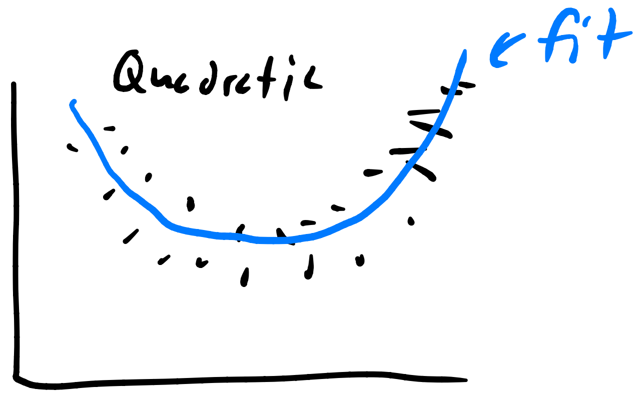
polynomial regression

- In project, you will explore fitting a polynomial regression model to one independent/dependent variable pair:



This will allow you to better fit nonlinear data — data that does not exhibit a linear relationship between independent and dependent variable.

- Implementing polynomial regression is simple.

Example: Fit model $\overbrace{c_0 + c_1 x_1 + c_2 x_1^2 + c_3 x_1^3}^{\text{Cubic eq.}}$

notice: only <u>one</u> independent var

- Take A with desired independent var x_1 (can be any one you want) and augment matrix with powers of x_1:

$$A = \begin{bmatrix} \vec{1} & \vec{x}_1 \end{bmatrix} \rightarrow \begin{bmatrix} \vec{1} & \vec{x}_1 & \vec{x}_1^2 & \vec{x}_1^3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & x_{1,1} & x_{1,1}^2 & x_{1,1}^3 \\ 1 & x_{1,2} & x_{1,2}^2 & x_{1,2}^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1,N} & x_{1,N}^2 & x_{1,N}^3 \end{bmatrix}$$

$\Rightarrow$ Solve with least squares algorithm of choice — e.g. Normal eqs, QR, etc.

example:

$$\vec{x}_1 = \begin{bmatrix} 2 \\ 4 \\ 1 \\ 3 \\ 5 \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} 15 \\ 5 \\ 2 \\ 11 \\ 4 \end{bmatrix}$$

fit the data to the cubic polynomial model:

$$y = c_0 + c_1 x_1 + c_2 x_1^2 + c_3 x_1^3$$

$$\overbrace{\begin{bmatrix} 15 \\ 5 \\ 2 \\ 11 \\ 4 \end{bmatrix}}^{\vec{y}} = \overbrace{\begin{bmatrix} 1 & 2 & 4 & 8 \\ 1 & 4 & 16 & 64 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 81 \\ 1 & 5 & 25 & 125 \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}}^{\vec{c}}$$

Let's use polynomial regression on actual data! [Boston housing data]

problem with polynomial regression: overfitting to data

[show slides]