Neutron Multiplicity: LANL W Covariance Matrix for Curve Fitting

James George Wendelberger[1]
[1]Los Alamos National Laboratory, Los Alamos, NM 87545

Abstract
In neutron multiplicity counting one may fit a curve by minimizing an objective function, χ_n^2. The objective function includes the inverse of an n by n matrix of covariances, W. The inverse of the W matrix has a closed form solution. In addition, W^{-1} is a tri-diagonal matrix. The closed form and tri-diagonal nature allows for a simpler expression of the objective function χ_n^2. Minimization of this simpler expression will provide the optimal parameters for the fitted curve.

Key Words: Neutron, multiplicity counting, tri-diagonal, uncertainty, Lehmer matrix, covariance matrix

1. The Calculation of χ^2

In Walston [6] we find reference to the χ^2 and the Lawrence Livermore National Laboratory (LLNL) W matrix. The Los Alamos National Laboratory (LANL) W matrix is the same as the LLNL W because the correlation is identical between the random variables of the gate size. The estimates of the correlation will have the same expected value and those of LLNL are more variable as the gate data from all except the largest gate size is not entirely utilized for each gate size.

In Walston [6] we find the i, j element of the n by n matrix W is

$$W_{i,j} = \rho_{i,j}\sigma_{Y_{2F}}(T_i)\sigma_{Y_{2F}}(T_j). \qquad (1)$$

Define the i, j element of the n by n matrix Σ as:

$$\Sigma_{i,j} = \begin{cases} \sigma_{Y_{2F}}(T_i), & i = j \\ 0, & i \neq j. \end{cases} \qquad (2)$$

Σ is an n by n diagonal matrix with zeros on the off diagonal.

The n by n correlation matrix P contains the correlations as:

$$P_{i,j} = \rho_{i,j}. \qquad (3)$$

The correlations are listed in Prasad, Snyderman, and Walston [5], and the correlation matrix, P, is in fact the Lehmer matrix, proposed by Lehmer [1], or

$$P_{i,j} = \frac{\min(i,j)}{\max(i,j).} \quad (4)$$

In matrix form:

$$W = \Sigma\, P\, \Sigma. \quad (5)$$

The inverse of W is:

$$W^{-1} = \Sigma^{-1} P^{-1}\, \Sigma^{-1}. \quad (6)$$

As

$$W\, W^{-1} = \Sigma\, P\, \Sigma\, \Sigma^{-1} P^{-1}\, \Sigma^{-1} = I = W^{-1}\, W. \quad (7)$$

The inverse of the Lehmer matrix P is a tri-diagonal matrix P^{-1}, originally solved in Lehmer, Smiley, Smiley, and Williamson [2], with entries:

$$P_{i,j}^{-1} = \begin{cases} 4i^3/(4i^2-1), & i = j \text{ and } i < n \\ n^2/(2n-1), & i = j = n \\ -\min(i,j)(\min(i,j)+1)/(2\min(i,j)+1), & |i-j| = 1 \\ 0, & |i-j| > 1. \end{cases} \quad (8)$$

P^{-1} is a tri-diagonal matrix with zeros on the off tri-diagonal. The Lehmer matrix has also been used to test the inversion of a tri-diagonal matrix, Newman and Todd [4] and for evaluation of matrix inversion programs Lewis [3].

The inverse of Σ is the n by n matrix Σ^{-1} where

$$\Sigma_{i,j}^{-1} = \begin{cases} 1/\sigma_{i,j}, & i = j \\ 0, & i \neq j. \end{cases} \quad (9)$$

For completeness use (6), (8) and (9) to write the i, j-th term of W^{-1} as

$$W_{i,j}^{-1} = \begin{cases} \Sigma_{i,i}^{-2} 4i^3/(4i^2-1), & i = j \text{ and } i < n \\ \Sigma_{n,n}^{-2}\, n^2/(2n-1), & i = j = n \\ -\Sigma_{i,i}^{-1}\Sigma_{ij,ji}^{-1} \min(i,j)(\min(i,j)+1)/(2\min(i,j)+1), & |i-j| = 1 \\ 0, & |i-j| > 1. \end{cases} \quad (10)$$

