A Relationship Between the One-Way MANOVA Test Statistic and the Hotelling Lawley Trace Test Statistic

Hasthika S. Rupasinghe Arachchige Don

Correspondence: Hasthika S. Rupasinghe Arachchige Don, Department of Mathematical Sciences, Appalachian State University, Boone, NC., 28607, USA.

Received: August 23, 2018 Accepted: September 28, 2018 Online Published: October 15, 2018

doi:10.5539/ijsp.v7n6p124 URL: https://doi.org/10.5539/ijsp.v7n6p124

Abstract

The One-Way MANOVA model is a special case of the multivariate linear model, and this paper shows that the One-Way MANOVA test statistic and the Hotelling Lawley trace test statistic are equivalent if the design matrix is carefully chosen.

Keywords: ANOVA, Linear Moldel

1. Introduction

We want to show that the One-Way MANOVA test statistic and the Hotelling Lawley trace test statistic are equivalent for a carefully chosen full rank design matrix. First we will describe the MANOVA model, and then the One-Way MANOVA model. The notation in this paper follows that used in Olive (2017) and closely follows Rupasinghe Arachchige Don (2017).

1.1 MANOVA

Multivariate analysis of variance (MANOVA) is analogous to an ANOVA, but there is more than one dependent variable. ANOVA tests for the difference in means between two or more groups, while MANOVA tests for the difference in two or more vectors of means.

The multivariate analysis of variance (MANOVA) model $\boldsymbol{y}_i = \boldsymbol{B}^T\boldsymbol{x}_i + \boldsymbol{\epsilon}_i$ for $i = 1, \ldots, n$ has $m \geq 2$ response variables $Y_1, \ldots Y_m$ and p predictor variables $x_1, x_2, \ldots, x_p$. The ith case is $(\boldsymbol{x}_i^T, \boldsymbol{y}_i^T) = (x_{i1}, ..., x_{ip}, Y_{i1}, ..., Y_{im})$. If a constant $x_{i1} = 1$ is in the model, then x_{i1} could be omitted from the case.

For the MANOVA model predictors are indicator variables. Sometimes the trivial predictor $\mathbf{1}$ is also in the model. The MANOVA model in matrix form is $\boldsymbol{Z} = \boldsymbol{XB} + \boldsymbol{E}$ and has $E(\boldsymbol{\epsilon}_k) = \mathbf{0}$ and $Cov(\boldsymbol{\epsilon}_k) = \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = (\sigma_{ij})$ for $k = 1, \ldots, n$. Also $E(\mathbf{e}_i) = \mathbf{0}$ while $Cov(\mathbf{e}_i, \mathbf{e}_j) = \sigma_{ij}\boldsymbol{I}_n$ for $i, j = 1, \ldots, m$. Then $\boldsymbol{B}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ are unknown matrices to be estimated.

$$\boldsymbol{Z} = \begin{pmatrix} Y_{1,1} & Y_{1,2} & \cdots & Y_{1,m} \\ Y_{2,1} & Y_{2,2} & \cdots & Y_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{n,1} & Y_{n,2} & \cdots & Y_{n,m} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 & \cdots & \boldsymbol{Y}_m \end{pmatrix} = \begin{pmatrix} \boldsymbol{y}_1^T \\ \vdots \\ \boldsymbol{y}_n^T \end{pmatrix}.$$

The $n \times p$ matrix $\boldsymbol{X}$ is not necessarily of full rank p, and

$$\boldsymbol{X} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1^T \\ \vdots \\ \boldsymbol{x}_n^T \end{pmatrix}$$

where often $\mathbf{v}_1 = \mathbf{1}$.

