

**Filtre de Kalman discret pour l'estimation des moyennes
inconnues de bruits blancs**

par

Becem Saidani

Mémoire présenté au Département de mathématiques
en vue de l'obtention du grade de maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, juin 2012



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-88826-1

Our file Notre référence

ISBN: 978-0-494-88826-1

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Le 8 juin 2012

*le jury a accepté le mémoire de Monsieur Becem Saidani
dans sa version finale.*

Membres du jury

Professeur Éric Marchand
Directeur de recherche
Département de mathématiques

Professeur Maxime Descoteaux
Codirecteur de recherche
Département d'informatique

Professeur François Dubeau
Codirecteur de recherche
Département de mathématiques

Professeur Taoufik Bouezmarni
Membre
Département de mathématiques

Professeur Jean-Marc Belley
Président rapporteur
Département de mathématiques

*À ma mère
mon père
et mes soeurs*

SOMMAIRE

Le filtre de Kalman consiste à estimer l'état d'un système dynamique évoluant au cours du temps à partir d'observations partielles et généralement bruitées. Typiquement, on dispose d'une suite (Y_1, Y_2, \dots, Y_n) d'observations, obtenues après un traitement préalable du signal brut au niveau des capteurs, telle que chaque observation Y_n est reliée à l'état inconnu X_n de façon linéaire. Le but sera d'estimer l'état X_n de façon optimale et récursive.

Dans ce mémoire on va étudier la théorie de base du filtre de Kalman sur des modèles statistiques avec des bruits de type additif (gaussien, exponentiel ...), ensuite nous allons proposer une adaptation du filtre de Kalman dans le cas où les bruits du modèle statistique sont à moyennes inconnues.

REMERCIEMENTS

Je voudrais commencer par remercier mon co-directeur de recherche le professeur François Dubeau pour m'avoir aidé, à travers ce mémoire, à donner le meilleur de moi-même. Je lui suis particulièrement reconnaissant pour sa disponibilité, les efforts qu'il a investi dans ce travail et pour le soutien financier tout au long de la réalisation de ce travail. C'est une chance mais surtout un réel plaisir de travailler avec lui.

Je voudrais également remercier les professeurs Éric Marchand et Maxime Desco-teaux qui furent très généreux de leur temps et m'ont permis de me familiariser avec une série de nouvelles notions. La qualité de leur encadrement est d'une grande valeur ajoutée à ce travail, mes remerciements vont également au département de mathématiques de l'université de Sherbrooke ainsi qu'au professeur Éric Marchand pour le soutien financier.

Je tient aussi à remercier les professeurs Jean-Marc Belley et Taoufik Bouezmarni d'avoir accepté d'évaluer mon travail.

Sur le plan personnel, je dédie ce mémoire à mes chers parents. Je les remercie pour leurs encouragements et leurs sacrifices. Tout ce que j'ai accompli à ce jour est

grâce à leur amour et leur soutien précieux. Je leur dois tout. Merci également à mes soeurs Maweheb et Wifek et mon beau-frère Noumeiri pour leurs encouragements incessants.

Becem Saidani
Sherbrooke, Mai 2012

TABLE DES MATIÈRES

SOMMAIRE	iii
REMERCIEMENTS	iv
TABLE DES MATIÈRES	vi
LISTE DES TABLEAUX	ix
LISTE DES FIGURES	ix
INTRODUCTION	1
CHAPITRE 1 — Estimateur linéaire optimal et récursif	3
1.1 Introduction	3
1.2 Espérance conditionnelle	4
1.3 Estimateur au sens des moindres carrés	5
1.4 Estimateur linéaire optimal	9

1.5	Estimateur linéaire optimal et récursif	12
1.6	Vecteurs aléatoires corrélés	15
1.7	Conclusion	19
CHAPITRE 2 — Filtre de Kalman discret		20
2.1	Introduction	20
2.2	Modèle	21
2.3	Développement	23
2.3.1	Algorithme	29
2.3.2	Innovation	29
2.3.3	Analyse de la mise à jour	31
2.4	Modèle à bruits corrélés	32
2.5	Exemple : Position-Vitesse	34
CHAPITRE 3 — Filtrage avec bruits de moyennes inconnues		40
3.1	Introduction	40
3.2	Modèle à bruits non centrés	41
3.3	Problème avec estimation des moyennes	42
3.4	Expérimentation	45
3.5	Calcul des erreurs de reconstruction	48
3.6	Conclusion	53

CHAPITRE 4 — Autres types de bruit	54
4.1 Introduction	54
4.2 Transformation de vecteur aléatoire	54
4.3 Type exponentiel	56
4.4 Type khi-deux	57
4.5 Type ricien	59
4.6 Exemples multidimensionnels	63
4.7 Conclusion	71
CONCLUSION	72

LISTE DES TABLEAUX

3.1	Variation de l'erreur sur l'estimation des moyennes et le signal pour les méthodes : sans variabilité, variance décroissante et variance constante pour (E_2) avec $\hat{Z}_{0 0} = 0$, $\tilde{P}_{0 0} = 10^4 I_3$, $\lambda = 45$ et $N = 50$	49
3.2	Décroissance de l'erreur sur l'estimation des moyennes et la reconstruction du signal avec (\tilde{K}) : $\hat{Z}_{0 0} = 0$, $\tilde{P}_{0 0} = 10^4 I_3$, $R_k = Q_k = (0.6)^2$, $\lambda = 45$ et $N = 50$	49
4.1	Erreur sur l'estimation des moyennes et la reconstruction du signal pour les exemples (E_1) , (E_2) et (E_3)	64

LISTE DES FIGURES

2.1	Rudolf Kalman, inventeur du filtrage de Kalman	21
2.2	Principe du filtre de Kalman	29
2.3	Trajectoire théorique vs Observations bruitées avec $\sigma'_i = 0.1$	36
2.4	Filtre de Kalman pour l'estimation du vecteur Position-Vitesse avec $\sigma'_i = 0.1$ et $\sigma_i = 0.01$. Graphique de gauche : trajectoire théorique (en bleu) et trajectoire estimée (en rouge). Graphique de droite : vitesse théorique (en bleu) et vitesse estimée (en rouge).	37
2.5	Trajectoire théorique vs Observations bruitées avec $\sigma'_i = 0.7$	37
2.6	Filtre de Kalman pour l'estimation du vecteur Position-Vitesse avec $\sigma'_i = 0.7$ et $\sigma_i = 0.01$. Graphique de gauche : trajectoire théorique (en bleu) et trajectoire estimée (en rouge). Graphique de droite : vitesse théorique (en bleu) et vitesse estimée (en rouge).	38
2.7	Trajectoire théorique vs Observations bruitées avec $\sigma'_i = \sqrt{2}$	38

2.8	Filtre de Kalman pour l'estimation du vecteur Position-Vitesse avec $\sigma'_i = \sqrt{2}$ et $\sigma_i = 0.01$. Graphique de gauche : trajectoire théorique (en bleu) et trajectoire estimée (en rouge). Graphique de droite : vitesse théorique (en bleu) et vitesse estimée (en rouge)	39
3.1	Reconstruction du signal avec (K) : $P_0 = 10^4$	43
3.2	Reconstruction du signal et estimation des moyennes inconnues via la méthode sans variabilité où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_3$	46
3.3	Reconstruction du signal et estimation des moyennes inconnues via la méthode avec variance décroissante où $\hat{Z}_{0 0} = 0$, $\tilde{P}_{0 0} = 10^4 I_3$ et $\lambda = 45$	47
3.4	Reconstruction du signal et estimation des moyennes inconnues via la méthode avec variance constante ($\Sigma_* = 0.04$) où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_3$	47
3.5	Reconstruction du signal et estimation des moyennes inconnues avec (\tilde{K}) : $\Sigma_{x,0} = 0$ et $\Sigma_{y,0} = 0$	51
3.6	Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits gaussiens où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_5$	52
4.1	Algorithmes (K) et (\tilde{K}) appliqués sur (E_1) avec des bruits de type exponentiel où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 * I_3$	57
4.2	Algorithmes (K) et (\tilde{K}) appliqués sur (E_2) avec des bruits de type khi-deux où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 * I_3$	59
4.3	Densité de probabilité de la loi de Rice pour $\sigma = 1$ et différentes valeurs de ν	61

4.4	Algorithmes (\tilde{K}) et (\tilde{K}) appliqués sur (E_3) avec des bruits de type riciens où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_3$	63
4.5	Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits de type exponentiel où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_5$	65
4.6	Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits de type khie-deux où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_5$	67
4.7	Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits de type riciens où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_5$	69
4.8	Algorithme (\tilde{K}) appliqué sur $(S_4^{2,1})$ où $\hat{Z}_{0 0} = 0$ et $\tilde{P}_{0 0} = 10^4 I_5$	70

INTRODUCTION

Dans de nombreuses applications telles qu'en traitement d'image, détection radar, finance, contrôle industriel, séismologie, le signal vérité n'est pas accessible directement et est noyé dans le signal observé [5, 1, 19, 14, 6]. Par exemple, pour le cas des radars où l'on désire suivre une cible, des mesures sur sa position, sa vitesse et son accélération sont disponibles à chaque instant mais avec énormément de perturbations dues aux erreurs de mesures. Le filtre de Kalman fait appel à la dynamique de la cible qui définit son évolution dans le temps pour obtenir de meilleures données, éliminant ainsi l'effet du bruit. Le développement de méthodes d'extraction du signal vérité est en conséquence un enjeu important pour les différentes applications et est à l'origine d'une littérature scientifique très riche en traitement du signal [3, 9, 4]. La problématique qui se pose alors est de savoir, au regard des mesures et des hypothèses sur le système, comment estimer au mieux le signal utile.

En 1960, R.E. Kalman a publié son fameux article [8] intitulé " A new Approach to Linear Filtering and Prediction Problems ". Ses recherches le mènent à y décrire un processus qui sera connu plus tard comme le filtre de Kalman. En effet ce filtre est un processus mathématique qui permet une meilleure estimation de l'état futur d'un système malgré l'imprécision de la modélisation et des mesures. En effet il vise à

estimer l'état d'un système en fonction de son état précédent, des commandes appliquées et des mesures bruitées. Le filtre de Kalman ne nécessite pas toutes les données passées pour produire une estimation à l'instant courant ce qui nous laisse deviner l'aspect récursif du filtre et rend son implémentation possible pour des applications en temps réel.

Dans ce mémoire, nous présentons en premier la théorie de base sur l'estimation d'une variable aléatoire par la mesure ou l'observation d'une autre variable aléatoire : l'estimation au sens des moindres carrés, l'estimateur linéaire optimal au sens des moindres carrés et l'estimateur linéaire optimal récursif au sens des moindres carrés. Dans le deuxième chapitre nous étudions le filtre de Kalman et plus précisément nous établissons les différentes équations nécessaires sur lesquels reposent ce filtre. Ensuite au chapitre trois, on adapte ce filtre pour estimer l'état d'un système dynamique lorsque les bruits sont non centrés et de moyennes inconnues. Finalement dans le dernier chapitre, nous donnons des applications du filtre de Kalman modifié sur des systèmes dynamiques avec des bruits blancs gaussiens, exponentiels, khi-deux et riciens.

CHAPITRE 1

Estimateur linéaire optimal et récuratif

1.1 Introduction

Nous décrivons dans cette partie quelques principes de base de la théorie de la prévision d'une variable aléatoire par la mesure ou l'observation d'une autre variable aléatoire. De façon générale, les variables considérées sont vectorielles et nous distinguerons la variable aléatoire (en lettre majuscule) de sa réalisation (en lettre minuscule). L'estimation optimale consiste à réaliser une estimation ponctuelle par l'optimisation d'un critère. Nous allons décrire essentiellement dans ce chapitre l'estimation au sens des moindres carrés. Cette dernière estimation conduit à la construction de l'estimateur linéaire optimal au sens des moindres carrés et à l'estimateur linéaire optimal récuratif au sens des moindres carrés [15, 10, 13, 4, 7, 17, 3].

1.2 Espérance conditionnelle

Définition 1.2.1. Soient X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{A}) telle que $\mathbb{E}(|X|) < +\infty$ et $\mathcal{B} \subset \mathcal{A}$ une sous-tribu de \mathcal{A} . L'unique variable aléatoire W \mathcal{B} -mesurable vérifiant les conditions :

- $\mathbb{E}(|W|) < +\infty$,
 - $\mathbb{E}(XZ) = \mathbb{E}(WZ)$ pour toute variable aléatoire Z \mathcal{B} -mesurable et bornée,
- est appelée l'espérance conditionnelle de X sachant \mathcal{B} et sera notée $\mathbb{E}(X|\mathcal{B})$.

Proposition 1.2.1. L'espérance conditionnelle a les propriétés suivantes :

1. *Linéarité :*

$$\mathbb{E}(\lambda X + \mu Y|\mathcal{B}) = \lambda \mathbb{E}(X|\mathcal{B}) + \mu \mathbb{E}(Y|\mathcal{B}) \quad \forall \lambda, \mu \in \mathbb{R}.$$

2. *Positivité :*

$$X \geq 0 \Rightarrow \mathbb{E}(X|\mathcal{B}) \geq 0 \text{ presque sûrement (p.s.)}$$

3. *Si X est \mathcal{B} -mesurable alors :*

$$\mathbb{E}(X|\mathcal{B}) = X \text{ p.s.}$$

4. *Si X est indépendante de \mathcal{B} alors*

$$\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(X) \text{ p.s.}$$

5. *Si Z est une variable aléatoire \mathcal{B} -mesurable et bornée alors*

$$\mathbb{E}(ZX|\mathcal{B}) = Z\mathbb{E}(X|\mathcal{B}) \text{ p.s.}$$

6. *Soit \mathcal{C} une sous-tribu de \mathcal{B} alors*

$$\mathbb{E}(\mathbb{E}(X|\mathcal{B})|\mathcal{C}) = \mathbb{E}(X|\mathcal{C}) \text{ p.s.}$$

7. $\mathbb{E}(\mathbb{E}(X|\mathcal{B})) = \mathbb{E}(X)$ p.s.

Remarque 1.2.1. Lorsque la tribu \mathcal{B} est engendrée par une variable aléatoire Y , on note $\mathbb{E}(X|\mathcal{B})$ par $\mathbb{E}(X|Y)$.

Définition 1.2.2. Étant données deux vecteurs aléatoires X et Y dont les espérances existent, on dit que X et Y sont corrélées si et seulement si $\text{Cov}(X, Y) \neq 0$.

Proposition 1.2.2. Soient X et Y deux variables aléatoires dont les espérances existent. Alors les variables aléatoires $X - \mathbb{E}(X|Y)$ et Y sont décorréliées.

Démonstration. On pose $W = X - \mathbb{E}(X|Y)$, pour démontrer le résultat précédent, il suffit de montrer que $\text{Cov}(W, Y) = 0$. En effet,

$$\begin{aligned}\text{Cov}(W, Y) &= \mathbb{E}(WY) - \mathbb{E}(W)\mathbb{E}(Y) \\ &= \mathbb{E}[(X - \mathbb{E}(X|Y))Y] - \mathbb{E}(X - \mathbb{E}(X|Y))\mathbb{E}(Y) \\ &= 0\end{aligned}$$

car $\mathbb{E}[(X - \mathbb{E}(X|Y))Y] = \mathbb{E}(XY) - \mathbb{E}[\mathbb{E}(X|Y)Y] = 0$ et $\mathbb{E}(X - \mathbb{E}(X|Y)) = 0$. \square

1.3 Estimateur au sens des moindres carrés

Soit X une variable aléatoire réelle. On va chercher une estimation \hat{x} de la réalisation x de X au sens des moindres carrés : l'estimation \hat{x} de x doit satisfaire le critère d'optimalité suivant :

$$\mathbb{E}[|X - \hat{x}|^2] \leq \mathbb{E}[|X - z|^2]$$

pour tout réel z . On peut vérifier que cette meilleure estimation \hat{x} est $\mathbb{E}(X)$.

Nous allons maintenant aborder une question plus difficile. Étant données deux variables aléatoires X et Y , le problème sera d'estimer la réalisation x de la variable aléatoire X , tout en disposant d'une information supplémentaire, à savoir que y est une réalisation de la variable aléatoire Y . Évidemment si X et Y sont indépendantes, l'information $Y = y$ n'a aucune utilité pour estimer la réalisation x de X , par contre si X et Y sont liées, est-il possible d'améliorer notre estimation initiale $\mathbb{E}(X)$?

Problème : Soient deux vecteurs aléatoires X et Y liés de dimension n et m respectivement et de densité conjointe $f_{X,Y}$. Étant donnée une réalisation y du vecteur aléatoire Y , on va chercher une estimation \hat{x} de la réalisation correspondante x de X vérifiant :

$$\mathbb{E}[\|X - \hat{x}\|^2 | Y = y] \leq \mathbb{E}[\|X - z\|^2 | Y = y] \quad (1.1)$$

pour tout vecteur $z \in \mathbb{R}^n$.

Théorème 1.3.1. *L'estimation \hat{x} de la réalisation x de X à partir de l'information $Y = y$ qui vérifie l'inégalité (1.1) est donnée par l'espérance conditionnelle :*

$$\hat{x} = \mathbb{E}(X | Y = y).$$

Démonstration. Pour démontrer que $\hat{x} = \mathbb{E}(X | Y = y)$ il suffit de montrer que $\mathbb{E}(X | Y = y)$ vérifie l'inégalité (1.1). En effet

$$\begin{aligned} \mathbb{E}[\|X - z\|^2 | Y = y] &= \mathbb{E}[X'X - 2z'X + z'z | Y = y] \\ &= \mathbb{E}[X'X | Y = y] - 2z'\mathbb{E}[X | Y = y] + z'z \\ &= \mathbb{E}[\|X\|^2 | Y = y] + \|z - \mathbb{E}(X | Y = y)\|^2 - \|\mathbb{E}(X | Y = y)\|^2, \end{aligned}$$

et cette dernière expression est minimale si et seulement si $z = \mathbb{E}(X | Y = y)$ et $\mathbb{E}[\|X\|^2 | Y = y]$ est finie □

Remarque 1.3.1. L'estimation \hat{x} est appelée estimation au sens des moindres carrés de x .

Remarque 1.3.2. L'estimation $\hat{x} = \mathbb{E}(X|Y = y)$ est en fait une réalisation du vecteur aléatoire $\hat{X}(Y) = \mathbb{E}(X|Y)$. $\hat{X}(Y)$ peut être vu comme un opérateur de prédiction qui prend comme valeur d'entrée une observation de Y et donne la valeur prédite \hat{x} de X . D'après la construction de $\hat{X}(Y)$ on a :

$$\mathbb{E}[\|X - \hat{X}(Y)\|^2 | Y = y] \leq \mathbb{E}[\|X - g(Y)\|^2 | Y = y]$$

pour toute fonction g de \mathbb{R}^m dans \mathbb{R}^n . Or on a :

$$\mathbb{E}^Y[\mathbb{E}[\|X - g(Y)\|^2 | Y = y]] = \mathbb{E}[\|X - g(Y)\|^2],$$

donc

$$\mathbb{E}[\|X - \hat{X}(Y)\|^2] \leq \mathbb{E}[\|X - g(Y)\|^2]$$

ce qui implique que l'estimateur construit $\hat{X}(Y)$ satisfait le critère d'optimalité suivant :

$$\mathbb{E}^{X,Y}[\|X - \hat{X}(Y)\|^2] \leq \mathbb{E}^{X,Y}[\|X - g(Y)\|^2]$$

pour toute fonction g de \mathbb{R}^m dans \mathbb{R}^n . Dans ce cas $\hat{X}(Y)$ est appelé estimateur au sens des moindres carrés, ou bien l'estimateur sans biais de variance minimale, de X étant donnée une observation de Y .

