

# Observation

FRÉDÉRIC ROTELLA  
Ecole Nationale d'Ingénieurs de Tarbes

rotella@enit.fr

## Table des matières

<b>1</b>	<b>Principe des observateurs</b>	<b>3</b>
<b>2</b>	<b>Utilisation en boucle fermée</b>	<b>5</b>
2.1	Principe de séparation . . . . .	6
2.2	Robustesse . . . . .	6
2.3	Forme RST du régulateur-observateur . . . . .	7
<b>3</b>	<b>Détermination du gain de l'observateur</b>	<b>8</b>
3.1	Cas monosortie ( $l = 1$ ) . . . . .	8
3.1.1	Calcul direct . . . . .	8
3.1.2	Formule de Bass et Gura . . . . .	9
3.2	Cas multisortie ( $l > 1$ ) . . . . .	10
3.3	Méthode LTR . . . . .	12
<b>4</b>	<b>Systèmes discrets</b>	<b>14</b>
4.1	Observateurs prédicteurs ou correcteurs . . . . .	14
4.1.1	Observateur - prédicteur . . . . .	15
4.1.2	Observateur - correcteur . . . . .	15
4.1.3	Remarques sur les observateurs de systèmes discrets . . . . .	16
<b>5</b>	<b>Simplification des observateurs</b>	<b>17</b>
5.1	Observateur d'ordre réduit . . . . .	18
5.1.1	Construction . . . . .	18
5.2	Observation d'une fonction linéaire de l'état . . . . .	20
5.2.1	Réduction de l'ordre de l'observateur . . . . .	20
5.2.2	Structure générale . . . . .	22
5.2.3	Détermination de l'observateur . . . . .	23

5.3	Utilisation en boucle fermée et robustesse . . . . .	25
5.3.1	Principe de séparation . . . . .	25
5.3.2	Forme RST . . . . .	26
5.3.3	Robustesse . . . . .	26
<b>6</b>	<b>Applications des observateurs</b>	<b>26</b>
6.1	Observateurs à entrées inconnues . . . . .	27
6.1.1	Estimation des entrées inconnues . . . . .	29
6.1.2	Mesures dépendant des entrées inconnues . . . . .	29
6.2	Détection de défauts . . . . .	30
<b>7</b>	<b>Annexe : Algorithme de Luenberger</b>	<b>31</b>
7.1	Forme canonique commandable . . . . .	32
7.2	Forme canonique observable . . . . .	34
<b>8</b>	<b>Annexe B : Formes canoniques de matrices</b>	<b>35</b>
8.1	Forme de Smith . . . . .	35
8.2	Décomposition en valeurs singulières . . . . .	37
<b>9</b>	<b>Annexe C : Inverses généralisées et systèmes linéaires</b>	<b>40</b>
9.1	Inverse généralisée . . . . .	40
9.2	Application aux systèmes linéaires . . . . .	41
9.3	Pseudo-inverse . . . . .	42
9.3.1	Notion de pseudo-inverse . . . . .	42
9.3.2	Formes explicites de la pseudo-inverse . . . . .	42
9.3.3	Quelques propriétés de la pseudo-inverse . . . . .	43
9.3.4	Algorithme de Greville-Albert-Sittler . . . . .	44
9.3.5	Application aux systèmes linéaires incompatibles . . . . .	44
9.3.6	Résolution d'un système linéaire . . . . .	44
	<b>Références</b>	<b>44</b>

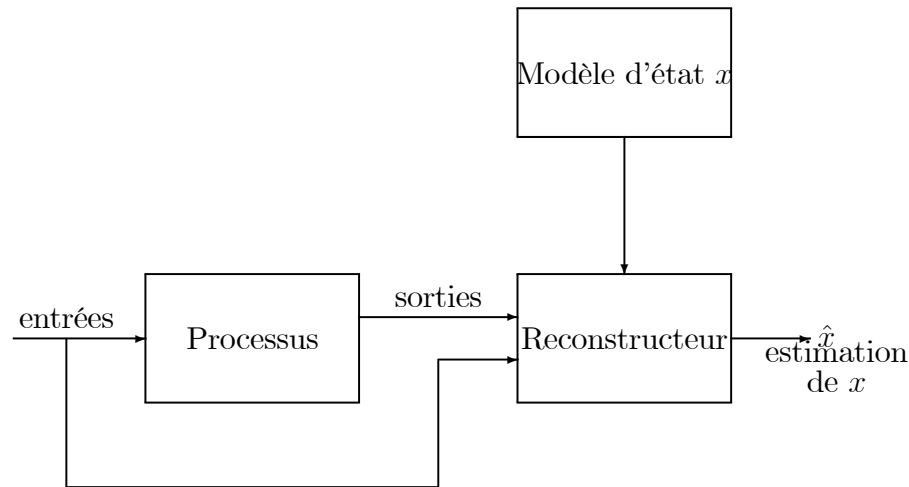


FIG. 1 – Principe d'un estimateur.

De nombreuses méthodes de commande des processus utilisent le principe du retour d'état (commande optimale, découplage, placement de pôles, ...). Comme dans la plupart des cas, les seules grandeurs accessibles du système sont les variables d'entrée et de sortie, il est nécessaire, à partir de ces informations, de reconstruire l'état du modèle choisi pour élaborer la commande. Un **reconstructeur d'état** ou **estimateur** est un système (figure 1) ayant comme entrées les entrées et les sorties du processus réel et dont la sortie est une **estimation** de l'état de ce processus.

Sous l'hypothèse de linéarité du modèle du processus, la structure de base de l'estimateur est toujours la même, mais sa réalisation dépendra du contexte choisi : continu ou discret, déterministe ou stochastique.

Dans le cas où ce modèle est un modèle déterministe, le reconstructeur d'état sera appelé **observateur**, ce qui sera l'objet de ce cours. Dans le cas de systèmes bruités, où interviennent des phénomènes aléatoires, nous parlerons alors de **filtre** et cela constitue l'objet du cours sur le filtre de Kalman.

## 1 Principe des observateurs

Soit un système continu décrit par l'équation d'état (déterministe car non bruitée) :

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \quad (1)$$

où  $u(t)$ ,  $y(t)$  et  $x(t)$  sont des vecteurs de dimension  $m$ ,  $l$  et  $n$  et représentent respectivement la commande, la sortie (mesurée) et l'état du système. Les matrices  $A$ ,  $B$ ,  $C$  et  $D$  sont des matrices constantes de dimensions convenables.

Comme l'état n'est en général pas accessible, l'objectif d'un observateur consiste, en vue de réaliser une commande par retour d'état, d'estimer cet état par une variable que nous noterons  $\hat{x}(t)$ . Cette estimation est réalisée par un système dynamique dont la sortie sera précisément  $\hat{x}(t)$  et l'entrée sera constituée de l'ensemble des informations disponibles, c'est-à-dire  $u(t)$  et  $y(t)$ . La structure d'un observateur est de la forme :

$$\begin{aligned}\dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) + L(y(t) - \hat{y}(t)), \\ \hat{y}(t) &= C\hat{x}(t) + Du(t),\end{aligned}$$

où apparaît clairement le terme correctif en fonction de l'erreur de reconstruction de la sortie,  $y(t) - \hat{y}(t)$ , et le gain de correction,  $L$ , appelé gain de l'observateur est à déterminer. Cette structure peut être écrite sous la forme :

$$\dot{\hat{x}}(t) = (A - LC)\hat{x}(t) + (B - LD)u(t) + Ly(t).$$

Si on considère l'erreur d'estimation :

$$\tilde{x}(t) = x(t) - \hat{x}(t),$$

on obtient :

$$\dot{\tilde{x}}(t) = (A - LC)\tilde{x}(t),$$

ce qui conduit à l'évolution de l'erreur d'estimation à partir de la condition initiale  $\tilde{x}(0) = x(0) - \hat{x}(0)$ , qui est non nulle de façon générale parce que l'état est *a priori* inaccessible :

$$\tilde{x}(t) = \exp[(A - LC)t] \tilde{x}(0).$$

Pour que l'observateur soit utilisable il est nécessaire que cette erreur tende vers 0 lorsque  $t$  augmente. Lorsque cette propriété est satisfaite l'observateur est dit asymptotique, mais il est évident que c'est une propriété nécessaire au fonctionnement correct d'un observateur. En conséquence il faudra choisir  $L$  telle que les valeurs propres de la matrice  $A - LC$  soient toutes à partie réelle strictement négative.

**Théorème 1** *Les valeurs propres de  $A - LC$  peuvent être fixées arbitrairement si et seulement si la paire  $(A, C)$  est observable, i.e. :*

$$\text{rang} \left( \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{(n-1)} \end{bmatrix} \right) = n.$$

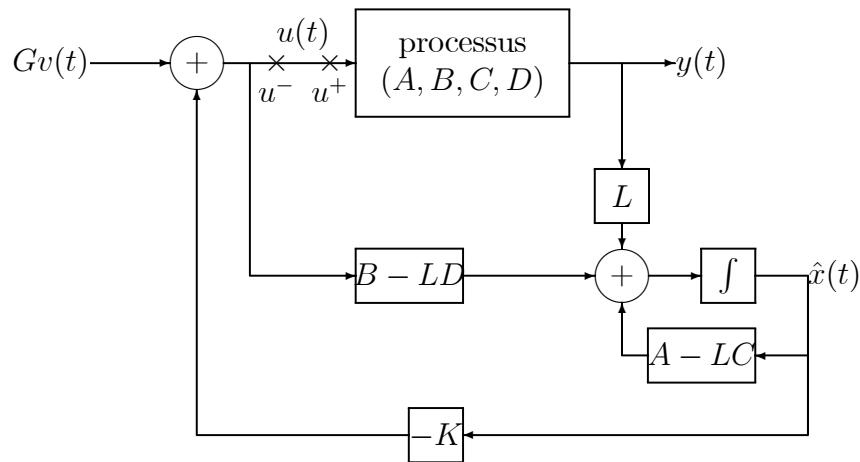


FIG. 2 – Commande par régulateur-observateur.

Dans ce cas, une grande liberté est laissée au choix des valeurs propres, mais en pratique on choisit une dynamique d'erreur plus rapide que celle du processus dans le cas d'une observation en boucle ouverte ou que celle désirée en boucle fermée. Cependant on ne peut les prendre infiniment grandes pour deux raisons essentielles : on ne peut utiliser que des gains réalisables et l'augmentation de la bande passante du reconstruteur ne permet plus de négliger les bruits qui deviennent prépondérants en haute fréquence.

## 2 Utilisation en boucle fermée

En supposant que tout l'état soit accessible, la commande par retour d'état que l'on veut mettre en œuvre est de la forme :

$$u(t) = Gv(t) - Kx(t),$$

où  $v(t)$  est la nouvelle entrée,  $K$  est le gain de retour d'état qui est défini par la structure désirée pour le système bouclé (*i.e.*  $A - BK$ ), et  $G$  un gain de prébouclage qui est calculé en général en fonction du gain statique désiré.

Comme l'état n'est pas accessible, la commande réellement mis en œuvre devient :

$$u(t) = Gv(t) - K\hat{x}(t), \quad (2)$$

et nous allons voir les conséquences d'utiliser la structure de commande par régulateur-observateur décrite dans la figure 2.

## 2.1 Principe de séparation

L'utilisation de (2) dans (1) conduit à l'équation d'état du système en boucle fermée :

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\tilde{x}}(t) \end{bmatrix} = \begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix} \begin{bmatrix} x(t) \\ \tilde{x}(t) \end{bmatrix} + \begin{bmatrix} BG \\ 0 \end{bmatrix} v(t),$$

$$y(t) = \begin{bmatrix} C - DK & DK \end{bmatrix} \begin{bmatrix} x(t) \\ \tilde{x}(t) \end{bmatrix} + DGv(t).$$

De ces relations on déduit deux conséquences :

- le principe de séparation : les dynamiques du système commandé par un régulateur-observateur en boucle fermée, *i.e.* les valeurs propres de la matrice :

$$\begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix}$$

sont constituées de la réunion de celles désirées en boucle fermée et celles de l'observateur. Ainsi on peut régler de façon indépendante le problème de la régulation et le problème de l'observation ;

- la non observabilité et non commandabilité de l'observateur : si l'on calcule le transfert obtenu à l'aide de la structure régulateur-observateur, on arrive à :

$$\begin{aligned} T_{\text{BF}}(p) &= DG + \begin{bmatrix} C - DK & DK \end{bmatrix} \times \\ &\quad \begin{bmatrix} pI - A + BK & -BK \\ 0 & pI - A + LC \end{bmatrix}^{-1} \begin{bmatrix} BG \\ 0 \end{bmatrix}, \\ &= DG + (C - DK)(pI - A + BK)^{-1}BG, \end{aligned}$$

c'est-à-dire le transfert que l'on aurait obtenu si on avait utilisé le retour d'état avec accès complet à l'état.

## 2.2 Robustesse

Toute utilisation d'un régulateur dynamique implique une dégradation des marges de stabilité, et un régulateur-observateur ne fait pas exception à cette règle. Nous allons voir dans ce paragraphe comment juger de la robustesse en stabilité du régulateur-observateur construit ce qui sera un guide dans le choix du gain de l'observateur. Si les marges obtenues, par un tracé fréquentiel, ne sont pas satisfaisantes il conviendra de modifier la dynamique de l'observateur.

Le transfert à calculer est celui obtenu lorsque l'on ouvre la boucle juste avant la commande que l'on applique au système, *i.e.* des variables  $u^+$  aux variables  $-u^-$ , sur la figure 2, lorsque le lien entre les deux ( $\times$ ) est ouvert.

