

# Aspects géométriques d'algorithmes de prédiction linéaire

François DESBOUVRIES

INT, Département SIM, 9 rue Charles Fourier, 91011 Évry, France  
desbou@int-evry.fr

**Résumé** – L'identité de Yule, en statistiques, et la loi des cosinus, en trigonométrie sphérique, sont une seule et même formule. Cette observation fournit une interprétation géométrique nouvelle de la régression linéaire et, par conséquent, des algorithmes de Schur et de Levinson-Szegö.

**Abstract** – An old identity of G. U. Yule among partial correlation coefficients is recognized as being equal to the cosine law of spherical trigonometry. This observation provides a new geometrical interpretation of (recursive) linear regression analysis and, as a consequence, of the classical algorithms of Schur and of Levinson-Szegö.

## 1 Introduction

En régression linéaire, on souhaite approximer, selon le critère d'erreur quadratique moyenne (e.q.m.), une variable aléatoire (v.a.) par une combinaison linéaire d'autres v.a. Soit  $\hat{X}_j^{1:n}$  l'estimée de  $X_j$  en fonction de  $\{X_i\}_{i=1}^n$  et  $\tilde{X}_j^{1:n} = X_j - \hat{X}_j^{1:n}$  l'erreur d'estimation. Le coefficient de corrélation partielle (ou parcor) de  $X_0$  et  $X_{n+1}$ , étant donné  $\{X_i\}_{i=1}^n$ , est défini par  $\rho_{0,n+1}^{1:n} = [E((\tilde{X}_0^{1:n})^2)]^{-1/2} E(\tilde{X}_0^{1:n} \tilde{X}_{n+1}^{1:n}) [E((\tilde{X}_{n+1}^{1:n})^2)]^{-1/2}$ . Il est borné par 1 en valeur absolue et est classiquement interprété comme le coefficient de corrélation entre  $X_0$  et  $X_{n+1}$ , une fois que l'influence de  $\{X_i\}_{i=1}^n$  a été soustraite. En 1907, G. U. Yule [1] a montré que les parcors pouvaient être calculés récursivement :

$$\rho_{0,n+1}^{1:n} = \frac{\rho_{0,n+1}^{1:n-1} - \rho_{0,n}^{1:n-1} \rho_{n,n+1}^{1:n-1}}{\sqrt{1 - (\rho_{0,n}^{1:n-1})^2} \sqrt{1 - (\rho_{n,n+1}^{1:n-1})^2}}.$$

Il se trouve que cette formule classique est algébriquement égale à la loi des cosinus en trigonométrie sphérique (TS) :

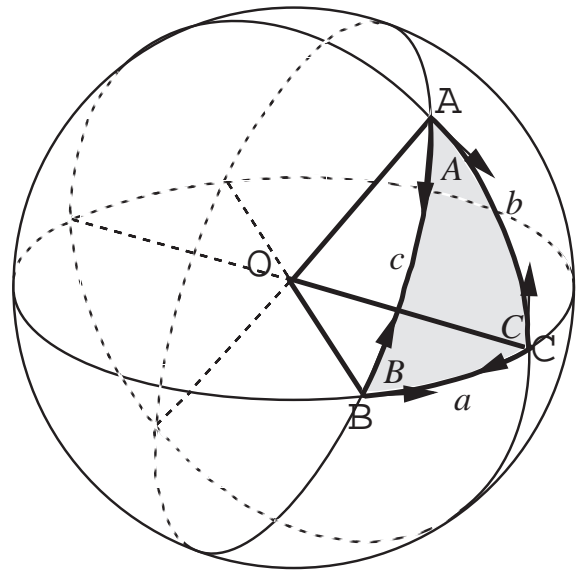
$$\cos A = \frac{\cos a - \cos b \cos c}{\sin b \sin c}, \quad (1)$$

qui donne l'angle d'un triangle sphérique en fonction de ses 3 côtés (voir figure). Cette observation établit un lien entre les statistiques et l'analyse des séries temporelles, d'une part, et la TS (une branche de la trigonométrie), d'autre part.

