

REVUE DE STATISTIQUE APPLIQUÉE

J. RAGOT

M. AUBRUN

Application de la régression orthogonale sous contrainte linéaire à un problème d'équilibrage de bilan-matière

Revue de statistique appliquée, tome 30, n° 2 (1982), p. 45-56

http://www.numdam.org/item?id=RSA_1982__30_2_45_0

© Société française de statistique, 1982, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

APPLICATION DE LA REGRESSION ORTHOGONALE SOUS CONTRAINTE LINEAIRE A UN PROBLEME D'EQUILIBRAGE DE BILAN-MATIERE

J. RAGOT et M. AUBRUN

RESUME

Après avoir rappelé le problème de l'estimation des paramètres à partir de mesures bruitées, on traite le cas particulier des modèles linéaires avec et sans contraintes sur les paramètres.

On applique cet estimateur à la résolution d'un problème d'équilibrage de bilan-matière pour un système industriel.

La plupart des systèmes physiques sont caractérisés par un certain nombre de variables mesurables; chercher un modèle qui représente ce processus consiste à déterminer une relation entre :

- une variable à expliquer Y^* (ou variable dépendante),
- n variables explicatives X_i^* (ou variables indépendantes).

De façon globale Y^*, X_1^*, \dots, X_n^* sont des vecteurs colonnes dont les composantes sont les valeurs prises par les variables correspondantes lors d'observations effectuées sur le système.

Lorsque le système est considéré comme linéaire, le modèle est mis sous la forme :

$$Y^* = X^* \alpha \quad (1)$$

$(N, 1) \quad (N, n) \quad (n, 1)$

avec : $X^* = (X_1 \dots X_n)$

et où : $Y^*, X_i^* \in R^N, \alpha \in R^n$

Pratiquement, les grandeurs X^* et Y^* sont inaccessibles: on ne dispose en effet que de mesures X et Y de ces grandeurs, telles que :

$$Y = Y^* + \epsilon_y$$

$$X = X^* + \epsilon_x$$

ϵ_y et ϵ_x , matrices de dimensions respectives; $N.1$ et $N.n$ qui représentent les erreurs de mesure. La relation (1) se traduit par :

$$y = X\alpha + \epsilon_y - \epsilon_x \alpha \quad (2)$$

où $X\alpha$ traduit la liaison fonctionnelle introduite par le modèle et $\epsilon = \epsilon_y - \epsilon_x \alpha$ les "fluctuations" dues aux erreurs de mesure.

Comment alors estimer les paramètres α ?

La méthode la plus simple consiste à rendre la norme du vecteur e minimale : cette façon de procéder revient à reporter sur la variable à expliquer y toutes les erreurs de mesure, ce qui revient à chercher la loi conditionnelle de la variable, à expliquer connaissant les variables explicatives.

Une méthode plus générale consiste à tenir compte simultanément des erreurs sur les variables X_i et Y : cette solution est présentée ici et porte le nom d'estimation des moindres carrés étendus.

1. L'ESTIMATION DES MOINDRES CARRÉS ETENDUS (REGRESSION ORTHOGONALE)

1.1. Formulation algébrique de l'estimation

Comme les erreurs de mesure affectent simultanément toutes les variables, on ne fait pas de distinction de notation entre la variable à expliquer et les variables explicatives.

Soit :

W^* la matrice $(N, n + 1)$ des grandeurs réelles

W la matrice $(N, n + 1)$ des grandeurs mesurées

où on a placé les vecteurs Y^* , Y et \hat{Y} (estimation de Y) en dernière colonne.

Le système physique est décrit par la relation linéaire :

$$W^* A^* = 0 \quad (3)$$

où le vecteur A^* des paramètres est de dimension $(n + 1)$.

