COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 21 JUIN 75

LA PREDICTION LINEAIRE MODIFIEE ET SON APPLICATION A LA MODELISATION DU SIGNAL DE PAROLE

C. J. GUEGUEN

E. N. S. T. - LABORATOIRE DE THEORIE DES SYSTEMES (D 005) - 46 rue BARRAULT - PARIS CEDEX 13 - 75634 (FRANCE)

RESUME

L'analyse d'un signal, en vue de l'extraction d'un faible nombre de paramètres pertinents, peut être réalisée en identifiant les coefficients d'un modèle qui, soumis à une excitation donnée, reproduit un signal de même type. C'est le cas de la prédiction linéaire (LP) et des méthodes autorégressives (ARMA) qui ont connu ces dernières années un puissant développement, en particulier, dans le domaine de l'analyse de la parole.

Ces méthodes donnent par l'analyse de la fonction de transfert en z du modèle, une estimation du spectre du signal lissée et analytiquement définie. Mais il est vite apparu que, si les fréquences sont correctement restituées (formants), le calcul des amortissements demeure précaire et peut conduire à des problèmes d'instabilité.

C'est pourquoi, nous introduisons une prédiction linéaire modifiée (MLP) en utilisant une normalisation différente des paramètres du modèle qui conduit à fixer les pôles sur le cercle unitaire avec amortissement nul. Le nombre des paramètres à ajuster est donc réduit aux fréquences. Par ailleurs, cette approche possède des liens directs avec l'analyse factorielle ce qui conduit à introduire des théorèmes généraux sur les vecteurs propres des matrices de Toeplitz. Ces observations permettent d'expliquer les problèmes de stabilité rencontrés dans la prédiction linéaire classique et de réduire le coût des calculs concernant les matrices de covariance dans diverses applications.

SUMMARY

A pertinent signal analysis, extracting a small number of adequate parameters from the waveforms, can be performed by modelling techniques. Using this approach, a parametric model, excited by convenients inputs, is adjusted in order to match the corresponding output with the given signal. This is the case of the Classical Linear Prediction method (CLP) or of the usual autoregressive techniques (ARMA) which have received considerable attention in the past few years, especially for speech analysis purposes.

The above methods give a smoothed estimation of the spectrum with a complete analytical description by means of the z-transfer function of the model. But it was rapidly noted that if the formant frequencies are precisely determined, the information about the dampings was not correctly retreived.

This is the reason why we introduce in this paper the modified linear prediction (MLP) using a different normalization of the model parameters. This normalization is shown to locate the poles of the model on the unit circle. Zero-damped frequencies are only to be computed. This approach can be also interpreted as a factor analysis of the signal samples. The eigen vectors of Toeplitz matrices are investigated in this study and several theorems of general interest introduced. The consequence is a major reduction of the computational burden in many problems involving covariance matrices.

LA PREDICTION LINEAIRE MODIFIEE ET SON APPLICATION A LA MODELISATION DU SIGNAL DE PAROLE

1 - INTRODUCTION

De nombreuses applications requièrent d'extraire d'un signal complexe un faible nombre de paramètres pertinents. C'est, en particulier, le cas de la reconnaissance des formes où un nombre trop réduit de paramètres décrivant la forme nécessite des algorithmes sophistiqués pour la reconnaissance et où un nombre trop élevé de données redondantes alourdit les méthodes de traitement. La parole est un signal riche et complexe pour lequel ce problème est crucial. C'est pourquoi ces dernières années, on a vu se développer dans ce domaine des méthodes d'analyse performantes se prêtant évidemment à une extension à de nombreux autres signaux [1].

Ces méthodes peuvent être regroupées sous le concept de modélisation ou d'identification de système. Dans cette approche, en effet, le signal est considéré comme la sortie d'un système dont on recherche un modèle approximatif. A cette fin, la sortie d'un modèle paramétrique résultant d'une excitation donnée est comparée au signal à analyser. Un critère de différence étant défini, le processus d'identification consiste à ajuster les coefficients du modèle de façon à minimiser le critère. Les paramètres du modèle sont alors représentatifs du signal étudié.