Le lien entre la TS et le domaine voisin du filtrage adaptatif MCRR avait précédemment été établi [2]. En fait, de telles analogies existent car les équations sont établies à partir d'identités projectives. C'est ainsi qu'en régression linéaire, on reconnaît que l'e.q.m. à minimiser est une distance, donc on applique le théorème de projection orthogonale dans l'espace de Hilbert engendré par les v.a. Introduire une nouvelle variable de régression revient à réactualiser un projecteur, et on peut décrire le problème dans un espace engendré par 3 vecteurs. Mais 3 vecteurs de norme unité délimitent un tétraèdre dans l'espace, et établir des identités projectives dans le tétraèdre

( $OA, OB, OC$ ) revient à établir des relations trigonométriques dans le triangle sphérique ( $A, B, C$ ) délimité par ce tétraèdre (voir figure).

L'article est organisé de la façon suivante. Dans le §2 nous introduisons brièvement le formalisme projectif. Dans l'optique du §3, nous passons à la régression sur  $n$  v.a.  $\{X_i\}_{i=0}^n$  et considérons des manipulations algébriques sur la matrice de covariance  $R_{0:n}$  et sur son inverse. Nous faisons le lien entre l'ajout (ou la soustraction) de v.a. dans le problème de régression, les compléments de Schur (CS) sur  $R_{0:n}$  et  $R_{0:n}^{-1}$ , et les lois des cosinus. Enfin au §3, nous nous plaçons dans le cas stationnaire. L'algorithme de Schur (resp. de Levinson-Szegö) est essentiellement constitué de 2 CS dans  $R_{0:n}$  (dans  $R_{0:n}^{-1}$ ), et reçoit de ce fait une interprétation trigonométrique dans le triangle sphérique (resp. dans le triangle sphérique polaire).



## 2 CS dans $R_{0:n} / R_{0:n}^{-1}$ et TS

La connexion entre algorithmes de régression linéaire et TS s'obtient par l'intermédiaire des CS, parce qu'algébriquement la récursion élémentaire avec pivot  $t_{k,k} : t_{i,j} \rightarrow t_{i,j} - t_{i,k}t_{k,k}^{-1}t_{k,j}$ , se ramène à une loi des cosinus par normalisation par  $\sqrt{t_{i,i}} \sqrt{t_{j,j}}$ .

### 2.1 Parcours, projections récursives et TS

Nos résultats géométriques sont basés sur les propriétés des projecteurs orthogonaux et peuvent donc être formalisés dans un espace de Hilbert quelconque. Le cadre naturel de cet article est cependant l'espace  $\mathbb{L}^2(\Omega, \mathcal{A}, P)$  des v.a. réelles, centrées, de carré intégrable, définies sur  $(\Omega, \mathcal{A}, P)$ , muni du produit scalaire  $(X, Y) = E(XY)$ .

Soit  $P_{\mathcal{M}}$  le projecteur orthogonal sur le s.e.v. de Hilbert  $\mathcal{H}(\mathcal{M})$  engendré par  $\mathcal{M}$ ,  $P_{\mathcal{M}}^{\perp} = I - P_{\mathcal{M}}$ ,  $\tilde{A}^{\mathcal{M}}$  la projection de  $A$  sur  $\mathcal{H}(\mathcal{M})$  et  $\tilde{A}^{\mathcal{M}} = A - \tilde{A}^{\mathcal{M}}$ . La notation  $\bar{X}$  signifie normalisation de norme unité :  $\bar{X} = X(X, X)^{-1/2}$ . Soit  $A$  et  $B$  deux éléments de  $\mathcal{H} \setminus \{0\}$  (resp. de  $\mathcal{H} \setminus \{\mathcal{M}\}$ ). Le coefficient de corrélation totale  $\rho_{A,B}$  (resp. partielle  $\rho_{A,B}^{\mathcal{M}}$ ) de  $A$  et  $B$  (resp. de  $A$  et  $B$ , relativement à un s.e.v. commun  $\mathcal{M}$ ) est défini comme  $\rho_{A,B} = (\bar{A}, \bar{B})$  (resp.  $\rho_{A,B}^{\mathcal{M}} = (\bar{A}^{\mathcal{M}}, \bar{B}^{\mathcal{M}})$ ).

