



PONDERATIONS VARIABLES POUR LES FILTRES EN TREILLIS ADAPTATIFS

E. MOISAN - P. COMON

CEPHAG - Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble
B.P. 46 - 38402 SAINT MARTIN D'HERES (FRANCE)

La littérature présente de plus en plus d'algorithmes en treillis à des fins de modélisation ou de filtrage. Parmi les solutions adaptatives exactes, nous nous sommes intéressés à la recherche de facteurs d'oubli variables dans le but d'adapter ensuite le treillis à d'éventuelles non-stationnarités du signal à traiter. Nous montrons d'abord que pour conserver une solution exacte obtenue par minimisation d'une erreur quadratique, l'ensemble des lois d'évolution possible du facteur d'oubli est assez restreint. Enfin, nous voyons qu'une pondération variable est en fait équivalente à un facteur d'oubli constant. Ces résultats sont confirmés par quelques simulations.

Recursive least squares ladder estimation algorithms have attracted much attention recently. This paper presents an adaptive least squares lattice using a varying exponential weighting, to track time-varying parameters of the observed data. It is shown that the forgetting factor has to follow some rules of evolution to respect the simple recursions of the usual adaptive lattice. Then, we show that using such a varying exponential weighting gives the same results that using a constant forgetting factor. These results are confirmed by simulations.

Depuis quelques années, les algorithmes en treillis relatifs à la prédiction linéaire adaptative se multiplient, en raison de nombreux domaines d'application : modélisation, analyse spectrale, filtrage... Parmi les solutions exactes, qui minimisent une erreur

quadratique moyenne du type $E_T = \sum_{t=p}^{T-q} e_t^2$, nous nous

sommes intéressés aux méthodes avec "fenêtre a priori", ce qui revient à supposer que le signal est nul avant le début de l'observation. L'avantage de cette méthode est qu'elle ne nécessite pas de connaissance a priori sur l'ordre du prédicteur en treillis, mais l'inconvénient est que la non-stationnarité à l'origine introduit une erreur d'initialisation. Pour résoudre ce problème, il existe plusieurs solutions :

- mémoire glissante de taille fixe M qui nécessite le stockage des M dernières observations, d'où un encombrement mémoire qui peut être important [1,2].
- pondération exponentielle : l'erreur à minimiser s'écrit alors $E_T = \sum_{t=1}^T h^{T-t} e_t^2$, où h est le facteur d'oubli, réel compris entre 0 et 1 [1,3].
- filtrage des observations avant la mise en oeuvre de l'algorithme en treillis [4].

Dans le cas d'une pondération exponentielle, le choix d'un facteur d'oubli correct résulte d'un compromis entre deux qualités incompatibles :

- algorithme rapidement adaptatif (h proche de 0, qui permet de s'affranchir rapidement de l'erreur d'initialisation).
- algorithme peu sensible au bruit (h proche de 1 permet d'utiliser une mémoire de grande taille, efficace dans le cas de données stationnaires bruitées).

Sans connaissance a priori sur le signal observé, nous avons cherché s'il était possible d'utiliser un facteur d'oubli variable au cours du temps, proche de 0 à l'initialisation pour lutter contre la non-stationnarité de l'observation, puis croissant vers 1 pour

lutter contre le bruit perturbateur. Dans la première partie, nous reprenons rapidement la théorie des treillis, selon une approche géométrique, afin de mettre en évidence les contraintes que doit satisfaire le facteur d'oubli, compatibles avec les formules usuelles de récurrence de l'algorithme adaptatif ([2], [3]).

Dans la deuxième partie, nous exploitons ces contraintes et obtenons les lois de variations des facteurs d'oubli. Nous montrons alors que les deux premières valeurs du facteur d'oubli déterminent entièrement son évolution mais que le résultat final est identique à celui obtenu avec un facteur d'oubli constant.

Nous présentons, en troisième partie, des simulations qui confirment le résultat théorique précédent.

Enfin, nous présentons rapidement, d'un point de vue purement géométrique, le problème à résoudre pour faire varier à volonté le facteur d'oubli.

