

Nouvelle présentation

de la méthode du maximum d'entropie

The maximum entropy method

B. PICINBONO et M. BARRET

Laboratoire des Signaux et Systèmes, ESE, Plateau de Moulon, 91192 GIF-SUR-YVETTE CEDEX (France)
 et Technopole de Metz, 57078, METZ, CEDEX (France)

RÉSUMÉ

La méthode d'analyse spectrale par le maximum d'entropie est reprise dans le but d'en donner un fondement partant des principes de base sur l'entropie et l'estimation. On montre en particulier que le maximum du débit d'entropie correspond au maximum de l'erreur de prédiction. Utilisant alors les coefficients de réflexion, on montre très simplement l'équivalence de la méthode avec celle du modèle autorégressif. On présente également une interprétation en termes de blancher, ce qui permet de discuter ce que donnerait le minimum d'entropie.

MOTS-CLÉS : Maximum d'entropie, modèles autorégressifs, prédiction, coefficients de réflexion.

SUMMARY

The method of maximum entropy spectral analysis is reviewed. The main purpose is to establish this method from some first principles on entropy and estimation theory. It is shown that the maximum of entropy rate corresponds to the maximum of prediction error. Using the reflection coefficients, the equivalence with the autoregressive method is directly established. An interpretation in terms of whitening is given and the minimum of entropy rate is also discussed.

KEY WORDS : Maximum of entropy, autoregressive models, prediction, reflection coefficients.

TABLE DES MATIÈRES

1. Introduction
 2. Rappels essentiels sur l'entropie
 - 2.1. Définition de l'entropie
 - 2.2. Entropie d'une loi normale
 - 2.3. Théorème de base sur l'entropie normale
 - 2.4. Entropie et transformations linéaires
 - 2.5. Entropie et transformation d'innovation
 - 2.6. Débit d'entropie
 3. Méthode du maximum d'entropie
 4. Interprétation de la méthode
 5. Remarques sur le maximum d'entropie et le minimum d'erreur de prédiction
 6. Conclusion
- Appendice

1. Introduction

La méthode du maximum d'entropie (ME) est très couramment utilisée en analyse spectrale et si l'on veut se dispenser des intermédiaires de calcul il suffit de se souvenir qu'elle est dans ses grandes lignes équivalente à la méthode de modélisation par un modèle autorégressif (AR). Les détails de calcul sont par contre souvent omis et quand ils sont fournis ils ne présentent pas, à notre avis, une cohérence et une simplicité que l'on est en droit d'attendre pour une méthode classique et sanctionnée par l'usage.

La présente note pédagogique se propose de combler cette lacune et si les résultats, sauf exception, ne sont pas originaux, la méthode proposée nous paraît de loin la plus simple possible. Si son exposé peut paraître un peu long, c'est que nous

avons voulu présenter toutes les étapes de la démarche en partant des « premiers principes », c'est-à-dire des définitions des concepts d'entropie et des résultats fondamentaux que l'on doit connaître en matière de description de signaux aléatoires. Ceci veut évidemment dire que certains lecteurs peuvent sauter des paragraphes dont le contenu leur est déjà très familier. Enfin, comme il s'agit d'une note pédagogique entièrement autonome, il ne nous a pas paru nécessaire de l'ornementer d'une longue liste de références sur tous les travaux concernant le maximum d'entropie mais ne cherchant pas à présenter la méthode dans sa globalité. Un seul article à notre connaissance a tenté une démarche de même type mais en empruntant un chemin entièrement différent et moins direct [1].

Les étapes de notre démarche sont les suivantes. Les principaux concepts sur l'entropie sont rappelés dans la section 2. Le résultat essentiel consiste à établir le lien entre l'entropie et l'erreur de prédiction à un pas et passé infini. Comme cette erreur s'exprime très simplement au moyen des coefficients de réflexion eux même fortement liés aux propriétés de modèles autorégressifs, la méthode du maximum d'entropie s'établit très directement, comme on le voit dans la section 3. Une interprétation de la méthode en terme de blancheur est présentée dans la section 4, ce qui conduit à examiner dans la section 5 ce qui serait le minimum d'entropie. Enfin un appendice présente les principaux résultats de théorie du signal utilisés dans la discussion.

