

# LDRP Institute of Technology and Research, Gandhinagar

## Machine Learning: Tools, Techniques and Applications

Date: 01.05.2018

Prof. (Dr.) Mahesh M. Goyani

### Formulas:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$\text{Cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

### For Linear Regression:

Estimated response variable:  $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

### For Multiple Linear Regression:

Estimated response variable:  $\hat{Y} = X\hat{\beta}$

$$\hat{\beta} = (X'X)^{-1}X'Y$$

**Problems:**

Given the following vectors:

$$X = [-1.27, -0.63, 0.00, 0.63, 1.27];$$

$$Y = [-1.79, -0.90, 0.0, 0.90, 1.79];$$

**Ex. 1:** Find the correlation between X and Y

**Ex. 2:** Find the covariance between X and Y

Given the following vectors:

$$X = [340, 230, 405, 325, 280, 195, 265, 300, 350, 310]; \%sale$$

$$Y = [71, 65, 83, 74, 67, 56, 57, 78, 84, 65];$$

**Ex. 3:** Find the Linear Regression model for independent variable X and dependent variable Y.

**Ex. 4:** Predict the value of y for x = 250. Also find the residual for y<sub>4</sub>.

Given the following vectors:

$$x_1 = [41.9, 43.4, 43.9, 44.5, 47.3, 47.5, 47.9, 50.2, 52.8, 53.2, 56.7, 57.0, 63.5, 65.3, 71.1, 77.0, 77.8];$$

$$x_2 = [29.1, 29.3, 29.5, 29.7, 29.9, 30.3, 30.5, 30.7, 30.8, 30.9, 31.5, 31.7, 31.9, 32.0, 32.1, 32.5, 32.9];$$

$$y = [251.3, 251.3, 248.3, 267.5, 273.0, 276.5, 270.3, 274.9, 285.0, 290.0, 297.0, 302.5, 304.5, 309.3, 321.7, 330.7, 349.0];$$

**Ex. 5:** Find the Multiple Linear Regression model for independent variables x<sub>1</sub>, x<sub>2</sub> and dependent variable Y. Find the model parameters.

**Ex. 6:** Predict the value of y for x<sub>1</sub> = 47 & x<sub>2</sub> = 31. Also find the residual for y<sub>5</sub>.