Displays for Statistics 5401/8401

Lecture 17

October 14, 2005

Christopher Bingham, Instructor

612-625-1024, kb@umn.edu
372 Ford Hall

Class Web Page

http://www.stat.umn.edu/~kb/classes/5401

One-way MANOVA Model

- *Data* consist of g *independent random samples* $\{\mathbf{y}_{ij}\}_{1\leq i\leq n_j}$ of sizes n_1, ..., n_g from g groups or populations
- The additive linear model is
  $\mathbf{y}_{ij} = (\boldsymbol{\mu} + \boldsymbol{\alpha}_j) + \{\boldsymbol{\varepsilon}_{ij}\}$, $j = 1,\ldots,g$
  $\mathbf{y}_{ij}$, $\boldsymbol{\mu}$, $\boldsymbol{\alpha}_j$ and $\boldsymbol{\varepsilon}_{ij}$ all $p \times 1$ and $E[\boldsymbol{\varepsilon}_{ij}] = \mathbf{0}$.

The other assumptions are:

- Equal variance matrices
  $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \ldots = \boldsymbol{\Sigma}_g = \boldsymbol{\Sigma} = [\sigma_{\ell m}]$
  with $\boldsymbol{\Sigma}_j = [\sigma_{\ell m}^{(j)}] = V[\boldsymbol{\varepsilon}]$ for group j.

  Equality of $\boldsymbol{\Sigma}$'s is strong condition:
  1. *Equal variances* among groups
  $\sigma_{\ell\ell}^{(1)} = \sigma_{\ell\ell}^{(2)} = \ldots = \sigma_{\ell\ell}^{(g)} = \sigma_{\ell\ell}$, $\ell = 1,\ldots,p$
  2. *Equal correlations* among groups
  $\rho_{\ell m}^{(1)} = \rho_{\ell m}^{(2)} = \ldots = \rho_{\ell m}^{(g)} = \rho_{\ell m}$, $1\leq\ell\neq m\leq p$
- Exact small sample inference requires that $\boldsymbol{\varepsilon}$ is $N_p(\mathbf{0}, \boldsymbol{\Sigma})$.

You can also parametrize the one-way MANOVA model in terms of group mean vectors

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu} + \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\mu}_g = \boldsymbol{\mu} + \boldsymbol{\alpha}_g$$

instead of a grand mean $\boldsymbol{\mu}$ and effects $\boldsymbol{\alpha}_j$:

$$\mathbf{y}_{ij} = \boldsymbol{\mu}_j + \boldsymbol{\varepsilon}_{ij}$$

$\mathbf{y}_{ij}$, $\boldsymbol{\mu}_j$, $\boldsymbol{\varepsilon}_{ij}$ all $p \times 1$.

MANACOVA - Multivariate ANACOVA

$$\mathbf{y}_{ij} = \boldsymbol{\mu} + Z_{ij,1}\boldsymbol{\beta}_1 + Z_{ij,2}\boldsymbol{\beta}_2 + \ldots + Z_{ij,k}\boldsymbol{\beta}_k + \boldsymbol{\alpha}_j + \boldsymbol{\varepsilon}_{ij}$$

- The Z's are covariates
- The $\boldsymbol{\beta}$'s don't differ among groups.
- $\boldsymbol{\Sigma} = V[\boldsymbol{\varepsilon}]$ is constant and doesn't depend on group or any of the Z_j's.

The standard approach to multivariate linear models assumes the same model for every variable.

Regression:

$$\mathbf{y}_i = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 Z_{i1} + \ldots + \boldsymbol{\beta}_k Z_{ik} + \boldsymbol{\varepsilon}_i$$

is equivalent to p univariate regressions

$$y_{i\ell} = \beta_{0\ell} + \beta_{1\ell} Z_{i1} + \ldots + \beta_{k\ell} Z_{ik} + \varepsilon_{i\ell}$$

$$\ell = 1, \ldots p$$

all with the same predictors.

2 factor MANOVA

$$\mathbf{y}_{ij} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + (\boldsymbol{\alpha\beta})_{ij} + \boldsymbol{\varepsilon}_{ij}$$

is equivalent to p univariate ANOVA models

$$y_{ij\ell} = \mu + \alpha_{i\ell} + \beta_{j\ell} + (\alpha\beta)_{ij\ell} + \varepsilon_{ij\ell}$$

all with the both main effects and interaction.

