Measurement 116 (2018) 140–145

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

An "ad-hoc" modified likelihood function applied to optimization of data analysis in atomic spectroscopy

Leonardo Bennun

Applied Physics Laboratory, Physics Department, University of Concepcion, Casilla 160-C, Concepcion, Chile

ARTICLE INFO

Keywords:
"Ad-hoc" likelihood function
Optimization of statistical results
Improving accuracy and uncertainty
Metrological traceability
Least squares methods
Spectral analysis
Evaluation of repeatability

ABSTRACT

In this paper we propose an "ad-hoc" construction of the Likelihood Function, in order to develop a data analysis procedure, to be applied in atomic and nuclear spectral analysis.

The classical Likelihood Function was modified taking into account the underlying statistics of the phenomena studied, by the inspection of the residues of the fitting, which should behave with specific statistical properties. This new formulation was analytically developed, but the sought parameter should be evaluated numerically, since it cannot be obtained as a function of each one of the independent variables. For this simple numerical evaluation, along with the acquired data, we also should process many sets of external data, with specific properties — This new random data should be uncorrelated with the acquired signal.

The statistical method developed was evaluated over computer simulated spectra and over an experimental example. The numerical evaluations of the calculated parameter applying this method, indicate an improvement in one order of magnitude over accuracy compared with those produced by least squares approaches. In the experimental application of this method, over the quantification of Mn in a spectrum of a IAEA's sample (Proficiency Ring Test, PTXRFIAEA12), we found an improvement over the precision of the results.

We still have to evaluate the improvement produced by this method over Detection and Quantitation Limits, in TXRF spectral analysis.

1. Introduction

In atomic and nuclear spectroscopies, almost all of the methods for spectral analysis are based on the least squares algorithms. These data processing and interpretation techniques are usually applied indistinctly to linear or non-linear physical systems. The great success of the application of this method is based essentially on a deep agreement between the underlying physical phenomena studied and the applied mathematical theory.

In most of the atomic and nuclear spectroscopic techniques, in order to obtain the sample/system characterization, we found a common sequence of events, which can be detailed as follows: (*a*) a generation of energetic particles required as an excitation source, (*b*) the de-excitation process of the sample, (*c*) the detection (usually trough a solid state detector) and (*d*) an acquisition process (typically the electronic chains are composed by a preamplifier, amplifier, and a multichannel analyzer). The intrinsic fluctuations of the excitation and de-excitation of the sample, being discrete events, are ruled by the Binomial Distribution. But, in the atomic or nuclear interactions, the Poisson distribution may be used as an approximation to the Binomial Distribution (population n large and probability of a given event, p small). Moreover, in the classical texts of Probability, the radioactive decay and nuclear decay reactions are used as iconic examples of the Poisson's Statistics. This probability also rules the characteristic backgrounds proper of the matrix of the sample. At this stage, we should mention that the Poisson's Distribution, when it is applied to a large number of events (usually, $n > 30$) is very well described by a Normal Distribution. So, the joint probability that describes all kind or sequence of atomic or nuclear events in the sample, is a Normal Distribution. The Monte Carlo methods apply this property in order to infer the average behavior of the particles in the physical system from the average behavior of simulated particles [1].

At each one of the rest of the steps that follow the electrical signal leaving the detector until it is processed in the multichannel analyzer, it is affected by characteristics fluctuations, like temperature and gain voltage variations, electrical environmental noise, small systematics errors, etc. Again, by applying the Central Limit theorem, we can assure that the Distribution that rule the acquired atomic or nuclear signal, at each channel in the multichannel analyzer, would be a Normal Distribution.

The Central Limit theorem is implicitly applied in the data processing of a large number of complex systems, which are studied with the

E-mail address: lbennun@udec.cl.

https://doi.org/10.1016/j.measurement.2017.10.061
Received 8 May 2017; Received in revised form 6 September 2017; Accepted 26 October 2017
Available online 27 October 2017
0263-2241/

least squares approach and many of their multiple variations [2]. These methodologies are applied with many strategies in computational mathematics, in order to optimize the results in scientific and engineering applications. Some of the areas of applications are: image and video processing, medical treatments, etc. [3].

In the specific case of spectral analysis, the application of the Maximum Likelihood formulation is almost perfectly suited, since it is constructed considering a Normal Distribution at each channel. But, a question remains: How good is the quality of the results, for a given parameter, obtained from a Maximum Likelihood estimation? What the Likelihood Function is computing is how likely the measured data is to have come from the distribution assuming a particular value for the hidden parameter; the more likely this is, the closer one would think that this particular choice for hidden parameter is to the true value. So, in atomic spectroscopies, the results obtained from least squares algorithms should be the best, and no method of improvement of the results could be proposed. Moreover, the properties of the obtained results were largely studied. The evaluated parameters from the least squares formulation are unbiased (in the limit of infinite measurements) and have minimum variance among all unbiased linear estimators. This means that the estimates "get us as close to the true unknown parameter values as we can get". For these reasons, an improvement on the quality of the results over those obtained by the least squares method seems to be unrealistic. Furthermore, these results are considered as the limit of highest quality, to which tend the results, for instance, obtained from the Neural Network approach [4,5] when they are applied to data analysis in atomic or nuclear spectroscopies, and related systems.

However, in another work [6] we devised a new smoothing method which was applied to simulated spectroscopic data, producing results with better accuracy than those obtained from the least squares approaches.

