J. M. BRAUN

Séries chronologiques multiples : recherche d'indicateurs

Revue de statistique appliquée, tome 21, nº 1 (1973), p. 81-106
<http://www.numdam.org/item?id=RSA_1973__21_1_81_0>

SÉRIES CHRONOLOGIQUES MULTIPLES : RECHERCHE D'INDICATEURS

J. M. BRAUN (1)

I – INTRODUCTION

Lorsqu'on étudie un phénomène économique comme la consommation mensuelle d'électricité H.T. en France, il est d'usage de la représenter par la chronique des consommations mensuelles totales. Ce choix présente deux gros inconvénients :

– on perd une partie de l'information concernant la structure du système ;

– les délais nécessaires à la connaissance des derniers résultats de cette série sont souvent assez longs.

Le premier point nous amène à caractériser le phénomène étudié par la série chronologique multiple des consommations mensuelles des différentes branches de l'industrie et du commerce. Sa visualisation nous fournira celle du phénomène de façon globale.

Le deuxième point pose le problème du choix d'une, ou plus généralement d'un petit groupe de séries, extraits de la série multiple et la représentant au mieux, en un sens que l'on précisera. Ce choix est fondamental dans l'étude de la conjoncture et dans la prévison à court terme où le praticien peut à la rigueur obtenir rapidement les résultats d'une, ou d'un petit nombre de branches, alors que s'il doit les attendre tous, l'échéance d'obtention risque d'être celle de sa prévision.

Nous proposons ici une méthode permettant de résoudre ce dernier problème.

II – CONSTRUCTION D'UN MODELE

1 – Généralités

Considérons les séries chronologiques des consommations mensuelles des

- - - - - - - - - - - -

(1) Travail effectué au Commissariat à l'Energie Atomique – Département de Protection – Groupe Statistique et Calcul du Service de Protection Sanitaire.

différentes branches de l'industrie et du commerce ; sommes-nous intéressés par leur "trend" ou bien par leurs taux d'accroissement durant chaque mois ?

Pour nous, c'est la structure du "comportement" durant chaque année des séries qui présente de l'intérêt : deux branches, tout en ayant même forme de "trend", peuvent avoir des "comportements" bien différents, l'une pouvant présenter son maximum annuel en hiver, l'autre en été.

Aussi sommes-nous conduit à caractériser notre phénomène par la série chronologique multiple "des taux d'accroissement mensuels" des différentes branches :

deux branches ayant mêmes taux d'accroissement mensuels seront donc considérées comme identiques.

NOTATIONS : on considère p branches dont on observe les taux d'accroissement durant n mois.

- $x^i(t)$ représente le taux d'accroissement de la branche i durant le mois t.
- $x^i = (x^i(1), x^i(2), \ldots, x^i(n)) \in \mathbf{R}^n$ représente la ième branche, i = 1 , . . . , p .
- $x(j) = (x^1(j), x^2(j), \ldots, x^p(j)) \in \mathbf{R}^p$ représente le jième mois, j = 1,. . . , n.

2 – Choix d'une représentation et d'une distance entre séries et mois

Le problème est ici de savoir dans quel cas nous pouvons considérer que deux branches sont identiques ; doivent-elles avoir mêmes taux d'accroissement ou bien a-t-on intérêt à considérer une relation d'équivalence moins restrictive ?

a) *Translation*

Considérons deux séries (c'est-à-dire deux vecteurs de $\mathbf{R}^n$) x^i et x^j liées par la relation :

$$x^j = x^i + b\,J \tag{1}$$

où b est une constante réelle et où J est le vecteur de $\mathbf{R}^n$ dont toutes les composantes sont égales à un. Il est clair que la connaissance de l'une détermine complètement l'autre qui n'apporte donc rien quant au phénomène étudié. Nous pouvons donc les considérer comme identiques et les représenter par le même point de $\mathbf{R}^n$.

Si $$\bar{x}^j = \frac{1}{n}\sum_{k=1}^{n} x^j(k) \qquad \text{et} \qquad \bar{x}^i = \frac{1}{n}\sum_{k=1}^{n} x^i(k),$$

alors $\bar{x}^j = \bar{x}^i + b$ d'après la relation (1)

et $$x^j - \bar{x}^j \,.\, J = x^i - \bar{x}^i \,.\, J$$

Donc, si nous centrons chaque série sur sa moyenne, ce qui a un sens ici puisque nous travaillons en taux d'accroissement, deux séries qui se déduisent l'une de l'autre par translation seront représentées par le même point.

Si nous appelons *individu t* le vecteur $x(t)$ de $\mathbf{R}^p$ de coordonnées $(x^i(t))_{i=1,\dots,p}$, le fait de centrer les séries a pour effet dans $\mathbf{R}^p$ (espace des individus) de prendre pour origine le point $\bar{x}$ dont les coordonnées sont les moyennes $\bar{x}^i$ de chaque série.

Dans tout ce qui suit, nous considérons les séries des taux d'accroissement comme *centrées sur leurs moyennes.*

b) *Homothétie*

Les séries étant centrées, considérons deux séries x^i et x^j liées par la relation :

$$x^j = a \, . \, x^i \, , a \in \mathbf{R} \, , \qquad (2).$$

Le même raisonnement que pour les translations montre qu'il n'y a pas lieu de les distinguer dans leur représentation dans $\mathbf{R}^n$.

Or

$$\mathrm{Var}\,(x^j) = s_j^2 = \frac{1}{n} \sum_{k=1}^{n} (x^j(k))^2 \ ,$$

donc d'après (2),

$$s_j^2 = a^2 \, . \, s_i^2 \ , \qquad \text{soit} \qquad a = \epsilon \, . \frac{s_j}{s_i} \, ,$$

où $\qquad \epsilon = 1 \quad \text{si} \quad a > 0 \, , -1 \quad \text{si} \quad a < 0 \, ,$

et

$$x^j \, . \frac{1}{s_j} = \epsilon \, . \, x^i \, . \frac{1}{s_i}$$

– Cas où $a > 0$:

alors :

$$x^j \, . \frac{1}{s_j} = x^i \, . \frac{1}{s_i}$$

Ainsi, en considérant les séries centrées réduites, deux séries qui se déduisent l'une de l'autre par le produit d'une translation et d'une homothétie "positive" sont représentées par le même point de $\mathbf{R}^n$.

– Cas où $a < 0$:

alors :

$$x^j \, . \frac{1}{s_j} = - \, x^i \, . \frac{1}{s_i} \ ,$$

par suite, si les séries sont centrées réduites, elles sont représentées par des points symétriques par rapport à l'origine.

Le fait de réduire les séries correspond dans $\mathbf{R}^p$, espace des individus, à transformer chaque vecteur individu par l'endomorphisme de matrice $D_{1/s}$ telle que :

$$(D_{1/s})_{ij} = \frac{\delta_{ij}}{s_i} \quad \text{où} \quad \delta_{ij} = \begin{cases} 1 & \text{si} \quad i = j \\ & 0 \text{ sinon} \end{cases}$$

A partir de maintenant, nous allons représenter les *séries centrées réduites* de manière à ce que deux séries se déduisant l'une de l'autre par une transformation affine :

$$x^j = a\,x^i + b\,.\,J$$

soient représentées par le même point, ou des points opposés, dans $\mathbf{R}^n$. Cela revient donc à considérer la relation d'équivalence $\mathcal{R}$ définie par :

$$x^i\,\mathcal{R}\,x^j \Leftrightarrow \exists\ (a\,,b) \in \mathbf{R}^2 / x^j = a\,x^i + b\,.\,J$$

et à travailler sur l'espace quotient.

Nous définissons un produit scalaire dans $\mathbf{R}^n$ par la donnée da la matrice diagonale $D = \frac{1}{n} I_n$, où I_n est la matrice identité d'ordre n. Ce produit scalaire n'est autre que celui de la "covariance", les poids affectés à chacune des observations étant tous égaux à $\frac{1}{n}$ (les séries x^i sont centrée-réduites)

$$< x^i\,, x^j > = {}^t x^i\,D\,x^j = D(x^i\,, x^j)\ (1)$$

$$= \frac{1}{n} \sum_{k=1}^{n} x^i(k)\,x^j(k)$$

$= \mathrm{cov}\,(x^i\,, x^j)$ puisque les séries sont centrées ;

$= \mathrm{corr}\,(x^i\,, x^j)$ puisque les séries sont réduites.

