Performance Guarantee of OMP

Goal: To analyze the performance of OMP, and determine conditions under which OMP will return the sparsest solution.

Setting: Assume that the system \( Ax = b \) has a sparse solution \( x \) with \( k_0 \) non-zeros, i.e., \( \| x \|_0 = k_0 \).

Assume that \( k_0 < \text{spark}(A)/2 \).

Consider a two-ortho system:

\[
A = [\Psi, \Phi],
\]

so that \( b \) is created by the first \( k_P \) columns of \( \Psi \), and the first \( k_\Phi \) columns of \( \Phi \), such that \( k_P + k_\Phi = k_0 \):

\[
b = \sum_{i=1}^{k_P} x_i \psi_i + \sum_{i=1}^{k_\Phi} x_i \phi_i
\]

Let \( S_P \) and \( S_\Phi \) be the sets of support indices.
Theorem (4.1)
For a system \( Ax = b \) with \( A = (\Phi, \Phi) \), where \( \Phi \) and \( \Phi \) are two orthogonal matrices of size \( n \)-by-\( n \), if a solution \( x \) exists such that it has \( k_p \) non-zeros in its first half, and \( k_q \) non-zeros in the second, and the two obey
\[
\max(k_p, k_q) < \frac{1}{2\mu(A)},
\]
then OMP run with \( \epsilon_0 = 0 \) is guaranteed to find \( x \) exactly in \( k_p + k_q \) steps.

Proof:
Let's start with iteration 0. Assume that \( \|a_j\|_2 = 1 \)
\[
\begin{align*}
\begin{cases}
    r^0 = b \\
    c(j) = \min_{j \neq i} \|a_j z_j - b\|^2 = \|b\|^2 - (a_j^T b)^2 \geq 0
\end{cases}
\end{align*}
\]
WLOG, assume that the largest entry of \( x \) is \( x_{i'} \). Therefore, if we want to choose an entry from the correct support, we would require that
(i) \( |Y_i^T b| > |Y_j^T b| \), \( \forall j \notin S_p \)
(ii) \( |Y_i^T b| > |\Phi_j^T b| \), \( \forall j \notin \Phi_q \)
Claim: if $k_p < \frac{1}{2\mu(A)}$ and $k_q < \frac{1}{2\mu(A)}$, then (i) holds.

\textbf{Pf.} \quad |\psi_1^T b| = \left| \sum_{i=1}^{k_p} x_i \psi_i^T \psi_i + \sum_{i=1}^{k_q} x_i \phi_i^T \phi_i \right|

= \left| x_i^\psi + \sum_{i=1}^{k_q} x_i \phi_i^T \phi_i \right| \quad \text{ (orthogonality of } \Phi) \n
\geq |x_i^\psi| - \left| \sum_{i=1}^{k_q} x_i \phi_i^T \phi_i \right| \quad \text{ (} |a + b| \geq |a| - |b| \text{)} \n
\geq |x_i^\psi| - \sum_{i=1}^{k_q} |x_i^\phi| |\psi_i^T \phi_i| \quad \text{ (} |\sum a_i| \leq \Sigma |a_i| \text{)} \n
\geq |x_i^\psi| - \mu(A) \sum_{i=1}^{k_q} |x_i^\phi| \n
\geq |x_i^\psi| - \mu(A) k_q |x_i^\psi| \quad \text{ (} x_i^\psi \text{ is the largest)} \n
= |x_i^\psi| (1 - k_q \mu(A)).

|\psi_j^T b| = \left| \sum_{i=1}^{k_p} x_i \psi_j^T \psi_i + \sum_{i=1}^{k_q} x_i \phi_j^T \phi_i \right|

= \left| \sum_{i=1}^{k_q} x_i \phi_j^T \phi_i \right| \quad \text{ (} j \notin S_p, \text{ so } \psi_j^T \psi_i = 0 \text{)} \n
\leq \sum_{i=1}^{k_q} |x_i^\phi| |\psi_j^T \phi_i| \n
\leq k_q |x_i^\phi| |\mu(A)|.

So \quad 1 - k_q \mu(A) > k_q \mu(A) \iff k_q < \frac{1}{2\mu(A)}.

