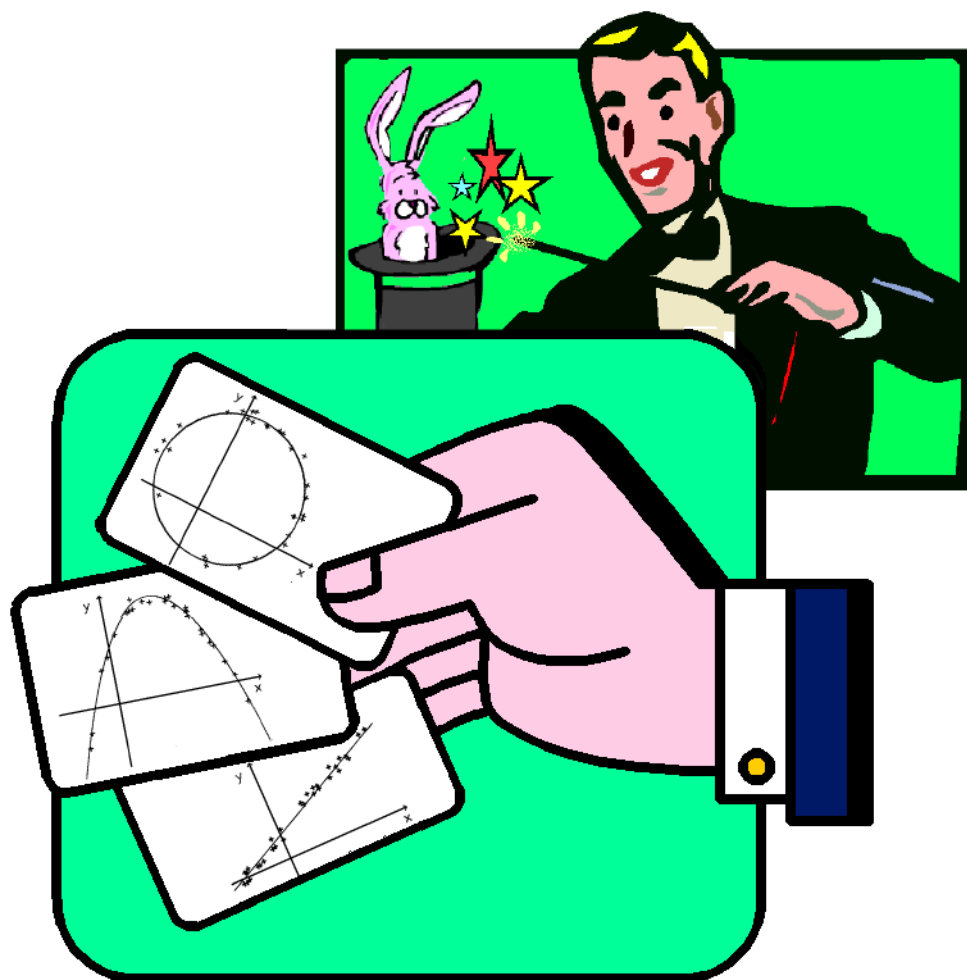


REGRESSIONS CONIQUES, QUADRIQUES

Régressions linéaires et apparentées, circulaire, sphérique



A l'origine de ce papier, on trouve la publication en 1982 de la méthode de régression circulaire [2].
La première partie du texte suivant a été publiée dans le magazine QUADRATURE n°63, pp.33-40, janvier 2007.
La deuxième partie a été publiée dans QUADRATURE n°65, pp.4-5, juillet 2007
Edité par EDP Sciences, 17 av. du Hoggar, PA de Courtaboeuf, 91944 Les ULIS, France
<http://www.edpsciences.org/quadrature/>

REGRESSIONS CONIQUES, QUADRIQUES

Régressions linéaires et apparentées, circulaire, sphérique

Résumé :

Les problèmes de régressions considérés dans ce papier sont du genre linéaire, c'est-à-dire se ramenant à une régression linéaire par de simples transformations analytiques.

Après un rappel des cas classiques de régressions linéaires, le cas spécifique de la *régression circulaire* est traité en détail.

L'étude est étendue aux coniques, puis aux quadriques, avec une attention particulière au cas de la *régression sphérique*.

Table des matières

Première partie :

1. Introduction
 2. Régression linéaire à un paramètre
 - 2.1. Moindres carrés des écarts d'ordonnées
 - 2.2. Moindres carrés des écarts d'abscisses
 - 2.3 Moindres carrés des écarts de distances
 3. Régression linéaire à deux paramètres
 - 3.1. Moindres carrés des écarts d'ordonnées
 - 3.2. Moindres carrés des écarts d'abscisses
 - 3.3 Moindres carrés des écarts de distances
 4. Régression linéaire, cas général
 5. Ajustement des paramètres d'une équation
 6. Cas de régressions non linéaires et/ou de critère autre que les moindres carrés
 7. Régression circulaire
 8. Discussion
- Annexe : Exemple d'application de la méthode de régression circulaire
- Références

Deuxième partie : Généralisations

- Généralisation aux coniques
- Généralisation aux quadriques
- Régression sphérique

Première partie

Résumé :

Après des généralités sur la notion de régression et un rappel des équations de régression linéaire, les formules de régression circulaire sont établies.

Abstract :

After some generalities related to the notion of fitting and a reminder of the linear fitting equations, the formulas for circular fitting are established.

1. INTRODUCTION

Alors que la régression linéaire est très répandue dans les ouvrages de statistique, il est surprenant que la régression circulaire n'y soit généralement pas spécifiquement traitée, bien que ce soit un problème qui n'est pas rare en pratique, ce qui sera illustré en citant quelques exemples.

Mais, au fait, que veut dire le mot "régression" ?

L'appellation *analyse de régression* provient d'une publication de F.Galton ayant pour titre : "Regression Towards Mediocrity in Hereditary Stature" (1885). Son étude statistique montrait que des parents de tailles fortement différentes de la moyenne avaient des enfants dont la taille tendait à *régresser* vers la moyenne de la population. Galton identifie la droite décrivant la relation entre la taille des parents et celle des enfants par "droite de régression" [1]. Le mot est resté, tout en évoluant vers le sens plus significatif d'ajustement ou d'optimisation.

En fait, il s'agit d'ajuster les paramètres d'une fonction pour qu'elle satisfasse, *au mieux selon certains critères*, à une ou des conditions imposées. Une traduction graphique serait, par exemple, la suivante : ajuster les coefficients d'un polynôme (ou d'une autre fonction) de telle sorte que sa courbe représentative passe *au plus près* de points donnés. La figure 1 en donne une illustration intuitive.

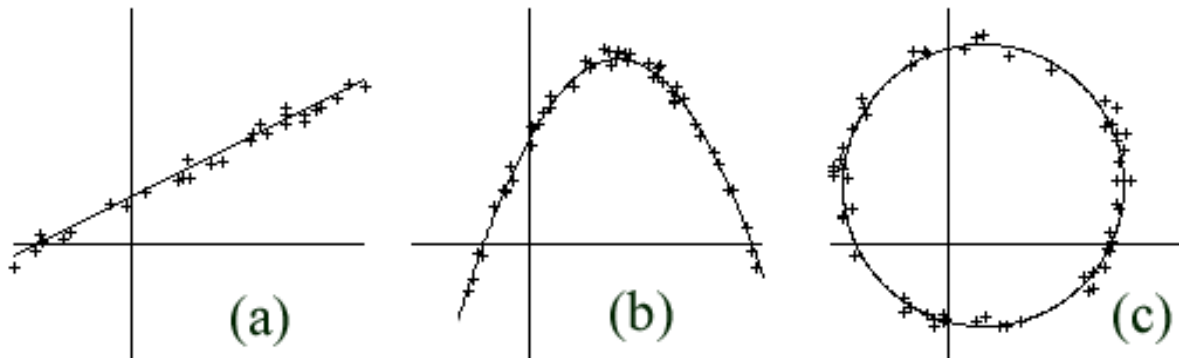


Figure 1 : Exemples de régressions. (a): linéaire ; (b): parabolique ; (c):circulaire.

