Multivariate Linear Regression Model Example (32950-24620)

Example Thunder Basin Antelope Study

The data consists of 8 observations for each year on four variables:

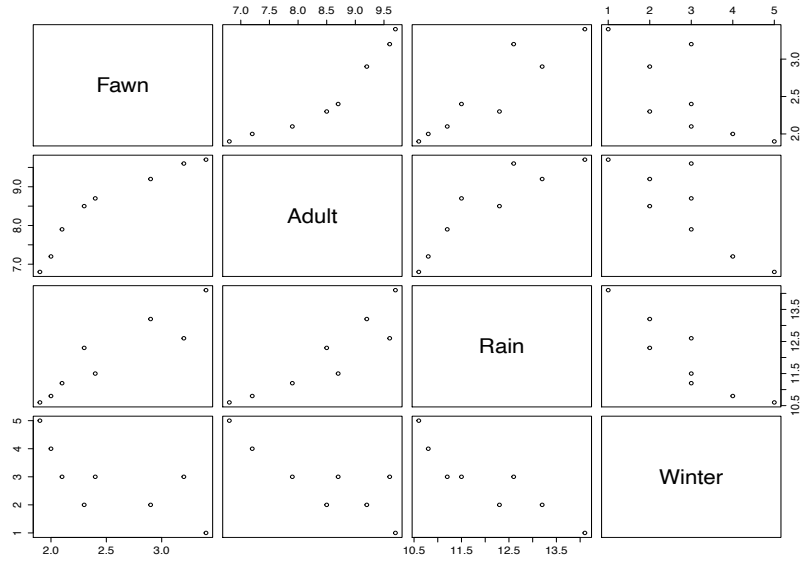
Fawn = spring fawn count per 100

Adult = size of adult antelope population per 100

Rain = annual precipitation (inches)

Winter = winter severity index (1 = mild, 5 = severe)

```
Deerdata=read.csv("antelope.csv")
Deerdata
  Fawn Adult Rain Winter
1  2.9   9.2 13.2      2
2  2.4   8.7 11.5      3
3  2.0   7.2 10.8      4
4  2.3   8.5 12.3      2
5  3.2   9.6 12.6      3
6  1.9   6.8 10.6      5
7  3.4   9.7 14.1      1
8  2.1   7.9 11.2      3
```


```
mfit = lm(cbind(Fawn,Adult)~Rain+Winter,data=Deerdata)
> summary(mfit,cor=T)
Response Fawn :
Call:
lm(formula = Fawn ~ Rain + Winter, data = Deerdata)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -5.7791     2.2139  -2.610  0.04765 *
Rain          0.6357     0.1511   4.207  0.00843 **
Winter        0.2269     0.1490   1.522  0.18842

Residual standard error: 0.2133 on 5 degrees of freedom
Multiple R-squared:    0.9,Adjusted R-squared:   0.86
F-statistic: 22.49 on 2 and 5 DF,  p-value: 0.003164

Correlation of Coefficients:
       (Intercept) Rain
Rain   -1.00
Winter -0.93        0.90

Response Adult :
Call:
lm(formula = Adult ~ Rain + Winter, data = Deerdata)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.4227     5.6421   0.075    0.943
Rain          0.6923     0.3851   1.798    0.132
Winter       -0.1067     0.3798  -0.281    0.790

Residual standard error: 0.5437 on 5 degrees of freedom
Multiple R-squared:  0.8175,Adjusted R-squared:  0.7445
F-statistic:  11.2 on 2 and 5 DF,  p-value: 0.01422

Correlation of Coefficients:
       (Intercept) Rain
Rain   -1.00
Winter -0.93        0.90
```


Compared with univariate regressions

```
> summary(lm(Fawn~Rain+Winter,data=Deerdata),cor=T)
Residuals:
       1        2        3        4        5        6        7        8
-0.165458  0.188313  0.006417 -0.193358  0.289080 -0.193312 -0.010695  0.079013
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -5.7791     2.2139  -2.610  0.04765 *
Rain          0.6357     0.1511   4.207  0.00843 **
Winter        0.2269     0.1490   1.522  0.18842
Residual standard error: 0.2133 on 5 degrees of freedom
Multiple R-squared:    0.9,Adjusted R-squared:   0.86
F-statistic: 22.49 on 2 and 5 DF,  p-value: 0.003164
Correlation of Coefficients:
       (Intercept) Rain
Rain   -1.00
Winter -0.93        0.90

> summary(lm(Adult~Rain+Winter,data=Deerdata),cor=T)
Residuals:
       1        2        3        4        5        6        7        8
-0.14820  0.63547 -0.27320 -0.22509  0.77389 -0.42804 -0.37800  0.04317
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.4227     5.6421   0.075    0.943
Rain          0.6923     0.3851   1.798    0.132
Winter       -0.1067     0.3798  -0.281    0.790
Residual standard error: 0.5437 on 5 degrees of freedom
Multiple R-squared:  0.8175,Adjusted R-squared:  0.7445
F-statistic:  11.2 on 2 and 5 DF,  p-value: 0.01422
Correlation of Coefficients:
       (Intercept) Rain
Rain   -1.00
Winter -0.93        0.90
```

