České vysoké učení technické v Praze
Fakulta dopravní

Czech Technical University in Prague
Faculty of Transportation Sciences

Ing. Evgenia Suzdaleva CSc.

Využití odhadu modelu směsi distribucí s datově-závislým modelem přepínání k odhadování spotřeby paliva silničního vozidla

Application of mixture estimation with data dependent switching model to estimation of a car fuel consumption

Summary

The presented habilitation lecture focuses on application of the recursive Bayesian mixture estimation theory to a car fuel consumption modeling. The specific algorithm is based on using a mixture of normal and categorical components with a data dependent dynamic model of switching.

The outline of the habilitation lecture includes the following items that will be presented: (i) motivation and briefly a state of the art in the discussed field; (ii) general problem formulation; (iii) a theory, specification of the problem formulation and a structural estimation algorithm; (iv) demonstration of results with real data.

Souhrn

Předkládaná habilitační přednáška se zaměřuje na použití teorie rekurzivního bayesovského odhadování modelu směsi distribucí k modelování spotřeby paliva automobilu. Specifický algoritmus je založen na použití směsi normálních a kategorických komponent s datově-závislým dynamickým modelem přepínání.

Obsah habilitační přednášky zahrnuje následující body, které budou prezentovány: (i) motivace a krátký současný stav v diskutované oblasti; (ii) obecná formulace problému; (iii) teorie, specifikace formulace problému a strukturální algoritmus odhadu; (iv) demonstrace výsledků s reálnými daty.

Klíčová slova
model směsi distribucí; datově-závislý dynamický model přepínání; aktivní komponenta; rekurzivní odhad; modelování smíšených dat

Keywords
mixture model; data dependent dynamic switching model; active component; recursive estimation; mixed data modeling

Contents

1 Introduction

1.1 Motivation

Reducing of fuel consumption and CO_2 emission is a significant problem concerning both economical and ecological parts of a society life. Nowadays the automotive industry invests a lot in development and support of various approaches in this area. The reasons are obvious: environment protection, the increasing price of oil, etc., see, for example, [1, 2]. With a gradual emergence of hybrid and electric vehicles [3] in the market a solution might seem to be found, as they promise significant fuel savings in exploitation. From an ecological viewpoint it also seems to be the most appropriate solution: with zero or minimal emissions they are suitable for low-emission zones established in some cities.

Nevertheless, conventional vehicles with combustion motors are still demanded in the market too. Firstly, purchase of a hybrid or an electric vehicle is still rather expensive (although in recent times the prices are reduced) that compensates fuel savings. However, other factors such as (i) the natural need of any new technology in refining and improving; (ii) a slowly appearing network of charging stations, especially out-of-town; (iii) unsuitability of domestic parking places for charging; (iv) significant environmental pollution during production and disposal of electric vehicles, etc., realistically predispose to exploitation of conventional vehicles too.

It leads to the fact that research in the area of eco-driving and reducing the fuel consumption are still highly desired and call for novel solutions. Obviously investigations in the field of eco-driving include the optimal control design as one of its main issues. However, a task of the fuel consumption modeling and its predicting (estimating) plays the important role in constructing the optimal eco-driving strategy. Moreover, it is a research problem, which joins conventional and hybrid and electric vehicles. Modeling an optimal eco-driving strategy is a task desired for all of them since: (i) conventional vehicles need it to reduce fuel consumption and emissions; (ii) hybrid vehicles should be driven optimally not to lose a benefit of the use of hybrid powertrain; (iii) electric vehicles need to model a travel range before recharging.

1.2 State of the art in eco-driving

Relevance of the discussed topic is confirmed of a series of studies in this area. For example, the work [4] investigates which driving characteristics (speed profile, gear changing, etc.) have the main effect on emissions and the fuel consumption. The paper [5] evaluates the long-term impact of an eco-driving training course by monitoring the driving behavior and the fuel consumption for several months before and after the course.

Studies dealing with modeling the fuel consumption can be found, e.g., in papers [6, 7, 2, 8, 9]. They are mostly devoted to algorithms based on a physical description of the fuel consumption, which takes into account surrounding traffic conditions.

Here, in this habilitation lecture, modeling the fuel consumption is taken in the Bayesian context with the use of data measured on a driven vehicle. The considered data-driven approach to the task is based on Bayesian recursive estimation algorithms [10, 11, 12, 13] of mixture models, whose components are assumed to describe individual regimes of a driving style. An attempt to construct a data-based description of a driven car as a system under Bayesian methodology has been already considered in [14]. However, it was not with the use of mixture models. Here, the issue is considered through a mixture of normal and categorical components.

1.3 State of the art in mixture estimation

Mixture based approaches are applied in various application fields, see, for instance, [15, 16, 17, 18, 19]. Papers dealing with the mixture estimation are mostly based on

- the iterative expectation-maximization (EM) algorithm [20], see, e.g., [21, 22, 23, 24, 25];
- the Variational Bayes (VB) approach [26, 27, 28];
- Markov Chain Monte Carlo (MCMC) techniques, e.g., [29, 30, 31].

Recursive Bayesian algorithms avoiding numerical computations and aimed at the real-time performance can be found in [11, 12]. With the application of [10, 32], they create a basis for a solution used in the presented habilitation lecture. The discussed solution provides an extension of the dynamic switching model proposed in [13] used for a specific case of mixed components.

