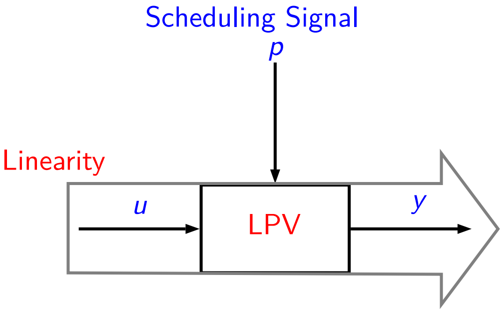
Kernelized Identification of Linear Parameter-Varying Models with Linear Fractional Representation

Manas Mejari, Dario Piga, Roland Tóth and Alberto Bemporad

IDSIA Lugano, TU/e Eindhoven, IMT Lucca

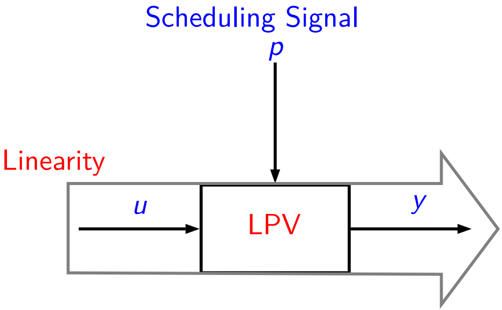
European Control Conference ECC 2019, Naples, Italy.
26 June, 2019

Linear Parameter-Varying (LPV) Concept



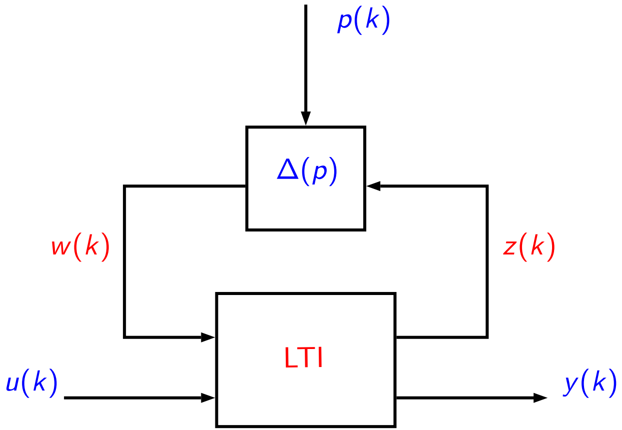
- Linear dynamic relation between input and output.
- Unlike Linear Time-Invariant (LTI), the relation changes according to a measurable time-varying 'scheduling signal' p.

Linear Parameter-Varying (LPV) Concept



- Linear dynamic relation between input and output.
- Unlike Linear Time-Invariant (LTI), the relation changes according to a measurable time-varying 'scheduling signal' p.

Linear Fractional Representations



Linear Fractional Representations

Forward LTI part is given by,

$$\left[\begin{array}{c} x(k+1) \\ \hline z(k) \\ y(k) \end{array}\right] = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \left[\begin{array}{c} x(k) \\ \hline w(k) \\ u(k) \end{array}\right],$$

The feedback path is represented by

$$w(k) = \Delta(p(k))z(k)$$

- Assumption: $\Delta(p(k))$ is known.
- Goal: Given $\{y(k), u(k), p(k)\}_{k=1}^{N}$, estimate the model parameter matrices $\{A, B1, \ldots, D_{22}\}$.

Linear Fractional Representations

Forward LTI part is given by,

$$\left[\begin{array}{c} x(k+1) \\ \hline z(k) \\ y(k) \end{array}\right] = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \left[\begin{array}{c} x(k) \\ \hline w(k) \\ u(k) \end{array}\right],$$

The feedback path is represented by

$$w(k) = \Delta(p(k))z(k)$$

- Assumption: $\Delta(p(k))$ is known.
- Goal: Given $\{y(k), u(k), p(k)\}_{k=1}^{N}$, estimate the model parameter matrices $\{A, B1, \ \ldots, \ D_{22}\}$.

Identification of LPV-LFR models

The developed approach consists of the following two stages:

S1 Estimation of $\hat{x}(k)$ using Kernel Canonical Correlation Analysis (KCCA).

S2 Estimation of latent variable $z(k)$ and model parameter matrices $\{A, \ldots, D_{22}\}$ solving non-linear least squares.

