LES CAHIERS DE L'ANALYSE DES DONNÉES

F. NAKHLÉ

Sur l'analyse d'un tableau de notes dédoublées

Les cahiers de l'analyse des données, tome 1, n° 3 (1976), p. 243-257

<http://www.numdam.org/item?id=CAD_1976__1_3_243_0>

© Les cahiers de l'analyse des données, Dunod, 1976, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Les Cahiers de l'Analyse des Données
Vol. I - 1976 - n° 3 - p. 243-257

SUR L'ANALYSE D'UN TABLEAU DE NOTES DÉDOUBLÉES

[DEDOU I]

Exposé des résultats de la thèse de F. Nakhlé (1)

Le travail de F. Nakhlé peut être divisé en quatre parties : § 1 : analyse de la correspondance obtenue en dédoublant un tableau de notes positives; § 2 changements d'échelles dans un tableau de notes, laissant invariants les facteurs; § 3 application au tableau des notes obtenues aux diverses épreuves par les candidats au concours d'admission à l'Ecole Polytechnique en 1970-71-72. Aux §§ 1 & 2, les résultats des calculs algébriques sont présentés sous forme de problèmes accompagnés de corrigé. Le § 3 donne les faits les plus saillants révélés par l'analyse du tableau de notes : les lecteurs curieux de ces faits, pourront les découvrir sans faire des §§ 1 & 2 une étude approfondie. Enfin on trouvera au § 4 un nouveau programme d'analyse des correspondances propre aux tableaux dédoublés : ce programme fondé sur les formules du § 1, vise à économiser la place en mémoire, et à réduire le temps de calcul donc le coût.

On expose dans ce cahier les résultats des §§ 1 et 2 de la thèse, en réservant de publier ultérieurement la suite du travail de F. Nakhlé.

1. *L'analyse d'un tableau de notes positives dédoublé* :

L'objet du problème qui suit est de réduire l'analyse d'un tableau de notes dédoublé à 2p colonnes à la diagonalisation d'une matrice p x p.

1.1. Enoncé du problème : on note

I : un ensemble d'individus (ou sujets) i; Card I = n;

Q : un ensemble de questions (ou d'épreuves); Card Q = p.

$a(q)$: la note maxima qu'un sujet peut obtenir à l'épreuve q; $a(q) > 0$.

$A = \Sigma\{a(q) \mid q \in Q\}$: le maximum que puisse atteindre le total des notes obtenues par un sujet à l'ensemble des questions.

$(k_0)_{IxQ} = \{k_0(i,q) \mid i \in I,\ q \in Q\}$; le tableau des notes obtenues par tous les sujets à toutes les questions; $k_0(i,q)$ est la note du sujet i à la question q et on a :

$\forall\, i \in I, \forall\, q \in Q : 0 \leq k_0(i,q) \leq a(q)$.

(1) *Docteur de 3ème cycle en statistique. Université Pierre et Marie Curie.*

$J = Q^+ \cup Q^-$: l'ensemble Q dédoublé : on a posé comme d'usage

$Q^+ = \{q^+ | q \in Q\}$; $Q^- = \{q^- | q \in Q\}$

$k_{IxJ} = \{k(i,j) | i \in I, j \in J\}$: le tableau de notes dédoublé; à chaque question q correspond une colonne q^+ contenant les notes initiales et une colonne complémentaire q^-: ce qu'on écrit :

$\forall i \in I, \forall q \in Q$:

$k(i,q^+) = k_0(i,q)$; $k(i,q^-) = a(q) - k_0(i,q^+)$

Comme à l'ordinaire on note encore : $k_0(q) = \Sigma\{k_0(i,q) | i \in I\}$; $k(i) = \Sigma\{k(i,j) | j \in J\}$; $k(j) = \Sigma\{k(i,j) | i \in I\}$; $k = \Sigma\{k(i,j) | i \in I, j \in J\}$ $f_i = k(i)/k$; $f_j = k(j)/k$ etc.

1.1.1° Exprimer en fonction des $k_0(q)$, $a(q)$, A, n les quantités $k(i)$, $k(q^+)$, $k(q^-)$ et k; ainsi que les f_i, f_{q^+}, f_{q^-}.

1.1.2° Montrer que les profils de deux colonnes complémentaires q^+ et q^- du tableau k_{IJ} sont liés au profil moyen par une relation de la forme :

$$f_{q^+} f_I^{q^+} + f_{q^-} f_I^{q^-} = r(q) f_I ,$$

où $r(q)$ est un nombre positif qu'on calculera. Que peut-on dire des contributions relatives apportées par un même facteur G_α issu du tableau k_{IJ} à deux variables complémentaires q^+ et q^- ? Quelle relation y-a-t-il entre $G_\alpha(q^+)$ et $G_\alpha(q^-)$?

1.1.3° Montrer que la distance distributionnelle entre deux lignes du tableau dédoublé k_{IJ} s'exprime en fonction des seules notes initiales $k_0(i,q)$ par une formule :

$$\| f_J^i - f_J^{i'} \|^2_{fJ} = \Sigma\{m(q)(k_0(i,q) - k_0(i',q))^2 | q \in Q\},$$

où les $m(q)$ sont des coefficients que l'on calculera en fonction des $k_0(q)$, $a(q)$, A, n.

1.1.4° On note désormais :

$$k_0'(i,q) = m(q)^{1/2} k_0(i,q)$$

$$k_0'^{iQ} = k_0'(i,Q) = \{k_0'(i,q) | q \in Q\} \in R^Q.$$

On considère d'une part, dans R_J muni de la métrique du χ^2 de centre f_J le nuage :

$$N(I) = \{(f_J^i, f_i) | i \in I\} \subset R_J;$$

(nuage usuel des profils des lignes); d'autre part, dans R^Q muni de la norme euclidienne $\|x^Q\|^2 = \Sigma\{(x^q)^2 | q \in Q\}$, le nuage :

$$N_0'(I) = \{(k_0'^{iQ}, 1/n) | i \in I\} \subset R^Q;$$

(i.e. les points $k_0'^{iQ}$ sont tous affectés de la même masse 1/n).

Comparer les moments principaux d'inertie des nuages N(I) et $N_0'(I)$.

Comparer aux facteurs $F_\alpha(i)$ issus de l'analyse du tableau de correspondance k_{IJ}, les coordonnées des points du nuage $N_0'(I)$ sur les axes principaux d'inertie de ce nuage.

