# Chapter 7. Multivariate Linear Regression Models.

Multivariate linear regression model is essentially several univariate linear regression models putting together, with the errors being related with each. Here univariate (multivariate) means the response variables are univariate (multivariate). The notations of this chapter may be nasty.

# 7.1. Univariate linear regression — a review

(i). Model description:

$$Y_i = \beta_0 + \beta_1 z_{i1} + \dots + \beta_r z_{ir} + \epsilon_i \quad i = 1, \dots, n_i$$

where, for the *i*-th subject,  $Y_i$  is the response,  $z_{i1}, ..., z_{ik}$  are covariates,  $\beta_0, ..., \beta_r$  are regression parameters and  $\epsilon_i$  is the error.

An alternative presentation of the model in matrix form is:

$$\mathbf{Y}_{n\times 1} = \mathbf{Z}_{n\times(r+1)}\beta_{(r+1)\times 1} + \epsilon_{n\times 1}$$

where

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} 1 & z_{11} & \cdots & z_{1r} \\ 1 & z_{21} & \cdots & z_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_{n1} & \cdots & z_{nr} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_r \end{pmatrix} \text{ and } \epsilon = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

(ii). Model assumptions:

The errors  $\epsilon_1, ..., \epsilon_n$  are iid  $\sim N(0, \sigma^2)$  and the covariates are either non-random or are independent of the errors.

(iii). The least squares estimation (LSE):

The LSE of  $\beta$  is

$$\hat{\beta} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y} = \operatorname{argmin}\{b \in R^{r+1} : \|\mathbf{Y} - \mathbf{Z}b\|^2\}$$

which is unbiased, maximum likelihood estimator of  $\beta$ . A geometric interpretation is that  $\mathbf{Z}\hat{\beta}$  is the projection of  $\mathbf{Y}$  in  $\mathbb{R}^n$  onto the linear space  $\{\mathbf{Z}b \in \mathbb{R}^n : b \in \mathbb{R}^{r+1}\}$ . The residuals are

$$\hat{\epsilon} = \begin{pmatrix} \hat{\epsilon}_1 \\ \vdots \\ \hat{\epsilon}_n \end{pmatrix} = \mathbf{Y} - \mathbf{Z}\hat{\beta}.$$

Then, for i = 1, ..., n,

$$\hat{\epsilon}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 z_{i1} - \dots - \hat{\beta}_r z_{ir}$$

The estimator of  $\sigma^2$  is

$$s^2 \equiv \frac{1}{n-r-1}\hat{\epsilon}'\hat{\epsilon} = \frac{1}{n-r-1}\sum_{i=1}^n\hat{\epsilon}_i^2,$$

which is an unbiased estimator.

(iv). Inferences:

Observe that

$$\hat{\beta} = \beta + (\mathbf{Z}'\mathbf{Z})^{-1}Z'\epsilon$$

Inference is based on the facts that

$$\hat{\beta} - \beta \sim MN(0, (\mathbf{Z}'\mathbf{Z})^{-1}\sigma^2), \qquad \frac{1}{\sigma^2} \sum_{i=1}^n \epsilon_i^2 = (n-r-1)\frac{s^2}{\sigma^2} \sim \chi_{n-r-1}^2$$

and that  $\hat{\beta}$  is independent of  $s^2$ .

Confidence interval for  $\beta_k$  at confidence level  $1 - \alpha$ :

$$\hat{\beta}_k \pm t_{n-r-1} (\alpha/2) s \sqrt{a_{kk}}$$

where  $a_{lm}$ ,  $0 \le l, m \le r$  are the entries of the matrix  $(\mathbf{Z}'\mathbf{Z})^{-1}$ , i.e.,

$$(\mathbf{Z}'\mathbf{Z})^{-1} = \begin{pmatrix} a_{00} & a_{01} & \cdots & a_{0r} \\ a_{10} & a_{11} & \cdots & a_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r0} & a_{r1} & \cdots & a_{rr} \end{pmatrix}.$$