- For the affine case ($D_{11} = 0$), this reduces to ordinary least-squares, followed by economic SVD.

Identification of LPV-LFR models

The developed approach consists of the following two stages:

S1 Estimation of $\hat{x}(k)$ using Kernel Canonical Correlation Analysis (KCCA).

S2 Estimation of latent variable $z(k)$ and model parameter matrices $\{A, \ \ldots, \ D_{22}\}$ solving non-linear least squares.

- For the affine case ($D_{11} = 0$), this reduces to ordinary least-squares, followed by economic SVD.

Identification of LPV-LFR models

The developed approach consists of the following two stages:

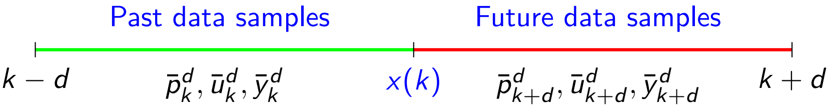
S1 Estimation of $\hat{x}(k)$ using Kernel Canonical Correlation Analysis (KCCA).

S2 Estimation of latent variable $z(k)$ and model parameter matrices $\{A, \ \ldots, \ D_{22}\}$ solving non-linear least squares.

- For the affine case $(D_{11} = 0)$, this reduces to ordinary least-squares, followed by economic SVD.

Kernel Canonical Correlation Analysis for state estimation

Key idea: State $x(k)$ is the minimal interface between past and future inputs u, outputs y, and scheduling data samples p.



$$x(k) = \underbrace{\varphi_f(\bar{p}_{k+d}^d)}_{\text{Unknown}} \begin{bmatrix} \bar{u}_{k+d}^d \\ \bar{y}_{k+d}^d \end{bmatrix} \quad \text{Future data}$$

$$x(k) = \underbrace{\varphi_p(\bar{p}_k^d)}_{\text{Unknown}} \begin{bmatrix} \bar{u}_k^d \\ \bar{y}_k^d \end{bmatrix} \quad \text{Past data}$$

Kernel Canonical Correlation Analysis for state estimation

Key idea: State $x(k)$ is the minimal interface between past and future inputs u, outputs y, and scheduling data samples p.

Past data samples | Future data samples

$k-d$ $\bar{p}_k^d, \bar{u}_k^d, \bar{y}_k^d$ $x(k)$ $\bar{p}_{k+d}^d, \bar{u}_{k+d}^d, \bar{y}_{k+d}^d$ $k+d$

$$x(k) = \underbrace{\varphi_f(\bar{p}_{k+d}^d)}_{\text{Unknown}} \begin{bmatrix} \bar{u}_{k+d}^d \\ \bar{y}_{k+d}^d \end{bmatrix} \quad \text{Future data}$$

$$x(k) = \underbrace{\varphi_p(\bar{p}_k^d)}_{\text{Unknown}} \begin{bmatrix} \bar{u}_k^d \\ \bar{y}_k^d \end{bmatrix} \quad \text{Past data}$$

Kernel Canonical Correlation Analysis for state estimation

Key idea: State $x(k)$ is the minimal interface between past and future inputs u, outputs y, and scheduling data samples p.

Past data samples | Future data samples

$k-d$ | $x(k)$ | $k+d$

$$\underbrace{\varphi_p(\bar{p}_k^d)}_{\text{Unknown}} \begin{bmatrix} \bar{u}_k^d \\ \bar{y}_k^d \end{bmatrix} \qquad \underbrace{\varphi_f(\bar{p}_{k+d}^d)}_{\text{Unknown}} \begin{bmatrix} \bar{u}_{k+d}^d \\ \bar{y}_{k+d}^d \end{bmatrix}$$

- State x(k) is estimated by maximizing the correlation between past and future data.

Kernel Canonical Correlation Analysis for state estimation

KCCA idea: Maximize the correlation between past and future data.

$$x(k) = \underbrace{\varphi_f(\bar{p}^d_{k+d})}_{\text{Unknown}} \underbrace{\begin{bmatrix} \bar{u}^d_{k+d} \\ \bar{y}^d_{k+d} \end{bmatrix}}_{\bar{z}^d_{k+d}} = \underbrace{\varphi_p(\bar{p}^d_k)}_{\text{Unknown}} \underbrace{\begin{bmatrix} \bar{u}^d_k \\ \bar{y}^d_k \end{bmatrix}}_{\bar{z}^d_k}$$

