JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

G. BAUMANN

Filtrage et prédiction d'une série chronologique

Journal de la société statistique de Paris, tome 116 (1975), p. 61-74
<http://www.numdam.org/item?id=JSFS_1975__116__61_0>

FILTRAGE ET PRÉDICTION D'UNE SÉRIE CHRONOLOGIQUE

We study the filtering and the prediction of chronological series by the technique of spectrum analysis. We see that even from a rather small number of information, we can do rather precise predictions and we show how important the part of the autocorrelation function is. We see that the best prediction we can do is got by the average value in reducing the average quadratic error to its minimum.

Wir studieren die Filtrierung und die Voraussagen von Zeitserien mit der Technik der Spektralanalyse. Wir stellen fest, dass man selbst mit verhältnismässig beschränkten Informationen Voraussagen mit einer gewissen Prezision machen kann und wir zeigen die bedeutende Rolle, die die Autocorrelation spielt. Wir stellen fest, dass die best mögliche Voraussage mit dem Durchschnittswert erhalten wird, in dem man den geringtsen quadratischen Durchschnittefehler berechnet.

En physique expérimentale, beaucoup de données sont fournies sous formes de séries chronologiques ou sous forme de mesures répétées. Ceci est notamment le cas en physique nucléaire et en théorie de l'information.

En plus, chaque mesure peut être entachée d'erreurs multiples, souvent impossibles à calculer à cause de la complexité des méthodes de mesure et des appareils utilisés.

Donc, pour analyser le signal $x(t)$ concernant le phénomène étudié, on ne dispose que du signal $z(t)$ tel que :

$$z(t) = x(t) + w(t)$$

$w(t)$ étant le bruit de fond et qui est généralement inconnu.

Pour ces séries, l'analyse spectrale fournit alors une excellente méthode d'étude permettant de mettre en évidence les propriétés du 1er et du 2^{e} ordre [1] [2].

Cependant, ces séries ne sont pas toujours des réalisations de processus stationnaires. On part alors d'un processus stationnaire connu; à partir de là, on va engendrer une série chronologique, dont on peut éliminer la tendance et que l'on peut alors soumettre à l'analyse spectrale. Nous allons voir en particulier comment peuvent évoluer ces séries de mesures dans le futur par des études de filtrage et des méthodes de prédiction d'avenir.

1. Prédiction d'avenir d'une série chronologique connaissant une seule valeur

Soit donc la série $x(t)$ telle que $x(t) = z(t) + w(t)$ avec $z(t)$ connue à l'instant t et supposons en plus que nous connaissons ses propriétés statistiques, donc que nous pouvons avoir, en particulier la fonction d'autocorrélation $C(\tau)$.

Nous cherchons donc $z(t+\tau)$.

Avec un opérateur linéaire, on a donc :

$$\tilde{z}(t+\tau) = A\, z(t)$$

Mais, dans cette prédiction, il faut de plus remplir la condition, à savoir :

$$E\,\{[z(t+\tau) - A\, z(t)]\}^2$$

minimum (nous considérons pour simplifier que $< z(t) > = 0$).

Ce que nous pouvons écrire :

$$E\,\{[z(t+\tau) - A\, z(t)]\, z(t)\} = 0$$

Nous pouvons maintenant introduire les fonctions d'autocorrélation $C(\tau)$ et $C(0)$ qui sont, par définition :

$$C(\tau) = E\,\{z(t+\tau)\, z(t)\}$$
$$C(0) = E\,\{z(t)^2\}$$

Donc, cette condition peut se mettre sous la forme :

$$C(\tau) - A\, C(0) = 0$$

Donc nous avons comme solution :

$$A = \frac{\tilde{z}\,(t+\tau)}{z\,(t)} = \frac{C\,(\tau)}{C\,(0)} \qquad \text{c'est-à-dire} \qquad \tilde{z}\,(t+\tau) = \frac{C\,(\tau)}{C\,(0)}\, z\,(t)$$

Cette prédiction statistique donne donc $z(t+\tau)$ avec une certaine précision, et nous pouvons calculer l'erreur à laquelle nous pouvons nous attendre, soit $\varepsilon(t)$ telle que :

$$\langle\, \varepsilon^2(t)\, \rangle = E\,\{[z(t+\tau) - \tilde{z}(t+\tau)]^2\} = E\,\{[z(t+\tau) - A\, z(t)]^2\}$$

et en tenant compte des fonctions d'autocorrélation, on a pour l'erreur $\langle\, \varepsilon^2(t)\, \rangle$:

$$\langle\, \varepsilon^2(t)\, \rangle = C(0) - 2A\, C(\tau) + A^2\, C(0)$$

Mais, si, en plus, on retient que $A = \frac{C\,(\tau)}{C\,(0)}$, on a finalement pour l'erreur :

$$\langle \varepsilon^2\,(t) \rangle = C\,(0) - \frac{C^2\,(\tau)}{C\,(0)}$$

On voit donc, que même avec des informations assez réduites, on peut faire des prédictions d'une certaine précision et le rôle important que joue ici encore la fonction d'autocorrélation.

Évidemment, si τ devient grand, $C(\tau)$ tend vers zéro et la prédiction $\tilde{z}(t+\tau)$ devient de moins en moins valable et tend vers une valeur nulle. La prédiction ne peut donc se faire avec une certaine précision que pour τ pas trop grand, donc pour l'avenir immédiat.