Considérons l'aspect récursif. Il est bien connu que

$$P_{\mathcal{M},A} = P_{\mathcal{M}} + P_{\tilde{A}^{\mathcal{M}}}, \quad P_{\tilde{A}^{\mathcal{M}},A} = P_{\mathcal{M}}^{\perp} - P_{\tilde{A}^{\mathcal{M}}} \quad (2)$$

Ces identités sont très importantes en filtrage adaptatif MCRR et de Kalman. À partir de (2), on obtient

$$\tilde{B}^{\mathcal{M},A} = \tilde{B}^{\mathcal{M}} - (\tilde{B}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})(\tilde{A}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})^{-1}\tilde{A}^{\mathcal{M}}, \quad (3)$$

ce qui donne la relation utile

$$(\tilde{B}^{\mathcal{M},A}, \tilde{B}^{\mathcal{M},A}) / (\tilde{B}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}}) = (1 - (\rho_{A,B}^{\mathcal{M}})^2). \quad (4)$$

Les relations de TS s'obtiennent à partir de la manipulation de 3 résidus de projection. De (3) nous avons

$$\begin{aligned} (\tilde{C}^{\mathcal{M},A}, \tilde{B}^{\mathcal{M},A}) &= (\tilde{C}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}}) \\ &- (\tilde{C}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})(\tilde{A}^{\mathcal{M}}, \tilde{A}^{\mathcal{M}})^{-1}(\tilde{A}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}}) \end{aligned} \quad (5)$$

En divisant par  $(\tilde{C}^{\mathcal{M}}, \tilde{C}^{\mathcal{M}})^{1/2}(\tilde{B}^{\mathcal{M}}, \tilde{B}^{\mathcal{M}})^{1/2}$ , et en utilisant (4), nous obtenons

$$\rho_{C,B}^{\mathcal{M},A} = \frac{\rho_{C,B}^{\mathcal{M}} - \rho_{C,A}^{\mathcal{M}}\rho_{A,B}^{\mathcal{M}}}{\sqrt{1 - (\rho_{C,A}^{\mathcal{M}})^2} \sqrt{1 - (\rho_{A,B}^{\mathcal{M}})^2}}, \quad (6)$$

qui est formellement égale à (1), après une identification immédiate des variables.

Enfin, notons brièvement que ce lien avec la TS peut être exploité en suivant la démarche adoptée en [2]. Certaines formules se transposent en effet du cadre de la TS à celui des statistiques. Cette façon de faire fournit de nouvelles relations entre coefficients de corrélation partielle, qui, bien que purement algébriques, ne sont pas intuitives, ou permet d'interpréter géométriquement des relations existantes. Par ex., la relation [1, formule (19)] :

$$\rho_{C,B}^{\mathcal{M}} = \frac{\rho_{C,B}^{\mathcal{M},A} + \rho_{C,A}^{\mathcal{M},B} \rho_{A,B}^{\mathcal{M},C}}{\sqrt{1 - (\rho_{C,A}^{\mathcal{M},B})^2} \sqrt{1 - (\rho_{A,B}^{\mathcal{M},C})^2}} \quad (7)$$

est la relation des cosinus dans le triangle polaire :

