

M.-M. THOMASSIN

K. BEN SALEM

**Estimation de la diagonale d'un tableau
carré d'après le critère de la trace minima ;
application à l'analyse des flux de population
à la Martinique entre 1975 et 1982**

Les cahiers de l'analyse des données, tome 17, n° 3 (1992),
p. 261-274

http://www.numdam.org/item?id=CAD_1992__17_3_261_0

© Les cahiers de l'analyse des données, Dunod, 1992, 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/>

**ESTIMATION DE LA DIAGONALE
D'UN TABLEAU CARRÉ
D'APRÈS LE CRITÈRE DE LA TRACE MINIMA;
APPLICATION À L'ANALYSE DES
FLUX DE POPULATION À LA MARTINIQUE
ENTRE 1975 ET 1982**

[TRAC. MANQ. FLUX]

*K. Ben SALEM**

*M.-M. THOMASSIN***

0 Introduction: problème d'estimation et données de flux

Dans un travail récent ([MARTINIQUE FLUX], in *CAD*, Vol XVI, n°3), l'un d'entre nous, (M.-M. TH.; en collaboration avec M. OUASSOU), a rappelé l'attention sur le fait que, dans une matrice de flux (en l'occurrence une matrice de flux migratoires de commune à commune entre deux recensements), les termes diagonaux manquent. On peut certes considérer que ce qui reste en place au point i (dans l'article cité: la population non migrante) constitue un flux de i vers i ; mais cette assimilation engendre une diagonale très chargée qui perturbe l'analyse, avec de fortes valeurs propres et des facteurs qui n'expriment nettement que l'association de tel ou tel point i avec lui-même.

Dans l'article cité, un raisonnement plausible a permis de calculer des termes diagonaux d'un ordre de grandeur acceptable, en fonction de l'ensemble des données disponibles (flux proprement dits et population totale); et l'analyse est très satisfaisante. Cependant l'autre auteur, (K. Ben S.), a proposé une méthode générale pour créer la diagonale d'un tableau carré sans traiter d'autres données que les termes extradiagonaux (i.e., dans l'exemple, les flux): le critère étant, en bref, que, dans le tableau carré ainsi complété, la trace soit aussi faible que possible, ce qui prémunit, *a fortiori*, contre la présence des très fortes valeurs propres indésirables (cf. [TRAC. MANQ.], in *CAD*, Vol XVII, n°1).

(*) Département des Sciences de l'Informatique; Faculté des Sciences de Tunis.

(**) Géographe à l'ORSTOM.

Dans le présent travail, on rappelle d'abord, au §1, le format des données traitées dans [MARTINIQUE FLUX] et la voie suivie pour créer des termes diagonaux acceptables. On expose ensuite, au §2, l'application, à un tableau carré, du critère de la trace minima; sans reprendre les démonstrations données dans [TRAC. MANQ.], mais en expliquant l'algorithme qui permet d'appliquer ce critère. Enfin, au §3, le critère de la trace minima est appliqué à l'analyse des flux à la Martinique; en produisant des résultats qui concordent avec ceux de [MARTINIQUE FLUX]; fait d'autant plus remarquable que la diagonale est ici créée en fonction des seuls flux.

1 Estimation des migrations intérieures à chacune des unités territoriales, d'après une tendance migratoire générale

Les recensements, effectués à la Martinique en 1975 et 1982, permettent de construire deux tableaux carrés Flx (un pour chacun des deux sexes: $x=h$, ou $x=f$), donnant l'état comparatif de la population de l'île aux deux recensements. Un tableau Flx croise avec lui-même l'ensemble I des 34 communes de l'île. De façon précise, on trouve, à l'intersection de la ligne i et de la colonne i' , le nombre $\text{Flx}(i, i')$ des habitants, (du sexe x), recensés, en 1982, dans la commune i , et, en 1975, dans la commune i' .

Les termes extradiagonaux mesurent des flux proprement dits; mais $\text{Flx}(i, i)$ est le nombre des habitants recensés dans la même commune i , lors de chacun des deux recensements de 1975 et 1982, i.e., principalement, des habitants sédentaires. On modifie les termes diagonaux de telle sorte qu'ils puissent être regardés comme une estimation vraisemblable de flux internes; le but n'étant pas de considérer, pour lui-même, le nombre des habitants ayant déménagé, une ou plusieurs fois, entre 1975 et 1982, pour se retrouver finalement dans la même commune, sinon à la même adresse; mais de décharger la diagonale du tableau Flx en sorte qu'elle ne perturbe pas l'analyse.

