

MODELE DE SIGNAUX AUTOREGRESSIF A GAIN
DEPENDANT DE L'ETAT

JM Vesin

Laboratoire de Traitement des Signaux
Ecole polytechnique Fédérale de Lausanne
1015 Lausanne, Suisse

RÉSUMÉ

ABSTRACT

Nous présentons dans cet exposé un modèle se composant d'une partie AR linéaire et d'un terme nonlinéaire, multiplicatif sur l'excitation, et fonction des valeurs passées du signal. Ce modèle apparaît comme une alternative aux modèles AR à excitation non gaussienne et aux modèles nonlinéaires "classiques". Nous dérivons une condition suffisante générale de stabilité et proposons une méthode sous-optimale d'estimation des paramètres. Nous présentons ensuite une application à des signaux réels.

In this paper we present a signal model composed of a linear AR part and of a nonlinear term, multiplicative on the excitation, and function of the past values of the signal. This model appears as an alternative to linear AR models with non Gaussian input and to "classical" nonlinear models. We derive a general sufficient condition for stability and we propose a sub-optimal parameter estimation scheme. We present then an application to real signals.

1. INTRODUCTION

Pour de nombreuses raisons les modèles de signaux développés initialement ont été le plus souvent linéaires à entrée gaussienne blanche. Il apparaît que l'abandon de l'une ou des deux hypothèses (linéarité et caractère gaussien) permet d'obtenir une modélisation plus fine de nombre de processus existants. Si l'excitation n'est plus supposée gaussienne, l'estimateur aux moindres carrés des paramètres d'un modèle AR ne constitue plus une bonne approximation de l'estimateur au maximum de vraisemblance si des longueurs de signal "raisonnables" sont disponibles. L'estimation au maximum de vraisemblance doit donc être utilisée directement, l'amélioration des performances se faisant au prix d'un surcroît de calcul (p.e. [1] excitation à distribution de type mélange gaussien). Une question non triviale posée par cette approche est celle du choix *a priori* de la distribution de l'excitation. De nombreux modèles nonlinéaires ont également été proposés dans la littérature, qu'ils soient "spécialisés" comme le modèle AR à seuil [2] ou plus généraux comme les modèles NARMAX [3] ou bilinéaire [4]. Il est à remarquer que tous ces modèles sont décrits, $\{x_n\}$ étant la série temporelle étudiée, par une équation du type:

$$x_n = F(x_{n-1}, \dots, x_{n-p}, \epsilon_{n-1}, \dots, \epsilon_{n-q}) + \epsilon_n \quad (1)$$

où $F(\cdot)$ est une fonction nonlinéaire, le plus souvent estimée par une technique aux moindres carrés, et $\{\epsilon_n\}$ le bruit exciteur. Ces modèles sont donc plus spécifiquement orientés à des fins de *prédiction* que de *modélisation*, ceci d'autant plus que, si la fonction $F(\cdot)$ est estimée à l'aide de polynômes, le modèle décrit

par l'éq.(1) ne permettra pas de générer un signal du fait qu'il sera le plus souvent instable [5].

Nous tentons dans cet exposé de concilier les deux approches non-gaussienne et nonlinéaire avec un modèle décrit par:

$$x_n = \sum_{i=1}^p a_i x_{n-p} + g(x_{n-1}) \epsilon_n \quad (2)$$

dans lequel $g(\cdot)$ est une fonction que nous supposerons positive sans perte de généralité et ϵ_n un bruit blanc gaussien de variance unité.

