Probabilistic Approach to Analysis of Death Traffic Accidents

Evgenia Suzdaleva* and Ivan Nagy**

*Department of Adaptive Systems
Institute of Information Theory and Automation
of the Academy of Sciences of the Czech Republic

**Faculty of Transportation Sciences, Czech Technical University, Prague

Outline of presentation

- Experiments with data from one of the roads of the Czech republic
- The available data sets contain measurable discrete-valued variables, which describe various conditions of the accident
- The result of the accident is a modeled variable (either with death or injury or only with material damage)
- We used the dynamic discrete state-space model
- Estimation of accidents results is proposed with the help of Bayesian filtering based on determined probabilities.

Motivation

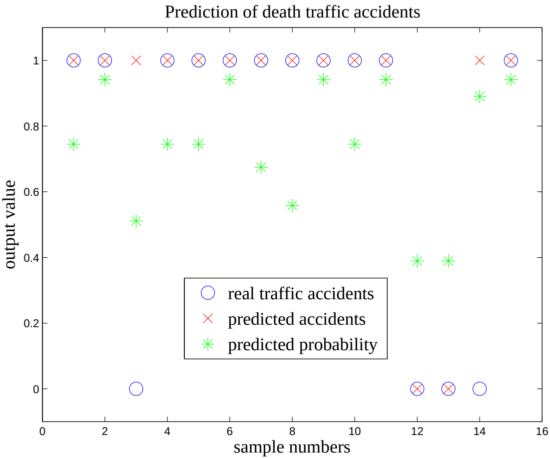
- Everyone knows that in our time of progress and high speed powerful cars the traffic accident statistics remains unfavorable and offensive: it is enough to see the everyday news.
- Modeling of traffic accidents and conditions under which they happened is important task to predict deathrate on roads under certain driving modes, speed, visibility etc.
- It gives a chance a driver to better regulate controllable variables such as speed and driving modes via advices of intelligent built-in car computers.

Introduction

- An accompanied problem – a lack of informative data and small amount of measurements. Surely, there are many other unmeasured variables, influencing the accidents.
- The data sets can be explored via various approaches to extract the information they can bring and understand relationships between variables. However, a choice of accurate model of the traffic accident is a difficult and ambiguous task, which requires an extra-use of expert knowledge.

Introduction

Previously the logistic regression was used. This is the most simple way of modeling, remained, however, in a static way.



Introduction

- Another way is to be aware of the unmeasured variables, or even to reduce the number of variables. The unmeasured and omitted variables, which are mostly periodical with the period one year, cause dynamics of the modeled variable.
- Thus, the dependence on some external variables can be substituted by considering the dynamic (state-space) model.

Data sets on the road II/114 in the Czech republic.

The modeled variable $x_t \in \{1, 0\}$ – result of the traffic accident, where 1 = accident with material damage, 0 = accident caused death or injury.

The measurements D_t contain the following information:

- $D_{1;t}$ is a daytime, where 1 – day, 2 – dawn, dusk, 3 – night;
- $D_{2;t}$ is visibility, where 1 – clear weather, 2 – fog, 3 – rain, 4 – snow;
- $D_{3;t}$ is a speed with 1 – normal, 2 – high;
- $D_{4;t}$ is a cause of accident, where 1 – high speed, 2 – wrong driving, 3 – wrong overtaking, 4 – other;
- $D_{5;t}$ is a type of accident: 1 – crash between cars, 2– contra-flow crash, 3 – crash with a fixed object, 4 – collision with an animal.

Preliminaries: Models

The observation model

$$f\left(y_t \middle| u_t, x_t\right),$$

relates the system output y_t to the system input u_t and the unobserved system state x_t at discrete time moments $t \in t^* \equiv \{0, \ldots, \mathring{t}\}$, where $\mathring{t}$ is the cardinality of the set t^* and $\equiv$ means equivalence.
The state evolution model

$$f\left(x_{t+1} \middle| u_t, x_t\right),$$

describes the evolution of the system state x_t.

Preliminaries: Bayesian filtering

Data updating

$$f\left(x_t \middle| d^t\right) \quad \propto \quad f\left(y_t \middle| u_t, x_t\right) f\left(x_t \middle| d^{t-1}\right),$$

($\propto$ means proportionality) incorporates the experience contained in the data d^t, where $d^t = (d_0, \ldots, d_t)$ and $d_t \equiv (y_t, u_t)$.

