

# Chapitre 1 : Généralités sur la l'identification

## I.1. Définition de l'identification :

- L'identification d'un système c'est la détermination de son modèle mathématique sur la base des observations expérimentales entrées/sorties. Le traitement mathématique des réponses graphiques du système est appelé **IDENTIFICATION**. Le modèle obtenu est dit de représentation.
- Identifier un procédé ou système consiste à proposer une structure entre son entrée et sa sortie et à déterminer à partir du couple entrée-sortie, les valeurs des paramètres du modèle. Le modèle ainsi trouvé doit, dans son domaine de validité, se comporter comme la réalité (physique) ou au moins s'en approcher au plus près.
- L'identification a pour but de déterminer un modèle mathématique, le plus souvent de représentation, du système étudié à partir d'un ensemble de mesure E/S du système.

**I.2. Type de modèles :** Il existe une multitude de types de modèles, selon les applications. Les plus populaires sont les modèles de connaissance et les modèles de représentation.

### + Les modèles de connaissance :

- Ce modèle est obtenu en écrivant toutes les équations différentielles qui régissent le fonctionnement du système (en se basant sur les lois de la physique de la chimie... (Newton, Kirchoff...)).
- Les paramètres d'un tel modèle ont une interprétation physique (température, vitesse, frottement, longueur, résistance électrique, ect.
- Difficultés de décrire fidèlement les phénomènes complexes
- Hypothèses simplificatrices
- Dilemme- précision-simplicité
- Un modèle simple est faux, un modèle compliqué est inutilisable.

### + Les modèles de représentation :

- Modèle déterminé à partir de données expérimentales (données entrée-sortie).
- Modèle dont la structure et les paramètres sont sans rapport avec le système réel.
- Les paramètres de ce modèle n'ont pas d'interprétation physique.

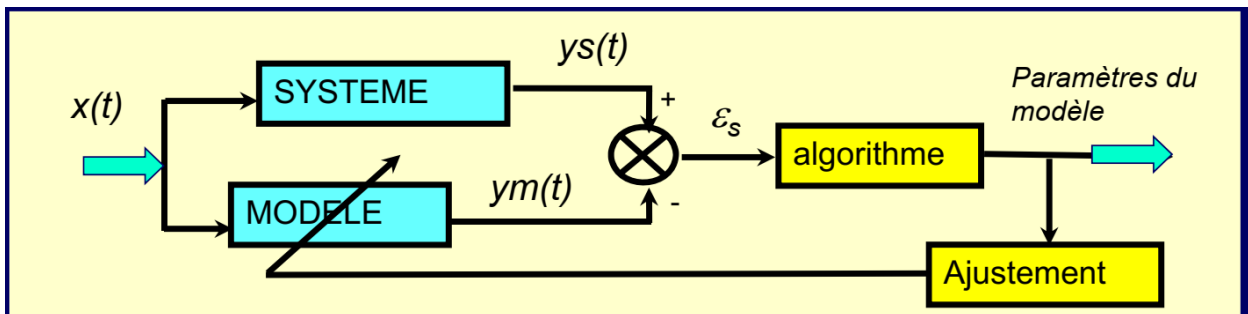
- La structure du modèle est fixée à priori, on parle de 'boite noire'.

### 1.3 : Classification des méthodes d'identification

**1.3.1. Méthode de base (réponse graphique)** Sur la base d'une connaissance à priori du système à identifier, on fixe une structure du modèle comportant des coefficients inconnus « boite noire » elles sont basées sur les réponses graphiques telles que :

- la réponse indicielle
- la réponse impulsionnelle
- la réponse à une rampe
- le diagramme de Bode et de Nyquist
- la méthode de Strejc
- la méthode de Broïda

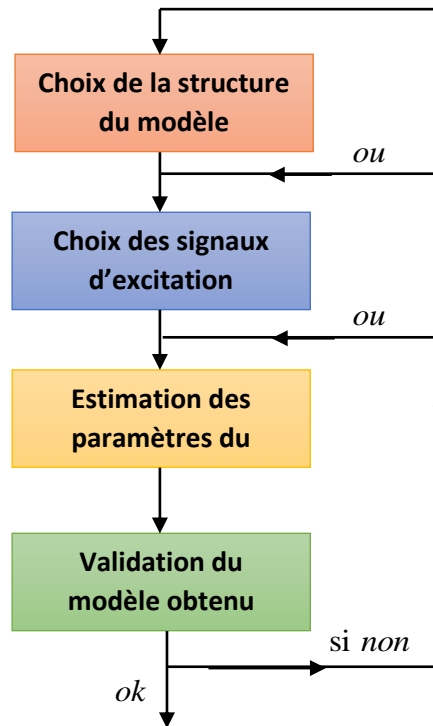
### 1.3.2. Méthode du modèle :



Ajuster manuellement ou automatiquement la structure ou les paramètres du modèle jusqu'à ce que  $\varepsilon_s \rightarrow \min$

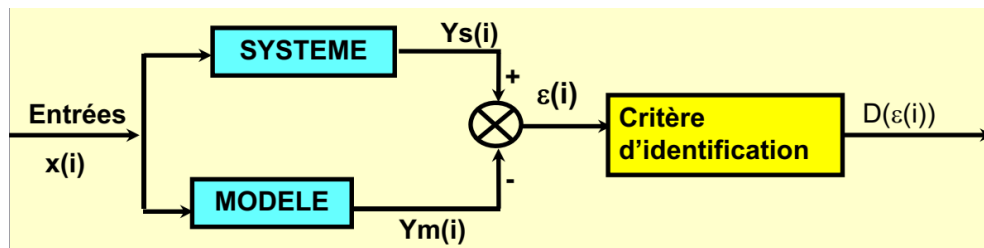
**1.3.3. Méthode statistique :** Basées sur les MC (moindres-carrés) : Elle consiste à la détermination des coefficients inconnus du modèle de façon que la différence entre les N sorties réelles du système et celles du modèle soit minimale selon un critère donné qu'on résout par un algorithme d'identification.

**1.4. Algorithme général d'identification**



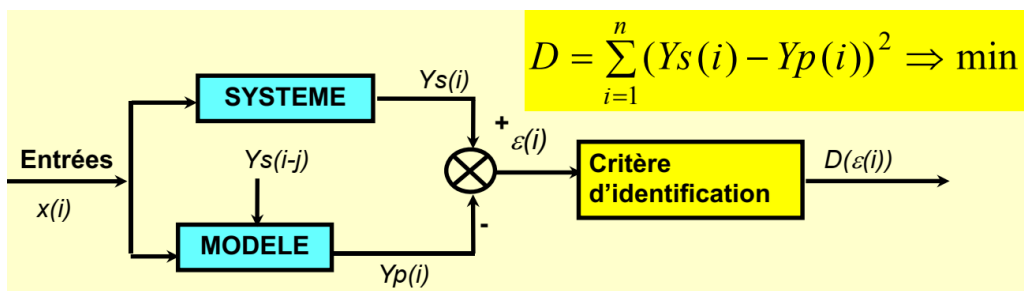
**1.5. Critères d'identification**

**1.5.1. Distance d'état :** Basée sur la différence entre la sortie du système et du modèle



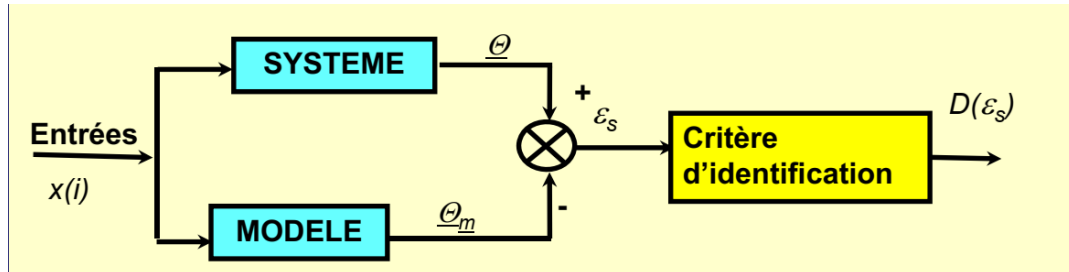
Critère d'identification à minimiser:  $D = \sum_{i=1}^n (Y_s(i) - Y_m(i))^2$

**1.5.2. Distance de prédiction :** Basée sur la différence entre la sortie du système et celle que prédit le modèle au même instant

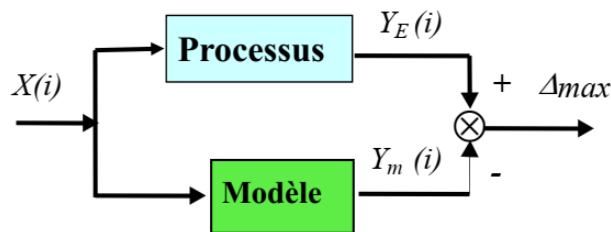


Critère d'identification à minimiser :  $D = \sum_{i=1}^n (Y_s(i) - Y_p(i))^2$

**1.5.3. Distance de structure :** Basée sur la différence entre les paramètres du système et ceux du modèle.



**1.6. Validation du modèle**



$$\Delta \max(i) = \left| \frac{Y_m \max - Y_E \max}{Y_E \max} \right| \cdot 100\% \leq \varepsilon_{\text{admissible}}$$

Il faut que l'erreur soit minimale dans les systèmes industriels

**1.7. Expérimentation (Choix du signal d'excitation :)**

Pour bien identifier, il faut bien exciter dans tout le spectre de fréquences susceptible de contenir des constantes de temps du système.

- Le signal  $\sin(\omega t)$  : parfait d'un point de vue spectre (balayage en fréquence) mais peu de systèmes acceptent ce genre d'entrées.
- Le signal  $\delta(t)$  : parfait du point de vue théorique, mais difficile de réaliser une bonne approximation de l'impulsion.
- Le signal  $u(t)$  : moins bon d'un point de vue spectral ( $u(f) = 1/2\delta(f) + j/21\pi f$ ), mais facile à implanter.
- Le signal  $b(t)$  : bruit blanc idéal d'un point de vue spectral mais comment le réaliser ?

Pour bien identifier il faut appliquer une entrée « riche » en fréquences

**Solution standard:** Séquence Binaire Pseudo-Aléatoires (SBPA)

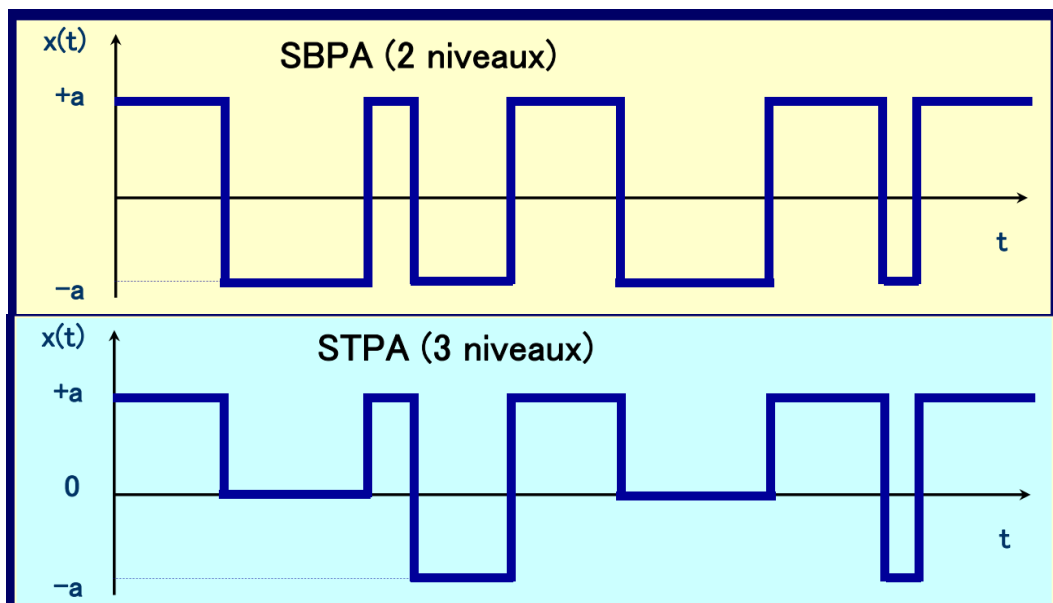
- Succession d'impulsions rectangulaires modulées en largeur
- très riches en fréquences – spectre uniforme de presque 0 à  $0.5f_e$
- Génération: à l'aide de registres à décalage bouclés
- Physiquement réalisable
- Amplitude limitée : ne pas trop perturber le processus, resté en linéaire

