

Filtrage Optimal de Volterra à Horizon Infini :

Application aux Systèmes Bilineaires.

André MONIN, Gérard SALUT.

Laboratoire d'Automatique et d'Analyse des Systèmes.
Centre National de la Recherche Scientifique.
7 Avenue du Colonel Roche 31077 TOULOUSE Cedex.

RÉSUMÉ

Une nouvelle technique de filtrage optimal non-linéaire des systèmes à réalisation bilinéaire est développée. Celle-ci réalise la projection (au sens du minimum de variance) de l'état à estimer sur une classe fermée de séries de Volterra finies à noyaux séparables. L'estimateur à horizon infini obtenu est réalisé par un système bilinéaire nilpotent de dimension finie. Les paramètres du filtre sont calculables hors ligne à partir de ceux du modèle du système bilinéaire à estimer. Ils font appel à la connaissance des moments d'ordre égal à deux fois le degré de la fonctionnelle de Volterra désirée. Deux applications sont traitées à titre d'illustration : l'estimation de la covariance du bruit dynamique d'un système linéaire (problème connu pour ne pas avoir de solution par l'approche linéaire) - l'estimation de l'état d'un système linéaire à perturbation non-gaussienne (de type sauts à occurrences poissonniennes) où l'amélioration des performances est très nettement mise en évidence par augmentation du degré du filtre (linéaire, linéaire+quadratique, linéaire+quadratique+cubique).

ABSTRACT

A new non-linear filtering technique for bilinear systems is developed. It produces the projection onto the closed class of finite Volterra series with separable kernels of arbitrary degree k . Such an optimal estimator is itself finitely realizable as a bilinear system. Its parameters are computable off line from those of the original system through higher moments of degree $2k$. Two simple illustrating example are being developed: the estimation of the covariance of the white noise driving a linear system — the filtering of a non-gaussian linear system (driven by a Poisson process).

1. Introduction.

Par définition, le filtrage optimal linéaire est incapable d'extraire les informations non-linéairement liées aux observations. C'est malheureusement le cas dans de nombreux problèmes pratiques. A titre d'exemple, citons le problème d'estimation de la covariance du bruit perturbant la dynamique d'un système linéaire connu pour être inobservable par un fonctionnelle linéaire des observations. De plus, pour un système linéaire mais non-gaussien, le filtre de Kalman est connu pour ne pas être le filtre optimal. Un filtre non-linéaire peut donc améliorer la qualité d'estimation. Dans toutes ces situations, il est nécessaire de travailler avec des fonctionnelles plus générales.

La classe la plus générale, de type probabiliste, est celle des estimateurs simplement mesurables par rapport à la tribu des observations et conduit à l'espérance conditionnelle, généralement irréalisable en dimension finie. D'autres classes, de type statistique, sont intermédiaires entre celles-ci et les fonctionnelles linéaires ; elles utilisent comme estimateur des fonctionnelles polynomiales (série de Volterra de degré fini k) dont l'optimalité est déterminée par l'autocorrélation de degré $2k$ des observations.

Il y eut plusieurs travaux dans cette direction, en automatique et traitement du signal, dont les plus connus sont ceux de Zadeh ([1]), Wiener ([2]) et Pugatchev ([3]) mais sans atteindre la

réalisabilité en dimension finie qui fit le succès des estimateurs linéaires de Wiener ([4]), Kalman ([5], [6]) et Bucy ([6]). Mentionnons de plus quelques résultats utilisant des fonctionnelles de Volterra sur un horizon fini, connus sous le nom de filtres transverses ([7], [8]). Ces filtres sont obtenus par projection orthogonale sur un passé limité des observations. Remarquons qu'une telle approche ne permet pas d'inclure comme cas particulier le filtre de Kalman qui lui travaille sur un horizon infini.

Dans le cas général, la fonctionnelle de Volterra n'est pas réalisable de par :

- la nature du système traité,
- la classe d'estimateurs utilisés.

En ce qui concerne le premier point, nous étudions la classe des systèmes à réalisation bilinéaire pour deux raisons. D'une part, celle-ci est relativement générale car dense dans celle des systèmes non-linéaires analytiques ([9], [10]) ; d'autre part, les moments de ce type de système sont récursifs en prédiction.

Concernant le deuxième point, nous considérons la classe des séries de Volterra finies à noyaux séparables, c'est à dire préservant la récursivité ([11], [12], [13], [14]). Plusieurs travaux antérieurs avaient utilisé de tels estimateurs, mais en fixant *a priori* certaines contraintes sur les noyaux ([15], [16], [17], [18]).

