SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Data Science and Big Data Analytics

Publisher's Home Page: https://www.svedbergopen.com/

Research Note Open Access

Generalized Least-Squares Ftting with Procedures for Uncorrelated Data of Constant Variance

Joachim W. Walewski[1*] and Thomas Metz[2]

[1]Division of Combustion Physics, Lund University, P.O. Box 118 221 00, Lund, Sweden. E-mail: joachim.walewski@gmail.com
[2]Trimble Jena GmbH, Carl-Zeiss-Promenade 10, 07745 Jena, Germany.

Article Info

Volume 1, Issue 3, November 2021
Received : 25 March 2021
Accepted : 21 September 2021
Published : 05 November 2021
doi: 10.51483/IJDSBDA.1.3.2021.1-5

Abstract

We show how generalized least-squares fitting, namely, the fitting of correlated data, can be carried out with algorithms for uncorrelated data of constant variance. Doing so requires only a simple linear transformation of the measurement data.

Keywords: *Constant variance, Correlated data, Curve fitting, Generalized least-squares fitting, Least squares, Least-squares fitting, Maximum-likelihood estimation, Ordinary least-squares fitting*

1. Introduction

Least-squares fitting of data is a recurring task in statistics and natural science. While there are many approaches and algorithms for performing least-squares fitting of uncorrelated data, only very few exist that support least-squares fitting correlated measurement data. This is not a major problem though, since—as we show in this application note—least-squares fitting of correlated measurement data can be translated into a least-squares fitting problem for uncorrelated data of constant variance. The only thing needed is a simple linear transformation of the data.

2. The Optimization Problem

Least-squares fitting can be described as follows. The relationship between the varaibles x_1, x_2, ... x_m is described by a model function.

$$x = f(x,\alpha) \quad ...(1)$$

where x is a column vector of the x-values and α is the vector of the model parameters. An example for f is

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 \end{pmatrix} \quad ...(2)$$

where f has two input variables x_1 and x_2, and one output variable x_3. This example is illustrated in Figure 1.

* Corresponding author: Joachim W. Walewski, Division of Combustion Physics, Lund University, P.O. Box 118221 00, Lund, Sweden.. E-mail: joachim.walewski@gmail.com

In this application note, we assume that the x-values are complex. However, our findings also hold for real-valued x's.

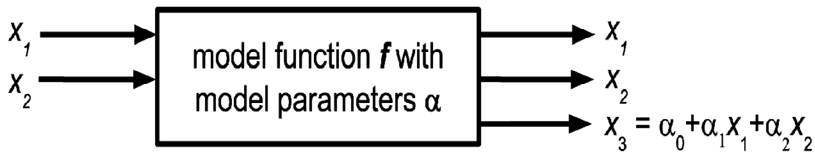

Figure 1: Illustration of the Model Function in Equation (2)

The problem to be solved is that the true x- and α- values are unknown. For instance, the x values can only be observed with additive noise. Generalized least-squares curve fitting provides a way of estimating the true values. The estimates are written as $\hat{x}$ and $\hat{\alpha}$ and they are inferred by minimizing Equation (3).

$$\chi^2 = \delta^\dagger C^{-1}\delta \quad \text{...(3)}$$

where

$$\delta := x - f(\hat{x},\hat{\alpha}) \quad \text{...(4)}$$

The elements of the matrix C are defined as:

$$C_{ij} := E(\{x_i - \varepsilon(x_i)\}\{xj - E(x_j)\}^*) \quad \text{... (5)}$$

for $1 \le i, j \le m$ [see, for instance (Taupin, 1988)]. ε stands for the expectation value. The minimization of Equation (3) is referred to as generalized least-squares fitting.

Note that the x-values in Equation (5) are modeled as stochastic variables. For more details on observed values and stochastic variables see Appendix A.

Also note that C_{ii} is the variance of x_i and that C_{ij} for $i \neq j$ indicates the correlation between x_i and x_j. Furthermore, note that if f is non-linear in x, f may have to be replaced with a slightly modified function. For more details on this see (Metz *et al.*, 2003).

While many analytical formulas and numerical procedures exist for the minimization of χ^2 for uncorrelated x-values ($C_{ij} = 0$ for $i \neq j$), few exist for correlated x-values. This is especially true for commercially available procedures, which only seem to tackle the minimization of χ^2 for uncorrelated x values with constant variance. In other words, they only minimize

$$\chi^2 = \delta^\dagger\delta \quad \text{...(6)}$$

The minimization of Equation (6) is referred to as ordinary least-squares fitting.

The next section discusses how generalised least-squares fitting problems can be transformed into ordinary least-squares fitting problems.

3. Translating Generalized Least-Squares Fitting into Ordinary Least-Squares Fitting

This transformation exploits the fact that C is hermitian [namely $C^\dagger = C$ according to Equation (5)]. See also (Wikipedia Contributors, 2021). Therefore, C can be factored into:

$$C = U\,V\,U^\dagger \quad \text{...(7)}$$

where U and V are constructed as follows (Wikipedia Contributors, 2020): solve the eignenvalue equation

$$Cu = vu \quad \text{...(8)}$$

The solutions are the vectors $u_1, u_2,...u_m$ and the real values $v_1, v_2 ... v_m$. Thereafter, construct U according to

$$U = (u_1, u_2,...u_m) \qquad ...(9)$$

and V according to

$$V_{ij} := \begin{cases} v_i & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \qquad ...(10)$$

Note that U is a unitary matrix, meaning $U^{-1} = U^{\dagger}$ (Wikipedia Contributors, 2020).

