

Chapitre VI

Méthodes d'identification

Version 1/12.11.2002

Chapitre 6. Méthodes d'identification

6.1 Méthodes d'identification basées sur le blanchissement de l'erreur de prédiction (type I)

6.1.1 Moindres carrés récurrents (M.C.R.)

6.1.2 Moindres carrés étendus (M.C.E.)

6.1.3 Maximum de vraisemblance récurrent (M.V.R.)

6.1.4 Erreur de sortie avec modèle de prédiction étendu (E.S.M.P.E.)

6.1.5 Moindres carrés généralisés (M.C.G.)

6.2 Validation des modèles identifiés avec les méthodes de type I

6.3 Méthodes d'identification basées sur la décorrélation du vecteur des observations et de l'erreur de prédiction (type II)

6.3.1 Variable instrumentale à modèle auxiliaire (VIMA)

6.3.2 Erreur de sortie à compensateur fixe (ESCF)

6.3.3 Erreur de sortie avec filtrage (adaptif) des observations (ESF(A)O)

6.4 Validation des modèles identifiés avec les méthodes de type II

6.5 Estimation de la complexité d'un modèle

6.5.1 Un exemple

6.5.2 Cas idéal (sans bruit)

6.5.3 Cas avec bruit

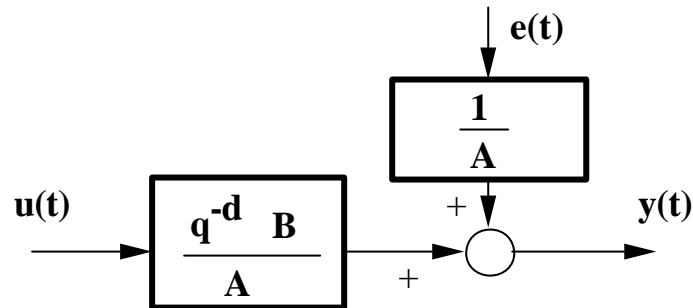
6.5.4 Critère d'estimation de complexité

6.6 Conclusion

6.7 Notes et indications bibliographiques

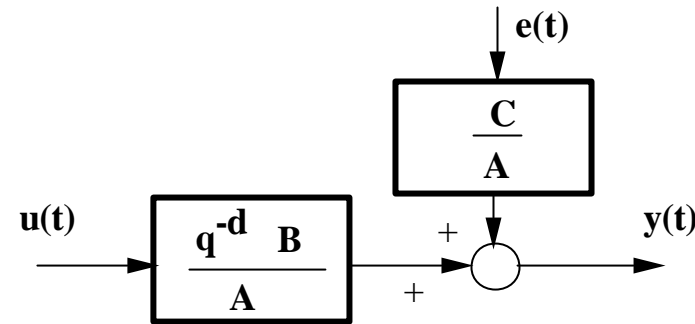
Structures « procédé + perturbation » et méthodes d'identification

$$S1: A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t) + e(t)$$



Moindres carrés récurrents (MCR)

$$S3: A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t) + C(q^{-1})e(t)$$

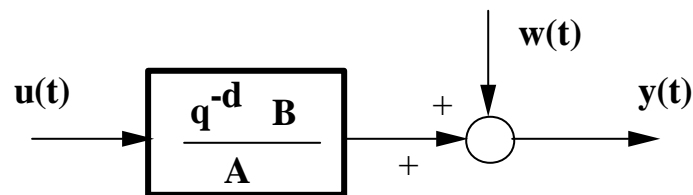


Moindres carrés étendus (MCE)

ES avec modèle de prédiction étendu (ESMPE)

Maximum de vraisemblance récursif (MVR)

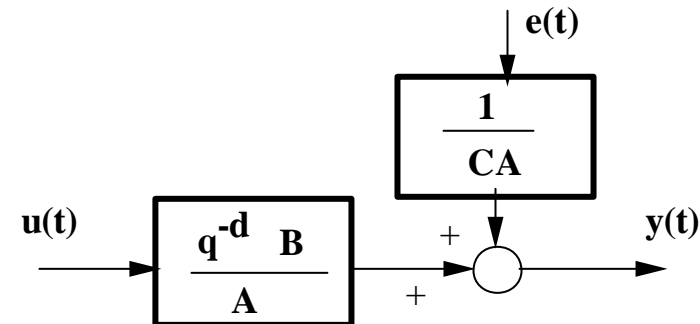
$$S2: A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t) + A(q^{-1})w(t)$$



Variable instrumentale...(VIMA)

Erreur de sortie (ESCF, ESFO, ESFAO)

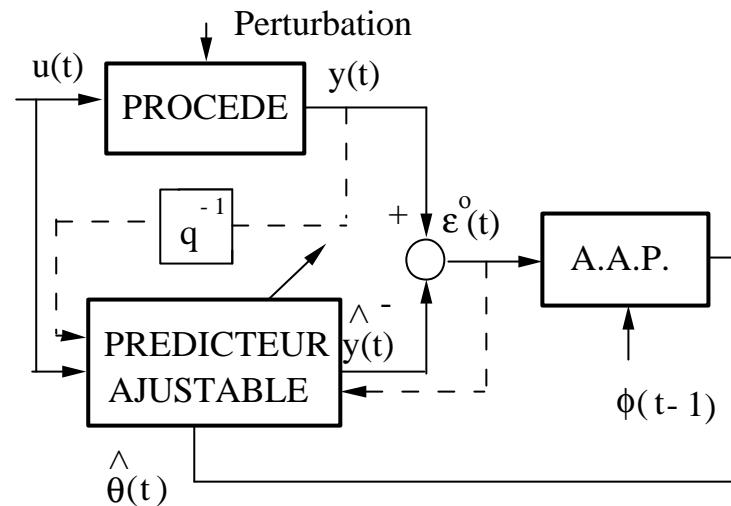
$$S4: A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t) + [1/C(q^{-1})]e(t)$$



Moindres carrés généralisés (MCG)

Méthodes d'identification récursive

Structure:



Algorithme
d'adaptation
paramétrique
(AAP)

$$\hat{\mathbf{q}}(t+1) = \hat{\mathbf{q}}(t) + F(t)\Phi(t)\mathbf{e}(t+1)$$

$$F^{-1}(t+1) = \mathbf{I}_1(t)F(t) + \mathbf{I}_2\Phi(t)\Phi(t)^T$$

$$0 < \mathbf{I}_1(t) \leq 1 ; 0 \leq \mathbf{I}_2(t) < 2 ; F(0) > 0$$

$$\mathbf{e}(t+1) = \frac{\mathbf{e}^0(t+1)}{1 + \Phi(t)^T F(t)\Phi(t)}$$

*Commun à toutes
les méthodes*

Méthodes d'identification basées sur le blanchissement de l'erreur de prédiction (type I)

