Signal Processing 108 (2015) 430–439

Contents lists available at ScienceDirect

Signal Processing

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

Drift removal by means of alternating least squares with application to Herschel data ☆

Lorenzo Piazzo [a,*], Pasquale Panuzzo [b], Michele Pestalozzi [c]

[a] DIET Department, Sapienza University of Rome, V. Eudossiana 18, 00184 Rome, Italy
[b] GEPI Laboratory, CNRS – University Paris Diderot, Meudon, France
[c] IAPS Institute, National Institute of Astrophysics, Rome, Italy

ARTICLE INFO

Article history:
Received 13 May 2014
Received in revised form
18 August 2014
Accepted 24 September 2014
Available online 16 October 2014

Keywords:
Least squares
Image formation
Infrared imaging

ABSTRACT

We consider the problem of reconstructing an image observed with a linear, noisy instrument, the output of which is affected by a drift too, causing a slowly varying deviation of the readouts from the baseline level. Since the joint estimation of the image and the drift, which is the optimal approach, is demanding for large data, we consider an alternative approach, where we remove the drift and the noise in two separate steps. In particular, we remove the drift by means of Least Squares (LS) and the noise by means of Generalised Least Squares (GLS). Moreover, we introduce an efficient drift removal algorithm, based on Alternating Least Squares (ALS), and carry out an analysis which proves convergence and gives geometrical insight. Finally, we apply the approach to the Herschel satellite data, discussing the performance and showing that nearly optimal results are achieved.

1. Introduction

Consider a linear instrument observing an image and producing a sequence of readouts. The image formation problem is that of reconstructing the image from the readouts. This problem is complicated by the impairments affecting the output, the first of which is the noise introduced by the instrument. A second impairment often arising in practice is a drift affecting the detector, causing a time varying deviation of the readouts from the baseline level. Such a problem is encountered, for example, in medical imaging [1,2], in electron microscopy [3] and in astronomical imaging [4,5]. Typically, the drift is slowly varying and can be modelled with smooth curves, depending on a few parameters. However the drift may be large, possibly larger than the useful signal, and must therefore be accounted for in order to obtain quality images.

The optimal approach to the image formation is that of jointly estimating the image and the drift, e.g. [2,4]. However the joint estimation is computationally demanding. Another option, e.g. [5], is to use a two steps approach: in the first step an estimate of the drift is produced and subtracted from the raw data, to obtain an updated data set which is, ideally, drift free; in the second step the updated data set is fed to a proper noise removal algorithm. This approach is less demanding and is studied in this paper where, specifically, we propose to use a Joint Least Square (JLS) method to remove the drift and a Generalised Least Squares (GLS) method to remove the noise. As we will see, the JLS preserves the useful signal and, under proper assumptions, completely removes the drift. As a result our approach produces near optimal images. Moreover, the proposed approach is more practical than the one of [5], where the drift estimation is carried out by means of an interactive method, requiring human intervention.

☆ This work was partially funded by the Sapienza University Research Projects no. C26A11825Y/2011 and C26A1324B9/2013.
* Corresponding author. Tel.: +39 06 44585480; fax: +39 06 4873300.
E-mail address: lorenzo.piazzo@uniroma1.it (L. Piazzo).

http://dx.doi.org/10.1016/j.sigpro.2014.09.039
0165-1684/

Even if the two steps approach simplifies the image formation, the drift estimation by means of the JLS remains challenging when the dimension of the data set is large. Therefore, we introduce an iterative drift removal algorithm, which, under proper assumptions, is capable of significantly reduce the computational burden. The algorithm is based on Alternating Least Squares (ALS) which is a well known optimisation method, e.g. [6], strictly linked with Alternating Projection (AP) [7]. The main contribution of the paper is an original analysis proving that the ALS algorithm converges to the JLS solution. The analysis also gives geometrical insight, showing that the drift removal amounts at a data projection onto an appropriate subspace.

As an additional contribution, we discuss the application of the approach to the data of the Herschel satellite, which is a space telescope launched by the European Space Agency (ESA) in year 2009 [8]. Therefore, we specialise the approach to Herschel data, modelling the drift as a polynomial, and discuss the performance, using both theoretical analysis and numerical examples, showing that the two steps approach yields essentially optimal results for such data and that the ALS is an efficient implementation of the JLS. Indeed ALS algorithms are exploited in several Herschel data reduction software [9–11].

