CHAPTER IX

# SOLUTIONS TO THE EXERCISES

## 9.3 Chapter 3

**1.** **(a)** Let us first examine the nature of the best approximation of $X$ by a matrix of rank 1. The eigenvalues of

$$X'X = \begin{bmatrix} 1 & r^2 \\ r^2 & 1 \end{bmatrix}$$

are the solutions of the characteristic equation

$$|I\lambda - X'X| = \begin{vmatrix} \lambda - 1 & -r^2 \\ -r^2 & \lambda - 1 \end{vmatrix} = \lambda^2 - 2\lambda + 1 - r^4 = 0,$$

which are

$$\lambda_1 = 1 + r^2 \quad \text{and} \quad \lambda_2 = 1 - r^2.$$

The diagonal matrix of singular values of $X$ is then

$$S = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} = \begin{bmatrix} \sqrt{1+r^2} & 0 \\ 0 & \sqrt{1-r^2} \end{bmatrix}. \tag{9.3.1}$$

The eigenvectors of $X'X$ (i.e., the right singular vectors of $X$) are the column vectors $v$ satisfying

$$X'Xv = \begin{bmatrix} 1 & r^2 \\ r^2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} (1 \pm r^2),$$

or

$$\left( \begin{bmatrix} 1 \pm r^2 & 0 \\ 0 & 1 \pm r^2 \end{bmatrix} - \begin{bmatrix} 1 & r^2 \\ r^2 & 1 \end{bmatrix} \right) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \pm r^2 & -r^2 \\ -r^2 & \pm r^2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We see immediately that the two components of $v$ must be either equal or of opposite sign. Normalizing them to have length 1, we obtain the orthogonal matrix

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}. \tag{9.3.2}$$

It follows that for some $n \times 2$ matrix $Q_1$, whose columns are orthogonal to each other and of length 1, the singular-value decomposition of $X$ is

$$\begin{aligned} X &= Q_1 S P' \\ &= Q_1 \begin{bmatrix} \sqrt{1+r^2} & 0 \\ 0 & \sqrt{1-r^2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \\ &= Q_1 \begin{bmatrix} \sqrt{\dfrac{1+r^2}{2}} & \sqrt{\dfrac{1+r^2}{2}} \\ \sqrt{\dfrac{1-r^2}{2}} & -\sqrt{\dfrac{1-r^2}{2}} \end{bmatrix}. \end{aligned} \tag{9.3.3}$$

The best approximation of $X$ by a matrix of rank 1 is obtained by replacing the matrix $S$ of (9.3.1) by the matrix

$$S_1 = \begin{bmatrix} s_1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{1+r^2} & 0 \\ 0 & 0 \end{bmatrix} \tag{9.3.4}$$

and then substituting $S_1$ for $S$ in (9.3.3) to obtain

$$\begin{aligned} X_{(1)} &= Q_1 S_1 P' \\ &= Q_1 \begin{bmatrix} \sqrt{1+r^2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \\ &= Q_1 \begin{bmatrix} \sqrt{\dfrac{1+r^2}{2}} & \sqrt{\dfrac{1+r^2}{2}} \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{9.3.5}$$

To obtain the expression for $Q_1$ we may may postmultiply (9.3.3) by $PS^{-1}$ to obtain $Q_1 = XPS^{-1}$. The $t$th row of $Q_1$ is

$$\begin{aligned} (q_{t1}, q_{t2}) &= (x_{t1}, x_{t2}) \begin{bmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \dfrac{1}{\sqrt{1+r^2}} & 0 \\ 0 & \dfrac{1}{\sqrt{1-r^2}} \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{x_{t1}+x_{t2}}{\sqrt{2}} & \dfrac{x_{t1}-x_{t2}}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \dfrac{1}{\sqrt{1+r^2}} & 0 \\ 0 & \dfrac{1}{\sqrt{1-r^2}} \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{x_{t1}+x_{t2}}{\sqrt{2(1+r^2)}} & \dfrac{x_{t1}-x_{t2}}{\sqrt{2(1-r^2)}} \end{bmatrix}. \end{aligned} \tag{9.3.6}$$

