Parameter Certainty Quantification in Nonlinear Models

Amit Ashkenazi[a], Dana Solav[a]

[a] *Faculty of Mechanical Engineering, Technion Institute of Technology, Haifa, Israel*

Abstract

Estimating model parameters from experimental data is a common practice across various research fields. For nonlinear models the parameters are estimated using an optimization algorithm that minimizes an objective function. Assessing the certainty of these parameter estimates is crucial to address questions such as "what is the probability the estimation error is smaller than 5%?", "is our experiment sensitive enough to estimate all parameters?", and "how much can we change each parameter while still fitting the data accurately?". Typically, the certainty levels are quantified using a linear approximation of the model. However, we show that in models that are highly nonlinear in their parameters or in the presence of large experimental errors, this method fails to capture the certainty levels accurately. To address these limitations, we present an alternative method based on the Hessian approximation of the objective function. We show that this method captures the certainty levels more accurately and can be derived geometrically. We demonstrate the efficacy of our approach through a case study involving a nonlinear hyperelastic material constitutive model and an application on a nonlinear model for the conductivity of electrolyte solutions. Despite its higher computational cost, we recommend adopting the Hessian approximation when accurate certainty levels are required in highly nonlinear models.

Keywords: Confidence regions, Experimental errors, Hessian-based approximation, Indifference regions, Nonlinear models, Parameter estimation

1. Introduction

Estimating model parameters is an important step in model analysis, serving as a foundation for subsequent investigations. Commonly, the goal of parameter estimation is to find a unique set of parameters that fits a model to the data. Ideally, it ensures these parameters can predict new, untested scenarios. When done well, for given experimental data and a chosen model, certain parameter sets will make the model's behavior match the data, increasing their likelihood. The certainty in these parameter estimates is equally important. Quantifying parameter certainty provides insights into the reliability and robustness of the model predictions.

The task of estimating the parameters and the certainty in them becomes particularly challenging in highly nonlinear models (e.g., hyperelastic solid material models, adsorption isotherm models, and growth of microorganisms models). The nonlinearity and lack of analytical solutions complicate the parameter estimation process. In such cases, we often need to compare experimental data with numerical predictions iteratively to find the model parameters. This process presents additional challenges in the parameter estimation [1]. This scenario is explored further in section 3.

The least squares method is frequently employed to determine material parameters and is widely regarded as the gold standard in the literature. This method obtains parameters that are normally distributed around the true parameters with minimal standard deviations [2]. This approach is based on the assumption that the measurement errors are uncorrelated and normally distributed, which is usually a reasonable approximation for many quantitative experiments.

The process of estimating model parameters begins with a model of the problem

$$\begin{aligned} y &= f(x_1, ..., x_m; \theta_1, ...\theta_p) \\ &= f(\boldsymbol{x}; \boldsymbol{\theta}) \end{aligned} \tag{1}$$

relating a response f (such as force reaction, deformation, yield of product) with values of explanatory variables $\boldsymbol{x}$ (such as pressure, temperature, time) and unknown model parameters $\boldsymbol{\theta}$. We then proceed to evaluate $n > p$ different states of $\boldsymbol{x}$ to determine the p model parameters $\boldsymbol{\theta}$.

1.1. Problem statement

We consider the unweighted least squares estimation method, which is widely used for parameter estimation [3, 4, 5]. This method requires defining the model's deviation from experimental data. Mathematically, this is achieved by formulating an objective function $F(\boldsymbol{\theta})$ that we aim to minimize:

$$F(\boldsymbol{\theta}) = \sum_{i=1}^{n} [y_i - f(\boldsymbol{x}_i; \boldsymbol{\theta})]^2 \tag{2}$$

The objective function is defined as the sum over all n data points of the squared difference between the experimentally measured results y_i, and the predicted results $f(\boldsymbol{x}_i; \boldsymbol{\theta})$, where $\boldsymbol{x}_i$ is a vector of n different combinations of explanatory variables. If $\boldsymbol{\theta}$ is close to the unknown true model parameters $\boldsymbol{\theta}^*$, and the model $f(\boldsymbol{x}_i; \boldsymbol{\theta})$ is exact, the value of $F(\boldsymbol{\theta})$ will approach a minimum [6].

Using numerical optimization algorithms to minimize the objective function obtains the best-fit parameters $\hat{\boldsymbol{\theta}}$. Assuming normally distributed experimental errors, the parameters $\hat{\boldsymbol{\theta}}$ obtained are also the maximum likelihood estimates [7]. Optimization algorithms typically yield precisely repeatable parameters, and so for an objective function with a single unique minimum, the differences in estimated $\hat{\boldsymbol{\theta}}$ are small, regardless of the initial guess. The true uncertainty in the parameter bounds arises from measurement errors in y_i, which in turn create uncertainty in $F(\boldsymbol{\theta})$. In this sense, if a parameter set $\tilde{\boldsymbol{\theta}} \neq \hat{\boldsymbol{\theta}}$ results in $F(\tilde{\boldsymbol{\theta}})$ that is only slightly larger than $F(\hat{\boldsymbol{\theta}})$, considering the measurement errors, then both $\hat{\boldsymbol{\theta}}$ and $\tilde{\boldsymbol{\theta}}$ can be considered reasonable estimations of the parameters. Let $\Delta F > 0$ represent the largest difference above the minimum objective value we consider insignificant. Therefore, we can consider $\hat{\boldsymbol{\theta}}$ as likely as any other value of $\boldsymbol{\theta}$ for which

$$\Omega_{\Delta F} \triangleq \{\boldsymbol{\theta} \in \Theta : F(\boldsymbol{\theta}) \leq \Delta F\} \tag{3}$$

where Θ is the parameter space. This gives the confidence (also known as certainty or indifference) region in Θ. Regions that satisfy (3) are called exact confidence regions. We will now limit our attention to regions with no local minima other than $\hat{\boldsymbol{\theta}} \in \Theta$. When F is continuous and $\hat{\boldsymbol{\theta}}$ is its unique unconstrained minimum, the confidence region is a simply connected space surrounding $\hat{\boldsymbol{\theta}}$ in the p-dimensional θ space. This region is bounded by $F(\boldsymbol{\theta}) = F(\hat{\boldsymbol{\theta}}) + \Delta F$, which is a p-dimensional hypersurface. Choosing the value of ΔF is not trivial, but once chosen, the analysis remains consistent.

