

Quelques considérations sur le filtre de Kalman discret

par

Hichem BELATTAR

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, mars 2017

Le 8 Mars 2017

*le jury a accepté le mémoire de Monsieur Hichem Belattar
dans sa version finale.*

Membres du jury

Professeur François Dubeau
Directeur de recherche
Département de mathématiques

Professeur Bernard Colin
Évaluateur interne
Département de mathématiques

Professeur Éric Marchand
Président-rapporteur
Département de mathématiques

SOMMAIRE

Le filtre de Kalman est un outil mathématique qui a pour principe d'estimer l'état d'un système dynamique à partir d'observations partielles et généralement bruitées. Il prend en entrée une suite d'observations telle que chaque observation soit reliée à un état inconnu précis de façon linéaire. Le but sera d'estimer ces états de manière optimale et récursive. Nous allons étudier dans ce mémoire la théorie du filtre de Kalman avec plusieurs types de bruits (gaussien, exponentiels, khi-deux). Nous développerons un nouveau type d'initialisation basé sur la théorie des moindres carrés. Enfin nous tenterons d'adapter le filtre de Kalman au cas où les moyennes des bruits du système dynamique sont inconnues.

Mots-clés : Kalman, filtre, moindres carrés, initialisation, modèle, bruit de moyenne inconnue

REMERCIEMENTS

La réalisation de ce mémoire a été possible grâce au concours de plusieurs personnes à qui je voudrais témoigner toute ma reconnaissance.

Je tiens d'abord à remercier mon directeur de recherche le professeur François Dubeau, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter ma réflexion.

Je désire aussi à remercier les professeurs Éric Marchand et Bernard Colin d'avoir accepté d'évaluer mon travail.

Je tiens aussi remercier le professeur Maxime Descoteaux qui m'a fait découvrir l'univers de l'imagerie médicale.

Je voudrais exprimer ma reconnaissance envers les amis et collègues pour leurs encouragements. Enfin, j'aimerais remercier mes parents Mourad et Radhia et ma soeur Katia qui m'ont apporté leur support moral et intellectuel tout au long de ma démarche.

Hichem Belattar

Sherbrooke, mars 2017

TABLE DES MATIÈRES

SOMMAIRE	iii
REMERCIEMENTS	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	viii
LISTE DES FIGURES	ix
INTRODUCTION	1
CHAPITRE 1 — Préliminaires mathématiques	3
1.1 Introduction	3
1.2 Matrice et déterminant	3
1.3 Estimation par la méthode des moindres carrés	6
CHAPITRE 2 — Filtre de Kalman	11

2.1	Introduction	11
2.2	Système dynamique et bruit blanc	12
2.3	Le modèle	12
2.4	Caractérisation orthogonale de l'estimateur à variance minimale	14
2.5	Suite d'innovations	18
2.6	Estimateur à variance minimale	22
2.7	Équations du filtre de Kalman	23
2.7.1	Initialisation	24
2.7.2	Prédiction-Correction	25
2.8	Exemple	32
2.8.1	Calcul des erreurs de reconstruction	33
2.9	Autres types de bruits	35
2.9.1	Transformation de vecteur aléatoire	36
2.9.2	Bruit suivant une loi exponentielle	36
2.9.3	Bruit suivant une loi Khi-deux	37
2.10	Conclusion	39
CHAPITRE 3 — Initialisation par la méthode des moindres carrés		40
3.1	Introduction	40
3.2	Initialisation par la méthode des moindres carrés	40
3.3	Exemple	42

3.4	Autres types de bruits	42
3.4.1	Bruit suivant une loi exponentielle	43
3.4.2	Bruit suivant une loi khi-deux	44
3.5	Conclusion	47
CHAPITRE 4 — Filtre de Kalman avec bruit de moyenne inconnue		48
4.1	Introduction	48
4.2	Modèle avec moyennes connues	48
4.3	Modèle avec moyennes inconnues	51
4.4	Autres types de bruits	58
4.4.1	Bruit suivant une loi exponentielle	59
4.4.2	Bruit suivant une loi khi-deux	59
4.5	Conclusion	65
CONCLUSION		66
BIBLIOGRAPHIE		67

LISTE DES TABLEAUX

2.1	Erreur quadratique moyenne sur $N = 30$ étapes et différentes valeurs de $P_{0,0}$	35
2.2	Décroissance de l'erreur de reconstruction du signal avec $\hat{x}_{0 0} = 0$, $P_{0,0} = 10^2$.	35

LISTE DES FIGURES

2.1	Reconstruction du signal avec $\hat{x}_{0 0} = 0$ et différentes valeurs de $P_{0,0}$	34
2.2	Filtre de Kalman avec des bruits suivant une loi exponentielle où $\hat{x}_{0 0} = 0$ et $P_{0,0} = 10^2$ ($Err^{(30)} = 0.1338$)	38
2.3	Filtre de Kalman avec des bruits suivant une loi de khi-deux où $\hat{x}_{0 0} = 0$ et $P_{0,0} = 10^2$ ($Err^{(30)} = 0.1294$)	39
3.1	Comparaison entre l'initialisation classique($\hat{x}_{0 0} = [0 \ 0]^T$, $P_{0,0} = 10^2 I_2$) et l'initialisation avec la méthode des moindres carrés	43
3.2	Erreur de reconstruction du signal avec initialisation classique et initiali- sation par la méthode des moindres carrés dans un cas multidimensionnel.	44
3.3	Reconstruction du signal avec l'initialisation classique ($\hat{x}_{0 0} = [0 \ 0]^T$, $P_{0,0} = 10^2 I_2$) et l'initialisation avec la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi exponentielle.	45
3.4	Erreur de reconstruction du signal avec initialisation classique et initiali- sation par la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi exponentielle.	45

3.5	Reconstruction du signal avec l'initialisation classique ($\hat{x}_{0 0} = [0 \ 0]^T$, $P_{0,0} = 10^2 I_2$) et l'initialisation avec la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi khi-deux.	46
3.6	Erreur de reconstruction du signal avec initialisation classique et initialisation par la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi khi-deux.	47
4.1	Filtre de Kalman avec moyennes des bruits non-nulles et connues ($\hat{x}_{0 0} = 0$ et $P_{0,0} = 10^2$)	51
4.2	Comparaison entre l'initialisation classique ($\hat{x}_{0 0} = 0$ et $P_{0,0} = 10^2$) et l'initialisation par la méthode des moindres carrés.	52
4.3	Erreur de reconstruction du signal avec initialisation classique et initialisation par la méthode des moindres carrés.	52
4.4	Reconstruction du signal et des moyennes des bruits avec initialisation classique($\hat{x}_{0 0} = [-2 \ -2 \ 0]^T$ et $P_{0,0} = 10^2 I_3$) et initialisation par la méthode des moindres carrés.	55
4.5	Erreur de reconstruction du signal et des moyennes des bruits avec initialisation classique et initialisation par la méthode des moindres carrés.	56
4.6	Reconstruction du signal et des moyennes des bruits avec initialisation classique($\hat{x}_{0 0} = [0 \ 0 \ 0 \ 0 \ 0]^T$ et $P_{0,0} = 10^2 I_5$) et initialisation par la méthode des moindres carrés dans un cas multidimensionnel.	57
4.7	Erreur de reconstruction du signal et des moyennes des bruits avec initialisation classique et initialisation par la méthode des moindres carrés dans un cas multidimensionnel.	58

4.8 Filtre de Kalman avec moyennes des bruits connues et inconnues suivant une loi exponentielle	60
4.9 Erreur de reconstruction avec moyennes des bruits connues et inconnues suivant une loi exponentielle	61
4.10 Filtre de Kalman avec moyennes des bruits connues et inconnues suivant une loi khi-deux	63
4.11 Erreur de reconstruction avec moyennes des bruits connues et inconnues suivant une loi khi-deux	64

INTRODUCTION

Le filtre de Kalman est un processus optimal d'estimation d'état appliqué à un système dynamique qui implique des perturbations aléatoires. Plus précisément, le filtre de Kalman fournit un algorithme récursif linéaire, non biaisé et d'erreur minimal pour estimer de façon optimale l'état inconnu d'un système dynamique à partir de données bruitées prises en temps réel. La théorie du filtre de Kalman a reçu une attention toute particulière dans de nombreux domaines tels que l'ingénierie, la robotique, la navigation et l'économie depuis les années 60. Avec le développement récent des ordinateurs à grande vitesse de calcul, le filtre de Kalman est devenu plus utile même pour des applications en temps réel très compliquées. Il doit son nom à Rudolf Kalman, éminent chercheur d'origine hongroise qui a publié son fameux article *A new approach to Linear Filtering and Prediction Problems* [8] qui rencontra un certain scepticisme au début, mais qui a par la suite été adopté par la plupart des chercheurs surtout après son application dans le programme spatial Apollo.

Dans ce mémoire, nous définirons dans le premier chapitre les outils mathématiques nécessaires pour bien aborder le filtre de Kalman. Nous présenterons entre autres le pseudo-inverse de Moore-Penrose et la théorie de base sur l'estimation d'une variable aléatoire par la mesure d'une autre variable aléatoire, qui est l'estimation au sens des moindres carrés. Dans le deuxième chapitre nous nous intéresserons au filtre de Kalman. Nous construirons les différentes équations nécessaires au bon fonctionnement du filtre classique et nous

définirons une transformation linéaire qui s'applique à des bruits non gaussiens (exponentiel, khi-deux). Nous introduirons dans le troisième chapitre l'initialisation du filtre par la méthode des moindres carrés en faisant bien entendu des comparaisons avec la méthode d'initialisation «classique». Finalement dans le dernier chapitre, nous étudierons le filtre de Kalman dans le cas où nous devons estimer l'état d'un système dynamique lorsque les bruits sont décentrés, de moyennes connues ou inconnues.

CHAPITRE 1

Préliminaires mathématiques

1.1 Introduction

Nous introduisons dans ce chapitre quelques résultats de base que nous utiliserons tout au long de ce mémoire. Nous commencerons par présenter l'inégalité de Schwarz, en passant par la trace, le pseudo-inverse de Moore-Penrose et enfin nous donnerons une définition de l'estimation par la méthode des moindres carrés.

1.2 Matrice et déterminant

Définition 1.1. Soit A une (n,n) -matrice symétrique et x un n -vecteur colonne. On dit que A est une matrice définie positive si

$$x^T Ax > 0 \quad \forall x \in \mathbb{R}^n, x \neq 0,$$

et qu'elle est non-négative si

$$x^T Ax \geq 0 \quad \forall x \in \mathbb{R}^n.$$

Définition 1.2. Soit A et B deux (n, n) -matrices de nombres réels, nous allons utiliser la notation

$$A > B,$$

si la matrice $A - B$ est définie positive, et

$$A \geq B,$$

si la matrice $A - B$ est non-négative.

Lemme 1.3. Soit P une (m, n) -matrice et Q une (m, l) -matrice, alors nous avons

$$(Q - PS)^T(Q - PS) \geq 0$$

pour toute (n, l) -matrice S , avec égalité si et seulement si $Q = PS$.

Lemme 1.4. (Inégalité de Schwarz) Soit P une (m, n) -matrice telle que $P^T P$ soit inversible, et soit Q une (m, l) -matrice. Alors

$$Q^T Q \geq (P^T Q)^T (P^T P)^{-1} (P^T Q).$$

Démonstration : Il suffit de prendre $S = (P^T P)^{-1} P^T Q$ dans le lemme 1.3. \square

Définition 1.5. Soit $A = [a_{ij}]$ une (n, n) -matrice. La trace de A , notée $\text{tr} A$, est la somme de ses éléments diagonaux

$$\text{tr} A = \sum_{i=1}^n a_{ii}.$$

Nous allons donner quelques propriétés.

Lemme 1.6. Soient A et B deux (n, n) -matrices, alors

$$\operatorname{tr} A^T = \operatorname{tr} A,$$

$$\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B,$$

et

$$\operatorname{tr}(\lambda A) = \lambda \operatorname{tr} A.$$

Définition 1.7. Soit A une (n, m) -matrice, alors

$$\|A\|_F = (\operatorname{tr} A^T A)^{1/2} = (\operatorname{tr} A A^T)^{1/2},$$

où $\|A\|_F$ est la norme dite de Frobenius.

Lemme 1.8. Soit A une (n, m) -matrice et B une (m, n) -matrice alors

$$\operatorname{tr} AB = \operatorname{tr} B^T A^T = \operatorname{tr} BA = \operatorname{tr} A^T B^T,$$

et

$$\operatorname{tr} A^T A = \sum_{i=1}^n \sum_{j=1}^m a_{ij}^2 = \|A\|_F^2,$$

Lemme 1.9. Soit A une (n, n) -matrice avec les valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$, alors

$$\operatorname{tr} A = \sum_{i=1}^n \lambda_i.$$

Démonstration : Il nous suffit d'écrire $A = U J U^{-1}$ où J est la forme canonique de Jordan de A et U est une matrice non singulière. Alors

$$\operatorname{tr} A = \operatorname{tr} U J U^{-1} = \operatorname{tr} U^{-1} U J = \operatorname{tr} J = \sum_{i=1}^n \lambda_i. \quad \square$$

Il résulte de ce lemme que si $A > 0$ alors $\operatorname{tr} A > 0$, et si $A \geq 0$ alors $\operatorname{tr} A \geq 0$. Notons que l'inverse n'est pas vrai.

Définition 1.10. (Pseudo-inverse de Moore-Penrose [2]) *Soit M une (n, n) -matrice de nombres réels. Le pseudo-inverse de M , noté M^+ , est une matrice qui satisfait les quatre critères suivants :*

- $MM^+M = M$,
- $M^+MM^+ = M^+$,
- $(MM^+)^T = MM^+$
- $(M^+M)^T = M^+M$.

La matrice pseudo-inverse de Moore-Penrose existe et est unique. En effet pour toute matrice M il n'y a qu'une seule matrice M^+ qui peut satisfaire les quatre propriétés de la définition.

Une matrice qui satisfait à la première condition est dite «inverse généralisé». Si la matrice satisfait aussi la deuxième condition, elle est appelée «inverse généralisée réflexive». L'inverse généralisée existe toujours, mais n'est généralement pas unique, et l'unicité du pseudo-inverse est la conséquence des deux dernières conditions.

1.3 Estimation par la méthode des moindres carrés

Prenons l'équation d'observation d'un système linéaire tel que les données observées soient contaminées par du bruit [4, 15]

$$v = Cx + \eta$$

où, x est le n -vecteur d'état du processus, v le q -vecteur des données et C la (q, n) -matrice est donnée. En règle générale, η est inconnu, mais nous supposons que c'est un bruit à moyenne nulle : $E(\eta) = 0$ et $Var(\eta) = R$ où R est une matrice définie positive [13].

Notre but est d'obtenir un estimateur «optimal» \hat{y} du vecteur d'état x à partir de l'information v . Quand les données sont contaminées par du bruit, nous minimisons la

quantité

$$F(y, W) = E(v - Cy)^T W (v - Cy) = E\|v - Cy\|_W^2$$

sur tous les n -vecteurs y où W est une (q, q) -matrice symétrique définie positive, appelée *matrice de poids*. Nous cherchons $\hat{y} = \hat{y}(W)$ telle que

$$F(\hat{y}, W) = \min_y F(y, W).$$

Nous avons

$$F(y, W) = E\|W^{\frac{1}{2}}v - W^{\frac{1}{2}}Cy\|^2 = E\|Q - Py\|^2,$$

tel que $Q = W^{\frac{1}{2}}v$ et $P = W^{\frac{1}{2}}C$. Comme

$$\begin{aligned} E\|Q - Py\|^2 &= E\|Q - P\hat{y} + P\hat{y} - Py\|^2 \\ &= E\|Q - P\hat{y}\|^2 + E\|P\hat{y} - Py\|^2 + 2E(Q - P\hat{y})^T P(\hat{y} - y) \\ &\geq E\|Q - P\hat{y}\|^2 \end{aligned}$$

puisque $E(Q - P\hat{y})^T P = 0$ et $E\|P\hat{y} - Py\|^2 \geq 0$.

En supposant que C^TWC soit inversible, nous en déduisons la solution optimale

$$\hat{y} = (P^T P)^{-1} P^T v = (C^T WC)^{-1} C^T W v.$$

Nous voulons de plus déterminer un *poids optimal* \hat{W} , considérons

$$F(\hat{y}, W) = E(v - C\hat{y})^T W (v - C\hat{y}).$$

Pour que cette quantité atteigne une valeur minimale il faudrait que W soit égale à la matrice zéro. Par conséquent, nous avons besoin d'une autre mesure pour déterminer \hat{W} .

