Multivariate Regression Generalized Likelihood Ratio Tests for FMRI Activation

Daniel B. Rowe

Division of Biostatistics
Medical College of Wisconsin

Technical Report 40

November 2002

Division of Biostatistics
Medical College of Wisconsin
8701 Watertown Plank Road
Milwaukee, WI 53226
Phone:(414) 456-8280

MEDICAL COLLEGE OF WISCONSIN

Multivariate Regression Generalized Likelihood Ratio Tests for FMRI Activation

Daniel B. Rowe
Medical College of Wisconsin
8701 Watertown Plank Rd.
Milwaukee, WI 53226
dbrowe@mcw.edu

Abstract

In neuroscience, an important research question to be investigated, is whether a region or regions of the brain are being activated when a subject is presented a stimulus. A few methods are in use to address this question but they do not jointly take into account the spatial relationship among the set of voxels under consideration. Multivariate regression can determine whether the set of voxels in one, or several regions of interest are related to the experimental paradigm, in addition to individual measures of activation which are different.

1 INTRODUCTION

In functional magnetic resonance imaging, the observed time courses in voxels are commonly analyzed using multiple linear regression. The typical analysis [9] is to model a voxel's observed time course as a linear function involving an intercept, a scaled time trend, and a scaled reference function. This typical multiple regression analysis assumes that the voxels are independent.

The multivariate regression model [10] is a generalization of the multiple regression model to vector valued observations with dependent elements (voxels). This generalization allows spatial dependencies between voxels to be incorporated into inferences of significance.

With multiple regression, inferences and hypotheses can only be evaluated on individual voxels. With multivariate regression, inferences and hypothesis can be evaluated on a set of voxels. That is, with multiple regression, one gets independently computed measures of the degree of activation in voxel 1, voxel 2, and so on. Multivariate regression gives a jointly computed measure of activation in a region of interest in addition to individual measures of activation.

2 UNIVARIATE MODEL

In the multiple linear regression model, at time point i, the observed hemodynamic response in voxel j is y_{ji},

$$y_{ji} \quad = \quad \beta_{0j} + \beta_{1j}x_{1i} + \cdots + \beta_{qj}x_{qi} \quad + \quad \epsilon_{ji} \tag{2.1}$$

which is written in terms of vectors as

$$\begin{array}{ccccccc} y_{ji} & = & (1, x_{1i}, \ldots, x_{qi}) & \begin{pmatrix} \beta_{0j} \\ \vdots \\ \beta_{qj} \end{pmatrix} & + & \epsilon_{ji} \\ y_{ji} & = & x_i' & \beta_j & + & \epsilon_{ji} \\ 1 \times 1 & & 1 \times (q+1) & (q+1) \times 1 & & 1 \times 1 \end{array} \tag{2.2}$$

where $i = 1, \ldots, n$ and $j = 1, \ldots, p$.

This model, for each voxel j and all n time points, is written in terms of vectors and matrices as

$$\begin{array}{ccccccc} \begin{pmatrix} y_{j1} \\ \vdots \\ y_{jn} \end{pmatrix} & = & \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} & \begin{pmatrix} \beta_{0j} \\ \vdots \\ \beta_{qj} \end{pmatrix} & + & \begin{pmatrix} \epsilon_{j1} \\ \vdots \\ \epsilon_{jn} \end{pmatrix} \\ Y_j & = & X & \beta_j & + & E_j. \\ n \times 1 & & n \times (q+1) & (q+1) \times 1 & & n \times 1 \end{array} \tag{2.3}$$

The errors of observation ϵ_{ji} are assumed to be independent and normally distributed with zero mean and variance σ_j^2. The likelihood is given by

$$p(Y_j|\beta_j, \sigma_j^2, X) = (2\pi)^{-\frac{n}{2}}(\sigma_j^2)^{-\frac{n}{2}} e^{-\frac{(Y_j - X\beta_j)'(Y_j - X\beta_j)}{2\sigma_j^2}}. \tag{2.4}$$

It can be shown [10] that the maximum likelihood estimate of the vector of regression coefficients $\hat{\beta}_j$ for each voxel j is

$$\hat{\beta}_j = (X'X)^{-1}X'Y_j, \tag{2.5}$$

that $\hat{\beta}_j$ is multivariate Student t distributed,

$$\hat{\beta}_j \sim t\left(n - q - 1, \beta_j, (n-q-1)^{-1}g_j(X'X)^{-1}\right), \tag{2.6}$$

and $\hat{\beta}_{kj}$ is univariate Student t distributed,

$$\hat{\beta}_{kj} \sim t\left(n - q - 1, \beta_{kj}, (n-q-1)^{-1}g_jW_{kk}\right) \tag{2.7}$$

where $g_j = (Y_j - X\beta_j)'(Y_j - X\beta_j)$ while W_{kk} is the kk^{th} element of $W = (X'X)^{-1}$. Note that $\hat{\sigma}_j^2 = g_j/n$ is the maximum likelihood estimate of σ_j^2.

