

Comparaison d'estimateurs de fréquence à complexité algorithmique réduite

Guy PLANTIER, Anthony SOURICE, Alain LE DUFF, Daniel SCHANG

École Supérieure d'Électronique de l'Ouest
4, rue Merlet de la Boulaye, B.P. 926, 49009 Angers Cedex 01, France
Tél : 02 41 86 67 67, Fax : 02 41 87 99 27
guy.plantier@eseo.fr, anthony.sourice@eseo.fr
alain.le_duff@eseo.fr, daniel.schang@eseo.fr

Résumé –

De nombreux algorithmes, basés sur une modélisation Auto Régressive du signal, ont été proposés pour des problèmes d'estimation de fréquence de signaux périodiques. Nous nous intéressons ici aux performances statistiques de tels estimateurs, et proposons des formules approchées du biais et de la variance des estimés. Les résultats obtenus permettent de mettre en évidence l'influence de la fréquence recherchée, du rapport signal sur bruit et du nombre de points sur les performances de l'estimateur.

Abstract –

A lot of techniques, based on autoregressive models, have been proposed to solve frequency estimation problems. We focus on the statistical behavior of those algorithms and propose approximate expressions for the bias and variance of estimation. By the way, we point out the influence of the values of the frequency to be estimated, the number of points and the signal to noise ratio on the performance of those estimators.

1 Introduction

La mise au point de certains capteurs passe souvent par l'estimation de la fréquence d'un signal périodique. Un très grand nombre de méthodes a été proposé dans la littérature pour résoudre ce problème [3], [4]. Cependant, si l'on souhaite réduire les coûts du capteur et des systèmes électroniques de traitement du signal, il est indispensable de sélectionner les algorithmes les plus simples, tout en conservant une précision acceptable des mesures. L'objectif du présent travail est de comparer les propriétés statistiques de quatre estimateurs de fréquence à complexité algorithmique réduite et de proposer pour chacun d'eux des formules approchées, mais simples à interpréter, donnant avec suffisamment de précision le biais et la variance d'estimation. Les quatre estimateurs étudiés sont basés sur les méthodes de l'autocorrélation biaisée de Yule Walker (B), de l'autocorrélation non biaisée (NB), de la covariance (COV) et de la covariance modifiée (CM).

De nombreux travaux ont montré l'efficacité des méthodes paramétriques pour l'estimation de fréquence [1]. Ces méthodes sont souvent supérieures aux méthodes dites classiques (périodogramme, Welch, etc...), lorsque l'on ne dispose que d'un faible nombre d'échantillons N . L'étude proposée en [6] montre que la fréquence estimée est fonction de la phase à l'origine et de N . Par ailleurs, la variance de la fréquence correspondant au maximum du spectre obtenu par analyse autorégressive d'ordre élevé du signal est approximativement inversement proportionnelle au carré du rapport signal sur bruit (RSB) et de N [5]. L'influence de la fréquence d'échantillonnage, sur le biais d'estimation, a été étudiée en [2].

Dans ce travail, nous proposons des résultats nouveaux concernant les propriétés statistiques des quatre méthodes d'estima-

tion précitées. Pour chacun des estimateurs, nous proposons une expression analytique approchée pour le biais et la variance de l'estimation, en fonction de la fréquence recherchée, du nombre de points utilisés et du rapport signal sur bruit.

2 Modèle du signal

L'étude des propriétés statistiques des quatre estimateurs cités précédemment (B, NB, COV et CM) est basée sur l'utilisation de N observations d'un signal réel $x(k)$, cosinusoidal, de fréquence F_0 , échantillonné à la fréquence F_e et noyé dans un bruit blanc additif gaussien $b(k)$:

$$x(k) = A \cos(2\pi f_0 k + \phi) + b(k), \quad (1)$$

où $k \in [0..N - 1]$, ϕ est une phase initiale aléatoire uniformément répartie dans $[0, 2\pi]$, $f_0 = \frac{F_0}{F_e}$ est la fréquence réduite du signal. On considère le cas où $b(k)$ est à moyenne nulle ($\mu_b = 0$) et de variance σ_b^2 . On notera par la suite $RSB = \frac{A^2}{2\sigma_b^2}$: le rapport Signal sur Bruit.

