(x_{k+1})}{\varphi_{Y_{k+1}}^{y_k}(y_{k+1})} \quad (3.3)$$

et comme (V_k) et (W_k) sont deux bruits blancs indépendants et indépendants de X_0 , alors conditionnellement à X_{k+1} la variable Y_{k+1} est indépendante des variables $\{Y_s : s \leq k\}$ et donc

$$\varphi_{Y_{k+1}}^{y_k, x_{k+1}}(y_{k+1}) = \varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1})$$

ce qui permet d'écrire

$$\begin{aligned} \varphi_{Y_{k+1}}^{y_k}(y_{k+1}) &= \int \varphi_{Y_{k+1}, X_{k+1}}^{y_k}(y_{k+1}, x_{k+1}) dx_{k+1} \\ &= \int \varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) \varphi_{X_{k+1}}^{y_k}(x_{k+1}) dx_{k+1} \end{aligned}$$

Et comme

$$\varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) = \varphi_{w_{k+1}}(y_{k+1} - h(x_{k+1}))$$

et que $W_k \sim \mathcal{N}(0, R_k)$ alors cette densité conditionnelle devient

$$\varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) = \frac{1}{(2\pi)^{\frac{m}{2}} |R_k|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(y_{k+1} - h(x_{k+1}))' R_{k+1}^{-1} (y_{k+1} - h(x_{k+1}))\right\}$$

D'où l'on déduit

$$\varphi_{X_{k+1}}^{y_{k+1}}(x_{k+1}) = \frac{\varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) \int \varphi_{X_{k+1}}^{x_k}(x_{k+1}) \varphi_{X_k}^{y_k}(x_k) dx_k}{\int \int \varphi_{Y_{k+1}}^{x_{k+1}}(y_{k+1}) \varphi_{X_{k+1}}^{x_k}(x_{k+1}) \varphi_{X_k}^{y_k}(x_k) dx_k dx_{k+1}} \quad (3.4)$$

■

Remarque 5 *L'équation (3.4) n'est analytiquement pas calculable, mais on peut la résoudre en utilisant des approximations de la solution par des méthodes numériques, (Voir [35] et [13]).*

En utilisant la formule de Taylor on peut linéariser le système d'espace d'état pour pouvoir appliquer le filtre de Kalman et obtenir un algorithme récursif sous optimal appelé le filtre de Kalman étendu (FKE) que l'on va décrire dans la sous section suivante.

3.1.1 Le filtre de Kalman étendu (FKE)

Considérons le modèle d'espace d'état (3.2) et supposons que b_{k+1} et h_{k+1} sont deux fonctions dérivables.

L'absence d'une solution analytique au problème de filtrage non linéaire à temps discret impose le recours à une méthode sous-optimale appelée le filtre de Kalman étendu qui consiste à remplacer les équations non linéaires du modèle d'espace d'état par leur développement de Taylor au premier ordre autour de l'estimateur le plus récent de l'état du système (c'est une version non linéaire du filtre de Kalman).

En utilisant développement de Taylor, on linéarise la fonction b_{k+1} autour du filtre $X_{k+1|k+1}$, on obtient

$$b_{k+1}(X) \approx b_{k+1}(X_{k+1|k+1}) + F_{k+1}(X - X_{k+1|k+1})$$

ensuite on linéarise la fonction h_{k+1} autour du prédicteur $X_{k+1|k}$, on trouve

$$h_{k+1}(X) \approx h_{k+1}(X_{k+1|k}) + H_{k+1}(X - X_{k+1|k})$$

Le système d'espace d'état est alors remplacé par :

$$\begin{cases} X_{k+1} &= F_{k+1}(X_k - X_{k/k}) + f_{k+1} + V_{k+1} \\ Y_{k+1} &= H_{k+1}(X_{k+1} - X_{k+1/k}) + h_{k+1} + W_{k+1} \end{cases}$$

où F_{k+1} et H_{k+1} sont les matrices Jacobiennes de b_{k+1} et h_{k+1} aux points $(X_{k+1/k+1})$ et $X_{k+1/k}$, respectivement, et $f_{k+1} = b_{k+1}(X_{k/k})$, $h_{k+1} = h_{k+1}(X_{k+1/k})$.

Il suffit alors d'appliquer les résultats concernant le filtre de Kalman à temps discret (§1, chap2) à ce nouveaux système pour déduire l'algorithme suivant

Étape prédiction

$$\begin{aligned} X_{k+1/k} &= b_{k+1}(X_{k/k}) \\ P_{k+1/k} &= F_{k+1}P_{k/k}F_{k+1}' + Q_{k+1} \end{aligned}$$

Étape Correction

$$\begin{aligned} X_{k+1/k+1} &= X_{k+1/k} + K_{k+1}[Y_{k+1} - h_{k+1}(X_{k+1/k})] \\ P_{k+1/k+1} &= (I - K_{k+1}H_{k+1})P_{k+1/k} \end{aligned}$$

où S_{k+1} et K_{k+1} sont, respectivement, la matrice de covariance du processus des innovations et la matrice de gain de Kalman tels que :

$$\begin{aligned} S_{k+1} &= H_{k+1}P_{k+1/k}H_{k+1}' + R_{k+1} \\ K_{k+1} &= P_{k+1/k}H_{k+1}'S_{k+1}^{-1} \end{aligned}$$

On choisit les initialisations $X_{0/0}$ et $P_{0/0}$ de telle sorte que la loi de X_0 de moyenne $X_{0/0}$ et de covariance $P_{0/0}$ soit une bonne approximation de celle de X_0 .

Remarques

- Le filtre de Kalman étendu n'est pas un filtre optimal (c'est un outil basé sur des approximations), ainsi les matrices $P_{k+1/k+1}$ et $P_{k+1/k}$ ne représentent pas la vraie covariance des états estimés.

- Les matrices F_k et H_k dépendent des estimateurs et par conséquent dépendent des mesures, alors le gain de Kalman et les matrices $P_{k+1/k+1}$ et $P_{k+1/k}$ ne peuvent pas être pré-calculés .

Exemple 6 *Considérant le système d'espace d'état suivant :*

$$\begin{aligned} X_{k+1} &= X_k^2 + V_{k+1} \\ Y_{k+1} &= X_{k+1}^3 + W_{k+1} \end{aligned}$$

avec $E(V_{k+1}) = E(W_{k+1}) = 0, Q_k = 1, R_k = 2, P_{0/0} = 1, X_{0/0} = 2$.

En appliquant le filtre FKE, on trouve que l'estimateur $X_{k+1/k+1}$ de l'état X_{k+1} étant donné les observations Y_1, \dots, Y_{k+1} et la matrice de covariance de l'erreur du filtrage $P_{k+1/k+1}$, sont donnés par

$$\begin{aligned} X_{k+1/k+1} &= X_{k/k}^2 + K_{k+1}[Y_{k+1} - (X_{k+1/k})^3] \\ P_{k+1/k+1} &= [1 - 3K_{k+1}(X_{k+1/k})^2]P_{k+1/k} \end{aligned}$$

où

$$\begin{aligned} K_{k+1} &= \frac{3P_{k+1/k}(X_{k+1/k})^2}{9(X_{k+1/k})^4P_{k+1/k} + 2} \\ P_{k+1/k} &= 4(X_{k/k})^2P_{k/k} + 1. \end{aligned}$$

3.2 Le filtrage non linéaire à temps continu

Rappelons que le problème du filtrage non linéaire à temps continu associé au modèle (3.1), consiste à déterminer la distribution conditionnelle de X_t sachant \mathcal{F}_t^Y :

$$\Pi_t(\phi) = E_p(\phi(X_t)/\mathcal{F}_t^Y), \phi \in \mathcal{C}_b^2.$$

Il s'agit alors de trouver une forme explicite de $\Pi_t(\phi)$ en terme de $(Y_s, s \leq t)$, c'est à dire qu'on cherche à exprimer $\Pi_t(\phi)$ comme solution d'une équation aux dérivées partielles stochastiques par rapport à Y_t . A cette fin nous envisageons deux approches :

1) L'approche changement de mesure

Cette approche consiste à construire une nouvelle mesure \tilde{P} sous laquelle Y_t est un mouvement brownien, et exprimer Π_t moyennant la formule de Striebel-Kallianpur en fonction du filtre non normalisé σ_t . Ce dernier filtre est solution d'une équation aux dérivées partielles stochastique (E.D.P.S) linéaire appelée équation de **Zakai**.

2) L'approche innovation

Cette approche (appelée souvent l'approche FKK (Fujisaki-Kallianpur-Kunita)) consiste à déterminer un mouvement brownien appelé processus d'innovation et en utilisant ce processus, on montre que le filtre Π_t est solution d'une E.D.P.S non linéaire appelée équation de **Kushner**.

3.2.1 Méthode de changement de mesure

Pour que l'intégrale d'Itô par rapport à Y_t soit bien définie, il faut montrer en utilisant le théorème de **Girsanov**, que le processus d'observation Y_t est un mouvement brownien sous une nouvelle mesure de probabilité.

Aussi nous commencerons ce sous paragraphe par le théorème suivant (théorème de Girsanov) qui montre comment une transformation convenable d'un mouvement brownien demeure encore un mouvement brownien sous l'action d'une nouvelle mesure de probabilité.

On a pour commencer le résultat suivant :

Lemme 9 Soit $\{\Omega, \mathcal{F}, P\}$ un espace de probabilité filtré par $\{\mathcal{F}_t, t \geq 0\}$, supposons que \tilde{P} est une autre mesure de probabilité équivalente à P et soit N la dérivée de Radon-Nykodim de \tilde{P} par rapport à P : $N = \frac{d\tilde{P}}{dP}$.

Alors d'une part, le processus $(N_t : t \geq 0)$ défini par

$$N_t = E_P(N/\mathcal{F}_t)$$

est une $(\mathcal{F}_t : t \geq 0)$ martingale et d'autre part, pour que le processus $\{X_t N_t\}$ soit une \mathcal{F}_t - martingale par rapport à P il faut et il suffit que $\{X_t\}$ soit une martingale par rapport à \tilde{P} .

Démonstration

La première partie du lemme est évidente.

Si maintenant $(X_t : t \geq 0)$ est une \tilde{P} martingale et $A \in \mathcal{F}_s (s \leq t)$, alors d'une part, on a

$$\begin{aligned} \int_A X_t d\tilde{P} &= \int_A E^{\mathcal{F}_s}(X_t) d\tilde{P}_s \\ &= \int_A E^{\mathcal{F}_s}(X_t) d\tilde{P} \end{aligned}$$

et comme (X_t) est une \tilde{P} martingale, alors on peut écrire

$$\int_A X_t d\tilde{P} = \int_A X_s d\tilde{P}_s$$

laquelle compte tenu de l'équivalence de \tilde{P} et P fait que

$$\int_A X_t d\tilde{P} = \int_A X_s N_s dP$$

D'autre part on a

$$\begin{aligned} \int_A X_t d\tilde{P} &= \int_A X_t N_t dP \\ &= \int_A E^{\mathcal{F}_s}(X_t N_t) dP \end{aligned}$$

et par l'unicité de l'espérance, on déduit que $\{X_t N_t\}$ est une P -martingale.