I - PREDICTION LINEAIRE - APPROCHE GEOMETRIQUE

Pour alléger les notations, l'étude est menée dans un cas scalaire, d'autant que l'extension au cas vectoriel est aisée. Ayant observé un processus y_t de l'instant $t = 1$ à l'instant $t = T - 1$, on cherche à construire le prédicteur linéaire d'ordre N du futur

$$\text{échantillon } y_t : \hat{y}_T = -\sum_{i=1}^N a_{N,i} y_{T-i}$$

L'erreur instantanée de prédiction directe est alors

$$e_{N,T} = y_T + \sum_{i=1}^N a_{N,i} y_{T-i}$$

Si on introduit un facteur d'oubli constant h, l'erreur quadratique à minimiser s'écrit :

$$E_{N,T} = h E_{N,T-1} + (1-h) e_{N,T}^2$$

$$= (1-h) \sum_{t=1}^T h^{T-t} e_{N,t}^2$$

Si de plus, ce facteur d'oubli varie au cours du temps, l'erreur quadratique devient :

$$E_{N,T} = h_T E_{N,T-1} + (1-h_T) e_{N,T}^2 \tag{1}$$

$$= \sum_{t=1}^{T-1} \left(\prod_{i=t+1}^T h_i \right) (1-h_t) e_{N,t}^2 + (1-h_T) e_{N,T}^2$$



soit, sous forme matricielle

$$E_{N,T} = [e_{N,1} \dots e_{N,T}] H_T [e_{N,1} \dots e_{N,T}]'$$

où la matrice diagonale H_T est définie positive si $0 < h_t < 1 \quad t = 1 \text{ à } T$

$$H_T = \text{diag}_T \{ h_T \dots h_3 h_2 (1-h_1), h_T \dots h_3 (1-h_2), \dots, h_T (1-h_{T-1}), 1-h_T \}$$

Les coefficients $a_{N,i} (i=1 \text{ à } N)$ qui minimisent cette erreur quadratique sont alors solutions de l'équation:

$$[a_{N,1} \ a_{N,2} \ \dots \ a_{N,N}] Y_{N,T} H_T Y'_{N,T} = -Y_{1:T} H_T Y'_{N,T}$$

où $Y_{1:T}$ est la matrice ligne $[Y_1 \ Y_2 \ \dots \ Y_T]$
 $Y_{N,T}$ est la matrice $(N \times T)$

$$\begin{bmatrix} 0 & Y_1 & \dots & Y_N & \dots & Y_{T-1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & Y_1 & \dots & Y_{T-N} \end{bmatrix}$$

La solution adaptative introduit également la prédiction rétrograde [3,5] à l'ordre N sur les mêmes échantillons :

$$\hat{Y}_{T-1-N} = - \sum_{i=1}^N b_{N,N+1-i} Y_{T-i}$$

erreur de prédiction rétrograde

$$r_{N,T-1} = Y_{T-1-N} + \sum_{i=1}^N b_{N,N+1-i} Y_{T-i}$$

Les coefficients $b_{N,j} (j=1 \text{ à } N)$ qui minimisent l'erreur quadratique

$$F_{N,T} = [0 \ r_{N,1} \ \dots \ r_{N,T-1}] H_T [0 \ r_{N,1} \ \dots \ r_{N,T-1}]'$$

sont solutions de l'équation :

$$[b_{N,N} \ \dots \ b_{N,1}] Y_{N,T} H_T Y'_{N,T} = - Y_{1:T}^{N+1} H_T Y'_{N,T}$$

$$\text{avec } Y_{1:T}^{N+1} = [0 \ \dots \ 0 \ Y_1 \ \dots \ Y_{T-N-1}]$$

matrice ligne, issue de $(N+1)$ décalages à droite de $Y_{1:T}$.

