

ANNALES DE L'I. H. P., SECTION B

J. PELLAUMAIL

Graphes et algorithme de calcul de probabilités stationnaires d'un processus markovien discret

Annales de l'I. H. P., section B, tome 26, n° 1 (1990), p. 121-143

http://www.numdam.org/item?id=AIHPB_1990__26_1_121_0

© Gauthier-Villars, 1990, tous droits réservés.

L'accès aux archives de la revue « *Annales de l'I. H. P., section B* » (<http://www.elsevier.com/locate/anihpb>) 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/>

Graphes et algorithme de calcul de probabilités stationnaires d'un processus markovien discret

par

J. PELLAUMAIL

I.N.S.A., 20, avenue des Buttes-de-Coësmes,
35043 Rennes Cedex, France

RÉSUMÉ. — Le but de cette étude est de donner un algorithme qui permet, dans certains cas, de calculer effectivement les « probabilités stationnaires » d'un processus de Markov homogène. Cet algorithme a été testé sur plusieurs exemples. Sa précision est évaluée de façon théorique à l'aide d'un produit de matrices « mélangeantes » : le calcul de ces matrices utilise fortement la théorie des graphes. Pour l'essentiel de la partie théorique, l'ensemble des états est supposé fini. Par contre, l'algorithme proposé est utilisable quand l'ensemble des états est infini dénombrable.

ABSTRACT. — Our aim is to give an algorithm which in certain cases allows the effective computation of the stationary probabilities of an homogeneous Markov process. The algorithm has been tested on several examples, its precision can be theoretically estimated with the aid of a product of "mixing" matrices: the computation of these matrices heavily uses graph theory. The space of states is supposed finite in most of the theoretical part; however the proposed algorithm can be used for a denumerably infinite space of states.

INTRODUCTION ET HYPOTHÈSES DE BASE

Soit E un ensemble fini. Soit g une fonction positive définie sur $(E \times E)$ telle que, pour tout élément e de E , on a $g(e, e) = 0$. On suppose que cette fonction g possède la propriété suivante :

(i) il existe un élément e_0 de E tel que, pour tout élément e' de E on peut trouver une séquence $(x_i)_{1 \leq i \leq n}$ telle que $e' = x_1, e_0 = x_n$ et, pour tout $i, 1 \leq i < n, g(x_i, x_{i+1}) > 0$.

Dans ce cas on sait qu'il existe une seule famille $(p_e)_{e \in E}$ de réels telle que les deux propriétés suivantes soient satisfaites :

(ii) pour tout élément e de E ,

$$p(e) \sum_{e' \in E} g(e, e') = \sum_{e' \in E} p(e') g(e', e),$$

(iii) $\sum_{e \in E} p(e) = 1$.

De plus cette famille est une famille de réels positifs. Les éléments de cette famille seront appelés les « probabilités stationnaires ».

Cette dénomination est justifiée par l'exemple suivant : considérons un processus markovien $(X_t)_{t \in T}$ qui évolue en temps continu et qui admet E comme ensemble des états. On suppose que la fonction g est telle que, pour tout couple (e, e') d'éléments de E avec $e \neq e'$, on a :

$$g(e, e') = \lim_{h \downarrow 0} \left\{ \frac{1}{h} \text{Proba} [X_{t+h} = e' \mid X_t = e] \right\}.$$

Alors la famille $(p_e)_{e \in E}$ est la famille des probabilités stationnaires associées à ce processus.

Notons toutefois que les hypothèses données plus haut correspondent à des situations en équilibre qui peuvent (apparemment) n'avoir aucun lien avec la théorie des probabilités (en économie notamment).

Le but de l'étude qui suit est de donner un algorithme qui permet, dans certains cas, de calculer effectivement les « probabilités stationnaires » $(p_e)_{e \in E}$. La forme générale de cet algorithme est donnée au paragraphe D.

Le lecteur qui ne s'intéresse qu'aux applications peut, dans un premier temps, se contenter de lire le paragraphe E — où on explique cet algorithme dans un cas particulier simple — et, éventuellement, parcourir l'exemple illustratif proposé au paragraphe F.

Le théorème « théorique » fondamental qui permet de prouver sous des hypothèses très larges la validité de cet algorithme, est donné au paragraphe C. Il utilise fondamentalement le théorème rappelé au paragraphe B. La « vitesse » de convergence de l'algorithme peut être minorée à l'aide des coefficients introduits au paragraphe A.

Avant de terminer cette introduction, il est important d'insister sur les points suivants: cet algorithme n'appartient pas à la vaste classe des méthodes semi-théoriques, intéressantes mais le plus souvent inutilisables. Au contraire, il s'agit d'un algorithme totalement nouveau, relativement simple à mettre en œuvre, qui converge avec une rapidité étonnante, pour lequel on a une estimation *a posteriori* de la précision et une vitesse théorique de convergence. Par contre, cet algorithme, sauf dans des cas particuliers, ne conduit pas à une formule close.

A. ÉCART ANGULAIRE

Les fonctions *m* introduites dans ce paragraphe jouent un rôle analogue aux coefficients de contraction de Birkhoff (*cf.*, par exemple [Sen]): en général, elles sont plus faciles à utiliser techniquement.

A.1. Notations et hypothèses

Pour tout couple (x, y) de matrices uni-colonnes $(n \times 1)$, on pose :

$$\|x\| := \sum_{i=1}^n |x_i|$$

$$d(x, y) = \|(x/\|x\| - y/\|y\|)\|.$$

On dira que $d(x, y)$ est l'écart angulaire entre x et y . On dira qu'une matrice est (strictement) positive si chaque terme de cette matrice est (strictement) positif.

On dit qu'une matrice a est primitive si a est positive et s'il existe k tel que a^k soit strictement positive.

On considère la situation suivante :

- n est un entier ($n > 0$).

- M est une famille de matrices $(n \times n)$ positives, stable pour le produit matriciel.

m est une application définie sur M , à valeurs dans $[0, 1]$, qui satisfait aux deux propriétés suivantes :

- (i) pour tout couple (a, b) d'éléments de M on a : $m(ab) \leq m(a)m(b)$;

- (ii) pour tout couple (x, y) de matrices uni-colonnes $(n \times 1)$ positives et pour tout élément a de M , on a :

$$d(ax, ay) \leq m(a)d(x, y).$$

A. 2. Premier exemple

n étant fixé, M est l'ensemble des matrices stochastiques ($n \times n$), c'est-à-dire l'ensemble des matrices positives a telles que, quel que soit j , $\sum_{i=1}^n a_{ij} = 1$. On sait que M est stable pour le produit matriciel.

Pour tout élément a de M , on pose :

$$m(a) := \sup_{x, y} \frac{1}{2} \sum_{i, j} \left| ax_i ay_j - ax_j ay_i \right| / \sum_{i, j} ax_i ay_i$$

cette borne supérieure étant considérée pour tous les couples (x, y) de matrices ($n \times 1$) positives non nulles (évidemment, ax_i désigne le i -ième terme de la matrice uni-colonne ax).

On vérifie facilement que l'on a A. 1 (i) et (ii). On a aussi $m(a) < 1$ pour toute matrice a strictement positive appartenant à M .

A. 3. Deuxième exemple

n étant fixé, M est l'ensemble des matrices ($n \times n$) positives. Pour tout élément a de M , on pose

$$m(a) = \sup_{i, j, u, v} |a_{i, u} a_{j, v} - a_{i, v} a_{j, u}| / (a_{i, u} a_{j, v} + a_{i, v} a_{j, u}).$$

Là encore, on vérifie facilement que les propriétés A. 1 (i) et (ii) sont satisfaites et que l'on a $m(a) < 1$ pour toute matrice a strictement positive appartenant à M .