Le problème qu'on se propose de résoudre consiste donc à déterminer les estimateurs \hat{A} de A^* et \hat{W} de W^* qui minimisent la somme des carrés des écarts entre l'estimation et la mesure :

$$\Phi = 1/2 \operatorname{Tr} \|\hat{W} - W\|^2 \quad (4)$$

sous la contrainte :

$$\hat{W} \hat{A} = 0 \quad (5)$$

Il correspond en fait à la méthode usuelle connue sous le nom de régression orthogonale (cf. paragraphe 1.4) ; on peut le résoudre en recherchant l'extrémum du Lagrangien :

$$L = \frac{1}{2} \operatorname{Tr} \|\hat{W} - W\|^2 + \lambda^T \hat{W} \hat{A} \quad (6)$$

où λ est le paramètre de Lagrange tel que $\lambda \in \mathbb{R}^N$.

La solution est la suivante :

– \hat{A} est le vecteur propre de $W^T W$ associé à sa plus petite valeur propre (qu'on désigne par s) :

$$(W^T W - sI) \hat{A} = 0 \quad (7)$$

– La matrice des estimations des grandeurs réelles est :

$$\hat{W} = W \left(I - \frac{\hat{A}\hat{A}^T}{\hat{A}^T\hat{A}} \right) \quad (8)$$

– Le critère résiduel est égal à :

$$\frac{s}{2} = \frac{1}{2} \frac{\hat{A}^T W^T W \hat{A}}{\hat{A}^T \hat{A}} \quad (9)$$

1.2. Cas du modèle explicite

Il peut être intéressant de différencier les variables explicatives de la variable à expliquer en utilisant les partitions suivantes :

$$W = (x \ y) \quad \hat{W} = (\hat{x} \ \hat{y})$$

$$A = \begin{pmatrix} a \\ -1 \end{pmatrix}$$

avec :

$$W^T W = \begin{pmatrix} X^T X & X^T y \\ y^T X & y^T y \end{pmatrix}$$

$$\hat{A}^T W^T W \hat{A} = (X\hat{a} - y)^T (X\hat{a} - y)$$

$$\hat{A}^T \hat{A} = 1 + \hat{a}^T \hat{a}$$

L'équation aux valeurs propres $(W^T W - sI) \hat{A} = 0$ se décompose sous la forme :

$$(X^T X - sI) \hat{a} - X^T y = 0 \quad (10)$$

$$-y^T X \hat{a} + (y^T y - s) = 0 \quad (11)$$

où :

$$s = \frac{(X\hat{a} - y)^T (X\hat{a} - y)}{1 + \hat{a}^T \hat{a}} \quad (12)$$

Quelques transformations élémentaires de ces relations donnent en particulier :

$$[(X^T + \hat{a}y^T) (y - X\hat{a})] = 0 \quad (13)$$

$$\left[X^T X - \frac{(y^T - \hat{a}^T X^T) (y - X\hat{a})}{1 + \hat{a}^T \hat{a}} \ I \right] \hat{a} = X^T y \quad (14)$$

Ces formes sont à comparer à celles que l'on obtient en effectuant la régression usuelle de la variable y par rapport aux variables X :

$$X^T (y - X\hat{a}) = 0 \quad (15)$$

$$X^T X \hat{a} = X^T y \quad (16)$$

1.3. Propriétés des écarts

– Les écarts $E = W - \hat{W}$ s'expriment avec la définition (8) des estimations sous la forme :

$$E = W \frac{\hat{A}\hat{A}^T}{\hat{A}^T\hat{A}} \quad (17)$$

– Les écarts E sont orthogonaux aux estimés :

$$E^T \hat{W} = 0$$

– Somme des carrés des écarts (SCE).

Compte tenu de (9), on a :

$$SCE = \text{Tr}(E^T E) = \frac{\hat{A}^T W^T W \hat{A}}{\hat{A}^T \hat{A}} \quad (18)$$

En notant par :

