

Estimateur explicite de modèles polynomiaux inverses par moindres carrés pondérés

Clément HUNEAU, François RIGAL, Jérôme IDIER

Laboratoire des Sciences du Numérique de Nantes
1, rue de la Noë, BP 92101, 44321 NANTES Cedex 31, France
clement.huneau@ls2n.fr

Résumé – Les modèles polynomiaux inverses constituent une classe de modèles dont l’estimation peut être considérée comme faiblement non linéaire. Plusieurs approches permettent de linéariser ce problème afin d’aboutir à un estimateur explicite. Nous proposons de formaliser ces approches au sein d’une famille d’estimateurs par moindres carrés pondérés. Sur la base de leur étude théorique, nous proposons un nouvel estimateur aux performances améliorées. Enfin, nous explorons le comportement de cet estimateur par des simulations d’un problème concret d’identification d’un modèle de débit sanguin cérébral.

Abstract – Inverse polynomial models are a class of models whose estimation is mildly non linear. Different methods allow linearisation of this problem in order to obtain a direct estimator. We propose a formalism to analyse these methods within a family of weighted least square estimators. Based on their theoretical study, we propose a new estimator with better performances. Finally, we explore the behavior of this new estimator through the simulation of a realistic identification problem of a cerebral blood flow model.

1 Introduction

1.1 Modèles polynomiaux inverses

Soit K observations positives et statistiquement indépendantes y_k , $k = 1, \dots, K$. On considère un modèle linéaire pour l’inverse de l’espérance :

$$\forall k, \quad E(y_k) = f_k^* = \frac{1}{\mathbf{x}_k^t \boldsymbol{\beta}^* + 1}, \quad (1)$$

avec des régresseurs à N dimensions $\mathbf{x}_k = (x_{k,1}, \dots, x_{k,N})^t$ et des paramètres inconnus $\boldsymbol{\beta}^* = (\beta_1^*, \dots, \beta_N^*)^t$ tels que $\mathbf{x}_k^t \boldsymbol{\beta}^* > -1$. La constante 1 au dénominateur est l’intercept, considéré connu dans ce modèle, qui est un cas particulier de modèle linéaire généralisé [1], avec une fonction de lien inversement linéaire aux paramètres. Son identification a été traitée sous le nom de *inverse polynomial* par Nelder [2]. Ce dernier a proposé un critère approché sur l’inverse du modèle, pondéré par le carré des observations. Dans la suite, nous généralisons ce critère en l’intégrant à une famille de critères des moindres carrés pondérés par une puissance des observations. De plus, nous montrons que l’estimateur proposé par Nelder est asymptotiquement biaisé, tandis qu’un choix différent de la pondération permet d’assurer la consistance.

1.2 Critère d’estimation

Considérant la fonction $f_k(\boldsymbol{\beta}) = (\mathbf{x}_k^t \boldsymbol{\beta} + 1)^{-1}$ non-linéaire en $\boldsymbol{\beta}$, le critère des moindres carrés classique s’écrit

$$J_K(\boldsymbol{\beta}) = \frac{1}{K} \sum_{k=1}^K (y_k - f_k(\boldsymbol{\beta}))^2. \quad (2)$$

Ce critère n’admet pas d’expression explicite pour le minimiseur en $\boldsymbol{\beta}$. Introduisons par ailleurs une famille spécifique de critères des moindres carrés pondérés, quadratiques en $\boldsymbol{\beta}$:

$$J_K(\boldsymbol{\beta}) = \frac{1}{K} \sum_{k=1}^K y_k^\alpha (y_k^{-1} - f_k^{-1}(\boldsymbol{\beta}))^2 \quad (3)$$

$$= \frac{1}{K} \sum_{k=1}^K y_k^\alpha (z_k - \mathbf{x}_k^t \boldsymbol{\beta})^2 \quad (4)$$

avec $z_k = y_k^{-1} - 1$, et α un hyperparamètre entier naturel. L’estimateur proposé dans [2] correspond au cas particulier $\alpha = 2$.

