## Linear Models

§ 8.3.2

$y = \beta_0 + \beta_1 x$

1. Select loss function

   quadratic loss

2. Write out total loss

$$L = \sum_{i=1}^{N} \left( y_i^{model} - y^{true} \right)^2$$

model contains unknown params ($\beta$)

3. Minimize loss

**Model:**

$$y = \beta_0 \Rightarrow \beta_0 = \frac{1}{N} \sum_{i=1}^{N} y_i^{true}$$

| $x$ | $y$ |
|---|---|
| 0.07 | -0.05 |
| 0.16 | 0.40 |
| 0.48 | 0.66 |
| 0.68 | 0.65 |
| 0.83 | 1.12 |

$$y = \beta_0 + \beta_1 x$$

$$-0.05 = \beta_0 + 0.07\beta_1 + \varepsilon_1$$
$$0.40 = \beta_0 + 0.16\beta_1 + \varepsilon_2$$
$$0.66 = \beta_0 + 0.48\beta_1 + \varepsilon_3$$
$$0.65 = \beta_0 + 0.68\beta_1 + \varepsilon_4$$
$$1.12 = \beta_0 + 0.83\beta_1 + \varepsilon_5$$

$$\Downarrow$$

$$\begin{pmatrix} -0.05 \\ 0.40 \\ 0.66 \\ 0.65 \\ 1.12 \end{pmatrix} = \begin{pmatrix} 1 & 0.07 \\ 1 & 0.16 \\ 1 & 0.48 \\ 1 & 0.68 \\ 1 & 0.83 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix}$$

$$\underline{y} = \underline{X}\,\underline{\beta} + \underline{\varepsilon}$$

model matrix

pseudoinverse

$$\underline{\beta} = \underline{X}^{+}\underline{y}$$

Matlab: pinv

If $\underline{X}$ is full rank

$$\underline{X}^{+} = (\underline{X}^T\underline{X})^{-1}\underline{X}^T$$

---

$$\underline{y} = \underline{X}\,\underline{\beta}$$
$$\underline{X}^T\underline{y} = \underline{X}^T\underline{X}\,\underline{\beta}$$
$$(\underline{X}^T\underline{X})^{-1}\underline{X}^T\underline{y} = (\underline{X}^T\underline{X})^{-1}\underline{X}^T\underline{X}\,\underline{\beta}$$
$$(\underline{X}^T\underline{X})^{-1}\underline{X}^T\underline{y} = \underline{\beta}$$

§ 9.3

$(x_1, x_2 \to y)$

height, weight, BP

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

$$\downarrow$$

$$\underline{y} = \underline{X}\,\underline{\beta}$$

$$\downarrow$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & x_{2,1} \\ 1 & x_{1,2} & x_{2,2} \\ \vdots & \vdots & \vdots \\ 1 & x_{1,N} & x_{2,N} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

| $x_1$ height | $x_2$ weight | $y$ MAP |
|---|---|---|
| 72 | 200 | 110 |
| 60 | 160 | 106 |
| 66 | 154 | 92 |
| 58 | 135 | 100 |

$$\begin{pmatrix} 110 \\ 106 \\ 92 \\ 100 \end{pmatrix} = \begin{pmatrix} 1 & 72 & 200 \\ 1 & 60 & 160 \\ 1 & 66 & 154 \\ 1 & 58 & 135 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

## Curvilinear models

Models are linear if they are linear in $\underline{\beta}$.

$$y = \beta_0 + \beta_1 x \implies \underline{y} = \underline{X}\,\underline{\beta}$$

weight $= f(\text{height})$

weight $\propto$ height$^2$

$$\begin{aligned} 200 &= \beta_0 + (72)^2 \beta_1 \\ 160 &= \beta_0 + (60)^2 \beta_1 \\ 154 &= \beta_0 + (66)^2 \beta_1 \\ 135 &= \beta_0 + (58)^2 \beta_1 \end{aligned} \implies \begin{pmatrix} 200 \\ 160 \\ 154 \\ 135 \end{pmatrix} = \begin{pmatrix} 1 & 72^2 \\ 1 & 60^2 \\ 1 & 66^2 \\ 1 & 58^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 \ldots$$

## Linearization §9.4

$$N(t) = N_0 e^{\mu t}$$

$$\log(N(t)) = \log(N_0 e^{\mu t})$$

$$\log(N(t)) = \log N_0 + \log e^{\mu t}$$

$$\underbrace{\log N(t)}_{y} = \underbrace{\log N_0}_{\beta_0} + \mu t$$

$$y = \beta_0 + \beta_1 t$$