International symposium on circuits and systems Kyoto Japan June 5-7 1985.

85-89C

SOURCE SEPARATION BY ADAPTIVE SINGULAR VALUE DECOMPOSITION

J. Vandewalle+, J. Vanderschoot++ and B. De Moor+ *

+ESAT Laboratory, Elec. Eng. Dept., ++Laboratory of G.I. Motility
Katholieke Universiteit Leuven

SUMMARY

When many signals composed of many sources (signals of different strength and noise) can be measured simultaneously a rather classical but unusual approach to the separation of source signals is to separate these according to directions of extremal energy. It has been applied to the extraction of the fetal ECG out of many cutaneous measurements containing maternal, fetal ECG and noise. A numerically reliable implementation requires singular value decomposition, which is quite time consuming. In order to reduce the computation time adaptive or partial, singular value decomposition algorithms are proposed based on the power method and Chebyshev iteration.

1. INTRODUCION

If a set of measured signals contains contributions from sources with different strengths, they can often be separated by projecting onto subspaces of maximal and minimal energy. This technique has been applied successfully to filter or separate the fetal electrocardiogram (ECG) from the maternal ECG in cutaneous measurements, and to filter out 50 Hz contributions, and, to eliminate noise in other problems in system theory. A numerically reliable algorithm to perform this separation requires the singular value decomposition SVD of the matrix of the measured signals. Since the classical SVD algorithm is too complex for real time processing, a new and faster adaptive algorithm is developped.

In the paper we first define the oriented energy of a vector signal and derive a number of its properties [1]. Then the source separation technique is derived and applied to the separation of the fetal ECG from cutaneous measurements [2]. A new adaptive algorithm based on the Chebyshev transformation is then presented which is much faster than the classical algorithm of Golub [3] and the adaptive SVD algorithm ASVD [4].

2. ORIENTED ENERGY OF A VECTOR SIGNAL

Suppose that A_{pxq} is a matrix, representing a set of q samples a_i from a p dimensional vector signal :

$$A = [a_1 \; a_2 \ldots \; a_q], \qquad (1)$$

---

* Supported by the Belgian I.W.O.N.L.

Then the energy of this set in the direction of a unit vector e can be defined as :

$$E\,[A] = \sum_{i=1}^{q} (e^t a_i)^2 = e^T AA\; e \qquad (2)$$

As is well known from classical matrix algebra, the symmetric matrix AA^T has p orthonormal eigenvectors $u_1, \ldots u_p$, and p non negative eigenvalues $\lambda_1, \ldots \lambda_p$. It is easy to show that

$$E_{u_i}[A] = \lambda_i \qquad (3)$$

are extremal values of the energy. Fig. 1 shows as an example, the polar plot of the oriented energy of a set of vector samples from a three dimensional signal space. In the next section it will be shown that these extremal energy directions are essential to the proposed source separation scheme. The foregoing discussion might

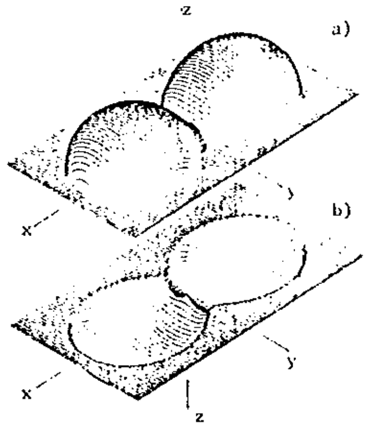

Fig.1. Polar plot of the oriented energy of a vector signal (p=3)

suggest to compute these directions by forming the covariance matrix AA^T and calculating the eigenvalue decomposition. However since the original matrix A is given, a numerically better approach [3] is to compute the singular value decomposition of A, [3] :

$$A_{pxq} = U_{pxp}\, \Sigma_{pxq}\, V^T_{qxq} \qquad p \leq q \qquad (4)$$

R7AC.3

in which U and V are orthonormal matrices, and Σ is a real pseudodiagonal nonnegative definite matrix. Column vectors of U, are called left singular basis vectors, and the diagonal elements of Σ are called singular values. Singular values and singular basis vectors are conventionally ordered so that

$$\sigma_1 \geq \sigma_2 \geq \dots \sigma_p \geq 0 \qquad (5)$$

It is easy to verify that the left singular basis vectors and singular values of A are equal to resp. the eigenvectors and the positive square roots of the eigenvalues of AA^T.

