$$\mathrm{Var} \le \frac{C\sigma^2 m}{n}$$

The bias term can be analyzed similarly

White noise ~~problem~~ model (filtering)

$\|f_*\|_k < \infty \qquad f_* : [0,T] \to \mathbb{C}$

periodic RKHS

Observe

$$dY_t = f_*(t)\,dt + \frac{\sigma}{\sqrt{n}}\, dB_t$$

$B_t = (B_t^1 + i B_t^2)/\sqrt{2}$ complex BM

Will assume wlog $T = 1$.

Reason for the scaling?

Imagine partitioning

$[0,1]$ into indervals

$$[0,1] = \bigcup_{\ell=0}^{N-1} \left[\frac{\ell}{N}, \frac{\ell+1}{N}\right] \sim I_\ell$$

$$Y_\ell := \cancel{\oint} \sigma N \cdot \frac{1}{|I_\ell|} \int_{I_\ell} dY_t =$$

$$= N \left\{ \int_{\ell/N}^{\frac{\ell+1}{N}} f^*(t)\,dt + \frac{\sigma}{\sqrt{n}} \int_{\ell/N}^{\ell+1/N} dB_t \right\}$$

$$Y_\ell \overset{\approx}{=} f^*(t_\ell) + \frac{\sigma}{\sqrt{n}} \cdot \sqrt{N}\, z_\ell \qquad z_\ell \sim N(0,1)$$

If n is large enough and $N = n$

$$Y_\ell \approx f^*(t_\ell) + \sigma z_\ell$$

~~Aszbef~~ Similarly to what we did for the discrete model, define

$$\tilde{Y}_q = \int_0^1 \varphi_q(t)^* \, dY_t \qquad q \in \mathbb{Z}$$

$$= \langle f^*, \varphi_q \rangle + \frac{\sigma}{\sqrt{n}} \int_0^1 \varphi_q(t)^* \, dB_t$$

$$= \theta_q + \frac{\sigma}{\sqrt{n}} g_q$$

g_q are normal with

$$\mathbb{E}\, g_q g_p^* = \int_0^1 \varphi_q^*(t) \varphi_p(t) \, dt = \delta_{p,q}$$

Ito isometry

$$\tilde{Y}_q = \theta_q + \frac{\sigma}{\sqrt{n}} g_q$$

$$\left\{ \underbrace{\sum_{q \in \mathbb{Z}} \frac{|\theta_q|^2}{\lambda_q^2}}_{= \|f^*\|_K^2} \leq R^2 \right\} =: \Theta_K(R)$$

Linear minimax risk

$$R_L(\rho,\sigma) = \inf_{M} \sup_{\theta \in \Theta_K(\rho)} \mathbb{E}\left\{ \|\theta - M\tilde{y}\|_{\ell_2}^2 \right\}$$

$$= \inf_{\mathcal{L}} \sup_{\|f\|_K \leq \rho} \mathbb{E}\left\{ \|f - \mathcal{L}\tilde{y}\|_{L^2}^2 \right\}$$

$$R_M(\rho,\sigma) = \inf_{\hat{\theta}} \sup_{\theta \in \Theta_K(\rho)} \mathbb{E}\left\{ \|\theta - \hat{\theta}(\tilde{y})\|_{\ell^2}^2 \right\}$$

$$= \inf_{\hat{f}_y} \sup_{\|f\|_K \leq \rho} \mathbb{E}\left\{ \|f - \hat{f}_y\|_{L^2}^2 \right\}$$

$$R(\rho,\sigma) = \rho^2 R\left(1, \frac{\sigma}{\rho}\right)$$

$$L \quad \rho^2 \bar{R}\left(\frac{\sigma}{\rho}\right)$$

<u>Lemma</u>

$$\bar{R}_L(\sigma) = \min_{\lambda \geq 0} \left\{ \lambda^2 + \frac{\sigma^2}{n} \sum_{i=1}^{\infty} \left(1 - \frac{\lambda}{\lambda_i}\right)_+^2 \right\}$$

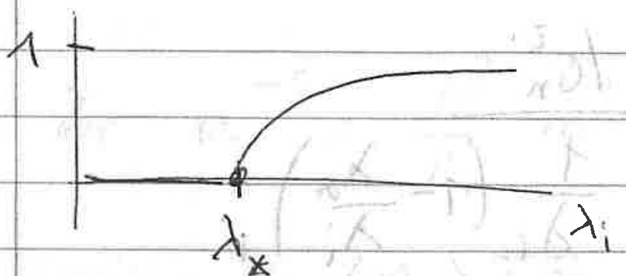
Minimum achieved at solution of

$$\lambda = \frac{\sigma^2}{n} \sum_{i=1}^{\infty} \frac{1}{\lambda_i} \left(1 - \frac{\lambda}{\lambda_i}\right)_+$$

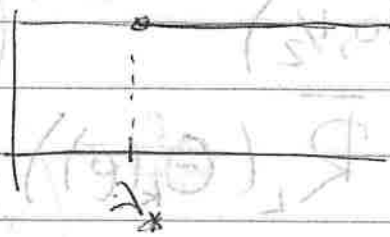
Further

$$\hat{\theta}(\tilde{y}) = M\tilde{y}$$

$$M = \mathrm{diag}\left(\left(1 - \frac{\lambda_*}{\lambda_i}\right)_+\right) \qquad \square$$

