Apprentissage de modèles dynamiques
Méthode de la VI

Marion Gilson
marion.gilson@univ-lorraine.fr

CRAN, Université de Lorraine – Polytech Nancy – France

Outline

Plan

Solution au biais asymptotique des l'estimateur des MC

On a vu en général (présence de bruit coloré) que l'estimateur des **MC est asymptotiquement biaisé.**

$\Longrightarrow$ Solutions au problème de biais asymptotique de l'estimateur des MC :

Choix d'autres structures de modèles pour le bruit (OE, BJ), 2 possibilités

1. Méthode itérative d'optimisation non linéaire
   - problème de minimum local
   - gourmand en calculs
   - sensible à l'initialisation
2. Méthode de la variable instrumentale
   - issue de l'estimateur des MC
   - s'appuie sur la régression linéaire (peu gourmand en calculs)
   - asymptotiquement non biaisé $\forall$ le bruit !

Solution au biais asymptotique des l'estimateur des MC

On a vu en général (présence de bruit coloré) que l'estimateur des **MC est asymptotiquement biaisé.**

$\Longrightarrow$ Solutions au problème de biais asymptotique de l'estimateur des MC :

Choix d'autres structures de modèles pour le bruit (OE, BJ), 2 possibilités

1. Méthode itérative d'optimisation non linéaire
   - ◇ problème de minimum local
   - ◇ gourmand en calculs
   - ◇ sensible à l'initialisation
2. Méthode de la variable instrumentale
   - ◇ issue de l'estimateur des MC
   - ◇ s'appuie sur la régression linéaire (peu gourmand en calculs)
   - ◇ asymptotiquement non biaisé $\forall$ le bruit !

Solution au biais asymptotique des l'estimateur des MC

On a vu en général (présence de bruit coloré) que l'estimateur des **MC est asymptotiquement biaisé.**

$\Longrightarrow$ Solutions au problème de biais asymptotique de l'estimateur des MC :

Choix d'autres structures de modèles pour le bruit (OE, BJ), 2 possibilités

1. Méthode itérative d'optimisation non linéaire
   - ◇ problème de minimum local
   - ◇ gourmand en calculs
   - ◇ sensible à l'initialisation
2. Méthode de la variable instrumentale
   - ◇ issue de l'estimateur des MC
   - ◇ s'appuie sur la régression linéaire (peu gourmand en calculs)
   - ◇ asymptotiquement non biaisé $\forall$ le bruit !

Solution au biais asymptotique des l'estimateur des MC

On a vu en général (présence de bruit coloré) que l'estimateur des **MC est asymptotiquement biaisé.**

$\Longrightarrow$ Solutions au problème de biais asymptotique de l'estimateur des MC :

Choix d'autres structures de modèles pour le bruit (OE, BJ), 2 possibilités

1. Méthode itérative d'optimisation non linéaire
   - ◇ problème de minimum local
   - ◇ gourmand en calculs
   - ◇ sensible à l'initialisation
2. Méthode de la variable instrumentale $\Longleftarrow$
   - ◇ issue de l'estimateur des MC
   - ◇ s'appuie sur la régression linéaire (peu gourmand en calculs)
   - ◇ asymptotiquement non biaisé $\forall$ le bruit !

Principe de la méthode de variable instrumentale

- Si le vecteur des paramètres "vrais" ($\mathcal{S}$) vérifie

$$y(t_k) = \varphi^T(t_k)\theta_0 + v(t_k)$$

- Estimateur des moindres carrés est donné par

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

- Estimateur convergent sous l'hypothèse de non corrélation entre le bruit $v(t_k)$ et le vecteur de régression, c'est à dire

$$E\left[\varphi(t_k)v(t_k)\right] = 0$$

- **Objectif nº1** : éliminer le biais de l'estimateur des MC lorsque la perturbation $v(t_k)$ n'est pas un bruit blanc
- Soit, un vecteur $z(t_k)$ appelé **instrument** ou **VI** défini de façon à ce que ses composantes soient **non corrélées** avec $v(t_k)$, c'est à dire

$$E\left[z(t_k)v(t_k)\right] = 0$$

Principe de la méthode de variable instrumentale

- Si le vecteur des paramètres "vrais" ($\mathcal{S}$) vérifie

$$y(t_k) = \varphi^T(t_k)\theta_0 + v(t_k)$$

- Estimateur des moindres carrés est donné par

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

- Estimateur convergent sous l'hypothèse de non corrélation entre le bruit $v(t_k)$ et le vecteur de régression, c'est à dire

$$E\left[\varphi(t_k)v(t_k)\right] = 0$$

- **Objectif n°1** : éliminer le biais de l'estimateur des MC lorsque la perturbation $v(t_k)$ n'est pas un bruit blanc
- Soit, un vecteur $z(t_k)$ appelé **instrument** ou **VI** défini de façon à ce que ses composantes soient **non corrélées** avec $v(t_k)$, c'est à dire

$$E\left[z(t_k)v(t_k)\right] = 0$$

Principe de la méthode de variable instrumentale

- Si le vecteur des paramètres "vrais" ($\mathcal{S}$) vérifie

$$y(t_k) = \varphi^T(t_k)\theta_0 + v(t_k)$$

- Estimateur des moindres carrés est donné par

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

- Estimateur convergent sous l'hypothèse de non corrélation entre le bruit $v(t_k)$ et le vecteur de régression, c'est à dire

$$E\left[\varphi(t_k)v(t_k)\right] = 0$$

- **Objectif n°1** : éliminer le biais de l'estimateur des MC lorsque la perturbation $v(t_k)$ n'est pas un bruit blanc
- Soit, un vecteur $z(t_k)$ appelé **instrument** ou **VI** défini de façon à ce que ses composantes soient **non corrélées** avec $v(t_k)$, c'est à dire

$$E\left[z(t_k)v(t_k)\right] = 0$$

Principe de la méthode de variable instrumentale

- Si le vecteur des paramètres "vrais" ($\mathcal{S}$) vérifie

$$y(t_k) = \varphi^T(t_k)\theta_0 + v(t_k)$$

- Estimateur des moindres carrés est donné par

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

- Estimateur convergent sous l'hypothèse de non corrélation entre le bruit $v(t_k)$ et le vecteur de régression, c'est à dire

$$\boxed{E\left[\varphi(t_k)v(t_k)\right] = 0}$$

- **Objectif nº1** : éliminer le biais de l'estimateur des MC lorsque la perturbation $v(t_k)$ n'est pas un bruit blanc
- Soit, un vecteur $z(t_k)$ appelé **instrument** ou **VI** défini de façon à ce que ses composantes soient **non corrélées** avec $v(t_k)$, c'est à dire