For example, ΔF can be calculated from the data distribution for a large enough number of data points n [7, 8, 9]. Through statistical estimation of the data distribution we obtain

$$\Delta F = \left(1 + \frac{p}{n-p} F^{1-\alpha}_{p,n-p}\right) F(\hat{\boldsymbol{\theta}}) \tag{4}$$

where we use a value from the F-distribution, denoted $F^{1-\alpha}_{p,n-p}$, with p and $n - p$ degrees of freedom. This value determines the certainty region, which is the range where we expect our estimates to fall with a certain level of confidence $100 \cdot (1 - \alpha)\%$ [8, p. 25]. However, from (4) we know that ΔF depends on the minimum value of our objective function. Consequently, we can only assess ΔF after determining $\hat{\boldsymbol{\theta}}$. Having established the problem and the need for accurate parameter estimation, we now focus on the approximations of the estimated parameters' confidence regions.

1.2. Approximations of confidence regions of the estimated parameters

Finding the exact confidence region boundaries is extremely computationally heavy, as it requires mapping out the entire objective function space. Consequently, we aim to use approximations for the objective function. In this section, we explore the classical approximation for the confidence region. Since most optimization algorithms use a local second-order approximation [10], it was deemed appropriate to use a second-order Taylor expansion of the

objective function about $\hat{\boldsymbol{\theta}}$ to approximate it [6, 8]. Considering the fact that the unconstrained minimum of F is at $\hat{\boldsymbol{\theta}}$, then $\frac{\partial F}{\partial \theta_j}\Big|_{\hat{\boldsymbol{\theta}}} = 0$, and

$$F(\theta_j) = F(\hat{\theta}_j) + \frac{1}{2}(\theta_k - \hat{\theta}_k)\frac{\partial^2 F}{\partial\theta_j\,\partial\theta_k}\Big|_{\hat{\boldsymbol{\theta}}}(\theta_j - \hat{\theta}_j) + ... \tag{5}$$

To simplify the approximation further, a classic linear approximation for the term of second derivatives, the Hessian $\boldsymbol{H} = \frac{\partial^2 F}{\partial\theta_j\,\partial\theta_k}$, was proposed [8, 9, 11], and implemented in solid mechanics (e.g. [12, 13]), ecology (e.g. [14, 15]), chemistry (e.g. [16, 17]) and more, suggesting that

$$\begin{aligned}\frac{\partial^2 F}{\partial\theta_j\,\partial\theta_k} &= 2\sum_{i=1}^{n}\left[\frac{\partial f_i}{\partial\theta_j}\frac{\partial f_i}{\partial\theta_k} - (y_i - f_i)\frac{\partial^2 f_i}{\partial\theta_j\,\partial\theta_k}\right] \\ &\approx 2\sum_{i=1}^{n}\frac{\partial f_i}{\partial\theta_j}\frac{\partial f_i}{\partial\theta_k} = 2\boldsymbol{J}(\boldsymbol{\theta})^T\boldsymbol{J}(\boldsymbol{\theta}) = 2\boldsymbol{V}(\boldsymbol{\theta})\end{aligned} \tag{6}$$

where $\boldsymbol{J}(\boldsymbol{\theta}) = \frac{\partial f_i}{\partial\theta_j}$ is the $n \times p$ Jacobian matrix of the model function. This approximation is based on at least one of two considerations: (a) if the model is linear in the parameters, then the Hessian of the model vanishes ($\frac{\partial^2 f_i}{\partial\theta_j\,\partial\theta_k} = 0$) and (6) is an equality; (b) we expect $(y_i - f_i)$ to be very small around $\hat{\boldsymbol{\theta}}$, thus the second term is negligible. It has been shown that (6) can be related to the covariance matrix [2, 4, 7, 9, 11], leading to the approximate confidence bounds of each parameter:

$$\theta_i = \hat{\theta}_i \pm \sqrt{\Delta F \boldsymbol{V}^{-1}_{ii}(\hat{\boldsymbol{\theta}})} \quad \text{(no summation on } i\text{)} \tag{7}$$

In cases where both assumptions do not hold, this approximation may lead to wrong confidence estimation. In this work, we consider models that are both nonlinear in their parameters, and do not capture the results extremely accurately.

To address this challenge, we propose using the actual Hessian matrix instead of the linear approximation. The Hessian coefficients can be obtained either theoretically, by differentiating the expressions for $F(\boldsymbol{\theta})$ twice with respect to the parameters, or numerically, by evaluating the sum of squares for a set of (at least) $N_c = (k+1)^2 \cdot p(p-1)/2 \sim O(p^2k^2)$ values, where k is the order of accuracy in the second order numeric derivation. We note that a numeric Hessian calculation is more computationally expensive than the numeric Jacobian approximation, which only requires $N_c = (k+1) \cdot p \sim O(pk)$ model evaluations.

