THÈSE

présentée

A LA FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE ROUEN

pour l'obtention du diplôme de

DOCTEUR de l'UNIVERSITÉ de ROUEN

(arrêté du 5 Juillet 1984)

Discipline: MATHÉMATIQUES

Spécialité: PROBABILITÉS

Eric BUSVELLE

IMMERSION DES SYSTÈMES STOCHASTIQUES ET FILTRES FINIS; FILTRAGE PARTICULAIRE

Soutenue le *21* Octobre *1991*

devant le jury composé de

José de SAM LAZARO, *Professeur à l'université de Rouen*	Président
Denis de BRUCQ, *Professeur à l'université de Rouen*	Rapporteur
Étienne PARDOUX, *Professeur à l'université de Provence*	Rapporteur
Claude DELLACHERIE, *Directeur de Recherche au C.N.R.S.*	Examinateur
Jean-Paul GAUTHIER, *Professeur à l'INSA de Rouen*	Examinateur

Daniel RAKOTOPARA, *Ingénieur de recherche à Shell Research* Examinateur

A mes Parents, sans qui…

Je tiens à exprimer ma profonde gratitude à Monsieur le Professeur Denis de Brucq, pour m'avoir guidé et conseillé tout au long de ce travail.

Monsieur le Professeur José de Sam Lazaro, qui me fait l'honneur de présider ce jury, a été le premier à me faire comprendre que les Mathématiques étaient une science vivante. Je tiens à lui manifester ma profonde reconnaissance.

Je remercie aussi Monsieur Claude Dellacherie qui m'a permis de réaliser ce travail en m'accueillant dans le laboratoire d'analyse et modèles stochastiques de Rouen qu'il dirige.

Je remercie tout particulièrement Monsieur Daniel Rakotopara, pour l'aide technique qu'il m'a apportée tout au long de mon travail mais aussi et surtout pour son amitié.

Mes remerciements sincères vont également à Monsieur le Professeur Étienne Pardoux qui a bien voulu participer au jury de cette thèse et en être un des rapporteurs.

Je tiens aussi à remercier Monsieur le Professeur Jean-Paul Gauthier pour avoir accepté de faire partie de ce jury et pour les nombreuses discussions enrichissantes que nous avons eues depuis un an.

Je tiens enfin à remercier toute l'équipe du laboratoire AMS de Rouen pour sa sympathie et sa disponibilité à mon égard, tous les membres du LACIS qui m'ont accueillis parmi eux pendant toute la durée de ma thèse, ainsi que tous les membres de l'équipe d'automatique du centre de recherche de Shell à Grand-Couronne avec lesquels j'ai maintenant le plaisir de travailler.

INTRODUCTION

CONTEXTE ET POSITION DU PROBLEME

La problématique du filtrage est d'estimer l'état d'un processus aléatoire, où mieux, de calculer sa loi à partir d'observations partielles et entachées d'erreur. Mais on veut aussi réaliser un suivi du processus d'état qui évolue dans le temps. On est donc conduit à rechercher un algorithme d'estimation qui, à partir des observations qu'il reçoit, fournit une estimation dynamique de l'état. Cet algorithme doit être implémentable sur un ordinateur, c'est à dire qu'il doit être réalisable en dimension finie. Concrètement, cela signifie qu'il doit être fondé sur des équations différentielles ordinaires, par opposition à des équations aux dérivées partielles ou des équations intégrales. Or, aussi bien pour un processus modélisé par des équations aux différences (temps discret) que par des équations différentielles (temps continu), il existe une équation - l'équation de Zakaï [1]- qui donne l'évolution de la loi conditionnellement aux observations. Mais cette équation est en dimension infinie. Elle possède par contre une propriété que l'on appreciera dans ce travail et qui est la récursivité. Cela signifie que les observations sont utilisées au fur et à mesure pour réactualiser la loi conditionnelle de l'état.

Nous appellerons système le couple état/observation, et nous dirons qu'un tel système admet un filtre de dimension finie* si on peut calculer la loi de l'état au vu des observations en fonction d'un nombre fini de paramètres calculables récursivement. Un système sera dit linéaire gaussien si l'état et l'observation sont des processus gaussiens donnés par des équations d'évolution linéaires. Ce type de systèmes possède de très bonnes propriétés, entre autre celle qui nous intéresse: les systèmes linéaires gaussiens

* Souvent noté FDF qui signifiera suivant le contexte "Filtre de Dimension Finie" ou "Finite-Dimensional Filter"

admettent des filtres de dimension finie. Les équations (différentielles et en nombre fini) qui donnent la loi conditionnelle cherchée sont les très célèbres équations de Kalman [2] qui datent de 1961. Le caractère gaussien de ces systèmes est en grande partie responsable de ce résultat. Nous verrons notamment que c'est par des propriétés de stabilité des lois gaussiennes par rapport à certaines transformations que l'équation de Zakaï dans le cas linéaire gaussien discret s'écrit en dimension finie. Cependant, nous montrerons aussi, prolongeant des travaux de Makowski [3] que des systèmes conditionnellement gaussiens (i.e. l'évolution de l'état dépend des observations) avec une condition initiale non gaussienne admettent également un filtre de dimension finie, bien qu'ils ne soient plus tout à fait gaussiens, ni tout à fait linéaires. Nous verrons aussi le lien étroit qui existe entre l'existence de filtres finis et le caractère exponentiel des lois probabilistes mises en jeu. En effet, il est connu en statistique que les échantillons de lois exponentielles admettent des résumés exhaustifs et que réciproquement, lorsque une statistique est exhaustive, c'est toujours par rapport à une loi exponentielle [4]. La correspondance avec notre problème est évidente et nous verrons que la propriété la plus importante et restrictive à l'existence des filtres finis est la finitude du filtre (et pas la récursivité) donc l'existence d'un résumé exhaustif de dimension finie. Nous exhiberons en particulier une famille de systèmes qui admettent un filtre fini dès que l'observation de ces systèmes suit une loi exponentielle.

Il a fallu très longtemps pour trouver des systèmes non linéaires admettant un filtre fini. C'est en fait Benes [5], 20 ans après Kalman, qui a caractérisé une telle classe de systèmes et qui a ainsi relancé l'interêt des mathématiciens pour les filtres finis. Malheureusement, depuis le résultat de Benes, les résultats les plus profonds ont été des résultats de non existence de filtre finis, tel celui de D. Michel et M. Chaleyat-Maurel [6]. De plus, les résultats d'existence concernent de faibles classes de systèmes et reposent souvent sur des hypothèses de finitude de la dimension d'une certaine algèbre de Lie [7], critères difficiles à vérifier en pratique. Par rapport à ces travaux, nous utiliserons une méthode - l'immersion - qui permet de donner des conditions suffisantes explicites à l'existence de filtres de dimension finie. Les démonstrations sont constructives et permettent d'exhiber de nouveaux systèmes non linéaires admettant des filtres finis. Nous appliquerons cette approche pour les systèmes en temps discret et en temps continu, en essayant de généraliser au maximum les résultats déjà parus sur l'immersion des systèmes stochastiques, dus en particulier à Lévine et Marino [8]. Les outils utilisés sont prôches de ceux utilisés par les automaticiens pour immerger les systèmes non linéaires dans des systèmes linéaires afin de contrôler les premiers en utilisant les techniques classiques attachées aux seconds [9,10]. Nous nous intéresserons particulièrement à l'immersion des systèmes dont l'état dépend de l'observation et nous donnerons des conditions suffisantes d'existence de filtres finis pour de tels systèmes [11,12]. Ces hypothèses seront cependant assez restrictives mais il en sera toujours ainsi avec l'existence des filtres finis. C'est pourquoi, devant les problèmes pratiques, on a du trouver d'autre méthodes pour résoudre le problème du filtrage d'un processus non linéaire.

Il est intéressant, compte tenu de ce qui vient d'être dit, de trouver des méthodes sous-optimales permettant de donner une approximation de la loi de l'état sachant l'observation. Nous nous sommes focalisés sur une méthode issue de la mécanique des fluides [13] et que l'on appelle méthode particulaire. Elle consiste à approcher la loi conditionnelle de l'état par un nombre fini de masses de Dirac pondérées. Ces "particules" suivent le flot de l'équation d'état, avec un facteur de correction qui tient compte des observations. Les coefficients de pondération sur chaque particule sont une mesure du degré de confiance que l'on peut avoir en ces dernières pour représenter effectivement l'état. L'algorithme qui régit cette méthode est issu de l'équation de Zakaï discrète interprétée d'une façon originale. Un résultat de convergence est démontré et plusieurs simulations sont exposées à la fin de ce travail. Notons que le filtrage particulaire concerne une très large classe de systèmes non linéaires et que la méthodologie utilisée, très simple d'approche, peut encore être généralisée.

[1] **M. Zakaï** "On the optimal filtering of diffusion processes" *Zeit. Wahr. Verwand. Geb 11 (1969) p230-243*

[2] **R. Kalman, R. Bucy** "New results in linear filtering and prediction theory" *Trans. ASME Ser. D. J. Basic Eng. 83 (1961) p95-108*

[3] **A. Makowski** "Results on the filtering problem for linear systems with non-gaussian initial conditions" *Proc. 21st IEEE Conf. Decision Cont., Orlando, (1982) p201-204*

[4] **L. Brown** "Sufficient statistics in the case of independent random variables" *Ann. Math. Stat. 35 (1964) p1456-1474*

[5] **V. Benes** "Exact finite dimensional filters for certain diffusions with non-linear drift" *Stochastics 5 (1981) p65-92*

[6] **D. Michel, M. Chaleyat-Maurel** "Des résultats de non-existence de filtres de dimension finie" *Stochastics 13 (1984) p83-102*

[7] **M. Hazewinkel, S.I. Marcus** "On Lie algebras and finite-dimensionely filtering" *Stochastics 7 (1982) p29-62*

[8] **J. Lévine, R. Marino** "Nonlinear system immersion, observers and finite-dimensional filters" *Syst. and Control Letters 7 (1986) p133-142*

[9] **S. Monaco, D. Normand-Cyrot** "Sur la subordination d'un système non linéaire discret à un système linéaire" *Outils et modèles mathématiques pour l'automatique, l'analyse des systèmes, et la théorie du signal, CNRS (1983)*

[10] **D. Claude, M. Fliess, A. Isidori** "Immersion directe et par bouclage d'un

système non linéaire dans un linéaire" *C.R.A.S. Paris Série 1 Vol. 296 (1983)*

[11] **E. Busvelle, D. Rakotopara, D. de Brucq** "A sufficient condition for the existence of finite dimensional filters for discrete-time non-linear systems" *29th IEEE Conf. Decision Cont. (1990) p234-239*

[12] **D. Rakotopara, E. Busvelle, H. Hammouri** "Immersion in conditionally gaussian systems and finite dimensional filters" *29th IEEE Conf. Decision Cont. (1990) p234-239*

[13] **P.A. Raviart** "An analysis of particle methods" *CEME course in numerical methods in fluid dynamics, COMO, Lecture Notes in Mathematics, Springer-Verlag (1983)*

SOMMAIRE*

* Certaines parties de cette thèse sont rédigées en anglais. Elles correspondent à des articles publiés où soumis pour publication

CHAPITRE I

AUTOUR DE LA DEFINITION D'UN FILTRE FINI

INTRODUCTION

Nous allons pour commencer donner des définitions précises des objets et des propriétés mentionnées dans l'introduction.

Nous supposerons toujours donné un espace de probabilité (Ω, A, P) où Ω est un espace Polonais muni de sa tribu borelienne, ce qui fait de (Ω, A) un espace mesurable standard.

Nous allons maintenant donner une définition de la notion de filtre de dimension finie. Cette définition concerne une famille très large de systèmes en temps discrets. Plus précisémment, soient $(X_t)_{t\in I\!N}$ et $(Y_t)_{t\in I\!N}$ deux processus à valeurs dans $(\underline{\underline{X}}, \mathcal{X})$ et dans $(\underline{\underline{Y}}, \mathcal{Y})$, qui représenteront respectivement l'état et l'observation du système étudié. Notons pour tout $t \in I\!N$

$$\mathcal{F}_t^X = \sigma(X_0, \ldots, X_t)$$

$$\mathcal{F}_t^Y = \sigma(Y_0, \ldots, Y_t)$$

$$\mathcal{F}_t = \mathcal{F}_t^X \vee \mathcal{F}_t^Y$$

$$\mathcal{F}_\infty = \sigma\left(\sum_{s=0}^{\infty} \mathcal{F}_s\right)$$

$$Y^t = (Y_0, Y_1, \ldots, Y_t)$$

On supposera toujours, à partir de cet instant, les conditions suivantes réalisées:

(c1) $\forall t \in I\!N \qquad X_{t+1} \perp \mathcal{F}_t / X_t$

(c2) $\forall t \in I\!N \qquad Y_t = h(X_t, V_t)$ où $(V_t)_{t\in I\!N}$ est une suite de variables aléatoires indépendantes, suite elle même indépendante du processus $(X_t)_{t\in I\!N}$.

La condition *(c1)* exprime le fait que $(X_t)_{t\in I\!N}$ est supposé $(\mathcal{F}_t)$-Markovien, et la condition *(c2)* signifie que pour tout entier t, Y_t ne dépend que de X_t.

Un exemple de tel système est un système qui s'écrit

$$\begin{cases} X_{t+1} = f(X_t) + g(X_t)W_t \\ Y_t = h(X_t) + V_t \end{cases}$$

où f, g, et h sont des fonctions définies sur $\underline{\underline{X}}$ et où (W_t), (V_t) et X_0 sont des variables aléatoires indépendantes.

1. FILTRE DE DIMENSION FINIE EN TEMPS DISCRET

Nous dirons d'un ensemble quelconque que c'est un espace de dimension finie s'il s'injecte dans un espace vectoriel de dimension finie. Ainsi, un tel espace est caractérisé par le fait que chacun de ses points est déterminé par un nombre fini de coordonnées. Mais l'ensemble peut être totalement dépourvu de structure. Nous allons maintenant donner la définition de l'objet qui sera le sujet principal de ce travail.

Définition 1.1

Nous dirons qu'un système admet un filtre de dimension finie $(Z_t)_{t\in I\!N}$ ssi il existe un espace $\underline{\underline{Z}}$ de dimension n finie, une suite d'applications

$$\varphi_0 : \underline{\underline{Y}} \to \underline{\underline{Z}} \text{ et } \varphi_t : \underline{\underline{Z}} \times \underline{\underline{Y}} \to \underline{\underline{Z}} \qquad \forall t \geq 1$$

et une suite d'applications

$$\psi_t : \underline{\underline{Z}} \times \underline{\underline{X}} \to [0,1] \qquad \forall t \geq 0$$

le tout tel que pour tout ensemble mesurable $\underline{\underline{X}}_a \in \underline{\underline{X}}$, on ait

$$P\left\{X_t \in \underline{\underline{X}}_a / \mathcal{F}_t^Y\right\} = \psi_t(Z_t, \underline{\underline{X}}_a)$$

Z_t étant une variable aléatoire définie récursivement par

$$Z_0 = \varphi_0(Y_0) \text{ et } \forall s \geq 1 \qquad Z_s = \varphi_s(Z_{s-1}, Y_s)$$

Cette définition est issue de la notion intuitive que l'on avait d'un F.D.F., à savoir l'existance d'une statistique exhaustive qui puisse être calculée récursivement. Mais elle n'est pas du tout intrinsèque et, donnée de cette façon, plus très intuitive. Nous allons donc donner quelques propriétés caractéristiques des F.D.F.

2. PROPRIÉTÉS INTRINSEQUES D'UN FILTRE FINI

Commençons par rappeller quelque propriétés sur l'espérance conditionnelle et l'indépendance conditionnelle.

Définition 2.1

Soient $\mathcal{A}$, $\mathcal{B}$, et $\mathcal{C}$ trois tribus. On dit que $\mathcal{A}$ est indépendante de $\mathcal{B}$ conditionnellement à $\mathcal{C}$ et on écrit $\mathcal{A} \perp \mathcal{B}/\mathcal{C}$ si par définition

$$\forall a \in \mathcal{A} \quad \forall b \in \mathcal{B} \qquad E\left[ab/\mathcal{C}\right] = E\left[a/\mathcal{C}\right] E\left[b/\mathcal{C}\right]$$

et cette définition est équivalente à

$$\forall b \in \mathcal{B} \qquad E\left[b/\mathcal{A} \vee \mathcal{C}\right] = E\left[b/\mathcal{C}\right]$$

Proposition 2.2

Supposons $\mathcal{A} \perp \mathcal{B}/\mathcal{C}$ et soit $\mathcal{D} \subseteq \mathcal{A} \vee \mathcal{C}$. On a a fortiori

$$\mathcal{D} \perp \mathcal{B}/\mathcal{C} \text{ et } \mathcal{A} \perp \mathcal{B}/\mathcal{C} \vee \mathcal{D}$$

Démonstration

La première est très simple. Soit $b \in \mathcal{B}$, vérifions que $E\left[b/\mathcal{C}\right] = E\left[b/\mathcal{C} \vee \mathcal{D}\right]$

(i) $E\left[b/\mathcal{C}\right]$ est bien $\mathcal{C} \vee \mathcal{D}$-mesurable

(ii) Soient $C \in \mathcal{C}$ et $D \in \mathcal{D}$

$$\begin{aligned}
\int_{C \cap D} E\left[b/\mathcal{C}\right] dP &= \int 1_D E\left[b/\mathcal{C}\right] dP \\
&= \int_C E\left[1_D/\mathcal{C}\right] E\left[b/\mathcal{C}\right] dP \\
&= \int_C E\left[1_D b/\mathcal{C}\right] dP = \int_C 1_D b dP = \int_C b dP
\end{aligned}$$

La seconde n'est guère plus difficile. Soit encore $b \in \mathcal{B}$,

$$E\left[b/\mathcal{C} \vee \mathcal{D} \vee \mathcal{A}\right] = E\left[b/\mathcal{C} \vee \mathcal{A}\right] = E\left[b/\mathcal{C}\right] = E\left[b/\mathcal{C} \vee \mathcal{D}\right]$$

Nous allons maintenant établir un résultat qui nous sera très utile par la suite (cf Mouchard et Rollin).

Proposition 2.3

Soient A, B, et C trois tribus, on a

$$\left.\begin{array}{ll} (i) & \mathcal{A} \perp \mathcal{B}/\mathcal{C} \\ (ii) & \mathcal{D} \perp \mathcal{B}/(\mathcal{A} \vee \mathcal{C}) \end{array}\right\} \iff (iii) \quad (\mathcal{A} \vee \mathcal{D}) \perp \mathcal{B}/\mathcal{C}$$

Démonstration

Supposons (i) et (ii) réalisés et soit $b \in \mathcal{B}$,

$$E\left[b/\mathcal{A} \vee \mathcal{C} \vee \mathcal{D}\right] = E\left[b/\mathcal{A} \vee \mathcal{C}\right] = E\left[b/\mathcal{C}\right]$$

et réciproquement, supposons (iii), il est clair que (i) est a fortiori vérifiée, il reste à constater que si $b \in \mathcal{B}$,

$$E\left[b/\mathcal{A} \vee \mathcal{C} \vee \mathcal{D}\right] \overset{\text{(iii)}}{=} E\left[b/\mathcal{C}\right] \overset{\text{(i)}}{=} E\left[b/\mathcal{C} \vee \mathcal{A}\right]$$

Remarque 2.4

On utilisera souvent cette proposition sous la forme suivante

$$\mathcal{C} \subseteq \mathcal{B} \text{ et } \left.\begin{array}{c} \mathcal{A} \perp \mathcal{B}/\mathcal{C} \\ \mathcal{D} \perp \mathcal{B}/\mathcal{A} \end{array}\right\} \quad \Longrightarrow \quad \mathcal{D} \perp \mathcal{B}/\mathcal{C}$$

Théorème 2.5

Soit Y une variable aléatoire à valeurs dans un espace mesurable $(\underline{\underline{Y}}, \mathcal{Y})$, et Z une variable aléatoire à valeurs dans $I\!R$ muni de sa tribu borélienne. Z est $\sigma(Y)$ mesurable si et seulement si il existe une application φ $\mathcal{Y}$-mesurable telle que

$$\forall \omega \in \Omega \qquad Z(\omega) = \varphi(Y(\omega))$$

On pourra trouver la preuve de ce résultat dans Lehmann.

Lemme 2.6

Soient X,Y et Z trois variables aléatoires, avec $Z \in \sigma(Y)$. Il y a équivalence des deux propositions suivantes:

(1) $X \perp Y/Z$

(2) $\exists \psi : \underline{\underline{Z}} \times \mathcal{X} \to I\!R$ mesurable en sa première variable, telle que

$$\forall \underline{\underline{X}}_a \in \mathcal{X} \qquad E\left[1_{\underline{\underline{X}}_a}(X)/Y\right] = \psi(Z, \underline{\underline{X}}_a)$$

Démonstration

(1) $\Longrightarrow$ (2)

Soit $\underline{\underline{X}}_a \in \mathcal{X}$

$$\begin{aligned} E\left[1_{\underline{\underline{X}}_a}(X)/Y\right] &= E\left[1_{\underline{\underline{X}}_a}(X)/Y, Z\right] \\ &= E\left[1_{\underline{\underline{X}}_a}(X)/Z\right] \end{aligned}$$

qui est bien une fonction de Z et de $\underline{\underline{X}}_a$.

(2) $\Longrightarrow$ (1)

Soit $\underline{\underline{X}}_a \in \mathcal{X}$

$$\begin{aligned}
E\left[1_{\underline{\underline{X}}_a}(X)/Z\right] &= E\left[E\left[1_{\underline{\underline{X}}_a}(X)/Y\right]/Z\right] \\
&= E\left[\psi(Z, \underline{\underline{X}}_a)/Z\right] \\
&= \psi(Z, \underline{\underline{X}}_a) = E\left[1_{\underline{\underline{X}}_a}(X)/Y, Z\right]
\end{aligned}$$

donc $X \perp Y/Z$

Théorème 2.7

Soit (X_t, Y_t) un système. Si $(Z_t)_{t \in I\!N}$ est un filtre de dimension finie, alors $(Z_t)_{t \in I\!N}$ est un processus $\mathcal{F}_t^Y$-Markovien tel que

$$\forall t \in I\!N \qquad X_t \perp Y^t / Z_t$$

Démonstration

Pour commencer, on a bien

$$\forall t \in I\!N \qquad X_t \perp Y^t / Z_t$$

De plus, pour tout $t \in I\!N$, on a

$$\begin{aligned}
Z_t &= \varphi_t(Y_t, Z_{t-1}) \\
&= \varphi_t(Y_t, \varphi_{t-1}(Y_{t-1}, Z_{t-2})) \\
&= \cdots \\
&= \Phi_t(Y^t)
\end{aligned}$$

donc $(Z_t)_{t \in I\!N}$ est $(\mathcal{F}_t^Y)$-adapté. Il reste à montrer que

$$\forall t \geq 1 \qquad Z_t \perp Y^{t-1} / Z_{t-1}$$

Soit t un entier strictement positif, on a par hypothèses

$$Y_{t+1} \perp Y^t / X_{t+1} \vee X_t$$

Puisque de plus

$$X_{t+1} \perp Y^t / X_t$$

on en déduit

$$Y_{t+1} \perp Y^t / X_t$$

Ceci couplé au fait que $X_t \perp Y^t / Z_t$ entraine d'après la remarque 2.4

$$Y_{t+1} \perp Y^t / Z_t$$

ce qui entraine

$$Y_{t+1} \vee Z_t \perp Y^t / Z_t$$

soit finallement

$$Z_{t+1} \perp Y^t / Z_t$$

puisque $Z_{t+1} = \varphi_t(Y_{t+1}, Z_t)$ et donc

$$Z_{t+1} \in \sigma(Y_{t+1}, Z_t)$$

Le résultat suivant est démontré, par exemple, dans Ikeda-Watanabe [1981].

Théorème 2.8

Soient X une variable aléatoire à valeurs dans $(\underline{\underline{X}}, \mathcal{X})$ polonais et Y une variable aléatoire à valeurs dans $(\underline{\underline{Y}}, \mathcal{Y})$. Il existe une famille $\left\{P(y, \underline{\underline{X}}_a); y \in \underline{\underline{Y}}, \underline{\underline{X}}_a \in X\right\}$ qui vérifie les trois conditions suivantes:

(1) $\forall y \in \underline{\underline{Y}} \quad P(y, .)$ est une probabilité sur $(\underline{\underline{X}}, \mathcal{X})$.

(2) $\forall \underline{\underline{X}}_a \in \mathcal{X} \quad P(., \underline{\underline{X}}_a)$ est une application mesurable de $(\underline{\underline{Y}}, \mathcal{Y})$ dans $[0,1]$ muni de sa tribu borélienne $\mathcal{B}_{[0,1]}$.

(3) $\forall \varphi \in L^1(\underline{\underline{X}})$

$$E\left[\varphi(X)/Y = y\right] = \int_{\underline{\underline{X}}} \varphi(x) P(y, dx) \qquad P_{\underline{\underline{Y}}} - \text{p.s.}$$

De plus, s'il existe une autre famille $\left\{Q(y, \underline{\underline{X}}_a); y \in \underline{\underline{Y}}, \underline{\underline{X}}_a \in \mathcal{X}\right\}$ vérifiant les trois conditions (1) à (3), alors il existe un ensemble $P_{\underline{\underline{Y}}}$-négligeable $\underline{\underline{N}}$ de $\mathcal{Y}$ tel que

$$\forall y \in \underline{\underline{Y}} - \underline{\underline{N}} \quad \forall \underline{\underline{X}}_a \in \mathcal{X} \qquad P\left(y, \underline{\underline{X}}_a\right) = Q\left(y, \underline{\underline{X}}_a\right)$$

Nous sommes maintenant en mesure de démontrer un résultat qui prouve que la propriété de récursivité qui apparaît dans la définition des filtres finis (ou la propriété correspondante de caractère Markovien dans la définition intrinsèque) n'est pas une contrainte primordiale. C'est en fait l'existence d'un paramètre exhaustif à tout instant t qui caractérise les systèmes admettant un filtre fini. En effet, on a le théorème suivant

Théorème 2.9

Soit (X_t, Y_t) un système, et supposons qu'il existe un processus $(Z_t)_{t \in I\!N}$ à valeurs dans un espace de dimension finie, $(\mathcal{F}_t^Y)$-adapté, tel que

$$\forall t \in I\!N \qquad X_t \perp Y^t / Z_t$$

alors le système admet un filtre fini.

Démonstration

Définition d'un filtre fini

Le théorème 2.8 nous assure de l'existence d'une famille

$$(P_{1,t}\left\{y^t, \underline{\underline{X}}_a\right\}; y^t \in \underline{\underline{Y}}^t, \underline{\underline{X}}_a \in \mathcal{X})$$

loi de probabilité régulière de X_t sachant Y^t, et d'une famille

$$(P_{2,t}\left\{z_t, \underline{\underline{X}}_a\right\}; z_t \in \underline{\underline{Z}}, \underline{\underline{X}}_a \in X)$$

loi de probabilité régulière de X_t sachant Z_t. Mais Z_t est $\mathcal{F}_t^Y$-adapté donc il existe une application mesurable $\Phi_t : \underline{\underline{Y}}^t \to \underline{\underline{Z}}$ telle que $Z_t = \Phi_t(Y^t)$ donc la seconde famille s'écrit aussi

$$(P_{2,t}\left\{\Phi_t(y^t), \underline{\underline{X}}_a\right\}; y^t \in \underline{\underline{Y}}^t, \underline{\underline{X}}_a \in \mathcal{X})$$

et elle est égale presque partout (au sens du théorème précédent) à la première famille, car elle vérifie les même conditions. Ainsi, il existe un ensemble $P_{\underline{\underline{Y}}^t}$-négligeable $\underline{\underline{N}}_t$ tel que

$$\forall y^t \in \underline{\underline{Y}}^t - \underline{\underline{N}} \qquad P_{2,t}\left\{\Phi_t(y^t), .\right\} = P_{1,t}\left\{y^t, .\right\}$$

L'application $z_t \to P_{2,t}\left\{z_t, .\right\}$ peut être rendue injective, quitte à quotienter $\underline{\underline{Z}}$ par $\mathcal{R}_t$ où par définition

$$z\mathcal{R}_t z' \iff P_{2,t}\left\{z, .\right\} = P_{2,t}\left\{z', .\right\}$$

On notera s_t la surjection canonique mesurable de $\underline{\underline{Z}}$ dans $\underline{\underline{Z}}/\mathcal{R}_t$. Il est clair que l'on peut définir la famille

$$(P_{2,t}\left\{\dot{z}_t, \underline{\underline{X}}_a\right\}; \dot{z}_t \in \underline{\underline{Z}}/R_t, \underline{\underline{X}}_a \in \mathcal{X})$$

Pour tout t, la variable aléatoire $\dot{Z}_t$ est $\mathcal{F}_t^Y$-mesurable puisqu'elle s'écrit simplement $\dot{Z}_t = s_t \circ \Phi_t(Y^t)$, et elle vérifie aussi

$$X_t \perp Y^t / \dot{Z}_t$$

En effet, $\forall \varphi \in L^1(\underline{\underline{X}})$, $\quad \forall y^t \in \underline{\underline{Y}}^t - \underline{\underline{N}}_t$,

$$\begin{aligned} E\left[\varphi(X_t)/Y^t = y^t\right] &= \int_{\underline{\underline{X}}} \varphi(x) P_{1,t}\left\{y^t, dx\right\} \\ &= \int_{\underline{\underline{X}}} \varphi(x) P_{2,t}\left\{s_t \circ \Phi_t(y^t), dx\right\} \\ &= E\left[\varphi(X_t)/\dot{Z}_t = s_t \circ \Phi_t(y^t)\right] \end{aligned}$$

Montrons que $(\dot{Z}_t)_{t\in I\!N}$ est un filtre de dimension finie, c'est à dire, compte tenu de ce qui a déjà été montré, que $\dot{Z}_t$ est une fonction de $\dot{Z}_{t-1}$ et de Y_t.

