

ANNALES DE L'I. H. P., SECTION B

M. TOURNARIE

Exploitation d'une observation isolée pour l'estimation de plusieurs inconnues

Annales de l'I. H. P., section B, tome 5, n° 1 (1969), p. 49-67

http://www.numdam.org/item?id=AIHPB_1969__5_1_49_0

© Gauthier-Villars, 1969, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section B » (<http://www.elsevier.com/locate/anihpb>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). 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/>

Exploitation d'une observation isolée pour l'estimation de plusieurs inconnues

par

M. TOURNARIE

Service de Physique du Solide et de Résonance Magnétique
Centre d'Études Nucléaires de Saclay
B. P. n° 2, 91-Gif-sur-Yvette.

RÉSUMÉ. — Voici une méthode pour améliorer l'évaluation de plusieurs inconnues à partir d'une seule observation. La méthode est stable même quand l'observation probable est une application non linéaire des inconnues. Ceci permet la résolution des systèmes linéaires sur petite calculatrice et le pilotage automatique des ensembles de mesure en temps réel.

ABSTRACT. — Here is a method to improve the estimate of several unknowns from single observation. The method is stable, even when the probable observation is a non linear application of the unknowns. It allows the resolve of linear systems on small computer and the automatic command management of the measuring units in real time.

PLAN

Préambule	50
Notations	51
Principe des estimations successives	51
Optimisation immédiate	54
Approximation du 2 ^e ordre	56
Stabilité	58
Système d'équations linéaires	60
Solution du système précédent	60
Convergence de la solution	61
Mesures à temps constant	62
Estimateur infinitésimal	63

Mesures à nombre constant de quanta	63
Estimateur univantique	65
Conclusion	65
Applications possibles	66

PRÉAMBULE

L'algorithme optimal [2] décrit la méthode pour obtenir les meilleures estimations z^* d'un système de n inconnues à partir de N observations indirectes. Pour cela on considère d'une part que les N perturbations des observations sont engendrées par une transformation (non linéaire éventuellement) inversible d'un N -tuple s_y^* suivant une loi normale, d'autre part que les n inconnues sont engendrées elles-mêmes par une transformation inversible (souvent non linéaire) d'un n -tuple s_x^* normal. L'algorithme optimal définit les meilleures estimations comme celles qui correspondent à la probabilité maximale du $(N + n)$ -tuple décrivant les N observations.

Lorsque les transformations sont linéaires, l'algorithme optimal conduit à évaluer z^* par une combinaison linéaire des N observations y^* . L'opérateur qui réalise cette combinaison est dit opérateur optimal. Il est fonction de l'estimation préalable des inconnues et de leur matrice de covariance.

En effet, on écrit alors :

$$z^* = z_0^* + X s_x^*$$

où z_0^* représente l'évaluation initiale des inconnues, et où X est la matrice représentative de la transformation engendrant les inconnues à partir du n -tuple s_x^* . D'où

$$s_x^* = X^{-1}[x^* - z_0^*].$$

Les observations s'écrivent $y^* = Ax^* + Ys_y^*$ où A est la matrice permettant de calculer l'espérance mathématique des observations à partir des inconnues, et Y est la matrice représentative de l'application linéaire engendrant les perturbations des observations à partir du N -tuple s_y^* . D'où :

$$s_y^* = Y^{-1}[y^* - Ax^*].$$

L'algorithme optimal cherche le z^* qui substitué à x^* maximise la probabilité de s_x^* et s_y^* , c'est-à-dire qui minimise $s_x^* s_x^* + s_y^* s_y^*$. On trouve

$$z^* = [I + X\tilde{X}\tilde{A}W\tilde{A}W]^{-1}X\tilde{X}\tilde{A}W[y^* - Az_0^*].$$

Si l'on fait une approximation linéaire de cas non linéaires, on peut penser que les résultats seront d'autant meilleurs que les estimations z seront voisines des hypothèses d'estimation *a priori*. Cette particularité se présente surtout lorsque les nouvelles observations apportent peu d'information. Or une observation seule apporte effectivement peu d'information. Mais une seule observation fait un système mal conditionné. Toutefois, on sait que l'opérateur optimal peut résoudre un système mal conditionné.

Nous nous sommes alors posé les questions suivantes :

1) Les estimations fournies par l'algorithme optimal peuvent-elles servir d'estimations préalables pour la résolution d'un nouveau système d'observations.

2) Le nouveau système d'observations peut-il être réduit à une seule observation sans que la dégénérescence du système conduise à des divergences irréductibles.

Nous montrerons dans cet article que la réponse est oui aux deux questions.

NOTATIONS

Dans la notation, le vecteur colonne x_1, x_2, \dots, x_n est écrit x . Le vecteur ligne composé des mêmes éléments est noté \bar{x} . Le produit scalaire de x et de y s'écrit $\bar{x}y$ et on peut noter que :

$$\bar{x}y \equiv \sum_i x_i y_i \equiv \sum_i y_i x_i \equiv y\bar{x}$$

La construction $x''y$, par contre, désigne une matrice dont les éléments ij sont $x_i y_j$, de sorte que $x''y \neq y''x$. Une lettre majuscule en caractères gras désigne une matrice. Il s'ensuit que $\mathbf{E}x$ est un vecteur colonne, $\bar{y}\mathbf{E}$ est un vecteur ligne et $\bar{y}\mathbf{E}x$ est un scalaire.

Nous réservons la lettre f pour désigner la fonction qui nous intéresse, z désigne l'ensemble de ses variables.