The curve fit uses parameters which minimize χ^2 where

$$\chi^2 = \mathbf{E}_{\mathbf{2F}}^{\mathrm{T}}\, W^{-1}\, \mathbf{E}_{\mathbf{2F}}. \quad (11)$$

Rewriting (11) by using (6) to expand W^{-1} yields:

$$\chi^2 = \mathbf{E}_{\mathbf{2F}}^{\mathrm{T}}\, \Sigma^{-1} P^{-1}\, \Sigma^{-1}\, \mathbf{E}_{\mathbf{2F}}. \quad (12)$$

Divide the residual $\mathbf{E_{2F}}$ by its corresponding standard deviation to create scaled residuals as:

$$\mathbf{E_{2F,\sigma}} = \Sigma^{-1}\,\mathbf{E_{2F}}. \tag{13}$$

In this way one may "absorb" the Σ^{-1} into the residual to compute the χ^2. Using (13) and (12) obtain:

$$\chi^2 = \mathbf{E}^{\mathrm{T}}_{\mathbf{2F},\sigma}\,\mathrm{P}^{-1}\;\mathbf{E_{2F,\sigma}}. \tag{14}$$

2. Example Correlation Matrix with n = 5

As an example let $\mathrm{n} = 5$ then we have:

$$\mathrm{P}_5 = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1 & 2/3 & 1/2 & 2/5 \\ 1/3 & 2/3 & 1 & 3/4 & 3/5 \\ 1/4 & 1/2 & 3/4 & 1 & 4/5 \\ 1/5 & 2/5 & 3/5 & 4/5 & 1 \end{bmatrix}. \tag{15}$$

Simplify fractions:

$$\mathrm{P}_5 = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1 & 2/3 & 2/4 & 2/5 \\ 1/3 & 2/3 & 1 & 3/4 & 3/5 \\ 1/4 & 2/4 & 3/4 & 1 & 4/5 \\ 1/5 & 2/5 & 3/5 & 4/5 & 1 \end{bmatrix}. \tag{16}$$

Evaluate fractions:

$$\mathrm{P}_5 = \begin{bmatrix} 1 & .5 & .3\bar{3} & .25 & .2 \\ .5 & 1 & .6\bar{6} & .5 & .4 \\ .3\bar{3} & .6\bar{6} & 1 & .75 & .6 \\ .25 & .5 & .75 & 1 & .8 \\ .2 & .4 & .6 & .8 & 1 \end{bmatrix}. \tag{17}$$

Using equation (8) to invert P_5 yields:

$$\mathrm{P}_5^{-1} = \begin{bmatrix} 4/3 & -2/3 & 0 & 0 & 0 \\ -2/3 & 32/15 & -6/5 & 0 & 0 \\ 0 & -6/5 & 108/35 & -12/7 & 0 \\ 0 & 0 & -12/7 & 256/63 & -20/9 \\ 0 & 0 & 0 & -20/9 & 25/9 \end{bmatrix}. \tag{18}$$

Factor (18) by 1 over 315 = 5 × 7 × 9 and P_5^{-1} may be written as:

$$P_5^{-1} = \frac{1}{315}\begin{bmatrix} 420 & -210 & 0 & 0 & 0 \\ -210 & 672 & -378 & 0 & 0 \\ 0 & -378 & 972 & -540 & 0 \\ 0 & 0 & -540 & 1280 & -700 \\ 0 & 0 & 0 & -700 & 875 \end{bmatrix}. \tag{19}$$

For $n = 5$, use (19) and (14) where E_{2F,σ_i} is the i-th scaled residual or i-th element of $\mathbf{E_{2F,\sigma}}$ to obtain:

$$\chi_5^2 = [\,E_{2F,\sigma_1} \quad E_{2F,\sigma_2} \quad E_{2F,\sigma_3} \quad E_{2F,\sigma_4} \quad E_{2F,\sigma_5}] \tag{20}$$

$$\times \frac{1}{315}\begin{bmatrix} 420 & -210 & 0 & 0 & 0 \\ -210 & 672 & -378 & 0 & 0 \\ 0 & -378 & 972 & -540 & 0 \\ 0 & 0 & -540 & 1280 & -700 \\ 0 & 0 & 0 & -700 & 875 \end{bmatrix}\begin{bmatrix} E_{2F,\sigma_1} \\ E_{2F,\sigma_2} \\ E_{2F,\sigma_3} \\ E_{2F,\sigma_4} \\ E_{2F,\sigma_5} \end{bmatrix}.$$

