Propagation of Errors for Matrix Inversion

M. Lefebvre[1], R.K. Keeler[1], R. Sobie[1,2], J. White[1,3]

Abstract

A formula is given for the propagation of errors during matrix inversion. An explicit calculation for a 2×2 matrix using both the formula and a Monte Carlo calculation are compared. A prescription is given to determine when a matrix with uncertain elements is sufficiently nonsingular for the calculation of the covariances of the inverted matrix elements to be reliable.

Key words: matrix inversion; error propagation; branching ratios

1 Introduction

There are many problems that involve solving a set of simultaneous linear equations. In many instances, the coefficients of the variables are uncertain and these uncertainties need to be taken into account in the final solution. The results presented in this paper are general and can be applied to any problem involving the solution of linear equations. Moreover the full covariance matrix is calculated. The use of the covariance formula is illustrated using an example where a number of branching ratios of a given particle were simultaneously measured.

A recent paper reporting results on the decay of the tau lepton (1) used a matrix inversion technique to solve for several branching ratios simultaneously. In that paper, the treatment of the statistical uncertainties assumed only diagonal errors. This paper develops a formula for the covariance of the inverse matrix elements. The treatment follows the propagation of errors formalism (2) for small errors. A full treatment of the covariance matrix is necessary because the off-diagonal errors can be significant.

[1]Department of Physics and Astronomy, P.O. Box 3055 University of Victoria, Victoria, BC, Canada, V8W-3P6

[2]Supported by the Institute of Particle Physics of Canada

[3]Present address: Department of Physics, Carleton University, Ottawa, Ontario, Canada

The fraction of time a particle decays into a given final state is defined to be the branching ratio. The branching ratios of a particle decaying into a set of m possible final states may be determined by finding m selection criteria that have an efficiency for selecting the desired decay channel. The selection criteria are usually not fully efficient so a set of linear equations results,

$$\begin{array}{rcl} \epsilon_{11}B_1 + \epsilon_{12}B_2 + \epsilon_{13}B_3 + \cdots + \epsilon_{1m}B_m & = & f_1 \\ \epsilon_{21}B_1 + \epsilon_{22}B_2 + \epsilon_{23}B_3 + \cdots + \epsilon_{2m}B_m & = & f_2 \\ \epsilon_{31}B_1 + \epsilon_{32}B_3 + \epsilon_{33}B_3 + \cdots + \epsilon_{3m}B_m & = & f_3 \\ & \vdots & \\ \epsilon_{m1}B_1 + \epsilon_{m2}B_2 + \epsilon_{m3}B_3 + \cdots + \epsilon_{mm}B_m & = & f_m \end{array} \tag{1}$$

where B_j is the $j^{\rm th}$ unknown branching ratio, f_i is the fraction of events chosen by the $i^{\rm th}$ selection from a measured sample with an efficiency ϵ_{ij}. The fractions are usually corrected for background using Monte Carlo estimates.

The efficiency is typically calculated by applying the same selections to a Monte Carlo sample of events. Thus, the efficiency is given by

$$\epsilon_{ij} = \frac{n_{ij}^{\rm MC}}{N_j^{\rm MC}} \; , \tag{2}$$

where $n_{ij}^{\rm MC}$ is the number of events of decay channel j selected by selection i and $N_j^{\rm MC}$ is the total number events of type j in the Monte Carlo sample. As can be seen from equation 2, the elements of ϵ_{ij} are positive definite and less than or equal to one. An uncertainty for each efficiency matrix element can be determined from the Monte Carlo statistics.

The simultaneous equations given by equation 1 can be written in matrix form,

$$\epsilon B = f \; . \tag{3}$$

The vector of branching ratios can be solved from the matrix equation,

$$B = \epsilon^{-1} f \; , \tag{4}$$

where ϵ^{-1} is the inverse of the efficiency matrix. The matrix ϵ must be nonsingular in order for the inverse to exist or equivalently the determinant must be nonzero. A singular ϵ signals an ill defined set of selections or event types.

The uncertainty of the branching ratios, B, can be written in terms of the uncertainties on the elements of the matrices f and ϵ^{-1}. The errors for f are determined from the data and the estimates of the background. The errors on ϵ^{-1} can be calculated in terms of the covariances of the elements of ϵ. The error formulae for ϵ^{-1} used in the literature are described in section 2. A more general treatment of the errors taking into account the full covariance matrix is presented in section 3 whilst the detailed derivations are given in Appendix A.

An analytical example based on matrices of order 2×2 is developed in Appendix B and a Monte Carlo study based on the 2×2 case is discussed in sections 4 and 5. The Monte Carlo studies will illustrate that the formula is only realistic when the efficiency matrix is sufficiently far away from being singular. The calculation of the covariances becomes more reliable as the ratio of the value of the determinant of the efficiency matrix, to the value of the uncertainty on the determinant becomes larger. A formula for the uncertainty of a determinant is derived in Appendix C. The paper concludes with some general observations.

2 Calculation of Errors

Errors are often estimated by ignoring the off-diagonal elements of the covariance matrix. This is the correct procedure if the quantities are independent of each other. A calculation of the uncertainty on the branching ratios, assuming the elements of the inverted matrix are statistically independent of each other, yields [4],

$$[\sigma_B]_i^2 = [\epsilon^{-1}]_{ij}^2 [\sigma_f]_j^2 + [\sigma_{\epsilon^{-1}}]_{ij}^2 [f]_j^2 \, , \tag{5}$$

where the uncertainties are denoted by σ and a sum over repeated indices is assumed unless otherwise noted. The errors on the elements of the inverted efficiency matrix, $\sigma_{\epsilon^{-1}}$, are calculated from the known errors on the efficiency matrix.

The uncertainties on the inverse efficiencies, have been determined (see for example reference (1)) by differentiating the matrix equation $\epsilon\epsilon^{-1} = I$ and then applying a matrix analogy to the error propagation formula to yield,

$$[\sigma_{\epsilon^{-1}}]_{ij}^2 = \mid [\epsilon^{-1}]_{im} [\sigma_\epsilon]_{mn} [\epsilon^{-1}]_{nj} \mid^2 \, , \tag{6}$$

[4] Square brackets have been used to separate the superscripts and subscripts used to identify the quantity of interest from any indices and mathematical operations.

which has the same form as a transformation to a new basis (3). Equation 6 is a reasonable approximation but is not correct. It also neglects any correlations between the elements of the inverse matrix, ϵ^{-1}, which can be significant, as shown in the next section.