$$\cos a = (\cos A + \cos B \cos C) / \sin B \sin C. \quad (8)$$

## 2.2 CS dans $R_{0:n}$ et $R_{0:n}^{-1}$

Nous nous intéressons maintenant aux matrices de covariance et à leurs inverses. Soit donc  $R_{0:n}$  la matrice de covariance de  $X_{0:n} = [X_0 \cdots X_n]^T$ . Nous allons utiliser des ensembles contigus d'indices et simplifions donc les notations. Par ex,  $\tilde{X}_i^{\mathcal{H}(\{X_m\}_{p \leq m \leq q})}$  devient  $\tilde{X}_i^{p:q}$  et  $\tilde{X}_i^{\mathcal{H}(\{X_m\}_{p \leq m \leq q}^{m \neq r})}$  devient  $\tilde{X}_i^{[p:q] \setminus r}$ . Des notations similaires sont adoptées pour les coefficients de corrélation. L'ordre des exposants  $p$  et  $q$  a une signification :  $\tilde{X}_i^{p:q}$ ,  $\rho_{i,j}^{p:q}$  (et plus tard  $R_{i,j}^{p:q}$ ,  $r_{i,j}^{p:q}$ ,  $P_{i,j}^{p:q}$  et  $p_{i,j}^{p:q}$ ) se réduisent respectivement à  $X_i$ ,  $\rho_{i,j}$ ,  $R_{i,j}$ ,  $r_{i,j}$ ,  $P_{i,j}$  et  $p_{i,j}$  si  $p > q$ . De cette façon la notation unifie les cas total et partiel, qu'il n'y a pas lieu de distinguer : un coefficient de corrélation totale est en effet un coefficient de corrélation partielle d'ordre 0 ; une remarque similaire s'applique plus loin aux CS (cf. Th. 1).

Nous rassemblons maintenant (et étendons quelque peu) un certain nombre de résultats connus [3] [4] [5] [6] concernant l'expression de  $R_{0:n}$  (resp. de  $R_{0:n}^{-1}$ ) en fonction de covariances des v.a.  $\{X_i\}_{i=0}^n$  (resp.  $\{\tilde{X}_i^{[0:n] \setminus i}\}_{i=0}^n$ ).

**Lemme 1** Soit  $R_{0:n} = (r_{i,j})_{i,j=0}^n$ , et soit  $P_{0:n} = R_{0:n}^{-1} = (p_{i,j})_{i,j=0}^n$ . Alors  $\forall 0 \leq i, j \leq n$ ,

$$\begin{aligned} r_{i,j} &= (X_i, X_j), \\ p_{i,j} &= \left( \frac{\tilde{X}_i^{[0:n] \setminus i}}{(\tilde{X}_i^{[0:n] \setminus i}, \tilde{X}_i^{[0:n] \setminus i})}, \frac{\tilde{X}_j^{[0:n] \setminus j}}{(\tilde{X}_j^{[0:n] \setminus j}, \tilde{X}_j^{[0:n] \setminus j})} \right). \end{aligned}$$

**Lemme 2** Soit  $R_{0:n} = (r_{i,j})_{i,j=0}^n$ , et soit  $P_{0:n} = R_{0:n}^{-1} = (p_{i,j})_{i,j=0}^n$ . Alors  $\forall 0 \leq i, j \leq n$ ,

$$\begin{aligned} \frac{r_{i,j}}{\sqrt{r_{i,i}r_{j,j}}} &= (\bar{X}_i, \bar{X}_j) = \rho_{i,j}, \\ \frac{p_{i,j}}{\sqrt{p_{i,i}p_{j,j}}} &= (\tilde{X}_i^{[0:n] \setminus i}, \tilde{X}_j^{[0:n] \setminus j}) = -\rho_{i,j}^{[0:n] \setminus i,j}. \end{aligned}$$

Rappelons que si  $M$  est partitionnée en  $\begin{bmatrix} AB \\ CD \end{bmatrix}$  avec  $A$  inversible, alors le CS  $(M/A)$  de  $A$  dans  $M$  est défini par  $(M/A) = D - CA^{-1}B$ . Le théorème suivant englobe et généralise le lemme 2 (qui correspond au cas  $p = -1$ ) :

**Théorème 1** Soit  $R_{0:n}^{0:p} = (r_{i,j}^{0:p})_{i,j=p+1}^n$  (resp.  $P_{0:n}^{0:p} = (p_{i,j}^{0:p})_{i,j=p+1}^n$ ) le CS de  $R_{0:p}$  dans  $R_{0:n}$  (de la sous-matrice supérieure gauche  $[I_{p+1} 0]P_{0:n}[I_{p+1} 0]^T$ , de taille  $(p+1) \times (p+1)$ , de  $P_{0:n}$  dans  $P_{0:n}$ ). Alors  $\forall p, -1 \leq p \leq n-1$ , et  $\forall i, j, p+1 \leq i, j \leq n$ ,

$$\frac{r_{i,j}^{0:p}}{\sqrt{r_{i,i}^{0:p}r_{j,j}^{0:p}}} = \rho_{i,j}^{0:p}, \quad \frac{p_{i,j}^{0:p}}{\sqrt{p_{i,i}^{0:p}p_{j,j}^{0:p}}} = -\rho_{i,j}^{[p+1:n] \setminus i,j}.$$

