AI 539/ECE 569 final project: Sparse regression-based hyperspectral unmixing

Zhanpei Fang
Oregon State University
fangzha@oregonstate.edu

Abstract

This final project report summarizes sparse regression-based methods for hyperspectral unmixing as described in Iordache, Bioucas-Dias and Plaza (2011), "Sparse Unmixing of Hyperspectral Data" [1] and other related papers. I present Bioucas-Dias's reformulation of the spectral unmixing problem as a sparse regression (SR) problem and describe several sparse unmixing algorithms. The authors assess these algorithms' potential across a number of metrics over simulated and real aerial HSI data; I summarize my understanding of their method and apply a selection of the authors' described methods to a different hyperspectral dataset, as well as providing a brief gloss on some new developments in the field since the paper's original publication.

1 Introduction

1.1 Hyperspectral images

Hyperspectral data 'cubes' consist of a stack of images across hundreds of spectral bands. For example, the AVIRIS Next-Generation (AVIRIS-NG) instrument measures the wavelength range from 380 nm to 2510 nm with a nominal spectral resolution of 5 nm. Each pixel is represented by a signature which contains spectral information from all of the surface materials as well as atmospheric effects which reflect light back into the sensor. Processing tasks which are often undertaken for hyperspectral remote sensing include dimensionality reduction, target or anomaly detection[1], change detection, classification, and spectral unmixing. This last problem, of estimating the fraction of pixel area covered by each material present in the scene, is the one which this paper concerns itself with.

Mixed pixels are common in remotely-sensed hyperspectral images due to:

- insufficient spatial resolution of the imaging spectrometer, such that the resulting spectral measurement is a composite of individual pure spectra which correspond to materials jointly occupying a single pixel at a macroscopic level; or
- intimate mixing effects which arise when distinct materials are mixed on a microscopic (*intimate*) level; as a consequence, increasing the spatial resolution cannot fully solve the problem, as intimate mixtures happen at microscopic scales complicating the analysis with nonlinear effects.

In addition to spectral mixing effects, other phenomena which may affect the process of analyzing remotely sensed hyperspectral data include atmospheric interference, multiple scattering, shadows and variable illumination conditions. The rich spectral resolution available from hyperspectral imagers

[1] Anomaly detection for methane plumes in hyperspectral imagery is mostly what I did at my job immediately prior to entering the PhD program, for example with matched filtering or Reed-Xiaoli algorithm with a constant false alarm rate assumption based on Mahalanobis distance.

as well as other information may be exploited to unmix hyperspectral pixels into their component endmembers, as this report will begin to describe.

1.2 Linear mixing model and spectral unmixing problem formulation

The general goal of linear spectral unmixing is to find a set of macroscopically pure spectral components (endmembers) which can be used to unmix all other pixels in the data, and estimate the fractional abundances of these endmembers in a mixed pixel collected by an hyperspectral imager.

The linear mixture model operates under the assumption that the spectral response of a pixel in a given spectral band is a linear combination of the all endmembers present in the pixel for that band. This can be written for a given pixel as follows:

$$y_i = \sum_{j=1}^{q} m_{ij}\alpha_j + n_i$$

where y_i is the measured value of the reflectance at spectral band i, m_{ij} is the reflectance of the jth endmember at that band i, α_j is the fractional abundance of the jth endmember and n_i is the noise term for that band. If the sensor collects at L spectral bands, this can be written as a matrix equation:

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

where $\mathbf{y}$ is a $L \times 1$ vector, $\mathbf{M}$ is a $L \times q$ matrix, $\boldsymbol{\alpha}$ is $q \times 1$, and $\mathbf{n}$ is $L \times 1$. In a typical hyperspectral unmixing scenario, the goal is to, given $\mathbf{Y} \equiv \{\mathbf{y}_i \in \mathbb{R}^L, i = 1, \ldots, n\}$ of n observed L-dimensional spectral vectors, estimate mixing matrix $\mathbf{M}$ and fractional abundances $\boldsymbol{\alpha}$ for each pixel in the scene. The physical constraints on this problem are:

- Abundance non-negativity constraint (ANC): $\boldsymbol{\alpha} \succeq \mathbf{0}$, i.e. the spectra of the endmembers cannot be negative as negative reflectances would be nonphysical;
- Abundance sum-to-one constraint (ASC): $\mathbf{1}^T\boldsymbol{\alpha} = 1$.