The $p \times m$ coefficient matrix is

$$\boldsymbol{B} = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,m} \\ \beta_{2,1} & \beta_{2,2} & \cdots & \beta_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{p,1} & \beta_{p,2} & \cdots & \beta_{p,m} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_m \end{pmatrix}.$$

The $n \times m$ error matrix is

$$\boldsymbol{E} = \begin{pmatrix} \epsilon_{1,1} & \epsilon_{1,2} & \cdots & \epsilon_{1,m} \\ \epsilon_{2,1} & \epsilon_{2,2} & \cdots & \epsilon_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{n,1} & \epsilon_{n,2} & \cdots & \epsilon_{n,m} \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_m \end{pmatrix} = \begin{pmatrix} \boldsymbol{\epsilon}_1^T \\ \vdots \\ \boldsymbol{\epsilon}_n^T \end{pmatrix}.$$

Each response variable in a MANOVA model follows an ANOVA model $\boldsymbol{Y}_j = \boldsymbol{X}\boldsymbol{\beta}_j + \boldsymbol{e}_j$ for $j = 1, \ldots, m$, where it is assumed that $E(\boldsymbol{e}_j) = \boldsymbol{0}$ and $Cov(\boldsymbol{e}_j) = \sigma_{jj}\boldsymbol{I}_n$.

MANOVA models are often fit by least squares. The least squares estimator $\hat{\boldsymbol{B}}$ of $\boldsymbol{B}$ is

$$\hat{\boldsymbol{B}} = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^- \boldsymbol{X}^T\boldsymbol{Z} = \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 & \hat{\boldsymbol{\beta}}_2 & \ldots & \hat{\boldsymbol{\beta}}_m \end{pmatrix}$$

where $\left(\boldsymbol{X}^T\boldsymbol{X}\right)^-$ is a generalized inverse of $\boldsymbol{X}^T\boldsymbol{X}$. If $\boldsymbol{X}$ has a full rank then $\left(\boldsymbol{X}^T\boldsymbol{X}\right)^- = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}$ and $\hat{\boldsymbol{B}}$ is unique.

The predicted values or fitted values are

$$\hat{\boldsymbol{Z}} = \boldsymbol{X}\hat{\boldsymbol{B}} = \begin{pmatrix} \hat{\boldsymbol{Y}}_1 & \hat{\boldsymbol{Y}}_2 & \cdots & \hat{\boldsymbol{Y}}_m \end{pmatrix} = \begin{pmatrix} \hat{Y}_{1,1} & \hat{Y}_{1,2} & \cdots & \hat{Y}_{1,m} \\ \hat{Y}_{2,1} & \hat{Y}_{2,2} & \cdots & \hat{Y}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{Y}_{n,1} & \hat{Y}_{n,2} & \cdots & \hat{Y}_{n,m} \end{pmatrix}.$$

The residuals are $\hat{\boldsymbol{E}} = \boldsymbol{Z} - \hat{\boldsymbol{Z}} = \boldsymbol{Z} - \boldsymbol{X}\hat{\boldsymbol{B}}$.

Finally,

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = \frac{(\boldsymbol{Z} - \hat{\boldsymbol{Z}})^T(\boldsymbol{Z} - \hat{\boldsymbol{Z}})}{n - p} = \frac{\hat{\boldsymbol{E}}^T\hat{\boldsymbol{E}}}{n - p}.$$

1.2 One Way MANOVA

Assume that there are independent random samples of size n_i from p different populations, or n_i cases are randomly assigned to p treatment groups. Let $n = \sum_{i=1}^{p} n_i$ be the total sample size. Also assume that m response variables $\boldsymbol{y}_{ij} = (Y_{ij1}, \ldots, Y_{ijm})^T$ are measured for the ith treatment group and the jth case. Assume $E(\boldsymbol{y}_{ij}) = \boldsymbol{\mu}_i$ and $Cov(\boldsymbol{y}_{ij}) = \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$.

The one way MANOVA is used to test $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_p$. Note that if $m = 1$ the one way MANOVA model becomes the one way ANOVA model. One might think that performing m ANOVA tests is sufficient to test the above hypotheses. But the separate ANOVA tests would not take the correlation between the m variables into account. On the other hand the MANOVA test will take the correlation into account.

