Applicability evaluation of Akaike's Bayesian information criterion to covariance modeling in the cross-section adjustment method

Shuhei Maruyama[1*], *Tomohiro* Endo[2], and *Akio* Yamamoto[2]

[1]Japan Atomic Energy Agency, Reactor Core and Plant System Evaluation Group, Fast Reactor Cycle System Research and Development Center, 4002 Narita-cho, Oarai-machi, Higashi-Ibaraki-gun, Ibaraki, Japan
[2]Nagoya University, Department of Applied Energy, Graduate School of Engineering, Furo-cho, Chikusa-ku, Nagoya, Japan

Abstract. The applicability of Akaike's Bayesian Information Criterion (ABIC) to the covariance modeling in the cross-section adjustment method has been investigated. In the conventional cross-section adjustment method, the covariance matrices are assumed to be true. However, this assumption is not always appropriate. To improve the reliability of the cross-section adjustment method, the estimation of the covariance model using the metric ABIC has been introduced, and the performance of ABIC has been investigated through simple numerical experiments. This paper derives the formula to efficiently evaluate ABIC which is represented by a lower rank matrix to enable numerical experiments with large samples in a realistic computation time. From the results of the numerical experiments, it has been confirmed that ABIC tends to select a covariance model with fewer hyperparameters and a smaller variance for the estimation error. However, it has also been found that this desirable property of ABIC will be lost when the structure of the covariance model is far from the true one.

1 Background

Every measurement and analysis values always have uncertainties. Namely, the data follow a probability distribution with parameters such as population mean and covariance. Unfortunately, we cannot know the true parameters. Therefore, when we analyze the data, we usually estimate and/or assume these parameters. The appropriateness of the covariance matrix set (e.g., nuclear-data covariance) by analysts has been widely discussed in the framework of Sub-Groups (SGs) under the Working Party on International Nuclear Data Evaluation Cooperation (WPEC) in OECD/NEA (e.g., [1‑4]). Recently, the discussion had been also done in Japan through the activity of the covariance data utilization and promotion working group organized in the JENDL committee [5].

In the cross-section adjustment method, covariance matrices are also used. One of the most important things for a reliable cross-section adjustment method is giving suitable covariance matrices close enough to the true covariance matrices. To judge the goodness of the covariance modeling, a metric is desirable. As a candidate for this metric, we focus on Akaike's Bayesian Information Criterion (ABIC) [6] which is one of the information criteria in Bayesian inference, because the cross-section adjustment method is often discussed within the framework of Bayesian inference.

In the conventional cross-section adjustment method, incorporation of the analysis method errors (errors due to the core calculation method, e.g., discretizing error in a deterministic code) as a covariance matrix still requires ad hoc treatment. In JAEA, the integral experimental database for fast reactors has been developed and the adjusted cross-section set ADJ2017 [7, 8] has been created based on this database. Many of the core characteristics in the database have been analyzed by a deterministic method. Therefore, the predicted core characteristics have non-negligible uncertainties with correlations due to some numerical approximations. However, evaluating the uncertainties and their correlations is still a challenging issue. In addition, there would be unknown uncertainties that experimenters and analysts of reactor physics experiments were not able to recognize.

In order to address the difficulties in the conventional adjustment method, we will try to incorporate ABIC into the adjustment method. ABIC is expected to work as a metric for evaluating the goodness of covariance matrix modeling related to the uncertainties. We aim to improve the reliability of the conventional cross-section adjustment method by ABIC. This paper investigates the applicability of ABIC through several numerical experiments using random sampling techniques.

We introduce ABIC and incorporate this into the conventional cross-section adjustment in the next section. The applicability is evaluated through the simple numerical experiments in section 3. The conclusion of this paper will be described in section 4.

* Corresponding author: maruyama.shuhei@jaea.go.jp

2 Methodology

Before we get to the main subject, we introduce the notations that will be used in this paper. The covariance matrix with the tilde (" ~ ") means a covariance model with hyperparameters determined later. The covariance matrix with a hat (" ^ ") means the selected covariance by determining the hyperparameters. These covariances are not always "true." The true covariance is represented by the characters without a tilde or a hat. For example,

$\mathbf{V}$: a true covariance matrix,

$\widetilde{\mathbf{V}}$: a covariance model with undetermined hyper-parameters,

$\widehat{\mathbf{V}}$: the covariance selected among a covariance model $\widetilde{\mathbf{V}}$ by determining hyperparameters.

