STAT 8004, Assignment 2

David Dobor

January 29, 2015

Question 1

Write out the following models of elementary/intermediate statistical analysis in the matrix form

$$\mathbf{Y} = \mathbf{X}\,\beta + \epsilon$$

(a) A one-variable quadratic polynomial regression model

$$y_i = \alpha_0 + \alpha_1 x_i + \alpha_2 x_i^2 + \epsilon_i$$

for $i = 1, 2, \ldots, 5$.

(b) A two-factor ANCOVA model without interactions

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma(x_{ijk} - \bar{x}) + \epsilon_{ijk}$$

for $i = 1, 2,\ j = 1, 2,$ and $k = 1, 2.$

Answer to Question 1

(a)

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{11}^2 \\ 1 & x_{12} & x_{12}^2 \\ 1 & x_{13} & x_{13}^2 \\ 1 & x_{14} & x_{14}^2 \\ 1 & x_{15} & x_{15}^2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$$

(b)

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & x_{111} - \bar{x} \\ 1 & 1 & 0 & 1 & 0 & x_{112} - \bar{x} \\ 1 & 1 & 0 & 0 & 1 & x_{121} - \bar{x} \\ 1 & 1 & 0 & 0 & 1 & x_{122} - \bar{x} \\ 1 & 0 & 1 & 1 & 0 & x_{211} - \bar{x} \\ 1 & 0 & 1 & 1 & 0 & x_{212} - \bar{x} \\ 1 & 0 & 1 & 0 & 1 & x_{221} - \bar{x} \\ 1 & 0 & 1 & 0 & 1 & x_{222} - \bar{x} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \gamma \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \end{bmatrix}$$

Question 2

Use `eigen()` function in `R` to compute the eigenvalues and eigenvectors of

$$\mathbf{V} = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{pmatrix}$$

Then use `R` to find an "inverse square root" of this matrix. That is, find a symmetric matrix $\mathbf{W}$ such that $\mathbf{WW} = \mathbf{V}^{-1}$

Answer to Question 2

$\mathbf{V}$ is a real symmetric 3×3 matrix and thus has 3 real eigenvalues / 3 orthogonal eigenvectors in $\mathbb{R}^3$; $\mathbf{V}$ is diagonalizable. We put the eigenvectors into the columns of matrix $\mathbf{Q}$ in what follows. [1]

Also, $\mathbf{V}$'s eigenvalues happen to be positive - this is a positive definite matrix.

Thus we can decompose $\mathbf{V}$ into the product $\mathbf{Q}\Lambda\mathbf{Q}^T$, where Λ is the diagonal matrix of eigenvalues and $\mathbf{Q}$ is the orthogonal matrix containing the eigenvectors in its columns. We then compute the powers of $\mathbf{V}$ as follows:

$$\begin{aligned} \mathbf{V}^{-1} &= \mathbf{Q}\Lambda^{-1}\mathbf{Q}^T \\ \mathbf{W} &= \mathbf{Q}\Lambda^{-1/2}\mathbf{Q}^T \end{aligned}$$

[1] In a more general case when $\mathbf{V}$ is not symmetric but still has 3 linearly independent eigenvectors, we would still be able to quickly compute the powers of $\mathbf{V}$ using the decomposition $\mathbf{V} = \mathbf{S}\Lambda\mathbf{S}^{-1}$, where the columns of $\mathbf{S}$ contain the eigenvectors of $\mathbf{V}$ and Λ is the diagonal matrix of eigenvalues. The computation would be similar to what follows, except that we would have to compute the inverse of $\mathbf{S}$ which is more computationally intensive than simply transposing $\mathbf{Q}$. In the symmetric case we are fortunate that $\mathbf{Q}^T = \mathbf{Q}^{-1}$ for this orthogonal $\mathbf{Q}$.

```
# Solution to question 2

# V, the given matrix:
V <- matrix(c(3, -1, 1, -1, 5, -1, 1, -1, 3), 3)

# Q (orthogonal here) contains eigenvectors as its columns:
Q <- eigen(V)$vectors

# D contains the inverse square roots of of V's eigenvalues
# on the diagonal, zeros elsewhere:
D <- diag(1/sqrt(eigen(V)$values))

# The desired `inverse square root` matrix:
W <- Q %*% D %*% t(Q)
```

An output produced by R:

```
> options("digits"=4)  #show fewer decimals
> print(W)

         [,1]    [,2]     [,3]
[1,]  0.61404 0.05637 -0.09306
[2,]  0.05637 0.46462  0.05637
[3,] -0.09306 0.05637  0.61404
```

Question 3

Consider the matrices

$$\mathbf{A} = \begin{pmatrix} 4 & 4.001 \\ 4.001 & 4.002 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 4 & 4.001 \\ 4.001 & 4.002001 \end{pmatrix}$$

Obviously, these matrices are nearly identical. Use R and compute the determinants and inverses of these matrices. (Even though the original two matrices are nearly the same, $\mathbf{A}^{-1} \approx -3\mathbf{B}^{-1}$. This shows that small changes in the in the elements of nearly singular matrices can have big effects on some matrix operations.)