Nous allons donner ici une définition précise de l'estimateur d'un vecteur aléatoire X à partir de l'observation d'un autre vecteur aléatoire Y .

Définition 1.3.1. Soient X et Y deux vecteurs aléatoires liés de dimension m et n respectivement. Par définition un estimateur de X à partir de l'observation Y est

une variable aléatoire noté $\widehat{X}(Y)$. Pour chaque observation y de Y , $\widehat{x} = \widehat{X}(y)$ est appelée une estimation de X .

Lemme 1.3.1. *L'estimateur $\widehat{X}(Y)$ possède les propriétés suivantes*

1. *L'estimateur $\widehat{X}(Y)$ est linéaire, c'est-à-dire pour toute matrice $A \in M_m(\mathbb{R})$ et tout vecteur $b \in \mathbb{R}^m$ on a*

$$\mathbb{E}(AX + b|Y) = A\mathbb{E}(X|Y) + b.$$

2. *L'estimateur $\widehat{X}(Y)$ est sans biais pour X , c'est-à-dire*

$$\mathbb{E}(\widehat{X}(Y)) = \mathbb{E}(X).$$

Démonstration.

1. Évident puisque l'espérance conditionnelle est linéaire.
2. Il suffit d'appliquer la Définition 1.2.1 avec $Z = 1$.

□

Proposition 1.3.2. *Si X et Y sont deux vecteurs aléatoires on a*

1. *Si $X = f(Y)$ p.s., alors $\mathbb{E}(X|Y) = X$ p.s.*
2. *Si X et Y sont indépendantes, alors $\mathbb{E}(X|Y) = \mathbb{E}(X)$ p.s.*

Démonstration. Il suffit d'utiliser 3 et 4 de la Proposition 1.2.1.

□

Remarque 1.3.3. *On va donner une interprétation de la Proposition 1.3.2.*

1. *La première propriété exprime que lorsque X dépend explicitement de Y alors l'observation de Y permet de connaître exactement X .*
2. *La deuxième propriété montre que si X et Y sont indépendants, l'observation de Y n'apporte rien de nouveau sur X .*

1.4 Estimateur linéaire optimal

Dans cette partie, nous allons résoudre le même problème cité dans la section précédente en imposant la **linéarité** de l'estimateur [15, 17]. Plus précisément si X et Y sont deux vecteurs aléatoires liés, on va construire l'estimateur linéaire optimal E^* de X en fonction de Y :

$$E^*(X|Y) = A_0Y + B_0$$

minimisant la variance de l'erreur d'estimation

$$\mathbb{E}(\|X - AY - B\|^2),$$

avec A_0 et B_0 sont des matrices réelles (déterministes) de tailles convenables.

Proposition 1.4.1. *Soient X et Y deux vecteurs aléatoires ayant les moyennes $\mathbb{E}(X) = m_X$ et $\mathbb{E}(Y) = m_Y$ respectivement et de matrices de covariances :*

$$\mathbb{E}[(X - m_X)(X - m_X)^T] = P_{XX}$$

$$\mathbb{E}[(Y - m_Y)(Y - m_Y)^T] = P_{YY}$$

$$\mathbb{E}[(X - m_X)(Y - m_Y)^T] = P_{XY}.$$

Si nous supposons de plus que P_{YY} est inversible, alors l'estimateur linéaire optimal de X en fonction de Y est

$$E^*(X|Y) = m_X + P_{XY}P_{YY}^{-1}(Y - m_Y),$$

et la matrice de covariance de l'erreur d'estimation est

$$\mathbb{E}[(X - E^*(X|Y))(X - E^*(X|Y))^T] = P_{XX} - P_{XY}P_{YY}^{-1}P_{XY}^T.$$

Démonstration. Posons $X_c = X - m_X$, $Y_c = Y - m_Y$ et $B_c = B - m_X + Am_Y$.

Nous avons

$$\mathbb{E}(\|X - AY - B\|^2) = \text{Tr} \mathbb{E}[(X_c - AY_c - B_c)(X_c - AY_c - B_c)^T].$$

Ainsi,

$$\begin{aligned} \mathbb{E}(\|X - AY - B\|^2) &= \text{Tr}[P_{XX} - P_{XY}A^T + AP_{XY}^T \\ &\quad + AP_{YY}A^T] + \|B_c\|^2. \end{aligned}$$

Lorsque P_{YY} est inversible, l'utilisation de l'égalité

$$\begin{aligned} P_{XX} - P_{XY}A^T + AP_{XY}^T + AP_{YY}A^T &= P_{XX} - P_{XY}P_{YY}^{-1}P_{XY}^T \\ &\quad + [A - P_{XY}P_{YY}^{-1}]P_{YY}[A - P_{XY}P_{YY}^{-1}]^T \end{aligned}$$

conduit à écrire

$$\begin{aligned} \mathbb{E}(\|X - AY - B\|^2) &= \text{Tr}[P_{XX} - P_{XY}P_{YY}^{-1}P_{XY}^T] + \|B_c\|^2 \\ &\quad + \text{Tr}[(A - P_{XY}P_{YY}^{-1})P_{YY}(A - P_{XY}P_{YY}^{-1})^T]. \end{aligned}$$

Cette expression est optimale lorsque

$$A = P_{XY}P_{YY}^{-1}$$

$$B_c = 0,$$

ce qui implique que

$$A_0 = P_{XY}P_{YY}^{-1}$$

$$B_0 = m_X + P_{XY}P_{YY}^{-1}m_Y.$$

En conséquence l'estimateur linéaire optimal est

$$E^*(X|Y) = m_X + P_{XY}P_{YY}^{-1}(Y - m_Y)$$

et la matrice de covariance de l'erreur d'estimation est

$$\mathbb{E}[(X - E^*(X|Y))(X - E^*(X|Y))^T] = P_{XX} - P_{XY}P_{YY}^{-1}P_{XY}^T.$$

□

Proposition 1.4.2. *Si X et Y sont deux vecteurs aléatoires alors*

1. *L'estimateur linéaire optimal $E^*(X|Y)$ est sans biais dans le sens que*

$$\mathbb{E}(E^*(X|Y)) = \mathbb{E}(X).$$

2. *L'erreur d'estimation $X - E^*(X|Y)$ et l'observation Y sont décorrélées, c'est-à-dire*

$$E[(X - E^*(X|Y))Y^T] = 0,$$

et de façon plus générale, l'erreur d'estimation est décorrélée avec toute fonction linéaire de Y , en particulier $E^(X|Y)$.*

3. *Pour toute matrice déterministe M de $M_n(\mathbb{R})$ et tout vecteur déterministe b de \mathbb{R}^n , l'estimeur optimal linéaire vérifie*

$$E^*(MX + b|Y) = ME^*(X|Y) + b.$$

Démonstration.

- (1) Nous avons

$$\begin{aligned} \mathbb{E}(E^*(X|Y)) &= \mathbb{E}[m_X + P_{XY}P_{YY}^{-1}(Y - m_Y)] \\ &= m_X + P_{XY}P_{YY}^{-1}\mathbb{E}(Y - m_Y) \\ &= \mathbb{E}(X). \end{aligned}$$

(2) En posant $X_c = X - m_X$ et $Y_c = Y - m_Y$ nous avons

$$\begin{aligned}
\mathbb{E}[(X - E^*(X|Y))Y^T] &= \mathbb{E}[(X - E^*(X|Y))(Y - m_Y)^T] \\
&= \mathbb{E}[(X_c - P_{XY}P_{YY}^{-1}Y_c)Y_c^T] \\
&= P_{XY} - P_{XY}P_{YY}^{-1}P_{YY} \\
&= 0.
\end{aligned}$$

(3) D'une part nous avons

$$ME^*(X|Y) + b = M[m_X + P_{XY}P_{YY}^{-1}(Y - m_Y)] + b,$$

et d'autre part

$$\begin{aligned}
E^*(MX + b|Y) &= m_{(MX+b)Y} + P_{(MX+b)Y}P_{YY}^{-1}(Y - m_Y) \\
&= Mm_X + b + P_{(MX+b)Y}P_{YY}^{-1}(Y - m_Y).
\end{aligned}$$

Or puisque $P_{(MX+b)Y} = MP_{XY}$, nous trouvons que :

$$\begin{aligned}
E^*(MX + b|Y) &= M[m_X + P_{XY}P_{YY}^{-1}(Y - m_Y)] + b \\
&= ME^*(X|Y) + b.
\end{aligned}$$

□

1.5 Estimateur linéaire optimal et récursif

Dans cette section nous essayons d'étendre le résultat de la section précédente pour calculer l'estimateur linéaire optimal à partir d'une suite de variable aléatoires [15]. Plus précisément si X, Y_1, Y_2, \dots, Y_n est une suite de variables aléatoires, nous construisons l'estimateur linéaire optimal de X en fonction de Y_1, Y_2, \dots, Y_n de façon récursive.

Proposition 1.5.1. Soient X , Y et Z trois vecteurs aléatoires tels que Y et Z sont décorrélés. L'estimateur linéaire optimal de X à partir de Y et Z est

$$E^* \{X|Y, Z\} = E^* \{X|Y\} + E^* \left\{ \tilde{X}_{|Y|Z} \right\}$$

avec $\tilde{X}_{|Y} = X - E^* \{X|Y\}$. La matrice de covariance de l'erreur d'estimation $\tilde{X}_{|Y,Z} = X - E^* \{X|Y, Z\}$ est donnée par

$$P_{\tilde{X}_{|Y,Z}, \tilde{X}_{|Y,Z}} = P_{XX} - P_{XY}P_{YY}^{-1}P_{XY}^T - P_{XZ}P_{ZZ}^{-1}P_{XZ}^T.$$

Démonstration. En posant $W^T = [X^T, Y^T]$ et $m_W = [m_X^T, m_Y^T]$, nous obtenons

$$E^* \{X|W\} = m_X + P_{XW}P_{WW}^{-1}(W - m_W).$$

Or Y et Z sont décorrélés et donc

$$\begin{aligned} P_{XW} &= [P_{XY}, P_{XZ}] \\ P_{WW} &= \begin{pmatrix} P_{YY} & 0 \\ 0 & P_{ZZ} \end{pmatrix} \end{aligned}$$

ce qui implique que

$$E^* \{X|W\} = m_X + P_{XY}P_{YY}^{-1}(Y - m_Y) + P_{XZ}P_{ZZ}^{-1}(Z - m_Z).$$

Or

$$\begin{aligned} E^* \{X|Y\} &= m_X + P_{XY}P_{YY}^{-1}(Y - m_Y) \\ \tilde{X}_{|Y} &= X - E^* \{X|Y\} \end{aligned}$$

alors

$$E^* \left\{ \tilde{X}_{|Y|Z} \right\} = E^* \{X|Z\} - m_X - P_{YZ}P_{ZZ}^{-1}E^*(Y - m_Y|Z).$$

Puisque on a supposé que Y et Z sont décorrélés, on obtient donc

$$E^* \{X|W\} = E^* \{X|Y\} + E^* \left\{ \tilde{X}_{|Y}|Z \right\}.$$

Nous déduisons que la matrice de covariance de l'erreur d'estimation $\tilde{X}_{|Y,Z}$ est donnée par :

$$P_{\tilde{X}_{|Y}, \tilde{X}_{|Y}} - P_{\tilde{X}_{|Y}, Z} P_{ZZ}^{-1} P_{\tilde{X}_{|Y}, Z}^T = P_{XX} - P_{XY} P_{YY}^{-1} P_{XY}^T - P_{XZ} P_{ZZ}^{-1} P_{XZ}^T,$$

ce qui implique que

$$P_{\tilde{X}_{|Y,Z}, \tilde{X}_{|Y,Z}} = P_{XX} - P_{XY} P_{YY}^{-1} P_{XY}^T - P_{XZ} P_{ZZ}^{-1} P_{XZ}^T.$$

□

Rappelons un résultat sur l'inversion d'une matrice par blocs [5].

Lemme 1.5.1. *Soit A une matrice par blocs définie par*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

où

- A_{11} et A_{22} sont des matrices inversibles de taille $n \times n$ et $m \times m$ respectivement.
- $(A_{11} - A_{12} A_{22}^{-1} A_{21})$ et $(A_{22} - A_{21} A_{11}^{-1} A_{12})$ sont des matrices inversibles.

Alors A est non singulière et son inverse est donné par

$$A^{-1} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

avec

$$\begin{aligned} c_{11} &= A_{11}^{-1} + A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1} \\ c_{12} &= -A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} \\ c_{21} &= -(A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1} \\ c_{22} &= (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}. \end{aligned}$$

1.6 Vecteurs aléatoires corrélés

Dans la suite nous généralisons le résultat de la Proposition 1.5.1 pour des vecteurs aléatoires corrélés. Cette idée a été présentée sans démonstration dans l'article "A Tutorial Introduction to Estimation and Filtering" de Ian. B. Rhodes [15]. Dans cette section, nous allons donner les démonstrations de ces résultats.

Proposition 1.6.1. *Soient X , Y et Z des vecteurs aléatoires tels que Y et Z sont possiblement corrélés. Alors*

$$E^* \{X|Y, Z\} = E^* \{X|Y, \tilde{Z}|_Y\}$$

avec $\tilde{Z}|_Y = Z - E^*(Z|Y)$.

Démonstration. Pour démontrer ce résultat il suffit de calculer $E^* \{X|Y, Z\}$ et $E^* \{X|Y, \tilde{Z}|_Y\}$. D'une part, il découle de la Proposition 1.4.2 que $\tilde{Z}|_Y = Z - E^*(Z|Y)$ est décorrélé avec Y et, en utilisant la Proposition 1.5.1, nous trouvons

$$\begin{aligned} E^* \{X|Y, \tilde{Z}|_Y\} &= E^* \{X|Y\} + E^* \{\tilde{X}|_Y|\tilde{Z}|_Y\} \\ &= m_X + P_{XY}P_{YY}^{-1}(Y - m_Y) + P_{\tilde{X}|_Y, \tilde{Z}|_Y}P_{\tilde{Z}|_Y}^{-1}(Z - E^*(Z|Y)) \end{aligned}$$

où

$$\begin{aligned} P_{\tilde{X}|_Y, \tilde{Z}|_Y} &= \mathbb{E}[(X - m_X)(\tilde{Z}|_Y)^T] \\ &= \mathbb{E}[(X - m_X)(Z - m_Z - P_{ZY}P_{YY}^{-1}(Y - m_Y))^T] \\ &= P_{XZ} - P_{XY}P_{YY}^{-1}P_{ZY}, \end{aligned}$$

et

$$\begin{aligned}
P_{\tilde{Z}|Y} &= \mathbb{E}[\tilde{Z}|Y \tilde{Z}|Y^T] \\
&= \mathbb{E}[(Z - m_Z - P_{ZY}P_{YY}^{-1}(Y - m_Y))(Z - m_Z - P_{ZY}P_{YY}^{-1}(Y - m_Y))^T] \\
&= P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{ZY} - P_{ZY}P_{YY}^{-1}P_{YZ} + P_{ZY}P_{YY}^{-1}P_{YY}P_{YY}^{-1}P_{ZY} \\
&= P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{ZY} - P_{ZY}P_{YY}^{-1}P_{YZ} + P_{ZY}P_{YY}^{-1}P_{ZY} \\
&= P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ}.
\end{aligned}$$

Donc

$$\begin{aligned}
E^* \{X|Y, \tilde{Z}|Y\} &= m_X + P_{XY}P_{YY}^{-1}(Y - m_Y) \\
&\quad + (P_{XZ} - P_{XY}P_{YY}^{-1}P_{ZY})(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}(Z - m_Z - P_{ZY}P_{YY}^{-1}(Y - m_Y)) \\
&= m_X + [P_{XY}P_{YY}^{-1} - (P_{XZ} - P_{ZY}P_{YY}^{-1}P_{ZY})(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}P_{ZY}P_{YY}^{-1}](Y - m_Y) \\
&\quad + (P_{XZ} - P_{ZY}P_{YY}^{-1}P_{ZY})(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}(Z - m_Z).
\end{aligned}$$

D'autre part en posant $W_1^T = [Y^T, Z^T]$ nous obtenons

$$E^* \{X|Y, Z\} = m_X + P_{XW_1^T}P_{W_1^TW_1^T}^{-1}(W_1^T - m_{W_1^T})$$

où

$$\begin{aligned}
P_{XW_1^T} &= [P_{XY} \ P_{XZ}] \\
P_{W_1^TW_1^T} &= \begin{pmatrix} P_{YY} & P_{YZ} \\ P_{ZY} & P_{ZZ} \end{pmatrix}.
\end{aligned}$$

En utilisant le Lemme 1.5.2, nous obtenons

$$P_{W_1^TW_1^T}^{-1} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

avec

$$\begin{aligned}
b_{11} &= P_{YY}^{-1} + P_{YY}^{-1}P_{YZ}(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}P_{ZY}P_{YY}^{-1} \\
b_{12} &= -P_{YY}^{-1}P_{YZ}(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1} \\
b_{21} &= -(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}P_{ZY}P_{YY}^{-1} \\
b_{22} &= (P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}.
\end{aligned}$$

Alors

$$E^* \{X|Y, Z\} = m_X$$

$$\begin{aligned}
&+ [P_{XY}P_{YY}^{-1} + P_{XY}P_{YY}^{-1}P_{YZ}(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}P_{ZY}P_{YY}^{-1} - P_{XZ}(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1} \\
&P_{ZY}P_{YY}^{-1}](Y - m_Y) \\
&- [P_{XY}P_{YY}^{-1}P_{YZ}(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1} + P_{XZ}(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}](Z - m_Z) \\
&= m_X + [P_{XY}P_{YY}^{-1} - (P_{XZ} - P_{ZY}P_{YY}^{-1}P_{ZY})(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}P_{ZY}P_{YY}^{-1}](Y - m_Y) \\
&+ (P_{XZ} - P_{ZY}P_{YY}^{-1}P_{ZY})(P_{ZZ} - P_{ZY}P_{YY}^{-1}P_{YZ})^{-1}(Z - m_Z) \\
&= E^* \{X|Y, \tilde{Z}_{|Y}\}.
\end{aligned}$$