3. SOURCE SEPARATION BASED ON EXTREMAL ENERGY DIRECTIONS

The separation problem that will be considered can be stated as follows. Given a set of p measurement signals being static linear combinations of a set of n unknown source signals, and corrupted by additive uncorrelated noise. The linear combination or transfer is only known qualitatively. How and under what conditions can the source signals be reconstruc-ted ? In matrix form one

$$M = TS + N, \qquad (6)$$

with a measurement signal matrix M, a noise signal matrix N, a source signal matrix S and a transfer matrix T. The principal idea of the method will be demonstrated on the basis of Fig. 2. Fig. 2.a. represents a set of vector samples in a two dimensional source signal space, and a polar plot of the oriented energy in this space. The transfer matrix maps the source signals into a three dimensional measurement space, see Fig. 2.b. The basis vectors of the source space are mapped into t_1 and t_2 respectively, t_1 and t_2 being column vectors of T. If there is no noise present, all measurement samples are located in the plane (t_1, t_2). Fig. 2.b. also shows the

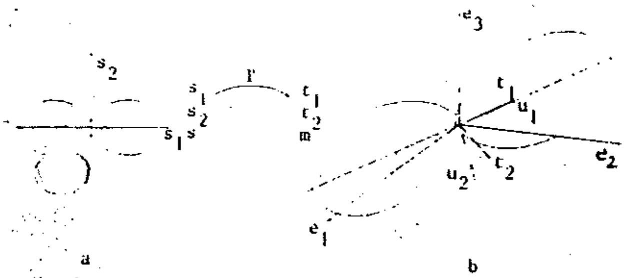

Fig.2. The transfer T of source signals s_1, s_2 to measurement signals.

first two left singular basis vectors u_1 and u_2 of the measurement matrix. The plane (u_1, u_2) will obviously coincide with (t_1, t_2). It can be shown that u_1 will be very close to t_1, if the energy of the first source signal is slightly larger than the energy of the second source signal. But if this is the case, u_2 will be almost orthogonal to t_1. If measurement samples are projected onto u_2, then this projection will be almost free of first source signal contributions. This is how a weak signal can be extracted in the presence of a stronger, disturbing signal. E.g. in the case of fetal ECG recordings, where a much stronger maternal ECG is present. But furthermore if t_1 can be made orthogonal to t_2 an exact separation is possible, since then $u_1 = t_1$ and $u_2 = t_2$. The next theorem is the basis for the signal separation algorithm in a general case.

Theorem : Given a measurement matrix M satisfying (6) with :
p = the number of measurement signals,
q = the number of vector samples,
n = the number of source signals,
$p < n < q$,
T, S, N matrices representing respectively pxn transfer, nxq source signals, pxq noise signals.
Let the singular value decomposition of M be given by :

$$M_{pxq} = U_{pxp} \Sigma_{pxp} V^T_{pxq} \qquad (7)$$

If (1) $SS^T = \text{diag}\,[\sigma^2_{s_1}, \sigma^2_{s_2}, \dots, \sigma^2_{s_n}]$

(2) $\sigma_{s_1} > \sigma_{s_2} \dots \sigma_{s_n} > 0$

(3) $T^T T = I_{nxn}$ (8)

(4) $NN^T = \sigma^2_N I$

(5) $N S^T = 0_{pxn}$

Then
(1) $\Sigma \cdot \Sigma^t = \text{diag}(\sigma^2_{s_1} + \sigma^2_N, \dots, \sigma^2_{s_n} + \sigma^2_N, \sigma^2_N, \dots, \sigma^2_N)$

(2) $U = T$ (9)

(3) $S = U^T M = \Sigma V^T$

The proof of this theorem is based on the uniqueness of the eigenvalue decomposition of the matrix MM^T, and will be omitted here. Let us instead translate the conditions (1) tot (5) to practical requirements.

(1) Source signals should not be correlated with each other for zero time shift. It is not obvious that this condition can be satisfied in all practical situations. However, if one is only interested in some source signals, it can be proven that only these source signals have to be uncorrelated mutually and with the other source signals.
(2) The source signals should have a different energy.
(3) The columns of the transfer matrix should constitute a set of n orthonormal basis vectors. In the biomedical problem this orthogonality can be influenced by appropriately choosing electrode positions.

(4) The noise signals should be uncorrelated with each other, and they should have equal energy.
(5) Noise signals should not be correlated with source signals.

Another important condition is that q should be sufficiently large so that SS^T, NN^T, NS^T are crosscorrelation estimators with a small variance. Let us now present the signal separation algorithm, and show the practical application for the problem of the fetal ECG extraction.