1.1.5° Exprimer en fonction de $k_0(q)$, $a(q)$, A, n, les coordonnées $G(q)$ dans R^Q du centre de gravité G du nuage $N_0'(I)$. Exprimer $G(q)$ en fonction de $k_0(q)$, $m(q)$, n. Exprimer comme une somme dont les termes sont fonction des $m(q)$, $k_0(q)$, $k_0(i,q)$, n, les produits d'inertie du nuage $N_0'(I)$ rapporté à son centre de gravité :

$$\sigma(q,q') = \Sigma\{(1/n)(k_0'(i,q)-G(q))(k_0'(i,q')-G(q'))\,|\,i \in I\}$$

1.1.6° On considère dans R^Q les vecteurs axiaux unitaires $\psi_\alpha^Q = \{\psi_\alpha(q)\,|\,q \in Q\}$ des axes principaux d'inertie définis par le système :

$$\forall\, q \in Q : \Sigma\{\sigma(q,q')\psi_\alpha(q')\,|\,q' \in Q\} = \lambda_\alpha\, \psi_\alpha(q),$$

avec la condition : $\Sigma\{(\psi_\alpha(q))^2\,|\,q \in Q\} = 1$.

Exprimer en fonction des $\psi_\alpha(q)$, $G(q)$, $k_0'(i,q)$ la valeur du facteur $F_\alpha(i)$ pour l'individu i.

1.1.7° Exprimer en fonction des n, λ_α, $\psi_\alpha(q)$, la valeur de la somme $S(q)$:

$$S(q) = \Sigma\{k_0'(i,q)F_\alpha(i)\,|\,i \in I\}$$

1.1.8° En faisant usage de la formule de transition, exprimer les valeurs des facteurs $G_\alpha(q^+)$, $G_\alpha(q^-)$ (issus de l'analyse du tableau dédoublé k_{IJ}) en fonction des $F_\alpha(i)$, puis des $\psi_\alpha(q)$. On aboutira (compte tenu des résultats de 7°) à une formule simple exprimant $G_\alpha(q^+)$ et $G_\alpha(q^-)$ en fonction des seuls λ_α, A, n, $k_0(q)$, $a(q)$, $\psi_\alpha(q)$.

1.1.9° Déduire des résultats précédents un schéma de calcul réduisant l'analyse factorielle du tableau de correspondance k_{IJ} (qui a 2p colonnes) à la diagonalisation d'une matrice carrée symétrique p x p.

1.2. Solution du problème :

1.2.1° Donnons sans démonstration les formules suivantes :

$k(i) = A$; $k = nA$; $f_i = 1/n$;

$k(q^+) = k_0(q)$; $k(q^-) = n\,a(q) - k_0(q)$;

$f_{q^+} = k_0(q)/(nA)$; $f_{q^-} = (n\,a(q) - k_0(q))/(nA)$.

1.2.2° Rappelons qu'on a par définition :

$f_j \times f_i^j = f_{ij} = k(i,j)/k$; d'où :

$$(f_{q^+} f_i^{q^+}) + (f_{q^-} f_i^{q^-}) = (k(i,q^+) + k(i,q^-))/k$$
$$= a(q)/(nA) = (a(q)/A)(1/n);$$

$f_{q^+}\ f_I^{q^+} + f_{q^-}\ f_I^{q^-} = (a(q)/A)\ f_I$.

Le nombre r(q) mentionné dans l'énoncé est donc r(q) = a(q)/A. Les profils des colonnes q^+ et q^- sont alignés avec le profil moyen f_I (conformément au principe du bras de levier). La contribution relative apportée par le facteur G_α à q^+ ou à q^- n'est autre que le carré du cosinus de l'angle formé par l'axe α et la droite joignant les trois profils $f_I^{q^+}$, f_I, $f_I^{q^-}$; et les facteurs sont liés par la relation linéaire :

$f_{q^+}\ G_\alpha(q^+) + f_{q^-}\ G_\alpha(q^-) = 0$.

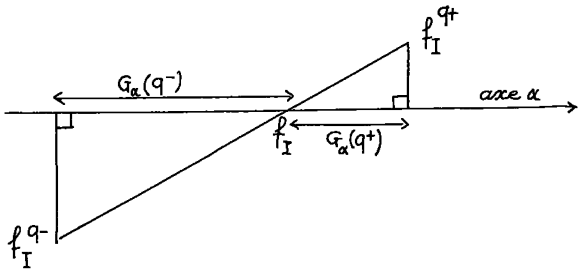

Figure 1 : Principe du bras de levier; on a supposé $f_{q+} = 2\ f_{q-}$

1.2.3° Appliquons à la correspondance k_{IJ} les formules usuelles, en tenant compte des résultats de 1.2.1° :

$$\|f_J^i - f_J^{i'}\|^2_{fJ} = \Sigma\{f_j^i - f_j^{i'})^2/f_j \mid j \in J\}$$
$$= \Sigma\{(f_{q^+}^i - f_{q^+}^{i'})^2/f_{q^+} \mid q \in Q\}$$
$$+ \Sigma\{(f_{q^-}^i - f_{q^-}^{i'})^2/f_{q^-} \mid q \in Q\};$$

$f_j^i = k(i,j)/k(i) = k(i,j)/A$;

$f_{q^+}^i - f_{q^+}^{i'} = -(f_{q^-}^i - f_{q^-}^{i'}) = (k_0(i,q) - k_0(i',q))/A$;

$$\|f_J^i - f_J^{i'}\|^2_{fJ} = \Sigma\{(k_0(i,q) - k_0(i',q))^2 A^{-2} (f_{q^+}^{-1} + f_{q^-}^{-1}) \mid q \in Q\}$$
$$= \Sigma\{m(q)(k_0(i,q) - k_0(i',q))^2 \mid q \in Q\};\ \text{où :}$$

$m(q) = A^{-2} \times ((1/f_{q^+}) + (1/f_{q^-}))$
$= A^{-2} \times (f_{q^+} + f_{q^-})/(f_{q^+}\ f_{q^-})$
$= n^2\ a(q)/(A \times k_0(q) \times (na(q) - k_0(q)))$.

1.2.4° Les deux nuages N(I) et $N_0^i(I)$ ont même triple : c'est le même ensemble de points I, avec les mêmes masses $f_i = (1/n)$, et les mêmes distances : $d^2(i,i') = \|f_J^i - f_J^{i'}\|^2$. Donc les moments principaux d'inertie sont les mêmes; et les facteurs $F_\alpha(i)$ issus de l'analyse du tableau de correspondance k_{IJ} ne sont autres que les coordonnées des points du

nuage $N_0^1(I)$ sur les axes principaux d'inertie de ce nuage.