□

Proposition 1.6.2. Soient X , Y et Z des vecteurs aléatoires tels que Y et Z sont corrélés. L'estimateur linéaire optimal de X à partir de Y et Z est

$$E^* \{X|Y, Z\} = E^* \{X|Y\} + E^* \{ \tilde{X}_{|Y} | \tilde{Z}_{|Y} \}$$

avec $\tilde{X}_{|Y} = X - E^*(X|Y)$ et $\tilde{Z}_{|Y} = Z - E^*(Z|Y)$. De plus, la matrice de covariance de l'erreur d'estimation $\tilde{X}_{|Y,Z}$ est donnée par

$$P_{\tilde{X}_{|Y,Z}, \tilde{X}_{|Y,Z}} = P_{\tilde{X}_{|Y}, \tilde{X}_{|Y}} - P_{\tilde{X}_{|Y}, \tilde{Z}_{|Y}} P_{\tilde{Z}_{|Y}, \tilde{Z}_{|Y}}^{-1} P_{\tilde{X}_{|Y}, \tilde{Z}_{|Y}}^T.$$

Démonstration. Dans le cas où Y et Z sont corrélés, nous pouvons se ramener au cas décorréolé par l'intermédiaire du vecteur aléatoire :

$$\tilde{Z}_{|Y} = Z - E^*(Z|Y)$$

qui est décorréolé avec Y . D'après la Proposition 1.6.1 nous obtenons alors

$$E^* \{X|Y, Z\} = E^* \{X|Y, \tilde{Z}_{|Y}\}.$$

Maintenant pour calculer $E^* \{X|Y, \tilde{Z}_{|Y}\}$ nous utilisons la Proposition 1.5.1. En effet puisque Y et $\tilde{Z}_{|Y}$ sont décorréolés, on obtient

$$E^* \{X|Y, \tilde{Z}_{|Y}\} = E^* \{X|Y\} + E^* \{\tilde{X}_{|Y}|\tilde{Z}_{|Y}\}.$$

Enfin, la matrice de covariance de l'erreur d'estimation est bien égale à :

$$P_{\tilde{X}_{|Y,Z}, \tilde{X}_{|Y,Z}} = P_{\tilde{X}_{|Y}, \tilde{X}_{|Y}} - P_{\tilde{X}_{|Y}, \tilde{Z}_{|Y}} P_{\tilde{Z}_{|Y}, \tilde{Z}_{|Y}}^{-1} P_{\tilde{X}_{|Y}, \tilde{Z}_{|Y}}^T.$$

□

Proposition 1.6.3. Soient $X, Y_1, Y_2, \dots, Y_{k+1}$ des vecteurs aléatoires, alors l'estimateur linéaire optimal de X à partir de la suite des vecteurs aléatoires Y_1, Y_2, \dots, Y_{k+1} est donné par :

$$\hat{X}_{|k+1} = \hat{X}_{|k} + E^* \{\tilde{X}_{|k}|\tilde{Y}_{k+1|k}\}$$

avec

$$\begin{aligned} \tilde{X}_{|k} &= X - \hat{X}_{|k} = X - E^* \{X|Y_1, Y_2, \dots, Y_k\}, \\ \tilde{Y}_{k+1|k} &= Y_{k+1} - \hat{Y}_{k+1} = Y_{k+1} - E^* \{Y_{k+1}|Y_1, Y_2, \dots, Y_k\}. \end{aligned}$$

La matrice de covariance de l'erreur d'estimation $\tilde{X}_{|k+1}$ est donnée par :

$$P_{\tilde{X}_{|k+1}, \tilde{X}_{|k+1}} = P_{\tilde{X}_{|k}, \tilde{X}_{|k}} - P_{\tilde{X}_{|k}, \tilde{Y}_{k+1|k}} P_{\tilde{Y}_{k+1|k}, \tilde{Y}_{k+1|k}}^{-1} P_{\tilde{X}_{|k}, \tilde{Y}_{k+1|k}}.$$

Démonstration. Pour démontrer ce résultat, il suffit de poser

$Y = (Y_0, \dots, Y_k)$, $Z = Y_{k+1}$ et d'appliquer la Proposition 1.6.2. □

1.7 Conclusion

Dans ce chapitre nous avons établi la théorie de base sur la prévision d'une variable aléatoire à l'aide d'une autre : l'estimation au sens des moindres carrés, l'estimateur linéaire optimal au sens des moindres carrés et l'estimateur linéaire optimal récursif au sens des moindres carrés.

CHAPITRE 2

Filtre de Kalman discret

2.1 Introduction

Les problèmes d'estimation de l'état d'un système dynamique peuvent être classés selon la quantité d'information disponible. En effet, considérons un système dynamique dont on possède un ensemble de mesures $Y(I_0, I_f)$ entre l'instant initial I_0 et l'instant final I_f . On peut chercher à estimer l'état du système dynamique X à un instant donnée τ que l'on notera $X(\tau|Y(I_0, I_f))$. Suivant la valeur de τ nous distinguons trois cas [15] :

- si $\tau < I_f$ il s'agit d'un problème de **lissage** (Smoothing) ;
- si $\tau = I_f$ il s'agit d'un problème de **filtrage** (Filtering) ;
- si $\tau > I_f$ il s'agit d'un problème de **prédiction** (Prédiction).

Ces problèmes consistent essentiellement à estimer l'état d'un système dynamique à partir d'observations généralement bruitées. En effet, pour ces problèmes, nous disposons à l'instant k d'une suite $Y_1, Y_2, Y_3, \dots, Y_k$ d'observations recueillies au niveau

des capteurs. Chaque observation Y_k est reliée à l'état inconnu X_k par une relation du type :

$$Y_k = H_k X_k + V_k$$

où H_k est la matrice de mesure et V_k est un bruit qui modélise l'erreur d'observation. Le but sera d'obtenir le plus d'information possible sur l'état du système X_l avec

$$\begin{cases} l < k & \text{pour le lissage} \\ l = k & \text{pour le filtrage} \\ l > k & \text{pour la prédiction.} \end{cases}$$

Dans ce chapitre on va résoudre le problème de filtrage à l'aide de l'approche du filtre de Kalman[15, 8, 14, 7, 3].

Paternité

Le filtre de Kalman doit son nom à Rudolf Kalman né à Budapest, en Hongrie, le 19 mai 1930. Ce dernier l'a inventé pour résoudre un problème de poursuite de trajectoire dans la préparation des missions Apollo en 1960.

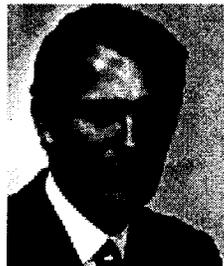


Figure 2.1 – Rudolf Kalman, inventeur du filtrage de Kalman

2.2 Modèle

Rappelons en premier lieu ce qu'on entend par système dynamique et bruits blancs.

Définition 2.2.1. *Un système dynamique est un modèle permettant de décrire l'évolution au cours du temps d'un ensemble d'objets en interaction dont l'état courant dépend juste de l'état précédent, des commandes appliquées et d'un processus aléatoire.*

Les équations décrivant le système sont appelées les équations d'état du système. Ces équations contiennent également ce que nous appelons du " bruit blanc "

Définition 2.2.2. *On appelle bruit blanc discret un processus aléatoire $\{X_n\}_{n \in \mathbb{N}}$ tels que $\mathbb{E}(X_n)$ existe et*

$$\begin{cases} \text{Var}(X_n) < \infty \text{ pour tout } n \in \mathbb{N} \\ \text{Cov}(X_n, X_l) = 0 \text{ si } n \neq l. \end{cases}$$

Le filtre de Kalman repose sur deux équations :

- l'équation d'état X_k du système ;
- l'équation de l'observation Y_k ;

où on note l'instant discret par un indice k . Typiquement, nous avons le système :

$$(S^{n,m}) \begin{cases} X_{k+1} = A_k X_k + B_k U_k + W_k & (1) \\ Y_k = H_k X_k + V_k & (2) \end{cases}$$

où

- X_k est le vecteur d'état du processus, de taille $n \times 1$, tel que l'état initial X_0 du processus est d'espérance et de variance connues, et décorrélé avec W_l et V_l pour $l = 0, 1, 2, \dots$;
- U_k est un vecteur déterministe, de taille $n \times 1$, qui représente la commande appliquée sur la dynamique de l'état X_k ;
- A_k est la matrice déterministe de transition de l'état, de taille $n \times n$, qui décrit la dynamique des X_k ;

- B_k est une matrice déterministe qui décrit l'application de la commande sur l'état X_k ;
- W_k est un bruit blanc qui modélise l'erreur du processus, de taille $n \times 1$, de moyenne connue et de matrice de covariance Q_k connue (semi-définie positive);
- Y_k est le vecteur d'observation à l'instant k , de taille $m \times 1$;
- H_k est la matrice déterministe de mesure de taille $m \times n$;
- V_k est un bruit blanc qui modélise l'erreur d'observation, de taille $m \times 1$, de moyenne connue et de matrice de covariance R_k connue (définie positive) et non corrélée avec W_l pour $l = 0, 1, 2, \dots$

Dans la suite en posant

$$\delta_{ik} = \begin{cases} 1 & \text{si } i = k \\ 0 & \text{si } i \neq k \end{cases}$$

Nous noterons :

- $\mathbb{E}(W_k W_i^T) = Q_k \delta_{ik}$ (de taille $n \times n$ et semi-définie positive);
- $\mathbb{E}(V_k V_i^T) = R_k \delta_{ik}$ (de taille $m \times m$ et définie positive);
- $\mathbb{E}(W_k V_i^T) = 0 \quad \forall k, i \in \mathbb{N}$.

L'objectif sera de déterminer une estimation optimale et récursive des états X_k à partir des mesures Y_k .

2.3 Développement

Le filtre de Kalman est un filtre optimal pour l'estimation de l'état du système, il garantit que l'erreur sur l'estimation est en moyenne nulle et de variance minimale [15, 10, 5]. Pour calculer cet estimateur après initialisation, nous procédons en deux

étapes :

- l'étape de prédiction ;
- l'étape de correction.

Nous avons donc :

Initialisation : À l'instant $k = 0$, nous n'avons aucune mesure disponible. La première mesure arrive à l'instant $k = 1$. X_0 , l'état du système à l'instant $k = 0$, est une variable aléatoire, il paraît logique en l'absence de mesure de l'estimer par son espérance. Nous avons alors :

$$\hat{X}_0 = \mathbb{E}(X_0)$$

et la matrice de covariance de l'erreur est :

$$P_0 = \mathbb{E}[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T].$$

Notez que si nous connaissons parfaitement l'état de départ du système nous prenons $\hat{X}_0 = X_0$ et dans ce cas P_0 est nulle.

Étape de prédiction (Évolution du système dynamique) : Nous nous plaçons à l'instant discret k . À cet instant, nous disposons d'une estimation initiale fondée sur la connaissance du processus et des observations jusqu'à l'instant précédent, c'est-à-dire $k - 1$. Nous chercherons un estimateur $\hat{X}_{k|k-1}$ de X_k sans tenir compte de la mesure Y_k , mais en utilisant uniquement l'équation (1) dans le système $(S^{n,m})$. Nous obtenons :

$$\hat{X}_{k|k-1} = A_{k-1}\hat{X}_{k-1|k-1} + B_{k-1}U_{k-1}.$$

Nous définissons l'erreur de cette estimation par :

$$\tilde{X}_{k|k-1} = X_k - \hat{X}_{k|k-1},$$

ainsi que la matrice de covariance de l'erreur :

$$P_{k|k-1} = E[(X_k - \hat{X}_{k|k-1})(X_k - \hat{X}_{k|k-1})^T] = A_{k-1}P_{k-1|k-1}A_{k-1}^T + Q_{k-1}.$$

Étape de correction (Prise en compte de la mesure) : Nous allons maintenant utiliser l'observation Y_k pour améliorer l'estimation $\hat{X}_{k|k-1}$ et obtenir un nouvel estimateur $\hat{X}_{k|k}$ de X_k . Pour cela nous utilisons la Proposition 1.6.3, nous obtenons :

$$\begin{aligned}\hat{X}_{k|k} &= \hat{X}_{k|k-1} + E^*(\tilde{X}_{k|k-1}|\tilde{Y}_{k|k-1}) \\ &= \hat{X}_{k|k-1} + P_{\tilde{X}_{k|k-1}, \tilde{Y}_{k|k-1}} P_{\tilde{Y}_{k|k-1}, \tilde{Y}_{k|k-1}}^{-1} (Y_k - \hat{Y}_{k|k-1}) \\ &= \hat{X}_{k|k-1} + K_k (Y_k - \hat{Y}_{k|k-1})\end{aligned}$$

où $K_k = P_{k|k-1}H_k^T(H_kP_{k|k-1}H_k^T + R_k)^{-1}$ est appelé le gain du filtre de Kalman.

Mais, puisque les observations Y_0, \dots, Y_{k-1} dépendent linéairement que de X_0 et $\{V_j\}_{j=0}^{k-1}$, nous avons alors :

$$\begin{aligned}\hat{Y}_{k|k-1} &= \mathbb{E}^*(Y_k|Y_0, \dots, Y_{k-1}) = \mathbb{E}^*(H_k X_k + V_k|Y_0, \dots, Y_{k-1}) \\ &= H_k \hat{X}_{k|k-1},\end{aligned}$$

ce qui implique que :

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (Y_k - H_k \hat{X}_{k|k-1}). \quad (2.1)$$

Nous pouvons alors définir l'erreur de cette estimation :

$$\tilde{X}_{k|k} = X_k - \hat{X}_{k|k},$$

ainsi que la matrice de covariance de l'erreur :

$$P_{k|k} = E[(X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^T].$$

Remarque 2.3.1. Le gain du filtre de Kalman K_k a les propriétés suivantes :

1. il existe toujours car la matrice R_k est définie positive ;
2. il tient compte des caractéristiques statistiques du bruit des mesures mais ne dépend pas des observations, donc il peut être calculé avant l'étape de correction.

Proposition 2.3.1. On a

1. $\mathbb{E}(\tilde{Y}_{k|k-1}\tilde{Y}_{k|k-1}^T) = H_k P_{k|k-1} H_k^T + R_k.$
2. $\mathbb{E}(\tilde{X}_{k|k-1}\tilde{Y}_{k|k-1}^T) = P_{k|k-1} H_k^T.$

Démonstration.

(1) Nous avons

$$\begin{aligned}\tilde{Y}_{k|k-1} &= Y_k - \hat{Y}_{k|k-1} \\ &= H_k \tilde{X}_{k|k-1} + V_k.\end{aligned}$$

Or V_k est décorrélé avec $\tilde{X}_{k|k-1}$ car $\tilde{X}_{k|k-1}$ dépend linéairement de $X_0, W_0, W_1, \dots, W_{k-2}$.

Donc

$$\mathbb{E}(\tilde{Y}_{k|k-1}\tilde{Y}_{k|k-1}^T) = H_k P_{k|k-1} H_k^T + R_k.$$

(2) Nous avons $\tilde{Y}_{k|k-1} = H_k \tilde{X}_{k|k-1} + V_k$, d'où $\mathbb{E}(\tilde{X}_{k|k-1}\tilde{Y}_{k|k-1}^T) = P_{k|k-1} H_k^T$, car $\tilde{Y}_{k|k-1}$ est décorrélé avec V_k . □

Proposition 2.3.2. La matrice de covariance de l'erreur $\tilde{X}_{k|k}$ est donnée par :

$$P_{k|k} = [I_n - K_k H_k] P_{k|k-1}.$$

Démonstration. Nous avons

$$P_{k|k} = \mathbb{E}[(X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^T]$$

or

$$\begin{aligned} X_k - \widehat{X}_{k|k} &= X_k - \widehat{X}_{k|k-1} - K_k(Y_k - H_k \widehat{X}_{k|k-1}) \\ &= \widetilde{X}_{k|k-1} - K_k(\widetilde{Y}_{k|k-1}). \end{aligned}$$

Ceci implique

$$\begin{aligned} P_{k|k} &= P_{k|k-1} - \mathbb{E}(\widetilde{X}_{k|k-1} \widetilde{Y}_{k|k-1}^T) K_k^T \\ &\quad - K_k \mathbb{E}(\widetilde{Y}_{k|k-1} \widetilde{X}_{k|k-1}^T) + K_k (H_k P_{k|k-1} H_k^T + R_k) K_k^T \\ &= P_{k|k-1} - K_k (H_k P_{k|k-1} H_k^T + R_k)^{-1} K_k^T \\ &= P_{k|k-1} - P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} H_k P_{k|k-1} \\ &= [I_n - K_k H_k] P_{k|k-1}. \end{aligned}$$

□

Remarque 2.3.2. La matrice de covariance $P_{k|k}$, relative à l'erreur d'estimation $\widetilde{X}_{k|k}$ de X_k comporte deux termes :

- $P_{k|k-1}$ est la matrice de covariance de l'erreur à l'issue de l'étape de prédiction qui a été obtenue sans l'utilisation de Y_k .
- $(-K_k H_k P_{k|k-1})$ exprime l'influence de Y_k et entraîne généralement une diminution de $P_{k|k}$, ce qui correspond à un gain en précision, en remplaçant K_k par son expression, nous obtenons :

$$-K_k H_k P_{k|k-1} = -P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} H_k P_{k|k-1}$$

et nous remarquons bien que ce dernier terme est de la forme $-MM^T \leq 0$.

En conséquence l'exploitation correcte d'une mesure Y_k , même de mauvaise qualité, conduit à une amélioration de l'estimateur courant $\widehat{X}_{k|k}$.

Rappelons un résultat important d'inversion de matrice [5].

Lemme 2.3.1. Soit A , B et C des matrices carrées telles que A et C soient inversibles. Alors :

$$(A + BCB^T)^{-1} = A^{-1} - A^{-1}B(C^{-1} + B^T A^{-1}B)^{-1}B^T A^{-1}.$$

Proposition 2.3.3. Si $R_k = \mathbb{E}(V_k V_k^T)$ est inversible nous avons :

1. $P_{k|k}^{-1} = P_{k|k-1}^{-1} + H_k^T R_k^{-1} H_k.$
2. $K_k = P_{k|k} H_k^T R_k^{-1}.$

Démonstration.