2.1 Review of conventional cross-section adjustment method

When true nuclear data set $\boldsymbol{T}$ follows a multivariate normal distribution with the mean $\boldsymbol{T_0}$ and covariance $\mathbf{M}$, the probability distribution of $\boldsymbol{T}$ is

$$\mathcal{P}(\boldsymbol{T}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{M}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\boldsymbol{T}-\boldsymbol{T_0})^{\mathrm{T}} \cdot \mathbf{M}^{-1}(\boldsymbol{T}-\boldsymbol{T_0})\right], \quad (1)$$

where n is the dimension of $\boldsymbol{T}$ and $\boldsymbol{T_0}$, $|\mathbf{X}|$ is the determinant of a matrix $\mathbf{X}$, and the superscripts -1 and T represent the inverse and transpose of the matrix respectively. When integral experimental data sets are obtained, the likelihood function is represented as follows:

$$\mathcal{P}(\boldsymbol{R_e}|\boldsymbol{T}) = \frac{1}{(2\pi)^{\frac{r}{2}}|\mathbf{V_e}+\mathbf{V_m}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}\left(\boldsymbol{R_e}-\boldsymbol{R_c}(\boldsymbol{T})\right)^{\mathrm{T}} \cdot (\mathbf{V_e}+\mathbf{V_m})^{-1}\left(\boldsymbol{R_e}-\boldsymbol{R_c}(\boldsymbol{T})\right)\right]. \quad (2)$$

$\boldsymbol{R_e}$ is a set of experimental values and $\mathbf{V_e}$ is the experimental covariance matrix. $\boldsymbol{R_c}(\boldsymbol{T})$ is a set of calculational values if the true nuclear data set $\boldsymbol{T}$ is given, and $\mathbf{V_m}$ is the covariance matrix due to calculation methods. Here, r denotes the dimension of $\boldsymbol{R_e}$ and $\boldsymbol{R_c}$. From Bayes' theorem,

$$\mathcal{P}(\boldsymbol{T}|\boldsymbol{R_e}) \propto \mathcal{P}(\boldsymbol{R_e}|\boldsymbol{T})\mathcal{P}(\boldsymbol{T}), \quad (3)$$

the posterior probability is represented as follows:

$$\mathcal{P}(\boldsymbol{T}|\boldsymbol{R_e}) \propto \exp\left[-\frac{1}{2}(\boldsymbol{T}-\boldsymbol{T_0})^{\mathrm{T}}\mathbf{M}^{-1}(\boldsymbol{T}-\boldsymbol{T_0}) -\frac{1}{2}\left(\boldsymbol{R_e}-\boldsymbol{R_c}(\boldsymbol{T})\right)^{\mathrm{T}}(\mathbf{V_e}+\mathbf{V_m})^{-1}\left(\boldsymbol{R_e}-\boldsymbol{R_c}(\boldsymbol{T})\right)\right]. \quad (4)$$

The cross-section and its covariance after adjustment are derived from the condition that maximizes the posterior probability of Eq. (4) as follows:

$$\boldsymbol{T'} = \boldsymbol{T_0} + \mathbf{M}\mathbf{G}^{\mathrm{T}}\mathbf{V_t}^{-1}\left(\boldsymbol{R_e}-\boldsymbol{R_c}(\boldsymbol{T_0})\right), \quad (5)$$

$$\mathbf{M'} = \mathbf{M} - \mathbf{M}\mathbf{G}^{\mathrm{T}}\mathbf{V_t}^{-1}\mathbf{G}\mathbf{M}, \quad (6)$$

where

$$\mathbf{V_t} \equiv \mathbf{G}\mathbf{M}\mathbf{G}^{\mathrm{T}} + \mathbf{V_e} + \mathbf{V_m}. \quad (7)$$

Here, $\mathbf{G}$ means nuclear-data sensitivity coefficients. Note that the linear approximation for $\boldsymbol{R_c}(\boldsymbol{T})$, *i.e.*,

$$\boldsymbol{R_c}(\boldsymbol{T}) = \boldsymbol{R_c}(\boldsymbol{T_0}) + \mathbf{G}(\boldsymbol{T}-\boldsymbol{T_0}), \quad (8)$$

is assumed to derive Eqs. (5) and (6). The concept of the cross-section adjustment method is shown in Fig. 1.