Time updating

$$f\left(x_{t+1} \middle| d^t\right) = \int f\left(x_{t+1} \middle| u_t, x_t\right) f\left(x_t \middle| d^t\right) \, dx_t,$$

fulfills the state prediction. The filtering does not depend on the control strategy $\{f(u_t|d^{t-1})\}_{t \in t^*}$ but on the generated inputs only. The prior pdf $f(x_0)$, which expresses the subjective prior knowledge on the state x_0, starts the recursions.

Mapping to scalar variables

Let's consider the discrete-valued variables x_t, y_t and u_t, where the time moments t are irregular ones, determined by a sequence of the traffic accidents, and the chosen variables from the data sets are exploited as a discrete output and a discrete input.

In general, x_t, y_t and u_t – vectors. For example, $x_t \equiv [x_{1;t}, \ldots, x_{\mathring{x};t}]'$ with finite, preferably small $\mathring{x}$, where each entry $x_{i;t}$ with $i = \{1, \ldots, \mathring{x}\}$ has a set of possible values $\{s_1, s_2, \ldots, s_n\}$ with finite number n.

The dimension of the variables is proposed to be reduced via a special mapping, which transforms the multivariate discrete variable to a scalar one.

Mapping to scalar variables

Principle of the mapping

For example, we have vector $x_t \equiv [x_{1;t}, x_{2;t}]'$ with $x_{1;t} \in \{0, 1\}$ and $x_{2;t} \in \{0, 1\}$.

The scalar state x_t has the set of the possible values $\{[0, 0]', [1, 0]', [0, 1]', [1, 1]'\}$,

which can be denoted as the new set $\{1, 2, 3, 4\}$.

Obviously the described mapping is more successful with originally small $\mathring{x}$ and n. Nevertheless, there is no increasing of computational complexity with greater $\mathring{x}$ and n. The similar reducing of dimension is applied to the multivariate output and input.

Mapping to scalar variables

Let's assume, that the scalar variables x_t, y_t and u_t (either due to the mapping or originally) with the finite set of possible discrete values have to be considered for the filtering.

Generally, we could use discrete multinomial distribution for our state-space model.

For simplicity of presentation, let's restrict a number of possible values by two values $\{0, 1\}$, which leads to Bernoulli distribution.

Bernoulli observation model

$$f(y_t|u_t, x_t) = \prod_{j,k,l \in \{0,1\}} \alpha_{j|kl}^{\delta(j,k,l;y_t,u_t,x_t)}, \text{ with } \sum_j \alpha_{j|kl} = 1, \alpha_{j|kl} \geq 0 \, \forall j.$$

where $\alpha_{j|kl}$ is a known probability of a certain value of y_t, conditioned on corresponding values of u_t and x_t, and δ is Kronecker delta.

	$y_t = 0$	$y_t = 1$
$u_t = 1, x_t = 1$	$\alpha_{0\|11}$	$\alpha_{1\|11}$
$u_t = 0, x_t = 1$	$\alpha_{0\|01}$	$\alpha_{1\|01}$
$u_t = 1, x_t = 0$	$\alpha_{0\|10}$	$\alpha_{1\|10}$
$u_t = 0, x_t = 0$	$\alpha_{0\|00}$	$\alpha_{1\|00}$

Bernoulli state evolution model

$$f\left(x_{t+1} \middle| u_t, x_t\right) = \prod_{i,k,l \in \{0,1\}} \beta_{i|kl}^{\delta(i,k,l;x_{t+1},u_t,x_t)}, \text{ with} \sum_i \beta_{i|kl} = 1, \beta_{i|kl} \geq 0 \,\forall\, l,$$

where $\beta_{i|kl}$ is a known probability of a value of x_{t+1}, conditioned on values of u_t and x_t.

	$x_{t+1} = 0$	$x_{t+1} = 1$
$u_t = 1, x_t = 1$	$\beta_{0\|11}$	$\beta_{1\|11}$
$u_t = 0, x_t = 1$	$\beta_{0\|01}$	$\beta_{1\|01}$
$u_t = 1, x_t = 0$	$\beta_{0\|10}$	$\beta_{1\|10}$
$u_t = 0, x_t = 0$	$\beta_{0\|00}$	$\beta_{1\|00}$

Prior distribution is chosen as the Bernoulli one

$$f\left(x_t \middle| d^{t-1}\right) = \prod_{l \in \{0,1\}} p_{l(t)}^{\delta(l;x_t)}, \text{ with } \sum_l p_{l(t)} = 1, p_{l(t)} \geq 0 \ \forall \ l,$$

where $p_{l(t)}$ is a probability of taking the value l by the state x_t, that has to be recursively estimated for the time $t+1$.

