Stimulus Identification from fMRI scans

Charles Zheng and Yuval Benjamini

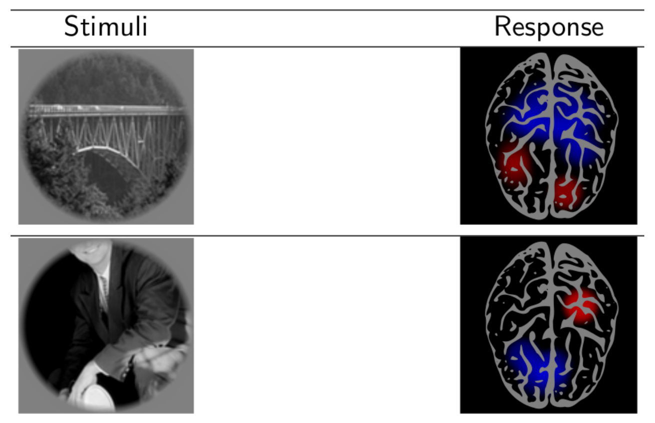
Stanford University

Setting

- Sequence of stimuli (pictures) shown at time $t = 1, \dots, T$
- Record subject's multivariate response $Y_t \in \mathbb{R}^p$
- Stimuli represented as *feature vector* $X_t \in \mathbb{R}^q$
- Linear model:

$$Y_{T \times p} = X_{T \times q} B_{q \times p} + E_{T \times p}$$

- E.g. Kay (2008)

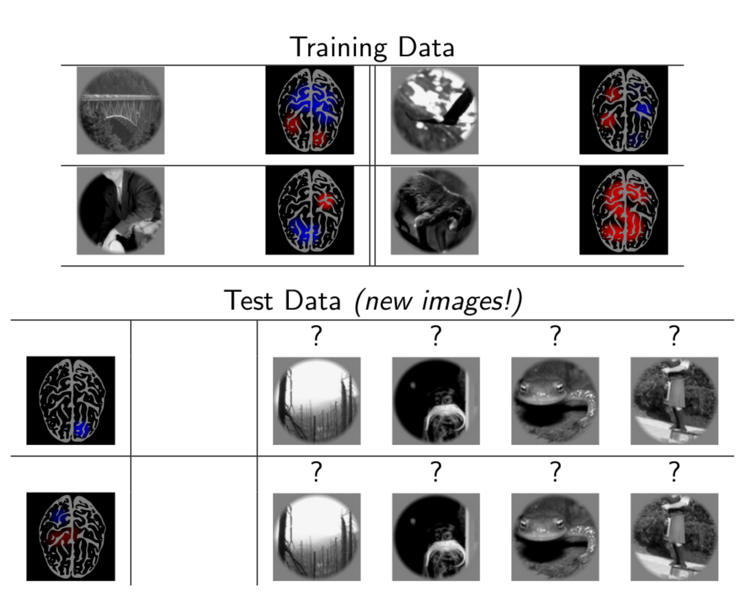


Identification

- Introduced in Kay (2008)
- Supervised learning task, validates the power of the linear model $Y = XB + E$
- Let S be a set of *new* stimuli (not in the training set) with features

$$\{x_1^{te}, \dots, x_\ell^{te}\}$$

- Scientist picks a stimulus i^* from S and measures the subject's reponse y^*
- Can the statistician *identify* the stimulus from y^*?



Objectives

- Develop new methodology for optimal identification

Key theoretical problems:

- Optimal identification given parameter estimates
- Estimation of model (e.g. linear model $Y = XB$)
- Estimation of noise $\Sigma_E = \text{Cov}(E)$

Optimality criteria:

- Bayes (average case) under parametric models
- Minimax regret under nonparametric models

We consider two methods: *maximum likelihood* and *empirical Bayes*, and compare to *Bayes risk*.

Maximum Likelihood

Procedure for identification of y^*, variants used in Kay (2008), Vu (2011)

- Obtain point estimates of coefficients B and noise covariance Σ_E
- E.g. B estimated using elastic net with CV (Zou 2005), shrinkage estimate for covariance

$$\hat{\Sigma}_E = \frac{1}{2}\hat{\text{Cov}}(Y - \hat{Y}) + \frac{1}{2}\text{diag}(\hat{\text{Cov}}(Y - \hat{Y}))$$

