Probabilistic Theory of Multi-Models

Miroslav Švítěk
Faculty of Transportation Sciences, Czech Technical University in Prague, Konviktská 20, 110 00
Prague 1, Czech Republic

Abstract:
The paper presents the mathematical theory of probabilistic multi-models composition set together by predefined probabilistic model components selected by designer. Each probabilistic component is composed of phase parameter that characterizes the composition rules of multi-models assemble and the mixture parameter that characterizes component significance to capture the features of studied system. The presented approach results in introduction of complex probabilities being represented in spectral area with similar attributes like e.g. Fourier series. Presented theory is exemplified on numerical example to demonstrate its practical use.

1. Introduction

The system modeling is extensive area of physical and technical disciplines bringing a lot of results in past. One of the main problems caused especially in large scale systems in e.g. telecommunication, transport, etc. area could be characterized like "curse of dimensionality". One way how to overcome the dimensionality problem consists in predefinition of partial low-dimensional "peaces of knowledge" represented by probabilistic model components. These model components should be known a priori, or chosen by model designer. The problem referred to in this paper is to answer the question how the set of predefined model components should be combined to catch the main features of studied system and how many parameters are necessary to perform the multi-models mixture.

The inspiration for the above defined problem came from quantum physics [4] where the backward problem against to our problem was identified - the precise set of particles' models is known from physical theory in advance but their interaction results in probabilistic transforms induced by context transformation [1,2] which involves mutual influences of physical parameters, e.g. link between position and momentum measurement.

2. Mathematical Theory of Multi-Models Composition

We start the presented mathematical theory by revision of well-known Bayes' formula for multi-model components where each model component is not known in advance and is a priori predefined by designer. This assumption will yield to introduction of probabilities conditioned on designer's decision which takes into account statistical deviation parameters that are equal to zero if designer's decision is accurate. In analogy with results achieved in [1] and [2] the multi-models components could be represented in complex area in same way like e.g. periodic signal by Fourier series. The analogy with Fourier series is right because the phases represent the rule of multi-model composition (it is structural parameter) and amplitude the significance of each model (it shows how often the model is used). The spectral representation of multi-models composition is demonstrated at the end of this chapter.

2.1 Representation of Multi-Models System

Let the sequence with \( m \) output values \( Y_j, j \in \{1,2,\ldots,m\} \) is represented by a set of \( n \) models \( P(Y_j|H_i), i \in \{1,2,\ldots,n\} \) and the models are changed over with probability \( P(H_i) \). Then according to well-known Bayes' formula the probability of \( j \)th output value could be computed as follows:

\[
P(Y_j) = \sum_{i=1}^{n} P(Y_j|H_i) \cdot P(H_i)
\]

The equation (1) holds only if we know both probabilities \( P(H_i) \) and the model components \( P(Y_j|H_i), i \in \{1,2,\ldots,n\} \). Model components \( P(Y_j|H_i) \) represent in our approach the partial knowledge of modeled system.

In practical situations the number of model components \( n \) is finite and it is often chosen as predefined set of multi-model components (mixtures) \( P(Y_j|H_i, C) \) where \( C \) denotes that the model component is conditioned on designer decision (letter \( C \) means context transition used in [1] in quantum physics representation). The probabilities \( P(H_i) \) mean the combination factors of model mixture.
In case the real model components \( P(Y_j|H_i) \) are same as designer's models \( P(Y_j|H_i,C) \), the equation (1) is fulfilled. In other case the Bayes's formula must be changed to remove the designer's decision (context transition C).

2.2 Composition of Multi-Model Components

With respect to inspiration of results achieved in quantum mechanics [1] the Bayes's formula (1) could be rewritten as following\(^1\) (here the backward process to the one in [1] studied\(^2\) and (2) is adopted for multi-models case):

\[
P(Y_j) = \sum_{i=1}^{n} P(Y_j|H_i,C) \cdot P(H_i) + 2 \sum_{k,l} P(Y_j|H_i,H_k,C) \cdot P(H_i,H_k) \cdot P(H_l) \cdot (4)
\]

where coefficients \( c_{k,l} \) are normalized statistical deviations that arise due to designer's decision-making (in [1] it means the transitions between reality and context C):

\[
c_{k,l} = \frac{1}{n-1} P(H_i) \cdot P(Y_j|H_i,C) - P(Y_j|H_l,C) \cdot P(H_l) = \frac{2}{n} \sum_{i=1}^{n} P(Y_j|H_i,C) - P(Y_j|H_l,C) \cdot P(H_l)
\]

If designer's decision is right the coefficients \( c_{k,l} \) yields to zero and equation (2) converge into equation (1).

