Applied Numerical Methods for Civil Engineering

CGN 3405 - 0002

Week 14: Curve Fitting

Xinyu Chen

Assistant Professor

University of Central Florida

Definition

- **Definition**: Linear regression is a model that estimates the relationship between a scalar response (dependent variable) and one or more explanatory variables (independent variable).
- Example:

$$\{x_i, y_i\}, \quad i = 1, 2, \ldots, n$$

  - Independent variables $x_i, i = 1, 2, \ldots, n$
  - Dependent variables $y_i, i = 1, 2, \ldots, n$
- Regression formula such as

$$y_i = ax_i + b + \varepsilon_i$$

Optimization

Minimizing the residuals of linear regression:

$$y_i = ax_i + b + \varepsilon_i$$

- Sum of squared residuals:

$$\min_{a,b} \quad \sum_{i=1}^{n} (y_i - (ax_i + b))^2$$

- Sum of absolute residuals:

$$\min_{a,b} \quad \sum_{i=1}^{n} |y_i - (ax_i + b)|$$

Sum of Squared Residuals

Given data points $\{x_i, y_i\}, i = 1, 2, \ldots, n$:

- Reformulate the optimization

$$\begin{aligned} & \min_{a,b} \quad \sum_{i=1}^{n} (y_i - (ax_i + b))^2 \\ \Rightarrow \quad & \min_{\boldsymbol{w}} \quad \sum_{i=1}^{n} \left(y_i - \boldsymbol{w}^\top \boldsymbol{z}_i\right)^2 \end{aligned}$$

  with an inner product between

$$\boldsymbol{w} = \begin{bmatrix} a \\ b \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{z}_i = \begin{bmatrix} x_i \\ 1 \end{bmatrix}$$

Objective Function with ℓ_2-Norm

Given data points $\{x_i, y_i\}, i = 1, 2, \ldots, n$:

- Introduce ℓ_2-norm:

$$\min_{\boldsymbol{w}} \quad \|\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{w}\|_2^2$$

  with

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n \qquad \text{and} \qquad \boldsymbol{Z} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \in \mathbb{R}^{n \times 2}$$

Objective Function with ℓ_2-Norm

Given data points $\{x_i, y_i\}, i = 1, 2, \ldots, n$:

- Introduce ℓ_2-norm:

$$\min_{\boldsymbol{w}} \quad \|\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{w}\|_2^2$$

  with

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n \qquad \text{and} \qquad \boldsymbol{Z} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \in \mathbb{R}^{n\times 2}$$

- **Gradient**: First-order derivative of objective function

$$\boldsymbol{g} = 2\boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{w} - \boldsymbol{y}) = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{Z}^\top\boldsymbol{Z}\boldsymbol{w} = \boldsymbol{Z}^\top\boldsymbol{y}$$

- **Closed-form solution**:

$$\boldsymbol{w} = (\boldsymbol{Z}^\top\boldsymbol{Z})^{-1}\boldsymbol{Z}^\top\boldsymbol{y}$$

Linear Regression

Linear regression with 10 data pairs $\{x_i, y_i\}, i = 1, 2, \ldots, 10$:

- Independent and dependent variables:

$$\boldsymbol{x} = \begin{bmatrix} 3.2 \\ 1.9 \\ 2.4 \\ 3.7 \\ 3.3 \\ 0.5 \\ 2.4 \\ 1.3 \\ 1.3 \\ 1.9 \end{bmatrix} \qquad \boldsymbol{y} = \begin{bmatrix} 2.6 \\ 2 \\ 2.2 \\ 2.8 \\ 2.6 \\ 1.2 \\ 2.3 \\ 1.6 \\ 1.6 \\ 1.8 \end{bmatrix} \qquad \Rightarrow \qquad \boldsymbol{Z} = \begin{bmatrix} 3.2 & 1 \\ 1.9 & 1 \\ 2.4 & 1 \\ 3.7 & 1 \\ 3.3 & 1 \\ 0.5 & 1 \\ 2.4 & 1 \\ 1.3 & 1 \\ 1.3 & 1 \\ 1.9 & 1 \end{bmatrix}$$

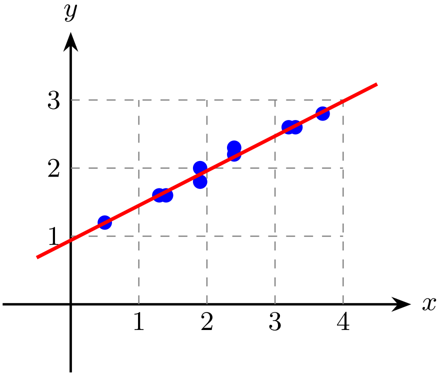
- The optimal coefficients are

$$\boldsymbol{Z}^\top \boldsymbol{Z} = \begin{bmatrix} 57.46 & 22 \\ 22 & 10 \end{bmatrix} \qquad \Rightarrow \qquad \boldsymbol{w} = (\boldsymbol{Z}^\top \boldsymbol{Z})^{-1} \boldsymbol{Z}^\top \boldsymbol{y} = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.51 \\ 0.94 \end{bmatrix}$$

