Matrix Algebra Solutions

1. With R…

(a)

```
> A <- matrix(c(2, 3, 1, -4), nrow=2, byrow=TRUE)
> A
     [,1] [,2]
[1,]    2    3
[2,]    1   -4
> b <- matrix(c(4, -9), nrow=2, byrow=FALSE)
> b
     [,1]
[1,]    4
[2,]   -9
> x <- solve(A) %*% b
> x
     [,1]
[1,]   -1
[2,]    2
```

(b)

```
> A <- matrix(c(2, 2, 3, 1, -1, 4, -2, 4, 2), nrow=3, byrow=TRUE)
> A
     [,1] [,2] [,3]
[1,]    2    2    3
[2,]    1   -1    4
[3,]   -2    4    2
> b <- matrix(c(7, -3, 0), nrow=3, byrow=TRUE)
> b
     [,1]
[1,]    7
[2,]   -3
[3,]    0
> x <- solve(A) %*% b
> x
     [,1]
[1,]    3
[2,]    2
[3,]   -1
```

(c)

```
> A <- matrix(c(2, 2, 3, 1, -1, 4, 3, 1, 7), nrow=3, byrow=TRUE)
> b <- matrix(c(7, -3, 4), nrow=3, byrow=TRUE)
> x <- solve(A) %*% b
```

```
Error in solve.default(A) :
  system is computationally singular: reciprocal condition number = 1.90324e-17
```

The matrix is singular. Note the last row can be expressed as a linear combination of the first two. So all three column vectors (and all three row vectors) of A lie in the same plane, and so any linear combination of them will also lie in the same plane. The three planes intersect on a line in this case, and so there are an infinite number of solutions of the form

x_1 = .25 – 11/4 * x_3,
x_2 = 13/4 + 5/4 * x_3.

2. (with R…)

```
> A <- matrix(c(1, 4, 2, 6, 3, -1), nrow=3, byrow=TRUE)
> B <- matrix(c(2, 2, 3, 2, 1, 7), nrow=3, byrow=TRUE)
> C <- matrix(c(3, 7, 1, 4, 7, 5), nrow=2, byrow=TRUE)
> A
     [,1] [,2]
[1,]    1    4
[2,]    2    6
[3,]    3   -1
> B
     [,1] [,2]
[1,]    2    2
[2,]    3    2
[3,]    1    7
> C
     [,1] [,2] [,3]
[1,]    3    7    1
[2,]    4    7    5
> A+B
     [,1] [,2]
[1,]    3    6
[2,]    5    8
[3,]    4    6
> A-B
     [,1] [,2]
[1,]   -1    2
[2,]   -1    4
[3,]    2   -8
> A %*% C
     [,1] [,2] [,3]
[1,]   19   35   21
[2,]   30   56   32
[3,]    5   14   -2
> C %*% A
     [,1] [,2]
[1,]   20   53
```

```
[2,]   33   53
> A %*% t(B)
     [,1] [,2] [,3]
[1,]   10   11   29
[2,]   16   18   44
[3,]    4    7   -4
> t(B) %*% A
     [,1] [,2]
[1,]   11   25
[2,]   27   13
```

#3. $$\begin{pmatrix} a_{22}/D & -a_{12}/D \\ -a_{21}/D & a_{11}/D \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

