MATRICE INTERGÉNÉRATIONNELLE ET ORDRE DE BRUHAT DANS LE GROUPE SYMÉTRIQUE

CHARLES TOROSSIAN

1. Introduction

Ce texte s'inspire largement de la thèse de Michel Martinez [3], ancien économiste à l'institut Rexecode [1] avec qui j'ai travaillé sur la démonstration exposée à la fin de cet article. L'essentiel des analyses économiques est issu directement de sa thèse.

L'objectif de cette séance de formation des inspecteurs stagiaires de la promotion 2010, est de montrer comment un problème d'économie sociale peut trouver une solution grâce à des mathématiques du domaine de l'algèbre, en l'occurrence la théorie des groupes. Ce qui nous intéresse ici, c'est donc d'analyser l'aller-retour intellectuel entre deux domaines assez éloignés ; détailler l'analyse de la problématique économique, expliciter des outils de modélisation adaptés. On tentera une interprétation mathématique en termes économiques et politiques et on montrera que le dialogue entre les différents champs scientifiques reste toujours fructueux.

2. Contexte économique selon la thèse de M. Martinez

Les utopies du monde occidental contemporain ont souvent avancé une société où chacun occuperait une position qu'il devrait à son seul mérite et non pas au milieu social dans lequel il est né. La mobilité sociale fascine depuis longtemps tous les observateurs de la société ; *La façon dont un homme peut s'élever dans la hiérarchie sociale et parvenir à la gloire, à la fortune, au pouvoir, est un des sujets les plus troublants pour les sociologues et des plus attirants pour le public* écrivait Alfred Sauvy, l'inventeur du mot *Tiers-Monde*.

La mobilité intergénérationnelle représente une entrée privilégiée de la sociologie et l'un des points forts de la discipline ; c'est aussi le cas en économie.

Dès le début des années 1960, aux Etats-Unis, le thème de l'égalité des chances a été l'objet d'importantes recherches empiriques. En France, ce thème a été introduit par Raymond Boudon (*L'Inégalité des chances*, 1973), partisan d'une explication de la mobilité sociale fondée sur les contraintes économiques qui s'exercent sur les familles, et Pierre Bourdieu et J.C. Passeron (*La Reproduction*, 1970) pour qui la sélectivité sociale est d'origine culturelle, inscrite au sein même de l'institution

---

Date: 15 Décembre 2010, Cours donné à l'ESEN aux IPR et IEN stagiaires.

1. Coe-Rexecode, Centre d'observation économique et de Recherche pour l'Expansion de l'économie et le Développement des Entreprises, 29 avenue Hoche 75008 Paris, www.coe-rexecode.fr .

scolaire. Clairement ces idées (ces croyances ?) sont encore vivantes.

Ce thème de l'égalité des chances est resté au centre des débats publics et fut considéré comme le principe essentiel de justice sociale : seul importait de réduire les freins à l'égalité des chances, ce qui justifiait qu'une place centrale soit accordée à l'école comme moyen d'atteindre l'idéal méritocratique. On peut constater que ces débats sont encore vifs aujourd'hui et qu'ils ont pris des formes plus variées ; on pourra penser aux diverses analyses du moment concernant l'accès des boursiers aux classes préparatoires aux grandes écoles.

Une question de fond est de savoir si la fluidité de la société française évolue. Sans entrer dans le débat de l'utilité d'une telle fluidité, on peut d'un point de vue méthodologique, distinguer trois approches de la mesure de la mobilité intergénérationnelle :

– l'approche statistique qui élabore de façon *ad hoc* des mesures de mobilité notamment à travers la matrice de transition intergénérationnelle à partir de laquelle on peut mesurer la vitesse de convergence du processus de Markov associé [2] ; une grande partie de la littérature sur la mesure de la mobilité intergénérationnelle se concentre exclusivement sur l'information contenue dans les matrices de transition. Toutefois il semble, et c'est la thèse soutenue par Michel Martinez, que *bonne approche de la mobilité intergénérationnelle doit s'efforcer de mesurer la mobilité relative.*
– une approche axiomatisante sur ces mesures de mobilité. Certains sociologues remarquent par exemple que la mobilité socio-économique sous-entend non seulement une notion de mouvement mais aussi une référence à la prédétermination des positions futures, compte tenu des positions occupées initialement (revenu ou éducation des parents par exemple).
– enfin une approche *welfariste* ou de *bien-être* dont l'hypothèse commune suppose que l'utilité marginale d'un euro de revenu pour les enfants issus de parents pauvres est supérieure à l'utilité marginale de ceux nés de parents plus riches. On peut évidemment discuter ce dernier point.

3. Outils mathématiques pour la mobilité

3.1. **La matrice de transition.** Il existe une vaste littérature sociologique sur la mobilité sociale considérée sous l'angle des catégories socioprofessionnelles (CSP).

Lorsqu'on étudie la mobilité intergénérationnelle, l'unité (du point de vue statistique) reste l'individu et l'on cherche à situer les individus par rapport à une origine sociale qu'ils doivent à leur famille. Habituellement, c'est le revenu ou la profession du père qui sert à caractériser le niveau social de la famille. Lorsqu'on rapproche des situations correspondant à des générations différentes, on parle de mobilité intergénérationnelle. En pratique, pour éviter des problèmes de générations, les individus que l'on compare sont supposés être du même âge.

L'approche par les matrices de transition est née de l'analyse des tableaux de contingence ou tableaux de mobilité. Les tableaux de mobilité mettent en relation

2. On verra au §3.3 des exemples de telles matrices.

		CSP du fils			
		A	B	C	Total
CSP père	A	$n_{1,1}$	$n_{1,2}$	$n_{1,3}$	$n_{1,+}$
	B	$n_{2,1}$	$n_{2,2}$	$n_{2,3}$	$n_{2,+}$
	C	$n_{3,1}$	$n_{3,2}$	$n_{3,3}$	$n_{3,+}$
	Total	$n_{+,1}$	$n_{+,2}$	$n_{+,3}$	n

FIGURE 1. Tableau de mobilité

les effectifs de l'origine et ceux de la destinée comme dans le tableau de Fig.1, où on n'a considéré que trois états sociologiques.

Parce que la notion de mobilité fait ainsi plus facilement sens, origines et destinées sont souvent représentées de manière identique. Par exemple, les CSP du père et du fils sont représentées par k mêmes catégories. Ces tableaux ont généralement deux dimensions $k \times k$.

En général, en sociologie, les positions initiales et finales correspondent à des positions sociales, communément, celles des pères et de leurs fils. Remarquons toutefois que les générations qui se trouvent aujourd'hui sur le marché du travail sont nées et ont grandi dans des milieux familiaux moins stables. C'est donc la notion même de classe sociale qui est remise en question. Remarquons aussi que les CSP ne sont plus l'alpha et l'omega et auraient à être remplacées par le triangle Temps-Culture-Argent mieux à même d'analyser l'impact sociétal.

