Stagewise Polytope Faces Pursuit for Recovery of Sparse Representations

Mark D. Plumbley and Marco Bevilacqua
Queen Mary University of London
{mark.plumbley, marco.bevilacqua}@elec.qmul.ac.uk

In sparse representations or compressed sensing, we are typically interested in finding a sparse vector $\mathbf{s}$ which satisfies an underdetermined system of linear equations $\mathbf{x} = \mathbf{A}\mathbf{s}$. We can do this by searching for the minimum ℓ_1 norm, or *Basis Pursuit* (BP) solution, which in standard form [2] is the left-hand side member of the following equality:

$$\min_{\tilde{\mathbf{s}}} \left\{ \mathbf{1}^T \tilde{\mathbf{s}} \;\middle|\; \mathbf{x} = \tilde{\mathbf{A}}\tilde{\mathbf{s}},\ \tilde{\mathbf{s}} \geq 0 \right\} = \max_{\mathbf{c}} \left\{ \mathbf{x}^T \mathbf{c} \;\middle|\; \tilde{\mathbf{A}}^T \mathbf{c} \leq \mathbf{1} \right\} \geq 0. \tag{1}$$

The right-hand side of (1) is the so called *dual* Linear Program (LP) which has an optimum $\mathbf{c}^*$ associated with any optimum $\tilde{\mathbf{s}}^*$ of the LP represented by the left-hand side of (1). In previous work [6] we introduced a greedy algorithm, *Polytope Faces Pursuit*, to search for solutions of (1), adopting a path-following approach through the relative interior of the faces of the polar polytope $P^* = \{\mathbf{c} \mid \tilde{\mathbf{A}}^T \mathbf{c} \leq \mathbf{1}\}$ associated with the dual LP problem [3]. The PFP algorithm is in the style of *Matching Pursuits* [5] in that it adds new basis vectors one at a time, each corresponding to a hyperplane of the polytope, but it uses an adjusted correlation criterion to determine which basis vector to add, and can switch out basis vectors as necessary.

Recently new 'Stagewise' algorithms have been introduced (e.g. [4] and [1]) that add several atoms per iteration. In the current work we similarly generalize our previous PFP algorithm to a stagewise version, *Stagewise Polytope Faces Pursuit*. In this algorithm at each iteration we select the $q > 1$ atoms with the largest adjusted correlations. From the polytope point of view, this corresponds to selecting the first q hyperplanes 'pierced'. We can also switch out several basis vectors as well. The algorithm steps can be summarized as follows:

1. Start at zero. ($\mathbf{c} = \mathbf{0}$, $\hat{\mathbf{x}} = \mathbf{0}$, $\mathbf{r} = \mathbf{x} - \hat{\mathbf{x}} = \mathbf{x}$)
2. Move in the direction of $\mathbf{x}$ until we 'pierce' q hyperplanes corresponding to duals of the basis vectors (this corresponds to selecting the largest q atoms of $\left\{\mathbf{a}_i^T \mathbf{r} / (1 - \mathbf{a}_i^T \mathbf{c})\right\}$).
3. Add a constraint (basis vector) corresponding to each of the pierced hyperplanes.
4. Now move $\mathbf{c}$ in direction of $\mathbf{x}$ projected onto the new constraint surface.
5. If movement would take us away from any faces (i.e. $\mathbf{s}_i < 0$), remove the corresponding constraints.
6. Repeat from step 2 until termination condition (e.g. residual or sparsity level).

We have applied this algorithm to compressed sensing problems taken from the Sparco toolkit. Fig. 1 shows the results of these as we vary the number of atoms added per iteration in Step 2 of the algorithm. With 1 atom per iteration we have the original PFP algorithm. We can see that, although the behaviour on the two examples is different as we add more atoms, typically we can achieve a factor of 3 speedup by using 7-10 atoms per iteration.

We have also compared our algorithm with other sparse recovery algorithms, and investigated some optimization refinements, such as the replacement of a pseudo-inverse computation, necessary to calculate the current position of the vertex $\mathbf{c}$, with a Cholesky n-column update (not shown).

Our Stagewise PFP algorithm has proved to be effective on compressed sensing tasks, since it can operate faster than the (single-atom) stepwise algorithm, while obtaining similar recovery errors. In further work we plan to compare the Cholesky-based pseudo-inverse with a conjugate gradient approach as used in the StOMP algorithm [4].

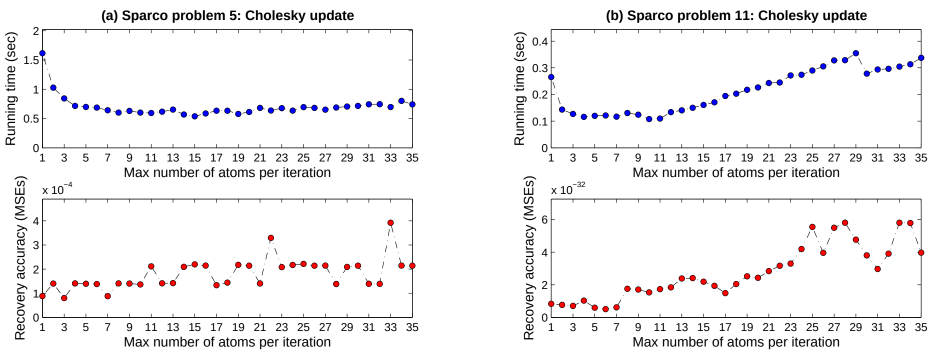

Figure 1: Running time and recovery accuracy against number of atoms added per iteration for (a) Sparco problem 5 (Cosine with Spikes) and (b) Sparco problem 11 (Gaussian amlitude spikes).

References

[1] T. Blumensath and M. E. Davies. Stagewise Weak Gradient Pursuits. Part I: Fundamentals and Numerical Studies. Submitted, 2008.

[2] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM J. on Sci. Comp.*, 20:33–61, 1998.

[3] D. L. Donoho. Neighborly Polytopes and Sparse Solutions of Underdetermined Linear Equations. Tech. Rep., Statistics Department, Stanford University, December 2004.

[4] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck. Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit. Tech. Rep. 2006-2, 2006.

[5] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. on Sig. Proc.*, 41:3397–3415, 1993.

[6] M. D. Plumbley. Recovery of Sparse Representations by Polytope Faces Pursuit. In *Proc. ICA 2006, Charleston, SC, USA*, pages 206–213, 2006.