Notons que le problème original était d'estimer le vecteur d'état x par $\hat{y}(W)$, donc il est naturel de considérer la mesure de l'erreur $x - \hat{y}(W)$. Étant donné que l'on n'a pas assez d'informations sur x et que seul le bruit est mesurable, cette mesure devrait être déterminée par la variance de l'erreur. Donc, nous allons minimiser $Var(x - \hat{y}(W))$ par rapport aux matrices définies positives W . En écrivant $\hat{y} = \hat{y}(W)$, nous avons

$$\begin{aligned} x - \hat{y} &= (C^T WC)^{-1} (C^T WC)x - (C^T WC)^{-1} C^T Wv \\ &= (C^T WC)^{-1} C^T W(Cx - v) \\ &= -(C^T WC)^{-1} C^T W\eta. \end{aligned}$$

Par conséquent

$$\begin{aligned} Var(x - \hat{y}) &= (C^T WC)^{-1} C^T W E(\eta\eta^T) WC (C^T WC)^{-1} \\ &= (C^T WC)^{-1} C^T W R W C (C^T WC)^{-1}. \end{aligned}$$

Nous obtenons ainsi cette quantité avec laquelle nous allons chercher un critère d'optimisation. Pour écrire cela comme le carré d'une matrice, nous avons besoin de la racine carrée positive de la matrice symétrique définie positive R . Prenons les valeurs propres de R , $\lambda_1, \lambda_2, \dots$, qui sont toutes strictement positives, écrivons $R = U^T diag[\lambda_1, \lambda_2, \dots]U$ où U est une matrice unitaire formée par les vecteurs propres normalisés associés aux valeurs propres λ_i pour R . Définissons $R^{1/2} = U^T diag[\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots]U$ pour avoir $(R^{1/2})(R^{1/2}) = R_k$. Il en résulte que

$$Var(x - \hat{y}) = Q^T Q,$$

où $Q = (R^{1/2})^T WC(C^T WC)^{-1}$. D'après l'inégalité de Schwarz et sous l'hypothèse que P est une (q, n) -matrice, nous avons

$$Q^T Q \geq (P^T Q)^T (P^T P)^{-1} (P^T Q).$$

Par conséquent, nous pouvons choisir $P = R^{-1/2}C$, de façon à obtenir

$$P^T P = C^T (R^{-1/2})^T (R^{-1/2}) C = C^T R^{-1} C.$$

Ainsi

$$P^T Q = C^T R^{-1/2} (R^{1/2})^T WC(C^T WC)^{-1} = C^T WC(C^T WC)^{-1} = I,$$

d'où

$$\begin{aligned} (P^T Q)^T (P^T P)^{-1} (P^T Q) &= (C^T R^{-1} C)^{-1} \\ &= Var(x - \hat{y}(R^{-1})). \end{aligned}$$

Par conséquent, $Var(x - \hat{y}(W)) \geq Var(x - \hat{y}(R^{-1}))$ pour toute matrice de poids symétrique définie positive W . Donc, la matrice de poids optimale est $\hat{W} = R^{-1}$, et l'estimation de la variance minimale de x en utilisant ce poids optimal est

$$\hat{x} := \hat{y}(R^{-1}) = (C^T R^{-1} C)^{-1} C^T R^{-1} v$$

Nous appelons \hat{x} l'estimation optimale par les moindres carrés de x . Étant l'image d'une transformation linéaire des données v , cela nous donne une estimation non biaisée de x .

Effectivement nous avons

$$\begin{aligned}
E(\hat{x}) &= (C^T R^{-1} C)^{-1} C^T R^{-1} E(v - Du) \\
&= (C^T R^{-1} C)^{-1} C^T R^{-1} E(Cx + \eta) \\
&= (C^T R^{-1} C)^{-1} C^T R^{-1} (CE(x) + 0) \\
&= E(x).
\end{aligned}$$

Considérons cette fois ci le cas où $C^T R^{-1} C$ n'est pas inversible. Dans ce cas nous donnons, sans démonstration le théorème suivant.

Théorème 1.11. [2] *Considérons le modèle*

$$v = Cx + \eta$$

où η est une variable aléatoire de moyenne nulle et de matrice de variance-covariance R .

La solution \hat{x} de la minimisation

$$F(\hat{x}, R^{-1}) = \min_x E \|v - Cx\|_{R^{-1}}^2,$$

est donnée par

$$\hat{x} = (C^T R^{-1} C)^+ C^T R^{-1} v$$

où $(C^T R^{-1} C)^+$ est la matrice pseudo-inverse de Moore-Penrose de $C^T R^{-1} C$. La matrice de variance-covariance de cette estimation est

$$E[(x - \hat{x})(x - \hat{x})^T] = (C^T R^{-1} C)^+.$$

CHAPITRE 2

Filtre de Kalman

2.1 Introduction

Les problèmes de signaux bruités sont assez fréquents dans beaucoup de domaines, comme par exemple le domaine de la navigation où le signal est noyé dans des erreurs de mesures et d'observations [5, 11]. Notre but est d'extraire le signal vérité du signal brut observé. Pour cela nous allons introduire dans ce chapitre l'algorithme du filtre de Kalman qui fait appel à la dynamique de la cible qui définit son évolution dans le temps et ainsi obtenir de meilleures données même lorsque la nature précise du système modélisé est inconnue [7, 1]. Typiquement, on dispose d'une suite d'observations, obtenues après un traitement préalable du signal brut au niveau des capteurs, telle que chaque observation est reliée à un état inconnu précis de façon linéaire. Le but sera d'estimer cet état d'une manière optimale et récursive [3].

2.2 Système dynamique et bruit blanc

Définition 2.1. (Système dynamique) *Un système dynamique est un modèle mathématique qui permet de décrire un ensemble de composants en interaction en temps réel, étant donné les états antérieurs et l'état précédent, ils ne dépendent que de l'état précédent, des commandes appliquées et d'un processus aléatoire [12].*

Définition 2.2. (Bruit blanc) *Un processus aléatoire φ_n , $n = 0, 1, \dots$ est un bruit blanc si*

$$\begin{cases} E(\varphi_n) &= 0 \\ Var(\varphi_n) &< \infty \\ Cov(\varphi_n, \varphi_l) &= 0 \quad si \quad n \neq l \end{cases}$$

2.3 Le modèle

Pour établir le filtre de Kalman, nous avons besoin de deux équations. Ces équations sont [8, 10] :

1. l'équation d'état qui donne l'évolution temporelle de l'état

$$y_{k+1} = A_k y_k + B_k u_k + \Gamma_k \xi_k;$$

2. l'équation d'observation qui nous fournit des informations sur l'état

$$w_k = C_k y_k + D_k u_k + \eta_k,$$

tel que $k = 0, 1, \dots$. Dans ces équations :

- y_k est le n -vecteur d'état du processus ;
- A_k est la (n, n) -matrice déterministe de transition de l'état ;

- u_k est le m -vecteur déterministe de commande ;
- B_k est la (n, m) -matrice déterministe qui relie la commande u_k à l'état y_{k+1} ;
- ξ_k est le p -vecteur de bruit blanc, qui modélise l'erreur du processus, avec une moyenne et une (p, p) -matrice de variance covariance Q_k connues ;
- Γ_k est la (n, p) -matrice déterministe qui relie le bruit blanc ξ_k à l'état y_{k+1} ;
- w_k est le q -vecteur d'observation ;
- C_k est la (q, n) -matrice déterministe de mesure ;
- D_k est la (q, m) -matrice déterministe qui relie la commande u_k à l'observation w_k ;
- η_k est le q -vecteur de bruit blanc, qui modélise l'erreur du processus, avec une moyenne et une (q, q) -matrice de variance covariance R_k connues ;

Ce système peut être décomposé en la somme de deux systèmes linéaires. En posant

$$\begin{cases} y_{k+1} &= z_{k+1} + x_{k+1} \\ w_k &= s_k + v_k \end{cases}$$

on peut écrire

1. un système linéaire déterministe

$$\begin{cases} z_{k+1} &= A_k z_k + B_k u_k \\ s_k &= C_k z_k + D_k u_k \end{cases}$$

2. et un système linéaire stochastique

$$\begin{cases} x_{k+1} &= A_k x_k + \Gamma_k \xi_k \\ v_k &= C_k x_k + \eta_k \end{cases}$$

Notre objectif est de déterminer une estimation optimale et récursive des états y_k à partir des observations w_k .

2.4 Caractérisation orthogonale de l'estimateur à variance minimale

Soient x un n -vecteur aléatoire et w un q -vecteur aléatoire. Nous définissons le «produit scalaire» $\langle x, w \rangle$ comme étant une (n, q) -matrice

$$\langle x, w \rangle = E[xw^T].$$

La racine carrée non-négative de $\langle w, w \rangle$ sera une (q, q) -matrice symétrique notée $\|w\|_q$ et on a

$$\|w\|_q^2 = \|w\|_q \|w\|_q^T = \langle w, w \rangle.$$

De la même manière, $\|x\|_n$ sera la racine carrée non-négative de $\langle x, x \rangle$.

L'erreur quadratique moyenne d'un estimateur \hat{y} de x est donnée par

$$\begin{aligned} E[\|x - \hat{y}\|^2] &= E[(x - \hat{y})^T(x - \hat{y})] \\ &= \text{tr}(E[(x - \hat{y})(x - \hat{y})^T]) \\ &= \text{tr}\langle x - \hat{y}, x - \hat{y} \rangle \\ &= \text{tr}\|x - \hat{y}\|_n^2. \end{aligned}$$

Donc chercher \hat{y} qui minimise la variance $E[\|x - \hat{y}\|^2]$ revient à chercher \hat{y} qui minimise la trace de $\|x - \hat{y}\|_n^2$.

Lemme 2.3. *Soit x un n -vecteur aléatoire. Alors*

$$\text{tr}\|x\|_n^2 = 0 \text{ si et seulement si } x = 0.$$

Démonstration : Nous avons

$$\operatorname{tr} \|x\|_n^2 = \operatorname{tr} \langle x, x \rangle = \operatorname{tr} E[xx^T] = E[x^T x] = E[\|x\|^2].$$

Donc

$$\operatorname{tr} \|x\|_n^2 = 0 \text{ si et seulement si } E[\|x\|^2] = 0,$$

mais

$$E[\|x\|^2] = 0 \text{ si et seulement si } E[x_i^2] = 0 \text{ pour } i = 1, \dots, n,$$

et

$$E[x_i^2] = 0 \text{ si et seulement si } x_i = 0,$$

d'où le résultat. \square

Maintenant, soit w_0, \dots, w_r des q -vecteurs aléatoires. Considérons le sous-espace vectoriel de \mathbb{R}^n généré par ces q -vecteurs donné par

$$Y(w_0, \dots, w_r) = \{y \in \mathbb{R}^n : y = \sum_{i=0}^r P_i w_i, P_i \text{ est une } (n, q)\text{-matrice}\}.$$

Le problème de minimisation que nous allons étudier est de déterminer un \hat{y} dans $Y(w_0, \dots, w_r)$ de telle sorte que

$$\operatorname{tr} \|x - \hat{y}\|_n^2 = \min \{\operatorname{tr} \|x - y\|_n^2 : y \in Y(w_0, \dots, w_r)\}.$$

Lemme 2.4. (Existence d'une solution) *Soit $x \in \mathbb{R}^n$ donné. Il existe un $\hat{y} \in Y(w_0, \dots, w_r)$ tel que*

$$\operatorname{tr} \|x - \hat{y}\|_n^2 = \min \{\operatorname{tr} \|x - y\|_n^2 : y \in Y(w_0, \dots, w_r)\}.$$

Démonstration : On sait que $\min \{ \text{tr} \|x - y\|_n^2 : y \in Y(w_0, \dots, w_r)\} \geq 0$ et Y est de dimension finie. Donc il existe un $\hat{y} \in Y$ tel que

$$\text{tr} \|x - \hat{y}\|_n^2 = \min \{ \text{tr} \|x - y\|_n^2 : y \in Y(w_0, \dots, w_r)\}. \quad \square$$

Lemme 2.5. (Caractérisation de la solution) *Pour $\hat{y} \in Y(w_0, \dots, w_r)$, nous avons*

$$\text{tr} \|x - \hat{y}\|_n^2 = \min \{ \text{tr} \|x - y\|_n^2 : y \in Y(w_0, \dots, w_r)\}.$$

si et seulement si

$$\langle x - \hat{y}, w_j \rangle = O_{n \times q} \text{ pour tout } j = 0, 1, \dots, r.$$

Démonstration : Soit $\lambda \in \mathbb{R}$ et $\Delta y \in Y(w_0, \dots, w_r)$, Nous aurons $\hat{y} + \lambda \Delta y \in Y(w_0, \dots, w_r)$ pour tout $\lambda \in \mathbb{R}$. Alors

$$\begin{aligned} \text{tr} \|x - (\hat{y} + \lambda \Delta y)\|_n^2 &= \text{tr} \langle x - (\hat{y} + \lambda \Delta y), x - (\hat{y} + \lambda \Delta y) \rangle \\ &= \text{tr} \langle (x - \hat{y}) - \lambda \Delta y, (x - \hat{y}) - \lambda \Delta y \rangle \\ &= \text{tr} \|x - \hat{y}\|_n^2 - 2\lambda \text{tr} \langle x - \hat{y}, \Delta y \rangle + \lambda^2 \text{tr} \|\Delta y\|_n^2. \end{aligned}$$

Si $\min \{ \text{tr} \|x - y\|_n^2 : y \in Y(w_0, \dots, w_r)\} = \text{tr} \|x - \hat{y}\|_n^2$, alors $\text{tr} \|x - (\hat{y} + \lambda \Delta y)\|_n^2 - \text{tr} \|x - \hat{y}\|_n^2 \geq 0$. Ainsi pour tout $\lambda \in \mathbb{R}$ nous aurons

$$-2\lambda \text{tr} \langle x - \hat{y}, \Delta y \rangle + \lambda^2 \text{tr} \|\Delta y\|_n^2 \geq 0.$$

Prenons $\lambda > 0$, divisons par λ et faisons décroître λ vers 0, nous aurons :

$$\text{tr} \langle x - \hat{y}, \Delta y \rangle \leq 0.$$

De même pour $\lambda < 0$, divisons par λ et faisons croître λ vers 0, nous aurons :

$$\text{tr} \langle x - \hat{y}, \Delta y \rangle \geq 0.$$

D'où

$$\operatorname{tr}\langle x - \hat{y}, \Delta y \rangle = 0.$$

Posons $\Delta y = P_j w_j$, alors

$$\operatorname{tr}[\langle x - \hat{y}, w_j \rangle P_j^T] = 0.$$

Prenons

$$P_j = \langle x - \hat{y}, w_j \rangle,$$

alors nous avons

$$\operatorname{tr}[\langle x - \hat{y}, w_j \rangle \langle x - \hat{y}, w_j \rangle^T] = 0.$$

Donc d'après le lemme 1.8

$$\langle x - \hat{y}, w_i \rangle = O_{n \times q}$$

pour tout $i = 0, \dots, q$.

