OSE SEMINAR 2011

System identification by support vector regression

AMIR SHIRDEL

CENTER OF EXCELLENCE IN
OPTIMIZATION AND SYSTEMS ENGINEERING
AT ÅBO AKADEMI UNIVERSITY

ÅBO DECEMBER 8 2011

OSE OPTIMIZATION AND SYSTEMS ENGINEERING

Åbo Akademi University

Content

- System identification
- Switching system
- Identification of switching system
- Support vector regression
- Sparse optimization
- Sparse optimization and SVR
- Simple example

System Identification

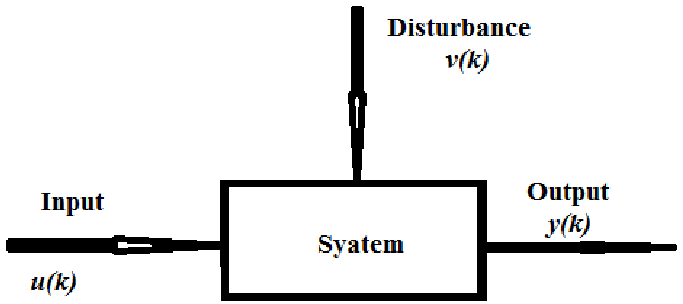
- System identification is the field of modeling dynamic systems from experimental to give a mathematical model which describes how the input,the output and the disturbances are related.

- For measured input-output data *{u(k),y(k)}*, the identification problem is :

Find model parameter θ such that $$y(k) = \varphi(k)^T \theta + v(k)$$

where

$$\varphi(k)^T = [y(k\text{-}r),\ldots,y(k\text{-}1),u(k),u(k\text{-}1),\ldots,u(k\text{-}r)]$$



Switching System

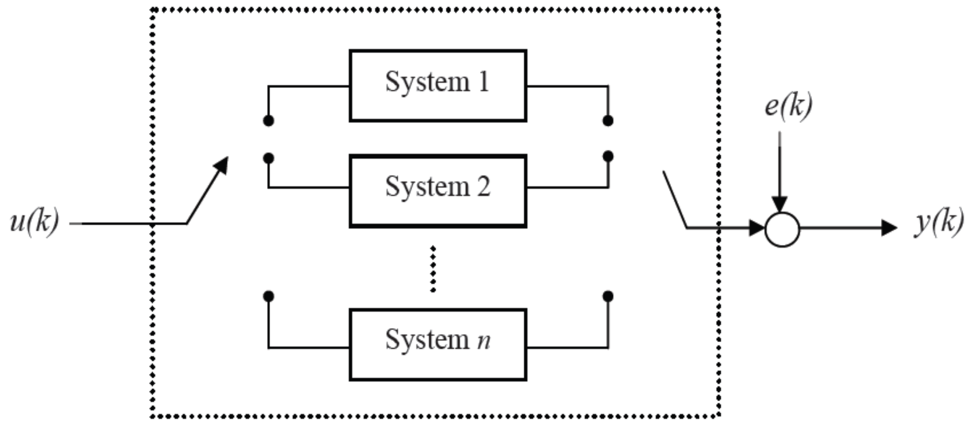
Described by two sets of states:

- discrete state: determines the active mode
- continuous state: evolves according to the dynamics of the active mode