Now we substitute (9.3.6) back in (9.3.5) to obtain the $t$th row of $X_{(1)}$:

$$\begin{aligned} \left(x_{t1}^{(1)}, x_{t2}^{(1)}\right) &= (q_{t1}, q_{t2}) \begin{bmatrix} \sqrt{\dfrac{1+r^2}{2}} & \sqrt{\dfrac{1+r^2}{2}} \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{x_{t1}+x_{t2}}{\sqrt{2(1+r^2)}} & \dfrac{x_{t1}-x_{t2}}{\sqrt{2(1-r^2)}} \end{bmatrix} \begin{bmatrix} \sqrt{\dfrac{1+r^2}{2}} & \sqrt{\dfrac{1+r^2}{2}} \\ 0 & 0 \end{bmatrix} \\ &= \left(\tfrac{1}{2}(x_{t1}+x_{t2}), \tfrac{1}{2}(x_{t1}+x_{t2})\right). \end{aligned} \tag{9.3.7}$$

Thus, the best approximation $X_{(1)}$ of $X$ consists of the matrix in which each of the two columns of $X$ has been replaced by the average of the two. □

**(b)** The Marquardt estimator of $\beta$ is $X_{(1)}^{\dagger} y$. From (9.3.5) and (9.3.6) we have

$$\begin{aligned} X_{(1)}^{\dagger} &= P S_{(1)}^{\dagger} Q_1' \\ &= \begin{bmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \dfrac{1}{\sqrt{1+r^2}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dfrac{x_{11}+x_{12}}{\sqrt{2(1+r^2}} & \cdots & \dfrac{x_{n1}+x_{n2}}{\sqrt{2(1+r^2}} \\ \dfrac{x_{11}-x_{12}}{\sqrt{2(1-r^2}} & \cdots & \dfrac{x_{n1}-x_{n2}}{\sqrt{2(1-r^2}} \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{x_{11}+x_{12}}{2(1+r^2)} & \cdots & \dfrac{x_{n1}+x_{n2}}{2(1+r^2)} \\ \dfrac{x_{11}+x_{12}}{2(1+r^2)} & \cdots & \dfrac{x_{n1}+x_{n2}}{2(1+r^2)} \end{bmatrix}. \end{aligned} \tag{9.3.8}$$

Accordingly, the Marquardt estimator is

$$\hat{\beta}_{(1)} = X^{\dagger}_{(1)} y = \frac{1}{2(1+r^2)} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \sum_{t=1}^{n} (x_{t1} + x_{t2}) y_t. \quad \square \tag{9.3.9}$$

**(c)** To obtain the least-squares estimator we compute the Moore-Penrose generalized inverse of $X$:

$$\begin{aligned} X^{\dagger} &= P S^{-1} Q_1' \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{1+r^2}} & 0 \\ 0 & \frac{1}{\sqrt{1-r^2}} \end{bmatrix} \begin{bmatrix} \frac{x_{11}+x_{12}}{\sqrt{2(1+r^2}} & \cdots & \frac{x_{n1}+x_{n2}}{\sqrt{2(1+r^2}} \\ \frac{x_{11}-x_{12}}{\sqrt{2(1-r^2}} & \cdots & \frac{x_{n1}-x_{n2}}{\sqrt{2(1-r^2}} \end{bmatrix}. \end{aligned} \tag{9.3.10}$$

Letting $\iota$ denote the column vector of two ones, the matrix each of whose rows is the average of the two rows of $X^{\dagger}$ may be written $\iota(\iota'\iota)^{-1}\iota X^{\dagger}$, which computes to

$$\begin{aligned} \iota(\iota'\iota)^{-1}\iota X^{\dagger} &= \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{1+r^2}} & 0 \\ 0 & \frac{1}{\sqrt{1-r^2}} \end{bmatrix} \begin{bmatrix} \frac{x_{11}+x_{12}}{\sqrt{2(1+r^2}} & \cdots & \frac{x_{n1}+x_{n2}}{\sqrt{2(1+r^2}} \\ \frac{x_{11}-x_{12}}{\sqrt{2(1-r^2}} & \cdots & \frac{x_{n1}-x_{n2}}{\sqrt{2(1-r^2}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{x_{11}+x_{12}}{2(1+r^2)} & \cdots & \frac{x_{n1}+x_{n2}}{2(1+r^2)} \\ \frac{x_{11}-x_{12}}{2(1-r^2)} & \cdots & \frac{x_{n1}-x_{n2}}{2(1-r^2)} \end{bmatrix}. \end{aligned} \tag{9.3.11}$$