Le lien avec la TS est alors immédiat :

**Corollaire 1** Après une normalisation adéquate, une étape de complémentation de Schur élémentaire (c.à.d. de rang 1) sur  $R_{0:n}^{0:p}$  (resp. sur  $P_{0:n}^{0:p}$ ) s'identifie à la loi des cosinus (1) (resp. la loi des cosinus dans le triangle polaire (8)) :  $\forall p \in [-1, 0, \dots, n-2]$ , et  $\forall i, j \in [p+2 \cdots n]$ ,

$$\begin{aligned} \rho_{i,j}^{0:p+1} &= \frac{\rho_{i,j}^{0:p} - \rho_{i,p+1}^{0:p}\rho_{p+1,j}^{0:p}}{\sqrt{1 - (\rho_{i,p+1}^{0:p})^2} \sqrt{1 - (\rho_{p+1,j}^{0:p})^2}}, \\ \rho_{i,j}^{[p+2:n] \setminus i,j} &= \frac{\rho_{i,j}^{[p+1:n] \setminus i,j} + \rho_{i,p+1}^{[p+1:n] \setminus i,p+1} \rho_{p+1,j}^{[p+1:n] \setminus p+1,j}}{\sqrt{1 - (\rho_{i,p+1}^{[p+1:n] \setminus i,p+1})^2} \sqrt{1 - (\rho_{p+1,j}^{[p+1:n] \setminus p+1,j})^2}}. \end{aligned}$$

### 3 Application aux algorithmes de Schur et de Levinson-Szegö

Nous identifions désormais  $\{X_i\}_{i=0}^n$  à  $\{X_{t-i}\}_{i=0}^n$ , où  $\{X_t, t \in \mathbf{Z}\}$  est un p.a. SSL à temps discret.  $R_{0:n}$  est donc Toeplitz, et nous posons  $r_{i,j} = r_{i-j}$ . Les parocs vérifient aussi une propriété d'invariance temporelle :  $\forall i, j, k, p, q \in \mathbf{Z}, p \leq q, i, j \notin \{p, \dots, q\}, \rho_{i,j}^{p:q} = \rho_{i+k,j+k}^{p+k:q+k}$ . Parmi l'ensemble des parocs, la fonction  $\{\rho(p) = \rho_{0,p}^{1:p-1}\}_{p \in \mathbf{N}^*}$  (avec  $\rho(1) = \rho_{0,1}$ ) est la fonction d'autocorrélation partielle du processus. Notre but est de donner une interprétation, en termes de géométrie sphérique, de 2 algorithmes de calcul de cette fonction, l'algorithme de Schur et l'algorithme de Levinson-Szegö.

Ces algorithmes sont bien connus dans la communauté signal en tant qu'algorithmes de prédiction / interpolation linéaires ; pourtant ils sont originellement issus de disciplines différentes, ainsi que nous le rappelons brièvement.

#### 3.1 un bref historique

En 1917, Schur a développé un algorithme récursif pour tester si une fonction donnée  $S(z) = \sum_{k=0}^{\infty} s_k z^k$  était analytique et bornée par 1 dans le disque unité [7]. De telles fonctions sont caractérisées par une séquence de paramètres de module inférieur à 1 (les paramètres de Schur) qui sont calculés récursivement à partir des coefficients  $s_k$  par un algorithme élégant. Or Caratheodory et Toeplitz ont montré que  $C(z) = c_0 + 2 \sum_{k=1}^{\infty} c_k z^k$  est analytique et de partie réelle positive pour  $|z| < 1$  si et seulement si les formes de Toeplitz  $\sum_{i,j=0}^n a_i b_j c_{j-i}$ , avec  $c_{-k} = c_k$ , sont positives pour tout  $n$ . Soit

$$S(z) = \frac{C(z) - c_0}{C(z) + c_0} \iff C(z) = c_0 \frac{1 + S(z)}{1 - S(z)} ; \quad (9)$$

puisque  $|S(z)| \leq 1 \iff \operatorname{Re}(C(z)) \geq 0$ , l'algorithme de Schur permet implicitement de tester si une forme de Toeplitz est positive.

