MDPI sciforum

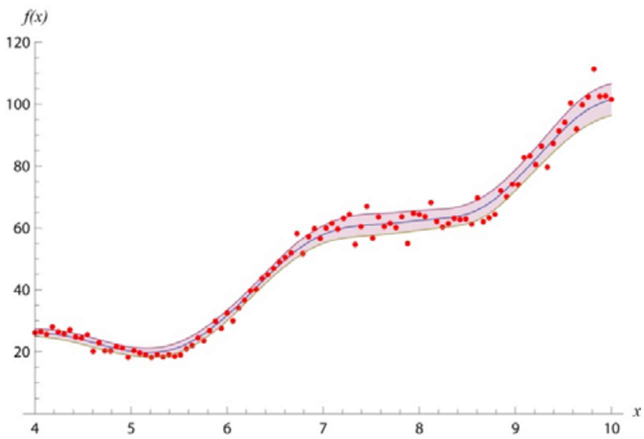
IWIMSM-04: Iberoamerican Workshop on Model. and Simulation Methods, Valencia, Spain, 2020

Gaussian method for smoothing experimental data

J.L. González-Santander[a], *G. Martín*[b]

[a] *Department of Mathematics. Universidad de Oviedo, 33007 Asturias, Spain*
[b] *Department of Experimental Sciences and Mathematics. Universidad Católica de Valencia, 46001 Valencia, Spain*

Graphical Abstract



Abstract

We provide a method for experimental data smoothing under a certain noise by using a statistical fitting considering gaussian weight functions. On the one hand, this method is quite useful when we have a large amount of experimental data, which are expected to approach an unknown theoretical curve. This allows us to find quite closely the derivative of the theoretical curve from the data and provides as well the error in the numerical integration of the data. On the other hand, the proposed method improves the typical smoothening of the time series of financial data and allows the calculation of the volatility as a function of time.

Introduction

We consider experimental data y_i as the prediction of a theoretical curve $y(x)$ at x_i, $i = 1, \ldots, n$, subjected to a random noise produced by measurement uncertainty, which is normally distributed.

Materials and Methods

If data is homocedastic with normal standard deviation σ, the fitting function is given by,

$$f(x,s) := \frac{\sum_{i=1}^{n} y_i\,\rho_i(x,s)}{\sum_{i=1}^{n} \rho_i(x,s)}, \qquad \rho_i(x,s) := \exp\left[\frac{x_i(2x-x_i)}{2s^2}\right],$$

where the parameter s is calculated by numerically solving the following equation:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[f(x_i,s)-y_i\right]^2} - \sigma = 0.$$

The derivative of the fitting function can be obtained from the data (x_i, y_i) as:

$$f'(x,s) = \frac{\mathrm{Cov}\langle x_i, y_i\rangle}{s^2} = \frac{\langle x_i y_i\rangle - \langle x_i\rangle\langle y_i\rangle}{s^2},$$

where we have adopted the notation:

$$\langle z_i\rangle = \frac{\sum_{i=1}^{n} z_i\rho_i(x,s)}{\sum_{i=1}^{n}\rho_i(x,s)}.$$

The confidence band of the fitting function is given by

$$\Delta f(x,s) := \sqrt{\frac{\sum_{i=1}^{n}\left[f(x_i,s)-y_i\right]^2\rho_i(x,s)}{\sum_{i=1}^{n}\rho_i(x,s)}}.$$

If data is heterocedastic with known standard deviation $\sigma(x)$, we calculate parameter s minimizing the following function:

$$G(s) := \int_{x_1}^{x_n}\left[\Delta f(x,s)-\sigma(x)\right]^2 dx.$$

If $\sigma(x)$ is unknown, we calculate s minimizing the function:

