Impact de méthodes non paramétriques sur l'évaluation numérique de la proximité de systèmes d'attente

Aicha Bareche [1] & Djamil Aïssani [1]

[1] *Unité de Recherche LaMOS (Modélisation et Optimisation des Systèmes), Université de Béjaia, Targa-Ouzamour, 06000, Algérie*
aicha_bareche@yahoo.fr ; lamos_bejaia@hotmail.com

Résumé. Dans ce travail, on montre l'impact des méthodes non paramétriques lors de l'évaluation numérique de l'erreur d'approximation des distributions stationnaires des systèmes d'attente $G/M/1$ et $M/M/1$, lorsque la fonction densité de la loi générale G des arrivées dans le système $G/M/1$ est inconnue et définie sur un support borné. Pour déterminer cette erreur, deux types de normes sont utilisés: la norme $\mathbf{L}_1$ et la norme poids de la méthode de stabilité forte. Un exemple numérique basé sur la simulation est présenté pour les deux cas de normes considérés et une étude comparative de résultats a été effectuée.

Mots-clés. Système d'attente, approximation, distribution stationnaire, stabilité forte, estimation non paramétrique, simulation.

Abstract. In this paper, we show the impact of nonparametric methods when evaluating numerically an approximation error between the stationary distributions of $G/M/1$ and $M/M/1$ queueing systems, and when the density function of the general law of arrivals G in the $G/M/1$ system is unknown and defined on a bounded support. To compute this error, we use two kinds of norms: the $\mathbf{L}_1$ norm and the weighted norm of the strong stability method. A numerical example based on a simulation study is presented for the two cases of the considered norms. A comparative study of the results has been provided.

Keywords. Queueing system, approximation, stationary distribution, strong stability, nonparametric estimation, simulation.

1 Introduction

Pour pallier les difficultés rencontrées dans l'obtention de solutions exactes et interprétables pour de nombreux systèmes d'attente, les analystes ont recours à des méthodes d'approximation. L'utilisation de ces méthodes permet d'approcher les caractéristiques d'un modèle complexe par celles d'un modèle plus simple. Il est intéressant dans ce cas de mesurer l'erreur commise lors de cette approximation.

La méthode de stabilité forte (Kartashov (1996)) permet de faire à la fois une analyse qualitative et quantitative de modèles stochastiques complexes. À la différence des autres approches de stabilité, cette méthode suppose que la perturbation du noyau de transition est petite par rapport à une certaine norme d'opérateurs (norme poids). Cette condition, beaucoup plus stricte que les conditions habituelles, permet d'obtenir des inégalités de stabilité avec un calcul exact de constantes et de fournir essentiellement de meilleures estimations pour les distributions stationnaires perturbées.

Dans ce travail, on s'intéresse à l'évaluation de l'erreur d'approximation des distributions stationnaires des systèmes $G/M/1$ et $M/M/1$ lorsque la fonction densité de la loi générale G des arrivées dans le système $G/M/1$ est inconnue et définie sur un support borné et doit être préalablement estimée par une méthode non paramétrique appropriée (voir Chen (2000), Schuster (1985)). Pour déterminer cette erreur, on utilise différentes normes, à savoir une norme $\mathbf{L}_1$ définie par Pedrono et Hellary (1983) et la norme poids de la stabilité forte (Kartashov (1996)). Un exemple numérique illustratif basé sur la simulation (sous Matlab 7.1) est présenté, permettent de réaliser une étude comparative.

2 Différentes normes considérées

2.1 Norme $\mathbf{L}_1$

Pour chercher la proximité de deux distributions stationnaires $\{\pi_i\}$ et $\{\alpha_i\}$ de deux systèmes d'attente, Pedrono et Hellary (1983) ont défini la métrique (norme $\mathbf{L}_1$) suivante:

$$\varepsilon = \max_{i\geq 0} |\pi_i - \alpha_i|. \tag{1}$$

Dans notre cas, la distribution stationnaire $\{\pi_i\}$ du système $M/M/1$ est donnée par: $\pi_i = (1-\rho)\rho^i$, $i = 0, 1, ...,$ où $\rho = \frac{\lambda}{\gamma}$ est la charge du système $M/M/1$, λ est le taux moyen de la durée des inter-arrivées et γ est la durée moyenne de service. La distribution stationnaire $\{\alpha_i\}$ du système $G/M/1$ est donnée par: $\alpha_i = (1-x)x^i$, $\forall i \geq 0$, où x est la solution unique (retrouvée numériquement par la méthode du point fixe) du système:

$$x = \int e^{-\gamma t(1-x)} g(t)dt, \tag{2}$$

avec g, la fonction densité de la loi générale G.