2.3 Definition of Complex Probabilistic Models

The achieved results in [2, 3] was taken into consideration in definition of complex probabilistic models. In [3] the general random variables taking \( 2^n \) different values ware (inductively) reduced to the case of dichotomous random variables. In [2] the composition procedure of known components (theoretical models of quantum components) was applied to identify the context transited output (measured output of whole set of quantum components). Firstly the quantum system is expected to be prepared under a complex of physical conditions marked by C and than the measurement/filtration is performed. The output is so affected by preparation procedure C that yields to wave probability interpretation together with reconstruction of quantum theory on the basis of the formula of total probability [2].

The extended Bayes' formula (2) can be rewritten into complex representation defined in following theorem:

**Theorem 1:**

Let \( n \) models \( P(Y_j|H_i,C) \), \( j \in \{1,2,...,m\} \) with \( m \) output values \( Y_j \), \( j \in \{1,2,...,m\} \) denotes that \( i \)th model component \( H_i \) is conditioned on designer's decision represented by parameter C and let \( P(H_i) \) represents the probability occurrence of \( H_i \) model, than the probability of \( j \)th output value \( P(Y_j) \) could be characterized by complex parameter \( \gamma_j \) with following properties:

\[
P(Y_j) = | \gamma_j |^2
\]

\[
\gamma_j = \sum_{i=1}^{n} c_{i,j} \cdot \gamma_i
\]

where \( c_{i,j} \) are computed as follows:

\[
c_{i,j} = \sqrt{P(Y_j|H_i,C) \cdot P(H_i) \cdot \alpha(i,j)}
\]

and parameters \( \alpha(i,j) \) are figured out through algorithm described bellow:

\[
\alpha(i,1) = 0,
\]

\[
\alpha(i,2) = \alpha(i,1) - \alpha(i,2),
\]

\[
\alpha(i,3) = \alpha(i,2) + \alpha(i,2) - \alpha(i,3),
\]

\[
\vdots
\]

\[
\alpha(i,n) = \alpha(i,n-1) + \alpha(i,n-1) - \alpha(i,n)
\]

\[
\alpha(i) = \arccos\left( \frac{P(Y_j|H_i,C) \cdot P(H_i)}{P(Y_j|H_i,C) \cdot P(H_i)} \right)
\]

\[
\alpha(i) = \arccos\left( \frac{P(Y_j|H_i,C) \cdot P(H_i)}{P(Y_j|H_i,C) \cdot P(H_i)} \right)
\]

\[
\alpha(i) = \arccos\left( \frac{P(Y_j|H_i,C) \cdot P(H_i)}{P(Y_j|H_i,C) \cdot P(H_i)} \right)
\]

\[
\alpha(i) = \arccos\left( \frac{P(Y_j|H_i,C) \cdot P(H_i)}{P(Y_j|H_i,C) \cdot P(H_i)} \right)
\]

\[
\alpha(i) = \arccos\left( \frac{P(Y_j|H_i,C) \cdot P(H_i)}{P(Y_j|H_i,C) \cdot P(H_i)} \right)
\]

\[
\alpha(i) = \arccos\left( \frac{P(Y_j|H_i,C) \cdot P(H_i)}{P(Y_j|H_i,C) \cdot P(H_i)} \right)
\]

\[
\alpha(i) = \arccos\left( \frac{P(Y_j|H_i,C) \cdot P(H_i)}{P(Y_j|H_i,C) \cdot P(H_i)} \right)
\]
Proof:
For proving original Theorem 1 the formula of total probability described in [2] was used for inspiration. In quantum mechanics the back-ward problem in comparison with this paper was introduced. This means that in quantum mechanics the precise components models \( P(Y_i|H_i) \) \( i=1,2,...,n \) were known (description of quantum states without interactions) and the context transited result \( P(Y_j|C) \) was computed (result covering interactions was marked as context C probability transition).