Multiply the last two terms of (20)

$$\chi_5^2 = [\,E_{2F,\sigma_1} \quad E_{2F,\sigma_2} \quad E_{2F,\sigma_3} \quad E_{2F,\sigma_4} \quad E_{2F,\sigma_5}] \tag{21}$$

$$\times \frac{1}{315}\begin{bmatrix} 420\,E_{2F,\sigma_1} - 210\,E_{2F,\sigma_2} \\ -210\,E_{2F,\sigma_1} + 672\,E_{2F,\sigma_2} - 378\,E_{2F,\sigma_3} \\ -378\,E_{2F,\sigma_2} + 972\,E_{2F,\sigma_3} - 540\,E_{2F,\sigma_3}E_{2F,\sigma_4} \\ -540\,E_{2F,\sigma_3} + 1280\,E_{2F,\sigma_4} - 700\,E_{2F,\sigma_5} \\ -700\,E_{2F,\sigma_4} + 875\,E_{2F,\sigma_5} \end{bmatrix}.$$

Multiply the two arrays in (21)

$$\chi_5^2 = \frac{1}{315}\left\{\begin{array}{c} 420\,E_{2F,\sigma_1}^2 - 210\,E_{2F,\sigma_1}E_{2F,\sigma_2} + \\ -210\,E_{2F,\sigma_1}E_{2F,\sigma_2} + 672\,E_{2F,\sigma_2}^2 - 378\,E_{2F,\sigma_2}E_{2F,\sigma_3} + \\ -378\,E_{2F,\sigma_2}E_{2F,\sigma_3} + 972\,E_{2F,\sigma_3}^2 - 540\,E_{2F,\sigma_3}E_{2F,\sigma_4} \quad + \\ -540\,E_{2F,\sigma_3}E_{2F,\sigma_4} + 1280\,E_{2F,\sigma_4}^2 - 700\,E_{2F,\sigma_4}E_{2F,\sigma_5} + \\ -700\,E_{2F,\sigma_4}E_{2F,\sigma_5} + 875\,E_{2F,\sigma_5}^2. \end{array}\right. \tag{22}$$

Collect like terms in (22)

$$\chi_5^2 = \frac{1}{315}\left\{\begin{array}{c} 420\,E_{2F,\sigma_1}^2 - 420\,E_{2F,\sigma_1}E_{2F,\sigma_2} + \\ 672\,E_{2F,\sigma_2}^2 - 756\,E_{2F,\sigma_2}E_{2F,\sigma_3} + \\ 972\,E_{2F,\sigma_3}^2 - 1080\,E_{2F,\sigma_3}E_{2F,\sigma_4} + \\ 1280\,E_{2F,\sigma_4}^2 - 1400\,E_{2F,\sigma_4}E_{2F,\sigma_5} + \\ 875\,E_{2F,\sigma_5}^2\,. \end{array}\right. \tag{23}$$

Multiply the terms in (23) and rearrange as

$$\chi_5^2 = \begin{cases} \frac{4}{3} E_{2F,\sigma_1}^2 + \frac{32}{15} E_{2F,\sigma_2}^2 + \frac{108}{35} E_{2F,\sigma_3}^2 + \frac{256}{63} E_{2F,\sigma_4}^2 + \frac{25}{9} E_{2F,\sigma_5}^2 + \\ -\frac{4}{3} E_{2F,\sigma_1} E_{2F,\sigma_2} - \frac{12}{5} E_{2F,\sigma_2} E_{2F,\sigma_3} - \frac{24}{7} E_{2F,\sigma_3} E_{2F,\sigma_4} - \frac{40}{9} E_{2F,\sigma_4} E_{2F,\sigma_5}. \end{cases} \tag{24}$$