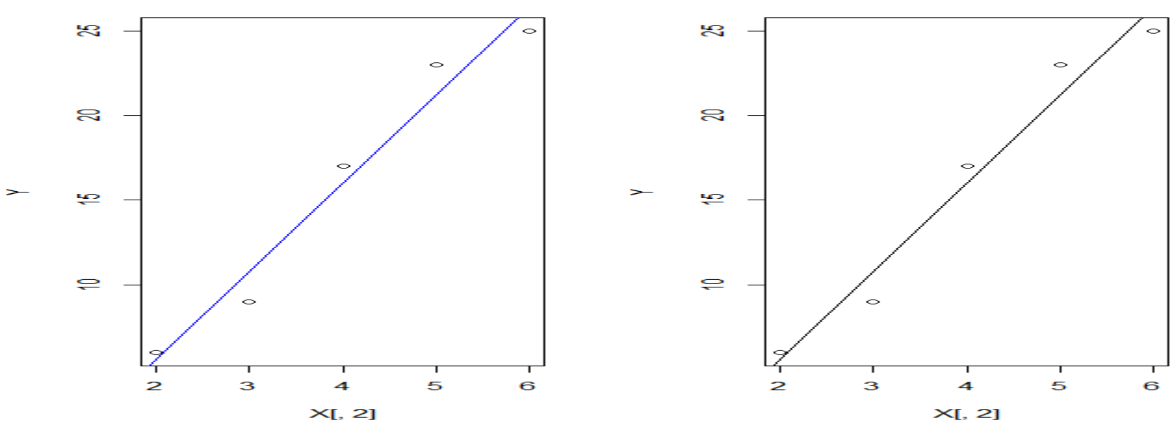
$$= \begin{pmatrix} \frac{a_{22}a_{11}}{D} - \frac{a_{12}a_{21}}{D} & \frac{a_{22}a_{12}}{D} - \frac{a_{12}a_{22}}{D} \\ \frac{-a_{21}a_{11}}{D} + \frac{a_{11}a_{21}}{D} & \frac{-a_{21}a_{12}}{D} + \frac{a_{11}a_{22}}{D} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11}a_{22} - a_{12}a_{21}} & \frac{a_{22}a_{12} - a_{12}a_{22}}{a_{11}a_{22} - a_{12}a_{21}} \\ \frac{a_{11}a_{21} - a_{21}a_{11}}{a_{11}a_{22} - a_{12}a_{21}} & \frac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

4.

```
> X <- matrix(c(1, 2, 1, 3, 1, 5, 1, 6, 1, 4), nrow=5, byrow=TRUE)
> Y <- matrix(c(6, 9, 23, 25, 17), nrow=5)
> X
     [,1] [,2]
[1,]    1    2
[2,]    1    3
[3,]    1    5
[4,]    1    6
[5,]    1    4
> Y
     [,1]
[1,]    6
[2,]    9
[3,]   23
[4,]   25
[5,]   17
> beta <- solve(t(X) %*% X) %*% t(X) %*% Y
> beta
     [,1]
[1,] -4.8
[2,]  5.2
> par(mfrow=c(1, 2))
> plot(Y ~ X[,2])
> abline(a=-4.8, b=5.2, col="blue")
> out <- lm(Y ~ X[,2])
> plot(Y ~ X[,2])
> abline(out)
```



They're the same!!!!!!!!!!!!!!!!!!

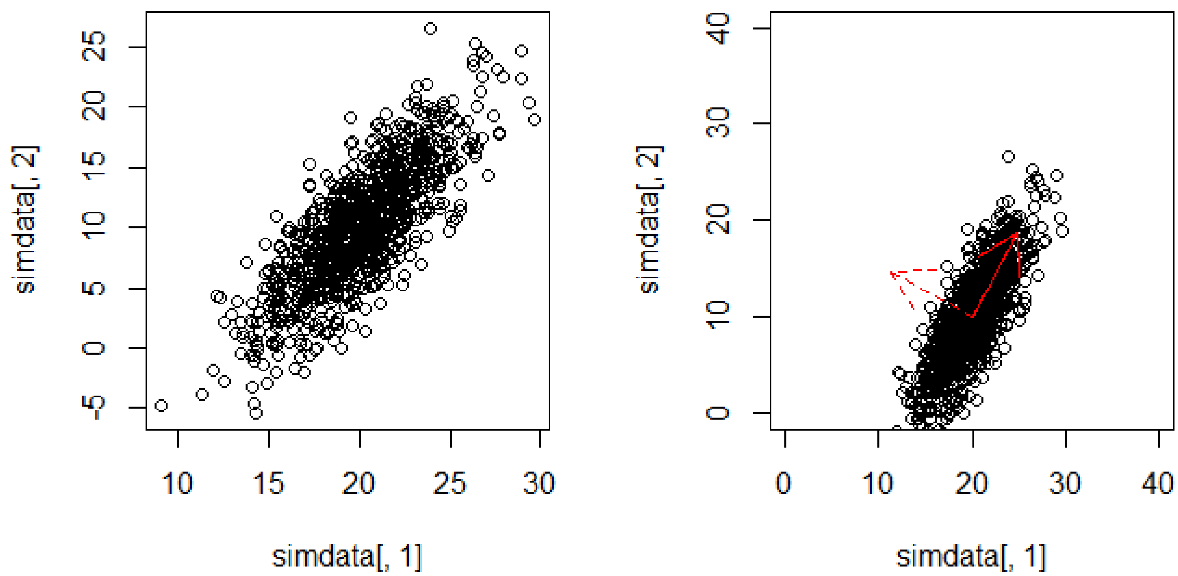
5. With R…

```
> diag(3)
     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1
> eigen(diag(3))
$values
[1] 1 1 1

$vectors
     [,1] [,2] [,3]
[1,]    0    0    1
[2,]    0    1    0
[3,]    1    0    0

> A <- matrix(c(1, 2, 3, 2, 2, 0, 2, 0, 5), nrow=3, byrow=TRUE)
> eigen(A)
$values
[1]  6.353111  2.858373 -1.211483

$vectors
          [,1]       [,2]       [,3]
[1,] 0.5426618 -0.3700921  0.8188170
[2,] 0.2493214 -0.8623110 -0.5099307
[3,] 0.8020953  0.3456177 -0.2636462

> B <- matrix(c(1, 1, 1, 2, 2, 2, 1, 2, 3), nrow=3, byrow=TRUE)
> eigen(B)
$values
[1]  5.236068e+00  7.639320e-01 -5.717676e-16

$vectors
           [,1]       [,2]       [,3]
[1,] -0.3162278 -0.3162278 -0.4082483
[2,] -0.6324555 -0.6324555  0.8164966
[3,] -0.7071068  0.7071068 -0.4082483
```