2. STABILITE DU MODELE

L'éq.(2) peut être réécrite sous la forme:

$$X_n = AX_{n-1} + g(X_{n-1}^T e_1) \epsilon_n \cdot e_1 \quad (3)$$

avec $X_n = [x_n, \dots, x_{n-p+1}]^T$, $e_1 = [1, 0, \dots, 0]^T$, et A la matrice compagnon du système. Du fait que $\{\epsilon_n\}$ est blanc, $\{X_n\}$ est clairement markovien. Le caractère gaussien de $\{\epsilon_n\}$ permet, si $g(\cdot)$ est supposée continue, d'appliquer le corollaire 5.4 de Tweedie [6] sur la récurrence des chaînes de Markov. Ce corollaire stipule que, γ_y étant définie par:

$$\gamma_y = E\{\|X_n\| - \|X_{n-1}\| \mid X_{n-1} = y\} \quad (4)$$

une condition suffisante pour la stabilité de (3) est l'existence de $\alpha > 0$ tel que $\gamma_y \leq 0$ pour tout y tel que $\|y\| > \alpha$. L'argument de l'espérance mathématique dans l'éq.(4) fournit l'inégalité suivante:



$$\begin{aligned} \|X_n\| - \|X_{n-1}\| &= \|X_{n-1}\| \left(\frac{\|X_n\|}{\|X_{n-1}\|} - 1 \right) \\ &\leq \|y\| \left(\frac{\|Ay\|}{\|y\|} + \frac{g(y^T e_1)}{\|y\|} |\varepsilon_n| - 1 \right) \end{aligned} \quad (5)$$

Le premier terme de l'expression entre parenthèses du terme de droite de l'inégalité (5) peut être rendu arbitrairement petit, si la transformation linéaire est contractive, c.à.d. si les racines de l'équation caractéristique de A sont à l'intérieur du cercle unité. Le deuxième terme peut également être rendu arbitrairement petit sous la condition:

$$\frac{g(x)}{x} \rightarrow 0, |x| \rightarrow \infty \quad (6)$$

ce qui se vérifie aisément si la norme utilisée est:

$$\|X_n\| = \sup_{k \in (n, n-p+1)} |x_k| \quad (7)$$

La somme entre parenthèses peut donc être rendue négative, son espérance mathématique négative ou nulle (et donc γ_y) si la partie AR linéaire du modèle vérifie la condition classique de stabilité et si $g(\cdot)$ vérifie l'éq.(6). Notons que si $g(\cdot)$ n'est pas continue le théorème de Tweedie n'est pas applicable. Mais si $g(\cdot)$ est bornée l'entrée du modèle décrit par l'éq.(2) sera statistiquement bornée en amplitude du fait du caractère gaussien de $\{\varepsilon_n\}$, et la stabilité dépendra uniquement de celle de la partie AR.

3. ESTIMATION DES PARAMETRES

3.1 Choix de la méthode d'estimation

L'estimation des paramètres du modèle de l'éq.(2) devrait théoriquement être faite selon un critère de maximum de vraisemblance sur les N échantillons (x_1, \dots, x_N) , ce qui peut ne pas être une tâche facile. Nous avons préféré utiliser une stratégie sous-optimale consiste à estimer tout d'abord la partie AR linéaire à l'aide d'une technique aux moindres carrés, puis à calculer les résidus et à estimer $g(\cdot)$ sur ceux-ci par maximum de vraisemblance. Dans cette optique, la fonction $g(\cdot)$ apparaît comme un moyen de corriger les résidus si ceux-ci sont manifestement non gaussiens mais décorrélés sans avoir à estimer leur densité de probabilité, et devrait permettre de synthétiser des signaux de même type à partir d'un séquence gaussienne blanche.

3.2 Estimation de la partie AR

Nous avons de plus employé ici une méthode aux moindres carrés particulière ayant l'intérêt de permettre l'adjonction éventuelle de termes non linéaires dans la partie AR (la condition de stabilité de la section 2. impliquant alors une condition de stabilité sur le système AR non linéaire. Nous nous limiterons dans le présent exposé à des termes linéaires). Il s'agit de la méthode d'identification de systèmes NARMAX présentée dans [7] et qui est basée sur le principe suivant:

si on suppose que le modèle peut être exprimé par:

$$x_n = \sum_{m=0}^M \phi_m p_m(n) + \rho_n \quad (8)$$

dans lequel les $p_m(n)$ représentent les diverses composantes, linéaires et non linéaires, de l'autorégression, c.à.d. $p_0(n)=1$,