2. Prédiction d'une série chronologique connaissant deux valeurs

Cherchons maintenant comment pourra être prolongé la suite $x(t) = z(t) + w(t)$ lorsqu'on connaît pour celle-ci deux valeurs $z(t)$ et $z(t-T)$. Alors, nous aurons au temps $t+\tau$:

$$\tilde{z}(t+\tau) = A - z(t) + B\, z(t-T)$$

mais avec la condition de rendre minimum le moment :

$$E\left\{\left[z(t+\tau)-(A\,z(t)+B\,z(t-T))\right]^2\right\}$$

c'est-à-dire que les moments suivants soient nuls, à savoir :

$$E\left\{\left[z(t+\tau)-(A\,z(t)+B\,z(t-T))\right]z(t)\right\}=0$$

et

$$E\left\{\left[z(t+\tau)-(A\,z(t)+B\,z(t-T))\right]z(t-T)\right\}=0$$

ou encore, en introduisant les fonctions de corrélation

$$\begin{cases} C(\tau)-A\,C(0)-B\,C(T)=0 \\ C(\tau+T)-A\,C(T)-BC\,(0)=0 \end{cases}$$

On peut ainsi calculer les coefficients de prédiction A et B :

$$A=\frac{\begin{vmatrix} C(\tau) & C(T) \\ C(\tau+T) & C(0) \end{vmatrix}}{\begin{vmatrix} C(0) & C(T) \\ C(T) & C(0) \end{vmatrix}}$$

$$B=\frac{\begin{vmatrix} C(0) & C(\tau) \\ C(T) & C(\tau+T) \end{vmatrix}}{\begin{vmatrix} C(0) & C(T) \\ C(T) & C(0) \end{vmatrix}}$$

Il faut évidemment calculer $C(\tau)$; $C(T)$ et $C(T+\tau)$. Cela peut se faire en prenant une série de Taylor à partir de $C(t)$.

3. Prédiction d'une série chronologique connaissant n valeurs

Si enfin nous connaissons les valeurs de $z(t)$ pour n points, à savoir $z(t)$, $z(t-T)$, $z(t-2T)$, ... $z(t-nT)$ nous pouvons prédire à $t+\tau$ par la suite :

$$\tilde{z}(t+\tau)=A_0\,z(t)+A_1\,z(t-T)+A_2\,z(t-2T)+\ldots+A_n\,z(t-t-nT)$$

donc rendant minimum le moment

$$E\{[z(t+\tau)-A_0\,z(t)-A_1\,z(t-T)-A_2\,z(t-2T)-\ldots-A_n\,z(t-nT)]^2\}$$

ce qui donne comme équations, tenant compte de la définition des fonctions d'autocorrélation :

$$\begin{aligned} &E\left\{(z(t))^2\right\}=C(0)\\ &E\{z(t)\quad z(t+T)\}=C(T)\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ &E\{z(t)\quad z(t+nT)\}=C(nT)\end{aligned}$$

donc

$$\begin{cases} A_0\,C(0)+A_1\,C(T)+A_2\,C(2T)+\ldots+A_n\,C(nT)-C(\tau)=0\\ A_0\,C(T)+A_1\,C(0)+A_2\,C(T)+\ldots+A_n\,C[(n-1)T]-C(\tau+T)=0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ A_0\,C(nT)+A_1\,C[(n-1)T]+A_2\,C[(n-2)T]+\ldots+A_n\,C(0)-C(\tau+nT)=0 \end{cases}$$

ou encore, sous forme matricielle :

$$\begin{bmatrix} C(0) & C(T) & C(2T) \ldots & C(nT) \\ C(T) & C(0) & C(T) \ldots & C((n-1)T) \\ C(2T) & C(T) & C(0) \ldots & C((n-2)T) \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ C(nT) & C((n-1)T) & C((n-2)T) \ldots & C(0) \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} C(\tau) \\ C(\tau+T) \\ C(\tau+2T) \\ \vdots \\ C(\tau+nT) \end{bmatrix}$$

ce qui permet le calcul des coefficients A_t.

Mais il faut bien remarquer que la prédiction dépend de la forme de la loi de corrélation et on peut montrer que, pour un processus poissonnien, le passé n'aide pas à prévoir l'avenir et que la seule valeur intéressante est la dernière valeur de la série, soit $z(t)$.

On dit qu'un tel système est sans mémoire.

Une telle suite dans laquelle le passé n'influe pas sur l'avenir constitue une chaîne de Markov et son évolution un processus de Markov. Mais il faut évidemment faire bien attention. Tout ce qui précède influe bien sur le calcul de z_n à l'instant t, c'est-à-dire z_n dépend de z_{n-1}, mais z_n contient toute l'information de la série. On peut encore dire que toute l'évolution du passé se trouve résumée dans le dernier état z_n. Donc il faut connaître z_n pour prévoir l'avenir, mais peu importe comment on est arrivé à z_n. On écrit dans ce cas :

$$P_r\,\{ z(t) \in \omega_1, z(t_1) = z_1, \ldots, z(t_n) = z_n \} = P_r\,\{z(t) \in \omega_1, z(t_n) = z_n \}$$

pour $\forall\ n$ et pour $\forall\ t_1, \ldots, t_n$ tel que $t_1 < t_2 < \ldots < t_n < t$.