2.2 Norme d'opérateurs de la stabilité forte (norme poids)

2.2.1 Critère de stabilité forte

Nous rappelons la définition de base de la stabilité forte.

Définition 2.1 *(Voir Kartashov (1996)) Une chaîne de Markov X de noyau de transition P et de distribution stationnaire π est dite υ-fortement stable par rapport à la norme $\|.\|_\upsilon$ (où $\|\alpha\|_\upsilon = \sum_{j\geq 0} \upsilon(j)\,|\alpha_j|$ pour toute mesure α), si $\|P\|_\upsilon < \infty$ et chaque noyau stochastique Q dans un certain voisinage $\{Q : \|Q-P\|_\upsilon < \epsilon \ \text{ pour } \ \epsilon > 0\}$ admet une unique distribution stationnaire $\mu = \mu(Q)$ telle que $\|\pi - \mu\|_\upsilon \to 0$ lorsque $\|Q-P\|_\upsilon \to 0$.*

2.2.2 Stabilité forte de $M/M/1$ après perturbation du flot des arrivées

a) Description des modèles

Considérons un système $G/M/1(FIFO,\, \infty)$ où les inter-arrivées sont indépendantes et distribuées avec une distribution générale $H(t)$ et les durées de service sont distribuées avec $E_\gamma(t)$ (exponentielle de moyenne $1/\gamma$). Soit X_n^* le nombre de clients juste avant l'arrivée du $n^{ième}$ client. Il est aisé de prouver que X_n^* forme une chaîne de Markov (Kartashov (1996)) avec un opérateur de transition P^*.
Considérons également un système $M/M/1\ (FIFO, \infty)$, ayant des arrivées de Poisson de taux λ et la même distribution de la durée de service que dans le système précédent. Il est connu que X_n (le nombre de clients dans le système avant la $n^{ième}$ arrivée), forme une chaîne de Markov avec un opérateur de transition P.
Supposons que le flot des arrivées du système $G/M/1$ est proche d'un flot de Poisson. Cette proximité est caractérisée par la métrique:

$$w = w(H, E_\lambda) = \int_0^\infty |H - E_\lambda|\,(dt) = \int_0^\infty |h - e_\lambda|\,(t)\,(dt)\,, \tag{3}$$

où $|a|$ est la variation de la mesure a, h et e_λ sont les fonctions densité respectives des distributions H et E_λ. Désignons par $\{\alpha_k\}$ et $\{\pi_k\}$ les distributions stationnaires des états de X_n^* et X_n. Nous avons alors:

$$\begin{cases} \alpha_k = \lim\limits_{n\to\infty} P(X_n^* = k), k = 0, 1, 2, ... \\ \pi_k = \lim\limits_{n\to\infty} P(X_n = k), k = 0, 1, 2, ... \end{cases} \tag{4}$$

b) Estimation de la stabilité forte

Théorème 2.2 *(Voir Bouallouche et Aïssani (2006)) Supposons que la charge $\frac{\lambda}{\gamma}$ du système $M/M/1$ est plus petite que 1. Alors, pour tout β tel que $1 < \beta < \frac{\gamma}{\lambda}$, la chaîne de Markov induite X_n est fortement v-stable, après une petite perturbation des temps inter-arrivées, pour $v(k) = \beta^k$.*
La marge entre les opérateurs de transition est donnée par: $\|P^* - P\|_v \leq (1+\beta)w$.
Si de plus, $w < \frac{(1-\rho)(\gamma-\lambda\beta)}{(1+\beta)(2\gamma-\lambda(1+\beta))}$, on a l'erreur entre les distributions stationnaires:

$$Err = \|\alpha - \pi\|_v \leq \frac{(1+\beta)(2\gamma - \lambda(1+\beta))(\gamma-\lambda)w}{\frac{(\beta-1)(\gamma-\lambda\beta)^3}{(\beta-1)\gamma+\lambda\beta} - (2\gamma-\lambda(1+\beta))(1+\beta)(\gamma-\lambda\beta)w}, \tag{5}$$

où w est définie dans (3), α et π sont définies dans (4) et, $\rho = \beta\frac{\lambda}{\gamma - \frac{\gamma}{\beta} + \lambda}$.