This confidence interval is based on

$$\frac{\hat{\beta}_k - \beta_k}{\sqrt{a_{kk}s}} \sim t_{n-r-1}, \qquad k = 0, ..., r.$$

One might also construct  $T^2$ -type confidence region at confidence level  $1 - \alpha$  for  $\beta$  as a vector of r + 1 dimension:

$$\left\{\beta \in R^{r+1} : (\hat{\beta} - \beta)'(\mathbf{Z}'\mathbf{Z})(\hat{\beta} - \beta)/s^2 \le (r+1)F_{r+1,n-r-1}(\alpha)\right\}$$

based on the fact

$$(\hat{\beta} - \beta)'(\mathbf{Z}'\mathbf{Z})(\hat{\beta} - \beta)/s^2 \sim (r+1)F_{r+1,n-r-1}$$

The  $T^2$ -type simultaneous confidence intervals for  $\beta_k, k = 0, ..., r$  at confidence level  $1 - \alpha$  are

for 
$$\beta_k$$
 :  $\hat{\beta}_k \pm F_{r+1,n-r-1}(\alpha) s \sqrt{a_{kk}}, \qquad k = 0, 1, ..., r$ 

Bonferroni's simultaneous confidence intervals for  $\beta_k, k = 0, ..., r$  at confidence level  $1 - \alpha$  are

for 
$$\beta_k$$
 :  $\hat{\beta}_k \pm t_{n-r-1}(\frac{\alpha}{2(r+1)})s\sqrt{a_{kk}}, \qquad k = 0, 1, ..., r$ 

(v). Prediction.

Let

$$Z_0 = \begin{pmatrix} 1\\ z_{01}\\ z_{02}\\ \vdots\\ z_{0r} \end{pmatrix}$$

be any given value of covariates. Let  $Y_0 = Z'_0\beta + \epsilon_0$  be the response of an experiment not yet carried out. We wish to predict the response  $Y_0$  or estimate its mean  $Z'_0\beta$  with accuracy justification. The predictor or estimator is naturally  $Z'_0\hat{\beta}$ , which is unbiased. Moreover,

$$Z'_{0}\hat{\beta} - Z'_{0}\beta \sim N\left(0, \ \sigma^{2}Z'_{0}(\mathbf{Z}'\mathbf{Z})^{-1}Z_{0}\right) \quad \text{and} \quad Z'_{0}\hat{\beta} - Y_{0} \sim N\left(0, \ \sigma^{2}[1 + Z'_{0}(\mathbf{Z}'\mathbf{Z})^{-1}Z_{0}]\right),$$

by noticing that

$$Z_0'\hat{\beta} - Z_0'\beta = Z_0'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}\epsilon \quad \text{and} \quad Z_0'\hat{\beta} - Y_0 = Z_0'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}\epsilon + \epsilon_0$$

Then, the prediction interval (or just call it confidence interval if you like) for  $Y_0$  at confidence level  $1 - \alpha$  is

$$Z'_0\hat{\beta} \pm t_{n-r-1}(\alpha/2)s\sqrt{1+Z'_0(\mathbf{Z}'\mathbf{Z})^{-1}Z_0}$$

And confidence interval for  $Z'_0\beta$ , the mean of  $Y_0$ , at confidence level  $1 - \alpha$  is

$$Z'_0 \hat{\beta} \pm t_{n-r-1} (\alpha/2) s \sqrt{Z'_0 (\mathbf{Z}' \mathbf{Z})^{-1} Z_0}$$

We note that this confidence interval still holds when the first component of  $Z_0$  takes any value, not just 1.