SCM la somme des carrés des mesures = $\text{Tr}(W^T W)$

SCR la somme des carrés des estimées = $\text{Tr}(\hat{W}^T \hat{W})$

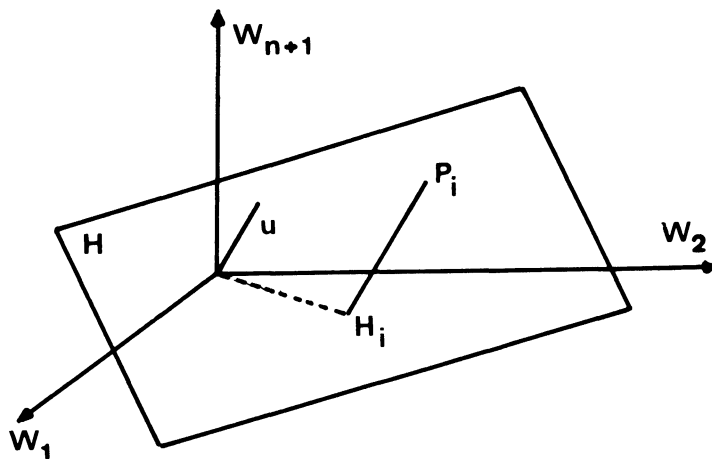
on a :
$$SCM = SCR + SCE \quad (19)$$

Comme dans le cas de la régression de y par rapport à X on peut, pour cet estimateur étendu, définir un "indice de qualité" du modèle par R tel que :

$$R^2 = \frac{SCR}{SCM} \quad (20)$$

1.4. Formulation géométrique

On représente dans l'espace des variables les observations P_i ($i = 1, \dots, N$).



L'hypothèse du modèle linéaire se traduit par l'existence d'un plan H de vecteur normal unitaire u représentatif du nuage de points constitué par les observations P_i .

L'équation du plan qui représente le nuage est donnée par les cosinus directeurs de u, que l'on peut par exemple obtenir en minimisant la somme des carrés des distances des mesures P_i à ce plan (ce qui correspond à la régression orthogonale, la régression usuelle revenant à minimiser la somme des carrés des écarts parallèlement à l'axe W_{n+1}).

Si W_i est la $i^{\text{ème}}$ ligne de la matrice des mesures W, donc formée par les composantes du point P_i , $W_i u$ est la distance de ce point au plan $(P_i H_i)$. Pour l'ensemble des points, la somme des carrés des distances au plan H est :

$$\Phi = \|Wu\|^2$$

Le vecteur unitaire u qui minimise Φ est le vecteur propre de $W^T W$ associé à sa plus petite valeur propre, la valeur de Φ étant égale à cette valeur propre.

2. CAS AVEC CONTRAINTE SUR LES PARAMETRES

Dans certains problèmes, les paramètres des modèles sont liés par des relations parfaitement connues ; on examine ce cas en le limitant aux contraintes de type égalité mise sous la forme :

$$T \cdot A = 0 \quad \text{où} \quad \dim. \text{ de } T = p(n+1) \quad (21)$$

avec :

$$p < n+1$$

$$\text{rang } T = p$$

Soit P le projecteur dans R^{n+1} sur le sous-espace d'équation $TA = 0$

$$P = I - T^T (TT^T)^{-1} T$$

Si l'on pose $W^* = W P$, on est ramené au problème précédent. Ainsi \hat{A} est vecteur propre de la matrice $W^{*T} W^*$ associé à sa plus petite valeur propre.

Il faut remarquer la nature singulière de la matrice P qui est de dimension $(n+1)(n+1)$, mais de rang p ; pour cette raison, la matrice $\hat{W}^T \hat{W}^*$ est également singulière et ses p dernières valeurs propres sont nulles.

Pour éviter cette singularité, il suffit de réduire la dimension du problème initial en ne prenant en compte dans le vecteur A que les composantes indépendantes. A cet effet, on utilise les partitionnements :

$$T = (T_1 : T_2) \quad \text{où} \quad T_1 : p \cdot (n+1-p) \quad (22)$$

$$T_2 : p \cdot p \text{ régulière}$$

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \quad \text{où} \quad A_1 : (n+1-p) \cdot 1 \quad (23)$$

$$W = (W_1 : W_2) \quad \text{où} \quad \begin{array}{l} A_2 : p \cdot 1 \\ W_1 : N \cdot (n + 1 - p) \\ W_2 : N \cdot p \end{array} \quad (24)$$

