Mathematical Transformation of Multidimensional Correlated Data into Uncorrelated Raman Spectra to Increase the Sensitivity of Identification with Silver Nanoparticles

Viktor M. Emelyanov 1*, Tatiana A. Dobrovolskaya 1, Viktor V. Yemelyanov 1

1 The Southwest State University; Faculty of Mechanics and Technology, Kursk, Russia vmemelianov@yandex.ru (V.M.E.); dobatatiana74@mail.ru(T.A.D.); wemelyanov@yandex.ru (V.V.Y.);
* Correspondence: vmemelianov@yandex.ru (V.M.E);
Scopus Author ID 56344103500
Received: 29.12.2021; Accepted: 24.01.2022; Published: 25.03.2022

Abstract: The article discusses approaches to translating correlated statistical data into uncorrelated form. The results of the transformation of mathematical models with silver nanoparticles and without nanoparticles into uncorrelated form with simultaneous solution of a system of equations with uncorrelated matrices are presented. Solutions of a system of multi-dimensional equations for determining the probability densities \( p_0 \) and \( p_1 \) are obtained. These mathematical models are based on the Raman polarization spectra of polyester fibers in recognizing silver nanoparticles, taking into account the polarization of laser radiation in two directions: X-across and Y-along the fibers. A method for increasing the resolution of the identification of silver nanoparticles on polyester fibers is proposed. When solving the system using nonlinear quadratic and XY differential equations of probability densities of distribution ellipses, the resolution of identification of silver nanoparticles \( p_0 \) and \( p_1 \) in the range \( 10^{-2} \) - \( 10^{-547} \) was obtained.

Keywords: colloidal silver nanoparticles; Raman spectra; multi-dimensional correlation equations; the intersection of distribution ellipses; transformation of correlation data; recognition reliability; normal two-dimensional distributions; resolution of nanoparticle identification.

© 2022 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

To ensure nano, pico - and molecular biotechnology requires accuracy and resolution in the range of \( 10^{-9} \) - \( 10^{-16} \) mathematical models of data processing and technical devices.

Research data [1-5] have shown the possibility of signal amplification using Raman spectroscopy up to \( 10^6 \) - \( 10^7 \) using graphene [6-20] and silicon nanofibers. To increase the resolution, especially when detecting cancer cells [21-24], bilirubin [9, 15], combinations of various metal nanoparticles Ag, Au, Cu [11] are used to obtain plasmon effects on silicon wafers with surface modification by inhomogeneities in the form of vertical and horizontal nanofibers [12].

The prospect of increasing the resolution is estimated - identifying interdependencies between multi-dimensional parameters when processing a large number of statistical data [25,26].
The paper [14] revealed the possibility of constructing mathematical models and solving the problem of obtaining accuracy up to $10^{-16} - 10^{-18}$ using the construction of systems and solving systems with multi-dimensional correlation mathematical equations and taking into account the interdependence of parameters.

According to [27-30], the accuracy of solving multi-dimensional correlation mathematical equations in $10^{-14} - 10^{-16}$ in automatic mode has been consistently obtained for a long time. However, the resolution is not sufficient since data is obtained only in the range $10^{-1} - 10^{-7}$. This is not acceptable since a resolution of $10^{-14} - 10^{-16}$ with the currently obtained accuracy is required when solving multi-dimensional correlation mathematical equations.

Many problems, in this case, are that the compilation and solution of systems of correlation equations requires a large number of such equations in the system, for example, up to 81 equations. Many equations are necessary to account for all significant 9 peaks of the Raman spectrum in two directions across X and along Y of the object of study. The compilation and solution of such several correlation equations in the system are impossible at this technical level. Therefore, a simplified compilation and solution of a system of such equations do not provide the required reliability.

In these studies, the Bayes hypothesis is used to multiply the obtained interdependent (correlation) probabilities:

$$P_0 = P_1/P_{1c} \cdot P_2/P_{2c} \cdot \ldots \cdot P_n/P_{nc},$$

where $P_0$ is the total probability of occurrence of events; $P_1/P_{1c}, P_2/P_{2c}, ..., P_n/P_{nc}$ are the probabilities of dependent (correlation) events [25,26].

