On Estimating the Parameters of the Complex fMRI Time Course Model

Daniel B. Rowe[1,2]
Department of Biophysics[1] and Division of Biostatistics[2]

Division of Biostatistics
Medical College of Wisconsin

Technical Report 46

May 2004

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

MEDICAL COLLEGE OF WISCONSIN

On Estimating the Parameters of the Complex fMRI Time Course Model

Daniel B. Rowe[1,2]*

Department of Biophysics[1] and Division of Biostatistics[2]

Medical College of Wisconsin

Milwaukee, WI USA

Abstract

[illegible]

*Corresponding Author: Daniel B. Rowe, Department of Biophysics, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, dbrowe@mcw.edu.

[illegible]

1 Introduction

Recently a complex time course model was introduced by Rowe and Logan [10] to determine functional brain activation. This model builds upon previous work [5, 7] in which pre magnitude complex valued voxel time courses were used to determine brain activation. This model showed improved power of detection at low signal-to-noise ratios and low contrast-to-noise ratios for three distinct thresholding methods [6].

Also recently, the Cramer-Rao lower bound (CRLB) of unbiased parameter estimators for a similar model in MRI was presented [11]. This complex MRI model required data for a single time point and pooled information from neighboring voxels. In contrast, the complex fMRI model uses repeated observations over time in each voxel where the magnitude of the response is allowed to vary according to a general linear model.

2 Models

A nonlinear multiple regression model with design matrix $X' = (x_1, \ldots, x_n)$ was introduced individually for each voxel [10] that includes a phase imperfection θ in which at time t

In fMRI, repeated measurements are tken over time while a subject is performing a task. In each voxel, we usually compute a measure of association between the observed time course and a preassigned reference function that characterizes the experimental paradigm. The typical method to compute activations [1, 2] is to use only the magnitude at time t denoted by r_t and written as

$$r_t = \left[(x_t'\beta \cos\theta + \eta_{Rt})^2 + (x_t'\beta \sin\theta + \eta_{It})^2\right]^{\frac{1}{2}} . \tag{2.2}$$

As previously outlined [10], the magnitude of a complex valued observation at time t is Ricean distributed [3, 8] and given by

$$p(r_t|x_t, \beta, \sigma^2) = \frac{r_t}{\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}\left[r_t^2 + (x_t'\beta)^2\right]\right\} \int_{\phi_t=-\pi}^{\pi} \frac{1}{2\pi} \exp\left\{\frac{x_t'\beta r_t}{\sigma^2}\cos(\varphi_t - \theta)\right\} d\phi_t . \tag{2.3}$$

where a general linear model is assumed and the integral factor often denoted $I_o(x_t'\beta r_t/\sigma^2)$ is the zeroth order modified Bessel function of the first kind. It is well known that for "large" SNR's the Ricean distribution of the magnitude r_t in Equation 2.3 is approximately normal with mean $x_t'\beta$ and variance σ^2. When the SNR is zero, the Ricean distribution is a Rayleigh distribution. It is intermediate SNR values that are of interest along with guidelines for what is considered a "large" SNR.

Maximum likelihood estimates (MLE's) of the parameters (β, σ^2) can be determined for both null and alternative hypotheses such as $H_0 : C\beta = \gamma$ versus $H_1 : C\beta \neq \gamma$ from

$$p(r|X, \beta, \sigma^2) = (\sigma^2)^{-n} \left(\prod_{t=1}^{n} r_t\right) \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{t=1}^{n} r_t^2 + \sum_{t=1}^{n} (x_t'\beta)^2\right)\right] \prod_{t=1}^{n} I_o(x_t'\beta r_t/\sigma^2) ,$$

or the logarithm of the likelihood

$$LL = -n\log(\sigma^2) + \sum_{t=1}^{n} \log r_t - \frac{1}{2\sigma^2}\left(\sum_{t=1}^{n} r_t^2 + \sum_{t=1}^{n} (x_t'\beta)^2\right) + \sum_{t=1}^{n} \log I_o(x_t'\beta r_t/\sigma^2)$$

