BULLETIN DE LA S. M. F.

J.-P. BENZÉCRI

Approximation stochastique dans une algèbre normée non commutative

Bulletin de la S. M. F., tome 97 (1969), p. 225-241
<http://www.numdam.org/item?id=BSMF_1969__97__225_0>

Bull. Soc. math. France,
97, 1969, p. 225 à 241.

APPROXIMATION STOCHASTIQUE DANS UNE ALGÈBRE NORMÉE NON COMMUTATIVE

PAR

JEAN-PAUL BENZÉCRI.

Nous avons proposé, pour l'analyse des correspondances [1], une méthode d'approximation stochastique. Nous voulons ici donner une justification théorique de cette méthode : nous montrerons que, sous des hypothèses assez générales, le processus d'approximation converge vers le résultat cherché; quant à la rapidité de la convergence, les premiers résultats des calculs numériques, dus à J.-P. Fénelon, indiquent qu'elle est satisfaisante (le programme de calcul, non publié, sera communiqué aux chercheurs qui nous écriront).

Nos démonstrations reposent sur des calculs de majorations dans les algèbres normées; c'est pourquoi nous commencerons par des rappels relatifs à ces algèbres.

1. Algèbres normées.

Soit E un espace vectoriel normé; l'espace $L(E, E)$ des applications linéaires de E dans lui-même est muni d'une norme ainsi définie :

Si $l \in L(E, E)$, on pose

$$\| l \| = \sup_{\{x \mid x \in E;\, \| x \| = 1\}} \| l(x) \|.$$

Ainsi, $L(E, E)$ est muni d'une structure d'algèbre normée, i. e. d'une structure d'espace vectoriel normé, et d'une multiplication associative (la composition des applications linéaires) compatible avec la norme au sens suivant :

$$\| l_1 . l_2 \| \leq \| l_1 \| \times \| l_2 \|;$$

---

[1] *Cf.* e. g., notre article dans les *Methodologics of Pattern Recognition,* Edited by M. S. WATANABE. — New York, Academic Press, 1969.

Par exemple, si E est R^n, muni de la norme L_1, somme des modules des coordonnées :

$$\| x \| = \sum | x^i |,$$

$L(E, E)$ est l'espace des matrices carrées $n \times n$, avec pour norme :

$$\| a \| = \sup_j \Big(\sum_i | a^i_j | \Big).$$

On sait que sont équivalentes toutes les normes dont peut être muni un espace vectoriel de dimension finie sur R : en particulier, quelle que soit la norme $\| \quad \|$ dont est muni $L(R^n, R^n)$ (ou toute autre algèbre $\mathfrak{A}$ de dimension finie), on peut trouver une forme quadratique définie positive Q et une constante $k < 1$, telles que

$$\forall\, l \in L : \quad k Q(l, l) \leqq \| l \|^2 \leqq Q(l, l) ;$$

[on notera encore $\| l \|^2_Q$, le produit scalaire $Q(l, l)$ de l par elle-même].

Soit $\mathfrak{A}$ une algèbre normée complète ($\mathfrak{A}$ est nécessairement complète si elle est de dimension finie); soient $A \in \mathfrak{A}$ et $S(X)$ une série formelle à coefficients réels de rayon de convergence supérieur à $\| A \|$. On peut définir $S(A)$ comme somme d'une série convergente. Si S est une série à termes positifs, on a de plus $\| S(A) \| \leqq S(\| A \|)$. En particulier, quel que soit A, on peut calculer $\exp(A)$, et on a

$$\| \exp(A) \| \leqq \exp(\| A \|).$$

Supposons l'espace vectoriel de dimension finie E munie d'une norme quadratique $\| x \|^2_Q = Q(x, x)$. Soit $l \in L(E, E)$, une application linéaire de E dans lui-même, symétrique relativement à Q, i. e. telle que

$$\forall\, x, y \in E : \quad Q(x, l(y)) = Q(l(x), y).$$

On sait que E admet une base orthonormée pour Q, formée de vecteurs propres de l. On en déduit que, pour la norme sur $L(E, E)$ associée à la norme $\| \quad \|_Q$ sur E, la norme de l n'est autre que le maximum $| \lambda_{\max}(l) |$ du module des valeurs propres des l. Si on suppose de plus (ce qui sera le cas en analyse factorielle des correspondances) que les valeurs propres de l sont toutes positives [i. e. $\forall\, x : Q(x, l(x)) \geqq 0$], on aura

$$\| \exp(l) \| = \exp | \lambda_{\max}(l) | = \exp(\| l \|).$$

[Si l admettait des valeurs propres négatives, on aurait, plus généralement, pour norme de $\exp(l)$, l'exponentielle de la plus grande valeur propre de l.]

Ces rappels faits, nous pouvons énoncer, et démontrer, un théorème sur la convergence de certains produits aléatoires dans les algèbres normées. C'est ce théorème que nous appliquerons à la recherche des vecteurs propres par approximation stochastique, et en particulier à l'analyse des correspondances.

2. **Produits d'éléments aléatoires d'une algèbre normée.**

THÉORÈME 1. — *Soit* $\{A(i)\}$ *une suite d'éléments aléatoires indépendants, indicés par l'entier naturel* $i \in N$, *distribués suivant la même loi* μ *dans une algèbre normée unitaire de dimension finie* $\mathcal{A}$ *(autrement dit,* $\{A(i)\}$ *est un élément aléatoire de l'espace* $\Omega = \mathcal{A}^N$, *muni de la mesure de probabilité* μ^N*). Notons* $E(A)$ *l'espérance mathématique de* A *pour la loi* μ. *Soit* γ *un nombre positif,* $5/6 < \gamma \leq 1$.

Supposons, de plus, que, pour la loi μ, *ait une espérance mathématique la variable aléatoire* $\| A \|^p$, *où* p *désigne un nombre réel donné supérieur à* $4/(\delta\gamma - \delta)$.

