LES CAHIERS DE L'ANALYSE DES DONNÉES

J. S. TIBEIRO

Trace et information mutuelle pour un cas modèle de tableau décomposé en blocs

Les cahiers de l'analyse des données, tome 17, n° 2 (1992), p. 197-202

<http://www.numdam.org/item?id=CAD_1992__17_2_197_0>

Les Cahiers de l'Analyse des Données
Vol XVII - 1992 - n° 2 - pp. 197-202.

TRACE ET INFORMATION MUTUELLE POUR UN CAS MODÈLE DE TABLEAU DÉCOMPOSÉ EN BLOCS

[INF. BLOCS.]

*J. S. TIBEIRO**

1 Introduction: le modèle considéré

De récents travaux, (cf. K. Ben SALEM, [TRAC. MANQ.], in *CAD*, Vol. XVII, n°1; et [INF. MANQ.]), ont appelé l'attention sur la comparaison entre la trace et l'information mutuelle, toutes deux utilisées comme mesure de la dépendance entre deux éléments aléatoires *i* et *j* liés par un loi de probabilité p_{IJ}, donnée sur le produit de deux ensembles finis I et J.

D'ailleurs, il est bien connu que, dans certains cas, un tableau décomposé en blocs rectangulaires a une structure particulière qui permet d'en ramener l'analyse à celle de tableaux plus simples. On propose ici le modèle de tableau décomposé en blocs le plus général sur lequel nous ayons pu démontrer des résultats, qui comprennent ceux publiés auparavant: nous considérerons, d'une part, le calcul de l'information mutuelle et de la trace; puis, d'autre part, le calcul, complet ou partiel, des facteurs.

On part de deux ensembles finis, I et J, chacun réunion d'une famille finie d'ensembles finis non vides, deux à deux d'intersection vide, indicée par $a \in A$ ou $b \in B$:

$$I = \cup \{I_a \mid a \in A\} \quad ; \quad \forall a, a' \in A : (a = a') \Leftrightarrow (I_a \cap I_{a'} = \emptyset) \ ;$$

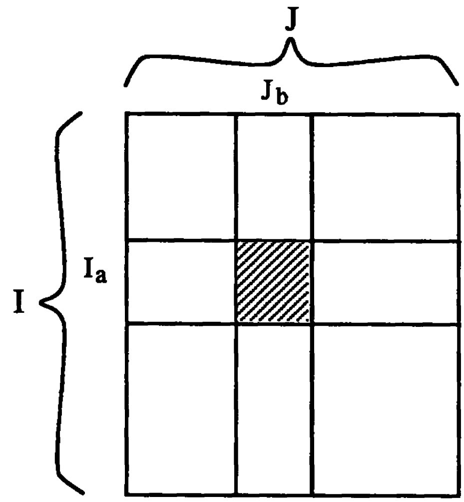
$$J = \cup \{J_b \mid b \in B\} \quad ; \quad \forall b, b' \in B : (b = b') \Leftrightarrow (J_b \cap J_{b'} = \emptyset) \ ;$$

un élément arbitraire de I_a ou J_b sera noté i_a ou j_b.

Chacun des ensembles I_a ou J_b est muni d'une loi de probabilité (mesure positive de masse totale unité) notée: p_{Ia} ou p_{Jb} ; et on a, sur chacun des produits

(*) Université de Sherbroke.

La restriction au bloc $I_a \times J_b$ de la loi $q_{I;J}$ est une mesure positive qui n'est autre que le produit par q_{ab} (masse du couple a,b pour la loi q_{AB}) de la loi $p_{Ia;Jb}$ donnée sur ce même produit $I_a \times J_b$.



$I_a \times J_b$, une loi de probabilité, notée $p_{Ia;Jb}$, dont les lois marginales sont les lois p_{Ia} et p_{Jb} spécifiées sur chacun des ensembles facteur; la masse d'un couple (i_a, j_b) sera simplement notée $p_{ia;jb}$.

Sur le produit $A \times B$ des ensembles finis qui donnent les indices des partitions, $\{I_a \mid a \in A\}$ et $\{J_b \mid b \in B\}$, de I et J, on a une loi de probabilité q_{AB}; la masse du couple (a, b) est notée q_{ab}; les lois marginales de q_{AB} seront généralement notées q_A, q_B; si, toutefois, les deux ensembles A et B coïncident, sans que q_{AA} soit symétrique, on peut distinguer les deux lois marginales en les notant $q_{A.}$ et $q_{.A}$. De même, la masse d'un élément pour la loi marginale est notée q_a ou q_b ; voire $q_{a.}$ ou $q_{.a}$.