Il faut donc bien distinguer le processus de Markov du cas particulier où :

$$P_r\,\{ z(t_1) \in \omega_1, z(t_1) = z_1, \ldots, z(t_n) = z_n \} = P_r(z(t) \in \rho_1)$$

qui est un cas particulier où la loi chronologique conditionnelle se confond avec la loi chronologique, à priori, et où donc le passé n'apporte aucune information sur l'avenir. Ce cas, bien que riche en propriétés, est dans la pratique assez peu intéressant, vu qu'il est extrêmement rare que l'on trouve remplies les conditions d'indépendance.

Enfin, on peut encore calculer une interpolation $z(t) \in z(0)$ et $z(t)$ et définir la prédiction optimum dans cet intervalle, soit $\tilde{P}$. On a :

$$\tilde{P} = \int_0^T \tilde{z}(t)\,\mathrm{d}t = \int_0^T C(t)\,\mathrm{d}t \cdot \frac{z(0) + z(T)}{C(0) + C(T)}$$

car on a, dans ces cas : $0 < t < T$

$$\tilde{z}(t) = A\,z(T) + B\,z(0)$$

avec

$$\left\{ \begin{aligned} A &= \frac{C(0)\,C(T-t) - C(T)\,C(t)}{C^2(0) - C^2(T)} \\ B &= \frac{C(0)\,C(t) - C(T-t)\,C(T)}{C^2(0) - C^2(T)} \end{aligned} \right.$$

4. Prédiction dans le cas ou l'on dispose de toute la série $x(t)$ sur l'ensemble de l'intervalle des temps

Si on connaît donc $x(t) = z(t) + w(t)$ sur $t \in (-\infty, +\infty)$, on n'est pas obligé d'injecter la série $z(t)$ dès le départ. On utilise un filtre linéaire G et nous avons alors :

$$\tilde{z}(t) = \int_{-\infty}^{+\infty} G(t - T)\,x(T)\,\mathrm{d}t$$

qui, avec $t - T = \tau$ donne :

$$\tilde{z}(t) = \int_{-\infty}^{+\infty} G(\tau)\, x(t-\tau)\, d\tau$$

La prédiction doit être telle que le moment :

$$E\{[z(t) - \tilde{z}(t)]^2\}$$

soit minimum (principe d'orthogonalité), ce qui peut encore s'écrire :

$$E\left\{\left[z(t) - \int_{-\infty}^{+\infty} G(\tau)\, x(t-\tau)\, d\tau\right] \wedge \right\} = 0$$

Si nous développons cette expression, nous obtenons 2 termes.

Nous pouvons exprimer $E\{z(t)\ x(\wedge)\}$, qui devient :

$$E\{z(t)\ x(\wedge)\} = C_{zx}(t - \wedge)$$

et ensuite le deuxième terme, qui devient :

$$\int_{-\infty}^{+\infty} G(\tau)\, E\{x(t-\tau)\, x(\wedge)\}\, d\tau = \int_{-\infty}^{+\infty} G(\tau)\, C_{xx}(t-\tau-\wedge)\, d\tau$$

et notre moment à annuler devient :

$$C_{zx}(t-\tau) = \int_{-\infty}^{+\infty} G(T)\, C_{xx}(t-\tau-\wedge)\, d\tau$$

qui est un produit de convolution où nous appelons C_{xx} la fonction d'autocorrélation pour $x = z + W$, donc :

$$C_{xx}(T) = E\{[z(t) + w(t)]\ [z(t+T) + w(t+T)]\}$$

et

$$C_{zx}(T) = E\{[z(t)]\ [z(t+T) + w(t+T)]\}$$

Si maintenant nous admettons qu'entre la série réelle et le « bruit ». il n'y a pas de corrélation, nous obtenons finalement :

$$C_{xx}(T) = C_{zz}(T) + C_{ww}(T)$$
$$C_{zx}(T) = C_{zz}(T)$$

Pour calculer $C_{zx}(t - \wedge)$, il faut prendre la transformée de Fourier du produit de convolution, donc :

$$\int_{-\infty}^{+\infty} C_{zz}(T)\, e^{i\omega t}\, dt = \int_{-\infty}^{+\infty} G(T)\, [C_{zz}(T) + C_{ww}(T)]\, e^{i\omega t}\, dt$$

$$S_z(\omega) = G(\omega)\, [S_z(\omega) + S_w(\omega)]$$

et enfin

$$G(\omega) = \frac{S_z(\omega)}{S_z(\omega) + S_w(\omega)}$$

sera le filtre à employer.

Il nous faut encore connaître l'erreur commise dans cette prédiction. Cette erreur est donnée par $\varepsilon(t)$ avec :

$$\langle \varepsilon^2(t) \rangle = E\left\{\left[z(t) - \int_{-\infty}^{+\infty} G(\tau)\, x(t-\tau)\, d\tau\right] z(t)\right\}$$

donc, en développant et en introduisant les autocorrélations

$$\langle \varepsilon^2(t) \rangle = C_z(0) - \int_{-\infty}^{+\infty} G(\tau)\, C_{xz}(\tau)\, d\tau$$

mais, en utilisant les transformations de Fourier pour les deux membres de cette équation, à savoir

$$C_z(0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_z(\omega)\, d\omega ,$$

$$\int_{-\infty}^{+\infty} G(\tau)\, C_{zx}(\tau)\, d\tau = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G(\omega)\, S_{zx}(\omega)\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G(\omega)\, S_z(\omega)\, d\omega$$