2. Confidence bounds by Hessian approximation

To address models that are both nonlinear in the parameters and do not capture the results very accurately due to either poor model definition or poor experimental results, we propose the Hessian approximation method. In this section, we derive the certainty region bounds for each parameter using geometrical considerations. The classical method to estimate parameter certainty, as shown in (7), is derived from statistical and algebraic considerations [6, 8] and relies on a linear approximation, which is invalid in our analysis.

Let F($\hat{\boldsymbol{\theta}}$) be the minimum of the objective function's hyperplane. Since we assume a quadratic approximation of the objective function, then F($\hat{\boldsymbol{\theta}}$) is situated at the minimum of a $p+1$-dimensional elliptic paraboloid. Consequently, (3) defines the surface of a p-dimensional hyperellipsoid with its center at $\hat{\boldsymbol{\theta}}$. From a geometrical standpoint, the certainty in each parameter can then simply be defined as the projection of the hyperellipsoid on each of the parameter axes. e define the curvatures around $\hat{\boldsymbol{\theta}}$ by the second derivative of the objective function with respect to the parameters, $\frac{\partial^2 F}{\partial\theta_j\,\partial\theta_k} = \boldsymbol{H}$ [8, 14]. Since F($\hat{\boldsymbol{\theta}}$) is a minimum, $\boldsymbol{H}(\hat{\boldsymbol{\theta}})$ is definite positive. Consequently, $\boldsymbol{H}(\hat{\boldsymbol{\theta}})$ is invertible, and we can consider the eigenvalues of the matrix as the main curvatures, with the eigenvectors representing their direction [18].

Now to compute the size of the projections, we start by translating the center of the hyperellipsoid to the origin, for simplicity of analysis. The equation describing an axis-aligned canonical hyperellipsoid is given by

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + ... + \frac{x_n^2}{a_n^2} = 1 \tag{8}$$

and in matrix notation it can be represented as

$$\boldsymbol{x}^T \boldsymbol{D} \boldsymbol{x} = 1 \tag{9}$$

where

$$\boldsymbol{D} = \begin{bmatrix} \frac{1}{a_1^2} & 0 & \dots \\ \vdots & \ddots & \\ 0 & & \frac{1}{a_n^2} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{10a,b}$$

To represent the problem's general hyperellipsoid, we will apply a rotation operation on $\boldsymbol{D}$. In this case, we use the objective function's Hessian matrix eigendecomposition to inform the rotation. Let $\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{\kappa}\boldsymbol{Q}^T$, so that the general hyperellipsoid will be defined by $\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T\boldsymbol{x} = 1$, or simply

$$\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} = 1 \tag{11}$$

Here, the matrix $\boldsymbol{A}$[1] represents the hyperellipsoid in the parameters' coordinate system. The lengths of the hyperellipsoid's semi-axes are retrieved from the main curvatures of the objective function κ_i around the target point $\hat{\boldsymbol{\theta}}$, and are defined as the opening of the parabolic bowl at a certain height $s^2 > F(\hat{\boldsymbol{\theta}})$. Therefore, for a parabolic function $p(x_i) = \frac{\kappa_i}{2} x_i^2$, the semi-axes lengths are

$$a_i = \sqrt{\frac{2s^2}{\kappa_i}} \tag{12}$$

Note that $s^2 = \Delta F - F(\hat{\boldsymbol{\theta}})$. Since we aim to find the minimum and maximum values of the hyperellipsoid projection on each axis, we will utilize Lagrange multipliers. Mathematically, we look for the extrema of the function $h(\boldsymbol{x}) = e_i^T \boldsymbol{x}$, where $e_i^T = (0, ..., 0, 1, 0, ..., 0)$ is the parameter axis on which we project the hyperellipsoid. The constraint is the surface of the hyperellipsoid $g(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - 1 = 0$. To maximize $h(\boldsymbol{x})$ subject to the constraint $g(\boldsymbol{x})$ we find the zero of the derivative of the Lagrangian function

$$\begin{aligned} \nabla \left[e_i^T \boldsymbol{x} + \lambda(\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} - 1) \right] = 0 \\ e_i^T \cdot I + \lambda \cdot 2\boldsymbol{A}\boldsymbol{x} = e_i + 2\lambda \boldsymbol{A}\boldsymbol{x} = 0 \end{aligned} \tag{13}$$

Solving for $\boldsymbol{x}$ we get $\boldsymbol{x} = -\frac{1}{2\lambda}\boldsymbol{A}^{-1}e_i = \alpha \boldsymbol{A}^{-1}e_i$. By substituting this into $g(\boldsymbol{x})$ we obtain

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \alpha e_i^T \boldsymbol{A}^{-1}\boldsymbol{A}\alpha\boldsymbol{A}^{-1}e_i = \alpha^2 e_i^T\boldsymbol{A}^{-1}e_i = 1 \tag{14}$$

which implies

$$\alpha^2 = \frac{1}{e_i^T \boldsymbol{A}^{-1} e_i} \tag{15a}$$

$$\Rightarrow \boldsymbol{x} = \pm \frac{\boldsymbol{A}^{-1}e_i}{\sqrt{e_i^T \boldsymbol{A}^{-1} e_i}} \tag{15b}$$

Hence, the projection of the hyperellipsoid on each axis is

$$x_i = e_i^T \boldsymbol{x} = \pm \sqrt{e_i^T \boldsymbol{A}^{-1} e_i} = \pm \sqrt{\boldsymbol{A}_{ii}^{-1}} \quad \text{(no summation on } i\text{)} \tag{16}$$