In this paper, we propose a modification in the construction of the Likelihood function, which leads to a remarkable improvement on the quality of the results provided by the least squares approaches. This modification was made taking into account the underlying statistics of the phenomena studied, by the inspection of the residues of the fitting, which should behave with specific statistical properties. This new formulation was analytically developed, but the calculated parameter should be evaluated numerically, since it cannot be obtained as a function of each one of the independent variables. For the required numerical evaluation, along with the acquired spectrum, we should process many sets of external data with specific properties. This arbitrary term is a random sequence of data which is uncorrelated with the acquired signal. It should be ruled by a Gaussian distribution, having mean value zero and standard deviation $\Delta = 1$.

This statistical method was evaluated using computer simulated spectra. The numerical estimations of the calculated parameter applying this method, indicate an improvement over accuracy, one order of magnitude better than those produced by the least squares approaches. The precision of the results is also enhanced but with smaller impact – it is reduced in a factor of 3 in our examples.

We also describe an experimental application of the developed method, over the quantification of Mn in a spectrum of a IAEA's sample (Proficiency Ring Test, PTXRFIAEA12), where we found an improvement over the precision of the results.

We still have to evaluate the improvement produced by this method over Detection and Quantitation Limits, in TXRF spectral analysis.

2. Theoretical

A Maximum Likelihood estimate for some hidden parameter γ (or parameters, plural) of some probability distribution is a number $\hat{\gamma}$ computed from an Independent and Identically Distributed sample (IID) $M_1, \ldots, M_n$ from the given distribution that maximizes something called the "Likelihood Function". Let's suppose that this distribution is governed by a probability density function (pdf) $G(X; \gamma_1, \ldots, \gamma_k)$, where the γ_i's are all hidden parameters. The Likelihood Function associated to this sample is:

$$L(M_1,\ldots,M_n) = \prod_{i=1}^{n} G(M_i,\gamma_1,\ldots,\gamma_k) \tag{1}$$

Note that in all cases the estimated values are represented by English letters while parameter values are represented by Greek letters.

If the distribution is $N(\mu,\sigma^2)$, the Likelihood Function is:

$$L(M_1,\ldots,M_n;\hat{\mu}_i,\hat{\sigma}_i^2) = \frac{1}{(2\pi)^{n/2}}\left(\frac{e^{-\frac{1}{2\hat{\sigma}_1^2}(M_1-\hat{\mu}_1)^2}}{\hat{\sigma}_1}\times\ldots\times\frac{e^{-\frac{1}{2\hat{\sigma}_n^2}(M_n-\hat{\mu}_n)^2}}{\hat{\sigma}_n}\right) \tag{2}$$

where the symbol "^" over the variables μ_i and σ_i^2 indicates that they are estimators of their real values.

In atomic spectroscopies (like TXRF, μSR-XRF, PIXE, etc.) the $\hat{\mu}_i$ values could be linearly related with an specific known function, that is, $\mu_i = \alpha F_i$. The function F can be understood as a particular perfectly defined signal, directly linked to the presence of a given element in the sample [7,8]. The Fi sequence take specific values at each channel i; and the α value is linearly related mainly with the abundance of this element, and others fundamental parameters [9]. In this case, the Maximum Likelihood yields [10]:

$$lnL = -\sum_{i=1}^{n}\frac{(m_i-\alpha F_i)^2}{2\sigma_i^2} + \sum_{i=1}^{n} ln\left(\frac{1}{\sqrt{2\pi}\,\sigma_i}\right) \tag{3}$$

At Eq. (3) the m_i values represent the discrete data acquired with the multichannel analyzer; and the σ_i values are their standard deviation, at each channel i. In this case, the total number of channels in the spectrum is n.

The most probable $\hat{\alpha}$ value with its confidence interval can be calculated from Eq. (3) as [10]:

$$\hat{\alpha} = \frac{\sum_{i=1}^{n}\frac{m_i F_i}{\sigma_i^2}}{\sum_{i=1}^{n}\frac{F_i^2}{\sigma_i^2}} \pm \frac{1}{\sqrt{\sum_{i=1}^{n}\frac{F_i^2}{\sigma_i^2}}} \tag{4}$$

The same result is obtained applying least squares minimization to this heteroscedastic system.

As it was discussed in the Introduction, in atomic and nuclear spectroscopies, it is advisable to mention that the results at Eq. (4) are obtained from an adequate representation of the studied physical system. Each one of the underlying processes are indeed well described by the probabilities applied (Normal Distributions) in the developed model. In this sense, the values obtained at Eq. (4) should be the best, and no method of improvement of the results could be proposed. If we ask: Could the quality (accuracy and uncertainty) of a measurement be improved beyond that the Maximum Likelihood Criterion establishes in atomic spectroscopy? Being these spectroscopies well described by the proposed statistical model, the answer should be negative.

But, let's inspect the residues of the adjustment, which also should behave with specific statistical properties. Once the α parameter is evaluated, the difference $\varepsilon_i = m_i - \alpha F_i$, should on the average: (*i*) have equal number/quantity of positive and negative values, (*ii*) since it is ruled by a Poisson statistics, its uncertainty at each channel i should be: $\Delta\varepsilon_i = \sqrt{\alpha F_i}$. Taking into account these elements we can propose an improved version of the likelihood function[1] [11,12,13,14] which includes the statistical properties of the studied system, as:

$$lnL = -\sum_{i=1}^{n}\frac{(m_i-\alpha F_i - bg_i\sqrt{\alpha F_i})^2}{\sigma_i^2} + \sum_{i=1}^{n} ln\left(\frac{1}{\sqrt{2\pi}\,\sigma_i}\right) \tag{5}$$

At Eq. (5) the term bg_i is an arbitrary random sequence of n datum, uncorrelated with the acquired signal. This data is ruled by a Gaussian

[1] There are many denominations about the "ad hoc" modifications over the likelihood function. It is not clear in which category this proposition should be included.