where the ε_i are i.i.d. $N(0, \sigma^2)$ and the x_{ij} are standardized so that for $j = 1, 2, \ldots, p-1$, $\sum_i x_{ij} = 0$ and $\sum_i x_{ij}^2 = c$. We now show that

$$\frac{1}{p} \sum_{j=0}^{p-1} \text{var}[\hat{\beta}_j] \qquad (3.22)$$

is minimized when the columns of $\mathbf{X}$ are mutually orthogonal.

From

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & \mathbf{0}' \\ \mathbf{0} & \mathbf{C} \end{pmatrix},$$

say, we have

$$\begin{aligned} \sum_{j=0}^{p-1} \text{var}[\hat{\beta}_j] &= \text{tr}(\,\text{Var}[\hat{\boldsymbol{\beta}}]) \\ &= \sigma^2 \left[\text{tr}(\mathbf{C}^{-1}) + \frac{1}{n} \right] \\ &= \sigma^2 \sum_{j=0}^{p-1} \lambda_j^{-1}, \qquad (3.23) \end{aligned}$$

where $\lambda_0 = n$ and λ_j $(j = 1, 2, \ldots, p-1)$ are the eigenvalues of $\mathbf{C}$ (A.1.6). Now the minimum of (3.23) subject to the condition $\text{tr}(\mathbf{X}'\mathbf{X}) = n + c(p-1)$, or $\text{tr}(\mathbf{C}) = c(p-1)$, is given by $\lambda_j =$ constant, that is, $\lambda_j = c$ $(j = 1, 2, \ldots, p-1)$. Hence there exists an orthogonal matrix $\mathbf{T}$ such that $\mathbf{T}'\mathbf{C}\mathbf{T} = c\mathbf{I}_{p-1}$, or $\mathbf{C} = c\mathbf{I}_{p-1}$, so that the columns of $\mathbf{X}$ must be mutually orthogonal. $\square$

This example shows that using a particular optimality criterion, the "optimum" choice of $\mathbf{X}$ is the design matrix with mutually orthogonal columns. A related property, proved by Hotelling (see Exercises 3e, No. 3), is the following: Given any design matrix $\mathbf{X}$ such that $\mathbf{x}^{(j)\prime}\mathbf{x}^{(j)} = c_j^2$, then

$$\text{var}[\hat{\beta}_j] \geq \frac{\sigma^2}{c_j^2},$$

and the minimum is attained when $\mathbf{x}^{(j)\prime}\mathbf{x}^{(r)} = 0$ (all r, $r \neq j$) [i.e., when $\mathbf{x}^{(j)}$ is perpendicular to the other columns].

EXERCISES 3e

Min $\sum_j \lambda_j^{-1}$ st $\sum_j \lambda_j =$ const

1. Prove the statement above that the minimum is given by $\lambda_j = c$ $(j = 1, 2, \ldots, p-1)$.

is $\lambda_j =$ const by Lagrange FONC. + convexity.

2. It is required to fit a regression model of the form

$$E[Y_i] = \beta_0 + \beta_1 x_i + \beta_2 \phi(x_i) \qquad (i = 1, 2, 3),$$

where $\phi(x)$ is a second-degree polynomial. If $x_1 = -1$, $x_2 = 0$, and $x_3 = 1$, find ϕ such that the design matrix $\mathbf{X}$ has mutually orthogonal columns.

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}$$

$\phi(-1) = 1$

$\phi(\cancel{0}) = -2$

$\phi(1) = 1$

$\phi(x) = 3x^2 - 2$

3. Suppose that $\mathbf{X} = (\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p-1)}, \mathbf{x}^{(p)}) = (\mathbf{W}, \mathbf{x}^{(p)})$ has linearly independent columns.

   (a) Using A.9.5, prove that Follows by Schur so

   $$\det(\mathbf{X}'\mathbf{X}) = \det(\mathbf{W}'\mathbf{W})\left(\mathbf{x}^{(p)\prime}\mathbf{x}^{(p)} - \underbrace{\mathbf{x}^{(p)\prime}\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{x}^{(p)}}_{\geq 0}\right).$$

   (b) Deduce that

   $$(*)\ \frac{\det(\mathbf{W}'\mathbf{W})}{\det(\mathbf{X}'\mathbf{X})} \geq \frac{1}{\mathbf{x}^{(p)\prime}\mathbf{x}^{(p)}}, \leftarrow \text{so LB}$$

   and hence show that $\text{var}[\hat{\beta}_p] \geq \sigma^2(\mathbf{x}^{(p)\prime}\mathbf{x}^{(p)})^{-1}$ with equality if and only if $\mathbf{x}^{(p)\prime}\mathbf{x}^{(j)} = 0$ $(j = 0, 1, \ldots, p-1)$.

   → iff implies equality $(*)$

   (Rao [1973: p. 236])

4. What modifications in the statement of Example 3.3 proved above can be made if the term β_0 is omitted? → No mean = 0

   eqn w/ var moment Cramer's rule: consider design $[\mathbf{W} \mid (0,0,0,1)']$

5. Suppose that we wish to find the weights β_i $(i = 1, 2, \ldots, k)$ of k objects. One method is to weigh each object r times and take the average; this requires a total of kr weighings, and the variance of each average is σ^2/r (σ^2 being the variance of the weighing error). Another method is to weigh the objects in combinations; some of the objects are distributed between the two pans and weights are placed in one pan to achieve equilibrium. The regression model for such a scheme is

   $$Y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon,$$

   where $x_i = 0$, 1, or -1 according as the ith object is not used, placed in the left pan or in the right pan, ε is the weighing error (assumed to be the same for all weighings), and Y is the weight required for equilibrium (Y is regarded as negative if placed in the left pan). After n such weighing operations we can find the least squares estimates $\hat{\beta}_i$ of the weights.

   (a) Show that the estimates of the weights have maximum precision (i.e., minimum variance) when each entry in the design matrix $\mathbf{X}$ is ± 1 and the columns of $\mathbf{X}$ are mutually orthogonal.

   (b) If the objects are weighed individually, show that kn weighings are required to achieve the same precision as that given by the optimal design with n weighings.

   (Rao [1973: p. 309])

β

$\text{Var}\,\hat{\beta}_i = \dfrac{\sigma^2}{x_i^T x_i}$ when also, min when $\|x_i\|^2$ max. → all vals ± 1