

Optical Imaging Analysis for Neural Signal Processing: A Tutorial

Andrew T. Sornborger, PhD

Department of Mathematics and Faculty of Engineering, University of Georgia
Athens, Georgia

Tutorial: Denoising and Frequency Localization of an Imaging Time Series

In this tutorial, we will cover the basic steps to follow when using a singular value decomposition (SVD) and Thomson F-test in order to reduce the amount of noise in a neural imaging data set (Mitra and Bokil, 2008). Figures 1 through 5 depict the steps involved.

It should be remembered that, in this type of analysis, an image is considered to be primarily a vector of values (Fig. 1, top). That is, the pixels that were originally arranged in two dimensions are arranged as elements of a vector, and the two-dimensionality is, for the moment, not depicted. A movie of imaging data taken over time is therefore described as a data matrix in which each row is an image vector, and time progresses row by row from the initial image to the final image (Fig. 1, bottom).

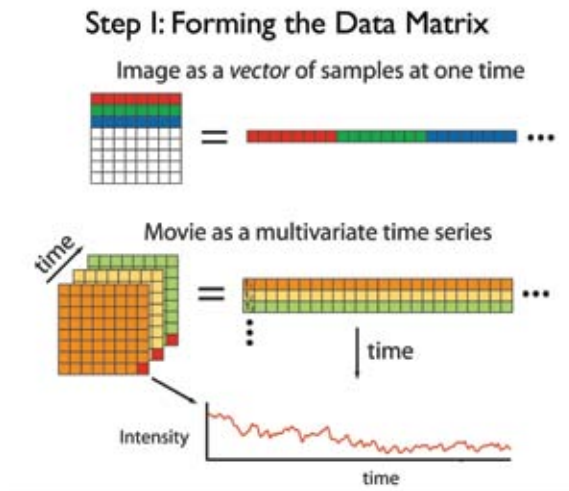


Figure 1. The first step in the analysis of optical imaging data is to list the pixels in the form of a vector (top panel). Then the image vectors are arranged sequentially in time, thereby forming a data matrix.

Therefore, if you are working with imaging data saved in a MATLAB.mat file, for example, in the variable DAT, where DAT has elements DAT (i, j, k) (with $i = 1 \dots nt$, $j = 1 \dots nx$ and $k = 1 \dots ny$), then with the command

```
X = reshape(DAT, [nt nx*ny]);
```

you can make a data matrix, X , out of your imaging data. This data matrix has nt rows and $nx \times ny$ columns.

When performing an analysis for the first time, it is often useful (and prudent) to plot the data matrix. To do this, first subtract the mean of each pixel time course (each column) from each time course and make a pseudo-color plot of the data. Much of the structure of the data (including possible measurement artifacts) may be seen in such a plot. Often, more of the dynamics may be seen in this way than by viewing a movie of the raw imaging data. We can do this for our data matrix, X , using the following commands:

```
% subtract mean pixel value from each pixel
X = X - repmat(mean(X, 1), [nt 1]);
% make pseudo-color plot of matrix
pcolor(X); shading flat;
```

Raw imaging data are typically noisy. Sources of noise can be either shot noise from fluctuations in the number of photons arriving at the detector or dark noise arising from thermal fluctuations in the measurement apparatus. Patterned measurement artifacts from the imaging device and changes in the physiology of the system being imaged can often be significant sources of noise as well.

The judicious use of an SVD can help reduce such noise in imaging data sets. An SVD may be thought of in many ways, one useful conceptualization being that the SVD takes the raw data matrix (Fig. 2, top row) and decomposes it into a sum of many submatrices (Fig. 2, bottom row). Each of the submatrices is the combination of an eigenvector and a time course. In turn, each row of the submatrix is proportional to the pixel values of the eigenvector, and each column

