EECS 16A Designing Information Devices and Systems I
Fall 2019 Discussion 14B

1. 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, $\mathbf{A} = \begin{bmatrix} \quad \end{bmatrix}$
- 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 $\vec{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}$.

2. One Magical Procedure (Fall 2015 Final)

Suppose that we have a vector $\vec{x} \in \mathbb{R}^5$ and an $N \times 5$ measurement matrix $\mathbf{M}$ defined by column vectors $\vec{c}_1, \ldots, \vec{c}_5$, such that:

$$\mathbf{M}\vec{x} = \begin{bmatrix} | & & | \\ \vec{c}_1 & \cdots & \vec{c}_5 \\ | & & | \end{bmatrix} \vec{x} \approx \vec{b}$$

We can treat the vector $\vec{b} \in \mathbb{R}^N$ as a noisy measurement of the vector $\vec{x}$, with measurement matrix $\mathbf{M}$ and some additional noise in it as well.

You also know that the true $\vec{x}$ is sparse – it only has two non-zero entries and all the rest of the entries are zero in reality. Our goal is to recover this original $\vec{x}$ as best we can.

However, your intern has managed to lose not only the measurements $\vec{b}$ but the entire measurement matrix $\mathbf{M}$ as well!

Fortunately, you have found a backup in which you have all the pairwise inner products $\langle \vec{c}_i, \vec{c}_j \rangle$ between the columns of $\mathbf{M}$ and each other as well as all the inner products $\left\langle \vec{c}_i, \vec{b} \right\rangle$ between the columns of $\mathbf{M}$ and the vector $\vec{b}$. Finally, you also know the inner product of $\vec{b}$ with itself, *i.e.* $\left\langle \vec{b}, \vec{b} \right\rangle$.

All the information you have is captured in the following table of inner products. (These are not the vectors themselves.)

$\langle \cdot, \cdot \rangle$	$\vec{c}_1$	$\vec{c}_2$	$\vec{c}_3$	$\vec{c}_4$	$\vec{c}_5$	$\vec{b}$
$\vec{c}_1$	2	0	1	-1	1	1
$\vec{c}_2$		2	1	-1	-1	-5
$\vec{c}_3$			2	0	-1	2
$\vec{c}_4$				2	-1	6
$\vec{c}_5$					2	-1
$\vec{b}$						29

(So, for example, if you read this table, you will see that the inner product $\langle \vec{c}_2, \vec{c}_3 \rangle = 1$, that the inner product $\left\langle \vec{c}_3, \vec{b} \right\rangle = 2$, and that the inner product $\left\langle \vec{b}, \vec{b} \right\rangle = 29$. By symmetry of the real inner product, $\langle \vec{c}_3, \vec{c}_2 \rangle = 1$ as well.)

Your goal is to find which entries of $\vec{x}$ are non-zero and what their values are.

(a) Use the information in the table above to answer which of the $\vec{c}_1, \ldots, \vec{c}_5$ has the largest magnitude inner product with $\vec{b}$.

Answer:

Reading off the table, $\boxed{\vec{c}_4}$ has the largest inner product with $\vec{b}$.

(b) Let the vector with the largest magnitude inner product with $\vec{b}$ be $\vec{c}_a$. Let $\vec{b}_p$ be the projection of $\vec{b}$ onto $\vec{c}_a$. Write $\vec{b}_p$ symbolically as an expression only involving $\vec{c}_a$, $\vec{b}$, and their inner products with themselves and each other.

Answer:

The magnitude of the projection is $\frac{\langle \vec{c}_a, \vec{b} \rangle}{\|\vec{c}_a\|}$, and the direction of the projection is $\frac{\vec{c}_a}{\|\vec{c}_a\|}$. Thus:

$$\vec{b}_p = \boxed{\frac{\left\langle \vec{c}_a, \vec{b} \right\rangle}{\langle \vec{c}_a, \vec{c}_a \rangle} \vec{c}_a}$$

(c) Use the information in the table above to find which of the column vectors $\vec{c}_1, \ldots, \vec{c}_5$ has the largest magnitude inner product with the residue $\vec{b} - \vec{b}_p$.

