Analytica Chimica Acta 705 (2011) 218–226

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

Deconvolution of pulse trains with the L_0 penalty

Johan de Rooi [a,b,*], Paul Eilers [b]

[a] Department of Bioinformatics, Erasmus Medical Center, Rotterdam, The Netherlands
[b] Department of Biostatistics, Erasmus Medical Center, Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 13 January 2011
Received in revised form 3 May 2011
Accepted 16 May 2011
Available online 24 May 2011

Keywords:
Spike trains
Spectra
Blind deconvolution
Ill-conditioned
Penalized regression

ABSTRACT

The output of many instruments can be modeled as a convolution of an impulse response and a series of sharp spikes. Deconvolution considers the inverse problem: estimate the input spike train from an observed (noisy) output signal. We approach this task as a linear inverse problem, solved using penalized regression. We propose the use of an L_0 penalty and compare it with the more common L_2 and L_1 penalties. In all cases a simple and iterative weighted regression procedure can be used. The model is extended with a smooth component to handle drifting baselines. Application to three different data sets shows excellent results.

1. Introduction

Many instruments produce signals that consist of a series of pulses. Examples are electrophoretic DNA sequencers, chromatographs, and spectrometers. Some biological signals have the same characteristics; an example is hormone release in the human body. The pulses have (more or less) equal shapes but different heights, and they may overlap. These output signals are the convolution of a series of true (input) spikes or diracs and the impulse response function. The task is to deduce the heights and positions of the spikes from the output signal.

Essentially there are two ways to approach this issue. The first is to search for local maxima to find peak positions, followed by summarizing the signal in their neighborhoods, to estimate peak heights. Examples of this approach are found in many places in the literature. We mention only a small selection. Yasui et al. [27] search for zeros of the first derivative, while Mariscotti [14] uses the second derivative. When computing derivatives, it is essential that proper noise filtering is first applied. Wavelets have been proposed as a tool for filtering by Coombes et al. [6] in this setting, but other filters are also possible. Du et al. [7] use a wavelet spectrum to locate peaks. It is also possible to apply a discrete Markov chain as done by Silagadze [22] and Morháč [16], these approaches results in a probability distribution targeting the location of peaks.

The second approach is to model signals as a convolution of a series of sharp input spikes and a constant impulse response. The task then is to estimate the input from the observed output signal. This is the deconvolution problem that has been studied in many fields of science. It is a so-called inverse problem, and it is generally very badly conditioned, which means that small changes in the observed signal or the impulse response lead to large changes in the estimated input. Conversely, many very different inputs are compatible with the observed output.

To address the bad condition various deconvolution algorithms have been proposed. An early solution is the van Cittert algorithm (see e.g. [11]) that was later improved in the form of the Gold algorithm (see e.g. [1]). Other iterative approaches are often based on the Richardson–Lucy algorithm or using the general class of expectation maximisation (EM) algorithms (see e.g. [13,24]). The EM algorithm iteratively redistributes the observed output, proportionally to the current estimate of the input. Averaging gives an improved estimate of the input, to be used in the next iteration.

The class of deconvolution algorithms also contains a branch of boosting algorithms. Cardot et al. [3] propose to use boosting to find an optimal set of input spikes. As a first step the locations of peak are estimated and subsequently renewed in an updating step, in the third stage peaks that are too close to each other are merged to one. Recently Morháč and Matoušek [18] proposed a boosted version of the Gold and Richardson–Lucy algorithms.

A general approach to ill-conditioned problems is the use of regularization: some form of penalty is imposed on the parameters of the model. We already referred to Li and Speed [13]. A familiar example in the chemometric literature is ridge regression [10],

* Corresponding author at: Department of Biostatistics, Erasmus Medical Center, Dr Molewaterplein 50, 3015 GE Rotterdam, The Netherlands. Tel.: +31107044505
E-mail addresses: j.derooi@erasmusmc.nl (J. de Rooi), p.eilers@erasmusmc.nl (P. Eilers).

0003-2670/$ – see front matter
doi:10.1016/j.aca.2011.05.030

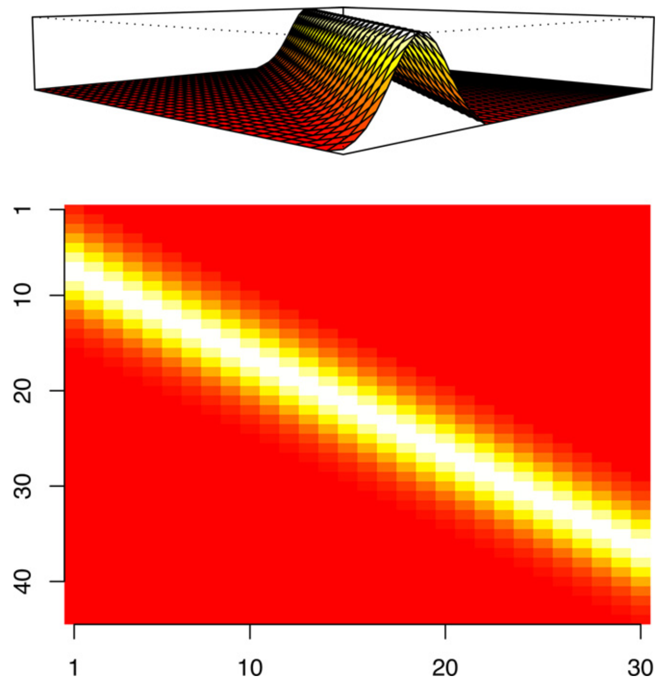

Fig. 1. A convolution matrix S with $m+(n-1)$ rows and m columns, in 3D in the upper panel and in the lower panel in 2D. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

