Supported by the Austrian Science Fund (FWF)

Bias and variance of Thomson multitaper estimator

Luis Daniel Abreu

NuHAG, Faculty of Mathematics
University of Vienna

Banff
26/06/2022

Der Wissenschaftsfonds.

Spectral density of a stationary process

- We want to learn the spectral density:

$$S(\xi) = \sum_{k \in \mathbb{Z}} r_k e^{2\pi i k \xi} \qquad\qquad r_k = \mathbb{E}(X_j X_{j+k}).$$

from $X(0), X(1), \ldots, X(N-1)$ (one realization).

Spectral density of a stationary process

- We want to learn the spectral density:

$$S(\xi) = \sum_{k \in \mathbb{Z}} r_k e^{2\pi i k \xi} \qquad r_k = \mathbb{E}(X_j X_{j+k}).$$

  from $X(0), X(1), \ldots, X(N-1)$ (one realization).

- Classical assumptions: $X(n)$ ergodic, $\mathbb{E}X_k = 0$, stationary.
- Challenges: Limited observations; Stochastic fluctuations; Single realization.

The periodogram

Data: samples: $X(0), X(1), \ldots, X(N-1)$ (one realization).
Spectral density:

$$S(\xi) = \lim_{N \longrightarrow \infty} \frac{1}{N} \mathbb{E} \left| \sum_{t=0}^{N-1} X(t) e^{2\pi i t \xi} \right|^2 .$$

The periodogram

Data: samples: $X(0), X(1), \ldots, X(N-1)$ (one realization).
Spectral density:

$$S(\xi) = \lim_{N \longrightarrow \infty} \frac{1}{N} \mathbb{E} \left| \sum_{t=0}^{N-1} X(t) e^{2\pi i t \xi} \right|^2 .$$

Periodogram estimator:

$$\widehat{S}(\xi) = \frac{1}{N} \left| \sum_{t=0}^{N-1} X(t) e^{2\pi i t \xi} \right|^2 .$$

The periodogram

Data: samples: $X(0), X(1), \ldots, X(N-1)$ (one realization).
Spectral density:

$$S(\xi) = \lim_{N \longrightarrow \infty} \frac{1}{N} \mathbb{E} \left| \sum_{t=0}^{N-1} X(t) e^{2\pi i t \xi} \right|^2 .$$

Periodogram estimator:

$$\widehat{S}(\xi) = \frac{1}{N} \left| \sum_{t=0}^{N-1} X(t) e^{2\pi i t \xi} \right|^2 .$$

The periodogram is (asymptotically) unbiased:

$$\mathrm{Bias}(\widehat{S}(\xi)) = \mathbb{E}\widehat{S}(\xi) - S(\xi) \longrightarrow 0,$$

as $N \longrightarrow +\infty$.

Performance of the periodogram

- Terrible variance. White gaussian noise case:

$$\mathrm{Var}(\hat{S}(\xi)) \approx S^2(\xi).$$

Performance of the periodogram

- Terrible variance. White gaussian noise case:

$$\mathrm{Var}(\hat{S}(\xi)) \approx S^2(\xi).$$

- Bias: creates non-zero values in non-existing frequencies

$$\mathbb{E}\widehat{S}(\xi) = S * \frac{1}{N} \left| \mathbf{D_N} \right|^2 (\xi).$$

- $\mathbf{D_N}$ poorly concentrated in $[-W, W]$
- Normally the observations are weighted with a *taper window.*

Data tapers

Given samples $X(0), X(1), \ldots, X(N-1)$, we consider a *taper* sequence $(D(t))_{t=0}^{N-1}$.
The tapered estimator is:

$$\widehat{S_d}(\xi) = \left| \sum_{t=0}^{N-1} D(t) X(t) e^{2\pi i t \xi} \right|^2 .$$

Data tapers

Given samples $X(0), X(1), \ldots, X(N-1)$, we consider a *taper* sequence $(D(t))_{t=0}^{N-1}$.
The tapered estimator is:

$$\widehat{S_d}(\xi) = \left| \sum_{t=0}^{N-1} D(t) X(t) e^{2\pi i t \xi} \right|^2 .$$

and

$$\mathbb{E}\widehat{S_D}(\xi) = S * \left| \mathcal{F}D \right|^2 (\xi),$$

where

$$\mathcal{F}D(\xi) = \sum_{t=0}^{N-1} e^{-2\pi i \xi t} D(t)$$

is the Finite Fourier Transform.

