

Properties of matrices 行列の特徴

ryamada

2016年12月22日

- 1 似ている行列 (相似) Matrix similarity
- 2 相似行列に共通する特徴量
 - 2.1 Rでの計算
- 3 Exercise 1
 - 3.1 Exercise 1-1
 - 3.2 Exercise 1-2
- 4 対称行列
- 5 Exercise 2
 - 5.1 Exercise 2-1
- 6 距離行列 分散共分散行列 対称行列 固有値が正 正定値行列
 - 6.1 距離行列
 - 6.2 分散共分散行列
 - 6.3 固有値がすべて正 正定値
- 7 多変量正規分布
 - 7.1 Rでの多変量正規乱数の発生
- 8 Exercise 3
 - 8.1 Exercise 3-1
 - 8.2 Exercise 3-3
- 9 多次元の任意な分布
 - 9.1 Exercise 4
 - 9.2 Exercise 4-1
 - 9.3 Exercise 4-2
 - 9.4 Exercise 4-3
 - 9.5 Exercise 4-4
 - 9.6 Exercise 4-5
 - 9.7 Exercise 4-6
 - 9.8 Exercise 4-7

1 似ている行列 (相似) Matrix similarity

When we use matrices for statistical analyses, we change bases so that “better” axes should grab data sets in simpler or better ways.

The change of bases are achieved with linear transformation. Therefore matrices that are transformed each other by linear transformation should be considered “similar”.

Among the similar matrices, the simplest one tends to represent data structure the “best” way from data interpretation standpoint.

The simplest one among the similar matrices is Jordan normal.

Or the simplest one is diagonal, if diagonalizable, and, in many cases the matrices in statistical analyses are diagonalizable.

統計解析に行列を使うとき、データセットが単純に理解できるような基底の取り直しをすることがよくある。

基底の取り直しは線形変換で行うので、線形変換で互に行きあえる行列同士は「相似(似ている)」関係にあると言う。

相互に相似な関係にある行列の中で、最も単純な行列を代表とすることがあり、それを使うことでデータ解釈が最もわかりやすくなる。

相似な行列の中で最も単純なのは、ジョルダン標準形のそれである。

対角化可能な行列なら、ジョルダン標準形は対角化行列である。実際、統計解析で現れる行列の多くは、対角化可能である。

任意の正方行列 M はジョルダン標準形 J と逆行列を持つ行列 R とを用いて、次のように表せる。

$$M = RJR^{-1}$$

$M' = R'JR'^{-1}$ となると、 M と M' は相似であると言う。

以下では、ジョルダン標準形が対角行列であるようなもの(対角化可能行列)に絞って話を進めるが、特に断らなければ、提示している性質等は対角化可能行列に限るものではない。

2 相似行列に共通する特徴量

相互に相似な行列は見かけは違うが、「本質的に同じ」なので、その「本質」に関する性質は共通である。

- 階数(Rank)
- 行列式(Determinant)
- トレース(跡) (Trace)
- 固有値(Eigen values)

2.1 Rでの計算

Rank, Determinant, Trace, Eigen valuesは次のように計算できる

```
library(Matrix)
d <- 4
M <- matrix(rnorm(d^2), d, d)
eigen(M)[[1]]
```

```
## [1] -1.1741118+0.9913840i -1.1741118-0.9913840i  0.5786552+0.7952877i
## [4]  0.5786552-0.7952877i
```

```
rankMatrix(M)[[1]]
```

```
## [1] 4
```

```
det(M)
```

```
## [1] 2.284221
```

```
sum(diag(M)) # Trace
```

```
## [1] -1.190913
```

3 Exercise 1

3.1 Exercise 1-1

逆行列を持つ行列Rを使って、対角行列Jを変換した行列M($M = RJR^{-1}$)がJと相似であることを、固有値が一致していることを確認して確かめよ。

そして、ランク、行列式、トレースが同一であることを確かめよ。

```
d <- 4
J <- diag(rnorm(d))
R <- matrix(rnorm(d^2), d, d)
M <- R %*% J %*% solve(R)
eigen(J)[[1]]
```

```
## [1] 0.4955596 0.4456923 -0.1680365 -0.9141442
```

```
eigen(M)[[1]]
```

```
## [1] -0.9141442 0.4955596 0.4456923 -0.1680365
```

```
rankMatrix(J)[[1]]
```

```
## [1] 4
```

```
rankMatrix(M)[[1]]
```

```
## [1] 4
```

```
det(J)
```

```
## [1] 0.03392731
```

```
det(M)
```

```
## [1] 0.03392731
```

```
sum(diag(J)) # Trace
```

```
## [1] -0.1409288
```

```
sum(diag(M)) # Trace
```

```
## [1] -0.1409288
```

3.2 Exercise 1-2

行列が相互に相似かどうかは、(すべての固有値が異なるならば)固有値のセットが等しいことで確認できる。
したがって相似な行列に共通する特徴は、固有値と紐づけて算出できるはずである。

以下を確かめよ。

$$\text{determinant} = \prod_{i=1}^n \lambda_i$$

$$\text{trace} = \sum_{i=1}^n \lambda_i$$

Exercise 1-3

d 個の異なる固有値を持つ行列を作り、それと相似な行列を複数作れ。

それらの固有値のセットが一致すること、トレース・行列式が一致することを確認せよ。

また、トレースの計算は、対角成分の和を取る方法と、固有値の和を取る方法の両方を実施せよ。行列式の計算も`det()`関数を使用する方法と固有値の積を取る方法の両方を実施せよ。

```
## [[1]]
## [1] -0.0004091244
##
## [[2]]
## [1] -0.0004091244
##
## [[3]]
## [1] -0.0004091244
##
## [[4]]
## [1] -0.0004091244
##
## [[5]]
## [1] -0.0004091244
##
## [[6]]
## [1] -0.0004091244
##
## [[7]]
## [1] -0.0004091244
##
## [[8]]
## [1] -0.0004091244
##
## [[9]]
## [1] -0.0004091244
##
## [[10]]
## [1] -0.0004091244
```

```
## [[1]]
## [1] -11.612590  8.736930  3.415908 -1.971686
##
## [[2]]
## [1] -2.47443331  1.55825946 -0.43920254 -0.07606178
##
## [[3]]
## [1] -0.6768818 -0.1819945 -0.8542280  0.2816662
##
## [[4]]
## [1]  1.5433476  0.1180999 -2.4963997 -0.5964859
##
## [[5]]
## [1]  0.6039185 -1.4078521 -0.4202507 -0.2072539
##
## [[6]]
## [1]  1.2912410 -5.6506502  2.5405990  0.3873721
##
## [[7]]
## [1]  0.6876451  1.0517309  0.5574392 -3.7282534
##
## [[8]]
## [1] -1.49928360 -0.43924960  0.58623485 -0.07913983
##
## [[9]]
## [1] -2.08933580 -0.07569291  0.66716061  0.06642992
##
## [[10]]
## [1] 12.6960258 -11.8824854 -0.5644387 -1.6805399
```

```
## [[1]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[2]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[3]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[4]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[5]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[6]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[7]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[8]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[9]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
##
## [[10]]
## [1] -1.46323305  0.12643222 -0.05256953 -0.04206782
```

4 対称行列

$M = M^T$ である正方行列を対称行列と言う。

対称行列は

$$M = RJR^{-1} = RJR^T$$

$$R^T = R^{-1}$$

$$RR^T = R^T R = I$$

と対角化で現れる行列 R が特別な性質(直交行列)を持つ。

確かめてみる。

```
d <- 5
n.iter <- 10
for(i in 1:n.iter){
  M <- matrix(rnorm(d^2), d, d)
  M <- M + t(M) # 対称行列化
  eigen.out <- eigen(M)
  print(round(eigen.out[[2]] %*% t(eigen.out[[2]]), 8))
}
```

```

##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0

```

```
## [3,] 0 0 1 0 0
## [4,] 0 0 0 1 0
## [5,] 0 0 0 0 1
```

5 Exercise 2

5.1 Exercise 2-1

対角行列と直交行列によって相似な関係にある行列を複数作り、それらが対称行列であることを確かめよ。

なお、直交行列をランダムに発生するには、`GPRotation::Rnadam.Start()`関数を用いるとよい。


```

## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"

```

```

##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
## [1] "zero?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
## [1] "unit?"
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0

