Multivariate Ordinary Least Squares Estimator

Gary Schurman MBE, CFA

October, 2012

We are tasked with estimating the unknown value of the dependent variable x based upon the known values of two explanatory variables y and z. We will define the actual value of the dependent variable x to be a function of the estimated value of x plus an error term. We will define $\hat{x}$ to be the estimated value of the dependent variable x. The equation for the estimated value of x is...

$$\hat{x} = \alpha + \beta_1 \, y + \beta_2 \, z \tag{1}$$

The actual value of the dependent variable x is the estimated value of x plus an error term. Using Equation (1) above the equation for the actual value of x where ϵ is the error term is...

$$\begin{aligned} x &= \hat{x} + \epsilon \\ &= \alpha + \beta_1 \, y + \beta_2 \, z + \epsilon \end{aligned} \tag{2}$$

We will develop the mathematics for the multivariate regression equation via Ordinary Least Squares (OLS) using the following hypothetical problem...

Our Hypothetical Problem

Our client owns a meat packing plant in the United States. We are tasked with developing a materials cost equation given the dollar value of annual sales and the dollar value of ending inventory. Using Equation (2) above as our guide the materials cost equation will take the following form...

$$\text{Materials Cost} = \alpha + \beta_1 \, \text{Sales} + \beta_2 \, \text{Ending Inventory} + \epsilon \tag{3}$$

The Appendix includes actual data for the U.S. Meat Packing Industry for the years 1958 through 1996. The time series data, means and covariances are presented in Appendix A, B and C, repectively. Our multivariate regression equation variables are as follows...

Table 1 - Regression Equation Variables

Variable	Description
x	Annual materials cost (in millions of dollars)
y	Annual sales (in millions of dollars)
z	End of year inventory (in millions of dollars)
α	Regression constant (in millions of dollars)
β_1	Regression coefficient for annual sales
β_2	Regression coefficient for ending inventory
ϵ	Error term

Question: Given that our client estimates annual sales to be \$50 million and ending inventory to be \$5 million what is the OLS estimate of annual materials costs?

The Equations for the Sum of Squared Errors and Derivatives

The estimation error is the difference between the actual value of the dependent variable and the estimated value of the dependent variable. Using Equations (1) and (2) above the equation for the estimation error is...

$$\begin{aligned} \epsilon &= x - \hat{x} \\ &= x - (\alpha + \beta_1 \, y + \beta_2 \, z) \end{aligned} \tag{4}$$

Using Equation (4) above the equation for the sum of squared errors (SSE) where the variable N is the number of observations is...

$$\begin{aligned}
SSE &= \sum_{i=1}^{N}\Bigg[\epsilon_i^2\Bigg] \\
&= \sum_{i=1}^{N}\Bigg[\Bigg\{x_i-(\alpha+\beta_1 y_i+\beta_2 z_i)\Bigg\}^2\Bigg] \\
&= \sum_{i=1}^{N}\Bigg[x_i^2+\alpha^2-2\alpha x_i+2\alpha\beta_1 y_i+2\alpha\beta_2 z_i-2\beta_1 x_i y_i-2\beta_2 x_i z_i+2\beta_1\beta_2 y_i z_i+\beta_1^2 y_i^2+\beta_2^2 z_i^2\Bigg] \\
&= \sum_{i=1}^{N}\Bigg[x_i^2\Bigg]+N\alpha^2-2\alpha\sum_{i=1}^{N}\Bigg[x_i\Bigg]+2\alpha\beta_1\sum_{i=1}^{N}\Bigg[y_i\Bigg]+2\alpha\beta_2\sum_{i=1}^{N}\Bigg[z_i\Bigg]-2\beta_1\sum_{i=1}^{N}\Bigg[x_i y_i\Bigg]-2\beta_2\sum_{i=1}^{N}\Bigg[x_i z_i\Bigg] \\
&+2\beta_1\beta_2\sum_{i=1}^{N}\Bigg[y_i z_i\Bigg]+\beta_1^2\sum_{i=1}^{N}\Bigg[y_i^2\Bigg]+\beta_2^2\sum_{i=1}^{N}\Bigg[z_i^2\Bigg] \qquad (5)
\end{aligned}$$

We want to set the values of regression Equation (3) parameters α, β_1 and β_2 such that the sum of squared errors as defined by Equation (5) is minimized. To minimize the sum of squared errors we need the derivatives of equation (5) with respect to α, β_1 and β_2. The plan is to set each derivative equation equal to zero and simultaneously solve for α, β_1 and β_2.

