Profile-Based Parametric Sensitivity Measures for Nonlinear Regression Models

H. Sulieman, J. McLellan and D.W. Bacon
Department of Chemical Engineering
Queen's University

Outline

- Motivation
- Existing Approaches
- Regression Models
- Profile-Based Sensitivity - Single Response Case
- Illustrative Example - Michaelis-Menten kinetics
- Profile-Based Sensitivity - Multi-Response Case
- Illustrative Example - Dow Chemical regression Benchmark problem
- Summary and Conclusions

Motivation

Parametric sensitivities describe the impact of perturbations in model parameter estimates on predicted responses.

Uses

- identification of influential parameters
- guide further experimentation

Sensitivities are in the context of -

- a specified model formulation
  - structure
  - parameterization
- a specified dataset
  - run conditions
  - designed experiments?

Regression Models

Uniresponse Model

- N runs

$$y_j = \eta_j(\Theta) + \varepsilon_j = f(\mathbf{x}_j, \Theta) + \varepsilon_j \;, \; j = 1, \ldots, N$$

$$\mathbf{y} = \eta(\Theta) + \varepsilon$$

Multiresponse Model

- N runs, M responses

$$y_{nm} = f_m(\mathbf{x}_n, \Theta) + \varepsilon_{nm} \quad , n = 1, \ldots, N; \; m = 1, \ldots, M$$

$$\mathbf{Y} = \mathbf{H}(\Theta) + \mathbf{Z}$$

Existing Approaches

- Marginal sensitivities
  - » first-order derivatives of predicted responses with respect to parameters $\frac{\partial \hat{y}_i}{\partial \theta_j}$
- Monte Carlo simulation
  - » random sampling of model parameters to generate input/output distributions
- FAST - Fourier Amplitude Sensitivity Test - Cukier et al. (1973), Cukier et al. (1978), Saltelli et al. (1999)
  - » impact of perturbations assessed using search curve in parameters
- correlations between parameter values typically ignored in these approaches

Profile-Based Sensitivity Coefficient (PSC)

- motivated by profiling algorithm of Bates and Watts (1988) for producing *profile trace* plots
  - » *profile traces* - plots of conditional estimate of one parameter vs. another parameter, with remaining parameters held at their conditional estimates
  - » indicates both nonlinearity and extent of correlation between parameter estimates
- *prediction parameter transformation*
  - » re-assign parameter to be prediction at a specific point
  - » profile trace of this parameter provides a graphical indication of sensitivity of prediction to perturbations in other parameters
  - » motivates definition of profile-based sensitivity

Profile-Based Sensitivity Coefficient

Definition - Uniresponse Case

(Sulieman, 1998, Sulieman et al., 2001)

» total derivative of the predicted response at a point with respect to a given parameter, with other parameter estimates held at their conditional least squares values

$$PSC_i(\mathbf{x}_0) = \frac{D\eta_0(\theta_i, \Theta_{-i}(\theta_i))}{D\theta_i} = \frac{\partial \eta_0}{\partial \theta_i} + \left.\frac{\partial \eta_0}{\partial \Theta_{-i}}\right|_{\tilde{\Theta}_{-i}} \frac{\partial \tilde{\Theta}_{-i}}{\partial \theta_i}$$

» cf., marginal sensitivity coefficient

$$MSC_i(\mathbf{x}_0) = \frac{\partial \eta_0}{\partial \theta_i}$$

Profile-Based Sensitivity Coefficient - Uniresponse

Structure of PSC

$$PSC_i(\mathbf{x}_0) = \frac{D\eta_0(\theta_i, \Theta_{-i}(\theta_i))}{D\theta_i} = \frac{\partial \eta_0}{\partial \theta_i} + \left.\frac{\partial \eta_0}{\partial \Theta_{-i}}\right|_{\tilde{\Theta}_{-i}} \frac{\partial \tilde{\Theta}_{-i}}{\partial \theta_i}$$

$$= \frac{\partial \eta_0}{\partial \theta_i} - \frac{\partial \eta_0}{\partial \Theta_{-i}} \left(\frac{\partial^2 S}{\partial \Theta_{-i} \partial \Theta'_{-i}} \right)^{-1} \left.\frac{\partial^2 S}{\partial \theta_i \partial \Theta'_{-i}}\right|_{\tilde{\Theta}_{-i}(\theta_i)}$$

$$= MSC(\mathbf{x}_0) + correction\ term$$

» correction term accounts for correlation between parameter estimates, nonlinearity

