Accelerate Evaluation for Lane Change Scenario using Gaussian Mixture Model and Monotonic Rare-event Set Learning

Zhiyuan Huang
Department of Industrial and
Operations Engineering
University of Michigan

Ding Zhao
University of Michigan
Transportation Research Institute

James Kirk
and Montgomery Scott
Starfleet Academy
San Francisco, California 96678–2391
Telephone: (800) 555–1212
Fax: (888) 555–1212

***Abstract*—The abstract goes here.**

I. Introduction

AV has been popular, evaluation important. Lane change scenario is interested to us, because (lane change crash data, times, percentage).

Other approaches.

In (REF), we proposed the Accelerated Evaluation concept and used the procedure to evaluate Automated Vehicles crash risks in natural driving environment. We mainly explored the application on models of interaction between AVs and human-controlled vehicles. In (REF), we modeled the frontal crash with a lead vehicle(HV or AV) using single variable Gaussian process. We also explored different approaches to tackle the evaluation of lane change scenario(REF).

In previous work, the results are based on the independence of variables(name?). We managed to decompose the density distribution and model the randomness using single variate distributions. In (REF), we used single parametric distribution to model the variables and we proposed the piecewise mixture distribution (REF) to increase the accuracy of model fitting. Cross entropy method was used to construct the importance sampling distribution. Since the independence of variables is concluded from data observation, it is not strictly proved. Ignoring the dependence between variables might lead to some estimation error. Moreover, the independence of variables is not a general condition. The requirement of independence limits the application popularity(better word needed) of the Accelerated Evaluation.

In this paper, we proposed an Accelerated Evaluation procedure using Gaussian Mixture Model for lane change scenario. (dependency) The GMM can model the dependency between variables. (accuracy) By increasing the number of components in GMM, the fitting...

Due to numbers of parameter, using cross entropy to construct IS distribution can be inefficient. (explain) We develop a new procedure based on the monotonicity property of the lane change problem. We derive a rare-event set learning procedure to construct IS distribution using both monotonicity property of rare-event sets and efficiency of change of measure for Gaussian distributions.

provide an upper and lower bound for the estimation.

Section II reviews the setting of the lane change scenario. In Section III, we review the truncated GMM fitting. The definition of monotonicity rare-event sets and the set learning algorithm are present in Section V. We introduce how to constructing an efficient Important Sampling distribution for GMM in Section IV and show the specific approach for monotonicity rare-events(monotone?). We present the simulation results for the lane change scenario in Section VII. Section VIII concludes the paper.

II. The Lane Change Scenario

The lane change scenario studied in this paper is the defined as the following: a human-controlled vehicle driving in front of an automated vehicle start to cut in the lane where the automated vehicle is running. We want to evaluate the risk of crash in this scenario. Since the initial condition for the scenario is controlled by the AV, we can take it as a random environment. Then we can use the Automated Vehicle model to simulate the output (crash or not) for a given initial condition.

To model the initial condition of lane changes, we extracted data from the Safety Pilot Model Deployment (SPMD) database (Bezzina2014). The database includes over 2 million miles of vehicle driving data collected from 98 cars over 3 years. we identify 403,581 lane change events and use 173,692 events with a negative range rate to build statistical models. Three key variables can capture the effects of gap acceptance of the lane changing vehicle: velocity of the lead vehicle (v_L), range to the lead vehicle (R_L) and time to collision (TTC_L). TTC_L was defined as:

$$TTC_L = -\frac{R_L}{\dot{R}_L}, \tag{1}$$

where $\dot{R}_L$ is the relative speed.

The simulation of Automated Vehicle model is based on Adaptive Cruise Control (ACC) and Autonomous Emergency

Braking (AEB) (Ulsoy2012a) regarding the surrounding environment. With an initial condition, the simulator returns an output of the scenario. We can considered as an event indicator function $I_\varepsilon(x)$ that returns 1 (crash) or 0 (safe) depending on the event of interest.

III. Gaussian Mixture Model Fitting

In this section, we review the truncated GMM fitting studied by Lee and Scott (REF).

