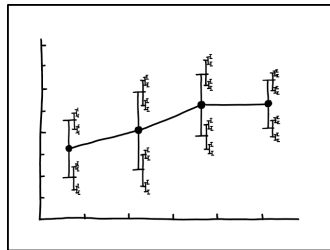


2 Error propagation

❖ transformation of uncertainties under change of variables

- define how to quantify uncertainty. . .
 - ... interval estimate or point estimate?
 - ▶ maximum possible range
 - “tolerance” in engineering
 - ▶ confidence level interval
 - from likelihood fits or bayesian analysis
 - ▶ variance or standard deviation
 - when dealing with random variables
- work out how the uncertainties transform
 - ▶ intervals: transformation of the endpoints
 - ▶ variances: “error propagation”



I DON'T KNOW HOW TO PROPAGATE
ERROR CORRECTLY, SO I JUST PUT
ERROR BARS ON ALL MY ERROR BARS.

source: https://imgs.xkcd.com/comics/error_bars.png

input from Joint Committee for Guides in Metrology →

WG 1 (JCGM 100:2008, Recommendation INC-1 (1980))

- The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories . . . : (A) those which are evaluated by statistical methods, (B) those which are evaluated by other means. . .
- The components in category A are characterized by the estimated variances s_i^2 (or the estimated “standard deviations” s_i) . . .
- The components in category (B) should be characterized by quantities u_j^2 , which may be considered as approximations to the corresponding variances, the existence of which is assumed. The quantities u_j^2 may be treated like variances and the quantities u_j like standard deviations. . .
- . . . The combined uncertainty and its components should be expressed in the form of standard deviations.

JCGM: treat all uncertainties as if they were standard deviations

Transformation $x \rightarrow y = R(x)$

❖ quantities to consider ... mind the hats!

- ▶ $\hat{x}$: measured values, array with components $\hat{x}_i$, $i = 1, \dots, n$
- ▶ x : true values that are estimated by $\hat{x}$
- ▶ $\Delta_k = \hat{x}_k - x_k = O(\Delta)$: actual deviation between measured and the true value
- ▶ $\langle \Delta_k \rangle = 0$: the measured values are unbiased, $\langle \hat{x} \rangle = x$
- ▶ $\langle \Delta_k \Delta_l \rangle = C_{kl} = O(\Delta^2)$: the covariance matrix of $\hat{x}$ is known
- ▶ $y = R(x)$: mapping from x to derived quantities y
- ▶ $\hat{y} = R(\hat{x})$: transformed measurements
- ▶ $C(\hat{y})$: covariance matrix of $\hat{y}$
- ▶ $\hat{C}(\hat{y})$: estimate for the covariance matrix of $\hat{y}$
- ▶ $\partial R(\hat{x})/\partial x$: derivative w.r.t argument x taken at $\hat{x}$

leading order treatment →

Expectation values of transformed variables

Taylor expansion for a single component $\hat{y}_k$ around x yields

$$\hat{y}_k = R_k(x) + \sum_i \frac{\partial R_k(x)}{\partial x_i} \Delta_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 R_k(x)}{\partial x_i \partial x_j} \Delta_i \Delta_j + O(\Delta^3)$$

- ▶ the derivatives are constants
- ▶ by definition: $R_k(x) = y_k$, $\langle \Delta_i \rangle = 0$ and $\langle \Delta_i \Delta_j \rangle = C_{ij}$

▶ expectation value

$$\langle \hat{y}_k \rangle = y_k + B_k + O(\Delta^3) \quad \text{with bias} \quad B_k = \frac{1}{2} \sum_{i,j} \frac{\partial^2 R_k(x)}{\partial x_i \partial x_j} C_{ij}$$

- ▶ for non-linear transformations $\hat{y} = R(\hat{x})$ is biased: $\langle \hat{y} \rangle - y = O(\Delta^2)$
- ▶ the leading order correction is known

problem: the point x where to take the 2nd derivative is not known →

- remember: Taylor expansion shows

$$f(x) = f(\hat{x}) + O(\Delta) \quad \text{with} \quad \Delta = x - \hat{x}$$

- application to bias calculation

$$\begin{aligned}\langle \hat{y}_k \rangle &= y_k + \frac{1}{2} \sum_{i,j} \frac{\partial^2 R_k(x)}{\partial x_i \partial x_j} C_{ij} + O(\Delta^3) \\ &= y_k + \frac{1}{2} \sum_{i,j} \left(\frac{\partial^2 R_k(\hat{x})}{\partial x_i \partial x_j} + O(\Delta) \right) C_{ij} + O(\Delta^3) \\ &= y_k + \frac{1}{2} \sum_{i,j} \frac{\partial^2 R_k(\hat{x})}{\partial x_i \partial x_j} C_{ij} + O(\Delta^3)\end{aligned}$$