A. 4. Cas $m(a) < 1$

Pour tout vecteur y , posons $y_k := a^k y$; puisque l'écart angulaire entre y_k et y_{k+1} est plus petit que $m(a)^k$, la suite de Cauchy (y_k) converge en direction et, à la limite, on a : $ax = \lambda x$ avec x et λ positifs; la direction x est unique et l'écart angulaire entre x et y_k (quel que soit l'«initialisation» y) est inférieur à $m(a)^k / (1 - m(a))$ (convergence «géométrique»). On sait donc *a priori* majorer le nombre d'itérations nécessaires pour connaître x avec une approximation imposée.

De plus, si $ax = \lambda x$ et si $d(y, ay) = \varepsilon$, on a $d(x, y) \leq \varepsilon / (1 - m(a))$: autrement dit, si y est une solution «approchée», on a une estimation de la «qualité» de cette approximation.. On a évidemment des propriétés analogues pour toute matrice primitive.

Malheureusement, en général, a est bien une matrice primitive, mais il faut choisir k très grand pour que a^k soit strictement positive et $1 - m(a^k)$

est très petit : les inégalités précédentes sont donc souvent peu utilisables en pratique.

A. 5. Produits matriciels

Dans la suite (notamment *cf.* C. 9, 2°), nous utiliserons essentiellement l'application m introduite en A. 3 pour étudier certains produits matriciels.

Plus précisément, on considérera des suites $(R_k)_{k>0}$ de matrices carrées positives; en général, on a, quel que soit k , $m(R_k) < 1 - \rho$ avec ρ non négligeable. On a donc, dans ce cas, $m(R_j R_{j+1} \dots R_{k-1}) < (1 - \rho)^{k-j}$. Pour $(1 - \rho)^{k-j}$ assez petit, le produit $R_j \dots R_k Z_k$ dépend donc peu de Z_k : ce point a un rôle crucial dans tout ce qui suit (*cf.* aussi [BeP]).

B. SOLUTION STATIONNAIRE

B. 1. Théorème

On considère les hypothèses générales données dans l'introduction.

Pour tout élément e de E , on définit $F(e)$, $H(e)$ et $s_e(g)$ de la façon suivante :

$$F(e) := E \setminus \{e\}.$$

$H(e)$ est l'ensemble des applications f , définies sur E , à valeurs dans E , telles que $f(e) = e$ et telles que $f^j(e') = e'$ implique $e' = e$.

$$s_e(g) := \sum_{h \in H(e)} \prod_{e' \in F(e)} g[e', h(e')].$$

On pose :
$$r(g) := \sum_{e \in E} s_e(g).$$

On a alors, pour tout élément e de E :

(i)
$$s_e(g) \sum_{e' \in F(e)} g(e, e') = \sum_{e' \in F(e)} s_{e'}(g) g(e', e).$$

Si $r(g)$ est différent de zéro, on a donc

$$p_e(g) = s_e(g) / r(g).$$

De plus, la condition (i) donnée dans l'introduction implique $r(g) > 0$.

B. 2. Preuve

1° Soit e un élément de E . Soit $U(e)$ l'ensemble des applications f de E dans E qui satisfont aux trois propriétés suivantes :

- (i) $f(e) \neq e$,
- (ii) il existe $j > 1$ tel que $f^j(e) = e$,

(iii) $f^h(e') = e'$ implique qu'il existe k tel que $f^k(e) = e'$.

On pose :

$$u(e, g) := \sum_{f \in U(e)} \prod_{e \in E} g(e, f(e)).$$

2° On vérifie facilement que f appartient à $U(e)$ si et seulement si il existe h élément de $H(e)$ tel que $f = h$ en restriction à $F(e)$: ce petit exercice de théorie des graphes est laissé au lecteur. On a donc :

$$u(e, g) = s_e(g) \sum_{e' \in F(e)} g(e, e').$$

3° Par ailleurs, f appartient à $U(e)$ si et seulement si il existe un couple (e', h) avec h élément de $H(e')$ tel que $f = h$ en restriction à $F(e')$ et $f(e') = e$, $e' \neq e$ (ce couple est unique). On a donc :

$$u(e, g) := \sum_{e' \in F(e)} s_{e'}(g) g(e', e)$$

La famille $(s_e(g))_{e \in E}$ satisfait donc à la condition (ii) donnée dans l'introduction. Si $r(g) \neq 0$, on a donc bien $p_e = s_e/r$.

4° Supposons que la condition (i) de l'introduction soit satisfaite relativement à l'élément fixé $e(0) = e_0$ de E . Il est alors facile, en raisonnant par récurrence, de construire une fonction h appartenant à $H(e_0)$ et telle que, pour tout élément e' de $F(e_0)$, on ait $g(e', h(e')) \neq 0$. Ceci implique $s_{e(0)}(g) > 0$ (si g est positive) et donc $r(g) > 0$.

B.3. Remarque

1° Compte tenu de la définition de $s_e(g)$ et $r(g)$, on peut, dans $H(e)$, ne considérer que les applications h qui, en plus des conditions (i) et (ii), satisfont à la condition suivante :

(v) pour tout élément e' de $F(e)$, $g(e', h(e')) > 0$.

2° On peut démontrer directement que $r(g)$ est exactement le déterminant qui intervient dans le système des équations (ii) et (iii) données dans l'introduction (cf. [BoM], [Ber] et [Alg]). On peut aussi démontrer l'unicité de la solution de ce système d'équations en utilisant le paragraphe A.

C. THÉORÈME FONDAMENTAL

C.1. Énoncé du théorème

On se place dans le cadre général indiqué dans l'introduction : notamment, $(p_e)_{e \in E}$ désigne la « probabilité stationnaire » associée à g .

De plus, on suppose qu'il existe une application γ définie sur E , à valeurs dans l'ensemble $(0, \dots, n)$ telles que $|\gamma(e) - \gamma(e')| > 1$ implique $g(e, e') = 0$.

Quel que soit k , $0 \leq k \leq n$, on pose :

$$E_k := E(k) := \{e : e \in E \text{ et } \gamma(e) = k\}.$$

Alors, il existe des matrices positives $Z_0, Z'_n, S'_{0,e}$ (pour e élément de E_0), $S''_{n,e}$ (pour e élément de E_n), $Z_k, Z'_k, R_k, R'_k, S_{k,e}$ pour $1 \leq k \leq n-1$ et e élément de E_k — dont les dimensions seront précisées au cours de la démonstration —, telles que les propriétés suivantes sont satisfaites :

(i) Z_{n-1} et $S''_{n,e}$ (resp. Z'_1 et $S'_{0,e}$) ne dépendent que de la restriction de $g(\dots)$ au domaine $(E_n \times E)$ (resp. $(E_0 \times E)$);

(ii) pour $1 \leq k \leq n-1$, R_k, R'_k et $S_{k,e}$ ne dépendent que de la restriction de $g(\dots)$ au domaine $(E_k \times E)$;

(iii) pour $1 \leq k \leq n-1$, $Z_{k-1} = R_k Z_k$ et $Z'_{k+1} = Z'_k R'_k$;

(iv) pour $1 \leq k \leq n-1$ et e élément de E_k , $p_e = Z'_k S_{k,e} Z_k$;

(v) pour e élément de E_0 (resp. E_n) $p_e = S'_{0,e} Z_0$ (resp. $p_e = Z'_n S''_{n,e}$).