From the Hessian eigendecomposition, it is evident that $\boldsymbol{\kappa} = \boldsymbol{Q}^T\boldsymbol{H}\boldsymbol{Q}$. So by substituting (12) into (9), we can now express $\boldsymbol{D}$ using the diagonal matrix of curvatures $\boldsymbol{\kappa}$:

$$\boldsymbol{D} = \frac{1}{2s^2}\boldsymbol{\kappa} = \frac{1}{2s^2}\mathbf{Q}^T\mathbf{H}\mathbf{Q} \tag{17}$$

from which we relate the generalized ellipsoid (11) to the Hessian. Therefore, $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T = \frac{1}{2s^2}\boldsymbol{H}$, so $\boldsymbol{A}^{-1} = 2s^2\boldsymbol{H}^{-1}$. And finally, we substitute this into (16). We obtain that the confidence bounds of each parameter i is estimated by

$$\theta_i = \hat{\theta}_i \pm s\sqrt{2\boldsymbol{H}_{ii}^{-1}(\hat{\boldsymbol{\theta}})} \quad \text{(no summation on } i\text{)} \tag{18}$$

[1] It is worth noting that the matrix $\boldsymbol{A}$ is both positive definite and symmetric. It is positive definite because its diagonal matrix is constructed of only positive values, and it is symmetric because $\boldsymbol{A}^T = (\boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T)^T = \boldsymbol{Q}(\boldsymbol{Q}\boldsymbol{D})^T = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T = \boldsymbol{A}$.

2.1. Alternative algebraic method

We present two alternative methods to derive the bounds of the confidence region. First, we follow the statistical approach using the linear approximation as a basis [14, 9], by simply substituting (6) into (7) we get

$$\boldsymbol{V}^{-1} \approx 2\boldsymbol{H}^{-1} \tag{19}$$

$$\theta_i = \hat{\theta}_i \pm \sqrt{2\Delta F \boldsymbol{H}^{-1}_{ii}(\hat{\boldsymbol{\theta}})} \quad \text{(no summation on } i\text{)} \tag{20}$$

Using the fact that $F(\hat{\boldsymbol{\theta}})$ might be substantial, unlike the approximation done in (7), we will use s^2 instead of ΔF to get (18). This method is more of a heuristic rather than a rigorous proof. Second, we follow the algebraic proposal derived by Bard (1974) [6, p. 171] in which he presents (3) with

$$|F(\boldsymbol{\theta}) - F(\hat{\boldsymbol{\theta}})| \leq s^2 \tag{21}$$

By substituting in (5) the confidence region is approximately defined by

$$\left| \frac{1}{2}(\theta_k - \hat{\theta}_k) \frac{\partial^2 F}{\partial\theta_j \, \partial\theta_k}\bigg|_{\hat{\boldsymbol{\theta}}} (\theta_j - \hat{\theta}_j) \right| \leq s^2 \tag{22}$$

Since $\boldsymbol{H}(\hat{\boldsymbol{\theta}})$ is at a minimum, it is definite positive and (22) becomes

$$\delta\theta_k \boldsymbol{H}_{jk}(\hat{\boldsymbol{\theta}}) \delta\theta_j \leq 2s^2 \tag{23}$$

We can now determine how much a parameter θ_i can be varied from $\hat{\theta}_i$ under this inequality. If we allow change only in one parameter, Equation (23) reduces to

$$\boldsymbol{H}_{ii}(\hat{\boldsymbol{\theta}}) \delta\theta_i^2 \leq 2s^2 \tag{24}$$

so by rearranging we get yet again (18). We note that this proof is valid only after isolating the change in one parameter, which limits its generality and restricts application to other scenarios. To illustrate the ability of this method to predict the certainty region, we now look at a specific example.

3. Numerical examples

In this section, we present a couple of numerical examples to illustrate the application and effectiveness of the different certainty approximation methods in various models. We aim to compare the Hessian-based and classical parameter uncertainty approximations across different scenarios, highlighting their strengths and limitations. We begin with a comparative analysis in hyperelastic material models, followed by an application to the conductivity of electrolyte solutions.

3.1. Comparing confidence approximations in hyperelastic models

We first investigate the relative merits of (18) and (7), as approximations of the confidence regions. Generally, we realize that:

1. The difference between these approximations and the exact confidence regions is typically small, as even non-linear models behave linearly sufficiently close to the minimum [9].
2. When the model accurately captures the experiments, no substantial difference is expected between these approximate confidence regions.

Therefore, we expect both estimations to be relatively close to each other, in many cases. However, there may be cases where the linear approximation does not accurately represent the exact region. To showcase such scenarios, we plot an example problem from the open dataset provided in Oddes and Solav (2023) [18]. In their paper, they studied the identifiability of the parameters of hyperelastic models estimated from indentation force-depth data. This problem

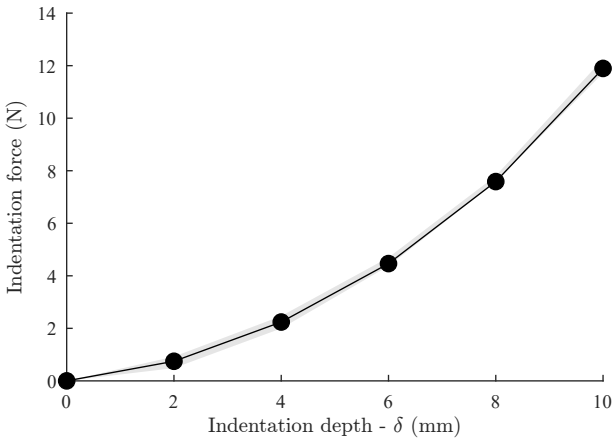


Figure 1: The indentation-force vs. indentation-depth curve results of Oddes and Solav (2023) [18] from a simulation using the Ogden material model with the true parameters $\boldsymbol{\theta}^*$. The shaded area represents the range of added random noise to the results.

