Prédiction linéaire

La prédiction linéaire fait partie intégrante des filtres adaptatifs basés sur les algorithmes de moindres carrés rapides. De plus, elle joue un rôle important dans de nombreuses applications, notamment l'analyse des signaux ou la compression.

Plan

- La matrice d'autocorrélation
- La prédiction linéaire avant
- La prédiction linéaire arrière
- Relation entre les prédicteurs avant et arrière
- Relations de récurrence sur l'ordre et l'algorithme de Levinson-Durbin
- Le filtre de prédiction linéaire en treillis
- Résumé

La matrice d'autocorrélation

Soit une série temporelle stationnaire réelle $x(n)$, le vecteur correspondant aux L données les plus récentes est:

$$\mathbf{x}_L(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-L+1) \end{bmatrix}^T,$$

et la matrice d'autocorrélation de dimension $L \times L$ est:

$$\begin{aligned} \mathbf{R}_L &= E\left\{\mathbf{x}_L(n)\mathbf{x}_L^T(n)\right\} \qquad (1) \\ &= \begin{bmatrix} r(0) & r(1) & \cdots & r(L-1) \\ r(1) & r(0) & \cdots & r(L-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(L-1) & r(L-2) & \cdots & r(0) \end{bmatrix}. \end{aligned}$$

Pour un vecteur de longueur $L+1$:

$$\mathbf{x}_{L+1}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-L) \end{bmatrix}^T,$$

la matrice d'autocorrélation de dimension $(L+1) \times (L+1)$ est:

$$
\begin{aligned}
\mathbf{R}_{L+1} &= E\left\{\mathbf{x}_{L+1}(n)\mathbf{x}_{L+1}^T(n)\right\} \qquad (2)\\
&= \begin{bmatrix} r(0) & r(1) & \cdots & r(L) \\ r(1) & r(0) & \cdots & r(L-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(L) & r(L-1) & \cdots & r(0) \end{bmatrix} \\
&= \begin{bmatrix} r(0) & \mathbf{r}^T \\ \mathbf{r} & \mathbf{R}_L \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{R}_L & \mathbf{r}_\mathrm{b} \\ \mathbf{r}_\mathrm{b}^T & r(0) \end{bmatrix},
\end{aligned}
$$

où

$$
\mathbf{r} = \begin{bmatrix} r(1) & r(2) & \cdots & r(L) \end{bmatrix}^T
$$

est un vecteur d'autocorrélation à L éléments et

$$
\mathbf{r}_\mathrm{b} = \begin{bmatrix} r(L) & r(L-1) & \cdots & r(1) \end{bmatrix}^T.
$$

On voit bien donc comment les deux matrices $\mathbf{R}_{L+1}$ et $\mathbf{R}_L$ sont reliées.

La prédiction linéaire avant

Le but de la prédiction linéaire avant est d'estimer la valeur d'un signal à l'instant n à partir de ses valeurs aux instants antérieurs $n-1,\ n-2, \cdots$.

L'erreur de prédiction linéaire avant s'écrit:

$$\begin{aligned} e_{\mathrm{a}}(n) &= x(n) - \sum_{l=1}^{L} a_{L,l} x(n-l) \\ &= x(n) - \mathbf{a}_L^T \mathbf{x}(n-1) \end{aligned} \tag{3}$$

où

$$\mathbf{a}_L = \left[\begin{array}{cccc} a_{L,1} & a_{L,2} & \cdots & a_{L,L} \end{array}\right]^T$$

est le vecteur prédicteur avant à L éléments, et:

$$\mathbf{x}(n-1) = \left[\begin{array}{cccc} x(n-1) & x(n-2) & \cdots & x(n-L) \end{array}\right]^T.$$