1.2.5° Calculons les coordonnées du centre de gravité du nuage $N_0^1(I)$:

$$G(q) = \Sigma\{f_i\ k_0^1(i,q) \mid i \in I\}$$

$$= \Sigma\{(1/n)\ m(q)^{1/2}\ k_0(i,q) \mid i \in I\} = m(q)^{1/2}\ k_0(q)/n$$

$$= (a(q)\ k_0(q)/(A(na(q) - k_0(q))))^{1/2}.$$

Dans le calcul des produits d'inertie du nuage $N_0^1(I)$ rapporté à son centre de gravité, on se souvient du théorème de Huyghens; il vient :

$$\sigma(q,q') = \Sigma\{(1/n)(k_0^1(i,q) - G(q))(k_0^1(i,q') - G(q')) \mid i \in I\}$$

$$= \Sigma\{(1/n)\ k_0^1(i,q)\ k_0^1(i,q') \mid i \in I\} - G(q)\ G(q')$$

$$= (m(q)\ m(q'))^{1/2}\ (\Sigma_i\{k_0(i,q)\ k_0(i,q')/n\} - k_0(q)\ k_0(q')n^{-2}).$$

1.2.6° Le facteur $F_\alpha(i)$ n'est autre que la coordonnée du point i du nuage $N_0^1(I)$ sur l'axe principal d'inertie de rang α (cf 1.2.4°) :

$$F_\alpha(i) = < \overrightarrow{GI} \cdot \psi_\alpha^Q > = \Sigma\{\psi_\alpha(q)(k_0^1(i,q) - G(q)) \mid q \in Q\},$$

où $< x^Q, y^Q > = \Sigma\{x^q\ y^q \mid q \in Q\}$ désigne le produit scalaire usuel de l'espace R^Q, espace ambiant du nuage $N_0^1(I)$.

1.2.7° Parce que sur I, $F_\alpha(i)$ est de moyenne nulle on a :

$$S(q) = \Sigma\{k_0^1(i,q)F_\alpha(i) \mid i \in I\} = \Sigma\{(k_0^1(i,q) - G(q))F_\alpha(i) \mid i \in I\}.$$

En remplaçant $F_\alpha(i)$ par sa valeur calculée en 1.2.6° il vient :

$$S(q) = \Sigma\{(k_0^1(i,q) - G(q))(k_0^1(i,q') - G(q'))\psi_\alpha(q') \mid i \in I,\ q' \in Q\},$$

somme double où l'on reconnait l'expression de $n\sigma(q,q')$:

$$S(q) = \Sigma\{n\sigma(q, q')\psi_\alpha(q') \mid q' \in Q\};$$

d'où si l'on se souvient de l'équation des axes d'inertie rappelée dans l'énoncé en 1.1.6° :

$$S(q) = n\lambda_\alpha\ \psi_\alpha(q).$$

On peut encore pour calculer S(q) rappeler l'expression du vecteur $\overrightarrow{GI}$ en somme de composantes portées par les axes principaux d'inertie.

$\overrightarrow{GI} = \Sigma_\beta\ F_\beta(i)\ \psi_\beta^Q$ (somme étendue à tous les axes); d'où :

$$k_0^1(i,q) = G(q) + \Sigma_\beta\ F_\beta(i)\ \psi_\beta(q);$$

$$S(q) = \Sigma\{F_\alpha(i)(G(q) + \Sigma_\beta\ F_\beta(i)\ \psi_\beta(q)) \mid i \in I\};$$

de cette somme les termes en $\beta \neq \alpha$ disparaissent parce que les facteurs sont deux à deux non corrélés; le terme en G(q) est nul parce que F_α est de moyenne nulle, il reste :

$$S(q) = \Sigma\{(F_\alpha(i))^2\ \psi_\alpha(q) \mid i \in I\} = n\lambda_\alpha\ \psi_\alpha(q);$$

car la variance de F_α est $\lambda_\alpha = (1/n)\ \Sigma\{(F_\alpha(i))^2 \mid i \in I\}$.

1.2.8° On applique la formule de transition pour calculer $G_\alpha(q^+)$ à partir des $F_\alpha(i)$:

$$G_\alpha(q^+) = \lambda_\alpha^{-1/2} \Sigma\{F_\alpha(i)\ k(i,q^+)/k(q^+) \mid i \in I\}$$
$$= \lambda_\alpha^{-1/2} \Sigma\{F_\alpha(i) k_0(i,q)/k_0(q) \mid i \in I\}$$
$$= \lambda_\alpha^{-1/2} \Sigma\{F_\alpha(i) k_0'(i,q)\ m(q)^{-1/2}/k_0(q) \mid i \in I\}$$
$$= \lambda_\alpha^{-1/2}\ m(q)^{-1/2}\ k_0(q)^{-1} \Sigma\{F_\alpha(i) k_0'(i,q) \mid i \in I\};$$

d'où il vient, en se souvenant de la valeur de m(q) calculée en 1.2.3°, et de la somme faite en 1.2.7° :

$$G_\alpha(q^+) = n\lambda_\alpha^{1/2}\ m(q)^{-1/2}\ k_0(q)^{-1}\ \psi_\alpha(q)$$
$$= (\lambda_\alpha A(na(q) - k_0(q))/(a(q)\ k_0(q)))^{1/2}\ \psi_\alpha(q);$$

$G_\alpha(q^-)$ se calcule à partir de $G_\alpha(q^+)$ en utilisant le principe du bras de levier (cf 1.2.2°) :

$$G_\alpha(q^-) = -\ (f_{q+}/f_{q-})\ G_\alpha(q^+)$$
$$= -\ (\lambda_\alpha A k_0(q)/(a(q)(na(q) - k_0(q))))^{1/2}\ \psi_\alpha(q).$$

1.2.9° On pourra suivre le schéma de calcul suivant :

a) Calcul de la matrice symétrique p x p des $\sigma(q,q')$ (cf 1.2.5°).

b) Diagonalisation de cette matrice; d'où les ψ_α^Q (cf 1.1.6°).

c) Calcul des $G_\alpha(j)$ c'est-à-dire des $G_\alpha(q^+)$ et $G_\alpha(q^-)$ (cf 1.2.8°).

d) Calcul des $F_\alpha(i)$: on pourra partir des $G_\alpha(j)$ et appliquer la formule de transition; il vient :

$$F_\alpha(i) = A^{-1}\ \lambda_\alpha^{-1/2}\ \Sigma\{k_0(i,j)\ G_\alpha(q^+) + (a(j) - k_0(i,j))\ G_\alpha(q^-) \mid q \in Q\}.$$