The derivative of equation (5) with respect to α is...

$$\begin{aligned}
\frac{\delta SSE}{\delta\alpha} &= 2N\alpha-2\sum_{i=1}^{N}\Bigg[x_i\Bigg]+2\beta_1\sum_{i=1}^{N}\Bigg[y_i\Bigg]+2\beta_2\sum_{i=1}^{N}\Bigg[z_i\Bigg] \\
&= 2N\alpha-2N\sum_{i=1}^{N}\Bigg[\frac{1}{N}x_i\Bigg]+2N\beta_1\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i\Bigg]+2N\beta_2\sum_{i=1}^{N}\Bigg[\frac{1}{N}z_i\Bigg] \\
&= 2N\alpha-2N\bar{x}+2N\beta_1\bar{y}+2N\beta_2\bar{z} \\
&= 2N\Bigg\{\alpha-\bar{x}+\beta_1\bar{y}+\beta_2\bar{z}\Bigg\} \qquad (6)
\end{aligned}$$

Note that in Equation (6) above $\bar{x}$ is the mean of the observed values of the dependent variable x, $\bar{y}$ is the mean of the observed values of the independent variable y and $\bar{z}$ is the mean of the observed values of the independent variable z.

The derivative of equation (5) with respect to β_1 is...

$$\begin{aligned}
\frac{\delta SSE}{\delta\beta_1} &= 2\alpha\sum_{i=1}^{N}\Bigg[y_i\Bigg]-2\sum_{i=1}^{N}\Bigg[x_i y_i\Bigg]+2\beta_2\sum_{i=1}^{N}\Bigg[y_i z_i\Bigg]+2\beta_1\sum_{i=1}^{N}\Bigg[y_i^2\Bigg] \\
&= 2N\alpha\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i\Bigg]-2N\sum_{i=1}^{N}\Bigg[\frac{1}{N}x_i y_i\Bigg]+2N\beta_2\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i z_i\Bigg]+2N\beta_1\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i^2\Bigg] \\
&= 2N\Bigg\{\alpha\bar{y}-\sum_{i=1}^{N}\Bigg[\frac{1}{N}x_i y_i\Bigg]+\beta_2\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i z_i\Bigg]+\beta_1\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i^2\Bigg]\Bigg\} \qquad (7)
\end{aligned}$$

The derivative of equation (5) with respect to β_2 is...

$$\begin{aligned}
\frac{\delta SSE}{\delta\beta_2} &= 2\alpha\sum_{i=1}^{N}\Bigg[z_i\Bigg]-2\sum_{i=1}^{N}\Bigg[x_i z_i\Bigg]+2\beta_1\sum_{i=1}^{N}\Bigg[y_i z_i\Bigg]+2\beta_2\sum_{i=1}^{N}\Bigg[z_i^2\Bigg] \\
&= 2N\alpha\sum_{i=1}^{N}\Bigg[\frac{1}{N}z_i\Bigg]-2N\sum_{i=1}^{N}\Bigg[\frac{1}{N}x_i z_i\Bigg]+2N\beta_1\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i z_i\Bigg]+2N\beta_2\sum_{i=1}^{N}\Bigg[\frac{1}{N}z_i^2\Bigg] \\
&= 2N\Bigg\{\alpha\bar{z}-\sum_{i=1}^{N}\Bigg[\frac{1}{N}x_i z_i\Bigg]+\beta_1\sum_{i=1}^{N}\Bigg[\frac{1}{N}y_i z_i\Bigg]+\beta_2\sum_{i=1}^{N}\Bigg[\frac{1}{N}z_i^2\Bigg]\Bigg\} \qquad (8)
\end{aligned}$$

The Equations for the Regression Coefficients

We want to minimize the sum of squared errors (SSE) so to solve for α, β_1 and β_2 in our regression Equation (3) we will set the derivative equations (6), (7) and (8) equal to zero and jointly solve for the three parameter values.