2. Rappels essentiels sur l'entropie

2.1. DÉFINITION DE L'ENTROPIE

Soit un vecteur aléatoire \mathbf{X} à n dimensions continu réel et centré et de densité de probabilité $p_{\mathbf{X}}(\mathbf{x})$. On appelle *entropie* de ce vecteur la quantité

$$(2.1) \quad H_{\mathbf{X}} = -E[\text{Log } p_{\mathbf{X}}(\mathbf{X})] \\ = - \int p_{\mathbf{X}}(\mathbf{x}) \text{Log } [p_{\mathbf{X}}(\mathbf{x})] d\mathbf{x}$$

que l'on suppose exister. L'entropie possède de nombreuses propriétés analysées dans les ouvrages de théorie de l'information, mais ce dont nous avons besoin pour exposer la méthode du ME se résume à très peu de choses. Notons que le logarithme apparaissant dans (2.1) est le logarithme népérien.

2.2. ENTROPIE D'UNE LOI NORMALE

Supposons que \mathbf{X} soit un vecteur aléatoire $N(\mathbf{0}, \mathbf{R})$. Sa densité s'écrit alors

$$(2.2) \quad p(\mathbf{x}) = (2\pi)^{-n/2} \{\det \mathbf{R}\}^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x} \right\}.$$

Le logarithme de cette quantité vaut

$$(2.3) \quad \text{Log } p(\mathbf{x}) = - (n/2) \text{Log } 2\pi \\ - (1/2) \text{Log } [\det \mathbf{R}] - (1/2) \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}.$$

Comme $\mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}$ peut s'écrire $\text{Tr}(\mathbf{x}\mathbf{x}^T \mathbf{R}^{-1})$, où Tr signifie la trace de la matrice carrée, on a

$$(2.4) \quad E(\mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}) = E[\text{Tr}(\mathbf{x}\mathbf{x}^T \mathbf{R}^{-1})] = \text{Tr}(\mathbf{I}) = n$$

où \mathbf{I} est la matrice identité. Utilisant alors (2.1) on obtient

$$(2.5) \quad H_{\mathbf{N}}(\mathbf{R}) = (n/2) \text{Log}(2\pi e) + (1/2) \text{Log}[\det(\mathbf{R})]$$

où \mathbf{R} est la matrice de covariance du vecteur \mathbf{X} apparaissant dans la définition de la loi normale (2.2). La quantité $H_{\mathbf{N}}(\mathbf{R})$ représente donc l'entropie d'une loi $N(\mathbf{0}, \mathbf{R})$, ou l'indice \mathbf{N} rappelle la loi normale. Pour simplifier le langage on dénomme $H_{\mathbf{N}}(\mathbf{R})$ l'*entropie normale*.

2.3. THÉORÈME DE BASE SUR L'ENTROPIE NORMALE

Proposition. Soit un vecteur aléatoire \mathbf{X} centré, de matrice de covariance \mathbf{R} et d'entropie $H_{\mathbf{X}}$. On a alors

$$(2.6) \quad H_{\mathbf{X}} \leq H_{\mathbf{N}}(\mathbf{R})$$

l'égalité ayant lieu si et seulement si \mathbf{X} est $N(\mathbf{0}, \mathbf{R})$.

