Cognitive Modelling

Modelling of face perception 2

Introduction

- We are modelling face perception
- Last time you ran an experiment and collected data
- Now we move on to data analysis and modelling

The experiment

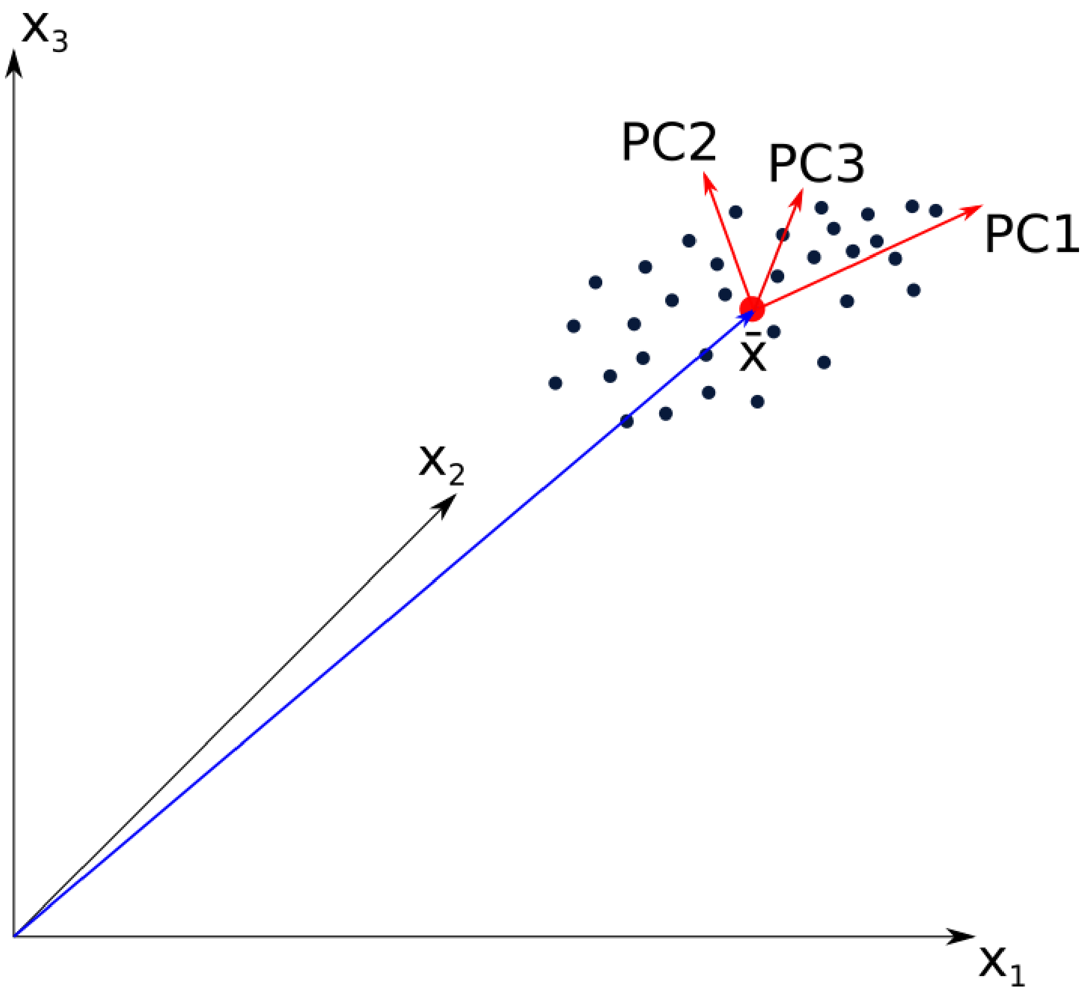
- Cognitive test about face perception
  - How do we recognize that a face is smiling?
  - What characterizes a masculine-looking vs a feminine-looking face?