PRINCIPE DES ESTIMATIONS SUCCESSIVES

THÉORÈME. — Dans un système linéaire, les matrices de covariance des estimations optimales des inconnues sont les mêmes, que les observations soient exploitées globalement ou successivement, lorsque les observations sont sans corrélation de second ordre.

Preuve. — Pour démontrer ce théorème nous allons décomposer l'ensemble y' des observations en deux sous-ensembles y'_1 et y'_2 .

Soient E_0 la matrice de covariance *a priori*, W la matrice de pondération et A la matrice du système linéaire.

On sait que la covariance optimale est dans le cas global :

$$E = [I + E_0 \tilde{A} W A]^{-1} E_0. \quad (1)$$

Aux sous-ensembles y'_1 et y'_2 correspondent les sous-matrices W_1 et W_2 ainsi que A_1 et A_2 . L'indépendance des observations permet d'écrire

$$W = \begin{bmatrix} W_1 & O \\ O & W_2 \end{bmatrix}. \quad (2)$$

Comme

$$A = [A_1 A_2] \quad (3)$$

on déduit

$$E = [I + E_0 \tilde{A}_1 W_1 A_1 + E_0 \tilde{A}_2 W_2 A_2]^{-1} E_0. \quad (4)$$

Dans le cas successif, on estimera d'abord :

$$E_1 = [I + E_0 \tilde{A}_1 W_1 A_1]^{-1} E_0 \quad (5)$$

puis

$$E_2 = [I + E_1 \tilde{A}_2 W_2 A_2]^{-1} E_1. \quad (6)$$

En mettant la première égalité dans la deuxième il vient :

$$E_2 = [I + [I + E_0 \tilde{A}_1 W_1 A_1]^{-1} E_0 A_2 W_2 A_2]^{-1} [I + E_0 \tilde{A}_1 W_1 A_1]^{-1} E_0. \quad (7)$$

En groupant les inverses, cela donne :

$$E_2 = [[I + E_0 \tilde{A}_1 W_1 A_1][I + [I + E_0 \tilde{A}_1 W_1 A_1]^{-1} E_0 A_2 W_2 A_2]]^{-1} E_0. \quad (8)$$

En effectuant le produit au dénominateur

$$E_2 = [I + E_0 A_1 W_1 A_1 + E_0 A_2 W_2 A_2]^{-1} E_0. \quad (9)$$

Ce qui montre que $E_2 = E$ C. Q. F. D. (10)

THÉORÈME. — Dans un système linéaire, les estimations optimales sont les mêmes que les observations soient exploitées globalement ou successivement.

Preuve. — Nous procéderons comme pour le théorème précédent.

z_0 désigne les estimations *a priori*,

z_1 les estimations après les premières observations,

z_2 les estimations après le second sous-ensemble d'observations.

On sait que

$$z' = z_0' + \mathbf{E}\tilde{\mathbf{A}}\mathbf{W}[y' - \mathbf{A}z_0'] \quad (11)$$

pour les estimations globales.

Les estimations successives sont :

$$z_1' = z_0' + \mathbf{E}_1\tilde{\mathbf{A}}_1\mathbf{W}_1[y_1' - \mathbf{A}_1z_0'] \quad (12)$$

$$z_2' = z_1' + \mathbf{E}_2\tilde{\mathbf{A}}_2\mathbf{W}_2[y_2' - \mathbf{A}_2z_1'] \quad (13)$$

En portant la première égalité (12) dans la seconde (13) il vient :

$$z_2' = z_0' + \mathbf{E}_1\mathbf{A}_1\mathbf{W}_1[y_1' - \mathbf{A}_1z_0'] + \mathbf{E}_2\mathbf{A}_2\mathbf{W}_2[y_2' - \mathbf{A}_2\{z_0' + \mathbf{E}_1\tilde{\mathbf{A}}_1\mathbf{W}_1[y_1' - \mathbf{A}_1z_0']\}] \quad (14)$$

soit :

$$z_2' = z_0' + [\mathbf{I} - \mathbf{E}_2\tilde{\mathbf{A}}_2\mathbf{W}_2\mathbf{A}_2]\mathbf{E}_1\tilde{\mathbf{A}}_1\mathbf{W}_1[y_1' - \mathbf{A}_1z_0'] + \mathbf{E}_2\tilde{\mathbf{A}}_2\mathbf{W}_2[y_2' - \mathbf{A}_2z_0'] \quad (15)$$

En multipliant à gauche les deux membres de l'égalité (6) par

$$\mathbf{I} + \mathbf{E}_1\mathbf{A}_2\mathbf{W}_2\mathbf{A}_2,$$

il vient :

$$[\mathbf{I} + \mathbf{E}_1\tilde{\mathbf{A}}_2\mathbf{W}_2\mathbf{A}_2]\mathbf{E}_2 = \mathbf{E}_1. \quad (16)$$

En effectuant le produit et en retirant $\mathbf{E}_1\tilde{\mathbf{A}}_2\mathbf{W}_2\mathbf{A}_2\mathbf{E}_2$ aux deux membres de l'égalité (16), il vient :

$$\mathbf{E}_2 = \mathbf{E}_1 - \mathbf{E}_1\tilde{\mathbf{A}}_2\mathbf{W}_2\mathbf{A}_2\mathbf{E}_2. \quad (17)$$

En mettant \mathbf{E}_1 en facteur et en transposant il vient :

$$[\mathbf{I} - \mathbf{E}_2\tilde{\mathbf{A}}_2\mathbf{W}_2\mathbf{A}_2]\mathbf{E}_1 = \mathbf{E}_2. \quad (18)$$