Alors, pour tout ε, *il existe un entier* $N(\varepsilon)$ *et une partie* $\Omega(\varepsilon)$ *de* Ω, *telle que la mesure de* $\Omega(\varepsilon)$ *soit supérieure à* $1 - \varepsilon$, *et que si* $\{A(i)\} \in \Omega(\varepsilon)$, $\forall\, n_1, n_2 > N(\varepsilon)$:

$$\left\| \prod_{n_1}^{n_2} \left(1 + \frac{A(i)}{i^\gamma}\right) - \exp\left(\int_{n_1}^{n_2} \frac{du}{u^\gamma} \times E(A)\right) \right\|$$
$$< \varepsilon \times \exp\left(\| E(A) \| \times \int_{n_1}^{n_2} \frac{du}{u^\gamma} \right).$$

Dans cette formule, comme dans tout ce qui suit, un produit ordonné d'éléments d'une algèbre associative mais non commutative sera noté :

$$\prod_{n_1}^{n_2} (X(i)) = X(n_2).X(n_2 - 1) \ldots X(n_1 + 1)$$

(produit égal à 1 si $n_1 = n_2$, non défini si $n_2 < n_1, \ldots$). Dans les sommes $\sum$, comme dans les produits, on appliquera la convention que la borne inférieure est exclue, tandis que la borne supérieure est incluse.

Donnons d'abord un lemme qui nous permettra de comparer des produits d'éléments d'une algèbre normée non commutative.

LEMME 0. — *Soit* $\{V_i\}$, $\{v_i\}$ *deux suites finies d'éléments de* $\mathcal{A}$, *indicées par* $i = 1, \ldots, q$ *(*q *quelconque). Notons :*

$$\prod (V_i) = V_q . V_{q-1} \ldots V_1 = \prod_0^q (V_i);$$

$$\prod (V_i + v_i) = (V_q + v_q) \ldots (V_1 + v_1) = \prod_0^q (V_i + v_i).$$

Supposons que, quel que soit i, on ait $1/2 \leq \| V_i \|$; *et que, de plus :*

$$\sum \| v_i \| \leq 1/2.$$

Alors on a

$$\left\| \prod (V_i) - \prod (V_i + v_i) \right\| \leq \prod (\| V_i \|) \times \left(4 \sum \| v_i \| \right).$$

Preuve du lemme. — Quels que soient les V_i et les v_i, on a l'inégalité

$$\left\| \prod (V_i) - \prod (V_i + v_i) \right\| \leq \prod (\| V_i \| + \| v_i \|) - \prod (\| V_i \|).$$

En effet, en bref, le membre de droite est une somme de termes positifs dont chacun majore la norme d'un terme de l'expression droite. On peut encore écrite :

$$\left\| \prod (V_i) - \prod (V_i + v_i) \right\| \leq \prod (\| V_i \|) \times \left[\prod \left(1 + \frac{\| v_i \|}{\| V_i \|} \right) - 1 \right];$$

d'où en tenant compte de l'hypothèse que $\| V_i \|$ est supérieur à $1/2$:

$$\left\| \prod (V_i) - \prod (V_i + v_i) \right\| \leq \prod (\| V_i \|) \times \left[\prod (1 + 2 \| v_i \|) - 1 \right].$$

D'autre part, quel que soit $\rho \in [0, 1]$, on a l'inégalité

$$1 + \rho \leq \exp(\rho) \leq 1 + 2\rho.$$

D'où, pour le produit, en utilisant la majoration de la somme des $\| v_i \|$:

$$\prod (1 + 2 \| v_i \|) \leq \exp\left(2 \sum \| v_i \| \right) \leq 1 + 4 \sum \| v_i \|.$$

L'inégalité annoncée en découle immédiatement.

Notre propos est de comparer les deux produits

$$\prod_{n_1}^{n_2} \left(1 + \frac{A(i)}{i^\gamma} \right)$$

et

$$\prod_{n_1}^{n_2} \exp\left(\int_{i-1}^{i} \frac{du}{u^\gamma} \times E(A) \right) = \exp\left(\int_{n_1}^{n_2} \frac{du}{u^\gamma} \times E(A) \right).$$

Pour cela, on groupera en paquets les termes des produits, les paquets étant choisis *assez petits* pour que l'on puisse assimiler

$$\prod \left(1 + \frac{A(i)}{i^\gamma} \right) \quad \text{et} \quad 1 + \sum \frac{A(i)}{i^\gamma}$$

(donc calculer à l'intérieur de chaque paquet comme si l'algèbre $\mathcal{A}$ était commutative), mais aussi, *assez grands*, pour que la loi des grands nombres nous permette d'assimiler

$$\sum \frac{A(i)}{i^\gamma} \quad \text{et} \quad \left(\sum \frac{1}{i^\gamma}\right) \times E(A).$$

Afin de s'assurer que la loi des grands nombres joue, avec une probabilité arbitrairement grande indépendamment des bornes n_1, n_2, il sera commode d'utiliser toujours une même division en paquets des nombres entiers positifs. Pour cela on construira, une fois pour toutes, une fonction $N(t)$, entière positive, de variable entière positive, satisfaisant aux conditions suivantes :

$$\forall t: \quad N(t) + N(t)^\theta < N(t+1) \leqq N(t) + 2(N(t)^\theta),$$

l'exposant positif θ étant choisi dans l'intervalle $]0, 1[$, tel qu'il satisfasse à des conditions qui apparaîtront dans la suite.

Ainsi un produit de n_1 à n_2 sera décomposé en produits partiels suivant la formule

$$\prod_{n_1}^{n_2} X_i = \prod_{N(t_2+1)}^{n_2} X_i \times \prod_{t_1}^{t_2} \left[\prod_{N(t)}^{N(t+1)} X_i\right] \times \prod_{n_1}^{N(t_1+1)} X_i;$$

les bornes t_1 et t_2 étant définies par les inégalités

$$N(t_1) < n_1 \leqq N(t_1 + 1),$$
$$N(t_2 + 1) \leqq n_2 < N(t_2 + 2).$$

Dans nos majorations, les termes non regroupés en paquets [i. e. les termes de rang inférieur ou égal à $N(t_1+1)$, et ceux de rang supérieur à $N(t_2+1)$] ne joueront qu'un rôle mineur; en revanche, les paquets, dont le nombre peut être arbitrairement grand, seront l'objet d'évaluations que nous n'avons pas su faire simplement !