Les marges de Fig.1 donnent la distribution des enquêtés en fonction de leurs origines et de leurs positions à l'enquête. On note $n_{i,+}$ l'effectif de la ligne i et $n_{+,j}$ l'effectif de la colonne j. Ce tableau d'effectifs permet de construire plusieurs tableaux de proportions. Ainsi

$$q_j = \frac{n_{+,j}}{n}$$

représente la proportion de la catégorie j parmi les enquêtés, tandis que $p_i = \frac{n_{i,+}}{n}$ représente la proportion des enquêtés qui ont un père dans la catégorie i.

Une transformation intéressante consiste à rapporter l'effectif de chaque case au total de la ligne. On note alors $p_{i,j}$ la probabilité que l'état j se réalise sachant l'origine i :

$$p_{i,j} = P(j|i) = \frac{n_{i,j}}{n_{i,+}}.$$

Le tableau dont les éléments sont les $p_{i,j}$ répond à la question ***de la destinée en fonction de l'origine***. C'est la probabilité que le fils évolue dans la catégorie j sachant que le père vivait dans la catégorie i.

Dans les approches descriptives utilisées en sociologie, c'est ce tableau carré, appelé matrice de transition $\mathbf{P}$, qui est utilisé pour répondre à l'interrogation sur *l'égalité des chances.*

3.2. **État limite.** Remarquons tout d'abord que les coefficients de la matrice $\mathbf{P}$ sont positifs et que
$$\sum_j p_{i,j} = \sum_j \frac{n_{i,j}}{n_{i,+}} = \frac{n_{i,+}}{n_{i,+}} = 1.$$

Supposons que la matrice intergénérationnelle soit constante au cours des générations. Cette hypothèse est évidemment absurde, mais elle sert ici d'alibi à une analyse mathématique intéressante.

Considérons la répartition en k catégories CSP d'une génération et posons nous la question de la répartition à la génération suivante. On dispose d'une matrice de transition $\mathbf{P}$ de taille $k \times k$.

Je note $X(0) = [X_1(0), \dots, X_k(0)]$ le vecteur de répartition, c'est à dire que $X_i(0) \geq 0$ et $\sum_i X_i(0) = 1$. Je dis qu'à la génération suivante la répartition est donnée par le vecteur ligne
$$X(1) = X(0) \cdot \mathbf{P},$$
produit de $X(0)$ avec la matrice $\mathbf{P}$.

Discussion 1. *Discuter cette affirmation.*

Indication : Quelle est la probabilité de l'évènement $[X(1) = j]$? On peut réfléchir à partir d'un arbre et on obtient
$$P([X(1) = j]) = \sum_i P([X(0) = i]) p_{i,j}.$$
□

On définit ainsi une suite $(X(n))_n$ par récurrence en posant $X(n+1) = X(n) \cdot \mathbf{P}$. Il s'agit de vecteurs de répartition (composantes positives de somme 1). On espère alors que la suite $(X(n))_n$ converge vers une répartition limite L. Un tel vecteur L est appelé état stable limite et on a $L = L \cdot \mathbf{P}$. Comme 1 est valeur propre de $\mathbf{P}$, c'est aussi une valeur propre pour $\mathbf{P}^t$ la transposée de $\mathbf{P}$, par conséquent il existe un état stable limite.

Bien-sûr un tel état limite n'est pas unique ; pensez à une matrice $\mathbf{P} = I_n$, pire il se peut que la suite ne converge pas ; pensez par exemple à une matrice de la forme $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Il est clair qu'il n'existe pas en général, pour cet exemple, d'état limite stable pour une donnée initiale différente [3] de $(1/2, 1/2)$.

Toutefois, si les données de $\mathbf{P}$ sont issues d'une analyse réelle, on peut espérer que 1 est valeur propre avec multiplicité 1 et que les autres valeurs sont en module strictement plus petites que 1, ce qui va assurer l'unicité de l'état limite.

Discussion 2. *Discuter cette affirmation.*

Indication : On peut trigonaliser la matrice, ou utiliser les projecteurs sur les espaces caractéristiques. Les puissances de la matrices P convergent (vers quoi ?).

---

3. Justifier cette affirmation.

3.3. **Deux exemples concrets de matrices de transition.** Illustrons notre propos de deux exemples.

Le premier tableau Fig.2 sont des données réelles pour la génération née en France entre 1920 et 1925 issues d'une enquête réalisée en 1982. Il s'agit de la matrice intergénérationnelle où les données sont exprimées en %.

		CSP du fils						
		Agrig.	Indé.	Cadre	Prof. Inter.	Employé	Ouv.	SP
CSP du père	Agriculteur	41,9	7,1	3,2	8,6	7,3	30,7	1,2
	Indépendant	2,9	28,9	15,1	19,8	9,2	23,1	1
	Cadre	1	8,8	53,5	22,1	6,3	6,6	1,7
	Prof. Intermédiaire	0,8	7,1	28,5	33,4	11,7	17,8	0,7
	Employé	1,7	8,5	17,1	25,2	17,9	28,5	1,1
	Ouvrier	3,1	8,2	5,6	17,9	11,3	53	0,9
	Sans Profession	13,6	11,1	11,4	17,5	10,3	35	1,1

FIGURE 2. Génération née en France entre 1920 et 1925

Le deuxième tableau Fig.3 est la matrice intergénérationnelle, issue d'une enquête effectuée en 2000 pour une génération de père nés en 1950 et 1955.

Comme on le constate ces matrices ne sont pas constantes au cours des générations (ici on a deux générations successives), ce qui met à mal l'hypothèse très simplificatrice faite plus haut.

		CSP du fils						
		Agrig.	Indé.	Cadre	Prof. Inter.	Employé	Ouv.	SP
CSP du père	Agriculteur	21,5	8,7	7,2	14,7	7, 6	39,6	0,7
	Indépendant	0,9	22,2	20,9	22	9,5	23, 6	0,9
	Cadre	0,8	7,8	51, 4	24, 4	7,5	7,1	1
	Prof. Intermédiaire	0,4	8,3	27,7	33	10,2	19,3	1,1
	Employé	0,5	7,6	17,5	27,2	16,7	29,4	1,1
	Ouvrier	0,6	7,9	7,4	20	11,2	42,9	10
	Sans Profession	3,9	9,8	15,7	22	10,8	36,9	0,9

FIGURE 3. Génération née en France entre 1950 et 1955

Discussion 3. *Discuter les différences notables sur ces deux matrices intergénérationnelles.*

Indication : L'agriculture n'est plus un débouché pour la génération issue de la seconde guerre mondiale.

Discussion 4. *Simulation de l'état stable sur ces deux matrices.*

Indication : Calculer les valeurs propres des matrices intergénérationnelle avec un logiciel (il s'agit quand même de matrice de taille 7×7). On approxime l'état stable en itérant la suite $X(n+1) = X(n) \cdot \mathbf{P}$.