Hypothesis (for each voxel) such as

$$\begin{array}{llll} H_0: & C_j\beta_j = \gamma_j & vs \quad H_1: & C_j\beta_j \neq \gamma_j \\ & \sigma_j^2 > 0 & & \sigma_j^2 > 0 \end{array} \tag{2.8}$$

can be evaluated where C_j is an $r \times (q+1)$ matrix of full row rank with linear constraints as rows and γ_j is an $r \times 1$ vector. This is done with the use of the

$$F = \frac{(C\hat{\beta}_j - \gamma_j)'[C(X'X)^{-1}C']^{-1}(C\hat{\beta}_j - \gamma_j)}{rg_j/(n-q-1)} \tag{2.9}$$

which under the null hypothesis follows an F-distribution with r and $n-q-1$ degrees of freedom. This statistic is derived (see appendix) from a likelihood ratio statistic. Under the null hypothesis, the likelihood is maximized subject to the constraint that $C_j\beta_j = \gamma_j$ using Lagrange multipliers.

For example, $H_0: \beta_{kj} = 0$ can be evaluated with $\gamma_{kj} = 0$, $C = (\ldots, 0, 1, 0, \ldots)$ is a $(q+1)$ dimensional zero row vector except a one in the k^{th} column, and either of the test statistics

$$t_{kj} = \frac{\hat{\beta}_{kj} - \gamma_{kj}}{[W_{kk}g_j/(n-q-1)]^{\frac{1}{2}}} \tag{2.10}$$

$$F_{kj} = \frac{(\hat{\beta}_{kj} - \gamma_{kj})^2}{W_{kk}g_j/(n-q-1)} \tag{2.11}$$

which are distributed as either univariate student t with $n-q-1$ degrees of freedom or F with 1 and $n-q-1$ numerator and denominator degrees of freedom respectively. In the above statistics which can be derived from a likelihood ratio statistic, g_j is computed under the alternative hypothesis. In order for the above statistics to be computable, $(X'X)$ has to be invertible and $n > q+1$

3 MULTIVARIATE MODEL

In the multivariate linear regression model, at time point i, the observed hemodynamic response in all p voxels is y_i,

$$\begin{pmatrix} y_{1i} \\ \vdots \\ y_{pi} \end{pmatrix} = \begin{pmatrix} \beta_{01} + \beta_{11}x_{1i} + \cdots + \beta_{q1}x_{qi} \\ \vdots \\ \beta_{0p} + \beta_{1p}x_{1i} + \cdots + \beta_{qp}x_{qi} \end{pmatrix} + \begin{pmatrix} \epsilon_{1i} \\ \vdots \\ \epsilon_{pi} \end{pmatrix} \tag{3.1}$$

which can be written as

$$\begin{array}{ccccccc}
(y_{1i},\ldots,y_{pi}) & = & (1,x_{1i},\ldots,x_{qi}) & \begin{pmatrix} \beta_{01} & \beta_{02} & \ldots & \beta_{0p} \\ \vdots & & & \\ \beta_{q1} & \beta_{q2} & \ldots & \beta_{qp} \end{pmatrix} & + & (\epsilon_{1i},\ldots,\epsilon_{pi}) \\
y_i' & = & x_i' & (B_0,B_1,\ldots,B_q)' & + & \epsilon_i' \\
y_i' & = & x_i' & B' & + & \epsilon_i' \\
1\times p & & 1\times(q+1) & (q+1)\times p & & 1\times p
\end{array} \tag{3.2}$$

where $i=1,\ldots,n$.

The model, for all p voxels and all n time points, is written as

$$\begin{array}{ccccccc}
\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} & = & \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} & \begin{pmatrix} \beta_{01} & \beta_{02} & \ldots & \beta_{0p} \\ \vdots & & & \\ \beta_{q1} & \beta_{q2} & \ldots & \beta_{qp} \end{pmatrix} & + & \begin{pmatrix} \epsilon_1' \\ \vdots \\ \epsilon_n' \end{pmatrix} \\
Y & = & X & B' & + & E. \\
n\times p & & n\times(q+1) & (q+1)\times p & & n\times p
\end{array} \tag{3.3}$$

Notice that if $p=1$, this reduces to the univariate regression model. Each row of Y, for example the i^{th}, is the observed values in all p voxels at time i and each column of Y, for example the j^{th} is the observed values at all n time points in voxel j.