On obtient pour $\varepsilon(t)$:

$$\langle \varepsilon^2(t) \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left[S_z(\omega) - \frac{S_z^2(\omega)}{S_z(\omega) + S_w(\omega)} \right] d\omega$$

donc,

$$\langle \varepsilon^2(t) \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{S_z(\omega)\, S_w(\omega)}{S_z(\omega) + S_w(\omega)}\, d\omega$$

qui donne l'erreur minimum à laquelle on doit s'attendre. Ceci ne prédit d'ailleurs en rien si le filtre est en pratique réalisable. Ce filtre dépend évidemment de la forme de $S_z(\omega)$ et de $S_w(\omega)$. Il faut encore signaler que si la série et le « bruit » occupent des bandes de fréquence différentes, ce filtre est très efficace et donne une excellente élimination.

Par contre, la prédiction est mauvaise si série et « bruit » ont des bandes de fréquence qui se recouvrent.

Ce filtre permet donc de mettre en évidence des cycles conjoncturels et de prévoir leur variation à condition qu'ils aient une période différente des cycles aléatoires ou des cyles saisonniers.

5. Fonctions de corrélation et erreurs de prédictions

Soit la variable aléatoire X_t qui prend les valeurs x_n au temps $t = n$ ($n = -\infty$, ..., -1, 0, 1, 2, $+\infty$) et cherchons les relations entre les possibilités de prédiction de s_n et la fonction d'autocorrélation $C(\tau)$.

5.1. *Prédiction et fonctions de corrélations*

Considérons le terme x_{n+1}. On peut écrire que :

$$x_{n+1} = \lambda x_n + \mu x_{n-1} + \nu x_{n-2} + \ldots + Z_{n+1}$$

$\lambda x_n + \mu x_{n-1} + \nu x_{n-2}$, le premier terme, constitue la partie entièrement prévisible de notre série tandis que le deuxième terme Z_{n+1} est une fonction totalement imprévisible, c'est-à-dire une fonction non corrélée dont nous connaissons deux propriétés :

$$\langle Z_n \rangle = 0 \quad \text{et} \quad \langle Z_n^2 \rangle = \sigma^2 = \text{Constante}$$

Nous prendrons donc pour valeur de x_{n+1}, tout simplement la prédiction :

$$x_{n+1} = \lambda x_n + \mu x_{n-1} + \nu x_{n-2}$$

mais il faut alors connaître la fonction de corrélation pour calculer l'erreur commise.

Reprenons alors toute la série à partir de $t = 0$, donc :

$$\begin{aligned} x_0 &= Z_0 \\ x_1 &= \lambda Z_0 + Z_1 \\ x_2 &= \lambda x_1 + \mu x_0 + Z_2 = (\lambda^2 + \mu)\, Z_0 + \lambda Z_1 + Z_0 \end{aligned}$$

$$\begin{aligned}
x_3 &= \lambda x_2 + \mu x_1 + \nu x_1 + Z_3 \\
&= (\lambda^3 + 2\lambda\mu + \nu)\, Z_0 + (\lambda^2 + \mu)\, Z_1 + \lambda Z_2 + Z_3 \\
x_4 &= \lambda x_3 + \mu x_2 + \nu x_1 + Z_4 \\
&= (\lambda^4 + 3\lambda^2\mu + 2\lambda\nu + \mu^2)\, Z_0 + (\lambda^3 + 2\lambda\mu + \nu)\, Z_1 + (\lambda^2 + \mu)\, Z_2 + \lambda Z_3 + Z_4 \\
&\dots\dots\dots\dots\dots\dots\dots\dots \\
&\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}$$

donc :

$$x_n = q_n Z_0 + q_{n-1} Z_1 + \dots + q_1 Z_{n-1} + q_0 Z_n$$

avec

$$\begin{aligned}
q_0 &= 1 & q_3 &= \lambda^3 + 2\lambda\mu + \nu \\
q_1 &= \lambda & q_4 &= \lambda^4 + 3\lambda^2\mu + 2\lambda\nu + \mu^2 \\
q_2 &= \lambda^2 + \mu & & \\
&\dots\dots\dots\dots & & \\
&\dots\dots\dots\dots & &
\end{aligned}$$

donc

$$q_{n+1} = \lambda q_n + \mu q_{n-1} + \nu q_{n-2}$$

constitue la loi de récurrence entre les coefficients. Cherchons alors les solutions de ces équations.

Si nous appelons r_1, r_2, r_3 les racines de l'équation

$$r^3 = \lambda r^2 + \mu r + \nu$$

nous obtenons pour q_n :

$$q_n = a_1 \cdot r_1^n + a_2 r_2^n + a_3 r_3^n$$

avec

$$\left\{\begin{aligned}
&a_1 + a_2 + a_3 = q_0 = 1 \\
&a_1 r_1 + a_2 r_2 + a_3 r_3 = q_1 = \lambda \\
&a_1 r_1^2 + a_2 r_2^2 + a_3 r_3^2 = q_2 = \lambda^2 + \mu
\end{aligned}\right.$$

