CONF- 9602101--9

Title:

SUBSPACE ANGLES:
A METRIC FOR COMPARISONS IN EEG AND MEG

RECEIVED
JUL 19 1996
OSTI

Author(s):

J. C. Mosher

Submitted to:

BIOMAGNETISM '96, Santa Fe, NM, February 1996

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Subspace Angles: A Metric for Comparisons in EEG and MEG

Mosher, J.C.

Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Introduction

In forward head modeling, various approximations are made in order to keep the problem tractable. Simplifications can yield models ranging from simple spherical models to multi-tessellated arbitrary surfaces in a boundary element model (BEM). Spherical head models differ in the number of shells and the assumed conductivities. Other assumptions in the BEM include the choice of basis sets, such as constant, linear, or quadratic variations of the voltages across the individual areal elements, or the selection of error-weighting method, such as collocation, Galerkin, or "direct" methods. Numerical versus analytic integration can also yield numerical differences. These differences in parameters and approximations can yield models whose external fields (EEG potentials or MEG magnetic fields) differ for the same internal source configuration. Quantitative measures are needed to determine if these differences are significant.

The conventional approach is to define the source to be an arbitrarily placed single current dipole, then compare the field pattern generated. If two forward models generate similar patterns for many different placements of the dipole, then the presumption is that the two models are comparable. Several different methods of comparison between models have been proposed, such as the residual variance (RV), which we briefly review. We then propose a new method that effectively performs a multidimensional correlation between the two models, yielding their *subspace angles*. The method is specifically presented for head model comparisons, but we will also briefly show that the MEG-MUSIC algorithm [1] can be easily explained by these subspace angles.

Methodology

The concept of subspace angles may be explained by some examples. Let us assume first a very simple example of the comparison of two vectors (one-dimensional matrices). Vectors a and b may for instance be the gain vectors for a constrained dipole orientation model, each generated by a different method. The goal is to quantitatively describe how the two vectors are alike or different. For the measure of *residual variance*, we compare the difference between the vectors, scaled by one of the norms, $RV(a, b) = (a - b)^T (a - b) / a^T a$ (alternatively, one may take the square root of this value for the *relative difference measure* RDM). This gives us a measure of how different b is from a in an average sense.

A problem with RV may arise when b is simply a scaled version of a , for instance $b = ca$. Then the RV is $(1 - c)^2$, indicating a significant difference between the two vectors. However, a scalar difference between two vectors may not be significant in some instances. For instance, in the EEG scalp potentials of a single shell spherical model, a scalar difference between two models may simply be due to a difference in assumed conductivities. The shape of the two vectors are otherwise identical, and a localization algorithm would perform identically for the two cases. The difference between the two models would arise in the amplitude of the assumed source, not in its localization. Similarly, one vector may simply be the negative of the other due to sign conventions, in other words with a scale factor of $c = -1$.

We would therefore like a measure that quantifies how similar the two patterns are, i.e., the correlation between the two. The correlation is found as the inner product between a and b , divided by their norms, $CORR(a, b) = (a^T b) / (a^T a)^{1/2} (b^T b)^{1/2}$. A correlation of unity indicates perfect agreement of the two waveforms, within a scale factor. The scale factor may easily be determined by the ratios of the norms, b/a , if so desired. A perfect negative correlation simply indicates a negative scale factor. The correlation and scale factor may be combined to generate the RV,

$$RV(a, b) = \left(1 - \frac{2b}{a} CORR(a, b) + \frac{b^2}{a^2} \right) \quad (1)$$

Thus we might view (1) as a method of partitioning the RV into two constituent components, the correlation and the scale factor. We will make one other adjustment to (1) by separating out the sign of the correlation, $SGN(a, b)$, such that we will use a definition of the absolute correlation ACORR and its sign separately,

$$RV(a, b) = \left(1 - \frac{2SGN(a, b)b}{a} ACORR(a, b) + \frac{b^2}{a^2} \right). \quad (2)$$

The (absolute) correlation between a and b has a geometric interpretation. Consider a and b to be two vectors in a multidimensional space. The correlation between a and b is the cosine of the angle between them, as

Figure 1. The correlation between vectors a and b is the cosine of the angle between them, θ . The absolute correlation between a and b is the cosine of the acute angle formed by the pair of intersecting lines. We can refer to the acute angle (θ or $\pi - \theta$) in a more general sense as the *subspace angle* between a and b .

shown in Fig. 1. If we think of a and b as lines in space (i.e., allowing both positive and negative directions), then by convention we consider the angle between the two lines to be the lesser of $\pi - \theta$ or θ . The absolute correlation is the cosine of this acute angle between these two lines. This acute angle is also known in a more general sense as the subspace angle between a and b .