On a

$$J_K(\boldsymbol{\beta}) = \frac{1}{K} (\mathbf{z} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{A}_\alpha (\mathbf{z} - \mathbf{X}\boldsymbol{\beta}) \quad (5)$$

avec $\mathbf{z} = (z_1, \dots, z_K)^t$, $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_K)^t$ et la matrice de pondération $\mathbf{A}_\alpha = \text{Diag}\{y_k^\alpha\}$. Le gradient de J s’annule pour $\boldsymbol{\beta}$ solution de l’équation normale :

$$(\mathbf{X}^t \mathbf{A}_\alpha \mathbf{X}) \boldsymbol{\beta} = \mathbf{X}^t \mathbf{A}_\alpha \mathbf{z} \quad (6)$$

Si la matrice $(\mathbf{X}^t \mathbf{A}_\alpha \mathbf{X})$ est inversible, on obtient un estimateur explicite

$$\hat{\boldsymbol{\beta}}_\alpha = (\mathbf{X}^t \mathbf{A}_\alpha \mathbf{X})^{-1} \mathbf{X}^t \mathbf{A}_\alpha \mathbf{z} \quad (7)$$

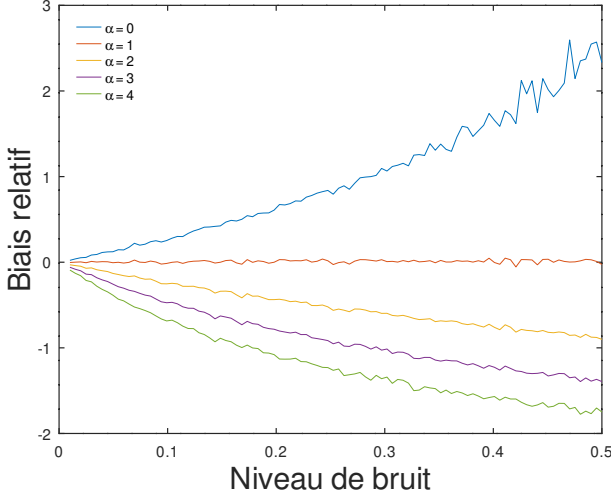


FIGURE 1 – Biais des estimateurs $\hat{\beta}_\alpha$, $N = 1$, sur $K = 100$ observations de loi gamma de moyenne f_k^* et de variance $(f_k^*)^2$.

Dans le cas sans bruit $y_k = f_k$, on a $\hat{\beta} = \beta^*$ quel que soit la valeur de α . L'étude de $\mathbb{E}(\hat{\beta}_\alpha)$ est difficile car $(\mathbf{X}^t \mathbf{A}_\alpha \mathbf{X})^{-1}$ dépend du bruit d'observation. Même dans le cas scalaire, $N = 1$, l'espérance du quotient ne donne pas d'expression évidente. Cependant, une simulation dans ce cas montre que certaines valeurs de α peuvent entraîner un biais (Figure 1). En particulier, l'estimateur $\hat{\beta}_0$ revient à un simple critère des moindres carrés sur l'inverse des observations, dégénéré pour les observations proches de 0. On peut aussi montrer que l'estimateur $\hat{\beta}_4$ correspond à un développement limité du modèle au premier ordre, mais on constate qu'il produit un biais important. L'estimateur $\hat{\beta}_2$ proposé dans l'article de Nelder présente lui aussi un biais relatif dépendant du niveau de bruit. Finalement, le comportement de l'estimateur $\hat{\beta}_1$ semble optimal du point de vue du biais dans cette simulation. A notre connaissance, il n'a jamais été proposé dans la littérature, c'est pourquoi nous proposons d'étudier théoriquement son biais asymptotique, puis de simuler son comportement dans un cadre applicatif d'identification d'un modèle de débit sanguin cérébral.

2 Propriétés asymptotiques

Ici, nous étudions la consistance de $\hat{\beta}_\alpha$ lorsque $K \rightarrow \infty$, pour $\alpha = 1$ ou $\alpha = 2$. Pour cela, nous adoptons l'hypothèse que les régresseurs \mathbf{x}_k sont des vecteurs aléatoires i.i.d. Cette hypothèse est peu réaliste en pratique, mais elle permet une analyse statistique simplifiée.