Answer to Question 3

An example R session:

```
> A <- matrix(c(4 , 4.001 , 4.001, 4.002), 2)
> B <- matrix(c(4 , 4.001 , 4.001, 4.002001), 2)

> det(A)  # nearly zer0
[1] -1e-06

> det(B)  # nearly zer0
[1] 3e-06

> Ainv <- solve(A)
> Binv <- solve(B)

> Ainv
         [,1]     [,2]
[1,] -4002000  4001000
[2,]  4001000 -4000000

> Binv
         [,1]     [,2]
[1,]  1334000 -1333667
[2,] -1333667  1333333

> 3 * Binv # check that this is approximately Ainv
         [,1]     [,2]
[1,]  4002001 -4001000
[2,] -4001000  4000000
```

Question 4

Write an `R` function to conduct projection, e.g. with name `project()`, so that the input is the given design matrix $\mathbf{X}$, and the output is the projection $\mathbf{P_X}$ for projecting a vector onto the column space of $\mathbf{X}$.

Answer to Question 4

```
project <- function(X) {
    # Computes the projection matrix onto the column space of X
    #
    # Args:
    #   X: a design matrix
    #
    # Returns:
    #   P: the projection matrix

    # library MASS' ginv() computes the generalized inverse:
    suppressPackageStartupMessages(library(MASS))

    P <- X %*% ginv(t(X) %*% X) %*% t(X)
}
```

Question 5

Consider the (non-full-rank) two-way "effect model" with interactions in the Example (d) in lecture.

(a) Determine which of the parametric functions below are estimable:

$$\alpha_1, \alpha_2 - \alpha_1, \mu + \alpha_1 + \beta_1 + \delta_{11}, \delta_{12}, \delta_{12} - \delta_{11} - (\delta_{22} - \delta_{21})$$

For those that are estimable, find $\mathbf{c}^T(\mathbf{X^T X})^{-}\mathbf{X}^T$, such that $\mathbf{c}^T(\mathbf{X^T X})^{-}\mathbf{X}^T\mathbf{Y}$ produces the estimate of $\mathbf{c}^T\beta$.

(b) For the parameter vector β written in the order used in class, consider the hypothesis $H_0 : \mathbf{C}\beta = \mathbf{0}$ for

$$\mathbf{C} = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \end{pmatrix}$$

Is this hypothesis testable? Explain.

Answer to Question 5

(a) The $\mathbf{Y} = \mathbf{X}\beta + \epsilon$ model is of the following form here:

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} \\ \epsilon_{112} \\ \epsilon_{121} \\ \epsilon_{122} \\ \epsilon_{211} \\ \epsilon_{212} \\ \epsilon_{221} \\ \epsilon_{222} \end{bmatrix}$$

In what follows we use the `project()` function from question # 4 to determine whether any vector $\mathbf{v} \in \mathbb{R}^9$ belongs to the *row* space of the design matrix $\mathbf{X}$ (or equivalently, whether it belongs to the *column* space of $\mathbf{X}^T$, denoted by $\mathbf{v} \in C(\mathbf{X}^T)$; please see the attached `R` code for computational details).

Denoting the projection matrix onto the column space of $\mathbf{X}^T$ by $\mathbf{P}$, we use the following criterion to determine whether $\mathbf{v} \in C(\mathbf{X}^T)$:

$$\mathbf{v} \in C(\mathbf{X}^T) \quad \text{iff} \quad \mathbf{P}\mathbf{v} = \mathbf{v}.$$

- α_1 : Is $(0,1,0,0,0,0,0,0,0)\begin{bmatrix}\mu\\\alpha_1\\\alpha_2\\\beta_1\\\beta_2\\\delta_{11}\\\delta_{12}\\\delta_{21}\\\delta_{22}\end{bmatrix}$ estimable?

  Or, equivalently, is $(0,1,0,0,0,0,0,0,0)$ in the row space of $\mathbf{X}$?