En utilisant au besoin une matrice pseudo-inverse [6] si le produit $Y_{N,T} H_T Y'_{N,T}$ n'est pas inversible, on aboutit à :

$$[e_{N,1} \ e_{N,2} \ \dots \ e_{N,T}] = Y_{1:T} (I_T - H_T Y'_{N,T} (Y_{N,T} H_T Y'_{N,T})^{-1} Y_{N,T})^{-1} Y_{N,T}$$

$$[0 \ r_{N,1} \ \dots \ r_{N,T-1}] = Y_{1:T}^{N+1} (I_T - H_T Y'_{N,T} (Y_{N,T} H_T Y'_{N,T})^{-1} Y_{N,T})^{-1} Y_{N,T}$$

où I_T est la matrice identité $(T \times T)$

soit, en post-multipliant par $H_T^{1/2}$:

$$[e_{N,1} \ e_{N,2} \ \dots \ e_{N,T}] H_T^{1/2} = Y_{1:T} H_T^{1/2} (I_T - P_{Y_{N,T}})$$

$$[0 \ r_{N,1} \ \dots \ r_{N,T-1}] H_T^{1/2} = Y_{1:T}^{N+1} H_T^{1/2} (I_T - P_{Y_{N,T}})$$

L'opérateur

$$P_{Y_{N,T}} = H_T^{1/2} Y'_{N,T} (Y_{N,T} H_T Y'_{N,T})^{-1} Y_{N,T} H_T^{1/2}$$

n'est autre que le projecteur orthogonal sur le sous-espace vectoriel engendré par les lignes du produit $Y_{N,T} H_T^{1/2}$.

Alors $Q_{Y_{N,T}} = I_T - P_{Y_{N,T}}$ est le projecteur sur l'espace complémentaire. Les équations du treillis résultent des propriétés de ces opérateurs et sont largement commentées dans la littérature ([1] à [6]).

Nous ne reprendrons ici que les points qui imposent certaines contraintes à la matrice H_T , c'est à dire au facteur d'oubli variable h_T .

I.1 - Récurrence sur l'ordre

Il s'agit de passer du projecteur orthogonal $P_{Y_{N,T}}$ sur le sous-espace vectoriel engendré par les

lignes de $Y_{N,T} \cdot H_T^{1/2}$ au projecteur $P_{Y_{N+1,T}}$ sur le sous-espace engendré par $Y_{N+1,T} \cdot H_T^{1/2}$:

$$Y_{N+1,T} \cdot H_T^{1/2} = \begin{bmatrix} Y_{N,T} \\ Y_{N+1} \\ 1:T \end{bmatrix} H_T^{1/2} = \begin{bmatrix} Y_{N,T} \cdot H_T^{1/2} \\ Y_{N+1} \cdot H_T^{1/2} \\ 1:T \end{bmatrix}$$

Ces deux quantités étant déjà définies, la récurrence sur l'ordre ne pose aucun problème particulier.

I.2 - Récurrence sur l'ordre et le temps

Il faut maintenant déduire le projecteur $P_{Y_{N+1,T+1}}$ de $P_{Y_{N,T}}$:

$$Y_{N+1,T+1} \cdot H_{T+1}^{1/2} = \begin{bmatrix} 0 & Y_{1:T} \\ \vdots & \vdots \\ 0 & Y_{N,T} \end{bmatrix} H_{T+1}^{1/2}$$

est à relier à $Y_{N,T} \cdot H_T^{1/2}$.

En reprenant la définition des matrices diagonales H_T , on peut écrire :

$$H_{T+1} = \begin{bmatrix} h_{T+1} \cdot H_T & 0 \\ \vdots & \vdots \\ 0 & \dots & 0 & 1-h_{T+1} \end{bmatrix} = \begin{bmatrix} d_{T+1} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & h_{T+1} H_T \cdot D_T \end{bmatrix}$$

où le coefficient d_{T+1} est égal à $h_{T+1} h_T \dots h_2 (1-h_1)$ la matrice diagonale D_T est égale à

$$D_T = \text{diag}_T \left\{ \frac{1-h_2}{h_2(1-h_1)}, \frac{1-h_3}{h_3(1-h_2)}, \dots, \frac{1-h_T}{h_T(1-h_{T-1})}, \frac{1-h_{T+1}}{h_{T+1}(1-h_T)} \right\}$$

ainsi

$$Y_{N+1,T+1} \cdot H_{T+1}^{1/2} = h_{T+1} \begin{bmatrix} 0 & Y_{1:T} \cdot H_T^{1/2} \cdot D_T^{1/2} \\ \vdots & \vdots \\ 0 & Y_{N,T} \cdot H_T^{1/2} \cdot D_T^{1/2} \end{bmatrix}$$