Can we teach a computer to model this kind of information?

Plan for today

- 9.00-10.00 Lecture
  - PCA for image data
  - Discrete answers to continuous scale
  - Feature selection
  - Linear regression
  - Generate samples from the model
- 10.00-12.00
  - Work on project

PCA – loadings in original space

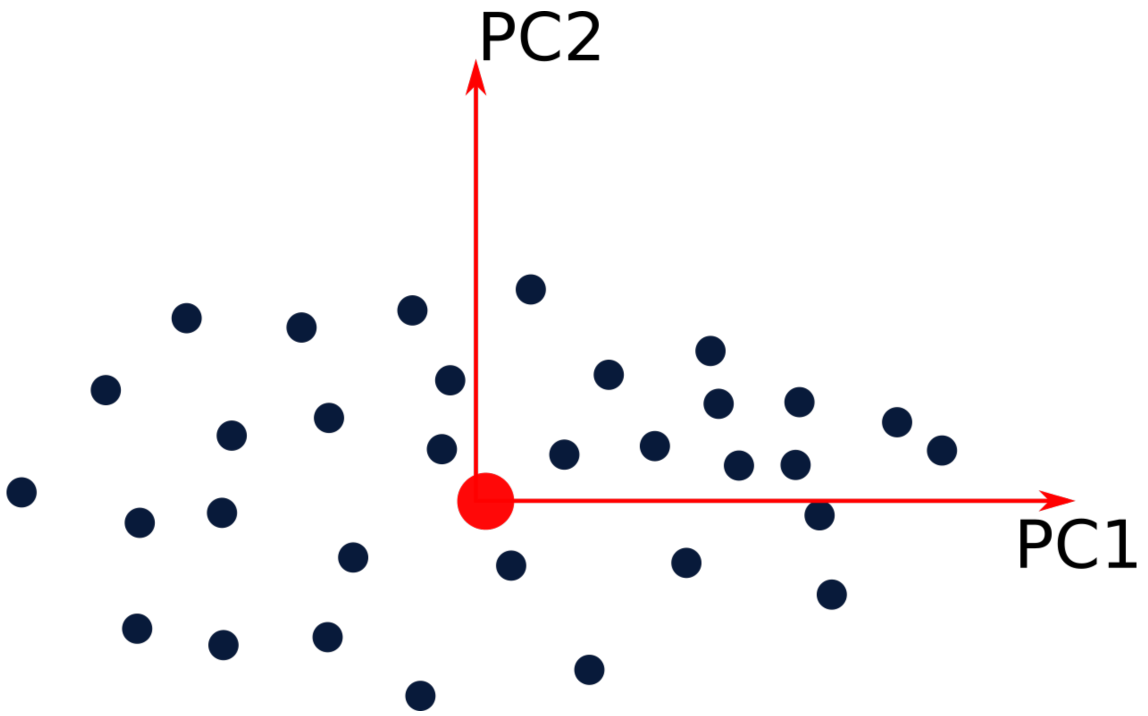


Remember to subtract the mean. In Matlab:

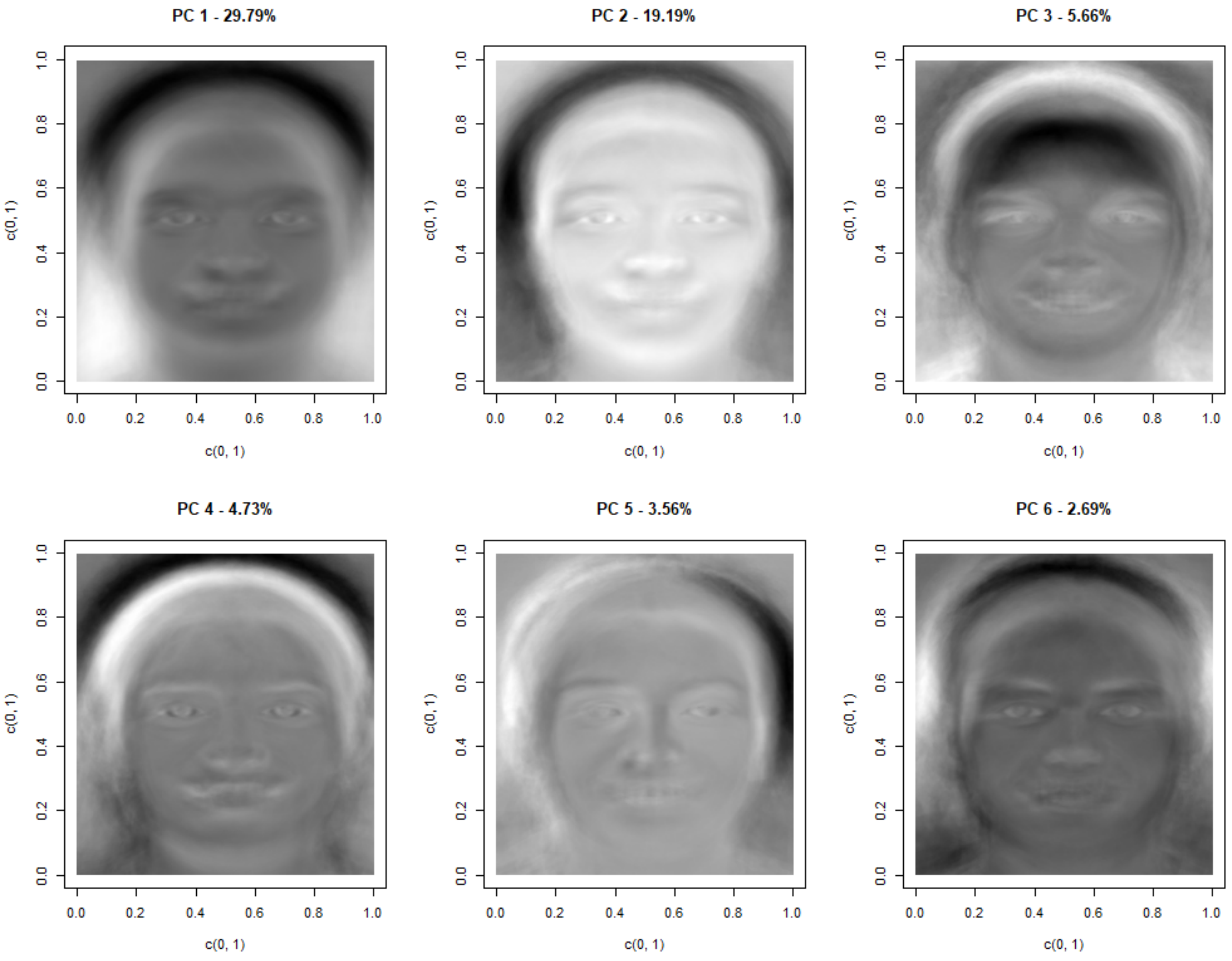
```
pca(image_data, 'Centered', true)
```