So if \quad k_q < \frac{1}{2\mu(A)}, \text{ then (i) holds.}
if the largest entry is $x_i^\phi$, then by using parallel argument we have

$$\text{if } k_p < \frac{1}{2\mu(A)}, \text{ then (i) holds.}$$

This concludes the claim.

Claim: if $k_p + k_{\phi} < \frac{1}{\mu(A)}$, then (ii) holds.

Pf: From the derivation above, we have

$$|\psi_i^T b| \geq |x_i^\psi| (1 - k_{\phi} \mu(A))$$

Let's compute $|\phi_j^T b|$

$$|\phi_j^T b| = \left| \sum_{i=1}^{k_p} x_i^\psi \phi_j^T \psi_i + \sum_{i=1}^{k_{\phi}} x_i^\phi \phi_j^T \phi_i \right|$$

$$= \left| \sum_{i=1}^{k_p} x_i^\psi \phi_j^T \psi_i \right|, \quad (j \notin S_\phi, \text{ so } \phi_j^T \phi_i = 0)$$

$$\leq \sum_{i=1}^{k_p} |x_i^\psi| \mu(A)$$

$$\leq |x_i^\psi| k_p \mu(A)$$

Therefore,

$$|\psi_i^T b| > |\phi_j^T b|$$

$$\Leftrightarrow |x_i^\psi| (1 - k_{\phi} \mu(A)) > |x_i^\psi| k_p \mu(A)$$

$$\Leftrightarrow k_p + k_{\phi} < \frac{1}{\mu(A)}$$

So if $k_p + k_{\phi} < \frac{1}{\mu(A)}$, then (ii) holds.
General Case

Theorem (4.3)

For a system $Ax = b$ ($A \in \mathbb{R}^{n \times m}$ full rank, $n < m$), if a solution $x$ exists obeying

$$\|x\|_0 < \frac{1}{2} \left(1 + \frac{1}{\mu(A)}\right),$$

then OMP with $\ell_0 = 0$ is guaranteed to find it exactly.

Proof: WLOG, assume that the sparsest solution has non-zeros in the first $k_0$ entries, and in the decreasing order of the values of $|x_j|$. Thus,

$$b = \sum_{i=1}^{k_0} x_i a_i.$$

In the first iteration, we have

$$r^0 = b$$

$$e(j) = \|b\|^2 - (a_j^T b)^2 \geq 0.$$

In order to choose one of the $k_0$ entries, we must require that

$$|a_i^T b| > |a_j^T b|, \quad \forall j > k_0.$$
\[ |a_i^T b| = \left| \sum_{i=1}^{k_0} x_i a_i^T a_i \right| \]
\[ = \left| x_1 \frac{1}{k_0} + \sum_{i=2}^{k_0} x_i a_i^T a_i \right|, \quad \text{assume } \|a_i\|_2 = 1 \]
\[ \geq \left| x_1 \right| - \left| \sum_{i=2}^{k_0} x_i a_i^T a_i \right|, \quad (|a+b| \geq |a|-|b|) \]
\[ \geq \left| x_1 \right| - \sum_{i=2}^{k_0} |x_i| |a_i^T a_i| \]
\[ \geq \left| x_1 \right| - \sum_{i=2}^{k_0} |x_i| \mu(A) \]
\[ \geq \left| x_1 \right| - \sum_{i=2}^{k_0} |x_1| \mu(A) \]
\[ = \left| x_1 \right| \left( 1 - \mu(A) (k_0 - 1) \right). \]

\[ |a_j^T b| = \left| \sum_{i=1}^{k_0} x_i a_j^T a_i \right| \]
\[ \leq \sum_{i=1}^{k_0} |x_i| |a_j^T a_i| \]
\[ \leq \sum_{i=1}^{k_0} |x_i| \mu(A) \]
\[ \leq \left| x_1 \right| \mu(A) k_0 \]

Therefore, \[ |a_i^T b| > |a_j^T b| \]
\[ \iff \quad 1 - (k_0 - 1) \mu(A) > k_0 \mu(A) \]
\[ \iff \quad k_0 < \frac{1}{2} \left( 1 + \frac{1}{\mu(A)} \right). \]