On cherchera à minimiser le critère suivant:

$$J(\mathbf{a}_L) = E\{e_{\mathrm{a}}^2(n)\}, \tag{4}$$

ce qui donne:

$$\begin{aligned}
\frac{\partial J(\mathbf{a}_L)}{\partial \mathbf{a}_L} &= 2E\left\{e_{\mathrm{a}}(n)\frac{\partial e_{\mathrm{a}}(n)}{\partial \mathbf{a}_L}\right\} \\
&= -2E\left\{e_{\mathrm{a}}(n)\mathbf{x}(n-1)\right\} \\
&= \mathbf{0}_{L\times 1},
\end{aligned} \tag{5}$$

finalement, on a:

$$\begin{aligned}
E\left\{\mathbf{x}(n-1)\mathbf{x}^T(n-1)\right\}\mathbf{a}_L &= E\left\{\mathbf{x}(n-1)x(n)\right\} \\
\mathbf{R}_L\mathbf{a}_L &= \mathbf{r}.
\end{aligned} \tag{6}$$

Le système précédent peut être formulé différemment en augmentant sa taille:

$$\begin{aligned}
\begin{bmatrix} r(0) & \mathbf{r}^T \\ \mathbf{r} & \mathbf{R}_L \end{bmatrix}\begin{bmatrix} 1 \\ -\mathbf{a}_L \end{bmatrix} &= \begin{bmatrix} E_{\mathrm{a},L} \\ \mathbf{0}_{L\times 1} \end{bmatrix} \\
\mathbf{R}_{L+1}\begin{bmatrix} 1 \\ -\mathbf{a}_L \end{bmatrix} &= \begin{bmatrix} E_{\mathrm{a},L} \\ \mathbf{0}_{L\times 1} \end{bmatrix},
\end{aligned} \tag{7}$$

où

$$E_{\mathrm{a},L} = r(0) - \mathbf{r}^T\mathbf{a}_L \tag{8}$$

est la puissance d'erreur de prédiction avant. En fait $E_{\mathrm{a},L} = J_{\min}$ (EQM minimale).

Ecrivons les erreurs de prédiction linéaire avant à l'ordre L et $L - i$:

$$
\begin{aligned}
e_{\mathrm{a},L}(n) &= x(n) - \sum_{l=1}^{L} a_{L,l}\, x(n-l), && (9) \\
e_{\mathrm{a},L-i}(n) &= x(n) - \sum_{l=1}^{L-i} a_{L-i,l}\, x(n-l). && (10)
\end{aligned}
$$

Le principe d'orthogonalité nous dit que:

$$
E\left\{ e_{\mathrm{a},L}(n) \mathbf{x}(n-1) \right\} = \mathbf{0}_{L \times 1}. \qquad (11)
$$

Pour $1 \leq i \leq L$, on peut facilement vérifier en utilisant (11) que:

$$
E\left\{ e_{\mathrm{a},L}(n) e_{\mathrm{a},L-i}(n-i) \right\} = 0. \qquad (12)
$$

Donc quand $L \to \infty$, $E\left\{ e_{\mathrm{a}}(n) e_{\mathrm{a}}(n-i) \right\} = 0$ et la suite $e_{\mathrm{a}}(n)$ est un bruit blanc. C'est pourquoi le filtre d'erreur de prédiction est aussi appelé blanchisseur.

La prédiction linéaire arrière

Le but de la prédiction linéaire arrière est d'estimer la valeur d'un signal à l'instant $n - L$ à partir de ses valeurs aux instants futurs $n, \; n-1, \; n-2, \cdots$.

L'erreur de prédiction linéaire arrière s'écrit:

$$\begin{aligned} e_{\mathrm{b}}(n) &= x(n-L) - \sum_{l=1}^{L} b_{L,l} x(n-l+1) \\ &= x(n-L) - \mathbf{b}_L^T \mathbf{x}(n) \end{aligned} \tag{13}$$

où

$$\mathbf{b}_L = \left[\begin{array}{cccc} b_{L,1} & b_{L,2} & \cdots & b_{L,L} \end{array} \right]^T$$

est le vecteur prédicteur arrière à L éléments, et:

$$\mathbf{x}(n) = \left[\begin{array}{cccc} x(n) & x(n-1) & \cdots & x(n-L+1) \end{array} \right]^T.$$