- Moindres carrés récurrents (MCR) (*voir chapitre 5, transp.15 à 19*)
- Moindres carrés étendus (MCE)
- Maximum de vraisemblance récursif (MVR)
- Erreur de sortie avec modèle de prédiction étendu (ESMPE)
- Moindres carrés généralisés (MCG)

Moindres carrés étendus (MCE)

Idée : *Identification du modèle du procédé et de la perturbation (ARMAX) pour pouvoir obtenir une erreur de prédiction blanche*

Procédé + perturbation (ARMAX):

$$y(t+1) = -a_1 y(t) + b_1 u(t) + c_1 e(t) + e(t+1)$$

Prédicteur optimal (paramètres connus) – voir Chapitre 4

$$\hat{y}(t+1) = -a_1 y(t) + b_1 u(t) + c_1 e(t) \quad \leftarrow \text{On remplace } e(t) \text{ par } e(t)$$

Erreur de prédiction (paramètres connus) : $\mathbf{e}(t+1) = y(t+1) - \hat{y}(t+1) = e(t+1)$

Prédicteur ajustable (paramètres inconnus):

$$\hat{y}^o(t+1) = -\hat{a}_1(t)y(t) + \hat{b}_1(t)u(t) + \hat{c}_1(t)\mathbf{e}(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t) \quad (\text{à priori})$$

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t), \hat{b}_1(t), \hat{c}_1(t)] \quad ; \quad \mathbf{f}(t)^T = [-y(t), u(t), \mathbf{e}(t)]$$

$$\hat{y}(t+1) = -\hat{a}_1(t+1)y(t) + \hat{b}_1(t+1)u(t) + \hat{c}_1(t+1)\mathbf{e}(t) = \hat{\mathbf{q}}(t+1)^T \mathbf{f}(t)^T \quad (\text{à posteriori})$$

Moindres carrés étendus (MCE)

Erreur de prédiction (paramètres inconnus)

$$\mathbf{e}^o(t+1) = y(t+1) - \hat{y}^o(t+1) \quad ; \quad \mathbf{e}(t+1) = y(t+1) - \hat{y}(t+1)$$

AAP: On utilise l'algorithme donné transparent 4

Rem.: La taille de $\hat{\mathbf{q}}$ et Φ augmente par rapport au moindres carrés

Cas général :

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t) \dots \hat{a}_{n_A}(t), \hat{b}_1(t) \dots \hat{b}_{n_B}(t), \hat{c}_1(t) \dots \hat{c}_{n_C}(t)]$$

$$\Phi(t)^T = [-y(t) \dots -y(t - n_A + 1), u(t-d) \dots u(t-d - n_B + 1), \mathbf{e}(t) \dots \mathbf{e}(t - n_C + 1)]$$

Propriétés:

- $e(t)$ tend asymptotiquement vers un bruit blanc (donc estimation paramétrique non biaisée en présence d'une excitation riche)

- Condition suffisante de convergence: $\left(\frac{1}{C(z^{-1})} - \frac{I_2}{2} \right) ; \quad 2 > I_2 \geq \max I_2(t)$

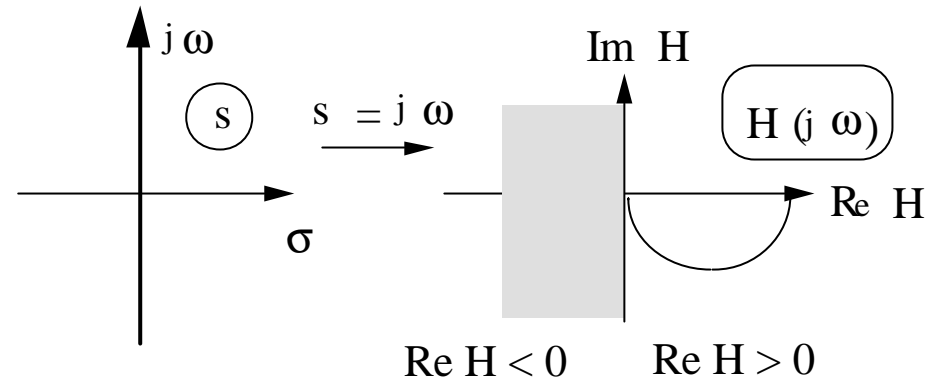
Peut expliquer la non-convergence
pour certains bruits

fonction de transfert strictement réelle positive

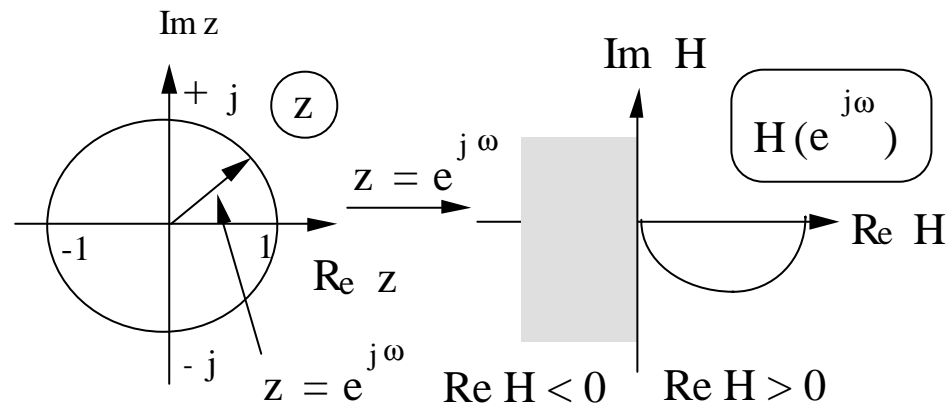
Voir fonction: *rels.sci(m)* sur le site web du livre

Fonction de transfert *strictement* réelle positive (SRP)

continu



discret



- asymptotiquement stable
- $\text{Re} H(e^{j\omega}) > 0$ pour tout $|e^{j\omega}| = 1$, ($0 < \omega < \pi$) (cas discret)

Maximum de vraisemblance récursif (MVR)

Idée: *La même que pour MCE + filtrage de $\mathbf{F}(t)$ par $C(q^{-1})$ (pour éliminer la condition SPR et pour une décorrélation plus rapide de \mathbf{F} et \mathbf{e})*