Now we consider a more general case of two matrices, A and B , each containing two columns and two or more rows. These matrices could represent the gain matrix relating a tangentially oriented dipole to the MEG sensor measurements, with each column representing one of the two tangential dipole components. The two matrices may represent two methods of generating this *forward* model. We would like to compare these two matrices to understand how they are similar, again using correlation analysis. We form the vector $a = Ax$, the linear combination of the columns of A , where $x = x[\cos\alpha, \sin\alpha]^T$ with magnitude x and a unit-norm two-dimensional orientation represented by α , $0 \leq \alpha < \pi$. Similarly, we form $b = By$, $y = y[\cos\beta, \sin\beta]^T$. We then calculate the absolute correlation between these two vectors, $\text{ACORR}(a, b) = \text{ACORR}(A, \alpha, B, \beta)$; considerations of the magnitudes x and y can be disregarded in the correlation, as can α and β greater than π . We generate a grid of values of α over the range 0 to π , and for each α we calculate the absolute correlation using β over the same grid. A coarse grid might yield the results shown in Table 1.

Table 1: Absolute correlations between a and b as a function of α and β .

	$\beta = 0$	36	72	108	144	Row Max
$\alpha = 0$ (degrees)	0.6735	0.1848	0.3722	0.7862	0.9027	0.9027
36	0.9406	0.6802	0.1610	0.4174	0.8388	0.9406
72	0.8484	0.9158	0.6327	0.1108	0.4545	0.9158
108	0.4322	0.8016	0.8628	0.5968	0.1034	0.8628
144	0.1492	0.3812	0.7633	0.8548	0.6218	0.8548
Column Max	0.9406	0.9158	0.8628	0.8548	0.9027	

We extract from this matrix of data the two vectors of data labeled "Column Max" and "Row Max." We are interested in two quantities. The first is the overall maximum correlation. We find this quantity as the maximum of either the Column Max or the Row Max. In the table presented, we see this occurs for $\alpha = 36$, $\beta = 0$ degrees, for a maximum correlation of 0.9406. Thus, based on this coarse table, the best way to correlate matrix A with matrix B is to form $a = A[\cos 36^\circ, \sin 36^\circ]^T$ and $b = B[\cos 0^\circ, \sin 0^\circ]^T$. The second quantity of interest is the minimum of the Column Max or Row Max, which occurs here for $\alpha = 144$, $\beta = 108$ degrees, for a correlation of 0.8548. This *minimax* correlation is the worst way to linearly combine A , since it yields a maximum correlation with B of only 85%, the *minimum* of the *maximum* correlations.

Of course, the grid in Table 1 is too coarse to draw such conclusions, and we should grid more finely. However, a far more elegant method bypasses the need to search over these grids and is easily extended to arbitrary dimensional matrices. The steps are [2] (p. 585),

1) Perform a singular value decomposition (SVD) of A , such that $A = U_A S_A V_A^T$. Similarly decompose B . Retain only those components of U_A and U_B that correspond to non-zero singular values, i.e., the number of columns in U_A and U_B correspond to their ranks, and the other matrices are square, with dimension equal to the ranks.

2) From $C = U_A^T U_B$, and decompose as $C = U_C S_C V_C^T$. Let r be the minimum of the ranks of A and B . Let $s_1, \dots, s_r$ be the r singular values extracted from the diagonal of S_C , $s_1 \geq \dots \geq s_r$.

3) Optionally form $U_a = U_A U_C$, $U_b = U_B V_C$, $Q_a = V_A S_A^{-1} U_C$, and $Q_b = V_B S_B^{-1} V_C$.

The largest singular value s_1 is the maximum absolute correlation between A and B , and the smallest singular value, s_r , is the minimax absolute correlation. The best way to arrange A is found in the first column of Q_a , $a_1 = Aq_{a1}$, which is best correlated with B when it is arranged as $b_1 = Bq_{b1}$. The first columns of U_a and U_b are

a_1 and b_1 normalized to unit-norm. Similarly, the worst way to arrange A is $a_r = Aq_{ar}$, which is best correlated with $b_r = Bq_{br}$ with a correlation of only s_r .

