Thomas Schanze*

Removing noise in biomedical signal recordings by singular value decomposition

Abstract: Noise reduction or denoising is the process of removing noise from a signal. If some signal properties are known linear filtering is often useful. Fourier, wavelet and similar transform approaches remove unwanted signal components in the codomain. For this, predefined eigenfunctions, e.g. wavelets, are used. Here we use singular value decomposition in order to compute a signal driven representation (eigendecompositon). By removing unwanted components of the representation the signal can be denoised. We introduce the new method, apply it to signals and discuss its properties.

1 Introduction

A still challenge in signal processing is to develop methods to reduce noise in signal recordings. If some properties of the signal under investigation are known, then simple linear filters can be useful and effective. A classical example is the removing of alternating current hum in ECG or EEG recordings by means of a notch filter. Another approach transforms the recorded signal in another domain, e.g. from time to frequency domain, where components related to noise, which comprises artefacts, interferences etc., can be easily removed in such a way, that the back transformation yields a purified signal. This type of denoising can be achieved by Fourier, wavelet or similar transforms [1, 2]. The basic procedure is as follows: transform the signal, attenuate or remove components of the transformed signal which are not related to the useful signal, and transform back.

Analysis, processing and representation of signals can be done by predefined basis vectors, e.g. sine waves, or by basis vectors constructed out of the investigated data. Well-known examples are Fourier and principal component analysis, where scalar products of signals with basis vectors are used.

Singular value decomposition (SVD) is a mathematical procedure to decompose a matrix in a product of three matrices, which can be rewritten as a sum of rank one matrices [3]. In addition, SVD can be regarded as a generalization of the eigendecomposition of positive semi-definite normal matrix. Thus it is related to principal components analysis (PCA), which is quite often used to analyse and represent data.

A digital image is an array of pixels. Such an array can be represented by a matrix, where the value of a matrix element codes information of the related pixel. SVD provides an interesting way for breaking a matrix, which may contain some useful data we are interested in, into simpler, meaningful pieces. Since a digital image can be represented by a matrix, the application of SVD in image processing is straight forward [4, 5], whereas the application of SVD in signal processing is not. Here we show how to use SVD in order to remove noise in biomedical signals and how to construct signal and noise subspaces for signal analyses and approximations.

2 Method

2.1 SVD

We give a short review of singular value decomposition, which is an important tool in modern linear algebra. The SVD was established for real square matrices in the 1870's by Beltrami and Jordan [6, p. 78]. Steward [7] provides the early history of SVD.

Suppose $\mathbf{A}$ is a real $m \times n$ matrix. The singular value decomposition (SVD) of $\mathbf{A}$ is a decomposition of the form

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}}, \tag{1}$$

where $\mathbf{U}$ is an orthogonal $m \times m$ matrix, $\mathbf{S}$ is a diagonal $m \times n$ matrix with non-negative entries, i.e. $s_{ii} \geq 0$, and $\mathbf{V}$ is an orthogonal $n \times n$ matrix. The columns of $\mathbf{U}$ and $\mathbf{V}$ are

_____
*Corresponding author: Thomas Schanze: Technische Hochschule Mittelhessen, Dept. Life Science Engineering, Wiesenstr. 14, D-35390 Gießen, Germany, e-mail: thomas.schanze@lse.thm.de

called left and right singular vectors, the diagonal elements s_{ii} of **S** are the singular values [3–8].

Equation (1) can also be written as a sum of rank one matrices:

$$\mathbf{A} = \sum_i s_{p,ii}\mathbf{u}_{p,i}\mathbf{v}_{p,i}^{\mathrm{T}}, \tag{2}$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the i-th columns of **U** and **V**. Note that $\mathbf{u}_i\mathbf{v}_i^{\mathrm{T}}$ is the outer product of $\mathbf{u}_i$ and $\mathbf{v}_i$, where $r \leq n$ is the number of positive singular values used. It was assumed here that $s_{ii} > 0$ for $i = 1, 2, \cdots, r$.

The diagonal values $s_{ii} \geq 0$ are said to make up the singular value spectrum, which can be easily visualized by a one-dimensional plot. The importance of a singular value is given by its magnitude. More specifically, the square of each singular value is proportional to the variance explained by each singular vector. This property is of importance when using SVD for denoising.