On cherchera à minimiser le critère suivant:

$$J\left(\mathbf{b}_L\right) = E\{e_{\mathrm{b}}^2(n)\}, \tag{14}$$

ce qui donne:

$$\begin{aligned}
\frac{\partial J(\mathbf{b}_L)}{\partial \mathbf{b}_L} &= 2E\left\{e_{\mathrm{b}}(n)\frac{\partial e_{\mathrm{b}}(n)}{\partial \mathbf{b}_L}\right\} \\
&= -2E\left\{e_{\mathrm{b}}(n)\mathbf{x}(n)\right\} \\
&= \mathbf{0}_{L\times 1},
\end{aligned} \tag{15}$$

finalement, on a:

$$\begin{aligned}
E\left\{\mathbf{x}(n)\mathbf{x}^T(n)\right\}\mathbf{b}_L &= E\left\{\mathbf{x}(n)x(n-L)\right\} \\
\mathbf{R}_L\mathbf{b}_L &= \mathbf{r}_{\mathrm{b}}.
\end{aligned} \tag{16}$$

Le système précédent peut être formulé différemment en augmentant sa taille:

$$\begin{aligned}
\begin{bmatrix} \mathbf{R}_L & \mathbf{r}_{\mathrm{b}} \\ \mathbf{r}_{\mathrm{b}}^T & r(0) \end{bmatrix}\begin{bmatrix} -\mathbf{b}_L \\ 1 \end{bmatrix} &= \begin{bmatrix} \mathbf{0}_{L\times 1} \\ E_{\mathrm{b},L} \end{bmatrix} \\
\mathbf{R}_{L+1}\begin{bmatrix} -\mathbf{b}_L \\ 1 \end{bmatrix} &= \begin{bmatrix} \mathbf{0}_{L\times 1} \\ E_{\mathrm{b},L} \end{bmatrix},
\end{aligned} \tag{17}$$

où

$$E_{\mathrm{b},L} = r(0) - \mathbf{r}_{\mathrm{b}}^T\mathbf{b}_L \tag{18}$$

est la puissance d'erreur de prédiction arrière. En fait $E_{\mathrm{b},L} = J_{\min}$ (EQM minimale).

Une propriété importante de la prédiction arrière est qu'elle fournit un ensemble de signaux non corrélés: les erreurs $e_{\mathrm{b},l}(n)$ pour les ordres successifs $0 \leq l \leq L$ ne sont pas corrélées. En effet, soit le vecteur:

$$\begin{bmatrix} e_{\mathrm{b},0}(n) \\ e_{\mathrm{b},1}(n) \\ e_{\mathrm{b},2}(n) \\ \vdots \\ e_{\mathrm{b},L-1}(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\mathbf{b}_1^T & 1 & 0 & \cdots & 0 \\ -\mathbf{b}_2^T & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\mathbf{b}_{L-1}^T & 1 & & & \end{bmatrix} \mathbf{x}_L(n),$$

ou encore

$$\mathbf{e}_\mathrm{b}(n) = \mathbf{T}_\mathrm{b}\mathbf{x}_L(n). \quad (19)$$

La matrice de covariance correspondante s'écrit:

$$E\left\{\mathbf{e}_\mathrm{b}(n)\mathbf{e}_\mathrm{b}^T(n)\right\} = \mathbf{T}_\mathrm{b}\mathbf{R}_L\mathbf{T}_\mathrm{b}^T. \quad (20)$$

Par définition, c'est une matrice symétrique. Le produit $\mathbf{R}_L\mathbf{T}_\mathrm{b}^T$ est une matrice triangulaire inférieure, à cause de l'équation (17) et la diagonale principale se compose des puissances des erreurs de prédiction $E_{\mathrm{b},l}$ $(0 \leq l \leq L-1)$. Mais $\mathbf{T}_\mathrm{b}$ est aussi une matrice triangulaire inférieure et donc le produit doit avoir la même structure; comme il doit être symétrique, ce ne peut être qu'une matrice diagonale, soit:

$$E\left\{\mathbf{e}_{\mathrm{b}}(n)\mathbf{e}_{\mathrm{b}}^T(n)\right\} = \mathrm{diag}\{E_{\mathrm{b},l}\}, \qquad (21)$$

et les erreurs de prédiction arrière sont décorrélées.