$$\boxed{E\left[z(t_k)v(t_k)\right] = 0}$$

Principe de la méthode de variable instrumentale

- Si le vecteur des paramètres "vrais" ($\mathcal{S}$) vérifie

$$y(t_k) = \varphi^T(t_k)\theta_0 + v(t_k)$$

- Estimateur des moindres carrés est donné par

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

- Estimateur convergent sous l'hypothèse de non corrélation entre le bruit $v(t_k)$ et le vecteur de régression, c'est à dire

$$E\left[\varphi(t_k)v(t_k)\right] = 0$$

- **Objectif n°1** : éliminer le biais de l'estimateur des MC lorsque la perturbation $v(t_k)$ n'est pas un bruit blanc
- Soit, un vecteur $z(t_k)$ appelé **instrument** ou **VI** défini de façon à ce que ses composantes soient **non corrélées** avec $v(t_k)$, c'est à dire

$$E\left[z(t_k)v(t_k)\right] = 0$$

Estimateur de la méthode de variable instrumentale

$$\boxed{E\left[z(t_k)v(t_k)\right] = 0}$$

- L'estimateur de VI est solution du système suivant

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)v(t_k) = 0 \;\text{ et }\; v(t_k) = y(t_k) - \varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)\left(y(t_k) - \varphi^T(t_k)\theta\right) = 0$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) - \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) = \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

Estimateur de la méthode de variable instrumentale

$$E\left[z(t_k)v(t_k)\right] = 0$$

- L'estimateur de VI est solution du système suivant

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)v(t_k) = 0 \;\text{ et }\; v(t_k) = y(t_k) - \varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)\left(y(t_k) - \varphi^T(t_k)\theta\right) = 0$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) - \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) = \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

Estimateur de la méthode de variable instrumentale

$$\boxed{E\left[z(t_k)v(t_k)\right] = 0}$$

- L'estimateur de VI est solution du système suivant

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)v(t_k) = 0 \;\text{ et }\; v(t_k) = y(t_k) - \varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)\left(y(t_k) - \varphi^T(t_k)\theta\right) = 0$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) - \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) = \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

Estimateur de la méthode de variable instrumentale

$$\boxed{E\left[z(t_k)v(t_k)\right] = 0}$$

- L'estimateur de VI est solution du système suivant

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)v(t_k) = 0 \;\text{ et }\; v(t_k) = y(t_k) - \varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)\left(y(t_k) - \varphi^T(t_k)\theta\right) = 0$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) - \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) = \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

Estimateur de la méthode de variable instrumentale

$$\boxed{E\left[z(t_k)v(t_k)\right] = 0}$$

- L'estimateur de VI est solution du système suivant

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)v(t_k) = 0 \ \text{ et } \ v(t_k) = y(t_k) - \varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)\left(y(t_k) - \varphi^T(t_k)\theta\right) = 0$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) - \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

$$\frac{1}{N}\sum_{k=1}^{N} z(t_k)y(t_k) = \sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\theta$$

D'où

$$\boxed{\hat{\theta}_{vi} = \left(\sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\right)^{-1}\left(\sum_{k=1}^{N} z(t_k)y(t_k)\right)}$$

Plan

Propriétés statistiques de la méthode de variable instrumentale

$$\hat{\theta}_{vi} = \left(\sum_{k=1}^{N} z(t_k)\varphi^T(t_k) \right)^{-1} \left(\sum_{k=1}^{N} z(t_k)y(t_k) \right)$$

- L'estimateur de VI est convergent ($p\lim_{N\to\infty} \hat{\theta}_{vi} = \theta_0$) ssi :
  1. $E\left[z(t_k)\varphi^T(t_k)\right]$ est non singulière
  2. $E\left[z(t_k)v(t_k)\right] = 0$
- La VI $z(t_k)$ doit donc être
  - ◇ suffisamment corrélée avec le vecteur de régression pour que la matrice (1) soit inversible
  - ◇ non corrélée avec le bruit $v(t_k)$
- L'estimateur de la variable instrumentale est donc asymptotiquement non biaisé en général, $\forall v(t_k)$
- Remarque : si $z(t_k) = \varphi(t_k)$, estimateur de la VI = estimateur des MC

Propriétés statistiques de la méthode de variable instrumentale

$$\hat{\theta}_{vi} = \left(\sum_{k=1}^{N} z(t_k)\varphi^T(t_k) \right)^{-1} \left(\sum_{k=1}^{N} z(t_k)y(t_k) \right)$$

- L'estimateur de VI est convergent ($p\lim_{N\to\infty} \hat{\theta}_{vi} = \theta_0$) ssi :
  1. $E\left[z(t_k)\varphi^T(t_k)\right]$ est non singulière
  2. $E\left[z(t_k)v(t_k)\right] = 0$
- La VI $z(t_k)$ doit donc être
  - ◇ suffisamment corrélée avec le vecteur de régression pour que la matrice (1) soit inversible
  - ◇ non corrélée avec le bruit $v(t_k)$
- L'estimateur de la variable instrumentale est donc asymptotiquement non biaisé en général, $\forall v(t_k)$
- Remarque : si $z(t_k) = \varphi(t_k)$, estimateur de la VI = estimateur des MC

Propriétés statistiques de la méthode de variable instrumentale

$$\hat{\theta}_{vi} = \left(\sum_{k=1}^{N} z(t_k)\varphi^T(t_k) \right)^{-1} \left(\sum_{k=1}^{N} z(t_k)y(t_k) \right)$$

- L'estimateur de VI est convergent ($p\lim_{N\to\infty} \hat{\theta}_{vi} = \theta_0$) ssi :
  1. $E\left[z(t_k)\varphi^T(t_k)\right]$ est non singulière
  2. $E\left[z(t_k)v(t_k)\right] = 0$

- La VI $z(t_k)$ doit donc être
  - ◇ suffisamment corrélée avec le vecteur de régression pour que la matrice (1) soit inversible
  - ◇ non corrélée avec le bruit $v(t_k)$

- L'estimateur de la variable instrumentale est donc asymptotiquement non biaisé en général, $\forall v(t_k)$

- Remarque : si $z(t_k) = \varphi(t_k)$, estimateur de la VI = estimateur des MC

Propriétés statistiques de la méthode de variable instrumentale

$$\hat{\theta}_{vi} = \left(\sum_{k=1}^{N} z(t_k)\varphi^T(t_k)\right)^{-1} \left(\sum_{k=1}^{N} z(t_k)y(t_k)\right)$$

- L'estimateur de VI est convergent ($p\lim_{N\to\infty} \hat{\theta}_{vi} = \theta_0$) ssi :
  1. $E\left[z(t_k)\varphi^T(t_k)\right]$ est non singulière
  2. $E\left[z(t_k)v(t_k)\right] = 0$

- La VI $z(t_k)$ doit donc être
  - ◇ suffisamment corrélée avec le vecteur de régression pour que la matrice (1) soit inversible
  - ◇ non corrélée avec le bruit $v(t_k)$

- L'estimateur de la variable instrumentale est donc asymptotiquement non biaisé en général, $\forall v(t_k)$

- Remarque : si $z(t_k) = \varphi(t_k)$, estimateur de la VI = estimateur des MC

Variable instrumentale sous-optimale

- Objectif indiqué : éliminer le biais asymptotique de l'estimateur des moindres carrés en choisissant une variable instrumentale décorrélée de $v(t_k)$.

