

# Confidence Interval for Nonlinear Regression

12/04/2024

We observe data  $(Y_i)_{i \in \llbracket 1, n \rrbracket}$ , which corresponds to the number of deaths per day.

**Assumption:**

The number of deaths per day follows a nonlinear model, such as an exponential regression:

$$Y_i = h_{\theta^*}(x_i) + \epsilon_i$$

with  $\theta^* \in \mathbb{R}^d$  the parameters of the regression and  $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ , i.i.d with  $\sigma$  unknown.

We want to find  $\theta$  and to compute confidence intervals on the predictions of the model.

Let

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} \quad \text{and} \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

We find  $\hat{\theta}$ , an estimator of  $\theta^*$  by minimizing the least square difference between  $h_{\theta}$  and the observed data:

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^d}{\operatorname{argmin}} \left( \sum_{i=1}^n (h_{\theta}(x_i) - Y_i)^2 \right)$$

As  $\hat{\theta}$  is close to  $\theta^*$ , we can linearize the objective function around  $\theta^*$ :

$$Y_i = h_{\theta^*}(x_i) + \epsilon_i$$

$$Y_i = h_{\hat{\theta}}(x_i) + (\theta^* - \hat{\theta})^T \nabla_{\theta} h_{\hat{\theta}}(x_i) + \epsilon_i$$

Let  $\tilde{Y}_i = Y_i - h_{\hat{\theta}}(x_i)$ ,  $\beta = \theta^* - \hat{\theta}$ , and  $A \in \mathbb{R}^{n \times d}$  such that  $A_{i,d} = \frac{\partial h_{\hat{\theta}}}{\partial \theta_d}(x_i)$ . We finally have:  $\tilde{Y} = A\beta + \epsilon$ , with  $\beta$  the parameter to estimate.

It is a linear regression problem that we can solve:

$$\hat{\beta} = (A^T A)^{-1} A^T \tilde{Y}$$

As  $\tilde{Y} \sim \mathcal{N}(0, \sigma^2 I_n)$ , we have:

$$\hat{\beta} \sim \mathcal{N}(0, \sigma^2 (A^T A)^{-1})$$

and so:

$$\sqrt{n}(\hat{\beta} - \beta^*) \longrightarrow \mathcal{N}(0, \sigma^2 (A^T A)^{-1})$$

As  $\beta^* = 0$ , we have:

$$\sqrt{n}(\hat{\theta} - \theta^*) \longrightarrow \mathcal{N}(0, \sigma^2 (A^T A)^{-1})$$

and so:

$$\hat{\theta} \sim \mathcal{N}\left(0, \frac{\sigma^2}{n} (A^T A)^{-1}\right)$$

Finally, as stated by the book "Bayesian Methods for Structural Dynamics and Civil Engineering", in the case of parameter estimation for Gaussian variables with the least squares method, we can approximate the covariance matrix (here  $\frac{\sigma^2}{n} (A^T A)^{-1}$ ) by the inverse of the Hessian of the objective function. We can also compute the unbiased estimator of  $\sigma^2$  by:

$$\hat{\sigma}^2 = \frac{1}{n-d} \sum_{i=1}^n (Y_i - h_{\hat{\theta}}(x_i))^2$$