Les matrices $Z_{n-1}, Z'_1, R_k, S'_{0,e}$ et $S''_{n,e}$ seront explicitées au cours de la démonstration. Ceci permet donc de calculer la famille $(p_e)_{e \in E}$ puisque l'on a :

1° si e appartient à E_0 :

$$p_e = S'_{0,e} R_1 R_2 \dots R_{n-1} Z_{n-1};$$

2° si e appartient à E_n :

$$p_e = Z'_1 R'_1 R'_2 \dots R'_{n-1} S''_{n,e};$$

3° si e appartient à E_k , $1 \leq k \leq n-1$:

$$p_e = Z'_1 R'_1 \dots R'_{k-1} S_{k,e} R_{k+1} \dots R_{n-1} Z_{n-1}.$$

C. 2. Convention et notation $\pi(h, A)$

Dans ce type d'études (comme au paragraphe B), les relations d'ordre total — liées à la notation matricielle classique — sur les ensembles des indices des matrices considérées n'interviennent pas. Si A et B sont deux ensembles finis, on dira donc que M est une « matrice indexée par $(A \times B)'$ » si M est un tableau de nombres indexé par $(A \times B)$. Si M et M' sont deux matrices indexées par $(A \times B)$ et $(B \times C)$ respectivement, MM' est la matrice indexée par $(A \times C)$, définie par $(MM')_{a,c} := \sum_{b \in B} M_{a,b} M'_{b,c}$. Il est clair que

ceci n'est qu'une commodité de convention et que, pour retrouver les notations classiques, il suffirait d'expliciter, pour chaque ensemble d'indices — A par exemple — une bijection entre cet ensemble A et l'ensemble $\{1, \dots, \text{card}(A)\}$.

Par ailleurs, et indépendamment de ce qui précède, si h est une application définie sur une partie de E qui contient l'ensemble A , on posera :

$$\pi(h, A) := \prod_{e \in A'} g(e, h(e)) \quad \text{où } A' := \{e \in E : e \in A, h(e) \neq e\}.$$

Enfin si B est un ensemble vide, on pose :

$$\sum_{h \in B} \pi(h, A) = 0.$$

C.3. Notations h' , E'_k , E''_k , W_k , W'_k

Dans ce qui suit, on fera souvent intervenir des applications h définies sur un ensemble fini et satisfaisant à la propriété suivante :

(i) $h^j(e) = e$ implique $h(e) = e$.

Dans ce cas, pour tout état e' , $h^j(e')$ est fixe à partir d'un certain rang : on peut donc poser $h'(e') = \lim_{j \rightarrow \infty} h^j(e')$.

On pose :

$$\begin{aligned} E'_k &:= \{e \in E_k : \exists e' \in E_{k+1} \text{ tel que } g(e', e) \neq 0\} \\ E''_k &:= \{e \in E_k : \exists e' \in E_{k-1} \text{ tel que } g(e', e) \neq 0\} \\ U_k &:= E_k \cup E'_{k-1} \cup E''_{k+1}. \end{aligned}$$

Soit W_k l'ensemble des applications w définies sur E''_{k+1} et à valeurs dans E'_k . Soit W'_k l'ensemble des applications w' définies sur E'_{k-1} et à valeurs dans E''_k . La matrice Z_k (resp. Z'_k) est indexée par W_k (resp. W'_k). La matrice R_k (resp. R'_k) est indexée par $(W_{k-1} \times W_k)$ (resp. $W'_k \times W'_{k+1}$). La matrice $S_k(e)$ est indexée par $(W'_k \times W_k)$. La matrice $S'_{0,e}$ (resp. $S'_{n,e}$) est indexée par W_0 (resp. W'_n).

C.4. Calcul de R_k et R'_k

On fixe w élément de W_k et w' élément de W_{k-1} . On se propose de définir $R_k(w', w)$.

On appelle $H(k, w', w)$ l'ensemble des applications h définies sur U_k , à valeurs dans U_k et qui satisfont aux propriétés suivantes :

- (i) $h^j(e) = e$ implique $h(e) = e$;
- (ii) $h(e) = e$ si et seulement si e appartient à E'_{k-1} ;
- (iii) e élément de E''_{k+1} implique $h(e) = w(e)$;
- (iv) e élément de E'_k implique $h'(e) = w'(e)$.

On pose alors :

$$R_k(w', w) := \sum_{h \in H(k, w', w)} \pi(h, E_k).$$

Le calcul de R'_k est tout à fait analogue en intervertissant tous les termes. Plus précisément, on fixe w élément de W'_k et w' élément de W'_{k+1} . On appelle $H'(k, w, w')$ l'ensemble des applications h définies sur U_k à valeurs dans U_k et qui satisfont aux propriétés suivantes :

- (i) $h^j(e) = e$ implique $h(e) = e$;
- (ii) $h(e) = e$ si et seulement si e appartient à E''_{k+1} ;
- (iii) e élément de E'_{k-1} implique $h(e) = w(e)$;
- (iv) e élément de E'_k implique $h'(e) = w'(e)$.

On pose alors :

$$R'_k(w, w') = \sum_{h \in H'(k, w, w')} \pi(h, E_k).$$

C. 5. Calcul de Z_{n-1} et Z'_1

Soit w un élément de W_{n-1} . Soit $H''(n, w)$ l'ensemble des applications h de $U'_n := (E'_{n-1} \cup E_n)$ dans U'_n qui satisfont aux propriétés suivantes :

- (i) $h^j(e) = e$ implique $h(e) = e$;
- (ii) $h(e) = e$ si et seulement si e appartient à E'_{n-1} ;
- (iii) e élément de E''_n implique $h'(e) = w(e)$.

On pose alors :

$$Z_{n-1}(w) := \sum_{h \in H''(n, w)} \pi(h, E_n).$$

Soit w un élément de W'_1 . Soit $H''(0, w)$ l'ensemble des applications h de $U''_0 := (E_0 \cup E'_1)$ dans U''_0 qui satisfont aux propriétés suivantes :

- (i) $h^j(e) = e$ implique $h(e) = e$;
- (ii) $h(e) = e$ si et seulement si e appartient à E'_1 ;
- (iii) e élément de E'_0 implique $h'(e) = w(e)$.

On pose alors :

$$Z'_1(w) := \sum_{h \in H''(0, w)} \pi(h, E_0).$$

C. 6. Calcul de $S_k(e)$

Soit $k, 1 \leq k \leq n$, et e' un élément de E_k . On fixe w' élément de W'_k et w élément de W_k . Soit $G(k, e', w', w)$ l'ensemble des applications h définies sur U_k à valeurs dans U_k et qui satisfont aux conditions suivantes :

- (i) $h^j(e) = e$ implique $h(e) = e$;
- (ii) $h(e) = e$ si et seulement si $e = e'$;
- (iii) e élément de E'_{k-1} implique $h(e) = w'(e)$;
- (iv) e élément de E'_{k+1} implique $h(e) = w(e)$.

On pose alors :

$$S_{k, e'}(w', w) := \sum_{h \in G(k, e', w', w)} \pi(h, E_k).$$

C.7. Calcul de $S'_{0, e}$ et $S''_{n, e'}$

Soit e' un élément de E_0 et w un élément de W_0 . Soit $J(e', w)$ l'ensemble des applications h définies sur $U''_0 = E_0 \cup E'_1$, à valeurs dans U''_0 et qui satisfont aux propriétés suivantes :

- (i) $h^j(e) = e$ implique $h(e) = e$;
- (ii) $h(e) = e$ si et seulement si $e = e'$;
- (iii) e élément de E'_1 implique $h(e) = w(e)$.