The first two matrices have rank 3. Matrix B has rank 2.

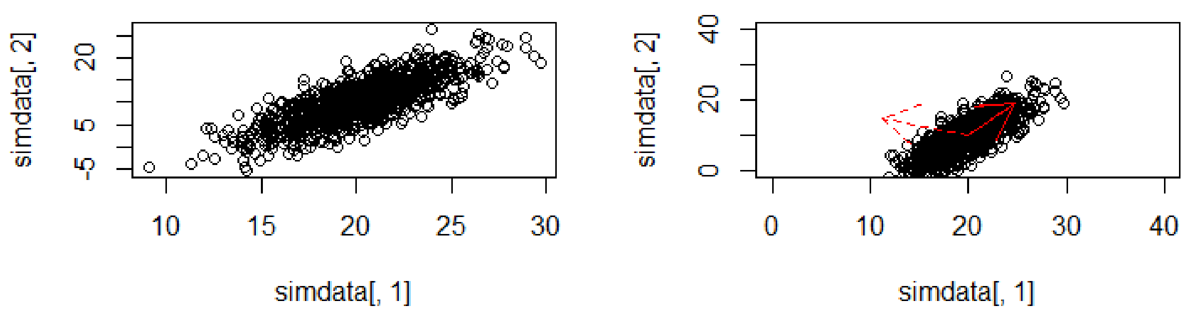
7.