La démonstration de la condition nécessaire se fait de la même manière. ■

Théorème 11 (*Caractérisation du mouvement Brownien - Lévy -*)

Soit $\{X_t\}$ un processus continu définie sur $\{\Omega, \mathcal{F}, (\mathcal{F}_t), P\}$.

$\{X_t\}$ est un mouvement brownien si et seulement si $\{X_t\}$ et $\{X_t^2 - t\}, t \geq 0$ sont deux martingales.

Démonstration Voir [20] ■

Théorème 12 (*Girsanov*)

Soit $(w_t : 0 \leq t \leq T)$ un mouvement brownien défini sur l'espace filtré $\{\Omega, \mathcal{F}, (\mathcal{F}_t), P\}$

et soit $g : \Omega \times [0, t] \rightarrow \mathbb{R}$ un processus prévisible, tel que :

$$\int_0^T |g_t|^2 dt < \infty \quad p.s$$

et

$$L_t(g) = \exp\left(-\int_0^t g_s dw_s - \frac{1}{2} \int_0^t |g_s|^2 ds\right)$$

Soit \tilde{P} la mesure de probabilité sur $\{\Omega, \mathcal{F}\}$ définie par : $d\tilde{P}_t = L_t(g) dP_t$. Si

$$E\left(\exp\left(\frac{1}{2} \int_0^T |g_t|^2 dt\right)\right) < \infty^1$$

alors le processus (\tilde{w}_t) défini par :

$$\tilde{w}_t \triangleq w_t + \int_0^t g_s ds$$

est un mouvement brownien sur $\{\Omega, \mathcal{F}, (\mathcal{F}_t), \tilde{P}\}$

Démonstration

Pour que (\tilde{w}_t) soit un \tilde{P} -mouvement brownien, il suffit de montrer selon le théorème (11) ci-dessus que \tilde{w}_t est une martingale continue et que $(\tilde{w}_t^2 - t)$ est une martingale.

En effet,

1. Il est évident que \tilde{w}_t est continu.

1. Cette condition est appelée la condition de Novikov

2. Pour montrer que (\tilde{w}_t) est une \tilde{P} -martingale il suffit d'appliquer le lemme (9) et de montrer que $(L_t\tilde{w}_t)$ est une P -martingale. En effet, la condition de Novikov assure que (L_t) est une martingale (voir [36]) et en appliquant la formule de ITÔ à $L_t(g)$ on obtient :

$$dL_t(g) = -L_t(g)g_tdw_t$$

et d'après la règle de différenciation de ITÔ appliqué au produit $(L_t\tilde{w}_t)$ on a

$$\begin{aligned} d\tilde{w}_tL_t(g) &= L_t(g)d\tilde{w}_t + \tilde{w}_t dL_t - L_t(g)g_tdt \\ &= L_t(g)dw_t + L_t(g)g_tdt - \tilde{w}_tL_t(g)g_tdw_t - L_t(g)g_tdt \\ &= L_t(g)(1 - \tilde{w}_tg_t)dw_t \end{aligned}$$

Et pour conclure que c'est une martingale, il suffit de montrer que $L_t(1 - \tilde{w}_tg_t) \in \mathbb{L}_2(dt \otimes dP)$, ce que nous admettrons dans la suite (voir [11]).

3. On applique la formule d'Itô à (\tilde{w}_t) , on trouve que

$$\begin{aligned} \tilde{w}_t^2 &= 2 \int_0^t \tilde{w}_s d\tilde{w}_s + t \\ \tilde{w}_t^2 - t &= 2 \int_0^t \tilde{w}_s d\tilde{w}_s \end{aligned}$$

en utilisant le même argument que la propriété 2, on peut conclure que $(\tilde{w}_t^2 - t)$ est bien une \tilde{P} -martingale.

■

Théorème 13 (*La formule de Bayes*)

Soit X une v.a intégrable sur $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ et \mathcal{G} la sous σ -algèbre de \mathcal{F} . Si \tilde{P} est une seconde mesure de probabilité équivalente à P avec $Z = \frac{d\tilde{P}}{dP}|_{\mathcal{G}}$, alors :

$$E_P(X/\mathcal{G}) = \frac{E_{\tilde{P}}(XZ/\mathcal{G})}{E_{\tilde{P}}(Z/\mathcal{G})}$$

où $E_{\tilde{P}}$ et E_P sont les espérances sous les mesures P et \tilde{P} respectivement .

Démonstration

Notons d'abord que la variable aléatoire Z est presque sûrement strictement positive et de plus sa \mathcal{G} -mesurabilité fait que pour tout $A \in \mathcal{G}$ on ait

$$\int_A \frac{E_{\tilde{P}}(XZ/\mathcal{G})}{E_{\tilde{P}}(Z/\mathcal{G})} dP = \int_A E_{\tilde{P}}(XZ/\mathcal{G}) d\tilde{P}$$

que l'on peut écrire encore sous la forme

$$\int_A \frac{E_{\tilde{P}}(XZ/\mathcal{G})}{E_{\tilde{P}}(Z/\mathcal{G})} dP = \int_A E_P(X/\mathcal{G}) dP$$

et l'unicité de l'espérance conditionnelle conduit à l'identité recherchée. ■

La distribution conditionnelle non-normalisée

Soit \tilde{P} une mesure non négative sur (Ω, \mathcal{F}) absolument continue par rapport à P de sorte que sa restriction \tilde{P}_t à $\mathcal{F}_t^{V,W} = \sigma\{V_s, W_s : s \leq t\}$ soit définie par sa dérivée de Radon-Nikodym par rapport à la restriction P_t de P sur $\mathcal{F}_t^{V,W}$, donnée par

$$L_t(h) = \exp\left(-\int_0^t h(X_s) dW_s - \frac{1}{2} \int_0^t |h(X_s)|^2 ds\right)$$

Sous l'hypothèse que h est bornée, on montre que le processus $L_t(h)$ est une martingale (voir [17]) de moyenne 1, ce qui assure d'une part que \tilde{P} est une mesure de probabilité et d'autre part le théorème de Girsanov montre que le processus des observations $(Y_t)_{t \geq 0}$ est un \tilde{P} -mouvement brownien. De plus, on montre ([1] page 55)

Proposition 9 *Sous la mesure de probabilité \tilde{P}_t , le mouvement brownien $(Y_t)_{t \geq 0}$ est indépendant de $(V_t)_{t \geq 0}$ et la loi de $(X_t)_{t \geq 0}$ sous \tilde{P} est la même que celle de $(X_t)_{t \geq 0}$ sous P .*

Théorème 14 (La formule de Striebel-Kallianpur)

Pour toute fonction $\phi \in \mathcal{C}_b^2$, on a

$$\Pi_t(\phi) = \frac{E_{\tilde{P}}(\phi(X_t)Z_t/\mathcal{F}_t^Y)}{E_{\tilde{P}}(Z_t/\mathcal{F}_t^Y)}$$

où

$$Z_t = L_t^{-1}(h) = \exp\left(\int_0^t h(X_s)dY_s - \frac{1}{2} \int_0^t |h(X_s)|^2 ds\right)$$

Démonstration

Ce résultat découle directement de la formule de Bayes, en remplaçant dans celle ci X , Z et \mathcal{G} par $\phi(X_t)$, Z_t et \mathcal{F}_t^Y respectivement. ■

On pose pour toute fonction $\phi \in \mathcal{C}_b^2(\mathbb{R})$:

$$\sigma_t(\phi) = E_{\tilde{P}}(\phi(X_t)Z_t/\mathcal{F}_t^Y)$$

appelée **la distribution conditionnelle non-normalisée** et lié au filtre optimal Π_t par la formule de Striebel-Kallianpur

$$\Pi_t(\phi) = \frac{\sigma_t(\phi)}{\sigma_t(1)}$$

Notons au passage que le processus (Z_t) satisfait l'équation différentielle stochastique

$$dZ_t = Z_t h(X_t) dY_t$$

Théorème 15 (L'équation de Zakai)

Pour toute fonction $\phi \in \mathcal{C}_b^2(\mathbb{R})$, le filtre non-normalisé $\sigma_t(\phi)$ est solution de l'équation aux dérivées partielles stochastique (E.D.P.S) suivante

$$\sigma_t(\phi) = \sigma_0(\phi) + \int_0^t \sigma_s(\mathcal{L}\phi) ds + \int_0^t \sigma_s(\mathcal{L}^1\phi) dY_s \quad (3.5)$$

où $\mathcal{L}^1\phi(x) = h(x)\phi(x)$ et \mathcal{L} un opérateur différentielle sur \mathcal{C}^2 défini par

$$\mathcal{L}\phi(x) = \frac{1}{2} f^2(x) \frac{d^2}{dx^2} \phi(x) + b(x) \frac{d}{dx} \phi(x).$$

Démonstration

La formule de Itô appliquée à $\Phi(X_t)$ fournit

$$d\phi(X_t) = \mathcal{L}\phi(X_t)dt + \frac{d}{dx}\phi(X_t)f(X_t)dV_t \quad (3.6)$$

et en appliquant la formule de différenciation au produit $\Phi(X_t).Z_t$, on trouve que

$$d(Z_t\phi(X_t)) = Z_t d\phi(X_t) + \phi(X_t)dZ_t + dZ_t d\phi(X_t)$$

Et comme (Y_t) et (V_t) sont indépendants, le troisième terme de cette égalité est nul et donc

$$d(Z_t\phi(X_t)) = Z_t \left[\mathcal{L}\phi(X_t)dt + \frac{d}{dx}\phi(X_t)f(X_t)dV_t \right] + \phi(X_t)Z_t h(X_t)dY_t$$

et cette dernière s'écrit encore

$$Z_t\phi(X_t) = \phi(X_0) + \int_0^t Z_s \mathcal{L}\phi(X_s)ds + \int_0^t Z_s [\phi(X_s)h(X_s)] dY_s + \int_0^t Z_s \frac{d}{dx}\phi(X_s)f(X_s)dV_s$$

En prenant l'espérance conditionnelle des deux membres de l'égalité et en appliquant le lemme de Fubini stochastique (Voir Annexe A), on obtient

$$\begin{aligned} \sigma_t(\phi) &= \sigma_0(\phi) + E_{\tilde{P}} \left(\int_0^t Z_s \mathcal{L}\phi(X_s)ds / \mathcal{F}_t^Y \right) + E_{\tilde{P}} \left(\int_0^t Z_s \phi(X_s)h(X_s)dY_s / \mathcal{F}_t^Y \right) \\ &+ E_{\tilde{P}} \left(\int_0^t Z_s \left(\frac{d\phi}{dx}(X_s) \right) f(X_s) dV_s / \mathcal{F}_t^Y \right) \\ &= \sigma_0(\phi) + \int_0^t \sigma_s(\mathcal{L}\phi)ds + \int_0^t \sigma_s(\mathcal{L}^1\phi)dY_s. \end{aligned}$$

■

Remarque 6 Si $\phi \equiv 1$, l'équation de Zakai compte tenu de la formule de Striebel - Kallianpur s'écrit sous la forme suivante :

$$\begin{aligned} \sigma_t(1) &= 1 + \int_0^t \sigma_s(h)dY_s \\ &= 1 + \int_0^t \sigma_s(1)\Pi_s(h)dY_s \end{aligned}$$

Et cette dernière admet une solution de la forme

$$\sigma_t(1) = \exp \left(\int_0^t \Pi_s(h)dY_s - \frac{1}{2} \int_0^t \Pi_s^2(h)ds \right). \quad (3.7)$$

D'ailleurs ce résultat suivant montre que $\sigma_t(1)$ est une dérivée de Radon - Nikodym.

