
EECS 16A
Spring 2020

Designing Information Devices and Systems I

Discussion 13B

1. Linearizing Different Problems

Notice that least squares can only be applied to linear systems. Suppose that we have a vector \vec{x} and a vector \vec{y} , and $\vec{y}[n] = f(\vec{x}[n])$. We would like to approximate f using least squares, where f is not necessarily a linear function.

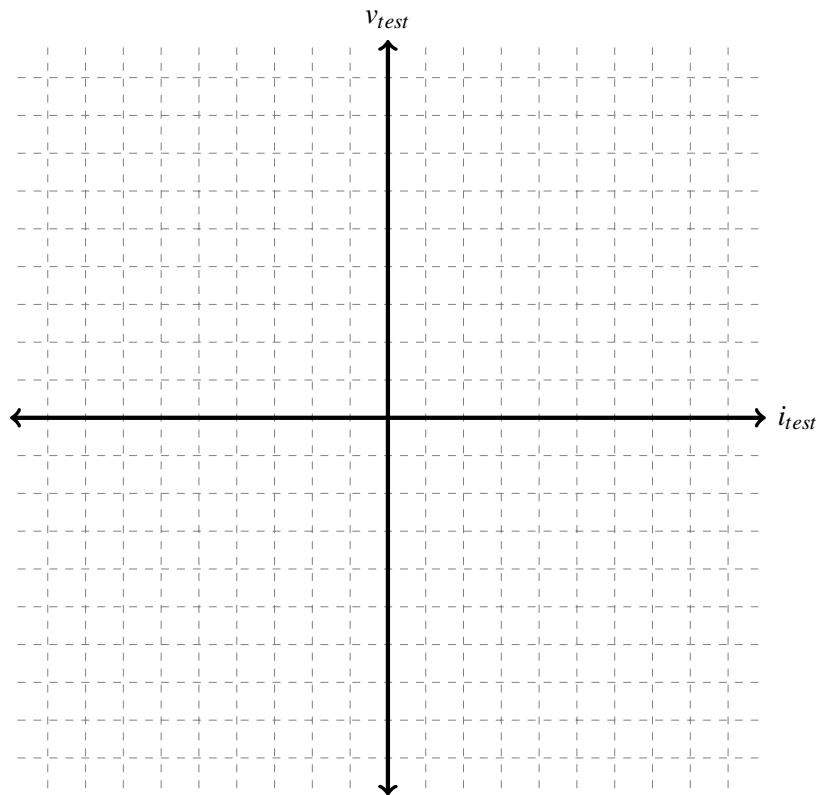
- (a) Let's begin with a linear approximation. We want to find some a such that $y = ax$. Set this up as a least squares problem. What are the elements in the matrix \mathbf{A} ?
- (b) Let's add a constant to the problem. Suppose that $y = ax + b$. Set this up as a least squares problem. What are the elements in the matrix \mathbf{A} ?
- (c) Suppose that $y = ax^2 + bx + c$. Set this up as a least squares problem. What are the elements in the matrix \mathbf{A} ?
- (d) Suppose that $y = ae^{bx}$. Set this up as a least squares problem. What are the elements in the matrix \mathbf{A} ?

2. Ohm's Law With Noise

We are trying to measure the resistance of a black box. We apply various i_{test} currents and measure the output voltage v_{test} . Sometimes, we are quite fortunate to get nice numbers. Oftentimes, our measurement tools are a little bit noisy, and the values we get out of them are not accurate. However, if the noise is completely random, then the effect of it can be averaged out over many samples. So we repeat our test many times:

| Test | i_{test} (mA) | v_{test} (V) |
|------|------------------------|-----------------------|
| 1 | 10 | 21 |
| 2 | 3 | 7 |
| 3 | -1 | -2 |
| 4 | 5 | 8 |
| 5 | -8 | -15 |
| 6 | -5 | -11 |

- (a) Plot the measured voltage as a function of the current.



- (b) Suppose we stack the currents and voltages to get $\vec{I} = \begin{bmatrix} 10 \\ 3 \\ -1 \\ 5 \\ -8 \\ -5 \end{bmatrix}$ and $\vec{V} = \begin{bmatrix} 21 \\ 7 \\ -2 \\ 8 \\ -15 \\ -11 \end{bmatrix}$. Is there a unique

solution for R ? What conditions must \vec{I} and \vec{V} satisfy in order for us to solve for R uniquely?

- (c) Ideally, we would like to find R such that $\vec{V} = \vec{I}R$. If we cannot do this, we'd like to find a value of R that is the *best* solution possible, in the sense that $\vec{I}R$ is as "close" to \vec{V} as possible. We are defining the sum of squared errors as a **cost function**. In this case the cost function for any value of R quantifies the difference between each component of \vec{V} (i.e. v_j) and each component of $\vec{I}R$ (i.e. i_jR) and sum up the squares of these "differences" as follows:

$$\text{cost}(R) = \sum_{j=1}^6 (v_j - i_jR)^2$$

Do you think this is a good cost function? Why or why not?

- (d) Show that you can also express the above cost function in vector form, that is,

$$\text{cost}(R) = \langle (\vec{V} - \vec{I}R), (\vec{V} - \vec{I}R) \rangle$$

Hint: $\langle \vec{a}, \vec{b} \rangle = \vec{a}^T \vec{b} = \sum_i a_i b_i$

- (e) Find \hat{R} , which is defined as the optimal value of R that minimizes $\text{cost}(R)$.

Hint: Use calculus. The optimal \hat{R} makes $\frac{d\text{cost}(\hat{R})}{dR} = 0$

- (f) On your original IV plot, also plot the line $v_{test} = \hat{R}i_{test}$. Can you visually see why this line “fits” the data well? How well would we have done if we had guessed $R = 3\text{ k}\Omega$? What about $R = 1\text{ k}\Omega$? Calculate the cost functions for each of these choices of R to validate your answer.
- (g) Now, suppose that we add a new data point: $i_7 = 2\text{ mA}$, $v_7 = 4\text{ V}$. Will \hat{R} increase, decrease, or remain the same? Why? What does that say about the line $v_{test} = \hat{R}i_{test}$?

3. 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,

$$\mathbf{M}\vec{x} \approx \vec{y} \quad (1)$$

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

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

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?
- (b) Why can we not apply the least squares process to obtain \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 , $\mathbf{A} = []$
 - while ($j \leq k$):
 - Compute the inner product for each vector in the set, \vec{m}_i , with \vec{e} : $c_i = \langle \vec{m}_i, \vec{e} \rangle$.
 - Column concatenate matrix \mathbf{A} with the column vector that had the maximum inner product value with \vec{e} , c_i : $\mathbf{A} = [\mathbf{A} \mid \vec{m}_i]$
 - Use least squares to compute \vec{x} given the \mathbf{A} for this iteration: $\vec{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \vec{y}$
 - Update the error vector: $\vec{e} = \vec{y} - \mathbf{A}\vec{x}$
 - 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} .

- (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.
- (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} .
- (g) We now have two non-zero entries for our vector, \vec{x} . Find the values of those two entries.

(Reminder: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$)