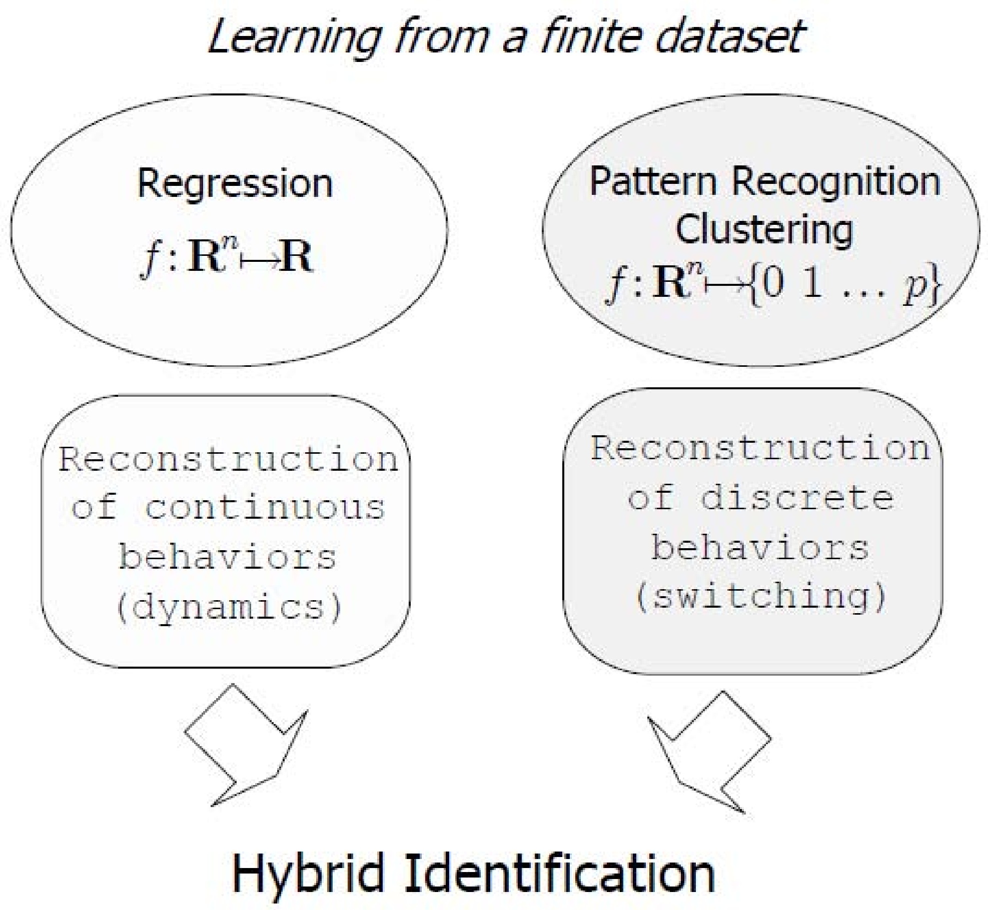
$$y(k) = \varphi(k)^T \theta_i + e(k) \quad , \quad i \in \{1, \ldots, n\}.$$

where

$$\varphi(k)^T = [y(k\text{-}r),\ldots,y(k\text{-}1),u(k),u(k\text{-}1),\ldots,u(k\text{-}r)]$$



Identification of switching system



Support vector regression

- Support vector regression tries to find parameter θ such that the maximum number of data points lie within the epsilon-wide insensitivity tube. [Vapnik, 1995].

