NeuroImage Processing Final Exam Solutions

Moo K. Chung
mkchung@wisc.edu

December 3, 2009

1. Group A consists of 3 people and group B consists of 2 people. Cortical thickness for group A is 1, 2 and 3 mm for some cortical region. Cortical thickness for group B is 1 and 2 mm. Estimate the thickness of group A using the least squares method. What is the sum of squared error (SSE) of your estimation? Set up a general linear model testing for equality of cortical thickness between the two groups, and compute the actual F-statistic value.

   Solution. Let λ be the cortical thickness of group A. Then we need to solve

   $$(1,2,3)' = (1,1,1)'\lambda.$$

   This is easily solved by noting that

   $$\lambda = [(1,1,1)(1,1,1)']^{-1}(1,1,1)(1,2,3)' = 2.$$

   The sum of squared errors (SSE) of the estimation is then

   $$\text{SSE } = (1-2)^2 + (2-2)^2 + (3-2)^2 = 2.$$

   We have the following linear model

   $$\texttt{thick}_i = \lambda + \beta \cdot \texttt{group}_i,$$

   where the dummy variable $\texttt{group}_i$ is 0 for group A and 1 for group B. So we have the group variable to be $(0,0,0,1,1)$. We test for the group effect, $\beta = 0$.

   For the reduced model ($\beta = 0$), the estimation $\widehat{\lambda}_0 = 1.8$. The corresponding $\text{SSE}_0 = 2.8$. For the full model, $\widehat{\lambda}_1 = 2, \widehat{\beta}_1 = -0.5$. The corresponding $\text{SSE}_1 = 2.5$. The F-stat is given by

   $$\frac{(SSE_0 - SSE_1)/1}{2.8/(5-1-1)} = 0.3214.$$

   The p-value is $1 - \texttt{fcdf}(\texttt{F}, 1, 3) = 0.6104$. If you used $(1,1,0,0,0)$ for the group variable, you get the same answers except for the estimation for the full model is slightly different at $\widehat{\lambda}_1 = 1.5, \widehat{\beta}_1 = 0.5$.

2. Given a T1-weighted MRI of brain, the proportions of gray matter and white matter in the image are 0.7 and 0.3 respectively. Then explain why we can model image intensity values as the mixture of two Gaussian components with mixing proportions 0.7 and 0.3.

   Solution. Then k-components mixture model on image intensity values assume image intensity values Y to come from k different distributions f_i with proportions p_i. This can be modeled by conditioning on a multinomial distribution. Another way of saying this is that the the k-components mixture model can be obtained by mixing samples obtained from distributions f_j with p_j proportions.

   Let X_j be an indicator variable for the j-th class such that $P(X_j = 1) = p_j$ and $P(X_j = 0) = 1 - p_j$. The collection of variables $X = (X_1, \cdots, X_k)$ form a multinomial distribution with parameters $(p_1, \cdots, p_k)$ if we have constraint $X_1 + \cdots + X_k) = 1$. The probability mass function of the multinomial distribution is given by

$$f(x_1, \cdots, x_k) = p_1^{x_1} \cdots p_k^{x_k}.$$

   We define a random variable Y conditionally on the event $X_j = 1$ such that $Y \sim f_j$ if $X_j = 1$. The conditional density $f(y|x_j = 1) = f_j$ is the distribution for the j-th class. Then the joint density is given by $f(x_j = 1, y) = p_j f_j(y)$ for each j, which can be compactly written as

$$f(x, y) = [p_1 f_1(y)]^{x_1} \cdots [p_k f_k(y)]^{x_k}.$$

   The marginal density of Y is trivially then

$$f(y) = \sum_x f(x, y) = \sum_{i=1}^{k} p_i f_i(y).$$

3. Evaluate kernel smoothing $K * I$ at the pixel position (2,2). Note that at (2,2), image intensity value is -2. Why heat kernel smoothing is used for cortical thickness data instead of Gaussian kernel smoothing?

   Solution. $K * I(2, 2) = -1/8$. Gaussian kernel is defined as $K_\sigma(p, q) \exp[-\frac{\|p-q\|^2}{2\sigma^2}]$, where $\|p - q\|$ is the Euclidian distance between p and q. Since a cortical surface is not Euclidian, the kernel is not defined along the cortical surface. So we need to refine the kernel using the geodesic distance rather than the Euclidian distance. Then the kernel becomes the heat kernel and we have heat kernel smoothing instead of Gaussian kernel smoothing.

4. Determine the fractional anisotropy (FA) index. The following MATLAB code is provided. What does FA value measure?

   Solution. FA measures the degree of anisotropy of water diffusion in white matter fibers. It varies from zero to one and obtains 0 when all eigenvalues are equal indicating isotropy of diffusion. It is defined as FA $= \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - \bar{\lambda})^2 + (\lambda_2 - \bar{\lambda})^2 + (\lambda_3 - \bar{\lambda})^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$. This definition normalizes FA values properly between 0 and 1. The definition of FA originally given in Basser and Pierpaoli (1998) does not normalize properly. Note that FA is 0 if all eigenvalues are identical and 1 if two eigenvalues are all zero. Based on this defintion, we obtain FA = 0.3536 for our problem.

5. Given the following two brain connectivity graphs A and B, write down the adjacency matrices. Write down the degree distributions. Compute the clustering coefficients. Which brain network is more complex?

   Solution. The adjacency matrix for the graph A is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

   while that of the graph B is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

   One can also define adjacency matrices in such a way that all the diagonal terms are 1. Obviously a given node is connected to itself by definition. Adjacency matrices are symmetric. The degrees computed by summing of either row or column vectors. At nodes $(1, 2, 3, 4)$, degrees are given by by $(1, 3, 2, 2)$ and $(1, 3, 1, 1)$. Hence the degree distributions are $(1, 2, 1)$ and $(3, 0, 1)$.

   There are many different definitions of *clustering coefficients*. We follow the defintion of Watts and Strogatz (1998) where the clustering coefficient of the i-th node is given by

$$C_i = \frac{\#\text{ of triangles connected to node } i}{\#\text{ of triples centered on node } i}.$$

   Triples are three connected nodes. For nodes with degree 0 and 1, we define $C_i = 0$. Then we have $(0, 1/3, 1, 1)$ and $(0, 0, 0, 0)$ for the graphs A and B. The clustering coefficient of a graph is given by

$$C = \frac{1}{n}\sum_{i=1}^{n} C_i.$$

   So we have $C_A = 7/14$ and $C_B = 0$. So the graph A is considered more complex.

6. Given a deformation field $(x, y)' \to (2x - 2y + 1, x + 4y - 1)'$ for warping 2D image A to another 2D image B, determine the displacement vector field. Determine the Jacobian determinant at (0,0). Is the pixel at (0,0) in image A shrinking or enlarging under the deformation? What is the main difference between deformation-based morphometry (DBM) and tensor-based morphometry (TBM)?

   Solution. The displacement vector field is simply given by

$$\mathbf{u} = (2x - 2y + 1, x + 4y - 1)' - (x, y)' = (x - 2y + 1, x + 3y - 1)'.$$

   The displacement gradient is given by

$$\frac{\partial \mathbf{u}}{\partial (x, y)} = \begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}.$$

The Jacobian matrix J is then

$$J = \begin{pmatrix} 2 & -2 \\ 1 & 4 \end{pmatrix}$$

and its determinant is 10. Hence we have huge volume enlargement. DBM uses the relative spatial position difference while TBM uses spatial derivatives in characterizing anatomical difference.