This is the same as (9.3.8). $\square$

(**d**) Since the least-squares estimator of $\beta$,

$$\tilde{\beta} = X^{\dagger} y = (X'X)^{-1} X' y, \tag{9.3.12}$$

is unbiased, its (matrix) mean-square error is the same as its variance,

$$\text{Risk}\{\tilde{\beta}\} = \sigma^2 (X'X)^{-1} = \sigma^2 \begin{bmatrix} 1 & r^2 \\ r^2 & 1 \end{bmatrix}^{-1} = \frac{\sigma^2}{1-r^4} \begin{bmatrix} 1 & -r^2 \\ -r^2 & 1 \end{bmatrix}. \tag{9.3.13}$$

Denoting (as above) by $\iota$ the column vector of two ones, the averaged estimator used by the investigator (which as we have just seen in part (b) above is the Marquard estimator) may be denoted

$$\hat{\beta}_{(1)} = \iota(\iota'\iota)^{-1}\iota'\tilde{\beta}. \tag{9.3.14}$$

Its deviation from $\beta$ is

$$\hat{\beta}_{(1)} - \beta = \iota(\iota'\iota)^{-1}\iota'(X'X)^{-1}X'\varepsilon - [I - \iota(\iota'\iota)^{-1}\iota']\beta; \tag{9.3.15}$$

consequently its matrix mean-square error is

$$\begin{aligned}
\text{Risk}\{\hat\beta_{(1)}\} &= \sigma^2\iota(\iota'\iota)^{-1}\iota'(X'X)^{-1}\iota(\iota'\iota)^{-1}\iota' \\
&\quad + [I-\iota(\iota'\iota)^{-1}\iota']\beta\beta'[I-\iota(\iota'\iota)^{-1}\iota'] \\
&= \frac{\sigma^2}{1-r^4}\begin{bmatrix}\frac12 & \frac12\\ \frac12 & \frac12\end{bmatrix}\begin{bmatrix}1 & -r^2\\ -r^2 & 1\end{bmatrix}\begin{bmatrix}\frac12 & \frac12\\ \frac12 & \frac12\end{bmatrix} \\
&\quad + \frac14\begin{bmatrix}(\beta_1-\beta_2)^2 & -(\beta_1-\beta_2)^2\\ -(\beta_1-\beta_2)^2 & (\beta_1-\beta_2)^2\end{bmatrix} \\
&= \begin{bmatrix}\dfrac{\sigma^2}{2(1+r^2)}+\dfrac{(\beta_1-\beta_2)^2}{4} & \dfrac{\sigma^2}{2(1+r^2)}-\dfrac{(\beta_1-\beta_2)^2}{4}\\ \dfrac{\sigma^2}{2(1+r^2)}-\dfrac{(\beta_1-\beta_2)^2}{4} & \dfrac{\sigma^2}{2(1+r^2)}+\dfrac{(\beta_1-\beta_2)^2}{4}\end{bmatrix}. \qquad (9.3.16)
\end{aligned}$$

The required expressions are given by (9.3.13) and (9.3.16). □

**(e)** Taking into account the fact that

$$\frac{1}{1+r^2}+\frac{1}{1-r^2}=\frac{2}{1-r^4},$$

we see that the difference between the two mean-square errors (9.3.13) and (9.3.15) is

$$\begin{aligned}
&\text{Risk}\{\tilde\beta\}-\text{Risk}\{\hat\beta\} \\
&= \begin{bmatrix}\dfrac{\sigma^2}{2(1-r^2)}-\dfrac{(\beta_1-\beta_2)^2}{4} & -\dfrac{\sigma^2}{2(1-r^2)}+\dfrac{(\beta_1-\beta_2)^2}{4}\\ -\dfrac{\sigma^2}{2(1-r^2)}+\dfrac{(\beta_1-\beta_2)^2}{4} & \dfrac{\sigma^2}{2(1-r^2)}-\dfrac{(\beta_1-\beta_2)^2}{4}\end{bmatrix}. \qquad (9.3.17)
\end{aligned}$$

Now the second row of (9.3.17) is simply the negative of the first row, hence the matrix is singular; therefore, it is positive-semidefinite if and only if the two (identical) diagonal elements are positive, i.e.,

$$\text{Risk}\{\tilde\beta\}\succ\text{Risk}\{\hat\beta\}\iff\frac{(\beta_1-\beta_2)^2}{2\sigma^2}<\frac{1}{1-r^2}. \qquad (9.3.18)$$

This is the required necessary and sufficient condition. □

**2.** From (9.3.1), the condition number of $X$ is

$$\kappa(X)=\sqrt{\frac{1+r^2}{1-r^2}}.$$