Ces résultats restaient donc perfectibles et nous donnons ici les noyaux optimaux en ne fixant *a priori* que degré et ordre



de séparabilité de la fonctionnelle de Volterra souhaitée. Nous déterminons ces noyaux de sorte que la série fournisse l'estimateur optimal dans cette classe au sens du minimum de variance. L'étude est effectuée pour les systèmes à temps discret. Nous montrons que pour des systèmes stochastiques à réalisation bilinéaire et observation linéaire, la mise en oeuvre de cet estimateur, c'est à dire le calcul récursif de ses paramètres est lui-même obtenu en dimension finie.

Finalement, nous présentons deux applications de cette technique : l'estimation de la covariance des bruits internes dans un système linéaire et l'estimation d'un système linéaire gouverné par des bruits non-gaussiens (ponctuels en l'occurrence).

2. Position du problème.

Considérons le système bilinéaire suivant :

$$\begin{aligned} x_t &= Fx_{t-1} + \sum_{i=2}^r (G_i x_{t-1} + g_i) w_i^1 \\ y_t &= Hx_t + w_t^1 \end{aligned} \quad (2.1)$$

L'état $x \in R^n$, w et v sont des bruits blancs discrets indépendants, l'observation $y \in R$. Nous noterons \hat{x} l'estimateur de x .

Remarquons qu'il n'y a pas de restriction à ne s'intéresser qu'à l'estimation de l'état et non à une fonction polynomiale de celui-ci. En effet, il est toujours possible de construire le système bilinéaire augmenté dans lequel le polynôme apparaît comme état de ce nouveau système.

3. Réalisabilité des séries de Volterra.

Soit y_t le signal d'observation. Soit \hat{x} le terme à réaliser. Une série de Volterra de degré d à la forme suivante :

$$\hat{x}_t = \sum_{k=1}^d \sum_{\tau_1=0}^t \dots \sum_{\tau_{k-1}=0}^{\tau_{k-1}} W_k(t, \tau_1, \dots, \tau_k) y_{\tau_1} \dots y_{\tau_k} \quad (3.1)$$

On peut montrer que cette série de Volterra est réalisable par un système bilinéaire si les noyaux W_k sont séparables ([9], [10]).

3.1. Propriété de séparabilité.

Un noyau W est dit séparable s'il peut être développé en une série finie de sommes de produits de fonctions monovariées. Ceci peut s'écrire sous la forme tensorielle suivante :

$$W(t, \tau_1, \dots, \tau_k)^{i_{k+1}} = \sum_{i_k=1}^{n_k} \dots \sum_{i_1=1}^{n_1} [K_{k+1}(t)]_{i_k}^{i_{k+1}} [K_k(\tau_k)]_{i_{k-1}}^{i_k} \dots [K_2(\tau_2)]_{i_1}^{i_2} [K_1(\tau_1)]_{i_1}^{i_1}$$

Les entiers n_1, \dots, n_k sont les différents ordres de séparabilité en les variables $(\tau_1, \tau_2), \dots, (\tau_k, t)$ respectivement ; n_{k+1} est la dimension de \hat{x} .

3.2. Réalisation dans l'espace d'état.

Proposition 1. Si les noyaux sont séparables comme dans (3.1) peut être réalisée par le système bilinéaire nilpotent (au sens de l'algèbre de Lie des champs de vecteur) suivant :

$$\begin{cases} \xi_1 = (\xi_1)_- + K_1 y & \xi_1 \in R^{n_1}, \xi_1(0) = 0 \\ \xi_2 = (\xi_2)_- + K_2 \xi_1 y & \xi_2 \in R^{n_2}, \xi_2(0) = 0 \\ \dots \\ \xi_k = (\xi_k)_- + K_k \xi_{k-1} y & \xi_k \in R^{n_k}, \xi_k(0) = 0 \\ \hat{x} = K_{k+1} \xi_k & \hat{x} \in R^{n_{k+1}} \end{cases} \quad (3.2)$$

où $\xi = \xi(t)$, $(\xi)_- = \xi(t-1)$, K_l représente le vecteur de composantes K_l^i et $K_p, \forall l = 2 \dots k$ sont par définition les matrices K_{i-1}^i . \square

