NORME INTERNATIONALE

ISO
7066-2

Première édition
1988-07-01

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ

Évaluation de l'incertitude dans l'étalonnage et l'utilisation des appareils de mesure du débit —

Partie 2: Relations d'étalonnage non linéaires

Assessment of uncertainty in the calibration and use of flow measurement devices —

Part 2: Non-linear calibration relationships

Numéro de référence
ISO 7066-2 : 1988 (F)

Avant-propos

L'ISO (Organisation internationale de normalisation) est une fédération mondiale d'organismes nationaux de normalisation (comités membres de l'ISO). L'élaboration des Normes internationales est en général confiée aux comités techniques de l'ISO. Chaque comité membre intéressé par une étude a le droit de faire partie du comité technique créé à cet effet. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO participent également aux travaux. L'ISO collabore étroitement avec la Commission électrotechnique internationale (CEI) en ce qui concerne la normalisation électrotechnique.

Les projets de Normes internationales adoptés par les comités techniques sont soumis aux comités membres pour approbation, avant leur acceptation comme Normes internationales par le Conseil de l'ISO. Les Normes internationales sont approuvées conformément aux procédures de l'ISO qui requièrent l'approbation de 75 % au moins des comités membres votants.

La Norme internationale ISO 7066-2 a été élaborée par le comité technique ISO/TC 30, *Mesure de débit des fluides dans les conduites fermées*.

L'attention des utilisateurs est attirée sur le fait que toutes les Normes internationales sont de temps en temps soumises à révision et que toute référence faite à une autre Norme internationale dans le présent document implique qu'il s'agit, sauf indication contraire, de la dernière édition.

Sommaire

Page

0 Introduction 1

1 Objet et domaine d'application 1

2 Références 1

3 Définitions 1

4 Symboles et abréviations 1

5 Ajustement de la courbe 2

5.1 Généralités 2

5.2 Méthodes de calcul 3

5.3 Choix du degré optimal d'ajustement 3

6 Incertitude 4

Annexes

A Méthodes de régression 5

B Ajustement des courbes par la méthode des polynômes orthogonaux 9

C Programme d'ordinateur pour la méthode des polynômes orthogonaux 11

D Exemples 17

E Méthode des différences finies 27

Page blanche

Évaluation de l'incertitude dans l'étalonnage et l'utilisation des appareils de mesure du débit —

Partie 2 : Relations d'étalonnage non linéaires

0 Introduction

L'ISO 7066-1 indiquait comment on pouvait faire passer une droite d'ajustement entre les valeurs d'étalonnage des appareils de mesurage du débit et calculer l'incertitude de cet étalonnage. La présente partie de l'ISO 7066 traite du cas où une droite ne permet pas de représenter les valeurs d'étalonnage.

1 Objet et domaine d'application

La présente partie de l'ISO 7066 indique comment faire passer une courbe exprimée sous la forme d'un polynôme de degré 2, 3 ou au-delà, par un ensemble non rectiligne[1] de valeurs d'étalonnage en utilisant comme critère la méthode des moindres carrés. Elle indique également comment évaluer l'incertitude à partir de la courbe d'étalonnage ainsi obtenue. La présente partie de l'ISO 7066 ne traite que du cas des polynômes à exposants entiers.

L'établissement de ce type de courbe et l'évaluation de l'incertitude correspondante étant généralement impossibles sans ordinateur, la présente partie de l'ISO 7066 pose comme hypothèse de travail que l'utilisateur en possède un. Dans la plupart des cas, il sera possible de se servir des routines normalisées de l'ordinateur. Il est également possible d'exploiter le programme FORTRAN indiqué en annexe C.

Des exemples d'application des méthodes préconisées figurent en annexe D.

Il n'est pas permis d'extrapoler au-delà de l'étendue des mesures effectuées. Les annexes A, B, C, D et E ne font pas partie intégrante de la présente partie de l'ISO 7066.