» uses identity for total derivative of $\left.\frac{\partial S(\theta_i, \Theta_{-i})}{\partial \Theta_{-i}}\right|_{\tilde{\Theta}_{-i}}$

where S is the sum of squares function

Profile-Based Sensitivity Coefficient - Uniresponse

Structure of PSC

$$PSC_i(\mathbf{x}_0) = \frac{D\eta_0(\theta_i, \Theta_{-i}(\theta_i))}{D\theta_i} = \frac{\partial \eta_0}{\partial \theta_i} + \left.\frac{\partial \eta_0}{\partial \Theta_{-i}}\right|_{\tilde{\Theta}_{-i}} \frac{\partial \tilde{\Theta}_{-i}}{\partial \theta_i}$$

$$= \frac{\partial \eta_0}{\partial \theta_i} - \frac{\partial \eta_0}{\partial \Theta_{-i}} \left(\frac{\partial^2 S}{\partial \Theta_{-i} \partial \Theta'_{-i}} \right)^{-1} \left.\frac{\partial^2 S}{\partial \theta_i \partial \Theta'_{-i}}\right|_{\tilde{\Theta}_{-i}(\theta_i)}$$

$$= MSC(\mathbf{x}_0) + correction\ term$$

- correction term accounts for correlation between parameter estimates, nonlinearity
- uses identity for total derivative of $\left.\frac{\partial S(\theta_i, \Theta_{-i})}{\partial \Theta_{-i}}\right|_{\tilde{\Theta}_{-i}}$

  where S is the sum of squares function

Profile-Based Sensitivity Coefficient - Uniresponse

Computational issues

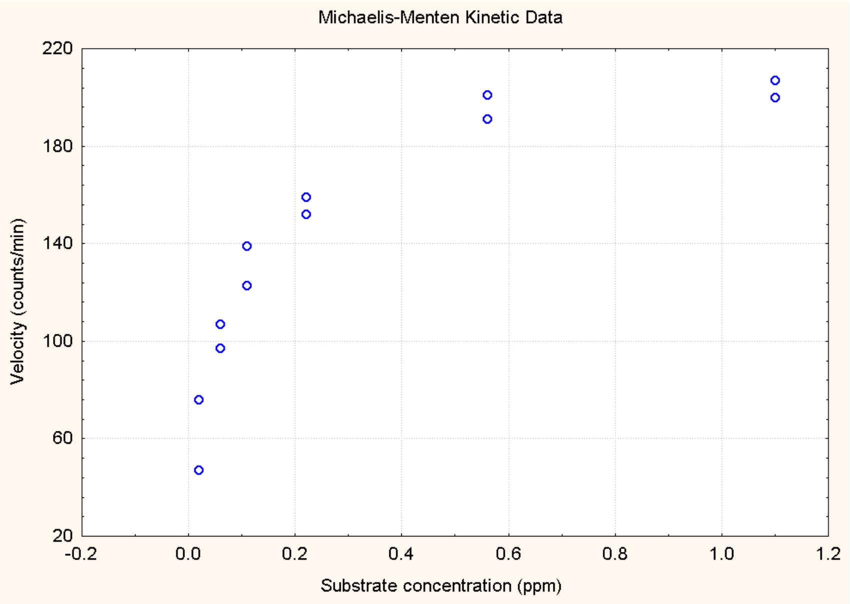
- scaling
  - » work with studentized parameters and predicted responses to remove scale dependence
    - centering by least squares estimate
    - scaling by standard error
  - » dimensionless sensitivity coefficients
- derivative values
  - » algebraic models - obtain directly
  - » differential equation models - via first- and second-order sensitivity equations
    - velocity and acceleration arrays are involved in PSC