La contrainte sur les paramètres s'explicité alors :

$$A_2 = -T_2^{-1} T_1 A_1 \quad (25)$$

et le modèle :

$$(\hat{W}_1 - \hat{W}_2 T_2^{-1} T_1) A_1 = 0 \quad (26)$$

Le Lagrangien (6) s'écrit alors :

$$L = 1/2 \text{Tr} \|\hat{W}_1 - W_1\|^2 + 1/2 \text{Tr} \|\hat{W}_2 - W_2\|^2 + \lambda^T (\hat{W}_1 - \hat{W}_2 T_2^{-1} T_1) A_1 \quad (27)$$

Il est extrémal quand :

$$\begin{aligned} L_{\hat{W}_1} &= \hat{W}_1 - W_1 + \lambda A_1^T &= 0 \\ L_{\hat{W}_2} &= \hat{W}_2 - W_2 - \lambda A_1^T T_1^T (T_2^T)^{-1} &= 0 \\ L_{\lambda} &= (\hat{W}_1 - \hat{W}_2 T_2^{-1} T_1) A_1 &= 0 \\ L_{A_1} &= (\hat{W}_1 - \hat{W}_2 T_2^{-1} T_1)^T \lambda &= 0 \end{aligned} \quad (28)$$

En posant :

$$\left. \begin{aligned} \tilde{W} &= W_1 - W_2 T_2^{-1} T_1 \\ P &= I + T_1^T (T_2^T)^{-1} T_2^{-1} T_1 \end{aligned} \right\} \quad (29)$$

on obtient :

$$\left(\tilde{W}^T \tilde{W} - \frac{A_1^T \tilde{W}^T \tilde{W} A_1}{A_1^T P A_1} P \right) A_1 = 0 \quad (30)$$

Avec la décomposition :

$$P = H H$$

et les changements :

$$B_1 = H A_1 \quad (31)$$

$$\bar{W} = \tilde{W} H^{-1}$$

l'équation (30) prend la forme :

$$\boxed{\left(\bar{W}^T \bar{W} - \frac{B_1^T \bar{W}^T \bar{W} B_1}{B_1^T B_1} \cdot I \right) B_1 = 0} \quad (32)$$

Ainsi B_1 est vecteur propre de la matrice $\bar{W}^T \bar{W}$ et comme le montre l'examen du critère, il est associé à la plus petite valeur propre de cette matrice.

3. APPLICATION A UN PROBLEME DE BILAN-MATIERE

Le problème de l'équilibrage des bilans-matière est longuement exposé dans la littérature spécialisée.

On ne rappelle ici que quelques définitions élémentaires à partir d'un exemple très simple décrit par la figure 1.

Les Q_{ei} représentent des flux d'entrée d'un produit ou d'une énergie, les Q_{si} les flux de sortie. L'équation de ces bilans s'écrit :

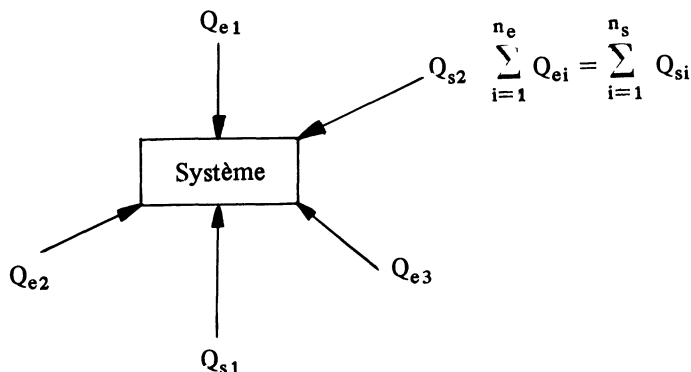


FIG. 1. — Bilan massique d'un système simple

D'une façon générale, un bilan traduit, pour un système en équilibre, des conservations de masse, de débit, d'énergie. Ainsi pour le système de la figure 1, supposé en équilibre, on a écrit l'égalité de la somme des débits entrants et celle des débits sortants; en réalité, la présence d'erreurs de mesure déséquilibre cette équation de bilan.