En combinant toutes ces données, on construit une loi de probabilité $q_{I;J}$ sur $I \times J$; la masse $q_{ia;jb}$ d'une paire (i_a, j_b) est donnée par le produit:

$$q_{ia;jb} = q_{ab}\, p_{ia;jb} \quad ; \quad q_{Ia;Jb} = q_{ab} \cdot p_{Ia;Jb} \quad ;$$

ainsi la restriction, $q_{Ia;Jb}$, de $q_{I;J}$ au bloc $A \times B$, a pour masse totale q_{ab} et est proportionnelle à $p_{Ia;Jb}$. On vérifie que les lois marginales q_I et q_J de la loi $q_{I;J}$ ont, de même, respectivement, pour restriction à un segment I_a et J_b le produit de la loi p donnée sur le segment par la masse marginale correspondante de q_{AB} :

$$q_{Ia} = q_a \cdot p_{Ia} \quad ; \quad q_{Jb} = q_b \cdot p_{Jb}$$

2 Calcul de l'information mutuelle

Pour plus de brièveté, nous reprenons, sans commentaire, les notations de [INF. MANQ.]:

$$H(q_I) = \Sigma\{-q_i \log(q_i) \mid i \in I\} \quad ; \quad H(q_J) = \Sigma\{-q_j \log(q_j) \mid j \in J\} \quad ;$$

$$\begin{aligned} Hm(q_{I;J}) &= H(q_I) + H(q_J) - H(q_{I;J}) \\ &= \Sigma\{-q_{i;j} \log(q_{i;j} / (q_i \,.\, q_j)) \mid i \in I\,;\, j \in J\} \quad ; \end{aligned}$$

dans cette dernière formule, on peut substituer, à la sommation double par rapport à *i* et *j*, une sommation double par rapport à l'ensemble des blocs, lequel est indicé par *a* et *b*; avec au sein de chaque bloc, une sommation double par rapport à ia et ib. Il vient:

$$Hm(q_{I;J}) = \Sigma\{Som(a, b) \mid a \in A\,;\, b \in B\} \quad ; \quad \text{où} \;:$$

$$Som(a, b) = \Sigma\{-q_{ia;ib} \log(q_{ia;ib} / (q_{ia} \,.\, q_{ib})) \mid i_a \in I_a\,;\, j_b \in J_b\} \;;$$

en tenant compte des formules du §1:

$$q_{ia;jb} = q_{ab}\, p_{ia;jb} \quad ; \quad q_{ia} = q_a\, p_{ia} \quad ; \quad q_{jb} = q_b\, p_{jb} \quad ;$$

puis en décomposant le logarithme du produit en somme de logarithmes, il vient pour expression de $S_{ab} = Som(a, b)$:

$$\begin{aligned} S_{ab} &= q_{ab}\,\Sigma\{-p_{ia;jb} \log((p_{ia;jb} / (p_{ia} \,.\, p_{jb})).(q_{ab} / (q_a \,.\, q_b))) \mid i_a \in I_a\,;\, j_b \in J_b\} \\ &= (-q_{ab}) \log(q_{ab} / (q_a \,.\, q_b))\, \Sigma\{p_{ia;jb} \mid i_a \in I_a\,;\, j_b \in J_b\} \\ &\qquad\qquad + q_{ab}\,\Sigma\{-p_{ia;jb} \log(p_{ia;jb} / (p_{ia} \,.\, p_{jb})) \mid i_a \in I_a\,;\, j_b \in J_b\} \quad ; \end{aligned}$$

de ce que $p_{Ia;Jb}$ est une loi de probabilité il résulte qu'est égale à un la somme $\Sigma\{p_{ia;jb} \mid i_a \in I_a\,;\, j_b \in J_b\}$; ce qui permet de simplifier l'expression de S_{ab} :

$$S_{ab} = (-q_{ab}) \log(q_{ab} / (q_a \,.\, q_b)) \;+\; q_{ab}\, Hm(p_{Ia;Jb}) \quad ;$$

et pour l'information mutuelle afférente à la loi $q_{I;J}$ sur $I \times J$:

$$\begin{aligned} Hm(q_{I;J}) &= \Sigma\{S_{ab} \mid a \in A\,;\, b \in B\} \\ &= Hm(q_{AB}) + \Sigma\{q_{ab}\, Hm(p_{Ia;Jb}) \mid a \in A\,;\, b \in B\} \quad ; \end{aligned}$$

il est normal de retrouver dans $Hm(q_{I;J})$ l'information mutuelle $Hm(q_{AB})$, car la loi q_{AB} est la loi quotient de $q_{I;J}$ par la relation d'équivalence définies par les partitions $\{I_a \mid a \in A\}$ et $\{J_b \mid b \in B\}$, de I et J; quant à la somme Σ des $q_{ab}\, Hm(p_{Ia;Jb})$, on peut l'interpréter comme l'espérance mathématique de l'information complémentaire transmise par *i* sur *j* (ou par *j* sur *i*) quand on passe entre I_a et J_b; mais cette interprétation ne nous paraît pas pouvoir offrir la base d'une véritable démonstration.

