Guidelines and Gotchas for Covariance Data Testing

D. Kent Parsons

Nuclear Data Team

LA-UR-20-29752, December 2020

Talk Outline

- Basics for Covariance Matrix Data
- Guidelines for Appropriate Use of Covariance Data
- "Gotchas" to watch out for when Using the Data

- We recommend covariance testing and are developing robust testing software at Los Alamos to do the testing.

Managed by Triad National Security, LLC for the U.S. Department of Energy's NNSA

Background Document

NUCLEAR DATA AND MEASUREMENTS SERIES

ANL/NDM-104

Some Thoughts on Positive Definiteness in the Consideration of Nuclear Data Covariance Matrices

by

L.P. Geraldo and D.L. Smith

January 1988

ARGONNE NATIONAL LABORATORY,
ARGONNE, ILLINOIS 60439, U.S.A.

Covariance Data in ENDF/B VIII.0

- Covariance data is found in the MF 31-40 sections of the Evaluated Data File
  - Corresponding to MF 1-10 (i.e., add 30)
  - e.g., Reaction Cross sections in MF 3, have covariances in MF 33
- Covariance data is in the old ENDF 6 format
  - Stored in (essentially) "single precision", 6-7 digits
  - A few numeric entries without an exponent might have 10-11 significant digits
  - (Future data formats will have more precision)

Covariance Data in ENDF/B VIII.0

- ENDF/B VIII.0 Covariance data includes:
  - Variances for each data point
    - *along the diagonal*
  - Covariances between different data points
    - *off-diagonal terms*
    - With respect to energy (for a given reaction or data of an isotope)
    - With respect to energy and to other reactions or data of the <u>same</u> isotope
    - (yet future) With respect to energy and other reactions or data of <u>different</u> isotopes

Covariance Matrices

- Symmetric
- Positive Definite (or at least, Semi-Positive Definite -- because of threshold reactions)
  - No Negative nor 0.0 Eigenvalues for Positive Definite
  - No Negative Eigenvalues for Semi-Positive Definite

- This also applies to the related Correlation Matrices and to the related Relative Covariance Matrices

Correlation Matrices

- Symmetric
- Positive Definite (or at least, Semi-Positive Definite)
  - No Negative nor 0.0 Eigenvalues for Positive Definite
  - No Negative Eigenvalues for Semi-Positive Definite

- Also, the values of the diagonal elements of the Correlation Matrix must be 1.0 and the off-diagonal matrix elements must be between -1.0 and 1.0.

Typical Applications of Covariance Data

- (0) First, Obtain Group-Wise Covariances from NJOY processing of the Evaluated Covariance data, Then:
- (1) For Steady State (linear first order)
  - Obtain group-wise linear first-order problem-dependent Sensitivities from MCNP ksen option (with tallies into the group structure) or SENSMG calculations
  - Use the “sandwich rule”:

    Sensitivity * Covariance Matrix = overall Uncertainty (i.e., a variance)
- (2) An Alternative Approach
  - Generate multiple realizations of the data from the Covariance Matrix
  - Run the analysis code with each set of input realization data
  - Collect statistics of the code outputs (at least 75x!!)

General Guidelines for Working with Covariance Data

- Preserve as much numerical precision as possible in the covariance data values
  - Future nuclear data formats will help with this
- If the values of the data being studied span many powers of 10 (e.g., fission chi's), consider using relative covariances
- The end result of the sandwich rule is a variance (the square of the standard deviation) and it is a positive number

What Could Go Possibly Wrong?

- Negative Eigenvalues in the Evaluated Covariance Data (e.g., see capture in Pu239 and mubar in U235)
  - If they are large and negative –
    - INVALID COVARIANCE DATA → return to the evaluator
  - If they are negative and close to 0.0
    - Probably a result of limited numerical precision
- After NJOY processing to get group-wise data
  - If the group structure is finer than the resolution of the evaluated covariance data → NJOY generates more small negative eigenvalues – “group expansion”
  - “group collapsing” is ok

How About Locating the Big Negative Eigenvalues?

- Using modern software tools, you can usually locate the general energy range of the problems in the covariance data
  - For 50x50 matrix with problems, try a 49x49, then a 48x48, etc.
  - Choleski Decomposition can also be used
  - *In the old days, determinants and minors were calculated by hand to locate the general energy range of the problem …*

What's a Few Small Negative Eigenvalues amongst Friends?