2. Transformations d'un tableau laissant invariant les facteurs :

On sait que l'influence d'une épreuve dans un concours n'est pas rigoureusement fixée par son coefficient : telle épreuve notée de 0 à 20, pourra compter plus qu'une autre notée de 0 à 40 si le correcteur de celle-ci ne sort qu'exceptionnellement de l'intervalle 18-30, tandis que son collègue chargé de celle-là n'hésite pas à attribuer 1 ou 19. Dans le total général qui détermine l'issue du concours, l'influence d'une épreuve est mieux exprimée par la variance de sa note que par son coefficient. En analyse des correspondances, un tableau de notes positives est traité après dédoublement; l'influence d'une épreuve q se mesure par la somme des contributions des points q^+ et q^- à l'inertie totale du nuage; on peut, de plus, chercher quelles transformations linéaires affectant une ou plusieurs notes laissent invariants les facteurs $F_\alpha(i)$ (quant à l'invariance du total général, elle ne pose qu'un problème trivial : la seule transformation permise étant alors d'ajouter, ou de retrancher, une constante aux notes de l'une des épreuves ou de plusieurs d'entre elles). Dans ce §, nous utilisons les résultats des questions 1°, 2°, 3° du § 1 pour étudier les transformations d'un tableau de notes dédoublé; nous avons ailleurs considéré les transformations des colonnes d'un tableau de forme plus générale, laissant invariants les facteurs sur l'ensemble I (*). Au § 3.4.1, les princi-

(*) *Les cahiers de l'Analyse des Données - Vol. I - 1976 n° 2 pp.127-136.*

paux résultats démontrés au § 2 seront rappelés pour être appliqués aux notes du concours de l'Ecole Polytechnique.

2.1. *Problème sur les transformations d'un tableau de notes dédoublé* :

2.1.0. *Notations adoptées* :

Les transformations que nous considérons affectent simultanément deux colonnes complémentaires q^+ et q^- qu'elles transforment en des colonnes nouvelles qn^+, qn^- suivant les formules :

$$k(i,qn^+) = \rho_q\ (k(i,q^+) + \xi_q\ a_q);$$

$$k(i,qn^-) = \rho_q\ (k(i,q^-) + \eta_q\ a_q).$$

Il revient au même de dire qu'à la note initiale $k_0(i,q) = k(i,q^+)$ on ajoute la constante ξ_q; qu'on considère l'intervalle de variation de cette note comme étant $(0,(1 + \xi_q + \eta_q)a_q)$ bien que le domaine de variation soit $(\xi_q\ a_q\ ,\ (1 + \xi_q)a_q)$ (ici on note a_q ce qui est noté $a(q)$ au § 1); qu'on effectue enfin un changement d'échelle de coefficient ρ_q. Ce qui s'écrit en formules :

$$k_0(i,qn) = \rho_q(k_0(i,q) + \xi_q a_q);$$

$$a(qn) = \rho_q(1 + \xi_q + \eta_q)\ a_q;$$

et est représenté sur la figure 2. Dans le cas de la figure, ρ_q , ξ_q , η_q sont positifs : mais il n'en est pas toujours ainsi : la condition imposée est seulement que le tableau initial ne fournisse que des notes

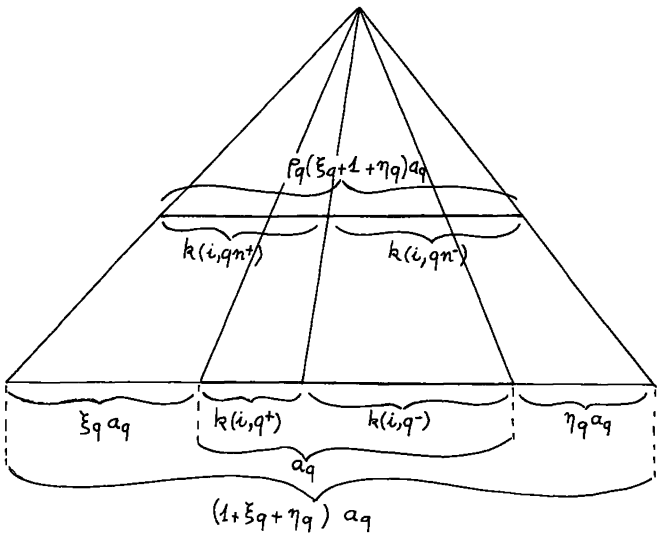

Figure 2: Changement simultané d'origines et d'échelles sur les notes d'un tableau dédoublé.

$k(i,qn^+)$ et $k(i,qn^-)$ positives. Cette restriction elle-même n'est pas absolument impérative : on sait que le programme d'analyse des correspondances tolère des k(i,j) négatifs, à condition que les totaux des colonnes et des lignes soient tous positifs; donc ici que :

$$0 < k(qn^+) = \rho_q(k_0(q) + n\,\xi_q a_q);$$

$$0 < k(qn^-) = \rho_q(n(1 + \eta_q)\ a_q - k_0(q));$$

$$0 < a_{qn} = \rho_q(1 + \xi_q + \eta_q)\ a_q\ ;$$

formules où le nombre des sujets de I est noté n, et où comme avec a(q) on a écrit a_{qn} pour a(qn).

Etant en analyse des correspondances où il importe peu de multiplier le tableau à analyser (tableau dédoublé) par une constante, on supposera, dans tout le problème sans restreindre la généralité, que A = 1 et que les transformations effectuées sur l'ensemble des colonnes respectent cette somme :

$$A = \Sigma\{a_q \mid q \in Q\} = \Sigma\{a_{qn} \mid q \in Q\} = A_n = 1$$

dans le langage des examens, les a_q sont les coefficients des épreuves, la condition $A = A_n = 1$ fixe cette somme à 1; ce qu'on obtiendra en substituant aux coefficients initiaux a_q les coefficients relatifs a_q/A.

Enfin dans un but de simplification on introduit les x_q , y_q , rapport des moyennes des notes du tableau dédoublé à leur amplitude possible de variation a_q :

$$x_q = k(q^+)/(n\,a_q);\quad y_q = k(q^-)/(n\,a_q) = 1 - x_q\ ;$$

et de même pour les nouvelles notes :

$$x_{qn} = k(qn^+)/(n\,a_{qn});\quad y_{qn} = k(qn^-)/(n\,a_{qn}).$$

Ces notations étant fixées, voici les questions proposées.