Figure	Random error		$\hat{\boldsymbol{\theta}}$	$\boldsymbol{H}$ estimation (18)	$\boldsymbol{V}$ estimation (7)	$F(\hat{\boldsymbol{\theta}})$
2a	1.9%	θ_1 θ_2	1.04 0.89	27 ± 2.8 kPa 17 ± 5.3	27 ± 2.9 kPa 17 ± 5.3	0.0025
2b	7.0%	θ_1 θ_2	0.85 1.32	22 ± 3.7 kPa 25 ± 11	22 ± 4.1 kPa 25 ± 12	0.0099
2c	7.9%	θ_1 θ_2	0.85 1.26	22 ± 47 kPa † 24 ± 155 †	22 ± 8.7 kPa † 24 ± 25 †	0.0438
2d	9.4%	θ_1 θ_2	1.19 0.53	31 ± 7.4 kPa † 10 ± 8.5 †	31 ± 8.1 kPa † 10 ± 10.1 †	0.0226

Table 1: Numerical values of confidence ranges for varying noises on the simulated force data. The table lists the values shown in figures 2a, 2b, 2c, 2d. We compare the resulting estimations to the true values $\boldsymbol{\theta}^* = \{26\text{ kPa}, 19\}$. † Values are the unconstrained estimations, as the error exceeds the range of the parameter space.

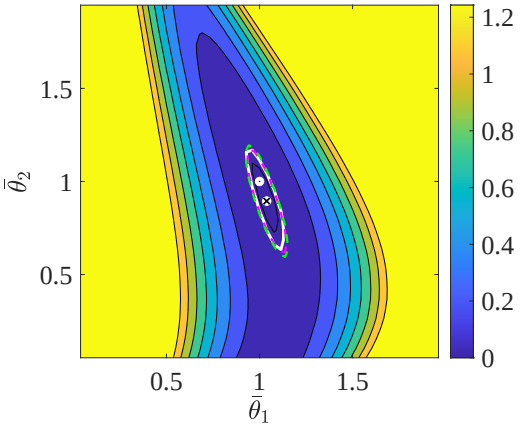
involves a highly nonlinear two-parameter Ogden hyperelastic material model. The constitutive relation between stresses and strains in the Ogden model is obtained using the strain-energy density function

$$\Psi^{OG} = \frac{\theta_1}{\theta_2^2}\left(\lambda_1^{\theta_2} + \lambda_2^{\theta_2} + \lambda_3^{\theta_2} - 3\right) \tag{25}$$

where λ_i is the stretch in the i-direction. For each combination of the two parameters Oddes and Solav (2023) [18] computed the force reactions from different indentation depths using a finite element simulation. The aim of the inverse analysis, in this case, is to estimate the Ogden material parameters from a given indentation depth-force response.

We use the center values of the parameter space, $\boldsymbol{\theta}^* = \{26\text{ kPa}, 19\}$, as the true values. To simulate different experimental results, we add a normally distributed random error of up to ±0.36 N to the force results obtained by the true material parameters, using MATLAB's `rand` function. This produces noisy y_i values to be used in the objective function (2) calculations. The baseline force response is shown in Figure 1 with the range of the noise responses in the shaded area. We then use the dataset of the simulated force reaction results to plot the entire objective function over the selected parameter space. There are in total 51×37 available simulation results in the dataset. Since we only have force values at discrete parameter sets, the Hessian and Jacobian matrices are calculated numerically using a central difference scheme with an accuracy of $k = 4$. Additionally, to find the global minimum we simply use MATLAB's `min` function In all simulations, the global minimum value $F(\hat{\boldsymbol{\theta}})$ is greater than zero, which allows us to construct exact and approximate confidence regions based on the desired confidence level (4).

To make a comparison between two parameters that don't share units, we present in Figure 2 a normalized parameter space $\{\bar{\theta}_1, \bar{\theta}_2\} = \theta_i/\theta_i^*$, with their normalized true values at $\bar{\boldsymbol{\theta}}^* = \{1, 1\}$. We use the Hessian at the minimum of



(a) Objective function values with small deviations of y_i from the true force values.

(b) Objective function values with large deviations of y_i from the true force values.

(c) Objective function of poor model fitting, with a large minimum value.

(d) Objective function values with very large deviations of y_i from the true force values, but good model fitting.

Figure 2: Contour plots depicting the objective function (2) using the force measurement values as reported for the Ogden material in Oddes and Solav (2023) [18] with an added random error. The white curve indicates the isoline representing a 95% certainty in the parameters. The magenta ellipse is the confidence region estimation using the Hessian $\boldsymbol{H}$ (18), and the green ellipse is the estimation using the Jacobian $\boldsymbol{V}$ (7). The $\otimes$ symbol marks the global minimum $\hat{\boldsymbol{\theta}}$, and the $\odot$ symbol marks the true parameters $\bar{\boldsymbol{\theta}}^* = \{1, 1\}$. The numerical values of the certainty ellipses are summarized in Table 1.

the parameter space, $\hat{\boldsymbol{\theta}} = \{\hat{\theta}_1, \hat{\theta}_2\}$, to determine the relative certainty in each parameter. To quantify the approximate certainty in each parameter we use (7) and (18) with the value of ΔF computed for 95% certainty. The parameter estimations and certainty bound results are summarized in Table 1.