2 Références

ISO 5168, *Mesure de débit des fluides — Calcul de l'erreur limite sur une mesure de débit.*[2]

ISO 7066-1, *Évaluation de l'incertitude dans l'étalonnage et l'utilisation des appareils de mesure de débit — Partie 1 : Relations d'étalonnage linéaires.*[3]

3 Définitions

Dans le cadre de la présente partie de l'ISO 7066, les définitions suivantes sont applicables.

3.1 méthode des moindres carrés : Méthode utilisée pour calculer les coefficients d'une équation lorsqu'on a choisi la forme d'équation la plus adaptée pour ajuster une courbe à des valeurs. Le principe de la méthode des moindes carrés est de minimiser la somme des carrés des écarts des valeurs observées par rapport à la courbe.

3.2 polynômes : Pour une variable x, une série de termes entiers ordonnés selon les puissances croissantes de x.

3.3 analyse de régression : Le processus de quantification de la dépendance d'une variable par rapport à une ou plusieurs autres variables.

NOTE — Beaucoup de programmes d'ordinateur utilisables pour l'ajustement de courbes possèdent le terme «régression» dans leur titre. Dans la présente partie de l'ISO 7066, les termes «régression» et «moindres carrés» peuvent être employés indifféremment.

3.4 écart-type : Racine carrée positive de la variance.

3.5 variance : Mesure de la dispersion basée sur la moyenne des carrés des écarts des valeurs d'une variable à sa valeur espérée.

4 Symboles et abréviations

b_j coefficient de x_j

C_{jb} élément de la matrice inverse

1) Les méthodes sont valables pour un ensemble rectiligne également.

2) Actuellement au stade de projet. (Révision de l'ISO 5168 : 1978.)

3) Actuellement au stade de projet.

$e_r(\ )$ incertitude aléatoire de la variable contenue entre parenthèses[1)]

$e_s(\ )$ incertitude systématique de la variable contenue entre parenthèses[1)]

$e(\hat{y}_c)$ incertitude totale du coefficient d'étalonnage[1)]

g_j coefficient du polynôme orthogonal de degré j

m degré du polynôme

n nombre de valeurs

$p_j(x)$ polynôme orthogonal de degré j

$s(\ )$ écart-type expérimental de la variable contenue entre parenthèses

s_r écart-type résiduel des valeurs autour de la courbe

t coefficient de Student

x variable indépendante

x^* valeur arbitraire spécifiée de x

$\bar{x}$ moyenne arithmétique des valeurs x_i

x_i valeur de x au $i^{\text{ème}}$ point de mesure

x_j $j^{\text{ème}}$ variable indépendante (en régression linéaire multiple)

x_{ji} valeur de x_j *au* $i^{\text{ème}}$ point de mesure

y variable dépendante

$\bar{y}$ moyenne arithmétique des valeurs y_i

$\hat{y}$ valeur de y obtenue par l'équation de la courbe d'ajustement

y_i valeur de y au $i^{\text{ème}}$ point de mesure

$\hat{y}_i$ valeur de $\hat{y}$ à $x = x_i$

ν nombre de degrés de liberté

5 Ajustement de la courbe

5.1 Généralités

Avant de chercher à ajuster les courbes polynômiales, il faut voir si un simple changement de la variable x ou de la variable y ou des deux variables ne pourrait pas effectivement linéariser les données et permettre l'utilisation des méthodes utilisant des courbes rectilignes décrites dans l'ISO 7066-1. Certains changements de variables sont présentés à titre de suggestion dans l'ISO 7066-1.

S'il n'est pas possible d'obtenir une droite, alors il faut trouver le degré et les coefficients de la fonction polynômiale qui représentent le mieux un ensemble de n paires de valeurs (x_i, y_i) relevées pendant l'étalonnage. Si l'on choisit, par exemple, le polynôme du deuxième degré, la courbe sera de la forme:

$$\hat{y} = b_0 + b_1 x + b_2 x^2 \qquad \ldots (1)$$

L'expression polynômiale générale est:

$$\hat{y} = b_0 + b_1 x + \ldots + b_j x^j + \ldots + b_m x^m$$

ou bien

$$\hat{y} = \sum_{j=0}^{m} b_j x^j \qquad \ldots (2)$$

En appliquant la règle des moindres carrés, on calcule les coefficients b_j pour minimiser la somme des carrés des écarts des points de mesure par rapport à la courbe:

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

où $\hat{y}_i$ est la valeur obtenue par l'équation (2) pour $x = x_i$.

