PUC-Rio - Certificação Digital Nº 0115595/CA

4
Métodos de identificação de subespaço

4.1
Introdução

Os métodos ou algoritmos de identificação de subespaço são métodos no domínio do tempo aplicáveis na identificação de parâmetros modais de sistemas mecânicos lineares invariantes no tempo com múltiplas entradas e múltiplas saídas (MIMO), usando como base o modelo em espaço de estados tipicamente usado em Dinâmica e Vibrações. Em geral, os métodos que serão estudados utilizam os dados de excitação e resposta da estrutura em estudo e precisam do cálculo prévio dos parâmetros de Markov (definidos no capítulo 2). Os algoritmos de subespaço que serão descritos neste capítulo são: o Eigensystem Realization Algorithm (ERA), sem tratamento especial dos dados; o ERA/DC empregando funções de correlação das matrizes de Hankel; o OKID/ERA no qual um observador de estado é introduzido para comprimir os dados e melhorar os resultados da identificação e o OKID/ERA/DC que além de observador usa funções de correlação. Estes algoritmos destacam-se pela utilização da decomposição em valores singulares (SVD) dos dados, obtendo subespaços ortogonais dos modos associados ao sistema e dos modos associados ao ruído.

4.2
Cálculo dos parâmetros de Markov

Supondo que os dados obtidos num experimento para identificação de parâmetros modais de um sistema mecânico têm uma realização da forma

$$\boldsymbol{x}(k+1) = \mathbf{A}\boldsymbol{x}(k) + \mathbf{B}\boldsymbol{u}(k) \tag{4-1}$$

$$\boldsymbol{y}(k) = \mathbf{C}\boldsymbol{x}(k) + \mathbf{D}\boldsymbol{u}(k) \tag{4-2}$$

e considerando que o sistema tem ordem $n = 2q$ (q graus de liberdade e com as dimensões das matrizes como na eq.(2-16) e eq.(2-17)); e supondo condições iniciais iguais a zero, $\boldsymbol{x}(0) = 0$, a seqüência obtida das equações (4-1 e 4-2) avaliadas para $k = 0,\ 1, \ldots,\ l-1$, sendo l o número de amostras, pode ser escrito como

$$\begin{aligned}
\boldsymbol{x}(0) &= 0, \\
\boldsymbol{y}(0) &= \mathbf{D}\boldsymbol{u}(0), \\
\boldsymbol{x}(1) &= \mathbf{B}\boldsymbol{u}(0), \\
\boldsymbol{y}(1) &= \mathbf{CB}\boldsymbol{u}(0) + \mathbf{D}\boldsymbol{u}(1), \\
\boldsymbol{x}(2) &= \mathbf{AB}\boldsymbol{u}(0) + \mathbf{B}\boldsymbol{u}(1), \\
\boldsymbol{y}(2) &= \mathbf{CAB}\boldsymbol{u}(0) + \mathbf{CB}\boldsymbol{u}(1) + \mathbf{D}\boldsymbol{u}(2), \\
&\vdots \\
\boldsymbol{x}(l-1) &= \sum_{i=1}^{l-1} \mathbf{A}^{i-1}\mathbf{B}\boldsymbol{u}(l-1-i), \\
\boldsymbol{y}(l-1) &= \sum_{i=1}^{l-1} \mathbf{CA}^{i-1}\mathbf{B}\boldsymbol{u}(l-1-i) + \mathbf{D}\boldsymbol{u}(l-1)
\end{aligned}$$

A relação entre os históricos no tempo discreto da entrada e a saída pode ser representada na seguinte equação matricial

$$[\boldsymbol{y}]_{m\times l} = [\mathbf{Y}]_{m\times rl}\,[\mathbf{U}]_{rl\times l} \tag{4-3}$$

sendo

$$\begin{aligned}
\boldsymbol{y} &= \begin{bmatrix} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(l-1) \end{bmatrix} \\
\mathbf{Y} &= \begin{bmatrix} \mathbf{D} & \mathbf{CB} & \mathbf{CAB} & \cdots & \mathbf{CA}^{l-2}\mathbf{B} \end{bmatrix}
\end{aligned}$$

$$\mathbf{U} = \begin{bmatrix} \boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{u}(0) & \boldsymbol{u}(1) & \cdots & \boldsymbol{u}(l-2) \\ & & \boldsymbol{u}(0) & \cdots & \boldsymbol{u}(l-3) \\ & & & \ddots & \vdots \\ & & & & \boldsymbol{u}(0) \end{bmatrix}$$

Na equação (4-3), observa-se que há $(m \times rl)$ incógnitas na matriz $\mathbf{Y}$, a qual contêm a seqüência de parâmetros de Markov, mas só há $(m \times l)$ equações. Se $r > 1$, a solução para $\mathbf{Y}$ não é única, mas para um sistema linear invariante no tempo de dimensão finita, $\mathbf{Y}$ tem que ser única.

No caso em que $\mathbf{A}$ é assintoticamente estável, tal que para um valor

PUC-Rio - Certificação Digital Nº 0115595/CA

grande p, $\mathbf{A}^k \approx 0$, $\forall\, k \geqslant p$; a eq.(4-3) pode ser aproximada por

$$[\boldsymbol{y}]_{m\times l} = [\mathbf{Y}]_{m\times r(p+1)}\,[\mathbf{U}]_{r(p+1)\times l} \tag{4-4}$$

sendo

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(l-1) \end{bmatrix}$$
$$\mathbf{Y} = \begin{bmatrix} \mathbf{D} & \mathbf{CB} & \mathbf{CAB} & \cdots & \mathbf{CA}^{p-1}\mathbf{B} \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} \boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(p) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{u}(0) & \boldsymbol{u}(1) & \cdots & \boldsymbol{u}(p-1) & \cdots & \boldsymbol{u}(l-2) \\ & & \boldsymbol{u}(0) & \cdots & \boldsymbol{u}(p-2) & \cdots & \boldsymbol{u}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \boldsymbol{u}(0) & \cdots & \boldsymbol{u}(l-p-1) \end{bmatrix}$$

Elegendo o número de amostras ou quantidade de dados medidos l, tal que $l \geqslant r(p+1)$, a solução para $\mathbf{Y}$ será única e os primeiros p parâmetros de Markov aproximadamente satisfazem a relação

$$\mathbf{Y} = \boldsymbol{y}\mathbf{U}^{\dagger} \tag{4-5}$$

sendo $\mathbf{U}^{\dagger}$ a pseudo inversa de $\mathbf{U}$ e o erro de aproximação decresce quando aumenta p.