3 The Covariance of the Elements of ϵ^{-1}

Matrix inversion is a nonlinear operation. It is always possible to write the inverse of a matrix in terms of the matrix of cofactors divided by the determinant (4). One sees explicitly in Appendix C that each element of an inverse matrix has elements of the original matrix in common. Therefore the inverse matrix elements clearly are correlated.

Consider the m^2 matrix elements ϵ_{ij} with uncertainties given in the most general case by the m^4 covariances $\mathrm{cov}(\epsilon_{\alpha\beta}, \epsilon_{ab})$. The inverse matrix elements ϵ_{ij}^{-1}, in general, have covariances $\mathrm{cov}(\epsilon_{\alpha\beta}^{-1}, \epsilon_{ab}^{-1})$, which can be written as,

$$\mathrm{cov}(\epsilon_{\alpha\beta}^{-1}, \epsilon_{ab}^{-1}) = \epsilon_{\alpha i}^{-1}\epsilon_{j\beta}^{-1}\epsilon_{ak}^{-1}\epsilon_{lb}^{-1}\mathrm{cov}(\epsilon_{ij}, \epsilon_{kl}) \; . \tag{7}$$

The full derivation of this equation is given in Appendix A (see equation A-19). The usual case where there are no correlations between the elements of the efficiency matrix (see equations A-20, A-21 and A-22) is given by

$$\mathrm{cov}\,(\epsilon_{ij}, \epsilon_{kl}) = [\sigma_\epsilon]_{ij}^2 \delta_{ik}\delta_{jl} \quad \text{(no summation)}. \tag{8}$$

Hence the full set of covariances of ϵ^{-1} are given by

$$\mathrm{cov}(\epsilon_{\alpha\beta}^{-1}, \epsilon_{ab}^{-1}) = ([\epsilon^{-1}]_{\alpha i}[\epsilon^{-1}]_{ai})[\sigma_\epsilon]_{ij}^2([\epsilon^{-1}]_{j\beta}[\epsilon^{-1}]_{jb}) \; , \tag{9}$$

where there is no sum in this case over repeated indices inside the parentheses. The variance of an element of the inverse efficiency matrix can be written as,

$$[\sigma_{\epsilon^{-1}}]_{\alpha\beta}^2 \equiv \mathrm{cov}(\epsilon_{\alpha\beta}^{-1}, \epsilon_{\alpha\beta}^{-1}) = [\epsilon^{-1}]_{\alpha i}^2[\sigma_\epsilon]_{ij}^2[\epsilon^{-1}]_{j\beta}^2 \; . \tag{10}$$

This equation is exact and replaces the approximate equation 6. Note that each term of equation 10 is squared before making the sum whereas in equation 6 the sum is done first.

The complete expression for the uncertainties on the branching ratios can then be calculated from,

$$\mathrm{cov}(B_i, B_j) = f_\alpha f_\beta \mathrm{cov}(\epsilon_{i\alpha}^{-1}, \epsilon_{j\beta}^{-1}) + \epsilon_{ik}^{-1}\epsilon_{jl}^{-1}\mathrm{cov}(f_k, f_l) \; , \tag{11}$$

where the measured fractions of events are often uncorrelated,

$$\mathrm{cov}(f_k, f_l) = [\sigma_f]_k^2 \delta_{kl} \quad \text{(no summation)}. \tag{12}$$

Equation 11 is a generalization of equation 5, including all of the covariances.

4 A Monte Carlo Study - Part 1

The properties of equation 9 have been studied using a Monte Carlo simulation of a 2×2 matrix. Relatively small errors have been used to satisfy the small error approximation used to derive the error propagation formula (2). The results of the Monte Carlo calculations directly can be compared to the analytic formulae developed in Appendix B.

Given an initial efficiency matrix ϵ, and the variance on each element, $[\sigma_\epsilon]_{ij}^2$, the m^{th} random instance of the matrix is generated by

$$[\epsilon_m]_{ij} = \epsilon_{ij} + \Gamma_m \cdot [\sigma_\epsilon]_{ij} \ , \tag{13}$$

where Γ_m is a normally distributed pseudorandom deviate with a mean of zero and a standard deviation of one. A set of N matrices are created, are then inverted. The covariances are calculated from

$$\mathrm{cov}(\epsilon^{-1}_{\alpha\beta}, \epsilon^{-1}_{ab}) = \langle \epsilon^{-1}_{\alpha\beta} \ \epsilon^{-1}_{ab} \rangle - \langle \epsilon^{-1}_{\alpha\beta} \rangle \langle \epsilon^{-1}_{ab} \rangle \ , \tag{14}$$

where $\langle \ \ \rangle$ is the mean of the quantity enclosed in the brackets.

A sample of $N = 10,000$ instances of the matrix $\epsilon \pm \sigma_\epsilon$ where,

$$\epsilon = \begin{pmatrix} 0.700 & 0.200 \\ 0.400 & 0.600 \end{pmatrix} \pm \begin{pmatrix} 0.007 & 0.002 \\ 0.004 & 0.006 \end{pmatrix} , \tag{15}$$

were generated with each element of ϵ normally distributed about the central values with the one percent errors, σ_ϵ, indicated. The inverse of the matrix ϵ is given by

$$\epsilon^{-1} = \begin{pmatrix} 1.765 & -0.588 \\ -1.177 & 2.059 \end{pmatrix} , \tag{16}$$

and the determinant, $\det \epsilon = 0.340$ The covariances can be calculated using only the inverse matrix, ϵ^{-1}, and the error matrix, σ_ϵ. The numerical values for the 16 covariances of ϵ^{-1}, calculated from the Monte Carlo and analytically using equation 9, are given in Table 1. The symmetry of the covariance