How to reduce variance keeping bias low

Tapering introduces Bias from spectral leackage:

$$Bias\left(\widehat{S}_D(\xi)\right) = \mathbb{E}\{\widehat{S}_D(\xi)\} - S(\xi)$$

Does not considerably reduce variance.

How to reduce variance keeping bias low

Tapering introduces Bias from spectral leackage:

$$Bias\left(\widehat{S}_D(\xi)\right) = \mathbb{E}\{\widehat{S}_D(\xi)\} - S(\xi)$$

Does not considerably reduce variance.
The minimal bias would be achieved by an idealized band-pass:

$$\mathcal{F}(D_W)(\xi) = \frac{1}{2W} 1_{[-W,W]}(\xi) \qquad Bias\left(\widehat{S}_{D_W}(\xi)\right) \lesssim W^2$$

Problem:
How to reduce variance keeping Bias close to the optimal W^2?

Slepian sequences

The *discrete prolate spheroidal sequences* $v_t^{(k)}(N,W)$ are the eigenvectors of the equation:

$$\sum_{n=0}^{N-1} \frac{\sin 2\pi W\,(t-n)}{\pi\,(t-n)} v_n^{(k)}(N,W) = \lambda_k(N,W) v_t^{(k)}(N,W).$$

Slepian sequences

The *discrete prolate spheroidal sequences* $v_t^{(k)}(N,W)$ are the eigenvectors of the equation:

$$\sum_{n=0}^{N-1} \frac{\sin 2\pi W\,(t-n)}{\pi\,(t-n)} v_n^{(k)}(N,W) = \lambda_k(N,W) v_t^{(k)}(N,W).$$

Their discrete Fourier transforms

$$U_k(N,W;\xi) = \mathcal{F}(v_t^{(k)}(N,W))\,(\xi)$$

satisfy the Integral equation

$$\int_{-W}^{W} \frac{\sin N\pi(\xi-\xi\prime)}{\sin \pi(\xi-\xi\prime)} U_k(N,W;\xi\prime) d\xi\prime = \lambda_k(N,W) U_k(N,W;\xi\prime),$$

$U_k(N,W;\xi)$ are called 'The Slepians' (*Discrete prolate spheroidal functions*).

Thomson multitaper estimator

Estimators tapered with *prolate* sequence k: $(v_t^{(k)}(N,W))_{t=0}^{N-1}$:

$$\widehat{S_k} = \left| \sum_{t=0}^{N-1} v_t^{(k)}(N,W) X(t) e^{-2\pi i \xi t} \right|^2$$

Thomson averaged estimator (multi-taper):

$$\widehat{S_K}(\xi) = \tfrac{1}{K} \sum_{k=1}^{K} \widehat{S_k}. \qquad \mathrm{Var}(\widehat{S_K}) \lesssim \frac{1}{K}$$

Thomson multitaper estimator

Estimators tapered with *prolate* sequence k: $(v_t^{(k)}(N,W))_{t=0}^{N-1}$:

$$\widehat{S_k} = \left| \sum_{t=0}^{N-1} v_t^{(k)}(N,W) X(t) e^{-2\pi i \xi t} \right|^2$$

Thomson averaged estimator (multi-taper):

$$\widehat{S_K}(\xi) = \tfrac{1}{K} \sum_{k=1}^{K} \widehat{S_k}. \qquad \mathrm{Var}(\widehat{S_K}) \lesssim \frac{1}{K}$$

(A., Romero 2017):

$$\mathrm{Bias}(\widehat{S}_{(K)}(\xi)) \lesssim W^2 + \frac{\log N}{K}.$$

Bias close to optimal (band-pass) W^2.
Agree with Thomson 1982 conjecture.