The fitting of Gaussian Mixture Model is well studied and generally used. In the lane change scenario, the data can be considered as truncated. For instance, the range of two cars cannot be negative. For this reason, we use the truncated GMM to model the initial condition.

For a d dimensional dataset with truncated as a hyper-rectangle with vertices s and t, observations y^n satisfies $s \le y^n \le t$. Note that we can have $s_i = -\infty$ or $t_i = \infty$, which indicates the ith coordinate is not truncated below or above.

For a truncated GMM with K components, we use observation to estimate parameters in

$$g(y) = \sum_{k=1}^{K} \eta_k g_k(y), \tag{2}$$

where η_k is the mixing weights, g_k is truncated Gaussian with support $[s,t]$, mean μ_k and covariance Σ_k. Similarly as the vanilla GMM fitting, we use an EM algorithm to estimate the parameters. With a proper initial value for the estimated parameters, we use the following algorithm to iterate for converged estimators.

The E-step is:

$$\langle z_k^n \rangle = \frac{\eta_k g_k(y^n)}{\sum_l \eta_l g_l(y^n)}, \tag{3}$$

where we define $\langle z_k^n \rangle = p[z_k^n = 1 | y^n]$ to denote the probability that y^n is generated from the kth component.

The M-step is also similar to the vanilla GMM fitting except some correction terms:

$$\hat{\eta}_k = \frac{1}{N} \sum_n \langle z_k^n \rangle, \tag{4}$$

$$\hat{\mu}_k = \frac{\sum_n \langle z_k^n \rangle y^n}{\sum_n \langle z_k^n \rangle} - m_k, \tag{5}$$

$$\hat{\Sigma}_k = \frac{\sum_n \langle z_k^n \rangle (y^n - \hat{\mu}_k)(y^n - \hat{\mu}_k)^T}{\sum_n \langle z_k^n \rangle} + H_k, \tag{6}$$

where

$$m_k = \mathcal{M}^1(0, \Sigma_k; [s - \mu_k, t - \mu_k]), \tag{7}$$

$$H_k = \Sigma_k - \mathcal{M}^2(0, \Sigma_k; [s - \mu_k, t - \mu_k]). \tag{8}$$

$\mathcal{M}^1(\mu, \Sigma; [a,b])$ and $\mathcal{M}^2(\mu, \Sigma; [a,b])$ denote the first and second moment of truncated Gaussian distribution with truncating vertices a, b, mean μ and covariance Σ.

The choice of number of components K can be determined by some criterion for goodness of fitting, for example, we can use the Bayesian Information Criterion (BIC).

IV. Importance Sampling Distribution for Gaussian Mixture Models

In this section, we first review the concept of Importance Sampling and some known Importance Sampling schemes for rare events with Gaussian distribution. We derive the scheme for GMM based on these background knowledge.

A. *Importance Sampling for Gaussian Distribution*

Importance Sampling is a technique to reduce the variance in simulation.

Consider a random vector x with distribution F and a rare event set $\varepsilon \subset \Omega$ on sample space Ω. Our goal is to estimate the probability of the rare event

$$P(X \in \varepsilon) = E[I_\varepsilon(X)] = \int I_\varepsilon(x) dF, \tag{9}$$

where the event indicator function is defined as

$$I_\varepsilon(x) = \begin{cases} 1 & x \in \varepsilon. \\ 0 & otherwise. \end{cases} \tag{10}$$

The crude Monte Carlo using the sample mean of $I_\varepsilon(x)$

$$\hat{P}(X \in \varepsilon) = \frac{1}{N} \sum_{n=1}^{N} I_\varepsilon(X_n), \tag{11}$$

where X_i's are drawn from distribution F.