Soit $\varphi \in L^1(\underline{\underline{X}})$

$$\begin{aligned} E\left[\varphi(X_t)/Y^{t-1},\dot{Z}_{t-1}\right] &= E\left[\varphi(X_t)/Y^{t-1}\right] \\ &= E\left[E\left[\varphi(X_t)/X_{t-1},Y^{t-1}\right]/Y^{t-1}\right] \\ &= E\left[E\left[\varphi(X_t)/X_{t-1}\right]/Y^{t-1}\right] \\ &= E\left[E\left[\varphi(X_t)/X_{t-1}\right]/Y^{t-1},\dot{Z}_{t-1}\right] \\ &= E\left[E\left[\varphi(X_t)/X_{t-1}\right]/\dot{Z}_{t-1}\right] \\ &= E\left[E\left[\varphi(X_t)/X_{t-1},\dot{Z}_{t-1}\right]/\dot{Z}_{t-1}\right] \\ &= E\left[\varphi(X_t)/\dot{Z}_{t-1}\right] \end{aligned}$$

donc $X_t \perp Y^{t-1}/\dot{Z}_{t-1}$ et en appliquant deux fois la proposition habituelle,

$$Y_t \perp Y^{t-1} \vee \dot{Z}_{t-1}/X_t \Longleftrightarrow \left\{ \begin{array}{r} \dot{Z}_{t-1} \perp Y_t/X_t \\ Y^{t-1} \perp Y_t/\dot{Z}_{t-1} \vee X_t \end{array} \right.$$

puis

$$\left. \begin{array}{r} Y^{t-1} \perp Y_t/\dot{Z}_{t-1} \vee X_t \\ X_t \perp Y^{t-1}/\dot{Z}_{-1} \end{array} \right\} \Longleftrightarrow X_t \vee Y_t \perp Y^{t-1}/\dot{Z}_{t-1}$$

ce qui entraine successivement

$$\begin{aligned} X_t \vee Y_t \perp Y^{t-1}/\dot{Z}_{t-1} &\Longrightarrow X_t \vee Y_t \perp Y^{t-1}/\dot{Z}_{t-1} \vee Y_t \\ &\Longrightarrow X_t \perp Y^{t-1}/\dot{Z}_{t-1} \vee Y_t \\ &\Longrightarrow X_t \perp Y^{t}/\dot{Z}_{t-1} \vee Y_t \end{aligned}$$

Il n'y a plus qu'à appliquer une dernière fois le théorème 2.8 au couple de variable aléatoires $(\dot{Z}_{t-1}, Y_t)$, qui est aussi $\mathcal{F}_t^Y$-mesurable puisque $(\dot{Z}_{t-1}, Y_t) = (s_{t-1} \circ \Phi_{t-1}(Y^{t-1}), Y_t)$. On en déduit l'existence d'une famille

$$(P_{3,t}\left\{(s_{t-1} \circ \Phi_{t-1}(y^{t-1}), y_t), \underline{\underline{X}}_a\right\}; y^t \in Y^t, \underline{\underline{X}}_a \in \mathcal{X})$$

telle que pour presque tout $y^t \in Y^t$, on ait

$$P_{3,t}\left\{(s_{t-1} \circ \Phi_{t-1}(y^{t-1}), y_t), .\right\} = P_{2,t}\left\{s_t \circ \Phi_t(y^t), .\right\}$$

Par injectivité, $\dot{z}_t$ est déterminé de façon unique, connaissant y_t et $\dot{z}_{t-1}$.

REFERENCES

[1] **Ikeda N., Watanabe S.**"Stochastic Differential Equations and Diffusion Processes" *North-Holland Publishing Company*

[2] **Liptzer R.Ch., Shyriaev An.**"Statistique des processus stochastiques" *Ed. Naouka*

[3] **Mouchard M., Rollin J.M.** "A Note on conditional independance with statistical application"*Statistica 4*

[4] **Szpirglas J.**"Sur les propriétés Markoviennes du processus de filtrage"*Proc. of the ENST-CNET, Paris, Springer-Verlag*

CHAPITRE II

ETUDE DES SYSTÈMES EN TEMPS DISCRET

1. EQUATION DE ZAKAI DISCRETE

Nous allons reprendre les hypothèses *(c1)* et *(c2)* du chapitre précédent mais légèrement élargies. On considère un processus $(X_t, Y_t)_{t \geq 0}$ avec par convention $Y_0 = 0$ et

(c'1) $X_t \perp (X^{t-1}, Y^{t-1})/(X_{t-1}, Y_{t-1})$

(c'2) $Y_t \perp (X^t, Y^{t-1})/(X_t, Y_{t-1})$

auxquelles nous ajouterons les deux hypothèses suivantes

(c3) $\mathcal{P}_{X_t}^{X_{t-1}, Y_{t-1}} = \{P_{X_t}^{X_{t-1}=x, Y_{t-1}=y}; x \in \underline{\underline{X}}, y \in \underline{\underline{Y}}\}$ pour tout $t \geq 1$ est dominée par μ (μ étant une mesure sur $\underline{\underline{X}}$) et $dP_{X_t}^{X_{t-1}=\xi, Y_{t-1}=y_{t-1}}(x) = \Pi_t(\xi, x) d\mu$

On omettra de répéter y_t dans les notations pour ne pas trop les surcharger. Les notations génériques P_A^B et f_A^B signifieront respectivement probabilité et densité de A sachant B.

(c4) $\mathcal{P}_{Y_t}^{X_t, Y_{t-1}} = \{P_{Y_t}^{X_t=x, Y_{t-1}=y}; x \in \underline{\underline{X}}, y \in \underline{\underline{Y}}\}$ pour tout $t \geq 1$ est dominée par ν (ν étant une mesure sur $\underline{\underline{Y}}$) et $dP_{Y_t}^{X_t=x, Y_{t-1}=y_{t-1}}(y) = \Phi_t(x, y) d\nu$.

Théorème 1.1

Si on note de façon abrégée $p_t(x) = f_{X_t}^{Y^t=y^t}(x)$, on a l'équation de type Zakaï suivante:

$$p_t(x) = \frac{1}{D_t} \Phi_t(x, y_t) \int_X \Pi_t(\xi, x) p_{t-1}(\xi) d\mu(\xi)$$

où D_t est un simple facteur de normalisation que l'on précisera dans la démonstration (élémentaire) de ce théorème.

Démonstration

Notons que ce théorème à été démontré dans l'article de Di Masi et Runggaldier [1] par une technique utilisant le théorème de Girsanov (en temps discret). Cette procédure est à rapprocher de la démonstration classique de ce théorème en temps continu. Pour notre part, nous démontrerons ce résultat en utilisant simplement des considérations Bayésiennes. Mais nous aurons à utiliser plus loin la version temps discret du théorème de Girsanov telle qu'elle a été énoncée dans l'article çi-dessus mentionné.

Soit $\varphi : \underline{\underline{X}} \to I\!R^+$ une fonction mesurable à valeurs numériques positives. On a pour tout $t \geq 1$

$$\begin{aligned}
E[\varphi(X_t)/Y^{t-1} = y^{t-1}] &= \int_{\underline{\underline{X}}} \varphi(x) dP_{X_t}^{Y^{t-1}=y^{t-1}}(x) \\
&= E[E[\varphi(X_t)/Y^{t-1} = y^{t-1}, X_{t-1}]Y^{t-1} = y^{t-1}] \\
&= \int_{\underline{\underline{X}}} \int_{\underline{\underline{X}}} \varphi(x) dP_{X_t}^{X_{t-1}=\xi, Y^{t-1}=y^{t-1}}(x) dP_{X_{t-1}}^{Y^{t-1}=y^{t-1}}(\xi) \\
&= \int_{\underline{\underline{X}}} \varphi(x) \int_{\underline{\underline{X}}} \Pi_t(\xi, x) dP_{X_{t-1}}^{Y^{t-1}=y^{t-1}}(\xi) d\mu(x)
\end{aligned}$$

Ce calcul prouve que $P_{X_t}^{Y^{t-1}} << \mu$ et que sa densité vaut

$$\int_{\underline{\underline{X}}} \Pi_t(\xi, x) dP_{X_{t-1}}^{Y^{t-1}=y^{t-1}}(\xi)$$

résultat que nous allons utiliser dans la suite.

Soient $\varphi : \underline{\underline{X}} \to I\!R^+$ et $\psi : \underline{\underline{Y}} \to I\!R^+$ deux fonctions mesurables à valeurs numériques positives. Notons $y^t = (y^{t-1}, y)$, on a

$$\begin{aligned}
E[\varphi(X_t)\psi(Y_t)/Y^{t-1} = y^{t-1}] &= \int_{\underline{\underline{X}} \times \underline{\underline{Y}}} \varphi(x)\psi(y) dP_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}}(x,y) \\
&= \int_{\underline{\underline{X}}} \varphi(x) \int_{\underline{\underline{Y}}} \psi(y) dP_{Y_t}^{X_t=x, Y_{t-1}=y_{t-1}}(y) dP_{X_t}^{Y^{t-1}=y^{t-1}}(x) \\
&= \int_{\underline{\underline{X}}} \varphi(x) \int_{\underline{\underline{Y}}} \psi(y) f_{Y_t}^{X_t=x, Y_{t-1}=y_{t-1}}(y) d\nu(y) f_{X_t}^{Y^{t-1}=y^{t-1}}(x) d\mu(x) \\
&= \int_{\underline{\underline{X}} \times \underline{\underline{Y}}} \varphi(x)\psi(y) f_{Y_t}^{X_t=x, Y_{t-1}=y_{t-1}}(y) f_{X_t}^{Y^{t-1}=y^{t-1}}(x) d\mu \otimes \nu(x,y)
\end{aligned}$$

Ceci entraine que la famille de probabilités

$$\mathcal{P}_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}} = \{P_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}} ; y^{t-1} \in \underline{\underline{Y}}^{t-1}\}$$

est dominée par la mesure produit $\mu \otimes \nu$. De plus, on a obtenu

$$f_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}}(x,y) = f_{Y_t}^{X_t=x, Y_{t-1}=y_{t-1}}(y) f_{X_t}^{Y^{t-1}=y^{t-1}}(x) \tag{1}$$

En intégrant cette fois tout d'abord en x, on obtient d'une part

$$\begin{aligned}
E[\varphi(X_t)\psi(Y_t)/Y^{t-1} = y^{t-1}] &= \int_{\underline{\underline{X}} \times \underline{\underline{Y}}} \varphi(x)\psi(y) dP_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}}(x,y) \\
&= \int_{\underline{\underline{Y}}} \psi(y) \int_{\underline{\underline{X}}} \varphi(x) dP_{X_t}^{Y^t=y^t}(x) dP_{Y_t}^{Y^{t-1}=y^{t-1}}(y)
\end{aligned} \tag{2}$$

et d'autre part, on peut toujours écrire

$$(3)\quad \begin{aligned} E[\varphi(X_t)\psi(Y_t)/Y^{t-1}=y^{t-1}] &= \int_{\underline{\underline{X}}\times\underline{\underline{Y}}} \varphi(x)\psi(y) f_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}}(x,y)d\mu\otimes\nu(x,y) \\ &= \int_{\underline{\underline{Y}}} \psi(y) \int_{\underline{\underline{Y}}} \varphi(x) \frac{f_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}}(x,y)}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y)} d\mu(x) f_{Y_t}^{Y^{t-1}=y^{t-1}}(y)d\nu(y) \end{aligned}$$

En identifiant (2) et (3), on obtient l'existence de la densité suivante, ainsi qu'une écriture de cette densité:

$$f_{X_t}^{Y^t=y^t}(x) = \frac{f_{(X_t,Y_t)}^{Y^{t-1}=y^{t-1}}(x,y)}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y)}$$

et grâce à (1), on déduit immédiatement la formule de type Bayesienne suivante

$$f_{X_t}^{Y^t=y^t}(x) = \frac{f_{Y_t}^{X_t=x,Y_{t-1}=y_{t-1}}(y) f_{X_t}^{Y^{t-1}=y^{t-1}}(x)}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y)}$$

Le dénominateur peut être considéré comme un facteur de normalisation, et on obtient finallement la formule

$$f_{X_t}^{Y^t=y^t}(x) = \frac{f_{Y_t}^{X_t=x,Y_{t-1}=y_{t-1}}(y) f_{X_t}^{Y^{t-1}=y^{t-1}}(x)}{\int_{\underline{\underline{X}}} f_{Y_t}^{X_t=\xi,Y_{t-1}=y_{t-1}}(y) f_{X_t}^{Y^{t-1}=y^{t-1}}(\xi)d\mu(\xi)}$$

et si l'on adopte les notations suivantes, pour tout entier t

$$p_t(x) = f_{X_t}^{Y^t=y^t}(x)$$

et

$$D_t = \int_{\underline{\underline{X}}} f_{Y_t}^{X_t=\xi,Y_{t-1}=y_{t-1}}(y) f_{X_t}^{Y^{t-1}=y^{t-1}}(\xi)d\mu(\xi)$$

et que l'on remarque que

$$f_{X_t}^{Y^{t-1}=y^{t-1}}(x) = \int_{\underline{\underline{X}}} f_{X_t}^{X_{t-1}=\xi,Y_{t-1}=y_{t-1}}(x) f_{X_{t-1}}^{Y^{t-1}=y^{t-1}}(\xi)d\mu(\xi)$$

alors on obtient bien la formule annoncée.

□

Remarque 1.2

Les systèmes qui satisfont les hypothèses de ce théorème peuvent être du type

$$\begin{cases} X_{t+1} = f(X_t, Y_t, W_t) \\ Y_{t+1} = h(X_t, Y_t, V_t) \end{cases}$$

pour peu que les variables aléatoires (W_t) et (V_t) soient de dimension suffisament grande afin de vérifier *(c3)* et *(c4)*.

Notons aussi que ce théorème est valable y compris en présence d'un facteur dépendant de l'état devant le bruit d'observation, lorsque l'équation d'observation est de la forme

$$y_{t+1} = h(x_t, y_t) + \nu(x_t, y_t)V_t$$

et ceci tant que l'hypothèse *(c4)* du théorème 1.1 est satisfaite, i.e. tant que le bruit d'observation n'est pas dégénéré. Cette remarque montre une différence fondamentale avec le cas du temps continu oú l'on sait qu'un facteur dépendant de l'état devant le bruit d'observation introduit une information déterministe sur le processus quadratique associé à X donc sur la loi de X, ce qui dégénère les équations du filtrage, telle que l'équation de Zakaï.

En notant que les densités qui apparaissent dans le théorème 1.1 sont directement issues des transitions qui définissent le système, on constate que cette équation qui met en jeu la densité conditionnelle recherchée est fondamentale. Il est évident que en toute généralité, cette équation n'admet pas de solution en dimension finie, en ce sens que pour connaître $p_t(x)$ en un point, il faut connaître intégralement la fonction p_{t-1} qui est un point d'un espace fonctionnel a priori de dimension infinie.

Notons que le facteur de normalisation peut être abandonné pour la plupart des calculs théoriques qui partent de cette équation pour déterminer p_t, comme le montre la remarque suivante.

Corollaire 1.3

Notons $q_0(x) = p_0(x) = f_{X_0}(x)$ et $\forall t \in I\!N^*$

$$q_t(x) = \Phi_t(x, y_t) \int_{\underline{\underline{X}}} \Pi_t(\xi, x) q_{t-1}(\xi) d\xi$$

alors q_t est appellée densité conditionnelle non normalisée de X_t sachant l'observation passée, et on a pour toute fonction φ de $\underline{\underline{X}}$ dans $I\!R$ bornée

$$E[\varphi(X_t)/Y^t = y^t] = \frac{< q_t, \varphi >}{< q_t, 1 >}$$

où nous notons $< .,. >$ le produit de dualité $L^1_\mu(\underline{\underline{X}}) - L^\infty_\mu(\underline{\underline{X}})$, i.e.

$$< f, g > = \int f(\xi) g(\xi) d\mu(\xi)$$

Ce résultat est une conséquence immédiate du théorème précédent. Notons que si les densités de probabilités q_t évoluent dans $L^2_\mu(\underline{\underline{X}})$ alors on peut écrire

$$q_t = K_t(q_{t-1})$$

où K_t est l'opérateur de Hilbert-Schmidt associé au noyau k_t défini par

$$k_t(x,\xi) = \Phi(x,y_t)\Pi(\xi,x) = f_{(X_t,Y_t)}^{X_{t-1}=\xi}(x,y_t)$$

et donc

$$q_t(x) = \int_{\underline{\underline{X}}} k_t(x,\xi)q_{t-1}(\xi)d\mu(\xi)$$

2. LE CAS ETAT DISCRET

L'étude particulière de ces cas est motivée par le fait que l'existence de filtres finis est beaucoup moins rare que dans le cas général. De plus, il n'est pas rare que l'on soit ammené en pratique à considérer de tels systèmes. Enfin, ces cas permettront d'obtenir des approximations de filtres finis dans le cas plus général d'espace d'état continu.

Supposons que l'espace d'état est fini, soit par exemple

$$\underline{\underline{X}} = \{0,1,...,p-1\}$$

alors la formule qui donne la densité conditionnelle de l'état sachant l'observation devient

$$p_t(x) = \frac{1}{D_t}\Phi(x,y_t)\sum_{\xi=0}^{n-1}\Pi(\xi,x)p_{t-1}(\xi)\mu(\xi)$$

ce qui se traduit plus simplement en termes de probabilités, en notant cette fois $p_t(x) = P\{X_t = x/Y^t = y^t\}$ (ce qui est obtenu en prenant comme loi μ la loi uniforme sur $\underline{\underline{X}}$)

$$p_t(x) = \frac{1}{D_t}\Phi(x,y_t)\sum_{\xi=0}^{n-1}P\{X_t = x/X_{t-1} = \xi, Y_{t-1} = y_{t-1}\}p_{t-1}(\xi)$$

et il suffit de connaître $(p_{t-1}(\xi))_{\xi=0,..,n-1}$ qui est un point d'un espace à n dimensions pour calculer p_t en tout point. On a donc toujours, dans ce cas, un filtre fini.

Nous allons généraliser ce résultat par un lemme en reprenant l'hypothèse *(c1)* à la place de *(c'1)* i.e. le processus $(X_t)_{t\in I\!N}$ est supposé indépendant du processus $(Y_t)_{t\in I\!N}$.

Lemme 2.1

Supposons qu'il existe n fonctions $h_1,...,h_n$ et n fonctions $g_1,...,g_n$ telles que pour tout couple $(x,\xi)\in\underline{\underline{X}}\times\underline{\underline{X}}$, on ait

$$f_{X_t}^{X_{t-1}=\xi}(x) = \sum_{i=1}^{n} g_i(x)h_i(\xi)$$

alors le système admet un filtre fini.

Démonstration

Appliquons le théorème 1.1

$$\begin{aligned}
p_t(x) &= \frac{1}{D_t}\Phi(x, y_t) \int_{\underline{\underline{X}}} f_{X_t}^{X_{t-1}=\xi}(x) p_{t-1}(\xi) d\mu(\xi) \\
&= \frac{1}{D_t}\Phi(x, y_t) \sum_{i=1}^{n} g_i(x) \int_{\underline{\underline{X}}} h_i(\xi) p_{t-1}(\xi) d\mu(\xi) \\
&= \frac{1}{D_t}\Phi(x, y_t) \sum_{i=1}^{n} C_{t-1}^i g_i(x)
\end{aligned}$$

Les n coefficients $(C_{t-1}^i)_{i=1,\dots,n}$ sont obtenus récursivement et permettent de connaître la loi de X_t conditionnellement à Y^t, ce qui constitue bien un filtre fini.

Concrétement, ce lemme s'applique de la façon suivante

Proposition 2.2

Supposons qu'il existe un processus $(A_t)_{t \in I\!N}$ à valeurs dans un ensemble fini $\underline{\underline{A}}$ et tel que pour tout entier naturel t, on ait

$$X_{t+1} \perp X_t / A_t$$

Alors il existe un filtre de dimension finie.

Cette proposition découle du lemme précédent, en remarquant que

$$f_{X_t}^{X_{t-1}=\xi}(x) = \sum_{a \in A} f_{X_t}^{A_{t-1}=a}(x) P\{A_{t-1} = a / X_{t-1} = \xi\}$$

On peut interpréter ce résultat de la façon suivante. Les coefficients du lemme s'écrivent

$$\begin{aligned}
C_t^i &= \int_{\underline{\underline{X}}} P\{A_{t-1} = a / X_{t-1} = \xi\} p_{t-1}(\xi) d\mu(\xi) \\
&= P\{A_{t-1} = a / Y^{t-1}\}
\end{aligned}$$

donc on a simplement

$$p_t(x) = \frac{1}{D_t}\Phi(x, y_t) \sum_{a \in A} f_{X_t}^{A_{t-1}=a}(x) P\{A_{t-1} = a / Y^{t-1}\}$$

avec la formule de récurrence

$$P\{A_t = a / Y^t\} = \int_{\underline{\underline{X}}} P\{A_t = a / X_t = \xi\} p_t(\xi) d\mu(\xi)$$

CHAPITRE II

3. LE CAS LINÉAIRE GAUSSIEN

C'est le cas classique, le premier ayant été traité avec succès, d'existence de filtre fini. C'est aussi l'un des seuls filtres finis qui est couramment utilisé en pratique. Il y a plusieurs façon de démontrer qu'un système linéaire gaussien admet un filtre fini. Indiquons succintement une de ces méthodes, qui utilise l'équation de Zakaï que l'on vient d'établir.

Considérons donc le système suivant

$$\begin{cases} X_{t+1} = A_t X_t + B_t W_t \\ Y_t = C_t X_t + D_t V_t \end{cases}$$

où $(W_t)_{t\in\mathbb{N}}$ et $(V_t)_{t\in\mathbb{N}}$ sont des vecteurs gaussiens centrés et indépendants dans respectivement $\mathbb{R}^r$ et $\mathbb{R}^s$. On supposera que X_0 suit une loi gaussienne indépendante, X_t évoluant dans $\mathbb{R}^p$, Y_t dans $\mathbb{R}^q$. Les matrices A_t, B_t, C_t, et D_t sont de dimensions respectives pp, pr, qp, et qs. On supposera que les matrices C_t sont toutes de rang q, cette hypothèse n'étant pas trop restrictive, puisqu'elle consiste simplement à interdire toute redondance sur l'observation.

Considérons deux lemmes, préparant le résultat final.

Lemme 3.1

Soient f et g deux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$, densités respectivement des lois $N(a, A)$ et $N(b, B)$ par rapport à la mesure de Lebesgue. Alors, le produit renormalisé de f et de g, c'est à dire la fonction

$$h(x) = \frac{\int_{\mathbb{R}} f(\xi)g(\xi)d\xi}{f(x)g(x)}$$

est la densité de la loi

$$\mathcal{N}(S(A^{-1}a + B^{-1}b), S)$$

où

$$S = (A^{-1} + B^{-1})^{-1}$$

Lemme 3.2 (dit lemme d'inversion matricielle)

Considérons la matrice de la forme $P^{-1} + C^T Q^{-1} C$. Si la matrice $Q + CPC^T$ est inversible, alors $P^{-1} + C^T Q^{-1} C$ est inversible et son inverse est donnée par $P - PC^T(CPC^T + Q)^{-1}CP$.

Ces deux lemmes se vérifient par un simple calcul. Il en est de même du résultat suivant, qui prouve l'existence d'un filtre fini dans le cas linéaire gaussien, et qui donne sa forme, particulièrement simple.

Théorème 3.3 (Kalman)

$$\mathcal{L}(X_{t+1}/Y^{t+1}) = \mathcal{N}(m_{t+1}, \Sigma_{t+1})$$

où les paramètres m_{t+1} et Σ_{t+1} sont obtenus à partir des équations récursives suivantes

$$\begin{aligned} m_{t+1} &= A_t m_t + K_{t+1}(y_{t+1} - C_{t+1}A_t m_t) \\ \Sigma_{t+1} &= (I - K_{t+1}C_{t+1})P_t \\ P_t &= A_t \Sigma_t A_t^T + B_t B_t^T \\ K_{t+1} &= P_t C_{t+1}^T (C_{t+1} P_t C_{t+1}^T + D_{t+1} D_{t+1}^T)^{-1} \end{aligned}$$

La démonstration consiste simplement à écrire l'équation de Zakaï et à appliquer succéssivement les deux lemmes. Notons que l'on démontrera plus loin l'existence d'un filtre de Kalman pour des systèmes linéaires beaucoup plus généraux que celui qui vient d'être présenté.

4. FILTRES FINIS ET LOIS EXPONENTIELLES

Nous allons montrer, reprenant des résultats de Sawitzki *[1981]* et de Runggaldier *[1989]* l'étroit lien qui existe entre les filtres finis et les lois exponentielles en temps discret.

Définition 4.1

Soit (Ω, A, μ) un espace de probabilité. Soit T une fonction mesurable sur (Ω, A, μ) à valeurs dans un espace de Hilbert E sur $\mathbb{R}$. Considérons

$$\Theta = \left\{ \theta \in E; \psi(\theta) = \log(\int_\Omega \exp(< \theta, T >) d\mu) < \infty \right\}$$

Alors, la vraissemblance $L(\theta) = \exp(-\psi(\theta) + < \theta, T >)$ pour $\theta \in \Theta$ défini le modèle exponentiel $(\Omega, A, (P_\theta)_{\theta \in \Theta})$ par

$$\frac{dP_\theta}{d\mu}(\omega) = L(\theta, \omega) = \exp(-\psi(\theta) + < \theta, T(\omega) >)$$

On supposera dans la suite Θ d'intérieur non vide, et on prendra en fait $E = \mathbb{R}^n$.

Nous allons éxaminer un cas particulièrement intéressent de systèmes admettant un filtre de dimension finie. Ce cas a été traité par Levine et Pignié *[1981]* dans le cas gaussien. Nous proposons une extension de leur résultat, en simplifiant considérablement leur démonstration.

On se place donc dans le cas d'un système à évolution déterministe, c'est à dire qu'il existe une fonction mesurable f de X dans X telle que $\forall t \in \mathbb{N}$

$$X_{t+1} = f(X_t)$$

Ainsi, seul l'état initial du système est aléatoire. On rencontre ce genre de systèmes dans les problèmes de détection de phase avec amplitude aléatoire mais constante.

Supposons que l'observation de l'état obéisse à une loi de type exponentiel, c'est à dire qu'il existe une application mesurable $h : X \to I\!R^n$ telle que

$$f_{Y_t}^{X_t=x}(y) = \exp(-\psi(x) + < h(x), T(y) >)$$

le rôle de T et de ψ ayant été précisé plus haut. On a alors le résultat suivant

Théorème 4.2

Soit H l'espace vectoriel fonctionnel engendré par l'ensemble de fonctions

$$H = \text{Span}\{h \circ f^k ; k \in I\!N\}$$

(où $f^0 = \text{Id}$ et $f^{k+1} = f \circ f^k$).

Si $\dim(H) = d < \infty$, alors le système admet un filtre fini.

