MATHÉMATIQUES ET SCIENCES HUMAINES

G. TH. GUILBAUD

Bribes méthodologiques (I)

Mathématiques et sciences humaines, tome 20 (1967), p. 33-46
<http://www.numdam.org/item?id=MSH_1967__20__33_0>

G. Th. GUILBAUD

BRIBES METHODOLOGIQUES (I)

Avertissement

Ce qu'on va lire est une conséquence de conversations entre psychologues et mathématiciens (conversations qui ont eu lieu en Octobre 1967 au Centre Condorcet). Ce n'est pas un compte-rendu: on a cherché à extraire des propos tenus une esquisse d'exposé, d'où il soit facile de tirer quelques leçons mathématiques.

La psychologie, la sociologie, la démographie, l'économie et bien d'autres disciplines encore, font usage de modèles où un certain état de choses est représenté par une distribution:

$$p_1 + p_2 + \dots + p_n = 1 \quad , \quad p_i \geqslant 0$$

les p_i peuvent être des probabilités, représentant par exemple les chances d'une réponse i parmi n possibles. Ce peuvent être aussi des proportions d'individus d'une population, p_i pour ceux qui se trouvent dans l'attitude cotée i.

Une telle distribution est censée évoluer au cours du temps (processus) et le passage (transition) d'une distribution à une autre est figuré par une application linéaire.

Dans ce cadre très large (où l'on reconnaît les modèles markoviens, des modèles d'apprentissages, des modèles démographiques, et bien d'autres) on peut s'intéresser à certaines classes de distributions. Dans le cas des conversations citées, il s'agissait de relations d'<u>ordre</u>.

Par exemple, on sait que la réponse numérotée i est <u>toujours</u> plus probable que celle qu'on numérote j.

Ou bien: la population des individus dans l'état i est <u>toujours</u> supérieure à celle qui se trouve dans l'état j.

Une question posée était: que peut-on dire des transitions qui conservent de telles relations ?

La réponse tient en peu de mots.

Mais il a paru intéressant de donner à la consultation un tour plus didactique - et d'allonger le discours, à titre d'expérience, pour amorcer un dialogue entre l'enseignement de la mathématique et celui de ce qu'on appelle quelquefois la méthodologie. D'où le titre: lequel se veut prémice d'une rubrique régulière, et appel aux collaborations futures.

BRIBES METHODOLOGIQUES: (I)

POUVOIR DES SIMPLEXES

1. ENSEMBLES ORDONNES ET NOTATIONS SIMPLICIALES

1.1.- Une relation d'ordre, notée comme de coutume:

$$\text{(b est avant a)} \qquad b > a \tag{1}$$

peut être transmise en langage ensembliste. C'est d'ailleurs l'un des modes de transmission d'information assez couramment utilisé, même dans la conversation; on dit d'abord de quoi on va parler: l'ensemble /a,b/; et on dit, ensuite, lequel des deux est l'aîné, soit /b/.

Finalement la donnée de l'ensemble:

$$/\ /\ a,\ b\ /\ ,\ /\ b\ // \tag{2}$$

en dit autant que (1)

Remarque 1 : Bien entendu, on pourrait aussi bien dire: / / a,b / , /a/ / - mais il faut choisir: ici nous choisirons toujours la convention (2).

Remarque 2 : L'introduction de cette technique en mathématiques formalisées semble être due à: C. KURATOWSKI, sur la notion d'ordre dans la théorie des Ensembles (Fundamenta Math. 1921, p. 161).

1.2.- On peut généraliser: un ordre <u>total</u> tel que

$$d > c > b > a \tag{3}$$

sera décrit par l'ensemble d'ensembles (emboités):

$$/\ /a,b,c,d/\ ,\ /b,c,d/\ ,\ /c,d/\ ,\ /d/\ ,\ \emptyset\ / \tag{4}$$

(si l'on a mis la partie vide, c'est seulement pour faire joli).

1.3.- Ce n'est guère plus difficile quand il s'agit d'un ordre partiel. On doit sélectionner des parties de l'ensemble donné (et ordonné) telles que si: $a < b$, alors toute partie contenant a doit contenir b.