Covariance Matrix ϵ^{-1}					
Monte Carlo $\times(10^{-4})$		Analytic $\times(10^{-4})$		Fractional Difference	
$\mathrm{cov}(\epsilon_{11}^{-1}, \epsilon_{ij}^{-1})$					
5.231	−2.312	5.269	−2.245	−0.007	0.029
−4.604	2.496	−4.490	2.514	0.025	−0.007
$\mathrm{cov}(\epsilon_{12}^{-1}, \epsilon_{ij}^{-1})$					
−2.312	1.643	−2.245	1.603	0.029	0.024
2.567	−2.642	2.514	−2.619	0.020	0.009
$\mathrm{cov}(\epsilon_{21}^{-1}, \epsilon_{ij}^{-1})$					
−4.604	2.567	−4.490	2.514	0.025	0.020
6.511	−5.229	6.413	−5.238	0.025	0.020
$\mathrm{cov}(\epsilon_{22}^{-1}, \epsilon_{ij}^{-1})$					
2.496	−2.642	2.514	−2.619	−0.007	0.009
−5.229	6.971	−5.238	7.172	−0.002	−0.029

Table 1: The covariance matrix for each element of ϵ^{-1} calculated by the Monte Carlo and calculated using the analytic formula. The last column is the fractional difference between the Monte Carlo and analytic calculation.

operation implies there should be 10 unique numbers as is apparent from the table.

Figure 1 shows a contour plot of the element ϵ_{12}^{-1} versus ϵ_{11}^{-1} generated from the Monte Carlo. The contours are approximate lines of constant probability density. The entries are distributed around the values of the calculated inverse matrix elements and are clearly correlated. The slope of the error ellipses has been calculated by substituting the analytical values of the covariances from Table 1 into a formula derived from reference (2). The calculated slope is in good agreement with the slope of the major axis of the error ellipses.

Recall that the efficiency matrix must be nonsingular in order to invert it. In the Monte Carlo calculation the individual elements of ϵ are varied and hence, if the variations are large enough, it is possible for the determinant to become zero. Figure 2a shows a histogram of the Monte Carlo calculation of the determinant. The mean value is 0.340 as expected with a root mean square deviation of 0.006 . The shape of the distribution is normal, see figure 2, also as expected given the determinant is the sum of products of

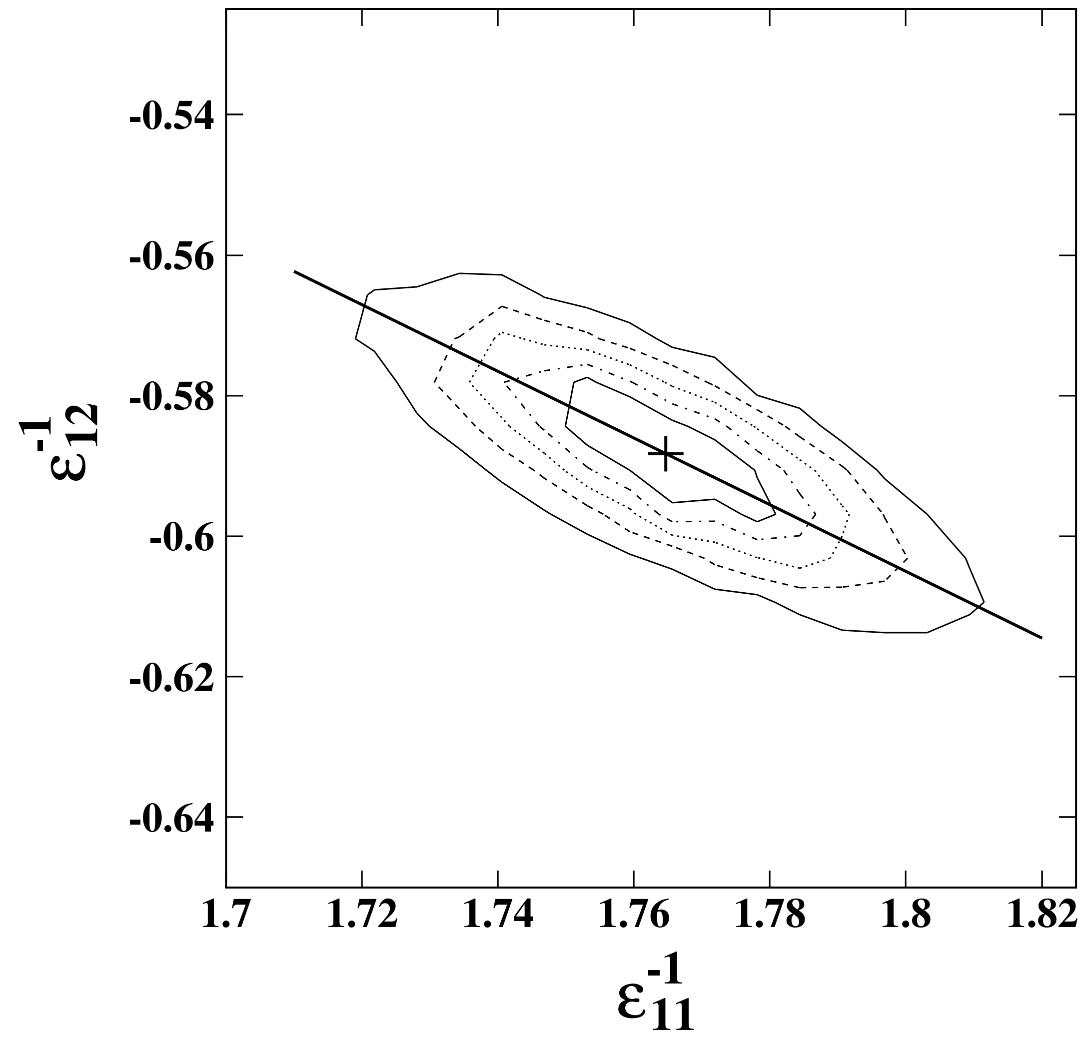

Figure 1: The 10,000 elements ϵ_{12}^{-1} and ϵ_{11}^{-1} generated in the Monte Carlo simulation form a two dimensional distribution shown here as a contour plot. The contours are at the level of 50, 100, 150, 200 and 250 counts per bin (each bin is $6.25 \times 10^{-3} \times 6.25 \times 10^{-3}$). The correlation between the two inverse matrix elements is clear. The data are centred around the exact calculation of the inverse, shown by the large "+" sign. The slope of the line is the theoretical calculation of the correlation coefficient. It is in good agreement with the slope of the axis of the approximately elliptical contours.

normally distributed quantities. In this particular case the mean value of the determinant is 56.7 standard deviations from zero. Increasing the errors or reducing the determinant will both produce a higher likelihood of the determinant fluctuating nearer to zero.

