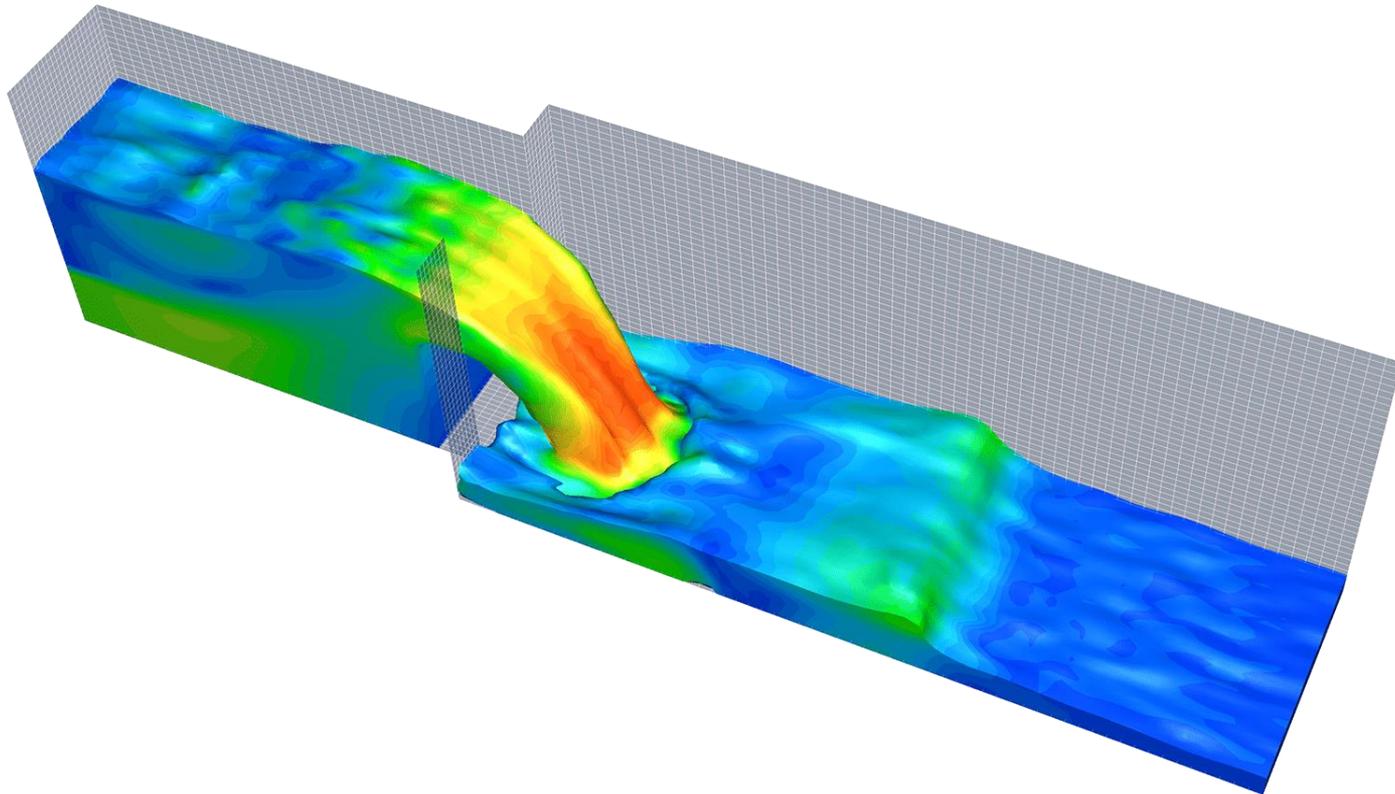


Modélisation numérique en génie chimique

GCH2535

Analyse de données expérimentales



**POLYTECHNIQUE
MONTRÉAL**

LE GÉNIE
EN PREMIÈRE CLASSE

Diapositives adaptées de :
David Vidal
Bruno Blais
François Bertrand



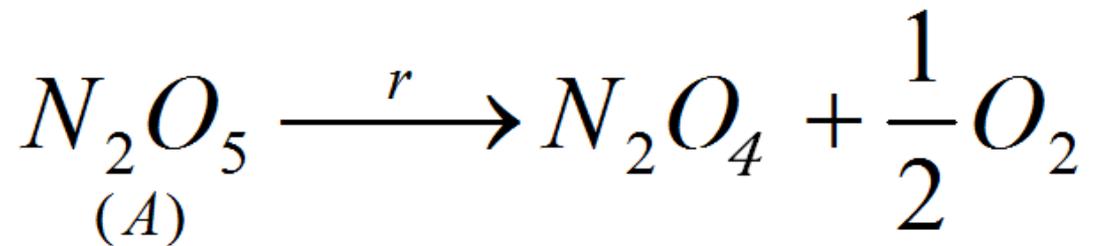
Semaine	Date	Séance	Notes	Sujet
2	Mercredi 8 Jan	1		EDP
2	Jeudi 9 Jan	2		EDP
3	Mercredi 15 Jan	3		EDP
3	Jeudi 16 Jan	4	Départ Devoir 1 (EDP)	EDP
4	Mercredi 22 Jan	5		EDP
4	Jeudi 23 Jan	6		TD#1
5	Mercredi 29 Jan	7	Remise Devoir 1 (EDP)	Intro
5	Jeudi 30 Jan	8		Intro / EDO
6	Mercredi 5 Fév	9	Partiel 1 (EDP)	CP#1
6	Jeudi 6 Fév	10		EDO
7	Mercredi 12 Fév	11	Départ Devoir 2 (MDF)	TD#2
7	Jeudi 13 Fév	12		MDF
8	Mercredi 19 Fév	13		MDF
8	Jeudi 20 Fév	14		MEF
9	Mercredi 26 Fév	15		TD#3 (LAB-MDF)
9	Jeudi 27 Fév	16		MEF
10	Mercredi 4 Mar		Relâche	
10	Jeudi 5 Mar		Relâche	
11	Mercredi 11 Mar	17	Remise Devoir 2 (MDF)	TD#4 (LAB-MEF)
11	Jeudi 12 Mar	18	Partiel 2 (EDO-MDF-MEF)	CP#2
12	Mercredi 18 Mar	19		Données exp.
12	Jeudi 19 Mar	20	Départ Devoir 3	Données exp.
13	Mercredi 25 Mar	21		Bilans
13	Jeudi 26 Mar	22		TD#5
14	Mercredi 1 Avr	23		Bilans
14	Jeudi 2 Avr	24		Opt./Rec.
15	Mercredi 8 Avr	25	Remise Devoir 3	Opt./Rec.
15	Jeudi 9 Avr	26		TD#6

Analyse de données expérimentales

- Régression
 - Régression simple
 - Régression pondérée
 - Régression multivariée
 - Critère des moindres carrés
 - Qualité de la régression
 - R et S
 - Régression non linéaire
 - Méthode de Gauss-Newton
- Interpolation
 - Polynômes
 - Splines

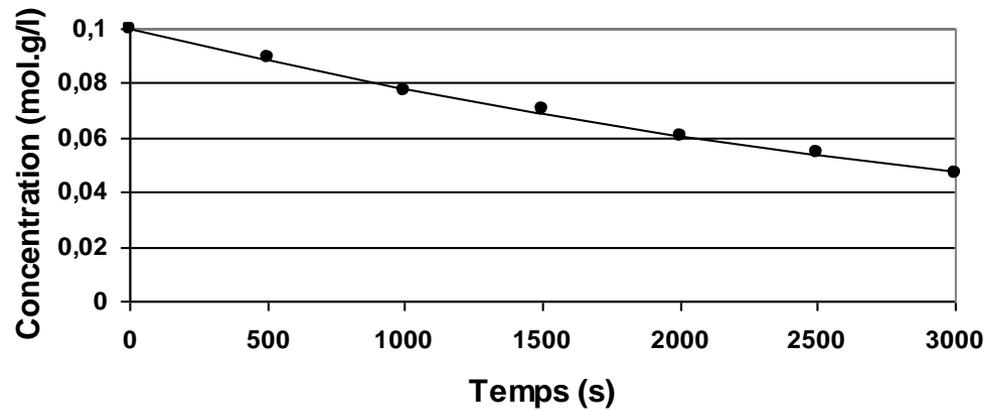
Régression simple au sens des moindres carrés

- Exemple: Décomposition de l'anhydride nitreux N_2O_5 en tétraoxyde de diazote à 313.3K



- On a les données expérimentales suivantes:

Temps (s)	Concentration de N_2O_5 (mol.g/l)
0	0.1000
500	0.0892
1,000	0.0776
1,500	0.0705
2,000	0.0603
2,500	0.0542
3,000	0.0471



On suppose que la réaction est du 1^{er} ordre:

$$r = KC_A \text{ (vitesse de réaction)}$$

$$\left\| \begin{array}{l} \frac{dC_A}{dt} = -r = -KC_A \\ C_A(t=0) = C_{A0} \end{array} \right.$$

Solution

$$\left\| \begin{array}{l} C_A = C_{A0} e^{-Kt} \\ \log\left(\frac{C_A}{C_{A0}}\right) = -Kt \end{array} \right.$$

Remarque: dans un diagramme semi-logarithmique de (C_A/C_{A0}) en fonction du temps, (C_A/C_{A0}) devrait être une droite de pente $-K$.

Question: Comment déterminer K?

Régression simple au sens des moindres carrés

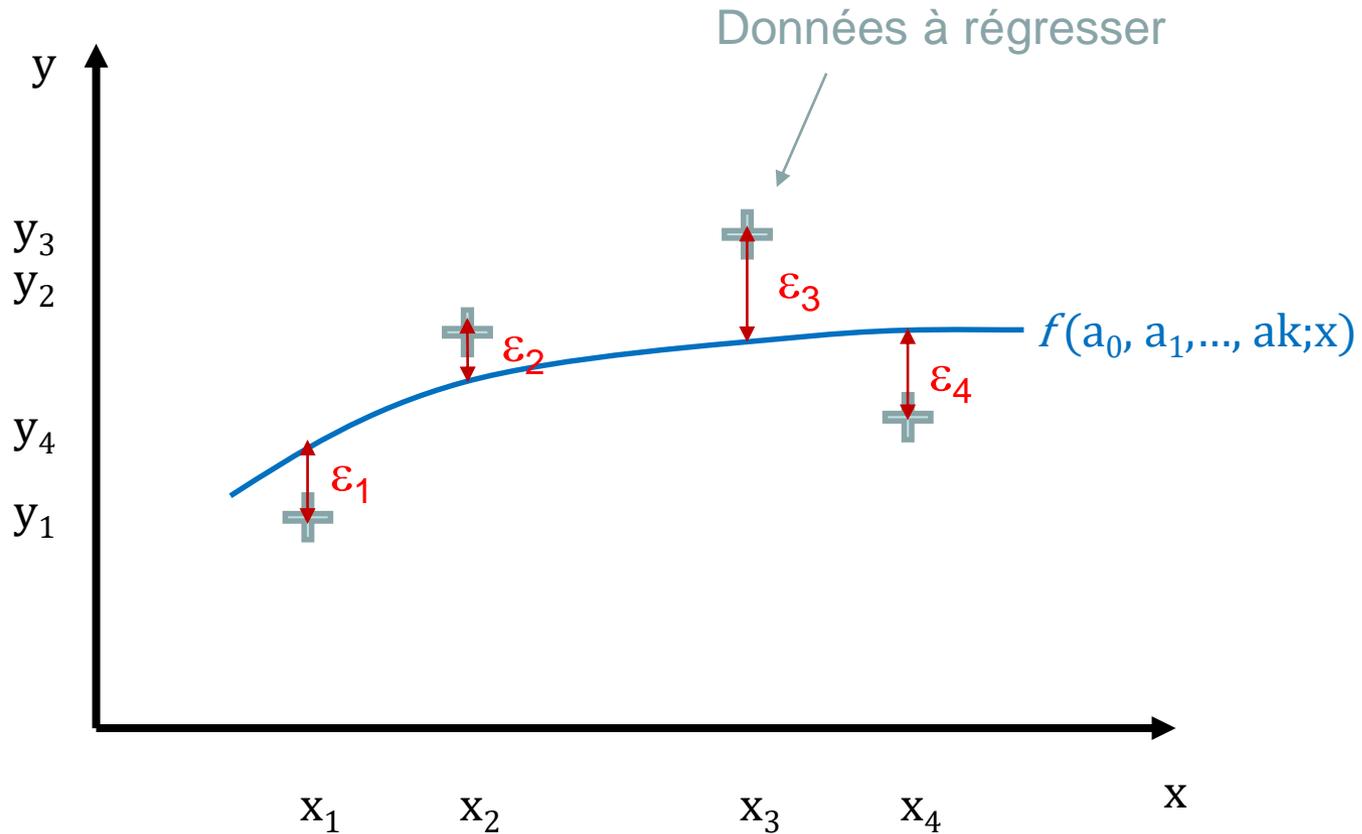
Supposons connus N points (x_i, y_i) , $i=1 \dots N$.

Considérons une fonction $f(a_0, a_1, \dots, a_k; x)$ où les paramètres a_i , $i = 0, \dots, k$ sont à déterminer de façon à minimiser le critère des moindres carrés:

$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^N \underbrace{|f(a_0, a_1, \dots, a_k; x_i) - y_i|}_{\varepsilon_i}^2 \quad (1)$$

erreur de la polynomiale de degré k au point i considéré
= résidu

Condition nécessaire pour un minimum: les dérivées partielles de $(\delta J / \delta a_i)$, $i = 0, \dots, k$ doivent être nulles. On a donc à résoudre un système de $(k+1)$ équations non linéaires à $(k+1)$ inconnues.



$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^4 \varepsilon_i^2 = (\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2)$$

Nous allons trouver les paramètres a_j de la polynomiale de degré k tel que la somme des erreurs au carré est minimum \rightarrow moindres carrés

$\rightarrow (\delta J / \delta a_i), i = 0, \dots, k$ doivent être nulles

Cas particulier

f est un polynôme de degré k avec $k+1$ paramètres:

$$f = a_0 + a_1x + a_2x^2 + \dots + a_kx^k$$

Dans ce cas, les équations à résoudre sont linéaires.

