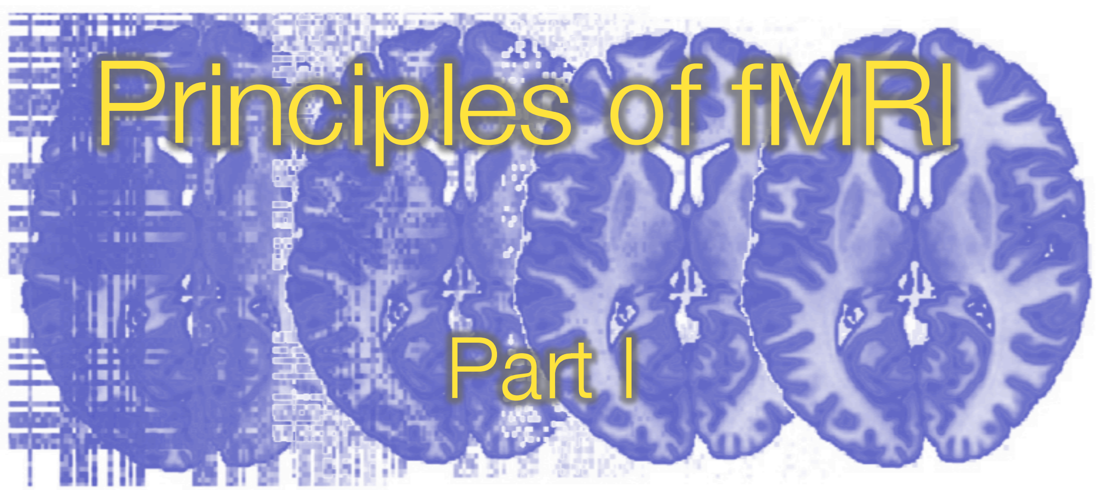


Principles of fMRI

Part I

Martin Lindquist
Department of Biostatistics
Johns Hopkins
Bloomberg School of Public Health

Tor Wager
Department of Psychology and
Neuroscience and the
Institute for Cognitive Science
University of Colorado, Boulder

@fMRIstats

Noise Models

GLM

A standard GLM can be written:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{V})$$

where

$$\underbrace{\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}}_{\text{fMRI Data}} = \underbrace{\begin{bmatrix} 1 & X_{11} & \cdots & X_{1p} \\ 1 & X_{21} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \cdots & X_{np} \end{bmatrix}}_{\text{Design matrix}} \times \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}}_{\text{Regression coefficients}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}}_{\text{Noise}}$$

V is the covariance matrix whose format depends on the noise model.

fMRI Noise

- Functional MRI data typically exhibit significant autocorrelation.
  - Caused by physiological noise and low frequency drift, that has not been appropriately modeled.
  - Typically modeled using either an AR(p) or an ARMA(1,1) process.
  - Single subject statistics are not valid without an accurate model of the noise.

AR(1) model

- Serial correlation can be modeled using a first-order autoregressive model, i.e.

$$\varepsilon_t = \phi\varepsilon_{t-1} + u_t \qquad u_t \sim N(0,\sigma^2)$$

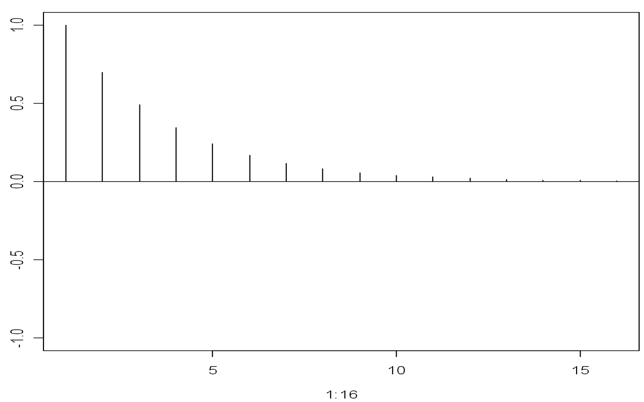
- The error term ε_t depends on the previous error term ε_{t-1} and a new disturbance term u_t.

AR(1) model

- The autocorrelation function (ACF) for an AR(1) process at lag *h*:

$$\rho(h) = \begin{cases} 1, & \text{if } h = 0, \\ \phi^{|h|} & \text{if } h \neq 0 \end{cases}$$

$\phi=0.7$



Error Term

- The format of $\mathbf{V}$ will depend on what noise model is used.

IID Case

$$
\mathbf{V} \propto \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}
$$

AR(1) Case

$$
\mathbf{V} \propto \begin{bmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{n-1} \\ \phi & 1 & \phi & \cdots & \phi^{n-2} \\ \phi^2 & \phi & 1 & \cdots & \phi^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \cdots & 1 \end{bmatrix}
$$

GLM Summary

model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \longrightarrow$$

estimate

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$$

fitted values

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$$

residuals

$$\begin{aligned}\mathbf{r} &= \mathbf{Y} - \hat{\mathbf{Y}} \\ &= (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1})\mathbf{Y} \\ &= \mathbf{R}\mathbf{Y}\end{aligned}$$

Estimating V

- In general the form of the covariance matrix is unknown, which means it has to be estimated.

