

Glossary of important terms for GW1876

April 21, 2018

Notes on mathematics symbology

It would be really nice if a totally consistent set of symbols was used in the mathematics of parameter estimation and uncertainty analysis. But....no such luck! So many researchers in different fields over a long time have contributed to the work used in this class. As a result, this glossary is an attempt to highlight some general symbology and clarify some terms.

First of all, though, one thing that is consistent (mostly!) is the general linear algebra notation used throughout the class.

scalar values Lowercase, non-bold font indicates a scalar (single) value: x , y , z

vectors lkjh

Glossary of terms and equations

Parameters Variable input values for models, typically representing system properties and forcings. Values to be estimated in the history matching process. Typically identified as k , p , or x (\mathbf{k} , \mathbf{p} or \mathbf{x} for multiple parameters in a vector).

Observation Measured system state values. These values are used to compare with model outputs collocated in space and time. The term is often used to mean *both* field measurements and outputs from the model. When referring to a measured value, observations are typically identified by the variables y or o (\mathbf{y} or \mathbf{o} for multiple parameters in a vector)

Modeled Equivalent A modeled value collocated in time and space with an observation. There are various ways to identify a single or multiple modeled equivalent values (and, to make things confusing, they are often *also* called "observations")

Single values

1. $f(x)$
2. $X(\beta)$
3. $M(p)$

Multiple values

1. $\mathbf{X}\beta$
2. \mathbf{Mp}
3. **NOBS** Number of observations/simulated equivalents in the inverse model setup
4. **NPAR** Number of adjustable input parameters in the inverse model setup

Forecasts Model outputs for which field observations are not available. Typically these values are simulated under an uncertain future condition.

Phi Objective function, defined as the weighted sum of squares of residuals. Phi (aka Φ) is typically calculated as

$$\Phi = \sum_{i=1}^n \left(\frac{y_i - f(x_i)}{w_i} \right)^2 \quad \text{or} \quad \Phi = (\mathbf{y} - \mathbf{Jx})^T \mathbf{Q}^{-1} (\mathbf{y} - \mathbf{Jx}) \quad (1)$$

Residuals The difference between observation values and modeled equivalents $r_i = y_i - f(x_i)$

Sensitivity The incremental change of an observation (modeled equivalent, actually) due to an incremental change in a parameter. Typically expressed as a finite-difference approximation of a partial derivative: $\frac{\partial y}{\partial x}$

Jacobian Matrix A matrix of the sensitivity of all observations in an inverse model to all parameters. This is often shown as a matrix by various names **X**, **J**, or **H**. Each element of the matrix is a single sensitivity value $\frac{\partial y_i}{\partial x_j}$ for $i \in NOBS, j \in NPAR$

Regularization A preferred condition pertaining to parameters, the deviation from which, elicits a penalty added to the objective function. This serves as a balance between the level of fit or “measurement Phi” (Φ_M) and the coherence with soft knowledge/previous conditions/prior knowledge/regularization (Φ_R). These terms can also be interpreted as the likelihood function and prior distribution in Bayes’ theorem (see below)

PHIMLIM A PEST input parameter the governs the strength with which regularization is applied to the objective function. A high value of PHIMLIM indicates a strong penalty for deviation from preferred parameter conditions while a low value of PHIMLIM indicates a weak penalty. The reason this “dial” is listed as a function of PHIM (e.g. Φ_M) is because it can then be interpreted as a limit on how well we want to fit the observation data.

FOSM *fill this in*

Gaussian (multivariate) The equation for Gaussian (Normal) distribution for a single variable (x) is

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \quad (2)$$

where μ is the mean and σ is the standard deviation The equation for a multivariate Gaussian for a vector of k variables (\mathbf{x}) is

$$f(\mathbf{x}|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} e^{-\frac{1}{2} (\mathbf{x}-\mu)^T \Sigma^{-1} (\mathbf{x}-\mu)} \quad (3)$$

where μ is a k -length vector of mean values, Σ is the covariance matrix, and $|\Sigma|$ is the determinant of the covariance matrix

Weight A value which a residual is divided by when constructing the sum of squared residuals. In principal, $w \approx \frac{1}{\sigma}$ where σ is an approximation of the expected error between model output and collocated observation values. While the symbol σ implies a standard deviation, it is important to note that measurement error only makes up a portion of this error. Other aspects such as structural error (e.g. inadequacy inherent in all models to perfectly simulate the natural world) also contribute to this expected level of error.

Weight Covariance matrix (correlation matrix)

Parametric uncertainty

Measurement noise

Structural (model) error

Monte Carlo Ensemble

Bayes’ Theorem

Posterior (multivariate distribution)

Schur Complement

Prior (multivariate distribution)

Likelihood (multivariate distribution)