Setting derivative Equation (6) equal to zero the value of α is...

$$\begin{aligned}\frac{\delta SSE}{\delta\alpha} &= 0\\ 2N\left\{\alpha - \bar{x} + \beta_1\bar{y} + \beta_2\bar{z}\right\} &= 0\\ \alpha &= \bar{x} - \beta_1\bar{y} - \beta_2\bar{z} \end{aligned} \tag{9}$$

To calculate the variance of the estimation error we will need an equation for α^2, which is...

$$\begin{aligned}\alpha^2 &= (\bar{x} - \beta_1\bar{y} - \beta_2\bar{z})^2\\ &= \bar{x}^2 - 2\beta_1\bar{x}\bar{y} - 2\beta_2\bar{x}\bar{z} + 2\beta_1^2\beta_2^2\bar{y}\bar{z} + \beta_1^2\bar{y}^2 + \beta_2^2\bar{z}^2\end{aligned} \tag{10}$$

Setting derivative Equation (7) equal to zero and using the value of α in equation (9) the equation for β_1 that minimizes the sum of squared errors is...

$$\begin{aligned}\frac{\delta SSE}{\delta\beta_1} &= 0\\ 2N\left\{\alpha\bar{y} - \sum_{i=1}^{N}\left[\frac{1}{N}x_iy_i\right] + \beta_2\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] + \beta_1\sum_{i=1}^{N}\left[\frac{1}{N}y_i^2\right]\right\} &= 0\\ 2N\left\{\left(\bar{x} - \beta_1\bar{y} - \beta_2\bar{z}\right)\bar{y} - \sum_{i=1}^{N}\left[\frac{1}{N}x_iy_i\right] + \beta_2\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] + \beta_1\sum_{i=1}^{N}\left[\frac{1}{N}y_i^2\right]\right\} &= 0\\ \bar{x}\bar{y} - \beta_1\bar{y}^2 - \beta_2\bar{y}\bar{z} - \sum_{i=1}^{N}\left[\frac{1}{N}x_iy_i\right] + \beta_2\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] + \beta_1\sum_{i=1}^{N}\left[\frac{1}{N}y_i^2\right] &= 0\\ \beta_1\left\{\sum_{i=1}^{N}\left[\frac{1}{N}y_i^2\right] - \bar{y}^2\right\} + \beta_2\left\{\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] - \bar{y}\bar{z}\right\} - \left\{\sum_{i=1}^{N}\left[\frac{1}{N}x_iy_i\right] - \bar{x}\bar{y}\right\} &= 0\\ \beta_1Var(y) + \beta_2Cov(y,z) - Cov(x,y) &= 0\end{aligned} \tag{11}$$

Setting derivative Equation (8) equal to zero and using the value of α in equation (9) the equation for β_2 that minimizes the sum of squared errors is...

$$\begin{aligned}\frac{\delta SSE}{\delta\beta_2} &= 0\\ 2N\left\{\alpha\bar{z} - \sum_{i=1}^{N}\left[\frac{1}{N}x_iz_i\right] + \beta_1\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] + \beta_2\sum_{i=1}^{N}\left[\frac{1}{N}z_i^2\right]\right\} &= 0\\ 2N\left\{\left(\bar{x} - \beta_1\bar{y} - \beta_2\bar{z}\right)\bar{z} - \sum_{i=1}^{N}\left[\frac{1}{N}x_iz_i\right] + \beta_1\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] + \beta_2\sum_{i=1}^{N}\left[\frac{1}{N}z_i^2\right]\right\} &= 0\\ \bar{x}\bar{z} - \beta_1\bar{y}\bar{z} - \beta_2\bar{z}^2 - \sum_{i=1}^{N}\left[\frac{1}{N}x_iz_i\right] + \beta_1\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] + \beta_2\sum_{i=1}^{N}\left[\frac{1}{N}z_i^2\right] &= 0\\ \beta_1\left\{\sum_{i=1}^{N}\left[\frac{1}{N}y_iz_i\right] - \bar{y}\bar{z}\right\} + \beta_2\left\{\sum_{i=1}^{N}\left[\frac{1}{N}z_i^2\right] - \bar{z}^2\right\} - \left\{\sum_{i=1}^{N}\left[\frac{1}{N}x_iz_i\right] - \bar{x}\bar{z}\right\} &= 0\\ \beta_2Var(z) + \beta_1Cov(y,z) - Cov(x,z) &= 0\end{aligned} \tag{12}$$

Solving for the Regression Equation Coefficients

We can rewrite Equations (11) and (12), respectively, as...

$$\beta_1 Var(y) + \beta_2 Cov(y,z) = Cov(x,y) \tag{13}$$
$$\beta_2 Var(z) + \beta_1 Cov(y,z) = Cov(x,z) \tag{14}$$

In the system of linear equations above we have two equations and two unknowns. We can solve for β_1 and β_2 using simple Algebra but we want to develop a method of solving for the betas when there are more than two equations and two unknowns. We know from Linear Algebra that we can represent any system of N linear equations with N unknowns as the following matrix:vector product...

$$\mathbf{A}\,\vec{\mathbf{b}} = \vec{\mathbf{c}} \tag{15}$$

...where the matrix $\mathbf{A}$ is an N by N square matrix and the vectors $\vec{\mathbf{b}}$ and $\vec{\mathbf{c}}$ are both row vectors with N rows. For our purposes the relevant matrix:vector product equation will take the following form...