Equilibrer un bilan consiste donc à détecter au mieux ces erreurs et à en déduire une estimation des grandeurs vraies qui, elles, vérifient les équations de bilan.

Dans la suite on ne considère que le cas d'erreurs de mesures considérées comme des réalisations de variables aléatoires, centrées, non corrélées entre elles et de fonction de répartition gaussienne.

Pour le type de système schématisé à la figure 1, on est très fréquemment confronté au problème suivant :

chaque flux de matière est caractérisé, en plus de son débit Q_i , par les concentrations C_{ij} en différents constituants ($j = 1, \dots, c$) (où c est le nombre de constituants considérés).

En régime d'équilibre, on doit avoir pour les grandeurs vraies :

$$\sum_{i=1}^n Q_i^* = 0 \quad (33)$$

$$\sum_{i=1}^n Q_i^* C_{ij}^* = 0 \quad j = 1, \dots, c \quad (34)$$

qui expriment la conservation des débits totaux et des débits par constituants (les débits entrants sont comptés positivement et les débits sortants, négativement).

En général, sur de tels systèmes, on dispose des mesures en concentration C_{ij} (obtenues par analyse chimique sur des échantillons, ou en ligne, par des analyseurs à fluorescence fonctionnant de façon quasi instantanée); par contre, il est très rare de disposer de toutes les mesures de débit Q_i pour des raisons technologiques (produits trop visqueux ou trop chargés en air). On se place dans cette configuration où la matrice des observations W est constituée par les mesures C_{ij} et où le vecteur A des paramètres est constitué par les variables débits.

Les équations (33) et (34) se reformulent :

$$TA^* = 0 \quad (35)$$

$$W^*A^* = 0 \quad (36)$$

Le problème d'équilibrage de bilan-matière se formule ainsi : déterminer l'estimateur (\hat{W}, \hat{A}) de (W^*, A^*) qui minimise la somme des carrés des écarts :

$$\Phi = 1/2 \text{ Tr } \|\hat{W} - W\| \quad (37)$$

sous les contraintes :

$$\hat{W} \hat{A} = 0 \quad (38)$$

$$T \hat{A} = 0 \quad (39)$$

La solution de ce problème a été donnée au paragraphe 2. Compte tenu de la forme particulière de la contrainte (33), on a ici :

$$T = (1, 1, \dots, 1) \quad (40)$$

qui se partitionne, compte tenu des notations du paragraphe 2, en :

$$T_1 = (1, 1, \dots, 1) \quad T_2 = (1) \quad (41)$$

On évalue P (équation 29) :

$$P = I + T_1^T T_1$$

(où T_1 est un vecteur ligne de dimension n), puis sa factorisation (équation 31) en posant :

$$H = I + h^T h \quad (42)$$

(où h est un vecteur ligne),

$$\begin{aligned} \text{d'où:} \quad 1 + T^T T_1 &= (1 + h^T h) (1 + h^T h) \\ &= I + h^T (2 + hh^T) h \end{aligned}$$

Comme hh^T est un scalaire, on obtient :

$$h \sqrt{2 + hh^T} = T_1 \quad (43)$$

qui montre que h et T_1 sont coplanaires.

Dans (43), si l'on pose $h = \alpha T_1$ on a: (44)

$$\alpha^4 \cdot n + 2 \alpha^2 - 1 = 0$$

d'où:
$$\alpha = \left(\frac{-1 + \sqrt{1+n}}{n} \right)^{1/2} \quad (45)$$

La matrice H est donc entièrement définie par (43), (44) et (45).