Corollaire 4 $\sigma_t(1)$ est la dérivée de Radon-Nikodym de la restriction de \tilde{P}_t à $(\mathcal{F}_t^Y = \sigma\{Y_s : s \leq t\})$ par rapport à la restriction de P_t sur \mathcal{F}_t^Y :

$$\sigma_t(1) = \frac{dP_t|_{\mathcal{F}_t^Y}}{d\tilde{P}_t|_{\mathcal{F}_t^Y}}. \quad (3.8)$$

Démonstration

Remarquons d'abord que la sous σ algèbre \mathcal{F}_t^Y est contenue dans la sous σ algèbre engendrée par $\mathcal{F}_t^{V,W}$. Par suite, pour tout $A \in \mathcal{F}_t^Y$, on a d'une part

$$P_t(A) = \int 1_A dP_t = E_{\tilde{P}}(1_A Z_t)$$

et d'autre part

$$E_{\tilde{P}}(1_A \sigma_t(1)) = E_{\tilde{P}}(1_A E_{\tilde{P}}(Z_t / \mathcal{F}_t^Y)) = P_t(A)$$

d'où l'on tire que

$$\sigma_t(1) = \frac{dP_t|_{\mathcal{F}_t^Y}}{d\tilde{P}_t|_{\mathcal{F}_t^Y}}.$$

■

3.2.2 L'approche innovation

Cette approche consiste en utilisant l'équation de Zakai et la formule de Striebel-Kallainpur, de montrer que le filtre normalisé Π_t est une solution d'une E.D.P.S. A cette fin, on commence par définir le processus d'innovations et établir certaines de ses propriétés.

Processus d'innovation

Définition 31 On appelle processus d'innovation le processus $(\nu_t)_{t \geq 0}$ défini par

$$\nu_t = Y_t - \int_0^t \Pi_s(h) ds$$

Lemme 10 Le processus d'innovation (ν_t) est un P -mouvement brownien adapté à la filtration (\mathcal{F}_t^Y) .

La démonstration résulte directement du théorème de Girsanov et des équations (3.7) et (3.8).

La distribution conditionnelle normalisée

Théorème 16 (*Kushner-Stratonovich*)

La distribution conditionnelle normalisée $\Pi_t(\phi)$ satisfait l'équation aux dérivées partielles stochastique suivante :

$$\Pi_t(\phi) = \Pi_0(\phi) + \int_0^t \Pi_s(\mathcal{L}\phi)ds + \int_0^t \{\Pi_s(\mathcal{L}^1\phi) - \Pi_s(h)\Pi_s(\phi)\}d\nu_s$$

où ν_t est le processus d'innovation.

Démonstration

Il suffit d'appliquer la règle de différenciation de Itô au produit $\frac{\sigma_t(\phi)}{\sigma_t(1)}$:

$$d\left(\frac{\sigma_t(\phi)}{\sigma_t(1)}\right) = \frac{d\sigma_t(\phi)}{\sigma_t(1)} + \sigma_t(\phi)d\left(\frac{1}{\sigma_t(1)}\right) - \sigma_t(\mathcal{L}^1\phi)\frac{\Pi_t(h)}{\sigma_t(1)}dt \quad (3.9)$$

Il s'agit alors de calculer $d(\sigma_t^{-1}(1))$ en appliquant la formule de Itô :

$$\begin{aligned} d\left(\frac{1}{\sigma_t(1)}\right) &= -\frac{1}{\sigma_t(1)^2}\sigma_t(1)\Pi_t(h)dY_t + \frac{1}{\sigma_t(1)^3}\sigma_t(1)^2\Pi_t(h)^2dt \\ &= -\frac{\Pi_t(h)}{\sigma_t(1)}dY_t + \frac{\Pi_t(h)^2}{\sigma_t(1)}dt \end{aligned}$$

En reportant cette dernière dans l'expression (3.9), on obtient

$$\begin{aligned} d\left(\frac{\sigma_t(\phi)}{\sigma_t(1)}\right) &= \frac{\sigma_t(\mathcal{L}\phi)dt + \sigma_t(\mathcal{L}^1\phi)dY_t}{\sigma_t(1)} + \sigma_t(\phi)\left(-\frac{\Pi_t(h)}{\sigma_t(1)}dY_t + \frac{\Pi_t(h)^2}{\sigma_t(1)}dt\right) \\ &\quad - \Pi_t(\mathcal{L}^1\phi)\Pi_t(h)dt \\ &= \Pi_t(\mathcal{L}\phi)dt + \{\Pi_t(\mathcal{L}^1\phi) - \Pi_t(\phi)\Pi_t(h)(dY_t - \Pi_t(h)dt)\} \end{aligned}$$

d'où l'on déduit

$$\Pi_t(\phi) = \Pi_0(\phi) + \int_0^t \Pi_s(\mathcal{L}\phi)ds + \int_0^t \{\Pi_s(\mathcal{L}^1\phi) - \Pi_s(h)\Pi_s(\phi)\}d\nu_s$$

■

Remarque 7 L'équation de Kushner est non linéaire en Π_t alors que l'équation de Zakai est linéaire en σ_t et plus simple à résoudre. D'ailleurs, on montre qu'une solution de l'équation de Zakai induit une solution de Kushner en utilisant la formule de Striebel- Kallianpur et réciproquement. De façon précise on a :

Corollaire 5 Si Π_t est solution de l'équation de Kushner, alors $\sigma_t(\phi) = \Pi_t(\phi)\alpha_t$, où

$$\alpha_t = \exp\left(\int_0^t \Pi_s(h(s))dY_s - \frac{1}{2} \int_0^t (\Pi_s(h))^2 ds\right)$$

est solution de l'équation de Zakai.

Réciproquement, si σ_t est solution de l'équation de Zakai, alors $\Pi_t(\phi) = \frac{\sigma_t(\phi)}{\sigma_t(1)}$ est solution de l'équation de Kushner.

Démonstration

1. Soit Π_t la solution de l'équation de Kushner, alors

$$d\Pi_t(\phi) = \Pi_t(\mathcal{L}\phi)dt + [\Pi_t(\mathcal{L}^1\phi) - \Pi_t(h)\Pi_t(\phi)](dY_t - \Pi_t(h)dt)$$

il est facile de vérifier que $d\alpha_t = \alpha_t\Pi_t(h)dY_t$ et par application de la formule de différenciation de Itô au produit $\Pi_t(\phi)\alpha_t$, on trouve que

$$\begin{aligned} d(\Pi_t(\phi)\alpha_t) &= \Pi_t(\phi)d\alpha_t + \alpha_t d\Pi_t(\phi) + \alpha_t\Pi_t(h)[\Pi_t(\mathcal{L}^1\phi) - \Pi_t(h)\Pi_t(\phi)]dt \\ &= \alpha_t\Pi_t(\phi)\Pi_t(h)dY_t + \alpha_t\Pi_t(\mathcal{L}\phi)dt + \alpha_t\Pi_t(\mathcal{L}^1\phi)dY_t \\ &\quad - \alpha_t\Pi_t(h)\Pi_t(\phi)dY_t - \alpha_t\Pi_t(\mathcal{L}^1\phi)\Pi_t(h)dt + \alpha_t\Pi_t(\phi)(\Pi_t(h))^2dt \\ &\quad + \alpha_t\Pi_t(\mathcal{L}^1\phi)\Pi_t(h)dt - \alpha_t\Pi_t(\phi)(\Pi_t(h))^2dt \end{aligned}$$

Soit

$$d(\Pi_t(\phi)\alpha_t) = \alpha_t\Pi_t(\mathcal{L}\phi)dt + \alpha_t\Pi_t(\mathcal{L}^1\phi)dY_t$$

D'où le résultat.

2. La réciproque, évidente d'après le théorème (16).

■

Remarque 8 Le problème de l'existence et de l'unicité des solutions des équations de Zakai et de Kushner-Stratonovich dans le cas où les mouvements browniens (V_t) et (W_t) sont indépendants a été étudié par plusieurs auteurs (voir [38] et [33]) et dans ([1] chap. 04) on trouve la démonstration dans le cas où ces browniens sont dépendants .

La densité conditionnelle normalisée et non normalisée

Considérons le système d'état

$$\begin{cases} dX_t &= b(X_t)dt + f(X_t)dV_t \\ dY_t &= h(X_t)dt + dW_t \end{cases}$$

avec les conditions initiales $Y_0 = 0$ et X_0 une variable aléatoire indépendante des mouvements browniens (W_t) et (V_t) admettant une densité de probabilité $\varphi_0 \in L^2(\mathbb{R})$.

Théorème 17 (Kushner) *Si le filtre normalisé $\Pi_t(\phi)$ admet une densité $\varphi_t(\cdot)$, alors $\varphi_t(\cdot)$ est solution de l'E.D.P.S suivante :*

$$d\varphi_t(x) = \mathcal{L}^* \varphi_t(x)dt + \{\mathcal{L}^1 \varphi_t(x) - \varphi_t(x)\Pi_t(h)\}d\nu_t \quad (3.10)$$

où \mathcal{L}^* est l'opérateur adjoint donné pour toute fonction $\psi \in C_b^2$ par

$$\mathcal{L}^* \psi(x) = \frac{1}{2} \frac{d^2}{dx^2} \{f^2(x)\psi(x)\} - \frac{d}{dx} \{b(x)\psi(x)\}$$

Démonstration

En utilisant une propriété des opérateurs adjoints, on a d'une part

$$\int_0^t \Pi_s(\mathcal{L}\phi)ds = \int_0^t \left(\int_{\mathbb{R}} \mathcal{L}\phi(x)\varphi_s(x)dx \right) ds = \int_0^t \left(\int_{\mathbb{R}} \phi(x)\mathcal{L}^*\varphi_s(x)dx \right) ds$$

et d'autre part

$$\int_0^t \Pi_s(\mathcal{L}^1\phi)d\nu_s = \int_0^t \left(\int_{\mathbb{R}} \mathcal{L}^1\phi(x)\varphi_s(x)dx \right) d\nu_s = \int_0^t \left(\int_{\mathbb{R}} \phi(x)\mathcal{L}^1\varphi_s(x)dx \right) d\nu_s$$

et de plus on a

$$\int_0^t \Pi_s(h)\Pi_s(\phi)d\nu_s = \int_0^t \left(\int_{\mathbb{R}} \phi(x)\varphi_s(x)\Pi_s(h)dx \right) d\nu_s$$

Appliquons le théorème de Fubini aux trois termes précédents et reportons ensuite chacun de ces termes dans l'équation de Kushner-Stratonovich, on obtient

$$\begin{aligned}
\Pi_t(\phi) &= \int_{\mathbb{R}} \phi(x) \varphi_0(x) dx + \int_{\mathbb{R}} \phi(x) \left(\int_0^t \mathcal{L}^* \varphi_s(x) ds \right. \\
&\quad \left. + \int_0^t \mathcal{L}^1 \varphi_s(x) d\nu_s + \int_0^t \varphi_s(x) \Pi_s(h) d\nu_s \right) dx \\
&= \int_{\mathbb{R}} \phi(x) \varphi_0(x) dx + \int_{\mathbb{R}} \phi(x) \left\{ \int_0^t \mathcal{L}^* \varphi_s(x) ds + \int_0^t (\mathcal{L}^1) \varphi_s(x) \right. \\
&\quad \left. - \varphi_s(x) \Pi_s(h) \right\} d\nu_s
\end{aligned}$$

d'où

$$\varphi_t(x) = \varphi_0(x) + \int_0^t \mathcal{L}^* \varphi_s(x) ds + \int_0^t (\mathcal{L}^1) \varphi_s(x) - \varphi_s(x) \Pi_s(h) d\nu_s$$

ce qui achève la démonstration. ■

Le théorème suivant présente la densité conditionnelle de la distribution non normalisée $\sigma_t(\phi)$ et sa relation avec la densité de la distribution normalisée $\Pi_t(\phi)$.