On a les relations :

$$\begin{aligned} u^-(t) &= Gv(t) - K\hat{x}(t), \\ \dot{\hat{x}}(t) &= (A - LC)\hat{x}(t) + (B - LD)u^-(t) + Ly(t), \end{aligned}$$

ce qui conduit au transfert entre  $y(t)$  et  $u^-(t)$  :

$$T_{y(t) \rightarrow u^-(t)} = -K(pI - A + LC + BK - LDK)^{-1}L.$$

Ainsi le transfert de  $u^+(t)$  vers  $-u^-(t)$  s'écrit :

$$T_{\text{BO}}(p) = K(pI - A + LC + BK - LDK)^{-1}L [D + C(pI - A)^{-1}B].$$

Notons que ce transfert est également à comparer au transfert idéal de la boucle ouverte que l'on obtiendrait si l'on avait accès à l'état, soit :

$$T_{\text{BO}}^*(p) = K(pI - A)^{-1}B,$$

Si l'on veut que  $T_{\text{BO}}(p)$  soit aussi proche que possible de  $T_{\text{BO}}^*(p)$  il faut choisir  $K$  et  $L$  tels que :

$$\begin{aligned} K(pI - A + LC + BK - LDK)^{-1}LC &\simeq K, \\ K(pI - A + LC + BK - LDK)^{-1}LD &\simeq 0. \end{aligned}$$

Le principe de détermination des gains reposant sur cette technique s'appelle le recouvrement de transfert de boucle (Loop Transfer Recovery) et sera décrit ultérieurement.

### 2.3 Forme RST du régulateur-observateur

En reprenant les calculs précédents on obtient que la commande fournie par le régulateur-observateur s'écrit sous la forme :

$$u(t) = T_v(p)v(t) - T_y(t)y(t),$$

où :

$$\begin{aligned} T_v(p) &= (I - K(pI - A + LC + BK - LDK)^{-1}(B - LD))G, \\ T_y(t) &= K(pI - A + LC + BK - LDK)^{-1}L. \end{aligned}$$

En notant  $\Delta(p) = \det(pI - A + LC + BK - LDK)$ , et  $M(p) = K [\text{Com}(pI - A + LC + BK - LDK)]$  on obtient la commande écrite sous une forme RST :

$$S(p)u(t) = T(p)v(t) - R(p)y(t),$$

avec :

$$\begin{aligned} S(p) &= \Delta(p)I, \\ T(p) &= (\Delta(p)I - M(p)(B - LD))G, \\ R(p) &= M(p)L. \end{aligned}$$

Dorénavant, et parce que cela est vérifié pour la plupart des processus nous supposons, pour simplifier les expressions, que  $D = 0$ . Dans le cas où  $D$  est non nul, une méthode simple permettant sans aucun changement d'utiliser les résultats qui suivent est de construire le nouveau vecteur de mesures  $Y(t) = y(t) - Du(t) = Cx(t)$ .

### 3 Détermination du gain de l'observateur

Comme nous l'avons vu le gain de l'observateur est fixé d'après le choix des valeurs propres de la matrice  $A - LC$ . Mais comme :

$$\det(pI - A + LC) = \det(pI - A^T + C^T L^T),$$

les méthodes que l'on utilise sont directement transposées des méthodes utilisées pour trouver un gain de retour d'état qui place les pôles du système bouclé. Dans ce qui suit nous noterons par  $a(p) = \det(pI - A) = p^n + \sum_{i=0}^{n-1} a_i p^i$  et  $\hat{a}(p)$  le polynôme représentant les dynamiques désirées pour l'observateur :

$$\hat{a}(p) = \prod_{i=1}^n (p - \lambda_i) = p^n + \sum_{i=0}^{n-1} \hat{a}_i p^i,$$

où les  $\lambda_i$  sont les valeurs propres désirées pour  $A - LC$ .

#### 3.1 Cas monosortie ( $l = 1$ )

Il s'agit de déterminer les  $n$  composantes du vecteur  $L$  :

$$L = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix}.$$

Cela peut être fait directement ou par la formule de Bass et Gura qui permet la construction d'un algorithme.

##### 3.1.1 Calcul direct

En développant  $\det(pI - A + LC)$  on obtient :

$$\det(pI - A + LC) = p^n + \sum_{i=0}^{n-1} f_i(l_1, \dots, l_n) p^i,$$

où les  $f_i(l_1, \dots, l_n)$  sont  $n$  formes linéaires en les inconnues  $l_1, \dots, l_n$ , c'est-à-dire que l'on peut écrire, pour  $i = 0, \dots, n-1$ ,  $f_i(l_1, \dots, l_n) = F_i L$ . Les  $n$  égalités :

$$i = 0, \dots, n-1, \quad f_i(l_1, \dots, l_n) = \hat{a}_i,$$

conduisent à :

$$FL = \hat{a},$$

où :

$$F = \begin{bmatrix} F_0 \\ F_1 \\ \vdots \\ F_{n-1} \end{bmatrix}, \quad \hat{a} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_{n-1} \end{bmatrix}.$$

On obtient donc le gain par :

$$L = F^{-1}\hat{a}.$$

### 3.1.2 Formule de Bass et Gura

On peut écrire :

$$\hat{a}(p) = \det(pI - A + LC) = a(p) \det(I - (pI - A)^{-1}LC),$$

et comme  $\det(I - MN) = \det(I - NM)$ , on obtient :

$$\hat{a}(p) = a(p) \det(1 - C(pI - A)^{-1}L) = a(p) - a(p)C(pI - A)^{-1}L.$$

Or la formule de Leverrier-Souriau donne :

$$\begin{aligned} a(p)(pI - A)^{-1} &= p^{n-1}I + (A + a_{n-1}I)p^{n-2} + \\ &\quad (A^2 + a_{n-1}A + a_{n-2}I)p^{n-3} + \cdots + \\ &\quad (A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I), \end{aligned}$$

ce qui permet d'obtenir, en notant :

$$\tilde{a} = \begin{bmatrix} \hat{a}_{n-1} - a_{n-1} \\ \hat{a}_{n-2} - a_{n-2} \\ \vdots \\ \hat{a}_0 - a_0 \end{bmatrix},$$

les relations :

$$\begin{aligned} \hat{a}_{n-1} - a_{n-1} &= CL, \\ \hat{a}_{n-2} - a_{n-2} &= CAL + a_{n-1}CL, \\ \hat{a}_{n-3} - a_{n-3} &= CA^2L + a_{n-1}CAL + a_{n-2}CL, \\ &\vdots \\ \hat{a}_0 - a_0 &= CA^{n-1}L + a_{n-1}CA^{n-2}L + \cdots + a_2CAL + a_1CL. \end{aligned}$$

Cet ensemble de relations se met sous la forme matricielle :

$$\tilde{a} = T_A O_{(A,C)} L,$$

où :

$$T_A = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ a_{n-1} & 1 & 0 & & 0 \\ a_{n-2} & a_{n-1} & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ a_1 & a_2 & \cdots & a_{n-1} & 1 \end{bmatrix}, \quad O_{(A,C)} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}.$$

Lorsque la paire  $(A, C)$  est observable,  $O_{(A,C)}$  est régulière et on obtient :

$$L = O_{(A,C)}^{-1} T_A^{-1} \tilde{a}.$$

### 3.2 Cas multisortie ( $l > 1$ )

Le cas à plusieurs sorties ne peut être traité directement car on arriverait à un système, toujours linéaire mais de  $n$  équations et à  $nl$  inconnues, les composantes de  $L$ . Pour éviter le problème du choix des inconnues principales et des inconnues secondaires de ce système surdéterminé, nous allons passer par une forme plus simple de l'équation d'état. Cette méthode peut être vue comme une généralisation de la formule de Bass et Gura.

Suivant l'annexe sur les formes canoniques, lorsque le système est observable, on peut, par les changements de variables :  $x_O = T_O x$  et  $y = H_O y_O$ , ramener l'équation d'état (1) sous la forme canonique observable de Luenberger :

$$\begin{aligned} \dot{x}_O(t) &= A_O x_O(t) + B_O u(t), \\ y_O(t) &= C_O x_O(t), \end{aligned} \quad (3)$$

où  $A_O = T_O A T_O^{-1}$ ,  $B_O = T_O B$ , et  $C_O = H_O^{-1} C T_O^{-1}$ , mais où les matrices  $A_O$ ,  $C_O$  et  $H_O$  possèdent les structures suivantes :

–  $A_O$  est formée de  $l^2$  sous-matrices  $A_O^{ij}$ ,  $i = 1$  à  $l$ ,  $j = 1$  à :

$$A_O = \begin{bmatrix} A_O^{11} & \cdots & A_O^{1l} \\ \vdots & & \vdots \\ A_O^{l1} & \cdots & A_O^{ll} \end{bmatrix},$$

qui sont de taille  $(\nu_i \times \nu_j)$  où les  $\nu_k$  sont les  $l$  indices d'observabilité qui vérifient  $\sum_{k=1}^l \nu_k = n$ . De plus, ces sous-matrices ont la forme :

$$A_O^{ii} = \begin{bmatrix} 0 & \cdots & 0 & \times \\ 1 & \ddots & \vdots & \times \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & 1 & \times \end{bmatrix}, \quad \text{et pour } i \neq j, \quad A_O^{ij} = \begin{bmatrix} 0 & \cdots & 0 & \times \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \times \end{bmatrix};$$

- $C_O$  est formée de  $l$  sous-matrices  $C_O^i$ ,  $i = 1$  à  $l$  :

$$C_O = [ C_O^1 \quad \cdots \quad C_O^l ],$$

qui sont de taille  $(l \times \nu_i)$  et qui ont des composantes toutes nulles sauf la composante  $(C_O^i)_{i,\nu_i}$  qui est égale à 1 ;

- $H_O$  est une matrice  $(l \times l)$  qui a pour structure :

$$H_O = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \times & 1 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ \times & \cdots & \times & 1 \end{bmatrix}.$$

Les composantes significatives des matrices (notées ici  $\times$ ) et les indices d'observabilité sont donnés par l'application de l'algorithme de Luenberger de construction de la forme canonique.

L'étape suivante consiste à écrire :

$$\begin{aligned} a(p) &= \det(pI - A + LC), \\ &= \det(pI - T_O^{-1}A_OT_O + LH_OC_OT_O), \\ &= \det(pI - A_O + L_OC_O), \end{aligned}$$

avec  $L_O = T_O L H_O$  et à organiser ce nouveau gain sous la forme :

$$L_O = \begin{bmatrix} L_O^{11} & \cdots & L_O^{1l} \\ \vdots & & \vdots \\ L_O^{l1} & \cdots & L_O^{ll} \end{bmatrix},$$

où les  $L_O^{ij}$ ,  $i = 1$  à  $l$ ,  $j = 1$  sont des vecteurs de dimensions  $(\nu_i \times 1)$ . Lorsque l'on réalise alors l'opération  $A_O - L_OC_O$  on s'aperçoit que l'on peut fixer les vecteurs  $L_O^{ij}$  par les règles suivantes :

- les vecteurs  $L_O^{ij}$  avec  $i \neq j$  sont égaux à la dernière colonne de la matrice  $A_O^{ij}$ , ce qui rend  $A_O$  diagonale par blocs ;
- on décompose le polynôme caractéristique désiré, soit  $\hat{a}(p)$  en  $l$  polynômes  $\hat{a}_i(p)$ , chacun de degré  $\nu_i$  :

$$\hat{a}(p) = \prod_{i=1}^l \hat{a}_i(p);$$

- les  $l$  relations :

$$\det(pI - A_O^{ii} + L_O^{ii} [ 0 \quad \cdots \quad 0 \quad 1 ]) = \hat{a}_i(p),$$

fournissent directement les  $L_O^{ii}$ .

Ayant obtenu  $L_O$ , on obtient  $L$  par :

$$L = T_O^{-1} L_O H_O^{-1}.$$

### 3.3 Méthode LTR

Une méthode de réglage du gain de retour d'état est basée sur la minimisation du critère :

$$J = \int_0^{\infty} (u^T R u + x^T Q x) dt, \quad (4)$$

où  $R$  et  $Q$  sont deux matrices définies positives et conduit au gain optimal :

$$K = R^{-1} B^T P,$$

où  $P$  est solution de l'équation algébrique de Riccati :

$$A^T P + P A - P B R^{-1} B^T P + Q = 0.$$

Par dualité on peut proposer une forme "optimale" pour le gain de l'observateur sous la forme :

$$L = \Pi C^T \Sigma^{-1},$$

où  $\Pi$  est solution de l'équation algébrique de Riccati :

$$A \Pi + \Pi A^T - \Pi C^T \Sigma^{-1} C \Pi + \Lambda = 0,$$

où  $\Sigma$  et  $\Lambda$  sont deux matrices définies positives. Dans le cadre du filtrage optimal on peut donner une interprétation à ces matrices, mais ce point de vue sert surtout de point de départ à la technique de recouvrement de transfert de boucle que nous avons évoqué plus haut. Cette technique est avantageuse car elle assure une bonne robustesse en stabilité, cependant elle n'est utilisable que sur les systèmes carrés à minimum de phase, c'est-à-dire sur ceux dont les zéros du transfert initial :

$$F(p) = C(pI - A)^{-1} B,$$

sont tous à partie réelle négative.