En d'autres termes:

$a < b$ sera traduit $a \Longrightarrow b$

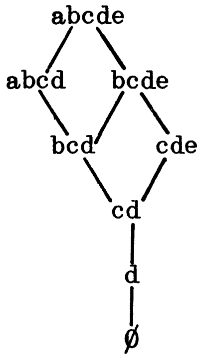
(la présence de a implique celle de b, ou autrement: pas de a sans b).

Exemple: l'ordre partiel : $a < b < c < d$
$\phantom{a < {}} e < c < d$

se traduira par le schéma d'implications:

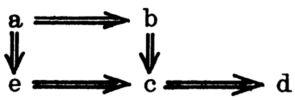
$$\begin{tikzcd} a \arrow[r, Rightarrow] & b \arrow[dr, Rightarrow] & & \\ & & c \arrow[r, Rightarrow] & d \\ & e \arrow[ur, Rightarrow] & & \end{tikzcd} \tag{5}$$

On en déduit la liste de <u>toutes</u> les parties de l'ensemble /a,b,c,d,e/ qui respectent ce schéma; il est commode de disposer ces parties en lattis:



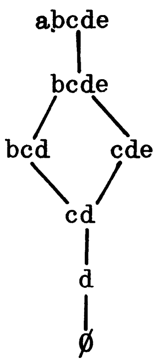
(6)

1.4.- Si, dans l'exemple précédent, on ajoute quelque liaison dans le schéma (5); pour fixer les idées, disons :



(5bis)

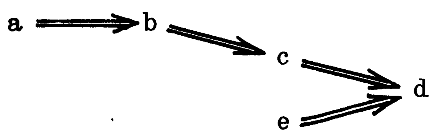
alors le lattis (6) devient:



(6bis)

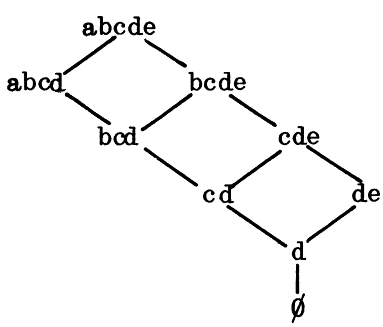
Inversement, si l'on enrichit (6), on vérifiera que l'on appauvrit (5).

Par exemple, à partir de



(5ter)

on construit



(6ter)

1.5.- Je laisse au lecteur le plaisir et le soin d'établir les propositions générales : A tout ordre correspond un lattis
Si l'ordre est total, on a une chaîne
Si l'ordre est vide, on a l'ensemble de toutes les parties (simplexe abstrait)
et de poursuivre l'étude amusante de cette correspondance.

2.- ECRITURES VECTORIELLES

2.1.- A toute partie d'un ensemble /a,b,c,d,e/ on peut faire correspondre la fonction indicatrice (dite encore caractéristique) ou vecteur indicateur (ou: indicateur, pour les intimes).

a	b	c	d	e
0	1	1	1	0

désigne: /b,c,d/

Au schéma (6) correspond le tableau :

a	b	c	d	e
1	1	1	1	1
1	1	1	1	0
0	1	1	1	1
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	0	0	1	0
0	0	0	0	0

2.2.- Si l'on appelle v l'un quelconque des indicateurs précédents, application de l'ensemble donné dans l'ensemble /0,1/, on vérifiera que la condition

$$b \Longrightarrow c$$

correspond à $v(b) < v(c)$
pourvu qu'on décide que : $0 < 0$ $1 < 1$ et $0 < 1$

Remarque 1 : les puristes, ici écrivent $\leqslant$

Remarque 2 : et les étudiants logiciens parlent de "Tables de vérité".

3.- DISTRIBUTIONS ET SIMPLEXES AFFINES

3.1.- Pour un statisticien, une application d'un ensemble donné (d'étiquettes) tel que:

$$E = /a,b,c,d,e/$$

dans un ensemble "numérique" (par exemple: les entiers naturels, ou bien les rationnels, ou encore: les réels) sera considérée comme une distribution. Soit:

	a	b	c	d	e	
(f)	0	17	4	31	28	(somme = 80)

ou bien :

	a	b	c	d	e	
(p)	0	22%	5%	38%	35%	(somme = 100%)

Pour le probabiliste, la coutume est de normer, c'est-à-dire de s'arranger pour que la somme soit égale à l'unité (encore que le langage familier n'ait pas abandonné l'expression "tant de chances sur tant", on constate que, de plus en plus, même les laïcs disent: "tant de chances sur cent", ce qui est une façon de normer). En réalité, il faut bien le dire, ce que les usagers ont l'intention de faire, c'est de confondre en une classe d'équivalence, toutes les distributions homothétiques entre elles: pour le statisticien, f et p ci-dessus diront la même chose.

