
La droite de Williamson : une technique de régression linéaire injustement oubliée

François Auger, Jean-Christophe Olivier

*IREENA (Institut de Recherche en Électrotechnique et Électronique de Nantes-Atlantique)
et IUT de Saint-Nazaire, département Mesures Physiques,
CRTT, 37 Bd de l'Université, BP 406, 44602 Saint Nazaire cedex, France
{francois.auger,jean-christophe.olivier}@univ-nantes.fr*

RÉSUMÉ. L'objectif de cet article est de rappeler l'existence de la méthode de régression linéaire de Williamson, qui est la seule à prendre en compte de manière rigoureuse les incertitudes des deux coordonnées des points de mesure. Après avoir rappelé les hypothèses et les résultats de la régression linéaire classique, le principe de cette méthode est rappelé, et des expressions analytiques exactes des coefficients de la droite de régression sont présentés (alors que ces coefficients sont obtenus par une méthode d'optimisation numérique itérative dans l'article original). Une étude statistique de la qualité de ces estimations est ensuite présentée, suivie d'un exemple d'utilisation de cette technique dans un contexte réel. Nous évoquons enfin la possibilité de construire des estimateurs récursifs.

ABSTRACT. The aim of this paper is to recall the Williamson approach for straight line fitting. In this pioneering work, the case where both coordinates are blurred with noise was considered. After a short review of the classical ordinary least squares regression, this method is first recalled. We then show that in a particular case, closed form expressions of the estimated coefficients may be found. The statistical properties of these estimations are studied, followed by an example of their use in a real situation. Finally, the design of recursive estimators is addressed.

MOTS-CLÉS : régression linéaire, estimateurs du maximum de vraisemblance, estimateurs récursifs de Kalman.

KEYWORDS: linear regression, maximum likelihood estimation, recursive Kalman estimators.

1. Introduction

Il est très fréquent en physique expérimentale que deux grandeurs mesurées (et donc entachées toutes les deux d'une certaine erreur de mesure) soient théoriquement liées par une relation linéaire. Les coefficients de cette droite sont alors trop souvent obtenus par la méthode de régression classique, qui suppose que seule une des coordonnées, l'ordonnée, est entachée d'un bruit de mesure. En 1968, J.H. Williamson [WIL 68] a formulé correctement ce problème d'estimation, et proposé une solution. Une rédaction trop dense et par endroits imprécise de cet article, difficilement accessible à des non-statisticiens, n'a semble-t-il malheureusement pas permis une large diffusion de cette approche, qui répond pourtant indiscutablement à des besoins réels. Des alternatives moins rigoureuses ont alors été proposées [MAN 84, GUL 89].

L'objectif de cet article est de rappeler le principe de la méthode de régression de Williamson. Des expressions analytiques exactes des coefficients de la droite de régression sont présentées, ainsi qu'une évaluation des variances des erreurs d'estimation de ces coefficients. Dans une première partie, on va rappeler les résultats de la régression linéaire classique [AFN 97, EUR 00, PRO 99], qui suppose qu'une seule des coordonnées est entachée d'un bruit de mesure. Ce rappel permettra de préciser les hypothèses sous-jacentes. Les expressions des estimations des coefficients et de la variance de l'erreur d'estimation sont présentées, ainsi que la manière dont elles sont obtenues. Puis on présente la formulation du problème de la régression linéaire de N points dont les deux coordonnées sont mesurées [WIL 68]. Dans une troisième partie, une seconde méthode d'estimation des coefficients basée sur l'utilisation d'un estimateur d'état récursif [LEW 86] est présentée.

