creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

문학석사학위논문

Event-related Potential analysis using elastic net logistic regression

엘라스틱 넷 로지스틱 회귀분석을 통한 사건 관련 전위 분석

2016 년 2 월

서울대학교 대학원
협동과정 인지과학 전공
이 우 종

Abstract

The objective of the thesis is to explore whether regularization techniques can be applied to ERP analysis, and which type of regularization is adequate. This thesis proposes elastic net regularization logistic regression as a good candidate of data analytic method for Event-Related Potential analysis (ERP). Specifically, regularization techniques are used to identify latency in ERP. Study 1 tested whether regularization logistic regression can classify latency using simulated ERP data. It showed that ridge and lasso could identify latency information. In study 2, the same analyses were applied to actual ERP data. Ridge regression can identify latency information wheras lasso cannot.

Keywords: Elastic net, Regularization, Event-related potential, ERP, Logistic regression

Student Number: 2013-22803

Contents

List of Figures

Chapter 1. Introduction

Nowadays, functional brain imaging techniques are used in many areas: cognitive sciences, including psychology, neuroscience, marketing, medicine and so on.

In business marketing, for instance, the researchers want to know what can be chosen from customers, more generally speaking, what the customers' preferences are. However, through traditional questionnaires and interview techniques, the customers' true preferences cannot be extracted. On the other side, it is believed to be that brains have people's inner-information, so marketers are interested in those neuro imaging techniques (Ariely & Berns, 2010). In cognitive sciences, what cognition is and how people think are the most interesting questions. Using functional brain imaging techniques, researchers measure brain activity during specific cognitive tasks, and they infer cognitive functions related to specific brain activity. Electroencephalogram (EEG) is one of the most popular methods. Especially when timing issues are important, EEG is more prefered to functional Magnetic Resonance Imaging (fMRI) or Positron Emission Tomography (PET), which have good spatial resolution, but not temporal resolution. EEG technique has some benefits compared to other brain imaging techniques. First, temporal resolution is better than fMRI or PET. This can be advantageous for measuring attention or other cognitive functions that

are characterized as maintaining short period time and lasting very briefly. Secend, EEG equipment is relatively cheap (Luck, 2005), which means that experiments using EEG, compared to fMRI or PET, can easily conducted. Using EEG, Event-Related Potential analysis can be applied. Event-related potential is, namely, a potential related to specific events. These events are combined to cognitive events. This anlaysis implies that the EEG measurements are the neural activities which are connected with specific perceptual, cognitive, and motor functions (Luck, 2005). In this context, we can assume that the cognitive functions can be extracted through analysing electical potentials.

Since Walter et al. (1964), ERP analysis was widely used in cognitive experiments. For example, Luck et al. (2000) investigated ERP components related to attention. In Luck et al. (1993), an attention component, known as N2pc, was discovered and reported. The other cognitive function, memory, was also studied. Rugg and Curran (2007) and D. Friedman and Johnson (2000) scrutinized recognition memory and memory encoding respectively. Language components are one of well-known ERP components. When sentences are semantically asynchronous, then negatively-going evoked potential, which is now called an N400 component, was identified (Kutas & Hillyard, 1980; Kutas & Federmeier, 2000). Otherwise, a P600 component is related to syntactical congruency. The P600 component, namely, is positive peak on about 600ms (Osterhout & Holcomb, 1992; Coulson, King, & Kutas, 1998).

As aforementioned, EEG has a good temporal characteristic. Not surprisingly, ERP time courses have two characteristics: latency and amplitude. In ERP analysis, it is important to know when the latency is, and how strong the amplitude is. In this respect, what cognitive functions contribute to ERP pattern and how strongly they affect those ERP waves are interesting questions. These latency and amplitude can be identified using regression analysis, which

can extract important timing related to specific tasks. Those information could be used as an index between task conditions.

1.1 Regression

1.1.1 Linear regression

We can consider simple linear regression model as follows:

$$y = a + bx + \varepsilon \tag{1.1}$$

In above formula, $\hat{a}, \hat{b}$ can be obtained by minimizing $\varepsilon^2, i.e., (y - a - bx)^2$ (OLS approach). Expanding equation 1.1, we can think that independent variables are more than two variables, i.e., x is two or more, not one.

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon. \tag{1.2}$$

that is,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\mathbf{Y} \quad = \quad \mathbf{X} \quad \beta \quad + \quad \varepsilon$$

where p means the number of independent variables, n means the number of observations, y is a dependent variable, β is regression coefficients, ε indicates

random errors. Likewise simple regression, $\hat{\mathbf{Y}} = \mathbf{X}\beta$, it can be also obtained by minimizing ε^2, that is, $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$

1.1.2 Logistic regression

Linear regression (which is called general linear model) is about continuous dependent variables. However, when the dependent variables are binary, general linear model cannot be applied. In that case, we can think of logistic regression model. The model can be described as follows:

$$\begin{aligned} logit(p) &= ln\frac{p}{1-p} \\ \mathbf{y} &= \pi(\mathbf{X}\beta) + \varepsilon \\ \pi(\mathbf{x}) &= \frac{1}{1+e^x} \\ \mathbf{y} &= \frac{1}{1+e^{-}\mathbf{X}\beta} + \varepsilon \end{aligned} \tag{1.3}$$

Different from linear regression, logistic regression cannot be estimated using ordinary least square methods. Estimation can be carried out through maximizing likelihood, which can be as follows:

$$\mathcal{L}(\mathbf{X}, y, \beta) = \prod_{i=1} \pi(\mathbf{X}\beta)^{y_i}(1-\pi(\mathbf{X}\beta))^{1-y_i} \tag{1.4}$$

In EEG data context, $\mathbf{X}$ is the matrix of time points measured during experiment, β is the coefficient vector for time points, $\mathbf{y}$ is the vector of experimental conditions. In this thesis, the question is that ERP latency can be extracted through regression analysis, and that latency can be used as index for cognitive function.

1.1.3 Regularization methods

Regularization methods, which are also known as penalization methods, give some penalty to model complexity . It can be useful for two reasons: interpreta-

tion and prediction. The most interpretable models can be accepted in scientific society. Also, predictability is about how well the model predicts future data, and considering predictability implies that dealing with overfitting problems, which are frequently occured as model complexity raises.

Ridge

Hoerl and Kennard (1970) proposed ridge regression as follows:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}(\|y - X\beta\|^2 + \lambda\|\beta\|^2) \tag{1.5}$$

Ridge regression minimizes the sum of squares of the coefficients subject to a bound on the $\ell 2$-norm, which can be constraints to the model.

Lasso

Unlike ridge estimation, lasso penalizes coefficients as zero (Tibshirani, 1996) using $\ell 1$-norm.

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}(\|y - X\beta\|^2 + \lambda\|\beta\|_1) \tag{1.6}$$

As equation (1.6) showed, the absoulte value of coefficients can be less than a constant. This penalty tends to make the coefficients be zero.

Elastic net

Elastic net is a combination of $\ell 1$ and $\ell 2$ penalty (Zou & Hastie, 2005). That is, this method combines ridge and lasso regularization.

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}}(\|y - \mathbf{X}\beta\|^2 + \lambda_2\|\beta\|^2 + \lambda_1\|\beta\|_1) \tag{1.7}$$

let $\alpha = \frac{\lambda_2}{(\lambda_2+\lambda_1)}$, then

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{X}\beta\|^2, \text{subject to } (1-\alpha)\|\beta\|_1 + \alpha\|\beta\|^2 \tag{1.8}$$

Elastic net regularization minimizes the residuals sum of squares subject to a bound on the $\ell 2$-norm and $\ell 1$-norm of the coefficients. This function $(1-\alpha)\|\beta\|_1 + \alpha\|\beta\|^2$ is called elastic net penalty. When α is 1, elastic net penalty is equal to $\ell 2$-norm (ridge), in contrast, when α becomes 0, then elastic net is the same as $\ell 1$-norm (lasso).

So far, some regression techniques were reviewed. Now, let's focus on how those analyses can be applied to EEG data, and ERP analysis. The main question of this thesis is whether logistic regression can be applied for identifying ERP latency, and regularization technique can be useful for those procedures.

Chapter 2. Study 1

Study 1 is about analysis of simulated EEG data. The question is whether elastic net logistic regression can indetify specific time point for the experimental conditions. For this question to be solved, the exact latency information (i.e., when the latency is) should be identified. Using simulation data, ERP latency can be manipulated, and the regression results will be compared with each other.

2.1 Simulation

EEG epoch data were generated based on Yeung, Bogacz, Holroyd, and Cohen (2004). The codes used in simulation were uploaded on the website[1]. For the convenience of analysis, only one phasic peak was used, and epochs were simulated by adding the peak onto the uncorrelated random noises. Target peak amplitude was 6 μV. Sampling rates were 1024Hz, epoch interval was 1000 milliseconds. The number of trials was 80 respectively, and two experimental conditions were used. Target peak latency was 700 milliseconds between the experimental conditions. Following plot is averaging of simulated ERP data.