Bayesian filtering with Bernoulli state-space model

$$f\left(x_{t+1}\middle|\, d^{t}\right) \propto \int f\left(x_{t+1}\middle|\, u_t, x_t\right) \left\{ \underbrace{f\left(y_t\middle|\, u_t, x_t\right) f\left(x_t\middle|\, d^{t-1}\right)}_{\propto f(x_t|d^t)} \right\} dx_t$$

$$\propto \sum_{x_t \in \{0,1\}} \prod_{i,k,l} \beta_{i|kl}^{\delta(i,k,l;x_{t+1},u_t,x_t)} \prod_{j,k,l} \alpha_{j|kl}^{\delta(j,k,l;y_t,u_t,x_t)} \prod_{l} p_{l(t)}^{\delta(l;x_t)} = \prod_{l} p_{l(t+1)}^{\delta(l;x_{t+1})},$$

where the updated probability of the value $x_{t+1} = l$ is calculated as

$$p_{l(t+1)} = \beta_{l|u_t0}\alpha_{y_t|u_t0}p_{0(t)} + \beta_{l|u_t1}\alpha_{y_t|u_t1}p_{1(t)},$$

and normalized after computation of the probabilities $\forall\ l$.
Extension up to multinomial models is straightforward

Experiments

- The proposed algorithm of the discrete-valued state filtering does not require high computational cost and enables to incorporate expert knowledge.
- That's the way it was decided to test the algorithm on the available data.
- The data sets have been rescaled until two possible values of the variables so that to be simplified and also adapted to Bernoulli model.
- Rescaling was done so that to choose the most informative data in the data sets.

Experiment 1 with scalar variables

The measured visibility $D_{2;t}$ was identified with the output y_t.

It was rescaled: $D_{2;t} \in \{1, 0\}$, where $1 \equiv$ clear visibility, and $0 \equiv$ the worse visibility with fog, rain or snow.

The measured speed $D_{3;t}$ – identified with the input u_t. It was rescaled: $D_{3;t} \in \{1, 0\}$, where $1 \equiv$ normal speed, while $0 \equiv$ high speed.

The prior probability of value $x_t = 0$ (i.e. the death accident) is chosen to be equal 0.5.

Experiment 1 with scalar variables

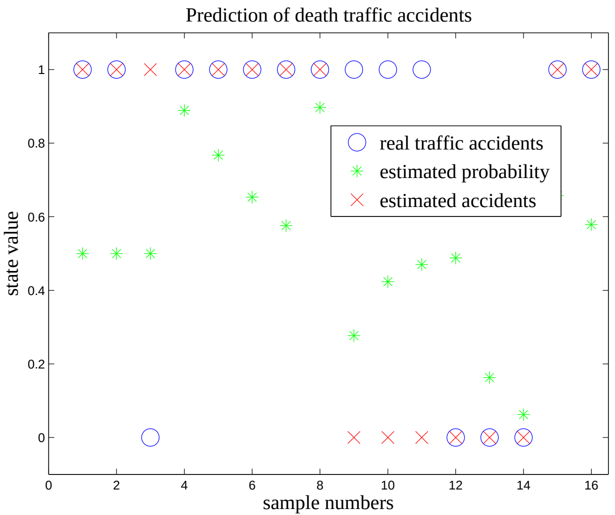
Probabilities for observation model

	$D_{2;t} = 0$	$D_{2;t} = 1$
$D_{3;t} = 1, x_t = 1$	0.68	0.32
$D_{3;t} = 0, x_t = 1$	0.49	0.51
$D_{3;t} = 1, x_t = 0$	0.04	0.96
$D_{3;t} = 0, x_t = 0$	0.45	0.55

Probabilities for state evolution model

	$x_{t+1} = 0$	$x_{t+1} = 1$
$D_{3;t} = 1, x_t = 1$	0.08	0.92
$D_{3;t} = 0, x_t = 1$	0.69	0.31
$D_{3;t} = 1, x_t = 0$	0.64	0.36
$D_{3;t} = 0, x_t = 0$	0.97	0.03

Experiment 1 with scalar variables



Experiment 2 with the mapping

The cause of the accident $D_{4;t}$ was identified with the output y_t.