2.1.1° Exprimer x_{qn}, y_{qn} en fonction des x_q , ρ_q , ξ_q , η_q.

2.1.2° La distance distributionnelle entre deux individus i et i' est donnée par la formule :

$$d^2(i,i') = \Sigma\{m(q)(k_0(i,q) - k_0(i',q))^2 \mid q \in Q\}\ :$$

exprimer les m(q) en fonction de x_q , a_q (on utilisera l'expression de m(q) trouvée au § 1.2.3° et se souviendra qu'on a posé $A = A_n = 1$).

2.1.3° Après transformation des notes on a les nouvelles distances :

$$dn^2(i,i') = \Sigma\{m(qn)(k_0(i,qn) - k_0(i',qn))^2 \mid q \in Q\}\ :$$

exprimer les m(qn) en fonction des a_q , x_q , ρ_q , ξ_q , η_q.

2.1.4° A quelle condition C doivent satisfaire les a_q , x_q , ρ_q , ξ_q , η_q pour que après transformation le tableau de notes dédoublé fournisse les mêmes facteurs $F_\alpha(i)$ qu'avant transformation (on supposera ici et dans la suite que $A = A_n$).

2.1.5° Sous la condition C, exprimer $G_\alpha(qn^+)$ en fonction de

$G_\alpha(q^+)$, x_q , ξ_q; de même exprimer $G_\alpha(qn^-)$ en fonction de $G_\alpha(q^-)$, y_q , η_q. Vérifier que le principe du levier pour qn^+ et qn^- résulte des expressions trouvées et du principe de levier pour q^+ et q^-. Comparer la somme des contributions absolues des points qn^+ et qn^- au facteur α, à la somme des contributions absolues des points q^+ et q^- à ce même facteur.

N.B. Les questions 6°, 7°, 8°, 9° sont indépendantes de la question 5°.

2.1.6° a) Convenons de dire que le profil d'une note q (plus exactement du couple complémentaire q^+, q^-) est inchangé si $\xi_q = \eta_q = 0$; que vaut alors ρ_q , sous la condition C ?

b) Supposons qu'une seule note ω ($\omega \in Q$) subisse une modification de profil : que vaut alors $a_{\omega n}$, sous la condition C ?

2.1.7° Soient (ρ_q , ξ_q , η_q) des coefficients définissant une transformation de deux colonnes complémentaires q^+, q^-. Trouver (ρ'_q , ξ'_q , η'_q) définissant la même transformation que (ρ_q , ξ_q , η_q) suivie d'une permutation des deux colonnes, i.e. telle que :

$$k(i,qn'^+) = \rho'_q(k(i,q^+) + \xi'_q a_q) = k(i,qn^-) = \rho_q(k(i,q^-) + \eta_q\, a_q)$$

$$k(i,qn'^-) = \rho'_q(k(i,q^-) + \eta'_q a_q) = k(i,qn^+) = \rho_q(k(i,q^+) + \xi_q\, a_q)$$

En déduire que l'on peut se restreindre à l'étude des transformations telles que $\rho_q > 0$.

2.1.8° On suppose vérifiée la condition C; exprimer ρ_q en fonction des x_q , ξ_q , η_q. Qu'en résulte-t-il pour ρ_q si ξ_q et η_q sont positifs ? On suppose ρ_q positif, et ξ_q , η_q négatifs : que peut-on dire de $(\rho_q - 1)$? Que peut-on dire des signes de ξ_q et η_q si ρ_q est positif et $a_q = a_{qn}$?

2.1.9° Est-il toujours possible de modifier l'ensemble d'un tableau de notes dédoublé, sans y introduire de notes négatives, et en respectant la condition C (et la normalisation $A = A_n = 1$) ?

2.2. *Solution du problème* :

2.2.1° On a :

$$k(qn^+) = \rho_q(k(q^+) + n\,\xi_q a_q) = \rho_q\, n(x_q + \xi_q)\, a_q;$$

$$a_{qn} = \rho_q(1 + \xi_q + \eta_q)a_q$$

$$x_{qn} = k(qn^+)/(n\,a_{qn}) = (x_q + \xi_q)/(1 + \xi_q + \eta_q)$$

$$y_{qn} = (y_q + \eta_q)/(1 + \xi_q + \eta_q) = (1 - x_q + \eta_q)/(1 + \xi_q + \eta_q)$$

2.2.2° Rappelant la formule du § 1.2.3°, et la condition A = 1, il vient :

$$m(q) = n^2 a_q/(A \times k_0(q) \times (n\,a(q) - k_0(q))$$

$$= n^2 a_q/(A \times k(q^+) \times k(q^-))$$

$$= n^2 a_q/(A \times n\, x_q\, a_q \times n\, y_q\, a_q)$$

$= 1/(x_q\ y_q\ a_q) = 1/(x_q(1 - x_q)a_q)$

2.2.3° On a de même :

$$m(qn) = 1/(x_{qn}\ y_{qn}\ a_{qn})$$
$$= (1 + \xi_q + \eta_q)/(\rho_q(x_q + \xi_q)(y_q + \eta_q)a_q)$$
$$= (1 + \xi_q + \eta_q)/(\rho_q(x_q + \xi_q)(1 - x_q + \eta_q)a_q)$$

2.2.4° Il suffit que la métrique du nuage I soit invariante.

$$\Sigma\{m(q)(k_0(i,q) - k_0(i',q))^2 | q \in Q\} =$$
$$\Sigma\{m(qn)(k_0(i,qn) - k_0(i',qn))^2 | q \in Q\};$$

or on a : $(k_0(i,qn) - k_0(i',qn)) = \rho_q(k_0(i,q) - k_0(i,q'))$, d'où pour la condition C :

$$\forall\ q \in Q : m(q) = (\rho_q)^2 m(qn);$$

ce qu'on récrira d'après 2.2.3° :

$$\forall\ q \in Q : (1 + \xi_q + \eta_q)\ x_q\ y_q\ \rho_q = (x_q + \xi_q)(y_q + \eta_q)$$
$$(1 + \xi_q + \eta_q)\ x_q(1 - x_q)\rho_q = (x_q + \xi_q)(1 - x_q + \eta_q).$$

