Accounting for correlations in measurement models

Adriaan M.H. van der Veen

Van Swinden Laboratorium, Delft, the Netherlands

17NRM05 – Examples of measurement uncertainty evaluation
M18 Workshop, LNE, Paris, France, 21–22 January 2020

Dutch Metrology Institute

Examples of Measurement Uncertainty Evaluation

Introduction

- Correlations (dependencies) between quantities arise when they depend on (at least) one other joint quantity
- It is essential to evaluate covariances as part of the evaluation of measurement uncertainty
- Ignoring covariances is the same as setting them to zero ...
- ... which may lead to serious over- or understating the measurement uncertainty

Contents

A simple case

- A calibrated volt meter is used to measure two voltages which are to be used in a subsequent calculation. What is the covariance between the two voltages?
  - $V_1 = \delta V_{\text{cal}} + \bar{V}_{\text{ind},1}$
  - $V_2 = \delta V_{\text{cal}} + \bar{V}_{\text{ind},2}$
- We assume the indications as independent and identically distributed (IID).
- The only common variable in the two measurement equations is δV_{cal}, so according to equation (F.2) in the GUM, $u(V_1, V_2) = u^2(\delta V_{\text{cal}})$
- The correlation coefficient is evaluated as $r(V_1, V_2) = \dfrac{u^2(\delta V_{\text{cal}})}{u(V_1)u(V_2)}$

Generalisation to multiple common quantities

- If there are multiple common input quantities,

$$u(X_1, X_2) = \sum_i \frac{\partial F}{\partial Q_i} \frac{\partial G}{\partial Q_i} u^2(Q_i)$$

- where F and G are the measurement equations relating X_1 and X_2 to the input quantities Q_i;
- only terms for which both partial derivatives are non-zero contribute (i.e., the ones concerning the common quantities).
- See also JCGM 100:2008 clause F.1.2

Buoyancy effect on a permeation tube in a chamber flushed with high-purity nitrogen

- Measurand of interest: permeation rate $\dot{m} = q_m = \mathrm{d}m/\mathrm{d}t$
- Buoyancy effect used to model temperature and pressure influence on weighing results
- Buoyancy effect modelled as

$$\delta m_{\mathrm{buo}} = V_{\mathrm{tube}} \delta\rho \qquad \text{where} \qquad \rho = \frac{\bar{M}}{V_{\mathrm{m}}}$$

- Molar volume of high-purity nitrogen V_{m} computed using a cubic polynomial (Soave-Redlich-Kwong equation-of-state)

Use of the Soave-Redlich-Kwong equation-of-state

Equation-of-state can be formulated as

$$V_m^3 + \alpha_1 V_m^2 + \alpha_2 V_m + \alpha_3 = 0$$

where

$$\alpha_1 = -\frac{RT}{p}; \alpha_2 = -b^2 - \frac{RT}{p} b + \frac{a}{p}; \alpha_3 = -\frac{ab}{p}$$

and a and b are coefficients dependent on, among other, the critical temperature and pressure.

Propagation of uncertainty can be performed using the law of propagation of uncertainty from GUM-S2 for (multivariate) implicit measurement models, enabling including the correlations between the α_i, i.e.,

$$u^2(V_\mathrm{m}) = \left(3V_\mathrm{m}^2 + 2\alpha_1 V_\mathrm{m} + \alpha_2\right)^{-2} \boldsymbol{C}_{\boldsymbol{\alpha}} \boldsymbol{U}_{\boldsymbol{\alpha}} \boldsymbol{C}_{\boldsymbol{\alpha}}^\mathrm{T}$$

Generalisation to multiple output quantities

- If there are
  - multiple quantities having correlations and/or
  - multiple measurement equations

  the law of propagation of uncertainty for multivariate measurement models (GUM-S2) is a convenient tool to calculate the covariances
- In the notation from the GUM, let $X_i = f_i(\boldsymbol{Q})$ for $i = 1 \dots N$, then $\boldsymbol{U_X} = \boldsymbol{C}\boldsymbol{U_Q}\boldsymbol{C}^\top$ provides a full covariance matrix $\boldsymbol{U_X}$, containing all covariances $u(X_i, X_j)$.
- The matrix $\boldsymbol{C}$ holds the sensitivity coefficients $\partial f_i / \partial Q_j$

