EECS 16A Designing Information Devices and Systems I
Spring 2021 Discussion 13A

1. Polynomial Fitting

Let's try an example. Say we know that the output, y, is a quartic polynomial in x. This means that we know that y and x are related as follows:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$$

We're also given the following observations:

x	y
0.0	24.0
0.5	6.61
1.0	0.0
1.5	-0.95
2.0	0.07
2.5	0.73
3.0	-0.12
3.5	-0.83
4.0	-0.04
4.5	6.42

(a) What are the unknowns in this question? What are we trying to solve for?

Answer:

The unknowns are a_0, a_1, a_2, a_3, and a_4. They are also what we are trying to solve for.

(b) Can you write an equation corresponding to the first observation (x_0, y_0), in terms of a_0, a_1, a_2, a_3, and a_4? What does this equation look like? Is it linear in the unknowns?

Answer:

Plugging (x_0, y_0) into the expression for y in terms of x, we get

$$24 = a_0 + a_1 \cdot 0 + a_2 \cdot 0^2 + a_3 \cdot 0^3 + a_4 \cdot 0^4$$

You can see that this equation is linear in a_0, a_1, a_2, a_3, and a_4.

(c) Now, write a system of equations in terms of a_0, a_1, a_2, a_3, and a_4 using *all of the observations*.

Answer:

Write the next equation using the second observation. You will now get:

$$6.61 = a_0 + a_1 \cdot (0.5) + a_2 \cdot (0.5)^2 + a_3 \cdot (0.5)^3 + a_4 \cdot (0.5)^4$$

And for the third:

$$0.0 = a_0 + a_1 \cdot (1) + a_2 \cdot 1^2 + a_3 \cdot 1^3 + a_4 \cdot 1^4$$

Do you see a pattern? Let's write the entire system of equations in terms of a matrix now.

$$\begin{bmatrix} 1 & 0 & 0^2 & 0^3 & 0^4 \\ 1 & 0.5 & (0.5)^2 & (0.5)^3 & (0.5)^4 \\ 1 & 1 & 1^2 & 1^3 & 1^4 \\ 1 & 1.5 & (1.5)^2 & (1.5)^3 & (1.5)^4 \\ 1 & 2 & 2^2 & 2^3 & 2^4 \\ 1 & 2.5 & (2.5)^2 & (2.5)^3 & (2.5)^4 \\ 1 & 3 & 3^2 & 3^3 & 3^4 \\ 1 & 3.5 & (3.5)^2 & (3.5)^3 & (3.5)^4 \\ 1 & 4 & 4^2 & 4^3 & 4^4 \\ 1 & 4.5 & (4.5)^2 & (4.5)^3 & (4.5)^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 24 \\ 6.61 \\ 0.0 \\ -0.95 \\ 0.07 \\ 0.73 \\ -0.12 \\ -0.83 \\ -0.04 \\ 6.42 \end{bmatrix}$$

(d) Finally, solve for a_0, a_1, a_2, a_3, and a_4 using IPython. You have now found the quartic polynomial that best fits the data!

Answer:

Let $\mathbf{D}$ be the big matrix from the previous part.

$$\vec{a} = (\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T\vec{y} = \begin{bmatrix} 24.00958042 \\ -49.99515152 \\ 35.0039627 \\ -9.99561772 \\ 0.99841492 \end{bmatrix}$$

It turns out that the actual parameters for the polynomial equation were:

$$\vec{a} = \begin{bmatrix} 24 \\ -50 \\ 35 \\ -10 \\ 1 \end{bmatrix}$$

(Remember that our observations were noisy.)

Therefore, we have actually done pretty well with the least squares estimate!

2. Orthogonal Subspaces

Two vectors are $\vec{x}$ and $\vec{y}$ are said to be orthogonal if their inner product is zero. That is $\langle \vec{x}, \vec{y} \rangle = 0$.

Two subspaces $\mathbb{S}_1$ and $\mathbb{S}_2$ of $\mathbb{R}^N$ are said to be orthogonal if all vectors in $\mathbb{S}_1$ are orthogonal to all vectors in $\mathbb{S}_2$. That is,

$$\langle \vec{v}_1, \vec{v}_2 \rangle = 0 \; \forall \vec{v}_1 \in \mathbb{S}_1, \vec{v}_2 \in \mathbb{S}_2.$$

(a) Recall that the *column space* of an $M \times N$ matrix $\mathbf{A}$ is the subspace spanned by the columns of $\mathbf{A}$ and that the *null space* of $\mathbf{A}$ is the subspace of all vectors $\vec{v}$ such that $\mathbf{A}\vec{v} = \vec{0}$.

Prove that the column space of $\mathbf{A}^T$ and null space of any matrix $\mathbf{A}$ are orthogonal subspaces. This can be denoted by $\text{Col}(\mathbf{A}^T) \perp \text{Null}(\mathbf{A}) \; \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.

Hint: Use the row interpretation of matrix multiplication.

Answer:

First, we denote the rows of $\mathbf{A}$ as $\vec{a}_1^T, \vec{a}_2^T, \ldots, \vec{a}_M^T$. Now consider any vector $\vec{v} \in \text{Null}(\mathbf{A})$ which means that $\mathbf{A}\vec{v} = \vec{0}$. Note that matrix multiplication can be viewed as many inner products between the rows of $\mathbf{A}$ and the vector $\vec{v}$.

$$\mathbf{A}\vec{v} = \begin{bmatrix} \langle \vec{a}_1, \vec{v} \rangle \\ \langle \vec{a}_2, \vec{v} \rangle \\ \vdots \\ \langle \vec{a}_M, \vec{v} \rangle \end{bmatrix} = \vec{0}$$

Therefore, any vector $\vec{v} \in \text{Null}(\mathbf{A})$ is orthogonal to all rows of $\mathbf{A}$. From the linearity of the inner product, it follows that $\vec{v}$ is orthogonal to any linear combination of the rows of $\mathbf{A}$ and thus, any vector in $\text{Null}(\mathbf{A})$ is orthogonal to any vector in $\text{Col}(\mathbf{A}^T)$, proving that $\text{Col}(\mathbf{A}^T) \perp \text{Null}(\mathbf{A})\ \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.

(b) Now prove that the column space and null space of $\mathbf{A}^T$ of any matrix $\mathbf{A}$ are orthogonal subspaces. This can be denoted by $\text{Col}(\mathbf{A}) \perp \text{Null}(\mathbf{A}^T)\ \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.

Answer:

We can define a new matrix $\mathbf{B} \triangleq \mathbf{A}^T$ and denote its rows as $\vec{b}_1^T, \vec{b}_2^T, \ldots, \vec{b}_N^T$. Using the same steps as in part (a), we can conclude that $\text{Col}(\mathbf{B}^T) \perp \text{Null}(\mathbf{B})\ \forall \mathbf{B} \in \mathbb{R}^{N \times M}$. Changing $\mathbf{B}$ back to $\mathbf{A}^T$ yields $\text{Col}(\mathbf{A}) \perp \text{Null}(\mathbf{A}^T)\ \forall \mathbf{A} \in \mathbb{R}^{M \times N}$, which is what we wanted to prove.