

GENERATION OF CROSS-CORRELATED RANDOM VARIABLES IN THE TYPE B EVALUATION OF UNCERTAINTY MEASUREMENT

Paweł Majda, Dominik Rozkrut

S u m m a r y

This study presents the algorithm for generating cross-correlated random variables. The algorithm uses the Cholesky decomposition. We discuss the advantages and disadvantages of the algorithm in terms of its application in the type B evaluation of measurement uncertainty. This study considered the determination of uncertainty in measuring the stiffness of machine tools for input variables suspected of having a significant degree of correlation. A strict solution was compared with the solution obtained with the use of the Monte Carlo method.

Keywords: uncertainty measurement, correlated variables, stiffness of machine tool

Zastosowanie generatora zmiennych losowych skorelowanych wzajemnie w analizie niepewności pomiaru metodą typu B

S t r e s z c z e n i e

W pracy przedstawiono algorytm generowania zmiennych losowych skorelowanych. Podstawą algorytmu jest dekompozycja macierzy Choleskiego. Prowadzono dyskusję wad i zalet jego użycia w odniesieniu do praktycznego zastosowania w analizie niepewności pomiarów metodą typu B. Rozważano możliwość wyznaczania niepewności wskaźnika sztywności maszyn technologicznych dla zmiennych wejściowych o znacznym stopniu korelacji. Porównano wyniki rozwiązania ścisłego z rozwiązaniem uzyskanym metodą Monte Carlo.

Słowa kluczowe: niepewność pomiarów, zmienne skorelowane, sztywność maszyn technologicznych

1. Introduction

This article was inspired by our observations made during the measurements of stiffness of machine tools at the Institute of Mechanical Technology, the West Pomeranian University of Technology. We noted that the input data obtained at various time intervals (e.g. one-day intervals) and used for the determination of the stiffness coefficient of a machine are subject to random fluctuations. The measurements concerned force and displacement caused by this force. Their

Address: Paweł MAJDA, DSc Eng., Szczecin University of Technology, Faculty of Mechanical Engineering, Institute of Manufacturing Engineering, 19 Piastów Av., 70-310 Szczecin, e-mail: Paweł.Majda@ps.pl, Dominik ROZKRUT, PhD, Szczecin University, Faculty of Economics and Management, Institute of Econometrics and Statistics, 64 Mickiewicza St., 71-101 Szczecin, e-mail: dominic@univ.szczecin.pl

analysis showed a significant cross-correlation which was probably caused by temperature fluctuations affecting the metrology equipment, i.e. signal conditioning and data acquisition systems. As the data acquisition and displacement forces used the same data acquisition card (DAQ), a question arose whether the correlation resulting from the previously described causes has a significant effect on measurement uncertainty.

Research stiffness of machine elements are often carried out without the measurement uncertainty analysis [1, 2]. Before planning an experiment aimed at experimental determination of the correlation of variables and potential detection and determination of corrections, it was decided to perform a type B evaluation of uncertainty of stiffness [3]. This approach saves time and cost of research. The problem boils down to a proper preparation of simulation before performing painstaking measurements. The solution is meant to establish whether it is important and reasonable to take into account correlated variables in the evaluation of uncertainty in the determination of stiffness of machine tools.

2. Model Generation of correlated random variables – Cholesky decomposition

The Cholesky decomposition can be used if there is a need to generate several sequences of correlated random variables [4, 5]. First, we must a priori assume the correlation coefficient between the variables and arrange them in a symmetric positive-definite matrix. For example, for three variables (of course there may be more) we have the following matrix:

$$\mathbf{K} = \begin{bmatrix} 1 & 0,2 & 0,5 \\ & 1 & 0,9 \\ sym & & 1 \end{bmatrix} \quad (1)$$

Elements on the main diagonal are equal one, while elements outside the diagonal are the assumed Pearson correlation coefficients between the variables. The next step involves the Cholesky decomposition of the matrix K into the following product:

$$\mathbf{K} = \mathbf{L}^T \cdot \mathbf{L} \quad (2)$$

where: L is a triangular matrix. After the determination of the matrix L, it is sufficient to calculate:

$$\mathbf{S} = \mathbf{Z} \cdot \mathbf{L} \quad (3)$$

where: S – matrix of correlated random variables, Z – matrix of uncorrelated random variables, L – Cholesky decomposition triangular matrix.

