

Chapitre 3.

Principe d'ajustement du modèle

III.1 Introduction :

Soit un ensemble de mesures expérimentales ($y_0, \dots, y_n$), on aimerait trouver un modèle ajustable (une loi affine "linéaire") qui doit suivre (ou reproduire) ces mesures.

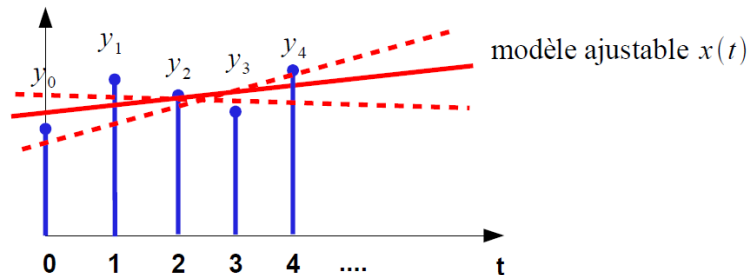


Figure III.1 Principe d'ajustement du modèle

Les différentes droites $x(t)$ (en rouge) sur la figure précédente représentent différents modèles (lois affines ajustables en ordonnée et en pente), elles sont ajustées au mieux pour suivre les mesures (points bleu). On constate que la "qualité" de l'ajustement diffère d'une droite à l'autre, ici le critère d'ajustement n'est pas clairement exprimé.

III.2 Critère d'évaluation de l'erreur de modélisation : "critère des moindres carrés"

Il existe plusieurs critères d'évaluation du modèle ajusté, parmi lesquels, le critère de l'erreur quadratique cumulée (ou le critère des moindres carrés).

Soient $y(t_i)$ les observations faites sur le système et $x(t_i, a_1, a_2, \dots)$ les valeurs prises par le modèles aux instants d'observation. Posons ε_i l'écart (ou l'erreur) entre les mesures et le modèle, alors on peut écrire :

$$y(t_i) = x(t_i, a_1, a_2, \dots) + \varepsilon_i$$

L'erreur quadratique cumulée est alors exprimée par :

$$\begin{aligned} \mathcal{E}(a_1, a_2, \dots) &= \sum_{i=1}^{n_2} \varepsilon_i^2 \\ \mathcal{E}(a_1, a_2, \dots) &= \sum_{i=1}^{n_2} (y(t_i) - x(t_i, a_1, a_2, \dots))^2 \end{aligned}$$

La minimisation de ce critère constitue l'objectif à atteindre : plus il est minimal, plus le modèle est parfaitement ajusté et les erreurs de modélisation sont minimales.

III.3 Minimisation du critère des moindres carrés

La figure suivante présente le graphe de l'erreur carrée $\mathcal{E}$ par rapport au modèle $(\hat{a}_1, \hat{a}_2, \dots)$.

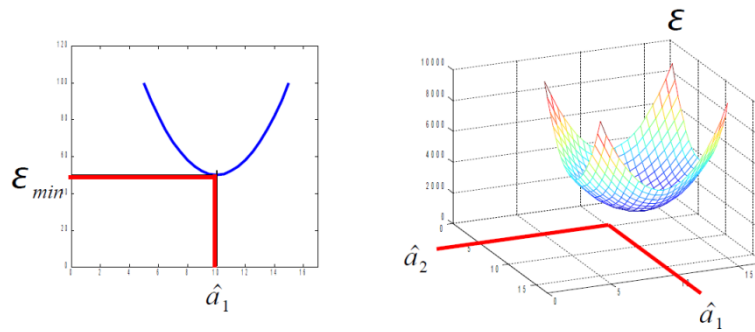


Figure III.2 Dépendances quadratique des erreurs par rapport au modèle

La Figure III.2 montre que le modèle présente une dépendance quadratique par rapport au critère quadratique de l'erreur. Si on suppose qu'il y a indépendance/orthogonalité des paramètres, alors le critère d'erreur présente un minimum aux dérivées de l'erreur par rapport aux paramètres :

$$\frac{\partial \mathcal{E}}{\partial a_1} = 0 \quad \frac{\partial \mathcal{E}}{\partial a_2} = 0 \quad \dots \quad \frac{\partial \mathcal{E}}{\partial a_k} = 0$$