- Estimating $\mathbf{V}$ depends on β's, and estimating β's depends on $\mathbf{V}$. Need iterative procedure.

- Methods for estimating variance components:
  - Method of moments
  - Maximum likelihood
  - Restricted maximum likelihood

Iterative Procedure

1. Assume that $\mathbf{V}=\mathbf{I}$ and calculate the OLS solution.

2. Estimate the parameters of $\mathbf{V}$ using the residuals.

3. Re-estimate the β values using the estimated covariance matrix $\hat{\mathbf{V}}$ from step 2.

4. Iterate until convergence.

Yule-Walker Estimates

- Assume ε_t is an AR(1) process, i.e.

$$\varepsilon_t = \phi \varepsilon_{t-1} + u_t \qquad t = 0, \pm 1, \ldots$$

where $u_t \sim WN(0,\sigma^2)$

- The Yule-Walker estimates are:

$$\hat{\phi} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} \qquad \hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}\hat{\gamma}(1)$$

Auto Covariance Function

MLE

- Maximum likelihood estimators (MLEs) are obtained by maximizing the log-likelihood:

$$l^{*}(\lambda) = -\frac{1}{2}\log\left(|\mathbf{V}|\right) - \frac{1}{2}\left(\mathbf{Y} - \mathbf{X}\hat{\beta}\right)^{T}\mathbf{V}^{-1}\left(\mathbf{Y} - \mathbf{X}\hat{\beta}\right)$$

where λ are parameters associated with $\mathbf{V}$.

ReML

- Restricted maximum likelihood (ReML) requires maximizing the restricted log-likelihood:

$$l^*(\lambda) = -\frac{1}{2}\log\left(|\mathbf{V}|\right) \boxed{-\frac{1}{2}\log\left(|\mathbf{X}^T\mathbf{V}\mathbf{X}|\right)} -\frac{1}{2}\left(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)^T\mathbf{V}^{-1}\left(\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right)$$

Extra ReML variance term

where λ are parameters associated with V.

ML vs ReML

- Maximum Likelihood
  - Maximize likelihood of data *y*
  - Used to estimate "mean" parameters β
  - But can produce biased estimates of variance

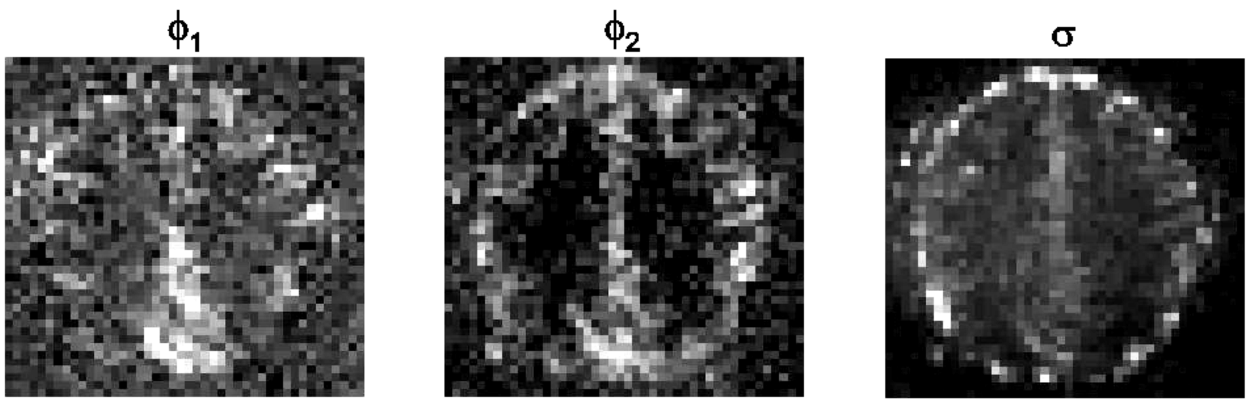
$$\hat{\sigma}^2_{\mathrm{ML}} = \frac{1}{n}\sum (y_i - \bar{y})^2$$

- Restricted Maximum Likelihood
  - Maximize likelihood of residuals $r = y - Xb$
  - Used to estimate variance parameters
  - Provides unbiased estimates

$$\hat{\sigma}^2_{\mathrm{ReML}} = \frac{1}{n-1}\sum (y_i - \bar{y})^2$$

Spatio-temporal Behavior

- The spatiotemporal behavior of these noise processes is complex.



Spatial maps of the model parameters from an AR(2) model estimated for each voxel's noise data.

End of Module

@fMRIstats