D'autre part, les formes de Toeplitz ont été étudiées indépendamment par Szegö, qui a été amené à introduire un ensemble de polynômes orthogonaux par rapport à une mesure positive sur le cercle unité. Ces polynômes obéissent à des récursions [8] où interviennent des paramètres de module inférieur à un, qui plus tard ont été reconnus comme étant égaux aux paramètres de Schur [9].

Dans les années 1940, les formes de Toeplitz ont connu un nouvel engouement du fait de leur occurrence naturelle dans la théorie de la prédiction et de l'interpolation des p.a. stationnaires qui se développait alors [10, ch. 10] [11]. En travaillant sur la solution de Wiener du problème de prédiction à temps continu, Levinson a proposé un algorithme [12] dont les récursions ne sont autres que celles de Szegö.

Ces algorithmes héritent de l'environnement naturel des disciplines mathématiques dont ils sont issus. Ils ont donc une interprétation naturelle classique en termes de géométrie non euclidienne hyperbolique (cf. §3.2). Du fait de leur lien avec la prédiction linéaire, ils reçoivent également une interprétation géométrique nouvelle, en termes de trigonométrie sphérique (cf. §3.3).

### 3.2 Interprétation géométrique non euclidienne (hyperbolique) classique

Prenons le cas de l'algorithme de Schur. Les transformations homographiques (ou de Möbius, ou linéaires fractionnaires) sont les fonctions :

$$z \mapsto z' = \frac{az + b}{cz + d}, \quad M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ inversible ;} \quad (10)$$

une propriété bien connue de ces fonctions est qu'elles transforment tout cercle du plan complexe (de rayon fini ou non) en un autre cercle de  $\mathbf{C}$  (de rayon fini ou non).

L'algorithme de Schur, qui consiste en la récursion

$$S_p(z) \mapsto S_{p+1}(z) = \frac{1}{z} \frac{S_p(z) - S_p(0)}{1 - S_p(0)S_p(z)}, \quad (11)$$

est fondé sur les propriétés du sous groupe des transformations homographiques qui préservent le cercle unité, c'est-à-dire des transformations de la forme :

$$z \mapsto z' = e^{i\theta} \frac{z - \alpha}{1 - \bar{\alpha}z}, \quad |\alpha| < 1, \quad (12)$$

associées à l'ensemble des matrices  $J$ -unitaires

$$M = \begin{bmatrix} e^{i\theta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\alpha \\ -\bar{\alpha} & 1 \end{bmatrix} \frac{1}{\sqrt{1 - |\alpha|^2}}, \quad (13)$$

c'est-à-dire vérifiant  $M^H J M = J$  avec  $J = \operatorname{diag}(1, -1)$ .

Or les transformations homographiques (12) sont des fonctions holomorphes très particulières, aux nombreuses propriétés. En particulier, elles jouent un rôle fondamental en géométrie non euclidienne de Lobachevski. Le cercle unité est en effet un modèle euclidien du plan hyperbolique  $\mathbb{H}_2$  (que ce soit le modèle conforme ou celui de Klein-Beltrami) [13] [14], dont les isométries sont précisément les transformations (12) [15] [16]. Cette géométrie est donc un cadre naturel de l'algorithme de Schur [17].

### 3.3 Interprétation géométrique non euclidienne (sphérique) nouvelle

Cette interprétation provient du lien (via les formes de Toeplitz définies positives) entre l'algorithme de Schur et la prédiction linéaire (cf. §3.1). Pour cette application de prédiction linéaire, la récursion (11) est initialisée via (9) par

$$S_0(z) = \frac{r_1 z + r_2 z^2 + \dots}{r_0 + r_1 z + r_2 z^2 + \dots}.$$

Dans ce cas, le paramètre de Schur  $S_p(0)$  est égal au coefficient de corrélation partielle  $\rho_{0,p}^{1:p-1}$ . Écrivons l'algorithme sous forme vectorielle [18] : pour  $p \geq 1$ ,

