

Covariance Matrix - Prob 10.16

Blake Baird

2/17/2022

A is an $n \times k$ matrix.

```
ones = function(n) matrix(rep(1,n)) # function to create a column vector of 1s
dat = cbind(c(2,5,6),c(8,22,8),c(9,10,11),c(22,44,66))
A = matrix(data=dat,ncol=4,nrow=3) # demonstration matrix
A
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    2    8    9   22
## [2,]    5   22   10   44
## [3,]    6    8   11   66
```

a) Give expression for column mean (k -vector), μ , in terms of **A**.

```
n = nrow(A)
k = ncol(A)
one_over_n = matrix((1/n) * ones(n)) # scalar-vector multiplication
mu = crossprod(A, one_over_n)        # matrix-vector multiplication
one_over_n
```

```
##      [,1]
## [1,] 0.3333333
## [2,] 0.3333333
## [3,] 0.3333333
```

```
mu # vector of column means
```

```
##      [,1]
## [1,] 4.333333
## [2,] 12.666667
## [3,] 10.000000
## [4,] 44.000000
```

b) Give expression for de-meaned **A**, **Atilde**.

$$\tilde{A} = A - 1_n \mu^T$$

```
Atilde = A - (ones(n) %*% t(mu))
Atilde # de-meaned A
```

```
##      [,1] [,2] [,3] [,4]
## [1,] -2.3333333 -4.666667 -1 -22
## [2,]  0.6666667  9.333333  0  0
## [3,]  1.6666667 -4.666667  1  22
```

c) The *covariance matrix* $\Sigma = (1/N)\tilde{A}^T \tilde{A}$

The Σ_{ii} terms are $\text{stdev}(a_i)^2$ where a_i are the column vectors of the original matrix, A .

The Σ_{ij} terms are $\text{stdev}(a_i)\text{stdev}(a_j)\rho_{ij}$ where ρ_{ij} is the correlation between columns i and j .

```
library(magrittr) # import the pipe operator %>%
```

```
# stdev = rms of de-meaned vector
stdev = function(x) {
  if (!is.vector(x)) {
    warning("stdev should only be used on vectors")
    return(NULL)
  }
  sqrt( sum( (x - mean(x))^2 ) / length(x) )
}
```

```
covar = (1/n) * crossprod(Atilde)
covar
```

```
##           [,1]      [,2]      [,3]      [,4]
## [1,]  2.888889  3.111111  1.333333  29.33333
## [2,]  3.111111 43.555556  0.000000  0.00000
## [3,]  1.333333  0.000000  0.666667  14.66667
## [4,] 29.333333  0.000000 14.666667 322.66667
```

```
for(i in 1:4) stdev(A[,i])^2 %>% print
```

```
## [1] 2.888889
## [1] 43.55556
## [1] 0.6666667
## [1] 322.6667
```

```
rho12 = covar[1,2] / ( diag(covar)[2]^0.5 * diag(covar)[1]^0.5 )
rho12
```

```
## [1] 0.2773501
```

```
stdev(A[,1])*stdev(A[,2])*rho12
```

```
## [1] 3.111111
```