# Linear, Weighted and Locally Weighted Regression

In this exercise you will compare three different regression techniques ($\mathbf{x} \in \mathbb{R}^N$ and $y \in \mathbb{R}$), namely:

1. **Regular Least Squares (RLS):**
   - Regressor: $y = w^T\mathbf{x} + b$
   - Optimisation: $w = (XX^T)^{-1}Xy$
2. **Weighted Least Squares (WLS):**
   - Regressor: $y = w^T\mathbf{x} + b$
   - Optimisation: $w = (ZZ^T)^{-1}Zv$ where $Z = XB^{1/2}$ and $v = B^{1/2}y$
3. **Locally Weighted Regression (LWR):**
   - Regressor: $y = \left(\sum_{i=1}^{M} \beta_i(\mathbf{x})y^i\right) / \left(\sum_{i=1}^{M} \beta_i(\mathbf{x})\right)$

     The beta is a kernel density function centred on a point $i$: $\beta_i(\mathbf{x}) = \exp(-\frac{1}{2}\|\mathbf{x}^i - \mathbf{x}\|^{\frac{1}{2}})$
   - Optimisation: no-optimisation, data driven.

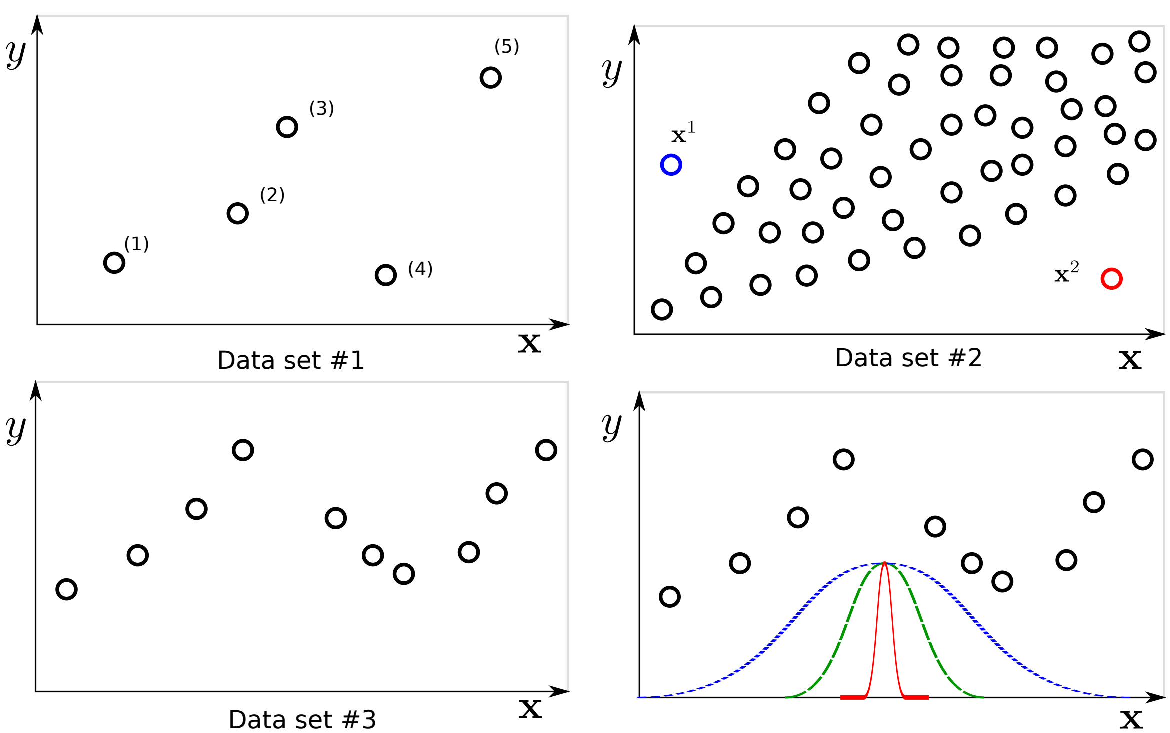


Figure 1: Three datasets, the black circles depict the data points. $\mathbf{x}$ is the input and $y$ is the output and we wish to estimate $y = f(\mathbf{x})$.

**A)** In Figure 1, three different datasets are given

1. Draw the solution that RLS would give you for datasets 1 to 3 (do not consider the colored points in dataset 2).

2. Given the set of weights $\boldsymbol{\beta} = [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0, \frac{1}{4}]$, apply WLS to dataset 1 and draw the resulting regression function.

3. What solution WLS would give for dataset 2, considering that the blue (point $\mathbf{x}^1$) and the red data point (point $\mathbf{x}^2$) are weighted with $\beta(\mathbf{x}) = \frac{1}{\mathbf{x}}$.