1. Il suffit d'appliquer le lemme précédent en posant :

$$A = P_{k|k-1}, \quad B = K_k \quad \text{et} \quad C = (H_k P_{k|k-1} H_k^T + R_k)^{-1}.$$

2. Nous avons $K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1}$, en utilisant le Lemme 2.3.1 nous obtenons :

$$\begin{aligned} (H_k P_{k|k-1} H_k^T + R_k)^{-1} &= R_k^{-1} - R_k^{-1} H_k (P_{k|k-1}^{-1} + H_k^T R_k^T H_k^T)^{-1} H_k^T R_k^{-1} \\ &= R_k^{-1} - R_k^{-1} H_k P_{k|k} H_k^T R_k^{-1}, \end{aligned}$$

ce qui implique que :

$$\begin{aligned} K_k &= P_{k|k-1} H_k^T R_k^{-1} - P_{k|k-1} H_k^T R_k^{-1} H_k P_{k|k} H_k^T R_k^{-1} \\ &= P_{k|k-1} H_k^T R_k^{-1} - P_{k|k-1} (P_{k|k}^{-1} - P_{k|k-1}^{-1}) P_{k|k} H_k^T R_k^{-1}. \end{aligned}$$

d'où le résultat :

$$K_k = P_{k|k} H_k^T R_k^{-1}.$$

□

2.3.1 Algorithme

En résumé l'algorithme proposé est le suivant

$$(K) \left\{ \begin{array}{l} \hat{X}_{0|0} = \mathbb{E}(X_0) \\ P_{0|0} = P_0 \\ \text{pour } k = 1, 2, 3, \dots, \text{ faire} \\ \hat{X}_{k|k-1} = A_{k-1}\hat{X}_{k-1|k-1} + B_{k-1}U_{k-1} \\ P_{k|k-1} = A_{k-1}P_{k-1|k-1}A_{k-1}^T + Q_{k-1} \\ K_k = P_{k|k-1}H_k^T(H_kP_{k|k-1}H_k^T + R_k)^{-1} \\ P_{k|k} = [I - K_kH_k]P_{k|k-1} \\ \hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k(Y_k - H_k\hat{X}_{k|k-1}) \end{array} \right.$$

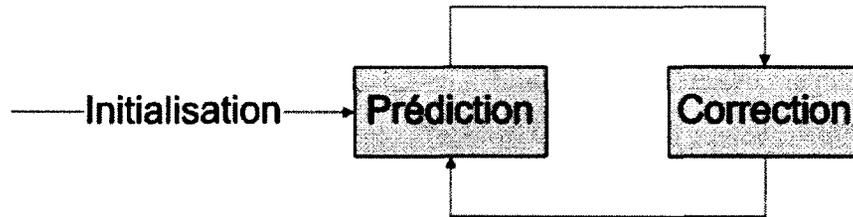


Figure 2.2 – Principe du filtre de Kalman

2.3.2 Innovation

Revenons maintenant sur le terme correctif utilisé dans l'équation de la mise à jour

(2.1) :

$$Y_k - H_k\hat{X}_{k|k-1}.$$

Ce terme s'appelle l'innovation (de l'observation) car il représente les informations supplémentaires apportées par l'observation à l'instant k par rapport aux observa-

tions passées jusqu'à l'instant $k - 1$. Dans la suite l'innovation sera notée par :

$$\tilde{Y}_{k|k-1} = Y_k - H_k \hat{X}_{k|k-1}.$$

Proposition 2.3.4. *L'innovation $\tilde{Y}_{k|k-1}$ est de moyenne nulle et de matrice de covariance*

$$H_k P_{k|k-1} H_k^T + R_k.$$

De plus l'innovation $\tilde{Y}_{k|k-1}$ est décorrélée des innovations précédentes c'est-à-dire

$$\text{Cov}(\tilde{Y}_{k|k-1}, \tilde{Y}_{j|j-1}) = 0, \forall j < k.$$

Démonstration. Nous avons

$$\begin{aligned} \tilde{Y}_{k|k-1} &= Y_k - H_k \hat{X}_{k|k-1} \\ &= H_k \tilde{X}_{k|k-1} + V_k. \end{aligned}$$

Ceci implique $\mathbb{E}(\tilde{Y}_{k|k-1}) = 0$ et $\text{Var}(\tilde{Y}_{k|k-1}) = H_k P_{k|k-1} H_k^T + R_k$ car V_k est décorrélé avec X_k et $\hat{X}_{k|k-1}$ (donc avec $\tilde{X}_{k|k-1}$).

Montrons maintenant que $\text{Cov}(\tilde{Y}_{k|k-1}, \tilde{Y}_{j|j-1}) = 0, \forall j < k$. En effet : $\tilde{Y}_{j|j-1}$ est une combinaison linéaire de Y_0, \dots, Y_{j-1}, Y_j , or $\tilde{Y}_{k|k-1}$ est décorrélé avec Y_0, \dots, Y_{k-1} , d'où le résultat $\text{Cov}(\tilde{Y}_{k|k-1}, \tilde{Y}_{j|j-1}) = 0, \forall j < k$. \square

Remarque 2.3.3. *Dans le cas où les bruits blancs du système $(S^{n,m})$ sont des bruits blancs gaussiens, l'innovation $\{\tilde{Y}_{j|j-1}, j \in \mathbb{N}\}$ est un processus aléatoire gaussien.*

2.3.3 Analyse de la mise à jour

À partir de l'équation de mise à jour

$$\begin{aligned}\hat{X}_{k|k} &= \hat{X}_{k|k-1} + K_k(Y_k - \hat{Y}_{k|k-1}) \\ &= \hat{X}_{k|k-1} + K_k(Y_k - H_k \hat{X}_{k|k-1}),\end{aligned}$$

nous voyons que :

- le premier terme de droite s'écrit encore $\hat{X}_{k|k-1}$, c'est l'étape de prédiction de X_k en utilisant uniquement le modèle, mais pas l'observation ;
- le second terme est la correction de la prédiction en utilisant les observations.

Nous pouvons aussi remarquer que le gain de Kalman varie selon la confiance que l'on peut accorder aux mesures. En effet le gain de Kalman peut s'écrire :

$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1},$$

- Si l'observation est de mauvaise qualité, c'est-à-dire si le bruit de mesure a une forte variance (R_k est élevé), et que nous avons confiance dans les estimations précédentes ($P_{k|k-1}$ est faible), le gain du filtre de Kalman

$$K_k \approx 0$$

et la mise à jour sera réduite à

$$\hat{X}_{k|k} \approx \hat{X}_{k|k-1}.$$

- Au contraire, si l'observation est très fiable (R_k est faible) et que nous avons des doutes sur les estimations précédentes, le gain du filtre de Kalman devient grand, de sorte que la mise à jour repose de façon prépondérante sur l'observation et très peu sur le modèle.

D'après ce qui précède nous pouvons conclure que le gain de Kalman fluctue et

- le gain K_k diminue, si l'estimation par le modèle devient plus précise ;
- le gain K_k augmente, si les observations deviennent plus précises.

2.4 Modèle à bruits corrélés

Dans le cas où les bruits de la dynamique de l'état et de l'observation sont corrélés, il est possible de reprendre toute l'étude précédente. Une façon d'établir les équations du filtre de Kalman consiste à se ramener au cas d'un modèle à bruits décorrélés par la construction d'un modèle équivalent. En effet, supposons qu'on a :

$$(S_1^{n,m}) \begin{cases} X_{k+1} = A_k X_k + B_k U_k + W_k \\ Y_k = H_k X_k + V_k \end{cases}$$

avec

$$\mathbb{E}(V_k W_l^T) = S_k^T \delta_{kl}.$$

Nous posons le nouveau vecteur de bruit de dynamique :

$$\bar{W}_k = W_k - S_k R_k^{-1} V_k,$$

pour lequel il est facile de vérifier que $\mathbb{E}(V_k \bar{W}_k) = 0$. L'utilisation de ce bruit, dans les équations du modèle initial $(S_1^{n,m})$, conduit directement à un nouveau modèle $(S_2^{n,m})$ équivalent à $(S_1^{n,m})$:

$$(S_2^{n,m}) \begin{cases} X_{k+1} = \bar{A}_k X_k + B_k U_k + \bar{G}_k Y_k + \bar{W}_k \\ Y_k = H_k X_k + V_k \end{cases}$$

avec $\bar{A}_k = A_k - S_k R_k^{-1} H_k$ et $\bar{G}_k = S_k R_k^{-1}$. Les propriétés statistiques des bruits de ce modèle sont telles que :

- $\mathbb{E}(V_k) = 0, \mathbb{E}(\overline{W}_k) = 0.$
- $\mathbb{E}(V_k V_l^T) = R_k \delta_{kl}$ et $\mathbb{E}(\overline{W}_k \overline{W}_l^T) = \overline{Q}_k \delta_{kl}$ avec $\overline{Q}_k = Q_k - S_k R_k^{-1} S_k^T.$
- $\mathbb{E}(\overline{W}_k V_l^T) = 0.$

Il suffit alors d'appliquer, sur le système $(S_2^{n,m})$, le filtre de Kalman discret défini dans le cas d'un modèle à bruits non corrélés. Nous obtenons alors :

- Estimation :

$$\begin{aligned}\widehat{X}_{k|k-1} &= \overline{A}_{k-1} \widehat{X}_{k-1|k-1} + B_{k-1} U_{k-1} + \overline{G}_{k-1} Y_{k-1} \\ P_{k|k-1} &= \overline{A}_{k-1} P_{k-1|k-1} \overline{A}_{k-1}^T + \overline{Q}_{k-1} \\ &= (A_{k-1} - S_{k-1} R_{k-1}^{-1} H_{k-1}) P_{k-1|k-1} (A_{k-1} - S_{k-1} R_{k-1}^{-1} H_{k-1})^T + Q_{k-1} - S_{k-1} R_{k-1}^{-1} S_{k-1}^T.\end{aligned}$$

- Correction :

$$\begin{aligned}\widehat{X}_{k|k} &= \widehat{X}_{k|k-1} + K_k (Y_k - H_k \widehat{X}_{k|k-1}) \\ P_{k|k} &= (I - K_k H_k) P_{k|k-1},\end{aligned}$$

avec $K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1}$ comme gain du filtre de Kalman. Dans ce cas l'algorithme du filtre de Kalman est donné par :

$$(K) \left\{ \begin{array}{l} \widehat{X}_{0|0} = \mathbb{E}(X_0) \\ P_{0|0} = P_0 \\ \text{pour } k = 1, 2, 3, \dots, \text{ faire} \\ \widehat{X}_{k|k-1} = A_{k-1} \widehat{X}_{k-1|k-1} + B_{k-1} U_{k-1} + S_{k-1} R_{k-1}^{-1} Y_{k-1} \\ P_{k|k-1} = (A_{k-1} - S_{k-1} R_{k-1}^{-1} H_{k-1}) P_{k-1|k-1} (A_{k-1} - S_{k-1} R_{k-1}^{-1} H_{k-1})^T \\ \quad + Q_{k-1} - S_{k-1} R_{k-1}^{-1} S_{k-1}^T \\ K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \\ P_{k|k} = [I - K_k H_k] P_{k|k-1} \\ \widehat{X}_{k|k} = \widehat{X}_{k|k-1} + K_k (Y_k - H_k \widehat{X}_{k|k-1}) \end{array} \right.$$

2.5 Exemple : Position-Vitesse

Pour illustrer la théorie développée précédemment, nous allons considérer un problème de poursuite. Nous désirons suivre une cible, en supposant que la dynamique de celle-ci est à peu près linéaire et de vitesse à peu près constante. Cela n'est jamais le cas en réalité (d'où le terme "à peu près") puisque la cible peut tourner légèrement, ou encore accélérer, et c'est pour cette raison qu'un modèle aléatoire se prête bien à la modélisation de la dynamique de la cible. Nous disposons d'un capteur qui relève périodiquement des informations sur la position de la cible (signal observé), la période d'observation est notée T . Le capteur n'est pas parfait ce qui induit des erreurs de mesure. Le but sera d'estimer la trajectoire et la vitesse du mobile en fonction du temps à l'aide du filtre de Kalman. Nous modélisons le problème sous forme de deux équations, appelées équation d'état (qui porte sur la dynamique de la cible) et équation d'observation (qui porte sur les mesures du capteur). Ici, on s'intéresse à la position de la cible dans le plan que nous notons par $\vec{x}(t) = (x_1(t), x_2(t))$ et à sa vitesse $\vec{v}(t) = (v_1(t), v_2(t))$, avec $t \in \mathbb{R}$. En effectuant un développement limité d'ordre 2, nous obtenons :

$$\begin{aligned}x_i((k+1)T) &= x_i(kT) + Tv_i(kT) + \frac{T^2}{2}\epsilon_{i,k} \\v_i((k+1)T) &= v_i(kT) + T\epsilon_{i,k},\end{aligned}$$

pour $i = 1, 2$. En posant comme vecteur d'état $X_k = [x_1(k), x_2(k), v_1(k), v_2(k)]^T$, nous obtenons l'équation d'état :

$$X_{k+1} = AX_k + BW_k \tag{2.2}$$

avec

$$A = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} \frac{T^2}{2} & 0 \\ 0 & \frac{T^2}{2} \\ T & 0 \\ 0 & T \end{bmatrix} \text{ et } W_k = \begin{bmatrix} \epsilon_{1,k} \\ \epsilon_{2,k} \end{bmatrix}$$

où W_k est un bruit blanc gaussien centré avec matrice de variance-covariance $\begin{bmatrix} \sigma_{1,k}^2 & 0 \\ 0 & \sigma_{2,k}^2 \end{bmatrix}$.

Puisque nous observons la position du mobile à chaque période T , l'équation d'observation est donnée par :

$$Y_k = HX_k + \xi_k \quad (2.3)$$

où

- $H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ est la matrice de mesure ;

- $\xi_k = \begin{bmatrix} \xi_{1,k} \\ \xi_{2,k} \end{bmatrix}$ est un bruit blanc gaussien centré avec matrice de variance-covariance $\begin{bmatrix} \sigma_{1,k}^2 & 0 \\ 0 & \sigma_{2,k}^2 \end{bmatrix}$.

Supposons que l'équation de la trajectoire (théorique) est de la forme :

$$\begin{cases} x_1(n) = n^2 \\ x_2(n) = \sin(2\pi f_0 n) \end{cases}$$

avec $f_0 = 0.005$. Cela implique que le vecteur vitesse (théorique) est :

$$\begin{cases} v_1(n) = 2n \\ v_2(n) = 2\pi f_0 \cos(2\pi f_0 n) \end{cases}$$

Simulation sous Matlab

Pour notre exemple, fixons $\sigma_{1,k} = \sigma_{2,k} = 0.01$. Pour différentes valeurs de $\sigma'_{1,k}$ et $\sigma'_{2,k}$ nous estimons la trajectoire et la vitesse à partir des observations bruitées grâce au filtre de Kalman, nous affichons dans les Figures 2.4, 2.6 et 2.8 à gauche la trajectoire théorique versus la trajectoire estimée (x_1 en abscisse et x_2 en ordonnée)

et à droite la vitesse théorique versus la vitesse estimée (v_1 en abscisse et v_2 en ordonnée), tandis que les Figures 2.3, 2.5 et 2.7 montrent la trajectoire théorique versus les observations bruitées (x_1 en abscisse et x_2 en ordonnée). Nous remarquons, d'après les Figures 2.4, 2.6 et 2.8 que la qualité d'estimation se dégrade au fur et à mesure qu'on augmente la variance du bruit dans les observations (voir les Figures 2.3, 2.5 et 2.7).

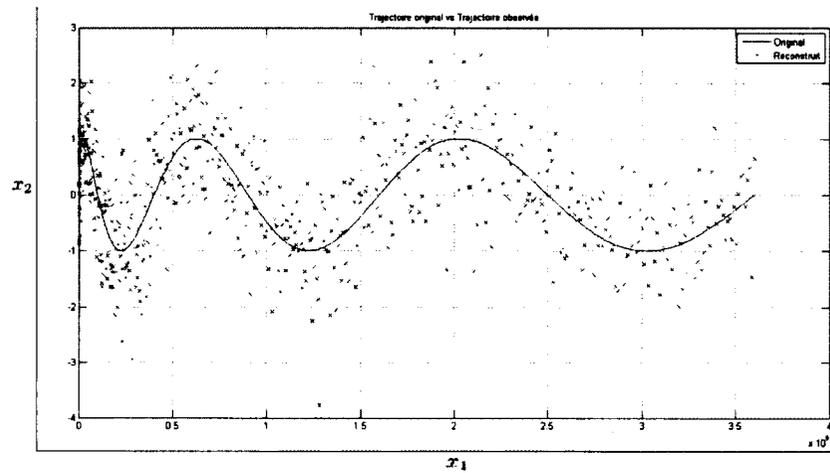


Figure 2.3 – Trajectoire théorique vs Observations bruitées avec $\sigma'_i = 0.1$.

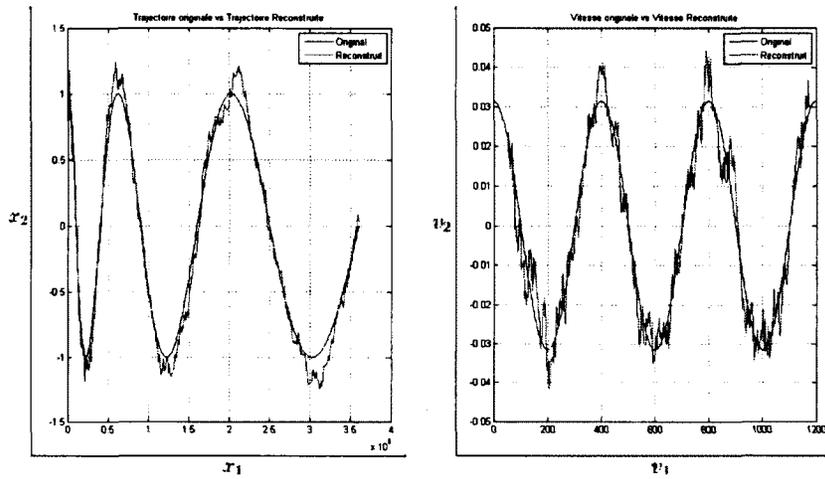


Figure 2.4 – Filtre de Kalman pour l'estimation du vecteur Position-Vitesse avec $\sigma'_i = 0.1$ et $\sigma_i = 0.01$. Graphique de gauche : trajectoire théorique (en bleu) et trajectoire estimée (en rouge). Graphique de droite : vitesse théorique (en bleu) et vitesse estimée (en rouge).

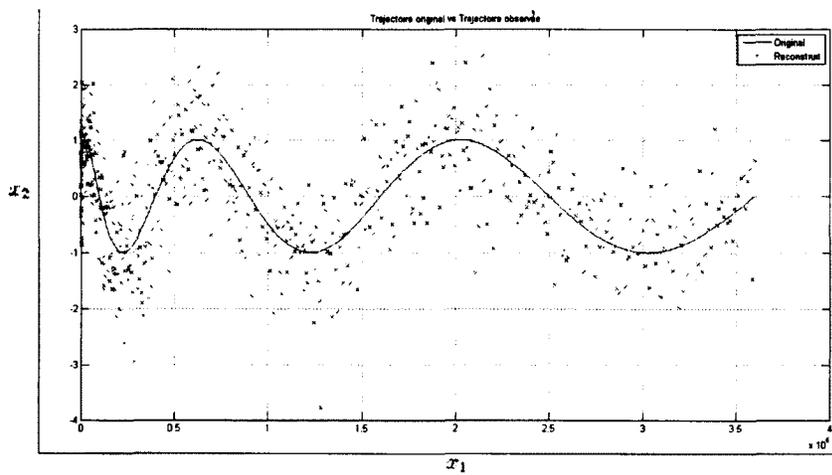


Figure 2.5 – Trajectoire théorique vs Observations bruitées avec $\sigma'_i = 0.7$.

