hw6 p1

For fractions:

```
library(MASS)
```

R Markdown

```
# Com
X <- cbind(c(2,2,2,0,-1,-2,-3),c(2,4,6,0,-4,-4,-4))
colnames(X) <- NULL
n <- dim(X)[1]
r <- dim(X)[2]
```

(1.b) Determine the proportion of total sample variance due to the first sample principale component.

```
# Compute sample mean and S
Ones <- rep(1,n)
x_sample_mean <- 1/n * t(X)%*%Ones
S <- 1/(n-1) * t(X - Ones%*%t(x_sample_mean))%*%(X - Ones%*%t(x_sample_mean))

# eigens
ev <- eigen(S)
eigen_values <- ev$values
V <- ev$vectors

prop <- eigen_values[1]/(S[1,1] + S[2,2])
cat( fractions(prop), '=', eigen_values[1], '/', (S[1,1] + S[2,2]))
```

```
## 0.975743 = 21.1411 / 21.66667
```

(1.c) Compare the contributions of the two variates to the determination of the first sample principal com-ponent based on loadings.

```
# We can see the Loadings in V
V
```

```
##           [,1]       [,2]
## [1,] 0.4297717 -0.9029376
## [2,] 0.9029376  0.4297717
```

(1.d) Compare the contributions of the two variates to the determination of the first sample principal com-ponent based on sample correlations.

```
# Compute the correlations
cat('Corr(Y_1, X_1) =', V[1,1]*sqrt(eigen_values[1]/S[1,1]))
```

```
## Corr(Y_1, X_1) = 0.9492716
```

```
cat(' --- Corr(Y_1, X_2) =', V[2,1]*sqrt(eigen_values[1]/S[2,2]))
```

```
##  --- Corr(Y_1, X_2) = 0.9971958
```

(1.e) Repeat (a-d) with the data standarized.

```
# Com
D <- cbind(c(1/sqrt(S[1,1]),0),c(0,1/sqrt(S[2,2])))
Z <- (X - Ones%*%t(x_sample_mean))%*%t(D)
colnames(Z) <- NULL

# Compute sample mean and S
z_sample_mean <- 1/n * t(Z)%*%Ones
S_z <- 1/(n-1) * t(Z - Ones%*%t(z_sample_mean))%*%(Z - Ones%*%t(z_sample_mean))

# eigens
ev_z <- eigen(S_z)
eigen_values_z <- ev_z$values
V_z <- ev_z$vectors

prop_z <- eigen_values_z[1]/(S_z[1,1] + S_z[2,2])
cat( fractions(prop_z), '=', eigen_values_z[1], '/', (S_z[1,1] + S_z[2,2]))
```

```
## 0.9615385 = 1.923077 / 2
```

```
# We can see the loadings in V
V_z
```

```
##           [,1]       [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

```
# Compute the correlations
cat('Corr(Y_1, Z_1) =', V_z[1,1]*sqrt(eigen_values_z[1]))
```

```
## Corr(Y_1, Z_1) = 0.9805807
```

```
cat(' --- Corr(Y_1, Z_2) =', V_z[2,1]*sqrt(eigen_values_z[1]))
```

```
##  --- Corr(Y_1, Z_2) = 0.9805807
```