Dis 13A EECS16A

tinyurl.com/miyuki-feedback

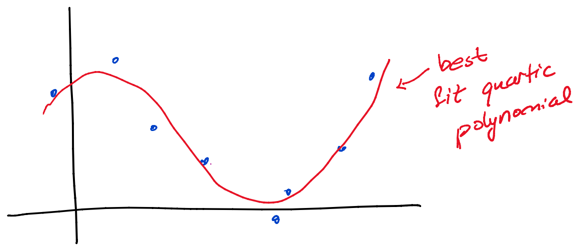
1. Polynomial Fitting

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:

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

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



* unique sol would give a curve that goes through each data point

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

$a_0, \ldots, a_4$ 5 unknowns

(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?

$(x_0, y_0) \rightarrow (0.0, 24.0)$

$$* \quad 24.0 = a_0 + a_1(0.0) + a_2(0.0)^2 + a_3(0.0)^3 + a_4(0.0)^4$$

(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.*

$$(x_1, y_1) \quad * \quad 6.61 = a_0 + a_1(0.5) + a_2(0.5)^2 + a_3(0.5)^3 + a_4(0.5)^4$$

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

$$\vdots$$

$$\underbrace{\begin{bmatrix} 24.0 \\ 6.61 \\ 0.0 \\ \vdots \\ 6.42 \end{bmatrix}}_{\vec{y}} = \underbrace{\begin{bmatrix} 1 & 0 & 0^2 & 0^3 & 0^4 \\ 1 & 0.5 & 0.5^2 & 0.5^3 & 0.5^4 \\ & & \vdots & & \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}}_{\vec{a}}$$

10×1 $\quad$ 10×5 $\quad$ 5×1

$\vec{y} = D\vec{a}$

↳ probably no soln

to find our best guess $\hat{\vec{a}}$

→ use least squares

$$\hat{\vec{a}} = (D^T D)^{-1} D^T \vec{y}$$

* if it was a unique sol

$\hat{\vec{a}} = \vec{a}$ using least squares formula

(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!

$$\hat{\vec{a}} = (D^T D)^{-1} D^T \vec{y} = \begin{bmatrix} 24.009 \\ -49.995 \\ 35.004 \\ -9.996 \\ 0.998 \end{bmatrix}$$

* this data was generated using $\vec{a} = \begin{bmatrix} 24 \\ -50 \\ 35 \\ -10 \\ 1 \end{bmatrix}$ but with noise!

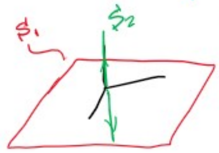
→ pretty good estimation!

$$\text{error} = \text{cost} = \|\vec{y} - D\hat{\vec{a}}\|^2$$

* if unique sol cost = 0

Ex: $S_1 = \text{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$ $S_2 = \text{span}\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$

$\vec{v}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$ $\vec{v}_2 = \begin{bmatrix}0\\0\\1\end{bmatrix}$



$\left\langle \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\rangle = 0$ $\left(\begin{bmatrix}-3\\0\\0\end{bmatrix} \cdot \begin{bmatrix}0\\0\\7\end{bmatrix} \right) = 0$

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

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,

$$\langle \vec{v}_1, \vec{v}_2 \rangle = 0 \; \underbrace{\forall}_{\text{for all}} \vec{v}_1 \overset{\text{in}}{\in} \mathbb{S}_1, \vec{v}_2 \in \mathbb{S}_2.$$

(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.

Def: orthogonality $\vec{x} \perp \vec{y}$ if $\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \vec{y} = 0$

$A^T = \begin{bmatrix} | & | & & | \\ \vec{\omega}_1 & \vec{\omega}_2 & \cdots & \vec{\omega}_n \\ | & | & & | \end{bmatrix}$ $C(A^T) = \text{span}\{\vec{\omega}_1, \vec{\omega}_2, \ldots, \vec{\omega}_n\}$

$N(A) =$ all $\vec{v}$ such that $A\vec{v} = \vec{0}$

$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} - & \vec{\omega}_1^T & - \\ - & \vec{\omega}_2^T & - \\ & \vdots & \\ - & \vec{\omega}_n^T & - \end{bmatrix}$

$\begin{bmatrix} 1 & 2 & 3 \\ & \vdots & \end{bmatrix} \begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix} 1\cdot 1 + 2\cdot 2 + 3\cdot 3 \\ \vdots \end{bmatrix} = \begin{bmatrix} [1\,2\,3]^T \begin{bmatrix}1\\2\\3\end{bmatrix} \\ \vdots \end{bmatrix}$

Strategies for Proofs

- write out mathematical defn for what you know & what you want to show
- try simple examples to find patterns
- manipulate the defns to get from what you know to what you want to show

$A\vec{v} = \vec{0}$

$\underset{M \times N}{\begin{bmatrix} - & \vec{\omega}_1^T & - \\ - & \vec{\omega}_2^T & - \\ & \vdots & \\ - & \vec{\omega}_n^T & - \end{bmatrix}} \underset{N \times 1}{\vec{v}} = \underset{M \times 1}{\begin{bmatrix} \vec{\omega}_1^T \vec{v} \\ \vec{\omega}_2^T \vec{v} \\ \vdots \\ \vec{\omega}_n^T \vec{v} \end{bmatrix}} = \begin{bmatrix} \langle \vec{\omega}_1, \vec{v} \rangle \\ \langle \vec{\omega}_2, \vec{v} \rangle \\ \vdots \\ \langle \vec{\omega}_n, \vec{v} \rangle \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

$\Rightarrow C(A^T) \perp N(A) \; \forall A \in \mathbb{R}^{M \times N}$

columnspace of A^T is orthogonal to nullspace of A for all $A \in \mathbb{R}^{M \times N}$

(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 $\text{Col}(\mathbf{A}) \perp \text{Null}(\mathbf{A}^T) \; \forall \mathbf{A} \in \mathbb{R}^{M \times N}$.

switched which one was transposed

rewrite as something we've already proved (i.e. part a)

define $B = A^T$ $B^T = A$

$C(A) = C(B^T)$ $N(A^T) = N(B)$

$C(A) \perp N(A^T) \Rightarrow \underbrace{C(B^T) \perp N(B)}_{\text{proved in part (a)}}$

then switch back to A