UNIVERSITY OF EDINBURGH
SCHOOL OF MATHEMATICS

Generalised Regression Models

GRM: Solutions 3 **Semester 1, 2022–2023**

---

1. The model may be written as

$$\begin{aligned} E(Y_1) &= \theta \\ E(Y_2) &= 2\theta - \phi \\ E(Y_3) &= \theta + 2\phi \end{aligned}$$

$$\text{var}(Y_i) = \sigma^2 \ \ (i = 1,2,3)$$

$$\text{cov}(Y_i, Y_j) = 0 \ \ (i \neq j).$$

This may be expressed in the form

$$E(\mathbf{Y}|X) = X\beta \qquad \text{var}(\mathbf{Y}|X) = \sigma^2 I$$

where

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}, \ \beta = \begin{pmatrix} \theta \\ \phi \end{pmatrix}, \ X = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Therefore, the least squares estimate of $\begin{pmatrix} \theta \\ \phi \end{pmatrix}$ is given by $\begin{pmatrix} \widehat{\theta} \\ \widehat{\phi} \end{pmatrix} = \widehat{\beta} = (X^T X)^{-1} X^T \mathbf{y}$.

$$(X^T X) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix}$$

$$(X^T X)^{-1} = \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}$$

and

$$(X^T \mathbf{y}) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} y_1 + 2y_2 + y_3 \\ -y_2 + 2y_3 \end{pmatrix}.$$

Thus,

$$\begin{pmatrix} \widehat{\theta} \\ \widehat{\phi} \end{pmatrix} = \widehat{\beta} = (X^T X)^{-1} X^T \mathbf{y} = \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} y_1 + 2y_2 + y_3 \\ -y_2 + 2y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{6}(y_1 + 2y_2 + y_3) \\ \frac{1}{5}(2y_3 - y_2) \end{pmatrix}.$$

The covariance matrix of the least squares estimator of β is given by

$$\text{var}(\widehat{\beta}) = (X^T X)^{-1} \sigma^2 = \sigma^2 \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}.$$

So $\widehat{\theta}$ and $\widehat{\phi}$ are uncorrelated, and the standard errors are given by the square-roots of the diagonal elements of this matrix, i.e.

$$\begin{aligned} SE(\widehat{\theta}) &= \sqrt{\text{var}(\widehat{\theta})} = \frac{\sigma}{\sqrt{6}} \\ SE(\widehat{\phi}) &= \sqrt{\text{var}(\widehat{\phi})} = \frac{\sigma}{\sqrt{5}} \end{aligned}$$

To estimate the standard errors, use the (unbiased) estimate of σ^2 given by

$$\widehat{\sigma}^2 = \frac{RSS}{n-p} = \frac{\mathbf{y}^T \mathbf{y} - \widehat{\beta}^T X^T \mathbf{y}}{3-2} = (y_1^2 + y_2^2 + y_3^2) - \left\{ \frac{(y_1 + 2y_2 + y_3)^2}{6} + \frac{(2y_3 - y_2)^2}{5} \right\}.$$

2. If $\mathbf{X}$ is $n \times p$ and of rank p and $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, then

$$\mathbf{H}^T = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T = \mathbf{H}$$

so that $\mathbf{H}$ is symmetric, and

$$\mathbf{H}^2 = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T = \mathbf{H},$$

so that $\mathbf{H}$ is idempotent. Note that if the product $\mathbf{AB}$ exists then $\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{A})$. Thus

$$\text{rank}(\mathbf{H}) \leq \text{rank}(\mathbf{X}) = p,$$

and (since $\mathbf{X} = \mathbf{HX}$)

$$p = \text{rank}(\mathbf{X}) = \text{rank}(\mathbf{HX}) \leq \text{rank}(\mathbf{H}),$$

so that

$$\text{rank}(\mathbf{H}) = \text{rank}(\mathbf{X}) = p.$$

3. (a) The model

$$\text{E}(Y_t\,|\,t) = \beta_0 + \beta_1 \cos\left(\frac{\pi t}{12}\right) + \beta_2 \sin\left(\frac{\pi t}{12}\right) \qquad (t = 1, 2, \ldots, 24)$$

may be fitted in R using the file Births.txt, e.g. using the commands

```
Births.dat <- read.table('Births.txt',header=T)
attach(Births.dat)
lm(Number_of_births~cos(Hour_ending_at/12*pi)+sin(Hour_ending_at/12*pi))
```

The regression coefficients for $\cos\left(\frac{\pi t}{12}\right)$ and $\sin\left(\frac{\pi t}{12}\right)$ are $\widehat{\beta}_1 = -8.79$ and $\widehat{\beta}_2 = 66.94$; also $\widehat{\beta}_0 = 392.5$.