**Exemple :** Génération d'une SBPA de longueur  $31=2^5-1$ 

Variation aléatoire de la largeur des impulsions à l'intérieur d'une longueur de séquence

**Éléments caractéristiques:**

- nombres de cellules (N)
- durée maximale d'une impulsion ( $t_{im}=Nt_e$ )
- longueur de la séquence ( $L=2N-1$ )



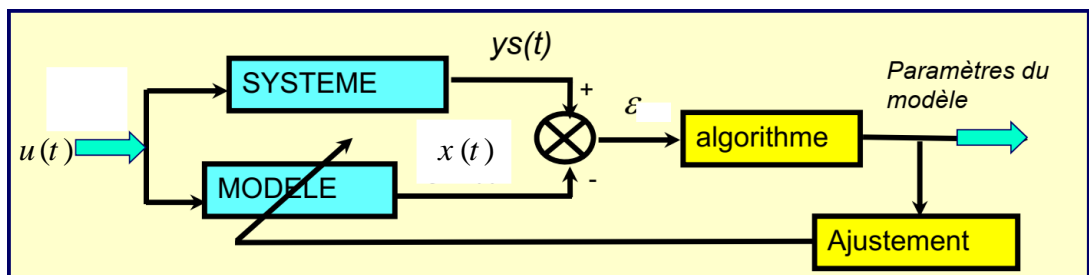
# Chapitre 1

## Identification des systèmes par la méthode des Moindres-Carrés

### Partie1 : Ajustement du modèle par les moindres-carres

#### 1. Positionnement du problème

Le problème d'identification que nous allons considérer est le suivant :



Soient :

- $y_s(t_i) = (y_{s1}, y_{s2}, \dots, y_{sp})$  les mesures prises par le système (les observations)
- $x(t_i, a_1, a_2, \dots, a_k)$  Les valeurs prises par le modèle aux instants d'observation avec  $(a_1, a_2, \dots, a_k)$  sont les paramètres du modèles
- Posons  $\varepsilon$  écart (erreur) entre les mesures et les valeurs du modèle tel que :

$$\varepsilon = y_s(t_i) - x(t_i, a_1, a_2, \dots, a_k)$$

- L'erreur quadratique cumulée est :  $E(a_1, a_2, \dots, a_k) = \sum_{i=1}^k \varepsilon_i^2 = \sum_{i=1}^k (y_s(t_i) - x(t_i, a_1, a_2, \dots))^2$

Le critère d'erreur présente un minimum pour:  $J = \frac{\partial E}{\partial a_i} = 0 \Rightarrow \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k = ?$  c'est-à-dire le

problème revient à résoudre les k équations suivantes pour trouver les paramètres  $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_k$

$$\frac{\partial E}{\partial a_1} = 0, \frac{\partial E}{\partial a_2} = 0, \dots, \frac{\partial E}{\partial a_k} = 0 \Rightarrow \hat{a}_1, \hat{a}_2, \dots, \hat{a}_k = ?$$

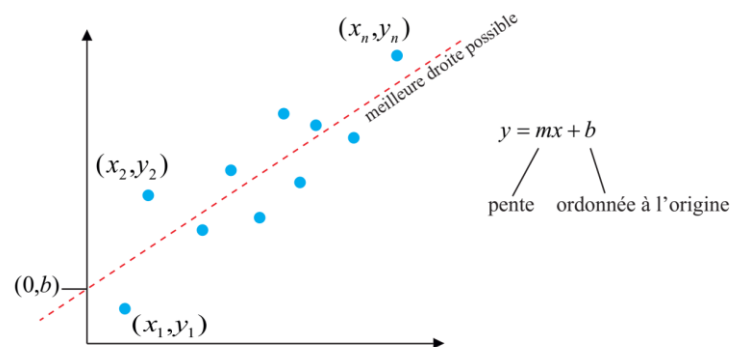
Le modèle le plus proche des mesures au sens du critère d'erreur est :

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

**Remarque :** Selon la nature du modèle (linéaire ou non-linéaire par rapport aux paramètres), la résolution des équations peut poser problème. Ce cours porte essentiellement sur les méthodes de base applicable aux modèles linéaires.

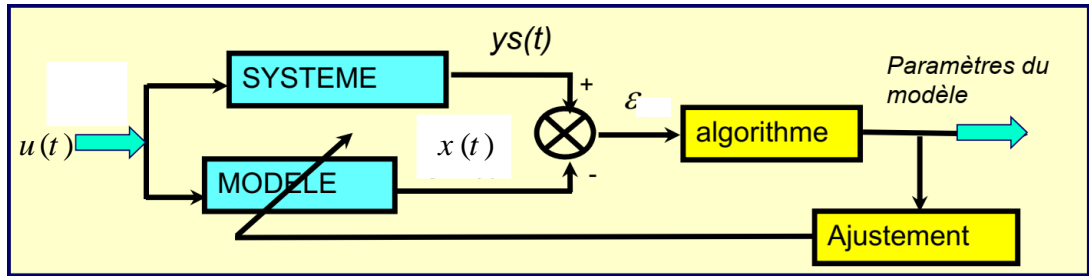
### 2. Méthode de moindres-carrés

Supposons des données expérimentales portées sur un graphique qui forment un nuage de points (de 1 jusqu'à  $n$ .)



l'objectif : On cherche à déterminer l'équation de la droite qui passe le plus près possible de l'ensemble des points.

- L'idée de la méthode ça va être de faire passer une droite au mieux par un certain nombre de points.
- Minimiser l'écart entre le point et la droite, ( $\varepsilon_i$  représente l'écart)
- Minimiser la somme des écarts au carrée
- Le but donc de savoir quelle est l'équation de la droite qui passe au mieux par les points  
 $a=?$   $b=?$



**❖ Choix du modèle :**

Soit un modèle  $x(t)$  de variable  $t$ , caractérisé par des paramètres  $\{a_1, a_2, \dots, a_k\}$

C'est-à-dire  $x(t, a_1, a_2, \dots, a_k)$

➤ **Modèle linéaire par rapport à ses paramètres**

- Le modèle est linéaire par rapport à ses paramètres  $\lambda_i$  s'il vérifie

$$x(t, \lambda_1 a_1 + \lambda_2 a_2) = \lambda_1 x(t, a_1) + \lambda_2 x(t, a_2)$$

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

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

Avec  $\{a_1, a_2, \dots, a_k\}$  paramètres du modèle et  $f_1(t), f_2(t), \dots, f_k(t)$  sont les "formants" du modèle

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

Le modèle étudié dans notre cas est de la forme :  $Y = X \theta + E$

Avec  $Y$  vecteur de sortie,

$X$  : vecteur des paramètres prises par le modèle

$E$  : vecteur d'erreur quadratique et  $\theta$  est le vecteur de paramètres

**3. Ecriture matricielle du problème des moindres-carrés :**

Soit  $x(t)$  un 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 : on peut écrire alors

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

ou encore



$$X = \begin{bmatrix} x(t_1) \\ x(t_2) \\ \dots \\ x(t_n) \end{bmatrix} = \underbrace{\begin{bmatrix} f_1(t_1) & f_2(t_1) & \dots & f_k(t_1) \\ f_1(t_2) & f_2(t_2) & \dots & f_k(t_2) \\ \dots & \dots & \dots & \dots \\ f_1(t_n) & f_2(t_n) & \dots & f_k(t_n) \end{bmatrix}}_H \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_k \end{bmatrix}}_{\theta} \Rightarrow X = H \theta$$

Avec  $H$  est la matrice/vecteur de régression

L'erreur à chaque instant  $t$  s'écrit comme suit:  $\varepsilon_i = y(t_i) - x(t_i)$

Alors l'erreur quadratique cumulée s'écrit:

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

Avec  $E$  : Erreur globale est  $E = Y - X$  et que  $X = H \theta$

Alors on peut écrire :  $E_Q = E^T E = (Y - H \theta)^T (Y - H \theta)$

La minimisation du critère se fait pour :

$$J = \frac{\partial E_Q}{\partial \theta} = \begin{bmatrix} \frac{\partial E_Q}{\partial a_1} \\ \frac{\partial E_Q}{\partial a_2} \\ \dots \\ \frac{\partial E_Q}{\partial a_k} \end{bmatrix} = 0$$

La solution optimale sera obtenue pour :

$$\frac{\partial E_Q}{\partial \theta} = -H^T (Y - H \theta) - (Y - H \theta)^T H$$

on obtient :

$$\begin{aligned} \frac{\partial E_Q}{\partial \theta} &= -2H^T (Y - H \theta) = 0 \\ \Rightarrow H^T Y - H^T H \theta &= 0 \\ \Rightarrow H^T H \theta &= H^T Y \\ \Rightarrow \underbrace{(H^T H)^{-1}}_I (H^T H) \theta &= \underbrace{(H^T H)^{-1}}_I H^T Y \end{aligned}$$

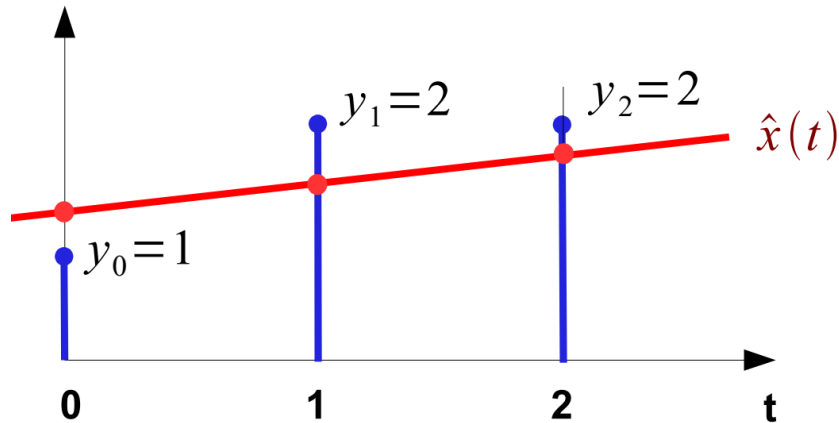
Alors la solution optimale est :

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

- L'erreur d'estimation :  $E_{est} = \hat{\theta} - \theta$

**4. Exemple d'application Modélisation de 3 mesures successives**

par un modèle affine de la forme  $x(t) = a + bt$



- 1) Déterminer les paramètres prisés par le modèle pour  $t = (0,1,2)$
- 2) Déterminer le vecteur des observation  $Y$  pour les mêmes instants précédents
- 3) En déduire alors la matrice  $H$
- 4) Calculer alors la solution optimale  $\hat{\theta} = ?$

**Solution**

<u>n</u>	<u>Abscisse</u>	<u>Mesure</u>	<u>Modèle</u>
1	$t_1 = 0$	$y_s(t_1) = 1$	$x(t_1) = a_1 + 0b$
2	$t_2 = 1$	$y_s(t_2) = 2$	$x(t_2) = a + b$
3	$t_3 = 2$	$y_s(t_3) = 2$	$x(t_3) = a + 2b$

**Ecriture du problème sous forme matricielle**

$$X = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}}_H \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\theta} = H\theta \quad \text{et} \quad Y = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

La solution optimale par la méthode du moindres-carrés

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

$$H^T H = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \Rightarrow (H^T H)^{-1} = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$$

$$H^T Y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

On obtient alors  $\hat{\theta} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = (H^T H)^{-1} H^T Y = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 7 \\ 3 \end{bmatrix}$

Finalement on trouve  $\hat{\theta} = \begin{pmatrix} \hat{a} = 7/6 \\ \hat{b} = 1/2 \end{pmatrix}$

Le modèle estimé devient alors,  $x(t) = \hat{a} + \hat{b}t = 7/6 + 0.5t$

Exercice:

Prenant un modèle polynomial d'un système non linéaire continu :

$$y(t) = \theta_0 + \theta_1 u(t) + \theta_2 u(t)^2$$

Les données de test sont

$$u(1) = 1, \quad u(2) = 2, \quad u(3) = 3, \quad u(4) = 4 \\ y(1) = 6, \quad y(2) = 17, \quad y(3) = 34, \quad y(4) = 57$$

Solution

$$\Phi = \begin{pmatrix} 1 & u(1) & u(1)^2 \\ 1 & u(2) & u(2)^2 \\ 1 & u(3) & u(3)^2 \\ 1 & u(4) & u(4)^2 \end{pmatrix} \quad Y = \begin{pmatrix} 6 \\ 17 \\ 34 \\ 57 \end{pmatrix} \quad \hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y$$

$$\theta = [1 \ 2 \ 3]$$

## 5. Analyse de la méthode des moindres-carrés

### 5.1. Pondération du critère J

**Principe :** Avec N mesures aux instants  $t_1, \dots, t_n$ , l'erreur s'écrit:

$$E = Y - X = Y - H \theta$$

- Cette équation matricielle comporte N lignes, toutes les erreurs de prédiction  $\varepsilon(t_1), \dots, \varepsilon(t_n)$  ne doivent pas nécessairement avoir la même importance dans le critère quadratique.
- Il est possible de les pondérer, par exemple comme suit:

$$E_Q = Q E = Q (Y - H \theta)$$

Où  $Q$  est une matrice de pondération de dimension  $(N \times N)$ .