On a en effet

$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^N |f(a_0, a_1, \dots, a_k; x_i) - y_i|^2$$

Donc

$$\frac{\delta J}{\delta a_j} = 2 \sum_{i=1}^N x_i^j (a_0 + a_1x_i + a_2x_i^2 + \dots + a_kx_i^k - y_i)$$

$$\text{Rappel: } \frac{\partial U^n(a)}{\partial a} = n U^{n-1}(a) \frac{\partial U(a)}{\partial a}$$

Et le système linéaire à résoudre de $k+1$ équations à $k+1$ inconnues est

$$\mathbf{A}a = b \quad (2)$$

Où

$$\mathbf{A}_{kl} = \mathbf{A}_{lk} = \sum_{i=1}^N x_i^{k+l-2}$$

$$\mathbf{a} = (a_0, a_1, \dots, a_k)^T$$

$$b_k = \sum_{i=1}^N x_i^{k-1} \cdot y_i$$

Par exemple, pour 6 inconnues, on a:

$$\begin{bmatrix} N & \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^4 & \sum_{i=1}^N x_i^5 \\ & \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^4 & \sum_{i=1}^N x_i^5 & \sum_{i=1}^N x_i^6 \\ & & \sum_{i=1}^N x_i^4 & \sum_{i=1}^N x_i^5 & \sum_{i=1}^N x_i^6 & \sum_{i=1}^N x_i^7 \\ & & & \sum_{i=1}^N x_i^6 & \sum_{i=1}^N x_i^7 & \sum_{i=1}^N x_i^8 \\ & & & & \sum_{i=1}^N x_i^8 & \sum_{i=1}^N x_i^9 \\ & & & & & \sum_{i=1}^N x_i^{10} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N x_i^2 y_i \\ \sum_{i=1}^N x_i^3 y_i \\ \sum_{i=1}^N x_i^4 y_i \\ \sum_{i=1}^N x_i^5 y_i \end{bmatrix}$$

sym.

Remarques

1. Le système matriciel (2) est en général très mal conditionné et le conditionnement se détériore lorsque que k augmente (en pratique $3 \geq k$).
2. On peut souvent par des transformations astucieuses linéariser certaines équations.

Exemple: Décomposition de l'anhydride nitreux

$$\log\left(\frac{C_A}{C_{A0}}\right) = -Kt$$

On pose

$$y = \log\left(\frac{C_A}{C_{A0}}\right)$$

et

$$\alpha = -K$$

On a donc

$$f = \alpha t \quad \longleftarrow \quad t \equiv x$$

Le paramètre α est donné par la minimisation de

$$J(\alpha) = \sum_{i=1}^N |\alpha t_i - y_i|^2$$

où les y_i sont les valeurs mesurées de C_A à $t=t_i$.

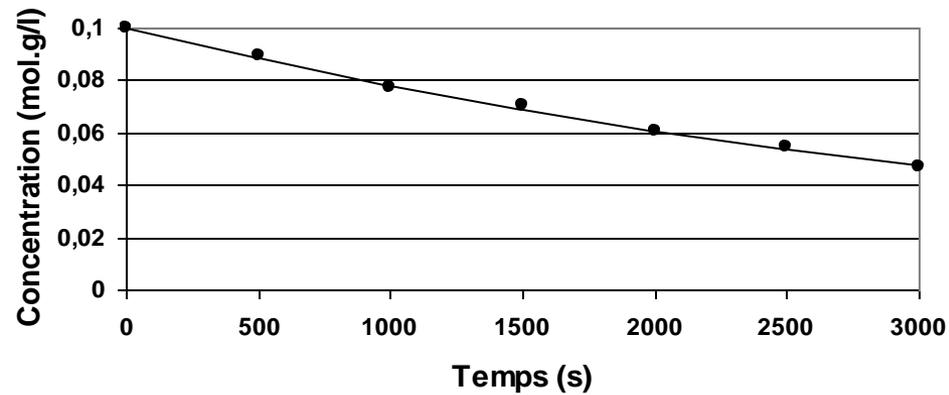
$$\frac{\delta J}{\delta \alpha} = 2 \sum_{i=1}^N t_i (\alpha t_i - y_i) = 0$$

$$\alpha \sum_{i=1}^N t_i^2 = \sum_{i=1}^N t_i y_i \rightarrow \alpha = \frac{\sum_{i=1}^N t_i y_i}{\sum_{i=1}^N t_i^2}$$

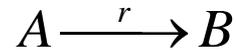
Application numérique

t	Ca	(Ca/Ca0)	ln(Ca/Ca0)	t*y	t*t	alfa
(s)	(mol/l)		y			(1/s)
0	0,1	1	0	0	0	
500	0,0892	0,892	-0,11428915	-57,1445732	250000	
1000	0,0776	0,776	-0,25360276	-253,602759	1000000	
1500	0,0705	0,705	-0,34955748	-524,336214	2250000	
2000	0,0603	0,603	-0,50583808	-1011,67616	4000000	
2500	0,0542	0,542	-0,61248928	-1531,22319	6250000	
3000	0,0471	0,471	-0,75289718	-2258,69155	9000000	

-5636,67446 22750000 -2,48E-04



Exemple: Soit une réaction chimique d'ordre n



$$r = K(C_A)^n \text{ (vitesse de réaction)}$$

$$\frac{dC_A}{dt} = -r = -KC_A^n \quad (3)$$

Pour laquelle on a des données expérimentales de C_A en fonction du temps (dC_A/dt). On aimerait déterminer l'ordre de la réaction. On tire de (3) que

$$\log\left(-\frac{dC_A}{dt}\right) = \log(K) + n \log(C_A)$$

Qu'on peut réécrire sous la forme

$$y = \alpha + nx$$

Établissons la relation fonctionnelle sous la forme $f=a_0+a_1x$. Le critère des moindres s'écrit

$$J(a_0, a_1) = \sum_{i=1}^N |(a_0 + a_1 x_i) - y_i|^2$$

L'application de la condition de minimisation donne

$$\frac{\delta J}{\delta a_0} = 2 \sum_{i=1}^N (a_0 + a_1 x_i - y_i) = 0$$

$$\frac{\delta J}{\delta a_1} = 2 \sum_{i=1}^N x_i (a_0 + a_1 x_i - y_i) = 0$$

Ce qui revient à

$$\sum_{i=1}^N (a_0 + a_1 x_i) = \sum_{i=1}^N y_i$$

$$\sum_{i=1}^N x_i (a_0 + a_1 x_i) = \sum_{i=1}^N x_i y_i$$

Sous forme matricielle nous obtenons

$$\begin{bmatrix} N & \sum_{i=1}^N x_i \\ \text{sym.} & \sum_{i=1}^N x_i^2 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_i y_i \end{Bmatrix} \quad (4)$$

La solution du système est

$$a_0 = \frac{1}{N} \left(\sum_{i=1}^N y_i - a_1 \sum_{i=1}^N x_i \right) \quad a_1 = \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2}$$



déterminant

Exemple

Considérons le tableau de valeurs expérimentales suivant

x	1.0	2.0	3.0	5.0
y	3.0	4.5	2.5	0.5

Question

Trouvez les paramètres de la régression linéaire simple par la méthode des moindres carrés.

Solution

Pour une régression linéaire simple nous avons à déterminer un système symétrique de deux équations à deux inconnues a_0 et a_1 tel que donné par (4). De plus nous avons 4 points expérimentaux. Ce qui revient à écrire

$$\left[\begin{array}{c} 4 \\ \text{sym.} \end{array} \begin{array}{c} \sum_{i=1}^4 x_i \\ \sum_{i=1}^4 x_i^2 \end{array} \right] \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^4 y_i \\ \sum_{i=1}^4 x_i y_i \end{Bmatrix}$$

Application numérique

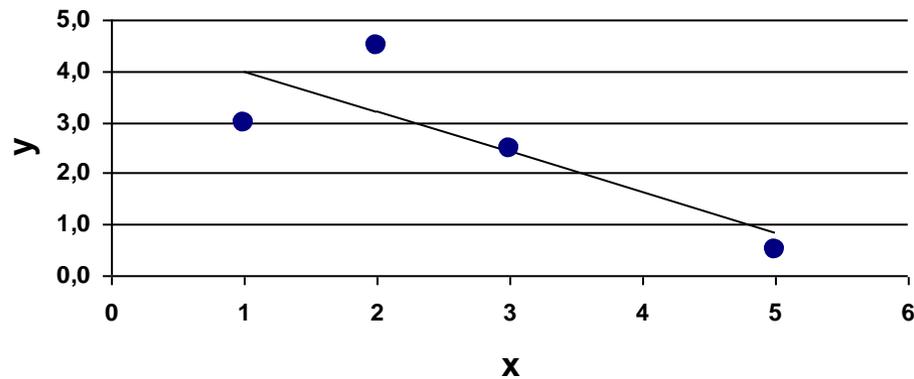
Données	x	y	x*x	x*y
1	1	3	1	3
2	2	4,5	4	9
3	3	2,5	9	7,5
4	5	0,5	25	2,5
Somme	11	10,5	39	22

Déterminant = 35

Solution

Ord. Origine a0= 4,78571429

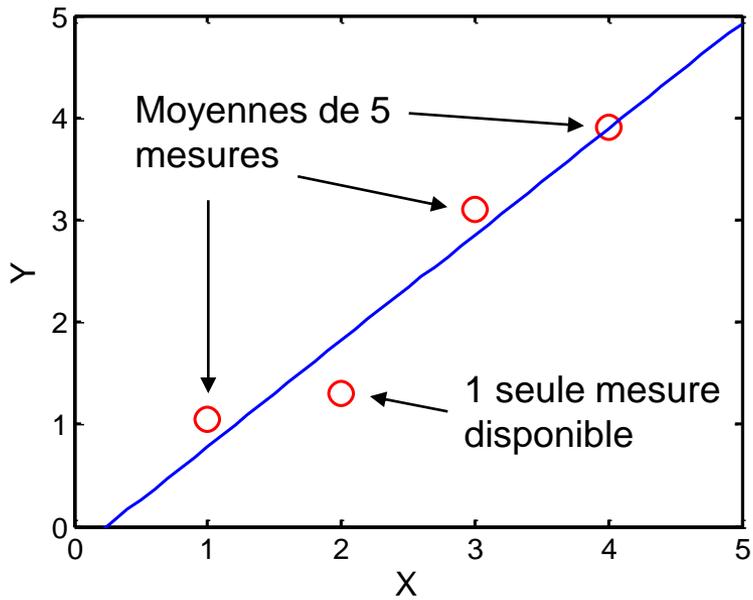
Pente a1= -0,78571429



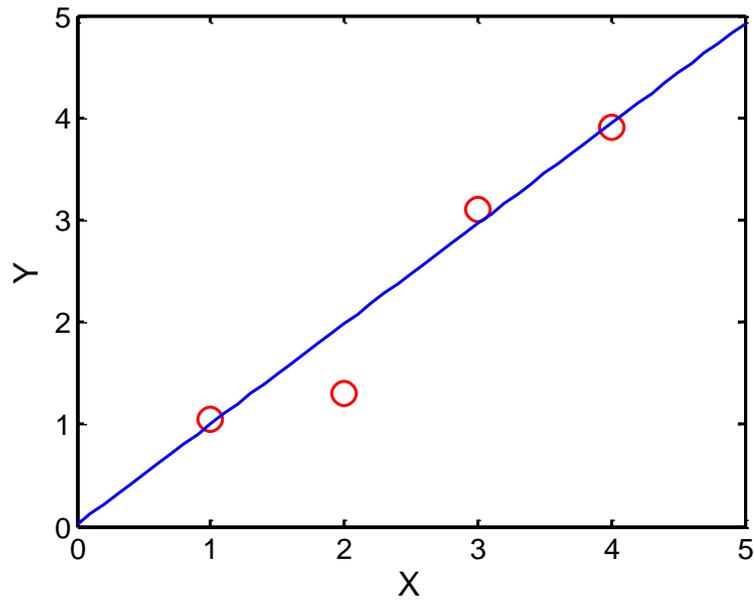
Régression pondérée au sens des moindres carrés

Régression simple au sens des moindres carrés pondérés

Régression simple



Régression pondérée



$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^4 w_i \varepsilon_i^2 = (5 \varepsilon_1^2 + 1 \varepsilon_2^2 + 5 \varepsilon_3^2 + 5 \varepsilon_4^2)$$

On donne moins d'importance aux données moins significatives

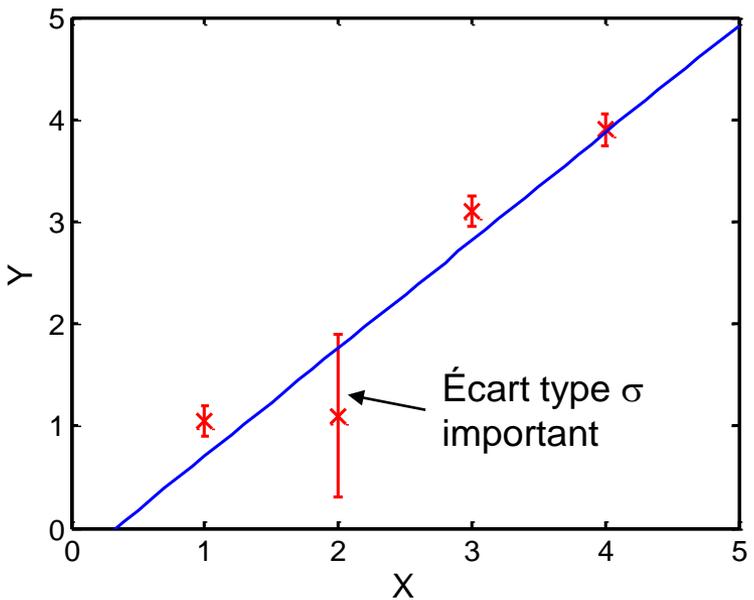
$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^N w_i |f(a_0, a_1, \dots, a_k; x_i) - y_i|^2$$