Problem can be formulated and solved using Least Square-Support Vector Machine (LS-SVM).

$$\max_{v_j, w_j} \sum_{k=1}^{N} \left(\gamma s_k r_k - \gamma_f \frac{1}{2} s_k^2 - \gamma_p \frac{1}{2} r_k^2 \right) - \frac{1}{2} v_j^\top v_j - \frac{1}{2} w_j^\top w_j$$

$$\text{s.t. } s_k = v_j^\top \varphi_f(\bar{p}^d_{k+d}) \bar{z}^d_{k+d},\ r_k = w_j^\top \varphi_p(\bar{p}^d_k) \bar{z}^d_k, \forall k = \mathbb{I}_1^N,$$

Kernel Canonical Correlation Analysis for state estimation

KCCA idea: Maximize the correlation between past and future data.

$$x(k) = \underbrace{\varphi_f(\bar{p}^d_{k+d})}_{\text{Unknown}} \underbrace{\begin{bmatrix} \bar{u}^d_{k+d} \\ \bar{y}^d_{k+d} \end{bmatrix}}_{\bar{z}^d_{k+d}} = \underbrace{\varphi_p(\bar{p}^d_k)}_{\text{Unknown}} \underbrace{\begin{bmatrix} \bar{u}^d_k \\ \bar{y}^d_k \end{bmatrix}}_{\bar{z}^d_k}$$

Problem can be formulated and solved using Least Square-Support Vector Machine (LS-SVM).

$$\max_{v_j, w_j} \sum_{k=1}^{N} \left(\gamma s_k r_k - \gamma_f \frac{1}{2} s_k^2 - \gamma_p \frac{1}{2} r_k^2 \right) - \frac{1}{2} v_j^\top v_j - \frac{1}{2} w_j^\top w_j$$
$$\text{s.t. } s_k = v_j^\top \varphi_f(\bar{p}^d_{k+d}) \bar{z}^d_{k+d}, \; r_k = w_j^\top \varphi_p(\bar{p}^d_k) \bar{z}^d_k, \forall k = \mathbb{I}_1^N,$$

Kernel Canonical Correlation Analysis for state estimation

The primal LS-SVM problem:

$$\max_{v_j, w_j} \sum_{k=1}^{N} \left(\gamma s_k r_k - \gamma_f \frac{1}{2} s_k^2 - \gamma_p \frac{1}{2} r_k^2 \right) - \frac{1}{2} v_j^\top v_j - \frac{1}{2} w_j^\top w_j$$

$$\text{s.t. } s_k = v_j^\top \varphi_f(\bar{p}_{k+d}^d) \bar{z}_{k+d}^d, \ \ r_k = w_j^\top \varphi_p(\bar{p}_k^d) \bar{z}_k^d, \forall k = \mathbb{I}_1^N,$$

The optimal LS-SVM dual variables satisfy generalized eigenvalue problem :

$$\begin{bmatrix} 0 & K_{\mathrm{pp}} \\ K_{\mathrm{ff}} & 0 \end{bmatrix} \begin{bmatrix} \eta_j \\ \kappa_j \end{bmatrix} = \lambda_j \begin{bmatrix} \gamma_f K_{\mathrm{ff}} + I & 0 \\ 0 & \gamma_p K_{\mathrm{pp}} + I \end{bmatrix} \begin{bmatrix} \eta_j \\ \kappa_j \end{bmatrix},$$

$$[K_{\mathrm{ff}}]_{l,m} = (\bar{z}_{l+d}^d)^\top \underbrace{\varphi_f^\top(\bar{p}_{l+d}^d) \varphi_f(\bar{p}_{m+d}^d)}_{\bar{k}(\bar{p}_{l+d}^d, \bar{p}_{m+d}^d)} \bar{z}_{m+d}^d,$$

Kernel Canonical Correlation Analysis for state estimation

The primal LS-SVM problem:

$$\max_{v_j, w_j} \sum_{k=1}^{N} \left(\gamma s_k r_k - \gamma_f \frac{1}{2} s_k^2 - \gamma_p \frac{1}{2} r_k^2 \right) - \frac{1}{2} v_j^\top v_j - \frac{1}{2} w_j^\top w_j$$

$$\text{s.t. } s_k = v_j^\top \varphi_f(\bar{p}_{k+d}^d) \bar{z}_{k+d}^d, \; r_k = w_j^\top \varphi_p(\bar{p}_k^d) \bar{z}_k^d, \forall k = \mathbb{I}_1^N,$$