Si nous appelons C_j la moyenne, sur tous les $t \in (-\infty, +\infty)$:

$$\begin{aligned}
C_j &= \langle x_{n-j} x_n \rangle \\
&= (q_0 Z_n + q_1 Z_{n-1} + \dots + q_n Z_0) \times (q_0 Z_{n-j} + q_1 Z_{n-1-j} + \dots + q_{n-j} Z_0)
\end{aligned}$$

mais nous savons que

$$\langle Z_n \, Z_m \rangle = 0 \qquad \text{si } n \neq m$$

et

$$\langle (Z_n)^2 \rangle = \sigma^2$$

ce qui donne :

$$\begin{aligned}
C_j &= G^2 \sum_{i=0}^{\infty} q_i \, q_{i+j} \\
&= G^2 \sum_{i=0}^{\infty} (a_1 r_1^i + a_2 r_2^i + a_3 r_3^i)\,(a_1 r_1^{i+j} + a_2 r_2^{i+j} + a_3 r_3^{i+j})
\end{aligned}$$

Nous allons maintenant introduire les relations suivantes :

$$\sum_{i=0}^{\infty} r_1^i \, r_2^i = \frac{1}{1 - r_1 r_2}$$

dans C_j, ce qui donne :

$$C_j = \sigma^2 \left(\frac{a_1^2}{1-r_1^2} + \frac{a_1\,a_2}{1-r_1\cdot r_2} + \frac{a_1\,a_3}{1-r_1\cdot r_3}\right) r_1^{|j|} + \sigma^2 \left(\frac{a_1\,a_2}{1-r_1\cdot r_2} + \frac{a_2^2}{1-r_2^2} + \frac{a_2\,a_3}{1-r_2 r_3}\right) r_2^{|j|}$$

$$+ \sigma^2 \left(\frac{a_1\,a_3}{1-r_1 r_3} + \frac{a_2\,a_3}{1-r_2 r_3} + \frac{a_3^2}{1-r_3^2}\right) r_3^{|j|}$$

qui est la loi de corrélation.

Mais rappelons-nous que r_1, r_2, r_3 sont les racines de l'équation en r^3, donc :

$$r^3 - (\lambda r^2 + \mu r + \nu) = (r - r_1)\ (r - r_2)\ (r - r_3)$$

et en développant, nous pouvons identifier et obtenir :

$$\begin{cases} \lambda = r_1 + r_2 + r_3 \\ \mu = -r_1 r_2 = r_1 r_3 - r_2 r_3 \\ \nu = r_1 r_2 r_3 \end{cases}$$

et, à l'aide des relations entre r_1, r_2, r_3 et a_1, a_2 a_3, on obtient :

$$\begin{aligned} a_1 &= r_1^2\ (r_2 - r_3)/\Delta \\ a_2 &= r_2^2\ (r_3 - r_1)/\Delta \\ a_3 &= r_3^2\ (r_1 - r_2)/\Delta \end{aligned}$$

avec

$$\Delta = r_1 r_2 (r_1 - r_2) + r_2 r_3 (r_2 - r_3) + r_3 r_1 (r_3 - r_1)$$

Conclusion

Nous obtenons donc la loi de corrélation en remplaçant dans C_j les coefficients a_1, a_2, a_3 par les valeurs définies ci-dessus.

Nous obtenons la prédiction par :

$$x_{n+1} = \lambda x_n + \mu x_{n-1} + \nu x_{n-2} + Z_{n+1}$$

en remplaçant les coefficients λ, μ et ν par les valeurs calculées ci-dessus.

5.2. *Erreurs sur la prédiction*

Il est clair, d'après la définition de la prédiction de

$$x_{n+1} = \lambda x_n + \mu x_{n-1} + \nu x_{n-2} + Z_{n+1}$$

que l'erreur est donnée par le terme aléatoire Z_{n+1} avec $\varepsilon(t)$ tel que

$$\langle\, \varepsilon_1^2(t)\, \rangle = \langle\, Z_{n+1}^2\, \rangle = \sigma^2$$

De même, la prédiction de x_{n+2} à partir des valeurs x_n, x_{n-1}, x_{n-2} est donnée par une combinaison Z_{n+1} et Z_{n+2}.

L'ensemble des deux relations suivantes :

$$\begin{cases} x_{n+1} = \lambda x_n + \mu x_{n-1} + \nu x_{n-2} + Z_{n+1} \\ x_{n+2} = \lambda x_{n+1} + \mu\, x_n + \nu x_{n-1} + Z_{n+2} \end{cases}$$

donne :

$$x_{n+2} = (\lambda^2 + \mu)\, x_n + (\lambda\mu + \nu)\, x_{n-1} + \lambda\nu x_{n-2} + \lambda Z_{n+1} + Z_{n+2}$$

mais Z_{n+1} et Z_{n+2} étant imprévisibles, on ne peut donc avoir que x_{n+2} comme meilleure prévision. On a :

$$\tilde{x}_{n+2} = (\lambda^2 + \mu)\, x_n + (\lambda\mu + \nu)\, x_{n-1} + \lambda\nu x_{n-2}$$

avec l'erreur $\varepsilon_2(t) = \lambda Z_{n+1} + Z_{n+2}$
dont le carré moyen donne :

$$\langle \varepsilon_2^2(t) \rangle = \lambda^2 \langle Z_{n+1}^2 \rangle + \langle Z_{n+2}^2 \rangle + \langle Z_{n+1} \quad Z_{n+2} \rangle \cdot 2\lambda$$

c'est-à-dire :

$$\langle \varepsilon_2^2(t) \rangle = \sigma^2 (1 + \lambda^2)$$

On peut alors calculer tous les termes suivants.