4.3 Eigensystem Realization Algorithm (ERA)

O Eigensystem Realization Algorithm (ERA) utiliza os dados de exitação e resposta de uma estrutura obtidos num experimento de vibrações para identificar a quádrupla de matrizes ($\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$) no tempo discreto.

Dados $\{\mathbf{Y}_0, \mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_k\}$, os parâmetros de Markov do sistema no tempo discreto, definidos na eq.(2-27), a realização do sistema começa com a construção da matriz de Hankel generalizada, ($\alpha m \times \beta r$), que tem posto n, sempre que $\alpha \geqslant n$ e $\beta \geqslant n$:

$$\mathbf{H}(k-1) = \begin{bmatrix} \mathbf{Y}_k & \mathbf{Y}_{k+1} & \cdots & \mathbf{Y}_{k+\beta-1} \\ \mathbf{Y}_{k+1} & \mathbf{Y}_{k+2} & \cdots & \mathbf{Y}_{k+\beta} \\ \vdots & \vdots & & \vdots \\ \mathbf{Y}_{k+\alpha-1} & \mathbf{Y}_{k+\alpha} & \cdots & \mathbf{Y}_{k+\alpha+\beta-2} \end{bmatrix} \tag{4-6}$$

PUC-Rio - Certificação Digital Nº 0115595/CA

Usando a definição dos parâmetros de Markov, a matriz de Hankel pode ser escrita como

$$\mathbf{H}(k-1) = \mathcal{P}_\alpha \mathbf{A}^{k-1} \mathcal{Q}_\beta \tag{4-7}$$

sendo, $\mathcal{P}_\alpha$ a matriz de observabilidade e $\mathcal{Q}_\beta$ a matriz de controlabilidade definidas como

$$\mathcal{P}_\alpha = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{\alpha-1} \end{bmatrix} \tag{4-8}$$

$$\mathcal{Q}_\beta = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \cdots & \mathbf{A}^{\beta-1}\mathbf{B} \end{bmatrix} \tag{4-9}$$

Fazendo k=1 na eq.(4-7) obtém-se a matriz $\mathbf{H}(0)$, a qual tem uma decomposição em valores singulares da forma

$$\mathbf{H}(0) = \mathbf{R}\Sigma\mathbf{S}^T \tag{4-10}$$

sendo, $\mathbf{R}(\alpha m \times \alpha m)$ e $\mathbf{S}(\beta r \times \beta r)$, matrizes ortonormais ($\mathbf{R}^T\mathbf{R} = \mathbf{I}_{\alpha m}$, $\mathbf{S}^T\mathbf{S} = \mathbf{I}_{\beta r}$). As colunas de $\mathbf{R}$ e $\mathbf{S}$ são os vetores singulares á esquerda e direita de $\mathbf{H}(0)$, respectivamente. A matriz Σ $(\alpha m \times \beta r)$ contém os valores singulares do sistema, σ_i $(i = 1, 2, \ldots, n)$. Teoricamente, ou seja, sem ruído nos parâmetros de Markov, para um sistema de ordem n, a matriz $\mathbf{H}(0)$ tem posto n e deve-se obter n valores singulares de $\mathbf{H}(0)$ diferentes de zero, tal que as matrizes Σ, $\mathbf{R}$ e $\mathbf{S}$ podem ser subdivididas da seguinte maneira

$$\Sigma = \begin{bmatrix} \Sigma_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \mathbf{R} = \begin{bmatrix} \mathbf{R}_n & \bar{\mathbf{R}} \end{bmatrix}, \ \mathbf{S} = \begin{bmatrix} \mathbf{S}_n & \bar{\mathbf{S}} \end{bmatrix} \tag{4-11}$$

sendo $\Sigma_n = diag(\sigma_1, \sigma_2, ..., \sigma_n)$ com $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$; $\mathbf{R}_n$ e $\mathbf{S}_n$ são as matrizes formadas por as n colunas de $\mathbf{R}$ e $\mathbf{S}$ respectivamente associadas aos n valores singulares σ_i. Portanto, a matriz $\mathbf{H}(0)$ pode ser escrita como

$$\mathbf{H}(0) = \mathbf{R}_n \Sigma_n \mathbf{S}_n^T \tag{4-12}$$

Usando a eq.(4-7) quando k=1, tem-se

$$\mathbf{H}(0) = \mathcal{P}_\alpha \mathcal{Q}_\beta \tag{4-13}$$

Comparando a eq.(4-12) e eq.(4-13), parece óbvio deduzir que $\mathcal{P}_\alpha$ está

PUC-Rio - Certificação Digital Nº 0115595/CA

relacionado com $\mathbf{R}_n$ e $\mathcal{Q}_\beta$ está relacionado com $\mathbf{S}_n^T$. Então, uma possível escolha "balanceada"é fazer

$$\mathcal{P}_\alpha = \mathbf{R}_n \Sigma_n^{1/2} \quad \text{e} \quad \mathcal{Q}_\beta = \Sigma_n^{1/2} \mathbf{S}_n^T. \tag{4-14}$$

É a partir dessa equivalência que podemos estimar as matrizes $\mathbf{B}$ e $\mathbf{C}$; a matriz de distribuição de entradas estimada $\hat{\mathbf{B}}$ são as r primeiras colunas do $\mathcal{Q}_\beta$ e a matriz de distribução de saídas estimada $\hat{\mathbf{C}}$ são as m primeiras linhas do $\mathcal{P}_\alpha$.

Fazendo k=2 na eq.(4-7) e substituindo $\mathcal{P}_\alpha$ e $\mathcal{Q}_\beta$ por suas equivalências, obtém-se que

$$\mathbf{H}(1) = \mathcal{P}_\alpha \mathbf{A} \mathcal{Q}_\beta = \mathbf{R}_n \Sigma_n^{1/2} \mathbf{A} \Sigma_n^{1/2} \mathbf{S}_n^T \tag{4-15}$$

então, uma realização mínima para o sistema pode ser estimada como

$$\begin{aligned} \hat{\mathbf{A}} &= \Sigma_n^{-1/2} \mathbf{R}_n^T \mathbf{H}(1) \mathbf{S}_n \Sigma_n^{-1/2} \\ \hat{\mathbf{B}} &= \Sigma_n^{1/2} \mathbf{S}_n^T \mathbf{E}_r \\ \hat{\mathbf{C}} &= \mathbf{E}_m^T \mathbf{R}_n \Sigma_n^{1/2} \\ \hat{\mathbf{D}} &= \mathbf{Y}_0 \end{aligned} \tag{4-16}$$

sendo, $\mathbf{E}_m^T = [\mathbf{I}_m\ \mathbf{O}_m \cdots \mathbf{O}_m]$ e $\mathbf{E}_r^T = [\mathbf{I}_r\ \mathbf{O}_r \cdots \mathbf{O}_r]$ ($\mathbf{O}_i$ é a matriz nula $i \times i$).