The situation when different variables have different models is called Seemingly Unrelated Regression or **SUR**. The best estimates are *not* least squares.

For all these models, the $\boldsymbol{\varepsilon}_i$'s or $\boldsymbol{\varepsilon}_{ij}$'s are assumed to have these properties in <u>decreasing order of importance</u> (most important first)

1 $E[\boldsymbol{\varepsilon}] = \mathbf{0}$

2 Independent cases (data matrix rows)

3 $V[\boldsymbol{\varepsilon}] = \boldsymbol{\Sigma}$ (constant variance)

4 $\boldsymbol{\varepsilon} \cong N_p(\mathbf{0}, \boldsymbol{\Sigma})$ Needed for "exact" small sample inference .

Most tests and confidence procedures related to elements of $\mathbf{B}$ are <u>resistant to non-normality</u> - they "work as advertised" adequately even with non-normal $\boldsymbol{\varepsilon}$'s.

The assumption that $E[\boldsymbol{\varepsilon}] = \mathbf{0}$ is really just a statement that the fixed part of the model is correct. That's why I list it as the most important assumption.

You can put any multivariate linear model (regression, MANOVA, MANACOVA) in the form of a multivariate linear regression (involving "dummy" variables for MANOVA and MANACOVA).

This means you can express *all* the models in the form

$$\mathbf{Y} = (\mathbf{ZB}) + \{\boldsymbol{\varepsilon}\}, \text{ N by p}$$

- $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_p]$, N by p matrix of response (dependent) variables
- $\mathbf{Z} = [\mathbf{Z}_0, \mathbf{Z}_1, \ldots, \mathbf{Z}_k]$ is a n by k+1 matrix of <u>predictor</u> (independent) variables, possibly including dummy variables
- $\mathbf{B} = [\beta_{j\ell}] = \begin{bmatrix} \boldsymbol{\beta}_0' \\ \boldsymbol{\beta}_1' \\ \cdots\cdot \\ \boldsymbol{\beta}_k' \end{bmatrix} = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_p]$

  is a k+1 by p <u>matrix</u> of coefficients.

Each row $\boldsymbol{\beta}_\ell'$ of $\mathbf{B}$ goes with a predictor Z_ℓ. Each column $\mathbf{b}_m$ of $\mathbf{B}$ goes with a response variable Y_m.

One-way MANOVA $\quad \mathbf{B} = \begin{bmatrix} \boldsymbol{\mu}' \\ \boldsymbol{\alpha}_1' \\ \boldsymbol{\alpha}_2' \\ \cdots \\ \boldsymbol{\alpha}_{g-1}' \\ \boldsymbol{\alpha}_g' \end{bmatrix}$

Linking with the general notation, $k = g$

$$\boldsymbol{\beta}_0 = \boldsymbol{\mu},\ \boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1,\ \ldots,\ ,\ \boldsymbol{\beta}_g = \boldsymbol{\alpha}_g$$

$$\mathbf{b}_\ell = \begin{bmatrix} \mu_\ell \\ \alpha_{1\ell} \\ \cdots\cdot \\ \alpha_{g\ell} \end{bmatrix},\ \ell = 1,\ \ldots,\ p$$

Caution: The $\mathbf{Z}$ matrix for this parameter matrix is not full rank. It is, if $\mathbf{B}$ omits the last row ($\boldsymbol{\alpha}_g'$).

Estimation

For normal errors, it turns out that the best way (maximum likelihood) to estimate $\mathbf{B}$ is by univariate ordinary least squares (OLS) for each column of $\mathbf{B}$

$$\mathbf{b}_\ell = [\beta_{0\ell}, \beta_{1\ell}, \beta_{2\ell}, \ldots, \beta_{k\ell}]',\ \ell = 1, \ldots, p.$$

separately.

The matrix formula for the univariate OLS estimates is

$$\hat{\mathbf{b}}_\ell \equiv [\hat{\beta}_{0\ell},\ \hat{\beta}_{1\ell}, \ldots,\ \hat{\beta}_{k\ell}]' = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}_\ell,$$
$$\ell = 1, \ldots, p$$

This assumes $\mathbf{Z}$ is of full rank so $\mathbf{Z}'\mathbf{Z}$ is invertible and the coefficients are all estimable.

