Multivariate Methods

Homework #2 (2014-2015), Answers

Q1: A three-stimulus brain in a two-stimulus world

Consider a toy functional imaging experiment, in which the brain has 3 pixels, and there are two stimuli. Say that stimulus 1 causes an activation of +2 units in pixel 1, and -1 unit in pixels 2 and 3; say that stimulus 2 causes an activation of +2 units in pixel 2, and -1 unit in pixels 1 and 3. So we have a 3×2 *data matrix* Y.

A. Compute its principal components, $Y = XB$, *with the columns of* X *orthonormal, and the rows of* B *orthogonal (but not necessarily orthonormal).*

$Y = \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix}$, so $YY^* = \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 5 & -4 & -1 \\ -4 & 5 & -1 \\ -1 & -1 & 2 \end{pmatrix}$. But it's easier to determine the eigenvalues of $Y^*Y = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$. We use the approach described under "Symmetry-an important practical issue": seek Z as the first p (row) eigenvectors of the $k\times k$ matrix Y^*Y, and then find $X = YZ^*\Lambda^{-1/2}$ and $B = \Lambda^{1/2}Z$, where Λ is the matrix with the eigenvalues of Y^*Y on the diagonal.

We could find the eigenvalues of Y^*Y by solving its characteristic equation, $\det(\lambda I - Y^*Y) = 0$. This is $\det\left(\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}\right) = \det\begin{pmatrix} \lambda-6 & 3 \\ 3 & \lambda-6 \end{pmatrix} = (\lambda-6)^2 - 9 = \lambda^2 - 12\lambda + 27 = (\lambda-9)(\lambda-3)$, which has roots $\lambda = 9$ and $\lambda = 3$. Or we could note that Y^*Y is symmetric under interchange of coordinates, so its eigenvectors must lie in subspaces that are preserved under interchange of coordinates, and therefore, must be proportional to $v_+ = (1,1)$ and $v_- = (1,-1)$ -- and determine the eigenvectors by computing $v_+Y^*Y = 3v_+$, and $v_-Y^*Y = 9v_-$. Noting that $v_- = (1,-1)$ and $v_+ = (1,1)$ have squared-lengths of 2, we find the matrix of orthonormalized row eigenvectors, $Z = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, and $\Lambda = \begin{pmatrix} 9 & 0 \\ 0 & 3 \end{pmatrix}$.

So $B = \Lambda^{1/2}Z = \begin{pmatrix} 3 & 0 \\ 0 & \sqrt{3} \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 3 & -3 \\ \sqrt{3} & \sqrt{3} \end{pmatrix}$.

$$X = YZ^*\Lambda^{-1/2} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^*\begin{pmatrix} 9 & 0 \\ 0 & 3 \end{pmatrix}^{-1/2} = \frac{1}{\sqrt{2}}\begin{pmatrix} 2 & -1 \\ -1 & 2 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1/3 & 0 \\ 0 & 1/\sqrt{3} \end{pmatrix}.$$

And

$$= \frac{1}{\sqrt{2}}\begin{pmatrix} 3 & 1 \\ -3 & 1 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} 1/3 & 0 \\ 0 & 1/\sqrt{3} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1/\sqrt{3} \\ -1 & 1/\sqrt{3} \\ 0 & -2/\sqrt{3} \end{pmatrix}$$

Verify that $Y = XB$, *where the columns of* X *orthonormal, and the rows of* B *orthogonal.*

$Y = XB$:

$$XB = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1/\sqrt{3} \\ -1 & 1/\sqrt{3} \\ 0 & -2/\sqrt{3} \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} 3 & -3 \\ \sqrt{3} & \sqrt{3} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 4 & -2 \\ -2 & 4 \\ -2 & -2 \end{pmatrix} = Y$$

X has orthonormal columns: $X^*X = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -2/\sqrt{3} \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1/\sqrt{3} \\ -1 & 1/\sqrt{3} \\ 0 & -2/\sqrt{3} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = I$.

B has orthogonal rows: $BB^* = \frac{1}{\sqrt{2}}\begin{pmatrix} 3 & -3 \\ \sqrt{3} & \sqrt{3} \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} 3 & \sqrt{3} \\ -3 & \sqrt{3} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 18 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 9 & 0 \\ 0 & 3 \end{pmatrix}$

Note that the first principal component (i.e, the first column of *X*) is a mixture of the two responses – and is in some sense more complicated than either of them.

Q2. Rotation of principal components. Same setup as Q1. Let's see if we can find a simple way to unmix these components. Let $\vec{u}_1 = \vec{x}_1 \cos\theta + \vec{x}_2 \sin\theta$, $\vec{u}_2 = -\vec{x}_1 \sin\theta + \vec{x}_2 \cos\theta$. *Since the* $\vec{u}_i$ *are a non-singular linear combination of the* $\vec{x}_i$, *they necessarily also account for the data matrix* Y. *We might consider a transformation to the* $\vec{u}_i$ *to be simpler if the coefficients in the* $\vec{u}_i$ *are smaller. The* $\vec{u}_i$, *like the* $\vec{x}_i$, *constitute the columns of a* 3×2 *matrix,* $U(\theta)$. *Is there a rotation* θ *that minimizes the sum of the squares of these 6 quantities? If so, find it; if not, explain why and suggest alternative strategies.*

The sum of the squares of the entries in $U(\theta)$ is independent of θ, and identical to that of X. The reason is that $U = XR_\theta$, where $R_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, a rotation matrix.Since $R_\theta R_\theta^* = I$,
$tr(UU^*) = tr\left(XR_\theta (XR_\theta)^*\right) = tr\left(XR_\theta R_\theta^* X^*\right) = tr(XX^*)$.

An alternative is to extremize the sum of the fourth powers of the entries of $U(\theta)$. This will seek components whose entries have either very large values, or very small values (i.e., a "sparse" representation). This is effectively the "varimax" procedure.