Le critère quadratique des moindres carrés dans ce cas s'écrit :

$$J(\theta) = E_Q^T E_Q = (Y - H \theta)^T Q^T Q (Y - H \theta)$$

Le vecteur des paramètres estimés devient alors

$$\begin{aligned}\frac{dJ(\theta)}{d\theta} = 0 &\Rightarrow -2H^T Q^T Q (Y - H\theta) = 0 \\ &\Rightarrow H^T Q^T Q Y = H^T Q^T Q H \theta \\ &\Rightarrow \hat{\theta} = (H^T Q^T Q H)^{-1} H^T Q^T Q Y\end{aligned}$$

### ❖ **Choix de la matrice de pondération ?**

- La matrice de pondération Q doit être Symétrique et Définie positive :  $Q = Q^T > 0$
- Un critère de choix indique que les mesures plus anciennes possèdent une influence réduite dans le calcul des paramètres actuels.
- L'influence des mesures plus anciennes est artificiellement réduite par l'introduction du facteur d'oubli  $\lambda$ , par exemple:

$$Q^T Q = \begin{bmatrix} \lambda^{N-1} & 0 & \dots & 0 \\ 0 & \ddots & \dots & 0 \\ \vdots & \vdots & \lambda^1 & \vdots \\ 0 & 0 & \dots & \lambda^0 \end{bmatrix}$$

avec  $\lambda \leq 1$  (souvent  $0.9 \leq \lambda \leq 0.99$ )

**Remarque :** Si Q est la matrice identité, on aura le critère des moindres carrés conventionnel

### **5.2. Biais de l'estimation des paramètres :**

- Le biais d'un estimateur est l'espérance de l'écart entre l'estimation et les vraies valeurs.

$$\text{Biais} = b = E(\hat{\theta} - \theta)$$

- L'estimateur sans biais fournit une valeur **non-décalée** par rapport à la vraie valeur
- $b = 0$  : Estimateur non biaisé (**pas d'erreur d'estimation**);
- $b \neq 0$  Estimateur biaisé (il y a une erreur d'estimation)

- Soit  $\theta$  la valeur théorique des paramètres.

- Donc le modèle parfait a pour valeur  $X = H\theta$
- Les mesures ont pour valeur  $Y = H\theta + v$  ( $v$  est le bruit de mesure)
- L'estimateur de MC  $\hat{\theta} = (H^T H)^{-1} H^T Y = (H^T H)^{-1} H^T (H\theta + v)$

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

$$\hat{\theta} = (H^T H)^{-1} (H^T H\theta + H^T v)$$

$$\hat{\theta} = \underbrace{(H^T H)^{-1} (H^T H)}_I \theta + (H^T H)^{-1} H^T v$$

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

Alors l'écart entre l'estimation et la vraie valeur des paramètres du modèle est

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

- L'espérance de l'erreur d'estimation est :  $E[\hat{\theta} - \theta] = E[(H^T H)^{-1} H^T v]$

- Pour un bruit de mesure non corrélé à H, l'espérance de l'erreur d'estimation devient :

$$E[\hat{\theta} - \theta] = (H^T H)^{-1} H^T E[v]$$

▪ Seul  $E(v) = 0$  peut être nul pour que l'estimation soit sans biais.

▪ Dans les cas usuels (bruit à valeur moyenne nulle ou bruit blanc), l'estimateur des moindres-carrés est **sans-biais** : la valeur estimée des paramètres est proche de la vraie valeur en moyenne.

▪ Si le bruit n'est pas blanc, cet estimateur est biaisé c'est-à-dire  $E[\hat{\theta} - \theta] \neq 0$

▪ Il fournit une estimation non-décalée par rapport aux vraies valeurs. Mais en général (en présence d'un bruit coloré), l'estimateur des MC **est biaisé**

### • Solutions au problème de biais asymptotique de l'estimateur des MC

1- Choix d'autres structures de modèles pour le bruit (bruit à valeur moyenne nulle)

2- Méthode de la variable instrumentale (VI)

### 5.3. Méthode de la variable instrumentale (VI)

- Dérivée de l'estimateur des MC

- S'appuie sur la régression linéaire (peu gourmand en calculs)

- Asymptotiquement non biaisé lorsque le bruit est coloré

$$E[\hat{\theta} - \theta] = (H^T H)^{-1} H^T E[v]$$

- Il est convergent si les perturbations  $v$  ne sont pas corrélées avec le vecteur de régression  $H$  (et donc avec la sortie  $y(t)$  car  $u(t)$  est déterministe)

- **But recherché** : éliminer le biais de l'estimateur des MC lorsque la perturbation n'est pas un bruit blanc.
- Le principe est de déterminer un vecteur  $Z$  appelée instrument ou variable instrumentale VI dont les composantes sont non-corrélées avec la perturbation  $v$ , c'est-à-dire  $E[Z^T v] = 0$
- Posons  $\hat{\theta} = (Z^T H)^{-1} Z^T Y$  où  $Z$  est la matrice instrumentale. (seulement on remplace  $H^T$  par  $Z^T$ )
- Quelles sont les conditions sur  $Z$  pour que l'équation précédente ait un sens ? En poursuivant le calcul :

$$\begin{aligned}\hat{\theta} &= (Z^T H)^{-1} Z^T Y = (Z^T H)^{-1} Z^T (H\theta + v) \\ &= \underbrace{(Z^T H)^{-1} Z^T H}_{I} \theta + (Z^T H)^{-1} Z^T v \\ \Rightarrow \hat{\theta} &= \theta + (Z^T H)^{-1} Z^T v\end{aligned}$$

Pour que  $E[\hat{\theta}] = \theta$  il faut :

- 1-  $E[Z^T H] \neq 0$  et
- 2-  $E[Z^T v] = 0$

L'estimateur de la variable instrumentale est donc asymptotiquement non biaisé en général même si  $v$  est un bruit coloré

### Choix de la variable instrumentale :

- **Variable instrumentale à observations retardées** : certains auteurs ont pensé à introduire une variable dans  $H$  qui consiste à retarder les sorties d'une valeur  $R$ . (ce point sera détaillé dans le chapitre 3)

### 5.4. Variance de l'estimation des paramètres:

❖ **La variance** : d'une variable scalaire est définie comme suit:

$$\sigma^2(\alpha) = E[(\alpha - \bar{\alpha})^2] \text{ avec } \bar{\alpha} = E[\alpha]$$

❖ **Matrice de variance/covariance** : d'une grandeur vectorielle  $A = [\alpha_1 \ \alpha_2 \ \dots \alpha_n]^T$

est définie comme suit :  $\Gamma_A = E \left[ (A - \bar{A})(A - \bar{A})^T \right]$  qui est équivalent à

$$\Gamma_A = E \left[ (A - \bar{A})(A - \bar{A})^T \right] = E \left( \begin{bmatrix} \alpha_1 - \bar{\alpha}_1 \\ \alpha_2 - \bar{\alpha}_2 \\ \dots \\ \alpha_n - \bar{\alpha}_n \end{bmatrix} \begin{bmatrix} \alpha_1 - \bar{\alpha}_1 & \alpha_2 - \bar{\alpha}_2 & \dots & \alpha_n - \bar{\alpha}_n \end{bmatrix} \right)$$

$$\Gamma_A = E \left( \begin{bmatrix} (\alpha_1 - \bar{\alpha}_1)^2 & (\alpha_1 - \bar{\alpha}_1)(\alpha_2 - \bar{\alpha}_2) & \dots & (\alpha_1 - \bar{\alpha}_1)(\alpha_n - \bar{\alpha}_n) \\ (\alpha_2 - \bar{\alpha}_2)(\alpha_1 - \bar{\alpha}_1) & (\alpha_2 - \bar{\alpha}_2)^2 & \dots & (\alpha_2 - \bar{\alpha}_2)(\alpha_n - \bar{\alpha}_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_n - \bar{\alpha}_n)(\alpha_1 - \bar{\alpha}_1) & (\alpha_n - \bar{\alpha}_n)(\alpha_2 - \bar{\alpha}_2) & \dots & (\alpha_n - \bar{\alpha}_n)^2 \end{bmatrix} \right)$$

❖ **Variance de l'estimateur des moindres-carrés** (pour un bruit centré):

on a l'écart entre l'estimation et la vraie valeur des paramètres du modèle est

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

Alors la variance de l'estimateur des moindres-carrés sera comme suit :

$$\Gamma_\theta = E \left[ (\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right]$$

$$\Gamma_\theta = E \left[ \left( (H^T H)^{-1} H^T v \right) \left( (H^T H)^{-1} H^T v \right)^T \right]$$

$$\Gamma_\theta = E \left[ (H^T H)^{-1} H^T v v^T H (H^T H)^{-T} \right]$$

Pour un bruit  $v$  non-corrélé à la matrice de modèle  $H$ , la matrice de variance/covariance de l'estimateur peut être écrite de la manière suivante

$$\Gamma_\theta = E \left[ (H^T H)^{-1} H^T v v^T H (H^T H)^{-T} \right]$$

❖ **Cas du bruit blanc sur les mesures**

- Chaque mesure est supposée entachée d'un bruit blanc additif
- Il n'y a pas de corrélation du bruit de mesure entre deux (02) mesures

$$\begin{aligned} E[vv^T] &= E \left( \begin{bmatrix} v_{t1} & 0 & \dots & 0 \\ 0 & v_{t1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & v_m \end{bmatrix} \begin{bmatrix} v_{t1} & 0 & \dots & 0 \\ 0 & v_{t1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & v_m \end{bmatrix}^T \right) \\ &= E \begin{bmatrix} v_{t1}^2 & 0 & \dots & 0 \\ 0 & v_{t2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & v_m^2 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix} \end{aligned}$$

Pour un bruit blanc constant sur chaque mesure, de variance  $\sigma_Y^2$ , le bruit sur l'estimation des paramètres du modèle est :

$$\Gamma_\theta = E \left[ (H^T H)^{-1} H^T \sigma_Y^2 H (H^T H)^{-T} \right]$$

$$\Gamma_\theta = \sigma_Y^2 \underbrace{(H^T H)^{-1} H^T H (H^T H)^{-T}}_I$$

$$\Gamma_\theta = \sigma_Y^2 (H^T H)^{-T}$$



**± Partie 2 : Moindres-carrés récurrents**

❖ **Inconvénient de la méthode de MC-simple :**

- L'estimation de paramètres par la méthode des moindres carrés simples présente un inconvénient majeur, la nécessité de calculer l'inverse d'une matrice, ce qui est long et parfois impossible sur un microcontrôleur.

❖ **Motivation :**

- Cette méthode recalcule une nouvelle estimation des paramètres à chaque instant sur la base de l'estimation précédente.
- Elle est très pratique dans le cas où l'on a un très grand nombre de données ou que l'on désire utiliser le modèle à l'instant k sans attendre la fin des mesures.