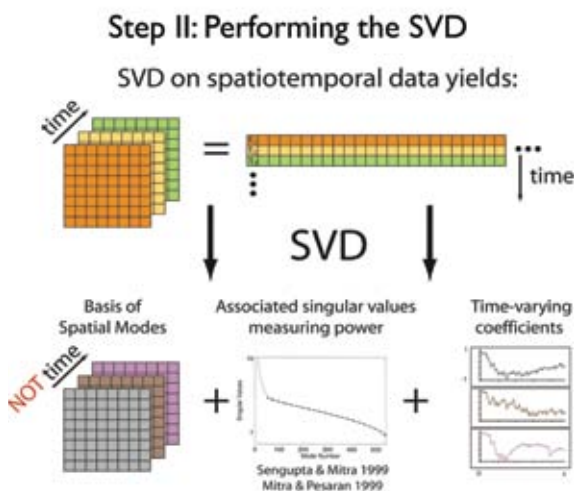


Figure 2. An SVD is performed on the data matrix, re-resulting in a set of eigenimages, singular values, and time-varying coefficients.

NOTES

of the submatrix is proportional to the time course. The overall weight of the submatrix formed in this way is given by the singular value associated with the eigenvector and the time course.

MATLAB has a routine for performing an SVD on a data matrix. To perform an SVD on our data matrix, we type

```
[u, s, v] = svd(X, 0);
```

The matrices u , s , and v that this routine outputs are the left eigenvectors (these are time-varying coefficients), the singular values, and the right eigenvectors (often called eigenimages), respectively, of the matrix X .

If you add up all the submatrices, the result is identically equal to the original data matrix. Sometimes, however, some of the submatrices may be identified as contributing only noise. In this case, they may be deleted, and the resulting reconstruction of the original data matrix will be less noisy. One commonly used method for choosing eigenvectors and time courses that are not part of the noise is to note the location of a pronounced “knee” in a plot of the singular values. This knee is used to identify a threshold beyond which eigenvectors are discarded.

To visualize the singular values, type the command

```
% plot singular values on
semilogarithmic plot
semilogy(diag(s));
```

To reconstruct the data matrix using only the first M eigenvectors, type

```
% reconstructed data matrix using only M
eigenvectors
Xrecon = u(:,1:M) * s(1:M,1:M) * v(:,1:M)';
```

You can use several different values for M and use a pseudo-color plot to compare the reconstruction with the original data.

The denoising process is depicted in Figure 3. First, the data matrix is formed (Fig. 3, top row). Next, an SVD is performed on the data (Fig. 3, middle row). Finally, the results from the SVD are examined and noisy eigenvectors discarded (Fig. 3, middle row; note the singular values to the left of the knee inside the red circle). The data matrix is then reconstructed by reversing the SVD decomposition (adding up all the nondiscarded submatrices), resulting in a denoised data set. Since it is hoped that only random noise was discarded in this process, SVD may also be viewed

as a data compression method. The retained eigenvectors and their time courses may be thought of as a low-dimensional summary of the data set.

The above method is a useful first step for denoising the neural imaging data. However, a number of possible problems can arise. One major problem is that some of the eigenvectors before the knee may be noisy. The reason for this phenomenon lies in the fact that the SVD is a blind method for analyzing data: The eigenvectors represent covarying information in the data set and nothing more, and the SVD does not contain any model of the signal. Therefore, for instance, patterned noise from the imaging apparatus (in which many pixels have covarying time courses) is likely to be depicted in the first few eigenvectors. Another problem is that small sources of covariance (possibly related to a signal of interest) may be discarded along with the eigenvectors below the knee.

One method for obtaining improved reconstructions of neural imaging data is to incorporate periodicity into the experimental design by using a periodic stimulus (Kalatsky and Stryker, 2003; Sornborger et al., 2005; Xu et al. 2008). By doing this, we are putting an extra “hook” in the data that can be used later to more accurately identify the information that is truly relevant to our experiment.