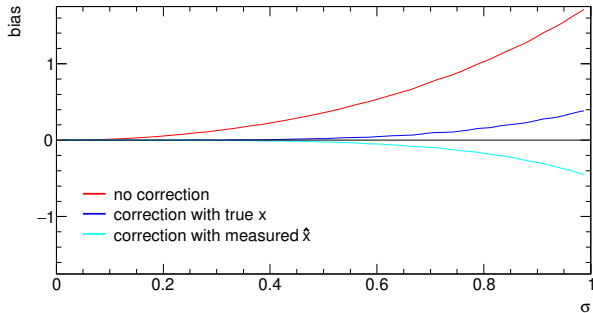
- ▶ uses that $C_{ij} O(\Delta) = O(\Delta^3)$
- ▶ it follows that taking the derivatives at the measured $\hat{x}$ is correct at $O(\Delta^3)$
- ▶ **however**, the leading order bias correction becomes a random variable

problems and exercises

Assume measurements $\hat{x}$, which are normally distributed around $x = 1$ with standard deviation σ . Consider the derived variable $\hat{y} = e^{\hat{x}}$. Determine numerically, with 10^6 values $\hat{x}$ for each configuration, ...

- the bias $\langle \hat{y} \rangle - e$ as a function of σ for $\sigma \in [0, 1]$
- the bias after subtracting the leading order correction B with ...
 - ▶ (a) derivatives taken at the true value $x = 1$: $B = e\sigma^2/2$
 - ▶ (b) derivatives taken at the measured values $\hat{x}$: $B = e^{\hat{x}}\sigma^2/2$

❖ numerical results



- ▶ correction with true x too small, correction with measured $\hat{x}$ over-compensates
- ▶ in both cases significant improvement
- ▶ leading order correction insufficient for large non-linearities

Transformation of the covariance matrix

❖ estimate $C(\hat{y}) = \langle (\hat{y}_k - y_k)(\hat{y}_l - y_l) \rangle$ of the transformed measurements

expansion of $\hat{y}_k$ around true values x : $\hat{y}_k = y_k + \sum_i \frac{\partial R_k(x)}{\partial x_i} \Delta_i + O(\Delta^2)$

- ▶ valid for $\hat{y}$ with and without bias correction since the bias is $O(\Delta^2)$
- ▶ use that the derivatives are constants and that $\langle \Delta_i \Delta_j \rangle = C_{ij}$

$$\begin{aligned}\hat{C}_{kl}(\hat{y}) &= \left\langle \sum_i \frac{\partial R_k(x)}{\partial x_i} \Delta_i \sum_j \frac{\partial R_l(x)}{\partial x_j} \Delta_j \right\rangle + O(\Delta^3) = \sum_{ij} \frac{\partial R_k(x)}{\partial x_i} \frac{\partial R_l(x)}{\partial x_j} C_{ij} + O(\Delta^3) \\ &= \sum_{ij} \left(\frac{\partial R_k(\hat{x})}{\partial x_i} + O(\Delta) \right) \left(\frac{\partial R_l(\hat{x})}{\partial x_j} + O(\Delta) \right) C_{ij} + O(\Delta^3) \\ &= \sum_{ij} \frac{\partial R_k(\hat{x})}{\partial x_i} \frac{\partial R_l(\hat{x})}{\partial x_j} C_{ij} + O(\Delta^3)\end{aligned}$$

- ▶ at $O(\Delta^3)$ taking the derivatives at $\hat{x}$ is fine, but makes $\hat{C}(\hat{y})$ a random object

❖ result in matrix notation

under a transformation $y = R(x)$ the covariance matrix transforms as

$$C(y) = M(x) \cdot C(x) \cdot M^T(x)$$

with Jacobian $M(x)$ and matrix elements $M_{ij} = \frac{\partial R_i}{\partial x_j}$.