## 7.2. Multivariate linear regression.

**Example 1.1** (continued) TEN AMERICAN COMPANIES Data containing three variables: sales, profit and asset, and ten companies as observations is given in Appendix A. It is reasonable to believe that the size of business, measured by sales, and the profitability, measured by profit, both depends on the size of the company, measured by asset. One might consider two linear regression models. One is sales regressed on asset and the other is profit regressed on asset. As the errors of the two models shall inevitably be related, this presents an example of multivariate (actually bivariate here) linear regression model by putting together the two univariate linear regression models, this can be jointly presented as

$$Y_{i1} = \beta_{01} + \beta_{11}z_i + \epsilon_{i1}, \qquad i = 1, ..., 10$$
$$Y_{i2} = \beta_{02} + \beta_{12}z_i + \epsilon_{i2}, \qquad i = 1, ..., 10$$

where  $Y_{i1}, Y_{i2}, z_i$  are sales, profit and asset of company *i*.

Another simple example is a multivariate linear regression model relating students academic performance measured by midterm and final exam scores with diligence measured by weekly time devoted to study and ingenuity measured by IQ score:

$$\begin{cases} Y_{i1} = \beta_{01} + \beta_{11} z_{i1} + \beta_{21} z_{i2} + \epsilon_{i1}, \\ Y_{i2} = \beta_{02} + \beta_{12} z_{i1} + \beta_{22} z_{i2} + \epsilon_{i2} \end{cases} \qquad i = 1, ..., n \end{cases}$$

where  $Y_{i1}$ ,  $Y_{i2}$   $z_{i1}$  and  $z_{i2}$  are, respectively, midterm exam score, final exam score, weekly hours of time devoted to study and IQ score of the *i*-th randomly selected student.

(i). Model description:

The *k*-th univariate linear regression model:

$$Y_{ik} = \beta_{0k} + z_{i1}\beta_{1k} + \dots + z_{ir}\beta_{rk} + \epsilon_{ik}, \qquad i = 1, \dots, n,$$

where  $Y_{ik}$  is the k-th response for subject i. Writing in matrix form:

$$\mathbf{Y}_{(k)} = \mathbf{Z}\beta_{(k)} + \epsilon_{(k)},$$

where

$$\mathbf{Y}_{(k)} = \begin{pmatrix} Y_{1k} \\ \vdots \\ Y_{nk} \end{pmatrix}, \quad \beta_{(k)} = \begin{pmatrix} \beta_{0k} \\ \vdots \\ \beta_{rk} \end{pmatrix} \quad \text{and} \quad \epsilon_{(k)} = \begin{pmatrix} \epsilon_{1k} \\ \vdots \\ \epsilon_{nk} \end{pmatrix}$$

The so called multivariate linear regression model is to put the m univariate linear regression model together as

$$\mathbf{Y}_{n \times m} = \mathbf{Z}_{n \times (r+1)} \beta_{(r+1) \times m} + \epsilon_{n \times m},$$

which is the same as

$$[\mathbf{Y}_{(1)} \vdots \cdots \vdots \mathbf{Y}_{(m)}] = \mathbf{Z}[\beta_{(1)} \vdots \cdots \vdots \beta_{(m)}] + [\epsilon_{(1)} \vdots \cdots \vdots \epsilon_{(m)}]$$

One distinct feature here is that the errors,  $\epsilon_{i1}, ..., \epsilon_{im}$ , within one observation/subject may be related with each other. This suggest that the *m* univariate linear regression models should be treated together as one model rather than separately treated as *m* individual univariate models.

(ii). Model assumptions:

The errors  $(\epsilon_{i1}, ..., \epsilon_{im})'$ , as random vectors of *m* dimension, are iid ~  $MN(0, \Sigma)$  and are independent of the covariates. Note that here  $\Sigma$  is an  $m \times m$  matrix.

