Signal Processing 142 (2018) 75–86

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

# Enhancement of variational mode decomposition with missing values

Guebin Choi[a], Hee-Seok Oh[a,*], Donghoh Kim[b]

[a] Department of Statistics, Seoul National University, Seoul 08826, Republic of Korea
[b] Department of Applied Mathematics, Sejong University, Seoul 05006, Republic of Korea

ARTICLE INFO

*Article history:*
Received 19 September 2016
Revised 6 July 2017
Accepted 11 July 2017
Available online 12 July 2017

*Keywords:*
Decomposition
Missing values
Periodogram
Self-consistency
Thresholding
Variational mode decomposition

ABSTRACT

This paper considers an improvement of variational mode decomposition (VMD) in the presence of missing values. VMD developed by Dragomiretskiy and Zosso (2014) efficiently decomposes a signal into some meaningful modes according to their frequency information. It is well known that VMD is useful for tone detection and denoising of noisy signals. However, VMD may not be efficient for analyzing missing data since it is based on discrete Fourier transform (DFT). This paper proposes a new VMD procedure that can effectively handle problems caused by missing values. The proposed method is based on an estimation of spectral density that reflects frequency information of a signal properly with removing the effects of missing samples; hence, it is able to produce stable decomposition results. Results from numerical studies including simulation study and real data analysis demonstrate the promising empirical properties of the proposed method.



## 1. Introduction

Variational mode decomposition (VMD) is a recently developed method for decomposing a signal when the signal consists of several modes according to different main frequencies. VMD has a solid theoretical background since it is based on spectrum-based analysis, and it is robust to noise compared to empirical mode decomposition [13]. In order to decompose each mode embedded in the signal, VMD first conducts discrete Fourier transform (DFT) for detecting frequency information of each mode, and then identifies several meaningful modes using the detected main frequencies. However, irregularity of observations due to missing values might distort the decomposition results of VMD. To be specific, when a signal is not uniformly sampled, the information of the signal observed in the time domain cannot be effectively transferred into the frequency domain through DFT; hence, VMD malfunctions in such a case.

For a particular example that illustrates the irregularity effect on the decomposition by VMD, we consider a sinusoidal signal of the form

$$f(t) = \cos(20\pi t) + \cos(100\pi t), \ \ 0 \le t \le 1.$$

The left panel of Fig. 1 shows the decomposition results by VMD for a complete dataset, which is regularly observed at time points $t_i = i/N, \ \ i = 0, \dots, N-1$. The corresponding periodogram detects two main frequencies properly, and thus, VMD successfully identifies two modes, the low-frequency component $\cos(20\pi t)$ and the high-frequency component $\cos(100\pi t)$. As shown, the reconstruction from two extracted modes recovers the signal $f$ effectively. The middle panel of Fig. 1 displays the decomposition results by applying VMD to a signal with missing values, where 70% of time points are missing at random over the entire domain. It seems that the information of main frequencies is not well detected by periodogram. Thus, VMD cannot decompose two components of the signal $f$ effectively. To cope with the aforementioned irregularity problem, as a simple way, missing values can be imputed by linearly interpolating observed signals. The right panel of Fig. 1 demonstrates the decomposition results by applying VMD to the imputed signal, which reduces some distortion caused by missing values. This imputed method is termed as LVMD. However, it seems that this remedy of missing values in the time domain is not enough to identify each mode.

As a previous work that is related to the current study, Choi et al. [1] developed an imputation method for VMD when a signal contains missing values, which is based on a combination of VMD and hierarchical (or $h$)-likelihood of Lee and Nelder [12]. This method is fully operated in the time domain. Further, in the perspective of using likelihood principle, the method of Choi et al. [1] can be considered as an extension of EM-algorithm. On the other hand, the proposed method in the current paper attempts to solve the problem of missing samples in the frequency domain, which can be well embedded in VMD framework.

* Corresponding author.
*E-mail addresses:* heeseok@stats.snu.ac.kr, heeseok.oh@gmail.com (H.-S. Oh).