Pour expliciter (31), il faut calculer H^{-1} ; un calcul classique montre que si:

$$H = I + h^T h$$

alors:
$$H^{-1} = I - h^T (1 + h h^T)^{-1} h \quad (46)$$

En définitive, à partir de la matrice d'observation $W = (W_1 \ W_2)$, la procédure de calcul est la suivante:

- Calcul de $\tilde{W} = W_1 - W_2 T_2^{-1} T_1$ (équation 29)

$$H = I + h^T h \quad (\text{équation 42, 44, 45})$$

$$H^{-1} = I - \frac{h^T h}{1 + h h^T} \quad (\text{équation 46})$$

$$\bar{W} = \tilde{W} H^{-1} \quad (\text{équation 31})$$

- Recherche du vecteur propre B_1 de $\bar{W}^T \bar{W}$ associé à sa plus petite valeur propre (32)

- Calcul de $A_1 = H^{-1} B_1$ (équation 31)

$$\text{Calcul de } A_2 = -T_2^{-1} T_1 A_1 \quad (\text{équation 25})$$

Exemple : Equilibrage de bilans-matière sur un séparateur.

Dans les circuits de traitement de minerais, le séparateur classe les particules en deux catégories: celles suffisamment petites pour être utilisées par les étages de purification ou d'enrichissement et celles qui doivent subir un retraitement mécanique (broyage). Un tel séparateur est schématisé à la figure 2; on y distingue la voie d'entrée ou alimentation et les voies de sortie désignées par surverse (fines particules) et sousverse (grosses particules).

Pour repérer la taille des grains dans un échantillon, on définit g_j le pourcentage en masse des grains dont la taille t est comprise entre t_j et t_{j+1} (les valeurs des t_j sont standardisées); les débits massiques de solides sont désignés par la lettre Q.

En régime d'équilibre, les grandeurs vraies doivent vérifier les équations de bilan-matière:

$$Q_1^* + Q_2^* + Q_3^* = 0 \quad (47)$$

$$Q_1^* g_{1j}^* + Q_2^* g_{2j}^* + Q_3^* g_{3j}^* = 0 \quad j = 1, \dots, n$$

où n représente le nombre de "tranches granulométriques".

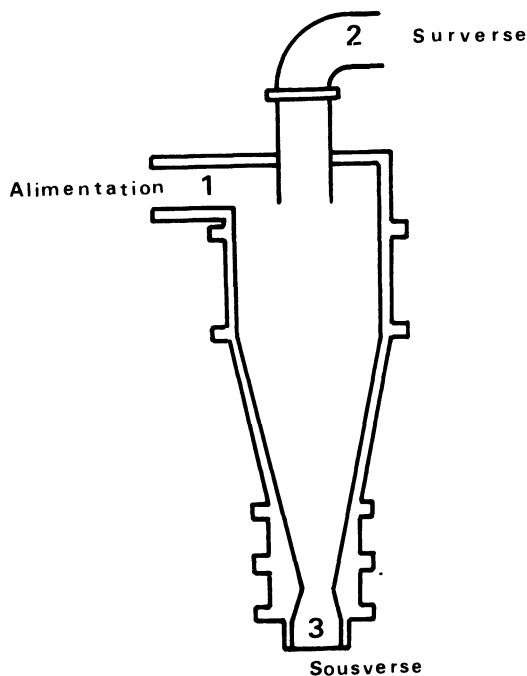


Fig. 2. — Séparateur

Dans la pratique, on définit souvent le coefficient de partage du débit ou encore le rendement poids :

$$r^* = \frac{Q_2^*}{Q_1^*} \quad (48)$$

Compte tenu de la relation de conservation des débits, on a :

$$1 + r^* = - \frac{Q_3^*}{Q_1^*} \quad (49)$$

l'équation de conservation des débits par tranche granulométrique s'écrit :

$$g_{1j}^* + g_{2j}^* r^* - (1 + r^*) g_{3j}^* = 0 \quad (50)$$

Ainsi, le rendement poids r^* peut s'évaluer par :

$$r^* = - \frac{g_{1j}^* - g_{3j}^*}{g_{2j}^* - g_{3j}^*} \quad (51)$$

On se place dans le cas où seules les mesures de granulométrie sont accessibles. La relation (51), appliquée à des mesures g_{ij} , donne en général des estimations aberrantes (par exemple supérieures à 1 ou de signes variables). Le tableau 1 est relatif à des mesures effectuées sur un hydrocyclone Krebs de 4 pouces; les trois premières colonnes donnent la granulométrie mesurée à l'alimentation, la surverse et la sousverse ; la quatrième colonne fait apparaître le rendement poids