Dans certains cas, le degré m du polynôme est déjà connu: par exemple, on peut savoir par expérience qu'une cubique ($m = 3$) représentera de façon satisfaisante les données d'étalonnage. Dans le cas contraire, le degré à retenir sera déterminé par accroissements successifs du degré du polynôme jusqu'à ce qu'un ajustement optimal soit obtenu (voir 5.3).

Cependant, si l'augmentation du degré au-delà d'une valeur raisonnable continue d'améliorer l'ajustement de façon significative (comme décrit en 5.3), c'est probablement qu'une loi polynômiale ne convient pas pour représenter la relation fonctionnelle en cause. De plus, si l'équation obtenue comporte trop de termes, la courbe peut présenter des oscillations aberrantes. Un exemple assez courant de ces aberrations se retrouve dans des données virtuellement constantes sur la presque totalité de l'étendue des x, mais qui présentent une variation extrêmement marquée à l'une des extrémités de l'étendue.

Dans de tels cas, il convient de diviser l'étendue en plusieurs sections (voir ISO 7066-1) dans chacune des desquelles l'ajustement pourra être obtenu soit par une droite soit par un polynôme de faible degré. On peut également changer l'une ou l'autre variable ou les deux pour obtenir une fonction linéaire ou une fonction polynômiale de faible degré; le changement de la variable indépendante en son inverse, $1/x$, donne dans certains cas une linéarité convenable.

Les méthodes des moindres carrés décrites dans la présente partie de l'ISO 7066 peuvent ne pas convenir si l'effet de l'incertitude aléatoire $e_r(x)$ des valeurs mesurées x_i n'est pas négligeable par rapport à celui de l'incertitude aléatoire $e_r(y)$ des

1) Dans certaines Normes internationales, les symboles U et E sont utilisés à la place de e.

valeurs de y. Si, comme dans l'ISO 7066-1 la pente[1] de la courbe d'étalonnage est toujours inférieure à 1/5 de $e_r(y) / e_r(x)$, les méthodes peuvent être considérées comme appropriées. Dans le cas contraire, il faut procéder à un traitement mathématique qui sort du cadre de la présente partie de l'ISO 7066. Donc, si la pratique normale d'étalonnage d'un débitmètre donné consiste à reporter les variables sur une courbe de telle manière qu'on ne puisse pas remplir la condition ci-dessus, il faut soit inverser le choix normal des abscisses et des ordonnées, soit renoncer à utiliser la présente partie de l'ISO 7066.

Si l'on change l'une ou l'autre variable avant l'ajustement, les incertitudes évoquées ci-dessus et plus bas (au chapitre 5) se rapporteront alors aux nouvelles variables. Et si, en raison du changement de la variable dépendante, l'incertitude aléatoire $e_r(y)$ ne peut plus être considérée comme constante sur toute l'étendue, il faut alors utiliser une méthode de moindres carrés pondérés. Les moindres carrés pondérés ne sont pas décrits dans la présente partie de l'ISO 7066, mais de nombreuses routines de bibliothèques d'ordinateurs disponibles dans le commerce permettent de pondérer les données.

5.2 Méthodes de calcul

Des routines normalisées d'ajustement des courbes par la méthode des moindres carrés sont disponibles sur la plupart des ordinateurs. La méthode d'ajustement d'une droite décrite dans l'ISO 7066-1 est communément désignée sous le terme de «régression linéaire» ou de «régression linéaire simple». La méthode équivalente utilisable pour les polynômes peut être qualifiée de régression polynômiale ou curviligne et est un type spécial de régression linéaire multiple. L'annexe A donne des détails plus précis sur la nature des méthodes de régression et leur utilisation.