Algorithms for spectral linear unmixing typically take a three-step approach:

1. Dimensionality reduction: identify the subspace spanned by the columns of $\mathbf{M}$
2. Endmember determination: identify the columns of $\mathbf{M}$
3. Inversion: For each pixel, identify the vector of proportional abundances $\boldsymbol{\alpha}$

1.2.1 Linear versus nonlinear spectral unmixing

Linear unmixing assumes that the macroscopically pure components are homogeneously distributed in separate patches within the field of view, so the spectra collected by the imaging spectrometer can be expressed as a linear combination of endmembers weighted by their abundances. Nonlinear unmixing describes mixed spectra (both physically and statistically) by assuming that the microscopically pure components are intimately mixed inside the pixel in a nonlinear fashion and part of the source radiation is multiply scattered before being collected at the sensor, thus one of the challenges is to derive the nonlinear function; this requires detailed *a priori* physical knowledge about the materials and object geometry. The authors [1] focus on linear unmixing due to its generality and ease of implementation, even though nonlinear unmixing may provide better results with respect to spectral characterization in many cases.

1.2.2 Disadvantages of endmember-based approach: pure pixel assumption

Up to the point of the paper [1]'s publication, the majority of unmixing algorithms had been designed under the pure pixel assumption. Spectral unmixing algorithms using the pure pixel assumption require the presence of pure pixels in the scene for endmember extraction; due to spatial resolution and mixing phenomena at macroscopic and microscopic levels, this assumption cannot be guaranteed in many real-world cases. In the case that the scene does not contain any pure signatures, the most feasible option is to use unmixing algorithms which do not assume the presence of pure signatures (i.e. minimum volume enclosing algorithms). Spectral unmixing algorithms without the

pure pixel assumption often find the minimum-volume simplex which encompasses all observations and generate endmember signatures, called *virtual* endmembers, which often do not relate to real physically meaningful signatures.

2 Methods

2.1 Reformulating spectral unmixing as a sparse regression problem

The authors of the paper, following [2], reformulate the spectral unmixing problem in terms of sparse reconstruction (SR) as follows. The SR approach allows the authors to address problems related to the unavailability of pure spectral signatures in the data by taking a semi-supervised approach, by taking advantage of the availability of spectral libraries taken through field measurements, known and organized in advance, and the typically very low number of endmembers present in the scene relative to said library's size.

Spectral vectors can be expressed as linear combinations of a few pure spectral signatures obtained from a (potentially very large) spectral library $\mathbf{A}$:

$$\mathbf{y} = \sum_{i \in I} \mathbf{a}_i \mathbf{x}_i = \mathbf{A}\mathbf{x}$$

The unmixing problem is then, given $\mathbf{y} \in \mathbb{R}^{L \times n}$ and $\mathbf{A} \in \mathbb{R}^{L \times m}$, to find the sparsest solution of $\mathbf{y} = \mathbf{A}\mathbf{x}$. The library $\mathbf{A} \in \mathbb{R}^{L \times m}$ where L is the number of spectral bands and m is the number of materials in the library, represents an underdetermined system ($L < m$). If $\mathbf{x} \in \mathbb{R}^m$ is the fractional abundance vector with respect to library $\mathbf{A}$ (where we say it's k-sparse if it has k non-zero components), the SR problem can be written as

$$\text{(SR)} \qquad \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \delta, \mathbf{x} \succeq \mathbf{0}, \mathbf{1}^T\mathbf{x} = 1 \tag{2}$$

Here $\|\mathbf{x}\|_0$ is the number of non-zero components of $\mathbf{x}$, and $\delta \geq 0$ is the error tolerance with respect to expected errors due to noise and modeling.

This method has the advantage of eliminating abundance estimation's need for a pure pixel assumption, which is helpful since most hyperspectral images have no pure pixels at all, and sidestepping the endmember extraction step (which uses algorithms such as vertex component analysis, pixel purity index, N-FINDR, orthogonal subspace projection, etc.) as well as obviating the need to estimate *a priori* the number of endmembers in a scene completely.