  $\longrightarrow$ Pas d'objectif de minimisation de la variance (dans un premier temps)

- Le vecteur de VI $z(t_k)$ est de même dimension que le vecteur de régression $\varphi(t_k)$ et se constitue de 2 parties (entrée/sortie) :
  - $\varphi(t_k) = \begin{bmatrix} -y(t_{k-1}) & \cdots & -y(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$
  - $u(t_k)$ est non corrélée à $v(t_k)$ contrairement à $y(t_k)$
  - La VI (sous-optimale) prend donc la forme :

    $$z(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

    avec $x(t_k)$ une version de $y(t_k)$ décorrélée de $v(t_k)$.

- **VI à modèle auxiliaire**

$$z(t_k) = \begin{bmatrix} -\hat{y}(t_{k-1}) & \cdots & -\hat{y}(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

avec $\hat{y}(t_k) = x(t_k) = G(q^{-1}, \hat{\theta}_{mc})u(t_k)$

Variable instrumentale sous-optimale

- Objectif indiqué : éliminer le biais asymptotique de l'estimateur des moindres carrés en choisissant une variable instrumentale décorrélée de $v(t_k)$.

  $\longrightarrow$ Pas d'objectif de minimisation de la variance (dans un premier temps)

- Le vecteur de VI $z(t_k)$ est de même dimension que le vecteur de régression $\varphi(t_k)$ et se constitue de 2 parties (entrée/sortie) :

  - $\varphi(t_k) = \begin{bmatrix} -y(t_{k-1}) & \cdots & -y(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$
  - $u(t_k)$ est non corrélée à $v(t_k)$ contrairement à $y(t_k)$
  - La VI (sous-optimale) prend donc la forme :

    $$z(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

    avec $x(t_k)$ une version de $y(t_k)$ décorrélée de $v(t_k)$.

- **VI à modèle auxiliaire**

  $$z(t_k) = \begin{bmatrix} -\hat{y}(t_{k-1}) & \cdots & -\hat{y}(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

  avec $\hat{y}(t_k) = x(t_k) = G(q^{-1}, \hat{\theta}_{mc})u(t_k)$

Variable instrumentale sous-optimale

- Objectif indiqué : éliminer le biais asymptotique de l'estimateur des moindres carrés en choisissant une variable instrumentale décorrélée de $v(t_k)$.
  $\longrightarrow$ Pas d'objectif de minimisation de la variance (dans un premier temps)

- Le vecteur de VI $z(t_k)$ est de même dimension que le vecteur de régression $\varphi(t_k)$ et se constitue de 2 parties (entrée/sortie) :
  - $\varphi(t_k) = \begin{bmatrix} -y(t_{k-1}) & \cdots & -y(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$
  - $u(t_k)$ est non corrélée à $v(t_k)$ contrairement à $y(t_k)$
  - La VI (sous-optimale) prend donc la forme :

$$z(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

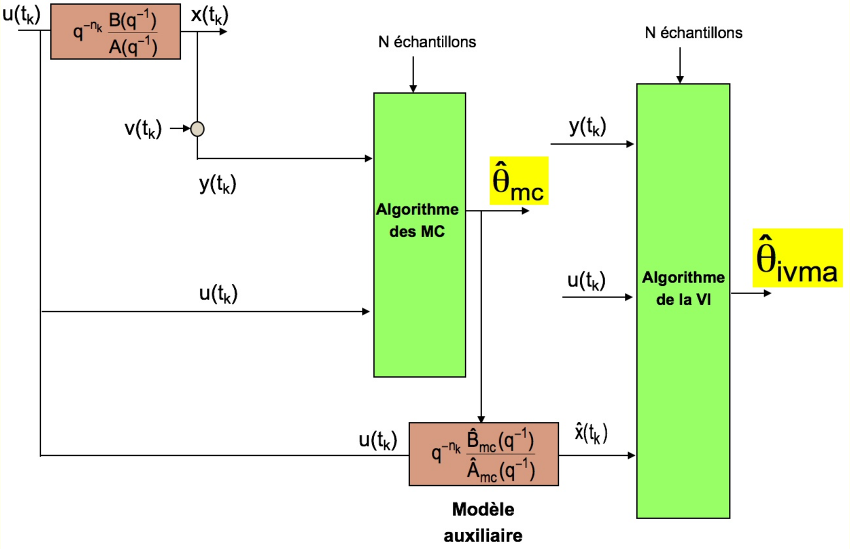
  avec $x(t_k)$ une version de $y(t_k)$ décorrélée de $v(t_k)$.

- **VI à modèle auxiliaire**

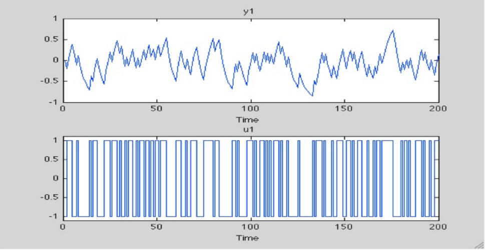
$$z(t_k) = \begin{bmatrix} -\hat{y}(t_{k-1}) & \cdots & -\hat{y}(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

avec $\hat{y}(t_k) = x(t_k) = G(q^{-1}, \hat{\theta}_{mc})u(t_k)$

VI à modèle auxiliaire



VI à modèle auxiliaire




```
B=[0 0.2];           % B(q)=0.2q-1
A=[1 -0.8];          % A(q)=1-0.8q-1
N=200;
S = idpoly(1,B,1,1,A);%y=B/Au+e
u=sign(randn(N,1));

% Simulation du jeu de données
randn('state',sum(100*clock));
e=0.2*randn(N,1);
y=sim(S,[u e]); % simulation de la sortie bruitée
dataS=iddata(y,u); % Objet de données avec y comme sortie et u comme entrée
idplot(dataS)

% 1. Estimation d'un modèle ARX par MC simple
Phi=[-y(1:N-1) u(1:N-1)];     % Matrice de régression
Y=y(2:N);                     % Régresseur
theta_mc=Phi\Y;               % Estimation par MC
theta_mc'
ans   -0.4675   0.1898        % voir aussi la fonction arx de la SID
                              % Marx=arx(dataS,[1 1 1])
```