The optimal LS-SVM dual variables satisfy generalized eigenvalue problem :

$$\begin{bmatrix} 0 & K_{\mathrm{pp}} \\ K_{\mathrm{ff}} & 0 \end{bmatrix} \begin{bmatrix} \eta_j \\ \kappa_j \end{bmatrix} = \lambda_j \begin{bmatrix} \gamma_f K_{\mathrm{ff}} + I & 0 \\ 0 & \gamma_p K_{\mathrm{pp}} + I \end{bmatrix} \begin{bmatrix} \eta_j \\ \kappa_j \end{bmatrix},$$

$$[K_{\mathrm{ff}}]_{l,m} = (\bar{z}_{l+d}^d)^\top \underbrace{\varphi_f^\top(\bar{p}_{l+d}^d) \varphi_f(\bar{p}_{m+d}^d)}_{\bar{k}(\bar{p}_{l+d}^d, \bar{p}_{m+d}^d)} \bar{z}_{m+d}^d,$$

Kernel Canonical Correlation Analysis for state estimation

The primal LS-SVM problem:

$$\max_{v_j, w_j} \sum_{k=1}^{N} \left(\gamma s_k r_k - \gamma_f \frac{1}{2} s_k^2 - \gamma_p \frac{1}{2} r_k^2 \right) - \frac{1}{2} v_j^\top v_j - \frac{1}{2} w_j^\top w_j$$

$$\text{s.t. } s_k = v_j^\top \varphi_f(\bar{p}_{k+d}^d) \bar{z}_{k+d}^d, \; r_k = w_j^\top \varphi_p(\bar{p}_k^d) \bar{z}_k^d, \forall k = \mathbb{I}_1^N,$$

The optimal LS-SVM dual variables satisfy generalized eigenvalue problem :

$$\begin{bmatrix} 0 & K_{\mathrm{pp}} \\ K_{\mathrm{ff}} & 0 \end{bmatrix} \begin{bmatrix} \eta_j \\ \kappa_j \end{bmatrix} = \lambda_j \begin{bmatrix} \gamma_f K_{\mathrm{ff}} + I & 0 \\ 0 & \gamma_p K_{\mathrm{pp}} + I \end{bmatrix} \begin{bmatrix} \eta_j \\ \kappa_j \end{bmatrix},$$

$$[K_{\mathrm{ff}}]_{l,m} = (\bar{z}_{l+d}^d)^\top \underbrace{\varphi_f^\top(\bar{p}_{l+d}^d) \varphi_f(\bar{p}_{m+d}^d)}_{\bar{k}(\bar{p}_{l+d}^d, \bar{p}_{m+d}^d)} \bar{z}_{m+d}^d,$$

Kernel Canonical Correlation Analysis for state estimation

- The state sequence $\hat{x}(k)$ is reconstructed (upto a similarity transform)

$$[\hat{x}(k)]_j = \eta_j^\top \begin{bmatrix} \bar{z}_{1+d}^{d\top} \bar{k}(\bar{p}_{1+d}^d, \bar{p}_{k+d}^d) \\ \vdots \\ \bar{z}_{N+d}^{d\top} \bar{k}(\bar{p}_{N+d}^d, \bar{p}_{k+d}^d) \end{bmatrix} \bar{z}_{k+d}^d.$$

- $\bar{k}(\bar{p}_{1+d}^d, \bar{p}_{k+d}^d)$ positive definite kernel functions.
  For eg., RBF kernel $\bar{k}(p_i, p_j) = c \, \exp\left(-\frac{\|p_i - p_j\|^2}{\sigma^2}\right)$.

Identification of LPV-LFR models

- Once $\hat{x}(k)$ is estimated, latent variable $z(k)$ and model parameter matrices $\Theta = \{A, \ \dots, \ D_{22}\}$ are obtained solving non-linear least squares problem.

$$\min_{Z,\Theta} \quad \mathcal{J}(Z,\Theta)$$

- Using co-ordinate descent algorithm:
  1. **Iterate for** $n = 1, \dots$
     1.1 $\Theta^n \leftarrow \mathrm{argmin}_{\Theta} \quad \mathcal{J}(Z^{n-1},\Theta)$
     1.2 $Z^n \leftarrow \mathrm{argmin}_{Z} \quad \mathcal{J}(Z,\Theta^n)$
  2. **Until** $\|Z^n - Z^{n-1}\| \leq \epsilon$ or $n = n_{\max}$