$$\begin{bmatrix} u_{p-1}^0 & v_{p-1}^1 \\ \vdots & \vdots \\ u_{p-1}^k & v_{p-1}^{k+1} \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 & -S_p(0) \\ -S_p(0) & 1 \end{bmatrix} = \begin{bmatrix} u_p^0 & 0 \\ u_p^1 & v_p^1 \\ \vdots & \vdots \\ u_p^k & v_p^k \\ \vdots & \vdots \end{bmatrix} \quad (14)$$

avec  $u_0^0 = r_0$  et  $u_0^i = v_0^i = r_i$  pour  $i > 0$ .

Du point de vue de la théorie de l'interpolation, la  $p^{\text{eme}}$  étape (14) de l'algorithme incorpore la nouvelle donnée  $r_p$

dans le problème d'extension de la fonction de covariance. Cette incorporation progressive des contraintes d'interpolation  $r_2, \dots, r_p, \dots$  correspond à l'incorporation progressive des v.a.  $\tilde{X}_{t-1}, \dots, \tilde{X}_{t-p+1}, \dots$  dans le problème de prédiction linéaire, et donc à la réactualisation progressive du projecteur associé (ceci n'est autre que l'interprétation classique de l'algorithme en termes d'orthogonalisation de Gram-Schmidt [19]). En effet, pour  $k \geq 0$ , les 2 récursions de la ligne numéro  $k$  de (14) sont 2 occurrences couplées de la même identité (5) :

$$\begin{bmatrix} (\tilde{X}_{t-p-k}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1}) & (\tilde{X}_{t-p-k}^{t-p+1:t-1}, \tilde{X}_t^{t-p+1:t-1}) \\ \left[ \begin{array}{cc} 1 & -\frac{(\tilde{X}_{t-p}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})}{(\tilde{X}_{t-p}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})} \\ -\frac{(\tilde{X}_{t-p-k}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})}{(\tilde{X}_{t-p-k}^{t-p+1:t-1}, \tilde{X}_{t-p}^{t-p+1:t-1})} & 1 \end{array} \right] \end{bmatrix} = \begin{bmatrix} (\tilde{X}_{t-p-k}^{t-p+1:t}, \tilde{X}_{t-p}^{t-p+1:t}) & (\tilde{X}_{t-p-k}^{t-p:t-1}, \tilde{X}_t^{t-p:t-1}) \end{bmatrix} \quad (15)$$

Les équations (15) sont des récursions de type CS dans  $R_{0:n}^{1:p-1} \stackrel{def}{=} (\tilde{X}_{t-i}^{t-p+1:t-1}, \tilde{X}_{t-j}^{t-p+1:t-1})_{i,j=0}^n$ . Le lien avec la TS résulte immédiatement du §2 :

**Théorème 2** *Après une normalisation adéquate, une étape élémentaire de l'algorithme de Schur s'identifie à la loi des cosinus (1) :  $\forall p \geq 1$ , et  $\forall k \geq 0$ ,*

$$\begin{bmatrix} \rho_{p+k,p}^{1:p-1}, \rho_{p+k,0}^{1:p-1} \end{bmatrix} \begin{bmatrix} 1 & -\rho_{p,0}^{1:p-1} \\ -\rho_{0,p}^{1:p-1} & 1 \end{bmatrix} \frac{1}{\sqrt{1 - (\rho_{0,p}^{1:p-1})^2}} = \begin{bmatrix} \rho_{p+k,p}^{0:p-1} \sqrt{1 - (\rho_{p+k,0}^{1:p-1})^2}, \rho_{p+k,0}^{1:p} \sqrt{1 - (\rho_{0,p}^{1:p-1})^2} \end{bmatrix} .$$

Pour finir, disons brièvement que le rôle joué par  $\{X_k\}_{k=0}^n$  dans l'algorithme de Schur peut être mis en parallèle avec celui joué par  $\{\tilde{X}_k^{[0:n]}\}_{k=0}^n$  dans l'algorithme de Levinson-Szegö. On montre alors que, convenablement normalisé, l'algorithme de Levinson-Szegö décroissant en ordre (resp. croissant en ordre) s'identifie à la loi des cosinus dans le triangle polaire (resp. la formule polaire des 5 éléments).