However, it is almost impossible to obtain interdependent probabilities for a large number of parameters and compose and solve a multi-dimensional system of up to 81 equations. In [14, 27, 28, 30], solutions were obtained only for two-dimensional correlation mathematical models and [29] for three-dimensional ones.

Therefore, in order to apply the Bayes hypothesis for a large number of parameters, it is necessary to mathematically transform all interdependent (correlation) parameters into an independent (uncorrelated) form. Then you can use the Bayes hypothesis for independent probabilities:

$$P_0 = P_1 \cdot P_2 \cdot P_3 \cdot \ldots \cdot P_n,$$

where $P_0$ is the total probability of occurrence of events; $P_1, P_2, P_3, ..., P_n$ are the probabilities of independent events.

Research in [31] aims to solve the problem of converting dependent (correlation) data into an independent form and vice versa from independent data into a dependent form. In these studies, the mathematical transformation into an independent form is performed simultaneously with the solution of a system of correlation multi-dimensional equations.

2. Materials and Methods

For clarity, the preliminary transformation of correlation data into an independent firm with the solution of the problem of detecting the intersection of ellipses of correlation and non-correlation data is shown in Figure 1 when generating correlation and non-correlation data according to the normal law of distribution ellipses using experimental data [31].
Figure 1. Mathematical transformation of the two-dimensional dependence (correlation) of the data in the independent (no correlation) when generated according to the normal law with the conservation and real parameters of the ellipses values of the peaks of the Raman spectrum of polyester fibers with silver nanoparticles (+ +) and without nanoparticles (· ·): (a) – 288 generate correlation data ellipses values of the peaks of the Raman spectrum; (b) – data conversion ellipses, Fig. 1a in a non-correlative form with the preservation of mathematical expectations and mean square deviations; (c) – the change in Fig. 1b during the generation of mean square deviations up to the intersection with one point of the ellipses of the Raman spectrum with silver nanoparticles and without nanoparticles; (d) – generation of 28800 independent data of Raman spectrum ellipses with silver nanoparticles and without nanoparticles with an 8-fold increase in the values of mean square deviations to the intersection with one point of the Raman spectrum ellipses with silver nanoparticles and without nanoparticles.

The transformation is based on obtaining a diagonal matrix of eigenvalues $\lambda 0 = \text{eigenvals}(r_{XY})$ from the correlation matrix $r_{XY}$ [31].

Finding the diagonal matrix of the eigenvalues of the correlation matrix is mathematically worked out in [31]. For example, for two-dimensional data for the correlation dependence on X and Y, when solving a specific problem of converting Raman spectra, we obtain:

$$\sum \Rightarrow \left( \begin{array}{cc} 1 & r_{XY1,j} \\ r_{XY1,j} & 1 \end{array} \right) \text{eigenvals} \left( \begin{array}{cc} 1 & r_{XY1,j} \\ r_{XY1,j} & 1 \end{array} \right) \rightarrow \left( \begin{array}{cc} 1 - r_{XY1,j} \\ r_{XY1,j} \end{array} \right) \left( \begin{array}{cc} 1 - r_{XY1,j} \\ r_{XY1,j} \end{array} \right) \sum 1 \Rightarrow \left( \begin{array}{cc} 1 - r_{XY1,j} & 0 \\ 0 & 1 + r_{XY1,j} \end{array} \right)$$
Then the transformation of mathematical models with silver nanoparticles and without nanoparticles into an uncorrelated form with simultaneous solution of a system of equations with uncorrelated matrices will have the following form for Ag9-12 (i=0; j=4):

\[
\sum_0 := \begin{pmatrix} 1-rXY0_{i,j} & 0 \\ 0 & 1+rXY0_{i,j} \end{pmatrix} \quad \sum_1 := \begin{pmatrix} 1-rXY1_{i,j} & 0 \\ 0 & 1+rXY1_{i,j} \end{pmatrix}
\]