2 Driven car as stochastic system

A driven car can be considered as a stochastic system, which in discrete time instants $t \in \{1, \ldots, T\}$ generates the following variables of a continuous nature:

- $y_{1;t}$ is the instantaneous fuel consumption [μl],
- $y_{2;t}$ is a vehicle speed [km/h],
- $y_{3;t}$ is a pressing the gas pedal [%],
- $y_{4;t}$ is the engine torque [Nm]
- and $y_{5;t}$ is the engine speed [rpm],

which are grouped in the vector $y_t = [y_{1;t}, y_{2;t}, y_{3;t}, y_{4;t}, y_{5;t}]'$. The discrete variable $z_t \in \{1, 2, \ldots, m_z\} \equiv z^*$ observed on the system is selection of gear during driving with the six-speed gearbox. It is assumed that the observed system works in several regimes different by a driving style, which means that the system is multimodal.

The presented habilitation lecture focuses on modeling and estimating at each time instant the first variable of the vector y_t – the instantaneous fuel consumption $y_{1;t}$. The considered problem can be generally formulated as follows.

2.1 General problem formulation

The considered problem includes the following subtasks:

- construct a specific data-based model of the fuel consumption;
- recursively estimate its parameters based on available data;
- verify the model using the real measurements.

Here, the problem is solved using the algorithm of estimation of mixtures of normal and categorical components with the dynamic categorical data-dependent model of switching. The solution is proposed in the submitted paper [33], co-authored by the habilitation candidate. Main ideas of the specific theory and the structural algorithm are given in the section below.

3 Mixtures with data dependent switching model

3.1 Specific models

Let's describe the observed system by a mixture of m_c components, where each i-th component, $\forall i \in \{1, \ldots, m_c\} \equiv c^*$, has the form of the joint pdf, which is decomposed according to the chain rule, see, e.g., [10] in the following way

$$f\left(y_t, z_t = l | y\left(t-1\right), z\left(t-1\right), \Theta, \beta, c_t = i\right)$$

$$= f\left(y_t | \underbrace{y(t-1), z(t)}_{\psi_t^y}, \Theta, c_t = i\right) f\left(z_t = l | \underbrace{z(t-1)}_{\psi_{t-1}^z}, \beta, c_t = i\right), \quad (1)$$

where $l \in z^*$, $i \in c^*$, and

- $f(y_t|\psi_t^y, \Theta, c_t = i) \equiv f(y_t|\psi_t^y, \Theta_i)$ is the i-th linear regression model
$$\mathcal{N}_y(\psi_t^y \theta_i, r_i) \quad (2)$$
with normal noise and with the regression vector
$$\psi_t^y = [z_t', y_{t-1}', z_{t-1}', \ldots, y_{t-m_y}', z_{t-m_y}', 1]' \quad (3)$$
of the memory length m_y and with parameters $\Theta_i = \{\theta_i, r_i\}$ existing for each $i \in c^*$, where θ_i is the collection of regression coefficients of the i-th component and r_i is the constant covariance matrix of the noise, and $\{\Theta_i\}_{i=1}^{m_c} \equiv \Theta$,
- $f(z_t = l|\psi_{t-1}^z = q, \beta, c_t = i) \equiv f(z_t = l|\psi_{t-1}^z = q, \beta_i)$, assumed to be independent on y_t, is the transition table with the regression vector $\psi_{t-1}^z = [z_{t-1}', \ldots, z_{t-m_z}']'$ of the memory length m_z, and

with the matrix parameter β_i, where $\{\beta_i\}_{i=1}^{m_c} \equiv \beta$, existing for each $i \in c^*$. The transition table has the form

$$f\left(z_t = l|\psi_{t-1}^z = q, \beta_i\right) \equiv \tag{4}$$

	$z_t = 1$	$z_t = 2$	$\cdots$	$z_t = m_z$
ψ_{t-1}^z denoted by 1	$(\beta_{1\|1})_i$	$(\beta_{2\|1})_i$	$\cdots$	$(\beta_{m_z\|1})_i$
ψ_{t-1}^z denoted by 2	$(\beta_{1\|2})_i$	$\cdots$	$\cdots$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
ψ_{t-1}^z denoted by m_ψ	$(\beta_{1\|m_\psi})_i$	$\cdots$	$\cdots$	$(\beta_{m_z\|m_\psi})_i$

where $q \in \{1, 2, \ldots, m_\psi\} \equiv \psi^*$, which is a set of possible configurations of the regression vector. Here β_i is a matrix of entries $(\beta_{l|q})_i$ with $l \in z^*$, $q \in \psi^*$ and $i \in c^*$, where $(\beta_{l|q})_i$ is the probability of $z_t = l$ conditioned by $\psi_{t-1}^z = q$ existing for the component i, and it holds

$$\left(\beta_{l|q}\right)_i \geq 0, \sum_{l \in z^*} \left(\beta_{l|q}\right)_i = 1, \ \forall l \in z^*, \forall q \in \psi^*, \ \forall i \in c^*, \tag{5}$$

- it is assumed that neither the y-part of the pdf (1) depends on β nor the z-part includes Θ.