$p_1(n)=x_{n-1}, \dots, p_m(n)=f(x_{n-s}, \dots, x_{n-q})$, le but recherché est d'obtenir une nouvelle représentation de la forme:

$$x_n = \sum_{m=0}^M c_m w_m(n) + \rho_n \quad (9)$$

dans laquelle les $w_m(n)$ sont orthogonaux. Du fait de cette orthogonalité il est facile d'obtenir les estimées \hat{c}_m des coefficients, et à partir de celles-ci les estimées $\hat{\phi}_m$ du modèle de l'éq.(8) (dans notre cas les \hat{a}_i). L'orthogonalité permet également pour chaque terme de calculer la réduction d'erreur quadratique moyenne qu'il entraîne, ce qui est utile pour sélectionner au moins approximativement l'ordre p de l'autorégression. En effet, les résidus n'étant pas gaussiens, un critère de type Akaike [8] n'est pas applicable.

3.3 Estimation paramétrique de $g(\cdot)$

Une fois que les estimées $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_p$ des paramètres AR ont été obtenues, les résidus sont calculés avec:

$$\rho_n = x_n - \sum_{i=1}^p \hat{a}_i x_{n-i} \quad n = p+1, \dots, N \quad (10)$$

Si nous supposons maintenant que ces résidus sont de la forme:

$$\rho_n = g(x_{n-1}) \varepsilon_n \quad (11)$$

et que $g(\cdot)$ est paramétrée par un ensemble θ , nous pouvons dire que, pour θ fixé, les $\{\rho_n\}$ sont des transformations linéaires des $\{\varepsilon_n\}$ gaussiens indépendants, et sont donc gaussiens, indépendants, de variance $g^2(x_{n-1})$. Leur densité de probabilité conjointe peut s'écrire:

$$p\{\rho_{p+1}, \dots, \rho_N | \theta\} = \prod_{i=p+1}^N \frac{1}{\sqrt{2\pi}g(x_{i-1}; \theta)} \exp\left(-\frac{\rho_i^2}{2g^2(x_{i-1}; \theta)}\right) \quad (12)$$

Une estimée $\hat{\theta}$ sera donc obtenue en maximisant la partie droite de l'éq.(12) ou encore, prenant l'opposé du logarithme de cette expression et en supprimant les constantes additives, en minimisant la fonction:

$$L(\theta) = \frac{1}{2} \sum_{i=p+1}^N \left[\frac{\rho_i^2}{g^2(x_{i-1}; \theta)} + 2 \log[g(x_{i-1}; \theta)] \right] \quad (13)$$

Notons qu'une paramétrisation de $g(\cdot)$ par une fonction polynômiale n'est guère judicieuse, un tel choix ne vérifiant pas la condition de stabilité développée à la section 2.

4. APPLICATION A DES SIGNAUX REELS

4.1 Présentation des signaux

Nous avons appliqué ce modèle à des signaux de $N=2000$ échantillons provenant de détecteurs de variation de température placés dans des salles contenant une source de chaleur ou possédant une ouverture sur l'extérieur. Ces signaux, visiblement non gaussiens (figures 1.a et 3.a), se caractérisent par la présence d'excursions en amplitude dues à l'intermittence des turbulences générées et un aspect clairement non symétrique dû à la prédominance de variations de température soit négatives soit positives.