\[
f(x, y) := \ln \left( \frac{1}{2\pi \sigma Y_j \sigma X_i} \sum_n \left[ \begin{array}{c} x - \text{MENX}_1 \\ y - 0 \end{array} \right] \right) + \frac{1}{2} \left[ \begin{array}{c} x - \text{MENX}_0 \\ y - 0 \end{array} \right] \sum_0^{-1} \left[ \begin{array}{c} x - \text{MENX}_0 \\ y - 0 \end{array} \right] \sum_1^{-1} \left[ \begin{array}{c} x - \text{MENX}_1 \\ y - 0 \end{array} \right] 1
\]

\[
g(x, y) := \ln \left( \frac{1}{2\pi \sigma Y_j \sigma X_i} \sum_n \left[ \begin{array}{c} x - \text{MENX}_0 \\ y - 0 \end{array} \right] \right) + \frac{1}{2} \left[ \begin{array}{c} x - \text{MENX}_0 \\ y - 0 \end{array} \right] \sum_0^{-1} \left[ \begin{array}{c} x - \text{MENX}_0 \\ y - 0 \end{array} \right] \sum_1^{-1} \left[ \begin{array}{c} x - \text{MENX}_0 \\ y - 0 \end{array} \right] 1
\]

x:=610.0 \quad y:=0

Given

\[f(x, y) = 0 \quad g(x, y) = 0\]

\[v2:= \text{Find}(x, y) \quad v2 = \left(615.635489, 0\right)\]

\[f(v2_0, v2_1) = 5.664475168578822 \times 10^{-14} \quad g(v2_0, v2_1) = 0\]

\[R0 := \left(\frac{v2_0 - \text{MENX}_0}{\sigma X_0} \right) \sum_0^{-1} \left(\frac{v2_1 - 0}{\sigma Y_0} \right) \sum_1^{-1} \left(\frac{v2_0 - \text{MENX}_0}{\sigma X_0} \right) \sum_0^{-1} \left(\frac{v2_1 - 0}{\sigma Y_0} \right)\]

\[R1 := \left(\frac{v2_0 - \text{MENX}_0}{\sigma X_0} \right) \sum_0^{-1} \left(\frac{v2_1 - 0}{\sigma Y_0} \right) \sum_1^{-1} \left(\frac{v2_0 - \text{MENX}_0}{\sigma X_0} \right) \sum_0^{-1} \left(\frac{v2_1 - 0}{\sigma Y_0} \right)\]

\[pQ0 := \frac{1}{2\pi \sigma Y_j \sigma X_i} e^{-\frac{1}{2} R0} \quad pQ1 := \frac{1}{2\pi \sigma Y_j \sigma X_i} e^{-\frac{1}{2} R1}\]

\[pQ0 = 1.4884714041498341 \times 10^{16} \quad pQ1 = 1.4884714041499187 \times 10^{16}\]

\[QAg := \frac{1}{pQ0, pQ1} \quad QAg = 6718301723580423 \quad QAg = 6.718 \times 10^{15}\]

The solution of the system of equations (3) gives the result:

\[pQ0=1.4884714041498341 \times 10^{16} \quad pQ1=1.4884714041499187 \times 10^{16}, \quad \text{which is acceptable both in accuracy and resolution.}\]
To verify the reliability, it is also necessary to evaluate the graphical representation of the intersection of the distribution ellipses of the Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles.

\[
\text{Figure 2. Validation by graphical representation of the intersection of ellipses of the distribution of Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles.}
\]

It is also necessary to check the likelihood of the intersection of the distribution ellipses with correlated and uncorrelated peak values in experiments on Raman spectra. Instead of binding to the axis MENY_0=0 and MENU_1=0 according to (3) Figure 2, we use the solution (4) Figure 3 with binding to the axis MENY_0 and MENU_1:

\[
f(x, y) := \left[ \ln \left( \frac{1}{(2\pi)^{0.5}} \right) \right] + \frac{1}{2} \sum_{i=1}^{\infty} \frac{x - \text{MENX}_i}{\sigma_{\Delta X_i}} \cdot \frac{y - \text{MENY}_i}{\sigma_{\Delta Y_i}} \cdot \sum_{j=1}^{\infty} \frac{x - \text{MENX}_j}{\sigma_{\Delta X_j}} \cdot \frac{y - \text{MENY}_j}{\sigma_{\Delta Y_j}}
\]