Hint: The linearity of inner products might prove useful.

Answer:

The inner product of $\vec{b} - \vec{b}_p$ with a vector $\vec{c}_i$ is:

$$\left\langle \vec{b} - \vec{b}_p, \vec{c}_i \right\rangle = \left\langle \vec{b}, \vec{c}_i \right\rangle - \frac{\left\langle \vec{c}_a, \vec{b} \right\rangle}{\langle \vec{c}_a, \vec{c}_a \rangle} \langle \vec{c}_a, \vec{c}_i \rangle$$

Finding the numerical values of the inner products:

$$\begin{matrix} \left\langle \vec{b}-\vec{b}_p, \vec{c}_1 \right\rangle & \left\langle \vec{b}-\vec{b}_p, \vec{c}_2 \right\rangle & \left\langle \vec{b}-\vec{b}_p, \vec{c}_3 \right\rangle & \left\langle \vec{b}-\vec{b}_p, \vec{c}_4 \right\rangle & \left\langle \vec{b}-\vec{b}_p, \vec{c}_5 \right\rangle \\ 4 & -2 & 2 & 0 & 2 \end{matrix}$$

Thus the vector with the highest inner product with the residue is: $\boxed{\vec{c}_1}$.

(d) Suppose that the vectors we found in parts (a) and (c) are $\vec{c}_a$ and $\vec{c}_c$. These correspond to the components of $\vec{x}$ that are non-zero, that is, $\vec{b} \approx x_a\vec{c}_a + x_c\vec{c}_c$. However, there might be noise in the measurements $\vec{b}$, so we want to find the least squares estimates $\widehat{x}_a$ and $\widehat{x}_c$. Write a matrix expression for $\begin{bmatrix} \widehat{x}_a \\ \widehat{x}_c \end{bmatrix}$ in terms of appropriate matrices filled with the inner products of $\vec{c}_a$, $\vec{c}_c$, $\vec{b}$.

Answer:

We use least squares to solve for $\begin{bmatrix} \widehat{x}_a \\ \widehat{x}_c \end{bmatrix}$. Let $\mathbf{A} = \begin{bmatrix} \vec{c}_a & \vec{c}_c \end{bmatrix}$. Using the least-squares formula,

$$\begin{aligned} \begin{bmatrix} \widehat{x}_a \\ \widehat{x}_c \end{bmatrix} &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{b} \\ &= \begin{bmatrix} \langle \vec{c}_a, \vec{c}_a \rangle & \langle \vec{c}_a, \vec{c}_c \rangle \\ \langle \vec{c}_c, \vec{c}_a \rangle & \langle \vec{c}_c, \vec{c}_c \rangle \end{bmatrix}^{-1} \begin{bmatrix} \left\langle \vec{c}_a, \vec{b} \right\rangle \\ \left\langle \vec{c}_c, \vec{b} \right\rangle \end{bmatrix} \end{aligned}$$

(e) Compute the numerical values of $\widehat{x}_a$ and $\widehat{x}_c$ using the information in the table.

Answer:

Substituting the previous expression with values from the table, we get: $x_1 = 2\frac{2}{3}, x_4 = 4\frac{1}{3}$.

$$\begin{aligned} \begin{bmatrix} \widehat{x}_4 \\ \widehat{x}_1 \end{bmatrix} &= \begin{bmatrix} \langle \vec{c}_4, \vec{c}_4 \rangle & \langle \vec{c}_4, \vec{c}_1 \rangle \\ \langle \vec{c}_1, \vec{c}_4 \rangle & \langle \vec{c}_1, \vec{c}_1 \rangle \end{bmatrix}^{-1} \begin{bmatrix} \left\langle \vec{c}_4, \vec{b} \right\rangle \\ \left\langle \vec{c}_1, \vec{b} \right\rangle \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{13}{3} \\ \frac{8}{3} \end{bmatrix} \end{aligned}$$

