Homework 2: Fun with Trends and Detrending

36-467, Fall 2018

Due at **12 noon** on **Thursday, 13** September 2018

AGENDA: Getting a feel for how linear smoothers work; understand properties of detrended data.

1. *What do splines like?* Refer back to the Kyoto cherry blossom data from Homework 1. Fit the spline again, and call it `kyoto.spline` (as in the solutions to HW 1).

   (a) (1) The value of λ (the penalty on curvature) selected by cross-validation is stored in the `$\lambda` element of `kyoto.spline`. What is it?

   (b) (6) `smooth.spline` unfortunately does not store the smoother matrix $\mathbf{w}$, though it does store its diagonal in the `$lev` component. We can recover the whole of $\mathbf{w}$ re-fitting the spline on artificial data. The following code will do it:

```
smoother.matrix <- function(a.spline, x) {
  n <- length(x)
  w <- matrix(0, nrow=n, ncol=n)
  for (i in 1:n) {
    y <- rep_len(0, n)  # Equivalent to rep(0, length.out=n) but faster
    y[i] <- 1
    w[,i] <- fitted(smooth.spline(x, y, lambda=a.spline$lambda))
  }
  return(w)
}
```

   (This code is also in a file on the class homepage.)

   Run this function on the spline you fit from the data, and a suitable choice `x`. Check that it isn't doing something obvious wrong by checking the dimensions of the resulting matrix (what should they be?), and by checking that the diagonal of this matrix matches `kyoto.spline$lev`. (For that latter, the `all.equal` function is helpful.) In this problem, be explicit about the code you are using.

   (c) (4) Fix any matrix $\mathbf{w}$, and let $\mathbf{e}_j$ be the $n \times 1$ matrix which is 0 everywhere, except in row j, where it is 1. Explain why $\mathbf{w}\mathbf{e}_j$ gives the j^{th} column of $\mathbf{w}$. Explain how this relates to the code in the previous problem.

(d) (5) Make a plot of all of the eigenvalues (not vectors) of the smoother matrix, in order. How many of them are > 0.95? How many are > 0.1? How many are > 0.01? Are any of them exactly zero?

(e) (5) Following the example in the notes for lectures 2 and 3, make a plot of the first ten eigenvectors of the smoother matrix. Describe the kinds of patterns in the data these eigenvectors capture.

(f) (3) Make a similar plot for the last ten eigenvectors. Describe the kinds of patterns they capture.

(g) (2) What sorts of patterns will show up in the fitted values of the spline (= the estimate of the trend)? What sorts of patterns will show up in its residuals (= the estimate of the fluctuations)?

(h) (3) The example in the notes for the previous problem plots the entries in the eigenvectors against their position ("index") in the vector. Here, different positions correspond to different calendar years. Redo the plot from problem 1e so the eigenvectors are plotted against the `Year.AD` variable. Does this change your description of any of the patterns?

(i) (3) Use the `contour` function to make a contour plot of the smoother matrix. (You will probably want to adjust some of the default settings of `contour`.) Explain how this plot relates to the idea that the spline is doing a kind of local averaging.

2. *Degrees of freedom and moving averages* Assume, for simplicity, no missing values in the data.

   (a) (5) Suppose we have one-dimensional data, and estimate the trend by averaging each observation with its k nearest neighbors. How many degrees of freedom does the smoother matrix $\mathbf{w}$ have?

   (b) (4) Assume we have two-dimensional data, laid out in a regular grid, and estimate the trend by averaging each observation with its k nearest neighbors in every direction. How many degrees of freedom does $\mathbf{w}$ have?

   (c) (1) Assume we have four-dimensional data (3D space + time), with measurements taken at irregular points and times, and we estimate the trend by averaging each observation with its k nearest neighbors in 4D. How many degrees of freedom does $\mathbf{w}$ have?

3. *The Yule-Slutsky effect* In this problem, assume that $X(t) = \mu(t) + \epsilon(t)$, that $\hat{\mu}(t) = \frac{1}{3}\sum_{i=t-1}^{t+1} X_i$, and that $\hat{\epsilon}(t) = X(t) - \hat{\mu}(t)$.

   (a) (3) Find an expression for $\mathbb{E}\left[\hat{\epsilon}(t)\right]$ in terms of the μ's.

   (b) (5) Find an expression for $\operatorname{Var}\left[\hat{\epsilon}(t)\right]$.

(c) (5) Find an expression for $\mathrm{Cov}\left[\hat{\epsilon}(t), \hat{\epsilon}(t+1)\right]$.

(d) (5) Now further assume that $\mathrm{Var}\left[\epsilon(t)\right] = \sigma^2$, and that $\mathrm{Cov}\left[\epsilon(t), \epsilon(s)\right] = 0$ when $t \neq s$. Find $\mathrm{Cov}\left[\hat{\epsilon}(t), \hat{\epsilon}(t+1)\right]$. Explain why the de-trended residuals are correlated, even though the true fluctuations are not.

4. *Detrending by differencing* We have focused on detrending by smoothing, where we first estimate the trend, and then subtract it off. An alternative procedure is to remove trends by taking *differences* between nearby observations. This problem explores how this works, in a situation where we have one coordinate, so our data can be written $X(t) = \mu(t) + \epsilon(t)$. Assume t is discrete, so it can be $1, 2, \ldots$, and that there are no missing values.

   Define $\Delta(t)$ as $X(t) - X(t-1)$.

   (a) (5) Write $\Delta(t)$ in terms of the μ's and ϵ's. Does it make sense to view $\Delta(t)$ as an estimate of $\epsilon(t)$?

   (b) (5) Find the expectation and variance of $\Delta(t)$ in terms of μ, and the variance and covariance of ϵ.

   (c) (5) Explain why $\Delta(t)$ can be said to be "detrended", if μ changes slowly.

   (d) (5) Find an expression for $\mathrm{Cov}\left[\Delta(t), \Delta(t+1)\right]$.

   (e) (5) *Yule-Slutsky again* Assume $\mathrm{Var}\left[\epsilon(t)\right] = \sigma^2$, $\mathrm{Cov}\left[\epsilon(t), \epsilon(s)\right] = 0$ (unless $t = s$). What is $\mathrm{Cov}\left[\Delta(t), \Delta(t+1)\right]$?

   (f) *Differencing random walks* In a **random walk** process, $X(t) = X(t-1) + \eta(t)$, and the η all have constant variance and are uncorrelated with each other.

      i. (2) Show that differencing a random walk gives us exactly the η's, i.e., that here $\Delta(t) = \eta(t)$.

      ii. (3) Explain how this relates to the previous parts of the problem. *Hints:* In a random walk, what is $\mathrm{Cov}\left[X(t), X(t-1)\right]$? How are the ϵ's related to the η's?

RUBRIC (10): The text is laid out cleanly, with clear divisions between problems and sub-problems. The writing itself is well-organized, free of grammatical and other mechanical errors, and easy to follow. Questions which ask for a plot or table are answered with both the figure itself and the command (or commands) use to make the plot. Plots are carefully labeled, with informative and legible titles, axis labels, and (if called for) sub-titles and legends; they are placed near the text of the corresponding problem. All quantitative and mathematical claims are supported by appropriate derivations, included in the text, or calculations in code. Numerical results are reported to appropriate precision. Code is properly integrated with a tool like R Markdown or knitr, and both the knitted file and the source file are submitted. The code is indented, commented, and uses meaningful names. All code is relevant, without dangling or useless commands. All parts of all problems are answered with coherent sentences, and raw computer code or output are only shown when explicitly asked for.