**1. Principe de la méthode :**

± **Calcul sur n mesures :** On dispose de n observations auxquelles correspondent n valeurs du modèle

$$Y_n = \begin{bmatrix} y(t_1) \\ y(t_1) \\ \dots \\ y(t_n) \end{bmatrix} \text{ et } X_n = H_n \theta = \begin{bmatrix} f_1(t_1) & f_2(t_1) & \dots & f_k(t_1) \\ f_1(t_2) & f_2(t_2) & \dots & f_k(t_2) \\ \dots & \dots & \dots & \dots \\ f_1(t_n) & f_2(t_n) & \dots & f_k(t_n) \end{bmatrix} \theta$$

La solution optimale calculée sur les n mesures est (par MC):

$$\hat{\theta} = (H_n^T H_n)^{-1} H_n^T Y_n$$

Elle minimise le critère (erreur quadratique) calculé sur ces n mesures :

$$E_{Q_n} = (Y_n - X_n)^T (Y_n - X_n)$$

± **Inconvénients de la méthode classique :** La dimension (nxk) de la matrice H croît avec le nombre n de mesures

$$\hat{\theta}_n = (H_n^T H_n)^{-1} H_n^T Y_n$$

L'ajout d'une (n+1) ème mesure impose de recommencer la totalité du calcul

± **Mise en évidence de la solution**

Utiliser la technique de calcul  $\hat{\theta}_n = R_n^{-1} Q_n$

Avec  $R_n = H_n^T H_n$  de dimension (kxk)

et  $Q_n = H_n^T Y_n$  de dimension (kx1)

**Ajout de la (n+1)ème mesure :**

On dispose de **n+1** observations auxquelles correspondent **n+1** valeurs du modèle :

$$Y_{n+1} = \begin{bmatrix} y(t_1) \\ y(t_1) \\ \dots \\ y(t_n) \\ y(t_{n+1}) \end{bmatrix} \quad X_{n+1} = H_{n+1} \theta = \begin{bmatrix} f_1(t_1) & f_2(t_1) & \dots & f_k(t_1) \\ f_1(t_2) & f_2(t_2) & \dots & f_k(t_2) \\ \dots & \dots & \dots & \dots \\ f_1(t_n) & f_2(t_n) & \dots & f_k(t_n) \\ f_1(t_{n+1}) & f_2(t_{n+1}) & \dots & f_k(t_{n+1}) \end{bmatrix} \theta$$

On pose  $h_{n+1}^T = [f_1(t_{n+1}) \quad f_2(t_{n+1}) \quad \dots \quad f_k(t_{n+1})]$

On dispose de **n+1** observations auxquelles correspondent **n+1** valeurs du modèle

$$Y_{n+1} = \underbrace{\begin{bmatrix} Y_n \\ y(t_{n+1}) \end{bmatrix}}_{Y_{n+1}}, \quad X_{n+1} = \underbrace{\begin{bmatrix} X_n \\ x(t_{n+1}) \end{bmatrix}}_{X_{n+1}} = \underbrace{\begin{bmatrix} H_n \\ h_{n+1}^T(t_{n+1}) \end{bmatrix}}_{H_{n+1}} \theta,$$

Posons 
$$\hat{\theta}_{n+1} = (H_{n+1}^T H_{n+1})^{-1} H_{n+1}^T Y_{n+1}$$

➤ Calcul de  $H_{n+1}^T H_{n+1}$

$$H_{n+1}^T H_{n+1} = \begin{pmatrix} H_n^T & h_{n+1} \end{pmatrix} \begin{pmatrix} H_n \\ h_{n+1}^T \end{pmatrix}$$

$$H_{n+1}^T H_{n+1} = H_n^T H_n + h_{n+1} h_{n+1}^T$$

➤ Calcul de  $H_{n+1}^T Y_{n+1}$

$$H_{n+1}^T Y_{n+1} = \begin{pmatrix} H_n^T & h_{n+1} \end{pmatrix} \begin{pmatrix} Y_n \\ y_{n+1} \end{pmatrix}$$

$$H_{n+1}^T Y_{n+1} = H_n^T Y_n + h_{n+1} y_{n+1}$$

**➤ Expression récursive de base (les étapes à suivre):**

$$\hat{\theta}_{n+1} = (H_{n+1}^T H_{n+1})^{-1} H_{n+1}^T Y_{n+1}$$

$$\hat{\theta}_{n+1} = (H_n^T H_n + h_{n+1} h_{n+1}^T)^{-1} (H_n^T Y_n + h_{n+1} y_{n+1}) = R_{n+1}^{-1} Q_{n+1}$$

1. Construire la matrice  $H$  au rang  $n$
2. Calculer  $R_n = H_n^T H_n$  de dimension  $(k \times k)$
3. Calculer  $Q_n = H_n^T Y_n$  de dimension  $(k \times 1)$
4. En déduire  $\hat{\theta}_n = R_n^{-1} Q_n$  estimation initiale au rang  $n$

- Itérer pour passer du rang  $n$  au rang  $n+1$  :

5. Déterminer  $h_{n+1}^T = [f_1(t_{n+1}) \quad f_2(t_{n+1}) \quad \dots \quad f_k(t_{n+1})]$
6. Calculer  $R_{n+1} = R_n + h_{n+1}h_{n+1}^T$  ( $k \times k$ ) puis  $Q_{n+1} = Q_n + h_{n+1}y_{n+1}$  ( $k \times 1$ )
7. En déduire  $\hat{\theta}_{n+1} = R_{n+1}^{-1}Q_{n+1}$

**Calcul récursif évolué :**

à partir de

$$\hat{\theta}_{n+1} = \underbrace{(H_n^T H_n + h_{n+1}h_{n+1}^T)^{-1}}_{P_{n+1}} (H_n^T Y_n + h_{n+1}y_{n+1})$$

- 1- Posons  $P_n = (H_n^T H_n)^{-1}$  et  $P_{n+1} = (H_n^T H_n + h_{n+1}h_{n+1}^T)^{-1}$
- 2- Quelle est la relation entre  $P_n$  et  $P_{n+1}$  ?
- 3- Maintenant on utilise le lemme d'inversion suivant :

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

Avec  $A = (H_n^T H_n)$ ,  $B = h_{n+1}$ ,  $C = I$  et  $D = h_{n+1}^T$

- 4.. Appliquons le lemme à  $P_{n+1}$ , on trouve

$$P_{n+1} = \left( \underbrace{H_n^T H_n}_A + \underbrace{h_{n+1} I h_{n+1}^T}_{B C D} \right)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

$$P_{n+1} = \underbrace{(H_n^T H_n)^{-1}}_{P_n} - \underbrace{(H_n^T H_n)^{-1}}_{P_n} h_{n+1} \left( I + h_{n+1}^T \underbrace{(H_n^T H_n)^{-1}}_{P_n} h_{n+1} \right)^{-1} h_{n+1}^T \underbrace{(H_n^T H_n)^{-1}}_{P_n}$$

En fonction de  $P_n$  on trouve ;

$$P_{n+1} = P_n - \underbrace{P_n h_{n+1} (1 + h_{n+1}^T P_n h_{n+1})^{-1}}_{K_{n+1}} h_{n+1}^T P_n$$

Maintenant on pose

$$K_{n+1} = P_n h_{n+1} (1 + h_{n+1}^T P_n h_{n+1})^{-1}$$

Alors on peut écrire  $P_{n+1}$  sous cette forme

$$P_{n+1} = P_n - K_{n+1} h_{n+1}^T P_n$$

Ou encore

$$P_{n+1} = (I - K_{n+1} h_{n+1}^T) P_n$$

alors le vecteurs estimé devient :

$$\begin{aligned} \hat{\theta}_{n+1} &= (H_n^T H_n + h_{n+1}h_{n+1}^T)^{-1} (H_n^T Y_n + h_{n+1}y_{n+1}) \\ &= (P_n - K_{n+1} h_{n+1}^T P_n) (H_n^T Y_n + h_{n+1}y_{n+1}) \end{aligned}$$

$$\hat{\theta}_{n+1} = P_n H_n^T Y_n + P_n h_{n+1} y_{n+1} - K_{n+1} h_{n+1}^T P_n H_n^T Y_n - K_{n+1} h_{n+1}^T P_n h_{n+1} y_{n+1}$$

### Récurrance finale :

- 1- Calculer  $K_{n+1} = P_n h_{n+1} (1 + h_{n+1}^T P_n h_{n+1})^{-1}$
- 2- Calculer le nouveau vecteur estimé  $\hat{\theta}_{n+1} = \hat{\theta}_n + K_{n+1} (y_{n+1} - h_{n+1}^T \hat{\theta}_n)$
- 3-  $P_{n+1} = (I - K_{n+1} h_{n+1}^T) P_n$

Résumé

$$K_{n+1} = P_n h_{n+1} (1 + h_{n+1}^T P_n h_{n+1})^{-1}$$
$$\hat{\theta}_{n+1} = \hat{\theta}_n + K_{n+1} (y_{n+1} - h_{n+1}^T \hat{\theta}_n)$$
$$P_{n+1} = (I - K_{n+1} h_{n+1}^T) P_n$$

Dans cette expression,

$K_{n+1}$  : est le gain matriciel de correction

### ➤ Autre expression de la récurrance finale :

$$\hat{\theta}_{n+1} = \hat{\theta}_n + P_{n+1} h_{n+1} (y_{n+1} - h_{n+1}^T \hat{\theta}_n)$$
$$P_{n+1} = P_n - \frac{P_n h_{n+1} h_{n+1}^T P_n}{1 + h_{n+1}^T P_n h_{n+1}}$$

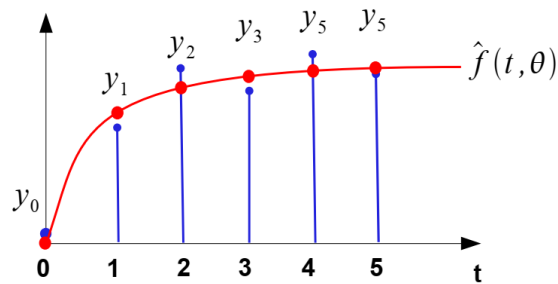
$P_{n+1}$  : est le gain matriciel de correction

**⚡ Partie 3 : Identification par moindres-carrés non linéaire**

**1- Position du problème**

Dans de nombreuses applications, la fonction modèle n'est pas linéaire par rapport à ses paramètres.

**Exemple :** réponse indicielle d'un circuit du 1er ordre



L'expression du modèle est :  $f(t, \theta) = K(1 - e^{-\frac{t}{\tau}})$

avec  $\theta = [K, \tau]^T$

- La dépendance vis à vis du paramètre  $\tau$  est non-linéaire, il n'est plus possible de résoudre le système par la méthode linéaire (pas de matrice H)

**⚡ Méthode de Gauss-Newton (méthode de gradient)**

**Principe :** La proposition est de partir d'une valeur initiale  $\theta_0$  des paramètres et de modifier itérativement la valeur de  $\theta$  d'un incrément  $\delta_\theta$  de façon à minimiser le critère d'erreur quadratique cumulée  $E_Q$  à chaque étape.

**⚡ Etape initiale :**

- Pour  $\theta_0 = [a_{10} \ a_{20} \ \dots \ a_{k0}]^T$  le modèle prend les valeurs  $F(t, \theta_0) = \begin{bmatrix} f(t_1, \theta_0) \\ f(t_2, \theta_0) \\ \vdots \\ f(t_n, \theta_0) \end{bmatrix}$

- L'erreur entre le modèle et les mesures est  $E = Y - F(t, \theta_0)$

- L'erreur quadratique cumulée est :  $E_Q(\theta_0) = E^T E = (Y - F(t, \theta_0))^T (Y - F(t, \theta_0))$

**Remarque :** généralement, l'erreur cumulée sera importante, les conditions initiales étant éloignées de la solution optimale.