Étant donné l'erreur ε définie dans (1) et la borne définie dans (5) exprimée dans le Théorème 2.2, il reste à calculer la solution x du système (2) et la distance de variation w définie dans (3), qui font toutes les deux intervenir la fonction densité inconnue de la loi générale G. Pour ce faire, des méthodes statistiques seront discutées par la suite.

3 Méthodes non paramétriques de densité considérées

Nous présentons dans cette section, un panorama très réduit des techniques d'estimation par noyaux. Soit $X_1, ..., X_n$ un échantillon issu d'une variable aléatoire X de densité de probabilité f et de distribution F. L'estimateur à noyau de Parzen-Rosenblatt (Parzen (1962)) de la densité f au point $x \in \mathbf{R}$, $f(x)$, est donné par:

$$f_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right), \tag{6}$$

où K est une densité symétrique appelée noyau, h_n est dit paramètre de lissage.

Plusieurs résultats sont connus dans la littérature quand la fonction densité est définie sur $\mathbf{R}$. Lorsque cette dernière est définie sur un support borné, sans correction, l'estimateur à noyau souffre d'effets de bord puisque il présente un biais au bord. Ce problème est dû à l'usage d'un noyau fixe qui assigne un poids en dehors du support quand le lissage est considéré près du bord.

Pour remédier à ce problème, Schuster (1985) suggère de créer l'image miroir des données de l'autre côté du bord puis d'appliquer l'estimateur (6) pour l'ensemble de données initiales et de nouvelles données créées. $f(x)$ est alors estimée, pour $x \geq 0$, par:

$$\tilde{f}_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} \left[K\left(\frac{x - X_j}{h_n}\right) + K\left(\frac{x + X_j}{h_n}\right)\right]. \tag{7}$$

Une autre simple idée pour éviter les effets de bord, est l'usage d'un noyau flexible, qui n'assigne jamais un poids en dehors du support de la fonction densité et qui corrige automatiquement et implicitement les effets de bord. On cite les estimateurs à noyaux asymétriques (Chen (2000)) donnés par la forme:

$$\hat{f}_b(x) = \frac{1}{n} \sum_{i=1}^{n} K(x, b)(X_i), \tag{8}$$

où b est le paramètre de lissage et le noyau asymétrique K peut être pris comme la densité d'une loi Gamma K_G avec les paramètres $(\frac{x}{b} + 1, b)$ donnée par:

$$K_G\left(\frac{x}{b} + 1, b\right)(t) = \frac{t^{x/b} e^{-t/b}}{b^{x/b+1} \Gamma(x/b + 1)}. \tag{9}$$

4 Exemple numérique – Simulation

On considère les quatre cas suivants:
Premier cas: Considérons un système $G/M/1$ tel que la fonction densité de la loi générale G est donnée par:

$$g(x) = \begin{cases} \frac{1}{2}e^{-x} + e^{-2x}, & \text{si} \quad x \geq 0; \\ 0, & \text{sinon.} \end{cases} \tag{10}$$

On génère un échantillon issu de la loi générale G (supposée hyper-exponentielle dans notre cas) avec la fonction densité $g(x)$ définie dans (10). On applique ensuite la méthode du noyau et les techniques de correction des effets de bord pour estimer cette densité en utilisant les différents estimateurs donnés dans les cas suivants:
Deuxième cas: Utilisation de l'estimateur à noyau $g_n(x)$ défini dans (6).
Troisième cas: Utilisation de l'estimateur image miroir $\tilde{g}_n(x)$ défini dans (7).
Quatrième cas: Utilisation de l'estimateur à noyau asymétrique $\hat{g}_b(x)$ défini dans (8) avec le noyau Gamma donné dans (9).
Pour les trois derniers cas, on prend: la taille de l'échantillon $n = 200$, le nombre de simulations $R = 100$. Dans tous les cas, on introduit le taux moyen de service $\gamma = 10$. Nous déterminons d'abord le taux moyen de la durée des inter-arrivées: $\lambda = \frac{1}{\int xg(x)dx}$.
Pour préciser l'erreur de proximité sur les distributions stationnaires ε définie dans (1) des systèmes $G/M/1$ et $M/M/1$, nous élaborons un algorithme approprié. Nous obtenons les résultats de la Table 1. Pour déterminer la distance de variation w et l'erreur d'approximation Err sur les distributions stationnaires, on élabore un algorithme (voir Bareche et Aïssani (2008) pour les étapes générales à suivre). Les résultas sont résumés dans la Table 2.