In summary, the sparse unmixing problem is then framed as finding the optimal subset of signatures in a (potentially very large) library which best models each pixel in a scene. This is a combinatorial problem which is sparse and difficult to solve; it requires efficient sparse regression techniques based on sparsity-inducing regularizers since the number of endmembers within a mixed pixel is so much smaller than the entire library. Due to these computational issues, fast algorithms are required to solve them, as described below. When implementing, the difference between the ground library and the image data also requires atmospheric correction to overcome.

2.1.1 Exact solutions

In the case where noise is zero and the ANC and ASC constraints are not enforced, the SR optimization problem is

$$(P_0): \quad \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \tag{3}$$

This problem is NP-hard, but can be approximated with greedy pursuit algorithms like orthogonal matching pursuit (OMP) [3] or basis pursuit (BP) [4].

2.1.2 Approximate solutions

If $\mathbf{n}$ in the observation model is nonzero and ANC/ASC constraints are not enforced, and we want to find an approximate solution to the SR problem above, convex approximations to P_0 can be obtained with constrained basis pursuit denoising (CBPDN). The relaxed problem is:

$$\text{(Convex approximation)} \quad \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \delta,\ \mathbf{x} \succeq \mathbf{0} \tag{4}$$

which is equivalent to the problem:

$$\text{(CSR)} \quad \min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{x} \succeq \mathbf{0} \tag{5}$$

A striking result is that in certain circumstances, depending on the coherence among the columns of matrix $\mathbf{A}$, BP(DN) yields the sparsest solution [5, 6].

2.2 Algorithms and authors' contribution

The optimization problems formulated for spectral unmixing have a huge number of variables which make them difficult to solve. Sparse regression requires fast algorithms, based on ADMM (alternating direction method of multipliers) which iteratively solves a complex problem by decomposing it into subproblems which are easier to solve. Efficient solvers for CBPDN include SUnSAL, the initial formulation proposed by Bioucas-Dias in 2009 [2], and CSUnSAL. The former casts SR-based unmixing in the form of a convex, constrained, $\ell_2 - \ell_1$ norm optimization problem, and proposes to solve it with ADMM.

The authors of [1] evaluate the performance of several (available and new at the time of publication) sparse and non-sparse algorithms for spectral unmixing applications using extensive experimental validation, systematically studying the application of CBPDN to SLU. They use six libraries $\mathbf{A}_1, \ldots, \mathbf{A}_6$ and test algorithms over simulated data (where endmembers are randomly selected from the libraries and fractional abundances are uniformly distributed over the simplex) and real data (the AVIRIS Cuprite flightlines, using a calibrated version of the USGS spectral library $\mathbf{A}_1$). The algorithms tested are:

- OMP: orthogonal matching pursuit (Algorithm 1 in their paper).
- SUnSAL: Spectral Unmixing via variable Splitting Augmented Lagrangian, solves equation 5. The ideal trade-off between the reconstruction error and the sparsity of the solution vector is targeted in each pixel. Solves (P_1) and (P_1^δ) problems from the paper.
- SUnSAL+: Solves (P_1^+) and $(P_1^{\delta+})$ problems from the paper.
- CSUnSAL: Constrained SUnSAL, solving equation 4. Solves (P_1) and (P_1^δ) problems from the paper.
- CSUnSAL+: Constrained SUnSAL solving the (P_1^+) and $(P_1^{\delta+})$ problems from the paper.
- Iterative spectral mixture analysis (ISMA), which iteratively finds a set of optimal endmembers by computing change in RMSE after attempting to reconstruct the original scene based on the estimated fractional abundances.

In the experiments with simulated data the authors do the calculation of approximate solutions without and with imposing the ASC constraint, and compare the unmixing algorithms with respect to computational complexity and influence of noise. The ability to obtain useful sparse solutions for an under-determined system of equations depends on the amount of coherence between the columns of the system matrix and sparseness of the original abundance fractions; the authors find (in their Table 1) that hyperspectral libraries unfortunately exhibit high mutual coherence, but hyperspectral mixtures are very sparse (the cardinality of the solution, i.e. number of materials participating in a mixed pixel, being on the order of $k = 5$ in many cases) which can mitigate the high coherence problem. The authors plot reconstruction errors at an SNR of 30 dB (Fig. 8 in their paper).