Dans le domaine de la parole, les premières méthodes de ce type, bien que formulées de manière différente, ont été introduites par ATAL et HANAUER [2] et se sont considérablement développées par la suite. La principale raison en est que cette approche peut être interprétée comme une analyse par synthèse. En plus de leur intérêt propre (synthèse de la parole, transmission à faible débit), ces techniques permettent un contrôle de la qualité des paramètres extraits en ce qui concerne la préservation du contenu informationnel.

Mais il est vite apparu que l'implantation des algorithmes de prédiction linéaire classique (CLP) conduisait, quelle que soit la technique de calcul, à des problèmes délicats concernant la stabilité du modèle déterminé. Suivant la précision et bien que le modèle soit théoriquement stable, il est nécessaire de replacer, de temps en temps, les pôles de la fonction de transfert en z à l'intérieur du cercle unité. Cette constatation traduit le fait que, bien que les fréquences des formants soient correctement déterminées, les amortissements correspondants demeurent très imprécis. Cette imprécision n'est pas d'une importance considérable dans de nombreuses applications où seule la fréquence des formants est prise en compte, mais dans ce cas, l'application de la méthode conduit à doubler dans le calcul le nombre des paramètres utiles.

C'est pourquoi nous introduisons dans la suite, la prédiction linéaire modifiée (MLP) qui est fondée sur une normalisation différente des paramètres inconnus. Cette approche peut être interprétée comme une analyse factorielle des échantillons du signal et soulève le problème du calcul des vecteurs propres d'une matrice de covariance (de Toeplitz). Les propriétés ici introduites de ces vecteurs permettent de montrer que les pôles de la fonction de transfert du modèle sont situés sur le cercle unité. Seules les fréquences sont donc à déterminer. Mais de plus, notre approche peut être considérée comme le cas limite de la prédiction linéaire classique quand l'erreur de prédiction tend vers zéro. Les problèmes de stabilité rencontrés trouvent alors une explication naturelle.

Il est à noter que les résultats présentés ici sur le support du signal vocal sont de portée générale (identification de systèmes, égalisation de lignes, ...) et que les propriétés factorielles énoncées sont fondamentales dans diverses applications.

2 - LA PREDICTION LINEAIRE CLASSIQUE

Soient $\{s_n\}$ et $\{e_n\}$ les échantillons de la sortie et de l'entrée. ATAL et HANAUER [2] introduisent pour représenter le modèle la série autorégressive :

$$a_o s_n + a_1 s_{n-1} + \dots + a_p s_{n-p} = e_n \quad \text{avec} \quad a_o = 1 \qquad (1)$$

En l'absence de mesures relatives à la source vocale, on est amené à supposer que l'excitation e_n est un bruit blanc gaussien à moyenne nulle. La relation (1) peut être interprétée comme une équation de prédiction de l'échantillon s_n à partir des mesures précédentes :

$$\hat{s}_n = -\sum_{i=1}^{p} a_i s_{n-i} \quad , \quad e_n = s_n - \hat{s}_n \qquad (2)$$

et e_n est alors l'erreur de prédiction instantanée. L'ajustement des coefficients a_i peut être ainsi réalisé au sens des moindres carrés de e_n.

Introduisant les notations :

$$\underline{s}_n^T = [s_n, s_{n-1}, \dots, s_{n-p}] \qquad \underline{a} = [a_o, a_1, \dots, a_p]$$

$$S = [\underline{s}_o, \underline{s}_1, \dots, \underline{s}_N] \quad \text{et} \quad V_p = SS^T$$

Le problème revient à minimiser le critère C (moindres carrés de e_n sur l'intervalle O, N) par rapport à $\underline{a}$:

$$C = \underline{a}^T V_p \underline{a} \qquad \text{avec} \qquad a_o = 1 \qquad (4)$$

Supposant de plus que N est assez grand, la matrice V_p d'éléments v_i, prend la forme de Toeplitz et la solution du problème d'optimisation est donnée par l'équation classique :

$$\begin{bmatrix} v_o & v_1 & \cdots & v_{p-1} \\ v_1 & \ddots & & \vdots \\ \vdots & & \ddots & v_1 \\ v_{p-1} & \cdots & v_1 & v_o \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix} \qquad (5)$$