Réiproquement, supposons que

$$\langle x - \hat{y}, w_j \rangle = O_{n \times q} \text{ pour } j = 0, \dots, r.$$

Comme nous avons

$$\operatorname{tr}\|x - (\hat{y} + \Delta y)\|_n^2 = \operatorname{tr}\|x - \hat{y}\|_n^2 - 2\operatorname{tr}\langle x - \hat{y}, \Delta y \rangle + \operatorname{tr}\|\Delta y\|_n^2,$$

en posant $\Delta y = \sum_{i=0}^r P_i w_i$, nous obtenons

$$\begin{aligned} -2\operatorname{tr}\langle x - \hat{y}, \Delta y \rangle &= -2\operatorname{tr}\langle x - \hat{y}, \sum_{i=0}^r P_i w_i \rangle, \\ &= -2\operatorname{tr}[\sum_{i=0}^r \langle x - \hat{y}, w_i \rangle P_i^T], \\ &= 0. \end{aligned}$$

Ainsi

$$\operatorname{tr}\|x - (\hat{y} + \Delta y)\|_n^2 \geq \operatorname{tr}\|x - \hat{y}\|_n^2,$$

nous obtenons que \hat{y} est bien le minimum et

$$\text{tr} \|x - \hat{y}\|_n^2 = \min \{ \text{tr} \|x - y\|_n^2 : y \in Y(w_0, \dots, w_r) \}. \quad \square$$

Lemme 2.6. (Unicité de la solution) \hat{y} est uniquement déterminé dans le sens où

$$\text{tr} \|x - \hat{y}\|_n^2 = \text{tr} \|x - y\|_n^2 \text{ si et seulement si } \hat{y} = y.$$

Démonstration : Comme précédemment, nous avons

$$\text{tr} \|x - y\|_n^2 = \text{tr} \|x - \hat{y}\|_n^2 + \text{tr} \|\hat{y} - y\|_n^2.$$

Donc pour que $\text{tr} \|x - \hat{y}\|_n^2 = \text{tr} \|x - y\|_n^2$, il faut et il suffit que $\text{tr} \|\hat{y} - y\|_n^2 = 0$ c'est-à-dire $\hat{y} - y = 0$ selon le lemme 2.3 d'où le résultat. \square

2.5 Suite d'innovations

Pour une suite de q -vecteurs d'observations donnée $\{v_j\}_{j=0}^k$, considérons :

$$\hat{y}_{j-1} = \begin{cases} 0 & \text{si } j = 0, \\ \sum_{i=0}^{j-1} \hat{P}_{j-1,i} v_i \in Y(v_0, \dots, v_{j-1}) & \text{si } j = 1, \dots, k, \end{cases}$$

où les $\hat{P}_{j-1,i}$ sont des (n, q) -matrices choisies de telle sorte que pour chaque j on ait

$$\text{tr} \|x_j - \hat{y}_{j-1}\|_n^2 = \min \{ \text{tr} \|x_j - y_{j-1}\|_n^2 : y_{j-1} \in Y(v_0, \dots, v_{j-1}) \}.$$

Définition 2.7. La suite de q -vecteurs d'innovations $\{z_j\}_{j=0}^k$ est donnée par

$$z_j = v_j - C_j \hat{y}_{j-1} \quad (j = 0, \dots, k).$$

Pour simplifier posons $\hat{e}_j = C_j(x_j - \hat{y}_{j-1})$. Ainsi

$$\begin{aligned} z_j &= v_j - C_j \hat{y}_{j-1} \\ &= C_j x_j + \eta_j - C_j \hat{y}_{j-1} \\ &= C_j(x_j - \hat{y}_{j-1}) + \eta_j \\ &= \hat{e}_j + \eta_j. \end{aligned}$$

Lemme 2.8. *Nous avons*

$$\langle \eta_l, \hat{e}_j \rangle = O_{q \times q} \quad \text{pour tout } l \geq j. \quad (2.1)$$

Démonstration : On a

$$\begin{aligned} \hat{e}_j &= C_j(x_j - \hat{y}_{j-1}) \\ &= C_j(A_{j-1}x_{j-1} + \Gamma_{j-1}\xi_{j-1} - \sum_{i=0}^{j-1} \hat{P}_{j-1,i}(C_i x_i + \eta_i)) \\ &\vdots \\ &= B^0 x_0 + \sum_{i=0}^{j-1} B_{ji}^1 \xi_i + \sum_{i=0}^{j-1} B_{ji}^2 \eta_i. \end{aligned}$$

Ainsi

$$\begin{aligned} \langle \eta_l, \hat{e}_j \rangle &= \langle \eta_l, x_0 \rangle B^{0T} + \sum_{i=0}^{j-1} \langle \eta_l, \xi_i \rangle B^{1T} + \sum_{i=0}^{j-1} \langle \eta_l, \eta_i \rangle B^{2T} \\ &= O_{q \times q}. \end{aligned}$$

□

Lemme 2.9. (Orthogonalité) *La suite d'innovations $\{z_j\}$ associée à la suite d'observations $\{v_j\}$ satisfait la propriété*

$$\langle z_j, z_l \rangle = (R_l + C_l \|x_l - \hat{y}_{l-1}\|_n^2 C_l^T) \delta_{jl}.$$

Démonstration : D'une part pour $j = l$ nous avons

$$\begin{aligned}
\langle z_l, z_l \rangle &= \langle \hat{e}_l + \eta_l, \hat{e}_l + \eta_l \rangle \\
&= \langle \hat{e}_l, \hat{e}_l \rangle + \langle \eta_l, \eta_l \rangle \\
&= C_l \|x_l - \hat{y}_{l-1}\|_n^2 C_l^T + R_l.
\end{aligned}$$

Pour le cas où $j \neq l$, supposons $j > l$, nous avons

$$\begin{aligned}
\langle z_j, z_l \rangle &= \langle \hat{e}_j + \eta_j, \hat{e}_l + \eta_l \rangle \\
&= \langle \hat{e}_j, \hat{e}_l + \eta_l \rangle + \langle \eta_j, \hat{e}_l \rangle + \langle \eta_j, \eta_l \rangle \\
&= \langle \hat{e}_j, z_l \rangle \\
&= \langle \hat{e}_j, v_l - C_l \hat{y}_{l-1} \rangle \\
&= \langle C_j(x_j - \hat{y}_{j-1}), v_l - C_l \sum_{i=0}^{l-1} \hat{P}_{l-1,i} v_i \rangle \\
&= C_j \langle x_j - \hat{y}_{j-1}, v_l \rangle - C_j \sum_{i=0}^{l-1} \langle x_j - \hat{y}_{j-1}, v_i \rangle \hat{P}_{l-1,i}^T C_l^T \\
&= O_{q \times q}.
\end{aligned}$$

car $\langle x_j - \hat{y}_{j-1}, v_i \rangle = 0$ pour $i = 0, \dots, j-1$. \square

Nous allons maintenant normaliser la suite d'innovations orthogonales en posant

$$e_j = \|z_j\|_q^{-1} z_j = \|\hat{e}_j + \eta_j\|_q^{-1} (\hat{e}_j + \eta_j).$$

Ainsi nous aurons une suite orthonormale $\{e_0, e_1, \dots\}$. Donc

$$\langle e_j, e_i \rangle = \begin{cases} O_{q \times q} & \text{si } j \neq l \\ I_{q \times q} & \text{si } j = l \end{cases}$$

Lemme 2.10. *Nous avons*

$$Y(e_0, \dots, e_k) = Y(v_0, \dots, v_k).$$

Démonstration :

Si $w \in Y(e_0, \dots, e_k)$, alors nous avons

$$w = \sum_{i=0}^k P_{w,i} e_i = \sum_{i=0}^k P_{w,i} \|z_j\|_q^{-1} (v_j - C_j (\sum_{i=0}^{j-1} \hat{P}_{j-1,i} v_i)),$$

et ainsi $w \in Y(v_0, \dots, v_k)$.

D'un autre côté, nous avons $w \in Y(v_0, \dots, v_k)$ et alors

$$w = \sum_{i=0}^k P_{w,i} v_i.$$

Mais

$$\begin{aligned} v_0 &= z_0 = \|z_0\|_q e_0 \\ v_1 &= z_1 + C_1 \hat{y}_0 = z_1 + C_1 \hat{P}_{0,0} v_0 = \|z_1\|_q e_1 + C_1 \hat{P}_{0,0} \|z_0\|_q e_0 \\ &\vdots \\ v_j &= z_j + C_j \hat{y}_{j-1} = \|z_j\|_q e_j + C_j \sum_{i=0}^{j-1} \hat{P}_{j-1,i} v_i \end{aligned}$$

et donc nous obtenons le résultat. \square

Comme $Y(e_0, \dots, e_k) = Y(v_0, \dots, v_k)$, nous avons que

$$v_j = \sum_{l=0}^j L_l e_l.$$

Donc, lorsque $j < k$, nous avons

$$\langle v_j, z_k \rangle = \langle v_j, e_k \rangle \|z_k\|_q = \sum_{l=0}^j L_l \langle e_l, e_k \rangle \|z_k\|_q = O_{q \times q}.$$

De même

$$\hat{y}_j = \sum_{l=0}^j \hat{P}_{ji} v_i,$$

et ainsi, lorsque $j < k$, nous avons

$$\langle \hat{y}_j, z_k \rangle = \sum_{i=0}^j \hat{P}_{ji} \langle v_i, z_k \rangle = \sum_{i=0}^j \hat{P}_{ji} \langle v_i, e_k \rangle \|z_k\|_q = O_{n \times q}.$$

2.6 Estimateur à variance minimale

Nous avons maintenant tous les outils nécessaires pour identifier l'estimation de variance minimale \hat{x}_k du vecteur d'état x_k .

Théorème 2.11. *Étant donné x_k , l'expression $\hat{x}_k = \sum_{i=0}^k \langle x_k, e_i \rangle e_i$ est l'estimateur de variance minimale de x_k vérifiant*

$$tr \|x_k - \hat{x}_k\|_n^2 = \min \{ tr \|x_k - y\|_n^2 : y \in Y(v_0, \dots, v_k) \}$$

Démonstration : Le développement de Fourier tronquée de x_k par rapport à la suite orthonormale $\{e_0, e_1, e_2, \dots\}$ est

$$\hat{x}_k = \sum_{i=0}^k \langle x_k, e_i \rangle e_i.$$

Par orthogonalité des e_i , nous avons

$$\begin{aligned} \langle \hat{x}_k, e_j \rangle &= \sum_{i=0}^k \langle x_k, e_i \rangle \langle e_i, e_j \rangle \\ &= \sum_{i=0}^k \langle x_k, e_i \rangle \delta_{ij} \\ &= \langle x_k, e_j \rangle \end{aligned}$$

et ainsi

$$\langle x_k - \hat{x}_k, e_j \rangle = O_{n \times q}, \quad j = 0, 1, \dots, k.$$

Comme $Y(e_0, \dots, e_k) = Y(v_0, \dots, v_k)$ nous obtenons donc notre résultat. \square

2.7 Équations du filtre de Kalman

Définition 2.12. *Nous pouvons classer les problèmes d'estimation de l'état d'un système dynamique selon la quantité d'information disponible.*

- Pour $j = k$, nous notons $\hat{x}_k = \hat{x}_{k|k}$ et nous appelons cette estimation, le processus de filtrage.
- pour $j < k$, nous définissons $\hat{x}_{k|j}$ la prédiction optimale de x_k et nous appelons cette estimation, le processus de prédiction.
- Pour $j > k$, nous définissons $\hat{x}_{k|j}$ l'estimation optimale de lissage x_k et nous appelons cette estimation, le processus de lissage.

Pour le problème de filtrage, $\hat{x}_k = \hat{x}_{k|k}$ est déterminé en utilisant les données v_0, \dots, v_k .

Nous allons développer une formule récursive qui va estimer $\hat{x}_k = \hat{x}_{k|k}$ à partir des prédictions $\hat{x}_{k|k-1}$ et $\hat{x}_{k-1} = \hat{x}_{k-1|k-1}$. Donc à chaque étape, nous n'utiliserons qu'une petite partie de l'information de nos données qui ne nécessite qu'un petit espace de stockage. En résumé, nous allons procéder en deux étapes :

- L'étape de prédiction.
- L'étape de correction.

C'est ce qu'on appelle généralement l'algorithme du filtre de Kalman.

Notations et remarques :

Nous avons

$$\begin{aligned}
v_j &= C_j x_j + \eta_j & \in \mathbb{R}^q \\
x_j &= A_{j-1} x_{j-1} + \Gamma_{j-1} \xi_{j-1} & \in \mathbb{R}^n \quad \text{c'est-à-dire que } x_j \in Y(\xi_{j-1}, \xi_{j-2}, \dots, \xi_0; x_0) \\
\hat{y}_{j-1} &= \sum_{i=0}^{j-1} \hat{P}_{j-1,i} v_i & \in \mathbb{R}^q \quad \text{c'est-à-dire que } \hat{y}_{j-1} \in Y(\eta_{j-1}, \dots, \eta_0; \xi_{j-1}, \dots, \xi_0; x_0) \\
\hat{e}_j &= C_j(x_j - \hat{y}_{j-1}) & \in \mathbb{R}^q \\
z_j &= \hat{e}_j + \eta_j & \in \mathbb{R}^q \\
e_j &= \|z_j\|_q^{-1} z_j & \in \mathbb{R}^q.
\end{aligned}$$

Nous en déduisons que

$$\begin{aligned}
v_j &\in Y(\eta_j; \xi_{j-1}, \dots, \xi_0; x_0), \\
z_j, e_j &\in Y(\eta_j, \dots, \eta_0; \xi_{j-1}, \dots, \xi_0; x_0), \\
\hat{e}_j &\in Y(\eta_{j-1}, \dots, \eta_0; \xi_{j-1}, \dots, \xi_0; x_0).
\end{aligned}$$

De plus nous avons déjà démontré que

$$\begin{aligned}
\langle \xi_k, e_j \rangle &= O_{p \times q} \quad j = 0, 1, \dots, k, \\
\langle \eta_k, e_j \rangle &= O_{q \times q} \quad j = 0, 1, \dots, k-1, \\
\langle x_j, \xi_j \rangle &= O_{n \times p} \quad \forall j, \\
\langle x_k, \eta_j \rangle &= O_{n \times q} \quad \forall j, \forall k.
\end{aligned}$$

2.7.1 Initialisation

Nous n'avons aucune mesure disponible au premier instant, donc il paraît logique, en l'absence de mesure, d'estimer le premier état par son espérance

$$\hat{x}_{0|0} = E(x_0)$$

et la matrice de variance covariance de l'erreur par

$$P_{0,0} = E[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T].$$

2.7.2 Prédiction-Correction

Passons maintenant à l'étape de prédition-correction, qui consiste à estimer l'état prédit $\hat{x}_{k|k-1}$ avec sa matrice de variance-covariance $P_{k,k-1}$, ainsi que l'état corriger $\hat{x}_{k|k}$ et sa matrice de variance-covariance $P_{k,k}$. Nous avons

$$\begin{aligned}\hat{x}_k &= \sum_{j=0}^k \langle x_k, e_j \rangle e_j \\ &= \sum_{j=0}^{k-1} \langle x_k, e_j \rangle e_j + \langle x_k, e_k \rangle e_k \\ &= \sum_{j=0}^{k-1} \{ \langle A_{k-1} x_{k-1}, e_j \rangle e_j + \langle \Gamma_{k-1} \xi_{k-1}, e_j \rangle e_j \} + \langle x_k, e_k \rangle e_k \\ &= A_{k-1} \sum_{j=0}^{k-1} \langle x_{k-1}, e_j \rangle e_j + \Gamma_{k-1} \sum_{j=0}^{k-1} \langle \xi_{k-1}, e_j \rangle e_j + \langle x_k, e_k \rangle e_k \\ &= A_{k-1} \hat{x}_{k-1} + \langle x_k, e_k \rangle e_k.\end{aligned}$$

Par conséquent, en définissant

$$\hat{x}_{k|k-1} = A_{k-1} \hat{x}_{k-1},$$

où $\hat{x}_{k-1} = \hat{x}_{k-1|k-1}$, nous obtenons

$$\hat{x}_k = \hat{x}_{k|k} = \hat{x}_{k|k-1} + \langle x_k, e_k \rangle e_k.$$

Notre but est de montrer qu'il existe une (n, q) -matrice constante G_k telle que

$$\langle x_k, e_k \rangle e_k = G_k(v_k - C_k \hat{x}_{k|k-1}).$$

Lemme 2.13. *Pour $j = 0, 1, \dots, k$, nous avons*

$$\langle v_k - c_k \hat{x}_{k|k-1}, e_j \rangle = \|z_k\|_q \delta_{kj}.$$

Démonstration :

Pour $k = j$, nous avons

$$\begin{aligned} \langle v_k - C_k \hat{x}_{k|k-1}, e_k \rangle &= \langle v_k - C_k(\hat{x}_{k|k} - \langle x_k, e_k \rangle e_k), e_k \rangle \\ &= \langle v_k, e_k \rangle - C_k \{ \langle \hat{x}_{k|k}, e_k \rangle - \langle x_k, e_k \rangle \} \\ &= \langle v_k, e_k \rangle - C_k \langle \hat{x}_{k|k} - x_k, e_k \rangle \\ &= \langle v_k, e_k \rangle \\ &= \langle z_k + C_k \hat{y}_{k-1}, \|z_k\|_q^{-1} z_k \rangle \\ &= \langle z_k, z_k \rangle \|z_k\|_q^{-1} + C_k \langle \hat{y}_{k-1}, z_k \rangle \|z_k\|_q^{-1} \\ &= \|z_k\|_q^2 \|z_k\|_q^{-1} + C_k O_{n \times q} \|z_k\|_q^{-1} \\ &= \|z_k\|_q. \end{aligned}$$

Pour $k < j$, nous avons

$$\begin{aligned} \langle v_k - C_k \hat{x}_{k|k-1}, e_j \rangle &= \langle C_k x_k + \eta_k - C_k(\hat{x}_{k|k} - \langle x_k, e_k \rangle e_k), e_j \rangle \\ &= C_k \langle x_k - \hat{x}_{k|k}, e_j \rangle + \langle \eta_k, e_j \rangle + C_k \langle x_k, e_k \rangle \langle e_k, e_j \rangle \\ &= O_{q \times q}. \end{aligned}$$

□

On a

$$\hat{x}_{k-1} = \hat{x}_{k-1|k-1} \in Y(e_0, \dots, e_{k-1})$$

et

$$\hat{x}_{k|k-1} = A_{k-1} \hat{x}_{k-1|k-1} \in Y(e_0, \dots, e_{k-1}),$$

$$C_k \hat{x}_{k|k-1} \in Y(e_0, \dots, e_{k-1}),$$

$$v_k - C_k \hat{x}_{k|k-1} \in Y(e_0, \dots, e_{k-1}).$$

Ainsi nous pouvons écrire

$$v_k - C_k \hat{x}_{k|k-1} = \sum_{i=0}^k M_i e_i$$

et pour $j = 0, 1, \dots, k$ nous avons

$$\langle v_k - C_k \hat{x}_{k|k-1}, e_j \rangle = \|z_k\|_q \delta_{kj} = \left\langle \sum_{i=0}^k M_i e_i, e_j \right\rangle = \sum_{i=0}^k M_i \langle e_i, e_j \rangle = M_j.$$