Let us suppose that we have complete set of models \( \{H_1,H_2,...,H_n\} \) with property:
\[
P(H_1 \cdot \ldots \cdot H_n) = 1
\]

Then equation (9) and (10) could be derived from probability rules:
\[
P(Y_j) = P(Y_j|H_1 \cdot \ldots \cdot H_n) = \frac{P(Y_j|H_1,H_2 \cdot \ldots \cdot H_n) + 2 \cdot \sqrt{P(Y_j|H_1,H_2 \cdot \ldots \cdot H_n) \cdot P(Y_j|H_2 \cdot \ldots \cdot H_n)}}{2 \cdot \sqrt{P(Y_j|H_1,H_2 \cdot \ldots \cdot H_n) \cdot P(Y_j|H_2 \cdot \ldots \cdot H_n)}}
\]
(8)

The proof could be easily done by substitution of (10) into (9).
If we suppose that
\[
P(Y_j) = \sqrt{P(Y_j|H_2,C) \cdot P(Y_j|H_2 \cdot \ldots \cdot H_n)}
\]
(11)
then the equation (9) could be rewritten into complex form:
\[
. \cdot \left( Y_j \right) = \sqrt{P(Y_j|H_2,C) \cdot P(Y_j|H_2 \cdot \ldots \cdot H_n)} + e^{-i \cdot \theta} \cdot \sqrt{P(Y_j|H_2 \cdot \ldots \cdot H_n)}
\]
(12)
\[
. \cdot (1) = \text{arccos} \left( \left( Y_j[H_1,H_2 \cdot \ldots \cdot H_n] \right) \right)
\]
(13)

Because \( . \cdot \left( Y_j \right) \) is complex value it could be represented by complex module and angle:
\[
. \cdot \left( Y_j \right) = \sqrt{P(Y_j|H_1 \cdot \ldots \cdot H_n)} \cdot e^{-i \cdot \alpha_{\theta}}
\]
(14)
\[
\alpha_{\theta} = \text{arccos} \left( \frac{\sqrt{P(Y_j|H_2,C) \cdot P(Y_j|H_2 \cdot \ldots \cdot H_n)} + \sqrt{P(Y_j|H_2 \cdot \ldots \cdot H_n)}}{\sqrt{P(Y_j|H_1 \cdot \ldots \cdot H_n)}} \right)
\]
(15)

In the same way as in (9) and (10) the following equations (16) and (17) could be written as second step of derived algorithm (the proof is same):
\[
P(Y,(H_1 \cdot \ldots \cdot H_n)) = P(Y(H_1,C) \cdot P(H_1) + P(Y,(H_1 \cdot \ldots \cdot H_n)) +
+ 2 \cdot \sqrt{P(Y,H_1,H_2 \cdot \ldots \cdot H_n) \cdot P(Y,(H_1,H_2 \cdot \ldots \cdot H_n))}
\]
(16)
\[
\cdot \cdot \left( Y_j \right) = \frac{P(Y,(H_1 \cdot \ldots \cdot H_n) \cdot P(Y,(H_1,H_2 \cdot \ldots \cdot H_n) \cdot P(Y,(H_1,H_2 \cdot \ldots \cdot H_n))}{2 \cdot \sqrt{P(Y,(H_1,C) \cdot P(H_1) \cdot P(Y,(H_1 \cdot \ldots \cdot H_n))}}
\]
(17)