**✚ Incrémentation :**

Dans cette étape on va modifier de la valeur de  $\theta_0$  d'un incrément  $\delta_\theta$  de façon à minimiser le critère d'erreur quadratique cumulée  $E_Q(\theta_0)$

- Pour  $\theta = \theta_0 + \delta_\theta$ , le modèle prend les valeurs  $F(t, \theta_0 + \delta_\theta)$  et l'erreur sera

$$E_Q(\theta_0 + \delta_\theta) = (Y - F(t, \theta_0 + \delta_\theta))^T (Y - F(t, \theta_0 + \delta_\theta))$$

- Pour se placer au minimum d'erreur, on choisit  $\delta_\theta$  tel que  $\frac{\partial E_Q(\theta_0 + \delta_\theta)}{\partial \delta_\theta} = 0$

$$\frac{\partial E_Q(\theta_0 + \delta_\theta)}{\partial \delta_\theta} = -2 \left[ \frac{\partial F(t, \theta_0 + \delta_\theta)}{\partial \delta_\theta} \right]^T \left( Y - \underbrace{F(t, \theta_0 + \delta_\theta)}_? \right) = 0$$

Calcul de  $F(t, \theta_0 + \delta_\theta)$  ?

Le développement de Taylor du 1er ordre du modèle permet d'approximer la nouvelle valeur du modèle à chaque instant d'observation :

$$f(t_i, \theta_0 + \delta_\theta) = f(t_i, \theta_0) + J(\theta_0) \delta_\theta$$

Avec  $J(\theta_0) = \nabla f(t_i, \theta_0)^T$  c'est le gradient de  $f$  en ligne.

L'extension à l'ensemble des points de calculs prend la forme matricielle suivante :

$$F(t, \theta_0 + \delta) = F(t, \theta_0) + J(\theta_0) \delta_\theta$$

Avec le gradient de  $f$  vaut  $J(\theta_0) = \nabla f(t_i, \theta_0)^T$ .

$$J(\theta_0) = \frac{\partial F(t, \theta_0 + \delta_\theta)}{\partial \delta_\theta} = \begin{bmatrix} \frac{\partial f(t_1)}{\partial a_{10}} & \frac{\partial f(t_1)}{\partial a_{20}} & \dots & \frac{\partial f(t_1)}{\partial a_{k0}} \\ \frac{\partial f(t_2)}{\partial a_{10}} & \frac{\partial f(t_2)}{\partial a_{20}} & \dots & \frac{\partial f(t_2)}{\partial a_{k0}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f(t_n)}{\partial a_{10}} & \frac{\partial f(t_n)}{\partial a_{20}} & \dots & \frac{\partial f(t_n)}{\partial a_{k0}} \end{bmatrix}, \text{ matrice Jacobienne de } f \text{ paramètres}$$

Maintenant, on calcule le minimum de l'erreur cumulée.

$$\begin{aligned} \frac{\partial E_Q(\theta_0 + \delta_\theta)}{\partial \delta_\theta} &= -2 \left[ \frac{\partial F(t, \theta_0 + \delta_\theta)}{\partial \delta_\theta} \right]^T \left( Y - \underbrace{F(t, \theta_0 + \delta_\theta)}_? \right) = 0 \\ &\Rightarrow -2J(\theta_0)^T (Y - F(t, \theta_0 + \delta_\theta)) = 0 \\ &\Rightarrow -2J(\theta_0)^T (Y - (F(t, \theta_0) + J(\theta_0) \delta_\theta)) = 0 \end{aligned}$$

On en déduit la valeur de l'accroissement  $\delta_\theta$  à faire sur les paramètres pour minimiser l'erreur:

$$\begin{aligned} \frac{\partial E_\theta(\theta_0 + \delta_\theta)}{\partial \delta_\theta} &= -2 \left[ \frac{\partial F(t, \theta_0 + \delta_\theta)}{\partial \delta_\theta} \right]^T \left( Y - \underbrace{F(t, \theta_0 + \delta_\theta)}_? \right) = 0 \\ -2J^T (Y - (F(t, \theta_0) + J\delta_\theta)) &= 0 \\ \Rightarrow J^T Y - J^T F(t, \theta_0) - J^T J \delta_\theta &= 0 \\ \Rightarrow J^T J \delta_\theta &= J^T (Y - F(t, \theta_0)) \end{aligned}$$

$$\Rightarrow \delta_\theta = (J^T(\theta_0)J(\theta_0))^{-1} J(\theta_0)^T (Y - F(t, \theta_0)) \dots \text{La valeur de l'incrément}$$

**✚ Itération :**

On itère en définissant les nouvelles valeurs des paramètres  $\theta_1 = \theta_0 + \delta_\theta$  et les nouvelles valeurs du modèle  $F(t, \theta_1)$ . La correction suivante à faire sera :

$$\delta_\theta = (J^T(\theta_1)J(\theta_1))^{-1} J(\theta_1)^T (Y - F(t, \theta_1))$$

**✚ Limitation (inconvenient) :** L'inversion de la matrice  $J^T J$  peut poser problème (matrice singulière).

Pour éviter ce blocage, l'algorithme a été modifié par Levenberg-Marquardt:

$$\delta_\theta = (J^T(\theta_1)J(\theta_1) + \lambda I)^{-1} J(\theta_1)^T (Y - F(t, \theta_1))$$

Le paramètre  $\lambda$  joue le rôle d'un amortissement de la correction; il doit être ajusté à chaque pas de calcul. Dans les cas simples, on peut se contenter d'un amortissement constant, dont la valeur initiale a été proposée par Marquardt:

$$\lambda_0 = \tau \max(J^T J)$$

Avec  $\tau$  est un paramètre de gain à choisir convenablement !

**Exemple d'application :**

Exemple d'application : soit un système du 1er ordre dont désire connaître les paramètres caractéristiques (gain et constante de temps) par l'observation de la réponse indicielle.

Temps	0	1	2	3	4	5
Les mesures $y(t)$	0.05	0.45	0.59	0.64	0.64	0.69

Le modèle de la réponse indicielle est  $f(t, \theta) = K \left( 1 - e^{-\frac{t}{\tau}} \right)$  ave  $\theta = [K \quad \tau]^T$

La matrice du Jacobien est construite à partir des dérivées partielles de  $f$  par rapport à  $\theta$ .

$$\frac{\partial f(t, \theta)}{\partial K} = 1 - e^{-\frac{t}{\tau}}, \quad \frac{\partial f(t, \theta)}{\partial \tau} = K$$

Contrairement au cas d'un modèle linéaire/paramètres, les dérivées partielles dépendent des paramètres eux-mêmes.

- Il faut fixer une valeur initiale de  $K$  et  $\theta$  pour donner une valeur à la matrice Jacobienne.  
Valeurs initiales proposées :  $K_0 = 1; \tau_0 = 1$ .
- La matrice Jacobienne  $J(\theta_0) = [ \quad ]$

• Méthode de Gauss-Newton

La matrice Jacobienne est :  $J(\theta_0) = \begin{bmatrix} 0. & 0. \\ 0.63 & -0.37 \\ 0.86 & -0.27 \\ 0.95 & -0.15 \\ 0.98 & -0.07 \\ 0.99 & -0.03 \end{bmatrix}$  les valeurs du modèle :  $F(\theta_0) = \begin{bmatrix} 0. \\ 0.63 \\ 0.86 \\ 0.95 \\ 0.98 \\ 0.99 \end{bmatrix}$

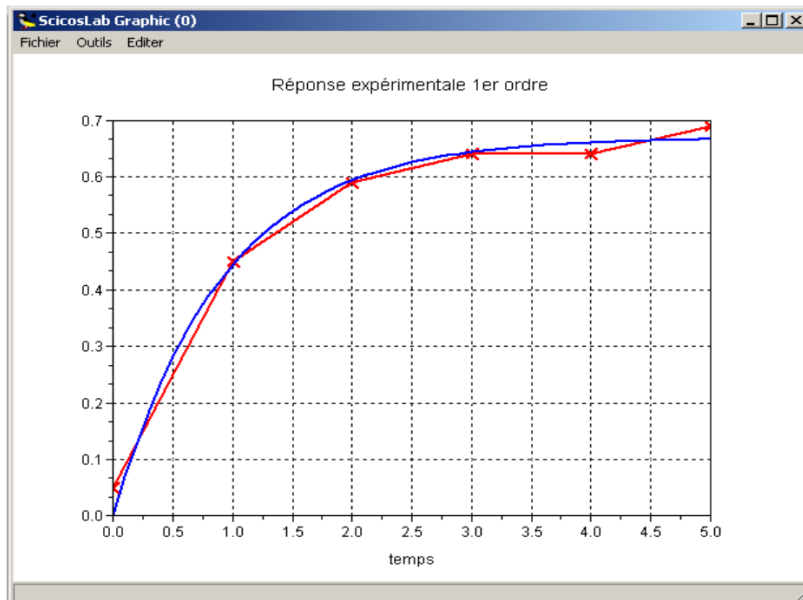
L'erreur quadratique cumulée est de :  $\mathcal{E} = 0.42$

L'incrément à faire sur les paramètres est :  $\delta_\theta = (J^T J)^{-1} J^T (Y - F(\theta_0)) = \begin{bmatrix} -0.33 \\ -0.06 \end{bmatrix}$

La nouvelle valeur des paramètres est :  $\theta_1 = \begin{bmatrix} 0.67 \\ 0.94 \end{bmatrix}$

Après 10 itérations , la valeur des paramètres converge vers :  $\theta_{10} = \begin{bmatrix} 0.67 \\ 0.92 \end{bmatrix}$

L'erreur quadratique cumulée est de :  $\mathcal{E} = 0.0035$



Attention : certaines valeurs initiales ne permettent pas à l'algorithme de converger (par exemple  $K = .1$  et  $tau = .1$  )!



# Chapitre 3

## Identification basée sur l'erreur de prédiction

### Introduction :

Dans ce chapitre, nous supposons que le modèle obtenu est un prédicteur, c'est à dire qu'il permet de calculer la sortie à l'instant  $i$  en fonction des entrées et des sorties réelles aux instants précédents  $u(i - k)$  et  $y(i - k)$ .

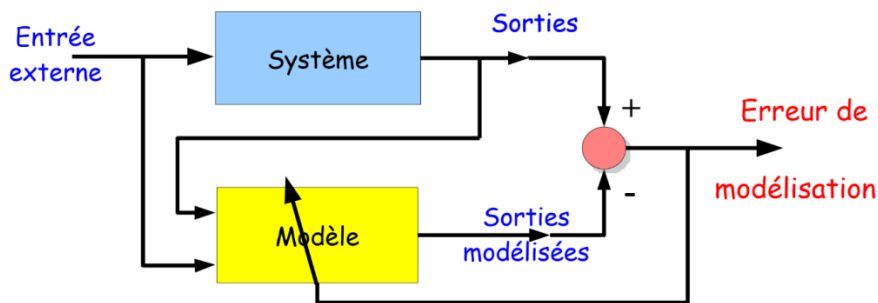
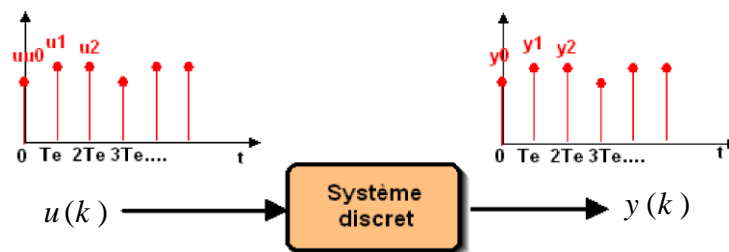


Figure 3.1. Principe d'identification basée sur l'erreur de prédiction

- Comme l'identification se fait en général par ordinateur, les mesures prélevées sont des échantillons, et le modèle est discret même si le système est continu.
- Le modèle doit traduire la relation entrée-sortie caractéristique du système.

