EECS 16A Designing Information Devices and Systems I
Spring 2019 Discussion 13B

In Fall 2018, this discussion was not held haha

1. **Orthogonal Matching Pursuit Lecture**

   Orthogonal Matching Pursuit (OMP) algorithm:

   Inputs:

   - A set of m songs, each of length n: $\mathbf{S} = \{\vec{S}_0, \vec{S}_1, ..., \vec{S}_{m-1}\}$
   - An n-dimensional received signal vector: $\vec{r}$
   - The sparsity level k of the signal
   - Some threshold, th. When the norm of the signal is below this value, the signal contains only noise.

   Outputs:

   - A set of songs that were identified, F, which will contain at most k elements
   - A vector $\vec{x}$ containing song messages $(x_1, x_2, \ldots)$, which will be of length k or less
   - An n-dimensional residual $\vec{y}$

   Procedure:

   - Initialize the following values: $\vec{y} = \vec{r}$, $j = 1$, k, $\mathbf{A} = \begin{bmatrix} \quad \end{bmatrix}$, $F = \{\emptyset\}$
   - while $((j \leq k)$ and $(\| \vec{y} \| \geq th))$:
     (a) Cross-correlate $\vec{y}$ with the shifted versions of all songs. Find the song index i and the shifted version of the song, $\vec{S}_i^N$, with which the received signal has the highest correlation value.
     (b) Add i to the set of song indices F.
     (c) Column concatenate matrix $\mathbf{A}$ with the correctly shifted version of the song: $\mathbf{A} = \begin{bmatrix} \mathbf{A} & | & \vec{S}_i^N \end{bmatrix}$
     (d) Use least squares to obtain the message value: $\vec{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{r}$
     (e) Update the residual value $\vec{y}$ by subtracting: $\vec{y} = \vec{r} - \mathbf{A}\vec{x}$
     (f) Update the counter: $j = j + 1$

2. **Orthogonal Matching Pursuit**

   Let's work through an example of the OMP algorithm. Suppose that we have a vector $\vec{x} \in \mathbb{R}^4$ that is sparse and we know that it has only 2 non-zero entries. In particular,

$$\mathbf{M}\vec{x} \approx \vec{y} \tag{1}$$

$$\begin{bmatrix} | & | & | & | \\ \vec{m}_1 & \vec{m}_2 & \vec{m}_3 & \vec{m}_4 \\ | & | & | & | \end{bmatrix} \vec{x} \approx \vec{y} \tag{2}$$

$$\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \approx \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix} \tag{3}$$

where exactly 2 of x_1 to x_4 are non-zero. Use Orthogonal Matching Pursuit to estimate x_1 to x_4.

(a) Why can we not solve for $\vec{x}$ directly?

Answer:

We cannot solve for $\vec{x}$ directly because we have three measurements (or equations) but four unknowns. Since our system is underdetermined, we cannot solve for the unique $\vec{x}$ directly.

(b) Why can we not apply the least squares process to obtain $\vec{x}$?

Answer:

Recall the least squares solution: $\vec{\hat{x}} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\vec{y}$. $\mathbf{M}^T\mathbf{M}$ is only invertible if it has a trivial null space, i.e., if $\mathbf{M}$ has a trivial null space. However, in this case, $\mathbf{M}$ is a 3×4 matrix, so there is at least one free variable, which means that its null space is non-trivial. Therefore, $\mathbf{M}^T\mathbf{M}$ is not invertible, and we cannot use least squares to solve for $\vec{x}$.

(c) Let us start by reviewing the OMP procedure,

Inputs:

- A matrix $\mathbf{M}$, whose columns, $\vec{m}_i$, make up a set of vectors, $\{\vec{m}_i\}$, each of length n
- A vector $\vec{y}$ of length n
- The sparsity level k of the signal

Outputs:

- A vector $\vec{x}$, that contains k non-zero entries.
- A error vector $\vec{e} = \vec{y} - \mathbf{M}\vec{x}$

Procedure:

- Initialize the following values: $\vec{e} = \vec{y}$, $j = 1$, k, $\mathbf{A} = \begin{bmatrix} \ \end{bmatrix}$
- while ($j \leq k$):
  i. Compute the inner product for each vector in the set, $\vec{m}_i$, with $\vec{e}$: $c_i = \langle \vec{m}_i, \vec{e} \rangle$.
  ii. Column concatenate matrix $\mathbf{A}$ with the column vector that had the maximum inner product value with $\vec{e}$, c_i: $\mathbf{A} = \begin{bmatrix} \mathbf{A} & | & \vec{m}_i \end{bmatrix}$
  iii. Use least squares to compute $\vec{x}$ given the $\mathbf{A}$ for this iteration: $\vec{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{y}$
  iv. Update the error vector: $\vec{e} = \vec{y} - \mathbf{A}\vec{x}$
  v. Update the counter: $j = j + 1$

(d) Compute the inner product of every column with the $\vec{y}$ vector. Which column has the largest inner product? This will be the first column of the matrix $\mathbf{A}$.