Switching the active components (1) is represented as the unmeasured random process $c_t = \{1, 2, \ldots, m_c\} \equiv c^*$, which is called *the pointer* (see, e.g., [12, 11]). It means that c_t points to the active component at time t. Here switching the components is described by the model

$$f\left(c_t = i|c_{t-1} = j, \underbrace{z\,(t-1)}_{\varphi_{t-1}=k}, \alpha\right), \ i, j \in c^*, \ k \in \varphi^*, \tag{6}$$

where $\varphi_{t-1} = [z'_{t-1}, \ldots, z'_{t-m_\alpha}]'$ is the regression vector with the memory length m_α, and the set φ^* includes all its possible configurations. It has the form similarly to (4), however, existing for each k-th configuration of the regression vector φ_{t-1} with $k = \{1, \ldots, m_\varphi\} \equiv \varphi^*$, where m_φ is the whole number of such configurations, i.e.,

	$c_t = 1$	$c_t = 2$	$\cdots$	$c_t = m_c$
$c_{t-1} = 1$	$(\alpha_{1\|1})_k$	$(\alpha_{2\|1})_k$	$\cdots$	$(\alpha_{m_c\|1})_k$
$c_{t-1} = 2$	$(\alpha_{1\|2})_k$		$\cdots$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$c_{t-1} = m_c$	$(\alpha_{1\|m_c})_k$		$\cdots$	$(\alpha_{m_c\|m_c)_k}$

The matrix-form parameter $\alpha \equiv \{\alpha_k\}_{k=1}^{m_\varphi}$ contains entries $(\alpha_{i|j})_k$, which are the probabilities of the pointer $c_t = i$ under condition that the previous pointer $c_{t-1} = j$ and $\varphi_{t-1} = k$, with the assumption of the form (5) for corresponding probabilities.

3.2 Problem specification

The problem is specified as follows: based on available data, estimate

- the parameter α of the switching model;
- the value of the pointer c_t, indicating the active component at time t;
- parameters Θ of normal components generating data y_t;
- parameters β of categorical components producing data z_t.

3.3 Summarized theory

The solution is based on construction of the joint pdf of variables to be estimated and application of Bayes rule and inspired by [10, 11, 12]. With the data denoted by $\Delta_t = \{y_t, z_t\}$ and the available data collection up to the time instant t denoted by $\Delta(t) = \{\Delta_0, \Delta_1, \ldots, \Delta_t\}$, including prior data Δ_0, the joint pdf of variables to be estimated is derived with the models (1), (4) and (6) as follows:

$$\underbrace{f\left(\Theta, \beta, c_t = i, c_{t-1} = j, \alpha | \Delta\left(t\right)\right)}_{posterior\ pdf}$$

$$\propto \underbrace{f\left(y_t, z_t = l, \Theta, \beta, c_t = i, c_{t-1} = j, \alpha | \Delta\left(t-1\right)\right)}_{joint\ pdf}$$

$$= \underbrace{f\left(y_t | \psi_t^y, \Theta, c_t = i\right) f\left(\Theta | \Delta\left(t-1\right)\right)}_{for\ \Theta\ estimation}$$

$$\times \underbrace{f\left(z_t = l | \psi_{t-1}^z = q, \beta, c_t = i\right) f\left(\beta | \Delta\left(t-1\right)\right)}_{for\ \beta\ estimation}$$

$$\times \underbrace{f\left(c_t = i | c_{t-1} = j, \varphi_{t-1} = k, \alpha\right) f\left(\alpha | \Delta\left(t-1\right)\right)}_{for\ \alpha\ estimation}$$

$$\times \underbrace{f\left(c_{t-1}=j|\Delta\left(t-1\right)\right)}_{pointer\ prior\ pdf}, \tag{7}$$

where $i, j \in c^*$, $l \in z^*$, $q \in \psi^*$ and $k \in \varphi^*$. This joint pdf is marginalized firstly over all parameters, which provides the posterior pdf joint for c_t and c_{t-1}, i.e.,

$$\underbrace{f(c_t=i, c_{t-1}=j|\Delta(t))}_{posterior\ pdf\ of\ c_t\ and\ c_{t-1}}$$

$$\propto$$

$$\int_{\Theta^*}\int_{\beta^*}\int_{\alpha^*} \underbrace{f\left(y_t, z_t=l, \Theta, \beta, c_t=i, c_{t-1}=j, \alpha|\Delta\left(t-1\right)\right)}_{(7)} d\Theta d\beta d\alpha$$

$$= \int_{\Theta^*} f\left(y_t|\psi_t^y, \Theta, c_t=i\right) f\left(\Theta|\Delta\left(t-1\right)\right)\, d\Theta$$

$$\times \int_{\beta^*} f\left(z_t=l|\psi_{t-1}^z=q, \beta, c_t=i\right) f\left(\beta|\Delta\left(t-1\right)\right)\, d\beta$$

$$\times \int_{\alpha^*} f\left(c_t=i|c_{t-1}=j, \varphi_{t-1}=k, \alpha\right) f\left(\alpha|\Delta\left(t-1\right)\right)\, d\alpha$$

$$\times \underbrace{f\left(c_{t-1}=j|\Delta\left(t-1\right)\right)}_{pointer\ prior\ pdf}. \tag{8}$$

Here the first integral is approximated by substituting the current measurement y_t and the previous point estimates of the parameters denoted by $\hat{\theta}_{i;t-1}$ and $\hat{r}_{i;t-1}$ to each i-th component. It provides a proximity of the current output y_t to the i-th component. The point estimates of parameters are obtained using the conjugate prior Gauss-inverse-Wishart distribution for each normal component via the Bayes rule, which leads to recursive updating the initially chosen statistics $V_{i;t-1}$ and $k_{i;t-1}$ for each $i \in c^*$ of the appropriate dimensions according to [10, 12].

Similarly, the second integral represents the probability of the current measurement z_t conditioned by the actual ψ_{t-1}^z taken from the previous-time point estimate denoted here by $\hat{\beta}_{i;t-1}$ (i.e., from the table) for each i-th component. The point estimates of the parameters of each i-th categorical component are obtained via the Bayes rule using the conjugate prior Dirichlet pdf according to [11] with the recomputable statistics denoted by $\vartheta_{i;t-1}$.