It was rescaled: $D_{4;t} \in \{1, 0\}$, where $1 \equiv$ the high speed caused the accident , and $0 \equiv$ all reasons concerned with a wrong driving.

Vector $D_t^r \equiv [D_{2;t}; \quad D_{3;t}]$, where $D_{2;t}$ is the rescaled visibility and $D_{3;t}$ – speed, was identified with the input u_t.

The mapping was applied to D_t^r to reduce dimension until a scalar.

It results in the following set of values: $D_t^r \in \{D1, D2, D3, D4\}$, where $D1 \equiv [1; \ 1]$, $D2 \equiv [0; \ 1]$, $D3 \equiv [1; \ 0]$, $D4 \equiv [0; \ 0]$.

Setting of probabilities for models was more complicated task than in the previous experiment.

Experiment 2 with the mapping

Probabilities for observation model

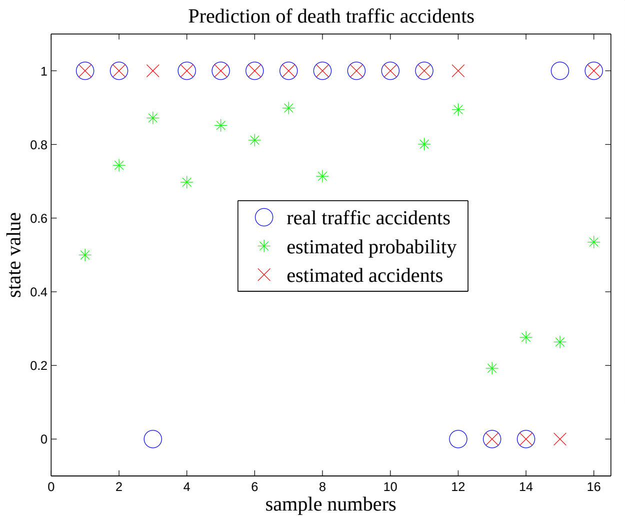
	$D_{4;t} = 0$	$D_{4;t} = 1$
$D_t^r = D1, x_t = 1$	0.87	0.13
$D_t^r = D2, x_t = 1$	0.73	0.27
$D_t^r = D3, x_t = 1$	0.04	0.96
$D_t^r = D4, x_t = 1$	0.11	0.89
$D_t^r = D1, x_t = 0$	0.96	0.04
$D_t^r = D2, x_t = 0$	0.72	0.28
$D_t^r = D3, x_t = 0$	0.18	0.82
$D_t^r = D4, x_t = 0$	0.14	0.86

Experiment 2 with the mapping

Probabilities for state evolution model

	$x_{t+1} = 0$	$x_{t+1} = 1$
$D_t^r = D1, x_t = 1$	0.04	0.96
$D_t^r = D2, x_t = 1$	0.21	0.79
$D_t^r = D3, x_t = 1$	0.83	0.17
$D_t^r = D4, x_t = 1$	0.82	0.18
$D_t^r = D1, x_t = 0$	0.96	0.04
$D_t^r = D2, x_t = 0$	0.94	0.06
$D_t^r = D3, x_t = 0$	0.12	0.88
$D_t^r = D4, x_t = 0$	0.7	0.3

Experiment 2 with the mapping



Remarks

- The quality of the results is very similar to the logistic regression. However, the last one was used in the static way, while the proposed filtering – in a dynamic one.
- The common problems for both approaches are a small amount of informative data and strong dependence on expert knowledge.
- The testing opens weak points of the algorithm: it is very sensitive to the setting of probabilities for the models.
- Probably, the estimation can be improved by better setting of probabilities.

A really important remark: the authors neither search nor invent the new optimal filtering algorithm for discrete models. This experimental work is one of items of the research of filtering with mixed (continuous and discrete) data.

Selected references

- M. Kárný, J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař,
  Optimized Bayesian Dynamic Advising: Theory and Algorithms,
  Springer, London, 2005.
- E. Suzdaleva,
  "Filtering with mixed continuous and discrete states: special case",
  Tech. Rep. 2246, ÚTIA AV ČR, Praha, January 2009,
  Draft of paper.
- E. Suzdaleva, I. Nagy, and L. Pavelková,
  "Bayesian filtering with discrete-valued state",
  in *Proceedings of 15th IEEE/SP Workshop on Statistical Signal Processing*, Cardiff, Wales, UK, Aug.31-Sept.3 2009,