International Journal of Bioelectromagnetism
Vol. 14, No. 1, pp. 46-50, 2012

Comparison of figures of merit in linear inverse problems

Bojana Petkovic[a], Roland Eichardt[a,b], Daniel Baumgarten[a,b], Uwe Graichen[a], Jens Haueisen[a,b], Luca Di Rienzo[c]

[a]*Institute of Biomedical Engineering and Informatics, Ilmenau University of Technology, Ilmenau, Germany*

[b]*Biomagnetic Center, Clinic for Neurology, University Hospital Jena, Jena, Germany*

[c]*Politecnico di Milano, Dipartimento di Elettrotecnica, 20133 Milano, Italy*

Correspondence: B. Petkovic, Institute of Biomedical Engineering and Informatics, Ilmenau University of Technology, Ilmenau, Germany

E-mail: bojana.petkovic@tu-ilmenau.de, phone +49 3677 69 1309, fax +49 3677 69 1311

Abstract. Conditioning is a very important factor to be considered when solving linear inverse problems. Many different measures of conditioning exist, but some of them are not always numerically stable or can indicate that the problem is ill-conditioned, even though in reality it is not. We theoretically consider and numerically compare four figures of merit: the condition number with respect to the L_2 norm, Skeel condition number, ratio of the largest and mean singular value of a matrix and a novel figure of merit, based on the linear dependency between rows in underdetermined linear inverse problems and between columns in overdetermined problems. Numerical simuations show that all the figures of merit have low values when the linear inverse problem is well-conditioned and their values increase if the condition decays. The newly proposed figure of merit has three advantages: it does not require singular value decomposition, it enables the comparison of different sensor arrays and it is independent of row and column scaling in underdetermined and overdetermind problems, respectively. These features of the proposed figure of merit are promising for further applications.

Keywords: Figures of merit; Ill-conditioning; Linear inverse problems

1. Introduction

In this paper, we consider a linear model describing the relationship between the three components of the magnetic moments of magnetic dipoles $\vec{m}$ and one component of the measurement values of the magnetic flux density $\vec{b}$, according to

$$L \cdot \vec{m} = \vec{b} \ , \tag{1}$$

where L is a lead field matrix, $L \in R^{n,3m}$. The lead field matrix contains information about the sensor setup and the source model.

When estimating the magnetic moments $\vec{m}$ from the magnetic flux density values $\vec{b}$, the condition of the linear inverse problem becomes very important. Many approaches exist, which try to improve this property of the linear inverse problem. The core measure, on which such improvement methods rely, is the figure of merit indicating the condition of L. We consider and compare the following figures of merit: condition number with respect to L_2 norm [1, 2], Skeel condition number [3] and the ratio between largest and mean singular values [4]. We also propose a novel figure of merit based on the dependency between rows of a lead field matrix in underdetermined linear inverse problems and dependency between columns of a matrix in overdetermined linear inverse problems.

2. Methods

2.1. Condition number CN

A condition number with respect to the L_2 norm, $CN(L)$, is defined as

$$CN(L) = \|L\| \cdot \|L^{+}\| \ , \tag{2}$$

where $\| \;\|$ denotes the L_2 norm.

This condition number is equal to the ratio of the largest and the smallest singular value of a matrix L. It is used as an indicator for the stability of the inversion process [1], or as *a priori* accuracy estimator for the inverse problem [2].

2.2. Skeel condition number

The Skeel condition number is first proposed for square matrices in [3] and then generalized to rectangular matrices in [4]. It is defined as

$$Skeel(L) = \left\| |L| \cdot |L^{+}| \right\|, \tag{3}$$

where $| \;|$ indicates that all the elements of the matrices L and L^{+} are replaced by their absolute values and $\| \;\|$ denotes the L_2 norm.

2.3. Figure of merit ρ

The condition of a matrix can also be measured using the ratio between the largest and the mean of all n singular values [5],

$$\rho(L) = \frac{\sigma_1(L)}{\frac{1}{n}\sum_{i=1}^{n}\sigma_i(L)}, \tag{4}$$

where σ_1 represents the largest singular value. In fact, the measure can also be interpreted as the average decay of the singular values, but inverted in order to get a similar behaviour as CN [5].

