

Maximum Entropy Method without False Peaks with Exact Numerical Equation

Fumihiko Ishiyama

NTT Network Technology Labs., Nippon Telegraph and Telephone Corp., Tokyo, Japan

E-mail: ishiyama.fumihiko@lab.ntt.co.jp

Abstract. The standard numerical maximum entropy method (MEM) still uses the Yule-Walker equation which contains rough approximation by Walker. The commonly used numerical equation contains additional modifications to reduce calculation cost. Nowadays, we have powerful computers and there is no reason to use the modified equation. We show that the drawbacks of MEM such as false peaks and peak splittings are from the modifications. They do not appear when using the exact numerical equation, even a given time series is fractional.

1. Introduction

The maximum entropy method (MEM) is a well-known and widely used method of time series analysis. We can obtain characteristic frequencies with high resolution from a smaller number of samples than Fourier analysis. However, it is known that an MEM power spectrum includes false peaks and peak splittings.

We show that the peak structures are from the modifications of the equations for numerical calculation, and that they do not appear without such modifications.

The numerical method for MEM that appears in textbooks was established about half a century ago, and the method was tuned for the computers of that time. The modifications reduce calculation cost, and they were practical at that time. Nowadays, we have more powerful computers, and the modifications are not necessary. We can avoid false peaks with the exact equation for numerical calculation.

In the following sections, we review the conventional MEM, present a method that does not cause false peaks, compare our method with conventional MEM, and conclude the paper.

2. Conventional MEM

There are at least three types of modification of the equations for numerical calculation written in textbooks. They are the periodic boundary condition by Walker [1], periodicity breaking modification by Itakura [2], and boundary closing modification with a window function.

Let us consider the autocorrelation coefficients for numerical calculation.

The autocorrelation coefficients $r_{j,k}$ without modifications are

$$r_{j,k} = \frac{1}{N-M} \sum_{n=0}^{N-1-M} D^{-(n+j)} S_{N-1} D^{-(n+k)} S_{N-1}, \quad (1)$$



where S_n is the time series for the analysis, M is the order of the analysis, N is the number of samples, and D is the shift operator, such that

$$D^m S_n = S_{n+m}. \quad (2)$$

The terms j and k run through 1 to M .

The modifications of Eq. 1 are as follows.

The periodic boundary condition

$$D^N S_n = S_{n+N} = S_n \quad (3)$$

is introduced assuming an infinite number of samples $N \rightarrow \infty$ [1]. Then, the autocorrelation coefficients R_j for finite N become

$$R_j = \frac{1}{N} \sum_{n=0}^{N-1} D^{-n} S_{N-1} D^{-[(n+j) \bmod N]} S_{N-1}. \quad (4)$$

The term j runs through 0 to $M - 1$. This modification was introduced as “a first approximation” to obtain rough results [1], and the autocorrelation matrix became a Toeplitz matrix. That is, being a Toeplitz matrix means that what we can obtain is a rough, approximated result.

This equation does not appear in textbooks because all the poles are on the unit circle and the situation is not suited for the Levinson-Durbin algorithm.

The numerical equation in textbooks is given by Itakura [2]. The periodicity breaking modification is introduced into Eq. 4, and the autocorrelation coefficients become

$$R'_j = \frac{1}{N} \sum_{n=0}^{N-1-j} D^{-n} S_{N-1} D^{-(n+j)} S_{N-1}. \quad (5)$$

This modification was introduced to obtain a stable filter for speech synthesis, and this modification decreases the accuracies in frequency characteristics. As the autocorrelation matrix is a Toeplitz matrix, a sinusoidal signal is replaced with a decay signal with this modification. That is, false decay rates are introduced with this modification, and they make the Levinson-Durbin algorithm applicable. This is the original aim of this modification, and zero suppression of the non-diagonal terms is the reason for retrofitting. In fact, Walker's equation Eq. 4 does not use zero suppression.

This modification has several variations, and a widely used one is

$$R''_j = \frac{1}{N-j} \sum_{n=0}^{N-1-j} D^{-n} S_{N-1} D^{-(n+j)} S_{N-1}, \quad (6)$$

which is written in a book of recipes [3].

The boundary closing modification

$$w_0 S_0 = w_{N-1} S_{N-1} = 0 \quad (7)$$

with a window function w_n is commonly used to close both ends of a time series. This modification accompanies the periodic boundary condition, and this brings continuity at the boundaries.

We call the condition without these modifications the “open boundary condition.” We do not close the boundaries nor assume periodicities.