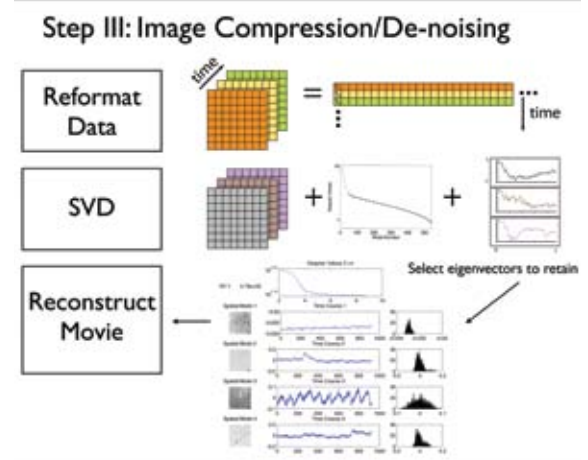


Figure 3. The data matrix may be reconstructed using only a subset of the eigenimages, their associated singular values, and time-varying coefficients. This reconstruction procedure can significantly reduce the noise in the data matrix.

In Figure 4, we show an SVD analysis of calcium imaging data taken of neural tissue that was stimulated periodically. Instead of the standard visualization in time (Fig. 4, top row, left panel), the frequency spectrum of

each left eigenvector (time-varying coefficient) was estimated using Thomson's multitaper spectral estimation method (Thomson, 1982) (top row, top right panel). Using Thomson's F-test, harmonic peaks in the spectrum at the stimulation frequency were detected and then estimated and extracted from the left eigenvectors (top row, bottom right panel). This process was performed for all left eigenvectors.

To calculate a multitaper estimate of the spectrum of the first 10 left eigenvectors and to visualize the information, we use the Chronux function "mtspectrumc" and a pseudo-color plot of the data:

```
% calculate frequency spectrum for a
left eigenvector
[S, f] = mtspectrum(u(:,1:10), params);
% plot the spectrum
pcolor([1:10], f, 10*log10(S)); shading flat;
```

To perform an F-test on the first 10 left eigenvectors, we use the Chronux function *f*-test:

```
% perform f-test on a left eigenvector of x
at frequency
% set sampling frequency parameter
params.Fs = 10;
% perform f-test
[Fval, A, fqs, sig, sd] = ftestc(u(:,1:10),
params);
```

This Chronux code outputs *Fval*, the value of the F-statistic for each frequency and channel; *A*, the complex line amplitude for each frequency and channel; *fqs*, the frequencies that the F-test was evaluated at; *sig*, the significance level of the test; and *sd*, the standard deviation of the amplitude.

A contour plot of the F-statistics across all frequencies for the first 10 left eigenvectors may be made with the following command:

```
% plot f-statistic for first 10 left
eigenvectors using contour plot
contour(Fval, [sig sig], 'k');
```

The frequencies at which significant harmonics were detected in the first left eigenvector may be output with the following command:

```
% frequencies of significant harmonics
fqs(find(Fval(:,1) > sig))
```

And the complex amplitudes of the significant harmonics are as follows:

```
% amplitudes of significant harmonics
A(find(Fval(:,1) > sig), 1)
```

Using the amplitude and frequency information from all the harmonics, we can reconstruct the statistically significant periodic part of the time series for the first left eigenvector:

```
% reconstruct periodic part
amps = A(find(Fval(:,1) > sig), 1);
fs = fqs(find(Fval(:,1) > sig));
sg = zeros(size(u(:,1)));
for I = 1:nh
    sg = sg + real(amps(i) * sin(2*pi*fs(i)*[1:nt]/
params.Fs) ...
+ amps(i)' * cos(2*pi*fs(i)*[1:nt]/
params.Fs));
end;
```

In general, we would loop through all the left eigenvectors in order to obtain all the periodic components in the data. This yields the periodic components for each of the first 10 left eigenvectors, $sg(i,j)$, where $i = 1 \dots nt$ and $j = 1 \dots 10$. Note that we could have investigated more than just the first 10 left eigenvectors; this is just an example. To reconstruct the periodic part of the entire data matrix, we simply put the pieces back together:

```
Xperiodrecon = sg(:,1:10) * s(1:10,1:10) * v(:,1:10)';
```

The results of the spectral and harmonic analyses of the first 10 time courses are depicted in Figure 4, bottom row, left panel. This panel shows a pseudo-color plot of the log-spectra of all time-varying coefficients plotted side by side. Note the bright yellow peaks in the third, fourth, and fifth time courses. In Figure 4 (bottom row, right panel), we plot contours of the harmonic peaks that were found across all the time-varying coefficients. Note the contours corresponding to the spectral peaks in the left panel.