2.2.5° La colonne $qn^+ = \{k(i,qn^+) | i \in I\}$ est une combinaison linéaire de la colonne $q^+ = \{k(i,q^+) | i \in I\}$ et de la colonne constante; d'où entre profils la relation linéaire :

$$f_I^{qn+} = (x_q f_I^{q+} + \xi_q\ f_I)/(x_q + \xi_q);$$

le profil f_I étant au centre de gravité du nuage, à l'origine des axes factoriels, on a pour les facteurs :

$$G_\alpha(qn^+) = x_q\ G_\alpha(q^+)/(x_q + \xi_q).$$

Le cas des colonnes qn^-, q^- est tout analogue à celui de qn^+, q^+; on a :

$$f_I^{qn-} = (y_q\ f_I^{q-} + \eta_q\ f_I)/(y_q + \eta_q)$$
$$G_\alpha(qn^-) = y_q\ G_\alpha(q^-)/(y_q + \eta_q).$$

Le principe du bras de levier s'énonce pour $\{q^+, q^-\}$ et $\{qn^+, qn^-\}$ respectivement :

$$x_q\ G_\alpha(q^+) + y_q\ G_\alpha(q^-) = 0;$$

$$\rho_q(x_q + \xi_q)\ G_\alpha(qn^+) + \rho_q(y_q + \eta_q)\ G_\alpha(qn^-) = 0;$$

or il résulte des formules trouvées pour les facteurs que l'on a .

$$x_q\ G_\alpha(q^+) + y_q\ G_\alpha(q^-) = (x_q + \xi_q)\ G_\alpha(qn^+) + (y_q + \eta_q)\ G_\alpha(qn^-);$$

d'où l'équivalence demandée.

Les diverses contributions absolues considérées sont données par les formules :

$$\text{contrabs}(q^+;\ \alpha) = f_{q+}\ G_\alpha(q^+)^2 = (x_q\ a_q/A)\ G_\alpha(q^+)^2$$

$$\text{contrabs}(q^-;\ \alpha) = (y_q\ a_q/a)\ G_\alpha(q^-)^2$$

$$\text{contrabs}(qn^+;\alpha) = f_{qn+}\ G_\alpha(qn^+)^2 = (\rho_q(x_q + \xi_q)a_q/A)\ G_\alpha(qn^+)^2$$

$$= (\rho_q\ x_q^2\ a_q/(A(x_q + \xi_q)))\ G_\alpha(q^+)^2$$

$$\text{contrabs}(qn^-;\alpha) = (\rho_q\ y_q^2\ a_q/(A(y_q + \eta_q)))\ G_\alpha(q^-)^2$$

En tenant compte du principe du bras de levier qui lie $G_\alpha(q^-)$ à $G_\alpha(q^+)$ il vient :

$$\text{contrabs}(q^+;\ \alpha) + \text{contrabs}(q^-;\ \alpha) =$$

$$(a_q/A)((1/x_q) + (1/y_q))\ x_q^2\ G_\alpha(q^+)^2;$$

$$\text{contrabs}(qn^+;\alpha) + \text{contrabs}(qn^-;\ \alpha) =$$

$$(\rho_q\ a_q/A)((x_q + \xi_q)^{-1} + (y_q + \eta_q)^{-1})x_q^2\ G_\alpha(q^+)^2.$$

L'égalité de ces deux sommes de contribution s'écrit :

$$\rho_q((x_q + \xi_q)^{-1} + (y_q + \eta_q)^{-1}) = x_q^{-1} + y_q^{-1};$$

compte tenu de ce que $x_q + y_q = 1$, cette relation est identique à la condition C trouvée en 4°.

Remarque : des calculs tout analogues à ceux qui précèdent permettent de trouver la condition C sans utiliser les résultats du § 1 (formules des m(q) et m(qn)). Voici comment. On a dans R_I d'une part le nuage N(J) des f_I^{q+}, f_I^{q-}; d'autre part le nuage N(Jn) des f_I^{qn+}, f_I^{qn-}. Il faut exprimer que ces deux nuages ont même matrice d'inertie. Soient par exemple U,V deux formes linéaires sur R_I s'annulant au centre f_I; e.g.:

$$U(p_I) = \Sigma_i\ U^i(p_i - f_i);\ V(p_I) = \Sigma_i\ V^i(p_i - f_i).$$

On veut écrire une relation entre les x_q , y_q , ξ_q , η_q , ρ_q , a_q , impliquant l'égalité :

$$f_{q+}\ U^{q+}\ V^{q+} + f_{q-}\ U^{q-}\ V^{q-} =$$

$$f_{qn+}\ U^{qn+}\ V^{qn+} + f_{qn-}\ U^{qn-}\ V^{qn-},$$

(formule où on a noté U^j, V^j pour $U(f_I^j)$, $V(f_I^j)$: e.g. U^{q-} pour $U(f_I^{q-})$; d'où résultera l'égalité des matrices d'inertie des deux nuages N(J) et N(Jn). Des relations entre profils utilisées ci-dessus :

$$x_q\ f_I^{q+} + y_q\ f_I^{q-} = f_I;$$

$$(x_q + \xi_q)\ f_I^{qn+} + (y_q + \eta_q)\ f_I^{qn-} = (1 + \xi_q + \eta_q)f_I;$$

$$f_I^{qn+} = x_q\ f_I^{q+}/(x_q + \xi_q) + \xi_q\ f_I/(x_q + \xi_q)$$

$$f_I^{qn-} = y_q\ f_I^{q-}/(y_q + \eta_q) + \eta_q\ f_I/(y_q + \eta_q)$$

il résulte que :

$$x_q\ U^{q+} + y_q\ U^{q-} = (x_q + \xi_q)\ U^{qn+} + (y_q + \eta_q)\ U^{qn-} = 0$$

$$U^{qn+} = U^{q+}\ x_q/(x_q + \xi_q)\ ;$$

$$U^{qn-} = U^{q-}\, y_q/(y_q + \eta_q);$$

et de même pour les V; d'où pour les produits U V :

$$f_{q+}\, U^{q+}\, V^{q+} + f_{q-}\, U^{q-}\, V^{q-} = (a_q/A)(x_q^{-1} + y_q^{-1})\, x_q^2\, U^{q+}\, V^{q+};$$

$$f_{qn+}\, U^{qn+} V^{qn+} + f_{qn-}\, U^{qn-}\, V^{qn-} =$$

$$(\rho_q a_q/A)((x_q + \xi_q)^{-1} + (y_q + \eta_q)^{-1});$$

d'où pour l'égalité de ces deux sommes la condition C trouvée au § 2.2.4° et déjà retrouvée ci-dessus à propos des contributions absolues.

2.2.6° a) Si $\xi_q = \eta_q = 0$, on a sous la condition C, $\rho_q = 1$; donc les notes complémentaires $k(i,q^+)$ et $k(i,q^-)$ sont inchangées.