Answer:

$$\left\langle \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}5\\1\\1\end{bmatrix} \right\rangle = 5$$

$$\left\langle \begin{bmatrix}0\\1\\2\end{bmatrix}, \begin{bmatrix}5\\1\\1\end{bmatrix} \right\rangle = 3$$

$$\left\langle \begin{bmatrix}2\\2\\0\end{bmatrix}, \begin{bmatrix}5\\1\\1\end{bmatrix} \right\rangle = 12$$

$$\left\langle \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}5\\1\\1\end{bmatrix} \right\rangle = 6$$

The third column has the largest inner product with $\begin{bmatrix}5\\1\\1\end{bmatrix}$, so $\mathbf{A} = \begin{bmatrix}2\\2\\0\end{bmatrix}$.

(e) Now, find the projection of $\vec{y}$ onto the columns of $\mathbf{A}$ (ie. $\text{proj}_{\text{Col}(\mathbf{A})}\vec{y} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{y}$). Use this to update the error vector.

Answer:

$$\vec{\hat{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{y} = \left(\begin{bmatrix}2 & 2 & 0\end{bmatrix} \begin{bmatrix}2\\2\\0\end{bmatrix} \right)^{-1} \begin{bmatrix}2 & 2 & 0\end{bmatrix} \begin{bmatrix}5\\1\\1\end{bmatrix} = \frac{1}{8} \cdot 12 = \frac{3}{2}$$

$$\text{proj}_{\text{Col}(\mathbf{A})}\vec{y} = \mathbf{A}\vec{\hat{x}} = \begin{bmatrix}2\\2\\0\end{bmatrix} \cdot \frac{3}{2} = \begin{bmatrix}3\\3\\0\end{bmatrix}$$

$$\vec{e} = \vec{y} - \text{proj}_{\text{Col}(\mathbf{A})}\vec{y} = \begin{bmatrix}5\\1\\1\end{bmatrix} - \begin{bmatrix}3\\3\\0\end{bmatrix} = \begin{bmatrix}2\\-2\\1\end{bmatrix}$$

(f) Now compute the inner product of every column with the new error vector. Which column has the largest inner product? This will be the second column of $\mathbf{A}$.

Answer:

$$\left\langle \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\-2\\1\end{bmatrix} \right\rangle = 2$$

$$\left\langle \begin{bmatrix}0\\1\\2\end{bmatrix}, \begin{bmatrix}2\\-2\\1\end{bmatrix} \right\rangle = 0$$

$$\left\langle \begin{bmatrix}2\\2\\0\end{bmatrix}, \begin{bmatrix}2\\-2\\1\end{bmatrix} \right\rangle = 0$$

$$\left\langle \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}2\\-2\\1\end{bmatrix} \right\rangle = 3$$

The fourth column has the largest inner product with $\begin{bmatrix}2\\-2\\1\end{bmatrix}$, so $\mathbf{A} = \begin{bmatrix}2 & 1\\2 & 0\\0 & 1\end{bmatrix}$.

(g) We now have two non-zero entries for our vector, $\vec{x}$. Find the values of those two entries. (Reminder: $\begin{bmatrix}a & b\\c & d\end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix}d & -b\\-c & a\end{bmatrix}$)

Answer:

$$\hat{\vec{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{y} = \left(\begin{bmatrix}2 & 2 & 0\\1 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 1\\2 & 0\\0 & 1\end{bmatrix}\right)^{-1}\begin{bmatrix}2 & 2 & 0\\1 & 0 & 1\end{bmatrix}\begin{bmatrix}5\\1\\1\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}$$

Therefore, $x_3 = 1$ and $x_4 = 2$, so $\vec{x} = \begin{bmatrix}0\\0\\1\\2\end{bmatrix}$.