For the real-data experiments they run the algorithms over the AVIRIS Cuprite dataset and show qualitative results. The authors conclude that their new ADMM-based algorithms are superior to previous methods both in accuracy as well as runtime.

2.3 Later work

2.3.1 Sparse regression with spatial regularization (total variation)

The authors of [7] propose a technique to take into account the rich spatial information present in a hyperspectral image. Total variation spatial regularization consists of using a total variation (TV) regularizer to enforce spatial homogeneity by including this term into the original objective function:

$$\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_{1,1} + \lambda_{TV}\text{TV}(\mathbf{X}) \quad \text{s.t.} \quad \mathbf{X} \succeq \mathbf{0}$$

where the TV term promotes pieceswise constant (or smooth) transitions in the fractional abundance of the same endmember among neighboring pixels:

$$\text{TV}(\mathbf{X}) \equiv \sum_{\{i,j\}\in\varepsilon} \|\mathbf{x}_i - \mathbf{x}_j\|_1$$

The resulting SUnSAL-TV combines the idea of sparse unmixing with that of exploiting spatial-contextual information present in the hyperspectral images by including the TV regularizer on top of the sparse unmixing formulation. Total variation spatial regularization produces spatially smooth abundance fractions which improve sparse unmixing performance even in very high noise conditions.

2.3.2 Exploiting group structure

The authors later [8] propose a framework to exploit the group structure of spectral libraries, often organized in groups of spectra which describe variations of the same material (so that the number of distinct materials in the library is smaller than the number of spectra), based on sparse group lasso and collaborative hierarchical lasso techniques, which impose sparsity on the endmember as well as the group level, per pixel and per groups of pixels, respectively. Sparse group lasso imposes sparsity at the individual and group levels, acting individually over each pixel of the image; the collaborative hierarchical lasso imposes sparsity across the pixels in the image given the specific structure of spectral libraries.

3 Experiment

I implement SUnSAL in Python, adapting the original MATLAB implementation by Bioucas-Dias [2], which is included in the code submission as a reference. The script solves the following problems:

1. Basis pursuit denoising with $\ell_2 - \ell_1$ (BPDN): $\lambda > 0$, ANC and ASC not enforced.
2. Constrained basis pursuit denoising with $\ell_2 - \ell_1$ (CBPDN): $\lambda > 0$, ANC enforced, ASC not enforced.
3. Constrained least squares (CLS): $\lambda = 0$, ANC enforced, ASC not enforced.
4. Fully constrained least squares (FCLS): $\lambda = 0$, ANC enforced, ASC not enforced. In this case the regularizer $\|\mathbf{x}\|_1$ is not used since it is constant.

Because intensive atmospheric correction felt probably out of scope–I didn't particularly want to futz around with radiative transfer models for this project–I found a hyperspectral dataset taken in-laboratory with endmembers by Zhao et al. [9], designed to help reduce the gap between theory and practice. The dataset authors systematically image in-laboratory scenes (see Figure 1) such as a printed checkerboard, mixed quartz sands, and reflection with a vertical board, with a push-broom hyperspectral camera in order to provide 36 mixtures with 256 wavelength bands, from 400-1000 nm, where the pure material spectra and material composition methods are known. This choice of dataset is intended to extend and complement the results of the paper in question [1], which like most unmixing papers uses synthetic and real airborne data (most commonly for unmixing papers, the AVIRIS Cuprite dataset, with spectra from the USGS spectral library), both of which have their drawbacks.

I also tried the DC1 and DC2 simulated datasets which are popular for sparse unmixing[2], which is composed of 240 spectral signatures from the USGS library; DC1 is of size (75,75) and simulated

[2]Copies of the DC1 and DC2 datasets borrowed from the codebase from this paper: [10].