Example - Uniresponse PSC

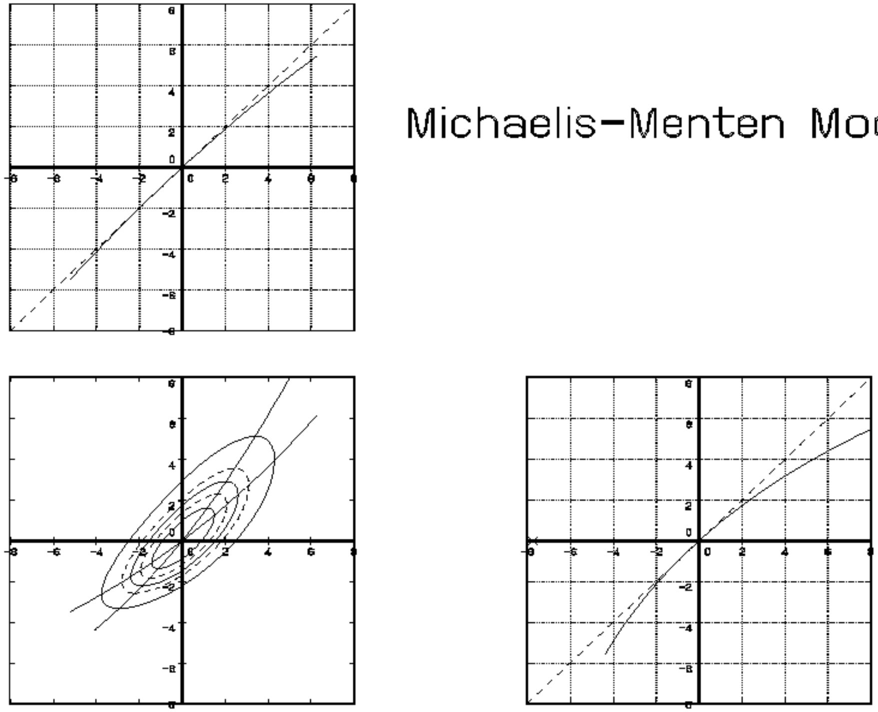
Michaelis-Menten model

» dataset from Bates and Watts (1988)

» model equation $$f(x,\theta) = \frac{\theta_1 x}{\theta_2 + x}$$



Profile traces

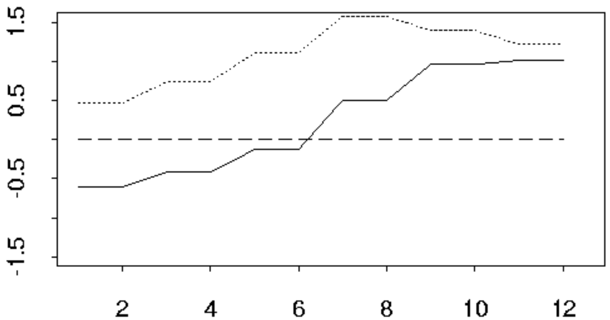


Michaelis-Menten Model

Example - Uniresponse PSC

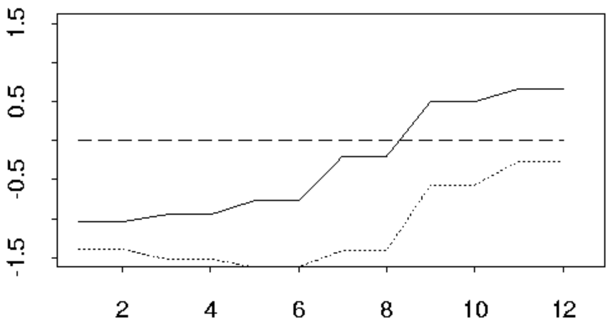
- Graphical summaries of PSC values

PSC_1, MSC_1



Solid line - PSC
Dashed line - MSC

PSC_2, MSC_2



Example - Unireponse PSC

Interpretation

- for θ_1
  - » MSC, PSC similar at high substrate
    - reduced correlation between parameter estimates
    - θ_1 defines asymptote at high concentration
  - » significant difference between MSC, PSC at mid-range
    - MSC indicates significant sensitivity, while PSC indicates negligible sensitivity
- for θ_2
  - » MSC, PSC similar at low substrate
    - reduced correlation between parameter estimates
    - θ_2 defines behaviour at low substrate concentrations
  - » PSC indicates more significant sensitivity at higher substrate - through parameter correlation

Multiresponse PSC

In this case, the parameter estimates are determined to minimize the Box-Draper determinant criterion:

$$d(\Theta) = \left|\mathbf{Z}'\mathbf{Z}\right|$$

The vector of predicted responses at a nominal run condition is denoted as:

$$\mathbf{H}_0(\Theta) = \left[f_1(\mathbf{x}_0, \Theta) \quad \ldots \quad f_M(\mathbf{x}_0, \Theta)\right]$$

