Wykorzystanie liczb losowych do obliczania niepewności pomiaru

Paweł Fotowicz

1. Wstęp

Dokument JCGM 101:2008 [1] wyznacza nowy standard w dziedzinie obliczania niepewności pomiaru. W miejsce prawa propagacji niepewności [2] wprowadza metodę propagacji rozkładów. Rozkład prawdopodobieństwa dla wielkości mierzonej nie jest przyjmowany a priori, lecz obliczany przy użyciu symulacji Monte Carlo poprzez matematyczny model wielkości wyjściowej. Miarą niepewności pomiaru jest przedział rozszerzenia, którego długość wyznacza odległość pomiędzy kwantylami tego rozkładu dla określonego prawdopodobieństwa rozszerzenia, na ogół 95 %.

2. Procedura Monte Carlo

Zalecana procedura [1] polega, w pierwszy kroku postępowania, na wyborze liczby próbkowań M. Liczba ta zależy od zakładanego prawdopodobieństwa rozszerzenia p. Przykładowo, dla zazwyczaj przyjmowanego $p = 95$ %, liczba ta wynosi $M = 10^4$. Następnie, na podstawie przyjętych rozkładów dla wielkości wejściowych generuje się zbiory możliwych wartości dla tych wielkości, każdy o liczebności równej M. Na podstawie tych wartości, wstawiając je do równania pomiaru, można wyznaczyć zbiór wartości dla wielkości wyjściowej. Sortując te wartości od najmniejszej do największej oraz przypisując im kolejne prawdopodobieństwa można wyznaczyć granice przedziału rozszerzenia. W przypadku symetrycznego rozkładu prawdopodobieństwa połowa tego przedziału wyznacza niepewność rozszerzoną.

Procedura powyższa zakłada oddzielne generowanie zbiorów dla każdej wielkości wejściowej i wykorzystanie równania pomiaru do wyznaczenia możliwych wartości dla wielkości wyjściowej. W pewnym, określonym przypadku można ten etap procedury pominąć i zastosować poniższe rozwiązanie.

3. Metoda obliczeniowa

Gdy równanie pomiaru jest liniowe lub linearyzowane, a wielkości wejściowe opisane są rozkładami: Studenta, normalnym, prostokątnym, trójkątnym lub trapezowym, zamiast generowania wielu zbiorów dla rozkładów wejściowych można generować tylko jeden zbiór wartości bezpośrednio dla wielkości wyjściowej określony równaniem:

$$y = u_i z_\mathrm{P} + \sqrt{\sum_{j=1,\, j\neq i}^{N} \left(\frac{t(\nu)}{k_\mathrm{N}} u_j \right)^2} \, z_\mathrm{N} \tag{1}$$

gdzie z_P i z_N to zmienne losowe o rozkładach standaryzowanych, odpowiednio prostokątnym P (0, 1) i normalnym N (0, 1), u_i to największy udział o rozkładzie prostokątnym, $t(\nu)$ to kwantyl rozkładu Studenta o liczbie stopni swobody ν, k_N to kwantyl rozkładu normalnego, a u_j to pozostałe udziały niepewności. Dla wielkości wejściowych opisanych innymi rozkładami niż Studenta należy przyjąć $t(\nu) = k_N$.

Równanie powyższe definiuje zmienną losową o rozkładzie płasko-normalnym (P*N), który jest splotem rozkładu prostokątnego z normalnym. Standaryzowany rozkład P*N (0, 1) można otrzymać stosując równanie:

$$z_{PN} = \frac{r\, z_P + z_N}{\sqrt{r^2 + 1}} \tag{2}$$

gdzie r to parametr rozkładu płasko-normalnego, będący ilorazem odchylenia standardowego σ_P rozkładu prostokątnego do odchylenia standardowego σ_N rozkładu normalnego:

$$r = \frac{\sigma_P}{\sigma_N} \tag{3}$$

Sortując zbiór możliwych wartości dla wielkości wyjściowej y, od najmniejszej do największej, i przypisując im kolejne prawdopodobieństwa możemy wyznaczyć dystrybuantę numeryczną jej rozkładu. Ponieważ rozkład wielkości wyjściowej jest symetryczny, to można określić granice probabilistycznie symetrycznego przedziału rozszerzenia jako:

$$y_{high} = G^{-1}\left(\frac{1+p}{2}\right) \quad \text{i} \quad y_{low} = G^{-1}\left(\frac{1-p}{2}\right) \tag{4}$$

gdzie $G^{-1}(\alpha)$ jest kwantylem rzędu α dystrybuanty numerycznej rozkładu wielkości wyjściowej. Na tej podstawie można wyznaczyć niepewność rozszerzoną dla określonego prawdopodobieństwa rozszerzenia:

$$U = \frac{y_{high} - y_{low}}{2} \tag{5}$$

jako połowę długości probabilistycznie symetrycznego przedziału rozszerzenia.

4. Własności metody obliczeniowej

Własności proponowanej metody numerycznej prześledźmy na kilku przykładach obliczeniowych. Jednym z nich niech będzie prosty model pomiaru przedstawiony w dokumencie [1], zawierający cztery składowe opisane rozkładami prostokątnymi, z których jedna ma dziesięciokrotnie większą niepewność standardową od pozostałych, mających niepewność standardową równą umownie jeden. Na podstawie przedstawionych tam wartości granicznych 95 % probabilistycznie symetrycznego przedziału rozszerzenia można wyznaczyć, że $U_{MCM} = 17{,}0$. Używając metody numerycznej, przy zastosowaniu równania (1), uzyskamy dla tych samych danych wejściowych niepewność rozszerzoną $U_{NM} = 17{,}1$. Dla porównania stosując metodę analityczną, opisaną w publikacji [3], otrzymamy $U_{AM} = 16{,}9$.