Each sub-figure in Figure 2 presents a randomly generated objective function space, handpicked to highlight different comparisons between the two certainty region approximations. In Figure 2a we present a small relative error between the generated y_i values and the true force reaction with an average error of 1.9%. As expected, both models are very good at approximating the exact confidence region when the discrepancies are small. In Figure 2b we note that at the larger approximate error of around 7.0%, the Hessian-based approximation more accurately captures the exact confidence region compared to the classical approximation, which slightly overestimates the uncertainty. For very large relative errors, where the exact confidence region is no longer elliptical, we observe either an underestimation of the confidence region by the classical approximation, as shown in Figure 2c, or an overestimation, as shown in Figure 2d. The two cases mainly differ by the minimal value of the objective function. In both cases, the Hessian-based confidence region approximation capture the exact confidence more precisely. Collectively, our results highlight the

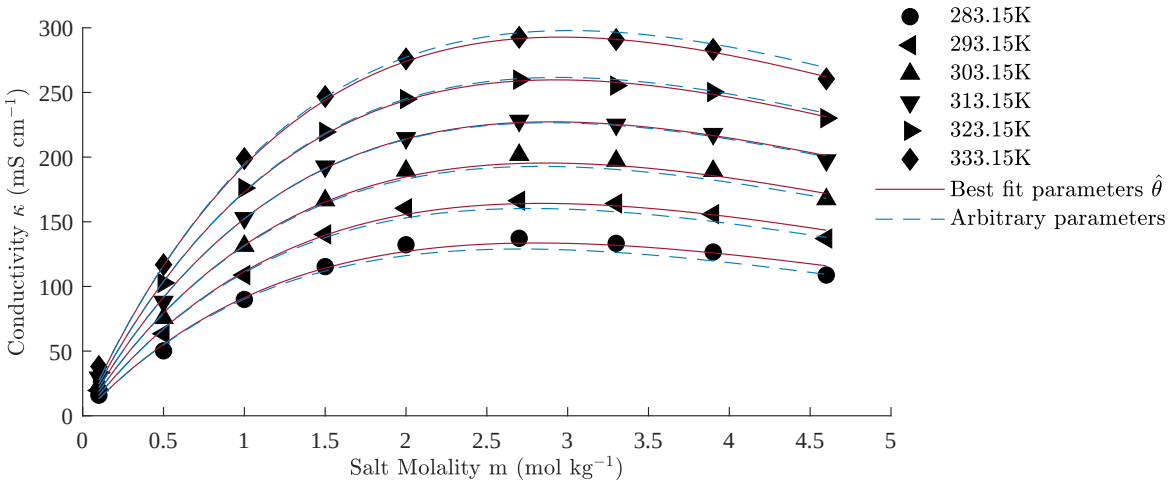


Figure 3: Experimental variation of conductivity κ with salt molality m at different temperatures T as reported by Zhang et al. (2020) [19]. Solid markers denote experimental data points, while the curves represent the model predictions from equation (26) with various parameter sets $\boldsymbol{\theta}$. The red curves represent the best-fit parameters $\hat{\boldsymbol{\theta}} = \{2.800, -663.3, 319.8, -635.4, 0.977\}$, and the blue dashed curves represent an arbitrary parameter set within the certainty bounds $\tilde{\boldsymbol{\theta}} = \{2.800, -663.3, 160, -160, 0.977\}$.

Method	θ_1	θ_2	θ_3	θ_4	θ_5	$F(\hat{\boldsymbol{\theta}})$
$\boldsymbol{H}$ estimation (18)	2.80 ± 0.06	−663 ± 18	-	-	0.977 ± 0.029	744
$\boldsymbol{V}$ estimation (7)	2.80 ± 0.16	−663 ± 50	319 ± 252	−635 ± 748	0.977 ± 0.034	

Table 2: Certainty bounds for the parameters of the model given the results of the $CaCl_2$-wPC-H_2O system with $w = 0.0425$. Dashed values indicate complete parameter insensitivity due to imaginary bound results.

strength of the Hessian based approximation in capturing the exact confidence regions.

3.2. Application to electrolyte solution conductivity

Having demonstrated the effectiveness of the Hessian method with a hyperelastic model, we now show an application. Zhang et al. (2020) [19] investigated the concentration and temperature dependence of the conductivity of electrolyte solutions. The relevant nonlinear model is

$$\kappa(m, T) = (\theta_1 T + \theta_2) m^{\theta_5} \exp\left(-\frac{\theta_3 m}{T - \theta_4}\right) \tag{26}$$

where κ is conductivity, m is electrolyte concentration, and T is temperature.

When quantifying the certainty, we symmetrically center the parameter space around the parameters that minimize the objective function, $\hat{\boldsymbol{\theta}}$, as identified by Zhang et al. (2020) [19]. We use the experimental results reported for the $CaCl_2$-wPC-H_2O system, with $w = 0.0425$ data. We use this data to calculate the residual value of the objective function at the minimum, $F(\hat{\boldsymbol{\theta}})$, to quantify the certainty of each reported parameter. The Hessian and Jacobian matrices are calculated numerically using a central difference scheme with an accuracy of $k = 4$. Initially, we evaluate the theoretical conductivity, and the objective function values over this 5-dimensional parameter space. To optimize computation time, the objective function is evaluated at $N_c = (4 + 1)^2 \cdot 5(5 - 1)/2 = 250$ discrete points used in the Hessian calculations. To quantify the approximate certainty in each parameter we use (7) and (18), with ΔF computed for 95% confidence. The experimental dataset comprised 54 results, each measured three times, resulting in $n = 54 \cdot 3$, with $p = 5$. We summarize the parameter estimates and their certainty bounds in Table 2.