where the penalty is on the sum of the squares of the regression coefficients. This is called the L_2 norm; generally the sum of absolute values (of the elements of a vector) to the power p is called the L_p norm. In recent years the Lasso [23], a penalty based on the L_1 norm, the sum of absolute values, has become popular in many applications. References can be found in the next section.

Penalties with a norm based on $p<1$ have received little attention. A main theoretical obstacle has been the fact that they leads to a non-convex optimization problem, in contrast to penalties with $p \geq 1$. Hence one cannot be sure of having found a global minimum. Another drawback is the lack of good practical algorithms. In this paper we propose regularized deconvolution using the L_0 penalty, and we show very good results using an algorithm based on repeated weighted regression. Apparently, in the limited context of pulse train deconvolution, a non-convex objective function is not a real problem.

In the next section we introduce the deconvolution framework, and we show the effects of regularization with different norms. There we assume that the impulse response is known. In practice only an approximation will be available, so we also consider "blind deconvolution": the estimation of both input and impulse response form one signal. Drifting baselines are quite common; we present two ways to handle them.

In Section 3 we present three applications to experimental data. Two of them are instrumental (electrophoretic DNA sequencing and gas chromatography), the third is a series of high-frequency measurements of concentrations of luteinizing hormone in human blood, which show strong pulsative behavior.

In the final section we discuss possible extensions and refinements.

2. The model

2.1. Convolution and deconvolution

Consider a (causal) discrete linear system with an input signal x, and a impulse response (or spread function) c, which incorporates

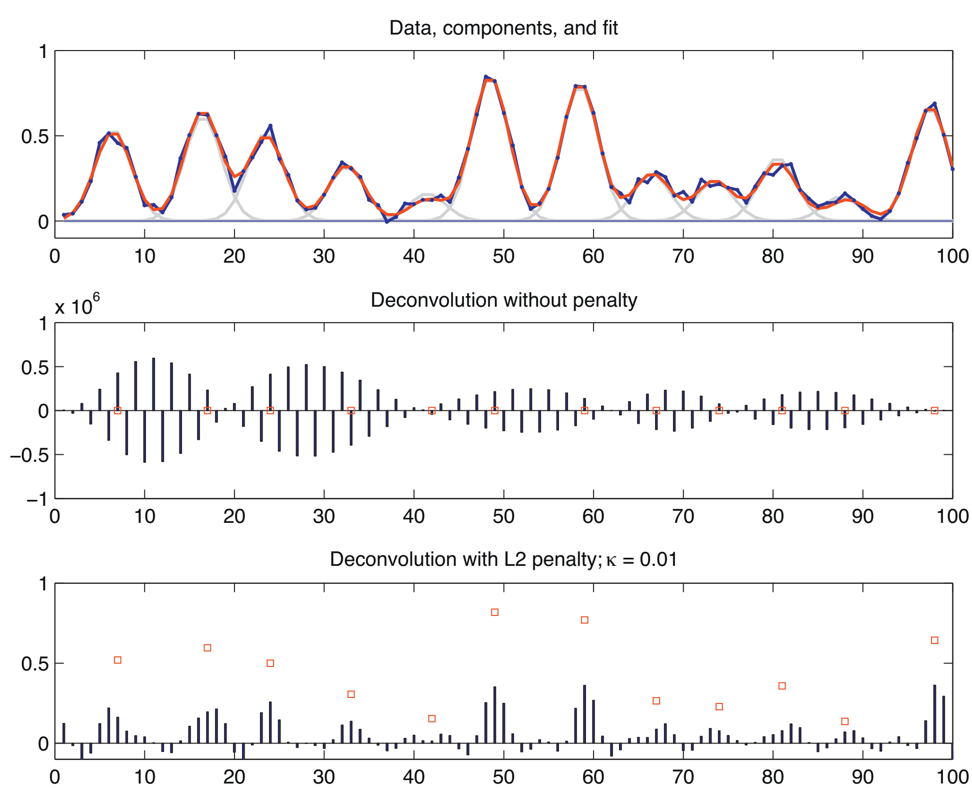

Fig. 2. Simulated data. Top panel: data (full blue line with dots), fit (full red line) and individual pulses (thick gray lines). Middle panel: input as estimated without a penalty. Bottom panel: input as estimated with an L_2 penalty. The small squares give the positions and the heights of the nonzero elements of the input used for the simulation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)