La meilleure prédiction pour x_{n+3} est $\tilde{x}_{n+3}$ avec

$$\tilde{x}_{n+3} = (\lambda^3 + 2\lambda\mu + \nu)\, x_n + (\lambda^2\mu + \lambda\nu + \mu^2)\, x_{n-1} + (\lambda^2\nu + \mu\nu)\, x_{n-2}$$

entraînant une erreur $\varepsilon_3(t)$ avec,

$$\langle \varepsilon_3^2(t) \rangle = \sigma^2 [1 + \lambda^2 + (\lambda^2 + \mu)^2]$$

c'est-à-dire qu'on peut facilement généraliser l'ensemble des relations. On peut alors prédire x_{n+j} en itérant j fois l'équation de prédiction

$$\lambda x_{n-j} + \mu x_{n-2j} + \nu x_{n-3j} + Z_{n+j}$$

et aussi une erreur $\varepsilon_j(t)$ telle que :

$$\langle \varepsilon_j^2 (t) \rangle = \sigma^2 \sum_{j=0}^{\infty} q_j^2$$

avec

$$\sum_{j=0}^{\infty} q_j^2 = \frac{a_1^2}{1 - r_1^2} + \frac{a_2^2}{1 - r_2^2} + \frac{a_3^2}{1 - r_3^2} + \frac{2a_1 a_2}{1 - r_1 r_2} + \frac{2a_1 a_3}{1 - r_1 r_3} + \frac{2a_2 a_3}{1 - r_2 r_3}$$

A la limite, on obtient donc :

$$\langle \varepsilon_\infty^2 \ (t) \rangle = C(0)$$

La meilleure prédiction possible est donc la valeur moyenne qui minimise l'erreur quadratique moyenne.

6. Itération des prédictions d'une série discrète

Essayons maintenant de généraliser les résultats de prédictions, c'est-à-dire prédire x_{n+1} à partir de l'ensemble des valeurs x_j. On obtient pour x_{n+1} la relation :

$$x_{n+1} = A_{01} x_n + A_{11} x_{n-1} + A_{21} x_{n-2} + \dots$$

relation dans laquelle les quantités A_{j1} sont des solutions de l'équation

$$\begin{bmatrix} C(0) & C(1) & C(2) & C(3) & \dots \\ C(1) & C(0) & C(1) & C(2) & \dots \\ C(2) & C(1) & C(0) & C(1) & \dots \\ C(3) & C(2) & C(1) & C(0) & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} A_{01} \\ A_{11} \\ A_{21} \\ A_{31} \\ \vdots \end{bmatrix} = \begin{bmatrix} C(1) \\ C(1+1) \\ C(1+2) \\ C(1+3) \\ \vdots \end{bmatrix} \qquad (1)$$

Les grandeurs A_{j1} peuvent alors être considérées comme les composantes d'un vecteur $\overrightarrow{A_1}$, et nous avons :

$$\tilde{x}_{n+1} = A_{01}x_n + A_{11}x_{n-1} + A_{21}x_{n-2} + A_{31}x_{n-3} + \dots$$

S'il existe une relation d'itération, on peut donc, également prédire x_{n+2} et on a :

$$\tilde{x}_{n+2} = A_{01}x_{n+1} + A_{11}x_n + A_{21}x_{n-1} + A_{31}x_{n-2} + \dots$$

qui s'écrit, compte tenu de $\tilde{x}_{n+1}$:

$$\tilde{x}_{n+2} = A_{02}x_n + A_{12}x_{n-1} + A_{22}x_{n-2} + A_{31}x_{n-3} + \dots$$

avec :

$$\begin{cases} A_{02} = A_{01}^2 + A_{11} \\ A_{12} = A_{01}A_{11} + A_{21} \\ A_{22} = A_{01}A_{21} + A_{31} \\ A_{32} = A_{01}A_{31} + A_{41} \end{cases}$$

Nous pouvons de même prévoir x_{n+3} en itérant 3 fois

$$\tilde{x}_{n+3} = A_{01}x_{n+2} + A_{11}x_{n+1} + A_{21}x_n + A_{31}x_{n-1} + \dots$$

et, tenant toujours compte des relations précédentes :

$$\tilde{x}_{n+3} = A_{03}x_n + A_{13}x_{n-1} + A_{23}x_{n-2} + A_{33}x_{n-3} + \dots$$

avec le coefficient donné par les relations :

$$\begin{cases} A_{03} = A_{01}^3 + A_{01}A_{11} + A_{11}A_{01} + A_{21} \\ A_{13} = A_{01}^2A_{11} + A_{01}A_{21} + A_{11}^2 + A_{31} \\ A_{23} = A_{01}^2A_{21} + A_{01}A_{31} + A_{11}A_{21} + A_{41} \\ A_{33} = A_{01}^2A_{31} + A_{01}A_{41} + A_{11}A_{31} + A_{51} \end{cases}$$

Nous allons maintenant montrer que la construction du vecteur $\overrightarrow{A_1}$ est itérative. Nous avons donc obtenu, à partir de $\overrightarrow{A_1}$ les vecteurs $\overrightarrow{A_2}$ $(A_{02}, A_{12}, A_{22}, A_{32})$ et $\overrightarrow{A_3}$ $(A_{03}, A_{13}, A_{23}, A_{33})$. Pour montrer que cette concentration est itérative, il faut donc porter les composantes de $\overrightarrow{A_2}$ et $\overrightarrow{A_3}$ dans (1) et vérifier l'égalité avec le deuxième membre.