The errors of observation ϵ_i are assumed to be independent and normally distributed with p dimensional zero mean vector and $p\times p$ positive definite covariance matrix Σ. This means that for each observation, which is a row in the left hand side of the model, there is a regression. Each row has its own regression complete with its own set of regression coefficients. It can be shown [10] that the estimate of the matrix of regression coefficients $\hat{B}'$ for all voxels is

$$\hat{B}' = (X'X)^{-1}X'Y, \tag{3.4}$$

that $\hat{B}$ is matrix Student t distributed,

$$\hat{B} \sim t\left(n-q-1, B, [(n-q-1)(X'X)]^{-1}, G\right), \tag{3.5}$$

that $\hat{B}_k$ is multivariate Student t distributed,

$$\hat{B}_k \sim t\left(n-q-p, B_k, (n-q-p)^{-1}W_{kk}G\right), \tag{3.6}$$

that $\hat{\beta}_j$ is multivariate Student t distributed,

$$\hat{\beta}_j \sim t\left(n-q-p, \beta_j, (n-q-p)^{-1}g_j(X'X)^{-1}\right), \tag{3.7}$$

and $\hat{\beta}_{kj} = \hat{B}_{jk}$ is univariate Student t distributed,

$$\hat{\beta}_{kj} \sim t\left(n-q-p, \beta_{kj}, (n-q-p)^{-1} W_{kk} g_j\right), \tag{3.8}$$

where $G = (Y - XB')'(Y - XB')$, g_j is its j^{th} diagonal element, while W_{kk} is the k^{th} diagonal element of $W = (X'X)^{-1}$. Note that $\hat{\Sigma} = G/n$ is the maximum likelihood estimate of Σ.

Notice that the estimate of the matrix of regression coefficients $\hat{B}$ does not depend on the (spatial) covariance between the voxels Σ or its estimate $\hat{\Sigma}$. So whether or not the voxels are spatially correlated, the estimate of the coefficient matrix is the same. However, the spatial correlation between the voxels does matter in the covariance matrix of the estimated coefficients. When we make significance statements about the coefficients, the spatial correlation matters!

Hypothesis regarding a particular coefficient (for all voxels or a subset of voxels) such as

$$\begin{array}{llllllll} H_0: & CB' & = & \Gamma & vs & H_1: & CB' & \neq \Gamma \\ & \Sigma & > & 0 & & & \Sigma & > 0 \end{array} \tag{3.9}$$

(i.e. $B_k = 0$) can be evaluated with the test statistic

$$F_k = \frac{n-q-p}{p} (\hat{B}_k - \Gamma_k)' W_{kk}^{-1} G^{-1} (\hat{B}_k - \Gamma_k) \tag{3.10}$$

which follows an F-distribution with p numerator and $n-q-p$ denominator degrees of freedom respectively. In order for this F statistic to be computable $(X'X)$ has to be invertible and $n \geq q+p$. Note that G is invertible if $p \leq n$ or the number of voxels is less than or equal to the number of time points [5]. In the above statistic, G is computed under the alternative hypothesis.

Note that if each of the elements of vector y_i were independent, then Σ would be diagonal. Since we assume that each voxel has its own distinct error term (i.e. its own $\sigma_j^2 = \Sigma_{jj}$), $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_p^2)$ is a matrix with nonzero variances along the diagonal and zero covariances off the diagonal.

After determining ROI activation via Equation 3.10, individual voxel activation can be determined via

$$t_{kj} = \frac{\hat{\beta}_{kj} - \gamma_{kj}}{[W_{kk} g_j/(n-q-p)]^{\frac{1}{2}}} \tag{3.11}$$