VI à modèle auxiliaire

```
% 2. Simulation du modèle auxiliaire
Bmc=[0 theta_mc(2)'];
Amc=[1 theta_mc(1)'];
Mmc = idpoly(Amc,Bmc);
xest=sim(Mmc,u); ; % simulation de la sortie du modèle auxiliaire
Z=[-xest(1:N-1) u(1:N-1)]; % Matrice de variable instrumentale
% 3. Estimation par variable instrumentale
theta_ivma=pinv(Z'*Phi)*Z'*Y;   % Estimation par IVMA
theta_ivma'
ans   -0.7565    0.1904
```

Plan

Estimateurs MC et VI

- **MC** : asymptotiquement biaisé en général (faible variance)

- **VIMA** : asymptotiquement non biaisé asymptotiquement
  - ◇ réponse à l'objectif n°1 (enlever le biais des MC)
  - ◇ MAIS à variance NON minimale

  $\Longrightarrow$ Estimateur sous optimal

- **Objectif n°2** : minimiser la variance de l'estimateur de VI à modèle auxiliaire

  Estimateur de la VI optimale `riv`

  - ◇ `riv` : Refined Instrumental Variable
  - ◇ non biaisé et à variance minimale

  $\Longrightarrow$ estimateur à utiliser en premier lieu !

- 2 versions suivant le modèle de bruit
  - ◇ structure BJ : cas général ($H(q^{-1},\theta) = \dfrac{C(q^{-1},\theta)}{D(q^{-1},\theta)}$) $\longrightarrow$ `riv`
  - ◇ structure OE : ($H(q^{-1},\theta) = 1$) $\longrightarrow$ `sriv`

Estimateurs MC et VI

- **MC** : asymptotiquement biaisé en général (faible variance)
- **VIMA** : asymptotiquement non biaisé asymptotiquement
  - ◇ réponse à l'objectif n°1 (enlever le biais des MC)
  - ◇ MAIS à variance NON minimale

$\Longrightarrow$ Estimateur sous optimal

- **Objectif n°2** : minimiser la variance de l'estimateur de VI à modèle auxiliaire

Estimateur de la VI optimale `riv`

  - ◇ `riv` : Refined Instrumental Variable
  - ◇ non biaisé et à variance minimale

$\Longrightarrow$ estimateur à utiliser en premier lieu !

- 2 versions suivant le modèle de bruit
  - ◇ structure BJ : cas général $(H(q^{-1},\theta) = \dfrac{C(q^{-1},\theta)}{D(q^{-1},\theta)}) \longrightarrow$ `riv`
  - ◇ structure OE : $(H(q^{-1},\theta) = 1) \longrightarrow$ `sriv`

Estimateurs MC et VI

- **MC** : asymptotiquement biaisé en général (faible variance)
- **VIMA** : asymptotiquement non biaisé asymptotiquement
  - ◇ réponse à l'objectif n°1 (enlever le biais des MC)
  - ◇ MAIS à variance NON minimale

  $\Longrightarrow$ Estimateur sous optimal
- **Objectif n°2** : minimiser la variance de l'estimateur de VI à modèle auxiliaire

  Estimateur de la VI optimale `riv`

  - ◇ `riv` : Refined Instrumental Variable
  - ◇ non biaisé et à variance minimale

  $\Longrightarrow$ estimateur à utiliser en premier lieu !
- 2 versions suivant le modèle de bruit
  - ◇ structure BJ : cas général ($H(q^{-1},\theta) = \dfrac{C(q^{-1},\theta)}{D(q^{-1},\theta)}$) $\longrightarrow$ `riv`
  - ◇ structure OE : ($H(q^{-1},\theta) = 1$) $\longrightarrow$ `sriv`

Estimateurs MC et VI

- **MC** : asymptotiquement biaisé en général (faible variance)

- **VIMA** : asymptotiquement non biaisé asymptotiquement
  - ◇ réponse à l'objectif n°1 (enlever le biais des MC)
  - ◇ MAIS à variance NON minimale

  $\Longrightarrow$ Estimateur sous optimal

- **Objectif nº2** : minimiser la variance de l'estimateur de VI à modèle auxiliaire

  Estimateur de la VI optimale `riv`

  - ◇ `riv` : Refined Instrumental Variable
  - ◇ non biaisé et à variance minimale

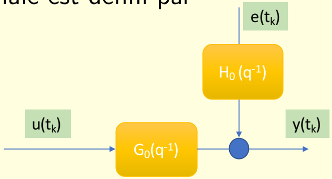
  $\Longrightarrow$ estimateur à utiliser en premier lieu !

- 2 versions suivant le modèle de bruit
  - ◇ structure BJ : cas général $\left(H(q^{-1},\theta) = \dfrac{C(q^{-1},\theta)}{D(q^{-1},\theta)}\right) \longrightarrow$ `riv`
  - ◇ structure OE : $(H(q^{-1},\theta) = 1) \longrightarrow$ `sriv`

Plan

VI optimale – Cas général

- On montre que l'estimateur de VI optimale est défini par



$$\begin{cases} z(t_k) = L(q^{-1})\tilde{\varphi}(t_k) \\ L(q^{-1}) = \dfrac{1}{H_0(q^{-1})A_0(q^{-1})} \end{cases}$$

$$y(t_k) = G_{\mathbf{0}}(q^{-\mathbf{1}})u(t_k) + H_{\mathbf{0}}(q^{-\mathbf{1}})e(t_k)$$

$$\text{et} \quad G_{\mathbf{0}}(q^{-\mathbf{1}}) = q^{-n_k}\frac{B_{\mathbf{0}}(q^{-\mathbf{1}})}{A_{\mathbf{0}}(q^{-\mathbf{1}})}$$

  - ◇ $H_0(q^{-1})$ est le modèle "vrai" du bruit
  - ◇ $\tilde{\varphi}(t_k)$ représente la version non bruitée du vecteur de régression

- **RIV optimale**

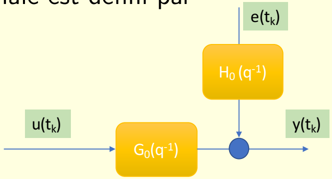
$$\tilde{\varphi}(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

avec $x(t_k) = G_0(q^{-1})u(t_k)$

- Remarques :
  - ◇ IVMA : $L(q^{-1}) = 1$ et $x(t_k) = G(q^{-1}, \hat{\theta}_{mc})u(t_k)$
  - ◇ **Problème** : $G_0(q^{-1})$ et $H_0(q^{-1})$ sont inconnus *a priori* ! !