The PSC is again defined as a total derivative, yielding a *vector* of PSC values in this instance:

$$\mathbf{PSC}_i(\mathbf{x}_0) = \frac{D\mathbf{H}'_0(\theta_i, \Theta_{-i}(\theta_i))}{D\theta_i} = \frac{\partial \mathbf{H}'_0}{\partial \theta_i} + \frac{\partial \mathbf{H}'_0}{\partial \Theta_{-i}} \frac{\partial \Theta_{-i}(\theta_i)}{\partial \theta_i}\Bigg|_{\tilde{\Theta}_{-i}}$$

$$= \mathbf{MSC}_i(\mathbf{x}_0) + \textit{correction vector}$$

Multiresponse PSC

Computational Issues

- scaling
  - » use similar studentization approach - centering by Box-Draper estimates and scaling by standard errors
- derivatives
  - » Hessian of determinant computed using identity due to Federov (1972)
  - » differential equation models - 1st and 2nd-order derivatives computed from sensitivity equations
    - efficiency issues
- conditioning
  - » dependencies can occur in response data, leading to singular residual matrix and spurious optima

Multiresponse PSC Example

Dow Chemical regression benchmark

(ref to be added)

- isothermal batch reactor
- dataset consists of concentration profiles over time for batches run at 3 different temperatures
    - » data for three species used - three response variables
- measurement times differ for each profile, and are sampled at non-uniform intervals
- considered by Biegler et al. (1986), Biegler et al. (1991), Guay and McLean (1995)

Example - Multiresponse PSC

Dow Chemical regression benchmark example - model

$$\frac{dy_1}{dt} = -k_2 y_1 y_2 A$$

$$\frac{dy_2}{dt} = -k_1 y_2 (x_2 + 2x_3 - x_4 - 2y_1 + y_2 - y_3) - k_2 y_1 y_2 A + k_{-1} \beta_1 (-x_3 + x_4 + y_1 - y_2) A$$

$$\frac{dy_3}{dt} = k_1 (x_3 - y_1 - y_3)(x_2 + 2x_3 - x_4 - 2y_1 + y_2 - y_3) + k_2 y_1 y_2 A - \frac{1}{2} k_{-1} \beta_2 y_3 A$$

where

$$k_i = k_{i0} \exp(\frac{E_i}{R} (\frac{1}{x_i} - \frac{1}{T_0}))\ , i = -1, 1, 2$$

$$A = \frac{-2x_3 + x_4 + 2y_1 - y_2 + y_3}{y_1 + \beta(-x_3 + x_4 + y_1 - y_2) + \beta_2 y_3}$$

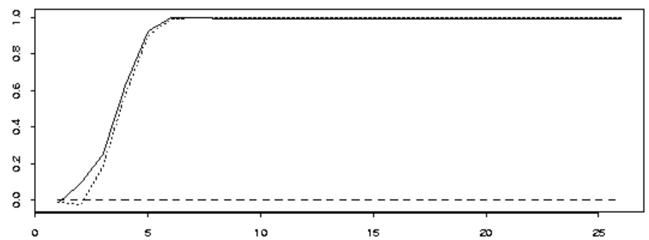
$$\beta_1 = \frac{K_1}{K_2}\ , \beta_2 = \frac{K_3}{K_2}$$

Example - Multiresponse PSC

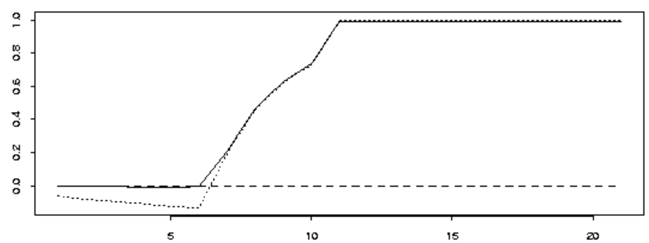
- responses - y_1, y_2, y_3
- parameters - k_{10}, E_1, k_{20}, E_2, k_{-10}, E_{-1}, β_1, β_2
- estimation
  - » using determinant criterion
  - » sensitivities obtained using Guay and McLean (1995) implementation of ODESSA routine of Kramer and Leis (1998)
  - » residual mean square has 73 degrees of freedom