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{16}$$

We will define matrix $\mathbf{A}$ to be a covariance matrix where the element $a_{1,1}$ is the variance of the independent variable y, the element $a_{2,2}$ is the variance of the independent variable z and the elements $a_{1,2}$ and $a_{2,1}$ are the covariances of the independent variables y and z. Our matrix $\mathbf{A}$ is therefore...

$$\mathbf{A} = \begin{bmatrix} Var(y) & Cov(y,z) \\ Cov(y,z) & Var(z) \end{bmatrix} \tag{17}$$

We will define vector $\vec{\mathbf{b}}$ to be a row vector of beta coefficients where element b_1 is the beta coefficient for the independent variable y and element b_2 is the beta coefficient for the independent variable z. Our vector $\vec{\mathbf{b}}$ is therefore...

$$\vec{\mathbf{b}} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \tag{18}$$

We will define vector $\vec{\mathbf{c}}$ to be a row vector of covariances where element c_1 is the covariance of the independent variable y with the dependent variable x and the element c_2 is the covariance of the independent variable z with the dependent variable x. Our vector $\vec{\mathbf{c}}$ is therefore...

$$\vec{\mathbf{c}} = \begin{bmatrix} Cov(y,x) \\ Cov(z,x) \end{bmatrix} \tag{19}$$

Using Equations (17), (18) and (19) above we can represent the system of linear equations (13) and (14) above as...

$$\begin{bmatrix} Var(y) & Cov(y,z) \\ Cov(y,z) & Var(z) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} Cov(y,x) \\ Cov(z,x) \end{bmatrix} \tag{20}$$

The row vector of beta coefficients is therefore the matrix:vector product of the inverse of matrix $\mathbf{A}$ and vector $\vec{\mathbf{c}}$. Using the time series data from Appendix A, B and C the beta coefficients for our regression equation are...

$$\begin{aligned} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} &= \begin{bmatrix} Var(y) & Cov(y,z) \\ Cov(y,z) & Var(z) \end{bmatrix}^{-1} \begin{bmatrix} Cov(y,x) \\ Cov(z,x) \end{bmatrix} \\ &= \begin{bmatrix} 225859160 & 3800172 \\ 3800172 & 66545 \end{bmatrix}^{-1} \begin{bmatrix} 196428819 \\ 3289338 \end{bmatrix} \\ &= \begin{bmatrix} 0.000000113 & -0.000006458 \\ -0.000006458 & 0.000383827 \end{bmatrix} \begin{bmatrix} 196428819 \\ 3289338 \end{bmatrix} \\ &= \begin{bmatrix} 0.9708 \\ -6.0097 \end{bmatrix} \end{aligned} \tag{21}$$

Per Equation (21) above the regression coefficients β_1 and β_2 are...

$$\beta_1 = 0.9708 \tag{22}$$
$$\beta_2 = -6.0097 \tag{23}$$

Using Equations (9), (22) and (23) and the time series data from Appendix B the regression constant α is equal to...

$$\begin{aligned}\alpha &= \bar{x} - \beta_1\bar{y} - \beta_2\bar{z}\\ &= 27894 - (0.9708)(32239) - (-6.0097)(735)\\ &= 1012\end{aligned} \tag{24}$$

Residuals

A residual is defined as the differences between the estimated value of x via Equation (1) above and the actual value of x. Residuals are captured in the moments of the error term ϵ. Using Equations (1) and (4) above the equation for the first moment of the distribution of the error term is...

$$\begin{aligned}\mathbb{E}\Big[\epsilon\Big] &= \mathbb{E}\Big[x - \alpha - \beta_1\,y - \beta_2\,z\Big]\\ &= \mathbb{E}\Big[x\Big] - \mathbb{E}\Big[\alpha + \beta_1\,y + \beta_2\,z\Big]\\ &= \mathbb{E}\Big[x\Big] - \mathbb{E}\Big[x\Big]\\ &= 0\end{aligned} \tag{25}$$

Using Equations (1) and (4) above the equation for the second moment of the distribution of the error term is...