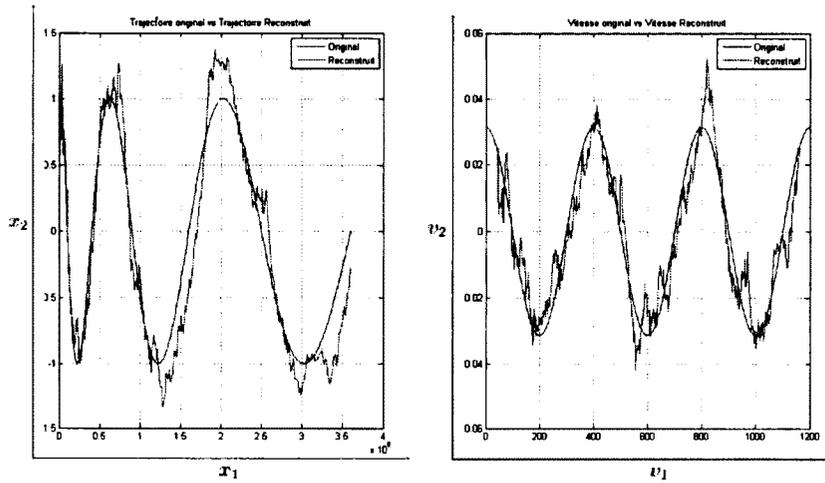


Figure 2.6 – Filtre de Kalman pour l'estimation du vecteur Position-Vitesse avec $\sigma'_i = 0.7$ et $\sigma_i = 0.01$. Graphique de gauche : trajectoire théorique (en bleu) et trajectoire estimée (en rouge). Graphique de droite : vitesse théorique (en bleu) et vitesse estimée (en rouge).

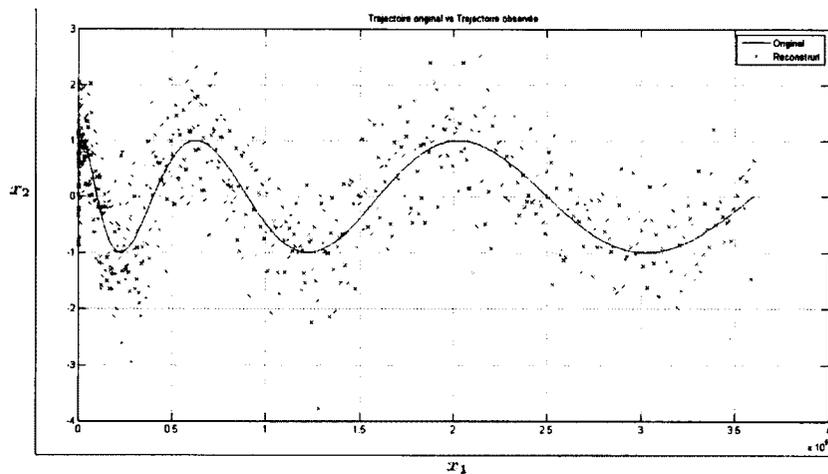


Figure 2.7 – Trajectoire théorique vs Observations bruitées avec $\sigma'_i = \sqrt{2}$.

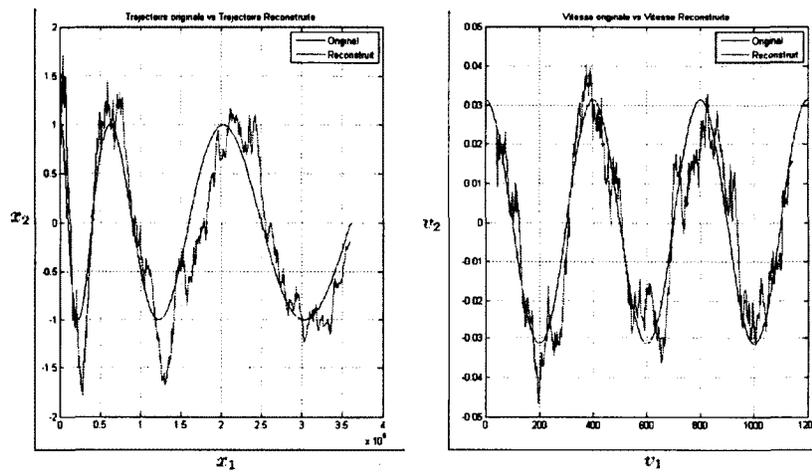


Figure 2.8 – Filtre de Kalman pour l'estimation du vecteur Position-Vitesse avec $\sigma'_i = \sqrt{2}$ et $\sigma_i = 0.01$. Graphique de gauche : trajectoire théorique (en bleu) et trajectoire estimée (en rouge). Graphique de droite : vitesse théorique (en bleu) et vitesse estimée (en rouge)

CHAPITRE 3

Filtrage avec bruits de moyennes inconnues

3.1 Introduction

Nous proposons dans ce chapitre une méthode de reconstruction de signal à l'aide du filtre de Kalman dans le cas où les bruits sont non centrés et de moyennes inconnues. L'idée est de développer un algorithme similaire à celui du filtre de Kalman qui permet de reconstruire le signal tout en estimant les moyennes inconnues des bruits.

3.2 Modèle à bruits non centrés

Considérons le problème classique d'estimation de signal à l'aide du filtre de Kalman lorsque les bruits sur la trajectoire et l'observation ne sont pas centrés c'est-à-dire :

$$(S^{n,m}) \quad \begin{cases} X_{k+1} = A_k X_k & + \epsilon_{x,k} \\ Y_{k+1} = H_{k+1} X_{k+1} & + \epsilon_{y,k+1} \end{cases}$$

où $X_k \in \mathbb{R}^n$ et $Y_k \in \mathbb{R}^m$ avec $\epsilon_{x,k}$, $\epsilon_{y,l}$ des vecteurs aléatoires de moyennes connues m_x et m_y et de variances connues Q_k et R_k respectivement tels que

$$\mathbb{E} [(\epsilon_{x,k} - m_x)(\epsilon_{y,l} - m_y)^T] = 0.$$

L'état initial X_0 du processus est d'espérance $\mathbb{E}(X_0)$ et de variance P_0 connues et est décorrélé des bruits $\epsilon_{x,k}$ et $\epsilon_{y,k}$. L'objectif du filtre de Kalman sera de déterminer, pour chaque instant k , une estimation optimale et récursive $\hat{X}_{k|k}$ de l'état X_k , ainsi que la matrice de covariance de l'erreur $P_{k|k}$, à partir des observations Y_1, Y_2, \dots, Y_k .

Pour ce problème, le filtre de Kalman donne l'algorithme suivant[15, 5] :

$$(K) \quad \left\{ \begin{array}{l} \hat{X}_{0|0} = \mathbb{E}(X_0) \\ P_{0|0} = P_0 \\ \text{pour } k = 1, 2, 3, \dots, \text{ faire} \\ \hat{X}_{k|k-1} = A_{k-1} \hat{X}_{k-1|k-1} + m_x \\ P_{k|k-1} = A_{k-1} P_{k-1|k-1} A_{k-1}^T + Q_{k-1} \\ K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \\ P_{k|k} = [I - K_k H_k] P_{k|k-1} \\ \hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (Y_k - (H_k \hat{X}_{k|k-1} + m_y)). \end{array} \right.$$

Cette méthode permet de donner une bonne estimation $\hat{X}_{k|k}$ du signal X_k lorsque les moyennes m_x et m_y sont connues exactement. Donnons une illustration qui montre

l'impact de la connaissance des moyennes sur l'estimation du signal.

Considérons l'exemple simple suivant :

$$(E) \quad \begin{cases} X_{k+1} = (0.9 + 0.2 * (-1)^k) * X_k + \epsilon_{x,k} \\ Y_{k+1} = 0.5 * X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

avec des bruits gaussiens $\epsilon_{x,k} \sim \mathcal{N}(2, (0.6)^2)$ et $\epsilon_{y,k} \sim \mathcal{N}(5, (0.6)^2)$.

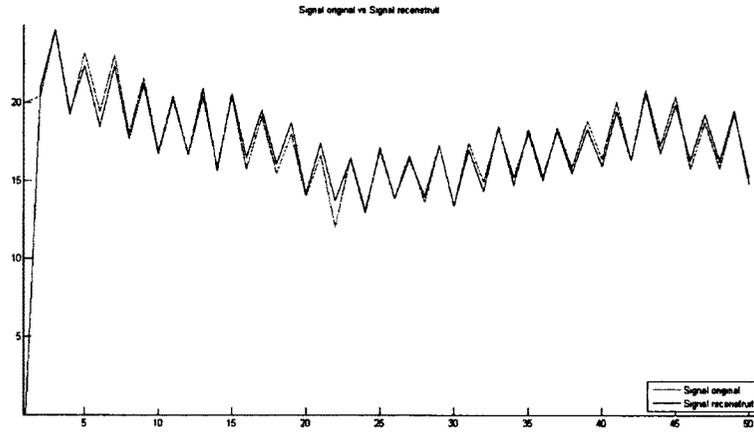
Utilisons l'algorithme (K) avec les moyennes exactes, c'est-à-dire $m_x = 2$ et $m_y = 5$. Le résultat, donné à la Figure 3.1(a), illustre bien que l'estimation du signal suit très bien ou reconstruit très bien le signal. Par contre si nous utilisons cet algorithme avec des valeurs arbitraires pour m_x et m_y au lieu des valeurs exactes, ici $m_x = 0$ et $m_y = 0$, la Figure 3.1(b) illustre bien que le signal n'est pas bien reconstruit. D'après cet exemple simple, nous concluons que la connaissance des moyennes est primordiale pour une estimation adéquate du signal à l'aide de l'algorithme (K). Dans la section suivante nous proposons une façon de reconstruire le signal tout en estimant les moyennes des bruits lorsque ces moyennes sont inconnues.

3.3 Problème avec estimation des moyennes

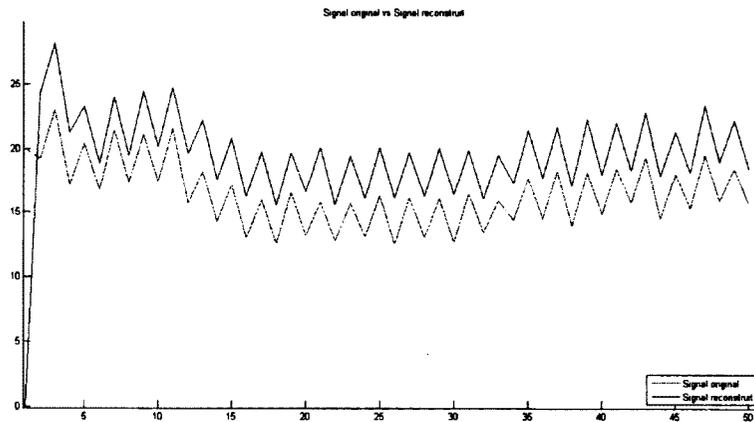
Considérons le système $(S^{n,m})$ en supposant que les moyennes m_x et m_y sont inconnues. Nous voulons simultanément estimer les deux moyennes inconnues et reconstruire le signal.

Le système $(S^{n,m})$ est équivalent au système suivant :

$$(S_1^{n,m}) \quad \begin{cases} X_{k+1} = A_k X_k + m_x + \tilde{\epsilon}_{x,k} \\ Y_{k+1} = H_{k+1} X_{k+1} + m_y + \tilde{\epsilon}_{y,k+1} \end{cases}$$



(a) Vraies moyennes : $m_x = 2$ et $m_y = 5$



(b) Moyennes arbitraires : $m_x = 0$ et $m_y = 0$

Figure 3.1 – Reconstruction du signal avec (K) : $P_0 = 10^4$.

avec les vecteurs aléatoires de bruits $\tilde{\epsilon}_{x,k}$ et $\tilde{\epsilon}_{y,k}$ de moyennes nulles et de variance R_k et Q_k respectivement. On modifie la dynamique du système en ajoutant les deux équations d'état suivantes :

$$\begin{cases} m_{x,k+1} = m_{x,k} + \zeta_{x,k} \\ m_{y,k+1} = m_{y,k} + \zeta_{y,k} \end{cases} \quad (3.1)$$

où

- $\varsigma_{x,k} \sim \mathcal{N}(0, \Sigma_{x,k})$ est non corrélé avec $\epsilon'_{x,l}$ et $\epsilon'_{y,l}$;
- $\varsigma_{y,k} \sim \mathcal{N}(0, \Sigma_{y,k})$ est non corrélé avec $\epsilon'_{x,l}$ et $\epsilon'_{y,l}$;
- $\varsigma_{x,k}$ et $\varsigma_{y,k}$ sont non corrélés .

Nous convenons que lorsque $\Sigma_{*,k} = 0$ alors $\varsigma_{*,k} \equiv 0$, c'est-à-dire que l'équation d'état pour la moyenne correspondante est sans bruit.

Le système $(S_1^{n,m})$ devient alors

$$(S_2^{n,m}) \quad \begin{cases} X_{k+1} = A_k X_k & + m_{x,k} & + \tilde{\epsilon}_{x,k} \\ m_{x,k+1} = & m_{x,k} & + \varsigma_{x,k} \\ m_{y,k+1} = & m_{y,k} & + \varsigma_{y,k} \\ Y_{k+1} = H_{k+1} X_{k+1} & + m_{y,k+1} & + \tilde{\epsilon}_{y,k+1} \end{cases}$$

qu'on peut réécrire

$$\begin{cases} \begin{bmatrix} X_{k+1} \\ m_{x,k+1} \\ m_{y,k+1} \end{bmatrix} = \begin{bmatrix} A_k & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} X_k \\ m_{x,k} \\ m_{y,k} \end{bmatrix} + \begin{bmatrix} \tilde{\epsilon}_{x,k} \\ \varsigma_{x,k} \\ \varsigma_{y,k} \end{bmatrix}, \\ Y_{k+1} = [H_{k+1} \ 0 \ I] \begin{bmatrix} X_{k+1} \\ m_{x,k+1} \\ m_{y,k+1} \end{bmatrix} + \tilde{\epsilon}_{y,k+1}. \end{cases}$$

En posant $Z_k = \begin{bmatrix} X_k \\ m_{1,k} \\ m_{2,k} \end{bmatrix}$, $\tilde{A}_k = \begin{bmatrix} A_k & I & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$, $\tilde{H}_k = [H_k \ 0 \ I]$, $\tilde{\epsilon}_k = \begin{bmatrix} \tilde{\epsilon}_{x,k} \\ \varsigma_{x,k} \\ \varsigma_{y,k} \end{bmatrix}$
 et $\tilde{Q}_k = \begin{bmatrix} Q_k & 0 & 0 \\ 0 & \Sigma_{x,k} & \\ 0 & 0 & \Sigma_{y,k} \end{bmatrix}$, nous obtenons la forme matricielle suivante :

$$(S_3^{n,m}) \quad \begin{cases} Z_{k+1} = \tilde{A}_k Z_k & + \tilde{\epsilon}_k \\ Y_{k+1} = \tilde{H}_{k+1} Z_{k+1} & + \tilde{\epsilon}_{y,k+1}. \end{cases}$$

Appliquons le filtre de Kalman au système (S₃), nous obtenons :

$$(\tilde{K}) \left\{ \begin{array}{l} \hat{Z}_{0|0} = \mathbb{E}(Z_0) \\ \tilde{P}_{0|0} = \tilde{P}_0 \\ \text{pour } k = 1, 2, 3, \dots, \text{ faire} \\ \hat{Z}_{k|k-1} = \tilde{A}_{k-1} \hat{Z}_{k-1|k-1} \\ \tilde{P}_{k|k-1} = \tilde{A}_{k-1} \tilde{P}_{k-1|k-1} \tilde{A}_{k-1}^T + \tilde{Q}_{k-1} \\ \tilde{K}_k = \tilde{P}_{k|k-1} \tilde{H}_k (\tilde{H}_k \tilde{P}_{k|k-1} \tilde{H}_k^T + R_k)^{-1} \\ \tilde{P}_{k|k} = [I - \tilde{K}_k \tilde{H}_k] \tilde{P}_{k|k-1} \\ \hat{Z}_{k|k} = \hat{Z}_{k|k-1} + \tilde{K}_k (Y_k - \tilde{H}_k \hat{Z}_{k|k-1}) \end{array} \right.$$

avec $\mathbb{E}(Z_0) = \begin{bmatrix} \mathbb{E}(X_0) \\ m_{x,0} \\ m_{y,0} \end{bmatrix}$ et $\tilde{P}_{0|0} = \begin{bmatrix} P_0 & 0 & 0 \\ 0 & \Sigma_{x,0} & 0 \\ 0 & 0 & \Sigma_{y,0} \end{bmatrix}$. Les valeurs initiales $m_{x,0}$ et $m_{y,0}$ étant arbitraires, nous initialisons leurs variances $\Sigma_{x,0}$ et $\Sigma_{y,0}$ à valeurs non nulles et possiblement assez grandes étant donné l'incertitude sur les vraies valeurs m_x et m_y .

3.4 Expérimentation

Pour les équations d'évolutions de moyennes inconnues (3.1) nous allons tester trois types d'équations :

– sans variabilité c'est-à-dire $\Sigma_{*,k} = 0$ et ainsi $\varsigma_{*,k} \equiv 0$ et

$$m_{*,k+1} = m_{*,k} \quad (3.2)$$

– avec variance décroissante c'est-à-dire

$$m_{*,k+1} = m_{*,k} + \varsigma_{*,k} \quad (3.3)$$

où $\varsigma_{*,k} \sim \mathcal{N}(0, \Sigma_{*,k})$, $\Sigma_{*,k} = \frac{1}{1+k^\lambda} \Sigma_{*,0}$ avec $\lambda \in \mathbb{R}_+$.

– avec variance constante c'est-à-dire

$$m_{*,k+1} = m_{*,k} + \varsigma_{*,k} \quad (3.4)$$

avec $\varsigma_{*,k} \sim \mathcal{N}(0, \Sigma_*)$.

Revenons à l'exemple (E) en utilisant ce schéma. Avec l'initialisation $m_{x,0} = 0$ et $m_{y,0} = 0$, et de relativement grandes variances sur ces valeurs, $\Sigma_{x,0} = 10^4$ et $\Sigma_{y,0} = 10^4$, nous obtenons les résultats illustrés aux Figures 3.2, 3.3 et 3.4. Nous observons qu'après quelques étapes nous obtenons à la fois une bonne reconstruction du signal tout en obtenant de bonnes estimations des moyennes $m_x \cong 2$ et $m_y \cong 5$.

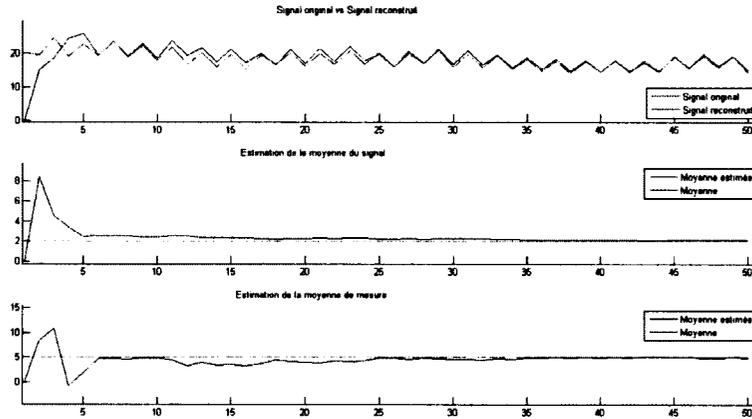


Figure 3.2 – Reconstruction du signal et estimation des moyennes inconnues via la méthode sans variabilité où $\hat{Z}_{0|0} = 0$ et $\hat{P}_{0|0} = 10^4 I_3$.