Alternatively, calculate using $\widehat{\beta} = (X^TX)^{-1}X^T\mathbf{y}$, where $X^TX = \text{diag}(n, \sum_t \cos^2\left(\frac{\pi t}{12}\right), \sum_t \sin^2\left(\frac{\pi t}{12}\right)) = \text{diag}(24, 12, 12)$ and $X^T\mathbf{y} = (\sum_t y_t, \sum y_t \cos\left(\frac{\pi t}{12}\right), \sum y_t \sin\left(\frac{\pi t}{12}\right))^T = (9421.0, -105.48, 803.25)^T$ giving the estimates $\widehat{\beta} = (9421/24, -105.48/12, 803.25/12)^T = (392.5, -8.79, 66.94)^T$.

The fitted model is thus

$$\widehat{Y}_t = 392.5 - 8.79\cos\left(\frac{\pi t}{12}\right) + 66.94\sin\left(\frac{\pi t}{12}\right) \qquad (t = 1, 2, \ldots, 24).$$

(b) The regression equation is equivalent to

$$\text{E}(Y_t\,|\,t) = \beta_0 + \gamma\cos\left\{\frac{\pi(t-\phi)}{12}\right\} \qquad (t = 1, 2, \ldots, 24)$$

and as

$$\cos\left(\frac{\pi(t-\phi)}{12}\right) = \cos\left(\frac{\pi t}{12} - \frac{\pi\phi}{12}\right) = \cos\left(\frac{\pi t}{12}\right)\cos\left(\frac{\pi\phi}{12}\right) + \sin\left(\frac{\pi t}{12}\right)\sin\left(\frac{\pi\phi}{12}\right)$$

we have

$$\beta_1 = \gamma\cos\left(\frac{\pi\phi}{12}\right), \quad \beta_2 = \gamma\sin\left(\frac{\pi\phi}{12}\right).$$

Thus, γ and ϕ may be expressed in terms of β_1 and β_2 as

$$\gamma = \sqrt{\beta_1^2 + \beta_2^2}, \quad \phi = \frac{12}{\pi}\arctan\left(\frac{\beta_2}{\beta_1}\right).$$

Hence the amplitude γ and the angle ϕ corresponding to the maximum expected number are estimated by

$$\widehat{\gamma} = \sqrt{\widehat{\beta}_1^2 + \widehat{\beta}_2^2} = 67.5, \qquad \widehat{\phi} = \frac{12}{\pi}\arctan\left(\frac{66.94}{-8.79}\right) = -5.50 \quad \text{or} \quad 6.50.$$

The latter time is obviously the correct one, and corresponds to the hour from 5:30 to 6:30, so the expected number is maximized at 6 a.m.

4. (a) If $y_{1\bullet}$ and $y_{2\bullet}$ denote the sums of the first m and the other $n-m$ responses and $\beta = (\mu \quad \delta)^T$ then

$$\mathbf{X}^T = \begin{pmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ 0 & \dots & 0 & 1 & \dots & 1 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{X} = \begin{pmatrix} n & n-m \\ n-m & n-m \end{pmatrix}, \quad \mathbf{X}^T\mathbf{y} = \begin{pmatrix} y_{1\bullet} + y_{2\bullet} \\ y_{2\bullet} \end{pmatrix}.$$

The estimates of μ and δ (given by the elements of $\widehat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$) are $\bar{y}_1$ and $\bar{y}_2 - \bar{y}_1$ respectively. The residual SS has $n-2$ degrees of freedom.

(b) Here

$$\beta = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{1}_{n_2} & \mathbf{0}_{n_2} \\ \mathbf{0}_{n_3} & \mathbf{0}_{n_3} & \mathbf{1}_{n_3} \end{pmatrix}, \quad \mathbf{X}^T\mathbf{X} = \begin{pmatrix} n_1 & 0 & 0 \\ 0 & n_2 & 0 \\ 0 & 0 & n_3 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{y} = \begin{pmatrix} y_{1\bullet} \\ y_{2\bullet} \\ y_{3\bullet} \end{pmatrix}.$$

The estimates of the μ_j are therefore the means of the three groups of responses, $\bar{y}_1$, $\bar{y}_2$, $\bar{y}_3$. The residual SS is the sum of squares within groups and has $n_1 + n_2 + n_3 - 3$ degrees of freedom.