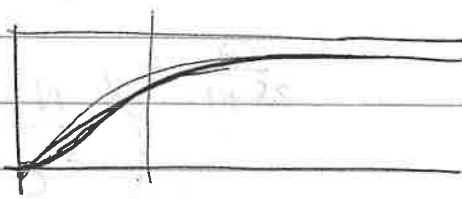


cf Proj



$\mathbb{1}_{\lambda_i \geq \lambda}$

Ridge



$$\frac{\lambda_i^2}{\lambda_*^2 + \lambda_i^2}$$

~~Theo~~

Theorem Define $\sigma_n^2 = \sigma^2/n$

$$\delta(\epsilon) = \lambda_{max}^2 \, e^{-\Lambda_* \epsilon^2/64}$$

$$\Lambda_* = \frac{\lambda_* / \sigma_n^2}{\max_i \frac{1}{\lambda_i}\left(1-\frac{\lambda_*}{\lambda_i}\right)_+}$$

Then $\forall \epsilon \in (0, 1/2)$

$$\bar{R}_M(\Theta_k(\rho)) \leq \bar{R}_L(\Theta_k(\rho)) \leq (1+c_0\epsilon)\bar{R}_M(\Theta_k(\rho)) + c_0\delta(\epsilon)$$

$\square$

I

$$\Lambda_* \approx \frac{\lambda_* n}{\sigma^2}$$

Let us consider the case $\lambda_\ell \approx \ell^{-s}$

$\lambda_* \approx \lambda_m$

$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i}\left(1-\frac{\lambda_*}{\lambda_i}\right)_+ \approx$$

$$\asymp \sum_{\ell=1}^{m} \frac{1}{\lambda_\ell} \asymp m^{1+s}$$

So the condition for λ_* implies

$$\lambda_* m^{-s} \asymp \frac{\sigma^2}{n} \cdot m^{1+s}$$

$$m \asymp \left(\frac{n}{\sigma^2}\right)^{\frac{1}{1+2s}}$$

$$\lambda_* \asymp \left(\frac{n}{\sigma^2}\right)^{-\frac{s}{1+2s}}$$

$$\Rightarrow R_L \asymp \left(\frac{n}{\sigma^2}\right)^{-\frac{2s}{1+2s}}.$$

$$\Lambda_* \asymp \frac{\lambda_*^2 n}{\sigma^2} \asymp \left(\frac{n}{\sigma^2}\right)^{-\frac{2s}{1+2s}} \left(\frac{n}{\sigma^2}\right) =$$

$$= \left(\frac{n}{\sigma^2}\right)^{\frac{1}{1+2s}}$$

$$\delta(\epsilon) = \exp\left\{-\frac{\epsilon^2}{64} n^{\frac{1}{1+2s}}\right\}, \quad R_M \asymp R_L$$

Multiresolution Analysis and Wavelet bases

Problem with Fourier analysis:

- Many natural signals ~~show~~ have ~~non-analiticities~~ non-smooth elements, eg edges in an image

- This leads to slowly decaying Fourier coefficients

- Example $f:[0,1] \to \mathbb{C}$ $f(x) = 1_{x \in [0,\frac{1}{2}]}$

$$a_q = \int_{[0,1]} e^{-2\pi i q x} f(x)\, dx = \int_0^{1/2} e^{-2\pi i q x}\, dx$$

$$= \frac{1}{2\pi i q}\left(e^{-\pi i q} - 1\right) = \frac{1}{2\pi i q} 1_{q \in \text{odd}}$$

$$= \begin{cases} 1/2 & q = 0 \\ 0 & q \text{ even} \\ \frac{1}{2\pi i q} & q \in \text{odd} \end{cases}$$

let us compute the RKHS

$$\|f\|_k^2 = \sum_{q\in\mathbb{Z}} \frac{|a_q|^2}{\lambda_q^2} \qquad \text{assume } \lambda_q, |a_q| \asymp q^{-s}$$

$$\asymp \sum_{q\in\mathbb{Z}} \frac{|q|^{-2}}{(|q|^{-s})^2} \asymp \sum_{n=1}^{\infty} n^{-2+2s}$$

$< \infty$ iff $2-2s > 1$ $\boxed{s < 1/2}$

In particular consider the problem of storing ~~(compre~~ in computer memory the signal f

A natural idea is to truncate Fourier series

$$T_m f = \sum_{|q|\leq m} a_q \varphi_q \qquad \varphi_q(t) := e^{2\pi i q t}$$

and storing the vector $(a_q)_{|q|\leq m}$

$$\|f - T_m f\|_{L^2}^2 = \sum_{|q|>m} a_q^2 \asymp \sum_{|q|>m} |q|^{-2} \asymp \frac{1}{m}$$

Need ~~$1/m^2$~~ ~~$1/\epsilon$ coeffici~~

$M(\epsilon) \asymp 1/\epsilon$ coefficients to achieve err ϵ

Similarly, if $f \in W^{s,2}_{per}([0,T], \mathbb{R})$

$$\|f - T_m f\|_{L^2}^2 = \sum_{|q|>m} |a_q|^2 = \sum_{|q|>m} |q|^{-2\alpha} \asymp m^{-2\alpha+1}$$

$$= \cancel{\sum_{\ell=0}^{\infty} \sum_{2^\ell m \le |q| \le 2^{\ell+1} m} |a_q|^2} = m^{-2s-6}$$

I can take $|a_q| = \cancel{|q|^{s-\frac{1}{2}}}\ |q|^{-\alpha}$

$\alpha > s + \frac{1}{2}$

For smoothness class s I need

$M(\epsilon) \asymp \left(\frac{1}{\epsilon}\right)^{1/2s}$ to achieve error ϵ.

However if we ~~kw~~ know f has only one discontinuity in $[0,T]$, ~~we~~ an is C^2 otherwise, we can use $O(1)$ bits to encode the discont. and achieve the C^2 rate otherwise.

[Wavelet bases achieve this type of goal <u>without</u> knowing a priory how many ~~dist~~ discontinuities etc.

Similar problems for denoising, estimation inverse problems etc].

MRA is a nice way to define / construct wavelets

Idea " localize in time and frequ