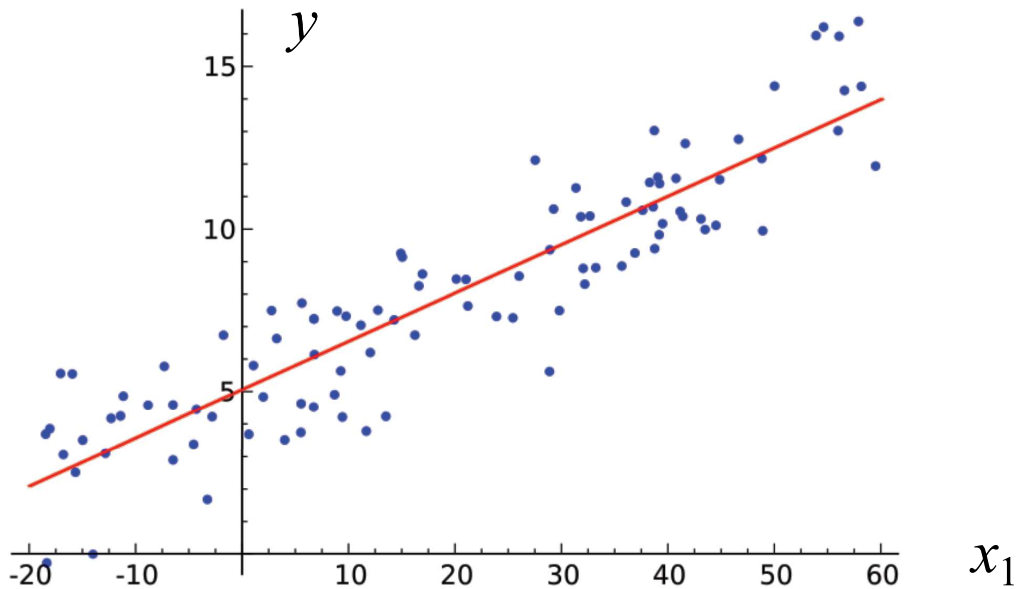
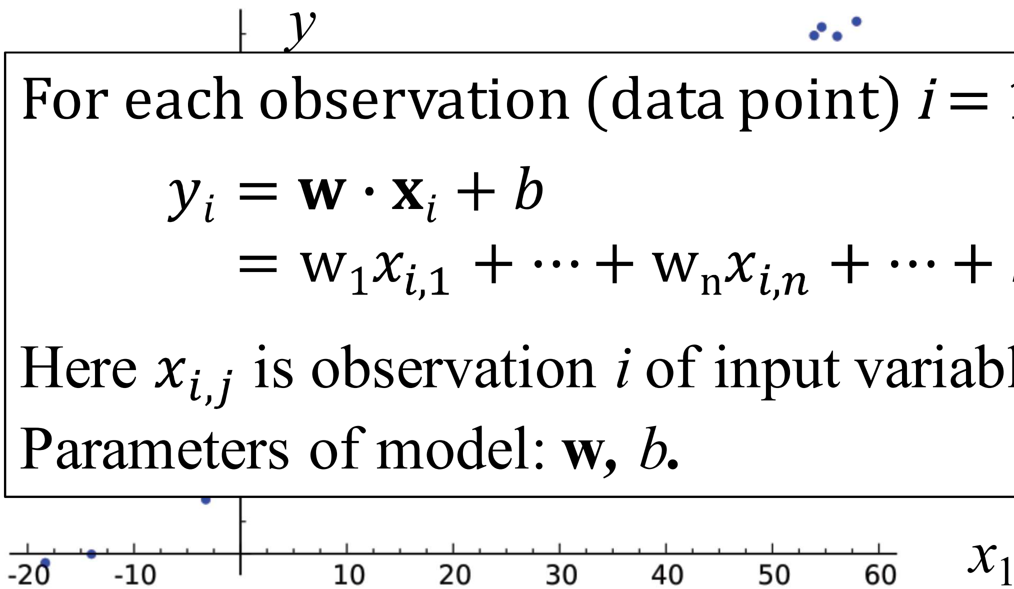
Linear Regression

- Uses a linear model to model relationship between dependent variable $y \in \mathbb{R}$, and input (independent) variables $x_1, \ldots, x_n \in \mathbb{R}^n$
- Is this supervised or unsupervised learning?