4.2 Paramétrisation de $g(\cdot)$

Comme nous l'avons dit plus haut, une paramétrisation polynômiale de $g(\cdot)$ n'est pas souhaitable. Nous avons utilisé dans le cadre de cette étude deux types particuliers de représentation basés sur des considérations heuristiques. En effet, supposons que la fonction $g(\cdot)$ soit de la forme:

$$g(x) = \begin{cases} c_1 & |x| < t \\ c_2 & |x| \geq t \end{cases} \quad (14)$$

avec $c_1 \ll c_2$. Dans le cas où t est relativement grand vis-à-vis de un, le signal généré sera pratiquement gaussien, car l'amplitude de x_n ne dépassera qu'exceptionnellement t . Par contre, si la valeur de t est diminuée, la probabilité que l'amplitude de x_n dépasse cette valeur sera non négligable, et ces excursions donneront lieu à l'apparition de pointes. D'autre part, si $g(\cdot)$ est non symétrique, le signal généré sera lui aussi non symétrique. Ces deux effets permettant de rendre compte de l'allure de nos signaux, nous avons utilisé deux représentations de $g(\cdot)$ les incorporant. La première est décrite par:

$$g(x) = \begin{cases} c_1, & x < -\frac{1}{2} \\ c_2, & -\frac{1}{2} \leq x \leq \frac{1}{2} \\ c_3, & x > \frac{1}{2} \end{cases} \quad (15)$$

la seconde par:

$$g(x) = g_1(x) \cdot g_2(x) \quad (16)$$

avec:

$$g_1(x) = \mu_1 - \alpha_1 \exp[-\beta_1 (x - \gamma_1)^2] \quad (17)$$

$$g_2(x) = \mu_2 + \alpha_2 \text{ArcTan}[\beta_2 (x - \gamma_2)] \quad (18)$$

La condition de positivité sur $g(\cdot)$ est facile à vérifier à partir des valeurs des paramètres dans les deux cas. La première représentation, non continue, est bornée. L'estimation de ses paramètres par la minimisation de $L(\theta)$ (éq.(13)) peut être réalisée par une méthode de type "recuit simulé" (simplifiée). La deuxième représentation vérifie la condition de stabilité de la section 2., et l'estimation des paramètres peut se faire par une technique numérique standard de type gradient.

4.3 Résultats

Pour les deux signaux les calculs ont été opérés sur les $N=2000$ échantillons disponibles. La méthode d'estimation de la partie AR présentée à la section 3.2 nous a conduit à sélectionner un ordre $p=5$ pour les deux signaux.

Nous avons ensuite synthétisé une séquence à partir des paramètres du signal de la figure 1.a, en utilisant la représentation de la fonction $g(\cdot)$ décrite par les éq.(16), (17) et (18) et un bruit blanc gaussien simulé (l'autre représentation donne des résultats tout à fait similaires). Le signal généré est montré en figure 1.b. A la figure 2.a se trouvent les estimées des fonctions d'autocorrélation des deux signaux réel et simulé, normalisées toutes deux par la puissance du signal réel, et à la figure 2.b leurs histogrammes calculés sur 50 intervalles. La même opération a été réalisée pour le signal de la figure 3.a en utilisant la représentation

de $g(\cdot)$ définie par l'éq.(15). Les résultats correspondants se trouvent respectivement aux figures 3.b, 4.a et 4.b. Les signaux générés ont bien approximativement les mêmes caractéristiques spectrales et statistiques que les signaux réels.

CONCLUSION

Nous avons présenté dans cet article un modèle de signal constituant une alternative possible aux modèles non linéaires classiques et aux modèles AR linéaires à excitation non gaussienne, et donnant de bons résultats sur des signaux réels. Ce modèle pourrait être amélioré en considérant une partie AR non linéaire et en affinant l'estimation de la fonction de gain $g(\cdot)$.

REMERCIEMENTS

Nous tenons à remercier la compagnie Cerberus, Zürich, Suisse, et le Comité d'Encouragement à la Recherche Scientifique Suisse (CERS) pour leur support.