A realização ($\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$) pode ser transformado nas coordenadas modais obtendo a realização ($\hat{\Lambda}$, $\hat{\Psi}^{-1}\hat{\mathbf{B}}$, $\hat{\mathbf{C}}\hat{\Psi}$, $\hat{\mathbf{D}}$). A matriz diagonal dos autovalores de $\hat{\mathbf{A}}$, $\hat{\Lambda}$, contém a informação dos amortecimentos modais e das freqüências naturais amortecidas. A matriz $\hat{\Psi}^{-1}\hat{\mathbf{B}}$ define as amplitudes modais iniciais e a matriz $\hat{\mathbf{C}}\hat{\Psi}$ os modos nos pontos de medição, sendo $\hat{\Psi}$ a matriz de autovetores de $\hat{\mathbf{A}}$. Todos os parâmetros modais de um sistema dinâmico podem ser identificados pela realização; de fato, os amortecimentos modais e freqüências naturais amortecidas são simplesmente a parte real e imaginária dos autovalores no $\hat{\Lambda}_c$, após a transformação do tempo discreto para o domínio no tempo contínuo.

4.4 ERA com correlação de dados (ERA/DC)

O ERA/DC (ERA/Data Correlations) também identifica uma realização mínima representada pela quádrupla ($\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$) no tempo discreto.

PUC-Rio - Certificação Digital Nº 0115595/CA

Diferente do ERA, o ERA/DC precisa construir a matriz de $(\varepsilon \times \varsigma)$ blocos, $\mathcal{H}(k)$, conhecida como a matriz Hankel de correlação em blocos definida como

$$\mathcal{H}(k) = \begin{bmatrix} \mathcal{R}_{hh}(k) & \mathcal{R}_{hh}(k+\eta) & \cdots & \mathcal{R}_{hh}(k+\varsigma\eta) \\ \mathcal{R}_{hh}(k+\eta) & \mathcal{R}_{hh}(k+2\eta) & \cdots & \mathcal{R}_{hh}(k+(\varsigma+1)\eta) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{R}_{hh}(k+\varepsilon\eta) & \mathcal{R}_{hh}(k+(\varepsilon+1)\eta) & \cdots & \mathcal{R}_{hh}(k+(\varepsilon+\varsigma)\eta) \end{bmatrix} \tag{4-17}$$

sendo, $\mathcal{R}_{hh}(k)$ a matriz de correlação de dados de ordem (αm) definida como

$$\mathcal{R}_{hh}(k) = \mathbf{H}(k)\mathbf{H}^T(0) \tag{4-18}$$

Se os ruídos nos parâmetros de Markov não forem correlacionados, a matriz de correlação $\mathcal{R}_{hh}(k)$ conterá menos ruído que a matriz de Hankel $\mathbf{H}(k)$. Usando a eq.(4-13) na eq.(4-18) obtém-se

$$\mathcal{R}_{hh}(k) = (\mathcal{P}_\alpha \mathbf{A}^k \mathcal{Q}_\beta)(\mathcal{P}_\alpha \mathcal{Q}_\beta)^T$$

fazendo $\mathcal{Q}_c = \mathcal{Q}_\beta \mathcal{Q}_\beta^T \mathcal{P}_\alpha^T$, a eq.(4-18) é reduzida para

$$\mathcal{R}_{hh}(k) = \mathcal{P}_\alpha \mathbf{A}^k \mathcal{Q}_c \tag{4-19}$$

então, usando a eq.(4-19) na eq.(4-17), a matriz $\mathcal{H}(k)$ pode-se escrever da forma

$$\mathcal{H}(k) = \begin{bmatrix} \mathcal{P}_\alpha \\ \mathcal{P}_\alpha \mathbf{A}^\eta \\ \vdots \\ \mathcal{P}_\alpha \mathbf{A}^{\varepsilon\eta} \end{bmatrix} \mathbf{A}^k \begin{bmatrix} \mathcal{Q}_c & \mathbf{A}^\eta \mathcal{Q}_c & \ldots & \mathbf{A}^{\varsigma\eta} \mathcal{Q}_c \end{bmatrix}$$

ou equivalentemente

$$\mathcal{H}(k) = \mathcal{P}_\varepsilon \mathbf{A}^k \mathcal{Q}_\varsigma \tag{4-20}$$

As matrizes $\mathcal{P}_\varepsilon$ e $\mathcal{Q}_\varsigma$ podem ser chamadas matrizes de observabilidade e controlabilidade de correlação em blocos de dimensão $(\alpha m(\varepsilon + 1) \times n)$ e $(n \times \alpha m(\varsigma + 1))$ respectivamente; e o número enteiro η é escolhido convenientemente para prevenir sobreposição entre blocos $\mathcal{R}_{hh}$ contíguos.

Para k=0 na eq.(4-20), a matriz Hankel de correlação em blocos se reduz à $\mathcal{H}(0) = \mathcal{P}_\varepsilon \mathcal{Q}_\varsigma$, logo após similar ao ERA, no ERA/DC se faz a fatorização de $\mathcal{H}(0)$ usando a decomposição em valores singulares e pegando

PUC-Rio - Certificação Digital Nº 0115595/CA

só os valores singulares significantes, tem-se que

$$\mathcal{H}(0) = \mathcal{P}_\varepsilon \mathcal{Q}_\varsigma = \mathbf{R}_n \Sigma_n \mathbf{S}_n^T \tag{4-21}$$

logo, é possível fazer a seguinte equivalência

$$\mathcal{P}_\varepsilon = \mathbf{R}_n \Sigma_n^{1/2} \quad \text{e} \quad \mathcal{Q}_\varsigma = \Sigma_n^{1/2} \mathbf{S}_n^T \tag{4-22}$$

as primeiras (αm) linhas do $\mathcal{P}_\varepsilon$ será a matriz de observabilidade $\mathcal{P}_\alpha$ e as primeiras (αm) colunas do $\mathcal{Q}_\varsigma$ será a matriz $\mathcal{Q}_c$.