À cette fin, on calcule, pour chacune des communes i , un coefficient d'émigration $\mu_x(i)$ tel que:

$$\forall ip \in I - \{i\} : \text{Flx}(ip, i) \approx \mu_x(i) \times \text{Totx}(ip) ;$$

en d'autres termes, tel que le nombre des habitants passés de i à ip puisse être assimilé au produit de la population totale, $\text{Totx}(ip)$, (du sexe x), de la commune ip de destination, par le coefficient, $\mu_x(i)$, propre à la commune de départ i . Le coefficient est calculé pour que la relation soit exacte, en moyenne; soit:

$$\mu_x(i) = \sum\{\text{Flx}(ip, i) \mid ip \in I; ip \neq i\} / \sum\{\text{Totx}(ip) \mid ip \in I; ip \neq i\} ;$$

on peut, pour faire image, dire qu'une fraction de la population de i s'évapore; puis retombe sur toutes les communes de l'île en proportion de la population de chacune de celles-ci; une partie des habitants retombant là-même d'où ils étaient partis.

Dans la matrice, Flx , $I \times I$, ainsi corrigée, la ligne i donne les flux de I vers i ; tandis que la colonne i donne les flux de i vers I ; afin d'avoir, sur une même ligne du tableau que l'on analyse, toute l'information afférente à une commune i , on adjoint à la matrice Flx sa transposée. De plus, on range, dans cinq colonnes, $\{xND, xzz, xFr, xÉt, xsd\}$, des informations complémentaires publiées avec le dépouillement du recensement (pour le sexe x):

- $Flx(i, xND)$: population résidant en i en 1982, et dont la situation en 1975 est Non Déterminée;
- $Flx(i, xzz)$: population résidant en i en 1982, et née après 1975;
- $Flx(i, xFr)$: population en i en 1982, et venue de France après 1975;
- $Flx(i, xÉt)$: population en i en 1982, et venue de l'Étranger après 1975;
- $Flx(i, xsd)$: population sédentaire calculée en retranchant, de la valeur initiale de $Flx(i, i)$, la valeur estimée ci-dessus, par le coefficient $\mu_x(i)$.

On a ainsi, pour l'ensemble des deux sexes, un tableau à 34 lignes, (ensemble I des communes), et 146 colonnes:

$$146 = 2 \times ((2 \times 34) + 5) ;$$

à l'analyse de ce tableau est consacré le §3 de [MARTINIQUE FLUX]: une partition de I en 9 classes est présentée sur une carte. De l'ensemble J des 146 colonnes, on retient une partition en 18 classes; et l'arbre de la CAH de I est étiqueté d'après le listage $iVACORJq$, afférent à ces deux partitions.

Dans le présent travail, on suit exactement la même voie; à ceci près que la diagonale de chacun des tableaux Flh et Flf est calculée par une méthode générale, fondée, comme on l'a dit, sur le critère de la trace minima. Plus précisément, on rend compte brièvement des résultats obtenus avec plusieurs variantes de cette méthode.

2 Calcul de valeurs diagonales pour un tableau carré d'après le critère de la trace minima

Puisqu'on trouve un exposé en forme dans [TRAC. MANQ.], §§2 & 5, on se bornera à présenter ici les calculs effectués comme une suite de procédures écrites en langage PASCAL; en justifiant, sans démonstration, la démarche adoptée.

Nous supposons donné un tableau carré $kji[j, i]$, de dimension $carti \times carti$, lu sur un support extérieur et logé en mémoire centrale; ce tableau n'est pas nécessairement symétrique: mais les calculs sont effectués en se référant au tableau symétrisé correspondant; lequel, toutefois, n'est pas créé explicitement, car les termes extradiagonaux de kji sont respectés tout au long des calculs; et, finalement, récrit sur le support extérieur, avec les termes diagonaux qui ont été calculés.

```

for j:=1 to carti do kji[j,j]:=0;ktot:=0;
for i:=1 to carti do begin tot:=0;
  for ip:=1 to carti do tot:=tot+kji[i,ip];
  for j:=1 to carti do tot:=tot+kji[j,i];
  tot:=tot/2;ktot:=ktot+tot;kji[0,i]:=tot;end;
for i:=1 to carti do kji[0,i]:=kji[0,i]/ktot;

```

Les termes diagonaux donnés initialement n'étant pas pris en compte, la diagonale de k_{ji} est mise à zéro; on range, sous un indice '0', la marge du symétrisé du tableau ainsi modifié; plus précisément, le profil de cette marge: car on divise par $ktot$.

```

bil:=0;i1:=0;
for i:=1 to carti do if (bil<kji[0,i]) then begin
  i1:=i;bil:=kji[0,i];end;

```

On désigne par $i1$, l'indice correspondant à la plus forte composante du profil de marge, et par b_{i1} cette composante.

On entre ici dans une étape, où les valeurs extradiagonales des $k_{ji}[j, i]$ ne jouent aucun rôle. Le calcul est conduit comme si les données provenaient d'une loi produit de deux lois marginales identiques: $a_i = \{a_i \mid i \in I\} = \{a_j \mid j \in J\}$, avec $\sum\{a_i \mid i \in I\} = 1$; les $k_{ji}[j, i]$ donnés, étant, pour $i \neq j$, les produits des $a_j \times a_i$ par un coefficient k convenable qui, toutefois, est inconnu. Selon ce schéma, le profil $\{b_i \mid i \in I\}$ est proportionnel à la distribution $\{a_i - (a_i)^2 \mid i \in I\}$, dont la masse totale est inférieure à 1. Il s'agit de retrouver a_i .