You can combine these into one matrix equation:

$$\hat{\mathbf{B}} = [\hat{\mathbf{b}}_1,\ \hat{\mathbf{b}}_2,\ \ldots,\ \hat{\mathbf{b}}_p]$$
$$= (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y},\ k+1 \text{ by } p$$

- $\hat{\mathbf{B}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}$ is a "clone" of the uni-variate formula, that is, it has the *same algebraic form*.
- $\hat{\mathbf{B}}$ maximizes the normal likelihood.

  If you do the math you find that the MLE $\hat{\mathbf{B}}$ minimizes the *determinant* of the residual cross product (RCP) matrix

  $$\det((\mathbf{Y}-\mathbf{ZB})'(\mathbf{Y}-\mathbf{ZB})) = \det(\text{RCP}).$$

  The matrix $\mathbf{Y} - \mathbf{ZB}$ consists of residuals from the regression

  Math shows that $\hat{\mathbf{B}}$ also minimizes all the diagonal elements of RCP, the residual sums of squares..

In the **SUR** situation (different models for different variables), although the maximum likelihood estimates minimize det(RCP), the solution isn't the same as the univariate least squares estimates.

Sampling distribution of $\hat{\mathbf{B}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}$

If you know the univariate facts you know a lot.

- Univariate LS estimates are unbiased ($E[\hat{\mathbf{b}}] = \mathbf{b}$) $\Rightarrow$ $\hat{\mathbf{B}}$ is *unbiased* ($E[\hat{\mathbf{B}}] = \mathbf{B}$).
- The variance matrix of a *column* $\hat{\mathbf{b}}_\ell$ of $\hat{\mathbf{B}}$ is (from the univariate result):

  $V[\hat{\mathbf{b}}_\ell] = \sigma_{\ell\ell}(\mathbf{Z}'\mathbf{Z})^{-1} = \sigma_{\ell\ell}\mathbf{C} = \sigma_{\ell\ell}[c_{ij}]$,
  where $\mathbf{C} = [c_{ij}] = (\mathbf{Z}'\mathbf{Z})^{-1}$, and
  $\sigma_{\ell\ell} = V[\varepsilon_\ell]$, $\ell = 1,\ldots,p$.
- The $(k+1)\times(k+1)$ matrix of covariances between elements in different *columns* of $\hat{\mathbf{B}}$ (coefficients for different variables) is

  $$\underset{p\times p}{\text{Cov}[\hat{\mathbf{b}}_\ell,\hat{\mathbf{b}}_m]} = E[(\hat{\mathbf{b}}_\ell - \mathbf{b}_m)(\hat{\mathbf{b}}_m - \mathbf{b}_m)'] = \sigma_{\ell m}(\mathbf{Z}'\mathbf{Z})^{-1} = \sigma_{\ell m}\mathbf{C},$$

  where $\sigma_{\ell m} = \text{Cov}[\varepsilon_\ell,\varepsilon_m]$, $\ell \neq m$

- Each element $\hat{\beta}_{j\ell}$ in column ℓ of $\hat{\mathbf{B}}$ is a linear combination of the elements of $\mathbf{Y}_\ell$.
- Each *column* $\hat{\mathbf{b}}_\ell$ (estimated coefficients for y_ℓ) is $N_{k+1}(\mathbf{b}_\ell, \sigma_{\ell\ell}(\mathbf{Z}'\mathbf{Z})^{-1})$
- Each *row* $\hat{\boldsymbol{\beta}}_j$ (estimated coefficients of Z_j for all y_ℓ's) is $N_p(\boldsymbol{\beta}_j, c_{jj}\boldsymbol{\Sigma})$.
- All the $p(k+1)$ elements $\hat{\beta}_{j\ell}$ together are multivariate normal $N_{p(k+1)}$.

What is the variance matrix of all $p(k+1)$ estimated coefficients $\hat{\beta}_{j\ell}$?