[1] http://www.cs.bris.ac.uk/home/rafal/phasereset/

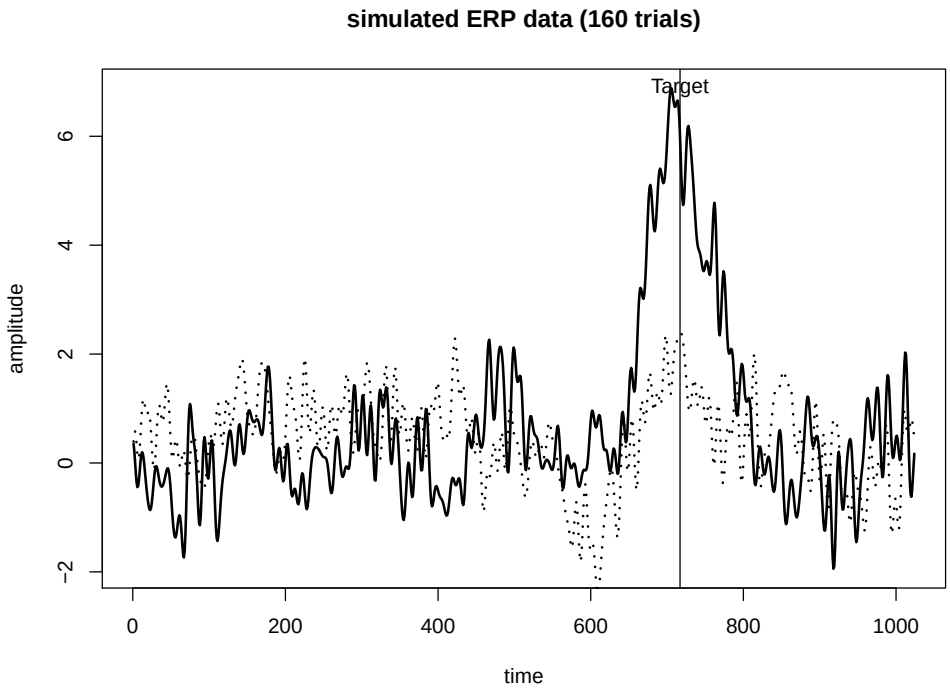


Figure 2.1: Mean of simulated ERP

2.2 Results and discussion

Several logistic regression analyses were applied to simulation data described above. Analyses were carried out using R version 3.2.3 (R Core Team, 2015) and R pacakge "glmnet" (J. Friedman, Hastie, & Tibshirani, 2010). λ (which is regularization parameter) was estimated through 10-fold cross-validation. As figure 2.2 showed, logistic regression results were summed up. The top of the figure is an usual logistic regression result, the middle is a ridge logistic regression result and the bottom shows a lasso logistic regression result. The solid line indicates the target ERP wave, and the dotted-line means the control ERP wave. The x-axis is time dimension and the y-axis indicates the amplitude of

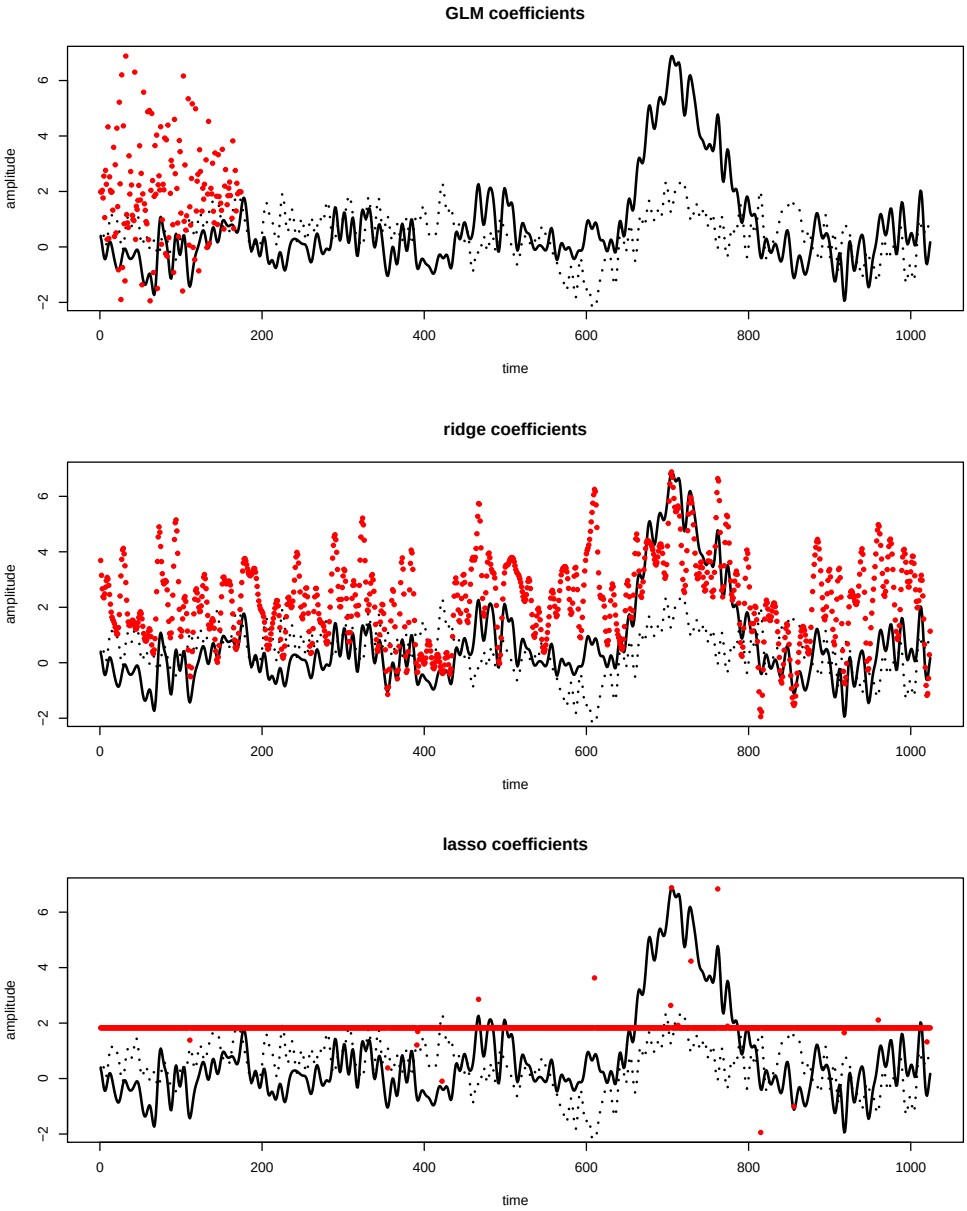


Figure 2.2: Comparison of logistic regression results

ERP waves. In each part of the figure, the dots mean the coefficients of logistic regression. Those coefficients are relative values (not absolute), therefore,

the dots indicate relative influence of the coefficients. In usual logistic regression, coefficients were not fully estimated. Because of high-dimensionality, the model cannot be solved using maximum likelihood estimation. The coefficients were estimated in the only number of observations, and the other independant variables were excluded. In contrast, ridge regularization classified all the time points. The maximum coefficient is 689ms. Otherwise, Lasso regularization estimated 18 coefficients and the others were all penalized as zero. Ridge and lasso could estimate important coefficients for speicific experimental condition. Especially, lasso was adequate parsimoniously. Only 18 time points were estimated, and those coefficients can be used as a marker for ERP latency.

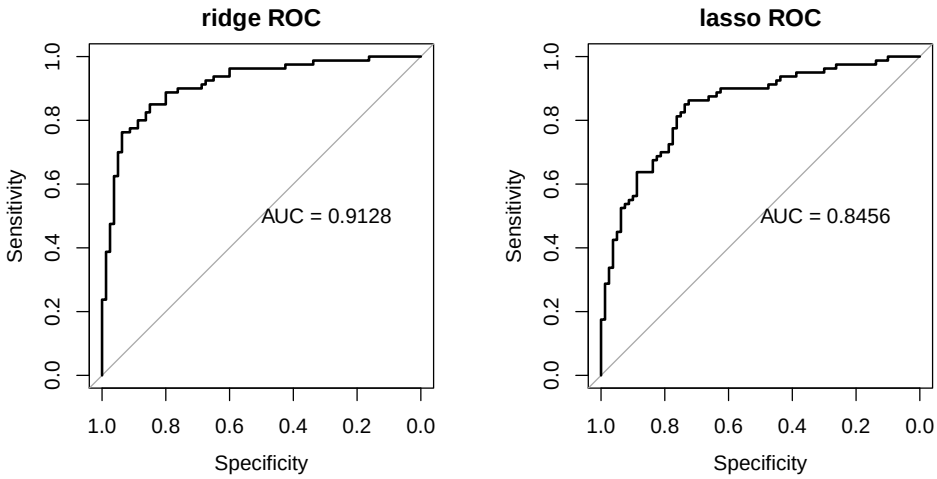


Figure 2.3: Simulated Ridge and Lasso ROC curves comparison

Figure 2.3 is ROC curve for ridge and lasso regression results. Ridge is more predictable than lasso.

Chapter 3. Study 2

In the previous study, elastic net regularized regression methods successfully classified ERP waves. For now, study 2 will investigate if previous results can be applied to human EEG data. Therefore, participants who received iconic memory task were included in study 2, and those data were analyzed using elastic net regression for the ERP pattern classification.

3.1 EEG acquisition

Data were acquired using Braintronics Brainbox EEG-1042 amplifier (Braintronics, Almere, The Netherlands). Reference channels were left and right mastoids. Sampling rates were 1024Hz, notch filter was off for minimizing a signal loss. Target electrode sites were 13 channels (Fp1, Fp2, Fz, F3, F4, FC3, FC4, Cz, Pz, P3, P4, Poz, Oz). Each was recorded according to the 10% positioning system (Nuwer et al., 1998).

3.2 Experiment design

For measuring EEG signals and analysing ERP waves, participants were asked to perform a specific cognitive task. Iconic memory has been known as a very breifly maintaining store of visual information, so the iconic memory task was suitable for ERP analysis. Within-subject design was chosen for minimizing individual differences, so all the subjects received every experimental conditions.