Donc $M_0 = M_1 = \dots = M_{k-1} = 0$ et $M_k = \|z_k\|_q$. Ainsi

$$v_k - C_k \hat{x}_{k|k-1} = M_k e_k = \|z_k\|_q e_k.$$

On pose

$$G_k = \langle x_k, e_k \rangle \|z_k\|_q^{-1},$$

alors nous obtenons

$$\langle x_k, e_k \rangle e_k = G_k \|z_k\|_q e_k = G_k (v_k - C_k \hat{x}_{k|k-1}).$$

En résumé

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + G_k \|z_k\|_q e_k = G_k (v_k - C_k \hat{x}_{k|k-1}).$$

Nous allons maintenant construire une formule récursive pour G_k . Nous avons

$$x_k - \hat{x}_{k|k} = A_{k-1} x_{k-1} + \Gamma_{k-1} \zeta_{k-1} - A_{k-1} \hat{x}_{k-1|k-1} - G_k (v_k - C_k A_{k-1} \hat{x}_{k-1|k-1}).$$

En utilisant $v_k = C_k x_k + \eta_k = C_k A_{k-1} x_{k-1} + C_k \Gamma_{k-1} \xi_k + \eta_k$, nous aurons

$$x_k - \hat{x}_{k|k} = (I - G_k C_k) A_{k-1} (x_{k-1} - \hat{x}_{k-1|k-1}) + (I - G_k C_k) \Gamma_{k-1} \xi_k - G_k \eta_k.$$

D'une part

$$\langle x_k - \hat{x}_{k|k}, v_k \rangle = 0.$$

Ainsi

$$\begin{aligned} 0 = \langle x_k - \hat{x}_{k|k}, v_k \rangle &= \langle (I - G_k C_k) A_{k-1} (x_{k-1} - \hat{x}_{k-1|k-1}) + (I - G_k C_k) \Gamma_{k-1} \xi_{k-1} - G_k \eta_k, \\ &\quad C_k A_{k-1} ((x_{k-1} - \hat{x}_{k-1|k-1}) + \hat{x}_{k-1|k-1}) + C_k \Gamma_{k-1} \xi_{k-1} + \eta_k \rangle \\ &= (I - G_k C_k) A_{k-1} \|x_{k-1} - \hat{x}_{k-1|k-1}\|_n^2 A_{k-1}^T C_k^T \\ &\quad + (I - G_k C_k) A_{k-1} \langle x_{k-1} - \hat{x}_{k-1|k-1}, \hat{x}_{k-1|k-1} \rangle A_{k-1}^T C_k^T \\ &\quad + (I - G_k C_k) A_{k-1} \langle x_{k-1} - \hat{x}_{k-1|k-1}, \xi_{k-1} \rangle \Gamma_{k-1}^T C_k^T \\ &\quad + (I - G_k C_k) A_{k-1} \langle x_{k-1} - \hat{x}_{k-1|k-1}, \eta_k \rangle \\ &\quad + (I - G_k C_k) \Gamma_{k-1} \langle \xi_{k-1}, x_{k-1} - \hat{x}_{k-1|k-1} \rangle A_{k-1}^T C_k^T \\ &\quad + (I - G_k C_k) \Gamma_{k-1} \langle \xi_{k-1}, \hat{x}_{k-1|k-1} \rangle A_{k-1}^T C_k^T \\ &\quad + (I - G_k C_k) \Gamma_{k-1} \langle \xi_{k-1}, \xi_{k-1} \rangle \Gamma_{k-1}^T C_k^T \\ &\quad + (I - G_k C_k) \Gamma_{k-1} \langle \xi_{k-1}, \eta_k \rangle \\ &\quad - G_k \langle \eta_k, x_{k-1} - \hat{x}_{k-1|k-1} \rangle A_{k-1}^T C_k^T \\ &\quad - G_k \langle \eta_k, \xi_{k-1} \rangle \Gamma_{k-1}^T C_k^T \\ &\quad - G_k \langle \eta_k, \hat{x}_{k-1|k-1} \rangle A_{k-1}^T C_k^T \\ &\quad - G_k \langle \eta_k, \eta_k \rangle. \end{aligned}$$

Comme

$$\begin{aligned}
\langle x_{k-1} - \hat{x}_{k-1|k-1}, \hat{x}_{k-1|k-1} \rangle &= O_{n \times n} \\
\langle x_{k-1} - \hat{x}_{k-1|k-1}, \xi_{k-1} \rangle &= O_{n \times p} \\
\langle x_{k-1} - \hat{x}_{k-1|k-1}, \eta_k \rangle &= O_{n \times q} \\
\langle \xi_{k-1}, \hat{x}_{k-1|k-1} \rangle &= O_{p \times n} \\
\langle \xi_{k-1}, \eta_{k-1} \rangle &= O_{p \times q} \\
\langle \eta_{k-1}, \hat{x}_{k-1|k-1} \rangle &= O_{q \times n},
\end{aligned}$$

nous en déduisons que

$$\begin{aligned}
O_{q \times q} &= (I - G_k C_k) A_{k-1} \|x_{k-1} - \hat{x}_{k-1|k-1}\|_n^2 A_{k-1}^T C_k^T \\
&\quad + (I - G_k C_k) \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T C_k^T - G_k R_k.
\end{aligned} \tag{2.2}$$

Définissons

$$P_{k,k} = \|x_k - \hat{x}_{k|k}\|_n^2$$

et

$$P_{k,k-1} = \|x_k - \hat{x}_{k|k-1}\|_n^2.$$

Nous avons

$$\begin{aligned}
P_{k,k-1} &= \|x_k - \hat{x}_{k|k-1}\|_n^2 \\
&= \|A_{k-1}(x_{k-1} - \hat{x}_{k-1|k-1}) + \Gamma_{k-1} \xi_{k-1}\|_n^2 \\
&= A_{k-1} \|x_{k-1} - \hat{x}_{k-1|k-1}\|_n^2 A_{k-1}^T + \Gamma_{k-1} \|\xi_{k-1}\|_q^2 \Gamma_{k-1}^T \\
&= A_{k-1} P_{k-1|k-1} A_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T.
\end{aligned}$$

En isolant G_k dans (2.2), nous aurons

$$\begin{aligned}
G_k [R_k + C_k (A_{k-1} P_{k-1|k-1} A_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T) C_k^T] \\
= [A_{k-1} P_{k-1|k-1} A_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T] C_k,
\end{aligned}$$

c'est-à-dire

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

d'où

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

puisque $R_k + C_k P_{k|k-1} C_k^T > 0$.

Nous souhaitons maintenant écrire $P_{k|k}$ d'une manière récursive. Ainsi

$$\begin{aligned} P_{k,k} &= \|x_k - \hat{x}_{k|k}\|_n^2 \\ &= \|x_k - (\hat{x}_{k|k-1} + G_k(v_k - C_k \hat{x}_{k|k-1}))\|_n^2 \\ &= \|x_k - \hat{x}_{k|k-1} - G_k(C_k x_k + \eta_k) + G_k C_k \hat{x}_{k|k-1}\|_n^2 \\ &= \|(I - G_k C_k)(x_k - \hat{x}_{k|k-1}) - G_k \eta_k\|_n^2 \\ &= (I - G_k C_k) \|x_k - \hat{x}_{k|k-1}\|_n^2 (I - G_k C_k)^T + G_k R_k G_k^T \\ &= (I - G_k C_k) P_{k,k-1} (I - G_k C_k)^T + G_k R_k G_k^T, \end{aligned}$$

puisque $\langle x_k - \hat{x}_{k|k-1}, \eta_k \rangle = O_{n \times q}$. Notons que

$$\begin{aligned} (I - G_k C_k) P_{k,k-1} (G_k C_k)^T &= P_{k,k-1} C_k^T G_k^T - G_k C_k P_{k,k-1} C_k^T G_k^T \\ &= G_k C_k P_{k,k-1} C_k^T G_k^T + G_k R_k G_k^T - G_k C_k P_{k,k-1} C_k^T G_k^T \\ &= G_k R_k G_k^T. \end{aligned}$$

Ainsi nous aurons

$$\begin{aligned} P_{k,k} &= (I - G_k C_k) P_{k,k-1} (I - G_k C_k)^T + (G_k C_k) P_{k,k-1} (I - G_k C_k)^T \\ &= (I - G_k C_k) P_{k,k-1}. \end{aligned}$$

Nous obtenons enfin les équations du filtre de Kalman pour la partie stochastique

$$\left\{ \begin{array}{l} \hat{x}_{0|0} = E(x_0) \\ P_{0,0} = E[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T] \\ \text{Pour } k = 1, 2, \dots, \text{ faire} \\ P_{k,k-1} = A_{k-1}P_{k-1,k-1}A_{k-1}^T + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^T \\ \hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} \\ G_k = P_{k,k-1}C_k^T(R_k + C_kP_{k,k-1}C_k^T)^{-1} \\ P_{k,k} = (I - G_kC_k)P_{k,k-1} \\ \hat{x}_{k|k} = \hat{x}_{k|k-1} + G_k(v_k - C_k\hat{x}_{k|k-1}) \end{array} \right.$$

Pour avoir le système linéaire déterministe / stochastique, nous allons utiliser la méthode de la superposition. Rappelons l'équation générale du filtre de Kalman

$$\left\{ \begin{array}{l} y_{k+1} = A_k y_k + B_k u_k + \Gamma_k \xi_k \\ w_k = C_k y_k + D_k u_k + \eta_k. \end{array} \right.$$

Avec le système déterministe

$$\left\{ \begin{array}{l} z_{k+1} = A_k z_k + B_k u_k \\ s_k = C_k z_k + D_k u_k. \end{array} \right.$$

Ainsi

$$\begin{aligned} (y_{k+1} - z_{k+1}) &= A_k(y_k - z_k) + \Gamma_k \xi_k \\ (w_k - s_k) &= C_k(y_k - z_k) + \eta_k. \end{aligned}$$

Nous posons

$$\left\{ \begin{array}{l} x_k = y_k - z_k \\ v_k = w_k - s_k. \end{array} \right.$$

Nous pouvons maintenant trouver les valeurs de $\hat{y}_{k|k-1}$ et $\hat{y}_{k|k}$

$$\begin{aligned}\hat{x}_{k|k-1} &= A_{k-1}\hat{x}_{k-1|k-1} \\ \hat{y}_{k|k-1} - z_k &= A_{k-1}(\hat{y}_{k-1|k-1} - z_{k-1}) \\ \hat{y}_{k|k-1} &= A_{k-1}\hat{y}_{k-1|k-1} + z_k - A_{k-1}z_{k-1} \\ &= A_{k-1}\hat{y}_{k-1|k-1} + B_{k-1}u_{k-1},\end{aligned}$$

et

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + G_k(v_k - C_k\hat{x}_{k|k-1}) \\ \hat{y}_{k|k} - z_k &= \hat{y}_{k|k-1} - z_k + G_k(w_k - s_k - C_k(\hat{y}_{k|k-1} - z_k)) \\ \hat{y}_{k|k} &= \hat{y}_{k|k-1} + G_k(w_k - D_ku_k - C_k\hat{y}_{k|k-1}).\end{aligned}$$

Nous obtenons l'algorithme du filtre de Kalman

$$\left\{ \begin{array}{l} \hat{y}_{0|0} = E(y_0) \\ P_{0,0} = E[(y_0 - \hat{y}_{0|0})(y_0 - \hat{y}_{0|0})^T] \\ \text{Pour } k = 1, 2, \dots, \text{ faire} \\ \quad P_{k,k-1} = A_{k-1}P_{k-1,k-1}A_{k-1}^T + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^T \\ \quad \hat{y}_{k|k-1} = A_{k-1}\hat{y}_{k-1|k-1} + B_{k-1}u_{k-1} \\ \quad G_k = P_{k,k-1}C_k^T(R_k + C_kP_{k,k-1}C_k^T)^{-1} \\ \quad P_{k,k} = (I - G_kC_k)P_{k,k-1} \\ \quad \hat{y}_{k|k} = \hat{y}_{k|k-1} + G_k(w_k - D_ku_k - C_k\hat{y}_{k|k-1}) \end{array} \right. \quad (2.3)$$

2.8 Exemple

Pour illustrer la théorie développée précédemment, nous considérerons l'exemple simple suivant :

$$\begin{cases} x_{k+1} &= (\sqrt{2} + (-1)^k)x_k + \xi_k \\ v_k &= 0.5x_k + \eta_k \end{cases} \quad (2.4)$$

avec des bruits gaussiens $\xi_k \sim N(0, (0.3)^2)$ et $\eta_k \sim N(0, (0.3)^2)$.

La particularité de cet exemple est que la matrice A_k change à chaque itération k , nous remarquerons aussi l'absence de paramètres déterministes (B_k, u_k et D_k).

Nous allons faire plusieurs tests en fixant la valeur de $E(x_0) = 0$ et faire varier les valeurs de $P_{0,0}$ en reprenant bien sûr le même échantillon.

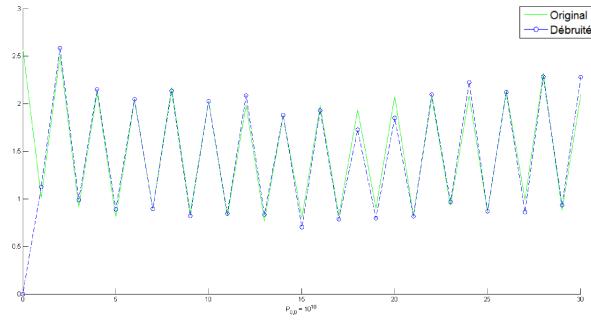
La méthode de simulation est simple et se divise en trois étapes [9, 6] :

- Générer la trajectoire bruité ;
- débruiter avec le filtre de Kalman ;
- comparer notre signal généré sans bruit et notre signal débruité.

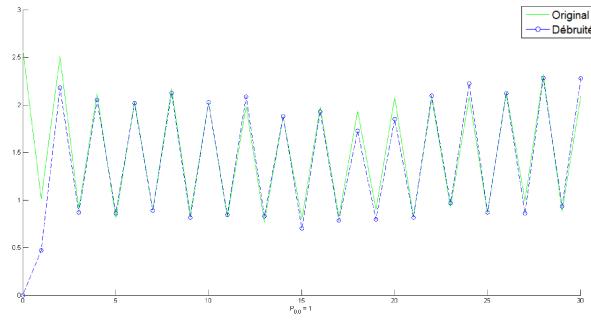
2.8.1 Calcul des erreurs de reconstruction

Nous allons calculer ici l'erreur de reconstruction du signal, pour cela nous allons définir l'erreur quadratique moyenne sur un échantillon de taille N , nous avons :

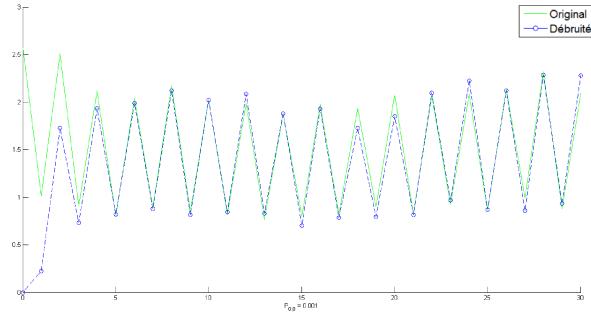
$$Err^{(N)} = \frac{\|x - \hat{x}\|_2}{\|x\|_2} = \sqrt{\frac{\sum_{k=1}^N |x_k - \hat{x}_{k|k}|^2}{\sum_{k=1}^N |x_k|^2}}$$



(a) $P_{0,0} = 10^{10}$



(b) $P_{0,0} = 1$



(c) $P_{0,0} = 0.001$

Figure 2.1 – Reconstruction du signal avec $\hat{x}_{0|0} = 0$ et différentes valeurs de $P_{0,0}$.

Nous allons résumer nos résultats dans le tableau suivant

$P_{0,0}$	0.001	1	10	10^2	10^{10}
$Err^{(30)}$ avant filtrage	0.5974	0.5974	0.5974	0.5974	0.5974
$Err^{(30)}$ après filtrage	0.3125	0.2951	0.2909	0.2876	0.2873

Tableau 2.1 – Erreur quadratique moyenne sur $N = 30$ étapes et différentes valeurs de $P_{0,0}$

Nous remarquerons ici deux choses essentielles. La première est que plus la valeur de $P_{0,0}$ est petite plus le débruitage met du temps à converger vers la vraie valeur du signal original. La deuxième chose à remarquer est que l'erreur obtenue avec $P_{0,0} = 10^{10}$ et $P_{0,0} = 1$ est à peu près identique, cela montre qu'à un certain degré il n'est plus vraiment nécessaire d'augmenter la valeur de $P_{0,0}$.