where $r = (r_1, \ldots, r_n)'$. When maximizing the log likelihood under the null hypothesis, the Lagrange multiplier constraint $\psi'(C\beta - \gamma)$ is added. If the parameter estimates under the null hypothesis are denoted $(\tilde{\beta}, \tilde{\sigma}^2)$ and those under the alternative hypothesis $(\hat{\beta}, \hat{\sigma}^2)$, then substituting back into the likelihood and the ratio λ of null over alternative leads to the approximately χ^2 distributed statistic

$$-2\log\lambda \quad = \quad 2n\log(\tilde{\sigma}^2/\hat{\sigma}^2) + \frac{1}{\tilde{\sigma}^2}\left(\sum_{t=1}^{n} r_t^2 + \sum_{t=1}^{n}(r_t'\tilde{\beta})^2 \quad - \frac{1}{\hat{\sigma}^2}\left(\sum_{t=1}^{n} r_t^2 + \right.\right.$$

Complex

$$
\begin{aligned}
\hat{\theta}_C &= \frac{1}{2}\tan^{-1}\left[\frac{\hat{\beta}_R'(X'X)\hat{\beta}_I}{(\hat{\beta}_R'(X'X)\hat{\beta}_R-\hat{\beta}_I'(X'X)\hat{\beta}_I)/2}\right] \\
\hat{\beta}_C &= \hat{\beta}_R\cos\hat{\theta}_C+\hat{\beta}_I\sin\hat{\theta}_C \\
\hat{\sigma}_C^2 &= \frac{1}{2n}\begin{pmatrix} y_R - X\hat{\beta}\cos\hat{\theta}_C \\ y_I - X\hat{\beta}\sin\hat{\theta}_C\end{pmatrix}'\begin{pmatrix} y_R - X\hat{\beta}\cos\hat{\theta}_C \\ y_I - X\hat{\beta}\sin\hat{\theta}_C\end{pmatrix} \\
\tilde{\theta}_C &= \frac{1}{2}\tan^{-1}\left[\frac{\hat{\beta}_R'\Psi(X'X)\hat{\beta}_I}{(\hat{\beta}_R'\Psi(X'X)\hat{\beta}_R-\hat{\beta}_I'\Psi(X'X)\hat{\beta}_I)/2}\right] \\
\tilde{\beta}_C &= \Psi\left[\hat{\beta}_R\cos\tilde{\theta}_C+\hat{\beta}_I\sin\tilde{\theta}_C\right] \\
\tilde{\sigma}_C^2 &= \frac{1}{2n}\begin{pmatrix} y_R - X\tilde{\beta}\cos\tilde{\theta}_C \\ y_I - X\tilde{\beta}\sin\tilde{\theta}_C\end{pmatrix}'\begin{pmatrix} y_R - X\tilde{\beta}\cos\tilde{\theta}_C \\ y_I - X\tilde{\beta}\sin\tilde{\theta}_C\end{pmatrix}
\end{aligned}
\tag{3.3}
$$

where $\Psi = I_{q+1}-(X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}C$, $\hat{\beta}_R = (X'X)^{-1}X'y_R$, and $\hat{\beta}_I = (X'X)^{-1}X'y_I$, while y_R and y_I are the $n \times 1$ vectors of real and imaginary observations. The parameters for the approximate Taylor series model are iteratively estimated [9, 10].

Alternative hypothesis estimators will be evaluated in terms of the deviation of their mean and variance from the true value and their CRLB. An estimator τ of θ is said to be unbiased if $E(\tau) = \theta$ for all θ [4]. The deviation of the mean of an estimator from the true value is called the bias, $b(\tau) = E(\tau) - \theta$. However, since the CRLB is for unbiased parmaeter estimators, the alternative hypothesis estimators for the variance are multiplied by $n/(n-q-1)$ for the magnitude-only models and $2n/(2n-q-2)$ for the complex model.