3 Calcul de la trace

Comme dans le calcul de l'information mutuelle, on substitue, à la sommation double par rapport à *i* et *j*, une sommation double par rapport à l'ensemble des blocs, lequel est indicé par *a* et *b*; avec au sein de chaque bloc, une sommation double par rapport à ia et ib:

$$\begin{aligned}\mathrm{Tr}(q_{I;J}) &= \Sigma\{(q_{i;j})^2 / (q_i \,.\, q_j) \mid i \in I \,;\, j \in J\} - 1\\ &= \Sigma\{((q_{i;j})^2 / (q_i \,.\, q_j)) \mid i \in I \,;\, j \in J\} - \Sigma\{q_{ab} \mid a \in A \,;\, b \in B\}\\ &= \Sigma\{T_{ab} \mid a \in A \,;\, b \in B\} \,;\end{aligned}$$

où on a noté:

$$\begin{aligned}T_{ab} &= \Sigma\{((q_{ia;jb})^2 / (q_{ia} \,.\, q_{jb})) \mid i_a \in I_a \,;\, j_b \in J_b\} - q_{ab} \,;\\ &= ((q_{ab})^2 / (q_a \,.\, q_b))\, \Sigma\{(p_{ia;jb})^2 / (p_{ia} \,.\, p_{jb}) \mid i_a \in I_a \,;\, j_b \in J_b\} - q_{ab}\\ &= ((q_{ab})^2 / (q_a \,.\, q_b))\, (\mathrm{Tr}(p_{Ia;Jb}) + 1) - q_{ab} \,;\end{aligned}$$

d'où pour la somme double qui exprime la trace cherchée:

$$\begin{aligned}\mathrm{Tr}(q_{I;J}) = {} & \Sigma\{((q_{ab})^2 / (q_a \,.\, q_b)) - q_{ab} \mid a \in A \,;\, b \in B\}\\ & + \Sigma\{((q_{ab})^2 / (q_a \,.\, q_b))\, \mathrm{Tr}(p_{Ia;Jb}) \mid a \in A \,;\, b \in B\} \,;\end{aligned}$$

dans la première somme, on reconnaît $\mathrm{Tr}(q_{AB})$; d'où la formule:

$$\mathrm{Tr}(q_{I;J}) = \mathrm{Tr}(q_{AB}) + \Sigma\{((q_{ab})^2 / (q_a \,.\, q_b))\, \mathrm{Tr}(p_{Ia;Jb}) \mid a \in A \,;\, b \in B\} \,;$$

comme au §2, on a un premier terme afférent à loi q_{AB}, la loi quotient de $q_{I;J}$; vient ensuite une somme de traces dont les coefficients de pondération ne sont autres que les termes de la somme par laquelle s'exprime $\mathrm{Tr}(q_{AB}) + 1$.

On sait (cf. [Inf. Tab.], §1.3.2) qu'au voisinage de l'indépendance on a l'équivalence:

$$\mathrm{Tr}(q_{I;J}) \approx 2\,\mathrm{Hm}(q_{I;J}) \;;$$

cette équivalence est compatible avec les formules de décomposition trouvées ici. Supposons, en effet, que tous les quotients $\{q_{ab}/(q_a\, q_b)\}$ et $\{p_{ia;jb}/(p_{ia}\, p_{jb})\}$ ne diffèrent de 1 que par une quantité du 1-er ordre; alors les traces $\mathrm{Tr}(q_{AB})$ et

$Tr(p_{Ia;Jb})$ sont des quantités du 2-ème ordre équivalentes; respectivement, à $2.Hm(q_{A;B})$ et $2.Hm(p_{Ia;Jb})$. Le coefficient $(q_{ab})^2 / (q_a \,.\, q_b)$ ne diffère de q_{ab} que par un terme du 1-er ordre; et finalement, dans la formule de décomposition de $Tr(q_{I;J})$, les termes (tous d'ordre 2) sont équivalents respectivement à ceux de $2.Hm(q_{I;J})$ à des termes du 3-ème ordre près.

4 Cas particuliers et recherche de facteurs pour le tableau décomposé en blocs

4.1 Tableau décomposé en blocs diagonaux

Le cas d'un tableau décomposé en blocs diagonaux est bien connu. Avec les notations introduites ici, il s'agit du cas particulier où A = B, la loi q_{AA} étant portée par la diagonale; dans ce cas, seuls comptent les blocs $p_{Ia;Ja}$, car il n'y a pas lieu d'introduire des blocs extra-diagonaux $p_{Ia;Ja'}$, dont le coefficient $q_{aa'}$ serait nul. La trace $Tr(q_{IJ})$ comprend, d'une part, $Tr(q_{AA})$ qui, pour un tableau diagonal, n'est autre que cardA-1; et, d'autre part, la somme des traces des cardA blocs $Tr(p_{Ia;Ja})$: car les coefficients $(q_{aa})^2 / (q_{a.} \times q_{.a})$ sont tous égaux à 1.