Calculons également la décroissance de l'erreur tout au long de l'estimation de l'état avant et après filtrage. le Tableau 2.2 illustre cette décroissance pour $N = 5, 10, 15, 20, 25, 30$.

N	5	10	15	20	25	30
$Err^{(N)}$ avant filtrage	0.9374	0.6891	0.5993	0.5987	0.5981	0.5974
$Err^{(N)}$ après filtrage	0.6568	0.4801	0.3074	0.2951	0.2893	0.2876

Tableau 2.2 – Décroissance de l'erreur de reconstruction du signal avec $\hat{x}_{0|0} = 0$, $P_{0,0} = 10^2$.

2.9 Autres types de bruits

Nous allons voir dans cette section des applications de l'algorithme de Kalman (2.3) avec des bruits pas nécessairement gaussiens. En effet, la distribution du bruit peut être quelconque.

2.9.1 Transformation de vecteur aléatoire

Proposition 2.14. *Prenons un vecteur aléatoire U qui suit une certaine loi de probabilité avec une moyenne μ et une matrice de variance-covariance Ω définie positive, qu'on note par $U \sim \mathcal{L}(\mu, \Omega)$. Nous pouvons appliquer des transformations linéaires sur U afin d'obtenir une nouvelle moyenne m et une nouvelle matrice de variance-covariance \sum définie positive. Il suffit de définir un second vecteur aléatoire V tel que*

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

avec $W = \sum^{\frac{1}{2}} \Omega^{-\frac{1}{2}}$ où $\sum^{\frac{1}{2}}$ désigne une racine carrée de \sum et $\Omega^{-\frac{1}{2}}$ désigne l'inverse de $\Omega^{\frac{1}{2}}$.

2.9.2 Bruit suivant une loi exponentielle

Définition 2.15. *On dit que le vecteur $\xi = (\xi_1(m_1), \xi_2(m_2), \dots, \xi_d(m_d))$ dans \mathbb{R}^d suit une loi exponentielle, et on le note par $\xi \sim \mathcal{E}(m)$ où $m = (m_1, m_2, \dots, m_d)$, si :*

- pour tout i , $\xi_i(m_i)$ est une variable aléatoire exponentielle de moyenne m_i ;
- les ξ_i sont indépendantes.

Nous allons considérer un problème d'estimation d'un signal à l'aide du filtre de Kalman lorsque les bruits de l'état et de l'observation sont de loi exponentielle :

$$\begin{cases} x_{k+1} &= A_k x_k + \xi_k \\ v_k &= C_k x_k + \eta_k \end{cases}$$

avec $\xi_k = m_x + W_1(\epsilon - m)$ et $\eta_k = m_y + W_2(\epsilon' - m')$ où $\epsilon \sim \mathcal{E}(m)$ et $\epsilon' \sim \mathcal{E}(m')$, nous aurons donc :

$$\begin{cases} \mathbb{E}(\xi_k) &= m_x \text{ et } \text{Var}(\xi_k) = W_1 \text{Var}(\epsilon) W_1^T \\ \mathbb{E}(\eta_k) &= m_y \text{ et } \text{Var}(\eta_k) = W_2 \text{Var}(\epsilon') W_2^T \end{cases}$$

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

Exemple

Reprenons l'exemple (2.4) mais cette fois-ci avec des bruits de loi exponentielle :

$$\begin{cases} x_{k+1} &= (\sqrt{2} + (-1)^k)x_k + \xi_k \\ v_k &= 0.5x_k + \eta_k \end{cases}$$

avec $\xi_k = 0.1(\epsilon - 1)$ et $\eta_k = 0.1(\epsilon' - 1)$ où $\epsilon \sim \mathcal{E}(1)$ et $\epsilon' \sim \mathcal{E}(1)$.

La Figure 2.2 nous montre l'application de l'algorithme du filtre de Kalman. Nous remarquerons qu'après quelques étapes nous obtenons une bonne reconstruction du signal.

2.9.3 Bruit suivant une loi Khi-deux

Définition 2.16. On dit que le vecteur $\xi = (T_1, T_2, \dots, T_d)$ dans \mathbb{R}^d suit une loi khi-deux, et on le note par $\epsilon \sim \mathcal{X}_m^2$ où $m = (m_1, m_2, \dots, m_d)$, si :

- pour tout i , $T_i \sim \mathcal{X}_{m_i}^2$ est une variable aléatoire du khi-deux avec un degré de liberté m_i .
- les T_i sont indépendantes.

Nous allons considérer un problème d'estimation d'un signal à l'aide du filtre de Kalman lorsque les bruits de l'état et de l'observation sont de loi khi-deux :

$$\begin{cases} x_{k+1} &= A_k x_k + \xi_k \\ v_k &= C_k x_k + \eta_k \end{cases}$$

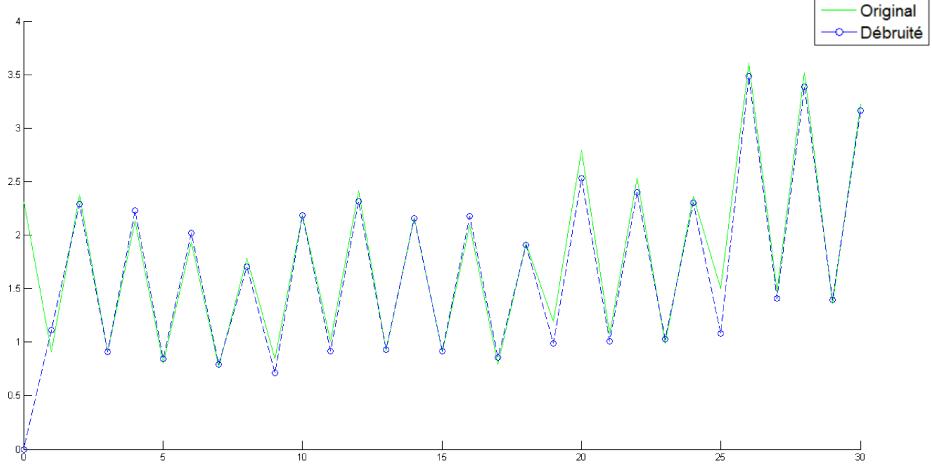


Figure 2.2 – Filtre de Kalman avec des bruits suivant une loi exponentielle où $\hat{x}_{0|0} = 0$ et $P_{0,0} = 10^2$ ($Err^{(30)} = 0.1338$)

avec $\xi_k = m_x + W_1(\epsilon - m)$ et $\eta_k = m_y + W_2(\epsilon' - m')$ où $\epsilon \sim \mathcal{X}_m^2$ et $\epsilon' \sim \mathcal{X}_{m'}^2$, nous aurons donc :

$$\begin{cases} \mathbb{E}(\xi_k) = m_x \text{ et } Var(\xi_k) = W_1 Var(\epsilon) W_1^T \\ \mathbb{E}(\eta_k) = m_y \text{ et } Var(\eta_k) = W_2 Var(\epsilon') W_2^T \end{cases}$$

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

Exemple

Reprendons l'exemple (2.4) mais cette fois-ci avec des bruits de loi khi-deux :

$$\begin{cases} x_{k+1} = (\sqrt{2} + (-1)^k)x_k + \xi_k \\ v_k = 0.5x_k + \eta_k \end{cases}$$

avec $\xi_k = \frac{0.1}{\sqrt{2}}(\epsilon - 1)$ et $\eta_k = \frac{0.1}{\sqrt{2}}(\epsilon' - 1)$ où $\epsilon \sim \mathcal{X}_1^2$ et $\epsilon' \sim \mathcal{X}_1^2$.

La Figure 2.3 nous montre l’application de l’algorithme du filtre de Kalman. Nous remarquerons qu’après quelques étapes nous obtenons une bonne reconstruction du signal.

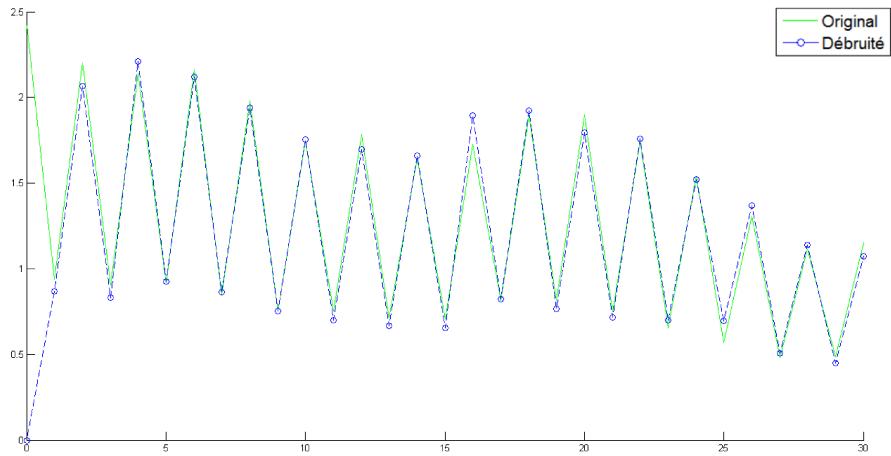


Figure 2.3 – Filtre de Kalman avec des bruits suivant une loi de khi-deux où $\hat{x}_{0|0} = 0$ et $P_{0,0} = 10^2$ ($Err^{(30)} = 0.1294$)

2.10 Conclusion

Les résultats précédents montrent que l’initialisation de $P_{0,0}$ et de $\hat{y}_{0|0}$ sont essentielles au bon fonctionnement de l’algorithme du filtre de Kalman. Est-ce que notre optimisation est optimale et est-ce qu’il y aurait une meilleure méthode pour trouver une estimation optimale de ces valeurs ? Nous essaierons de répondre à ces questions au prochain chapitre.

CHAPITRE 3

Initialisation par la méthode des moindres carrés

3.1 Introduction

Est-ce que notre initialisation est optimale ? Est-ce qu'il y aurait une meilleure méthode pour estimer ces valeurs ? Nous nous sommes posés cette question et allons essayer d'y répondre dans ce chapitre. Pour ce faire nous allons développer un système d'équations basé sur la méthode des moindres carrés et faire des comparaisons entre cette méthode et la méthode classique en utilisant des exemples.

3.2 Initialisation par la méthode des moindres carrés

Considérons le problème classique d'estimation de signal à l'aide du filtre de Kalman :

$$\begin{cases} y_{k+1} &= A_k y_k + B_k u_k + \Gamma_k \xi_k \\ w_k &= C_k y_k + D_k u_k + \eta_k. \end{cases}$$

Dans le cas classique, pour passer de $k = 0$ à $k = 1$, le processus d'initialisation requiert d'avoir $\hat{y}_{0|0} = E(y_0)$ et $P_{0,0} = P_0$. Au lieu d'utiliser ces valeurs initiales, nous pouvons utiliser l'observation initiale pour initialiser le processus.

Nous avons des observations pour $k = 1, 2, \dots$. Nous allons utiliser la première observation $w_1 = C_1 y_1 + D_1 u_1 + \eta_1$ pour estimer $\hat{y}_{1|1}$ et $P_{1,1}$ avec la méthode des moindres carrés. Dans ce cas nous n'avons pas besoin d'information pour $k = 0$. En utilisant le Théorème 1.11, nous aurons

$$\hat{y}_{1|1} = (C_1^T R_1^{-1} C_1)^+ C_1^T R_1^{-1} (w_1 - D_1 u_1).$$

et donc

$$P_{1,1} = \text{Var}(\hat{y}_{1|1} - y_1) = (C_1^T R_1^{-1} C_1)^+.$$

Nous aurons un nouvel algorithme pour le filtre de Kalman :

$$\left\{ \begin{array}{l} \hat{y}_{1|1} = (C_1^T R_1^{-1} C_1)^+ C_1^T R_1^{-1} (w_1 - D_1 u_1) \\ P_{1,1} = (C_1^T R_1^{-1} C_1)^+ \\ \text{Pour } k = 2, 3, \dots, \text{ faire} \\ P_{k,k-1} = A_{k-1} P_{k-1,k-1} A_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \\ \hat{y}_{k|k-1} = A_{k-1} \hat{y}_{k-1|k-1} + B_{k-1} u_{k-1} \\ G_k = P_{k,k-1} C_k^T (R_k + C_k P_{k,k-1} C_k^T)^{-1} \\ P_{k,k} = (I - G_k C_k) P_{k,k-1} \\ \hat{y}_{k|k} = \hat{y}_{k|k-1} + G_k (w_k - D_k u_k - C_k \hat{y}_{k|k-1}) \end{array} \right.$$

3.3 Exemple

Pour illustrer la théorie développée précédemment, nous allons considérer l'exemple bidimensionnel suivant :

$$\begin{cases} x_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} x_k + \xi_k \\ v_k = [1 \ 0] x_k + \eta_k \end{cases}$$

où

ξ_k est un vecteur aléatoire gaussien dans \mathbb{R}^2 de moyenne nulle et matrice de variance-covariance $\begin{bmatrix} 0.6 & 0 \\ 0 & 0.6 \end{bmatrix}$

η_k est une variable aléatoire gaussienne de moyenne nulle et de variance 0.6

Nous remarquons dans cet exemple que nous n'avons qu'une seule observation pour deux états ; dans ce cas le filtre de Kalman va débruiter le premier état et reconstruire le deuxième sans l'observer.

Nous remarquerons aussi dans la Figure 3.2 que l'erreur de reconstruction par la méthode des moindres carrés est plus petite que celle de la méthode classique pour les deux états.

3.4 Autres types de bruits

Nous allons voir dans cette section des applications de l'algorithme du filtre de Kalman avec des bruits pas nécessairement gaussiens et avec les deux types d'initialisations.

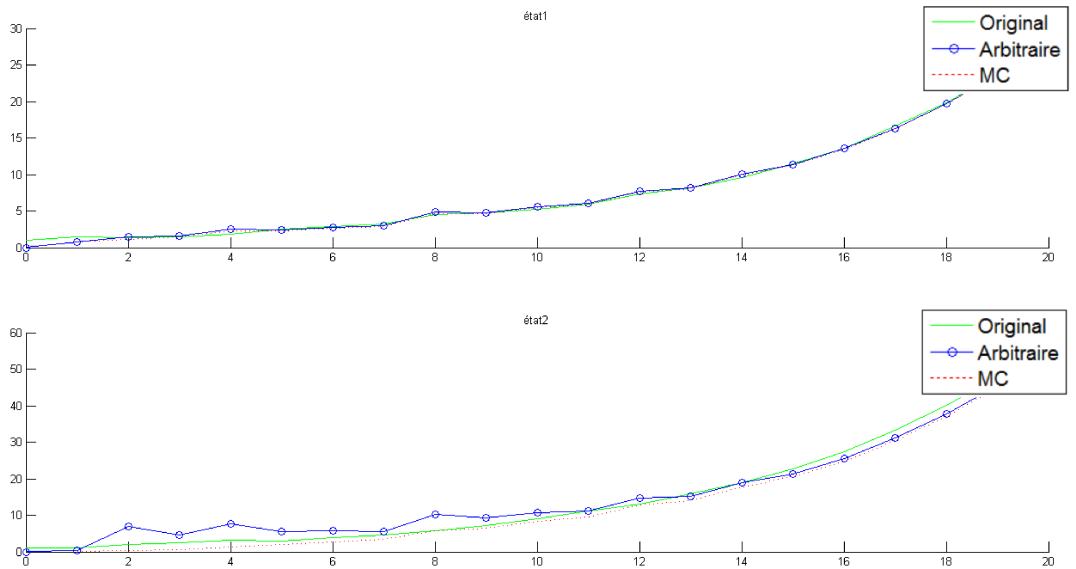


Figure 3.1 – Comparaison entre l’initialisation classique($\hat{x}_{0|0} = [0 \ 0]^T$, $P_{0,0} = 10^2 I_2$) et l’initialisation avec la méthode des moindres carrés

3.4.1 Bruit suivant une loi exponentielle

Reprenons l’exemple (2.4) mais cette fois-ci avec des bruits de loi exponentielle :

$$\begin{cases} x_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} x_k + \xi_k \\ v_k = \begin{bmatrix} 1 & 0 \end{bmatrix} x_k + \eta_k \end{cases}$$

Avec $\xi_k = \begin{bmatrix} 0.1(\epsilon - 1) & 0 \\ 0 & 0.1(\epsilon' - 1) \end{bmatrix}$ et $\eta_k = 0.1(\epsilon'' - 1)$ où $\epsilon \sim \mathcal{E}(1)$, $\epsilon' \sim \mathcal{E}(1)$ et $\epsilon'' \sim \mathcal{E}(1)$.