D'autre part, puisque:

$$\mathbf{T}_{\mathrm{b}}\mathbf{R}_L\mathbf{T}_{\mathrm{b}}^T = \mathrm{diag}\{E_{\mathrm{b},l}\}, \qquad (22)$$

en prenant l'inverse:

$$\mathbf{T}_{\mathrm{b}}^{-T}\mathbf{R}_L^{-1}\mathbf{T}_{\mathrm{b}}^{-1} = \mathrm{diag}\{E_{\mathrm{b},l}\}^{-1}, \qquad (23)$$

et finalement:

$$\mathbf{R}_L^{-1} = \mathbf{T}_{\mathrm{b}}^T\mathrm{diag}\{E_{\mathrm{b},l}\}^{-1}\mathbf{T}_{\mathrm{b}}. \qquad (24)$$

C'est la décomposition triangulaire, dite de Cholesky, de la matrice d'autocorrélation inverse.

Relation entre les prédicteurs avant et arrière

Soit la matrice carrée d'ordre L, dite co-identité, suivante:

$$\mathbf{J}_L = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

qui retourne les composantes d'un vecteur. On peut facilement vérifier que:

$$\mathbf{R}_L \mathbf{J}_L = \mathbf{J}_L \mathbf{R}_L. \tag{25}$$

La matrice $\mathbf{R}_L$ est donc aussi symétrique par rapport à la seconde diagonale, elle est dite doublement symétrique ou persymétrique.

Rappelons que:

$$\mathbf{R}_L \mathbf{b}_L = \mathbf{r}_\mathrm{b}. \tag{26}$$

En multipliant par la matrice co-identité $\mathbf{J}_L$:

$$\begin{aligned} \mathbf{J}_L\mathbf{R}_L\mathbf{b}_L &= \mathbf{J}_L\mathbf{r}_\mathrm{b} \\ \mathbf{R}_L\mathbf{J}_L\mathbf{b}_L &= \mathbf{r} = \mathbf{R}_L\mathbf{a}_L, \end{aligned} \tag{27}$$

comme $\mathbf{R}_L$ est supposée inversible, il vient:

$$\mathbf{a}_L = \mathbf{J}_L\mathbf{b}_L. \tag{28}$$

D'autre part:

$$\begin{aligned} E_{\mathrm{b},L} &= r(0) - \mathbf{r}_\mathrm{b}^T\mathbf{b}_L \\ &= r(0) - \mathbf{r}_\mathrm{b}^T\mathbf{J}_L\mathbf{J}_L\mathbf{b}_L \\ &= r(0) - (\mathbf{J}_L\mathbf{r}_\mathrm{b})^T\mathbf{a}_L \\ &= r(0) - \mathbf{r}^T\mathbf{a}_L = E_{\mathrm{a},L} = E_L \end{aligned} \tag{29}$$

Pour un signal d'entrée stationnaire, les puissances d'erreur de prédiction avant et arrière sont égales et les coefficients sont les mêmes, mais dans l'ordre inverse.

Relations de récurrence sur l'ordre et l'algorithme de Levinson-Durbin

On mettra en évidence des relations entre les équations de prédiction linéaire à l'ordre $L-1$ et L.

Soient $\mathbf{r}_{\mathrm{a},L}$ et $\mathbf{r}_{\mathrm{b},L}$ les vecteurs suivants:

$$\begin{aligned}\mathbf{r}_{\mathrm{a},L} &= \begin{bmatrix} r(1) & r(2) & \cdots & r(L) \end{bmatrix}^T,\\ \mathbf{r}_{\mathrm{b},L} &= \mathbf{J}_L \mathbf{r}_{\mathrm{a},L},\end{aligned}$$

et soit l'équation:

$$\begin{bmatrix} \mathbf{R}_L & \mathbf{r}_{\mathrm{b},L} \\ \mathbf{r}_{\mathrm{b},L}^T & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ -\mathbf{a}_{L-1} \\ 0 \end{bmatrix} = \begin{bmatrix} E_{L-1} \\ \mathbf{0}_{(L-1)\times 1} \\ K_L \end{bmatrix}, \quad (30)$$

où le scalaire K_L représente la quantité:

$$\begin{aligned} K_L &= r(L) - \mathbf{a}_{L-1}^T \mathbf{r}_{\mathrm{b},L-1} \\ &= r(L) - \mathbf{a}_{L-1}^T \mathbf{J}_{L-1} \mathbf{r}_{\mathrm{a},L-1}. \end{aligned} \quad (31)$$

Le facteur:

$$k_L = \frac{K_L}{E_{L-1}} \qquad (32)$$

est appelé coefficient de réflexion.