### 1. Modèle de récurrence (équation de récurrence):

Soit le système suivant d'entrée  $u(k)$  et de sortie  $y(k)$



La relation entre la séquence d'entrée  $u(k)$  et la séquence de sortie  $y(k)$  est linéaire et définie par un ensemble de coefficients constants:

$$\begin{aligned}
 y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n) \\
 = b_0 u(k) + b_1 u(k-1) + b_2 u(k-2) + \dots + b_m u(k-m)
 \end{aligned}
 \tag{1}$$

**2. Fonction de transfert discrète :**

En appliquant la transformée en Z à chaque membre de l'équation, on obtient la fonction de transfert discrète du système : sachant que  $Z^{-1}(y(k-q)) \rightarrow +z^{-q}Y(z)$

$$\begin{aligned}
 Y(z) + a_1 z^{-1} Y(z) + a_2 z^{-2} Y(z) + \dots + a_n z^{-n} Y(z) \\
 = b_0 U(z) + b_1 z^{-1} U(z) + b_2 z^{-2} U(z) + \dots + b_m z^{-m} U(z) \\
 \Rightarrow (1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}) Y(z) = (b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}) U(z)
 \end{aligned}$$

La fonction de transfert devient alors :

$$G(z) = \frac{Y(z)}{U(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}}
 \tag{2}$$

**Exemple de modèle de récurrence discrète :**

**Exemple 1 :** Soit un système du premier ordre de fonction de transfert :

$$G(p) = \frac{Y(z)}{U(z)} = \frac{K}{P + \alpha}$$

Ce système correspond à une équation différentielle de la forme :

$$\dot{y}(t) + \alpha y(t) = Ku(t)$$

**On utilise l'approximation la dérivée par Euler :**

Pour une discrétisation régulière, on estime la dérivée d'une fonction  $f'$  par l'une des formules suivantes pour une période d'échantillonnage  $T_e$  :

$$1) f' = \frac{f(k+1) - f(k)}{T_e},$$

$$2) f' = \frac{f(k) - f(k-1)}{T_e},$$

$$3) f' = \frac{f(k+1) - f(k-1)}{2T_e}.$$

Si la fonction  $f$  est deux fois dérivable, on peut alors estimer sa dérivée seconde  $f''$  par la formule suivante :

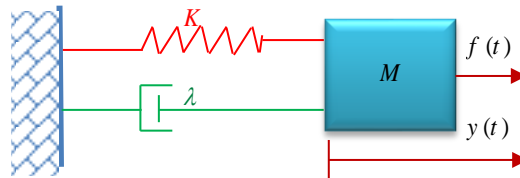
$$f'' = \frac{f(k+1) - 2f(k) + f(k-1)}{T_e^2}.$$

On obtient alors:  $\frac{y(k+1) - y(k)}{T_e} + \alpha y(k) = Ku(k)$

Le système discrétisé est décrit par la récurrence :  $y(k+1) = (1 - \alpha T_e) y(k) + K T_e u(k)$

Ou encore,  $y(k) = (1 - \alpha T_e) y(k-1) + K T_e u(k-1)$

**Exemple 2 : (Système masse- ressort- amortisseur) :** soit le système suivant



$f(t)$  la force exercée sur la masse  $M$  et  $y(t)$  la position de la masse par rapport à l'équilibre.

- 1- Déterminer l'équation différentielle de ce système
- 2- Calculer la sa fonction de transfert (en suppose que les CI sont nulles)  $G(p) = \frac{F(p)}{Y(p)}$
- 3- A l'aide de l'approximation de Euler (deux fois dérivable), déterminer l'équation de récurrence (aux différences) de ce système (en suppose que la période d'échantillonnage égale à  $t_{i+1} - t_i = T_e$ )

**Solution :**

**1)** En appliquant le principe fondamental de la dynamique sur la masse  $M$  soumise à l'action du ressort ( $-Ky(t)$ ), de l'amortisseur ( $-\lambda \frac{dy(t)}{dt}$ ) et à la force  $f(t)$  on obtient

l'équation suivante:  $M \frac{d^2y(t)}{dt^2} + \lambda \frac{dy(t)}{dt} + Ky(t) = f(t)$

**2)** la fonction de transfert  $G(p) = \frac{1}{Mp^2 + \lambda p + K}$

**3)** l'équation aux différences

On a  $M \frac{d^2y(t)}{dt^2} + \lambda \frac{dy(t)}{dt} + Ky(t) = f(t)$

$\Rightarrow M \frac{y(k+1) - 2y(k) + y(k-1)}{T_e^2} + \lambda \frac{y(k+1) - y(k)}{T_e} + Ky(k) = f(k)$

$\Rightarrow \left( \frac{M}{T_e^2} + \frac{\lambda}{T_e} \right) y(k+1) - \left( \frac{2M}{T_e^2} + \frac{\lambda}{T_e} + K \right) y(k) + \frac{M}{T_e^2} y(k-1) = f(k)$

**2. Classe de modèle linéaire (Structures de modèles)**

L'objectif de l'identification est de trouver les paramètres d'un modèle entrée-sortie. Mais, en pratique, la sortie mesurée des procédés est généralement bruitée. Cela est dû soit à l'effet des perturbations aléatoires agissant à différents endroits du procédé, soit à des bruits de mesure.

Il existe différentes structures de modèles linéaires. La différence se situe généralement dans la modélisation de la perturbation et la présence ou non d'une entrée externe.

Dans ce qui suit, on s'intéressera à quatre structures de modèles les plus utilisées : **ARX**, **ARMAX**, Modèle à erreur de sortie (**Output Error**), modèle de **Box-Jenkins**.

Le problème est d'identifier les paramètres d'un modèle entrée-sortie dont sa structure est représentée sur la figure suivante :

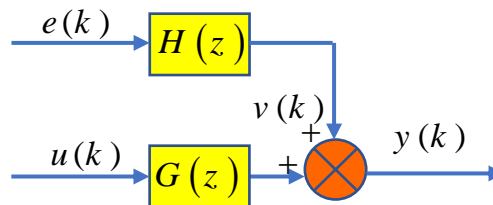


Figure 3.2 : Système linéaire avec bruit.

Avec

$y(k)$  : La sortie du système,

$u(k)$  : L'entrée du système (entrée déterministe)

$e(k)$  : Une perturbation de type bruit blanc (processus stochastique non prédictible de moyenne nulle).

- Les coefficients des polynômes en  $z^{-1}$  (dans la fonction de transfert) sont les paramètres à estimer.
- L'ensemble des effets des bruits et perturbations sont représentées par le signal stochastique  $v(k)$ , lui-même étant généré avec une dynamique  $H(z)$  par le signal également stochastique  $e(k)$  (bruit blanc de moyenne nulle "variable exogène"),

A partir de cette figure, on peut déduire le modèle entrée-sortie sous la forme suivante:

$$Y(z) = G(z)U(z) + H(z).E(z) \tag{3}$$

**➤ Prédicteur optimal** : Soit le modèle entrée-sortie dont sa structure est représentée sur la figure suivante :

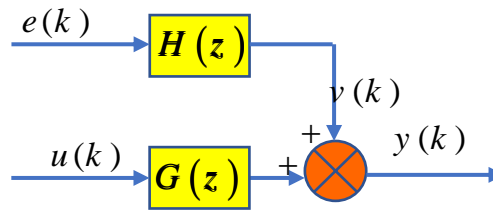


Figure 3.2 : Système linéaire avec bruit.

D'où le modèle entrée-sortie:

$$Y(z) = G(z)U(z) + H(z).E(z) \tag{4}$$

- On cherche à prédire la sortie  $y(k)$  à l'instant  $k$  sachant que la sortie et l'entrée sont connues aux instants précédents.
- On prend comme estimation de la sortie son espérance mathématique conditionnelle, soit :

$$\hat{y}(k | k-1) = E[y(k) | k-1] \tag{5}$$

Le prédictiveur est définie par :

$$\hat{y}(k | k-1) = H^{-1}(z)G(z)u(k) + (1 - H^{-1}(z))y(k) \tag{6}$$

Ou encore

$$H(z)\hat{y}(k | k-1) = G(z)u(k) + (H(z) - 1)y(k) \tag{7}$$

**✚ Erreur de prédiction**

L'erreur de prédiction peut être calculée (en utilisant (6)) comme suit :

$$\begin{aligned} \hat{y}(k | k-1) &= H^{-1}(z)G(z)u(k) + y(k) - H^{-1}(z)y(k) \\ \Rightarrow y(k) - \hat{y}(k | k-1) &= -H^{-1}(z)G(z)u(k) + H^{-1}(z)y(k) \\ e(k) = y(k) - \hat{y}(k | k-1) &= -H^{-1}(z)G(z)u(k) + H^{-1}(z)y(k) \end{aligned} \tag{8}$$

**2.1. Structure (modèle) ARX : (AR="Auto-Régressive", X="exogènes"**

Dans ce cas, le bruit  $e(k)$  perturbe la sortie brute de la fonction de transfert  $G(z)$  du système via

la dynamique.  $H(z) = \frac{1}{A(z)}$

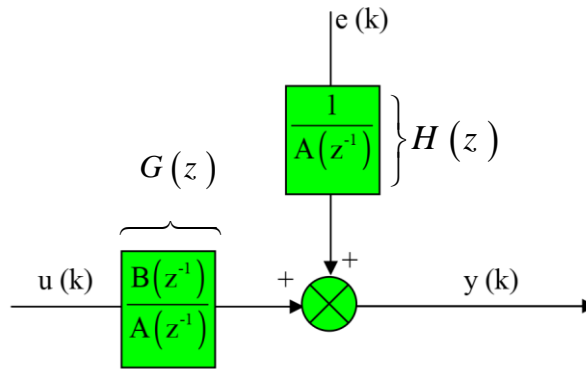


Figure 3.3. Modèle de structure ARX

Le modèle ARX est un modèle entrée-sortie de la forme générale suivante:

$$A(z)Y(z) = B(z)U(z) + E(z) \tag{9}$$

Avec

$$A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_nz^{-n}$$

$$B(z) = b_1z^{-1} + b_2z^{-2} + \dots + b_mz^{-m}$$

Ou encore à partir de la structure, on obtient :

$$y(k) = \frac{B(z)}{A(z)}u(k) + \frac{1}{A(z)}e(k)$$

D'où la structure du modèle ARX suivante :

$$A(z)y(k) = B(z)u(k) + e(k) \tag{10}$$

**Démonstration :** à partir de l'équation (4) en remplaçant  $H(z) = \frac{1}{A(z)}$  et  $G(z) = \frac{B(z)}{A(z)}$ ,

on obtient :

$$\begin{aligned} Y(z) &= G(z)U(z) + H(z).E(z) \\ \Rightarrow Y(z) &= \frac{B(z)}{A(z)}U(z) + \frac{1}{A(z)}E(z) \\ \Rightarrow A(z)Y(z) &= B(z)U(z) + E(z) \end{aligned}$$

Le modèle (9) peut être décrit sous la forme suivante :

$$\begin{aligned} &\underbrace{y(k) + a_1y(k-1) + a_2y(k-2) + \dots + a_ny(k-n)}_{\text{Partie Autoregressive}} \\ &= \underbrace{b_1u(k-1) + b_2u(k-2) + \dots + b_mu(k-m)}_{\text{Partie Exogène}} + \underbrace{e(k)}_{\text{Bruit}} \tag{11} \\ \Rightarrow y(k) + \sum_{i=1}^n a_i y(k-i) &= e(k) + \sum_{j=1}^m b_j u(k-j) \end{aligned}$$

En appliquant la transformée de Z au modèle (11), on obtient ;

$$Y(z) + a_1 z^{-1} Y(z) + a_2 z^{-2} Y(z) + \dots + a_n z^{-n} Y(z) = b_1 z^{-1} U(z) + b_2 z^{-2} U(z) + \dots + b_m z^{-m} U(z) + E(z)$$

Ou encore

$$\underbrace{(1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n})}_{A(z)} Y(z) = \underbrace{(b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m})}_{B(z)} U(z) + E(z)$$

Ce qui équivalent à :

$$A(z)Y(z) = B(z)U(z) + E(z)$$

### 2.1.1. Prédicteur optimal du modèle ARX

Sachant que l'équation générale du prédicteur optimale est

$$\hat{y}(k|k-1) = H^{-1}(z)G(z)u(k) + (1 - H^{-1}(z))y(k)$$

Pour un modèle ARX on prend  $H(z) = \frac{1}{A(z)}$  et  $G(z) = \frac{B(z)}{A(z)}$

Alors le prédicteur du modèle ARX devient

$$\hat{y}(k|k-1) = A(z) \frac{B(z)}{A(z)} u(k) + (1 - A(z))y(k)$$

$$\hat{y}(k|k-1) = B(z)u(k) + (1 - A(z))y(k) \tag{12}$$

**Démonstration :** L'équation (11) permet d'écrire la relation suivante :

$$\underbrace{y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n)}_{\text{Partie Autoregressive}} = \underbrace{b_1 u(k-1) + b_2 u(k-2) + \dots + b_m u(k-m)}_{\text{Partie Exogène}} + \underbrace{e(k)}_{\text{Bruit}}$$

