EE 378A Statistical Signal Processing	Professor Andrea Montanari

Homework 1

Due on: 04/18/2022

Problem 1

Equivalence between different types of estimators
Adapted from Exercise 1.11 from A. Tsybakov, Introduction to Nonparametric Estimation
Consider the regression model under the following assumptions:

(i) We consider the nonparametric regression model

$$Y_l = f(t_l) + \sigma z_l,\ l \in \{0, \dots, n-1\},$$

where f is a function from $[0,1]$ to $\mathbb{C}$. The random variables z_l are i.i.d. complex Gaussian with $z_l \sim \mathcal{N}_{\mathbb{C}}(0,1)$ and $t_l = l/n$ for $l \in \{0, \dots, n-1\}$.

(ii) $(\phi_q(t))_{q\in\mathbb{Z}}$ is the trigonometric basis:

$$\phi_q(t) = e^{i2\pi qt}$$

(iii) The Fourier coefficients $b_q = \int_0^1 f(t)\phi_q(t)dt$ of f satisfy

$$\sum_{q\in\mathbb{Z}} |b_q| < \infty$$

The smoothing spline estimator $f_n^{\mathrm{sp}}(t)$ is defined as a solution of the following minimization problem:

$$f_n^{\mathrm{sp}} = \operatorname*{argmin}_{f\in W} \left[\frac{1}{n} \sum_{l=0}^{n-1} |Y_l - f(t_l)|^2 + \kappa \int_0^1 |f''(t)|^2 \, dt \right], \tag{1}$$

where $\kappa > 0$ is a smoothing parameter and W is one of the sets of functions defined below.

(a) First suppose that W is the set of all the functions $f : [0,1] \to \mathbb{C}$ such that f' is absolutely continuous. Prove that the estimator f_n^{sp} reproduces polynomials of degree ≤ 1 if $n \geq 2$ (i.e. if $f(t) = \alpha t + \beta$ for some $\alpha, \beta \in \mathbb{C}$ and $\sigma = 0$ then $f_n^{\mathrm{sp}}(t) = f(t)$).

(b) Suppose next that W is the set of all the functions $f : [0,1] \to \mathbb{C}$ such that

(i) f' is absolutely continuous and

(ii) the periodicity condition is satisfied: $f(0) = f(1), f'(0) = f'(1)$.

Prove that the minimization problem (1) is equivalent to:

$$\min_{\{b_q\}_{q\in\mathbb{Z}}} \sum_{q\in\mathbb{Z}} \left(-2\mathrm{Re}\left(\hat{\theta}_q b_q^*\right) + |b_q|^2 \left(\kappa |a_q|^2 + 1\right) \left[1 + O(n^{-1})\right] \right), \tag{2}$$

where b_q are the Fourier coefficients of f, $\hat{\theta}_q = \frac{1}{n}\sum_{l=0}^{n-1} Y_l \phi_q^*(t_l)$, and a_q is defined as $a_q = -(2\pi q)^2$.

Bonus: prove that the term $O(n^{-1})$ is uniform in $\{b_q\}_{q\in\mathbb{Z}}$, namely that there exists a constant $C > 0$ that does not depend on $\{b_q\}_{q\in\mathbb{Z}}$ and the modulus of the term is bounded by C/n.

(c) Assume now that the term $O(n^{-1})$ in (2) is negligible. Formally replacing it by 0, find the solution of (2) and conclude that the periodic spline estimator is approximately equal to a weighted projection estimator:

$$f_n^{\mathrm{sp}}(x) \approx \sum_{q \in \mathbb{Z}} \lambda_q \hat{\theta}_q \phi_q(t)$$

with weights λ_q written explicitly.

(d) Use (c) to show that for sufficiently small κ the spline estimator f_n^{sp} is approximated by the kernel estimator:

$$f_n(t) = \frac{1}{nh} \sum_{l=0}^{n-1} Y_l K\left(\frac{t_l - t}{h}\right),$$

where $h = \kappa^{1/4}$ and K is the Silverman kernel:

$$K(u) = \int_{-\infty}^{\infty} \frac{\cos(2\pi t u)}{1 + (2\pi t)^4} dt.$$