3. χ^2 for General n

The previous example of computing χ_5^2 with $n = 5$ motivates a general solution of minimizing χ_n^2 by generalizing the steps used to create (24). To determine χ_n^2 we sum the appropriate terms of the tri-diagonal P^{-1} matrix. Define $\chi_n^2(i,j)$:

$$\chi_n^2(i,j) = \begin{cases} \frac{4i^3}{4i^2-1} E_{2F,\sigma_i}^2, & i = j \text{ and } i < n \\ \frac{n^2}{2n-1} E_{2F,\sigma_n}^2, & i = j = n \\ -2\frac{i(i+1)}{2i+1} E_{2F,\sigma_i} E_{2F,\sigma_{i+1}}, & j = i + 1 \text{ and } i < n \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

Utilizing the terms defined in (25) and (14) we obtain the general expression:

$$\chi_n^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \chi_n^2(i,j)\,. \tag{26}$$

Eliminating the zero terms in (26) yields:

$$\chi_n^2 = \sum_{i=1}^{n} \chi_n^2(i,i) + \sum_{i=1}^{n-1} \chi_n^2(i,i+1). \tag{27}$$

Substitution of (25) in (27) yields:

$$\chi_n^2 = \sum_{i=1}^{n-1} \frac{4i^3}{4i^2-1} E_{2F,\sigma_i}^2 + \frac{n^2}{2n-1} E_{2F,\sigma_n}^2 - 2\sum_{i=1}^{n-1} \frac{i(i+1)}{2i+1} E_{2F,\sigma_i} E_{2F,\sigma_{i+1}}\,. \tag{28}$$

Minimizing (28) with respect to the parameters that define $\mathbf{E_{2F,\sigma}}$ provides the fitted curve while accounting for the correlation between the various gate lengths.

4. Enhancements and Future Work

In order to reduce the computational load fewer than n points may be included in the fit. Excluding points does not affect the tri-diagonal nature of the resulting W^{-1} matrix. One may derive a similar formula to (28) which excludes various points in the fit. Excluding points may decrease the computational load of the fit and if not done wisely it may also decrease the quality of the fit.

One advantage of exclusion of points is to increase the numerical stability of the solution. This is done by improving the condition number of the W matrix. [4] demonstrate that the condition number of the Lehmer matrix is greater than n and less than $4n^2$. Reducing the size of the matrix by excluding the largest points may directly improve the condition number. It is hypothesized that exclusion of points, which may not be the largest, may also result in a better condition number.

Future research may include choosing only a small number of points to use in fitting the curve. Mark Smith-Nelson, personal communication, has suggested only including points distributed more or less uniformly throughout the x-axis region as well as including points where the curvature is highest. One may select these fewer than n points in an optimal manner with an appropriate statistical experimental design.

The analysis used in Prasad, Snyderman, and Walston [5] may be used to provide a similar and direct Monte Carlo estimation of the LANL W matrix. In this way directly confirming the Lehmer functional form of the LANL P correlation matrix.

References

[1] D. H. Lehmer. E 710. The American Mathematical Monthly, 53(2):97, 1946. URL http://www.jstor.org/stable/2305464.

[2] D. H. Lehmer, D. M. Smiley, M. F. Smiley, and J. Williamson. E 710. The American Mathematical Monthly, 53(9):534–535, 1946. ISSN 00029890, 19300972. URL http://www.jstor.org/stable/2305078.

[3] J. W. Lewis. *Inversion of tridiagonal matrices.* Numerische Mathematik, 38(3):333–345, 1982. ISSN 0945-3245. doi: 10.1007/BF01396436. URL http://dx.doi.org/10.1007/BF01396436.

[4] M. Newman and J. Todd. *The evaluation of matrix inversion programs.* Journal of the Society for Industrial and Applied Mathematics, 6(4):466–476, 1958.

[5] M. Prasad, N. Snyderman, and S. Walston. *Neutron time interval distributions with background neutrons.* Technical report, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2016.

[6] S. Walston. *Thoughts on uncertainties in the moments formalism of the statistical theory of fission chains.* Technical report, LLNL-TR-645582, Lawrence Livermore National Laboratory (LLNL), Livermore, CA, 2013.