2.4. Rows dependency RD

Geometrical interpretation of the linear dependency between a matrix row/column and a space spanned by the all other rows/columns has been used before. The angle between a row of a matrix and a space spanned by all other rows is employed to estimate the degree of information this particular row adds to the other rows of a matrix [6]. If many rows add little or no additional information, the matrix is ill-conditioned. An approximate expression for the condition number CN of a well-scaled matrix in terms of the minimum angle between a column vector of a matrix and a linear subspace spanned by the remaining columns is derived in [7]. An interesting inequality is given in [8], showing that either the matrix is not well-scaled or the columns of a matrix are nearly dependent if the CN increases.

Let $L = \left[\vec{l}_1, \vec{l}_2, ..., \vec{l}_n\right]^{\mathrm{T}}$ be a lead field matrix of an underdetermined problem, where $\vec{l}_i$ is the i^{th} row vector of L and n corresponds to the number of sensors. By computing the mean value of the angles among all $\vec{l}_i$, we get a figure of merit of rows dependency, RD:

$$RD = \frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left| 90^{\mathrm{o}} - \left| \mathrm{Cos}^{-1}\left(\frac{\vec{l}_i \cdot \vec{l}_j}{\|\vec{l}_i\| \cdot \|\vec{l}_j\|}\right)\right|\right|}{\frac{n!}{2!(n-2)!}}. \tag{5}$$

In order to indicate a well-conditioned matrix by a small value, as in the figures of merit presented above, a subtraction of each obtained angle from 90^{o} is imposed in Eq. 5. $| \;|$ indicates that the absolute values of the angles and of the differences are used. $\| \;\|$ denotes the L_2 norm. $RD = 0^{\mathrm{o}}$ corresponds to the set of linearly independent rows, while a set of parallel vectors is denoted by $RD = 90^{\mathrm{o}}$. In overdetermined problems, this figure of merit calculates angles between columns.

2.5. Numerical simulations

We investigated the influence of the number of sensors in an array on the figures of merit presented above. The number of sensors in the array is increased from 5×5 to 15×15, while keeping a constant grid of 15×15 for the three component magnetic dipoles. Dipoles are placed in the area $(x_{\min}, x_{\max}) = (-0.02\,\mathrm{m}, 0.22\,\mathrm{m})$, $(y_{\min}, y_{\max}) = (-0.02\,\mathrm{m}, 0.22\,\mathrm{m})$, $0.09\,\mathrm{m}$ underneath a sensor array.

All the sensors are uniformly oriented along the $z-$direction and placed in the area $(x_{\min}, x_{\max}) = (0\,\mathrm{m}, 0.2\,\mathrm{m})$, $(y_{\min}, y_{\max}) = (0\,\mathrm{m}, 0.2\,\mathrm{m})$.

In a second simulation, we examined the influence of extensions of the source grid in both x and y directions on the presented figures of merit. We considered the grid of 15×15 for the three component magnetic dipoles, $0.05\,\mathrm{m}$ below the sensor array. The sensor area $(x_{\min}, x_{\max}) = (0\,\mathrm{m}, 0.2\,\mathrm{m})$, $(y_{\min}, y_{\max}) = (0\,\mathrm{m}, 0.2\,\mathrm{m})$ contains 10×10 sensors uniformly oriented in $z-$direction. The extension of the source grid is varied from $-0.04\,\mathrm{m}$ (smaller than sensor area) to $+0.04\,\mathrm{m}$ (larger than sensor area).

3. Results

The dependencies of the condition number CN, rows dependency RD, figure of merit ρ, and the Skeel condition number $Skeel$, on the number of sensors in x and y direction are presented in Fig.1a, Fig.1b, Fig.1c and Fig.1d, respectively. CN changes from 30.80 to 7.82×10^8, RD from 12.53 to 17.33, $Skeel$ from 37.96 to 9.00×10^8 and ρ from 3.18 to 32.22. The increase in CN and $Skeel$ is seven orders of magnitudes larger compared to RD and ρ. The figures CN and $Skeel$ show a similar slope, when increasing the number of sensors.