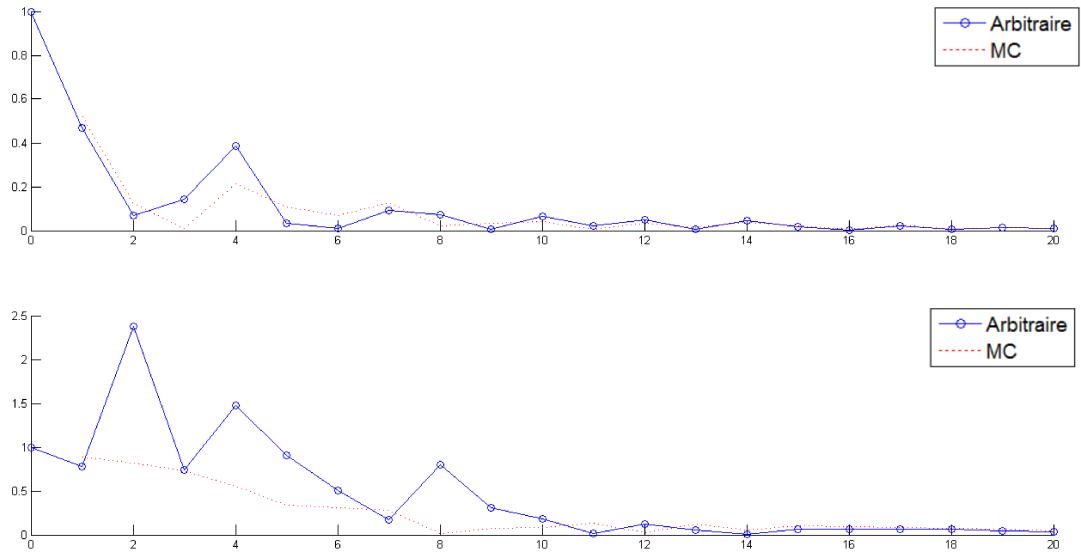


Figure 3.2 – Erreur de reconstruction du signal avec initialisation classique et initialisation par la méthode des moindres carrés dans un cas multidimensionnel.

La Figure 3.3 nous montre l’application de l’algorithme du filtre de Kalman. Nous remarquerons qu’après quelques étapes nous obtenons une bonne reconstruction du signal, nous remarquerons aussi que dans ce cas précis l’initialisation par les moindres carrés donne une meilleure estimation.

3.4.2 Bruit suivant une loi khi-deux

Reprendons l’exemple (2.4) mais cette fois-ci avec des bruits de loi khi-deux :

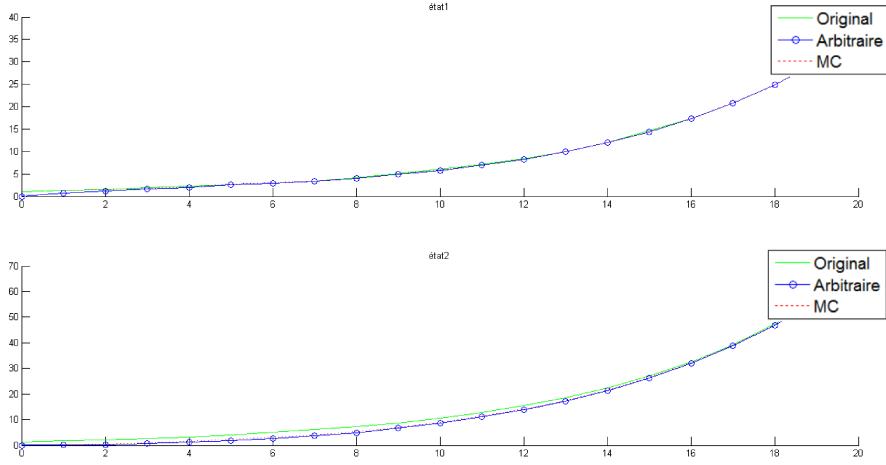


Figure 3.3 – Reconstruction du signal avec l’initialisation classique ($\hat{x}_{0|0} = [0 \ 0]^T$, $P_{0,0} = 10^2 I_2$) et l’initialisation avec la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi exponentielle.

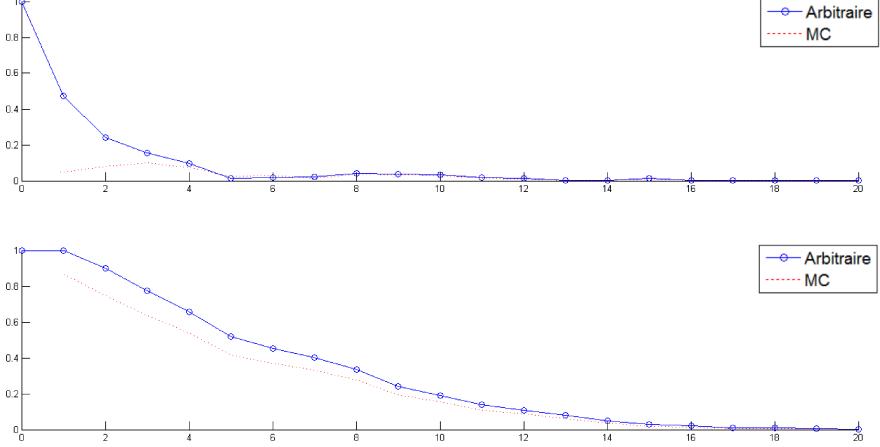


Figure 3.4 – Erreur de reconstruction du signal avec initialisation classique et initialisation par la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi exponentielle.

$$\begin{cases} x_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} x_k + \xi_k \\ v_k = [1 \ 0] x_k + \eta_k \end{cases}$$

avec $\xi_k = \begin{bmatrix} \frac{0.1}{\sqrt{2}}(\epsilon - 1) & 0 \\ 0 & \frac{0.1}{\sqrt{2}}(\epsilon' - 1) \end{bmatrix}$ et $\eta_k = \frac{0.1}{\sqrt{2}}(\epsilon'' - 1)$ où $\epsilon \sim \mathcal{X}_1^2$, $\epsilon' \sim \mathcal{X}_1^2$ et $\epsilon'' \sim \mathcal{X}_1^2$.

La Figure 3.5 nous montre l’application de l’algorithme du filtre de Kalman. Nous remarquerons qu’après quelques étapes nous obtenons une bonne reconstruction du signal, nous remarquerons aussi que l’initialisation par les moindres carrés donne une meilleure estimation.

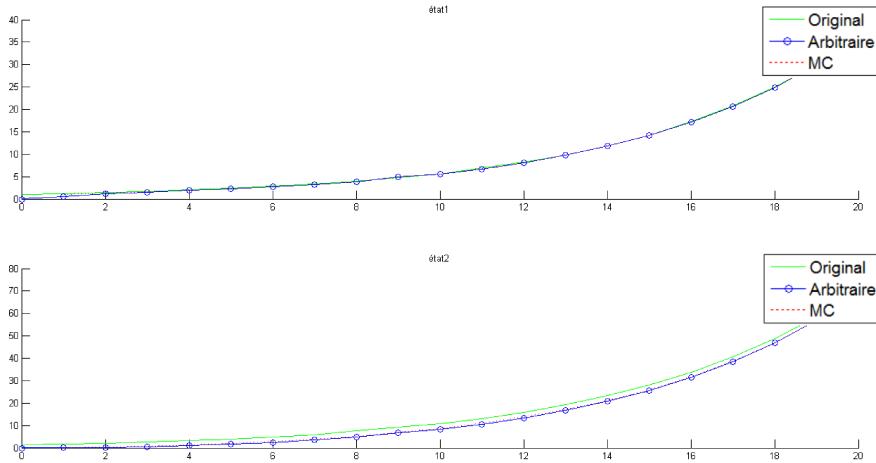


Figure 3.5 – Reconstruction du signal avec l’initialisation classique ($\hat{x}_{0|0} = [0 \ 0]^T$, $P_{0,0} = 10^2 I_2$) et l’initialisation avec la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi khi-deux.

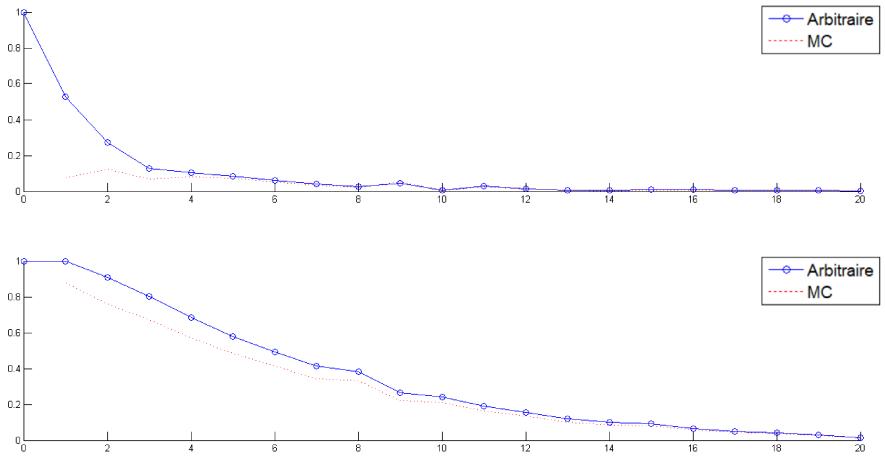


Figure 3.6 – Erreur de reconstruction du signal avec initialisation classique et initialisation par la méthode des moindres carrés dans un cas multidimensionnel avec des bruits suivant une loi khi-deux.

3.5 Conclusion

Nous nous sommes intéressés dans ce chapitre au problème d’initialisation du filtre de Kalman, nous avons constaté que la méthode des moindres carrés est efficace et reste pour la plupart du temps plus préférable à une initialisation arbitraire.

CHAPITRE 4

Filtre de Kalman avec bruit de moyenne inconnue

4.1 Introduction

Nous allons proposer dans ce chapitre une méthode pour reconstruire le signal à l'aide du filtre de Kalman avec des bruits de moyennes inconnues non-nulles. Nous allons modifier le filtre de façon à reconstruire le signal et estimer les moyennes inconnues des bruits en même temps [3, 14].

4.2 Modèle avec moyennes connues

Prenons le filtre que nous avons développé au Chapitre 2 qui nous permet d'estimer le signal, mais cette fois-ci avec des bruits non centrés mais de moyennes connues :

$$\begin{cases} x_{k+1} &= A_k x_k + \Gamma_k \xi_k \\ v_k &= C_k x_k + \eta_k \end{cases}$$

avec ξ_k et η_k des vecteurs aléatoires de moyennes connues m_x et m_v et de matrices de variance covariance Q_k et R_k respectivement.

Nous allons encore une fois donner l'estimation $\hat{x}_{k|k}$ de l'état x_k , mais en prenant en compte les moyennes des bruits de l'état et de l'observation [3, 12], ce qui nous donne :

$$\left\{ \begin{array}{l} \text{Utiliser (IN) à } k_0 \text{ pour avoir} \\ \\ \hat{x}_{k_0|k_0} \quad \quad \quad \text{et } P_{k_0,k_0} \\ \\ \text{Pour } k = k_0 + 1, k_0 + 2, \dots, \text{ faire} \\ \\ P_{k,k-1} = A_{k-1} P_{k-1,k-1} A_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \\ \hat{x}_{k|k-1} = A_{k-1} \hat{x}_{k-1|k-1} + \Gamma_{k-1} m_x \\ G_k = P_{k,k-1} C_k^T (R_k + C_k P_{k,k-1} C_k^T)^{-1} \\ P_{k,k} = (I - G_k C_k) P_{k,k-1} \\ \hat{x}_{k|k} = \hat{x}_{k|k-1} + G_k (v_k - (C_k \hat{x}_{k|k-1} + m_v)) \end{array} \right. \quad (4.1)$$

avec

$$(IN) \left\{ \begin{array}{ll} \text{processus classique : } & k_0 = 0 \\ & \begin{cases} \hat{x}_{0|0} = E(x_0) \\ P_{0,0} = E[(x_0 - \hat{x}_{0|0})(x_0 - \hat{x}_{0|0})^T] \end{cases} \\ \\ \text{Moindres carrés : } & k_0 = 1 \\ & \begin{cases} \hat{x}_{1|1} = (C_1^T R_1^{-1} C_1)^+ C_1^T R_1^{-1} v_1 \\ P_{1,1} = (C_1^T R_1^{-1} C_1)^+ \end{cases} \end{array} \right.$$

Cette méthode nous donne une bonne estimation dans le cas où les moyennes sont connues.

Pour illustrer la théorie développée précédemment, nous allons considérer l'exemple simple suivant avec cette fois-ci les moyennes des bruits connues et non-nulles :

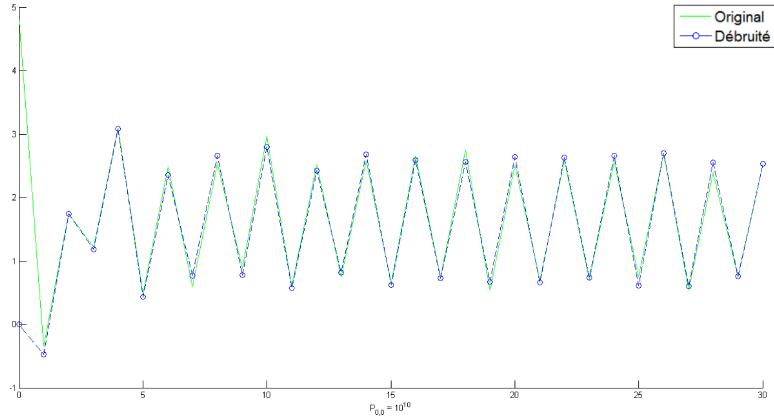
$$\begin{cases} x_{k+1} &= (0.2 + (0.7 * (-1)^k))x_k + \xi_k \\ v_k &= 0.5x_k + \eta_k \end{cases} \quad (4.2)$$

Avec des bruits gaussiens $\xi_k \sim N(2, (0.3)^2)$ et $\eta_k \sim N(5, (0.3)^2)$.

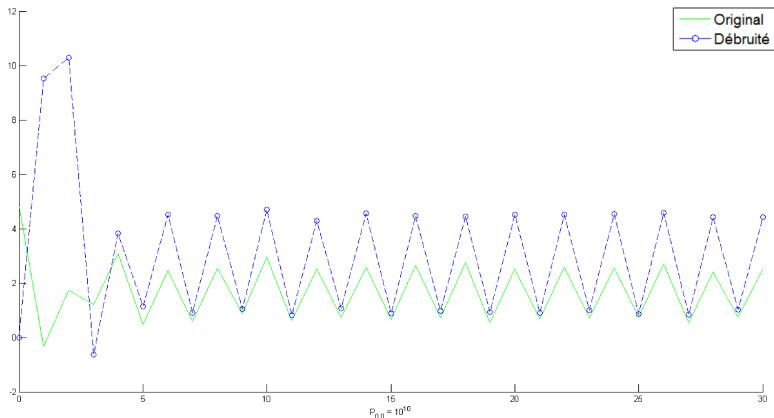
Nous obtenons un bon résultat en utilisant l'algorithme (4.1) avec les vraies moyennes ($m_x = 2$ et $m_v = 5$), nous voyons bien que l'estimation reconstruit très bien le signal (Figure 4.1a). Par contre si nous utilisons des moyennes arbitraires, comme dans la Figure 4.1b, où $m_x = 0$ et $m_v = 0$, nous constatons que le signal n'est pas du tout bien reconstruit. Nous pouvons donc en conclure que l'algorithme (4.1) n'est efficace que si nous connaissons exactement les valeurs des moyennes des bruits.

Comparons maintenant dans ce cas précis l'initialisation classique avec l'initialisation par la méthode des moindres carrés. Cela nous donne le résultat résumé à la Figure 4.2. Nous remarquerons ici que globalement l'erreur de reconstruction avec initialisation par la méthode des moindres carrés est plus petite que celle de l'initialisation classique.

Nous allons proposer dans la partie suivante une technique pour reconstruire notre signal tout en estimant les moyennes des bruits.



(a) $m_x = 2$ et $m_v = 5$



(b) $m_x = 0$ et $m_v = 0$

Figure 4.1 – Filtre de Kalman avec moyennes des bruits non-nulles et connues ($\hat{x}_{0|0} = 0$ et $P_{0,0} = 10^2$)

4.3 Modèle avec moyennes inconnues

Le but de cette section va être de reconstruire le signal tout en estimant les moyennes inconnues m_x et m_v . Considérons le système suivant :

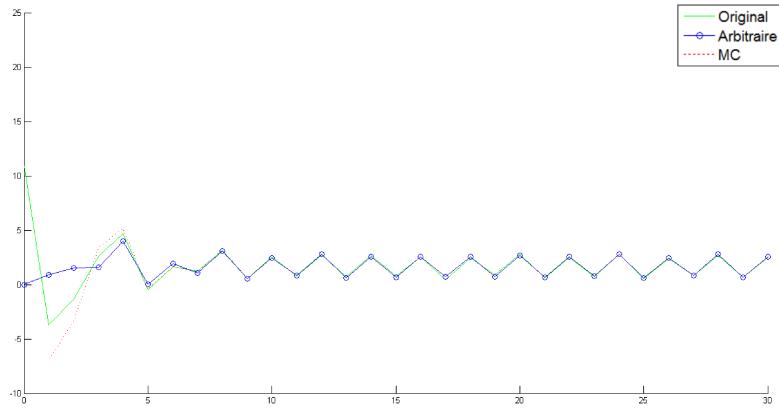


Figure 4.2 – Comparaison entre l’initialisation classique ($\hat{x}_{0|0} = 0$ et $P_{0,0} = 10^2$) et l’initialisation par la méthode des moindres carrés.