#### (iii). Estimation:

The least squares criterion for the multivariate linear regression model is to minimize over  $\beta$ 

$$\sum_{k=1}^{m} \sum_{i=1}^{n} \left[ Y_{ik} - (\beta_{0k} + z_{i1}\beta_{1k} + \dots + z_{ir}\beta_{rk}) \right]^2 = \sum_{k=1}^{m} \|\mathbf{Y}_{(k)} - \mathbf{Z}\beta_{(k)}\|^2.$$

The LSE of  $\beta$  is denoted as

$$\hat{\beta} = [\hat{\beta}_{(1)} \vdots \cdots \vdots \hat{\beta}_{(m)}].$$

Notice that

$$\min_{\beta} \sum_{k=1}^{m} \|\mathbf{Y}_{(k)} - \mathbf{Z}\beta_{(k)}\|^2 \sum_{k=1}^{m} \min_{\beta_{(k)}} \|\mathbf{Y}_{(k)} - \mathbf{Z}\beta_{(k)}\|^2,$$

i.e.,

$$\min_{\beta} \sum_{k=1}^{m} \sum_{i=1}^{n} (Y_{ik} - [\beta_{0k} + \sum_{l=1}^{r} z_{il}\beta_{lk}])^2 = \sum_{k=1}^{m} \left[ \min_{\beta_{(k)}} \sum_{i=1}^{n} (Y_{ik} - [\beta_{0k} + \sum_{l=1}^{r} z_{il}\beta_{lk}])^2 \right].$$

Hence  $\hat{\beta}_{(k)}$  is the same as the LSE of  $\beta_{(k)}$  of the univariate linear regression model:  $\mathbf{Y}_{(k)} = \mathbf{Z}\beta_{(k)} + \epsilon_{(k)}$ . Therefore,

$$\begin{aligned} \boldsymbol{\beta}_{(k)} &= (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_{(k)}, \qquad k = 1, ..., m, \quad and \\ \hat{\boldsymbol{\beta}} &= [\hat{\boldsymbol{\beta}}_{(1)} \vdots \cdots \vdots \hat{\boldsymbol{\beta}}_{(m)}] = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'[\mathbf{Y}_{(1)} \vdots \cdots \vdots \mathbf{Y}_{(m)}]\mathbf{Y} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y} \end{aligned}$$

which is formally the same as the expression of the LSE of univariate linear regression model. The residuals are

$$\hat{\epsilon}_{n \times m} = \mathbf{Y}_{n \times m} - \mathbf{Z}_{n \times (r+1)} \hat{\beta}_{(r+1) \times m}.$$

and the estimator of  $\Sigma$  is

$$\mathbf{S} \equiv \begin{pmatrix} s_{11} & \cdots & s_{1m} \\ \vdots & \vdots & \vdots \\ s_{m1} & \cdots & s_{mm} \end{pmatrix} = \frac{1}{n-r-1} \hat{\epsilon}' \hat{\epsilon}.$$

### (iv). Inference with $\beta_{(k)}$ .

It is often of interest to consider inference with regression parameter, say  $\beta_{(k)}$ , for one single univariate regression model. All theory, estimation and inference procedures with univariate regression model as presented in Section 7.1 apply here. For example,

$$\hat{\beta}_{(k)} - \beta_{(k)} \sim MN(0, \sigma_{kk} (\mathbf{Z}'\mathbf{Z})^{-1})$$

where  $\sigma_{kl}$  is the (k, l)-th entry of  $\Sigma$ .

$$\frac{(\hat{\beta}_{(k)} - \beta_{(k)})(\mathbf{Z}'\mathbf{Z})(\hat{\beta}_{(k)} - \beta_{(k)})}{s_{kk}} \sim (r+1)F_{r+1,n-r-1}$$

where  $s_{kl}$  is the (k, l)-th entry of **S**. Then, a confidence region at confidence level  $1 - \alpha$  for  $\beta_{(k)}$ , as a vector in  $\mathbb{R}^{r+1}$ , can be constructed as

$$\Big\{\beta_{(k)} \in R^{r+1} : \frac{(\hat{\beta}_{(k)} - \beta_{(k)})(\mathbf{Z}'\mathbf{Z})(\hat{\beta}_{(k)} - \beta_{(k)})}{s_{kk}} \sim (r+1)F_{r+1,n-r-1}(\alpha)\Big\}.$$

For more detail, such as confidence intervals for  $\beta_{jk}$ , simultaneous confidence intervals for  $\beta_{0k}, ..., \beta_{rk}$ , please refer to Section 7.1 about univariate linear regression models.