La prédiction linéaire s'intègre donc dans les méthodes d'ajustement de modèles autorégressifs (éventuellement à moyenne glissante)(ARMA) du type introduit depuis 1946 par BARLETT [3] dans l'étude de certaines séries temporelles. L'équation (5) connue sous le nom de Yule-Walker a été considérablement étudiée quant aux propriétés de l'estimation (WALD [4], HANNAN [5], DURBIN [6]) et diverses méthodes de résolution ont été proposées (WIGGINS, ROBINSON [7], HO, LEE [8]). Si le système est représentable par le simple modèle autorégressif (1), l'estimation au sens des moindres carrés coïncide avec celle du maximum de vraisemblance, les $\hat{a}_i$ solutions de (5) sont asymptotiquement sans biais et normalement distribuées [9][10]. Ce type de problème reçoit toujours une attention considérable à l'heure actuelle [11][12].

Dans le cadre de la parole, ATAL et HANAUER [2] font appel à la procédure de Cholewski pour résoudre l'équation (5). Une contribution importante est fournie par ITAKURA et SAITO [13] en utilisant l'algorithme introduit dès 1946 par LEVINSON [14]. Par une approche différente, GUEGUEN et CARAYANNIS [15] mettent en oeuvre la formulation du filtre de Kalman pour aboutir à une estimation récursive. Ces diverses méthodes ont fait l'objet d'études approfondies et d'applications [16], [17], [18].

3 - LA PREDICTION LINEAIRE MODIFIEE

3.1 - Formulation

Quelque soit le processus de calcul, il apparaît que la formulation précédente pose de sérieux problèmes de sensibilité en ce qui concerne la stabilité du modèle déduit. L'analyse de signaux synthétiques issus d'un modèle connu montre que, si les fréquences sont correctement restituées, il n'en va pas de même pour les amortissements correspondants. En général, les pôles du modèle sont situés trop près du cercle unitaire ce qui entraîne un spectre où les formants sont dotés de bandes passantes faibles. Les informations concernant les amortissements étant sujets à caution, il paraît intéressant de développer des méthodes visant à ne déterminer que les fréquences des formants.

L'hypothèse essentielle sur laquelle repose la prédiction linéaire et les méthodes autorégressives introduites est que le coefficient a_o dans (1) est unité. Cette hypothèse fixe l'ordre du modèle à p et entraîne une solution unique pour le problème d'optimisation (4). Dans la suite, nous introduirons, sous le nom de "prédiction linéaire modifiée", le problème suivant : Minimiser le critère C avec la nouvelle contrainte :

$$C = \underline{a}^T V_p \underline{a} \qquad \underline{a}^T \underline{a} = 1 \qquad (6)$$

Associant le paramètre de Lagrange λ à la contrainte, on obtient les conditions nécessaires d'optimalité sur le Lagrangien :

$$\mathcal{L} = \underline{a}^T V_p \underline{a} + \lambda (1 - \underline{a}^T \underline{a})$$

$$\frac{\delta \mathcal{L}}{\delta \underline{a}} = 2 V_p \underline{a} - 2 \lambda \underline{a} \qquad \frac{\delta \mathcal{L}}{\delta \lambda} = 1 - \underline{a}^T \underline{a}$$

D'où l'estimation optimale $\hat{\underline{a}}$ est le vecteur propre associé à la valeur propre minimale :

$$V_p \hat{\underline{a}} = \lambda \hat{\underline{a}} \qquad (7)$$

et la valeur correspondante de l'optimum :

$$C = \hat{\underline{a}} V_p \hat{\underline{a}} = \lambda \hat{\underline{a}}^T . \hat{\underline{a}} = \lambda \qquad (8)$$

3.2 - Relation avec l'analyse factorielle

Une façon très directe de paramétriser un signal serait de mémoriser des ensembles de p échantillons consécutifs pour représenter leur évolution temporelle typique. Chaque ensemble de paramètres ainsi prélevé contribue à former un nuage d'occurrences d'apprentissage dont l'analyse factorielle peut être réalisée [19]. L'analyse en composantes principales consiste, en particulier, à déterminer les axes principaux d'inertie de ce nuage, c'est-à-dire les vecteurs propres de la matrice : $SS^T = V_p$.