La procédure LTR consiste à poser :

$$\begin{aligned} \Lambda &= B B^T + \nu^2 M, \\ \Sigma &= \nu^2 N, \end{aligned}$$

où  $M$  et  $N$  sont définies positives et  $\nu$  un petit paramètre que l'on va faire tendre vers 0. Dans ces conditions l'équation algébrique de Riccati devient :

$$A \Pi + \Pi A^T - \Pi C^T (\nu^2 N)^{-1} C \Pi + B B^T + \nu^2 M = 0,$$

et le gain :

$$L = \Pi C^T (\nu^2 N)^{-1}. \quad (5)$$

On en déduit :

$$\Pi C^T N^{-1} C \Pi = \nu^2 (A \Pi + \Pi A^T + B B^T) + \nu^4 M,$$

soit :

$$\lim_{\nu \rightarrow 0} \Pi = 0 \text{ et } \lim_{\nu \rightarrow 0} L(\nu^2 N)L^T = BB^T.$$

Ainsi :

$$\lim_{\nu \rightarrow 0} L = \frac{1}{\nu} BUN^{-1/2},$$

où  $U$  est une matrice carrée orthogonale,  $UU^T = I$ , et  $N^{1/2}$  est une racine carrée de  $N$ ,  $N = N^{1/2}(N^{1/2})^T$ .

Remarque : il existe une infinité de racine carrées d'une matrice définie positive. On peut cependant en proposer une d'expression simple. En effet,  $N$  étant définie positive, il existe une matrice orthogonale,  $V$ , et une matrice diagonale positive,  $\Delta = \text{diag}\{d_i, d_i > 0\}$ , telles que  $N = V\Delta V^T$ . En posant  $\sqrt{\Delta} = \text{diag}\{\sqrt{d_i}\}$ , on peut écrire :

$$N^{1/2} = V\sqrt{\Delta}V^T.$$

Dans ces conditions, il vient :

$$\lim_{\nu \rightarrow 0} L = \frac{1}{\nu} BW\sqrt{\Delta}^{-1}V^T,$$

où  $W$  est une matrice orthogonale arbitraire.

Maintenant il s'agit de montrer que lorsque  $\nu$  tend vers 0, le transfert  $T_{\text{BO}}(p)$  :

$$T_{\text{BO}}(p) = K(pI - A + LC + BK)^{-1}LC(pI - A)^{-1}B,$$

tend vers le transfert idéal  $T_{\text{BO}}^*(p) = K(pI - A)^{-1}B$ . Pour ce faire, posons :

$$\Phi = (pI - A + BK)^{-1}.$$

Alors, l'utilisation du lemme d'inversion matricielle :

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

conduit à écrire :

$$T_{\text{BO}}(p) = K\Phi L[I + C\Phi L]^{-1}F(p).$$

Lorsque  $\nu$  tend vers 0, le transfert  $T_{\text{BO}}(p)$  tend vers :

$$\begin{aligned} & K\Phi \frac{1}{\nu} BUN^{-1/2} [I + C\Phi \frac{1}{\nu} BUN^{-1/2}]^{-1} F(p) \\ &= K\Phi BUN^{-1/2} [\nu I + C\Phi BUN^{-1/2}]^{-1} F(p) \\ &\longrightarrow K\Phi B [C\Phi B]^{-1} F(p). \end{aligned}$$

D'autre part, toujours en utilisant le lemme d'inversion matricielle, on a :

$$\Phi = (pI - A)^{-1} - (pI - A)^{-1}B[I + T_{\text{BO}}^*(p)]^{-1}K(pI - A)^{-1},$$

ce qui conduit aux relations :

$$\begin{aligned} K\Phi B &= T_{\text{BO}}^*(p)[I + T_{\text{BO}}^*(p)]^{-1}, \\ C\Phi B &= F(p)[I + T_{\text{BO}}^*(p)]^{-1}. \end{aligned}$$

Ainsi, lorsque  $F(p)$  est d'inverse stable :

$$K\Phi B[C\Phi B]^{-1}F(p) = T_{\text{BO}}^*(p).$$

En résumé la procédure LTR se déroule suivant les étapes :

1. On calcule  $K$  par la minimisation d'un critère linéaire quadratique (4). Cela n'est pas obligatoire mais cela permet d'assurer de bonnes marges de robustesse *a priori*.
2. On calcule un gain optimal (5) avec  $\Lambda = BB^T + \nu^2 M$ ,  $\Sigma = \nu^2 N$  où  $\nu$  est fixé à une valeur arbitraire.
3. On trace les caractéristiques fréquentielles obtenues avec  $T_{\text{BO}}(j\omega)$ .
4. On mesure les marges de robustesse. Si elles sont satisfaisantes on garde les gains trouvés sinon on revient à la deuxième étape en diminuant  $\nu$ .

## 4 Systèmes discrets

### 4.1 Observateurs prédicteurs ou correcteurs

Considérons le système discret stationnaire défini par :

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \\ y_k &= Cx_k, \end{aligned}$$

où  $k \in \mathbb{N}$ ,  $x_k \in \mathbb{R}^n$  est l'état du modèle,  $u_k \in \mathbb{R}^m$ , l'entrée du processus,  $y_k \in \mathbb{R}^l$ , la sortie mesurée du processus, et  $A$ ,  $B$  et  $C$  sont des matrices de dimensions adaptées.

On construit un observateur pour ce système par le principe énoncé dans le cas des systèmes continus, c'est-à-dire :

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + G_k,$$

où  $G_k$  est un terme de correction, seulement ici nous distinguerons plusieurs types de corrections donc d'observateurs suivant que la mesure est disponible à l'instant précédent l'estimation ou à l'instant de l'estimation. Dans le premier cas nous construirons un observateur-prédicteur donnant une estimation de l'état à l'instant  $k + 1$  d'après ce que l'on connaît à l'instant  $k$ , nous noterons  $\hat{x}_{k+1/k}$  cette prédiction. Dans le deuxième cas, nous construirons un observateur-correcteur, l'estimation de l'état sera alors notée  $\hat{x}_{k+1/k+1}$ .

#### 4.1.1 Observateur - prédicteur

Dans ce premier cas on envisage  $G_k = L^p[y_k - \hat{y}_{k/k-1}]$ , où  $L^p$  est un gain, ce qui conduit au reconstruteur d'état défini par :

$$\hat{x}_{k+1/k} = A\hat{x}_{k/k-1} + Bu_k + L^p[y_k - \hat{y}_{k/k-1}],$$

avec  $\hat{y}_{k/k-1} = C\hat{x}_{k/k-1}$ . Ce qui peut s'écrire également :

$$\hat{x}_{k+1/k} = [A - K^p C]\hat{x}_{k/k-1} + Bu_k + L^p y_k.$$

Cette structure est identique à celle obtenue dans le cas des systèmes continus. En notant  $\tilde{x}_{k+1/k}$  l'erreur de reconstruction  $x_{k+1} - \hat{x}_{k+1/k}$ , il vient :

$$\tilde{x}_{k+1/k} = [A - L^p C]\tilde{x}_{k/k-1}.$$

Le reconstruteur est asymptotique si le gain  $L^p$  est choisi tel que les valeurs propres de la matrice  $A - L^p C$  sont, en module, inférieures à l'unité (on a alors une matrice de Schur). Dans le cas où la paire  $(A, C)$  est observable, ces valeurs propres peuvent être fixées arbitrairement à l'aide des techniques abordées dans la partie précédente.

#### 4.1.2 Observateur - correcteur

Dans le cas où  $y_{k+1}$  est connue, on peut envisager que  $G_k = L^c[y_{k+1} - \bar{y}_{k+1}]$ , où  $L^c$  est un gain et  $\bar{y}_{k+1}$  représente l'estimation de la sortie à l'instant  $k + 1$  qui peut être réalisée de deux façons différentes : par anticipation ou par prédiction. L'estimation de  $x_{k+1}$  est fournie par le reconstruteur :

$$\hat{x}_{k+1/k+1} = A\hat{x}_{k/k} + Bu_k + L^c[y_{k+1} - \bar{y}_{k+1}].$$

**Anticipation** Si on suppose connue  $\hat{x}_{k+1/k+1}$ , on pose dans ce cas :

$$\bar{y}_{k+1} = C\hat{x}_{k+1/k+1},$$

ce qui conduit à écrire l'observateur, sous la forme :

$$\hat{x}_{k+1/k+1} = [I + L^c C]^{-1}[A\hat{x}_{k/k} + Bu_k + L^c y_{k+1}].$$

L'erreur de reconstruction,  $\tilde{x}_{k+1/k+1} = x_{k+1} - \hat{x}_{k+1/k+1}$ , est définie par l'équation récurrente :

$$\tilde{x}_{k+1/k+1} = [I + L^c C]^{-1} A \tilde{x}_{k/k}.$$

Pour obtenir un reconstruteur asymptotique,  $L^c$  doit être choisie telle que  $[I + L^c C]^{-1} A$  soit une matrice de Schur. Cela n'est pas toujours possible et c'est ce qui rend cette forme d'observateur peu utilisée en pratique.

**Prédiction** Dans ce cas,  $\bar{y}_{k+1}$  est la meilleure estimation que l'on puisse faire de la sortie à partir de l'information disponible avant le calcul de  $\hat{x}_{k+1/k+1}$ . On pose donc :

$$\bar{y}_{k+1} = C[A\hat{x}_{k/k} + Bu_k],$$

ce qui conduit à la deuxième forme de restructeur-correcteur :

$$\hat{x}_{k+1/k+1} = [I - L^c C][A\hat{x}_{k/k} + Bu_k] + L^c y_{k+1}.$$

Avec les mêmes notations que précédemment, l'erreur de reconstruction est régie par l'équation :

$$\tilde{x}_{k+1/k+1} = [I - L^c C]A\tilde{x}_{k/k},$$

Si le gain  $L^c$  est choisi tel que  $[I - L^c C]A$  soit de Schur, le restructeur obtenu sera asymptotique. Lorsque la paire  $(A, CA)$  est observable, il est toujours possible de trouver  $L^c$  correspondant à des dynamiques fixées *a priori* pour l'erreur. Cependant, on peut remarquer que l'on a :

$$O_{(A, CA)} = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^n \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} A = O_{(A, C)} A,$$

soit, si la paire  $(A, C)$  est observable, d'après les inégalités de Sylvester,  $\text{rang}(O_{(A, CA)}) = \text{rang} A$ . Ainsi cet observateur ne sera pas utilisé lorsque  $A$  a une valeur propre nulle, car on a une perte d'observabilité.

#### 4.1.3 Remarques sur les observateurs de systèmes discrets

Quel que soit le type d'observateur choisi, le principe de séparation énoncé dans le cas des systèmes continus reste valide : commande par retour d'état et observateur (prédicteur ou correcteur) peuvent être déterminés séparément.

Comme dans le cas continu, deux approches existent pour le choix du gain du restructeur d'état, l'une basée sur la notion de placement de pôles, l'autre basée sur la minimisation d'un critère quadratique associé, la procédure LTR pouvant également être appliquée.

Cependant, dans le cas des systèmes discrets, il existe un choix particulièrement intéressant qui consiste à déterminer  $L$  tel que la matrice  $A - LC$  soit nilpotente, c'est à dire que toutes ses valeurs propres soient nulles. En effet, dans ce cas on obtient un observateur à réponse pile où l'erreur est exactement nulle au bout de  $\max\{\nu_i\}$  pas où les  $\nu_i$  sont les indices d'observabilité. Notons la différence avec un observateur asymptotique où l'erreur d'estimation tend vers 0. Pour calculer le gain  $L$  correspondant on reprend la technique développée dans le cas continu et il suffit de prendre  $L_O$  telle que toutes les composantes significatives de  $A_O - L_O C_O$  soient nulles.

Une autre technique mais qui ne demande pas de passage par la forme canonique observable consiste à faire les remarques suivantes :

$$Y_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix} = O_{(A,C)}x_k + MU_k,$$

où :

$$U_k = \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+n-2} \end{bmatrix}, \quad M = \begin{bmatrix} 0 & \cdots & 0 \\ CB & \ddots & \vdots \\ \vdots & \ddots & 0 \\ CA^{n-2}B & \cdots & CB \end{bmatrix},$$

soit, si le système est observable :

$$x_k = O_{(A,C)}^{-1}(Y_k - MU_k).$$

Cependant, cet observateur n'étant pas réalisable, il convient de remarquer également que :

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_{k-1}, \\ &= A^2x_{k-2} + ABu_{k-2} + Bu_{k-1}, \\ &\vdots \\ &= A^{n-1}x_{k-n+1} + \sum_{i=1}^{n-1} A^{i-1}Bu_{k-i}, \\ &= A^{n-1}x_{k-n+1} + \Psi U_{k-n+1}, \end{aligned}$$

où  $\Psi = [A^{n-2}B \ \cdots \ AB \ B]$ . Or  $x_{k-n+1} = O_{(A,C)}^{-1}(Y_{k-n+1} - MU_{k-n+1})$ , ce qui conduit à l'observateur réalisable :

$$x_k = A^{n-1}O_{(A,C)}^{-1}Y_{k-n+1} + \left(\Psi - A^{n-1}O_{(A,C)}^{-1}M\right)U_{k-n+1}.$$

## 5 Simplification des observateurs

Dans les parties précédentes, nous avons déterminé des systèmes observateurs de même dimension que l'état du système à reconstruire. Nous allons montrer que l'on peut construire, en réalité, des reconstituteurs d'ordre inférieur. Plusieurs principes peuvent être utilisés. Nous regarderons successivement : le reconstituteur réduit de Luenberger qui estime la partie non accessible de l'état et les observateurs de fonctionnelles linéaires de l'état qui permettent d'estimer directement la commande lorsque celle-ci est un retour d'état linéaire. Nous verrons que si dans le premier cas l'observateur est assez facile à déterminer, il n'en est pas de même dans le deuxième. Cependant cette dernière solution est très intéressante car elle permet d'obtenir des observateurs de taille beaucoup plus petite (la minimalité est encore un problème ouvert) et donc d'augmenter la robustesse d'une régulation par régulateur-observateur.