Intuition behind

Thomson's heuristics: the spectral window

$$\frac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2$$

resembles the ideal band-pass

$$\frac{1}{2W}1_{[-W,W]}(\xi).$$

Intuition behind

Thomson's heuristics: the spectral window

$$\frac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2$$

resembles the ideal band-pass

$$\frac{1}{2W}1_{[-W,W]}(\xi).$$

Thus, the bias

$$\mathbb{E}\widehat{S}_{(K)}(\xi)-S(\xi)=\widehat{S}_{(K)}*\tfrac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2-S(\xi).$$

stays close to the band-pass case

$$\mathbb{E}\widehat{S}_{D_W}(\xi)-S(\xi)=S*\frac{1}{2W}1_{[-W,W]}(\xi)-S$$

The accumulation property of Slepians

The approximation

$$\frac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2 \to \frac{1}{2W}1_{[-W,W]}(\xi)$$

Ressembles the Pythagoras identity

$$\sin^2 x + \cos^2 x = 1.$$

Accumulation property: energy profiles are complementary. Average becomes flat.

The accumulation property

The approximation

$$\frac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2 \to \frac{1}{2W}1_{[-W,W]}(\xi)$$

Ressembles the Pythagoras identity

$$\sin^2 x + \cos^2 x = 1.$$

Accumulation property: energy profiles are complementary.
Average becomes flat.

Accumulation in Pythagoras theorem

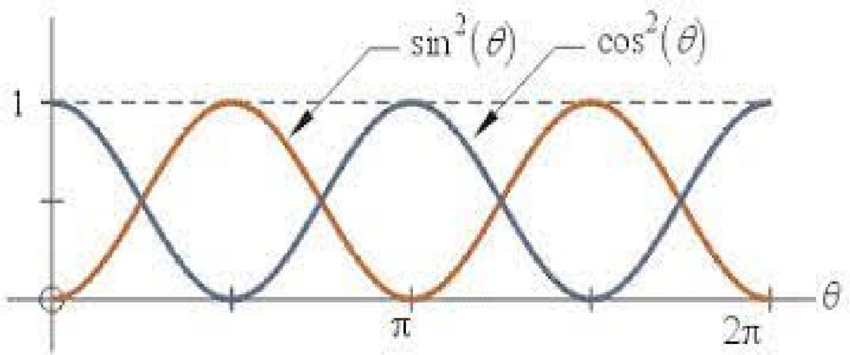
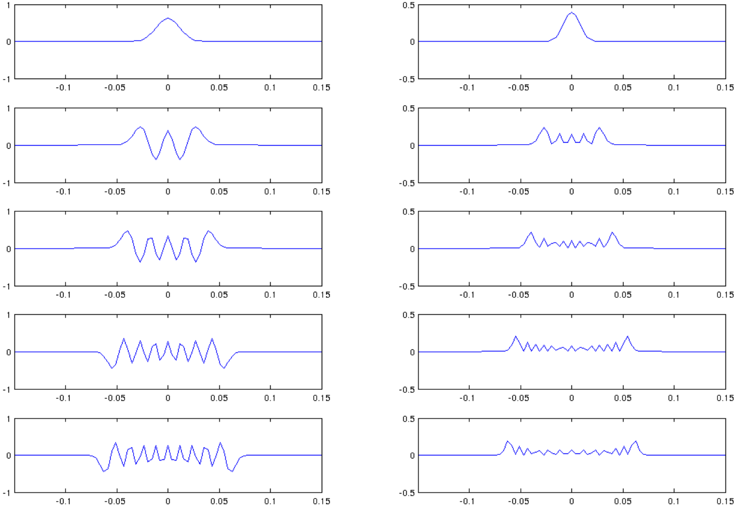
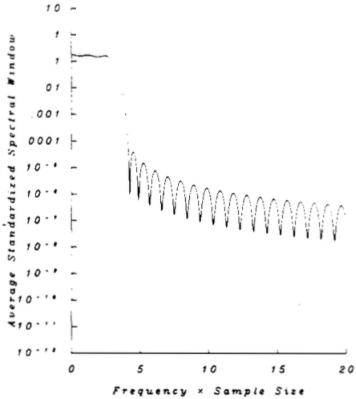


Figure: Energy allocation of sin^2x in the line complements energy allocation of cos^2x; Maxima and zeros of $\sin x$ and $\cos x$ are interlaced.