Théorème 18 *Si la distribution conditionnelle non-normalisée σ_t admet une densité $q_t(\cdot)$, alors $q_t(x)$ satisfait l'équation aux dérivées partielles stochastique suivante :*

$$dq_t(x) = \mathcal{L}^* q_t(x) dt + \mathcal{L}^1 q_t(x) dY_t, q_0 \in L^2(\mathbb{R})$$

De plus

$$\varphi_t(x) = \frac{q_t(x)}{\int_{\mathbb{R}} q_t(y) dy}$$

où $\varphi_t(\cdot)$ est la densité de la distribution Π_t .

Démonstration

Réécrivons l'équation linéaire de Zakai en fonction de la densité $q_t(\cdot)$

$$\sigma_t(\phi) = \int_{\mathbb{R}} \phi(x) q_0(x) dx + \int_0^t \left(\int_{\mathbb{R}} \mathcal{L} \phi(x) q_s(x) dx \right) ds + \int_0^t \left(\int_{\mathbb{R}} \mathcal{L}^1 \phi(x) q_s(x) dx \right) dY_s$$

et en utilisant le théorème de Fubini et la propriété des opérateurs adjoints, cette équation devient :

$$\begin{aligned}
\sigma_t(\phi) &= \int_R \phi(x)q_0(x)dx + \int_0^t \left(\int_{\mathbb{R}} \phi(x)\mathcal{L}^*q_s(x)dx \right) ds + \int_0^t \left(\int_{\mathbb{R}} \phi(x)\mathcal{L}^1q_s(x)dx \right) dY_s \\
&= \int_R \phi(x)q_0(x)dx + \int_{\mathbb{R}} \phi(x) \left(\int_0^t \mathcal{L}^*q_s(x)ds \right) dx + \int_{\mathbb{R}} \phi(x) \left(\int_0^t \mathcal{L}^1q_s(x)dY_s \right) dx \\
&= \int_{\mathbb{R}} \phi(x) \left(q_0(x) + \int_0^t \mathcal{L}^*q_s(x)ds + \int_0^t \mathcal{L}^1q_s(x)dY_s \right) dx
\end{aligned}$$

d'où l'on déduit que :

$$q_t(x) = q_0(x) + \int_0^t \mathcal{L}^*q_s(x)ds + \int_0^t \mathcal{L}^1q_s(x)dY_s$$

Maintenant en utilisant la formule de Kallianpur -Striebel et les relations

$$\sigma_t(\phi) = \int_{\mathbb{R}} \phi(x)q_t(x)dx \text{ et } \Pi_t(x) = \int_R \phi(x)\varphi_t(x)dx$$

on déduit que

$$\varphi_t(x) = \frac{q_t(x)}{\int_{\mathbb{R}} q_t(y)dy}.$$

■

3.2.3 Exemple

On montre dans cet exemple que l'approche innovation appliquée aux modèles linéaires permet de déduire le filtre de Kalman-Bucy.

A cette fin, considérons le modèle d'espace d'état linéaire gaussien suivant :

$$\begin{cases} dX_t &= bX_tdt + f dV_t \\ dY_t &= hX_tdt + dW_t \end{cases}$$

où b , f et h sont des constantes.

Posons pour tout $k \in \mathbb{N}$, $\phi(x) = x^k$, alors

$$\Pi_t(\phi) = \Pi_t(x^k)$$

Rappelons que la distribution conditionnelle de X_t sachant \mathcal{F}_t^Y est normale de moyenne

$$\widehat{X}_t = E(X_t / \mathcal{F}_t^Y) = \Pi_t(x)$$

et de variance

$$\widehat{\sigma}_t^2 = E((X_t - \widehat{X}_t)^2 / \mathcal{F}_t^Y) = \Pi_t(x^2) - \Pi_t^2(x)$$

Pour $k = 1$ (c'est-à-dire $\phi(x) = x$), on a d'après l'équation de Kushner-Stratonovich :

$$\Pi_t(x) = \Pi_0(x) + b \int_0^t \Pi_s(x) ds + h \int_0^t (\Pi_s(x^2) - \Pi_s^2(x))(dY_s - h\Pi_s(x) ds) \quad (3.11)$$

et pour $k = 2$, on a

$$\Pi_t(x^2) = \Pi_0(x^2) + \int_0^t (f^2 + 2b\Pi_s(x^2)) ds + \int_0^t (h\Pi_s(x^3) - h\Pi_s(x)\Pi_s(x^2))(dY_s - h\Pi_s(x) ds) \quad (3.12)$$

Soit X une variable aléatoire telle que $X \sim \mathcal{N}(\mu, \sigma^2)$, rappelons que le moment d'ordre trois² de X est donné par :

$$E(X^3) = \mu^3 + 3\mu\sigma^2.$$

Par analogie, on montre aussi que

$$\begin{aligned} \Pi_t(x^3) &= \Pi_t^3(x) + 3\Pi_t(x)\widehat{\sigma}_t^2 \\ &= \Pi_t^3(x) + 3\Pi_t(x)\Pi_t(x^2) - 3\Pi_t^3(x) \end{aligned} \quad (3.13)$$

où $\Pi_t(x^3) = E(X_t^3 / \mathcal{F}_t^Y)$

En utilisant les équations (3.11) et (3.12), on obtient :

$$\begin{aligned} d\widehat{\sigma}_t^2 &= d(\Pi_t(x^2) - \Pi_t^2(x)) = d\Pi_t(x^2) - 2\Pi_t(x)d\Pi_t(x) - [h(\Pi_t(x^2) - \Pi_t^2(x))]^2 dt \\ &= \{f^2 + 2b\Pi_t(x^2)\} dt + \{h\Pi_t(x^3) - h\Pi_t(x)\Pi_t(x^2)\} (dY_t - h\Pi_t(x) dt) \\ &\quad - 2\Pi_t(x)\{b\Pi_t(x) dt + h(\Pi_t(x^2) - \Pi_t^2(x))(dY_t - h\Pi_t(x) dt)\} - [h(\Pi_t(x^2) - \Pi_t^2(x))]^2 dt \end{aligned}$$

2. le moment d'ordre n est donné par : $E(X^n) = (n-1)\sigma^2 E(X^{n-2}) + \mu E(X^{n-1})$.

et tenant compte de l'équation (3.13) on trouve que

$$\begin{aligned} d\hat{\sigma}_t^2 &= \{f^2 + 2b\Pi_t(x^2)\}dt - 2b\Pi_t^2(x)dt - h^2(\Pi_t(x^2) - \Pi_t^2(x))^2dt \\ &= \{f^2 + 2b(\Pi_t(x^2) - \Pi_t^2(x)) - h^2(\Pi_t(x^2) - \Pi_t^2(x))^2\}dt. \end{aligned}$$

D'où l'on retrouve l'équation déterministe de Riccati :

$$\frac{d\hat{\sigma}_t^2}{dt} = f^2 + 2b\hat{\sigma}_t^2 - h^2(\hat{\sigma}_t^2)^2$$

Alors on peut écrire l'équation (3.11) sous la forme

$$\begin{aligned} d\hat{X}_t &= b\hat{X}_t dt + h\hat{\sigma}_t^2(dY_t - h\hat{X}_t dt) \\ &= (b - h^2\hat{\sigma}_t^2)\hat{X}_t dt + h\hat{\sigma}_t^2 dY_t \end{aligned}$$

et nous reconnaissons à travers cette dernière relation, l'équation du filtre de Kalman-Bucy dont la solution est donnée par

$$\hat{X}_t = \hat{X}_0 \exp(b - h^2\hat{\sigma}_t^2)t + \int_0^t \exp(b - h^2\hat{\sigma}_s^2)(t - s)h\hat{\sigma}_s^2 dY_s.$$

Deuxième partie

Applications

Introduction

La recherche en filtrage est motivée par de nombreuses applications dans des domaines variées tels le traitement de signal, finance, météorologie, océanographie et navigation ...etc. Ce problème est en effet assez général puisqu'il s'agit d'estimer l'état du système dynamique à partir d'observations bruitées.

Nous décrivons dans cette partie trois applications du filtrage. La première porte sur l'application du filtrage linéaire à temps continu en traitement de signal, la deuxième sur l'application et la simulation du filtrage linéaire à temps discret à un problème de navigation et la troisième est une application du filtrage non linéaire à temps discret à un problème de finance.

Chapitre 4

Application du filtrage linéaire à temps continu au traitement de signal

La théorie du filtrage joue un rôle très important en traitement de signal. La modélisation des circuits électriques a été largement étudiée, notamment par W.Kampowsky [37] (Classification et la simulation numérique des circuit électriques avec un bruit blanc) C.Penski [5] (Méthode numérique pour les E.D.S avec un bruit blanc) et R.Rezayen & R.Farnoosh [32] ont appliqué le filtre de Kalman-Bucy au problème du circuit RC.

Dans ce chapitre, on présente une application directe du filtre de Kalman-Bucy au problème du circuit électrique RL (E.Kolarova [8]). On commence d'abord par quelques définitions concernant le filtrage dans le domaine de traitement de signal .

Définition d'un filtre en traitement de signal

Très utilisé en électronique, un filtre est un circuit électronique qui réalise une opération de traitement de signal. Il atténue certaines composantes d'un signal et

en laisse passer d'autres. Les filtres servent à éliminer ou favoriser une bande de fréquences.