Valeur propre 1	1.00000
Valeur propre 2	0.50864
Valeur propre 3	0.37679
Valeur propre 4	0.00051
Valeur propre 5	0.15136
Valeur propre 6	0.20350
Valeur propre 7	0.05617

FIGURE 4. Valeurs propres de la matrice intergénérationnelle des pères nés entre 1920 et 1925

J'ai choisi un calculateur sur Internet et j'ai trouvé pour la première matrice comme valeurs propres (tronquées à 5 décimales) le tableau Fig. 4

Elles sont de module plus petites que 1, ce qui était prévisible [4] et 1 est valeur propre ce qui était aussi prévisible. Il faut remarquer que les autres valeurs propres sont plus petites qu'environ $0,5$ ce qui assure une convergence rapide vers l'état stable donné par les valeurs en % dans le tableau Fig. 5.

Agriculteur	3,2%
Indépendant	10,2%
Cadre	25,1%
Prof. Intermédiaire	23,2%
Employé	10,5%
Ouvrier	26,6 %
Sans Prof	1,1%

FIGURE 5. Etat stable de la matrice intergénérationnelle des pères nés entre 1920 et 1925

.

Pour la deuxième matrice intergénérationnelle je trouve comme valeurs propres le tableau Fig.6.

Valeur propre 1	1.00000
Valeur propre 2	0.43744
Valeur propre 3	-0.04850
Valeur propre 4	0.20412
Valeur propre 5	0.10220
Valeur propre 6	0.05175
Valeur propre 7	0.13897

FIGURE 6. Valeurs propres de la matrice intergénérationnelle des pères nés entre 1950 et 1955

Les remarques sur les valeurs propres sont les mêmes. L'état stable est donné par le tableau Fig. 7.

4. Cela provient du fait que la somme des lignes vaut 1.

Agriculteur	0,9 %
Indépendant	9,3%
Cadre	27,0%
Prof. Intermédiaire	25,4 %
Employé	10,3%
Ouvrier	23,8 %
Sans Prof	3,2%

FIGURE 7. Etat stable de la matrice intergénérationnelle des pères nés entre 1950 et 1955

La conclusion est toutefois spectaculaire et c'est la force de l'outil matriciel ; ces calculs nous montrent que l'état stable des deux matrices est semblable (excepté pour l'agriculture [5]). Ainsi on peut conclure que malgré les changements apparents, la distribution stable n'a pas beaucoup évoluée.

3.4. **La mobilité parfaite.** La situation de référence pour juger d'une table de mobilité est la situation de mobilité parfaite, correspondant à une situation de totale fluidité, c'est à dire d'indépendance statistique entre origines et destinées. L'indépendance statistique entre destinée et origine signifie que les cellules du tableau des effectifs de mobilité sont telles que

$$N_{i,j} = \frac{n_{i,+}n_{+,j}}{n}.$$

Un tel tableau correspond donc à l'hypothèse de mobilité parfaite ou d'égalité des chances. Pour mesurer la distance à la mobilité parfaite, les sociologues utilisent les outils statistiques utilisés classiquement pour tester l'indépendance de deux variables qualitatives ou quantitatives. La mesure la plus usitée est le $\chi(2)$ de Pearson.

Discussion 5. *Rappeler comment on effectue un tel test.*

Indication : Si la matrice de effectifs est de taille 5×5 par exemple. On dispose donc de $4 \times 4 = 16$ degrés de libertés [6]. On calcule la matrice $N_{i,j}$ comme définie ci-dessus et on cherche à comparer selon un test de $\chi(2)$ la distance entre ces deux matrices. On fixe un seuil α, par exemple 5% ou 1% et on calcule la distance

$$\sum \frac{(N_{i,j} - n_{i,j})^2}{N_{i,j}}.$$

Si la distance calculée est supérieure à $\chi_\alpha(16)$ alors on rejette l'hypothèse de mobilité parfaite. □

4. La mobilité relative selon la thèse de M. Martinez

4.1. **Quelques éléments d'analyse économique.** Dans la pratique, les sociétés observées n'ont jamais une structure sociale identique, mais connaissent des différences ou des transformations sociales importantes. On doit donc se demander comment s'articulent les changements globaux de la société avec les changements de statut que connaissent les individus ou dynasties qui la composent, en général de père en fils. Il est inévitable que les transformations dans la répartition des différents

5. Remarquez le basculement entre Agriculture et Sans Profession.
6. Pour une matrice de taille $n \times p$ on disposerait de $(n-1) \times (p-1)$ degrés de libertés.

groupes professionnels ou de revenus correspondent à des cas de mobilité. Si beaucoup d'hommes nés en 1955-1960 n'appartiennent pas à la même classe sociale que leur père au même âge au contraire de la génération née en 1920-25, c'est essentiellement parce que la structure sociale a changé ; beaucoup moins d'agriculteurs, moins d'artisans et de commerçants, beaucoup plus de cadres et de professions intermédiaires.

D'autres phénomènes comme l'élévation du niveau moyen d'éducation, l'accroissement du nombre de cadres, la hausse du niveau de vie d'une génération à l'autre impliquent également un degré forcément non nul de mobilité, d'où le nom de mobilité *forcée ou mobilité structurelle.*

Toutefois l'échelle sociale n'est pas la même aujourd'hui qu'il y a trente ans. On voudrait distinguer de manière naturelle dans la mobilité des individus, le différentiel par rapport à la mobilité structurelle. Ainsi la mobilité *totale* peut s'analyser (se décomposer) en une mobilité structurelle, et une mobilité *nette ou relative.*

Notons que l'on peut observer quantité de mobilité structurelle, même en l'absence de changements dans la structure sociale (c'est-à-dire dans la distribution des positions sociales). Il en est ainsi des discordances entre structure scolaire et structure sociale.

Une autre critique du distingo entre mobilité relative et mobilité structurelle tient au lien existant entre la mobilité (relative) et l'efficacité productive.

Détaillons cette affirmation qui reste au cœur de la politique éducative du moment : toute amélioration de l'efficacité productive améliore la distribution finale des résultats et exerce donc un effet sur la mobilité structurelle. Autrement dit, lorsque la mobilité relative et l'efficacité économique sont endogènes, ce que supposent de nombreux modèles économiques, il en est de même pour la mobilité relative et la mobilité structurelle. Dans la plupart des circonstances, on peut s'attendre à ce que les freins à la mobilité désincitent les enfants talentueux à produire suffisamment d'effort, engendrant ainsi une mauvaise allocation des talents ; dans un tel cas, plus une société est mobile, plus le revenu moyen est élevé.

On peut cependant envisager des schémas de mobilité où la situation qui conduit à la plus forte efficacité productive est l'immobilité ; par exemple, si l'on suppose que les talents des enfants dépendent fortement de ceux de leurs parents et si le quotient intellectuel (QI) des enfants est hérité de celui des parents et si ce QI conditionne l'efficacité productive (c'est une hypothèse pour l'argument décrit ici et non une affirmation !), il existe alors un dilemme entre efficacité et mobilité. Ce n'est qu'en l'absence totale de corrélation entre les talents des parents et ceux de leurs enfants que ce dilemme disparaît et que la mobilité est nécessaire à une allocation efficace des talents.