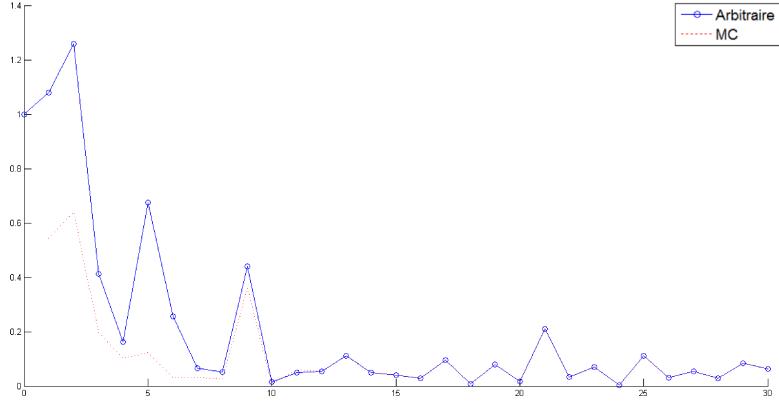


Figure 4.3 – Erreur de reconstruction du signal avec initialisation classique et initialisation par la méthode des moindres carrés.

$$\begin{cases} x_{k+1} &= A_k x_k + \Gamma_k m_x + \Gamma_k \tilde{\xi}_k \\ v_{k+1} &= C_{k+1} x_{k+1} + m_v + \tilde{\eta}_{k+1} \end{cases}$$

avec $\tilde{\xi}_k$ et $\tilde{\eta}_k$ des bruits de moyennes nulles et de variance R_k et Q_k respectivement . On

modifie la dynamique du système en ajoutant les deux équations suivantes :

$$\begin{cases} m_{x,k+1} = m_{x,k} + \varepsilon_k \\ m_{v,k+1} = m_{v,k} + \varsigma_k \end{cases}$$

Avec $\varepsilon_k \sim N(0, \Sigma_{x,k})$ et $\varsigma_k \sim N(0, \Sigma_{v,k})$.

Nous pouvons donc réécrire le système :

$$\begin{cases} x_{k+1} = A_k x_k + \Gamma_k m_{x,k} + \Gamma_k \tilde{\xi}_k \\ m_{x,k+1} = m_{x,k} + \varepsilon_k \\ m_{v,k+1} = m_{v,k} + \varsigma_k \\ v_{k+1} = C_{k+1} x_{k+1} + m_{v,k+1} + \tilde{\eta}_{k+1} \end{cases}$$

et donc

$$\begin{cases} \begin{bmatrix} x_{k+1} \\ m_{x,k+1} \\ m_{v,k+1} \end{bmatrix} = \begin{bmatrix} A_k & \Gamma_k & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ m_{x,k} \\ m_{v,k} \end{bmatrix} + \begin{bmatrix} \Gamma_k & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \tilde{\xi}_k \\ \varepsilon_k \\ \varsigma_k \end{bmatrix} \\ v_{k+1} = [C_{k+1} \ 0 \ I] \begin{bmatrix} x_{k+1} \\ m_{x,k+1} \\ m_{v,k+1} \end{bmatrix} + \tilde{\eta}_{k+1} \end{cases}$$

En posant $z_k = \begin{bmatrix} x_k \\ m_{x,k} \\ m_{v,k} \end{bmatrix}$, $\tilde{A}_k = \begin{bmatrix} A_k & \Gamma_k & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$, $\tilde{\Gamma}_k = \begin{bmatrix} \Gamma_k & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$, $\tilde{C}_{k+1} = [C_{k+1} \ 0 \ I]$, $\tilde{\epsilon}_{z,k} = \begin{bmatrix} \tilde{\xi}_k \\ \varepsilon_k \\ \varsigma_k \end{bmatrix}$, $\tilde{Q}_k = \begin{bmatrix} Q_k & 0 & 0 \\ 0 & \Sigma_{x,k} & 0 \\ 0 & 0 & \Sigma_{v,k} \end{bmatrix}$ et $\tilde{R}_k = R_k$ nous obtenons le système d'équations suivant :

$$\begin{cases} z_{k+1} = \tilde{A}_k z_k + \tilde{\Gamma}_k \tilde{\epsilon}_{z,k} \\ v_{k+1} = \tilde{C}_{k+1} z_{k+1} + \tilde{\eta}_{k+1} \end{cases}$$

Nous aurons le nouvel algorithme :

$$\left\{ \begin{array}{l} \text{Utiliser (IN) à } k_0 \text{ pour avoir} \\ \\ \hat{z}_{k_0|k_0} \quad \text{et} \quad \tilde{P}_{k_0,k_0} \\ \\ \text{Pour } k = k_0 + 1, k_0 + 2, \dots, \text{ faire} \\ \\ \tilde{P}_{k,k-1} = \tilde{A}_{k-1} \tilde{P}_{k-1,k-1} \tilde{A}_{k-1}^T + \tilde{\Gamma}_k \tilde{Q}_{k-1} \tilde{\Gamma}_k^T \\ \\ \hat{z}_{k|k-1} = \tilde{A}_{k-1} \hat{z}_{k-1|k-1} \\ \\ \tilde{G}_k = \tilde{P}_{k,k-1} \tilde{C}_k^T (\tilde{R}_k + \tilde{C}_k \tilde{P}_{k,k-1} \tilde{C}_k^T)^{-1} \\ \\ \tilde{P}_{k,k} = (I - \tilde{G}_k \tilde{C}_k) \tilde{P}_{k,k-1} \\ \\ \hat{z}_{k|k} = \hat{z}_{k|k-1} + \tilde{G}_k (v_k - (\tilde{C}_k \hat{z}_{k|k-1})) \end{array} \right. \quad (4.3)$$

avec

$$\left\{ \begin{array}{ll} \text{processus classique : } & k_0 = 0 \\ & \left\{ \begin{array}{l} \hat{z}_{0|0} = \begin{bmatrix} E(x_0) \\ m_{x,0} \\ m_{v,0} \end{bmatrix} \\ \tilde{P}_{0,0} = \begin{bmatrix} P_0 & 0 & 0 \\ 0 & P_{x,0} & 0 \\ 0 & 0 & P_{v,0} \end{bmatrix} \end{array} \right. \\ \\ \text{Moindres carrés : } & k_0 = 1 \\ & \left\{ \begin{array}{l} \hat{z}_{1|1} = (\tilde{C}_1^T \tilde{R}_1^{-1} \tilde{C}_1)^+ \tilde{C}_1^T \tilde{R}_1^{-1} v_1 \\ \tilde{P}_{1,1} = (\tilde{C}_1^T \tilde{R}_1^{-1} \tilde{C}_1)^+ \end{array} \right. \end{array} \right. \quad (\text{IN})$$

Exemple :

Reprenons l'exemple (4.2), mais cette fois-ci avec des bruits de moyennes inconnues :

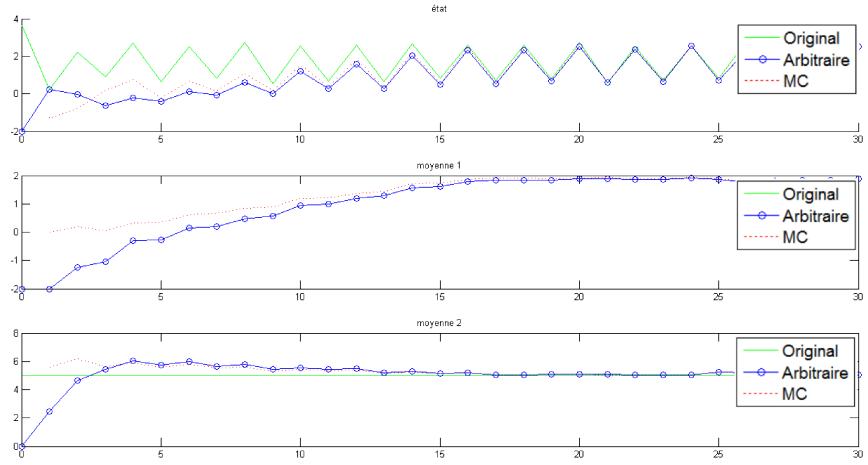


Figure 4.4 – Reconstruction du signal et des moyennes des bruits avec initialisation classique($\hat{x}_{0|0} = [-2 \ -2 \ 0]^T$ et $P_{0,0} = 10^2 I_3$) et initialisation par la méthode des moindres carrés.

Nous voyons bien dans la Figure 4.4 que dans les deux cas d'initialisation le signal converge vers les vraies valeurs, mais nous remarquerons aussi que la méthode des moindres carrés est plus efficace que la méthode classique.

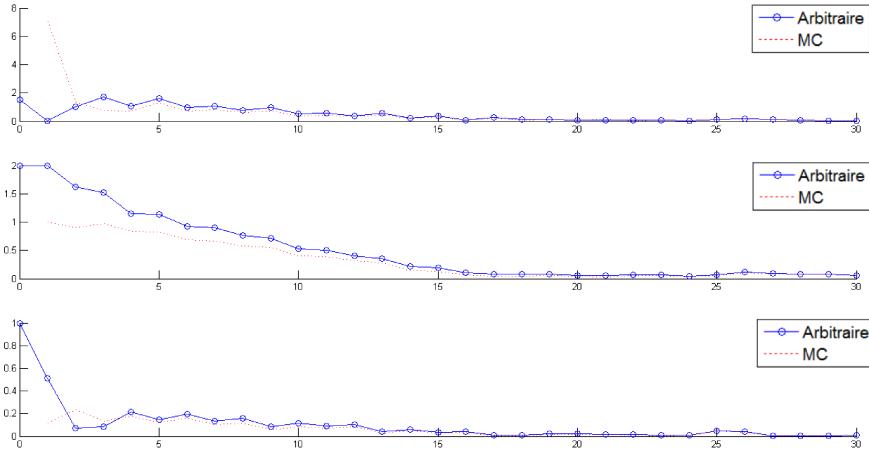


Figure 4.5 – Erreur de reconstruction du signal et des moyennes des bruits avec initialisation classique et initialisation par la méthode des moindres carrés.

Prenons un autre exemple avec cette fois-ci deux états

$$\begin{cases} x_{k+1} = \begin{bmatrix} 1.01 & 0.1 \\ 0.2 & 1.1 \end{bmatrix} x_k + \xi_k \\ v_k = [1 \ 0] x_k + \eta_k \end{cases}$$

où

ξ_k est un vecteur aléatoire gaussien dans \mathbb{R}^2 de moyenne $(5, 7)$ et matrice de variance-covariance $\begin{bmatrix} 0.6 & 0 \\ 0 & 0.6 \end{bmatrix}$

η_k est une variable aléatoire gaussienne de moyenne 10 et de variance 0.6

Nous aurons ici à reconstruire le signal des deux états, les deux moyennes des états et enfin la moyenne de l'observation.

Analysons la Figure 4.6 qui représente la reconstruction du signal et des moyennes des bruits avec initialisation classique et initialisation par la méthode des moindres carrés. Nous remarquerons que dans les deux cas, le signal est bien reconstruit, mais la méthode

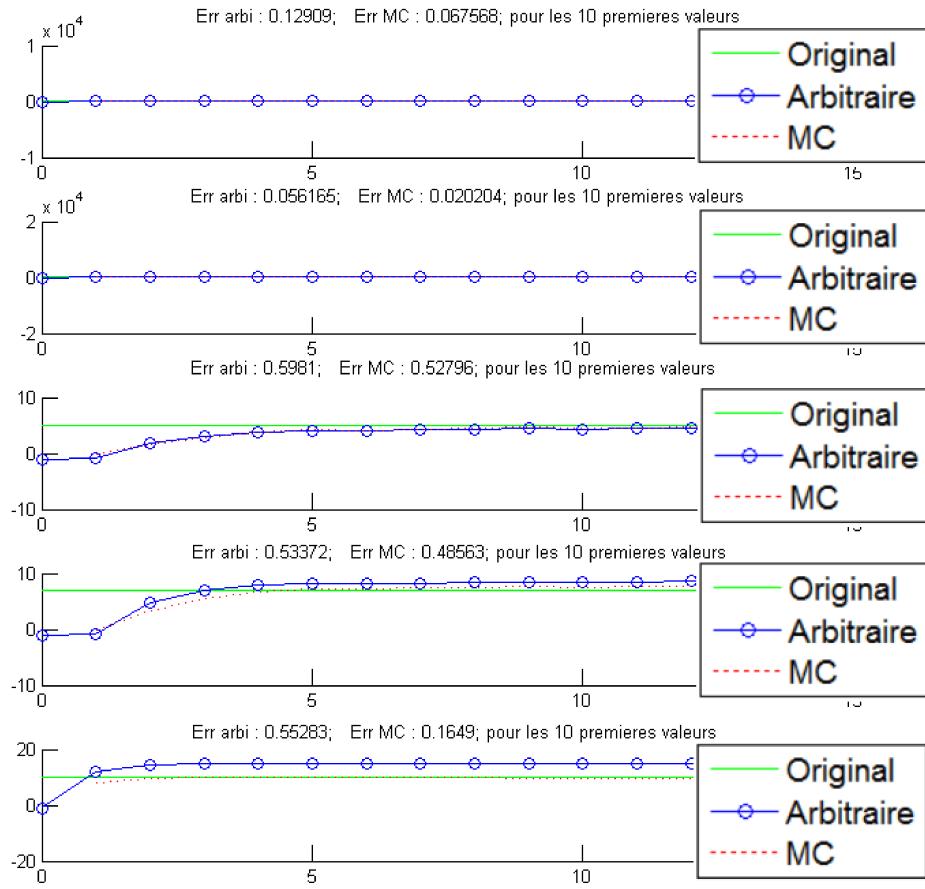


Figure 4.6 – Reconstruction du signal et des moyennes des bruits avec initialisation classique($\hat{x}_{0|0} = [0 \ 0 \ 0 \ 0]^T$ et $P_{0,0} = 10^2 I_5$) et initialisation par la méthode des moindres carrés dans un cas multidimensionnel.

des moindres carrés est plus efficace que la méthode classique, la Figure 4.7 représente les erreurs de reconstruction du signal et des moyennes des bruits avec les deux types d'initialisation.

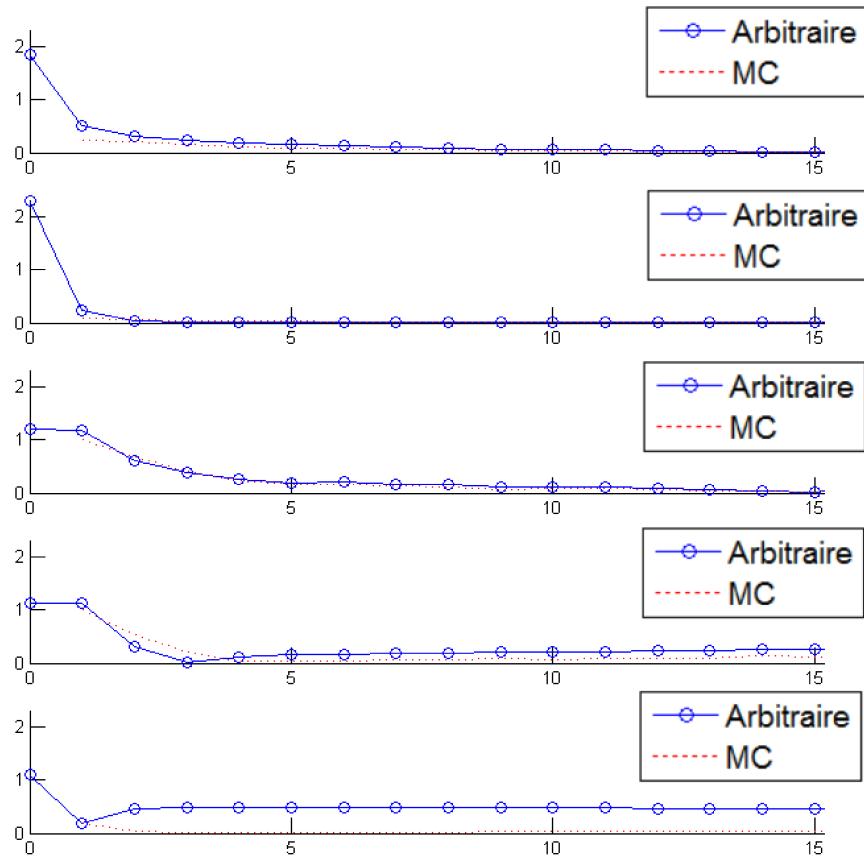


Figure 4.7 – Erreur de reconstruction du signal et des moyennes des bruits avec initialisation classique et initialisation par la méthode des moindres carrés dans un cas multidimensionnel.