Sans perte de généralité, on peut considérer que tous les ordres de séparabilité sont égaux. Avec cette hypothèse, on a :

Proposition 2. Si $\forall i, j = 1 \dots k, n_i = n_j$, le système (3.2) peut être réalisé par :

$$\begin{cases} z_1 = A_1(z_1)_- + A_0 y \\ z_2 = A_2(z_2)_- + z_1 y \\ \dots \\ z_k = A_k(z_k)_- + z_{k-1} y \\ \hat{x} = z_k \end{cases} \quad (3.3)$$

\square

Preuve : Le changement de variables suivant :

$$\forall i = 1 \dots k, z_i = K_{k+1} K_k \dots K_{i+1} \xi_i \text{ et } \hat{x} = z_k \quad (3.4)$$

donne le résultat. \blacksquare

Remarquons que dans cette représentation, les paramètres $A_i(t)$ apparaissent linéairement dans l'expression de l'estimateur \hat{x} . Ceci a pour objet de faciliter le calcul d'optimisation de ces paramètres dans la suite de l'exposé.

4. Optimisation de l'estimateur.

Nous utiliserons les notations suivantes : φ^T pour la transposée de φ et $E[\varphi]$ pour l'espérance mathématique.

Pour de nouveau simplifier l'exposé, nous nous limiterons, sans perte de généralité, à la génération d'un seul noyau de degré k .

Il s'agit de minimiser l'erreur quadratique d'estimation $J_t = E[(x_t - \hat{x}_t)^T (x_t - \hat{x}_t)]$ par rapport aux paramètres $\{A_i(\tau), \forall i = 0 \dots k, \forall \tau \geq 0\}$ et ceci à chaque pas $t \geq 0$. En vertu du principe d'optimalité, rappelons que si les paramètres sont optimaux sur les intervalles $[0, t-1]$ et $[0, t]$, ils doivent alors être optimaux sur l'intervalle $[t-1, t]$. Ceci nous amène au calcul récursif des paramètres : étant donnés les $\{A_i(\tau), \forall i = 0 \dots k, \forall \tau = 0 \dots t-1\}$, il s'agit de trouver les $\{A_i(t), \forall i = 0 \dots k\}$ qui minimisent le critère J_t . On peut remarquer que J_t est la fonction de Bellman sur l'horizon passé du problème d'optimisation posé ici.

Proposition 3. Les paramètres du filtre optimal vérifient le système d'équations linéaires suivant :

$$\forall i = 0 \dots k, \sum_{j=0}^k A_j E[(z_j)_- (z_j)_-^T y^{2k-i-j}] = E[x(z_i^T)_- y^{k-i}] \quad (4.1)$$

\square

Il apparaît donc que le calcul des paramètres fait appel à l'inversion de ce système linéaire précédé par le calcul des espérances $E[(z_j)_-(z_j)_T y^{2k-i-j}]$ et $E[x(z_j)_- y^{k-i}]$. Ces quantités sont calculables en dimension finie grâce au caractère bilinéaire du modèle de l'observation y_t .

Proposition 4. Si le système qui engendre le signal d'observation y_t est bilinéaire, alors les espérances des équations (4.1) sont réalisables en dimension finie. □

Éléments de preuve : A l'étape t , les valeurs des $A_i(t-1)$ sont supposées connues. En vertu de l'équation (3.3), il est clair que le degré de z_i est i en y (cela signifie que z_i est un polynôme de degré i des $y_\tau, \tau = 0 \dots t$). De plus, l'observation y étant linéaire en x , z_i est aussi un polynôme en x de degré i . Par conséquent, le degré de $E[(z_j)_-(z_j)_T y^{2k-i-j}]$ est égal à $2k$ en x alors que le degré de $E[x(z_j)_- y^{k-i}]$ est $k+1$. Le modèle étant bilinéaire, le degré d'un polynôme quelconque en x est conservé en prédiction. L'objet du paragraphe suivant est de donner les formules de calcul explicite des paramètres du filtre.

5. L'algorithme.

Le calcul de l'évolution des espérances étant particulièrement difficile, il est nécessaire de récrire le modèle sous une forme purement bilinéaire.

