

Omówienie międzynarodowego dokumentu JCGM 102:2011 dotyczącego wyrażania niepewności pomiaru

dr Paweł Fotowicz (Zakład Promieniowania i Drgań, GUM)

Dokument JCGM 102:2011 jest rozwinięciem metodyki opracowania danych pomiarowych przedstawionej w opracowaniu JCGM 101:2008. Dotyczy wielowymiarowego modelu pomiaru, czyli takiego w którym występuje dowolna liczba wielkości wyjściowych. Wielkości te są wzajemnie skorelowane, gdyż zależą od tych samych wielkości wejściowych. Dokument przedstawia prawo propagacji niepewności w postaci macierzowej. Uogólnia też zastosowanie metody Monte Carlo w celu numerycznego wyznaczania wspólnego rozkładu prawdopodobieństwa dla wielkości wyjściowej wielowymiarowego modelu pomiaru. Na ich podstawie można wyznaczyć obszar rozszerzenia, będący odpowiednikiem przedziału rozszerzenia dla jednowymiarowego modelu pomiaru, który odpowiada określonemu prawdopodobieństwu. Obszar ten może przybierać postać hiperelipsy lub hiperprostokąta. Dokument przedstawia również procedurę obliczeniową wyznaczania najmniejszego obszaru rozszerzenia.

The document describes a generalization of the Monte Carlo method for measurement models having any number of input quantities and any number of output quantities. Two approaches are considered for treating such models. The first approach is a generalization of the Guide uncertainty framework. The second is a Monte Carlo method as an implementation of the propagation of distributions. Guidance is also given on the determination of a coverage region for the output quantities of a multivariate model, the counterpart of a coverage interval for a single scalar output quantity, corresponding to a stipulated coverage probability. The guidance includes the provision of coverage regions that take the form of hyper-ellipsoids and hyper-rectangles. A calculation procedure is also described for obtaining an approximation to the smallest coverage region.

Wprowadzenie

Wspólny Komitet ds. Przewodników w Metrologii JCGM (Joint Committee for Guides in Metrology), działający pod egidą Międzynarodowego Biura Miar, wydał w 2011 roku kolejny dokument [1] dotyczący wyrażania niepewności pomiaru. Jest on kontynuacją i uogólnieniem podejścia przedstawionego w dokumencie [2], opracowanym w 2008 roku. Oba te materiały w znaczący sposób rozwijają metodykę opracowania danych pomiarowych, opisaną w klasycznym już dla metrologów dokumencie [3], opartą na prawie propagacji niepewności.

Omawiany dokument dotyczy wielowymiarowego modelu pomiaru, czyli takiego w którym występuje dowolna liczba wielkości wyjściowych. Wielkości te są wzajemnie skorelowane, gdyż zależą od tych samych wielkości wejściowych. Materiał przedstawia uogólnienie podejścia GUM (Guide Uncertainty Measurement), służącego do wyznaczania estymat

wielkości wyjściowych, związanych z nimi niepewności standardowych i kowariancji par tych wielkości. Wielkości wejściowe i wyjściowe mogą być rzeczywiste, jak i zespolone. Dokument uogólnia zastosowanie metody Monte Carlo w celu numerycznego wyznaczania wspólnego rozkładu prawdopodobieństwa dla wielkości wyjściowej wielowymiarowego modelu pomiaru. Na jego podstawie można wyznaczyć obszar rozszerzenia, będący odpowiednikiem przedziału rozszerzenia dla jednowymiarowego modelu pomiaru, który odpowiada określonemu prawdopodobieństwu. Obszar ten może przybierać postać hiperelipsy lub hiperprostokąta. Dokument przedstawia również procedurę obliczeniową wyznaczania najmniejszego obszaru rozszerzenia.