## 5.1 Observateur d'ordre réduit

### 5.1.1 Construction

Cet observateur, appelé **observateur de Luenberger**, utilise le fait que la matrice d'observation soit de rang plein. Cette hypothèse n'est pas restrictive car il suffit alors d'éliminer les composantes de la sortie redondantes. Soit le système (1) avec  $D = 0$ , et où après une permutation des variables d'état, les matrices  $A$ ,  $B$ ,  $C$ , et le vecteur  $x$  sont de la forme :

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \\ C = [C_1 \quad C_2], \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

où  $\text{rang } C_1 = l$ ,  $A_{11}$  et  $C_1 \in \mathbb{R}^{l \times l}$ ,  $x_1 \in \mathbb{R}^l$ , et  $B_1 \in \mathbb{R}^{l \times m}$ .

Le changement de variables :

$$\bar{x} = \begin{bmatrix} C_1 & C_2 \\ 0 & I_{n-l} \end{bmatrix} x,$$

transforme le système (1) en :

$$\begin{aligned} \dot{\bar{x}}(t) &= \bar{A}\bar{x}(t) + \bar{B}u(t), \\ y(t) &= \bar{C}\bar{x}(t), \end{aligned} \tag{6}$$

où :

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} \bar{B}_1 \\ B_2 \end{bmatrix}, \quad \bar{C} = [I_m \quad 0].$$

Ces matrices sont définies par :

$$\begin{aligned} \bar{A}_{11} &= [C_1 A_{11} + C_2 A_{21}] C_1^{-1}, \\ \bar{A}_{12} &= -\bar{A}_{11} C_2 + [C_1 A_{12} + C_2 A_{22}], \\ \bar{A}_{21} &= A_{21} C_1^{-1}, \\ \bar{A}_{22} &= A_{22} - A_{21} C_1^{-1} C_2, \\ \bar{B}_1 &= C_1 B_1 + C_2 B_2. \end{aligned}$$

L'équation d'état (6) s'écrit donc sous la forme :

$$\begin{aligned} \dot{\bar{x}}_1(t) &= \bar{A}_{11}\bar{x}_1(t) + \bar{A}_{12}\bar{x}_2(t) + \bar{B}_1 u(t), \\ \dot{\bar{x}}_2(t) &= \bar{A}_{21}\bar{x}_1(t) + \bar{A}_{22}\bar{x}_2(t) + B_2 u(t), \\ y(t) &= \bar{x}_1(t), \end{aligned} \tag{7}$$

où la sortie correspond aux  $l$  premières composantes d'état, elles n'ont donc pas à être reconstruites. La première équation de (7) peut alors être considérée comme une mesure  $\xi(t)$  dépendant de  $\bar{x}_2(t)$ , variables d'état à reconstruire, et de  $u(t)$  :

$$\xi(t) = \dot{\bar{x}}_1(t) - \bar{A}_{11}\bar{x}_1(t) = \bar{A}_{12}\bar{x}_2(t) + \bar{B}_1 u(t).$$

Suivant le principe de construction des observateurs, on peut proposer comme reconstruteur de  $\bar{x}_2(t)$ , le vecteur  $\hat{v}(t)$  défini par :

$$\begin{aligned}\dot{\hat{v}}(t) &= \bar{A}_{21}\bar{x}_1(t) + \bar{A}_{22}\hat{v}(t) + B_2u(t) + \bar{L}[\xi(t) - \hat{\xi}(t)], \\ \bar{x}_1(t) &= y(t), \quad \hat{\xi}(t) = \bar{A}_{12}\hat{v}(t) + \bar{B}_1u(t).\end{aligned}$$

L'inconvénient de cette structure est de nécessiter, pour élaborer la mesure  $\xi(t)$ , la dérivation de la sortie réelle  $y(t)$ . De façon à contourner cette difficulté, on définit la variable :

$$\hat{z}(t) = \hat{v}(t) - \bar{L}y(t),$$

ce qui permet d'obtenir :

$$\begin{aligned}\dot{\hat{z}}(t) &= \dot{\hat{v}}(t) - \bar{K}\dot{\bar{x}}_1(t), \\ &= \bar{A}_{21}\bar{x}_1(t) + \bar{A}_{22}\hat{v}(t) + B_2u(t) \\ &\quad - \bar{L}[\bar{A}_{11}\bar{x}_1(t) + \bar{A}_{12}\hat{v}(t) + \bar{B}_1u(t)],\end{aligned}\tag{8}$$

où n'apparaît aucune dérivation de la sortie. En tenant compte du fait que  $\hat{v}(t) = \hat{z}(t) + \bar{L}y(t)$  et  $\bar{x}_1(t) = y(t)$ , (8) se réécrit sous la forme :

$$\begin{aligned}\dot{\hat{z}}(t) &= \bar{M}\hat{z}(t) + \bar{N}u(t) + \bar{P}y(t), \\ \bar{M} &= \bar{A}_{22} - \bar{L}\bar{A}_{12}, \\ \bar{N} &= B_2 - \bar{L}\bar{B}_1, \\ \bar{P} &= \bar{A}_{21} + \bar{A}_{22}\bar{L} - \bar{L}\bar{A}_{11} - \bar{L}\bar{A}_{12}\bar{L}.\end{aligned}\tag{9}$$

Cette équation d'état définit un observateur réduit (d'ordre  $n - l$ ) pour le système (7), la variable  $\bar{x}_2(t)$  étant reconstruite par :

$$\hat{v}(t) = \hat{z}(t) + \bar{L}y(t),$$

où  $\bar{L}$  est la matrice  $((n - l) \times l)$  de gain de cet observateur. Notons l'erreur d'observation  $e(t) = \bar{x}_2(t) - \hat{v}(t)$ , il vient :

$$\dot{e}(t) = \bar{M}e(t),$$

Ainsi, lorsque la paire  $(\bar{A}_{22}, \bar{A}_{12})$  est observable les valeurs propres de  $\bar{M}$  peuvent être fixées arbitrairement par un choix convenable de  $\bar{L}$ .

**Théorème 2** *Si  $(A, C)$  est observable, alors  $(\bar{A}_{22}, \bar{A}_{12})$  est observable.*

Démonstration : Il suffit de montrer que si  $(\bar{A}, \bar{C})$  est observable, alors  $(\bar{A}_{22}, \bar{A}_{12})$  l'est également. Si  $(\bar{A}, \bar{C})$  est observable, on a :

$$\forall s, \text{rang} \begin{bmatrix} sI_n - \bar{A} \\ \bar{C} \end{bmatrix} = \text{rang} \begin{bmatrix} sI_l - \bar{A}_{11} & -\bar{A}_{12} \\ -\bar{A}_{21} & sI_{n-l} - \bar{A}_{22} \\ I_l & 0 \end{bmatrix} = n,$$

ce qui est équivalent à :

$$\forall s, \text{rang} \begin{bmatrix} \bar{A}_{12} \\ sI_{n-l} - \bar{A}_{22} \end{bmatrix} = n - l,$$

donc  $(\bar{A}_{22}, \bar{A}_{12})$  est une paire observable.

Dans le cas où le système initial est observable, on peut appliquer les techniques de détermination du gain à la paire  $(\bar{A}_{22}, \bar{A}_{12})$  pour déterminer le gain  $\bar{L}$ . L'observateur réduit du système (1) a finalement la structure suivante, en notant  $\hat{x}_1$  et  $\hat{x}_2$  les estimations des vecteurs  $x_1$  et  $x_2$  :

$$\dot{\hat{z}}(t) = \bar{M}\hat{z}(t) + \bar{N}u(t) + \bar{P}y(t),$$

où  $\bar{M}$ ,  $\bar{N}$  et  $\bar{P}$  sont définies par les relations (9) et :

$$\begin{aligned} \hat{x}_1(t) &= C_1^{-1}[(I - C_2\bar{L})y(t) - C_2\hat{z}(t)], \\ \hat{x}_2(t) &= \hat{z}(t) + \bar{L}y(t). \end{aligned}$$

## 5.2 Observation d'une fonction linéaire de l'état

Dans la plupart des cas, l'observation de l'état d'un système est réalisée pour construire une loi de commande linéaire de retour d'état de la forme  $Kx$ . On peut donc penser à construire un reconstruteur (appelé **observateur de fonctionnelle linéaire**) estimant directement la variable  $w(t) = Kx(t)$  sans utiliser l'étape intermédiaire d'estimation de l'état. Nous verrons dans ce paragraphe que cela apporte une simplification, par diminution de l'ordre, de la structure de l'observateur utilisé.

### 5.2.1 Réduction de l'ordre de l'observateur

Si on considère le cas des systèmes mono-entrées,  $m = 1$ , la fonctionnelle linéaire à estimer est une forme linéaire de l'état,  $w = kx$ ,  $k \in \mathbb{R}^n$ . L'utilisation de la méthode de calcul du gain indique que l'observation de l'état d'un système multi-sorties se fait par l'intermédiaire de  $l$  sous-systèmes chacun observant une partie  $x^i$  de l'état et dont les ordres sont donnés par les  $l$  indices d'observabilité  $\nu_i$ ,  $i \in \{1, \dots, l\}$ , définis à l'annexe sur l'algorithme de Luenberger. L'utilisation du principe de réduction d'ordre détaillé dans la partie précédente permet d'amener l'ordre de ces sous-systèmes à  $\nu_i - 1$ . La décomposition de la variable sous la forme :

$$w = \sum_{i=1}^l k_i x^i,$$

conduit à la structure de principe de la figure 3. .

En considérant le transfert de chacun des observateurs, on peut écrire :

$$i = \{1, \dots, l\}, D_i(p)w_i(t) = N_u^i(p)u(t) + N_y^i(p)y(t),$$

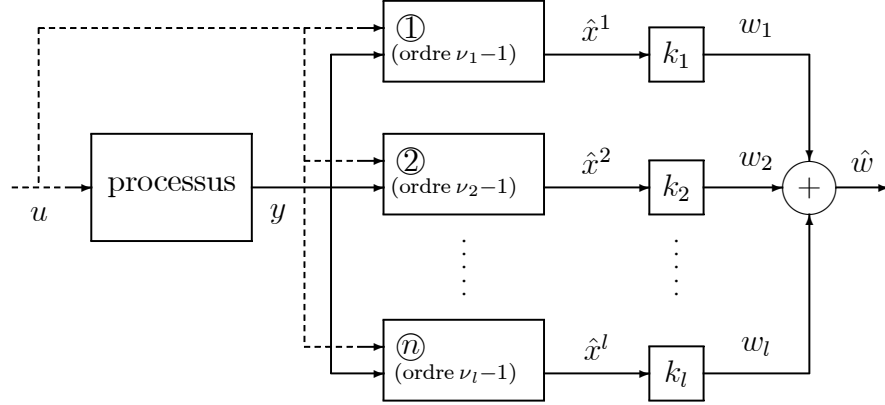


FIG. 3 – Utilisation d'observateurs réduits.

où  $\deg D_i(p) = \nu_i - 1$ ,  $\deg N_u^i(p) \leq \deg D_i(p)$  et  $\deg N_y^i(p) \leq \deg D_i(p)$ . Comme les modes de ces observateurs réduits peuvent être fixés arbitrairement, on peut choisir le gain de l'observateur tel que :

$$\forall i \in \{1, \dots, l\}, i \neq \nu, \exists P_i(s) \in \mathbb{R} \langle p \rangle,$$

$$\deg P_i(p) = \nu - \nu_i \text{ et } D_\nu(p) = P_i(p)D_i(p),$$

où  $\mathbb{R} \langle p \rangle$  représente l'ensemble des polynômes en la variable  $p$  à coefficients réels et  $\nu$  l'index d'observabilité du système défini par :

$$\nu = \max_{i=1}^l \{\nu_i\}.$$

Ainsi, il vient :

$$\begin{aligned} D_\nu(p)\hat{w}(t) &= D_\nu(p) \sum_{i=1}^l w_i(t) = \sum_{i=1}^l D_\nu(p)w_i(t), \\ &= \sum_{i=1}^l P_i(p)D_i(p)w_i(t) = \sum_{i=1}^l P_i(p) [N_u^i(p)u(t) + N_y^i(p)y(t)], \\ &= \left[ \sum_{i=1}^l P_i(p)N_u^i(p) \right] u(t) + \left[ \sum_{i=1}^l P_i(p)N_y^i(p) \right] y(t), \\ &= N_u(p)u(t) + N_y(p)y(t), \end{aligned}$$

où  $\deg D_\nu(p) = \nu - 1$ ,  $\deg N_u(p) \leq \deg D_\nu(p)$ , et  $\deg N_y(p) \leq \deg D_\nu(p)$ .

Ceci indique que l'observateur d'une forme linéaire peut être réalisé par un observateur d'ordre  $\nu - 1$ . Dans le cas où  $\nu - 1 < n - l$ , ce résultat implique une réduction de l'ordre de l'observateur par rapport à l'emploi d'un observateur réduit.

