THRESHOLDING SCHEMES FOR CEPSTRAL ANALYSIS

Ciprian Doru Giurcăneanu and Seyed Alireza Razavi

Department of Signal Processing
Tampere University of Technology, Finland
ciprian.giurcaneanu@tut.fi, alireza.razavi@tut.fi

ABSTRACT

In this talk, we discuss how the periodogram can be smoothed by thresholding the estimated cepstral coefficients. Three different approaches are considered for the selection of the threshold: (i) a carefully designed uniformly most powerful unbiased test (UMPUT) [1]; (ii) the Bayesian information criterion (BIC) [1]; (iii) the criterion based on Kolmogorov structure function (KSF) [2].

1. INTRODUCTION

Our talk is focused on estimating the spectrum $\Phi(\omega)$ of a stationary, discrete time, real-valued signal from the measurements $y_0, \dots, y_{N-1}$. With the convention that $\omega_p = (2\pi p)/N$, $p \in \{0, \dots, N-1\}$, are the Fourier frequency grid points, we use the notation Φ_p for $\Phi(\omega_p)$. The estimate of the spectrum at point ω_p is [3]:

$$\hat{\Phi}_p = \frac{1}{N} \left| \sum_{t=0}^{N-1} y_t \exp(-i\omega_p t) \right|^2,$$

where $i = \sqrt{-1}$.

We assume that N, the number of samples, is even and we take $M = N/2 + 1$. Under the hypothesis that $\min\{\Phi_p, \hat{\Phi}_p\} > 0$ for all p, the first M cepstral coefficients and their estimates are given by [1]:

$$\begin{aligned} c_j &= \frac{1}{N} \sum_{p=0}^{N-1} \ln(\Phi_p) \exp(i\omega_j p), \\ \hat{c}_j &= \frac{1}{N} \sum_{p=0}^{N-1} \ln(\hat{\Phi}_p) \exp(i\omega_j p) + \gamma \delta_{j,0}, \end{aligned}$$

where $j \in \{0, \dots, M-1\}$ and $\gamma = 0.577216\dots$ is the Euler constant. The Kronecker indicator $\delta_{j,0}$ takes value one if $j = 0$, and otherwise takes value zero. The rest of the coefficients can be obtained without difficulties because $c_{N-j} = c_j$ and $\hat{c}_{N-j} = \hat{c}_j$ for $j \in \{1, \dots, M-2\}$.

Next we investigate how the periodogram can be smoothed by thresholding the estimated cepstral coefficients [1, 4, 5]. The interested reader can find in [6] more results on the statistical properties of the spectra obtained by cepstral nulling.

2. CEPSTRAL NULLING

Based on the distributional properties of $\hat{\mathbf{c}} = [\hat{c}_0 \dots \hat{c}_{M-1}]^\top$, the following thresholding scheme was introduced in [1]:

$$\check{c}_j = \begin{cases} 0, & |\hat{c}_j| < \mu \left[\mathbf{C}(j+1, j+1)/N\right]^{1/2}, \\ \hat{c}_j, & \text{otherwise,} \end{cases}$$

where $j \in \{0, \dots, M-1\}$ and $\mathbf{C}(j+1, j+1)$ is the $(j+1)$-th diagonal element of the matrix

$$\mathbf{C} = \frac{\pi^2}{6} \begin{bmatrix} 2 & & & & 0 \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ 0 & & & & 2 \end{bmatrix}.$$

We outline below three different methods which have been proposed for the selection of μ.

UMPUT [1] - The formula is derived by combining a uniformly most powerful unbiased test (UMPUT) [7] with some empirical evidence, and it is given by:

$$\mu_{\text{UMPUT}} = (5 - \mathrm{I}_{\text{st}}) + \frac{N - 128}{1920}.$$

Its application is restricted to data sets for which N is an integer power of two and $128 \leq N \leq 2048$. The value of I_{st} depends on the type of the signal and is chosen as follows:

- $\mathrm{I}_{\text{st}} = 1$ for broadband signal with small dynamic range;
- $\mathrm{I}_{\text{st}} = 2$ for broadband signal with medium dynamic range;
- $\mathrm{I}_{\text{st}} = 3$ for narrowband signal with large dynamic range.

