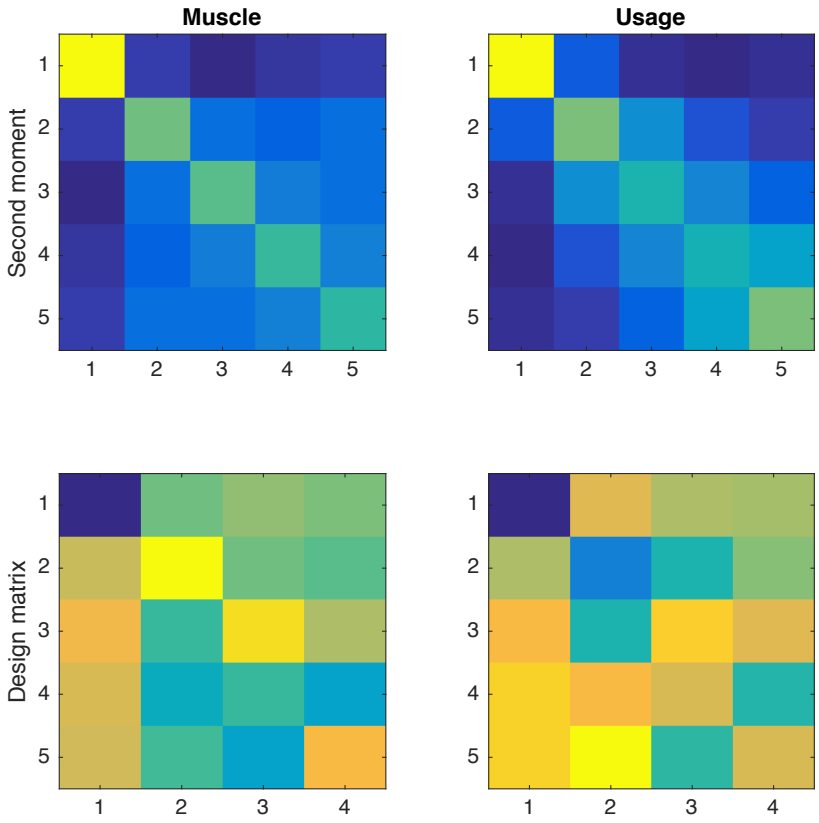
Homework 6 - Encoding models

1. Build an encoding model

Visualize both the model feature matrix. Where are the models similar and where are they distinct?



The first row shows the second moment matrix for both models (G), the second one the model feature matrices. As can be seen, the two models share the first and most important feature, which codes the thumb vs. other fingers.

2. Evaluate the encoding model without regularization

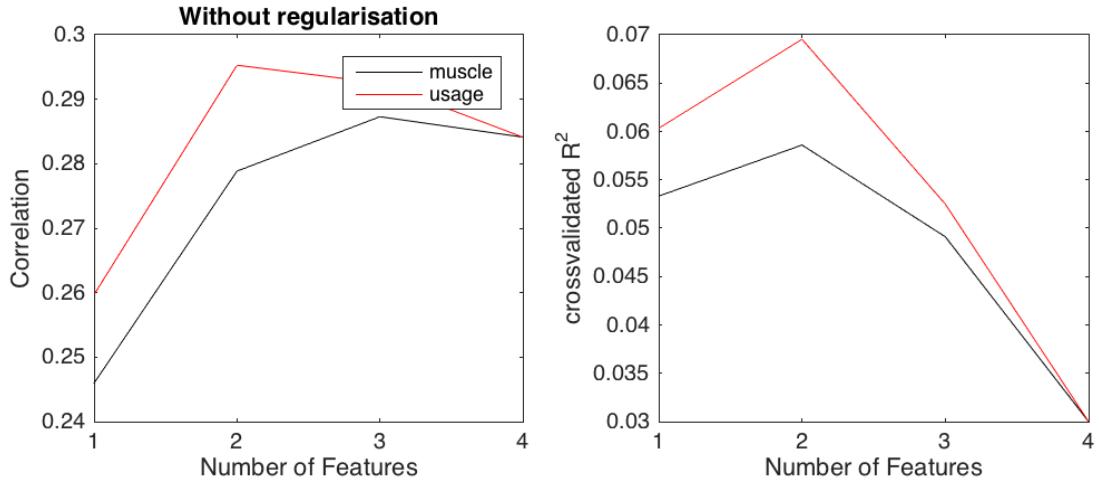
Write the function that estimates the W using the data from 7 runs (i.e. the average data).
Make sure you remove the mean across fingers for each voxel within each run first. Then use the correlation criterion to evaluate the model fit.
Evaluate the muscle model and the Usage model, using 1-4 eigenvalues. So you are really evaluating a total of 8 different models. Do it separately for each subject.
Make a bar or line graph to show the average predictive correlation, averaged across participants.
What do you observe? How do you explain this?

```
function [R,R2] = encoding_test(Y,X,lambda);
% Y = data
% X = Design matrix
% lambda  = regularisation coefficient for L2 (set to zero for no
prior)


% Subtract the mean pattern for each run
Y = bsxfun(@minus,Y,mean(Y,1));
N = 5;
numFact = size(X,2);

% Run encoding model
for i=1:8
    indx = [1:8];
    indx(i) = [];                       % training indices
    W  = inv(X'*X+eye(numFact)*lambda)*X'*mean(Y(:,:,indx),3); %
Estimation of weights
    Yp(:,:,i) = X*W;                            % Prediction activity
patterns
end;

% Evaluate using correlation and R2 measure
R = corr2(Y(:),Yp(:));
R2 = 1-sum((Y(:)-Yp(:)).^2)/sum(Y(:).^2);
```

