Multivariate Data Analysis

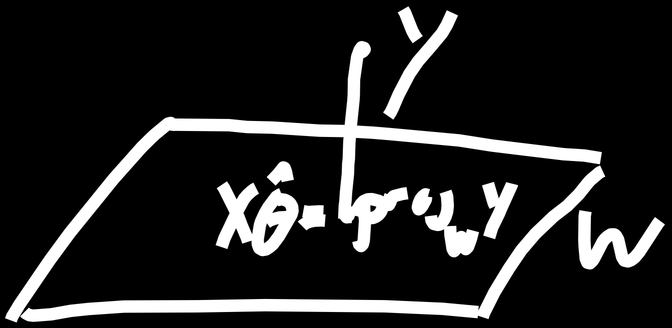
Jakob Söhl

$Y \in \mathbb{R}^n \qquad X \in \mathbb{R}^{n \times p} \qquad \theta \in \mathbb{R}^p \qquad \varepsilon \in \mathbb{R}^n$

$Y = X\theta + \varepsilon$

known known unknown

$W = \operatorname{Col} X$



$Y - X\hat{\theta} \perp \operatorname{Col} X$

$\Rightarrow (X_j)^T (y - X\hat{\theta}) = 0$ for all columns X_j of X

$\Rightarrow X^T (y - X\hat{\theta}) = 0$ $\quad X \in \mathbb{R}^{n \times p}$

$\Rightarrow X^T y = X^T X \hat{\theta}$

If X has full rank . $\hat{\theta} = (X^T X)^{-1} X^T y$

X full rank requires $p \leq n$

$$\mathrm{Cov}(y) = \mathbb{E}\left[(y-\mathbb{E}[y])(y-\mathbb{E}(y))^T\right]$$

$$= \mathbb{E}\left[y y^T - \mathbb{E}(y)\, y^T - y\, \mathbb{E}[y]^T + \mathbb{E}(y)\mathbb{E}[y]^T\right]$$

$$= \mathbb{E}[y y^T] - \mathbb{E}[y]\, \mathbb{E}[y]^T$$

$$X_2 = V \sim N(0,1) \qquad U, V \sim N(0,1)$$

$$\mathbb{E}[X_1] = \sqrt{1-\rho^2}\,\underbrace{\mathbb{E}[U]}_{=0} + \rho\,\underbrace{\mathbb{E}[V]}_{=0} = 0$$

$$\mathrm{Var}(X_1) = \mathrm{Var}\left(\sqrt{1-\rho^2}\,U\right) + 2\,\mathrm{Cov}\left(\sqrt{1-\rho^2}\,U, \rho V\right)$$

$$+ \mathrm{Var}(\rho V)$$

$$= (1-\rho^2)\underbrace{\mathrm{Var}(U)}_{=1} + 2\sqrt{1-\rho^2}\,\rho\,\underbrace{\mathrm{Cov}(U,V)}_{=0} + \rho^2\underbrace{\mathrm{Var}(V)}_{=1}$$

$$= 1-\rho^2+\rho^2 = 1$$

$X_2 \sim N(0,1)$

$\mathbb{E}[X_1] = 0$, $Var(X_1) = 1$, X_1 is normal as linear combination of normal random variables (RVs) U and V

$\Rightarrow X_1 \sim N(0,1)$

$Corr(X_1, X_2) = \rho$

X, Y normal RVs

X, Y independent $\iff$ X, Y uncorrelated

($\mathrm{Cov}(X, Y) = 0$)

For general RVs

X, Y independent $\implies$ X, Y uncorrelated

Converse is not true in general

Example $Z \sim N(0,1)$, $X = Z$, $Y = Z^2$

not independent, but uncorrelated

$$y = X\theta + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I)$$

$$\hat{\theta} = (X^T X)^{-1} X^T y = (X^T X)^{-1} X^T X \theta + (X^T X)^{-1} X^T \varepsilon = \theta + \underline{(X^T X)^{-1} X^T \varepsilon}$$

$$\mathbb{E}[\hat{\theta}] = (X^T X)^{-1} X^T \underbrace{\mathbb{E}[y]}_{= X\theta} = (X^T X)^{-1} (X^T X) \theta = \theta$$

$$Cov(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T]$$

$$= \mathbb{E}[(X^T X)^{-1} X^T \varepsilon \varepsilon^T X (X^T X)^{-1}]$$

$$= (X^T X)^{-1} X^T \underbrace{\mathbb{E}[\varepsilon \varepsilon^T]}_{= \sigma^2 I} X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}$$

$$\boxed{\hat{\theta} \sim N(\theta, \sigma^2 (X^T X)^{-1})}$$