Cette formule peut être réécrite comme suit :

$$y(k) = \underbrace{(b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m})}_{B(z)} u(k) + \underbrace{(-a_1 z^{-1} - a_2 z^{-2} - \dots - a_n z^{-n})}_{1-A(z)} y(k) + e(k)$$

$$y(k) = B(z)u(k) + (1 - A(z))y(k) + e(k) \tag{13}$$

**NB : A partir de (10), on peut écrire**

$$\begin{aligned} A(z)y(k) + y(k) &= B(z)u(k) + y(k) + e(k) \\ \Rightarrow y(k) &= B(z)u(k) + (1 - A(z))y(k) + e(k) \\ &= (13) \end{aligned}$$

En introduisant cette équation (13) dans l'expression de l'espérance, on aura:

$$\hat{y}(k|k-1) = E[B(z)u(k) + (1 - A(z))y(k) + e(k)]$$

Comme  $e(k)$  est considéré comme un bruit blanc d'espérance nulle c'est-à-dire  $E[e(k)] = 0$

On aura donc  $\hat{y}(k|k-1) = B(z)u(k) + (1 - A(z))y(k)$  (la même que (12))

### 2.1.2. Erreur de prédicteur du modèle ARX

On à l'expression générale de l'erreur de prédiction

$$e(k) = -H^{-1}(z)G(z)u(k) + H^{-1}(z)y(k)$$

Dans ce cas on prend  $H(z) = \frac{1}{A(z)}$  et  $G(z) = \frac{B(z)}{A(z)}$ , on aura alors l'expression suivante :

$$e(k) = -B(z)u(k) + A(z)y(k) \tag{14}$$

### 2.1.3. Prédicteur optimal du modèle ARX sous forme d'une régression linéaire

Le prédicteur optimal du modèle ARX est particulièrement simple à calculer car il définit une régression linéaire que l'on écrit généralement sous la forme :

$$\hat{y}(k) = \varphi(k)\theta \tag{15}$$

Avec

$$\varphi(k) = [-y(k-1) \quad -y(k-2) \quad \dots \quad -y(k-n) \quad u(k-1) \quad u(k-2) \quad \dots \quad u(k-m)]$$

$$\theta = [a_1 \quad a_2 \quad \dots \quad a_n \quad b_1 \quad b_2 \quad \dots \quad b_m]^T$$

$\varphi(k)$  est le régresser

**Démonstration :** à partir de (11) permet d'écrire la relation suivante :



$$y(k) = -a_1y(k-1) - a_2y(k-2) - \dots - a_ny(k-n) + b_1u(k-1) + b_2u(k-2) + \dots + b_mu(k-m) + e(k)$$

Avec  $e(k) = y(k) - \hat{y}(k|k-1)$ , on trouve alors,

$$y(k) = -a_1y(k-1) - a_2y(k-2) - \dots - a_ny(k-n) + b_1u(k-1) + b_2u(k-2) + \dots + b_mu(k-m) + y(k) - \hat{y}(k|k-1)$$

Ce qui équivaut à

$$\hat{y}(k|k-1) = -a_1y(k-1) - a_2y(k-2) - \dots - a_ny(k-n) + b_1u(k-1) + b_2u(k-2) + \dots + b_mu(k-m)$$

Cette expression peut être exprimée sous une formule matricielle comme suit

$$\hat{y}(k) = \underbrace{\begin{bmatrix} -y(k-1) & -y(k-2) & \dots & -y(k-n) & u(k-1) & u(k-2) & \dots & u(k-m) \end{bmatrix}}_{\varphi} \begin{bmatrix} a_1 \\ a_1 \\ \vdots \\ a_n \\ b_1 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}_{\theta}$$

Qui est équivalent à  $\hat{y}(k) = \varphi(k)\theta$

### 2.1.4. Critère quadratique du prédicteur ARX

L'expression de l'erreur de prédiction est la suivante :  $e(k) = y(k) - \hat{y}(k|k-1)$

L'erreur quadratique s'exprime donc par :

$$J = \sum_{k=1}^N [e(k)]^2 = \sum_{k=1}^N [y(k) - \hat{y}(k|k-1)]^2 = \sum_{k=1}^N [y(k) - \varphi^T(k)\theta]^2 \tag{16}$$

$\theta$  est obtenu pour  $\frac{dJ}{d\theta} = 0$

### ✚ Méthode des moindres carrés :

Le problème est de chercher le vecteur  $\theta$  des paramètres qui minimise l'erreur quadratique  $J$ . La solution consiste à annuler la dérivée première du critère quadratique.

En supposant que l'inverse existe, on montre que :

$$\hat{\theta}(N) = \left( \sum_{k=1}^N [\varphi^T(k) \varphi(k)] \right)^{-1} \sum_{k=1}^N [\varphi^T(k) y(k)]$$

On a le critère quadratique est de la forme suivante :

Pour  $k = 2, \dots, n$ , on pose  $Y_N = \begin{bmatrix} y(1) \\ \vdots \\ y(n) \end{bmatrix}$

Le critère de moindre carrée (16) devient

$$J = \sum_{k=1}^N [y(k) - \varphi(k)\theta]^2 = (Y_N - \Phi_N \theta)^T (Y_N - \Phi_N \theta) \quad (17)$$

$\theta$  est obtenu pour  $\frac{dJ}{d\theta} = 0 \Rightarrow \frac{d}{d\theta} = -2\Phi_N^T (Y_N - \Phi_N \theta)^T = 0 \Rightarrow \Phi_N^T Y_N - \Phi_N^T \Phi_N \theta = 0$

$$\hat{\theta} = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T Y_N \quad (18)$$

**Exemple 3.1.** Système du premier ordre soumis à un échelon d'amplitude 1

Soit les valeurs relevées pour  $y(k)$  sont :

$$Y(k) = [0 \ 0.4 \ 0.7 \ 0.95 \ 1 \ 0.98 \ 0.97 \ 0.99 \ 1.02]'$$

Le modèle choisi est :  $y(k) = -a_1 y(k-1) + b_1 u(k-1)$  avec  $\theta = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}$

- L'entrée est un signal échelon de même taille que la mesure  $Y(k)$  c'est-à-dire

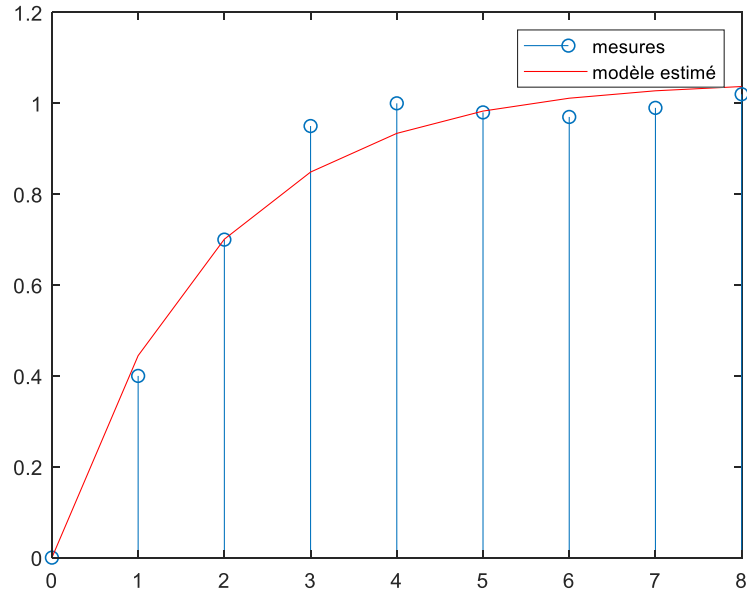
$u(k) = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]'$  ou simplement  $U = \text{ones}(\text{length}(Y), 1)$  ;

`% construction de la matrice phi(k) puis calcul de la pseudo-inverse theta`

```
for i = 2:size(Y,1)
phi(i,:) = [-Y(i-1) U(i-1)];
end
theta = inv(phi*phi')*phi*Y
```

on aura donc le résultat suivant :  $\theta = \begin{bmatrix} -0.577 \\ 0.444 \end{bmatrix}$

Le modèle est représenté en rouge sur la figure suivante.



**2.2. Structure (modèle) ARMAX : (ARMAX= AutoRégressif à Moyenne Ajustée et à variables eXogènes)**

on offre comparativement à la structure ARX un degré de liberté supplémentaire pour modéliser la dynamique des perturbations  $e(k)$  en les faisant intervenir sur le système avec la fonction de

$$\text{transfert } H(z) = \frac{V(z)}{E(z)} = \frac{C(z)}{A(z)}$$

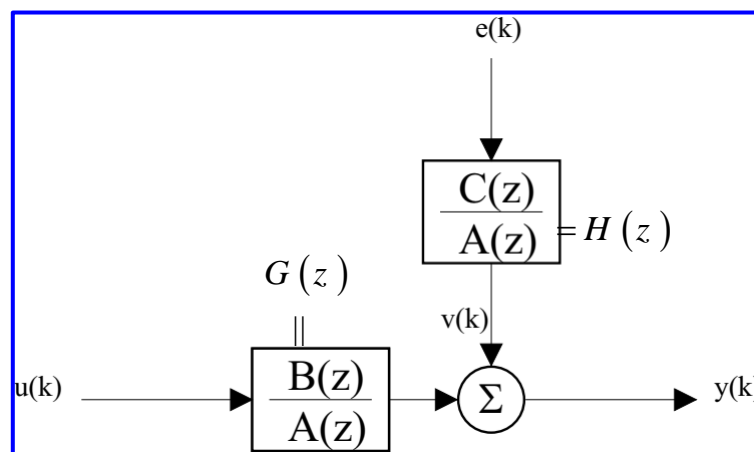


Figure 3.4. Modèle de structure ARMAX

Le modèle ARMAX est un modèle entrée-sortie de la forme générale suivante:

$$A(z)Y(z) = B(z)U(z) + C(z)E(z) \tag{19}$$

Avec

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}$$

$$B(z) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}$$

$$C(z) = c_1 z^{-1} + c_2 z^{-2} + \dots + c_j z^{-j}$$

Ou encore à partir de la structure, on obtient :

$$y(k) = \frac{B(z)}{A(z)}u(k) + \frac{C(z)}{A(z)}e(k)$$

D'où la structure du modèle ARMAX suivante :

$$A(z)y(k) = B(z)u(k) + C(z)e(k) \tag{20}$$

**Démonstration :** Le modèle (19) peut être décrit sous la forme suivante :

$$\begin{aligned} & \underbrace{y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n)}_{\text{Partie Autoregressive}} \\ &= \underbrace{b_1 u(k-1) + b_2 u(k-2) + \dots + b_m u(k-m)}_{\text{Partie Exogène}} + \underbrace{e(k) + c_1 e(k-1) + c_2 e(k-2) + \dots + c_j e(k-j)}_{\text{Bruit}} \tag{21} \\ \Rightarrow y(k) + \left( \sum_{i=1}^n a_i y(k-i) \right) &= \left( \sum_{j=1}^m b_j u(k-j) \right) + e(k) + \left( \sum_{l=1}^j c_l e(k-l) \right) \end{aligned}$$

En appliquant la transformée de Z au modèle (21), on obtient ;

$$\begin{aligned} Y(z) + a_1 z^{-1} Y(z) + a_2 z^{-2} Y(z) + \dots + a_n z^{-n} Y(z) \\ = b_1 z^{-1} U(z) + b_2 z^{-2} U(z) + \dots + b_m z^{-m} U(z) \\ + c_1 z^{-1} C(z) + c_2 z^{-2} C(z) + \dots + c_j z^{-j} C(z) \end{aligned}$$

Ou encore

$$\begin{aligned} \underbrace{(1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n})}_{A(z)} Y(z) &= \\ \underbrace{(b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m})}_{B(z)} U(z) + \\ \underbrace{(c_1 z^{-1} + c_2 z^{-2} + \dots + c_j z^{-j})}_{C(z)} E(z) \end{aligned}$$

Ce qui équivalent à :  $A(z)Y(z) = B(z)U(z) + C(z)E(z)$

Ce modèle est proche du modèle ARX, il s'utilise dans les mêmes cas. Il permet de créer un modèle de bruit un peu réaliste.