Avec la prédiction linéaire arrière, il vient de même:

$$\begin{bmatrix} r(0) & \mathbf{r}_{\mathrm{a},L}^T \\ \mathbf{r}_{\mathrm{a},L} & \mathbf{R}_L \end{bmatrix} \begin{bmatrix} 0 \\ -\mathbf{b}_{L-1} \\ 1 \end{bmatrix} = \begin{bmatrix} K_L \\ \mathbf{0}_{(L-1)\times 1} \\ E_{L-1} \end{bmatrix}, \qquad (33)$$

en multipliant les deux membres par le facteur k_L, on a:

$$\mathbf{R}_{L+1} \begin{bmatrix} 0 \\ -k_L\mathbf{b}_{L-1} \\ k_L \end{bmatrix} = \begin{bmatrix} k_L^2 E_{L-1} \\ \mathbf{0}_{(L-1)\times 1} \\ K_L \end{bmatrix}. \qquad (34)$$

En soustrayant ensuite (34) de (30), on obtient:

$$\mathbf{R}_{L+1} \begin{bmatrix} 1 \\ k_L\mathbf{b}_{L-1} - \mathbf{a}_{L-1} \\ -k_L \end{bmatrix} = \begin{bmatrix} E_{L-1}(1-k_L^2) \\ \mathbf{0}_{L\times 1} \end{bmatrix}.$$

En supposant $\mathbf{R}_{L+1}$ invertible, la solution unique au système précédent est le prédicteur avant à l'ordre L, d'où les équations de récurrence suivantes:

$$\mathbf{a}_L = \begin{bmatrix} \mathbf{a}_{L-1} \\ 0 \end{bmatrix} - k_L \begin{bmatrix} \mathbf{b}_{L-1} \\ -1 \end{bmatrix}, \tag{35}$$

$$E_L = E_{L-1}(1 - k_L^2), \tag{36}$$

$$a_{L,L} = k_L. \tag{37}$$

Puisque la puissance de l'erreur de prédiction est toujours positive, quel que soit l'ordre, on en déduit:

$$|k_l| \leq 1, \ \forall l \geq 1, \tag{38}$$

et

$$0 \leq E_l \leq E_{l-1}, \ \forall l \geq 1. \tag{39}$$

En itérant sur la puissance, on a:

$$E_L = r(0) \prod_{l=1}^{L} (1 - k_l^2). \tag{40}$$

L'ensemble des équations (35)-(37) représente l'algorithme de Levinson-Durbin. La complexité pour résoudre l'équation de la prédiction linéaire à l'ordre L à l'aide de l'algorithme de Levinson-Durbin est de l'ordre L^2 au lieu de L^3.

Table 1: L'algorithme de Levinson-Durbin.

Initialisation: $E_0 = r(0)$

For $1 \leq l \leq L$:

$$k_l = \frac{1}{E_{l-1}} \left[r(l) - \mathbf{a}_{l-1}^T \mathbf{J}_{l-1} \mathbf{r}_{\mathrm{a},l-1} \right]$$

$$\mathbf{a}_l = \begin{bmatrix} \mathbf{a}_{l-1} \\ 0 \end{bmatrix} - k_l \mathbf{J}_l \begin{bmatrix} -1 \\ \mathbf{a}_{l-1} \end{bmatrix}$$

$$E_l = E_{l-1}(1 - k_l^2)$$

Le filtre de prédiction linéaire en treillis

Les coefficients k_l établissent des relations directes entre les erreurs de prédiction avant et arrière pour les ordres successifs.