(c) If vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ contain the first m and the other $n-m$ values x_i and $\beta = (\alpha_1 \quad \alpha_2 \quad \beta)^T$ then

$$\mathbf{X} = \begin{pmatrix} \mathbf{1}_m & \mathbf{0}_m & \mathbf{x}_1 \\ \mathbf{0}_{n-m} & \mathbf{1}_{n-m} & \mathbf{x}_2 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{X} = \begin{pmatrix} m & 0 & \sum_1 x_i \\ 0 & n-m & \sum_2 x_i \\ \sum_1 x_i & \sum_2 x_i & \sum x_i^2 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{y} = \begin{pmatrix} y_{1\bullet} \\ y_{2\bullet} \\ \sum x_i y_i \end{pmatrix},$$

where $\sum_1$ and $\sum_2$ denote summation over the first m and the remaining $n-m$ values of i. The first two normal equations, $\mathbf{X}^T\mathbf{X}\widehat{\beta} = \mathbf{X}^T\mathbf{y}$, give $\widehat{\alpha}_1 = \bar{y}_1 - \widehat{\beta}\bar{x}_1$ and $\widehat{\alpha}_2 = \bar{y}_2 - \widehat{\beta}\bar{x}_2$. Substituting these expressions into the third equation and solving for $\widehat{\beta}$ gives

$$\widehat{\beta} = \frac{\sum_1 (x_i - \bar{x}_1) y_i + \sum_2 (x_i - \bar{x}_2) y_i}{\sum_1 (x_i - \bar{x}_1)^2 + \sum_2 (x_i - \bar{x}_2)^2}.$$

There are 3 parameters in β, so the residual SS has $n-3$ degrees of freedom.

5. (a)

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{X} = \begin{pmatrix} n & \sum_i x_i & \sum_i x_i^2 \\ \sum_i x_i & \sum_i x_i^2 & \sum_i x_i^3 \\ \sum_i x_i^2 & \sum_i x_i^3 & \sum_i x_i^4 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{y} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \\ \sum_i x_i^2 y_i \end{pmatrix}.$$

The residual sum of squares has $n-3$ degrees of freedom.

(b) Here $\beta = (\alpha \quad \beta_1 \quad \beta_2)^T$ and

$$\mathbf{X} = \begin{pmatrix} \mathbf{1}_m & \mathbf{x}_1 & \mathbf{0}_m \\ \mathbf{1}_{n-m} & \mathbf{0}_{n-m} & \mathbf{x}_2 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{X} = \begin{pmatrix} n & \sum_1 x_i & \sum_2 x_i \\ \sum_1 x_i & \sum_1 x_i^2 & 0 \\ \sum_2 x_i & 0 & \sum_2 x_i^2 \end{pmatrix}, \quad \mathbf{X}^T\mathbf{y} = \begin{pmatrix} \sum y_i \\ \sum_1 x_i y_i \\ \sum_2 x_i y_i \end{pmatrix},$$

where $\sum_1$ and $\sum_2$ denote summation over the first m and the remaining $n-m$ values of i. [The equations for α, β_1 and β_2 are not easy to simplify.] The residual SS has $n-3$ degrees of freedom.

6. Putting $\mathbf{A}_{11} = a$, $\mathbf{A}_{12} = \mathbf{b}^T$, $\mathbf{A}_{21} = \mathbf{b}$, $\mathbf{A}_{22} = c\mathbf{I}_p$ in (9.3) and (9.4) of *Useful Matrix Results* gives

$$\mathbf{A}^{-1} = \begin{pmatrix} d & -c^{-1}d\mathbf{b}^T \\ -c^{-1}d\mathbf{b} & c^{-1}\left(\mathbf{I}_p + c^{-1}d\,\mathbf{b}\mathbf{b}^T\right) \end{pmatrix},$$

where $d = (a - c^{-1}\mathbf{b}^T\mathbf{b})^{-1}$, and

$$|\mathbf{A}| = c^p\,(a - c^{-1}\mathbf{b}^T\mathbf{b})$$

$\mathbf{A}$ is singular if c or $a - c^{-1}\mathbf{b}^T\mathbf{b}$ is zero.

