Lecture 6:

Assignment 6

Part 1:

* linear models

Part 2:

* bakery dataset

splitting dataset:

1. always (re)split your dataset
2. random state ---🡪 different splits
3. how to split?
4. 90 training/10% testing
5. 50/50
6. 10% training/90% testing

N-fold validation

LINEAR MODELS: means it is linear in coefficients

let W =(a,b,c) be the weights of a quadratic polynomial

f(x, W) = a\*x\*\*2 + bx + c

f(x, W1) = a1\*x\*\*2 + b1\*x + c1

f(x, W2) = a2\*x\*\*2 + b2\*x + c2

f(x,W1) + f(x, W2) = (a1+a2)\*x\*\*2 +

(b1+b2)\*x + (c1+c2) = f(x, W1+W2)

Easy to show that for any c1 and c2

f(x, c1W1 + c2W2) = c1\*f(x, W1) + x2\*f(x,W2)

Note on polynomials:

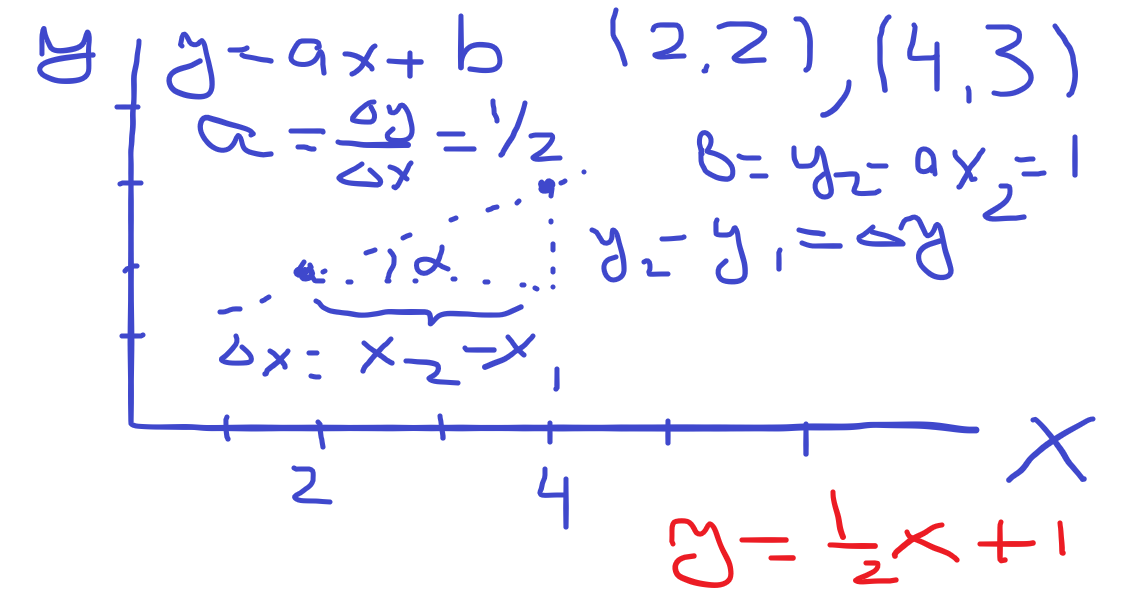
if we have (n+1) points

(x1,y1),(x2,y2), ….(xn,yn) then

there exists a polynomial P(x) of degree at most n and this polynomial is unique

(polynomial interpolation)

Example:



y = ax + b

---------------

2a + b = 2

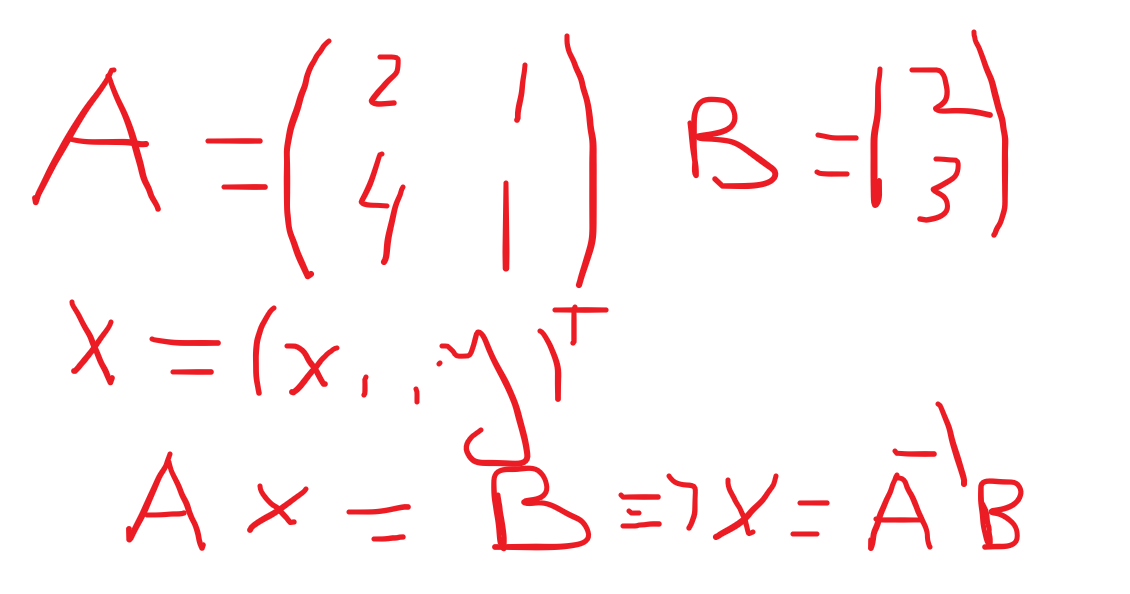
4a + b = 3

---------------

2a = 1 ---------🡪 a = ½

substitute into first one --🡪 b = 1

In matrix format



One more example: 3 points

P1 = (5, 0), P2 = (6,1), P3 = (7, 4)

y = ax\*\*2 + bx + c

25a + 5b + c = 0 P1

36a + 6b + c = 1 P2

49a + 7b + c = 4 P3

compute P2-P1 and P3 – P2

11a + b = 1 (equation 1.1)

13a + b = 3 (equation 1.2)

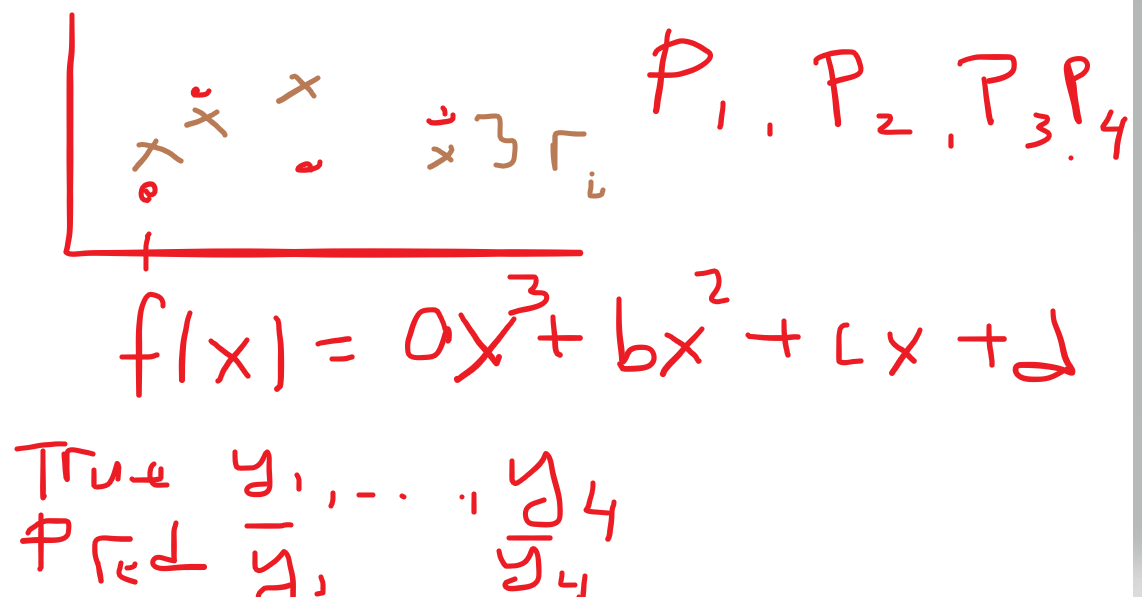
2a = 2 ----------------------🡪 a =1

compute b = -10

substitute a,b in equation 1 for point 1 and get c = 25

Final equation:

y = x\*\*2 – 10\*x + 25 =



Notes: if we know the degree of a polynomial to use, then we can compute the coefficients and do the prediction

if we do not know what degree to use but would like to avoid very high degrees, we can use “regularization”

what does this mean:

try to fit with degree “n” polynomial but you add a penalty term to the loss function that will add more loss to higher degrees