TD 3 – Inverse problems and approximation

Mathematics of data

02/10/24

Exercise 1.

Part 1. Let $f_0 \in \mathbb{R}^n$ be a discrete signal. We observe the signal $y \coloneqq f_0 \star h + \varepsilon$, where $h \in \mathbb{R}^n$ is a low-pass filter and $\varepsilon \in \mathbb{R}^n$ is noise. For some $\lambda \in \mathbb{R}_+^*$, we look for

$$\operatorname{argmin}_{f \in \mathbb{R}^n} \frac{1}{2} \|f \star h - y\|^2 + \frac{\lambda}{2} \|f\|^2 .$$

1. Using the Fourier decomposition, prove that for an optimal $f \in \mathbb{R}^n$,

   $$\hat{f}_k = \frac{\overline{\hat{h}_k}}{|\hat{h}_k|^2 + \lambda} \hat{y}_k,$$

   where $\overline{\hat{h}_k}$ denotes the complex conjugate of $\hat{h}_k$.

2. Why does $\lambda > 0$ improve the deconvolution in presence of noise?

Part 2. For $f \in \mathbb{R}^n$, denote $Gf \coloneqq (f_i - f_{i-1})_i$ (considering indexes modulo n).

1. What is the adjoint operator $G^\top : \mathbb{R}^n \to \mathbb{R}^n$ for the canonical inner product? (i.e. $\langle Gf, u\rangle = \langle f, G^\top u\rangle$)
2. Show that G, $G^\top$ and $L \coloneqq GG^\top$ are discrete convolution operators and give their associated filters g, $\tilde{g}$ and ℓ.
3. Compute the discrete Fourier coefficients $\hat{g}$. Express $\hat{\tilde{g}}$ and $\hat{l}$ as a function of $\hat{g}$.

Part 3. We now consider the problem

$$\operatorname{argmin}_{f \in \mathbb{R}^n} \frac{1}{2} \|f \star h - y\|^2 + \frac{\lambda}{2} \|Gf\|^2 .$$

1. Using the Fourier decomposition and the results from part 2, prove that for an optimal $f \in \mathbb{R}^n$,

   $$\hat{f}_k = \frac{\overline{\hat{h}_k}}{|\hat{h}_k|^2 + \lambda \hat{\ell}_k} \hat{y}_k.$$

2. How does this expression differ from part 1?

Exercise 2. For an arbitrary $y \in [0, 1]$, consider $f \coloneqq 1_{[y,1]} \in \mathrm{L}^2([0, 1])$, the indicator function of the interval $[y, 1]$. Denote $\langle f, g\rangle \coloneqq \int_0^1 fg$ the inner product and $\|f\|^2 \coloneqq \langle f, f\rangle$.

1. For $M \in \mathbb{N}^*$, denote $\theta_k \coloneqq \sqrt{M} 1_{[\frac{k}{M}, \frac{k+1}{M}[}$ for $0 \le k < M$. Show that $(\theta_k)_k$ is an orthonormal family and give the expression for the linear approximation

   $$f_M \coloneqq \sum_k \langle f, \theta_k\rangle \theta_k.$$

2. Bound $\|f - f_M\|$ as a function of M, independently of y, using a bound as sharp as possible.
3. Denote $\theta \coloneqq 1_{[0,1]}$ and $\psi \coloneqq 1_{[0,1/2[} - 1_{[1/2,1[}$ (it is called the Haar wavelet). Denote, for $j \le 0$, and $0 \le n < 2^{-j}$ the wavelet functions as $\psi_{j,n} \coloneqq \dfrac{1}{2^{j/2}} \psi(2^{-j}x - n)$. Draw the wavelets $\psi_{-1,1}$ and $\psi_{-2,3}$.

4. For some $j_{\min} < 0$, show that

$$\{\theta\} \cup \{\psi_{j,n} : 0 \geq j \geq j_{\min} \text{ and } 0 \leq n < 2^{-j}\}$$

is an orthogonal family. What is the space spanned by this family ?

5. For each j, what is the set Σ_j of index n where $\langle f, \psi_{j,n} \rangle$ is non-zero? For these $n \in \Sigma_j$, bound $|\langle f, \psi_{j,n} \rangle|$ as a function of j.

6. For $T > 0$ define the non-linear approximation of f as

$$\hat{f} \coloneqq \langle f, \theta \rangle \theta + \sum_{|\langle f, \psi_{j,n} \rangle| > T} \langle f, \psi_{j,n} \rangle \psi_{j,n}.$$

Bound as a function of T the number M of non-zero coefficients

$$M \coloneqq |\{(j,n) : |\langle f, \psi_{j,n} \rangle| > T\}|.$$

7. Defining $j_0 \coloneqq \lfloor \log_2(T^2) \rfloor$ a cutoff scale, we define an approximation using

$$\tilde{f}_T \coloneqq \langle f, \theta \rangle \theta + \sum_{j \geq j_0, n \in \Sigma_j} \langle f, \psi_{j,n} \rangle \psi_{j,n}.$$

Show that $\left\| f - \hat{f}_T \right\| \leq \| f - \tilde{f}_T \|$.

8. Bound $\| f - \tilde{f}_T \|$ as a function of j_0 and then as a function of M . Compare the decay with M of $\| f - f_M \|$ and $\left\| f - \hat{f}_T \right\|$.

If you have finished all the exercises, you can move on to the TP3 on github.com/vcastin/teaching

Solutions

Exercise 1.