2.2.6° b) La condition C et l'invariance de tous les profils sauf celui de ω entraîne :

$$\forall\, q \in Q : q \neq \omega \Rightarrow \rho_q = 1;\ a_{qn} = a_q;$$

mais alors on a aussi par différence $a_{\omega n} = a_\omega$, (car $\Sigma\, a_q = \Sigma a_{qn} = 1$). Indépendamment de la condition $A = A_n = 1$, nous venons de démontrer que si l'on veut modifier les coefficients relatifs des notes (que certains a_{qn}/A_n diffèrent des a_q/A) en respectant la métrique, on doit modifier le profil d'au moins deux notes.

2.2.7° Après identification, il vient :

$$(\rho'_q\,,\ \xi'_q\,,\ \eta'_q) = (-\rho_q\,,\ -1-\eta_q\,,\ -1-\xi_q)$$

l'équivalence des deux transformations est expliquée sur la figure 3. On se restreindra donc au cas $\rho_q > 0$, car le cas $\rho_q < 0$ fournit les mêmes tableaux à une permutation près des deux colonnes complémentaires q^+ et q^-. On peut encore dire que cette permutation est réalisée par la transformation $(\rho_q, \xi_q, \eta_q) = (-1, -1, -1)$: on verra au § 2.2.10° que les transformations $(\rho\,, \xi\,, \eta)$ forment un groupe.

2.2.8° Notons $\rho_q = 1 + \delta_q$; et supprimons l'indice q pour alléger l'écriture du calcul; on a :

$$(1 + \xi + \eta)\, xy(1 + \delta) = (x + \xi)\ (y + \eta)$$

$$(1 + \xi + \eta)\, xy\delta = (x + \xi)(y + \eta) - (1 + \xi + \eta)\, xy$$

$$(1 + \xi + \eta)\, xy\delta = \xi y(1 - x) + \eta x(1 - y) + \xi\eta$$

$$(1 + \xi + \eta)\, xy\delta = \xi y^2 + \eta x^2 + \xi\eta$$

$$\delta = (\xi(1 - x)^2 + \eta x^2 + \xi\eta)/((1 + \xi + \eta)\ x\ (1 - x))$$

Sous la condition C, ξ et η positifs implique donc ρ supérieur à 1.

Supposons maintenant $\rho > 0$; $\xi \leq 0$, $\eta \leq 0$; $\xi + \eta \neq 0$. On a nécessairement :

$$0 < a_{nq} = (1 + \xi_q + \eta_q)\ \rho_q a_q$$

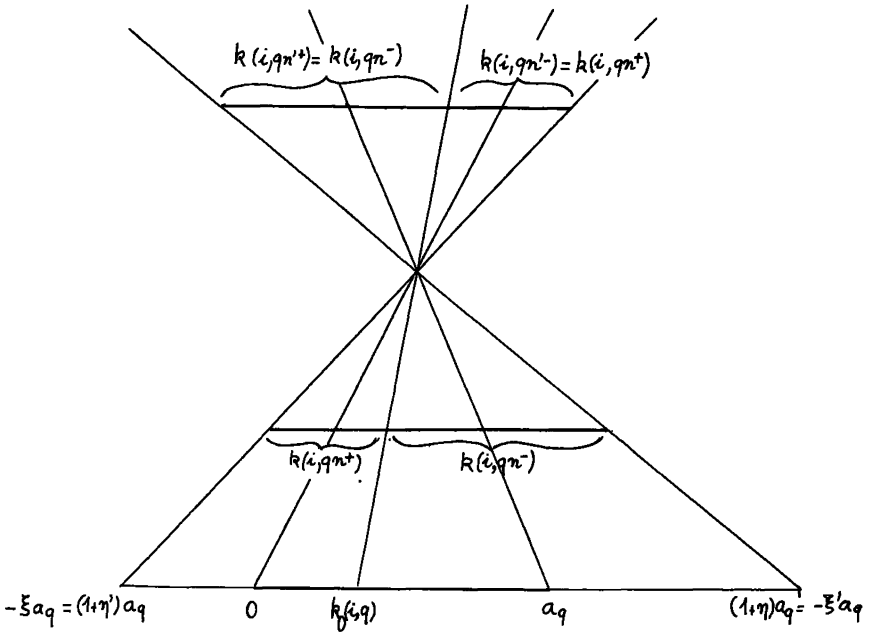

Figure 3 : permutation de deux colonnes complémentaires et changement de signe de ρ ; sur la figure on a posé : ξ, η positifs ; ρ = 1/2 ; d'où ρ' = -1/2 ; ξ' = -(1+η) < -1 ; η' = -(1+ξ) < -1 .

donc $(1 + \xi_q + \eta_q) > 0$; et le signe de δ est celui du numérateur :

$$\xi(1 - x)^2 + \eta x^2 + \xi\eta = (\xi + \eta)x^2 - 2\xi x + \xi + \xi\eta;$$

puisque $x_q = k_0(q)/(na(q))$ est compris entre 0 et 1, il suffit d'étudier dans cet intervalle la variation du numérateur : celui-ci est maximum pour $x = \xi/(\xi + \eta)$ on a alors :

$$(\xi + \eta)(\xi/(\xi + \eta))^2 - 2\,\xi(\xi/(\xi + \eta)) + \xi + \xi\eta =$$

$$= \xi\eta(1 + \xi + \eta)/(\xi + \eta) < 0;$$

et δ est négatif; donc $\rho \in (0,1)$. Nous redémontrerons ce résultat en 2.2.10°.

Si ρ_q est positif et que ξ_q et η_q sont de même signe ε, on a donc (sous la condition C) $(\rho_q - 1)$ et $a_{qn} - a_q = ((1 + \xi_q + \eta_q)\rho_q - 1)a_q$ du même signe ε. Par conséquent $\rho_q > 0$ et $a_{qn} = a_q$ (coefficient relatif inchangé) implique que ξ_q et η_q soient de signes contraires. On peut encore remarquer que si l'on a simultanément la condition C et $(a_{qn} = a_q)$, on a :

$x_q y_q = (x_q + \xi_q)(y_q + \eta_q)$; d'où :

$\eta_q = - \xi_q y_q/(x_q + \xi_q)$;

donc $(\xi_q > 0)$ implique $(\eta_q < 0)$; et de même $(\eta_q > 0)$ implique $(\xi_q < 0)$.