Para k=1 na eq.(4-20) e substituindo $\mathcal{P}_\varepsilon$ e $\mathcal{Q}_\varsigma$ por suas equivalências, obtém-se

$$\mathcal{H}(1) = \mathcal{P}_\varepsilon \mathbf{A} \mathcal{Q}_\varsigma = \mathbf{R}_n \Sigma_n^{1/2} \mathbf{A} \Sigma_n^{1/2} \mathbf{S}_n^T \tag{4-23}$$

então uma mínima realização de $(\mathbf{A}, \mathcal{Q}_c, \mathcal{P}_\alpha)$ é

$$\begin{aligned} \hat{\mathbf{A}} &= \Sigma_n^{-1/2} \mathbf{R}_n^T \mathcal{H}(1) \mathbf{S}_n \Sigma_n^{-1/2} \\ \mathcal{Q}_c &= \Sigma_n^{1/2} \mathbf{S}_n^T \mathbf{E}_\gamma \\ \mathcal{P}_\alpha &= \mathbf{E}_\gamma^T \mathbf{R}_n \Sigma_n^{1/2} \end{aligned} \tag{4-24}$$

sendo a matriz $\mathbf{E}_\gamma^T = [\mathbf{I}_\gamma \ \mathbf{O}_\gamma \cdots \mathbf{O}_\gamma]$; $\mathbf{I}_\gamma$ e $\mathbf{O}_\gamma$ são as matrizes identidade e nula de ordem $\gamma = \alpha m$ respectivamente.

Usando a eq.(4-13), a matriz de controlabilidade calcula-se da seguinte maneira

$$\mathcal{Q}_\beta = \mathcal{P}_\alpha^\dagger \mathbf{H}(0) = \left[\mathbf{E}_\gamma^T \mathbf{R}_n \Sigma_n^{1/2}\right]^\dagger \mathbf{H}(0) \tag{4-25}$$

o símbolo † indica a pseudo-inversa. As matrizes $\mathbf{B}$ e $\mathbf{C}$ podem ser agora estimadas a partir de $\mathcal{Q}_\beta$ e $\mathcal{P}_\alpha$ como no ERA, obtendo-se a realização para $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}})$ seguinte

$$\begin{aligned} \hat{\mathbf{A}} &= \Sigma_n^{-1/2} \mathbf{R}_n^T \mathcal{H}(1) \mathbf{S}_n \Sigma_n^{-1/2} \\ \hat{\mathbf{B}} &= \left[\mathbf{E}_\gamma^T \mathbf{R}_n \Sigma_n^{1/2}\right]^\dagger \mathbf{H}(0) \mathbf{E}_r \\ \hat{\mathbf{C}} &= \mathbf{E}_m^T \mathbf{E}_\gamma^T \mathbf{R}_n \Sigma_n^{1/2} \\ \hat{\mathbf{D}} &= \mathbf{Y}_0 \end{aligned} \tag{4-26}$$

sendo $\mathbf{E}_m$ e $\mathbf{E}_r$ definidos como em (4-16). Igual que no ERA, a realização ($\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$) pode ser transformado nas coordenadas modais obtendo a realização ($\hat{\Lambda}$, $\hat{\Psi}^{-1}\hat{\mathbf{B}}$, $\hat{\mathbf{C}}\hat{\Psi}$, $\hat{\mathbf{D}}$) sendo assim identificados todos os parâmetros modais do sistema dinâmico.

PUC-Rio - Certificação Digital Nº 0115595/CA

4.5 Identificação com observador /filtro de Kalman (OKID)

O ERA e ERA/DC usa os parâmetros de Markov do sistema diretamente obtidos dos dados de entrada e saída, mas os cálculos na identificação de ($\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$) podem ter pouca precisão quando os parâmetros de Markov estão contaminados com ruído; além disso se o sistema tem decaimento lento, ou seja se $\mathbf{A}^k$ não converge para $\mathbf{0}$ rapidamente, é necessário uma maior quantidade de dados o que implica em mais tempo computacional; usando OKID (Observer/Kalman filter identification), um observador de estado é introduzido para obter um modelo discreto no espaço de estados estável, conseguindo comprimir os dados e melhorar os resultados da identificação. Sob certas condições, o observador introduzido pode trabalhar como um filtro de Kalman, permitindo a identificação da quádrupla ($\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$) de um modelo discreto em espaço de estados que considera os ruídos no processo (nas entradas) e os ruídos nas medições (nas saídas) como o modelo estocastico descrito pelas equações (3-1) e (3-2).

4.5.1 Modelo em espaço de estados com observador

Somando e subtraindo o termo $\boldsymbol{G}\boldsymbol{y}(k)$ no lado direito da equação de estado eq.(4-1) temos

$$\begin{aligned}\boldsymbol{x}(k+1) &= \mathbf{A}\boldsymbol{x}(k) + \mathbf{B}\boldsymbol{u}(k) + \boldsymbol{G}\boldsymbol{y}(k) - \boldsymbol{G}\boldsymbol{y}(k) \\ &= (\mathbf{A} + \boldsymbol{G}\mathbf{C})\boldsymbol{x}(k) + (\mathbf{B} + \boldsymbol{G}\mathbf{D})\boldsymbol{u}(k) - \boldsymbol{G}\boldsymbol{y}(k)\end{aligned}$$

definindo,

$$\begin{aligned}\bar{\mathbf{A}} &= (\mathbf{A} + \boldsymbol{G}\mathbf{C}) \\ \bar{\mathbf{B}} &= [(\mathbf{B} + \boldsymbol{G}\mathbf{D}) \quad -\boldsymbol{G}\,] \\ \boldsymbol{v}(k) &= \begin{bmatrix} \boldsymbol{u}(k) \\ \boldsymbol{y}(k) \end{bmatrix}\end{aligned} \tag{4-27}$$

o novo modelo discreto em espaço de estados com observador será

$$\boldsymbol{x}(k+1) = \bar{\mathbf{A}}\boldsymbol{x}(k) + \bar{\mathbf{B}}\boldsymbol{v}(k) \tag{4-28}$$

$$\boldsymbol{y}(k) = \mathbf{C}\boldsymbol{x}(k) + \mathbf{D}\boldsymbol{u}(k) \tag{4-29}$$

A matriz $\boldsymbol{G}\,(n \times m)$ é o ganho do observador e pode ser escolhido arbitrariamente. $\boldsymbol{x}(k)$ é considerado agora como um vetor de observador de estado e os parâmetros de Markov como parâmetros de Markov do observador.

PUC-Rio - Certificação Digital Nº 0115595/CA

Considerando que no processo de identificação existem incertezas introduzidas no sistema devido ao ruído no processo e ruído nas medições, a equação de estado com observador e a equação de medição no tempo discreto terão a forma de observador seguinte

$$\hat{\boldsymbol{x}}(k+1) = \mathbf{A}\hat{\boldsymbol{x}}(k) + \mathbf{B}\boldsymbol{u}(k) - \boldsymbol{G}\left[\boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k)\right] \tag{4-30}$$

$$\hat{\boldsymbol{y}}(k) = \mathbf{C}\hat{\boldsymbol{x}}(k) + \mathbf{D}\boldsymbol{u}(k) \tag{4-31}$$

ou também

$$\hat{\boldsymbol{x}}(k+1) = \bar{\mathbf{A}}\hat{\boldsymbol{x}}(k) + \bar{\mathbf{B}}\boldsymbol{v}(k) \tag{4-32}$$

$$\hat{\boldsymbol{y}}(k) = \mathbf{C}\hat{\boldsymbol{x}}(k) + \mathbf{D}\boldsymbol{u}(k) \tag{4-33}$$

sendo $\hat{\boldsymbol{x}}(k)$ o vetor de observador de estado estimado e $\hat{\boldsymbol{y}}(k)$ o conhecido vetor de medição estimado.