```

```
## [2,] 0 1 0 0 0
## [3,] 0 0 1 0 0
## [4,] 0 0 0 1 0
## [5,] 0 0 0 0 1
## [1] "zero?"
##      [,1] [,2] [,3] [,4] [,5]
## [1,] 0 0 0 0 0
## [2,] 0 0 0 0 0
## [3,] 0 0 0 0 0
## [4,] 0 0 0 0 0
## [5,] 0 0 0 0 0
## [1] "unit?"
##      [,1] [,2] [,3] [,4] [,5]
## [1,] 1 0 0 0 0
## [2,] 0 1 0 0 0
## [3,] 0 0 1 0 0
## [4,] 0 0 0 1 0
## [5,] 0 0 0 0 1
## [1] "zero?"
##      [,1] [,2] [,3] [,4] [,5]
## [1,] 0 0 0 0 0
## [2,] 0 0 0 0 0
## [3,] 0 0 0 0 0
## [4,] 0 0 0 0 0
## [5,] 0 0 0 0 0
```

6 距離行列 分散共分散行列 対称行列 固有値が正 正定値行列

行列には相似という間柄があり、それは固有値で決まることが分かった。

また、相似な行列の間で保存される性質があることもわかった。

また、 RMR^{-1} という変換が相似な行列を作ることもわかった。

また対象行列の場合 $RMR^{-1} = RMR^T$ と書け、 R が直交行列であることもわかった。

これらは、一般的な行列の話であったが、この節では、統計解析でよく登場する特別な行列を取り上げ、その場合に固有値・相似行列の特徴がどのような性質を持つかを確認する。

6.1 距離行列

n 個の要素があって、それらの要素ペアの距離を行列とすることがある。

要素間の関連を調べるときの基本的なデータ構造であり

距離の場合、自身から自身への距離は0である。自身以外への距離は正である。

AからBへの距離とBからAへの距離は同じなので、距離行列は、次の性質を持つ。

- 対角成分が0
- 対称
- 非対角成分は正

このような行列では、以下のような特徴がある。

```
d <- 5
M <- matrix(abs(rnorm(d^2)), d, d)
diag(M) <- 0
M <- M+t(M) # 対称にする

eigen.out <- eigen(M)
trace <- sum(diag(M)) # zero
trace2 <- sum(eigen.out[[1]])
determinant <- det(M)
determinant2 <- prod(eigen.out[[1]])

eigen.out[[1]] # Negative and positive eigen values
```

```
## [1]  5.5789604  0.1556252 -0.7320391 -1.2424011 -3.7601454
```

```
round(eigen.out[[2]] %*% t(eigen.out[[2]]), 6) # 固有値ベクトルが作る行列は直交
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
```

6.2 分散共分散行列

k 個の変数、 n 個のサンプルがあるとき、 $k \times k$ の分散共分散行列が算出できる。

対角成分は k 個の変数の分散である。トレースは k 個の変数の分散の和である。

k 個の変数の分散の和が固有値の和である。

対称行列である。これは距離行列の場合と同じ。

```
k <- 5
n <- 100
X <- matrix(0, n, k)
X[,1] <- rnorm(n)
for(i in 2:k){
  X[,i] <- rnorm(n) + X[,i-1]
}
M <- cov(X)
M
```

```
##      [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] 0.9237420 0.8361656 0.8443777 0.8773991 0.7178514
## [2,] 0.8361656 1.7435079 1.9513828 1.8298230 1.6240981
## [3,] 0.8443777 1.9513828 3.0667371 2.9675999 2.6337225
## [4,] 0.8773991 1.8298230 2.9675999 3.8318982 3.4805980
## [5,] 0.7178514 1.6240981 2.6337225 3.4805980 4.1203933
```

```
round(M-t(M), 8) # 対称行列
```

```
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    0    0    0    0    0
## [2,]    0    0    0    0    0
## [3,]    0    0    0    0    0
## [4,]    0    0    0    0    0
## [5,]    0    0    0    0    0
```

```
det(M)
```

```
## [1] 0.698088
```

```
sum(diag(M))
```

```
## [1] 13.68628
```

```
eigen.out <- eigen(M)
round(eigen.out[[2]] %*% t(eigen.out[[2]]), 8)
```

```
##      [, 1] [, 2] [, 3] [, 4] [, 5]
## [1,]    1    0    0    0    0
## [2,]    0    1    0    0    0
## [3,]    0    0    1    0    0
## [4,]    0    0    0    1    0
## [5,]    0    0    0    0    1
```

```
eigen.out[[1]]
```

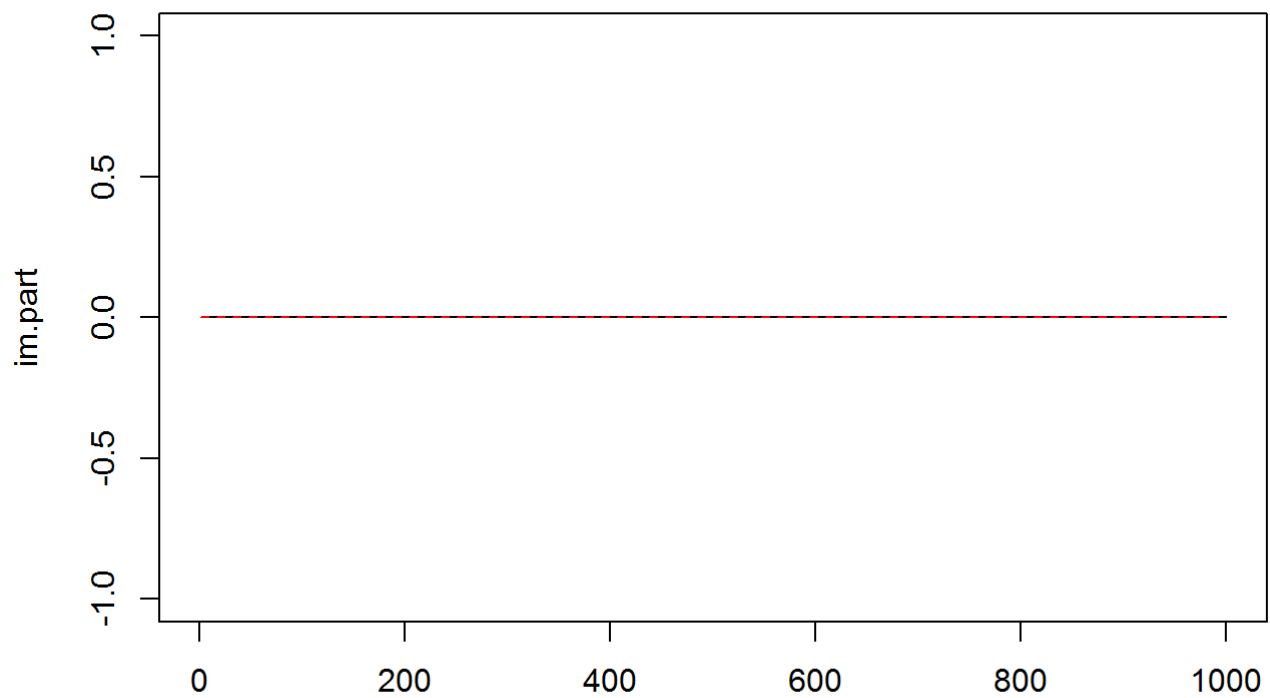
```
## [1] 11.0961899  1.4065392  0.6052861  0.3876234  0.1906399
```