There a neat mathematical notation you can use to describe the variance matrix of all $p\times(k+1)$ elements $\hat{\beta}_{j\ell}$:

Let

$$\mathbf{b} \equiv \text{vec}(\mathbf{B}) = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \dots \\ \mathbf{b}_p \end{bmatrix} = [\mathbf{b}_1', \mathbf{b}_2', \dots, \mathbf{b}_p']'$$

be the length $p(k+1)$ vector obtained by stringing the columns $\mathbf{b}_\ell$ of $\mathbf{B}$ one after the other. Similarly, let

$$\hat{\mathbf{b}} \equiv \text{vec}(\hat{\mathbf{B}}) = [\hat{\mathbf{b}}_1' \;\; \hat{\mathbf{b}}_2' \; \dots \; \hat{\mathbf{b}}_p']'.$$

Then

- $\hat{\mathbf{b}}$ is $N_{p(k+1)}(\mathbf{b}, \boldsymbol{\Sigma}\otimes(\mathbf{Z}'\mathbf{Z})^{-1})$,
  where the $p(k+1)$ by $p(k+1)$ matrix $V[\hat{\mathbf{b}}] = \boldsymbol{\Sigma}\otimes(\mathbf{Z}'\mathbf{Z})^{-1}$ is the *Kronecker product* of $\boldsymbol{\Sigma}$ and $(\mathbf{Z}'\mathbf{Z})^{-1}$.

Vocabulary: When **A** is a M by N matrix and **B** is a m by n matrix, their Kronecker product is the M×m by N×n matrix

$$\mathbf{A}\otimes\mathbf{B} \equiv \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} \ldots a_{1N}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} \ldots a_{2N}\mathbf{B} \\ \cdots & \cdots \\ a_{M1}\mathbf{B} & a_{M2}\mathbf{B} \ldots a_{MN}\mathbf{B} \end{bmatrix},$$

MacAnova example using **kronecker()**:

```
Cmd> a <- matrix(run(4),2); a # arbitrary 2 by 2 matrix
(1,1)            1             3      M = 2, N = 2
(2,1)            2             4

Cmd> b <- matrix(vector(1,1,1, 1,-1,0), 3); b # 3 by 2 matrix
(1,1)            1             1      m = 3, n = 2
(2,1)            1            -1
(3,1)            1             0

Cmd> kronecker(a,b) # macro distributed with MacAnova
WARNING: searching for unrecognized macro kronecker near
kronecker(
(1,1)            1             1 :            3             3
(2,1)            1            -1 :            3            -3
(3,1)            1 a[1,1]*b    0 :            3 a[1,2]*b    0
..................................................................
(4,1)            2             2 :            4             4
(5,1)            2            -2 :            4            -4
(6,1)            2 a[2,1]*b    0 :            4 a[2,2]*b    0

Cmd> dim(kronecker(a,b)) # 2*3 by 2*2 matrix
(1)            6             4
```


Facts

- $(\mathbf{A}\otimes\mathbf{B})^{-1} = \mathbf{A}^{-1}\otimes\mathbf{B}^{-1}$
- $\hat{\mathbf{b}}'V[\hat{\mathbf{b}}]^{-1}\hat{\mathbf{b}} = \text{tr } \mathbf{\Sigma}^{-1}\hat{\mathbf{B}}'(\mathbf{Z}'\mathbf{Z})^{-1}\hat{\mathbf{B}}$
  $= \text{sum of diagonals of } \mathbf{\Sigma}^{-1}\hat{\mathbf{B}}'(\mathbf{Z}'\mathbf{Z})^{-1}\hat{\mathbf{B}}$

Application

Suppose $\mathbf{S} = \hat{\mathbf{\Sigma}}$ estimates $\mathbf{\Sigma}$. Then

$$T^2 = \hat{\mathbf{b}}'\hat{V}[\hat{\mathbf{b}}]^{-1}\hat{\mathbf{b}} = \text{tr } \mathbf{S}^{-1}(\hat{\mathbf{B}}'(\mathbf{Z}'\mathbf{Z})^{-1}\hat{\mathbf{B}})$$

is a form of Hotelling's T^2 statistic that tests H_0: $\mathbf{B} = \mathbf{0}$, that is

$$H_0: \beta_{j\ell} = 0,\ j = 0, \ldots, k,\ \ell = 1, \ldots, p$$

Under wide conditions, in large samples, the null distribution of T^2 is $\chi^2_{p(k+1)}$.

There is no easy exact small sample distribution as there is for the two-sample and paired Hotelling T^2 statistics.

