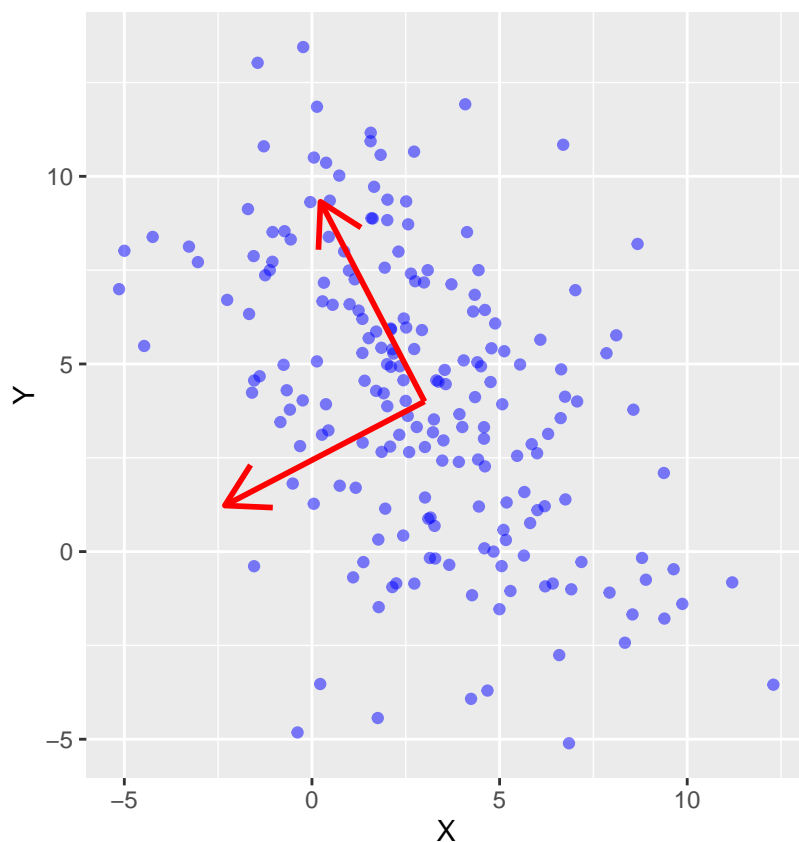# 102HW3

JING LI

2/2/2017

1.

(a)

```r
library(MASS)
mu <- matrix(c(3,4), ncol = 1)
sigma <- matrix(c(9, -5, -5, 16), nrow = 2)
eigen(sigma)
```

```
## $values
## [1] 18.603278  6.396722
##
## $vectors
##              [,1]        [,2]
## [1,] -0.4618104 -0.8869787
## [2,]  0.8869787 -0.4618104
```

```r
library(ggplot2)
points <- mvrnorm(n=200, mu = mu, Sigma = sigma)
df <- data.frame(points)
colnames(df) = c("X","Y")
ggplot(data = df, aes(x=X, y=Y)) + geom_point(color="blue", alpha=0.5) + annotate(geom = 'segment', x=3
```
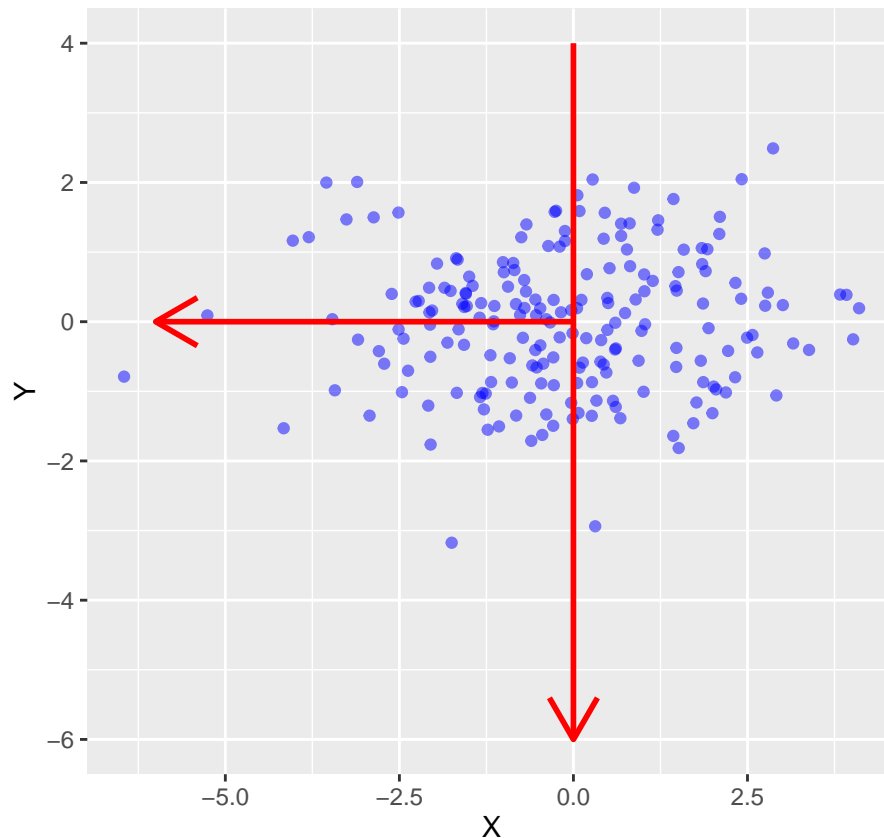
(b)

```r
mu <- matrix(c(0,0), ncol = 1)
sigma <- matrix(c(3, 0, 0, 1), nrow = 2)
eigen(sigma)
```

```
## $values
## [1] 3 1
##
## $vectors
##      [,1] [,2]
## [1,]   -1    0
## [2,]    0   -1
```

```r
points <- mvrnorm(n=200, mu = mu, Sigma = sigma)
df <- data.frame(points)
colnames(df) = c("X","Y")
ggplot(data = df, aes(x=X, y=Y)) + geom_point(color="blue", alpha=0.5) + annotate(geom = 'segment', x=0
```


