HYPERSPECTRAL UNMIXING WITH SPARSE GROUP LASSO

Marian-Daniel Iordache[1,2], *José M. Bioucas-Dias*[2], and *Antonio Plaza*[1]

[1]Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications,
[2]Instituto de Telecomunicações, Instituto Superior Técnico, TULisbon, 1900-118, Lisboa, Portugal.
University of Extremadura, E-10071 Caceres, Spain.

ABSTRACT

Sparse unmixing has been recently introduced as a mechanism to characterize mixed pixels in remotely sensed hyperspectral images. It assumes that the observed image signatures can be expressed in the form of linear combinations of a number of pure spectral signatures known in advance (e.g., spectra collected on the ground by a field spectroradiometer). Unmixing then amounts to finding the optimal subset of signatures in a (potentially very large) spectral library that can best model each mixed pixel in the scene. In available spectral libraries, it is observed that the spectral signatures appear organized in groups (e.g. different alterations of a single mineral in the U.S. Geological Survey spectral library). In this paper, we explore the potential of the sparse group lasso technique in solving hyperspectral unmixing problems. Our introspection in this work is that, when the spectral signatures appear in groups, this technique has the potential to yield better results than the standard sparse regression approach. Experimental results with both synthetic and real hyperspectral data are given to investigate this issue.

1. INTRODUCTION

Sparse linear regression [1], [2] has recently been introduced in hyperspectral unmixing [3] to circumvent the need to obtain image-derived endmembers, which is often a difficult and ill-posed problem. In sparse linear regression, the observed vectors (hyperspectral vectors) are approximated with a linear combination of a "small" number or regressors (spectral signatures). The regressors weights (fractional abundances) are obtained by minimizing an objective function, often containing a quadratic data term and a sparsity-inducing regularizer, usually the ℓ_1-norm.

In many hyperspectral applications, the linear combinations of spectral signatures tend to be organized in groups. This is so either because not all combinations of materials are possible, or because the spectral library (the collection of spectral signatures) is built in order to account for variability. In the latter case, a single spectral signature may be replaced with a few slightly different ones such that small spectral variations can be modeled by linear combinations of these spectral signatures. In this work we are interested in the latter case. When the spectral signatures appear in groups, then the group lasso [4] and the sparse group lasso [5] regression techniques have the potential to yield better results than the standard sparse regression approach, as the former methods enforce sparseness on the groups instead of the singleton variables.

With the above rationale in mind, the aim of this work is twofold: i) to illustrate the potential of the sparse group lasso technique [5] in hyperspectral unmixing problems, and ii) to introduce a new optimization algorithm to solve efficiently the convex optimization problem underlying the sparse group lasso regression technique. This algorithm constitutes a generalization of the sparse unmixing by variable splitting and augmented Lagrangian (SUNSAL) introduced in [6]. Both SUNSAL and the new algorithm that we propose in this work are instances of the methodology introduced in [7].

2. THE LINEAR MODEL

A spectral library is a collection of spectra of pure materials, generally acquired in the field or in the laboratory. These spectra are collected in the form of a $B \times n$ matrix $\mathbf{A}$, where B is the number of bands and n is the number of available pure spectra. In this paper, we model an observed spectrum according to the linear mixing model, which assumes that the observed spectrum is a linear combination of the spectra of the pure materials (*endmembers*) present in the considered pixel. For a given pixel, the linear model can be expressed as

$$\mathbf{y} = \mathbf{M}\alpha + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is the observed spectrum, $\mathbf{M}$ is the mixing matrix containing the spectra of the endmembers, α is a vector of fractional abundances, and $\mathbf{n}$ is a vector collecting the errors that affect the measurement process (e.g., noise). The acquisition process imposes two constraints on the fractional

This work has been supported by the European Community's Marie Curie Research Training Networks Programme under contract MRTN-CT-2006-035927, Hyperspectral Imaging Network (HYPER-I-NET). Funding from the Spanish Ministry of Science and Innovation (HYPERCOMP/EODIX project, reference AYA2008-05965-C04-02) is also gratefully acknowledged.

abundances: the abundance non-negativity constraint (ANC) and the abundance sum-to-one constraint (ASC). The ASC is prone to criticisms in real scenarios, due to the so-called scale variability of the signatures. This is why, hereinafter, we use only the ANC. With these ideas in mind, we can now reformulate (1) to express $\mathbf{y}$ in terms of the matrix $\mathbf{A}$ as

$$\mathbf{y} = \mathbf{A}\mathbf{f} + \mathbf{n}, \tag{2}$$

where $\mathbf{f}$ is the new vector of fractional abundances. Due to the fact that only a few of the signatures contained in $\mathbf{A}$ are actually contributing to the observed spectrum, $\mathbf{f}$ is sparse, *i.e.*, it contains zero components. An important indicator regarding the difficulty to infer correct sparse solutions for a linear and overdetermined system of equations is the so-called *mutual coherence* [8], [9], [10], [11], defined as the largest cosine between any two columns of $\mathbf{A}$. It was shown [8] that the quality of the solution of a linear system of equations decreases when the mutual coherence increases. The mutual coherence of hyperspectral libraries is very close to one, which means that the problem is hard to solve, both in nature and from a numerical point of view.