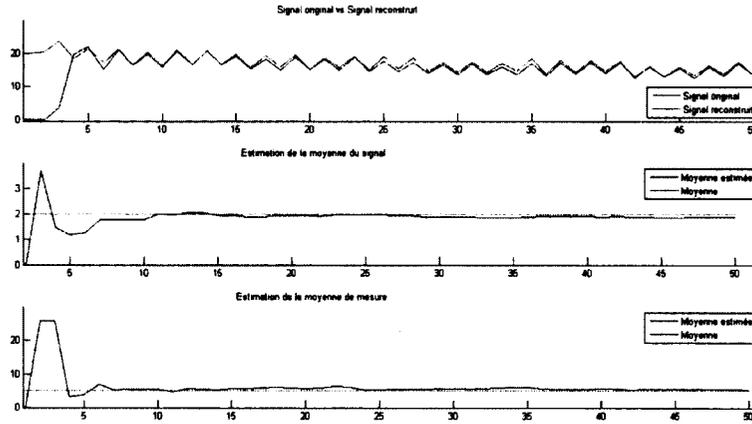


Figure 3.3 – Reconstruction du signal et estimation des moyennes inconnues via la méthode avec variance décroissante où $\hat{Z}_{0|0} = 0$, $\tilde{P}_{0|0} = 10^4 I_3$ et $\lambda = 45$.

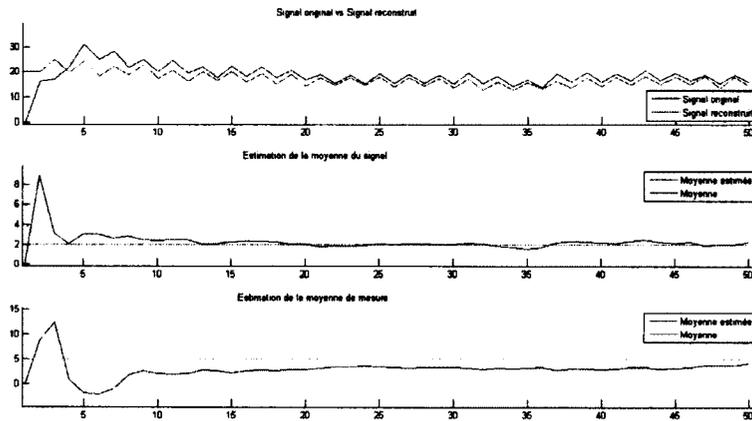


Figure 3.4 – Reconstruction du signal et estimation des moyennes inconnues via la méthode avec variance constante ($\Sigma_* = 0.04$) où $\hat{Z}_{0|0} = 0$ et $\tilde{P}_{0|0} = 10^4 I_3$.

3.5 Calcul des erreurs de reconstruction

Nous donnons ici quelques calculs d'erreurs de reconstruction du signal et des moyennes afin de comparer les trois types d'équations citées ci-dessus. Nous utilisons un échantillon de taille 500. Plus précisément nous calculons la moyenne de l'erreur de reconstruction du signal et les moyennes sur l'erreur d'estimation des moyennes inconnues en reproduisant 500 fois la dynamique du système avec initialisation aléatoire de l'état initial X_0 du système.

Nous posons :

- Moyenne de l'erreur relative sur l'estimation de la moyenne m_* (pour $* = x, y$).
Pour chaque répétition j nous calculons l'erreur relative de moindres carrées sur les 10 étapes $k = N - 9, \dots, N$. Nous avons

$$err_*^{(N)}(j) = \frac{\sqrt{\frac{1}{10} \sum_{k=N-9}^N |m_* - m_{*,k}(j)|^2}}{|m_*|}$$

et nous prenons la moyenne de ces erreurs relatives

$$\overline{err}_*^{(N)} = \frac{1}{500} \sum_{j=1}^{500} err_*^{(N)}(j).$$

- Moyenne de l'erreur relative sur le signal. Comme pour les moyennes, pour chaque répétition j nous calculons l'erreur relative de moindres carrées du signal sur les 10 étapes $k = N - 9, \dots, N$. Nous avons

$$Err^{(N)}(j) = \frac{\|X(j) - \widehat{X}(j)\|_2}{\|X(j)\|_2} = \sqrt{\frac{\sum_{k=N-9}^N |X_k(j) - \widehat{X}_{k|k}(j)|^2}{\sum_{k=N-9}^N |X_k(j)|^2}}$$

et

$$Moy^{(N)} = \frac{1}{500} \sum_{j=1}^{500} Err^{(N)}(j).$$

Le Tableau 3.1 donne les erreurs de reconstruction et d'estimation des moyennes et du signal après $N = 50$ étapes du filtre pour notre exemple (E).

	R_k	1	0.6^2	0.1^2	1	0.1^2
	Q_k	1	0.6^2	0.1^2	0.1^2	1
Sans variance	$\overline{err}_x^{(50)}$	0.0902	0.0420	0.0182	0.0521	0.1152
	$\overline{err}_y^{(50)}$	0.1003	0.0600	0.0092	0.1602	0.0495
	$Moy^{(50)}$	0.0901	0.0308	0.0102	0.0542	0.0480
Variance décroissante	$\overline{err}_x^{(50)}$	0.0901	0.0425	0.0110	0.0552	0.1123
	$\overline{err}_y^{(50)}$	0.0998	0.0598	0.0120	0.1623	0.0402
	$Moy^{(50)}$	0.0902	0.0509	0.0162	0.0556	0.0496
variance constante ($\Sigma_x^2 = 0.04$)	$\overline{err}_x^{(50)}$	0.1598	0.1202	0.0892	0.0905	0.1409
	$\overline{err}_y^{(50)}$	0.1803	0.1490	0.0930	0.1417	0.0820
	$Moy^{(50)}$	0.1609	0.1285	0.0899	0.1010	0.1093

Tableau 3.1 – Variation de l'erreur sur l'estimation des moyennes et le signal pour les méthodes : sans variabilité, variance décroissante et variance constante pour (E_2) avec $\hat{Z}_{0|0} = 0$, $\tilde{P}_{0|0} = 10^4 I_3$, $\lambda = 45$ et $N = 50$.

Calculons également la décroissance de l'erreur tout au long de l'estimation des moyennes inconnues et la reconstruction du signal pour les deux méthodes sans variance et variance décroissante. Le Tableau 3.2 illustre cette décroissance pour $N = 10, 15, 20, 25, 30, 35, 40, 45, 50$.

	N	10	15	20	25	30	35	40	45	50
Sans variabilité	$\overline{err}_1^{(N)}$	0.2378	0.1424	0.0909	0.0888	0.0760	0.0610	0.0520	0.0497	0.0418
	$\overline{err}_2^{(N)}$	0.3545	0.2915	0.2015	0.1536	0.1109	0.0905	0.0813	0.0684	0.0598
	$Moy^{(N)}$	0.1068	0.0885	0.0750	0.0662	0.0599	0.0455	0.0404	0.0392	0.0306
variance décroissante	$\overline{err}_1^{(N)}$	0.2480	0.1434	0.1010	0.0884	0.0758	0.0633	0.0567	0.0478	0.0423
	$\overline{err}_2^{(N)}$	0.3714	0.3002	0.2050	0.1534	0.1107	0.1004	0.0822	0.0696	0.0597
	$Moy^{(N)}$	0.1168	0.0965	0.0898	0.0792	0.0708	0.0675	0.0595	0.0555	0.0508

Tableau 3.2 – Décroissance de l'erreur sur l'estimation des moyennes et la reconstruction du signal avec (\tilde{K}) : $\hat{Z}_{0|0} = 0$, $\tilde{P}_{0|0} = 10^4 I_3$, $R_k = Q_k = (0.6)^2$, $\lambda = 45$ et $N = 50$.

D'après les Tableaux 3.1 et 3.2 on remarque bien que la méthode sans variabilité est la plus appropriée pour la reconstruction du signal et l'estimation des moyennes inconnues comme le montrent les Figures 3.2, 3.3 et 3.4.

Finalement, notons que si on utilise la méthode sans variance et que nous initialisons (\tilde{K}) avec une matrice de covariance du type $\tilde{P}_{0|0} = \begin{bmatrix} P_0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, c'est-à-dire $\Sigma_{x,0} = 0$ et $\Sigma_{y,0} = 0$, nous allons obtenir le même résultat que celui donné par l'algorithme du filtre de Kalman (K) appliqué directement sur le système (S). En effet, puisque nous avons

$$\tilde{K}_k = \tilde{P}_{k|k-1} \tilde{H}_k^T (\tilde{H}_k \tilde{P}_{k|k-1} \tilde{H}_k^T + R_k)^{-1}$$

et que $\tilde{P}_{k|k-1}$ est de la forme

$$\tilde{P}_{k|k-1} = \begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix}.$$

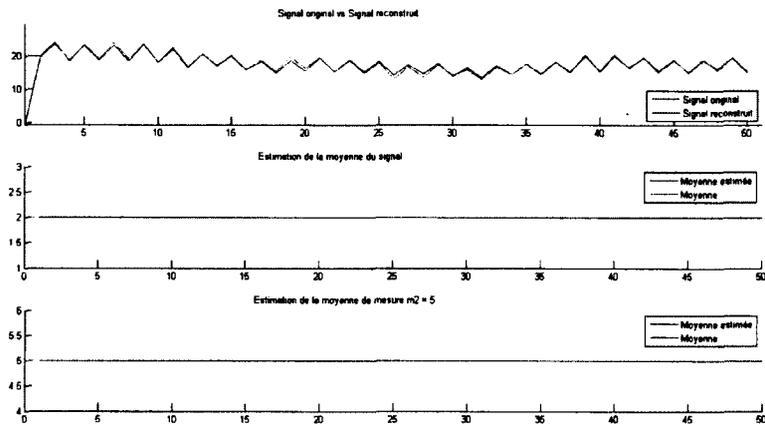
nous obtenons que $\tilde{K}_k = \begin{bmatrix} K_k \\ 0 \\ 0 \end{bmatrix}$, ce qui implique, d'après (\tilde{K}) , que

$$\hat{Z}_{k|k} = \hat{Z}_{k|k-1} + \begin{bmatrix} K_k \\ 0 \\ 0 \end{bmatrix} (Y_k - \tilde{H}_k \hat{Z}_{k|k-1}).$$

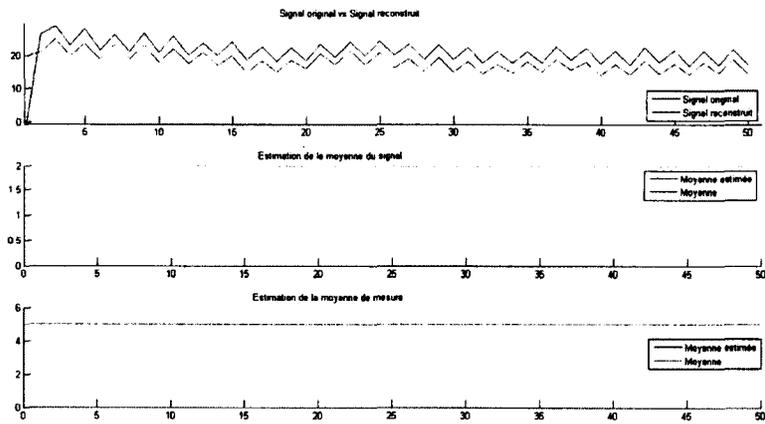
Ainsi d'après ce qui précède nous remarquons que :

$$\begin{cases} m_{x,k+1} = m_{x,k} = m_{x,0} \\ m_{y,k+1} = m_{y,k} = m_{y,0}. \end{cases}$$

Cette remarque est illustrée en comparant les Figures 3.1 et 3.5 qui donnent les mêmes résultats à l'aide des méthodes (K) et (\tilde{K}) sur l'exemple (E).



(a) Vraies moyennes : $m_{x,0} = 2 = m_x$ et $m_{y,0} = 5 = m_y$



(b) Moyennes quelconques : $m_{x,0} = 0$ et $m_{y,0} = 0$

Figure 3.5 – Reconstruction du signal et estimation des moyennes inconnues avec $(\tilde{K}) : \Sigma_{x,0} = 0$ et $\Sigma_{y,0} = 0$.

Remarque 3.5.1. On va donner un exemple multidimensionnel pour appliquer l'al-

gorithme (\tilde{K}). Considérons le système dynamique dans \mathbb{R}^2

$$(S^{2,1}) \quad \begin{cases} X_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} X_k + \epsilon_{x,k} \\ Y_{k+1} = [1 \ 2] X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

avec

- $\epsilon_{x,k}$ étant un vecteur aléatoire gaussien dans \mathbb{R}^2 de moyenne $(5, 7)$ et de matrice de variance-covariance $\begin{bmatrix} 0.36 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$
- $\epsilon_{y,k}$ étant une variable aléatoire gaussienne de moyenne 10 et de variance 0.3.

La Figure 3.6 montre l'application de l'algorithme (\tilde{K}) sur le système $(S^{2,1})$ pour estimer le vecteur d'état (qui est dans \mathbb{R}^2), la moyenne de l'état et de la mesure.

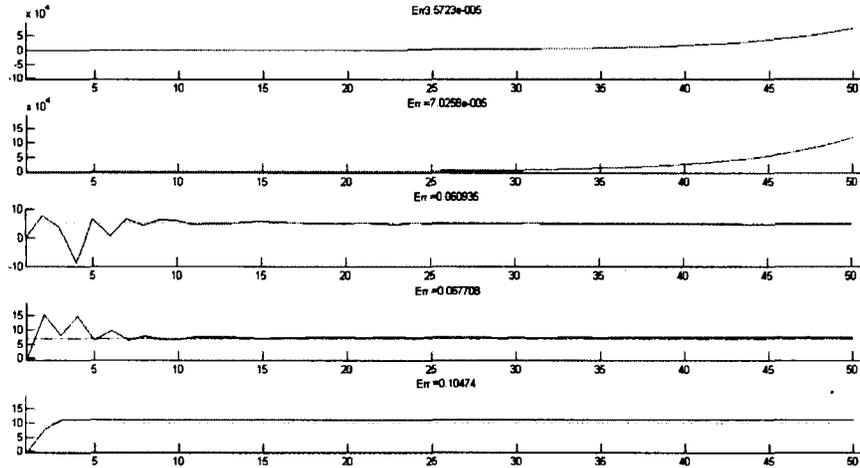


Figure 3.6 – Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits gaussiens où $\hat{Z}_{0|0} = 0$ et $\tilde{P}_{0|0} = 10^4 I_5$.

3.6 Conclusion

Dans ce chapitre nous avons proposé une méthode pour résoudre le problème de filtrage à l'aide de l'approche du filtre de Kalman dans le cas où les moyennes des bruits sont inconnues. Cette méthode nous a conduit vers un algorithme similaire au filtre de Kalman qui estime à la fois le signal et les moyennes inconnues.

CHAPITRE 4

Autres types de bruit

4.1 Introduction

Dans ce chapitre nous donnons des applications de l'algorithme (\tilde{K}) sur le problème $(S^{n,m})$ avec des bruits blancs pas nécessairement gaussiens. En effet, dans les équations du modèle, la distribution du bruit peut être quelconque. Ceci vient du fait que le filtre de Kalman reste valable pour des bruits blancs non gaussiens avec espérances et variances connues.

4.2 Transformation de vecteur aléatoire

Rappelons un résultat sur les matrices symétriques semi-définies positives [5].

Lemme 4.2.1. *Soit A une matrice symétrique réelle semi-définie positive alors il*

existe une unique matrice symétrique réelle semi-définie positive B tel que

$$A = B^2;$$

la matrice B est appelée une racine carrée de A et sera notée par $A^{\frac{1}{2}}$.

Démonstration. A est une matrice symétrique réelle alors elle est diagonalisable via une matrice de passage orthogonale P et une matrice diagonale D qui contient les valeurs propres de A :

$$A = PDP^T.$$

où $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ avec λ_i sont les valeurs propre de A pour tout $i \in \{1, 2, \dots, n\}$ Mais puisque on a supposé que A est positive alors les valeurs propres de A sont positives ou nulles, ceci nous laisse écrire :

$$A = PD^{\frac{1}{2}}D^{\frac{1}{2}}P^T = PD^{\frac{1}{2}}P^T PD^{\frac{1}{2}}P^T = B^2.$$

Où $B = PD^{\frac{1}{2}}P^T$ et $D^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})$ □

Proposition 4.2.1. *Supposons qu'on a un vecteur aléatoire U qui suit une certaine loi de moyenne μ et de matrice de variance-covariance Σ_0 définie positive, qu'on note par $U \sim \mathcal{L}(m, \Sigma_0)$, il est possible de faire des transformations linéaires sur U afin d'obtenir une certaine moyenne m et une certaine matrice de variance-covariance Σ (définie positive) données. Il suffit de définir un second vecteur aléatoire V tel que :*

$$V = m + W(U - \mu)$$

avec $W = \Sigma^{\frac{1}{2}}\Sigma_0^{-\frac{1}{2}}$ où $\Sigma^{\frac{1}{2}}$ désigne la racine carrée de Σ et $\Sigma_0^{-\frac{1}{2}}$ désigne l'inverse de $\Sigma_0^{\frac{1}{2}}$.

4.3 Type exponentiel

Définition 4.3.1. On dira que le vecteur aléatoire $\epsilon = (\epsilon_1(m_1), \epsilon_2(m_2), \dots, \epsilon_d(m_d))$ dans \mathbb{R}^d suit une loi exponentielle, et on le note par $\epsilon \sim \mathcal{E}(m)$ où $m = (m_1, m_2, \dots, m_d)$, si :

- pour tout i , $\epsilon_i(m_i)$ est une variable aléatoire exponentielle de moyenne m_i ;
- les ϵ_i sont deux à deux indépendantes.