Nous verrons plus loin que cette condition est en fait nécessaire et suffisante, sous certaines conditions de régularité. Avant de démontrer ce théorème, énonçons le corollaire suivant, qui est la version gaussienne du théorème

Corollaire 4.3

Soit le système

$$\begin{cases} X_{t+1} = f(X_t) \\ Y_t = h(X_t) + \eta(X_t)V_t \end{cases}$$

Ce système admet un filtre de dimension finie dès que

$$\dim(\text{Span}(H)) < \infty$$

Démonstration du théorème

Le seul élément aléatoire du processus est X_0. Calculons donc la vraissemblance de X_0 à l'instant t.

$$L(x_0) = \exp\left\{ -\sum_{s=0}^{t} \psi(x_s) + \sum_{s=0}^{t} < h(x_s), T(y_s) > \right\}$$

Soit $g_1, ..., g_d$ une base de H: $\forall k \in I\!N \qquad h \circ f^k = \sum_{i=1}^{d} \lambda_k^i g_i$ donc

$$\begin{aligned} L(x_0) &= \exp\left\{ -\sum_{s=0}^{t} \psi(x_s) + \sum_{s=0}^{t} < \sum_{i=1}^{d} \lambda_s^i g_i(x_0), T(y_s) > \right\} \\ &= \exp\left\{ -\sum_{s=0}^{t} \psi(x_s) + \sum_{i=1}^{d} < g_i(x_0), \sum_{s=0}^{t} \lambda_s^i T(y_s) > \right\} \end{aligned}$$

et le théorème de factorisation nous indique que

$$z_t = (\sum_{s=0}^{t} \lambda_s^i T(y_s))_{i=1\ldots d}$$

est un résumé exhaustif, et il est de dimension dn. On constate que c'est bien un filtre (linéaire) puisque l'on a simplement

$$\begin{aligned} z_t &= \left(\sum_{s=0}^{t-1} \lambda_s^i T(y_s)\right)_{i=1\ldots d} + (\lambda_t^i T(y_t))_{i=1\ldots d} \\ &= z_{t-1} + (\lambda_t^i T(y_t))_{i=1\ldots d} \end{aligned}$$

Le théorème que nous allons énoncer, dù à Ferrante et Runggaldier [2] est une réciproque du théorème 4.2, lorsque la dimension du filtre est 1. Cette condition très restrictive pourra être supprimée, au prix de conditions supplémentaires sur la régularité des densités mises en jeu. Le théorème est cependant intéressant, dans la mesure où c'est l'un des seuls théorèmes qui donne une condition nécessaire et suffisante d'existence de filtre de dimension finie pour une classe très large de systémes.

Théorème 4.4

Sous les hypothèses habituelles, supposons en outre que $P_{X_t}^{Y^{t-1}}$ et $P_{Y_t}^{X_t}$ aient des densités continues strictement positives, et ceci telle que les applications

$$y \to \log \frac{f_{Y_t}^{X_t=x}(y)}{f_{Y_t}^{X_t=\xi}(y)}$$

et

$$y^{t-1} \to \log \frac{f_{X_t}^{Y^{t-1}=y^{t-1}}(x)}{f_{X_t}^{Y^{t-1}=y^{t-1}}(\xi)}$$

ne soient pas constantes pour au moins un couple de points (x, ξ) de X^2. S'il existe un filtre de dimension 1 tel que les applications $\varphi_t : Z \times Y \to Z \subseteq I\!R$ soient séparément continues, alors $P_{X_t}^{Y^{t-1}}$ et $P_{Y_t}^{X_t}$ sont des familles de probabilités exponentielles (unidimensionnelles).

REFERENCES

[1] **Di Masi G.B., Runggaldier W.J.**"On measure transformations for combined filtering and parameter etimation in discrete time"*Syst. and Cont. Letters*

[2] **Ferrante M., Runggaldier W.J.**"On necessary conditions for the existence of finite dimensional filters in discrete-time"*Syst. and Cont. Letters 14 (1990) p63-69*

CHAPITRE II

[3] **Jazwinski A. H.**"Stochastic Filtering Theory"*Academic Press*

[4] **Sawitzki G.** "Exact filtering in exponential families: discrete-time"*Lect. Notes in Cont. and Info. Sci. 16 (1979) p554-558*

[5] **Sawitzki G.** "Finite dimensional filter systems in discrete-time"*Stochastics 5 (1981) p107-114*

3

CHAPITRE III

A SUFFICIENT CONDITION FOR THE EXISTENCE OF FINITE DIMENSIONAL FILTERS FOR DISCRETE-TIME NONLINEAR SYSTEMS

ABSTRACT

The main difficulty in finding a nonlinear control is that a feedback must be built which depends on the state. It is then necessary to estimate the state knowing the observation. The resulting filtering problem is such that a way must be found to resume observations in a finite number of parameters which are been calculated recursively. When one can compute such parameters, the system is said to admit a finite dimensional filter (F.D.F.). For example, any linear system has a F.D.F.

The technique used here is to steer our nonlinear model into a linear one which is parametrised by observations. For linear systems, one uses the Kalman filter to estimate the mean and variance of the state. This filter is optimal if the initial condition is gaussian.

But generally , for physical reasons, it is know that the state can't be outside of a given subset of the state space. So the initial condition is not gaussian. In this work, we show that it is possible to construct an optimal filter which takes into account the previous information.

We show that a F.D.F. exists for any linear system in discrete-time parameterised by the observation, even if the initial condition is not Gaussian. We use this result to give a necessary condition for the existence of a F.D.F. for a large class of nonlinear systems.

This approach has been used also in continuous time, with the same results.

INTRODUCTION

CHAPITRE III

The aim of nonlinear filtering is to control the state of a hidden and stochastic process starting from a partial observation with noise. This follow-up should be achieved through data processing in order to be able to provide all the information available on the process state at any time, taking all the previous observations as a basis. Theoretically, remembering the observation in its rough form should allow this goal to be achieved but the amount of data to remember increases with time until it saturates the computer memory. Besides, too many observations make the specific information on the real state of the process difficult to interpret. As a consequence, it is essential to structure these observations and summarize them in a finite number of parameters.

However, this data reduction should not cause any loss in the useful data on the present state. What is sought after is a sufficient summary of the observation. In order that calculations could be made during the process, these parameters should be updated at each new observation. The aim is then to try and summarize the observation in a finite number of recursively-computable parameters. When this operation is feasible for a given problem, one says that the system admits a finite dimensional filter (FDF). The existence of an FDF is essential when aiming at following the process very carefully, since, in that case only, calculations will be implemented in a computer. The FDF then provides a means of structuring the information, which may help to better understand the nature of the process. Thus, for a Gaussian process, the mean and the variance are sufficient parameters which better characterise the process than a non-structured series of observed values.

The aim is to characterise the systems which admit FDFs and, in that case, to build the filter explicitly. One of the main classes of systems admitting finite filters includes the Gaussian linear filters, and the currently used filter is the most famous Kalman-Bucy (1961) filter. Linear systems can be used in practice only when the process is in steady state. This linearisation - of a priori nonlinear systems - is carried out precisely because linear systems admit FDFs.

During this linearisation, the following hypothesis is essential for the application of the Kalman-Bucy filter: the initial law of the state process should be Gaussian. This hypothesis is often transgressed in practice as, for physical reasons, the state is often limited or restricted to a given area of the state space. The a priori knowledge on the state process is, in practice, neglected. It is shown in this paper that conditionally Gaussian systems - which form a larger class than nonlinear systems - admit an FDF even when the initial law of the process is not Gaussian. This leads us to optimise filtering by attaching more importance to the physical constraints placed upon the state variables. It is also a preliminary result for the study on nonlinear systems.

It is well-known that such systems do not often admit FDFs. However, an example mentioned in this paper reveals the existence of a class of nonlinear systems which admit finite dimensional filters, the latter being built explicitly. The conditions required for this system to belong to this class can be checked easily. This result can then be used without difficulty in practice. The technique employed is immersion, i.e. injective mapping. However, while the usual objective is to immerse a nonlinear system into a linear one, we shall try to convert the nonlinear system into a conditionally Gaussian nonlinear system with a non-Gaussian initial condition. We shall then take the first result obtained as a basis for our conclusions.

Characterising the nonlinear systems which admit FDFs, or exhibiting classes of such systems enables us to consider typically nonlinear behaviours and to filter such systems precisely. It should be noted that the class of systems which admit finite filters, although restricted, contains the linear systems already in general use. Besides, it can be remarked that the technique employed is constructive and allows the filtering algorithm to be written explicitly. The result may then be used for real systems, all the more since the dimension of the obtained filter is not disproportionate to that of the system state, while the result is optimal.

We will consider the following non-linear system

$$(\Sigma) \qquad \begin{cases} x_{t+1} = x_t + f(x_t, y_t) + g(x_t, y_t) w_t \\ y_{t+1} = h(x_t, y_t) + v_t \end{cases}$$

where $(x_t)_{t \in I\!N}$ is a $I\!R^p$-valued process, $(y_t)_{t \in I\!N}$ is a $I\!R^q$-valued process, and (W_t) and (V_t) are sequences of independent unit normal variables.

More generally, let $(x_t)_{t \in I\!N}$ a $I\!R^p$-valued process and for any $t \in I\!N$, let us define $y_{t+1} = h(x_t) + v_t$ where h is an application from $I\!R^p$ to $I\!R^q$, and where (v_t) is a sequence of independent unit normal variables.

One of the aims of the nonlinear filtering is to compute the conditional measure of the state x_t when one observes $(y_s; 0 \leq s \leq t)$. The only way to compute easily this probability measure is to find a FDF. We say that a system admits a FDF if and only if there exists some positive integer n and

-a function $\varphi_0 : I\!R^q \to I\!R^n$ and a sequence of functions $\varphi_t : I\!R^n \times I\!R^q \to I\!R^n \qquad t \geq 1$,

-a sequence of functions $\psi_t : I\!R^n \times \mathcal{B}(I\!R^p) \to I\!R \qquad t \geq 0$, such that for any Borel set B in $I\!R^p$,

$$P\{x_t \in B / \sigma(y_s; s \leq t)\} = \psi_t(\xi_t, B)$$

where ξ_t is recursively computable by

$$\xi_0 = \varphi_0(y_0)$$

and

$$\forall s \geq 1 \qquad \xi_s = \varphi_s(\xi_{s-1}, y_s)$$

This definition means that the conditional measure of x_t is a function of a boundary number of recursively computable parameters.

CHAPITRE III

In second part of our paper, we study the following particular system

$$(\Sigma) \qquad \begin{cases} x_{t+1} = \alpha(y_t) + A(y_t)x_t + B(y_t)w_t + D(y_t)v_t \\ y_{t+1} = Cx_t + v_t \end{cases}$$

We know that if the initial condition x_0 is Gaussian, the distribution of x_t knowing $(y_1, \cdots, y_t)$ is given by the Kalman filter which is a FDF. In most cases, the initial condition x_0 is not Gaussian, this implies that the Kalman filter cannot be used directly to solve the problem. We show that, in this case, a FDF exists.

In the third part, we study the immersion of (Σ) into a certain class of systems $(\bar{\Sigma})$. We will give necessary and sufficient conditions which indicate when the system

$$\begin{cases} x_{t+1} = x_t + f(x_t, y_t) + g(x_t, y_t)w_t \\ y_{t+1} = h(x_t, y_t) + v_t \end{cases}$$

can be immersed into $(\bar{\Sigma})$.

We will establish in the fourth part some necessary conditions, easily verified, under which a FDF exists. We will construct this FDF explicitly.

1. LINEAR SYSTEMS WITH NON GAUSSIAN INITIAL CONDITIONS

In this section, we will show that the following system

$$\begin{cases} x_{t+1} = A_t(y_t)x_t + B_t(y_t)w_t + D_t(y_t)v_t \\ y_{t+1} = C_t(y_t)x_t + v_t \end{cases}$$

with x_0 non Gaussian admits a FDF, and we will use Girsanov's theorem to transform this system in a conditionally Gaussian system. The proof is very similar to the proof of the theorem without the correlation between the state and the observation [1,8,9].

Lemme 1.1

Soit X, Y un vecteur gaussien de moyenne m_X, m_Y et de variance

$$\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}$$

On a

$$E[X/Y] = m_X + \alpha(Y - m_Y)$$
$$\Gamma(X/Y) = \Sigma_{XX} - \alpha\Sigma_{YY}\alpha^T$$

où α est tel que $\Sigma_{XY} = \alpha\Sigma_{YY}$ i.e. $\alpha = \Sigma_{XY}\Sigma_{YY}^{-}$ en notant Σ_{YY}^{-} une pseudo-inverse de Σ_{YY}.

Démonstration

Nous allons simplement identifier deux fonctions caractéristiques.

$$(1) \quad E\left[\exp\left\{i\left(u^T X + v^T Y\right)\right\}\right] = \exp\left\{-\frac{1}{2}(u^T, v^T)\Sigma\begin{pmatrix} u \\ v \end{pmatrix} + i(u^T m_X + v^T m_Y)\right\}$$

d'une part, et d'autre part

$$(2) \qquad E\left[\exp\left\{i\left(u^T X + v^T Y\right)\right\}\right] = E\left[\exp\left\{iv^T Y\right\} E\left[\exp\left\{iu^T X\right\}/\sigma(Y)\right]\right]$$

Montrons que l'on peut écrire

$$E\left[\exp\left\{iu^T X\right\}/\sigma(Y)\right] = \exp\left\{-\frac{1}{2}u^T \Sigma_{X/Y} u + iu^T m_{X/Y}\right\}$$

avec

$$m_{X/Y} = m_X + \alpha(Y - m_Y)$$

et

$$\Sigma_{X/Y} \text{ indépendante de Y}$$

ce qui montrera le caractère gaussien de la loi de X sachant Y. Sous cette hypothèse, (2) se réécrit

$$\begin{aligned}
(2) &= \exp\left\{-\frac{1}{2}u^T \Sigma_{X/Y} u + iu^T m_X - iu^T \alpha m_Y\right\} E\left[\exp\left\{i(v^T + u^T\alpha)Y\right\}\right] \\
&= \exp\left\{i(u^T m_X + v^T m_Y) - \frac{1}{2}\left[u^T \Sigma_{X/Y} u + (v^T + u^T\alpha)\Sigma_{YY}(v + \alpha^T u)\right]\right\}
\end{aligned}$$

En identifiant (1) et (2), il reste

$$\begin{cases} \Sigma_{XY} = \alpha \Sigma_{YY} \\ \Sigma_{X/Y} = \Sigma_{XX} - \alpha \Sigma_{YY} \alpha^T \end{cases}$$

□

Proposition 1.2

Let $(W_t)_{t \in \mathbb{N}}$ and $(V_t)_{t \in \mathbb{N}}$ be two independent standard -i.e. the covariance matrix is identity- Gaussian white noises, $\mathbb{R}^r$-valued and $\mathbb{R}^q$-valued respectively. Let Σ be the conditionally Gaussian system

$$\begin{cases} x_{t+1} = \alpha_t(y^{t+1}) + A_t(y^t)x_t + B_t(y^t)w_t \\ y_{t+1} = C_t(y^t)x_t + D_t(y^t)v_t \end{cases}$$

where we have noted $y^t = (y_1, \ldots, y_t)$. We will denote the transpose of the matrix M by M^T. Suppose $\mathcal{L}(x_0) = \mathcal{N}(m_0, \Sigma_0)$, then $\mathcal{L}(x_t/y^t) = \mathcal{N}(m_t, \Sigma_t)$ with

$$\begin{aligned}
m_{t+1} &= \alpha_t(y^{t+1}) + A_t(y^t)m_t + K_t(y_t - C_t(y^t)m_t) \\
\Sigma_{t+1} &= A_t(y^t)\Sigma_t A_t^T(y^t) + B_t(y^t)B_t^T(y^t) - K_t P_t K_t^T
\end{aligned}$$

and

$$K_t = A_t(y^t)\Sigma_t C_t^T(y^t)P_t^-$$
$$P_t = C_t(y^t)\Sigma_t C_t^T(y^t) + D_t(y^t)D_t^T(y^t)$$

where P_t^- satisfy $\Sigma_t C_t^T(y^t)P_t^- P_t = \Sigma_t C_t^T(y^t)$, i.e. P_t^- is a pseudo-inverse of P_t.

Demonstration

To show this result, we will suppose that, conditionnally to y^t, the random variable (x_t, y_t) is gaussian. The second equation of the system show that (x_t, y_{t+1}) knowing y^t is gaussian with mean

$$(m_t, C_t(y^t)m_t)$$

and covariance

$$\begin{aligned}
\Gamma &= E\left[\begin{pmatrix} X_t - m_t \\ Y_{t+1} - C_t(y^t)m_t \end{pmatrix} \begin{pmatrix} X_t - m_t \\ Y_{t+1} - C_t(y^t)m_t \end{pmatrix}^T / y^t\right] \\
&= E\left[\begin{pmatrix} \Sigma_t & \Sigma_t C_t^T(y^t) \\ C_t(y^t)\Sigma_t & C_t(y^t)\Sigma_t C_t^T(y^t) + R_t(y^t) \end{pmatrix}\right]
\end{aligned}$$

where $R_t(y^t) = D_t(y^t)D_t^T(y^t)$. Set also $P_t = C_t(y^t)\Sigma_t C_t^T(y^t) + R_t(y^t)$. We will use the lemma 1.1 which give us the law of a gaussian vector knowing some of their coordinates

$$\begin{aligned}
E\left[x_t/y^{t+1}\right] &= m_t + \Sigma_t C_t^T(y^t)P_t^-(y_{t+1} - C_t(y^t)m_t) \\
\Gamma\left[x_t/y^{t+1}\right] &= \Sigma - \Sigma_t C_t^T(y^t)P_t^- P_t P_t^- C_t(y^t)\Sigma_t
\end{aligned}$$

Therefore x_{t+1} is the sum of two gaussian random variables and a function of y^{t+1}, and

$$\begin{aligned}
m_{t+1} &= \alpha_t(y^{t+1}) + A_t(y^t)\left(m_t + \Sigma_t C_t^T(y^t)P_t^-(y_{t+1} - C_t(y^t)m_t)\right) \\
\Sigma_{t+1} &= A_t(y^t)\left(\Sigma_t - \Sigma_t C_t^T(y^t)P_t^- P_t P_t^- C_t(y^t)\Sigma_t\right)A_t^T(y^t) + B_t(y^t)B_t^T(y^t)
\end{aligned}$$

and so, if we note $K_t = A_t(y^t)\Sigma_t C_t^T(y^t)P_t^-$ then we have

$$\begin{aligned}
m_{t+1} &= \alpha_t(y^{t+1}) + A_t(y^t)m_t + K_t\left(y_{t+1} - C_t(y^t)m_t\right) \\
\Sigma_{t+1} &= A_t(y^t)\Sigma_t A_t^T(y^t) - K_t P_t K_t^T + B_t(y^t)B_t^T(y^t)
\end{aligned}$$

□

We will need also the two following lemma

Lemma 1.2

Let $(\Omega, \mathcal{F})$ a measured space with a probability P, X a random variable on this space and $\mathcal{A}$ a σ-algebra s.t. $\mathcal{A} \in \mathcal{F}$. Let Q be a probability defined by its density with respect to P, i.e. $dQ = L.dP$.Then

$$E_Q[X/\mathcal{A}] = \frac{E_P[X.L/\mathcal{A}]}{E_P[L/\mathcal{A}]} \qquad P-\text{a.s.}$$

Demonstration

Let

$$\Theta = \{E_P[L/\mathcal{A}] = 0\}$$

Then Θ is A-measurable and

$$\begin{aligned} Q(\Theta) &= E_Q[1_\Theta] = E_P[1_\Theta.L] \\ &= E_P[1_\Theta E_P[L/\mathcal{A}]] = 0 \end{aligned}$$

Therefore $E_Q[X/\mathcal{A}]$ can take any value on Θ and we can suppose for our proof without lost of generality that $A \in \Omega - \Theta$

$$\begin{aligned} E_Q\left[1_A \frac{E_P[X.L/\mathcal{A}]}{E_P[L/\mathcal{A}]}\right] &= E_P\left[1_A.L\frac{E_P[X.L/\mathcal{A}]}{E_P[L/\mathcal{A}]}\right] \\ &= E_P[E_P[1_A.X.L/\mathcal{A}]] \\ &= E_P[1_A.X.L] = E_Q[1_A.X] \end{aligned}$$

□

Lemma 1.3

Let $(B_t)_{t\in\mathbb{N}}$ be a Gaussian white noise process, $\mathbb{R}^n$-valued, adapted to the filtration $(\mathcal{A}_t)_{t\in\mathbb{N}}$, with covariance matrix R_t, and $(H_t)_{t\in\mathbb{N}}$ a $(\mathcal{A}_t)$-predictable(*) $\mathbb{R}^n$-valued process. Then the process $(G_t)_{t\in\mathbb{N}}$ defined by $G_t = B_t - H_t$ is a Gaussian white noise process with the same covariance matrix R_t, adapted to the same filtration $(\mathcal{A}_t)_{t\in\mathbb{N}}$, under a probability measure Q, defined by its marginals

$$dQ_t = L_t dP = \prod_{s=0}^{t} \exp\left\{H_s^T R_s^{-1} B_s - \frac{1}{2} H_s^T R_s^{-1} H_s\right\} dP$$

The proof of this lemma consists in verifying that L_t is an $(\mathcal{A}_t)$-martingale, and then in using Kolmogorov's theorem to prove the existence of Q. More precisely, let's show that

$$L_t = \prod_{s=0}^{t} \exp\left\{H_s^T R_s^{-1} B_s - \frac{1}{2} H_s^T R_s^{-1} H_s\right\}$$

is an $\mathcal{A}_t$-martingale

$$\begin{aligned} E_P[L_t/\mathcal{A}_{t-1}] &= E_P\left[\prod_{s=0}^{t} \exp\left\{H_s^T R_s^{-1} B_s - \frac{1}{2} H_s^T R_s^{-1} H_s\right\} /\mathcal{A}_{t-1}\right] \\ &= L_{t-1} E_P\left[\exp\{\frac{1}{2} B_t^T R_t^{-1} B_t - \frac{1}{2}(B_t - H_t)^T R_t^{-1}(B_t - H_t)\}/\mathcal{A}_{t-1}\right] \\ &= L_{t-1} \int \exp\{\frac{1}{2} b^T R_t^{-1} b\} \exp\{-\frac{1}{2}(b - H_t)^T R_t^{-1}(b - H_t)\} \\ &\qquad \frac{1}{\sqrt{(2\pi)^n.\det(R_t)}} \exp\{-\frac{1}{2} b^T R_t^{-1} b\} db \\ &= L_{t-1} \end{aligned}$$

(*) i.e. for any $t > 0$ $\quad H_t \in \mathcal{A}_{t-1}$

which is true for any $t \in \mathbb{N}$, if we note $L_{-1} = 1$ and $\mathcal{A}_{-1}$ the trivial σ-algebra.

According to the definition of Q_t, it is clear that the compatibility condition of the Kolmogorov's theorem is satisfy. Therefore, there exist a probability Q defined on $\mathbb{R}^{\mathbb{N}}$ which admit Q_t as marginals. We must show that G_t is a gaussian white noise for Q. In order to do this, we will calculate its Fourier transform. At first, since L_t is a P- martingale and thanks to the lemma 1, we have for any functional f $\mathbb{C}$-valued

$$\begin{aligned}
E_Q\left[f(B_t)/\mathcal{A}_{t-1}\right] &= E_{Q_t}\left[f(B_t)/\mathcal{A}_{t-1}\right] \\
&= \frac{E_P\left[f(B_t).L_t/\mathcal{A}_{t-1}\right]}{E_P\left[L_t/\mathcal{A}_{t-1}\right]} \\
&= E_P\left[f(B_t)\exp\{H_t^T R_t^{-1} B_t - \frac{1}{2} H_t^T R_t^{-1} H_t/\mathcal{A}_{t-1}\right]
\end{aligned}$$

so for any complexe variable θ

$$\begin{aligned}
E_Q\left[\exp\{i\theta^T G_t\}\right] &= E_Q\left[E_{Q_t}\left[\exp\{i\theta^T(B_t - H_t)\}/\mathcal{A}_{t-1}\right]\right] \\
&= E_Q\left[\exp\{-i\theta^T H_t\} E_{Q_t}\left[\exp\{i\theta^T B_t\}/\mathcal{A}_{t-1}\right]\right]
\end{aligned}$$

but

$$\begin{aligned}
&E_{Q_t}\left[\exp\{i\theta^T B_t\}/\mathcal{A}_{t-1}\right] \\
&\quad = E_{Q_t}\left[\exp\{i\theta^T B_t\}\exp\{H_t^T R_t^{-1} B_t - \frac{1}{2} H_t^T R_t^{-1} H_t\}/\mathcal{A}_{t-1}\right] \\
&\quad = \exp\{-\frac{1}{2} H_t^T R_t^{-1} H_t\} E_{Q_t}\left[\exp\{i(\theta - iR_t^{-1} H_t) B_t\}/\mathcal{A}_{t-1}\right] \\
&\quad = \exp\{-\frac{1}{2} H_t^T R_t^{-1} H_t\}\exp\{-\frac{1}{2}(\theta - iR_t^{-1} H_t)^T R_t^{-1}(\theta - iR_t^{-1} H_t)\} \\
&\quad = \exp\{-\frac{1}{2}\theta^T R_t^{-1}\theta + i\theta^T H_t\}
\end{aligned}$$

and so

$$\begin{aligned}
E_Q\left[\exp\{i\theta^T G_t\}\right] &= E_Q\left[\exp\{-i\theta^T H_t\}\exp\{-\frac{1}{2}\theta^T R_t^{-1}\theta + i\theta^T H_t\}\right] \\
&= \exp\{-\frac{1}{2}\theta^T R_t^{-1}\theta\}
\end{aligned}$$

Theorem 1.4

Let Λ be the following system

$$(\Lambda) \qquad \begin{cases} x_{t+1} = A_t(y_t)x_t + B_t(y_t)w_t + \overline{D}_t(y_t)v_t \\ y_{t+1} = C_t(y_t)x_t + D_t(y_t)v_t \end{cases}$$

for $t \geq 0$ and suppose $\mathcal{L}(x_0) = \mu_0$. Suppose that $(v_t)_{t\geq 0}$ and $(w_t)_{t\geq 0}$ are two independent Gaussian white noises, independent of x_0, and suppose that for any t and for any y, $D_t(y)$ is an inversible matrix. Then Λ admits a filter with dimension $2p(p+1)$.