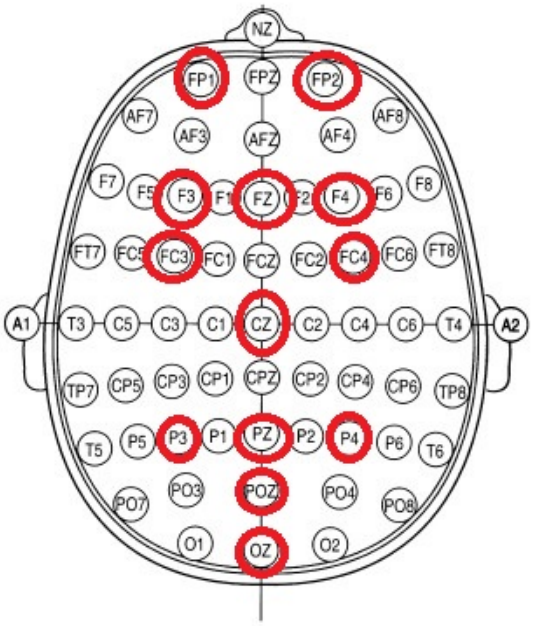
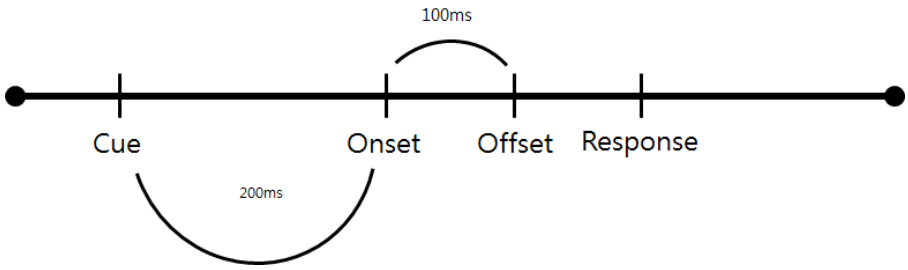
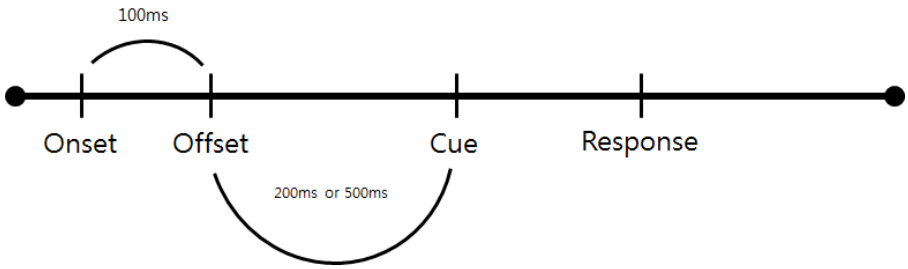


Figure 3.1: EEG electrode channels used in experiment (10% system)

The experimental conditions consisted mainly of two factors. One was a stimulus reporting type, and the other was a cue type. The stimulus reporting type was 2-level. Color and Number stimulus reporting types were required to report. When the color reporting condition, participants were asked to report what the color was, and when number reporting condition, then asked to report what the number itself (digit) was. The cue type was 3 levels, which was pre-cue 200ms, post-cue 200ms, and post-cue 500ms respectively. Pre-cue meant the cue appeared before the stimuli were presented, while the post-cue was that the stimuli were presented before the cue appeared. Each condition was 80 trials. The number of total trials was 480 trials, the experiment was splited into 4 sessions for minimizing participants' fatigue.



(a) Pre-cue condition



(b) Post-cue condition

Figure 3.2: Overview of experimental design

3.2.1 Stimuli

Stimuli used in the experiment were a single array of eight items. Each item was a colored number. Numbers consisted of 2, 4, 7 and 8, which were chosen for decreasing similarities among digit shape. And red, green, blue, and yellow were chosen for color stimuli.

Figure 3.3 is an example of iconic memory array used in the experiment. Stimuli were presented on 8-orientation, each was 45-degree interval, and target locations were randomly assigned. Stimuli were presented with Matlab 2013a Psychophysics Toolbox extensions 3.0.11 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007)

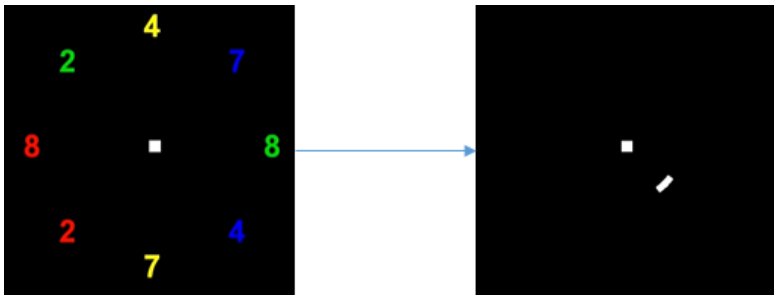


Figure 3.3: Iconic memory array

3.2.2 Participants

12 participants performed iconic memory task with EEG electrodes. 7 were males, the others were females. All participants had normal vision and no one had color blindness. Participants were recruited from Functional Brain Imaging Laboratory in Seoul National University.

3.2.3 Procedure

When participants came to the laboratory, they were provided an informed consent. In the beginning of each trial, when a box-shaped fixation point present, the participants fixed at that point. After that, the cue or the stimuli were presented depending on the condition. When the cue (or the stimuli) appeared, the participant should determine what the target stimulus was.

3.3 Results

3.3.1 Preprocessing

The recorded EEG data were bandpass filtered from 0.1Hz to 50Hz using eegfilt function implemented in EEGLAB (Delorme & Makeig, 2004). Using Electrooculography (EOG) channels, trials including eye blink or bad eye movement were rejected. Incorrect trials were also excluded, only correct trials were used

in analysis. Trials were classified by epochs, then averaged according to the experimental conditions.

3.3.2 ERP analysis

ERP analysis was conducted using R version 3.2.3 (R Core Team, 2015) and R pacakge "glmnet" (J. Friedman et al., 2010). Preprocessed EEG epochs were averaged by the experimental conditions. In each condition, baseline was corrected using pre-stimulus interval. The following figure 3.4, 3.5, 3.6, 3.7 will show averaged ERP data by conditions.