5.1. Représentation purement bilinéaire d'un système bilinéaire.

Nous utiliserons par la suite la notation tensorielle classique qui permet d'omettre le symbole de sommation lorsqu'un indice apparaît deux fois. Avec cette convention, le modèle (2.1) peut être écrit sous la forme suivante :

$$x^i = (F_j^i + G_{j,i}^i w^j) x^j + g_i^i w^i, \forall i, j = 1 \dots n, \forall l = 1 \dots r \quad (5.1)$$

Proposition 5. Le système bilinéaire (5.1) peut être représenté par le système purement bilinéaire suivant :

$$\begin{cases} x^i = \bar{G}_{j,i}^i x^j w^i \\ y = \bar{H}_i x^i w^i \end{cases} \quad \square(5.2)$$

Ajoutons une nouvelle composante au vecteur d'état x , soit $x^0 = 1$, et une nouvelle composante au vecteur de bruits w , soit $w^0 = 1$. Il est clair que l'équation (5.1) prend alors la forme annoncée. Il en va de même pour l'équation d'observation. ■

5.2. Calcul récursif des moments.

La Proposition 3. indique que le calcul des paramètres du filtre exige la connaissance des autocorrelations de l'état x au moins jusqu'à l'ordre $2k$. Par conséquent, il est nécessaire de connaître les moments du bruit w également jusqu'à l'ordre $2k$. Définissons :

$$Q^{i_1 \dots i_r} = E[w^{i_1} \dots w^{i_r}] \quad (5.3)$$

En utilisant le modèle bilinéaire (5.2) de x , d'après l'équation (4.1), le calcul des paramètres fait appel à l'actualisation à chaque pas de l'ensemble suivant d'espérances :

$$\begin{aligned} A(t) &= \{E[z_i^i z_j^j x^{k_1} \dots x^{k_{2k-i-j}}]\}_{i,j=1}^k \\ B(t) &= \{E[z_i^i x^{k_1} \dots x^{k_{2k-i+1}}]\}_{i=1}^k \end{aligned} \quad (5.4)$$

$$C(t) = \{E([x^{k_1} \dots x^{k_{2k}}], E[x^{k_1} \dots x^{k_{2k+1}}])\}$$

Les éléments de $A(t)$ s'obtiennent pour $i_1, i_2 = 1 \dots k$ et $p = 2k - i_1 - i_2$ par l'ensemble d'équations suivant :

$$E[z_{i_1}^{m_1} z_{i_2}^{m_2} x^{k_1} \dots x^{k_p}] = \sum_{p_1=0}^{i_1} \sum_{p_2=0}^{i_2} A_{p_1, q_1}^{m_1} A_{p_2, q_2}^{m_2} \quad (5.5)$$

$$E[(z_{p_1}^{q_1})_-(z_{p_2}^{q_2})_ x^{k_1} \dots x^{k_p} y^{i_1+i_2-p_1-p_2}]$$

avec, pour $q = i_1 + i_2 - p_1 - p_2$:

$$\begin{aligned} E[(z_{p_1}^{q_1})_-(z_{p_2}^{q_2})_ x^{k_1} \dots x^{k_p} y^q] &= \bar{G}_{i_1, j_1}^{k_1} \dots \bar{G}_{i_p, j_p}^{k_p} \bar{H}_{l_1, m_1} \dots \bar{H}_{l_p, m_p} \\ \bar{G}_{r_1, s_1}^{l_1} \dots \bar{G}_{r_p, s_p}^{l_p} E[(z_{p_1}^{q_1})_-(z_{p_2}^{q_2})_ x^{i_1} \dots x^{i_p} x^{r_1} \dots x^{r_p}] & Q^{j_1 \dots j_p m_1 \dots m_p s_1 \dots s_p} \end{aligned}$$

De manière analogue, les espérances de $B(t)$ et $C(t)$ sont récursives.

6. Applications.

Remarquons tout d'abord que pour le filtre de Kalman les coefficients sont calculés hors ligne. De plus, l'espérance de l'erreur quadratique de filtrage est calculable *a priori*.

Pour illustrer l'efficacité de la technique présentée, nous nous intéressons à deux applications :

- L'estimation de la covariance d'un système linéaire
- L'estimation de l'état d'un système linéaire mais à bruits non-gaussiens.