2. La régression linéaire classique

Le cas classique est fréquemment traité dans la littérature. Il correspond à un cas où les deux grandeurs physiques ne sont pas traitées de la même manière : l'une (l'abscisse, ou "*variable explicative*") est supposée *déterministe* et l'autre (l'ordonnée, ou "*variable expliquée*") est supposée *aléatoire*, car entachée d'un bruit de mesure additif w_i . Si \mathcal{X}_i et \mathcal{Y}_i désignent les valeurs "vraies" liées par une relation linéaire de la forme $\mathcal{Y}_i = a\mathcal{X}_i + b$ et si X_i et Y_i sont les données fournies pour le point i , alors $X_i = \mathcal{X}_i$ et $Y_i = \mathcal{Y}_i + w_i$. Si on suppose que tous les w_i sont des variables aléatoires gaussiennes de valeurs moyennes nulles et de même variance σ_Y^2 , de densité de probabilité

$$\rho(w_i) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-w_i^2/(2\sigma_Y^2)}, \quad [1]$$

et si les bruits de mesure sont supposés indépendants les uns des autres, alors le vecteur W constitué de l'ensemble des N bruits de mesure a pour densité de probabilité

$$\rho(W) = \prod_{i=1}^N \rho(w_i) = \frac{1}{(2\pi\sigma_Y^2)^{N/2}} e^{-\frac{1}{2\sigma_Y^2} \sum_{i=1}^N w_i^2}. \quad [2]$$

En écrivant alors $w_i = Y_i - \mathcal{Y}_i = Y_i - aX_i - b$ et en écrivant les données sous forme matricielle, on en déduit alors l'expression de la densité de probabilité du vecteur Y (constitué de l'ensemble des N ordonnées mesurées), dont le vecteur des abscisses X et les coefficients a et b sont des *paramètres déterministes* :

$$\rho(Y; X, a, b) = \frac{1}{(2\pi\sigma_Y^2)^{N/2}} e^{-\frac{1}{2\sigma_Y^2} \sum_{i=1}^N (Y_i - a X_i - b)^2} \quad [3]$$

$$\text{soit } \rho(Y; X, C) = \frac{1}{(2\pi\sigma_Y^2)^{N/2}} e^{-\frac{1}{2\sigma_Y^2} (Y - H C)^T (Y - H C)}, \quad [4]$$

$$\text{avec } Y = H C + W, \quad Y^T = [Y_1 \ Y_2 \ \dots \ Y_N], \quad C^T = [a \ b]$$

$$\text{et } H^T = \begin{bmatrix} X_1 & X_2 & \dots & X_N \\ 1 & 1 & \dots & 1 \end{bmatrix}, \quad W^T = [w_1 \ w_2 \ \dots \ w_N],$$

où A^T désigne la transposée de la matrice A . Pour estimer a et b , un estimateur du maximum de vraisemblance [KAY 93] est utilisé. Il correspond à chercher les valeurs des coefficients pour lesquelles les valeurs Y_i mesurées sont les plus probables, c'est à dire pour lesquelles la densité de probabilité (ou son logarithme) est maximale :

$$\frac{\partial}{\partial C} [\ln(\rho(Y; X, C))] = 0 \iff \hat{C}(Y; X) = (H^T H)^{-1} H^T Y \quad [5]$$

$$\iff \begin{cases} \hat{a} = \frac{N S_{XY} - S_X S_Y}{N S_{XX} - S_X^2} \\ \hat{b} = \frac{S_{XX} S_Y - S_X S_{XY}}{N S_{XX} - S_X^2} = \frac{S_Y - \hat{a} S_X}{N} \\ \hat{Y}_i = \hat{a} X_i + \hat{b} \end{cases} \text{ avec } \begin{cases} S_X = \sum_{i=1}^N X_i \\ S_Y = \sum_{i=1}^N Y_i \end{cases} \quad [6]$$

$$S_{XX} = \sum_{i=1}^N X_i^2, \quad S_{YY} = \sum_{i=1}^N Y_i^2, \quad \text{et } S_{XY} = \sum_{i=1}^N X_i Y_i \quad [7]$$

Sous les hypothèses indiquées, \hat{a} et \hat{b} , parfois appelés *estimateurs des moindres carrés ordinaires (ordinary least squares)*, constituent des estimateurs linéaires en Y , non biaisés (leur espérance mathématique est égale à leur valeur vraie) et efficaces : tout autre estimateur fournirait des estimations de moindre qualité, car de variances plus grandes que celles de $\hat{a}(Y; X)$ et $\hat{b}(Y; X)$:

$$\sigma_{\hat{a}}^2 = \frac{N \sigma_Y^2}{N S_{XX} - S_X^2} \quad \text{et} \quad \sigma_{\hat{b}}^2 = \frac{S_{XX} \sigma_Y^2}{N S_{XX} - S_X^2} \quad [8]$$