$$\begin{aligned}\mathbb{E}\Big[\epsilon^2\Big] &= \mathbb{E}\Big[\Big(x - \alpha - \beta_1\,y - \beta_2\,z\Big)^2\Big]\\ &= \mathbb{E}\Big[x^2 + \alpha^2 - 2\alpha x + 2\alpha\beta_1 y + 2\alpha\beta_2 z - 2\beta_1 xy - 2\beta_2 xz + 2\beta_1\beta_2 yz + \beta_1^2 y^2 + \beta_2^2 z^2\Big]\\ &= \mathbb{E}\Big[x^2 + \Big(\bar{x}^2 - 2\beta_1\bar{x}\bar{y} - 2\beta_2\bar{x}\bar{z} + 2\beta_1\beta_2\bar{y}\bar{z} + \beta_1^2\bar{y}^2 + \beta_2^2\bar{z}^2\Big) - 2\Big(\bar{x} - \beta_1\bar{y} - \beta_2\bar{z}\Big)x\\ &\quad + 2\Big(\bar{x} - \beta_1\bar{y} - \beta_2\bar{z}\Big)\beta_1 y + 2\Big(\bar{x} - \beta_1\bar{y} - \beta_2\bar{z}\Big)\beta_2 z - 2\beta_1 xy - 2\beta_2 xz + 2\beta_1\beta_2 yz + \beta_1^2 y^2 + \beta_2^2 z^2\Big]\\ &= \mathbb{E}\Big[x^2 + \bar{x}^2 - 2\beta_1\bar{x}\bar{y} - 2\beta_2\bar{x}\bar{z} + 2\beta_1\beta_2\bar{y}\bar{z} + \beta_1^2\bar{y}^2 + \beta_2^2\bar{z}^2 - \Big(2\bar{x} - 2\beta_1\bar{y} - 2\beta_2\bar{z}\Big)x\\ &\quad + \Big(2\beta_1\bar{x} - 2\beta_1^2\bar{y} - 2\beta_1\beta_2\bar{z}\Big)y + \Big(2\beta_2\bar{x} - 2\beta_1\beta_2\bar{y} - 2\beta_2^2\bar{z}\Big)z - 2\beta_1 xy - 2\beta_2 xz + 2\beta_1\beta_2 yz + \beta_1^2 y^2 + \beta_2^2 z^2\Big]\\ &= \mathbb{E}\Big[x^2 + \bar{x}^2 - 2\beta_1\bar{x}\bar{y} - 2\beta_2\bar{x}\bar{z} + 2\beta_1\beta_2\bar{y}\bar{z} + \beta_1^2\bar{y}^2 + \beta_2^2\bar{z}^2 - 2\bar{x}^2 + 2\beta_1\bar{x}\bar{y} + 2\beta_2\bar{x}\bar{z} + 2\beta_1\bar{x}\bar{y} - 2\beta_1^2\bar{y}^2\\ &\quad - 2\beta_1\beta_2\bar{y}\bar{z} + 2\beta_2\bar{x}\bar{z} - 2\beta_1\beta_2\bar{y}\bar{z} - 2\beta_2^2\bar{z}^2 - 2\beta_1\mathbb{E}\Big[xy\Big] - 2\beta_2\mathbb{E}\Big[xz\Big] + 2\beta_1\beta_2\mathbb{E}\Big[yz\Big] + \beta_1^2\mathbb{E}\Big[y^2\Big] + \beta_2^2\mathbb{E}\Big[z^2\Big]\\ &= \Big(\mathbb{E}\Big[x^2\Big] - \bar{x}^2\Big) + \beta_1^2\Big(\mathbb{E}\Big[y^2\Big] - \bar{y}^2\Big) + \beta_2^2\Big(\mathbb{E}\Big[z^2\Big] - \bar{z}^2\Big) - 2\beta_1\Big(\mathbb{E}\Big[xy\Big] - \bar{x}\bar{y}\Big) - 2\beta_2\Big(\mathbb{E}\Big[xz\Big] - \bar{x}\bar{z}\Big)\\ &\quad + 2\beta_1\beta_2\Big(\mathbb{E}\Big[yz\Big] - \bar{y}\bar{z}\Big)\\ &= Var(x) + \beta_1^2 Var(y) + \beta_2^2 Var(z) - 2\beta_1 Cov(x,y) - 2\beta_2 Cov(x,z) + 2\beta_1\beta_2 Cov(y,z)\end{aligned} \tag{26}$$

Using Equation (25) above the mean of the error term ϵ is...

