Application of the LASSO algorithm for fitting the multiexponential data of the NMR relaxometry

Vladimir Bochkarev$^{1,c}$, Vladimir Tyurin$^1$, Andrey Savinkov$^1$, Bulat Gizatullin$^2$

$^1$Kazan Federal University, Kremlevskaya 18, Kazan 420008, Russia
$^2$Current address: Ilmenau Technical University, Unterpoerlitzer 38, Ilmenau 98693, Germany

E-mail: vbochkarev@mail.ru

Abstract. The problem of time series fitting by the sum of several exponentials with different decay parameters often arises analyzing the data of a physical experiment. In particular, such a problem arises determining the nuclear transverse relaxation times from the spin-echo decay NMR data. Mathematically, the problem can be formulated as a solution of the Fredholm integral equation of the first kind with an exponential kernel, or an equivalent system of linear equations. To solve this equation, various regularization methods are used, based on the $L_2$-norm in most cases. We report the application of the $L_1$-regularization for the problem, using the LASSO algorithm. A comparison was made between the results of $L_1$- and $L_2$-regularization on model time series, which were the sum of several exponential decays with noise. Also the $L_1$- and $L_2$-regularization were applied to analyze the $^1$H NMR experimental data of the spin-echo decays of n-hexane adsorbed in the porous volume of aluminium poor MFI-type zeolites known as ZSM-5. It was found out that the $L_1$-regularization is an effective method for determination of the transverse relaxation times from NMR data.

1. Introduction

The main goal of the most pulsed NMR experiments to measure the evolution of nuclear transverse or longitudinal magnetization is to determine the nuclear spin-spin or spin-lattice relaxation rates in the spin system, that is, the problem is reduced to determining the decrements of multi-exponential decays in the measured NMR data. The nuclear transverse magnetization data can be acquired using a conventional CPMG pulse train that generates the spin-echo decay array $A(\tau)$, where $\tau$ is the spin echo time samples. The decay of the spin echo signal occurs due to processes of the nuclear spin-spin relaxation, which can be described by the $T_2$ times with corresponding amplitudes $S(T_2)$. The amplitudes $S(T_2)$ are related to the parts of molecules with the spin-spin relaxation times $T_2$ in the spin system. The accurate definition of the former is important to study of dynamical processes such as a molecular exchange and others. Measured spin-echo decay data $A(\tau)$ relates to $S(T_2)$ via the Fredholm integral equation of the first kind with an exponential kernel:

$$
A(\tau) = \int_0^\infty e^{-\tau/T_2} S(T_2) dT_2,
$$

where $E(\tau)$ is the experimental noise. Equation (1) can be approximated by a discretized system of the linear algebraic equations which can be represented in a matrix form as $A = KS$, where $K(\tau,T_2) = e^{-\tau/T_2}$ is the exponential kernel.
The solution of the system of linear equations, as well as the initial equation (1), is an incorrectly posed problem, because small changes in the acquired data $A(\tau)$ can lead to significant changes in the calculated distribution of the spin-spin relaxation times $S(T_2)$. As a result, regularization is used to find a stable solution. In addition, the solutions are reduced by applying some restrictions. First, this is the non-negativity of the amplitudes $S(T_2)$ of the $T_2$ times; and secondly, $T_2$ times are restricted in a certain range of values:

$$ S(T_2) \geq 0, \quad T_{2,\text{min}} \leq T_2 \leq T_{2,\text{max}} $$

The most common numerical method is based on the Tikhonov’s $L_2$-norm regularization (see Ref. [1,2] and others). In the framework of this approach, a vector $S(T_2)$ is adopted as a solution of (1) if it minimizes the following expression (taking into account the restrictions (2)):

$$ \| K S - A \|_2^2 + \lambda \| S \|_2^2 $$

The quantity $\| S \|_2^2 = \int S^2(T_2) dT_2$ in Eq.(3) does not have a specific physical meaning. On the contrary, the $L_1$-norm $\| S \|_1 = \int |S(T_2)| dT_2$ is proportional to the number of protons contributing to the spin echo signal. Therefore, the requirement of boundedness of the $L_1$-norm seems to be physically justified. When using the $L_1$-regularization a vector $S(T_2)$ is adopted as a solution minimizing the following expression:

$$ \| K S - A \|_1^1 + \lambda \| S \|_1 $$

As a rule, $L_1$-regularization gives rise to sparse solutions in which there are a small number of non-zero elements. This can be especially important in analysis of the spin echo decays, since real physical systems usually are characterized by limited number of different spin-spin relaxation times. Authors of Ref. [3] have applied the $L_1$-regularization to estimate the distribution of the transverse relaxation rate in the experimental and model spin echo decay data.