Dans ce travail, nous proposons une analyse autorégressive d'ordre 2 de $x(k)$ basée sur l'équation aux différences suivante [3], [4] :

$$x(k) + a_1 x(k-1) + a_2 x(k-2) = n(k), \quad (2)$$

où $n(k)$ représente l'erreur de prédiction du modèle. Ce filtre prédictif est caractérisé par la fonction de transfert en z suivante :

$$G(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}}.$$

L'estimation de la fréquence est alors donnée par l'argument

des pôles complexes de $G(z)$:

$$\hat{f}_0 = \frac{1}{4} + \frac{1}{2\pi} \arcsin \left(\frac{\hat{a}_1}{2\sqrt{\hat{a}_2}} \right), \quad (3)$$

où \hat{a}_1 et \hat{a}_2 désignent les estimations des coefficients autorégressifs a_1 et a_2 .

3 Démarche méthodologique

Les coefficients \hat{a}_1 et \hat{a}_2 estimés à partir des mesures de $x(k)$ sont obtenus par la résolution d'un système d'équation du type Yule Walker d'ordre 2 :

$$R \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = -\underline{r}, \quad (4)$$

où R et \underline{r} désignent respectivement la matrice et le vecteur d'autocorrélation du signal $x(k)$.

Les différents algorithmes étudiés diffèrent entre eux par les moyens de calculer les termes R et \underline{r} .

L'objectif de ce travail est de trouver une expression approchée du biais et de la variance d'estimation pour chacune des méthodes citées. Plus précisément, nous nous intéressons au cas où f_0 est proche de 0.25, où $\epsilon = \frac{1}{N}$ est faible et où le rapport signal sur bruit est élevé. Dans ces conditions, le terme $\frac{\hat{a}_1}{2\sqrt{\hat{a}_2}}$ est proche de 0, et la fonction arcsinus de l'équation (3) peut être approchée par

$$\arcsin(x) \approx x + O(x^3).$$

L'équation (3) devient alors :

$$\hat{f}_0 \approx \frac{1}{4} + \frac{1}{4\pi} \left(\frac{\hat{a}_1}{\sqrt{\hat{a}_2}} \right). \quad (5)$$

Dans ce travail, nous proposons pour chacune des 4 méthodes précitées, le calcul analytique des estimations de l'autocorrélation de $x(k)$. Les résultats font apparaître un terme déterministe correspondant à l'autocorrélation théorique de $x(k)$, auquel s'ajoute un terme aléatoire dépendant de N , ϕ et $b(k)$. Ces erreurs d'estimation se traduisent par une solution aléatoire de l'équation (4) qui, introduite en (5), donne pour f_0 une estimation bruitée.

Ainsi, nous obtenons pour les quatre algorithmes, un développement limité simple de la fréquence d'estimation \hat{f}_0 , en fonction des données du problème : f_0 , N , RSB, ϕ et $b(k)$.

Nous présentons ci-après l'ensemble de cette démarche dans le cas de la méthode de l'autocorrélation biaisée. La transposition de ce travail aux cas NB, COV et CM suit la même logique et n'est pas détaillée par la suite.

4 Étude du biais

4.1 Méthode de l'autocorrélation Biaisée (B)

Le système à résoudre, équivalent à l'équation (4), est le suivant [3] :

$$\begin{bmatrix} \hat{\varphi}_x^B(0) & \hat{\varphi}_x^B(1) \\ \hat{\varphi}_x^B(-1) & \hat{\varphi}_x^B(0) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = - \begin{bmatrix} \hat{\varphi}_x^B(1) \\ \hat{\varphi}_x^B(2) \end{bmatrix} \quad (6)$$

où $\hat{\varphi}_x^B(k)$ représente l'estimation biaisée de l'autocorrélation en k :

$$\begin{aligned} \hat{\varphi}_x^B(k) &= \frac{1}{N} \sum_{l=0}^{N-|k|-1} x(l)x(k+l) & \text{pour } k \geq 0 \\ \hat{\varphi}_x^B(k) &= \hat{\varphi}_x^B(-k) & \text{pour } k < 0. \end{aligned}$$