Définition d'un circuit RL (le cas déterministe)

Considérons le système électrique de la figure (4.1), ce système représente un circuit électrique résistance-inductance (RL), constitué d'une résistance R (en Ohms) et d'une inductance L (en Henrys) et d'une force électromotrice U (en Volts). L'équation donnant la quantité I du courant électrique I (Ampères) dans un circuit R.L est :

$$L \frac{dI}{dt} + RI = U$$

où

I : l'intensité du courant électrique.

L : l'inductance.

R : La résistance totale du circuit.

U : La tension aux bornes du montage.

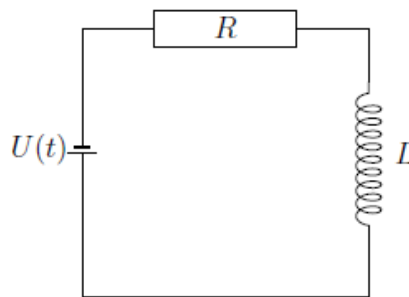


FIGURE 4.1 – Le circuit RL

4.0.4 Position du Problème

La modélisation de ce genre de système (RL) par des équations différentielles ordinaires ignore les effets stochastiques. En ajoutant des éléments aléatoires dans ces

équations différentielles, on obtient une équation différentielle stochastique E.D.S. L'équation du problème du circuit RL est décrite par la formule suivante :

$$LI'(t) + RI(t) = \sigma\xi(t)$$

ξ est une force électromagnétique qui fluctue rapidement, généré par le bruit thermique assimilé à un bruit blanc et σ une constante. Le problème électrique peut être formaliser par l' E.D.S suivante :

$$dI(t) = -\frac{R}{L}I(t)dt + \frac{\sigma}{L}dV(t), \quad I(0) = I_0. \quad (4.1)$$

où $V(t)$ est un processus de Wiener.

$\frac{R}{L}$:le coefficient de friction.

$\frac{\sigma}{L}$:le coefficient de diffusion.

I_0 : le courant initial (c'est une variable aléatoire telle que $E(I_0) < \infty, var(I_0) < \infty$).

4.0.5 Application du filtre de Kalman-Bucy

On fournit quelques mesures du courant continu jusqu'à l'instant $s \leq t$, on obtient l'équation d'observation suivante :

$$dY(t) = I(t)dt + dW(t). \quad (4.2)$$

où $W(t)$ est un mouvement brownien indépendant de $V(t)$.

Revenant maintenant au problème du filtrage linéaire à temps continu décrit par le modèle (4.1), (4.2). En appliquant le filtre de kalman-Bucy (l'équation (2.14) chap 2,§ 2), on trouve :

$$d\hat{I}(t) = \left(-\frac{R}{L} - P(t)\right)\hat{I}(t)dt + P(t)dY(t). \quad \hat{I}(0) = E(I_0) = 0. \quad (4.3)$$

où $P(t) = E[I(t) - \hat{I}(t)]^2$ satisfait l'équation déterministe de Riccati (2.15) :

$$P'(t) = -\frac{2R}{L}P(t) - P^2(t) + \frac{\sigma^2}{L^2}, \quad P(0) = E(I_0^2) = A^2. \quad (4.4)$$

Pour résoudre l'équation de Riccati, on pose :

$$P(t) = M(t) - \frac{R}{L}; P'(t) = M'(t); P^2(t) = M^2(t) - 2\frac{R}{L}M(t) + \frac{R^2}{L^2}$$

on obtient une équation différentielle ordinaire de $M(t)$ à coefficients séparables, que l'on écrit

$$M'(t) = \frac{R^2 + \sigma^2}{L^2} - M^2(t)$$

alors

$$\frac{dt}{dM(t)} = \frac{1}{\gamma^2 - M^2}$$

où $\gamma = \frac{\sqrt{R^2 + \sigma^2}}{L}$. Après quelques manipulations ([8]), on montre que cette équation admet la solution suivante :

$$\ln \left| \frac{M(t) - \gamma}{M(t) + \gamma} \right| = C - 2\gamma t.$$

$$\text{où } C = \ln \left| \frac{A^2 - \frac{R}{L} - \gamma}{A^2 - \frac{R}{L} + \gamma} \right|.$$

Concernant la valeur de $P(t)$, on a :

$$\frac{M(t) - \gamma}{M(t) + \gamma} = \exp^{(C-2\gamma t)}$$

ou encore

$$M(t) = \gamma \left(\frac{1 + \exp^{(C-2\gamma t)}}{1 - \exp^{(C-2\gamma t)}} \right)$$

Par suite, en remplaçant dans l'expression de $P(t)$ donnée ci-dessus, on trouve que

$$P(t) = \gamma \left(\frac{1 + \exp^{C-2\gamma t}}{1 - \exp^{C-2\gamma t}} \right) - \frac{R}{L}$$

On substitue cette expression de $P(t)$ dans (4.3), on déduit l'E.D.S du filtrage :

$$d\hat{I}(t) = \gamma \left(\frac{1 + \exp^{C-2\gamma t}}{1 - \exp^{C-2\gamma t}} \right) \hat{I}(t)dt + \gamma \left(\frac{1 + \exp^{C-2\gamma t}}{1 - \exp^{C-2\gamma t}} - \frac{R}{L} \right) dY(t)$$

pour des grands valeurs de t on trouve $M(t) \simeq \gamma$ et $P(t) = \gamma - \frac{R}{L}$ ce qui implique

$$d\hat{I}(t) = -\frac{\sqrt{R^2 + \sigma^2}}{L} \hat{I}(t)dt + \frac{\sqrt{R^2 + \sigma^2} - R}{L} dY(t)$$

on peut résoudre cette équation par application directe de la formule d'Itô pour une fonction : $g(t, x) = x \exp^{\gamma t} = x \exp^{\frac{\sqrt{R^2 + \sigma^2}}{L} t}$

$$d \left(\exp^{\frac{\sqrt{R^2 + \sigma^2}}{L} t} \hat{I}(t) \right) = \left[\frac{\sqrt{R^2 + \sigma^2}}{L} \exp^{\frac{\sqrt{R^2 + \sigma^2}}{L} t} \hat{I}(t) dt + \exp^{\frac{\sqrt{R^2 + \sigma^2}}{L} t} d\hat{I}(t) \right]$$

ainsi

$$\exp^{\gamma t} \hat{I}(t) = \int_0^t \left(\gamma - \frac{R}{L} \right) \exp^{\gamma s} dY(s)$$

donc

$$\hat{I}(t) = \left(\frac{\sqrt{R^2 + \sigma^2} - R}{L} \right) \int_0^t \exp^{\frac{\sqrt{R^2 + \sigma^2}}{L} (s-t)} dY(s)$$

c'est la solution du problème du filtrage du circuit électrique RL.

Chapitre 5

Application du filtrage linéaire à temps discret à un problème de navigation et simulation

Le filtre de Kalman a été proposé par Rudolf Kalman en 1960 pour résoudre un problème de poursuite de trajectoire dans la préparation des missions Apollo. Depuis cette année et en raison du développement du calcul numérique le filtre de Kalman a fait l'objet de plusieurs recherches dans l'ingénierie et en particulier dans le domaine de navigation pour déterminer la position d'un mobile ainsi que quelques autres informations concernant le déplacement de ce mobile.

L'objectif de ce chapitre est d'appliquer le filtre de Kalman à un simple exemple de navigation pour estimer la position d'un mobile. Moyennant le logiciel MATLAB, on effectue une simulation de la position de ce mobile à partir des mesures et on compare la valeur mesurée avec la vraie valeur et la valeur estimée.

5.1 Description du problème

On veut estimer la position d'un mobile qui se déplace selon un axe Ox . Pour cela à chaque instant k , $k = 1, 2, \dots, N$ nous disposons uniquement d'un capteur de position fournissant une mesure bruitée de celle-ci . Le mouvement de ce mobile est caractérisé à l'instant k par sa position P_k et sa vitesse V_k , aussi on utilise comme vecteur d'état $X(k) = [P_k \quad V_k]^t$ où t est la transposé.

Pour appliquer le filtre de Kalman, il faut d'abord et avant tout modéliser le problème en fonction des variables à estimer et des mesures des capteurs.

Le modèle d'espace d'état associé au problème présenté ci-dessus, est donné par les équations suivantes :

$$\begin{cases} X_{k+1} &= F X_k + Z_k \\ Y_k &= G X_k + W_k, \end{cases} \quad (5.1)$$

où X_k , Y_k sont respectivement l'état et l'observation à l'instant k , F et G sont des matrices, Z_k est le bruit d'état et W_k est le bruit de mesure (il vient de l'instrument de mesure) .

Supposons que Z_k et W_k sont deux bruits gaussiens indépendants tels que

$$Z_k \sim \mathcal{N}(0, Q_k) \quad \text{et} \quad W_k \sim \mathcal{N}(0, R_k).$$

Pour une période d'échantillonnage T , on peut écrire le modèle 5.1 sous la forme suivante :

$$\begin{cases} X_{k+1} &= \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} X_k + Z_k \\ Y_k &= \begin{bmatrix} 1 & 0 \end{bmatrix} X_k + W_k, \end{cases} \quad (5.2)$$

Le vecteur X contient toute l'information concernant l'état présent du système, mais on peut pas le mesurer directement. En effet, on mesure Y qui est en fonction de X plus un bruit et on utilise cette mesure pour estimer l'état en appliquant le filtre de Kalman.

5.2 Application du filtre de Kalman

Rappelons que l'algorithme de Kalman consiste à deux étapes

Étape prédiction

$$\begin{cases} X_{k+1|k} = FX_{k|k} \\ P_{k+1|k} = FP_{k|k}F^T + Q_k \end{cases}$$

Étape correction

$$\begin{cases} X_{k|k} = X_{k|k-1} + K_k(Y_k - GX_{k|k-1}) \\ P_{k|k} = (I - K_kG)P_{k|k-1} \end{cases}$$

où K_k est le gain de Kalman donné par

$$K_k = P_{k|k-1}G^T(R_k + GP_{k|k-1}G^T)^{-1}$$

L'objectif est d'appliquer cet algorithme au problème présenté ci-dessus.

5.3 Simulation avec Matlab

L'algorithme de simulation qui a été décrit dans la section précédente, a été implanté par le logiciel Matlab .