3. Our method

The equation for our method

$$\arg \min_{a_1, a_2, \dots, a_M} \sum_{n=0}^{N-1-M} \left[D^{-n} \left(1 - \sum_{m=1}^M a_m D^{-m} \right) S_{N-1} \right]^2 \quad (8)$$

is the conventional autoregressive model with the method of least squares, and the difference is in the process of numerical calculation [4, 5, 6]. We use Eq. 1 instead of Eqs. 5 or 6, and do not use window functions. That is, we use the open boundary condition. We obtain the prediction coefficients a_m by solving Eq. 8 with the appropriate process of numerical calculation.

Consequently, we factorize the equation

$$1 - \sum_{m=1}^M a_m D^{-m} = \prod_{m=1}^M (1 - x_m D^{-1}) \quad (9)$$

with the obtained prediction coefficients a_m , and obtain the complex constants x_m .

The complex constants x_m correspond to the oscillation modes of the time series S_n . The time series S_n is expanded with the oscillation modes x_m , and the corresponding characteristic frequencies f_m and decay rates λ_m are obtained from the oscillation modes x_m as follows.

$$\begin{aligned} S_n &\simeq \sum_{m=1}^M c_m x_m^n \\ &= \sum_{m=1}^M c_m e^{n(2\pi i f_m + \lambda_m) \Delta T} \end{aligned} \quad (10)$$

The complex coefficients c_m are the complex amplitudes of the oscillation modes x_m , and ΔT is the interval of sampling.

Note that the decay rates λ_m are not always negative [4, 5, 6]. That is, we also treat growing modes, so the word “decay” is not quite appropriate. Nevertheless, we use the word “decay” for simplicity.

We apply the method of least squares

$$\arg \min_{c_1, c_2, \dots, c_M} \sum_{n=0}^{N-1} \left[D^{-n} S_{N-1} - \sum_{m=1}^M c_m x_m^{\frac{N-1}{2}-n} \right]^2 \quad (11)$$

to obtain the values of the complex amplitudes c_m . As the decay rates λ_m are not zeros in general, we choose the amplitudes c_m at the center of the time series S_n .

The equation for the MEM power spectrum $P(f)$ is written as

$$\begin{aligned} P(f) &= \frac{1}{\left| 1 - \sum_{m=1}^M a_m e^{-2\pi i m f \Delta T} \right|^2} \\ &= \frac{1}{\left| \prod_{m=1}^M (1 - e^{2\pi i (f_m - f) \Delta T + \lambda_m \Delta T}) \right|^2}, \end{aligned} \quad (12)$$

where a_m are the prediction coefficients and f_m and λ_m are the frequencies and the decay rates that correspond to the poles.

Obviously, the equation does not contain the parameters that correspond to the real intensities. Therefore, we introduce an equation for the power spectrum that reflects the real intensities of the oscillation modes x_m [6].

We consider the amplitude spectrum

$$F_m(f) = \left| \frac{c_m \lambda_m}{2\pi i(|f_m| - f) - |\lambda_m|} \right| \quad (13)$$

of each oscillation mode x_m , which has the maximum value

$$F_m(|f_m|) = |c_m|. \quad (14)$$

We use $|f_m|$ and $|\lambda_m|$ to keep their values non-negative. Then, summing up the amplitude spectrum $F_m(f)$ of each oscillation mode x_m , we obtain the equation for the power spectrum

$$P(f) = \left(\sum_{m=1}^M F_m(f) \right)^2, \quad (15)$$

that reflects the intensities of the oscillation modes x_m . Note that each peak includes some contribution from other peaks, and the value of each peak becomes larger than $|c_m|$ [6].

4. Analysis of fractional time series

We applied our method and the conventional MEM to a fractional sinusoidal time series

$$S_n = \sin 2\pi n \Delta T, \quad (16)$$

with the number of samples $N = 40, 43, 49$ and 55 , and interval of sampling $\Delta T = 1/20$. We used a small number of samples to emphasize the false peaks.

We show the time series and time widths for the analysis in figure 1, assuming $t = n\Delta T$. The case $N = 40$ corresponds to the periodic case, and the others are the fractional cases.

We used Eqs. 6 and 12 for the conventional MEM, and Eqs. 1 and 15 for our method. We used the order for the analysis $M = 8$ in common, and did not apply any window functions. The results are shown in figure 2.