- Identify the parameters θ in

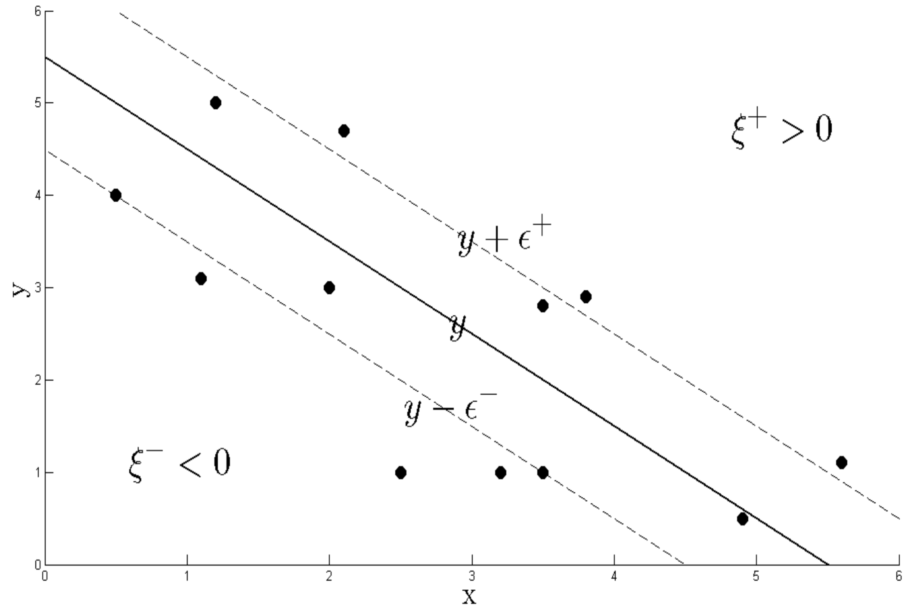
$$y_k = \varphi_k^T \theta$$

- By solving

$$\min \frac{1}{2}\|\theta\|^2 + C\sum_{k=1}^{N}(\xi_k + \xi_k^*)$$

subject to

$$\begin{cases} y_k - \varphi_k^T\theta \le \varepsilon + \xi_k \\ \varphi_k^T\theta - y_k \le \varepsilon + \xi_k^* \\ \xi_k, \xi_k^* \ge 0 \end{cases}$$



- The constant C > 0 determines the trade-off between regularization term and the amount up to which deviations larger than ε are tolerated.

Support vector regression

Introducing the associated Lagrangian and the saddle-point condition for θ gives the dual cost :

$$L := \frac{1}{2}\|\theta\|_2^2 + C\sum_{k=1}^{N}(\xi_k + \xi_k^*) - \sum_{k=1}^{N}(\eta_k\xi_k + \eta_k^*\xi_k^*) - \sum_{k=1}^{N}\alpha_k(\varepsilon + \xi_k - y_k + \varphi_k^T\theta) - \sum_{k=1}^{N}\alpha_k^*(\varepsilon + \xi_k^* + y_k - \varphi_k^T\theta)$$

$$\alpha_k^{(*)}, \eta_k^{(*)} \geq 0$$

It follows from the saddle point condition that the partial derivatives of *L* with respect to the primal variables(w, b, ξ_i, ξ_i^*) have to vanish for optimality.This gives the dual maximization problem:

maximize

$$-\frac{1}{2}\sum_{k,l=1}^{N}(\alpha_k - \alpha_k^*)(\alpha_l - \alpha_l^*)\varphi_k\varphi_l + \sum_{k=1}^{N} y_k(\alpha_k - \alpha_k^*) - \varepsilon\sum_{k=1}^{N}(\alpha_k + \alpha_k^*)$$

subject to

$$\sum_{k=1}^{N}(\alpha_k - \alpha_k^*) = 0, \alpha_k, \alpha_k^* \in [0, C]$$

The optimal parameters are given by:

$$\theta = \sum_{k=1}^{N}(\alpha_k - \alpha_k^*)\varphi_k$$

Sparse optimization

- Look for simple approximate solution of optimization problem.

$$\min_u \quad \|y - Au\|_2^2$$

such that

number of nonzero elements in $u \prec M$

- LASSO is a sparse optimization method for linear regression with L1-penalty at the cost of least squares fit.[Tibshirani 1996]

$$\min_u \quad \|y - Au\|_2^2 + \lambda \|u\|_1$$

- Shrinks some coefficients and sets others to zero.

Sparse optimization and SVR

- Dual problem in SVR can be written as a LASSO problem for sparse optimization:

$$\min_u = \|y - Au\|_2^2 + \varepsilon \|u\|_1$$

subject to $-C \leq u_k \leq C$,where

$$u_k = \alpha_k - \alpha_k^*$$

- Sparse solution implies that Lagrange multipliers vanish
  - many constraints are satisfied
  - solution corresponds to one mode of switching system(with high probability)

Sparse optimization and SVR

- **Proposed method for identification of switching systems**

Algorithm for finding one mode at a time using sparse optimization:

Step 1. Determine parameters of mode number *i* using SVR by applying sparse optimization with reweighting to dual problem