À cette fin, comme l'explique [TRAC. MANQ.], on peut se baser sur la valeur de la composante a_{i1} ; dont l'indice, $i1$, est celui de la composante maxima, b_{i1} , du profil b_i . En effet, a_{i1} étant fixé, on se fonde sur la relation:

$$\forall i \in I - \{i1\} : (a_i - (a_i)^2) / b_i = (a_{i1} - (a_{i1})^2) / b_{i1} \quad ,$$

pour calculer les autres a_i : il s'agit simplement de résoudre une équation du 2-ème degré, dont les deux racines ont pour somme 1; la racine inférieure à 1/2 pouvant seule convenir, puisque a_{i1} est plus grand que tous les autres a_i , comme b_{i1} est plus grand que tous les autres b_i .

```

procedure sigmer;begin
  dh:=(a11-sqr(a11))/bil;tot:=a11;
  for i:=1 to carti do if not(i=i1) then
    tot:=tot+(0.5*(1-sqrt(1-(4*kji[0,i]*dh)))));end;

```

On aura fait pour a_{i1} un heureux choix, si est égale à 1 la somme, $\sum(a_{i1})$, de a_{i1} et de tous les autres a_i déterminés d'après a_{i1} ; ce qu'on peut vérifier par la procédure 'sigmer'. Au premier abord, le lecteur peut nous croire engagé dans des tâtonnements incertains: mais la voie est sûre, parce qu'on a démontré que $\Sigma(a)$ est une fonction convexe, telle que $\Sigma(0)=0$, $\Sigma(1)=1$, et qui a, entre 0 et 1, un maximum supérieur à 1.

```

tot0:=0;
for ip:=1 to 9 do begin ail:=ip/10;sigmer;
  if ((tot0<1) and (tot=1)) then aail:=ail;
  if ((tot0<1) and (1<tot)) then begin
    aail:=(ip-(tot-1)/(tot-tot0))/10;end;
  tot0:=tot;end;

```

Le bloc d'instructions ci-dessus montre comment on peut déterminer la valeur aa_{i1} pour laquelle $\Sigma(aa_{i1}) = 1$: on essaie, avec un pas de 0.1, les 9 valeurs de a_{i1} intermédiaires entre 0 et 1; quand Σ traverse la valeur 1, aa_{i1} peut être située; et on en calcule une valeur approchée en assimilant à une ligne brisée la courbe représentative de Σ .

Il va sans dire que ce calcul approché peut facilement être amélioré; mais, d'une part, pour une courbe convexe, l'assimilation à une ligne brisée n'expose pas à de graves erreurs; d'autre part, les 9 valeurs intermédiaires de Σ ont, en fait, été retenues, pour vérifier que le dessin de la courbe était conforme au schéma qui illustre la fin du §5 de [TRAC. MANQ.]. Le résultat a été ensuite précisé en prenant un pas de 0.01: l'expérimentation, sur les données du recensement de la Martinique ou sur des cas modèles, cf. *infra*, a montré que la précision était adéquate.

```

procedure tabler;begin ail:=aail;
dh:=(ail-sqr(ail))/bil;ch:=ktot*bil/(ail-sqr(ail));
kji[i1,i1]:=sqr(ail);
for i:=1 to carti do if not(i=i1) then begin
  res:=1-(4*kji[0,i]*dh);if (res<0) then res:=0;
  kji[i,i]:=sqr(0.5*(1-sqrt(res)));end;
for i:=1 to carti do kji[i,i]:=ch*kji[i,i];end;

```

La procédure 'tabler' attribue une diagonale au tableau kji d'après la valeur optima, aa_{i1} , trouvée pour a_{i1} . On reprend les calculs de la procédure "sigmer", mais, de plus, on utilise un coefficient 'ch' pour mettre à l'échelle les valeurs (a_i^2) de telle sorte que $(a_{i1} - (a_{i1})^2)$ prenne la valeur voulue: $ktot \times b_{i1}$, (marge du tableau initial à diagonale nulle).

Le tableau ainsi créé est conservé; mais on peut aller plus avant dans la construction de la diagonale en tenant compte, maintenant, du détail des termes extradiagonaux, sans sortir, toutefois de la considération de la matrice symétrisée. Comme l'explique [TRAC. MANQ.], on ajoute à la diagonale, présentement proportionnelle à $\{(a_i)^2 \mid i \in I\}$, un terme de plus, proportionnel à $a_I = \{a_i \mid i \in I\}$; le coefficient de ce terme étant choisi pour que la trace soit aussi faible que possible.