- $//x^i//^2 = {}^t x^i\,D\,x^i = \frac{1}{n} \sum_{k=1}^{n} (x^i(k))^2$

  $= \mathrm{Var}\,(x^i)$, (séries centrées)

  $= 1$ (séries réduites)

- $d^2(x^i\,, x^j) = //x^i - x^j//^2 = \mathrm{Var}\,(x^i - x^j)$

  $= 2\,(1 - \mathrm{corr}\,(x^i\,, x^j))$

---

(1) En confondant dans nos notations la matrice (tenseur) D et la forme bilinéaire qu'elle définit.

Dans R^p, espace des individus, les caractères "séries" étant centrés réduits, nous prenons le produit scalaire classique défini par la matrice identité I_p. Deux individus sont une distance nulle s'ils ont mêmes coordonnées.

En résumé, nous allons :

– *représenter dans R^n les p branches* par les p vecteurs de leurs taux d'accroissement mensuels centrés réduits ;

– considérer R^n muni du produit scalaire de la "covariance" ;

– *représenter dans R^p les n individus "mois"* par les n vecteurs des taux d'accroissement centrés réduits dans p branches ;

– considérer R^p muni du produit scalaire classique défini par I_p.

3 – ANALYSE EN COMPOSANTES PRINCIPALES SUR LES SERIES

Compte tenu du choix des représentations des séries et des mois, pour obtenir une série expliquant au mieux la variabilité du phénomène, il suffit de retenir la première composante principale du groupe de p séries. Cette dernière n'expliquant qu'une partie de la variabilité totale, on peut considérer la deuxième composante principale qui, parmi les combinaisons linéaires des p séries orthogonales à la première composante principale, explique au mieux la variabilité restante, et ainsi de suite. Toutefois, pour des raisons d'interprétations et de rapidité d'obtention, l'utilisateur préférerait sélectionner, pour expliquer au mieux le phénomène, un ou des séries parmi le groupe initial.

Pour extraire un sous-groupe de p_0 ($p_0 < p$) séries représentatives du phénomène, l'idéal serait d'en trouver p_0 admettant les mêmes composantes principales que le groupe initial ; à défaut, nous prendrons les p_0 séries dont les composantes principales sont "le plus proche" de celles du groupe initial. Afin de quantifier cette proximité, nous allons, dans le paragraphe III, remplacer le groupe et les sous-groupes par des opérateurs, puis définir une distance entre opérateurs, nulle si les groupes de variables dont ils proviennent ont mêmes composantes principales (cf. Y. ESCOUFIER).

III – CONSTRUCTION ET ETUDE DE L'OPERATEUR (1)

1 – Définition

Soit U_p l'opérateur de R^n dans R^n associé à un groupe de p séries $(x^i)_{i=1,\dots,p}$(2) et tel que :

$$\forall\, y \in R^n , \qquad U_p(y) = \frac{1}{p} \sum_{i=1}^{p} D(x^i , y)\, x^i$$

----

(1) Nous nous plaçons ici dans le cas descriptif ; pour une étude "dans la population" voir ESCOUFIER (2).

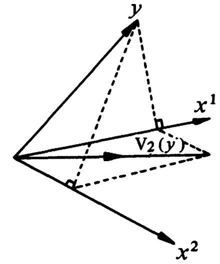
(2) Pour notre étude les séries sont centrées réduites, toutefois la théorie des opérateurs reste valable dans le cas où on n'a pas réduit.

Remarque : dans le cas d'une seule série x^1 centrée réduite l'opérateur V_1 associé à x^1 est un projecteur,

$$V_1(y) = D(x^1, y) x^i,$$

par contre, dès que $p \geqslant 2$, ce n'est plus vrai, ainsi si $p = 2$,

$$V_2(y) = \frac{1}{2}[D(x^1, y) x^1 + D(x^2, y) x^2]$$



2 – Propriétés

Théorème 1

L'opérateur U_p est un opérateur linéaire, D – symétrique, positif.

La linéarité est évidente, montrons qu'il est D – symétrique :

$$z \in \mathbf{R}^n, < U_p(y), z > = D(U_p(y), z) = \frac{1}{p}\sum_{i=1}^{p} D(x^i, y) D(x^i, z)$$

$$= < y, U_p(z) > \qquad \text{CQFD}$$

U_p admet donc un système de vecteurs propres D – orthonormés –

$$< U_p(y), y > = \frac{1}{p}\sum_{i=1}^{p} D(x^i, y) D(x^i, y) \geqslant 0,$$

U_p est donc positif.

Théorème 2

Les vecteurs propres de U_p associés aux valeurs propres $\mu \neq 0$ sont les composantes principales du groupe $(x^i)_{i=1,\dots,p}$ "associées" aux valeurs propres $\lambda = \mu \cdot p$ de la matrice de variance-covariance de $(x^i)_{i=1,\dots,p}$.

Soit W le sous-espace vectoriel de $\mathbf{R}^n$ engendré par les x^i, x^i, $i = 1, \dots, p$ et soit $W^\perp$ le sous-espace D – orthogonal à W dans $\mathbf{R}^n$, alors

$$\mathbf{R}^n = W \oplus W^\perp$$

$W^\perp$ est contenu dans le noyau de U_p, donc les vecteurs propres de U_p associés à des valeurs propres non nulles se trouvent dans W. Soit v un de ces vecteurs,

$$v = \sum_{i=1}^{p} \alpha_i x^i$$

et

$$U_p(v) = \frac{1}{p} \sum_{j=1}^{p} D(x^j, v)\, x^j = \mu v$$

Or :

$$D(x^j, v) = \sum_{i=1}^{p} \alpha_i \,.\, D(x^j, x^i)$$

donc :

$$U_p(v) = \frac{1}{p} \sum_{j=1}^{p} \sum_{i=1}^{p} \alpha_i \,.\, D(x^j, x^i)\, x^j = \mu \,.\, v = \mu \sum_{i=1}^{p} \alpha_j \,.\, x^j$$

et pour chaque j , $\frac{1}{p} \sum_{i=1}^{p} \alpha_i \,.\, D(x^j, x^i) = \mu \,.\, \alpha_j$, (1)

or $D(x^j, x^i) = \left(\sum\right)_{ij}$ où $\sum$ est la matrice de variance-covariance du groupe $(x^i)_{i=1,\dots,p}$ que l'on suppose centré. (1) peut donc s'écrire :

pour chaque $\quad j , \sum_{i=1}^{p} \alpha_i \,.\, \left(\sum\right)_{ji} = (p \,.\, \mu)\, \alpha_j$

ce qui traduit le fait que le vecteur $\alpha = (\alpha_1, \dots, \alpha_p)$ est vecteur propre de $\sum$ associé à la valeur propre $\lambda = p \,.\, \mu$; or, nous savons que si $\sum_{i=1}^{p} \alpha_i^2 = 1$, le vecteur $v = \sum_{i=1}^{p} \alpha_i x^i$ est une composante principale du groupe $(x^i)_{i=1,\dots,p}$.

Corollaire 1

Les valeurs propres de U_p sont positives ou nulles. Soient $(c^i)_{i=1,\dots,p}$ les composantes principales de $(x^i)_{i=1,\dots,p}$, alors

$$\forall y \in R^n , U_p(y) = \frac{1}{p} \sum_{i=1}^{p} D(c^i, y)\, c^i$$

L'opérateur U_p étant D – symétrique et positif a des valeurs propres positives ou nulles.