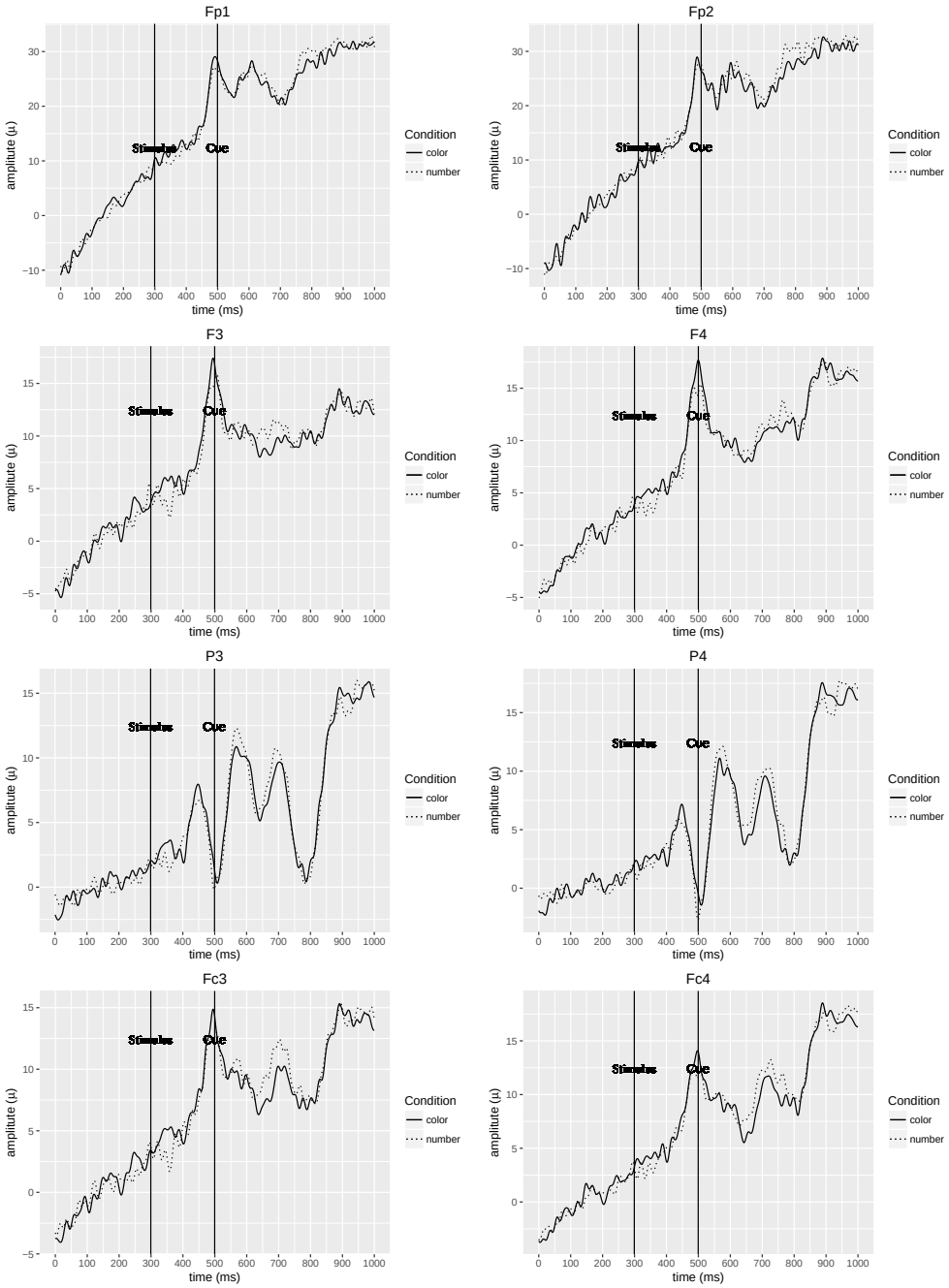


Figure 3.4: Overall post-cue 200ms averaged group ERP data

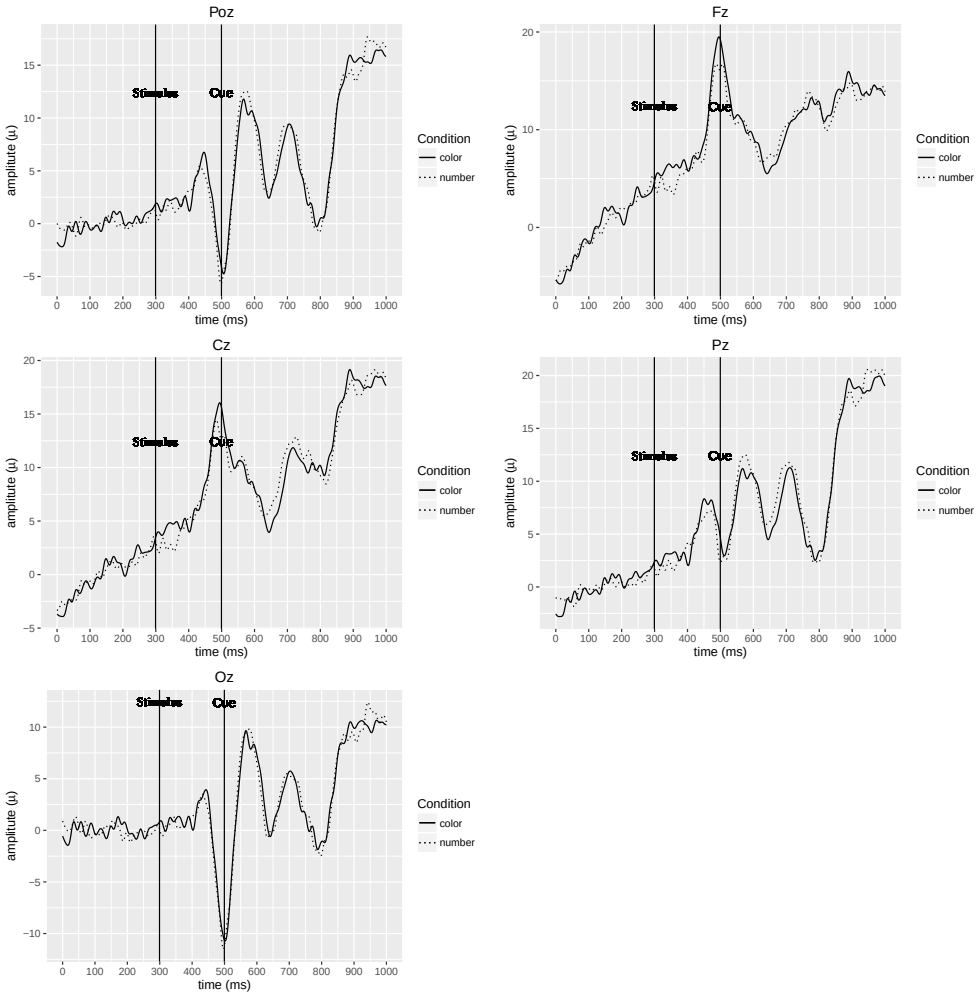


Figure 3.5: Overall post-cue 200ms averaged group ERP data (cont.)

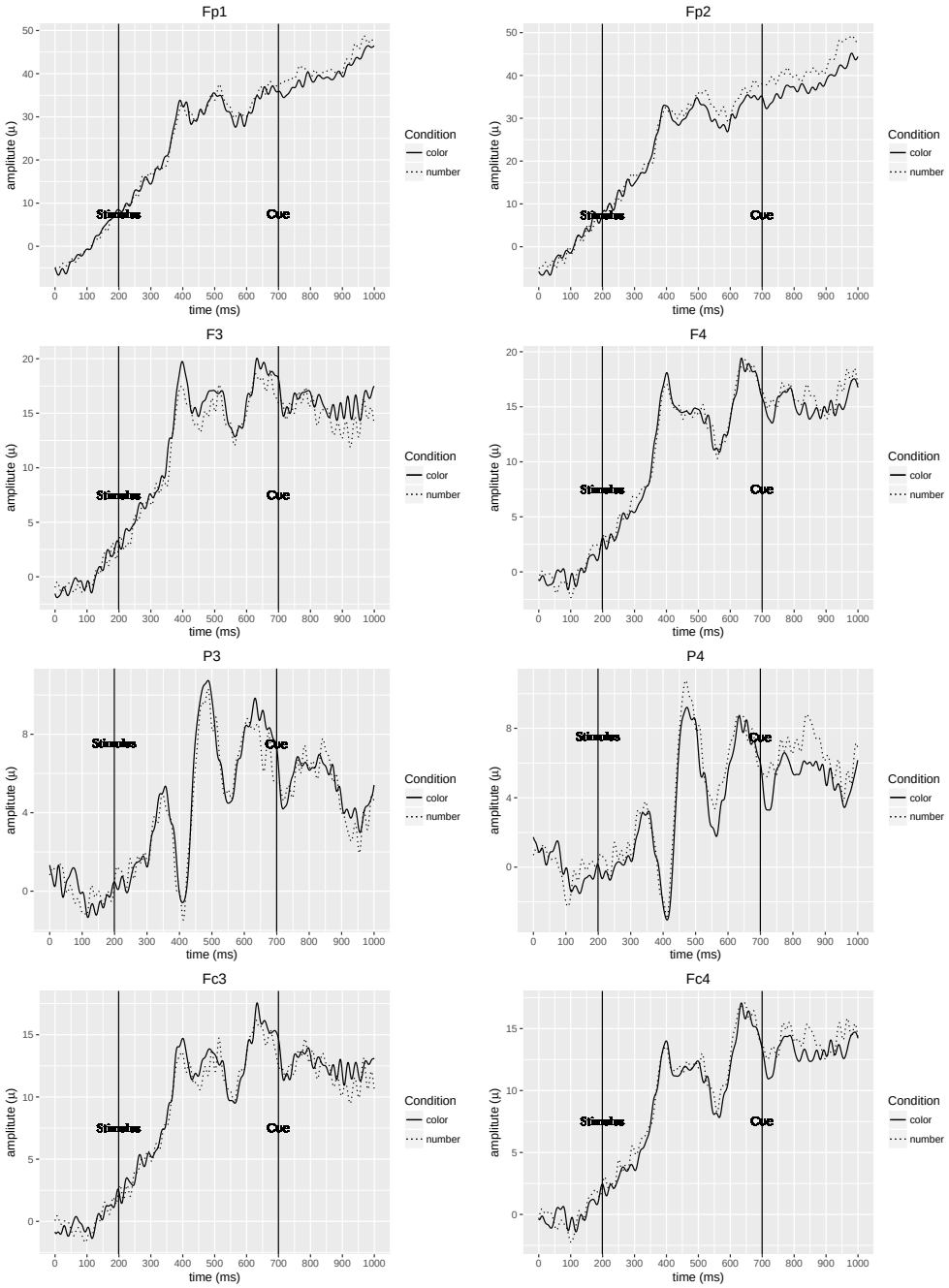


Figure 3.6: Overall post-cue 500ms averaged group ERP data

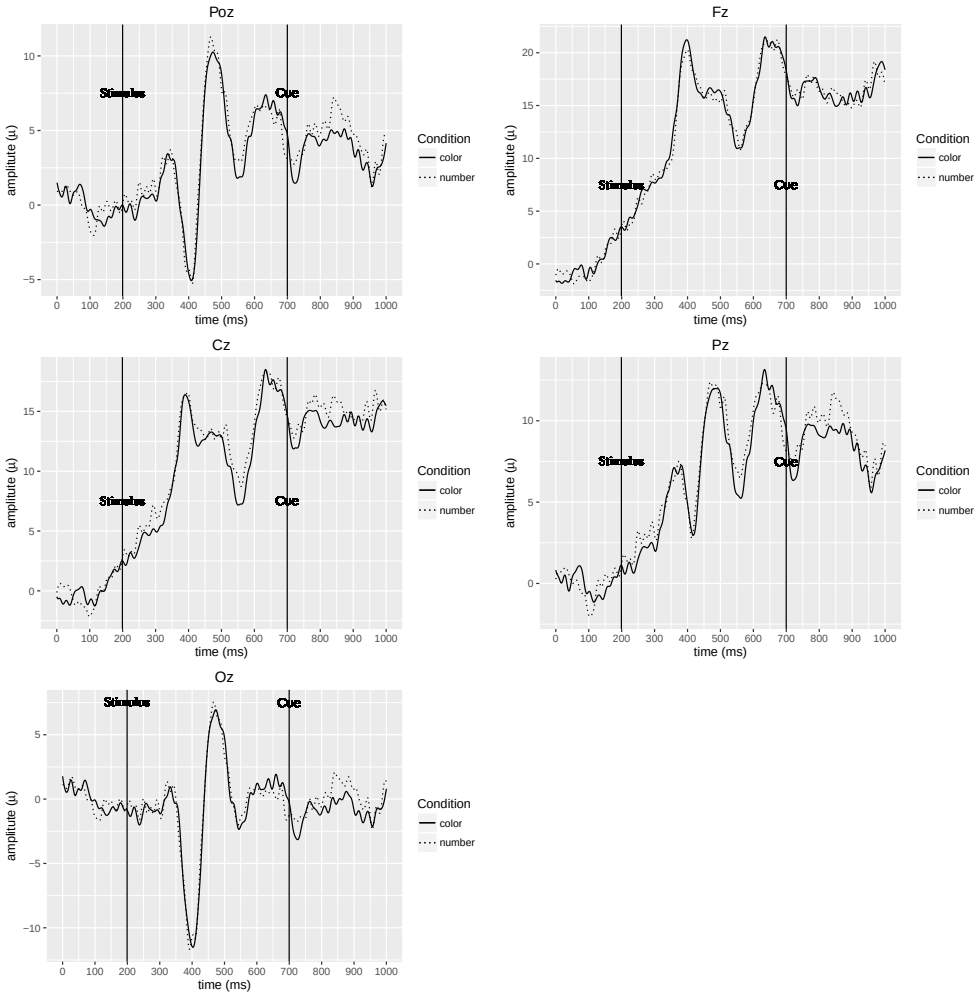
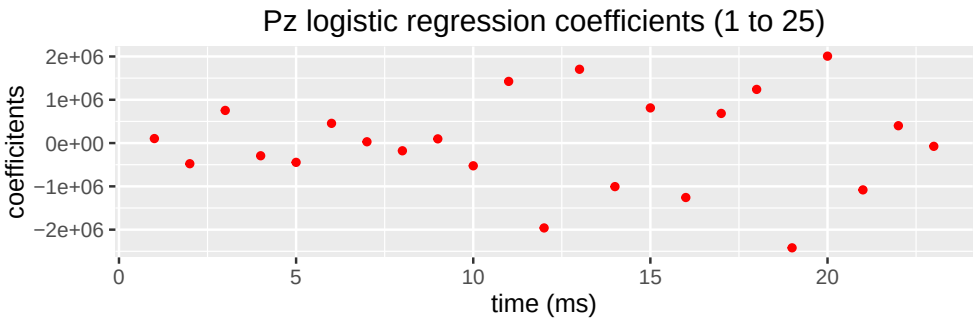


Figure 3.7: Overall post-cue 500ms averaged group ERP data (cont.)