Et, de même, les distributions

a	b	c	d	e
1	1	0	1	0

ou bien

a	b	c	d	e
1/3	1/3	0	1/3	0

ou bien

a	b	c	d	e
100	100	0	100	0

(il ne s'agira donc pas, à proprement parler, d'espace vectoriel, mais d'espace projectif).

3.2.- Une distribution normée sera donc une application

$$p : E \longrightarrow R$$

telle que

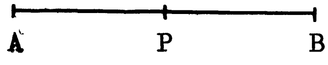
$$p(x) \geqslant 0$$

et

$$\sum_{x \in E} p(x) = 1$$

Si E n'a que deux éléments, à chaque distribution on fera correspondre un point P sur un segment

$AP = p$ $\qquad$ $PB = q$ $\qquad$ $p + q =$ l'unité AB

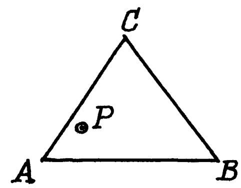


Si E possède trois éléments a,b,c, on figurera par trois points A,B,C, les trois distributions de base:

$$A \begin{pmatrix} a & b & c \\ 1 & 0 & 0 \end{pmatrix} \qquad B \begin{pmatrix} a & b & c \\ 0 & 1 & 0 \end{pmatrix} \qquad C \begin{pmatrix} a & b & c \\ 0 & 0 & 1 \end{pmatrix}$$

Toute autre distribution p est une combinaison linéaire convexe (coefficients positifs de somme unité).

$$P \begin{pmatrix} a & b & c \\ p & q & r \end{pmatrix} = pA + qB + rC$$



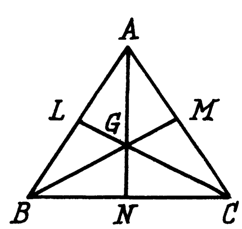
P est barycentre

Pour quatre points: tétraèdre, et au delà: ce que Poincaré appelait tétraèdre généralisé, et que l'on appelle aujourd'hui simplexe géométrique (ou affine, pour être plus précis).

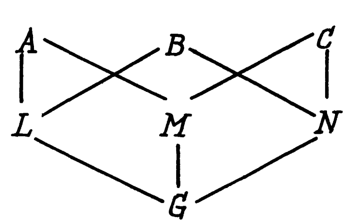
3.3.- Reprenons notre ensemble à trois éléments E = /a,b,c/ à chacune de ses parties on peut faire correspondre une fonction indicatrice.

	a	b	c
A	1	0	0
B	0	1	0
C	0	0	1
L	1	1	0
M	1	0	1
N	0	1	1
G	1	1	1

Si nous adoptons la convention du 3.1. qui considère comme équivalentes deux distributions homothétiques, on aura la figure:



(décomposition barycentrique: les milieux et le centre de gravité où se rencontrent les médianes).
Simplexe affine image du simplexe abstrait:



(il manque la partie vide, bien entendu)
Exercice: faire les dessins pour quatre éléments.

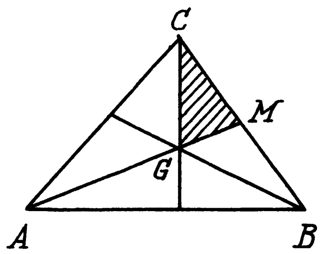
4.- STRUCTURES D'ORDRE ET DISTRIBUTIONS

4.1.- Soit un ensemble d'étiquettes $E = /a,b,c,d,\ldots/$ et une distribution p. Les valeurs p(a), p(b), p(c), etc... (qui sont des réels non-négatifs) sont appelées, selon l'opportunité, des poids (ou pondérations), des probabilités, ou ce qu'on voudra. Leur ordre peut être considéré comme une information essentielle dans le contexte traité. Par exemple, ce peut être intéressant de constater que:

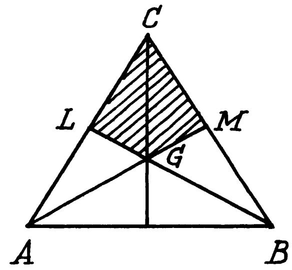
$$p(a) < p(b) < p(c) < \text{etc}\ldots$$

ou toute autre structure ordinale qu'on pourra imaginer. C'est à cette rhétorique là que se rattache la notion de "mode", les caractères: unimodal, bimodal, etc...