### 2.2.1. Prédicteur optimal du modèle ARMAX

Sachant que l'équation générale du prédicteur optimal est

$$\hat{y}(k|k-1) = H^{-1}(z)G(z)u(k) + (1 - H^{-1}(z))y(k)$$

Pour un modèle ARMAX on prend  $H(z) = \frac{C(z)}{A(z)}$  et  $G(z) = \frac{B(z)}{A(z)}$

Alors le prédicteur du modèle ARMAX devient

$$\hat{y}(k|k-1) = \frac{B(z)}{C(z)}u(k) + \left(1 - \frac{A(z)}{C(z)}\right)y(k)$$

Après simplification, on aura le prédicteur suivant :

$$\hat{y}(k|k-1) = \frac{B(z)}{C(z)}u(k) + \left(1 - \frac{A(z)}{C(z)}\right)y(k)$$

$$C(z)\hat{y}[k|k-1] = B(z)u(k) + [C(z) - A(z)]y(k) \tag{22}$$

### 2.2.2. Erreur de prédicteur du modèle ARMAX

On a l'expression générale de l'erreur de prédiction

$$e(k) = -H^{-1}(z)G(z)u(k) + H^{-1}(z)y(k)$$

Dans ce cas on prend  $H(z) = \frac{C(z)}{A(z)}$  et  $G(z) = \frac{B(z)}{A(z)}$ , on aura alors l'expression suivante :

$$e(k) = -\frac{B(z)}{C(z)}u(k) + \frac{A(z)}{C(z)}y(k) \tag{23}$$

### 2.2.3. Prédicteur optimal du modèle ARMAX sous forme d'une régression linéaire

Le prédicteur optimal du modèle ARMAX est particulièrement simple à calculer car il définit une régression linéaire que l'on écrit généralement sous la forme :

$$\hat{y}(k) = \varphi(k)\theta \tag{24}$$

Avec

$$\varphi(k) = [-y(k-1) \quad \dots \quad -y(k-n) \quad u(k-1) \quad \dots \quad u(k-m) \quad e(k-1) \quad e(k-2) \quad \dots \quad \hat{e}(k-j)]$$

$$\theta = [a_1 \quad a_2 \quad \dots \quad a_n \quad b_1 \quad b_2 \quad \dots \quad b_m \quad c_1 \quad c_2 \quad \dots \quad c_j]^T$$

### 2.2.4. Critère quadratique du prédicteur ARMAX

L'expression de l'erreur de prédiction est la suivante :  $e(k) = y(k) - \hat{y}(k|k-1)$

L'erreur quadratique s'exprime donc par :

$$J = \sum_{k=1}^N [e(k)]^2 = \sum_{k=1}^N [y(k) - \hat{y}(k|k-1)]^2 = \sum_{k=1}^N [y(k) - \varphi(k)\theta]^2$$

$\theta$  est obtenu pour  $\frac{dJ}{d\theta} = 0$

**✚ Méthode des moindres carrés :**

Pour  $k = 1, \dots, N$ , on pose  $Y_N = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}$

$$\varphi_N(k) = [-y(k-1) \quad \dots \quad -y(k-n) \quad u(k-1) \quad \dots \quad u(k-m) \quad e(k-1) \quad e(k-2) \quad \dots \quad \hat{e}(k-j)]$$

Le critère de moindre carrée est

$$\hat{\theta} = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T Y_N$$

**2.3. Structure (modèle) OE : (OE= Output Error)**

Le modèle OE (Output Error) Calcule l'estimation de l'erreur de prédiction d'un modèle d'erreur de sortie. Sa forme générale est :

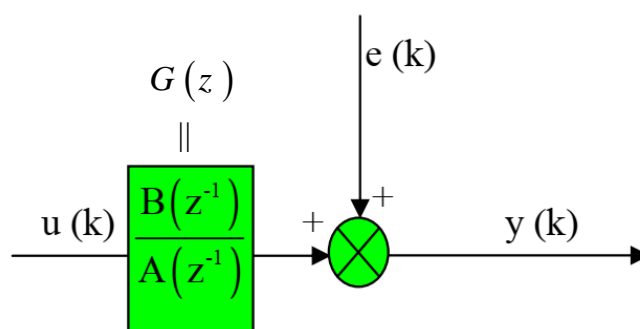


Figure 3.5. Modèle de structure OE

Le modèle ARMAX est un modèle entrée-sortie de la forme générale suivante:

$$A(z)Y(z) = B(z)U(z) + A(z)E(z) \tag{25}$$

Avec

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}$$

$$B(z) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}$$

Ou encore à partir de la structure, on obtient :

$$y(k) = \frac{B(z)}{A(z)} u(k) + e(k)$$

D'où la structure du modèle OE suivante :

$$A(z)y(k) = B(z)u(k) + A(z)e(k) \tag{26}$$

**Démonstration :** Le modèle (26) peut être décrit sous la forme suivante :

$$\begin{aligned} & \underbrace{y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n)}_{\text{Partie Autoregressive}} \\ & = \underbrace{b_1 u(k-1) + b_2 u(k-2) + \dots + b_m u(k-m)}_{\text{Partie Exogène}} + \underbrace{e(k) + a_1 e(k-1) + a_2 e(k-2) + \dots + a_n e(k-n)}_{\text{Bruit}} \tag{27} \\ \Rightarrow y(k) + \sum_{i=1}^n a_i y(k-i) &= \sum_{j=1}^m b_j u(k-j) + \sum_{i=1}^n a_i e(k-i) \end{aligned}$$

En appliquant la transformée de Z au modèle (27), on obtient ;

$$\begin{aligned} Y(z) + a_1 z^{-1} Y(z) + a_2 z^{-2} Y(z) + \dots + a_n z^{-n} Y(z) \\ = b_1 z^{-1} U(z) + b_2 z^{-2} U(z) + \dots + b_m z^{-m} U(z) \\ E(z) + a_1 z^{-1} C(z) + a_2 z^{-2} C(z) + \dots + a_n z^{-n} C(z) \end{aligned}$$

Ou encore

$$\begin{aligned} & \underbrace{(1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n})}_{A(z)} Y(z) = \\ & \underbrace{(b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m})}_{B(z)} U(z) + \\ & \underbrace{(1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n})}_{A(z)} E(z) \end{aligned}$$

Ce qui équivaut à :

$$A(z)Y(z) = B(z)U(z) + C(z)E(z)$$

### 2.3.1. Prédicteur optimal du modèle OE sous forme d'une régression linéaire

Le prédicteur optimal du modèle OE est particulièrement simple à calculer car il définit une régression linéaire que l'on écrit généralement sous la forme :

$$\hat{y}(k) = \varphi(k)\theta \tag{24}$$

Avec

$$\varphi(k) = [-y(k-1) \quad \dots \quad -y(k-n) \quad u(k-1) \quad \dots \quad u(k-m) \quad e(k-1) \quad e(k-2) \quad \dots \quad e(k-n)]$$

$$\theta = [a_1 \quad a_2 \quad \dots \quad a_n \quad b_1 \quad b_2 \quad \dots \quad b_m \quad a_1 \quad a_2 \quad \dots \quad a_n]^T$$

### 2.3.2. Critère quadratique du prédicteur OE

#### ✚ Méthode des moindres carrés :

Pour  $k = 1, \dots, N$ , on pose  $Y_N = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}$

$$\varphi_N(k) = [-y(k-1) \quad \dots \quad -y(k-n) \quad u(k-1) \quad \dots \quad u(k-m) \quad e(k-1) \quad e(k-2) \quad \dots \quad e(k-n)]$$

Le critère de moindre carrée est

$$\hat{\theta} = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T Y_N$$

### 2.4. Structure (modèle) Box-Jenkins

Le modèle de Box-Jenkins est le modèle complet par excellence, où la dynamique de

l'entrée est différente à celle du bruit. Ici  $H(z) = \frac{V(z)}{E(z)} = \frac{C(z)}{A(z)}$

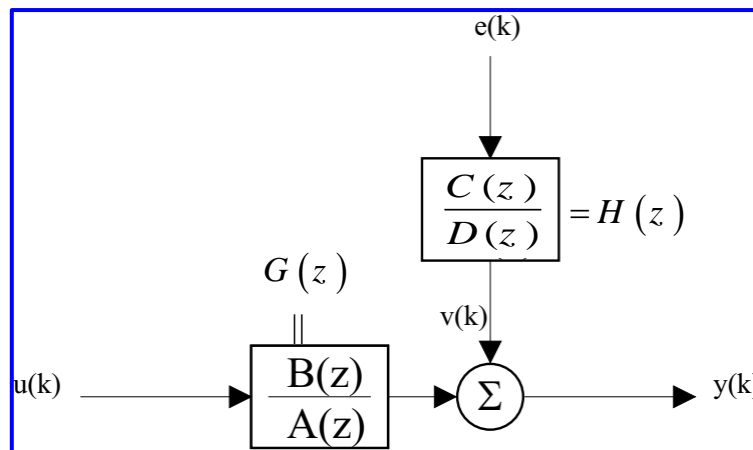


Figure 3.4. Modèle de structure BoxJ

Le modèle BoxJ est un modèle entrée-sortie de la forme générale suivante:



$$Y(z) = \frac{B(z^{-1})}{\underbrace{A(z^{-1})}_{G(z^{-1})}} U(z) + \frac{C(z^{-1})}{\underbrace{D(z^{-1})}_{H(z^{-1})}} E(z) \quad (25)$$

$$A(z^{-1})D(z^{-1})Y(z) = D(z^{-1})B(z^{-1})U(z) + A(z^{-1})C(z^{-1})E(z)$$

Avec

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}$$

$$B(z) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_m z^{-m}$$

$$C(z) = c_1 z^{-1} + c_2 z^{-2} + \dots + c_j z^{-j}$$

$$D(z) = 1 + d_1 z^{-1} + d_1 z^{-2} + \dots + d_l z^{-l}$$

Où encore à partir de la structure, on obtient :

$$y(k) = \frac{B(z)}{A(z)} u(k) + \frac{C(z)}{D(z)} e(k)$$

D'où la structure du modèle Boxj suivante :

$$A(z)D(z)y(k) = B(z)D(z)u(k) + A(z)C(z)e(k) \quad (26)$$

### 2.4.1. Prédicteur optimal du modèle BoxJ

Sachant que l'équation générale du prédicteur optimale est

$$\hat{y}(k|k-1) = H^{-1}(z)G(z)u(k) + (1 - H^{-1}(z))y(k)$$

Pour un modèle BoxJ on prend  $H(z) = \frac{C(z)}{D(z)}$  et  $G(z) = \frac{B(z)}{A(z)}$

La structure du modèle BoxJ conduit à un prédicteur non linéaire par rapport au vecteur de paramètres :

$$\hat{y}(k|k-1) = \frac{D(z)}{C(z)} \frac{B(z)}{A(z)} u(k) + \left(1 - \frac{D(z)}{C(z)}\right) y(k) \quad (27)$$

L'erreur de prédiction pour le modèle BJ s'écrit comme suit :

La forme générale :

$$e(k) = y(k) - \hat{y}(k) = -H^{-1}(z)G(z)u(k) + H^{-1}(z)y(k) - H^{-1}(z)(y(k) - G(z)u(k))$$

Pour un modèle BoxJ on prend  $H(z) = \frac{C(z)}{D(z)}$  et  $G(z) = \frac{B(z)}{A(z)}$

On trouve l'erreur de prédiction comme suit

$$\varepsilon = y(k) - \hat{y}(k) = \frac{D(z)}{C(z)} \left( y(k) - \frac{B(z)}{A(z)} u(k) \right)$$

Soient :  $v(k) = \frac{B(z^{-1})}{A(z^{-1})} u(k)$  et  $z(k) = y(k) - v(k)$

On peut montrer que

$$\hat{y}(k) = (C(z^{-1}) - 1)e(k) + B(z^{-1})u(k) + (1 - D(z^{-1}))z(k) + (1 - A(z^{-1}))v(k)$$

Dans ce cas le prédicteur est représenté comme suit :

$$\hat{y}(k) = \varphi^T(k)\theta$$

avec

$$\varphi(k) = [u(k-1) \quad \dots \quad u(k-n_u) \quad e(k-1) \quad \dots \quad e(k-n_e) \quad -z(k-1) \quad \dots] \\ [-z(k-n_z) \quad -v(k-1) \quad \dots \quad -v(k-n_v)]$$