3.3.3 Logistic regression

Using the above ERP results, logistic regression was applied for estimating ERP latency cofficients. As mentioned in the previous section, ERP data have high-dimensional structures, so it cannot be identifiable precisely. Estimated coefficients were the number of observatrions (in this case, 24), the other independent variables were excluded for identifying coefficients.

(a) Logistic coefficients



(b) Logistic coefficients (1 to 25)

Figure 3.8: Estimated logistic regression coefficients

3.3.4 Lasso logistic regression

Unlike the simulation data results, The only one variable was selected and the other variables were penalized as zero.

Figure 3.9: Lasso-logistic regression coefficients

The lasso parameter (ℓ_1) was found by performing 10-folds cross validation. As the figure 3.9 showed, all but one time points were penalized as zero, and only one time point 675ms was estimated, which coefficient value was $1.7419585 \times 10^{-17}$.

3.3.5 Ridge logistic regression

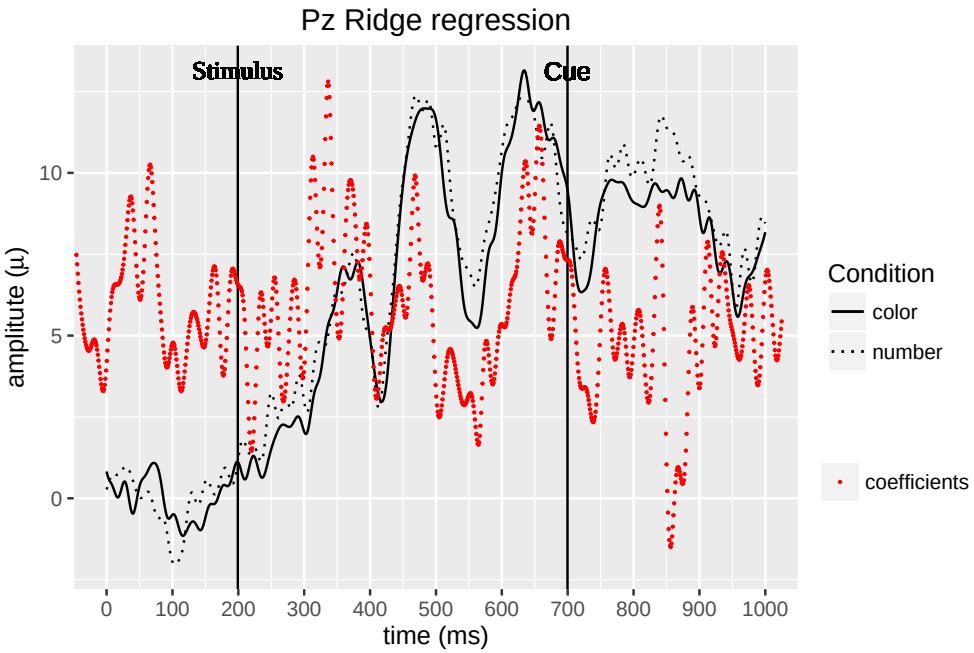


Figure 3.10: Ridge coefficients

Using ridge estimator, time coefficients were acquired. The x-axis means each time points, and the y-axis indicates the amplitude of ERP. The solid line shows the ERP wave of the color reporting condition, whereas the dotted line means the number reporting experimental condition, and the dots on figure 3.10 is the coefficients for time points. In each channel, different time points were estimated as predictor variables for specific conditions. For examples, figure 3.10 shows time point 367ms is the most probable classfication point in Pz channel in post-cue 500ms condition. The following pages are results of elastic net logistic regression with coefficients.

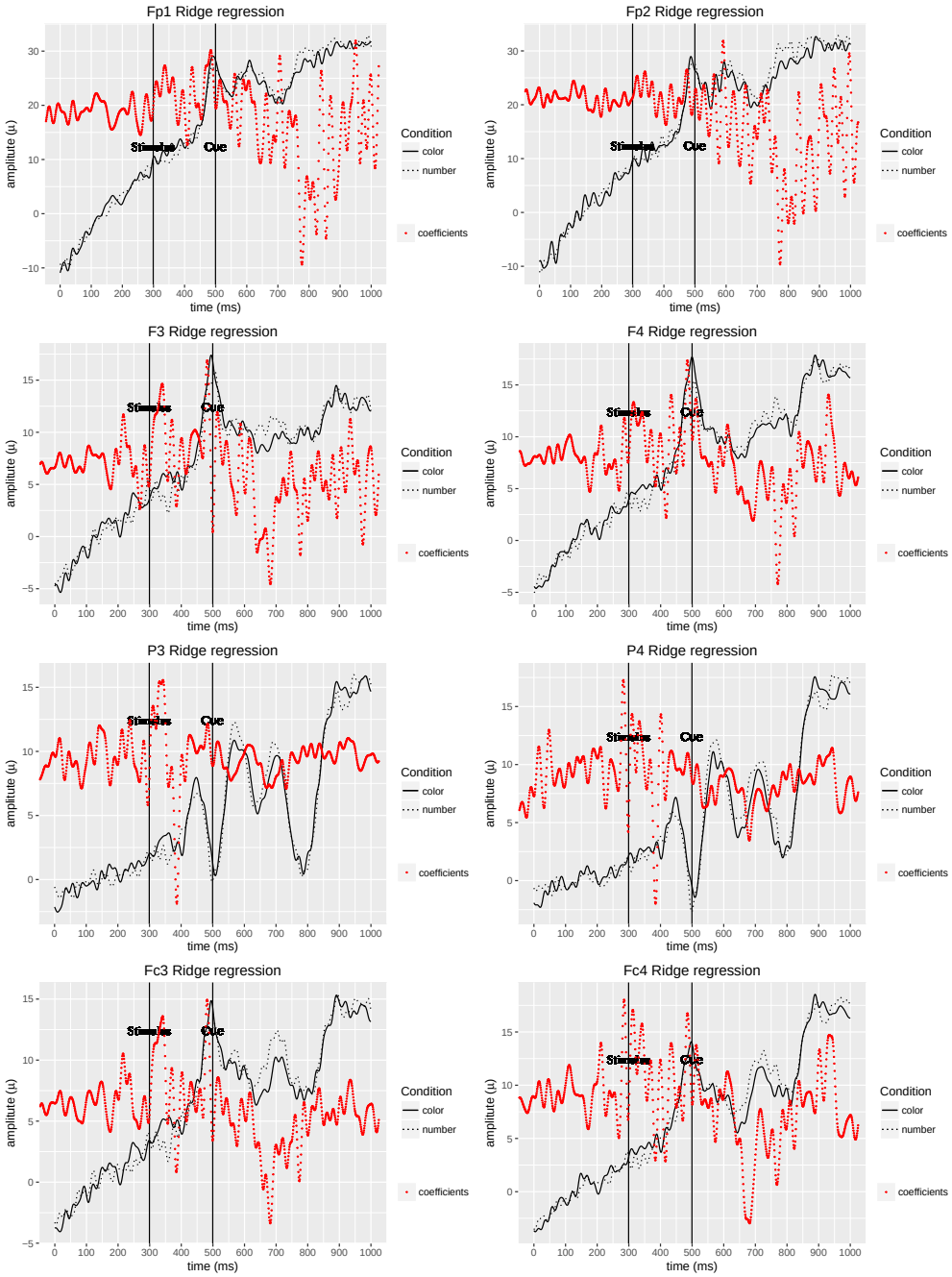


Figure 3.11: Overall post-cue 200ms ridge regression ERP data

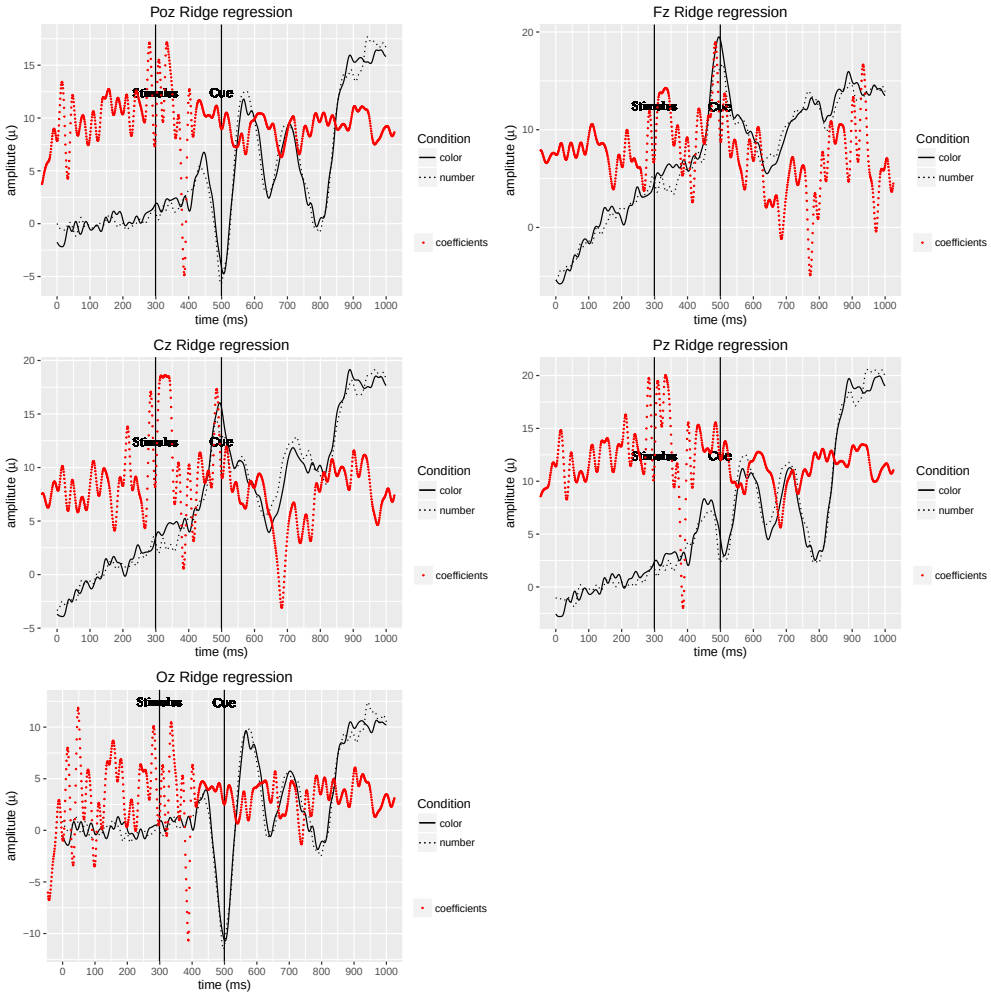


Figure 3.12: Overall post-cue 200ms ridge regression ERP data (cont.)