Le travail réalisé consiste à simuler le fonctionnement du filtre de Kalman en vu de vérifier quelques performances, mais avant de simuler l'algorithme, il est important de bien initialiser le filtre, on pose :

$$X_{0|0} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ et } P_{0|0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Pour $Q_k = \begin{bmatrix} 0.0005 \\ 0.1 \end{bmatrix} \begin{bmatrix} 0.0005 \\ 0.1 \end{bmatrix}^T$ et $N = 60$, on a simulé l'algorithme de Kalman pour trois valeurs différentes de l'erreur de mesure $R_k = 1, R_k = 0.1, R_k = 0.01$.On a obtenu les résultats suivants

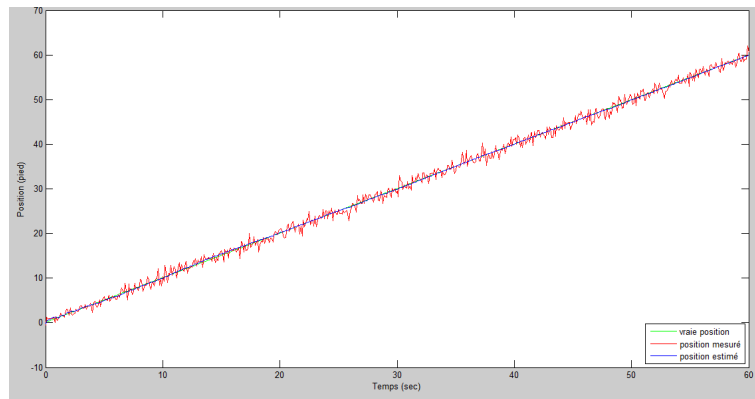


FIGURE 5.1 – Estimation de la position du mobile pour $W_k \sim \mathcal{N}(0, 1)$

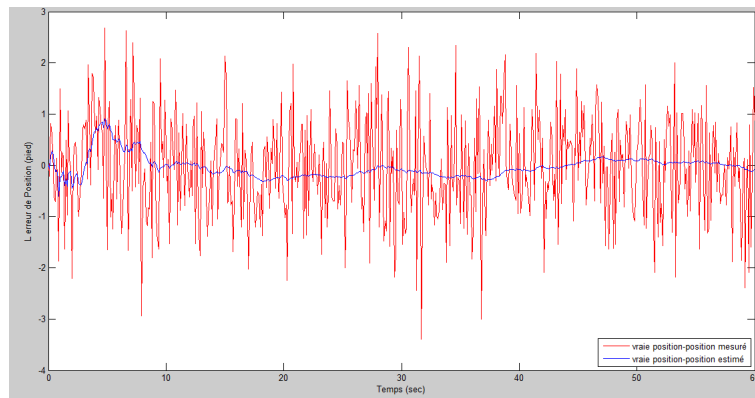


FIGURE 5.2 – Erreur de mesure et erreur d'estimation pour $W_k \sim \mathcal{N}(0, 1)$

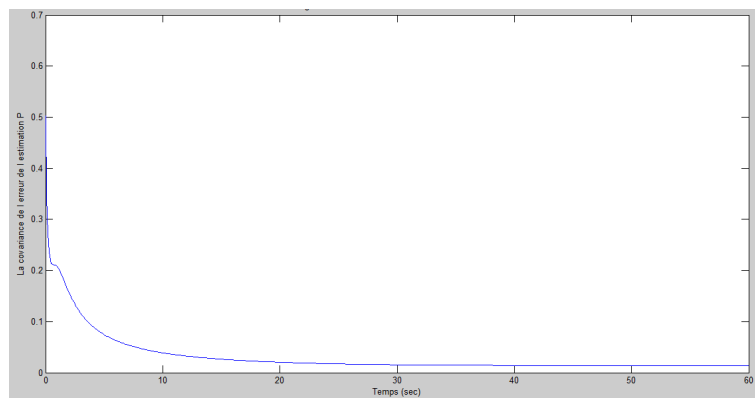


FIGURE 5.3 – La variance de l'erreur d'estimation pour $W_k \sim \mathcal{N}(0, 1)$

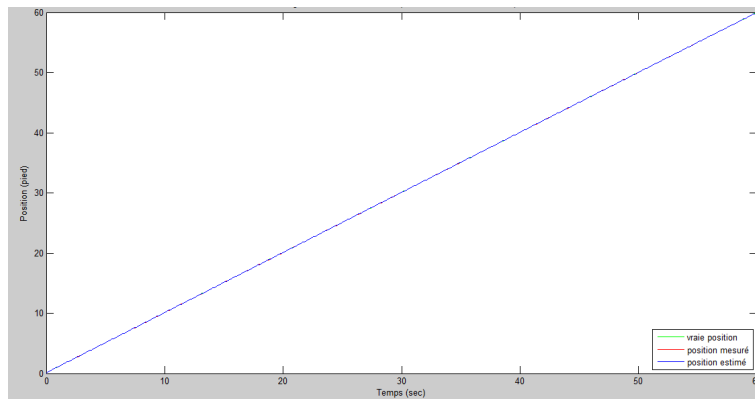


FIGURE 5.4 – Estimation de la position du mobile pour $W_k \sim \mathcal{N}(0, 0.1)$

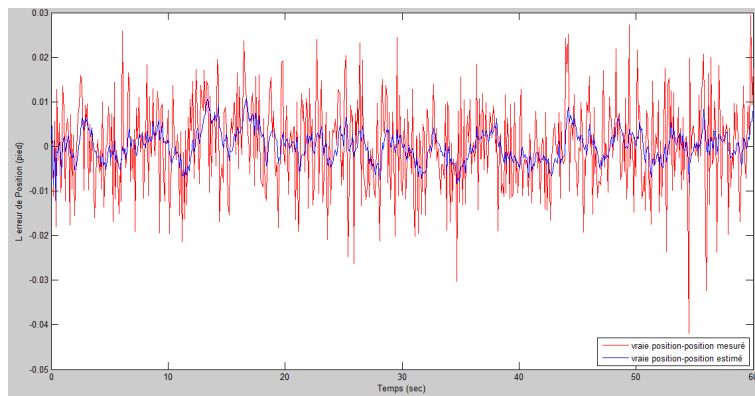


FIGURE 5.5 – Erreur de mesure et erreur d'estimation pour $W_k \sim \mathcal{N}(0, 0.1)$

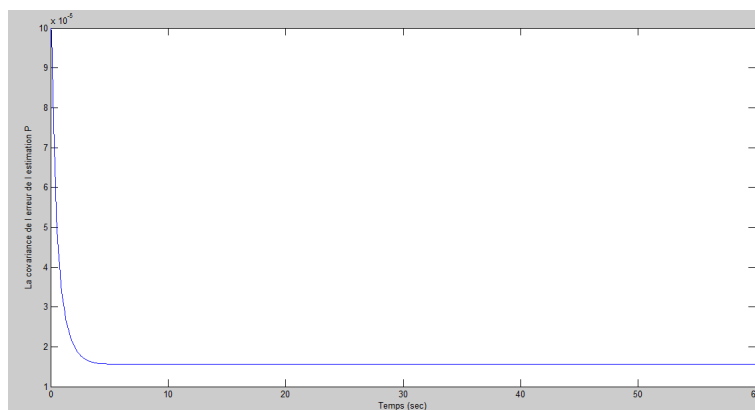


FIGURE 5.6 – La variance de l'erreur d'estimation pour $W_k \sim \mathcal{N}(0, 0.1)$

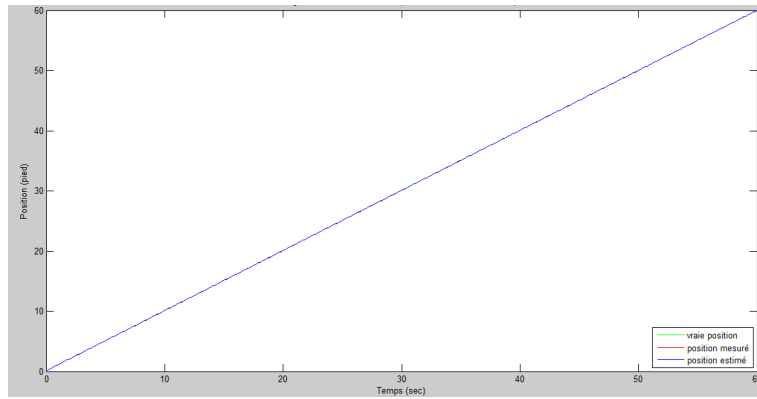


FIGURE 5.7 – Estimation de la position du mobile pour $W_k \sim \mathcal{N}(0, 0.01)$

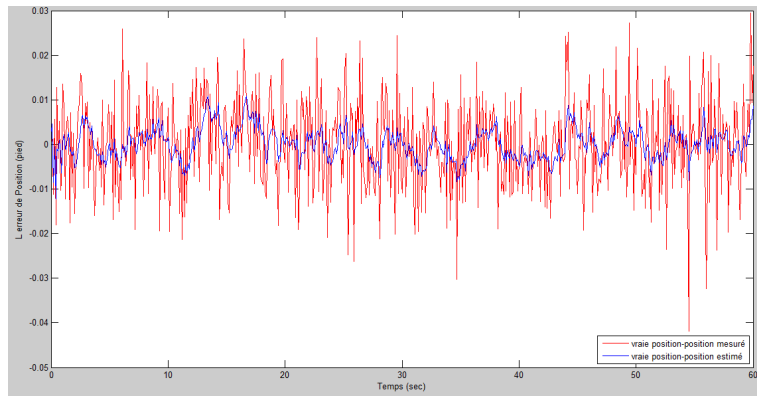


FIGURE 5.8 – Erreur de mesure et erreur d'estimation pour $W_k \sim \mathcal{N}(0, 0.01)$

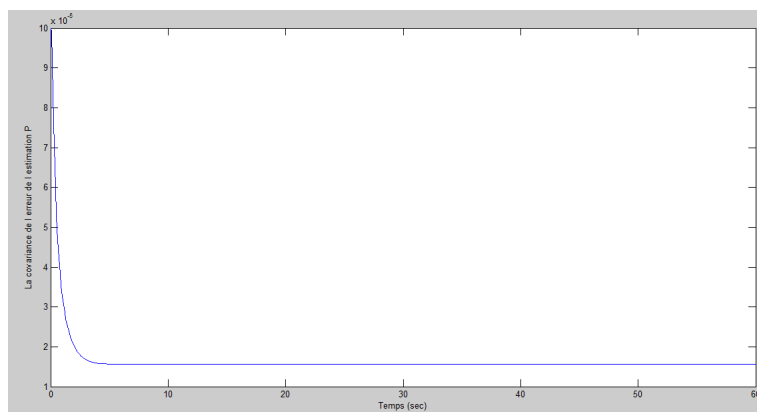


FIGURE 5.9 – La variance de l'erreur d'estimation pour $W_k \sim \mathcal{N}(0, 0.01)$

On remarque que l'évolution de la position du mobile est corrigée à chaque itération, alors que les observations disponibles sont bruitées au fur et à mesure du déplacement du mobile.

Un des résultats du filtre de Kalman énoncés dans la chapitre (02) c'est que le gain de Kalman K_k diminue si l'estimation par le modèle devient plus précise et il augmente si les observations deviennent plus précises et c'est ce qu'on a trouvé évidemment dans la simulation, en effet :

$R_k = 1$	$R_k = 0.1$	$R_k = 0.01$
$K_k = \begin{pmatrix} 0.0437 \\ 0.0098 \end{pmatrix}$	$K_k = \begin{pmatrix} 0.1319 \\ 0.0932 \end{pmatrix}$	$K_k = \begin{pmatrix} 0.3600 \\ 0.8000 \end{pmatrix}$

En remarque aussi d'après les figures (5.9), (5.6) et (5.3) que la variance de l'erreur d'estimation s'approche de zéro à chaque fois que l'estimation devient plus précise.