### 5.2.2 Structure générale

L'étude du cas particulier précédent nous a permis de mettre en évidence la réduction de dimensionnalité apportée par la notion d'observateurs de formes linéaires. Dans cette partie, nous allons préciser, dans le cas général d'une fonctionnelle linéaire,  $l \geq 1$ , la structure de l'observateur utilisé. Pour estimer la fonctionnelle linéaire,  $Kx$ ,  $K \in \mathbb{R}^{m \times n}$ , on utilise un observateur de Luenberger dont la structure générale est :

$$\begin{aligned} \dot{z}(t) &= Dz(t) + Hu(t) + Ey(t), \\ w(t) &= Pz(t) + Vy(t), \end{aligned} \quad (10)$$

où  $z(t) \in \mathbb{R}^p$ ,  $w(t) \in \mathbb{R}^m$  et  $D, E, H, P, V$  sont des matrices de dimensions convenables. Notons que l'observateur réduit est de cette forme. La sortie  $w(t)$  de cet observateur reconstruira asymptotiquement la fonctionnelle linéaire si l'on a :

$$\lim_{t \rightarrow \infty} [w(t) - Kx(t)] = 0. \quad (11)$$

D'après la relation (11) et compte tenu de l'équation de sortie de (10), si  $w(t)$  est une estimation de  $Kx(t)$ , alors  $z(t)$  sera une estimation d'une autre forme linéaire de l'état, par exemple  $Tx(t)$ . Soient les erreurs d'estimation :

$$\begin{aligned} e(t) &= w(t) - Kx(t), \\ \epsilon(t) &= z(t) - Tx(t), \end{aligned}$$

il vient :

$$\begin{aligned} \dot{e}(t) &= Pe(t) + [PT + VC - K]x(t), \\ \dot{\epsilon}(t) &= D\epsilon(t) + [DT - TA + EC]x(t) + [H - TB]u(t). \end{aligned}$$

Si les relations suivantes :

1.  $D$  est une matrice de Hurwitz,
2.  $TA - DT = EC$ ,
3.  $PT + VC = K$ ,
4.  $H = TB$ ,

sont vérifiées, alors  $\lim_{t \rightarrow \infty} e(t) = 0$ , et (10) constitue bien un observateur asymptotique de la fonctionnelle linéaire. Lorsque  $D$  est fixée, et possède des valeurs propres différentes de celles de  $A$  alors la deuxième relation est soluble pour tout  $E$ , et dans ce cas le problème de la construction de l'observateur réside dans la détermination du triplet  $(E, P, V)$  tel que soit vérifiée la relation 3 (la dernière donnant directement  $G$ ), le problème principal étant de construire un observateur de fonctionnelle linéaire minimal, c'est-à-dire où  $p$  est le plus petit possible. Nous allons regarder quelques méthodes permettant de répondre à cette question.

Auparavant, montrons que ces relations sont automatiquement vérifiées dans le cas d'un observateur réduit. En supposant, ce qui ne restreint pas ce résultat mais simplifie les notations, supposons que  $C = [ I \ 0 ]$  et  $K = [ K_1 \ K_2 ]$  alors l'observateur réduit correspond à :

$$\begin{aligned} D &= A_{22} - LA_{12}, & H &= B_2 - LB_1, \\ E &= A_{21} + A_{22}L - LA_{11} - LA_{12}L, \\ P &= K_2, & V &= K_1 + K_2L, & T &= [ -L \ I ], \end{aligned}$$

ce qui donne :

$$\begin{aligned} TA &= [ A_{21} - LA_{11} \ A_{22} - LA_{12} ], & DT &= [ LA_{12}L - A_{22}L \ A_{22} - LA_{12} ], \\ TB &= B_2 - LB_1, & PT &= [ -K_2L \ K_2 ], & VC &= [ K_1 + K_2L \ 0 ], \end{aligned}$$

qui montre que les trois dernières relations sont vérifiées, la première ne dépendant que du choix de  $L$ . Cette remarque permet d'affirmer que l'ordre d'un observateur de fonctionnelle linéaire sera compris entre  $\nu - 1$  et  $n - m$ .

### 5.2.3 Détermination de l'observateur

**Cas d'une forme linéaire ( $m = 1$ )** Dans ce cas, suivant ce qui précède, on a  $p = \nu - 1$ . Soit  $\lambda_i, i \in \{1, \dots, p\}$  l'ensemble des valeurs propres de  $D$  que l'on peut choisir arbitrairement. La méthode, dite de Murdoch, consiste à décomposer  $D$ , à l'aide d'une matrice  $U$  régulière, sous la forme :

$$D = U\Delta U^{-1}$$

où  $\Delta = \text{diag}_{i=1}^p[\lambda_i]$ . Posons les matrices :

$$R = U^{-1}T = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix}, \text{ et } M = U^{-1}E = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_p \end{bmatrix},$$

la deuxième condition nécessaire s'écrit alors :

$$RA - \Delta R = MC,$$

soit, pour  $i \in \{1, \dots, p\}$  :

$$r_i = m_i C (A - \lambda_i I_n)^{-1}. \quad (12)$$

Si, de plus, on impose pour  $P$  la structure :

$$P = 1_p U^{-1},$$

où  $1_p = [1, 1, \dots, 1] \in \mathbb{R}^p$ , la troisième condition se met sous la forme :

$$VC + 1_p R = L,$$

soit, d'après (12) :

$$VC + \sum_{j=1}^p m_j C(A - \lambda_j I_n)^{-1} = L.$$

La multiplication, par  $\prod_{i=1}^p (A - \lambda_i I_n)$ , de chacun des membres de cette égalité, permet de la mettre sous la forme matricielle :

$$X\mathcal{A} = \mathcal{B}, \quad (13)$$

où  $X = [V, m_1, \dots, m_p]$  est un vecteur inconnu de dimension  $\nu l$  et :

$$\mathcal{B} = L \prod_{i=1}^p (A - \lambda_i I_n),$$

$$\mathcal{A} = \begin{bmatrix} C \prod_{i=1}^p (A - \lambda_i I_n) \\ C \prod_{i=2}^p (A - \lambda_i I_n) \\ \vdots \\ C \prod_{\substack{i=1 \\ i \neq j}}^p (A - \lambda_i I_n) \\ \vdots \\ C \prod_{i=1}^{p-1} (A - \lambda_i I_n) \end{bmatrix}.$$

Si la paire  $(A, C)$  est observable, et si les  $\lambda_i$  sont toutes distinctes et ne sont pas valeurs propres de  $A$  alors  $\text{rang} \mathcal{A} = n$  et la solution de ce système existe toujours (elle est unique pour  $\nu m = n$ ) et elle s'écrit :

$$X = \mathcal{B}\mathcal{A}^+,$$

où  $\mathcal{A}^+$  est la pseudo-inverse de  $\mathcal{A}$  :

$$\mathcal{A}^+ = \mathcal{A}^T (\mathcal{A}\mathcal{A}^T)^{-1}.$$

La connaissance de  $X$  permet alors de calculer tous les paramètres de l'observateur (10).

**Cas général** ( $m > 1$ ) Dans le cas quelconque la détermination de l'ordre minimal de l'observateur est délicate, et bien souvent on construira un observateur réduit mais non minimal. On peut appliquer la méthode précédente avec l'équation linéaire (13) à résoudre où la dimension de  $X$  est  $m \times (p + 1)l$ , mais le problème réside dans le choix de  $p$ . En effet une valeur trop petite peut conduire à un système non soluble, cependant certains algorithmes, comme celui de Sirisena permettent d'accéder à cette valeur minimale et de déterminer les paramètres de l'observateur minimal.

Dans le cas où l'on cherche à construire rapidement un observateur (non forcément minimal) de la fonctionnelle linéaire :

$$Lx = \begin{bmatrix} L_1 x \\ L_2 x \\ \vdots \\ L_m x \end{bmatrix}, \quad L_i \in \mathbb{R}^n, \quad i = 1, \dots, m,$$

on peut utiliser une méthode interactive basée sur l'algorithme utilisé dans le cas où  $m = 1$ . Cette méthode consiste à construire un ensemble de  $m$  observateurs décrits par les équations :

$$\begin{aligned} i &\in \{1, \dots, m\}, \\ \dot{z}_i(t) &= D_i z_i(t) + G_i u(t) + E_i y(t) + \sum_{j=1}^{i-1} K_{ij} z_j(t), \\ w_i(t) &= P_i z_i(t) + V_i y(t), \end{aligned}$$

le  $i$ -ième observateur estimant la forme linéaire  $L_i x$ . L'intérêt de cette méthode est de définir cette série d'observateurs de façon récursive : le  $(i+1)$ -ème étant construit par l'application de la méthode précédente au système  $(A, B, C_i)$  avec :

$$C_i = \begin{bmatrix} C \\ T_1 \\ \vdots \\ T_i \end{bmatrix},$$

où les  $T_j$ ,  $j \leq i$ , sont les matrices apparaissant aux étapes précédentes de détermination des observateurs 1 à  $i$ . La dimension de l'observateur complet est inférieure à celle que l'on obtiendrait avec  $m$  observateurs de formes linéaires en parallèle. En effet, l'index d'observabilité  $\nu_i$  à l'étape  $i$  est inférieur ou égal à celui de l'étape précédente, car  $T_{i-1}$  contient au moins une ligne linéairement indépendante des lignes de  $C$  ou de  $T_j$ ,  $j \in \{1, \dots, i-1\}$  (sinon la forme linéaire  $L_{i-1}$  est inutile). On obtient donc un observateur de taille réduite que l'on peut déterminer rapidement.

## 5.3 Utilisation en boucle fermée et robustesse

### 5.3.1 Principe de séparation

Dans le cas où les 4 relations qui caractérisent un observateur asymptotique de fonctionnelle linéaire sont vérifiées nous allons voir que le principe de séparation s'applique toujours. En effet, dans ce cas, nous avons les relations de bouclage :

$$\begin{aligned} \dot{\epsilon}(t) &= D\epsilon(t), \\ w(t) &= P\epsilon(t) + Kx(t), \\ u(t) &= Gv(t) - w(t). \end{aligned}$$

Le système bouclé admet donc comme état  $X(t) = \begin{bmatrix} x(t) & \epsilon(t) \end{bmatrix}$  et comme équation d'état :

$$\begin{aligned} \dot{X}(t) &= \begin{bmatrix} A - BK & -BP \\ 0 & D \end{bmatrix} X(t) + \begin{bmatrix} BG \\ 0 \end{bmatrix} v(t), \\ y(t) &= \begin{bmatrix} I & 0 \end{bmatrix} X(t), \end{aligned}$$

ce qui permet de reprendre le principe de séparation dans notre cas : les pôles en boucle fermée et ceux de l'observateur peuvent être réglés de façon indépendante.

### 5.3.2 Forme RST

La commande à réaliser est de la forme :

$$u(t) = Gv(t) - Pz(t) - Vy(t),$$

où  $z(t)$  est régi par l'équation :

$$\dot{z}(t) = Dz(t) + H(Gv(t) - Pz(t) - Vy(t)) + Ey(t),$$

soit le transfert :

$$(pI - D + HP)z(t) = HGv(t) + (E - HV)y(t).$$

La commande  $u$  s'écrit donc :

$$u = [I - P(pI - D + HP)^{-1}H] Gv(t) - [V + P(pI - D + HP)^{-1}(E - HV)] y(t).$$

En notant  $\Delta(p) = \det(pI - D + HP)$  et  $M(p) = P[\text{Com}(pI - D + HP)]^T$ , on obtient la commande sous forme RST :

$$S(p)u(t) = T(p)v(t) - R(p)y(t),$$

avec :

$$\begin{aligned} S(p) &= \Delta(p)I, \\ T(p) &= [\Delta(p)I - M(p)H] G, \\ R(p) &= [\Delta(p)V + M(p)(E - HV)]. \end{aligned}$$

### 5.3.3 Robustesse

A partir des expressions précédentes il est facile de calculer le transfert de boucle entre  $u^+$  et  $-u^-$ , lorsque  $v(t) \equiv 0$  et lorsque l'on ouvre la régulation juste avant le système. On obtient :

$$T_{\text{BO}}(p) = [V + P(pI - D + HP)^{-1}(E - HV)] C(pI - A)^{-1}B,$$

dont le tracé fréquentiel permet de mesurer les marges de robustesse.

## 6 Applications des observateurs

Outre la reconstruction de l'état pour élaborer une commande par retour d'état, dont nous avons abondamment parlé dans les chapitres précédents, nous allons voir ici une autre application importante des observateurs en diagnostic et détection. Dans cet optique on utilise l'observateur pour générer des résidus permettant d'élaborer une décision dans un étage de surveillance et diagnostic du système lorsque des perturbations ou des défauts apparaissent sur

un processus. On a en effet des variables qui agissent sur le système mais qui ne peuvent être mesurées, et l'objectif consiste ici à construire des résidus qui, selon le cas, doivent être sensibles aux défauts ou insensibles aux perturbations, pour ne citer que deux exemples. Nous verrons que ces observateurs utilisent les principes de base des observateurs de Luenberger avec des conditions structurelles supplémentaires. Nous n'utiliserons ici que des observateurs d'ordre complet, les techniques de réduction de la taille des observateurs pouvant bien sûr être utilisées.

Le principe général utilisé consiste à partir du modèle :

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) + \Phi(t), \\ y(t) &= Cx(t) + \Psi(t),\end{aligned}$$

où, en plus des termes habituels,  $\Phi(t)$  et  $\Psi(t)$  représentent, par exemple, l'effet des perturbations, d'entrées non mesurées, des défauts de capteurs ou d'actionneurs. Le but ici étant de détecter ces influences, on utilise un observateur de Luenberger du résidu :

$$\begin{aligned}\dot{z}(t) &= Dz(t) + Hu(t) + Ey(t), \\ r(t) &= Pz(t) + Vy(t),\end{aligned}$$

où le résidu  $r(t)$  doit vérifier, si on souhaite être insensible à cette influence,  $\lim_{t \rightarrow \infty} r(t) = 0$ , et  $\lim_{t \rightarrow \infty} r(t) \neq 0$ , dans le cas contraire. D'autre part, nous ne décrivons que le cas des systèmes continus en gardant bien à l'esprit que tout ce que nous verrons peut être traduit dans un cadre discret.