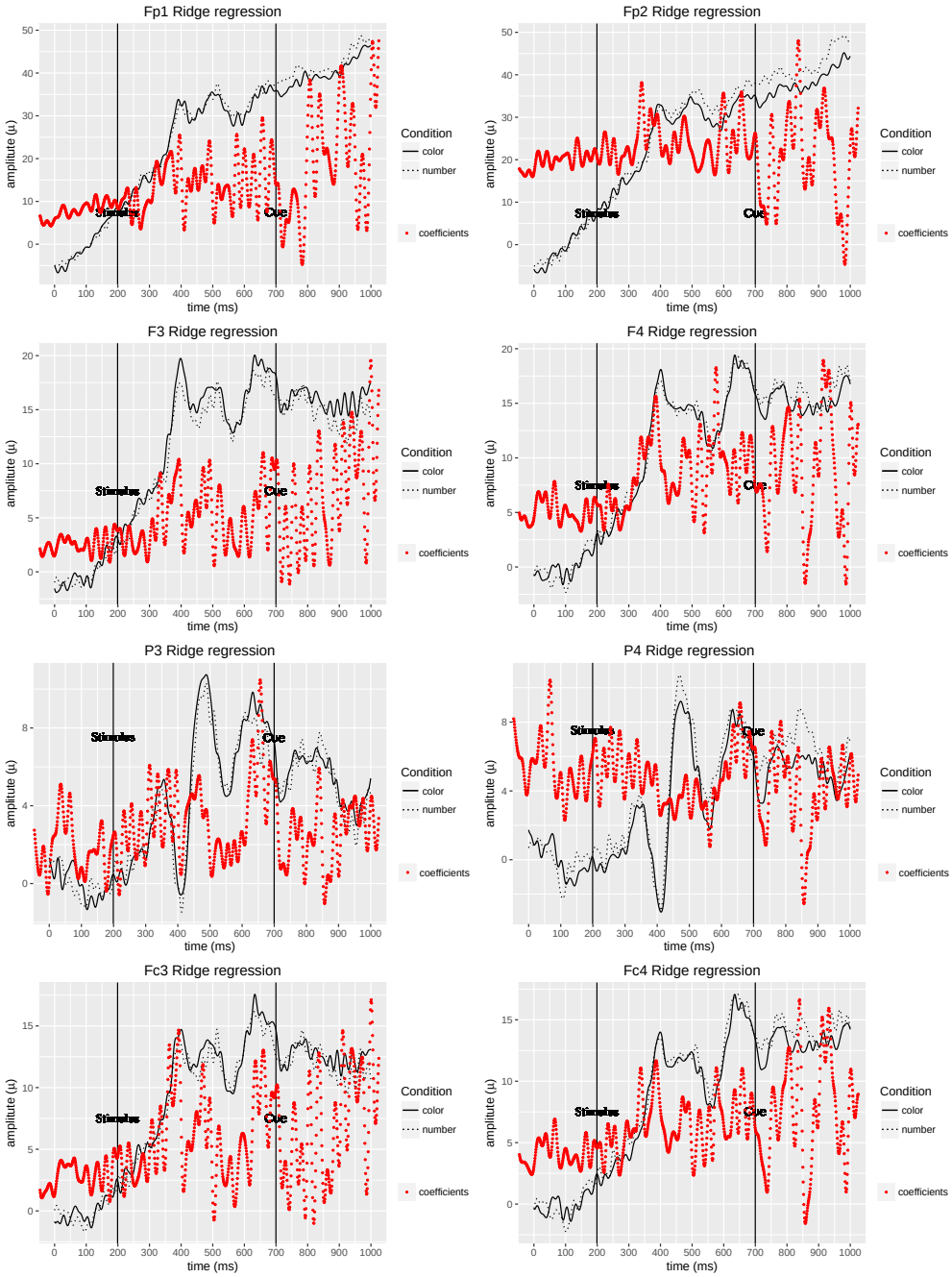


Figure 3.13: Overall post-cue 500ms ridge regression ERP data

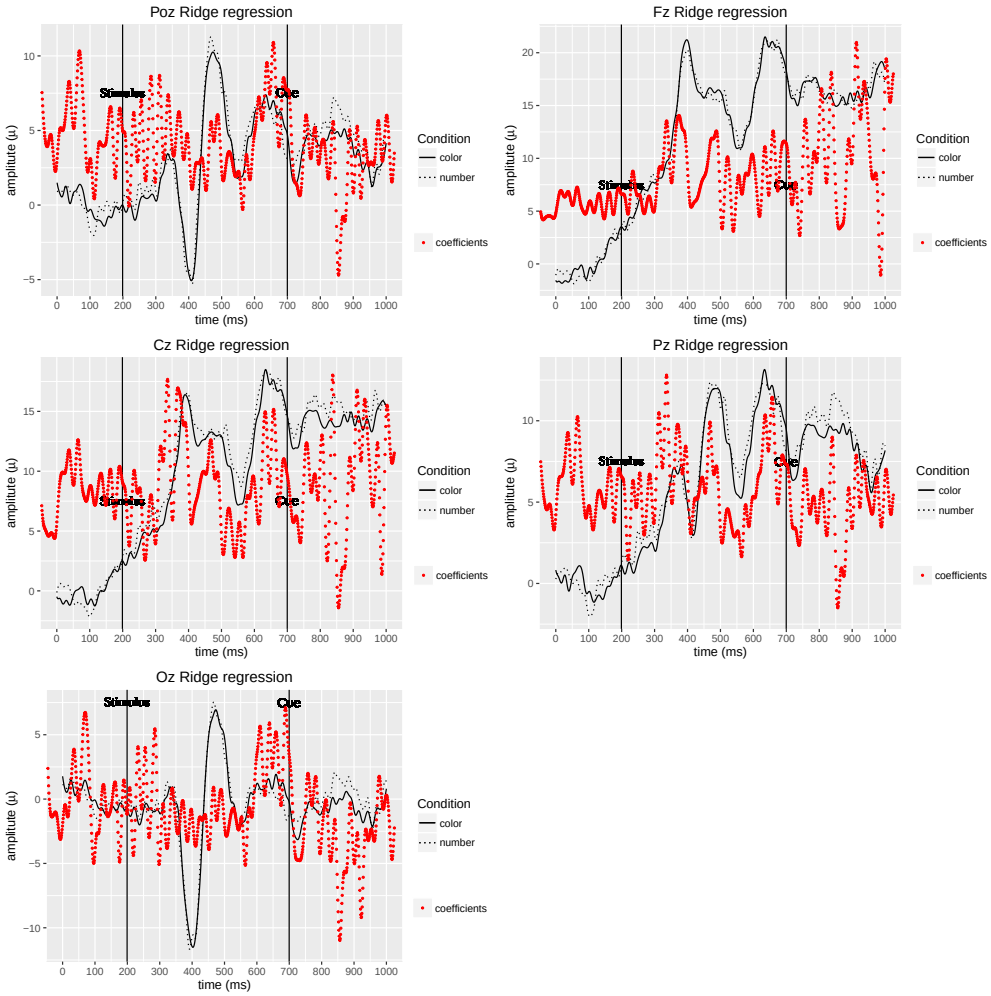


Figure 3.14: Overall post-cue 500ms ridge regression ERP data (cont.)