2.

(a)

```r
X <- as.matrix(iris[1:50,1:4])
X <- scale(X, scale = FALSE)
Sx = var(X)
Sx
```

```
##              Sepal.Length Sepal.Width Petal.Length Petal.Width
## Sepal.Length   0.12424898 0.099216327   0.016355102 0.010330612
```

```
## Sepal.Width    0.09921633 0.143689796  0.011697959 0.009297959
## Petal.Length   0.01635510 0.011697959  0.030159184 0.006069388
## Petal.Width    0.01033061 0.009297959  0.006069388 0.011106122
```

(b)

```
EP <- eigen(Sx)
V <- EP$vectors
Lambda <- EP$values
V
```

```
##              [,1]       [,2]       [,3]        [,4]
## [1,] -0.66907840  0.5978840  0.4399628 -0.03607712
## [2,] -0.73414783 -0.6206734 -0.2746075 -0.01955027
## [3,] -0.09654390  0.4900556 -0.8324495 -0.23990129
## [4,] -0.06356359  0.1309379 -0.1950675  0.96992969
```

```
Lambda
```

```
## [1] 0.236455690 0.036918732 0.026796399 0.009033261
```
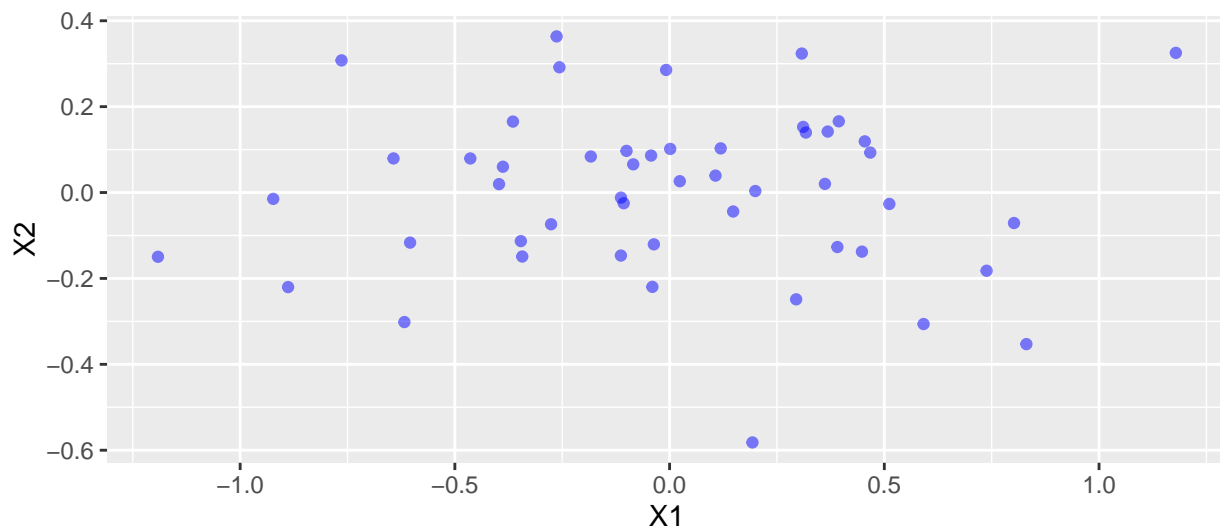
(c)

```
X[1:2,] %*% V
```

```
##          [,1]        [,2]       [,3]        [,4]
## 1 -0.1068424 -0.02489398 0.08216974 -0.03454175
## 2  0.3940472  0.16586593 0.13148092 -0.01755119
```

(d)

```
df2 <- data.frame(X %*% V)
colnames(df2) <- c('X1','X2','X3','X4')
ggplot(data = df2, aes(x=X1,y=X2)) + geom_point(color='blue',alpha=0.5) + coord_fixed()
```



```
cor(df2$X1, df2$X2)
```

```
## [1] 8.729423e-16
```

(e)

```
cumsum(Lambda)/sum(Lambda)
```

```
## [1] 0.7647237 0.8841229 0.9707854 1.0000000
```