À côté des méthodes de régression classiques, on peut également utiliser la méthode des polynômes orthogonaux décrite à l'annexe B. Cette méthode est particulièrement bien adaptée aux cas où l'on ne connaît pas à l'avance le degré d'ajustement. L'annexe C fournit à cet égard un programme d'ordinateur.

Lorsqu'on ne dispose pas d'un ordinateur et que les valeurs de x sont réparties de façon uniforme, on peut utiliser la méthode des différences finies (voir annexe E) qui donne une indication immédiate du degré d'ajustement approprié pour représenter les données. On peut également calculer les coefficients du polynôme représentant ces données mais il ne s'agit plus là du polynôme des moindres carrés et la méthode de calcul des valeurs de l'incertitude n'entre pas dans le cadre de la présente partie de l'ISO 7066.

5.3 Choix du degré optimal d'ajustement

On cherche la représentation la mieux adaptée en essayant des valeurs croissantes du degré m, soit jusqu'à un maximum spécifié, soit jusqu'à ce qu'on n'observe plus aucune amélioration significative. On calcule ensuite l'écart-type résiduel s_r pour chaque degré (s_r est la racine carrée de la variance résiduelle) à l'aide de l'équation suivante:

$$s_r^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 / (n - m - 1) \qquad \ldots (3)$$

où $\hat{y}_i$ est la valeur donnée par l'expression polynômiale [équation (2)] pour $x = x_i$.

NOTE — s_r^2 est l'équivalent de $s^2(y,x)$ utilisé dans l'ISO 7066-1.

Le degré m doit toujours être inférieur au nombre de points de mesure n.

Si les données peuvent être correctement représentées par un polynôme de degré m, s_r décroîtra de façon significative jusqu'à ce qu'on atteigne ce degré m. Par la suite, s_r demeurera sensiblement constant. En général, toutefois, il n'est pas évident de trouver à partir de quel degré la diminution de s_r cesse d'être significative et il convient de prendre un test objectif de signification pour aider à définir le degré optimal d'ajustement.

On considère que l'augmentation du degré de $m-1$ à m représente une amélioration statistiquement satisfaisante de l'ajustement si le nouveau coefficient, b_m, diffère de façon significative de zéro, c'est-à-dire si $b_m + t_{95}\, s(b_m)$ et $b_m - t_{95}\, s(b_m)$, (limites de confiance à 95 % de b_m) n'inclut pas zéro.

Cette condition peut s'exprimer sous la forme:

$$\left| \frac{b_m}{s(b_m)} \right| > t_{95}$$

où t_{95} est le coefficient de Student pour un niveau de confiance de 95 % à $\nu = n - m - 1$.

La valeur de t_{95} en fonction du nombre de degrés de liberté ν peut se calculer par la relation empirique suivante:

$$t_{95} = 1{,}96 + 2{,}36/\nu + 3{,}2/\nu^2 + 5{,}2/\nu^{3{,}84} \qquad \ldots (4)$$

Pour le coefficient polynômial orthogonal g_m (voir annexe B), cette condition est:

$$\left| \frac{g_m}{s(g_m)} \right| > t_{95}$$

Les expressions des variances des coefficients $s^2(b_m)$ et $s^2(g_m)$ sont données en annexe A et annexe B respectivement.

Il est important de tester l'effet d'une augmentation du degré à un échelon au moins au-delà de celui qui ne fait plus apparaître aucune amélioration significative car il s'avère souvent que, soit seuls les termes impairs, soit seuls les termes pairs, engendrent cette amélioration significative.

D'un point de vue statistique, le degré le plus élevé apportant une amélioration significative de l'ajustement au niveau de confiance de 95 % peut être considéré comme le degré optimal. Toutefois, avant de choisir ce degré comme le plus apte à repré-

1) «Pente» signifie ici la dérivée $d\hat{y} / dx = b_1 + 2b_2x + \ldots$.

senter les données, il faut considérer d'autres facteurs parmi lesquels: la forme attendue de la courbe, l'aspect pratique d'avoir une forme fonctionnelle qui ne soit pas trop complexe, l'étendue qu'il est nécessaire de représenter et la précision recherchée.