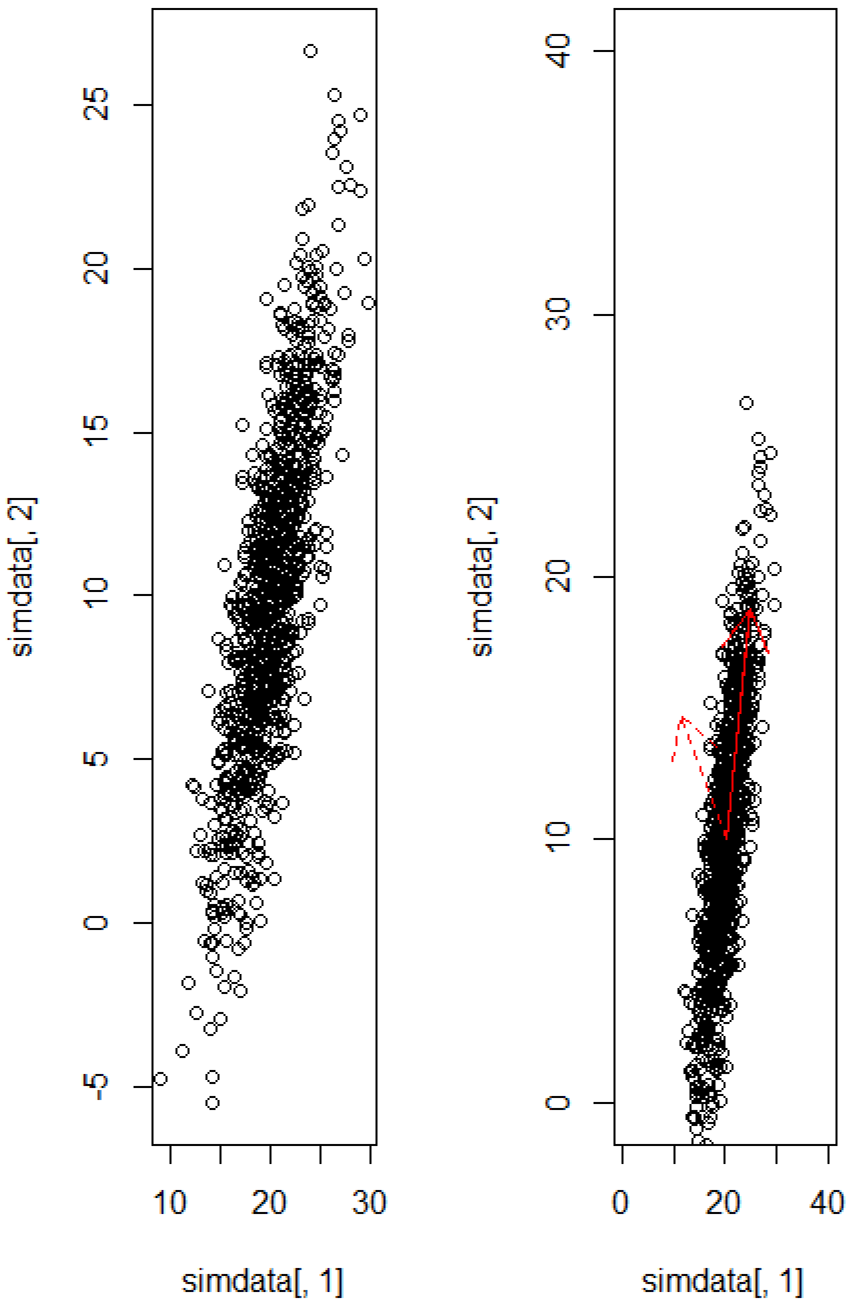
```
> library(MASS)
> mu <- matrix(c(20, 10), byrow=FALSE)
> sigma2 <- matrix(c(9, 12, 12, 25), byrow=TRUE, nrow=2)
> mu
     [,1]
[1,]   20
[2,]   10
> sigma2
     [,1] [,2]
[1,]    9   12
[2,]   12   25
> simdata <- mvrnorm(n=1000, mu=mu, Sigma=sigma2, empirical = FALSE)
> head(simdata)
         [,1]      [,2]
[1,] 25.59492 13.646068
[2,] 15.41152  1.647309
[3,] 20.69279 10.436536
[4,] 21.52288 13.144593
[5,] 23.73087  9.193433
[6,] 20.21615 11.841881
> plot(simdata[,1], simdata[,2])
> plot(simdata[,1], simdata[,2], xlim=c(0,40), ylim=c(0,40))
> windows()
> par(mfrow=c(1,2))
> plot(simdata[,1], simdata[,2])
> plot(simdata[,1], simdata[,2], xlim=c(0,40), ylim=c(0,40))
> xbar <- mean(simdata[,1])
> ybar <- mean(simdata[,2])
> eig <- eigen(sigma2)
> arrows(0+xbar, 0+ybar, x1=xbar+10*eig$vectors[1,1], y1=ybar+10*eig$vectors[2,1], col="red", lty=20)
> arrows(0+xbar, 0+ybar, x1=xbar+10*eig$vectors[1,2], y1=ybar+10*eig$vectors[2,2], col="red", lty=20)
```



Note that since the variance-covariance matrix is symmetric, the eigenvectors will be orthogonal. Also, the eigenvectors R gives are orthonormal, meaning they are orthogonal and have length 1. I've multiplied their lengths by 10 so you can see them better.

The scaling has a dramatic effect on whether or not the eigenvectors look orthogonal (see graphs below).