A partir de la définition de l'erreur de prédiction avant à l'ordre L:

$$e_{\mathrm{a},L}(n) \quad = \quad x(n) - \mathbf{a}_L^T \mathbf{x}(n-1) \qquad (41)$$

et de la récurrence sur les coefficients:

$$\mathbf{a}_L \quad = \quad \begin{bmatrix} \mathbf{a}_{L-1} \\ 0 \end{bmatrix} - k_L \begin{bmatrix} \mathbf{b}_{L-1} \\ -1 \end{bmatrix},$$

il vient:

$$e_{\mathrm{a},L}(n) = e_{\mathrm{a},L-1}(n) - k_L[-\mathbf{b}_{L-1}^T \ 1]\mathbf{x}(n-1). \qquad (42)$$

L'erreur de prédiction arrière à l'ordre L est:

$$e_{\mathrm{b},L}(n) \quad = \quad x(n-L) - \mathbf{b}_L^T \mathbf{x}(n), \qquad (43)$$

et à l'ordre $L-1$:

$$\begin{aligned} e_{\mathrm{b},L-1}(n) &= x(n-L+1) - \mathbf{b}_{L-1}^T \mathbf{x}_{L-1}(n) \\ &= [-\mathbf{b}_{L-1}^T \; 1]\mathbf{x}(n). \end{aligned} \tag{44}$$

Par suite, ces erreurs de prédiction s'expriment par:

$$e_{\mathrm{a},L}(n) = e_{\mathrm{a},L-1}(n) - k_L e_{\mathrm{b},L-1}(n-1), \tag{45}$$

$$e_{\mathrm{b},L}(n) = e_{\mathrm{b},L-1}(n-1) - k_L e_{\mathrm{a},L-1}(n). \tag{46}$$

La structure correspondante est donnée a la figure 1; c'est la section de filtre en treillis et un filtre complet d'ordre L est constitué d'un ensemble de L telles sections en cascade. Bien sûr, au départ, $e_{\mathrm{a},0}(n) = e_{\mathrm{b},0}(n) = x(n)$.

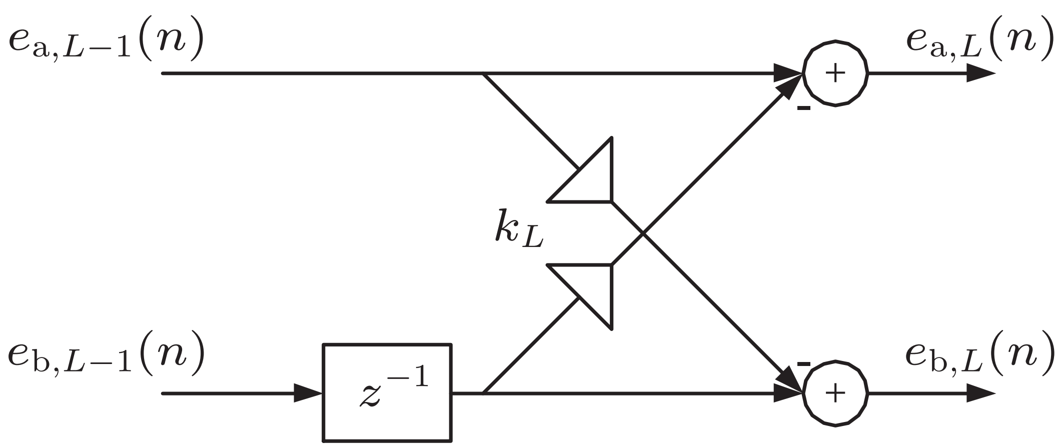

Figure 1: Cellule de filtre de prédiction en treillis.