2.2 SVD and denoising

In image processing the SVD is often used to reduce noise. This can be easily obtained by attenuating or zeroing suitable small singular values.

Digital images can be represented by matrices [4, 5]. Let **A** be an $m \times n$ matrix representing a grey image or single colour channel of a colour image. Assume that S has $r = \min(m,n)$ positive singular values. If so, we find $r = \mathrm{rank}(\mathbf{A})$. Thus **A** can be expressed by r vectors $\mathbf{u}_i$, each of dimension m, r singular values s_{ii}, and r vectors $\mathbf{v}_i$, each of dimension n.

Rearranging the singular values of **A** in descending order, i.e. $s_{ii} \geq s_{kk}$ for all $k > i$, and selecting the first $p \leq r$ singular values and related left and right singular vectors leads to

$$\mathbf{A}_p = \mathbf{U}_p\mathbf{S}_p\mathbf{V}_p^{\mathrm{T}} = \sum_{i=1}^{p} s_{p,ii}\mathbf{u}_{p,i}\mathbf{v}_{p,i}^{\mathrm{T}}. \tag{3}$$

The rank of this matrix is $\mathrm{rank}(\mathbf{A}_p) = p \leq r$. If $p < r$ then $\mathbf{A}_p$ is a low-rank approximation of **A**. Note that this ordering is useful, since large singular values are more important for the signal representation than small singular values.

If we assume that small singular values are related to noise, we can use this assumption to reduce noise. Suppose that the first singular values are above a pre-defined threshold, i.e. $s_{ii} \geq c$ for $1 \leq i \leq p$, thus $s_{ii} < c$ for $p < i \leq r$. This means that we can separate the matrix **A** in two matrices: the matrix of $\mathbf{A}_p$ is related to useful or de-noised image data, and, consequently, $\mathbf{A}_{r-p} = \mathbf{A} - \mathbf{A}_p$ is related to noise. Thus we can separate an image in useful data and
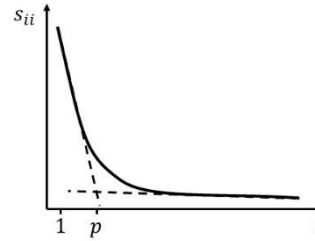
Figure 1: Scree plot of singular values arranged in descending order to estimate the number p required to compute $\mathbf{A}_p$.

noise subspaces. It is obvious, that this approach can be easily extended to the soft-thresholding method introduced by Donoho [9] and to single or multiple rank analyses and representations.

2.3 Bijective data mapping and components selection

If we want to use SVD in signal processing, e.g. to process a one-channel ECG recording, we have to map the data. Assume that we have a single channel time discrete recording $x = (x_1 \quad \cdots \quad x_t)^{\mathrm{T}}$. If $m \cdot n = t$ holds, we map **x** to **A** by using a bijective transformation, i.e. $\mathbf{A} = \mathrm{Tr}\{\mathbf{x}\}$, and compute $\mathbf{A}_p = \mathbf{U}_p\mathbf{S}_p\mathbf{V}_p^{\mathrm{T}}$. The reconstruction or approximation can be obtained by the inverse transformation: $\mathbf{x} = \mathrm{Tr}^{-1}\{\mathbf{A}_p\}$.

A straight forward idea is to map the indices. For the forward transformation we have

$$(\mathbf{A})_{ij} = x_k : \begin{cases} i = 1 + (k-1)\bmod m \\ j = 1 + (k-1)\operatorname{div} m \end{cases} \tag{4}$$

with modulo and integer division. For the reverse or backward transformation we find

$$x_{p,k} = (\mathbf{A}_p)_{ij} : k = i + m \cdot (j-1). \tag{5}$$

The problems remaining are to find a proper mapping, i.e. finding values m and n that obey $m \cdot n = t$, and to find the number p in order to compute the denoised approximation $\mathbf{A}_p$ (eq. 3) or equivalently, $x_{p,k}$ (eq. 5). One solution is to try out several factorisations and different p values. It is clear, that for this a criterion must be defined. A similar problem occurs in PCA [10]: How to select the number of the PCs? A well-known solution bases on the scree plot developed by Cattell [11], which we adopt as follows: arrange the singular values in descending order and plot s_{ii} against i, look at which value of i the slopes of two lines approximating the data are steep for small i and not steep for large i, and estimate the value of i which is related to the point of intersection of the lines ("elbow"). This value

can be used to estimate the number of singular values and vectors, i.e. components, in order to obtain a denoised signal approximation or representation $\mathbf{A}_p$ or rather $x_{p,k}$. **Figure 1** sketches this approach. It is obvious, that the smaller p, the more compact is the (denoised) signal representation. We note that the achieved signal decomposition is an eigen-decomposition.