## 6.1 Observateurs à entrées inconnues

Considérons le système modélisé par l'équation d'état :

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) + Rf(t), \\ y(t) &= Cx(t),\end{aligned}$$

où  $f(t)$  est une entrée inconnue (non mesurée) et  $R$  une matrice constante. Utilisons l'observateur :

$$\begin{aligned}\dot{z}(t) &= Dz(t) + Hu(t) + Ey(t), \\ \hat{x}(t) &= z(t) - My(t).\end{aligned}$$

En introduisant la matrice  $P = I + MC$  et l'erreur d'observation  $\tilde{x}(t) = x(t) - \hat{x}(t)$ , on obtient :

$$\begin{aligned}\tilde{x}(t) &= Px(t) - z(t), \\ r(t) &= y(t) - \hat{y}(t) = C\tilde{x}(t), \\ &= CPx(t) - Cz(t), \\ &= (I + CM)y(t) - Cz(t),\end{aligned}$$

cette dernière expression étant celle utilisée pour générer les résidus. Pour étudier le comportement asymptotique du résidu on obtient :

$$\begin{aligned}\dot{\tilde{x}}(t) &= P\tilde{x}(t) - \dot{z}(t), \\ &= D\tilde{x}(t) + (PA - DP - EC)x(t) + (PB - H)u(t) + PRf(t).\end{aligned}$$

Pour que  $\lim_{t \rightarrow \infty} r(t) = 0$ , pour toutes les entrées, connues et inconnues, et pour tout état initial, il est nécessaire que les conditions suivantes soient vérifiées :

1.  $D$  est une matrice de Hurwitz;
2.  $PA - DP = EC$ ;
3.  $H = PB$ ;
4.  $PR = 0$ .

Il s'agit donc de trouver un ensemble de matrices  $(D, E, M)$  telles que soient vérifiées les 4 contraintes précédentes. Pour résoudre ce système on traite d'abord la relation 4, ce qui avec  $P = I + MC$ , s'écrit :

$$MCR = -R.$$

D'après l'annexe C, cette équation est soluble si et seulement si :

$$\text{rang}(CR) = \text{rang}(R),$$

et dans ce cas une solution possible (non unique) s'écrit :

$$M = -R(CR)^+,$$

où  $A^+$  désigne la pseudoinverse de  $A$ . Dans le cas où  $CR$  est de rang plein en lignes, l'expression de  $(CR)^+$  est  $R^T C^T (CRR^T C^T)^{-1}$  et on obtient :

$$M = -RR^T C^T (CRR^T C^T)^{-1}.$$

Après avoir choisi  $D$ , la valeur de  $M$  conduit à déterminer  $E$  comme solution du système :

$$EC = A - D + MCA - DMC = F.$$

Ce système est soluble si et seulement  $F^T \in \text{Im}(C^T)$ , et lorsque cette condition est vérifiée on obtient une solution possible sous la forme :

$$E = FC^+,$$

Comme généralement  $C$  est de rang plein en lignes,  $C^+ = C^T (CC^T)^{-1}$ .

### 6.1.1 Estimation des entrées inconnues

Dans certains cas il peut être intéressant d'estimer les entrées inconnues  $f(t)$ . Pour obtenir cette estimation on peut utiliser la décomposition en valeurs singulières de  $R$  :

$$R = U_R \Sigma_R V_R^T,$$

puis en posant  $\phi(t) = V_R^T f(t)$  obtenir avec  $x(t) \simeq \hat{x}(t)$  :

$$\Sigma_R \phi(t) = U_R^T (\hat{\dot{x}}(t) - Az(t) + AMy(t) - Bu(t)).$$

Comme :

$$P\dot{\hat{x}}(t) = \dot{z}(t) = Dz(t) + Hu(t) + Ey(t),$$

si  $P$  est inversible, on obtient :

$$\Sigma_R \phi(t) = U_R^T ((P^{-1}D - A)z(t) + (AM + P^{-1}E)y(t) + (P^{-1}H - B)u(t)).$$

On pourra estimer toutes les entrées inconnues si  $\Sigma_R$  est de rang plein en colonnes.

### 6.1.2 Mesures dépendant des entrées inconnues

Les développements précédents ont supposé que les entrées inconnues n'intervenaient pas sur les mesures. Dans le cas où la sortie  $y(t)$  dépend de ces entrées sous la forme :

$$y(t) = Cx(t) + Sf(t),$$

on utilise la décomposition en valeurs singulières de  $S$  :

$$S = U_S \Sigma_S V_S^T,$$

où, en notant  $s = \text{rang}(S)$  :

$$\Sigma_S = \begin{bmatrix} \text{diag}_{i=1}^s \{\sigma_{S,i}\} & 0 \\ 0 & 0 \end{bmatrix}.$$

En notant :

$$\begin{aligned} \phi(t) &= V_S^T f(t) = \begin{bmatrix} \phi_1(t) \\ \phi_2(t) \end{bmatrix}, & \gamma(t) &= U_S^T y(t) = \begin{bmatrix} \gamma_1(t) \\ \gamma_2(t) \end{bmatrix}, \\ \bar{C} &= U_S^T C, & \bar{R} &= R V_S, \end{aligned}$$

où  $\phi_1(t)$  et  $\gamma_1(t)$  sont des vecteurs de dimension  $s$ , l'équation d'état initiale s'écrit :

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + \bar{R}_1 \phi_1(t) + \bar{R}_2 \phi_2(t), \\ \gamma_1(t) &= \bar{C}_1 x(t) + [\text{diag}_{i=1}^s \{\sigma_{S,i}\}] \phi_1(t), \\ \gamma_2(t) &= \bar{C}_2 x(t). \end{aligned}$$

La relation intermédiaire fournissant :

$$\phi_1(t) = [\text{diag}_{i=1}^s \{\sigma_{s,i}\}]^{-1} (\gamma_1(t) - \bar{C}_1 x(t)),$$

on obtient le modèle sans entrées inconnues sur les mesures :

$$\begin{aligned} \dot{x}(t) &= [A - \bar{R}_1 [\text{diag}_{i=1}^s \{\sigma_{s,i}\}]^{-1} \bar{C}_1] x(t) \\ &\quad + [B \quad \bar{R}_1 [\text{diag}_{i=1}^s \{\sigma_{s,i}\}]^{-1}] \begin{bmatrix} u(t) \\ \phi_1(t) \end{bmatrix} + \bar{R}_2 \phi_2(t), \\ \gamma_2(t) &= \bar{C}_2 x(t). \end{aligned}$$

Il est bien évident que l'on peut se ramener à un tel modèle lorsque  $s$  est inférieur strictement à la dimension de  $y(t)$ , ce qui, en d'autres termes signifie que l'on peut trouver des combinaisons linéaires de mesures qui sont indépendantes des entrées inconnues.

## 6.2 Détection de défauts

Bien que basé sur un principe similaire aux observateurs à entrées inconnues, l'objectif est ici complètement différent puisque les résidus doivent être sensibles à ces défauts et même doivent permettre de les retrouver et dans ce cas on parle alors de l'isolation des défauts.

Considérons le modèle :

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B(u(t) + b(t)), \\ y(t) &= Cx(t) + m(t), \end{aligned}$$

où  $b(t)$  représente une défaillance des actionneurs,  $m(t)$ , une défaillance des capteurs. L'utilisation d'un observateur d'ordre plein :

$$\begin{aligned} \dot{\hat{x}}(t) &= (A - LC)\hat{x}(t) + Bu(t) + Ly(t), \\ \hat{y}(t) &= C\hat{x}(t), \end{aligned}$$

conduit au résidu :

$$\begin{aligned} \dot{\tilde{x}}(t) &= (A - LC)\tilde{x}(t) + Bb(t) - Lm(t), \\ \tilde{y}(t) &= C\tilde{x}(t) + m(t). \end{aligned}$$

Si  $L$  est telle que  $A - LC$  soit une matrice de Hurwitz, le résidu  $\tilde{y}(t)$  tend bien vers 0 en l'absence de défauts. Le transfert entre les défauts et les résidus s'écrit :

$$\tilde{y}(t) = C(pI - A + LC)^{-1} Bb(t) + [I - C(pI - A + LC)^{-1} L] m(t),$$

ce qui, en tenant compte du lemme d'inversion :

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

peut se mettre sous la forme :

$$[I + C(pI - A)^{-1}L] \tilde{y}(t) = C(pI - A)^{-1}Bb(t) + m(t).$$

À partir de cette relation on peut construire :

- en l'absence de défaut actionneur, le système permettant d'isoler les défauts capteurs à partir des résidus :

$$\begin{aligned} \dot{\eta}(t) &= A\eta(t) + L\tilde{y}(t), \\ m(t) &= C\eta(t) + \tilde{y}(t); \end{aligned}$$

- en l'absence de défaut capteur, le système permettant d'isoler les défauts actionneurs à partir des résidus est construit à partir de la variable  $\xi(t)$  définie par :

$$\xi(t) = C(pI - A)^{-1}Bb(t).$$

Cette variable est estimée à partir du même système que précédemment, c'est-à-dire :

$$\begin{aligned} \dot{\eta}(t) &= A\eta(t) + L\tilde{y}(t), \\ \xi(t) &= C\eta(t) + \tilde{y}(t), \end{aligned}$$

et l'estimation d'une défaillance capteur est donnée par l'inversion du modèle initial du système à savoir :

$$\beta(t) = \Delta(p) [C(pI - A)^{-1}B]^{-1} \xi(t),$$

où  $\Delta(p) = \text{diag} \left\{ \frac{1}{(p+\alpha)^i} \right\}$  est la matrice de filtres telle que  $\Delta(p) [C(pI - A)^{-1}B]^{-1}$  soit bicausale.  $\beta(t)$  n'est pas une estimation des défauts mais plutôt d'un filtrage des défauts, cependant le caractère diagonal de  $\Delta(p)$  permet l'isolation.

## 7 Annexe : Algorithme de Luenberger

Cet algorithme permet de mettre sous une forme canonique, commandable ou observable suivant le cas, une équation d'état d'un système commandable et observable :

$$\begin{aligned} \dot{x} &= Ax + Bu, \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \\ y &= Cx, \quad y \in \mathbb{R}^l. \end{aligned}$$

La construction de la forme observable étant duale de la construction de la forme commandable, nous détaillerons la première et obtiendrons la deuxième par transposition.

## 7.1 Forme canonique commandable

Le test préliminaire de commandabilité à effectuer s'écrit :

$$\text{rang} [ B \ AB \ A^2B \ \dots \ A^{n-1}B ] = n,$$

et si cette condition est vérifiée, la construction de la forme canonique commandable demande les étapes suivantes.

1. Dans la matrice de commandabilité :

$$C_{(A,B)} = [ B \ AB \ A^2B \ \dots \ A^{n-1}B ],$$

en allant de gauche à droite ( $\rightarrow$ ), on sélectionne une colonne si elle n'est pas linéairement dépendante des précédentes.

2. En notant  $B$  sous la forme :

$$B = [ b_1 \ b_2 \ \dots \ b_{m-1} \ b_m ]$$

où les  $b_i$  sont des vecteurs de dimension  $n$ , on calcule les  $m$  indices de commandabilité  $\mu_i$ , pour  $i = 1, \dots, m$ , comme les plus petit entiers  $k$  tels que  $A^k b_i$  soit linéairement dépendante des colonnes précédentes. On calcule également les indices cumulés de commandabilité  $\sigma_i$  :

$$i = 1, \dots, m, \quad \sigma_i = \sum_{j=1}^i \mu_j.$$

3. On réordonne les colonnes sélectionnées au pas 1 pour construire la matrice :

$$V = b_1 \ Ab_1 \ \dots \ A^{\mu_1-1}b_1 \ b_2 \ Ab_2 \ \dots \ A^{\mu_2-1}b_2 \ \dots \ \dots \ b_m \ Ab_m \ \dots \ A^{\mu_m-1}b_m .$$

4. On calcule  $V^{-1}$ , d'où l'on extrait les  $\sigma_i$ -ièmes lignes. Soient  $L_i$ ,  $i = 1, \dots, m$ , ces lignes.
5. On construit la matrices de changement de variables :

$$T_C = \begin{bmatrix} L_1 \\ L_1 A \\ \vdots \\ L_1 A^{\mu_1-1} \\ L_2 \\ L_2 A \\ \vdots \\ L_2 A^{\mu_2-1} \\ \vdots \\ L_m \\ L_m A \\ \vdots \\ L_m A^{\mu_m-1} \end{bmatrix} .$$



où  $H_C$  est la matrice  $(m \times m)$  formée des lignes significatives de  $B_C$  :

$$H_C = \begin{bmatrix} 1 & \times & \times & \cdots & \times \\ 0 & 1 & \times & \cdots & \times \\ & & \vdots & & \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix},$$

on arrive finalement à l'équation d'état canonique commandable :

$$\begin{aligned} \dot{x}_C &= A_C x_C + \bar{B}_C u_C, \\ y &= C_C x_C, \end{aligned}$$

avec :

$$\bar{B}_C = B_C H_C^{-1} = \begin{bmatrix} & & & & \uparrow \\ & & & & \mu_1 \\ & 1 & 0 & \cdots & 0 \\ & & & & \downarrow \\ & & & & \uparrow \\ & & & & \mu_2 \\ & & & & \downarrow \\ & & & \vdots & \\ & & & & \uparrow \\ & & & & \mu_m \\ 0 & \cdots & 0 & 0 & 1 \\ & & & & \downarrow \end{bmatrix}.$$

## 7.2 Forme canonique observable

Le test préliminaire d'observabilité à effectuer s'écrit :

$$\text{rang} \begin{bmatrix} A \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} = n.$$

et si cette condition est vérifiée, la construction de la forme canonique observable sera construite en partant du système dual :

$$\begin{aligned} \dot{\zeta} &= A^T \zeta + C^T v, \\ \eta &= B^T \zeta, \end{aligned}$$

sur lequel l'application de l'algorithme précédent conduit à la forme canonique commandable du système dual. Une dernière transposition conduit à la forme canonique observable du système

initial :

$$\begin{aligned} \dot{x}_0 &= A_0 x_0 + B_0 u, \\ y_0 &= \bar{C}_0 x_0, \end{aligned}$$

où  $x_0 = T_0 x$ ,  $y_0 = H_0 y$ , et :

$$A_0 = T_0 A T_0^{-1} = \begin{bmatrix} 1 & \times & & \times & & \times & \uparrow \\ & \times & & \times & \dots & \times & \nu_1 \\ & \ddots & \vdots & \vdots & & \vdots & \downarrow \\ & & 1 & \times & & \times & \uparrow \\ & & \times & \times & & \times & \nu_2 \\ & & \times & 1 & \times & \times & \downarrow \\ & & \vdots & \ddots & \vdots & \vdots & \uparrow \\ & & \times & & 1 & \times & \nu_l \\ & & \times & & \times & \times & \downarrow \\ & & \vdots & & \vdots & \dots & \uparrow \\ & & \times & & \times & 1 & \nu_l \\ & & \times & & \times & \dots & \downarrow \\ & & \vdots & & \vdots & \ddots & \uparrow \\ & & \times & & \times & 1 & \nu_l \\ & & \times & & \times & & \downarrow \\ \leftarrow \nu_1 \rightarrow & & \leftarrow \nu_2 \rightarrow & & \leftarrow \nu_l \rightarrow & & \end{bmatrix},$$

$$B_0 = T_0 B,$$

$$\bar{C}_0 = H_0^{-1} C T_0^{-1} = \begin{bmatrix} & & & & & & & 0 \\ & & & & & & & 0 \\ & & & & & & & \vdots \\ & & & & & & & 0 \\ & & & & & & & 1 \\ \leftarrow \nu_1 \rightarrow & & \leftarrow \nu_2 \rightarrow & & \leftarrow \nu_l \rightarrow & & & \end{bmatrix}.$$

Les indices  $\nu_i$ ,  $i = 1, \dots, l$ , mis en évidence pour la construction de cette forme s'appellent les indices d'observabilité.