Démonstration : [2], p. 108. Appelons $q(\mathbf{x})$ la densité de probabilité de \mathbf{X} et $p(\mathbf{x})$ la densité donnée par (2.2). En raison de (2.3) la moyenne de $\text{Log}[p(\mathbf{X})]$ est la même si on la calcule à l'aide des densités $p(\mathbf{x})$ ou $q(\mathbf{x})$ puisque la matrice de covariance est la même sous les deux lois. On en déduit

$$(2.7) \quad \int q(\mathbf{x}) \text{Log}[p(\mathbf{x})] d\mathbf{x} = \\ = \int p(\mathbf{x}) \text{Log}[p(\mathbf{x})] d\mathbf{x} = -H_{\mathbf{N}}(\mathbf{R}).$$

Il en résulte donc

$$(2.8) \quad H_{\mathbf{X}} - H_{\mathbf{N}}(\mathbf{R}) = \int q(\mathbf{x}) \text{Log}[p(\mathbf{x})/q(\mathbf{x})] d\mathbf{x}.$$

Utilisant alors l'inégalité $\text{Log } x \leq x - 1$, l'égalité n'ayant lieu que si $x = 1$, on obtient

$$(2.9) \quad H_{\mathbf{X}} - H_{\mathbf{N}}(\mathbf{R}) \leq 0$$

l'égalité n'ayant lieu que si $q(\mathbf{x}) = p(\mathbf{x})$, c'est-à-dire si \mathbf{X} est $N(\mathbf{0}, \mathbf{R})$.

2.4. ENTROPIE ET TRANSFORMATIONS LINÉAIRES

Associons au vecteur \mathbf{X} de densité $p_{\mathbf{X}}(\mathbf{x})$ le vecteur \mathbf{Y} défini par $\mathbf{Y} = \mathbf{A}\mathbf{X}$, où \mathbf{A} est une matrice inversible. On sait alors que \mathbf{Y} est une variable aléatoire de densité

$$(2.10) \quad p_{\mathbf{Y}}(\mathbf{y}) = p_{\mathbf{X}}(\mathbf{A}^{-1}\mathbf{y}) |\det(\mathbf{A})|^{-1}.$$

Appliquant alors la définition (2.1) on trouve immédiatement

$$(2.11) \quad H_{\mathbf{Y}} = H_{\mathbf{X}} + \text{Log} |\det(\mathbf{A})|.$$

En particulier les transformations inversibles et telles que $\det(A) = 1$ donnent

$$(2.12) \quad H_Y = H_X.$$

2.5. ENTROPIE ET TRANSFORMATION D'INNOVATION

Au vecteur X de composantes X_1, X_2, \dots, X_n , on peut associer le vecteur \tilde{X} de composantes $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n$ constitué par les innovations successives qui forment des variables aléatoires décorrélatées. Plus précisément on a $\tilde{X}_1 = X_1$ et

$$(2.13) \quad \tilde{X}_i = X_i - \text{EL}[X_i | X_1, X_2, \dots, X_{i-1}]$$

où EL représente la prédiction linéaire en moyenne quadratique de X_i à l'aide de X_1, X_2, \dots, X_{i-1} . D'après le principe d'orthogonalité \tilde{X}_i est non corrélée à X_1, X_2, \dots, X_{i-1} donc à $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_{i-1}$. Il est clair que $\tilde{X} = AX$, où A est une matrice triangulaire dont les éléments diagonaux sont égaux à un. Il en résulte que $\det(A) = 1$ et d'après (2.12)

$$(2.14) \quad H_{\tilde{X}} = H_X.$$

Dans le cas où X est normal, \tilde{X} l'est aussi et la covariance de \tilde{X} vaut

$$(2.15) \quad R = \text{diag}(\eta_i^2)$$

qui, en raison de l'indépendance des \tilde{X}_i , est une matrice diagonale faisant intervenir les erreurs de prédiction

$$(2.16) \quad \eta_i^2 = E(\tilde{X}_i^2).$$

L'entropie normale (2.5) peut donc s'écrire à l'aide des erreurs de prédiction par

$$(2.17) \quad H_N(R) = (n/2) \text{Log}(2\pi e) + (1/2) \sum_{i=1}^n \text{Log}(\eta_i^2).$$