Un autre cas de figure où les freins à la mobilité peuvent accroître la richesse nationale se produit lorsque certaines positions sociales, accessibles alors indépendamment de l'origine sociale, assurent des rentes économiques ; En présence de rentes, les individus les plus talentueux sont incités à se les approprier, délaissant d'autres activités économiques où leurs talents seraient socialement plus utiles. Au contraire, lorsque des freins à la mobilité - comme l'exclusivité de certaines positions à l'aristocratie ou à des castes- empêchent les individus talentueux d'accaparer ces rentes, ils deviennent socialement plus utiles et augmentent la richesse nationale.

Discussion 6. *Donner des exemples de telles situations dans l'histoire.*

Indication : La science soviétique, le développement de la bourgeoisie économique au 18-ème et 19-ème siècle en France.

D'une manière générale, si l'on suppose qu'il existe des liens endogènes entre les schémas de la mobilité et la distribution finale des caractéristiques étudiées, la distinction entre mobilité structurelle et mobilité relative devient alors délicate.

Michel Martinez, malgré ces critiques suppose que c'est la mobilité relative qui mérite intérêt. Il suggère par exemple que *la panne de l'ascenseur social* peut tenir à deux raisons ; la première est de nature relative et tiendrait au fait que les *privilégiés* profiteraient plus des changements économiques. La seconde serait d'ordre structurel. Dès lors que les progrès de productivité sont devenus moins importants avec le ralentissement économique depuis les trente glorieuses, la tendance du bien-être général s'est infléchie au point que les probabilités d'accès à une meilleure position sociale se sont dégradées, notamment pour les jeunes générations.

4.2. **Matrices bistochastiques.** On construit les matrices de *transition relative* selon la méthode suivante. Si la variable d'intérêt est ordinale (revenus, éducation...), on classera les individus par ordre croissant et un individu sera repéré par son n-quantile dans la distribution de cette variable.

Dans un tel cadre d'analyse, la matrice de transition $\mathbf{P}$ devient simplement une matrice bistochastique, dont la somme des lignes et des colonnes est égale à 1.

Une matrice M de taille n est dite bistochastiques si

$$\begin{cases} \forall j \quad \sum\limits_{1\leq i\leq n} m_{i,j} = 1, \\ \forall i \quad \sum\limits_{1\leq j\leq n} m_{i,j} = 1 \\ \forall i,j \quad m_{i,j} \geq 0 \end{cases} \tag{1}$$

Ces matrices correspondent à l'enveloppe convexe des matrices de permutations P_σ, où σ est une bijection de l'ensemble $[|1;n|]$. Rappelons que P_σ est la matrice dont les entrées sont nulles sauf aux termes $(\sigma(j), j)$ où elle vaut 1. Ce polytope est appelé polytope de Birkhoff et noté $\mathcal{B}$.

La matrice la plus fréquente est la *matrice de transition à quartiles.*

4.3. **Exemples de matrices de transition à quartiles.** Donnons quelques exemples de matrices de transition à quartiles, issues de véritables enquêtes :

4.4. **Transformations élémentaires.** Toute transformation sur la matrice de transition relative, issue par exemple d'une politique économique précise, doit laisser invariant le caractère bistochastique de la matrice de transition $\mathbf{P}$.

On suppose que nos matrices sont de taille $n \times n$.

		Fils			
		Q1	Q2	Q3	Q4
Père	Q1	32,9	31,2	21,4	14,5
	Q2	37,9	22,3	25,1	14,7
	Q3	18,6	34,1	22,7	24,6
	Q4	10,6	12,4	30,8	46,2

FIGURE 8. Matrice intergénérationnelle en Italie en 1987

		Fils			
		Q1	Q2	Q3	Q4
Père	Q1	28,5	27,9	21,4	22,2
	Q2	29,7	24,2	23,4	22,7
	Q3	23,4	25,8	25,6	25,2
	Q4	18,4	22,1	29,6	29,9

FIGURE 9. Matrice intergénérationnelle au Canada en 1986

		Fils			
		Q1	Q2	Q3	Q4
Père	Q1	39,2	25,5	21,6	13,7
	Q2	34	27,5	22,8	15,7
	Q3	16,7	30,8	25,5	27
	Q4	10,1	16,2	30,1	43,6

FIGURE 10. Matrice intergénérationnelle aux USA en 1987

		Fils			
		Q1	Q2	Q3	Q4
Père	Q1	38	25	22	15
	Q2	30	29	22	19
	Q3	19	29	27	25
	Q4	13	17	29	41

FIGURE 11. Matrice intergénérationnelle au Royaume-Uni en 1996

		Fils			
		Q1	Q2	Q3	Q4
Père	Q1	26	28,4	26,4	19,2
	Q2	29	25,7	23,4	21,9
	Q3	25,4	21,3	27,1	26,2
	Q4	19,6	24,6	23,1	32,7

FIGURE 12. Matrice intergénérationnelle au Canada en 1994

Les matrices les plus simples qui pourraient opérer sur les matrices bistochastiques sont par exemple les matrices $E_{i,j,k,l}(\epsilon)$ avec $i < j \leq n$ et $k < l \leq n$ dont les coefficients sont tous nuls sauf quatre d'entre eux

$$e_{i,k} = -\epsilon, \;\; e_{i,l} = \epsilon, \;\; e_{j,k} = \epsilon, \;\; e_{j,l} = -\epsilon,$$

$$E_{i,j,k,l}(\epsilon) = \begin{array}{l|cc|} \text{position} & k & l \\ \hline i & -\epsilon & \epsilon \\ j & \epsilon & -\epsilon \\ \hline \end{array}$$

On peut penser qu'une transformation élémentaire

$$\mathbf{P} \to \mathbf{P} + E_{i,j,k,l}(\epsilon),$$

résulte d'une politique économique concrète ciblée sur les classes i, j des pères et k, l des fils. Par exemple on peut augmenter les droits d'entrée à l'Université en fonction des salaires des pères.

Discussion 7. *Interpréter le cas $\epsilon > 0$ et $\epsilon < 0$*

Indication : Si les quartiles correspondent à des revenus croissants, par exemple, on peut considérer qu'une transformation à $\epsilon > 0$ a augmenté la fluidité intergénérationnelle, car la probabilité pour le fils d'être dans la classe k a diminué et celle d'être dans la classe l a augmenté. □

Lorsque ϵ est positif, une telle transformation est dite ***dédiagonalisante***.

4.5. **Indice de mobilité.** On souhaite définir, à partir des matrices de transition à quartiles, un indice de mobilité relative $M(\mathbf{P})$, où $\mathbf{P} \mapsto M(\mathbf{P})$ est une fonction à valeurs réelles définie sur le polytope des Birkhoff.