Le fait que l'on ne considère que des lois de probabilités sur $I \times I$ ayant une marge fixée a_I joue ici un rôle essentiel. De façon précise, soit p_{II} la loi de probabilité associée au tableau kji que l'on vient de construire: on considère les

```

for i:=1 to carti do begin tot:=0;
  for ip:=1 to carti do tot:=tot+kji[i,ip];
  for j:=1 to carti do tot:=tot+kji[j,i];
  kji[0,i]:=tot/2;end;
tr1:=0;tr2:=0;
for i:=2 to carti do for j:=1 to i-1 do
  tr2:=tr2+(sqr(kji[j,i]+kji[i,j]))/(kji[0,i]*kji[0,j]);
tr2:=tr2/2;
for i:=1 to carti do
  tr2:=tr2+sqr((kji[i,i]-kji[0,i])/kji[0,i]);
for i:=1 to carti do
  tr1:=tr1+((kji[i,i]-kji[0,i])/kji[0,i]);
if not(tr1*tr2=0) then begin tt:=-tr1/tr2;tt:=(1-tt)/tt;
  for i:=1 to carti do kji[i,i]:=kji[i,i]+(tt*kji[0,i]);end;

```

lois de la forme: $f_{II}(t) = (1-t) p_{II} + t \cdot a_1 \times a_1$. La trace de $f(t)$ est une fonction quadratique de t , en sorte qu'il suffit de déterminer les coefficients des termes de degré 2 et de degré 1 de cette fonction pour avoir la dérivée, qui est une fonction linéaire; laquelle s'annule au minimum cherché.

Les calculs sont compliqués par le fait que, d'une part, on opère, non sur p_{II} , mais sur le tableau kji lui-même; et que, d'autre part, il convient de modifier seulement les termes diagonaux; à l'exclusion des termes extradiagonaux, imposés au départ: d'où l'introduction du tableau $p_{II} + (t/(1-t)) a_1 \times a_1$.

Dans le bloc d'instructions proposé ici, on se réfère d'abord à un tableau symétrique $sji[j, i]$ donné par la formule:

$$\forall (i, j) \in I \times I, i \neq j \Rightarrow sji[j, i] = t \times (kji[j, i] + kji[i, j]) / 2 ;$$

$$\forall i \in I : sji[i, i] = (t \times kji[i, i]) + ((1 - t) \times kji[0, i]) ,$$

où $kji[0, i]$ désigne la marge du symétrisé de $kji[j, i]$.

La trace de ce tableau est calculée, suivant la formule générale, comme la somme (diminuée de 1) des carrés des termes divisés par le produit des marges, soit ici:

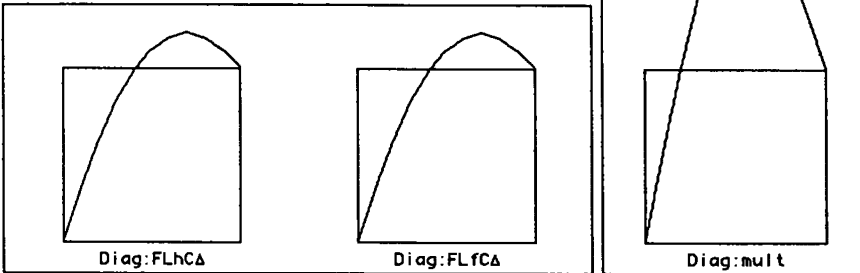
$$\text{trace} = \sum \{ sji[j, i]^2 / (kji[0, i] \times kji[0, j]) \mid i \in I, j \in J \} - 1 ;$$

plus précisément, il suffit de calculer, pour \sum , les coefficients, $tr2$ et $2 \cdot tr1$, des termes du 2-ème et du 1-er degré en t .

Le maximum de la trace est atteint pour la valeur $t = -tr1/tr2$; et l'on garde sur un support externe un tableau kji dont les termes extradiagonaux sont conservés, tandis qu'aux termes diagonaux des valeurs sont affectées par la formule:

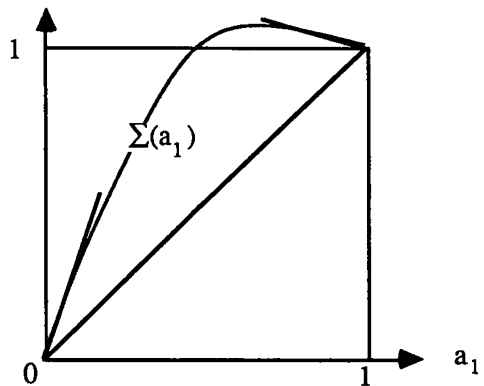
$$kji[i, i] := kji[i, i] + ((1-t) / t) \times kji[0, i] .$$

Recherche de la diagonale pour un tableau de flux



Les graphiques ci-dessus correspondent au graphique de [TRAC. MANQ.], §5, reproduit ci-dessous: la variation de $\sum(a_{i1})$, en fonction de a_{i1} , est donnée dans trois cas: d'une part, les données statistiques du recensement pour chacun des deux sexes; d'autre part, le cas modèle où l'on part d'un tableau (9×9) qui n'est autre que la table de multiplication: $kji[j, i] = i \times j$; le programme retrouvant la diagonale: $kji[i, i] = i^2$.

On a également expérimenté sur la table de multiplication à 10 lignes et colonnes; ainsi que sur une table de multiplication dont les termes extradiagonaux sont modifiés, sans que toutefois les marges en soient perturbées. Dans ce dernier cas, joue effectivement l'optimisation de la trace par adjonction, à la diagonale, d'un terme proportionnel à la marge; ce qui permet de vérifier le jeu du programme.