Densité théorique et ses différents estimateurs	$g(x)$	$g_n(x)$	$\tilde{g}_n(x)$	$\hat{g}_b(x)$
Solution x du système (2)	0.1475	0.0668	0.1223	0.1364
Erreur sur les distributions stationnaires ε	0.0142	0.0665	0.0220	0.0170

Table 1: Erreur de proximité ε selon la norme $\mathbf{L}_1$ avec les différents estimateurs

Densité théorique et ses différents estimateurs	$g(x)$	$g_n(x)$	$\tilde{g}_n(x)$	$\hat{g}_b(x)$
Distance de variation w	0.0711	0.2104	0.0895	0.0792
Erreur sur les distributions stationnaires Err	0.21		0.35	0.26

Table 2: Erreur de proximité Err selon la norme poids avec les différents estimateurs

Discussion et conséquence: On remarque dans la Table 1 que les erreurs d'approximation ε obtenues par les estimateurs image miroir et à noyau asymétrique

Gamma sont meilleures que celle obtenue par la méthode du noyau classique. On remarque aussi dans la Table 2 que l'erreur d'approximation Err est donnée en utilisant les estimateurs image miroir et à noyau asymétrique Gamma, contrairement au cas de l'estimateur à noyau classique, où cette erreur n'a pas pu être déterminée.
En comparant les résultats des Tables 1 et 2, on remarque que les conclusions tirées se rejoignent pour les deux normes considérées, dans le sens que la méthode du noyau présente des insuffisances et que les techniques de correction des effets de bord sont plus appropriées dans notre cas d'étude. Cependant, on note que l'erreur Err (en norme poids) est assez importante comparée à l'erreur ε (en norme $\mathbf{L}_1$). Cela est dû, dans un premier temps, à la différence entre les deux norme. De plus, lors de l'application de la stabilité forte, on rajoute à l'erreur commise par l'estimation non paramétrique l'erreur résultante de la perturbation effectuée dans ce cas. Ce constat peut suggérer d'opter pour la norme $\mathbf{L}_1$ pour l'évaluation numérique de la proximité de deux systèmes d'attente quelconques. Cela serait facile si on avait à traiter de systèmes simples pour lesquels on peut calculer aisément les distributions stationnaires. Cependant, dans la pratique, la plupart des systèmes n'ont pas de résultats analytiques sous forme explicite (exemple de $G/G/1$).

5 Conclusion

L'étude comparative réalisée entre les résultats obtenus par les différentes méthodes non paramétriques avec les deux cas de normes considérés montre l'impact de celles qui prennent en compte la correction des effets de bord pour préciser l'erreur d'approximation désirée entre les systèmes considérés ($G/M/1$ et $M/M/1$). Il serait intéressant de considérer les résultats de ce travail pour l'approximation du système plus complexe $G/G/1$.

Bibliographie

[1] Bareche, A. et Aïssani, D. (2008), Kernel density in the study of the strong stability of the $M/M/1$ queueing system, *Operations Research Letters*, 36(5), 535–538.
[2] Bouallouche, L. et Aïssani, D. (2006), Measurement and Performance of the Strong Stability Method, *Theory of Probability and Mathematical Statistics*, 72, 1–9.
[3] Chen, S. X. (2000), Probability Density Function Estimation Using Gamma Kernels, *Ann. Inst. Statis. Math.*, 52, 471–480.
[4] Kartashov, N. V. (1996), *Strong Stable Markov Chains*, TbiMC Scientific Publishers, VSPV, Utrecht.
[5] Parzen, E. (1962), On estimation of a probability density function and mode, *Ann. Math. Stat.*, 33, 1065–1076.
[6] Pedrono, R. et Hellary, J. M. (1983), *Recherche Opérationnelle*, Herman ed., Paris.
[7] Schuster, E. F. (1985), Incorporating support constraints into nonparametric estimation of densities, *Commun. Statist. Theory Meth.*, 14, 1123–1136.