Covariance estimation from multivariate linear model

$$\begin{bmatrix} Fawn \\ Adult \end{bmatrix} = \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} + Rain \begin{bmatrix} \beta_{11} \\ \beta_{21} \end{bmatrix} + Winter \begin{bmatrix} \beta_{12} \\ \beta_{22} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

Residuals:

```
 mfit$residuals
          Fawn       Adult
1 -0.165457518 -0.14820089
2  0.188313098  0.63547035
3  0.006417185 -0.27320137
4 -0.193358111 -0.22509230
5  0.289079731  0.77389303
6 -0.193311876 -0.42804419
7 -0.010695469 -0.37799828
8  0.079012960  0.04317365
```

Estimated $\hat{\Sigma}$ in $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \sim N_p(0_p, \Sigma)$

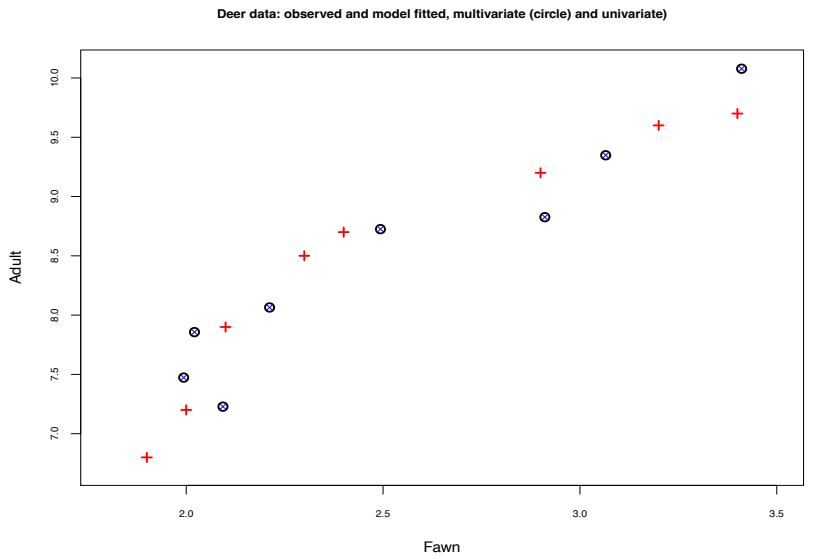
```
> cov(mfit$residuals)
           Fawn      Adult
Fawn   0.03250865 0.07141079
Adult 0.07141079 0.21113864
```

The response variables are strongly correlated:

```
> cor(mfit$residuals)
           Fawn     Adult
Fawn   1.0000000 0.8619469
Adult 0.8619469 1.0000000
```


MANOVA (Sequential ANOVA and MANOVA)

```
> manova(cbind(Fawn,Adult) ~ Rain + Winter, data=Deerdata)

                Rain   Winter Residuals
Fawn            1.941965 0.105475  0.227561
Adult           6.598701 0.023327  1.477971
Deg. of Freedom        1        1         5

Residual standard errors: 0.2133357 0.5436857 (Estimated effects may be unbalanced)

> manova(cbind(Fawn,Adult) ~ Winter + Rain, data=Deerdata)

                Winter     Rain Residuals
Fawn            1.241897 0.805543  0.227561
Adult           5.666436 0.955593  1.477971
Deg. of Freedom        1        1         5


> summary(manova(cbind(Fawn,Adult) ~ Rain + Winter, data=Deerdata),test="Wilks")
          Df   Wilks approx F num Df den Df   Pr(>F)
Rain       1 0.09831  18.3439      2      4 0.009665 **
Winter     1 0.29085   4.8764      2      4 0.084594 .
Residuals  5
> summary(manova(cbind(Fawn,Adult) ~ Winter + Rain, data=Deerdata),test="Wilks")
          Df   Wilks approx F num Df den Df  Pr(>F)
Winter     1 0.15457   10.939      2      4 0.02389 *
Rain       1 0.14004   12.281      2      4 0.01961 *


> summary(manova(cbind(Fawn,Adult) ~ Rain + Winter, data=Deerdata))
          Df  Pillai approx F num Df den Df   Pr(>F)
Rain       1 0.90169  18.3439      2      4 0.009665 **
Winter     1 0.70915   4.8764      2      4 0.084594 .
```

Fit one explanatory variable only

```
mfit1 = lm(cbind(Fawn,Adult)~Rain,data=Deerdata)
> summary(mfit1)

Response Fawn :

Call: lm(formula = Fawn ~ Rain, data = Deerdata)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.63251    0.87591  -3.005  0.02384 *
Rain         0.42845    0.07244   5.915  0.00104 **

Residual standard error: 0.2356 on 6 degrees of freedom
Multiple R-squared:  0.8536,Adjusted R-squared:  0.8292
F-statistic: 34.99 on 1 and 6 DF,  p-value: 0.001039

Response Adult :

Call: lm(formula = Adult ~ Rain, data = Deerdata)

Residuals:
    Min      1Q  Median      3Q     Max
-0.5147 -0.2992 -0.1627  0.2522  0.7057

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.0571     1.8597  -0.568  0.59037
Rain          0.7898     0.1538   5.135  0.00215 **

Residual standard error: 0.5002 on 6 degrees of freedom
Multiple R-squared:  0.8147,Adjusted R-squared:  0.7838
F-statistic: 26.37 on 1 and 6 DF,  p-value: 0.002146
```