L'historique de ce genre de problèmes, ajouté au fait qu'ils soient le plus souvent traités dans les ouvrages de statistiques, pourraient faire croire que les questions d'ajustement et d'optimisation relèvent uniquement du domaine des statistiques. Ce serait oublier beaucoup d'autres cas. Par exemple, dans un problème technologique nécessitant de *calculer les paramètres d'une fonction de telle sorte que sa courbe représentative satisfasse à un certain critère de proximité par rapport à des points donnés*, il n'est pas fait mention de dispersions dues

à des déviations aléatoires ou à des mesures imprécises. Il peut tout aussi bien s'agir de points disposés ainsi volontairement, selon des coordonnées bien déterminées, pour des raisons techniques bien précises, peu importe lesquelles.

Le présent papier se place donc dans un cadre plus large que celui des statistiques proprement dites, auxquelles il peut s'appliquer aussi bien évidemment.

Mais encore faut-il préciser ce que veut dire "au plus près".

Soient n points $(x_1, y_1), \dots (x_k, y_k), \dots (x_n, y_n)$ et une famille de fonctions $f(x; a_1, a_2, \dots, a_p)$ dépendant de p paramètres. Si la courbe représentative d'une fonction $f(x)$ passe par le point (x_k, y_k) , l'équation $y_k = f(x_k)$ est satisfaite. Il se présente alors trois cas très différents :

- soit $n = p$: Les n équations $y_k = f(x_k)$ à n inconnues (a_1, a_2, \dots, a_n) constituent un système dont la résolution conduit généralement à une solution (cas linéaire) ou éventuellement à plusieurs solutions (cas non linéaire). Cette solution, ou chacune de ces solutions, définit une fonction dont la courbe représentative passe exactement par les n points donnés.
- soit $n < p$: on a moins d'équations que d'inconnues. On peut fixer arbitrairement $(p-n)$ paramètres parmi les (a_1, a_2, \dots, a_p) ce qui ramène au cas précédent en ce qui concerne le calcul des n paramètres restants.
- soit $n > p$, le seul cas où l'on parle de "régression" : Sauf cas particulier, il est hors de question que l'on puisse trouver une fonction de la famille à p paramètres qui satisfasse simultanément à un nombre d'équations n supérieur au nombre p d'inconnues. On se contente alors de rechercher des valeurs de (a_1, a_2, \dots, a_p) telles que les n équations, qui ne peuvent plus être exactement satisfaites, le soient approximativement, avec le moins d'écart possible. La courbe représentative de la fonction $f(x; a_1, a_2, \dots, a_p)$ qui passe dans le "nuage" des n points donnés *au plus proche* de ces points.

Il convient maintenant de traduire par un critère mathématique clair la notion assez floue de "satisfaire *approximativement* un système d'équations" ou de "qui passe *au plus près* d'un ensemble de points". Le critère le plus usité est celui des moindres carrés : minimiser la somme des carrés des écarts entre la valeur calculée $f(x_k)$ de la fonction et sa valeur donnée y_k :

$$\mathcal{E}_{(a_1, \dots, a_i, \dots, a_p)}^2 = \sum_{k=1}^n (f(x_k) - y_k)^2$$

Il s'agit donc de calculer (a_1, a_2, \dots, a_p) tels que \mathcal{E}^2 soit minimum. Le ou les extremums de cette fonction de (a_1, a_2, \dots, a_p) sont recherchés en égalant à 0 les dérivées partielles respectives. On obtient un système de p équations à p inconnues :

$$\left\{ \frac{\partial \mathcal{E}^2}{\partial a_i} = 2 \sum_{k=1}^n (f(x_k) - y_k) \left(\frac{\partial f}{\partial a_i} \right)_{x=x_k} = 0 \right\}_{i=1 \text{ à } i=p}$$

La résolution conduit généralement à une solution (cas de système linéaire), ou éventuellement à plusieurs solutions (cas non linéaire). Il convient finalement de vérifier que l'extremum obtenu est bien un minimum et, dans le cas où l'on obtient plusieurs minimums, choisir la solution correspondant au plus petit d'entre eux.

Néanmoins, on peut se poser la question : pourquoi les moindres carrés ?

Au lieu de minimiser $\sum_{k=1}^n (f(x_k) - y_k)^2$, pourquoi ne pas minimiser $\sum_{k=1}^n |f(x_k) - y_k|$? , ou $\sum_{k=1}^n \sqrt{|f(x_k) - y_k|}$? , ou plus généralement pourquoi ne pas minimiser $\sum_{k=1}^n \Phi(f(x_k) - y_k)$? avec Φ une fonction paire et telle que $\Phi(0) = 0$.

En fait, d'une façon très générale, on ne voit pas de raison en faveur exclusivement de la méthode des moindres carrés. Certes, cette méthode a souvent l'avantage de simplifier les calculs. De plus, on cite classiquement le cas où elle correspond à l'estimateur du maximum de vraisemblance dans un modèle précis où x est la variable explicative, y la variable dépendante et où on suppose que le terme d'erreur aléatoire a une forme bien précise.

Mais en général, rien ne prouve, à priori, que la méthode des moindres carrés soit la plus appropriée dans le cadre de tel ou tel problème concret. En principe, on devrait donc choisir, au cas par cas, la méthode de régression la plus appropriée pour satisfaire le critère d'ajustement spécifique au problème que l'on traite.

De plus, il convient de conserver à l'esprit que le terme *moindres carrés* ne s'applique pas toujours à une somme de carrés de distances, d'abscisses ou d'ordonnées, mais parfois à des sommes de carrés d'expressions mathématiques n'ayant pas de signification géométrique.

On comprendra mieux ces généralités en considérant quelques cas typiques :
 Nous verrons d'abord, les régressions linéaires qui sont les applications les plus courantes et que l'on retrouve dans de très nombreux ouvrages, par exemple [1, 4, 5].
 Puis nous passerons à l'objet final de l'étude : La méthode de régression circulaire, qui sera décrite ici, a été publiée originalement dans [2].

2. REGRESSION LINEAIRE à un paramètre (Figure 2)

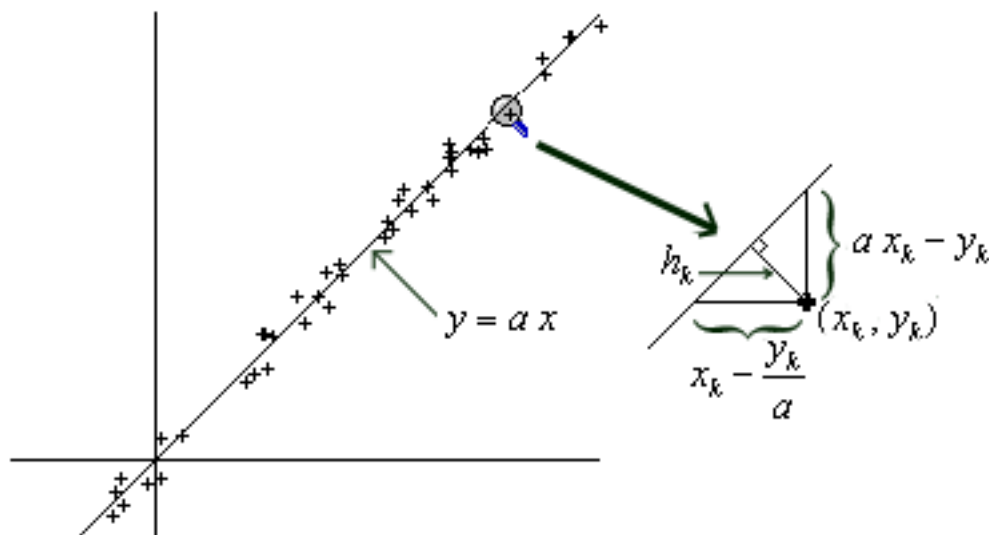


Figure 2 : Régression linéaire (paramètre a)

2.1. Moindres carrés des écarts d'ordonnées :

On cherche à ajuster le paramètre a de la fonction $f(x)=ax$ de telle sorte que la somme des carrés des écarts d'ordonnées soit minimum :

$$\mathcal{E}_{(a)}^2 = \sum_{k=1}^n (a x_k - y_k)^2$$
$$\frac{\partial \mathcal{E}^2}{\partial a} = 2 \sum_{k=1}^n (a x_k - y_k) x_k = 0 \quad \text{d'où : } a = \frac{\sum_{k=1}^n x_k y_k}{\sum_{k=1}^n (x_k)^2}$$