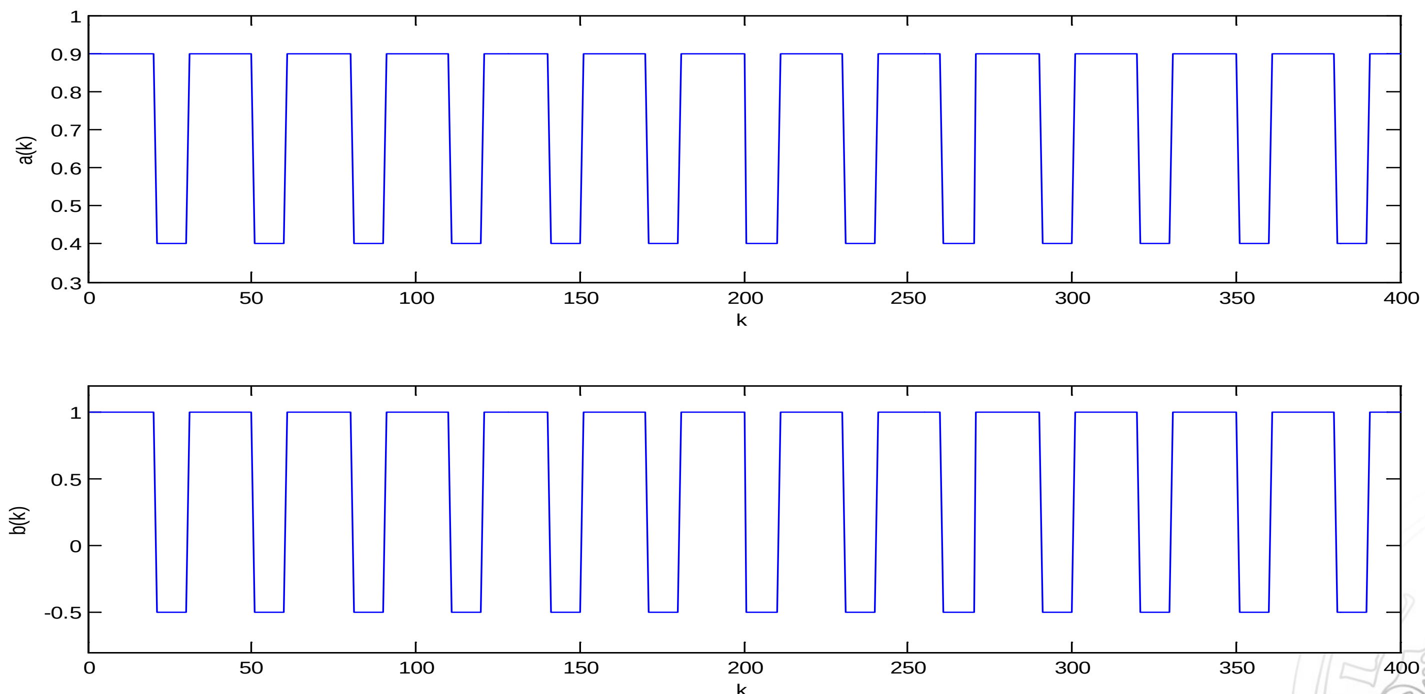
Step 2. Eliminate the data points associated with the mode found in step 1 and continue from step 1 until all modes have been found.

Example

Consider the simple system

$$y(k) = a(k)y(k\text{-}1) + b(k)u(k) + e(k);$$

$$where \begin{cases} a(k) = 0.9; \quad b(k) = 1; & for \;\; k = (30n + 1{:}30n + 20) \\ & \qquad\qquad n = 0,1,\ldots,N \\ a(k) = 0.4; \quad b(k) = \text{-}0.5 & for \;\; k = (30n + 21{:}30(n+1)) \end{cases}$$