Noisy signal

Denoised signal
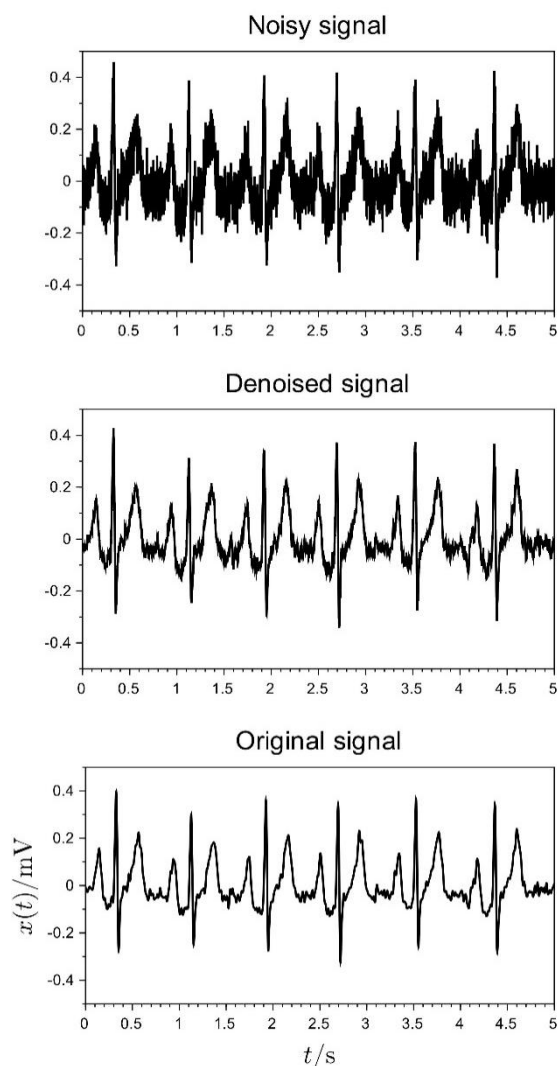
Original signal

Figure 2: Application of the SVD-based denoising approach to an ECG recording. The upper time-course is the Gaussian white noise contaminated original signal, the denoised signal is shown in the middle, the time-course at the bottom is the original signal. Note that the shown signals are segments only: 5 s out of 20 s. The ECG was amplified and sampled with 1,000 samples/sec and bandpass-filtered (0.5-38.5 Hz).
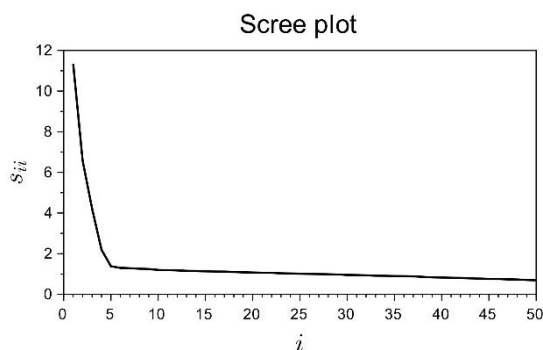
Scree plot

Figure 3: Scree plot of the singular values computed for the noisy data partly shown in Figure 2. The plot indicates that the four largest singular values and the related singular vectors should be used for denoising.