Let $\bar{\boldsymbol{y}} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} \boldsymbol{y}_{ij}/n$ be the overall mean. Let $\bar{\boldsymbol{y}}_i = \sum_{j=1}^{n_i} \boldsymbol{y}_{ij}/n_i$. Several $m \times m$ matrices will be useful. Let $\boldsymbol{S}_i$ be the sample covariance matrix corresponding to the ith treatment group. Then the within sum of squares and cross products matrix is $\boldsymbol{W} = (n_1 - 1)\boldsymbol{S}_1 + \cdots + (n_p - 1)\boldsymbol{S}_p = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (\boldsymbol{y}_{ij} - \bar{\boldsymbol{y}}_i)(\boldsymbol{y}_{ij} - \bar{\boldsymbol{y}}_i)^T$. Then $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} = \boldsymbol{W}/(n - p)$. The treatment or between sum of squares and cross products matrix is

$$\boldsymbol{B}_T = \sum_{i=1}^{p} n_i(\bar{\boldsymbol{y}}_i - \bar{\boldsymbol{y}})(\bar{\boldsymbol{y}}_i - \bar{\boldsymbol{y}})^T.$$

The total corrected (for the mean) sum of squares and cross products matrix is $\boldsymbol{T} = \boldsymbol{B}_T + \boldsymbol{W} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} (\boldsymbol{y}_{ij} - \bar{\boldsymbol{y}})(\boldsymbol{y}_{ij} - \bar{\boldsymbol{y}})^T$. Note that $\boldsymbol{S} = \boldsymbol{T}/(n - 1)$ is the usual sample covariance matrix of the $\boldsymbol{y}_{ij}$ if it is assumed that all n of the $\boldsymbol{y}_{ij}$ are iid so that the $\boldsymbol{\mu}_i \equiv \boldsymbol{\mu}$ for $i = 1, ..., p$.

The one way MANOVA model is $\boldsymbol{y}_{ij} = \boldsymbol{\mu}_i + \boldsymbol{\epsilon}_{ij}$ where $\boldsymbol{\epsilon}_{ij}$ are iid with $E(\boldsymbol{\epsilon}_{ij}) = \boldsymbol{0}$ and $Cov(\boldsymbol{\epsilon}_{ij}) = \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$. The summary one way MANOVA table is shown bellow.

Source	matrix	df
Treatment or Between	$\boldsymbol{B}_T$	$p - 1$
Residual or Error or Within	$\boldsymbol{W}$	$n - p$
Total (Corrected)	$\boldsymbol{T}$	$n - 1$

There are three commonly used test statistics to test the above hypotheses. Namely,

1. Hotelling Lawley trace statistic: $U = tr(\boldsymbol{B}_T \boldsymbol{W}^{-1}) = tr(\boldsymbol{W}^{-1}\boldsymbol{B}_T)$.

2. Wilks' lambda: $\Lambda = \dfrac{|\boldsymbol{W}|}{|\boldsymbol{B}_T + \boldsymbol{W}|}$.

3. Pillai's trace statistic: $\boldsymbol{V} = tr(\boldsymbol{B}_T \boldsymbol{T}^{-1}) = tr(\boldsymbol{T}^{-1}\boldsymbol{B}_T)$.

If the $\boldsymbol{y}_{ij} - \boldsymbol{\mu}_j$ are iid with common covariance matrix $\boldsymbol{\Sigma_\epsilon}$, and if H_0 is true, then under regularity conditions Fujikoshi (2002) showed

1. $(n - m - p - 1)U \xrightarrow{D} \chi^2_{m(p-1)}$,

2. $-[n - 0.5(m + p - 2)]log(\Lambda) \xrightarrow{D} \chi^2_{m(p-1)}$, and

3. $(n - 1)V \xrightarrow{D} \chi^2_{m(p-1)}$.

Note that the common covariance matrix assumption implies that each of the p treatment groups or populations has the same covariance matrix $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma_\epsilon}$ for $i = 1, ..., p$, an extremely strong assumption. Kakizawa (2009) and Olive et al. (2015) show that similar results hold for the multivariate linear model. The common covariance matrix assumption, $\text{Cov}(\boldsymbol{\epsilon}_k) = \boldsymbol{\Sigma_\epsilon}$ for $k = 1, ..., n$, is often reasonable for the multivariate linear regression model.