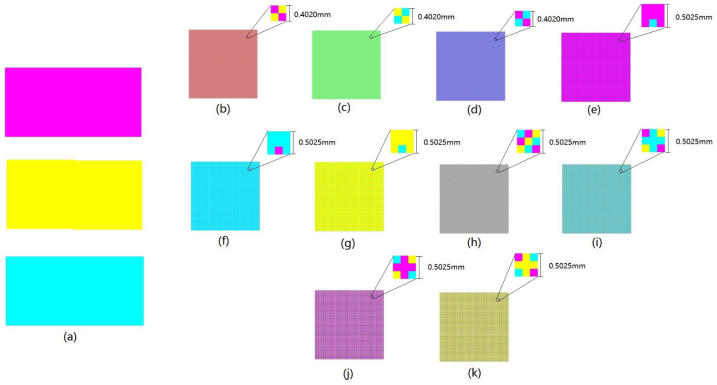

Fig. 2: Illustration of Scene I. (a): Three pure color blocks (magenta, yellow and cyan). (b) to (k): Checkerboards consisting of $0.2010mm \times 0.2010mm$ or $0.1675mm \times 0.1675mm$ squares with different layouts to form ten mixtures. Zoomed-in part of each subfigure illustrate the sizes and arrangements of color blocks.

Figure 1: Reproduction of Figure 2 from the Zhao et al. [9] paper which describes how one of the three scenes from their laboratory HSI dataset is set up.

using a linear mixing model with five endmembers selected from the library, and abundance maps composed of five rows of square regions uniformly distributed over the spatial dimension. DC2 is of size (100, 100) and simulated using a linear mixing model with nine endmembers; the abundance maps are sampled from a Dirichlet distribution centered at a Gaussian random field to have piecewise smooth maps with steep transitions.

4 Results and discussion

4.1 Laboratory dataset

The predicted total abundances for each mixture for each scene are reproduced below. For this run I set $\lambda = 0$ and did not enforce the ANC or ASC constraints, which was the only combination of options which gave a close-to-reasonable solution (other solutions blew up). The abundance ground truths for these scenes and mixtures are reproduced from the Zhao et al. paper [9].

	scene	mixture	magenta	yellow	cyan
2	scene1	1	41.425	64.0788	-5.50376
3	scene1	2	51.4741	1.15299	47.3729
4	scene1	3	4.27308	44.2666	51.4603
8	scene1	4	2.59883	88.4444	8.95673
9	scene1	5	2.40764	9.03772	88.5546
6	scene1	6	87.7156	5.72777	6.55659
5	scene1	7	31.0325	38.3799	30.5876
0	scene1	8	22.6823	22.6959	54.6218
1	scene1	9	22.4993	58.2851	19.2156
7	scene1	10	53.9775	25.6912	20.3313

	scene	mixture	magenta	yellow	cyan
0	scene2	1	-7.04232	30.7105	17.4422
2	scene2	2	22.0616	3.43199	74.1481
1	scene2	3	-7.67986	20.2544	14.9134
6	scene2	4	-5.42088	43.4765	15.9579
5	scene2	5	14.2839	2.80634	81.6006

```
|  3 | scene2  |         6 |  38.7996  |  0.947416 | 60.7462 |
|  4 | scene2  |         7 | -15.0949  | 32.0733   | 41.3851 |
|  9 | scene2  |         8 |  10.3156  | 39.5402   | 45.079  |
|  8 | scene2  |         9 | -16.503   | 24.6229   | 40.9011 |
| 12 | scene2  |        10 |  -1.57824 | 26.8009   | 36.7238 |
| 11 | scene2  |        11 |  -6.08881 | 21.6601   | 34.3152 |
|  7 | scene2  |        12 | -20.1703  | 18.865    | 52.2895 |
| 10 | scene2  |        13 |  -2.37912 | 12.2249   | 46.4819 |
| 13 | scene2  |        14 |  -2.17881 | 33.5014   | 34.8665 |
+----+---------+-----------+-----------+-----------+---------+
```

```
+----+---------+-----------+-----------+----------+-----------+
|    | scene   |   mixture |   magenta |   yellow |      cyan |
|----+---------+-----------+-----------+----------+-----------|
|  4 | scene3  |         1 |  38.0883  |  60.0129 |  1.89876  |
|  5 | scene3  |         2 |  19.8704  |  63.1615 | 16.9681   |
|  6 | scene3  |         3 |  52.166   |  16.5263 | 31.3077   |
|  2 | scene3  |         4 |  71.2343  |  27.5462 |  1.21956  |
|  3 | scene3  |         5 |  30.4442  |  47.3572 | 22.1986   |
|  1 | scene3  |         6 |  78.547   |  20.7278 |  0.725164 |
|  0 | scene3  |         7 |   9.05204 |  13.0575 | 77.8904   |
|  7 | scene3  |         8 |   2.55175 | 101.479  | -4.03036  |
+----+---------+-----------+-----------+----------+-----------+
```