6.1. Estimation d'une covariance.

Considérons un système linéaire scalaire dont la covariance du bruit de dynamique est inconnue et fluctuante. Si on définit l'écart type σ de ce bruit comme un nouvel état, le système complet peut être représenté par le système bilinéaire suivant :

$$\begin{cases} \sigma_t = \sigma_{t-1} + \varepsilon w_t^1 \\ x_t = f x_t + \sigma_t w_t^2 \end{cases} \quad y_t = x_t + v_t \quad (6.1)$$

où w^1, w^2 sont des bruits blancs normalisés et indépendants alors que v est un bruit blanc indépendant de variance R . Les fluctuations de σ sont réglées par le paramètre ε . Il est nécessaire de considérer des fonctionnelles au moins quadratiques des observations pour estimer la variance de ce bruit $\sigma^2 = Q$. Nous nous limiterons ici à un estimateur quadratique de Q .

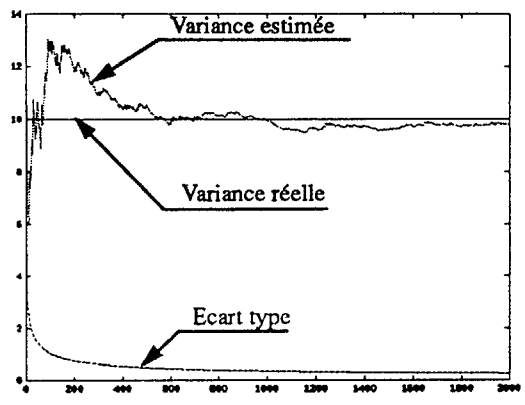


Figure 6.1

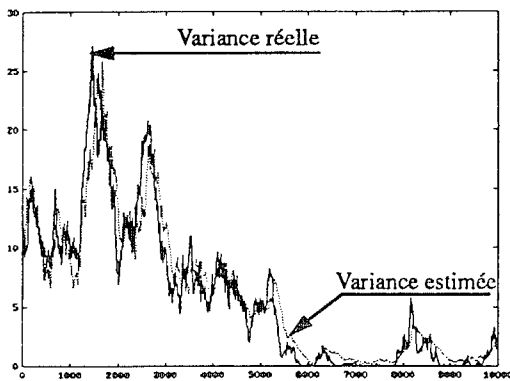


Figure 6.2

6.2. Bruit non-gaussien.

Considérons un système linéaire scalaire à perturbation blanche mais non-gaussienne. Le signal d'observation est linéaire à bruit additif blanc gaussien.

$$\begin{aligned} x_t &= f x_t + \pi_t \\ y_t &= x_t + v_t \end{aligned} \quad (6.2)$$

Dans ce cas, le filtre de Kalman est sous optimal. Par conséquent, un filtre de Volterra peut améliorer la qualité d'estimation de l'état comme le démontre l'exemple suivant. Soit π_t un processus à sauts défini de la manière suivante : $\pi_t = g(\Delta N - \lambda)$ où N est un processus de comptage de Poisson de fréquence λ . La comparaison entre trajectoires est visible sur la Figure 6.3. Sur la Figure 6.4, sont comparés les erreurs quadratiques moyennes des estimateur linéaire et de Volterra.

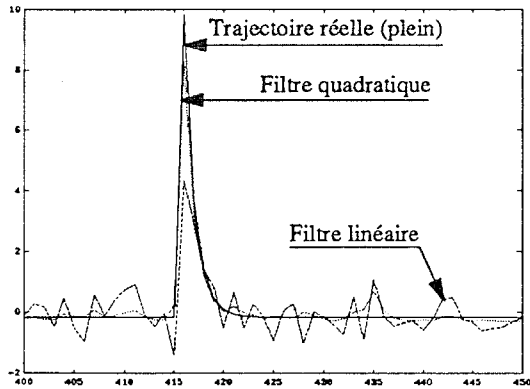
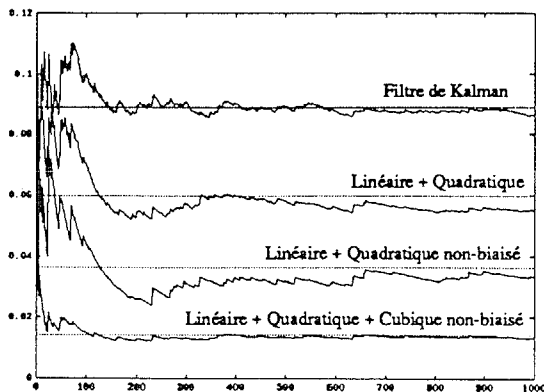


Figure 6.3.

Figure 6.4. $f = 0.1, g = 3, \lambda = 0.1, R = 0.1$.