En vue d'appliquer le lemme 0, convenons de noter

$$V_i = \exp\left(\int_{i-1}^{i} \frac{du}{u^\gamma} \times E(A)\right),$$

$$V_i + v_i = 1 + \frac{A(i)}{i^\gamma},$$

$$W_t = \prod_{N(t)}^{N(t+1)} V_i = \exp\left(\int_{N(t)}^{N(t+1)} \frac{du}{u^\gamma} \times E(A)\right),$$

$$W_t + w_t = \prod_{N(t)}^{N(t+1)} (V_i + v_i).$$

Pour que le lemme 0 s'applique, il faut que $\| W_t \|$ soit bornée inférieurement par $1/2$: ceci sera assuré, à partir d'une certaine valeur de t, si on a $W_t \xrightarrow[t \to \infty]{} 1$; c'est bien le cas si θ est strictement inférieur à γ, car

$$\int_X^{X+2X^\theta} \frac{du}{u^\gamma} = \frac{1}{1-\gamma} X^{1-\gamma}[(1 + X^{\theta-1})^{1-\gamma} - 1] \approx X^{\theta-\gamma}.$$

Dans la suite, nous supposerons donc que $\theta < \gamma$; de plus, nous noterons N_0 un entier tel que, si $N(t) > N_0$, $\| W_{t-2} \| > 1/2$.

Ceci posé, on peut énoncer une proposition, qui nous donnera le plan de la démonstration du théorème 1 :

PROPOSITION 1. — *Supposons que* $N_0 < n_1 < n_2$, *et que, de plus*

$$\Delta = \sum_{N(t_2+1)}^{n_2} \| v_i \| + \sum_{t_1}^{t_2} \| w_t \| + \sum_{n_1}^{N(t_1+1)} \| v_i \| \leq 1/2;$$

alors on aura

$$\left\| \prod_{n_1}^{n_2} (V_i + v_i) - \prod_{n_1}^{n_2} (V_i) \right\| \leq 4\Delta \prod_{n_1}^{n_2} (\| V_i \|).$$

La proposition 1 est une conséquence immédiate du lemme 0.

Posons

$$u_t = \sum_{N(t)}^{N(t+1)} \| v_i \|,$$

le théorème 1 sera démontré, si on établit que, sous les hypothèses du théorème 1, u_t tend presque sûrement vers zéro et $\sum \| w_t \|$ converge presque sûrement. Plus précisément, on démontrera la proposition 2 :

PROPOSITION 2. — *Sous les hypothèses du théorème* 1, *quel que soit* ε, *il existe une partie* $\Omega(\varepsilon)$ *de* Ω, *de mesure supérieure à* $1 - \varepsilon$, *telle que pour les suites* $\{ A(i) \} \in \Omega(\varepsilon)$, u_t *tende uniformément vers zéro quand* $t \to 0$, *et* $\sum \| w_t \|$ *converge uniformément.*

Il est clair que le théorème 1 résulte facilement de la proposition 2 : car, soit N un entier satisfaisant aux conditions suivantes :

$$N_0 < N \qquad (cf.\ \text{supra}),$$
$$\forall \{ A(i) \} \in \Omega(\varepsilon) : \quad N < N(t+2) \Rightarrow u_t < \frac{\varepsilon}{12},$$
$$\forall \{ A(i) \} \in \Omega(\varepsilon) : \quad \sum_{N < N(t)} \| w_t \| < \frac{\varepsilon}{12}.$$

On aura bien, avec une probabilité supérieure à $1 - \varepsilon$ [ce sera vrai pour les suites de $\Omega(\varepsilon)$], pour tout $n_1, n_2 > N$:

$$\left\| \prod_{n_1}^{n_2} (V_i + v_i) - \prod_{n_1}^{n_2} (V_i) \right\| \leqq \varepsilon \prod_{n_1}^{n_2} (\| V_i \|) \leqq \varepsilon \exp\left(\| E(A) \| \times \int_{n_1}^{n_2} \frac{du}{u^\gamma} \right);$$

car, alors, le Δ de la proposition 1 sera majoré par $3 \frac{\varepsilon}{12} = \frac{\varepsilon}{4}$.

Pour achever de démontrer le théorème 1, il ne nous reste plus qu'à construire une partie $\Omega(\varepsilon)$ satisfaisant aux conditions de la proposition 2. Pour cela, nous construisons d'abord une partie $\Omega_1(\varepsilon)$, de mesure $\geqq (1 - \varepsilon/2)$, sur laquelle soit satisfaite une certaine majoration des $\| A(i) \|$. Nous utiliserons le lemme suivant :

LEMME 1. — *Pour que* $\| A(n) \| / n^\beta$ $(\beta > 0)$, *tende vers zéro presque sûrement, il suffit que, pour la* loi μ, $\| A \|$ *admette un moment* σ_p, *d'ordre p, tel que* $p\beta > 1$.

En effet, on a

$$\text{Prob}\{ \| A(n) \| > h n^\beta \} \leqq \frac{\sigma_p}{h^p . n^{p\beta}} = h(n);$$

$h(n)$ est le terme général d'une série convergente; donc la probabilité

$$\text{Prob}\{ \forall\, n > N : \| A(n) \| < h n^\beta \} \leqq 1 - \sum_{n > N} h(n)$$

tend vers 1 quand N tend vers l'infini.