TABLE I: Abundance ground-truth of Scene I.

	Magenta	Yellow	Cyan
Mixture 1	50.00%	50.00%	0
Mixture 2	0	50.00%	50.00%
Mixture 3	50.00%	0	50.00%
Mixture 4	88.89%	0	11.11%
Mixture 5	11.11%	0	88.89%
Mixture 6	0	88.89%	11.11%
Mixture 7	33.33%	33.33%	33.33%
Mixture 8	22.22%	22.22%	55.56%
Mixture 9	55.56%	22.22%	22.22%
Mixture 10	22.22%	55.56%	22.22%

TABLE II: Abundance ground-truth of Scene II.

	Red	Green	Blue	White
Mixture 1	50%	50%	0	0
Mixture 2	0	0	50%	50%
Mixture 3	70%	30%	0	0
Mixture 4	30%	70%	0	0
Mixture 5	0	0	70%	30%
Mixture 6	0	0	30%	70%
Mixture 7	33.33%	33.33%	33.33%	0
Mixture 8	0	33.33%	33.33%	33.33%
Mixture 9	50%	20%	30%	0
Mixture 10	25%	25%	25%	25%
Mixture 11	40%	20%	20%	20%
Mixture 12	50%	0	50%	0
Mixture 13	33.33%	0	33.33%	33.33%
Mixture 14	20%	40%	20%	20%

TABLE III: Abundance ground-truth of Scene III.

	Magenta	Yellow	Cyan
Mixture 1	50.00%	0	50.00%
Mixture 2	22.22%	22.22%	55.56%
Mixture 3	50.00%	50.00%	0
Mixture 4	88.89%	0	11.11%
Mixture 5	33.33%	33.33%	33.33%
Mixture 6	1	0	0
Mixture 7	0	1	0
Mixture 8	0	0	1

I think the results on the lab data did not work so well precisely because the library is *not* sparse, consisting of only the 3 pure endmembers which were included in the dataset, so these results make some sense.

4.2 Simulated data

The signal to reconstruction error is computed in decibels as $\text{SRE}(\mathbf{X}, \hat{\mathbf{X}}) = 10 \log_{10} \|\mathbf{X}\|_F / \|\mathbf{X} - \hat{\mathbf{X}}\|_F$ and reported below at a SNR of 20 dB. I'm not fully sure what to make of the negative SREs but the one for the BPDN variant of the problem is at least reasonable.

```
running over dataset: DC1
SRE for lambda=0, anc=False, asc=False: 11.025841
```

```
SRE for lambda=1e-6, anc=False, asc=False (BPDN): 11.025841
SRE for lambda=1e-6, anc=True, asc=False (CBPDN): -13.046763
SRE for lambda=0, anc=True, asc=False (CLS): -13.046763
SRE for lambda=0, anc=True, asc=True (FCLS): -14.016166
running over dataset: DC2
SRE for lambda=0, anc=False, asc=False: 8.899101
SRE for lambda=1e-6, anc=False, asc=False (BPDN): 8.899101
SRE for lambda=1e-6, anc=True, asc=False (CBPDN): -11.034189
SRE for lambda=0, anc=True, asc=False (CLS):  -11.034189
SRE for lambda=0, anc=True, asc=True (FCLS): -11.123086
```

5 Conclusions

The paper authors [1] show that the sparse regression framework has strong potential for linear hyperspectral unmixing, tailor new regression criteria to cope with the high coherence of hyperspectral libraries, and develop optimization algorithms for the above criteria. I enjoyed learning about sparse regression in this context, given my familiarity with the curse-of-dimensionality challenges of this type of dataset and other common analysis challenges in hyperspectral imaging.