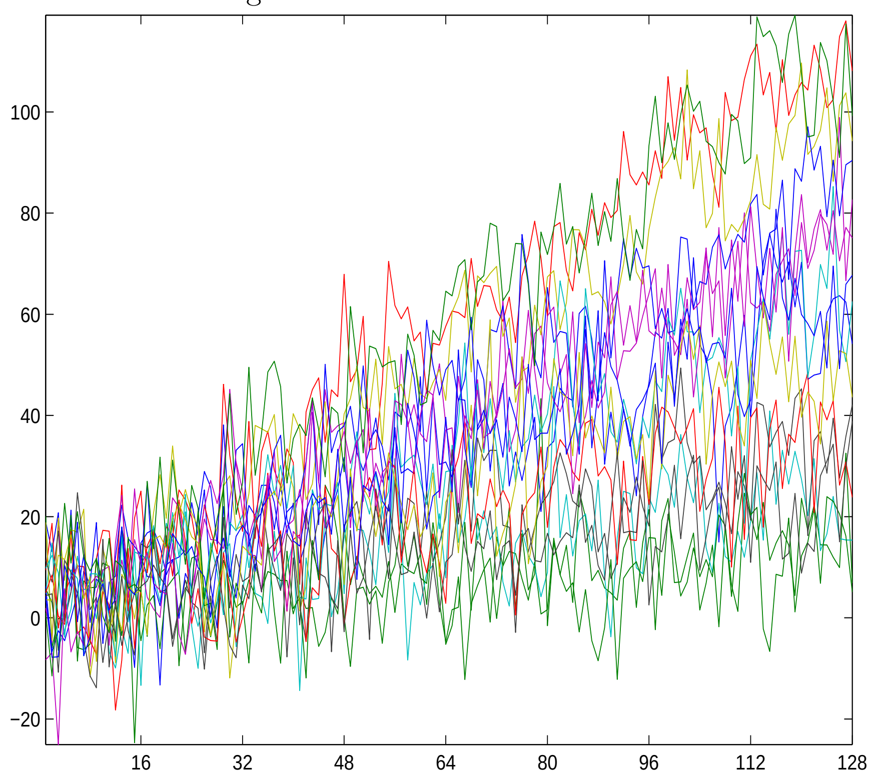
$$F_{kj} = \frac{(\hat{\beta}_{kj} - \gamma_{kj})^2}{W_{kk} g_j/(n-q-p)} \tag{3.12}$$

which are distributed as either univariate student t with $n-q-p$ degrees of freedom or F with 1 numerator and $n-q-p$ denominator degrees of freedom respectively from Equation 3.8.

4 SIMULATION

From an FMRI experiment, a 4×4 ROI is selected from a single slice. For this ROI, simulated FMRI data is constructed. The voxels in the ROI are numbered sequentially from top left to bottom right and stacked in increasing numerical order. The simulated data consists of $n = 128$ data points for $p = 16$ voxels. The simulated data is generated according to Equation 3.3 where the design matrix X is a $n \times 3$ matrix whose first column is an n dimensional vector of ones, the second column is at an n dimensional vector of the first counting numbers, and the third column is an n dimensional vector consisting of eight replicates of eight ones then eight negative ones. The true regression coefficient matrix B is given in Table 2. The voxels were assumed to have an AR(1) spatial correlation Σ as in Table 3 and the observation errors were randomly generated multivariate normal variates with zero mean and covariance matrix Σ. Values for the variance and corelation were selected to be $\sigma^2 = 64$ and $\rho = 0.25$. The simulated voxel time courses are displayed in Figure 1 and sample error covariance matrix is given in Table 5.

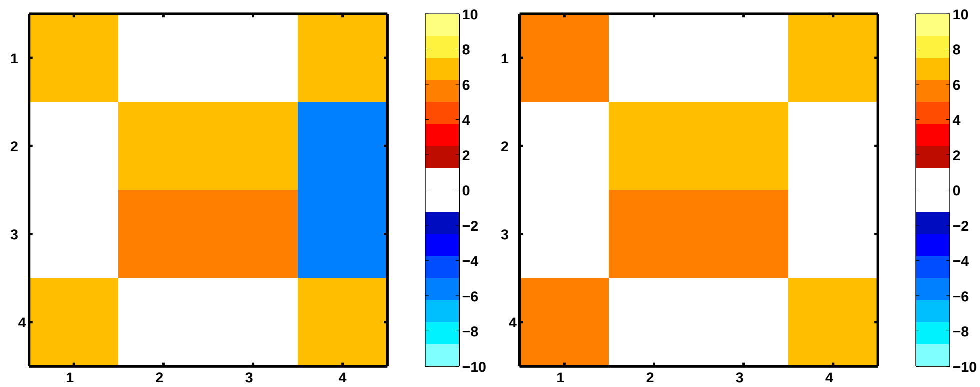
Figure 1: Simulated time courses.



A multivariate regression was performed and the estimated regression coefficients are displayed in Table 2. An F-test was performed using Equation 3.10

to jointly determine if the vector of coefficients corresponding to the reference function is zero or not. This is a test to determine if there was activation in the ROI due to the presentation of the stimulus. The F-statistic was $F = 43.7380$ with a critical value of $F_{10^{-6},16,128-2-16} = 4.4614$. This was also computed assuming that G was diagonal and resulted in $F = 28.2180$ with the same critical value. It is concluded that there is activation in the ROI due to the presentation of the stimulus. Recall that an ANOVA for equality of means is followed up by post hoc tests to determine which means are different if the null hypothesis is rejected. Similarly, the multivariate hypothesis test for the reference function coefficients is followed up by univariate regression tests to determine which voxels are active if the null hypothesis is rejected.