Demonstration

Let $\mathcal{F}_t$ be the σ-algebra generated by $\mathcal{F}^X$, $\mathcal{F}^Y$, $\mathcal{F}^V$ and $\mathcal{F}^W$. The system can be rewritten in the form

$$\begin{cases} x_{t+1} = & \overline{D}_t(y_t)y_{t+1} + \left(A_t(y_t) - \overline{D}_t(y_t)C_t(y_t)\right) x_t \\ & + B_t(y_t)w_t \\ y_{t+1} = & C_t(y_t)x_t + D_t(y_t)v_t \end{cases}$$

and we have

$$x_t = \Phi(t,0)x_0 + \sum_{s=0}^{t-1} \Phi(t,s+1)B_s(y_s)w_s + \sum_{s=0}^{t-1} \Phi(t,s+1)\overline{D}_s(y_s)y_{s+1}$$

where Φ is a function of the observation defined for every $(t,s), t \geq s$ by

$\phi(t,t) = I$

$\forall 0 \leq s < t$

$$\Phi(t,s) = \Phi(t,s+1)\left[A_s(y_s) - \overline{D}_s(y_s)C_s(y_s)\right]$$

It is easy to verify that

$$\forall s,\tau,t \geq 0 \qquad \Phi(t,s) = \Phi(t,\tau)\Phi(\tau,s)$$

Let

$$\xi_t^0 = \Phi(t,0)x_0$$

and

$$\xi_t^1 = \sum_{s=0}^{t-1} \Phi(t,s+1)B_s(y_s)w_s + \sum_{s=0}^{t-1} \Phi(t,s+1)\overline{D}_s(y_s)y_{s+1}$$

hence we have

$$\begin{aligned} \xi_{t+1}^0 &= \Phi(t+1,0)x_0 = \Phi(t+1,t)\Phi(t,0)x_0 \\ &= \left(A_t(y_t) - \overline{D}_t(y_t)C_t(y_t)\right)\xi_t^0 \end{aligned}$$

with $\xi_0^0 = x_0$ and

$$\begin{aligned} \xi_{t+1}^1 = & \sum_{s=0}^{t} \Phi(t+1,s+1)\left(B_s(y_s)w_s + \overline{D}_s(y_s)y_{s+1}\right) \\ = & \alpha_t(y_t,y_{t+1}) + \left(A_t(y_t) - \overline{D}_t(y_t)C_t(y_t)\right)\xi_t^1 + B_t(y_t)w_t \end{aligned}$$

with $\xi_0^1 = 0$ where $\alpha_t(y_t,y_{t+1}) = \overline{D}_t(y_t)y_{t+1}$.
Since $x_t = \xi_t^0 + \xi_t^1$, we have

$$y_{t+1} = C_t(y_t)(\xi_t^0 + \xi_t^1) + v_t = C_t(y_t)\xi_t^1 + D_t(y_t)\tilde{v}_t$$

where $D_t(y_t)\tilde{v}_t = D_t(y_t)v_t + C_t(y_t)\xi_t^0$. The process $C_t(y_t)\xi_t^0$ is $\mathcal{F}_t$-predictable, hence according to lemma 1.3, $D_t(y_t)\tilde{v}_t$ is a Gaussian white noise with covariance matrix

$$R_t(y_t) = D_t(y_t)D_t^T(y_t)$$

under Q where Q is defined by its projections

$$dQ_t = L_t dP$$

with

$$L_t = \exp\left\{-\sum_{s=0}^{t} x_0^T \Phi^T(s,0) C_s^T(y_s) R_t^{-1}(y_s) D_s(y_s) \tilde{v}_s\right.$$
$$\left.+\frac{1}{2}\sum_{s=0}^{t} x_0^T \Phi^T(s,0) C_s^T(y_s) R_t^{-1}(y_s) C_s(y_s) \Phi(s,0) x_0\right\}$$

Let

$$M_t = \sum_{s=0}^{t} \Phi^T(s,0) C_s^T(y_s) R_s^{-1}(y_s) C_s(y_s) \Phi(s,0)$$

and

$$\xi_t^2 = \sum_{s=0}^{t} \Phi^T(s,0) C_s^T(y_s) R_s^{-1}(y_s) D_s(y_s) \tilde{v}_s$$

we have

$$\xi_{t+1}^2 = \xi_t^2 + \Phi^T(t+1,0) C_{t+1}^T(y_{t+1}) R_{t+1}^{-1}(y_{t+1}) D_{t+1}(y_{t+1}) \tilde{v}_{t+1}$$

with $\xi_0^2 = 0$. Therefore (ξ_t^1, ξ_t^2) satisfies a conditionally Gaussian system with Gaussian initial condition, which has the same form as the system studied in proposition 1. Since $\Phi(t+1,0) = \left(A_t(y_t) - \overline{D}_t(y_t) C_t(y_t)\right) \Phi(t,0)$, it is sufficient to know y_t to calculate parameters m_t and Σ_t recursively, which are respectively the mean and the covariance matrices of (ξ_t^1, ξ_t^2). Since x_0 is independent of (ξ_t^1, ξ_t^2) under Q, the distribution of $(\xi_t^0, \xi_t^1, \xi_t^2)$ can be computed recursively as a function of a finite number of parameters. Indeed, a classic formula illustrates to us that we can compute $E_P[f(x_t)|y^t]$ for any integrable function f, by calculating

$$\frac{E_Q\left[f(\Phi(t,0)x_0 + \xi_t^1) \exp\left\{x_0^T \xi_t^2\right\} \exp\left\{-\frac{1}{2} x_0^T M_t x_0\right\} \mid y^t\right]}{E_Q\left[\exp\left\{x_0^T \xi_t^2\right\} \exp\left\{-\frac{1}{2} x_0^T M_t x_0\right\} \mid y^t\right]}$$

□

Remark 1.5

Consider the following system

$$\begin{cases} x_{t+1} = A_t(y_t) x_t + B_t(y_t) \\ y_{t+1} = C_t(y_t) x_t \end{cases}$$

We suppose that the output is a deterministic function of the state.

$$\begin{aligned} y_{t+1} &= C_t(y_t) A_{t-1}(y_{t-1}) x_{t-1} + B_{t-1}(y_{t-1}) w_{t-1} \\ &= C_t(y_t) A_{t-1}(y_{t-1}) A_{t-2}(y_{t-2}) x_{t-2} \\ &\quad + C_t(y_t) A_{t-1}(y_{t-1}) B_{t-2}(y_{t-2}) w_{t-2} \\ &= C_t(y_t) A_{t-1}(y_{t-1}) \cdots A_{t-r}(y_{t-r}) x_{t-r} \\ &\quad + C_t(y_t) A_{t-1}(y_{t-1}) \cdots A_{t-r+1}(y_{t-r+1}) B_{t-r}(y_{t-r}) w_{t-r} \end{aligned}$$

Set $\overline{y}_t = y_{t+r}$ therefore $\overline{y}_{t-r} = y_t$. We can write

$$\begin{cases} x_{t+1} = A_t(\overline{y}_{t-r})x_t + B_t(\overline{y}_{t-r}) \\ \overline{y}_{t+1} = \overline{C}_t(\overline{y}_t, \ldots, \overline{y}_{t-r})x_t + \overline{D}_t(\overline{y}_t, \ldots, \overline{y}_{t-r})w_t \end{cases}$$

Let v_t be the gaussian process adapted to the filtration $\mathcal{F}_Y^t$) and defined by

$$v_t = \overline{D}_t(\overline{y}_t, \ldots, \overline{y}_{t-r})w_t$$

Let $\overline{w}_t$ be the projection of w_t over the space which is orthogonal to v_t, that is to say

$$\begin{aligned} \overline{w}_t &= \mathrm{Proj}^{\perp}_{v_t^{\perp}}(w_t) \\ &= (I - D_t^T(D_tD_t^T)^{-1})D_tw_t \end{aligned}$$

then the system can be written in the following form

$$\begin{cases} x_{t+1} = (A_t - B_tQ_t\overline{C}_t)x_t + B_tQ_t\overline{y}_{t+1} + B_t\overline{w}_t \\ \overline{y}_{t+1} = \overline{C}_tx_tv_t \end{cases}$$

where $Q_t = D_t^T(D_tD_t^T)^{-1}$ and this system satisfies the hypothesis of the previous theorem.

2. FILTRE FINI PAR IMMERSION DANS UN SYSTEME LINEAIRE

La notion d'immersion à déjà donné, surtout en automatique non linéaire, de nombreux résultats. Une technique d'immersions des systèmes déterministes en temps discret à été élaborée par S. Monaco et D. Normand Cyrot. C'est cette technique que nous allons utiliser pour trouver des conditions nécessaires et suffisantes d'existence d'immersions dans des systèmes Gaussien ou conditionnellement Gaussien. Nous allons donc définir l'exponentielle tensorielle.

Soit f une fonction analytique de $\mathbb{R}^p$ dans $\mathbb{R}^p$. On note L_f l'opérateur différentiel du premier ordre défini par:

$$L_f = \sum_{i=1}^{p} f_i \frac{\partial}{\partial x_i}$$

Notons que l'application qui à f associe L_f est linéaire. Soit g une autre fonction analytique de $\mathbb{R}^p$ dans $\mathbb{R}^p$. On a, bien sûr

$$L_f \circ L_g = \sum_{i,j=1}^{p} f_i g_j \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i,j=1}^{p} f_i \frac{\partial}{\partial x_i} g_j \frac{\partial}{\partial x_j}$$

Introduisons sur ces opérateurs une autre opération; le produit tensoriel défini par

$$L_f \otimes L_g = \sum_{i,j=1}^{p} f_i g_j \frac{\partial^2}{\partial x_i \partial x_j}$$

et on notera $L_f^{\otimes 0} = Id$ et $L_f^{\otimes n+1} = L_f \otimes L_f^{\otimes n}$ pour tout n de $\mathbb{N}$.

Enfin, on définit l'exponentielle tensorielle par:

$$\Delta_f = \sum_{n=0}^{p} \frac{1}{n!} L^{\otimes n}$$

Il est immédiat de vérifier que $\otimes$ est commutatif, associatif, et distributif par rapport à l'addition. On tire de ces remarques élémentaires la formule qui justifie le nom de Δ

$$\Delta_{f+g} = \Delta_f \otimes \Delta_g$$

L'une des propriétés de l'exponentielle tensorielle qui nous sera très utile est la suivante, simple réécriture de la formule de Taylor

$$\Delta_f(h) \mid_x = h(x + f(x))$$

On peut ainsi écrire par exemple

$$\Delta_f \circ \Delta_g(h) \mid_x = \Delta_g(h) \mid_{x+f(x)} = h(x + f(x) + g(x + f(x)))$$

Nous considérerons dans la suite tout système non linéaire qui s'écrit, quand t parcourt $\mathbb{N}$:

$$\begin{cases} x_{t+1} = x_t + f(x_t) + \sum_{l=1}^{r} g_l(x_t) w_t^l + \sum_{l=r+1}^{r+q} g_l(x_t) v_t^l \\ y_{t+1} = h(x_t) + v_t \end{cases}$$

où - $x_t \in \mathbb{R}^p$, $y_t \in \mathbb{R}^q$, x_0 etant une variable aléatoire de loi μ_0.

- $f : \mathbb{R}^p \to \mathbb{R}^p$, $g_l : \mathbb{R}^p \to \mathbb{R}^p$, $l = 1, \cdots, r+q$, et $h : \mathbb{R}^p \to \mathbb{R}^q$ sont des fonctions analytiques.
- $w : \mathbb{N} \to \mathbb{R}^r$ et $v : \mathbb{N} \to \mathbb{R}^q$ sont des bruits blancs Gaussien.On définit les espaces vectoriels sur $\mathbb{R}$ suivants:

$$\forall k \in \{1, ..., q\} \quad O_k^{d(k)} = \mathrm{Span}_{\mathbb{R}} \left\{ \Delta_f^{j-1}(h_k); j \in \{1, ..., d(k)\} \right\}$$

$$\text{et} \quad O^{\bar{d}} = \sum_{k=1}^{q} O_k^{d(k)}$$

où $\bar{d} = (d(1), ..., d(q))$ est pour le moment un vecteur quelconque de $(\mathbb{N}^*)^q$.

Un système linéaire est un système de la forme particulière

$$\begin{cases} x_{t+1} & = Ax_t + \sum_{l=1}^{r} B_l w_t^l + \sum_{l=r+1}^{r+q} B_l v_t^l \\ y_{t+1} & = Cx_t + v_t \end{cases}$$

Un tel système admet un filtre fini, puisqu'il est de la forme décrite au paragraphe précédent, avec la particularité que les matrices ne dépendent pas encore de l'observation. Il est temps d'introduire la principale notion de ce chapitre. Soit (Σ) un système non linéaire. On peut représenter ce système de la façon suivante:

$$y_{t+1} = \Sigma_t(x_0, w^t, v^t) \qquad t \in I\!N$$

où Σ_t pour $t \in I\!N$ sont des applications mesurables. L'existence des Σ_t provient trivialement de la forme d'un système non linéaire telle qu'elle à été définie.

Définition 2.1

Soient (Σ^1) et (Σ^2) deux systèmes stochastiques. Nous dirons que (Σ^1) s'immerge dans (Σ^2) ssi il existe une immersion $\tau : X_1 \to X_2$ telle que pour tout $t \in I\!N$

$$\Sigma^1(x_0, w^t, v^t) = \Sigma^2(\tau(x_0), w^t, v^t) \qquad P_t - \text{p.s.}$$

où P_t est la loi produit des $2t+3$ variables aléatoires indépendantes (x_0, w^t, v^t).

Rappellons qu'une immersion d'une variété analytique X_1 dans une variété analytique X_2 est un difféomorphisme de X_1 dans $\tau(X_1)$.

Définition 2.2

Le système linéaire

$$\begin{cases} x_{t+1} & = Ax_t + \sum_{l=1}^{r} B_l w_t^l + \sum_{l=r+1}^{r+q} B_l v_t^l \\ y_{t+1} & = Cx_t + v_t \end{cases}$$

est dit de la forme de Brunowski ssi les matrices A, B, et C s'écrivent sous la forme

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ a_{11}^1 & & & \cdots & a_{1d_1}^1 & \cdots & a_{q1}^1 & & & \cdots & a_{qd_q}^1 \\ \vdots & & & & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ a_{11}^q & & & \cdots & a_{1d_1}^q & \cdots & a_{q1}^q & & & \cdots & a_{qd_q}^q \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11}^1 & b_{11}^2 & \cdots & \cdots & b_{11}^r \\ b_{12}^1 & b_{12}^2 & \cdots & \cdots & b_{12}^r \\ \vdots & \vdots & & & \vdots \\ b_{1d_1}^1 & b_{1d_1}^2 & \cdots & \cdots & b_{1d_1}^r \\ \vdots & \vdots & & & \vdots \\ b_{q1}^1 & b_{q1}^2 & \cdots & \cdots & b_{q1}^r \\ b_{qd_q}^1 & b_{qd_q}^2 & \cdots & \cdots & b_{qd_q}^r \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots & \vdots & & & \\ 0 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \end{pmatrix}$$

ù les $a_{i,j}^k$ et les $b_{i,j}^k$ sont des constantes fixées.

Théorème 2.3

Une condition nécessaire et suffisante pour qu'un système non linéaire (Σ) s'immerge dans un système linéaire (L) est que

a) $\exists d = (d(1), ..., d(q)) \in (I\!N^*)^q \quad \forall k \in \{1, ..., q\}$

$$\Delta_f^{d(k)}(h_k) \in O^{\bar{d}}$$

On notera $\bar{d} = \sum_{k=1}^q d(k)$

b) $\forall k \in \{1, ..., r\} \quad \exists K = \text{Cte} \quad \forall \Phi \in O^{\bar{d}}$

$$\Delta_f \otimes L_{g_k}(\Phi) = K$$

c) $\forall m \geq 2 \forall k(1), ..., k(m) \in \{1, ..., r\} \forall \Phi \in O^{\bar{d}}$

$$\Delta_f \otimes L_{g_{k(m)}} \otimes \cdots \otimes L_{g_{k(1)}}(\Phi) = 0$$

Démonstration

Nous allons adapter la démonstration de S Monaco et D Normand-Cyrot à notre situation stochastique.

Condition suffisante:

Soit donc le système non linéaire suivant

$$\begin{cases} x_{t+1} = x_t + f(x_t) + \sum_{l=1}^{r} g_l(x_t) w_t^l + \sum_{l=r+1}^{r+q} g_l(x_t) v_t^l \\ y_{t+1} = h(x_t) + v_t \end{cases}$$

Nous supposons qu'il satisfait les hypothèses usuelles ainsi que les hypothèses a), b) et c) du théorème. Posons

$$\tau(x) = \begin{pmatrix} \tau_1(x) \\ \vdots \\ \tau_q(x) \end{pmatrix} \text{et} \tau_k(x) = \begin{pmatrix} \tau_{k,1}(x) \\ \vdots \\ \tau_{k,d(k)}(x) \end{pmatrix}$$

où $\tau_{i,j}(x) = \Delta_f^{j-1}(h_i) \mid_x$ pour $i = 1, ..., q$ et $j = 1, ..., d(i)$. Montrons que τ est bien une immersion dans un système linéaire, et pour cela, il suffit de montrer que le processus $z(t)$ défini par $z(t) = \tau(x(t))$ est solution d'un système stochastique linéaire. Pour commencer, la sortie $y(t)$ s'écrit bien pour $k \in \{1, ..., q\}$

$$y_k(t) = h_k(x(t)) + v(t) = \tau_{k,1}(x(t)) + v(t) = z_{k,1}(t) + v(t)$$

De plus, $\forall i \in \{1, ..., q\}$ et $\forall j \in \{1, ..., d(i)\}$, on a

$$\begin{aligned} z_{i,j}(t+1) = \tau_{i,j}(x(t+1)) &= \Delta_f^{j-1}(h_i) \mid_{x(t+1)} \\ &= \Delta_{f+\Sigma g \bar{w}(t)} \circ \Delta_f^{j-1}(h_i) \mid_{x(t)} \\ &= \Delta_f^{j}(h_i) \mid_{x(t)} + \sum_{l=1}^{r+q} \bar{w}_l(t) \Delta_f \otimes L_{g_l} \circ \Delta_f^{j-1}(h_i) \mid_{x(t)} \end{aligned}$$

où l'on a posé pour plus de facilités

$$\bar{w}(t) = \begin{pmatrix} w(t) \\ v(t) \end{pmatrix}$$

Nous avons utilisé pour la dernière égalité l'hypothèse c). En utilisant a) et b), nous savons que nous pouvons écrire

$$\Delta_f^{d(k)}(h_k) \mid_{x(t)} = \sum_{l=1}^{d(k)} a_{k,l} \Delta_f^{l-1}(h_k) \mid_{x(t)} = \sum_{l=1}^{d(k)} a_{k,l} z_{k,l}(t)$$

pour tout k de 1 à q, et

$$\Delta_f \otimes L_{g_l} \circ \Delta_f^{j-1}(h_i) \mid_{x(t)} = b_{,j}^{l}$$

pour $i = 1, ..., q$, $j = 1, ..., d(i)$, et $l = 1, ..., r+q$. Il s'ensuit que
Si $j < d(i)$

$$z_{i,j}(t) = z_{i,j+1}(t) + \sum_{l=1}^{r+q} b_{i,j}^{l} \bar{w}_l(t)$$

et que si $j = d(i)$

$$z_{i,d(i)}(t) = \sum_{l=1}^{d(i)} a_{i,l} z_{i,l}(t) + \sum_{l=1}^{r+q} b_{i,j}^{l} \bar{w}_l(t)$$

Finallement, le système vérifié par z est bien un système linéaire, et il est même de la forme de Brunovski.

Condition nécessaire:

Nous allons maintenant supposer que le système s'immerge dans un système linéaire de dimension n du type

$$\begin{cases} x_{t+1} = Ax_t + \sum_{l=1}^{r} B_l w_t^l + \sum_{l=r+1}^{r+q} B_l v_t^l \\ \breve{y}_{t+1} = Cx_t + v_t \end{cases}$$

où les matrices A, B, et C sont de forme quelconque. Montrons que ce système satisfait les hypothèses d'immersibilité, ce qui nous permettra d'ailleurs d'illustrer la signification des hypothèses sur un exemple. Nous avons donc ici, pour tout $i \in \{1, ..., q\}$ et tout $j \in I\!N$

$$\Delta_A^j(C_i) = C_i A^j$$

donc la première hypothèse du théorème est satisfaite en tant que conséquence du théorème de Cayley-Hamilton. D'autre part, $\forall i \in \{1, ..., q\}$, $\forall j \in I\!N$, $\forall k \in \{1, ..., r+q\}$

$$\begin{aligned}
\Delta_A \otimes L_{B_k} \circ \Delta_A^j(C_i) \mid_z &= \sum_{n=0}^{\infty} \frac{1}{n!} L_A^{\otimes n} \otimes L_{B_k}(C_i A^j z) = L_{B_k}(C_i A^j z) \\
&= \sum_{l=1}^{q} B_{l,k} \frac{\partial}{\partial z_l} \sum_{m=1}^{n} (C_i A^j)_m x_m \\
&= \sum_{l=1}^{q} B_{l,k} (C_i A^j)_l = C_i A^j B_k = \text{Cte}
\end{aligned}$$

Enfin, il est clair que la dernière condition est satisfaite, car dans un système linéaire, les dérivée secondes sont nulles. On constate que l'on trouve ici les coefficients de la matrice de Toeplitz associée au système linéaire, matrice qui à une grande importance dans la théorie des systèmes linéaires. Ainsi, les grandeurs qui figurent dans les hypothèses du théorème présentent des analogies certaines avec les grandeurs classiques des systèmes linéaires.

Puisque tout système linéaire s'immerge dans un système linéaire de forme de Brunowski, par transitivité de la relation "immersibilité", on peut supposer que le système non linéaire initial s'immerge dans un système linéaire de la forme de Brunowski. Nous avons $z(0) = \tau(x(0))$ donc $\forall k \in \{1, ..., q\}$

$$y_k(1) = h_k(x(0)) + v(0) \text{et}$$

$$\breve{y}_k(1) = C_k z(0) + v(0) = \tau_{k,1}(x(0)) + v(0)$$

or $y_k = \breve{y}_k \quad P_0 -$ p.s. donc Supp(μ_0) étant d'intérieur non vide, on a

$$\tau_{k,1}(x) = h_k(x)$$

en tout point x de X par analyticité de h et de τ. Regardons ce qui se passe à l'instant suivant

$$y_k(2) = h_k(x(1)) + v(1) = \Delta_{f+g\bar{w}(0)}(h_k) \mid_{x(0)} + v(1)$$

$$\breve{y}_k(2) = z_{k,1}(1) + v(1) = z_{k,2}(0) + \sum_{l=1}^{r+q} B_{k,l} \bar{w}_l(0) + v(1)$$

donc

$$\Delta_{f+g\bar{w}(0)}(h_k) \mid_{x(0)} = \tau_{k,2}(x(0)) + \sum_{l=1}^{r+q} B_{k,l} \bar{w}_l(0)$$

soit encore

$$\sum_{m\geq 0}\sum_{l(1)=1}^{r+q}\cdots\sum_{l(m)=1}^{r+q}\left(\frac{\bar{w}_{l(m)}(0)\cdots\bar{w}_{l(1)}(0)}{m!}\right)\Delta_f\otimes L_{g_{l(m)}}\otimes\cdots\otimes L_{g_{l(1)}}(h_k)\mid_{x(0)}$$

$$=\tau_{k,2}(x(0))+\sum_{l=1}^{r+q}B_{k,l}\bar{w}_l(0)$$

et en identifiant, on obtient à nouveau

$$\tau_{k,2}(x)=\Delta_f(h_k)\mid_x$$

et les conditions

$$\Delta_f\otimes L_{g_l}(h_k)\mid_x=B^l_{k,1}\quad\text{et}\quad\Delta_f\otimes L_{g_{l(m)}}\otimes\cdots\otimes L_{g_{l(1)}}(h_k)\mid_x=0$$

On continue ainsi jusqu'à l'instant $d(k)$ qui nous donnera similairement les dernières conditions, à savoir

$$\sum_{m\geq 0}\sum_{l(1),\ldots,l(m)=1}^{r+q}\left(\frac{\bar{w}_{l(m)}(0)\cdots\bar{w}_{l(1)}(0)}{m!}\right)\Delta_f\otimes L_{g_{l(m)}}\otimes\cdots\otimes L_{g_{l(1)}}\circ\Delta_f^{d(k)-1}(h_k)\mid_{x(0)}$$

$$=\sum_{l=1}^{d(k)}a_{k,l}\tau_{k,l}(x(0))+\sum_{l=1}^{r+q}B_{k,l}\bar{w}_l(0)$$

soit finallement

$$\Delta_f^{d(k)}(h_k)\mid_x=\sum_{l=1}^{d(k)}A_{k,l}\tau_{k,l}(x)=\sum_{l=1}^{d(k)}A_{k,l}\Delta_f^{l-1}(h_k)\mid_x$$

$$\Delta_f\otimes L_{g_l}\circ\Delta_f^{d(k)-1}(h_k)\mid_x=B^l_{k,d(k)}$$

et

$$\Delta_f\otimes L_{g_{l(m)}}\otimes\cdots\otimes L_{g_{l(1)}}\circ\Delta_f^{d(k)-1}(h_k)\mid_x=0$$

soit donc toutes les conditions du théorème.

3. FINITE FILTER BY IMMERSION INTO A CONDITIONNALLY GAUSSIAN SYSTEM

We say that a system Σ can be immersed into a system $\overline{\Sigma}$ if and only if there exists an analytic function τ from $\mathbb{R}^p\times\mathbb{R}^q$ to $\mathbb{R}^p$ such that the input-output map of $\overline{\Sigma}$ initialized in $\tau(x_0,y_0)$ is the same as the input-output map of Σ initialized in x_0.

Let us first introduce some notation and definitions. We will consider nonlinear systems of the form

$$\begin{cases} x_{t+1} = x_t + f(x_t, y_t) + g(x_t, y_t)w_t + \tilde{g}(x_t, y_t)v_t \\ y_{t+1} = h(x_t, y_t) + v_t \end{cases}$$

where

$x_t \in \mathbb{R}^p$, $y_t \in \mathbb{R}^q$, x_0 and y_0 are points of $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively.

$f : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^p$, $g : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^p \times \mathbb{R}^r$, $\tilde{g} : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^p \times \mathbb{R}^q$, and $h : \mathbb{R}^p \to \mathbb{R}^q$ are analytic functions.

$w : \mathbb{N} \to \mathbb{R}^r$ and $v : \mathbb{N} \to \mathbb{R}^q$ are any deterministic functions

We note

$$c : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q \text{ the projection on } \mathbb{R}^q$$

$$F = \begin{pmatrix} f \\ h - c \end{pmatrix} : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^p \times \mathbb{R}^q$$

$$G = \begin{pmatrix} g & \tilde{g} \\ 0 & Id \end{pmatrix} : \mathbb{R}^p \times \mathbb{R}^q \to (\mathbb{R}^p \times \mathbb{R}^q) \times (\mathbb{R}^r \times \mathbb{R}^q)$$

and we define the $\mathbb{R}$-vector spaces $\forall k \in \{1, \ldots, q\}$

$$\mathcal{O}_k^{d(k)} = \mathrm{Span}_{\mathbb{R}} \left\{ \Delta_F^{j-1}(c_k); \quad j \in \{1, \ldots, d(k)\} \right\}$$

and

$$\mathcal{O}^{\bar{d}} = \sum_{k=1}^{q} \mathcal{O}_k^{d(k)}$$

where $d = (d(1), \ldots, d(q))$ is any vector of $(N^*)^q$. Let c be a continuous function from $\mathbb{R}^n$ to I, where I is an open set in $\mathbb{R}^q$, and let

$$c^* \left(\mathcal{C}^\omega(I, \mathbb{R}) \right)$$

the set of continuous functions from $\mathbb{R}^n$ to $\mathbb{R}$ written as $f \circ c$ where f is an analytic function from I to $\mathbb{R}$. This functional space is a $\mathbb{R}$-module. For any vector space E of numerical functions, we denote

$$E \otimes_{\mathbb{R}} c^* \left(\mathcal{C}^\omega(I, \mathbb{R}) \right)$$

the tensor product of E by the $\mathbb{R}$-module $c^* \left(\mathcal{C}^\omega(I, \mathbb{R}) \right)$, which is the set of linear combinations of products of functions in E by functions in $c^* \left(\mathcal{C}^\omega(I, \mathbb{R}) \right)$.

Definition 3.1

We say that the system

$$\begin{cases} z(t+1) = & \alpha(y(t)) + A(y(t))z(t) + B(y(t))w(t) \\ & + D(y(t))v(t) \\ y(t+1) = & Cz(t) + v(t) \end{cases}$$

is a canonical system if and only if matrices A, B, C, and D and the vector α are of the form

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ a^1_{11} & & & \cdots & a^1_{1d_1} & \cdots & a^1_{q1} & & & \cdots & a^1_{qd_q} \\ \vdots & & & & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ a^q_{11} & & & \cdots & a^q_{1d_1} & \cdots & a^q_{q1} & & & \cdots & a^q_{qd_q} \end{pmatrix}$$

$$\alpha^T = \begin{pmatrix} 0 & \cdots & 0 & \alpha^1 & 0 & \cdots & \cdots & 0 & \cdots & 0 & \alpha^q \end{pmatrix}$$

$$B = \begin{pmatrix} b^1_{11} & b^2_{11} & \cdots & \cdots & b^r_{11} \\ b^1_{12} & b^2_{12} & \cdots & \cdots & b^r_{12} \\ \vdots & \vdots & & & \vdots \\ b^1_{1d_1} & b^2_{1d_1} & \cdots & \cdots & b^r_{1d_1} \\ \vdots & \vdots & & & \vdots \\ b^1_{q1} & b^2_{q1} & \cdots & \cdots & b^r_{q1} \\ b^1_{qd_q} & b^2_{qd_q} & \cdots & \cdots & b^r_{qd_q} \end{pmatrix}$$

$$D = \begin{pmatrix} d^1_{11} & d^2_{11} & \cdots & \cdots & d^q_{11} \\ d^1_{12} & d^2_{12} & \cdots & \cdots & d^q_{12} \\ \vdots & \vdots & & & \vdots \\ d^1_{1d_1} & d^2_{1d_1} & \cdots & \cdots & d^q_{1d_1} \\ \vdots & \vdots & & & \vdots \\ d^1_{q1} & d^2_{q1} & \cdots & \cdots & d^q_{q1} \\ d^1_{qd_q} & d^2_{qd_q} & \cdots & \cdots & d^q_{qd_q} \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots & \vdots & & & \\ 0 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0 \end{pmatrix}$$

where α^k, a_{ij}^k and b_{ij}^k are analytic functions from $I\!R^q$ to $I\!R$

The form of the system is classical. When all matrices are constant, any linear system can be rewritten in this form. But this is not true in our case.