3. La régression linéaire lorsque les deux variables sont bruitées

Il n'est pas rare que la disparité de traitement entre les deux grandeurs X et Y , rappelée au paragraphe précédent, soit davantage justifiée par un argument de facilité que par un argument de réalité. Bien souvent, ces deux grandeurs correspondent à des mesures toutes les deux entachées de bruits additifs indépendants : $X_i = \mathcal{X}_i + v_i$ et $Y_i = \mathcal{Y}_i + w_i$. Si les calculs sont un peu plus complexes, la démarche [WIL 68] demeure la même : on suppose que les erreurs de mesure v_i et w_i sont des variables aléatoires gaussiennes de valeurs moyennes nulles et de variances respectives $\sigma_{\mathcal{X}_i}^2$ et $\sigma_{\mathcal{Y}_i}^2$ (que l'on supposera tout d'abord distinctes), de densité de probabilité conjointe

$$\rho(v_i, w_i) = \rho(v_i) \rho(w_i) = \frac{1}{2\pi\sigma_{X_i}\sigma_{Y_i}} e^{-\frac{1}{2}\left(\frac{v_i^2}{\sigma_{X_i}^2} + \frac{w_i^2}{\sigma_{Y_i}^2}\right)}, \quad [9]$$

et que les erreurs de mesure successives sont indépendantes les unes des autres. Les vecteurs V et W constitués de l'ensemble des N bruits de mesure ont alors pour densité de probabilité conjointe

$$\rho(V, W) = \prod_{i=1}^N \rho(v_i, w_i) = \frac{1}{(2\pi)^N \prod_{i=1}^N \sigma_{X_i} \sigma_{Y_i}} e^{-\frac{1}{2} \sum_{i=1}^N \left(\frac{v_i^2}{\sigma_{X_i}^2} + \frac{w_i^2}{\sigma_{Y_i}^2} \right)} \quad [10]$$

En écrivant alors $v_i = X_i - \mathcal{X}_i$ et $w_i = Y_i - \mathcal{Y}_i = Y_i - a\mathcal{X}_i - b$, on en déduit l'expression de la densité de probabilité conjointe des vecteurs X et Y , dont le vecteur des N abscisses "vraies" \mathcal{X} et les coefficients a et b sont des *paramètres déterministes* :

$$\rho(X, Y; a, b, \mathcal{X}) = \frac{1}{(2\pi)^N \prod_{i=1}^N \sigma_{X_i} \sigma_{Y_i}} e^{-\frac{1}{2} \sum_{i=1}^N \left(\frac{(X_i - \mathcal{X}_i)^2}{\sigma_{X_i}^2} + \frac{(Y_i - a\mathcal{X}_i - b)^2}{\sigma_{Y_i}^2} \right)} \quad [11]$$

Pour construire des estimateurs de ces $N + 2$ paramètres, la technique du maximum de vraisemblance peut encore être utilisée. En dérivant par rapport à \mathcal{X}_i , on obtient

$$\frac{\partial}{\partial \mathcal{X}_i} [\ln(\rho(X, Y; a, b, \mathcal{X}))] = 0 \iff \begin{cases} \hat{\mathcal{X}}_i = X_i + \frac{\hat{a} \lambda_i^2}{\hat{a}^2 \lambda_i^2 + 1} (Y_i - \hat{a} X_i - \hat{b}) \\ \hat{\mathcal{Y}}_i = Y_i - \frac{1}{\hat{a}^2 \lambda_i^2 + 1} (Y_i - \hat{a} X_i - \hat{b}) \end{cases}, \quad [12]$$

avec $\lambda_i = \sigma_{X_i}/\sigma_{Y_i}$. On peut vérifier que si $\lambda_i = 0$, on retrouve $\hat{\mathcal{X}}_i = X_i$ et $\hat{\mathcal{Y}}_i = \hat{a} X_i + \hat{b}$. En dérivant par rapport à b on obtient, après quelques calculs, une équation,

