

Propagation of Uncertainty

Two routes to the root-sum-of-squares formula, and how they fit together

Lecture Notes

Overview

For a derived quantity $y = f(x_1, \dots, x_n)$ with uncorrelated inputs, the standard propagation formula is the root-sum-of-squares (RSS) of sensitivity-weighted uncertainties:

$$u_y = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} u_{x_i} \right)^2}. \quad (1)$$

Two derivations are commonly quoted:

1. Compute the variance of the output and assume the inputs are uncorrelated.
2. Taylor-expand f about the input means and drop higher-order terms.

Both are valid, but they are not two independent proofs. For a nonlinear f , Route 1 cannot be completed on its own, and Route 2 does not produce a variance on its own. Each supplies one of the two assumptions behind (1):

Step	Assumption introduced
Truncated Taylor expansion (Route 2)	f is nearly <i>linear</i> over the input spread
Variance of a linear combination (Route 1)	inputs are <i>uncorrelated</i> (diagonal covariance)

These notes derive each route, then show how they combine.

1 Setup

Model each input as a random variable X_i with

$$E[X_i] = \mu_i, \quad \text{Var}[X_i] = \sigma_i^2, \quad \text{Cov}[X_i, X_j] = \sigma_{ij},$$

and the output as $Y = f(X_1, \dots, X_n)$. In measurement practice μ_i is replaced by the best estimate $\hat{x}_i$ and σ_i by the standard uncertainty u_{x_i} . Write the deviations from the mean as

$$\delta_i = X_i - \mu_i, \quad E[\delta_i] = 0, \quad E[\delta_i^2] = \sigma_i^2, \quad E[\delta_i \delta_j] = \sigma_{ij}.$$

The goal is $u_y^2 = \text{Var}[Y] = E[(Y - E[Y])^2]$.

2 Route 1: The variance of the output

1a. Start from the definition

$$\text{Var}[Y] = \mathbb{E}[(Y - \mathbb{E}[Y])^2].$$

To evaluate this we need $Y - \mathbb{E}[Y]$ in terms of the δ_i . That is only straightforward when f is linear, so take that case first.

1b. Linear f

Let

$$Y = a_0 + \sum_{i=1}^n a_i X_i.$$

By linearity of expectation, $\mathbb{E}[Y] = a_0 + \sum_i a_i \mu_i$, so

$$Y - \mathbb{E}[Y] = \sum_{i=1}^n a_i (X_i - \mu_i) = \sum_{i=1}^n a_i \delta_i.$$

1c. Square and take the expectation

Squaring a sum gives a double sum:

$$(Y - \mathbb{E}[Y])^2 = \left(\sum_i a_i \delta_i \right) \left(\sum_j a_j \delta_j \right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \delta_i \delta_j.$$

Taking expectations term by term, $\mathbb{E}[\delta_i \delta_j] = \sigma_{ij}$ (with $\sigma_{ii} = \sigma_i^2$):

$$\text{Var}[Y] = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \sigma_{ij} = \sum_{i=1}^n a_i^2 \sigma_i^2 + 2 \sum_{i < j} a_i a_j \sigma_{ij}.$$

This is *exact* for linear f and needs no assumption about the shape of the input distributions – only their second moments.

1d. Assume uncorrelated inputs

If $\sigma_{ij} = 0$ for $i \neq j$ (diagonal covariance), the cross terms vanish:

Result 1: linear f , uncorrelated inputs

$$\text{Var}[Y] = \sum_{i=1}^n a_i^2 \sigma_i^2 \quad \implies \quad u_y = \sqrt{\sum_{i=1}^n (a_i u_{x_i})^2}.$$

1e. Where Route 1 stops

For a nonlinear f , $\text{Var}[f(X_1, \dots, X_n)]$ depends on the full joint distribution of the inputs, not just their means and covariances, and in general has no closed form. There are also no coefficients a_i to plug in. Route 1 by itself therefore gives the RSS formula only for linear functions. To go further we need to make f linear.

