1) $\mathbf{X} = (X_{(1)}, X_{(2)}, \ldots, X_{(m)})$ is orthogonal, so $X_{(i)}^t X_{(j)} = 0$ for $i \neq j$ and 1 for $i = j$.
$tr(\mathbf{P}) = tr(\mathbf{X}\mathbf{X}^t) = tr(\mathbf{X}^t\mathbf{X}) = tr\left((X_{(i)}^t X_{(j)})_{i,j}\right) = tr(I_{m*m}) = m$

2) Since A is symmetric, $A = KDK^t$, where is K is orthogonal and D is the diagonal matrix with elements, $d_1, \ldots, d_n$.
$\because \lim\limits_{k\to\infty} KA^kK^t = K(\lim\limits_{k\to\infty} A^k)K^t$and K is orthogonal.
$\therefore \lim\limits_{k\to\infty} KA^kK^t = 0 \Leftrightarrow \lim\limits_{k\to\infty} A^k = 0$

$\lim\limits_{k\to\infty} KA^kK^t = \lim\limits_{k\to\infty} (KAK^t)^k = \lim\limits_{k\to\infty} D^t = 0$
$\Leftrightarrow \lim d_i^k = 0 \qquad i = 1, \ldots, n \Leftrightarrow |d_i| < 1 \qquad i = 1, \ldots, n$
$\therefore \lim\limits_{k\to\infty} A^k = 0 \Leftrightarrow |d_i| < 1 \qquad i = 1, \ldots, n$

3)(a)

$\forall x \in \mathbb{R}^n$
$x^t(A - C)x = x^t(A - B)x + x^t(B - C)x$
$\because x^t(A - B)x > 0$ and $x^t(B - C)x > 0 \qquad$ since$A \rhd B, B \rhd C$
$x^t(A - C)x > 0 \Rightarrow A \rhd C$

(b)

$\because A \rhd B$ and $B \rhd A$
$\forall x \in \mathbb{R}^n, \quad x^t(A - B)x \geq 0$ and $x^t(B - A)x \geq 0$
$\Rightarrow x^t(A - B)x = 0 \quad \forall x \in \mathbb{R}^n \Rightarrow A = B$

4)(a)Let $A = SS^t$ and $B = UU^t$ are the Cholesky decomposition. Then $A^{-1} = S^{-t}S^{-1}$ and $B^{-1} = U^{-t}U^{-1}$.
Assume $x^tBx = 1$, then $x^tAx \geq 1$ since $A \rhd B$. This means at the minimum point, the term is still true. So differentiae $x^tAx - \lambda(x^tBx - 1)$, we get:

$$Ax = \lambda Bx, \Rightarrow x^tAx = \lambda x^tBx = \lambda$$

According to lecture2 page 122, $\lambda y = U^{-1}AU^{-t}y = U^{-1}SS^tU^{-t}y$, where eigenvalue $\lambda \geq 1$.
For $x^tA^{-1}x = 1$, now we calculate the minimum value of $x^tB^{-1}x$. Then $\mu y = (S^{-t})^{-1}B^{-1}(S^{-t})^{-t}y = S^tU^{-t}U^{-1}Sy$. Note $S^tU^{-t}U^{-1}S$ and $U^{-1}AU^{-t} = U^{-1}SS^tU^{-t}$ have the same eigenvalues (AB and BA have the same eigenvalues). So eigenvalue $\mu \geq 1 \quad \forall \mu$. Thus $x^tB^{-1}x - x^tA^{-1}x \geq 0$, namely

$B^{-1} \rhd A^{-1}$.

(b) In the notation of (a), $U^{-1}SS^tU^{-t}$ and $SS^tU^{-t}U^{-1} = AB^{-1}$ have the same eigenvalues. Thus $det(AB^{-1}) = det(U^{-1}SS^tU^{-t}) \geq 1 \Rightarrow det(A) \geq det(B)$.

$C = A - B$ is a semi-positive definite matrix and suppose $\exists i, \quad c_{ii} < 0$. Let $x = e_i$, then $x^t(A-B)x < 0$, which contradicts the condition. So $\forall i, \quad c_{ii} \geq 0 \Rightarrow tr(A-B) \geq 0 \Rightarrow tr(A) \geq tr(B)$.