Usando a definição do erro do estado estimado, eq.(3-5), a equação que governa o erro do estado estimado pode ser escrito como

$$\begin{aligned}
\boldsymbol{e}(k+1) &= \boldsymbol{x}(k+1) - \hat{\boldsymbol{x}}(k+1) \\
&= \mathbf{A}\boldsymbol{x}(k) + \mathbf{B}\boldsymbol{u}(k) - \left[\mathbf{A}\hat{\boldsymbol{x}}(k) + \mathbf{B}\boldsymbol{u}(k) - \boldsymbol{G}\left[\boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k)\right]\right] \\
&= (\mathbf{A} + \boldsymbol{G}\mathbf{C})\boldsymbol{e}(k) \quad \text{, (usando eq.(4-29) e eq.(4-33))}
\end{aligned}$$

então,

$$\boldsymbol{e}(k+1) = \bar{\mathbf{A}}\boldsymbol{e}(k) \tag{4-34}$$

Se $\bar{\mathbf{A}}$ é assintoticamente estável, então, para um grande k, o vetor de estado estimado $\hat{\boldsymbol{x}}(k)$ aproxima ao valor real do vetor de estado $\boldsymbol{x}(k)$. Teoricamente, é possível escolher um ganho $\boldsymbol{G}$ para que o erro do estado estimado decresça rapidamente. Na presença de ruído no processo e ruído nas medições, em condições ideais, o observador mais "rápido"é o filtro de Kalman, como será mostrado mais adiante.

4.5.2
Cálculo dos parâmetros de Markov do observador

A relação entre os históricos no tempo discreto da entrada e a saída para as equações (4-28) e (4-29), considerando condições iniciais iguais a zero, pode ser representada pela seguinte equação matricial

$$[\boldsymbol{y}]_{m\times l} = \left[\bar{\mathbf{Y}}\right]_{m\times[(m+r)(l-1)+r]} \left[\mathbf{V}\right]_{[(m+r)(l-1)+r]\times l} \tag{4-35}$$

sendo l o número de amostras ou dados medidos, e

$$\boldsymbol{y} = \left[\begin{array}{ccccccc} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(p) & \cdots & \boldsymbol{y}(l-1) \end{array} \right]$$

$$\bar{\mathbf{Y}} = \left[\begin{array}{ccccccc} \mathbf{D} & \mathbf{C}\bar{\mathbf{B}} & \mathbf{C}\bar{\mathbf{A}}\bar{\mathbf{B}} & \cdots & \mathbf{C}\bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}} & \cdots & \mathbf{C}\bar{\mathbf{A}}^{l-2}\bar{\mathbf{B}} \end{array} \right]$$

$$\mathbf{V} = \left[\begin{array}{ccccccc} \boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(p) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{v}(0) & \boldsymbol{v}(1) & \cdots & \boldsymbol{v}(p-1) & \cdots & \boldsymbol{v}(l-2) \\ & & \boldsymbol{v}(0) & \cdots & \boldsymbol{v}(p-2) & \cdots & \boldsymbol{v}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \boldsymbol{v}(0) & \cdots & \boldsymbol{v}(l-p-1) \\ & & & & & \ddots & \vdots \\ & & & & & & \boldsymbol{v}(0) \end{array} \right]$$

a matriz $\bar{\mathbf{Y}}$ contêm a seqüência de matrizes que serão chamados como parâmetros de Markov do observador, definidos como

$$\bar{\mathbf{Y}}_k = \begin{cases} \mathbf{D}, & k = 0 \\ \mathbf{C}\bar{\mathbf{A}}^{k-1}\bar{\mathbf{B}}, & k \geqslant 1\,. \end{cases} \tag{4-36}$$

Quando um ganho adequado de observador é usado, é possível obter um sistema com observador que tenha decaimento rápido dos seus parâmetros de Markov; por exemplo, usar um observador que aproxime os polos de $\bar{\mathbf{A}}$ ao origem no plano complexo [22, 35] tal que $\mathbf{C}\bar{\mathbf{A}}^{k-1}\bar{\mathbf{B}} = 0$ para $k \geqslant p$. Ainda quando os dados estão contaminados com ruído, é possível obter $\mathbf{C}\bar{\mathbf{A}}^{k-1}\bar{\mathbf{B}} \approx 0$ para $k \geqslant p$, sendo p suficientemente grande, e a eq.(4-35) pode ser reduzida para

$$[\boldsymbol{y}]_{m\times l} = \left[\bar{\mathbf{Y}}\right]_{m\times[(m+r)p+r]} [\mathbf{V}]_{[(m+r)p+r]\times l} \tag{4-37}$$

sendo,

$$\boldsymbol{y} = \left[\begin{array}{ccccccc} \boldsymbol{y}(0) & \boldsymbol{y}(1) & \boldsymbol{y}(2) & \cdots & \boldsymbol{y}(p) & \cdots & \boldsymbol{y}(l-1) \end{array} \right]$$

$$\bar{\mathbf{Y}} = \left[\begin{array}{ccccc} \mathbf{D} & \mathbf{C}\bar{\mathbf{B}} & \mathbf{C}\bar{\mathbf{A}}\bar{\mathbf{B}} & \cdots & \mathbf{C}\bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}} \end{array} \right]$$

$$\mathbf{V} = \left[\begin{array}{ccccccc} \boldsymbol{u}(0) & \boldsymbol{u}(1) & \boldsymbol{u}(2) & \cdots & \boldsymbol{u}(p) & \cdots & \boldsymbol{u}(l-1) \\ & \boldsymbol{v}(0) & \boldsymbol{v}(1) & \cdots & \boldsymbol{v}(p-1) & \cdots & \boldsymbol{v}(l-2) \\ & & \boldsymbol{v}(0) & \cdots & \boldsymbol{v}(p-2) & \cdots & \boldsymbol{v}(l-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & \boldsymbol{v}(0) & \cdots & \boldsymbol{v}(l-p-1) \end{array} \right]$$

PUC-Rio - Certificação Digital Nº 0115595/CA

então, para um $l \geqslant (r+m)p + r$, os primeiros p parâmetros de Markov do observador satisfazem a seguinte equação

$$\bar{\mathbf{Y}} = \boldsymbol{y}\mathbf{V}^{\dagger} \tag{4-38}$$

sendo $\mathbf{V}^{\dagger}$ a pseudo inversa de $\mathbf{V}$ e o erro de aproximação decresce quando aumenta p.