The third integral is a computation of the point estimate of the parameter α using the previous time instant statistics (here denoted by

$\gamma_{k;t-1}$) of the switching model for the actual value of the regression vector $\varphi_{t-1} = k$ with $k \in \varphi^*$.

The prior pointer pdf $f\left(c_{t-1} = j | \Delta\left(t-1\right)\right)$ expresses the probability of each component activity at previous time instant $t-1$. Initially, it is chosen and then it is updated into the posterior pdf of the current pointer c_t by marginalization of the result (8) over the values of c_{t-1}

$$\underbrace{f(c_t = i|\Delta(t))}_{posterior\ pointer\ pdf} \propto \sum_{j \in c^*} \underbrace{f(c_t = i, c_{t-1} = j|\Delta(t))}_{(8)} = w_{i;t},\ i, j \in c^*, \quad (9)$$

which is denoted by $w_{i;t}$, and it is the i-th entry of the m_c-dimensional weighting vector w_t. At the time instant t, $w_{i;t}$ is the updated probability of the i-th component activity with respect to the currently measured data items y_t and z_t. The index of the maximum entry of the vector w_t denotes the point estimate of the pointer c_t, i.e., the component declared to be active at time t.

This briefly summarized theory leads to recursive updating the statistics for estimation of parameters Θ and β respectively as follows (based on [12])

$$\begin{aligned} V_{i;t} &= V_{i;t-1} + w_{i;t} \left[\begin{array}{c} y_t \\ \psi_t^y \end{array} \right] \left[\begin{array}{c} y_t \\ \psi_t^y \end{array} \right]', && (10) \\ \kappa_{i;t} &= \kappa_{i;t-1} + w_{i;t}, && (11) \\ (\vartheta_{l|q})_{i;t} &= (\vartheta_{l|q})_{i;t-1} + \delta\left(l, q;\ z_t, \psi_{t-1}^z\right) w_{i;t}, && (12) \end{aligned}$$

$\forall i \in c^*$ and $l \in z^*$, $q \in \psi^*$, and from which the point estimates $\hat{\theta}_{i;t}$ and $\hat{r}_{i;t}$ are recomputed according to

$$\hat{\theta}_{i;t} = V_1^{-1} V_y, \quad \hat{r}_{i;t} = \frac{V_{yy} - V_y' V_1^{-1} V_y}{\kappa_{i;t}} \quad (13)$$

with the help of partition

$$V_{i;t} = \left[\begin{array}{cc} V_{yy} & V_y' \\ V_y & V_1 \end{array} \right], \quad (14)$$

where V_{yy} is the square matrix of the dimension k_y of the vector y_t, V_y' is k_y-dimensional column vector and V_1 is scalar [10]. The point estimate $(\hat{\beta}_{l|q})_{i;t}$ is computed for each categorical part of the i-th component as follows [11]

$$(\hat{\beta}_{l|q})_{i;t} = \frac{(\vartheta_{l|q})_{i;t}}{\sum_{s=1}^{m_c} (\vartheta_{s|q})_{i;t}},\ \ l \in z^*,\ q \in \psi^*. \quad (15)$$

The posterior pdf $f(c_t = i, c_{t-1} = j|\Delta(t))$, obtained in (9), is joint for c_t and c_{t-1}. It is denoted by $W_{i,j;t}$ and used in the update of statistics for estimation of the parameter α

$$(\gamma_{i|j})_{k;t} = (\gamma_{i|j})_{k;t-1} + \delta\,(k;\,\varphi_{t-1})\,W_{i,j;t}, \quad \forall i, j \in c^*, \quad k \in \varphi^*, \qquad (16)$$

from which the point estimate of the parameter α at the time instant t is computed similarly to (15) for the actual k

$$(\hat{\alpha}_{i|j})_{k;t} = \frac{(\gamma_{i|j})_{k;t}}{(\sum_{s=1}^{m_c} \gamma_{s|j})_{k;t}}. \qquad (17)$$

The detailed derivations and explanations can be found in [33]. The structural algorithm is provided below.

3.4 Algorithm

Initial part (for t=1)

- Specify m_c components (1) with their normal and categorical parts.
- Specify the switching model (6).
- Set the initial statistics $V_{i;t}$, $\kappa_{i;t}$ and $\vartheta_{i;t}$ for each $i \in c^*$.
- Set the initial statistics of the switching model $\gamma_{k;t}$ $\forall k \in \varphi^*$.
- Using these initial statistics, compute the initial point estimates $\hat{\theta}_{i;t}$, $\hat{r}_{i;t}$, $\hat{\beta}_{i;t}$ and $\hat{\alpha}_{k;t}$ of all parameters and for all components.
- Set the initial weighting vector w_t.

On-line part (for t=2,3,...)

1. Measure new data y_t and z_t.
2. For each $i \in c^*$ substitute y_t and the point estimates $\hat{\theta}_{i;t-1}$ and $\hat{r}_{i;t-1}$ into each normal component. Construct the m_c-dimensional vector from obtained proximities from all components.
3. For each $i \in c^*$ take the probability $(\hat{\beta}_{l|q})_{i;t-1}$ for the current with $z_t = l$ and $\psi^z_{t-1} = q$. Construct the m_c-dimensional vector from results from all components.

4. According to (8), multiply entry-wise the resulted vectors from two previous steps, the prior weighting vector w_{t-1} and the point estimate matrix $\hat{\alpha}_{k;t-1}$ for the actual k.
5. The result of this entry-wise multiplication is the matrix with entries $W_{i,j;t}$. Normalize this matrix.
6. Perform the summation of the above normalized matrix over rows and obtain the vector with updated entries $w_{i;t}$ according to (9).
7. Update all statistics according to (10)–(12) and (16).
8. Recompute the point estimates of all parameters according to (13), (15) and (17) and use them for Step 1 of the on-line part of the algorithm.