$$\frac{\partial}{\partial b} [\ln(\rho(X, Y; a, b, \mathcal{X}))] = 0 \iff \sum_{i=1}^N \frac{Y_i - \hat{\mathcal{Y}}_i}{\sigma_{Y_i}^2} = \sum_{i=1}^N \frac{Y_i - \hat{a} X_i - \hat{b}}{\hat{a}^2 \lambda_i^2 + 1} = 0, \quad [13]$$

dont on peut déduire une expression de \hat{b} en fonction de \hat{a} et des mesures X_i et Y_i . Si les λ_i sont constants, cette équation se ramène simplement à $\hat{b} = (S_Y - \hat{a} S_X)/N$. Enfin, en dérivant par rapport à a , on obtient

$$\frac{\partial}{\partial a} [\ln(\rho(X, Y; a, b, \mathcal{X}))] = 0 \iff \sum_{i=1}^N \frac{\hat{\mathcal{X}}_i (Y_i - \hat{\mathcal{Y}}_i)}{\sigma_{Y_i}^2} = 0 \quad [14]$$

Dans le cas général, la solution de cette équation n'a pas d'expression simple, et doit être approchée par des algorithmes numériques. Dans le cas où $\sigma_{X_i} = \sigma_X$ et $\sigma_{Y_i} = \sigma_Y$, elle se réduit par contre à une équation du second degré

$$\lambda^2 S_{xy} \hat{a}^2 + (S_{xx} - \lambda^2 S_{yy}) \hat{a} - S_{xy} = 0, \quad [15]$$

avec $S_{xx} = S_{XX} - S_X^2/N$, $S_{xy} = S_{XY} - S_X S_Y/N$ et $S_{yy} = S_{YY} - S_Y^2/N$. Cette équation ayant un discriminant positif, elle possède deux racines réelles, de signes

opposés puisque leur produit $(-1/\lambda^2)$ est négatif. En s'appuyant sur le fait que a doit être de même signe que S_{xy} , on en déduit que

$$\hat{a} = \frac{-(S_{xx} - \lambda^2 S_{yy}) + \sqrt{(S_{xx} - \lambda^2 S_{yy})^2 + 4\lambda^2 S_{xy}^2}}{2\lambda^2 S_{xy}} \quad [16]$$

On peut vérifier que lorsque $\lambda = 0$, l'expression 16, à notre connaissance originale, se ramène à $\hat{a} = S_{xy}/S_{xx}$. Comme le suggère Mandel [MAN 84], cette valeur peut être approchée numériquement en calculant la limite des suites $\hat{a}_{n+1} = (S_{xy} + \alpha_n S_{yy})/(S_{xx} + \alpha_n S_{xy})$ et $\alpha_{n+1} = \lambda^2 \hat{a}_{n+1}$, initialisées par $\alpha_0 = 0$. Leur convergence est généralement très rapide. Pour connaître les caractéristiques de cet estimateur et déterminer sa zone de pertinence, une étude statistique a été effectuée sur $N_r = 50000$ réalisations des bruits de mesure. Les valeurs $a = 10$, $b = 3$, $\mathcal{X}_i = i$ ont été choisies arbitrairement. Les figures 1.a et 1.b présentent respectivement les valeurs moyennes $\sum_{j=1}^{N_r} \hat{a}_j/N_r$, $\sum_{j=1}^{N_r} \hat{b}_j/N_r$ et les écarts quadratiques moyens $\sum_{j=1}^{N_r} (\hat{a}_j - a)^2/N_r$, $\sum_{j=1}^{N_r} (\hat{b}_j - b)^2/N_r$ obtenus en prenant $\lambda = 1$, $\sigma_Y = 0.6$ et en faisant varier le nombre de points de mesure N . Elles montrent que le biais de l'estimateur proposé est faible lorsque $N > 5$ (celui de l'estimateur classique étant plus important, surtout pour l'estimateur de b) et que les variances de \hat{a} et \hat{b} ont une décroissance en $1/N^3$ et $1/N$ respectivement. Les figures 1.c et 1.d présentent respectivement les valeurs moyennes et les écarts quadratiques moyens obtenus en prenant $N = 15$, $\lambda = 1$ et en faisant varier σ_Y^2 de 10^{-4} à 10. Elles montrent que le biais de l'estimateur proposé est sensiblement indépendant de σ_Y (celui de l'estimateur classique devenant important lorsque $\sigma_Y > 1$) et que les variances de \hat{a} et \hat{b} sont sensiblement proportionnelles à σ_Y^2 . Enfin, les figures 1.e et 1.f présentent respectivement les valeurs moyennes et les écarts quadratiques moyens obtenus en prenant $N = 15$, $\sigma_Y = 1$ et en faisant varier λ de 10^{-3} à 1 (si $\lambda > 1$, il semble préférable de permuter les données X_i et Y_i). Elles montrent que lorsque λ augmente (> 0.1), l'erreur d'estimation de l'estimateur classique devient importante, et les variances de \hat{a} et \hat{b} deviennent sensiblement proportionnelles à λ^2 . Cette étude montre donc que l'estimateur proposé est particulièrement intéressant lorsque $5 \leq N \leq 30$, $\sigma_Y^2 > 0.1$ et $\lambda > 0.1$. Elle montre aussi que si le nombre de points de mesure est supérieur à 50 et si les deux bruits de mesure sont faibles, l'estimateur classique fournira des résultats de bonne qualité.