Avec :

$$\frac{\partial^2 \mathcal{E}}{\partial a_1^2} > 0 \quad \frac{\partial^2 \mathcal{E}}{\partial a_2^2} > 0 \quad \dots \quad \frac{\partial^2 \mathcal{E}}{\partial a_k^2} > 0$$

La résolution des équations précédentes nous donne des valeurs du modèle dites optimale par rapport au critère quadratique :

$$\frac{\partial \mathcal{E}}{\partial a_i} = 0 \quad \xrightarrow{\text{Résolution des } k \text{ équations}} \quad \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k$$

Et l'obtention du modèle le plus proche des mesures :

$$\hat{x}(t) = x(t, \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k)$$

La résolution de telles équations peut être plus au moins difficile à mettre en œuvre, selon le modèle voulu (linéaire, non linéaire...etc.). Dans ce qui suit, nous n'allons considérer que les méthodes d'obtention de modèles linéaires.

III.4 Modèle linéaire

Soit un modèle de variable explicative $x(t)$, caractérisé par un jeu de paramètres $p = \{a_1, a_2 \dots a_k\}$. Le modèle est linéaire par rapport à ses paramètres s'il vérifie :

$$x(t, \lambda_1 p_1 + \lambda_2 p_2) = \lambda_1 x(t, p_1) + \lambda_2 x(t, p_2)$$

Un modèle linéaire par rapport à ses paramètres a une expression du type :

$$x(t) = a_1 f_1(t) + a_2 f_2(t) + \dots$$

Où les $f_1(t), f_2(t) \dots$ sont les formants du modèle.

Exemples :

$$x(t) = a_0 + a_1 t + a_2 t^2 \dots \quad x(t) = A \cos(10t) + B \sin(10t)$$

III.5 Ajustement des paramètres du modèle linéaire

On dispose de n mesures des sorties (expérimentales) d'un système ($y(t_1), y(t_2) \dots y(t_n)$). On désire ajuster les paramètres du modèle linéaire :

$$x(t) = a_1 f_1(t) + a_2 f_2(t) + \dots + a_k f_k(t)$$

pour qu'ils correspondent aux mesures expérimentales.

Ecrivons les expressions du modèle aux instants de mesures : ($x(t_1), x(t_2) \dots x(t_n)$), ainsi que le critère quadratique d'erreur cumulée à optimiser :

$$\mathcal{E}(a_1, a_2 \dots) = \sum_{i=1}^n (y(t_i) - x(t_i))^2$$

Le tableau suivant résume les données du problème :

n	<i>abscisse</i>	<i>mesure</i>	<i>modèle</i>
1	t_1	$y(t_1)$	$x(t_1) = a_1 f_1(t_1) + a_2 f_2(t_1) + \dots + a_k f_k(t_1)$
2	t_2	$y(t_2)$	$x(t_2) = a_1 f_1(t_2) + a_2 f_2(t_2) + \dots + a_k f_k(t_2)$
...	...	...	
n	t_n	$y(t_n)$	$x(t_n) = a_1 f_1(t_n) + a_2 f_2(t_n) + \dots + a_k f_k(t_n)$

Nous pouvons donc écrire :

$$\mathcal{E} = \sum_{i=1}^n [y(t_i) - a_1 f_1(t_i) - a_2 f_2(t_i) - \dots - a_k f_k(t_i)]^2$$

Les dérivées du critère quadratique $\mathcal{E}$ sont données par :

$$\begin{aligned}
\frac{\partial \mathcal{E}}{\partial a_1} &= - 2 \sum_{i=1}^n f_1(t_i) [y(t_i) - a_1 f_1(t_i) - a_2 f_2(t_i) - \dots - a_k f_k(t_i)] \\
\frac{\partial \mathcal{E}}{\partial a_2} &= - 2 \sum_{i=1}^n f_2(t_i) [y(t_i) - a_1 f_1(t_i) - a_2 f_2(t_i) - \dots - a_k f_k(t_i)] \\
&\dots \\
\frac{\partial \mathcal{E}}{\partial a_k} &= - 2 \sum_{i=1}^n f_k(t_i) [y(t_i) - a_1 f_1(t_i) - a_2 f_2(t_i) - \dots - a_k f_k(t_i)]
\end{aligned}$$