Figure 2: Marginal univariate (left) and multivariate (right) activations.



Marginal univariate and multivariate voxel t-statistic activations are computed according to Equations 2.10 and 3.11 and presented in Table 4. These marginal maps which are displayed in Figure 2 are thresholded at critical values $t_{1-\alpha/2,128-2-1} = 5.1465$ and $t_{1-\alpha/2,128-2-16} = 5.1830$ where $\alpha = 10^{-6}$. It is evident from these individual coefficient tests that the general activation in the ROI is in the pattern of a cross as by design except for a couple of negative activations for the independent t-statistics.

The above simulation procedure was replicated 10000 times to evaluate the distributions of the statistics. In Figure 3 histograms of the F-statistics for ROI activation are displayed assuming both dependent (full G) and independent (diagonal G) voxels. The mean and standard deviation of these histograms are $(34.3935, 5.5572)$ for the dependent voxels (left) and $(31.1633, 3.1888)$ for the independent voxels (right).

Further, t-statistics histograms of sixteen voxels (which are not presented here) were made. The means and standard deviations are given in Table 4.

Figure 3: Histograms of ROI F-statistics for dependent (left) and independent (right) voxels.

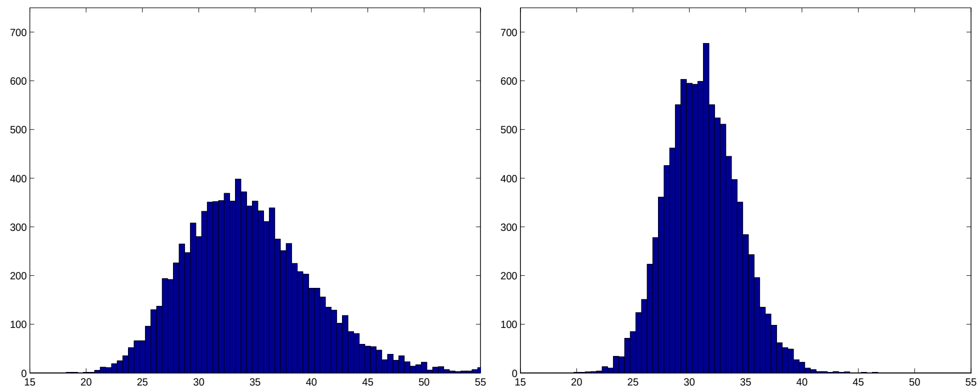


Table 1: Computed means and standard deviations for the marginal multivariate (top) and univariate (bottom) t-statistics.

$\bar{t}, \bar{\sigma}_t$	1	2	3	4
1	7.0776, 1.1159	1.4174, 1.0083	1.4040, 0.9997	7.0476, 1.1057
2	-4.2406, 1.0397	7.0732, 1.0915	7.0812, 1.0986	-4.2425, 1.0398
3	-4.2433, 1.0520	7.0762, 1.1071	7.0853, 1.0978	-4.2435, 1.0451
4	7.0769, 1.0917	1.4239, 1.0127	1.4260, 1.0219	7.0800, 1.1069
$\bar{t}, \bar{\sigma}_t$	1	2	3	4
1	6.6394, 1.0468	1.3297, 0.9458	1.3171, 0.9378	6.6113, 1.0372
2	-3.9780, 0.9753	6.6353, 1.0239	6.6427, 1.0306	-3.9799, 0.9754
3	-3.9806, 0.9869	6.6380, 1.0386	6.6466, 1.0298	-3.9807, 0.9804
4	6.6387, 1.0241	1.3357, 0.9500	1.3377, 0.9587	6.6417, 1.0383

The correlation matrix between the 10000 sets of t-statistics (from both methods) of length 16 was computed and resulted in the same 16×16 correlation structure as the random error from which the data was generated.

5 CONCLUSION

The multivariate regression model was presented with the univariate regression model as a special case. The matrix Student T distribution of the matrix of estimated regression coefficients was presented along with the multivariate Student t distributions of any row or colum in addition to the univariate Student t distribution of any element. Multivariate tests of hypothesis were presented for all the coefficients of a particular independent variable (for example a reference function). The multivariate test of hypothesis was used to determine whether a region of interest was activated or not. This multivariate test of hypothesis was followed up by univariate tests of hypothesis to determine the particular voxels which were active. These tests of hypothesis were illustrated on a simulated FMRI data set for a region of interest.