Proposition 3.2

We consider a nonlinear system $\bar{\Sigma}$. Let c be the matrix $(\ {}_qO^p \quad {}_qI^q\ )$. Necessary and sufficient conditions for $\bar{\Sigma}$ to be immersed in the canonical system as mentioned above are

a) $\exists \bar{d} = (d(1), \ldots, d(q)) \in (I\!N^*)^q \quad \forall k \in \{1, \ldots, q\}$

$$\Delta_F^{d(k)}(c_k) \in (\mathcal{O}^{\bar{d}} \oplus I\!R) \otimes_{I\!R} c^* (\mathcal{C}^\omega(I, I\!R))$$

We will denote $d = \sum_{k=1}^q d(k)$

b) $\forall k \in \{1, \ldots, r+q\}$

$$\Delta_F \otimes L_{G_k}(\mathcal{O}^{\bar{d}}) \in c^* (\mathcal{C}^\omega(I, I\!R))$$

c) $\forall m \geq 2 \quad \forall k(1), \ldots, k(m) \in \{1, \ldots, r+q\}$

$$\Delta_F \otimes L_{G_{k(m)}} \otimes \cdots \otimes L_{G_{k(1)}}(\mathcal{O}^{\bar{d}}) = \{0\}$$

Demonstration

Necessity

Suppose that the nonlinear system can be immersed in the canonical system

$$\begin{cases} z(t+1) = & \alpha(y(t)) + A(y(t))z(t) + B(y(t))w(t) \\ & + D(y(t))v(t) \\ y(t+1) = & Cz(t) + v(t) \end{cases}$$

Then, equivalently, the system

$$\begin{cases} x(t+1) = & x(t) + f(x(t), \xi(t)) + g(x(t), \xi(t))w(t) \\ & + \tilde{g}(x(t), \xi(t))v(t) \\ \xi(t+1) = & h(x(t), \xi(t)) + v(t) \\ y(t) = & \xi(t) \end{cases}$$

which can be rewritten in the short form

$$\begin{cases} \bar{x}(t+1) = \bar{x}(t) + F(\bar{x}(t)) + G(\bar{x}(t))\bar{w}(t) \\ y(t) = c\bar{x}(t) \end{cases}$$

where $\bar{x}(t) = \begin{pmatrix} x(t) \\ \xi(t) \end{pmatrix}$ and $\bar{w}(t) = \begin{pmatrix} w(t) \\ v(t) \end{pmatrix}$ can be immersed into the linear system of the canonical form

$$\begin{cases} \bar{z}(t+1) = \bar{\alpha}(y(t)) + \bar{A}(y(t))\bar{z}(t) + \bar{B}(y(t))\bar{w}(t) \\ y(t) = \bar{C}\bar{z}(t) \end{cases}$$

which is another form of the first system where we have noted $\forall k \in \{1, \ldots, q\}$

$$\bar{z}_{k1}(t) = y_k(t)$$

and $\forall i \in \{1, \ldots, d(k)\}$

$$\bar{z}_{k,i+1}(t) = z_{ki}(t)$$

For any input $\bar{w}(t) = (w(t), v(t))^T$, the output $(y(t))$ is the same for the two systems.
Let us define the particular input

$\bar{w}(0) = \bar{w}$ as a vector in $I\!R^r \times I\!R^q$, $w(t) = 0 \quad \forall t > 0$

If the system is initialised with

$$\bar{x} = \begin{pmatrix} x \\ \xi \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

we can write for all $k \in \{1, \ldots, q\}$ fixed and $\forall\ 1 \leq t \leq d(k)$

$$\begin{aligned} y_k(t) &= \Delta_{F+G\bar{w}} \circ \Delta_F^{t-1}(c_k)|_{\bar{x}} \\ &= \sum_{m \geq 0} \sum_{k(1)=1}^{r+q} \cdots \sum_{k(m)=1}^{r+q} \frac{\bar{w}_{k(1)} \cdots \bar{w}_{k(m)}}{m!} \\ &\qquad\qquad \Delta_F \otimes L_{G_{k(m)}} \otimes \cdots \otimes L_{G_{k(1)}} \circ \Delta_F^{t-1}(c_k)|_{\bar{x}} \end{aligned}$$

and on the other hand, the linear system results in

$$\begin{aligned} y_k(t) &= \bar{z}_{k1}(t) = \bar{z}_{k2}(t-1) = \cdots = \bar{z}_{kt}(1) \\ &= \begin{cases} \tau_{k,t+1}(\bar{x}) + \displaystyle\sum_{l=1}^{r+q} \bar{b}^l_{k,t}(y)\bar{w}_l \quad \text{if } t < d(k) \text{ and} \\ \alpha^k(y) + \displaystyle\sum_{i=1}^{q} \sum_{j=1}^{d(i)} a^k_{ij}(y)\tau_{ij}(\bar{x}) \\ \qquad\qquad + \displaystyle\sum_{l=1}^{r+q} \bar{b}^l_{k,d(k)}(y)\bar{w}_l \text{ if not.} \end{cases} \end{aligned}$$

Hence we have

For $m \geq 2$

$$\Delta_F \otimes L_{G_{k(m)}} \otimes \cdots \otimes L_{G_{k(1)}} \circ \Delta_F^{t-1}(c_k)|_{\bar{x}} = 0$$

For $m = 1$

$$\Delta_F^t \otimes L_{G_l} \circ \Delta_F^{t-1}(c_k)|_{\bar{x}} = \bar{b}^l_{k,t}(c)|_{\bar{x}}$$

and for $m = 0$

$$\Delta_F^t(c_k)|_{\bar{x}} = \tau_{k,t+1}(\bar{x}) \qquad \text{if } t < d(k) \text{ and else}$$

$$\Delta_F^{d(k)}(c_k)|_{\bar{x}} = \alpha^k(y) + \sum_{i=1}^{q}\sum_{j=1}^{d(i)} a_{ij}^k(y)\tau_{ij}(\bar{x})$$
$$= \alpha^k(y) + \sum_{i=1}^{q}\sum_{j=1}^{d(i)} a_{ij}^k(y)\Delta_F^{j-1}(c_i)|_{\bar{x}}$$

Sufficiency

Let us rewrite the initial system

$$\begin{cases} x(t+1) = x(t) + f(x(t),\xi(t)) + g(x(t),\xi(t))w(t) \\ \qquad\qquad + \tilde{g}(x(t),\xi(t))v(t) \\ \xi(t+1) = h(x(t),\xi(t)) + v(t) \\ y(t) = \xi(t) \end{cases}$$

in the form
$$\begin{cases} \bar{x}(t+1) = \bar{x}(t) + F(\bar{x}(t)) + G(\bar{x}(t))\bar{w}(t) \\ y(t) = c\bar{x}(t) \end{cases}$$

where we have defined $\bar{x}(t) = \begin{pmatrix} x(t) \\ \xi(t) \end{pmatrix}$ and $\bar{w}(t) = \begin{pmatrix} w(t) \\ v(t) \end{pmatrix}$, like previously. Let $\tau_{ij}(\bar{x}) = \Delta_F^{j-1}(c_i)|_{\bar{x}}$, and $\forall i \in \{1,\ldots,q\}$

$$\tau_i(\bar{x}) = \begin{pmatrix} \tau_{i1}(\bar{x}) \\ \vdots \\ \tau_{id(i)}(\bar{x}) \end{pmatrix} \text{ and } \tau(\bar{x}) = \begin{pmatrix} \tau_1(\bar{x}) \\ \vdots \\ \tau_q(\bar{x}) \end{pmatrix}$$

At first, we will show that Σ can be immersed into

$$\begin{cases} z(t+1) = A(y(t))z(t) + B(y(t))\bar{w}(t) \\ y(t) = Cz(t) \end{cases}$$

where matrices A, B, and C are in the usual form and $z(t) = \tau(\bar{x}(t))$. We have

$$y_i(t)(=\xi_i(t)) = c_i(\bar{x}(t)) = z_{i1}(t) = C_i z(t)$$

but we must verify the linearity of the evolution of $\tau(\bar{x}(t))$. According to a), $\forall k \in \{1,\ldots,q\}$

$\exists \alpha^k$ continuous functions from $I\!R^q$ to $I\!R$ and

$\exists a_{ij}^k$ continuous functions from $I\!R^q$ to $I\!R$ such that

$$\Delta_F^{d(k)}(c_k) = \alpha^k(y) + \sum_{i=1}^{q}\sum_{j=1}^{d(i)} a_{ij}^k(y)\Delta_F^{j-1}(c_i)$$

and by b), $\forall k \in \{1,\ldots,r\}$ $\quad\forall i \in \{1,\ldots,q\}$ $\quad\forall j \in \{1,\ldots,d(i)\}$

$\exists b_{ij}^k$ continuous functions from $I\!R^q$ to $I\!R$ such that

$$\Delta_F \otimes L_{G_k} \circ \Delta_F^{j-1}(c_i) = b_{ij}^k(c)$$

Therefore, letting $i \in \{1, \ldots, q\}$ and $j \in \{1, \ldots, d(i)\}$, we have

$$\begin{aligned}
z_{ij}(t+1) = \tau_{ij}(\bar{x}(t+1)) &= \Delta_F^{j-1}(c_i)|_{\bar{x}(t+1)} \\
&= \Delta_{F+G\bar{w}} \circ \Delta_F^{j-1}(c_i)|_{\bar{x}(t)} \\
&= \Delta_F^j(c_i)|_{\bar{x}(t)} + \sum_{k=1}^{r+q} \bar{w}_k(t)\Delta_F \otimes L_{G_k} \circ \Delta_F^{j-1}(c_i)|_{\bar{x}(t)} \\
&= \Delta_F^j(c_i)|_{\bar{x}(t)} + \sum_{k=1}^{r+q} b_{ij}^k(y(t))\bar{w}_k(t) \\
&= z_{i,j+1}(t) + \sum_{k=1}^{r+q} b_{ij}^k(y(t))\bar{w}_k(t) \qquad \text{if } j < d(i) \text{ and else}
\end{aligned}$$

$$z_{id(i)}(t+1) = \sum_{i'=1}^{q} \sum_{j'=1}^{d(i')} a_{i'j'}^i(y(t))z_{i'j'}^i(t) + \sum_{k=1}^{r+q} b_{ij}^k(y(t))\bar{w}_k(t)$$

Hence we have $z(t+1) = A(y(t))z(t) + B(y(t))\bar{w}(t)$. But also $\forall k \in \{1, \ldots, r+q\}$ $\quad \forall i \in \{1, \ldots, q\}$ we have

$$\Delta_F \otimes L_{G_k}(c_i)|_{\bar{x}} = \Delta_F \otimes L_{G_k}(y_i) = \delta_{k-r}(i)$$

where δ_j is the function, defined on $\mathbb{Z}$, such that $\delta_j(j) = 1$ and $\delta_j(i) = 0$ everywhere else.

Therefore , we can write the system in the form

$$\begin{cases}
\tilde{z}(t+1) = \tilde{\alpha}(y(t)) + \tilde{A}(y(t))\tilde{z}(t) + \tilde{B}(y(t))w(t) \\
\qquad\qquad + \tilde{D}(y(t))v(t) \\
y(t+1) = \tilde{C}\tilde{z}(t) + v(t)
\end{cases}$$

□

Exemple 3.3

Le système non linéaire

$$\begin{cases}
x_1(t+1) = x_2(t) - x_1(t)^2 \\
x_2(t+1) = y(t)x_1(t) + x_2(t) - x_1(t)^2 + (x_2(t) - x_1(t)^2)^2 + y(t)w(t) \\
y(t+1) = x_1(t) + v(t)
\end{cases}$$

s'immerge dans le système linéaire

$$\begin{cases}
z_1(t+1) = z_2(t) \\
z_2(t+1) = y(t)z_1(t) + z_2(t) + y(t)w(t) \\
y(t+1) = z_1(t) + v(t)
\end{cases}$$

par l'immersion τ définie par

$$\tau(x,y) = \begin{pmatrix} x_1 \\ x_2 - x_1^2 \end{pmatrix}$$

We can now prove the following theorem

Theorem 3.4

Consider the nonlinear system

$$\begin{cases} x_{t+1} = x_t + f(x_t,y_t) + g(x_t,y_t)w_t + \tilde{g}(x_t,y_t)v_t \\ y_{t+1} = h(x_t,y_t) + v_t \end{cases}$$

where $x_t \in \mathbb{R}^p$, $y_t \in \mathbb{R}^q$, $(w_t)_{t\geq 0}$ is a $\mathbb{R}^r$-valued Gaussian white noise, (v_t) is a $\mathbb{R}^q$-valued Gaussian white noise, and x_0 is a random variable with probability measure μ_0. We suppose that the random variables (w_t), (v_t) and x_0 are independents, and that the functions f, g, $\tilde{g}$, and h are analytic functions. Suppose

a) $\exists \bar{d} = (d(1),\ldots,d(q)) \in (\mathbb{N}^*)^q \quad \forall k \in \{1,\ldots,q\}$

$$\Delta_F^{d(k)}(c_k) \in (\mathcal{O}^{\bar{d}} \oplus \mathbb{R}) \otimes_{\mathbb{R}} c^*\left(\mathcal{C}^\omega(I,\mathbb{R})\right)$$

b) $\forall k \in \{1,\ldots,r+q\}$

$$\Delta_F \otimes L_{G_k}(\mathcal{O}^{\bar{d}}) \in c^*\left(\mathcal{C}^\omega(I,\mathbb{R})\right)$$

c) $\forall m \geq 2 \quad \forall k(1),\ldots,k(m) \in \{1,\ldots,r+q\}$

$$\Delta_F \otimes L_{G_{k(m)}} \otimes \cdots \otimes L_{G_{k(1)}}(\mathcal{O}^{\bar{d}}) = \{0\}$$

d) The analytic function τ defined by

$$\tau_{ij}(x,y) = \Delta_f^{j-1}(h_i)|_{x,y}$$

is an injection.
Then the system admits a FDF

Demonstration

This is a corollary of theorem 1 and proposition 2. Therefore, letting $z_t = \tau(x_t,y_t)$, we have for any Borel set in $\mathbb{R}^p$

$$P\{x_t \in B|\mathcal{F}_t^y\} = P\{z_t \in \tau(B)|\mathcal{F}_t^y\}$$

and hence, since the system (z_t,y_t) admits a FDF, (x_t,y_t) admits a FDF. $\square$

Exemple 3.5

On considère le système non linéaire suivant

$$y(t) = A.\sin(2\pi\eta t + \phi) + V(t)$$

où

- A et ϕ sont des variables aléatoires de lois quelconques,
- η est une constante fixée ou une variable aléatoire observée.

Alors ce système admet un filtre de dimension finie.

4. FILTRE FINI PAR IMMERSION ET BOUCLAGE DYNAMIQUE

Dans cette partie, nous allons commencer par considerer des systèmes non linéaires en temps discret déterministes de la forme générale

$$\begin{cases} x_{t+1} = x_t + f(x_t) + g(x_t)u_t \\ \quad y_t = h(x_t) \end{cases}$$

Nous supposerons comme d'habitude les fonctions analytiques.

Définition 4.1

On appelle bouclage par retour d'état l'opération qui consiste à faire le changement de variable suivant:

$$u_t = \alpha(x_t) + \beta(x_t)v_t$$

où $\alpha : I\!R^p \to I\!R^r$ et $\beta : I\!R^p \to I\!R^r \times I\!R^r$ sont des applications analytiques, β étant toujours de rang plein. Dans le cas particulier où β est l'identité, nous dirons que le bouclage est simple, et nous nous intéresserons en fait uniquement à ce type de bouclage, pour des raisons que nous verrons ultérieurement.

Définition 4.2

On appelle nombre caractéristique l'entier d qui est tel que

$$\forall n \in I\!N^* \quad \forall x \in I\!R^p \quad \forall r \in I\!N \qquad r < d \Longrightarrow \Delta_f \otimes L_g^{\otimes n} \circ \Delta^r(h)|_x = 0$$
$$\forall n \in I\!N^* \quad \forall x \in U \quad \text{ouvert dense de } I\!R^p, \qquad \Delta_f \otimes L_g^{\otimes n} \circ \Delta^d(h)|_x \neq 0$$

Nous verrons ultérieurement que d représente l'inertie de la sortie par rapport à l'entrée, soit en termes plus précis que d est le plus petit entier tel que y_{d+1} soit affecté par u_0. Notons que d peut être nul, ou infini. On a cependant le théorème suivant, adapté de D. Normand-Cyrot et S. Monaco:

Théorème 4.3

Si $d < \infty$ et si $\forall x \in I\!R^p$

$$\Delta_f \otimes L_g^{\otimes n} \circ \Delta_f^d(h)|_x = 1 \text{ si } n = 1 \text{ et } 0 \text{ si } n > 1$$

alors le système non linéaire initial peut être immergé dans un système linéaire par un bouclage par retour d'état simple.

Démonstration

Nous allons reproduire la démonstration de D. Normand-Cyrot et S. Monaco. Posons $\xi_t = \tau(x_t)$ où $\tau(x)$ est l'application analytique de $I\!R^p$ dans $I\!R^{d+1}$ dont la $(r+1)^{ieme}$ coordonnée est $\Delta_f^r(h)|_x$ et montrons que l'on peut écrire $\xi_{t+1} = A\xi_t + Bv_t$. En effet, $\forall r \in \{0, ..., d\}$,

$$\Delta_f^r(h)|_{x_{t+1}} = \Delta_{f+gu_t} \circ \Delta_f^r(h)|_{x_t} = \sum_{n=0}^{\infty} \frac{u_t^n}{n!} \Delta_f \otimes L_g^{\otimes n} \circ \Delta_f^r(h)|_{x_t}$$
$$= \Delta_f^{r+1}(h)|_{x_t} + \sum_{s=1}^{\infty} \frac{u_t^n}{n!} \Delta_f \otimes L_g^{\otimes n} \circ \Delta_f^r(h)|_{x_t}$$

Si $r < d$, le second terme est nul par définition de d.
Si $r = d$, le second terme vaut u_t par hypothèse du théorème.

Posant $u_t = -\Delta_f^{d+1}(h)|_{x_t} + v_t$, on obtient bien un système linéaire en v_t. Nous allons maintenant utiliser ce résultat pour des systèmes stochastiques.

Nous allons pouvoir faire sur les systèmes stochastiques une opération analogue au bouclage par retour d'état simple, en utilisant la version discrète du théorème de Girsanov que nous avons vu précedemment.

Théorème 4.4

Considérons le système suivant

$$\begin{cases} x_{t+1} = x_t + f(x_t) + g(x_t)w_t \\ y_t = h(x_t) + v_t \end{cases}$$

où (w_t) et (v_t) sont deux bruits blancs à valeurs dans $I\!R$, indépendants. On suppose que la loi de x_0 est indépendante de (w_t) et de (v_t). Soit d le nombre caractéristique associé à ce système. Supposons que

1) $d < \infty$ et $\Delta_f \otimes L_g^{\otimes n} \circ \Delta_f^d(h)|_x = 1$ si $n = 1$, et 0 si $n > 1$.
2) l'immersion τ est injective et $\forall x \in I\!R^p \quad \det(\tau'(x)) \neq 0$

alors le système admet un filtre fini sous une probabilité $\tilde{P}$ équivalente à P.

Démonstration

Posons $\tilde{w}_t = w_t + \Delta_f^{d+1}(h)|_x$ et introduisons la probabilité

$$d\tilde{P} = \exp\left\{\sum_{s=0}^{t} -\Delta_f^{d+1}(h)|_{x_s}\tilde{w}_s + \frac{1}{2}\sum_{s=0}^{t}\left(\Delta_f^{d+1}(h)|_x^2\right)^2\right\} dP$$

Sous $\tilde{P}$, le système s'écrit

$$\begin{cases} x_{t+1} = x_t + (f(x_t) - \Delta_f^{d+1}(h)|_{x_s} g(x_t)) + g(x_t)\tilde{w}_t \\ y_t = h(x_t) + v_t \end{cases}$$

et il s'immerge donc par l'application τ dans un système linéaire, les conditions du théorème précédent étant ici satisfaites. Le système linéaire en question n'a pas une condition initiale gaussienne, mais d'après le théorème de Makowski, $\tau(x_t)$ est bien solution d'un système admettant un filtre de dimension finie: on peut paramétrer $L(\tau(x_t)/\mathcal{F}_Y^t)$ par un nombre fini de paramètres calculables récursivement en temps. Mais connaître cette loi, c'est connaître $L(x_t/\mathcal{F}_Y^t)$ puisque τ est injective et que son Jacobien ne s'annulle pas sur $\mathbb{R}^p$ Donc le système admet bien un filtre de dimension finie sous la probabilité $\tilde{P}$.

Exemple 4.5

Pour fixer les idées, le système suivant, où p=2, satisfait les hypothèses du théorème:

$$\begin{cases} x_1(t+1) = x_1(t) + x_2(t) + k(x_1(t)) \\ x_2(t+1) = x_2(t) + l(x_1(t), x_2(t)) + w(t) \\ y(t) = x_1(t) + v(t) \end{cases}$$

Il est cependant plus intéressant de pouvoir revenir à la probabilité initiale. Cela peut se faire dans certains cas, en particulier s'il y a absence de bruit d'observation et que la condition initiale est connue.

Théorème 4.6

Considérons le système

$$\begin{cases} x_{t+1} = x_t + f(x_t) + g(x_t)w_t \\ y_t = h(x_t) \end{cases}$$

où (w_t) est un bruit blanc à valeurs dans $\mathbb{R}$. On suppose ici que x_0 est une constante connue. Soit d le nombre caractéristique associé à ce système. Supposons que

1) $d < \infty$ et $\Delta_f \otimes L_g^{\otimes n} \circ \Delta_f^d(h)|_x = 1$ si $n = 1$, et 0 si $n > 1$.
2) l'immersion τ est injective et $\forall x \in \mathbb{R}^p \qquad \det(\tau'(x)) \neq 0$
3) $\Delta_f^{d+1}(h) \in h^*(C^\omega(\mathbb{R}, \mathbb{R}))$

alors le système admet un filtre de dimension 2p+1.

Démonstration

Posons

$$d\tilde{P} = \exp\left\{\sum_{s=0}^{t} -\Delta_f^{d+1}(h)|_{x_s}\tilde{w}_s + \frac{1}{2}\sum_{s=0}^{t}\left(\Delta_f^{d+1}(h)|_x^2\right)^2\right\} dP$$

avec

$$\tilde{w}_t = w_t + \Delta_f^{d+1}(h)|_{x_t}$$

Sous $\tilde{P}$, $(\tilde{w}_t)$ est un bruit blanc gaussien et le système s'écrit

$$\begin{cases} x_{t+1} = x_t + (f(x_t) - \Delta_f^{d+1}(h)|_{x_t} g(x_t))^2 + g(x_t)\tilde{w}_t \\ \quad y_t = h(x_t) \end{cases}$$

donc il s'immerge par τ dans un système du type suivant:

$$\begin{cases} z_{t+1} = Az_t + Bw_t \\ \quad y_t = Cz_t \end{cases}$$

avec $z_0 = \tau(x_0)$. Par hypothèse, il existe une fonction ϕ telle que

$$\Delta_f^{d+1}(h)|_{x_t} = \phi(h)\,|_{x_t} = \phi(y_t)$$

Ajoutons l'équation d'état suivante

$$\xi_{t+1} = \xi_t + \frac{1}{2}\phi^2(y_t) - \phi(y_t)\tilde{w}_t$$

avec $\xi_0 = 0$. Le système obtenu est linéaire conditionnellement gaussien, donc admet un filtre de Kalman, ce qui se traduit du fait de l'injectivité de l'immersion par l'existence d'un filtre de dimension finie pour le système initial.

RÉFÉRENCES

[1] **C.D.de Benito, K.A.Loparo,** "Nonlinear Filtering for Discrete-Time Linear Systems with Non-Gaussian Initial Data," *IEEE Trans. on Aut. Cont.* 34, Juin *1989* 7 pp.781-783

[2] **Busvelle E., Rakotopara D., de Brucq D.** "A sufficient condition for the existence of finite dimensional filters for discrete-time non-linear systems," *Proceedings of the* 29^{th} *Conf. on Dec. and Cont. (1990) p234–239*

[3] **D. de Brucq, D. Rakotopara,** "Equations de filtrage non linéaire pour une loi d'état conditionnellement gaussienne," Colloque *AFCET Toulouse*, p225-231

[4] **D. de Brucq, G. Foliot,** "Théorie du signal" Masson

[5] **Lévine J., Pignié G.,** "Exact finite dimensional filter for a class of nonlinear discrete-time filter," Stochastics 18 (1986) p97-132

[6] **S.Monaco, D.Normand-Cyrot,** "Sur la subordination d'un système non linéaire discret à un système linéaire," *Actes du Colloque National CNRS*, RCP567

[7] **S.Monaco, D.Normand-Cyrot,** "The immersion under feedback of a multidimensional discrete-time nonlinear system into a linear system," *int. J. of Cont.*

[8] **R.B.Sowers, A.M.Makowski,** "Discrete-time filtering for linear systems in correlated noise with non-Gaussian initial conditions," *Proceedings of the 23rd Annual Conference on Information Sciences and Systems,* The Johns Hopkins University, Baltimore (MD), March *1989* p308-313

[9] **R.B.Sowers, A.M.Makowski,** "Discrete-time filtering for linear systems in correlated noise with non-Gaussian initial conditions: Formulas and asymptotics," *Robust Control of Linear Systems and Nonlinear Control,* Eds. E.M.A. Kaaskoek, J.H. van Schuppen and A.C.M. Ran, Birkhäuser Boston Inc., Cambridge (MA), Proceedings of the 1989 International Symposium on the Mathematical Theory of Networks and Systems, Amsterdam (The Netherlands), June *1989*, in Press.

CHAPITRE IV

FILTRE DE DIMENSION FINIE
SYSTÈMES EN TEMPS CONTINU

INTRODUCTION

Pour commencer, regardons un cas très simple où il s'agit en fait d'une "erreur de modélisation", c'est à dire qu'en prenant les bonnes variables d'état, on a un système linéaire. Pour cela, nous allons chercher à quelles conditions un système non linéaire peut se ramener à un système linéaire par un difféomorphisme. Ainsi, considérons un système non linéaire de la forme

$$\begin{cases} dx_t = f(x_t)dt + g(x_t)dw_t \\ dy_t = h(x_t)dt + dv_t \end{cases}$$

On cherche une transformation de x_t en z_t telle que le système en z_t soit linéaire, c'est à dire telle qu'il existe des matrices A, B et C telles que

$$\begin{cases} dz_t = Az_tdt + Bdw_t \\ dy_t = Cz_tdt + dv_t \end{cases}$$

Supposons par exemple que h soit bijective, ce qui est le cas en pratique lorsque l'observation de l'état est totale (et non plus partielle). On cherche alors à ramener le système non linéaire sous la forme

$$\begin{cases} dz_t = Az_tdt + Bdw_t \\ dy_t = Cz_tdt + dv_t \end{cases}$$

et par conséquent, le changement de variable ne peut être que

$$z_t = h(x_t)$$

Cherchons les conditions que doivent satisfaire f et g pour que l'on puisse trouver A et B. La formule de Itô donne

$$\begin{aligned} dz_t = dh(x_t) &= h'(x_t)dx_t + \frac{1}{2}h''(x_t)d < x, x >_t \\ &= h'(x_t)f(x_t)dt + g(x_t)dw_t + \frac{1}{2}h''(x_t)g^2(x_t)dt \\ &= h'(x_t)f(x_t)dt + \frac{1}{2}h''(x_t)g^2(x_t)dt + h'(x_t)g(x_t)dw_t \end{aligned}$$

qui doit être à rapprocher de

$$dz_t = Az_tdt + Bdw_t$$

En identifiant parties drift et parties martingales, on obtient les conditions suffisantes suivantes

$$h'(x)f(x) + \frac{1}{2}h''(x)g^2(x) = Ah(x)$$
$$h'(x)g(x) = B$$

En dérivant la seconde équation, et en substituant h'' dans la première, on obtient les conditions équivalentes

$$h''(x)g(x) + h'(x)g'(x) = 0$$
$$h'(x)f(x) - \frac{1}{2}h'(x)g'(x)g(x) = Ah(x)$$

la seconde équation se réécrivant plus simplement

$$Bf(x) = \frac{1}{2}Bg'(x)g(x) = Ag(x)h(x)$$

d'où l'expression de f en fonction de g et de h

$$f(x) = g(x)(AB^{-1}h(x) + \frac{1}{2}g'(x))$$

On a ainsi démontré le résultat suivant

Lemme 0.1

Le système non linéaire

$$\begin{cases} dx_t = f(x_t)dt + g(x_t)dw_t \\ dy_t = h(x_t)dt + dv_t \end{cases}$$

est difféomorphe au système linéaire

$$\begin{cases} dz_t = Az_tdt + Bdw_t \\ dy_t = Cz_tdt + dv_t \end{cases}$$

dès qu'il existe des constantes B et C telles que

$$h(x) = B(\int g^{-1}(x)dx + C)$$
$$f(x) = g(x)(AB^{-1}h(x) + \frac{1}{2}g'(x))$$

et le difféomorphisme est donné par

$$z = h(x)$$

Le principal intérêt de ce résultat est qu'il permet de donner un certain nombre d'exemples de systèmes non linéaires admettant des filtres finis, mais ces exemples sont des cas d'écoles. Il faudra travailler un peu plus pour arriver à des résultats exploitables. Notons tout de même que nous avons fait des hypothèses très fortes, par exemple en supposant que l'observation du système était totale. Chercher des conditions d'existence de difféomorphisme linéarisant un système non linéaire dans un cas général est beaucoup plus difficile. Mais notre but est de trouver des conditions suffisantes pour immerger un système non linéaire dans un linéaire, ce qui est plus général, car un difféomorphisme est un cas particulier d'immersion injective. Dans toute la suite, nous aurons besoin d'utiliser les résultat suivant.