As noted earlier the inverse matrix elements all have a common factor of $1/|\epsilon|$. Therefore values of the determinant near zero produce very large values for the inverse matrix elements. Moreover, the reciprocal of a normal distribution is not normally distributed but is instead asymmetric with a tail towards large values. In figure 2b one sees that any tails are insignificant for determinants many standard deviations away from zero.

Figure 3 shows an identical analysis for the matrix,

$$\epsilon' = \begin{pmatrix} 0.400 & 0.500 \\ 0.400 & 0.600 \end{pmatrix} \pm \begin{pmatrix} 0.004 & 0.005 \\ 0.004 & 0.006 \end{pmatrix} . \tag{17}$$

In this case $\det \epsilon' = 0.040$ and its root mean square deviation, due to the one percent errors, is 0.004 . Now the distribution of the reciprocal determinants, seen in Fugure 3b, is clearly skewed. Table 2 lists the values of the covariance matrix calculated for each element of the inverse matrix, ϵ'^{-1}, using the Monte Carlo method, equation 14 and the analytic expression, equation 9.

In Table 1 the percent difference between the Monte Carlo and the analytic formula has a mean value of approximately one percent averaged over the 10 independent values of the covariance. This is consistent with the precision expected for a 10,000 event Monte Carlo and demonstrates that the analytic expression reproduces the covariances well for the far from singular case. However, the same mean for the values in Table 2 is eleven percent. Moreover, the magnitude of the Monte Carlo value is larger than the analytic value for each covariance. The analytic expression becomes less accurate and no longer appropriate to use as the matrix becomes closer to being singular.

5 A Monte Carlo Study - Part 2

In order to explore the range of uncertainty on how well the analytic expression models the covariances, the Monte Carlo simulation described in Part 1 was modified. Instead of specifying a particular 2×2 matrix for analysis, 5000 2×2 matrices were generated by choosing the elements of each matrix from

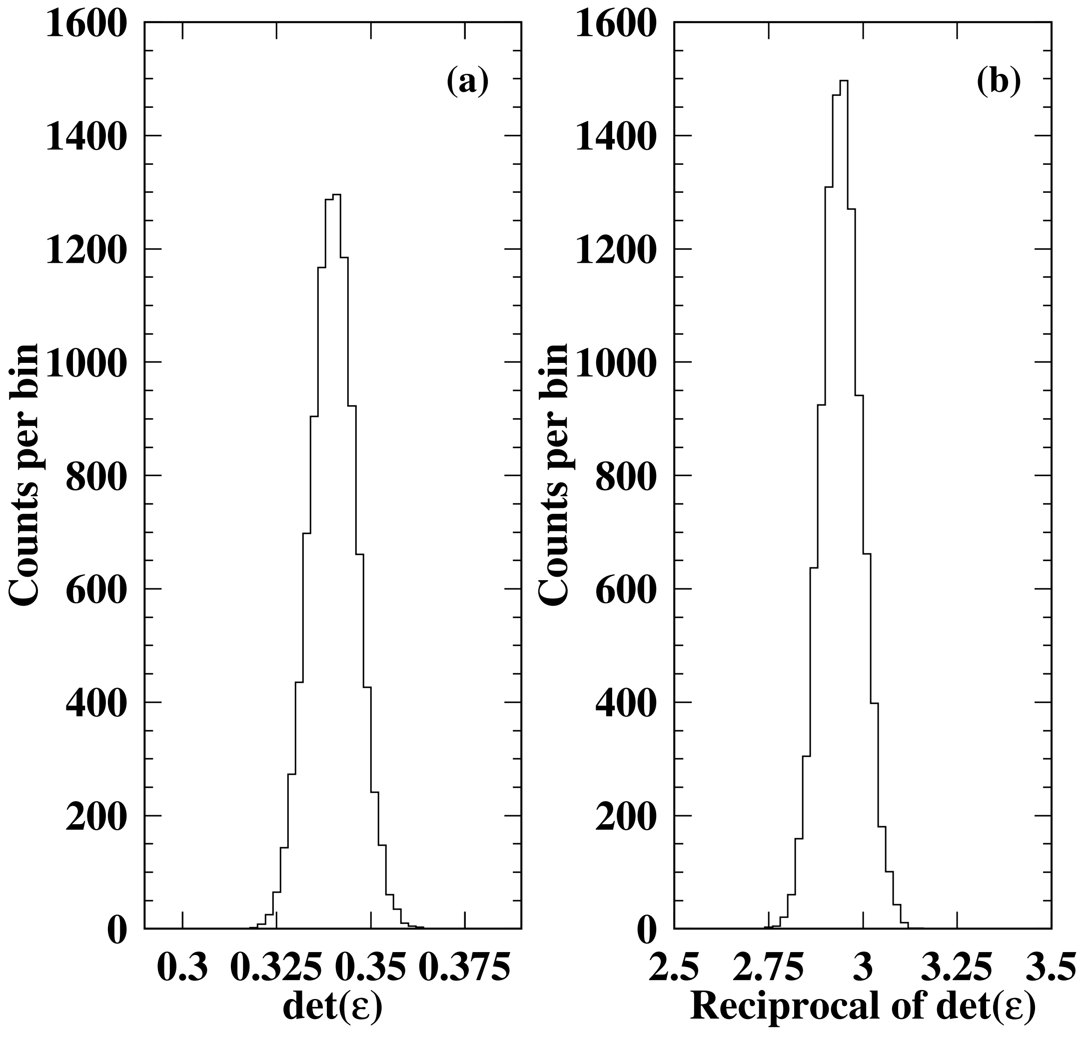

Figure 2: The histogram labelled (a) shows the determinant of the matrix ϵ described in the text. The values of the determinant are clearly not near zero and consequently the histogram of the reciprocal of the determinant shown in (b) is well behaved.