Rappels :

- $D(c^i, c^i) = //c^i//^2 = \mathrm{Var}(c^i) = \lambda_i$ ième valeur propre de $\sum$

• $D(x^i, c^j) = \lambda_j \cdot \alpha_j^i$ où α_j^i est la ième coordonnée du jième vecteur propre normé $\left(\sum_{i=1}^{p} (\alpha_j^i)^2 = 1\right)$ de Σ

Posons

$$\varphi^j = \frac{c^j}{\sqrt{\lambda_j}}, \ j = 1, \ldots, p \quad \text{alors} \quad //\varphi^j//^2 = 1$$

$$x^i = \sum_{j=1}^{p} D(x^i, \varphi^j)\, \varphi^j$$

$$= \sum_{j=1}^{p} \sqrt{\lambda_j} \cdot \alpha_j^i \cdot \varphi^j$$

$$U_p(y) = \frac{1}{p} \sum_{i=1}^{p} D(x^i, y)\, x^i$$

$$= \frac{1}{p} \sum_{i=1}^{p} \sum_{j=1}^{p} \sqrt{\lambda_j}\, \alpha_j^i \cdot D(\varphi^j, y)\, x^i$$

$$= \frac{1}{p} \sum_{j=1}^{p} \sqrt{\lambda_j} \cdot D(\varphi^j \cdot y) \sum_{i=1}^{p} \alpha_j^i \cdot x^i$$

or

$$\sum_{i=1}^{p} \alpha_j^i \cdot x^i = c^j$$

Donc

$$U_p(y) = \frac{1}{p} \sum_{j=1}^{p} \sqrt{\lambda_j} \cdot D(\varphi^j, y)\, c^i$$

Soit

$$U_p(y) = \frac{1}{p} \sum_{j=1}^{p} D(c^j, y)\, c^j$$

De ceci il découle que :

Corollaire 2

Le groupe $(x^i)_{i=1,\ldots,p}$ et le groupe de ses composantes principales $(c^j)_{j=1,\ldots,p}$ ont les mêmes opérateurs associés.

Ainsi à chaque groupe de séries de $\mathbf{R}^n$, on sait faire correspondre un opérateur linéaire D – symétrique *positif* de rang $\leqslant$ n. Toutefois, nous devons remarquer que si U_p est un opérateur de ce type, $V_p = - U_p$ ne le sera plus du fait que V_p est alors un opérateur négatif. Les valeurs propres de V_p seront donc négatives ou nulles, ce qui nous interdit d'espérer trouver un groupe de séries de $\mathbf{R}^n$ dont V_p soit l'opérateur associé au sens que nous avons défini.

Ceci est lourd de conséquence ; en effet, étant donné U_p associé au groupe de p séries $(x^i)_{i=1,\ldots,p}$, nous aurions aimé pouvoir associer un groupe de séries à l'opérateur $V_p = a \cdot U_p$, $a \in \mathbf{R}$. La remarque précédente permet d'affirmer que si $a < 0$, ce n'est pas possible. Par contre, si $a > 0$, il est évident que V_p est l'opérateur associé au groupe de séries $(\sqrt{a} \cdot x^i)_{i=1,\ldots,p}$.

Corrolaire 3

Si U_p est l'opérateur associé au groupe de séries $(x^i)_{i=1,\ldots,p}$ de R^n, alors si $a > 0$, l'opérateur $V_p = a \cdot U_p$ est associé au groupe $(\sqrt{a} \cdot x^i)_{i=1,\ldots,p}$; par contre si $a < 0$, il n'existe pas de groupe de séries dont V_p soit l'opérateur associé au sens de III.1.

Ceci restreint le champ d'application de la méthode qui consiste à remplacer un groupe de séries par un opérateur. En effet, étant donné plusieurs groupes de séries, une idée naturelle consiste à remplacer chaque groupe par un opérateur et faire ensuite des analyses sur les opérateurs, or, par exemple, dans le cas d'une analyse en composantes principales sur opérateurs, lorsqu'on obtiendra les "opérateurs composantes principales", dans la majorité des cas on ne pourra pas trouver de groupe de séries dont "l'opérateur composante principale" soit l'opérateur associé au sens de III.1.

3 – Définition d'une distance entre opérateurs

Soit $\mathcal{B}$ la classe des opérateurs linéaires D – symétriques de $\mathbf{R}^n$.

$$\forall\, U_i , V_j \in \mathcal{B} , \quad \text{posons} \quad \langle\langle U_i , V_j \rangle\rangle = \sum_{k=1}^{n} D(U_i(e_k) , V_j(e_k)), \qquad (1),$$

où $(e_k)_{k=1,\ldots,n}$ est une base D – orthonormée de $\mathbf{R}^n$. (1) définit bien un produit scalaire, c'est une forme hermitienne positive. (1) est indépendant de la base D – orthonormée choisie. Pour le démontrer, il suffit d'en prendre une autre, la démonstration repose alors sur le fait que la matrice de passage est orthogonale.

Considérons l'ensemble des vecteurs propres D – orthonormés $(v^k)_{k=1,\ldots,p_i}$, de U_i associés aux valeurs propres $(\lambda_k)_{k=1,\ldots,p_i}$ différentes de zéro ; en le complétant par une base D – orthonormée du noyau de U_i, on obtient une base D – orthonormée de $\mathbf{R}^n$, par suite

$$\langle\langle U_i , V_j \rangle\rangle = \sum_{k=1}^{p_i} D(U_i(v^k) , V_j(v^k))$$

$$= \sum_{k=1}^{p_j} \lambda_k \cdot D(v^k, V_j(v^k))$$

Soit maintenant, l'ensemble des vecteurs propres D – orthonormés $(w^k)_{k=1,\dots,p_j}$ de V_j associés aux valeurs propres $(\mu_k)_{k=1,\dots,p_j}$ différentes de zéro. En le complétant par des vecteurs D – orthonormés du noyau de V_j, on obtient une autre base D – orthonormée de $\mathbf{R}^n$, et

$$v^k = \sum_{l=1}^{p_j} D(v^k, w^l)\, w^l + h \qquad \text{où} \qquad h \in N(V_j)$$

$$V_j(v^k) = \sum_{l=1}^{p_j} D(v^k, w^l)\, V_j(w^l)$$

$$= \sum_{l=1}^{p_j} D(v^k, w^l) \cdot \mu_l \cdot w^l$$

Donc $$\boxed{\langle\langle U_i, V_j\rangle\rangle = \sum_{k=1}^{p_j} \sum_{l=1}^{p_j} \lambda_k \cdot \mu_l \cdot D^2(v^k, w^l)}$$

On en déduit $|||U_i|||^2 = \sum_{k=1}^{p_j} \lambda_k^2$ Or $d^2(U_i, V_j) = |||U_i - V_j|||^2$

donc $$\boxed{d^2(U_i, V_j) = \sum_{k=1}^{p_i} \lambda_k^2 + \sum_{l=1}^{p_j} \mu_l^2 - 2 \sum_{k=1}^{p_i} \sum_{l=1}^{p_j} \lambda_k \cdot \mu_l \cdot D^2(v^k, w^l)} \qquad (2)$$

De (2), ou du corollaire 2, on déduit :

Corollaire 4

Deux groupes de séries ayant mêmes composantes principales ont des opérateurs associés confondus.

Remarque : Les opérateurs associés à deux groupes de séries étant *positifs,* les λ_k et les μ_l le sont aussi. Donc

$$\langle\langle U_i, V_j\rangle\rangle \geqslant 0$$

Nous allons maintenant énoncer une propriété d'algèbre, démontrée à l'annexe 1, permettant de calculer de façon plus économique la distance entre deux opérateurs.

Supposons que nous ayons deux groupes de séries $(x^i)_{i=1,\dots p_i}$ et $(y^j)_{j=1,\dots,p_j}$ où x^i et $y^j \in \mathbf{R}^n$, quels que soient i et j. Soient $(\varphi^k)_{k=1,\dots,p_i}$

l'ensemble des composantes principales D – orthonormées de $(x^i)_{i=1,\dots,p_i}$ "associées" aux valeurs propres $(\lambda_k)_{k=1,\dots,p_i}$, de même soient $(\Psi^k)_{k=1,\dots,p_j}$, l'ensemble des composantes principales D – orthonormées de $(y^j)_{j=1,\dots,p_j}$ "associées" aux valeurs propres $(\mu_k)_{k=1,\dots,p_j}$, alors

Proposition 1

$$\sum_{i=1}^{p_i} \sum_{j=1}^{p_j} D^2(x^i , y^j) = \sum_{i=1}^{p_i} \sum_{j=1}^{p_j} \lambda_i \mu_j D^2(\varphi^i , \Psi^j)$$

Soient les groupes de séries $(x^i)_{i=1,\dots,p_i}$ et $(y^j)_{j=1,\dots,p_j}$ et Σ_1 , Σ_2 leurs matrices de variances-covariances, alors d'après le théorème 2 si les $(\lambda_k)_{k=1,\dots,p_i}$ (resp les $(\mu_l)_{l=1,\dots,p_j}$) sont les valeurs propres de Σ_1 $\Big($resp Σ_2 , les $\left(\frac{\lambda_k}{p_i}\right)_{k=1,\dots,p_i}$ $\left(\text{resp les } \left(\frac{\mu_l}{p_j}\right)_{l=1,\dots,p_j}\right)$ sont les valeurs propres de l'opérateur associé U_{p_i} (resp V_{p_j}), les vecteurs propres étant les composantes principales, donc d'après la proposition 1

$$\boxed{\begin{aligned} &\langle\langle U_{p_i} , V_{p_j}\rangle\rangle = \frac{1}{p_i \cdot p_j} \sum_{k=1}^{p_i} \sum_{l=1}^{p_j} D^2(x^k , y^l) \\ &\text{et} \quad |||U_{p_i}|||^2 = \frac{1}{p_i^2} \sum_{k=1}^{p_i} \sum_{l=1}^{p_j} D^2(x^k \; x^l) \end{aligned}}$$

IV – RECHERCHE DE SERIES "RESUMES"

1 – Définitions

Etant donné la série chronologique multiple à p composantes caractérisant le phénomène étudié, nous appelons "série résumée" la série extraite la plus représentative.