\[
g(x, y) := \left[ \ln \left( \frac{1}{(2\pi)^{0.5}} \right) \right] + \frac{1}{2} \sum_{i=1}^{\infty} \frac{x - \text{MENX}_i}{\sigma_{\Delta X_i}} \cdot \frac{y - \text{MENY}_i}{\sigma_{\Delta Y_i}} \cdot \sum_{j=1}^{\infty} \frac{x - \text{MENX}_j}{\sigma_{\Delta X_j}} \cdot \frac{y - \text{MENY}_j}{\sigma_{\Delta Y_j}}
\]

\[
x := 610.0 \quad y := 3000
\]

**Given**

\[
f(x, y) = 0 \quad g(x, y) = 0
\]

\[
v_2 := \text{Find}(x, y) \quad v_2 = 615.936996
\]

https://biointerfaceresearch.com/
The solution of the system of equations (4) gives the result:

\[ pQ_0 = 1.245798043704045 \times 10^{-16} \]

\[ pQ_1 = 1.2457980437041158 \times 10^{-16} \]

which is acceptable both in accuracy and resolution.

To confirm, it is also necessary to evaluate the intersections of the distribution ellipses of the Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles with the initial experimental correlated data with the results already obtained for (3) – (4):

\[ \sum_0 = \begin{pmatrix} 1 & r_{XY0_{i,j}} \\ r_{XY0_{i,j}} & 1 \end{pmatrix} \quad \sum_1 = \begin{pmatrix} 1 & r_{XY1_{i,j}} \\ r_{XY1_{i,j}} & 1 \end{pmatrix} \]

**Figure 3.** Graphical validation based on the image of the intersection of ellipses of the distribution of Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles with reference to the MENY0 - MENY1 axis.
Given 

\[ f(x, y) = 0 \quad g(x, y) = 0 \]

\[ v_2 := \text{Find}(x, y) \quad v_2 = \left( \frac{449.955654}{2471.429168} \right) \]

\[ f(v_{20}, v_{21}) = -3.504467994814062 \times 10^{15} \quad g(v_{20}, v_{21}) = -5.142505384808134 \times 10^{-18} \]

\[ R_0 := \left( \frac{v_{20} - \text{MENX}_{0_i}}{\sigma \Delta X_{0_i}} \quad \frac{v_{21} - \text{MENY}_{0_j}}{\sigma \Delta Y_{0_j}} \right) \sum_{i}^{-1} \left( \frac{v_{20} - \text{MENX}_{0_i}}{\sigma \Delta X_{0_i}} \quad \frac{v_{21} - \text{MENY}_{0_j}}{\sigma \Delta Y_{0_j}} \right) \]

\[ R_1 := \left( \frac{v_{20} - \text{MENX}_{1_i}}{\sigma \Delta X_{1_i}} \quad \frac{v_{21} - \text{MENY}_{1_j}}{\sigma \Delta Y_{1_j}} \right) \sum_{i}^{-1} \left( \frac{v_{20} - \text{MENX}_{1_i}}{\sigma \Delta X_{1_i}} \quad \frac{v_{21} - \text{MENY}_{1_j}}{\sigma \Delta Y_{1_j}} \right) \]

\[ pQ_0 := \frac{1}{\left(2 \pi \left( \sum_0 \right)^{0.5} \right)^{e^{-1}}} \quad pQ_1 := \frac{1}{\left(2 \pi \left( \sum_1 \right)^{0.5} \right)^{e^{-1}}} \quad pQ_0 = 0.01100562484527942 \quad pQ_1 = 0.01100562484527937 \quad QAg := \frac{1}{pQ_0} \quad QAg = 90.863 \]

To verify the reliability, it is also necessary to evaluate the graphical representation of the intersection of the ellipses of the distribution of Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles with the initial experimental correlated data:

https://doi.org/10.33263/BRIAC132.139 

https://biointerfacereasearch.com/
3. Results and Discussion

The general results of solving the problem for all peaks from \(i=0...8\) to \(j=0...8\) in total 81 equations with mathematically transformed uncorrelated statistical data of the Raman spectrum of polyester fiber with silver nanoparticles are shown in the following figures.