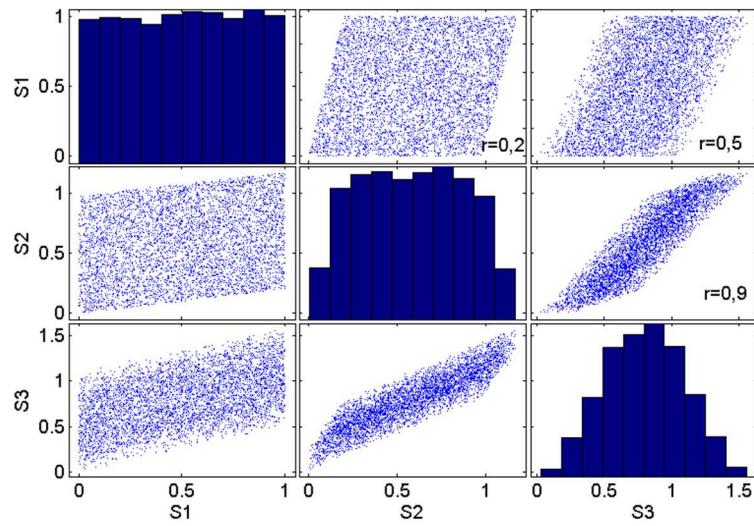


Fig. 1. Matrix of scatter plots by group with the use of the Cholesky decomposition
for input data with a rectangular distribution

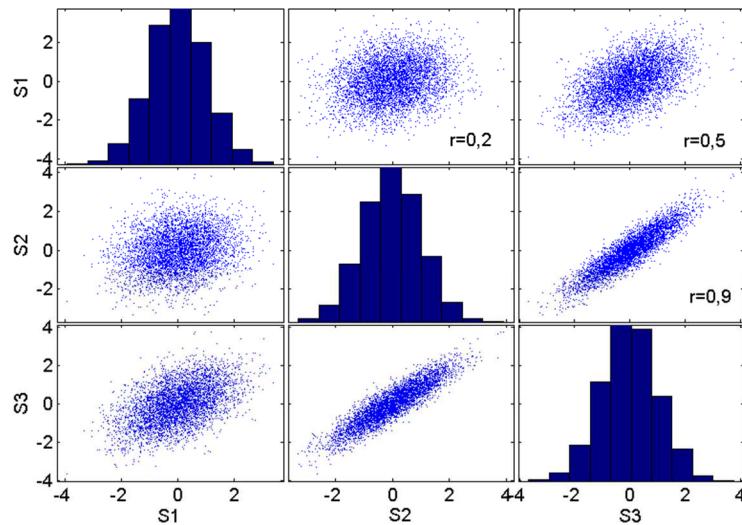


Fig. 2. Matrix of scatter plots by group with the use of Cholesky decomposition
for input data with a normal distribution $N(0,1)$

For example, Fig. 1 and 2 show the visualization of the effects of generating the elements of matrix S for two cases. In the first case, the input data for the matrix Z contain generated pseudo-random sequences of random variables with a rectangular distribution in the range of 0 to 1, and in the second case a variable with a normal distribution N (0,1). The population size was assumed to be 5000. Elements on the main diagonal are histograms of variables obtained at the output of the correlated variables generator with the use of Cholesky decomposition matrix K.

In the type B evaluation of uncertainty measurement one must make assumptions about the type of probability distribution of measurement results [3]. This is particularly important if we do not know the type of probability distribution of measurement errors. In such a situation it is usually assumed that the variables can be characterized by a rectangular distribution. In this study, the generation of variables with the use of the Cholesky decomposition showed that the type of probability distribution of output variables may differ from the distribution of input variables. This is particularly evident in Figure 1, where the conversion of input data with a rectangular distribution into a set of correlated variables results in different types of distributions. If at the input to the algorithm random variables have a normal distribution, at the output correlated variables also have a normal distribution (Fig. 2). Therefore, with regard to type B evaluation of measurement uncertainty, this method has a significant drawback despite its advantageous simplicity; it does not allow a reliable control of probability distribution of the generated correlated variables.

3. Example of measurement uncertainty analysis taking into account correlated variables

3.1. Strict solution

The example illustrates the evaluation of the combined uncertainty of stiffness coefficient E based on the knowledge of the results of measurement of the force F and displacement δ . It is assumed that the maximum permissible errors of input data are known: MPE(F) = 20 N, MPE(δ) = 2 μm .

$$E = \frac{F}{\delta} = \frac{400}{20} = 20 \frac{\text{N}}{\mu\text{m}} \quad (4)$$

With the assumption that input variables are independent and have a rectangular distribution, the combined standard uncertainty would be as follows:

$$u_c(E) = \sqrt{\left(\frac{1}{\delta}\right)^2 \cdot \left(\frac{MPE(F)}{\sqrt{3}}\right)^2 + \left(-\frac{F}{\delta^2}\right)^2 \cdot \left(\frac{MPE(\delta)}{\sqrt{3}}\right)^2} = 1,29 \cdot \frac{\text{N}}{\mu\text{m}} \quad (5)$$

Taking into account the correlation with the a priori assumed correlation coefficient $r(F, \delta) = 0.9$, it is:

$$\begin{aligned} u_c(E^r) &= \sqrt{\left(\frac{1}{\delta}\right)^2 \cdot \left(\frac{MPE(F)}{\sqrt{3}}\right)^2 + \left(-\frac{F}{\delta^2}\right)^2 \cdot \left(\frac{MPE(\delta)}{\sqrt{3}}\right)^2 + \dots} \\ &\sqrt{2 \cdot \frac{1}{\delta} \cdot \left(-\frac{F}{\delta^2}\right) \cdot \frac{MPE(F)}{\sqrt{3}} \cdot \frac{MPE(\delta)}{\sqrt{3}} \cdot r(F, \delta)} = 0,68 \cdot \frac{\text{N}}{\mu\text{m}} \end{aligned} \quad (6)$$

Therefore the rejection of the correlation of input variables leads to a significant overestimation of uncertainty of the discussed measurement.