2.1 Biais asymptotique de $\hat{\beta}_1$

On peut commencer par remarquer que $\forall k$, on a :

$$\begin{aligned} \mathbb{E}[\mathbf{x}_k(1 - y_k)] &= \mathbb{E}[\mathbf{x}_k \mathbb{E}(1 - y_k | \mathbf{x}_k)] \\ &= \mathbb{E}(f_k^* \mathbf{x}_k \mathbf{x}_k^t) \beta^* \end{aligned}$$

et de la même façon,

$$\mathbb{E}(y_k \mathbf{x}_k \mathbf{x}_k^t) = \mathbb{E}(f_k^* \mathbf{x}_k \mathbf{x}_k^t).$$

D'après la loi forte des grands nombres, on montre alors :

$$\frac{1}{K} \mathbf{X}^t \mathbf{A}_1 \mathbf{z} = \frac{1}{K} \sum_{k=1}^K \mathbf{x}_k (1 - y_k) \xrightarrow[K \rightarrow \infty]{p.s.} \mathbb{E}(f_k^* \mathbf{x}_k \mathbf{x}_k^t) \beta^* \quad (8)$$

$$\frac{1}{K} \mathbf{X}^t \mathbf{A}_1 \mathbf{X} = \frac{1}{K} \sum_{k=1}^K y_k \mathbf{x}_k \mathbf{x}_k^t \xrightarrow[K \rightarrow \infty]{p.s.} \mathbb{E}(f_k^* \mathbf{x}_k \mathbf{x}_k^t). \quad (9)$$

Finalement, d'après le *continuous mapping theorem* [5, Theorem 2.3] :

$$\hat{\beta}_1 \xrightarrow[K \rightarrow \infty]{p.s.} \beta^*. \quad (10)$$

On en conclut que le choix $\alpha = 1$ donne un estimateur consistant.

2.2 Biais asymptotique de $\hat{\beta}_2$

Le cas où $\alpha = 2$ donne $\forall k$,

$$\begin{aligned} \mathbb{E}[\mathbf{x}_k (y_k - y_k^2)] &= \mathbb{E}[\mathbf{x}_k \mathbb{E}(y_k - y_k^2 | \mathbf{x}_k)] \\ &= \mathbb{E}[\mathbf{x}_k (f_k^* (1 - f_k^*) - \sigma_k^2)] \\ &= \mathbb{E}[(f_k^*)^2 \mathbf{x}_k \mathbf{x}_k^t \beta^* - \sigma_k^2 \mathbf{x}_k] \\ &= \mathbb{E}[(f_k^*)^2 + \sigma_k^2 \mathbf{x}_k \mathbf{x}_k^t] \beta^* - \mathbb{E} \left(\frac{\sigma_k^2}{f_k^*} \mathbf{x}_k \right) \end{aligned}$$

ainsi que $\mathbb{E}(y_k^2 \mathbf{x}_k \mathbf{x}_k^t) = \mathbb{E}[(f_k^*)^2 + \sigma_k^2 \mathbf{x}_k \mathbf{x}_k^t]$, avec $\sigma_k^2 = \text{Var}(y_k)$. On en conclut :

$$\hat{\beta}_2 \xrightarrow[K \rightarrow \infty]{p.s.} \beta^* - \mathbb{E}[(f_k^*)^2 + \sigma_k^2 \mathbf{x}_k \mathbf{x}_k^t]^{-1} \mathbb{E} \left(\frac{\sigma_k^2}{f_k^*} \mathbf{x}_k \right). \quad (11)$$

L'estimateur $\hat{\beta}_2$ est donc asymptotiquement biaisé si

$$\mathbb{E} \left(\frac{\sigma_k^2}{f_k^*} \mathbf{x}_k \right) \neq 0,$$

ce qui est vrai en particulier lorsque les composantes de \mathbf{x}_k sont de signe constant $\forall k$.

2.3 Normalité asymptotique $\hat{\beta}_1$

Le critère (3) s'écrit aussi

$$\begin{aligned} J_K &= \frac{1}{K} \sum_{k=1}^K y_k^\alpha (y_k^{-1} - f_k^{-1})^2 \\ &= \frac{1}{K} \sum_{k=1}^K (y_k f_k^{-2} - 2f_k^{-1}) + \text{cte.} \end{aligned} \quad (12)$$