3. GROUP SPARSE REGRESSION

Let assume that our library is a $B \times n$ matrix which can be partitioned into L groups with n_i elements in the group i. The notation $\mathbf{A}_i$ and α_i denotes, respectively, the spectral signatures and the fractional abundances corresponding to the group i. With this notation in place, the sparse group lasso (SGL) criterion adapted to our problem is

$$\begin{aligned} \min_{\mathbf{f}} \quad & \left\| \mathbf{y} - \sum_{l=1}^{L} \mathbf{A}_l \mathbf{f}_l \right\|_2^2 + \lambda_1 \sum_{l=1}^{L} \|\mathbf{f}_l\|_2 + \lambda_2 \|\mathbf{f}\|_1 , \quad (3) \\ \text{s.t:} \quad & \mathbf{f} \geq 0. \end{aligned}$$

where λ_1 and λ_2 are regularization parameters attached to the mixed $\ell_{2,1}$ norm $\sum_{l=1}^{L} \|\mathbf{f}_l\|_2$ and to the ℓ_1 norm $\|\mathbf{f}\|_1$, respectively. Compared with the SGL introduced in [5], we add the ANC. Accordingly, we designate the criterion (3) non-negative constrained sparse group lasso (NCSGL). Notice that the classical non-negative constrained least squares (NCLS) and the lasso subject to non-negativity, which we name non-negative constrained lasso (NCL) correspond, respectively, to setting $(\lambda_1 = 0, \lambda_2 = 0)$ and $(\lambda_1 = 0, \lambda_2 > 0)$.

To solve the optimization problem (3), we use the group SUNSAL (GSUNSAL) algorithm which is an ellaboration of the SUNSAL algorithm, introduced in [6]. GSUNSAL accounts for the $\ell_{2,1}$ regularizer in addition to the ℓ_1 and the non-negativity regularizers. In fact, GSUNSAL is an instance of the methodology introduced in [7] for solving ℓ_2 plus a linear combination of convex regularizers, based on the alternative direction method of multipliers (ADMM) [12]. For libraries not larger than a few thousand spectral signatures, GSUNSAL is very fast because the linear systems that appear in the ADMM iterations are solved efficiently by precomputing the inverse of the system matrix.

As a final note we emphasize that, by proper selection of the regularization parameters λ_1 and λ_2, GSUNSAL can be used to solve not only NCSGL, but also NCLS and NCL optimization problems.

4. RESULTS WITH SYNTHETIC DATA

In this section we give a brief illustration of the NCSGL criterion using simulated hyperspectral data. We built a dictionary $\mathbf{A}$ with $n = 450$ columns of size $B = 225$ and $L = 45$ groups of size 10. For each group, the elements of $\mathbf{A}$ are positive, line-wise independent and column-wise dependent, with correlation factor of 0.5, thus modeling the similarity of the spectral signatures inside a group as referred to above. The signal-to-noise (SNR) ratio, defined as $\mathbb{E}\|\mathbf{A}\mathbf{f}\|^2/\mathbb{E}\|\mathbf{n}\|^2$, was set to SNR = 30 dB in our simulation. The number of active groups was set to $k = 5$. Each group has five components set to value 0.0222, two components set to value 0.0444, and three components set to value 0. The regularization parameters λ_1 and λ_2 were hand-tuned for optimal performance.

Figure 1 shows the estimated fractional abundances for NCLS (top row), NCL (middle row), NCGSL (bottom row). The corresponding ℓ_1 norm of the errors is 0.14, 0.11, and 0.07, respectively. Perhaps more important than the obtained ℓ_1 error is the fact that NCLS generated the sparsest solution, as can be inferred from the number of non-zero components: 279 for NCLS, 163 for NCL, and 91 for NCGSL. This is a consequence of using the $\ell_{2,1}$ regularizer defined for groups; when a group is not selected, all its components are set to zero. On the other hand, when a group is selected, the ℓ_1 regularizer enforces sparseness on its components.