分散共分散行列は、対象行列であって、対角成分はすべて正。

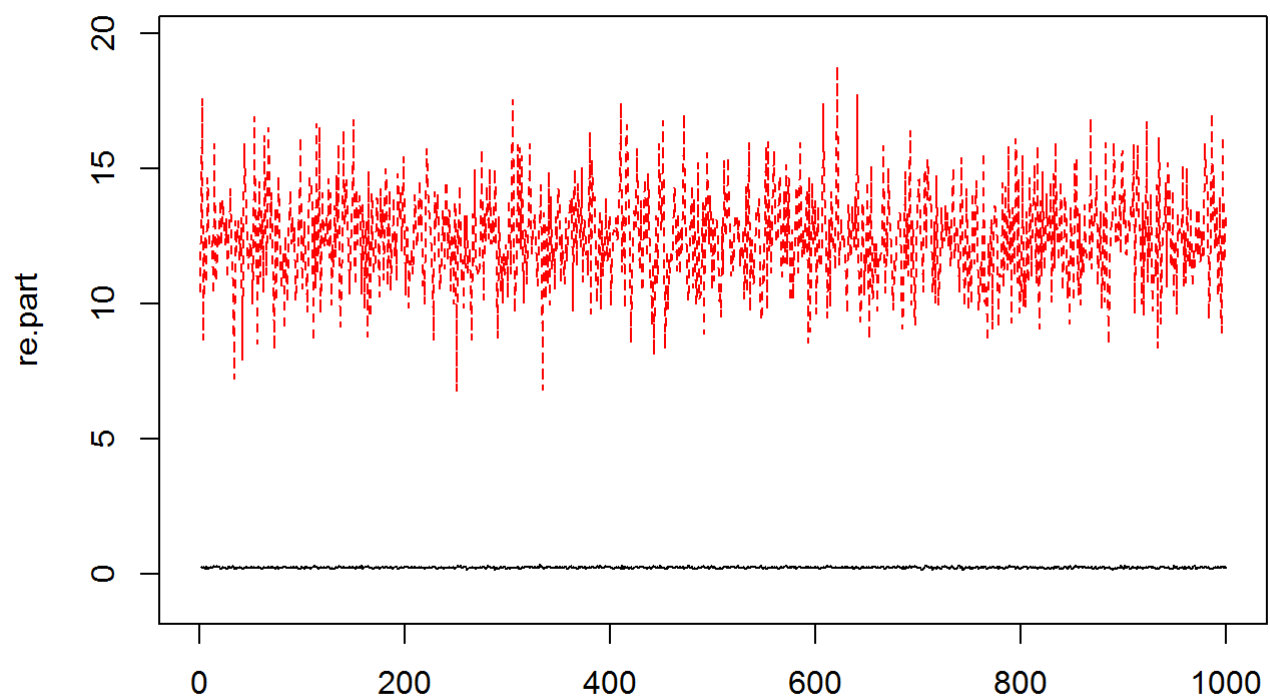
また、以下に示すように、固有値もすべて正。

```
n.iter <- 1000
im.part <- matrix(0,n.iter,2)
re.part <- matrix(0,n.iter,2)
for(i in 1:n.iter){
  d <- sample(2:50,1)
  X <- matrix(0,n,k)
  X[,1] <- rnorm(n)
  for(j in 2:k){
    X[,j] <- rnorm(n) + X[,j-1]
  }
  M <- cov(X)
  M
  eigen.out <- eigen(M)
  im.part[i,] <- range(Im(eigen.out[[1]]))
  re.part[i,] <- range(Re(eigen.out[[1]]))
}

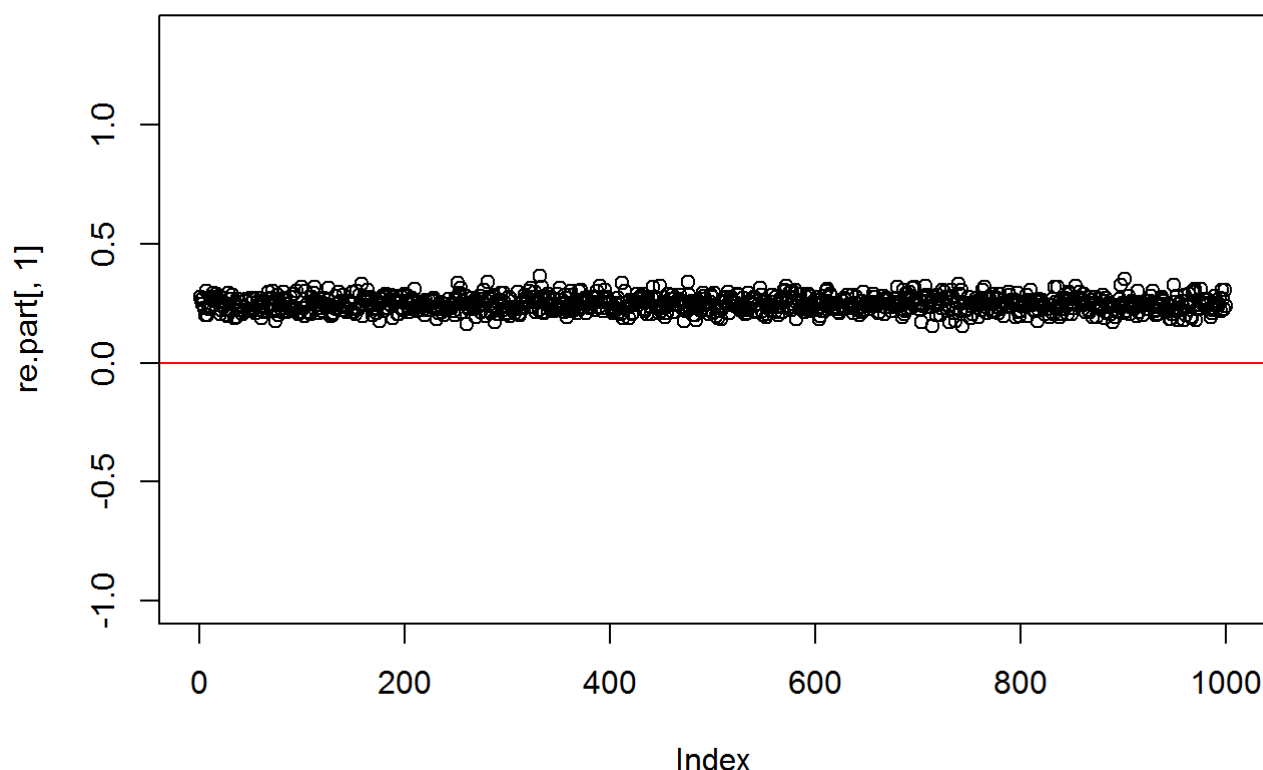
matplot(im.part, type="l")
```



```
matplot(re.part, type="l", ylim=c(-1, max(re.part)+1))
```



```
plot(re.part[, 1], ylim=c(-1, max(re.part[, 1]+1)))
abline(h=0, col=2)
```



```
min(re.part[, 1])
```

```
## [1] 0.1546466
```

6.3 固有値がすべて正 正定値

固有値がすべて正であるような対称行列は正定値行列と呼ばれている。

正定値行列 M は、任意の 0 でないベクトル v に対して

$$v^T M v > 0$$

となる。

正定値行列の特殊例でもある、単位行列の場合は、以下のように、「ベクトルの長さの二乗」である。

これが 0 ベクトルなら 0 となり、そうでなければ正になる、というのは、ベクトルの「長さ」を定義するのに適切な性質を持つ行列を正定値行列と言う、というようにも理解できる。

逆に言えば、統計解析で登場する行列は、この性質を満たすものが多いこととなる。

$$v^T M v = v^T v$$

このような M は、ベクトルの長さの定め方でもあり、 $v^T M v$ の値を決めるという意味では、「計量」を定めるとも言える。

対称行列の固有値分解 $M = RJR^{-1} = RJR^T$ で、 V は回転行列の性質を持つから、
対角成分がすべて正である対角行列と回転行列から作る行列は正定値行列であることになる。

```
library(GPArotation)
d <- 5
#V <- Random.Start(d)
V <- matrix(rnorm(d^2), d, d)
lambdas <- abs(eigen(V)$values)
M <- V %*% diag(lambdas) %*% solve(V)
eigen(M)[[1]]
```

```
## [1] 1.2801155 0.9444490 0.8374103 0.3624414 0.2057298
```

```
lambdas
```

```
## [1] 0.9444490 0.2057298 0.8374103 1.2801155 0.3624414
```

```
n.vec <- 1000
vecs <- matrix(rnorm(n.vec*d), nrow=d)
```

```
range(diag(t(vecs) %*% M %*% vecs)) # t(v) M vのみを取り出し、そのrangeを確認。すべて正
```

```
## [1] -16.65512 37.21324
```

7 多変量正規分布

多変量正規分布は、統計解析でよく使われる。

$$N(\mu, \Sigma)$$

とパラメタ表現される。

変数の数 n 個の多変量正規分布の場合、 μ は長さ n のベクトルであり、各変数の平均値を表す。

Σ は $n \times n$ 行列であり、正定値行列であることになっている。

正定値行列は、対角化可能で、 $\Sigma = RJR^{-1}$ と対角化したときに、 R が直交行列 ($R^{-1} = R^T$) であるようなものであった。