2.2. Moindres carrés des écarts d'abscisses :

Il est également possible de minimiser la somme des carrés des écarts d'abscisses, ce qui correspond à l'ajustement du paramètre $(1/a)$ de la fonction $f(y)=(1/a)y$:

$$\mathcal{E}_{(1/a)}^2 = \sum_{k=1}^n ((1/a) y_k - x_k)^2$$
$$\frac{\partial \mathcal{E}^2}{\partial (1/a)} = 2 \sum_{k=1}^n ((1/a) y_k - x_k) y_k = 0 \quad \text{d'où : } a = \frac{\sum_{k=1}^n (y_k)^2}{\sum_{k=1}^n x_k y_k}$$

2.3. Moindres carrés des écarts de distances :

Enfin, on peut chercher à minimiser la somme des carrés des distances entre les points et la droite :

$$\mathcal{E}_{(a)}^2 = \sum_{k=1}^n (h_k)^2 \quad \text{avec : } (h_k)^2 = \frac{1}{a^2 + 1} (a x_k - y_k)^2$$
$$\frac{\partial \mathcal{E}^2}{\partial (a)} = \frac{2}{(a^2 + 1)^2} \left(a \sum_{k=1}^n (x_k)^2 + (a^2 - 1) \sum_{k=1}^n x_k y_k - a \sum_{k=1}^n (y_k)^2 \right) = 0$$
$$\text{d'où : } a = c + \sqrt{c^2 + 1} \quad \text{avec } c = \frac{\sum_{k=1}^n (y_k^2 - x_k^2)}{2 \sum_{k=1}^n x_k y_k}$$

3. REGRESSION LINEAIRE à deux paramètres (Figure 3).

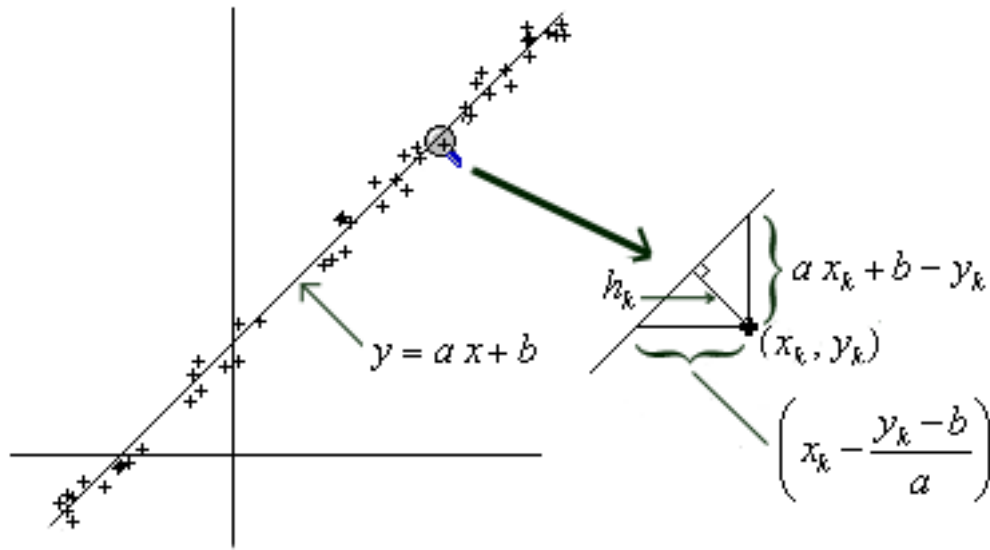


Figure 3 : Régression linéaire (paramètres a et b)

3.1. Moindres carrés des écarts d'ordonnées :

On cherche à ajuster les paramètres a et b de la fonction $f(x)=ax+b$ de telle sorte que la somme des carrés des écarts d'ordonnées soit minimum :

$$\mathcal{E}_{(a,b)}^2 = \sum_{k=1}^n ((a x_k + b) - y_k)^2$$

$$\begin{cases} \frac{\partial \mathcal{E}^2}{\partial a} = 2 \sum_{k=1}^n ((a x_k + b) - y_k) x_k = 0 \\ \frac{\partial \mathcal{E}^2}{\partial b} = 2 \sum_{k=1}^n ((a x_k + b) - y_k) = 0 \end{cases}$$

$$a = \frac{n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k}{n \sum_{k=1}^n (x_k)^2 - \left(\sum_{k=1}^n x_k \right)^2} ; \quad b = \frac{\sum_{k=1}^n y_k \sum_{k=1}^n (x_k)^2 - \sum_{k=1}^n x_k \sum_{k=1}^n x_k y_k}{n \sum_{k=1}^n (x_k)^2 - \left(\sum_{k=1}^n x_k \right)^2}$$

3.2. Moindres carrés des écarts d'abscisses :

On cherche à ajuster les paramètres a et b de la fonction $f(y)=(y-b)/a$ de telle sorte que la somme des carrés des écarts d'abscisses soit minimum. les formules sont les mêmes que les précédentes en en permutant x, y et en posant $\alpha=1/a$ et $\beta=-b/a$:

$$a = \frac{n \sum_{k=1}^n (y_k)^2 - \left(\sum_{k=1}^n y_k \right)^2}{n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k} ; \quad b = \frac{\sum_{k=1}^n y_k \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n (y_k)^2}{n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k}$$

2.3. Moindres carrés des écarts de distances :

Enfin, on peut chercher à minimiser la somme des carrés des distances entre les points et la droite :

$$\mathcal{E}_{(a)}^2 = \sum_{k=1}^n (h_k)^2 \quad \text{avec :} \quad (h_k)^2 = \frac{1}{a^2 + 1} ((a x_k + b) - y_k)^2$$

D'où le système d'équations ayant a et b pour inconnues :

$$\begin{cases} \frac{\partial \mathcal{E}^2}{\partial a} = \frac{2}{(a^2 + 1)^2} \left(a \sum_{k=1}^n x_k^2 - a \sum_{k=1}^n y_k^2 + (a^2 - 1) \sum_{k=1}^n x_k y_k + b(1 - a^2) \sum_{k=1}^n x_k + 2ab \sum_{k=1}^n y_k - n a b^2 \right) = 0 \\ \frac{\partial \mathcal{E}^2}{\partial b} = \frac{2}{a^2 + 1} \left(a \sum_{k=1}^n x_k - \sum_{k=1}^n y_k + n b \right) = 0 \end{cases}$$

En tirant b de la seconde équation et en reportant dans la première, on obtient après simplification une équation du second degré. Des explications plus détaillées sont données dans [5, p.1046]. La résolution conduit au résultat suivant :

$$\text{avec } c = \frac{1}{2} \frac{n \sum_{k=1}^n (y_k)^2 - \left(\sum_{k=1}^n y_k \right)^2 - n \sum_{k=1}^n (x_k)^2 + \left(\sum_{k=1}^n x_k \right)^2}{n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k}.$$

$$\text{on trouve : } a = c + \sqrt{c^2 + 1} \quad \text{et} \quad b = \frac{1}{n} \left(\sum_{k=1}^n y_k - a \sum_{k=1}^n x_k \right)$$

4. REGRESSION LINEAIRE (cas général)

Etant données p fonctions : $F_1(x), F_2(x), \dots, F_p(x)$, on considère la fonction $f(x)$ suivante

$$f(x) = a_1 F_1(x) + a_2 F_2(x) + \dots + a_p F_p(x) = \sum_{i=1}^p a_i F_i(x)$$

La fonction $f(x)$ comporte p paramètres (a_1, a_2, \dots, a_p). Ce sont eux que l'on va ajuster pour que sa courbe représentative passe au plus près des n points $(x_1, y_1), \dots, (x_k, y_k), \dots, (x_n, y_n)$.

Il s'agit encore d'une régression linéaire car les p paramètres apparaissent linéairement dans la formulation de la fonction, bien que les $F_1(x), F_2(x), \dots, F_p(x)$ soient quelconques, donc généralement non linéaires en fonction de x .

