1. Monotone Smoothing and Infant Growth Data
   The `onechild` object in the `fda` library contains the length of a new-born’s tibia, measured daily. We will be interested in both the original length and in acceleration, so you should penalize the fourth derivative throughout this question.

   (a) Fit the data by least squares, choosing $\lambda$ by GCV. Is your function monotone increasing?
   (b) Try increasing $\lambda$ until the derivative of the function is always positive at the observation times. How well does this fit the data?
   (c) Re-fit with a monotone smooth; choose $\lambda$ to provide a reasonable fit to the data.
   (d) Plot the velocity curves; what is the maximum rate of growth?

2. PCA of Handwriting Data
   The `handwrit` object in the `fda` library contains the handwriting data analyzed in class. We will investigate performing functional principle components analysis on it.

   (a) First, smooth these data with a B-spline basis and a second-derivative penalty. Select a reasonable number of knots given the nature of the data and the number of observations. You may choose a smoothing parameter as any reasonable value. You should expect this to be small.
   (b) Carry out a principle components analysis of the vertical co-ordinate. Does this analysis support the contention that the most important source of variation is the over-all size of the letters? Are the other components interpretable?
   (c) Now conduct a principle components analysis on the bivariate data. How many components are necessary to explain 90% of the variation? Interpret the leading components, including a plot of the mean writing with variation in this components around it.
   (d) Extract the vertical direction from the bivariate principal components, are these different from the components found from those found by the univariate analysis? Are they close to being orthogonal – you can evaluate this on a grid, or use the function `inprod`.
   (e) Conduct a bivariate principal components analysis using just the first 19 data points. Use these components to try to reconstruct the vertical direction of the 20th as though you have only observed the horizontal direction. Is your reconstruction better than just predicting the mean?
3. **Pincforce Data**

The object `pinch` contains 20 replications of a subject pinching between their thumb and forefinger. For each replicate, the force of the pinch was recorded at 151 time points.

(a) Conduct a principal components analysis of these data. How many components do you need to recover 90% of the variation? Do the components appear satisfactory?

(b) Try a smoothed PCA analysis. Choose the smoothing parameter by cross-validation. Plot the cross-validation curve. Plot the new smoothed principal components. Does this appear to be more satisfactory? Can you interpret the principle components.

(c) Apply a varimax rotation to the smoothed principle components. Does this rotation change your interpretation?

4. **Medfly Data**

The medfly data have been a popular dataset for functional data analysis. You can find them in the `medfly.Rdata` file on the class website. The medfly data consist of records of the number of eggs laid by 50 fruit flies on each of 31 days, along with each individual’s total lifespan.

(a) Smooth the data for the number of eggs, choosing the smoothing parameter by GCV. Plot the smooths.

(b) Conduct a principal components analysis using these smooths. Are the components interpretable? How many do you need to retain to recover 90% of the variation. If you believe that smoothing the PCA will help, do so.

(c) Perform a functional linear regression to predict the total lifespan of the fly from their egg laying. Choose a smoothing parameter by cross validation, and plot the co-efficient function along with confidence intervals.

(d) Conduct a permutation test for the significance of the regression. Calculate the $R^2$ for your regression.

(e) Try a linear regression of lifespan on the principal component scores from your analysis. What is the $R^2$ for this model? Does `lm` find that the model is significant? Reconstruct and plot the co-efficient function for this model along with confidence intervals. How does it compare to the model obtained through functional linear regression?

5. **Your own data**

Conduct a principal components analysis of your own data. How many components do you need to explain the data well? Do they uncover interpretable patterns? Is using smoothing helpful in this case? If canonical correlation analysis is appropriate, or you have a natural scalar response, carry out this analysis.