これを利用して、多変量正規分布からの乱数発生を試みる。

7.1 Rでの多変量正規乱数の発生

```
library(mvtnorm)
```

```
## Warning: package 'mvtnorm' was built under R version 3.4.3
```



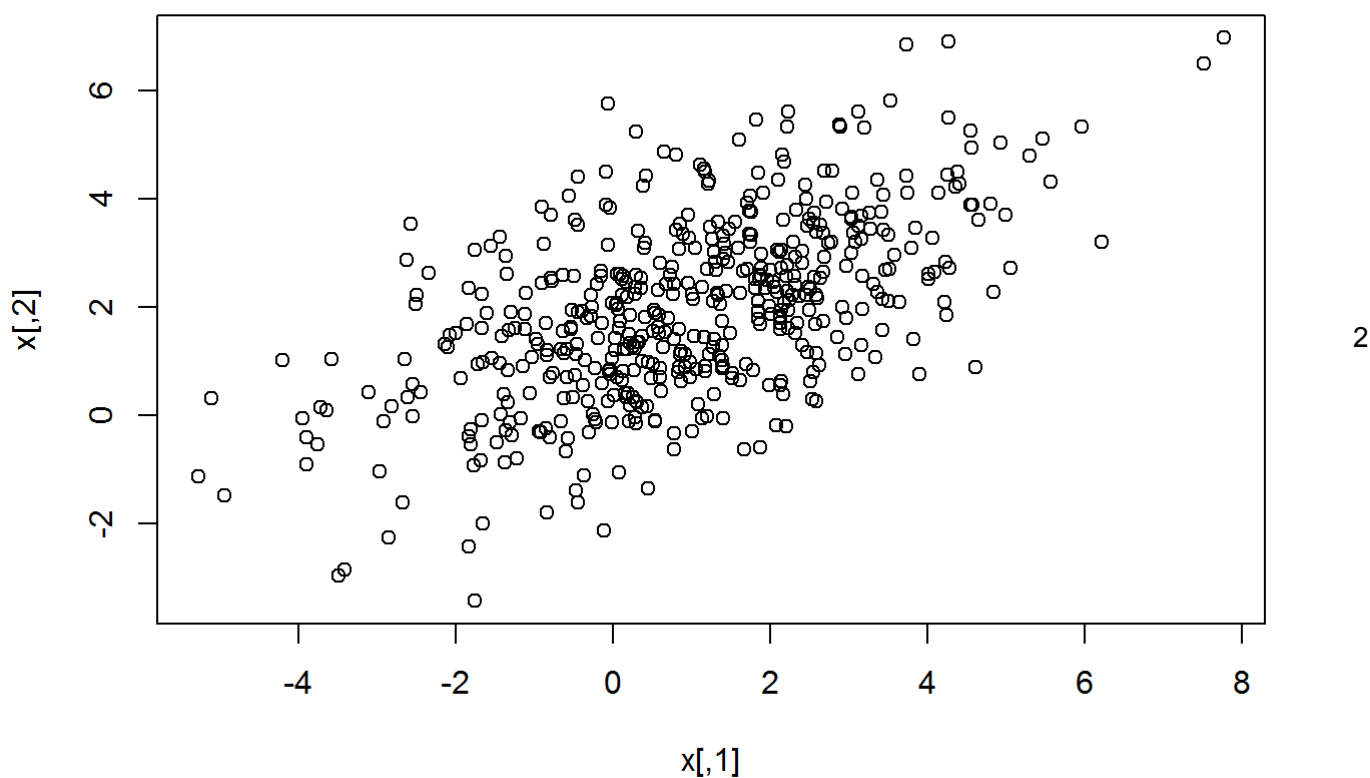
```
#help(rmvnorm)
sigma <- matrix(c(4, 2, 2, 3), ncol=2)
x <- rmvnorm(n=500, mean=c(1, 2), sigma=sigma)
colMeans(x)
```

```
## [1] 0.9370393 1.9506145
```

```
var(x)
```

```
##          [,1]      [,2]
## [1,] 4.048278 1.947546
## [2,] 1.947546 2.793245
```

```
plot(x)
```



変数標準正規分布は、3変数の平均がともに0で、分散共分散行列が 2×2 の単位行列であるようなものである。

```
d <- 2
m <- rep(0, d)
sigma <- diag(rep(1, d))
m
```

```
## [1] 0 0
```

```
sigma
```

```
##      [,1] [,2]  
## [1,]    1    0  
## [2,]    0    1
```

```
n <- 10000  
x <- rmvnorm(n=n, mean=m, sigma=sigma)  
apply(x, 2, mean)
```

```
## [1] 0.008866332 0.005640604
```

```
apply(x, 2, var)
```

```
## [1] 0.9806272 0.9927339
```

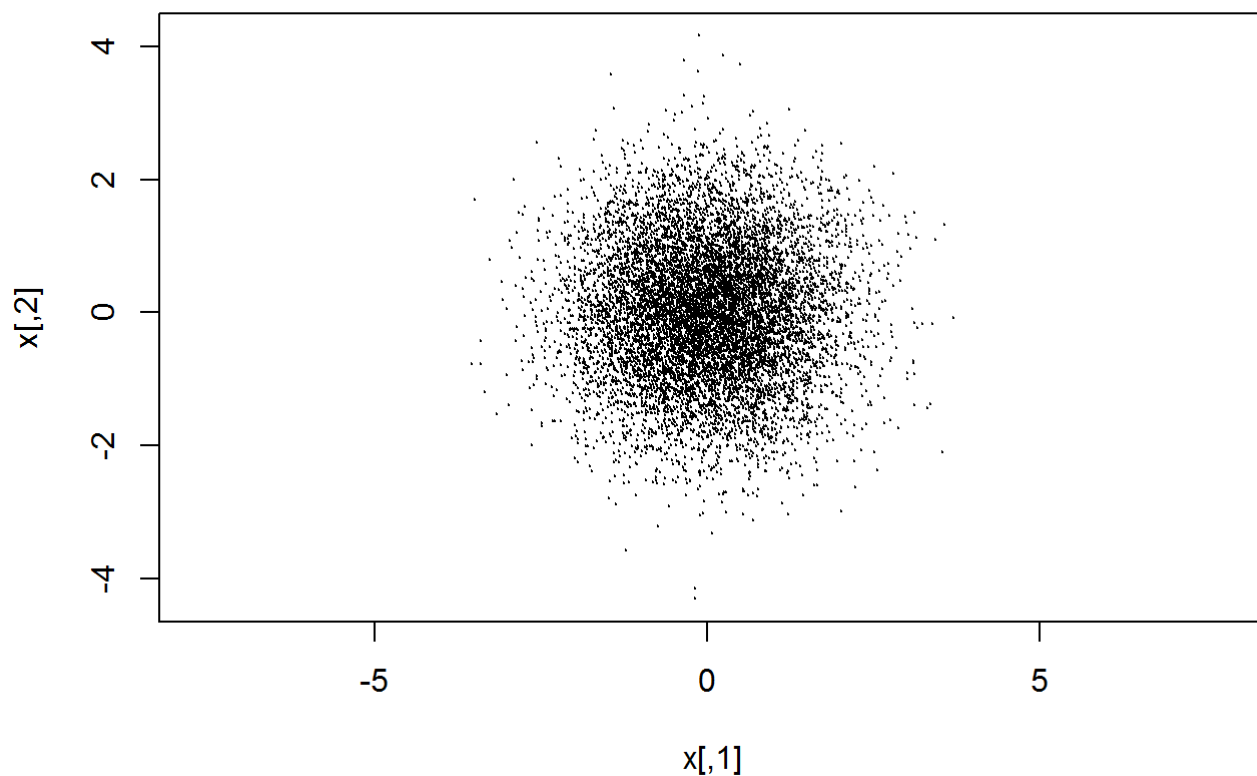
```
cov(x)
```

```
##      [,1]      [,2]  
## [1,] 0.980627229 -0.001679324  
## [2,] -0.001679324 0.992733870
```

```
round(cov(x), 8)
```

```
##      [,1]      [,2]  
## [1,] 0.98062723 -0.00167932  
## [2,] -0.00167932 0.99273387
```

```
plot(x, asp=TRUE, pch=20, cex=0.1)
```