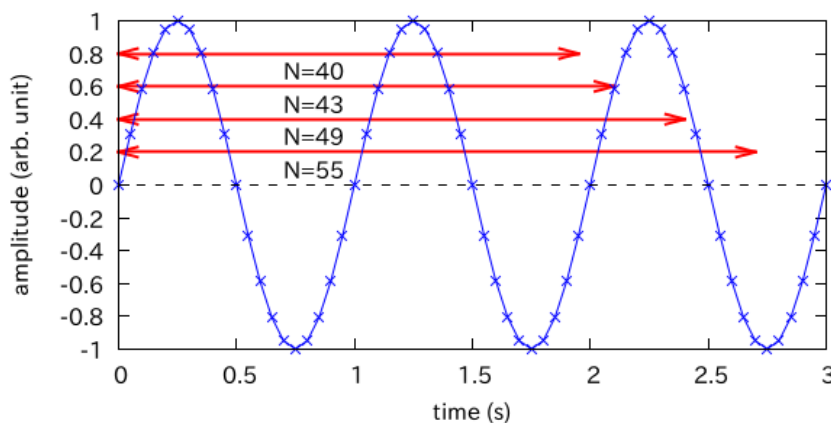


Figure 1. Time series for analysis. Various fractional time series were taken to evaluate modifications on conventional MEM.

The conventional MEM (figure 2(a)) showed various false peaks depending on the number of samples N . Even the periodic case $N = 40$ had false peaks. This is because the periodicity breaking modification [2] replaces the sinusoidal signal with a decay signal. The decay rates of

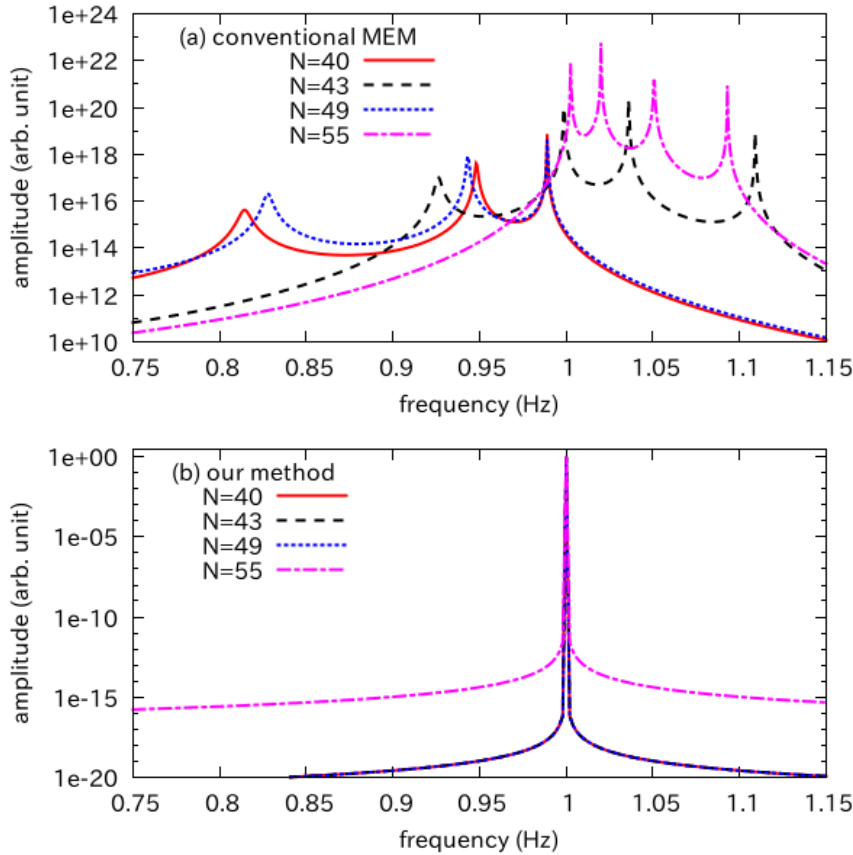


Figure 2. Power spectra obtained with (a) conventional MEM, and (b) our method for various N . Conventional MEM showed various peak structures caused by modifications in numerical calculation. We obtained exact frequency and intensity from all fractional time series with our method.

the peaks for the case $N = 40$ were obtained from their prediction coefficients a_m by using Eqs. 9 and 10, and major ones were $-4 \cdot 10^{-3} \text{ s}^{-1}$ for 0.95 Hz, and $-7 \cdot 10^{-4} \text{ s}^{-1}$ for 0.99 Hz.

The intensities of the peaks on the spectra had extremely large values because of the small values of the decay rates of the crowded peaks.