On peut constater que l'espérance du terme général :

$$\mathbb{E}(y_k f_k^{-2}) - \mathbb{E}(2f_k^{-1}) = \mathbb{E}[f_k^* (f_k^{-1} - (f_k^*)^{-1})^2] \quad (13)$$

est minimum pour $\beta = \beta^*$. L'estimateur $\hat{\beta}_1$ rentre donc dans la catégorie des critères de contraste (ou M-estimateurs) [5]. Cela

nous permet de conclure qu'il est asymptotiquement normal tel que :

$$\sqrt{K}(\hat{\beta}_1 - \beta^*) \xrightarrow{K \rightarrow \infty} \mathcal{N}(0, \Sigma) \quad (14)$$

avec la covariance :

$$\Sigma = \mathbb{E}(f_k^* \mathbf{x}_k \mathbf{x}_k^t)^{-1} \mathbb{E} \left[\left(\frac{\sigma_k}{f_k^*} \right)^2 \mathbf{x}_k \mathbf{x}_k^t \right] \mathbb{E}(f_k^* \mathbf{x}_k \mathbf{x}_k^t)^{-1}. \quad (15)$$

Dans la suite, nous comparons le comportement des deux estimateurs $\hat{\beta}_1$ et $\hat{\beta}_2$ avec l'estimateur des moindres carrés non linéaires obtenu par minimisation itérative de (2) et avec le maximum de vraisemblance (ML), sur des données simulées à partir d'un modèle réaliste de débit sanguin.

3 Application au débit sanguin cérébral

La modélisation du débit sanguin cérébral [3] est un cas réaliste pour recourir à un modèle polynomial inverse. En effet, la résistance vasculaire est régulée à plusieurs niveaux, amont (artères) ou aval (artérioles), avec des dynamiques différentes (Figure 2, haut).

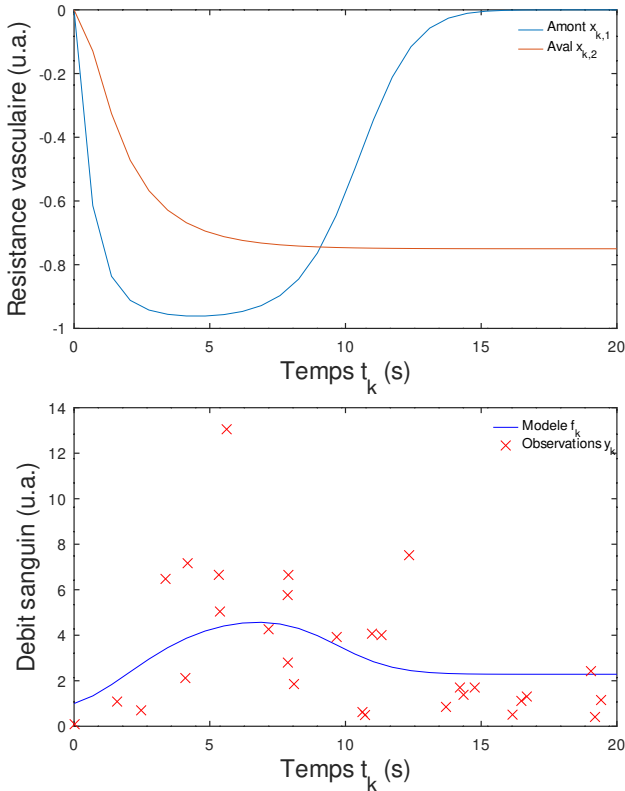


FIGURE 2 – Modèle de résistance vasculaire en série entraînant une variation du débit sanguin. (haut) Taux de variation des résistances vasculaires $x_{k,1}$, $x_{k,2}$ suivant des dynamiques différentes. (bas) Débit résultant f_k observé par y_k à des temps t_k aléatoires.

On considère ici que le débit relatif f dans le vaisseau dépend non-linéairement des variations de résistances vasculaires

$x_{k,1}$ et $x_{k,2}$ suivant un modèle du type :

$$f(t_k) = \frac{1}{\beta_1 x_{k,1}(t_k) + \beta_2 x_{k,2}(t_k) + 1} \quad (16)$$

où t_k sont les instants d'échantillonnage. À $t = 0$ les variations de résistances sont nulles, donc le débit relatif est $f(0) = 1$. Les régresseurs ont une expression supposée connue.

Dans un premier temps, afin de vérifier l'hypothèse de régresseurs stochastiques i.i.d., on suppose un échantillonnage aléatoire des t_k suivant une loi uniforme entre 0 et 20 secondes (Figure 2, bas).