Pour que le projecteur $P_{Y_{N+1,T+1}}$ se déduise de $P_{Y_{N,T}}$ selon la formule habituelle des treillis, il

semble nécessaire que la matrice D_T soit scalaire. C'est l'hypothèse faite :

$$D_T = \text{diag}_T \{ s_T, s_T, \dots, s_T \} \quad \text{où } s_T \text{ est un coefficient positif.}$$

I.3 - Récurrence sur le temps

Cette dernière récurrence, moins immédiate que les deux précédentes, résulte des propriétés du vecteur $\pi = [0 \ \dots \ 01]$. On montre en effet que :

$$Q_{Y_{N,T+1}} \pi = \begin{bmatrix} Q_{Y_{N,T-1}} & 0 \\ 0 & \dots & 0 \end{bmatrix} \quad \text{si } H_T \text{ et } H_{T-1} \text{ vérifient}$$

$$H_T = \begin{bmatrix} h_T \cdot H_{T-1} & 0 \\ \vdots & \vdots \\ 0 & \dots & 0 & 1-h_T \end{bmatrix} \quad \text{condition toujours satisfaite dans le cas qui nous intéresse}$$

II - ETUDE DU FACTEUR D'OUBLI

On vient de voir que les formules simples de récurrence sur le temps d'une part, puis sur l'ordre et le temps d'autre part, ne sont préservées que si la matrice H_{T+1} se décompose facilement en fonction de H_T selon les relations suivantes :



$$H_{T+1} = \begin{bmatrix} h_{T+1} \cdot H_T & 0 \\ \vdots & \vdots \\ 0 & \dots & 0 & 1-h_{T+1} \end{bmatrix} \text{ pour la r\u00e9currence sur le temps}$$

$$H_{T+1} = \begin{bmatrix} d_{T+1} & 0 & \dots & 0 \\ 0 & \vdots & \vdots & \vdots \\ 0 & \dots & h_{T+1} \cdot s_T \cdot H_T & \end{bmatrix} \text{ pour la r\u00e9currence sur l'ordre et le temps}$$

Ces relations sont toujours v\u00e9rifi\u00e9es dans le cas d'un facteur d'oubli constant ($h_{T+1} \cdot s_T = 1$ quel que soit T). Elles imposent, lorsque l'on fait varier le facteur d'oubli que :

$$D_T = \text{diag}_T \left\{ \frac{1-h_2}{h_2(1-h_1)}, \frac{1-h_3}{h_3(1-h_2)}, \dots, \frac{1-h_{T+1}}{h_{T+1}(1-h_T)} \right\}$$

$$= s_T \text{diag}_T \{ 1, 1, \dots, 1 \}$$

c'est \u00e0 dire :

$$\frac{1-h_{i+1}}{h_{i+1}(1-h_i)} = \frac{1-h_{j+1}}{h_{j+1}(1-h_j)} \quad (2)$$

quels que soient i et j compris entre 1 et T.