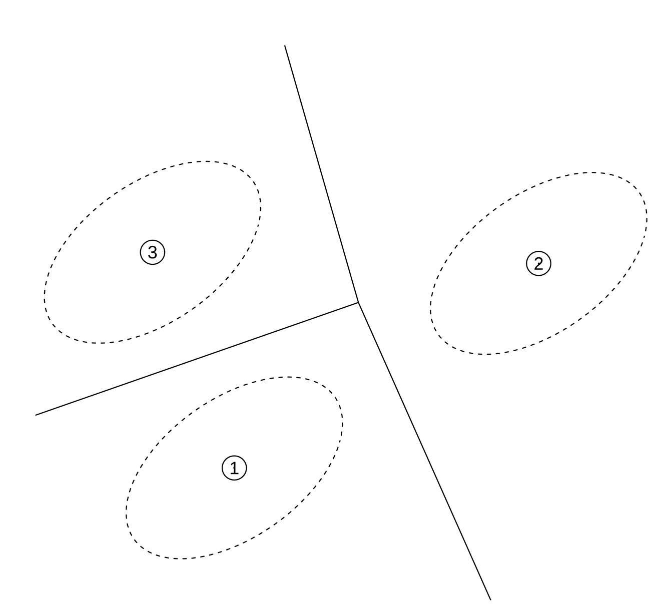
  where $\hat{\text{Cov}}$ denotes sample covariance
- Obtain predicted means for test stimuli

$$\hat{\mu}_i^{te} = (x_i^{te})^T B$$

- Identify the stimulus i^* by

$$i^* = \text{argmin}_i (\hat{\mu}_i^{te} - y^*)^T \hat{\Sigma}_E^{-1} (\hat{\mu}_i^{te} - y^*)$$

- Results in *linear decision boundaries*



What is Empirical Bayes?

- General approach for statistical problems (e.g. Efron 2004)
- Start with a hierarchical Bayes model with hyperparameters
- *Estimate* hyperparameters from data, e.g. maximizing marginal likelihood

Empirical Bayes

Model

- Noise $E_t \sim N(0, \Sigma_E)$ iid
- Coefficients $B_i \sim N(0, \sigma_i^2 I)$ for $i = 1, \dots, p$
- X non-random

Estimate hyperparameters

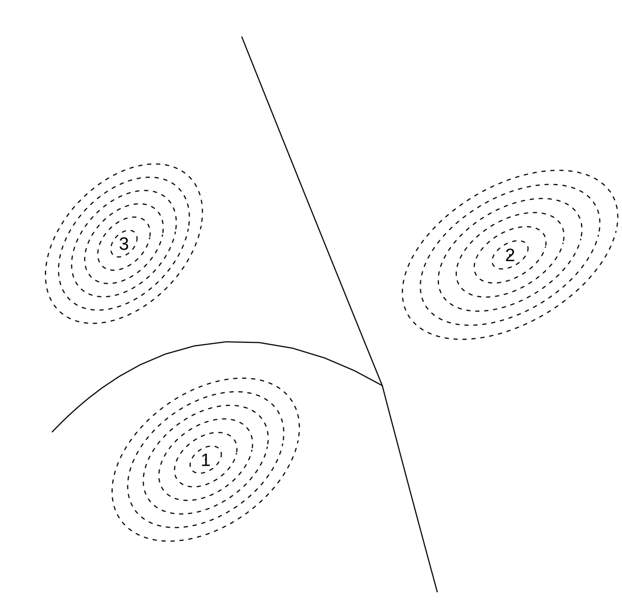
- Use *eigenprism* (Janson 2015) to estimate $\theta_i^2 = ||B_i||^2$ for $i = 1, \dots, p$
- Set $\sigma_i^2 = \hat{\theta}_i^2 / q$
- Estimate $\hat{B}$ as posterior mean given estimated σ_i^2
- Estimate Σ_E using residuals (same as in Maximum Likelihood)

Compute posterior

- Closed-form expressions for posteriors of B, μ_i^{te}
- Computational bottleneck: inverting the $pq \times pq$ covariance matrix of $\vec{B}$

Apply Bayes rule

- *Uncertainty* in B is reflected as *added noise*
- Result: posterior $\text{Cov}(y^* | i^*)$ varies, hence *quadratic boundaries*



Simulation Results

Ongoing Work

- Estimate noise covariance based on correlation structure of fMRI data (e.g. spatial correlation)
- Apply methods to data of Kay (2008)

References

Acknowledgements

Nam mollis tristique neque eu luctus. Suspendisse rutrum congue nisi sed convallis. Aenean id neque dolor. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.

Contact Information

- Web: http://www.university.edu/smithlab
- Email: john@smith.com
- Phone: +1 (000) 111 1111