D'ailleurs, il suffit de replacer la question dans le contexte géométrique d'un système d'axes $OX_1, \dots, OX_i, \dots, OX_p$ gradués selon les fonctions $X_i = F_i(x)$ pour se ramener à l'ajustement linéaire d'un hyperplan dans cet espace à p dimensions.

Il est essentiel de ne pas confondre avec le cas d'une fonction $f(x; a_1, a_2, \dots, a_p)$ où les paramètres ne figurent pas en tant que simples coefficients, au quel cas nous parlerions de régression non-linéaire.

La méthode des moindres carrés s'applique de la même façon :

$$\mathcal{E}_{(a_1, \dots, a_i, \dots, a_p)}^2 = \sum_{k=1}^n (f(x_k) - y_k)^2 = \sum_{k=1}^n \left(\left(\sum_{i=1}^p a_i F_i(x_k) \right) - y_k \right)^2$$

Le système linéaire de p équations aux p inconnues (a_1, \dots, a_p) est obtenu en écrivant que les dérivées partielles respectives sont nulles :

$$\frac{\partial}{\partial a_j} \sum_{k=1}^n \left(\sum_{i=1}^p a_i F_i(x_k) - y_k \right)^2 = 2 \sum_{k=1}^n \left(\left(\sum_{i=1}^p a_i F_i(x_k) - y_k \right) F_j(x_k) \right) = 0$$

Pour condenser les écritures, posons :

$$B_{i,j} = \sum_{k=1}^n F_i(x_k) F_j(x_k) \quad \text{et} \quad C_i = \sum_{k=1}^n y_k F_i(x_k)$$

$$\begin{pmatrix} B_{1,1} & B_{2,1} & \dots & B_{p,1} \\ B_{1,2} & B_{2,2} & \dots & B_{p,2} \\ \dots & \dots & \dots & \dots \\ B_{1,p} & B_{2,p} & \dots & B_{p,p} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_p \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_p \end{pmatrix} \rightarrow (B)(A) = (C)$$

Si une fonction F_i est constante soit $F_i(x_k) = c$, on aura $B_{i,i} = n c^2$

Le système est résolu par $(A) = (B)^{-1} (C)$, ce qui donne les coefficients ajustés de la fonction non linéaire $f(x)$.

Ceci est bien connu et couvre de nombreux cas. Par exemple, si $F_i(x) = x^i$ il s'agit d'une régression dite "polynomiale de degré p ", bien que ramenée à un système linéaire. C'est le cas de la régression dite "parabolique", $f(x) = ax^2 + bx + c$ qui correspond à $F_1(x) = 1$, $F_2(x) = x$, $F_3(x) = x^2$, $a_1 = c$, $a_2 = b$ et $a_3 = a$.

Autre exemple : La régression improprement dite "logarithmique" correspondant au cas $f(x) = a_1 \ln(x) + a_2$ est en réalité une régression linéaire avec $F_1(x) = \ln(x)$ et $F_2(x) = 1$ ce qui apparaît de toute évidence dans un système avec l'axe des abscisses gradué logarithmiquement : la courbe de régression y est une droite.

5. AJUSTEMENT DES PARAMETRES D'UNE EQUATION

Nous allons maintenant considérer un problème un peu différent et qui sera d'une grande utilité pour l'application à la régression circulaire.

Etant donné une famille de fonctions $F_0(x,y), F_1(x,y), F_2(x,y), \dots, F_p(x,y)$, on considère l'équation suivante :

$$F_0(x, y) + a_1 F_1(x, y) + a_2 F_2(x, y) + \dots + a_p F_p(x, y) = 0$$

Il s'agit d'ajuster des paramètres (a_1, a_2, \dots, a_p) de telle sorte que des couples donnés $(x_1, y_1), \dots (x_k, y_k), \dots (x_n, y_n)$ la satisfasse au mieux. L'égalité ne peut pas être rigoureuse pour les n points donnés si $n > p$ ce qui est le cas considéré ici. Par conséquent, si l'on reporte dans cette équation le couple (x_k, y_k) , on observera un écart ε_k :

$$\varepsilon_k = F_0(x_k, y_k) + a_1 F_1(x_k, y_k) + \dots + a_p F_p(x_k, y_k) = F_0(x_k, y_k) + \sum_{i=1}^p a_i F_i(x_k, y_k)$$

Le critère considéré sera le minimum de la somme des carrés de ces écarts :

$$\varepsilon_{(a_1, \dots, a_p)}^2 = \sum_{k=1}^n \left(F_0(x_k, y_k) + \sum_{i=1}^p a_i F_i(x_k, y_k) \right)^2$$

Il serait inutile de répéter ce qui vient d'être écrit au paragraphe 4. En effet, on voit que le remplacement de y_k par $-F_0(x_k, y_k)$ et le remplacement des $F_i(x_k)$ par les $F_i(x_k, y_k)$ conduit au même système d'équations linéaires. Les mêmes formules permettent donc de calculer les (a_1, a_2, \dots, a_p) .

Bien évidemment, ceci peut être présenté d'une façon plus structurée dans des espaces à p dimensions. De plus, on voit les généralisations possibles à des fonctions comportant plus de variables, c'est à dire $F_i(x_k, y_k, z_k, t_k, \dots)$ au lieu de $F_i(x_k, y_k)$.

6. CAS DE REGRESSIONS NON LINEAIRES ET/OU DE CRITERE AUTRE QUE LES MOINDRES CARRES

Le principe général qui a été sommairement exposé en introduction reste valable. Par contre, la mise en œuvre devient ardue et la résolution n'est en général pas possible littéralement. Ce sont alors les méthodes numériques qu'il faut employer, avec tous leurs aléas dans le calcul des solutions d'équations non linéaires (bien que des progrès importants aient été faits et se poursuivent). Ce genre de problèmes et les algorithmes de calcul qui s'y rapportent suscitent de nombreuses études.

7. REGRESSION CIRCULAIRE (Figure 4)

La régression circulaire entrerait dans le cadre des cas difficiles mentionnés au paragraphe 6 si on considérait l'équation du cercle sous la forme d'une fonction de x :

$$y = f(x) = b \pm \sqrt{R^2 - (x-a)^2}$$

avec R = rayon du cercle et (a, b) coordonnées de son centre.

On dispose donc de trois paramètres ajustables (R, a, b) pour faire passer le cercle au plus près des n points donnés $(x_1, y_1), \dots (x_k, y_k), \dots (x_n, y_n)$.

Comme on l'a vu, l'application directe de la méthode des moindres carrés avec la fonction $y=f(x)$ correspondant au cercle et qui n'est pas linéaire par rapport à deux des paramètres (R et a) poserait de très grandes difficultés. Néanmoins, on peut contourner l'obstacle. Dans l'approche sensiblement différente que nous allons voir, nous nous éloignons de la simple notion de carré de distance entre deux points et cela n'est pas sans conséquence. Nous en discuterons plus tard.