On pose alors :

$$S'_{0, e'}(w) := \sum_{h \in J(e', w)} \pi(h, E_0).$$

Soit e' un élément de E_n et w un élément de W'_n . Soit $J'(e', w)$ l'ensemble des applications h définies sur $U'_n = E_n \cup E'_{n-1}$, à valeurs dans U'_n et qui satisfont aux propriétés suivantes :

- (i) $h^j(e) = e$ implique $h(e) = e$;
- (ii) $h(e) = e$ si et seulement si $e = e'$;
- (iii) e élément de E'_{n-1} implique $h(e) = w(e)$.

On pose alors :

$$S''_{n, e'}(w) := \sum_{h \in J'(e', w)} \pi(h, E_n).$$

C.8. Vérification du théorème C.1

1° La vérification de ce théorème est typique de la théorie des graphes. Une vérification formelle complète serait longue, fastidieuse et quasi illisible. En fait, pour se convaincre de l'exactitude du résultat, une étape essentielle est d'observer à l'aide de dessins les divers cas possibles en commençant par le cas où E_k a deux ou trois éléments.

2° L'idée de base est évidemment de vérifier que si on développe les formules 1°, 2° et 3° données dans le théorème C.1, on obtient exactement les mêmes produits que dans le théorème B.1. Nous allons expliquer le rôle de la matrice R_k . Celui de R'_k est analogue. Les matrices Z_{n-1} et Z'_1 correspondent aux raccords «aux bords» quand l'état e étudié n'est pas au bord; dans le cas où e est au bord, le raccordement s'effectue par l'intermédiaire des matrices $S'_{0, e}$ ou $S''_{n, e'}$. Si e n'est pas au bord, le raccordement en e s'effectue par l'intermédiaire de la matrice $S_k(e)$.

3° Le rôle des diverses matrices que l'on vient d'évoquer est facile à comprendre et à vérifier une fois que l'on a compris le «cœur» du théorème c'est-à-dire le rôle des matrices R_k . Pour faciliter l'expression, on suppose que les ensembles E_k sont «dessinés» sur un plan vertical (un tableau noir), à chaque ensemble E_k correspond une droite verticale, les droites les plus «à gauche» correspondant aux plus petites valeurs de k .

4° On fixe k . Les éléments h de $H(k, w', w)$ sont, «en simplifiant», les restrictions des fonctions h qui interviennent dans B.1 quand l'état e considéré est à gauche de k . On a une situation de type markovien mais dans un cadre tout à fait différent du cadre usuel.

5° Plus précisément, un élément w de W_k contient l'«information minimale» quand on va de la droite vers la gauche pour savoir quels sont les restrictions qui seront ultérieurement (c'est-à-dire plus à gauche) «autorisées».

Évidemment, quand on est en k , w élément de W_k correspond à l'information minimale qui provient de la droite et w' élément de W'_k correspond à l'information minimale qu'il faut transmettre à la gauche. La façon la plus simple de constater que w et w' contiennent suffisamment d'information est justement de regarder ce qui se passe en k .

6° Plus précisément, w et w' étant donnés, les fonctions globales h (définies sur E tout entier) et qui apparaissent dans le théorème B.1 associées au couple (w, w') sont celles qui sont compatibles avec ce couple au sens suivant : pour la fonction globale h , quand on considère un élément x de E_{k+1} , soit s la borne inférieure des entiers t tels que $h^t(x)$ appartienne à E_k (s est fini puisque le «point terminal» e est à gauche); h est compatible avec w si $h^s(x) := h^s(x) = w(x)$ (quand on part à droite, w détermine où on revient).

La compatibilité de h et de w' se définit de même mais, évidemment, relativement à R_k , les rôles de w et de w' ne sont pas du tout symétriques (cf. le 5° ci-dessus). Le couple (w, w') étant donné, R_k correspond aux produits $\pi(h, E_k)$ qui apparaissent dans le théorème B.1 et qui sont associés à des fonctions h compatibles avec w et w' .

7° Le point important, et un peu «miraculeux», est que, finalement, on aboutisse à un simple produit de matrices (positives).

C.9. Remarques

1° Pour certains petits réseaux simples, il est facile d'expliciter les diverses matrices qui interviennent dans le théorème C.1 : file M/PH/1, file PH/M/ r , cas où on peut utiliser la méthode des coupures (cf. [Lem]), etc. Il est également possible d'expliciter ces diverses matrices pour la file PH/PH/1 sous réserve d'utiliser — pour les états fictifs liés aux départs et ceux liés aux arrivées — une modélisation avec pivot (cf. [Pel]) et sous

réserve que le nombre d'états fictifs soit petit: en effet, s'il y a x états fictifs pour chaque loi les matrices R_k et R'_k sont indexées par $(W_k \times W_k)$ et $(W'_k \times W'_k)$ avec $\text{card}(W_k) = x^x = \text{card}(W'_k)$.

2° Par contre, dès que les réseaux sont complexes et de grande dimension, il devient vite difficile d'explicitier les matrices R_k, \dots ; l'intérêt du théorème C.1 repose alors sur les propriétés des produits de matrices positives. Or, même quand elles sont de grande dimension, les matrices R_k et R'_k sont, en général, très fortement «mélangeantes» en ce sens que les coefficients $m(R_k)$ et $m(R'_k)$ introduits au paragraphe A sont nettement inférieurs à 1. Notamment (cf. A.5), la relation:

$$p_e = Z'_1 R'_1 \dots R'_{k-1} S_{k,e} R_{k+1} \dots R_{n-1} Z_{n-1}$$

pour e élément de E^k montre que p_e dépend peu des «conditions aux bords» Z'_1 et Z_{n-1} quand les matrices R_j sont suffisamment mélangeantes et quand k et $(n-k)$ sont assez grands. Cette propriété est cruciale relativement aux algorithmes étudiés dans les paragraphes suivants.

De plus, si $R_k = R$ est fixe pour $k \geq j$ et si n est très grand (ou même «infini»), le produit $R_j R_{j+1} \dots R_{n-1} Z_{n-1} = R^{n-1} Z_{n-1}$ ne dépend presque pas de Z_{n-1} et est «presque» le vecteur propre positif associé à la plus grande valeur propre de R (cf. paragraphe G).

D. ALGORITHME DE BASE

D.1. Hypothèses et notations

On considère l'ensemble suivant de données et d'hypothèses: $n_1, n_2, m, n_3, n_4, n_5, n_6$ sont des entiers strictement positifs.

A est une matrice ($n_2 \times n_2$); pour $0 \leq j \leq m$, K_1 est une matrice ($n_1 \times n_2$); pour $0 < j < m-1$ et $1 \leq u \leq n_5$, G_{uj} est une matrice ($n_3 \times n_1$); pour $1 \leq k \leq m$ et $1 \leq v \leq n_6$, $H_{v,k}$ est une matrice ($n_4 \times n_1$); pour $0 \leq j < m$, F_j est une matrice ($n_1 \times n_1$); G_0 est une matrice ($n_3 \times n_1$) et H_m est une matrice ($n_4 \times n_1$).

Étant donné j et k , $0 < j < k < m$, et Y une matrice ($n_2 \times 1$), on pose:

$$f(j, k, X) := (K_j Y, K_{j+1} Y, \dots, K_{k-1} Y, K_k Y).$$

Dans tout ce qui suit, Y est une matrice «inconnue» ($n_2 \times 1$). On posera $n(5) := n_5$ et $n(6) := n_6$.