- ▣ derivatives taken at $\hat{x}$
- ▣ if $M(x)$ can be inverted then no information is lost in the transformation
- ▣ when chaining transformations one finds:

$$y = S(R(x)) \quad \text{and} \quad M_{ij} = \sum_{k=1}^n \frac{\partial S_i}{\partial R_k} \frac{\partial R_k}{\partial x_j} \quad \text{or} \quad M = M(R) \cdot M(x)$$

- ▶ linear error propagation is consistent and accurate to $O(\Delta^3)$
- ▶ identical results if a transformation is done in one or in several steps

problems and exercises

Consider two uncorrelated gaussian random variables x and y with mean values $\langle x \rangle = 2$, $\langle y \rangle = 1$ and standard deviation $\sigma = 0.1$. Two derived variables are $p = x \cdot y$ and $q = x/y$, with the true (co)variances $C_{pp} \approx 0.0501$, $C_{pq} \approx -0.0313$ and $C_{qq} \approx 0.0538$.

Generate 10^6 pairs of measurements and use error propagation to determine for each pair the estimated covariance matrix elements. Compare the distributions of those estimates to the respective true values.

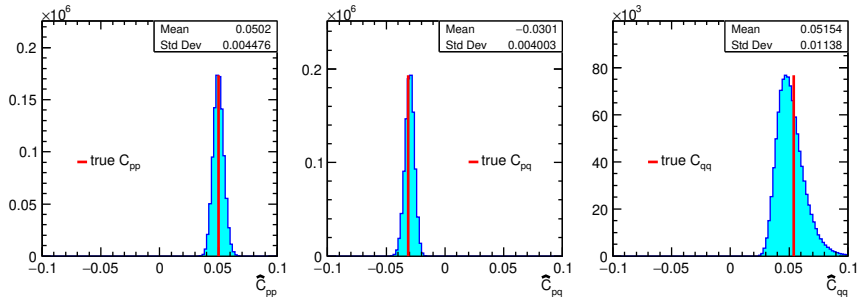
The covariance matrix of x and y is σ^2 times the unit matrix. Error propagation then gives

$$C = \sigma^2 M M^T \quad \text{with} \quad M = \begin{pmatrix} y & x \\ 1/y & -x/y^2 \end{pmatrix}$$

and thus

$$C_{pp} = \sigma^2(x^2 + y^2) \quad , \quad C_{pq} = \sigma^2 \left(1 - \frac{x^2}{y^2} \right) \quad \text{and} \quad C_{qq} = \frac{\sigma^2}{y^2} \left(1 + \frac{x^2}{y^2} \right)$$

► distributions



- (co)variance estimates are random variates that scatter around the true value
- the scatter is of $O(\sigma^3)$
- the average of the estimates approaches the true values

Example for error propagation in N dimensions

❖ error of the mean value of N IID values x_i

- covariance matrix of the measurements with σ the standard deviation of each x_i

$$C_{ij} = \sigma^2 \delta_{ij}$$

- mean value (a linear function of the measurements)

$$y = \frac{1}{N} \sum_{i=1}^N x_i$$

- variance of the mean value (exact result)

$$\sigma^2(y) = \sum_{i,j=1}^N \frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j} C_{ij} = \sum_{i,j=1}^N \frac{1}{N^2} \sigma^2 \delta_{ij} = \frac{\sigma^2}{N^2} \sum_{i=1}^N 1 = \frac{\sigma^2}{N}$$

→ for IID measurements the error of the arithmetic mean is $\sigma/\sqrt{N}$

Summary on error propagation

- transformations of covariance matrices are ...
 - ▶ exact for linear and approximate for non-linear transformations
 - ▶ always consistent when chaining transformations

$$C_{kl}(y) = \sum_{i,j=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial y_l}{\partial x_j} C_{ij}(x) \quad \text{for scalar } y \text{ and uncorrelated } x_i: \quad \sigma^2(y) = \sum_{i=1}^n \left(\frac{\partial y}{\partial x_i} \right)^2 \sigma^2(x_i)$$

- non-linear transformations
 - ▶ induce a bias B of $O(\Delta^2)$
 - ▶ the leading order bias correction is known and recommended to be applied

$$B_k = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 y_k}{\partial x_i \partial x_j} C_{ij}(x) \quad \text{scalar } x \text{ and } y: \quad B = \frac{1}{2} \frac{\partial^2 y}{\partial x^2} \sigma^2(x)$$

- all derivatives can be taken at the measured values
- approximate result are accurate with errors of $O(\Delta^3)$