Les paramètres optimaux $(\hat{a}_1, \hat{a}_2, \dots, \hat{a}_k)$ sont obtenus en résolvant le problème :

$$\frac{\partial \mathcal{E}}{\partial a_1} = 0 \quad \frac{\partial \mathcal{E}}{\partial a_2} = 0 \quad \dots \quad \frac{\partial \mathcal{E}}{\partial a_k} = 0 \quad \rightarrow \quad \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k$$

III.6 Présentation homogène du problème

L'équation de minimisation est présentée comme suit :

$$\begin{aligned}
a_1 \sum_{i=1}^n f_1^2(t_i) + a_2 \sum_{i=1}^n f_1(t_i) f_2(t_i) + \dots + a_k \sum_{i=1}^n f_1(t_i) f_k(t_i) &= \sum_{i=1}^n f_1(t_i) y(t_i) \\
a_1 \sum_{i=1}^n f_1(t_i) f_2(t_i) + a_2 \sum_{i=1}^n f_2^2(t_i) + \dots + a_k \sum_{i=1}^n f_2(t_i) f_k(t_i) &= \sum_{i=1}^n f_2(t_i) y(t_i) \\
&\dots \\
a_1 \sum_{i=1}^n f_1(t_i) f_k(t_i) + a_2 \sum_{i=1}^n f_1(t_i) f_2(t_i) + \dots + a_k \sum_{i=1}^n f_k^2(t_i) &= \sum_{i=1}^n f_k(t_i) y(t_i)
\end{aligned}$$

Avec :

$$\begin{aligned}
\text{Vecteur des paramètres à identifier } \underline{\theta} &= \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_k \end{pmatrix} & \text{Vecteur des mesures pondérées } \underline{Q} &= \begin{pmatrix} \sum_{i=1}^n f_1(t_i) y_i \\ \sum_{i=1}^n f_2(t_i) y_i \\ \dots \\ \sum_{i=1}^n f_k(t_i) y_i \end{pmatrix} \\
R &= \begin{pmatrix} \sum_{i=1}^n f_1^2(t_i) & \sum_{i=1}^n f_1(t_i) f_2(t_i) & \dots & \sum_{i=1}^n f_1(t_i) f_k(t_i) \\ \sum_{i=1}^n f_2(t_i) f_1(t_i) & \sum_{i=1}^n f_2^2(t_i) & \dots & \sum_{i=1}^n f_2(t_i) f_k(t_i) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^n f_k(t_i) f_1(t_i) & \sum_{i=1}^n f_k(t_i) f_2(t_i) & \dots & \sum_{i=1}^n f_k^2(t_i) \end{pmatrix} & \text{Matrice de pondération symétrique}
\end{aligned}$$

Les équations de minimisation peuvent donc s'écrire sous la forme matricielle :

$$R\theta = Q$$

La résolution de cette équation matricielle donne la solution optimale pour le modèle :

$$\hat{\theta} = R^{-1}Q$$

Remarque importante : La matrice R doit être inversible.

III.7 Ecriture matricielle du problème des moindres-carrés

Soit $x(t)$ le modèle linéaire par rapport à ses paramètres :

$$x(t) = a_1 f_1(t) + a_2 f_2(t) + \dots + a_k f_k(t)$$

Soit X l'ensemble des valeurs prises par le modèle et Y les mesures :

$$\underline{X} = \begin{pmatrix} x(t_1) \\ x(t_2) \\ \dots \\ x(t_n) \end{pmatrix} = \begin{pmatrix} a_1 f_1(t_1) + a_2 f_2(t_1) + \dots + a_k f_k(t_1) \\ a_1 f_1(t_2) + a_2 f_2(t_2) + \dots + a_k f_k(t_2) \\ \dots \\ a_1 f_1(t_n) + a_2 f_2(t_n) + \dots + a_k f_k(t_n) \end{pmatrix} \quad \underline{Y} = \begin{pmatrix} y(t_1) \\ y(t_2) \\ \dots \\ y(t_n) \end{pmatrix}$$