Pour d\u00e9terminer les lois de variation du facteur d'oubli, compatibles avec les formules usuelles de r\u00e9currence du treillis, il nous faut reprendre l'\u00e9quation (2) qui peut se mettre sous la forme :

$$\frac{1-h_{i+1}}{h_{i+1}} = \frac{1-h_i}{1-h_i} \frac{1-h_2}{h_2} \quad i = 1 \text{ \u00e0 } T$$

posons $u_i = \frac{h_i}{1-h_i} > 0$ puisque $0 < h_i < 1$

l'\u00e9quation pr\u00e9c\u00e9dente devient

$$u_{i+1} = \frac{u_2}{1+u_1} (1+u_i) \quad i = 1 \text{ \u00e0 } T \quad (3)$$

Si $\frac{u_2}{1+u_1} < 1$, la suite des u_i converge vers une limite v telle que

$$v = (1+v) \frac{u_2}{1+u_1} \quad \text{soit } v = \frac{u_2}{1+u_1-u_2}$$

c'est \u00e0 dire si $h_2 \frac{1-h_1}{1-h_2} < 1$, la suite des h_i tend vers

$$h_0 = h_2 \frac{1-h_1}{1-h_2} \text{ suppos\u00e9 compris entre 0 et 1}$$

On obtient finalement :

$$h_i = 1 - \frac{1-h_2 \frac{1-h_1}{1-h_2}}{1 + \frac{h_1-h_2}{h_2 (1-h_1)^2} \left(h_2 \frac{1-h_1}{1-h_2} \right)^i}$$

Il est clair que les deux premi\u00e8res valeurs h_1 et h_2 d\u00e9terminent enti\u00e8rement l'\u00e9volution du facteur

d'oubli vers sa valeur finale $h_2 \frac{1-h_1}{1-h_2}$ \u00e0 condition que celle-ci soit bien inf\u00e9rieure \u00e0 1.

La figure 6 illustre un exemple de facteur d'oubli variant selon la loi pr\u00e9c\u00e9dente ($h_1=0,1$ et $h_2=0,5$). Mais on peut montrer qu'utiliser un facteur d'oubli suivant une telle loi d'\u00e9volution est \u00e9quivalent \u00e0 utiliser une pond\u00e9ration constante \u00e9gale \u00e0 la valeur

$$\text{finale } h_2 \frac{1-h_1}{1-h_2}.$$

En effet, regardons l'erreur quadratique \u00e0 minimiser :

$$d'apr\u00e8s (1) E_{N,T} = (1-h_T) e^2_{N,T} + \sum_{t=1}^T (1-h_t) e^2_{N,t} \left(\prod_{i=t+1}^T h_i \right)$$

et en utilisant la contrainte (2), qui peut s'\u00e9crire

$$h_{i+1}(1-h_i) = h_2 \frac{1-h_1}{1-h_2} \quad (1-h_{i+1}) \quad i=1 \text{ \u00e0 } T$$

$$\text{alors } (1-h_t) \prod_{i=t+1}^T h_i = (1-h_T) \left(h_2 \frac{1-h_1}{1-h_2} \right)^{T-t}$$

ce qui correspond bien \u00e0 un facteur d'oubli constant

$$h_2 \frac{1-h_1}{1-h_2}.$$

Les simulations qui suivent confirment parfaitement ces r\u00e9sultats.

III - ETUDE EXPERIMENTALE

Le premier exemple illustre l'erreur d'initialisation en fonction du facteur d'oubli constant. Le signal trait\u00e9 est une sinuso\u00efde \u00e0 0,12 Hz, non bruit\u00e9e, \u00e9chantillonn\u00e9e \u00e0 1 Hz et observ\u00e9e sur 40 points. La figure 1 repr\u00e9sente les coefficients du pr\u00e9dicteur d'ordre 2, en fonction du nombre de points du signal, pour $h = 0,96$, tandis que la figure 2 est relative \u00e0 un facteur d'oubli plus faible $h = 0,2$. On constate que l'algorithme converge beaucoup plus vite dans le deuxi\u00eame cas.

Le second exemple est l'un de ceux trait\u00e9s en [2]. Il s'agit d'un processus autor\u00e9gressif d'ordre 2 :

$$y_T = 1,8 y_{T-1} - 0,95 y_{T-2} + 0,3 b_T$$

o\u00f9 b_T est un bruit blanc, centr\u00e9, norm\u00e9.