1° On suppose que le système $AY = 0$ a une seule solution non nulle à une constante multiplicative près.

2° On suppose que le système $AY = 0$ est équivalent à l'ensemble des trois conditions suivantes:

- (a) quel que soit j , $0 \leq j < m$, $K_{j+1} Y = F_j K_j Y$;
- (b) $G_0 K_0 Y = 0$;

(c) $H_m K_m Y = 0$.

3° Quels que soient j, k, u et v , avec $0 < j < k < m$, $1 \leq u \leq n_5$ et $1 \leq v \leq n_6$, le système des trois conditions suivantes a, à une constante multiplicative près, une seule solution non nulle (relativement aux inconnues $K_i, Y, j \leq i \leq k$):

(a) quel que soit $i, j \leq i < k, K_{i+1} Y = F_i K_i Y$;

(b) $G_{uj} K_j Y = 0$;

(c) $H_{v,k} K_k Y = 0$.

On pose :

$$f_{j,k,u,v} = f(j,k,Y)$$

où Y est une solution de ce système d'équations (avec, évidemment, $f_{j,k,u,v}$ non nulle).

4° Quels que soient j et $k, 0 < j < k < m$, si Y est une solution non nulle du système $AY = 0$, il existe une famille de nombres *positifs* $(c_{u,v})$, $1 \leq u \leq n(5), 1 \leq v \leq n(6)$ tels que

$$f(j,k,Y) = \sum_{u=1}^{n(5)} \sum_{v=1}^{n(6)} c_{u,v} f_{j,k,u,v}$$

Attention. — Il n'y a aucune corrélation entre les notations introduites dans ce paragraphe D et celles introduites au paragraphe C.

D.2. Remarques

1° Cet ensemble d'hypothèses peut sembler un peu compliqué et contraignant pour un lecteur non averti. En fait, on a très souvent (pour ne pas dire presque toujours) un tel système d'hypothèses quand on cherche à calculer les probabilités stationnaires d'un système markovien fini. Rappelons aussi qu'une situation simplifiée sera étudiée au paragraphe E.

2° L'hypothèse 2° est satisfaite à chaque fois que la matrice A (associée au générateur infinitésimal) est « tridiagonale par blocs ».

3° Les hypothèses jumelées 3° et 4° signifient, en simplifiant, que, quels que soient j et k fixés, $0 < j < k < m$, on peut trouver une famille de contraintes à gauche ($G_{uj} K_j Y = 0$) et une famille de contraintes à droite ($H_{v,k} K_k Y = 0$) telles que :

— d'une part, pour chaque couple (u, v) de contraintes, on ait une seule solution (sous-système ergodique) : on dira alors qu'on a des contraintes compatibles;

— d'autre part, cet ensemble de contraintes est suffisamment riche pour constituer une « base » en ce sens que la solution du système globale est une combinaison « convexe » des demi-droites vectorielles associées à ces contraintes.

En général, au niveau des exemples, ces contraintes sont assez faciles à expliciter : elles reviennent à introduire « artificiellement » une famille de « conditions aux bords extrémales » en procédant comme dans la démonstration du théorème E. 3.

D. 3. Algorithme

Supposons que l'on ait choisi j et k , $0 < j < k < m$; ce choix sera expliqué au paragraphe suivant.

Soit u , $1 \leq u \leq n_5$. Soit $(x_w)_{w \in W}$ une base du sous-espace vectoriel $G_{u,j}(X) = 0$: le noyau de $G_{u,j}$ ne peut pas être réduit à zéro [cf. 3° (b)].

Pour tout élément w de W , on calcule :

$$x'_w := F_k F_{k-1} \dots F_{j+1} F_j x_w.$$

Ensuite, pour tout entier v , $1 \leq v \leq n_6$, on calcule une famille de nombres $(z_w)_{w \in W}$ telle que $H_{v,k}(\sum_{w \in W} z_w x'_w) = 0$.

Une telle famille existe et est unique à une constante multiplicative près (hypothèse 3°). On pose $S_{u,v} := \sum_{w \in W} z_w x_w$.

On sait alors (hypothèse 4°) que, si Y est la solution du système global, $K_j Y$ appartient au « convexe » \mathcal{C} engendré par la famille $S_{u,v}$, $1 \leq u \leq n_5$, $1 \leq v \leq n_6$: or, si on connaît $K_j Y$, on connaît aussi $K_i Y$ pour $i \geq j$.

Le problème est alors de choisir j et k en sorte que :

- d'une part le convexe \mathcal{C} engendré par les $S_{u,v}$ soit de diamètre petit (relativement aux solutions du système global);
- d'autre part, les calculs qui précèdent soient compatibles avec la précision de l'ordinateur utilisé : notamment, il ne faut pas que le convexe évoqué ci-dessus soit trop petit!
- enfin que les matrices $K_i Y$ soient négligeables pour $i < j$.

Notons que, aux erreurs de calculs près (lesquelles ne sont pas à négliger), on sait que la solution appartient au convexe \mathcal{C}' associé à \mathcal{C} : on a donc une estimation *a posteriori* de la précision de la méthode utilisée (pour $K_i Y$, $i \geq j$) ce qui est assez rare pour de tels types de calculs.

D. 4. Choix de j et k

Le choix précis dépend des cas particuliers considérés. On ne peut donc ici que proposer des considérations empiriques concernant ce choix à partir d'une situation simplifiée qui sera revue complètement au paragraphe E qui suit.

On suppose donc que la matrice inconnue Y est la probabilité stationnaire d'un système markovien fini ergodique et que A est tridiagonale par

blocs. Le système $AY=0$ équivaut alors aux conditions «aux bords» et à la famille de relations matricielles $B'_i X_i = B_i X_{i+1} + B''_i X_{i-1}$; en posant $(K_i Y)^t = (X'_i, X'_{i+1})$ on se ramène à l'écriture proposée précédemment.

Soit n_7 un entier strictement positif. On suppose que X_i correspond aux états pour lesquels il y a i clients dans une station (ou $n_7 - i$) avec $0 \leq i \leq n_7$. Si les états correspondant à $i=0$ (resp. $i=n_7$) ont une probabilité globalement non négligeable, il faut choisir $j=0$ (resp. $k=n_7$) : dans ce(s) cas, l'algorithme proposé précédemment se simplifie considérablement.

En fait, les cas les plus importants en pratique sont ceux pour lesquels les états associés à $i=0$ et $i=n_7$ ont une probabilité négligeable: la probabilité d'avoir la station vide est négligeable et le nombre total de clients peut être infini donc n_7 est une intermédiaire technique qu'il faut choisir en sorte que la probabilité de dépasser n_7 soit négligeable. C'est dans ce cas que la méthode ici proposée est particulièrement efficace (pour ne pas dire que c'est la seule méthode fiable pour des systèmes non triviaux). On choisit alors j et k en sorte que les probabilités associées aux états $i=j$ et $i=k$ soient :

- d'une part suffisamment faibles pour que le poids de ce qui se passe en ces états soit faible : un minimum de réflexion montre que, en général, ceci implique que le «convexe» introduit précédemment à un diamètre petit; le théorème fondamental du paragraphe C précise ce point;

- d'autre part pas trop faibles : notamment si la précision du calculateur utilisé est ε et s'il y a N calculs à faire, il faut que le diamètre du convexe soit nettement supérieur à $N\varepsilon$: un minimum de réflexion, là encore, montre que l'on effectue des calculs qui n'ont aucun sens en partant d'un convexe trop petit qui conduira à une sécurité fallacieuse.