PCA space

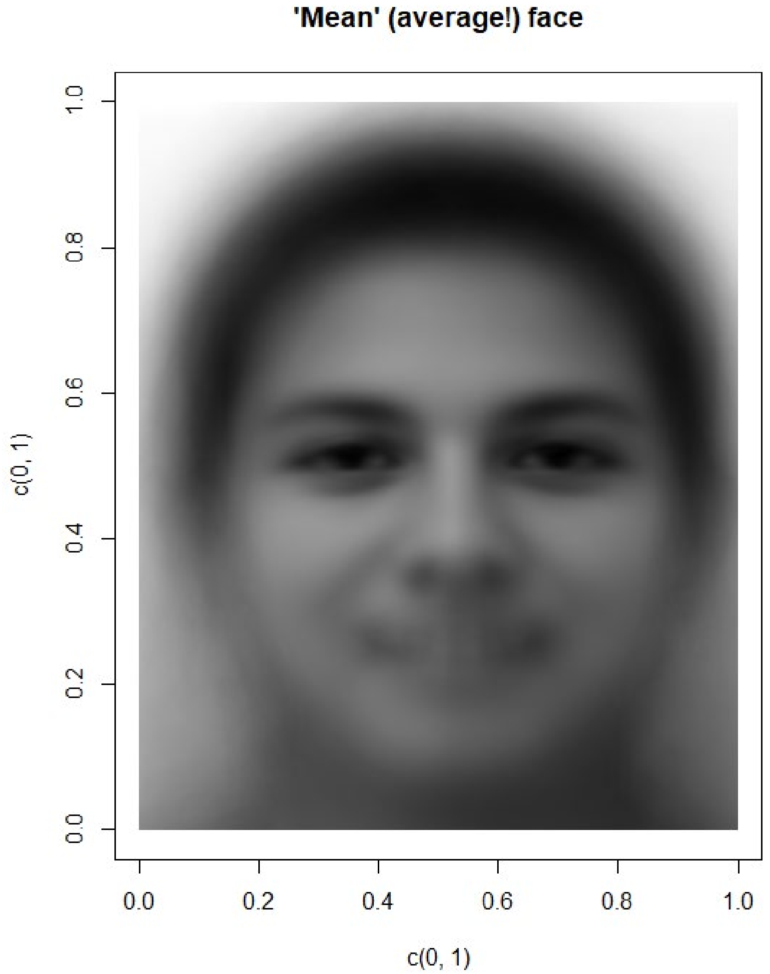
- Each sample (image) is represented as a linear combination of the components, where the coefficients are the **PCA scores**.