Ici **poids** = nombre de données

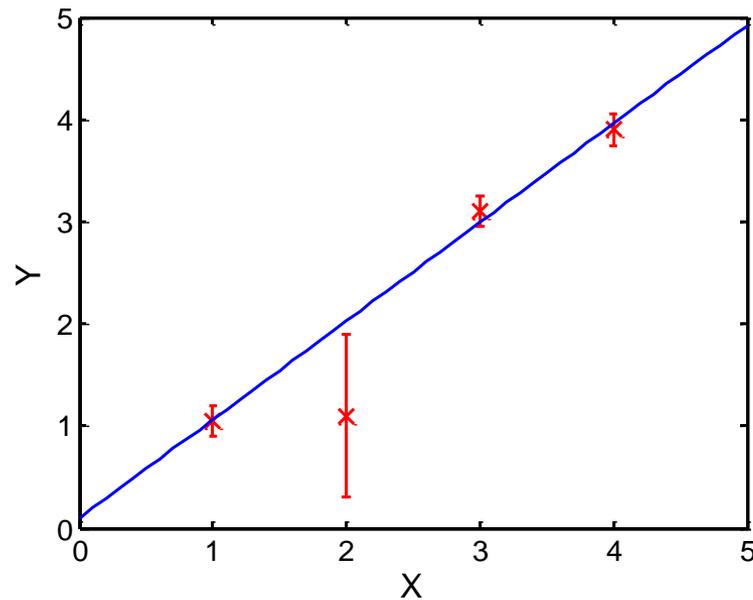
Régression simple au sens des moindres carrés pondérés

Autre exemple

Régression simple



Régression pondérée



$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^4 w_i \varepsilon_i^2 = \left(\frac{1}{\sigma_1^2} \varepsilon_1^2 + \frac{1}{\sigma_2^2} \varepsilon_2^2 + \frac{1}{\sigma_3^2} \varepsilon_3^2 + \frac{1}{\sigma_4^2} \varepsilon_4^2 \right)$$

On donne moins d'importance aux données moins significatives

$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^N w_i |f(a_0, a_1, \dots, a_k; x_i) - y_i|^2$$

Ici **poids** = $\frac{1}{\sigma_i^2}$ = inverse de la variance

Conditionnement d'une matrice

Lors de la résolution d'un système matriciel, l'arithmétique flottante est telle que la solution approchée n'est généralement pas égale à la solution exacte.

Exemple

Soit le système matriciel suivant:

$$\begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{Bmatrix} = \begin{Bmatrix} 32 \\ 23 \\ 33 \\ 31 \end{Bmatrix}$$

Qui admet pour solution exacte

$$x = (1 \quad 1 \quad 1 \quad 1)^T$$

Si on remplace le second membre b par b^* comme suit

$$b^* = (32.06 \quad 22.87 \quad 33.07 \quad 30.89)^T$$

La solution exacte est

$$x^* = (9.19 \quad -12.59 \quad 4.49 \quad -1.09)^T$$

$$\frac{\|b - b^*\|_\infty}{\|b\|_\infty} = 0.3\% \rightarrow \frac{\|x - x^*\|_\infty}{\|x\|_\infty} = 1359\%$$

C'est un signe que la matrice A est mal conditionnée **par rapport à l'erreur expérimentale**. Les erreurs d'arrondis sur les données expérimentales sont désastreuses pour un tel système.

Remarque

$$\|x\|_\infty = \mathit{Max}_i |x_i|$$

Définition du conditionnement d'une matrice

Soit A une matrice non singulière ($\det A \neq 0$). Il existe un nombre $\text{cond}(A) \geq 1$ tel que si $Ax=b$ et $r=Ax^*-b$ où x^* est une approximation de la solution du système, alors

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

$$\frac{\|r\|_{\infty}}{\|b\|_{\infty}} \rightarrow \frac{\|x - x^*\|_{\infty}}{\|x\|_{\infty}} \quad \text{lorsque} \quad \text{cond}(A) \rightarrow 1$$

En pratique, on peut détecter un problème de conditionnement si lorsqu'on modifie légèrement soit A ou b , la solution change de façon significative.

Qualité d'une régression

Pour chaque donnée (x_i, y_i) , $i=1 \dots N$ on a:

$$y_i = f(x_i) + \varepsilon_i$$

On peut mesurer la qualité de la régression de plusieurs façons:

Écart-type S

$$S = \sqrt{\frac{\sum_{i=1}^N \varepsilon_i^2}{\nu}}$$

ν est le nombre de degrés de liberté.

Pour un polynôme de degré k , on a: $\nu = N - (k+1)$.

$$\hat{S} = \frac{S}{\Delta y} \quad \Delta y = \left| \underset{i}{Max} y_i - \underset{i}{Min} y_i \right|$$

Coefficient de corrélation r

$$r = \sqrt{1 - \frac{\sum_{i=1}^N \varepsilon_i^2}{S_y}}$$

$$S_y = \sum_{i=1}^N (y_i - \bar{y})^2$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

Remarques

1. \hat{S} est généralement un meilleur indicateur que S car il tient compte de l'intervalle Δy des valeurs mesurées (Figure 1).
2. S est généralement un meilleur indicateur que r car il est moins affecté par la forme de la courbe de régression (Figure 2).
3. La meilleure façon de mesurer la qualité d'une régression c'est souvent de procéder graphiquement (Figure 3).

Remarques

1. \hat{S} est généralement un meilleur indicateur que S car il tient compte de l'intervalle Δy des valeurs mesurées (Figure 1).

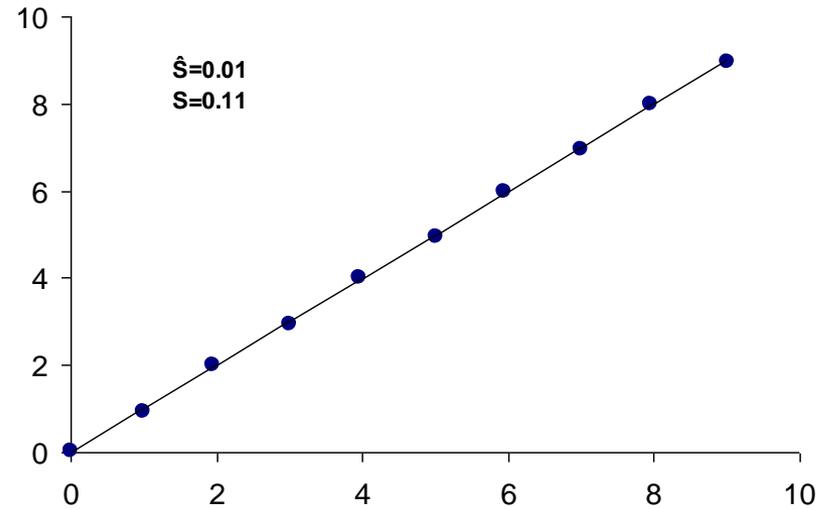
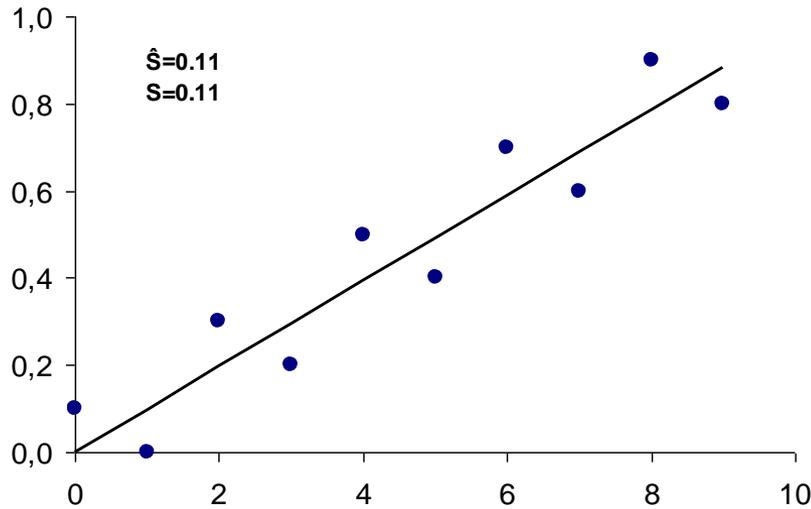


Figure 1. S et \hat{S}

Remarques

2. S est généralement un meilleur indicateur que r car il est moins affecté par la forme de la courbe de régression (Figure 2).

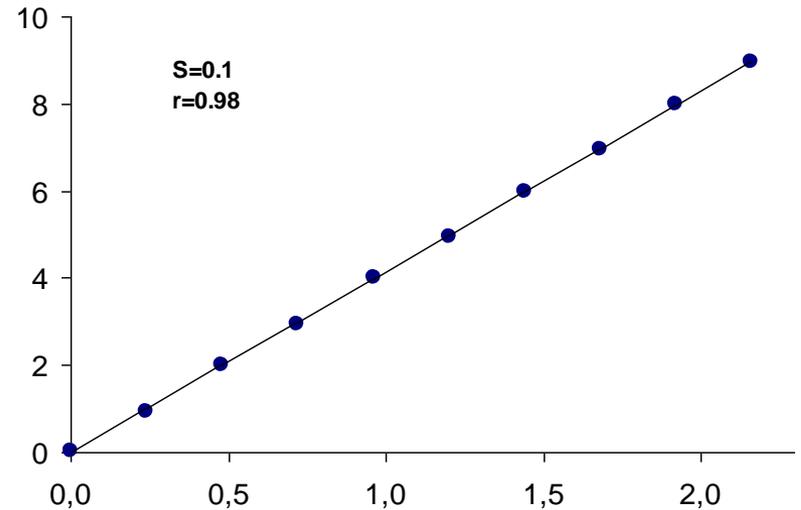
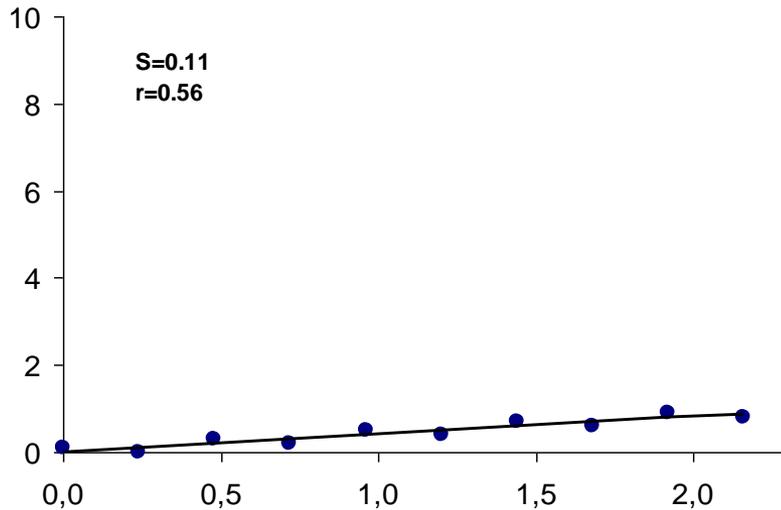


Figure 2. S et r

Remarques

3. La meilleure façon de mesurer la qualité d'une régression c'est souvent de procéder graphiquement (Figure 3). ET de tracer les résidus.

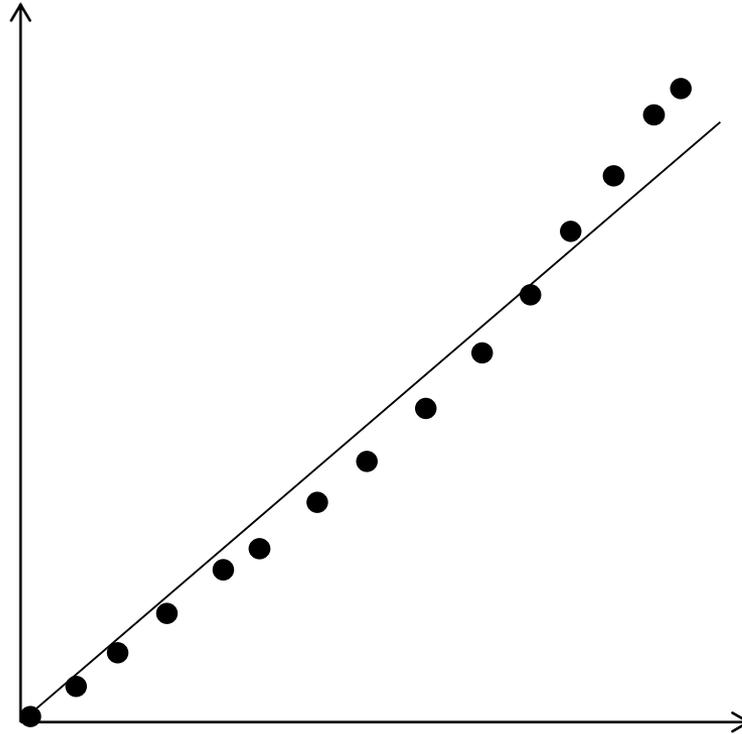


Figure 3. Parabole ou droite?

Exemple:

Considérons le jeu de données suivant:

x	y
0.0	0.0000
0.5	0.9330
1.0	1.0000
2.0	1.0718
3.0	1.1161
4.0	1.1487

Question: déterminer le polynôme $y=f(x)$ par la méthode de régression linéaire dont la matrice est de type Vandermonde.