En particulier, étant donné ε positif arbitrairement petit, et β (dont le choix sera précisé par la suite), on peut trouver un entier $N_1(\varepsilon)$ et un domaine $\Omega_1(\varepsilon)$ de mesure supérieure à $1 - \varepsilon/2$, tels que

$$\{ A(n) \} \in \Omega_1(\varepsilon);\ N_1(\varepsilon) < n \quad \Rightarrow \quad \| A(n) \| < n^\beta.$$

Pour assurer la convergence uniforme de la série des $\| w_t \|$, nous devrons restreindre $\Omega_1(\varepsilon)$, mais on peut déjà vérifier que, si $\beta + \theta < \gamma$, u_t tend uniformément vers zéro pour les suites de $\Omega_1(\varepsilon)$. En effet, on a

$$u_t = \sum_{N(t)}^{N(t+1)} \left\| 1 + \frac{A(i)}{i^\gamma} - \exp\left(\int_{i-1}^{i} \frac{du}{u^\gamma} \times E(A) \right) \right\|$$

$$\leqq \sum_{N(t)}^{N(t+1)} \frac{\| A(i) \|}{i^\gamma} + \| W_t - 1 \|.$$

On a vu plus haut que W_t [qui ne dépend pas de la suite $\{ A(i) \}$, mais seulement de $E(A)$ et de t] tend vers 1 : il nous reste donc à vérifier (ce qui est clair) que, sur $\Omega_1(\varepsilon)$, tend uniformément vers zéro la somme

$$\sum_{N(t)}^{N(t+1)} \frac{\|A(i)\|}{i^\gamma} \leqq \sum_{N(t)}^{N(t+1)} i^{\beta-\gamma} \leqq 2N(t)^{\theta+\beta-\gamma}$$

[car il y a au plus $2N(t)^\theta$ termes, chacun inférieur à $N(t)^{\beta-\gamma}$].

Reste maintenant à étudier la série des $\|w_t\|$: pour cela, nous intercalerons entre W_t et $W_t + w_t$ plusieurs intermédiaires. Nous poserons

$$W_t^0 = W_t = \exp\left(\int_{N(t)}^{N(t+1)} \frac{du}{u^\gamma} \times E(A)\right),$$

$$W_t^1 = 1 + \int_{N(t)}^{N(t+1)} \frac{du}{u^\gamma} \times E(A),$$

$$W_t^2 = 1 + \sum_{N(t)}^{N(t+1)} \frac{1}{i^\gamma} \times E(A),$$

$$W_t^3 = 1 + \sum_{N(t)}^{N(t+1)} \frac{A(i)}{i^\gamma},$$

$$W_t^4 = W_t + w_t = \prod_{N(t)}^{N(t+1)} \left(1 + \frac{A(i)}{i^\gamma}\right).$$

Posons encore, pour $s = 1, 2, 3, 4$,

$$w_t^s = W_t^s - W_t^{s-1};$$

il est clair que l'on a

$$w_t = w_t^1 + w_t^2 + w_t^3 + w_t^4.$$

Pour démontrer la proposition 2, donc le théorème 1, il nous suffira de restreindre $\Omega_1(\varepsilon)$ en un domaine $\Omega(\varepsilon)$, de mesure $\geqq 1 - \varepsilon$, sur lequel les quatre séries de forme générale w_t^s ($s = 1, 2, 3, 4$) convergent uniformément et absolument.

Pour les termes w_t^s, nous obtiendrons des majorations en fonction, non de t, mais de $N(t)$; nous utiliserons alors le lemme suivant :

LEMME 2. — *Soit une série, dont le terme général x_t satisfait à l'inégalité $0 \leqq x_t \leqq N(t)^{-\rho}$; la série converge absolument si $\theta + \rho > 1$ (θ étant le paramètre introduit pour définir la taille des paquets de termes).*

Pour prouver le lemme on va majorer

$$y_n = \sum_{2^n < N(t) \leqq 2^{n+1}} x_t,$$

cette somme comprend au plus $2^{n(1-\theta)}$ termes (les entiers de 2^n à 2^{n+1} forment au plus $2^{n(1-\theta)}$ paquets dont chacun comprend plus de $2^{n\theta}$ entiers), chacun inférieur à $2^{-n\rho}$; on a donc

$$y_n \leqq 2^{n(1-\theta-\rho)};$$

d'où, puisque $\theta + \rho > 1$, majoration par une série géométrique convergente.

Les termes w_t^1 et w_t^2 ne dépendent pas de la suite $\{A(i)\}$; il est facile de voir que moyennant certaines conditions imposées à θ et γ, la série des w_t^1 et celle des w_t^2 convergent absolument. On a

$$\| w_t^1 \| = \left\| \exp\left(\int_{N(t)}^{N(t+1)} \frac{du}{u^\gamma} \times E(A)\right) - 1 - \int_{N(t)}^{N(t+1)} \frac{du}{u^\gamma} \times E(A) \right\|.$$

Or l'intégrale est un $O(N(t)^{\theta-\gamma})$; de plus, on a

$$0 \leqq u \leqq 1 \quad \Rightarrow \quad (\exp(u) - 1 - u) < u^2;$$

d'où, pour $\| w_t^1 \|$, une majoration par un $o(N(t)^{2\theta-2\gamma})$; i. e., pour t supérieur à un t_0 convenable,

$$\| w_t^1 \| < K/N(t)^{2\gamma-2\theta},$$

où K est une constante; la convergence de la série des $\| w_t^1 \|$ en résulte, d'après le lemme 2, si on a

$$2\gamma - 2\theta + \theta = 2\gamma - \theta > 1.$$

Passons au terme w_t^2, on a

$$\| w_t^2 \| \leqq \| E(A) \| \times \left[\int_{N(t)}^{N(t+1)} \frac{du}{u^\gamma} - \sum_{N(t)}^{N(t+1)} \frac{1}{i^\gamma} \right]$$

ou [car, avec les conventions posées pour les sommes, le terme de rang $N(t)$ est exclu] :

$$\| w_t^2 \| \leqq \| E(A) \| \, N(t)^{-\gamma};$$

la convergence de la série des $\| w_t^2 \|$ en résulte d'après le lemme 2, si on a $\theta + \gamma > 1$.