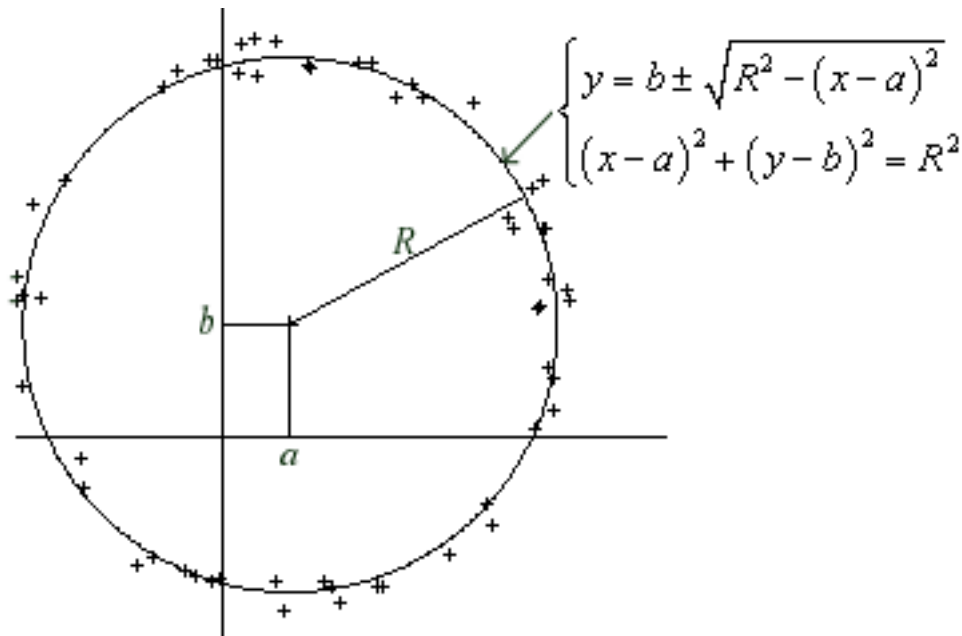


Figure 4 : Régression circulaire (3 paramètres : a , b et R)

Considérons une autre forme de présentation de l'équation du cercle (dans laquelle les paramètres R , a , b sont actuellement inconnus) :

$$R^2 - (x-a)^2 - (y-b)^2 = 0$$

Cette équation ne pourra pas être rigoureusement satisfaite pour les n points donnés. Par conséquent, si l'on reporte dans l'équation les coordonnées (x_k, y_k) d'un point, on observera un écart ε_k :

$$\varepsilon_k = R^2 - (x_k - a)^2 - (y_k - b)^2$$

$$\varepsilon_k = 2ax_k + 2by_k + c - x_k^2 - y_k^2 \quad \text{avec} \quad c = R^2 - a^2 - b^2$$

Le critère considéré sera le minimum de la somme des carrés de ces écarts :

$$\mathcal{E}_{(a,b,c)}^2 = \sum_{k=1}^n (2a x_k + 2b y_k + c - x_k^2 - y_k^2)^2$$

On reconnaît le problème d'ajustement des paramètres d'une équation qui a été considéré au paragraphe 5. La correspondance est : $F_0(x,y) = -x^2 - y^2$; $F_1(x,y) = 2x$; $F_2(x,y) = 2y$; $F_3(x,y) = 1$; $a_1 = a$; $a_2 = b$; $a_3 = c$.

Tout en suivant la méthode qui a été indiquée, nous allons donner le détail du développement, ce qui constitue l'objectif final du présent papier.

$$\begin{cases} \frac{\partial \mathcal{E}^2}{\partial a} = 4 \sum_{k=1}^n (2a x_k + 2b y_k + c - x_k^2 - y_k^2) x_k = 0 \\ \frac{\partial \mathcal{E}^2}{\partial b} = 4 \sum_{k=1}^n (2a x_k + 2b y_k + c - x_k^2 - y_k^2) y_k = 0 \\ \frac{\partial \mathcal{E}^2}{\partial c} = 2 \sum_{k=1}^n (2a x_k + 2b y_k + c - x_k^2 - y_k^2) = 0 \end{cases}$$

Ce qui donne le système de 3 équations linéaires :

$$\begin{pmatrix} 2 \sum x_k^2 & 2 \sum x_k y_k & \sum x_k \\ 2 \sum x_k y_k & 2 \sum y_k^2 & \sum y_k \\ 2 \sum x_k & 2 \sum y_k & n \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum x_k^3 + \sum x_k y_k^2 \\ \sum y_k^3 + \sum x_k^2 y_k \\ \sum x_k^2 + \sum y_k^2 \end{pmatrix}$$

Par substitution de c , tiré de la troisième équation et reporté dans les deux premières, le système se réduit à :

$$\begin{pmatrix} \delta_{20} & \delta_{11} \\ \delta_{11} & \delta_{02} \end{pmatrix} \begin{pmatrix} 2a \\ 2b \end{pmatrix} = \begin{pmatrix} \delta_{30} + \delta_{12} \\ \delta_{03} + \delta_{21} \end{pmatrix}$$

dans lequel on a posé, pour simplifier les écritures :

$$(1) \begin{cases} \delta_{11} = n \sum_{k=1}^n x_k y_k - \sum_{k=1}^n x_k \sum_{k=1}^n y_k & ; \\ \delta_{20} = n \sum_{k=1}^n x_k^2 - \left(\sum_{k=1}^n x_k \right)^2 & ; \quad \delta_{02} = n \sum_{k=1}^n y_k^2 - \left(\sum_{k=1}^n y_k \right)^2 \\ \delta_{30} = n \sum_{k=1}^n x_k^3 - \sum_{k=1}^n x_k^2 \sum_{k=1}^n x_k & ; \quad \delta_{03} = n \sum_{k=1}^n y_k^3 - \sum_{k=1}^n y_k \sum_{k=1}^n y_k^2 \\ \delta_{21} = n \sum_{k=1}^n x_k^2 y_k - \sum_{k=1}^n x_k^2 \sum_{k=1}^n y_k & ; \quad \delta_{12} = n \sum_{k=1}^n x_k y_k^2 - \sum_{k=1}^n x_k \sum_{k=1}^n y_k^2 \end{cases}$$

La résolution du système de deux équations linéaires donne les coordonnées (a,b) du centre du cercle :

$$(2) \begin{cases} a = \frac{(\delta_{30} + \delta_{12})\delta_{02} - (\delta_{03} + \delta_{21})\delta_{11}}{2(\delta_{20}\delta_{02} - \delta_{11}^2)} \\ b = \frac{(\delta_{03} + \delta_{21})\delta_{20} - (\delta_{30} + \delta_{12})\delta_{11}}{2(\delta_{20}\delta_{02} - \delta_{11}^2)} \end{cases}$$

Puis, en revenant à la 3ième équation du système initial, on obtient le rayon du cercle :

$$(3) \begin{cases} c = \frac{1}{n} \left(\sum_{k=1}^n x_k^2 + \sum_{k=1}^n y_k^2 - 2a \sum_{k=1}^n x_k - 2b \sum_{k=1}^n y_k \right) \\ R = \sqrt{c + a^2 + b^2} \end{cases}$$

En résumé, la méthode est la suivante :

- Calculer les δ par les formules (1)
- Calculer les coordonnées du centre $(a \text{ et } b)$ par les formules (2)
- Calculer c et le rayon R par les formules (3)

8. DISCUSSION

La méthode de régression qui a été suivie dans le cas du cercle est un cas particulier de la théorie générale des ajustements que l'on trouve abondamment exposée dans la littérature.

En fait, il s'agit d'un ajustement par les moindres carrés, mais non au sens traditionnel de "minimum de la somme des carrés des distances" entre le cercle et les points. Sa réelle signification est celle que nous avons vu au paragraphe 5 : un critère de "moindres carrés" est bien respecté, mais ce n'est pas exactement celui auquel on s'attend traditionnellement. La comparaison est la suivante :

$$\varepsilon_{k \text{ (utilisé)}}^2 = \left(R^2 - (x_k - a)^2 - (y_k - b)^2 \right)^2 \quad ; \quad \varepsilon_{k \text{ "traditionnel"}}^2 = \left(R - \sqrt{(x_k - a)^2 + (y_k - b)^2} \right)^2$$

Bien que $\varepsilon_{k \text{ (utilisé)}}^2$ ne soit pas égal à $\varepsilon_{k \text{ "traditionnel"}}^4$ il est quand même de dimension quatre, ce qui peut faire craindre une forte sensibilité aux éventuels points aberrants. Néanmoins, en pratique, ce possible défaut ne s'est pas manifesté dans les applications qui sont citées plus loin.

La liberté prise par rapport au critère traditionnel est-elle rédhibitoire ? Plusieurs arguments et observations conduisent à penser que ce n'est pas forcément le cas :

- On pourrait arguer que, dans le cas de la régression linéaire traditionnelle, c'est à dire l'ajustement d'une droite dans un nuage de points, on se sert la plupart du temps de la formule donnée au paragraphe 3.1. Cette formule correspond bien à un critère de moindre carrés, mais ce ne sont pas les carrés des distances *minimum* des points à la droite. Il faudrait utiliser les formules du paragraphe 3.3 qui sont nettement plus compliquées. Beaucoup ne le font pas et se contentent donc, parfois sans s'en rendre compte, d'un critère qui n'est pas rigoureusement celui du "minimum de la somme des carrés des distances" des points à la droite. Cela n'empêche pourtant pas ceux qui utilisent les formules "simplifiées" d'en être satisfait.