Nous l'obtiendrons en remplaçant la série multiple, ainsi que les séries composantes, par leurs opérateurs associés et en choisissant celle dont l'opérateur est le plus proche de V, opérateur associé à la série multiple.

De la même manière, nous rechercherons le sous-groupe de $p_0 (p_0 < p)$, extrait de la série multiple, le plus représentatif.

2 – Etude de la signification de la proximité entre la série multiple et la série résumée

Par définition $\forall y \in \mathbf{R}^n \, , \mathrm{V}(y) = \frac{1}{\mathrm{p}} \sum_{\mathrm{i}=1}^{\mathrm{p}} \mathrm{D}(x^{\mathrm{i}} \, , y) \, x^{\mathrm{i}}$.

D'après le corollaire 2,

$\mathrm{V}(y) = \frac{1}{\mathrm{p}} \sum_{\mathrm{i}=1}^{\mathrm{p}} \mathrm{D}(c^{\mathrm{i}} \, , y) c^{\mathrm{i}}$ où les c^{i} sont les composantes principales du groupe $(x^{\mathrm{i}})_{\mathrm{i}=1,\ldots,\mathrm{p}}$.

soit $\mathrm{U}_{1\mathrm{i}}$ l'opérateur associé à x^{i},

$$\forall \, \mathrm{y} \in \mathbf{R}^n \, , \mathrm{U}_{1\mathrm{i}}(y) = \mathrm{D}(x^{\mathrm{i}} \, , y) x^{\mathrm{i}}$$

$$\mathrm{d}^2(\mathrm{V} \, , \mathrm{U}_{1\mathrm{i}}) = |||\mathrm{V}|||^2 + |||\mathrm{U}_{1\mathrm{i}}|||^2 - \frac{2}{\mathrm{p}} \sum_{\mathrm{k}=1}^{\mathrm{p}} \mathrm{D}^2(c^{\mathrm{k}} \, , x^{\mathrm{i}}), \quad \text{or} \quad |||\mathrm{U}_{1\mathrm{i}}|||^2 = 1$$

puisque les séries sont réduites, donc rechercher la série x^{i_0} qui minimise $\mathrm{d}(\mathrm{V} \, , \mathrm{U}_{1\mathrm{i}})$ revient à chercher celle qui maximise

$$\sum_{\mathrm{k}=1}^{\mathrm{p}} \mathrm{D}^2(c^{\mathrm{k}} \, , x^{\mathrm{i}})$$

$$= \sum_{\mathrm{k}=1}^{\mathrm{p}} \lambda_{\mathrm{k}} \, . \, \mathrm{D}^2(\varphi^{\mathrm{k}} \, , x^{\mathrm{i}}) \quad \text{où} \quad \varphi^k = \frac{c^{\mathrm{k}}}{\sqrt{\lambda_{\mathrm{k}}}} \, , \; ||\varphi^{\mathrm{k}}||^2 = 1$$

$$= \sum_{\mathrm{k}=1}^{\mathrm{p}} \lambda_{\mathrm{k}} \, . \, (\mathrm{corr} \, (\varphi^{\mathrm{k}} \, , x^{\mathrm{i}}))^2 \quad , \quad (1)$$

$$\leqslant \lambda_1 \, . \sum_{\mathrm{k}=1}^{\mathrm{p}} \mathrm{D}^2(\varphi^{\mathrm{k}} \, , x^{\mathrm{i}}) = \lambda_1 \qquad \text{plus grande valeur propre.}$$

Donc, si on se permet de chercher une série combinaison linéaire des séries du groupe, la première composante principale réduite constitue le meilleur résumé possible parmi les combinaisons linéaires des p séries composantes de la série multiple. D'autre part, dans le cas où $\frac{\lambda_1}{\sum_{\mathrm{i}} \lambda_{\mathrm{i}}}$ est proche de 1, il est évident, qu'à défaut de φ^1, la série résumée sera la plus proche de φ^1, c'est-à-dire la plus corrélée en valeur absolue avec φ^1.

Il est donc clair qu'avec cette méthode, on trouve la série la plus proche du groupe initial (série multiple) au sens des premières composantes principales ;

dans le cas où $\frac{\lambda_1}{\sum_i \lambda_i}$ est proche de un, on trouve la plus proche au sens de la première composante principale. Ainsi, si dans notre exemple on travaille sur les consommations mensuelles, la première composante principale traduit le trend global et explique une forte proportion de la variabilité ; on trouvera alors la série la plus proche au sens du trend.

3 – Recherche de groupes de séries "résumés"

Comme en régression "pas à pas", pour des problèmes de coût de calcul, on va utiliser une méthode "ascendante" ou bien "descendante" ou les deux.

Méthode ascendante

Etant donné la série "résumé" x^{i_0}, on va chercher, parmi les séries restantes, celle qui avec x^{i_0} constitue le groupe de deux le plus proche de la série multiple

Soit U_{2i} l'opérateur associé au groupe de deux séries :

$$x^{i_0}\ , x^i\ , d^2(V\ , U_{2_i}) = |||V|||^2 + |||U_{2_i}|||^2 - \frac{2}{2 \cdot p} \sum_{k=1}^{p} (D^2(x^k\ , x^{i_0}) + D^2(x^k\ , x^i))$$

Posons $$A(1\ , 1) = D^2(x^{i_0}\ , x^{i_0})\ .$$

Alors $$//U_{2i}//^2 = A(2\ , 2) = \frac{1}{4}[A(1\ , 1) + 2\,D^2(x^{i_0}\ , x^i) + D^2(x^i\ , x^i)]$$

De même posons :

$$A(p\ , 1) = \frac{1}{p} \sum_{k=1}^{p} D^2(x^k\ , x^{i_0})\ ,$$

alors $$A(p\ , 2) = \frac{1}{2} A(p\ , 1) + \frac{1}{p \cdot 2} \sum_{k=1}^{p} D^2(x^k\ , x^i)\ ,$$

et $$d^2(V\ , U_{2i}) = ///V///^2 + A(2\ , 2) - 2\,A(p\ , 2)\ .$$

Nous allons donc rajouter la série, choisie parmi les séries restantes, qui est la moins corrélée avec x^{i_0}, déjà choisie, et la plus corrélée avec les composantes de la série multiple.

Plus généralement, si nous connaissons le meilleur groupe de m $(m < p - 1)$ séries construites comme précédemment et si nous notons $(y^i)_{i=1,\ldots,m}$ les séries de ce groupe, nous aurons pour le groupe de $(m + 1)$

$$d^2(V\ , U_{(m+1)i}) = V^2 + A(m + 1)\ , m + 1) - 2\,A(p\ , m + 1)$$

où

$$A(m+1\,,m+1) = \frac{m^2}{(m+1)^2}\,A(m\,,m) + \frac{2}{(m+1)^2}\sum_{k=1}^{m} D^2(y^k\,,x^i)$$

$$+\frac{1}{(m+1)^2}\,D^2(x^i,x^i),$$

x^i étant la série que l'on se propose d'introduire si elle minimise $d^2(V\,,U_{(m+1)i})$.

$$A(p\,,m+1) = \frac{m}{m+1}\,A(p\,,m) + \frac{1}{p\,.\,(m+1)}\sum_{k=1}^{p} D^2(x^k\,,x^i)$$

Nous obtenons ainsi une méthode itérative assez économique sur le plan calcul. Toutefois, il est fondamental d'observer que cette méthode "ascendante" n'est valable que si m est petit devant p, sinon on doit faire la méthode "descendante" qui consiste à enlever du groupe de p séries celle dont l'opérateur associé est le plus éloigné de V ; et ainsi de suite.