Notons alors ceci: toute structure d'ordre, totale ou partielle, concernant les p(x) désigne une partie du simplexe, et cette partie est convexe.



$p(a) < p(b) < p(c)$



$p(a) < p(c)$
$p(b) < p(c)$

La démonstration est facile: si p et q sont deux distributions satisfaisant au même système d'inégalités, il en va de même de toute combinaison linéaire $hp + kq$ à coefficients positifs.

Invoquons alors un principe bien connu de la géométrie affine: celui qui consiste à définir un convexe (borné) par des points extrêmes (ou sommets).

Ainsi toutes les distributions p qui vérifient:

$$p(a) < p(c) \quad \text{et} \quad p(b) < p(c)$$

peuvent être obtenues en combinant les extrêmes:

C	0	0	1
L	1	0	1
M	0	1	1
G	1	1	1

(combinaisons linéaires à coefficients positifs, bien entendu).

4.2.- La recherche des distributions extrêmes pour une structure d'ordre donnée, le lecteur l'aura déjà compris, n'est autre que le problème déjà traité aux paragraphes 1.2 et 1.3 ci-dessus.

Pour le cas d'un ordre total, c'est vraiment facile.
A l'ordre total

$$p(a) < p(b) < p(c) < p(d) \qquad (*)$$

correspond l'ensemble des extrêmes :

a	b	c	d
1	1	1	1
0	1	1	1
0	0	1	1
0	0	0	1

(scalogramme, comme on dit parfois),
ou bien, si l'on tient à normer :

1/4	1/4	1/4	1/4
0	1/3	1/3	1/3
0	0	1/2	1/2
0	0	0	1

qui sont respectivement solutions des équations :

$$0 < pa = pb = pc = pd$$
$$0 = pa < pb = pc = pd$$
$$0 = pa = pb < pc = pd$$
$$0 = pa = pb = pc < pd$$

(ci-dessus le système d'inéquations (*) aurait peut-être dû être écrit avec $\leqslant$ plutôt que $<$.

4.3. - Pour le cas général, on s'assurera que la recherche des extrêmes procède bien selon le même algorithme que dans les paragraphe 1. Pour cela, on notera d'abord que les extrêmes sont obligatoirement ou bien des sommets, ou bien des milieux d'arêtes, ou bien des centres de faces, etc.. Nous appellerons "centres" du simplexe affine les points dont les pondérations non nulles sont toutes égales. Il suffit alors de chercher quels arrangements de zéros et de uns conviennent à la structure d'ordre donnée. C'est-à-dire, très exactement, quelles parties de l'ensemble sont compatibles avec cet ordre: comme on a fait en 1.3. et 1.4. ci-dessus.

5. - TRANSITIONS MARKOVIENNES ISOTONES

5.1. - On sait que c'est A.A. Markov qui a introduit ce type de modèle probabiliste: transformation linéaire d'une distribution.

L'aspect proprement probabiliste sera ici passé sous silence; ce qui nous attachera, c'est le fait de considérer une transformation linéaire de l'objet appelé distribution.

A tout vecteur p on fait correspondre un vecteur q par:

$$q = Tp$$

et T est linéaire (c'est-à-dire sauvegarde la structure vectorielle).

Ce qui implique:

$$q(a) = T(aa)\,p(a) + T(ab)\,p(b) + \text{etc.}$$
$$q(b) = T(ba)\,p(a) + T(bb)\,p(b) + \text{etc.}$$
etc.

et l'on écrit T sous forme matricielle:

$$\begin{matrix} T(aa) & T(ab) & \dots \\ T(ba) & T(bb) & \dots \\ \dots & \dots & \dots \end{matrix}$$

5.2.- Toute transformation linéaire ne convient pas. Pour que Tp soit une distribution lorsque p l'est, il faut et suffit que la proposition soit vérifiée pour les extrêmes:

$$\begin{matrix} A: & 1 & 0 & 0 & 0 & \dots \\ B: & 0 & 1 & 0 & 0 & \dots \\ C: & 0 & 0 & 1 & 0 & \dots \\ D: & 0 & 0 & 0 & 1 & \dots \end{matrix}$$

c'est-à-dire que TA, TB, TC, ... soient des distributions, ce qui revient à exiger que leurs composants soient des nombres positifs. Il en résulte que tous les termes de la matrice T doivent être positifs.