Chapitre 6

L'application du filtrage à la finance

Une dernière application du filtrage et notamment celle du filtre de Kalman porte sur le domaine de la finance (voir [18], [35] et [29]).

Dans leur fameuse équation du prix d'une option, Black et Scholes (1973) ont proposé que la volatilité σ^2 du prix S_t d'un actif sous-jacent est constante, tel que :

$$dS_t = S_t(\nu dt + \sigma dW_t)$$

où W_t est un mouvement brownien standard, ν est une constante et σ représente l'écart-type de $\frac{dS_t}{S_t}$.

Black et Scholes ont fait appel au concept de volatilité inconditionnelle pour définir leur équation, et dans ce cas elle est facile à estimer et à interpréter tant qu'elle reste constante. Malheureusement, la déviation standard des rendements d'un actif financier n'est pas constante; le niveau de la volatilité d'aujourd'hui peut être différent de celui d'hier, ce qui nous pousse à penser à une manière d'estimer cette volatilité instantanée (à l'instant courant) qui n'est pas observable, appelée *volatilité stochastique*.

6.1 Position du problème

On suppose que y_t est décrit par le processus stochastique suivant

$$y_t = \sigma_t u_t \quad \text{où} \quad u_t \sim \mathcal{N}(0, 1) \quad (6.1)$$

où σ_t^2 est la volatilité de y_t .

Les modèles de la volatilité stochastique (SV) spécifient le logarithme de la volatilité comme étant un processus linéaire stochastique. Une version plus simple est décrite par le modèle AR(1)

$$\log(\sigma_t^2) = \alpha + \beta \log(\sigma_{t-1}^2) + w_t \quad (6.2)$$

où $\alpha > 0$, $0 \leq \beta < 1$, $w_t \sim \mathcal{N}(0, \sigma_w^2)$ et u_t et w_t sont indépendants pour tout s et t . Soit θ le vecteur des paramètres $(\alpha, \beta, \sigma_w)$. L'utilisation pratique des modèles à volatilité stochastique nécessite l'estimation de ces paramètres en utilisant la fonction de vraisemblance. Mais cette estimation est délicate car la volatilité n'est pas une variable d'état observable.

La méthode du Quasi-Maximum de vraisemblance (QML) qui a été développée par Nelson (1988), Harvey (1994) et Ruiz (1994) permet de contourner ce problème. Elle transforme d'abord le modèle à volatilité stochastique en forme d'espace d'état linéaire

$$\log y_t^2 = \log \sigma_t^2 + \log u_t^2$$

On définit $Y_t = \log y_t^2$; $x_t = \log \sigma_t^2$ et $v_t = \log u_t^2$ ce qui conduit au système d'équation suivant :

$$\begin{cases} x_t = \alpha + \beta x_{t-1} + w_t \\ Y_t = x_t + v_t \end{cases} \quad (6.3)$$

on remarque que v_t n'est pas gaussienne. L'application du filtre de Kalman à ce modèle conduit à un estimateur non optimal des paramètres, du fait que la vraisemblance est une fonction de v_t laquelle est non gaussienne. Cette méthode ne se base pas sur la vraie vraisemblance, et c'est pour cela qu'on l'appelle l'estimation du Quasi-Maximum de vraisemblance.

6.2 Application du filtrage non linéaire au modèle à volatilité stochastique

Pour pouvoir traiter ce type de problème on applique le filtre non linéaire au modèle (6.1), (6.2). On adopte l'approche de Watanabe (voir[35]) qui a proposé un filtre non linéaire pour déterminer vraisemblance exacte des observations.

On note $x_t = \log \sigma_t^2$ et $\underline{y}_t = (y_1, y_1, \dots, y_t)$. Pour alléger l'écriture, notons $\varphi(./.)$ la densité conditionnelle, alors d'après les équations (6.2) et (6.1), on a :

$$\begin{cases} \varphi(x_t/x_{t-1}, \underline{y}_{t-1}) &= \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left[-\frac{(x_t - \alpha - \beta x_{t-1})^2}{2\sigma_w^2}\right] \\ \varphi(y_t/x_t, \underline{y}_{t-1}) &= \frac{1}{\sqrt{2\pi \exp(x_t)}} \exp\left[-\frac{y_t^2}{2\exp(x_t)}\right] \end{cases} \quad (6.4)$$

Les équations de ce filtre non linéaire à temps discret sont données par les formules suivantes :

Prédiction :

$$\varphi(x_t/\underline{y}_{t-1}) = \int_{-\infty}^{+\infty} \varphi(x_t/x_{t-1}, \underline{y}_{t-1}) \varphi(x_{t-1}/\underline{y}_{t-1}) dx_{t-1} \quad (6.5)$$

Correction où bien la mise-à-jour :

$$\varphi(x_t/\underline{y}_t) = \frac{\varphi(y_t/x_t, \underline{y}_{t-1}) \varphi(x_t/\underline{y}_{t-1})}{\varphi(y_t/\underline{y}_{t-1})} \quad (6.6)$$

où

$$\varphi(y_t/\underline{y}_{t-1}) = \int_{-\infty}^{+\infty} \varphi(y_t/x_t, \underline{y}_{t-1}) \varphi(x_t/\underline{y}_{t-1}) dx_t \quad (6.7)$$

Rappelons que ce filtre est difficile à résoudre analytiquement. Kitagawa en 1987 a proposé une technique pour approximer ce filtre et l'idée de base de cette méthode est d'approximer les densités de probabilité pour chaque période par des fonctions linéaires par morceaux. En utilisant la règle du trapèze¹, on approxime les équations (6.5), (6.6) et (6.7), pour $i = 0, 1, \dots, N$, on obtient :

1. $\int_a^b f(x) dx \approx \frac{1}{2} \sum_{i=1}^n (x_{i+1} - x_i) (f(x_{i+1}) + f(x_i))$ pour $i \in [0, n]$

Prédiction

$$\begin{aligned}
\varphi(x_t^i/\underline{y}_{t-1}) &= \int_{-\infty}^{+\infty} \varphi(x_t^i/x_{t-1}, \underline{y}_{t-1})\varphi(x_{t-1}/\underline{y}_{t-1})dx_{t-1} \\
&\approx \frac{1}{2} \sum_{n=1}^N (x_{t-1}^{(n)} - x_{t-1}^{(n-1)}) [\varphi(x_t^i/x_{t-1}^{(n-1)}, \underline{y}_{t-1})\varphi(x_{t-1}^{(n-1)}/\underline{y}_{t-1}) \\
&+ \varphi(x_t^i/x_{t-1}^{(n)}, \underline{y}_{t-1})\varphi(x_{t-1}^{(n)}/\underline{y}_{t-1})] \tag{6.8}
\end{aligned}$$

Correction :

$$\varphi(x_t^i/\underline{y}_t) = \frac{\varphi(y_t/x_t^i, \underline{y}_{t-1})\varphi(x_t^i/\underline{y}_{t-1})}{\varphi(y_t/\underline{y}_{t-1})} \tag{6.9}$$

où

$$\begin{aligned}
\varphi(y_t/\underline{y}_{t-1}) &\approx \frac{1}{2} \sum_{n=1}^N (x_t^{(n)} - x_t^{(n-1)}) [\varphi(y_t/x_t^{(n-1)}, \underline{y}_{t-1})\varphi(x_t^{(n-1)}/\underline{y}_{t-1}) \\
&+ \varphi(y_t/x_t^{(n)}, \underline{y}_{t-1})\varphi(x_t^{(n)}/\underline{y}_{t-1})] \tag{6.10}
\end{aligned}$$

Il suffit de déterminer la distribution de la variable d'état initial, les équations (6.8), (6.9) et (6.10) peuvent être résolus récursivement. Selon Watanabe [35], cette distribution est normale de moyenne $\frac{\alpha}{1-\beta}$ et de variance $\frac{\sigma_w^2}{1-\beta^2}$.

La Log-vraisemblance est alors définie par :

$$\log L = \log[\varphi(\underline{y}_t)] = \sum_{t=1}^T \log[\varphi(y_t/\underline{y}_{t-1})].$$

Contrairement à la méthode d'estimation QML, cette méthode donne la log-vraisemblance exacte tant que la perte causée par l'approximation est négligeable. Les paramètres du modèle à volatilité stochastique sont alors estimés en maximisant la log-vraisemblance exacte.

Annexe A

Théorème de Fubini stochastique

Lemme 11 (Théorème de Fubini stochastique)

Supposons que f et g sont deux processus prévisibles sur $(\Omega, \mathcal{F}, \mathcal{F}_t^{V,W}, \tilde{P})$ tels que :

$$E_{\tilde{P}}\left(\int_0^t |f_s| ds\right) < \infty, E_{\tilde{P}}\left(\int_0^t |g_s|^2 ds\right) < \infty$$

alors

1. $E_{\tilde{P}}\left(\int_0^t f_s ds / \mathcal{F}_t^Y\right) = \int_0^t E_{\tilde{P}}(f_s / \mathcal{F}_s^Y) ds$
2. $E_{\tilde{P}}\left(\int_0^t g_s dY_s / \mathcal{F}_t^Y\right) = \int_0^t E_{\tilde{P}}(g_s / \mathcal{F}_s^Y) dY_s$
3. $E_{\tilde{P}}\left(\int_0^t g_s dV_s / \mathcal{F}_t^Y\right) = 0$

Démonstration

1. Supposons que f est une fonction simple ;

$$f(s) = \sum_{i=1}^k f_i 1_{]a_i, b_i]}(s)$$

où pour tout $i = 1, 2, \dots, k$, $]a_i, b_i]$ est une subdivision de $[0, t]$ et f_i est $\mathcal{F}_{a_i}^{\{V,W\}}$ mesurable. Soit

$$\mathcal{F}_{s,t}^Y = \sigma(Y_u - Y_s; s \leq u \leq t).$$

Notons que :

$$\mathcal{F}_t^Y = \mathcal{F}_{a_i}^Y \vee \mathcal{F}_{a_i,t}^Y \equiv \sigma(\mathcal{F}_{a_i}^Y \cup \mathcal{F}_{a_i,t}^Y)$$

et comme Y_t est un \tilde{P} -mouvement brownien à accroissements indépendants, alors

$$\begin{aligned} E_{\tilde{P}}\left(\int_0^t f_s ds / \mathcal{F}_t^Y\right) &= \sum_{i=1}^k E_{\tilde{P}}(f_i(b_i - a_i) / \mathcal{F}_t^Y) \\ &= \sum_{i=1}^k E_{\tilde{P}}(f_i / \mathcal{F}_{a_i}^Y \vee \mathcal{F}_{a_i,t}^Y)(b_i - a_i) \\ &= \sum_{i=1}^k E_{\tilde{P}}(f_i / \mathcal{F}_{a_i}^Y)(b_i - a_i) \\ &= \int_0^t E_{\tilde{P}}(f_s / \mathcal{F}_s^Y) ds \end{aligned}$$

si $f \geq 0$, on prend f comme la limite d'une suite croissante de fonctions simples, on aura $E_{\tilde{P}}(\int_0^t f_s ds / \mathcal{F}_t^Y) = \int_0^t E_{\tilde{P}}(f_s / \mathcal{F}_s^Y) ds$ d'après le théorème de convergence monotone et pour $f = f^+ - f^-$ on conclut le résultat d'après la linéarité.