Suivant une tradition héritée de la sociologie pour laquelle la probabilité de mouvement entre classes sociales est donnée par la somme des éléments situés en dehors de la diagonale, on peut légitimement penser qu'une transformation dédiagonalisante accroît la mobilité. Ainsi on souhaite que pour $\epsilon > 0$, $i < j$ et $k < l$ on ait (tant que l'on reste dans le polytope $\mathcal{B}$) :

$$M(\mathbf{P} + E_{i,j,k,l}(\epsilon)) \geq M(\mathbf{P}). \tag{2}$$

Par ailleurs, la matrice $\mathbf{P} = I_n$, relève de l'immobilité totale, tant dis que la matrice $\frac{1}{n}\mathbf{1}_n = \begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix}$ dont toutes les entrées valent $\frac{1}{n}$ représente une mobilité parfaite, car le devenir des fils est indépendant de la condition des pères. On demande donc que l'indice à construire vérifie

$$M(\mathbf{P}) \geq M(I_n), \tag{3}$$

et

$$M(\mathbf{P}) \leq M(\frac{1}{n}\mathbf{1}_n). \tag{4}$$

Exemple d'un tel indice : $M(\mathbf{P}) = \frac{n}{n-1}\big(1 - \frac{1}{n}\sum p_{i,i}\big)$.

Cette mesure, largement utilisée en sociologie, dénombre la proportion d'individus qui changent de classe sociale d'une génération à l'autre. Il existe bien d'autre indices, basés par exemple sur les valeurs propres de la matrice **P** ou utilisant l'entropie de **P**.

5. Comparaison de matrices bistochastiques et marches positives

On cherche à privilégier maintenant l'utilisation de matrices dédiagonalisantes. En effet on imagine que l'action de telles matrices illustre une politique de discrimination positive de nature sociale. On cherche donc un critère de lecture sur les matrices intergénérationnelles relatives qui assure que l'on puisse suivre une telle politique.

Une matrice **P** est jugée plus mobile qu'une matrice **Q** si elle peut être obtenue à partir d'un nombre fini de transformations dédiagonalisantes.

En toute logique, c'est vers des résultats de dominance que doit tendre une théorie de la mesure de la mobilité ou de l'égalité des chances. Faute de quoi, il faut s'attendre à ce que plusieurs indices de mobilité classent différemment deux sociétés données, sans que l'on parvienne clairement à identifier les raisons de telles différences de hiérarchisation.

Dans une telle optique, la matrice de transition

$$\begin{pmatrix} 0 & \dots & 0 & 1 \\ \vdots & & \iddots & 0 \\ 0 & \iddots & & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}.$$

va jouer le role de la matrice la plus favorable ; cette situation paradisiaque [7] peut sans doute se discuter, notamment par rapport à une mesure de mobilité telle qu'on l'a définie précédemment.

5.1. Matrice cumulante.

Matrice cumulante. Définissons la matrices cumulées d'une matrice M par $T(M) = (t_{i,j})$ où on a les relations

$$t_{i,j}(M) = \sum_{\substack{1\leq k\leq i \\ 1\leq l\leq j}} m_{k,l}.$$

On définit maintenant une relation d'ordre les matrices bistochastiques :

$$M \leq_{\mathcal{B}} N \Longleftrightarrow T(M) \geq T(N),$$

où la relation d'ordre partielle pour des matrices $A = (a_{i,j}) \geq B = (b_{i,j})$ signifie $\forall 1 \leq i, j \leq n,\ a_{i,j} \geq b_{i,j}$.

La relation d'ordre induite sur les matrices de permutations définit une relation d'ordre partiel sur le groupe symétrique S_n, des bijections de l'ensemble $\{1, \cdots, n\}$.

7. Les premiers seront les derniers.

La matrice associée à la permutation $w_o = (n, n-1, \cdots, 2, 1)$ est

$$P_{w_o} = \begin{pmatrix} 0 & \dots & 0 & 1 \\ \vdots & & \iddots & 0 \\ 0 & \iddots & & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}.$$

C'est la plus grande pour l'ordre défini ci-dessus tandis que la matrice la plus petite est la matrice identité I_n :

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & \dots & 1 \end{pmatrix}.$$

On peut montrer que cet ordre correspond à l'ordre de Bruhat dans le groupe symétrique S_n [2].

5.2. **Marches positives ou théorème de la discrimination positive.** Cette partie s'inspire de description de marches positives pour l'ordre de Bruhat que l'on trouve dans l'article de P. Magyar [2]. L'adaptation à ce problème économique est dûe à l'auteur de cette note.

On veut montrer que si M, N sont deux matrices bistochastiques avec

$$M \leq_{\mathcal{B}} N$$

alors il existe une marche finie de M vers N dans le polytope $\mathcal{B}$ par des mouvements correspondant à des translations par des matrices $E_{i,j,k,l}(\epsilon)$ $(\epsilon > 0)$ avec $i < j$ et $k < l$. Ce résultat est le *théorème de la discrimination positive* de Michel Martinez.

Rappelons que la matrice $E_{i,j,k,l}(\epsilon)$ a ses coefficients tous nuls sauf quatre d'entre eux

$$e_{i,k} = -\epsilon, \;\; e_{i,l} = \epsilon, \;\; e_{j,k} = \epsilon, \;\; e_{j,l} = -\epsilon,$$

$$E_{i,j,k,l}(\epsilon) = \begin{array}{l|cc|} \text{position} & k & l \\ \hline i & -\epsilon & \epsilon \\ j & \epsilon & -\epsilon \\ \hline \end{array}.$$

Ces matrices ont des cumulées de la forme :

$$T(E_{i,j,k,l}(\epsilon)) = \begin{array}{l|ccccccc|}
\text{position} & & k & & & l-1 & l & \\
\hline
 & & \vdots & & & & \vdots & \\
i & \cdots & -\epsilon & -\epsilon & \cdots & -\epsilon & 0 & \cdots \\
 & & \vdots & & & & \vdots & \\
j-1 & & -\epsilon & -\epsilon & \cdots & -\epsilon & 0 & \\
j & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots \\
 & & \vdots & & & & \vdots & \\
\hline
\end{array}$$

Elles ont tendance à diminuer par blocs la valeur des coefficients des matrices cumulées.

Introduisons un ordre partiel $\leq_{par}$ sur les couples (i,j) défini par :

$$(i,j) \leq_{par} (k,l) \Longleftrightarrow i \leq j \text{ et } k \leq l.$$

Par rapport aux positions sur un tableau $n \times n$, le couple (i,j) est NO par rapport à (k,l).

Rappelons ce qu'est l'ordre lexicographique sur les couples, noté $\leq_{lexi}$:

$$(i,j) \leq_{lexi} (k,l) \Longleftrightarrow (i < k) \text{ ou } (i = k \text{ et } j \leq l).$$

Enfin si $\leq$ est une relation d'ordre (quelconque) on note $<$ la relation définie par

$$x < y \Longleftrightarrow x \leq y \text{ et } x \neq y.$$

5.3. **Description de la procédure.** On peut dans un premier temps sauter cette section qui aborde la démonstration du théorème de la discrimination positive.