Si l'on exige en plus la normalisation (somme des composantes égales à l'unité) il en résulte que les sommes des termes d'une même colonne doivent être égales à l'unité.

Tout cela est bien connu.

5.3.- Il peut être utile de savoir reconnaître, parmi les transitions markoviennes T, quelles sont celles qui conservent telle ou telle structure d'ordre.

Commençons par un cas simple (ordre total).

Supposons qu'on s'intéresse aux distributions telles que:

$$p(a) < p(b) < p(c) < p(d) \qquad (0)$$

L'ensemble de ces distributions est celui de toutes les combinaisons linéaires possibles des quatre extrêmes:

$$\begin{matrix} 1 & 1 & 1 & 1 & \text{ou} & A + B + C + D \\ 0 & 1 & 1 & 1 & \text{ou} & B + C + D \\ 0 & 0 & 1 & 1 & \text{ou} & C + D \\ 0 & 0 & 0 & 1 & \text{ou} & D \end{matrix}$$

Si p est une combinaison de ces quatre vecteurs, Tp sera la même combinaison de leurs transformés par l'opérateur T.

Il faut donc et suffit, que les quatre vecteurs

$$TA + TB + TC + TD$$
$$TB + TC + TD$$
$$TC + TD$$
$$\text{et} \quad TD$$

vérifient la condition d'ordre (0).

Mais on aura remarqué que TA, TB, TC et TD sont les quatre colonnes de la matrice T.

On peut donc exprimer la condition, cherchée de la façon suivante:

1) Ecrire la matrice T: (TA, TB, TC, TD)
2) Faire les sommes de colonnes
   (TA+TB+TC+TD, TB+TC+TD, TC+TD, TD)
3) vérifier que chaque colonne est monotone selon la condition (0).

5.4.- Le critère trouvé fournit une règle de construction, que nous décrivons au moyen d'un exemple numérique.

On va choisir successivement les vecteurs-colonne TD, puis TC, puis TB, enfin TA.

Nous normerons sur cent.

Prenons $$TD = \begin{pmatrix} 10 \\ 20 \\ 30 \\ 40 \end{pmatrix}$$

qui est bien monotone, comme il faut.

Il faut maintenant choisir une seconde colonne TC, qui n'a pas besoin d'être monotone, pourvu que la somme TC + TD le soit. Par exemple

$$TC = \begin{pmatrix} 25 \\ 20 \\ 25 \\ 30 \end{pmatrix} \quad \text{donne} \quad TC + TD = \begin{pmatrix} 35 \\ 40 \\ 55 \\ 70 \end{pmatrix}$$

et par conséquent convient.

Pour TB on peut prendre $\begin{pmatrix} 30 \\ 30 \\ 20 \\ 20 \end{pmatrix}$ car $TB + TC + TD = \begin{pmatrix} 65 \\ 70 \\ 75 \\ 90 \end{pmatrix}$

enfin $TA = \begin{pmatrix} 35 \\ 30 \\ 25 \\ 10 \end{pmatrix}$ convient encore puisque

$$TA + TB + TC + TD = \begin{pmatrix} 100 \\ 100 \\ 100 \\ 100 \end{pmatrix}$$

et que, nous avons décidé d'interpréter les inégalités au sens large (lisons $\leqslant$ au lieu de $<$)

(voir à ce sujet la remarque 1 du paragraphe 2.2).

Ainsi la matrice

$$T = \begin{pmatrix} .35 & .30 & .25 & .10 \\ .30 & .30 & .20 & .20 \\ .25 & .20 & .25 & .30 \\ .10 & .20 & .30 & .40 \end{pmatrix}$$

est de celles qui conservent la monotonie d'une distribution.

On peut se rendre compte de la façon dont elle modifie une distribution en calculant son effet sur quelques cas particuliers.

$$
\begin{array}{ll}
\bar{p} = (0,0,0,\ 100) & Tp = (10,20,30,40) \\
p = (5,10,25,60) & Tp = (17,21,28,34) \\
p = (15,20,25,40) & Tp = (22,23,26,29) \\
p = (20,25,25,30) & Tp = (24,24,25,27) \\
p = (25,25,25,25) & Tp = (25,25,25,25)
\end{array}
$$

5.5. - On peut présenter le même calcul autrement.