The paper is organised as follows. In Section 2 we introduce the data model and discuss the image formation problem. In Section 3 the two steps approach and the ALS algorithm are presented. The ALS algorithm is analysed in Section 4 and the application of the approach to Herschel data is discussed in Section 5. In Section 6 we investigate the performance, by presenting results obtained from simulated and true Herschel data. The conclusions are given in Section 7.

Notation: In the paper we use lowercase letters to denote vectors and uppercase letters to denote matrices. We use a superscript T to denote matrix or vector transposition, e.g. A^T. We use $|v|^2 = v^T v$ to denote the squared magnitude of the column vector v and say that two vectors v and w are orthogonal if $v^{T}w=0$. We use $\mathbf{E}[.]$ to denote expectation. Given a vector v with V elements v_i for $i=1,\ldots,V$, its average is $g=\frac{1}{V}\sum_{i=1}^{V} v_i$ and we denote by $\mathbf{v}$ the vector obtained from v by subtracting the average element-wise, i.e. $\mathbf{v}_i = v_i - g$ for $i=1,\ldots,V$.

2. Preliminaries

Consider an image of M pixels, represented by an $M \times 1$ vector m. The image is observed with an instrument producing $N \gg M$ readouts, represented by an $N \times 1$ *data* vector d. Assuming a linear, noisy instrument, the data vector is

$$d = Pm + n = s + n$$

where P is an $N \times M$, full-rank matrix, which depends on the instrument and on the observation protocol, n is a zero-mean, random *noise* vector and we introduced the *signal* vector $s = Pm$ representing the ideal, noiseless output. Given this set up, the image formation (or noise removal) problem is that of producing an estimate of m, denoted by $\overline{m}$, knowing d, P and the statistics of n. This is a classical problem having several established solutions. A simple and effective one is based on Least Squares (LS) and is summarised in the following.

Since for a generic image x the ideal output is Px, we can minimise $|d-Px|^2$ for varying x and obtain the image producing the output closest to the actual data vector. This image is the LS estimate of m and is given by [12]

$$\overline{m} = (P^T P)^{-1} P^T d. \tag{1}$$

LS estimation is an effective technique when the noise is white. Indeed, in white Gaussian noise it yields the Maximum Likelihood (ML) estimate. On the contrary, when the noise is correlated, LS performs poorly. In this case, by denoting by $C = E\{nn^T\}$ the noise covariance matrix, we can use a Generalised LS approach (GLS) [13] and minimise $|C^{-\frac{1}{2}}(d-Px)|^2$ for varying x. The solution is the GLS estimate given by

$$\overline{m} = (P^T C^{-1} P)^{-1} P^T C^{-1} d \tag{2}$$

which is the ML solution if the noise is Gaussian.

We now assume that the instrument introduces a drift too. The drift is represented with an $N \times 1$ *drift* vector $y = Xa$, where X is a given, full rank, $N \times K$ matrix, termed the drift matrix, and a is a $K \times 1$ coefficients vector, with $K \ll N$. Then the data vector is

$$d = Pm + Xa + n = s + y + n, \tag{3}$$

where X and P are known, full-rank matrices, m and a are deterministic, unknown vectors, and n is a zero-mean random vector. The last equation is a general and linear model, that will be used in this paper and is exploited, for example, in Diffuse Optical Imaging (DOI) [2] and functional Magnetic Resonance Imaging (fMRI) [1]. It is also well suited for space-born infrared data, as we see later. By constructing an appropriate drift matrix, the model can be specialised to several practical drift types, including polynomial, sinusoidal, piece-wise constant and piece-wise linear. We give an example in Section 5.1.

In the presence of the drift, both the LS and GLS estimates cannot be used directly, since the results are biased by the drift. However the approaches can be extended to the joint estimation of the image and coefficients. Indeed, by introducing the $N \times (M+K)$ matrix $A = [P, X]$ and the $(M+K) \times 1$ vector $z = [m^T, a^T]^T$, the data vector can be written as

$$d = Az + n$$

and the problem becomes that of estimating z. Then, by exploiting the LS approach and assuming that A is full rank, we obtain

$$\overline{z} = \begin{bmatrix} \overline{m} \\ \overline{a} \end{bmatrix} = (A^T A)^{-1} A^T d \tag{4}$$

which will be called the Joint LS (JLS) estimate and is the ML solution in white, Gaussian noise. When the noise is correlated, by exploiting the GLS approach, we obtain

$$\overline{z} = \begin{bmatrix} \overline{m} \\ \overline{a} \end{bmatrix} = (A^T C^{-1} A)^{-1} A^T C^{-1} d$$

which will be called the Joint GLS (JGLS) estimate and is the ML estimate in Gaussian noise.