1.3 Hotelling Lawley Trace Test

Hotelling Lawley trace test statistic Hotelling (1951); Lawley (1938), and the asymptotic distribution $(n-m-p-1)U \xrightarrow{D} \chi^2_{m(p-1)}$ by Fujikoshi (2002) are widely used. Olive et al. (2015) explains the large sample theory of the Wilks' Λ, Pillai's trace, and Hotelling Lawley trace test statistics and gives two theorems to show that the Hotelling Lawley test generalizes the usual partial F test for $m = 1$ response variable to $m \geq 1$ response variables.

2. Method

2.1 A Relationship Between the One-Way MANOVA Test and the Hotelling Lawley Trace Test

An alternative method for One-Way MANOVA is to use the model $\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{E}$ where

$$\boldsymbol{Y}_{ij} = \begin{pmatrix} Y_{ij1} \\ \vdots \\ Y_{ijm} \end{pmatrix} = \boldsymbol{\mu}_i + \boldsymbol{e}_{ij}, \text{ and } \qquad \text{E[Y}_{\text{ij}}] = \boldsymbol{\mu}_{\text{i}} = \begin{pmatrix} \mu_{ij1} \\ \vdots \\ \mu_{ijm} \end{pmatrix}$$

for $i = 1, \ldots, p$ and $j = 1, \ldots, n_i$ Then $\boldsymbol{X}$ is a full rank where the ith column of $\boldsymbol{X}$ is an indicator for group $i - 1$ for $i = 2, \ldots, p$.

$$\boldsymbol{X} = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{1}$$

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{\mu}_p^T \\ (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_p)^T \\ \vdots \\ (\boldsymbol{\mu}_{p-1} - \boldsymbol{\mu}_p)^T \end{pmatrix}$$ and let $\boldsymbol{L} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_{p-1} \end{pmatrix}$. Note that $\boldsymbol{Y}_{ij}^T = \boldsymbol{\mu}_i^T + \boldsymbol{e}_{ij}^T$.

Then

$$\boldsymbol{X}^T\boldsymbol{X} = \begin{pmatrix} n & n_1 & n_2 & \cdots & n_{p-1} \\ n_1 & n_1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ n_{p-2} & 0 & \cdots & n_{p-2} & 0 \\ n_{p-1} & 0 & \cdots & 0 & n_{p-1} \end{pmatrix} \tag{2}$$

and

$$\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1} = \frac{1}{n_p}\begin{pmatrix} 1 & -1 & -1 & \cdots & -1 \\ -1 & 1+\frac{n_p}{n_1} & 1 & \cdots & 1 \\ \vdots & & \ddots & & \vdots \\ -1 & 1 & \cdots & 1+\frac{n_p}{n_{p-2}} & 1 \\ -1 & 1 & \cdots & 1 & 1+\frac{n_p}{n_{p-1}} \end{pmatrix}. \tag{3}$$

Then the least squares estimator $\hat{\boldsymbol{B}}$ of $\boldsymbol{B}$,

$$\hat{\boldsymbol{B}} = \begin{pmatrix} \bar{\mathbf{y}}_p^T \\ (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_p)^T \\ \vdots \\ (\bar{\mathbf{y}}_{p-1} - \bar{\mathbf{y}}_p)^T \end{pmatrix}, \quad \text{and} \quad \boldsymbol{L}\hat{\boldsymbol{B}} = \begin{pmatrix} (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_p)^T \\ (\bar{\mathbf{y}}_2 - \bar{\mathbf{y}}_p)^T \\ \vdots \\ (\bar{\mathbf{y}}_{p-1} - \bar{\mathbf{y}}_p)^T \end{pmatrix}.$$

Then $\boldsymbol{L}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{L}^T$ becomes

$$\boldsymbol{L}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{L}^T = \frac{1}{n_p}\begin{pmatrix} 1+\frac{n_p}{n_1} & 1 & 1 & \cdots & 1 \\ 1 & 1+\frac{n_p}{n_2} & 1 & \cdots & 1 \\ \vdots & & \ddots & & \vdots \\ 1 & 1 & \cdots & 1 & 1+\frac{n_p}{n_{p-1}} \end{pmatrix}. \tag{4}$$