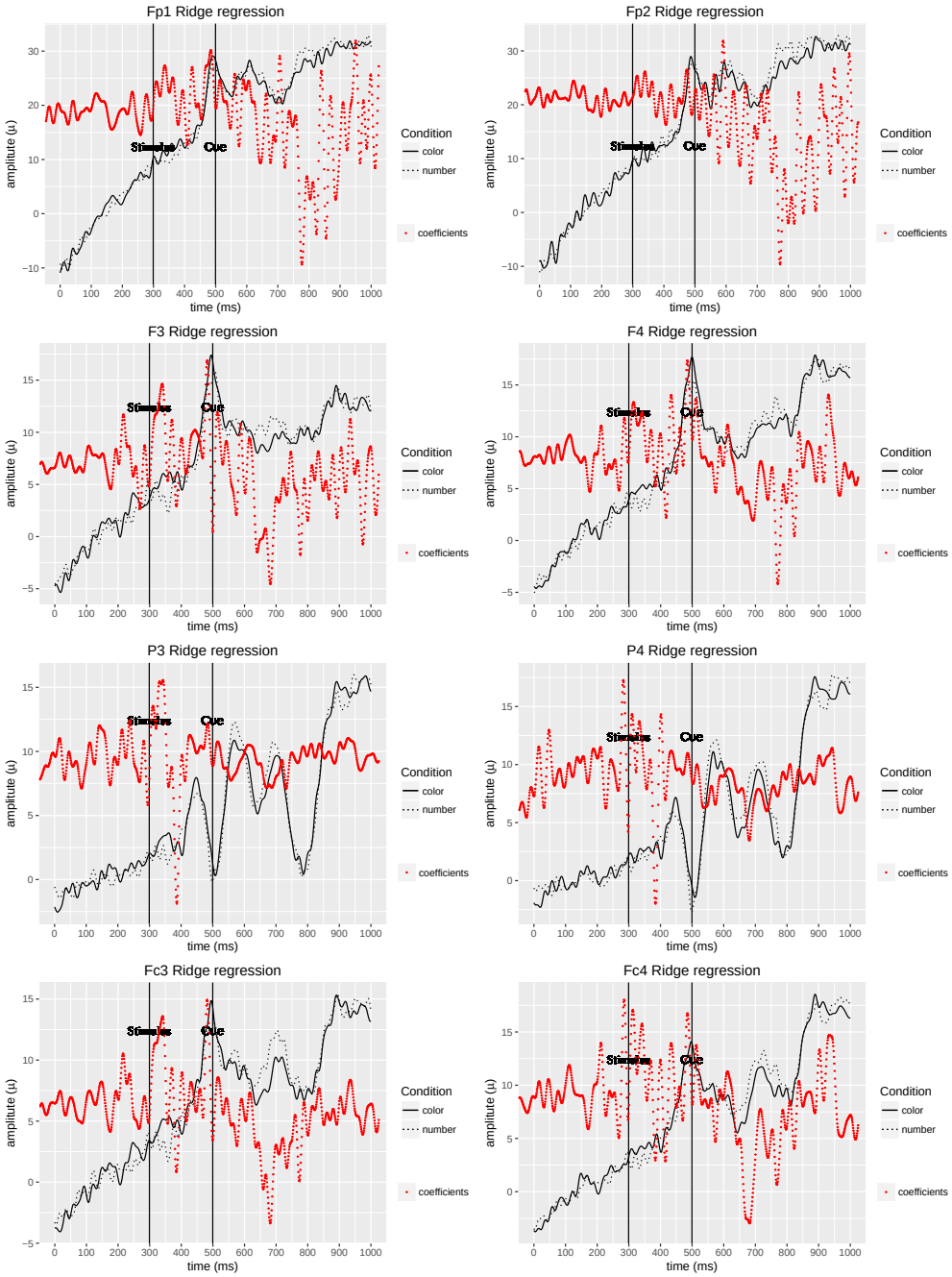


Figure 3.15: Overall post-cue 200ms lasso regression ERP data

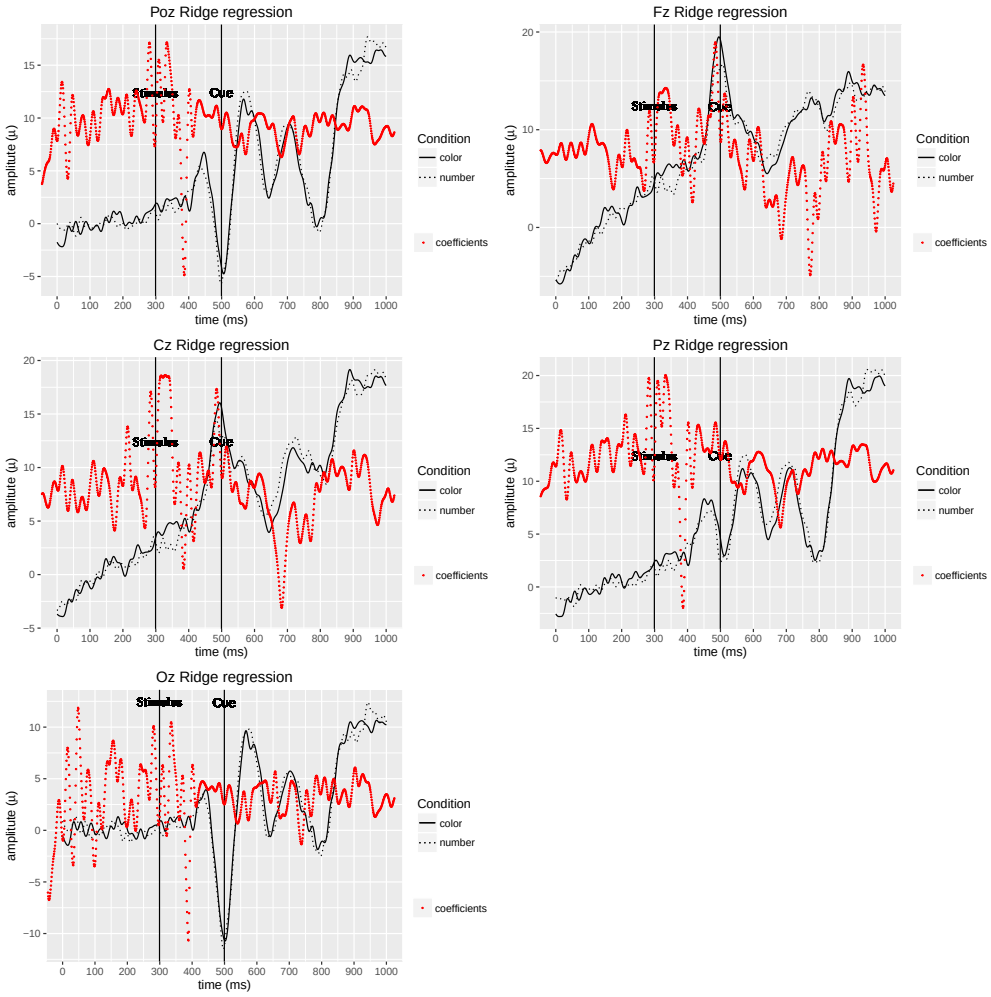


Figure 3.16: Overall post-cue 200ms lasso regression ERP data (cont.)