Pour évaluer ces facteurs, il est toujours conseillé de tracer des graphiques représentant les données et les courbes possibles. Ces graphiques permettent également de mettre en lumière d'autres problèmes éventuels et notamment, si le degré est trop faible, l'incapacité de la courbe à représenter la tendance réelle des données, les valeurs présumées $\hat{y}$ pouvant présenter une erreur systématique sur une partie de l'étendue. Si le degré est trop élevé, la courbe peut être ajustée à des données dispersées, et non à la tendance sous-jacente.

Les exemples donnés en annexe D illustrent une application de certains de ces principes.

6 Incertitude

La composante aléatoire de l'incertitude au niveau de confiance de 95 % sur la valeur présumée ŷ, est donnée par la formule:

$$e_r(\hat{y}) = t_{95}\, s(\hat{y})$$

où $s(\hat{y})$ est la racine carrée de la variance $s^2(\hat{y})$ de $\hat{y}$.

Les expressions de $s^2(\hat{y})$ figurent en annexes A et B ; en général $s^2(\hat{y})$ peut s'exprimer sous la forme d'une fonction polynômiale de x de degré $2m$. Il est important de vérifier que le calcul de $s^2(\hat{y})$ intègre suffisamment de chiffres significatifs pour éviter les grosses erreurs d'arrondissage résultant de soustractions.

À noter que l'estimation de l'incertitude apportée par $e_r(\hat{y})$ ne sera valable que dans la mesure où le polynôme choisi représente une bonne approximation de la relation fonctionnelle vraie entre y et x.

Les limites de confiance aléatoires à 95 % pour la valeur vraie de y sont définies par:

$$y \pm e_r(\hat{y})$$

Comme dans l'ISO 7066-1, l'incertitude sur le coefficient d'étalonnage est donnée par:

$$e(\hat{y}_c) = [e_r^2(\hat{y}) + e_s^2(\hat{y})]^{1/2}$$

où $e_s(\hat{y})$ est la composante systématique de l'incertitude sur $\hat{y}$.

NOTE — Dans la version révisée de l'ISO 5168 (en préparation), figurent des directives sur l'utilisation de l'addition linéaire ou de la combinaison quadratique des erreurs aléatoires et systématiques.

Si la variable dépendante a été transformée, toutes les incertitudes ci-dessus se rapportent à la variable transformée.

Annexe A

Méthodes de régression

(Cette annexe ne fait pas partie intégrante de la norme.)

A.1 Introduction

Des méthodes de régression utilisables pour l'ajustement des courbes sont disponibles sous diverses appellations dans les routines normalisées des bibliothèques d'ordinateurs. La documentation jointe à ces programmes suppose en général un certain niveau de connaissances en matière d'analyse de régression. L'objet de la présente annexe est de décrire ces méthodes d'un point de vue général et de fournir une terminologie de l'ajustement de courbes par régression facilitant l'exploitation de ces programmes.

La technique de régression la plus couramment rencontrée est, exception faite de la régression linéaire simple, la régression linéaire multiple. L'ajustement de courbes s'effectue à l'aide d'un type spécial de régression linéaire multiple appelé «régression polynômiale» ou «curviligne». Si l'on ne dispose pas de programme de régression polynômiale, on peut procéder par régression linéaire multiple mais cette méthode est moins commode. Les méthodes «pas à pas» et de «régression pas à pas descendante» constituent des modalités spéciales de régression linéaire multiple qui peuvent être utilisées.

A.2 Régression linéaire multiple

Dans ce qui suit, le signe de sommation Σ signifie, sauf annotation contraire, $\sum_{i=1}^{n}$.

Une variable dépendante y est supposée en relation linéaire avec m variables indépendantes $x_1, x_2, \dots, x_m$ suivant l'équation

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + U \qquad \dots (5)$$

où

β_0 et β_m sont les coefficients inconnus de régression;

U est une mesure des effets aléatoires responsables de l'écart de linéarité de la dépendance de y par rapport aux variables indépendantes m.