Portons maintenant l'égalité précédente (18) dans (15) pour faire disparaître \mathbf{E}_1 , il vient :

$$z_2' = z_0' + \mathbf{E}_2\tilde{\mathbf{A}}_1\mathbf{W}_1[y_1' - \mathbf{A}_1z_0'] + \mathbf{E}_2\tilde{\mathbf{A}}_2\mathbf{W}_2[y_2' - \mathbf{A}_2z_0']. \quad (19)$$

Or, \mathbf{E}_2 peut se mettre en facteur et vaut \mathbf{E} d'après le théorème (10) précédent. D'autre part l'autre facteur vaut, d'après la décomposition des sous-matrices :

$$\tilde{\mathbf{A}}_1\mathbf{W}_1[y_1' - \mathbf{A}_1z_0'] + \tilde{\mathbf{A}}_2\mathbf{W}_2[y_2' - \mathbf{A}_2z_0'] = \mathbf{A}\mathbf{W}[y' - \mathbf{A}z_0']. \quad (20)$$

Donc :

$$z_2' = z'. \quad \text{C. Q. F. D.} \quad (21)$$

OPTIMISATION IMMÉDIATE

Ainsi chaque observation peut donner lieu à une modification linéaire immédiate optimale des estimations, ce qui permet de choisir au mieux les conditions de l'observation suivante.

Les nouvelles estimations tirées de la $k^{\text{ième}}$ observation s'écrivent en fonction des anciennes

$$z_k^* = z_{k-1}^* + b_0^* + y_k b_1^*. \quad (22)$$

La matrice d'erreur se calcule en admettant que :

$$x^* = z_{k-1}^* + Xs^* \quad (X \text{ dépend de } k) \quad (23)$$

et que l'appareil fournit des observations dont la valeur probable est une fonction non linéaire des inconnues :

$$\bar{y}_k = a_k(x^*). \quad (24)$$

La valeur observée diffère de la valeur probable d'une quantité aléatoire σs où σ est une constante et s une variable aléatoire d'espérance nulle et de variance unité

$$y_k = a_k(z_{k-1}^* + Xs^*) + \sigma s. \quad (25)$$

En introduisant (25) dans (22) on obtient donc :

$$z_k^* = z_{k-1}^* + b_0^* + [a_k(z_{k-1}^* + Xs^*) + \sigma] b_1^*. \quad (26)$$

La relation liant les nouvelles estimations aux anciennes peut donc se mettre sous la forme :

$$z^* = z^* + b_0^* + f b_1^* \quad (27)$$

où f est une fonction scalaire dépendant de la dernière observation :

$$f = a_k(z_{k-1}^* + Xs^*) + \sigma s. \quad (28)$$

Nous allons calculer la nouvelle matrice de covariance de z^* . Pour cela il nous faut $z^* - x^*$, ce qu'on obtient en retirant x^* des deux côtés de la relation (27) précédente

$$z^* - x^* = z^* - x^* + b_0^* + f b_1^*. \quad (29)$$

Or, d'après (23) $z^* - x^* = -Xs^*$ où $XX^* = E$ par définition.

On a donc :

$$z^* - x^* = -Xs^* + b_0^* + f b_1^*. \quad (30)$$

La matrice de covariance est la matrice probable correspondant à :

$$[z' - x'] [z - x] = Xs''s\tilde{X} - [b_0''s\tilde{X} + Xs''b_0] - f[b_1''s\tilde{X} + Xs''b_1] + b_0''b_0 + f[b_0''b_1 + b_1''b_0] + f^2b_1''b_1. \quad (31)$$

On sait que

$$\tilde{s}'' = \mathbf{0} \quad \tilde{s}''\tilde{s} = \mathbf{I} \quad (32) \quad (33)$$

f dépend de la dernière observation, donc de s par l'intermédiaire de x . $\tilde{f}\tilde{s}$ n'est donc pas forcément nul.

Donc :

$$E' = \overline{[z' - x'] [z - x]} = E - [b_1''\tilde{f}\tilde{X} + X\tilde{f}\tilde{s}''b_1] + b_0''b_0 + \tilde{f}[b_0''b_1 + b_1''b_0] + \tilde{f}^2b_1''b_1 \quad (34)$$

Nous allons choisir b_0^* et b_1^* pour rendre minimale toute forme quadratique construite sur E' . C'est-à-dire résoudre :

$$\frac{\partial}{\partial b_0^*} uE'u = \mathbf{0} \quad (35 a)$$

$$\frac{\partial}{\partial b_1^*} uE'u = \mathbf{0} \quad (35 b)$$

1^{re} équation. — En laissant les termes indépendants de b_0^* , il reste :

$$\frac{\partial}{\partial b_0^*} uE'u = \frac{\partial}{\partial b_0^*} [b_0^*u''ub_0^* + 2\tilde{f}b_0^*u''ub_1^*] = \mathbf{0} \quad (36)$$

ou encore :

$$\frac{\partial}{\partial b_0^*} uE'u = u''u[b_0^* + \tilde{f}\tilde{b}_1^*] = \mathbf{0}. \quad (37)$$

Cette équation sera satisfaite si :

$$b_0^* = -\tilde{f}\tilde{b}_1^*. \quad (38)$$

2^e équation. — En remplaçant dans (34) b_0^* par son expression (38) en fonction de b_1^* , il vient :

$$E' = E - [b_1''\tilde{f}\tilde{X} + X\tilde{f}\tilde{s}''b_1] + [\tilde{f}^2 - \tilde{f}^2]b_1''b_1. \quad (39)$$