En partant de la matrice dont les colonnes sont respectivement:

$$TA + TB + TC + TD, \quad TB + TC + TD, \quad TC + TD\ , \quad TD$$

soit, dans notre exemple numérique

$$\begin{pmatrix} 100 & 65 & 35 & 10 \\ 100 & 70 & 40 & 20 \\ 100 & 75 & 55 & 30 \\ 100 & 90 & 70 & 40 \end{pmatrix} = M$$

c'est une matrice qu'on peut dire "monotone": les termes vont en croissant dans chaque colonne en descendant, et dans chaque ligne de droite à gauche.

Pour calculer la matrice T à partir de M, il suffit de retrancher chaque colonne de la précédente. Cette opération est une opération linéaire, on peut l'écrire en forme matricielle.

$$\begin{vmatrix} 100 & 65 & 35 & 10 \\ 100 & 70 & 40 & 20 \\ 100 & 75 & 55 & 30 \\ 100 & 90 & 70 & 40 \end{vmatrix} \times \begin{vmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 35 & 30 & 25 & 10 \\ 30 & 30 & 20 & 20 \\ 25 & 20 & 25 & 30 \\ 10 & 20 & 30 & 40 \end{vmatrix}$$

d'où la formule générale:

$$T = M\ .\ (I - J)$$

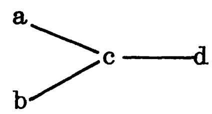
où l'on a posé:

$$I = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \qquad J = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

On a aussi $M = T\ (I + J + J^2 + J^3\ )$ parce que $J^4 = 0$

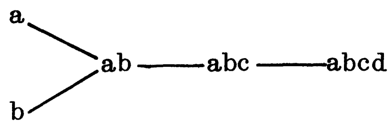
5.6. - Un autre exemple (ordre partiel).

$p\ (a) > p\ (c)$
$p\ (b) > p\ (c)$
$p\ (c) > p\ (d)$



(1)

traduction simpliciale:



(2)

Eléments extrêmes:

		a	b	c	d
A	=	1	0	0	0
B	=	0	1	0	0
A + B	=	1	1	0	0
A + B + C	=	1	1	1	0
A + B + C + D	=	1	1	1	1

(3)

L'opérateur T doit être tel que

TA , TB , TA + TB , TA + TB + TC , et TA + TB + TC + TD soient des distributions vérifiant la condition imposée (1). Il suffit de lire les colonnes de la matrice T.

Les deux premières colonnes TA et TB seront choisies arbitrairement et indépendamment, en respectant (1).

Par exemple:

25	25
45	30
20	20
10	15

TA + TB sera automatiquement du même type.

On choisira TC de façon à réaliser la condition (1) pour TA + TB + TC.

$$25 + 35 + p$$
$$45 + 30 + q$$
$$20 + 20 + r$$
$$10 + 15 + s$$

c'est-à-dire $35 + q > r$, $20 + p > r$, $15 + r > s$

Prenons par exemple:

20
15
35
40

(on voit que la 3ème colonne n'a pas besoin de vérifier (1): elle peut s'écarter du type d'ordre choisi, et la 4ème encore davantage).

Enfin pour la 4ème colonne, on calcule la somme TA + TB + TC + TD

$$25 + 35 + 20 + \ldots$$
$$45 + 30 + 15 + \ldots$$
$$20 + 20 + 35 + \ldots$$
$$10 + 15 + 40 + \ldots$$

et l'on choisit TD.

Comme précédemment rien n'empêche d'écrire en notation matricielle la transformation linéaire qui fournit la matrice (TA, TB, TC, TD) à partir de la matrice (TA, TB, TA + TB + TC, TA + TB + TC + TD). On trouvera encore une formule telle que:

$$T = M\,(I - K) \quad \text{avec } K^3 = 0$$

5.7. - Autres exercices: la condition d'ordre est de type unimodal, soit, par exemple:

$p(a) < p(b) < p(c) < p(d) > p(e) > p(f)$; on trouvera, pour le treillis simplicial correspondant une jolie forme, bien régulière. Les conditions sur les colonnes de T s'écrivent aisément.

*   *

*