2.6. DÉBIT D'ENTROPIE

Supposons maintenant que les composantes de X soient les valeurs d'un signal aléatoire centré stationnaire à temps discret. On appelle ([2], p. 18) débit d'entropie (entropy rate) de ce signal la quantité

$$(2.18) \quad h_X = \lim_{n \rightarrow \infty} (1/n) H_X.$$

Calculons alors le débit d'entropie d'un signal normal. Pour celui-ci on peut appliquer (2.17), quel que soit n . Mais on a de plus

$$(2.19) \quad \lim_{i \rightarrow \infty} \eta_i^2 = \varepsilon^2$$

où ε^2 représente l'erreur de prédiction à un pas et passé infini du signal considéré. Si cette erreur n'est

pas nulle, ce qui caractérise un signal aléatoire non déterministe [3], on a

$$(2.20) \quad h_N = (1/2) \text{Log}(2\pi e) + (1/2) \text{Log}(\varepsilon^2).$$

Ainsi le débit d'entropie d'un signal normal est une fonction strictement croissante de l'erreur de prédiction à un pas et passé infini.

Cette erreur de prédiction peut s'exprimer au moyen de la fonction de corrélation du signal ou plutôt au moyen de la densité spectrale $\Gamma(\nu)$ par l'expression bien connue

$$(2.21) \quad \varepsilon^2 = \exp \left\{ \int_{-1/2}^{+1/2} \text{Log} \{ \Gamma(\nu) \} d\nu \right\}.$$

Mais pour la suite il est beaucoup plus commode d'utiliser les coefficients de réflexion k_i liés à l'algorithme de Levinson de résolution des équations de prédiction. On sait alors (voir si nécessaire l'appendice) que l'erreur de prédiction est donnée par

$$(2.22) \quad \varepsilon^2 = \gamma[0] \prod_{i=1}^{\infty} (1 - k_i^2).$$

L'algorithme de calcul des coefficients k_i montre que pour obtenir k_i , $1 \leq i \leq r$, il suffit de connaître les corrélations $\gamma[p]$ pour $0 \leq p \leq r$. Cette remarque est essentielle pour la suite.

3. Méthode du maximum d'entropie

Supposons que les $r+1$ premières valeurs $\gamma[i]$, $0 \leq i \leq r$, de la fonction de corrélation d'un signal stationnaire $X[n]$ soient connues. Peut-on alors déterminer l'ensemble des valeurs de la fonction de corrélation $\gamma[p]$, $p > r$, de sorte que le *débit d'entropie soit maximum* ?

Il convient tout d'abord de noter que le problème posé n'a de sens que si les coefficients $\gamma[i]$, $0 \leq i \leq r$, satisfont des conditions leur permettant d'être les $r+1$ premières valeurs d'une fonction de corrélation d'un signal aléatoire régulier.

Une condition nécessaire et suffisante pour qu'il en soit ainsi est que la matrice de Toeplitz symétrique dont la première ligne vaut $\gamma[0], \gamma[1], \dots, \gamma[r]$ soit définie positive. La démonstration de ce résultat ne jouant pas un rôle central dans la suite est donnée en appendice.

La méthode du ME est fondée sur le théorème suivant.

Théorème : La fonction de corrélation dont les $r+1$ premières valeurs sont connues et qui conduit au maximum du débit d'entropie est celle d'un signal normal (ou gaussien) autorégressif d'ordre r , $AR(r)$, dont le vecteur de régression est déduit des valeurs connues de la fonction de corrélation.

Démonstration : Le caractère normal provient du résultat du paragraphe 2.3. En effet il n'est pas possible que le maximum du débit d'entropie soit atteint pour un signal non normal puisqu'à tout signal non normal on peut associer le signal normal

possédant la même fonction de corrélation et dont l'entropie, donc le débit d'entropie est plus grand.

