Multivariate Analysis

Homework #2 (2008) Answers

This is a simple example of ICA applied to the cocktail party problem. We will generate a mixture of signals and then attempt to unmix them.

Q1. First, let's generate the signals. We take two non-Gaussian sources, each of which can assume a value of -1 or 1 with a probability of 0.5. That is, a "typical" example of four samples is represented by the matrix $S = \begin{pmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}$. *(Geometrically, these are positioned at the four corners of a square.) Our measured time series are represented by the columns of* $Y = SM$. *To keep things really simple, we'll assume that the mixing matrix M is the identity,* $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, *so* $Y = S$.

Now, (forgetting that $Y = S$*) we want to try to unmix Y to recover the sources S. We first apply principal components to write* $Y = XA$, *and then we will look for choices of A that maximize the non-Guassian-ness of the columns* $X = YA^{-1}$

That is, we look for the eigenvectors and eigenvalues of $Y^T Y = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$. *This is the identity, so any pair of orthonormal vectors can serve for A. So we can write*

$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, *and* $A^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$.

Given this setup, what rotations (i.e., what values of θ*) maximize the kurtosis-squared of the columns of* YA^{-1}*? What rotations maximize the negentropy (minimize the entropy) of the columns of* YA^{-1}*?*

Solution.

$YA^{-1} = \begin{pmatrix} -\cos\theta + \sin\theta & -\cos\theta - \sin\theta \\ -\cos\theta - \sin\theta & \cos\theta - \sin\theta \\ \cos\theta + \sin\theta & -\cos\theta + \sin\theta \\ \cos\theta - \sin\theta & \cos\theta + \sin\theta \end{pmatrix}$. The kurtosis of a set of values u_i (whose mean is 0) is defined by $\kappa = \dfrac{<u^4> - 3<u^2>^2}{<u^2>^2} = \dfrac{m_4 - 3<m_2>^2}{<m_2>^2}$. So we need to find the second and fourth moments and of each column. Note that the columns are identical except for their order, so it suffices to calculate it for one column.

$$m_2 = \frac{1}{4}\left((\cos\theta + \sin\theta)^2 + (-\cos\theta + \sin\theta)^2 + (\cos\theta - \sin\theta)^2 + (-\cos\theta - \sin\theta)^2\right)$$

from which it follows that

$$m_2 = \frac{1}{4}\left(4\cos^2\theta + 4\sin^2\theta\right) = 1.$$

$$m_4 = \frac{1}{4}\left((\cos\theta + \sin\theta)^4 + (-\cos\theta + \sin\theta)^4 + (\cos\theta - \sin\theta)^4 + (-\cos\theta - \sin\theta)^4\right)$$

from which it follows that

$$m_4 = \cos^4\theta + 6\cos^2\theta\sin^2\theta + \sin^4\theta.$$

Using

$$\sin^2\theta = 1 - \cos^2\theta,$$

this becomes

$m_4 = \cos^4\theta + 6\cos^2\theta(1-\cos^2\theta) + (1-\cos^2\theta)^2 = 1 + 4\cos^2\theta - 4\cos^4\theta$, and, with one more trig identity,

$$m_4 = 1 + 4\cos^2\theta - 4\cos^4\theta = 2 - (1 - 2\cos^2\theta)^2 = 2 - \cos^2(2\theta).$$

So $\kappa = \frac{m_4 - 3 < m_2 >^2}{< m_2 >^2} = -1 - \cos^2 2\theta$. This is furthest from 0 when $\cos^2 2\theta$ is 1 or -1, i.e., when 2θ is 0 or π, i.e., when θ is 0, $\pi/2$, π, or $3\pi/2$. At these points, $\kappa = -2$.

So ICA based on kurtosis has “singled out” four specific rotations. Each of these rotations keep the square aligned to the axis, so ICA based on kurtosis has found the unmixing, and would reject any other rotation as suboptimal. For example, $\theta = \pi/4$ leads to $\kappa = -1$, and $\theta = \pi/6$ leads to $\kappa = -5/4$.

Now, negentropy.

To calculate negentropy (or entropy) of a column, we need to know the probability distribution of its values.

When θ is in $\{0, \pi/2, \pi, 3\pi/2\}$, each column contains only two values, +1 or -1, and each occurs 2/4 times. So for these θ's, the negentropy is

$$NE = \frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2} = -\log 2.$$

When θ is in $\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$, each column contains three values: 0 occurs 2/4 times, and $\pm\sqrt{2}$ each occur $1/4$ times. (Think of rotating a square by 45 deg, and projecting the points onto one axis.) So for these θ's, the negentropy is

$$NE = \frac{1}{4}\log\frac{1}{4} + \frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} = -\frac{3}{2}\log 2.$$

For all other values of θ, each column contains four distinct values. (Rotate a square by a "generic" angle, and project onto an axis. All four corners project onto different points.) So for these θ's, the negentropy is

$$NE = \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4} = -\log 4 = -2\log 2 .$$

So negentropy is also maximized when θ is in $\{0, \pi/2, \pi, 3\pi/2\}$. Interestingly, negentropy picks out a 45-deg rotation as "second-best", and considers all other rotations equally bad – while the kurtosis criterion considers the all the 45-deg rotation to be the worst. But both criteria pick out the right rotation as being the best unmixer.