Passons au terme w_t^4 : nous allons voir que, moyennant certaines conditions sur γ, β, θ, il y a convergence uniforme de la série des $\| w_t^4 \|$ sur le domaine $\Omega_1(\varepsilon)$. On a

$$w_t^4 = \prod_{N(t)}^{N(t+1)} \left(1 + \frac{A(i)}{i^\gamma}\right) - 1 - \sum_{N(t)}^{N(t+1)} \frac{A(i)}{i^\gamma},$$

d'où la majoration :

$$\| w_t^4 \| \leqq \prod_{N(t)}^{N(t+1)} \left(1 + \frac{\| A(i) \|}{i^\gamma}\right) - 1 - \sum_{N(t)}^{N(t+1)} \frac{\| A(i) \|}{i^\gamma};$$

ou, si $N(t) > N_1(\varepsilon)$ et si $\{ A(i) \} \in \Omega_1(\varepsilon)$ (*cf.* lemme 1),

$$\| w_t^4 \| \leq \prod_{N(t)}^{N(t+1)} (1 + i^{\beta-\gamma}) - 1 - \sum_{N(t)}^{N(t+1)} i^{\beta-\gamma};$$

on peut encore écrire

$$\| w_t^4 \| \leq \exp\left(\sum_{N(t)}^{N(t+1)} i^{\beta-\gamma} \right) - 1 - \sum_{N(t)}^{N(t+1)} i^{\beta-\gamma};$$

$$\| w_t^4 \| \leq \exp(2N(t+1)^{\theta+\beta-\gamma}) - 1 - 2N(t+1)^{\theta+\beta-\gamma}.$$

On sait que, si $0 \leq u < 1$, $\exp(u) - 1 - u < u^2$; d'où, si $\theta + \beta < \gamma$, une majoration de $\| w_t^4 \|$ par un $O(N(t+1)^{2\theta+2\beta-2\gamma})$; d'où la convergence absolue uniforme dans $\Omega_1(\varepsilon)$ de la série des w_t^4 si on a (*cf.* lemme 2) :

$$2\gamma - 2\theta - 2\beta + \theta > 1,$$
$$2\gamma > 1 + 2\beta + \theta$$

(inégalité qui implique $\theta + \beta < \gamma$, puisque $\gamma \leq 1$).

Reste le terme w_t^3. Nous aurons achevé la preuve de la proposition 2, donc du théorème 1, si nous construisons un domaine $\Omega_2(\varepsilon)$, de mesure $\geq 1 - \varepsilon/2$, et tel que, pour $\{ A(i) \} \in \Omega_2(\varepsilon)$, et pour t supérieur à un certain entier $t_2(\varepsilon)$, on ait

$$\| w_t^3 \| = \left\| \sum_{N(t)}^{N(t+1)} \frac{A(i) - E(A)}{i^\gamma} \right\| \leq \frac{1}{N(t)^\lambda},$$

où λ désigne un nombre réel positif satisfaisant à l'inégalité du lemme 2 ($\theta + \lambda > 1$), et à d'autres conditions qui apparaîtront dans la suite. En effet, sur $\Omega_2(\varepsilon)$, $\| w_t^3 \|$ sera uniformément convergent; et donc, d'après ce qui a été dit de w_t^1, w_t^2, w_t^3, w_t^4, sur $\Omega(\varepsilon) = \Omega_1(\varepsilon) \cap \Omega_2(\varepsilon)$ (intersection qui est de mesure $\geq 1 - \varepsilon$), la série des w_t sera absolument et uniformément convergente.

Ici, afin d'utiliser le fait que la variance d'une somme de variables aléatoires indépendantes est égale à la somme des variances de ces variables, nous supposerons l'algèbre $\mathcal{A}$ munie d'une norme quadratique $\| \; \|_Q$, équivalente à la norme $\| \; \|$ et de plus supérieure à celle-ci. Posons

$$D_Q(A) = E(\| A - E(A) \|_Q^2);$$

on a

$$D_Q\left(\sum_{N(t)}^{N(t+1)} \frac{A(i)}{i^\gamma} \right) = D_Q(A) \times \sum_{N(t)}^{N(t+1)} \frac{1}{i^{2\gamma}} \leq D_Q(A) \times 2N(t)^{\theta-2\gamma},$$

d'où l'on déduit, par l'inégalité de Čebyšev-Bienaymé :

$$\mathrm{Prob}\left(\| w_t^3 \|_Q > \frac{1}{N(t)^\lambda}\right) \times N(t)^{-2\lambda} < 2D_Q(A).N(t)^{\theta-2\gamma},$$

$$\mathrm{Prob}\left(\| w_t^3 \| > \frac{1}{N(t)^\lambda}\right) < 2D_Q(A).N(t)^{\theta-2\gamma+2\lambda}.$$

C'est là le terme général d'une série convergente si on a (*cf.* lemme 2) :

$$2\gamma - 2\lambda - \theta + \theta > 1,$$

$$\gamma - \lambda > 1/2.$$

A cette condition, on pourra déterminer $t_2(\varepsilon)$ tel que

$$\mathrm{Prob}\left(\exists\, t > t_2 : \| w_t^3 \| > \frac{1}{N(t)^\lambda}\right) < 2D_Q(A)\sum_{t>t_2} N(t)^{\theta-2\gamma+2\lambda} < \frac{\varepsilon}{2}.$$

Le domaine, de mesure $> 1 - \varepsilon/2$, ensemble des suites $\{ A(i) \}$, telles que

$$\forall\, t > t: \quad \| w_t^3 \| < \frac{1}{N(t)^\lambda}$$

sera un $\Omega_2(\varepsilon)$ convenable.

La démonstration du théorème est maintenant achevée, à ceci près qu'il nous faut collationner les inégalités imposées à γ, θ, β, λ. D'abord λ doit satisfaire aux deux inégalités

$$\theta + \lambda > 1; \qquad \gamma - \lambda > 1/2,$$

l'existence d'un λ convenable sera donc assurée si on a $1 - \theta < \gamma - 1/2$ (il suffira de prendre un λ entre ces deux bornes). Il nous reste donc, entre β, γ, θ, deux inégalités

$$\frac{3}{2} < \gamma + \theta; \qquad 1 + \theta + 2\beta < 2\gamma.$$

Comme β est nécessairement positif, on a, pour θ,

$$\frac{3}{2} - \gamma < \theta < 2\gamma - 1,$$

inégalité qui ne peut être satisfaite que si $\gamma > 5/6$. Pour plus petit que l'on prenne θ dans les limites permises, β sera inférieur à $(6\gamma - 5)/4$, quantité dont il pourra être arbitrairement voisin. On voit pourquoi nous avons besoin de l'existence d'un moment σ_p avec $4 < p(6\gamma - 5)$, c'est pour que l'on puisse avoir (*cf.* lemme 1) $p\beta > 1$.