Réponse

À partir de la matrice de Vandermonde on construit un polynôme qui interpole tous les points expérimentaux. Pour le présent exemple nous avons 6 points expérimentaux, le polynôme à calculer est donc de degré 5 qu'on exprime comme suit:

$$f(a_0, a_1, \dots, a_5) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5$$

Le critère des moindres carrés s'exprime comme suit:

$$J(a_0, a_1, \dots, a_5) = \sum_{i=1}^N [y_i - (a_0 + a_1x_i + a_2x_i^2 + a_3x_i^3 + a_4x_i^4 + a_5x_i^5)]^2$$

La détermination de $a(a_0, a_1, \dots, a_5)$ revient à minimiser le critère J:

$$\frac{\partial J}{\partial a_i} = 0$$

$$\frac{\partial J}{\partial a_0} = -2 \sum_{i=1}^N [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_i^5)] = 0$$

$$\frac{\partial J}{\partial a_1} = -2 \sum_{i=1}^N x_i [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_i^5)] = 0$$

$$\frac{\partial J}{\partial a_2} = -2 \sum_{i=1}^N x_i^2 [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_i^5)] = 0$$

$$\frac{\partial J}{\partial a_3} = -2 \sum_{i=1}^N x_i^3 [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_i^5)] = 0$$

$$\frac{\partial J}{\partial a_4} = -2 \sum_{i=1}^N x_i^4 [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_i^5)] = 0$$

$$\frac{\partial J}{\partial a_5} = -2 \sum_{i=1}^N x_i^5 [y_i - (a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_i^5)] = 0$$

Le système matriciel résultant s'exprime comme suit:

$$\begin{bmatrix}
 N \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^4 & \sum_{i=1}^N x_i^5 \\
 \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^4 & \sum_{i=1}^N x_i^5 & \sum_{i=1}^N x_i^6 \\
 & \sum_{i=1}^N x_i^4 & \sum_{i=1}^N x_i^5 & \sum_{i=1}^N x_i^6 & \sum_{i=1}^N x_i^7 \\
 & & \sum_{i=1}^N x_i^6 & \sum_{i=1}^N x_i^7 & \sum_{i=1}^N x_i^8 \\
 & & & \sum_{i=1}^N x_i^8 & \sum_{i=1}^N x_i^9 \\
 & & & & \sum_{i=1}^N x_i^{10}
 \end{bmatrix}
 \begin{bmatrix}
 a_0 \\
 a_1 \\
 a_2 \\
 a_3 \\
 a_4 \\
 a_5
 \end{bmatrix}
 =
 \begin{bmatrix}
 \sum_{i=1}^N y_i \\
 \sum_{i=1}^N x_i y_i \\
 \sum_{i=1}^N x_i^2 y_i \\
 \sum_{i=1}^N x_i^3 y_i \\
 \sum_{i=1}^N x_i^4 y_i \\
 \sum_{i=1}^N x_i^5 y_i
 \end{bmatrix}$$

sym.

Une autre façon d'exprimer le système matriciel est d'avoir recours à la forme quadratique plus compacte

$$[\mathbf{X}]^T [\mathbf{X}] \{\mathbf{a}\} = [\mathbf{X}]^T \{\mathbf{y}\}$$

où

$$[\mathbf{X}] = \left[\begin{array}{cccccc} \{1\} & \{\mathbf{x}\} & \{\mathbf{x}^2\} & \{\mathbf{x}^3\} & \{\mathbf{x}^4\} & \{\mathbf{x}^5\} \end{array} \right]$$

Mais comme $[\mathbf{X}]$ est une matrice carré de type Vandermonde, on peut réécrire le système d'équation plus simplement:

$$[\mathbf{X}] \{\mathbf{a}\} = \{\mathbf{y}\}$$

La matrice $[\mathbf{X}]$ est mieux conditionnée (ou moins mal) que $[\mathbf{X}]^T [\mathbf{X}]$.

Application numérique

matrice X des données

1. 0. 0. 0. 0. 0.
1. 0.5 0.25 0.125 0.0625 0.03125
1. 1. 1. 1. 1. 1.
1. 2. 4. 8. 16. 32.
1. 3. 9. 27. 81. 243.
1. 4. 16. 64. 256. 1024.

vecteur Y des données

0.
0.933
1.
1.0718
1.1161
1.1487

matrice A du système

1. 0. 0. 0. 0. 0.
1. 0.5 0.25 0.125 0.0625 0.03125
1. 1. 1. 1. 1. 1.
1. 2. 4. 8. 16. 32.
1. 3. 9. 27. 81. 243.
1. 4. 16. 64. 256. 1024.

second membre b du système

0.
0.933
1.
1.0718
1.1161
1.1487

solution a

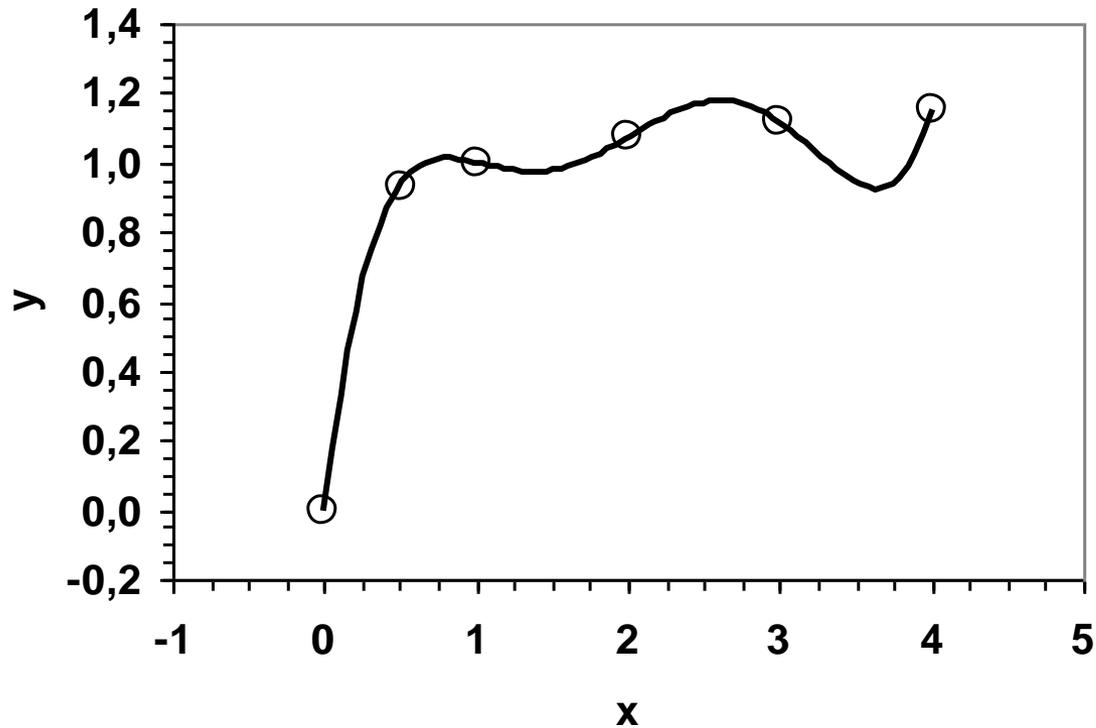
1 0.
2 3.6394269
3 -4.76558869
4 2.78323333
5 -0.725982738
6 0.0689111905

La courbe passe
par l'origine

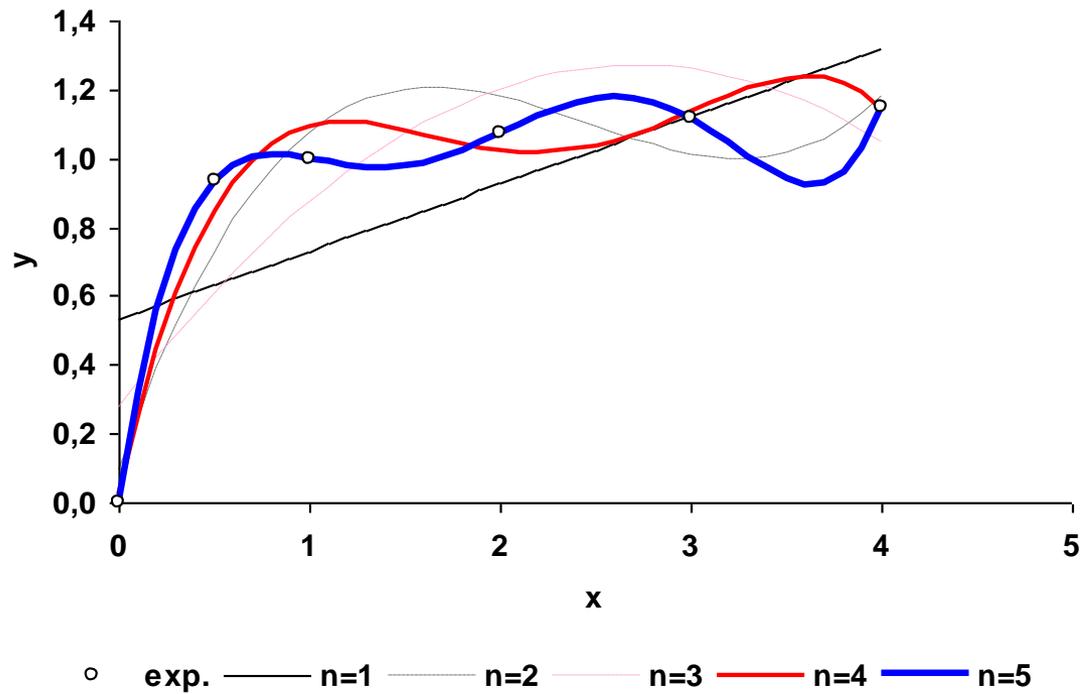
Finalement, nous obtenons

$$f(x) = 3.639x - 4.766x^2 + 2.783x^3 - 0.726x^4 + 0.069x^5$$

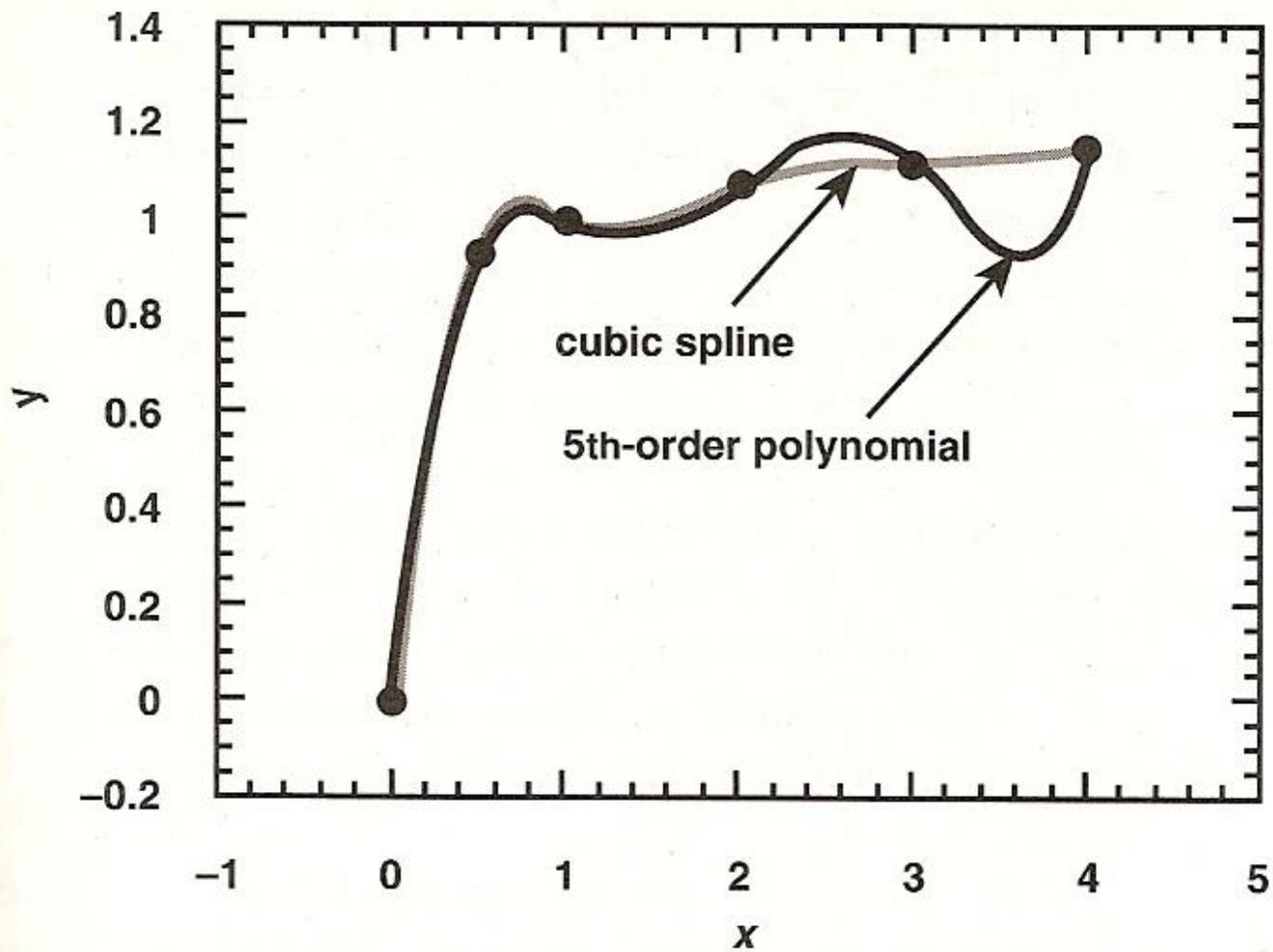
Les données et le polynôme ainsi construit sont représentés sur le graphique ci-dessous:



et met clairement en évidence le caractère oscillant du polynôme.
Question: de quelle façon peut-on lisser la courbe de tendance?



n	$f(x)$
1	$0.5347 + 0.1963x$
2	$0.2777 + 0.7303x - 0.1344x^2$
3	$0.0919 + 1.6225x - 0.7402x^2 + 0.1006x^3$
4	$0.0236 + 2.4817x - 1.9646x^2 + 0.6204x^3 - 0.0667x^4$
5	$3.639x - 4.766x^2 + 2.783x^3 - 0.726x^4 + 0.069x^5$



Régression multivariée au sens des moindres carrés

Régression multivariée au sens des moindres carrés

La régression simple peut être étendue au cas multivarié de façon naturelle.