REFERENCES

- [1] D. Sengupta and S.M. Kay, "Efficient Estimation of Parameters for Non-Gaussian Autoregressive Processes," IEEE Trans. Acoust., Speech, and Signal Processing, Vol. ASSP-37, No. 6, pp. 785-794, June 1989.
- [2] H. Tong and K.S. Lim, "Threshold Autoregression, Limit Cycles and Cyclical Data," J. Roy Statist. Soc. Ser. B, Vol. 42, pp. 245-292, 1980.
- [3] S. Chen and S.A. Billings, "Representation of Non-linear Systems: the NARMAX Model," Int. J. Control, Vol. 49, pp. 1013-1032, 1989.
- [4] T. Subba Rao and M.M. Gabr, An Introduction to Bispectral Analysis and Bilinear Time Series Models, Springer-Verlag, Berlin, 1984.
- [5] S. Chen and S.A. Billings, "Modelling and Analysis of Non-linear Time Series," Int. J. Control, Vol. 50, No 6, pp. 2151-2171, 1989.
- [6] R.L. Tweedie, "Sufficient Conditions for Ergodicity and Recurrence of Markov Chains on a General State Space," Stochastic Process. Appl., Vol. 3, pp. 385-403, 1975.
- [7] M. Korenberg, S.A. Billings, Y.P. Liu and P.J. McIlroy, "Orthogonal Parameter Estimation Algorithm for Non-linear Stochastic Systems," Int. J. Control, Vol.48, No. 1, pp. 193-210, 1988.
- [8] H. Akaike, "A New Look at the Statistical Model Identification," IEEE Trans. Autom. Contr., Vol. AC-19, pp. 716-723, 1974.