Sposób zapisu

Wielowymiarowy model pomiaru może być zapisany w postaci:

$$Y = f(X) \quad (1)$$

gdzie $Y = (Y_1, \dots, Y_m)^T$ jest wielowymiarową wielkością wyjściową, $f = (f_1, \dots, f_m)^T$ jest wielowymiarową funkcją pomiaru, a X reprezentuje N wielkości wejściowych $(X_1, \dots, X_N)^T$. Estymatą X jest $x = (x_1, \dots, x_N)^T$, macierz o wymiarze $N \times 1$. Macierzą kowariancji U_x wielkości wejściowych jest macierz o wymiarze $N \times N$. Estymatą Y jest $y = (y_1, \dots, y_m)^T$, macierz o wymiarze $m \times 1$. Macierzą kowariancji U_y wielkości wyjściowych jest macierz o wymiarze $m \times m$.

Zasady postępowania

Główne etapy postępowania to: opis wielkości (*Formulation*), obliczenia (*Propagation*) i zapis wyniku (*Summarizing*). Opis wielkości dotyczy:

- 1) definicji wielkości wyjściowej Y jako wektora,
- 2) określenie wielkości wejściowej X , od której zależy wielkość wyjściowa Y ,
- 3) zbudowanie modelu pomiaru lub funkcji pomiaru f określającej relacje pomiędzy X i Y ,
- 4) przyjęcie rozkładów prawdopodobieństwa (na ogół normalnych lub prostokątnych) dla składowych wielkości wejściowej X , a dla skorelowanych par składowych X wspólnych funkcji gęstości prawdopodobieństwa.

Obliczenia polegają na realizacji zasady propagacji rozkładów składowych wielkości wejściowej X , poprzez model pomiaru, w celu otrzymania wspólnego rozkładu dla wielkości wyjściowej Y . Zapis wyniku polega na przedstawieniu:

- 1) wartości oczekiwanej Y jako estymaty y wielkości wyjściowej,
- 2) macierzy kowariancji U_y wielkości Y ,
- 3) obszaru rozszerzenia dla wielkości wyjściowej Y przy określonym prawdopodobieństwie.

Podstawa obliczeniowa niepewności (GUF)

Propagacja niepewności jest uogólnieniem prawa propagacji niepewności stosowanego w dziedzinie jednowymiarowych modeli pomiaru. Opiera się na rachunku macierzowym, wykorzystując estymatę $x = (x_1, \dots, x_N)^T$ wielkości wejściowej X oraz macierz kowariancji U_x związaną z estymatą x , zawierającą kowariancje $u(x_i, x_j)$ wielkości x_i, x_j .

Propagacja niepewności dla jawnych wielowymiarowych modeli pomiaru polega na określeniu relacji pomiędzy wielkością wyjściową $Y = (Y_1, \dots, Y_m)^T$ a wielkością wejściową $X = (X_1, \dots, X_N)^T$ w postaci $Y = f(X)$, gdzie $f = (f_1, \dots, f_m)^T$ oznacza wielowymiarową funkcję pomiaru. Dla estymaty x wielkości X estymata wielkości Y dana jest zależnością $y = f(x)$. Macierz kowariancji o wymiarze $m \times m$, związana z estymatą y , ma postać:

$$U_y = \begin{bmatrix} u^2(y_1) & \dots & u(y_1, y_m) \\ \vdots & \ddots & \vdots \\ u(y_m, y_1) & \dots & u^2(y_m) \end{bmatrix} \quad (2)$$

i jest obliczana na podstawie równania:

$$U_y = C_x U_x C_x^T \quad (3)$$

gdzie C_x jest macierzą wrażliwości o wymiarze $m \times N$, dana zależnością:

$$C_x = \begin{bmatrix} \frac{\partial f_1}{\partial X_1} & \dots & \frac{\partial f_1}{\partial X_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial X_1} & \dots & \frac{\partial f_m}{\partial X_N} \end{bmatrix} \quad (4)$$

