Homework 6 - Encoding models

Data

The data is exactly the same as for homework 5 (data set 5, see description there). In this homework, we are building and evaluating encoding models for the activation patterns across the 5 fingers.

We are trying to compare two different encoding model for this data:
Muscle model: The 4 "muscle" features are the first 4 principle components of the muscle activation patterns observed during each of the 5 finger movements. For example. The first component contrast thumb movements with all the other movements. This implies that the thumb uses different musculature from the remaining fingers. The subsequent columns then capture the remaining differences between fingers. Because they are the components from a PCA, they are not easily interpreted.
Usage model: The first 4 principle components, derived from the natural statistic of hand movements (Ingram et al.) - correlations of MCP velocities - i.e. the idea that M1 encodes movements that are executed in every day life together also together in similar areas.

To define the models, dataset_6 now includes dataset_5, but also the model definition. The Models are stored in the structure array model with the fields:

Model(i).name = name of Model
Model(i).X = 5*4 Model feature (design matrix) indicating the value for each finger (row) for each feature (column).
Model(i).G = Predicted second moment matrix (in correlation format)
Model(i).G_cent = Centered predicted second moment matrix

1. Build an encoding model
Visualize the model feature matrices for the two models. Where are the models similar and where are they distinct?

2. Evaluate the encoding model without regularization
Write the function that estimates the the voxel-feature weights (W) using the data from 7 runs. This is most easily done by averaging the activity estimates across the 7 runs first. Make sure you remove the mean activity for each voxel across fingers within each run first.

Evaluate the muscle model and the Usage model, each time using 1-4 eigenvalues. So you are really evaluating a total of 8 different models. Do it separately for each subject. Use the correlation criterion and the R2 to evaluate the model fit. Make a bar or line graph to show the average predictive correlation and cross-validated R2, averaged across participants.
What do you observe? How do you explain your observations?

3. Evaluate the encoding model with regularization
Write another function that estimates W using Ridge regression with a regularization factor of lambda = 0.1.
Again, evaluate the 8 models (muscle / usage + 1-4 factors), using the same correlation between real and predicted patterns as evaluation criterion. How do the results change from 3? Why?