Enfin, les figures 1.g et 1.h présentent les valeurs moyennes des estimations de a et b calculées en utilisant une valeur de λ différente de la valeur réelle $\lambda_{\text{vrai}} = 1$. Ces figures montrent qu'il n'est pas nécessaire de connaître la valeur exacte de λ pour obtenir de bonnes estimations. Une valeur approchée, de préférence par excès, fournira des résultats satisfaisants.

4. Un exemple d'utilisation

Une application de l'estimateur proposé au test inter-laboratoire présenté dans [MAN 84] permettrait d'illustrer ses possibilités d'utilisation et de comparer les approches et les outils proposés. Nous avons choisi de présenter ici son utilisation pour la validation d'une nouvelle carte électronique de mesure des courants de sortie d'un on-

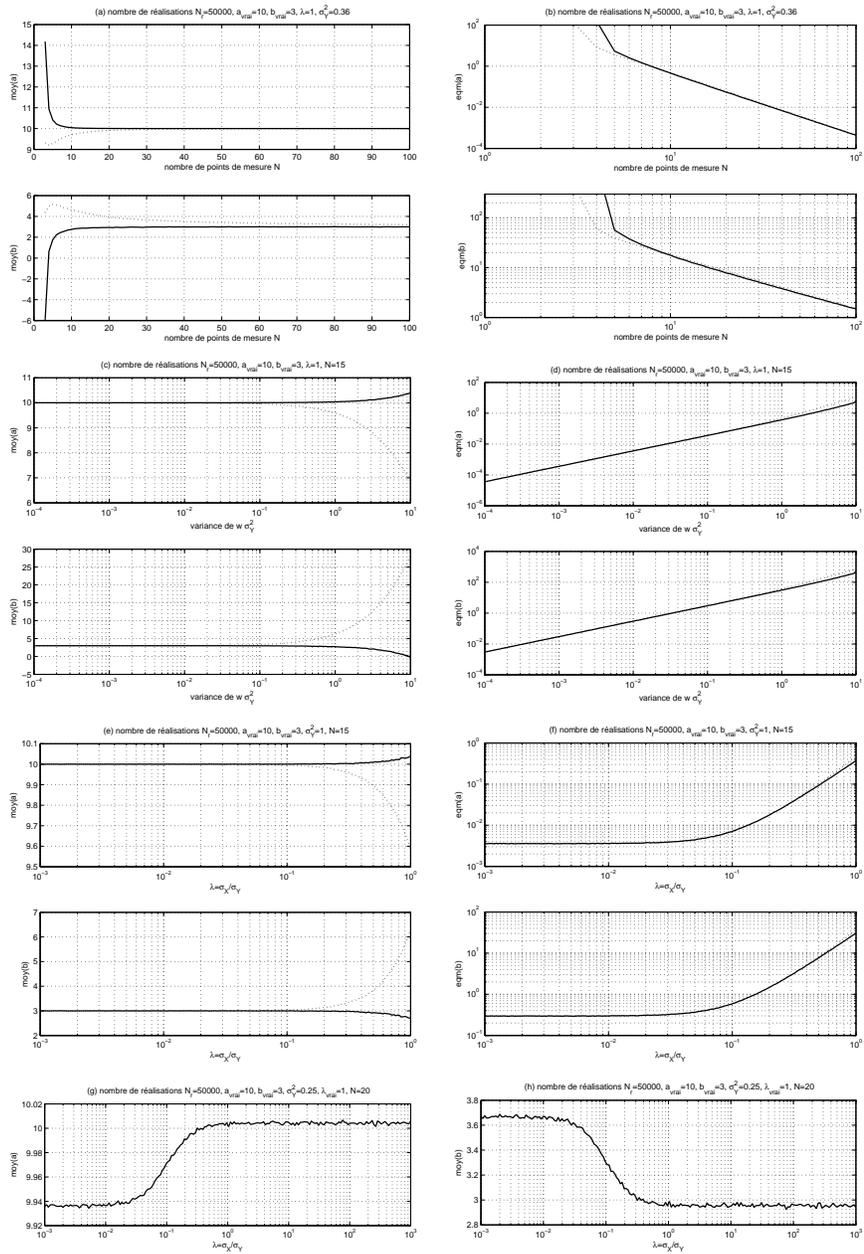


Figure 1 – Résultats de l'étude statistique expérimentale de l'estimateur proposé.

daleur de puissance triphasé, conçue au laboratoire afin de mesurer des courants allant jusqu'à 100 A avec une bande passante d'au moins 100 kHz. Pour cela, des capteurs LEM LAH 100-P ont été utilisés. Les résultats de mesure obtenus avec ce système ont été comparés à ceux obtenus avec une pince de courant Tektronix A 6303 associée à un module amplificateur AM 503B. Le tableau ci-dessous regroupe les mesures effectuées :

X_i (A)	-79.900	-60.440	-42.600	-23.100	-0.118
Y_i (A)	-79.940	-60.710	-43.350	-23.410	-0.281
X_i (A)	0.286	20.580	37.280	60.580	79.460
Y_i (A)	0.013	20.460	37.580	60.750	80.290