Para condições iniciais diferente de zero, é necessário fazer uma outra formulação para eliminar a influencia das condições iniciais no cálculo dos parâmetros de Markov do observador. Usando a eq.(4-28) obtém-se a seguinte seqüência

$$\begin{aligned}
\boldsymbol{x}(k+1) &= \bar{\mathbf{A}}\boldsymbol{x}(k) + \bar{\mathbf{B}}\boldsymbol{v}(k), \\
\boldsymbol{x}(k+2) &= \bar{\mathbf{A}}\boldsymbol{x}(k+1) + \bar{\mathbf{B}}\boldsymbol{v}(k+1) \\
&= \bar{\mathbf{A}}^2\boldsymbol{x}(k) + \bar{\mathbf{A}}\bar{\mathbf{B}}\boldsymbol{v}(k) + \bar{\mathbf{B}}\boldsymbol{v}(k+1), \\
&\vdots \\
\boldsymbol{x}(k+p) &= \bar{\mathbf{A}}\boldsymbol{x}(k+p-1) + \bar{\mathbf{B}}\boldsymbol{v}(k+p-1) \\
&= \bar{\mathbf{A}}^p\boldsymbol{x}(k) + \bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}}\boldsymbol{v}(k) + \bar{\mathbf{A}}^{p-2}\bar{\mathbf{B}}\boldsymbol{v}(k+1) + \cdots + \bar{\mathbf{B}}\boldsymbol{v}(k+p-1),
\end{aligned}$$

e substituindo na eq.(4-29) obtém-se

$$\begin{aligned}
\boldsymbol{y}(k+p) &= \mathbf{C}\boldsymbol{x}(k+p) + \mathbf{D}\boldsymbol{u}(k+p) \\
&= \mathbf{C}\bar{\mathbf{A}}^p\boldsymbol{x}(k) + \mathbf{C}\bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}}\boldsymbol{v}(k) + \mathbf{C}\bar{\mathbf{A}}^{p-2}\bar{\mathbf{B}}\boldsymbol{v}(k+1) + \cdots \\
&\quad + \mathbf{C}\bar{\mathbf{B}}\boldsymbol{v}(k+p-1) + \mathbf{D}\boldsymbol{u}(k+p)
\end{aligned}$$

O conjunto de equações obtidas para $k = 0, 1, \ldots, (l-p-1)$ pode ser escrita na forma matricial como

$$\bar{\boldsymbol{y}} = \mathbf{C}\bar{\mathbf{A}}^p\boldsymbol{x} + \bar{\mathbf{Y}}\bar{\mathbf{V}} \tag{4-39}$$

sendo,

$$\begin{aligned}
\bar{\boldsymbol{y}} &= \begin{bmatrix} \boldsymbol{y}(p) & \boldsymbol{y}(p+1) & \cdots & \boldsymbol{y}(l-1) \end{bmatrix} \\
\boldsymbol{x} &= \begin{bmatrix} \boldsymbol{x}(0) & \boldsymbol{x}(1) & \cdots & \boldsymbol{x}(l-p-1) \end{bmatrix} \\
\bar{\mathbf{Y}} &= \begin{bmatrix} \mathbf{D} & \mathbf{C}\bar{\mathbf{B}} & \mathbf{C}\bar{\mathbf{A}}\bar{\mathbf{B}} & \cdots & \mathbf{C}\bar{\mathbf{A}}^{p-1}\bar{\mathbf{B}} \end{bmatrix}
\end{aligned}$$

PUC-Rio - Certificação Digital Nº 0115595/CA

$$\bar{\mathbf{V}} = \begin{bmatrix} \boldsymbol{u}(p) & \boldsymbol{u}(p+1) & \cdots & \boldsymbol{u}(l-1) \\ \boldsymbol{v}(p-1) & \boldsymbol{v}(p) & \cdots & \boldsymbol{v}(l-2) \\ \boldsymbol{v}(p-2) & \boldsymbol{v}(p-1) & \cdots & \boldsymbol{v}(l-3) \\ \vdots & \vdots & & \vdots \\ \boldsymbol{v}(0) & \boldsymbol{v}(1) & \cdots & \boldsymbol{v}(l-p-1) \end{bmatrix}$$

Quando $\bar{\mathbf{A}}^p \approx \mathbf{0}$ para p suficientemente grande, o termo $\mathbf{C}\bar{\mathbf{A}}^p\boldsymbol{x}$ que representa a influência das condições iniciais, pode ser considerado desprezível e a eq.(4-39) pode ser aproximada como

$$\bar{\boldsymbol{y}} = \bar{\mathbf{Y}}\bar{\mathbf{V}} \tag{4-40}$$

a qual tem uma solução de mínimos quadrados seguinte

$$\bar{\mathbf{Y}} = \bar{\boldsymbol{y}}\bar{\mathbf{V}}^T \left[\bar{\mathbf{V}}\bar{\mathbf{V}}^T\right]^{-1} \tag{4-41}$$

sempre que $\left[\bar{\mathbf{V}}\bar{\mathbf{V}}^T\right]^{-1}$ exista, caso contrário usa-se

$$\bar{\mathbf{Y}} = \bar{\boldsymbol{y}}\bar{\mathbf{V}}^{\dagger}. \tag{4-42}$$

4.5.3 Cálculo dos parâmetros de Markov do sistema

A seqüência de parâmetros de Markov do observador, $\bar{\mathbf{Y}}_i$, podem ser representados como uma partição matricial da seguinte maneira

$$\begin{aligned} \bar{\mathbf{Y}}_0 &= \mathbf{D} \\ \bar{\mathbf{Y}}_k &= \mathbf{C}\bar{\mathbf{A}}^{k-1}\bar{\mathbf{B}} \\ &= \left[\bar{\mathbf{Y}}_k^{(1)} \quad -\bar{\mathbf{Y}}_k^{(2)}\right] \; ; \, k = 1,\, 2,\, 3, \ldots \end{aligned} \tag{4-43}$$

os índices (1) e (2) indicam as partições de $\bar{\mathbf{Y}}_k$ definidas como:

$$\begin{aligned} \bar{\mathbf{Y}}_k^{(1)} &= \mathbf{C}\left(\mathbf{A} + \boldsymbol{G}\mathbf{C}\right)^{k-1}\left(\mathbf{B} + \boldsymbol{G}\mathbf{D}\right) \\ \bar{\mathbf{Y}}_k^{(2)} &= \mathbf{C}\left(\mathbf{A} + \boldsymbol{G}\mathbf{C}\right)^{k-1}\boldsymbol{G} \end{aligned} \tag{4-44}$$