1. CONDITION INITIALE NON GAUSSIENNE

Les théorèmes suivants montrent qu'un système linéaire avec une condition initiale non gaussienne admet un filtre fini.

Théorème 1.1 (Théorème de Makowski)

Considérons le système linéaire sur $I\!R^N$ suivant où $x_t \in I\!R^p$ et $y_t \in I\!R^q$

$$\begin{cases} dx_t = Ax_t dt + BdW_t \\ dy_t = Cx_t dt + dV_t \end{cases}$$

où $(w)_t$ et $(v)_t$ sont deux processus indépendants entre eux et indépendants de la variable aléatoire x_0 que l'on suppose de loi quelconque μ_0 sur $I\!R^p$. Ce système admet un filtre de dimension 2p.

Esquisse de la démonstration

Ecrivons

$$x_t = \Phi(t,0)x_0 + \int_0^t \Phi(t,s)BdW_s$$

où $\Phi(t,s)$ est le noyau de Green associé à la partie dynamique du système.

Posons
$\xi_t^1 = \int_0^t \Phi(t,s) B dW_s$ et $\xi_t^2 = \Phi(t,0)x_0$.

On obtient le système

$$\begin{cases} d\xi_t^1 = A\xi_t^1 dt \quad + \quad BdW_t & \xi_0^1 \ = \ 0 \\ d\xi_t^2 = A\xi_t^2 dt & \xi_0^2 \ = x_0 \\ dy_t = C(\xi_t^1 + \xi_t^2)dt + dV_t & y_0 \ = \ 0 \end{cases}$$

Il reste à poser

$$d\tilde{V}_t = C\xi_t^2 dt + dV_t$$

$(\tilde{V}_t)$ est un mouvement brownien indépendant de x_0 et de (W_t) sous la loi $\tilde{P}$ déterminée par l'exponentielle de Girsanov

$$d\tilde{P} = \exp\left\{ - \int_0^t \Phi(s,0).{}^tC.d\widetilde{V}_s + \frac{1}{2}\int_0^t \Phi(s,0).{}^tC.C.\Phi(s,0).x_0 ds \right\} dP.$$

On obtient ainsi un système linéaire sous la loi $\tilde{P}$, et ξ_t^1 est indépendant de ξ_t^2 sous $\tilde{P}$ conditionnellement à $\sigma(y_s; s \leq t)$. Il reste à revenir à la loi d'origine P, ce qui est techniquement possible ici.

Citons un autre théorème, du à Haussmann et Pardoux, qui étend le résultat précédent, et que nous avons vu en temps discret dans le chapitre III.

Théorème 1.2

Considérons le système suivant

$$\begin{cases} dx_t = \alpha(y_t) + A(y_t)dt + B(y_t)dw_t \\ dy_t = Cx_t dt + dv_t \end{cases}$$

On suppose que α, A, et B sont des applications analytiques à valeurs dans les espaces adéquats, (w_t) et (v_t) étant deux processus de Wiener indépendants à valeurs dans respectivement $I\!R^r$ et $I\!R^q$. Supposons que x_O est une variable aléatoire de loi μ_O, indépendante des deux processus de Wiener précédents. Ce système admet un filtre de dimension finie.

Nous serons aussi ammené à utiliser ce théorème dont la démonstration figure dans l'article de Haussmann-Pardoux. Notons que dans cette article est démontré un théorème sensiblement plus général, qui lui aussi nous sera utile par la suite.

Définition 2.1

Soient Σ et Σ' deux systèmes différentiels stochastiques et notons X et X' les espaces d'états respectifs de ces deux systèmes. Nous dirons que le système Σ peut s'immerger dans le système Σ' s'il existe une application $\tau : X \to X'$ telle que

$$Y(t, W, V, X_0) = Y'(t, W, V, \tau(X_0)) \qquad P_{(W,V)} - \text{p.s.}$$

Notre but sera en général d'immerger un système non linéaire donné dans un système linéaire, ou plus généralement dans un système qui admet un filtre de dimension finie. Il est déjà très fructueux de voir que l'observation est la sortie d'un système linéaire. Mais si τ est une application injective qui envoie un système donné dans un système admettant un filtre de dimension finie, alors le système initial admet un filtre de dimension finie. Ce résultat sera constament utilisé, et mérite donc d'être énoncé sous la forme d'une proposition

Proposition 2.2

Si un système différentiel stochastique Σ s'immerge par une application injective τ dans un système admettant un filtre de dimension finie, alors Σ admet un filtre de dimension finie.

Démonstration

Notons X_t la variable d'état du premier système et Z_t celle du second. Soit P_t la loi de X_t sachant F_Y^t et Q_t celle de Z_t sachant F_Y^t. Soit $\varphi \in C_b(Z)$

$$\begin{aligned}\int_Z \varphi(z) dQ_t(z) &= E\left[\varphi(Z_t)/F_Y^t\right] \\ &= E\left[\varphi(\tau(X_t))/F_Y^t\right] \\ &= \int_X \varphi \circ \tau(x) dP_t(x) \\ &= \int_X \varphi(z) d(\tau \circ P_t)(z)\end{aligned}$$

donc Q_t est la probabilité image de P_t par τ. Mais τ étant injective et analytique, il existe une application analytique θ telle que $\theta \circ \tau$ soit l'identité sur X. Soit donc $\psi \in C_b(X)$. Le théorème de transfert nous indique encore que

$$\begin{aligned}\int_X \psi(x) d(\theta \circ Q_t)(x) &= \int_X \psi \circ \theta(z) dQ_t(z) \\ &= \int_X \psi \circ \theta(z) d(\tau \circ P_t)(z) \\ &= \int_X \psi \circ \theta \circ \tau(x) dP_t(x) \\ &= \int_X \psi(x) dP_t(x)\end{aligned}$$

Il s'ensuit que si Q_t est caractérisée par un nombre fini de paramètres calculables récursivement, il en sera de même de P_t puisque $P_t = \theta \circ Q_t$.

Avant de continuer, précisons quelques notations.

Notations 2.3

Nous noterons $L_f h$ la dérivée de Lie de la fonction h dans la direction du champ de vecteur f. Si f s'écrit en coordonnées locales

$$f = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i}$$

alors

$$(L_f h)(x) = \sum_{i=1}^{n} f_i(x) \frac{\partial h}{\partial x_i}(x)$$

D'autre part, on adoptera la notation naturelle $L_f^0 h = h$ et $\forall \nu > 0 \quad L^\nu h = L_f L^{\nu-1} h$. Pour illustrer ces notations, calculons $L_M N$ où M et N sont des applications linéaires respectivement de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et de $\mathbb{R}^n$ dans $\mathbb{R}$. Soit $x \in \mathbb{R}^n$

$$L_M N \mid_x = \sum_{i=1}^{n} (Mx)_i \frac{\partial N}{\partial x_i} = \sum_{i=1}^{n} N_i \sum_{j=1}^{n} M_{ij} x_j = NMx$$

Nous utiliserons cette remarque ultérieurement.

Notations 2.4

Soit h une application à valeurs dans $\mathbb{R}^n$. Nous noterons $h^*(C^\omega(\mathbb{R}^n, \mathbb{R}))$ l'anneau des fonctions qui peuvent s'écrire $\varphi \circ h$ où φ est une fonction analytique de $\mathbb{R}^n$ dans $\mathbb{R}$. Pour tout espace vectoriel E de fonctions numériques, nous noterons

$$E \otimes_{\mathbb{R}} h^*(C^\omega(\mathbb{R}^n, \mathbb{R}))$$

le produit tensoriel de E par le $\mathbb{R}$-module $h^*(C^\omega(\mathbb{R}^n, \mathbb{R}))$ qui est par définition l'ensemble des combinaisons linéaires de produits de fonctions de E par des fonctions de $h^*(C^\omega(\mathbb{R}^n, \mathbb{R}))$.

Definition 2.5

Nous dirons que le système Σ suivant, défini sur l'espace des états $\mathbb{R}^p$ et l'espace des observations $\mathbb{R}^q$

$$\begin{cases} dX_t = AX_t dt + BdW_t \\ dY_t = CX_t dt + dV_t \end{cases}$$

où X_0 est une variable aléatoire indépendante de processus de Wiener (W_t, V_t) qui évolue dans $\mathbb{R}^r \times \mathbb{R}^q$, est un système de Brunowski ssi il existe un q-uplet $(p_1, \ldots, p_q) \in (\mathbb{N}^*)^q$ tel que les matrices A, B, et C puissent s'écrire sous la forme suivante

$$A=\begin{pmatrix}
0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0\\
0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0\\
\vdots & & & & & & \vdots & & & & \\
0 & 0 & 0 & \cdots & 1 & \cdots & 0 & 0 & 0 & \cdots & 0\\
a^1_{11} & & & \cdots & a^1_{1d_1} & \cdots & a^1_{q1} & & & \cdots & a^1_{qd_q}\\
\vdots & & & & & & \vdots & & & & \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 1 & 0 & \cdots & 0\\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 1 & \cdots & 0\\
\vdots & & & & & & \vdots & & & & \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 1\\
a^q_{11} & & & \cdots & a^q_{1d_1} & \cdots & a^q_{q1} & & & \cdots & a^q_{qd_q}
\end{pmatrix}$$

$$B=\begin{pmatrix}
b^1_{11} & b^2_{11} & \cdots & \cdots & b^r_{11}\\
b^1_{12} & b^2_{12} & \cdots & \cdots & b^r_{12}\\
\vdots & \vdots & & & \vdots\\
b^1_{1d_1} & b^2_{1d_1} & \cdots & \cdots & b^r_{1d_1}\\
\vdots & \vdots & & & \vdots\\
b^1_{q1} & b^2_{q1} & \cdots & \cdots & b^r_{q1}\\
b^1_{qd_q} & b^2_{qd_q} & \cdots & \cdots & b^r_{qd_q}
\end{pmatrix}$$

$$C=\begin{pmatrix}
1 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0\\
0 & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0\\
\vdots & & & & \vdots & \vdots & & & \\
0 & 0 & \cdots & 0 & \cdots & 1 & 0 & \cdots & 0
\end{pmatrix}$$

où tous les coefficients sont des reels fixés. Nous noterons parfois ce système $\Sigma_{(p_1,\ldots,p_q)}$.

Théorème 2.6

Le système non linéaire

$$\begin{cases} dx_t = f(x_t)dt + g(x_t)\circ dw_t \\ dy_t = h(x_t)dt + dv_t \end{cases}$$

s'immerge dans un système linéaire si et seulement si

a. Si $E = \mathrm{Span}_{\mathbb{R}}(L^j_f h_i; 1 \leq i \leq q, j \geq 0)$, $\mathrm{Dim}(E) < \infty$

b. $L_g L^j h_i = Cte \qquad \forall 1 \leq k \leq r \quad \forall L^j h_i \in E$

Remarque 2.7

Les hypothèses de ce théorème sont, pour un système donné, faciles à vérifier. Elles se distinguent en ce sens des théorèmes classiques qui donnent des hypothèses sur la finitude de la dimension d'une algèbre de Lie. De plus, nous étendrons ce théorème à un cas où on ne peut définir l'algèbre de Lie associée au système (cas conditionnellement gaussien) et donc où ce type de théorème prend tout son intérêt.

Démonstration

Nous allons recopier ici le démonstration de Lévine et Marino.

Condition suffisante

Supposons donc que $E = \mathrm{Span}_{I\!R}(L_f^j h_i; 1 \leq i \leq q, 0 \leq j \leq p_i - 1)$ Posons $\tau_{ij}(x) = L_f^{j-1} h_i|_x$, $\tau_i = \begin{pmatrix} \tau_{i1} \\ \vdots \\ \tau_{ip_i} \end{pmatrix}$, et $\tau = \begin{pmatrix} \tau_i \\ \vdots \\ \tau_q \end{pmatrix}$. Nous allons montrer que τ est l'immersion qui convient. On va montrer pour cela que le processus $z(t) = \tau(x_t)$ possède une dynamique linéaire. Il est déjà clair que

$$\forall 1 \leq i \leq q \qquad y_i(t) = h_i(x_t) + v_t = C_i z(t) + v_t$$

où C_i^* est le i$^{\text{ème}}$ vecteur de la base canonique de $I\!R^q$. Ainsi, on a bien

$$y(t) = Cz(t) + v_t$$

et de plus, $\forall 1 \leq i \leq q \quad \forall 1 \leq j \leq p_i$

$$\begin{aligned} dz_{ij}(t) &= d\tau_{ij}(x_t) = d\left(L_f^{j-1} h_i|_{x_t}\right) \\ &= \mathrm{Grad}_x \left(L_f^{j-1} h_i\right)_{x_t} \circ dx_t \\ &= \mathrm{Grad}_x \left(L_f^{j-1} h_i\right)_{x_t} \circ \left(f(x_t)dt + \sum_{k=1}^{r} g_k(x_t) \circ dw_t^k\right) \\ &= L_f^j h_i|_{x_t} dt + \sum_{k=1}^{r} L_{g_h} L_f^{j-1} h_i|_{x_t} \circ dw_t^k \end{aligned}$$

et de deux choses l'une:

$$j < p_i : \qquad dz_{ij}(t) \qquad = z_{i,j+1}(t)dt + \textstyle\sum_{k=1}^r b_{ij}^k dw_t^k$$

où

$$j = p_i : \qquad dz_{ip_i}(t) \quad = \textstyle\sum_{k=1}^r \sum_{l=1}^{p_k} a_{kl}^i z_{kl}(t)dt + \sum_{k=1}^r b_{ip_i}^k dw_t^k$$

et de toute façon, ceci s'écrit sous la forme matricielle

$$dz(t) = Az(t)dt + Bdw_t$$

Notons que le système linéaire obtenu est de la forme de Brunowski.

Condition nécessaire

Pour commencer, nous allons vérifier qu'un système linéaire vérifie les hypothèses, ce qui est la moindre des choses. Soit donc

$$\begin{cases} dx_t = Ax_t dt + Bdw_t \\ dy_t = Cx_t dt + dv_t \end{cases}$$

Pour un tel système,

$$\begin{aligned} E &= \mathrm{Span}_{I\!R}\left\{L_A^j C_i; 1 \leq i \leq q, j \geq 0\right\} \\ &= \mathrm{Span}_{I\!R}\left\{C_i A^j; 1 \leq i \leq q, j \geq 0\right\} \end{aligned}$$

est de dimension finie, ce qui nous est montré par le théorème de Cayley-Hamilton, et $\forall 1 \leq k \leq r \quad \forall L^j C_i \in E$

$$L_{B_k} L_A^j C_i = C_i A^j B_k = \mathrm{Cte}$$

Donc d'après la première partie de la démonstration, tout système linéaire s'immerge dans un système de forme de Brunowski. Donc si un système quelconque s'immerge dans un système linéaire, par transitivité de la relation "s'immerge dans", ce système s'immerge dans un système de forme de Brunowski. On peut verifier que la condition est nécessaire en supposant que le système non linéaire s'immerge dans un système de la forme de Brunowski.

De par la définition d'une inmmersion, on a $\forall 1 \leq k \leq p$

$$h_k(x(t)) = z_{k1}(t)$$

et par dérivation

$$\begin{aligned} L_f h_k(x(t)) &= z_{k2}(t) \\ L_g h_k(x(t)) &= B_{k1} \end{aligned}$$

puis

$$\begin{aligned} L_f^2 h_k(x(t)) &= z_{k3}(t) \\ L_g L_f h_k(x(t)) &= B_{k2} \end{aligned}$$

$$\ldots \text{etc} \ldots$$

Finallement, puisque $Z_{ij}(t) = L_f^{j-1} h_i(x(t))$, on a

$$L_f^{p_k} h_k(x(t)) = \sum_{i=1}^{q} \sum_{j=1}^{p_i} a_{ij}^k Z_{ij}(t)$$

$$L_g L_f^{p_k - 1} h_k(x(t)) = B_{kp_k}$$

$$\mathrm{Dim}\ \mathrm{Span}_{I\!R}\left\{L_f^j h_i; 1 \leq i \leq q, j \geq 0\right\} < \infty$$

et

$$L_{g_k} L_f^j h_i = \mathrm{Cte} \qquad \forall 1 \leq k \leq r \quad \forall 1 \leq i \leq q \quad \forall j \in I\!N$$

3. FINITE-DIMENSIONAL FILTERS FOR CONTINUOUS TIME SYSTEMS

Let $(\Omega, \mathcal{F}, P)$ be a standard probability space, and in this space W and V two independent Wiener processes,respectively $\mathbb{R}^m$-valued and $\mathbb{R}^s$-valued. Let $(\mathcal{F}_t)$ be the filtration associated with the couple (W_t, V_t). If X_t is a semi-martingale, we denote dX_t his Stratonovich differential. Let $(X_t, Y_t) \in \mathbb{R}^n \times \mathbb{R}^s$ be the signal/observation couple, solution of the following differential stochastic system (Σ)

$$(\Sigma) \qquad \left\{ \begin{array}{ll} dX_t= & f(X_t,Y_t)dt + g(X_t,Y_t)dW_t \\ dY_t = & h(X_t,Y_t)dt + dV_t \end{array} \right.$$

We suppose that (Σ) satisfies classical hypotheses to admit a strong solution, and that the vector fields f, g, and h are analytics. For $(p_1, ..., p_s) \in \mathbb{N}^s$, we can define a class of stochastic differential systems denoted $\mathcal{O}_{(p_1,...,p_s)}$ and defined as

$$(\bar{\Sigma}_{p_1,...,p_s}) \qquad \left\{ \begin{array}{ll} dZ_t= & (A(\bar{Y}_t)Z_t + \varphi(\bar{Y}_t))dt + B(\bar{Y}_t)dW_t \\ dY_t= & C\ Z_t dt + dV_t \end{array} \right.$$

where

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 0 \\ a^1_{11} & & & \cdots & a^1_{1p_1} & \cdots & \cdots & a^1_{s1} & & & \cdots & a^1_{sp_s} \\ \vdots & & & & \vdots & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots & & & \vdots & & & & \\ 0 & 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & 0 & 0 & \cdots & 1 \\ a^s_{11} & & & \cdots & a^s_{1p_1} & \cdots & \cdots & a^s_{s1} & & & \cdots & a^s_{sp_s} \end{pmatrix}$$

$$\varphi^* = (\,0 \quad \cdots \quad 0 \quad \varphi^1 \quad 0 \quad \cdots \quad \cdots \quad 0 \quad \cdots \quad 0 \quad \varphi_s\,)$$

$$B = \begin{pmatrix} b^1_{11} & b^2_{11} & \cdots & \cdots & b^r_{11} \\ b^1_{12} & b^2_{12} & \cdots & \cdots & b^r_{12} \\ \vdots & \vdots & & & \vdots \\ b^1_{1p_1} & b^2_{1p_1} & \cdots & \cdots & b^r_{1p_1} \\ \vdots & \vdots & & & \vdots \\ b^1_{s1} & b^2_{s1} & \cdots & \cdots & b^r_{s1} \\ b^1_{sp_s} & b^2_{sp_s} & \cdots & \cdots & b^r_{sp_s} \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 & & 0 & 0 & \cdots & 0 & & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & & 1 & 0 & \cdots & 0 & & 0 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots & & & & & \vdots & & & \\ 0 & 0 & \cdots & 0 & & 0 & 0 & \cdots & 0 & & 1 & 0 & \cdots & 0 \end{pmatrix}$$

where $r = m + s$ and a_{ij}^k, b_{ij}^k, and φ^k are analytic functions of y. Using the result established by in [14], we know that the conditional distribution of Z_t given $\sigma(Y_s;\ s \leq t)$ admits a finite dimensional filter.

Definition 3.1

The differential stochastic system (Σ) can be *immersed* in $(\bar{\Sigma})$ at X_0 if there exists an application $\tau : I\!R^n \to I\!R^m$ such that

$$Y(t, W, V, X_0) = \bar{Y}(t, W, V, \tau(X_0)) \qquad P_{(W,V)} - \text{p.s.}$$

The aim of our work is to characterize differential stochastic systems which can be immersed in $(\bar{\Sigma}_{p_1,\dots,p_s})$, and to conclude the existence of finite dimensional filters for (Σ).

Let $\bar{C}(x, y) = y$ and let us denote

$$E_k^i = Span_{I\!R} \left\{ L_F^j \bar{C}_i,\ j \leq k \right\}$$

where

$$F = \begin{pmatrix} f(x, y) \\ h(x, y) \end{pmatrix}$$

$$E = \sum_{i=1}^{s} E_{p_i - 1}^i$$

$$\overline{E} = E \oplus I\!R$$

Let $\bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$ and $E \otimes_{I\!R} \bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$ as in the previous section.

Theorem 3.2

The stochastic differential system (Σ) can be immersed in $(\Sigma_{p_1,\dots,p_s})$ if and only if

i) $\forall i \in \{1, \dots, s\}$

$$L_F^{p_i} \bar{C}_i \in E \otimes_{I\!R} \bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$$

ii) for $i = 1, 2$

$$L_{G_i}(E) \subset \bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$$

where

$$G_1 = \begin{pmatrix} g(x, y) \\ O \end{pmatrix} \text{ and } G_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Demonstration

Let us denote

$$\check{X}(t) = \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}$$

and

$$\check{W}(t) = \begin{pmatrix} W(t) \\ V(t) \end{pmatrix}$$

Sufficiency

Let

$$\tau_{ij}(\check{X}) = L_F^{j-1}\bar{C}_i(\check{X}) \qquad \forall i \in \{1, \ldots, s\} \quad \forall j \in \{1, ..., p_i\}$$

and

$$\tau_i = \begin{pmatrix} \tau_{i1} \\ \vdots \\ \tau_{in_i} \end{pmatrix} \qquad \tau = \begin{pmatrix} \tau_1 \\ \vdots \\ \tau_s \end{pmatrix}$$

$$Z_{ij}(t) = \tau_{ij}(\check{X}(t))$$

We have $Y_i(t) = \bar{C}_i\check{X}(t) = \tau_{i1}(\check{X}(t)) = Z_{i1}(t)$, hence the observation is linear. We have also, about the state $\forall i \in \{1, \ldots, s\} \qquad \forall j \in \{1, \ldots, p_i\}$

$$\begin{aligned} dZ_{ij}(t) &= d\left(L_F^{j-1}\bar{C}_i(\check{X}(t))\right) \\ &= L_F^j\bar{C}_i(\check{X}(t))dt + L_{G_i}L_F^{j-1}\bar{C}_i(\check{X}(t))d\check{W}(t) \\ &= Z_{i,j+1}(t)dt + B_{ij}(Y(t))d\check{W}(t) \quad \text{if } j < p_i \text{ and} \\ &= \varphi_i(Y(t)) + \sum_{k=1}^{s}\sum_{l=1}^{p_k} a_{kl}^i(Y(t))Z_{kl}(t) + B_{in_i}(Y(t))d\check{W}(t) \text{ if } j = p_i \end{aligned}$$

Therefore, we have shown that (Σ) can be immersed by an analytical application τ in $(\bar{\Sigma}_{p_1,\ldots,p_s})$. Let us remark that

$$dY_i(t) = dZ_{i1}(t) = Z_{i2}(t)dt + \sum_{k=1}^{m+s} L_{G_k}\bar{C}_i(\check{X}(t))d\check{W}^k(t)$$

if $k \leq m$ then $L_{G_k}\bar{C}_i(\check{X}(t)) = 0$ and if $k > m$ then $L_{G_k}\bar{C}_i(\check{X}(t)) = \delta_{(k-m),i}$

Necessity

According to the definition of the immersion, we must have for every $k \in \{1, \ldots, s\}$

$$\bar{C}_k(\check{X}(t)) = Z_{k1}(t)$$

Hence, by derivation, we have

$$L_F\bar{C}_k(\check{X}(t)) = Z_{k2}(t)$$

$$L_G\bar{C}_k(\check{X}(t)) = B_{k1}(Y(t))$$

then

$$L_F^2\bar{C}_k(\check{X}(t)) = Z_{k3}(t)$$

$$L_G L_F\bar{C}_k(\check{X}(t)) = B_{k2}(Y(t))$$

and so on... Finally, since $Z_{ij}(t) = L_F^{j-1}\bar{C}_i(\check{X}(t))$, we have

$$L_F^{p_k}\bar{C}_k(\check{X}(t)) = \sum_{i=1}^{s}\sum_{j=1}^{p_i} a_{ij}^k(Y(t))Z_{ij}(t) + \varphi_k(Y(t))$$

$$L_G L_F^{p_k-1}\bar{C}_k(\check{X}(t)) = B_{kp_k}(Y(t))$$

Therefore

$$L_F^{p_k}\bar{C}_k \in E \otimes_{I\!R} \bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$$

and

$$L_G L_F^{j-1}\bar{C}_k \in \bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$$

□

Theorem 3.3

A sufficient condition for a nonlinear system (Σ) to admit a finite dimensional filter is that

$i)\forall i \in \{1, \ldots, s\}$

$$L_F^{p_i}\bar{C}_i \in \overline{E} \otimes_{I\!R} \bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$$

ii) for i=1,2

$$L_{G_i}(E) \subset \bar{C}_k \in \bar{C}^*(\mathcal{C}^\omega(I\!R^s, I\!R))$$

iii) The map

$$\tau = \begin{pmatrix} \tau_{11} \\ \vdots \\ \tau_{sp_s} \end{pmatrix}$$

defined by

$$\forall i = 1, \cdots, s \quad \forall j = 1, \cdots, p_i \qquad \tau_{ij}(\check{x}) = L_F^j\bar{C}_i(\check{x})$$

is injective.

Using theorem 4.1 [14] and theorem 3.2, it is easy to prove the existence of a finite dimensional filter.

4. DE L'USAGE DU THÉORÈME DE GIRSANOV

Nous allons nous intéresser dans ce paragraphe au système stochastique non linéaire (Σ) du type

$$\begin{cases} dx_t = f(x_t)dt + g(x_t)odW_t \\ dy_t = h(x_t)dt + dV_t \end{cases}$$

et nous montrerons que, sous certaines conditions, il existe une loi de probabilité sous laquelle la loi conditionnelle de x_t sachant $\mathcal{F}_t^Y = \sigma\{y_\tau; \tau \leq t\}$ admet un filtre de dimension finie (F.D.F.) ; en d'autres termes, cette loi admet une densité paramétrée, les paramètres étant en nombre fini et étant calculables en temps réel par un procédé récursif. Nous utiliserons pour cela des techniques courantes en automatique non linéaire.

Considérons le système déterministe multi-entrées, multi-sorties (Σ_o) suivant

$$\begin{cases} \dot{x}_t = f(x_t) + \sum_{j=1}^{e} g_j(x_t)u_t^j \\ y_t = h(x_t) \end{cases}$$

où l'état x du système appartient à une variété analytique $\underline{X}$ de dimension finie n, les champs de vecteurs $f, g_1, \ldots, g_e : \underline{X} \longrightarrow T_{\underline{X}}$ étant analytiques, ainsi que la fonction $h : \underline{X} \longrightarrow \mathbb{R}^s$. Les entrées $u^1, \ldots, u^e$ seront supposées continues.

Nous dirons que (Σ_o) peut s'immerger dans un système linéaire $(\tilde{\Sigma}_o)$ s'il existe une application $T : \underline{X} \longrightarrow \mathbb{R}^n$ telle que si $Z_t = T(x_t)$, Z_t soit solution du système linéaire $(\tilde{\Sigma}_o)$, et ceci quelle que soit l'entrée. En d'autres termes, les deux systèmes (Σ_o) et $(\tilde{\Sigma}_o)$ doivent avoir le même comportement entrée-sortie (cf. Claude, Fliess & Isidori).

Commençons par introduire quelques notations et définitions. On notera $\mathcal{L}_f h$ la dérivée de Lie de la fonction h dans la direction du champ de vecteurs f. Si f s'écrit en coordonnées locales

$$f = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i}$$

alors

$$(\mathcal{L}_f h)(x) = \sum_{i=1}^{n} f_i(x) \frac{\partial h}{\partial x_i}(x).$$

D'autre part, on définit $\mathcal{L}_f^0 h = h$ et, pour tout $\nu > 0$, $\mathcal{L}_f^\nu h = \mathcal{L}_f \mathcal{L}_f^{\nu-1} h$.