Une étude des caractéristiques métrologiques des deux systèmes de mesure a permis de retenir la valeur $\lambda = 3/16$. Dans un tel contexte, où la donnée X est bruitée et où le nombre de points de mesure est faible, l'étude statistique précédente a montré que la méthode proposée offrait une meilleure estimation. On peut donc penser que les valeurs $\hat{a} = 1.00591733$ et $\hat{b} = -0.05788270$ obtenues avec la méthode proposée conduisent à une meilleure estimation de l'erreur de gain et de l'erreur de décalage que les valeurs $\hat{a} = 1.00591624$ et $\hat{b} = -0.05788357$ obtenues avec l'estimateur classique.

5. Estimateurs récursifs

Dans certains cas, il peut être intéressant de disposer des estimations de a et b au fur et à mesure que de nouveaux points de mesure sont apportés. Ceci peut être obtenu en rendant évolutives les sommations des expressions 6 et 7, en utilisant soit une fenêtre de taille croissante, soit une fenêtre glissante de taille constante, soit un facteur d'oubli. Une autre possibilité consiste à utiliser un estimateur récursif construit à partir d'un modèle d'état. En notant $x_1[n] = a$, $x_2[n] = b/a$, $x_3[n] = \mathcal{X}_n + b/a = X_n + b/a - v_n$, on peut considérer que ces variables évoluent de manière aléatoire

$$X_{[n+1]} = A X_{[n]} + B X_{n+1} + V_{[n+1]}, \quad \text{avec} \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad [17]$$

$$B^T = (0 \ 0 \ 1), \quad \text{et} \quad X^T = (x_1 \ x_2 \ x_3), \quad [18]$$

et sont associées à une équation d'observation non linéaire déduite de $Y_{n+1} = a(\mathcal{X}_n + b/a) + w_{n+1}$

$$Y_{[n+1]} = Y_{n+1} = \mathcal{H}(X_{[n+1]}) + w_{[n+1]}, \quad \text{avec} \quad \mathcal{H}(X) = x_1 x_3. \quad [19]$$