Méthode 1 [4]. Soit alors $G_r(\{\gamma[i]\})$ la classe de tous les signaux gaussiens qui ont la même fonction de corrélation $\gamma[i]$ pour $0 \leq i \leq r$. Tous ces signaux ont donc la même erreur de prédiction à un pas et passé fini r , notée ε_r^2 . Leur erreur à passé infini ε^2 satisfait évidemment $\varepsilon^2 \leq \varepsilon_r^2$. Or ε_r^2 est l'erreur à passé infini du signal de $G_r(\{\gamma[i]\})$ autorégressif d'ordre r . Ce signal est donc, d'après (2.20), celui de $G_r(\{\gamma[i]\})$ à débit d'entropie maximum.

Méthode 2. Tous les signaux de $G_r[\{\gamma[i]\}]$ ont les mêmes coefficients de réflexion k_i , $1 \leq i \leq r$ en raison de la remarque faite à la fin du paragraphe précédent. L'erreur à passé infini donnée par (2.22) est donc maximum si tous les coefficients k_i , $i > r$, sont nuls, ce qui caractérise un signal AR(r).

Remarque : Le lecteur peut se demander où se trouve dans ceci la méthode d'analyse spectrale par le maximum d'entropie. Il est clair qu'elle s'en déduit directement, car la densité spectrale d'un modèle AR(r) se déduit directement du vecteur de régression lui-même calculé à partir des valeurs connues de la fonction de corrélation. Pour plus de détails on peut se reporter à l'appendice.

4. Interprétation de la méthode

La méthode du maximum d'entropie pourrait tout autant s'appeler méthode du maximum de l'erreur de prédiction, d'après (2.20). Ceci peut paraître surprenant, tant on est habitué à minimiser l'erreur de prédiction, et nécessite donc quelques commentaires.

Notons tout d'abord que l'erreur de prédiction à un pas et passé infini ε^2 ne pouvant être supérieure à la variance $\gamma[0]$ du signal considéré, on a

$$(4.1) \quad \varepsilon^2 \leq \sigma^2 = \gamma[0].$$

La borne supérieure est atteinte lorsque le signal $X[n]$ est blanc, c'est-à-dire quand

$$(4.2) \quad \gamma[p] = \sigma^2 \delta[p].$$

Il est évident que dans le cas normal le même type d'inégalité est valable pour le débit d'entropie h_N en raison de (2.20), et c'est pourquoi nous nous contenterons dans toute la suite de parler de ε^2 , dénommée plus simplement « erreur de prédiction ».

Si donc on reprend le problème de base de la méthode du ME avec $r = 0$, c'est-à-dire si l'on cherche le signal à erreur de prédiction maximum connaissant simplement $\gamma[0]$, on trouve qu'il s'agit du bruit blanc normal de variance $\gamma[0]$.

On peut alors se demander ce que serait la solution du problème concernant à rechercher un minimum de l'erreur de prédiction plutôt qu'un maximum. Le minimum de ε^2 est évidemment nul et correspond à un signal dit déterministe et de variance donnée. On sait qu'il y a une infinité de tels signaux, les plus

connus étant ceux possédant un spectre se limitant à un nombre arbitraire de raies spectrales. On sait même que pour qu'un signal soit déterministe il est nécessaire que sa densité spectrale s'annule au moins pour une fréquence, ce qui est dans une certaine mesure l'opposé de la blancheur [3].

On voit donc que le critère du ME ou de l'erreur de prédiction maximum peut s'interpréter comme un critère de blancheur maximum. De plus chercher le minimum de l'erreur de prédiction conduit à une infinité de solutions ayant toutes une densité spectrale nulle au moins en une fréquence.

Cette interprétation se généralise pour r quelconque. Le signal, solution unique du problème du maximum d'entropie, peut être considéré comme le plus blanc possible, compte-tenu des données sur sa fonction de corrélation. Il y a par contre une infinité de signaux différents dont les $r + 1$ premières valeurs de la fonction de corrélation sont données et conduisant à une erreur de prédiction nulle. Ces signaux déterministes sont à l'opposé de la blancheur et vont maintenant être examinés.