As outlined in the Appendix, the CRLB for the magnitude-only normal activation model is found to be

$$
CRLB_N = \begin{array}{c} \\ \beta \\ \sigma^2 \end{array}\begin{array}{c} \begin{array}{cc} \quad\beta & \qquad\sigma^2 \end{array} \\ \begin{pmatrix} \sigma^2(X'X)^{-1} & 0 \\ 0 & 2\sigma^4/n \end{pmatrix}\end{array}, \tag{3.4}
$$

for the magnitude-only Taylor series approximation to the Ricean distribution model to be

$$
CRLB_T = \begin{array}{c} \\ \beta \\ \sigma^2 \end{array}\begin{array}{c} \begin{array}{cc} \qquad\qquad\beta & \qquad\qquad\sigma^2 \end{array} \\ \begin{pmatrix} \sigma^2(X'X)^{-1}\left[I - \frac{\sigma^2}{2}(X'X)^{-1}\sum_{t=1}^{n} x_t/(x_t'\beta)\right]^{-1} & 0 \\ 0 & 2\sigma^4/n \end{pmatrix}\end{array}, \tag{3.5}
$$

and for the complex activation model of Rowe and Logan (2004) it can also be found to be

$$CRLB_C = \begin{array}{c} \\ \beta \\ \sigma^2 \\ \theta \end{array} \begin{array}{c} \begin{array}{ccc} \beta & \sigma^2 & \theta \end{array} \\ \begin{pmatrix} \sigma^2(X'X)^{-1} & 0 & 0 \\ 0 & \sigma^4/n & 0 \\ 0 & 0 & \sigma^2/\beta'(X'X)\beta \end{pmatrix} \end{array} . \quad (3.6)$$

The mean and variance of the unbiased parameter estimates and activation statistics from the simulations were computed and plotted in Figures 1 and 2 for the various SNR's. Note that the mean and variance of the parameter estimates in the magnitude-only normal (in red) and the Taylor series (in cyan) models deviate from the true value as the SNR decreases while the parameter estimates in the complex model (in blue) achieve their correct value and remain fairly constant.

It can be seen in Figure 1 that the the estimated intercept coe cient or baseline is unbiased for the magnitude-only normal model to about an SNR of 7.5 and the magnitude-only Taylor model down to an SNR of about 2.5. Further, the for the reference function coe cient β_2 and the variance σ^4 in the magnitude-only normal model are unbiased only to an SNR of approximately 10 while they are unbiased for the Taylor approximate model to an SNR of about 5.

Also included are the true parameter values and the CRLB's for the magnitude-only normal, the magnitude-only Taylor approximation, and complex models, in green, yellow, and magenta, respectively in Figure 2. The CRLB's are not exactly attained even for large SNR's because the estimators are only asymptotically e cient. Also note that the error variance estimates are approximately twice as large for the magnitude-only models than for the complex model, even for the largest value of SNR. This factor of two disparity is as stated by the CRLB's.

Additionally, the mean and variance of the $-2\log\lambda$ activation statistics for testing the hypothesis that $\beta_2 = 0$ are included in Figures 1 and 2 respectively. The mean and variance of the activation statistics are uniformly lower for the magnitude-only models than for the complex model of Rowe and Logan with a disparity that levels o as the SNR increases.

Figure 1: Plot of estimated parameter means with varying SNR.

5 Conclusions

A previously proposed complex data fMRI model of Rowe and Logan (2004) as an alternative to the typical magnitude-only normal data model was outlined along with a model that uses a Taylor series approximation in the Ricean distribution. Maximum likelihood parameter estimates were also described. The CRLB for the variance of the observation variance was half as large in the complex data model as it is in the magnitude-only models.

Simulations were performed for several SNR's and the parameters for both models estimated along with an activation statistic. The mean and variance of the estimated parameter values and activation statistics were computed and compared with the true values and CRLB's where applicable.

It was found that the complex model performed extremely well at estimating the true parameter values and achieving its CRLB's even for very low SNR. The magnitude-only models did not perform as well as the complex model. Additionally, even for very large SNR's, the variance of the error variance was twice as large for the magnitude-only models as for the complex model. These results indicate that using the complex data model instead of approximations to the Ricean distribution of the magnitude-only data are more useful at low SNR.