Dans cette expérimentation, il est apparu que la précision du calcul de la valeur optimale de a_{i1} était suffisante; et qu'en modifiant ultérieurement la diagonale, on diminuait quelque peu la trace.

Mais c'est dans l'interprétation des résultats obtenus sur les données réelles du recensement de la Martinique qu'il convient de chercher une justification pratique de nos calculs de données manquantes, fondés sur le critère de la trace minima: tel est l'objet du § qui suit.

Diverses constructions des termes diagonaux $k_{ji}[i, i]$

comm	diag+marg%	diag+ μ marg	$\partial\mu$	∂	év	Tséd
01	1.47544e+0	1.47111e+0	1	0	0	807
02	7.17882e+0	7.16518e+0	7	2	3	1873
03	8.72729e+0	8.71216e+0	9	3	4	2056
04	1.04304e+0	1.03985e+0	1	0	0	435
05	2.10152e+0	2.09572e+0	2	0	0	596
06	3.52984e+0	3.52121e+0	4	1	1	944
07	1.23039e+0	1.22669e+0	1	0	0	530
08	3.92770e+0	3.91839e+0	4	1	0	565
09	1.99099e+1	1.98901e+1	20	10	9	4035
10	1.04146e+2	1.04186e+2	104	75	21	4447
11	2.52288e+3	2.54444e+3	2523 (44)	2381	1748	31181
12	1.75034e+2	1.75171e+2	175	136	76	8430 ...

3 Analyse des flux de population entre communes d'après un tableau complété suivant le critère de la trace minima

Au §3.1, nous comparons les diverses constructions de la diagonale appliquées au tableau Flh, (34×34) des flux de sujets masculins entre communes, de 1975 à 1982. On a expliqué, au §1, comment, les tableaux carrés Flh et Flf étant munis d'une diagonale, on construit un tableau (34×146) dont chaque ligne renferme l'ensemble des informations élaborées afférentes à une commune. De la construction adoptée pour la diagonale, dépendent les résultats obtenus pour la classification des 34 communes et celle des 146 modalités. Ces résultats sont considérés respectivement aux §§3.2 et 3.3.

3.1 Les diverses constructions de la diagonale

Le tableau ci-dessus donne les éléments diagonaux attribués aux 12 premières communes, de 01, Le Prêcheur, à 12, Le Lamentin, en passant par la capitale, 11, Fort de France.

Le sigle 'év' renvoie au modèle utilisé dans [MARTINIQUE FLUX]: modèle à propos duquel on a, pour faire image, parlé d'évaporation. Dans ' ∂ ', le critère de la trace minima est appliqué une seule fois, comme pour reconstituer la diagonale d'une loi produit, sans ajouter à la diagonale de terme proportionnel à la marge. Le sigle ' $\partial\mu$ ' évoque, au contraire, les deux applications successives du critère.

Afin d'apprécier la précision du calcul de la loi marginale (cf. §2), on a donné deux variantes de ' $\partial\mu$ ': par interpolation de Σ avec 9 valeurs intermédiaires, ou en utilisant un pas de 1/100 (sigle '%'). Ces variantes ne diffèrent guère que si l'on considère l'expression exacte des valeurs calculées: ce qui n'a pas de signification dans notre tableau, où il s'agit du recensement de sujets individuels!

Tout ce qui apparaît à la lecture de ce tableau est que, de 'év' à ' ∂ ' puis à ' $\partial\mu$ ', les valeurs inscrites sur la diagonale croissent, tout en restant du même ordre de grandeur; qu'on comparera à celui de la population Tséd, sédentaire, i.e. recensée dans la commune deux fois de suite, en 1975 et en 1982.

Population par résidence et sexe au 9/3/82 & au 1/1/75

trace :	1.133e+0 ('év' : évaporation)									
rang :	1	2	3	4	5	6	7	8	9	10
lambda :	1974	1464	1162	973	743	680	590	521	434	414 e-4
taux :	1742	1292	1026	859	655	600	521	459	383	365 e-4
cumul :	1742	3034	4059	4918	5574	6174	6694	7153	7537	7902 e-4

Population par résidence et sexe au 9/3/82 & au 1/1/75

trace :	1.127e+0 ('ð' : diagonale créée sans terme de marge)									
rang :	1	2	3	4	5	6	7	8	9	10
lambda :	1952	1460	1126	932	741	687	625	532	469	446 e-4
taux :	1732	1295	1000	827	657	610	554	472	416	396 e-4
cumul :	1732	3027	4027	4854	5511	6121	6675	7147	7564	7960 e-4

Population par résidence et sexe au 9/3/82 & au 1/1/75

trace :	1.095e+0 ('ðμ%' : diagonale avec terme de marge; pas = 1/100)									
rang :	1	2	3	4	5	6	7	8	9	10
lambda :	2058	1540	1233	976	796	618	575	465	408	370 e-4
taux :	1880	1407	1126	892	727	565	525	425	373	338 e-4
cumul :	1880	3287	4413	5305	6032	6596	7122	7546	7919	8257 e-4

Dans la présente étude, la diagonale est calculée, non pour suppléer des données manquantes, mais afin de donner aux données disponibles une forme telle que les résultats de l'analyse factorielle et de la classification automatique s'interprètent clairement. C'est de ce point de vue qu'il convient de comparer les diverses méthodes employées.