Définition 4.1

Les nombres caractéristiques associés au système (Σ_o) sont les nombres entiers $n_1, \ldots, n_s$ où, pour tout $k \in \{1, \ldots, s\}$, n_k est défini par

$$\forall j \in \{1,\ldots,e\}\ \forall \nu \in I\!N \qquad \nu < n_k \Longrightarrow \mathcal{L}_{g_j}\mathcal{L}_f^{\nu} h_k \equiv 0$$

et

$$\exists j \in \{1,\ldots,e\}\mathcal{L}_{g_j}\mathcal{L}_f^{n_k} h_k \not\equiv 0$$

Notons que ces nombres ne sont pas nécessairement finis et qu'ils peuvent être nuls.

Définition 4.2

On appelle bouclage par retour d'état l'opération qui consiste à faire le changement de variables suivant

$$\forall j \in \{1,\ldots,e\} u_t^j = \alpha_j(x_t) + \sum_{i=1}^{e} \beta_{ji}(x_t) v_t^i$$

les fonctions α_j, β_{ji} étant analytiques et la matrice $\beta(x) = (\beta_{ji}(x))$ étant inversible quel que soit x dans $\underline{X}$.

Les $v^j, j \in \{1, ..., e\}$, sont donc les nouvelles entrées. Le système obtenu après bouclage par retour d'état s'écrit alors

$$\begin{aligned} \dot{x}_t &= \left(f(x_t) + \sum_{j=1}^{e} g_j(x_t)\alpha_j(x_t) \right) + \sum_{i,j=1}^{e} g_j(x_t)\beta_{ji}(x_t) v_t^i \\ y_t &= h(x_t) \end{aligned}$$

Ce type de changement de variable est très naturel en automatique, le contrôle étant une fonction connue. Pour un système stochastique, ce type de changement de variable peut para^itre beaucoup plus artificiel.

Notons que si le système (Σ_o) est observable, le système obtenu par bouclage par retour d'état reste observable (cf Hermann & Krener).

Nous sommes maintenant en mesure d'énoncer le théorème suivant.

Théorème 4.3 (Claude-Fliess-Isidori)

Si les nombres caractéristiques $n_1, \ldots, n_s$ du système (Σ_o) sont finis, une condition suffisante pour que le bouclage $u_t = \alpha(x_t) + \beta(x_t) v_t$ permette l'immersion de (Σ_o) dans un système linéaire est que

$$\forall k \in \{1,\ldots,s\} \quad \left\{ \begin{array}{l} \exists \Gamma_k \in I\!R^{n_k+1}\ \forall x \in \underline{X} \quad \Delta_k(x)\alpha(x) + \delta_k(x) = \Gamma_k T^k(x) \\ \\ \exists \gamma_k \in I\!R^{e} \quad \forall x \in \underline{X} \quad \Delta_k(x)\beta(x) = \gamma_k \end{array} \right.$$

où $\Delta_k(x) \begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^{n_k} h_k(x) \\ \cdots \\ \mathcal{L}_{g_e}\mathcal{L}_f^{n_k} h_k(x) \end{pmatrix}$, $\delta_k(x) = \mathcal{L}_f^{n_k+1} h_k(x)$ et $T^k(x) = \begin{pmatrix} \mathcal{L}_f^0 h_k(x) \\ \cdots \\ \mathcal{L}_f^{n_k} h_k(x) \end{pmatrix}$

Esquisse de la démonstration

Elle est constructive, en ce sens que nous pouvons exhiber l'application T.

Posons $Z_t = T(x_t)$, T étant défini, pour tout x de $\underline{X}$, par $T(x) = \begin{pmatrix} T^1(x) \\ \cdots \\ T^s(x) \end{pmatrix}$

où chaque bloc $T^k(x)$, pour $k = 1, \ldots, s$, vaut $\begin{pmatrix} T^{k,0}(x) \\ \cdots \\ T^{k,n_k}(x) \end{pmatrix}$ avec $T^{k,\ell}(x) = \mathcal{L}_f^\ell h_k(x)$, ℓ allant de 0 à n_k.

Si on note $Z_t^k = T^k(x_t)$ et $Z_t^{k,\ell} = T^{k,\ell}(x_t)$, pour $k \in \{1, ..., s\}$ et $\ell \in \{0, ..., n_k\}$, alors on peut facilement vérifier que le système vérifié par Z est linéaire, ceci sous les conditions du théorème et par simple bouclage. Après calcul, on obtient le système linéaire suivant

$\forall k \in \{1, ..., s\}$

$$\dot{Z}_t^k = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & & & 1 \\ \Gamma_{k,0} & \cdots & & \cdots & \Gamma_{k,n_k} \end{pmatrix} Z_t^k + \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \cdots & & & \cdots \\ \cdots & & & \cdots \\ \gamma_{k,1} & \cdots & \cdots & \gamma_{k,e} \end{pmatrix} \begin{pmatrix} v_t^1 \\ \vdots \\ v_t^e \end{pmatrix}$$

Remarque 4.4

Regardons plus particulièrement le cas où $\Delta_k(x) = \Delta_k$. Il est clair qu'il suffit alors de choisir $\beta(x) = I$ et donc, il ne reste qu'à trouver une fonction α telle que

$\forall k \in \{1, ..., s\} \quad \forall x \in \underline{X} \quad \Delta_k.\alpha(x) = \Gamma_k.T^k(x) - \delta_k(x)$.

Si une telle fonction existe, chacune de ses composantes α_j s'écrira comme combinaison linéaire des $\mathcal{L}_f^i h_k(x)$, $i \in \{0, ..., n_k + 1\}$.

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité. Soient (W_t) un mouvement brownien à valeurs dans $\mathbb{R}^e$ et (V_t) un autre mouvement brownien, indépendant de (W_t) et à valeurs dans $\mathbb{R}^s$. On considère le système stochastique non linéaire (Σ) suivant

$$\begin{cases} dx_t = f(x_t)dt + g(x_t) \circ dW_t \\ dy_t = h(x_t)dt + dV_t \end{cases}$$

où f,g, et h sont définies comme dans le premier paragraphe. On supposera en outre que f et g vérifient les conditions habituelles pour que l'équation d'évolution admette une unique solution (cf Kallianpur p. 105). La notation $\circ$ désigne ici la différentielle au sens de Stratonovitch, la différentielle au sens de Itô n'étant pas particulièrement maniable avec le calcul sur les variétés en particulier, et la géométrie différentielle en général.

Pour commencer, définissons ce que nous recherchons.

Définition 4.5

Le système (Σ) admet un filtre de dimension finie d si et seulement si il existe une diffusion ξ_t évoluant sur une variété $V \subset I\!R^d$ suivant l'équation différentielle stochastique

$$d\xi_t = a(\xi_t)dt + b(\xi_t)dy_t$$

et des fonctions $f_t : \underline{X} \times V \longrightarrow I\!R$ telles que $f_t(x, \xi_t)$ soient des densités (éventuellement non normalisées) de la loi conditionnelle de x_t sachant $\sigma(y_\tau; \tau \leq t)$.

Un exemple typique est le filtre de Kalman-Bucy lorsque le système (Σ) est linéaire avec une loi initiale gaussienne. On sait alors que la loi conditionnelle de x_t sachant $\sigma(y_s; s \leq t)$ est gaussienne de moyenne m_t et de matrice de covariance Σ_t. La covariance est solution d'une équation de Ricatti indépendante de l'observation, m_t est solution d'une équation stochastique linéaire du type indiqué ci-dessus et ainsi le filtre est de dimension n.

On associe à (Σ) le système (Σ_o) étudié précédemment. Supposons que les conditions du théorème 4.3 soient satisfaites avec Δ_k constante pour tout k de $\{1, \cdots, s\}$. Le théorème de Girsanov nous permet, par un changement de probabilités, d'effectuer une opération analogue au bouclage par retour d'état (mais avec $\beta = I$). Ceci sera détaillé dans le dernier théorème.

Nous allons maintenant donner trois lemmes techniques s'appliquant à α qui seront nécessaires pour pouvoir appliquer le théorème de Girsanov dans le théorème principal.

Lemme 4.6

Supposons que, pour au moins un k de $\{1, \ldots, s\}$,

$$\forall i \in \{0, \ldots, n_k + 1\} \quad \exists K_i > 0 \qquad \| \mathcal{L}_f^i h_k(x) \|^2 \leq K_i(1 + \| x \|^2)$$

alors

$$\exists K > 0 \qquad \| \alpha(x) \|^2 \leq K(1 + \| x \|^2)$$

Lemme 4.7

$$\forall T \in I\!R_+ \qquad \int_0^T \| \alpha(x_t) \|^2 \, dt < \infty$$

Lemme 4.8

Supposons qu'en plus des hypothèses habituelles sur f et g pour que l'équation stochastique $dx_t = f(x_t)dt + g(x_t)odW_t$ ait une unique solution forte, on suppose que

a) $\exists \eta_0 > 0 \qquad E\left[\exp \eta_0 \parallel x_0 \parallel^2\right] < \infty$

b) $\exists K > 0 \quad \forall x \in X \qquad 0 < \parallel g(x) \parallel \leq K$

En restant sous l'hypothèse du lemme 1, $E\left[L_t\right] = 1$ où L_t est l'exponentielle de Girsanov c'est à dire par définition

$$L_t = \exp\left\{\sum_{j=1}^{e} \int_0^T \alpha_j(x_t)dW_t^j - \frac{1}{2}\int_0^T \parallel \alpha(x_t) \parallel^2 dt\right\}$$

La démonstration de ce lemme peut être adaptée de Kallianpur pour le cas où x évoule dans $\mathbb{R}$, mais elle peut aussi être démontrée directement en dimension finie.

Démonstration

Le théorème de Novikov nous indique que si

$$\exists \epsilon > 0 \quad \exists C > 0 \quad \forall s \in [0,t] \qquad E\left[\exp\left\{\epsilon \parallel \alpha(X_s) \parallel^2\right\}\right] \leq C$$

alors $E\left[L_t\right] = 1$, ce que l'on veut montrer. Il suffit donc de trouver ϵ et C. En dimension 1 et sous les hypothèses a) et b), un lemme de Kallianpur [*1981*] montre l'existence d'une constante positive η telle que

$$E\left[\exp\left\{\eta \sup_{s<t} \parallel X_s \parallel^2\right\}\right] < \infty$$

Soit alors $\epsilon = \frac{\eta}{K}$, on a pour tout s de $[0,t]$

$$\begin{aligned} E\left[\exp\left\{\epsilon \parallel \alpha(X_s) \parallel^2\right\}\right] &\leq E\left[\exp\left\{\eta(1+ \parallel X_s \parallel^2)\right\}\right] \\ &\leq e^{\eta} E\left[\exp\left\{\eta \parallel X_s \parallel^2\right\}\right] \\ &\leq e^{\eta} E\left[\exp\left\{\eta \sup_{s<t} \parallel X_s \parallel^2\right\}\right] \quad = C \end{aligned}$$

Il suffit donc de montrer que même en dimension quelconque, on a encore un tel η. On notera dans la suite de cette démonstration

$$|X|_t = \sup_{s<t} \parallel X_s \parallel$$

On a par définition

$$X_t = X_0 + \int_0^t f(X_s)ds + \int_0^t g(X_s)dW_s$$

Notons

$$M_t = \int_0^t g(X_s)dW_s \text{ et } M_t^k = \int_0^t g_k(X_s)dW_s = \sum_{i=1}^{r} \int_0^t g_{ki}(x_s)dW_s^i$$

Alors

$$
\begin{aligned}
< M^k >_t &= \sum_{i,j=1}^{r} \int_0^t g_{ki}(X_s) g_{kj}(X_s) d < W^i, W^j >_s \\
&= \sum_{i=1}^{r} \int_0^t g_{ki}^2(X_s) ds \quad = \int_0^t \| g_k(X_s) \|^2 \, ds \quad \leq K^2 t
\end{aligned}
$$

De plus, la majoration exponentielle nous donne

$$P\{\sup_{s<t} \| M_s^k \| \geq \alpha\} \leq 2 \exp\left\{ \frac{-\alpha}{2K^2 t} \right\}$$

$$P\{|M|_t \geq \alpha\} \geq \sum_{k=1}^{r} P\{\sup_{s<t} \| M_s^k \| \geq \frac{\alpha}{\Gamma}\} \leq 2 \exp\left\{ \frac{-\alpha^2}{2K^2 r t} \right\}$$

Par conséquent, pour tout ϵ positif,

$$
\begin{aligned}
E\left[\exp\left\{\epsilon |M|_t^2\right\}\right] &\leq \sum_{n \in I\!N} (n+1) P\{\exp\{\epsilon |M|_t^2 \geq n\}\} \\
&\leq \sum_{n \in I\!N} (n+1) P\{|M|_t^2 \geq \frac{\log n}{\epsilon}\} \\
&\leq 2 \sum_{n \in I\!N} (n+1) \exp\left\{ \frac{-\log(n)}{2\epsilon r K^2 t} \right\} \\
&\leq 2 \sum_{n \in I\!N} \frac{(n+1)}{n^{\frac{1}{2\epsilon r K^2 t}}} \quad < \infty
\end{aligned}
$$

dès que $\epsilon < \dfrac{1}{4rK^2 t}$. Passons au processus $X_t : \forall s \in [0, t]$

$$
\begin{aligned}
\| X_s \| &\leq \| X_0 \| + |M|_t + \| \int_0^t f(X_s) ds \| \\
&\leq \| X_0 \| + |M|_t + \int_0^t \| f(X_s) \| \, ds
\end{aligned}
$$

mais f est L-Lipschitzienne donc

$$
\begin{aligned}
\| f(x) \| &\leq \| f(x) - f(0) \| + \| f(0) \| \\
&\leq \| f(0) \| + L \| x \|
\end{aligned}
$$

donc

$$
\begin{aligned}
\| X_t \| &\leq \| X_0 \| + |M|_t + t \| f(0) \| + L \int_0^t \| X_s \| \, ds \\
&\leq (\| X_0 \| + |M|_t + t \| f(0) \|) \exp\{Lt\}
\end{aligned}
$$

d'après le lemme de Groonwall. Il en résulte

$$
\begin{aligned}
\eta \| X_t \|^2 &\leq \eta (\| X_0 \| + |M|_t + t \| f(0) \|)^2 \exp\{2Lt\} \\
&\leq 3\eta \left(\| X_0 \|^2 + |M|^2 + t^2 \| f(0) \|^2\right) \exp\{2Lt\}
\end{aligned}
$$

soit en écrivant $C_1 = t^2 \parallel f(0) \parallel^2$ et $C_2 = 3\exp\{2Lt\}$, et du fait que la majoration est uniforme en t,

$$\begin{aligned} E\left[\exp\left\{\eta|X|_t^2\right\}\right] &\leq E\left[\exp\left\{\eta C_1\left(\parallel X_0 \parallel^2 + |M|_t^2 + C_2\right)\right\}\right] \\ &\leq C_3 E\left[\exp\left\{\eta C_1 \parallel X_0 \parallel^2\right\}\exp\left\{\eta C_1 |M|_t^2\right\}\right] \\ &\leq C_3 E\left[\exp\left\{2\eta C_1 \parallel X_0 \parallel^2\right\}\right]^{\frac{1}{2}} E\left[\exp\left\{2\eta C_1 |M|_t^2\right\}\right]^{\frac{1}{2}} \end{aligned}$$

d'après l'inégalité de Schwartz, et en prenant η assez petit dans les deux espérances, on obtient bien

$$E\left[\exp\left\{\eta|X|_t^2\right\}\right] < \infty$$

Nous pouvons maintenant énoncer le principal résultat de ce paragraphe.

Théorème 4.9

Considérons le système (Σ) suivant

$$\begin{cases} dx_t = f(x_t)dt + g(x_t)odW_t \\ dy_t = h(x_t)dt + dV_t \end{cases}$$

et les conditions

(i) (Σ_o), le système déterministe associé à (Σ), est observable

(ii) (Σ_o) s'immerge par bouclage de type $u_t = \alpha(x_t) + v_t$ dans un système linéaire et l'application T vérifie

$$\forall x \in \underline{X} \qquad \det(T'(x)) \neq 0$$

(iii) les conditions techniques suivantes sont vérifiées

a) $\exists \eta_0 > 0 \qquad E\left[\exp \eta_0 \parallel x_0 \parallel^2\right] < \infty$

b) $\exists K > 0 \quad \forall x \in X \qquad 0 < \parallel g(x) \parallel \leq K$

c) $\forall i \in \{0, \dots, n_k + 1\} \quad \exists K_i > 0 \qquad \parallel \mathcal{L}_f^i h_k(x) \parallel^2 \leq K_i(1 + \parallel x \parallel^2)$

Si les conditions (i),(ii), et (iii) sont vérifiées, alors il existe une probabilité $\tilde{P}$ équivalente à P sous laquelle (Σ) admet un F.D.F.

Esquisse de la démonstration

D'après le lemme 4.8, on peut affirmer que l'exponentielle de Girsanov L_T est une P-martingale telle que $E_P[L_T] = 1$ et donc que $\widetilde{P}$ définie par $d\tilde{P} = L_T dP$ est bien une probabilité. Le théorème de Girsanov dit alors que sous $\widetilde{P}$, $\widetilde{W}_t$ défini par

$$d\tilde{W}_t = dW_t - \alpha(x_t)dt$$

est un mouvement brownien. L'équation d'évolution de l'état du système s'écrit alors

$$dx_t = (f(x_t) + g(x_t)\alpha(x_t))dt + g(x_t)od\widetilde{W}_t$$

cette équation admettant toujours une solution forte unique, d'après le lemme 4.6. Mais, d'après la seconde hypothèse du théorème 4.3, ce système s'immerge dans un système linéaire c'est à dire qu'il existe une application T telle que, si $Z_t = T(x_t)$, on a

$$\begin{cases} dZ_t = AZ_t dt + B \circ d\tilde{W}_t \\ dy_t = CZ_t dt + dV_t \end{cases}$$

avec

$$Z_0 = T(x_0)$$

Le théorème de Makowski indique qu'un tel système, linéaire avec une loi initiale non gaussienne, admet un F.D.F. . Pour pouvoir en conclure le résultat annoncé, il reste à remarquer que T est une bijection de $\underline{X}$ sur $T(\underline{X})$. En effet, (Σ_o) étant observable par hypothèse, et compte tenu de la forme explicite de l'immersion, T est injective (cf. Fliess & Kupka). Puisqu'enfin, $det(T'(x)) \neq 0$, pour tout $x \in \underline{X}$, T est bijective et on peut écrire la loi conditionnelle de x_t sachant $\mathcal{F}_t^Y$ en fonction de la loi conditionnelle de Z_t sachant $\mathcal{F}_t^Y$, qui ne dépend que d'un nombre fini de paramètres, ceci sous la loi $\tilde{P}$, et ainsi le système (Σ) admet bien un F.D.F. sous $\tilde{P}$.

Le fait de savoir que le système admet un F.D.F. sous une certaine loi est intéressant car, même si on ne connait pas exactement la loi $\widetilde{P}$, on peut approcher l'exponentielle de Girsanov et trouver ainsi une approximation de F.D.F.

Cependant l'existence d'un véritable F.D.F. ne peut être prouvée que si on peut revenir sous la loi d'origine P. Ceci sera possible dans certains cas. On a par exemple le corollaire suivant.

Corollaire 4.10

Supposons que le système (Σ) se présente sous la forme

$$\begin{cases} dx_t = f(x_t)dt + g(x_t) \circ dW_t \\ y_t = h(x_t) \end{cases}$$

ce qui est souvent le cas en pratique, les hypothèses du théorème 2 étant réunies, en y ajoutant quelques hypothèses techniques dont la suivante qui impose que la matrice (constante)

$$\begin{pmatrix} \mathcal{L}_{g_1}\mathcal{L}_f^{n_1}h_1 & \cdots & \cdots & \mathcal{L}_{g_e}\mathcal{L}_f^{n_1}h_1 \\ \vdots & & & \vdots \\ \mathcal{L}_{g_1}\mathcal{L}_f^{n_s}h_s & \cdots & \cdots & \mathcal{L}_{g_e}\mathcal{L}_f^{n_s}h_s \end{pmatrix}$$

soit de rang plein.
Alors, s'il existe une fonction φ telle que, pour tout x de $\underline{X}$, on ait $\alpha(x) = \varphi(h(x))$, le système admet un F.D.F. sous la loi d'origine P.

Démonstration (abrégée)

On pose $U_0 = 0$ et $dU_t = -\frac{1}{2} \parallel \varphi(Y_t) \parallel^2 dt + \varphi(Y_t)dW_t$. Le système dont l'état est le couple (Z_t, U_t) et l'observation Y_t est conditionnellement gaussien. De plus, pour tout k de $\{1, \ldots, q\}$, la matrice $CA^{n_k}B$ est inversible. Le système conditionnellement gaussien admet donc un filtre fini (de type Kalman Bucy), obtenu après dérivation de l'observation (technique habituelle en cas de dégénerescence de l'observation). Enfin, on se ramène à la loi P en écrivant la formule de Kallianpur-Striebel puis à X_t par le difféomorphisme T.

5. IMMERSION DANS UN SYSTÈME DE BENES

Nous allons maintenant donner une autre façon d'utiliser le théorème précédent: nous allons donner des conditions suffisantes pour qu'un système s'immerge dans un système admettant un filtre fini, mais pas nécéssairement linéaire. Plus précisémment, les systèmes en question sont des systèmes qui possèdent un drift non linéaire qui dérive d'un potentiel. Ces systèmes ont été découverts par Benes dans un célèbre article de 1981. Commençons par regarder le cas mono-dimensionnel, et tout d'abord, rappelons le théorème de Benes dans sa forme la plus simple

Théorème 5.1

Considérons un système du type

$$\begin{cases} dx_t = f(x_t) + dw_t \\ dy_t = x_t + v_t \end{cases}$$

où (w_t) et (v_t) sont deux processus de Wiener indépendants qui évoluent sur $\mathbb{R}$ et x_0 est un point fixé de $\mathbb{R}$ (et $y_0 = 0$).
Si $f'(x) + f^2(x) = ax^2 + bx + c$ avec $a \geq -1$ alors le système admet un filtre de dimension finie.

La démonstration de ce théorème repose essentiellement sur une technique de changement de jauge, i.e. un changement de probabilité sur le bruit d'état. Le système est ainsi linéarisé, donc il admet un filtre fini sous une probabilité équivalente à la probabilité de départ et sous l'hypothèse sur f, l'exponentielle de Girsanov admet une forme simple qui permet de revenir à la probabilité initiale. C'est justement ce qui nous manquait dans le théorème précédent. Il faut remarquer ici, comme l'a fait Lévine, que la notion de transformation de jauge de Benes est identique à celle de bouclage par retour d'état simple.

Si un système non linéaire donné s'immerge par changement de probabilité dans un système linéaire et que la fonction de bouclage (que l'on a appellée α) dérive d'un potentiel à la manière de Benes, alors c'est que le système s'immerge dans un système de type Benes.

Nous allons donner un exemple avant d'énoncer le théorème:

Exemple 5.2

$$\begin{cases} dx_t = \left(x_t + \dfrac{1}{x_t}\right) dt + \sqrt{1+x_t^2}\, dw_t \\ dy_t = \log\left(x_t + \sqrt{1+x_t^2}\right) dt + dv_t \end{cases}$$

qui s'immerge (par un difféomorphisme) dans

$$\begin{cases} d\xi_t = \dfrac{\exp(\xi_t) + \exp(-\xi_t)}{\exp(\xi_t) - \exp(-\xi_t)} dt + dw_t \\ dy_t = \xi_t dt + dv_t \end{cases}$$

qui est un système de type Benes.

Nous allons utiliser le théorème suivant, qui est une généralisation du théorème de Benes et qui un cas particulier du résultat de Haussmann-Pardoux.

Théorème 5.3

Considérons le système suivant

$$\begin{cases} dx_t = (Ax_t + \varphi(x_t))\, dt + B dw_t \\ dy_t = Cx_t dt + dv_t \end{cases}$$

(w_t) et (v_t) étant deux processus de Wiener indépendants à valeurs dans respectivement $I\!R^r$ et $I\!R^q$. Supposons que x_O est une variable aléatoire de loi μ_O, indépendante des deux processus de Wiener précédents.

On suppose aussi, suivant Benes, que φ, application analytique, est telle que il existe une application Φ vérifiant

$$BB^T \Phi_x = \varphi$$

$$2\dot{\Phi} + \mathrm{tr}(B^T B \Phi_{xx}) + 2^t\Phi_x(Ax + \alpha) + |B^T \Phi_x|^2 = x^T \Sigma x + \Gamma^T x + \delta$$

$$\Sigma + C^T C \geq 0$$

Ce système admet un filtre de dimension finie.

A partir de ce théorème, nous pouvons énoncer le résultat suivant

Théorème 5.4

Considérons le système Σ suivant

$$\begin{cases} dx_t = f(x_t)dt + g(x_t) \circ dW_t \\ dy_t = h(x_t)dt + dV_t \end{cases}$$

avec les hypothèses habituelles.

Si les conditions (i), (ii), et (iii) sont satisfaites, alors le système admet un filtre de dimension finie, car il s'immerge dans un système de type Benes.

(i) Le système déterministe Σ associé à $\check{\Sigma}$ est observable

(ii) Σ s'immerge par bouclage de type $u_t = \alpha(x_t) + v_t$ dans un système linéaire, et l'application T vérifie

$$\forall x \in X \qquad \det(T'(x)) \neq 0$$

(iii) Les conditions techniques suivantes sont vérifiées:

a) $\exists \eta_0 > 0 \qquad E\left[\exp\left\{\eta_0 \parallel x_0 \parallel^2\right\}\right] < \infty$

b) $\exists K > 0 \quad \forall x \in X \qquad 0 < \parallel g(x) \parallel \leq K$

b) $\forall i \in \{0, \ldots, n_k + 1\} \quad \exists K_i > 0 \parallel L^i h_k(x) \parallel^2 \leq K_i(1 + \parallel x \parallel^2)$

(iv) La fonction φ définie par $\varphi(\xi) = \alpha \circ T(\xi)$ vérifie les hypothèses du théorème précédent.

RÉFÉRENCES

[1] **Benes V.E.** Exact finite dimensional filters for certain diffusions with nonlinear drift. *Stochastics 5 (1981) p65–92*

[2] **D. Bossanne, D. Rakotopara, J.P. Gauthier** Local and global immersion into linear systems up to output injection. Submitted to *Math. for Control, signals and systems*

[3] **E. Busvelle, D. Rakotopara, D. de Brucq** De l'usage du théorème de Girsanov et de ses applications *Compte-rendu des séminaires de Mathématiques 1988-89 de Rouen, Presses Universitaires de Rouen*

[4] **Claude D., Fliess M., Isidori A.** Immersion directe et par bouclage d'un système non linéaire dans un linéaire. *C.R. Acad. Sc. Paris Série I* (1983) 296 p237–240

[5] **Cohen de Lara M., Lévine J.** Deterministic feedback linearization, Girsanov transformations and finite- dimensional filters *Syst. & Cont. Letters 13 (1989) p81–92*

[6] **H. Hammouri, J.P. Gauthier** The time varying linearization up to output injection. 28^{th}*CDC Tampa, florida, U.S.A.*, December *13-15, 1989*

[7] **U.G.Haussmann, E.Pardoux:** A Conditionally Almost Linear Filtering Problem with Non-Gaussian Initial Condition. *Stochastics* **23** *1988* p241–275

[8] **Ikeda N., Watanabe S.** *Stochastic Differential Equations and Diffusion Processes* (1981) North-Holland Publishing Company

[9] **Kallianpur G.** *Stochastic Filtering Theory* (1980) Springer-Verlag

[10] **A.J. Krener, A. Isidori** Linearization by output injection and nonlinear observers. *Syst. and Cont. Let.* **3** *1985* p47–52

[11] **A.J.Krener, W. Respondek** Nonlinear observers with linearizable error dynamics. *SIAM Journal on control and Opt.* **23** *1985* p139–149

[12] **Lévine** Finite dimensional filters for a class of nonlinear systems and immersion in a linear system *SIAM J. Cont. & Optimization 25 (1987) 6*

[13] **Lévine** Finite dimensional realizations of stochastic P.D.E.'s and application to filtering *Preprint*

[14] **Levine J., Marino R.** Non linear system immersion, observers and finite-dimensional filters *Syst. & Contr. Letters* **7** 1986) p133–142

[15] **Makowski A. M.,** Results on filtering problem for linear systems with non-Gaussian initial conditions. *Proc. 21st IEEE Conference on Decision and Control, Orlando, FL* (1982) p101–104

[16] **Pardoux E.** Filtrage non linéaire et équations aux dérivées partielles stochastiques associées *Cours de St Flour 1989, à paraître chez Springer-Verlag*

[17] **D. Rakotopara, E. Busvelle, H. Hammouri** Immersion in conditionally gaussian systems and finite dimensional filters. 29^{th}*IEEE Conf. on Dec. and Cont.*

[18] **Wong W.S.** New classes of finite dimensional nonlinear filters *Syst. & Cont. Letters 3 (1983) p155–164*

[19] **Yau S.S.T.** Recent results on nonlinear filtering: new class of finite dimensional filters *proceedings of the* 29^{th}*IEEE Conf. on Dec. and Cont.*

PARTICLE FILTERING DRIVEN BY OBSERVATION

E.BUSVELLE*, **D.RAKOTOPARA***, **D.de BRUCQ**†

* SHELL Research, Centre de Recherche de Grand-Couronne,
76530 Grand-Couronne, FRANCE

† LACIS-ITEPEA, Universit de Rouen, BP118
76134 Mont Saint Aignan Cedex, FRANCE

INTRODUCTION

The principle of a particle method is to approach a probability law (in general absolutely continuous w.r.t. Lebesgue measure) by a weighted sum of Dirac measure. When applied to filtering, this just consists in approaching the law of the current state knowing observations by a particular weighted sum of Dirac distributions. This kind of method is well adapted to the case where the dimension of the state is large, because in this case, one usually uses the Monte-Carlo method to compute the conditional expectation

$$E\left[\phi(X_t) \mid \mathcal{F}_t^Y\right] = \int \phi(x) p_t(x) dx$$

and this method is a particle method.

