Composite Hypothesis Testing via Support Vector Regression

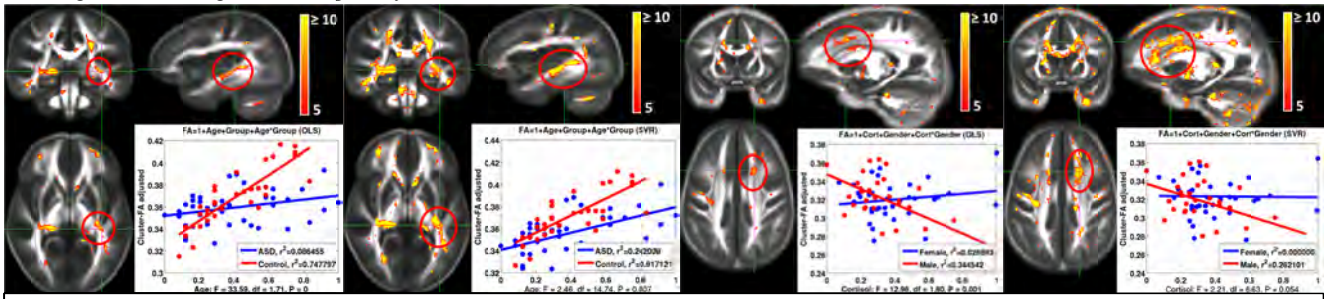
Nagesh Adluru[1], Richard J. Davidson[1], and Andrew L. Alexander[1]
[1]*University of Wisconsin-Madison, Madison, WI, United States*

Introduction: Composite hypothesis testing or "signal detection" is one of the basic concepts involved in statistical analyses of neuroimaging data. For example, the popular voxel based analyses (VBA), region-of-interest analyses, tract specific analyses, tensor-based morphometry all rely on this basic concept to detect group differences or effects of other physiological or behavioral measures. Generalized likelihood ratio tests are very commonly used to reject null-hypothesis. In these tests the data is projected onto general linear models (GLM), defined using the diagnostic/physiological variables, that represent null and alternate hypotheses and the ratio of the energies of the projected data is compared against a critical threshold. Under normal assumptions in neuroimaging setting these critical thresholds are usually chosen based on either F or t distributions and hence these are called F or t tests respectively. The projection of data on to these GLMs is most commonly performed using ordinary least squares (OLS) which minimizes least squared residuals. In this paper we propose to do the projection by minimizing ϵ-insensitive L1 residuals as used in support vector regression (ϵ-SVR). This allows for more robust projections of the data allowing for better sensitivity to signal and as a result also allow for better control of confounding effects of nuisance parameters used in the models. We demonstrate the potential advantages of the proposed approach in VBA using diffusion tensor imaging (DTI) data from an autism study as well as an emotion-regulation study.