Propagacja niepewności dla niejawnych wielowymiarowych modeli pomiaru polega na określeniu relacji pomiędzy wielkością wyjściową $Y = (Y_1, \dots, Y_m)^T$ a wielkością wejściową $X = (X_1, \dots, X_N)^T$ w postaci $h(Y, X) = 0$, gdzie $h = (h_1, \dots, h_m)^T$. Estymata x wielkości X i estymata y wielkości Y spełnia równanie $h(y, x) = 0$. Macierz kowariancji o wymiarze $m \times m$, związana z estymatą y , jest obliczana z równania:

$$C_y U_y C_y^T = C_x U_x C_x^T, \quad (5)$$

gdzie C_y jest macierzą wrażliwości o wymiarze $m \times m$ zawierającą pochodne cząstkowe $\partial h_i / \partial Y_j$, a C_x jest macierzą wrażliwości o wymiarze $m \times N$, zawierającą pochodne cząstkowe $\partial h_i / \partial X_j$. Wszystkie pochodne będą obliczane dla $X = x$ i $Y = y$.

Formalnie, macierze kowariancji U_x i U_y są powiązane zależnością:

$$U_y = C U_x C^T, \quad (6)$$

gdzie:

$$C = C_y^{-1} C_x \quad (7)$$

jest macierzą wrażliwości o wymiarze $m \times N$.

Obszar rozszerzenia dla wektorowej wielkości wyjściowej

Rozważane są w szczególności dwa rodzaje obszaru rozszerzenia:

- a) hiperelipsa lub elipsa wielowymiarowa o liczbie wymiarów m , która to dana jest zależnością:

$$(\boldsymbol{\eta} - \mathbf{y})^T \mathbf{U}_y^{-1} (\boldsymbol{\eta} - \mathbf{y}) = k_p^2, \quad (8)$$

gdzie k_p jest stałą określającą wielowymiarowy obszar rozszerzenia z prawdopodobieństwem p , a gdy $m = 1$ to $(\eta - y)^2 = k_p^2 u_y^2$, co daje $\eta = y \pm k_p u_y$,

- b) hiperprostokąt lub prostokąt wielowymiarowy o liczbie wymiarów m , centrowany estymatą \mathbf{y} z bokami równymi długości oddzielnie określanych przedziałów rozszerzenia dla każdego z wymiarów Y_j . Przedziały te określane są dla prawdopodobieństwa rozszerzenia $q = 1 - (1 - p)/m$. Wielowymiarowy rozkład prostokątny określa obszar rozszerzenia dla wielkości \mathbf{Y} z prawdopodobieństwem p .

Ogólnie istnieje wiele dowolnych obszarów rozszerzenia dla \mathbf{Y} o określonym prawdopodobieństwie. Każdy z nich oparty jest na dystrybucji \mathbf{G} , czyli zbiorze M punktów \mathbf{y}_r przypadkowo losowanych z rozkładu wielkości \mathbf{Y} , które mogą być otrzymane przy zastosowaniu metody Monte Carlo:

- a) hipereliptyczny obszar rozszerzenia, który będzie najmniejszym obszarem rozszerzenia dla \mathbf{Y} , gdy dobrym przybliżeniem \mathbf{Y} jest rozkład normalny,
 b) hiperprostokątny obszar rozszerzenia, który jest najprostszą interpretacją tego obszaru, ale nadmierny,
 c) najmniejszy obszar rozszerzenia, który nie ma szczególnej geometrycznej definicji i jest otrzymywany jako stopień przybliżenia zależny od M .

Najmniejszy obszar rozszerzenia

Postępowanie przy wyznaczaniu najmniejszego obszaru rozszerzenia:

- a) zbudowanie hiperprostokątnego obszaru w przestrzeni wielkości wyjściowej,
 b) podzielenie tego prostokątnego obszaru na sieć mniejszych prostokątów,
 c) przypisanie każdej wartości wielkości wyjściowej do tego mniejszego prostokąta, który ją zawiera,
 d) użycie części wartości wielkości wyjściowych przypisanych do każdego prostokąta jako przybliżenia

nia prawdopodobieństwa, że \mathbf{Y} należy do tego prostokąta,

- e) wylistowanie prostokątów w porządku malejącego prawdopodobieństwa,
 f) utworzenie skumulowanej sumy prawdopodobieństw dla tych wylistowanych prostokątów, gdy suma ich nie jest mniejsza od prawdopodobieństwa p ,
 g) przyjęcie zbioru odpowiednich prostokątów do zdefiniowania najmniejszego obszaru rozszerzenia.