It can be shown that the inverse of the above matrix is

$$\left[\boldsymbol{L}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{L}^T\right]^{-1} = \frac{1}{n}\begin{pmatrix} n_1(n-n_1) & -n_1n_2 & -n_1n_3 & \cdots & -n_1n_{p-1} \\ -n_1n_2 & n_2(n-n_2) & -n_2n_3 & \cdots & -n_2n_{p-1} \\ \vdots & & \ddots & & \vdots \\ -n_1n_{p-1} & -n_2n_{p-1} & \cdots & & n_{p-1}(n-n_{p-1}) \end{pmatrix}.$$

For convenience, write $\left[\boldsymbol{L}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{L}^T\right]^{-1} =$

$$\frac{1}{n}\begin{pmatrix} -n_1^2 & -n_1n_2 & -n_1n_3 & \cdots & -n_1n_{p-1} \\ -n_1n_2 & -n_2^2 & -n_2n_3 & \cdots & -n_2n_{p-1} \\ \vdots & & \ddots & & \vdots \\ -n_1n_{p-1} & -n_2n_{p-1} & \cdots & & -n_{p-1}^2 \end{pmatrix} + \begin{pmatrix} n_1 & 0 & 0 & \cdots & 0 \\ 0 & n_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & n_{p-1} \end{pmatrix}.$$

Then

$$\left(\boldsymbol{L}\hat{\boldsymbol{B}}\right)^T \left[\boldsymbol{L}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{L}^T\right]^{-1}\left(\boldsymbol{L}\hat{\boldsymbol{B}}\right) =$$

$$-\frac{1}{n}\sum_{i=1}^{p-1}\sum_{j=1}^{p-1} n_i n_j(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_j - \bar{\mathbf{y}}_p)^T + \sum_{i=1}^{p-1} n_i(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)^T = \boldsymbol{H}.$$

Let $\boldsymbol{X}$ be as in (1). Then the multivariate linear regression Hotelling Lawley test statistic for testing $H_0 : \boldsymbol{LB} = \boldsymbol{0}$ versus $H_0 : \boldsymbol{LB} \neq \boldsymbol{0}$ has

$$U = tr(\boldsymbol{W}^{-1}\boldsymbol{H}).$$

One-Way MANOVA is used to test $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_p$. The One-Way MANOVA Hotelling Lawley test statistic for testing for above hypotheses is

$$U = tr(\boldsymbol{W}^{-1}\boldsymbol{B}_T)$$

where

$$\boldsymbol{W} = (n-p)\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\epsilon}} \quad \text{and} \quad \boldsymbol{B}_{\mathrm{T}} = \sum_{\mathrm{i}=1}^{\mathrm{p}} \mathrm{n_i}(\bar{\mathbf{y}}_{\mathrm{i}} - \bar{\mathbf{y}})(\bar{\mathbf{y}}_{\mathrm{i}} - \bar{\mathbf{y}})^{\mathrm{T}}.$$

Theorem 1. *The One-Way MANOVA and the multivariate linear regression Hotelling Lawley trace test statistics are the same for the design matrix as in (1).*

To show that the above two test statistics are equal, it is sufficient to prove that $\boldsymbol{H} = \boldsymbol{B}_T$. First we will prove two special cases and then give the proof for the theorem.

Proof. **Special case I:** $p = 2$ (Two group case)

Consider $\boldsymbol{H}$.

$\boldsymbol{H} = -\frac{1}{n}n_1 n_1(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T + n_1(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T$. Since $n = n_1 + n_2$,

$\boldsymbol{H} = -\frac{1}{n}(nn_1 - n_1 n_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T + n_1(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T$

$\boldsymbol{H} = -n_1(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T + \frac{n_1 n_2}{n}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T + n_1(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T$

$\boldsymbol{H} = \frac{n_1 n_2}{n}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T$.

Now consider $\boldsymbol{B}_T$ with $p = 2$.