7. If $\mathbf{y}$ denotes the $2n$-vector $(y_{11} \ldots y_{n1}\; y_{12} \ldots y_{n2})^T$ then the least squares estimates are those for a linear model

$$\mathrm{E}(\mathbf{Y}\,|\,\mathbf{X}) = \mathbf{X}\beta$$

in which

$$\beta = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{pmatrix} \quad \text{and} \quad \mathbf{X} = \begin{pmatrix} 1 & 0 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & x_{n1} & x_{n2} \\ 0 & 1 & x_{12} & -x_{11} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & x_{n2} & -x_{n1} \end{pmatrix}.$$

This gives

$$\mathbf{X}^T\mathbf{X} = n \begin{pmatrix} 1 & 0 & \bar{x}_1 & \bar{x}_2 \\ 0 & 1 & \bar{x}_2 & -\bar{x}_1 \\ \bar{x}_1 & \bar{x}_2 & t & 0 \\ \bar{x}_2 & -\bar{x}_1 & 0 & t \end{pmatrix}, \quad \mathbf{X}^T\mathbf{y} = \begin{pmatrix} n\bar{y}_1 \\ n\bar{y}_2 \\ u \\ v \end{pmatrix}$$

where

$$\bar{x}_1 = \frac{1}{n}\sum x_{i1}, \;\; \bar{x}_2 = \frac{1}{n}\sum x_{i2}, \;\; \bar{y}_1 = \frac{1}{n}\sum y_{i1}, \;\; \bar{y}_2 = \frac{1}{n}\sum y_{i2},$$

and

$$\begin{aligned} t &= \frac{1}{n}\sum_i (x_{i1}^2 + x_{i2}^2) \\ u &= \sum_i (x_{i1}\,y_{i1} + x_{i2}\,y_{i2}) \\ v &= \sum_i (x_{i2}\,y_{i1} - x_{i1}\,y_{i2}). \end{aligned}$$

Thus, the *normal equations*

$$(\mathbf{X}^T\mathbf{X})\widehat{\beta} = \mathbf{X}^T\mathbf{y}$$

may be written as

$$n \begin{pmatrix} 1 & 0 & \bar{x}_1 & \bar{x}_2 \\ 0 & 1 & \bar{x}_2 & -\bar{x}_1 \\ \bar{x}_1 & \bar{x}_2 & t & 0 \\ \bar{x}_2 & -\bar{x}_1 & 0 & t \end{pmatrix} \begin{pmatrix} \widehat{\alpha}_1 \\ \widehat{\alpha}_2 \\ \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{pmatrix} = \begin{pmatrix} n\bar{y}_1 \\ n\bar{y}_2 \\ u \\ v \end{pmatrix}.$$

Note that

$$(\mathbf{X}^T\mathbf{X})^{-1} = \frac{1}{n(t - \bar{x}_1^2 - \bar{x}_2^2)} \begin{pmatrix} t & 0 & -\bar{x}_1 & -\bar{x}_2 \\ 0 & t & -\bar{x}_2 & \bar{x}_1 \\ -\bar{x}_1 & -\bar{x}_2 & 1 & 0 \\ -\bar{x}_2 & \bar{x}_1 & 0 & 1 \end{pmatrix}$$

and thus the least squares estimates are

$$\begin{aligned} \widehat{\beta} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \\ &= \frac{1}{n(t - \bar{x}_1^2 - \bar{x}_2^2)} \begin{pmatrix} t & 0 & -\bar{x}_1 & -\bar{x}_2 \\ 0 & t & -\bar{x}_2 & \bar{x}_1 \\ -\bar{x}_1 & -\bar{x}_2 & 1 & 0 \\ -\bar{x}_2 & \bar{x}_1 & 0 & 1 \end{pmatrix} \begin{pmatrix} n\bar{y}_1 \\ n\bar{y}_2 \\ u \\ v \end{pmatrix} \\ &= \frac{1}{n(t - \bar{x}_1^2 - \bar{x}_2^2)} \begin{pmatrix} nt\bar{y}_1 - \bar{x}_1 u - \bar{x}_2 v \\ nt\bar{y}_2 - \bar{x}_2 u + \bar{x}_1 v \\ u - n\bar{x}_1\bar{y}_1 - n\bar{x}_2\bar{y}_2 \\ v - n\bar{x}_2\bar{y}_1 + n\bar{x}_1\bar{y}_2 \end{pmatrix}. \end{aligned}$$