L'estimateur de Kalman [LEW 86, GRE 08] déduit de ce modèle non linéaire observable comprend une première phase de prédiction,

$$\begin{cases} X_p[n+1] = A X_e[n] + B X_{n+1} \\ P_p[n+1] = A P_e[n] A^T + Q \end{cases} \quad \text{avec} \quad Q = \begin{pmatrix} q_a & 0 & 0 \\ 0 & q_{b/a} & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix}, \quad [20]$$

suivie d'une phase de correction de la prédiction à l'aide de la mesure. Dans le cas d'un filtre de Kalman étendu, cette phase est déduite d'une approximation de Taylor au premier ordre de $H(X)$ autour de la prédiction :

$$K[n+1] = P_p[n+1] H[n+1]^T (H[n+1] P_p[n+1] H[n+1]^T + 1)^{-1} \quad [21]$$

$$X_e[n+1] = X_p[n+1] + K[n+1] (Y[n+1] - H(X_p[n+1])) \quad [22]$$

$$P_e[n+1] = P_p[n+1] - K[n+1] H[n+1] P_p[n+1] \quad [23]$$

$$H[n+1] = [x_{3p}[n+1] \quad 0 \quad x_{1p}[n+1]] \quad [24]$$

L'équation de mesure étant bilinéaire, il est également possible de concevoir et de mettre en œuvre le filtre de Kalman non linéaire optimal, en s'appuyant par exemple sur des approximations de Stirling au second ordre [AUG 04]. Le manque de place ne permet cependant pas de présenter ici cet estimateur et de fournir ses caractéristiques statistiques.

6. Conclusion

L'objectif principal de cet article était de rappeler la technique d'estimation des coefficients d'une droite de régression lorsque les deux variables sont entachées d'un bruit additif. Du cas général étudié par Williamson, nous avons extrait un cas particulier, correspondant à incertitudes constantes en X et en Y . En délivrant des estimations plus pertinentes que l'estimateur basique lorsque le nombre de points de mesure est faible et lorsque les données en X sont significativement bruitées, l'estimateur proposé peut contribuer à l'amélioration dans des situations concrètes de la qualité des estimations, à coût constant. Un outil logiciel basé sur les résultats présentés ici renforcerait encore davantage cette contribution.

7. Bibliographie

- [AFN 97] AFNOR, *Métrologie : gérer et maîtriser les équipements de mesure*, 1997.
- [AUG 04] AUGER F., SAHEB T., LORON L., « Utilisation des approximations de Stirling du premier et du second ordre pour l'évaluation de l'estimation et de l'incertitude-type d'une grandeur indirectement mesurée », *Colloque C2I 2004, volume 1*, Éditions Hermès, 2004, p. 37-44.
- [EUR 00] EURACHEM, CITAC, *Quantifying Uncertainty in Analytical Measurement*, EURACHEM / CITAC, 2nd édition, 2000.
- [GRE 08] GREWAL M., ANDREWS A., *Kalman theory, theory and practice using MATLAB*, Wiley, 3rd édition, 2008.
- [GUL 89] GULL S., « Bayesian data analysis : straight-line fitting », *J. Skilling ed, "Maximum entropy and Bayesian methods"*, Kluwer Academic Press, 1989, p. 511–518.
- [KAY 93] KAY S. M., *Fundamentals of Statistical Signal Processing : Estimation Theory*, Prentice Hall, 1993.
- [LEW 86] LEWIS R., *Optimal Estimation with an Introduction to Stochastic Control Theory*, John Wiley & Sons, Inc., 1986.
- [MAN 84] MANDEL J., « Fitting straight lines when both variables are subject to error », *Journal of quality technology*, vol. 16, n° 1, 1984, p. 1-13.
- [PRO 99] PROTASSOV K., *Probabilités et incertitudes dans l'analyse des données expérimentales*, Presses Universitaires de Grenoble, 1999.
- [WIL 68] WILLIAMSON J., « Least-squares fitting of a straight line », *Canadian Journal of Physics*, vol. 46, 1968, p. 1845-1847.