En introduisant l'expression de $x(k)$ dans $\hat{\varphi}_x^B(k)$, il vient :

$$\begin{aligned} \hat{\varphi}_x^B(0) &= \frac{A^2}{2} + \frac{A^2}{2N} \frac{k_1}{\sin(\omega_0)} + \sigma_b^2 + b_1 \\ \hat{\varphi}_x^B(1) &= \frac{A^2}{2} \frac{N-1}{N} \cos(\omega_0) + \frac{A^2}{2N} \frac{k_2}{\sin(\omega_0)} + b_2 \\ \hat{\varphi}_x^B(2) &= \frac{A^2}{2} \frac{N-2}{N} \cos(2\omega_0) + \frac{A^2}{2N} \frac{k_3}{\sin(\omega_0)} + b_3 \end{aligned}$$

avec $\omega_0 = 2\pi f_0$ et $k_i = \cos[\omega_0(N-1) + 2\phi] \sin((N+1-i)\omega_0)$.

Les k_i et les b_i sont des variables aléatoires à moyenne nulle, caractérisant les erreurs d'estimation des autocorrélations. Les k_i dépendent de la phase aléatoire et du nombre de points N , et les b_i , indépendants des k_i , sont relatifs à la présence du bruit additif.

Un développement limité de la fréquence estimée \hat{f}_0 autour de $f_0 = 0.25$, de $\epsilon = 0$ et de $x = \frac{\text{RSB}}{1+\text{RSB}} = 1$ est donné par

$$\begin{aligned} \hat{f}_0 \approx & \frac{1}{4} + (f_0 - \frac{1}{4}) \left(\frac{\text{RSB}}{1+\text{RSB}} - \frac{1}{N} \right) \\ & - \frac{1}{2\pi} \left(\frac{k_2}{N} + \frac{2b_2}{A^2} \frac{\text{RSB}}{1+\text{RSB}} \right). \end{aligned} \quad (7)$$

Cette expression fait apparaître un terme déterministe et un terme aléatoire dépendant uniquement des termes à moyenne nulle l_2 et k_2 . Le biais de \hat{f}_0 , défini par $\text{Biais}_B = E[\hat{f}_0] - f_0$, est alors donné par :

$$\text{Biais}_B \approx \left(f_0 - \frac{1}{4} \right) \left(\frac{\text{RSB}}{1+\text{RSB}} - \frac{1}{N} - 1 \right). \quad (8)$$

Une analyse statistique de type Monte Carlo a permis de valider cette expression approchée. La FIGURE 1 montre par exemple les résultats obtenus pour un rapport signal sur bruit de 10dB. Ce résultat fait apparaître une très bonne concordance entre les simulations de Monte Carlo et la formule approchée pour une large plage de valeurs de f_0 et de N .

4.2 Méthode de l'autocorrélation Non Biaisée (NB)

La démarche proposée précédemment reste quasiment identique pour le cas de l'autocorrélation non biaisée. Le système à résoudre dans ce cas est le suivant [3] :

$$\begin{bmatrix} \hat{\varphi}_x^{NB}(0) & \hat{\varphi}_x^{NB}(1) \\ \hat{\varphi}_x^{NB}(-1) & \hat{\varphi}_x^{NB}(0) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = - \begin{bmatrix} \hat{\varphi}_x^{NB}(1) \\ \hat{\varphi}_x^{NB}(2) \end{bmatrix} \quad (9)$$

où $\hat{\varphi}_x^{NB}(k)$ désigne l'estimation non biaisée de l'autocorrélation en k donnée par

$$\begin{aligned} \hat{\varphi}_x^{NB}(k) &= \frac{1}{N-|k|} \sum_{l=0}^{N-|k|-1} x(l)x(k+l) & \text{pour } k \geq 0 \\ \hat{\varphi}_x^{NB}(k) &= \hat{\varphi}_x^{NB}(-k) & \text{pour } k < 0. \end{aligned}$$