En laissant les termes indépendants de \mathbf{b}_1 , il vient :

$$\frac{\partial}{2\partial \mathbf{b}_1} \mathbf{u}' \mathbf{E}' \mathbf{u}' = \frac{\partial}{2\partial \mathbf{b}_1} [2' \mathbf{b}_1 \mathbf{u}'' \mathbf{X} \widetilde{\mathbf{f}} \widetilde{\mathbf{s}} + [\widetilde{f}^2 - \widetilde{f}^2]' \mathbf{b}_1 \mathbf{u}'' \mathbf{u} \mathbf{b}_1'] = \mathbf{0}' \quad (40)$$

ou encore :

$$\frac{\partial}{2\partial \mathbf{b}_1} \mathbf{u}' \mathbf{E}' \mathbf{u}' = \mathbf{u}'' \mathbf{u} [\mathbf{X} \widetilde{\mathbf{f}} \widetilde{\mathbf{s}} + [\widetilde{f}^2 - \widetilde{f}^2] \mathbf{b}_1'] = \mathbf{0}' \quad (41)$$

Cette équation sera satisfaite quel que soit \mathbf{u}'' si

$$\mathbf{b}_1' = \frac{1}{\widetilde{f}^2 - \widetilde{f}^2} \mathbf{X} \widetilde{\mathbf{f}} \widetilde{\mathbf{s}} \quad (42)$$

Ce qui donne en portant (42) dans (38)

$$\mathbf{b}_0' = - \frac{f}{\widetilde{f}^2 - \widetilde{f}^2} \mathbf{X} \widetilde{\mathbf{f}} \widetilde{\mathbf{s}} \quad (43)$$

et en portant dans \mathbf{E}' et dans \mathbf{z}'

$$\mathbf{E}' = \mathbf{E} - \frac{1}{\widetilde{f}^2 - \widetilde{f}^2} \mathbf{X} \widetilde{\mathbf{f}} \widetilde{\mathbf{s}} \widetilde{\mathbf{s}}' \widetilde{\mathbf{f}}' \mathbf{X} \quad (44)$$

$$\mathbf{z}' = \mathbf{z}' + \frac{f - \widetilde{f}}{\widetilde{f}^2 - \widetilde{f}^2} \mathbf{X} \widetilde{\mathbf{f}} \widetilde{\mathbf{s}} \quad (45)$$

APPROXIMATION AU DEUXIÈME ORDRE

Il est rare qu'on puisse exprimer explicitement les formes analytiques de f , \widetilde{f}^2 , $\widetilde{f}\widetilde{s}$. Il faut donc faire souvent appel à des développements numériques. Mais ces derniers ne peuvent pas être poussés très loin sinon ils deviennent extrêmement lourds et demandent la connaissance des 4^{es} moments des fonctions de probabilité des inconnues, lesquels sont la plupart du temps mal connus. Par contre les développements poussés jusqu'au deuxième ordre inclus sont simples et ne font appel qu'à la matrice de covariance des inconnues. f se développe alors comme

$$f = a_0 + \sigma_{s_0} + \mathbf{a}' \mathbf{X} \mathbf{s}' + \frac{1}{2} \mathbf{s}' \widetilde{\mathbf{X}} \mathbf{A} \mathbf{X} \mathbf{s}' + 0(s^3). \quad (46)$$

En se rappelant que

$$\overline{s_0} = 0, \quad \overline{s} = \mathbf{0} \quad \text{et} \quad \overline{s\tilde{X}AXs} = \text{Tra}(\mathbf{AE}), \quad (47)$$

on voit que :

$$\overline{f} = a_0 + \frac{1}{2} \text{Tra}(\mathbf{AE}) + 0(s^3). \quad (48)$$

D'autre part

$$f^2 = a^2_0 + 2a_0\sigma s_0 + 2a_0\mathbf{aXs} + a_0\mathbf{s\tilde{X}AXs} + \sigma^2 s^2_0 + 2\sigma s_0\mathbf{s\tilde{X}a} + \mathbf{aXs}\mathbf{s\tilde{X}a} + 0(s^3). \quad (49)$$

Comme

$$s^2_0 = 1 \quad (50), \quad \overline{s_0s} = \mathbf{0} \quad (51) \quad (50) \quad (51)$$

et

$$\overline{s\tilde{X}s} = \quad (52), \quad \text{d'où} \quad \overline{\mathbf{aXs}\mathbf{s\tilde{X}a}} = \mathbf{aEa} \quad (53) \quad (52) \quad (53)$$

on voit que :

$$\overline{f^2} = a^2_0 + a_0 \text{Tra}(\mathbf{AE}) + \sigma^2 + \mathbf{aEa} + 0(s^3). \quad (54)$$

Or d'après (48)

$$f^2 = a^2_0 + a_0 \text{Tra}(\mathbf{AE}) + 0(s^3). \quad (55)$$

Donc

$$\overline{f^2} - \overline{f}^2 = \sigma^2 + \mathbf{aEa} + 0(s^3) \quad (56)$$

D'autre part, en multipliant (46) par \mathbf{Xs} , on obtient

$$\mathbf{Xfs} = a_0\mathbf{Xs} + \sigma a_0 s_0 \mathbf{Xs} + \mathbf{Xs}\mathbf{s\tilde{X}a} + 0(s^3) \quad (57)$$

Donc

$$\mathbf{X}\overline{fs} = \mathbf{Ea} + \mathbf{0}(s^3) \quad (58)$$

En laissant tomber les termes d'ordre égal ou supérieur à s^3 , il reste

$$\mathbf{E}' \cong \mathbf{E} - \frac{\mathbf{Ea}\mathbf{aE}}{\sigma^2 + \mathbf{aEa}} \quad (59)$$

$$z' \cong z + \frac{y_k - a_0 - \frac{1}{2} \text{Tra}(\mathbf{AE})}{\sigma^2 + \mathbf{aEa}} \quad (60)$$

STABILITÉ

Nous allons démontrer que la méthode est stable. Nous entendons par là :

1° que les formes quadratiques construites avec le même vecteur quelconque \mathbf{u} sur deux matrices de covariance successives ne sont jamais croissantes, c'est-à-dire que $\mathbf{u} \mathbf{E}_k \mathbf{u} \leq \mathbf{u} \mathbf{E}_{k-1} \mathbf{u}$;

2° qu'il existe toujours au moins un vecteur \mathbf{v} tel que les formes quadratiques construites avec ce vecteur sur deux matrices de covariance successives, soient décroissantes, c'est-à-dire que $\mathbf{v} \mathbf{E}_k \mathbf{v} < \mathbf{v} \mathbf{E}_{k-1} \mathbf{v}$.