Considérons une fonction $f(a_0, \dots, a_k; x_1, x_2)$ où les paramètres $a_i, i=0, 1, \dots, k$ sont à déterminer de façon à minimiser le critère des moindres carrés:

$$J(a_0, \dots, a_k) = \sum_{i=1}^N |f(a_0, \dots, a_k; x_{1i}, x_{2i}) - y_i|^2$$

Où (x_{1i}, x_{2i}, y_i) est obtenu expérimentalement. Les paramètres $a_i, i=0, 1, \dots, k$ sont obtenus en résolvant le système de $(k+1)$ équations non linéaires

$$\frac{\delta J}{\delta a_j} = 0, \quad \forall i = 0, 1, \dots, k$$

Cas particulier: la régression linéaire multiple

$$f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2$$

$$J(a_0, a_1, a_2) = \sum_{i=1}^N |y_i - a_0 - a_1 x_{1i} - a_2 x_{2i}|^2$$

Des conditions nécessaires de minimisation

$$\frac{\delta J}{\delta a_0} = -2 \sum_{i=1}^N (y_i - a_0 - a_1 x_{1i} - a_2 x_{2i}) = 0$$

$$\frac{\delta J}{\delta a_1} = -2 \sum_{i=1}^N x_{1i} (y_i - a_0 - a_1 x_{1i} - a_2 x_{2i}) = 0$$

$$\frac{\delta J}{\delta a_2} = -2 \sum_{i=1}^N x_{2i} (y_i - a_0 - a_1 x_{1i} - a_2 x_{2i}) = 0$$

on en déduit que a_0 , a_1 et a_2 sont la solution du système linéaire

$$Na_0 + a_1 \sum_{i=1}^N x_{1i} + a_2 \sum_{i=1}^N x_{2i} = \sum_{i=1}^N y_i$$

$$a_0 \sum_{i=1}^N x_{1i} + a_1 \sum_{i=1}^N x_{1i}^2 + a_2 \sum_{i=1}^N x_{1i}x_{2i} = \sum_{i=1}^N x_{1i}y_i$$

$$a_0 \sum_{i=1}^N x_{2i} + a_1 \sum_{i=1}^N x_{2i}x_{1i} + a_2 \sum_{i=1}^N x_{2i}^2 = \sum_{i=1}^N x_{2i}y_i$$

ou encore sous forme matricielle

$$\left[\begin{array}{ccc}
 N & \sum_{i=1}^N x_{1i} & \sum_{i=1}^N x_{2i} \\
 \sum_{i=1}^N x_{1i} & \sum_{i=1}^N x_{1i}^2 & \sum_{i=1}^N x_{1i}x_{2i} \\
 \sum_{i=1}^N x_{2i} & \sum_{i=1}^N x_{2i}x_{1i} & \sum_{i=1}^N x_{2i}^2
 \end{array} \right] \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^N y_i \\ \sum_{i=1}^N x_{1i}y_i \\ \sum_{i=1}^N x_{2i}y_i \end{Bmatrix} \quad (5)$$

Écriture matricielle sous forme quadratique

Le système d'équation pour une régression linéaire n-multivariée s'écrit

$$[\mathbf{A}]\{\mathbf{a}\} = \{\mathbf{b}\}$$

ou bien encore sous la forme quadratique suivante:

$$[\mathbf{X}]^T [\mathbf{X}]\{\mathbf{a}\} = [\mathbf{X}]^T \{\mathbf{Y}\}$$

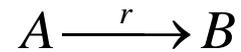
avec

$$[\mathbf{X}] = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{n1} \\ 1 & x_{12} & x_{22} & \dots & x_{n2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1N} & x_{2N} & \dots & x_{nN} \end{bmatrix} \quad \{\mathbf{Y}\} = \begin{Bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_N \end{Bmatrix}$$

un vecteur colonne correspond à une catégorie de données expérimentales

Exemple*

Soit la réaction réversible



et les données expérimentales suivantes

Température (K)	C_A (mol.g/l)	r (mol.g/l.s)
373	1.00	1.508
395	0.923	2.936
365	1.15	1.293
400	0.87	3.242
405	1.05	4.566
388	0.75	1.899
410	0.55	2.780
380	0.65	1.255

On suppose que la vitesse de réaction est donnée par:

$$r = k_0 e^{-(E/RT)} C_A^n \quad (6)$$

où E est l'énergie d'activation et R la constante des gaz parfaits égale à 1.9872cal/g.mol.K.

Question

On veut estimer k_0 , E et n à l'aide d'une régression linéaire multiple.

Marche à suivre pour la résolution

Logarithme de (6) $\Rightarrow \log(r) = \log(k_0) - \frac{E}{RT} + n \log(C_A)$

On pose le changement de variable suivant

$$y = \log(r); \quad x_1 = -\frac{1}{RT}; \quad x_2 = \log(C_A)$$

D'où

$$y = \log(k_0) + Ex_1 + nx_2$$

Application numérique

matrice des donnees

```
1. -0.00134911691  0.  
1. -0.00127397622 -0.0801260445  
1. -0.00137868659  0.139761942  
1. -0.00125805152 -0.139262067  
1. -0.00124252002  0.0487901642  
1. -0.00129696033 -0.287682072  
1. -0.00122736733 -0.597837001  
1. -0.00132426475 -0.430782916
```

vecteur des donnees

```
0.41078427  
1.07704811  
0.2569651  
1.17619042  
1.51863755  
0.641327432  
1.02245093  
0.227135573
```

matrice du systeme

```
8. -0.0103509437 -1.34713799  
-0.0103509437  1.3412672E-005  0.00170131533  
-1.34713799  0.00170131533  0.673471962
```

second membre du systeme

```
6.33053938  
-0.00803473645  
-1.03369332
```

solution

```
1  13.8117775  
2  9933.07765  
3  0.999862862
```

$$x_1 = -\frac{1}{RT}$$

$$x_2 = \log(C_A)$$

$$y = \log(r)$$

$$k_0 = 0.996 \times 10^6$$

$$E = 9933 \text{ cal/mol.g}$$

$$n \approx 1$$

Régression non linéaire au sens des moindres carrés

Régression non linéaire au sens des moindres carrés

Dans le cas où l'équation à ajuster n'est ni linéaire ni polynomiale et qu'il n'est pas possible de la transformer en une équation linéaire, on doit faire appel aux techniques de régression non linéaire.

Exemple

Équation d'Antoine qui permet d'estimer la pression de vapeur P_V en fonction de la température T d'un fluide:

$$P_V = 10^{A+B/(T+C)} \quad P_V [=] \text{mm Hg}$$
$$\log_{10} P_V = A + B/(T + C) \quad (B < 0) \quad T [=] \text{°C}$$

La régression non linéaire au sens des moindres carrés nous conduit à un système de 3 équations non linéaires à 3 inconnues A , B et C . Il existe plusieurs façons de résoudre ce système; par exemple la méthode de Newton. Nous présentons ici une variante appelée la méthode de Gauss-Newton qui est très utilisée en statistiques.

Méthode de Gauss-Newton

Supposons connus N points (x_i, y_i) , $i=1, 2, \dots, N$.

Considérons une fonction $f(a_0, \dots, a_k; x)$ et le critère des moindres carrés:

$$J(a_0, a_1, \dots, a_k) = \sum_{i=1}^N |f(a_0, a_1, \dots, a_k; x_i) - y_i|^2$$

Linéarisons f à l'aide d'un développement limité de Taylor d'ordre 1 autour de

$$(a_0^*, a_1^*, \dots, a_k^*) = \underset{\sim}{a}^*$$

Il vient

$$f(a_0, a_1, \dots, a_k; x) \approx f(a_0^*, a_1^*, \dots, a_k^*; x) + \sum_{j=0}^k \frac{\delta f}{\delta a_j} \bigg|_{\underset{\sim}{a}^*; x} (a_j - a_j^*)$$

En substituant à f son développement limité, on obtient le critère des moindres carrés

$$J_*(a_0, a_1, \dots, a_k; x) = \sum_{i=1}^N \left| f(\tilde{a}^*; x_i) + \left[\sum_{j=0}^k \frac{\delta f}{\delta a_j} (a_j - a_j^*) \right] - y_i \right|^2$$

Dont la minimisation conduit à la résolution d'un système linéaire de rang $(k+1)$

Algorithme de Gauss-Newton

0. Étant donné $\tilde{a}^{(0)} = \tilde{a}^*$ (estimé initial)

1. Pour $n=0, 1, \dots$, jusqu'à convergence

1.1 On cherche $\tilde{a}^{(n)} = (a_0^{(n)}, a_1^{(n)}, \dots, a_k^{(n)})$ en minimisant $J_*(\tilde{a}^{(n)})$

1.2 on pose $\tilde{a}^* = \tilde{a}^{(n)}$

Remarques

- La méthode de Gauss-Newton ne converge pas nécessairement. Le choix de l'estimé initial est important
1. Pour faciliter la convergence on peut relaxer l'étape 1.2:

$$\underset{\sim}{a}^* = \underset{\sim}{a}^{(n-1)} + \omega(\underset{\sim}{a}^{(n)} - \underset{\sim}{a}^{(n-1)})$$
$$\omega \geq 0$$

Une stratégie efficace consiste à choisir $\omega \in [0,1]$ de façon à veiller à ce que J_* diminue à chaque itération.

2. On pourrait remplacer les dérivées partielles par des formules de dérivation:

$$\left. \frac{\delta f}{\delta a_i} \right|_{\underset{\sim}{a}^*; x} \approx \frac{f(\underset{\sim}{a}_0^*, \dots, \underset{\sim}{a}_i^* + h, \dots, \underset{\sim}{a}_k^*; x) - f(\underset{\sim}{a}_0^*, \dots, \underset{\sim}{a}_k^*; x)}{h}$$

ce qui donne un algorithme plus simple mais, en règle générale, moins robuste.

Exemple: application de la méthode de Gauss-Newton pour une fonction à deux paramètres B_1 et B_2 :

$$J(B_1, B_2) = \sum_{i=1}^N |y_i - f(B_1, B_2, x_i)|^2 \quad (I)$$

Appliquons le théorème de Taylor pour linéariser la fonction f par rapport aux paramètres B_1 et B_2 autour de $B^* = \{ B_1^*, B_2^* \}$:

$$f(B_1, B_2, x_i) \approx f(B_1^*, B_2^*, x_i) + \left. \frac{\partial f}{\partial B_1} \right|_{B^*} (B_1 - B_1^*) + \left. \frac{\partial f}{\partial B_2} \right|_{B^*} (B_2 - B_2^*) \quad (II)$$

On injecte (II) dans (I), il vient:

$$J(B_1, B_2) = \sum_{i=1}^N \left| y_i - f(B_1^*, B_2^*, x_i) - \left(\left. \frac{\partial f}{\partial B_1} \right|_{B^*} \right)_i (B_1 - B_1^*) - \left(\left. \frac{\partial f}{\partial B_2} \right|_{B^*} \right)_i (B_2 - B_2^*) \right|^2$$

Le principe de minimisation de J par rapport à B_1 et B_2 s'exprime comme suit:

$$\frac{\partial J}{\partial B_1} = \frac{\partial}{\partial B_1} \left(\sum_{i=1}^N \left| y_i - f(B_1^*, B_2^*, x_i) - \left(\frac{\partial f}{\partial B_1} \Big|_{B^*} \right)_i (B_1 - B_1^*) - \left(\frac{\partial f}{\partial B_2} \Big|_{B^*} \right)_i (B_2 - B_2^*) \right|^2 \right) = 0$$

$$\frac{\partial J}{\partial B_2} = \frac{\partial}{\partial B_2} \left(\sum_{i=1}^N \left| y_i - f(B_1^*, B_2^*, x_i) - \left(\frac{\partial f}{\partial B_1} \Big|_{B^*} \right)_i (B_1 - B_1^*) - \left(\frac{\partial f}{\partial B_2} \Big|_{B^*} \right)_i (B_2 - B_2^*) \right|^2 \right) = 0$$

Ou encore

$$-2 \sum_{i=1}^N \left(y_i - f(B_1^*, B_2^*, x_i) - \left(\frac{\partial f}{\partial B_1} \Big|_{B^*} \right)_i (B_1 - B_1^*) - \left(\frac{\partial f}{\partial B_2} \Big|_{B^*} \right)_i (B_2 - B_2^*) \right) \left(\frac{\partial f}{\partial B_1} \Big|_{B^*} \right)_i = 0$$

$$-2 \sum_{i=1}^N \left(y_i - f(B_1^*, B_2^*, x_i) - \left(\frac{\partial f}{\partial B_1} \Big|_{B^*} \right)_i (B_1 - B_1^*) - \left(\frac{\partial f}{\partial B_2} \Big|_{B^*} \right)_i (B_2 - B_2^*) \right) \left(\frac{\partial f}{\partial B_2} \Big|_{B^*} \right)_i = 0$$

En pratique, on part d'un estimé initial

$$B^* = (B_1, B_2)^*$$

On évalue les dérivées premières en chaque point expérimental

$$\left(\frac{\partial f}{\partial B_1} \Big|_{B^*} \right)_i \quad \left(\frac{\partial f}{\partial B_2} \Big|_{B^*} \right)_i$$

On construit le système d'équations non linéaire de type Gauss-Newton:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{Bmatrix} B_1 \\ B_2 \end{Bmatrix} = \begin{Bmatrix} b_1 \\ b_2 \end{Bmatrix}$$

On résout pour B_1 et B_2 et on calcul la norme de convergence:

$$\|B - B^*\|_{\infty} \leq \varepsilon \text{ ou } \frac{\|B - B^*\|_{\infty}}{\|B^*\|_{\infty}} \leq \varepsilon$$

Si le critère de convergence n'est pas satisfait, mettre à jour $B^*=B$.

L'estimé initial peut être déterminant dans le processus de convergence de la méthode.