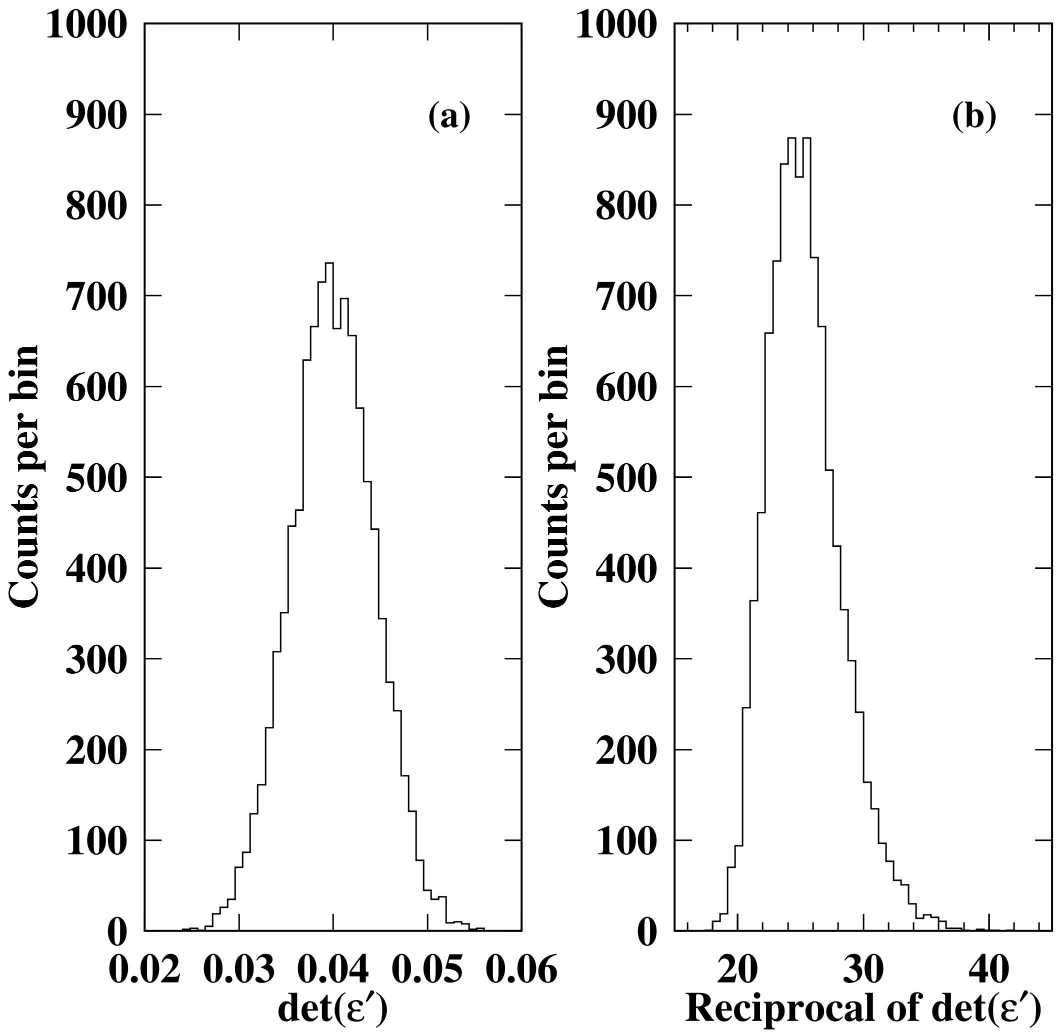

Figure 3: The histogram labelled (a) shows the determinant of the matrix ϵ described in the text. The values of the determinant approach zero and consequently an asymmetry is observed in the histogram, labelled (b), of the reciprocal distribution.

Covariance Matrix ϵ'^{-1}					
Monte Carlo $\times(10^{-4})$		Analytic $\times(10^{-4})$		Fractional Difference	
$\text{cov}(\epsilon'^{-1}_{11}, \epsilon'^{-1}_{ij})$					
2.812	−2.554	2.498	−2.269	0.112	0.112
−2.040	1.855	−1.815	1.650	0.110	0.110
$\text{cov}(\epsilon'^{-1}_{12}, \epsilon'^{-1}_{ij})$					
−2.554	2.337	−2.269	2.078	0.112	0.111
1.855	−1.699	1.650	−1.513	0.110	0.110
$\text{cov}(\epsilon'^{-1}_{21}, \epsilon'^{-1}_{ij})$					
−2.040	1.855	−1.815	1.650	0.110	0.110
1.492	−1.358	1.330	−1.210	0.109	0.109
$\text{cov}(\epsilon'^{-1}_{22}, \epsilon'^{-1}_{ij})$					
1.855	−1.699	1.650	−1.513	0.110	0.111
−1.358	1.244	−1.210	1.110	0.109	0.108

Table 2: The covariance matrix for each element of ϵ'^{-1} calculated by the Monte Carlo and calculated using the analytic formula. The last column is the fractional difference between the Monte Carlo and analytic calculation.

a uniform distribution in the interval $[0, 1]$. Each of the randomly generated matrices was then analyzed using exactly the same procedure previously discussed. One percent errors were used as before.

