
À LA RECHERCHE DU PLAN DÉTERMINANTAL OPTIMAL

Kim ANTUNEZ (*), Vincent LOONIS (**)

(*) Insee, Direction de la diffusion et de l'action régionale

(**) Insee, Direction de la méthodologie et de la coordination statistique et internationale

kim.antunez@insee.fr, vincent.loonis@insee.fr

Mots-clés : Échantillonnage déterminantal, équilibrage, échantillonnage spatial, optimisation semi-définie, recuit simulé

Domaine concerné : Théorie des sondages amont ; Échantillonnages particuliers

Résumé

Loonis et Mary (2019) [1] introduisent une nouvelle famille de plans de sondages paramétrée par les matrices hermitiennes contractantes, dont ils étudient les propriétés théoriques et pratiques : les plans déterminantaux. Ils montrent qu'une grande partie des propriétés, à distance finie ou asymptotiques, d'un plan déterminantal découle directement des caractéristiques de la matrice K à laquelle il est associé. Pour un jeu de probabilité d'inclusion simple fixé, ils identifient une sous-famille de plans déterminantaux conduisant empiriquement à une qualité d'équilibrage sur une variable supérieure à celle observée pour d'autres méthodes concurrentes.

L'extension à plusieurs variables d'équilibrage ou à l'échantillonnage spatial peut se faire dans le cadre de l'optimisation non linéaire, non convexe, semi-définie, c'est-à-dire l'optimisation dans l'ensemble des matrices hermitiennes contractantes à diagonale, et éventuellement spectre, fixés. La recherche des solutions pour de tels problèmes s'avère extrêmement complexe.

Loonis (2021) [2] propose une paramétrisation des matrices hermitiennes contractantes par un ensemble de paramètres indépendants transformant le problème d'optimisation semi-définie en un problème non contraint. L'application, qui à une valeur du paramètre associe une matrice hermitienne, est continue mais non différentiable pour certaines composantes, ce qui incite à mobiliser des algorithmes d'optimisation stochastiques dont le recuit-simulé (Kirpatrik et al. (1983) [3]). La très grande taille du vecteur de paramètres laisse cependant présager des difficultés liées à la *malédiction de la dimension* (Bellman (1959) [4]), c'est-à-dire des difficultés spécifiques apparaissant quand la dimension du vecteur de paramètres augmente.

Dans cette présentation, nous étudions la performance, en fonction de la taille de la population et celle de l'échantillon, du recuit-simulé pour la recherche de plans déterminantaux équilibrés et spatiaux optimaux. Nous procédons par simulation et comparons les résultats à ceux connus pour d'autres méthodes.

Bibliographie

- [1] Loonis, V., & Mary, X. (2019). Determinantal sampling designs. *Journal of Statistical Planning and Inference*, 199, 60-88.
- [2] Loonis, V. (2021). Construire tous les plans de sondage déterminantaux, Colloque francophone sur les sondages, Bruxelles, article soumis.
- [3] Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). *Optimization by simulated annealing*. *Science*, 220(4598), 671-680.
- [4] Bellman, R., & Kalaba, R. (1959). On adaptive control processes. *IRE Transactions on Automatic Control*, 4(2), 1-9.