A APPENDIX

A.1 Univariate Likelihood Ratio

The likelihood ratio statistic is computed by maximizing the likelihood $p(Y_j|\beta_j, \sigma_j^2, X)$ with respect to β_j and σ_j^2 under the null and alternative hypotheses. Denote the maximized values under the null hypothesis by $(\tilde{\beta}_j, \tilde{\sigma}_j^2)$ and those under the alternative hypothesis as $(\hat{\beta}_j, \hat{\sigma}_j^2)$. These maximized values are then substituted into the likelihoods and the ratio

$$\lambda_j = \frac{p(Y_j|\tilde{\beta}_j, \tilde{\sigma}_j^2, X)}{p(Y_j|\hat{\beta}, \hat{\sigma}_j^2, X)} \tag{A.1}$$

$$= \frac{(2\pi)^{-\frac{n}{2}}(\tilde{\sigma}_j^2)^{-\frac{n}{2}} e^{-\frac{(Y_j - X\tilde{\beta}_j)'(Y_j - X\tilde{\beta}_j)}{2\tilde{\sigma}_j^2}}}{(2\pi)^{-\frac{n}{2}}(\hat{\sigma}_j^2)^{-\frac{n}{2}} e^{-\frac{(Y_j - X\hat{\beta}_j)'(Y_j - X\hat{\beta}_j)}{2\hat{\sigma}_j^2}}} \tag{A.2}$$

$$\lambda_j^{-\frac{2}{n}} = \frac{(Y_j - X\tilde{\beta}_j)'(Y_j - X\tilde{\beta}_j)}{(Y_j - X\hat{\beta}_j)'(Y_j - X\hat{\beta}_j)} \tag{A.3}$$

$$= \frac{g_j + (\hat{\beta}_j - \tilde{\beta}_j)'(X'X)(\hat{\beta}_j - \tilde{\beta}_j)}{g_j} \tag{A.4}$$

$$\left(\frac{n-q-1}{r}\right)(\lambda_j^{-\frac{2}{n}} - 1) = \frac{(C\hat{\beta}_j - \gamma_j)'[C(X'X)^{-1}C']^{-1}(C\hat{\beta}_j - \gamma_j)}{rg_j/(n-q-1)} \tag{A.5}$$

taken. In this, equivalence of the second term in the numerator of the next to last equation and the numerator of the last equation can be shown. The statistic $\left(\frac{n-q-1}{r}\right)(\lambda_j^{-\frac{2}{n}} - 1)$ follows an F-distribution with r and $n - q - 1$ degrees of freedom. If $C = (0, I)$, then the second term in the numerator can be shown to be equal to the sum of squares of the reduced model minus the sum of squares of the full model.

A.2 Multivariate Likelihood Ratio

Similarly as in the univariate case, likelihood ratio statistic is computed by maximizing the likelihood $p(Y|B,\Sigma,X)$ with respect to (B,Σ) under the two hypotheses to obtain null and alternative estimates $(\tilde{B},\tilde{\Sigma})$ and $(\hat{B},\hat{\Sigma})$. The likelihood ratio is

$$\lambda = \frac{p(Y|\tilde{B},\tilde{\Sigma},X)}{p(Y|\hat{B},\hat{\Sigma},X)} \tag{A.6}$$

$$= \frac{(2\pi)^{-\frac{np}{2}}|\tilde{\Sigma}|^{-\frac{n}{2}}e^{-\frac{1}{2}tr\tilde{\Sigma}^{-1}(Y-X\tilde{B}')(Y-X\tilde{B}')'}}{(2\pi)^{-\frac{np}{2}}|\hat{\Sigma}|^{-\frac{n}{2}}e^{-\frac{1}{2}tr\hat{\Sigma}^{-1}(Y-X\hat{B}')(Y-X\hat{B}')'}} \tag{A.7}$$

$$\lambda^{-\frac{2}{n}} = \frac{|(Y-X\tilde{B}')(Y-X\tilde{B}')'|}{|(Y-X\hat{B}')(Y-X\hat{B}')'|} \tag{A.8}$$

$$= |G+(\hat{B}-\tilde{B})(X'X)(\hat{B}-\tilde{B})'|/|G| \tag{A.9}$$

It should be noted that when $C=(0,I)$, $\Lambda=\lambda^{\frac{2}{n}}$ is distributed as Wilks' Lambda with the second term in the denominator being the sum of squares of the reduced model minus the sum of squares of the full model. Further, it can be shown that when $C=(0,I)$,

$$(\hat{B}-\tilde{B})(X'X)(\hat{B}-\tilde{B})' = (C\hat{B}'-\Gamma)'[C(X'X)^{-1}C']^{-1}(C\hat{B}'-\Gamma) \tag{A.10}$$

so that,

$$\lambda^{-\frac{2}{n}} = |G+(C\hat{B}'-\Gamma)'[C(X'X)^{-1}C']^{-1}(C\hat{B}'-\Gamma)|/|G|. \tag{A.11}$$