2. Si g est une fonction simple :

$$g(s) = \sum_{i=1}^k g_i 1_{]a_i, b_i]}(s)$$

où g_i est $\mathcal{F}_{a_i}^{\{V, W\}}$ mesurable ($i = 1, \dots, k$). En utilisant encore une fois l'indépendance des accroissements de Y_t , on obtient

$$\begin{aligned} E_{\tilde{P}}\left(\int_0^t g_s dY_s / \mathcal{F}_t^Y\right) &= \sum_{i=1}^k E_{\tilde{P}}(g_i(Y_{b_i} - Y_{a_i}) / \mathcal{F}_t^Y) \\ &= \sum_{i=1}^k E_{\tilde{P}}(g_i / \mathcal{F}_t^Y)(Y_{b_i} - Y_{a_i}) \\ &= \sum_{i=1}^k E_{\tilde{P}}(g_i / \mathcal{F}_{a_i}^Y)(Y_{b_i} - Y_{a_i}) \\ &= \int_0^t E_{\tilde{P}}(g_s / \mathcal{F}_s^Y) dY_s \end{aligned}$$

Dans le cas générale on peut approximer g par une suite de processus simples g^n telle que $|g_s^n| \leq |g_s|, p.s \forall s \leq t$

$$E_{\tilde{P}} \left| \int_0^t g_s^n dY_s \right|^2 \leq E_{\tilde{P}} \int_0^t |g_s|^2 ds < \infty,$$

et comme $\{\int_0^t g_s^n dY_s; n \geq 1\}$ est uniformément intégrable, alors :

$$\begin{aligned} E_{\tilde{P}} \left(\int_0^t g_s dY_s / \mathcal{F}_t^Y \right) &= \lim_{n \rightarrow \infty} E_{\tilde{P}} \left(\int_0^t g_s^n dY_s / \mathcal{F}_t^Y \right) \\ &= \lim_{n \rightarrow \infty} \int_0^t E_{\tilde{P}} (g_s^n dY_s / \mathcal{F}_t^Y) \\ &= \int_0^t E_{\tilde{P}} (g_s / \mathcal{F}_s^Y) dY_s \end{aligned}$$

3. Supposons que g est simple. $V_{b_i} - V_{a_i}$ est indépendant de $\mathcal{F}_t^Y \vee \sigma(g_i)$ et on a

$$\begin{aligned} E_{\tilde{P}} \left(\int_0^t g_s dV_s / \mathcal{F}_t^Y \right) &= \sum_{i=1}^k E_{\tilde{P}} (g_i (V_{b_i} - V_{a_i}) / \mathcal{F}_t^Y) \\ &= \sum_{i=1}^k E_{\tilde{P}} [E_{\tilde{P}} (g_i (V_{b_i} - V_{a_i}) / \mathcal{F}_t^Y \vee \sigma(g_i)) / \mathcal{F}_t^Y] \\ &= 0 \end{aligned}$$

pour tout g , on fait de même que (2).

■

Bibliographie

- [1] Alan Bain. Dan Gisan. (2009). Fundamentals of Stochastic Filtering. Springer Science And Business Media.
- [2] Alan Bain. Stochastic Calculus
- [3] Andrew H. Jazwinski.(1970). Stochastic Process and Filtering Theory. Analytical Mechanics Associates. Inc Seabook. Maryland Academic Press. New York and Landon.
- [4] Bernt Oksendal.(2000). Stochastic Differential Equations. An Introduction With Applications. First Edition, Springer. New York.
- [5] C.Penski.(2000).A new numerical method for S.D.E and its application in circuit simulation, *com and appl.Math*,115.461-470.
- [6] Dan Simon.(2001).Kalman Filtering.Embedded Systems Programming,pp 72-79
- [7] D.Alazard.(2006). Introduction au filtre de Kalman. Notes de cours. Exercices corrigés. Sessions Matlab. Version2.
- [8] Edita Kalarova.(2007). An application of the Kalman-Bucy filter to electrical circuit.Brno University.
- [9] Gérard Alengrin.(1974). La Théorie du Filtrage Non Linéaire et ses Applications au traitement du Signal et à L'identification En Automatique.Thèse du Doctorat En Science Physique. Université Paul Sabatier. Toulouse.

- [10] Ghenshiro Kitagawa.(1991). A Non Linear Smoothing Method For Time Series Analysis, *Statistica Sinica*. The Institute of Statistical Mathematics and University of Taylor. 371-388.
- [11] Hui-Hsiung. (2006). *Introduction to Stochastic Integration*. Springer
- [12] H.Kunita.(1982). *Stochastic Partial Differential Equation Connected With Non-Linear Filtering, Non Linear Filtering And Stochastic Control*. Lecture Notes In Mathematics 972. Springer Verlag , Berlin.
- [13] H.Tanizaki, Roberto. S.Mariano.(1998). Nonlinear and Non-Gaussian State Space Modeling With Monte Carlo Simulations. *Journal of Econometrics* 83. 263 - 290.
- [14] I.Karatzas, S.E Shreve.(1991).*Brownian motion and stochastic calculus*,2nd ed.Springer Verlag.
- [15] Jason L. Speyer. Walter H.Chung.(2008).*Stochastic Processes, Estimation, and Control*. Advances in Design and Control Siam.
- [16] Jean Claude Bertin. Roger Ceschi.(1998). *Processus Stochastiques Et Filtrage De Kalman*. Collection Traitement De Signal. Hermes.
- [17] Jie Xiong.(2009).*An Introduction To Stochastic Filtering Theory*. Oxford University Press. Inc. New York.
- [18] Lakhdar Aggoun. Robert Elliot.(2004). *Mesure Theory and Filtering Introduction With Applications*. Cambridge University Press.
- [19] Lili.M,Huiran.W,Jinguang.C.(2010). Analysis of Kalman Filter with Correlated Noises under Different Dependence.*Journal of Information & Computational Science*, pp.1147-1154.
- [20] M.Fujisaki, G.Kallianpur, H.Kunita. (1972). Stochastic differential equation for nonlinear filtering problem. *Osaka Journal of Mathematics* 9. 19-40.
- [21] M.H Davis. (1977). *Linear Estimation And Stochastic Control*. Chapman And Hall.

- [22] M.H Davis.(1984).Stochastic Control And Non Linear Filtering. Tata Institute Of Fondamental research, Bombay.
- [23] M.I.Ribeiro.(2004). Kalman and Extented Kalman Filters : Concept, Dérivation and Properties.Institute for Systems and Robotics.Portugal.
- [24] Mohinder S. Grewal, Angus.P Andrews.(2001). Kalman Filtering : Theory and Practice Using Matlab. Second Edition. John Wiley & Sons, Inc.
- [25] Peter.J Brokwell ,Richard A.Davis.(1996). Time Series : Theory And Methods.Second Edition
- [26] Peter.J Brokwell ,Richard A.Davis.(2002). Introduction toTime Series and Forecasting.Second Edition
- [27] Peter.S.Mybeck.(1982). Stochastic Models Estimation And Control, Volume 2.Mathematics In Science and Engineering Volume141-2. Academic Press.
- [28] Philip.E.Potter.(2004). Stochastic Integration And Differential Equations,Second Edition , Springer-Verlag, Berlin Heidelberg.
- [29] Ramaprasad Bhar.(2010). Stochastic Filtering with Application in Finance.World Scientific Publishing.
- [30] Ramon van Handel. (2007). Stochastic calculs,Filtering and Stochastic Control. Lecture notes. Spring.
- [31] Richard Bronson.(1994).Equation différentielles Méthodes et application.Cours et problème.Séries Shaum
- [32] Rezaeyan.F.Farnoosh.(2009).Stochastic Differential Equations and application of the Kalman-Bucy filter in the modeling of R.C circuit.Applied Mathematical Sciences.Vol4,2010,n 23,1119-1127.
- [33] R.S. Lipster, A.N Shirayev.(1977). Statistics Of Random Processes I.General Theory .
- [34] Tomas. *Björk*.(2003). Arbitrage Theory in Continuous Time. Second Edition. Xford University Press.

- [35] Toshiaki, Watanabe.(1999). A Non Linear Approach to Stochastic Volatility Models With An Application to Daily Stock Returns. *J.Appl.Econ.*14 :101- 121.
- [36] Vincenzo Capasso.David Bakstein.(2005). An Introduction to Continuous-Time Stochastic Processus. *Birkhäuser* Baston(Theory,Models, and Applications to Finance,Biology, and Medicine).
- [37] W.Kampowsky,P.Rentrop,W.Schmidt.(1992).Classification and numerical simulation of electric circuit surveys *math.Indust*,23-65.
- [38] G.Kallianpur. (2004). Non Linear Stochastic Filtering : A Brief Survey. *Journal of Mathematical Sciences*, Vol.120, No 3, pp 13646-1372.

Résumé

L'objet de ce travail porte sur l'estimation de l'état d'un système dynamique linéaire et non linéaire et sur quelques applications du problème du filtrage dans différents domaines. Après un exposé concis des notions fondamentales du calcul stochastique, différentes approches au problème du filtrage ont été envisagées selon la nature temporelle et l'aspect linéaire ou non linéaire du système. Dans le cas d'un système dynamique linéaire, on retrouve moyennant l'approche innovation le filtre optimal de Kalman et en particulier le filtre de Kalman-Bucy dans la cas du systèmes gaussiens. Dans le cas où le système dynamique est non linéaire, le filtre à temps discret est déduit du filtre de Kalman après linéarisation du système et dans le cas à temps continu le filtre optimal est obtenu suivant l'approche changement de mesure d'abord et ensuite l'approche innovation. Trois applications pour illustrer l'intérêt et l'efficacité de la théorie du filtrage, ont été initiées en traitement de signal, en navigation et en finance.

Mots-Clefs : Système dynamique stochastique, le filtre de Kalman, équation de Zakai, équation de Kushner .

Abstract

The object of this work concerns on the estimate of the state of linear and non-linear dynamic system and some applications of the problem of filtering in various fields. After a concise talk of the basic concepts of stochastic calculation, various approaches to the problem of filtering have been considered according to temporal nature and the linear or not linear aspect of the system. In the case of a linear dynamic system, we use the approach innovation to find the optimal kalman filter and in particular we find the kalman-Bucy filter in the case of the Gaussian systems. When the dynamic system is nonlinear, the discrete time filter is deduced from the Kalman filter after linearizing of the system and in the case of continuous time, the optimal filter is obtained according to the approach change of measurement initially and then the approach innovation. Three applications to illustrate the interest and the effectiveness of filtering theory were initiated in signal processing, navigation and finance.

Keywords : Stochastic dynamic system, the Kalman filter, Zakai equation, Kushner equation.