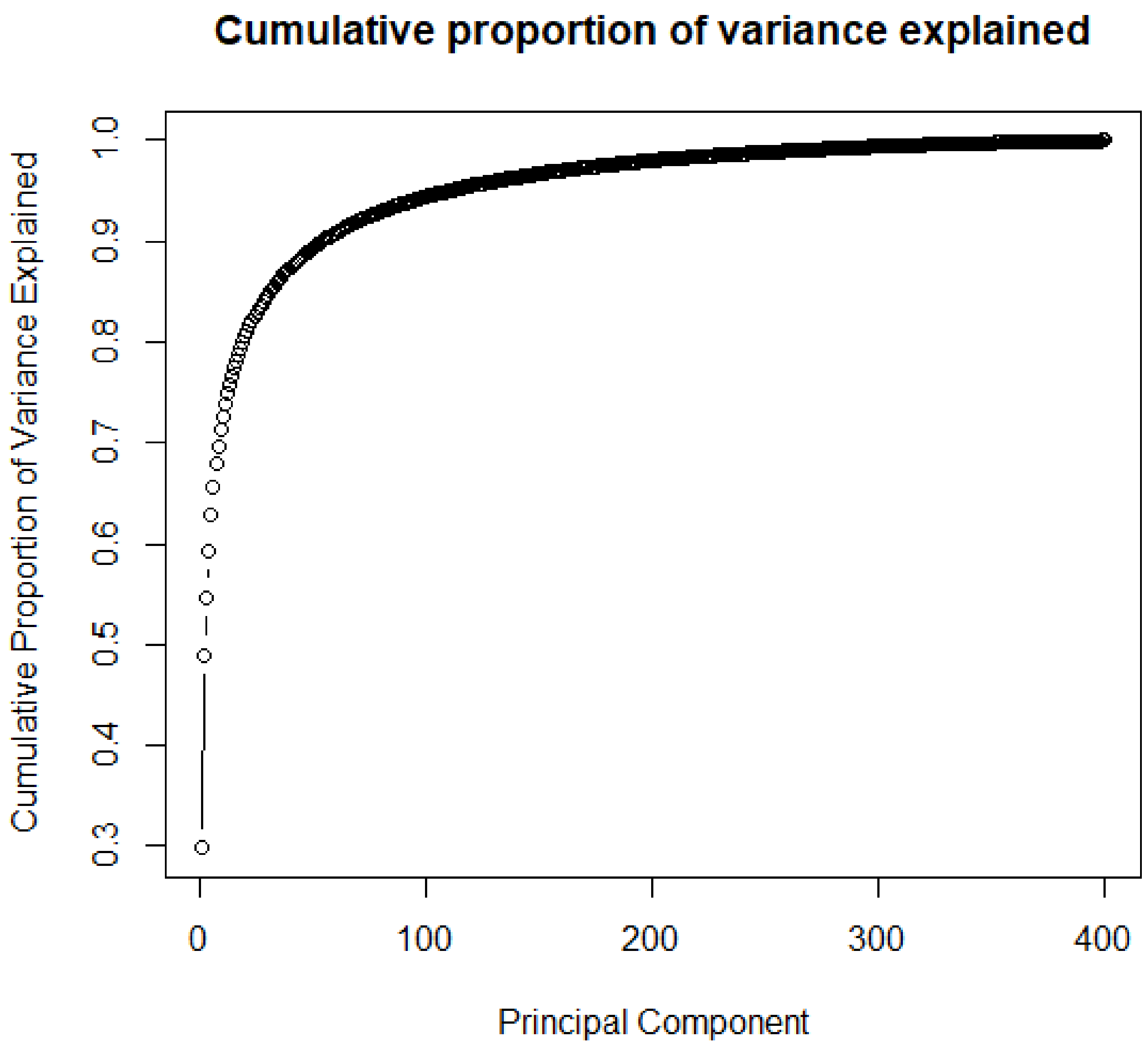


Example (Brazilian dataset)



Don't forget the mean!




Cumulative proportion of variance explained
Cumulative Proportion of Variance Explained
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
100
200
300
400
Principal Component

PCA for image data

- A **sample** in our space is **an image** that is presented during the experiment.

- To run PCA, we transform each image in a long vector of pixels (as a row), and stack the rows.
- Each column corresponds to a pixel location in the images.

Tips

- Reduce resolution if it’s high (target ~ 300 * 300)
- Transform to greyscale
- Limit yourself to components explaining 95% of total variance

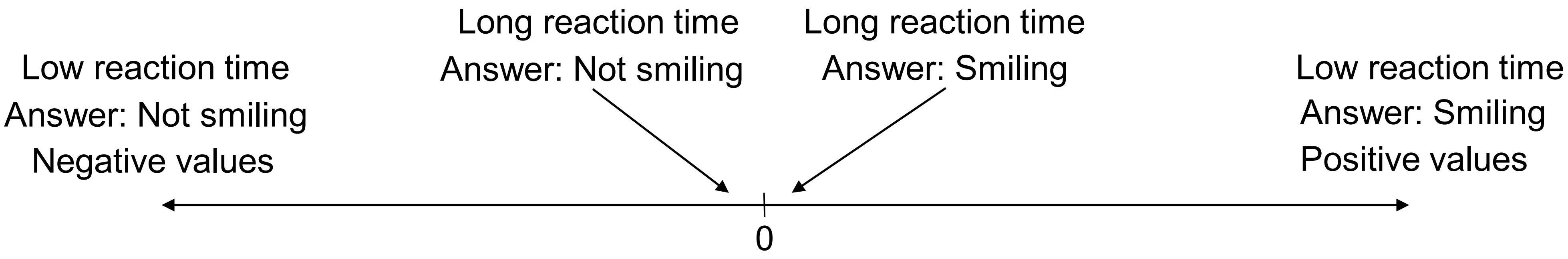
From binary answers to a continuous scale

- During the experiment we collected
  - Binary answers
  - Reaction times

- We want to build a linear regression, so we need to convert the answers to a continuous scale.