Les figures 3 \u00e0 5 repr\u00e9sentent les coefficients du pr\u00e9dicteur direct d'ordre 2 en fonction du nombre de points du signal. La figure 3 est relative \u00e0 un facteur d'oubli variable, dont l'\u00e9volution est donn\u00e9e en figure 6. La figure 4 est issue d'une pond\u00e9ration constante $h = 0,6$ et la figure 5 d'un facteur d'oubli constant $h = 0,9$. La remarque pr\u00e9c\u00e9dente reste vraie mais, surtout, on constate bien que les r\u00e9sultats obtenus avec un facteur d'oubli variable ($h_1 = 0,1$;

$h_2 = 0,5$; $h_2 \frac{1-h_1}{1-h_2} = 0,9$) sont semblables \u00e0 ceux obtenus pour $h = 0,9$.

IV - CONCLUSION

Dans un premier temps, nous avons \u00e9tabli \u00e0 quel type de variation doit ob\u00e9ir le facteur d'oubli d'un algorithme adaptatif en treillis pour que la solution soit la solution exacte obtenue par minimisation d'une erreur quadratique. Dans un deuxi\u00eame temps, nous avons montr\u00e9 qu'un facteur d'oubli qui suit une telle loi de variation conduit aux m\u00eames r\u00e9sultats qu'un facteur d'oubli constant. Le probl\u00e8me, tel que nous l'avons pos\u00e9, n'a donc sans doute pas de solution simple.

En effet, pour faire varier le facteur d'oubli \u00e0 volont\u00e9, toujours dans le souci d'adapter le treillis \u00e0 d'\u00e9ventuelles non-stationnarit\u00e9s du signal observ\u00e9, il semble qu'il faille revenir au projecteur orthogonal $P_{Y_{N,T}}$ lui-m\u00eame. En particulier, il faudrait pour

voir d\u00e9duire le projecteur orthogonal sur le sous-espace vectoriel engendr\u00e9 par les lignes du produit $Y_{N,T} \cdot H_T^{1/2} \cdot D_T^{1/2}$ du projecteur $P_{Y_{N,T}}$ et de la matrice D_T .

Est-ce possible sans accro\u00eetre consid\u00e9rablement le volume des calculs du treillis ? C'est le probl\u00e8me que nous \u00e9tudions maintenant.

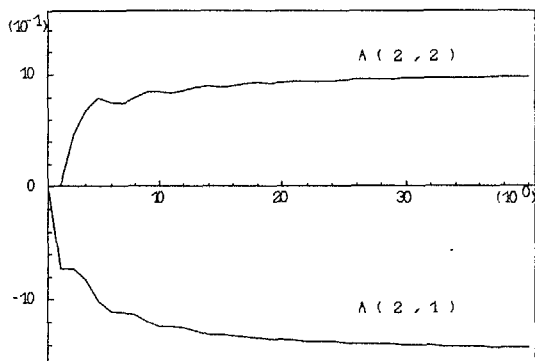


Fig. 1 : Coefficients du prédicteur (sinusoïde h=0,9)

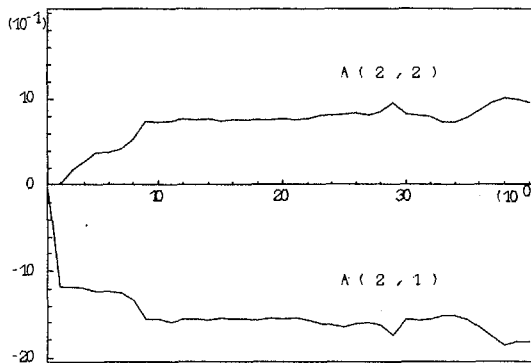


Fig.5 : Coefficients du prédicteur (processus AR h=0,9)