Note that $\bar{\mathbf{y}} = (n_1\bar{\mathbf{y}}_1 + n_2\bar{\mathbf{y}}_2)/n$ and

$\boldsymbol{B}_T = n_1(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}})(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}})^T + n_2(\bar{\mathbf{y}}_2 - \bar{\mathbf{y}})(\bar{\mathbf{y}}_2 - \bar{\mathbf{y}})^T$

$\boldsymbol{B}_T = \frac{n_1}{n^2}(n\bar{\mathbf{y}}_1 - n_1\bar{\mathbf{y}}_1 - n_2\bar{\mathbf{y}}_2)(n\bar{\mathbf{y}}_1 - n_1\bar{\mathbf{y}}_1 - n_2\bar{\mathbf{y}}_1)^T +$
$\frac{n_2}{n^2}(n\bar{\mathbf{y}}_2 - n_1\bar{\mathbf{y}}_1 - n_2\bar{\mathbf{y}}_2)(n\bar{\mathbf{y}}_2 - n_1\bar{\mathbf{y}}_1 - n_2\bar{\mathbf{y}}_2)^T$

$\boldsymbol{B}_T = \frac{n_1 n_2^2}{n^2}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T + \frac{n_1^2 n_2}{n^2}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T$

$\boldsymbol{B}_T = \frac{n_1 n_2}{n}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)^T$.

Therefore $\boldsymbol{B}_T = \boldsymbol{H}$ when $p = 2$.

□

Proof. **Special case II:** $n_i = n_1 \ \forall i = 1, \ldots, p$

$$\boldsymbol{H} = -\frac{1}{n}\sum_{i=1}^{p-1}\sum_{j=1}^{p-1} n_i n_j(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_j - \bar{\mathbf{y}}_p)^T + \sum_{i=1}^{p-1} n_i(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)^T.$$

Note that the i, j running from 1 through $p-1$ and i, j running from 1 through p would yield the same $\boldsymbol{H}$. Therefore $\boldsymbol{H}$ can be written as

$$\boldsymbol{H} = -\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_j - \bar{\mathbf{y}}_p)^T + \sum_{i=1}^{p} n_i(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)^T.$$

Now consider the double sum in $\boldsymbol{H}$. Note that $n = n_1 p$ and

$$-\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j (\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_j - \bar{\mathbf{y}}_p)^T = \frac{-n_1^2}{n_1 p}\sum_{i=1}^{p}\sum_{j=1}^{p}\left(\bar{\mathbf{y}}_i\bar{\mathbf{y}}_j^T - \bar{\mathbf{y}}_i\bar{\mathbf{y}}_p^T - \bar{\mathbf{y}}_p\bar{\mathbf{y}}_j^T + \bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T\right)$$

$$= \frac{n_1}{p}\left[-\sum_{i=1}^{p}\sum_{j=1}^{p}\left(\bar{\mathbf{y}}_i\bar{\mathbf{y}}_j^T\right) + p\left(\sum_{i=1}^{p}\bar{\mathbf{y}}_i\right)\bar{\mathbf{y}}_p^T + p\bar{\mathbf{y}}_p\left(\sum_{j=1}^{p}\bar{\mathbf{y}}_j^T\right) - p^2\bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T\right]. \tag{5}$$

Now consider the rest of $\boldsymbol{H}$,

$$n_1\sum_{i=1}^{p}(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)^T = n_1\sum_{i=1}^{p}\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - n_1\left(\sum_{i=1}^{p}\bar{\mathbf{y}}_i\right)\bar{\mathbf{y}}_p^T - n_1\bar{\mathbf{y}}_p\left(\sum_{i=1}^{p}\bar{\mathbf{y}}_i^T\right) + n_1 p\bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T. \tag{6}$$

Therefore by (5) and (6), it is clear that

$$\boldsymbol{H} = n_1\sum_{i=1}^{p}\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - \frac{n_1}{p}\sum_{i=1}^{p}\sum_{j=1}^{p}\bar{\mathbf{y}}_i\bar{\mathbf{y}}_j^T. \tag{7}$$

Now consider

$$\boldsymbol{B}_T = n_1\sum_{i=1}^{p}(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})^T. \tag{8}$$

