Data Fitting

As an application, we introduce the orthogonal projection to data fitting problem.

Data Fitting

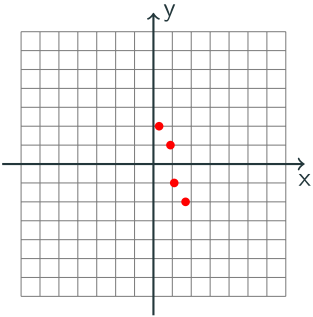
Excercise. Suppose one has the following data.

	X=0.3	X=0.9	X=1.1	X=1.7
Y	2	1	-1	-2

You want to find a function $D(x) = ax + b$ that fit these data in a way such that the error:

$$E(D) := ||D(0.3) - 2||^2 + ||D(0.9) - 1||^2 + ||D(1.1) - (-1)||^2 + ||D(1.7) - (-2)||^2$$

is minimal for your choice of function D.



Data Fitting

Model:

$$D(x) = ax + b$$

Regression Error:

$$E(D) := ||D(0.3)-2||^2+||D(0.9)-1||^2+||D(1.1)-(-1)||^2+||D(1.7)-(-2)||^2$$

Idea: We translate the problem into a geometric problem

- Data $\iff$ A point,
- Model $\iff$ point on a subspace;
- Error $\iff$ Distance.

Data as a point in a space

We may think the data

$X = 0.3$	$X = 0.9$	$X = 1.1$	$X = 1.7$
2	1	-1	-2

as a given vector in $\mathbb{R}^4$. This space is representing all possible observations.

$$\begin{pmatrix} 2 \\ 1 \\ -1 \\ -2 \end{pmatrix}$$

Model as subspace

The model is always selected as what the perfect data should be. In our model $D(x) = ax + b$. In terms of a and b, the perfect data can be listed as follows

$X = 0.3$	$X = 0.9$	$X = 1.1$	$X = 1.7$
$0.3a + b$	$0.9a + b$	$1.1a + b$	$1.7a + b$

For each a and b, it associates to a point, but the collection of all such points gives a subset

$$\left\{ \begin{pmatrix} 0.3a + b \\ 0.9a + b \\ 1.1a + b \\ 1.7a + b \end{pmatrix} : \text{ for } a, b \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} : \text{ for } \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \right\}$$

$$= \text{ Col } \begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix}$$

Model as subspace

With this understanding, the Error

$$E(D) := ||D(0.3) - 2||^2 + ||D(0.9) - 1||^2 + ||D(1.1) - (-1)||^2 + ||D(1.7) - (-2)||^2$$

is the distance of

$$\begin{pmatrix} 2 \\ 1 \\ -1 \\ -2 \end{pmatrix}$$

to a given point on the

$$\text{Col} \begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix}$$

Our first step is to formulate orthogonal projections to the perfect data space

Model as subspace

$$\begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix} \left(\begin{pmatrix} 0.3 & 0.9 & 1.1 & 1.7 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix} \right)^{-1} \begin{pmatrix} 0.3 & 0.9 & 1.1 & 1.7 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix} \begin{pmatrix} 5 & 4 \\ 4 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 0.3 & 0.9 & 1.1 & 1.7 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix}}_{\text{keep left factor}} \begin{pmatrix} -0.7 & -0.1 & 0.1 & 0.7 \\ 0.95 & 0.35 & 0.15 & -0.45 \end{pmatrix}$$

Model as subspace

Therefore, the best data is given by

$$\begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix} \begin{pmatrix} -0.7 & -0.1 & 0.1 & 0.7 \\ 0.95 & 0.35 & 0.15 & -0.45 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 3 \end{pmatrix}$$

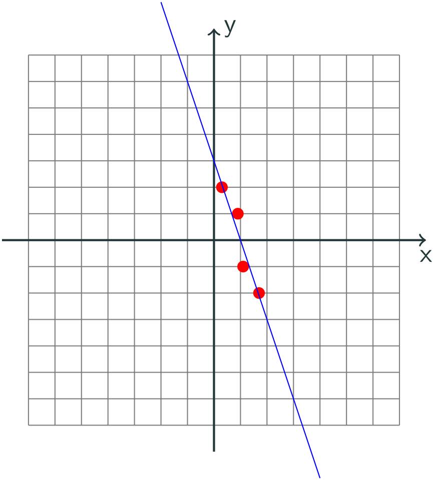
Note that this is of the form

$$\begin{pmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 1.1 & 1 \\ 1.7 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

for $a = 3$, $b = 3$. Therefore, the best regression is

$$D(x) = -3x + 3.$$

Model as subspace



The best fitting function for the printed data.