8 Exercise 3

8.1 Exercise 3-1

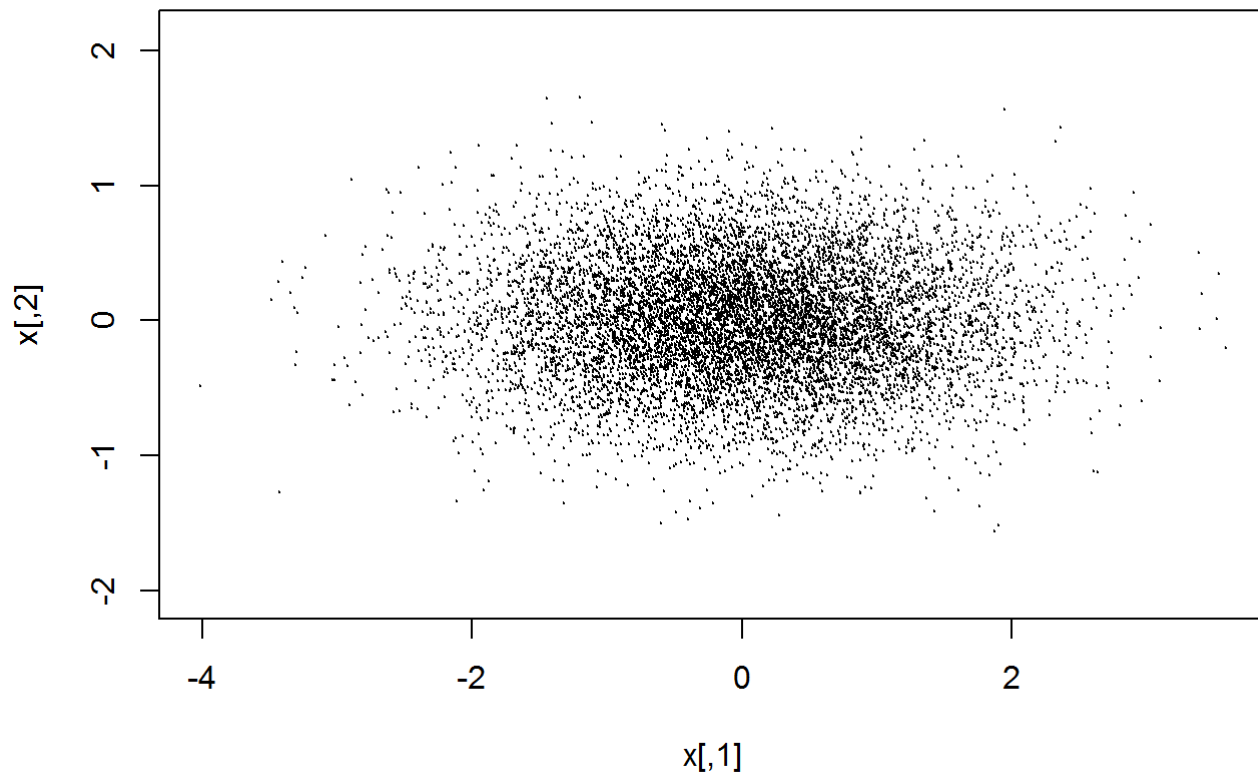
2変量正規分布において、各変数の分散を(1,1)からずらして、乱点を発生せよ。なお、分散共分散行列は対角行列のままとする。

```
## [1] 9.462167e-03 -4.070587e-05
```

```
## [1] 0.9951902 0.1943384
```

```
##           [, 1]      [, 2]
## [1, ] 0.9951902 -0.00308122
## [2, ] -0.00308122 0.19433838
```

```
##           [, 1]      [, 2]
## [1, ] 0.9861931 0.0000000
## [2, ] 0.0000000 0.1943617
```



Exercise 3-2

4-1で用いたSigmaと相似な行列で対称なものは正定値行列であるから、`rmvnorm::rmvnorm()`のsigma引数として適当なはずである。

そのようなsigmaを作り、乱点を発生せよ。

```
R <- Random.Start(d)
sigma3 <- R %*% sigma2 %*% t(R)
x <- rmvnorm(n=n, mean=m, sigma=sigma3)
apply(x, 2, mean)
```

```
## [1] 0.004337685 -0.004068588
```

```
apply(x, 2, var)
```

```
## [1] 0.2968694 0.8813722
```

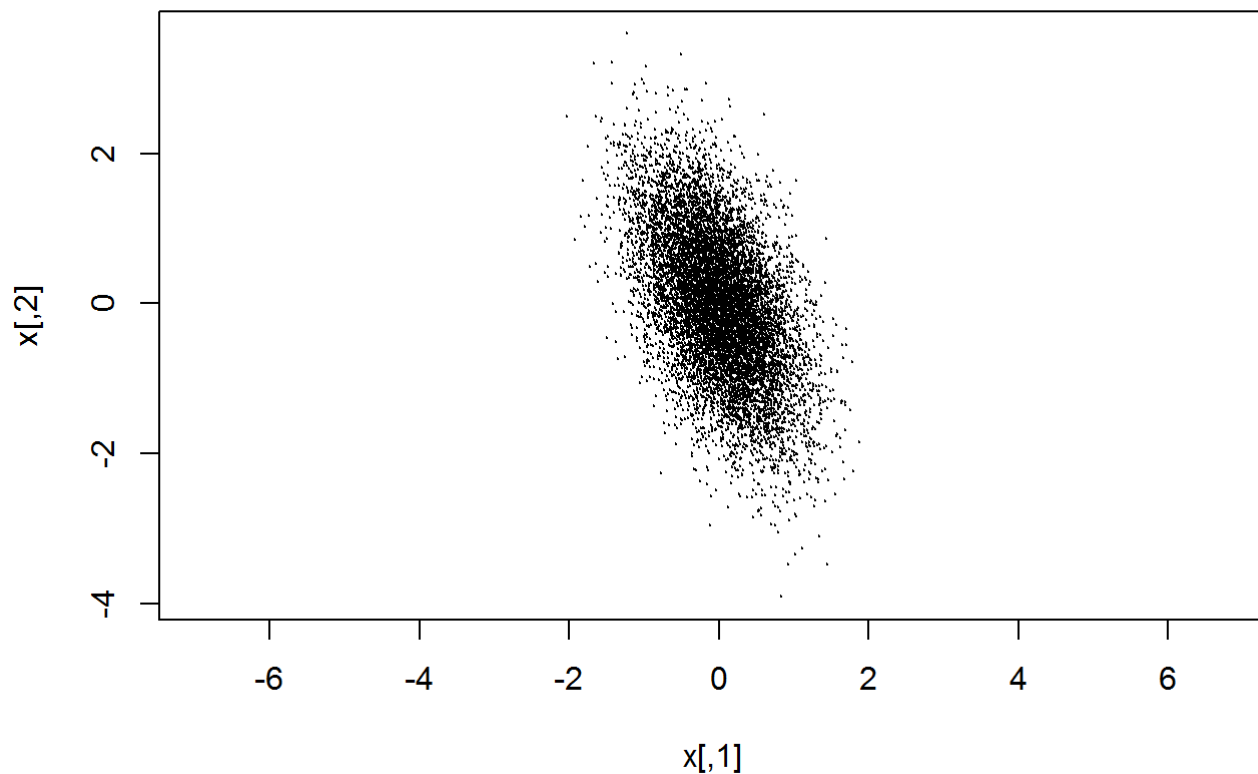
```
round(cov(x), 8)
```

```
##           [,1]      [,2]
## [1,] 0.2968693 -0.2644093
## [2,] -0.2644093 0.8813722
```

```
sigma3
```

```
##           [, 1]      [, 2]
## [1, ]  0.2928167 -0.2612783
## [2, ] -0.2612783  0.8877382
```

```
plot(x, asp=TRUE, pch=20, cex=0.1)
```