Linear Regression

Linear regression with 10 data pairs $\{x_i, y_i\}, i = 1, 2, \ldots, 10$:

$$y = 0.51x + 0.94$$



Linear Regression with Optimization

```python
import numpy as np

x = np.array([3.2, 1.9, 2.4, 3.7, 3.3, 0.5, 2.4, 1.3, 1.4,
    1.9])
y = np.array([2.6, 2,  2.2, 2.8, 2.6, 1.2, 2.3, 1.6, 1.6,
    1.8])

n = x.shape[0]
Z = np.ones((n, 2))
Z[:, 0] = x
w = np.linalg.inv(Z.T @ Z) @ Z.T @ y
print(w)
```

Linear Regression with Optimization

Optimization problem:

$$\min_{\boldsymbol{w}} \quad \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{w}\|_2^2$$

- **Gradient**: First-order derivative of objective function

$$\boldsymbol{g} = \boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{w} - \boldsymbol{y})$$

- **Hessian matrix**: Second-order derivative of objective function

$$\boldsymbol{H} = \boldsymbol{Z}^\top\boldsymbol{Z} \qquad \text{(constant matrix)}$$

Linear Regression with Optimization

Newton's method:

- Update gradient:
$$\boldsymbol{g}_k = \boldsymbol{Z}^\top \boldsymbol{Z}\boldsymbol{w}_k - \boldsymbol{Z}^\top \boldsymbol{y}$$
- Update variables:
$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k - \boldsymbol{H}^{-1}\boldsymbol{g}_k$$
- Why do we have exact solution at the first iteration if we start with $\boldsymbol{w}_0 = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$?

Linear Regression with Optimization

Newton's method:

- Update gradient:

$$\boldsymbol{g}_k = \boldsymbol{Z}^\top \boldsymbol{Z}\boldsymbol{w}_k - \boldsymbol{Z}^\top \boldsymbol{y}$$

- Update variables:

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k - \boldsymbol{H}^{-1}\boldsymbol{g}_k$$

- Why do we have exact solution at the first iteration if we start with $\boldsymbol{w}_0 = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$?

$$\begin{aligned} \boldsymbol{w}_1 &= \boldsymbol{w}_0 - \boldsymbol{H}^{-1}\boldsymbol{g}_0 \\ &= \boldsymbol{w}_0 - (\boldsymbol{Z}^\top \boldsymbol{Z})^{-1}(\boldsymbol{Z}^\top \boldsymbol{Z}\boldsymbol{w}_0 - \boldsymbol{Z}^\top \boldsymbol{y}) \\ &= \mathbf{0} - (\boldsymbol{Z}^\top \boldsymbol{Z})^{-1}(\boldsymbol{Z}^\top \boldsymbol{Z}\mathbf{0} - \boldsymbol{Z}^\top \boldsymbol{y}) \\ &= (\boldsymbol{Z}^\top \boldsymbol{Z})^{-1}\boldsymbol{Z}^\top \boldsymbol{y} \end{aligned}$$

This is just the closed-form solution!

Sum of Absolute Residuals

Given data points $\{x_i, y_i\}, i = 1, 2, \ldots, n$:

- Equivalent to ℓ_1-norm:

$$\min_{\boldsymbol{w}} \quad \sum_{i=1}^{n} \left| y_i - \boldsymbol{w}^\top \boldsymbol{z}_i \right|$$

- Optimization reformulation:

$$\begin{aligned} \min_{\boldsymbol{w},\boldsymbol{c}} \quad & \sum_{i=1}^{n} c_i \\ \text{s.t.} \quad & -c_i \le y_i - \boldsymbol{w}^\top \boldsymbol{z}_i \le c_i \quad && i = 1, 2, \ldots, n \\ & c_i \ge 0 && i = 1, 2, \ldots, n \end{aligned}$$

Multiple Linear Regression

- Multiple linear regression is a generalization of simple linear regression to the case of more than one independent variable.
- Independent variables $x_{i1}, x_{i2}, \ldots, x_{id}$, $i = 1, 2, \ldots, n$
- Dependent variables y_i, $i = 1, 2, \ldots, n$
- Formula:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_d x_{id} + \varepsilon_i$$

Multiple Linear Regression

- Multiple linear regression is a generalization of simple linear regression to the case of more than one independent variable.
- Independent variables $x_{i1}, x_{i2}, \ldots, x_{id}$, $i = 1, 2, \ldots, n$
- Dependent variables y_i, $i = 1, 2, \ldots, n$
- Formula:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_d x_{id} + \varepsilon_i$$

- Equivalent form:

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

with

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n \quad \boldsymbol{Z} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1d} \\ 1 & x_{21} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nd} \end{bmatrix} \in \mathbb{R}^{n \times (d+1)} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_d \end{bmatrix} \in \mathbb{R}^d$$

Multiple Linear Regression

Optimization:

$$\min_{\boldsymbol{\beta}} \quad \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{Z}\boldsymbol{\beta}\|_2^2$$

- **Gradient**: First-order derivative of objective function

$$\boldsymbol{g} = \boldsymbol{Z}^\top(\boldsymbol{Z}\boldsymbol{\beta} - \boldsymbol{y}) = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{Z}^\top\boldsymbol{Z}\boldsymbol{\beta} = \boldsymbol{Z}^\top\boldsymbol{y}$$

- **Closed-form solution**:

$$\boldsymbol{\beta} = (\boldsymbol{Z}^\top\boldsymbol{Z})^{-1}\boldsymbol{Z}^\top\boldsymbol{y}$$

Quick Summary

Today's Class:

- Definition of linear regression
- Optimization of linear regression
- Multiple linear regression