3.1 Régresseurs X stochastiques i.i.d.

Nous avons étudié statistiquement les performances des deux estimateurs $\hat{\beta}_1$ et $\hat{\beta}_2$ sur des observations y_k simulées avec différents nombres d'échantillons K . Les y_k sont i.i.d. suivant une loi gamma d'espérance f_k^* , avec $\beta = (0.25, 0.75)^{-1}$ et de variance $\frac{(f_k^*)^2}{2}$. On estime la norme des biais relatifs définie par $\|\text{Diag}(\beta^*)^{-1} \mathbb{E}(\hat{\beta}_\alpha) - \mathbf{1}\|_2$ et celle des écarts-types relatifs définie de manière analogue.

En accord avec nos résultats théoriques, on constate des biais relatifs beaucoup plus faibles pour $\hat{\beta}_1$ (Figure 3). De plus, l'étude empirique de l'écart-type ne montre pas de différence sensible entre $\alpha = 1$ et $\alpha = 2$. Cela conforte le choix de $\hat{\beta}_1$ comme meilleur estimateur explicite. On remarque aussi que $\hat{\beta}_1$ semble strictement meilleur que les moindres carrés non linéaire (NLS) et très proche de l'optimalité de l'estimateur ML.

L'hypothèse d'échantillonnage aléatoire n'est pas courante en acquisition biomédicale. Dans la suite, nous étudions empiriquement le même problème d'identification en relâchant cette hypothèse.

3.2 Régresseurs X déterministe

Ici, nous avons simulé le cas plus conventionnel d'un échantillonnage régulier entre 0 et 20 secondes, alors que le reste du protocole de simulation est inchangé. On observe les mêmes comportements dans ce second cas (Figure 4). Une démonstration formelle de la consistance de $\hat{\beta}_1$ sous des hypothèses plus générales que l'hypothèse de régresseurs i.i.d. est à l'étude.

4 Conclusion

Dans cet article, nous avons défini une famille d'estimateurs linéaires explicites pour des problèmes non-linéaires particuliers de type polynomiaux inverses. Pour ces problèmes, cette famille généralise, par le réglage d'un paramètre, d'autres estimateurs classiques approchant les moindres carrés non linéaires. Nous avons d'abord montré théoriquement la consistance d'un estimateur original issu de cette famille, sous l'hypothèse de régresseurs stochastiques i.i.d.. Nos résultats sur simulation confirment que cet estimateur est performant dans un cas concret réa-

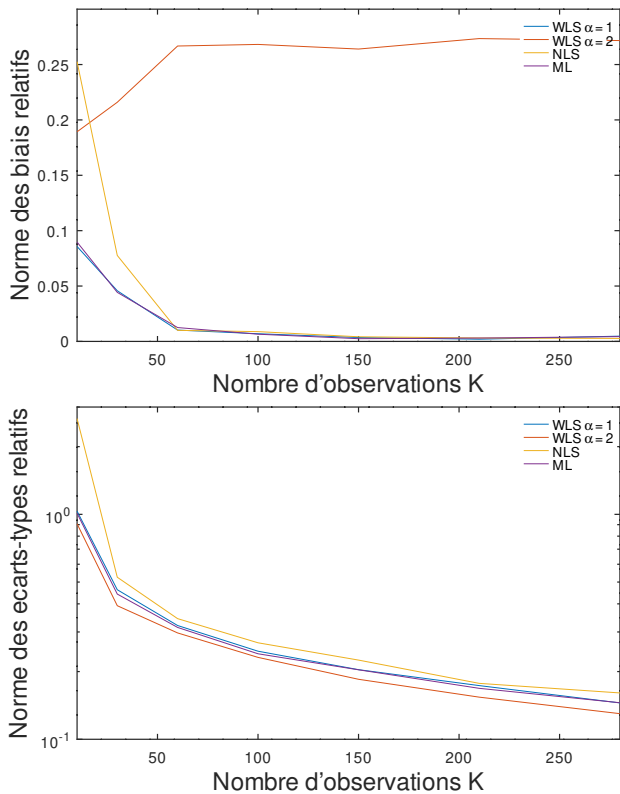


FIGURE 3 – Normes des biais et écart-type dans le cas de régresseurs i.i.d., estimées sur 500 réalisations. WLS : moindres carrés pondérés, NLS : moindres carrés non-linéaire, ML : maximum de vraisemblance.

liste. Enfin, des résultats empiriques indiquent que l’hypothèse d’indépendance sur les régresseurs n’est pas nécessaire.

