Statistical Techniques for Parameter Estimation

``It will be to little purpose to tell my Reader, of how great Antiquity the playing of dice is.'' John Arbuthnot, Preface to *Of the Laws of Chance*, 1692.

Statistical Model

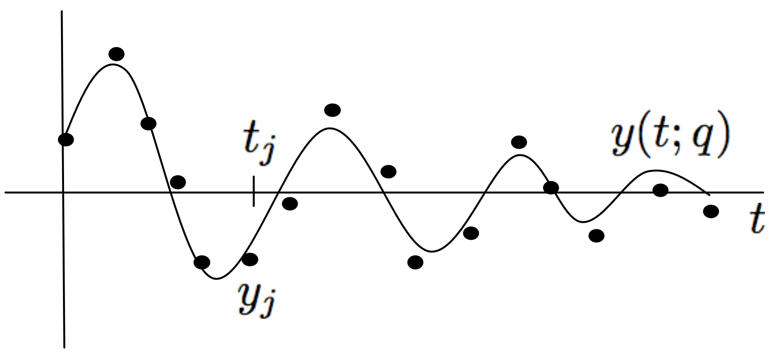
Observation Process: There are errors and noise in the model so consider

$$y_j = y(t_j; q) + \varepsilon_j$$

where y_j denotes data collected at times $t_j, j = 1, \cdots, n$ and $y(t_j; q)$ are observed model values specified by

$$\frac{dx}{dt} = Ax(t; q) + F(t)$$

$$y(t; q) = Cx(t; q)$$



Strategy: Assume ε_j are random variables and that y_j is a realization of a random variable Y_j

estimates $\hat{q}$ and $\hat{\sigma}^2$ of the mean and variance.

Statistical Model:

$$Y_j = y(t_j; q_0) + \varepsilon_j \qquad \text{(Discrete)}$$

or

$$Y(t) = y(t; q_0) + \varepsilon(t) \qquad \text{(Continuous)}$$

Strategy: Treat Q as a random variable for which we seek to find estimates $\hat{q}$ and $\hat{s}^2$ of the true parameter q_0 and the variance of ε_j.

Nonlinear Ordinary Least Squares (OLS) Results

Motivation: See the linear theory in Lecture 10 on "Aspects of Probability and Statistics." These results are analogous to those summarized on Slides 30-31.

Reference: See the paper "An inverse problem statistical methodology summary" by H.T. Banks, M. Davidian, J.R. Samuels, Jr., and K.L. Sutton in the References

Assumptions: $E(\varepsilon_j) = 0$, ε_j iid with $\mathrm{var}(\varepsilon_j) = \sigma_0^2$

Least Squares Estimator and Estimate: Note that $E(Q) = q_0$

$$Q = \arg\min_{q \in \mathcal{Q}} \sum_{j=1}^{n} [Y_j - y(t_j; q)]^2$$

$$\hat{q} = \arg\min_{q \in \mathcal{Q}} \sum_{j=1}^{n} [y_j - y(t_j; q)]^2$$

Variance Estimator and Estimate:

$$S^2 = \frac{1}{n-p} \sum_{j=1}^{n} [Y_j - y(t_j; Q)]^2$$

$$\hat{s}^2 = \frac{1}{n-p} \sum_{j=1}^{n} [y_j - y(t_j; \hat{q})]^2$$

Nonlinear Ordinary Least Squares (OLS) Results

Covariance Estimator and Estimate:

$$\text{cov}(Q) = \sigma_0^2 \left[\chi^T(q_0)\chi(q_0)\right]^{-1}$$

$$\widehat{\text{cov}(Q)} = s_0^2 \left[\chi^T(\hat{q})\chi(\hat{q})\right]^{-1}$$

Here the sensitivity matrix is defined by

$$\chi_{jk}(q) = \frac{\partial y(t_j;q)}{\partial q_k} = C\frac{\partial x(t_j;q)}{\partial q_k} \approx \frac{y(t_j;q+h_k) - y(t_j;q)}{|h_k|}$$

where h_k is a $p-$vector with a nonzero entry only in the k^{th} component

Nonlinear Ordinary Least Squares (OLS) Results

Statistical Properties: (if $\varepsilon_k \sim N(0, \sigma_0^2)$ or in the limit $n \to \infty$)