None of the matrix inversions actually failed. However, many cases involved determinants close to zero. Figure 4(a) shows the fractional difference between the Monte Carlo and the analytic calculation for $\text{cov}(\epsilon^{-1}_{11}, \epsilon^{-1}_{21})$,

$$\frac{\text{cov}(\epsilon^{-1}_{11}, \epsilon^{-1}_{21})_{\text{MC}} - \text{cov}(\epsilon^{-1}_{11}, \epsilon^{-1}_{21})_{\text{Analytic}}}{\text{cov}(\epsilon^{-1}_{11}, \epsilon^{-1}_{21})_{\text{MC}}}, \tag{18}$$

plotted against the determinant of ϵ divided by its uncertainty, $\det \epsilon/\sigma_{|\epsilon|}$. The number of sigmas the determinant may be from zero can be shown to be ± 70.7 for one percent errors. The fractional differences between the Monte Carlo and analytic evaluations are distributed around approximately zero for large numbers of sigma. However, near 10 sigmas the Monte Carlo calculated covariance becomes very large and the distribution becomes centred around one. Under the scatter plot is figure 4(b). This is a plot of the mean value of the fractional difference plotted against the number of sigmas the determinant is away from zero. The error bars are the root mean square deviation of the distribution. The means for $-10. < \det \epsilon/\sigma_{|\epsilon|} < 10.$ are not plotted. The largest deviation, for the bin 10–20, is less than four percent. The sigmas are approximately two percent. Hence, the analytic formula does a very good job of estimating the covariance as long as the determinant of the matrix is at least 10 sigmas from zero. A simple expression to calculate the error on a determinant is given in Appendix C.

6 Conclusions

The covariances of inverse matrix elements are nonzero in general. Therefore, the propagation of errors for formulae that depend on inverted matrices requires using the covariances of the inverse matrix elements. A concise formula is developed in the small error limit for the covariance of the inverted matrix elements. It is shown to work for nonsingular matrices. In particular the determinant must not become zero when the matrix elements are allowed to vary within their uncertainties.

An attempt was made to study how well the formula estimates the covariances for a random set of 2×2 matrices with positive elements between zero and one. On the average the formula was accurate to better than four percent with a root mean square deviation equal to two percent for matrices that have determinants more than ten sigmas from zero.

7 Acknowledgements

This project was supported by grants from the Natural Science and Engineering Research Council of Canada.

References

[1] The OPAL Collaboration, K. Ackerstaff et al., *Eur. Phys. J.* **C4**(1998)193–206.

[2] The Particle Data Group, C. Caso et al., Section 29, *Eur. Phys. J.* **C3**(1998)1–794.

[3] L. Lyons, Statistics for Nuclear and Particle Physics, Cambridge University Press, 1989.

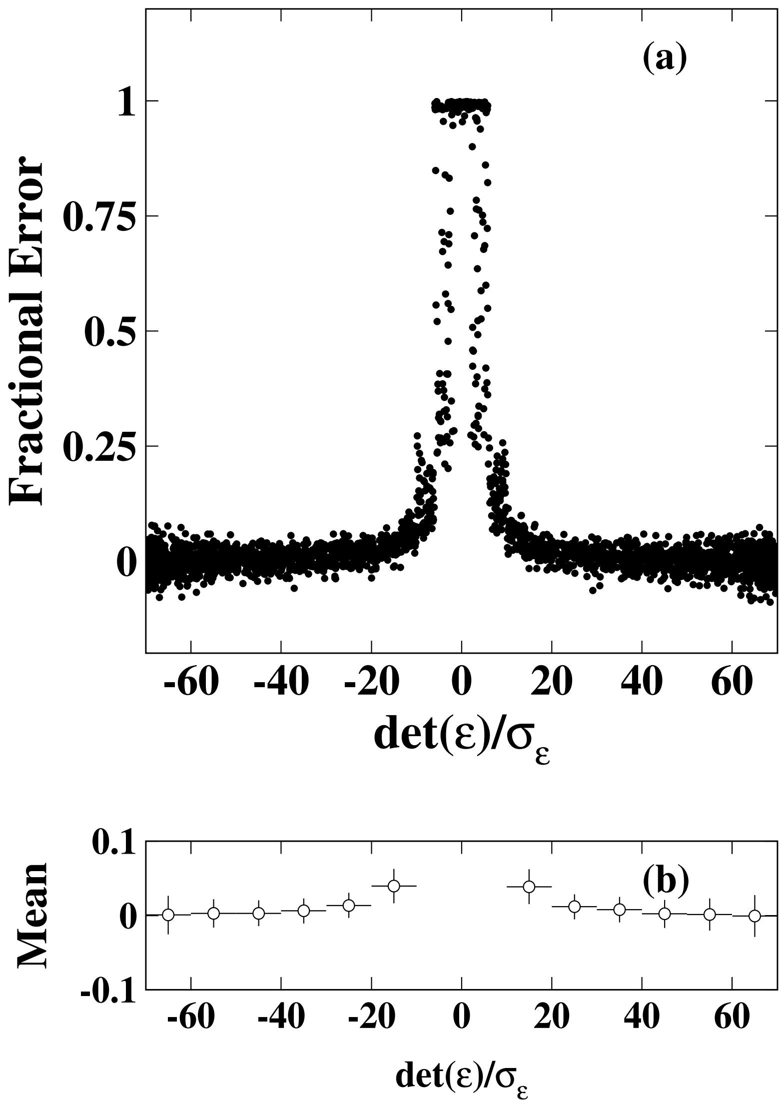

Figure 4: The fractional difference between the Monte Carlo and analytic calculations as a function of $\det \epsilon/\sigma_{|\epsilon|}$ for $\mathrm{cov}(\epsilon_{11}^{-1}, \epsilon_{21}^{-1})$ is plotted in part (a). The fractional difference is normalized by the Monte Carlo value. In part (b) the mean and sigma are plotted for slices of distribution (a). There are 14 slices of $\det \epsilon/\sigma_{|\epsilon|}$ between -70 and 70. The two centre slices are not plotted.

[4] S. Lipschutz Theory and Problems of Linear Algebra, Schaum's Outline Series, 1968.

A Derivation of the Covariance

The formal solution to the system of real linear equations,

$$Y_i = A_{ij} X_j \qquad i, j = 1 \ldots, n \tag{A-1}$$

where X_i are the unknowns and A is a real $n \times n$ matrix, is given by

$$X_i = A^{-1}_{ij} Y_j \, . \tag{A-2}$$

A sum over repeated indices is assumed unless otherwise noted. The uncertainties on the elements of A are assumed to be known and given by

$$\mathrm{cov}(A_{ij}, A_{\alpha\beta}) \, , \tag{A-3}$$

the covariance of A_{ij} and $A_{\alpha\beta}$. If each element of A is an independent random quantity, then only the terms $\mathrm{cov}(A_{ij}, A_{ij})$ are nonzero. In other words, each element of the matrix will have an associated variance.

