Lecture 12: Univariate Regression

Course Teacher: Md. Shariful Islam Bhuyan

Regression

- Regression as function approximation

$$y \approx h(\mathbf{x}) \qquad y_i = h(\mathbf{x}_i) + \epsilon_i \qquad \epsilon_i = y_i - h(\mathbf{x}_i)$$

- True loss, expected generalization loss or risk is a functional

$$\mathcal{L}_{true}(h) = \mathbb{E}_{P(\mathbf{x},y)}\big[L\big(y, h(\mathbf{x})\big)\big] = \int L\big(y, h(\mathbf{x})\big)P(\mathbf{x}, y)\, d\mathbf{x}\, dy = \sum_{(\mathbf{x},y)} L\big(y, h(\mathbf{x})\big)P(\mathbf{x}, y)$$

- Empirical risk function (objective function in non-probabilistic setting)

$$\mathcal{L}_{emp}(h) = \frac{1}{N}\sum_{i=1}^{N} L\big(y_i, h(\mathbf{x}_i)\big)$$

- Empirical Risk Minimization

$$h^* = \arg\min_{h \in \mathcal{H}} \mathcal{L}_{emp}(h)$$

Univariate Linear Regression

- *Single variable linear regression*

$$f(x) = w_0 + w_1 x \qquad L\big(y, f(x)\big) = (y - w_0 - w_1 x)^2$$

$$(w_0, w_1)^* = \arg\min_f \mathcal{L}_{emp}(f) = \arg\min_{(w_o, w_1)} \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2$$

$$= \arg\min_{(w_o, w_1)} \left\| \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} \right\|^2 = \arg\min_{(w_o, w_1)} \|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2$$

$$\frac{\partial \mathcal{L}_{emp}}{\partial w_0} = \frac{2}{n} \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)\,(-1) = 0 = \sum_{i=1}^{n} \epsilon_i$$

$$\frac{\partial \mathcal{L}_{emp}}{\partial w_1} = \frac{2}{n} \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)\,(-x) = 0 = \sum_{i=1}^{n} \epsilon_i x_i$$

Univariate Polynomial Regression

- *Single variable polynomial regression*

$$f(x) = w_0 + w_1 x + w_2 x^2 + \cdots + w_d x^p = \mathbf{w}^{\mathrm{T}}\mathbf{x} \qquad \mathbf{w} \in \mathbb{R}^{\mathrm{d}}$$

$$L\big(y, f(\mathbf{x})\big) = \left(y - \mathbf{w}^{\mathrm{T}}\mathbf{x}\right)^2$$

$$\mathbf{w}^* = \arg\min_f \mathcal{L}_{emp}(f) = \arg\min_{\mathbf{w}\in\mathbb{R}^{\mathrm{d}}} \sum_{i=1}^{n} \left(y_i - \mathbf{w}^{\mathrm{T}}\mathbf{x}_i\right)^2$$

$$= \arg\min_{\mathbf{w}\in\mathbb{R}^{\mathrm{d}}} \left\| \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^p \\ 1 & x_2 & x_2^2 & \cdots & x_2^p \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^p \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} \right\|^2 = \arg\min_{\mathbf{w}\in\mathbb{R}^{\mathrm{d}}} \|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2$$