Pour étudier la convergence du processus de recherche des facteurs, nous utiliserons une généralisation du théorème 1.

THÉORÈME 2. — *Soit* $\{A(i), B(i)\}$ *une suite aléatoire de couples d'éléments de l'algèbre normée* $\mathfrak{A}$. *Sur la loi de cette suite [cette loi est une mesure de probabilité sur* $\Omega^2 = (\mathfrak{A}^2)^N$*] faisons deux hypothèses :*

1° *La suite* $\{A(i)\}$ *est distribuée conformément aux hypothèses du théorème* 1.

2° *Chacun des* $\|B(i)\|$ *a une loi (mesure sur R) qui possède un moment d'ordre q donné, inférieur ou égal à une constante donnée* σ_q. *Notons* δ *un nombre réel tel que* $(\delta - 1) > \frac{1}{q}$.

Alors, pour tout ε, *il existe un entier* $N(\varepsilon)$ *tel que soit supérieure à* $1 - \varepsilon$ *la probabilité que*

$$\forall n_1, n_2 > N(\varepsilon) : \quad \left\| \prod_{n_1}^{n_2} \left(1 + \frac{A(i)}{i^\gamma} + \frac{B(i)}{i^\delta}\right) - \exp \int_{n_1}^{n_2} \frac{du}{u^\gamma} \times E(A) \right\|$$
$$< \varepsilon \exp\left(\|E(A)\| \times \int_{n_1}^{n_2} \frac{du}{u^\gamma} \right).$$

La démonstration du théorème 2 est analogue à celle du théorème 1 : outre les majorations déjà faites, on doit s'assurer que la série de terme général $(\|B(i)\|\, i^\delta)$ est uniformément convergente. Grâce à l'existence d'un moment d'ordre q pour $\|B(i)\|$, on peut, d'après le lemme 1, pour tout η trouver un $N(\eta)$ tel que soit supérieure à $(1 - \eta)$ la probabilité que

$$\forall i > N(\eta) : \quad \|B(i)\| < i^{\beta'},$$

où β' est un nombre inférieur à $\delta - 1$. Il en résulte aussitôt que la série des $\|B(i)\|/i^\delta$ converge uniformément avec une probabilité arbitrairement grande; d'où le théorème 2.

3. Application à l'analyse factorielle.

On suppose maintenant que $\mathfrak{A}$ est l'algèbre $L(E, E)$ des endomorphismes d'un espace vectoriel E, de dimension finie. L'espace E est muni d'une norme quadratique $\|x\|_Q^2 = Q(x, x)$; $L(E, E)$ est muni de la norme correspondante (*cf.* § 1). On considère une suite aléatoire $\{A(i)\}$ d'éléments indépendants distribués suivant la même loi μ; de l'espérance mathématique $E(A)$ (qui est une application linéaire de E dans lui-même), on suppose qu'elle est symétrique relativement à Q, et a toutes ses valeurs positives ou nulles; on note $\lambda_1, \lambda_2, \ldots$, la suite, rangée dans l'ordre strictement décroissant, des valeurs propres de $E(A)$; on note $F_1, F_2, \ldots$, la suite des sous-espaces propres (deux à deux orthogonaux) relatifs à ces valeurs propres; on note S_1 le supplémentaire orthogonal de F_1 (S_1 est somme directe de $F_2, F_3, \ldots$); on note $F_1(x)$ la projection orthogonale sur F_1 d'un vecteur $x \in E$.

A partir d'un vecteur, ou condition initiale, $X \in E$, on construit le processus

$$X(0) = X; \qquad X(n) = \prod_{0}^{n} \left(1 + \frac{A(i)}{i^{\gamma}}\right) \times X.$$

On escompte que, quand n tend vers l'infini, la direction de $X(n)$ tende vers une demi-droite de F_1, et, de plus, le rapport $\|X(2n)\|/\|X(n)\|$ tende vers $\exp(\lambda_1 . \mathrm{Ln}\, 2)$ cela est, en effet, le cas si $A(i)$ est certain et égal à $E(A)$, pourvu que $X(0)$ n'appartienne pas à S_1, ... [Un processus analogue, mais comportant orthogonalisation, doit de même fournir les p premiers vecteurs de $E(A)$: nous y reviendrons à la fin de ce paragraphe.] Nous allons démontrer le théorème de convergence suivant :

THÉORÈME 3. — *Supposons satisfaites les conditions du théorème; supposons, de plus, que quels que soient l'entier i et l'hyperplan $H \subset E$, soit nulle la probabilité que $X(i) \in H$; alors $X(n)$ converge presque sûrement en direction, vers F_1; et, si $\gamma = 1$, le rapport $\|X(2n)\|/\|X(n)\|$ tend presque sûrement vers* $\exp(\lambda_1 . \mathrm{Ln}\, 2)$.

Avant toute démonstration, une remarque sur la condition que $\forall\, i, H : P(X(i) \in H) = 0$. Si on fait choix d'une loi de probabilité pour la condition initiale $X(0)$, les lois de $X(1)$, $X(2)$, ... sont déterminées par la loi μ de A. Sous des hypothèses très générales relativement à μ, on sera assuré que, si la loi de $X(0)$ a une densité (ou même sans cela...), il en sera de même de la loi de $X(i)$, pour tout i; d'où résulte immédiatement que $\forall\, i, H : P(X(i) \in H) = 0$; et le théorème 3 assure le succès de l'approximation stochastique.

Dans la démonstration, afin de calculer sur des vecteurs bornés, on posera

$$Y(i) = \exp\left(-\lambda_1 \int_1^i \frac{du}{u^{\gamma}}\right) \times X(i),$$

et on donnera des conditions pour que, quand i tend vers l'infini, $Y(i)$ tende vers un vecteur non nul de F_1.

