

Résumé et Analyse de l'algorithme cyclique pour l'estimation du maximum de vraisemblance utilisant le complément de Schur

KABLAM EDJABROU ULRICH BLANCHARD

25 Mai 2025

Table des matières

1	Introduction	3
2	Résumé de l'article	3
3	Notions principales	4
3.1	Matrice symétrique définie positive	4
3.2	Maximum de vraisemblance (MLE)	4
3.3	Complément de Schur	5
3.4	Optimisation sous contraintes	5
3.5	Méthodes classiques comparées	5
4	Description de l'algorithme cyclique	5
4.1	Principe de l'algorithme cyclique proposé	5
4.2	Fonctionnement détaillé	6
4.3	Autres notions liées à l'algorithme	6
4.3.1	Loi ou distribution multinomiale (cas particulier : loi binomiale)	6
4.3.2	Gradient	6
4.3.3	Newton-Raphson (approfondissement)	7
5	Applications et résultats	7
6	Discussion	7

7	Conclusion	7
8	Références	8
9	Annexes	8
9.1	Liens pour approfondir les notions	8
9.2	Questions liées	8

1 Introduction

L'estimation paramétrique à partir de données observées est une démarche centrale en statistique et en modélisation, trouvant des applications dans une multitude de disciplines scientifiques. La méthode du maximum de vraisemblance (MLE) [?, ?, ?] s'impose comme une technique d'optimisation numérique de référence, permettant d'identifier les paramètres d'un modèle probabiliste qui maximisent la plausibilité des observations. Cette maximisation, qu'elle soit contrainte ou non, s'opère souvent sur le logarithme de la fonction de vraisemblance pour des raisons pratiques, notamment dans la modélisation de répartitions catégorielles via la loi multinomiale. Si, dans des cas simples, l'estimation des probabilités de classe découle directement des fréquences observées, la réalité des applications introduit fréquemment une dépendance à des paramètres auxiliaires eux-mêmes soumis à des contraintes.

La présence de ces paramètres auxiliaires contraints complexifie significativement l'estimation du maximum de vraisemblance. Les méthodes d'optimisation traditionnelles, telles que les algorithmes basés sur la méthode de Newton, se heurtent alors à des défis computationnels majeurs, en particulier lorsque le nombre de paramètres est élevé ou que les modèles présentent des structures de dépendance complexes. Le calcul et l'inversion répétés de matrices Hessiennes de grande dimension deviennent alors un goulot d'étranglement, limitant l'applicabilité et l'efficacité de ces approches classiques dans de nombreux contextes pratiques.

Face à ces limitations, le présent document se propose d'analyser une alternative prometteuse : un algorithme cyclique pour l'estimation du maximum de vraisemblance sous contraintes, développé par N'Guessan et Geraldo, qui tire parti des propriétés du complément de Schur. L'hypothèse sous-jacente est que cette approche peut significativement simplifier le processus d'optimisation, notamment en évitant les calculs matriciels lourds, et ainsi offrir une meilleure efficacité computationnelle par rapport aux méthodes établies. Pour vérifier cette hypothèse, nous examinerons en détail les fondements théoriques de cet algorithme, son fonctionnement étape par étape, et nous évaluerons ses performances, en particulier dans le contexte d'applications concrètes telles que l'analyse de données en sécurité routière.

2 Résumé de l'article

Voici un résumé détaillé du document "A cyclic algorithm for maximum likelihood estimation using Schur complement" d'Assi N'Guessan et Issa Cherif Geraldo, avec explications des notions principales, références citées,

et des liens vers les concepts pour approfondir.

Résumé du document original

Ce travail propose un nouvel algorithme cyclique pour l'estimation du maximum de vraisemblance sous contraintes, utilisant le complément de Schur pour résoudre le problème d'optimisation. L'algorithme est conçu pour des situations où les paramètres inconnus peuvent être séparés en deux groupes, ce qui permet une résolution alternée (par blocs) et explicite des équations d'estimation, évitant ainsi le calcul coûteux de la matrice Hessienne habituellement requis dans les méthodes classiques comme Newton-Raphson ou BFGS.

Les auteurs démontrent que leur approche est plus rapide que les méthodes traditionnelles, notamment dans des cas pratiques de grande dimension ou de contraintes complexes, et l'appliquent à l'évaluation de politiques de sécurité routière à travers des simulations numériques.