Considérons d'abord $\overrightarrow{A_1}$. L'équation (1) donne alors :

$$\begin{cases} A_{01}C(0) + A_{11}C(1) + A_{21}C(2) + A_{31}C(3) + \dots = C(1) \\ A_{01}C(1) + A_{11}C(0) + A_{21}C(1) + A_{31}C(2) + \dots = C(2) \\ A_{01}C(2) + A_{11}C(1) + A_{21}C(0) + A_{31}C(1) + \dots = C(3) \\ A_{01}C(3) + A_{11}C(2) + A_{21}C(1) + A_{31}C(0) + \dots = C(4) \end{cases}$$

Puis le vecteur $\overrightarrow{A_2}$ pour lequel l'équation donne alors :

$$\begin{cases} A_{02}C(0) + A_{12}C(1) + A_{22}C(2) + A_{32}C(3) + \dots = C(1) \\ A_{02}C(1) + A_{12}C(0) + A_{22}C(1) + A_{32}C(2) + \dots = C(2) \\ A_{02}C(2) + A_{12}C(1) + A_{22}C(0) + A_{32}C(1) + \dots = C(3) \\ A_{02}C(3) + A_{12}C(2) + A_{22}C(1) + A_{32}C(0) + \dots = C(4) \end{cases}$$

Tenant compte des relations précédentes, on peut alors vérifier que :

$$\begin{aligned} &A_{01}C(0) + A_{11}C(1) + A_{21}C(2) + A_{31}C(3) - C(1) \\ &\qquad = A_{02}C(0) + A_{12}C(1) + A_{22}C(0) + A_{32}C(3) - C(2) \end{aligned}$$

et ainsi de suite pour toutes les lignes.

Nous avons donc :

$$||C||\,[A_2] = \begin{bmatrix} C_2 \\ C_3 \\ C_4 \end{bmatrix}$$

et de même :

$$||C||\,[A_3] = \begin{bmatrix} C_3 \\ C_4 \\ C_5 \end{bmatrix}$$

Nous pouvons donc, à partir de $\overrightarrow{A_1}$, construire tous les vecteurs $\overrightarrow{A_2}$, $\overrightarrow{A_3}$, ce qui permet de généraliser la prédiction d'une série discrète.

7. Prédiction d'une variable aléatoire continue

Nous avons une série chronologique

$$x(t) = z(t) + w(t)$$

dont nous avons pu extraire la vraie série $z(t)$ et dont nous connaissons toutes les valeurs de $t = -\infty$ à $t = 0$. Cherchons alors la prédiction de $z(t)$ pour $t = \tau$, donc $z(\tau)$. Nous prendrons pour $z(\tau)$

$$\tilde{z}(\tau) = \int_0^\infty G(\tau, T)\, z(-T)\, \mathrm{d}T$$

Il nous faut donc déterminer $G(\tau, T)$ satisfaisant à la condition d'orthogonalité, c'est-à-dire, pour tout temps t pour lequel la série est connue, donc $t(-\infty, 0)$. On doit donc avoir :

$$E\left\{\left[z(\tau) - \int_0^\infty G(\tau, T)\, z(-T)\, \mathrm{d}T\right] z(t)\right\} = 0$$

Avec les définitions des fonctions d'autocorrélation :

$$E\{\, z(\tau)\ z(-t)\,\} = C(\tau + t)$$
$$E\{\, z(-T)\ z(-t)\,\} = C(t - T)$$

la condition d'orthogonalité devient :

$$\int_0^\infty G(\tau, T)\, C(t - T)\, \mathrm{d}T = C(\tau + t)$$

C'est l'équation de Wiener Hopf dont nous allons indiquer une méthode de résolution utilisant les arguments de la « physique » du phénomène, et qui est due à Wiener [3].

Nous allons introduire la fonction de transfert $T(\omega)$ du spectre de fréquence $S(\omega)$ du signal, donc :

$$|T(\omega)|^2 = |S(\omega)|$$

qui peut définir le filtre, mais pas sa phase.

Considérons maintenant le filtre de fonction de transfert $T^{-1}(\omega)$.

Après l'application de ce filtre, le spectre $S(\omega)$ devient alors :

$$S(\omega)\ |T(\omega)^{-}|^2 = 1$$

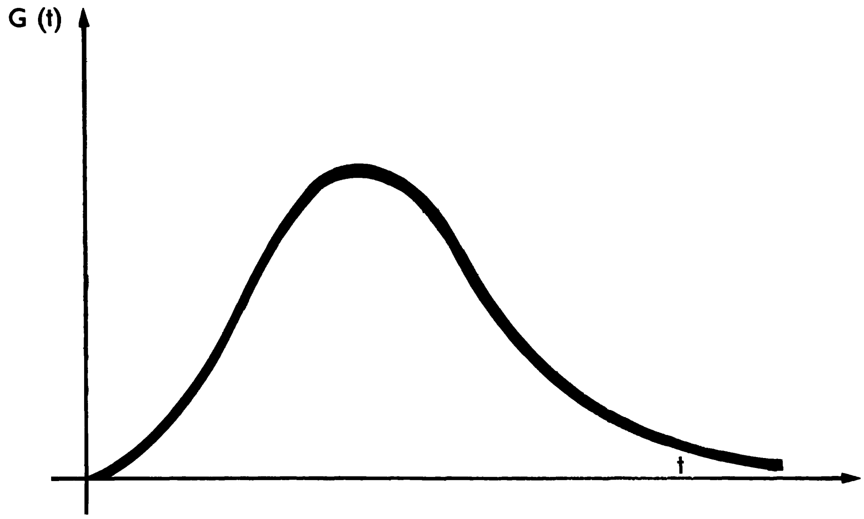
Et maintenant, appliquons de nouveau le filtre $T(\omega)$, ce qui nous donne comme information :