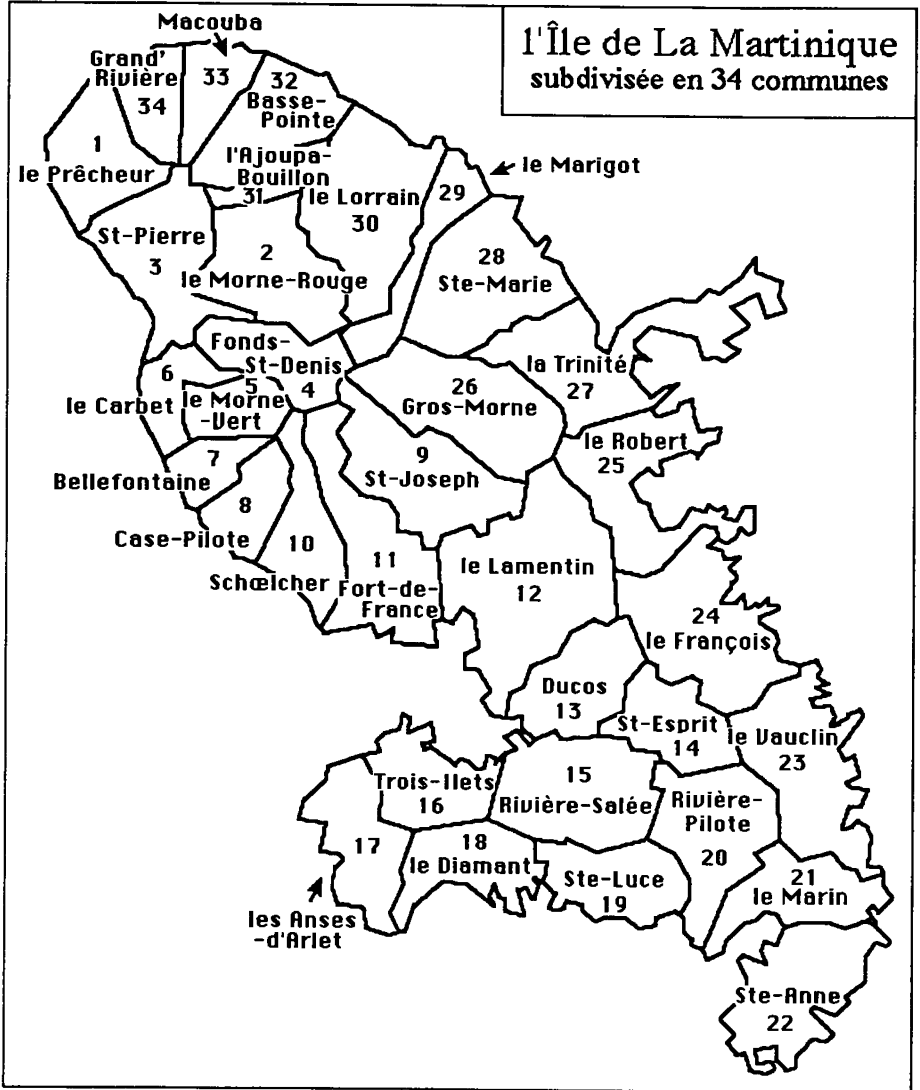
Le tableau (34 × 146) résulte d'une construction complexe, dont le calcul de la diagonale des flux n'est qu'une étape: on considère donc d'abord l'analyse d'un tableau carré (34 × 34). Bien qu'au §2, le critère de la trace minima soit appliqué par référence à un tableau symétrique, on donne ici le tableau des valeurs propres et taux issus de l'analyse du tableau dissymétrique des flux.

Entre les tableaux afférents à 'év' et à 'ð', (construction de la diagonale en une étape), il y a peu de différence. Mais de 'ð' à 'ðμ%', on voit que la trace diminue tandis que les valeurs propres et les taux de rang 1 à 5 augmentent: sans qu'on introduise d'inertie parasite, le cumul des 5 premiers taux passe de 55% à 60%. On peut attendre de telles modifications un progrès dans l'interprétation. Quant au choix d'un pas de 1/10 ou de 1/100 pour fixer la loi marginale, il n'a pas d'effet sensible.

3.2 Classification des communes

Dans ce §, comme dans le suivant, on se borne à considérer les versions 'ð' et 'ðμ%' de la construction de la diagonale; en comparant les résultats avec ceux déjà publiés pour 'év'.

La grande similitude des résultats ne justifie pas qu'on crée de nouvelles cartes, après celles publiées dans [MARTINIQUE FLUX]: mais afin d'interpréter aisément, en terme de cohérence spatiale, les différences entre variantes, on a reproduit le fond de carte de l'île, subdivisée en communes.