8.2 Exercise 3-3

分散共分散行列を RJR^{-1} にて変換すると、乱点分布が回転し、楕円の軸が斜めになった。

実際、分散共分散行列の変換に用いた直交行列 R を利用して、乱点を変換すると、楕円の軸を正すことができる。

以下のコードを読み、変換した分散共分散行列に従う2変量正規分布の軸が補正されていることを、コードの各行の処理にコメントを加えることで確認せよ。

```
R <- Random.Start(d)
sigma3 <- R %*% sigma2 %*% t(R)
x <- rmvnorm(n=n, mean=m, sigma=sigma3)
apply(x, 2, mean)
```

```
## [1] 0.007693247 0.010971247
```

```
apply(x, 2, var)
```

```
## [1] 0.9831017 0.1957530
```

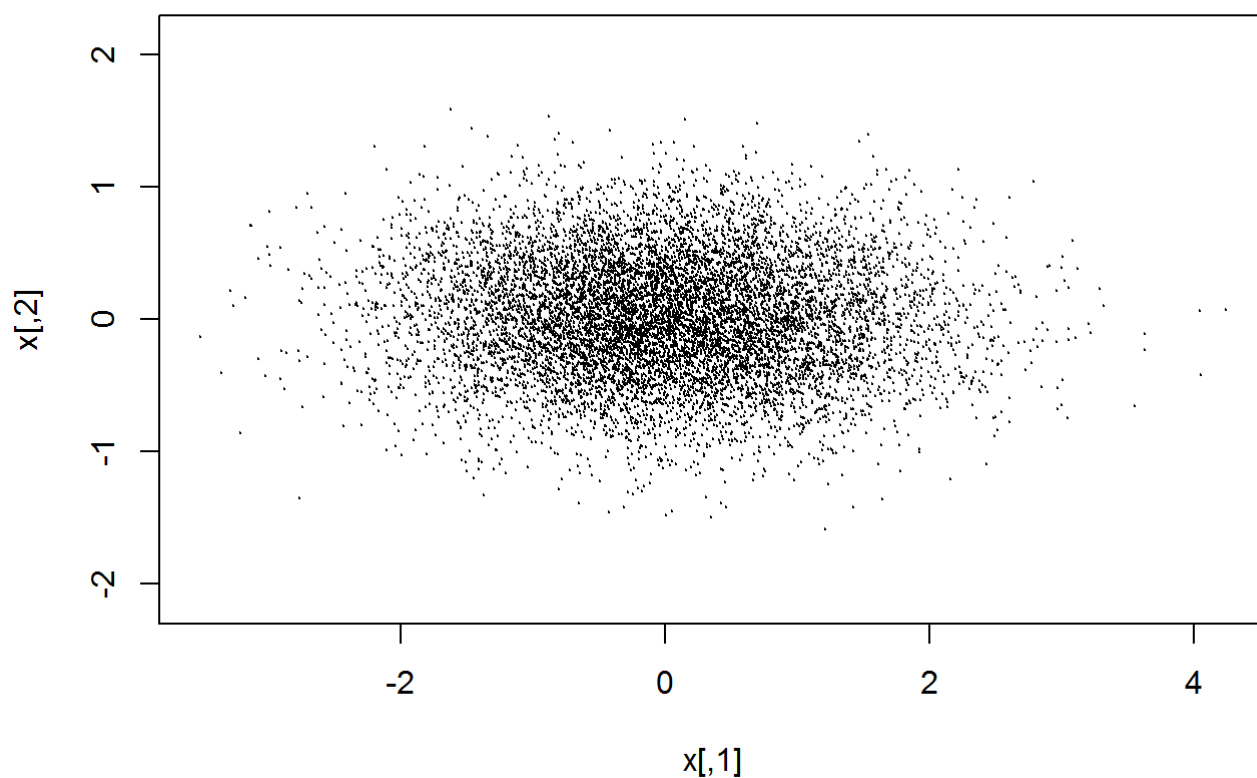
```
cov(x)
```

```
##           [, 1]      [, 2]
## [1, ] 0.98310167 -0.02002558
## [2, ] -0.02002558 0.19575299
```

```
round(cov(x), 8)
```

```
##           [, 1]      [, 2]
## [1, ] 0.98310167 -0.02002558
## [2, ] -0.02002558 0.19575299
```

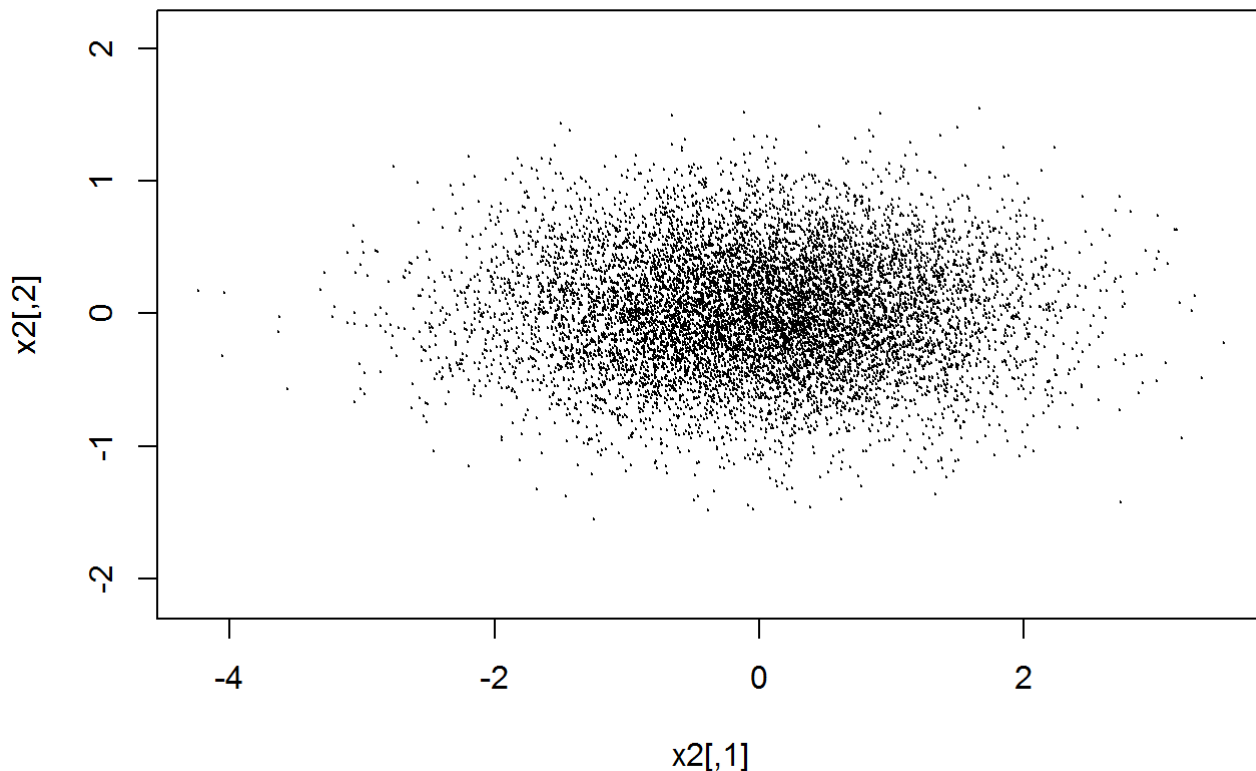
```
plot(x, asp=TRUE, pch=20, cex=0.1)
```



```
x2 <- t(t(R) %*% t(x))
cov(x2)
```

```
##           [, 1]      [, 2]
## [1, ] 0.9836103474 0.0005076271
## [2, ] 0.0005076271 0.1952443077
```

```
plot(x2, asp=TRUE, pch=20, cex=0.1)
```



