# Lecture 17: Conditional Variance Regression

Instructor: Prof. Shuai Huang

Industrial and Systems Engineering

University of Washington

# Conditional variance regression model

- **Heteroscedasticity** refers to the phenomenon that the variance of the response variable may also change

- This leads to the following model:

$$y = \boldsymbol{\beta}^T \boldsymbol{x} + \epsilon_{\boldsymbol{x}},$$

- and $\epsilon_{\boldsymbol{x}}$ is the error term that is a normal distribution with varying variance:

$$\epsilon_{\boldsymbol{x}} \sim N\left(0, \sigma_{\boldsymbol{x}}^2\right).$$

# Parameter estimation ($\sigma_x^2$ is known)

- If we have known the $\sigma_x^2$, this will lead to the following scheme for parameter estimation of the unknown regression parameters. The likelihood function is:

$$-\frac{n}{2}\ln 2\pi - \frac{1}{2}\sum_{n=1}^{N}\log \sigma_{x_n}^2 - \frac{1}{2}\sum_{n=1}^{N}\frac{\left(y_n - \boldsymbol{\beta}^T x_n\right)^2}{\sigma_{x_n}^2}.$$

- As we have known $\sigma_x^2$, the parameters to be estimated only involve the last part of the likelihood function. Thus, we estimate the parameters that minimize

$$\frac{1}{2}\sum_{n=1}^{N}\frac{\left(y_n - \boldsymbol{\beta}^T x_n\right)^2}{\sigma_{x_n}^2}.$$

- This could be written in the matrix form as

$$\min_{\boldsymbol{\beta}}(\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta})^T\mathbf{W}(\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta}),$$

- where $\mathbf{W}$ is a diagonal matrix with its diagonal elements as $\mathbf{W}_{nn} = \frac{1}{\sigma_{x_n}^2}$.

- And we can get that $\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}\boldsymbol{Y}$.

# Parameter estimation ($\sigma_x^2$ is unknown)

We propose the following steps:

- 1. Initialize $\hat{\sigma}_{x_n}^2$ for $n = 1, 2, \ldots, N$, by any reasonable approach including the random generation of values.

- 2. Build a regression model for the mean of the response variable using the weighted LS estimator. Estimate $\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{W}\boldsymbol{Y}$ and get $\hat{y}_n = \widehat{\boldsymbol{\beta}}^T\boldsymbol{x}_n$.

- 3. Derive the residuals $\hat{\varepsilon}_n = y_n - \hat{y}_n$.

- 4. Build a regression model, e.g., using the kernel regression which is a nonparametric method, to fit $\hat{\varepsilon}_n^2$ using $\boldsymbol{x}_n$ for $n = 1, 2, \ldots, N$.

- 5. Predict $\hat{\sigma}_{x_n}^2$ for $n = 1, 2, \ldots, N$ using the fitted regression model in Step 3.

- 6. Repeat Step 2 – Step 5 until convergence or satisfaction of a stopping criteria (could be a fixed number of iterations or small change of parameters).

# R lab

- Download the markdown code from course website

- Conduct the experiments

- Interpret the results

- Repeat the analysis on other datasets