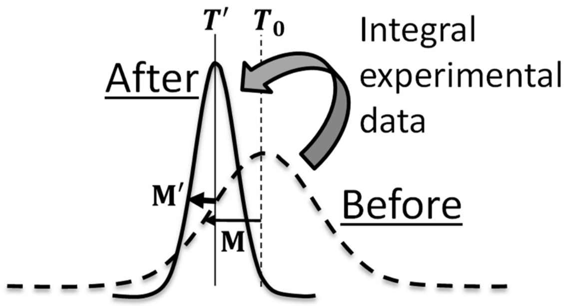

Fig. 1. Concept of cross-section adjustment

2.2 Akaike's Bayesian information criterion

In the conventional cross-section adjustment method, the covariance matrices $\mathbf{M}$, $\mathbf{V_e}$, and $\mathbf{V_m}$ used for the adjustment are assumed to be true. However, these are not always appropriate. Therefore, let us assume the covariance matrices have hyperparameters that are determined to suit the observed data. To find a better covariance model by tuning the hyperparameters, we propose to use ABIC in this paper. As mentioned in the previous section, since the setting of the covariance matrix $\mathbf{V_m}$ due to a calculation method is a challenging issue in the conventional adjustment method, only $\mathbf{V_m}$ is assumed to be unknown and to be the estimation target using ABIC in this study. In other words, $\mathbf{M}$ and $\mathbf{V_e}$ in this study are assumed to be true.

If the covariance model $\widetilde{\mathbf{V}}_\mathbf{m}$ has unknown positive hyperparameters $p_1^2, p_2^2, \cdots, p_q^2$, the likelihood function of Eq. (2) is represented as follows:

$$\mathcal{P}\left(\boldsymbol{R_e}|\boldsymbol{T}; p_1^2, p_2^2, \cdots, p_q^2\right) \propto \exp\left[-\frac{1}{2}\left(\boldsymbol{R_e}-\boldsymbol{R_c}(\boldsymbol{T})\right)^{\mathrm{T}} \cdot \left(\mathbf{V_e}+\widetilde{\mathbf{V}}_\mathbf{m}(p_1^2, p_2^2, \cdots, p_q^2)\right)^{-1}\left(\boldsymbol{R_e}-\boldsymbol{R_c}(\boldsymbol{T})\right)\right], \quad (9)$$

where q is the number of hyperparameters. As shown later (subsection 3.1.3), we prepare positive hyperparameters expressed as squared values to guarantee the positive definiteness of the covariance matrix $\widetilde{\mathbf{V}}_\mathbf{m}$. As a metric of the goodness of an inference model including the covariance modeling, we employ an information criterion called ABIC,

$$\mathrm{ABIC} = -2\ln l\left(p_1^2, p_2^2, \cdots, p_q^2\right) + 2q. \quad (10)$$

Here, l is the marginal likelihood defined as

$$l\left(p_1^2, p_2^2, \cdots, p_q^2\right) = \int d\boldsymbol{T}\ \mathcal{P}\left(\boldsymbol{R_e}|\boldsymbol{T}; p_1^2, p_2^2, \cdots, p_q^2\right)\mathcal{P}(\boldsymbol{T}). \quad (11)$$

ABIC consists of the two terms related to l and q. The marginal likelihood l is proportional to the generation probability of the observed data (integral experimental data set $\boldsymbol{R_e}$). By choosing a good inference model that leads to this probability, *i.e.*, a suitable covariance model, ABIC is decreased. On the other hand, the number of hyperparameters q works penalty in terms of the complexity of an inference model. Therefore, ABIC prefers a simpler model having fewer hyperparameters. In a previous study, a cross-section