Similarly the uncertainties on the elements of Y are assumed to be known and given by $\mathrm{cov}(Y_i, Y_j)$. Again, if the elements of Y are independent, the off-diagonal terms will be zero.

The most general form of the covariance for the X_i is given by the error propagation formula,

$$\mathrm{cov}(X_i, X_j) = \frac{\partial X_i}{\partial A^{-1}_{\alpha\beta}} \frac{\partial X_j}{\partial A^{-1}_{ab}} \mathrm{cov}(A^{-1}_{\alpha\beta}, A^{-1}_{ab}) + \frac{\partial X_i}{\partial Y_k} \frac{\partial X_j}{\partial Y_l} \mathrm{cov}(Y_k, Y_l) \, . \tag{A-4}$$

The partial derivatives are given by:

$$\frac{\partial X_i}{\partial A^{-1}_{\alpha\beta}} = \delta_{\alpha i} Y_\beta \, , \tag{A-5}$$

$$\frac{\partial X_j}{\partial A^{-1}_{ab}} = \delta_{aj} Y_b \, , \tag{A-6}$$

$$\frac{\partial X_i}{\partial Y_k} = A^{-1}_{ik} \, , \tag{A-7}$$

$$\frac{\partial X_j}{\partial Y_l} = A^{-1}_{jl} \, , \tag{A-8}$$

where δ_{ij} is the Kronecker delta. Therefore, after relabelling summed-over indices, equation A-4 becomes

$$\mathrm{cov}(X_i, X_j) = Y_\alpha Y_\beta \mathrm{cov}(A^{-1}_{i\alpha}, A^{-1}_{j\beta}) + A^{-1}_{ik} A^{-1}_{jl} \mathrm{cov}(Y_k, Y_l) \ . \tag{A-9}$$

The only unknown quantity is $\mathrm{cov}(A^{-1}_{i\alpha}, A^{-1}_{j\beta})$.

The inverse matrix elements can be considered as functions of the original matrix elements,

$$A^{-1}_{\alpha\beta} = A^{-1}_{\alpha\beta}(A_{ij}) \ . \tag{A-10}$$

The error propagation formula yields

$$\mathrm{cov}\left(A^{-1}_{\alpha\beta}, A^{-1}_{ab}\right) = \frac{\partial A^{-1}_{\alpha\beta}}{\partial A_{ij}} \frac{\partial A^{-1}_{ab}}{\partial A_{kl}} \mathrm{cov}\left(A_{ij}, A_{kl}\right) \ . \tag{A-11}$$

Hence the terms $\partial A^{-1}_{\alpha\beta}/\partial A_{ij}$ are required.

Consider the identity,

$$A^{-1}_{ij} A_{jk} = \delta_{ik} \ . \tag{A-12}$$

Taking the derivative with respect to $A_{\alpha\beta}$ yields,

$$\frac{d\delta_{ik}}{dA_{\alpha\beta}} = A^{-1}_{ij} \frac{\partial A_{jk}}{\partial A_{\alpha\beta}} + A_{jk} \frac{\partial A^{-1}_{ij}}{\partial A^{-1}_{ab}} \frac{\partial A^{-1}_{ab}}{\partial A_{\alpha\beta}} = 0 \ . \tag{A-13}$$

Making the substitutions,

$$\frac{\partial A_{jk}}{\partial A_{\alpha\beta}} = \delta_{j\alpha}\delta_{k\beta} \tag{A-14}$$

and

$$\frac{\partial A^{-1}_{ij}}{\partial A^{-1}_{ab}} = \delta_{ia}\delta_{jb} \tag{A-15}$$

yields

$$A^{-1}_{i\alpha}\delta_{k\beta} + A_{bk} \frac{\partial A^{-1}_{ib}}{\partial A_{\alpha\beta}} = 0 \ . \tag{A-16}$$

Contracting with $A^{-1}_{k\gamma}$ and using $A_{bk}A^{-1}_{k\gamma} = \delta_{b\gamma}$ gives

$$\frac{\partial A^{-1}_{i\gamma}}{\partial A_{\alpha\beta}} = -A^{-1}_{i\alpha} A^{-1}_{\beta\gamma} \ , \tag{A-17}$$

which can be relabelled as

$$\frac{\partial A^{-1}_{\alpha\beta}}{\partial A_{ij}} = -A^{-1}_{\alpha i} A^{-1}_{j\beta} \ . \tag{A-18}$$

Hence, equation A-11 becomes

$$\mathrm{cov}(A^{-1}_{\alpha\beta}, A^{-1}_{ab}) = A^{-1}_{\alpha i} A^{-1}_{j\beta} A^{-1}_{ak} A^{-1}_{lb} \mathrm{cov}(A_{ij}, A_{kl})\,. \tag{A-19}$$

This is the fundamental formula of this paper.

An important case is when matrix elements A_{ij} are uncorrelated. If each element of A_{ij} has associated with it a variance $[\sigma_A]^2_{ij}$ then

$$\mathrm{cov}\,(A_{ij}, A_{kl}) = [\sigma_A]^2_{ij}\delta_{ik}\delta_{jl} \quad \text{(no summation).} \tag{A-20}$$

Note that a corrected version of equation 6 now can be calculated as the variance $[\sigma_{A^{-1}}]^2_{\alpha\beta}$ for each element of $A^{-1}_{\alpha\beta}$,

$$[\sigma_{A^{-1}}]^2_{\alpha\beta} \equiv \mathrm{cov}(A^{-1}_{\alpha\beta}, A^{-1}_{\alpha\beta}) = [A^{-1}]^2_{\alpha i}[\sigma_A]^2_{ij}[A^{-1}]^2_{j\beta}\,. \tag{A-21}$$

Finally, the full covariance formula for uncorrelated A_{ij} is given by

$$\mathrm{cov}(A^{-1}_{\alpha\beta}, A^{-1}_{ab}) = (A^{-1}_{\alpha i} A^{-1}_{ai})[\sigma_A]^2_{ij}(A^{-1}_{j\beta} A^{-1}_{jb}) \tag{A-22}$$

where there is no sum inside the parentheses. For an $n \times n$ matrix there are n^4 covariances. From symmetry of the covariance operation one sees that there are $n^2(n^2+1)/2$ independent covariance terms of which n^2 are diagonal, i.e. variances, and $n^2(n^2-1)/2$ are “off-diagonal”.