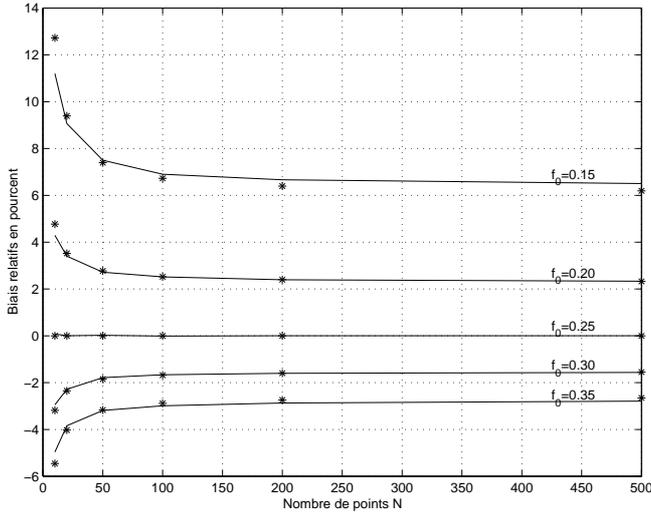


FIG. 1 – *Biases relatifs en pourcent de l'estimation de fréquence par la méthode de l'autocorrélation biaisée, pour un rapport signal sur bruit de 10dB : analyse de Monte Carlo (courbe -) et biais approché par l'équation (8) (courbe *).*

Les résultats obtenus par le développement limité montrent que

$$\text{Biais}_{NB} \approx \left(f_0 - \frac{1}{4} \right) \left(-\frac{7}{8} + \frac{3}{4} \frac{\text{RSB}}{1 + \text{RSB}} + \frac{1}{8} \frac{\text{RSB}^2}{(1 + \text{RSB})^2} \right). \quad (10)$$

4.3 Méthode de la Covariance (COV)

Le système à résoudre est le suivant [3]:

$$\begin{bmatrix} \hat{C}_x^{COV}(1,1) & \hat{C}_x^{COV}(1,2) \\ \hat{C}_x^{COV}(2,1) & \hat{C}_x^{COV}(2,2) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = - \begin{bmatrix} \hat{C}_x^{COV}(1,0) \\ \hat{C}_x^{COV}(2,0) \end{bmatrix},$$

avec,

$$\hat{C}_x^{COV}(k,l) = \frac{1}{N-2} \sum_{n=2}^{N-1} x(n-k)x(n-l).$$

Le développement limité obtenu est donné par :

$$\text{Biais}_{COV} \approx -\frac{1}{1 + \text{RSB}} \left(f_0 - \frac{1}{4} \right) \quad (11)$$

4.4 Méthode de la Covariance Modifiée (CM)

Le système d'équation à résoudre est le suivant [3]:

$$\begin{bmatrix} \hat{C}_x^{CM}(1,1) & \hat{C}_x^{CM}(1,2) \\ \hat{C}_x^{CM}(2,1) & \hat{C}_x^{CM}(2,2) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix} = - \begin{bmatrix} \hat{C}_x^{CM}(1,0) \\ \hat{C}_x^{CM}(2,0) \end{bmatrix},$$

avec,

$$\hat{C}_x^{CM}(k,l) = \frac{1}{2(N-2)} \left[\sum_{n=2}^{N-1} x(n-k)x(n-l) + \sum_{n=0}^{N-3} x(n+k)x(n+l) \right].$$

Soit :

$$\hat{C}_x^{CM}(k,l) = \frac{1}{2} \left[\hat{C}_x^{COV}(k,l) + \hat{C}_x^{COV}(2-k,2-l) \right].$$

Pour cette méthode, les résultats obtenus sont équivalents à ceux obtenus dans le cas de la méthode de la covariance. Ainsi :

$$\text{Biais}_{CM} \approx \text{Biais}_{COV} \quad (12)$$

La FIGURE 2 permet de confronter les tests de Monte Carlo aux résultats du biais de l'estimateur basé sur le calcul des covariances modifiées pour un signal de rapport signal sur bruit de 5dB. L'analyse de Monte Carlo conforte l'équation (12), démontrant par la même occasion l'influence négligeable du nombre de points utilisés sur le biais de l'estimateur pour cette méthode.