The Importance Sampling [**?**] technique is derived from

$$E[I_\varepsilon(X)] = \int I_\varepsilon(x) dF = \int I_\varepsilon(x) \frac{dF}{dF^*} dF^*, \tag{12}$$

which gives the estimator using the sample mean of the above expectation (use ref)

$$\hat{P}(X \in \varepsilon) = \frac{1}{N} \sum_{n=1}^{N} I_\varepsilon(X_n) \frac{dF}{dF^*}, \tag{13}$$

where X_i's are generated from F^*, which has the same support with F. We note that this is an unbiased estimation of $P(X \in \varepsilon)$. By appropriately selecting F^*, the evaluation procedure obtains an estimation with smaller variance. F^* is called the IS distribution. For simplicity, we define the likelihood function

$$L(x) = \frac{dF(x)}{dF^*(x)}. \tag{14}$$

The analysis of asymptotic efficiency (REF) is a benchmark to determine whether an IS distribution is proper. Let $Z = L(x) I_\varepsilon(x)$ be an IS estimator, it is weakly efficient if

$$\lim_{\varepsilon \to \infty} \frac{\log\left(E_{F^*}[Z^2]\right)}{\log\left(E_{F^*}[Z]\right)} = 2, \tag{15}$$

where $\varepsilon \to \infty$ denotes that the rare event set ε diverges to ∞ in a suitable sense (e.g., $\inf_{x \in \varepsilon} \|x\|_2 \to \infty$).

When x follows Gaussian distribution with mean μ and covariance matrix Σ and the rare event set ε satisfies certain assumptions, there is a simple scheme that obtains an weakly efficient IS distribution. Here, we introduce the scheme with assumptions satisfied in the lane change scenario.

For a convex set rare event set ε, we define

$$a^* = \arg\max_{a\in\varepsilon} \phi(a;\mu,\Sigma) \tag{16}$$

to be the dominating point of ε on $\phi(\cdot;\mu,\Sigma)$, where $\phi(\cdot;\mu,\Sigma)$ is the density function for Gaussian distribution with mean μ and covariance matrix Σ. The dominating point contributes the highest density among all points in ε. By shifting the mean μ of the Gaussian distribution to a^*, we can obtain an IS distribution that provides a weakly effecient estimator for $P(x\in\varepsilon)$ (REF).

B. IS Scheme for Gaussian Mixture Model and Union of Convex Rare Event Sets

Based on the IS scheme for Guassian distribution on convex rare event set, we propose an IS scheme for Gaussian Mixture Model on the union of convex rare event sets.

Assume x follows a Gaussian Mixture Model with k components, the density of x is

$$f(x) = \sum_{i=1}^{k} p_i\phi(x;\mu_i,\Sigma_i). \tag{17}$$

The rare event set ε is consisted with l convex sets, we denote as $\varepsilon = \cup_{j=1}^{l}\varepsilon_j$.

For each convex set ε_j, we find the dominating point of ε_j on the Gaussian component $\phi(x;\mu_i,\Sigma_i)$ by

$$a_{ij}^* = \arg\max_{a\in\varepsilon_j} \phi(a;\mu_i,\Sigma_i). \tag{18}$$

We propose to use the IS distribution as following:

$$f^*(x) = \sum_{i=1}^{k}\sum_{j=1}^{l} p_i q_j \phi(x;a_{ij},\Sigma_i), \tag{19}$$

where the q_j can be arbitrary positive number that satisfies $\sum_{j=1}^{l} q_j = 1$. We can use $q_j = 1/l$.

V. Monotonic Rare Event Sets Learning

In this section, we first define the monotonicity for rare event sets. Then we propose a learning algorithm to obtain an outer approximation set and an inner approximation set of a rare event set based on the monotonic property.

the result set is a strict inner/outer set, can provide bound for the estimation.

A. Definition of Monotonic Rare Event Sets

For a rare event set ε on d dimensional space, if $x_1 \in \varepsilon$ and $x_1 \le x_2$ (or $x_1 \ge x_2$) implies that $x_2 \in \varepsilon$, we call the set ε non-decreasing (or non-increasing).

For example, in the lane change scenario, if a crash occurs for initial condition (v_L, TTC_L, R_L), then if any of these variable is smaller, we can determine that a crash will happen. The set for crash is non-increasing. This can be explained intuitively, because a smaller v_L, TTC_L and R_L means that there is less room for the following Automated Vehicle to make adjustment.