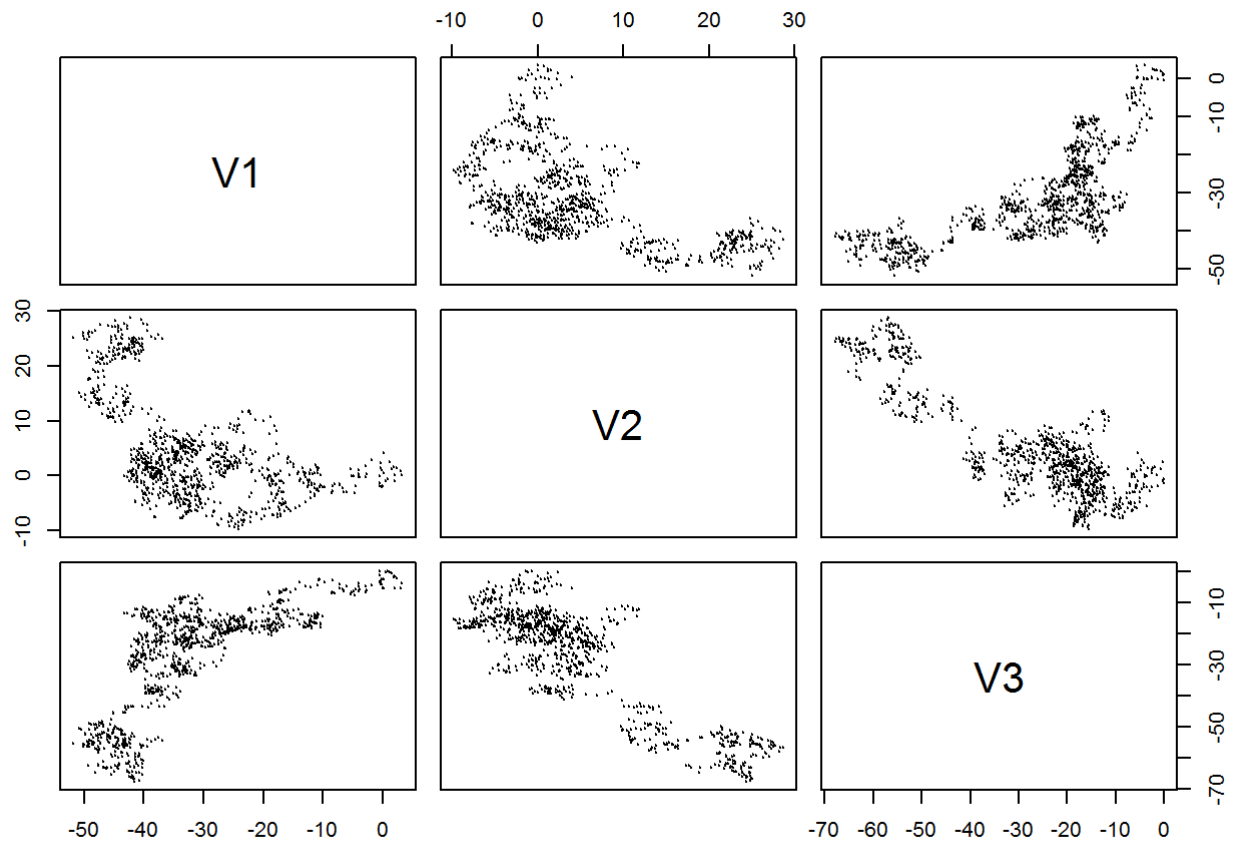
9 多次元の任意な分布

9.1 Exercise 4

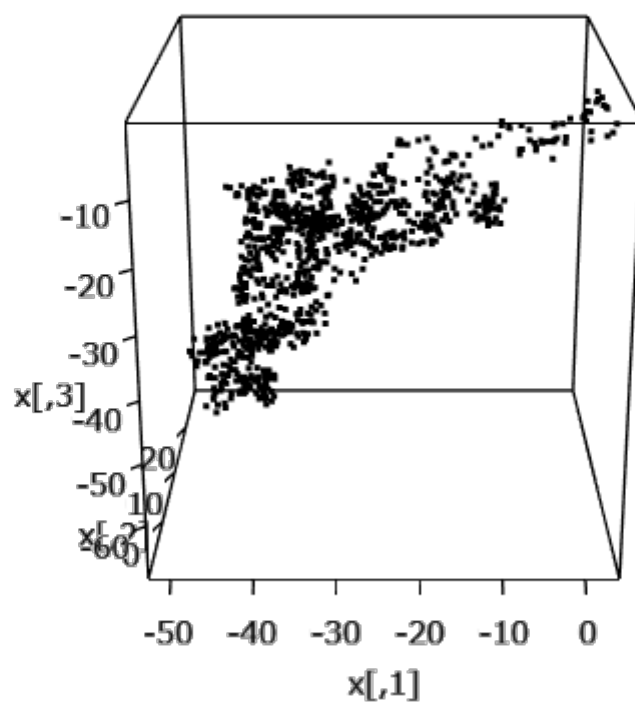
次のようにして生成される、多変量の観察データ x があったとする。

```
d <- 3
n <- 1000
x <- matrix(rnorm(d*n), ncol=d)
x <- apply(x, 2, cumsum)

plot(as.data.frame(x), pch=20, cex=0.1)
```



```
plot3d(x)
```



9.2 Exercise 4-1

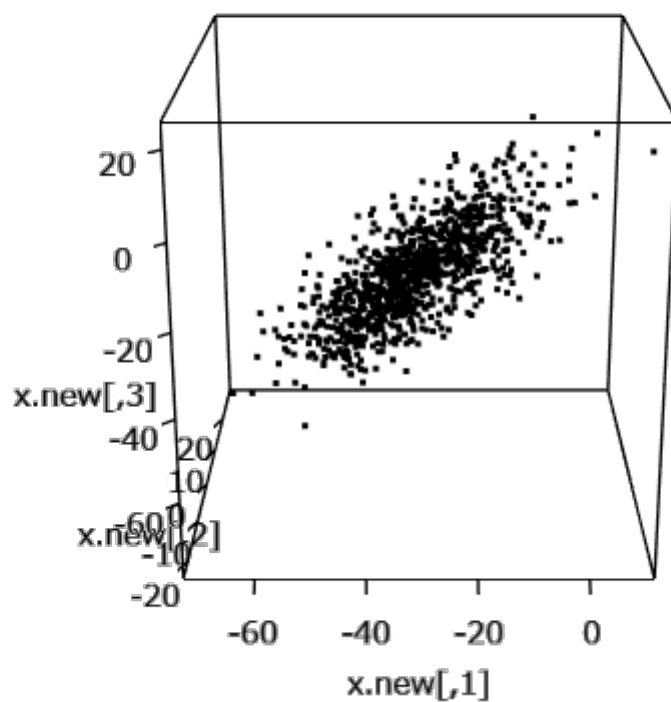
xの各変数の標本平均を求めよ。xの分散共分散行列を求めよ。

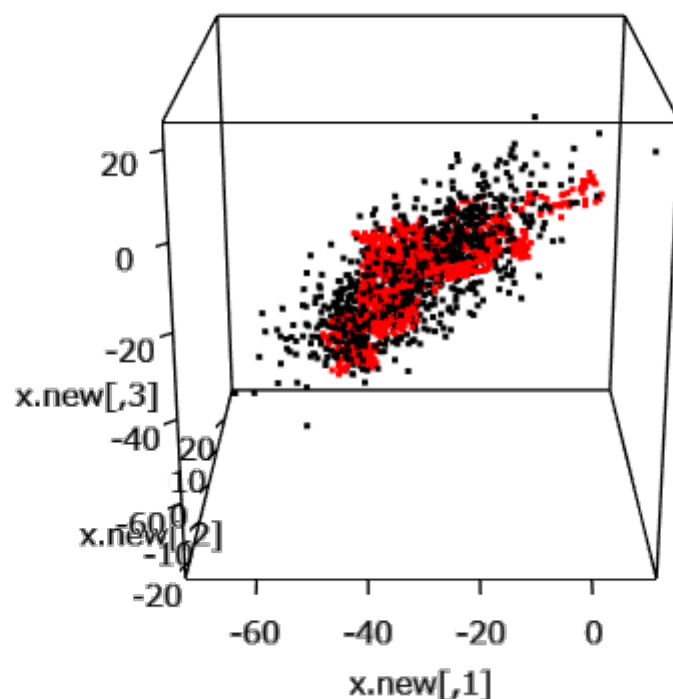
```
## [1] -31.912746  4.501103 -27.623215
```

```
##           [,1]      [,2]      [,3]
## [1,] 129.75955 -57.61427 140.0478
## [2,] -57.61427  84.32553 -131.9892
## [3,] 140.04775 -131.98923 279.8988
```