Une interprétation analogue consisterait à considérer les $\{s_n\}$ comme un processus stochastique stationnaire avec :

$$E(s_n) = 0 \qquad \text{et} \qquad E(s_n s_{n-i}) = v_i$$

et d'en effectuer le développement de Karhumen-Loeve. Cette transformation orthonormale destinée à engendrer des variables non corrélées, consiste à extraire les vecteurs propres de la matrice de covariance V_p.

La relation avec la prédiction linéaire modifiée est alors évidente puisqu'il s'agit aussi d'extraire les vecteurs propres d'une matrice de Toeplitz. Cependant, il est à noter que d'une manière usuelle dans l'analyse factorielle, la matrice S comporte un grand nombre de colonnes pour un plus faible nombre de ligne et l'on s'intéresse alors aux valeurs propres dominantes. Dans la MLP, l'intérêt est inverse puisque p est faible ($\simeq 10$) pour un nombre d'occurrences relativement grand ($\simeq 200$). Ces deux aspects sont cependant complémentaires et peuvent être utilisés concurremment avec profit.

4 - PROPRIETES FACTORIELLES DES MATRICES DE TOEPLITZ

Le problème central dans la méthode est d'effectuer un calcul rapide et précis des facteurs d'une matrice V_p sous la forme de Toeplitz. L'importance de cette question dans de nombreuses disciplines (filtre de Wiener, expansion de Karhumen-Loeve) a déjà suscité de nombreuses études [20] [21] qui n'ont cependant pas abouti à des résultats complets.

En application de la formule de Schur, l'équation caractéristique de V_p peut-être calculée par récurrence pour obtenir les valeurs propres (positives) :

Théorème 1 : Excepté pour les racines de $\Psi_{k-1}(\lambda)$, on a :

$$\Psi_k(\lambda) = \Psi_{k-1}(\lambda)\left[(v_o - \lambda) + v^T (V_{k-1} - \lambda I)^{-1} v\right] \qquad (9)$$

avec $v^T = [v_1, v_2, \ldots, v_{k-1}]$

Si l'on suppose toutes les racines de $\Psi_p(\lambda)$ distinctes, on a évidemment :

Théorème 2 : Les vecteurs propres a_i de V_p sont orthogonaux deux à deux.

Sans les mêmes hypothèses, nous introduisons [22] les théorèmes suivants sur les vecteurs propres de V_p.

Soit J la matrice qui a pour éléments $\delta_{i,\, p+2-i}$, on a :

Théorème 3 : Si $\underline{a}$ est vecteur propre associé à λ, il en est de même pour $J\underline{a}$.

La démonstration est immédiate en vérifiant les identités :

$$V_p = J\, V_p\, J \qquad \text{et} \qquad J^2 = I \qquad (10)$$

d'où

$$Va = \lambda a \quad \text{entraîne} \quad V(Ja) = \lambda(Ja)$$

Théorème 4 : Pour toute valeur propre (simple) et vecteur propre, on a l'une des égalités :

$$\underline{a} + J\underline{a} = 0 \qquad \text{ou} \qquad \underline{a} - J\underline{a} = 0$$

En effet, $\underline{a}$ et $J\underline{a}$ associés à la même valeur propre simple doivent être colinéaires et les valeurs propres de J sont +1 ou -1.

On en déduit l'important théorème de décomposition suivant pour toute matrice de covariance (ou plus généralement vérifiant (10)).

Supposant que V_p est de dimension paire 2m, on a les matrices partitionnées :

$$V_p = \begin{bmatrix} V & JU \\ UJ & V \end{bmatrix} \quad \text{et} \quad T = \begin{bmatrix} I & J \\ J & -I \end{bmatrix} \qquad (T^2 = I)$$

avec

$$V = V_{m-1} \quad \text{et} \quad U = \begin{bmatrix} v_1 & v_2 & \cdots & v_{m-1} \\ v_2 & & \cdot^{\cdot^{\cdot}} & \\ \vdots & \cdot^{\cdot^{\cdot}} & & \\ v_{m-1} & & & v_p \end{bmatrix}.$$

Par vérification directe, on démontre :

Théorème 5 : On a le changement de base :

$$TV_pT = \text{Diag}\,[V + JUJ,\ V - U]$$

Il en résulte que les vecteurs propres de V_p se déduisent par des transformations simples de ceux des deux matrices de dimension m (V + U) et (V - U). Il est à noter que ce résultat s'oppose à l'algorithme rapide introduit par FUKUNAGA [21].