When $k = 0$, $T^2 \cong (pf_e/(f_e-p+1))F_{p,f_e-p+1}$

When $p = 1$, $T^2 \cong F_{k+1,f_e}$

Unbiased estimate of Σ

Define the p by p ***error matrix***

$$\mathbf{E} = \sum_{1\le i\le N}(\mathbf{y}_i - \hat{\mathbf{y}}_i)(\mathbf{y}_i - \hat{\mathbf{y}}_i)' = (\mathbf{Y}-\mathbf{Z}\hat{\mathbf{B}})'(\mathbf{Y}-\mathbf{Z}\hat{\mathbf{B}})$$

where $\hat{\mathbf{y}}_i = \hat{\mathbf{B}}'\mathbf{z}_i = (\mathbf{z}_i'\hat{\mathbf{B}})'$ is the predicted value based on $\mathbf{z}_i'$, (row i of $\mathbf{Z}$).

- $\mathbf{Y} - \mathbf{Z}\hat{\mathbf{B}}$ is the matrix of <u>least squares residuals</u>.
- $\mathbf{E}$ is the multivariate analogue of SS_e in univariate ANOVA and regression. To get a formula for $\mathbf{E}$, replace $(...)^2$ in a formula for SS_e by $(...)(...)'$.
- $e_{\ell\ell} = \sum_{1\le i\le N}(y_{i\ell} - \hat{y}_{i\ell})^2 = SS_e^{(\ell)}$ (ANOVA residual <u>sum of squares</u> for y_ℓ)
- $e_{\ell m} = e_{m\ell} = \sum_{1\le i\le N}(y_{i\ell} - \hat{y}_{i\ell})(y_{im} - \hat{y}_{im})$ (residual <u>sum of products</u> for y_ℓ and y_m)

Johnson and Wichern use $\mathbf{W}$ (for $\mathbf{W}$ithin) instead of $\mathbf{E}$ in some contexts.

The minimum number of linearly independent parameter *vectors*, each of length p, required in the model is r = rank($\mathbf{Z}$). If $\mathbf{Z}$ is of full rank, r = k+1. Thus at least r×p parameters are required in all.

Define

- $\mathbf{S} \equiv (1/f_e)\mathbf{E} = (1/f_e)\sum_{1\le i\le N}(\mathbf{y}_i - \hat{\mathbf{y}}_i)(\mathbf{y}_i - \hat{\mathbf{y}}_i)'$ where
- $f_e = N - r$ ($f_e = N-k-1$ for full rank $\mathbf{Z}$)

Facts:

- $E[\mathbf{S}] = \mathbf{\Sigma} \Rightarrow \mathbf{S}$ is an <u>unbiased</u> estimate of $\mathbf{\Sigma}$.
- When $\mathbf{y}$ is MVN with $V[\mathbf{y}] = \mathbf{\Sigma}$, $\mathbf{E}$ is $W_p(f_e, \mathbf{\Sigma})$ ($\sigma^2\chi^2_{f_e}$ when p = 1)

$\mathbf{S}$ is multivariate analog of the univarate

$$s^2 = (1/f_e)\sum_{1\le i\le N}(y_i - \hat{y}_i)^2$$

MacAnova MANOVA Example

```
Cmd> irisdata <- read("","t11_05",quiet:T)
Read from file "TP1:Stat5401:Data:JWData5.txt"

Cmd> varieties <- factor(irisdata[,1])
```

Using **factor()** is essential to mark **varieties** as a categorical variable rather than a quantitative variable.

```
Cmd> y <- irisdata[,-1] # strip off variety numbers

Cmd> list(varieties,y)
varieties       REAL   150   1     FACTOR with 3 levels
y               REAL   150   4     N = 150, p = 4

Cmd> manova("y=varieties") # like anova()
Model used is y=varieties
WARNING: summaries are sequential
                         SS and SP Matrices
                DF
CONSTANT         1
            SepLen       SepWid       PetLen       PetWid
SepLen      5121.7       2679.8       3293.9       1051.2
SepWid      2679.8       1402.1       1723.4       550.01
PetLen      3293.9       1723.4       2118.4       676.06
PetWid      1051.2       550.01       676.06       215.76
varieties        2
            SepLen       SepWid       PetLen       PetWid
SepLen      63.212      -19.953       165.25       71.279
SepWid     -19.953       11.345       -57.24      -22.933    = H = B
PetLen      165.25       -57.24        437.1       186.77
PetWid      71.279      -22.933       186.77       80.413
ERROR1         147
            SepLen       SepWid       PetLen       PetWid
SepLen      38.956        13.63       24.625        5.645
SepWid       13.63       16.962       8.1208       4.8084    = E = W
PetLen      24.625       8.1208       27.223       6.2718
PetWid       5.645       4.8084       6.2718       6.1566
```

This is default **manova()** output when $p \leq 5$.