Metoda Monte Carlo (MCM)

Metoda MCM może być realizowana krok po kroku:

- a) wybór liczby próbek M ,
 b) wygenerowanie M wektorów poprzez losowanie z funkcji gęstości prawdopodobieństwa, przypisanych wielkościom wejściowym X_j ,
 c) dla każdego takiego wektora wyznaczenie odpowiadających mu wartości \mathbf{Y} , uzyskując M wartości wektorowej wielkości wyjściowej,
 d) przyjęcie jako reprezentacji dystrybuanty \mathbf{G} zbioru M wartości wektorowych \mathbf{Y} ,
 e) użycie \mathbf{G} do utworzenia estymaty \mathbf{y} wielkości \mathbf{Y} i macierzy kowariancji \mathbf{U}_y związanej z \mathbf{y} ,
 f) użycie \mathbf{G} do utworzenia odpowiedniego obszaru rozszerzenia dla \mathbf{Y} z określonym prawdopodobieństwem p .

Zalecana procedura Monte Carlo

Skuteczność procedury MCM, wyznaczającej estymatę \mathbf{y} wielkości wyjściowej \mathbf{Y} , związaną z nią macierz kowariancji \mathbf{U}_y oraz obszar niepewności \mathbf{Y} , zależy od użycia adekwatnie dużej liczby M próbek. Procedura prowadzi do:

- a) estymaty $\mathbf{y} = (y_1, \dots, y_m)^T$ wielkości wektorowej \mathbf{Y} ,
 b) wektora $\mathbf{u}(\mathbf{y}) = (u(y_1), \dots, u(y_m))^T$ niepewności standardowych związanych z estymatami,
 c) macierzy \mathbf{R}_y , o wymiarze $m \times m$, współczynników korelacji $r_{ij} = r(y_i, y_j)$ związanych z parami estymat,
 d) współczynnika rozszerzenia k_p zdefiniowanego dla obszaru rozszerzenia w postaci m wymiarowej hiperelipsy, dla którego oczekuje się, że wartości te mogą być zgodne z określoną tolerancją numeryczną.

Tolerancja numeryczna: $\delta = \frac{1}{2}10^l$ związana z wielkością z , gdy wielkość ta wyrażana jest w postaci: $z = c \times 10^l$, c to mantysa z określoną liczbą cyfr znaczących, a l to wykładnik.

Zalecana procedura postępowania:

- przyjmij n_{dig} jako małą liczbę całkowitą,
- przyjmij $M = \max(J, 10^4)$, gdzie $J \geq 100/(1-p)$ to najmniejsza liczba całkowita,
- przyjmij $h = 1$ jako pierwszy krok postępowania,
- przeprowadź symulację Monte Carlo na próbie M wartości,
- używając M wartości wielkości wyjściowej wektorowej y_1, \dots, y_M oblicz $y^{(h)}$, $u(y^{(h)})$, $R_y^{(h)}$ i $k_p^{(h)}$ jako estymatę Y , w postaci niepewności standardowych, macierzy korelacji, współczynnika rozszerzenia dla 100p % obszaru rozszerzenia,
- jeżeli $h \leq 10$ to powiększ h o jeden i powtórz procedurę od kroku d),
- dla $j = 1, \dots, m$ oblicz odchylenie standardowe s_{y_j} związane ze średnią estymat $y_j^{(1)}, \dots, y_j^{(h)}$ wielkości Y_j na podstawie:

$$s_{y_j}^2 = \frac{1}{h(h-1)} \sum_{r=1}^h (y_j^{(r)} - y_j)^2, \text{ gdzie } y_j = \frac{1}{h} \sum_{r=1}^h y_j^{(r)} \quad (9)$$