5 - COMPARAISON DES DEUX APPROCHES

Par le simple jeu de la normalisation différente des paramètres à estimer et malgré l'identité des critères, la CLP et la MLP se trouvent douées de propriétés différentes en ce qui concerne le nombre de solutions, la stabilité du modèle, ... La relation entre les deux approches peut être explicitée et éclaire les problèmes d'instabilité de la CLP.

5.1 - Nombre des solutions

La condition d'optimalité (7) montre l'existence de (p+1) extremums dont les seuls retenus sont relatifs aux valeurs propres minimales. Pour la parole sans pré-accentuation, l'expérience montre que pour p = 10, ces valeurs sont petites et peu différentes entre elles. Il en résulte que chacun des vecteurs propres correspondants fournit autant de coefficients de prédiction a_i de qualité presqu'égale. Par opposition à la prédiction linéaire classique où la solution est unique, les estimations correspondantes forment un sous-espace de faible sensibilité pour les performances. Des erreurs de transmission ou des erreurs d'arrondis sur les a_i maintenant le vecteur $\underline{a}$ dans ce sous-espace n'auront donc que peu d'influence sur la qualité du système.

5.2 - Stabilité du modèle

La stabilité du modèle est démontrée théoriquement pour la CLP. Il n'en va pas de même dans la pratique où le modèle peut se révéler instable dans 5/100 des cas. On est donc contraint dans cette éventualité d'imposer au système un amortissement artificiel (recalage des pôles de la fonction de transfert).

Les propriétés des vecteurs propres de V_p exposées au théorème 4, imposent que les pôles du modèle sont les zéros du polynome A(z) :

$$A(z) = \sum_{i=0}^{p} a_i \, z^{-i} \qquad A^*(z) = z^{-p} A(z^{-1})$$

qui vérifie avec son polynome aux inverses l'une des égalités :

$$A(z) = \pm A^*(z)$$

Moyennant une hypothèse supplémentaire, il en résulte :
Théorème 6 : Avec la contrainte $\underline{a}^T\underline{a} = 1$ (MLP), les pôles du modèle sont situés sur le cercle unité.

La prédiction linéaire modifiée réalise donc la synthèse du signal par superposition de diverses sinusoïdes non amorties. Le modèle est à l'exacte limite de la stabilité. Ce résultat pouvait d'ailleurs être attendu. En effet, les seules données du problème sont la matrice de covariance et la contrainte de normalisation qui sont en MLP indépendantes du sens du temps.

5.3 - Nombre de paramètres à estimer

A égalité d'ordre pour le modèle, le théorème 6 montre que, en MLP, le nombre de paramètres à déterminer est moitié puisqu'il ne s'agit que des fréquences. Ces paramètres sont, d'ailleurs, dans la plupart des cas, les seuls à rechercher étant donnée l'incertitude sur les amortissements. La MLP représenterait alors une forte réduction du nombre de calculs.

Il faut cependant se garder d'une telle affirmation étant donnée la nature très abstraite du critère d'optimisation adopté dans les deux cas. La détermination de l'ordre du modèle doit être remise en question sur la base de critères plus globaux relatifs à l'application (qualité de la parole synthétique, par exemple). Ces considérations conduisent, en général, à un ordre supérieur pour la MLP.

5.4 - Relation entre les solutions

La qualité de la MLP est directement liée à la présence d'une valeur propre quasi-nulle dans V_p. En notant par λ le critère commun aux deux approches, celles-ci peuvent être comparées par :

CLP : $V_p\underline{a} = \lambda\underline{u}$ MLP : $V_p\underline{a} = \lambda\underline{a}$

avec $u^T = [1 0 \ldots 0]$

On voit donc que, quand $\lambda \to 0$, les deux solutions deviennent identiques. Ce qui signifie que plus la qualité de la prédiction est grande (en particulier, quand l'ordre p augmente), plus les pôles se rapprochent du cercle unité en CLP. Des phénomènes d'imprécision numérique peuvent alors entraîner l'instabilité.

