DATA130013: Homework 2

Please hand in the pdf of your solution via elearning on March 21, 2022.

1. Shumway's book (4th ed.) Problems 2.6, 2.7, 2.8, 2.10, and 2.11.

2. Cubic spline is piecewise cubic polynomial within each interval, and the function itself, its first and second derivatives are all continuous at the knots. A natural cubic spline adds additional constraints, namely that the function is linear beyond the boundary knots.

   Suppose that g is the natural cubic spline interpolant to the data $\{x_i, y_i\}_{i=1}^n$ where $a < x_1 < \cdots < x_n < b$. Let $\widetilde{g}$ be any other differentiable function on $[a, b]$ that interpolates the same data set.

   (a) Let $h(x) = \widetilde{g}(x) - g(x)$. Use integration by parts and the fact that g is a natural cubic spline to show that

   $$\int_a^b g''(x)h''(x)\,dx = -\sum_{j=1}^{n-1} g'''(x_j^+)\{h(x_{j+1}) - b(x_j)\} = 0.$$

   (b) Hence show that

   $$\int_a^b \widetilde{g}''(t)^2\,dt \geq \int_a^b g''(t)^2\,dt,$$

   and that equality can only hold if h is identically zero in $[a, b]$.

   (c) Consider the penalized least squares problem

   $$\min_f \left\{ \sum_{i=1}^n (y_i - f(x_i))^2 + \lambda \int_a^b f''(t)^2\,dt \right\}.$$

   Use (b) to argue that the minimizer must be a cubic spline with knots at each of the x_i.