- Many of the "canned" statistical software routines in MATLAB, Python, R, etc. will NOT work in the presence of ANY negative eigenvalues
  - Covariance to Correlation Conversion
  - Multi-Variate Generation of Correlated Random Samples
- For the sandwich rule, it is very easy to calculate a "silent" wrong answer due to a problematic covariance matrix
  - Unless the final answer comes out negative …

General Gotcha's

- SENSMG and MCNP ksen have reverse conventions for group ordering. (SENSMG IS high E to low E, MCNP and NJOY are low E to high E.) Be sure that the sensitivity vectors and the covariance matrix in the sandwich rule have consistent group ordering!
- Relative Sensitivities (produced by SENSMG and MCNP) go with Relative Covariances in the sandwich rule;   *(alternatively, you may use Absolute Sensitivities with Absolute Covariances)*

How About Repairing Covariance Data with small Negative Eigenvalues?

- Several "ad hoc" methods exist:
  – Find the (valid) covariance matrix "closest" to the bad covariance matrix
  – "Ridge Correction" – add a small epsilon to all of the diagonal elements, thus increasing all of the eigenvalues by epsilon
  – Ignore the small negative eigenvalues and their eigenvectors while generating correlated random samples
- Active research area …

Well, What Else Could Go Wrong?

- Certain constraints exist on some nuclear data:
  - For Fission Chi's in a group structure, the sum = 1.0
    - Leads to the constraint that any "row sum" or "column sum" of the fission chi covariance matrix = 0.0
    - *(BTW, This constraint may also be enforced on the sensitivity vectors for the sandwich rule, but do NOT apply the constraint to both at the same time!)*
  - For the Total Cross Section = the sum of the partials
    - Leads to group-wise constraints on the covariance matrix
      - The covariance of the total cross section = the sum of the covariances of the partial cross sections
- Generation of valid random samples requires these constraints

Anything Else that Could Go Wrong?

- Mubar (cosine of the lab scattering angle) is strictly limited to values between -1.0 and 1.0
- In ENDF/B VIII.0 data for O-16, the uncertainty on mubar is many, many times larger than this very specific limited range.

- If relative uncertainties of a cross section are > than about 33%, then negative random samples will be generated for “normal” distributions … codes don’t allow negative cross sections or fission chi’s.

How much covariance data is there?

- 557 isotopes in ENDF/B VIII.0;
- 4227 individual covariance data sections
  - ~181 have covariance data for the total cross section
  - ~214 have covariance data for elastic scattering
  - ~211 have covariance data for (n,2n)
  - ~71 have covariance data for fission, prompt and total nu, and fission chi's
  - ~108 have mubar covariances
  - … (inelastic and inelastic by levels, n3n, other reactions)

Summary (Part 1)

- **What codes are available at LANL for testing covariances?** We are working on a robust code system which will automate the process of testing.
- **What are you testing?** We are testing for symmetry, positive or semi-positive definite (by eigenvalues), valid entries in the correlation matrix, covariance constraints on fission chi's and total cross sections, realistic values for mubar, and large (say ~33%) relative uncertainties in cross sections or fission chi's
- **What do you do with questionable covariances?** For large negative eigenvalues or non-physical uncertainties for mubar, we send them back to the evaluator. For small negative eigenvalues, we have tried various "ad hoc" approaches and we are still investigating

Summary (Part 2)

- **How Much Time to Test all ENDF/B VIII.0 Covariances?** We don't know yet, but there are 557 isotopes and 4227 covariance data sets in ENDF/B VIII.0. Note that we intend to focus on the evaluated data, since NJOY processing is not really the source of the negative eigenvalues (except for "group expansion"). Of course, we will also check NJOY produced covariance matrices. Once our robust testing code is written, it shouldn't take very long to run (and rerun) the analysis.

Summary (Part 3)

- **What Developments are needed for testing to satisfy your users?** We need a better feel for the effects of the violations of covariance rules and constraints. For example, can we be just a little bit off and still be ok? Just how sensitive are the output results to each rule and constraint? Other questions would be: How can we detect trouble in a sandwich rule result? Or what methods of repairing bad covariances work the best?

Los Alamos
NATIONAL LABORATORY
EST. 1943