Posons: $\underline{H} = \begin{pmatrix} f_1(t_1) & f_2(t_1) & \dots & f_k(t_1) \\ f_1(t_2) & f_2(t_2) & \dots & f_k(t_2) \\ \dots & \dots & \dots & \dots \\ f_1(t_n) & f_2(t_n) & \dots & f_k(t_n) \end{pmatrix}$ $\underline{\theta} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_k \end{pmatrix} \Rightarrow \underline{X} = \underline{H} \underline{\theta}$

L'erreur quadratique cumulée s'écrit :

$$\mathcal{E} = \sum_{i=1}^n (y(t_i) - x(t_i))^2 = \sum_{i=1}^n \varepsilon_i^2 \quad \underline{\mathcal{E}} = (\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n) \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix}$$

Sachant que : $\underline{\varepsilon} = \underline{Y} - \underline{X}$ et que $\underline{X} = \underline{H} \underline{\theta}$, l'erreur cumulée s'écrit :

$$\underline{\mathcal{E}} = (\underline{Y} - \underline{H} \underline{\theta})^T (\underline{Y} - \underline{H} \underline{\theta})$$

Et la minimisation du critère se fait pour :

$$\frac{\partial \mathcal{E}}{\partial \underline{\theta}} = \begin{pmatrix} \frac{\partial \mathcal{E}}{\partial a_1} \\ \frac{\partial \mathcal{E}}{\partial a_2} \\ \dots \\ \frac{\partial \mathcal{E}}{\partial a_k} \end{pmatrix} = \underline{0}$$

La solution optimale sera obtenue pour .

$$\frac{\partial \mathcal{E}}{\partial \underline{\theta}} = -H^T (\underline{Y} - H \underline{\theta}) - (\underline{Y} - H \underline{\theta})^T H$$

En remarquant que $\mathcal{E}$ est scalaire, on obtient :

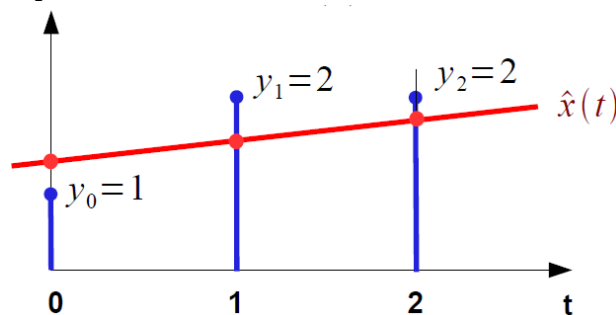
$$\frac{\partial \mathcal{E}}{\partial \underline{\theta}} = -2 H^T (\underline{Y} - H \underline{\theta})$$

Et la solution optimale est donc :

$$\hat{\underline{\theta}} = (H^T H)^{-1} H^T \underline{Y}$$

Exercice d'application : Modélisation de 3 mesures successives par un modèle affine de la forme : $x(t) = a + bt$.

- 1) Déterminer les valeurs prises par le modèle pour $t = 0, 1, 2$.
- 2) Construire le tableau des valeurs prises par le modèle aux instants de mesure, le vecteur des observations.
- 3) En déduire les matrices/vecteurs Y, H et $\theta = (a, b)^T$
- 4) Calculer la solution optimale $\hat{\theta}$



Solution : On a : $x(t) = a + bt = af_1(t) + bf_2(t)$, avec : $f_1(t) = 1$ et $f_2(t) = t$

$$\text{Donc } H = \begin{bmatrix} f_1(t_1) & f_2(t_1) \\ f_1(t_2) & f_2(t_2) \\ f_1(t_3) & f_2(t_3) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, Y = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

$$\hat{\underline{\theta}} = (H^T H)^{-1} H^T \underline{Y} \rightarrow \hat{\underline{\theta}} = \left(\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.83 \\ 0.5 \end{bmatrix}$$

Donc : $a = 0.83$ et $b = 0.5000$.

Ou en utilisant Matlab : `>> H=[1 0;1 1;1 2] ; Y=[1;2;2] ; tetas=inv(H'*H)*H'*Y`

tetas = [0.8333

0.5000]