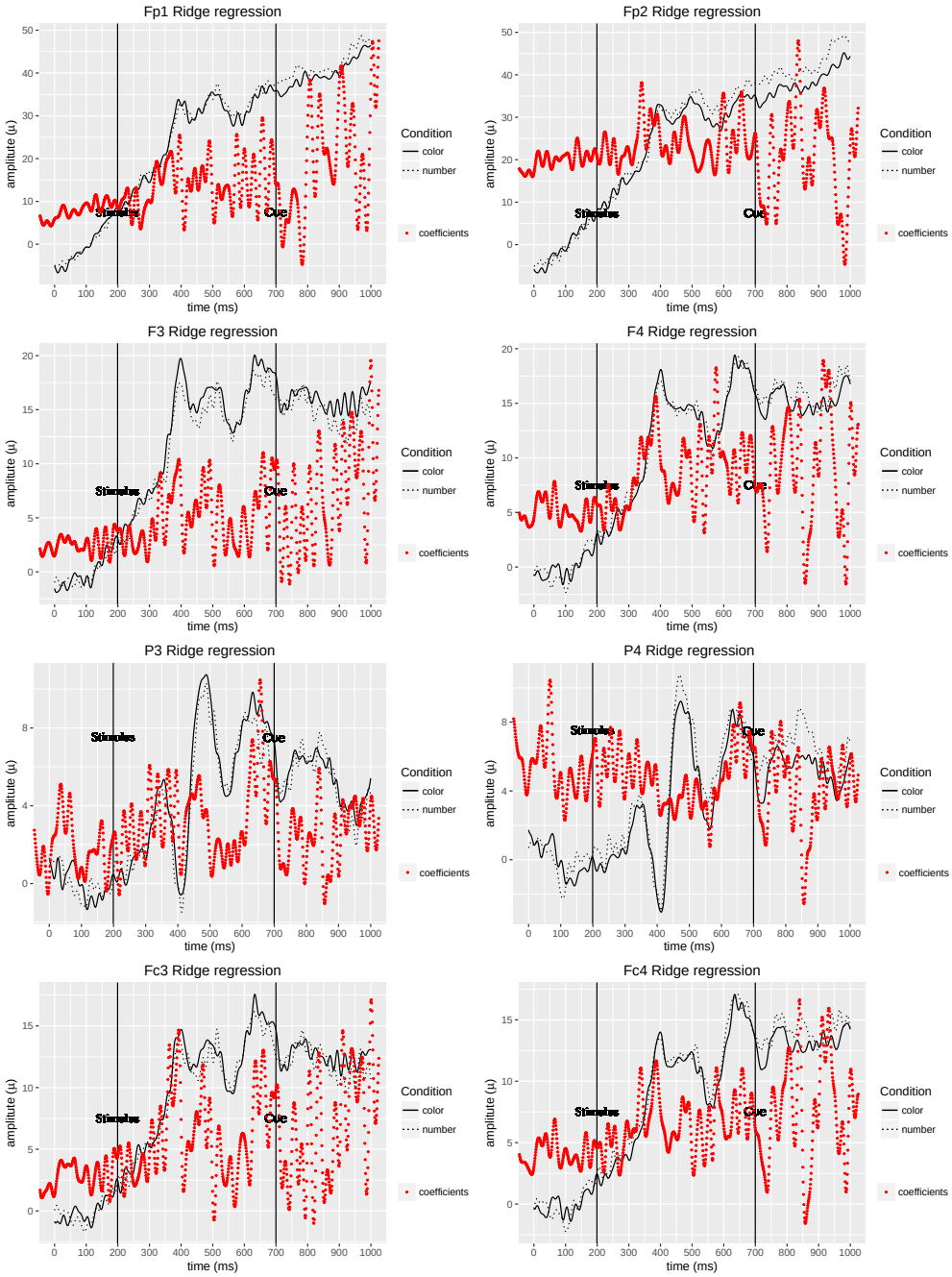


Figure 3.17: Overall post-cue 500ms lasso regression ERP data

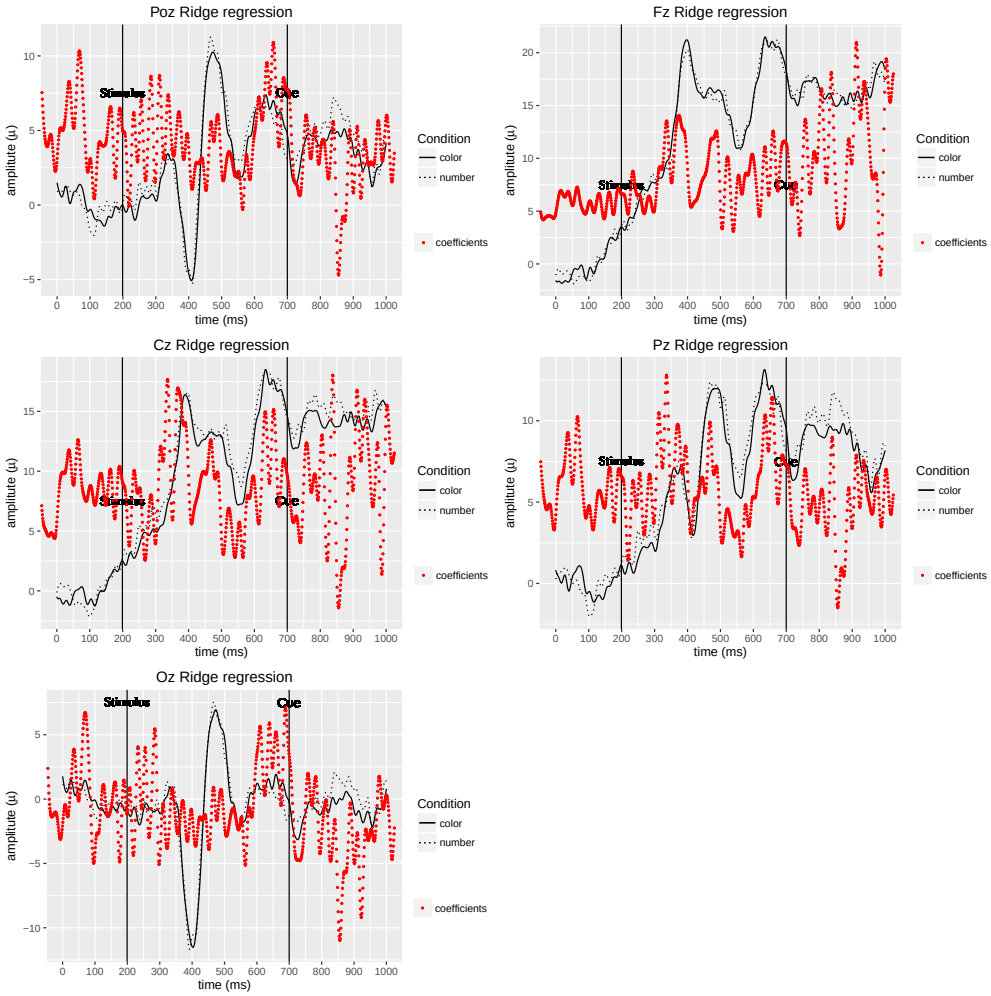


Figure 3.18: Overall post-cue 500ms lasso regression ERP data (cont.)

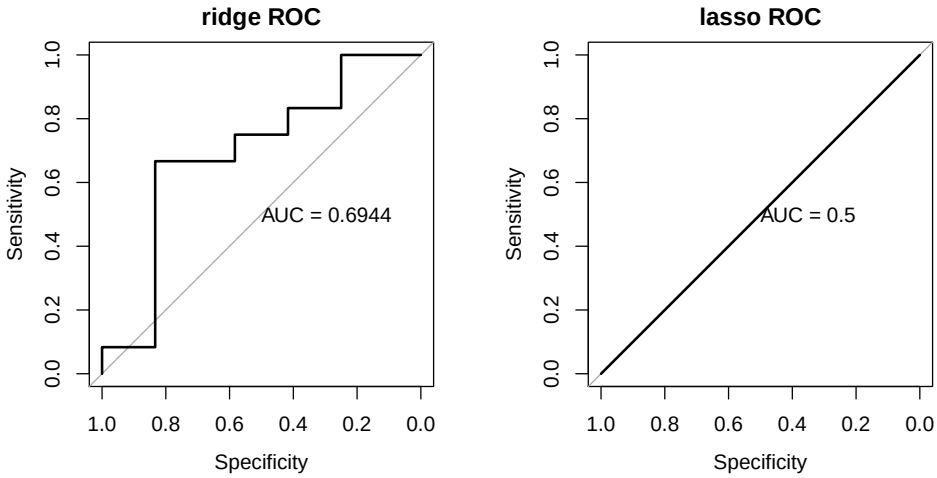


Figure 3.19: Ridge and Lasso ROC curves comparison

Figure 3.19 is ROC curves for ridge and lasso regression results. Although both of ROC curves showed lower AUCs than the simulation results, ridge was slightly better than lasso.

3.3.6 Discussion

As expected, logistic regression was not fully identifiable. In logistic regression, coefficients were obtained only in the number of observations, and the others were all excluded. Essentially, EEG data are high-dimensional, so they don't have inverse matrix of data x, in that situation, there has no sufficient degrees of freedom to estimate the model. And lasso also not fully estimated. As Zou and Hastie (2005) pointed out, when data are highly correlated, lasso selects variables before penalizing a model. Since EEG data are inter-correlated, only one variable was chosen in this study. In contrast, the ridge stably estimated coefficients. As figure 3.10 showed, 358 time point variable is the most classifiable ERP time point. That is, after the stimuli appeared in 163ms latency, ERP-wave might be changed due to a stimuli array. The participants were

provided with the stimuli array on 200ms, and they should maintain the array until the cue appeared. A potential differences between stimuli onset timing and cue onset timing might occur due to the experimental conditions. If differences between the onset timings would exist, it might be due to the maintained stimuli. As a result, the ERP latency might be related to iconic memory function. In sum, how much individual time points affect ERP waves respectively was main question in this study. Ridge could estimate all the coefficients, which can be used for identifying time points of the iconic memory function in the ERP waves.

Chapter 4. Conculsion

4.1 Conclusion and Limitations

In those studies, elastic net regularization regression analysis was applied to EEG data. Event-Related Potential has amplitude and latency. There might be two questions: when the latency is, and how strong that amplitude is. Those quesitons are important to analyze ERP components. In this respect, elastic net regularization technique is good for identifying the latency. When the number of independent variables outnumbers the number of observations ($p>n$), ridge is better than lasso, because lasso chose n variables before it fully saturates (Zou & Hastie, 2005). The result in lasso can be explained as not fully saturated coefficients. And study 1 showed lasso and ridge were similarly identified time points at the phasic peak. In contrast, in study 2, lasso estimator selected only one variable, and the result was not consistent with ridge. This result can be expained as highly correlated situation. In simulation, only one phasic peak was considered, therefore, random noise was uncorrelated. In contrast, in study 2, real EEG data not only had random noise, but confounded with systematic noise. For instance, laboratory situation, participants selection bias, demand characteristics and other covariates would affect data. Those noises may affect systemically the EEG waves, so it could make very highly correlated circumstances.

4.2 Further research

Another possible scenario is linear mixed-effect model with regularization. EEG data have multi-level structures. As such, first-level is trials, second-level is idividuals, third-level is groups. That is, EEG data have not only individual variablity, but trial-level variability. So when we fit EEG data, inter-trial and inter-personal variability should also be considered. For examples, we can think of as follows:

$$
\begin{aligned}
&\text{Level 1 (trial-level):} \\
&\qquad \mathbf{Y}_{ti} = \pi_{0i} + \pi_{1i}\mathbf{X}_{ii} + \epsilon_{ti} \\
&\text{Level 2 (person-level):} \\
&\qquad \pi_{pi} = \beta_{p0} + \sum_{q=1}^{Q_0} \beta_{pq}\mathbf{X}_{qi} + \gamma_{pi},
\end{aligned}
\tag{4.1}
$$

Level 1 is about trial level, level 2 is person level. π is coefficients, t means trial points, i means individual. Using linear mixed-effect model, we can consider that π is about time points, it can vary among trials in the same participant. If those inter-trial, inter-personal, and even inner-trial variability would be considered, much more information could be extracted. Thus, the model may discover unseen ERP data patterns.

References

Ariely, D., & Berns, G. S. (2010). Neuromarketing: the hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*, *11*(4), 284–292.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433–436.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and cognitive processes*, *13*(1), 21–58.

Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9–21.

Friedman, D., & Johnson, R. (2000). Event-related potential (erp) studies of memory encoding and retrieval: a selective review. *Microscopy research and technique*, *51*(1), 6–28.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, *33*(1), 1.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, *4*(12), 463–470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Luck, S. J. (2005). An introduction to the event-related potential technique (cognitive neuroscience).

Luck, S. J., Fan, S., Hillyard, S., et al. (1993). Attention-related modulation of sensory-evoked brain activity in a visual search task. *Cognitive Neuroscience, Journal of*, *5*(2), 188–195.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in cognitive sciences*, *4*(11), 432–440.

Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J.-M., Hinrichs, H., ... Rappelsburger, P. (1998). Ifcn standards for digital recording of clinical eeg. *Electroencephalography and clinical Neurophysiology*, *106*(3), 259–261.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of memory and language*, *31*(6), 785–806.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437–442.

R Core Team. (2015). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in cognitive sciences*, *11*(6), 251–257.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Walter, W. G., Cooper, R., Aldridge, V., McCallum, W., & Winter, A. (1964). Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. *Nature*(203), 380–4.

Yeung, N., Bogacz, R., Holroyd, C. B., & Cohen, J. D. (2004). Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods. *Psychophysiology*, *41*(6), 822–832.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320.

초록

본 논문의 목적은 정규화 기법이 사건 관련 전위를 분석하는 데에 사용될 수 있는지 알아보는 것이고, 어떤 종류의 정규화 기법이 적절한지 확인하는 것이다. 이 논문에서는 엘라스틱 넷 정규화 로지스틱 회귀가 사건 관련 전위 분석에 적합한 방법인지 제안하였다. 특히, 정규화를 통해 ERP에서의 latency 정보를 확인하는 데에 사용할 수 있는지 알아보았다. 연구 1에서는 시뮬레이션 결과를 이용하여 사건 관련 전위를 로지스틱 회귀분석의 관점에서 latency 정보를 구별할 수 있는지를 살펴보았다. 연구 2에서는 실제 참가자를 대상으로 획득한 사건 관련 전위가 시뮬레이션 결과와 같이 산출될 수 있는지 확인하였다. 연구 1에서 릿지와 라쏘 두 가지 정규화 방법을 통해 분석을 시행한 결과, 두 가지 방법에서 적절한 latency 정보를 확인할 수 있었다. 연구 2에서는 릿지 정규화를 통해서는 적절한 latency 정보를 확인할 수 있었으나, 라쏘에서는 그렇지 못 하였다.

주요어: 엘라스틱 넷, 정규화, 사건 관련 전위, ERP, 로지스틱 회귀

학번: 2013-22803