4 Application to fuel consumption estimation

This section demonstrates the application of the above approach to estimating the fuel consumption, see Section 2.

4.1 Data

Real data measured on a vehicle during driving each 1 second are used. Their whole number is about 75000, but here for a better illustration results are shown for 1500 data items. According to Section 2, the five-dimensional vector y_t includes the variables

- $y_{1;t}$ is the instantaneous fuel consumption [μl],
- $y_{2;t}$ is a vehicle speed [km/h],
- $y_{3;t}$ is a pressing the gas pedal [%],
- $y_{4;t}$ is the engine torque [Nm]
- and $y_{5;t}$ is the engine speed [rpm].

The categorical variable z_t represents values of the gear chosen during driving. Originally, its measured values belong to the set

$$\{-1, 0, 1, 2, 3, 4, 5, 6\},$$

where 0 denotes the neutral gear and -1 is the reverse gear. To avoid the unbalance of frequencies of the gear values, their number is reduced. The reason is as follows. To obtain data, which cover as many working modes as possible, the measurements were taken in different traffic situations: from a relatively calm economic driving on the highway and a mixed driving on the first and second class roads till driving through several villages. The major part of the taken data represents driving with a relatively high speed out of city, which more corresponds to higher values of the gear. Because of this, the lower values of the gear (except the neutral gear) were observed only rarely, which leads to the mentioned unbalance. It means that the rare measurements could be perceived as outliers. To avoid this, it is advantageous to group the gear values so that to justify the frequencies. Here it is done so that $\{-1, 0\} \rightarrow 1$, $\{1, 2, 3\} \rightarrow 2$, $4 \rightarrow 3$, $5 \rightarrow 4$, $6 \rightarrow 5$. In this way the set of possible values of the discrete variable z_t is $z^* \equiv \{1, 2, 3, 4, 5\}$. Now it is relatively balanced.

4.2 Model construction

The fuel consumption model is constructed according to Section 3.1 using (1), where the regression vector of the normal components (2) has the form $\psi_t^y = [y(t-1),\ z_{t-1},\ 1]'$. Notice that the regression vector ψ_t^y includes delayed values of the gear selection z_t.

The categorical parts (4) of components (1) with the regression vector $\psi_{t-1}^z = [z_{t-1},\ s_{t-1}]'$ contains the already mentioned gear selection z_t and the discretized speed $y_{2;t}$ denoted here by $s_t \in \{1, 2, 3, 4\}$.

The pointer c_t (see the switching model (6)) is assumed to classify data among various driving styles (for instance, eco-driving, sporty driving and mix of them, or driving on highway, in city and on out-of-town roads, etc.). However, c_t is unmeasured and has to be estimated from the data. Here, the experiment with $m_c = 3$ is demonstrated. Nevertheless, other number of components can be chosen. The regression vector of the switching model (6) involves the past pointer c_{t-1} and the discretized instantaneous fuel consumption $y_{1;t-1}$ denoted by $f_{t-1} \in \{1, 2, 3\}$.

4.3 Initialization

150 data items are used to obtain initial estimates of all parameters according to the initial part of Algorithm 3.4. Then the online part of the algorithm is tested on 1500 data.

4.4 Results

Maximal entries of the weighting vector actualized during the online estimation define the currently active component (here a driving style). They are stored as the point estimates of the pointer. There are no measured values of the driving style, so the pointer estimates cannot be compared with real data. However, together with the parameter point estimates they can be used for 0-step-ahead prediction (estimation) of the fuel consumption by their substituting into corresponding components. These results are shown in Figure 1, which compares real values of all entries of the vector y_t with results of these predictions from normal components. The first subplot shows results for the fuel consumption.

Figure 2 demonstrates prediction of the gear selection from categorical components.

In the case of the multivariate variable such as y_t it is difficult to show clusters detected in the data space. That's why Figure 3 shows them as several pairs plotted against each other. In all the figures dots always visualize clusters of original data, while a rest of symbols correspond to the predicted values of variables.

4.5 Discussion

The demonstrated results of application of the algorithm to real data confirm its functionality. Several comments should be made on the application of the algorithm:

- The initialization of the algorithm plays a significant role in the successful estimation. Here the smaller part (150) of prior data was enough for a good start of the algorithm.
- Another important aspect is the suitability of the used model. Here, the first order of regression components seems to be sufficiently suitable for describing the data.