Let

$$\bar{\boldsymbol{Y}} = \begin{pmatrix} \bar{\mathbf{y}}_1^T \\ \bar{\mathbf{y}}_2^T \\ \vdots \\ \bar{\mathbf{y}}_p^T \end{pmatrix}. \quad \text{Then} \quad \boldsymbol{B}_{\mathrm{T}} = \mathrm{n}_1\left[\bar{\boldsymbol{Y}}^{\mathrm{T}}\bar{\boldsymbol{Y}} - \frac{1}{\mathrm{p}}\bar{\boldsymbol{Y}}^{\mathrm{T}}\mathbf{1}\mathbf{1}^{\mathbf{T}}\bar{\boldsymbol{Y}}\right].$$

Therefore, $\boldsymbol{B}_T$ becomes

$$\boldsymbol{B}_T = n_1\sum_{i=1}^{p}\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - \frac{n_1}{p}\sum_{i=1}^{p}\sum_{j=1}^{p}\bar{\mathbf{y}}_i\bar{\mathbf{y}}_j^T. \tag{9}$$

From (8) and (9) $\boldsymbol{B}_T = \boldsymbol{H}$.

□

Proof. **General case:**

$$\boldsymbol{H} = -\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_j - \bar{\mathbf{y}}_p)^T + \sum_{i=1}^{p} n_i(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)^T.$$

First consider the double sum in $\boldsymbol{H}$.

$$-\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_j - \bar{\mathbf{y}}_p)^T =$$

$$-\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j\bar{\mathbf{y}}_i\bar{\mathbf{y}}_j^T + \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j\bar{\mathbf{y}}_i\bar{\mathbf{y}}_p^T + \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j\bar{\mathbf{y}}_p\bar{\mathbf{y}}_j^T - \frac{1}{n}\bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T\sum_{i=1}^{p}\sum_{j=1}^{p} n_i n_j \tag{10}$$

$$-\frac{1}{n}\sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\sum_{j=1}^{p} n_j\bar{\mathbf{y}}_j^T + \frac{1}{n}\sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\sum_{j=1}^{p} n_j\bar{\mathbf{y}}_p^T + \frac{1}{n}\bar{\mathbf{y}}_p\sum_{i=1}^{p} n_i\sum_{j=1}^{p} n_j\bar{\mathbf{y}}_j^T - \frac{1}{n}\bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T n^2$$

$$-\frac{1}{n}n\bar{\mathbf{y}}n\bar{\mathbf{y}}^T + \frac{1}{n}\sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i n\bar{\mathbf{y}}_p^T + \frac{1}{n}\bar{\mathbf{y}}_p n\sum_{j=1}^{p} n_j\bar{\mathbf{y}}_j^T - n\bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T$$

$$-n\bar{\mathbf{y}}\bar{\mathbf{y}}^T + \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}_p^T + \bar{\mathbf{y}}_p\sum_{j=1}^{p} n_j\bar{\mathbf{y}}_j^T - n\bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T. \tag{11}$$

Now consider the rest of $\boldsymbol{H}$,

$$\sum_{i=1}^{p} n_i(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)(\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_p)^T = \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}_p^T - \bar{\mathbf{y}}_p\sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i^T + n\bar{\mathbf{y}}_p\bar{\mathbf{y}}_p^T. \tag{12}$$

Therefore by (11) and (12)

$$\boldsymbol{H} = \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - n\bar{\mathbf{y}}\bar{\mathbf{y}}^T. \tag{13}$$

Now consider

$$\boldsymbol{B}_T = \sum_{i=1}^{p} n_i(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})^T$$

$$\boldsymbol{B}_T = \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}^T - \bar{\mathbf{y}}\sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i^T + \bar{\mathbf{y}}\bar{\mathbf{y}}^T\sum_{i=1}^{p} n_i$$

$$\boldsymbol{B}_T = \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - n\bar{\mathbf{y}}\bar{\mathbf{y}}^T - \bar{\mathbf{y}}n\bar{\mathbf{y}}^T + n\bar{\mathbf{y}}\bar{\mathbf{y}}^T$$

$$\boldsymbol{B}_T = \sum_{i=1}^{p} n_i\bar{\mathbf{y}}_i\bar{\mathbf{y}}_i^T - n\bar{\mathbf{y}}_i\bar{\mathbf{y}}^T. \tag{14}$$

(13) and (14) proves that $\boldsymbol{H} = \boldsymbol{B}_T$.