- Last time we discussed that:
  - A short reaction time corresponds to certainty in the answer
  - A long reaction time corresponds to uncertainty

From binary answers to a continuous scale



From binary answers to a continuous scale

- Different approaches are possible

- Need to **normalize** by test person because of individual differences
  - Some test persons are generally faster or slower than others

- Also need to consider **outliers**: reaction times above 2-3 seconds
  - They can correspond to the test person getting distracted
  - Best to remove these trials

From binary answers to a continuous scale

- You can look at the paper in the assignment folder for inspiration
  - For each test person, consider their reaction times $t_1, \ldots, t_k$
  - Normalize the reaction times: subtract mean and divide by standard deviation

$$t'_i = \frac{t_i - \mathrm{mean}(t_j)}{\mathrm{std}(t_j)}$$

  - Pool all the normalized times together, split into two answer classes, and map each set linearly
    - from 0 to 1 “smiling”
    - -1 to 0 for “not smiling”

From binary answers to a continuous scale

- Alternative method:
  - For each test person, consider the reaction times $t_1, \ldots, t_k$
  - divide the reaction times by the individual maximum
  - Subtract the value from 1 to obtain

$$t'_i = 1 - \frac{t_i}{\max_1^k(t_j)}$$

  - Multiply the “not smiling” answers by -1

$$s_i = \begin{cases} t'_i, & \text{if ans} = \text{smiling} \\ -t'_i, & \text{otherwise} \end{cases}$$

From binary answers to a continuous scale

Other approaches are possible!

Final step: **aggregate** all the scores for each photo (average)

Linear regression

We want to build a linear regression model

image score ~ PCA-scores

How is this going to work?

How many predictors are there?

What kind of information do the Principal Components encode?

Linear regression

What kind of information do the Principal Components encode?

- Directions with large amounts of variance in the dataset
- The PCs know nothing about our experiment
  - Most likely, we need to select the ones that are useful to predict the score

Feature selection

There are many methods available

- Minimum effort:

    choose the first *k* components

- A bit more effort:
    - Stepwise feature selection with stop criterion:
        - Information Criteria: AIC, BIC
        - RMSE
        - …

Stepwise regression / 1

- Iterative progress to select the predictor variables in a regression model

- At each step, variables are added/removed based on a criterion.
- It has a stop condition. Generally stop when any possible next step does not improve model.

- Often Cross-Validation is used as a criterion (use different training and testing subsets in the iteration)

Feature selection for the homework

- For this project, since you have limited time, it's recommended that you use a standard Matlab function called **sequentialfs**

  https://se.mathworks.com/help/stats/sequentialfs.html

- If sequentially selects features until there is no improvement in prediction.
- Has a built-in Cross-validation step

sequentialfs

- Please make sure that you understand what `sequentialfs` is doing

- You'll need to spend a paragraph in your report to explain it.

- You also need to define a **criterion** function, used in the iterative process to determine whether to add a feature.
  - Example: RMSE

sequentialfs

To use sequentialfs, your code will look something like:

```
criterionfunction = @(xtrain, ytrain, xtest, ytest) DEFINE CRITERION FUNCTION

opt = statset('display','iter');

inmodel = sequentialfs(criterionfunction, model_data, scores', ...
    'cv', 10, ...
    'options',opt,...
    'direction','forward');
```

Example: chosen components

Generate new samples from the model

- Let's say that you end up using k predictors

- Once you have trained your model, you obtain coefficients

  $\beta \in \mathbb{R}^k$ slope

  $\delta \in \mathbb{R}$ intercept

  such that $y = \beta x + \delta$ + error

  where $x \in \mathbb{R}^k$ is the independent variable, the PCA scores.

How to find a sample that gives you a specific value y_0 ?