Procédé + perturbation (ARMAX):

$$y(t+1) = -a_1 y(t) + b_1 u(t) + c_1 e(t) + e(t+1)$$

Prédicteur ajustable (paramètres inconnus):

$$\hat{y}^o(t+1) = -\hat{a}_1(t)y(t) + \hat{b}_1(t)u(t) + \hat{c}_1(t)\mathbf{e}(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t) \quad (\text{à priori})$$

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t), \hat{b}_1(t), \hat{c}_1(t)] \quad ; \quad \mathbf{f}(t)^T = [-y(t), u(t), \mathbf{e}(t)]$$

Erreur de prédiction (a priori): $\mathbf{e}^o(t+1) = y(t+1) - \hat{y}^o(t+1)$

Estimation de $C(q^{-1})$ à t : $\hat{C}(t, q^{-1}) = 1 + \hat{c}_1(t)q^{-1}$

$$\Phi(t)^T = \mathbf{f}_f(t)^T = \frac{1}{\hat{C}(t, q^{-1})} [-y(t), u(t), \mathbf{e}(t)] = \left[\frac{-y(t)}{\hat{C}(t, q^{-1})}, \frac{u(t)}{\hat{C}(t, q^{-1})}, \frac{\mathbf{e}(t)}{\hat{C}(t, q^{-1})} \right]$$

AAP: On utilise l'algorithme donné transparent 4

Maximum de vraisemblance récursif (MVR)

- Le filtrage des observations, ne peut se faire que pour $\hat{C}(t, q^{-1})$ stable. Un test de stabilité doit être incorporé (existe dans WinPIM et dans les routines du site web)
- L'initialisation se fait par les MCE (pour obtenir une estimation de C)
Horizon d'initialisation : (5 à 8) x nb. des paramètres
- Transition graduelle de MCE vers MVR en utilisant un « facteur de contraction » (force les racines à l'intérieur du cercle unité)

$$\hat{C}(t, q^{-1}) = 1 + \hat{c}_1(t)q^{-1} \rightarrow \hat{C}(t, q^{-1}) = 1 + \mathbf{a}\hat{c}_1(t)q^{-1}; 0 \leq \mathbf{a} \leq 1$$

Facteur de contraction variable: $\mathbf{a}(t) = \mathbf{a}_0\mathbf{a}(t-1) + 1 - \mathbf{a}_0; \lim_{t \rightarrow \infty} \mathbf{a}(t) = 1$

Choix WinPIM: $0.5 \leq \mathbf{a}_0 = \mathbf{a}(0) \leq 0.9$

- Cas général:

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t) \dots \hat{a}_{n_A}(t), \hat{b}_1(t) \dots \hat{b}_{n_B}(t), \hat{c}_1(t) \dots \hat{c}_{n_C}(t)]$$

$$\Phi(t)^T = \frac{1}{\hat{C}(t, q^{-1})} [-y(t) \dots -y(t - n_A + 1), u(t-d) \dots u(t-d - n_B + 1), \mathbf{e}(t) \dots \mathbf{e}(t - n_C + 1)]$$

Erreur de sortie avec modèle de prédiction étendu (ESMPE)

Origine: *Extension des méthodes d'erreur de sortie.*

C'est une variante des MCE

Procédé + perturbation (ARMAX):

$$y(t+1) = -a_1 y(t) + b_1 u(t) + c_1 e(t) + e(t+1)$$

Prédicteur ajustable des MCE:

$$\hat{y}^o(t+1) = -\hat{a}_1(t) y(t) + \hat{b}_1(t) u(t) + \hat{c}_1(t) e(t) \pm \hat{a}_1(t) \hat{y}(t)$$

Prédicteur ajustable de ESMPE:

$$\hat{y}^o(t+1) = -\hat{a}_1(t) y(t) + \hat{b}_1(t) u(t) + \hat{h}_1(t) e(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t) \quad ; \quad \hat{h}_1(t) = \hat{c}_1(t) - \hat{a}_1(t)$$

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t), \hat{b}_1(t), \hat{c}_1(t)] \quad ; \quad \Phi(t) = \mathbf{f}(t)^T = [-\hat{y}(t), u(t), e(t)]$$

↖ Au lieu de $y(t)$ dans les MCE

Erreur de prédiction (a priori): $\mathbf{e}^o(t+1) = y(t+1) - \hat{y}^o(t+1)$

AAP: On utilise l'algorithme donné transparent 4

Erreur de sortie avec modèle de prédiction étendu (ESMPE)

Cas général:

$$\hat{\mathbf{q}}(t)^T = \left[\hat{a}_1(t) \dots \hat{a}_{n_A}(t), \hat{b}_1(t) \dots \hat{b}_{n_B}(t), \hat{h}_1(t) \dots \hat{h}_{n_C}(t) \right]$$

$$\Phi(t)^T = \left[-\hat{y}(t) \dots -\hat{y}(t - n_A + 1), u(t - d) \dots u(t - d - n_B + 1), \mathbf{e}(t) \dots \mathbf{e}(t - n_C + 1) \right]$$

- Même propriétés asymptotiques que les MCE
- Performances meilleures que les MCE (en général) pour fichiers E/S avec un nombre limité de données (rejet plus rapide du biais)

Voir fonction: *xoloe.sci(.m)* sur le site web du livre

Moindres carrés généralisés (MCG)

Idée: *Obtenir une erreur de prédiction « blanche » pour une perturbation $\left[1 / C(q^{-1})\right]e(t)$*

Procédé + perturbation (ARARX):

$$y(t+1) = -a_1 y(t) + b_1 u(t) + \frac{e(t+1)}{1 + c_1 q^{-1}}$$

$$\mathbf{a}(t+1) = (1 + a_1 q^{-1})y(t+1) - b_1 u(t) = \frac{e(t+1)}{1 + c_1 q^{-1}} \longrightarrow (1 + c_1 q^{-1})\mathbf{a}(t+1) = e(t+1)$$

Prédicteur optimal (paramètres connus)

$$\hat{y}(t+1) = -a_1 y(t) + b_1 u(t) - c_1 \mathbf{a}(t)$$

Erreur de prédiction (paramètres connus) : $\mathbf{e}(t+1) = y(t+1) - \hat{y}(t+1) = e(t+1)$

Moindres carrés généralisés (MCG)

Prédicteur ajustable (paramètres inconnus):

$$\hat{y}^o(t+1) = -\hat{a}_1(t)y(t) + \hat{b}_1(t)u(t) - \hat{c}_1(t)\mathbf{a}(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t)$$