Normalisation of a composition

- The composition of any material, expressed in fractions always adds up to a constant
- Hence, these fractions are always correlated and the uncertainty of the sum of all fractions is always zero
- Suppose we have N raw amount fractions $x_1, \ldots, x_N$
- Then the composition expressed in amount fractions is given by $y_i = x_i / \sum_j x_j$
- From this expression, it is clear that any y_i depends on all x_j
- Other cases lead to other uncertainty structures, but still there are covariances to be taken into account

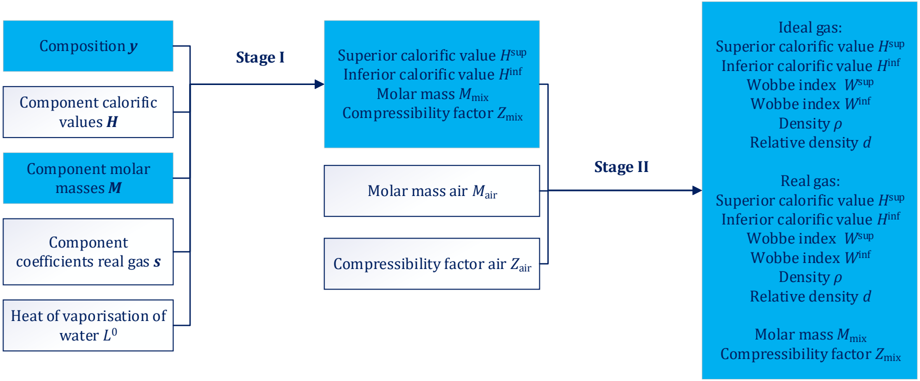
Effect of ignoring covariances in calculation of natural gas properties

Table: Natural gas properties calculated from ISO 6976:2016, example 3

	Ignoring correlations			Including correlations		
	x	$u(x)$	$u_{\mathrm{rel}}(x)$	x	$u(x)$	$u_{\mathrm{rel}}(x)$
H	937.14	0.63	0.067 %	937.14	0.38	0.040 %
$\bar{M}$	18.035	0.014	0.076 %	18.035	0.006	0.035 %
Z	0.997 57	0.000 05	0.005 %	0.997 57	0.000 05	0.005 %

Standard uncertainties calculated for the superior calorific value (H) and molar mass ($\bar{M}$) vastly different; effect of ignoring correlations differs on a case-by-case basis

Multistage measurement models with correlated input quantities: natural gas properties (ISO 6976)



Multistage measurement models with correlated input quantities: natural gas properties (ISO 6976)

- Use of law of propagation of uncertainty from GUM-S2 provides a computationally efficient way to propagate measurement uncertainty
- Appendix B of ISO 6976:2016 provides, without matrix calculus, all expressions to calculate the uncertainties associated with the natural gas properties
- It does not provide expressions for the covariances between natural gas properties, . . .
- . . . but these are often required in subsequent applications (such as calculating measurement errors).
- Duly propagating uncertainty in this area requires substantial mathematical skills, often not available to the readership of the standard.

Autocorrelation

- JCGM 100 presumes that correlations can be dealt with by expressing input quantities into a further set of (uncorrelated) input quantities
- Temperature readings of the thermocouple are *serially* correlated due to the working of the heater of the thermal bath
- A time series analysis is an appropriate way to deal with this form of correlation

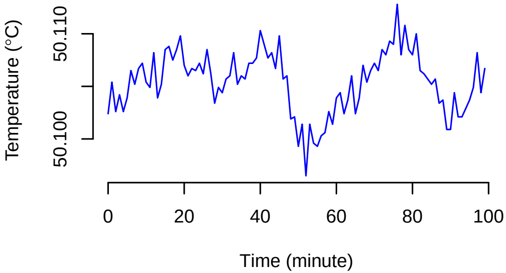


Figure: Readings of a thermocouple measuring the temperature of a thermal bath (JCGM 103:202n)

Autocorrelation (continued)

- In this instance, an auto-regression of order 2 is an appropriate model for the data
- A naive type A evaluation of standard uncertainty would yield $u(t) = 0.0003\,°\mathrm{C}$, whereas an ARIMA(2,0,0) model provides $u(t) = 0.0010\,°\mathrm{C}$, thus about three times larger (see JCGM 103:202n)