Fourier transform of Slepian sequences





(Accumulation, low spectral leakage. Thomson, Proc. IEEE, 82)

Bias and MSE estimates

Exceeding bias is measured by the approximation

$$\frac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2 \to \frac{1}{2W}1_{[-W,W]}(\xi)$$

Theorem (A., Romero, 2017)

Let $K := \lfloor 2NW \rfloor$. *Then*

$$\left\| \tfrac{1}{K}\sum_{m=1}^{K}|U_k(N,W;.)|^2 - \frac{1}{2W}1_{[-W,W]}(.) \right\|_{L^1} \lesssim \frac{\log N}{K}.$$

Bias and MSE estimates

Exceeding bias is measured by the approximation

$$\frac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2 \to \frac{1}{2W}1_{[-W,W]}(\xi)$$

Theorem (A., Romero, 2017)

Let $K := \lfloor 2NW \rfloor$. *Then*

$$\left\| \tfrac{1}{K}\sum_{m=1}^{K}|U_k(N,W;.)|^2 - \frac{1}{2W}1_{[-W,W]}(.) \right\|_{L^1} \lesssim \frac{\log N}{K}.$$

Consequently,

$$\mathrm{Bias}(\widehat{S}_{(K)}) \lesssim W^2 + \frac{\log N}{K}.$$

Bias and MSE estimates

Exceeding bias is measured by the approximation

$$\frac{1}{K}\sum_{m=1}^{K}|U_k(N,W;\xi)|^2 \to \frac{1}{2W}1_{[-W,W]}(\xi)$$

Theorem (A., Romero, 2017)

Let $K := \lfloor 2NW \rfloor$. *Then*

$$\left\| \tfrac{1}{K}\sum_{m=1}^{K}|U_k(N,W;.)|^2 - \frac{1}{2W}1_{[-W,W]}(.)\right\|_{L^1} \lesssim \frac{\log N}{K}.$$

Consequently,

$$\mathrm{Bias}(\widehat{S}_{(K)}) \lesssim W^2 + \frac{\log N}{K}.$$

and

$$\mathrm{MSE}(\widehat{S}_{(K)}) = \mathrm{Bias}(\widehat{S}_{(K)})^2 + \mathrm{Var}(\widehat{S}_{(K)}) \lesssim W^4 + \frac{\log^2 N}{K^2} + \frac{1}{K}$$

Averages of eigenvalues

Key step of the proof. Show that with $K := \lfloor 2NW \rfloor$,

$$\left| 1 - \frac{1}{K} \sum_{k=0}^{K-1} \lambda_k(N, W) \right| \lesssim \frac{\log N}{K}.$$

Averages of eigenvalues

Key step of the proof. Show that with $K := \lfloor 2NW \rfloor$,

$$\left| 1 - \frac{1}{K} \sum_{k=0}^{K-1} \lambda_k(N,W) \right| \lesssim \frac{\log N}{K}.$$

- The bound $\frac{\log N}{K}$ is optimal.
- Previous estimates used individual bounds on $\lambda_k(N,W)$.
- Our approach deals with the whole average as a single object.

Accumulation: a common phenomenon

- The accumulation property holds for eigenfunctions in other geometries.
- Time-frequency domains, spherical domains, circular symmetric domains.
- The accumulation depends only on the space (General domains: Andén-Romero)
- The accumulation property holds for point densities of some determinantal point processes (Ginibre: 'The circular law').

References

D. J. Thomson. Spectrum estimation and harmonic analysis. Proc. IEEE, 70(9):1055–1096, September 1982.

L. D. Abreu and J. L. Romero. MSE estimates for multitaper spectral estimation and off-grid compressive sensing. IEEE Trans. Inf. Theory, 63(12):7770–7776, December 2017..

Thanks!