5. Remarques sur le minimum d'entropie ou le minimum d'erreur de prédiction

Partons toujours de (2.22) dans laquelle $\gamma[0]$ et les k_i , $1 \leq i \leq r$, sont donnés. Cherchons alors les conditions permettant d'avoir une erreur de prédiction ε^2 nulle. On voit immédiatement qu'il y a une infinité de solutions, puisque $\varepsilon^2 = 0$ dès qu'au moins un $k_j = 1$. Sans entrer dans le détail de calculs qui peuvent être trouvés dans [3], indiquons qu'il y a deux grands types de signaux déterministes ou à erreur de prédiction nulle : ceux uniquement à raies spectrales et ceux dont le spectre ne se limite pas à des raies.

a) *Signaux à raies spectrales.* Supposons qu'on ait $k_m^2 = 1$ et que $k_j^2 < 1$ pour $j < m$. On a alors $\varepsilon^2 = 0$ et de plus on peut montrer que le signal correspondant possède une densité spectrale se limitant à m raies spectrales [5]. Si ce signal satisfait les conditions posées dans la méthode du maximum d'entropie, les coefficients de réflexion k_i vérifient $k_i^2 < 1$ pour $i \leq r$. En conséquence on a $m > r$ et le nombre minimum de raies spectrales associé est donc $r + 1$. Notons que l'hypothèse des raies spectrales est celle utilisée dans la méthode d'analyse spectrale de Pisarenko. En général on superpose à ce signal un bruit supposé blanc. Dans le résultat obtenu le signal à raies spectrales constitue la partie déterministe du développement de Wold, le bruit blanc étant la partie purement aléatoire.

b) *Autres signaux déterministes.* Il s'agit de signaux pour lesquels $\varepsilon^2 = 0$ sans qu'aucun coefficient de réflexion k_i soit tel que $k_i^2 = 1$. Il suffit pour cela que le produit infini apparaissant dans (2.22) soit nul, ce qui peut se faire de bien des manières.

Plus précisément on a $\varepsilon^2 = 0$ si et seulement si la série de terme général k_i^2 diverge. Partons en effet de (2.22) en supposant pour simplifier que $\gamma[0] = 1$.

On a alors

$$(5.1) \quad \text{Log}(\epsilon^2) = \sum_{i=1}^n \text{Log}(1 - k_i^2).$$

Si k_i^2 tend vers zéro, alors $\text{Log}(1 - k_i^2)$ est équivalent à $-k_i^2$. Il en résulte que la série (5.1) et celle de terme général k_i^2 sont de même nature ayant des termes généraux de même signe. Réciproquement si k_i^2 ne tend pas vers zéro les deux séries divergent.

Les signaux correspondants ont une densité spectrale s'annulant au moins en une fréquence. Ils peuvent être à bande limitée, mais ceci n'est pas nécessaire [3].

6. Conclusion

Partant de la définition de l'entropie d'une variable aléatoire vectorielle nous avons établi la méthode du maximum d'entropie. Celle-ci est fondée sur deux résultats : l'entropie d'un vecteur aléatoire de matrice de corrélation donnée est maximum si ce vecteur possède une loi de probabilité normale ; le débit d'entropie d'un signal aléatoire stationnaire est relié par une fonction croissante à l'erreur de prédiction à un pas et passé infini. On en déduit alors de façon très simple que si les r premières valeurs de la fonction de corrélation d'un signal sont connues, son débit d'entropie est maximum quand ce signal est gaussien et autorégressif d'ordre r . Le vecteur de régression associé se calcule à partir des valeurs connues de la fonction de corrélation en utilisant les équations normales. A l'opposé le minimum du débit d'entropie correspond à un signal aléatoire à erreur de prédiction nulle, c'est-à-dire de type déterministe, dont la densité spectrale doit s'annuler au moins en une fréquence.