## 4 Quelques remarques en guise de conclusion

Nous savons depuis Klein que les géométries non euclidiennes hyperbolique et sphérique forment, avec le modèle euclidien, les 3 seules géométries projectives à courbure constante. Or les algorithmes de Schur / Levinson-Szegö s'interprètent dans ces deux géométries non euclidiennes, ce qui peut sembler contradictoire. Une réconciliation est possible au travers du lien, connu depuis Gauss, entre certaines transformations homographiques (10) (celles pour lesquelles  $M$  est unitaire) et les rotations sur la sphère.

En effet, ce lien est classiquement établi grâce à la projection stéréographique qui identifie le plan complexe (augmenté du point  $\infty$ ), assimilé à  $\mathbb{R}^2$  et immergé dans  $\mathbb{R}^3$ , à la sphère de rayon 1 et de centre 0 [15] [16]. Mais ces relations purement algébriques sont valables dans tout espace de dimension 3, et donc, par ex., dans celui engendré par les 3 v.a.  $\tilde{X}_t^{t-p+1:t-1}$ ,  $\tilde{X}_{t-p}^{t-p+1:t-1}$  et  $\tilde{X}_{t-p-k}^{t-p+1:t-1}$ , dont la manipulation forme le cœur des récursions de Schur. Des remarques similaires s'appliquent à l'algorithme de Levinson-Szegö.

## Références

- [1] G. U. Yule, On the theory of correlation for any number of variables, treated by a new system of notations, *Proc. Roy. Soc.* 79A (1907), 182-93
- [2] F. Desbouvries, Trigonométrie sphérique, identité de Yule entre PARCORs et algorithmes MCRR, *Traitement du Signal* 13.4 (1996), 303-18
- [3] S. S. Wilks, "Mathematical statistics", Wiley 1962
- [4] H. Cramer, "Mathematical methods of statistics", 13<sup>th</sup> ed., Princeton Univ. press, Princeton, NJ, 1974
- [5] T. W. Anderson, "An introduction to multivariate statistical analysis", 2<sup>nd</sup> ed., Wiley, New York, 1984
- [6] J. Whittaker, "Graphical models in applied multivariate statistics", Wiley, Chichester, 1990
- [7] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, *J. Reine Angewandte Math.* 147 (1917), 205-32; et 148 (1918), 122-45
- [8] G. Szegö, "Orthogonal polynomials", American Mathematical Society, Providence, 1939
- [9] N. I. Akhiezer, "The classical moment problem and some related questions in analysis", Hafner publishing Company, New York, 1965
- [10] U. Grenander and G. Szegö "Toeplitz forms and their applications", Univ. of California Press 1958
- [11] T. Kailath, A view of three decades of linear filtering theory, *IEEE Tr. IT* 20.2 (1974), 146-181
- [12] N. Levinson, The Wiener r. m. s. (root mean square) error criterion in filter design and prediction, *Journal of Mathematics and Physics*, 25.4 (1947), 261-78
- [13] H. S. M. Coxeter, "Introduction to geometry", 2<sup>nd</sup> edition, Wiley, 1969
- [14] J. G. Ratcliffe "Foundations of hyperbolic manifolds", Springer-Verlag, New York 1994
- [15] B.A. Dubrovin, A.T. Fomenko and S.P. Novikov, "Modern geometry - Methods and applications", Part I, Springer Verlag, New York 1993
- [16] C. Caratheodory, "Theory of functions of a complex variable", Chelsea Pub. Company, New York 1954
- [17] C. Chang and T. T. Georgiou, Geometric aspects of the Carathéodory extension problem, *Linear algebra and its applications* 203-04 (1994), 209-51
- [18] T. Kailath, A theorem of I. Schur and its impact on modern signal processing, in "Operator Theory - 18" (I. Gohberg, ed.), pp. 9-30, Birkäuser, Basel, 1986
- [19] B. Friedlander, Lattice filters for adaptive processing, *Proceedings of the IEEE* 70.8 (1982), 829-68