Hypothesis matrix

$$\mathbf{H} = \mathbf{B} = \sum_{1\leq j\leq g} n_j(\overline{\mathbf{y}}_{.j} - \overline{\mathbf{y}}_{..})(\overline{\mathbf{y}}_{.j} - \overline{\mathbf{y}}_{..})'$$

This generalizes the univariate formula

$$SS_h = SSB = \sum_{1\leq j\leq g} n_j(\overline{y}_{.j} - \overline{y}_{..})^2$$

Error matrix is multiple of pooled variance matrix estimate

$$\mathbf{E} = \mathbf{W} = \sum_{1\leq j\leq g}(n_j - 1)\mathbf{S}_j$$

$$\mathbf{S} = \mathbf{S}_{pooled} = (N - g)^{-1}\sum_{1\leq j\leq g}(n_j - 1)\mathbf{S}_j$$

This generalizes the univariate formula

$$s_{pooled}^2 = (N - g)^{-1}\sum_{1\leq j\leq g}(n_j - 1)s_j^2$$

MacAnova computes variables **DF**, **RESIDUALS** and **SS** just as **anova()** and **regress()** do.

```
Cmd> list(DF, RESIDUALS, SS)
DF              REAL    3     (labels)
RESIDUALS       REAL    150   4     (labels)
SS              REAL    3     4     4     (labels)

Cmd> DF # computed by manova(); same as anova() DF
    CONSTANT   varieties      ERROR1
           1           2         147
```

```
Cmd> SS # 3 by 4 by 4 array; also computed by manova()
                      SepLen       SepWid       PetLen       PetWid
CONSTANT  SepLen      5121.7       2679.8       3293.9       1051.2
          SepWid      2679.8       1402.1       1723.4       550.01
          PetLen      3293.9       1723.4       2118.4       676.06
          PetWid      1051.2       550.01       676.06       215.76
varieties SepLen      63.212      -19.953       165.25       71.279
          SepWid     -19.953       11.345       -57.24      -22.933
          PetLen      165.25       -57.24        437.1       186.77
          PetWid      71.279      -22.933       186.77       80.413
ERROR1    SepLen      38.956        13.63       24.625        5.645
          SepWid       13.63       16.962       8.1208       4.8084
          PetLen      24.625       8.1208       27.223       6.2718
          PetWid       5.645       4.8084       6.2718       6.1566
```

`SS` is a 3 dimensional array, with the first subscript indexing matrices.

```
Cmd> list(SS) # SS is a three dimension array
SS               REAL    3     4     4     (labels)

Cmd> e <- SS[3,,]; e # third matrix E; SS[2,,] is H
                SepLen       SepWid       PetLen       PetWid
ERROR1 SepLen   38.956        13.63       24.625        5.645
       SepWid    13.63       16.962       8.1208       4.8084
       PetLen   24.625       8.1208       27.223       6.2718
       PetWid    5.645       4.8084       6.2718       6.1566
```


The diagonal elements of `SS[j,,]` are the univariate SS:

```
Cmd> ss <- SS # save it

Cmd> anova("{y[,3]} = varieties") # univariate ANOVA
Model used is {y[,3]} = varieties
                DF           SS           MS
CONSTANT         1       2118.4       2118.4
varieties        2        437.1       218.55
ERROR1         147       27.223      0.18519

Cmd> ss[,3,3] # 3rd diagonal element of matrices
                      PetLen
CONSTANT  PetLen      2118.4
varieties PetLen       437.1
ERROR1    PetLen      27.223
```

MacAnova computes MANOVA as multivariate regression with dummy variables with values 0, 1 and -1. You can see what they are using through `modelinfo()`. Here is an example with "toy" data, g = 3, p = 3, N = 10.

```
Cmd> a <- factor(1,1,1,2,2,2,3,3,3,3) # n_1=3, n_2=3, n_3=4

Cmd> Y <- matrix(rnorm(30),10) # N = 10, p = 3

Cmd> manova("Y = a", silent:T)

Cmd> xvariables() # gets the actual Z matrix used
 (1,1)           1            1            0
 (2,1)           1            1            0
 (3,1)           1            1            0
 (4,1)           1            0            1
 (5,1)           1            0            1
 (6,1)           1            0            1
 (7,1)           1           -1           -1
 (8,1)           1           -1           -1
 (9,1)           1           -1           -1
(10,1)           1           -1           -1
```