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t), \hat{b}_1(t), \hat{c}_1(t)] \quad ; \quad \mathbf{f}(t)^T = [-y(t), u(t), -\mathbf{a}(t)]$$

Estimation de $\mathbf{a}(t)$:

$$\begin{aligned} \mathbf{a}(t) &= \hat{A}(t, q^{-1})y(t) - q^{-d} \hat{B}(t, q^{-1})u(t) \\ &= (1 + \hat{a}_1(t)q^{-1})y(t) - \hat{b}_1(t)u(t-1) \end{aligned}$$

Erreur de prédiction (a priori): $\mathbf{e}^o(t+1) = y(t+1) - \hat{y}^o(t+1)$

AAP: On utilise l'algorithme donné transparent 4 avec : $\Phi(t) = \mathbf{f}(t)$

Cas général:

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t) \dots \hat{a}_{n_A}(t), \hat{b}_1(t) \dots \hat{b}_{n_B}(t), \hat{c}_1(t) \dots \hat{c}_{n_C}(t)]$$

$$\Phi(t)^T = [-y(t) \dots -y(t-n_A+1), u(t-d) \dots u(t-d-n_B+1), -\mathbf{a}(t) \dots -\mathbf{a}(t-n_C+1)]$$

Moindres carrés généralisés (MCG)

Propriétés:

- *$e(t)$ tend asymptotiquement vers un bruit blanc (donc estimation paramétrique non biaisée en présence d'une excitation riche)*

- *Condition suffisante de convergence: $\left(C(z^{-1}) - \frac{I_2}{2} \right)$; $2 > I_2 \geq \max I_2(t)$*

fonction de transfert strictement réelle positive

Cette méthode s'utilise dans le cas des perturbations ayant un spectre fréquentiel étroit (ex.: perturbation presque harmonique).
(la modélisation ARARX requiert dans ce cas moins de paramètres dans $C(q^1)$ que la modélisation ARMAX)

Validation des modèles identifiés avec les méthodes de type I

Méthodes d'identification basées sur le blanchissement de l'erreur de prédiction

- Moindres carrés récurrents (MCR)
 - Moindres carrés étendus (MCE)
- Les méthodes:
- Maximum de vraisemblance récursif (MVR)
 - Erreur de sortie avec modèle de prédiction étendu (ESMPE)
 - Moindres carrés généralisés (MCG)

Principe:

Si la structure « procédé + perturbation » est *correcte*:

- structure perturbation: $(e(t), C(q^{-1})e(t), \text{ou } e(t)/C(q^{-1}))$
- degrés des polynômes: $A(q^{-1}), B(q^{-1}), C(q^{-1}), \text{retard } d$

alors $\mathbf{e}(t)$ $\xrightarrow{\text{asympt.}}$ bruit blanc $(E\{\mathbf{e}(t)\mathbf{e}(t-i)\} = 0; i \neq 0)$

Validation des modèles identifiés avec les méthodes de type I

Méthode de validation:

- 1) Constitution d'un fichier E/S pour le modèle identifié
(avec la même entrée que pour le procédé réel)
- 2) Constitution d'un fichier des erreurs de prédiction
- 3) Test de « blancheur » sur la séquence d'erreur de prédiction résiduelle

Test de « blancheur »

$\{e(t)\}$: séquence centrée des erreurs de prédiction résiduelles
(centré = valeurs mesurées – valeur moyenne)

On calcule:

$$R(0) = \frac{1}{N} \sum_{t=1}^N e^2(t) \quad ; \quad RN(0) = \frac{R(0)}{R(0)} = 1$$

$$R(i) = \frac{1}{N} \sum_{t=1}^N e(t)e(t-i) \quad ; \quad RN(i) = \frac{R(i)}{R(0)} \quad ; \quad i = 1, 2, 3 \dots i_{\max} \quad ; \quad i_{\max} = \max(n_A, n_B + d)$$

Valeurs théoriques: $RN(0) = 1$; $RN(i) = 0$; $i > 0$ (voir aussi Chapitre 4)

Situation réelle:

- Nombre fini de données
- Erreurs résiduelles de structure (ordre, nonlinéarités, bruits)
- Objectif: obtenir des « bons » modèles simples

Critère de validation (N = nombre de données):

$$RN(0) = 1 \quad ; \quad |RN(i)| \leq \frac{2,17}{\sqrt{N}} \quad ; \quad i \geq 1$$

Test de « blancheur »

- Le critère de validation a été défini en tenant compte du test de blancheur pour une séquence bruit blanc gaussien de longueur N et en considérant un intervalle de confiance de 3%
- Quelques valeurs: $N = 128, |RN(i)| \leq 0.192; N = 256, |RN(i)| \leq 0.136$
- Testé extensivement en pratique
- D'autres intervalles de confiance peuvent être considérés (voir tableau 6.2.1 du livre)
- Test pratique: $|RN(i)| \leq 0.15 \quad i = 1, \dots, i_{\max}$

Remarques:

- Critère de validation « trop bon » \longrightarrow simplification possible du modèle
- A « complexité » égale on choisit le modèle qui conduit aux plus petits $|RN(i)|$

Méthodes d'identification basées sur la décorrélation du vecteur des observations et de l'erreur de prédiction (type II)

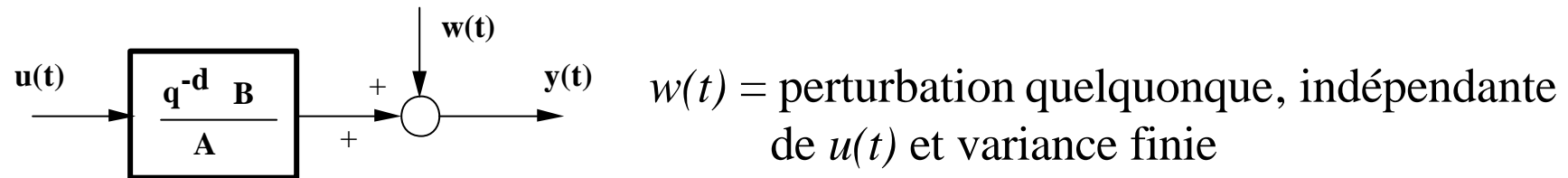
- Variable instrumentale à modèle auxiliaire (VIMA)
- Erreur de sortie à compensateur fixe (ESCF)
- Erreur de sortie avec filtrage des observations (ESFO)
- Erreur de sortie avec filtrage adaptatif des observations (ESFAO)

Variable instrumentale à modèle auxiliaire (VIMA)