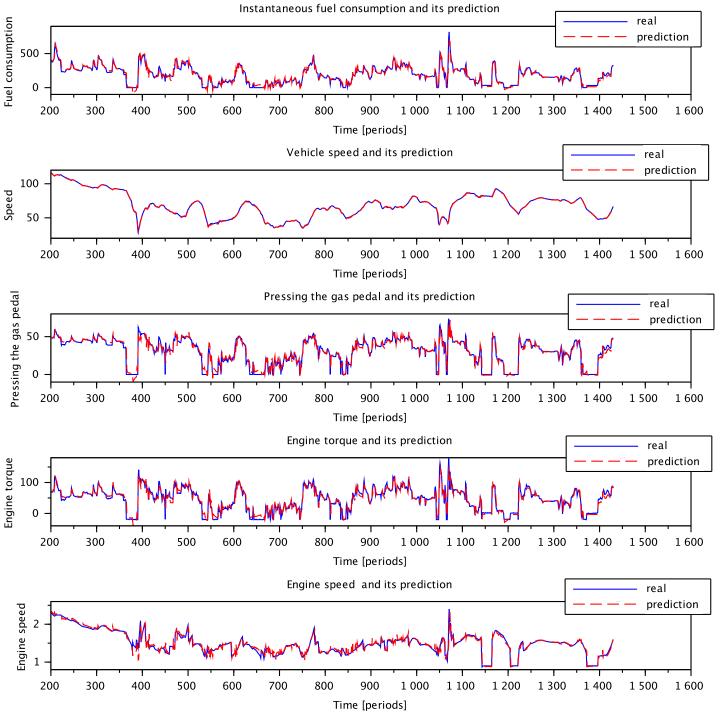

Figure 1: Predictions of continuous outputs
Notice that the red predicted values obtained by substituting the point estimates into components repeat the course of the blue real values. The measuring period is 1 second. It explains why the vehicle speed shown in the second subplot changes rather slowly unlike the rest of variables.

- The algorithm is not limited by the presented specific case of mixed components. Other pdfs with reproducible statistics can be used.

5 Conclusion

The main theoretical contribution of the presented habilitation lecture is the demonstration that a systematic approach to the Bayesian mixture estimation theory applied to regression models in [13] and state-space

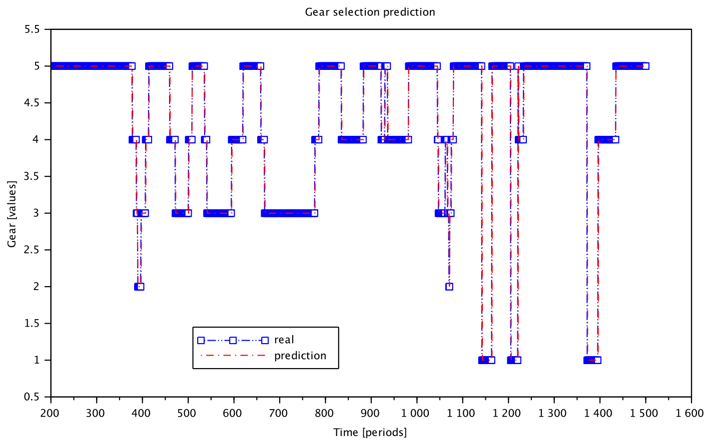

Figure 2: Prediction of gear selection
Each time when the red dotted line does not correspond to the blue square line means a wrong estimate. Here the whole course of the estimation is shown, which has only 38 wrong estimates for the whole number 1500 data.

models in [34] can be enriched by a mixture of normal and categorical components, saving such its main advantages as: (i) recursiveness that enables a real-time performance; (ii) one-pass elaboration of the data sample and (iii) orientation at explicit solutions with exploitation of numerical procedures only in those parts which cannot be computed analytically. A non-trivial model of dynamic switching extended by categorical measurements is considered.

From the practical point of view the habilitation lecture has been devoted to the problem of estimation of the fuel consumption using the mentioned theory. The obtained results look promising. The considered data-based model of the fuel consumption can be taken as a part of more complex problems such as its multi-step prediction and the control design arisen in the eco-driving field.

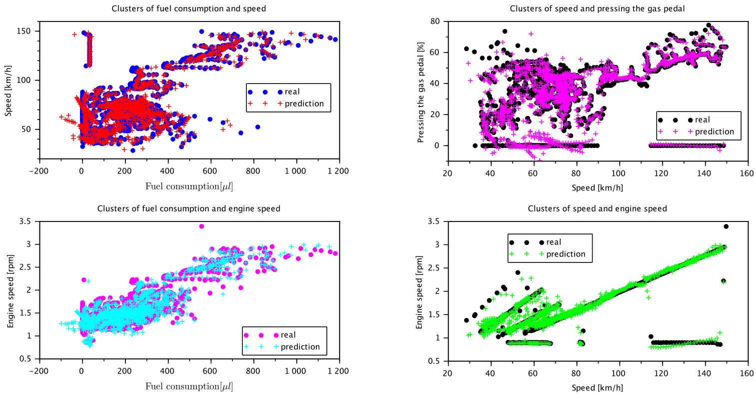

Figure 3: Clusters of continuous outputs
Two first figures compare clusters of the fuel consumption plotted against the vehicle and the engine speeds. The third and the fourth figures plot clusters of the vehicle speed against the pressing the gas pedal and the engine speed. Notice how the predicted clusters overlap the real values.

5.1 Future plans

In the discussed area there is still a series of open problems both from the theoretical and practical point of view. The following issues are expected to be especially fruitful as the further development of the mixture-based clustering and classification theory:

1. The estimation algorithms will be extended for different combinations of components and the switching model. Namely, the dynamic data-dependent switching model will be investigated in a combination with different types of components: regression and state-space models, categorical, uniform, exponential and other (for instance, general triangular) distributions. The components can be both static and dynamic. From the clustering viewpoint, the combination of dynamic switching model and static components is seen as the most promising.

2. Multi-step-ahead mixture prediction within the considered context is a separate task planned to be solved. It is expected that

the dynamic mixture prediction algorithm will be a significant contribution in the field of classification-related problems.

3. Extension of the switching model up to several delayed values in the regression vector is a further planned task.

The mentioned issues are expected to be solved during the project GAČR 15-03564S “Clustering and Classification Using Recursive Mixture Estimation” led by the habilitation candidate. The project began in 2015 and will be solved until 2017.

As regards practical plans of the work, a specific application to real problems in the transportation domain will be considered. There are also two fundamental areas: prediction (for the prediction itself or for the control task) and classification (for modeling multimodal systems or the data classification itself). Obviously, every concrete application situation requires specific settings, which mostly means tailoring an existing approach or even a new direction in the theory.