## 8 Annexe B : Formes canoniques de matrices

### 8.1 Forme de Smith

Considérons les matrices élémentaires régulières suivantes qui traduisent des opérations élémentaires sur les lignes (resp. colonnes) d'une matrice, lorsqu'elles la multiplient à gauche



**Théorème 3** *Toute matrice  $A(m \times n)$  de rang  $r$  est équivalente à sa forme de Smith  $S(m \times n)$  (17).*

C'est-à-dire qu'il existe deux matrices régulières  $P$  et  $Q$ , non uniques telles que  $S = PAQ$ . Ce principe s'étend utilement aux matrices polynomiales (cf. chap. 11). Comme  $\text{rang } A = \text{rang } S$ , on en déduit que deux matrices de mêmes dimensions sont équivalentes si et seulement si elles ont même rang.

## 8.2 Décomposition en valeurs singulières

Plus utilisée en pratique que la décomposition sous forme de Jordan ou de Smith car plus robuste d'un point de vue numérique, la décomposition en valeurs singulières d'une matrice met en évidence une matrice diagonale qui lui est équivalente. Elle jouera un rôle fondamental dans la généralisation au cas multivariable du lieu fréquentiel d'un système linéaire. Dans toute la suite  $\|\cdot\|$  désignera la norme spectrale  $\|\cdot\|_2$  :

$$\|A\|_2 = \sqrt{\rho_{A^*A}},$$

où  $\rho_A$  désigne le plus grand module obtenu sur l'ensemble des valeurs propres de  $A$

**Théorème 4** *Pour toute matrice complexe  $A(m \times n)$  il existe deux matrices unitaires  $U(m \times m)$  et  $V(n \times n)$ , telles que :*

$$U^*AV = \Sigma,$$

où  $\Sigma$  est une matrice réelle  $(m \times n)$  diagonale,  $\Sigma = \text{diag}\{\sigma_1, \dots, \sigma_p\}$ , avec  $p = \min\{m, n\}$ , et :

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0.$$

Les  $\sigma_i$  apparaissant dans ce théorème sont appelées les valeurs singulières de  $A$  d'où le nom de la décomposition en valeurs singulières de  $A$  sous la forme :

$$A = U\Sigma V^*.$$

Comme :

$$A^*A = V\Sigma^T\Sigma V^*, \text{ et } AA^* = U\Sigma\Sigma^T U^*,$$

où  $\Sigma^T\Sigma = \text{diag}\{\sigma_1^2, \dots, \sigma_p^2, 0, \dots, 0\}$  est une matrice  $(n \times n)$  et  $\Sigma\Sigma^T = \text{diag}\{\sigma_1^2, \dots, \sigma_p^2, 0, \dots, 0\}$ , est une matrice  $(m \times m)$ , où  $p = \min(m, n)$ , les valeurs singulières non nulles sont les racines carrées des valeurs propres non nulles de la matrice  $A^*A$  mais aussi celles de la matrice  $AA^*$ . Notons  $U = [u_1 \ \dots \ u_m]$  et  $V = [v_1 \ \dots \ v_n]$ , alors les  $u_i$  et les  $v_i$ , d'après les expressions précédentes reçoivent une interprétation en termes de vecteurs propres de  $AA^*$  ou  $A^*A$  :

– pour  $i = 1, \dots, n$  :

$$A^*Av_i = \sigma_i^2 v_i,$$

– pour  $i = 1, \dots, m$  :

$$AA^*u_i = \sigma_i^2 u_i.$$

Ainsi lorsque  $X$  et  $Y$  sont des matrices unitaires de dimensions convenables on a :

$$\sigma_i(XAY) = \sigma_i(A).$$

Comme on a également :

$$AV = U\Sigma, \text{ et } A^*U = V\Sigma^T,$$

on obtient les relations suivantes, avec  $p = \min(m, n)$  :

– pour  $i = 1, \dots, p$  :

$$Av_i = \sigma_i u_i, \text{ et } A^*u_i = \sigma_i v_i,$$

– pour  $i = p + 1, \dots, \max(m, n)$  :

$$Av_i = 0, \text{ et } A^*u_i = 0.$$

Géométriquement, les valeurs singulières d'une matrice sont les longueurs des demi-axes de l'hyper-ellipsoïde défini par :

$$\{y, y = Ax, x \in \mathbb{C}^n, \|x\| = 1\},$$

et  $v_i$  leur direction. Ainsi  $v_1$  est la direction dans laquelle  $\|Ax\|$  est la plus grande sur  $\|x\| = 1$  et  $v_n$ , la direction dans laquelle  $\|Ax\|$  est la plus petite.

**Théorème 5** Soient  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = 0$ , alors, en désignant par  $\text{span}$  l'espace vectoriel engendré par un ensemble de vecteurs :

1.  $\text{rang } A = r$ ;
2.  $\mathcal{N}(A) = \text{span}\{v_{r+1}, \dots, v_n\}$  et  $[\mathcal{N}(A)]^\perp = \text{span}\{v_1, \dots, v_r\}$ ;
3.  $\mathcal{I}(A) = \text{span}\{u_1, \dots, u_r\}$  et  $[\mathcal{I}(A)]^\perp = \text{span}\{u_{r+1}, \dots, u_m\}$ ;
4. En posant  $U_r = [u_1 \ \dots \ u_r]$ ,  $V_r = [v_1 \ \dots \ v_n]$ , et  $\Sigma_r = \text{diag}\{\sigma_1, \dots, \sigma_p\}$ , on a :

$$A = \sum_{i=1}^r \sigma_i u_i v_i^* = U_r \Sigma_r V_r^*.$$

5.

$$\sum_{i=1}^r \sigma_i^2 = \text{trace}(A^*A)$$

Le résultat suivant est intéressant car il permet de fournir la meilleure approximation d'une matrice par une matrice de rang donnée :

**Théorème 6** Soit  $k < \text{rang } A$ , et  $A_k = \sum_{i=1}^k \sigma_i u_i v_i^*$ , alors :

$$\min_{B, \text{rang}(B) \leq k} \|A - B\| = \|A - A_k\| = \sigma_{k+1}.$$

Pour toutes matrices  $A$  et  $B$ , dont les tailles sont compatibles, on a :

$$\begin{aligned} \sigma_{i+j+1}(A+B) &\leq \sigma_{i+1}(A) + \sigma_{j+1}(B), \\ \sigma_{i+j+1}(AB) &\leq \sigma_{i+1}(A)\sigma_{j+1}(B). \end{aligned}$$

Notons  $\bar{\sigma}(A)$  et  $\underline{\sigma}(A)$  la plus grande et la plus petite valeur singulière d'une matrice  $A$ , d'après la partie précédente on a :

$$\bar{\sigma}(A) = \|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2,$$

et :

$$\underline{\sigma}(A) = \min_{\|x\|_2=1} \|Ax\|_2.$$

On dispose des propriétés suivantes :

– si  $\lambda(A)$  est une valeur propre de  $A$ , alors :

$$\underline{\sigma}(A) \leq |\lambda(A)| \leq \bar{\sigma}(A);$$

– si  $A^{-1}$  existe, alors :

$$\underline{\sigma}(A) = \frac{1}{\bar{\sigma}(A^{-1})} = \frac{1}{\|A^{-1}\|_2};$$

– pour tout réel  $\alpha$  :

$$\bar{\sigma}(\alpha A) = |\alpha| \bar{\sigma}(A), \underline{\sigma}(\alpha A) = |\alpha| \underline{\sigma}(A);$$

–

$$\text{cond}_2(A) = \frac{\bar{\sigma}(A)}{\underline{\sigma}(A)}.$$

Lorsque les tailles sont compatibles :

$$|\bar{\sigma}(A) - \bar{\sigma}(B)| \leq \bar{\sigma}(A+B) \leq \bar{\sigma}(A) + \bar{\sigma}(B),$$

$$\max \{ \underline{\sigma}(A)\bar{\sigma}(B), \bar{\sigma}(A)\underline{\sigma}(B) \} \leq \bar{\sigma}(AB) \leq \bar{\sigma}(A)\bar{\sigma}(B),$$

$$\max \{ \underline{\sigma}(A) - \bar{\sigma}(B), \underline{\sigma}(B) - \bar{\sigma}(A) \} \leq \underline{\sigma}(A+B) \leq \min \{ \underline{\sigma}(A) + \bar{\sigma}(B), \bar{\sigma}(A) + \underline{\sigma}(B) \},$$

$$\max \{ 0, \bar{\sigma}(A) - 1 \} \leq \bar{\sigma}(A+I), \max \{ 0, \underline{\sigma}(A) - 1 \} \leq \underline{\sigma}(A+I),$$

$$\max \{ \bar{\sigma}(A), \bar{\sigma}(B) \} \leq \bar{\sigma} \begin{bmatrix} A & B \end{bmatrix} \leq \sqrt{2} \max \{ \bar{\sigma}(A), \bar{\sigma}(B) \}.$$

Dans le cas où on travaille sur l'ensemble des matrices réelles, la décomposition en valeurs singulières s'écrit :

$$A = U\Sigma V^T,$$

où  $\Sigma$  conserve le même sens qu'auparavant et  $U$  et  $V$  sont deux matrices orthogonales, c'est-à-dire qu'elles vérifient :

$$\begin{aligned} UU^T &= U^T U = I, \\ VV^T &= V^T V = I. \end{aligned}$$

## 9 Annexe C : Inverses généralisées et systèmes linéaires

### 9.1 Inverse généralisée

Une matrice régulière  $A(n \times n)$  admet une inverse unique  $A^{-1}$  telle que,  $AA^{-1} = A^{-1}A = I$ , ce qui donne la relation :  $AA^{-1}A = A$ . Par extension, l'inverse généralisée d'une matrice quelconque  $A$  est la matrice  $X$  qui vérifie la relation :

$$AXA = A.$$

Soit  $S$  la forme de Smith d'une matrice  $A$  de rang  $r$  et les matrices régulières  $P$  et  $Q$  telles que  $A = PSQ$ .

**Théorème 7**  $Q^{-1}S^T P^{-1}$  est une inverse généralisée de  $A$ .

De façon plus générale on peut montrer que la forme d'une inverse généralisée  $X$  de  $A(m \times n)$  peut être construite à partir de sa forme de Smith par  $X = Q^{-1}TP^{-1}$ , où  $T$  est une matrice  $(n \times m)$  qui est de la forme :

$$T = \begin{bmatrix} I_r & X_1 \\ X_2 & X_3 \end{bmatrix},$$

où  $X_1, X_2$  et  $X_3$  sont des matrices quelconques de dimensions convenables.

**Théorème 8** Toute inverse généralisée  $X$  de  $A$ , possède les propriétés suivantes :

1.  $\mathcal{I}(AX) = \mathcal{I}(A)$ ;
2.  $\mathcal{I}((XA)^T) = \mathcal{I}(A^T)$ ;
3.  $\mathcal{I}(I - XA) = \mathcal{N}(A)$ ;
4.  $AX$  et  $XA$  sont idempotentes.

## 9.2 Application aux systèmes linéaires

Dans beaucoup d'applications, on doit résoudre le système linéaire  $Ax = y$ , où  $A(m \times n)$  n'est pas carrée ou bien a un déterminant nul. Le système peut alors admettre plusieurs solutions ou aucune. Dans le cas où ce système est soluble, on dit qu'il est compatible.

**Théorème 9** *Le système  $Ax = y$  est soluble si et seulement si  $\text{rang}(A) = \text{rang}([A \ y])$ .*

Supposons qu'il existe une matrice  $X(n \times m)$  telle que  $AXA = A$ , alors :

$$x = Xy + (I - XA)z,$$

où  $z$  est un vecteur arbitraire, sera une solution du système si et seulement si :

$$AXy = y.$$

Cette dernière relation constitue une relation de compatibilité du système.