Then equations (12), (13), (14) and (15) could be rewritten into forms expressed in (18), (19), (20) and (21) using the same methodology as the second algorithm step:
\[
. \cdot \left( Y_j \right) = \sqrt{P(Y,(H_2,C) \cdot P(H_2) + e^{-i \cdot \theta} \cdot \sqrt{P(Y,(H_2 \cdot \ldots \cdot H_n)}}
\]
(18)
\[
. \cdot \cdot (2) = \text{arccos} \left( \left( Y_j[H_2,H_3 \cdot \ldots \cdot H_n] \right) \right)
\]
(19)
\[
. \cdot \cdot \left( Y_j \right) = \sqrt{P(Y,(H_2 \cdot \ldots \cdot H_n))} \cdot e^{-i \cdot \alpha_{\theta}}
\]
(20)
\[
\alpha_{\theta} = \text{arccos} \left( \frac{\sqrt{P(Y,H_2,C) \cdot P(H_2) + \sqrt{Y,(H_2,H_3 \cdot \ldots \cdot H_n)}}}{\sqrt{P(Y,(H_2 \cdot \ldots \cdot H_n))}} \right)
\]
(21)

The procedure described in (16) - (21) could be generalized for i-step of algorithm as it is presented in (22) - (27):
\[
P(Y,(H_1 \cdot \ldots \cdot H_n)) = P(Y(H_i,C) \cdot P(H_i) + P(Y,(H_1 \cdot \ldots \cdot H_n)) +
+ 2 \cdot \sqrt{P(Y,H_{i-1},H_i \cdot \ldots \cdot H_n) \cdot P(Y,(H_{i-1},H_i \cdot \ldots \cdot H_n))}
\]
(22)
\[
\cdot \cdot \left( Y_j \right) = \frac{P(Y,(H_{i-1},H_i \cdot \ldots \cdot H_n) \cdot P(Y,(H_{i-1},H_i \cdot \ldots \cdot H_n))}{2 \cdot \sqrt{P(Y,(H_i,C) \cdot P(H_i) \cdot P(Y,(H_{i-1},H_i \cdot \ldots \cdot H_n))}}}
\]
(23)
\[
. \cdot \cdot \left( Y_j \right) = \sqrt{P(Y,(H_{i+1},C) \cdot P(H_{i+1}) + e^{-i \cdot \theta} \cdot \sqrt{P(Y,(H_{i+1} \cdot \ldots \cdot H_n)}}
\]
(24)
\[
. \cdot \cdot \cdot (i) = \text{arccos} \left( \left( Y_j[H_i,H_{i+1} \cdot \ldots \cdot H_n] \right) \right)
\]
(25)
\[ \hat{Y}_j = \sqrt{P(Y_j|H_1, ..., H_n)} e^{i\alpha_j(i)} \]  

(26)

\[ \alpha_j(i) = \arccos \left( \frac{P(Y_j|H_1, C) e^{i\beta_{1j(i)}} P(H_1) + \cdots + P(Y_j|H_n, C) e^{i\beta_{nj(i)}} P(H_n)}{P(Y_j|H_1, ..., H_n)} \right) \]  

(27)

By combining different expressions of \( \hat{Y}_j \) described above the complex representations of partial component \( \hat{Y}_j \) for \( i \in \{1,2,\ldots,n\} \) together with expression of their phase \( \alpha_j(i) \) could be derived as follows:

\[ \hat{Y}_j = \sqrt{P(Y_j|H_1, C)} P(H_1) e^{i\alpha_j(i)} \]  

From (28) the algorithm for phase representation \( \alpha_j(i) \) of complex models probabilities could be easily derived:

\[ \alpha_j(1) = 0 \]

\[ \alpha_j(2) = \alpha_j(1) \]

\[ \alpha_j(3) = \alpha_j(1) + \alpha_j(2) \]

\[ \alpha_j(n) = \alpha_j(n-1) + \alpha_j(n-1) \]  

(29)

Theorem 1 is thus proved.

2.5 Spectral Representation of Complex Multi-Models Probabilities

In a same way as the periodic signal is represented in spectral area by well-known Fourier series (amplitude and phase of each harmonic component) the complex probabilities of multi-models could be also represented in complex plain where module has analogy with probability (amplitude reflects the rate of model component occurrence in long time period) and phase describes how the multi-models should be composed to describe the real system (phase reflect overlapping of multi-models selected a priori by designer).

Because Theorem 1 is independent on models selection \( P(Y_j|H_1, C) \) i.e. \( \{1,2,\ldots,n\} \), these models could be chosen in advance to cover whole range of probabilistic area. The selection of models \( P(Y_j|H_1, C) \) has analogy with frequency selection in Fourier transform. The parameters \( P(H_1) \) and \( \alpha_j(i) \) could be estimated from real data sample (like amplitude and phase in Fourier transform). Generally the more multi-models the lesser phase (for greater magnitude of statistical deviations, the hyperbolic or hyper-trigonometric transform could be applied as is cited in [1]).