Les coefficients de réflexion peuvent être caractérisés d'une manière statistique. Soit l'intercorrélation suivante:

$$E\{e_{\mathrm{a},L}(n)e_{\mathrm{b},L}(n-1)\} = \\ r(L+1) - \mathbf{b}_L^T\mathbf{r}_{\mathrm{a},L} - \mathbf{a}_L^T\mathbf{J}_L\mathbf{r}_{\mathrm{a},L} + \mathbf{a}_L^T\mathbf{R}_L\mathbf{b}_L.$$

En raison de l'équation de la prédiction arrière:

$$\mathbf{R}_L\mathbf{b}_L = \mathbf{r}_{\mathrm{b},L} = \mathbf{J}_L\mathbf{r}_{\mathrm{a},L},$$

la somme des deux derniers termes de cette intercorrélation est nulle, d'où:

$$E\{e_{\mathrm{a},L}(n)e_{\mathrm{b},L}(n-1)\} = r(L+1) - \mathbf{b}_L^T\mathbf{r}_{\mathrm{a},L} = K_{L+1},$$

et finalement:

$$k_L = \frac{E\{e_{\mathrm{a},L-1}(n)e_{\mathrm{b},L-1}(n-1)\}}{E_{L-1}}. \tag{47}$$

Les coefficients k_L de réflexion représentent l'intercorrélation normalisée des erreurs de prédiction avant et arrière.

Le coefficient k_L est lié aux L zéros z_l du filtre RIF de prédiction d'erreur d'ordre L, dont la fonction de transfert s'écrit:

$$A_L(z) = 1 - \sum_{l=1}^{L} a_{L,l} z^{-l} = \prod_{l=1}^{L} \left(1 - z_l z^{-1}\right), \quad (48)$$

et puisque $k_L = a_{L,L}$, on a:

$$k_L = (-1)^{L+1} \prod_{l=1}^{L} z_l. \quad (49)$$

Le filtre $A_L(z)$ est à phase minimale, c'est-à-dire que $|z_l| \leq 1$.

Nous allons montrer que le filtre

$$A_L(z) = 1 - \sum_{l=1}^{L} a_{L,l} z^{-l} \tag{50}$$

est à phase minimale. Pour simplifier les notations, posons: $w_l = -a_{L,l}$, avec $w_0 = 1$. Le polynôme devient:

$$A_L(z) = \sum_{l=0}^{L} w_l z^{-l}. \tag{51}$$

Soit le vecteur suivant:

$$\mathbf{w} = \begin{bmatrix} w_0 & w_1 & \cdots & w_L \end{bmatrix}^T,$$

nous savons que:

$$\mathbf{R}_{L+1}\mathbf{w} \quad = \quad \begin{bmatrix} E_{\mathrm{a},L} \\ \mathbf{0}_{L\times 1} \end{bmatrix}. \tag{52}$$

Soit λ une racine quelconque du polynôme, on en déduit que:

$$A_L(z) = (1 - \lambda z^{-1}) \sum_{l=0}^{L-1} c_l z^{-l}, \text{ avec } c_0 = 1. \tag{53}$$

Remarquer que λ peut être un nombre complexe, donc les coefficients c_l sont en général complexes.

Ainsi, le vecteur **w** peut encore s'écrire:

$$\mathbf{w} = \mathbf{c} - \lambda\tilde{\mathbf{c}}, \tag{54}$$

où

$$\mathbf{c} = \begin{bmatrix} 1 & c_1 & c_2 & \cdots & c_{L-1} & 0 \end{bmatrix}^T = \begin{bmatrix} \mathbf{c}'^T & 0 \end{bmatrix}^T,$$
$$\tilde{\mathbf{c}} = \begin{bmatrix} 0 & 1 & c_1 & c_2 & \cdots & c_{L-1} \end{bmatrix}^T = \begin{bmatrix} 0 & \mathbf{c}'^T \end{bmatrix}^T.$$

En remplaçant (54) dans (52), on obtient:

$$\mathbf{R}_{L+1}\mathbf{c} \quad = \quad \lambda\mathbf{R}_{L+1}\tilde{\mathbf{c}} + \begin{bmatrix} E_{\mathrm{a},L} \\ \mathbf{0}_{L\times 1} \end{bmatrix}. \tag{55}$$

Maintenant, si on multiplie à gauche par $\tilde{\mathbf{c}}^H$ des deux côtés de l'expression précédente, on a:

$$\tilde{\mathbf{c}}^H\mathbf{R}_{L+1}\mathbf{c} \quad = \quad \lambda\tilde{\mathbf{c}}^H\mathbf{R}_{L+1}\tilde{\mathbf{c}}. \tag{56}$$