$$\text{mean of } \epsilon = \mathbb{E}\Big[\epsilon\Big] = 0 \tag{27}$$

Using Equations (25) and (26) above the variance of the error term ϵ is...

$$\begin{aligned}
\text{variance of } \epsilon &= \mathbb{E}\left[\epsilon^2\right] - \left(\mathbb{E}\left[\epsilon\right]\right)^2 \\
&= Var(x) + \beta_1^2 Var(y) + \beta_2^2 Var(z) - 2\beta_1 Cov(x,y) - 2\beta_2 Cov(x,z) + 2\beta_1\beta_2 Cov(y,z) - 0 \\
&= Var(x) + \beta_1^2 Var(y) + \beta_2^2 Var(z) - 2\beta_1 Cov(x,y) - 2\beta_2 Cov(x,z) + 2\beta_1\beta_2 Cov(y,z) \qquad (28)
\end{aligned}$$

The Solution To Our Problem

Using Equations (24) and (21) above the OLS estimate of annual materials costs is...

$$\begin{aligned}
\text{Materials Cost} &= \alpha + \beta_1 \,\text{Sales} + \beta_2 \,\text{Ending Inventory} \\
&= 1012 + (0.9708)(50,000,000) + (-6.0097)(5,000,000) \\
&= 18,493,079 \qquad (29)
\end{aligned}$$

Using Equation (28) above and the time series data in Appendix C the variance of the error term ϵ is...

$$\begin{aligned}
\text{Error variance} &= Var(x) + \beta_1^2 Var(y) + \beta_2^2 Var(z) - 2\beta_1 Cov(x,y) - 2\beta_2 Cov(x,z) + 2\beta_1\beta_2 Cov(y,z) \\
&= 171,212,813 + (0.9708)^2(225,859,160) + (-6.0097)^2(66,545) - (2)(0.9708)(196,428,819) - (2)(-6.0097)(3,289 \\
&= 285,349 \qquad (30)
\end{aligned}$$

Using Equations (29) and (30) above the confidence interval for our estimate within two standard deviations is...

$$\begin{aligned}
\text{Confidence interval} &= 18,493,079 \pm \sqrt{285,249} \\
&= 18,493,079 \pm 534 \qquad (31)
\end{aligned}$$

Appendix

A. Meat Packing Plant data for the United States from 1958 to 1996. Table data is for Standard Industrial Code 2011 and is in millions of dollars.

Year	Materials	Sales	End Inv
1958	10230.10	11950.70	408.10
1959	9939.10	11788.40	370.10
1960	9890.80	11806.20	381.60
1961	10047.30	11916.80	395.30
1962	10508.80	12468.30	411.10
1963	10507.30	12412.60	407.80
1964	10835.90	12950.80	415.90
1965	11852.10	13882.50	435.60
1966	13015.10	15011.80	451.20
1967	13329.50	15553.70	478.00
1968	13845.70	16260.90	512.30
1969	15510.20	17940.20	524.20
1970	15711.10	18408.10	460.90
1971	15785.40	18802.40	494.40
1972	20120.10	23003.40	555.90
1973	24073.30	27311.50	681.70
1974	25090.70	28834.60	662.30
1975	27225.60	31341.80	741.80
1976	27977.00	32392.60	676.80
1977	27239.10	31208.20	700.10
1978	33963.50	38198.70	892.50
1979	38073.10	43191.30	964.60
1980	37762.50	42962.00	996.50
1981	39235.80	44570.00	908.80
1982	39048.60	44853.60	891.60
1983	37507.70	42774.60	940.80
1984	38738.30	44277.70	932.30
1985	36637.20	42553.50	856.00
1986	36680.20	42384.50	830.90
1987	40302.90	45536.60	832.20
1988	41700.60	47333.20	931.10
1989	41122.40	46542.00	929.00
1990	44462.70	51069.20	995.20
1991	43311.90	49326.20	949.20
1992	43586.40	50434.40	1054.30
1993	45735.80	53240.30	1182.70
1994	42045.70	50443.70	1104.00
1995	42456.70	51314.40	1178.60
1996	42762.60	51088.60	1125.90

B. Meat Packing Plant data (1958-1996) averages. In millions of dollars.

	Materials	Sales	EndInv
Mean	27894	32239	735

C. Meat Packing Plant data (1958-1996) covariance matrix. In millions of dollars.

	Materials	Sales	EndInv
Materials	171212813	196428819	3289338
Sales		225859160	3800172
EndInv			66545