http://dx.doi.org/10.1016/j.sigpro.2017.07.007
0165-1684/

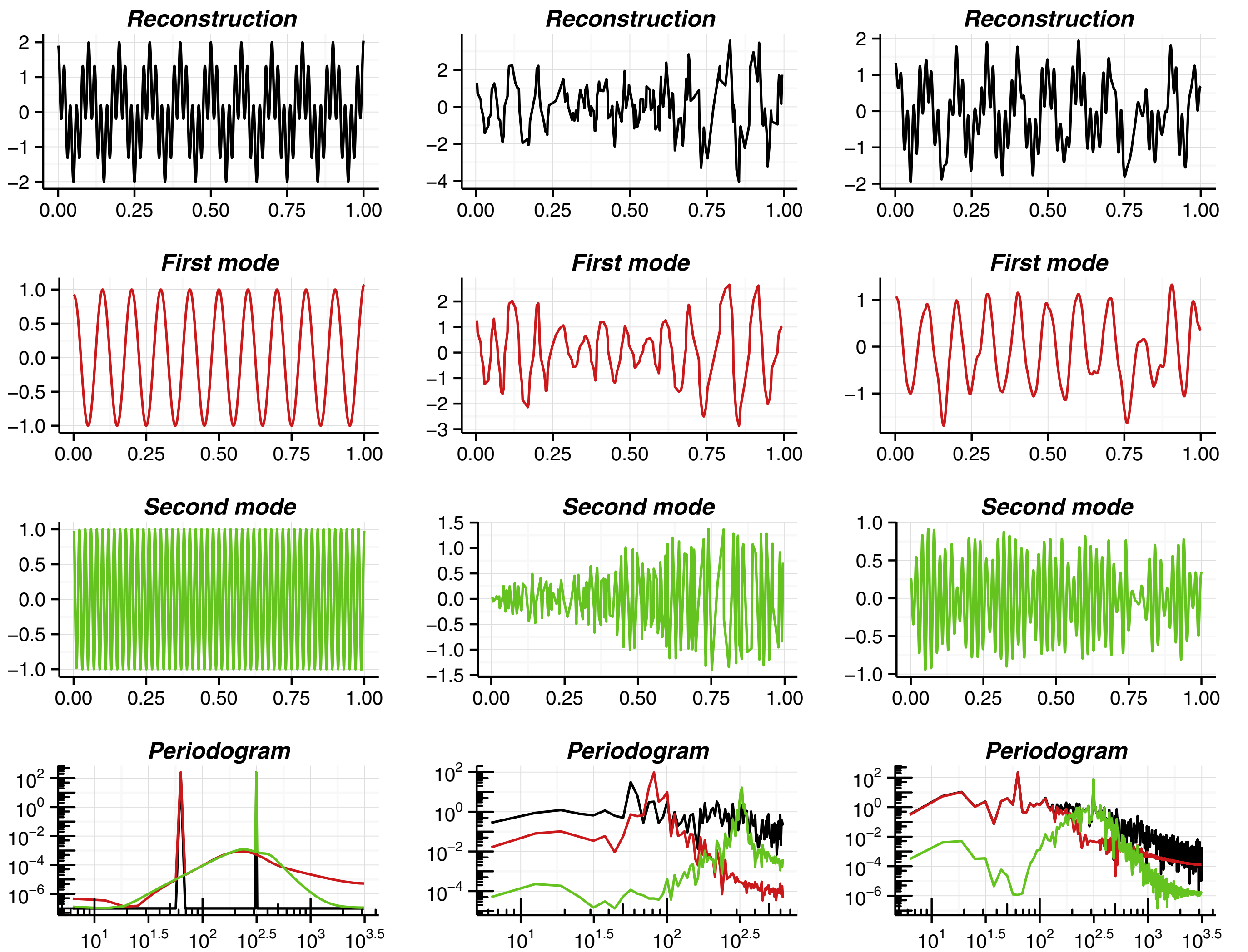


**Fig. 1.** From left to right, VMD results for complete data from $f(t) = \cos(20\pi t) + \cos(100\pi t)$, VMD results for 70% missing data, and LVMD results based on linear interpolation for 70% missing data.

In this paper, we propose a new method to improve the conventional VMD in the presence of missing values, termed *iterative thresholded VMD* (ITVMD) by treating the missing sample effect in the frequency domain. DFT does not work properly with missing values. Thus, the resulting frequency information is distorted, which perturbs the whole VMD procedure; hence, VMD cannot produce stable decomposition results. To overcome this problem, we suggest a thresholding procedure of periodogram that alleviates the missing value problem and eventually detects the frequency information of each mode correctly. In addition, we present a practical algorithm of ITVMD that reflects the adjustment of the missing sample effects under the framework of VMD algorithm. Furthermore, we consider a nonparametric estimation of spectral density by wavelet transform-based thresholding that extends the scope of signals for decomposition.

The remaining of this paper is organized as follows: Section 2 reviews the conventional VMD. In Section 3, the proposed ITVMD is presented with pointing out problems caused by missing values. Section 4 discusses various numerical examples for evaluating the proposed method. Section 5 provides a real-world signal analysis. A nonparametric estimation of spectral density via wavelet transform is discussed in Section 6. Finally, concluding remarks are addressed in Section 7.

## 2. Review: variational mode decomposition

In this section, we would like to briefly review VMD and its algorithm for a better presentation of our method in Section 3 as well as a self-contained material. The ultimate goal of VMD is to decompose a signal $f$ into several modes $u_k$ named intrinsic mode function (IMF). The definition of IMF is as follows.

**Definition 2.1.** Intrinsic mode function $u_k$ is an amplitude-modulated-frequency-modulated (AM-FM) signal of the form

$$u_k(t) = A_k(t)\cos(\theta_k(t)),$$

where $\theta_k(t)$ is a non-decreasing time-varying phase and $A_k(t)$ denotes a non-negative time-varying envelope function. Further, the instantaneous frequency is defined as $\omega_k(t) = \theta_k'(t)$, which is non-negative as well, and both the envelope function and the instantaneous frequency vary much slower than the phase $\theta_k(t)$.

Suppose that a signal $f(t)$ consists of $K$ IMFs as

$$f(t) = \sum_{k=1}^{K} u_k(t) = \sum_{k=1}^{K} A_k(t)\cos(\theta_k(t)).$$

VMD is a process of estimating IMFs $u_1, u_2, \ldots, u_K$ and the corresponding frequencies $\omega_1, \omega_2, \ldots, \omega_K$. For extracting IMFs from the