**Figure 4.** Graphical representation of the intersection of ellipses of the distribution of Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles according to the initial experimental correlation data.

**Figure 5.** Matrix of the results of assessing the reliability of the intersection of the ellipses of the distribution of Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles according to statistical data converted to uncorrelated form.

**Figure 6.** Data of logarithmic values of the reliability of the intersection of ellipses of the distribution of Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles according to the results of conversion to uncorrelated form.
The total value of the sum of logarithms of confidence with silver nanoparticles and without nanoparticles according to the results of conversion to uncorrelated form:

\[
SQL := \sum_{j=0}^{8} \sum_{i=0}^{8} Q_{i,j} \quad SQL = 547.123 \quad SQL1 := \frac{SQL}{81} \quad SQL1 = 6.755
\]

Figure 7. The logarithm of the confidence of the intersection of the ellipses of the distribution of Raman spectra of polyester fibers with silver nanoparticles and without nanoparticles in columns \( j=0-8 \) of the matrix of logarithmic confidence values Figure 6.

4. Conclusions

The use of a mathematical method for converting statistical correlation data into an uncorrelated form while simultaneously solving a system of equations for multi-dimensional Raman spectra of polyester fibers without nanoparticles and with silver nanoparticles makes it possible to increase the sensitivity of nanosilver identification up to \( 10^{547} \) times and, consequently, increase the resolution of the method.

The method uses the possibilities of applying the Bayes hypothesis when multiplying the probability densities in the intersection of the ellipses of the distribution of the values of the peaks of the Raman spectra of polyester with silver nanoparticles without nanoparticles with transformed correlated statistical data in an uncorrelated form.

The transformation of correlated statistical data into an uncorrelated form is carried out on the basis of finding diagonal matrices from eigenvalues (numbers) correlation matrices used to compile systems of equations.

This study selected correlation matrices of two-dimensional parameters: \( X \) - across the fibers and \( Y \) - along the fibers. The transformation and solution equations system is designed to replace the traditional ellipse equation with a vector-matrix form using a diagonal matrix of eigenvalues, which allows uncorrelated transformations and simultaneous rotations and displacement of the coordinate axes of the ellipses to the origin of the coordinates MENY0 and MENY1. As a result, the values of probability densities for the transformed uncorrelated transformed data for the peak with \( i=0 \) and \( j=4 \) were obtained: \( pQ1=1.4884714041498341 \times 10^{-16} \) or the inverse value (confidence) \( QA_g=6.718 \times 10^{15} \) compared to the original correlated data: \( pQ1=0.01100562484527942 \) with the inverse value (Q-factor) \( QA_g=90.863 \). These results for comparison are also obtained in this article.

A system of equations has also been compiled and solved for conversion to an uncorrelated form with the ellipses of distributions bound to the axes: MENY0, MENX0 –
MENY1, MENX1. The values obtained are: \(pQ1 = 1.2457980437041158 \times 10^{-16}\) or the inverse value (confidence) \(QAg = 8.027 \times 10^{15}\).

A matrix of all 81 values of \(Q\) and \(\log Q\) is also obtained. Since when multiplying all the independent terms of the matrix \(Q\), we get \(10^{547}\), i.e., more than \(10^{307}\) than the computing capabilities of a conventional computer, we have to use logarithmic values of the terms of the matrix \(\log Q\). In this case, all the values of the terms are added for the \(\log Q\) matrix, and the sum value 547 is obtained. For one member of this matrix, we get an average value of 6.755 or \(10^{6.755} = 5.683 \times 10^6\).

Further research aims to compile and solve a system of equations for converting correlated data into uncorrelated data for different concentrations of deposited colloidal silver: 5% and 11%, to compare with 17% of the already proven research method in this work.

**Funding**

This research received no external funding.

**Acknowledgments**

This research has no acknowledgment.

**Conflicts of Interest**

The authors declare no conflict of interest.

**References**