VI optimale – Cas général

- On montre que l'estimateur de VI optimale est défini par



$$\begin{cases} z(t_k) = L(q^{-1})\tilde{\varphi}(t_k) \\ L(q^{-1}) = \dfrac{1}{H_0(q^{-1})A_0(q^{-1})} \end{cases}$$

$$y(t_k) = G_{\mathbf{0}}(q^{-\mathbf{1}})u(t_k) + H_{\mathbf{0}}(q^{-\mathbf{1}})e(t_k)$$

et $G_{\mathbf{0}}(q^{-\mathbf{1}}) = q^{-n_k}\dfrac{B_{\mathbf{0}}(q^{-\mathbf{1}})}{A_{\mathbf{0}}(q^{-\mathbf{1}})}$

  - ◇ $H_0(q^{-1})$ est le modèle "vrai" du bruit
  - ◇ $\tilde{\varphi}(t_k)$ représente la version non bruitée du vecteur de régression

- **RIV optimale**

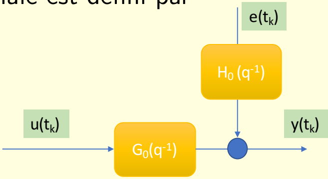
$$\tilde{\varphi}(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

avec $x(t_k) = G_0(q^{-1})u(t_k)$

- Remarques :
  - ◇ IVMA : $L(q^{-1}) = 1$ et $x(t_k) = G(q^{-1}, \hat{\theta}_{mc})u(t_k)$
  - ◇ **Problème** : $G_0(q^{-1})$ et $H_0(q^{-1})$ sont inconnus *a priori* ! !

VI optimale – Cas général

- On montre que l'estimateur de VI optimale est défini par



$$\begin{cases} z(t_k) = L(q^{-1})\tilde{\varphi}(t_k) \\ L(q^{-1}) = \dfrac{1}{H_0(q^{-1})A_0(q^{-1})} \end{cases}$$

$$y(t_k) = G_{\mathbf{0}}(q^{-\mathbf{1}})u(t_k) + H_{\mathbf{0}}(q^{-\mathbf{1}})e(t_k)$$

et $G_{\mathbf{0}}(q^{-\mathbf{1}}) = q^{-n_k}\dfrac{B_{\mathbf{0}}(q^{-\mathbf{1}})}{A_{\mathbf{0}}(q^{-\mathbf{1}})}$

  - ◇ $H_0(q^{-1})$ est le modèle "vrai" du bruit
  - ◇ $\tilde{\varphi}(t_k)$ représente la version non bruitée du vecteur de régression

- **RIV optimale**

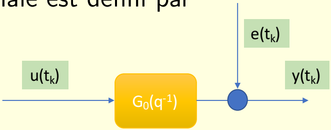
$$\tilde{\varphi}(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

avec $x(t_k) = G_0(q^{-1})u(t_k)$

- Remarques :
  - ◇ IVMA : $L(q^{-1}) = 1$ et $x(t_k) = G(q^{-1}, \hat{\theta}_{mc})u(t_k)$
  - ◇ **Problème** : $G_0(q^{-1})$ et $H_0(q^{-1})$ sont inconnus *a priori* ! !

VI optimale – Cas d'un modèle OE

- On montre que l'estimateur de VI optimale est défini par



$$\begin{cases} z(t_k) = L(q^{-1})\tilde{\varphi}(t_k) \\ L(q^{-1}) = \dfrac{1}{A_0(q^{-1})} \end{cases}$$

$$y(t_k) = G_{\mathbf{0}}(q^{-\mathbf{1}})u(t_k) + e(t_k)$$

$$\text{et} \quad G_{\mathbf{0}}(q^{-\mathbf{1}}) = q^{-n_k}\frac{B_{\mathbf{0}}(q^{-\mathbf{1}})}{A_{\mathbf{0}}(q^{-\mathbf{1}})}$$

  - ◇ $\tilde{\varphi}(t_k)$ représente la version non bruitée du vecteur de régression

- **SRIV optimale**

$$\tilde{\varphi}(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

$$\text{avec} \quad x(t_k) = G_0(q^{-1})u(t_k)$$

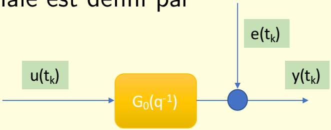
- **Problème** : $G_0(q^{-1})$ est inconnu *a priori* ! !

- **Solution** : utiliser une procédure itérative qui ajuste une estimée initiale des paramètres jusqu'à convergence vers l'estimée optimale

VI optimale – Cas d'un modèle OE

- On montre que l'estimateur de VI optimale est défini par

$$\begin{cases} z(t_k) = L(q^{-1})\tilde{\varphi}(t_k) \\ L(q^{-1}) = \dfrac{1}{A_0(q^{-1})} \end{cases}$$



$$y(t_k) = G_0(q^{-1})u(t_k) + e(t_k)$$

$$\text{et} \quad G_0(q^{-1}) = q^{-n_k}\frac{B_0(q^{-1})}{A_0(q^{-1})}$$

  - ◇ $\tilde{\varphi}(t_k)$ représente la version non bruitée du vecteur de régression

- **SRIV optimale**

$$\tilde{\varphi}(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

avec $x(t_k) = G_0(q^{-1})u(t_k)$

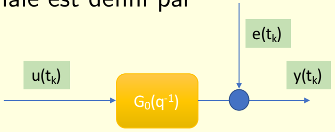
- **Problème** : $G_0(q^{-1})$ est inconnu *a priori* ! !

- **Solution** : utiliser une procédure itérative qui ajuste une estimée initiale des paramètres jusqu'à convergence vers l'estimée optimale

VI optimale – Cas d'un modèle OE

- On montre que l'estimateur de VI optimale est défini par

$$\begin{cases} z(t_k) = L(q^{-1})\tilde{\varphi}(t_k) \\ L(q^{-1}) = \dfrac{1}{A_0(q^{-1})} \end{cases}$$



$$y(t_k) = G_{\mathbf{0}}(q^{-\mathbf{1}})u(t_k) + e(t_k)$$

$$\text{et} \quad G_{\mathbf{0}}(q^{-\mathbf{1}}) = q^{-n_k}\frac{B_{\mathbf{0}}(q^{-\mathbf{1}})}{A_{\mathbf{0}}(q^{-\mathbf{1}})}$$

  - ◇ $\tilde{\varphi}(t_k)$ représente la version non bruitée du vecteur de régression

- **SRIV optimale**

$$\tilde{\varphi}(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

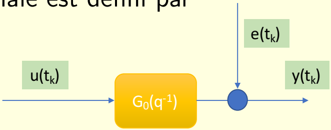
avec $x(t_k) = G_0(q^{-1})u(t_k)$

- **Problème** : $G_0(q^{-1})$ est inconnu *a priori* ! !

