Statistical Methods (127001 SS21) – Assignment 2

Dr. Nicolas Schuck
`schuck[at]mpib-berlin[dot]mpg[dot]de`

July 16, 2021

Notes Please submit your answers as a single pdf until August 15th to Nico (see email above).

Problem 1

Download subjects 1-4 of the Haxby data (see documentation of `datasets.fetch_haxby()` from `nilearn`). Subset the data to include all available classes, except `rest`. Perform suitable data preprocessing, including at least run-wise detrending and standardisation. Using the preprocessed data, set up a leave-one-run-out multi-class classification scheme.

a) (10 points) Compare performance of a multinomial logistic regression with a linear SVM either using one-vs-rest or a one-vs-one multi class classification. Make a graph that shows the mean decoding accuracy for each approach and the variability between subjects.

b) (10 points) Please report the decoding accuracy (percent true positives) for the individual classes and make a corresponding figure. Can some classes be decoded better than others?

c) (10 points) Using a linear SVM and the multi-class approach that worked best in **a**), set up a nested cross validation scheme in which you first find the best C parameter, and then decode using that C. Please submit your code of the nested cross validation. You can use any function that is part of `nilearn` or `sklearn`.

Problem 2

You want to find weights that minimize some error function, such as $E_D(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\left(t_n - \mathbf{w}^\top \mathbf{x}_n\right)^2$ (i.e., a sum of squares error function; n is the number of examples, t is the target). To regularize, you add a weight penalty term $E_W(\mathbf{w})$ of the form

$$E_W(\mathbf{w}) = \frac{1}{2}\sum_{j=1}^{M}|w_j|^q \tag{1}$$

i.e. the sum of the exponentiated absolute values of the weights. Note that q is a parameter that determines the exponent, and M is the number of weights. In the total error function, $E(\mathbf{w})$, the penality term is weighted by a parameter λ

$$E(\mathbf{w}) = E_D(\mathbf{w}) + \lambda E_W(\mathbf{w}) \tag{2}$$

such that the error function becomes

$$E(\mathbf{w}) = E_D(\mathbf{w}) + \frac{\lambda}{2}\sum_{j=1}^{M}|w_j|^q \tag{3}$$

a) (10 points) Describe how λ and q affect regularisation, using either your own words, or graphs that illustrate the behavior of $\lambda E_W(\mathbf{w})$ as a function of the two parameters and the weights. Hint: you don't need to worry about $E_D(\mathbf{w})$, and you can make simplifying assumptions about the size of $\mathbf{w}$.