New Data Analysis Method for Single-dish Spectroscopy Based on Gaussian Filtering

Takeshi Shakunaga*(1), Hajime Ezawa(1)(2), Wataru Kawasaki(1), Renaud Miel(1),
Takeshi Nakazato(1), Kazuhiko Shimada(1), Akira Yoshino(1), and Kanako Sugimoto(1)
(1) National Astromical Observatory of Japan, Tokyo, Japan; e-mail: takeshi.shakunaga@nao.ac.jp
(2) The Graduate University for Advanced Studies, SOKENDAI, Tokyo, Japan

Methods to evaluate the baseline and emission/absorption line features of ALMA single-dish spectroscopic observation data are discussed. In ALMA project, Pipeline has been developed based on the standard analysis package CASA [1], aiming to automatically process the observed data as a goal. Pipeline has been widely utilized for astronomical data reduction, as well as for quality assurance of the observed data prior to delivery to the observers. For spectroscopic single-dish observation, keys of data reduction process include baseline subtraction and emission/absorption line identification. The Pipeline process is developed with various heuristics to cover majority of the observational data, nevertheless the current process does not converge as expected for certain extreme cases. We propose to introduce Gaussian-based algorithms, which have been widely used for computer vision. The algorithms are expected to improve baseline subtraction and line identification to cover more cases compared to the conventional method, and may have various potential applications for astronomical data analysis in future.

Emission/absorption line peaks are basically detected by zero-crossings of first derivative of spectral signal. In order to suppress noises, spectral signal is convolved with Gaussian function before differentiation. When input signal and Gaussian kernel are denoted by $f(x)$ and $g(x)$, respectively, and $*$ indicates convolution, $\frac{d}{dx}\{g(x) * f(x)\}$ is implemented by $\frac{d}{dx}g(x) * f(x)$. Since zero-crossing detection is also easily implemented and $\frac{d^2}{dx^2}g(x) * f(x)$ is used for validation, line peaks are efficiently detected by Gaussian filtering. Line widths are also efficiently estimated by simple analysis after line peak detection using baseline estimation. On the other hand, ripples often appear in spectral signal, where spectral line detection gets harder than ordinary conditions. In those situations, Gaussian averaging is effective for spectral line detection, where the standard deviation, σ, of Gaussian function is optimized for a given data. Regarding the Gaussian average as a baseline, spectral lines are detected by Gaussian filtering. A fundamental example of line peak detection is as shown in Fig.1. Additional examples for tough situations, where large ripples appear, for examples, will be discussed in the conference.

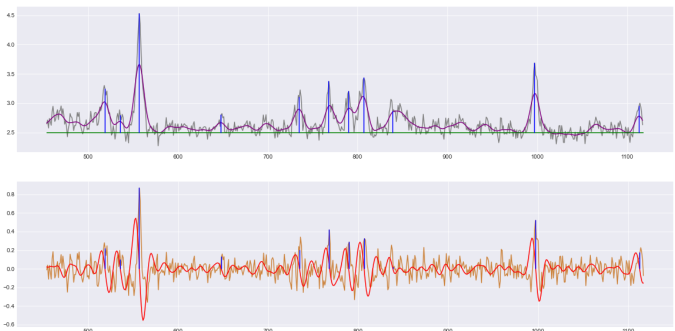

Figure 1. In above graphs, x- and y-axes indicate frequency and intensity of spectrum: (1) In the upper graph, given input signal $f(x)$ is as shown in gray, and its Gaussian average with optimum σ is as shown in purple. Baseline is estimated as shown in green. (2) Using the optimum σ, $f(x) - g(x) * f(x)$ is generated as shown in brown in the lower graph. Then, the first derivative of Gaussian-subtracted signal is computed as shown in red, and its zero-crossings are listed up as the emission line candidates. (3) In all the candidates, reliable ones are selected by the second derivative of Gaussian-subtracted signal. Selected lines in this domain are as shown in blue. (4) In the original signal domain, selected lines are as shown in blue in the upper graph.

References

[1] The CASA Team, et al., " CASA, the Common Astronomy Software Applications for Radio Astronomy ", PASP 134, 114501 (17pp), November 2022, doi: 10.1088/1538-3873/ac9642.