- **Solution** : utiliser une procédure itérative qui ajuste une estimée initiale des paramètres jusqu'à convergence vers l'estimée optimale

VI optimale – Cas d'un modèle OE

- On montre que l'estimateur de VI optimale est défini par



$$\begin{cases} z(t_k) = L(q^{-1})\tilde{\varphi}(t_k) \\ L(q^{-1}) = \dfrac{1}{A_0(q^{-1})} \end{cases}$$

$$y(t_k) = G_{\mathbf{0}}(q^{-\mathbf{1}})u(t_k) + e(t_k)$$

$$\text{et} \quad G_{\mathbf{0}}(q^{-\mathbf{1}}) = q^{-n_k}\frac{B_{\mathbf{0}}(q^{-\mathbf{1}})}{A_{\mathbf{0}}(q^{-\mathbf{1}})}$$

  - ◇ $\tilde{\varphi}(t_k)$ représente la version non bruitée du vecteur de régression

- **SRIV optimale**

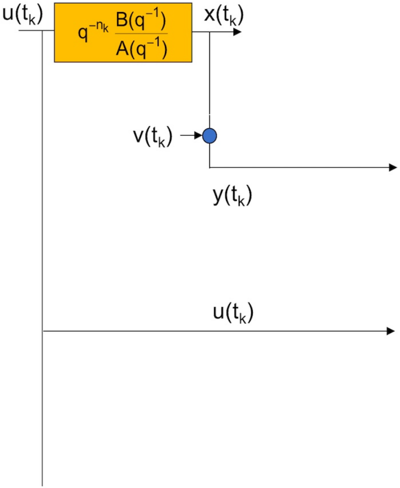
$$\tilde{\varphi}(t_k) = \begin{bmatrix} -x(t_{k-1}) & \cdots & -x(t_{k-n_a}) & u(t_{k-1}) & \cdots & u(t_{k-n_b-n_k-1}) \end{bmatrix}$$

$$\text{avec} \quad x(t_k) = G_0(q^{-1})u(t_k)$$

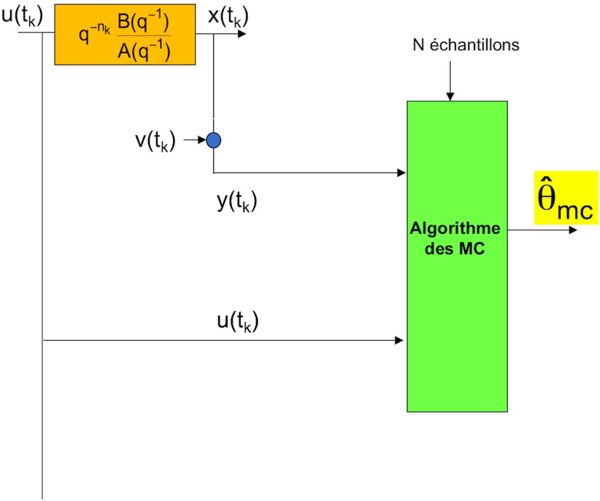
- **Problème** : $G_0(q^{-1})$ est inconnu *a priori* ! !

- **Solution** : utiliser une procédure itérative qui ajuste une estimée initiale des paramètres jusqu'à convergence vers l'estimée optimale