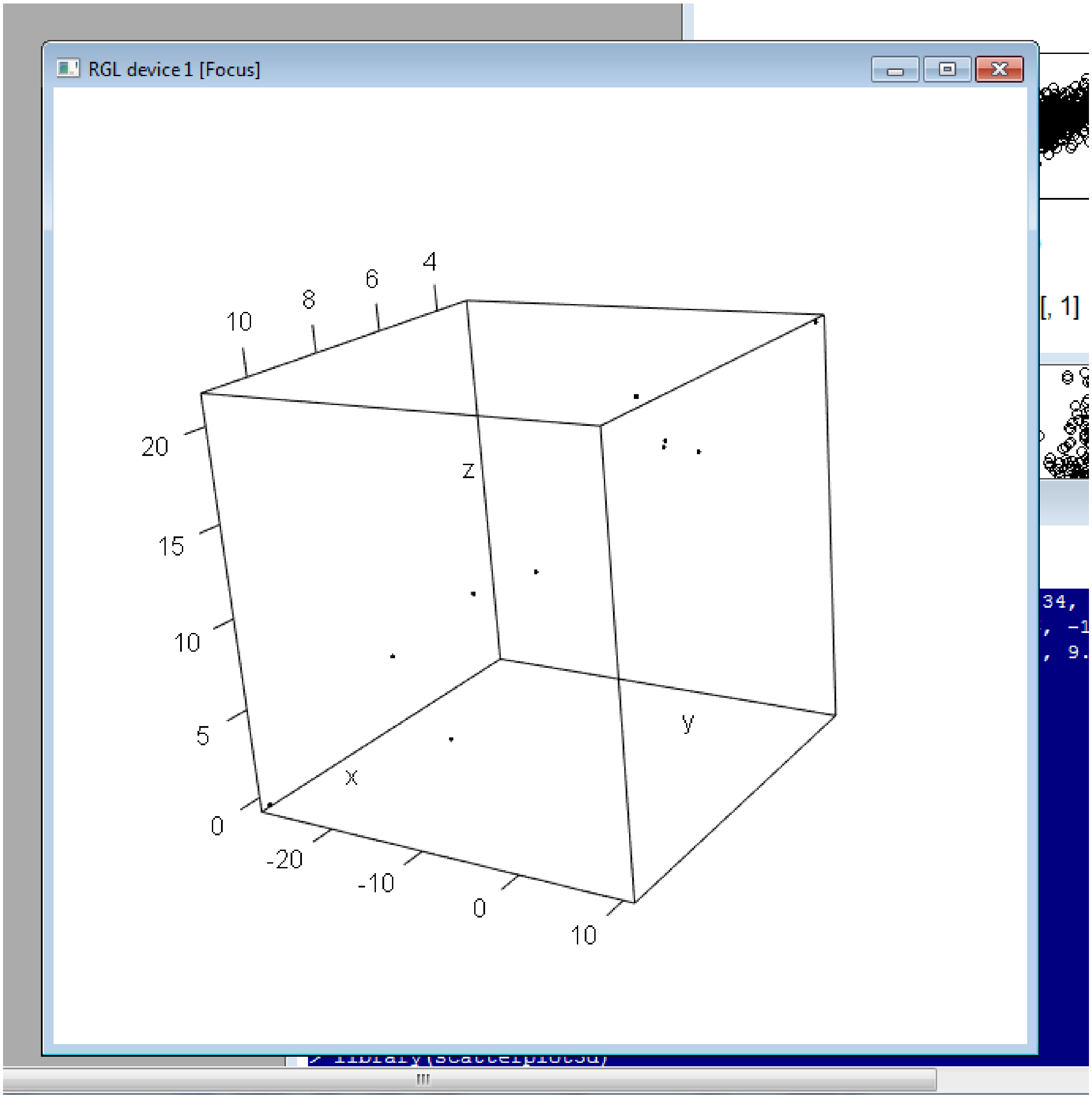




Now, the main thing we want to notice here is that the eigenvectors of the covariance matrix point in the directions of the most variation. Can you see this? This will be especially important when we talk about principle components analysis later in the course.