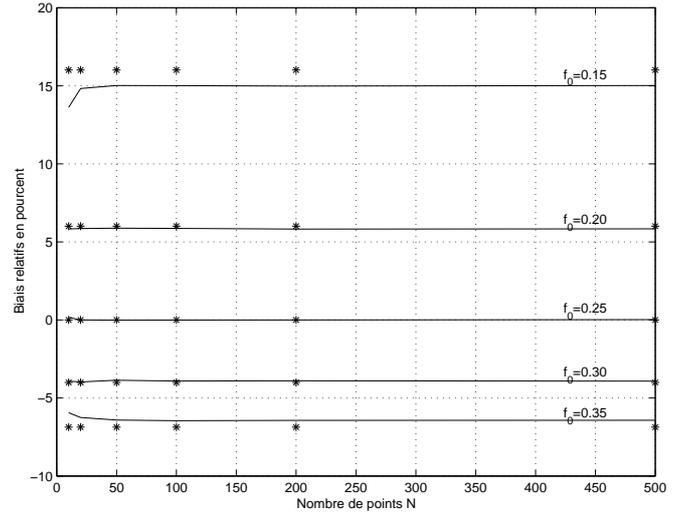


FIG. 2 – *Biases relatifs en pourcent de l'estimation de fréquence par la méthode de la covariance modifiée pour un rapport signal sur bruit de 5dB : analyse de Monte Carlo (courbe -) et biais approché par l'équation (12) (courbe *).*

Au terme de cette première partie du travail, nous disposons pour chacune des quatre méthodes des formules approchées, simples mais précises du biais de l'estimation. Celles-ci nous permettent de caractériser chacun de ces algorithmes dans des configurations données. Les résultats montrent que la méthode de l'autocorrélation biaisée est beaucoup plus sensible au nombre de points N . Enfin, il apparaît clairement que la fréquence recherchée influence notablement sur la qualité de l'estimation. En particulier, nous notons que le biais devient nul pour $f_0 = \frac{1}{4}$, ce qui a été montré auparavant dans [2].

5 Étude de la variance

Nous nous proposons de déterminer une formule approchée de la variance dans le cas optimal où la fréquence est proche de 0.25.

Pour la méthode de l'autocorrélation biaisée, le calcul de la variance de l'estimation de f_0 à partir de l'expression approchée (7) donne :

$$\text{Var}_B \approx \frac{1}{4\pi^2} \left(\frac{\text{Var}_{k_2}}{N^2} + \frac{4\text{Var}_{b_2}\text{RSB}^2}{A^4(1 + \text{RSB})^2} \right),$$

où les termes Var_{b_2} et Var_{k_2} désignent respectivement les variances des termes indépendants k_2 et b_2 de l'expression (7).

La détermination de ces deux termes pour $f_0 \approx 0.25$ donne :

$$\text{Var}_B \approx \frac{1}{4\pi^2 N} \left(\frac{1}{2N} + \frac{1}{(1 + \text{RSB})^2} \right) \quad (13)$$

L'analyse de la méthode de l'autocorrélation non biaisée aboutit à un résultat identique :

$$\text{Var}_{NB} \approx \text{Var}_B \quad (14)$$

L'étude des méthodes de covariance et de covariance modifiée donne :

$$\text{Var}_{COV} \approx \text{Var}_{CM} \approx \frac{1}{4\pi^2 N} \frac{1}{(1 + \text{RSB})^2} \left(1 + \frac{\text{RSB}}{N} \right). \quad (15)$$

Une analyse statistique de la variance des différents estimateurs a été menée. Celle-ci a permis de valider les résultats obtenus ci-dessus : les variances des estimateurs notés B et NB sont équivalentes. Il en va de même pour les variances des estimateurs COV et CM. De plus, pour chacune des quatre méthodes, l'influence de la fréquence recherchée f_0 peut être négligée en première approximation. En effet, les tests de Monte Carlo ont montré que les variances des estimations pour des fréquences différentes de $f_0 = 0.25$ sont au plus supérieures de 2dB à celles données par les formules (13) et (15). La FIGURE 3 présente les variances d'estimation obtenues par analyse de Monte Carlo pour un rapport signal sur bruit de 10dB et une fréquence $f_0 = 0,25$, pour les méthodes de l'autocorrélation biaisée et de la covariance, ainsi que les développements limités associés.