- On pourrait aussi arguer que le critère des moindres carrés n'a pas n'est pas forcément le plus approprié (sauf dans des cas classiques qui ont été évoqués en introduction) et que d'autres critères peuvent être recherchés selon le cas que l'on traite. Souvent, il est important que tous les écarts restent dans une fourchette de tolérance pour l'acceptation d'un produit fabriqué. Dans ce cas, le critère des moindres carrés n'est pas le meilleur, puisque c'est le plus grand des écarts qui est à considérer. Bien que ne répondant pas strictement à ce critère, on a pu constater qu'à ce point de vue la considération de la somme des écarts à la puissance quatre donne de meilleurs résultats que la somme des écarts à la puissance deux. Ceci intervient peut-être dans le succès de la méthode de régression circulaire faisant l'objet de la présente discussion.

- Egalement en introduction, il avait été signalé que le critère des moindres carrés des distances à une droite simplifie les calculs. Dans le cas de la régression circulaire, le même avantage de simplicité des calculs guide vers la méthode qui a été exposée au paragraphe 7.

Tous ces arguments n'ont pas de valeur universelle et sont à prendre, ou non, en considération selon les cas. On voit qu'il est difficile de conclure définitivement sur un plan purement théorique.

En pratique, l'usage qui en a été fait pour l'analyse statistique de certains diagrammes d'impédance [2] (un exemple est présenté en annexe), s'est avéré satisfaisant et a eu des applications dans le domaine de la R.&D., ainsi qu'en contrôle-qualité.

Ultérieurement, la méthode a été communiquée à plusieurs demandeurs pour des applications diverses (par exemple : caractérisation des paramètres de trajectoires circulaires en cinématique, positionnement de système optique rotatif, etc.). Les commentaires reçus en retour ont été favorables et aucune contre-indication n'a été signalée à ce jour.

ANNEXE : EXEMPLE

Il est parfois utile de disposer d'un exemple numérique lorsque l'on écrit un programme de calcul, afin de le vérifier. A cet effet, la figure 5 donne des résultats graphiques et numériques calculés avec les formules des δ_{ij} : (1), puis a, b : (2) et R : (3) du §.7, ainsi que :

$$x_{cal} = a + R \frac{x_k - a}{\sqrt{(x_k - a)^2 + (y_k - b)^2}} \quad ; \quad y_{cal} = b + R \frac{y_k - b}{\sqrt{(x_k - a)^2 + (y_k - b)^2}}$$

Cet exemple est retranscrit d'après un cas réel de mesures sur un réseau résistif-capacitif fabriqué de façon à ce que son impédance nominale soit proche de 2000 Ω à basse fréquence et 500 Ω à haute fréquence et possédant certaines propriétés particulières : équivalence avec la mise en parallèle d'une résistance pure et d'un composant spécial, '*phasor*'[2], à déphasage indépendant de la fréquence. Cette propriété particulière, désignée par '*phasance*', ainsi que d'autres termes techniques relatifs au même sujet et qu'il serait trop long de définir ici, ont fait l'objet d'une présentation synoptique dans un article antérieur de Quadrature [3] auquel on pourra se reporter si nécessaire.

En application pratique, les mesures de $Z' = x$ et de $-y = Z''$ (autrement dit, les mesures d'impédance complexe $Z = Z' + iZ''$) sont suivies d'une analyse statistique qui sort du

cadre du présent exposé. Nous nous limitons ici au calcul à posteriori de l'écart quadratique moyen (EQM) au sens "traditionnel" (§.8), dont la valeur est portée sur la figure 5 :

$$\varepsilon_{\text{quadratique.moy.}} = \sqrt{\frac{1}{n} \sum_{k=1}^n (\varepsilon_k)^2} \quad \text{avec} \quad \varepsilon_k = \left| R - \sqrt{(x_k - a)^2 + (y_k - b)^2} \right|$$

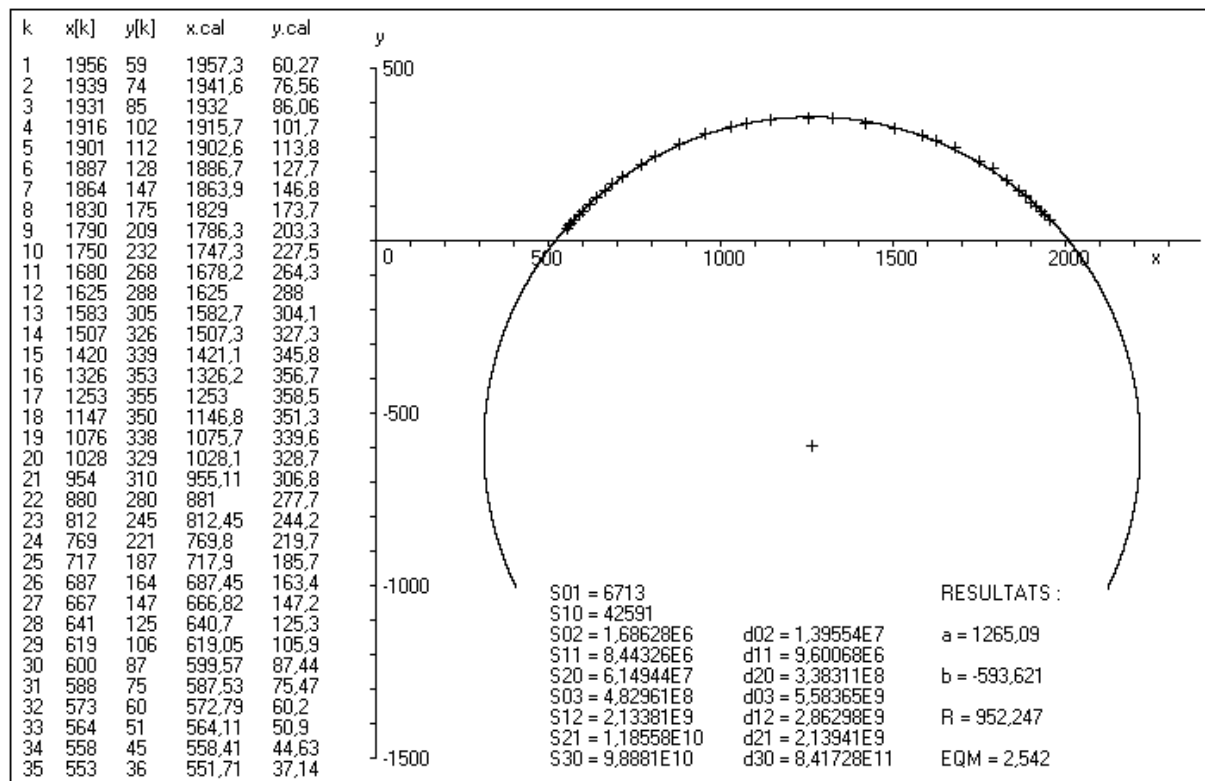


Figure 5 : Exemple numérique de régression circulaire.

REFERENCES :

- [1]: G.Baillargeon, "Probabilités, statistiques et techniques de régression", Edit. SMG (Quebec), p.433-434, 1989.
- [2]: J.Jacquelin, "Use of fractional derivatives to express the properties of energy storage phenomena in electrical networks" (1982), App.7 : *Statistical analysis of semi-circular impedance diagrams*. pp.70-79, Ré-édit. 1990 par Laboratoires de Marcoussis, Route de Nozay, F91460 Marcoussis.
- [3] J.Jacquelin, "La dérivation fractionnaire", Quadrature No.40, pp.10-12, 2000.
- [4]: F.Reinhardt, H.Soeder, "Atlas des mathématiques", *Méthode des moindres carrés*, Le Livre de Poche, Librairie Générale Française, p.313, 1997.
- [5]: E.W.Weisstein, "CRC Concise Encyclopedia of Mathematics", *Least Squares fitting*, pp.1045-1049, Chapman & Hall, N.-Y., 1999.