8.

```
> x<- c(3.04, 4.55, 7.47, 6.33, 5.00, 5.2, 9.66, 11.11, 8.34, 7.1)
> y<- c(10.48, 1.76, -19.38, .5, -.63, .08, -16.88, -27.66, -11.95, -9.05)
> z<- c(21.34, 15.14, -.1, 19.26, 15.60, 16.14, 7.31, -.51, 9.81, 10.06)
> cbind(x, y, z)
        x      y     z
 [1,]  3.04  10.48 21.34
 [2,]  4.55   1.76 15.14
 [3,]  7.47 -19.38 -0.10
 [4,]  6.33   0.50 19.26
 [5,]  5.00  -0.63 15.60
 [6,]  5.20   0.08 16.14
 [7,]  9.66 -16.88  7.31
 [8,] 11.11 -27.66 -0.51
 [9,]  8.34 -11.95  9.81
[10,]  7.10  -9.05 10.06
> library(rgl)
Warning message:
package 'rgl' was built under R version 3.1.0
> library(scatterplot3d)
Warning message:
package 'scatterplot3d' was built under R version 3.1.0
> plot3d(x, y, z)
> X <- cbind(x, y, z)
> cov(X)
          x          y         z
x   6.098133 -27.29360 -15.35686
y -27.293600 137.21282  85.02753
z -15.356856  85.02753  56.64756
> eig <- eigen(cov(X))
> eig
$values
[1] 196.16400844   3.75239996   0.04210938

$vectors
           [,1]       [,2]       [,3]
[1,]  0.1623784  0.4883704  0.8573958
[2,] -0.8345383 -0.3956526  0.3834122
[3,] -0.5264781  0.7777875 -0.3433183
```

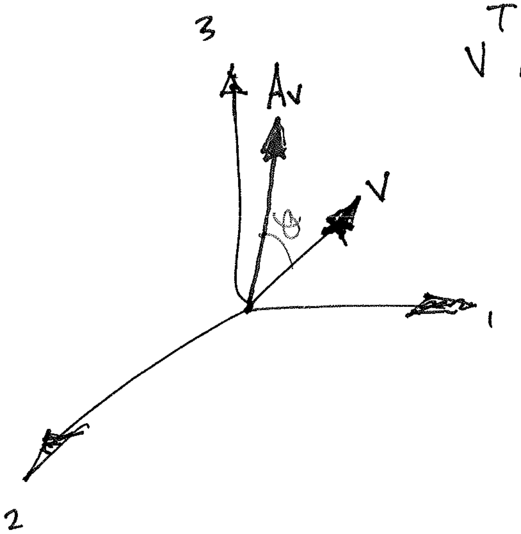


The eigenvectors of the covariance matrix point along the directions of the most variation. For example, the third eigenvector (0.8573958, 0.3834122, -0.3433183) seems to point in the direction of the most variation… remember this vector starts at the origin, and mind the axis scaling here. You can rotate the graph around in R and see this.

$\boxed{9}$ Let $A_{n\times n}$ be positive semidefinite. Then for any vector $v_{n\times 1}$, $v^T A v \geq 0$ (this is the defining characteristic of positive semidefinite matrices).

Note that Av is $n\times 1$ (a column vector).

Then $v^T A v$ can be thought of as the dot product of v^T and Av, and so



$$v^T A v = v^T \cdot (Av)$$

θ is the angle between v and Av.

$$= \|v^T\| \, \|Av\| \cos\theta$$

$$\geq 0.$$

by the positive semidefinite assumption

Thus, since $\|v^T\| \geq 0$, $\|Av\| \geq 0$, it must be that $\cos\theta \geq 0$ which holds iff $\theta \in [-\pi/2, \pi/2]$.

10. Let $\underset{n\times 1}{X}$ denote a random vector:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}.$$

These are random variables

Also, let m denote the expectation or mean vector:

$$m = E[X] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_n] \end{bmatrix}.$$

Then the covariance matrix σ_X^2 can be written as

$$\sigma_X^2 = \overbrace{E\left[(X-m)(X-m)^T\right]}^{\text{Remember this is an } n\times n \text{ matrix:}} = \begin{bmatrix} Var(X_1) & Cov(X_1,X_2) & \cdots \\ Cov(X_2,X_1) & Var(X_2) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Let v be any arbitrary vector of dimension $n\times 1$. Then

$$v^T \sigma_X^2 v = v^T E\left[(X-m)(X-m)^T\right] v$$

Since v is a constant vector...

$$= E\left[\underbrace{v^T(X-m)}_{U^T}\underbrace{(X-m)^T v}_{U_{n\times 1}}\right]$$

$$= E\left[U^T \cdot U\right]$$

Now U is a random vector, but ANY vector (random or not) dotted with itself is non negative. So $E[U^T\cdot U] \geq 0$.

In FACT, $E[U^T U] = 0$ iff $U^T U = 0$ with probability of 100%. So in many cases, σ_X^2 is positive definite meaning $v^T \sigma_X^2 v \boxed{>} 0$.