(v). Prediction.

Let

$$Y_0 = \beta' Z_0 + \epsilon_0 = \beta' \begin{pmatrix} 1\\z_{01}\\\vdots\\z_{0r} \end{pmatrix} + \begin{pmatrix} \epsilon_{01}\\\vdots\\\epsilon_{0m} \end{pmatrix}$$

be a new experiment with covariate  $Z_0$  which has not been carried out. We are interested in predicting its response  $Y_0$ , an *m*-vector, or estimating its mean  $\beta' Z_0$ , with accuracy justification. The predictor of  $Y_0$  or estimator of its mean  $\beta' Z_0$  is naturally  $\hat{\beta}' Z_0$ . To compute the variance, observe that, for any non-random *n*-vector  $V = (v_1, ..., v_n)'$ ,

$$\operatorname{var}(\epsilon' V) = \operatorname{var}(\sum_{i=1}^{n} v_i \epsilon_i) = \sum_{i=1}^{n} v_i^2 \operatorname{var}(\epsilon_i) = \sum_{j=1}^{n} v_i^2 = \|V\|^2 \Sigma = V' V \Sigma,$$

where  $\epsilon_i$ , an *m*-vector, is the *i*-row of  $\epsilon$  representing the *m* errors for the *i*-th subject/observation. Let  $V = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}Z_0$ . Then,

$$\operatorname{var}(Y_0 - \hat{\beta}' Z_0) = \operatorname{var}(\epsilon_0 - (\hat{\beta} - \beta)' Z_0) = \operatorname{var}(\epsilon_0) + \operatorname{var}((\hat{\beta} - \beta)' Z_0)$$
$$= \Sigma + \operatorname{var}([(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\epsilon]' Z_0) = \Sigma + \operatorname{var}(\epsilon'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} Z_0) = \Sigma + [Z_0'(\mathbf{Z}'\mathbf{Z})^{-1} Z_0]\Sigma$$

For simplicity of notation, let

$$\kappa = Z_0' (\mathbf{Z}'\mathbf{Z})^{-1} Z_0.$$

In summary

$$\begin{aligned} \operatorname{var}(Y_0 - \hat{\beta}_0' Z_0) &= \Sigma(1+\kappa), \quad \operatorname{var}(\beta' Z_0 - \hat{\beta}_0' Z_0) = \kappa \Sigma \\ &\frac{1}{\sqrt{1+\kappa}} \Sigma^{-1/2} (Y_0 - \hat{\beta}_0' Z_0) \sim MN(0, I_m) \end{aligned}$$
  
and 
$$\begin{aligned} &\frac{1}{\sqrt{\kappa}} \Sigma^{-1/2} (\beta' Z_0 - \hat{\beta}_0' Z_0) \sim MN(0, I_m). \end{aligned}$$

Moreover,

$$T^{2} \equiv \frac{1}{\kappa} [\hat{\beta}' Z_{0} - \beta' Z_{0}]' \mathbf{S}^{-1} [\hat{\beta}' Z_{0} - \beta' Z_{0}] \sim \frac{(n-r-1)m}{n-r-m} F_{m,n-r-m}$$
  
and 
$$\frac{1}{1+\kappa} [Y_{0} - \beta' Z_{0}]' \mathbf{S}^{-1} [Y_{0} - \beta' Z_{0}] \sim \frac{(n-r-1)m}{n-r-m} F_{m,n-r-m},$$

which provide theoretical foundation for the following inference procedures: (1). Confidence region for  $\beta' Z_0 = E(Y_0)$  at level  $1 - \alpha$ :

$$\left\{\beta' Z_0 \in R^m : [\hat{\beta}' Z_0 - \beta' Z_0]' \mathbf{S}^{-1} [\hat{\beta}' Z_0 - \beta' Z_0] \le \kappa \frac{(n-r-1)m}{n-r-m} F_{m,n-r-m}(\alpha)\right\}$$