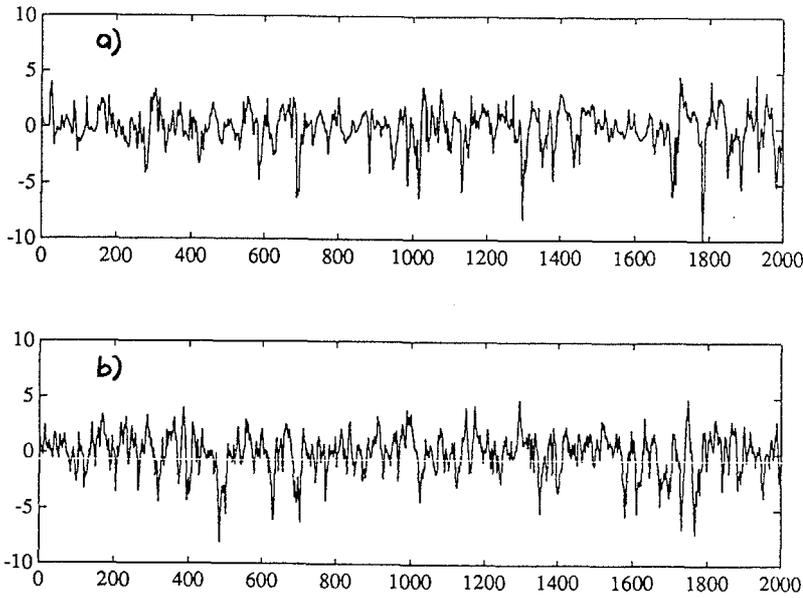


Figure 1. a) Signal réel b) Signal simulé

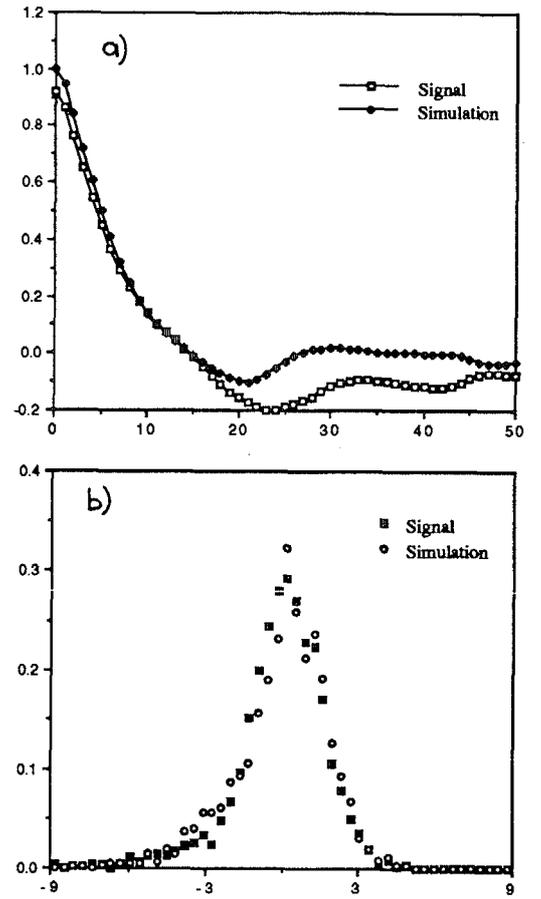


Figure 2. a) Estimées des autocorrélations b) Histogrammes

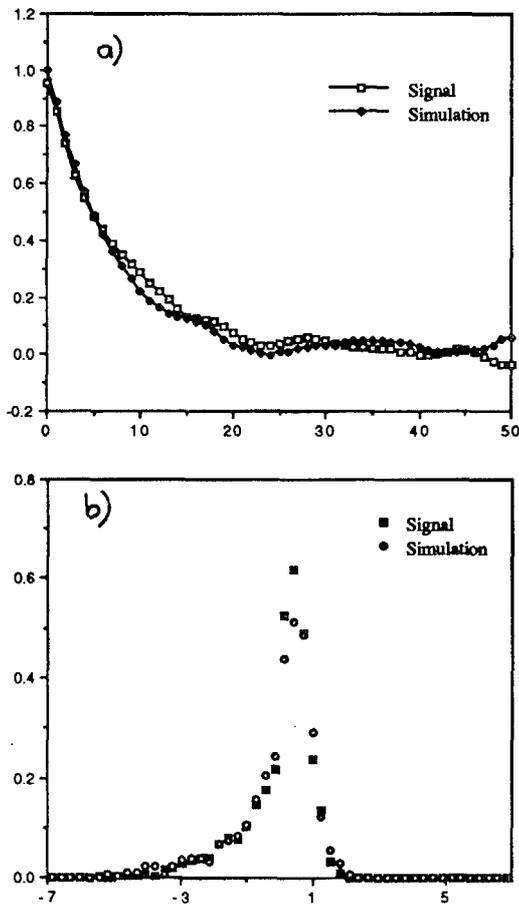


Figure 4. a) Estimées des autocorrélations b) Histogrammes

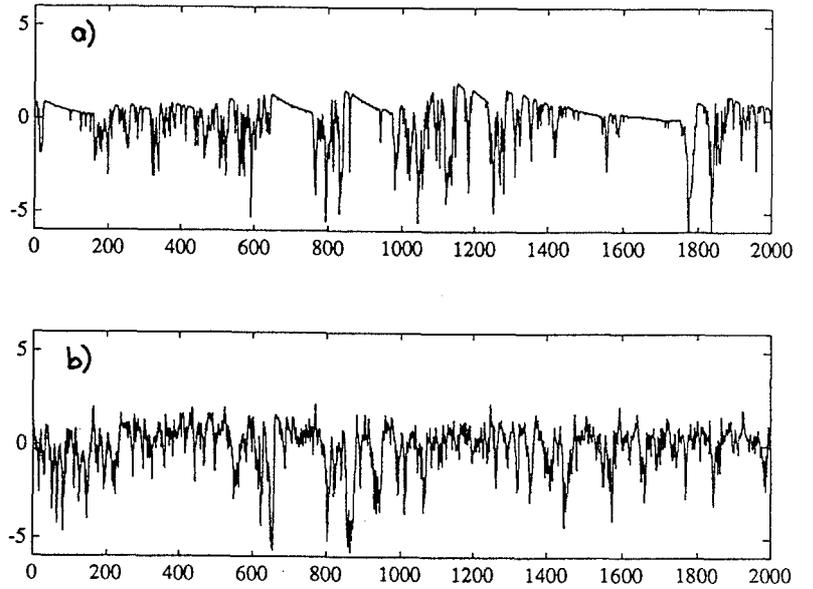


Figure 3. a) Signal réel b) Signal simulé