Une perspective directe consiste à confirmer théoriquement que l’efficacité asymptotique de notre estimateur est proche de l’optimalité du ML. Dans les cas pour lesquels le coût du calcul itératif de la solution ML est prohibitif, notre estimateur apparaît comme un choix optimal. Dans les autres cas, il peut servir de point initial pour le calcul du ML.

Par ailleurs, une autre hypothèse sous-jacente ici est la positivité des observations y_k , ce qui exclut par exemple le cas de bruits additifs gaussiens. Dans le cas d’observations sans contrainte de signe, lorsque α est impair, la pondération du critère par y_k^α entraîne que les termes associés aux observations négatives sont concaves. Néanmoins, cette situation ne pose pas de problème asymptotique lorsque les échantillons négatifs sont minoritaires, ce qu’il est possible d’expliciter dans un cadre général.

En conclusion, cet article met en avant un nouvel estimateur explicite dont nous n’avons pas trouvé mention dans la littérature. Au-delà du modèle de débit sanguin présenté ici, cet estimateur peut être appliqué à une classe de problèmes non-linéaires que l’on rencontre assez couramment, comme par exemple, l’identification du modèle classique de cinétique de réaction enzymatique de Michaelis-Menten [4]. Une de nos perspectives sera de considérer l’estimateur $\hat{\beta}_1$ dans ce contexte,

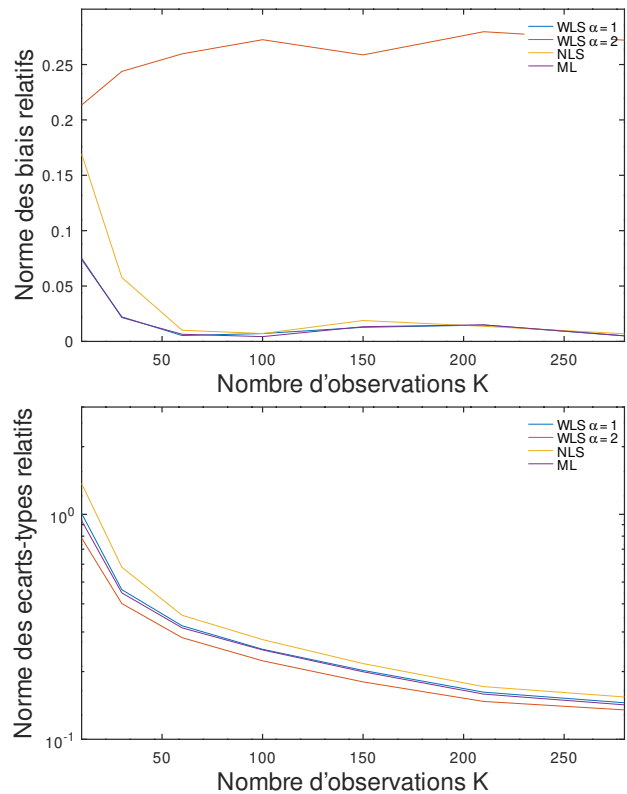


FIGURE 4 – Normes des biais et écart-type dans le cas de régresseurs déterministes, estimées sur 500 réalisations. WLS : moindres carrés pondérés, NLS : moindres carrés non-linéaire, ML : maximum de vraisemblance.

en confrontant ses performances aux solutions classiques dans ce domaine.

Références

- [1] P. McCullagh et J. A. Nelder, *Generalized Linear Models*, 2e éd., New York : Routledge, 1989.
- [2] J. A. Nelder, « Inverse Polynomials, a Useful Group of Multi-Factor Response Functions », *Biometrics*, vol. 22, n° 1, p. 128-141, 1966.
- [3] D. Mornas, « Modélisation paramétrique de l’hyperémie fonctionnelle en IRMf ASL comme un mélange de sources », thèse de doctorat, École Centrale de Nantes, 2021. <https://tel.archives-ouvertes.fr/tel-03560852>
- [4] C. T. Goudar, J. R. Sonnad, et R. G. Duggleby, « Parameter estimation using a direct solution of the integrated Michaelis-Menten equation », *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, vol. 1429, n° 2, p. 377-383, 1999.
- [5] A. W. van der Vaart, *Asymptotic Statistics*. Cambridge University Press, 1998.