Generate new samples from the model

One way to do so is to choose a vector x:

$$x = \alpha\beta, \qquad \alpha \in \mathbb{R}$$

And so:

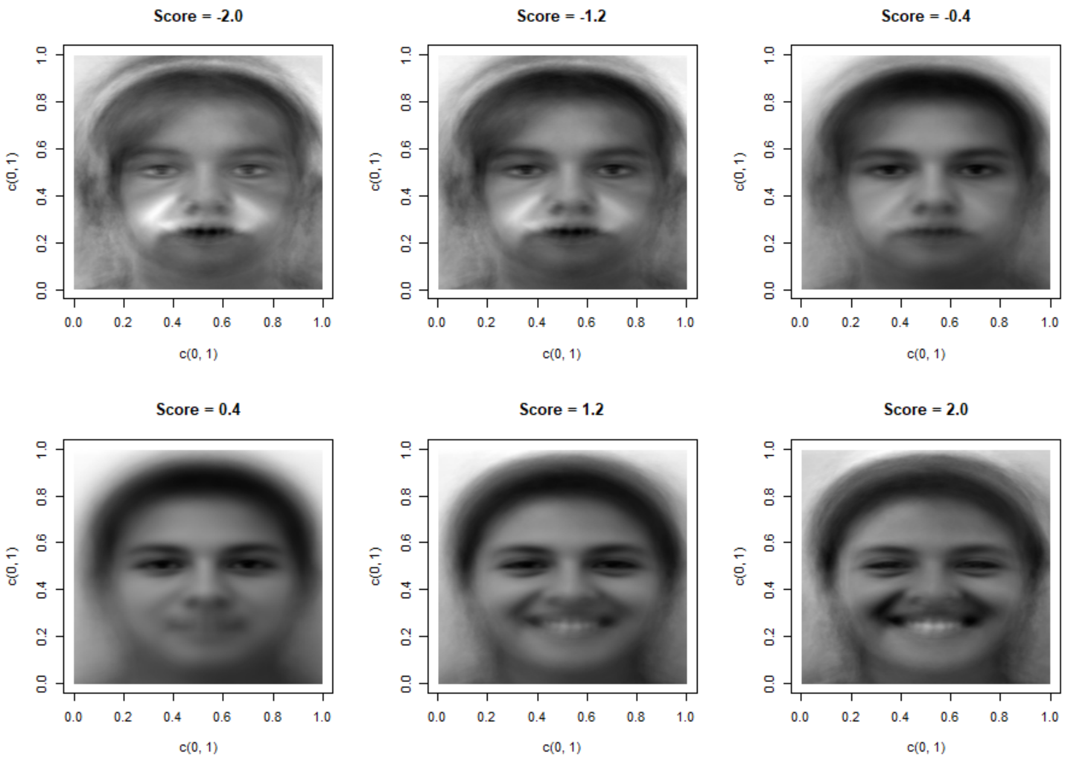
$$y = (\alpha\beta)\beta + \delta = \alpha||\beta||^2 + \delta$$

So, for a value y_0 we can write:

$$\alpha = \frac{y_0 - \delta}{||\beta||^2}$$

Generate new samples from the model

- This way we can choose a set of values y_i along the continuous scale for our experiment’s variable, and generate corresponding faces $x_i = \alpha_i \beta$



Remember to add the mean back!

Comparison of generated faces

Score -0.4 | Score 0.0 | Score 0.4 | Score 0.8 | Score 1.2

Naïve model (first 20 components)

Score -0.4 | Score 0.0 | Score 0.4 | Score 0.8 | Score 1.2

Model with selected features (15 chosen)

Other ideas for the report

- If you have time and want to do something extra, you could try looking into the following questions:
  - What faces have the highest / lowest values for your chosen scale?
  - Display the components chosen in your final model
  - What is the cumulative percentage of the variance explained by the first *k* components?
  - How are reaction times distributed?