The absence of values for the parameters θ_3 and θ_4 in the Hessian approximation of the certainty bounds is particularly noteworthy. This is a result of negative values in the square root of equation (18), indicating insensitivity of the

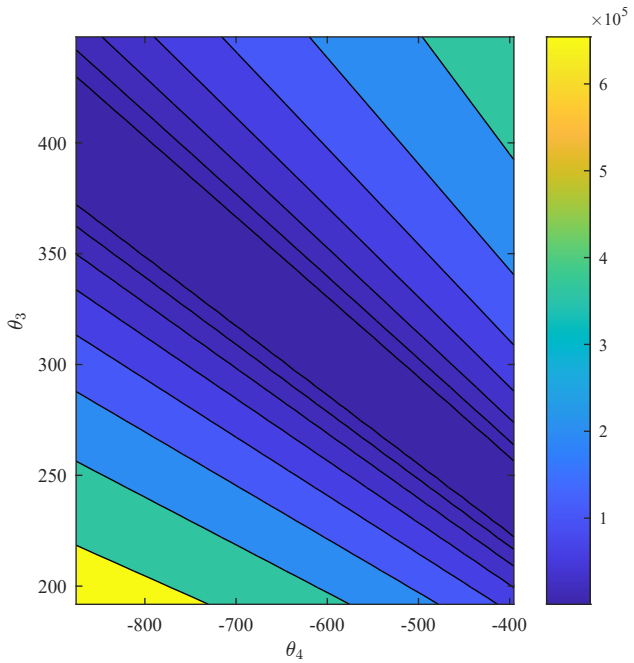


Figure 4: Two-dimensional slice of the objective function around the minimum shown in the $\{\theta_3, \theta_4\}$ plane.

objective function to these parameters. An analysis of the resulting eigenvectors and eigenvalues of the Hessian matrix confirms that the model exhibits insensitivity to a linear combination of these two parameters. Geometrically, this implies that the confidence hyperellipsoid is elongated along that direction. Figure 4 presents a two-dimensional slice of the objective function in the θ_3, θ_4 plane, revealing a "valley" structure that explains the unidentifiability of these parameters. This observation is consistent with the classical approximation, which assigns exceedingly large certainty bounds to θ_3 and θ_4, rendering their specific values effectively meaningless. Zhang et al. (2020) [19] reported large inconsistent variability in these two parameters across different systems, reinforcing our observation regarding their certainty. Figure 3 demonstrates this in practice, comparing the fit using the best-fit parameters reported by Zhang et al. (2020) [19] with a fit obtained by arbitrarily varying the insensitive parameters. Additionally, the model's nonlinearity introduces significant discrepancies between the estimations derived from the two methods. The Hessian-based certainties are tighter, suggesting that the classical approximation overestimates the true bounds.

4. Discussion

In this study, we investigate methods to quantify the certainty bounds of the estimated parameters in highly nonlinear models. For relatively small errors, we observe the classical method of approximating certainty regions aligns with the full Hessian method, even for highly nonlinear models. This alignment explains why the classical approach, which relies on a linear approximation of the Hessian, has been widely adopted: it is computationally efficient and relatively accurate. However, our results show that as experimental errors increase, this method falls short. In contrast, the full Hessian matrix approximation provides a more accurate estimation of parameter uncertainties, particularly in those cases where the model is highly nonlinear and the residuals are significant.

We show in Equation (6) that the classical method of approximating certainty regions relies strictly on the model, disregarding any information from the actual experimental results. This implies that only the sensitivity of the model to the parameters is determined. If the model fit is not perfect, a Hessian approximation must be used, as it also incorporates possible model errors into the analysis of the certainty region. Additionally, the example in section 3.1, which assumes the model is exact, demonstrates the advantages of the Hessian approximation in those cases as well. Consequently, our analysis further highlights the Hessian approximation's reliability in parameter bound estimation.

To quantify the bounds, we must address probability. The largest insignificant difference in the objective function value, ΔF, is crucial for giving meaning to the confidence region. The smaller it is, the greater the certainty. We note that ΔF decreases with the increase of the number of measurements n. Examining Equation (4), while the F-distribution value is relatively stable for large numbers of n (around 3-3.5 at 95% confidence), the ratio $(p/(n-p))$ significantly impacts the overall value. Therefore, sufficient data collection is important in increasing certainty in the parameters. Notably, the Hessian approximation, which incorporates more of this data, benefits more from repeated tests than the classical approach, though at a higher computational cost.

In some cases, the certainty bounds of two or more parameters are extensive. This does not imply that any combination of these parameters will yield statistically insignificant results from one another, but rather that there exists a broad range within which these parameters may be applicable. For instance, Figure 4 demonstrates that while not every combination in the parameter space is viable, selecting one parameter at random allows the second to be determined through a linear combination, thereby producing results comparable to the best-fit parameters. Although the presented method cannot identify these combinations, it highlights the need for further investigation. We recommend consulting the works of Gamage et al. (2011) [20] and Oddes and Solav (2023) [18], which explain the optimality criterion of the $\boldsymbol{H}$ or $\boldsymbol{V}$ matrices through their eigendecomposition.

The primary limitation of the Hessian estimation method is the need for the Hessian matrix to be invertible. Though $\hat{\boldsymbol{\theta}}$ should be at least in some neighborhood of the minimum [6], making the Hessian definite positive, this condition is not always met. Particularly when the model includes parameters with minimal influence on the results. In such cases, alternative strategies may be needed. These include excluding problematic data or using other experiments to obtain reliable parameter estimates [4].

One more difficulty in real situations is to choose the optimal increment for the finite differences approximating the second derivative. In most cases an analytical objective function is not available, and this issue must be addressed. However, given that this work focuses on an exemplary situation, a detailed discussion on this matter is unnecessary here. Nevertheless, the numerical accuracy of any approximation can be determined through various methods (e.g. [14, 21]). Addressing these difficulties will give better results from the Hessian method.

Lastly, the Hessian method works well for estimating parameter certainty when measurement errors are random, as it can handle these fluctuations and provide reliable estimates. However, it struggles with systematic errors, which consistently skew results due to biases in the measurement process or equipment. To address systematic errors, it's important to use thorough calibration and validation experiments.