- $Q \sim N_p\left(q_0, \sigma_0^2(\chi^T(q_0)\chi(q_0))^{-1}\right)$

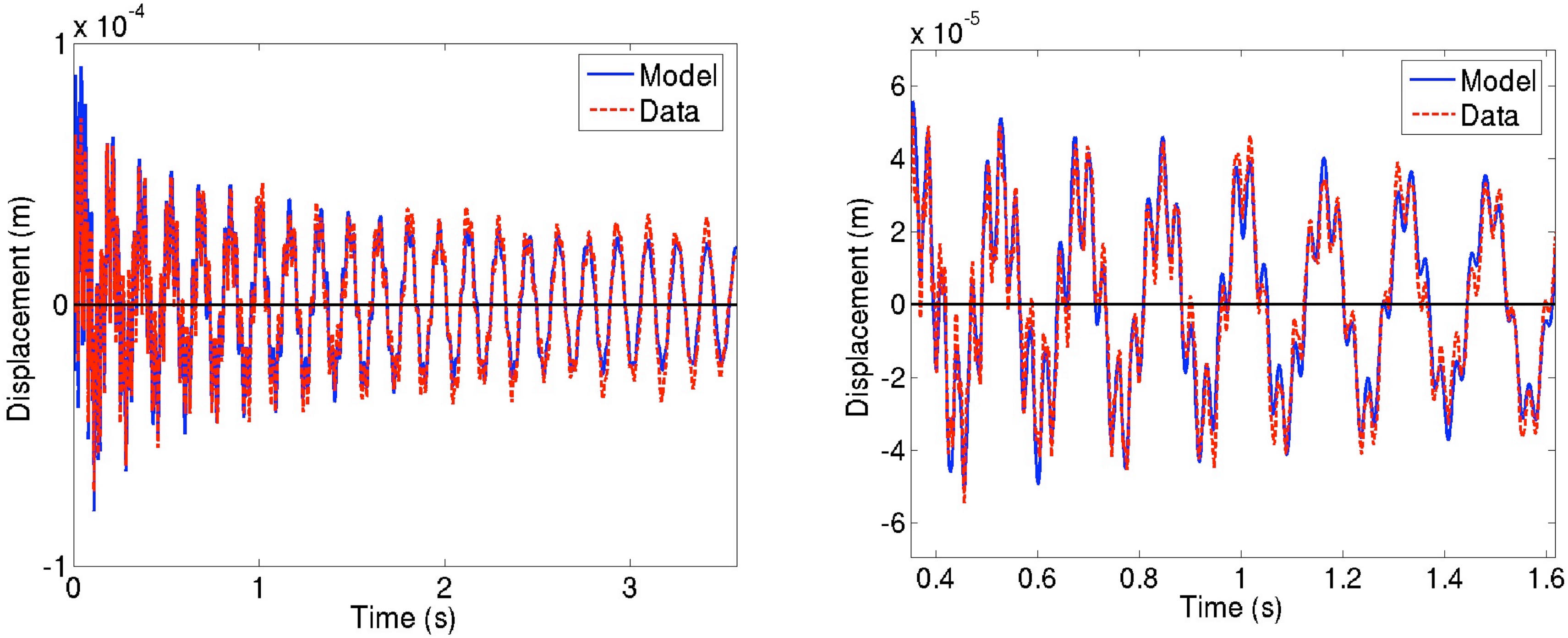
- The $(1-\alpha) \times 100\%$ confidence interval for $\hat{q}_k$ is

$$(\hat{q}_k - t_{n-p,1-\alpha/2} SE_k(\hat{q}), \hat{q}_k + t_{n-p,1-\alpha/2} SE_k(\hat{q}))$$