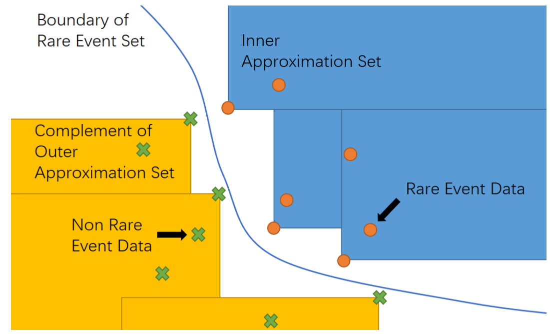

Fig. 1. An illustration of monotonic set learning.

B. Learning Algorithm

Since our goal for learning the approximation set of rare event is to construct an IS distribution, we want the approximated set to satisfy the assumption in the scheme we proposed in Section IV. Therefore, we want the approximation set to be an union of convex sets.

Let us consider a non-decreasing rare event set ε on a d dimensional space. We denote $\mathcal{S}_1 = \{a^1,...,a^{n_1}\}$ as the observed data in the rare event set and $\mathcal{S}_0 = \{b^1,...,b^{n_0}\}$ as the observed data not contained by the rare event set. We use $\mathcal{S}_1$ to construct an inner approximation set $\underline{\varepsilon}$ and use $\mathcal{S}_0$ to construct an outer approximation set $\bar{\varepsilon}$.

For each data point a_i in $\mathcal{S}_1$, we construct a set $\mathcal{I}_i = \{x : x \ge a_i\}$. By the definition of non-decreasing set, we have $\mathcal{I}_i \subset \epsilon$. Therefore, we have $\underline{\varepsilon} = \cup_{i=1}^{n_1}\mathcal{I}_i \subset \epsilon$ as the inner approximation set of ε.

For each data point b_j in $\mathcal{S}_0$, we construct sets $\mathcal{O}_{jk} = \{x : x^k \ge b_j^k\}$ for $k = 1,...,d$, where x^k denotes the kth element in x and b_j^k denotes the kth element in b_j. We have $\epsilon \subset \cup_{k=1}^{d}\mathcal{O}_{jk}$, which indicates that $\epsilon \subset \cup_{j=1}^{n_0}\cup_{k=1}^{d}\mathcal{O}_{jk}$. We define $\bar{\varepsilon} = \cup_{j=1}^{n_0}\cup_{k=1}^{d}\mathcal{O}_{jk}$ as the outer approximation set of ε.

We note that $\mathcal{I}_i$'s and $\mathcal{O}_{jk}$'s are convex, the approximation sets are unions of convex sets. We also note that using the minimums of $\mathcal{S}_1$ or maximums of $\mathcal{S}_0$ (also known as Pareto fronts) of the dataset provides the same inner or outer approximation set (minimum and maximum in a multi objective optimization sense.)

Additionally, to use the result for non-decreasing sets, we can simply flip all coordinates of data (take negative value) when we have a non-increasing set. For sets that is non-decreasing on some of the coordinates and non-increasing on the rest coordinates, we can flip the value of data those non-increasing coordinates. The approximation sets obtained will still be union of convex sets after flipping any coordinate.

VI. Scheme for Constructing IS Distribution for a Gaussian Mixture Model with a Monotonic Rare Event Set

Combining the discussion in Sections IV and V, we propose an iterated procedure that provides IS distributions for a Gaussian Mixture Model with a monotonic rare event set.

Here, we consider a non-decreasing rare event set ε on d dimensional space and the variable vector x is generated from a k components GMM with density (17). To clarify the notations, for each Gaussian component i, we construct a set of dominating points $\mathcal{A}_I^i$ using the inner approximation set $\underline{\varepsilon}$ and $\mathcal{A}_O^i$ using the inner approximation set $\bar{\varepsilon}$ for $i = 1, ..., k$. $|\mathcal{A}|$ denotes the number of elements in the set $\mathcal{A}$. $\underline{\varepsilon}$ and $\bar{\varepsilon}$ are constructed based on $\mathcal{S}_1$, which contains the minimums of observed data in ε, and $\mathcal{S}_0$, which contains the maximums of observed data that not in ε. The procedure iterates to update these sets, which will provide two IS distributions f_I^* and f_O^* based on the inner approximation and the outer approximation of ε respectively.