Régression non linéaire

$$\begin{bmatrix} \sum_{i=1}^N \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_i^2 & \sum_{i=1}^N \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_i \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_i \\ \text{sym.} & \sum_{i=1}^N \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_i^2 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} b_0 \\ b_1 \end{Bmatrix}$$

$$\begin{Bmatrix} b_0 \\ b_1 \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^N (y_i - f_i^*) \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_i + a_0^* \sum_{i=1}^N \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_i^2 + a_1^* \sum_{i=1}^N \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_i \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_i \\ \sum_{i=1}^N (y_i - f_i^*) \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_i + a_0^* \sum_{i=1}^N \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_i \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_i + a_1^* \sum_{i=1}^N \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_i^2 \end{Bmatrix}$$

Organisation matricielle – forme quadratique

$$[\mathbf{X}(\mathbf{a}^*)]^T [\mathbf{X}(\mathbf{a}^*)] \{\mathbf{a}\} = [\mathbf{X}(\mathbf{a}^*)]^T \{\mathbf{c}(\mathbf{a}^*)\}$$

$$[\mathbf{X}(\mathbf{a}^*)] = \left[\left\{ \frac{\partial \mathbf{f}}{\partial \mathbf{a}_0} \Big|_{\mathbf{a}^*} \right\} \quad \left\{ \frac{\partial \mathbf{f}}{\partial \mathbf{a}_1} \Big|_{\mathbf{a}^*} \right\} \right] = \left[\begin{array}{c} \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_1 \\ \vdots \\ \left(\frac{\partial f}{\partial a_0} \Big|_{a^*} \right)_N \end{array} \quad \begin{array}{c} \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_1 \\ \vdots \\ \left(\frac{\partial f}{\partial a_1} \Big|_{a^*} \right)_N \end{array} \right]$$

$$\{\mathbf{c}(\mathbf{a}^*)\} = \{\mathbf{y}\} - \{\mathbf{f}(\mathbf{a}^*)\} + a_0^* \left\{ \frac{\partial \mathbf{f}}{\partial \mathbf{a}_0} \Big|_{\mathbf{a}^*} \right\} + a_1^* \left\{ \frac{\partial \mathbf{f}}{\partial \mathbf{a}_1} \Big|_{\mathbf{a}^*} \right\}$$

Exemple*

Considérons les données expérimentales obtenues avec une solution de 2% de polyisobutylène dans du primol 355.

Taux de cisaillement (s⁻¹)	Viscosité (Pa.s)
0.0137	3220
0.0274	2190
0.0434	1640
0.0866	1050
0.137	766
0.274	490
0.434	348
0.866	223
1.370	163
2.74	104
4.34	76.7
5.46	68.1
6.88	58.2

Exemple*

On considère que la viscosité est de type loi de puissance

$$\eta = K\dot{\gamma}^{\beta-1}$$

où

K : indice de consistance

$\dot{\gamma}$: taux de cisaillement

β : indice de pseudo-plasticité

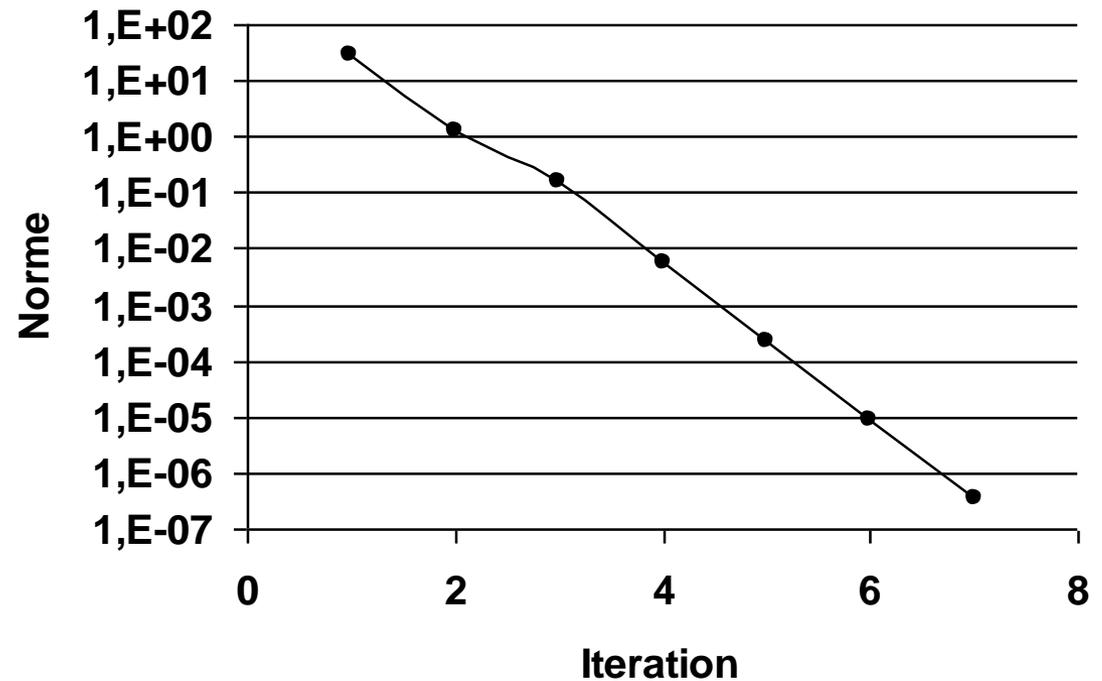
$$\left\{ \begin{array}{l} \textit{rhéofluidifiant} \beta < 1 \\ \textit{Newtonien} \beta = 1 \\ \textit{rhéoépaississant} 1 < \beta \end{array} \right\}$$

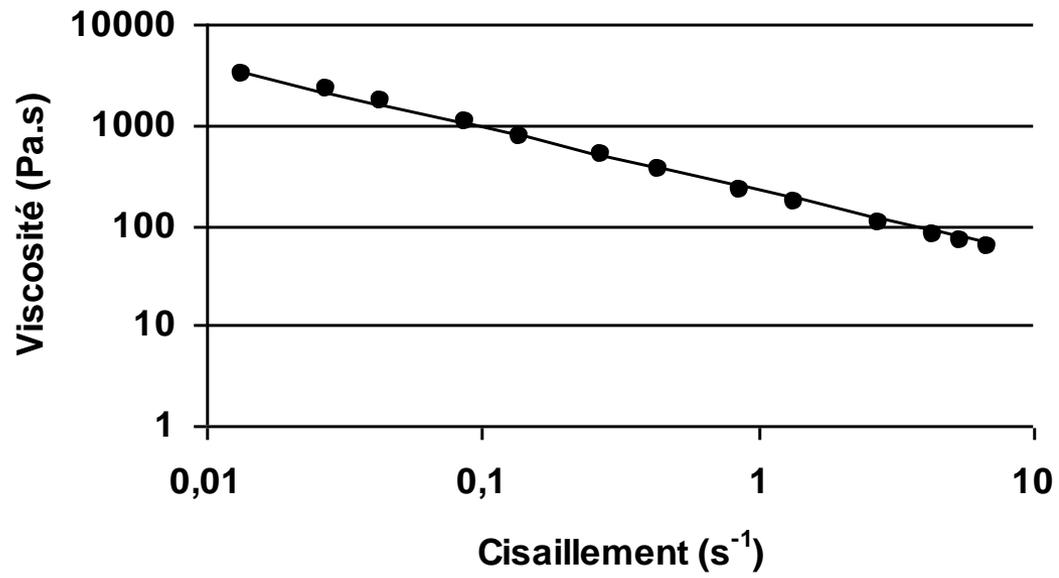
Question: déterminer le couple rhéologique $\mathbf{a}=(K, \beta)$ par une méthode de régression non linéaire de Gauss-Newton en considérant la solution initiale $\mathbf{a}^* = (K, \beta)^* = (200. , 0.4)$.

Solution fluide loi de puissance

Convergence

Solution initiale	
K	200
β	0,4
Solution à itération	7
K	228,324338
β	0,37971997
Précision	1,00E-06





Rappel sur le calage des données par interpolation

- En génie chimique, il est fréquent d'avoir à ajuster sur des données une loi empirique ou encore d'avoir à déterminer la valeur des paramètres d'une loi théorique de façon à reproduire les résultats d'une expérience.
- Interpolation à l'aide du polynôme de Lagrange
- Interpolation à l'aide de splines

Approximation à l'aide du polynôme d'interpolation

Une autre façon de faire passer une courbe par des points (données expérimentales) consiste à utiliser le polynôme d'interpolation.

Soient $(n+1)$ points (x_i, y_i) , $i=0, \dots, n$. On cherche un polynôme de degré n tel que $P_n(x_i) = y_i$, $i=0, \dots, n$. Une solution consiste à utiliser le polynôme de Lagrange:

$$P_n(x) = \sum_{i=0}^n y_i L_i(x)$$

où les coefficients L_i , appelés coefficient d'interpolation de Lagrange sont donnés par

$$L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{(x - x_j)}{(x_i - x_j)}$$

et vérifient

$$L_i(x_j) = \delta_{ij}$$

Exemple

Le tableau suivant contient une partie d'une table donnant l'enthalpie de la vapeur saturée en fonction de la température

Température (°F)	Enthalpie de la vapeur saturée (BTU/lb)
220	1153.4
240	1160.6
260	1167.4
280	1173.8

On nous demande:

- d'évaluer l'enthalpie de vapeur saturée à $T=252$ °F à l'aide d'un polynôme d'interpolation quadratique;
- de comparer le résultat de (a) avec la méthode de régression linéaire d'ordre 2.

a) Pour ce faire, on utilise les données à 240 °F, 260 °F et 280 °F. On a alors

$$P_2(x) = \sum_{i=0}^2 y_i L_i(x)$$

$$= y_0 L_0(x) + y_1 L_1(x) + y_2 L_2(x)$$

$$= 1160.6 L_0(x) + 1167.4 L_1(x) + 1173.8 L_2(x)$$

où

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 260)(x - 280)}{(240 - 260)(240 - 280)} =$$

$$L_0(x) = 1.25 \times 10^{-3} x^2 - 0.675x + 91$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 240)(x - 280)}{(260 - 240)(260 - 280)} =$$

$$L_1(x) = -2.5 \times 10^{-3} x^2 + 1.3x - 168$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 240)(x - 260)}{(280 - 240)(280 - 260)} =$$

$$L_2(x) = 1.25 \times 10^{-3} x^2 - 0.625x + 78$$

d'où

$$\begin{aligned} P_2(x) = & \\ & 1160.6[1.25 \times 10^{-3} x^2 - 0.675x + 91] + \\ & 1167.4[-2.5 \times 10^{-3} x^2 + 1.3x - 168] + \\ & 1173.8[1.25 \times 10^{-3} x^2 - 0.625x + 78] \end{aligned}$$

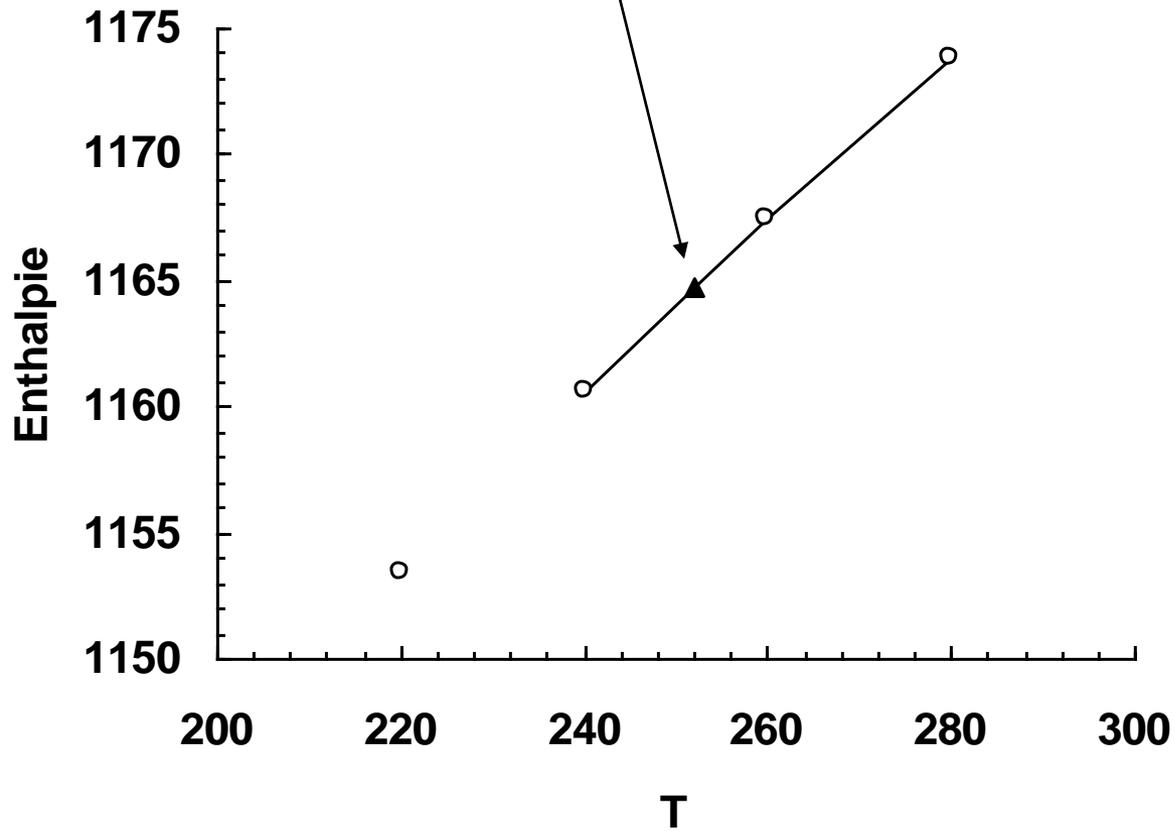
soit après simplification

$$P_2(x) = -5.0 \times 10^{-4} x^2 + 0.59x + 1047.8$$

L'enthalpie au point d'interpolation d'intérêt est donnée par

$$P_2(x = 252) = -5.0 \times 10^{-4} \times 252^2 + 0.59 \times 252 + 1047.8$$

Finalemment, $P_2(T=252 \text{ }^\circ\text{F}) = 1164.7 \text{ BTU/lb}$



Danger du polynôme d'interpolation

Il n'est pas recommandé d'utiliser des polynômes d'interpolation de degré élevé ($n > 3$) en raison de leur comportement oscillant (formation de plis voir figure 1).