7. Conclusion.

Dans ce travail, nous avons traité le filtrage polynomial à

horizon infini pour les systèmes stochastiques bilinéaires. L'estimation s'effectue dans la classe des séries de Volterra finies à noyaux séparables. Le problème consiste à déterminer ces noyaux de sorte que la série correspondante fournisse l'estimateur optimal au sens du minimum de variance. On montre alors que ce dernier est obtenu récursivement et réalisable en dimension finie pour les systèmes bilinéaires.

Deux exemples types ont été traités : l'estimation de la variance d'un bruit de dynamique par un filtre quadratique — l'estimation de l'état d'un système linéaire à perturbations ponctuelles par des filtres quadratiques puis cubiques. Dans ce cas, l'amélioration de la précision d'estimation avec le degré de la fonctionnelle utilisée est très nettement mise en évidence par la simulation.

8. Références.

- [1] L. A. ZADEH, *Optimum Non-linear Filters*. Jnl. of Applied Physics, Vol 24, n° 4, pp. 396-404, 1953,
- [2] N. WIENER. *Non-linear Problems In Random Theory*. M.I.T. Press, 1958.
- [3] V. S. PUGATCHEV. *Theory Of Random Functions And Its Applications To Control Problems*. Oxford, Pergamon Press, 1965.
- [4] N. WIENER. *Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. M.I.T. Press, 1949.
- [5] R.E. KALMAN. *A New Approach to Linear Filtering of Prediction Problems*. Trans. A.S.M.E. Jnl. of Basic Engineering, Vol 82D, 1960.
- [6] R.E. KALMAN, R. S. BUCY. *New Results in Linear Prediction and Filtering Theory*. Trans. A.S.M.E. Jnl. of Basic Engineering, Vol 83D, 1961.
- [7] J.M. CIOFFI, T. KAILATH. *Fast Recursive Least-squares Transversal Filters For Adaptive Filtering*. I.E.E.E. Trans. on Acoustics, Speech & Signal Processing, ASSP-32, n°2, pp. 304-313, 1984.
- [8] P. DUVAUT. *Le Filtrage de Wiener Linéaire-Quadratique à horizon fini - Application à la Prédiction*. Traitement du signal, Ed. GRETSI, Vol. 6, n°3, pp. 153-159, 1989.
- [9] M. FLIESS. *Séries de Volterra et Séries Formelles Non-Commutatives*. C. R. Acad. Sc. Paris, t. 280, 1975.
- [10] H. J. SUSSMAN. *Semigroup Representation, Bilinear Approximation of Input-Output Maps and Generalized Inputs*. Math. Syst. Theory, G, Marchesini & S, K, Mitter Eds, Lecture Notes in Economics and Mathematical Systems, Vol. 131, Springer Verlag, New-York, 1976.
- [11] R. W. BROCKETT. *Volterra Series and Geometric Control Theory*. Automatica, Vol. 12, pp. 167-176, 1976,
- [12] P. E. CROUCH. *Dynamical Realizations of Finite Volterra Series*. Siam J. of Control & Optimization, Vol. 19, n° 2, 1981.
- [13] J. W. RUGH. *Non-linear System Theory: The Volterra/Wiener Approach*. Johns Hopkins Univ. Press, 1981.
- [14] M. SCHETWEN. *The Volterra And Wiener Theories of Non-linear Systems*. John Wiley, New-York, 1980.
- [15] G. SALUT. *Identification Optimale des Systèmes Linéaires Stochastiques*. Thèse d'état, Université Paul Sabatier, Toulouse, 1976.
- [16] G. SALUT. *Filtrage Linéaire et Projections Réalisables*. Outils et modèles mathématiques pour l'automatique, Edition du C.N.R.S., 1983.
- [17] K. BERBICHE, G. SALUT. *Filtrage Polynomial par Corrélations d'un Système Bilineaire à Temps Discret*. A.P.I.I., Vol. 25, n°6, 1991.
- [18] K. BERBICHE, A. MONIN, G. MONTSENY, G. SALUT. *Filtrage Linéaire Quadratique des Systèmes Bilineaires à Temps Continu*. II-ème Colloque Maghrébin sur les modèles numériques de l'ingénieur, Rabat, Maroc, 1989.
- [19] P. R. BELANGER. *Estimation of Noise Covariance Matrices for a Linear Time Varying Stochastic Process*. 5-th I.F.A.C. World Congress, Paris, France, 1972.
- [20] A. H. JAZWINSKI. *Stochastic Processes and Filtering Theory*. Academic Press, New-York, 1970.