Dans la formule de l'information mutuelle $Hm(q_{IJ})$, les coefficients q_{aa} qui affectent les $Hm(p_{Ia;Ja})$ ne sont pas égaux à 1: c'est leur somme qui vaut 1. Et le terme $Hm(q_{AA})$, qui n'est autre que la néguentropie de la loi portée par la diagonale de q_{AA} (ou encore de l'une ou l'autre des lois marginales, égales entre elles), n'est pas déterminé par cardA: il est seulement majoré par log(cardA).

L'analyse factorielle d'un tableau décomposé en blocs diagonaux se ramène à celle des $p_{Ia;Ja}$: mais, plutôt que de répéter des résultats classiques, nous préférons les déduire de propositions plus générales.

4.2 Facteurs de q_{AB} pour q_{IJ}

L'analyse de q_{AB} fournit des couples de facteurs normalisés (φ^A, φ^B) associés à une valeur propre λ; on vérifie, en appliquant la formule de transition, que ces facteurs normalisés se retrouvent pour q_{IJ} associés à la même valeur propre; les fonctions (φ^A, φ^B), définies sur les ensembles quotients (A, B) induisant des fonctions (φ^I, φ^J), avec $\varphi^{ia} = \varphi^a$; $\varphi^{jb} = \varphi^b$. Dans le cas où q_{AA} est diagonal (§4.1) ces facteurs ne sont autres que les fonctions de moyenne nulle sur A; la valeur propre étant 1.

4.3 Facteurs des $p_{Ia;Jb}$ pour q_{IJ}

Il n'est pas vrai, en général, que les facteurs issus des $p_{Ia;Jb}$ se retrouvent pour q_{IJ}, même avec une autre normalisation et une autre valeur propre.

Mais q_{IJ} hérite des facteurs issus de $p_{Ia;Jb}$ dans le cas particulier où les bloc $p_{Ia;Jb'}$, d'indice $b'\neq b$, ainsi que les blocs $p_{Ia';Jb}$, d'indice $a'\neq a$, sont de trace nulle (i.e. sont chacun produit de ses lois marginales; ou, ce qui revient au même, sont affectés de coefficients, $q_{ab'}$ ou $q_{a'b}$, nuls).

De façon précise, soit $(\varphi^{Ia}, \varphi^{Jb})$ un couple de facteurs normalisés, issu de $p_{Ia;Jb}$, associé à la valeur propre λ; les fonctions (φ^I, φ^J) obtenues en prolongeant $(\varphi^{Ia}, \varphi^{Jb})$ par 0 sur I et J, et en les multipliant respectivement par $q_a^{-1/2}$ et $q_b^{-1/2}$, sont de moyenne nulle et variance 1 sur I et J. De plus (φ^I, φ^J) est un couple de facteurs normalisés pour q_{IJ} associé à la valeur propre $((q_{ab})^2/(q_a.q_b))\lambda$; pour démontrer ce résultat, il suffit de considérer que la transition $q_I{}^J$ a même effet sur φ^I, que la transition $p_{Ia}{}^{Jb}$ appliquée à φ^{Ia}, avec toutefois, en plus, le coefficient (q_{ab}/q_a); et de même pour l'action de $q_J{}^I$ et $p_{Jb}{}^{Ia}$ sur φ^J et φ^{Jb}, le coefficient étant ici (q_{ab}/q_b).

Cette proposition permet d'achever l'analyse d'un tableau décomposé en blocs diagonaux; elle a le même effet dans le cas, plus général, où A=B, le tableau q_{AA} étant de forme quelconque; mais les $p_{Ia;Ja'}$ tous triviaux (i.e. de trace nulle), à l'exception des $p_{Ia;Ja}$; lesquels, précisément fournissent les facteurs de q_{IJ}, autres que ceux issus de q_{AA}.

Ce dernier cas est proposé en exercice dans [Graph. Corr.] §2.5; et une solution détaillée se trouve dans [BLOC. LAT. SIMP.].

Références bibliographiques

[BLOC. LAT. SIMP.], in *Pratique de l'Analyse des Données*; Vol.2, IV n°12.

[Graph. Corr.], in *L'Analyse des Données*, TIIB n°10.

[INF. MANQ.], in *CAD*, Vol. XVII, n°2, (1992).

[Inf. Tab], in *L'Analyse des Données*, TIB n°5.

[TRAC. MANQ.], in *CAD*, Vol. XVII, n°2, (1992).