Appendice

PRÉDICTION ET COEFFICIENTS DE RÉFLEXION

Soit un signal aléatoire $X[n]$ réel centré et stationnaire à temps discret et de fonction de corrélation $\gamma[k]$. On appelle prédiction linéaire à un pas et passé fini p l'estimation linéaire en moyenne quadratique de $X[n]$ sous la forme

$$(A.1) \quad \hat{X}_p[n] = \mathbf{a}_p^T \mathbf{X}[n] = a_1 X[n-1] + a_2 X[n-2] + \dots + a_p X[n-p].$$

Le vecteur \mathbf{a}_p est dénommé vecteur de régression, pour des raisons qui seront plus claires ultérieurement. D'après (A.1), ce vecteur définit un filtre transverse causal dont les a_i constituent la réponse percussionnelle.

Le calcul de \mathbf{a}_p se fait en appliquant le principe d'orthogonalité précisant que $X[n] - \hat{X}_p[n]$ est orthogonal à $\mathbf{X}[n]$, ce qui donne l'équation normale

$$(A.2) \quad \mathbf{R}_p \mathbf{a}_p = \mathbf{c}_p$$

où \mathbf{R}_p et \mathbf{c}_p sont définis par

$$(A.3) \quad \mathbf{R}_p = \text{E}[\mathbf{X}\mathbf{X}^T]; \quad \mathbf{c}_p = \text{E}[X(n)\mathbf{X}].$$

Ainsi la matrice \mathbf{R}_p est la matrice de covariance du vecteur aléatoire \mathbf{X} : c'est donc une matrice symétrique et définie non négative. Comme les éléments de matrice R_{ij} valent $\gamma[i-j]$, on voit que \mathbf{R}_p est une matrice de Toeplitz symétrique définie par sa première ligne, soit

$$(A.4) \quad \mathbf{R}_p = \text{T}\{\gamma[0], \gamma[1], \dots, \gamma[p-1]\}.$$

Le vecteur \mathbf{c}_p , appelé vecteur de corrélation, peut s'écrire

$$(A.5) \quad \mathbf{c}_p = \{\gamma[1], \gamma[2], \dots, \gamma[p]\}^T.$$

Enfin l'erreur de prédiction à un pas et passé fini qui est la variance de l'innovation $X[n] - \hat{X}_p[n]$ s'écrit

$$(A.6) \quad \epsilon_p^2 = \gamma[0] - \mathbf{a}_p^T \mathbf{c}_p.$$

La résolution de (A.2) peut se faire de manière récursive au moyen de l'algorithme de Levinson qui s'écrit

$$(A.7) \quad \mathbf{a}_{m-1}^m = \mathbf{a}_{m-1} - k_m \mathbf{a}_{m-1}^{(-)}$$

$$(A.8) \quad a_m^m = k_m.$$

Dans cette équation \mathbf{a}_{m-1} est le vecteur de régression calculé à l'étape $m-1$, $\mathbf{a}^{(-)}$ est le vecteur déduit de \mathbf{a} en changeant l'ordre des composantes et \mathbf{a}_{m-1}^m le vecteur obtenu en prenant les $m-1$ premières composantes de \mathbf{a}_m , la dernière étant a_m^m . Le coefficient k_m est dénommé coefficient de réflexion et se calcule aussi par un algorithme récursif. De plus l'erreur de prédiction prend la forme récursive

$$(A.9) \quad \epsilon_i^2 = \epsilon_{i-1}^2 (1 - k_i^2)$$

ce qui implique que $k_i^2 \leq 1$. Enfin il convient de noter que si $k_i^2 < 1$, $1 \leq i < p$, l'algorithme de Levinson peut être inversé, ce qui permet de calculer les vecteurs \mathbf{a}_p ou \mathbf{c}_p à partir d'un vecteur de réflexion k_p .

En faisant croître p indéfiniment on réalise la prédiction à un pas et passé infini $\hat{X}[n]$. On sait qu'il s'agit d'un bruit blanc de variance ϵ^2 , erreur de prédiction, qui compte tenu de (A.9) est donnée par (2.22).