D.5. Liaison avec le théorème fondamental

On suppose donc qu'on a les hypothèses du théorème fondamental.

On suppose que la matrice X_k est la matrice des «probabilités stationnaires» $(p_e)_{e \in E(k)}$; les équations d'équilibre associées aux états éléments de E_i sont de la forme $B_i X_{i-1} + B'_i X_i + B''_i X_{i+1} = 0$. On est donc bien dans la situation expliquée en D.1 en posant $(K_i Y)^t = (X'_i, X'_{i+1})$. Au niveau des exemples, le choix des contraintes (cf. D.1.3°) s'effectue facilement : le problème «théorique» est de montrer que le convexe engendré par les $S_{u,v}$ a un diamètre qui tend vers zéro quand n tend vers l'infini. Si on revient aux notations du théorème fondamental, ceci revient exactement à dire que X_k dépend peu (modulo la constante de normalisation) de Z_{n-1} et Z'_1 (qui sont associées aux «conditions aux bords») : (cf. C.9.2°). C'est donc une propriété classique de type «ergodique» mais considérée dans un cadre tout à fait nouveau.

E. SITUATION DE BASE

E. 1. Introduction

Au paragraphe D, l'algorithme a été présenté sous la forme à laquelle on aboutit en général quand on utilise un ou plusieurs changements de variables (*cf.* E. 5). Nous allons maintenant expliquer le cas où il n'y a pas lieu d'effectuer de tels changements de variables. Cette situation simplifiée permettra de mieux comprendre comment on peut introduire les contraintes extrémales (*cf.* D. 1. 3°). Pour faciliter la compréhension, on se limite à un cadre simple.

E. 2. Hypothèses et notations

On considère les hypothèses générales données en introduction. De plus, comme dans le théorème C. 1, on suppose qu'il existe une partition finie $(E_k)_{0 \leq k \leq n}$ de E telle que $g(e, e') = 0$ si $e \in E_j$, $e' \in E_k$ et $|j - k| > 1$. On pose $n' = \text{card}(E_k)$.

Soit X_k la matrice uni-colonne dont les termes sont les « probabilités stationnaires » p_e pour $e \in E_k$. Les équations d'équilibre associées à ces mêmes états appartenant à E_k peuvent s'écrire :

$$B'_k X_k = B_k X_{k+1} + B''_k X_{k-1}$$

où B_k , B'_k , B''_k sont des matrices carrées.

Nous allons *supposer* que la matrice B_k est *inversible* quel que soit k , $0 \leq k \leq (n-1)$. Du point de vue technique, on suppose aussi que le poids relatif de X_0 n'est pas trop faible.

On a :

$$\begin{aligned} X_1 &= B_0^{-1} B'_0 X_0 \\ X_{k+1} &= B_k^{-1} (B'_k X_k - B''_k X_{k-1}). \end{aligned}$$

Il existe donc une matrice $(n' \times n')$ D telle que

$$X_n = DX_0.$$

Du point de vue technique, en général, il n'y a pas lieu d'explicitier les matrices B_k , B'_k et B''_k : il suffit d'implémenter l'algorithme qui permet de calculer X_{k+1} en fonction de X_k . Pour calculer D il suffit alors de calculer, à l'aide de cet algorithme, les valeurs de X_n associées aux n' initialisations (de X_0) correspondant aux n' colonnes de la matrice unité.

E. 3. Théorème

*La matrice D est inversible et son inverse D^{-1} est une matrice positive. De plus, pour tous les exemples que nous avons testés (*cf.* [Alg], [Ast], [BeP]),*

[Khm], [Oum], [Sag], etc.), le coefficient de contraction de Birkhoff de D^{-1} tend vers zéro quand n tend vers l'infini. Si ce coefficient est suffisamment faible, chaque colonne de D^{-1} donne, à une constante multiplicative près, une valeur approchée de X_0 .

Preuve. — 1° Soit u un élément de E_n . Soit g' la fonction positive définie sur $(E \times E)$ telle que :

- d'une part $g'(e, e') = g(e, e')$ sauf si $e' = u$ et si e appartient à E_n ;
- d'autre part $g'(e, u) = K + g(e, u)$ si e appartient à E_n où K est une constante.

Soit p' la « probabilité stationnaire » associée à g' ; cette famille p' existe et est unique (cf. la condition (i) de l'introduction). Soit $Y'(u, K)$ (resp. $X'(u, K)$) la matrice uni-colonne associée à cette solution p' et dont les termes sont les probabilités p'_e pour e élément de E_n (resp. E_0). Quels que soient u et K , on a

$$Y'(u, K) = DX'(u, K).$$

2° Lorsque K tend vers l'infini, le rapport $(Y'(u, K))_v / (Y'(u, K))_u$ tend vers zéro pour $v \neq u$ en appelant $(Y'(u, K))_v$ le terme p'_v associé à l'état v ($v \in E_n$). Pour K assez grand, les matrices uni-colonnes $Y'(u, K)$ sont donc linéairement indépendantes (quand u varie) : l'espace image associé à D est donc de dimension n' et la matrice D est inversible.

3° De plus, pour toutes les matrices $Y'(u, K)$, la matrice $X'(u, K) = D^{-1} Y'(u, K)$ est à termes positifs. Quand K varie, les combinaisons linéaires convexes des matrices $(Y'(u, K))_{u \in E(n)}$ parcourent tout l'octant strictement positif de \mathbb{R}^d (cf. 2°) : la matrice D^{-1} est donc à termes positifs.

4° Les diverses colonnes de D^{-1} sont les valeurs de X_0 associées aux initialisations « extrémales » de X_n . Or, le théorème fondamental C.1 montre que, en général, X_0 dépend peu en direction des valeurs de $g(e, e')$ pour $e \in E_n$ quand n est assez grand. Les diverses colonnes de D^{-1} sont donc, alors voisines en direction. Toutefois, cf. F. 5.

5° La fin du théorème est évidente.

E. 4. Diamètre du convexe

1° Soit Z_i la i -ième colonne de D^{-1} ; posons :

$$\delta := \sup_{i, j} d(Z_i, Z_j)$$

d étant défini comme en A. 1.

δ est le diamètre du convexe auquel appartient X_0 quelles que soient les « contraintes » en E_n . Le théorème fondamental C. 1 permet de calculer

une majoration *a priori* de δ . En fait, en général la valeur de δ est très nettement inférieure à celle donnée par cette majoration.

2° Attention : cette convergence étonnamment rapide du convexe quand n tend vers l'infini impose quelques précautions. Lorsque le diamètre δ de D^{-1} est de l'ordre, ou, *a fortiori*, inférieur, à l'ordre des erreurs de calcul, cela n'a évidemment plus de sens d'augmenter n . Ceci est facile à tester techniquement de la façon suivante : une fois que D^{-1} est calculée, on prend comme valeurs initiales de X_0 les n' colonnes de D^{-1} et on calcule les n' valeurs associées de X_n ; soit I' la matrice constituée des n' colonnes X_n ainsi obtenues et normalisées. S'il n'y avait pas d'erreurs de calculs la matrice I' serait égale à la matrice unité $I = DD^{-1}$. L'écart entre I et I' donne donc une idée très précise de l'ordre de grandeur des erreurs de calculs liées à l'algorithme expliqué ici.

3° Lorsque ces erreurs de calculs peuvent être négligées relativement à δ , on sait que la solution stationnaire cherchée p appartient au convexe engendré par les n' solutions associées aux n' colonnes de D^{-1} (quelles que soient les « contraintes » en E_n). On a donc une estimation *a posteriori* de la précision du résultat obtenu.