Last but not least, it is also necessary to mention the pedagogical meaning of the presented lecture and the whole habilitation thesis. The systematic development of the theory is important from this point of view. Students obviously will not derive the basic algorithm, but after understanding the theory they will be able to further improve it, test and expand into new models. Thus they may be directly involved in the highly complex scientific research in different levels.

6 References

[1] J.N. Barkenbus. Eco-driving: An overlooked climate change initiative. *Energy Policy*, 38(2), 2010, pages 762–769.

[2] M. Sivak, B.Schoettle. Eco-driving: Strategic, tactical, and operational decisions of the driver that influence vehicle fuel economy. *Transport Policy*, 22, 2012, pages 96–99.

[3] G. Pistoia. *Electric and Hybrid Vehicles. Power Sources, Models, Sustainability, Infrastructure and the Market.* Elsevier, 2010, ISBN: 978-0-444-53565-8.

[4] E. Ericsson. Independent driving pattern factors and their influence on fuel-use and exhaust emission factors. *Transportation Research Part D: Transport and Environment.* 6(5), 2001, pages 325–345.

[5] B.Beusen, S. Broekx, T. Denys, C. Beckx, B. Degraeuwe, M. Gijsbers, K. Scheepers, L. Govaerts, R.Torfs, L.Int Panis. Using onboard logging devices to study the longer-term impact of an eco-driving course. *Transportation Research Part D: Transport and Environment.* 14(7), 2009, pages 514–520.

[6] Y. Saboohi, H. Farzaneh. Model for developing an eco-driving strategy of a passenger vehicle based on the least fuel consumption. *Applied Energy*, volume 86, issue 10, October 2009, pages 1925–1932.

[7] C. Raubitschek, N. Schütze, E. Kozlov, and B. Bäker. Predictive Driving Strategies under Urban Conditions for Reducing Fuel Consumption based on Vehicle Environment Information. *Proceedings of IEEE Forum on Integrated and Sustainable Transportation Systems*, Vienna, Austria, June 29–July 1, 2011, pages 13–19.

[8] M.Barth and K. Boriboonsomsin. Energy and emissions impacts of a freeway-based dynamic eco-driving system. *Transportation Research Part D: Transport and Environment* 14 (6), 2009, pages 400–410.

[9] I. Ben Dhaou. Fuel estimation model for ECO-driving and ECO-routing. *Proceedings of IEEE Intelligent Vehicles Symposium (IV).* Baden-Baden, Germany, 2011, June 5–9, pages 37–42.

[10] V. Peterka. "Bayesian system identification," in *Trends and Progress in System Identification*, P. Eykhoff, Ed. Oxford: Pergamon Press, 1981, pp. 239–304.

[11] M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař. *Optimized Bayesian Dynamic Advising: Theory and Algorithms.* London, Springer, 2006.

[12] M. Kárný, J. Kadlec, E.L. Sutanto. Quasi-Bayes estimation applied to normal mixture. In *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*, Eds: J. Rojíček, M. Valečková, M. Kárný, K. Warwick, pp. 77–82, CMP '98 /3./, Prague, CZ, 07.09.1998–09.09.1998.

[13] I. Nagy, E. Suzdaleva, M. Kárný and T. Mlynářová. Bayesian estimation of dynamic finite mixtures. *Int. Journal of Adaptive Control and Signal Processing.* 25(9), 2011, pages 765–787.

[14] E. Suzdaleva and I. Nagy. Data-based speed-limit-respecting eco-driving system. *Transportation Research Part C*, 2014, vol.44, pages 253–264.

[15] Byung-Jung Park, Yunlong Zhang, Dominique Lord. Bayesian mixture modeling approach to account for heterogeneity in speed data. *Transportation Research Part B: Methodological.* 44 (5), 2010, pages 662–673.

[16] Yajie Zou, Yunlong Zhang, Dominique Lord. Analyzing different functional forms of the varying weight parameter for finite mixture of negative binomial regression models. *Analytic Methods in Accident Research.* 1 (0), 2014, pages 39–52.

[17] Yu, Jie. A nonlinear kernel Gaussian mixture model based inferential monitoring approach for fault detection and diagnosis of chemical processes. *Chemical Engineering Science.* 68 (1), 2012, pages 506–519.

[18] Yu, Jie. A particle filter driven dynamic Gaussian mixture model approach for complex process monitoring and fault diagnosis. *Journal of Process Control.* 22(4), 2012, pages 778–788.

[19] Yu, Jianbo. Fault detection using principal components-based Gaussian mixture model for semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing.* 24(3), 2011, pages 432–444.

[20] M. R. Gupta and Y. Chen. *Theory and Use of the EM Method. (Foundations and Trends(r) in Signal Processing)*. 2011, Now Publishers Inc.

[21] Boldea O, Magnus JR. Maximum Likelihood Estimation of the Multivariate Normal Mixture Model. *Journal Of The American Statistical Association*. 2009, 104(488), pages 1539–1549.

[22] Cuesta-Albertos JA, Matran C, Mayo-Iscar A. Robust estimation in the normal mixture model based on robust clustering. *Journal Of The Royal Statistical Society Series B-Statistical Methodology*. 2008, 70(Part 4), pages 779–802.

[23] Wang H, Luo B, Zhang Q, Wei S. Estimation for the number of components in a mixture model using stepwise split-and-merge EM algorithm. *Pattern Recognition Letters*. 2004, 25(16), pages 1799–1809.