Erreur d'interpolation

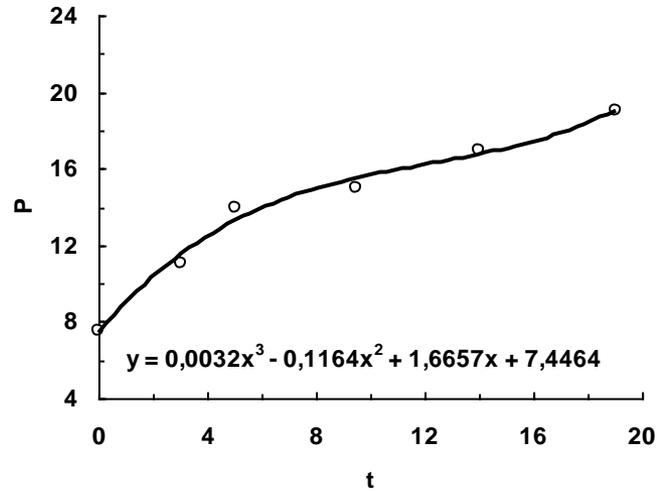
L'erreur commise en utilisant un polynôme d'interpolation est donnée par le résultat suivant: on suppose que la fonction $f(x)$ à approcher est $(n+1)$ fois dérivable. Alors on peut montrer que si les points de collocation $x_0 < x_1 < \dots < x_n$, il existe un point $\xi \in]x_0, x_n[$ tel que

$$E_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x - x_i)$$

avec

$$P_n(x) = \sum_{i=0}^n f(x_i) \cdot L(x_i)$$

n=3



n=5

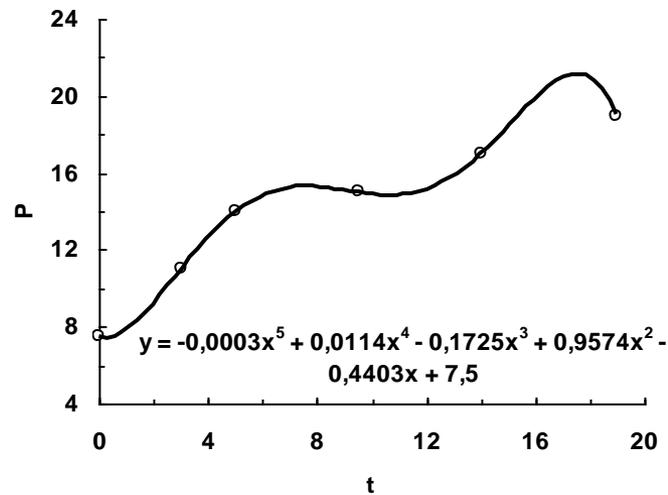
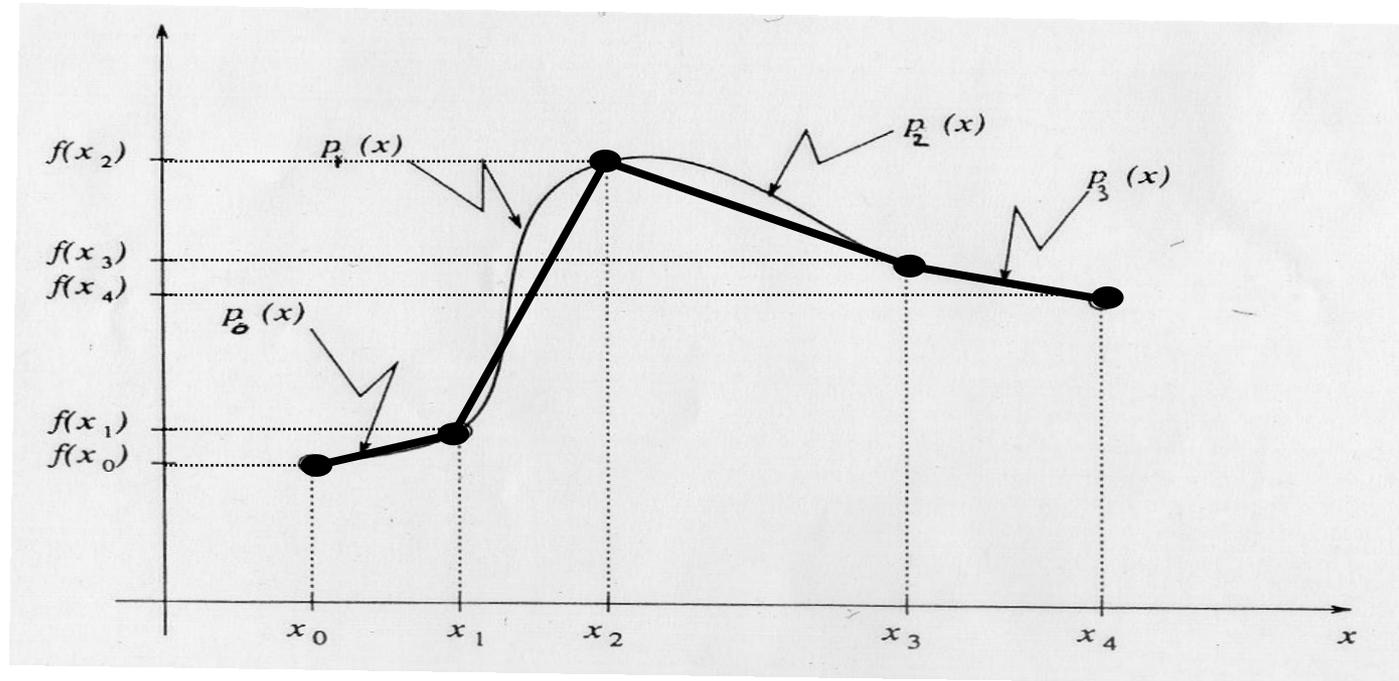


Figure 1: Inconvénient d'un polynôme d'ordre élevé.

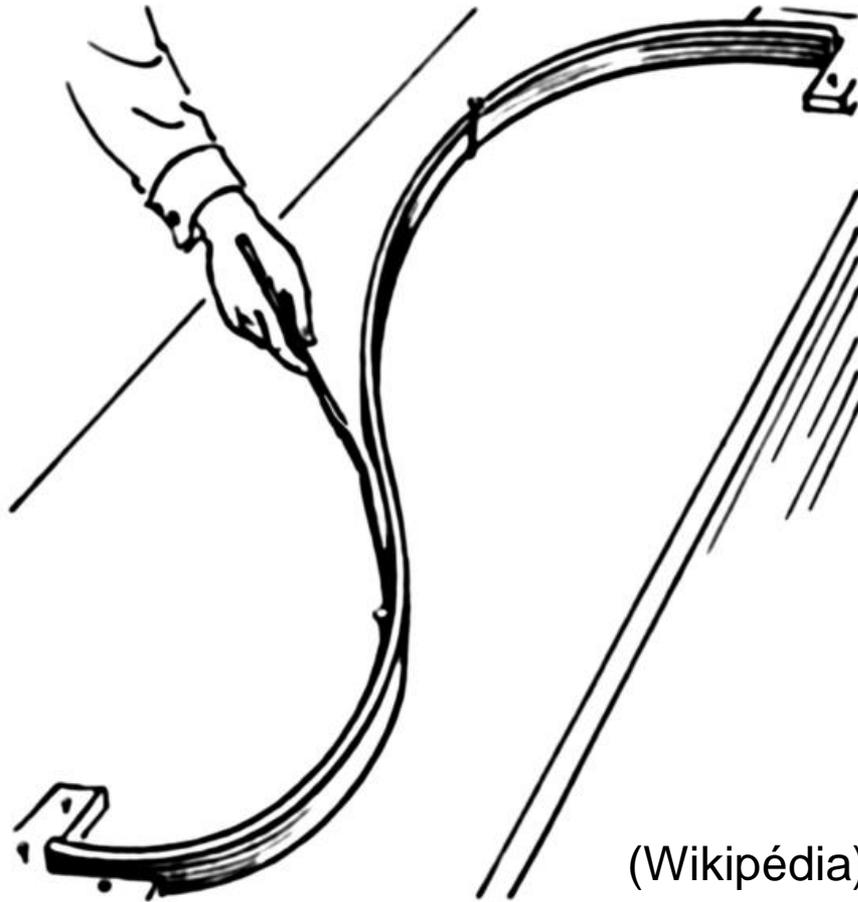
Splines

Idée de base: Utiliser l'interpolation par morceaux et approcher la fonction ou les données expérimentales par un ensemble de polynômes tout en contrôlant la régularité aux nœuds en forçant l'égalité des dérivées 1^e et 2^e aux mêmes points.



NB: Une spline — mot anglais se prononçant « splaïne », mais souvent prononcé à la française. Le terme français correspondant est une **cerce**, mais l'usage a consacré le terme anglais.

Cerce



Cerce: latte souple en bois utilisée par les menuisiers et charpentiers et servant à tracer les courbes *quelconques* (Courbes harmonieuses passant par un certain nombre de points mais ne pouvant être tracées à l'aide du compas).

Construction de la spline

- Soit une fonction F (ou un ensemble de points expérimentaux) définie sur un intervalle $[a, b]$ et un ensemble de nœuds
 - $a = x_0 < x_1 < \dots < x_n = b$
- On appelle spline cubique, interpolant F aux nœuds, une fonction $p(x)$ vérifiant:

- $p(x)|_{[x_i, x_{i+1}]} = p_i(x), \quad \forall i = 0, 1, \dots, n-1$

- $p(x_i) = F(x_i), \quad \forall i = 0, 1, \dots, n$

- $p_{i+1}(x_{i+1}) = p_i(x_{i+1}), \quad \forall i = 0, 1, \dots, n-2$

- $p'_{i+1}(x_{i+1}) = p'_i(x_{i+1}), \quad \forall i = 0, 1, \dots, n-2$

- $p''_{i+1}(x_{i+1}) = p''_i(x_{i+1}), \quad \forall i = 0, 1, \dots, n-2$

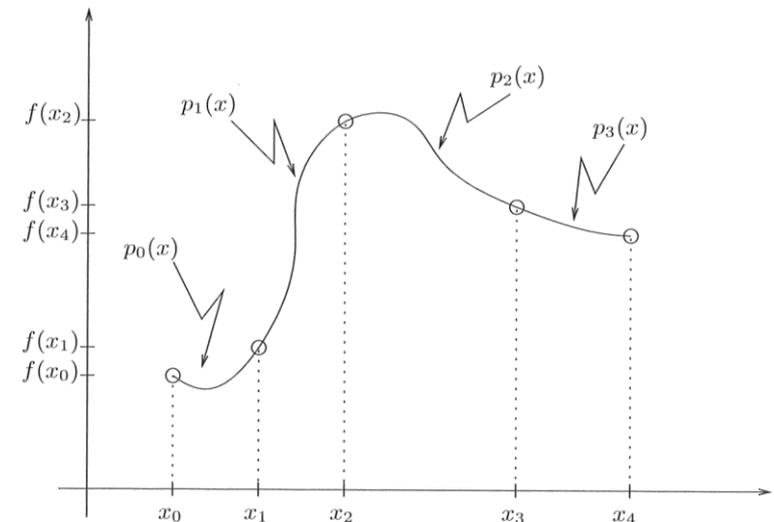


Figure 5.7 – Splines cubiques : ($n = 4$) polynômes de degré 3

- Spline libre:

- $p''(x_0) = p''_0(x_0) = 0$ et $p''(x_n) = p''_{n-1}(x_n) = 0$

- Spline forcée

- $p'(x_0) = p'_0(x_0) = F'(x_0)$ et $p'(x_n) = p'_{n-1}(x_n) = F'(x_n)$

On propose d'utiliser une spline cubique qui utilise dans chaque intervalle (h_i) un polynôme de la forme:

$$y = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3$$

$$y' = b_i + 2c_i(x - x_i) + 3d_i(x - x_i)^2$$

$$y'' = 2c_i + 6d_i(x - x_i)$$

Notes:

- La spline forcée donne de meilleurs résultats car elle tient compte de la géométrie de la fonction F;
- Pour obtenir les a_i , b_i , c_i et d_i , on doit résoudre un système de $4n$ inconnues;
- En pratique, on détermine les coefficients par la méthode suivante:

1) On pose:

$$h_i = x_{i+1} - x_i; \quad \forall i = 0, \dots, n-1$$

$$\alpha_{i+1} = \frac{3}{h_{i+1}}(a_{i+2} - a_{i+1}) - \frac{3}{h_i}(a_{i+1} - a_i); \quad \forall i = 0, \dots, n-2$$

Avec: $a_i = p(x_i) = F(x_i)$, $i = 0, \dots, n-1$

1) On pose

$$- h_i = x_{i+1} - x_i; \quad \forall i = 0, \dots, n-1$$

$$- \alpha_{i+1} = \frac{3}{h_{i+1}}(a_{i+2} - a_{i+1}) - \frac{3}{h_i}(a_{i+1} - a_i); \quad \forall i = 0, \dots, n-2$$

Avec: $a_i = p(x_i) = F(x_i)$, $i = 0, \dots, n-1$

2) On pose aussi $a_n = F(x_n)$, $b_n = p'(x_n)$ et $c_n = p''(x_n)/2$

3) On peut alors montrer que les c_i sont obtenus en résolvant (libre, forcée)

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & 0 & 0 & 0 \\ 0 & h_1 & 2(h_1 + h_2) & h_2 & 0 & 0 \\ 0 & 0 & \dots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ \vdots \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2h_0 & h_0 & 0 & 0 & 0 & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & 0 & 0 & 0 \\ 0 & h_1 & 2(h_1 + h_2) & h_2 & 0 & 0 \\ 0 & 0 & \dots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ 0 & 0 & 0 & 0 & h_{n-1} & 2h_{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ \vdots \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 3/h_0(a_1 - a_0) - 3F'(x_0) \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ 3F'(x_n) - 3/(h_{n-1}(a_n - a_{n-1})) \end{bmatrix}$$

4) On peut alors calculer les coefficients b_i et d_i :

$$- \quad b_i = \frac{1}{h_i}(a_{i+1} - a_i) - \frac{h_i}{3}(2c_i + c_{i+1})$$

$$- \quad d_i = \frac{1}{3h_i}(c_{i+1} - c_i)$$

$$i = 0, \dots, n-1$$

$$\mathbf{L} = \begin{bmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & 0 \\ & l_{32} & l_{33} & & \\ & & \ddots & \ddots & \\ 0 & & & l_{n,n-1} & l_{n,n} \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} 1 & u_{12} & & & \\ & 1 & u_{23} & & 0 \\ & & \ddots & \ddots & \\ & & & 1 & u_{n-1,n} \\ 0 & & & & 1 \end{bmatrix}$$

On peut aussi pour résoudre un système $Ax=b$, où A est une matrice tri-diagonale, combiner la factorisation à la détermination de x . On obtient ainsi l'algorithme de Thomas (Tri Diagonal Matrix Algorithm ou TDMA) que l'on peut établir très facilement.

