Live Interaction #1:

14th January 2024

E-masters Next Generation Wireless Technologies

EE902 Advanced ML Techniques for Wireless Technology

Linear regression:

	TV	Radio	Newspaper	Sales
0	230.1	37.8	69.2	22.1
1	44.5	39.3	45.1	10.4
2	17.2	45.9	69.3	9.3
3	151.5	41.3	58.5	18.5
4	180.8	10.8	58.4	12.9

- Sales: $y(k)$ - **Response**
- Advertising:$x_1(k)$,$x_2(k)$,… - **Regressors, Explanatory variable**

$$y(k) = h_0 + h_1 x_1(k) + h_2 x_2(k) + \cdots + h_N x_N(k) + \epsilon(k)$$

- h_i: Regression coefficients
- h_0: Bias
- $\epsilon(k)$: Noise, Model error
- Model for prediction

$$y(k) = h_0 + h_1 x_1(k) + h_2 x_2(k) + \cdots + h_N x_N(k)$$

- To estimate the regression coefficients

$$y(k) = h_0 + h_1 x_1(k) + h_2 x_2(k) + \cdots + h_N x_N(k) + \epsilon(k)$$

$$= \underbrace{[1 \quad x_1(k) \quad \dots \quad x_N(k)]}_{\bar{\mathbf{x}}^T(k)} \underbrace{\begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_N \end{bmatrix}}_{\bar{\mathbf{h}}} + \epsilon(k)$$

$$y(k) = \bar{\mathbf{x}}^T(k)\bar{\mathbf{h}} + \epsilon(k)$$

- How to learn?
- What to learn?
- We have to learn the model. Characterized by the ***regression coefficients***.

$$\bar{\mathbf{h}} = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_N \end{bmatrix}$$

- **Training data:**

$$y(1), \bar{\mathbf{x}}^T(1)$$
$$y(2), \bar{\mathbf{x}}^T(2)$$
$$\vdots$$
$$y(M), \bar{\mathbf{x}}^T(M)$$

- **Supervised learning**: Responses and inputs are available in the training data.
- Linear regression is a supervised learning technique.
- What is the relation between M,N

$$M \gg N$$

- M is the number of **data points**.
- N is the model order.

$$\underbrace{\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(M) \end{bmatrix}}_{\bar{\mathbf{y}}} = \underbrace{\begin{bmatrix} \bar{\mathbf{x}}^T(1) \\ \bar{\mathbf{x}}^T(2) \\ \vdots \\ \bar{\mathbf{x}}^T(M) \end{bmatrix}}_{\substack{\mathbf{X} \\ M\times(N+1)}} \bar{\mathbf{h}} + \underbrace{\begin{bmatrix} \epsilon(1) \\ \epsilon(2) \\ \vdots \\ \epsilon(M) \end{bmatrix}}_{\bar{\boldsymbol{\epsilon}}}$$

$$\bar{\mathbf{y}} = \mathbf{X}\bar{\mathbf{h}} + \bar{\boldsymbol{\epsilon}}$$

- $\mathbf{X}$: Tall matrix
- To determine $\bar{\mathbf{h}}$

$$\min \left\| \bar{\mathbf{y}} - \mathbf{X}\bar{\mathbf{h}} \right\|^2$$

$$\bar{\mathbf{h}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\bar{\mathbf{y}}$$

- $\mathbf{X}^\dagger = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$: Pseudo-inverse
- Example:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \bar{\mathbf{y}} = \begin{bmatrix} -1 \\ 2 \\ -2 \\ -3 \end{bmatrix}$$

$$\mathbf{X}^T\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix}$$

$$(\mathbf{X}^T\mathbf{X})^{-1} = \frac{1}{20}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}$$

$$(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\bar{\mathbf{y}}$$

$$= \frac{1}{20}\begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ -2 \\ -3 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3}{2} \\ -1 \end{bmatrix}$$