3 Notions principales

3.1 Matrice symétrique définie positive

Soit $A = (a_{ij})_{1 \leq i, j \leq n}$ une matrice réelle. A est dite :

1. Symétrique si :

$A = A^T$. Autrement dit pour tous les indices entiers i, j : $a_{i,j} = a_{j,i}$.

2. Définie positive :

Pour tout vecteur non nul $X \in \mathbb{R}^n$, on a $X^T A X > 0$.

<https://www.bibmath.net/dico/index.php?action=affiche&quoi=.m/matsympos.html>

3.2 Maximum de vraisemblance (MLE)

Méthode statistique pour estimer les paramètres d'un modèle probabiliste en maximisant la probabilité d'observer les données recueillies.

Généralement, on maximise la log-vraisemblance sous contraintes (par exemple, la somme des probabilités de classes égale à 1).

Voir la définition détaillée de la méthode du maximum de vraisemblance :

https://fr.wikipedia.org/wiki/Maximum_de_vraisemblance

3.3 Complément de Schur

Outil d'algèbre linéaire permettant de simplifier l'inversion de matrices partitionnées, crucial ici pour obtenir une solution explicite sans calculer la Hessienne entière.

En savoir plus sur le complément de Schur : <https://www.cis.upenn.edu/~jean/schur-comp.pdf>

Vidéo sur le Schur Complement : https://www.youtube.com/watch?v=3d_-klrMZLw&t=19s

3.4 Optimisation sous contraintes

Problème où l'on maximise (ou minimise) une fonction objectif sous des contraintes d'égalité ou d'inégalité sur les paramètres.

Ici, les contraintes sont linéaires (par exemple, somme des probabilités égale à 1) et de bornes (chaque probabilité entre 0 et 1).

3.5 Méthodes classiques comparées

- **Newton-Raphson** : nécessite le calcul de la Hessienne, coûteux en grande dimension.
Lien : https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Newton
- **Quasi-Newton (BFGS)** : approche la Hessienne via le gradient, plus efficace mais encore coûteuse.
Lien : https://en.wikipedia.org/wiki/Broyden%E2%80%93Fletcher%E2%80%93Goldfarb%E2%80%93Shanno_algorithm
- **Nelder-Mead** : méthode sans dérivées, adaptée à certains cas mais moins performante en présence de contraintes complexes.
Lien : https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Nelder-Mead

4 Description de l'algorithme cyclique

4.1 Principe de l'algorithme cyclique proposé

L'algorithme partitionne les paramètres en deux groupes, alternant leur estimation. Il utilise le complément de Schur pour obtenir la solution du sous-système linéaire à chaque étape, sans recourir à la Hessienne. Il converge plus rapidement dans les simulations, notamment pour des données de type multinomiale et des applications en sécurité routière.

4.2 Fonctionnement détaillé

L'algorithme cyclique proposé par N'Guessan et Geraldo fonctionne de la manière suivante :

- **Décomposition** : Les paramètres inconnus du modèle de maximum de vraisemblance sont divisés en (au moins) deux groupes (blocs).
- **Cycle d'estimation** :
 - *Étape A* : Estimer le premier groupe de paramètres en considérant le second groupe comme fixe (en utilisant leurs valeurs actuelles). Le complément de Schur est utilisé ici pour simplifier cette étape.
 - *Étape B* : Estimer le second groupe de paramètres en considérant le premier groupe comme fixe (en utilisant les valeurs nouvellement mises à jour de l'étape A).
- **Répétition** : Répéter les étapes A et B (le cycle) jusqu'à ce que les estimations des paramètres convergent (ne changent plus significativement).

4.3 Autres notions liées à l'algorithme

4.3.1 Loi ou distribution multinomiale (cas particulier : loi binomiale)

Vidéo d'explication (Anglais) : <https://www.youtube.com/watch?v=BZxCIkSkMgo>

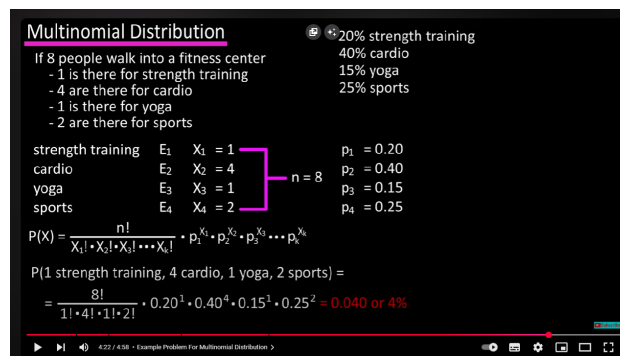


FIGURE 1 – Illustration de la loi multinomiale.