On réécrit tout d'abord le système matriciel sous la forme

$$\begin{array}{rcl} & d_1 x_1 + e_1 x_2 & = b_1 \\ c_2 x_1 + d_2 x_2 + e_2 x_3 & = b_2 \\ & \vdots & \vdots \\ c_n x_{n-1} + d_n x_n & = b_n \end{array}$$

Algorithme de Thomas

Stratégie: on utilise la $i^{\text{ème}}$ équation pour éliminer la $i^{\text{ème}}$ inconnue.

0. $\beta_1 = d_1$ et $v_1 = b_1/\beta_1 (=b_1/d_1)$

1. Pour $i=2,3,\dots,n$

1.1 $\beta_i = d_i - (c_i e_{i-1}/\beta_{i-1})$

1.2 $v_i = (b_i - c_i v_{i-1})/\beta_i$

2. $x_n = v_n$

3. Pour $i=n-1,n-2,\dots,1$

3.1 $x_i = v_i - (e_i x_{i+1}/\beta_i)$

Remarques

1. Algorithme de Thomas : CPU $\sim O(n)$

2. Factorisation directe : CPU $\sim O(n^3)$

+ remontée triangulaire: CPU $\sim O(n^2)$

Programme TDMA

```
subroutine tdma(a,b,c,d,x,m)
implicit real*8 (a-h,o-z)

dimension a(m),b(m),c(m),d(m),x(m)
c   a : diagonale inférieure
c   b : diagonale principale
c   c : diagonale supérieure
c   d : second membre
c
c---- factorisation
      do i=2,m
                alpha=-b(i)/a(i-1)
                d(i)=d(i)+alpha*d(i-1)
                a(i)=a(i)+alpha*c(i-1)
      enddo

C
C---- résolution par remontée
      x(m)=d(m)/a(m)

      do i = m-1,1,-1
                x(i)=(d(i)-c(i)*x(i+1))/a(i)
      enddo

      return
end
```

Exemple

Équilibre liquide/vapeur d'un mélange eau/éthanol

Fraction molaire éthanol dans la phase liquide	Fraction molaire éthanol dans la vapeur
0.0000	0.0000
0.0190	0.1700
0.0721	0.3891
0.0966	0.4375

Question: Quelle est la fraction molaire de l'éthanol dans la phase liquide qui correspond à une fraction molaire de l'éthanol dans la vapeur de 33%?

Construction d'une spline cubique libre.

Taille des 3 intervalles:

$$h_0 = (0.17 - 0.00) = 0.17$$

$$h_1 = (0.3891 - 0.17) = 0.2191$$

$$h_2 = (0.4375 - 0.3891) = 0.0484$$

Coefficients a_i :

$$a_0 = y_0 = 0.00$$

$$a_1 = y_1 = 0.019$$

$$a_2 = y_2 = 0.0721$$

$$a_3 = y_3 = 0.0966$$

On a aussi

$$2(h_0 + h_1) = 0.7782$$

$$2(h_1 + h_2) = 0.535$$

Les coefficients c_i ($i=1, \dots, n=3$) sont obtenus par la résolution du système linéaire suivant (voir expression générale de la matrice des coefficients c_i):

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & 0 \\ 0 & h_1 & 2(h_1 + h_2) & h_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{Bmatrix} = \begin{Bmatrix} 0 \\ \alpha_1 \\ \alpha_2 \\ 0 \end{Bmatrix}$$

Comme

$$\alpha_{i+1} = \frac{3}{h_{i+1}}(a_{i+2} - a_{i+1}) - \frac{3}{h_i}(a_{i+1} - a_i), i = 0, 1, \dots, n-2$$

alors:

$$\alpha_1 = \frac{3}{h_1}(a_2 - a_1) - \frac{3}{h_0}(a_1 - a_0) = 0.3918$$

$$\alpha_2 = \frac{3}{h_2}(a_3 - a_2) - \frac{3}{h_1}(a_2 - a_1) = 0.7915$$

Après simplification (comment?), on obtient alors les 2 équations

$$0.7782.c_1 + 0.2191.c_2 = 0.3918$$

$$0.2191.c_1 + 0.535.c_2 = 0.7915$$

admettant pour solution

$$c_1 = 0.0983$$

$$c_2 = 1.4393$$

Question: quelle méthode de résolution peut-on utiliser?

Par ailleurs nous savons que les coefficients b_i et d_i sont déterminés par les relations de récurrence suivantes:

$$b_i = \frac{1}{h_i} (a_{i+1} - a_i) - \frac{h_i}{3} (2c_i + c_{i+1}), i = 0, 1, 2, \dots, n-1$$

$$d_i = \frac{1}{3h_i} (c_{i+1} - c_i), i = 0, 1, 2, \dots, n-1$$

On en tire que pour $i=1$, nous avons:

$$b_1 = \frac{1}{h_1} (a_2 - a_1) - \frac{h_1}{3} (2c_1 + c_2) = 0.12287$$

$$d_1 = \frac{1}{3h_1} (c_2 - c_1) = 2.0402$$

Comme $x=0.33 \in [0.17, 0.3891]$ on en déduit que $P(0.33)=P_1(0.33)$.

Or

$$P_1(x) = a_1 + b_1(x - x_1) + c_1(x - x_1)^2 + d_1(x - x_1)^3$$

D'où:

$$P_1(x = 0.33) = 0.019 + 0.12287(0.33 - 0.17) + 0.0983(0.33 - 0.17)^2 + 2.0402(0.33 - 0.17)^3$$

$$P_1(x = 0.33) = 0.05047$$

soit à peu près 5%.

Splines: avantages

1. permettent de représenter des fonctions non lisses ou variant fortement
2. exploitables avec un nombre important de points expérimentaux
3. permettent de dériver une fonction
4. permettent d'intégrer une fonction...

...et inconvénients

1. plusieurs fonctions d'approximation sur un même intervalle
2. problème de fermeture aux bords
3. extrapolation hasardeuse

Remarques finales:

1. Il existe d'autres types:
 - Quadratiques
 - B-Splines, ne passant pas nécessairement par les points. Les points servent à contrôler les formes de la courbes (variantes: courbes de Bézier)
 - NURBS: B-Splines rationnelles qui sont utilisées pour représenter les surfaces dans les logiciels de CAO

Ce problème peut-il être linéarisé ?

$$\eta = K \dot{\gamma}^{\beta-1}$$



$$\log(\eta) = \log(K \dot{\gamma}^{\beta-1})$$



$$\log(\eta) = (\beta - 1) \log(\dot{\gamma}) + \log(K)$$



$$Y = a X + b$$

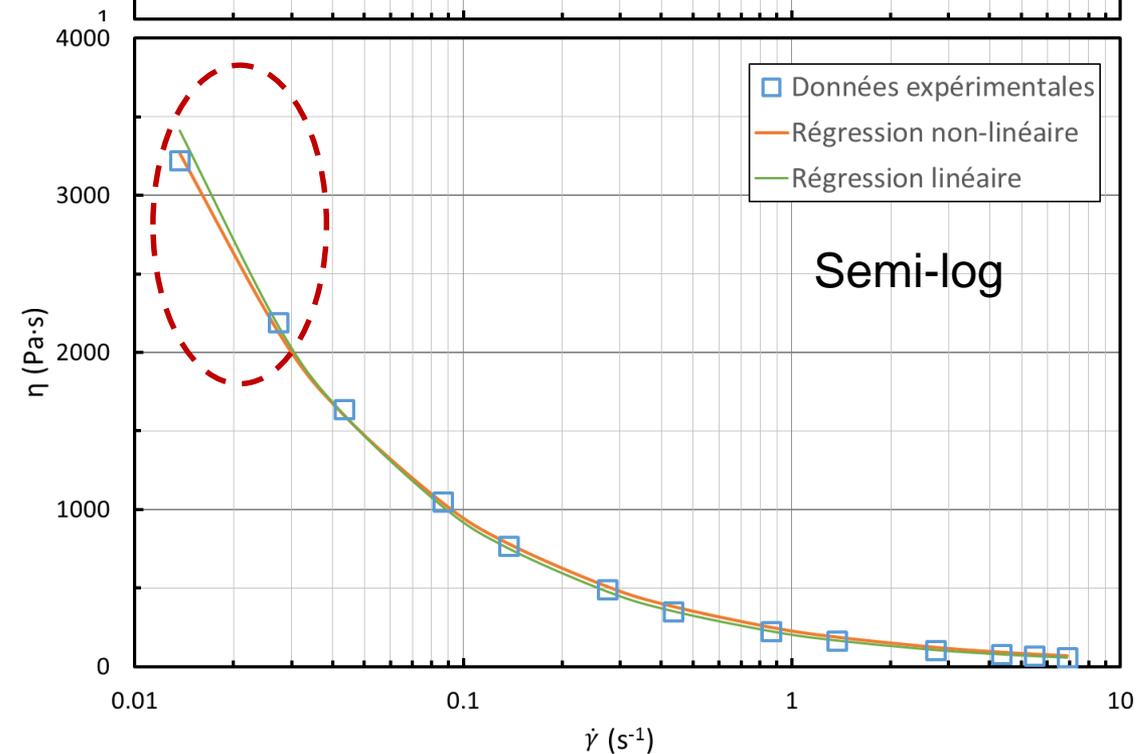
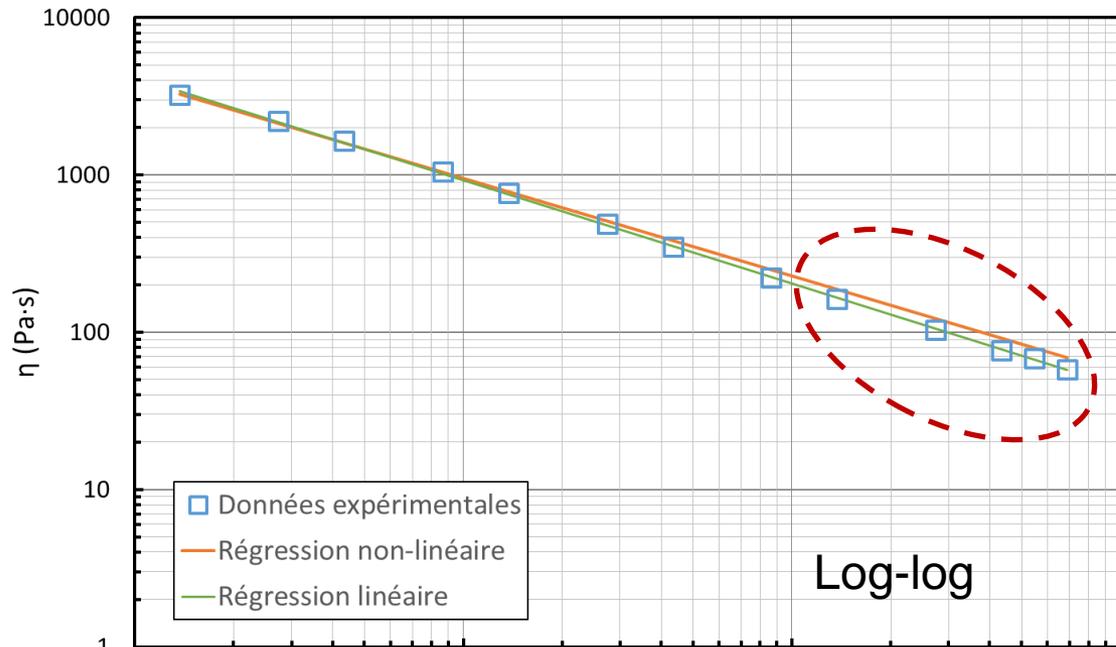


$$\begin{Bmatrix} b \\ a \end{Bmatrix} = \begin{bmatrix} 13 & -11.316 \\ -11.316 & 63.390 \end{bmatrix}^{-1} \times \begin{Bmatrix} 76.586 \\ -101.790 \end{Bmatrix} = \begin{Bmatrix} 5.320 \\ -0.656 \end{Bmatrix}$$



$$K = e^{5.320} = 204.4 \text{ Pa} \cdot \text{s}^{\beta}$$
$$\beta = 1 - 0.656 = 0.344$$

≠ régression non-linéaire



La régression linéaire minimise:

$$\min \sum_{i=1}^N |\log(f(\dot{\gamma}_i)) - \log(\eta_i)|^2$$

Alors que la régression non-linéaire minimise:

$$\min \sum_{i=1}^N |f(\dot{\gamma}_i) - \eta_i|^2$$

Si les valeurs de η varient sur plusieurs ordres de grandeur, alors la procédure peut engendrer des différences notables mais faibles