The problem (4) taking into account the restrictions (2) does not have an analytical solution. However, it is possible to find a solution by numerical methods, in particular, using the algorithm LASSO [4]. In this paper, we successfully applied the $L_1$-regularization using the LASSO algorithm (as part of the scikit-learn library [5] of the programming language Python) to find the distribution of spin-spin relaxation times from the experimental $^1H$ NMR data. The data were acquired by the CPMG train in the sample of Al-poor ZSM-5 zeolite. Since the porous volume of the sample filled with $n$-hexane consists of pores of very different sizes (from micropores $\sim 5 \cdot 10^{-10}$ m to macropores $\sim 10^{-5}$ m), several different spin-spin relaxation times of $^1H$ nuclei are expected in the system. The standard deviations of $T_2$ values as well as the corresponding contributions obtained from the model data by $L_1$- and $L_2$-regularization were compared.

2. Results and Discussion

The $L_1$-norm regularization was tested on experimental data measured by $^1H$ NMR in the sample of Al-poor ZSM-5 zeolite saturated with n-hexane. The porous medium of ZSM-5 zeolites, having the MFI structure, is rather complex and includes pores of three different types (micropores, mesopores and macropores) that may lead to different dynamics and behavior of adsorbed molecules in the sample. As a result, the spin-echo decay of $^1H$ nuclei measured in the n-hexane saturated sample may contain several contributions characterized by different values of the $T_2$ time. Detailed description of the porous media of the ZSM-5 zeolites is available in Ref.[6-8] and others. Powder of Al-poor ZSM-5 zeolite (Si/Al ration is 25) was obtained from Lulea University of Technology (Sweden). The main preparation procedure of the zeolite powder is described in Ref.[9]. Liquid n-hexane was added to the powder with subsequent vacuuming. Amount of adsorbed n-hexane was controlled by weighting on the standard balance. Content of n-hexane was 668 mg/g (40.1% wt.) in the sample.

Pulsed NMR spectrometer Proton 20M (Chromatec ©) was used to study the $^1H$ magnetic relaxation. The measurements were performed at radio-frequency 20 MHz, temperature of the sample
was 310±0.1 K. Spin-echo decay of \(^1\)H nuclei of adsorbed n-hexane was measured by CPMG pulse train with \(\tau = 50 \mu s\). Also, FID was recorded after first \(\pi/2\) rf-pulse during 50 \(\mu s\) after the dead time period of 14 \(\mu s\) to reach the solid-like component with \(T_2 \approx 10 \mu s\).

Figure 1a presents estimations of the distribution of spin-spin relaxation times for the Al-poor ZSM-5 saturated with n-hexane sample obtained in three ways: L\(_2\)-regularization using CONTIN algorithm [10] (green curve), L\(_2\)-regularization (blue curve) using the scikit-learn library and L\(_1\)-regularization using LASSO algorithm implemented to the scikit-learn library (red curve). The spin-echo decays recovered from the corresponding distributions of the spin-spin relaxation times \(S(T_2)\) (Fig.1a) are presented in Fig. 1b. The experimental data are presented as well. It can be seen that all three numerically reconstructed spin echo decays lie rather close to the experimental one.

![Figure 1](image)

Figure 1. a) Distribution of \(T_2\) times in the Al-poor ZSM-5 zeolite saturated with n-hexane, obtained with: L\(_2\)-regularization using CONTIN algorithm (green curve), L\(_2\)-regularization using the scikit-learn library (blue curve) and L\(_1\)-regularization using LASSO algorithm implemented to the scikit-learn library (red curve); b) Experimental spin-echo decay data (black circles) and the model spin-echo decays obtained from the appropriate distributions of \(T_2\) times presented on the plot a).

The \(T_1-T_2\) joint distribution was studied in the sample of Al-free ZSM-5 zeolite (known as silicalite-1), fully saturated by n-hexane by the two dimensional \(T_1-T_2\) \(^1\)H NMR relaxometry [11]. The results of the study allows to identify all four peaks in \(S(T_2)\), observed in Fig.1a. The first peak with the smallest \(T_2\) value is most likely related to n-hexane molecules in the channels of the silicalite-1 crystal (peak “A” in Fig.3a in Ref.[11]). Second and third rather weak peaks in Fig.1a with values of \(T_2\) ~0.6 ms and ~3 ms are related to the molecules of n-hexane in porous volume of Al-poor ZSM-5 zeolite attributed to mesopores with typical size of about 2-50 nm and relative small total porous volume of ~ 0.5% (peaks “B” and “D” in Fig.3a in Ref.[11]). Last peak in Fig.1a with \(T_2 \approx 20 \text{ ms}\) can be attributed to the n-hexane molecules in the intercrystallite volume of the sample (peak “C” in Fig.3a in Ref.[11]). Thus, all three applied regularization methods give sufficiently reasonable estimate of the \(T_2\) time distribution \(S(T_2)\) in the spin system.