Signal separation algorithm

1. Compute the SVD of a measurement matrix M, containing q vector samples.
2. Compute the projections onto the extremal energy directions $u_1, \dots u_n$. These projections are estimates of the corresponding source signals :

$$s_i = u_i^T M \approx s_i \qquad 1 \leq i \leq n \qquad (10)$$

re s_i is the i-th row of S.

3. the transfer is time invariant, these projections are still good estimates of the corresponding source signal for further time instants.

A result of the application of this algorithm to the extraction of a fetal ECG from skin electrodes is shown in Fig. 3. Fig. 3.a. shows the 6 measurement signals forming the matrix M_{6x1000}.

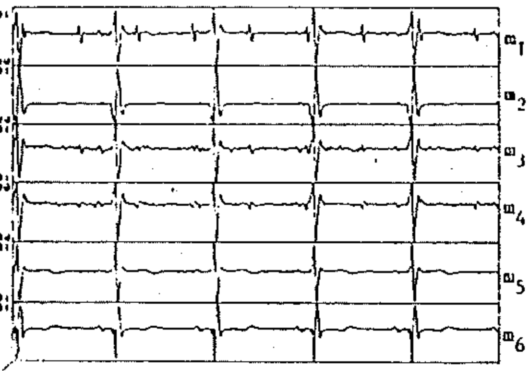

Fig.3a.Cutaneous recordings over an interval of 4 s., sampling rate 250 Hz., τ_0=10 ms., f_0=70 Hz, 37-th week of pregnancy.

Fig. 3.b. shows the 6 projections onto the extremal energy directions of M. It is clear from this figure that the separation is quite good. The dimension of the subspace generated by the maternal ECG is three (consistent with the so called dipole hypothesis of cardiology). The 4-th projection contains almost only FECG contributions. The lower energy directions 5 and 6 only show noise. The third step of the algorithm is illustrated in Fig. 3.c. These are projections onto u_4 of the 1000 measurement vector samples subsequent to those of Fig. 3.a. This step only requires 6 multiply and add operations per vector sample. As one can see, computations could be reduced if not all extremal energy directions were calculated. This means that a partial SVD, resulting in u_4 in this case, would suffice. If a small time variance of the transfer matrix is expected, then an adaptive SVD algorithm is of considerable value.

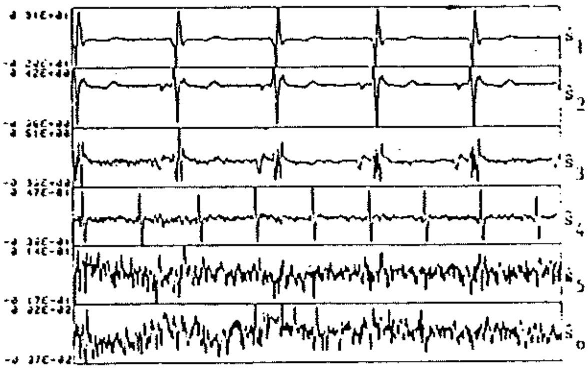

Fig.3b. Projections of the measurements of fig.3a onto $u_1,\dots,u_6$ resulting from the SVD of those 1000 measurements samples.

Fig.3c.Projection of the 1000 measurements samples subsequent to those of fig.1a.(not shown) onto the same u_4 as for fig.1b.

THE COMPUTATION OF THE PARTIAL OR ADAPTIVE SINGULAR VALUE DECOMPOSITION

In this section we discuss some numerically reliable algorithms for the SVD of real matrices. The choice of a suitable algorithm for the computation of the SVD depends upon the conditions that are imposed. If one needs a full decomposition (i.e. all singular triplets) of a matrix of moderate dimensions, Golub's algorithm [3] is very efficient and numerically reliable. Fully tested and documentated software is available. However, there exist many applications in which the matrices involved satisfy specific conditions : They may be large and structured, such that the entries preferably remain unchanged during computations. When they are of low rank only a partial SVD is necessary. (e.g. model reduction techniques). A supplementary aspect may be that the matrices may be slowly time-varying. Hence, it is natural to develop an adaptive algorithm that uses the SVD computations of a previous time step as an initial value for updating the actual SVD. Both observations hold for many signal separation problems such as FECG. Two recently developped algorithms are now briefly considered.

A. ASVD : the power method for SVD

This method, developped independently by Staar [1] and Shlien [6], and worked out in [5] can be summarized as follows : Suppose that k first triplets of an mxn matrix A are already computed

$$U_k = [u_1 \ \dots \ u_k] \quad (11.a)$$

$$V_k = [v_1 \ \dots \ v_k] \quad (11.b)$$

$$\Sigma_k = \text{diag}\ [\sigma_1 \ \dots \ \sigma_k] \quad (11.c)$$

then the k+1-th triplet is computed via an iterative process.