5. RESULTS WITH REAL DATA

The scene used in our real data experiments is the well-known AVIRIS Cuprite data set, available online in reflectance units[1]. This scene has been widely used to validate the performance of endmember extraction algorithms. The portion used in experiments corresponds to a 204×151-pixel subset of the sector labeled as f970619t01p02_r02_sc03.a.rfl in the online data. The scene comprises 224 spectral bands between 0.4 and 2.5 μm, with nominal spectral resolution of 10 nm. Prior to the analysis, bands 1–2, 105–115, 150–170, and 223–224 were removed due to water absorption and low SNR in those bands, leaving a total of 188 spectral bands. The Cuprite site is well understood mineralogically, and has several exposed minerals of interest, all included in the USGS library considered in experiments, denoted splib06[2] and released in September 2007. In our experiments, we use spectra

[1] http://aviris.jpl.nasa.gov/html/aviris.freedata.html

[2] http://speclab.cr.usgs.gov/spectral.lib06

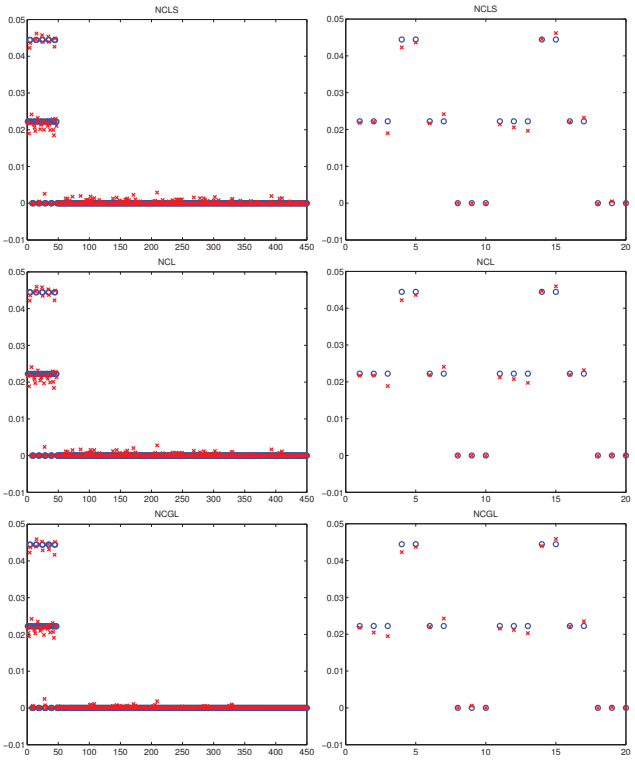

Fig. 1. Estimated fractional abundances for NCLS (top row), NCL (middle row), NCGSL (bottom row); the corresponding ℓ_1 norm of the errors is 0.14, 0.11, and 0.07, respectively. Following the same order, the number of non-zero components is 279, 163, and 91. The plots in the right column are a zoom of the corresponding plots in the left column for the first 20 components.

obtained from this library as input to GSUnSAL. The spectral library built in this way contains 240 materials (minerals) organized in 55 groups. Each group contains a number of spectra varying between 1 and 17. As for the considered image, we removed the (same) noisy bands prior to the unmixing process. The mutual coherence of the spectral library is very close to one, which makes the problem very difficult to solve. The parameters used in this experiment were empirically set to: $\lambda_1 = 0.05$ and $\lambda_2 = 0.01$.

For comparative purposes, we use a mineral map produced in 1995 by USGS, in which the Tetracorder 3.3 software product [13] was used to map different minerals present in the Cuprite mining district[3]. It should be noted that the Tetracorder map is only available for hyperspectral data collected in 1995, while the publicly available AVIRIS Cuprite data was collected in 1997. Therefore, a direct comparison between the 1995 USGS map and the 1997 AVIRIS data is not possible. However, the USGS map serves as a good indicator for qualitative assessment.

Fig. 2 shows a qualitative comparison between the classification maps extracted from the Tetracorder map and the fractional abundances inferred by GSUnSAL in the considered image, for four different minerals (alunite, buddingtonite, chalcedony and montmorillonite). As it can be seen in Fig. 2, the unmixing results show a good distribution of the minerals of interest. The fractional abundances obtained with GSUnSAL are generally higher in the regions assigned to the respective materials. Additionally, we mention that, in the unmixing results, the average number of endmembers with abundances higher than 0.05, per pixel, is 4.16, while the average number of groups with (total) abundances higher than 0.05, per pixel, is 4.90. The small difference between the two values is due to the fact that GSUnSAL method enforces the sparseness both at group and individual level, which means that, inside the selected groups, the algorithm uses a minimum number of members to explain the data. This result is in line with the information provided by the Tetracorder classification map, in which the four selected endmembers are quite dominant in the scene.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we gave experimental evidence of the potential of the sparse group lasso technique in solving hyperspectral unmixing problems. We further introduced GSUNSAL, a new algorithm to solve the NCGSL optimization problem. GSUnSAL is able to infer with good accuracy the abundance fractions both in simulated and real environment. The advantage of the method is that, compared to the iterative methods which consider all the members of the spectral library as potential endmembers, it can drop at once entire groups of materials from the solution, along the iterations. Although the ob-