However, obtained results do not allow to compare quantitatively the regularization methods, since exact values of the dynamical parameters of the spin system such as the spin-spin relaxation times and the corresponding contributions in the spin echo decay signal are not well defined. For the quantitative comparison, statistical modeling was performed. For this purpose, a spin echo decay signal was generated, which is the sum of four exponential components with the values of the times \(T_2\) and the contributions corresponding to the values obtained from the experimental data by the L\(_1\)-regularization using method LASSO (see columns 1 and 2 in Table 1). Then the normal white noise, which is close to that observed in the experiment, was added to the generated spin echo decay. The L\(_1\)-regularization using LASSO algorithm and L\(_2\)-regularization were applied to the model data. The quantitative comparison of the obtained central values of the \(T_2\) times and the corresponding contributions to the spin echo decay signal (which are the integral values under the peaks in \(T_2\)) with the model values are given in Table 1.
Table 1. Standard deviations of the $T_2$ times and the corresponding contributions (denoted as “I” in the table) to the model spin-echo decay obtained from the $L_1$- and $L_2$-regularization.

| Number of component | $T_2$, ms | I, a.u. | Standard deviation of $T_2$, ms | Standard deviation of I, a.u. |
|---------------------|-----------|--------|-------------------------------|----------------------------|
|                     | L₁        | L₂     | L₁                             | L₂                          |
| 1                   | $1.62\times10^{-2}$ | 160.6  | $4.80\times10^{-5}$            | 10.5                        | 164.0                       |
| 2                   | 0.613     | 133.5  | $2.77\times10^{-2}$            | 0.12                        | 4.07                        | 8.57                        |
| 3                   | 3.275     | 67.7   | 0.35                          | 1.63                        | 6.26                        | 29.9                        |
| 4                   | 13.22     | 1175.6 | 4.14                          | 4.61                        | 5.23                        | 19.03                       |

As it was shown above both $L_1$- and $L_2$- regularization methods are rather successful in the qualitative estimations of the spin-spin relaxation times in the spin system, but the $L_1$-norm regularization gives rise to the narrowest $T_2$-peaks. Moreover, analysis of model data showed that $L_2$-norm regularization tends to systematic displacement and broadening of the $T_2$-peaks. Also, when the noise level increases, there is a tendency for the 3rd and 4th components in the model spin echo decay to merge if the $L_2$-regularization is used, but the $L_1$-regularization estimate allows a stable separation of these components. Thus, $L_1$-norm regularization using LASSO algorithm causes rather narrow peaks in $S(T_2)$. Furthermore, their positions are most consistent with the a priori values.

3. Conclusion

In this paper, we compared the efficiency of $L_2$ and $L_1$-regularization (using LASSO algorithm) methods when the distribution of the spin-spin relaxation times was estimated from the spin-echo decay data. Both experimental data for the Al-poor zeolite sample and model data were analyzed. The application of $L_1$-regularization using LASSO algorithm gives better results, in particular, the broadening of the $T_2$-peaks is less typical for it, and the integral contributions of the $T_2$-components are better reproduced. Thus, the $L_1$-regularization using LASSO algorithm is suggested as a promising method for the analysis of NMR-relaxometry data.

Acknowledgements:

This research was financially supported by the Russian Government Program of Competitive Growth of Kazan Federal University, state assignment of Ministry of Education and Science, grant agreement № 34.5517.2017/6.7.

References

[1] Tikhonov A N and Arsenin V Y 1977 Solution of Ill-posed Problems (Washington: Winston & Sons)
[2] Kroeker R M and Henkelman R M 1986 J. Magn. Reson. 69 218–235
[3] Berman P, Levi O, Parmet Y, Saunders M and Weisman Z 2013 Concepts in Magnetic Resonance Part A 42A(3) 72-88
[4] Tibshirani R 1996 Journal of the Royal Statistical Society. Series B 58 (1) 267
[5] Pedregosa F et al. 2011 Scikit-learn: Machine Learning in Python JMLR 12 2825-2830
[6] Meier W M and Olson D H 1987 Atlas of Zeolite Structure Types (London: Butterworth)
[7] Olson D H, Kokotailo G T, Lawton S L and Meler W M 1981 J. Phys. Chem. 85, 2238
[8] Zhang L, van Laak A N C, de Jongh P E and de Jong K P 2010 Zeolites and catalysis: synthesis, reactions and applications, edited by Čejka J, Corma A, Zones S (Wiley-VCH Verlag CmbH & Co. KGaA).
[9] Mostowicz R, Crea F and Nagy J 1993 Zeolites 13 678
[10] Provender S W 1982 Computer Physics Communications 27 213-227
[11] Gizatullin B I, Savinkov A V, Shipunov T V, Melnikova D L, Dorogitzky M M and Skirda V D 2018 Magnetic Resonance in Solids (electronic journal) 20 (1) 18102