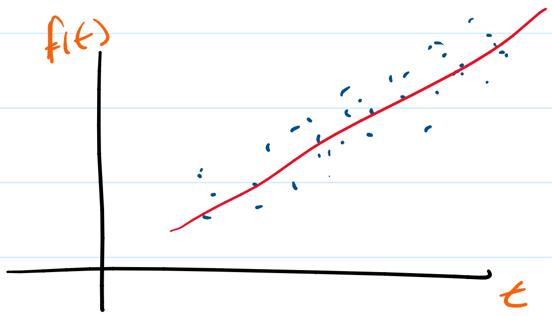
Friday, October 20, 2023 1:45 PM

What if I have data that has been collected and I need to fit a function to it?



Do we want to do interpolation?

No don't, at best it will be wild.

How do we mathematically approximate trends?

For simplicity let's assume we are building a linear fit. $p(x) = a_0 + a_1 x$

Goal: find a_0 & a_1 to make p the best fit for the data.

What do we mean by best fit?

We look at norms.

Given data $\{(x_i, y_i)\}_{i=0}^{n}$

1- Minimax

minimize $E_\infty(a_0, a_1) = \min \max_i \{|y_i - (a_0 + a_1 x_i)|\}$

2- l_1-minimization - minimize the l_1-norm

$$\min E_1(a_0, a_1) = \sum_{i=0}^{n} |y_i - (a_0 + a_1 x_i)|$$

3- least square or l_2-minization

$$\text{minimize } E_2(a_0, a_1) = \min\left(\sum_{i=0}^{n} |y_i - (a_0 + a_1 x_i)|^2 \right)^{1/2}$$

We Consider the least squares problem.
How do we find a_0 & a_1?

let's consider the function

$$g_2(a_0, a_1) = (E_2(a_0, a_1))^2$$

$$= \sum_{i=0}^{n} (y_i - (a_0 + a_1 x_i))^2$$

This is an upward facing parabola.
So we look for where the derivatives $=0$.

$$0 = \frac{dg}{da} = -2 \sum^{n} (y_i - a_0 - a_1 x_i)$$

$$0 = \frac{dg}{da_0} = -2 \sum_{i=0} (y_i - a_0 - a_1 x_i)$$

$$=$$