Considérons un problème d'estimation du signal à l'aide du filtre de Kalman lorsque les bruits sur la trajectoire et l'observation ne sont pas centrés (à moyennes inconnues) et de type exponentiel c'est-à-dire :

$$(S_1^{n,m}) \quad \begin{cases} X_{k+1} = A_k X_k & + \epsilon_{x,k} \\ Y_{k+1} = H_{k+1} X_{k+1} & + \epsilon_{y,k+1} \end{cases}$$

avec $\epsilon_{x,k} = m_x + W_1(\epsilon - m)$ et $\epsilon_{y,k} = m_y + W_2(\epsilon' - m')$ où $\epsilon \sim \mathcal{E}(m)$ et $\epsilon' \sim \mathcal{E}(m')$, ce qui implique que :

$$\begin{cases} \mathbb{E}(\epsilon_{x,k}) = m_x & \text{et} & \text{Var}(\epsilon_{x,k}) = W_1 \text{Var}(\epsilon) W_1^T \\ \mathbb{E}(\epsilon_{y,k}) = m_y & \text{et} & \text{Var}(\epsilon_{y,k}) = W_2 \text{Var}(\epsilon') W_2^T; \end{cases} \quad (4.1)$$

$$\text{avec } \text{Var}(\epsilon) = \begin{bmatrix} m_1^2 & 0 & \dots & 0 \\ 0 & m_2^2 & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & m_n^2 \end{bmatrix} \quad \text{et} \quad \text{Var}(\epsilon') = \begin{bmatrix} m_1'^2 & 0 & \dots & 0 \\ 0 & m_2'^2 & \dots & \vdots \\ \vdots & \dots & \ddots & 0 \\ 0 & \dots & 0 & m_m'^2 \end{bmatrix}$$

Exemple

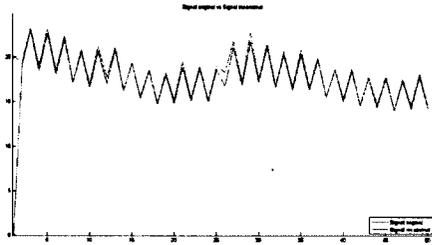
Considérons l'exemple simple suivant :

$$(E_1) \quad \begin{cases} X_{k+1} = (0.9 + 0.2 * (-1)^k) * X_k & + \epsilon_{x,k} \\ Y_{k+1} = 0.5 * X_{k+1} & + \epsilon_{y,k+1} \end{cases}$$

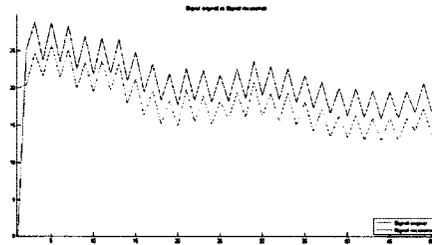
avec $\epsilon_{x,k} = 2 + 0.6(\epsilon - 1)$ et $\epsilon_{y,k} = 5 + 0.6(\epsilon' - 1)$ où $\epsilon \sim \mathcal{E}(1)$ et $\epsilon' \sim \mathcal{E}(1)$.

La Figure 4.1 montre l'application de l'algorithme (\tilde{K}) sur le système (E_1) pour

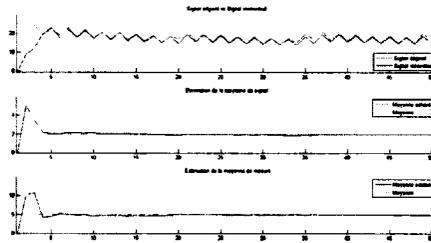
estimer le vecteur d'état, la moyenne de l'état et de la mesure. On remarque qu'après quelques étapes nous obtenons à la fois une bonne reconstruction du signal tout en obtenant de bonnes estimations des moyennes $m_x \cong 2$ et $m_y \cong 5$.



(a) Vraies moyennes : $m_x = 2$ et $m_y = 5$
(erreur = 0.0161)



(b) Moyennes quelconques : $m_x = 0$ et $m_y = 0$
(erreur = 0.2061)



(c) Algorithme (\tilde{K}) appliqué sur (E_1) avec $m_x = 2$, $m_y = 5$, $\hat{Z}_{0|0} = 0$, $\hat{P}_{0|0} = 10^4 I_3$, $m = m' = 1$ et $S_1 = S_2 = \sqrt{0.3}$. (erreur = 0.0404)

Figure 4.1 – Algorithmes (K) et (\tilde{K}) appliqués sur (E_1) avec des bruits de type exponentiel où $\hat{Z}_{0|0} = 0$ et $\hat{P}_{0|0} = 10^4 * I_3$.

4.4 Type khi-deux

Définition 4.4.1. On dira que le vecteur aléatoire $\epsilon = (T_1, T_2, \dots, T_d)$ dans \mathbb{R}^d suit une loi de khi-deux, et on le note par $\epsilon \sim \chi^2(m)$ où $m = (m_1, m_2, \dots, m_d)$, si :

- pour tout i , $T_i \sim \chi_{m_i}^2$ est une variable aléatoire khi-deux de paramètre m_i ;
- les χ_i^2 sont deux à deux indépendants.

Considérons un problème d'estimation du signal à l'aide du filtre de Kalman lorsque les bruits sur la trajectoire et l'observation ne sont pas centrés et de type khi-deux c'est-à-dire :

$$(S_2^{n,m}) \quad \begin{cases} X_{k+1} = A_k X_k + \epsilon_{x,k} \\ Y_{k+1} = H_{k+1} X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

avec $\epsilon_{x,k} = m_x + W_1(\epsilon - m)$ et $\epsilon_{y,k} = m_y + W_2(\epsilon' - m')$ où $\epsilon \sim \chi^2(m)$ et $\epsilon' \sim \chi^2(m')$, ce qui implique que :

$$\begin{cases} \mathbb{E}(\epsilon_{x,k}) = m_x & \text{et} & \text{Var}(\epsilon_{x,k}) = W_1 \text{Var}(\epsilon) W_1^T \\ \mathbb{E}(\epsilon_{y,k}) = m_y & \text{et} & \text{Var}(\epsilon_{y,k}) = W_2 \text{Var}(\epsilon') W_2^T; \end{cases} \quad (4.2)$$

$$\text{avec } \text{Var}(\epsilon) = \begin{bmatrix} 2m_1 & 0 & \dots & 0 \\ 0 & 2m_2 & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & 2m_n \end{bmatrix} \quad \text{et} \quad \text{Var}(\epsilon') = \begin{bmatrix} 2m'_1 & 0 & \dots & 0 \\ 0 & 2m'_2 & \dots & \vdots \\ \vdots & \dots & \ddots & 0 \\ 0 & \dots & 0 & 2m'_m \end{bmatrix}$$

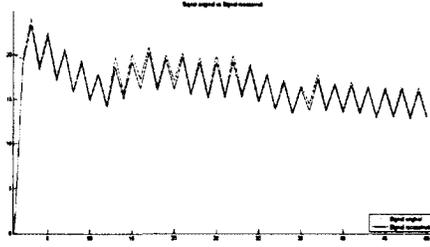
Exemple

Considérons l'exemple simple suivant :

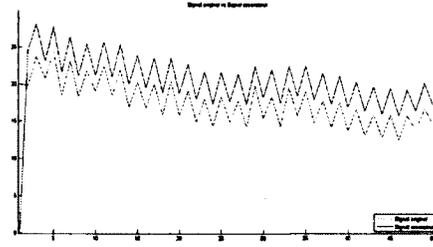
$$(E_2) \quad \begin{cases} X_{k+1} = (0.9 + 0.2 * (-1)^k) * X_k + \epsilon_{x,k} \\ Y_{k+1} = 0.5 * X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

$\epsilon_{x,k} = 2 + \frac{0.6}{\sqrt{2}}(\epsilon - 1)$ et $\epsilon_{y,k} = 5 + \frac{0.6}{\sqrt{2}}(\epsilon' - 1)$ où $\epsilon \sim \chi^2(1)$ et $\epsilon' \sim \chi^2(1)$.

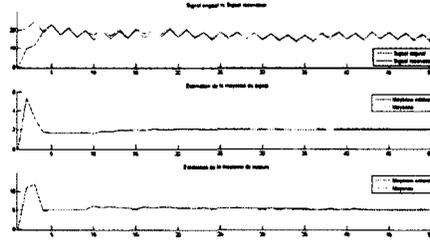
La Figure 4.2 montre l'application de l'algorithme (\tilde{K}) sur le système (E_2) pour estimer le vecteur d'état, la moyenne de l'état et de la mesure. On remarque qu'après quelques étapes nous obtenons à la fois une bonne reconstruction du signal tout en obtenant de bonnes estimations des moyennes $m_x \cong 2$ et $m_y \cong 5$.



(a) Vraies moyennes : $m_x = 2$ et $m_y = 5$
(erreur = 0.0147)



(b) Moyennes quelconques : $m_x = 0$ et $m_y = 0$
(erreur = 0.2141)



(c) Algorithme (\tilde{K}) appliqué sur (E_2) avec $m_x = 2$, $m_y = 5$, $\tilde{Z}_{0|0} = 0$, $\tilde{P}_{0|0} = 10^4 I_3$, $m = m' = 1$ et $W_1 = W_2 = 0.15$. (erreur = 0.04)

Figure 4.2 – Algorithmes (K) et (\tilde{K}) appliqués sur (E_2) avec des bruits de type khi-deux où $\hat{Z}_{0|0} = 0$ et $\tilde{P}_{0|0} = 10^4 * I_3$.

4.5 Type ricien

Définition 4.5.1. *Étant donnée deux variables aléatoires indépendantes X et Y tel que $(X, Y) \sim \mathcal{N}(\mu, \sigma^2 I_2)$, alors*

1. $X^2 + Y^2 \sim \sigma^2 \chi_2^2(\frac{\mu\mu'}{\sigma^2})$ où $\chi_2^2(\sigma^2)$ désigne une loi de khi-deux avec deux degrés de liberté et paramètre de décentralité σ .
2. $R = \sqrt{X^2 + Y^2}$ suit une loi dite de Rice avec paramètre $(\nu = \|\mu\|, \sigma)$, notée $Rice(\nu, \sigma)$.

On peut démontrer que la densité de la distribution de la loi de Rice est [12, 16] :

$$f(x|\nu, \sigma) = \begin{cases} \frac{x}{\sigma^2} \exp\left(\frac{-(x^2 + \nu^2)}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma}\right) & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

où $I_0(\cdot)$ est la fonction de Bessel modifiée de première espèce et d'ordre 0 définie par :

$$I_0(x) = \sum_{k=0}^{+\infty} \frac{\left(\frac{x^2}{4}\right)^k}{(k!)^2}.$$

L'espérance et la variance de la distribution ricienne sont données par :

$$\begin{cases} \mathbb{E}(R) &= \sigma \sqrt{\frac{\pi}{2}} L_{\frac{1}{2}}\left(\frac{-\nu^2}{2\sigma^2}\right) \\ \mathbb{E}(R^2) &= \nu^2 + 2\sigma^2 \\ \text{Var}(R) &= \mathbb{E}(R^2) - (\mathbb{E}(R))^2 = \nu^2 + 2\sigma^2 - \frac{\pi\sigma^2}{2} L_{\frac{1}{2}}^2\left(\frac{-\nu^2}{2\sigma^2}\right) \end{cases}$$

où $L_{\frac{1}{2}}(\cdot)$ représente un polynôme de Laguerre défini par :

$$L_{\frac{1}{2}}(x) = e^{\frac{x}{2}} \left[(1-x)I_0\left(\frac{-x}{2}\right) - xI_1\left(\frac{-x}{2}\right) \right]$$

avec

$$I_1(x) = \frac{x}{2} \sum_{k=0}^{+\infty} \frac{\left(\frac{x^2}{4}\right)^k}{k!(k+1)!}.$$

La Figure 4.3 illustre la densité ricienne pour $\sigma = 1$ et différentes valeurs de ν

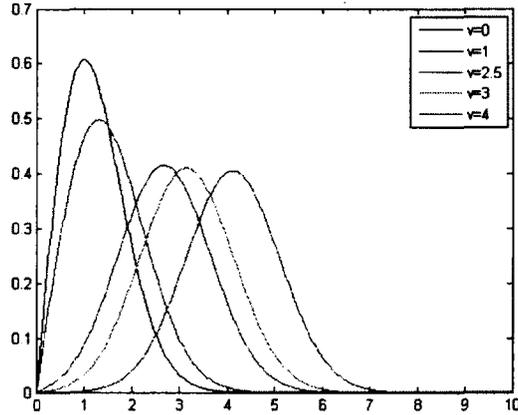


Figure 4.3 – Densité de probabilité de la loi de Rice pour $\sigma = 1$ et différentes valeurs de ν

Définition 4.5.2. On dit que le vecteur aléatoire $\epsilon = (Rice_1(\nu_1, \sigma_1), Rice_2(\nu_2, \sigma_2), \dots, Rice_d(\nu_d, \sigma_d))$ dans \mathbb{R}^d suit une loi de rice, et on le note par $\epsilon \sim Rice(\nu, \sigma)$ où $\nu = (\nu_1, \nu_2, \dots, \nu_d)$ et $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_d)$, si :

- pour tout i , $Rice_i(\nu_i, \sigma_i)$ est une variable aléatoire ricienne de paramètres ν_i et σ_i ;
- les $Rice(\nu_i, \sigma_i)$ sont deux à deux indépendants.

Considérons un problème d'estimation du signal à l'aide du filtre de Kalman lorsque les bruits sur la trajectoire et l'observation ne sont pas centrés et de type riciens c'est-à-dire :

$$(S_3^{n,m}) \quad \begin{cases} X_{k+1} = A_k X_k & + \epsilon_{x,k} \\ Y_{k+1} = H_{k+1} X_{k+1} & + \epsilon_{y,k+1} \end{cases}$$

avec $\epsilon_{x,k} = m_x + W_1(Rice(\nu, \sigma) - \mathbb{E}(Rice(\nu, \sigma)))$ et $\epsilon_{y,k} = m_y + W_2(Rice(\nu', \sigma') - \mathbb{E}(Rice(\nu', \sigma')))$, ce qui implique que :

$$\begin{cases} \mathbb{E}(\epsilon_{x,k}) = m_x & \text{et} & \text{Var}(\epsilon_{x,k}) = S_1 \text{Var}(Rice(\nu, \sigma)) S_1^T \\ \mathbb{E}(\epsilon_{y,k}) = m_y & \text{et} & \text{Var}(\epsilon_{y,k}) = S_2 \text{Var}(Rice(\nu', \sigma')) S_2^T \end{cases} \quad (4.3)$$

Exemple

Considérons l'exemple simple suivant :

$$(E_3) \quad \begin{cases} X_{k+1} &= (0.9 + 0.2 * (-1)^k) * X_k + \epsilon_{x,k} \\ Y_{k+1} &= 0.5 * X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

avec

$$\begin{cases} \epsilon_{x,k} &= m_x + 3.68(Rice(0.3, 0.2) - \mathbb{E}(Rice(0.3, 0.2))) \\ \text{Var}(Rice(0.3, 0.2)) &= 0.0265 \\ \text{Var}(\epsilon_{x,k}) &= 3.68 * 0.0265 * 3.68 = 0.36 \end{cases}$$

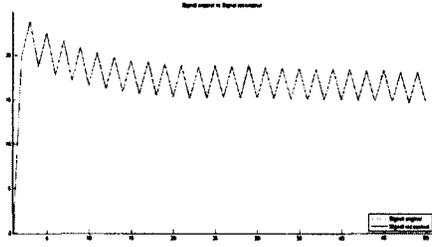
et

$$\begin{cases} \epsilon_{x,k} &= m_y + 3.68(Rice(0.3, 0.2) - \mathbb{E}(Rice(0.3, 0.2))) \\ \text{Var}(Rice(0.3, 0.2)) &= 0.0265 \\ \text{Var}(\epsilon_{x,k}) &= 3.68 * 0.0265 * 3.68 = 0.36 \end{cases}$$

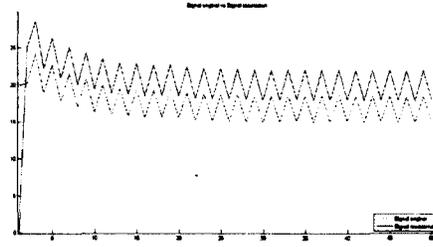
La Figure 4.4 montre l'application de l'algorithme (\tilde{K}) sur le système (E_3) pour estimer le vecteur d'état, la moyenne de l'état et de la mesure. On remarque qu'après quelques étapes nous obtenons à la fois une bonne reconstruction du signal tout en obtenant de bonnes estimations des moyennes $m_x \cong 2$ et $m_y \cong 5$.

Remarque 4.5.1. *Le tableau suivant résume l'erreur sur la reconstruction du signal et l'estimation des moyennes inconnues pour les exemples (E_1), (E_2) et (E_3).*

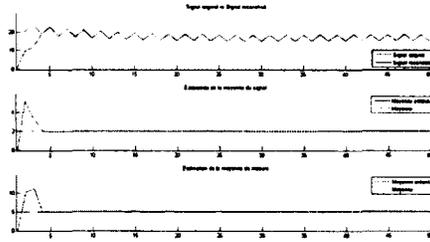
Le Tableau 4.1 illustre bien qu'il y a une bonne reconstruction du signal pour les différents types de bruits. Cela nous laisse dire que l'algorithme (\tilde{K}) ne tient pas compte de la distribution du bruit et converge si les bruits utilisés dans le système dynamique sont de type additifs et à variances finies et connues.



(a) vraies moyennes : $m_x = 2$ et $m_y = 5$
(erreur = 0.0037)



(b) moyennes quelconques : $m_x = 0$ et $m_y = 0$
(erreur = 0.2081)



(c) Algorithme (\tilde{K}) appliqué sur (E_3) avec $m_x = 2$, $m_y = 5$, $\hat{Z}_{0|0} = 0$ et $\tilde{P}_{0|0} = 10^4 I_3$, $\nu = 0.3$, $\sigma = 0.2$, $\nu' = 0.3$, $\sigma' = 0.2$ et $W_1 = W_2 = 0.3$.
(erreur = 0.0105)

Figure 4.4 – Algorithmes (K) et (\tilde{K}) appliqués sur (E_3) avec des bruits de type riciens où $\hat{Z}_{0|0} = 0$ et $\tilde{P}_{0|0} = 10^4 I_3$.

4.6 Exemples multidimensionnels

Dans cette section nous donnons des exemples multidimensionnelles de l'algorithme (\tilde{K}) avec des bruits de type exponentiel, khie-deux et ricien.

Premier exemple. Considérons le système dynamique dans \mathbb{R}^2 avec des bruits de

	$Var(\epsilon_{x,k})$	0.36
	$Var(\epsilon_{y,k})$	0.36
Type exponentiel	$\overline{err}_x^{(50)}$	0.0302
	$\overline{err}_y^{(50)}$	0.0202
	$Moy^{(50)}$	0.0404
Type khi-deux	$\overline{err}_x^{(50)}$	0.0295
	$\overline{err}_y^{(50)}$	0.0265
	$Moy^{(50)}$	0.04
Type ricien	$\overline{err}_x^{(50)}$	0.0245
	$\overline{err}_y^{(50)}$	0.028
	$Moy^{(50)}$	0.0285

Tableau 4.1 – Erreur sur l'estimation des moyennes et la reconstruction du signal pour les exemples (E₁), (E₂) et (E₃).

type exponentiel

$$(S_1^{2,1}) \quad \begin{cases} X_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} X_k + \epsilon_{x,k} \\ Y_{k+1} = [1 \ 2] X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

avec

- $\epsilon_{x,k}$ est un vecteur aléatoire de type exponentiel dans \mathbb{R}^2 de moyenne $[5, 7]'$ et de matrice de variance-covariance $\Sigma = \begin{bmatrix} 0.36 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$, en effet

$$\epsilon_{x,k} = m_x + W_1(\epsilon - m)$$

avec

$$\begin{cases} m_x = \begin{bmatrix} 5 \\ 7 \end{bmatrix} \\ \epsilon \sim \mathcal{E}(5, 6) \text{ avec } \Sigma_0 = \begin{bmatrix} 25 & 0 \\ 0 & 36 \end{bmatrix} \\ W_1 = \Sigma^{\frac{1}{2}} \Sigma_0^{-\frac{1}{2}} = \begin{bmatrix} 0.1184 & 0.0162 \\ 0.0194 & 0.0728 \end{bmatrix} \end{cases}$$

- $\epsilon_{y,k}$ est une variable aléatoire de type exponentiel de moyenne 10 et de variance $\Sigma' = 0.3$, en effet

$$\epsilon_{y,k} = m_y + W_2(\epsilon' - m')$$

avec

$$\begin{cases} m_y = 10 \\ \epsilon' \sim \mathcal{E}(1) \text{ et } \Sigma'_0 = 1 \\ W_2 = \Sigma'^{\frac{1}{2}} \Sigma'_0^{-\frac{1}{2}} = \sqrt{0.3} \end{cases}$$

La Figure 4.5 montre l'application de l'algorithme (\tilde{K}) sur le système $(S_1^{2,1})$ pour estimer le vecteur d'état, la moyenne de l'état (qui sont dans \mathbb{R}^2) et la moyenne de la mesure. On remarque qu'après quelques étapes nous obtenons à la fois une bonne reconstruction du signal tout en obtenant de bonnes estimations des moyennes $m_x \cong \begin{bmatrix} 5 \\ 7 \end{bmatrix}$ et $m_y \cong 10$.