Basic confidence interval for one coefficient

A multivariate linear model can always be put in the form

$$\mathbf{Y} = \mathbf{ZB} + \boldsymbol{\varepsilon},\ E[\boldsymbol{\varepsilon}] = \mathbf{0},\ V[\boldsymbol{\varepsilon}] = \boldsymbol{\Sigma}$$

$$\mathbf{Y} \text{ and } \boldsymbol{\varepsilon}\ n \text{ by } p,\ \mathbf{Z}\ N \text{ by } k+1,$$

$$\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_p] = [\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k]' \ k+1 \text{ by } p$$

Let $\mathbf{C} = [c_{ij}] = (\mathbf{Z}'\mathbf{Z})^{-1}$. Then

$$V[\hat{\mathbf{b}}_\ell] = \sigma_{\ell\ell}\mathbf{C},\ \ell = 1, \ldots, p$$

In particular

$$V[\hat{\beta}_{j\ell}] = c_{jj}\sigma_{\ell\ell},\ j = 0, \ldots, k,\ \ell = 1, \ldots, p$$

The <u>estimated</u> standard error of $\hat{\beta}_{j\ell}$ is

$$\widehat{SE}[\hat{\beta}_{j\ell}] = \sqrt{\{c_{jj}\hat{\sigma}_{\ell\ell}\}}$$

where $\hat{\sigma}_{\ell\ell}$, is the MSE for y_ℓ, and is a diagonal element of $\hat{\boldsymbol{\Sigma}} = \mathbf{S} = (1/f_e)\mathbf{E}$.

MacAnova You can use `secoefs()` to retrieve all $\hat{\beta}_{j\ell}$'s and all $\widehat{SE}[\hat{\beta}_{j\ell}]$.

There are several different $100(1 - \alpha)\%$ confidence intervals for a coefficient $\beta_{j\ell}$, both "ordinary" (non-simultaneous) and simultaneous.

All have the form

$\beta_{\ell j} = \hat{\beta}_{\ell j} \pm K_\alpha\sqrt{\{c_{jj}\hat{s}_{\ell\ell}\}}$, with constant K_α which depends on the type of interval

- *Single* non-simulatneous <u>large sample</u> confidence interval has $K_\alpha = z(\alpha/2)$
- *Single* non-simulatneous confidence interval with normal or near normal errors has $K_\alpha = t_{f_e}(\alpha/2)$.
- *Simultaneous* intervals for *all* $M = (k+1)p$ coefficients by Bonferronizing Student's t by M:

  $K_\alpha = z((\alpha/M)/2)$ or $K_\alpha = t_{f_e}((\alpha/M)/2)$.

Example

```
Cmd> manova("y=varieties",silent:T)

Cmd> coefs()#describes most recent regress(), anova(), manova()
component: CONSTANT      Least squares estimates of μ

         SepLen      SepWid      PetLen      PetWid
(1)      5.8433      3.0573       3.758      1.1993   μ̂'
component: varieties     Least squares of variety effects
         SepLen      SepWid      PetLen      PetWid
(1)    -0.83733     0.37067      -2.296    -0.95333   α̂'1
(2)    0.092667    -0.28733       0.502     0.12667   α̂'2
(3)     0.74467   -0.083333       1.794     0.82667   α̂'3

Cmd> stats <- secoefs(); stats
component: CONSTANT      Estimates and their standard errors
  component: coefs
         SepLen      SepWid      PetLen      PetWid
(1)      5.8433      3.0573       3.758      1.1993
  component: se
         SepLen      SepWid      PetLen      PetWid
(1)    0.042032    0.027735    0.035137     0.01671
component: varieties
  component: coefs   Alphahats
         SepLen      SepWid      PetLen      PetWid
(1)    -0.83733     0.37067      -2.296    -0.95333
(2)    0.092667    -0.28733       0.502     0.12667
(3)     0.74467   -0.083333       1.794     0.82667
  component: se      Standard errors of alphahats
         SepLen      SepWid      PetLen      PetWid
(1)    0.059443    0.039224    0.049691    0.023631
(2)    0.059443    0.039224    0.049691    0.023631
(3)    0.059443    0.039224    0.049691    0.023631
```

(to be continued)