For example, if we have system with two output components \( Y \). \( \{0,1\} \) and we chose a priori the number of models \( n=4 \), then models could be for example defined as follows:

<table>
<thead>
<tr>
<th>MODEL IDENTIFICATION</th>
<th>H_1</th>
<th>H_2</th>
<th>H_3</th>
<th>H_4</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(Y = 1</td>
<td>H_1, C)</td>
<td>0.8</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>P(Y = 0</td>
<td>H_1, C)</td>
<td>0.2</td>
<td>0.4</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Tab.1 Predefined probabilistic multi-models components for \( Y \). \( \{0,1\} \) and \( n=4 \).

Than complex amplitude and phase representation could be expressed as in Fig.1.

![Fig.1 Spectral representation of multi-models complex probabilities](image_url)

3. Numerical example

Let two values time series \( Y \). \( \{0,1\} \) is composed from mixture of three time series (component time series) described by probabilities \( P(Y|H_1), P(Y|H_2), P(Y|H_3) \)
where each component time series occur with probabilities $P(H_1), P(H_2)$ and $P(H_3)$. The probabilities $P(Y|H_i), i \in \{1,2,3\}$ are defined in Tab.2.

The probabilities $P(H_i)$ were chosen:

$$P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}$$

(30)

<table>
<thead>
<tr>
<th>MODEL IDENTIFICATION</th>
<th>$H_1$</th>
<th>$H_2$</th>
<th>$H_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(Y = 1</td>
<td>H_1)$</td>
<td>0.9</td>
<td>0.5</td>
</tr>
<tr>
<td>$P(Y = 0</td>
<td>H_1)$</td>
<td>0.1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Tab.2 Real time series components $P(Y|H_1), i \in \{1,2,3\}$

By using equation (6) together with algorithm (7) the following complex components could be numerically calculated:

$$Y(1) = \sqrt{P(Y = 1|H_1, C)P(H_1)} \cdot e^{i\theta(1)} = 0.5164$$

(31)

$$Y(1) = \sqrt{P(Y = 1|H_2, C)P(H_2)} \cdot e^{i\theta(2)} = 0.4472 \cdot e^{0.5166}$$

(32)

$$Y(1) = \sqrt{P(Y = 1|H_3, C)P(H_3)} \cdot e^{i\theta(3)} = 0.4830 \cdot e^{0.4012}$$

(33)

$$Y(0) = \sqrt{P(Y = 0|H_1, C)P(H_1)} \cdot e^{i\theta(1)} = 0.2582$$

(34)

Based on equation (5) the two complex parameters could be fond as follows:

$$Y(1) = Y(1) + Y(1) = Y(1) + Y(1) = 0.7746 \cdot e^{0.7637}$$

(37)

(38)

where probabilities of falling one or zero could be computed:

$$P(Y = 1) = \frac{Y(1)^2}{0.6} = 0.6$$

(39)

$$P(Y = 0) = \frac{Y(0)^2}{0.4} = 0.4$$

(40)

The outcomes (39), (40) agree with the result achieved with the knowledge of model components given in Tab.1 and by using of Bayes' formula (1):

$$P(Y = 1) = \frac{Y(1)^2}{0.6} = 0.6$$

(31)

(32)

(33)

(34)

The above mentioned numerical example shows that the theory of multi-models composition is feasible. In practical analyzes the amplitudes and phases of model components will be estimated from real time series.

4. Conclusion

The theory of probabilistic multi-models composition is the starting point of large scale systems modeling by complex probabilities representing the partial knowledge of studied systems. The results could be used for finding decision-making or control strategies of such systems, etc.

The presented methodology was shown only for case of output time series modeling but it could be extended to dynamical system modeling [5, 6, 7] where partial system knowledge reflect predefined low dimensional models with predefined control strategies. Composition of dynamical system components could yield to new algorithm of real system control strategy.
References