Idée: *Créer un nouveau vecteur des observations qui soit corrélé avec les variables non bruitées mais non corrélé avec le bruit*

Procédé + perturbation:

$$A(q^{-1})y(t) = q^{-d} B(q^{-1})u(t) + w'(t+1) \quad \text{où:} \quad w'(t+1) = A(q^{-1})w(t)$$



Prédicteur ajustable (MCR):

$$\hat{y}^0(t+1) = -\hat{a}_1(t)y(t) + \hat{b}_1(t)u(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t)$$

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t), \hat{b}_1(t)] \quad ; \quad \mathbf{f}(t)^T = [-y(t), u(t)]$$

Erreur de prédiction (a priori): $\mathbf{e}^o(t+1) = y(t+1) - \hat{y}^o(t+1)$

Variable instrumentale à modèle auxiliaire (VIMA)

Modèle de prédiction auxiliaire (générateur de la variable instrumentale):

$$y_{IV}(t) = -\hat{a}_1(t)y_{IV}(t-1) + \hat{b}_1(t)u(t-1)$$

La sortie prédite dépend des prédictions précédentes et non pas des mesures précédentes comme dans le prédicteur des MCR

Nouveau vecteur des observations (instrumental):

$$\Phi(t)^T = \mathbf{f}_{IV}(t)^T = [-y_{IV}(t), u(t)]$$

AAP: On utilise l'algorithme donné transparent 4

Cas général :

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t) \dots \hat{a}_{n_A}(t), \hat{b}_1(t) \dots \hat{b}_{n_B}(t)]$$

$$\Phi(t)^T = \mathbf{f}_{IV}(t)^T = [-y_{IV}(t), -y_{IV}(t-1), \dots, u(t-d), u(t-d-1), \dots]$$

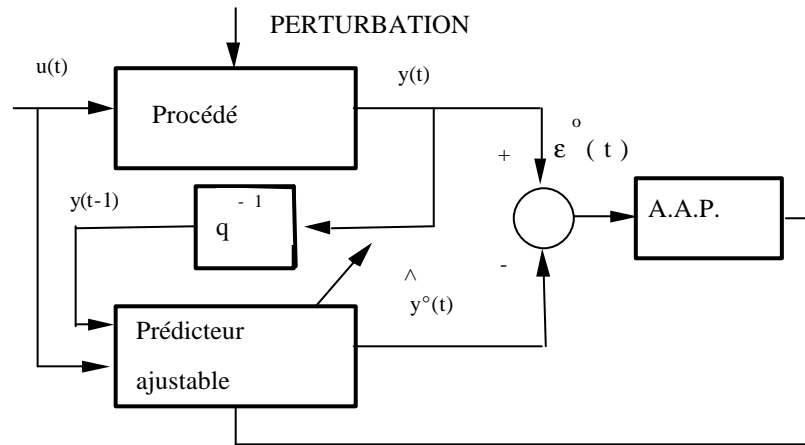
$$y_{IV} \text{ est engendré par: } \hat{A}(t, q^{-1})y_{IV}(t) = q^{-d}\hat{B}(t, q^{-1})u(t)$$

- Identification sans biais des paramètres du modèle du procédé sans identifier le modèle de la perturbation (utile pour bruits non gaussien)
- Nécessite l'initialisation par les MCR.

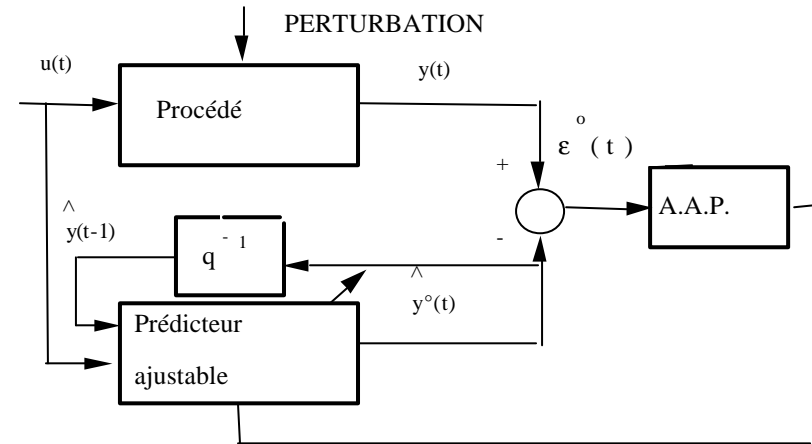
Horizon d'initialisation: (5 à 8) x nb. des paramètres

Erreur de sortie à compensateur fixe (ESCF)

Idée: Remplacer dans le prédictor des MCR, $y(t)$ (qui est bruité) par la prédiction $\hat{y}(t)$ (qui dépendra asymptotiquement que de l'entrée)



Moindres Carrés Récurrents (M.C.R.)



Erreur de Sortie

(E.S.)

Procédé + perturbation:

$$A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t) + w'(t+1) \quad \text{où:} \quad w'(t+1) = A(q^{-1})w(t)$$

$w(t)$ = perturbation quelconque, indépendante de $u(t)$ et variance finie

Erreur de sortie à compensateur fixe (ESCF)

Prédicteur ajustable (Erreur de sortie):

$$\hat{y}^o(t+1) = \hat{a}_1(t)\hat{y}(t) + \hat{b}_1(t)u(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t)$$

↖ au lieu de $y(t)$ (MCR)

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t), \hat{b}_1(t)] ; \quad \Phi(t)^T = \mathbf{f}(t)^T = [-\hat{y}(t), u(t)]$$

$$\hat{y}(t+1) = \hat{\mathbf{q}}(t+1)^T \mathbf{f}(t) \Rightarrow \hat{y}(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t-1)$$

Erreur de prédiction

$$\mathbf{e}^o(t+1) = y(t+1) - \hat{y}^o(t+1) \quad ; \quad \mathbf{e}(t+1) = y(t+1) - \hat{y}(t+1)$$

AAP: On utilise l'algorithme donné transparent 4 avec : $\Phi(t) = \mathbf{f}(t)$

Cas général:

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t) \dots \hat{a}_{n_A}(t), \hat{b}_1(t) \dots \hat{b}_{n_B}(t)]$$

$$\Phi(t) = \mathbf{f}(t)^T = [-\hat{y}(t), -\hat{y}(t-1), \dots, -\hat{y}(t-n_A+1), u(t-d), \dots, u(t-d-n_B+1)]$$

Voir fonction: *oloe.sci(.m)* sur le site web du livre

Erreur de sortie à compensateur fixe (ESCF)