Une méthode plus coûteuse, mais qui fournit de meilleurs résultats pour m voisin de $\frac{p}{2}$ est de faire les deux méthodes, ascendante et descendante, et de choisir le meilleur des deux groupes de m séries obtenus.

4 – Exemple

Recherche de séries "résumé" pour le phénomène : "consommation mensuelle d'électricité H T en France de janvier 1956 à décembre 1969, mai 1968 ayant été retiré et les points aberrants corrigés", que l'on caractérise par les taux d'accroissement mensuels de 17 branches de l'industrie et du commerce : 1 houillères (1) (01), 2 pétrole (03, 04), 3 gaz (052), 4 mines (11, 12), 5 sidérurgie (13, 14, 151, 153, . . . , 158), 6 transformation des métaux (21 , 22 , . . ., 29), 7 extraction de minéraux divers (32, 35, 361, 362, 31, 33, 34, 365), 8 bâtiment et T.P. (37, 38), 9 chimie (411, 412, 413, 42, 431, . . . , 435, 40, 44 (sauf 440), 45, 48), 10 verre (46), 11 caoutchouc (47), 12 production et transformation de produits agricoles et alimentaires (groupe V), 13 filatures et tissages (61), 14 papier et carton (65), 15 industries annexes des textiles, cuir, bois, industries diverses (62, 63, 64, 66, 67, 68), 16 transports (groupe VII), 17 commerce et administration (groupe VIII et IX).

Pour des raisons déjà exposées, nous avons considéré les taux d'accroissement $\left(\frac{x^i(t) - x^i(t-1)}{x^i(t)}\right)_{i,t}$ centrés réduits.

---

(1) Pour chaque branche, le premier numéro est notre identificateur, les suivants sont ceux de la nomenclature E.D.F. (1969).

La proximité entre opérateurs s'interprétant en termes de proximités avec les composantes principales, il nous semble utile de d'abord faire une analyse en composantes principales, les individus étant les mois, les caractéres étant les branches numérotées de 1 à 17.

Analyse en composantes principales

- Etude de la matrice de corrélation

Valeurs propres	Contributions à l'inertie	Contributions cumulées
9,356	0,5504	0,5504
2,105	0,1238	0,6742
1,732	0,1019	0,7761

- Représentation des branches dans le plan des 2 premières composantes principales

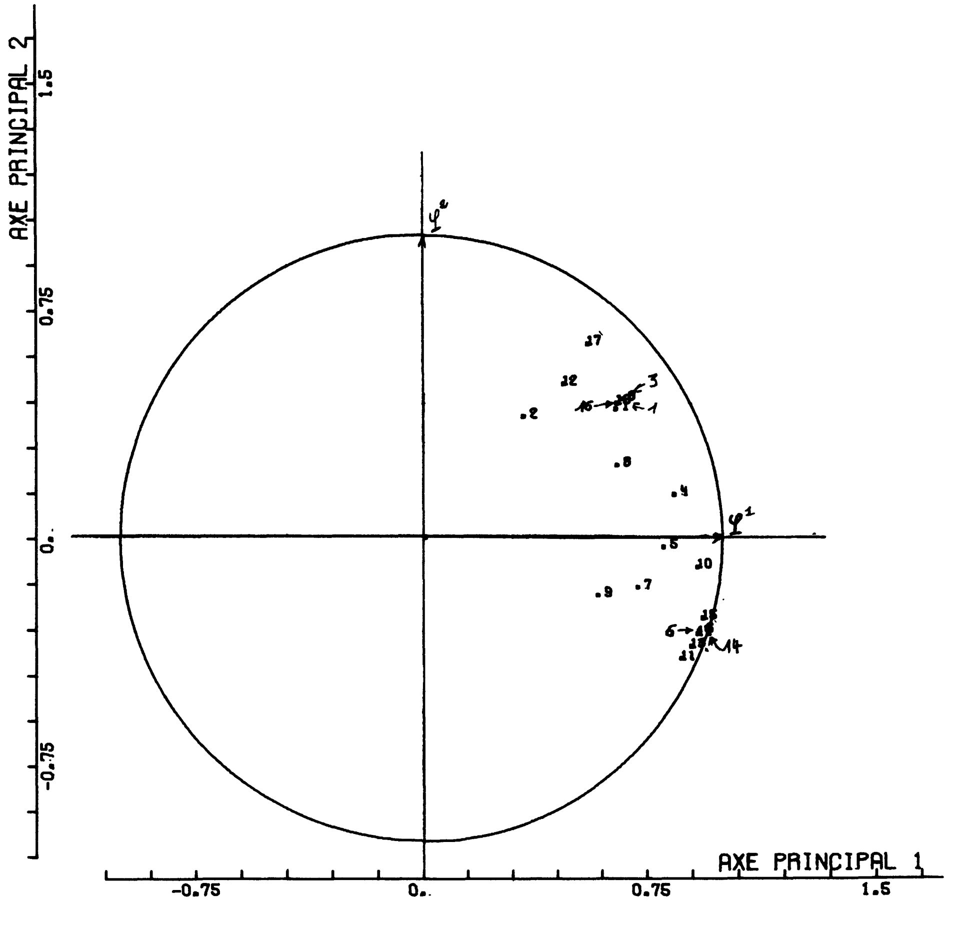


Figure 1 – A.C.P. sur les taux d'accroissement des 17 branches.

La première composante principale traduit un effet "taille" et met en évidence les branches (11, 13, 14, 6, 15, 5, 10) les plus corrélées entre elles. Ces dernières ont un comportement saisonnier analogue surtout caractérisé par la permanence des mois extrêmes, août et septembre, Parmi elles, celles (11, 13, 6, 14, 15) qui ont les plus fortes corrélations (> 91 %) ont un taux d'accroissement très faible en août et très fort en septembre (fig. 2 (1)) comparativement aux autres mois, ce qui est certainement la cause de cette forte liaison.

Les résultats de l'analyse, et la série résumée qui en découle, devront donc être interprétées avec précaution. Pour une exploitation pratique des résultats, le traitement devra être refait en éliminant ces deux mois extrêmes.

Sur la deuxième composante principale, les branches se regroupent suivant leurs saisonnalités et plus particulièrement suivant leurs extrêmums annuels. Ainsi le groupe (11, 13, 6, 14, 15) s'oppose au groupe (17, 12, 16, 1, 3) dont les branches n'ont pas d'extrêmums annuels en août et septembre (fig. 3) ; la branche 5 (fig. 4) qui en a seulement certains durant ces deux mois, n'est que très faiblement prise en compte

En vue de la construction de la série "résumé" il est fondamental d'observer que les branches (11, 13, 6, 14, 15) sont bien représentées dans le plan des deux premières composantes principales. Par conséquent, leurs projections sur les autres seront faibles. La série résumée étant celle qui maximise la quantité :

$$\sum_{k=1}^{p} \lambda_k \cdot D^2(\varphi^k , x^i) \leqslant \lambda_1 , \tag{1}$$

appartiendra donc nécessairement à ce groupe. Dans celui-ci, la branche 15 étant la plus correlée avec la première composante principale devrait être série résumée. Parmi les 17 branches, c'est elle qui représente le mieux le phénomène au sens de la première composante principale, c'est-à-dire ici de la "taille" et du comportement saisonnier en août et septembre.

- - - - - - - - - - - - - -

(1) Sur les graphiques des différentes séries, les mois sont numérotés de 1(FE 1956) à 165 (DE 1969), les mois de mai et juin 1968 ayant été sautés. D'autre part, certains points "mois" ont aussi été repérés par deux lettres : JA, FE, MR, AV, MA, JN, JL, AO, SE, OC, NO, DE.