[24] Hong Zeng and Yiu-ming Cheung. Learning a mixture model for clustering with the completed likelihood minimum message length criterion. *Pattern Recognition*. 47(5), 2014, pages 2011–2030.

[25] S.K. Ng and G.J. McLachlan. Mixture models for clustering multilevel growth trajectories. *Computational Statistics & Data Analysis*. 71(0), 2014, pages 43–51.

[26] McGrory, C. A. and Titterington, D. M. Variational Bayesian analysis for hidden Markov models. *Australian & New Zealand Journal of Statistics*. 51, 2009, pages 227âĂŞ244.

[27] Šmídl, V. and Quinn, A. *The Variational Bayes Method in Signal Processing*. Springer, 2005.

[28] Z. Ghahramani, G. E. Hinton, Variational learning for switching state-space models, Neural Computation 12 (4) (2000) 831–864.

[29] S. Frühwirth-Schnatter, Finite Mixture and Markov Switching Models, 2nd Edition, Springer New York, 2006.

[30] S. Früwirth-Schnatter, Fully bayesian analysis of switching gaussian state space models, Annals of the Institute of Statistical Mathematics 53 (1) (2001) 31–49.

[31] R. Chen, J. S. Liu, Mixture kalman filters, J. R. Statist. Soc. B 62 (2000) 493–508.

[32] D. M. Titterington, A.F.M. Smith and U.E. Makov, Statistical analysis of finite mixture distributions. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley, 1985, ISBN 9780471907633.

[33] E. Suzdaleva and I. Nagy. Recursive Mixture Estimation with Data Dependent Dynamic Model of Switching. *Applied Mathematical Modelling.* 2015, submitted.

[34] I. Nagy and E. Suzdaleva. Mixture estimation with state-space components and Markov model of switching. *Applied Mathematical Modelling.* 37(24), 2013, pages 9970–9984.

7 Ing. Evgenia Suzdaleva CSc.

Born: 10.9.1977 E-mail: suzdaleva@fd.cvut.cz

Work experience

10/2014 – present CTU in Prague, Faculty of Transportation Sciences, Department of Applied Mathematics, *assistant professor*

6/2004 – present The Institute of Information Theory and Automation of the CAS, *researcher*

9/2001 – 2/2003 Department of System Analysis and Operations Research, Siberian State Aerospace University, Krasnoyarsk, Russia, *researcher and lecturer*

11/2000 – 6/2001 Siberian Institute of Business, Control and Psychology, Krasnoyarsk, Russia, *exercises*

8/1999 – 6/2002 State Research Institute of Control Systems, Wave Processes and Technology, Krasnoyarsk, Russia, *researcher*

Education

CSc. 2002 Department of System Analysis and Operations Research, Siberian State Aerospace University, Krasnoyarsk, Russia, *subject:* System analysis, control and information processing

Ing. 2000 Department of System Analysis and Operations Research, Siberian State Academy, Krasnoyarsk, *subject:* System analysis

Research focus

Recursive Bayesian algorithms in data analysis; modeling, estimation and control of dynamic stochastic systems; systems with mixed data

Recent grants

GAČR 15-03564S “Clustering and Classification Using Recursive Mixture Estimation”, 2015 - 2017, *project leader*

TAČR TA01030123 “Optimization of the ecology of driving based on the currently measured data”, 2011–2013, *project team member*

GAČR 201/06/P434 “Urban Traffic Feedback Control”, 2006–2008, *project leader*

MDČR 1F43A/003/120, “Transportation control in the centers of historical cities”, 2004–2005, *project team member*

Selected teaching experience

lectures Stochastic systems

exercises Stochastic systems, probability

students Bc.(3 defended, 3 current), PhD (1 current – supervisor-expert)

Selected recent publications (whole number 63)

[1] E. Suzdaleva, I.Nagy. Data-based speed-limit-respecting eco-driving system. *Transportation Research Part C*, vol.44 (2014), p. 253-264.

[2]. I. Nagy, E. Suzdaleva. Mixture estimation with state-space components and Markov model of switching. *Applied Mathematical Modelling*, vol.37, 24 (2013), p. 9970-9984.

[3]. E. Suzdaleva, I. Nagy. Online soft sensor for hybrid systems with mixed continuous and discrete measurements. *Computers and Chemical Engineering*, vol.36, 10 (2012), p. 294-300.

[4]. E. Suzdaleva, I. Nagy. Recursive state estimation for hybrid systems. *Applied Mathematical Modelling*, vol.36, 4 (2012), p. 1347-1358.

[5]. I. Nagy, E. Suzdaleva, M. Kárný. Bayesian estimation of mixtures with dynamic transitions and known component parameters. *Kybernetika*, vol.47, 4 (2011), p. 572-594.

[6]. I. Nagy, E. Suzdaleva, M. Kárný, T. Mlynářová. Bayesian estimation of dynamic finite mixtures. *International Journal of Adaptive Control and Signal Processing*, vol.25, 9 (2011), p. 765-787

[7]. E. Suzdaleva. Prior knowledge processing for initial state of Kalman filter. *International Journal of Adaptive Control and Signal Processing*, vol.24, 3 (2010), p. 188-202.

[8]. E. Suzdaleva, I. Nagy, L. Pavelková Lenka, T. Mlynářová. Double Optimization of Fuel Consumption and Speed Tracking. In *Proceedings of the 11th IFAC International Workshop on Adaptation and Learning in Control and Signal Processing*, p. 305-310, Caen, FR, July 2013.

[9]. I. Nagy, E. Suzdaleva, T. Mlynářová. Optimization of Driving Based on Currently Measured Data. In *Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems*, p. 2088-2093, The Hague, NL, October 2013.