Sur les n ensembles d'observations:

$(y_i, x_{1i}, x_{2i}, \dots, x_{mi}) \qquad i = 1, 2, \dots, n$

les estimations des coefficients de régression sont:

$b_0, b_1, \dots, b_m$

de telle sorte que l'estimation $\hat{y}$ de la valeur vraie correspondant au $i^{\text{ème}}$ ensemble d'observations des variables indépendantes donne:

$$\hat{y}_i = b_0 + b_1 x_{1i} + \dots + b_m x_{mi} \qquad \dots (6)$$

En appliquant la méthode des moindres carrés pour minimiser $\Sigma\,(y_i - \hat{y}_i)^2$, on obtient un ensemble de $m + 1$ équations simultanées communément appelées «équations normales»:

$$\begin{aligned} &nb_0 + \Sigma(x_{1i})b_1 + \Sigma(x_{2i})b_2 + \dots + \Sigma(x_{mi})b_m = \Sigma y_i \\ &\Sigma(x_{1i})b_0 + \Sigma(x_{1i})^2 b_1 + \dots + \Sigma(x_{1i}x_{mi})b_m = \Sigma(x_{1i}y_i) \\ &\Sigma(x_{mi})b_0 + \Sigma(x_{mi}x_{1i})b_1 + \dots + \Sigma(x_{mi})^2 b_m = \Sigma(x_{mi}y_i). \end{aligned} \qquad \dots (7)$$

Ces équations peuvent alors être résolues pour les $m + 1$ inconnues $b_0, b_1, \dots, b_m$.

A.3 Régression polynômiale (curviligne)

Lorsque la relation entre deux variables n'est pas linéaire mais peut être ajustée à une fonction polynômiale de la forme:

$$\hat{y} = b_0 + b_1 x + b_2 x^2 + \ldots + b_m x^m$$

on dit que l'on a une régression polynômiale ou curviligne de y en x. Cette régression peut être traitée comme une régression linéaire multiple en remplaçant des variables indépendantes $x_1, x_2, \ldots, x_m$ par $x, x^2, \ldots, x^m$.

Aux chapitres A.4 et A.5, toute expression d'une régression linéaire multiple peut être transformée en une expression de régression polynomiale équivalente si l'on remplace la $i$ème variable indépendante x_j par x^j et les valeurs correspondantes, de x_{ji} par x_i^j.

A.4 Calcul des coefficients et des variances

Considérons l'équation de régression linéaire multiple de degré $m = 2$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 \qquad \ldots (8)$$

qui équivaut, dans le cas de la régression polynômiale à

$$\hat{y} = b_0 + b_1 x + b_2 x^2 \qquad \ldots (9)$$

Si l'on applique le critère des moindres carrés, on obtient les équations normales

$$n b_0 + \Sigma(x_{1i}) b_1 + \Sigma(x_{2i}) b_2 = \Sigma(y_i) \qquad \ldots (10)$$

$$\Sigma(x_{1i}) b_0 + \Sigma(x_{1i})^2 b_1 + \Sigma(x_{1i} x_{2i}) b_2 = \Sigma(x_{1i} y_i) \qquad \ldots (11)$$

$$\Sigma(x_{2i}) b_0 + \Sigma(x_{2i} x_{1i}) b_1 + \Sigma(x_{2i})^2 b_2 = \Sigma(x_{2i} y_i) \qquad \ldots (12)$$