9.3 Exercise 4-2

4-1 で求めた平均値ベクトルと分散共分散行列とに基づく、多変量正規乱数を発生し、ペアワイズにプロットせよ。





9.4 Exercise 4-3

回転行列を用いて、 x を回転し(軸を取り直し)、その標本平均と分散共分散行列とを求め、4-1のそれと比較せよ。

各変数の平均も分散も異なることを確かめよ。

```
## [1] "mean x"
```

```
## [1] -31.912746  4.501103 -27.623215
```

```
## [1] "mean x rotated"
```

```
## [1]  14.921155 -39.554673  -3.809494
```

```
## [1] "cov x"
```

```
##           [,1]      [,2]      [,3]
## [1,] 129.75955 -57.61427 140.0478
## [2,] -57.61427  84.32553 -131.9892
## [3,] 140.04775 -131.98923 279.8988
```

```
## [1] "cov x rotated"
```

```
##      [,1]      [,2]      [,3]
## [1,] 47.06362 16.88131 21.41078
## [2,] 16.88131 401.27447 89.98056
## [3,] 21.41078 89.98056 45.64576
```

9.5 Exercise 4-4

分散は異なるが、分散の和は変わっていないことを、分散共分散行列のトレースを計算することで確かめよ。
同様に、分散共分散行列の固有値の和としてトレースを計算することでも確かめよ。

```
## [1] 493.9838
```

```
## [1] 493.9838
```

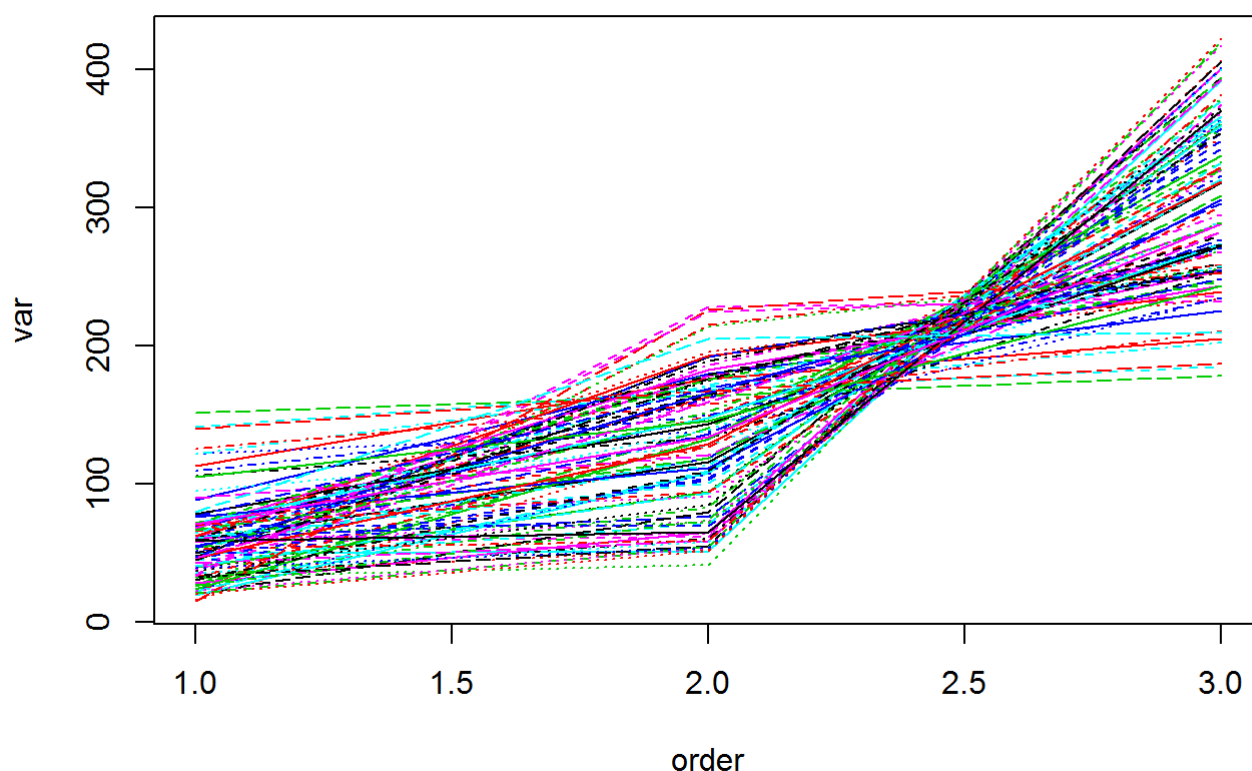
```
## [1] 493.9838
```

```
## [1] 493.9838
```

9.6 Exercise 4-5

データセットxの軸の取り方を回転することによって変えても、各軸の分散の和は変わらない。分散の和は変わらないが、分散の総和の分配具合は変わる。

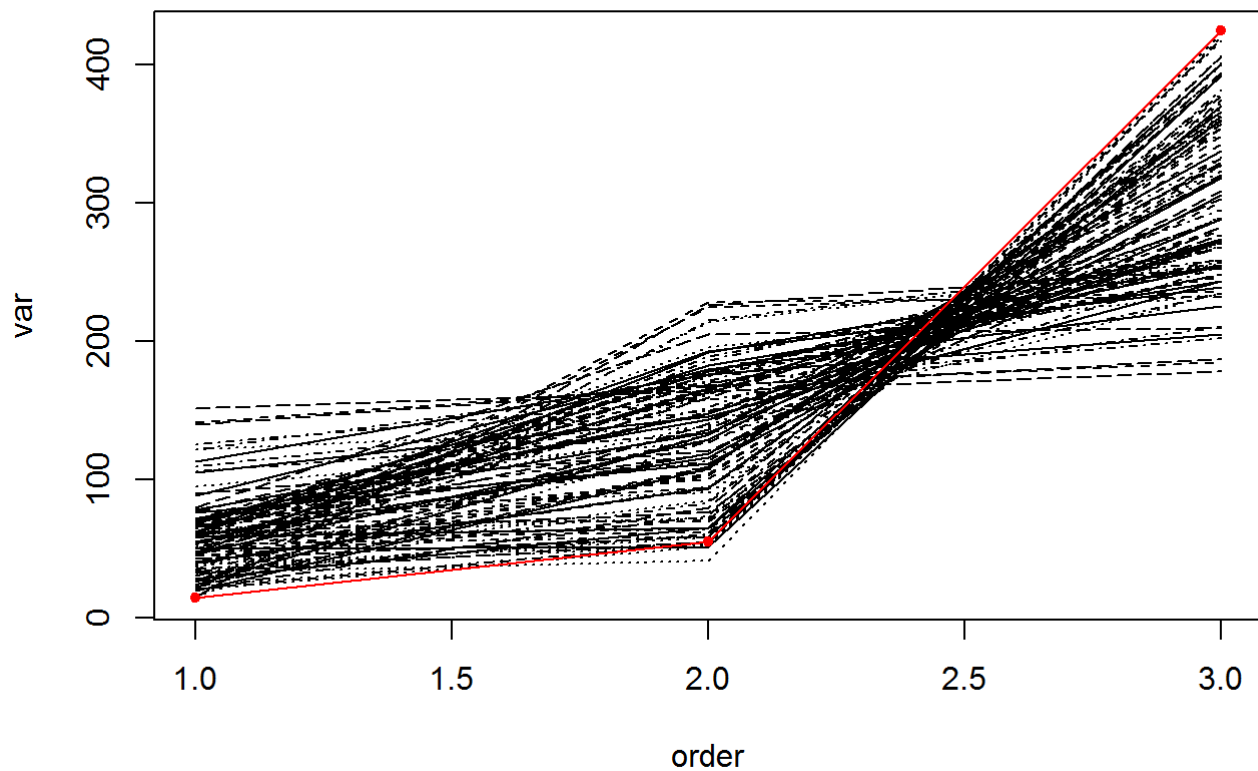
そのことを、多数の回転行列をランダムに発生させ、データセットを回転した上で各軸方向の分散を計算し、その分散の値を大小順序でソートしてプロットし、分配具合が変わることを具体的に確認せよ。



9.7 Exercise 4-6

x の分散共分散行列の固有値を求めることで、分散共分散行列を対角化するような回転をしたときの、各軸の分散を求めることができる。

その分散を4-4のプロットに重ねてプロットせよ。



9.8 Exercise 4-7

4-4,4-5と同様のプロットを変数の数を増やして実施せよ。

```
d <- 10
n <- 1000
x <- matrix(rnorm(d*n), ncol=d)
x <- apply(x, 2, cumsum)

plot(as.data.frame(x[, 1:5]), pch=20, cex=0.1)
```