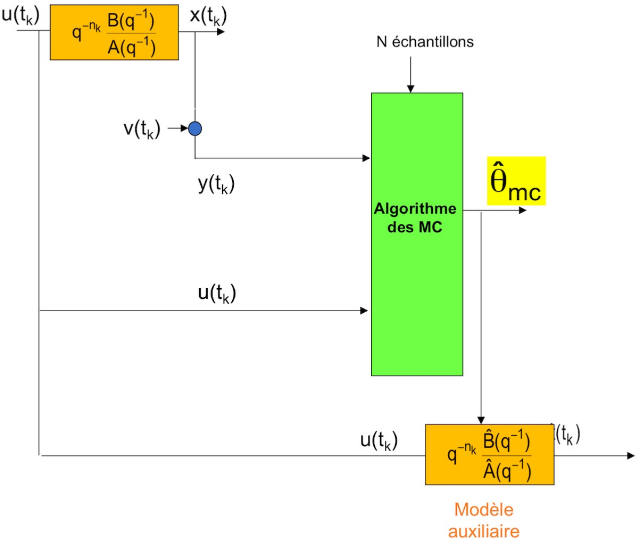
VI optimale sriv



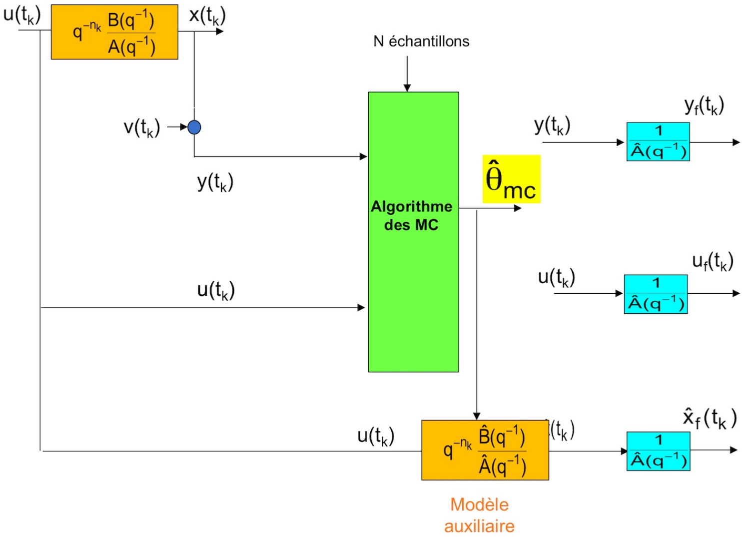
VI optimale sriv



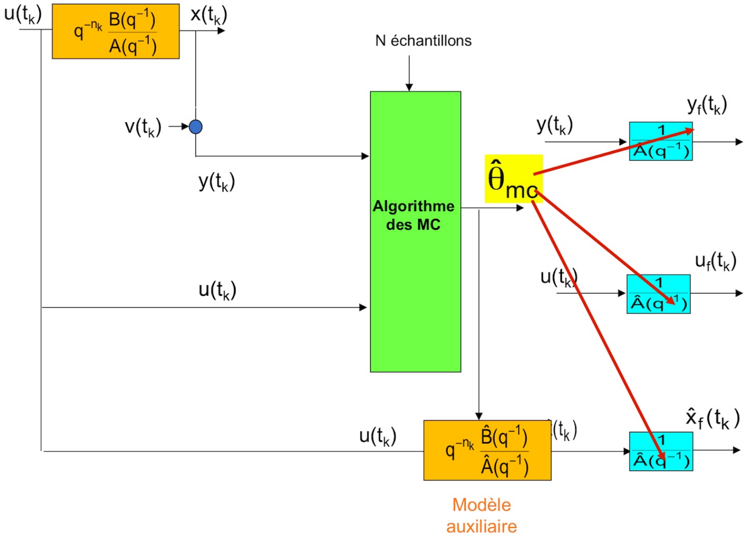
VI optimale sriv



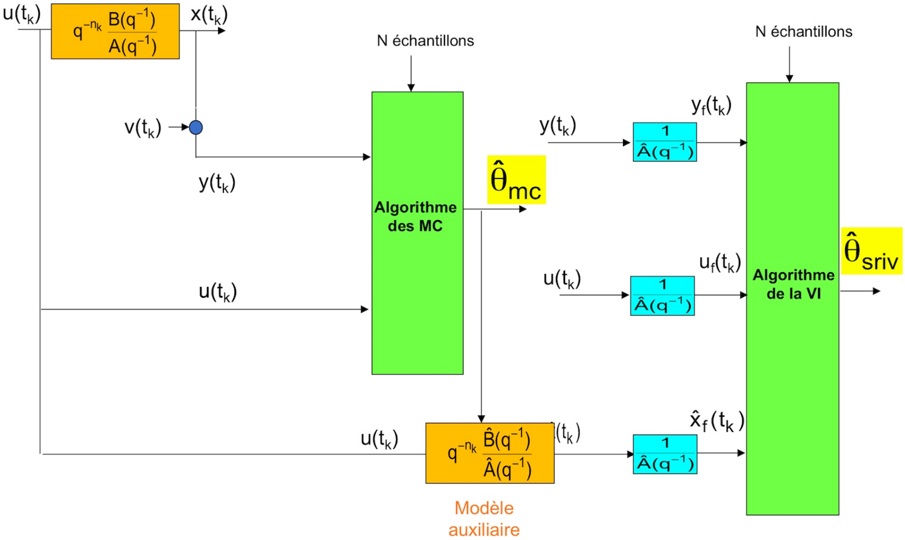
VI optimale sriv



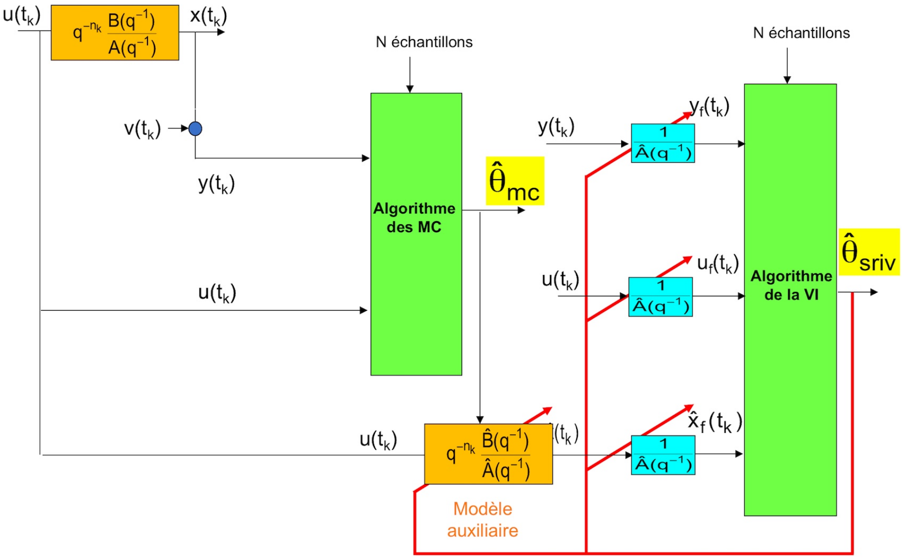
VI optimale sriv



VI optimale sriv



VI optimale sriv



VI optimale sriv et riv

A noter !

- L'algorithme `sriv` est disponible dans la **boîte à outils CONTSID** (`http://www.contsid.cran.univ-lorraine.fr`) librement téléchargeable sur le net

- **Ne pas confondre** avec la fonction `iv4` de la SID (toolbox Matlab) dans laquelle une version approchée de la VI optimale est implantée mais beaucoup moins performante !

- Pour identifier un modèle OE : utiliser l'algorithme `sriv`
- Pour identifier un modèle BJ : utiliser l'algorithme `riv`

- **Suite** : exercice

VI optimale sriv et riv

A noter !

- L'algorithme `sriv` est disponible dans la **boîte à outils CONTSID** (`http://www.contsid.cran.univ-lorraine.fr`) librement téléchargeable sur le net
- **Ne pas confondre** avec la fonction `iv4` de la SID (toolbox Matlab) dans laquelle une version approchée de la VI optimale est implantée mais beaucoup moins performante !
- Pour identifier un modèle OE : utiliser l'algorithme `sriv`
- Pour identifier un modèle BJ : utiliser l'algorithme `riv`
- **Suite** : exercice

VI optimale sriv et riv

A noter !

- L'algorithme `sriv` est disponible dans la **boîte à outils CONTSID** (`http://www.contsid.cran.univ-lorraine.fr`) librement téléchargeable sur le net
- **Ne pas confondre** avec la fonction `iv4` de la SID (toolbox Matlab) dans laquelle une version approchée de la VI optimale est implantée mais beaucoup moins performante !
- Pour identifier un modèle OE : utiliser l'algorithme `sriv`
- Pour identifier un modèle BJ : utiliser l'algorithme `riv`
- **Suite** : exercice

VI optimale sriv et riv

A noter !

- L'algorithme `sriv` est disponible dans la **boîte à outils CONTSID** (`http://www.contsid.cran.univ-lorraine.fr`) librement téléchargeable sur le net
- **Ne pas confondre** avec la fonction `iv4` de la SID (toolbox Matlab) dans laquelle une version approchée de la VI optimale est implantée mais beaucoup moins performante !
- Pour identifier un modèle OE : utiliser l'algorithme `sriv`
- Pour identifier un modèle BJ : utiliser l'algorithme `riv`
- **Suite** : exercice

VI optimale sriv et riv

A noter !