Pour prouver la première partie nous devons montrer préalablement que \mathbf{E}_k est définie positive. Comme \mathbf{E}_0 a été choisie définie positive on utilisera une preuve par récurrence.

THÉORÈME. — \mathbf{E}_k est définie positive si \mathbf{E}_{k-1} est définie positive.

Preuve. — Étudions la forme quadratique $\mathbf{u} \mathbf{E}_k \mathbf{u}$ et montrons qu'elle est positive. D'après (59), on obtient :

$$\mathbf{u} \mathbf{E}_k \mathbf{u} = \mathbf{u} \mathbf{E}_{k-1} \mathbf{u} - \frac{\mathbf{u} \mathbf{E}_{k-1} \mathbf{a}_k \mathbf{a}_k \mathbf{E}_{k-1} \mathbf{u}}{\sigma^2 + \mathbf{a}_k \mathbf{E}_{k-1} \mathbf{a}_k} \quad (61)$$

En mettant le premier terme du membre de droite au même dénominateur que le second terme, il vient :

$$\mathbf{u} \mathbf{E}_k \mathbf{u} = \frac{\sigma^2}{\sigma^2 + \mathbf{a}_k \mathbf{E}_{k-1} \mathbf{a}_k} \mathbf{u} \mathbf{E}_{k-1} \mathbf{u} + \frac{\mathbf{u} \mathbf{E}_{k-1} \mathbf{u} \mathbf{a}_k \mathbf{E}_{k-1} \mathbf{a}_k - [\mathbf{u} \mathbf{E}_{k-1} \mathbf{a}_k]^2}{\sigma^2 + \mathbf{a}_k \mathbf{E}_{k-1} \mathbf{a}_k} \quad (62)$$

Or

$$\sigma^2 > 0$$

$$\mathbf{u} \mathbf{E}_{k-1} \mathbf{u} > 0 \quad \text{car } \mathbf{E}_{k-1} \text{ est définie positive,}$$

et

$$\mathbf{a}_k \mathbf{E}_{k-1} \mathbf{a}_k > 0 \quad \text{pour la même raison.}$$

Le premier terme du membre de droite est donc positif.

Le second ne peut être négatif, car selon l'inégalité de Schwartz généralisée

$$\mathbf{u} \mathbf{E}_{k-1} \mathbf{u} \mathbf{a}_k \mathbf{E}_{k-1} \mathbf{a}_k > [\mathbf{u} \mathbf{E}_{k-1} \mathbf{a}_k]^2$$

Donc

$$\mathbf{u} \mathbf{E}_k \mathbf{u} > 0$$

C. Q. F. D.

THÉORÈME. — Les formes quadratiques construites avec le même vecteur sur des matrices de covariance successives ne sont jamais croissantes.

Preuve. — En effet, d'après (59)

$$\dot{\mathbf{u}}\mathbf{E}_{k-1}\dot{\mathbf{u}} - \dot{\mathbf{u}}\mathbf{E}_k\dot{\mathbf{u}} = \frac{[\dot{\mathbf{u}}\mathbf{E}_{k-1}\dot{\mathbf{a}}_k]^2}{\sigma^2 + \dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k} \quad (63)$$

$$\begin{aligned} \sigma^2 + \dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k &> 0 && \text{puisque } \mathbf{E}_{k-1} \text{ est définie positive} \\ (\dot{\mathbf{u}}\mathbf{E}_{k-1}\dot{\mathbf{a}}_k)^2 &\geq 0 && \text{puisque c'est le carré d'un scalaire.} \end{aligned}$$

Donc

$$\dot{\mathbf{u}}\mathbf{E}_{k-1}\dot{\mathbf{u}} - \dot{\mathbf{u}}\mathbf{E}_k\dot{\mathbf{u}} \geq 0$$

ou

$$\dot{\mathbf{u}}\mathbf{E}_k\dot{\mathbf{u}} \leq \dot{\mathbf{u}}\mathbf{E}_{k-1}\dot{\mathbf{u}}$$

C. Q. F. D.

La deuxième partie du théorème de stabilité s'énonce :

THÉORÈME. — Il existe au moins un vecteur tel que les formes quadratiques construites avec lui sur deux matrices de covariance successives sont décroissantes.