3.2. Monte Carlo simulation

In the next step, for the case presented in the previous section, measurement uncertainty was determined using the Monte Carlo simulation [6]. The calculations were performed for the variables treated as independent at different numbers of steps $n = 20, 50, 100, 1000, 5000$ and 10000 in loop randomization.

$$\begin{aligned} &\text{for } i = 1:n \\ &\Delta F = \text{rand} \cdot MPE(F) \\ &\Delta \delta = \text{rand} \cdot MPE(\delta) \\ &E_i^{MC} = \frac{F + \Delta F}{\delta + \Delta \delta} \\ &\text{end} \end{aligned} \quad (7)$$

where rand is a pseudo-random number from a population with a rectangular distribution with a range from -1 to 1 ; its value at each step of the loop is random regardless of F and δ .

In contrast, calculations in loop randomization taking into account the correlation of variables were performed for:

$$\begin{aligned} &\text{for } i = 1:n \\ &\Delta F = Fr_i \cdot MPE(F) \\ &\Delta \delta = \delta r_i \cdot MPE(\delta) \end{aligned}$$

$$E_i^{MCr} = \frac{F + \Delta F}{\delta + \Delta \delta}$$

end

(8)

where: F_i and δ_i – the elements of correlated vectors generated in an algorithm presented in p. 2 for the assumed a priori correlation coefficient $r(F, \delta) = 0.9$. The elements of these vectors take values between -1 and 1, and have rectangular distributions. Other symbols denote the same as above.

Example results of calculations and percentage differences between a Monte Carlo and strict solution (p. 3.1) for various values are shown in Table 1.

Table 1. The comparison of Monte Carlo and strict solutions in evaluating the uncertainty of stiffness coefficient determination, with or without taking into account the correlation of input variables

Parameters	Parameters of Monte Carlo method					
	20	50	100	1000	5000	10000
strict solution E , N/ μ m	20					
mean E^{MC} or E^{MCr}	20.4	20.15	20.06	20.18	20.08	20.10
% difference	2	0.7	0.3	0.9	0.4	0.5
strict solution $u_c(E)$, N/ μ m	1.29					
standard deviation E^{MC}	1.091	1.256	1.361	1.344	1.288	1.296
% difference	-15	-2.6	5.5	4.2	-0.2	0.5
strict solution $u_c(E')$, N/ μ m	0.68					
standard deviation E^{MCr}	0.725	0.701	0.691	0.690	0.689	0.683
% difference	6.6	3.1	1.6	1.5	1.3	0.4

If the differences between Monte Carlo and exact solutions are less than 10%, the results of calculations are considered to be satisfactory. In this approach, a satisfactory solution was achieved already at 50 repetitions of loop randomization. In the worst case a little over 3% difference was obtained. A further increase in the number of repetitions in the Monte Carlo simulation did not result in a significant improvement in the accuracy of calculations.

4. Conclusion

The article presents two methods of generating cross-correlated variables for the type B evaluation of measurement uncertainty. In addition to the specific proposals presented in the previous chapters, the whole article can be summarized as follows:

The presented example shows an easy method of determination of analytical solutions for the stiffness coefficient and its measurement uncertainty. The Monte

Carlo simulation was used only as a validation of solutions obtained with the use of correlated input variables.

It should be emphasized that the estimation of stiffness of machine tools often requires considering systems with many degrees of freedom, without a simple analytic dependency for the determination of their stiffness. Therefore for complex computational models, simulations with the use of the randomization methods analyzed in this study seem to be the only appropriate alternative for the estimation of measurement uncertainty. We show their usefulness and accuracy even if cross-correlated variables are included in the analysis.

References

- [1] K. KONOWALSKI: Experimental research and modeling of normal contact stiffness and contact damping of machined joint surfaces. *Advances in Manufacturing Science and Technology*, **33**(2009)3, 53-68.
- [2] P. GRUDZIŃSKI, K. KONOWALSKI: Studies of flexibility of a steel adjustable foundation chock. *Advances in Manufacturing Science and Technology*, **36**(2012)2, 79-90.
- [3] Wyrażanie niepewności pomiaru, przewodnik (Expressing of measurement uncertainty, a guide). Główny Urząd Miar, Warsaw 1999. (in Polish).
- [4] E. ANDREW SMITH, P. BARRY RYAN, JOHN S. EVANS: The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk. *Risk Analysis*, **12**(1992)4, 467-474.
- [5] L. IMAN RONALD, W.J. CONOVER: A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics-Simulation and Computation*, **11**(2002)3, 311-334.
- [6] J. ŚLĄDEK: Dokładność pomiarów współrzędnościovych. Wydawnictwo Politechniki Krakowskiej, Kraków 2011.

Received in May 2015