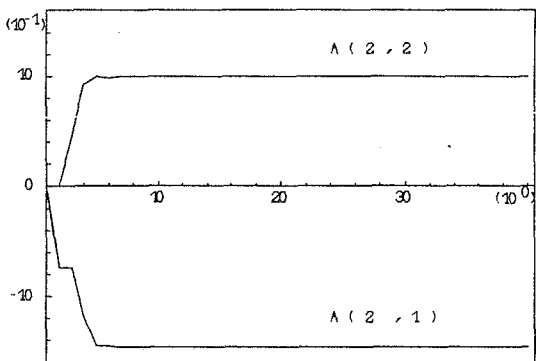


Fig.2 : Coefficients du prédicteur (sinusoïde h=0,2)

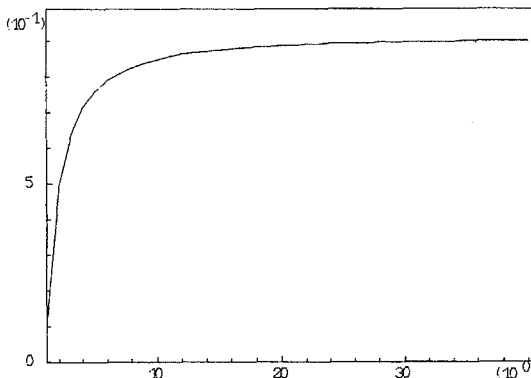


Fig.6 : facteur d'oubli variable (h₁=0,1 h₂=0,5 h₀=0,9)

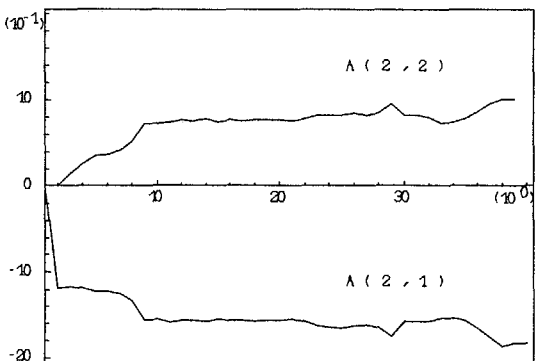


Fig.3 : Coefficients du prédicteur (processus AR h variable)

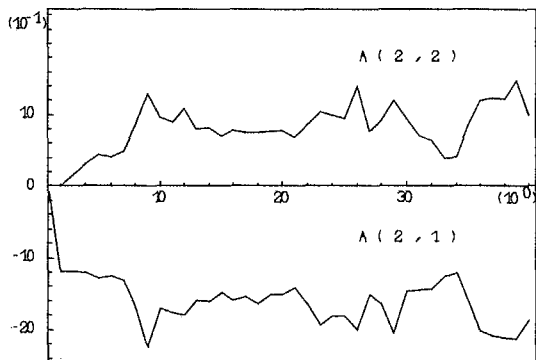


Fig.4 : Coefficients du prédicteur (processus AR h=0,6)

BIBLIOGRAPHIE

- [1] B.PORAT, T.KAILATH "Normalized lattice algorithms for least-squares FIR system identification" IEEE Trans. on ASSP, vol. 31, n°1. pp 122, 128. Février 1983.
- [2] B.FRIEDLANDER "Lattice filters for adaptive processing". Proc. IEEE, vol. 70, n° 8. pp 829, 867. Août 1982.
- [3] D.T.L. LEE, M.MORF "Recursive square-root ladder estimation algorithms" IEEE Int. Conf. ASSP (DENVER, CO), pp. 1005, 1017, Avril 1980.
- [4] J.MAKHOUL, L.K.COSELL "Adaptive lattice analysis of speech" IEEE Trans. on ASSP. Vol. 29, n° 3, pp 654, 658. Juin 1981.
- [5] H.LEV-ARI, T.KAILATH, J.CIOFFI "Least squares adaptive lattice and transversal filters : a unified geometric theory" IEEE Trans. on Info. Theory. Vol. IT 30, n° 2, pp 222, 236. Mars 1984.
- [6] B.PORAT, B.FRIEDLANDER, M.MORF "Square root covariance ladder algorithms" IEEE Trans. on Auto Control. Vol. AC 27, n° 4, pp 813,829. Août 1982.