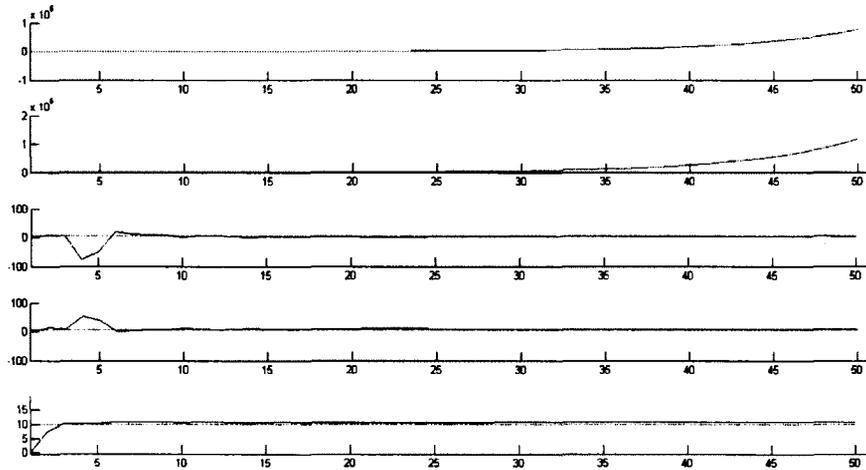


Figure 4.5 – Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits de type exponentiel où $\hat{Z}_{0|0} = 0$ et $\hat{P}_{0|0} = 10^4 I_5$.

Deuxième exemple. Considérons le système dynamique dans \mathbb{R}^2

$$(S_2^{2,1}) \quad \begin{cases} X_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} X_k + \epsilon_{x,k} \\ Y_{k+1} = [1 \ 2] X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

avec

- $\epsilon_{x,k}$ est un vecteur aléatoire de type khie-deux dans \mathbb{R}^2 de moyenne $[5, 7]'$ et de matrice de variance-covariance $\Sigma = \begin{bmatrix} 0.36 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$, en effet

$$\epsilon_{x,k} = m_x + W_1(\epsilon - m)$$

avec

$$\begin{cases} m_x = \begin{bmatrix} 5 \\ 7 \end{bmatrix} \\ \epsilon \sim \chi^2(5, 6) \text{ avec } \Sigma_0 = \begin{bmatrix} 10 & \\ 0 & 12 \end{bmatrix} \\ W_1 = \Sigma^{\frac{1}{2}} \Sigma_0^{-\frac{1}{2}} = \begin{bmatrix} 0.4187 & 0.0687 \\ 0.0687 & 0.3087 \end{bmatrix} \end{cases}$$

- $\epsilon_{y,k}$ est une variable aléatoire de type khie-deux de moyenne 10 et de variance $\Sigma' = 0.3$, en effet

$$\epsilon_{y,k} = m_y + W_2(\epsilon' - m')$$

avec

$$\begin{cases} m_y = 10 \\ \epsilon' \sim \chi^2(1) \text{ avec } \Sigma'_0 = 2 \\ W_2 = \Sigma'^{\frac{1}{2}} \Sigma'_0^{-\frac{1}{2}} = \sqrt{0.15} \end{cases}$$

La Figure 4.6 montre l'application de l'algorithme (\tilde{K}) sur le système $(S_2^{2,1})$ pour estimer le vecteur d'état (qui est dans \mathbb{R}^2), la moyenne de l'état et de la mesure. On remarque qu'après quelques étapes nous obtenons à la fois une bonne reconstruction du signal tout en obtenant de bonnes estimations des moyennes $m_x \cong \begin{bmatrix} 5 \\ 7 \end{bmatrix}$ et

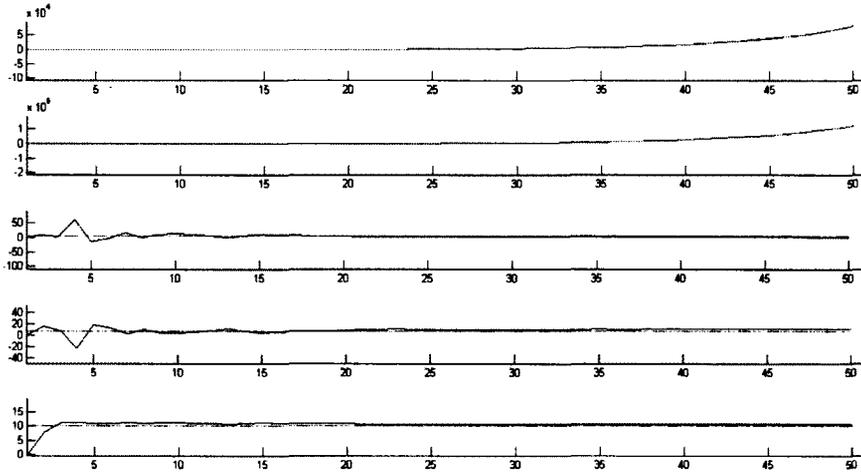


Figure 4.6 – Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits de type khie-deux où $\hat{Z}_{0|0} = 0$ et $\tilde{P}_{0|0} = 10^4 I_5$.

$m_y \cong 10$.

Troisième exemple. Considérons le système dynamique dans \mathbb{R}^2

$$(S_2^{2,1}) \quad \begin{cases} X_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} X_k + \epsilon_{x,k} \\ Y_{k+1} = [1 \ 2] X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

avec

- $\epsilon_{x,k}$ est un vecteur aléatoire de type ricien dans \mathbb{R}^2 de moyenne $[5, 7]'$ et de matrice de variance-covariance $\Sigma = \begin{bmatrix} 0.36 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$, en effet

$$\epsilon_{x,k} = m_x + W_1(\epsilon - m)$$

avec

$$\begin{cases} m_x = \begin{bmatrix} 5 \\ 7 \end{bmatrix} \\ \epsilon \sim \text{Rice}(\nu, \sigma) \text{ où } \nu = \begin{bmatrix} 0.3 \\ 0.3 \end{bmatrix}, \sigma = \begin{bmatrix} 0.23 \\ 0.155 \end{bmatrix} \text{ et } \Sigma_0 = \begin{bmatrix} 0.0362 & 0 \\ 0 & 0.0198 \end{bmatrix} \\ W_1 = \Sigma_0^{-\frac{1}{2}} = \begin{bmatrix} 3.4414 & 0.2185 \\ 0.5651 & 0.9810 \end{bmatrix} \end{cases}$$

- $\epsilon_{y,k}$ est une variable aléatoire de type ricien de moyenne 10 et de variance $\Sigma' = 0.3$, en effet

$$\epsilon_{y,k} = m_y + W_2(\epsilon' - m')$$

avec

$$\begin{cases} m_y = 10 \\ \epsilon' \sim \text{Rice}(\nu', \sigma') \text{ où } \nu' = 0.3, \sigma' = 0.2 \text{ et } \Sigma'_0 = 0.0294 \\ W_2 = \Sigma'_0^{-\frac{1}{2}} = 3.8925 \end{cases}$$

La Figure 4.7 montre l'application de l'algorithme (\tilde{K}) sur le système $(S_2^{2,1})$ pour estimer le vecteur d'état (qui est dans \mathbb{R}^2), la moyenne de l'état et de la mesure. On remarque qu'après quelques étapes nous obtenons à la fois une bonne reconstruction du signal tout en obtenant de bonnes estimations des moyennes $m_x \cong \begin{bmatrix} 5 \\ 7 \end{bmatrix}$ et $m_y \cong 10$.

Remarque 4.6.1. *L'algorithme (\tilde{K}) est capable de reconstruire et estimer les moyennes inconnues dans le cas où les bruits dans l'équation d'évolution et l'équation de mesure n'ont pas forcément la même distribution. En effet, considérons le système dynamique dans \mathbb{R}^2 où les bruits dans l'équation d'évolution sont de type exponentiel or que le bruit sur l'équation de mesure est de type khie-deux.*

$$(S_4^{2,1}) \quad \begin{cases} X_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} X_k + \epsilon_{x,k} \\ Y_{k+1} = [1 \ 2] X_{k+1} + \epsilon_{y,k+1} \end{cases}$$

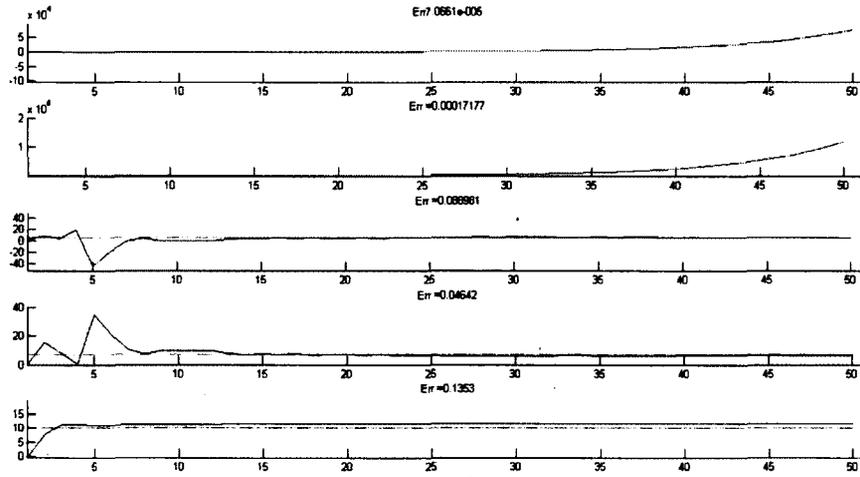


Figure 4.7 – Algorithme (\tilde{K}) appliqué sur $(S^{2,1})$ avec des bruits de type riciens où $\hat{Z}_{0|0} = 0$ et $\hat{P}_{0|0} = 10^4 I_5$.

avec

- $\epsilon_{x,k}$ est un vecteur aléatoire de type exponentiel dans \mathbb{R}^2 de moyenne $[5, 7]'$ et de matrice de variance-covariance $\Sigma = \begin{bmatrix} 0.36 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$, en effet

$$\epsilon_{x,k} = m_x + W_1(\epsilon - m)$$

avec

$$\left\{ \begin{array}{l} m_x = \begin{bmatrix} 5 \\ 7 \end{bmatrix} \\ \epsilon \sim \mathcal{E}(5, 6) \text{ avec } \Sigma_0 = \begin{bmatrix} 25 & 0 \\ 0 & 36 \end{bmatrix} \\ W_1 = \Sigma^{\frac{1}{2}} \Sigma_0^{-\frac{1}{2}} = \begin{bmatrix} 0.1184 & 0.0162 \\ 0.0194 & 0.0728 \end{bmatrix} \end{array} \right.$$

- $\epsilon_{y,k}$ est une variable aléatoire de type khi-deux de moyenne 10 et de variance 0.3,

en effet

$$\epsilon_{y,k} = m_y + W_2(\epsilon' - m')$$

avec

$$\begin{cases} m_y = 10 \\ \epsilon' \sim \chi^2(1) \text{ avec } \Sigma'_0 = 2 \\ W_2 = \Sigma'^{\frac{1}{2}} \Sigma'_0^{-\frac{1}{2}} = \sqrt{0.15} \end{cases}$$

La Figure 4.8 montre l'application de l'algorithme (\tilde{K}) sur le système ($S_4^{2,1}$) pour estimer le vecteur d'état, la moyenne de l'état (qui sont dans \mathbb{R}^2) et la moyenne de la mesure.

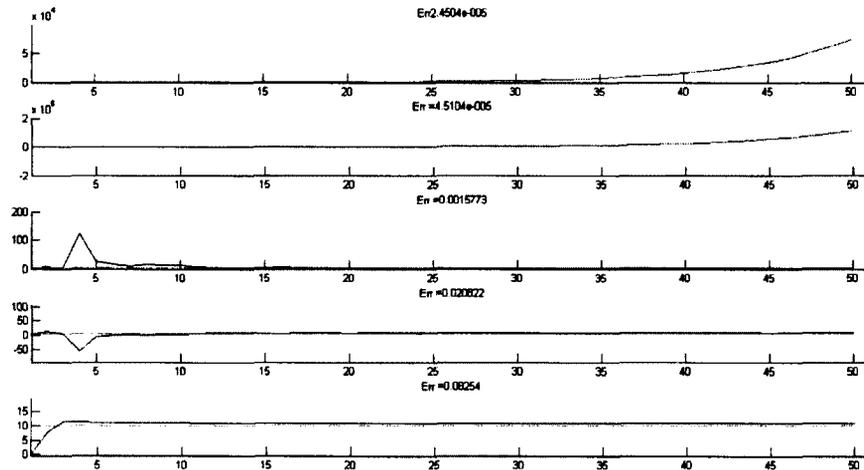


Figure 4.8 – Algorithme (\tilde{K}) appliqué sur ($S_4^{2,1}$) où $\hat{Z}_{0|0} = 0$ et $\tilde{P}_{0|0} = 10^4 I_5$.

4.7 Conclusion

Dans ce chapitre nous avons appliqué l'algorithme du filtre de Kalman modifié (\tilde{K}) sur des modèles statistiques linéaires avec différents type de bruit (gaussien, exponentiel, khi-deux et ricien) afin de le comparer avec le filtre de Kalman standard (\tilde{K}).

CONCLUSION

Le filtre de Kalman est une méthode récursive visant à estimer des paramètres d'un système évoluant dans le temps à partir de mesures bruitées. La force de ce filtre est sa capacité de prédiction des paramètres et de rectification des erreurs, non seulement des mesures, mais aussi du modèle lui-même. En effet, pour appliquer un filtre de Kalman, il faut modéliser de manière linéaire le système pour lequel on veut estimer les paramètres. Dans une méthode d'estimation classique (par exemple, la méthode des moindres carrés), une simple erreur dans la modélisation du système entraîne forcément une erreur au niveau de l'estimation. La force du filtre de Kalman est d'intégrer un terme d'imprécision sur le modèle lui-même, ce qui lui permet de donner des estimations correctes malgré les erreurs de modélisation (à condition que ces erreurs restent raisonnables). Un autre point fort du filtre de Kalman est sa capacité à déterminer l'erreur moyenne de son estimation. En effet, le filtre de Kalman fournit un vecteur contenant les paramètres estimés, mais aussi une matrice de variance-covariance de l'erreur. Cette matrice nous renseigne donc sur la précision de l'estimation, ce qui peut être utile dans de nombreuses applications.

Le filtre est même capable de reconstruire le signal et d'estimer les moyennes inconnues des bruits dans le cas où les moyennes des bruits sont inconnues. Une des limites du filtre de Kalman est le cas où la modélisation est trop approximative, le filtre n'est

pas assez performant et l'erreur des estimations ne convergera pas assez rapidement. Une autre limite importante d'une telle méthode est que le filtre de Kalman permet de prendre en compte uniquement un modèle de bruit additif, cette restriction limite l'utilisation du filtre de Kalman.

Dans ce mémoire nous avons résolu le problème de reconstruction du signal dans le cas où les bruits sont de type additif et à moyennes non nulles et inconnues. Nous espérons que ce travail sera utile en imagerie médicale et plus précisément en IRM (imagerie par résonance magnétique) pour le débruitage d'un signal corrompu par un bruit ricien non additif. Des travaux ont déjà été faits dans le cas où le bruit est gaussien additif [6], mais le défi est d'adapter le filtre de Kalman au cas d'un bruit qui n'est pas additif.

Bibliographie

- [1] D. Alazard. Introduction au filtre de Kalman
[http ://personnel.supaero.fr/alazard-daniel/Pdf/cours-Kalman.pdf](http://personnel.supaero.fr/alazard-daniel/Pdf/cours-Kalman.pdf).
- [2] A. Benavoli, L. Chisci et A. Farina. Fibonacci sequence, golden section, Kalman filter and optimal control. *Signal Processing*, 89(2009), Pages 1483 - 1488.
- [3] C. A. Bozzo. *Le filtrage optimal*, Technique et Documentation, Paris, 1983.
- [4] R.G. Brown et P. Y. C. Hwang. *Introduction to Random Signals and Applied Kalman Filtering*, John Wiley and Sons, New york, 1992.
- [5] C. K. Chui et G. Chen. *Kalman Filtering with Real-Time Applications*, Springer, Berlin, 2009.
- [6] R. Deriche, J. Calder, M. Descoteaux. Optimal real-time Q-ball imaging using regularized kalman filtering with Incremental Orientation Sets. *Medical Image Analysis*, 13(2009), Pages 564-579
- [7] M. S.Grewal, A. P.Andrews. *Kalman Filtering : Theory and Practice using Matlab*, Wiley IEEE-Press, London, 2008

- [8] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the AMSE, Part D, Journal of Basic Engineering*, 82(1960), Pages 35-45.
- [9] F. LeGland. Introduction au filtrage en temps discret, 2006
<http://www.irisa.fr/aspi/legland/rennes-1/cours.pdf>.
- [10] D. G. Luenberger. *Optimization by Vector Spaces Methods*, John Wiley and Sons, New york, 1969.
- [11] P. S. Maybeck. *Stochastic Models, Estimation and Control*. Mathematics in Science and Engineering, San Francisco, 1979
- [12] M. Descoteaux. *High Angular Resolution Diffusion MRI : From Local Estimation to Segmentation and Tractography*. PhD Thesis, Universite de Nice - Sophia Antipolis, 2008
- [13] A. Papoulis et S. U. Pillai. *Probability, Random Variables and Stochastic Processes*, Mc Graw Hill, New York, 2002
- [14] J.C. Radix. *Filtrage et lissage statistiques optimaux linéaires*, Cepadus-éditions, Toulouse, 1984.
- [15] I. B. Rhodes. A tutorial introduction to estimation and filtering. *IEEE Transaction on Automatic Control*, 16(1971), Pages 688 - 706.
- [16] S. O. Rice. *Mathematical Analysis of Random Noise*, Bell System Technical Journal, 24(1945), pages 46-156

- [17] S. M. Ross. *Initiation aux probabilités*, Presses Polytechniques et Universitaires Romandes, 2003.

- [18] D. Simon. *Optimal State Estimation : Kalman, H-Infinity and Nonlinear Approaches*, John Wiley and Sons, New york, 2006.

- [19] G. Welch et G. Bishop. *An Introduction to the Kalman Filter*. Department of computer science, University of North Carolina at Chapel Hill, 2006.