Deuxième partie : Généralisations

La publication dans *QUADRATURE* n°63, pp.33-40, de l'article "Régression circulaire", a suscité des questions quant aux possibles généralisations, soit dans le plan (ellipse, parabole, hyperbole), soit dans l'espace (sphère et plus généralement les quadriques).

En effet, allant au delà de la régression circulaire, le procédé une fois adapté à l'optimisation des paramètres de courbes ou de surfaces quadratiques, présente un intérêt pratique dans des domaines tels que l'étude de trajectoires de particules, l'imagerie médicale, ...

Ce bref complément à l'article précité se propose d'en expliciter les équations de base.

GENERALISATION AUX CONIQUES :

Il s'agit de la famille de courbes planes satisfaisant l'équation :

$$a_{0,2} y^2 + a_{2,0} x^2 + a_{1,1} x y + a_{0,1} y + a_{1,0} x + 1 = 0$$

$$\sum_{i,j} a_{i,j} x^i y^j = 0 \quad \begin{cases} 0 \leq i, j \leq 2 \\ i + j \leq 2 \\ a_{0,0} = 1 \end{cases}$$

Les données sont les coordonnées de n points expérimentaux ($n \geq 5$) : (x_1, y_1) , ..., (x_k, y_k) , ..., (x_n, y_n) . Les inconnues sont cinq coefficients $a_{i,j}$ que l'on cherche à optimiser de telle sorte que l'équation soit approximativement et au mieux, satisfaite.

La méthode générale indiquée dans la section V de l'article cité, se traduit par le système suivant, avec les sommes de $k=1$ à n calculées préalablement :

$$\begin{pmatrix} \sum y_k^4 & \sum x_k^2 y_k^2 & \sum x_k y_k^3 & \sum y_k^3 & \sum x_k y_k^2 \\ \sum x_k^2 y_k^2 & \sum x_k^4 & \sum x_k^3 y_k & \sum x_k^2 y_k & \sum x_k^3 \\ \sum x_k y_k^3 & \sum x_k^3 y_k & \sum x_k^2 y_k^2 & \sum x_k y_k^2 & \sum x_k^2 y_k \\ \sum y_k^3 & \sum x_k^2 y_k & \sum x_k y_k^2 & \sum y_k^2 & \sum x_k y_k \\ \sum x_k y_k^2 & \sum x_k^3 & \sum x_k^2 y_k & \sum x_k y_k & \sum x_k^2 \end{pmatrix} \begin{pmatrix} a_{0,2} \\ a_{2,0} \\ a_{1,1} \\ a_{0,1} \\ a_{1,0} \end{pmatrix} = - \begin{pmatrix} \sum y_k^2 \\ \sum x_k^2 \\ \sum x_k y_k \\ \sum y_k \\ \sum x_k \end{pmatrix}$$

Ce qui peut s'écrire sous une forme condensée :

$$\left[\sum_{k=1}^n x_k^{i+\lambda} y_k^{j+\mu} \right] [a_{i,j}] = - \left[\sum_{k=1}^n x_k^\lambda y_k^\mu \right] \quad \begin{cases} 0 \leq i, j, \lambda, \mu \leq 2 \\ 0 < i + j \leq 2 \\ 0 < \lambda + \mu \leq 2 \end{cases}$$

La résolution de ce système linéaire donne les paramètres de la conique recherchée.

GENERALISATION AUX QUADRIQUES :

Il s'agit de la famille de surfaces en 3D, satisfaisant l'équation :

$$\sum_{i,j,h} a_{i,j,h} x^i y^j z^h = 0 \quad \begin{cases} 0 \leq i, j, h \leq 2 \\ i + j + h \leq 2 \\ a_{0,0,0} = 1 \end{cases}$$

Les données sont les coordonnées de n points expérimentaux ($n \geq 9$) : $(x_1, y_1, z_1), \dots, (x_k, y_k, z_k), \dots, (x_n, y_n, z_n)$. Les inconnues sont neuf coefficients $a_{i,j,h}$ que l'on cherche à optimiser de telle sorte que l'équation soit approximativement et au mieux, satisfaite.

La forme matricielle du système de neuf équations s'écrit similairement à ce qui précède :

$$\left[\sum_{k=1}^n x_k^{i+\lambda} y_k^{j+\mu} z_k^{h+\nu} \right] [a_{i,j,h}] = - \left[\sum_{k=1}^n x_k^\lambda y_k^\mu z_k^\nu \right] \quad \begin{cases} 0 \leq i, j, k, \lambda, \mu, \nu \leq 2 \\ 0 < i + j + h \leq 2 \\ 0 < \lambda + \mu + \nu \leq 2 \end{cases}$$

La résolution de ce système linéaire donne les paramètres de la quadrique recherchée.

LA REGRESSION SPHERIQUE :

Ce cas particulier de quadrique présente un intérêt spécifique pour des problèmes technologiques en 3D, au même titre que la régression circulaire en 2D.

L'équation de la sphère comporte quatre paramètres (les coordonnées de son centre : a, b, c et son rayon R) :

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = R^2$$

que l'on écrit : $(x^2 + y^2 + z^2) + xA_1 + yA_2 + zA_3 + A_0 = 0$

Avec les relations suivantes, correspondant aux notations du paragraphe précédent :

$$A_0 = a^2 + b^2 + c^2 - R^2 \quad ; \quad A_1 = -2a \quad ; \quad A_2 = -2b \quad ; \quad A_3 = -2c$$

$$a_{2,0,0} = a_{0,2,0} = a_{0,0,2} = \frac{1}{A_0} \quad ; \quad a_{1,1,0} = a_{1,0,1} = a_{0,1,1} = 0$$

$$a_{1,0,0} = \frac{A_1}{A_0} \quad ; \quad a_{0,1,0} = \frac{A_2}{A_0} \quad ; \quad a_{0,0,1} = \frac{A_3}{A_0} \quad ; \quad a_{0,0,0} = 1$$

et en posant : $\rho_k^2 = x_k^2 + y_k^2 + z_k^2$ le système à résoudre se réduit à :

$$\begin{pmatrix} n & \sum x_k & \sum y_k & \sum z_k \\ \sum x_k & \sum x_k^2 & \sum x_k y_k & \sum x_k z_k \\ \sum y_k & \sum x_k y_k & \sum y_k^2 & \sum y_k z_k \\ \sum z_k & \sum x_k z_k & \sum y_k z_k & \sum z_k^2 \end{pmatrix} \begin{pmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{pmatrix} = - \begin{pmatrix} \sum \rho_k^2 \\ \sum \rho_k^2 x_k \\ \sum \rho_k^2 y_k \\ \sum \rho_k^2 z_k \end{pmatrix}$$

La résolution donne A_0, A_1, A_2 et A_3 , puis les coordonnées du centre de la sphère :

$$a = -\frac{1}{2} A_1, b = -\frac{1}{2} A_2, c = -\frac{1}{2} A_3 \text{ et son rayon : } R = \sqrt{a^2 + b^2 + c^2 - A_0}$$

Remarque :

Sur la base générale qui a été exposée, de nombreux cas particuliers peuvent se greffer, si des contraintes complémentaires sont imposées, par exemple un (ou des) paramètre(s) manquant(s) dans l'équation de la conique ou de la quadrique. En particulier, c'est le cas lorsqu'on impose que la conique ou la quadrique passe par l'origine du système d'axes.

Par exemple, dans le cas de la régression sphérique, si l'on impose que la sphère passe exactement par l'origine, cela correspond au cas où le paramètre A_0 est absent (ce qui n'est pas la même chose que A_0 présent et supposé égal à zéro : la régression donnerait une valeur de A_0 voisine de zéro, mais pas exactement nulle).

La méthode doit alors être adaptée au cas par cas, ce qui ne pose pas de difficulté, en éliminant de l'équation initiale le (ou les) paramètre(s) inutile(s) et en reconstruisant le système d'équations et la matrice correspondante, dont la taille se trouve réduite.

Troisième partie : Cas particuliers.