La partition en 9 classes, adoptée dans [MARTINIQUE FLUX], peut être ici retrouvée dans l'analyse 'd': ici comme là, le schéma le plus facile à interpréter sur la carte est obtenu en prenant des nœuds spécifiés appropriés; et non les 8 nœuds les plus hauts. De façon précise, avec 'd', on subdivise la classe 58, mais non les classes 59 et 60.

c | Partition en 9 classes : Sigles des communes de la classe c

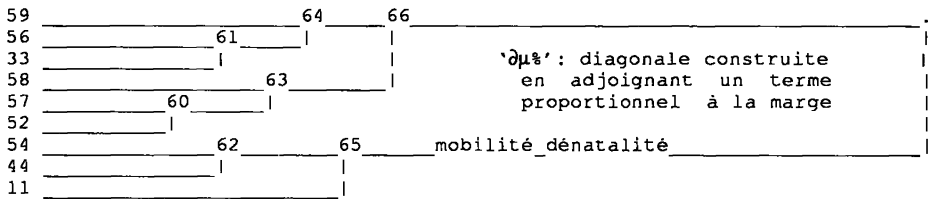
59	01	02	04	07	05	06	03	flux interne	264+++	274++++

56	32	31	29	34	30			flux interne	259+++	269++ 272+
33	33							->32,34	233+++++	272+

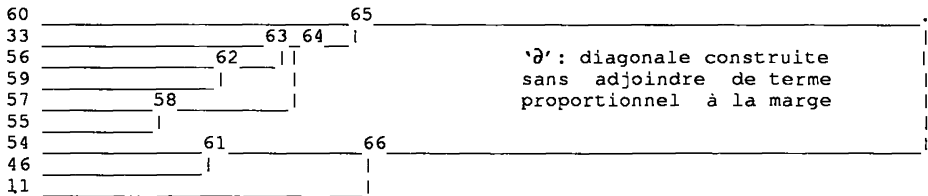
58	27	26	28	09	25			flux interne	239+	266+ 272+

57	22	21	20	23	14	24			268+	249+ 272+
52	19	18	16	17	15			flux interne	256+++	262+++

54	12	13							252+++	272- 270+++
44	08	10								272--- 270+++++
11	11									272-- 270+



Dans la version ' $\partial\mu\%$ ', on obtient, en retenant les 8 nœuds les plus hauts, une partition qui ne diffère de celle déjà vue qu'en ce que la commune 07 rejoint, opportunément, la région du Nord-Ouest. Ci-dessus, anticipant sur le §3.3, on a étiqueté l'arbre de la CAH en terme de classes de modalités.



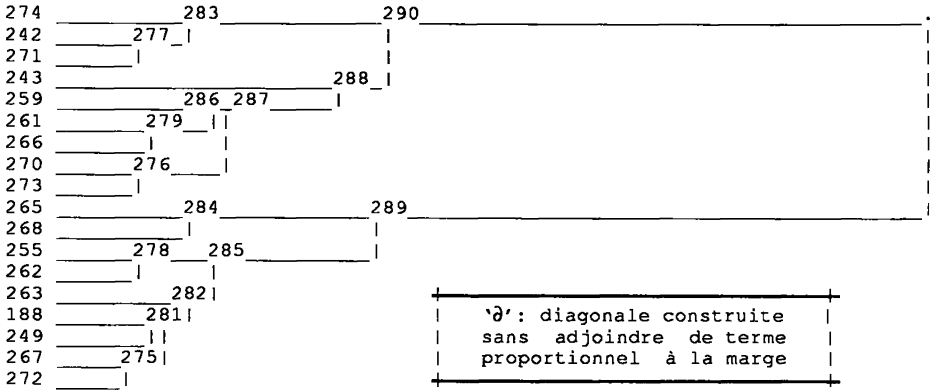
c | Partition en 9 classes: Sigles des communes de la classe c

60	01	02	04	05	06	03	

33	33						
56	31	29	34	30	32		
59	27	28	26	07	09	25	(07, Centre-N-Ouest, avec la côte C-N-Est!)

57	21	22	20	23	14	24	
55	17	15	19	18	16		

54	13	12					
46	08	10					
11	11						



c	Partition en 18 classes : Sigles des modalités de la classe c									
274	01&	∅04	04&	02&	∅01	∅02	05&	∅05	∅07	
242	06&	∅06								
271	∅03	03&								
-----290										
243	34&	32&								
-----288										
259	∅34	31&	33&	∅33						
-----286										
261	∅32	30&	∅30							
266	26&	∅26	28&	∅28	∅25	25&				
-----287										
270	29&	27&	∅27							
273	∅zz	∅sd								
=====										
265	24&	∅24	∅13	∅14	∅12					
268	∅10	∅08	∅11							
-----289										
255	17&	∅17	∅15	15&						
262	∅18	∅19	19&	14&	∅16	18&	16&			
-----285										
263	∅Ét	∅09	∅Fr							
-----282										
188	22&									
249	21&	∅21								
-----281										
267	∅22	20&	∅20	23&	∅23	13&				267-272 semble peu
272	∅31	∅29	11&	09&	12&	08&	07&	10&	∅ND	cohérent sur la carte

3.3 Classification des modalités des flux

Comme précédemment, on note $k[i, xj]$ le nombre des individus du sexe x , qui sont venus dans la commune i en provenance de j ; ce même nombre s'écrit encore $k[j, iX]$. Puisque les couples de modalités, $\{fi, hi\}$ ou $\{iF, iH\}$, s'agrègent à un très bas niveau, on les désigne par un sigle unique: $\&i$, ou $i\&$.