where $SE_k(\hat{q}) = \sqrt{\widehat{\text{cov}(Q)}_{kk}}\ ,\ k = 1, \cdots, p$

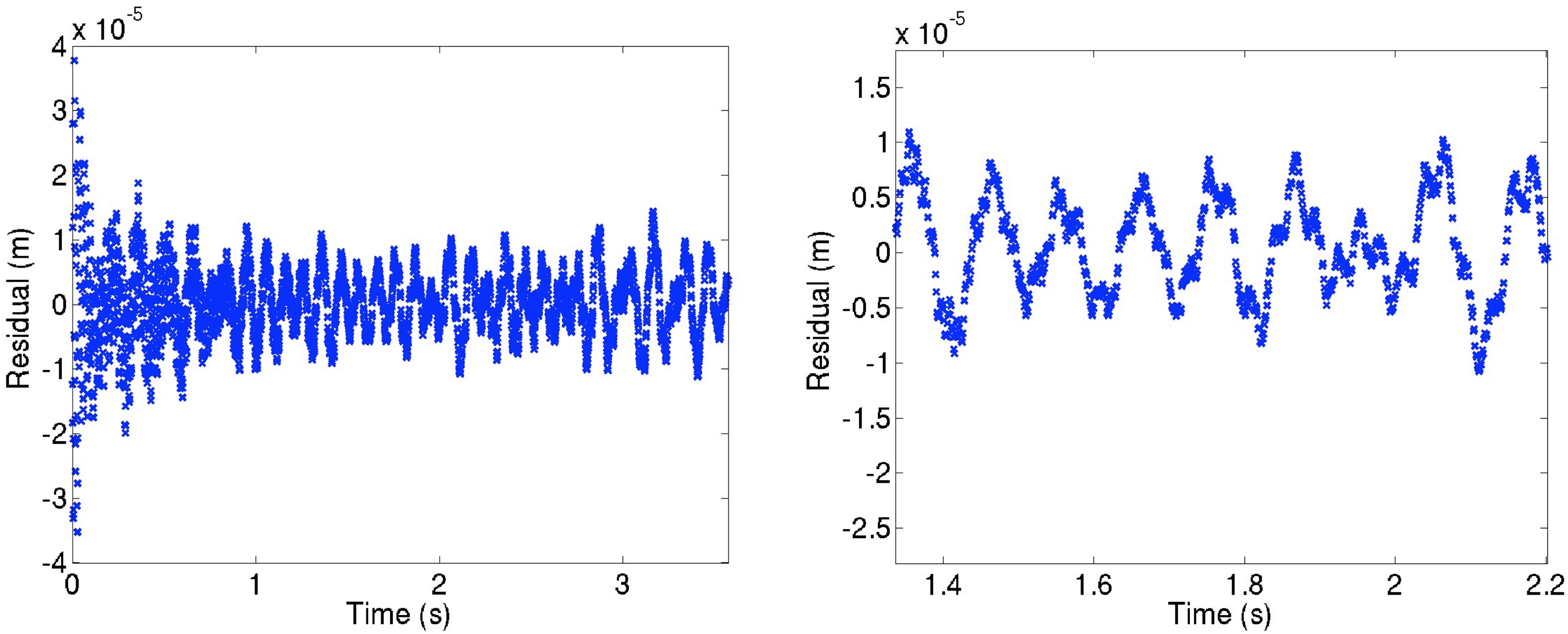
Generalized Least Squares (GLS) Motivation

Model Fit to Beam Data:



Generalized Least Squares (GLS) Motivation

Residual Plots:



Observation: Residuals (and hence errors) are not iid.

Strategy: Consider a statistical model where the errors are model-dependent

$$Y_j = y(t_j; q_0)(1 + \varepsilon_j)$$

Generalized Least Squares (GLS)

Note: Under the assumption that $E(\varepsilon_j) = 0$ and var$(\varepsilon_j) = \sigma_0^2$, it follows that

$$E(Y_j) = y(t_j; q_0)$$

$$\text{var}(Y_j) = \sigma_0^2 y^2(t_j; q_0)$$

Idea: Consider a weighted least squares estimator

$$Q_{GLS} = \arg\min_{q \in \mathcal{Q}} \sum_{j=1}^{n} w_j [Y_j - y(t_j; q)]^2$$

where

$$w_j = y^{-2}(t_j; Q_{GLS})$$

Algorithm: See Banks, Davidian, Samuels, Sutton, 2008

Note: The GLS does NOT changes the properties of the underlying model