What I would have liked to do given extra time: I would have liked to try the method with some of the other commonly-benchmarked-against aerial hyperspectral datasets (Indian Pines, Jasper Ridge, Botswana, Pavia, Salinas, Kennedy Space Center, Moffett Field, etc.) using an atmospheric correction routine, efficiently computed for AVIRIS and similar data with `isofit` [11], a set of utilities developed by a JPL team which fits surface, atmospheric and instrument models to the aerial data (with AVIRIS in mind) and admits a number of radiative transfer models including MODTRAN, LibRadTran and 6S. I would have liked to systematically implement the five other methods tested by the authors in their paper, as well as later methods such as SUnSAL-TV [7], the group lasso regularizer described, CLSUnSAL (collaborative sparse unmixing, which solves $\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_{2,1}$ subject to $\mathbf{X} \succeq \mathbf{0}$, $\mathbf{1}^T\mathbf{x} = 1$, where $\|\mathbf{X}\|_{2,1} = \sum_{k=1}^{m} \|\mathbf{x}^k\|_2$ is the $\ell_{2,1}$ norm), etc. and compare against the vanilla method. It would further be interesting and worthwhile to compare the results of these unmixing algorithms on aerial datasets versus those from space-borne imagers such as the Hyperion mission, or some of the new commercial hyperspectral or "superspectral" satellite datasets now coming online from companies such as Pixxel, Planet, and expected data from Landsat Next.

In the years since its first formulation, SUnSAL has formed the robust basis for diverse algorithms such as those which account for spatial homogeneity, data collaborativity, structured dictionaries, non-convex formulations, library pruning and data mismatch alleviation and more; as well as newer machine-learning approaches, for example SUnCNN [10] which combines SUnSAL with a convolutional neural network architecture; while remaining a competitive method in sparse regression unmixing [12] over a decade since its initial formulation.

References

[1] Marian-Daniel Iordache, José M Bioucas-Dias, and Antonio Plaza. Sparse unmixing of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2014–2039, 2011.

[2] José M Bioucas-Dias and Mário AT Figueiredo. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pages 1–4. IEEE, 2010.

[3] Yagyensh Chandra Pati, Ramin Rezaiifar, and Perinkulam Sambamurthy Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers*, pages 40–44. IEEE, 1993.

[4] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43(1):129–159, 2001.

[5] David L Donoho, Michael Elad, and Vladimir N Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on Information Theory*, 52(1):6–18, 2005.

[6] Emmanuel J Candes and Justin Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Foundations of Computational Mathematics*, 6(2):227–254, 2006.

[7] Marian-Daniel Iordache, José M Bioucas-Dias, and Antonio Plaza. Total variation spatial regularization for sparse hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11):4484–4502, 2012.

[8] Marian-Daniel Iordache, José M Bioucas-Dias, and Antonio Plaza. Hyperspectral unmixing with sparse group lasso. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 3586–3589. IEEE, 2011.

[9] Min Zhao, Jie Chen, and Zhe He. A laboratory-created dataset with ground truth for hyperspectral unmixing evaluation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2170–2183, 2019.

[10] Behnood Rasti and Bikram Koirala. Suncnn: Sparse unmixing using unsupervised convolutional neural network. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2021.

[11] David R Thompson, Kerry Cawse-Nicholson, Zachary Erickson, Cédric G Fichot, Christian Frankenberg, Bo-Cai Gao, Michelle M Gierach, Robert O Green, Daniel Jensen, Vijay Natraj, et al. A unified approach to estimate land and water reflectances with uncertainties for coastal imaging spectroscopy. *Remote Sensing of Environment*, 231:111198, 2019.

[12] Mario Parente and Marian-Daniel Iordache. Sparse unmixing of hyperspectral data: The legacy of SUnSAL. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 21–24. IEEE, 2021.