Identification of LPV-LFR models

- Once $\hat{x}(k)$ is estimated, latent variable $z(k)$ and model parameter matrices $\Theta = \{A, \ \ldots, \ D_{22}\}$ are obtained solving non-linear least squares problem.

$$\min_{Z,\Theta} \quad \mathcal{J}(Z, \Theta)$$

- Using co-ordinate descent algorithm:
  1. **Iterate for** $n = 1, \ldots$
     1.1 $\Theta^n \leftarrow \operatorname{argmin}_\Theta \quad \mathcal{J}(Z^{n-1}, \Theta)$
     1.2 $Z^n \leftarrow \operatorname{argmin}_Z \quad \mathcal{J}(Z, \Theta^n)$
  2. **Until** $\|Z^n - Z^{n-1}\| \leq \epsilon$ or $n = n_{\max}$

Identification of LPV-LFR models: Affine case

- Special case: affine dependency on scheduling, i.e., $D_{11} = 0$,
  - $A(p(k)) = A_0 + A_1 p_1(k) + A_2 p_2(k) + \cdots + A_{n_p} p_{n_p}(k)$
- $\Delta(p(k)) = \varphi(p(k)) I$, i.e., the feedback block Δ has a diagonal. structure
- $\{A, \ \ldots, \ D_{22}\}$ are estimated solving ordinary least-squares followed by SVD.

Identification of LPV-LFR models: Affine case

- Special case: affine dependency on scheduling, i.e., $D_{11} = 0$,
  - $A(p(k)) = A_0 + A_1 p_1(k) + A_2 p_2(k) + \cdots + A_{n_p} p_{n_p}(k)$
- $\Delta(p(k)) = \varphi(p(k))I$, i.e., the feedback block Δ has a diagonal. structure
- $\{A, \ \ldots, \ D_{22}\}$ are estimated solving ordinary least-squares followed by SVD.

Numerical Example

Data generating system

$$\left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 1.073 \\ -0.1 & 0.7 & 0.816 & 1.075 \\ \hline 0.524 & -0.625 & -0.5 & 0.5 \\ 0.443 & 0.060 & 0.5 & 0.5 \end{array}\right],$$

$$w(k) = p(k)z(k).$$

- Training data $N = 400$ and SNR$= 21$ dB.
- Input u a white-noise process with uniform distribution $\mathcal{U}(-1, 1)$.
- The scheduling signal is $p(k) = 0.5\sin(k) + \delta(k)$, $\delta(k)$ is a random variable with uniform distribution $\mathcal{U}(-0.5, 0.5)$.

Numerical Example

- For state estimation: KCCA with RBF kernels $\bar{k}(p_i, p_j) = c \, \exp\left(-\frac{\|p_i - p_j\|^2}{\sigma^2}\right)$, $c = 2$ and $\sigma = 10.5$.
- Past and future window length is $d = 3$, LS-SVM parameters: $\gamma_p = \gamma_f = 500$
- The total computation time to construct the state sequence $\hat{x}(k)$ is 8.7 sec.
- The co-ordinate descent algorithm is run for $n = 250$ iterations, each iteration takes 0.3 sec.
- Affine case: the computation time to solve the least squares problem followed by SVD is 1.01 sec.

Numerical Example

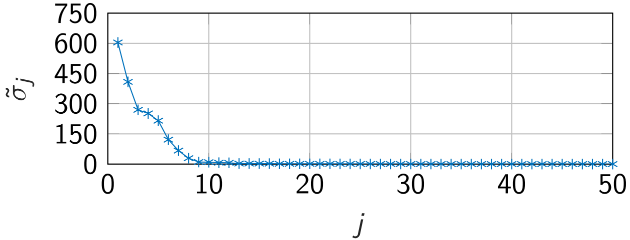


Figure: Singular values of the SVD problem

	General case	Affine case
BFR	92.71 %	96.05 %
VAF	99.74 %	99.84 %

Identification of LPV-LFR models

- ✓ Direct identification of LPV models with Linear Fractional Representation
- ✓ Parametric identification of state-space LPV models with rational scheduling dependencies.
- ✓ Computationally efficient solution for estimating affine scheduling dependent models.
- Estimation of the feedback block.
- Investigate efficient algorithms for solving non-liner least squares in the second stage.

Thank You