The procedure is present as the following:

1) Initialize $A_I^i = A_O^i = \{\mu_i\}$ and $\mathcal{S}_1 = \mathcal{S}_0 = \emptyset$.
2) Construct the sampling distribution
$$f_I^*(x) = \sum_{i=1}^{k} \sum_{a \in \mathcal{A}_I^i} p_i \frac{1}{|\mathcal{A}_I^i|} \phi(x; a, \Sigma_i) \tag{20}$$
and
$$f_O^*(x) = \sum_{i=1}^{k} \sum_{a \in \mathcal{A}_O^i} p_i \frac{1}{|\mathcal{A}_O^i|} \phi(x; a, \Sigma_i) \tag{21}$$
3) Sample N data points $D = \{x_1, ..., x_N\}$ from the density
$$f(x) = \rho f_I^*(x) + (1-\rho) f_O^*(x), \tag{22}$$
where we can pick ρ on $[0, 1]$. We can use $\rho = 1/2\ I_{\{\mathcal{S}_1 \neq \{\mu_i\}\}}$.
4) Input D to the simulator $I_\varepsilon(x)$ and use the outcome to update S_1 and S_0. We add the new data points to S_1 and S_0 regarding the outcome of $I_\varepsilon(x)$. Then we discard non-minimum data points in S_1 and non-maximum data points in S_0.
5) For each Gaussian component i, we use each data points $a \in S_1$ to solve
$$\max_x \ \phi(x; \mu_i, \Sigma_i) \text{ subject to } x \geq a. \tag{23}$$
We obtain $|S_1|$ solutions and the solution set is our new $\mathcal{A}_I^i$.
6) For each Gaussian component i, we use each data points $b \in S_0$ to solve
$$\max_x \ \phi(x; \mu_i, \Sigma_i) \text{ subject to } x_m \geq b_m, \tag{24}$$
for $m = 1, ..., d$. $x_m \geq b_m$ denotes the mth element of x is greater or equals to the mth element of b. We obtain $d|S_0|$ solutions and the solution set is our new $\mathcal{A}_O^i$.
7) Iterate from 2) to 6).

Truncated: no asymptotic on truncated coordinate.

VII. Simulation Analysis on Lane Change Scenario

In this section, we present the result in GMM fitting, IS distribution constructed and show simulation results.

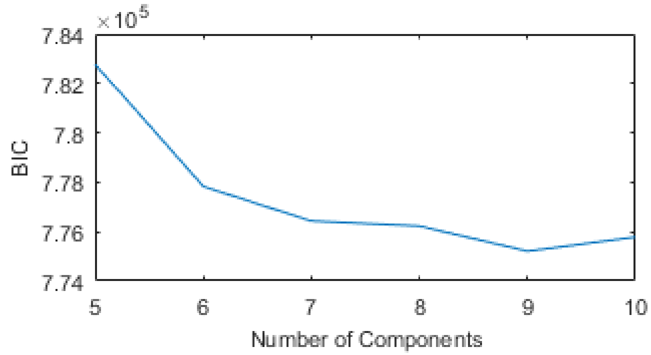

Fig. 2. The BIC regarding to different number of components K.

A. Truncated Gaussian Mixture Model Fitting

Since the truncated Gaussian Mixture Model fitting is a trivial fitting exercise, here we only present the selection of the number of components K. Fig. 2 shows the BIC regards to different K. We could observe that we obtain a local minimum at $K = 9$, where it means that the model with $K = 9$ provides a balance between the number of parameters and the fitting.

We note that the different scale of the variables might cause some numerical issues in implementing the algorithm presented in Section III, we can normalize (subtract by marginal mean and then divided by marginal standard deviation) the data before we fit the model.

In the lane change scenario, since the variables are truncated, we need to make some adjustment in the procedure. Note that if we directly use (23) and (24) for truncated variables, we might obtain a solution that out of the range of interest. In that case, the corresponding sampling distribution component

VIII. Conclusion

Acknowledgment

The authors would like to thank...

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.