Vidéo (Anglais) : The Binomial Distribution and the Multinomial Distribution
<https://www.youtube.com/watch?v=UXB9eeMZfwo>

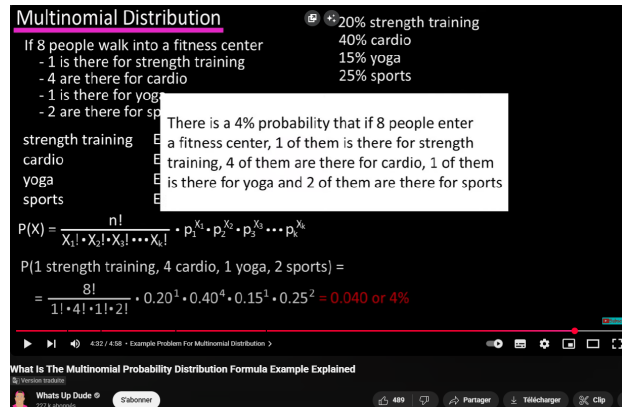


FIGURE 2 – Illustration complémentaire.

4.3.2 Gradient

Comment calculer le gradient d'une fonction : <https://www.youtube.com/watch?v=JiM2iR8gN4M>

4.3.3 Newton-Raphson (approfondissement)

Newton-Raphson Formula And Derivation | Part 1 of 2 : <https://www.youtube.com/watch?v=YS1370YMLFw>

5 Applications et résultats

L'algorithme est appliqué à la modélisation de données d'accidents de la route, pour estimer l'effet moyen d'une mesure de sécurité et les risques associés. Les simulations montrent une meilleure performance par rapport aux méthodes classiques, en particulier en termes de vitesse de convergence et de robustesse aux choix initiaux.

6 Discussion

Analysez les avantages, limites et perspectives de l'approche.

7 Conclusion

L'article propose une alternative efficace et explicite aux méthodes classiques d'estimation sous contraintes, en tirant parti du complément de Schur et

d'une démarche cyclique. Cette approche est particulièrement adaptée aux problèmes de grande dimension et à contraintes complexes, comme illustré dans le domaine de la sécurité routière.

Pour toute question sur une notion spécifique, consultez les liens ci-dessus pour des explications détaillées.

8 Références

- Méthode du maximum de vraisemblance : [1–3] dans le document (références internes à l'article)
- Revue complète des méthodes d'optimisation classiques
- 5–8 Description de l'algorithme BFGS
- Méthode de Nelder–Mead
- 10–12 Adaptations des méthodes classiques au maximum de vraisemblance sous contraintes
- 13 , algorithmes MM (Minorization–Maximization)
- Travaux antérieurs des auteurs sur la résolution par systèmes non linéaires
- 17–19 Propriétés du complément de Schur

9 Annexes

9.1 Liens pour approfondir les notions

- **Maximum de vraisemblance** : https://fr.wikipedia.org/wiki/Maximum_de_vraisemblance
- **Complément de Schur** : https://fr.wikipedia.org/wiki/Compl%C3%A9ment_de_Schur
- **Optimisation sous contraintes** : https://fr.wikipedia.org/wiki/Programmation_math%C3%A9matique
- **Méthode de Newton-Raphson** : https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Newton
- **Méthode BFGS** : https://en.wikipedia.org/wiki/Broyden%E2%80%93Fletcher%E2%80%93Goldfarb%E2%80%93Shanno_algorithm
- **Méthode Nelder-Mead** : https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Nelder-Mead
- **Algorithmes MM** : https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm#Relation_to_MM_algorithms

9.2 Questions liées

- Quelles sont les notions clés du complément de Schur et leurs applications ?
- Où puis-je trouver des ressources détaillées sur la décomposition en blocs de l'opérateur A ?
- Comment le complément de Schur facilite la résolution de systèmes linéaires complexes ?
- Quelles références académiques ou liens pour approfondir la théorie du complément de Schur ?
- Existe-t-il des exemples concrets ou études de cas illustrant l'utilisation du complément de Schur ?