$$\delta = U \sigma \delta' \qquad ...(11)$$

with

$$\sigma_{ij} := \begin{cases} \sqrt{v_i} & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \qquad ...(12)$$

transforms Equation (3) into Equation (6). To demonstrate this, we insert Equation (11) into Equation (3):

$$\begin{aligned} \chi^2 &= \delta^{\dagger} C^{-1} \delta \\ &= (U\sigma\,\delta')^{\dagger} C^{-1} (\mathrm{U}\sigma\,\delta') \\ &= (\delta')^{\dagger} \sigma^{\dagger} U^{\dagger} C^{-1} U \sigma \delta' \end{aligned} \qquad ... (13)$$

Using Equation (7), this yields'

$$\begin{aligned} \chi^2 &= (\delta')^{\dagger} \sigma U^{\dagger} (UVU^{\dagger})^{-1} U \sigma \delta' \\ &= (\delta')^{\dagger} \sigma \underbrace{U^{\dagger}\left(U^{\dagger}\right)^{-1}}_{1} V^{-1} \underbrace{U^{-1}U}_{1} \sigma \delta' \\ &= (\delta')^{\dagger} \sigma V^{-1} \sigma\,\delta' \end{aligned} \qquad ...(14)$$

where 1 is the identity matrix.

Remembering that $V = \sigma\sigma$ [see Equation (12)], we get

$$\begin{aligned} \chi^2 &= (\delta')^{\dagger} \underbrace{\sigma\sigma^{-1}}_{1} \underbrace{\sigma^{-1}\sigma}_{1} \delta' \\ &= (\delta')^{\dagger} \delta' \end{aligned} \qquad ...(15)$$

which indeed is isomorphic to Equation (6) (Wikipedia Contributors, 2021).

4. How to Use this Result in Practice?

From Equation (11) and Equation (4), we get

$$\begin{aligned} \delta' &= \sigma^{-1} U^{-1} \delta \\ &= \sigma^{-1} U^{-1} x - \sigma^{-1} U^{-1} f(\hat{x}, \hat{\alpha}) \\ &= \underbrace{\sigma^{-1} U^{-1} x}_{x'} - \underbrace{\sigma^{-1} U^{-1} f(\hat{x}, \hat{\alpha})}_{f'(\hat{x}', \hat{\alpha}')} \end{aligned} \qquad ...(16)$$

The last row exploits the fact that $\delta' = x' - f'(\hat{x}', \hat{\alpha})$.

From Equation (16), the following recipe for fitting correlated data with algorithms for ordinary least-squares fitting can be inferred.

1. Take your measurement data x and calculate

$$x' = \sigma^{-1} U^{\dagger} x \qquad ... (17)$$

2. Calculate a new model function

$$f'(x',\alpha) = \sigma^{-1} U^{\dagger} f(U\sigma\,\hat{x}',\alpha) \quad \text{...(18)}$$

3. Minimize

$$(\delta')^{\dagger}\delta' \quad \text{...(19)}$$

with

$$\delta' = x' - f'(x',\alpha) \quad \text{...(20)}$$

This step yields the estimates $\hat{x}'$ and $\hat{\alpha}$

4. Calculate the estimates $\hat{x}$ with

$$\hat{x} = U\sigma\,\hat{x}' \quad \text{...(21)}$$

5. Conclusion

This application note provides a simple recipe for how to translate generalized least-squares curve fitting into least-squares curve fitting of uncorrelated data with constant variance.

Acknowledgment

The authors thank the Center for Combustion Science and Technology, CE-COST, for financial support.

References

Metz, T., Walewski, J., and Kaminski, C.F. (2003). Maximum-likelihood Curve Fitting Scheme for Experiments with Pulsed Lasers Subject to Intensity Fluctuations. *Applied Optics*, 42, 1551-1563.

Taupin, D. (1988). Probabilities, Data Reduction and Error Analysis in the Physical Sciences. *Les editions de physique*, Les Ulis.

Wikipedia Contributors. (2020). Eigendecomposition of a Matrix—Wikipedia, the Free Encyclopedia [Online at https://en.wikipedia.org/w/index.php?title=Eigendecomposition_of_a_matrix&oldid=1012755753; accessed 2021-03-20].

Wikipedia Contributors. (2021a). Hermitian Matrix—Wikipedia, the Free Encyclopedia [Online at https://en.wikipedia.org/w/index.php?title=Hermitian_matrix&oldid=999254083; accessed 11-January-2021].

Wikipedia Contributors. (2021b). Isomorphism—Wikipedia, the Free Encyclopedia [Online at https://en.wikipedia.org/w/index.php?title=Isomorphism&oldid=1007734513; accessed 20-March-2021].

Appendix

Stochastic Variables and Observations

The reader should be aware of the distinction between stochastic variables and observations. Stochastic variables are a mathematical concept for representing quantities that are subject to random variation. Stochastic variables are associated with a probability law, a so-called distribution function, and one can assign expectation values, variances, co-variances and so on to them.

Observations, on the other hand, are concrete numbers that are observed in, for instance, measurements, and they are typed into, say, a computer when performing some kind of evaluation.

Many textbooks on statistical inference use capital letters such as X for stochastic variables and small letters such as x for observations. In this work, we do not make this formal distinction in order to keep the formalism simple.

From the context, it will be clear what is meant. For instance, expectation values, variances etc. are only defined for stochastic variables, whereas a minimization of Equation (6) can only be done for concrete numbers, namely, the observations x.

Statistical inference is about estimating the values of unknown quantities. These estimates are obtained from the data of an experiment, so they are numbers you calculate from the observations.

Cite this article as: Joachim W. Walewski and Thomas Metz (2021). Generalized Least-Squares Ftting with Procedures for Uncorrelated Data of Constant Variance. Developing an Algorithmic Culture. *International Journal of Data Science and Big Data Analytics,* 1(3), 1-5. doi: 10.51483 / IJDSBDA.1.3.2021.1-5.