**Théorème 10** *Pour le système compatible  $y = Ax$ , en notant  $A^{\{1\}}$  une inverse généralisée de  $A$ , l'ensemble des solutions est donné par :*

$$x = A^{\{1\}}y + (I_n - A^{\{1\}}A)z,$$

où  $z$  est un vecteur arbitraire de  $\mathbb{R}^n$ .

Considérons le système linéaire matriciel :

$$AXB = C,$$

où  $A(m \times n)$ ,  $B(p \times q)$  et  $C(m \times q)$  sont données et  $X(n \times p)$  est une matrice inconnue à déterminer. Comme :

$$AXB = AA^{\{1\}}AXB B^{\{1\}}B,$$

une condition de compatibilité s'écrit :

$$AA^{\{1\}}CB^{\{1\}}B = C.$$

**Théorème 11** *La solution générale du système compatible  $AXB = C$ , s'écrit :*

$$X = A^{\{1\}}CB^{\{1\}} + Y - A^{\{1\}}AYBB^{\{1\}},$$

où  $Y$  est une matrice arbitraire ( $n \times p$ ).

Soit  $A\{1\}$  l'ensemble des inverses généralisées d'une matrice  $A$ , l'application du résultat précédent à l'équation :

$$AXA = A,$$

permet de caractériser  $A\{1\}$  à partir d'un de ses éléments. En effet, comme la solution générale en est :

$$X = A^{\{1\}}AA^{\{1\}} + Y - A^{\{1\}}AYA A^{\{1\}},$$

il suffit de poser  $Y = A^{\{1\}} + Z$ , pour obtenir :

$$A\{1\} = \{A^{\{1\}} + Z - A^{\{1\}}AZAA^{\{1\}}, Z \in \mathbb{R}^{n \times m}\}.$$

## 9.3 Pseudo-inverse

### 9.3.1 Notion de pseudo-inverse

Comme l'inverse généralisée n'est pas unique, on cherche des matrices plus particulières vérifiant les 4 propriétés :

$$\begin{aligned} \text{P1} & : AXA = A; \\ \text{P2} & : XAX = X; \\ \text{P3} & : (AX)^T = AX; \\ \text{P4} & : (XA)^T = XA. \end{aligned}$$

A partir de la forme de Smith de  $A$  (17),  $A = PSQ$ , toute matrice vérifiant les deux premières propriétés est de la forme  $Q^{-1}TP^{-1}$ , où  $T$  est de la forme :

$$T = \begin{bmatrix} I_r & X_1 \\ X_2 & X_2X_1 \end{bmatrix}.$$

Toute matrice  $X$  vérifiant ces 4 propriétés est une pseudo-inverse de  $A$  (on dit aussi inverse au sens de Moore-Penrose), que l'on notera  $A^+$ .

**Théorème 12**  $A^+$  existe et est unique.

### 9.3.2 Formes explicites de la pseudo-inverse

Soit  $A = FG$  une décomposition de rang maximal de  $A$ , alors :

$$\begin{aligned} A^+ & = G^T(F^T A G^T)^{-1} F^T, \\ & = G^T(GG^T)^{-1}(F^T F)^{-1} F^T. \end{aligned}$$

Soit la décomposition de  $A(m \times n)$  de rang  $r$  sous la forme :

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

où  $A_{11}$  est régulière de rang  $r$ . Notons :

$$A_1 = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}, \quad A_2 = [ A_{11} \quad A_{12} ].$$

Comme nécessairement  $A_{22} = A_{21}A_{11}^{-1}A_{12}$ , on a :

$$A^+ = A_2^T(A_1^T A A_2^T)^{-1} A_1^T.$$

Soit la décomposition en valeurs singulières de  $A$ ,  $A = U\Sigma V^T$ , où  $U$  et  $V$  sont orthogonales :

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_r = \text{diag}\{\sigma_1, \dots, \sigma_r\},$$

les  $\sigma_i$  étant les valeurs singulières non nulles de  $A$  telles que  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ . Alors, d'après les propriétés précédentes, on aura :

$$A^+ = V\Sigma^+U^T,$$

avec :

$$\Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

### 9.3.3 Quelques propriétés de la pseudo-inverse

1. Si  $A$  est régulière, alors  $A^+ = A^{-1}$ .
2. Si  $A$  est de rang plein en colonnes :  $A^+ = (A^T A)^{-1} A^T$ , et on obtient  $A^+ A = I_n$ , mais  $AA^+ \neq I_m$ .
3. Si  $A$  est de rang plein en lignes :  $A^+ = A^T (AA^T)^{-1}$ , et on obtient  $AA^+ = I_m$ , mais  $A^+ A \neq I_n$ .
4. Pour toute matrice  $A$  on a l'égalité suivante :

$$A^+ = (A^T A)^+ A^T = A^T (AA^T)^+.$$

5. Soit  $FG$ , une factorisation de rang maximal de  $A$ , alors  $A^+ = G^+ F^+$ . Dans le cas particulier où  $r = 1$ , on peut écrire :

$$A^+ = \frac{A^T}{F^T A G^T} = \frac{A^T}{G A^T F}.$$

6. Il faut noter qu'en général  $(AB)^+ \neq B^+ A^+$ .
7. Si  $P$  et  $Q$  sont deux matrices unitaires alors :

$$(PAQ)^+ = Q^{-1} A^+ P^{-1} = Q^T A^+ P^T.$$

8. Si  $P$  et  $Q$  sont seulement régulières alors  $(PAQ)^+ \neq Q^{-1} A^+ P^{-1}$ . En effet, d'après ce qui précède on a seulement :

$$(PAQ)^+ = Q^T G^T (G Q Q^T G^T)^{-1} (F P^T P F)^{-1} F P^T.$$

9. Soit  $A(m \times n) = FPG$ , où  $\text{rang } F(m \times r) = \text{rang } G(r \times n) = r = \text{rang } A$  et  $P(r \times r)$  régulière. Alors  $A^+ = G^+ P^{-1} F^+$ .
10. Pour toute matrice  $A$ ,  $(A^+)^+ = A$ ,  $(A^+)^T = (A^T)^+$ , et on dispose des formules d'Albert et Sittler :

$$A^+ = \lim_{\epsilon \rightarrow 0} (A^T A + \epsilon I)^{-1} A^T = \lim_{\epsilon \rightarrow 0} A^T (AA^T + \epsilon I)^{-1}.$$

### 9.3.4 Algorithme de Greville-Albert-Sittler

Une méthode efficace de calcul itératif de la pseudo-inverse d'une matrice est donné par l'algorithme de Greville (ou sous une forme équivalente par l'algorithme d'Albert et Sittler) qui ne demande pas de traitement préalable de la matrice. Cet algorithme est très utile pour trois raisons :

- il permet de calculer de façon récurrente la pseudo-inverse d'une matrice en un nombre fini de pas ;
- il est indépendant du conditionnement de la matrice (pas besoin de factorisation de la matrice) ;
- il ne demande pas d'inversion de matrices, seules des quantités scalaires sont à inverser.

C'est donc un algorithme que l'on implante aisément sur calculateur et il peut même être utilisé pour calculer l'inverse d'une matrice régulière. Sa structure est décrite dans l'organigramme de la figure dans lequel  $A$  désigne une matrice  $(m \times n)$ ,  $a_k$  la  $k$ -ième colonne de  $A$  et  $A_k$  la matrice formée des  $k$  premières colonnes de  $A$ .

### 9.3.5 Application aux systèmes linéaires incompatibles

Lorsque le système linéaire  $y = Ax$  est incompatible, il n'a pas de solution exacte. On cherche alors une solution de ce système au sens des moindres carrés, *i.e.* celle qui minimise  $\|y - Ax\|_2$ . On peut remarquer que résoudre un système compatible au sens des moindres carrés c'est obtenir sa solution exacte, on peut donc généraliser à tout système la notion de solution.

**Théorème 13**  $\hat{x}$  est la solution du système  $y = Ax$  au sens des moindres carrés si et seulement si  $\hat{x}$  est solution du système toujours compatible  $A^T y = A^T Ax$ .

Le système linéaire  $A^T y = A^T Ax$  s'appelle le système normal associé au système  $y = Ax$ .

**Théorème 14** La solution au sens des moindres carrés du système linéaire  $y = Ax$  s'écrit :

$$\hat{x} = A^+ y + (I - A^+ A)z,$$

où  $z$  est un vecteur quelconque et  $A^+$  la pseudo-inverse de  $A$ . De plus, la solution unique de norme euclidienne minimale est obtenue avec  $z = 0$ , soit :

$$\bar{x} = A^+ y.$$

### 9.3.6 Résolution d'un système linéaire

D'après tout ce que l'on vient de dire, la solution du système linéaire  $y = Ax$  s'écrit :

$$A^+ y + (I - A^+ A)z,$$

et si le système est compatible cette solution est exacte, sinon elle ne l'est qu'au sens des moindres carrés. Le fait de prendre  $z = 0$  conduit à une certaine simplification.

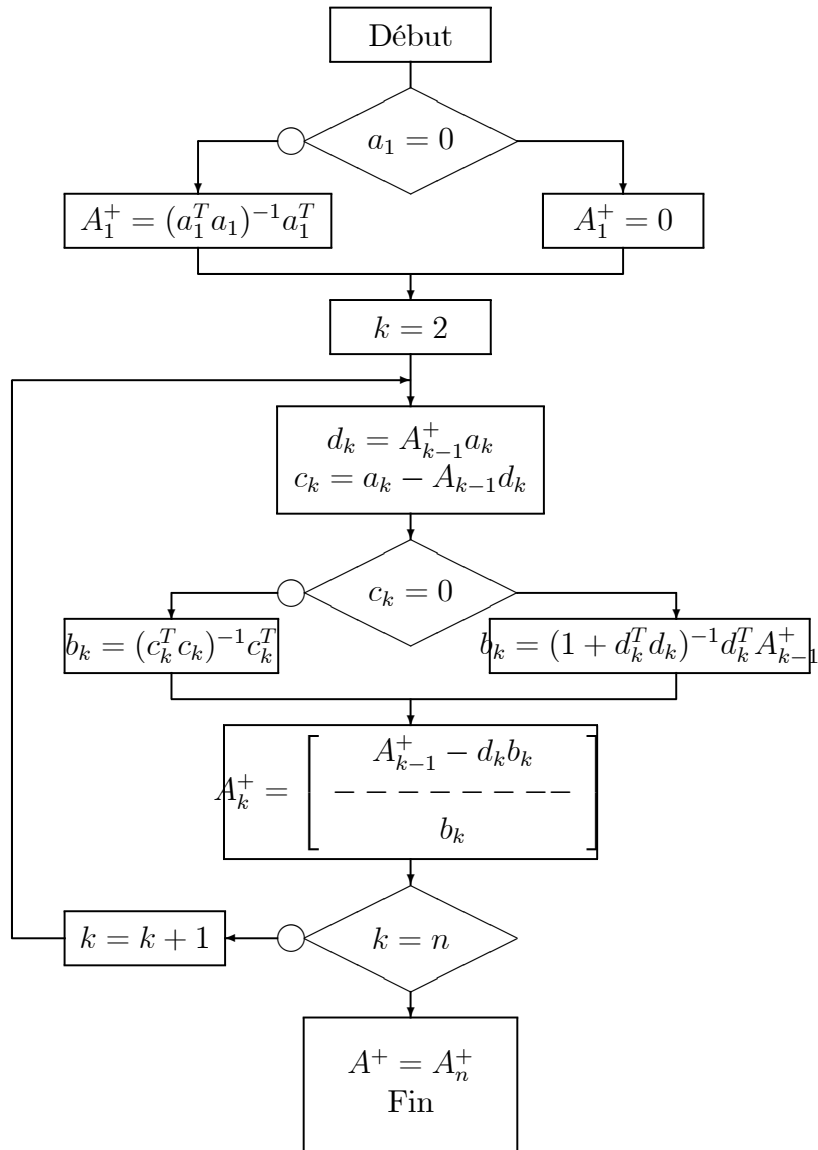


FIG. 4 – Algorithme de Greville du calcul de  $A^+$ .

## Références

- [1] Aström, K., Wittenmark, B., *Computer controlled systems*, Prentice Hall, 1984.
- [2] Borne, P., Dauphin-Tanguy, D., Richard, J.P., Rotella, F., Zambettakis, I., *Commande et optimisation des processus*, Technip, 1992.
- [3] Doyle, J.C., Stein, G., Robustness with observers, *IEEE Trans. Aut. Control*, AC-24, pp. 607–611, 1979.
- [4] Friedland, B., Observers, chap. 37, pp. 607–618, *The Control Handbook*, Levine, W.S., ed., CRC Press, 1996.
- [5] Kailath, T., *Linear systems*, Prentice-Hall, 1980.
- [6] Lewis, F.L., Syrmos, V.L., *Optimal control*, John Wiley & Sons, 1995.
- [7] Luenberger, D.G., Canonical forms for linear multivariable systems, *IEEE Trans. Aut. Control*, AC-12, pp. 290–293, 1967.
- [8] Luenberger, D.G., An introduction to observers, *IEEE Trans. Aut. Control*, AC-16, pp. 596–602, 1971.
- [9] Maquin, D., Ragot, J., *Diagnostic des systèmes linéaires*, Hermès, 2000.
- [10] Murdoch, Observer design for a linear functional of the state vector, *IEEE Trans. Aut. Control*, AC-18, pp.308–310, 1973.
- [11] O'Reilly, J., *Observers for linear systems*, Academic Press, 1983.
- [12] Rotella, F., Borne, P., *Théorie et pratique du calcul matriciel*, Technip, 1995.
- [13] Sirisena, H.R., Minimal order observers for linear functions of a state vector, *Int. J. Control*, 29, n.2, pp. 235–254, 1979.