5. Conclusion

In conclusion, we explore methods to determine the confidence in estimated parameters of highly nonlinear models, with statistical significance. We focus on the relative advantages of the Hessian approximation over the classical linear approximation method. Our findings demonstrate that while the classical method is effective for small discrepancies and low experimental errors, it fails as errors increase. The Hessian approximation, on the other hand, consistently provides more accurate and reliable confidence regions, particularly in complex, nonlinear models. Further, this approach not only enhances the reliability of the predictions but also provides deeper insights into the material properties. However, it is important to note that the Hessian approximation is more computationally intensive.

References

[1] A. Adessina, J.-F. Barthélémy, F. Lavergne, A. Ben Fraj, Effective elastic properties of materials with inclusions of complex structure, International Journal of Engineering Science 119 (2017) 1–15. doi:10.1016/j.ijengsci.2017.03.015.

[2] J. R. Wolberg, Data analysis using the method of least squares: extracting the most information from experiments, Springer, Berlin ; New York, 2006.

[3] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, Journal of the Society for Industrial and Applied Mathematics (1963) 431–441.

[4] S. Hartmann, R. R. Gilbert, Identifiability of material parameters in solid mechanics, Archive of Applied Mechanics 88 (2018). doi:10.1007/s00419-017-1259-4.

[5] S. S. Fayad, E. M. C. Jones, D. T. Seidl, P. L. Reu, J. Lambros, On the Importance of Direct-Levelling for Constitutive Material Model Calibration using Digital Image Correlation and Finite Element Model Updating, Experimental Mechanics 63 (2023) 467–484. doi:10.1007/s11340-022-00926-7.

[6] Y. Bard, Nonlinear Parameter Estimation, Academic Press, 1974.

[7] G. Box, G. Coutie, Application of digital computers in the exploration of functional relationships, Proceedings of the IEE - Part B: Radio and Electronic Engineering 103 (1956) 100–107. doi:10.1049/pi-b-1.1956.0020.
[8] G. A. F. Seber, C. J. Wild, Nonlinear regression, Wiley series in probability and statistics, Wiley-Interscience, Hoboken, NJ, 2003.
[9] D. De Pauw, Optimal experimental design for calibration of bioprocess models: a validated software toolbox, Ph.D. thesis, Universiteit Gent, 2005.
[10] S. S. Rao, Engineering optimization: theory and practice, 4th ed., John Wiley & Sons, Hoboken, N.J, 2009.
[11] K. Sidik, J. N. Jonkman, A comparison of the variance estimation methods for heteroscedastic nonlinear models, Statistics in Medicine (2016). doi:10.1002/sim.7024.
[12] N. Connesson, N. Briot, P. Y. Rohan, P. A. Barraud, S. A. Elahi, Y. Payan, Bilayer Stiffness Identification of Soft Tissues by Suction, Experimental Mechanics 63 (2023) 715–742. doi:10.1007/s11340-023-00946-x.
[13] J. Tang, W. Liu, Y. Mao, Y. Peng, Y. Zhang, S. Hou, Hierarchically goal-oriented prediction of skeletal muscle tissue constitutive behavior considering histological characteristics, International Journal of Engineering Science (2023) 103955. doi:10.1016/j.ijengsci.2023.103955.
[14] S. Marsili-Libelli, S. Guerrizio, N. Checchi, Confidence regions of estimated parameters for ecological systems, Ecological Modelling 165 (2003) 127–146. doi:10.1016/S0304-3800(03)00068-1.
[15] F. R. Pinto, C. F. C. Marcellos, C. Manske, A. Gomes Barreto Jr, Statistical analysis of parameters and adsorption isotherm models, Environmental Science and Pollution Research (2024). doi:10.1007/s11356-023-31820-x.
[16] S. Da Ros, K. Curran, Modelling and parameter estimation of diethyl phthalate partitioning behaviour on glass and aluminum surfaces, Chemosphere (2021) 131414. doi:10.1016/j.chemosphere.2021.131414.
[17] F. Martinez-Jimenez, M. P. de Arruda Ribeiro, C. R. Sargo, J. L. Ienczak, E. R. Morais, A. C. da Costa, Dynamic modeling application to evaluate the performance of spathaspora passalidarum in second-generation ethanol production: Parametric dynamics and the likelihood confidence region, Industrial & Engineering Chemistry Research (2021) 13822–13833. doi:10.1021/acs.iecr.1c02299.
[18] Z. Oddes, D. Solav, Identifiability of soft tissue constitutive parameters from in-vivo macro-indentation, Journal of the Mechanical Behavior of Biomedical Materials 140 (2023) 105708. doi:10.1016/j.jmbbm.2023.105708.
[19] W. Zhang, X. Chen, Y. Wang, L. Wu, Y. Hu, Experimental and modeling of conductivity for electrolyte solution systems, ACS Omega 5 (2020) 22465–22474. URL: https://doi.org/10.1021/acsomega.0c03013. doi:10.1021/acsomega.0c03013, publisher: American Chemical Society.
[20] T. P. Babarenda Gamage, V. Rajagopal, M. Ehrgott, M. P. Nash, P. M. F. Nielsen, Identification of mechanical properties of heterogeneous soft bodies using gravity loading, International Journal for Numerical Methods in Biomedical Engineering 27 (2011) 391–407. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.1429. doi:10.1002/cnm.1429, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cnm.1429.
[21] S. Avril, S. Evans (Eds.), Material Parameter Identification and Inverse Problems in Soft Tissue Biomechanics, volume 573 of *CISM International Centre for Mechanical Sciences*, Springer International Publishing, Cham, Switzerland, 2017. doi:10.1007/978-3-319-45071-1.