CHE384, From Data to Decisions: Measurement, Uncertainty, Analysis, and Modeling

Lecture 79
Propagation of Uncertainty

Chris A. Mack
Adjunct Associate Professor

http://www.lithoguru.com/scientist/statistics/

The Measurement Model

- Consider a measurement model

Measurement Output $\longrightarrow y = f(x_1, x_2, \ldots)$

- Some of the inputs represent the thing we want to measure, but some are nuisance variables
  - Ex: Measure length with a steel ruler; temperature is a nuisance variable
- How do variations in the x_i propagate to variations in y?
  - Called the propagation of uncertainty

Propagation of Uncertainty

- There are three common ways to propagate uncertainty from inputs to an output
  - Propagation of pdfs (often difficult to do)
  - Taylor Series (most common, but easy to do wrong)
  - Monte Carlo simulations (often best approach for complex measurement models)

Propagation of PDFs (1D example)

- Consider $Y = f(X)$ where X is a random variable with known pdf $P_X(x)$. What is the pdf of Y?

$$P_Y(y) = \left|\frac{dx}{dy}\right| P_X(x)$$

$$where\ x = f^{-1}(y)$$

- Note: $f^{-1}(y)$ can have multiple roots

Taylor Series Approach

- Expand the function as a Taylor series

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

$$Let\ \Delta y = f(x) - f(a) \qquad \Delta x = x - a \qquad a = \bar{x}$$

$$\Delta y = \frac{df}{dx}\Delta x + \frac{1}{2!}\frac{d^2 f}{dx^2}\Delta x^2 + \cdots + \frac{1}{n!}\frac{d^n f}{dx^n}\Delta x^n$$

- For K different input variables,

$$\Delta y = \sum_{j=0}^{\infty}\left(\frac{1}{j!}\left[\sum_{k=1}^{K}\Delta x_k \frac{\partial}{\partial x_k}\right]^j f\right)$$

Taylor Series Approach

- Turn Δx and Δy into σ_x and σ_y

Third moment about the mean

$$E[\Delta y^2] = \sigma_y^2 = \left(\frac{df}{dx}\right)^2 \sigma_x^2 + \left(\frac{df}{dx}\right)\left(\frac{d^2 f}{dx^2}\right)\mu_3 + \left[\frac{1}{4}\left(\frac{d^2 f}{dx^2}\right)^2 + \frac{1}{3}\left(\frac{df}{dx}\right)\left(\frac{d^3 f}{dx^3}\right)\right]\mu_4 + \cdots$$

- The partial derivatives are evaluated at the mean value of x

Taylor Series Approach

- For two input variables x_1 and x_2,

$$\sigma_y^2 = \left(\frac{df}{dx_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{df}{dx_2}\right)^2 \sigma_{x_2}^2 + 2\left(\frac{df}{dx_1}\right)\left(\frac{df}{dx_2}\right) cov(x_1, x_2) + \cdots$$

  - When the input errors are small and the slopes are not, we can ignore higher order terms
- When x_1 and x_2 are independent,

$$\sigma_y^2 = \left(\frac{df}{dx_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{df}{dx_2}\right)^2 \sigma_{x_2}^2$$

Some Examples

- Assume small, independent errors
  - Keep only linear terms, ignore covariance
- $y = ax_1 + bx_2, \;\; \sigma_y^2 = a^2\sigma_{x_1}^2 + b^2\sigma_{x_2}^2$
- $y = x_1 x_2$, assuming $\bar{x}_i$ >> σ_{x_1}

$$\left(\frac{\sigma_y}{\bar{y}}\right)^2 = \left(\frac{\sigma_{x_1}}{\bar{x}_1}\right)^2 + \left(\frac{\sigma_{x_2}}{\bar{x}_2}\right)^2$$

- $y = \ln(ax), \;\; \sigma_y^2 = \sigma_x^2/\bar{x}^2$

Using a Micrometer

(images from Wikimedia Commons)

Micrometer Example

- Sources of uncertainty:
  - Scale (calibration) error: 3 μm, assumed to be uniformly distributed (s = range/$\sqrt{3}$ = 1.73 μm)
  - Zeropoint error: 2 μm (s = 1.15 μm)
  - Anvil parallelism: s = 0.58 μm
  - Temperature different between micrometer and object: 3 ºC, leading to s = 0.61 μm
  - Measurement repeatability: s = 2 μm
- Combined standard uncertainty = 3.0 μm

Lecture 79: What have we learned?

- What are the major difficulties of Bayesian regression?
- Explain how frequentist and Bayesian regression concepts can merge in a “hybrid” form
- Explain how this “hybrid” form relates to ridge regression