We designate θ_1 as the *principal subspace angle*, such that $\cos\theta_1 = s_1$ and similarly designate the other subspace angles. The matrices U_a and U_b are each orthogonal, i.e., each column is orthogonal to all other columns within the matrix, and each column is unit-norm. The columns of U_a and U_b are known as the *principal vectors*. In the example presented, we find that the maximum correlation is indeed 95.23% and the minimax is 86.1%, corresponding to subspace angles of 18 and 31 degrees. Since our matrices were rank 2, there are only two subspace angles.

Results

In a companion paper [5], we generate the forward EEG or MEG model for a dipole in a sphere, but using different approaches of the boundary element method (BEM). The differences in our approaches focus on the polynomial approximation assumed across each tessellated triangle, the error weighting used to control the fit, and the inclusion of an isolated skull approach. The subspace angle was one of a couple metrics used to quantify the differences between models. In this paper, we present another use of subspace angles, for the comparison of different EEG models. In Table 2, we present the subspace angles for four EEG head models, as compared to a fifth EEG head model.

Table 2: Comparison of Different Spherical Head Models Using Subspace Angles.

Head Model	Radii (fraction of 88 mm)	Conductivities (mho/meter)	Location (mm) (% True Location)	Minimax	
				Correlation (%)	Scalar Ratio
de Munck ([7])	[0.63, 0.88, 0.91, 0.95, 1]	[5, 2, 3, 0.0040, 1]	[0, 0, 54]	Base Model	1
Stok ([6], [7])	[0.84, 0.8667, 0.9467, 1]	[0.33, 1, 0.0042, 0.33]	[0.12, 0.16, 51.4] (95.3%)	100.00	12.7
Rush/Driscoll ([6]) Ary ([7])	[0.87, 0.92, 1]	[0.33, 0.0042, 0.33]	[0.22, 0.25, 48.8] (90.3%)	100.00	13.9
Cuffin & Cohen 1979([6])	[0.8977, 0.9205, 0.9659, 1]	[0.33, 1, 0.0042, 0.33]	[0.3, 0.2, 43.7] (80.9%)	99.96	16.9
Single Shell	1	1	[0.61, 0.44, 31.7] (58.8%)	99.94	6.96

The head model parameters were extracted from [6] and [7], as indicated, and all were applied to a 88 mm outer spherical radius model. The number of shells ranged from 1 to 5, and 528 EEG sensors were spaced about 1 cm apart on the outer hemisphere. The de Munck head model was selected as the base model, and a dipole was placed just 1.4 mm from its innermost shell. The other models were then used in a three-dimensional Nelder-Meade Simplex minimization algorithm, with the goal of minimizing the minimax (third) subspace angle with the de Munck model, by changing the location of the dipole. The results confirm that each single dipole model has a corresponding dipole location in the other head models, and in particular in the single shell model. The minimax moment orientation was in every case z-directed, but the minimax correlation was nonetheless 99.9% or better over all five models. Although the results of Table 2 are presented for a single dipole location, other locations yielded similar results. For example, we used the Nelder-Meade Simplex algorithm to minimize the minimax subspace angle between the Single Shell model and the Cuffin & Cohen model, for 1000 dipoles spread mostly uniformly with depth. This minimization function converged well in all cases to locations 99% or better correlated.

The percentages in column 4 indicate the percentage of the radial distance to the dipole versus the true distance. The locations are almost, but not exactly on the z-axis, which is also the radial line through the true location. If the search were restricted to the radial line through the true location, then these percentages would be equivalent to the "Ary" correction factor. In general, the single dipole correction factor is a function of eccentricity and reveals how different head models may be interchanged by adjusting the dipole's depth between models. Two synchronous dipole models will in general not use the same correction factors as one dipole; see [6], [7] for further discussions on improvements for the multiple dipole cases. In these prior works, empirical results were often generated by fixing not only the base models location, but its orientation as well. Using subspace angles, we do not have to fix the orientation; indeed, we let the subspace angles select the worst orientations.

Discussion

The measure presented is based on the subspace angles between two matrices [2], a multi-dimensional extension of correlation analysis. A singular value decomposition is used to extract the *principal angles* and *principal vectors* between two subspaces, such as the subspaces spanned by two competing head models. If two head models

are mostly similar, then we find that the largest principal angles are still small (the minimax correlations are still high), which can be interpreted to mean that no dipoles exist in one head model that cannot be adequately modeled by another dipole in the other head model, irrespective of dipole orientation. By not performing Step 3, we can rapidly calculate just the singular values in Step 2 for literally thousands of comparison points between two models. By examining the minimax correlations, we can identify those comparison points where the two models yield the poorest comparisons, then investigate just these points further to understand the differences between the two models.