REGRESSION POUR DEUX DROITES :

L'équation d'une courbe constituée de deux droites $y = p_1x + q_1$ et $y = p_2x + q_2$ est :

$$(p_1x + q_1 - y)(p_2x + q_2 - y) = 0$$

$$y^2 + p_1p_2x^2 - (p_1 + p_2)xy - (q_1 + q_2)y + (p_1q_2 + p_2q_1)x + q_1q_2 = 0$$

Il s'agit donc d'une conique dégénérée.

L'équation d'une conique non dégénérée voisine serait, avec λ voisin de 1 :

$$y^2 + p_1p_2x^2 - (p_1 + p_2)xy - (q_1 + q_2)y + (p_1q_2 + p_2q_1)x + \lambda q_1q_2 = 0$$

Conformément aux notations précédemment utilisées (page 16) :

$$a_{0,2}y^2 + a_{2,0}x^2 + a_{1,1}xy + a_{0,1}y + a_{1,0}x + 1 = 0$$

$$a_{0,2} = \frac{1}{\lambda q_1q_2} ; \quad a_{2,0} = \frac{p_1p_2}{\lambda q_1q_2} ; \quad a_{1,1} = -\frac{p_1 + p_2}{\lambda q_1q_2}$$

$$a_{0,1} = -\frac{q_1 + q_2}{\lambda q_1q_2} ; \quad a_{1,0} = \frac{p_1q_2 + p_2q_1}{\lambda q_1q_2}$$

Ainsi que nous l'avons vu, les valeurs approchées de ces coefficients sont calculés par régression linéaire ($1 \leq k \leq n$) :

$$\begin{pmatrix} a_{0,2} \\ a_{2,0} \\ a_{1,1} \\ a_{0,1} \\ a_{1,0} \end{pmatrix} = \begin{pmatrix} \sum y_k^4 & \sum x_k^2 y_k^2 & \sum x_k y_k^3 & \sum y_k^3 & \sum x_k y_k^2 \\ \sum x_k^2 y_k^2 & \sum x_k^4 & \sum x_k^3 y_k & \sum x_k^2 y_k & \sum x_k^3 \\ \sum x_k y_k^3 & \sum x_k^3 y_k & \sum x_k^2 y_k^2 & \sum x_k y_k^2 & \sum x_k^2 y_k \\ \sum y_k^3 & \sum x_k^2 y_k & \sum x_k y_k^2 & \sum y_k^2 & \sum x_k y_k \\ \sum x_k y_k^2 & \sum x_k^3 & \sum x_k^2 y_k & \sum x_k y_k & \sum x_k^2 \end{pmatrix}^{-1} \begin{pmatrix} -\sum y_k^2 \\ -\sum x_k^2 \\ -\sum x_k y_k \\ -\sum y_k \\ -\sum x_k \end{pmatrix}$$

On en déduit :

$$p_1 = \frac{-a_{11} + \sqrt{a_{11}^2 - 4a_{20}a_{02}}}{2a_{02}} ; \quad p_2 = \frac{-a_{11} - \sqrt{a_{11}^2 - 4a_{20}a_{02}}}{2a_{02}}$$

$$q_1 = \frac{a_{1,0} + a_{0,1}p_1}{a_{0,2}(p_2 - p_1)} ; \quad q_2 = -\frac{a_{1,0} + a_{0,1}p_2}{a_{0,2}(p_2 - p_1)} ; \quad \lambda = \frac{1}{a_{0,2}q_1q_2}$$

$$x_c = \frac{q_2 - q_1}{p_1 - p_2} ; \quad y_c = \frac{p_2q_1 - p_1q_2}{p_1 - p_2}$$

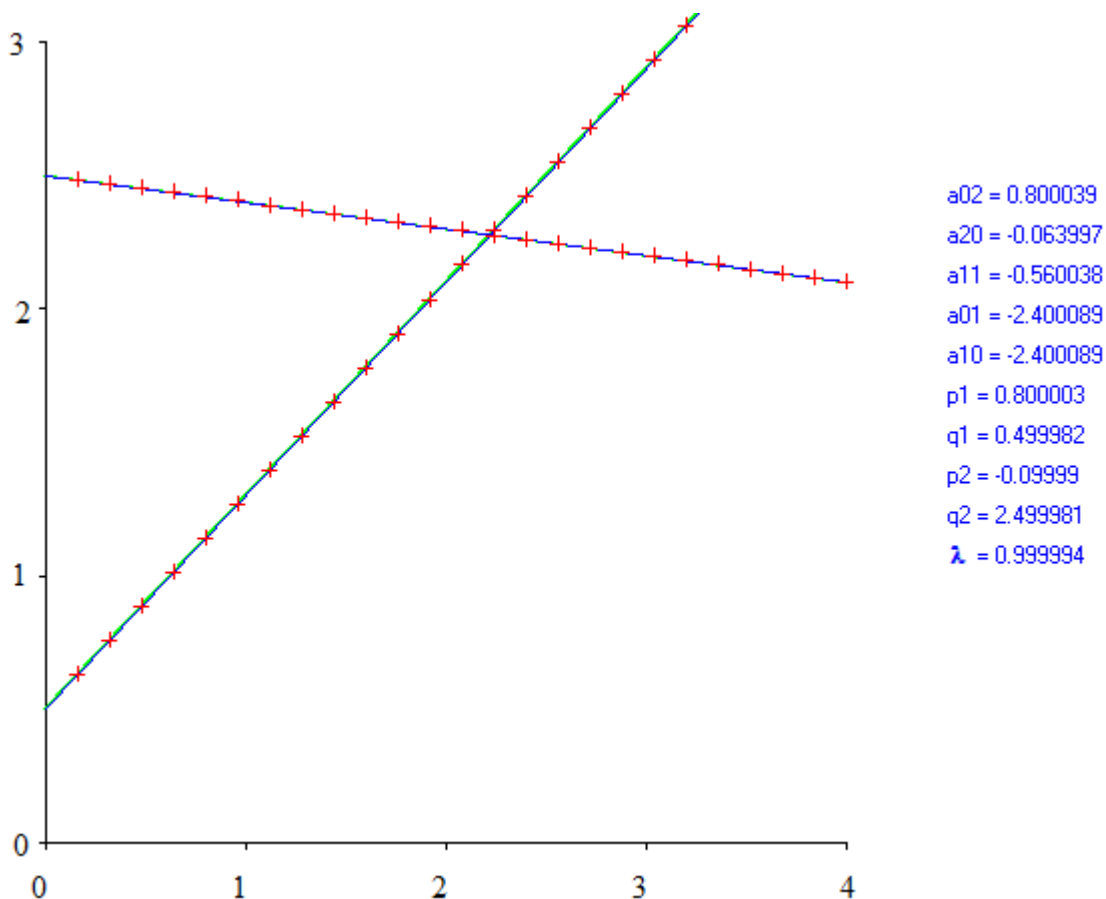
(x_c, y_c) est le centre de l'hyperbole, le point de concours des asymptotes.

Exemple numérique :

A partir d'un même jeu de données, on a appliqué des dispersions de plus en plus grandes, conduisant aux résultats représentés sur les cinq figures suivantes. La première figure est l'exemple exempt de dispersion.

Pour chaque exemple, le résultat du calcul de régression, selon les formules précédentes, est inscrit en bleu.

La courbe d'équation $a_{0,2} y^2 + a_{2,0} x^2 + a_{1,1} x y + a_{0,1} y + a_{1,0} x + 1 = 0$ est tracée en vert. C'est une hyperbole dont les asymptotes sont tracées en bleu.



Dans ce premier cas, sans dispersion, l'hyperbole est confondue avec ses asymptotes, ce qui donne donc un résultat quasiment exact. On constate bien que λ est très voisin de 1.

Dans les cas suivants, plus la dispersion des points donnés est grande, plus λ s'éloigne de 1 et plus l'hyperbole s'écarte des asymptotes. Si l'on considère que les asymptotes constituent une approximation des deux droites recherchées, cette approximation est d'autant moins bonne.

On en conclura que cette méthode de régression pour deux droites ne convient que dans les cas où les données numériques ne sont pas affectées de trop grande dispersion. Néanmoins, dans des cas de forte dispersion, elle pourrait servir à calculer une solution grossière, utile pour initialiser des calculs itératifs, au lieu de devoir recourir à des estimations initiales généralement difficiles.