La méthode traditionnellement utilisée pour résoudre ces équations normales implique de calculer l'inverse de la matrice 3 × 3 des coefficients b_0, b_1 et b_2. Si les éléments de cette matrice inverse sont

$$\begin{bmatrix} C_{00} & C_{01} & C_{02} \\ C_{10} & C_{11} & C_{12} \\ C_{20} & C_{21} & C_{22} \end{bmatrix}$$

alors

$$\begin{aligned} b_0 &= C_{00} \Sigma y_i + C_{01} \Sigma(x_{1i} y_i) + C_{02} \Sigma(x_{2i} y_i) \\ b_1 &= C_{10} \Sigma y_i + C_{11} \Sigma(x_{1i} y_i) + C_{12} \Sigma(x_{2i} y_i) \\ b_2 &= C_{20} \Sigma y_i + C_{21} \Sigma(x_{1i} y_i) + C_{22} \Sigma(x_{2i} y_i) \end{aligned} \qquad \ldots (13)$$

ou, sous forme généralisée,

$$b_j = \sum_{k=0}^{m} \left[C_{jk} \, \Sigma(x_{ki} y_i) \right]$$

où $x_{ki} = 1$ pour $k = 0$.

À noter que, puisque la matrice des équations normales est symétrique, la matrice inverse est également symétrique.

Les variances des coefficients sont:

$$s^2(b_0) = s_r^2 C_{00}$$
$$s^2(b_1) = s_r^2 C_{11}$$
$$s^2(b_2) = s_r^2 C_{22}$$

la variance résiduelle s_r^2 étant donnée comme en 5.3 par l'équation:

$$s_r^2 = \frac{\Sigma(y_i - \hat{y})^2}{n - m - 1}$$

La matrice inverse étant symétrique, on obtient:

$$\begin{aligned} C_{01} &= C_{10} \\ C_{02} &= C_{20} \\ C_{12} &= C_{21} \end{aligned} \qquad \ldots (14)$$

Ces termes diagonaux servent à calculer les covariances[1)] entre les coefficients b_j (la covariance étant notée COV):

$$\begin{aligned} \mathrm{COV}(b_0, b_1) &= s_r^2 C_{01} \\ \mathrm{COV}(b_0, b_2) &= s_r^2 C_{02} \\ \mathrm{COV}(b_1, b_2) &= s_r^2 C_{12} \end{aligned} \qquad \ldots (15)$$

Aux valeurs spécifiées $x_1 = x_1^*$ et $x_2 = x_2^*$, la valeur prédite par l'équation de régression est:

$$\hat{y} = b_0 + b_1 x_1^* + b_2 x_2^* \qquad \ldots (16)$$

La variance de cette valeur $\hat{y}$ est donnée par:

$$s^2(\hat{y}) = s_r^2 \left[C_{00} + C_{11}(x_1^*)^2 + C_{22}(x_2^*)^2 + 2C_{01}x_1^* + 2C_{02}x_2^* + 2C_{12}x_1^*x_2^* \right] \qquad \ldots (17)$$

Le facteur 2 apparaît parce que $C_{jk} = C_{kj}$ pour chaque couple de valeurs j et k.

La formule générale est:

$$s^2(\hat{y}) = s_r^2 \sum_{j=0}^{m} \sum_{k=0}^{m} (C_{jk}\, x_j^* x_k^*)$$

où $x_j^*, x_k^* = 1$ pour $j, k = 0$. ... (18)

Pour une régression polynômiale, $x_j^* = (x^*)^j$ et $x_k^* = (x^*)^k$, et ainsi

$$s^2(\hat{y}) = s_r^2 \sum_{j=0}^{m} \left[\sum_{k=0}^{m} C_{jk}(x^*)^{j+k} \right]$$

L'adaptation de cette expression à un polynôme de degré $2m$ donne:

$$s^2(\hat{y}) = s_r^2 \sum_{j=0}^{m} \left[\left(\sum_{k=0}^{j} C_{k,j-k} \right) (x^*)^j \right] + s_r^2 \sum_{j=m+1}^{2m} \left[\left(\sum_{k=j-m}^{m} C_{k,j-k} \right) (x^*)^j \right] \qquad \ldots (19)$$

1) La covariance de deux coefficients indique l'effet d'une variation d'un coefficient sur la valeur de l'autre. On appelle variance, covariance ou matrice de variance-covariance, le produit de la matrice inverse par le scalaire s_r^2.