3 Results

We applied the method, implemented under Scilab 5.5.2/6.0, to simulated data, e.g. sinusoidal wave with superimposed noise, and found that our approach prospects reasonable denoising properties. Of higher interest is the application of the method to biomedical signals. Five recordings of ECG and EEG were analysed. **Figure 2** shows a signal segment of a lead II ECG recorded from the author. The signal was acquired with a Biopac MP 36-system and standard ECG-electrodes. The bottom picture shows the time-course of a segment of the original recording. The picture at the top shows the Gaussian white noise contaminated signal. The time-course in the middle shows the denoised signal. Note that the introduced SVD-denoising method preserves the morphologies of P-waves, QRS-complexes and T-waves, despite distinctive noise removal. For SVD-denoising 20,000 data points were segmented into 400 vectors each of length 50, thus we have $m = 50$, $n = 400$ and $t = 20,000$. The estimation of the number of singular values and related singular vectors, i.e. p, to compute $\mathbf{A}_p$ was done with a scree plot of the singular values. **Figure 3** shows the scree plot. The selected value is $p = 4$. Thus the original signal is recovered and represented by four vectors of length $m = 50$, four singular values, and four vectors of length $n = 400$. Therefore the total denoised signal is represented by 1,804 data values, which is significantly smaller than the 20,000 data values of the original signal. This shows that denoising and efficient representation of biomedical signals can be achieved simultaneously with the presented method. Similar results were obtained for other ECG recordings and also for EEG recordings.

4 Discussion

We introduced a method to denoise a time-signal by using the SVD. For this the signal was transformed into matrix by a bijective mapping. The obtained decomposition can be represented by a sum of rank one matrices each weighted with the respective singular value. The selection of these matrices uses the scree plot-method. The signal obtained by backtransformation is a denoised approximation of the wanted signal.

SVD is a standard tool in modern linear algebra that enables a matrix factorisation [3-7]. This tool produces singular values and related left and right singular vectors (eigenvectors) that can be selected to approximate the signal. The presented method is a signal driven denoising approach, because it performs an eigendecomposition of a signal bijectively mapped into a matrix. In other words, the method "finds" the subspace related to the wanted signal and the subspace related to noise.

The method was tested with simulated data and biomedical signals. The results show that the method enables denoising in combination with efficient data representation. This is useful in biomedical signal processing applications, when large amounts of data must be stored.

A still weak point of the method is the selection of the singular values and the related singular vectors to establish a "compressive denoising" method. Another point that must be addressed in future work is the bijective transformation of the signal representing vector into a matrix and vice versa in order to apply SVD, i.e. the conditions that will lead to an efficient denoising and efficient or sparse representation must be figured out.

One-dimensional vector methods, e.g. Fourier or wavelet transformation, are well studied and a common tool in signal processing [1, 2]. SVD is a two-dimensional method which is widespread and well-studied mathematically [3-7]. Considering our approach, we note that we applied a two-dimensional method to remapped one-dimensional data. This approach might be extended by using higher order singular value decompositions [12].

To conclude, our results indicate that the presented approach might be useful in biomedical signal processing.

References

[1] Graps A. An introduction to wavelets. IEEE Computational Science and Engineering 1995; 2:50–61.

[2] Chirona L, van Agthoven MA, Kieffera B, Rolando C, Delsuc M-A. Efficient denoising algorithms for large experimental datasets and their applications in Fourier transform ion cyclotron resonance mass spectrometry. PNAS 2014; 111:1385–1390.

[3] Golub GH, Reinsch C. Singular value decomposition and least squares solutions. Numer Math 1970; 14:403-420.

[4] Andrews, HC, Patterson CL. Singular value decomposition (SVD) image coding. IEEE Transactions of Communications 1976; 24:425–431.

[5] Jain, AK. Advances in mathematical models for image processing. Proceedings of the IEEE 1981; 69:502–528.

[6] Mac Duffee, CC. The theory of matrices. Berlin: Verlag von Julius Springer 1933.

[7] Steward GW. The early history of the singular values decomposition. Siam Review 1993; 35: 551–566.

[8] Sadek RA. SVD based image processing applications: state of the art, contributions and research challenges. International Journal of Advanced Computer Science and Applications 2012; 3:26–34.

[9] Donoho DL. De-noising by soft-thresholding. IEEE Transactions on Information Theory 1995; 41:613-627.

[10] Jolliffe IT. Principal component analysis. 2nd ed. New York: Springer-Verlag 2002.

[11] Cattell RB. The scree test for the number of factors. Multivariate Behavioral Research 1966; 1:245–276.

[12] De Lathauwer L, De Moor B,Vandewalle J. A multilinear singular value decomposition. SIAM J Matrix Anal Applicat 2000; 21:1253–1278.