$$G(s) := \sum_{j=1}^{n-1}\left[\Delta f(m_j,s)-\sigma(m_j,s)\right]^2,$$

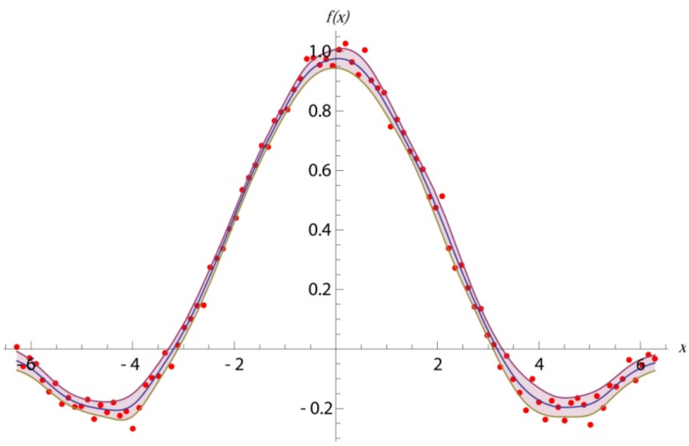
where

$$\sigma(m_j,s) = \frac{\left|y_{j+1}-y_j-f'(m_j,s)\Delta x_j\right|}{\sqrt{2}}, \quad m_j = \frac{x_{j+1}+x_j}{2}, \quad \Delta x_j = x_{j+1}-x_j.$$

Results and Discussion

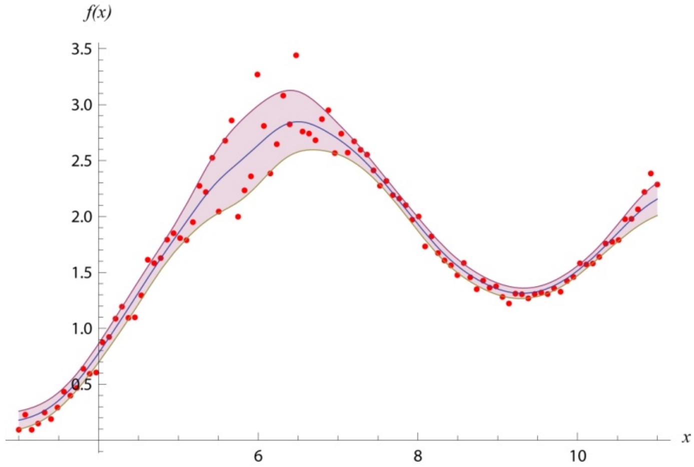
Theoretic function: $y(x) = \mathrm{sinc}(x)$

Data: $x \in [-2\pi, 2\pi]$, $n = 100$, $\sigma = 0.03$.

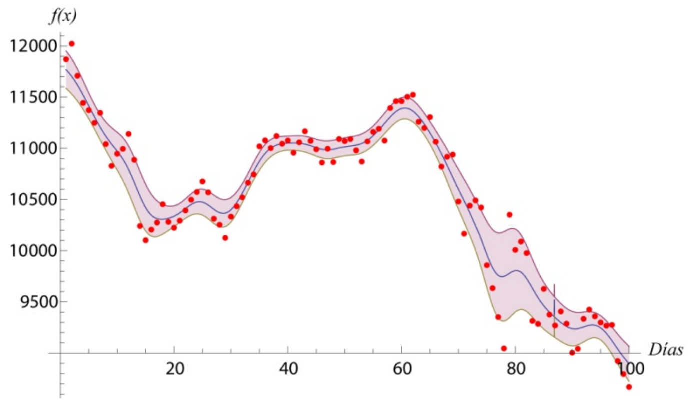


Theoretic function: $y(x) = \log x + \cos x$.

Data: $x \in [3,11]$, $n = 100$, $\sigma(x) = 0.05\left(\frac{3}{(x-6)^2 + 0.3} + 1\right)$.



Daily data at the close of the IBEX 35 index, from January 18 to June 8, 2010.



References

[1] J. Stoer and R. Bulirsch, Introduction to Numerical Analysis, SpringerVerlag, New York, 1980.

[2] N. Ahmed, T. Natarajan, K. R. Rao, Discrete Cosine Transform, IEEE Trans. Computers C-23 (1974) 90-93.

[3] J. Xiuyu Cong, Historical Development of Central Limit Theorem, 2003.

[4] J. Spanier, K. B. Oldham, An atlas of functions, Hemisphere Publising Corporation, 1987.

[5] E. W. Weisstein, Sinc Function. MathWorld—A Wolfram Web Resource.
http://mathworld.wolfram.com/SincFunction.html

[6] J. F. Kenney, E. S. Keeping, Moving Averages §14.2 en Mathematics of Statistics, Pt. 1, 3rd ed. Princeton, NJ: Van Nostrand, pp. 221-223, 1962.

[7] A. Korganoff, Méthodes de Calcul Numérique, Tome 1, Dunod, París, 1961

[8] A. Korganoff and M. Pavel-Parvu, Méthodes de Calcul Numérique, Tome 2, Dunod, París, 1961