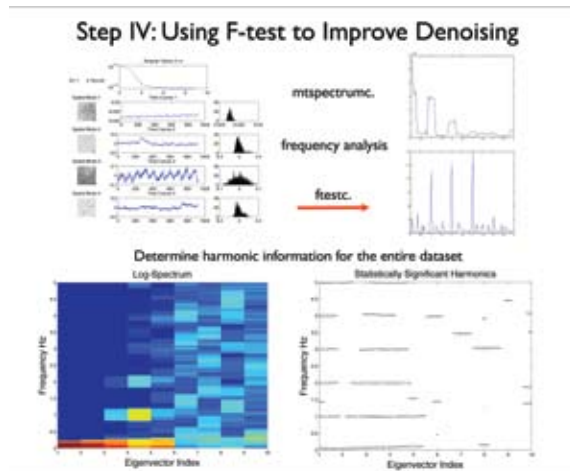


Figure 4. When a periodic stimulus/response experimental paradigm has been used, Thomson's spectral analysis may be used to visualize periodicity in the data set, and Thomson's F-test may be used to detect and estimate harmonics in the data.

NOTES

Figure 5 depicts the complete analysis process. After construction of the data matrix, we perform the SVD. Then, using the F-test, we extract statistically significant periodic responses in the data. Next, only the sinusoids with frequencies that are at multiples of the stimulus frequency are extracted. A reconstructed data set is then made of just the periodic response.

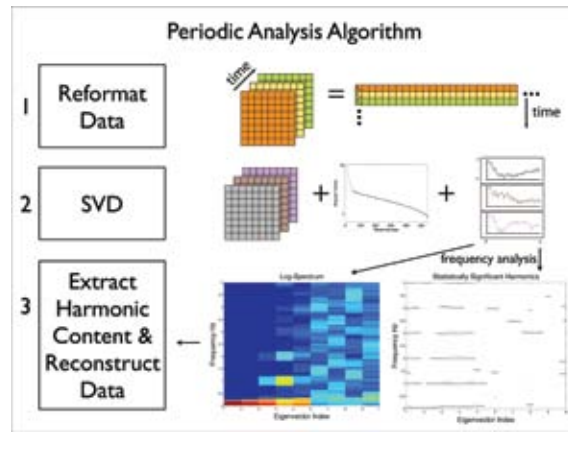


Figure 5. By combining statistically significant harmonics information with the associated eigenimages, a new data matrix may be reconstructed that represents only the periodic response to the stimulus.

Conclusion

Denoising data using the above methods can be extremely useful for improving the signal-to-noise ratio of neural imaging data. We have discovered that data features that would otherwise be missed can often be found using these methods. In the data set used in the above example, clear evidence was found of functional neuronal projections from one neuron to another within the neural tissue that was imaged. Without applying our methods, this functional information would not have been detectable.

References

- Kalatsky VA, Stryker MP (2003) New paradigm for optical imaging: temporally encoded maps of intrinsic signal. *Neuron* 38:529-545.
- Mitra PP, Bokil H (2008) *Observed brain dynamics*. Oxford, UK: Oxford UP.
- Sengupta AM, Mitra PP (1999) Distributions of singular values for some random matrices. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 60:3389-3392.
- Sornborger A, Yokoo T, Delorme A, Sailstad C, Sirovich L (2005) Extraction of the average and differential dynamical response in stimulus-locked experimental data. *J Neurosci Methods* 141:223-229.

Thomson, DJ (1982) Spectrum estimation and harmonic analysis. *Proc IEEE* 70:1055-1096.

Xu J, Sornborger AT, Lee JK, Shen P (2008) *Drosophila TRPA channel modulates sugar-stimulated neural excitation, avoidance and social response*. *Nat Neurosci* 11:676-682.