Preuve. — Il suffit de trouver un tel vecteur pour prouver le théorème. $\dot{\mathbf{a}}_k$ répond à cette propriété. En effet d'après (59) :

$$\dot{\mathbf{a}}_k\mathbf{E}_k\dot{\mathbf{a}}_k = \dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k - \frac{[\dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k]^2}{\sigma^2 + \dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k} \quad (64)$$

ou encore :

$$\dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k - \dot{\mathbf{a}}_k\mathbf{E}_k\dot{\mathbf{a}}_k = \frac{[\dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k]^2}{\sigma^2 + \dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k} \quad (65)$$

Or

$$\dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k > 0$$

Donc

$$\sigma^2 + \dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k > 0$$

et

$$[\dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k]^2 > 0.$$

Le membre de droite de (65) est donc positif, d'où :

$$\dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k - \dot{\mathbf{a}}_k\mathbf{E}_k\dot{\mathbf{a}}_k > 0$$

ou

$$\dot{\mathbf{a}}_k\mathbf{E}_k\dot{\mathbf{a}}_k < \dot{\mathbf{a}}_k\mathbf{E}_{k-1}\dot{\mathbf{a}}_k$$

C. Q. F. D.

SYSTÈME D'ÉQUATIONS LINÉAIRES

Un système de n équations linéaires à n inconnues peut s'écrire

$$y = Ax \tag{66}$$

soit encore

$$y_k = a_k x \quad k = 1, 2, \dots, n \tag{67}$$

où les n vecteurs a_k sont linéairement indépendants.

On sait que la solution est $x = A^{-1}y$.

L'application de l'algorithme optimal fournit une méthode simple de résolution où la matrice A^{-1} n'est pas calculée directement.

On utilise une matrice E_k qu'il suffit au départ de choisir parmi l'ensemble des matrices définies positives. Cette matrice est modifiée après chaque équation (observation certaine)

$$E_k = T_k E_{k-1} \tag{68}$$

en utilisant l'information apportée par cette équation.

On part donc d'une valeur hypothétique z_0 des inconnues et d'une matrice de covariance E_0 . Chaque observation est considérée comme certaine, c'est-à-dire que

$$\sigma_k^2 = 0 \tag{69}$$

SOLUTION DU SYSTÈME PRÉCÉDENT

La solution fournie par l'algorithme optimal est donc dans ce cas particulier

$$z_k = \left[-\frac{E_{k-1} a_k a_k}{a_k E_{k-1} a_k} \right] z_{k-1} + \frac{y_k}{a_k E_{k-1} a_k} E_{k-1} a_k \tag{70}$$

$$E_k = \left[\frac{E_{k-1} a_k a_k}{a_k E_{k-1} a_k} \right] E_{k-1} \tag{71}$$

ce qui montre que

$$T_k = I - \frac{E_{k-1} a_k a_k}{a_k E_{k-1} a_k} \tag{72}$$

CONVERGENCE DE LA SOLUTION

La méthode obtient la solution exacte en n équations. Pour prouver cela nous aurons besoin de montrer que $E_k \dot{a}_k = 0$.

THÉORÈME :

$$E_k \dot{a}_k = 0 \quad (73)$$

Preuve. — D'après (59)

$$E_k \dot{a}_k = E_{k-1} \dot{a}_k - \frac{E_{k-1} \dot{a}_k^* a_k^* E_{k-1} \dot{a}_k}{a_k^* E_{k-1} \dot{a}_k} \quad (74)$$

Le deuxième terme du membre de droite se simplifie par $\dot{a}_k E_{k-1} \dot{a}_k$.
Donc :

$$E_k \dot{a}_k = E_{k-1} \dot{a}_k - E_{k-1} \dot{a}_k = 0$$

C. Q. F. D.

Montrer que la solution exacte est obtenue en n estimations revient à montrer que E_n est nulle.

THÉORÈME :

$$E_k \dot{a}_p = 0 \quad \text{si} \quad k \geq p > 0. \quad (75)$$

Preuve. — Le théorème appliqué à la $k^{\text{ième}}$ estimation donne

$$E_k \dot{a}_k = 0$$

à la $(k-1)^{\text{ième}}$

$$E_k \dot{a}_{k-1} = T_k E_{k-1} \dot{a}_{k-1} = 0 \quad (76)$$

à la $(k-2)^{\text{ième}}$

$$E_k \dot{a}_{k-2} = T_k T_{k-1} E_{k-2} \dot{a}_{k-2} = 0 \quad (77)$$

à la première

$$E_k \dot{a}_1 = T_k T_{k-1} \dots T_2 E_1 \dot{a}_1 = 0 \quad (78)$$

Donc

$$E_k \dot{a}_p = 0 \quad \text{si} \quad 0 < p \leq k$$

ce qui prouve que

$$\text{Rang}(E_k) \leq n - k$$

THÉORÈME :

$$E_n = 0 \quad \text{puisque} \quad 0 \leq \text{Rang}(E_n) \leq 0 \quad (79)$$

Autre preuve. — On a n relations du type $\mathbf{E}_n \mathbf{a}_k^* = \mathbf{0}$ $k = 1, 2, \dots, n$. C'est donc que chacune des n lignes de \mathbf{E}_n est nulle ou orthogonale à \mathbf{a}_k^* . Or les n \mathbf{a}_k^* sont linéairement indépendant. C'est donc que ces lignes sont nulles.

MESURES A TEMPS CONSTANT

Un mode courant d'observation consiste à mesurer le nombre m de quanta reçus pendant un temps donné τ .

Soit α le flux en quanta par unité de temps. α dépend des inconnues x^* d'une manière connue et c'est en déterminant différents flux dépendant différemment de x^* qu'on mesure indirectement x^* .