Modèles autorégressifs

Un signal autorégressif d'ordre r , AR(r), est caractérisé par le fait que l'erreur de prédiction atteint sa valeur asymptotique pour $p = r$. Autrement dit, tous les coefficients de réflexion k_i sont nuls pour $i > r$. En poursuivant alors l'algorithme de Levinson (A.7) et (A.8), on voit que les vecteurs de régression d'ordre supérieur à r ont leurs dernières composantes nulles. On a donc $\hat{X}_p = \hat{X}$, prédiction à un pas et passé infini, et comme l'innovation est un bruit blanc de variance ϵ^2 on peut écrire un signal AR(r) sous la forme

$$(A.10) \quad X[n] = \mathbf{a}^T \mathbf{X}[n] + U[n]$$

où \mathbf{a} est le vecteur de régression, $\mathbf{X}[n]$ le vecteur de composantes $X[n-i]$, $1 \leq i \leq r$, et $U[n]$ un bruit blanc de variance ε^2 .

On en déduit en particulier la densité spectrale qui peut se mettre sous la forme $\Gamma(e^{2\pi i\nu})$ avec

$$(A.11) \quad \Gamma(z) = \varepsilon^2 H(z) H(z^{-1})$$

$$(A.12) \quad H(z) = \left[1 - \sum_{i=1}^r a_i z^{-i} \right]^{-1}.$$

Enfin on peut noter que la fonction de corrélation d'un modèle AR(r) se calcule aisément. Les $r+1$ premières valeurs se déduisent des vecteurs équivalents \mathbf{a} ou \mathbf{k} . Les valeurs suivantes satisfont

$$(A.13) \quad \gamma(m) = a_1 \gamma[m-1] + a_2 \gamma[m-2] + \dots + a_r \gamma[m-r].$$

Pour achever ceci montrons un des résultats utilisés dans le texte.

Proposition : Une condition nécessaire et suffisante pour qu'une suite de nombres $\gamma[i]$, $0 \leq i \leq r$, soient les $r+1$ valeurs d'une fonction de corrélation d'un signal régulier est que la matrice de Toeplitz symétrique dont la première ligne est $\gamma[0]$, $\gamma[1]$, ..., $\gamma[r]$, soit définie positive.

Démonstration : La condition est évidemment nécessaire. En effet la matrice considérée est la matrice de covariance du vecteur aléatoire de composantes $X[n]$, $X[n-1]$, ..., $X[n-r]$ qui est définie positive. Réciproquement partant d'une telle matrice et du vecteur \mathbf{c} qu'on lui associe on peut résoudre (A.2)

et déduire une suite de coefficients k_i tous inférieurs à un. En prolongeant alors la suite par des k_i nuls, on obtient un signal autorégressif d'ordre r . On a ainsi une fonction de corrélation dont les $r+1$ premières valeurs sont données.

Manuscrit reçu le 2 avril 1990

Remerciements

Nous remercions tous ceux qui ont contribué à l'amélioration de ce texte, et en particulier C. Bendjaballah, P. Brémaud, J. L. Lacoume, R. Levernier et B. Lumeau.

BIBLIOGRAPHIE

- [1] A. PAPOULIS, « Maximum entropy and spectral estimation: a review », *IEEE Trans. Acoust. Speech, Signal Processing*, ASSP-29, pp. 1176-1186, 1981.
- [2] T. BERGER, *Rate distortion theory*, Prentice Hall, 1971.
- [3] B. PICINBONO and J. M. KERILIS, « Some properties of prediction and interpolation errors », *IEEE Trans. Acoust. Speech, Signal Processing*, ASSP-36, pp. 525-531, 1988.
- [4] J. GRANDELL, M. HAMRUD and P. TOLL, « A remark on the correspondence between the maximum entropy method and the autoregressive model », *IEEE Trans. Inform. Theory*, IT-30, pp. 377-380, 1980.
- [5] B. PICINBONO and M. BENIDIR, « Some properties of lattice autoregressive filters », *IEEE Trans. Acoust. Speech, Signal Processing*, ASSP-34, pp. 342-349, 1986.