When $C=(0,\ldots,0,1,0,\ldots,0)$ is a row vector of zeros with a 1 in the k^{th} position, it can also be shown that

$$\left(\frac{n-q-p}{p}\right)(\lambda^{-\frac{2}{n}}-1) = \left(\frac{n-q-p}{p}\right)W_{kk}^{-1}(\hat{B}_k-\Gamma_k)'G^{-1}(\hat{B}_k-\Gamma_k) \tag{A.12}$$

follows an F-distribution with p numerator and $n-q-p$ denominator degrees of freedom.

Whenever it can be shown that Equation A.10 is true, the likelihood ratio statistic in Equation A.11 is true and can be used for significance evaluation. Additional hypotheses which we may wish to test include

$$C = (0,\ldots,0,I,-I,0,\ldots,0) \tag{A.13}$$

$$\text{and} \tag{A.14}$$

$$C = (0,\ldots,0,e',-e',0,\ldots,0) \tag{A.15}$$

where e is a column vector of ones.

Table 2: True (top) and estimated (bottom) regression coefficient matrix.

B'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0.2	0.7	0.4	0.3	0.9	0.4	0.5	0.2	0.9	0.1	0.5	0.1	0.6	0.4	0.4	0.8
1	0.5	0.1	0.9	0.2	0.6	0.8	0.3	0.7	0.1	0.3	0.5	0.6	0.4	0.2	0.5	0.9
2	5.0	1.0	1.0	5.0	-3.0	5.0	5.0	-3.0	-3.0	5.0	5.0	-3.0	5.0	1.0	1.0	5.0
$\hat{B}'$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	3.7	2.4	1.0	0.8	1.6	0.8	0.2	0.2	2.9	-0.9	0.3	-1.5	0.4	-0.5	-0.2	0.5
1	0.5	0.1	0.9	0.2	0.6	0.8	0.3	0.7	0.1	0.3	0.5	0.6	0.4	0.2	0.5	0.9
2	5.1	0.8	0.5	5.1	-2.6	5.0	4.6	-3.7	-2.8	4.3	4.4	-4.0	4.7	0.9	1.7	5.0

Table 3: ROI voxels covariance matrix.

Σ/σ^2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	ρ	0	0	ρ	0	0	0	0	0	0	0	0	0	0	0
2	ρ	1	ρ	0	0	ρ	0	0	0	0	0	0	0	0	0	0
3	0	ρ	1	ρ	0	0	ρ	0	0	0	0	0	0	0	0	0
4	0	0	ρ	1	0	0	0	ρ	0	0	0	0	0	0	0	0
5	ρ	0	0	0	1	ρ	0	0	ρ	0	0	0	0	0	0	0
6	0	ρ	0	0	ρ	1	ρ	0	0	ρ	0	0	0	0	0	0
7	0	0	ρ	0	0	ρ	1	ρ	0	0	ρ	0	0	0	0	0
8	0	0	0	ρ	0	0	ρ	1	0	0	0	ρ	0	0	0	0
9	0	0	0	0	ρ	0	0	0	1	ρ	0	0	ρ	0	0	0
10	0	0	0	0	0	ρ	0	0	ρ	1	ρ	0	0	ρ	0	0
11	0	0	0	0	0	0	ρ	0	0	ρ	1	ρ	0	0	ρ	0
12	0	0	0	0	0	0	0	ρ	0	0	ρ	1	0	0	0	ρ
13	0	0	0	0	0	0	0	0	ρ	0	0	0	1	ρ	0	0
14	0	0	0	0	0	0	0	0	0	ρ	0	0	ρ	1	ρ	0
15	0	0	0	0	0	0	0	0	0	0	ρ	0	0	ρ	1	ρ
16	0	0	0	0	0	0	0	0	0	0	0	ρ	0	0	ρ	1

Table 4: Marginal univariate (left) and multivariate (right) t-statistics.

t	1	2	3	4
1	6.6334, 6.2227	1.2038, 1.1292	0.6442, 0.6043	7.4922, 7.0283
2	-3.8636, -3.6244	7.2984, 6.8465	6.9349, 6.5055	-5.4322, -5.0958
3	-3.5956, -3.3730	5.6255, 5.2772	5.7558, 5.3994	-5.2393, -4.9149
4	6.3367, 5.9444	1.3136, 1.2322	2.4626, 2.3101	7.1131, 6.6727

Table 5: Estimated error covariance matrix.