"Small" and "adequate" must be interpreted relative to a given noise level. Our experiences in dipole localization errors [3] [4] indicate that a minimax correlation angle of less than 18 degrees (correlation above 0.95) should indicate that two models are mostly indistinguishable, given the noise conditions we see in EEG and MEG. These principal angles emphasize the *pattern* of the measurements each model can make at the sensors, regardless of the actual dipole intensity or orientation. If model *A* requires a dipole ten times stronger and at a moment angle 45 degrees rotated from that of model *B*, then we can still obtain a principal angle of zero degrees (perfect correlation) if the *shape* of the measurements is identical between the two models. In other words, correlation is scale independent, and using subspace angles, we achieve a performance measure independent of dipole orientation and intensity.

In inverse modeling, with the uncertainties of the true orientation of the equivalent current dipole, and with the approximation of the gains by uncertain conductivities, the emphasis in model comparison is properly placed on this shape, which leads directly to localization. We can, however, also easily extract orientation and gain differences from this same procedure, as shown in Step 3. The principal vectors of each subspace can be directly interpreted to indicate orientation differences between head models, and the gain difference between two models is simply found as a power ratio.

The subspace angle is also useful in the inverse problem. The MEG-MUSIC algorithm [1] simply generates the subspace angle between the signal subspace extracted from the data and the subspace spanned by the desired model. As discussed in [1], if each dipole has a time series linearly independent of the other dipolar time series, then we can decompose the spatio-temporal data matrix to extract a signal subspace. Each dipole's spatial model will be correlated with this signal subspace. We therefore form the gain matrix (forward model) for a dipole and calculate the subspace angles between the gain matrix and the signal subspace. In MEG with a spherical head model, the gain matrix is rank 2, so there are at most only two subspace angles. A MUSIC metric of 1.0 is the same as a minimum subspace angle of zero degrees, indicating that a dipole of some fixed orientation (found from Step 3) matches the data. If the second angle, corresponding to the minimax correlation, is also near 1.0, then we have indications of a "rotating" dipole, as discussed in [1].

We therefore scan over a grid of candidate dipole locations, looking for correlations near unity with the signal subspace. Peaks near unity can be refined by further gridding, or the subspace angle metric may simply be used in a directed search, such as a Nelder-Mead Simplex algorithm. If the rank of the signal subspace is underestimated, we may destroy the correlation of the dipolar model with the data. On the other hand, if we somewhat overestimate the rank, then it is unlikely that we will falsely find a correlation between a dipole model and the extraneous subspace components, since those components are drawn from a randomly oriented noise subspace. Thus we generally recommend that the signal subspace be somewhat overestimated.

Although in [1] the source model was limited to dipoles, the MUSIC method is general for any desired parametric source model, such as pairs of dipoles or combined dipole and quadrupoles. We simply generate the subspace angles between the signal subspace and the source model.

References

- [1] Mosher JC, Lewis PS, and Leahy RM, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans. Biomedical Eng.*, 1992, 39:541 – 557.
- [2] Golub GH, Van Loan CF, *Matrix Computations*, second edition, Johns Hopkins University Press, 1984.
- [3] Mosher JC, Lewis PS, Leahy RM, *Biomagnetism: Fundamental Research and Clinical Applications*, Proceedings of the Biomag '93 Vienna Conference, Elsevier, 1995, 330-334.
- [4] Mosher JC, Spencer ME, Leahy RM, Lewis PS, *Electroenceph. and clin. Neurophys.* 1993, 86:303 – 320.
- [5] Mosher JC, Chang, CH, Leahy RM, "Comparison of the constant and linear boundary element method for EEG and MEG forward modeling" this volume.
- [6] Berg P, Scherg M, "A fast method for forward computation of multiple-shell spherical head models," *EEG Journal*, 90 (1994) 58-64.
- [7] Zhang Z, "A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres," *Phys. Med. Biol.*, (1995) 335-349.

Acknowledgment

This work was supported by the National Institute of Mental Health grant no. R01-MH53213 and the National Eye Institute grant no. R01-EY08610-04.