(2). Prediction region (or just call it confidence region) for  $Y_0$  at level  $1 - \alpha$ :

$$\Big\{Y_0 \in R^m : [Y_0 - \beta' Z_0]' \mathbf{S}^{-1} [Y_0 - \beta' Z_0] \sim (1 + \kappa) \frac{(n - r - 1)m}{n - r - m} F_{m, n - r - m}(\alpha)\Big\}.$$

(3). Simultaneous confidence intervals for  $\beta'_{(k)}Z_0 = E(Y_{0k}), k = 1, ..., m$  at confidence level  $1 - \alpha$ :

for 
$$\beta'_{(k)}Z_0$$
:  $\hat{\beta}'_{(k)}Z_0 \pm \sqrt{s_{kk} \cdot \kappa \frac{(n-r-1)m}{n-r-m}F_{m,n-r-m}(\alpha)}$   $k = 1, ..., m$ 

30

(4). Simultaneous prediction/confidence intervals for  $Y_{0k}$ , k = 1, ..., m at confidence level  $1 - \alpha$ :

for 
$$Y_{0k}$$
:  $\hat{\beta}'_{(k)}Z_0 \pm \sqrt{s_{kk}(1+\kappa)\frac{(n-r-1)m}{n-r-m}}F_{m,n-r-m}(\alpha)$   $k = 1,...,m.$ 

**Example 1.1** (continued) TEN AMERICAN COMPANIES Sales  $(Y_1)$  and profit  $(Y_2)$  are two responses regressed on the covariate asset (Z). The multivariate linear regression is the two univariate linear regression models putting together:

$$Y_{i1} = \beta_{01} + \beta_{11}z_i + \epsilon_{i1}, \qquad i = 1, ..., 10$$
  
$$Y_{i2} = \beta_{02} + \beta_{12}z_i + \epsilon_{i2}, \qquad i = 1, ..., 10$$

An expression in matrix form is

$$\begin{pmatrix} Y_{11} & Y_{12} \\ \vdots & \vdots \\ Y_{n1} & Y_{n2} \end{pmatrix} = \begin{pmatrix} 1 & z_1 \\ \vdots & \vdots \\ 1 & z_n \end{pmatrix} \begin{pmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \end{pmatrix} + \begin{pmatrix} \epsilon_{11} & \epsilon_{12} \\ \vdots & \vdots \\ \epsilon_{n1} & \epsilon_{n2} \end{pmatrix}.$$
$$\mathbf{Y} = \mathbf{Z}\beta + \epsilon.$$

In the notations used above, n = 10, m = 2 and r = 1. And

$$\mathbf{Z}'\mathbf{Z} = \begin{pmatrix} 10 & 812.484\\ 812.484 & 92837.4 \end{pmatrix} \text{ and } (\mathbf{Z}'\mathbf{Z})^{-1} = \begin{pmatrix} 0.34609 & -0.00303\\ -0.00303 & 0.0000373 \end{pmatrix}$$

(1). The least squares estimator of  $\beta$ :

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_{01} & \hat{\beta}_{02} \\ \hat{\beta}_{11} & \hat{\beta}_{12} \end{pmatrix} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y} = \begin{pmatrix} 21.09 & 1.683 \\ 0.507 & 0.015 \end{pmatrix}$$

(2) Variance estimation:

The estimator of  $\Sigma$ , the 2 × 2 matrix of the variance of the error, is

$$\mathbf{S} = \begin{pmatrix} 262.85 & 2.72\\ 2.72 & 0.822 \end{pmatrix}.$$

It's easy to compute

$$\mathbf{S}^{-1} = \begin{pmatrix} 0.00394 & -0.01304 \\ -0.01304 & 1.2597 \end{pmatrix}.$$

(3) Inference with regression parameters of one univariate regression model.