La fonction de probabilité de m est :

$$p(m) = \frac{(\alpha\tau)^m e^{-\alpha\tau}}{m!}. \quad (80)$$

On en déduit facilement que

$$a_0 = \overline{m} = \alpha\tau \quad (81)$$

et

$$\overline{m^2} - \overline{m}^2 = \alpha\tau \quad (82)$$

Donc

$$[\mathbf{a}^*]_k = \tau \frac{\partial \alpha}{\partial x_k} \quad \text{ce qu'on écrit} \quad \mathbf{a}^* = \tau \boldsymbol{\alpha} \quad (83)$$

$$[\mathbf{A}]_{kl} = \tau \frac{\partial^2 \alpha}{\partial x_k \partial x_l} \quad \text{ce qu'on écrit} \quad \mathbf{A} = \tau \boldsymbol{\Psi} \quad (84)$$

enfin

$$\sigma^2 = q\alpha\tau \quad (q : \text{quantum unité}) \quad (85)$$

En portant (81), (83), (84) et (85) dans (59) et (60), et en simplifiant par τ il vient :

$$\delta \mathbf{E} = - \frac{\tau \mathbf{E} \boldsymbol{\alpha} \boldsymbol{\alpha} \mathbf{E}}{q\alpha + \tau \boldsymbol{\alpha} \mathbf{E} \boldsymbol{\alpha}^*} \quad (86)$$

$$\delta z^* = \frac{m - \alpha\tau - \frac{\tau}{2} \text{Tra} (\boldsymbol{\Psi} \mathbf{E})}{q\alpha + \tau \boldsymbol{\alpha} \mathbf{E} \boldsymbol{\alpha}^*} \mathbf{E} \boldsymbol{\alpha}^* \quad (87)$$

ESTIMATEUR INFINITÉSIMAL

Les formules précédentes s'appliquent dans les meilleures conditions au cas limite d'un accroissement continu. On a alors la correspondance suivante

$$\begin{aligned}\tau &\rightarrow dt \\ m &\rightarrow dy\end{aligned}$$

D'où les formules infinitésimales

$$d\mathbf{E} = -\frac{\mathbf{E}\boldsymbol{\alpha}^*\boldsymbol{\alpha}\mathbf{E}}{\alpha} dt \quad (88)$$

$$dz^i = \frac{\mathbf{E}\boldsymbol{\alpha}^*}{\alpha} \left[dy - \left[\alpha + \frac{1}{2} \text{Tra} (\boldsymbol{\Psi}\mathbf{E}) \right] dt \right] \quad (89)$$

α et $\boldsymbol{\alpha}^*$ dépendent de z^i et des paramètres expérimentaux qui peuvent varier dans le temps. Ces formules ne peuvent donc en général être intégrées sous cette forme, sauf en des cas tellement schématiques qu'ils sont dépourvus d'intérêt pratique.

MESURES A NOMBRE CONSTANT DE QUANTA

Pour s'approcher le plus possible des conditions de l'estimateur infinitésimal, on peut chercher à diminuer le temps τ donné. En dessous d'une certaine valeur (de l'ordre de $1/\alpha$), il arrivera souvent qu'aucun quantum ne sera reçu, ce qui est une situation peu satisfaisante. On est donc naturellement conduit à adopter un autre mode d'observation qui consiste à mesurer le temps t qui s'écoule entre un nombre $\mu + 1$ donné de quanta.

Cet intervalle de temps est une fonction aléatoire dont la densité de probabilité est :

$$p(t)dt = \frac{(\alpha t)^{\mu-1} e^{-\alpha t}}{(\mu-1)!} \alpha dt \quad (90)$$

La grandeur observée est

$$y = t \quad (91)$$

La valeur moyenne de t est :

$$\bar{t} = \int_0^\infty t p(t) dt = \frac{1}{(\mu-1)! \alpha} \int_0^\infty (\alpha t)^\mu e^{-\alpha t} \alpha dt = \frac{\mu!}{(\mu-1)! \alpha} = \frac{\mu}{\alpha} \quad (92)$$

donc

$$a = \frac{\mu}{\alpha} \quad (93)$$

On en déduit

$$[a']_k = \frac{\partial}{\partial x_k} \left(\frac{\mu}{\alpha} \right) = -\frac{\mu}{\alpha^2} \frac{\partial \alpha}{\partial x_k} \quad (94)$$

soit

$$a' = -\frac{\mu}{\alpha^2} \alpha' \quad (95)$$

De même

$$[A]_{kl} = \frac{\partial^2}{\partial x_k \partial x_l} \left[\frac{\mu}{\alpha} \right] = \frac{\partial}{\partial x_k} \left[-\frac{\mu}{\alpha^2} \frac{\partial \alpha}{\partial x_l} \right] = \frac{2\mu}{\alpha^3} \frac{\partial \alpha}{\partial x_k} \frac{\partial \alpha}{\partial x_l} - \frac{\mu}{\alpha^2} \frac{\partial^2 \alpha}{\partial x_k \partial x_l}$$

soit

$$A = \frac{2\mu}{\alpha^3} \alpha'' \alpha - \frac{\mu}{\alpha^2} \Psi \quad (96)$$

Enfin

$$\begin{aligned} \overline{t^2} &= \int_0^\infty t^2 p(t) dt = \frac{1}{(\mu-1)! \alpha^2} \int_0^\infty (\alpha t)^{\mu+1} e^{-\alpha t} \alpha dt = \frac{(\mu+1)!}{(\mu-1)! \alpha^2} \\ &= \frac{\mu(\mu+1)}{\alpha^2} \end{aligned} \quad (97)$$

D'où

$$\overline{t^2} - \bar{t}^2 = \frac{q\mu}{\alpha^2} \quad (98)$$

soit

$$\sigma^2 = \frac{q\mu}{\alpha^2} \quad (99)$$

Dans (60), il faut évaluer $\text{Tra}(\mathbf{AE})$

$$\text{Tra}(\mathbf{AE}) = \frac{2\mu}{\alpha^3} \alpha' \mathbf{E} \alpha' - \frac{\mu}{\alpha^2} \text{Tra}(\Psi \mathbf{E}) \quad (100)$$