6 - CONCLUSION

L'étude des propriétés des modèles autorégressifs soumis à une normalisation particulière des paramètres a été développée. Cette normalisation définissant la "prédiction linéaire modifiée" est naturelle et permet d'établir des liens directs avec l'analyse factorielle. Les théorèmes d'usage général développés à cette occasion sur les vecteurs propres des matrices de Toeplitz ont permis d'en expliciter les propriétés. Une comparaison avec l'établissement d'un modèle autorégressif par prédiction linéaire classique dans le cadre de l'analyse de la parole montre certains avantages sensibles de cette nouvelle formulation dont diverses extensions sont encore à l'étude.

REFERENCES

/1/ M. MORF, G. SIDHU, T. KAILATH : Some new algorithms for recursive estimation in constant, linear, discrete-time systems
IEEE Trans. on AC, Vol AC-19, n° 4, pp 315-323, 1974

/2/ B. ATAL, S. HANAUER : Speech analysis and synthesis by linear prediction of the speech wave
JASA, Vol 41, pp 293-309, 1967

/3/ M. BARLETT : On the theoretical specification and sampling properties of auto-correlated time series
J. Royal Stat. Soc., Vol B-8, pp 27-40, 1946

/4/ H. WOLD : The analysis of stationary time series
Alnquist and Nickseel, Uppsala, 1954

/5/ E. HANNAN : Time series analysis
Methuen, London, 1960

/6/ J. DURBIN : The fitting of time series models
Rev. Inst. Int. Stat., Vol 28, pp 233-244, 1960

/7/ R. WIGGINS, E. A. ROBINSON : Recursive solution to the multichannel filtering problem
J. Geophys. Res., Vol 7, pp 1885-1891, 1965

/8/ Y. HO, R. LEE : Identification of linear dynamic systems

Inform. Control, Vol 2, pp 99-110, 1965

/9/ H. MANN, A. WALD : On the statistic treatment of linear stochastic equations
Econometrica, Vol 11, pp 173-220, 1943

/10/ W. GERSCH : Estimation of the autoregressive parameters of a mixed autoregressive moving average time series
IEEE Trans. on AC, Vol AC-15, n° 5, pp 583-588, 1970

/11/ W. GERSCH : Least square estimates of structural system parameters using covariance function data
IEEE Trans. on AC, Vol AC-19, n° 6, pp 898-903, 1974

/12/ D. GRAUPE, D. KRAUSE, J. MOORE : Identification of autoregressive moving average parameters of time series
IEEE Trans. on AC, Vol AC-20, n° 1, pp 104-107, 1975

/13/ F. ITAKURA, S. SAITO : Digital filtering techniques for speech analysis and synthesis
7 Int. Congress on Acoustics, 25 C 1, Budapest, 1971

/14/ N. LEVINSON : The Wiener RMS (Root mean square) error criterion in filter design and prediction
J. Math. and Phy. Vol 25, n° 1, pp 262-278, 1946

/15/ C. GUEGUEN, G. CARAYANNIS : Analyse de la parole par filtrage optimal de Kalman
Automatisme, Tome 18, n° 3, pp 99-105,1973

/16/ R. VISWHANATHAN, J. MAKHOUL : Current issues in linear predictive speech compression
EASCON Conf., invited paper, Washington, 1974

/17/ J. MARKEL, A. GRAY : On autocorrelation equations as applied to speech analysis
IEEE Trans. on Audio. and Elec., Vol AU-21, n°2, pp 69-79, 1973

/18/ A. SAGE, J. MELSA : Development of a configuration concept of speech digitizer based on adaptive estimation techniques
Report Information and Control Science Center, S.M.U., Dallas

/19/ J. P. BENZECRI et Al. : L'analyse de donnêes
Dunod, Tome 1 et 2, Paris, 1973

/20/ U. GRANADER, G. SZEGO : Toeplitz forms and their applications
California monographs in Math. Science, U.C.L.A. Press, 1958

/21/ K. FUKUNAGA : Introduction to statistical pattern recognition
Academic Press, Chap. 8, pp 241-255, New York, 1972

/22/ C. GUEGUEN : Note sur certaines propriêtês des matrices de covariance et leur application â la modêlisation des signaux stationnaires
C.R. Ac. Sci. Paris, scêance du 3 mars 1975