To characterize a particle method, it is sufficient to give some rules like

1) how to calculate weights of particles

2) how to move particles in the state space

In this paper, we want to study nonlinear systems with discrete-time observations. We will consider at first that the state is also in discrete-time.

Many authors ([5],[6]) have studied systems of the following form

$$\begin{cases} X_{t+1} &= f(X_t) \\ Y_t &= h(X_t) + V_t \end{cases}$$

The only unknown variable is the initial state X_0, and we suppose that the law of X_0 is $P_0 << \lambda$. This system is equivalent to

$$Y_t = h \circ f^t(X_0) + V_t$$

and one can check that if we set $q_0 = \frac{dP_0}{dx}$ and

$$q_t(x) = f_{Y_t}^{X_0=x}(y_t) q_{t-1}(x)$$

then q_t is a non-normalised density of the law of X_0 knowing $\mathcal{F}_Y^t$, and

$$E\left[\phi(X_t) \mid \mathcal{F}_Y^t\right] = \frac{\int \phi(f^t(x))q_t(x)dx}{\int q_t(x)dx}$$

Suppose that we approach P_0 by a weighted sum of Dirac distributions, i.e.

$$P_0 \cong \sum_{i=1}^{N} a_0^i \ \delta_{\{x_i\}}$$

then the previous equation gives us an algorithm to compute weights

$$a_t^i = a_{t-1}^i f_{Y_t}^{X_0=x_i}(y_t)$$

So we can write

$$E\left[\phi(x_t) \mid \mathcal{F}_Y^t\right] \cong \sum_{i=1}^{N} a_t^i \ \phi \circ f^t(x_i)$$

and then calculate any moment of the conditional law, or any other quantity which depends on this law, like confidence intervals.

In the first part of the next paragraph, we will give a general theorem of simulation, then a first result of filtering by particle method. We will also give a more interesting method for practical cases, and we will show that this method converge in a certain sense. Then we will give some simulation results based on a chemical problem of estimation of kinetic constants in a chemical reaction.

The originality of this paper is that our algorithm takes into account observations to simulate random particle movements.

1. FILTERING USING SIMULATION

At first, we will give a simulation theorem. For applications of this classical result, see for example [1].

Theorem 1.1

Let $(\Omega, \mathcal{A}, \nu)$ be a measurable set, and let f be a density of probability on $(\underline{X}, \mathcal{X}, \mu)$ given by

$$f(x) = \int_\Omega g_\omega(x)\rho_\omega(x)h(\omega)d\nu(\omega)$$

where

- $g_.(.)$ is a transition of probability from $(\Omega, \mathcal{A})$ to $(\underline{X}, \mathcal{X})$
- $\rho_\omega(x) \in]0,1]$ for almost any ω in Ω and almost any x in $\underline{X}$
- $h \in L_+^1(\Omega, \mathcal{A}, \nu)$ (i.e. h is $\mathbb{R}^+$-valued and we will note $\nu(h) = \int h d\nu$)

Let $\mathcal{P}$ be the following algorithm

(1) Generate ω with respect to $\frac{h(\omega)}{\nu(h)}$
(2) Generate u w.r.t. $\mathcal{U}[0,1]$
(3) Generate x w.r.t. $g_\omega(x)$
(4) If $\rho_\omega(x) < u$, then go to (1), else return x
Then $\mathcal{P}$ generates random variables X w.r.t. the density $f(x)$.

The proof of this classical result can be found in R.Y. Rubinstein ([8]).

To simplify, we will study a nonlinear filtering problem with linear observation of the following form

$$\begin{cases} X_{t+1} &= f(X_t) + g(X_t)W_t \\ Y_t &= C_t X_t + V_t \end{cases}$$

where $(W_t)_{t\in I\!N}$ and $(V_t)_{t\in I\!N}$ are two independants gaussian white noises. The limitation to a linear observation function is not necessary for the proof of our theorem but is a simplification when one want to implement our algorithm, as we will see later. A particle method for this kind of system when the state is $I\!R$-valued is studied in [3]. The conditional law P_t is given by the well-known Zakai equation ([2])

$$p_t(x) = \frac{1}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y_t)} f_{Y_t}^{X_t=x}(y_t) \int_{\underline{X}} f_{X_t}^{X_{t-1}=\xi}(x) p_{t-1}(\xi) d\xi$$

where p_t is the density of P_t w.r.t. Lebesgue measure i.e.

$$dP_t(x) = p_t(x)dx$$

We suppose in this paper that the following hypothesis is verified:

(H) *The density p_t exist for any $t > 0$*

We can find in [2] some conditions which implies this hypothesis.

The Zakai equation can be written more simply by setting

$$\begin{aligned} g_t(\xi, x) &= \frac{f_{Y_t}^{X_t=x}(y_t) f_{X_t}^{X_{t-1}=\xi}(x)}{f_{Y_t}^{X_{t-1}=\xi}(y_t)} = f_{X_t}^{X_{t-1}=\xi, Y_t=y_t}(x) \\ K_t &= \frac{\sup_{u\in\underline{X}} \left(f_{Y_t}^{X_{t-1}=u} \right)}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y_t)} \\ \rho_t(\xi) &= \frac{f_{Y_t}^{X_{t-1}=\xi}(y_t)}{\sup_{u\in\underline{X}} \left(f_{Y_t}^{X_{t-1}=u} \right)} \\ h_t(\xi) &= K_t p_{t-1}(\xi) \end{aligned}$$

where we have noted $Y^t = (y_1, \ldots, y_t)$

in the following form

$$p_t(x) = \int g_t(\xi, x) \rho_t(\xi) h_t(\xi) d\xi$$

where g_t, ρ_t, and h_t satisfy hypotheses of the previous theorem, for any $y^t \in \underline{Y}^t$. If it was possible to have an infinite-range sample of random variables with the law p_{t-1}, then algorithm $\mathcal{P}$ would give us a sample with the same range of random variables with the law p_t. This sample permitted us to calculate for any function $\phi \in L^1(\underline{X})$

$$E\left[\phi(X_t) \mid \mathcal{F}_Y^t\right]$$

almost surely. In practice, we can only compute a finite number N_{t-1} of random variables at time $t-1$, which permits us to calculate $N_t \leq N_{t-1}$ random variables at time t. So there exists a fatal time T for which $N_T = 0$. We will now give a detailed algorithm $\mathcal{P}_N$ and the proof of the convergence of this algorithm when N increase to infinity.

For any N in $I\!N$, let $\mathcal{P}_N$ be the following algorithm

Initialisation

$(x_0^n)_{I\!N}$ is a sample of infinite range such that for any n in $I\!N$, x_0^n is a random variable w.r.t. the law P_0.

$(e_0^n)_{I\!N}$ is defined by $\quad \forall n \in I\!N \quad e_0^n = 1.$

$$\chi_N^n = \begin{cases} 1 & \text{if } n \leq N \\ 0 & \text{elsewhere} \end{cases}$$

Loop

For any $n \in I\!N$, x_t^n is a random variable w.r.t. the density $g_t(x_{t-1}^n, .)$.

For any $n > 0$, e_t^n is equal to 1 with probability $\rho_t(x_{t-1}^n)$ if $e_{t-1}^n = 1$ and 0 in other cases, and e_t^0 is always equal to 1.

e_t^n represents the state of the n^{th} particle at time t

$e_t^n = 1$ means that the particle is alive,

$e_t^n = 0$ means that the particle is dead.

The random probability given by the algorithm $\mathcal{P}_N$ is defined by the empirical formula

$$P_t^N = \frac{\sum_{n \in I\!N} \chi_N^n e_t^n \delta_{\{x_t^n\}}}{\sum_{n \in I\!N} \chi_N^n e_t^n}$$

For any given N, the algorithm is feasible since χ_N^n is a finite number which corresponds to the number of simulations needed to calculate the evolution of particles from $t-1$ to t. These simulations are very easy to do, because random variables are Gaussian, (thanks to the fact that the observation function is linear) so we can use a classical algorithm like Sibuya ([4]) to do this.

Theorem 1.2

If P_t^N is the empirical law generated by $\mathcal{P}_N$, then

$$P_t^N \to P_t \qquad \text{weakly a.s.}$$

The notation $P_t^N \to P_t \quad$ weakly a.s. means that the family of probability $(P_t^N)_{I\!N}$ converges in a weak sense to the probability P_t almost surely, i.e. for any bounded and continuous function φ

$$< \varphi, P_t^N > \to < \varphi, P_t > \qquad \text{a.s.}$$

where $< \varphi, P >= \int \varphi dP$.

Proof

The first step of the demonstration consists in verifying that, for any $n \geq 1$,

$$\mathcal{L}(x_t^n \mid e_t^n = 1) = P_t$$

We will show this by using a recursive technique for $n = 1$ since all (x_t^n, e_t^n) for $n \geq 1$ have the same law. We have $\mathcal{L}(x_0^1 \mid e_0^1 = 1) = \mathcal{L}(x_0^1) = P_0$. Suppose that $\mathcal{L}(x_{t-1}^1 \mid e_{t-1} = 1) = P_{t-1}$, and let $\underline{X}^1$ be a borel set of $\underline{X}$. All computations are made conditionally to y^t. We have

$$P\{x_t^1 \in \underline{X}^1 \mid e_t^1 = 1\} = \frac{P\{x_t^1 \in \underline{X}^1 \text{ and } e_t^1 = 1 \mid e_{t-1}^1 = 1\}}{P\{e_t^1 = 1 \mid e_{t-1}^1 = 1\}}$$

The numerator $P\{x_t^1 \in \underline{X}^1$ and $e_t^1 = 1 \mid e_{t-1}^1 = 1\}$ can be rewritten formally

$$\int_{\underline{X}} P\{x_t^1 \in \underline{X}^1 \text{ et } e_t^1 = 1 \mid x_{t-1}^1 = \xi \text{ et } e_{t-1}^1 = 1\}$$

$$P\{x_{t-1}^1 = \xi \mid e_{t-1}^1 = 1\} d\xi$$

and with a rigorous form

$$\int_{\underline{X}} \int_{\underline{X}^1} g_t(\xi, x) dx \ \rho_t(\xi) p_{t-1}(\xi) d\xi \quad = \frac{1}{K_t} \int_{\underline{X}^1} p_t(x) dx$$

The denominator can be easely computed by setting $\underline{X}^1 = \underline{X}$ and is equal to $\left(\frac{1}{K_t}\right)$. Finally, we obtain

$$P\{x_t^1 \in \underline{X}^1 \mid e_t^1 = 1\} = P_t\{\underline{X}^1\}$$

The second step of the demonstration consists in applying the strong law of large numbers. To do this, let us remark that

$$< \varphi, P_t^N >= \frac{\sum_{n \in I\!N} \chi_N^n e_t^n \varphi(x_t^n)}{\sum_{n \in I\!N} \chi_N^n e_t^n} = \frac{< \varphi, Q_t^N >}{< 1, Q_t^N >}$$

where

$$Q_t^N = \frac{\sum_{n \in \mathbb{N}} \chi_N^n e_t^n \delta_{\{x_t^n\}}}{N \ P\{e_t^1 = 1\}}$$

For any $n \geq 1$, we have

$$\begin{aligned} E\left[e_t^n \varphi(x_t^n)\right] &= E\left[e_t^n \varphi(x_t^n) \mid e_t^n = 1\right] P\{e_t^n = 1\} \\ &\quad + \ E\left[e_t^n \varphi(x_t^n) \mid e_t^n = 0\right] P\{e_t^n = 0\} \\ &= E\left[\varphi(x_t^n) \mid e_t^n = 1\right] P\{e_t^n = 1\} \\ &= P\{e_t^1 = 1\} \ < \varphi, P_t > \end{aligned}$$

Therefore, according to the strong law of large numbers

$$\begin{aligned} < \varphi, Q_t^N > &= \frac{1}{N}\varphi(x_t^0) + \frac{1}{N}\sum_{n=1}^{N} \frac{e_t^n}{P\{e_t^1 = 1\}}\varphi(x_t^n) \\ &\rightarrow < \varphi, P_t > \quad \text{a.s.} \end{aligned}$$

and $< 1, Q_t^N > \rightarrow < 1, P_t > = 1$ a.s. in the same way.

So $< \varphi, P_t^N > \rightarrow < \varphi, P_t >$ a.s. □

2. PARTICLES DRIVEN BY OBSERVATIONS

In this part, our aim is to solve the problem of death of particles. Hence we will replace coefficients e_t^n which take their values in $\{0, 1\}$ by coefficients b_t^n which take their values in $\mathbb{R}_+$. These numbers will represent the degree of confidence in each particle.

Let (Σ) be the following system

$$(\Sigma) \qquad \left\{ \begin{array}{rcl} X_{t+1} &=& f(X_t) + g(X_t)W_t \\ Y_t &=& C_t X_t + V_t \end{array} \right.$$

Suppose that (W_t) and (V_t) are two sequences of independent Gaussian variables, with covariance matrices identity and Q respectively. Suppose too that X_0 follow the law P_0. We define an algorithm $\mathcal{P}$ which consists in a way in simulating trajectory of particles x_t^n and weight coefficients b_t^n. The law truncated at N particles given by $\mathcal{P}$ is noted P_t^N and defined by

$$P_t^N = \frac{\sum_{n=1}^{N} b_t^n \delta_{\{x_t^n\}}}{\sum_{n=1}^{N} b_t^n} = \sum_{n=1}^{N} a_t^{N,n} \delta_{\{x_t^n\}}$$

where $a_t^{N,n}$ is the relative weight of the n^{th} particle

$$a_t^{N,n} = \frac{b_t^n}{\sum_{m=0}^{N} b_t^m}$$

Initialisation

x_0^n is a random variable w.r.t. P_0.

$b_0^n = 1$ by definition.

Loop

x_t^n is a Gaussian variable w.r.t.$P_{X_t}^{X_{t-1}=x_{t-1}^n, Y_t=y_t}$

b_t^n is defined by

$$b_t^n = b_{t-1}^n \; f_{Y_t}^{X_{t-1}=x_{t-1}^n}(y_t)$$

We remark that this algorithm is easy to implement on a computer, in particular on parallel computer.

Theorem 2.1

Let us consider the system (Σ). Let P_t be the conditional law of X_t knowing $\mathcal{F}_t^Y$. If P_t^N represents the law given by the algorithm $\mathcal{P}$, truncated to N particles, then we have

$$P_t^N \to P_t \qquad \text{weakly a.s.}$$

Proof

Set

$$g_t(\xi, x) = f_{X_t}^{X_{t-1}=\xi, Y_t=y_t}(x)$$

and

$$h_t(\xi) = \frac{f_{Y_t}^{X_{t-1}=\xi}(y_t)}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y_t)}$$

All calculus are made for a given trajectory of the observation. P_0^N is the empirical law, obtained by a N-sample of law P_0. It is clear that P_0^N converges weakly almost surely to P_0. Suppose by induction that P_{t-1}^N converges to P_{t-1} weakly a.s. We will note to simplify $a_t^{N,n} = a_t^n$. Let $\varphi \in C_b(\underline{X})$

$$\mathcal{E}_t^N(\varphi) = \mathrm{E}\left[\varphi(X_t) \mid \mathcal{F}_t^Y\right] - \sum_{n=1}^{N} a_t^n \varphi(x_t^n)$$

$$= \int_{\underline{X}} \int_{\underline{X}} \varphi(x) g_t(\xi, x) h_t(\xi) \left(p_{t-1}(\xi) - p_{t-1}^N(\xi)\right) d\xi dx$$

$$+ \sum_{n=1}^{N} \int_{\underline{X}} \varphi(x) g_t(x_{t-1}^n, x) \left(h_t(x_{t-1}^n) a_{t-1}^n - a_t^n\right) dx$$

$$+ \sum_{n=1}^{N} a_t^n \left(\int_{\underline{X}} \varphi(x) f_{X_t}^{Y_t=y_t, X_{t-1}=x_{t-1}^n}(x) dx - \varphi(x_t^n) \right)$$

then

$$\mathcal{E}_t^N(\varphi) = \mathcal{E}_{t-1}^N\left(\int_{\underline{X}} \varphi(x) g_t(\xi, x) h_t(\xi) dx\right)$$

$$+\sum_{n=1}^{N} a_{t-1}^n \int_{\underline{X}} \varphi(x) g_t(x_{t-1}^n, x) dx \ f_{Y_t}^{X_{t-1}=x_{t-1}^n}(y_t)$$

$$\left(\frac{1}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y_t)} - \frac{1}{\sum_{n=1}^{N} a_{t-1}^n f_{Y_t}^{X_{t-1}=x_{t-1}^n}(y_t)}\right)$$

$$+\sum_{n=1}^{N} a_t^n \left(\mathrm{E}\left[\varphi(X_t) \mid X_{t-1} = x_{t-1}^n, Y_t = y_t\right] - \varphi(x_t^n)\right)$$

The first term converges to 0, by induction and since the function of x is bounded and continuous. The last term can be rewritten in the following form

$$\frac{\sum_{n=1}^{N} b_t^n \left(\mathrm{E}\left[\varphi(X_t) \mid X_{t-1} = x_{t-1}^n, Y_t = y_t\right] - \varphi(x_t^n)\right)}{\sum_{n=1}^{N} b_t^n}$$

After normalisation of numerator and denominator by $\frac{1}{N}$ and by the strong law of large numbers, it appears that this term converges almost surely to

$$\frac{\mathrm{E}\left[b_t^1 \left(\mathrm{E}\left[\varphi(X_t) \mid X_{t-1} = x_{t-1}^1, Y_t = y_t\right] - \varphi(x_t^1)\right)\right]}{\mathrm{E}\left[b_t^1\right]}$$

Conditionally to $\sigma_{t-1} = \sigma(x_s^1; 0 \leq s \leq t-1) \vee \mathcal{F}_t^Y$,
b_t^1 is a constant, and then we obtain

$$\mathrm{E}\left[\left(\mathrm{E}\left[\varphi(X_t) \mid X_{t-1} = x_{t-1}^1, Y_t = y_t\right] - \varphi(x_t^1)\right) \mid \sigma_{t-1}\right]$$

which is equal to zero. For the middle term, which corresponds to the error of normalisation, we remark that

$$\sum_{n=1}^{N} a_{t-1}^n f_{Y_t}^{X_{t-1}=x_{t-1}^n}(y_t) = \int_{\underline{X}} f_{Y_t}^{X_{t-1}=\xi}(y_t) p_{t-1}^N(\xi) d\xi$$

which converges almost surely, by the induction hypothesis, to

$$\int_{\underline{X}} f_{Y_t}^{X_{t-1}=\xi}(y_t) p_{t-1}(\xi) d\xi = f_{Y_t}^{Y^{t-1}=y^{t-1}}(y_t)$$

and so,

$$\frac{1}{f_{Y_t}^{Y^{t-1}=y^{t-1}}(y_t)} - \frac{1}{\sum_{n=1}^{N} a_{t-1}^n f_{Y_t}^{X_{t-1}=x_{t-1}^n}(y_t)}$$

converges almost surely to 0. □

We will see a direct application of this result in the next section. If we want to extend this method for other cases, The main difficulty is to find a way to simulate explicitely laws for more general form of systems. However, it is possible to apply our method to the more general system

$$\begin{cases} x_{t+1} = f(t, x_t) + g(t, x_t)w_t \\ \theta_{t+1} = k \text{ with probability } p_k(t, x_t, \theta_t) \\ y_t = C(t, \theta_t)x_t + \nu(\theta_t)v_t \end{cases}$$

This kind of model correspond to a nonlinear system with change of model, and is outside of the previous theory. However, it is easy to check that we can apply the same algorithm and that these non-Gaussian laws can be simulated.

If the state is a diffusion process (in continuous time) then the system can be written in the following form

$$\begin{cases} dX_t &= f(X_t)dt + g(X_t)dW_t \\ Y_k &= h(X_{t_k}) + V_k \end{cases}$$

Suppose that there exists a strong solution for the first equation, and that all the laws are absolutely continuous with respect to the Lebesgue measure. For the algorithm $\mathcal{P}$, the only difference is in the way to generate a random variable w.r.t. the law $P_{X_{t_k}}^{X_{t_{k-1}}=\xi, Y^k=y^k}$. But the Bayes formula gives

$$f_{X_{t_k}}^{X_{t_{k-1}}=\xi, Y_k=y_k}(x) = \frac{f_{Y_k}^{X_{t_k}=x}(y_k)\, f_{X_{t_k}}^{X_{t_{k-1}}=\xi}(x)}{\int_{\underline{X}} f_{Y_k}^{X_{t_k}=u}(y_k) f_{X_{t_k}}^{X_{t_{k-1}}=\xi}(u)du}$$

The law which must be simulated is the product of a density by a bounded function. We can apply the rejection method to obtain x_{t_k} knowing $x_{t_{k-1}}$ This is an application of Theorem 1. The evolution of the diffusion can be simulated by the results in [7]. Theorem 3 is always applicable.

3. APPLICATIONS

Application 3.1

We consider a chemical reaction given by

$$\begin{cases} A & \xrightarrow{\lambda} C \\ A + B & \xrightarrow{\mu} D \end{cases}$$

where λ and μ are kinetic constants to estimate. Outputs are

- a = concentration of A
- b = concentration of B

λ and μ are functions of temperature. Equations of this reaction can be written in the following form

$$\begin{cases} \dot{a}_t = -\lambda_t a_t - \mu_t a_t b_t \\ \dot{b}_t = -\mu_t a_t b_t \\ d\lambda_t = \sigma dV_t \\ d\mu_t = \sigma dW_t \end{cases}$$

where a and b are measured in discrete time. The system can be write in the following generic form

$$\begin{cases} x_{t_{k+1}} = x_{t_k} + \int_{t_k}^{t_{k+1}} f(x_s)ds + \int_{t_k}^{t_{k+1}} Bdw_s \\ y_k = Cx_{t_k} + v_k \end{cases}$$

with

$$x = (\ a\ ,\ b\ ,\ \lambda\ ,\ \mu\ )^T$$

$$f(x) = (\ -\lambda a - \mu ab\ ,\ -\mu ab\ ,\ 0\ ,\ 0\ )^T$$

and

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \sigma & 0 \\ 0 & \sigma \end{pmatrix}$$

We propose the following discretization scheme for the continuous time equation

$$x_{t_{k+1}} = \Phi(t_k, t_{k+1}, x_{t_k}) + \frac{\partial \Phi(t_k, t_{k+1}, x_{t_k})}{\partial x} B\sqrt{t_{k+1} - t_k} w_k$$

where $\Phi(s,t,x)$ is the solution of

$$\begin{cases} \dot{x}_t = f(x_t) \\ x_s = x \end{cases}$$

at time t.

The right part of this scheme is the first order developpement of

$$\Phi(t_k, t_{k+1}, x_{t_k} + \int_{t_k}^{t_{k+1}} Bdw_s)$$

which come naturally from the diffusion equation. A classical theorem of probability (see e.g. Gikhman-Skorokhod [11]) shows that our scheme converge in law to the solution of the diffusion equation when the step of the discretization goes to zero.

Our principal goal is to estimate λ,μ, and their confidence intervals.

If we solve equations, we can see that for each particle ξ_t at time t and for each weight b_t, we have, thanks to the algorithm of Theorem 3

- **prediction** between t_k and t_{k+1}

$$\begin{aligned}
\dot{\xi}_t &= f(\xi_t) \\
\dot{P}_t &= F(\xi_t)P_t \ + \ P_t F(\xi_t)^T \ + \ Q \\
\dot{b}_t &= 0
\end{aligned}$$

- **Correction** at time t_k and t_{k+1}

$$\begin{aligned}
\xi_{t_k} &= \xi_{t_k^-} \ + \ P_{t_k^-} C_{t_k^-}^T (C_{t_k^-} P_{t_k^-} C_{t_k^-}^T + R)^{-1} (y_k - C_{t_k} x_{t_k^-}) \\
&\quad + \ \left(P_{t_k^-} - P_{t_k^-} C_{t_k^-}^T (C_{t_k^-} P_{t_k^-} C_{t_k^-}^T + R)^{-1} C_{t_k^-} P_{t_k^-} \right) \bar{w}_k \\
P_{t_k} &= BB^T (t_{k+1} - t_k) \\
b_{t_k} &= b_{t_{k-1}} \frac{\exp\left(-\frac{1}{2} (y_k - C_{t_k} x_{t_k^-})^T (C_{t_k^-} P_{t_k^-} C_{t_k^-}^T + R)^{-1} (y_k - C_{t_k} x_{t_k^-}) \right)}{\sqrt{2\pi . det(C_{t_k^-} G(\xi_{t-1}) C_{t_k^-}^T + R)}}
\end{aligned}$$

and $\bar{w}_k$ is a Gaussian white noise. Numerical calculus are made with these formulas. We have plot logarithms of kinetic constants (black line), estimation of these constants (dotted line), and confidence intervals (grey lines), i.e. at any time t, the probability for the true value of the kinetic constant to be between the upper and the lower curve is equal to 0.90, on both figures 2.1 and 2.2.

CHAPITRE V

REFERENCES

[1] **P. Bratley, B.L. Fox, L.E. Schrage** *A Guide to Simulation* Springer-Verlag *1983*

[2] **G.B. Di Masi, W.J. Runggaldier** On Measure Transformations for Combined Filtering and Parameter Estimation in Discrete Time. *Systems and Control Letters* **2** (*1982*) n°1 pp. 57–62

[3] **F. Campillo** La mthode de Gauss-Galerkin en filtrage non linaire. M^2AN *Modlisation mathmatique et analyse numrique, AFCET* **20** (*1986*) n°2 pp. 203–223

[4] **G. Grancher** Simulation de lois de probabilit. *Rapport technique LAMS, URA* n° *1378 1987*

[5] **F. Le Gland, F. Campillo** Application du filtrage non linaire en trajectographie passive. 12ième *colloque GRETSI, 1989*

[6] **E. Pardoux** Filtrage non linaire et quations aux drives partielles stochastiques associes. *Cours au sminaire d't de St-Flour 1989*

[7] **E. Pardoux, D. Talay** Discretization and Simulation of Stochastic Differential Equations. *Acta Applicandae Mathematicae* **3** (*1985*) pp. 23–47

[8] **M.H.A. Davis** New Approach to Filtering for Nonlinear Systems. *IEEE Proceedings, Vol. D-128 1981* pp. 166–172

[9] **F. LeGland** Monte carlo Methods in Nonlinear Filtering *Proceedings* 23rd *IEEE CDC (Las Vegas) 1984* pp. 31–32

[10] **R.Y. Rubinstein** Simulation and the Monte Carlo Method. *1981*

[11] **Gikhman-Skorokhod** Introduction to the theory of random processes. *1965*

Cette thèse est consacrée à la recherche de systèmes admettant un *filtre de dimension finie* . Les outils utilisés sont issus de la théorie des probabilités et de la géométrie différentielle. Il est surtout donné des conditions suffisantes d'existence de filtres finis pour des systèmes non linéaires, condition évidemment très fortes, mais *explicites et constructives* . Une utilisation simple et originale de l'équation de Zakaï discrète est ensuite présentée, qui permet d'approcher numériquement sa solution pour une très large classe de systèmes. Des simulations de ces différentes approches ont été effectuées et sont présentées à la fin de cette étude.