2.2.9° Pour être assuré de ne pas introduire de notes négatives par des transformations linéaires $(\rho_q , \xi_q , \eta_q)$, il faut que quel que soit $k_o(i,q) \in (0,a_q)$, on ait $k(i,qn^+)$ et $k(i,qn^-)$ positifs. Or on a :

$k(i,qn^+)/a_q = \rho_q((k_o(i,q)/a_q) + \xi_q) \in (\rho_q\xi_q , \rho_q(1 + \xi_q))$;

$k(i,qn^-)/a_q = \rho_q(((a_q - k_o(i,q))/a_q) + \eta_q) \in (\rho_q\eta_q , \rho_q(1 + \eta_q))$;

il faut donc qu'on ait :

soit $\rho_q > 0$; $\xi_q \geq 0$; $\eta_q \geq 0$;

soit $\rho_q < 0$; $\xi_q \leq - 1$; $\eta_q \leq - 1$;

d'après la remarque du § 2.2.7° on se restreindra aux transformations $\rho_q > 0$, $\xi_q \geq 0$, $\eta_q \geq 0$.

Or d'après 2.2.8° on a alors $\rho_q > 1$ et

$a_{qn} = a_q(1 + \xi_q + \eta_q)\ \rho_q \geq a_q$;

l'égalité $a_{qn} = a_q$ n'étant réalisée (sous la condition C) que si $(\rho_q , \xi_q , \eta_q) = (1,0,0)$, c'est-à-dire si les colonnes q^+ et q^- sont inchangées. Ainsi toute modification du tableau (compatible avec la condition C et la positivité) entraînerait un accroissement strict de la somme $A = \Sigma a_q$, contrairement à l'hypothèse faite $A_n = A = 1$. (Sans introduire cette hypothèse, on aurait obtenu : $\forall\ q$: $(a_{qn}/A_n) \geq (a_q/A)$). Il n'existe donc pas de modification du tableau compatible avec la condition C et la positivité, si chacune des notes $k_0(i,q)$ varie dans toute l'amplitude de l'intervalle $(0,a_q)$ *a priori* permis (i.e. si les correcteurs notent franchement de 0 à 20).

Mais supposons que l'une des notes q varie dans un sous-intervalle strict de $(0,a_q)$:

$(Xa_q , (1 - Y)a_q) \subset (0,a_q)$,

où les nombres X et Y sont positifs non tous deux nuls. On pourra alors poser $\xi_q = - X$, $\eta_q = - Y$; $\rho_q = (x_q + \xi_q)(y_q + \eta_q)/((1 + \xi_q + \eta_q)x_q y_q)$; et on aura alors $a_{nq} < a_q$ (inégalité stricte); d'où la possibilité d'effectuer sur les autres notes q' des transformations $(\rho_{q'} , \xi_{q'} , \eta_{q'})$ avec $\xi_{q'} , \eta_{q'}$ strictement positifs (mais assez petits pour que les accroissements $(a_{nq'} - a_{q'})$ pour $q' \neq q$, compensent exactement la diminution $(a_q - a_{nq})$. Ou encore on modifiera seulement la note q, en faisant en sorte que :

$(x + \xi)(y + \eta) = xy$; $\rho(1 + \xi + \eta) = 1$; $- X \leq \xi$; $- Y \leq \eta$,

(formules où l'indice q est omis), i.e. que $a_{nq} = a_q$ et que soit respectée la positivité des $k(i,qn^+)$ et $k(i,qn^-)$. Supposons e.g. $X \neq 0$, on pourra prendre $\xi < 0$, la plus forte valeur absolue permise pour ξ négatif étant X; soit $\xi = -X$; on aura alors, conformément au principe général (cf § 2.2.8°) η positif : $\eta = Xy/(x - X)$.

2.2.10° Etudions brièvement la loi de groupe des transformations décrites par les paramètres (ρ, ξ, η). Soit à composer deux de ces transformations : nous cherchons une formule :

$$(\rho', \xi', \eta') \circ (\rho, \xi, \eta) = (\rho'', \xi'', \eta'').$$

Considérons les codages successifs des notes $k(i,q^+)$ et $k(i,q^-)$, (notées ci-dessous simplement k^+ et k^-); et aussi de a_q (noté a) : on a :

$$k^+ \to \rho(k^+ + \xi a) \to \rho'(\rho(k^+ + \xi a) + \xi'\rho(1 + \xi + \eta)a) = \rho''(k^+ + \xi''a);$$

$$k^- \to \rho(k^- + \eta a) \to \rho'(\rho(k^- + \eta a) + \eta'\rho(1 + \xi + \eta)a) = \rho''(k^- + \eta''a);$$

$$a \to \rho(1 + \xi + \eta)a \to \rho\rho'(1 + \xi + \eta)(1 + \xi' + \eta')a = \rho''(1+\xi''+\eta'')a.$$

En identifiant, il vient :

$$(\rho', \xi', \eta') \circ (\rho, \xi, \eta) = (\rho'\rho,\ \xi + \xi'(1 + \xi + \eta),\ \eta + \eta'(1+\xi+\eta)).$$

La transformation identité (élément neutre du groupe) est (1,0,0). La permutation des colonnes k^+ et k^- (on l'a dit au § 2.2.7°) a pour paramètres (- 1, - 1, - 1): on a :

$$(-1, -1, -1) \circ (\rho, \xi, \eta) = (-\rho, -(1+\eta), -(1+\xi)) = (\rho,\xi,\eta)(-1,-1,-1)$$

La transformation inverse $(\rho, \xi, \eta)^{-1}$ est $(\rho^{-1}, -\xi/(1 + \xi + \eta), -\eta/(1 + \xi + \eta))$:

$$(\rho, \xi, \eta) \circ (\rho^{-1}, -\xi/(1 + \xi + \eta), -\eta/(1 + \xi + \eta)) = (1,0,0)$$

$$(\rho^{-1}, -\xi/(1 + \xi + \eta), -\eta/(1 + \xi + \eta)) \circ (\rho, \xi, \eta) = (1,0,0).$$

Supposons (cf 2.2.8°) ρ positif, ξ et η négatifs, $(1 + \xi + \eta)$ positif; on a alors pour la transformation inverse $(\rho',\xi',\eta') = (\rho,\xi,\eta)^{-1}$:

$$\rho' = 1/\rho \ ; \ \xi', \eta' \text{ positifs;}$$

si (ρ, ξ, η) satisfait à la condition C relativement au tableau initial (conserve la métrique) il en est de même de son inverse relativement au tableau transformé (dont il permet le retour au tableau initial); donc (cf 2.2.8°) ρ' est supérieur à 1, tandis que $\rho = 1/\rho'$ est inférieur à 1.