Linear Regression

- Uses a linear model to model relationship between dependent variable $y \in \mathbb{R}$, and input (independent) variables $x_1, \ldots, x_n \in \mathbb{R}^n$



For each observation (data point) $i = 1, \ldots, m$:

$$
\begin{aligned}
y_i &= \mathbf{w} \cdot \mathbf{x}_i + b \\
&= \mathrm{w}_1 x_{i,1} + \cdots + \mathrm{w}_\mathrm{n} x_{i,n} + \cdots + b
\end{aligned}
$$

Here $x_{i,j}$ is observation i of input variable j.

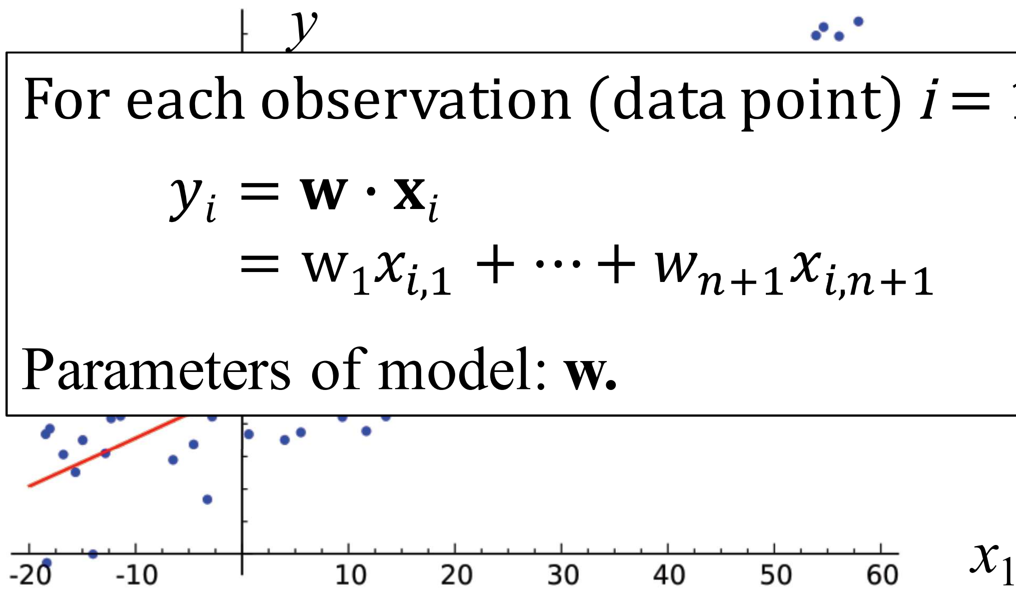
Parameters of model: $\mathbf{w}$, b.

Linear Regression

- Can simply the model by adding additional input that is always one:

  $x_{i,n+1} = 1 \qquad i = 1, \ldots, m$

- The corresponding parameter in w is called the **intercept**.



For each observation (data point) $i = 1, \ldots, m$:

$$
\begin{aligned}
y_i &= \mathbf{w} \cdot \mathbf{x}_i \\
&= \mathrm{w}_1 x_{i,1} + \cdots + w_{n+1} x_{i,n+1}
\end{aligned}
$$