Multiresponse PSC Example

T=40^{o}C

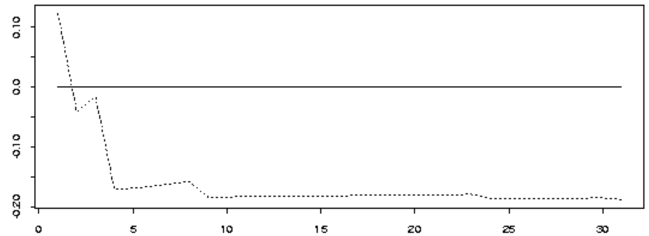


T=67^{o}C



PSC's describing sensitivity of predicted concentration of y_3 to β_1

T=100^{o}C



Observation Number

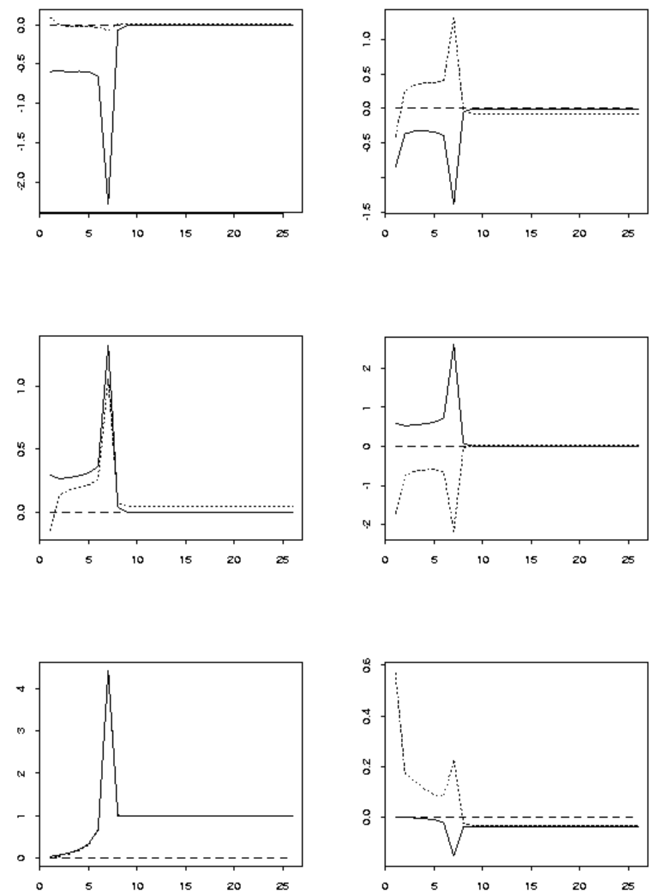
Multiresponse PSC Example

PSC's describing sensitivity of predicted concentration of y_2 to selected parameters at 40^{o}C

k_{10} k_{20}

k_{-10} E_{-1}

β_1 β_2



Observation Number

Example - Multiresponse PSC

- interpretation - y3 to beta1 at three temps
  - » close agreement of msc, psc at low, moderate T, but changes dramatically at high T - marginal sensitivities negligible at high T, but psc indicates significant sensitivities - consequence of parameter correlation and model behaviour over three temperatures
- y2 to params at 40 C
  - » general pattern - significant sensitivities at low times (early in batch) - where bulk of concentration changes occur - not surprising - msc’s indicate sensitivities where psc’s are negligible - suggestion that correlation counteracts marginal sensitivities … - discrepancies over early part of run - where major changes are occurring

Summary and Conclusions

- Proposed new measures of parametric sensitivity that account for correlation between parameter estimates, nonlinearity in estimation - profile-based sensitivity coefficients (PSC's)
- PSC's provide more accurate reflection of impact of parameter perturbation when all other parameters are adjusted to provide best fit - more accurate reflection of joint estimation
- PSC's defined as total derivative of predicted response with respect to parameter of interest
    - » for uni-response, via least squares criterion
    - » for multi-response, via determinant criterion
- quantitative measure that can be combined with graphical summaries - profile traces
- illustrated using uni-response algebraic model, multi-response differential equation model

Acknowledgements

We would like to acknowledge -

1) the assistance of Prof. Martin Guay through the use of his multi-response estimation routines;

2) the financial support of NSERC.