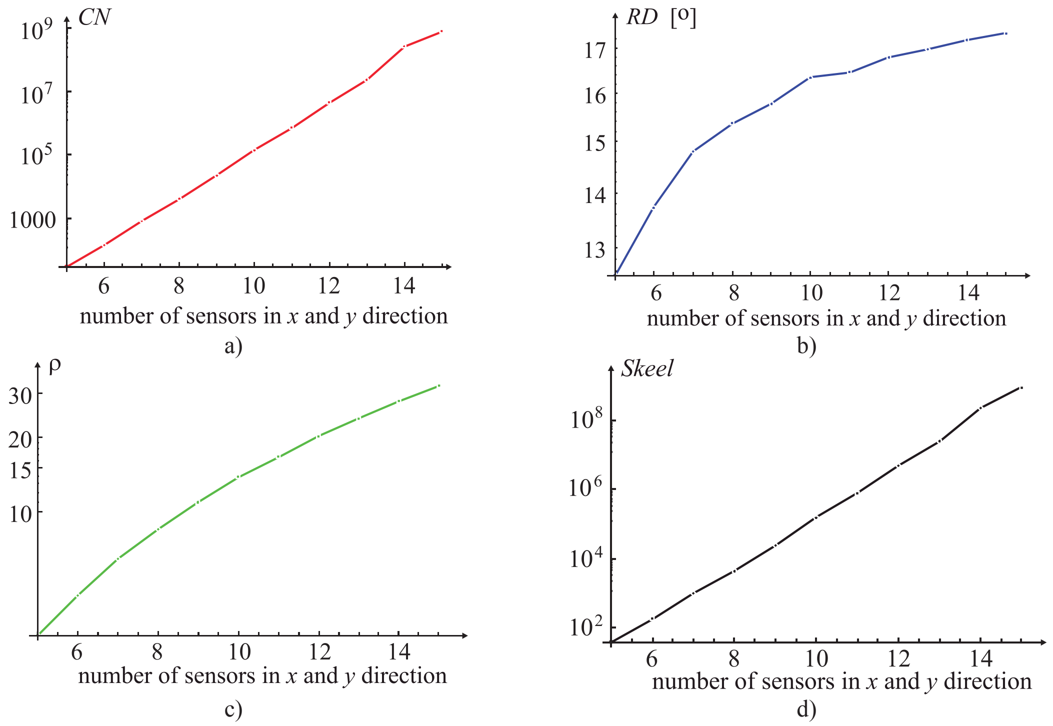

Figure 1. *Dependence of measures of conditioning of a lead field matrix: a) Condition number with respect to L_2 norm, CN, b) Rows dependency RD, c) Figure of merit ρ and d) Skeel condition number.*

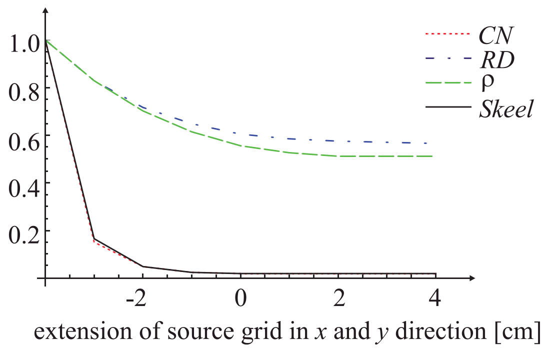

Figure 2. *The effect of increasing (positive values on abscissa) or reducing (negative values on abscissa) the square source grid area beyond the sensor area.*

Fig. 2 shows the dependencies of the condition number CN, rows dependency RD, figure of merit ρ, and Skeel condition number $Skeel$ on the extension of the source grid below the sensor area from $-0.04\,\mathrm{m}$ (smaller than sensor area) to $+0.04\,\mathrm{m}$ (larger than sensor area). The measures are normalized by their corresponding values for a source grid extension of $-0.04\,\mathrm{m}$. All the figures are dropping when extending the source grid. CN and $Skeel$ show considerably larger decays than RD and ρ.

4. Discussion

Conditioning is a very important concept to be considered when solving linear inverse problems. Therefore, we compared four measures of conditioning: condition number with respect to the L_2 norm, Skeel condition number, ratio between the largest and mean singular value and a newly formulated measure of angles between the rows of a matrix.

The condition number CN is well-known. However, it has several drawbacks [9]. First, it ignores a structure of a matrix with respect to scaling and/or sparsity. When the elements in one row or column of a matrix are significantly different in scale from the other elements, a condition number CN can be very high [10]. But, high condition number does not always mean a large sensitivity to error. This situation is known as artificial ill-conditioning [11].