On part de $M <_{\mathcal{B}} N$, c'est à dire $T(M) \geq T(N)$ et $T(M) \neq T(N)$. On a noté $t_{i,j}(M)$ les coefficients des matrices cumulées. Comme les coefficients des cumulées qui se trouvent en dernière ligne ou en dernière colonne ont des valeurs fixées, les conditions significatives pour la relation d'ordre sont pour $1 \leq i \leq n-1$ et $1 \leq j \leq n-1$

$$t_{i,j}(M) \geq t_{i,j}(N),$$

dont l'une de ces inégalités est stricte.

La procédure doit veiller à garder les coefficients des matrices positifs à chaque étape et faire en sorte que ces inégalités deviennent de plus en plus des égalités.

5.3.1. *Etape 0.* On suppose $M \neq N$.

5.3.2. *Etape 1 : choix d'une position minimale.* Considérons l'ordre lexicographique pour les positions (k,l). Choisissons la première position (i_0, j_0) dans l'ordre lexicographique, pour laquelle on a $t_{i,j}(M) > t_{i,j}(N)$. C'est donc la première position où les matrices cumulées diffèrent. Cette position est donc en dehors de la dernière ligne et en dehors de la dernière colonne.

Pour tout $1 \leq i < i_o, 1 \leq j \leq n$ et pour $i = i_0, 1 \leq j < j_0$, on a donc

$$m_{i,j} = n_{i,j}.$$

Par ailleurs on a

$$m_{i_0,j_0} - n_{i_0,j_0} = t_{i_0,j_0}(M) - t_{i_0,j_0}(N) > 0, \tag{5}$$

et donc $m_{i_0,j_0} > n_{i_0,j_0}$ et $m_{i_0,j_0} > 0$.

5.3.3. *Etape 2 : choix d'un rectangle SE.* Considérons un rectangle de positions $[i_0, p] \times [j_0, q]$ en position SE avec $i_0 < p \leq n$ et $j_0 < q \leq n$ (on rappelle que l'on a $i_0, j_0 \leq n - 1$), pour lequel on a

$$t_{i,j}(M) > t_{i,j}(N), \quad \forall\ (i,j) \in [i_0, p-1] \times [j_0, q-1].$$

Choisissons un tel rectangle maximal, c'est à dire que tout rectangle de position SE le contenant strictement ne vérifie pas la condition imposée. Cela revient à prendre la position du coin SE de ce rectangle, maximale pour l'ordre $\leq_{par}$. Un tel rectangle maximal n'est pas forcément unique car la relation d'ordre est partiel.

5.3.4. *Etape 3 : choix d'une position non nulle dans le rectangle SE décalé.* On considère le rectangle SE décalé

$$[i_0 + 1, p] \times [j_0 + 1, q].$$

Montrons qu'il existe une position (i, j) dans ce rectangle pour laquelle on a $0 < m_{i,j}$.

Supposons le contraire, c'est à dire $\forall (i,j) \in [i_0 + 1, p] \times [j_0 + 1, q]$ on a $m_{i,j} = 0$.

Comme le rectangle de l'étape 2 est maximal, tout rectangle SE le contenant strictement, ne vérifie pas la condition imposée. Donc il existe (p, a) avec $j_o \leq a \leq q - 1$ et il existe (b, q) avec $i_0 \leq b \leq p - 1$ pour lesquels on a

$$\begin{cases} t_{p,a}(M) = t_{p,a}(N) \\ t_{b,q}(M) = t_{b,q}(N) \end{cases} \tag{6}$$

(rappelons que la matrice cumulée de M domine celle de N).

$$M = \begin{array}{l|cccccccc|}
\text{position} & & j_0 & j_0+1 & & a & & q-1 & q \\
\hline
 & & \vdots & \vdots & & & & \vdots & \vdots \\
i_0 & \cdots & m_{i_0,j_0} & m_{i_0,j_0+1} & \cdots & & \cdots & m_{i_0,q-1} & m_{i_0,q} \\
i_0+1 & & m_{i_0+1,j_0} & 0 & \cdots & 0 & \cdots & 0 & 0 \\
 & & \vdots & \vdots & & & & \vdots & \vdots \\
b & & 0 & \cdots & & 0 & \cdots & 0 & 0 \\
 & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
p-1 & & m_{p-1,j_0} & 0 & \cdots & 0 & \cdots & 0 & 0 \\
p & \cdots & m_{p,j_0} & 0 & \cdots & 0 & \cdots & 0 & 0 \\
 & & \vdots & & & & & & \\
\hline
\end{array}$$

La position (b, a) est dans le rectangle de stricte dominance de l'étape 2. On a donc

$$t_{b,a}(M) > t_{b,a}(N)$$

Par ailleurs on a

$$t_{p,q}(M) = t_{p,a}(M) + t_{b,q}(M) - t_{b,a}(M) + \sum_{\substack{b+1 \leq i \leq p \\ a+1 \leq j \leq q}} m_{i,j},$$

d'où l'on tire $t_{p,q}(M) = t_{p,a}(M) + t_{b,q}(M) - t_{b,a}(M)$ à cause de la nullité du dernier terme.

Par ailleurs on a

$$t_{p,q}(N) = t_{p,a}(N) + t_{b,q}(N) - t_{b,a}(N) + \sum_{\substack{b+1 \leq i \leq p \\ a+1 \leq j \leq q}} n_{i,j},$$

d'où l'on tire $t_{p,a}(N) + t_{b,q}(N) - t_{b,a}(N) \leq t_{p,q}(N)$ à cause de la positivité du dernier terme.

Des équations (6) et de $t_{b,a}(M) > t_{b,a}(N)$ on déduit

$$\begin{aligned} (7) \quad t_{p,q}(M) = t_{p,a}(M) + t_{b,q}(M) - t_{b,a}(M) = \\ t_{p,a}(N) + t_{b,q}(N) - t_{b,a}(M) < t_{p,a}(N) + t_{b,q}(N) - t_{b,a}(N) \leq t_{p,q}(N), \end{aligned}$$

ce qui est absurde.

Il existe donc une position non nulle dans ce rectangle SE décalé. On choisit une position la plus NO, c'est à dire minimale pour l'ordre $\leq_{par}$. On a donc

$$m_{i,j} = 0 \ \ \forall (i,j) \ \in [i_0+1, i_1] \times [j_0+1, j_1], (i,j) \neq (i_1, j_1).$$

Une telle position n'est pas forcément unique car l'ordre $\leq_{par}$ n'est que partiel.

5.3.5. *Etape 4 : Choix du mouvement autorisé maximal.* On dispose de deux positions (i_0, j_0) et (i_1, j_1) pour lesquelles les entrées de la matrice M sont non nulles.

On consitère alors la matrice de déplacement autorisée

$$E_{i_0,i_1,j_0,j_1}(\epsilon).$$

La matrice $M + E_{i_0,i_1,j_0,j_1}(\epsilon)$ est dans le polytope $\mathcal{B}$ pour ϵ petit car on a m_{i_o,j_o}, m_{i_1,j_1} strictement positifs et m_{i_o,j_1}, m_{i_0,j_1} strictement plus petits que 1.