Le théorème 3 résulte immédiatement de la proposition suivante :

PROPOSITION 3. — *Reprenons les hypothèses et les notations du théorème* 1; *notons, de plus,*

$$\Omega(\varepsilon, k) = \bigcap_{p \geq 1} \Omega(\varepsilon/k^p),$$

où k est une constante supérieure à 1 [on sait alors que $\Omega(\varepsilon, k)$ est de mesure supérieure à $1 - (k\varepsilon/(k-1))$]; soit n, entier $\geq N(\varepsilon)$; alors, quel que soit $\{A(i)\} \in \Omega(\varepsilon, k)$, l'application linéaire

$$L(m) = \exp\left(-\lambda_1 \int_n^m \frac{du}{u^{\gamma}}\right) \times \prod_{n}^{m} \left(1 + \frac{A(i)}{i^{\gamma}}\right)$$

tend vers une limite quand m tend vers l'infini; cette application limite l est une surjection de E sur F_1; de plus, le noyau de l [*ensemble des x tels que* $l(x) = 0$] *est un supplémentaire S_l de F_1; tout vecteur x de S_l fait avec F_1 un angle dont le cosinus est inférieur à ε, autrement dit,*

$$\forall x \in S_1 : \quad \| F(x) \| / \| x \| < \varepsilon.$$

Pour déduire le théorème 3 de la proposition 3, il suffit de remarquer que si tout hyperplan (donc *a fortiori* pour toute sous-variété linéaire S_1 strictement incluse dans E) est de mesure nulle pour la loi de $X(n)$ [donc pour celle de $Y(n)$], $Y(i)$ converge vers un vecteur non nul de F_1 avec une probabilité supérieure à $1 - (k\varepsilon/(k-1))$; or ε peut être choisi arbitrairement petit, et k arbitrairement grand.

La proposition 3 va résulter des deux lemmes suivants :

LEMME 3. — *Soit $n > N(\varepsilon)$; $Y \in E$; $\varepsilon' > \varepsilon$. On peut trouver $m(Y, n, \varepsilon')$ tel que si $Y(n) = Y$:*

$$\forall \{A(i)\} \in \Omega(\varepsilon), \quad \forall i > m : \quad \| Y(i) - F_1(Y) \| < \varepsilon' \| Y \|.$$

Preuve du lemme 3. — Notons, pour $i \geqq n$,

$$F(i) = \exp\left(\int_n^i \frac{du}{u^\gamma} \times (E(A) - \lambda_1)\right) Y$$

[ici, $E(A) - \lambda_1$ désigne l'application linéaire somme de $E(A)$ et de l'homothétie de rapport $-\lambda_1$]. D'après le théorème 1, on a

$$\forall i \geqq n : \quad \| Y(i) - F(i) \| < \varepsilon \| Y \|;$$

d'autre part, quand i tend vers l'infini, $F(i)$ tend vers $F_1(Y)$; le lemme 3 en résulte immédiatement.

LEMME 4. — *Soient $n > N(\varepsilon)$, $k > 1$ quelconque, $Y \in E$ tel que $\| F_1(Y) \| / \| Y \| > \varepsilon$; alors, si $\{A(i)\} \in \Omega(\varepsilon, k)$, et si $Y(n) = Y$, $Y(i)$ tend, quand i tend vers l'infini, vers un vecteur non nul de F_1.*

Preuve du lemme 4. — Soit $\varepsilon' > \varepsilon$, et tel que

$$\varepsilon' \| Y \| / \| F_1(Y) \| = \sin\theta < 1$$

(un tel ε' existe par hypothèse). Il résulte du lemme 3 que, pour $i > m(Y, n, \varepsilon')$, $Y(i)$ a une norme supérieure à $\| Y \| (1 - \sin\theta)$, et fait avec $F_1(Y)$, donc avec F_1, un angle inférieur à θ. Soit p un entier assez grand pour que (ε/k^p) soit inférieur à $\cos\theta$; soit n' supéreur à $N(\varepsilon/k^p)$

et à $m(Y, n, \varepsilon')$: il résulte du lemme 3 que, pour $i > m(Y(n'), n', \varepsilon'/k^p)$, le vecteur $Y(i)$ fait avec $F_1(Y(n'))$, donc *a fortiori* avec le sous-espace F_2, un angle dont le sinus est inférieur à

$$(\varepsilon'/k^p) \times \frac{\| Y(n') \|}{\| F_1(Y(n')) \|} < \varepsilon' \cos\theta/k^p.$$

Le lemme 4 résulte de ce que p peut être pris arbitrairement grand.

On peut maintenant achever de démontrer la proposition 3. Choisissons une base de E formée de vecteurs $\{ s_k \}$ tels que $\| F_1(s_k) \| / \| s_k \|$ soit supérieur à ε. D'après le lemme 4, les images par $L(m)$ des vecteurs de cette base tendent, quand m tend vers l'infini, vers des limites (non nulles) appartenant à F_1 : ceci montre déjà que $L(m)$ tend vers une application linéaire limite l, et que $l(E) \subset F_1$. Pour montrer que l'image de l est F_1 tout entier, il suffit de remarquer que la restriction de l à F_1 est automorphisme de F_1, car (d'après le lemme 4) $l(x)$ ne peut être nul pour aucun vecteur x de F_1. Que tout vecteur du noyau $\delta(l)$ fasse avec F_1 un angle dont le cosinus est inférieur à ε résulte aussi, immédiatement du lemme 4.

Pour finir, traitons brièvement de la détermination simultanée de plusieurs directions propres (ou sous-espaces propres) de $E(A)$.

A partir d'un p-uple de vecteurs, $\{ X \} = \{ X_1, \ldots, X_q, \ldots, X_p \} \in E^p$, on construit le processus

$$\begin{aligned}
\{ X(0) \} &= \{ X_1(0), \ldots, X_p(0) \} = \{ X_q(0) \}, \\
\{ X'(n) \} &= \left\{ \prod_0^n \left(1 + \frac{A(i)}{i^\gamma} \right) X_q(0) \right\}, \\
\{ X(n) \} &= \text{orthogonalisé de } \{ X'(n) \},
\end{aligned}$$

l'orthogonalisation étant faite au sens de Schmidt-Hilbert, sans normalisation [i. e. $X_1(n) = X'_1(n), \ldots, X_q(n) = X'_q(n) +$ une combinaison des $X_h(n)$, $h < q$]. Si, par exemple, les p premières valeurs propres de $E(A)$ sont simples, on escompte que les $X_q(n)$ tendent en direction vers la $q^{\text{ième}}$ direction principale de $E(A)$ et que les quotients $\| X_q(2n) \| / \| X_q(n) \|$ tendent vers $\exp(\lambda_q . \mathrm{Ln}\, 2)$.