On a alors

$$\delta \mathbf{E} = -\frac{\frac{\mu^2}{\alpha^4} \mathbf{E} \alpha'' \alpha \mathbf{E}}{\frac{q\mu}{\alpha^2} + \frac{\mu^2}{\alpha^4} \alpha' \mathbf{E} \alpha'} \quad (101)$$

$$\delta z' = -\frac{t - \frac{\mu}{\alpha} - \frac{\mu}{\alpha^3} \alpha' \mathbf{E} \alpha' + \frac{\mu}{2\alpha^2} \text{Tra}(\Psi \mathbf{E})}{\frac{q\mu}{\alpha^2} + \frac{\mu^2}{\alpha^4} \alpha' \mathbf{E} \alpha'} \frac{\mu}{\alpha^2} \mathbf{E} \alpha' \quad (102)$$

En multipliant haut et bas par α^4/μ^2 il vient :

$$\delta \mathbf{E} = - \frac{\mathbf{E} \boldsymbol{\alpha}'' \boldsymbol{\alpha} \mathbf{E}}{\frac{q \alpha^2}{\mu} + \boldsymbol{\alpha} \mathbf{E} \boldsymbol{\alpha}'} \quad (103)$$

$$\delta z' = \frac{\alpha - \frac{\alpha^2 t}{\mu} + \frac{1}{\alpha} \boldsymbol{\alpha} \mathbf{E} \boldsymbol{\alpha}' - \frac{1}{2} \text{Tra} (\boldsymbol{\Psi} \mathbf{E})}{\frac{q \alpha^2}{\mu} + \boldsymbol{\alpha} \mathbf{E} \boldsymbol{\alpha}'} \mathbf{E} \boldsymbol{\alpha}' \quad (104)$$

ESTIMATEUR UNIQUANTIQUE

Les formules précédentes fournissent les modifications que l'on peut apporter aux estimations et à leur matrice de covariance dès réception d'un seul quantum, il suffit de faire

$$\mu = q = 1 \text{ quantum} \quad (105)$$

En définissant

$$\mathbf{I}' = \frac{1}{\alpha} \boldsymbol{\alpha}' \quad (106)$$

et

$$\mathbf{L} = \frac{1}{\alpha} \boldsymbol{\Psi} \quad (107)$$

il vient

$$\delta \mathbf{E} = - \frac{\mathbf{E} \mathbf{I}''' \mathbf{E}}{1 + \mathbf{I}' \mathbf{E} \mathbf{I}'} \quad (108)$$

$$\delta z' = \left[1 - \frac{\frac{\alpha t}{q} + \frac{1}{2} \text{Tra} (\mathbf{L} \mathbf{E})}{1 + \mathbf{I}' \mathbf{E} \mathbf{I}'} \right] \mathbf{E} \mathbf{I}' \quad (109)$$

CONCLUSION

On a appliqué l'algorithme optimal au problème de l'exploitation d'une observation isolée pour l'estimation de plusieurs inconnues. On a vu qu'il est possible de construire un opérateur simple à calculer fournissant immé-

diatement une réestimation approchée des inconnues d'un système d'équations non linéaires.

Cette méthode présente les avantages suivants :

a) Elle permet l'exploitation des informations qu'on possède préalablement. Cette information est introduite dans le système sous la forme des espérances mathématiques des inconnues et de la matrice X de covariance des écarts autour des espérances mathématiques.

Les relations linéaires approchées existant entre les inconnues sont ainsi prises en compte.

b) Cette méthode donne une meilleure estimation des inconnues que la méthode des moindres carrés puisqu'on traite une masse d'informations plus importante.

c) Les relations linéaires certaines entre les inconnues lorsqu'il en existe, sont fidèlement traduites par la matrice X , dont le rang devient inférieur à l'ordre. Il n'est donc pas nécessaire d'employer un formalisme spécial pour traduire chaque système de relations linéaires dans chaque cas.

d) Enfin, le principal avantage consiste évidemment en la possibilité de pouvoir obtenir une réestimation immédiate de toutes les inconnues.

e) Les calculs sont simples à effectuer (du moins lorsque les dérivées sont simples à calculer) car ils ne font pas intervenir d'inversion de matrice. Et même, par une écriture convenable du programme de calcul, on peut éviter les produits de matrices.

APPLICATIONS POSSIBLES

Il nous semble donc que cette méthode pourrait s'appliquer très simplement dans deux cas :

a) *Résolution de systèmes linéaires* sur petites calculatrices. L'emploi de la formule (70) n fois répétée demande en effet très peu de place en mémoire. Par contre le nombre des opérations est plus élevé que dans le cas d'une résolution simultanée.

b) *Pilotage automatique* des ensembles de mesure.

Nous entendons par « pilotage » le fait d'agir sur les paramètres de commande d'une part en avançant vers un but précis (par exemple détermination d'une des inconnues avec le maximum d'exactitude), d'autre part, en tenant compte à chaque commande des difficultés qui se présentent, c'est-à-dire en exploitant immédiatement l'observation qui vient d'avoir

lieu. Cette définition est suffisante pour montrer que la méthode doit présenter un intérêt pour le pilotage automatique des ensembles de mesure, car, bien qu'approximative, elle permet un calcul rapide et peu encombrant de l'état actuel de chaque mesure.

BIBLIOGRAPHIE

- [1] M. TOURNARIE, Actes du 4^e Congrès International de Cybernétique, Namur, 1964.
- [2] M. TOURNARIE, Actes du colloque sur les calculs cristallographiques, Grenoble, 1965.

(Manuscrit reçu le 12 septembre 1968).