A MLE's for Taylor Model

The logarithm of the likelihood for the Taylor series approximation of the Ricean distribution of magnitude-only observations is

$$LL = -\frac{n}{2}\log(2\pi) + \frac{1}{2}\sum_{t=1}^{n}\log r_t - \frac{n}{2}\log(\sigma^2) - \frac{n}{2}\sum_{t=1}^{n}\log(x_t'\beta) - \frac{1}{2\sigma^2}(r - X\beta)'(r - X\beta) \quad (A.1)$$

Unrestricted MLE's

Maximizing the likelihood in Equation A.1 with respect to the parameters is the same as maximizing the logarithm of the likelihood with respect to the parameters and yields

$$\begin{aligned}\frac{\partial LL}{\partial \beta} &= -\frac{1}{2\sigma^2}[2(X'X)\beta - 2X'r] - \frac{1}{2}\sum_{t=1}^{n}\frac{1}{x_t'\beta}x_t \\ \frac{\partial LL}{\partial \sigma^2} & \quad -\frac{2n}{2}\frac{1}{\sigma^2} + \frac{h}{2}\frac{1}{(\sigma^2)^2}\end{aligned}$$

where $h = (r - X\beta)'(r - X\beta)$. By setting these derivatives equal to zero and solving, we get the MLE's under the unrestricted model given in Equation 3.2.

Restricted MLE's

Maximizing this likelihood with respect to the parameters is the same asRestr(E91(observ)60(ations)-3

B Cramer-Rao Lower Bounds

The CRLB for the variance of an unbiased estimate of a model parameter requires the second derivatives of the log likelihoods LL with respect to the model parameters.

B.1 Normal Model

The second derivatives of the log likelihood are

$$
\begin{aligned}
H(1,1) = \frac{\partial^2 LL}{\partial\beta\partial\beta'} &= -\frac{1}{2\sigma^2}[2(X'X)] \\
H(2,2) = \frac{\partial^2 LL}{\partial(\sigma^2)^2} &= -\frac{n}{2}(-1)(\sigma^2)^{-2} - 2\frac{h}{2}(\sigma^2)^{-3} \\
H(1,2) = \frac{\partial^2 LL}{\partial\beta\partial\sigma^2} &= [X'r - (X'X)\beta](-1)(\sigma^2)^{-2} \; .
\end{aligned}
$$

The symmetric Hessian matrix H is formed from the second derivatives and it is seen to be negative definite and therefore the estimated values from the first derivatives are maxima and not minima. The Fisher information matrix I is -$E(H|\beta,\sigma^2)$, that is, the expectation of the Hessian matrix with respect to r. The CRLB is the inverse of the Fisher information matrix. By taking expectations of the block elements of the Hessian matrix, the CRLB in Equation 3.4 is found.

B.2 Taylor Model

The second derivatives of the log likelihood are

$$
\begin{aligned}
H(1,1) = \frac{\partial^2 LL}{\partial\beta\partial\beta'} &= -\frac{1}{2\sigma^2}[2(X'X)] - \frac{1}{2}\sum_{t=1}^{n} x_t(-1)(x_t'\beta)^{-2}x_t' \\
H(2,2) = \frac{\partial^2 LL}{\partial(\sigma^2)^2} &= -\frac{n}{2}(-1)(\sigma^2)^{-2} - 2\frac{h}{2}(\sigma^2)^{-3} \\
H(1,2) = \frac{\partial^2 LL}{\partial\beta\partial\sigma^2} &= [X'r - (X'X)\beta](-1)(\sigma^2)^{-2} \; .
\end{aligned}
$$

The symmetric Hessian matrix H is formed from the second derivatives and it is seen to be negative definite and therefore the estimated values from the first derivatives are maxima and not minima. The Fisher information matrix I is -$E(H|\beta,\sigma^2)$, that is, the expectation of the Hessian matrix with respect to r. The CRLB is the inverse of the Fisher information matrix. By taking expectations of the block elements of the Hessian matrix, the CRLB in Equation 3.5 is found.

B.3 Complex Model

The second derivatives of the log likelihood are

$$H(1,1) = \frac{\partial^2 LL}{\partial\beta\partial\beta'} = -\frac{1}{2\sigma^2}[2(X'X)]$$

$$H(2,2) = \frac{\partial^2 LL}{\partial\theta^2}$$

References