4. Draw the solutions of LWR for dataset 3 with each of the given kernels (see Figure 1, *Bottom right*).

**B)** Your lab (Lab 1) is studying a rare type of particles. Using particles with different sizes your lab took measurements of their speed. You wanted more data so you asked a cooperating lab (Lab 2) to share their measurements with you (figure 2). Lab 1 was using a measuring instrument with the Gaussian error $e_1 \sim \mathcal{N}(0, 10)$, while the error of Lab 2 measurements was $e_2 \sim \mathcal{N}(0, 20)$. You want to find out what's the linear relation between the speed of a particle and its size. Which regression method should you use, and how would you use it?

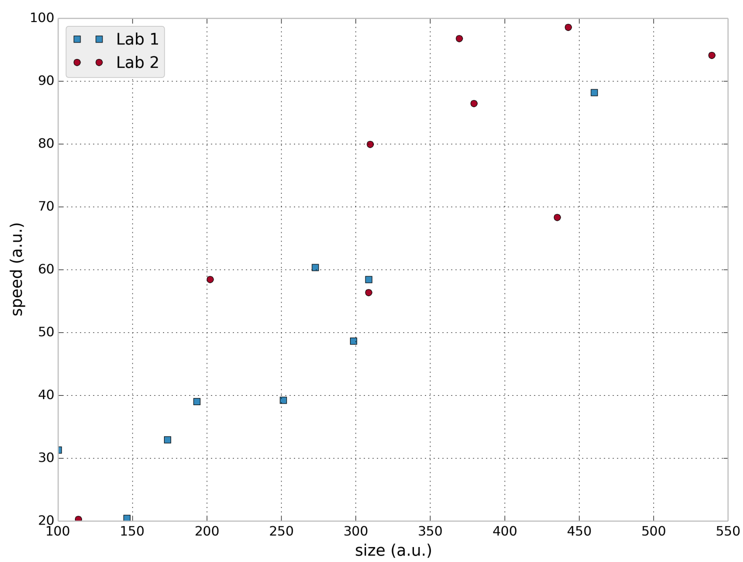


Figure 2: The datapoints collected from the two labs.

## Least Squares

In the lecture you have covered linear function estimators of the following form:

$$y^i = f(\mathbf{x}^i; \mathbf{w}, b) = \mathbf{w}^{\mathrm{T}}\mathbf{x}^i + b \tag{1}$$

where $\mathbf{w} \in \mathbb{R}^N$ and $\mathbf{x} \in \mathbb{R}^N$ are $(N \times 1)$ column vectors, $b$ is the scalar intercept and $y$ is the predictor.

Given you have a set of $M$ data points, $X = [\mathbf{x}^1, \ldots, \mathbf{x}^i, \ldots, \mathbf{x}^M]$, and associated predictors, $\mathbf{y} = [y^1, \ldots, y^i, \ldots, y^M]$. Consider the Sum of Squared Error (SSE) as your loss function and derive the optimal choice of parameters of the linear regressor for the **bivariate** case:

$$y^i = wx^i + b \tag{2}$$

$$SSE = \sum_{i=1}^{M} \left(y^i - f(x^i)\right)^2 = \sum_{i=1}^{M} e_i^2 \tag{3}$$

where $e_i$ is the error between the target and predicted value.

## Control of Robotic Manipulator (to be done at home)

Consider the 3 degree of freedom, $\mathbf{q} = \{q_1, q_2, q_3\}$, robotic arm in Fig. 3. The vector $\mathbf{q}$ denotes the current joints' position while $\mathbf{x} \in \mathbb{R}^2$ is the location in the 2D space of the tip of the robotic arm, also known as end-effector. The position of the end-effector is connected to the joints' position

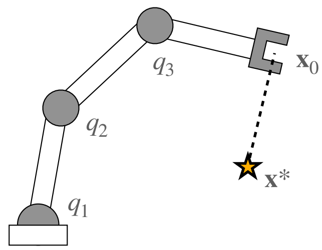


Figure 3: 3 degree of freedom robotic manipulator.

through the forward kinematics equation $\mathbf{x} = \phi(\mathbf{q})$.

**A)** We are interested in generating a joints velocity vector, $\dot{\mathbf{q}}$, that would move the end-effector of the robot from the current location $\mathbf{x}_0$ towards the goal location $\mathbf{x}^*$. Show that the optimal $\dot{\mathbf{q}}$ is the solution of a least-square linear (unweighted) regression of the form $\mathbf{w} = (XX^T)^{-1}X^T\mathbf{y}$ (Hint: the derivative of the forward kinematics with respect to the joints' position is $J(\mathbf{q}) = \frac{\partial\phi(\mathbf{q})}{\partial\mathbf{q}}$, also know as Jacobian).

**B)** We would like to move the first joint of the robot without changing the current location of the end-effector. Is there any other joints velocity vector $\dot{\mathbf{q}}$, solution of the linear regression problem derived in the previous step, that would achieve this?