The condition number CN is computed using only the largest and the smallest singular value, its computation might be numerically instable and it essentially depends on the smallest singular value. These drawbacks are partially overcome using the ratio between largest and the mean singular values of a matrix, defined as the figure of merit ρ. It also represents an average decay of the singular values, but inverted in order to get similar behaviour as CN. It is numerically more robust to compute than the condition number CN. In addition, it does not essentially depend on the smallest singular value and it includes information on all singular values of a lead field matrix.

Computation of the condition number CN requires both a solution space norm and a residual space norm. Different from CN, the Skeel condition number $\left\| |L| \cdot |L^+| \right\|$ is defined entirely in terms of the solution space norm since the matrix $|L| \cdot |L^+|$ is a mapping of the solution space onto itself. If used in overdetermined problems, the Skeel condition number defined as $\left\| |L| \cdot |L^+| \right\|$ is invariant under column scaling. In the case of underdetermined problems, the Skeel defined as $\left\| |L^+| \cdot |L| \right\|$ is invariant under row scaling. The Skeel condition number is equal to one for any matrix where only the entries $l_{i,i}$ are non-zero. Therefore, an overestimation of the ill-conditioning using the condition number CN in the case of a matrix containing non-zero entries only on positions i,i is overcome using the Skeel condition number.

A dependency between rows in underdetermined linear inverse problems is defined by a figure of merit RD. This measure of conditioning is calculated as the mean value of the angles between all the rows of a lead field matrix. The computation of RD does not require the singular value decomposition of a matrix and it enables the comparison of lead field matrices of different sizes. Furthermore, multiplication of all elements of a row by the same value influences only the norm of a row vector, but not the angles to other rows. Thus, RD is independent of row scaling. In overdetermined problems, this figure of merit calculates angles between columns and it is independent under column scaling. These features of the proposed figure RD are promising for further applications.

Similar to CN, the low values of RD, ρ and $Skeel$ indicate a well conditioned linear inverse problem. We are aware of the influence of a matrix dimension on the condition number CN. Our future work will focus on comparison of these figures of merit keeping the dimensions of a matrix constant.

Acknowledgements

This study was supported in parts by DFG GK 1567, TAB 2008VF0001 / 2008FE9048 under participation of the European Union within the European Funds for Regional Development.

References

[1] Rouve L., Schmerber L, Chadebec O, Foggia A. Optimal Magnetic Sensor Location for Spherical Harmonic Identification Applied to Radiated Electrical Devices, *IEEE Transactions on Magnetics*, 42(4): 1167-1170, 2006.

[2] Lee C, An K, Kim B, Im C, Jung H. Estimation of solution accuracy in magnetoencephalography by using the condition number from a leadfield matrix. In proceedings of the World Congress on Medical Physics and Biomedical Engineering 2006, 2710-2713.

[3] Skeel R. Scaling for numerical stability in Gaussian elimination. *J. Assoc. Comput. Mach*., 26(3): 494-526, 1979.

[4] Demmel J, Higham N. Improved error bounds for underdetermined system solvers, *SIAM J. Matrix Anal. Appl*, 14(1): 1-14, 1993.

[5] Eichardt R, Baumgarten D, Petković B, Wiekhorst F, Trahms L, Haueisen J. Adapting source grid parameters to improve the condition of the magnetostatic linear inverse problem of estimating nanoparticle distributions. In proceedings of the 11th Workshop on Optimization and Inverse Problems in Electromagnetism – OIPE 2010, 108-109.

[6] Sabatier P. On geophysical inverse problems and constraints, *Journal of Geophysics*, 43: 115-137, 1977.

[7] Chiang C, Chandler J. An approximate equation for the condition numbers of well-scaled matrices. In Proceedings of the IAJC-IJME International Conference, paper 036, 2008.

[8] Martinez J. A bound for the Moore-Penrose pseudoinverse of a matrix, *Commentationes Mathematicae Universitatis Carolinae*, 20(2): 357-360, 1979.

[9] Higham N. A survey of componentwise perturbation theory in numerical linear algebra. In: Gautschi,W. (ed.) Mathematics of Computation 1943–1993: A Half Century of Computational Mathematics. Proceedings of Symposia in Applied Mathematics, vol. 48, pp. 49–77. American Mathematical Society, Providence (1994)

[10] Stewart G. Collinearity and least squares regression, Statistical Science, 1987, 2(1): 68-100.

[11] Gentle J. Numerical Linear Algebra for Applications in Statistics. Springer, New York, 1998.