4.4 Autres types de bruits

Nous allons voir dans cette section des applications de l'algorithme du filtre de Kalman avec des bruits pas nécessairement gaussiens, avec les deux types d'initialisations et la reconstruction des moyennes inconnues.

4.4.1 Bruit suivant une loi exponentielle

Reprendons l'exemple (4.2), mais cette fois-ci avec des bruits de loi exponentielle :

$$\begin{cases} x_{k+1} &= (0.2 + (0.7 * (-1)^k))x_k + \xi_k \\ v_k &= 0.5x_k + \eta_k \end{cases}$$

Avec $\xi_k = 2 + 0.6(\epsilon - 1)$ et $\eta_k = 5 + 0.6(\epsilon' - 1)$ où $\epsilon \sim \mathcal{E}(1)$ et $\epsilon' \sim \mathcal{E}(1)$.

La Figure 4.8 nous montre l'application de l'algorithme du filtre de Kalman dans les cas où les moyennes sont connues (Figure 4.8a) et inconnues (Figure 4.8b), avec les bruits suivant une loi exponentielle. Nous remarquerons après quelques étapes que nous obtenons une bonne reconstruction du signal. La Figure 4.9 représente l'erreur de reconstruction avec moyennes des bruits connues et inconnues, nous remarquerons que l'initialisation par les moindres carrés donne une meilleure estimation.

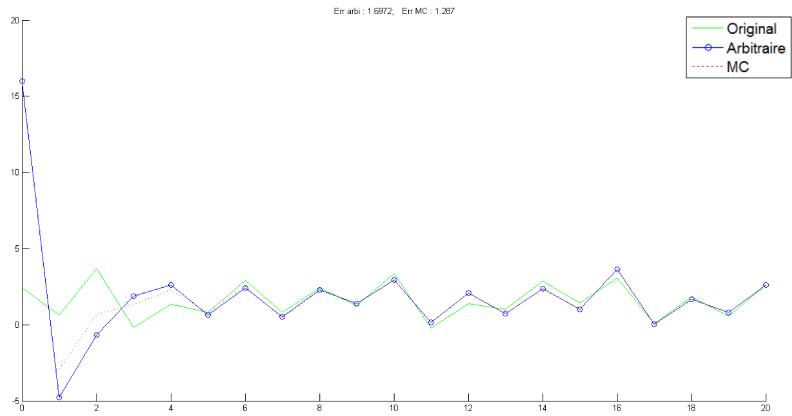
4.4.2 Bruit suivant une loi khi-deux

Reprendons l'exemple (4.2), mais cette fois-ci avec des bruits de loi khi-deux :

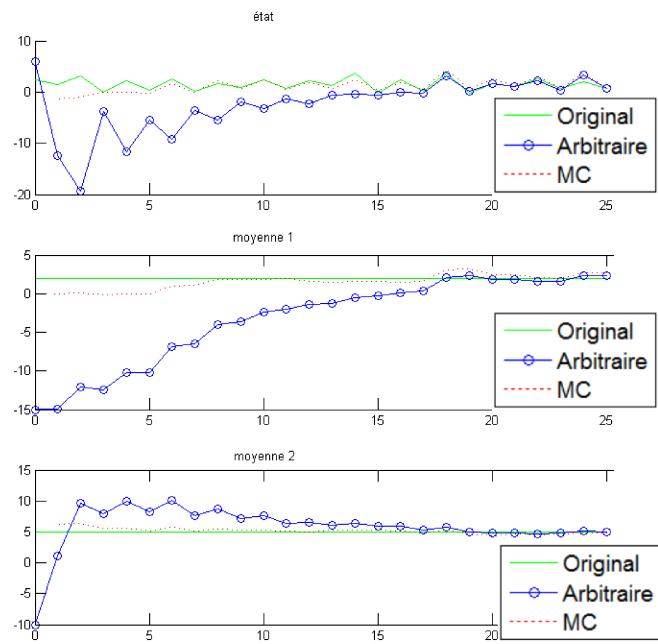
$$\begin{cases} x_{k+1} &= (0.2 + (0.7 * (-1)^k))x_k + \xi_k \\ v_k &= 0.5x_k + \eta_k \end{cases}$$

Avec $\xi_k = 2 + \frac{0.6}{\sqrt{2}}(\epsilon - 1)$ et $\eta_k = 5 + \frac{0.6}{\sqrt{2}}(\epsilon' - 1)$ où $\epsilon \sim \mathcal{X}_1^2$ et $\epsilon' \sim \mathcal{X}_1^2$.

La Figure 4.10 nous montre l'application de l'algorithme du filtre de Kalman dans les cas où les moyennes sont connues (Figure 4.10a) et inconnues (Figure 4.10b), avec les bruits suivant une \mathcal{X}_1^2 . Nous remarquerons après quelques étapes que nous obtenons une bonne reconstruction du signal. La Figure 4.11 représente l'erreur de reconstruction avec

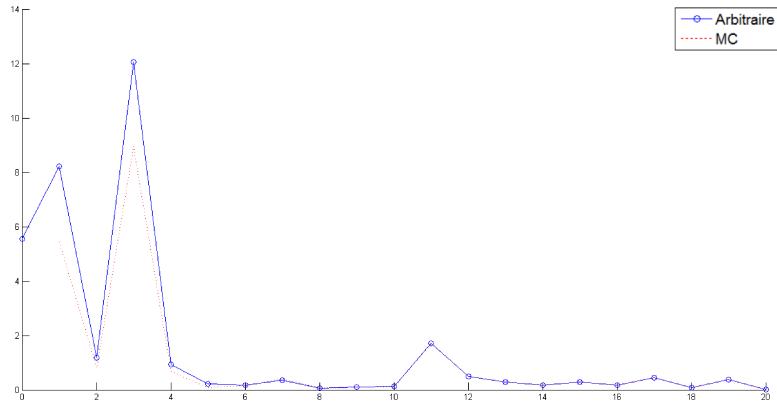


(a) moyennes connues $m_x = 2$ et $m_v = 5$

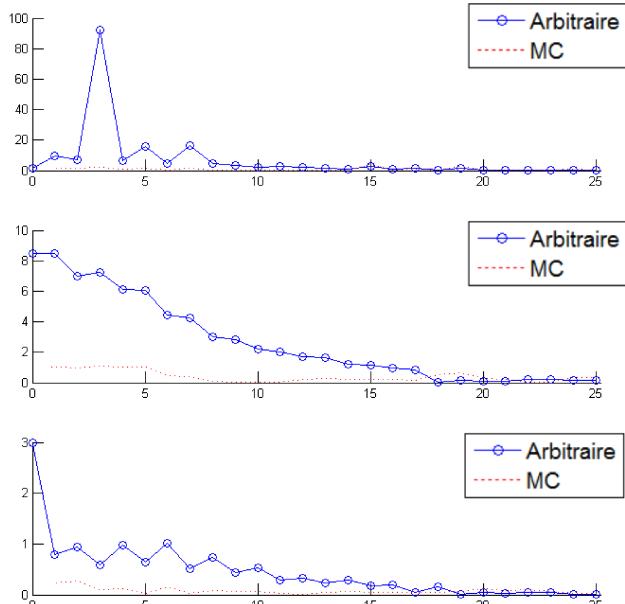


(b) moyennes inconnues

Figure 4.8 – Filtre de Kalman avec moyennes des bruits connues et inconnues suivant une loi exponentielle



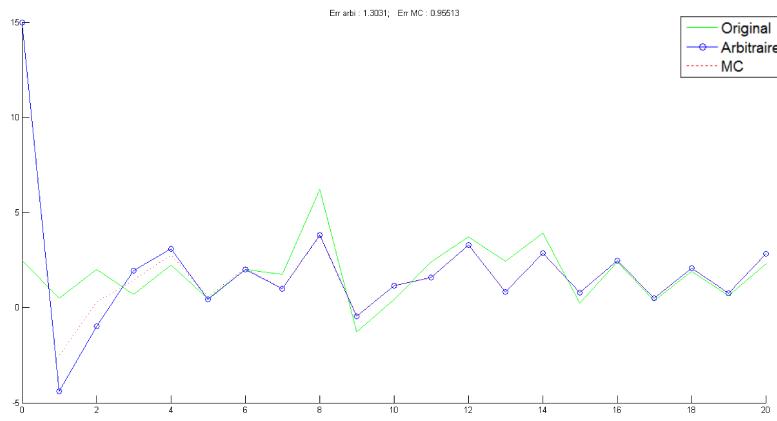
(a) moyennes connues $m_x = 2$ et $m_v = 5$



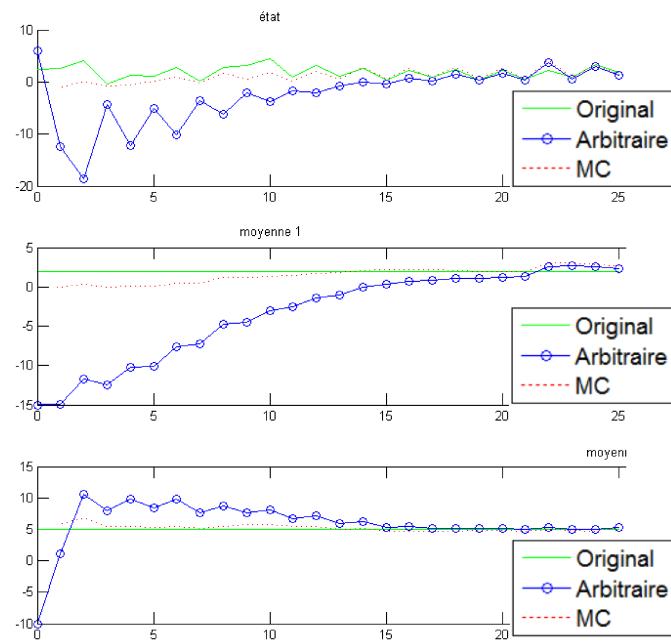
(b) moyennes inconnues

Figure 4.9 – Erreur de reconstruction avec moyennes des bruits connues et inconnues suivant une loi exponentielle

moyennes des bruits connues et inconnues, nous remarquerons que l'initialisation par les moindres carrés donne une meilleure estimation.

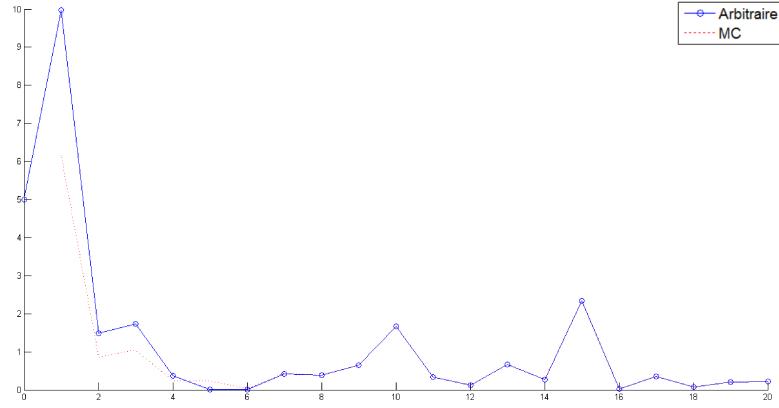


(a) moyennes connues $m_x = 2$ et $m_v = 5$

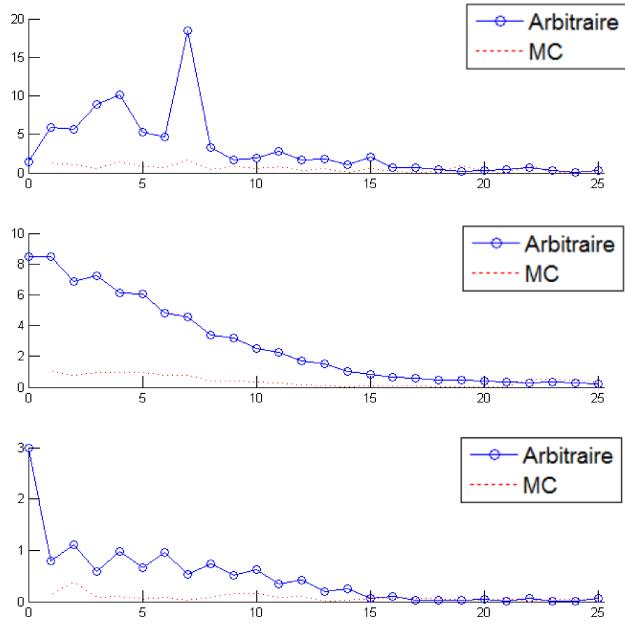


(b) moyennes inconnues

Figure 4.10 – Filtre de Kalman avec moyennes des bruits connues et inconnues suivant une loi khi-deux



(a) moyennes connues $m_x = 2$ et $m_v = 5$



(b) moyennes inconnues

Figure 4.11 – Erreur de reconstruction avec moyennes des bruits connues et inconnues suivant une loi khi-deux

4.5 Conclusion

Nous avons analysé dans ce chapitre le problème de filtrage à l'aide du filtre de Kalman dans le cas où les moyennes des bruits sont connues et inconnues. Nous avons proposé une méthode qui estime à la fois le signal et les moyennes inconnues avec deux types d'initialisations.

CONCLUSION

L'estimation séquentielle d'un processus caché à partir d'un processus observé continue d'être l'objet d'une littérature scientifique abondante. Nous nous sommes concentrés dans ce mémoire sur le filtre de Kalman et avons constaté sa puissance concernant la reconstruction du signal à partir de mesures bruitées. L'un des points forts de ce filtre est sa récursivité, en effet il peut estimer en temps réel les paramètres de l'état et cela grâce à sa capacité de prédiction des paramètres et de corrections de l'erreur. Un autre point fort, la faculté de donner des estimations correctes malgré des erreurs d'approximations, cela grâce au terme d'imprécision qui est intégré directement dans le modèle.

Nous avons proposé une méthode de reconstruction du signal dans le cas où nous devions estimer la moyenne des bruits apparus dans l'équation d'état et dans l'équation d'observation simultanément. Nous avons aussi pu reconstruire un signal dont le bruit ne suit pas forcément une loi gaussienne. 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. Nous avons aussi développé une nouvelle technique pour initialiser notre filtre qui repose sur le principe des moindres carrés.

Bibliographie

- [1] A. Benavoli, L. Chisci, and A. Farina. Fibonacci Sequence, Golden Section, Kalman Filter and Optimal Control. *Signal Processing*, 89 :1483–1488, 2009.
- [2] J. S. Chipman. Linear Restrictions, Rank Reduction, and Biased Estimation in Linear Regression. *Linear Algebra and its Applications*, 289 :55–74, 1998.
- [3] C. K. Chui and G. Chen. *Kalman Filtering with Real-Time Applications*. Springer, Berlin, 2009.
- [4] S. A. Van de Geer. Least Square Estimation. *Encyclopedia of Statistics in Behavioral Science*, 2 :1041–1045, 2005.
- [5] R. Deriche, J. Calder, and M. Descoteaux. *Optimal Real-Time Q-Ball Imaging Using Regularized Kalman Filtering with Incremental Orientation Sets*. 2009.
- [6] M. S. Grewal and A. P. Andrews. *Kalman Filtering : Theory and Practice Using Matlab*. John Wiley and Sons, New York, 2001.
- [7] J. Humpherys, P. Redd, and J. West. A Fresh Look at the Kalman Filter. *SIAM Review*, 54(4) :801–823, 2012.
- [8] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(series D) :35–45, 1960.
- [9] P. Kim. *Kalman Filter for Beginners with MATLAB Examples*. A-JIN Publisher Company, 2011.

- [10] D. G. Luenberger. Optimization by Vector Spaces Methods. John Wiley and Sons, New York, 1969.
- [11] J. G. Malcolm, M. E. Shenton, and Y. Rathi. Filtered Multitensor Tractography. *IEEE Transaction on Medical Imaging*, 29(9) :1664–1675, 2010.
- [12] I. B. Rhodes. A Tutorial Introduction to Estimation and Filtering. *IEEE Transactions on Automatic Control*, 16(6) :688–706, 1971.
- [13] S. M. Ross. *Initiation aux probabilités*. Presses Polytechniques et Universitaires Romandes, Lausane, 2003.
- [14] B. Saidani. Filtre de Kalman discret pour l'estimation des moyennes de bruits blancs. Mémoire de maîtrise, Département de Mathématiques, Université de Sherbrooke, 2012. (<http://savoirs.usherbrooke.ca/handle/11143/5748>).
- [15] L. R. Weill and P. N. De Land. *The Kalman Filter : an Introduction to the Mathematics of Linear Least Mean Square Recursive Estimation*, volume 17. Fullerton, 1986.