$\hat{\Sigma}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	71.9	23.4	12.3	-7.6	11.0	-7.0	8.4	-2.6	3.9	-2.0	3.8	-0.5	-16.1	-9.9	4.4	-3.7
2	23.4	59.9	7.8	-5.6	-7.4	14.5	1.1	1.1	8.3	4.4	-0.9	2.5	-4.2	-7.5	-4.4	-7.3
3	12.3	7.8	66.7	19.6	4.6	0.2	13.1	3.2	10.7	0.0	1.6	10.3	7.7	-2.2	1.0	-6.6
4	-7.6	-5.6	19.6	57.1	0.8	1.7	-1.0	15.4	-1.8	-4.1	3.1	17.3	2.4	-6.9	-6.0	-6.8
5	11.0	-7.4	4.6	0.8	55.9	18.6	3.3	-2.6	19.5	5.4	5.4	-10.7	-0.2	-5.6	2.0	-3.9
6	-7.0	14.5	0.2	1.7	18.6	57.3	16.9	2.7	11.7	17.5	1.7	-5.2	-0.9	4.7	-2.4	3.8
7	8.4	1.1	13.1	-1.0	3.3	16.9	55.3	23.3	7.0	2.3	19.5	4.7	-5.7	-3.1	3.5	4.2
8	-2.6	1.1	3.2	15.4	-2.6	2.7	23.3	57.8	-1.9	3.3	10.8	18.7	2.8	-3.0	9.2	3.5
9	3.9	8.3	10.7	-1.8	19.5	11.7	7.0	-1.9	76.5	28.9	1.8	-10.4	13.3	-4.2	-3.3	-3.3
10	-2.0	4.4	0.0	-4.1	5.4	17.5	2.3	3.3	28.9	71.5	25.0	-2.7	0.9	11.4	4.4	-3.5
11	3.8	-0.9	1.6	3.1	5.4	1.7	19.5	10.8	1.8	25.0	73.0	16.1	4.5	1.8	17.7	-0.6
12	-0.5	2.5	10.3	17.3	-10.7	-5.2	4.7	18.7	-10.4	-2.7	16.1	71.7	-6.1	1.1	-1.2	14.0
13	-16.1	-4.2	7.7	2.4	-0.2	-0.9	-5.7	2.8	13.3	0.9	4.5	-6.1	67.3	14.4	1.7	-0.9
14	-9.9	-7.5	-2.2	-6.9	-5.6	4.7	-3.1	-3.0	-4.2	11.4	1.8	1.1	14.4	59.2	15.9	5.6
15	4.4	-4.4	1.0	-6.0	2.0	-2.4	3.5	9.2	-3.3	4.4	17.7	-1.2	1.7	15.9	58.9	22.5
16	-3.7	-7.3	-6.6	-6.8	-3.9	3.8	4.2	3.5	-3.3	-3.5	-0.6	14.0	-0.9	5.6	22.5	60.9

References

[1] Lee J. Bain and Max Engelhardt. *Introduction to Probability and Mathematical Statistics.* PWS-Kent Publishing Company, Boston, Massachusetts, USA, second edition, 1992.

[2] Peter Bandettini, Andrzej Jesmanowicz, Eric Wong, and James Hyde. Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, 30, 1993.

[3] Robert W. Cox. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29, 1996.

[4] Robert W. Cox, Andrzej Jesmanowicz, and James S. Hyde. Real-time functional magnetic resonance imaging. *Magnetic Resonance in Medicine*, 33, 1995.

[5] Arjun K. Gupta and D. K. Nagar. *Matrix Variate Distributions.* Chapman & Hall/CRC, Boca Raton, Florida, USA, 2000.

[6] David A. Harville. *Matrix Algebra From A Statisticians Perspective.* Springer-Verlag Inc., New York, NY, USA, 1997.

[7] Robert V. Hogg and Allen T. Craig. *Introduction to Mathematical Statistics.* Macmillan, New York, NY, USA, fourth edition, 1978.

[8] Raymond H. Myers. *Classical and Modern Regression with Applications.* Duxbury, Pacific Grove, CA, USA, 1990.

[9] Daniel B. Rowe. Computing FMRI activations: Coefficients and t-statistics by detrending and multiple regression. Technical Report 39, Division of Biostatistics, Medical College of Wisconsin, Milwaukee, WI 53226, 2002.

[10] Daniel B. Rowe. *Multivariate Bayesian Statistics.* CRC Press, Boca Raton, FL, USA, 2002.

[11] B. Douglas Ward. Deconvolution analysis of FMRI time series data. AFNI 3dDeconvolve Documentation, Medical College of Wisconsin, July 2000.