$$|T(\omega)|^2 = S(\omega)$$

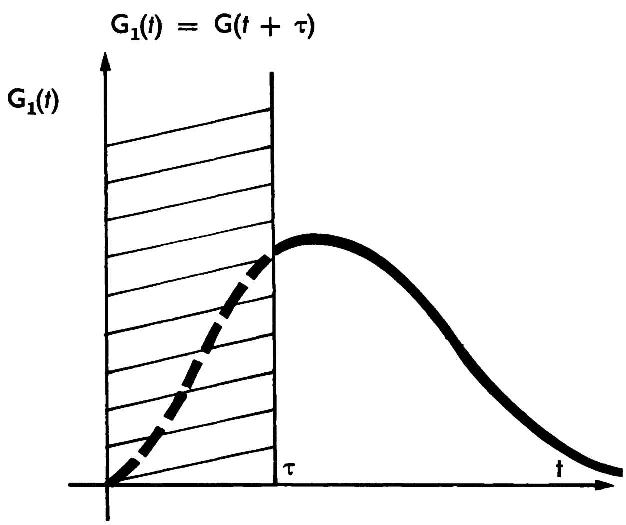
On récupère donc la série chronologique.

Mais ce que nous cherchons, c'est la série au temps $t = \tau$.
Nous pouvons définir $G(t)$ par sa transformée de Fourier.

$$G(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} T(\omega)\, e^{\iota\omega t}\, d\omega$$



Mais si nous voulons prédire à $t = \tau$, il nous faut $G(t + \tau)$, donc :



donc, pour obtenir $S(\tau)$, nous allons combiner les deux filtres suivants :

$$T^{-1}(\omega) \quad \text{puis} \quad T_1(\omega)$$

tel que :

$$T_1(\omega) = \int_{-\infty}^{+\infty} G_1(t)\, e^{-\iota\omega t}\, dt,$$

car $T_1(\omega)$ redonne la série $S(\omega)$, mais en avançant du temps τ la réponse.

Évidemment, il y a aura à tenir compte d'une introduction d'erreurs, mais les erreurs proviennent entièrement de $T(\omega)$ appliqué pendant l'intervalle $(0,\tau)$.

Nous pouvons alors donner une évaluation de l'erreur, soit ε.

On a :

$$\langle \varepsilon^2(\tau) \rangle \propto \int_0^\tau G_1^2(t)\, dt$$

Pour obtenir le coefficient de proportionnalité, on applique la propriété que pour $\tau \to \infty$

$$\langle \varepsilon^2(t) \rangle \to \sigma^2,$$

moyenne quadratique du signal de l'information. Donc :

$$\langle \varepsilon^2(\tau) \rangle = \frac{\sigma^2}{\int_0^\infty G_1^2(t)\, dt} \int_0^\tau G_1^2(t)\, dt$$

Appliquons maintenant cette théorie à quelques formes simples de série chronologique $z(t)$ pour lesquels on connaît le spectre de puissance $S(\omega)$.

Admettons que l'on connaisse une forme analytique de $S(\omega)$ et que l'on puisse écrire :

$$S(\omega) = T(\omega) \qquad T^*(\omega)$$

et que $S(\omega)$ n'a pas de pôles dans la partie inférieure du plan complexe des fréquences ω.

Cherchons alors à déterminer $T(\omega)$ que nous pouvons mettre sous la forme :

$$T(\omega) = \sum_j A_j (\alpha_j + i\omega)^{-1}$$

avec, en plus, $\alpha_j > 0$.

On a alors, avec la variable de Laplace $z = i\omega$

$$T(z) = \sum_j A_j (\alpha_j + z)^{-1}$$

qui donne :

$$R(t) = \sum_j A_j e^{-\alpha_j t}$$

Donc, on obtient pour la transformée de Laplace de $G_1(t)$:

$$S_1(z) = \sum_j \frac{A_j}{z + \alpha_j} e^{-\alpha_j \tau}$$

et la fonction de transfert à réaliser est donc :

$$\bar{s}(z) = \frac{S_1(z)}{S(z)}$$

C'est-à-dire qu'on applique d'abord le filtre $S^{-1}(z)$ qui détruit les corrélations, puis le filtre $S_1(z)$ qui reconstitue le signal mais au temps τ de prédiction.

G. Baumann

Laboratoire de physique théorique
Université de Nancy I

RÉFÉRENCES

[1] GRANGER C.-W. et HATANAKA M. — « Spectral analysis of economic time series ». Princeton N. J. Princeton University Press, 1964.

[2] THOMAS J.-B. — « An introduction to statistical communication theory ». John Wiley, New York, 1969.

[3] WIENER N. — « Extrapolation, interpolation and smoothing of stationary time series ». The technologie Press of the M. I. T., Cambridge, 1949.