De plus, toute combinaison linéaire à coefficients positifs des probabilités stationnaires p — par exemples les diverses marginales — appartient au convexe associé à cette combinaison linéaire, lequel peut avoir un diamètre nettement inférieur à δ .

E. 5. Liaison avec le paragraphe D

La situation considérée dans ce paragraphe E est assez générale : dans la relation $B'_k X_k = B_k X_{k+1} + B''_k X_{k-1}$ le plus souvent l'une des matrices B_k ou B''_k est inversible. Toutefois, il y a souvent lieu de généraliser les hypothèses de ce paragraphe E dans deux directions (*cf.* paragraphe D).

D'une part, très souvent, on peut caractériser X_k par une matrice Y_k de dimension nettement inférieure à celle de X_k (*cf.* exemple F) : on utilise alors à la fois l'algorithme expliqué ici et la technique du changement de variable (*cf.* [Alg], [BDK], [BoP], [Pe], etc.).

D'autre part, si le poids relatif de X_k est faible tant que $k \leq k'$, avec k' « grand », il n'est pas techniquement possible de partir de X_0 (*cf.* E. 4. 2°). Il faut alors se limiter aux matrices X_k dont le poids n'est pas négligeable et utiliser des contraintes à la fois à droite et à gauche.

F. FILE AVEC PRIORITÉ

F. 1. Introduction

L'algorithme présenté au paragraphe D a été testé sur plusieurs exemples (*cf.* [Alg], [Ast], [BeP], [Khm], [Oum], etc.). Pour illustrer ce qui précède, nous allons expliquer un exemple (*cf.* [Sag]) qui est à la fois facile à exposer et suffisamment complexe pour mettre en évidence l'intérêt de la méthode. Pour être précis, nous allons donner les valeurs des divers paramètres et les formes des lois étudiées, etc., mais, évidemment, le programme est adapté à l'étude de cet exemple pour des lois et des paramètres quelconques sauf à être limité par la puissance de l'ordinateur utilisé.

Cet exemple relève de la théorie des files d'attente, essentiellement par commodité d'exposition. On aurait pu, tout aussi bien, prendre un exemple de réseau de Pétri puisque, dans l'algorithme considéré, tous les taux $g(e, e')$ peuvent dépendre librement des états e et e' . Plus précisément, l'adéquation de l'algorithme dépend des couples (e, e') pour lesquels $g(e, e') = 0$.

F. 2. Hypothèses

Attention : Les notations utilisées dans ce paragraphe sont indépendantes de celles utilisées aux paragraphes précédents.

On considère une file unique avec deux classes de clients. La classe 1 est prioritaire avec préemption. Le délai A entre deux arrivées suit une loi non exponentielle; quand il y a une arrivée, il y a une chance sur deux que ce soit un client de classe 1 (et de même pour la classe 2). Le délai D entre deux départs d'un client de classe 2 suit une loi non exponentielle. Par contre, on suppose que la loi de service de la classe 2 est exponentielle; quand il y a préemption, c'est, en général, une simplification raisonnable; le taux de service de la classe 2 (quand il n'y a pas de client de classe 1) vaut d_n si n est le nombre de clients de classe 2. On a choisi $d_n = d_2 + n d'_2$ (le terme $n d'_2$ correspond à l'impatience).

On pose

$$f(a, s) := a/(s + a), g(a, s, c) := ac^2/(s + ac)$$

et

$$\mathcal{L}_a(s) := f(a, s)((1 + f^2(a, s))(1 + g(a, s, c))/2(1 + c).$$

Le terme $f + f^3$ donne un mélange d'une loi exponentielle et d'une loi Erlang 3; le terme $(1 + g)$ donne un mélange entre deux lois que l'on peut choisir de taux très différents en faisant varier c .

La transformée de Laplace de la variable aléatoire A (resp. D) a été choisie égale à $\mathcal{L}_a(s)$ [resp. $\mathcal{L}_{d_1}(s)$].

Le nombre d'états fictifs associé à A (resp. D) sera appelé m (resp. m'). Pour l'exemple ci-dessus on a $m = m' = 4$.

On supposera dans la suite que la modélisation par états fictifs utilisée est une modélisation avec pivot (cf. [Pel]): au niveau des arrivées, cela diminue considérablement la taille mémoire; au niveau des départs, c'est indispensable pour avoir la « convergence du convexe ». Rappelons que la modélisation avec pivot *remplace* la technique bien connue de la « chaîne incluse ». L'état pivot associé aux arrivées (resp. départs) est choisi correspondant à $i = 1$ (resp. $j = 1$).

Enfin, on suppose qu'il y a des départs de clients de classe 1 par impatience: ces départs suivent une loi exponentielle de taux $k d'_1$ (si k est le nombre de clients de classe 1).

F.3. Algorithme

Soit $p(i, j, k, n)$ la probabilité qu'il y ait k (resp. n) clients de classe 1 (resp. 2) et que l'état fictif lié aux arrivées (resp. départs) soit i (resp. j).

On vérifie qu'on est dans la situation du paragraphe D relativement à n . De plus, la probabilité d'avoir $n = 0$ n'est pas négligeable ce qui simplifie notablement l'algorithme. On voit que, si on connaît le vecteur $U_0 := p(i, 1, 0, 0)$, les équations d'équilibre permettent de calculer directement toutes les autres valeurs $p(i, j, k, n)$.

L'algorithme est alors le suivant: on calcule le vecteur $U_N := p(i, 1, 0, N)$ en fonction des m valeurs de base positives extrémales (matrice unité) de U_0 . On obtient ainsi une matrice R carrée $m \times m$. L'inverse R^{-1} de cette matrice R est une *matrice positive* qui donne les m initialisations extrémales u de U_0 (le « convexe » \mathcal{C}). On calcule les diverses valeurs $p_u(i, j, k, n)$ associées aux m valeurs initiales de u . On sait que l'erreur finale est inférieure à E où :

$$E := \sup_{2 \leq u \leq m} \sum_{i, j, k, n} |p_1(i, j, k, n) - p_u(i, j, k, n)|$$

sous réserve que toutes les valeurs $p_u(i, j, k, n)$ soient positives (ce qui est facile à tester) et que l'ordre de grandeur des erreurs de calcul soit inférieur à E . On notera que le calcul de p_u , quand u varie, peut être mené en parallèle: ce point n'a pas été exploité, les logiciels utilisés n'étant pas adaptés au calcul vectoriel.

F.4. Nombre de calculs et taille mémoire

On suppose que l'on peut se restreindre à $k \leq K$ et qu'on arrête l'algorithme à $n = N$. Le nombre total d'états est alors $Nm(1 + Km')$.

La taille mémoire totale nécessaire pour implémenter l'algorithme expliqué précédemment est inférieure à $4Kmm'$ (pour $K \geq 5$).

L'ensemble des calculs comprend moins de $10(mm')^3$ NK multiplications, autant d'additions et un peu moins de tests.

Pour ce type d'exemple, ce sont les valeurs N et K qui peuvent être grandes : il est donc important d'avoir une taille mémoire proportionnelle à K et un temps de calcul proportionnel à NK.

F.5. Attention

En F.2, on a indiqué qu'on pouvait calculer le vecteur $U_N := p(i, 1, 0, N)$ en fonction du vecteur $U_0 := p(i, 1, 0, 0)$. *A fortiori*, on peut calculer le vecteur $V_N := (i, j, 0, N)$ en fonction du vecteur $V_0 := (i, j, 0, 0)$. Si on a choisi une modélisation avec pivot pour la loi des inter-arrivées, seuls les termes de V_0 associés à U_0 sont différents de zéro. Et pourtant, dans ce cas, on ne peut pas prendre les initialisations associées à V_0 : le convexe ne converge pas !

