Theorem 1 revisited: $N < k\binom{m}{k}^2$?

Charles Garfinkle

June 9, 2015

Abstract

In "When can dictionary learning uniquely recover sparse data from subsamples?" (HS2011) it was shown that there exists a set of $N = m$ 1-sparse vectors (the canonical basis vectors) with a unique 1-sparse coding. It was then proven in general that when $k < m$ there exist collections of $N = k\binom{m}{k}^2$ k-sparse vectors with unique k-sparse codings. This formula only reduces to $N = m^2 > m$ when $k = 1$, however. Can we find a better general formula for N that reduces to $N = m$ when $k = 1$? (e.g. $N = k\binom{m}{k}$ or perhaps even $N = km$?)

1 Proof by induction

Sets constructed as follows enable a proof by induction for the uniqueness of their sparse codes. Let $k > 0$ and suppose we are given a set of k-sparse vectors $\mathcal{A}_k \subset \Re^m$ for which any matrix $A \in \Re^{n\times m}$ satisfying the k-sparse spark condition:

$$Aa_1 = Aa_2 \implies a_1 = a_2 \text{ for all } k\text{-sparse } a_1, a_2$$

generates a set of vectors $\mathcal{Y}_k = A\mathcal{A}_k = \{Aa : a \in \mathcal{A}_k\}$ with a unique k-sparse coding (i.e. if $\forall y \in \mathcal{Y}_k, y = Bb$ for some $B \in \Re^{n\times m}$ and k-sparse b, then $A = BPD$ for some permutation matrix P and invertible diagonal matrix D) . Suppose $k' > k$ and we have some procedure for constructing a set $\mathcal{A}_{k'} \supseteq \mathcal{A}_k$ of k'-sparse vectors with the following property: for any $A \in \Re^{n\times m}$ satisfying the k'-sparse spark condition, any k'-sparse coding for $\mathcal{Y}_{k'} = A\mathcal{A}_{k'} \supseteq \mathcal{Y}_k$ is also a k-sparse coding for $\mathcal{Y}_k$. Then $\mathcal{Y}_{k'}$ must have a unique k'-sparse coding. (Otherwise we are in contradiction with the uniqueness of the k-sparse code for $\mathcal{Y}_k$ generated by any such A, since if A satisfies the k'-sparse condition, then it also satisfies the k-sparse condition for any $k < k'$.)

If such a procedure can be found in general for $k' = k + 1$, then we can construct a sequence of sets $(\mathcal{A}_k)_{k=1,\ldots,m}$ in this way, starting with some $\mathcal{A}_1$ for which $\mathcal{Y}_1 = A\mathcal{A}_1$ has a unique 1-sparse coding for all measurement matrices A satisfying the 1-sparse spark condition. Then Theorem 1 holds with the general formula $N(k) = |\mathcal{A}_k|$. (The exact form of $N(k)$ will depend on the procedure, e.g. if $|\mathcal{A}_1| = m$ and $|\mathcal{A}_k|$ increases by m for every increment in k, then $N(k) = km$.)

It has already been shown in HS2011 that $\mathcal{A}_1 = \{e_i\}_{i=1,\ldots,m}$ is such that $A\mathcal{A}_1$ has a unique 1-sparse coding for any A satisfying the 1-sparse spark condition. We need now to define a procedure for which a unique k-sparse coding for $(\mathcal{A}_k)$ implies a unique $(k+1)$-sparse coding for $\mathcal{A}_{k+1}$). As described previously, this implication reduces to showing that any $(k+1)$-sparse coding for $\mathcal{A}_{k+1}$ must actually also be a k-sparse coding for $\mathcal{A}_k \subset \mathcal{A}_{k+1}$ (since it is therefore unique by the induction hypothesis).

A guess like this might work...

$$
\begin{aligned}
\mathcal{A}_1 &= \{e_i\}_{i=1,\ldots,m} \\
\mathcal{A}_2 &= \mathcal{A}_1 \cup \{e_i + e_{(i+1)\text{mod}m}\}_{i=1,\ldots,m} \\
&\ldots \\
\mathcal{A}_{m-1} &= \mathcal{A}_{m-2} \cup \{e_i + e_{(i+1)\text{mod}m} + \ldots + e_{(i+m-2)\text{mod}m}\}_{i=1,\ldots,m}
\end{aligned}
$$

Otherwise, we could try to construct one on the fly...something like the following, described for the k=2 case. Start with $\mathcal{A}_1 = \{e_i\}_{i=1,\ldots,m}$ and suppose we have generated measurements $\mathcal{Y}_1 = A\mathcal{A}_1$ with some matrix $A \in \Re^{n\times m}$ satisfying the spark condition for 2-sparse vectors. We want $\mathcal{A}_2$ to contain vectors that force all the vectors in $\mathcal{Y}_1 \subset \mathcal{Y}_2$ to have 1-sparse codes whenever $\mathcal{Y}_2 = A\mathcal{A}_2$ has a 2-sparse coding.

Of the set of possible dictionaries $B \in \Re^{n\times m}$ that yield 2-sparse codes for all the vectors in $\mathcal{Y}_1$, only a subset of these 2-sparse codes are in fact 1-sparse codes. These are the dictionaries we want to force to be the only possibilities through our construction of $\mathcal{Y}_2$. The set of undesirable dictionaries can perhaps be partitioned in some intelligent way (e.g. if we have m alternate code vectors $b \in \Re^m$ for the vectors in $\mathcal{A}_1$, at least one of which is 2-sparse, then if these code vectors span $\Re^m$ there must be at least two of them that have overlapping support; we can partition by the m possibilities for this support index). For each partition we can construct a 2-sparse vector to add to $\mathcal{A}_2$ which somehow disqualifies all dictionaries in that partition as possible alternate sparse codings of $\mathcal{Y}_2$ (e.g. creates a contradiction). Once all partitions are eliminated we are left only with those that yield 1-sparse codes for $\mathcal{A}_1$ and 2-sparse codes for $\mathcal{A}_2$, which is what we want in order to prove uniqueness by induction.