3 Route 2: Taylor expansion

2a. Expand about the means

Expand f about $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$:

$$Y = f(\boldsymbol{\mu}) + \sum_{i=1}^n c_i \delta_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n c_{ij} \delta_i \delta_j + \dots, \quad c_i = \left. \frac{\partial f}{\partial x_i} \right|_{\boldsymbol{\mu}}, \quad c_{ij} = \left. \frac{\partial^2 f}{\partial x_i \partial x_j} \right|_{\boldsymbol{\mu}}.$$

The c_i are the *sensitivity coefficients*.

2b. Truncate at first order

If the deviations δ_i are small compared with the scale over which f curves, the quadratic and higher terms are negligible:

$$Y \approx f(\boldsymbol{\mu}) + \sum_{i=1}^n c_i \delta_i.$$

This is the tangent-plane approximation (Figure 1). It has exactly the form of the linear model in 1b, with $a_0 = f(\boldsymbol{\mu}) - \sum_i c_i \mu_i$ and $a_i = c_i$.

2c. The mean of the output

Since $E[\delta_i] = 0$,

$$E[Y] \approx f(\boldsymbol{\mu}), \quad Y - E[Y] \approx \sum_{i=1}^n c_i \delta_i.$$

To first order, the best estimate of y is f evaluated at the best estimates of the inputs.

2d. Where Route 2 stops

Step 2b gives a *linear approximation of the output*, not its variance. To get a number for u_y you must still square and average – which is Route 1.

4 Putting the two routes together

Substitute 2c into 1c ($a_i \rightarrow c_i$):

$$\text{Var}[Y] \approx E\left[\left(\sum_i c_i \delta_i\right)^2\right] = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \sigma_{ij}.$$

Result: the law of propagation of uncertainty

General (correlated) form:

$$u_y^2 = \sum_{i=1}^n c_i^2 u_{x_i}^2 + 2 \sum_{i < j} c_i c_j u_{x_i x_j} = \mathbf{c}^\top \boldsymbol{\Sigma}_x \mathbf{c}, \quad c_i = \frac{\partial f}{\partial x_i}.$$

Uncorrelated inputs ($\boldsymbol{\Sigma}_x$ diagonal):

$$u_y = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} u_{x_i} \right)^2}.$$

The division of labour is:

	Route 2 (Taylor)	Route 1 (variance)
Does	replaces f by its tangent plane	variance of a linear combination
Assumes	higher-order terms negligible	covariance known; diagonal for RSS
Exact when	f is linear	always (for linear f)
Supplies	the coefficients $c_i = \partial f / \partial x_i$	the sum of squares

The “error” version of the same argument. Textbooks that work with a single error rather than a variance write $\delta y \approx \sum_i c_i \delta x_i$ (Route 2), square it to get $\delta y^2 \approx \sum_i \sum_j c_i c_j \delta x_i \delta x_j$, and then argue that the cross terms average to zero for independent errors. The “averaging” step is taking an expectation, so this is the same combination of routes written with different notation.

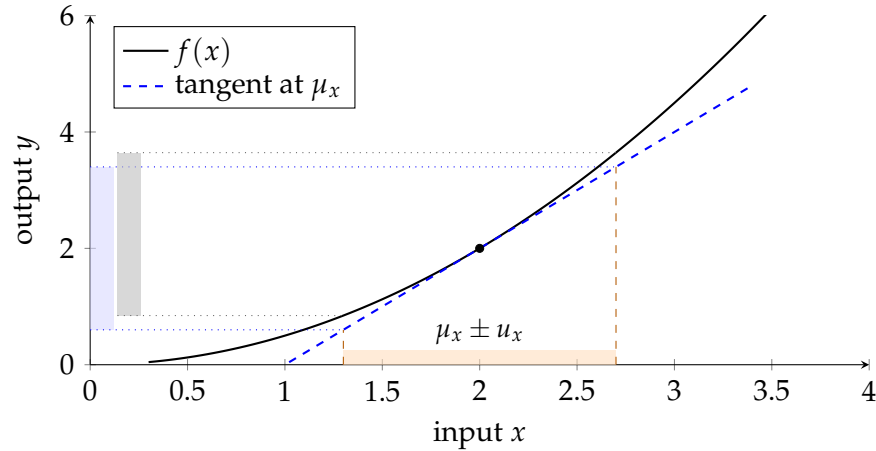


Figure 1: First-order propagation for $f(x) = x^2/2$ at $\mu_x = 2$, $u_x = 0.7$. The input band (orange) maps through the tangent line to a symmetric output band of half-width $|f'(\mu_x)| u_x = 1.4$ (blue). The true map (grey) is asymmetric and shifted upward; the difference is what the dropped second-order terms describe.