- L'algorithme `sriv` est disponible dans la **boîte à outils CONTSID** (`http://www.contsid.cran.univ-lorraine.fr`) librement téléchargeable sur le net

- **Ne pas confondre** avec la fonction `iv4` de la SID (toolbox Matlab) dans laquelle une version approchée de la VI optimale est implantée mais beaucoup moins performante !

- Pour identifier un modèle OE : utiliser l'algorithme `sriv`
- Pour identifier un modèle BJ : utiliser l'algorithme `riv`

- **Suite** : exercice

Exo complet

```
B=[0 0.2]; A=[1 -0.8];   % B(q)=0.2q-1 et A(q)=1-0.8q-1
N=500;                   % Nombre d'échantillons
S = idpoly(1,B,1,1,A);   %y=B/Au+e : le modèle vrai est un modèle OE
u=sign(randn(N,1));      % signal d'entrée
 Matrice_param_arx=[]; Matrice_param_ivma=[]; Matrice_param_sriv=[];
for i=1:200  % 200 réalisations de bruit
           randn('state',sum(100*clock));
           e=0.2*randn(N,1);
           y=sim(S,[u e]);               % simulation de la sortie bruitée
           dataS= iddata(y,u);           % Objet de données ayant y comme sortie et u comme en
    % 1. Estimation d'un modèle ARX par MC simple
           Marx=arx(dataS,[1 1 1]);      % Estimation par MC du modèle ARX
           Matrice_param_arx=[Matrice_param_arx;  [Marx.A(2:end)  Marx.B(2:end)]];
    % 2. Simulation du modèle auxiliaire
           xest=sim(Marx,u);             % simulation de la sortie du modèle auxiliaire
           Phi=[-y(1:N-1) u(1:N-1)];     % Matrice de régression
           Y=y(2:N);
           Z=[-xest(1:N-1) u(1:N-1)];    % Matrice de variable instrumentale
   % 3. Estimation par variable instrumentale à MA
           theta_ivma=pinv(Z'*Phi)*Z'*Y;     % Estimation par IVMA du modèle ARX
           Matrice_param_ivma=[Matrice_param_ivma;  theta_ivma'];
   %    Estimation par SRIV
           Msriv=sriv(dataS,[1 1 1]);    % Estimation par SRIV du modèle OE
           Matrice_param_sriv=[Matrice_param_sriv;  [Msriv.F(2:end)  Msriv.B(2:end)]];
end
```

Exo complet

% Vecteur de paramètres « vrais »
% [a1 b1] = [-0.8 0.2]

% Valeur moyenne des paramètres estimés
mean(Matrice_param_arx)
-0.5939 0.2016 % Paramètre $\hat{a}_1$ biaisé – ***Estimateur des MC : biaisé***

mean(Matrice_param_ivma)
-0.7975 0.2004 % ***Estimateur IVMA : non biaisé***

mean(Matrice_param_sriv)
-0.8005 0.1993 % ***Estimateur SRIV : non biaisé***

% Ecart-type des paramètres estimés
std(Matrice_param_arx)
0.0200 0.0100

std(Matrice_param_ivma)
0.0227 0.0108 % ***Estimateur IVMA à variance non minimale***

std(Matrice_param_sriv)
0.0097 0.0075 % ***Estimateur SRIV à variance minimale***

Bilan

- **Méthode des Moindres Carrés**
  - simple, solution analytique, peu gourmand en calculs
  - asymptotiquement biaisée en général

- **Solutions au problème de biais de l'estimateur des MC**

  Choix d'autres structures de modèles pour le bruit

  - Méthode itérative d'optimisation non linéaire
    - asymptotiquement non biaisée en général
    - mais pas très robuste en pratique : très sensible à l'initialisation (problème des minima locaux), gourmands en calculs
    - OE & PEM dans la boîte à outils SID de Matlab
  - Méthode itérative de la variable instrumentale (VI)
    - asymptotiquement non biaisée en général
    - peu gourmand en calculs (s'appuie sur la régression linéaire)
    - robuste en pratique (vis à vis du bruit) et peu sensible à l'initialisation
    - SRIV & RIV dans la boîte à outils CONTSID

Bilan

- **Méthode des Moindres Carrés**
  - simple, solution analytique, peu gourmand en calculs
  - asymptotiquement biaisée en général

- **Solutions au problème de biais de l'estimateur des MC**

  Choix d'autres structures de modèles pour le bruit

  - Méthode itérative d'optimisation non linéaire
    - asymptotiquement non biaisée en général
    - mais pas très robuste en pratique : très sensible à l'initialisation (problème des minima locaux), gourmands en calculs
    - OE & PEM dans la boîte à outils SID de Matlab

  - Méthode itérative de la variable instrumentale (VI)
    - asymptotiquement non biaisée en général
    - peu gourmand en calculs (s'appuie sur la régression linéaire)
    - robuste en pratique (vis à vis du bruit) et peu sensible à l'initialisation
    - SRIV & RIV dans la boîte à outils CONTSID

Bilan

- **Méthode des Moindres Carrés**
  - simple, solution analytique, peu gourmand en calculs
  - asymptotiquement biaisée en général

- **Solutions au problème de biais de l'estimateur des MC**

  Choix d'autres structures de modèles pour le bruit

  - Méthode itérative d'optimisation non linéaire
    - asymptotiquement non biaisée en général
    - mais pas très robuste en pratique : très sensible à l'initialisation (problème des minima locaux), gourmands en calculs
    - OE & PEM dans la boîte à outils SID de Matlab

  - Méthode itérative de la variable instrumentale (VI)
    - asymptotiquement non biaisée en général
    - peu gourmand en calculs (s'appuie sur la régression linéaire)
    - robuste en pratique (vis à vis du bruit) et peu sensible à l'initialisation
    - SRIV & RIV dans la boîte à outils CONTSID

Bilan

- **Méthode des Moindres Carrés**
  - ◇ simple, solution analytique, peu gourmand en calculs
  - ◇ asymptotiquement biaisée en général

- **Solutions au problème de biais de l'estimateur des MC**

  Choix d'autres structures de modèles pour le bruit

  - ◇ Méthode itérative d'optimisation non linéaire
    - asymptotiquement non biaisée en général
    - mais pas très robuste en pratique : très sensible à l'initialisation (problème des minima locaux), gourmands en calculs
    - OE & PEM dans la boîte à outils SID de Matlab
  - ◇ Méthode itérative de la variable instrumentale (VI)
    - asymptotiquement non biaisée en général
    - peu gourmand en calculs (s'appuie sur la régression linéaire)
    - robuste en pratique (vis à vis du bruit) et peu sensible à l'initialisation
    - SRIV & RIV dans la boîte à outils CONTSID