Os parâmetros de Markov do sistema podem ser calculados à partir

PUC-Rio - Certificação Digital Nº 0115595/CA

dos parâmetros de Markov do observador usando as seguintes relações:

$$\begin{aligned}\mathbf{Y}_0 &= \bar{\mathbf{Y}}_0 = \mathbf{D} \\ \mathbf{Y}_k &= \bar{\mathbf{Y}}_k^{(1)} - \sum_{i=1}^{k} \bar{\mathbf{Y}}_i^{(2)} \mathbf{Y}_{k-i} \quad ; k = 1, \dots, p \\ \mathbf{Y}_k &= -\sum_{i=1}^{p} \bar{\mathbf{Y}}_i^{(2)} \mathbf{Y}_{k-i} \quad ; k = p+1, \dots, \infty\end{aligned} \tag{4-45}$$

Dado que há $(p+1)$ parâmetros de Markov do observador, das equações (4-45) pode-se deduzir que há só p parâmetros de Markov do sistema independentes e os restantes $\mathbf{Y}_k$, para $k \geqslant p+1$, serão uma combinação linear dos primeiros.

Uma vez obtidos os parâmetros de Markov do sistema, pode-se usar o ERA ou ERA/DC para a identificação dos parâmetros modais do sistema dinâmico, as duas variantes obtidas são conhecidas como OKID/ERA e OKID/ERA/DC.

PUC-Rio - Certificação Digital Nº 0115595/CA

4.5.4 Relação do modelo na forma do observador com o filtro de Kalman

Usando as equações (3-18) e (3-19) do filtro de Kalman em estado estável definidas no capítulo 3, é possível expressar a equação de estado na forma

$$\begin{aligned}\hat{\boldsymbol{x}}(k+1) &= \mathbf{A}\hat{\boldsymbol{x}}(k) + \mathbf{B}\boldsymbol{u}(k) + \boldsymbol{K}\epsilon(k) \\ &= \mathbf{A}\hat{\boldsymbol{x}}(k) + \mathbf{B}\boldsymbol{u}(k) + \boldsymbol{K}\left[\boldsymbol{y}(k) - (\mathbf{C}\hat{\boldsymbol{x}}(k) + \mathbf{D}\boldsymbol{u}(k))\right] \\ &= (\mathbf{A} - \boldsymbol{K}\mathbf{C})\hat{\boldsymbol{x}}(k) + (\mathbf{B} - \boldsymbol{K}\mathbf{D})\boldsymbol{u}(k) + \boldsymbol{K}\boldsymbol{y}(k)\end{aligned}$$

definindo,

$$\begin{aligned}\tilde{\mathbf{A}} &= (\mathbf{A} - \boldsymbol{K}\mathbf{C}) \\ \tilde{\mathbf{B}} &= [(\mathbf{B} - \boldsymbol{K}\mathbf{D}) \quad \boldsymbol{K}] \\ \boldsymbol{v}(k) &= \begin{bmatrix} \boldsymbol{u}(k) \\ \boldsymbol{y}(k) \end{bmatrix}\end{aligned} \tag{4-46}$$

o filtro de Kalman em estado estável para um sistema representado pelo modelo discreto estocástico em espaço de estados pode ser escrito como

$$\hat{\boldsymbol{x}}(k+1) = \tilde{\mathbf{A}}\hat{\boldsymbol{x}}(k) + \tilde{\mathbf{B}}\boldsymbol{v}(k) \tag{4-47}$$

$$\boldsymbol{y}(k) = \mathbf{C}\hat{\boldsymbol{x}}(k) + \mathbf{D}\boldsymbol{u}(k) + \epsilon(k) \tag{4-48}$$

Comparando as equações (4-47) e (4-48) com as equações (4-32) e (4-33) respectivamente, observa-se que elas são equivalentes se $\boldsymbol{G} = -\boldsymbol{K}$ e terão os mesmos parâmetros de Markov; e teoricamente se $\epsilon(k) = 0$ serão equivalentes das equações (4-28) e (4-29). Mas, algumas condições têm que ser satisfeitas para que o ganho de observador seja equivalente ao ganho do filtro de Kalman, as quais serão deduzidas no seguinte análise.

Considerando as equações (4-47) e (4-48), de maneira similar que eq.(4-39), pode-se obter a seguinte equação matricial

$$\bar{\boldsymbol{y}} = \mathbf{C}\tilde{\mathbf{A}}^p \hat{\boldsymbol{x}} + \tilde{\mathbf{Y}}\bar{\mathbf{V}} + \epsilon \tag{4-49}$$

$\bar{\boldsymbol{y}}$ e $\bar{\mathbf{V}}$ são matrizes como definidas em eq.(4-39), e

$$\begin{aligned} \hat{\boldsymbol{x}} &= \begin{bmatrix} \hat{\boldsymbol{x}}(0) & \hat{\boldsymbol{x}}(1) & \cdots & \hat{\boldsymbol{x}}(l-p-1) \end{bmatrix} \\ \tilde{\mathbf{Y}} &= \begin{bmatrix} \mathbf{D} & \mathbf{C}\tilde{\mathbf{B}} & \mathbf{C}\tilde{\mathbf{A}}\tilde{\mathbf{B}} & \cdots & \mathbf{C}\tilde{\mathbf{A}}^{p-1}\tilde{\mathbf{B}} \end{bmatrix} \\ \epsilon &= \begin{bmatrix} \epsilon(p) & \epsilon(p+1) & \cdots & \epsilon(l-1) \end{bmatrix} \end{aligned}$$

sendo l a longitude dos dados medidos.