- oblicz powyższe statystyki dla $u(y^{(h)})$, λ_{max} i $k_p^{(h)}$,
- użyj wszystkich wartości funkcji modelu pomiaru $h \times M$, aby określić $u(y)$, R_y i k_p ,
- dla $j = 1, \dots, m$ oblicz tolerancję numeryczną δ związaną z $u(y_j)$,
- oblicz tolerancję numeryczną związaną z macierzą współczynników korelacji R_y ,
- oblicz tolerancję numeryczną związaną ze współczynnikiem rozszerzenia k_p ,
- jeżeli dla każdego $j = 1, \dots, m$ któraś z wartości statystyk: $2s_{y_j}$, $2s_{u(y_j)}$, $2s_{\lambda_{\text{max}}}$, $2s_{k_p}$ przekracza odpowiednią tolerancję numeryczną, to powiększ h o jeden i wróć do kroku d),
- stwierdziwszy, że wszystkie obliczenia są stabilne, użyj wszystkich wartości wielkości wyjściowej wektorowej ($h \times M$), aby wyznaczyć: y , U , oraz k_p dla 100p % obszaru rozszerzenia.

Walidacja obliczeń

Dokument przedstawia zalecany sposób postępowania przy walidacji obliczeń tradycyjną metodą, sprowadzającą się do dwóch działań:

- zastosowanie propagacji niepewności w celu uzyskania estymaty y^{GUF} , niepewności standardowych $u(y^{\text{GUF}})$, macierzy korelacji R_y^{GUF} i współczynnika rozszerzenia k_p^{GUF} dla 100p % obszaru rozszerzenia w postaci m wymiarowej hiperelipsy,
- zastosowanie zalecanej procedury obliczeniowej dla metody Monte Carlo w celu otrzymania odpowiednio: y^{MCM} , $u(y^{\text{MCM}})$, R_y^{MCM} i k_p^{MCM} .

Następnie należy sprawdzić, czy otrzymane wartości estymat, niepewności standardowych, współczynników korelacji i rozszerzenia metodą GUF (Guide Uncertainty Framework) i MCM (Monte Carlo Method) zgadzają się co do ustalonych tolerancji numerycznych. W tym celu oblicza się:

$$d_{y_j} = |y_j^{\text{GUF}} - y_j^{\text{MCM}}| \text{ dla } j = 1, \dots, m \quad (10)$$

$$d_{u(y_j)} = |u(y_j^{\text{GUF}}) - u(y_j^{\text{MCM}})| \text{ dla } j = 1, \dots, m \quad (11)$$

$$d_{\lambda_{\text{max}}} = |\lambda_{\text{max}}^{\text{GUF}} - \lambda_{\text{max}}^{\text{MCM}}| \quad (12)$$

$$d_{k_p} = |k_p^{\text{GUF}} - k_p^{\text{MCM}}| \quad (13)$$

Jeżeli wartości te są nie większe niż odpowiednie tolerancje numeryczne, to można uznać obliczenia wykonane metodą propagacji niepewności za zwalidowane.

Zakończenie

Dokument [1] jak i pozostałe opracowania, w tym [2] i [3], wydane pod wspólnym tytułem „Evaluation of measurement data”, dostępne są na stronie internetowej Międzynarodowego Biura Miar (BIPM) w dziale dotyczącym Przewodników w Metrologii (Guides in Metrology). Dokument zawiera również szereg przykładów obliczeniowych mających zastosowanie w metrologii.

Literatura

- Supplement 2 to the “Guide to the expression of uncertainty in measurement” – *Extension to any number of output quantities*. JCGM 102:2011.
- Supplement 1 to the “Guide to the expression of uncertainty in measurement” – *Propagation of distributions using a Monte Carlo method*. JCGM 101:2008.
- Guide to the expression of uncertainty in measurement. JCGM 100:2008. ISO 1995 (*Wyrażanie niepewności pomiaru*. Przewodnik. Wydawca Główny Urząd Miar 1999).