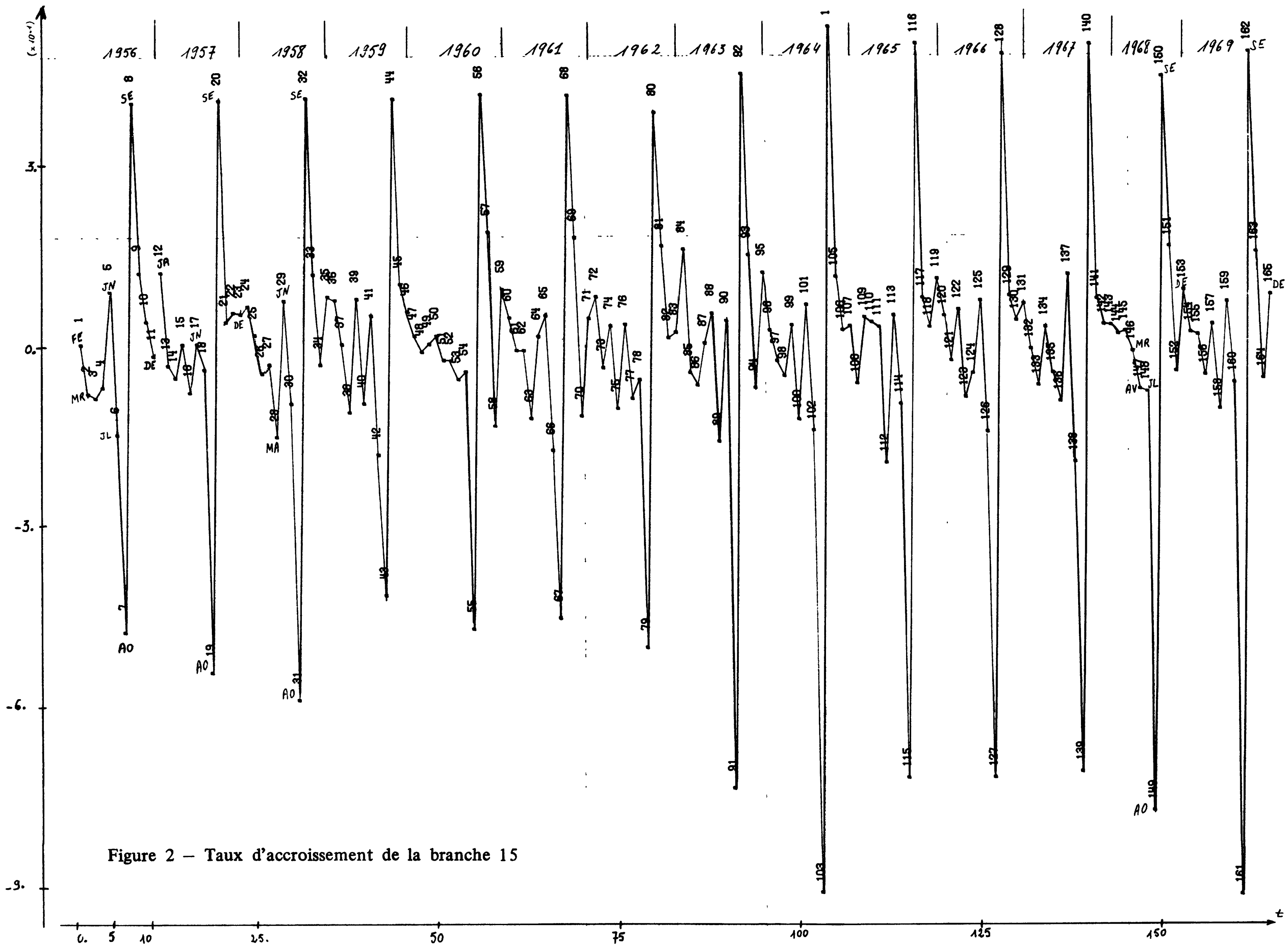


Figure 2 – Taux d'accroissement de la branche 15

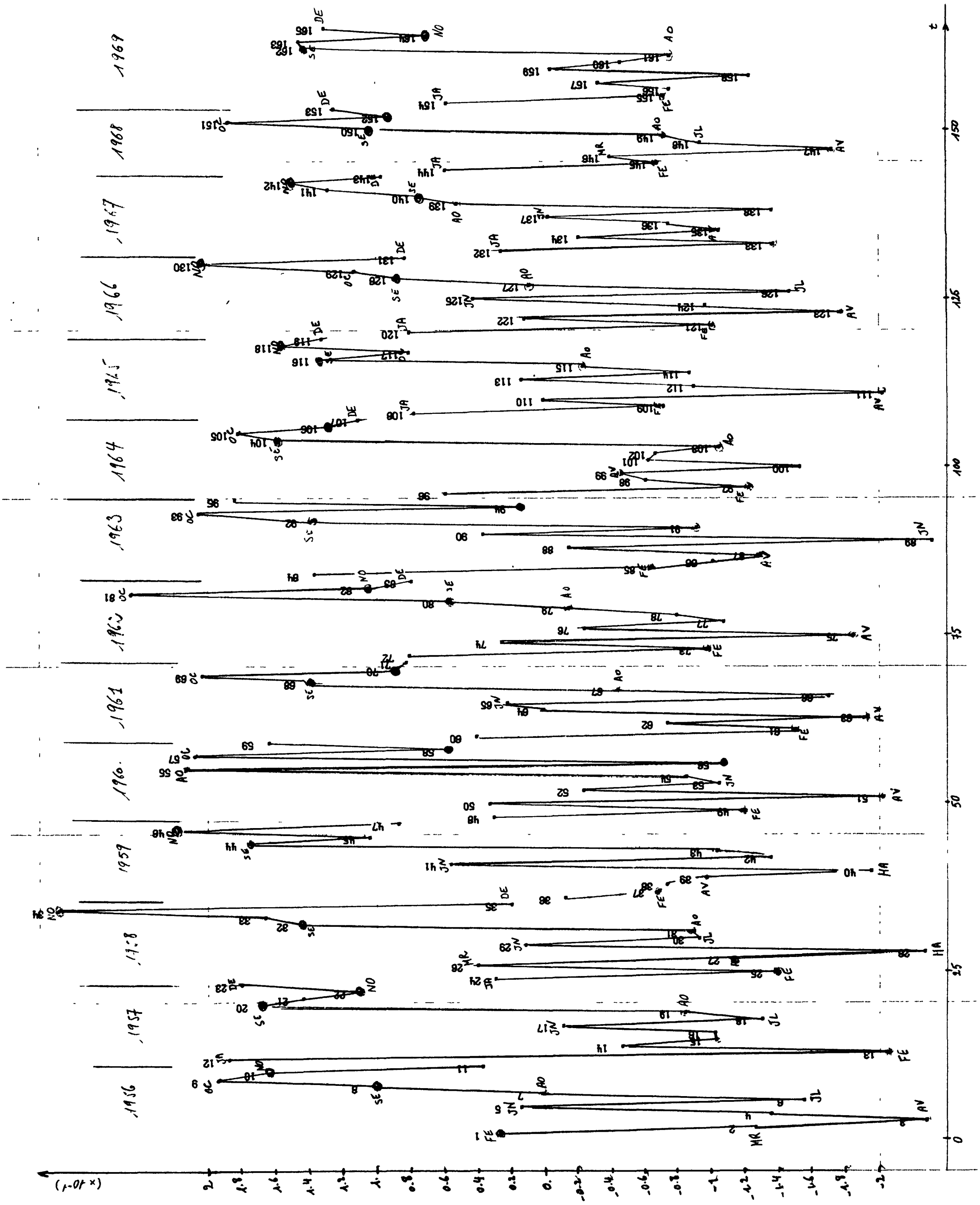


Figure 3 — Taux d'accroissement de la branche 17

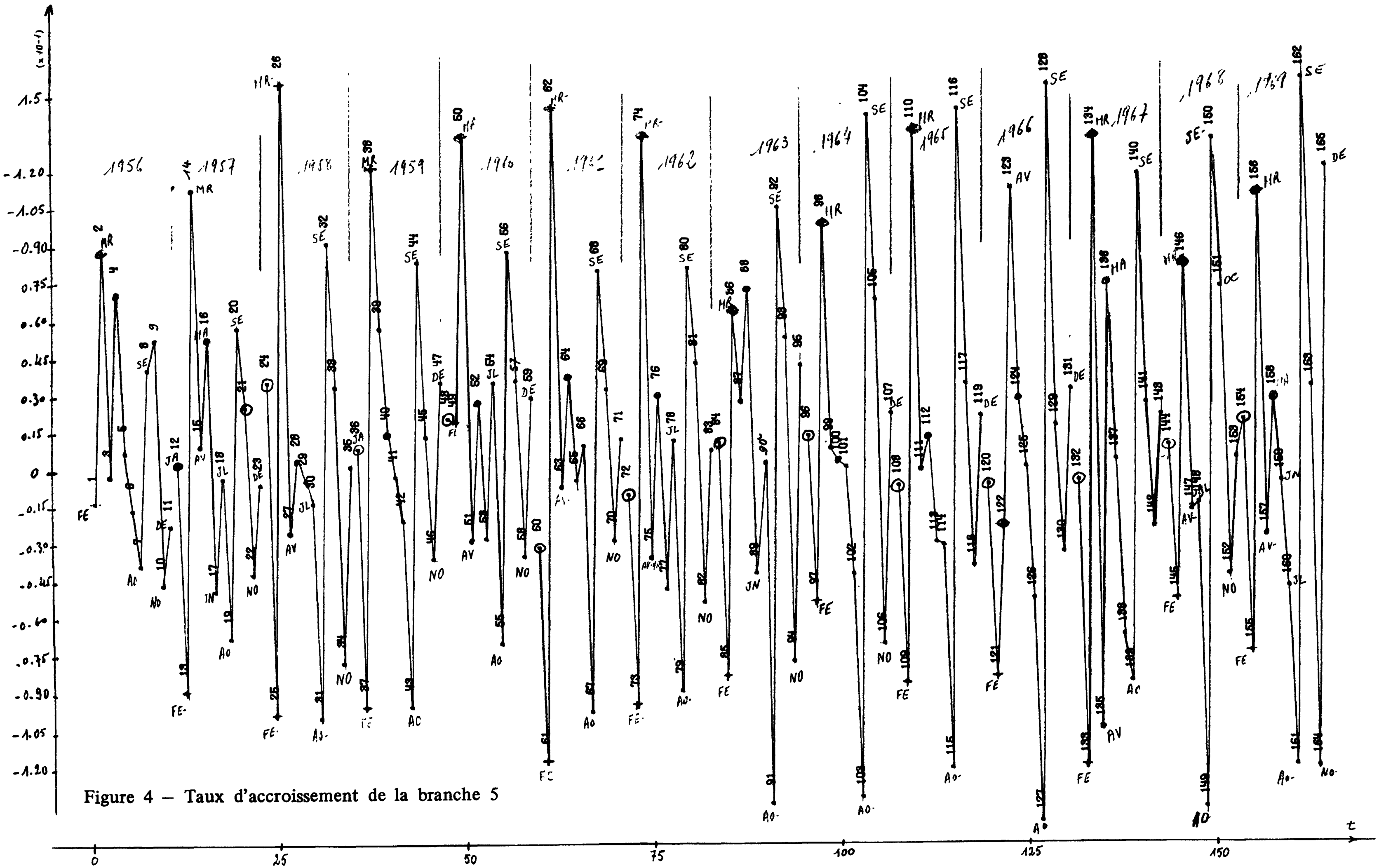


Figure 4 – Taux d'accroissement de la branche 5

Voici les distances (1) et. les produits scalaires entre l'opérateur V associé au groupe des 17 séries et les opérateurs $(U_{1i})_{i=1,\ldots,17}$ associé à chacune d'entre elles :

Branches	*Produits scalaires*	Distances
1 Houillères	0,26236	0,81021
2 Pétrole	0,12643	1,0821
3 Gaz	0,28413	0,76668
4 Mines	0,39457	0,54580
5 Sidérurgie	0,37500	0,58493
6 Transformation des métaux	0,47723	0,38047
7 Extract. de minéraux	0,30545	0,72404
8 Bât. et T.P.	0,25999	0,81496
9 Chimie	0,23634	0,86226
10 Verre	0,41450	0,46022
11 Caoutchouc	0,47884	0,42805
12 Produits agricoles et alimentaires	0,17696	0,98101
13 Filatures et tissages	0,45398	0,42697
14 Papier – carton	0,47135	0,39224
15 Ind. annexes des text., Bois, cuir	0,48635	0,36224
16 Transports	0,26493	0,80507
17 Commerce et administration	0,23063	0,87368

(1) Le programme nécessaire à ces calculs a été écrit en FORTRAN IV. Son encombrement est d'environ 33 K ; dans le cas de notre exemple le temps CP a été de 05,81 sec sur IBM 360/91.

Comme le laissait prévoir l'analyse en composantes principales, la branche 15 doit être prise comme série "résumé" puisque son opérateur associé est le plus proche de V.

Nous avons déjà vu que, lorsque les séries étaient réduites, la série "résumé" était celle qui maximisait

$$\sum_{k=1}^{17} \lambda_k \cdot D^2(\varphi^k , x^i)$$

Or

$$<< V , U_{1,15}>> = \frac{1}{17} \sum_{k=1}^{17} \lambda_k \cdot D^2(\varphi^k , x^{15})$$

$$\leqslant \frac{\lambda_1}{17} = \frac{9,356}{17} \quad \text{d'après} \qquad (1)$$

$$= 0,5504$$

Il est donc possible d'apprécier la qualité de la série "résumé" comparativement au cas idéal de la première composante principale en considérant le rapport

$$\frac{<<V , U_{1,15}>>}{\frac{\lambda_1}{17}} = \frac{0,48635}{0,5504} = 88\ \%$$

La qualité parait bonne, toutefois on ne doit pas perdre de vue la disparité observée entre les mois d'août, septembre et les autres.

Groupes "résumés" obtenus par la méthode "ascendante"

Effectif	Groupe "résumé"	Distance	Produit scalaire
2	15-5	0,17591	0,43068