□

2.2 Cell Means Model

We can get the same result for the cell means model which is defined for $\boldsymbol{X}$ and $\boldsymbol{B}$ given below.

$$\boldsymbol{X} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} \boldsymbol{\mu}_1^T \\ \vdots \\ \boldsymbol{\mu}_p^T \end{pmatrix} \quad \textit{and} \qquad \boldsymbol{L} = \begin{pmatrix} \boldsymbol{I}_{p-1} & -\mathbf{1} \end{pmatrix}$$

$$\hat{\boldsymbol{B}} = \begin{pmatrix} \bar{\mathbf{y}}_1^T \\ \vdots \\ \bar{\mathbf{y}}_p^T \end{pmatrix}, \qquad \boldsymbol{L}\hat{\boldsymbol{B}} = \begin{pmatrix} (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_p)^T \\ (\bar{\mathbf{y}}_2 - \bar{\mathbf{y}}_p)^T \\ \vdots \\ (\bar{\mathbf{y}}_{p-1} - \bar{\mathbf{y}}_p)^T \end{pmatrix}.$$

Then $\boldsymbol{X}^T\boldsymbol{X} = diag\left(n_1, \ldots, n_{p-1}\right)$ and $(\boldsymbol{X}^T\boldsymbol{X})^{-1} = diag\left(\frac{1}{n_1}, \ldots, \frac{1}{n_{p-1}}\right)$.

Then $\boldsymbol{L}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{L}^T$ becomes

$$\boldsymbol{L}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{L}^T = \frac{1}{n_p}\begin{pmatrix} 1+\frac{n_p}{n_1} & 1 & 1 & \cdots & 1 \\ 1 & 1+\frac{n_p}{n_2} & 1 & \cdots & 1 \\ \vdots & & \ddots & & \vdots \\ 1 & 1 & \cdots & 1 & 1+\frac{n_p}{n_{p-1}} \end{pmatrix}. \tag{15}$$

Notice that the matrix equation (15) is exactly same as (4). This is an indication that Theorem 1 does not depend on the full rank design matrix.

3. Conclusions

This work mathematically proved that the One-Way MANOVA test statistic and the Hotelling Lawley trace test statistic are in fact the same. The proof consisted of two special cases and the general case. This result indicates that one can use the One-Way MANOVA test statistic and the Hotelling Lawley trace test statistic alternatively if the design matrix is carefully chosen.

Acknowledgements

The author thanks Dr. David J. Olive and Dr. Lasanthi C. R. Pelawa Watagoda for some comments on this paper.

References

Fujikoshi, Y. (2002). Asymptotic expansions for the distributions of multivariate basic statistics and one-way MANOVA tests under nonnormality. *Journal of Statistical Planning and Inference, 108*(1), 263 - 282. https://doi.org/10.1016/S0378-3758(02)00313-0

Hotelling, H. (1951). *A generalized T test and measure of multivariate dispersion.* In J. Neyman (Ed.), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (pp. 23-41). Berkeley: University of California Press.

Kakizawa, Y. (2009). Third-order power comparisons for a class of tests for multivariate linear hypothesis under general distributions. *Journal of Multivariate Analysis, 100*(3), 473 - 496. https://doi.org/10.1016/j.jmva.2008.06.002

Lawley, D. N. (1938). A generalization of Fisher's z-test. *Biometrika, 30*, 180 - 187. https://doi.org/10.1093/biomet/30.1-2.180

Olive, D. J. (2017). *Robust Multivariate Analysis.* Springer International Publishing. https://doi.org/10.1007/978-3-319-68253-2

Olive, D. J., Pelawa, W. L. C. R., & Rupasinghe, A. D. H. S. (2015). Visualizing and Testing the Multivariate Linear Regression Model. *International Journal of Statistics and Probability, 4*(1), 126. https://doi.org/10.5539/ijsp.v4n1p126

Rupasinghe Arachchige Don, H. S. (2017). *Bootstrapping Analogs of the One Way MANOVA Test. (Ph.D. Thesis)*, Southern Illinois University, USA, at (https://opensiuc.lib.siu.edu/dissertations/1425/).