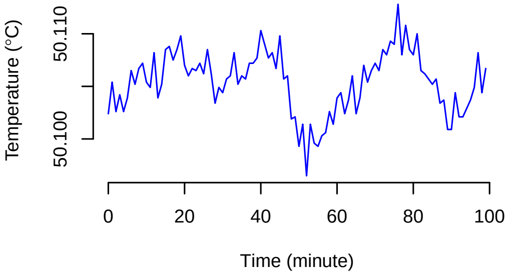


Figure: Readings of a thermocouple measuring the temperature of a thermal bath (JCGM 103:202n)

Regression

- Coefficients of a regression model are almost always correlated
- Consider the generalised linear model $y = \sum_i a_i \phi_i(x)$ where $\phi_i(x)$ is some function of i,
- the measurement model is (often) implicit, relating the a_i to a set of pairwise (x_j, y_j) values and corresponding uncertainties; there can be correlations between pairs of (x_j, x_k), (y_j, y_k) and (x_j, y_j) too.
- Any good software package provides a full covariance matrix
- Use of multiple coefficients in a subsequent calculation requires information of the corresponding covariances
- See also example 9.5 in GUM-S2 (temperature from a resistance thermometer)

Monte Carlo method and Bayesian inference

- Correlations between input quantities modelled by either
  - expressing these in independent quantities and assigning probability density functions
  - assigning a joint probability density function (such as the multivariate normal, t, or Dirichlet distribution)
- Not for all univariate probability density functions there is a multivariate counterpart (e.g., rectangular distribution)
- GUM-S2, section 7.6 provides an expression for obtaining the covariance matrix $\boldsymbol{U}_{\boldsymbol{y}}$ associated with the output vector $\boldsymbol{y}$
- Correlation matrix can be obtained as $\boldsymbol{R}_{\boldsymbol{y}} = \boldsymbol{D}\boldsymbol{U}_{\boldsymbol{y}}\boldsymbol{D}$ where $\boldsymbol{D} = \mathrm{diag}\{u^{-1}(y_1), \ldots, u^{-1}(y_M)\}$
- Approaches for Bayesian inference are very similar (including for assigning (joint) prior probability density functions)

Bayesian EIV regression

- From a Markov Chain Monte Carlo calculation, the coefficients of a straight line (errors-in-variables regression) are obtained
- If the coefficients are to be communicated, the associated covariance matrix should be provided as well
- This covariance matrix can obtained from the MCMC samples of the calculation as described in GUM-S2
- MCMC examples can be extracted from the calculation

Table: Results from fitting the Bayesian model for EIV for standard addition

	mean	sd	2.5 %	97.5 %
a	10.28	8.22	1.16	30.52
b	26.75	1.30	23.69	29.07
$r_{x,1}$	0.01	1.00	−1.95	1.97
$r_{x,2}$	−0.02	1.00	−1.98	1.94
$r_{x,3}$	0.01	1.00	−1.95	1.97
$r_{y,1}$	0.02	0.75	−0.79	1.83
$r_{y,2}$	0.37	0.87	−0.58	2.35
$r_{y,3}$	0.03	1.18	−2.04	2.49
y_1	0.39	0.35	0.04	1.24

Concluding remarks

- Evaluating covariances is an essential part of any uncertainty evaluation; they can increase or decrease the computed uncertainty
- When properly established, a measurement model articulates the dependencies between variables
- It is not always necessary to express the measurand(s) into a set of uncorrelated input quantities
- Working with multivariate methods (GUM-S2) is often the easier choice, but requires familiarity with matrix calculus
- Monte Carlo method and MCMC also provide means to extract information about covariances between output quantities

Acknowledgements

- Prof. Maurice Cox (NPL) for making the request for this contribution and participation in developing the outline
- Merima Čaušević (IMBIH) and Dr. Heleen Meuzelaar (VSL) for their contributions in developing the buoyancy correction for the permeation tube
- Dr. Iris de Krom (VSL) and Gerard Nieuwenkamp (VSL) for their aid in developing the Bayesian evaluation method for standard addition

EMUE Examples of Measurement Uncertainty Evaluation

EMPIR EURAMET

The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States