3	15-5-17	0,08316	0,36399
4	15-5-17-6	0,06863	0,39230
5	15-5-17-6-12	0,04868	0,34923
6	15-5-17-6-12-4	0,039671	0,35679
7	15-5-17-6-12-4-9	0,03406	0,33958

V – RECHERCHE DE SERIES "PILOTES"

1 – Introduction

Nous nous proposons tout d'abord de rechercher la branche qui, à l'époque (t – i), (i $\leqslant$ 12), est la plus représentative du phénomène à l'époque t. Nous l'appellerons "série pilote à i mois" dans le cas mensuel. Puis comme pour les séries "résumée", nous rechercherons des groupes de "séries pilotes à i mois".

La méthode consiste à considérer deux groupes de séries, l'un décalé de i mois par rapport à l'autre ; puis à en remplacer un par son opérateur, ainsi que chaque série de l'autre. Il suffit alors de choisir la branche en retard de i mois dont l'opérateur est le plus proche de celui associé à l'autre groupe.

Nous considérons p branches dont les taux d'accroissement ont été observés durant n mois.

Notations :

- ${}_1x^j = ({}_1x^j(i+1), \ldots, {}_1x^j(n)) \in R^{n-i}$ représente les taux d'accroissement menseuls centrés – réduits de la branche j du mois (i + 1) au mois n.
- V est l'opérateur associé au groupe $({}_1x^j)_j = 1, \ldots, p$
- ${}_2x^j = ({}_2x^j(1), \ldots, {}_2x^j(n-i)) \in R^{n-i}$ représente les taux d'accroissement mensuels centrés réduits de la branche j en retard de i mois sur ${}_1x^j$.
- U'_{1j} est l'opérateur associé à ${}_2x^j$
- V' l'opérateur associé au groupe $({}_2x^j)_{j=1,\ldots,q}$
- $D = \dfrac{1}{n-i} I_{n-i}$

2 – Recherche de la série pilote à i mois, puis du groupe

Par définition,

$$d^2(V, U'_{1j}) = |||V|||^2 + |||U'_{1j}|||^2 - \frac{2}{p}\sum_{k=1}^{p} D^2({}_1x^k, {}_2x^j) \qquad \text{et} \qquad |||U'_{1j}||| = 1$$

puisque les séries sont réduites. La série pilote à i mois sera donc celle qui minimise $d^2(V, U'_{1j})$, ce qui correspond à maximiser

$$\sum_{k=1}^{p} D^2({}_1x^k, {}_2x^j)$$

Pour calculer ces quantités, il suffit de former la matrice $\sum$, ((p + q) (p + q)), de corrélation des (p + q) séries, soit :

$$\Sigma = \left(\begin{array}{c|c} A & C \\ \hline C & B \end{array}\right)$$

où A est la matrice de corrélation des séries $({}_1x^k)_{k=1,\ldots,p}$, B celle de $({}_2x^j)_{j=1,\ldots,q}$ et C celle d'autocorrélation d'ordre i, et d'élever tous les éléments de Σ au carré.