Model as subspace

Excercise. Find the best parabola fit $b(t) = C + Dt + Et^2$ to the data points

$$(t_1, b_1) = (-1, 5), \qquad (t_2 b_2) = (0, 2), \qquad (t_3 b_3) = (1, 1), \qquad (t_4 b_4) = (2, 2).$$

in the sense of minimizing

$$||E||^2 = (b_1 - b(t_1))^2 + (b_2 - b(t_2))^2 + (b_3 - b(t_3))^2 + (b_4 - b(t_4))^2$$

Model as subspace

Solution. Consider the point of obeserved data

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 1 \\ 2 \end{pmatrix} \in \mathbb{R}^4$$

The map $b(t) \longmapsto \begin{pmatrix} b(-1) \\ b(0) \\ b(1) \\ b(2) \end{pmatrix}$ maps perfect data into a subspace into $\mathbb{R}^4$.

The error function

$$||E||^2 = (b_1 - b(t_1))^2 + (b_2 - b(t_2))^2 + (b_3 - b(t_3))^2 + (b_4 - b(t_4))^2$$

describe the distance of an observed data to the subspace of perfect data. **Therefore, we may find perfect data by the orthogonal projection formula.**

Model as subspace

For this, we need to describe the subspace of perfect data, which is all possible

$$\vec{v}\begin{pmatrix} b(-1) \\ b(0) \\ b(1) \\ b(2) \end{pmatrix} = \begin{pmatrix} C - D + E \\ C \\ C + D + E \\ C + 2D + 4E \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} C \\ D \\ E \end{pmatrix}$$

for any $b(t) = C + Dt + Et^2$.

Our goal is to find the orthogonal projection P_W to the perfect data space

$$W = \mathrm{Col}\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

Model as subspace

Of course, you may find it by the orthogonal projection formula

$$P_W = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix}$$

And the perfect data closest to the observed data is given by

$$P_W \vec{v} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \underbrace{\left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ 1 \\ 2 \end{pmatrix}}_{\text{Calculate this part you will get } (C,D,E)^T}$$

We omit the rest steps here.

Model as subspace

However, there is a **Cool** method for this problem.

Model as subspace

First trick, from orthogonal complement.

$$W = \mathrm{Col}\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

is 3-dimensional. Hard. But $W^{\perp}$ is 1-dimensional! Good!

- Instead of finding P_W, find $P_{W^\perp}$ and so $P_W = I - P_{W^\perp}$

Model as subspace

Second trick, By lagurange interpolation polynomial

$$P_W(\vec{v}) = \begin{pmatrix} b_{\text{answer}}(-1) \\ b_{\text{answer}}(0) \\ b_{\text{answer}}(1) \\ b_{\text{answer}}(2) \end{pmatrix}$$

do not directly give us the C, D, E, but it we know the value of $b(t)$ at $t = -1, 0, 1, 2$. We may use **Lagrange Interpolation Polynomial**

$$b(t) = b(-1)\frac{x(x-1)}{(-1-0)(-1-1)} + b(0)\frac{(x+1)(x-1)}{(0+1)(0-1)} + b(1)\frac{(x+1)x}{(1+1)(1-0)}$$

Model as subspace

Solution. Let us do. Firstly , by solving the equation, we find

$$W^{\perp} = \mathrm{Null}\begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 4 \end{pmatrix} = \mathrm{Col}\begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}$$

Then

$$P_{W^{\perp}} = \frac{\begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}\begin{pmatrix} 1 & -3 & 3 & -1 \end{pmatrix}}{\begin{pmatrix} 1 & -3 & 3 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ -3 \\ 3 \\ -1 \end{pmatrix}} = \frac{1}{20}\begin{pmatrix} 1 & -3 & 3 & -1 \\ -3 & 9 & -9 & 3 \\ 3 & -9 & 9 & -3 \\ -1 & 3 & -3 & 1 \end{pmatrix}$$

Model as subspace

Therefore,

$$P_W = \frac{1}{20}\begin{pmatrix} 19 & 3 & -3 & 1 \\ 3 & 11 & 9 & -3 \\ -3 & 9 & 11 & 3 \\ 1 & -3 & 3 & 19 \end{pmatrix}$$

So

$$\begin{pmatrix} b_{\text{answer}}(-1) \\ b_{\text{answer}}(0) \\ b_{\text{answer}}(1) \\ b_{\text{answer}}(2) \end{pmatrix} = P_W \begin{pmatrix} 5 \\ 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

Just by looking at the data, we observe

$$b_{\text{answer}} = (x-1)^2 + 1$$