Example

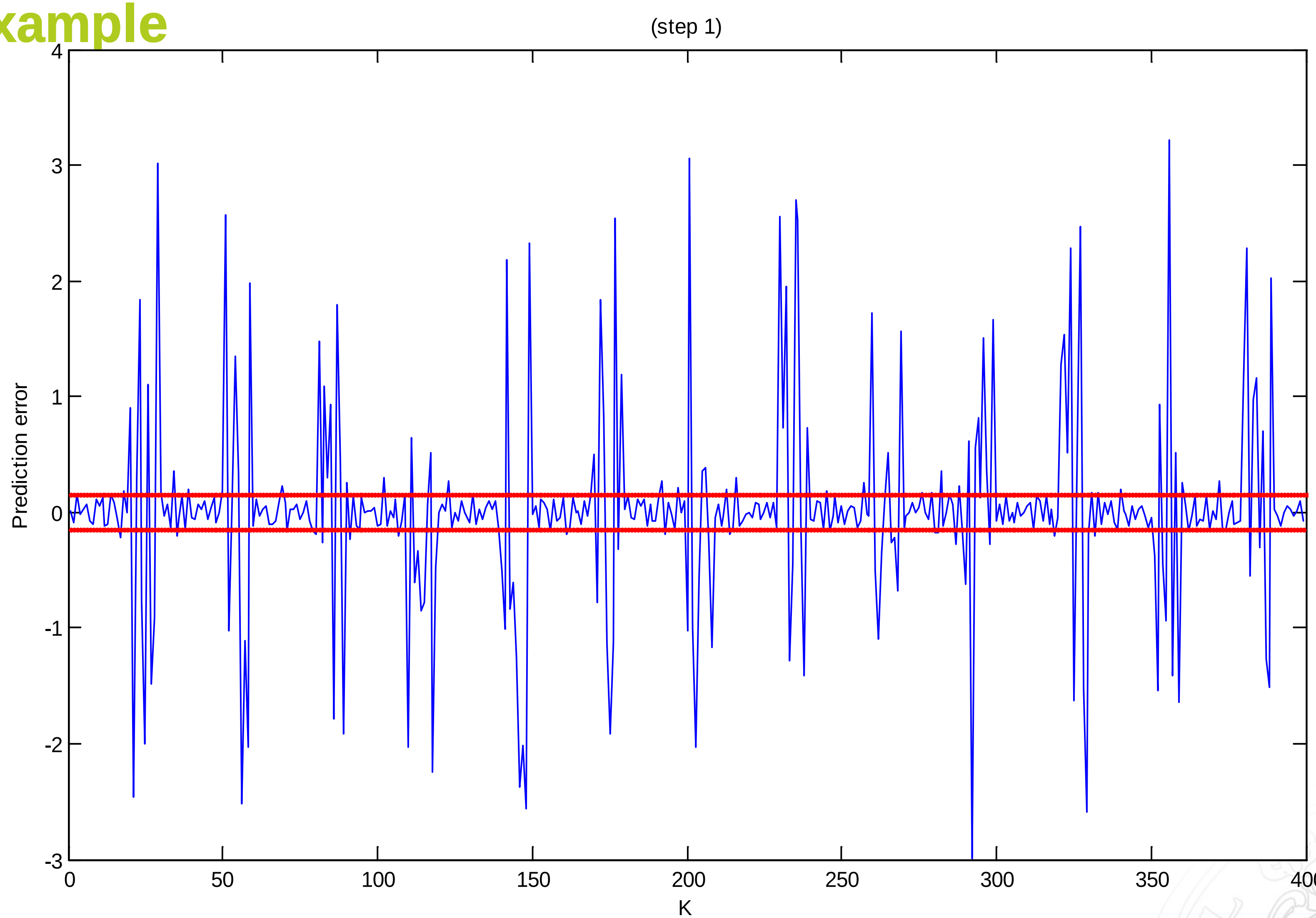
U(k) and e(k) are normally distributed with standard deviation 1 and 0.1.

SVR design parameters C = *100 and* ε = *0.15,*

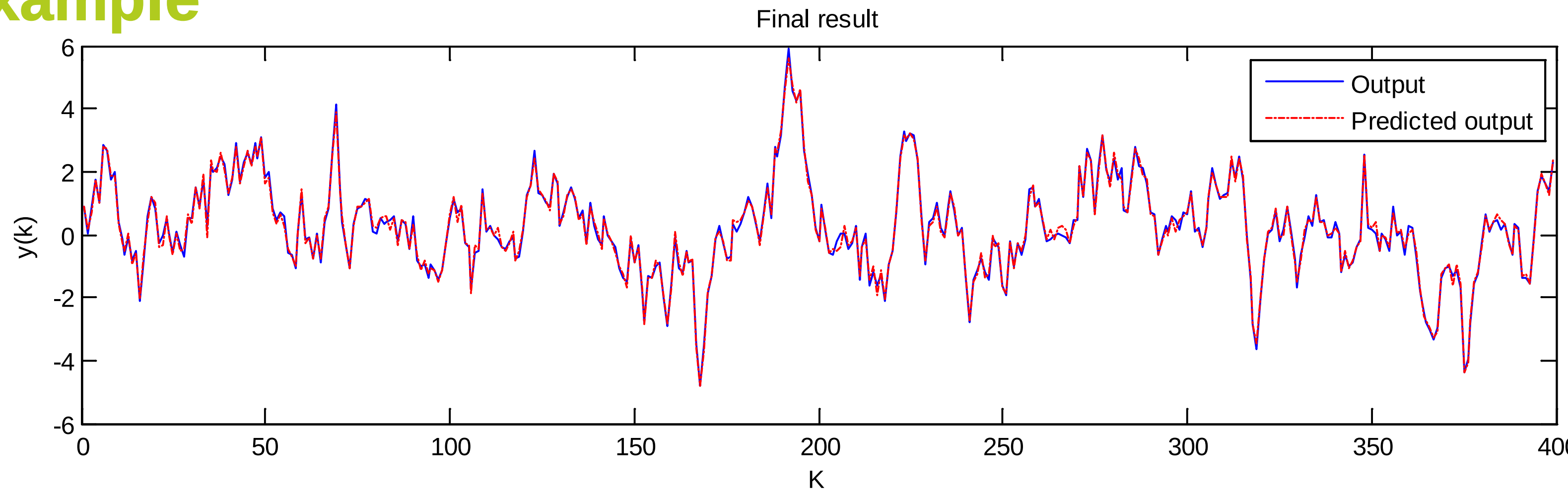
The following parameter estimates were obtained:

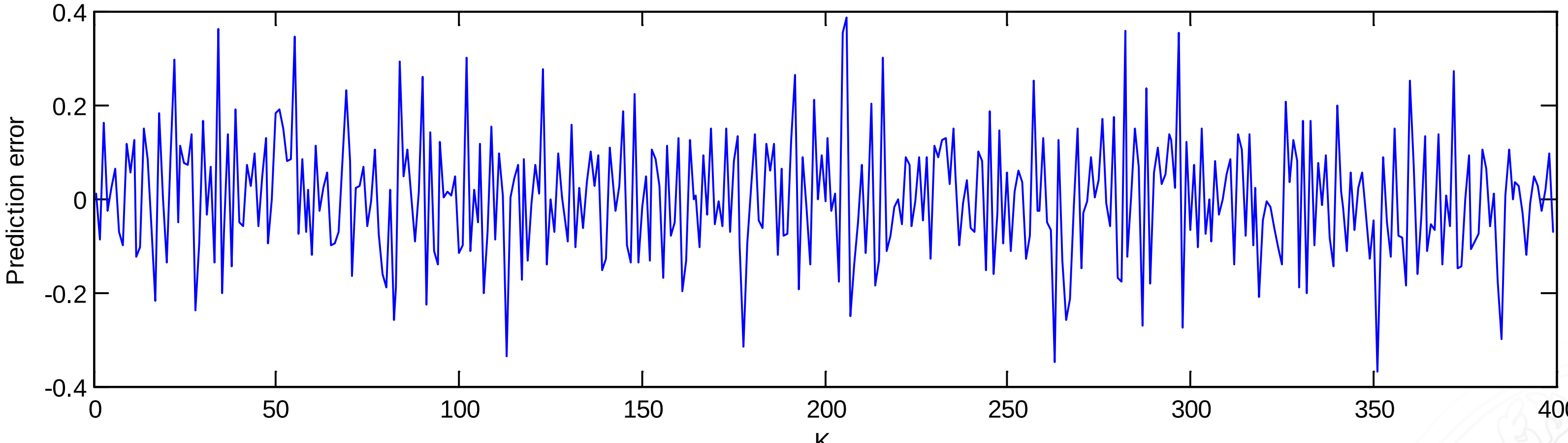
$$\hat{a}_1 = 0.8901; \qquad \hat{b}_1 = 1.0065$$
$$\hat{a}_2 = 0.3930; \qquad \hat{b}_2 = -0.4976$$

Example


(step 1)
Prediction error
K
4
3
2
1
0
-1
-2
-3
0
50
100
150
200
250
300
350
400

Example





RMSE=0.1283 Noise RMSE=0.1085

Discussion:

We have presented a method for identification of linear switching system by using SVR and LASSO sparse optimization.The method gives us acceptable results for finding the modes and the parameter values of each modes.

Future work:

We are trying to generalize kernel method for identification of nonlinear switching system.

Thank you for your attention!

Questions?