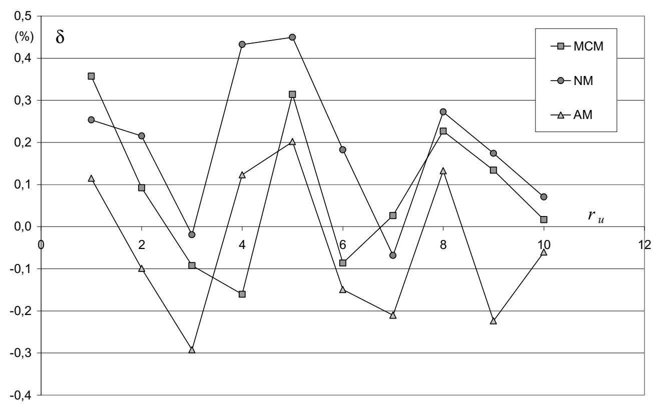
Powiększając w przykładowym modelu pomiaru liczbę składowych do dziesięciu można również wyznaczyć niepewność powyższymi metodami. Wyniki obliczeń odnoszono do wyniku obliczeń na bazie splotu dziesięciu rozkładów prostokątnych, wykonanych przy użyciu szybkiej transformaty Fouriera [4] i przedstawiono w postaci błędów metod w postaci:

$$\delta = \frac{U - U_{\text{splot}}}{U_{\text{splot}}} \tag{6}$$

gdzie U i U_{splot} to niepewności rozszerzone dla prawdopodobieństwa p = 95 % obliczane rozpatrywaną metodą oraz metodą splotu rozkładów. Przykładowe wyniki obliczeń reprezentuje rys. 1. Wartości błędów obliczono w funkcji parametru:

$$r_u = \frac{|u_i|}{\sqrt{u_c^2 - u_i^2}} \tag{7}$$

gdzie u_c to złożona niepewność standardowa. Dla każdej składowej przyjęto niepewność standardową równą jeden, a dla składowej dominującej u_i większą od jednego. We wszystkich przypadkach obliczeniowych błąd metod nie przekraczał 1 %.



Rys. 1. Przykładowe błędy względne metod obliczania niepewności rozszerzonej dla prawdopodobieństwa 95 %. MCM – metoda Monte Carlo, NM – metoda numeryczna, AM – metoda analityczna

Podobne postępując, jak w przykładzie z dokumentu [1] dla równania z czterema składowymi, można również przyjąć dla każdej z nich inny rozkład prawdopodobieństwa: Studenta, normalny, trójkątny i prostokątny. Można wykonać obliczenia, gdy wszystkim czterem składowym przypisuje się jednakową niepewność standardową równą umownie jeden oraz gdy każda z nich staje się dominująca z niepewnością dziesięciokrotnie większą od pozostałych. Przykładowe wyniki obliczeń przedstawia tabela 1. W obliczeniach przyjęto rozkład Studenta z liczbą stopni swobody ν = 2. W przypadku obliczeniowym A brak jest dominacji składowej, a w pozostałych przypadkach obliczeniowych dominują odpowiednio: w B składowa o rozkładzie Studenta, w C składowa o rozkładzie normalnym, w D składowa o rozkładzie trójkątnym, a w E składowa o rozkładzie prostokątnym.

Tabela 1. Przykładowe wyniki obliczenia niepewności rozszerzonej dla różnych rozkładów i udziałów składowych w liniowym równaniu pomiaru

Metoda	MCM	NM	AM
A	5,42	5,48	5,45
B	43,21	43,07	43,16
C	20,20	20,18	20,26
D	19,88	19,86	19,84
E	17,38	17,47	17,26

5. Wnioski

Zaproponowana metoda numeryczna zapewnia podobną dokładność obliczeniową niepewności rozszerzonej i tym samym przedziału rozszerzenia jak metoda Monte Carlo i metoda analityczna. Przy obliczeniach nie ma konieczności generowania oddzielnie rozkładów dla każdej wielkości wejściowej, bowiem z równania pomiaru (1) od razu uzyskujemy zbiór wartości dla wielkości wyjściowej na podstawie przypisanych rozkładów poszczególnym składowym i wartości ich udziałów. Równanie pomiaru powinno być liniowe lub linearyzowane. Pozwala na natychmiastowe obliczenia niepewności dla dowolnej liczby składowych. Łatwo się implementuje do tabeli budżetu niepewności, przy użyciu arkusza kalkulacyjnego.

Literatura

1. Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method. JCGM 101:2008.
2. Evaluation of measurement data – Guide to the expression of uncertainty in measurement. JCGM 100:2008.
3. Fotowicz P.: *An analytical method for calculating a coverage interval*. Metrologia, vol. 43 (2006).
4. Korczyński M. J., Hetman A., Fotowicz P.: *Fast Fourier Transformation – An Approach to Coverage Interval Calculation vs. Approximation Methods*. AMUEM 2005.