La convergence de l'algorithme va résulter des théorèmes 2 et 3, grâce à des considérations d'algèbre extérieure.

A toute application linéaire B de E dans lui-même est associée une application linéaire ${}^p\!\bigwedge B$ de ${}^p\!\bigwedge E$ dans lui-même; on pose

$${}^p\!\bigwedge B(x_1 \wedge \ldots \wedge x_p) = B(x_1) \wedge \ldots \wedge B(x_p);$$

si B est de la forme $1 + \frac{A(i)}{i^\gamma}$, on peut écrire ${}^p\bigwedge B$, comme une somme d'applications linéaires de ${}^p\bigwedge E$, dans lui-même :

$$1 + \sum_{q=1}^{p} \frac{{}^{p,q}A}{i^{\gamma q}},$$

où les ${}^{p,q}A$ se calculent en fonction de A seulement; par exemple, on a

$${}^{p,1}A(x_1 \wedge \ldots \wedge x_h \wedge \ldots \wedge x_p) = \sum_h x_1 \wedge \ldots \wedge Ax_h \wedge \ldots \wedge x_p;$$

et ${}^{p,p}A$ n'est autre que ${}^p\bigwedge A$.

Supposons A diagonalisable, avec pour vecteurs propres les x_j, relatifs aux valeurs propres λ_j : ${}^p\bigwedge A$ sera diagonalisable avec, pour vecteurs propres, les produits p à p des x_j (i. e. les

$$x_{j_1} \wedge \ldots \wedge x_{j_q} \wedge \ldots \wedge x_{j_p},$$

où $j_1 < \ldots < j_q < \ldots < j_p$) et, pour valeurs propres, les produits $\lambda_{j_1} \times \ldots \times \lambda_{j_q} \times \ldots \times \lambda_{j_p}$.

L'application linéaire ${}^{p,1}B$ a mêmes vecteurs propres que ${}^p\bigwedge B$, avec pour valeurs propres $\lambda_{j_1} + \ldots + \lambda_{j_p}$. Les deux applications linéaires $\left(\text{de } {}^p\bigwedge E \text{ dans lui-même}\right)$

$$\exp({}^{p,1}B), \quad {}^p\bigwedge \exp(1 + B)$$

sont égales, ayant encore les mêmes vecteurs propres avec, pour valeurs propres, les $\exp(\lambda_{j_1} + \ldots + \lambda_{j_p})$.

Associons maintenant, à une suite $\{ A(i) \}$ d'endomorphismes de E, une suite d'endomorphismes de ${}^p\bigwedge E$:

$${}^p\bigwedge\left(1 + \frac{A(i)}{i^\gamma}\right) = 1 + \frac{{}^{p,1}A(i)}{i^\gamma} + \frac{{}^{p,2}A(i)}{i^{2\gamma}} + \ldots + \frac{{}^{p,p}A(i)}{i^{p\gamma}}.$$

Soit $\{ A(i) \}$ aléatoire et satisfaisant aux conditions du théorème 1; $\left\{ {}^p\bigwedge\left(1 + \frac{A(i)}{i^\gamma}\right) \right\}$ satisfait aux conditions du théorème 2 si on pose $\delta = 2\gamma$ et

$$\frac{B(i)}{i^{2\gamma}} = \frac{{}^{p,2}A(i)}{i^{2\gamma}} + \ldots + \frac{{}^{p,p}A(i)}{i^{p\gamma}}.$$

Il en résulte que (sous des hypothèses de non-concentration sur un hyperplan analogues à celles du théorème 3) la suite des p-vecteurs ${}^pX(i)$, définis par

$$\begin{aligned} {}^pX(0) &= X_1(0) \wedge \ldots \wedge X_p(0), \\ {}^pX(i) &= \left[{}^p\bigwedge\left(1 + \frac{A(i)}{i^\gamma}\right)\right]^p X(i-1) \\ &= X_1(i) \wedge \ldots \wedge X_p(i) \end{aligned}$$

converge presque sûrement vers une direction propre relative à la plus grande valeur propre de

$$E({}^{p,1}A) = {}^{p,1}E(A)$$

[i. e. si, par exemple, les valeurs propres de $E(A)$ sont toutes simples, vers la direction de ${}^p\bigwedge E$ définie par le produit extérieur des p premiers vecteurs propres de $E(A)$] et, de plus, le quotient

$$\| {}^pX(2n) \| / \| {}^pX(n) \|$$

(la norme de E s'étend à ${}^pE$) tend vers exp (Ln 2.$(\lambda_1 + \ldots + \lambda_p)$) [i. e. vers l'exponentielle du produit par Ln 2 de la plus grande valeur propre de ${}^{p,1}E(A)$].

Revenons enfin à la suite des p-uples de vecteurs $X(i)$. Soit $q \leqq p$; le sous-espace linéaire engendré par les q premiers vecteurs $X_1(i), \ldots, X_q(i)$ du p-uple est complètement caractérisé par le produit extérieur ${}^qX(i)$ de ces q vecteurs; il résulte de ce qu'on vient de voir que (pour $q = p$, et plus généralement pour tout $q \leqq p$), le sous-espace engendré par les $X_1(i), \ldots, X_q(i)$ tend presque sûrement vers le sous-espace engendré par les q premières directions propres de $E(A)$ et que, de plus, le rapport de volumes

$$\frac{\prod \| X_1(2n) \| x \ldots x \| X_q(2n) \|}{\| X_1(n) \| x \ldots x \| X_q(n) \|} = \frac{\| {}^qX(2n) \|}{\| {}^qX(n) \|}$$

tend presque sûrement vers exp (Ln 2.$(\lambda_1 + \ldots + \lambda_q)$). Ceci suffit à établir pour l'algorithme d'approximation stochastique toutes les propriétés de convergence désirées.

(Texte reçu le 4 février 1969.)

Jean-Paul Benzécri,
Professeur à la Faculté des Sciences de Paris,
Institut de Statistique,
9, quai Saint-Bernard,
75-Paris 05.