Erreur d'adaptation (optionnel):

$$\mathbf{n}(t+1) = D(q^{-1})\mathbf{e}(t+1) \quad ; \quad \mathbf{n}^o(t+1) = \mathbf{e}^o(t+1) + \sum_{i=1}^{n_D} d_i \mathbf{e}(t+1-i)$$

$$D(q^{-1}) = 1 + d_1 q^{-1} + \dots + d_{n_D} q^{-n_D}$$

AAP: On utilise l'algorithme donné transparent 4 avec : $\mathbf{e}^o(t) = \mathbf{n}^o(t)$

Propriétés:

- *Identification sans biais des paramètres du modèle du procédé sans identifier le modèle de la perturbation (utile pour bruits non gaussiens)*
- *Condition suffisante de convergence: $\left(\frac{D(z^{-1})}{A(z^{-1})} - \frac{\mathbf{I}_2}{2} \right)$; $2 > \mathbf{I}_2 \geq \max \mathbf{I}_2(t)$*

fonction de transfert *strictement réelle positive*

- *Pour $n_A \geq 2$ et gain d'adaptation décroissant le compensateur $D(q^{-1})$ n'est pas nécessaire*
- *Degré du compensateur $D(q^{-1})$: $n_D \leq n_A - 1$ ou n_A*

Erreur de sortie avec filtrage des observations (ESFO)

Prédicteur ajustable (Erreur de sortie):

$$\hat{y}^o(t+1) = \hat{a}_1(t)\hat{y}(t) + \hat{b}_1(t)u(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t)$$

$$\hat{\mathbf{q}}(t)^T = [\hat{a}_1(t), \hat{b}_1(t)] ; \quad \mathbf{f}(t)^T = [-\hat{y}(t), u(t)]$$

$$\hat{y}(t+1) = \hat{\mathbf{q}}(t+1)^T \mathbf{f}(t) \Rightarrow \hat{y}(t) = \hat{\mathbf{q}}(t)^T \mathbf{f}(t-1)$$

Erreur de prédiction:

$$\mathbf{e}^o(t+1) = y(t+1) - \hat{y}^o(t+1) \quad ; \quad \mathbf{e}(t+1) = y(t+1) - \hat{y}(t+1)$$

Filtrage des observations:

Filtre: $L(q^{-1}) = \hat{A}(q^{-1})$ ← Estimation du polynôme $A(q-1)$

$$\Phi(t) = \mathbf{f}_f(t) = \frac{1}{\hat{A}(q^{-1})} \mathbf{f}(t)$$

AAP: On utilise l'algorithme donné transparent 4 avec : $\Phi(t) = \mathbf{f}_f(t)$

- *Condition suffisante de convergence:* $\left(\frac{\hat{A}(z^{-1})}{A(z^{-1})} - \frac{\mathbf{I}_2}{2} \right) ; \quad 2 > \mathbf{I}_2 \geq \max \mathbf{I}_2(t)$

fonction de transfert *strictement réelle positive*

Voir fonction: *foloe.sci(.m)*
sur le site web du livre

Erreur de sortie avec filtrage adaptatif des observations (ESFAO)

Utilise un filtre adaptatif des observations au lieu d'un filtre fixe
(prend avantage de l'amélioration de l'estimation de $A(q^{-1})$ au cours du temps)

Filtrage des observations:

Filtre: $L(t, q^{-1}) = \hat{A}(t, q^{-1})$ ← Estimation du polynôme $A(q^{-1})$ fournie par l'algorithme

$$\Phi(t) = \mathbf{f}_f(t) = \frac{1}{\hat{A}(t, q^{-1})} \mathbf{f}(t)$$

Voir fonction: *afoloe.sci(.m)*
sur le site web du livre

Initialisation:

$$\hat{A}(0, q^{-1}) = \hat{A}_0(q^{-1}) \quad (\text{fourni par un autre algorithme})$$

ou:

$$\hat{A}(0, q^{-1}) = 1 \quad (\text{plus simple et efficace})$$

Cette variante élimine en général les problèmes éventuels liés à la condition SPR de convergence des autres algorithmes (ESCF, ESFO)

Validation des modèles identifiés avec les méthodes de type II

Méthodes d'identification basées sur la décorrélation du vecteur des observations et de l'erreur de prédiction

- Les méthodes:
- Variable instrumentale à modèle auxiliaire (VIMA)
 - Erreur de sortie à compensateur fixe (ESCF)
 - Erreur de sortie avec filtrage des observations (ESFO)
 - Erreur de sortie avec filtrage adaptatif des observations (ESFAO)

Principe:

- Si la perturbation est indépendante de l'entrée
- Si la structure du modèle du procédé est correcte
(degrés des polynômes: $A(q^{-1})$, $B(q^{-1})$, retard d)

$$\text{Alors: } E\{\mathbf{e}(t)\hat{y}(t-i)\} \approx \frac{1}{N} \sum_{t=1}^N \mathbf{e}(t)\hat{y}(t-i) = 0 \quad ; \quad i = 1, 2, 3, \dots$$

avec $\hat{y}(t)$ donné par (prédicteur erreur de sortie):

$$\hat{A}(q^{-1})\hat{y}(t) = q^{-d}\hat{B}(q^{-1})u(t)$$

Validation des modèles identifiés avec les méthodes de type II

Méthode de validation:

- 1) Constitution d'un fichier E/S pour le modèle identifié (avec la même entrée que pour le procédé réel)
- 2) Constitution des fichiers: $\{y(t)\}, \{\hat{y}(t)\}, \{e(t)\}$
- 3) Test de « décorrélation » entre la séquence d'erreur de prédiction résiduelle et la séquence des sorties prédites

Test de « décorrélation »

$\{e(t)\}$: séquence centrée des erreurs de prédiction résiduelles

On calcule:

$$R(i) = \frac{1}{N} \sum_{t=1}^N e(t) \hat{y}(t-i) \quad ; \quad i = 0, 1, 2, \dots, i_{\max} \quad ; \quad i_{\max} = \max(n_A, n_B + d)$$

$$RN(i) = \frac{R(i)}{\left[\left(\frac{1}{N} \sum_{t=1}^N \hat{y}^2(t) \right) \left(\frac{1}{N} \sum_{t=1}^N e^2(t) \right) \right]^{1/2}} \quad ; \quad i = 0, 1, 2, \dots, i_{\max}$$

Remarque: $RN(0) \neq 1$

Valeurs théoriques: $RN(i) = 0; i = 0, 1, 2 \dots i_{\max}$

- Nombre fini de données

Situation réelle:

- Erreurs résiduelles de structure (ordre, non linéarités, bruits)
- Objectif: obtenir des « bons » modèles simples

Critère de validation (N = nombre de données):

$$RN(0) = 1 \quad ; \quad |RN(i)| \leq \frac{2,17}{\sqrt{N}} \quad ; \quad i \geq 1 \quad \text{ou:} \quad |RN(i)| \leq 0.15; \quad i = 1, \dots, i_{\max}$$

Validation comparative des modèles identifiés

- Comparaison entre modèles identifiés par les méthodes de type I :
Test de blancheur
- Comparaison entre modèles identifiés par les méthodes de type II :
Test de décorrélation
- Comparaison entre modèles identifiés par des méthodes de type I et modèles identifiés par des méthodes de type II:
1) *Utilisation du prédicteur d'erreur de sortie pour engendrer $\{\hat{y}(t)\}$*
2) *Test de décorrélation*

Remarque:

Il est souvent utile de comparer les résultats obtenus avec ceux donnés par les MCR

Estimation de la complexité d'un modèle

Exemple 1

Modèle (hypothèse) :

$$y(t) = -a_1 y(t-1) + b_1 u(t-1)$$

$$\text{Ordre: } n = \max(n_A, n_B + d) = 1$$

Construisons la matrice suivante:

$$\begin{bmatrix} y(t) & \vdots & -y(t-1) & u(t-1) \\ y(t-1) & \vdots & -y(t-2) & u(t-2) \\ y(t-2) & \vdots & -y(t-3) & u(t-3) \end{bmatrix} = [Y(t) \quad R(1)] \quad (*)$$

↑

$Y(t)$

↑

$R(1) = R(\hat{n})$

$$\text{rang}[Y(t) \quad R(1)] = \text{rang} \begin{bmatrix} -a_1 y(t-1) + b_1 u(t-1) & \vdots & -y(t-1) & u(t-1) \\ -a_1 y(t-2) + b_1 u(t-2) & \vdots & -y(t-2) & u(t-2) \\ -a_1 y(t-3) + b_1 u(t-3) & \vdots & -y(t-3) & u(t-3) \end{bmatrix} = 2 (< 3)$$

Si l'ordre du modèle est 1 alors (*) n'est pas de rang plein

Estimation de la complexité d'un modèle

Exemple 2

Modèle (hypothèse):
$$y(t) = -\sum_{i=1}^2 a_i y(t-i) + \sum_{i=1}^2 b_i u(t-i) \quad (\text{Ordre} = 2)$$

$$\text{rang} \begin{bmatrix} y(t) & \vdots & -y(t-1) & u(t-1) \\ y(t-1) & \vdots & -y(t-2) & u(t-2) \\ y(t-2) & \vdots & -y(t-3) & u(t-3) \end{bmatrix} = 3 \quad (\text{rang plein}) \quad (*)$$

\uparrow
 $Y(t)$

\uparrow
 $R(1) = R(\hat{n})$

$$\begin{bmatrix} y(t) & \vdots & -y(t-1) & -y(t-2) & u(t-1) & u(t-2) \\ y(t-1) & \vdots & -y(t-2) & -y(t-3) & u(t-2) & u(t-3) \\ y(t-2) & \vdots & -y(t-3) & -y(t-4) & u(t-3) & u(t-4) \\ y(t-3) & \vdots & -y(t-4) & -y(t-5) & u(t-4) & u(t-5) \\ y(t-4) & \vdots & -y(t-5) & -y(t-6) & u(t-5) & u(t-6) \end{bmatrix} = [Y(t) \quad R(2)]$$

\uparrow
 $Y(t)$

\uparrow
 $R(2) = R(\hat{n})$

Rang < 5 (c'est 4)
car $Y(t)$ est une
combinaison linéaire
des colonnes 2, 3, 4, 5

Estimation de la complexité d'un modèle (sans bruit)

Idée: *rechercher le nombre de colonnes à rajouter afin d'exprimer $Y(t)$ comme une combinaison linéaire des autres colonnes*

$$\text{Modèle : } y(t) = -\sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n b_i u(t-i) \quad \boxed{n = ?}$$

$$R(\hat{n}) = \begin{bmatrix} -y(t-1) & \cdots & -y(t-\hat{n}) & u(t-1) & \cdots & u(t-\hat{n}) \\ -y(t-2) & \cdots & -y(t-\hat{n}-1) & u(t-2) & \cdots & u(t-\hat{n}-1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -y(t-N) & \cdots & -y(t-\hat{n}-N+1) & u(t-N) & \cdots & u(t-\hat{n}-N+1) \end{bmatrix} = \begin{bmatrix} \mathbf{f}(t-1)^T \\ \mathbf{f}(t-2)^T \\ \vdots \\ \mathbf{f}(t-N)^T \end{bmatrix}$$

$$\mathbf{f}(t-j)^T = [-y(t-j), \dots, -y(t-\hat{n}-j+1), u(t-j), \dots, u(t-\hat{n}-j+1)]; j=1, \dots, N \quad (**)$$

$$Y(t)^T = [y(t), y(t-1), \dots, y(t-N+1)] \quad (***)$$

On teste le rang de : $[Y(t) \quad \vdots \quad R(\hat{n})]$

- si $\hat{n} < n$, $[Y(t) \quad \vdots \quad R(\hat{n})] = \text{rang plein } (2\hat{n}+1)$
- si $\hat{n} \geq n$, $[Y(t) \quad \vdots \quad R(\hat{n})] = n$ 'est pas de rang plein ($2n$ au lieu de $2\hat{n}+1$)

Estimation de la complexité d'un modèle (avec bruit)

Modèle:
$$y(t) = -\sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n b_i u(t-i) + w(t)$$
 n = ?