In contrast, the false peaks disappeared with our method for all fractional cases (figure 2(b)), and the errors of the obtained frequencies were less than 10^{-8} Hz . The floor levels of the spectra are given by the weak oscillation modes which correspond to the errors in the numerical calculation. Their amplitudes were less than 10^{-9} , and we recommend removing the modes for further use.

We show the extrapolated time series obtained with the conventional MEM and our method in figure 3. We used the periodic case $N = 40$, obtained the prediction coefficients a_m of the two methods, applied Eqs. 9 and 11, and plotted Eq. 10 for $-5 \leq t \leq 20$.

Our method with its open boundary condition reproduced sinusoidal time series. In contrast, the extrapolated time series obtained with the conventional MEM became a decay time series with a beat structure corresponding to the peaks shown in figure 2 (a). The approximate period of the beat was about 25 seconds, and the period corresponding to the beat was between 0.95 Hz and 0.99 Hz.

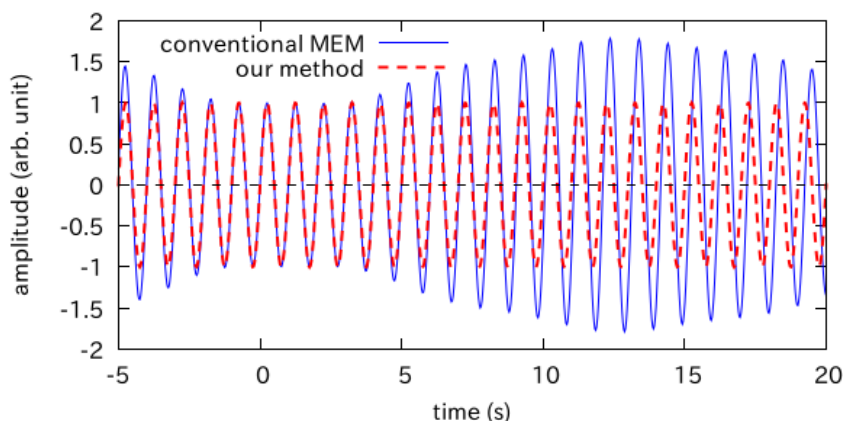


Figure 3. Extrapolated time series for case $N = 40$ shown in figure 1. Time series for analysis corresponds to $0 \leq t < 2$. Time series was extrapolated to $-5 \leq t < 0$ and $2 \leq t \leq 20$. Periodicity breaking modification of conventional MEM replaced sinusoidal time series with decay time series accompanying beating structure. Our method with its open boundary condition, reproduced sinusoidal time series.

5. Conclusion

We pointed out that the drawbacks of the conventional MEM such as false peaks and peak splittings are caused by the modifications of the equations for numerical calculation. The modifications are the periodic boundary condition, periodicity breaking modification, and boundary closing modification. The aims and the effects of the modifications are as follows.

The periodic boundary condition was introduced as “a first approximation” to simplify the unmodified equation and to obtain rough results with the simplified equation.

The periodicity breaking modification was added to the above modification to obtain a stable filter for speech synthesis, and this modification decreased the accuracies of frequency characteristics.

The boundary closing modification with a window function is a common modification that accompanies the periodic boundary condition to bring continuity at the boundaries.

These modifications are hidden behind the theoretical equations, and provide results inconsistent with the theoretical expectation.

We applied the MEM without the above modifications, which we call open boundary condition, to various fractional sinusoidal time series, and showed that we can obtain exact power spectra from all the various fractional time series. All of them have no false peaks nor peak splittings.

The standard numerical method of MEM still uses the Yule-Walker equation, Toeplitz matrix and Levinson-Durbin algorithm to reduce calculation cost. However, computers today are powerful enough to avoid the approximations, and it is the time to refresh the numerical method.

References

- [1] Walker G 1931 *Proc. Roy. Soc. Ser. A* **131** pp 518-532
- [2] Itakura F 1975 *IEEE Trans. Acoust., Speech, Signal Process.* **23** pp 67-72
- [3] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1986 *Numerical Recipes: The Art of Scientific Computing*. (Cambridge: Cambridge University Press)
- [4] Ishiyama F, Okugawa Y and Takaya K 2017 *Proc. 13th IEEE Colloq. Signal Process. Appl. (Penang)* (New York: IEEE) pp 44-46
- [5] Ishiyama F 2017 *Proc. 17th IEEE Int. Symp. Signal Process. Info. Tech. (Bilbao)* (New York: IEEE) pp 1-6
- [6] Ishiyama F 2018 *IEICE Trans. Fundamentals (Japanese Edition)* **J101-A** pp 36-45