B An Explicit Two Dimensional Example

The two dimensional case is very instructive because it is simple enough to derive analytically. Consider

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{B-1}$$

where

$$A^{-1} = \frac{1}{\mid A \mid} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{B-2}$$

and $\mid A \mid = ad - bc$ is the determinant of A. The matrix elements of A have uncertainties given by

$$\sigma_A = \begin{pmatrix} \sigma_a & \sigma_b \\ \sigma_c & \sigma_d \end{pmatrix}. \tag{B-3}$$

The uncertainty on $\alpha = A^{-1}_{11}$, for example, can be calculated using the error propagation formula,

$$\sigma^2_\alpha = (\frac{\partial\alpha}{\partial a})^2\sigma^2_a + (\frac{\partial\alpha}{\partial b})^2\sigma^2_b + (\frac{\partial\alpha}{\partial c})^2\sigma^2_c + (\frac{\partial\alpha}{\partial d})^2\sigma^2_d\,. \tag{B-4}$$

The partial derivatives can be calculated from equation B-2 as for example,

$$\frac{\partial\alpha}{\partial a} = \frac{\partial}{\partial a}\left(\frac{d}{ad-bc}\right) = \frac{\partial}{\partial a}\left(\frac{-d^2}{(ad-bc)^2}\right) = -\alpha^2\,. \tag{B-5}$$

For completeness the required derivatives are:

$$\frac{\partial\alpha}{\partial a} = -\alpha^2\,,\ \frac{\partial\alpha}{\partial b} = -\alpha\gamma\,,\ \frac{\partial\alpha}{\partial c} = -\alpha\beta\,,\ \frac{\partial\alpha}{\partial d} = -\beta\gamma\,, \tag{B-6}$$

and therefore

$$\sigma^2_\alpha = \alpha^4\sigma^2_a + \alpha^2\gamma^2\sigma^2_b + \alpha^2\beta^2\sigma^2_c + \beta^2\gamma^2\sigma^2_d\,. \tag{B-7}$$

This result can be compared directly with the results from equation A-21, where $\alpha = \beta = 1$,

$$\mathrm{cov}(A^{-1}_{11}, A^{-1}_{11}) \equiv [\sigma_{A^{-1}}]^2_{11} = [A^{-1}]^2_{1i}[\sigma_A]^2_{ij}[A^{-1}]^2_{j1}\,, \tag{B-8}$$

which on substitution for the elements of A^{-1} and σ_A yields equation B-7.

Similarly an explicit example of an "off-diagonal" covariance can be calculated,

$$\mathrm{cov}(\alpha,\beta) = \frac{\partial\alpha}{\partial a}\frac{\partial\beta}{\partial a}\sigma^2_a + \frac{\partial\alpha}{\partial b}\frac{\partial\beta}{\partial b}\sigma^2_b + \frac{\partial\alpha}{\partial c}\frac{\partial\beta}{\partial c}\sigma^2_c + \frac{\partial\alpha}{\partial d}\frac{\partial\beta}{\partial d}\sigma^2_d\,, \tag{B-9}$$

where the partial derivatives with respect to beta are given by,

$$\frac{\partial\beta}{\partial a} = -\alpha\beta\,,\ \frac{\partial\beta}{\partial b} = -\alpha\delta\,,\ \frac{\partial\beta}{\partial c} = -\beta^2\,,\ \frac{\partial\beta}{\partial d} = -\beta\delta\,. \tag{B-10}$$

Therefore the covariance of α and β is given by,

$$\mathrm{cov}(\alpha,\beta) = \alpha^3\beta\sigma^2_a + \alpha^2\gamma\delta\sigma^2_b + \alpha\beta^3\sigma^2_c + \beta^3\gamma\delta\sigma^2_d\,. \tag{B-11}$$

This result is identical with equation A-22 where $\alpha = \beta = 1$, $a = 1$ and $b = 2$. Note, that in general, the covariance is not zero.

C Covariances of the Determinant

Let A be an $n \times n$ matrix with elements A_{ij}. The determinant is given by,

$$\det A = |A| = \epsilon_{\alpha_1 \alpha_2 \ldots \alpha_n} A_{1\alpha_1} A_{2\alpha_2} \ldots A_{n\alpha_n} \, , \tag{C-1}$$

where $\epsilon_{\alpha_1 \alpha_2 \ldots \alpha_n} = 1$ for cyclic permutations of α_i, a factor of -1 is applied each time any two indices are exchanged, hence $\epsilon_{\alpha_1 \alpha_2 \ldots \alpha_n}$ is zero if any two indices are the same. This can be rewritten by defining new matrices by removing from A the elements of its i^{th} row and j^{th} column. The determinant of the remaining $(n-1) \times (n-1)$ matrix is called (4) the minor of A_{ij}. If the minor of A_{ij} is denoted by M_{ij} then the cofactor of A_{ij} is given by,

$$C_{ij} = (-1)^{i+j} M_{ij}. \tag{C-2}$$

The matrix of all cofactors is C. The determinant of A can now be written as,

$$\det A = A_{ij} C_{ij} = A_{ji} C_{ji}, \tag{C-3}$$

for i fixed.

The error on $\det A$ is given by,

$$\sigma^2_{|A|} = \frac{\partial \det A}{\partial A_{ab}} \frac{\partial \det A}{\partial A_{\alpha\beta}} \mathrm{cov}(A_{ab}, A_{\alpha\beta}) \tag{C-4}$$

From equation C-3 one gets,

$$\frac{\partial \det A}{\partial A_{ab}} = C_{ab}, \tag{C-5}$$

since $i = a$ is fixed and C_{ij} is independent of A_{ab} for all j. Therefore, one obtains the very compact result,

$$\sigma^2_{|A|} = C_{ab} C_{\alpha\beta} \mathrm{cov}(A_{ab}, A_{\alpha\beta}). \tag{C-6}$$