Part 1

1. Since the discrete Fourier basis is orthogonal, we can separate the problem into the following subproblems:
$$\arg\min_{\hat{f}_k} \frac{1}{2}|\hat{f}_k \hat{h}_k - \hat{y}_k|^2 + \frac{\lambda}{2}|\hat{f}_k|^2.$$
At optimality we have
$$\hat{f}_k|\hat{h}_k|^2 - \hat{y}_k\overline{\hat{h}_k} + \lambda \hat{f}_k = 0,$$
So
$$\hat{f}_k = \frac{\overline{\hat{h}_k}}{|\hat{h}_k|^2 + \lambda}\hat{y}_k.$$
2. When $\lambda = 0$, $\left\|\hat{f}_k - \hat{f}_{0,k}\right\| = \left|\frac{\hat{\varepsilon}_k}{\hat{h}_k}\right|$. The noise is thus amplified for small values of $\hat{h}_k$, which is the case in high-frequency since h is a low-pass filter.

Part 2

1. By change of variable,
$$\langle Gf, u\rangle = \sum_i (f_i - f_{i-1})u_i = \sum_i f_i u_i - \sum_i f_i u_{i+1} = \sum_i f_i(u_i - u_{i+1})$$
So $G^\top u = (u_i - u_{i+1})_i$.
2. $g = [1, -1, \dots 0, 0]$, $\tilde{g} = [1, 0 \dots 0, -1]$, and $\ell = [2, -1, 0 \dots 0, -1]$. One can notice $\tilde{g}_i = g_{-i}$ and $\ell_i = g_i + \tilde{g}_i = g_i \star \tilde{g}_i$
3. $\hat{g}_k = 1 - e^{-\frac{2ik\pi}{n}}$, $\hat{\tilde{g}}_k = 1 - e^{\frac{2ik\pi}{n}} = \overline{\hat{g}_k}$, and $\hat{\ell}_k = |\hat{g}_k|^2 = 4\sin(k\pi/n)$.

Part 3

1. Since the discrete Fourier basis is orthogonal, we can separate the problem into the following subproblems:
$$\arg\min_{\hat{f}_k} \frac{1}{2}|\hat{f}_k \hat{h}_k - \hat{y}_k|^2 + \frac{\lambda}{2}|\hat{f}_k \hat{g}_k|^2.$$
At optimality we have
$$\hat{f}_k|\hat{h}_k|^2 - \hat{y}_k\overline{\hat{h}_k} + \lambda \hat{f}_k \hat{\ell}_k = 0,$$
So
$$\hat{f}_k = \frac{\overline{\hat{h}_k}}{|\hat{h}_k|^2 + \lambda\hat{\ell}_k}\hat{y}_k.$$
2. This penalizes the high frequencies.

Exercise 2.

1. Let $I_k \coloneqq \mathbf{1}_{[\frac{k}{M}, \frac{k+1}{M}[}$. We have $\theta_k = \sqrt{M} I_k$, so on the one hand, $\langle \theta_k, \theta_k\rangle = M\int_{I_k} 1 = 1$ and on the other hand, if $k \neq k'$, $\langle \theta_k, \theta_{k'}\rangle = 0$ because of disjoint supports.
Let $x \in [0, 1[$ and (k, k') such that $x \in I_k$ and $y \in I_{k'}$. We denote $y = \frac{k'+t}{M}$, with $t \in [0, 1[$.
$$f_M(x) = \langle f, \theta_k\rangle\, \theta_k(x) = \sqrt{M}\,\langle f, \theta_k\rangle = M\int_{I_k} f.$$
$$\begin{cases} \text{If } k < k', f_M(x) = 0 \\ \text{If } k = k', f_M(x) = M(1-t)/M = 1 - t \\ \text{If } k > k', f_M(x) = 1 \end{cases}$$

2. We have
$$\|f - f_M\|^2 = \int_{I_k} |f(x) - (1-t)|^2 \, \mathrm{d}x = t(1-t)/M \leq \frac{1}{4M}.$$
3. Drawing
4. If the scales j and j' are different then the inner product vanishes. If $j = j'$ but $k \neq k'$ then the supports are disjoint.

   The space of functions which are constant on a grid of scale $2^{j_{\min}-1}$. Indeed it is included in that space and has $2^{1-j_{\min}}$ dimensions.
5. Let $I_{j,k}$ be the support of $\psi_{j,k}$. For a given j, let k be such that $y \in I_{j,k}$. Then $\Sigma_j = \{k\}$.
$$|\langle f, \psi_{j,n} \rangle| \leq 2^{-j/2} \int_{I_{j,k}} |f||\psi(2^{-j} \cdot -n)| \leq 2^{j/2}$$
   because $|f|, |\psi| \leq 1$
6. One has $|\langle f, \psi_{j,n} \rangle| < T$ if $2^{j/2} < T$ ie $j < \log_2(T^2)$ so that $M \leq \lceil |\log_2(T^2)| \rceil$
7. $\hat{f}_T$ is the best M-term approximation, since $\tilde{f}_T$ uses fewer than M coefficients one has the result.
8. One has
$$\left\|f - \tilde{f}_T\right\|^2 = \sum_{j_{\min} \leq j < j_0, n \in \Sigma_j} |\langle f, \psi_{j,n} \rangle|^2 \leq \sum_{j_{\min} \leq j < j_0, n \in \Sigma_j} (2^{j/2})^2 = 2^{j_0} - 2^{j_{\min}} \leq 2^{j_0} \leq T^2 < 2^{-M}.$$
   The non-linear error decay exponentially fast.