

Data Exploration

11.02.2020

HW 01

$$X_i = \begin{bmatrix} X_i^1 \\ X_i^2 \\ \vdots \\ X_i^m \end{bmatrix} \quad i = \overline{1 \dots n}$$

Sample mean vector

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i - \text{[number]} - \text{vector}$$

Sample covariance matrix

$$V = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^T$$

T - transpose

$$\begin{bmatrix} X_i^1 \\ \vdots \\ X_i^m \end{bmatrix} \cdot [X_i^1 \dots X_i^m]$$

What canonical form of information would you suggest to represent the seq X_i in order to compute the Sample mean Vector and the sample covariance matrix?

Observations for Sample mean vector

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i;$$

The result \bar{X} depends on count of vectors, and the sum of received vectors. So, for this SNV, we need to store n , and the sum

(n, S)

$$\begin{bmatrix} X_1^1 \\ \vdots \\ X_1^m \end{bmatrix} + \begin{bmatrix} X_2^1 \\ \vdots \\ X_2^m \end{bmatrix}$$