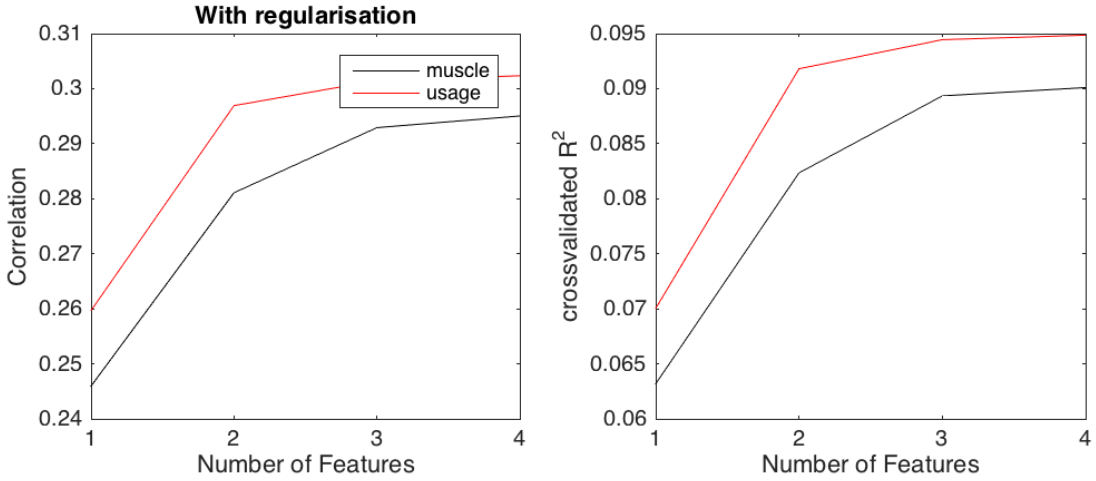


The Figure shows the average predictive correlation and cross validated R2 for the two models (muscle / usage) using 1-4 numbers of features (eigenvectors). For 1-3 number of features the muscle model is superior. For 4 features both are identical using either criterion. This is because they now span the same subspace. 4 instead of 5 features are sufficient to cover the whole space, as we removed the mean activity pattern across all conditions, which removes one df. For the R2 criterion, the best performance is reached with 2 features, meaning more features overfit the data.

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?

We simply our function with a regularization coefficient of 0.1:



The figure shows the average predictive accuracy for the two models, depending on the number of features involved. For 1 feature the results are very close to without regularization. For more features, the predictive accuracies are higher, and the models do not converge to the same point. This is because each eigenvector is now properly weighted (ensured by the ridge prior). That means that even though the two models span the same subspace, their predicted distributions differ.