5) For multivariate Gaussian, $M_{\mathbf{X}}(t) = e^{\mu^t t} + \frac{1}{2}t^t\Sigma t$. For univariate Gaussian, $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t}$. Distributions and moment generating functions are one-to-one map.

$\Rightarrow \forall k \in \mathbb{R}, \alpha \in \mathbb{R}^n, \quad M_{\mathbf{X}}(k\alpha) = \mathbb{E}e^{k\alpha^t\mathbf{X}} = e^{\mu^t\alpha k + \frac{1}{2}(\alpha^t\Sigma\alpha)k^2} = M_{\alpha^t\mathbf{X}}(k)$. So $\forall \alpha, \quad \alpha^t\mathbf{X} \sim N(\mu^t\alpha, \alpha^t\Sigma\alpha)$

$\Leftarrow \forall t \in \mathbb{R}^n, \quad M_{t^t\mathbf{X}}(1) = e^{t^t\mu + \frac{1}{2}t^t\Sigma t} = M_{\mathbf{X}}(t)$ So $\mathbf{X} \sim N_{\mathbf{X}}(\mu, \Sigma)$

6)Let A be a $n*n$ positive definite. A $= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, where A_{11} is a $m*m$ $(m < n)$ matrix.

Let x_{1*m} be nonzero and $y = (x^t_{1*m}, 0^t_{1*(n-m)})^t$, then $y^tAy = x^tA_{11}x > 0$. Thus, A_{11} is positive definite. So is A_{22} in similar method. When $m = 1$, $a_{11} = A_{11} > 0$. And A_{22} is positive definite, so its element in first row and first column, a_{22} is greater than 0. Apply the above method to A_{22}. In this way ,we get $a_{ii} > 0, \quad i = 1, \ldots, n$.

Let $a = (\sqrt{a_{11}}, \frac{a_{12}}{\sqrt{a_{11}}}, \ldots, \frac{a_{1n}}{\sqrt{a_{11}}})^t$, $A_{12} = (a_{12}, \ldots, a_{1n})$ and $A_{21} = A^t_{12}$. Then,

$$A - aa^t = \left(\begin{array}{c|ccc} 0 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \\ 0 & & & \end{array}\right)$$

where, $A' = A_{22} - A_{21}a^{-1}_{11}A_{12}$

Let $y = (-\frac{A_{12}}{a_{11}}x, x^t)^t$, where $x \in \mathbb{R}^{n-1}$, then $y^tAy = x^t(A_{22} - A_{21}a^{-1}_{11}A_{12})x > 0$. So $A_{22} - A_{21}a^{-1}_{11}A_{12}$ is positive definite. Then we could apply the same method to $A_{22} - A_{21}a^{-1}_{11}A_{12}$. In this way, we construct a upper triangular matrix T such that $A = T^tT$.

Suppose there are two upper triangular matrices T and S, which satisfy Cholesky decomposition.

1. Since $T^tT = S^tS = A$, then $t^2_{11} = s^2_{11} = a_{11}$. Then $t_{11} = s_{11} > 0$.
2. $\because t_{11}t_{1i} = s_{11}s_{1i} = a_{1i} \therefore t_{1i} = s_{1i}, \quad i = 2, \ldots, n$
3. $\because t^2_{12} + t^2_{22} = s^2_{12} + s^2_{22} \therefore t_{22} = s_{22} > 0$
4. keep applying the same technique until $t_{nn} = s_{nn} > 0$.

So $T = S$.