Pour que le convexe converge, il faut que les initialisations correspondent à des « contraintes compatibles » créées artificiellement au bord considéré : la méthode à utiliser est analogue à celle proposée en E.3. Il n'est évidemment pas possible de donner une formulation générale précise de cette technique.

F.6. Précision

Compte tenu de la faible taille mémoire nécessaire et de la grande stabilité de l'algorithme proposé, le programme a pu être testé sur un micro-ordinateur de faible puissance. Plus précisément, il a été testé sur un Goupil 3 standard (non muni de processeur de calcul en option) sous turbo-pascal standard.

Par exemple, il a été testé pour le système suivant de paramètres :

- $m = m' = 4$ (comme indiqué en F.2),
- $K = 10$ (nombre maximum de clients de classe 1),
- $N = 30$ (nombre maximum de clients de classe 2).

(on a donc 4920 états).

$$a = 1, \quad c = 1/3, \quad a' = 0,2, \quad d_1 = 2, \quad d'_1 = 0,2, \quad d_2 = 20, \quad d'_2 = 2/3.$$

Pour ce système de paramètres la norme « infinie » η (*i.e.* la borne supérieure du terme général) de $(I - I')$ (*cf.* E.4.2°) vaut $6 \cdot 10^{-6}$ (c'est le niveau des erreurs de calcul). L'erreur E (*cf.* la fin de F.3) vaut $4 \cdot 10^{-3}$.

C'est un exemple pour lequel l'approximation par un processus continu (processus de diffusion) n'est pas valable : les diverses probabilités sont

beaucoup trop contrastées. Par exemple les probabilités marginales associées à $k=0$, $k=1$ et $k=10$ valent respectivement 0,76, 0,2 et $8 \cdot 10^{-11}$. De même, les probabilités marginales associées $n=0$, $n=1$ et $n=10$ valent respectivement 0,93, 0,05 et $8 \cdot 10^{-5}$.

Enfin, on a augmenté le contraste concernant les lois des variables aléatoires A et D en prenant $c=1/9$ et $c=1/27$; la fiabilité des calculs a beaucoup moins diminué qu'on aurait pu le penser intuitivement puisque η vaut respectivement $3 \cdot 10^{-4}$ et $2,6 \cdot 10^{-3}$; de même E vaut respectivement 0,013 et 0,05.

G. CAS PARTICULIERS

G.1. Cas de matrices fixes

En général, les matrices R_k introduites au paragraphe C ou, équivalentement, les matrices F_k introduites au paragraphe D dépendent fortement de k . Dans ce cas et sauf quelques exemples très particuliers, il ne semble pas possible d'obtenir la solution stationnaire sous la forme de « formule close ».

Par contre, dans le cas très particulier où ces matrices (R_k ou F_k) sont fixes pour k « assez grand », la solution stationnaire peut être explicitée, soit complètement, soit partiellement à partir de matrices auxiliaires « calculables ». Cette situation a été extensivement étudiée par Neuts (*cf.* [Neu]). Notons que, même dans cette situation étudiée par Neuts, si les matrices auxiliaires évoquées ci-dessus sont difficiles à calculer effectivement, l'algorithme proposé au paragraphe D reste en général le plus efficace avec l'inconvénient qu'il s'agit d'un algorithme et non d'une « formule ».

De plus (*cf.* E.4), dès que n est assez grand, les probabilités pour k petit dépendent peu des taux en E_n : il n'y a donc pas d'inconvénients, *de ce point de vue*, à supposer que les matrices sont fixes à partir d'un certain rang. Mais, en même temps, ceci montre que l'étude de la partie « fixe » n'est pas forcément le point crucial.

G.2. Cas bidiagonal par blocs

Comme cela est indiqué en D.2.2°, l'hypothèse D.1.2° signifie que la matrice A associée au système global est tridiagonale par blocs; dans certains cas particuliers importants, on peut, par des transformations simples, se ramener à une matrice bidiagonale par blocs. Il est alors inutile d'effectuer la transformation donnée au paragraphe C et l'algorithme

proposé au paragraphe D peut être notablement simplifié. Ce cas est étudié dans [BeP].

RÉFÉRENCES

- [Alg] ALGERENNES, *Probabilités stationnaires pour des systèmes markoviens discrets*, rapport I.N.R.I.A., n° 633, 1987.
- [Ast] M. ASTOUATI, Méthodes des convexes pour le calcul de la probabilité stationnaire d'un réseau à deux stations à lois de services exponentielles, *Thèse de magistère*, Alger, 10 janvier 1985.
- [BeP] R. BELLAMINE et J. PELLAUMAIL, Opérateurs positifs et probabilités stationnaires, *R.A.I.R.O.*, Recherche Opérationnelle, vol. 23, n° 3, 1989.
- [BDK] P. BOYER, A. DUPUY et A. KHELLADI, *Méthodes de résolution d'un système linéaire markovien*, preprint, 1987.
- [Ber] C. BERGE, *Théorie des graphes et ses applications*, Dunod, 1967.
- [BoM] R. BOTT et J. P. MAYBERRY, *Matrices and trees*, Economic activity analysis, Wiley, 1954.
- [BoP] P. BOYER et J. PELLAUMAIL, *Deux files d'attente à capacité limitée en tandem*, Rapport I.R.I.S.A., n° 147, 1981.
- [Boy] P. BOYER, Évaluation des performances des systèmes réparés, *Thèse*, Rennes, 1982.
- [DER] I. F. DUFF, A. M. ERISMAN et J. K. REID, Direct methods for sparse matrices, *Oxford sciences publications*, 1986.
- [GeP] E. GELENBE et G. PUJOLLE, *Introduction aux réseaux de files d'attente*, Eyrolles, Paris, 1982.
- [GLM] E. GELENBE, J. LABETOULLE, R. MARIE, M. METIVIER, G. PUJOLLE et W. STEWART, *Réseaux de files d'attente*, Monographie A.F.C.E.T., 1980.
- [Khl] A. KHELLADI, *Changement de variable et probabilités stationnaires*, preprint.
- [Klm] Z. KHEMSI, *File d'attente avec deux classes de clients et arrivées non exponentielles*, preprint.
- [Lem] B. LEMAIRE, Théorème de conservation et blocage dans les files d'attente, *R.A.I.R.O.*, Recherche Opérationnelle, 1978, p. 395-399.
- [Len] L.-M. LE NY, *Probabilités stationnaires d'une file d'attente à deux seuils*, preprint.
- [Neu] M. F. NEUTS, *Matrix-geometric solutions in stochastic models*, The Johns Hopkins University Press, London, 1981.
- [Oum] B. OUMEHDI, Modélisation avec pivot d'un réseau de files d'attente à lois d'arrivée et de départ non exponentielles, *Thèse de magistère*, 1985.
- [Pel] J. PELLAUMAIL, Modélisation avec pivot pour une loi générale, *R.A.I.R.O.*, Recherche Opérationnelle, vol. 19, n° 3, 1985.
- [Sag] H. SAGGOU, *File d'attente avec deux classes de clients, les lois de départs et d'arrivées étant non exponentielles*, preprint.
- [Sen] E. SENETA, *Non-negative matrices and Markov chains*, Springer-Verlag, 1980.

(Manuscrit reçu le 27 juin 1988;
corrigé en juin 1989.)