Lecture 1

Regression

- Curve fitting
  - Given a set of points, try ot learn a function to describe them
  - Given a value x, we can predict the corresponding value y
  - Not just for straight line fitting

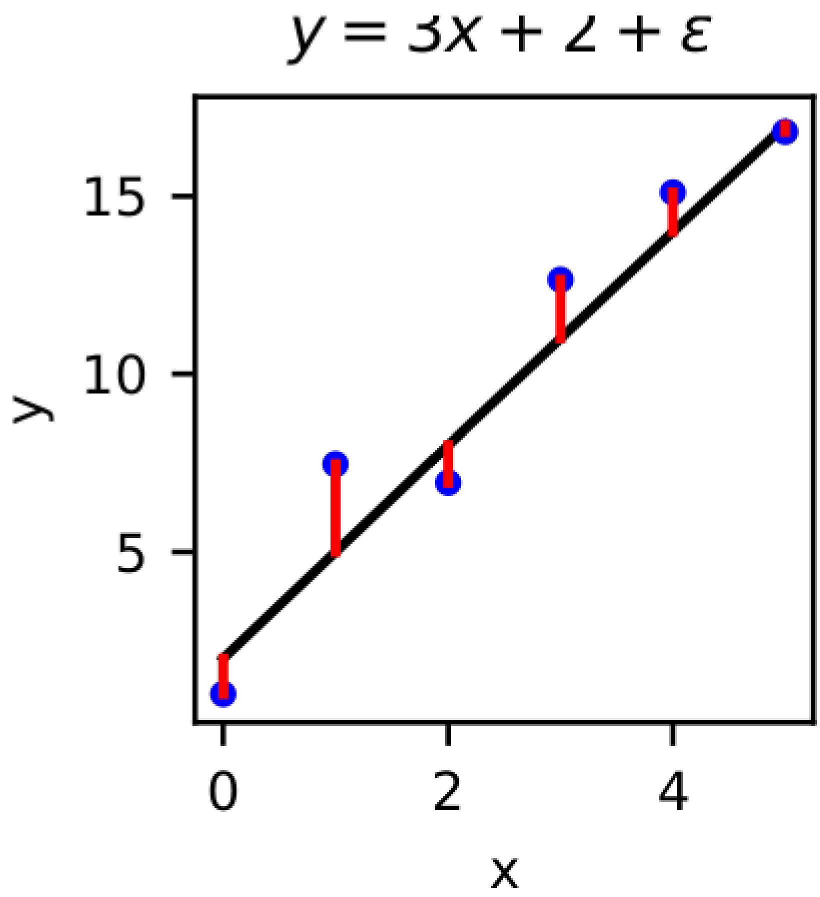
Simple example

Let us consider a simple *linear* example with 1 independent & 1 dependent variable $$D = {(x_1,y_1),..., (x_n,y_n)} = {(x_i,y_i)}_{i=1}^N$$ Model the relationship between x and y with the function $L(\textbf{w},x)$, s.t $y \approx f(\textbf{w},x)$ Measurements of y, subject ot noise are defined by, $$y_i = f(\textbf{w},w) + \epsilon$$ Where ϵ is a random number drawn from some continuous probability density function. Goal is to find some $\textbf{w}$ that solves the above equation

First, let us approach this as a optimisation problem in which the objective is to find the value of **w** (denoted **w***) that minimises some *loss* or objective function $L(\textbf{w})$

$$\textbf{w}* = \argmin_w L(\textbf{w}) $$

Intuitively, $L(\textbf{w})$ should be designed to capture the difference between the data and the predictions of the model, and seek to minimise this. One common choice for $L(\textbf{w})$ is *least-squares error*. Given our dataset D and modelling function $f(\textbf{w},x)$, we construct $\forall d \in D$ a residual error defined as: $$ r_i(\textbf{w}) = y_i - f(\textbf{w}, x_i)$$


$y = 3x + 2 + \varepsilon$
y
x
0
2
4
5
10
15