TABLEAU 1
Mesures et rendement poids

Alimentation	Surverse	Sousverse	Rendement poids × 100
14,8	37,5	13,0	7,3
2,0	4,4	1,8	7,7
1,7	3,5	1,5	10,0
2,3	5,8	2,0	7,9
3,0	7,9	1,8	19,0
3,8	20,0	3,5	1,8
7,3	2,1	6,5	18,0
12,9	11,2	14,0	39,0
22,0	5,9	22,6	3,6
19,9	1,1	22,5	12,0
9,3	0,3	9,7	4,0

calculé par la formule (51). Le tableau 2 donne les estimations granulométriques obtenues après "équilibrage" des bilans-matière en appliquant l'algorithme des moindres carrés étendus ; comme pour le tableau des données mesurées, le quatrième colonne indique le rendement poids dont les différentes estimations sont maintenant parfaitement cohérentes.

TABLEAU 2
Mesures et rendement poids

Alimentation	Surverse	Sousverse	Rendement poids × 100
18,88	36,14	10,27	33,2
2,43	4,25	1,52	33,2
2,99	3,40	1,30	33,2
2,92	5,60	1,58	33,2
3,54	7,72	1,45	33,2
7,20	19,17	1,23	33,2
5,84	2,58	7,47	33,2
13,01	11,17	13,93	33,2
18,81	6,96	24,72	33,2
17,0	2,06	24,44	33,2
7,55	0,89	10,87	33,2

CONCLUSION

L'estimateur des moindres carrés étendus est fondé sur l'hypothèse de mesures toutes affectées d'erreurs, sans distinguer la variable à expliquer des variables explicatives. On l'a comparé à l'estimateur obtenu par régression de la variable à expliquer par rapport aux variables explicatives, où toutes les erreurs de mesure sont attribuées à la seule variable à expliquer.

Outre l'approche plus réaliste quant à la répartition des erreurs, sa mise en œuvre est simple puisque les calculs ne font appel qu'à un programme standard de recherche de valeur et vecteur propres.

Enfin, la portée de cet estimateur a été étendue au traitement de cas où l'on doit prendre en compte des contraintes linéaires sur les paramètres.

BIBLIOGRAPHIE

- [1] J. BRENOT (1977). – Contribution à la pratique du modèle linéaire : qualité, protection et estimation biaisée. Thèse de 3^e Cycle, Paris, 88 pages.
- [2] H.J. BRITT and R.H. LUECKE (1973). – The estimation of parameters in non-linear implicit models. *Technometrics*, vol. 15, n° 2, p. 233-247.
- [3] D. MARQUARDT (1970). – Generalized inverse, ridge regression, biased linear estimation and non linear estimation. *Technometrics*, vol. 12, n° 3, p. 591-612.
- [4] D.R. POWELL and J.R. MacDONALD (1972). – A rapidly convergent iterative method for the solution of the generalized non-linear least-squares problem, *The Computer Journal*, vol. 15, p. 148-155.
- [5] W.H. SOUTHWELL (1969). – Fitting experimental data. *J. of Comp. Phys.*, vol. 4, p. 465-474.
- [6] W.H. SOUTHWELL (1976). – Fitting data to non-linear functions with uncertainties in all measurement variables. *The Computer Journal*, vol. 19, n° 1, p. 69-73.
- [7] W.J. WHITEN (1977). – Multiple curve fitting. International report n° 84. Julius Kruttschnitt Mineral Research Centre. Department of Mining and Metallurgical Engineering, University of Queensland.
- [8] J.H. WILLIAMSON (1968). – Least-squares fitting of a straight line. *Can. J. of Phys.*, vol. 46, p. 1845-1847.
- [9] D. YORK (1966). – Least-squares fitting of a straight line. *Can. J. of Phys.*, vol. 44, p. 1079-1086.