  Answer: **No**, because $\mathbf{Pv} \neq \mathbf{v}$.

- $\alpha_2 - \alpha_1$: Is $(0,-1,1,0,0,0,0,0,0)\begin{bmatrix}\mu\\\alpha_1\\\alpha_2\\\beta_1\\\beta_2\\\delta_{11}\\\delta_{12}\\\delta_{21}\\\delta_{22}\end{bmatrix}$ estimable?

  Or, equivalently, is $(0,-1,1,0,0,0,0,0,0)$ in the row space of $\mathbf{X}$?

  Answer: **No**, because $\mathbf{Pv} \neq \mathbf{v}$.

- $\mu + \alpha_1 + \beta_1 + \delta_{11}$: Is $(1,1,0,1,0,1,0,0,0)\begin{bmatrix}\mu\\\alpha_1\\\alpha_2\\\beta_1\\\beta_2\\\delta_{11}\\\delta_{12}\\\delta_{21}\\\delta_{22}\end{bmatrix}$ estimable?

  Or, equivalently, is $(1,1,0,1,0,1,0,0,0)$ in the row space of $\mathbf{X}$?

  Answer: **Yes**, because $\mathbf{Pv} = \mathbf{v}$.

The following linear combination of rows gives $(1,1,0,1,0,1,0,0,0)$:

$$(1,0,0,0,0,0,0,0) \times \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} = (1,1,0,1,0,1,0,0,0)$$

- δ_{12} : Is $(0,0,0,0,0,0,1,0,0) \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \end{bmatrix}$ estimable?

  Or, equivalently, is $(0,0,0,0,0,0,1,0,0)$ in the row space of $\mathbf{X}$?

  Answer: **No**, because $\mathbf{Pv} \neq \mathbf{v}$.

- $\delta_{12} - \delta_{11} - (\delta_{22} - \delta_{21})$: Is $(0,0,0,0,0,-1,1,1,-1) \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \end{bmatrix}$ estimable?

  Or, equivalently, is $(0,0,0,0,0,-1,1,1,-1)$ in the row space of $\mathbf{X}$?

  Answer: **Yes**, because $\mathbf{Pv} = \mathbf{v}$.

The row combination is:

$$(-1,0,1,0,0,1,0,-1) \times \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} = (0,0,0,0,0,-1,1,1,-1)$$

(b) For H_0 to be testable, each hypothesis shoud be testable (and also the $rank(\mathbf{C}) = 2$). In this case, since $\alpha_1 - \alpha_2$ is not esitmable, H_0 is *not* testable.

Appendix

```r
# This script helps answer questions for problem 5, assignment 2.
#
# To test whether a vector belongs to the rowspace of a design
# matrix, project it onto that space (Equivalently, project it onto
# the column space of X^T). If the projection is the same as the
# vector being  projected, then it's in the space; else it's not.

X = matrix(  # The design matrix
    c(1, 1, 1, 1, 1, 1, 1, 1,
      1, 1, 1, 1, 0, 0, 0, 0,
      0, 0, 0, 0, 1, 1, 1, 1,
      1, 1, 0, 0, 1, 1, 0, 0,
      0, 0, 1, 1, 0, 0, 1, 1,
      1, 1, 0, 0, 0, 0, 0, 0,
      0, 0, 1, 1, 0, 0, 0, 0,
      0, 0, 0, 0, 1, 1, 0, 0,
      0, 0, 0, 0, 0, 0, 1, 1),
    nrow=8,
    ncol=9)

# load the project() function from problem 4:
source("project.R")
P <- project(t(X))

#-------------------------------------------------
# question: is alpha1 estimable?
alpha1 <- matrix(c(0, 1, 0, 0, 0, 0, 0, 0, 0))
print(P %*% alpha1)         #not estimable

#-------------------------------------------------
# question: is alpha2 - alpha1 estimable:
alpha2_1 <- matrix(c(0, -1, 1, 0, 0, 0, 0, 0, 0))
print(P %*% alpha2_1)       # not estimable

#-------------------------------------------------
mu_and_others <- matrix(c(1, 1, 0, 1, 0, 1, 0, 0, 0))
print(P %*% mu_others)      # estimable

#-------------------------------------------------
delta1_2 <- matrix(c(0, 0, 0, 0, 0, 0, 1, 0, 0))
P %*% delta1_2

#-------------------------------------------------
deltas  <- matrix(c(0, 0, 0, 0, 0, -1, 1, 1, -1))
P %*% deltas                # estimable
```