5 When first order is not enough

Keeping the second-order term in 2a and taking expectations gives a correction to the mean:

$$E[Y] \approx f(\mu) + \frac{1}{2} \sum_i \sum_j c_{ij} \sigma_{ij}.$$

For independent, Gaussian (or symmetric) inputs, the GUM (JCGM 100, §5.1.2) gives the next terms in the variance:

$$u_y^2 \approx \sum_i c_i^2 u_{x_i}^2 + \sum_i \sum_j \left[\frac{1}{2} \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 + \frac{\partial f}{\partial x_i} \frac{\partial^3 f}{\partial x_i \partial x_j^2} \right] u_{x_i}^2 u_{x_j}^2.$$

Example 1: $y = x^2$, $X \sim \mathcal{N}(\mu, \sigma^2)$

Here $f' = 2\mu$, $f'' = 2$, $f''' = 0$.

	First order	Second order	Exact
$E[Y]$	μ^2	$\mu^2 + \sigma^2$	$\mu^2 + \sigma^2$
$\text{Var}[Y]$	$4\mu^2\sigma^2$	$4\mu^2\sigma^2 + 2\sigma^4$	$4\mu^2\sigma^2 + 2\sigma^4$

The relative size of the neglected term is $2\sigma^4/(4\mu^2\sigma^2) = \frac{1}{2}(\sigma/\mu)^2$, so first order is fine when $u_x \ll |\mu|$. At $\mu = 0$ the first-order result is $u_y = 0$, which is plainly wrong: the slope vanishes and all the spread comes from curvature.

Example 2: $y = x_1 x_2$, independent inputs

$c_1 = \mu_2$, $c_2 = \mu_1$, $c_{12} = 1$. First order gives

$$u_y^2 \approx \mu_2^2 u_1^2 + \mu_1^2 u_2^2 \iff \left(\frac{u_y}{y} \right)^2 \approx \left(\frac{u_1}{x_1} \right)^2 + \left(\frac{u_2}{x_2} \right)^2,$$

the familiar “add relative uncertainties in quadrature” rule for products. The second-order term adds $u_1^2 u_2^2$, which is the exact result, and is negligible when both relative uncertainties are small.

Rule of thumb

First order is adequate when, for each input, the change in slope across the input’s spread is small compared with the slope itself: $|c_{ii}| u_{x_i} \ll |c_i|$. When it is not, use the second-order terms above or propagate distributions by Monte Carlo (GUM Supplement 1, JCGM 101).

6 Remarks

Correlation matters in both directions. Positive correlation can inflate or cancel uncertainty depending on the signs of the c_i . For $y = x_1 - x_2$ measured with the same instrument, a shared calibration offset gives $u_{x_1 x_2} = \rho u_1 u_2$ with $\rho \approx 1$, and

$$u_y^2 = u_1^2 + u_2^2 - 2\rho u_1 u_2 \rightarrow (u_1 - u_2)^2,$$

so the common offset cancels. Applying RSS here would overstate u_y . For $y = x_1 + x_2$ the same correlation makes RSS understate it.

Evaluation point. The c_i are evaluated at the best estimates $\hat{x}_i$, since μ_i is unknown. This is part of the first-order approximation.

From u_y to an interval. The propagated u_y is a *standard* uncertainty. To report $y = \hat{y} \pm U$ with coverage γ , use $U = k u_y$. If some u_{x_i} come from small samples (Type A, $\nu_i = N_i - 1$), the coverage factor is the Student- t quantile $k = t_{\gamma, \nu_{\text{eff}}}$, with the effective degrees of freedom from the Welch–Satterthwaite formula

$$\nu_{\text{eff}} = \frac{u_y^4}{\sum_{i=1}^n \frac{(c_i u_{x_i})^4}{\nu_i}}.$$

This connects propagation back to the z - and t -interval derivations.