INTRODUCTION TO LINEAR MODELS AND REGRESSION

Moutushi Chatterjee, notes by Ramdas Singh

Fourth Semester

Contents

Chapter 1

RANDOM VECTORS

1.1 Definitions and Moments

January 8th.

Recall that for a random variable X, the r-th *raw moment* is defined as

$$\mu_r' = \mathbb{E}[X^r] \tag{1.1}$$

provided the expectation exists. Here, $\mu_1' = \mathbb{E}[X] = \mu$ is the mean. The r-th *central moment* is defined as

$$\mu_r = \mathbb{E}[(X-\mu)^r]. \tag{1.2}$$

Similarly, for a set of sample points $x_1, x_2, \ldots, x_n$, the r-th *sample raw moment* is defined as

$$m_r' = \frac{1}{n}\sum_{i=1}^{n} x_i^r \tag{1.3}$$

and the r-th *sample central moment* is defined as

$$m_r = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^r \tag{1.4}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean. Note that $\mu_2 = \mathrm{Var}(X)$ is the variance of X and $\dfrac{\mu_3^2}{\mu_2^3}$ is the *skewness* of X. There is also *coefficient of variation* defined as $\dfrac{\sigma}{\mu}$ where $\sigma = \sqrt{\mathrm{Var}(X)}$ is the standard deviation of X. For a sample, it is defined as $\dfrac{s}{\bar{x}}$ where $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ is the sample standard deviation. The *correlation coefficient* between x_i's and y_i's is defined as

$$\frac{\frac{1}{n}\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum(x_i-\bar{x})^2 \cdot \frac{1}{n}\sum(y_i-\bar{y})^2}} \tag{1.5}$$

It is also known as the product-moment correlation coefficient.

These definitions can be extended to random vectors as follows. For a random vector $\mathbf{X} = (X_1, X_2, \ldots, X_p)^T$, the *mean vector* is defined as

$$\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}] = \begin{pmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_p] \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}. \tag{1.6}$$

The *dispersion matrix*, or covariance matrix, is defined as

$$\mathbf{\Sigma} = \mathbb{E}[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T] = \begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_p) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2,X_p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_p,X_1) & \mathrm{Cov}(X_p,X_2) & \cdots & \mathrm{Var}(X_p) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{pmatrix}. \tag{1.7}$$

Similar to the correlation coefficient for random variables, we define the *correlation matrix* for a random vector $\mathbf{X}$ as

$$\boldsymbol{R} = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{pmatrix} \tag{1.8}$$

Define D_σ as the diagonal matrix $\mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_p)$. Then, we have the relation

$$\boldsymbol{R} = D_\sigma^{-1}\mathbf{\Sigma}D_\sigma^{-1}. \tag{1.9}$$

If $\mathbf{Z} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$ is a random vector where $\mathbf{X}$ is of dimension p and $\mathbf{Y}$ is of dimension q, then we simply have

$$\mathbb{E}[\mathbf{Z}] = \begin{pmatrix} \mathbb{E}[\mathbf{X}] \\ \mathbb{E}[\mathbf{Y}] \end{pmatrix}, \quad \mathrm{Cov}(\mathbf{Z}) = \begin{pmatrix} \mathbf{\Sigma}_{\mathbf{X}} & \mathbf{\Sigma}_{\mathbf{XY}} \\ \mathbf{\Sigma}_{\mathbf{XY}}^\top & \mathbf{\Sigma}_{\mathbf{Y}} \end{pmatrix}. \tag{1.10}$$

As for linear transformations, if $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ where $\mathbf{A}$ is a $q \times p$ matrix and $\mathbf{b} \in \mathbb{R}^q$, then

$$\mathbb{E}[\mathbf{Y}] = \mathbf{A}\mathbb{E}[\mathbf{X}] + \mathbf{b}, \quad \mathrm{Cov}(\mathbf{Y}) = \mathbf{A}\,\mathrm{Cov}(\mathbf{X})\mathbf{A}^\top. \tag{1.11}$$

In particular, if $\mathbf{a} = \mathbf{A}^\top$ is a p-dimensional vector, then Y is a univariate random variable written as $Y = \mathbf{a}^\top\mathbf{X} + b$.

Index