D'où:

$$|\tilde{\mathbf{c}}^H\mathbf{R}_{L+1}\mathbf{c}|^2 \quad = \quad |\lambda|^2(\tilde{\mathbf{c}}^H\mathbf{R}_{L+1}\tilde{\mathbf{c}})^2. \tag{57}$$

En utilisant l'inégalité de Schwartz:

$$|\tilde{\mathbf{c}}^H \mathbf{R}_{L+1} \mathbf{c}|^2 \leq (\tilde{\mathbf{c}}^H \mathbf{R}_{L+1} \tilde{\mathbf{c}})(\mathbf{c}^H \mathbf{R}_{L+1} \mathbf{c}). \tag{58}$$

Or:

$$\begin{aligned}\tilde{\mathbf{c}}^H \mathbf{R}_{L+1} \tilde{\mathbf{c}} &= \begin{bmatrix} 0 & \mathbf{c}'^H \end{bmatrix} \begin{bmatrix} r(0) & \mathbf{r}^T \\ \mathbf{r} & \mathbf{R}_L \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{c}' \end{bmatrix} \\ &= \mathbf{c}'^H \mathbf{R}_L \mathbf{c}',\end{aligned} \tag{59}$$

de même:

$$\begin{aligned}\mathbf{c}^H \mathbf{R}_{L+1} \mathbf{c} &= \begin{bmatrix} \mathbf{c}'^H & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_L & \mathbf{r}_\mathrm{b} \\ \mathbf{r}_\mathrm{b}^T & r(0) \end{bmatrix} \begin{bmatrix} \mathbf{c}' \\ 0 \end{bmatrix} \\ &= \mathbf{c}'^H \mathbf{R}_L \mathbf{c}',\end{aligned} \tag{60}$$

d'où: $\tilde{\mathbf{c}}^H \mathbf{R}_{L+1} \tilde{\mathbf{c}} = \mathbf{c}^H \mathbf{R}_{L+1} \mathbf{c}$ et l'inégalité de Schwartz devient:

$$|\tilde{\mathbf{c}}^H \mathbf{R}_{L+1} \mathbf{c}|^2 \leq (\tilde{\mathbf{c}}^H \mathbf{R}_{L+1} \tilde{\mathbf{c}})^2. \tag{61}$$

D'après (57), on voit bien que $|\lambda|^2 \leq 1$.

Résumé

Erreur de prédiction linéaire avant:

$$e_{\mathrm{a}}(n) = x(n) - \mathbf{a}_L^T \mathbf{x}(n-1).$$

La minimisation du critère:

$$J(\mathbf{a}_L) = E\{e_{\mathrm{a}}^2(n)\}$$

donne

$$\mathbf{R}_L \mathbf{a}_L = \mathbf{r}.$$

Puissance d'erreur de prédiction avant:

$$E_{\mathrm{a},L} = r(0) - \mathbf{r}^T \mathbf{a}_L.$$

Erreur de prédiction linéaire arrière:

$$e_{\mathrm{b}}(n) = x(n-L) - \mathbf{b}_L^T \mathbf{x}(n).$$

La minimisation du critère:

$$J(\mathbf{b}_L) = E\{e_{\mathrm{b}}^2(n)\}$$

donne

$$\mathbf{R}_L\mathbf{b}_L = \mathbf{r}_{\mathrm{b}}.$$

Puissance d'erreur de prédiction arrière:

$$E_{\mathrm{b},L} = r(0) - \mathbf{r}_{\mathrm{b}}^T\mathbf{b}_L.$$

Relation entre les prédicteurs avant et arrière:

$$\begin{aligned} \mathbf{a}_L &= \mathbf{J}_L\mathbf{b}_L, \\ E_{\mathrm{a},L} &= E_{\mathrm{b},L}. \end{aligned}$$

Relation de récurrence sur l'ordre:

$$\mathbf{a}_L = \begin{bmatrix} \mathbf{a}_{L-1} \\ 0 \end{bmatrix} - k_L \begin{bmatrix} \mathbf{b}_{L-1} \\ -1 \end{bmatrix}.$$