[3]http://speclab.cr.usgs.gov/PAPERS/tetracorder/FIGURES/fig9b.cuprite95.tgif.2.2um_map.gif

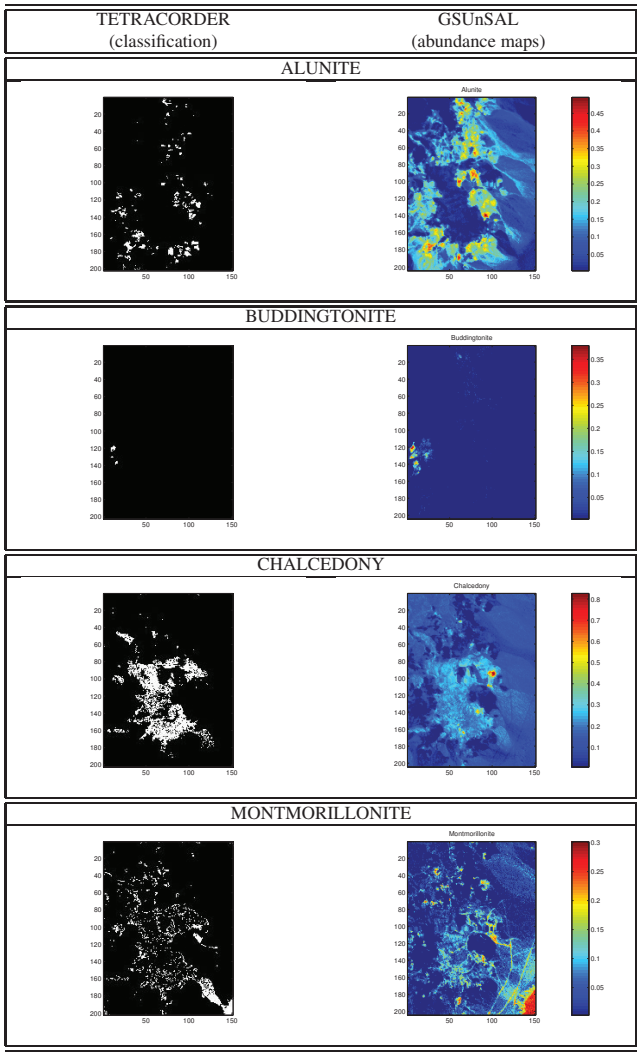

Fig. 2. Qualitative comparison between the fractional abundance maps estimated by GSUnSAL and the classification maps produced by USGS Tetracorder for the 204×151-pixel AVIRIS Cuprite scene.

tained results are very promising, further experimentation is needed in order to fully substantiate our contributions. Also, despite the fact that the algorithm is fast compared to other methods, a possible direction in our future research is the parallel implementation of the algorithm, possibly on GPUs.

7. REFERENCES

[1] B.K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM journal on computing*, vol. 24, no. 2, pp. 227–234, 1995.

[2] D.L. Donoho, M. Elad, and V.N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[3] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, 2011, in press.

[4] M. Y.and Lin M. Yuan, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B*, vol. 68, no. 1, pp. 49–67, 2006.

[5] R. J. Friedman, T. Hastie and Tibshirani, "A note on the group lasso and a sparse group lasso," *preprint*, 2010.

[6] J.M. Bioucas-Dias and M.A.T. Figueiredo, "Alternating Direction Algorithms for Constrained Sparse Regression: Application to Hyperspectral Unmixing," *Proceedings of the IEEE/GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS'2010)*, Reykjavik, Iceland, 2010.

[7] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, 2010, in press.

[8] A.M. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations," *Information Theory, IEEE Transactions on*, vol. 54, no. 11, pp. 4813–4820, 2008.

[9] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, pp. 34–81, 2009.

[10] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[11] E.J. Candès, J.K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207, 2006.

[12] J. Eckstein and D. Bertsekas, "On the Douglas Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 5, pp. 293–318, 1992.

[13] R. N Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *Journal of Geophysical Research*, vol. 108 [E12], 5131, 2003.