Inferences about  $\beta_{01}$ ,  $\beta_{11}$ ,  $\beta_{(1)} = (\beta_{01}, \beta_{11})'$ , predicting sales for a given company with known asset size or estimating mean sales for companies with a given size of asset are all only related with the first *univariate* linear regression model. Likewise, inferences about  $\beta_{02}$ ,  $\beta_{12}$ ,  $\beta_{(2)} = (\beta_{02}, \beta_{12})'$ , predicting profit for a given company with known asset size or estimating mean profit for companies with a given size of asset are all only related with the second *univariate* linear regression model. Then, inference methods reviewed in Section 7.1 apply. For example, a confidence interval at confidence level 95% for  $\beta_{11}$  is

$$\hat{\beta}_{11} \pm t_{10-2}(0.025)\sqrt{0.0000373 \times s_{11}} = 0.507 \pm t_8(.025)\sqrt{.0000373 \times 262.85} = .507 \pm .22835$$

and a confidence region at confidence level 95% for  $\beta_{(1)} = (\beta_{01}, \beta_{11})'$  is

$$\{\beta_{(1)} \in R^2 : (\hat{\beta}_{(1)} - \beta_{(1)})'(\mathbf{Z}'\mathbf{Z})(\hat{\beta}_{(1)} - \beta_{(1)}) \le 2344\}.$$

(4). Prediction of both sales and profit.

Suppose company ABC has 50 billion dollars of asset. We wish to predict both sales and profit of the company with accuracy justification. Set  $Z_0 = (1 \ 50)'$  and denote by  $Y_{01}$  and  $Y_{02}$  the yet unknown sales and profit of the company. The estimator is

$$\hat{\beta}' Z_0 = \begin{pmatrix} \hat{\beta}_{01} + 50\hat{\beta}_{11} \\ \hat{\beta}_{02} + 50\hat{\beta}_{12} \end{pmatrix} = \begin{pmatrix} 46.45 \\ 2.45 \end{pmatrix}$$

Note that  $\kappa = 0.1364$ . A prediction region at confidence level 95% is

$$\left\{ (S, P) : \left[ \begin{pmatrix} S \\ P \end{pmatrix} - \begin{pmatrix} 46.45 \\ 2.45 \end{pmatrix} \right]' \mathbf{S}^{-1} \left[ \begin{pmatrix} S \\ P \end{pmatrix} - \begin{pmatrix} 46.45 \\ 2.45 \end{pmatrix} \right] \le 12.31 \right\}$$

where S stands for sales and P for profit of the company. And simultaneous prediction intervals at (nominal) confidence level 95% is

for sales : 
$$46.45 \pm 58.87$$
  
for profit :  $2.45 \pm 3.18$ .

On the other hand, suppose we are interested in providing an estimation and inference for the mean sales and profit for all companies with asset 50 billion dollars, which is

$$\begin{pmatrix} EY_{01} \\ EY_{02} \end{pmatrix} = \begin{pmatrix} \beta_{01} + 50\beta_{11} \\ \beta_{02} + 50\beta_{12} \end{pmatrix}$$

The estimator is still same as (46.45, 2.45)'. A confidence region at confidence level 95% is

$$\left\{ (S, P) : \left[ \begin{pmatrix} S \\ P \end{pmatrix} - \begin{pmatrix} 46.45 \\ 2.45 \end{pmatrix} \right]' \mathbf{S}^{-1} \left[ \begin{pmatrix} S \\ P \end{pmatrix} - \begin{pmatrix} 46.45 \\ 2.45 \end{pmatrix} \right] \le 1.477 \right\}$$

And simultaneous confidence intervals at (nominal) confidence level 95% is

for mean sales :  $46.45 \pm 19.70$ for mean profit :  $2.45 \pm 1.102$ .