7)

$$q_1 = \frac{v_1}{\|v_1\|} = (0.2822163, 0.5644325, 0.7525767, 0.1881442)$$
$$q_2 = \frac{v_2 - < v_2, q_1 > q_1}{\|v_2 - < v_2, q_1 > q_1\|} = (0.6531533, -0.6473902, 0.1498410, 0.3630764)$$
$$q_3 = \frac{v_3 - < v_3, q_1 > q_1 - < v_3, q_2 > q_2}{\|v_3 - < v_3, q_1 > q_1 - < v_3, q_2 > q_2\|} = (0.45151564, -0.02620403, 0.07256501, -0.88892142)$$

8)

$$\frac{1}{\sqrt{2}} \begin{pmatrix} V & V \\ U & -U \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & -D \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} V^t & U^t \\ V^t & -U^t \end{pmatrix} = \frac{1}{2} \begin{pmatrix} VD & -VD \\ UD & UD \end{pmatrix} \begin{pmatrix} V^t & U^t \\ V^t & -U^t \end{pmatrix}$$
$$= \frac{1}{2} \begin{pmatrix} 0 & 2VDU^t \\ 2UDV^t & 0 \end{pmatrix} = \begin{pmatrix} 0 & A^t \\ A & 0 \end{pmatrix}$$

9) (a)

```r
#read matrix
A = matrix(c(10, 7, 8, 7,
             7, 6, 6, 5,
             8, 6, 10, 9,
             7, 5, 9, 10), nrow = 4, byrow = TRUE)
eigen(A)
```

```
## $values
## [1] 30.4375830  4.0468404  0.8764022  0.6391745
##
## $vectors
##            [,1]       [,2]       [,3]       [,4]
## [1,] -0.5263856  0.5551497  0.4771728 -0.4324732
## [2,] -0.3915365  0.4740296 -0.3372886  0.7129036
## [3,] -0.5490379 -0.2839068 -0.6609934 -0.4254904
## [4,] -0.5178559 -0.6216926  0.4707763  0.3517006
```

(b) The accuracy is the l_2 norm of the last two iterated eigenvector.

```r
initial_pos = matrix(rnorm(4), ncol=1)
error = 1
accurate = 1e-4
itera = 0

last_pos = initial_pos
while(error >= accurate)
{
  current_pos = A %*% last_pos
  current_pos = current_pos / sqrt(sum(current_pos^2))
  error = sqrt(sum((current_pos-last_pos)^2))
  last_pos = current_pos
  itera = itera + 1
}
current_pos = A %*% last_pos
lambda = sqrt( sum(current_pos^2) / sum(last_pos^2) )
print(list("eigenvalue" = lambda, "eigenvector" = current_pos,
           "iteration" = itera))
```

```
## $eigenvalue
## [1] 30.43758
##
## $eigenvector
##          [,1]
## [1,] 16.02190
## [2,] 11.91742
## [3,] 16.71139
## [4,] 15.76229
##
## $iteration
## [1] 7
```

c)

```r
svd(A)
```

```
## $d
## [1] 30.4375830  4.0468404  0.8764022  0.6391745
##
## $u
##            [,1]       [,2]       [,3]       [,4]
## [1,] -0.5263856 -0.5551497  0.4771728 -0.4324732
## [2,] -0.3915365 -0.4740296 -0.3372886  0.7129036
## [3,] -0.5490379  0.2839068 -0.6609934 -0.4254904
## [4,] -0.5178559  0.6216926  0.4707763  0.3517006
##
## $v
##            [,1]       [,2]       [,3]       [,4]
## [1,] -0.5263856 -0.5551497  0.4771728 -0.4324732
## [2,] -0.3915365 -0.4740296 -0.3372886  0.7129036
## [3,] -0.5490379  0.2839068 -0.6609934 -0.4254904
## [4,] -0.5178559  0.6216926  0.4707763  0.3517006
```

d)

```r
svd(A[1:3,])
```

```
## $d
## [1] 26.084899  2.835989  0.731587
##
## $u
##            [,1]       [,2]        [,3]
## [1,] -0.6181017  0.4778392 -0.62419547
## [2,] -0.4604058  0.4235486  0.78014941
## [3,] -0.6371630 -0.7695949  0.04179621
##
## $v
##            [,1]       [,2]       [,3]
## [1,] -0.5559217  0.5594071 -0.6103705
## [2,] -0.4183311  0.4473206  0.7686105
## [3,] -0.5397329 -0.4696578  0.1439266
## [4,] -0.4739604 -0.5161294 -0.1263764
```