A partir de la série pilote obtenue, on forme des groupes de séries pilotes par la même méthodologie que pour les groupes de séries résumées.

3 – Exemple

Comme au (IV-4), nous considérons le phénomène consommation mensuelle d'électricité HT en France. Avec nos notations du (V-1), le groupe un est constitué des 17 branches du (IV-4), par contre nous ajoutons au deuxième groupe la série des taux d'accroissement mensuels de la consommation HT totale observés des mois 1 à (n – i), cette dernière série étant centrée – réduite. Nous allons chercher dans le groupe deux la série pilote à i = 1 mois, puis des groupes de séries pilotes à un mois.

Les données de base étant les mêmes qu'au (IV-4), l'étude qui va suivre est une première approche du problème. Elle devrait être suivie, en pratique, d'une nouvelle étude ne tenant pas compte des mois litigieux d'août et septembre.

Voici les distances et les produits scalaires entre les opérateurs associés aux séries du groupe deux et l'opérateur associé au groupe un.

Branches	$d^2(V, U'_{Ij})$	$\langle\langle V, U'_{Ij}\rangle\rangle$
1 Houillères	1,2852	0,02610
2 Pétrole	1,2224	0,05751
3 Gaz	1,2761	0,03066
4 Mines	1,3083	0,01453
5 Sidérurgie	1,2683	0,03455
6 Transform. des métaux	1,2145	0,06147
7 Extract. de minéraux	1,3054	0,01600
8 Bât. et T.P.	1,3021	0,01764
9 Chimie	1,2241	0,05663
10 Verre	1,2652	0,03612
11 Caoutchouc	1,1477	0,09484

Branches	$d^2(V, U'_{lj})$	$<<V, U'_{lj}>>$
12 Produit agric. et alimentaires.	1,2080	0,06468
13 Filatures et tissages	1,2422	0,04759
14 Papier – carton	1,2327	0,05234
15 Ind. annexes des text., bois, cuir	1,2566	0,04038
16 Transports	1,1495	0,09396
17 Commerce et administration	1,2445	0,04646
18 Consommation totale	1,2543	0,04157

La branche 11 est série pilote à un mois. Si des études ultérieures, en particulier dans les mois d'août et septembre, la confirmaient dans ce rôle, on pourrait l'utiliser comme un indicateur pour la consommation d'électricité HT durant le mois qui suit la dernière consommation connue de cette branche.

Groupes de séries pilotes à un mois par la méthode "Ascendante"

Effectif	Groupe – résumé	Distance	Produit scalaire
2	11-2	0,69516	0,07618
3	11-2-17	0,56554	0,066271
4	11-2-17-9	0,51771	0,06386
5	11-2-17-9-12	0,48840	0,06402
6	11-2-17-9-12-16	0,48077	0,06901

Remarque : $d^2(V, V') = 0,60281$

o

o o

BIBLIOGRAPHIE

(1) T.W. ANDERSON. – "An introduction to multivariate Analysis" Wiley.

(2) Y. ESCOUFIER – Thèse – Faculté de Montpellier, 1970.

(3) "Analyse des données multidimensionnelles". – Centre d'Etudes Economiques d'Entreprises (C3E – 116 Bd Péreire – Paris XVII°).

ANNEXE

Proposition 1 :

$$\sum_{i=1}^{p_i} \sum_{j=1}^{p_j} D^2(x^i , y^j) = \sum_{i=1}^{p_i} \sum_{j=1}^{p_j} \lambda_i \mu_j D^2(\varphi^i , \varphi^j)$$

Soit :

$$x^i = \sum_{k=1}^{p_i} D(x^i , \varphi^k) \varphi^k$$

$$= \sum_{k=1}^{p_i} \sqrt{\lambda_k}\, \alpha_k^i\, \varphi^k \qquad \text{où} \qquad \alpha_k^i \qquad \text{est la ième coordonnée}$$

du kième vecteur propre normé $\left(\sum_{k=1}^{p_i} (\alpha_k^i)^2 = 1 \right)$ de $\sum_1$ matrice de variance covariance des x^i associé à λ_k.

De même $y^j = \sum_{l=1}^{p_j} \sqrt{\mu_l}\, \beta_l^j\, \Psi^l$ ou β_l^j est la j ème coordonnée du lème vecteur propre de $\sum_2$ associé à μ_l.

$(\alpha_k)_{k=1,\ldots,p_i}$ forme un système orthonormé, de même pour $(\beta_k)_{k=1,\ldots,p_j}$

Donc

$$\sum_{i=1}^{p_i} \sum_{j=1}^{p_j} D^2(x^i , y^j) = \sum_i \sum_j \left[\sum_k \sum_l \sqrt{\lambda_k} \cdot \sqrt{\mu_l} \cdot \alpha_k^i \cdot \beta_l^j \cdot D(\varphi^k , \Psi^l) \right]^2$$

or

$$\left[\sum_k \sum_l\right]^2 = \left[\sum_k \sum_l \alpha_k^i . \beta_l^i . D(\sqrt{\lambda_k} . \varphi^k , \sqrt{\mu_l} . \Psi^l)\right]^2$$

Soit A la matrice $(p_i \times p_j)$ telle que

$$(A)_{kl} = D(\sqrt{\lambda_k} . \varphi^k , \sqrt{\mu_l} . \Psi^l) ,$$

et soient : α^i le vecteur $(p_i \times 1)$ tel que ${}^t\alpha^i = (\alpha_1^i , \ldots , \alpha_{p_i}^i)$

β^j le vecteur $(p_j \times 1)$ tel que ${}^t\beta^j = (\beta_1^j , \ldots , \beta_{p_j}^j$

Alors

$$\left[\sum_k \sum_l\right]^2 = ({}^t\alpha^i . A . \beta^j)^2 = ({}^t\alpha^i . A . \beta^j) ({}^t\beta^j . {}^tA . \alpha^i)$$

et

$$\sum_i \sum_j ({}^t\alpha^i . A . \beta^j)^2 = \sum_i {}^t\alpha^i A \left(\sum_j \beta^j . {}^t\beta^j\right) {}^tA . \alpha^i$$

or

$\sum_j \beta^j . {}^t\beta^j = I$ puisque les $(\beta_k)_{k=1,\ldots,p_j}$ sont orthonormées.

Par suite

$$\sum_i \sum_j D^2(x^i , y^j) = \sum_i {}^t\alpha^i . A . {}^tA . \alpha^i , \qquad (1)$$

Soit $B = A . {}^tA$, B est une matrice symétrique (p_i , p_i) .

$${}^t\alpha^i . B . \alpha^i = \sum_{k=1}^{p_i} (\alpha_k^i)^2 (B)_{kk} + 2 \sum_{k<l} \alpha_k^i . \alpha_l^i . (B)_{kl}$$

(1) $$= \sum_i \sum_k (\alpha_k^i)^2 (B)_{kk} + 2 \sum_i \sum_{k<l} \alpha_k^i . \alpha_l^i . (B)_{kl}$$

$$= \sum_k \left(\sum_i (\alpha_k^i)^2\right) (B)_{kk} + 2 \sum_{k<l} \left(\sum_i \alpha_k^i . \alpha_l^i\right) (B)_{kl}$$

$= \sum_k (B)_{kk}$ puisque les $(\alpha_k)_{k=1,\ldots,p_i}$ sont orthonormés.

$= Tr(B) = Tr(A . {}^tA)$

or $(B)_{kk} = \sum_l (A)_{kl} . (A)_{kl}$, soit $Tr(B) = \sum_{k,l} ((A)_{kl})^2$ donc

(1) $$= \sum_k \sum_l \lambda_k . \mu_l . D^2(\varphi^k , \Psi^l)$$ C.Q.F.D.