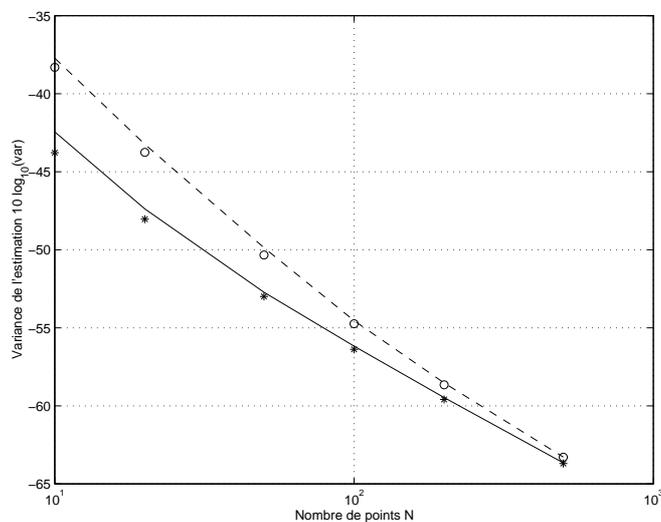


FIG. 3 – *Variances d'estimation de la fréquence f_0 obtenues par tirages de Monte Carlo pour les méthodes de l'autocorrélation biaisée (courbe - -) et de la covariance (courbe -) et variances approchées par les équations (13) et (15) proposées pour l'autocorrélation biaisée (courbe ◊) et la covariance (courbe *). Le rapport signal sur bruit est de 10dB.*

Ces résultats montrent que la variance de l'estimation de la fréquence f_0 par la méthode de la covariance modifiée est de l'ordre de 10^{-6} pour un nombre de points égal à 200, et pour un rapport signal sur bruit de 10dB. Une précision de cet ordre ne peut pas être obtenue par des méthodes classiques telles que le périodogramme, pour un nombre de points aussi faible.

6 Conclusion

L'étude qui est présentée dans cet article nous permet de proposer des formules approchées du Biais et de la variance d'estimation de quatre estimateurs de fréquence basés sur une modélisation autorégressive du signal. Ces formules analytiques simples et précises nous permettent d'étudier le comportement de ces estimateurs par rapport aux données du problème : fréquence recherchée, nombre de points utilisés et rapport Signal sur Bruit. En outre, ces résultats permettent de comparer les qualités et les défauts des différentes techniques d'estimation.

Les résultats obtenus ont été validés par une analyse statistique de Monte Carlo pour de nombreuses configurations, et montrent que les formules analytiques proposées sont valables pour une assez large gamme de variation des paramètres du problème : f_0 , N , RSB.

L'étude effectuée permet de redémontrer l'influence de la fréquence recherchée (f_0) et donc de la fréquence d'échantillonnage (F_e) sur les propriétés statistiques de l'estimation, mise en évidence auparavant dans [2]. Plus précisément, le biais d'estimation est nul quelque soit la méthode pour une fréquence $f_0 = 0.25$. De plus, le biais des estimateurs basés sur le calcul de l'autocorrélation biaisée dépend du nombre de points N utilisés, contrairement aux méthodes NB, COV et CM.

Par ailleurs, de façon très grossière, la variance de la méthode des covariances est inférieure de 10dB environ à celle de la méthode des autocorrélations pour des rapports signal sur bruit de l'ordre de 20dB. Pour des valeurs de N et de RSB grandes avec $N \gg \text{RSB}$, l'expression (15) se réduit à

$$\frac{1}{4\pi^2 N \text{RSB}^2}.$$

Ce résultat est en accord avec les remarques formulées par Sakai dans [5].

Les formules obtenues peuvent être utilisées pour la mise au point de capteurs faible coût ayant des précisions calculées *a priori*, et permettent de donner un intervalle de confiance pour les mesures effectuées.

Références

- [1] F. Castanié A. Ducasse, C. Mailhes. Estimation de fréquences : Panorama des méthodes paramétriques. *Traitement du signal*, 15:149 – 162, 1998.
- [2] F. Castanié A. Ducasse, C. Mailhes. Estimation fréquentielle et échantillonnage : Les bons choix. *Dix-septième colloque GRETSI*, pages 235 – 238, 1999.
- [3] S. M. Kay. *Modern Spectral Estimation*. Prentice Hall, 1988.
- [4] S. L. Marple. *Digital Spectral Analysis with Applications*. Prentice Hall, 1987.
- [5] H. Sakai. Statistical properties of AR spectral analysis. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-27:402 – 409, 1979.
- [6] D. N. Swingler. Frequency errors in MEM processing. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-28:257 – 259, 1980.