Parameters of model: **w.**

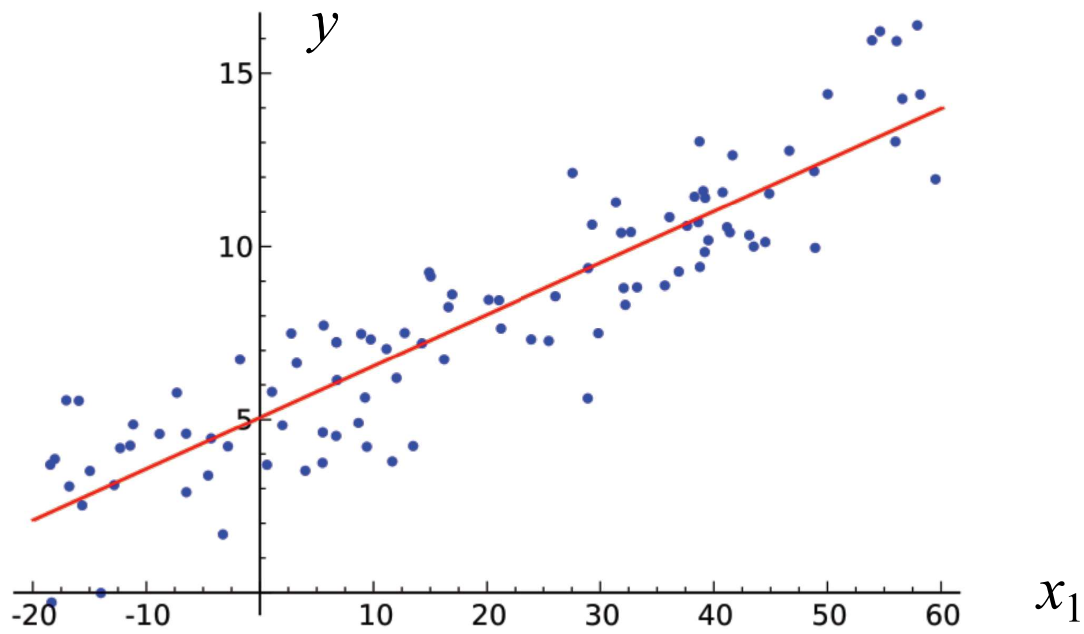
Linear Least Squares Regression

- Define an objective function or loss function to optimize the model
- One loss function: least squares (“least squared error”).

$$E = \sum_{i=1}^{m} (\mathbf{w} \cdot \mathbf{x}_i - y_i)^2$$

Parameters of model: **w.**

- What is E for the 2D line fitting case at right? (blackboard)
- How to minimize E?



Linear Least Squares Regression

Set derivatives of objective function with respect to parameters equal to zero.

$$\frac{\partial E}{\partial w_j} = \frac{\partial}{\partial w_j} \sum_{i=1}^{m} (\mathbf{w} \cdot \mathbf{x}_i - y_i)^2 = 0$$

$$2 \sum_{i=1}^{m} (\mathbf{w} \cdot \mathbf{x}_i - y_i) x_{ij} = 0$$

$$\sum_{i=1}^{m} \left(\sum_{k=1}^{n} x_{ik} w_k - y_i \right) x_{ij} = 0$$

Normal equations:

$$\sum_{i=1}^{m} \sum_{k=1}^{n} x_{ik} x_{ij} w_k = \sum_{i=1}^{m} x_{ij} y_i$$

Linear Least Squares Regression

- Normal equations in matrix form:

$$(\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{w} = \mathbf{X}^{\mathrm{T}}\boldsymbol{y}$$

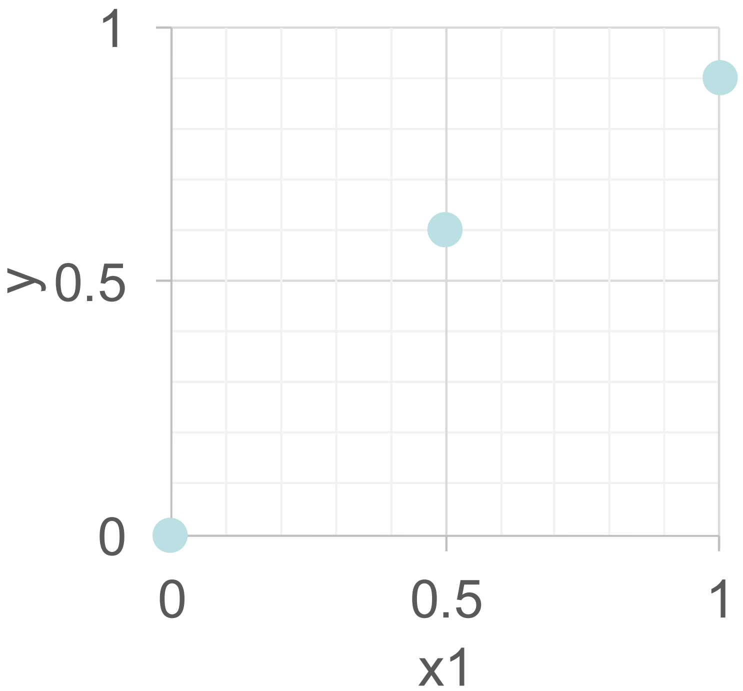
$$\mathbf{w} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{y}$$

X is the matrix with x_{ij} being observation i of input variable j.

y is the vector of dependent variable (output) observations.

Linear Least Squares Example

- Suppose we have three observations (m=3) of one input variable x_1:



x_1	y
0	0
0.5	0.6
1	0.9

Linear Least Squares Example

- Suppose we have three observations (m=3) of one input variable x_1.
- Add additional constant variable x_2:

x_1	x_2	y
0	1	0
0.5	1	0.6
1	1	0.9

- $\mathbf{X} = \begin{bmatrix} 0 & 1 \\ 0.5 & 1 \\ 1 & 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 0 \\ 0.6 \\ 0.9 \end{bmatrix}$, so $\mathbf{w} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{y} = \begin{bmatrix} 0.9 \\ 0.05 \end{bmatrix}$

Linear Least Squares Example