Cas: $w(t) =$ bruit blanc

Rem.: *Le test de rang s'apparente à la recherche d'un vecteur qui permet d'exprimer $Y(t)$ comme une combinaison linéaire des colonnes de $R(n)$, c'est à dire on cherche:*

$$J_{MC}(\hat{n}) = \min_{\hat{q}} \frac{1}{N} \left\| Y(t) - R(\hat{n}) \hat{q}_{\hat{n}} \right\|^2 \quad (\diamond)$$

Mais ce critère n'est rien d'autre que le critère des moindres carrés car il peut s'exprimer en tenant compte de (**) et (***):

$$J_{MC}(\hat{n}) = \min_{\hat{q}} \frac{1}{N} \sum_{i=t-N+1}^t \left[y(i) - \hat{q}_{\hat{n}}^T(t) \mathbf{f}(i-1) \right]^2 \quad ; \quad \hat{q}_{\hat{n}}^T = \left[\hat{a}_1, \dots, \hat{a}_{\hat{n}}, \hat{b}_1, \dots, \hat{b}_{\hat{n}} \right]$$

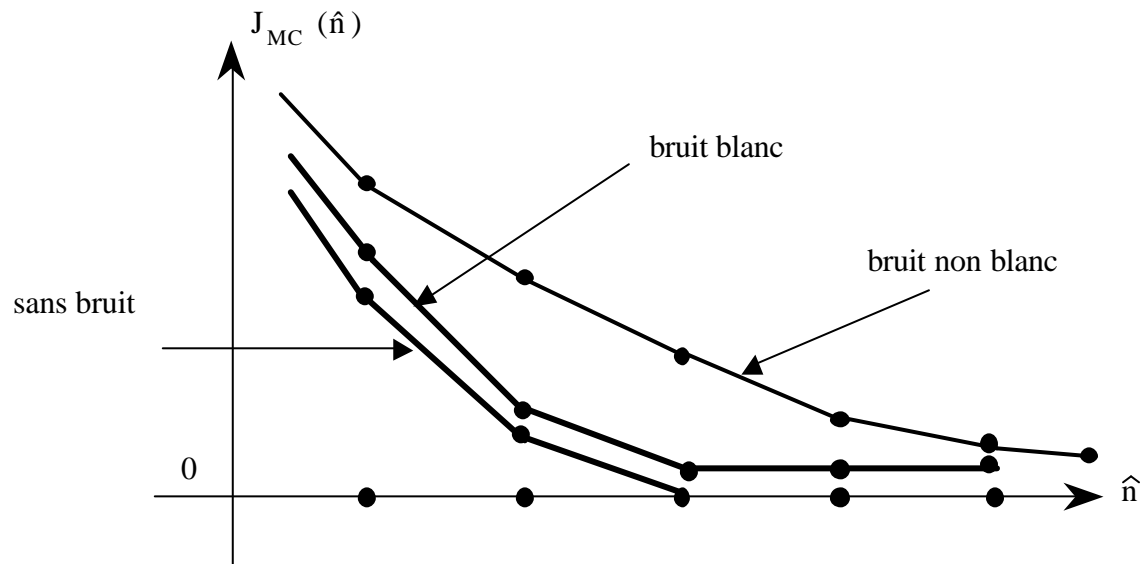
$\hat{n} \geq n$ {

- ▶ Le critère devient nul en absence de bruit
- ▶ Prend une valeur constante si $w(t)$ est un bruit blanc

Estimation de la complexité d'un modèle (avec bruit)

Cas : $w(t) = \text{bruit non blanc}$

La décroissance (\blacklozenge) ne tend pas franchement à zéro. \rightarrow Surestimation de l'ordre.



Il faut donc mettre en oeuvre un algorithme qui donne une estimation non biaisée

Solution: technique des « variables instrumentales » Voir détails pg.329-330 du livre
(remplacement des sorties bruitées par des entrées retardées)

Nouvelle quantité de test:
$$J_{IV}(\hat{n}) = \min_{\hat{\mathbf{q}}} \frac{1}{N} \left\| Y_{IV}(t) - R_{IV}(\hat{n}) \hat{\mathbf{q}} \right\|^2 \quad (\blacklozenge \blacklozenge)$$

Critère d'estimation de complexité

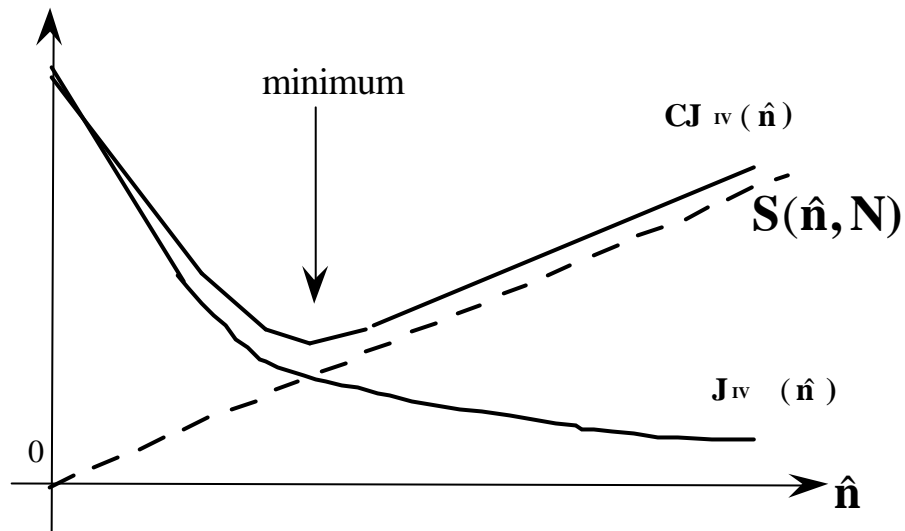
Objectif: *estimer des modèles d'ordre réduit*

Idée: *rajouter dans(♦♦) un terme qui pénalise la complexité*

Critère: $CJ_{IV}(\hat{n}) = J_{IV}(\hat{n}) + S(\hat{n}, N)$ avec: $S(\hat{n}, N) = \frac{2\hat{n} \log(N)}{N}$

L'estimation de l'ordre sera donnée par:

$$\hat{n} = \min_{\hat{n}} CJ_{IV}(\hat{n})$$



Après l'estimation de n on peut appliquer une procédure similaire pour estimer n_A , n_B , d

Voir les fonctions: *estorderls.sci(.m)*, *estorderiv.sci(.m)* sur le site web du livre.

Le logiciel WinPIM(Adaptech) utilise ce critère pour l'estimation de l'ordre.

Quelques remarques récapitulatives

- Algorithmes d'estimation paramétrique *récurifs* ou *non récurifs*
- Deux types de méthodes d'estimation paramétrique basées sur:
 1. Blanchissement de l'erreur de prédiction
 2. Décorrélation : vecteur d'observation/erreur de prédiction
- Deux techniques de validation statistique des modèles:
 1. Test de « blancheur » sur l'erreur de prédiction résiduelle
 2. Test de « décorrélation » entre sortie prédite et erreur de prédiction
- Valeur du seuil de validation pour les tests: $\frac{2.17}{\sqrt{N}}$ (N –nb. de données)
(valeur typique)
- L'ordre des modèles à identifier peut être estimé à partir des données E/S
- Programmes des algorithmes téléchargeables à partir du site web