Comme le rectangle de l'étape 2 est strictement dominant, la matrice cumulée $T(M + E_{i_0,i_1,j_0,j_1}(\epsilon))$

$$\begin{array}{l|ccccccc}
\text{position} & & j_0 & & & j_1 - 1 & j_1 & \\
\hline
 & & \vdots & & & & \vdots & \\
i_0 & \cdots & t(M)_{i_0,j_0} - \epsilon & \cdots & \cdots & t(M)_{i_0,j_1-1} - \epsilon & t(M)_{i_0,j_1} & \cdots \\
 & & \vdots & & & & \vdots & \\
 & & \vdots & & & & \vdots & \\
i_1 - 1 & & t(M)_{i_1-1,j_0} - \epsilon & \cdots & \cdots & t(M)_{i_1-1,j_1-1} - \epsilon & t(M)_{i_1-1,j_1} & \\
i_1 & \cdots & t(M)_{i_1,j_0} & \cdots & \cdots & t(M)_{i_1,j_1-1} & t(M)_{i_1,j_1} & \cdots \\
 & & \vdots & & & & \vdots &
\end{array}$$

domine encore $T(N)$ pour ϵ petit. On choisit alors

$$t_0 = \max\{\epsilon, \ M + E_{i_0,i_1,j_0,j_1}(\epsilon) \in \mathcal{B}, M + E_{i_0,i_1,j_0,j_1}(\epsilon) \leq_{\mathcal{B}} N\}.$$

On note $m_{i,j}(t_0)$ les coefficients de la matrice $M(t_0) = M + E_{i_0,i_1,j_0,j_1}(t_0)$.

La signification du temps t_0 est que l'une des trois possibilités suivantes est vérifiée :

(1) il existe $(i,j) \in [i_0, i_1 - 1] \times [j_0, j_1 - 1]$ pour lequel on a
$$t(M(t_0))_{i,j} = t(M)_{i,j} - t_0 = t(N)_{i,j}$$
(2) $m_{i_1,j_1}(t_0) = m_{i_1,j_1} - t_0 = 0$

(3) $m_{i_0,j_0}(t_0) = m_{i_0,j_0} - t_0 = 0.$

Remarquons maintenant que l'on a (équation (5))
$$(m_{i_0,j_0} - t_0) - n_{i_0,j_0} = t_{i_0,j_0}(M(t_0)) - t_{i_0,j_0}(N) \geq 0$$
ce qui montre que la troisième possibilité ne peut pas se produire avant que la première possibilité ne soit vérifiée.

De même si le coefficient m_{i_1,j_0} (resp. m_{i_0,j_1}) vaut 1 alors le coefficient m_{i_0,j_0} est nul (resp. m_{i_1,j_1}). Donc il est inutile de discuter sur ces valeurs.

5.3.6. *Etape 5 : retour de la procédure.* On procède dans cet ordre :

1– On a l'égalité $m_{i_0,j_0}(t_0) = n_{i_0,j_0}$:
$\longrightarrow$

On retourne à l'étape 1. Les matrices $M(t_0)$ et N ont des coefficients égaux pour toutes les positions lexicographiques $\leq_{lexi}$ (i_o, j_0). Au bout d'un nombre fini de boucles de ce type on a $M = N$. Le processus est alors fini.

SINON

2–Il existe $(i,j) \in [i_0, i_1 - 1] \times [j_0, j_1 - 1]$, $(i,j) \neq (i_0, j_0)$ avec $t_{i,j}(M(t_0)) = t_{i,j}(N)$:
$\longrightarrow$

On retourne à l'étape 2 avec la matrice $M(t_0)$ (on saute l'étape 1 car les matrices M et $M(t_0)$ ont les mêmes positions minimales (i_0, j_0)).

Le nouveau rectangle de positions de l'étape 2 reste dans $[i_0, n] \times [j_0, n]$ et le coin SE ne doit pas se situer dans le rectangle de positions $[i_1, n] \times [j_1, n]$, sinon il contiendrait le rectangle $[i_0, i_1 - 1] \times [j_0, j_1 - 1]$ censé vérifier la condition de stricte dominance, ce qui n'est pas vrai car on a
$$(i,j) \in [i_0, i_1 - 1] \times [j_0, q - 1].$$
Au bout d'un nombre fini de boucles de ce type, le rectangle de positions de l'étape 2 se réduit à $[i_0, i_0 + 1] \times [j_0, j_0 + 1]$ car la surface autorisée diminue strictement à chaque étape. On est alors sorti de la boucle car la condition ci-dessus correspond à une condition de type 1.

SINON

3–Le coefficient $m_{i_1,j_1}(t_0)$ s'annule :
$\longrightarrow$

On retourne à l'étape 3 pour la matrice $M(t_0)$ (on saute l'étape 1 et l'étape 2 car les matrices M et $M(t_0)$ ont même position minimale (i_0, j_0) et même rectangle de positions SE $[i_0, p] \times [j_0, q]$).

Notre sous-matrice SE décalée a un zéro de plus. Au bout d'un nombre fini de boucles de ce type, on se retrouve avec un matrice SE décalée qui n'a que des zéros, ce qui est exclu comme on l'a démontré à l'étape 3. On doit donc sortir de la boucle.

5.4. **Conclusion du théorème de la discrimination positive.** La procédure ci-dessus, décrit une marche finie à l'intérieur du polytope, par des matrices dédiagonalisantes de M vers N.

6. Travaux pratiques

Reprenons le cas des matrices de transition à quartiles du Royaume Uni en 1996 Fig.11 et du Canada 1994 Fig.12 et calculons les matrices cumulantes associées (les unités sont exprimées en %)

Canada=

26	54,4	80,8	100
55	109,1	158,9	200
80,4	155,8	232,7	300
100	200	300	400

GB=

38	63	85	100
68	122	166	200
87	170	241	300
100	200	300	400

Dans le cas présent, la cumulante GB domine la cumulante Canada, ce qui n'est pas pour nous étonner ! Nous pouvons conclure que la situation intergénérationnelle du Canada aussure une plus grande mobilité que cette du Royaume Uni. Toutefois, je ne crois pas qu'il fallait beaucoup de mathématiques pour arriver à cette conclusion aussi banale !

L'algorithme de la section précédente ne présente pas de difficulté de bords (qui rend l'algorithme aussi difficile) et les positions choisies sont toujours issues d'un mouvement de taille 2×2. Les mouvements sont au nombre de 9.

Discussion 8. *La solution mathématique proposée vous semble-t-elle applicable dans la vraie vie ?*

6.1. **Programmation.** Écrire un programme qui donne la marche positive à suivre, connaissant les matrices de transition. Appliquer l'algorithme dans les autres situations issues des matrices de transition à quartiles données dans ces notes.

7. Ordre de Bruhat dans le groupe S_n

Pour les curieux, voici quelques mots sur l'ordre de Bruhat dans le groupe symétrique S_n. Cette section est indépendante et peut être écartée pour une première lecture. On pourra se référer au livre [1] pour apprendre des choses intéressantes sur les groupes de Coxeter.