BIC [1] - In this case, the choice of the threshold relies on the Bayesian information criterion (BIC) from [8]. The expression of μ is:

$$\mu_{\text{BIC}} = 1 + (\ln M)^{1/2}.$$

KSF [2] - The value of μ is selected such that to minimize the Kolmogorov structure function (KSF) as it is formulated in [9]. To give the expression of the KSF, we need some preparations. For $j \in \{0, \ldots, M-1\}$, we define

$$x_j = \hat{c}_j \left[N/\mathbf{C}(j+1, j+1)\right]^{1/2}$$

and

$$q(x_j) = \begin{cases} 0, & |x_j| < 1+\ell \\ \mathrm{sgn}(x_j)\left(1 + 2\ell + 2\ell\left\lfloor \frac{|x_j|-1-\ell}{2\ell} \right\rfloor\right), & \text{otherwise,} \end{cases}$$

where $\ell \geq 0$ is a parameter. The operator $\mathrm{sgn}(\cdot)$ is used for the signum function. We also use the notation $\lfloor u \rfloor$ for the largest integer less than or equal to the real-valued number u.

We define $\eta = \{j : 0 \leq j \leq M-1, q(x_j) \neq 0\}$. The cardinality of η is denoted by k, and for the sake of simplicity, we assume that $0 < k < M$. Let

$$\begin{aligned} \eta &= \{j_0, \ldots, j_{k-1}\}, \\ \mathbf{z} &= [x_{j_0}, \ldots, x_{j_{k-1}}]^\top, \\ \tilde{\mathbf{z}} &= [q(x_{j_0}), \ldots, q(x_{j_{k-1}})]^\top. \end{aligned}$$

With the convention that $h_{\mathbf{c}}(\ell)$ denotes the KSF, we have:

$$h_{\mathbf{c}}(\ell) = L_\eta(\ell) + L_{\tilde{\mathbf{z}}}(\ell) + D_{\mathbf{x}}(\ell),$$

where

$$\begin{aligned} L_\eta(\ell) &= \min\left\{L_\eta^{\mathcal{A}}, L_\eta^{\mathcal{B}}(\ell)\right\}, \\ L_\eta^{\mathcal{A}} &= \ln(2^M - 2), \\ L_\eta^{\mathcal{B}}(\ell) &= \ln\binom{M}{k} + \ln k + \ln[1 + \ln(M-1)], \\ L_{\tilde{\mathbf{z}}}(\ell) &= \frac{k}{2}\ln\left[\frac{||\tilde{\mathbf{z}}||^2/k}{\ell^2}\frac{\pi\exp(1)}{2}\right] + \frac{1}{2}\ln k, \\ D_{\mathbf{x}}(\ell) &= \frac{1}{2}\sum_{j=0}^{M-1}\left[x_j - q(x_j)\right]^2. \end{aligned}$$

The following approximation is useful for reducing the computational burden [2, 10]:

$$\begin{aligned} \ln\binom{M}{k} \approx\ & (M+\frac{1}{2})\ln M - (k+\frac{1}{2})\ln k \\ & -(M-k+\frac{1}{2})\ln(M-k) - \frac{\ln(2\pi)}{2}. \end{aligned}$$

The parameter ℓ^* is chosen from a pre-defined set of nonnegative numbers such that to minimize the KSF, or equivalently,

$$\ell^* = \arg\min_\ell h_{\mathbf{c}}(\ell).$$

This leads to the following threshold to be used in cepstral nulling:

$$\mu_{\mathrm{KSF}} = 1 + \ell^*.$$

In our talk, we will discuss some of the properties of the KSF thresholding. Additionally, we will resort to simulations for comparing the performance of various thresholding schemes.

3. ACKNOWLEDGMENT

This work was supported by the Academy of Finland, Project Nos. 113572, 118355, 134767 and 213462.

4. REFERENCES

[1] P. Stoica and N. Sandgren, "Total-variance reduction via thresholding: Application to cepstral analysis," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 66–72, 2007.

[2] C.D. Giurcăneanu and S.A. Razavi, "On the use of Kolmogorov structure function for periodogram smoothing," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, Texas, USA, 2010*, pp. 3966 – 3969.

[3] P. Stoica and R. Moses, *Spectral analysis of signals*, Prentice Hall, 2005.

[4] E. Gudmundson, N. Sandgren, and P. Stoica, "Automatic smoothing of periodograms," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, 2006*, vol. 3, pp. III: 504–507.

[5] P. Stoica and N. Sandgren, "Smoothed nonparametric spectral estimation via cepstrum thresholding - introduction of a method for smoothed nonparametric spectral estimation," *IEEE Signal Processing Magazine*, vol. 23, no. 6, pp. 34–45, 2006.

[6] T. Gerkmann and R. Martin, "On the statistics of spectral amplitudes after variance reduction by temporal cepstrum smoothing and cepstral nulling," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4165 –4174, 2009.

[7] E. L. Lehmann, *Testing statistical hypotheses*, John Wiley & Sons, 1970.

[8] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, Mar. 1978.

[9] J. Rissanen, *Information and Complexity in Statistical Modeling*, Springer Verlag, 2007.

[10] T. Roos, P. Myllymäki, and J. Rissanen, "MDL denoising revisited," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3347–3360, 2009.