Methods: Data: For the autism study, data acquired from 78 male subjects were used: 42 high function subjects with autism spectrum disorders (ASD) and 36 controls group-matched for age, handedness and IQ. The data were acquired on a Siemens Trio 3.0 Tesla Scanner with an 8-channel, receive-only head coil using a single-shot, spin-echo, EPI pulse sequence and SENSE parallel imaging (undersampling factor of 2). Diffusion-weighted images were acquired in 12 non-collinear diffusion encoding directions with diffusion weighting factor b=1000s/mm^2 in addition to a single reference image (b=0). For the emotion-regulation study data from 64 eighteen-year-old adolescents were used. Cortisol (Cort) was obtained from salivary samples collected over 3 consecutive days when they were 4.5 years of age. Cort is an important steroid hormone implicated in the stress response, serving as an important measure in studies of emotion-regulation and anxiety [1]. The diffusion weighted images were acquired on a GE 3.0 Tesla scanner using 48 non-collinear diffusion encoding directions with diffusion weighting factor of b=1000s/mm^2 in addition to eight b=0 images. Image processing: Eddy current related distortion and head motion of each data set were corrected using FSL software package [2] and distortions from field in-homogeneities were corrected using field maps for the emotion-regulation study. For both the studies, brain tissue was extracted using the brain extraction tool (BET), also part of the FSL. The tensor elements were calculated using non-linear estimation using CAMINO [3]. It is important to establish spatial correspondence of voxels among all the subjects before performing VBA. Hence, a state-of-the-art diffusion tensor image registration DTI-TK [4] was used for spatial normalization of the subject data. All voxel based analyses were performed on spatially normalized 1mm^3 isotropic volumes with a data resolution of 192×224×144. Key ideas for robust projection: Let us assume we want to project n measurements (e.g. data from n different subjects) at a voxel using a GLM with p independent parameters, i.e. $Y = X\beta + \varepsilon$, where $Y \in R^{n\times 1}$, $X \in R^{n\times p}$ and $\beta = R^{p\times 1}$. Let β' be the estimated parametric model. Then, $Y' = X\beta'$ is the projected signal. The projection matrix is defined as a matrix H such that $Y' = HY$. H plays a crucial role in determining the degrees of freedom (dof) of the F distribution used for F-tests. In case of OLS, $H = X(X^TX)^{-1}X$ while in case of ϵ-SVR $H = \frac{1}{n}[X\beta'(1/Y)^T]$. Since ϵ-SVR estimates β by minimizing $|Y - X\beta|_\epsilon$ with a $\|\beta\|^2$ regularization [5], this leads to robust projections [6] and leads to improvements in the dof estimation [7,8] and hence improved F-stats. Implementation: The proposed hypothesis testing framework was implemented by integrating two popular MATLAB software packages: SurfStat [9] and LIBSVM [10]. SurfStat allows intuitive representation of GLMs using higher level representations known as model formulas. For example, if one wants to study the effect of group and age on fractional anisotropy (FA) in the brain by covarying for Gender, one could design the GLM as FA=1+Age+Group+Gender, where Age, Group and Gender are simply MATLAB arrays wrapped by a function called *term*. Comparisons: To demonstrate the advantage of using ϵ-SVR over OLS, we examine the following two different GLMs, one for each study: 1) FA=β_0+β_1Age+β_2Group+β_3Age×Group, 2) FA=β_0+β_1Cort+β_2Gender+β_3Cort×Gender, where the corresponding null-hypothesis are β_3=0 for each of the model. The first GLM aims at measuring the group difference between ASD and Controls in terms of the effect of Age on FA, while the second aims at measuring the effect of Cort on FA. We estimate the above GLMs voxelwise at each voxel in the white matter mask, which is defined as the set of voxels whose population-mean FA>0.2 and visualize the$F$-statistic maps as well as regression plots in some clusters of effects.

Results: It can be observed in the Fig. 1, that both OLS (left) and ϵ-SVR (right) show similar regions of significance but ϵ-SVR obtains higher F-stats (see colorbar). Also ϵ-SVR produces higher F-stat maps in spatially more contiguous regions (encircled in red) thus enabling biologically more meaningful results. The two-left regression plots show the regression between the mean FA (adjusted for Age and Group as nuisance covariates) and Age in the encircled cluster on the cingulum bundle. As can be seen in the plots ϵ-SVR not only obtains higher F-stats but also accounts for the nuisance covariance (especially in the ASD group) more accurately. The regression plots on the two-right are between the mean FA (adjusted for Gender and Cort) in the superior-frontal projections of white matter tracts and Cort.

Discussion: Support vector machines have been used in neuroimage analyses but mostly in the context of classification and not in the context of composite hypothesis testing. Our proposed approach of using ϵ-SVR results in robustness of projection of data onto the GLMs. The benefits can be seen both in terms of obtaining higher F-statistics and also in addressing the nuisance covariance. The implementation is made by integrating popularly used software packages for a more direct impact of the presented work. To our best knowledge, this is the first attempt to apply a very successful loss function used in machine learning community to the GLM framework for statistical significance testing in neuroimage analyses.



w the

References: [1] Smider et al. Child Dev 73, 2002; [2] Smith et al. NIMG 23, 2004; [3] Cook et al. ISMRM, 2006; [4] Zhang et al. MIA 10, 2006; [5] Smola et al. Stats & Comp. 14, 2002; [6] Huber J.H.Wiley & Sons, 1981; [7] Welch Biometrika, 1947; [8] Satterthwaite Biometrics, 1946; [9] Worsley et al. NIMG 47, 2009; [10] Chang et al. ACM Trans. Intel. Sys. Tech. 2, 2011.