7.1. **Présentation, longueur.** On note, pour $i = 1, \ldots, n-1$, $s_i = (i, i+1)$ les transpositions fondamentales. Elles engendrent S_n et vérifient les relations suivantes : $s_i^2 = 1$, $s_i s_j = s_j s_i$ si $|i-j| \geq 2$ et $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$ pour $1 \leq i < n$. Ces dernières relations sont dites relations de tresses.

Toute permutation w s'écrit comme produit de tels s_i. Il n'y a pas une unique façon de le faire. Cependant une expression sera dite réduite, si l'écriture en question utilise un nombre minimal de générateurs. Ce nombre est appelé la longueur et noté $\ell(w)$. Par exemple la permutation $w_o = (n, n-1, \ldots, 3, 2, 1)$ est de longueur $\frac{n(n-1)}{2}$. C'est la permutation de plus grande longueur [8].

Une observation facile est la suivante ; si $w = s_{p_1} s_{p_2} \cdots s_{p_k}$ une expression *réduite* de w, alors $s_{p_1} s_{p_2} \cdots s_{p_i}$ est une expression réduite pour tout $1 \leq i \leq k$.

On note $\mathcal{T}$ l'ensemble des conjugués des générateurs s_i. Clairement $\mathcal{T}$ est l'ensemble des transpositions (i, j).

En fait la longueur $\ell(w)$ vaut exactement le nombre d'inversions de w. En particulier on a $\ell(sw) = \ell(w) \pm 1$ pour s un générateur.

7.2. **Echange et simplification.** Il y a deux propriétés à connaître, la propriété d'échange et la propriété de simplification.

Propriété d'échange : Soit $w = s_{p_1} s_{p_2} \cdots s_{p_k}$ une expression *réduite* de w et t une transposition. Alors les points suivants sont équivalents,

i- $\ell(tw) < l(w)$
ii- $tw = s_{p_1} s_{p_2} \cdots \widehat{s_{p_i}} \cdots s_{p_k}$
iii- $t = s_{p_1} s_{p_2} \cdots \widehat{s_{p_i}} \cdots s_{p_2} s_{p_1}$.

On a une propriété analogue avec les produits wt. Attention le lemme d'échange ne dit pas que la longueur de tw est $k-1$, mais que tw s'écrit comme le produit de $k-1$ générateurs, expression non nécessairement réduite. La propriété suivante précise ce qui se passe dans un tel cas.

Propriété de simplification : Soit $w = s_{p_1} s_{p_2} \cdots s_{p_k}$ une expression non réduite avec $\ell(w) < k$. Alors il existe i, j tels que $1 \leq i < j \leq k$ et $w = s_{p_1} \cdots \widehat{s_{p_i}} \cdots \widehat{s_{p_j}} \cdots s_{p_k}$.

On déduit de la *propriété de simplification* : on peut, à partir d'une expression non réduite, obtenir une expression réduite en enlevant un nombre pair de générateurs.

7.3. **Définition de l'ordre de Bruhat.** Soit u, v deux éléments de S_n. On note $u \xrightarrow{t} v$ si $ut = v$ avec t une transposition dans $\mathcal{T}$ et $\ell(u) < \ell(v)$. En d'autres mots, la longueur augmente en multipliant à droite par une transposition (pas forcément un générateur).

On note $u \to v$ si on a $u \xrightarrow{t} v$ pour un certain t. Enfin $u \leq v$ signifie qu'il existe une suite
$$u = u_0 \to u_1 \to \cdots \to u_k = v.$$
Par exemple si $w = s_{p_1} s_{p_2} \cdots s_{p_k}$ est une expression réduite alors on a
$$e \to s_{p_1} \to s_{p_1} s_{p_2} \to \cdots \to s_{p_1} s_{p_2} \cdots s_{p_k}.$$

8. Justifier les affirmations sur w_o.

Remarquons alors que $u < v$ implique que $\ell(u) < \ell(v)$ et $u < ut$ si et seulement si $\ell(u) < \ell(ut)$.

7.4. **Propriété des sous-mots.** Les propriétés d'échange et de simplification nous assurent alors que u peut être obtenu comme sous mot d'une expression réduite de ut. Plus généralement $u \leq v$ implique que u est un sous-mot d'une expression réduite de v (et en vérité de toute expression réduite de v).

Enfin si $u \leq v$ alors on peut trouver une chaîne

$$u = u_0 \to u_1 \to \cdots \to u_k = v,$$

avec $\ell(u_{i+1}) = \ell(u_i) + 1$.

Cette dernière propriété pourrait laisser croire que l'on a $u_i \to u_{i+1}$ en multipliant par un générateur. La situation est plus compliquée :

Supposons que $u \xrightarrow{(a,b)} v$ avec $t = (a, b)$ une transposition avec $a < b$, alors $\ell(v) = \ell(u) + 1$ si et seulement si $u(a) < u(b)$ et il n'existe pas c tel que $a < c < b$ et $u(a) < u(c) < u(b)$. En d'autres termes la sous-matrice de coin Nord-Ouest $(u(a), a)$ et de coin Sud-Est $(u(b), b)$ est remplie de 0 (exceptés les deux coins cités). La sous-matrice correspondante pour v, a un coin Sud-Ouest $(u(b), a)$ et un coin Nord-Est $(u(a), b)$.

7.5. **Représentation graphique de l'ordre de Bruhat.** Soit w une permutation et considérons la matrice de permutation associée $\mathbf{P}_w$. C'est un tableau $n \times n$ avec pour entrées [9] $(w(j), j)$. La matrice cumulante est obtenue en comptant le nombre de 1 dans les rectangles Nord-Ouest, c'est à dire le cardinal de l'ensemble $w[|1; j|] \cap [|1, i|]$, que l'on note $w[i, j]$. On peut montrer alors que $v \leq w$ si et seulement si pour tout $1 \leq i, j \leq n$ on a

$$v[i, j] \geq w[i, j],$$

c'est à dire si la cumulante de la matrice $\mathbf{P}_v$ domine la cumulante de la matrice $\mathbf{P}_w$. Ainsi la domination des cumulantes définit l'ordre de Bruhat dans le groupe symétrique.

Références

[1] Anders Bjorner & Francesco Brenti (Auteur) *Combinatorics of Coxeter Groups*, Grauate Texts in Mathematics, 2005.

[2] Peter Magyar, *Bruhat Order for Two Flags and Line*, Journal of Algebraic Combinatorics **21** (2005).

[3] Michel Martinez, *Une contribution à la mesure de la mobilité intergénérationelle*, Ecole des hautes études en sciences sociales, Thèse de Doctorat 2004.

Inspection générale de, l'Education nationale, 107, rue de Grenelle, F-75007 Paris
E-mail address: `charles.torossian@education.gouv.fr`

9. Pour la numérotation des matrices, *ie.* la position $(1, 1)$ est la plus nord-ouest.