1. Initial guess $u_{k+1}^{(o)} = u_k$

iteration i

2. $v_{k+1}^{(2i+1)} = A^t . u_{k+1}^{2i} / \| A^t u_{k+1}^{2i} \|$ (12)

3. $\sigma_{k+1}^{2i+1} = \| A^t . u_{k+1}^{2i} \|$ (13)

4. $u_{k+1}^{2i+1} = A . v_{k+1}^{2i+1} / \| A v_{k+1}^{2i+1} \|$ (14)

5. $\sigma_{k+1}^{2i+2} = \| A v_{k+1}^{2i+1} \|$ (15)

6. the test for convergence

$$\| u_{k+1}^{2i+2} - u_{k+1}^{2i} \| < \varepsilon \quad (16)$$

is satisfied start the recursion for triplet k+2 (deflate), if not, go to 2.

In [5] the number of multiplications for each triplet, convergence rates, reductions in the orthogonalization, initial guess strategies, deflation techniques and acceleration algorithms are studied.

B. Chebyshev iteration

This is a very efficient technique based on the recursion of the Chebyshev polynomials. From (4) one obtains $A.A^t = U.\Sigma.\Sigma^t.U^t$ (17)

One can prove that for every polynomial or analytic function f of a matrix AA^t

$$f(AA^t) = V\ f(\Sigma.\Sigma^t)\ V^t \qquad f(\Sigma\Sigma^t) = \text{diag}[f\ (\sigma_i^2)]$$

Now the Chebyshev polynomials $T_n(x)$ satisfy

$$T_0(x) = 1 \qquad T_1(x) = x \quad (20.a,b)$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (20.c)$$

$$|T_n(x)| > |T_{n-1}(x)| > 1 \text{ if } x \in (-\infty,-1) \cup (1,\infty) \quad (20.d)$$

$$|T_n(x)| \leq 1 \quad \text{if} \quad x \in [-1,1] \quad (20.e)$$

If we now consider the following iterative procedure, with v_n the n-th iteration vector

1. Initialize v_0

$$v_1 = f(AA^t)\ v_0 \quad (21)$$

2. Iterate $v_{n+1} = 2f(AA^t)\ v_n - v_{n-1}$ (22)

3. Test for convergence

$$\| v_{n+1} - v_n \| < \varepsilon \quad (23)$$

if not, go to 2.

The result of this iterative procedure is a vector with components along those left singular vectors of A of which the corresponding singular value of f (AA^t) is larger than 1. The largest, smallest or intermediary singular values are computed by choosing suitable f. In [7], the number of iterations is derived. It can be proven that this iteration is generically considerably faster than the ordinary power method and Golub's algorithm [3,4]. Reliable refinements, deflation techniques and combination with previously developped acceleration techniques are actually under study.

CONCLUSION

A signal filtering technique has been described which is not based on the frequency content but rather on the strength of contribution to multiple signals. It has proven to be useful in eliminating maternal ECG in abdominal recordings of the fetal ECG. An algorithm for adaptive singular value decomposition is presented which can compute efficiently a part of the SVD on-line.

REFERENCES

[1] J. Staar, "Concepts for reliable modeling of linear systems with application to on-line identification of multivariable state space description", Doct. Th., Kath. Univ. Leuven, June 1982.

[2] J. Vanderschoot, J. Vandewalle, J. Janssens, W. Sansen and G. Vantrappen, "Extraction of weak bioelectrical signals by means of singular value decomposition", "Analysis and Optimization of Systems", Proc. of Sixth Int. Conf. on An. and Opt. of Systems, Nice, June 1984, Springer Verlag, Berlin, pp. 434

[3] G.H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions", Numer. Math., Vol. 14, 1970, pp. 403-420.

[4] J. Dongarra et al. "Linpack User's Guide", Philadelphia, SIAM 1979.

[5] J. Vandewalle, J. Staar, B. De Moor and J. Lauwers, "An adaptive singular value decomposition algorithm and its application to adaptive realization", "Analysis and Optimization of systems", Proc. of Sixth Int. Conf. on An. and Opt. of Systems, Nice, June 1984, Springer Verlag, Berlin, pp. 33-47.

[6] S. Shlien, "A method for computing the partial SVD", IEEE Trans. on pattern analysis and mach. intell., Vol. PAMI-14, no. 6, Nov. 1982.

[7] B. De Moor and J. Vandewalle, "An adaptive singular value decomposition algorithm based on generalized Chebyshev recursions", Proc. Conf. on Math. in Signal Processing, Univ. of Bath, Essex, England, Sept. 85.