Pós-multiplicando (4-49) por $\bar{\mathbf{V}}^T$,

$$\bar{\mathbf{V}}^T = \begin{bmatrix} \boldsymbol{v}_p^T & \boldsymbol{v}_{p-1}^T & \cdots & \boldsymbol{v}_0^T \end{bmatrix} \tag{4-50}$$

sendo

$$\boldsymbol{v}_i = \begin{bmatrix} \boldsymbol{v}(i) & \boldsymbol{v}(i+1) & \cdots & \boldsymbol{v}(l-p-1+i) \end{bmatrix}; \quad i = 0, 1, 2, \ldots, p$$

a equação (4-49) pode ser escrito como

$$\bar{\boldsymbol{y}}\bar{\mathbf{V}}^T - \tilde{\mathbf{Y}}\bar{\mathbf{V}}\bar{\mathbf{V}}^T = \mathbf{C}\tilde{\mathbf{A}}^p \begin{bmatrix} \hat{\boldsymbol{x}}\boldsymbol{v}_p^T & \hat{\boldsymbol{x}}\boldsymbol{v}_{p-1}^T & \cdots & \hat{\boldsymbol{x}}\boldsymbol{v}_0^T \end{bmatrix} + \begin{bmatrix} \epsilon\boldsymbol{v}_p^T & \epsilon\boldsymbol{v}_{p-1}^T & \cdots & \epsilon\boldsymbol{v}_0^T \end{bmatrix} \tag{4-51}$$

Observa-se que os termos $\epsilon\boldsymbol{v}_i^T$, no lado direito da equação (4-51), podem ser decompostos como

$$\epsilon\boldsymbol{v}_i^T = \sum_{j=0}^{l-p-1} \epsilon(p+j)\boldsymbol{v}^T(i+j) = \sum_{k=p}^{l-1} \epsilon(k)\boldsymbol{v}^T(k-p+i); \quad i = 0, 1, \ldots, p \tag{4-52}$$

dividindo este termo por $(l-p)$, é possível aplicar a propriedade ergódica

PUC-Rio - Certificação Digital Nº 0115595/CA

sempre que $l \rightarrow \infty$, achando a seguinte equivalência

$$\begin{aligned}
\lim_{l\rightarrow\infty} \frac{1}{l-p}\left(\epsilon \boldsymbol{v}_i^T\right) &= \lim_{l\rightarrow\infty} \frac{1}{l-p}\sum_{k=p}^{l-1} \epsilon(k)\boldsymbol{v}^T(k-p+i); \quad k>p \\
&= E\left[\epsilon(k)\boldsymbol{v}^T(k-p+i)\right]
\end{aligned} \tag{4-53}$$

Similarmente, para os termos $\hat{\boldsymbol{x}}\boldsymbol{v}_i^T$,

$$\lim_{l\rightarrow\infty} \frac{1}{l-p}\left(\hat{\boldsymbol{x}}\boldsymbol{v}_i^T\right) = E\left[\hat{\boldsymbol{x}}(k)\boldsymbol{v}^T(k+i)\right] \tag{4-54}$$

Logo, dividindo os termos na eq.(4-51) por $(l-p)$ e avaliando o limite quando $l \rightarrow \infty$, obtém-se

$$\lim_{l\rightarrow\infty} \frac{1}{l-p}\left[\bar{\boldsymbol{y}}\bar{\mathbf{V}}^T - \tilde{\mathbf{Y}}\bar{\mathbf{V}}\bar{\mathbf{V}}^T\right]$$

$$\begin{aligned}
&= \mathbf{C}\tilde{\mathbf{A}}^p \lim_{l\rightarrow\infty} \frac{1}{l-p}\left[\hat{\boldsymbol{x}}\boldsymbol{v}_p^T \quad \hat{\boldsymbol{x}}\boldsymbol{v}_{p-1}^T \quad \cdots \quad \hat{\boldsymbol{x}}\boldsymbol{v}_0^T\right] \\
&\quad + \lim_{l\rightarrow\infty} \frac{1}{l-p}\left[\epsilon\boldsymbol{v}_p^T \quad \epsilon\boldsymbol{v}_{p-1}^T \quad \cdots \quad \epsilon\boldsymbol{v}_0^T\right]
\end{aligned} \tag{4-55}$$

$$\begin{aligned}
&= \mathbf{C}\tilde{\mathbf{A}}^p \; E\left[\hat{\boldsymbol{x}}(k)\boldsymbol{v}^T(k+p) \quad \hat{\boldsymbol{x}}(k)\boldsymbol{v}^T(k+p-1) \quad \cdots \quad \hat{\boldsymbol{x}}(k)\boldsymbol{v}^T(k)\right] \\
&\quad + E\left[\epsilon(k)\boldsymbol{v}^T(k) \quad \epsilon(k)\boldsymbol{v}^T(k-1) \quad \cdots \quad \epsilon(k)\boldsymbol{v}^T(k-p)\right]
\end{aligned} \tag{4-56}$$

Sob a condição que seja escolhido um p suficientemente grande, tal que $\mathbf{C}\tilde{\mathbf{A}}^p \approx 0$, e se o observador escolhido satisfaz que

$$\tilde{\mathbf{Y}} = \bar{\boldsymbol{y}}\bar{\mathbf{V}}^T\left[\bar{\mathbf{V}}\bar{\mathbf{V}}^T\right]^{-1} \tag{4-57}$$

no limite $l \rightarrow \infty$, então, a eq.(4-56) é reduzida para

$$0 = E\left[\epsilon(k)\boldsymbol{v}^T(k) \quad \epsilon(k)\boldsymbol{v}^T(k-1) \quad \cdots \quad \epsilon(k)\boldsymbol{v}^T(k-p)\right] \tag{4-58}$$

para $k > p$. Usando a definição de $\boldsymbol{v}(k)$ da eq.(4-27), a eq.(4-58) implica as seguintes condições de ortogonalidade

$$\begin{aligned}
E\left[\epsilon(k)\boldsymbol{u}^T(k-i)\right] &= \mathbf{0}; \quad i = 0, 1, \ldots, p; \quad k>p \\
E\left[\epsilon(k)\boldsymbol{y}^T(k-j)\right] &= \mathbf{0}; \quad j = 1, 2, \ldots, p; \quad k>p
\end{aligned} \tag{4-59}$$

As equações (4-59) satisfazem os mesmos princípios de ortogonalidade que no caso do filtro de Kalman descritas nas equações (3-15). Assim, se o observador escolhido satisfaz a solução de quadrados mínimos dada pelas

PUC-Rio - Certificação Digital Nº 0115595/CA

equações (4-41), (4-42) ou (4-57), todas elas produziram as mesmas relações entrada-saída que o filtro de Kalman em estado estável para as equações estocasticas (3-1) e (3-2); ou seja, $\boldsymbol{G} = -\boldsymbol{K}$, sempre que o número de amostras, l, seja suficientemente grande e o erro por truncar os termos $(\mathbf{C}\bar{\mathbf{A}}^p\boldsymbol{x})$, para p suficientemente grande seja desprezível.

Na prática, devido a possíveis não linearidades, ou quando os ruídos no processo e nas medições não são perfeitamente brancos, o filtro obtido não será o filtro de Kalman, mas um observador que minimiza o resíduo do filtragem $\epsilon(k)$.

PUC-Rio - Certificação Digital Nº 0115595/CA