Les partitions présentées ici sont définies par les 17 nœuds les plus hauts.

264	_____	281	_____	59+++++	_____	290	_____	
274	_____		_____		_____		_____	
220	11--	_____	_____	_____	_____	_____	_____	
233	11--	_____	33+++++	_____	288	_____	_____	
259	_____	280	_____	286	_____		280: 56++++	
269	_____	_____	_____	_____	_____		269: 56++++	
239	_____	275	_____	285		239: 58+++++	275: 58++++	
266	_____	_____	_____	_____	_____		286: 11-, exceptés 269, 266	
272	58++	57+	_____	11--	_____		272: = cdg	
270	54++	44++++	_____	_____	_____		289	
256	_____	276	_____	284	_____	287		276: 52+++++
262	_____	_____	_____	_____	_____	_____	_____	
174	_____	283	_____	_____	_____		174: 57+++++	
249	_____	_____	_____	_____	_____		249: 57++++	
268	_____	_____	_____	_____	_____		268: 57+++	
252	_____	_____	_____	_____	_____		252: 54+++++	
248	44++	_____	282	_____	_____		282: 11+++	
273	11++++	_____	_____	_____	_____		_____	

'μ%': diagonale construite en adjoignant un terme proportionnel à la marge

c	Partition en 18 classes : Sigles des modalités de la classe c									
264	01&	04	04&	02&	05&	01	02			
274	06&	06	05	07	03			<->Nord-Ouest		
220	03&									
-----290										
233	34&	32&								
-----288										
259	34	31&	33F	33						
269	29&	31	29	30	32	33H	30&	<->Nord-Nord-Est		
-----286										
239	27&	27								
266	28&	28	26&	26	25	25&		<->Centre-Est-Nord		
-----285										
272	zz	hsd						Natalité et Sédentarité		
=====										
270	08	10	11					<-F de F		
-----289										
256	17&	17	f18	15	15&					
262	19	19&	14&	16	h18	16&	18&	<->Sud-Ouest		
-----284										
174	22&									
249	21&	21						Pointe-Sud		
-----283										
268	22	20&	20	23&	23	24&	24	<->Sud-Est		
252	13	14	12					<- Centre-Ouest		
-----287										
248	09	Fr								
273	ND	11&	13&	09&	12&	Et	07&	08&	10&	->Baie de F de F

Il y a, entre les deux versions des différences minimales: mais, ici encore, c'est en ajoutant à la diagonale un terme proportionnel à la marge qu'on obtient les classes dont la cohérence spatiale est la plus satisfaisante. L'arbre de la partition ainsi obtenue est étiqueté en terme de classes de communes (cf. §3.2); '¶' signale, sur le tableau, les couples, {33F, 33H} {f18, h18}, qui ne sont pas dans une même classe, mais dans des classes voisines; des flèches suggèrent un résumé des classes en terme de flux, entrant ou sortant, afférents à une région.

4 Conclusions et perspectives

Quant à l'analyse des flux migratoires, le présent article confirme les résultats du travail précédent: l'accord étant d'autant plus remarquable qu'on utilise ici une méthode mathématique générale qui supplée une diagonale sans autre donnée que celle des termes extradiagonaux d'une matrice carrée.

Quant à la méthode, ainsi validée, il apparaît préférable d'ajuster à la diagonale un terme complémentaire proportionnel à la marge d'abord reconstruite. Il reste souhaitable de traiter d'autres données: notamment des matrices de graphe, dont l'analyse, sans une diagonale bien choisie, n'aboutit pas à une typologie utile des sommets.

Références bibliographiques

K. Ben SALEM: "Le critère de la trace minima pour estimer les données manquantes dans un tableau de correspondance", [TRAC. MANQ.], in *CAD*, Vol XVII n°1, pp. 97-112; (1992).

M. OUASSOU, M.-M. THOMASSIN: "Flux de population à la Martinique entre 1975 et 1982: analyse des données et représentations sur la carte", [MARTINIQUE FLUX], in *CAD*, Vol XVI n°3, pp. 361-370; (1991).

J.-P. & F. BENZÉCRI, M.-M. THOMASSIN: "Utilisation par le logiciel MacSAIF de fonds de carte digitalisés créés par divers logiciels", [CARTE DIGIT.], in *CAD*, Vol XVI n°3, pp. 333-340; (1991).