행렬 Matrix Algebra

1

행렬(변수와 개체 관측치)대수는 컴퓨터 연산의 기본이다.

연립방정식

정의 위키피디아

연립 방정식(simultaneous equation) 2개 이상의 미지수를 포함하는 2개 이상의 방정식의 조합을 의미한다.

연립 방정식은 일반적으로 대입법과 가감법을 이용해서 특정 문자들을 소거함으로써 식을 간편화해서 해를 구한다.

미지수의 개수가 n개이고, 최고 차수가 m일 때, 그 연립 방정식을 연립 n원 m차 방정식이라고 한다.

연립방정식 해 구하기

1) $\begin{matrix} 3x + 2y = 3 \\ x - y = 1 \end{matrix}$ (한 개의 해) - 가감법(substraction method)

2) $\begin{matrix} 2x + y = 8 \\ 4x + 2y = 16 \end{matrix}$ (해 많음 - 부정) 조건 $\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$

3) $\begin{matrix} x - y = 3 \\ 2x - 2y = 6 \end{matrix}$ (해 없음 - 불능) 조건 $\frac{a}{a'} = \frac{b}{b'} \neq \frac{c}{c'}$

(가감법)

1) $\begin{matrix} 2x - 2y - z = 5 \\ x + y - 2z = 1 \\ x - z = 4 \end{matrix}$

2) $\begin{matrix} x + y + z = 1 \\ 2x + y + 2z = 4 \\ 4x + 3y + 4z = 7 \end{matrix}$

행렬대수 matrix algebra

정의

$$A_{n\times p} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

- 간편식 $\left\{A_{ij}\right\}$ - 첨자 i는 행, j는 열을 나타냄
- 행 차수 n, 열 차수 p인 행렬, 차수 (nXp)인 행렬 - 기호 대문자 A, B, ...
- 열의 차수가 1인 행렬을 열 벡터(column vector), 행의 차수가 1인 행렬을 행 벡터(row vector) - 일반적으로 벡터라 함은 열 벡터를 의미하고 소문자와 아래 바를 사용 - 기호 $\underline{a}$, $\underline{b}$, ...
- 행과 열의 차수 모두가 1인 경우 행렬은 scalar (스칼라)이다. 즉 -2, 3 등의 실수는 행렬에서는 스칼라라 한다. - 기호 a, b, ...

예제

$$A_{3\times 3} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ 8 & 9 & 10 \end{bmatrix}, B_{3\times 2} = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ -3 & 4 \end{bmatrix}, \underline{a}_{1\times 2} = \begin{bmatrix} 2 & 3 \end{bmatrix}, \underline{b}_{3\times 1} = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$$

특수한 행렬

정방 sqaure 행렬

행과 열의 차수가 동일한 행렬 - S_n

대각 diagonal 행렬

- 정방행렬 중 대각원소를 제외한 모든 원소가 0인 경우 - D_n
- 대각원소의 합을 trace라 한다. $tr(D) = \sum_{i}^{n} d_{ii}$

항등 identiry 행렬

- 대각행렬 중 대각원소를 모두 1인 경우 - I_n
- 항등행렬의 trace는 n이고 숫자의 1과 같은 역할을 한다.

전치 Transpose

행렬 $A_{n\times p}$ 에서 행과 열을 바꾸는 것을 전치라 하고 A'(혹은 A^T)로 나타내며 차수는 ($p \times n$).

$$A' = \begin{bmatrix} 1 & 4 & 8 \\ 2 & 5 & 9 \\ 3 & 7 & 10 \end{bmatrix}, \underline{b}' = [5\ 1\ 4]$$

대칭 symetric 행렬

원 행렬과 대칭 행렬이 동일한 행렬 - $A' = A$ (예) 항등행렬

```
> B=matrix(c(1,-1,2,3,-3,4),byrow=T,nrow=3,ncol=2)
> B; t(B)
     [,1] [,2]
[1,]    1   -1
[2,]    2    3
[3,]   -3    4
     [,1] [,2] [,3]
[1,]    1    2   -3
[2,]   -1    3    4
```

```
> A=matrix(c(1,2,3,4,5,7,8,9,10),byrow=T,nrow=3,ncol=3)
> A; t(A)
     [,1] [,2] [,3]
[1,]    1    2    3
[2,]    4    5    7
[3,]    8    9   10
     [,1] [,2] [,3]
[1,]    1    4    8
[2,]    2    5    9
[3,]    3    7   10
```

행렬의 합(차) Addition / Subtraction

- 대응하는 원소의 합/차로 계산된다.
- 합(차) 계산이 가능하려면 두 행렬의 차수는 동일해야 한다. - 그러므로 예제에서 $A \pm B$는 성립하지 않는다.

$$C = \begin{bmatrix} 1 & 2 \\ 3 & 5 \\ -4 & -6 \end{bmatrix} \Rightarrow B + C = \begin{bmatrix} 2 & 1 \\ 5 & 8 \\ -7 & -2 \end{bmatrix}$$

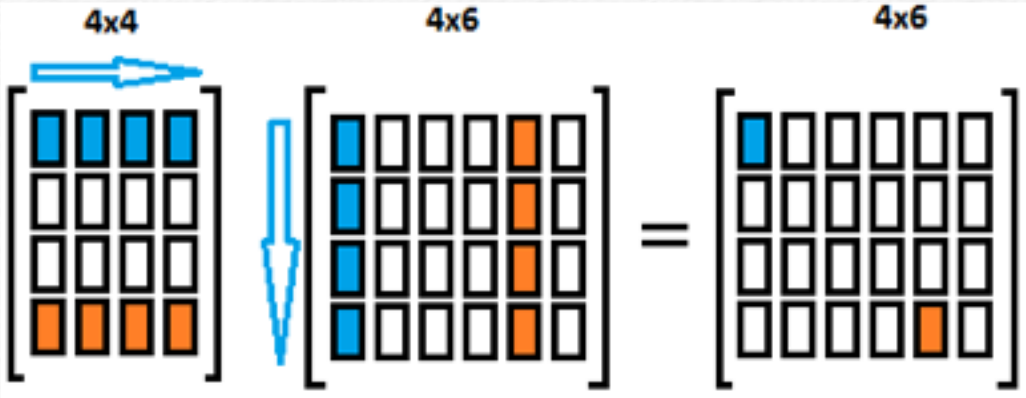
행렬의 곱 Multiplication

행렬에서 곱의 conformable(적합) 조건은 앞 행렬(벡터)의 열의 차수와 뒤 행렬 행의 차수가 동일해야 한다. 곱의 결과는 (앞 행렬의 행 차수)x(뒤 행렬의 열의 차수)인 행렬(벡터, 스칼라)이다.

- 행렬A[i번째 행]x행렬B[j번째 열]=(AB)행렬[(i, j) 원소]

$$A_{n\times p} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix},\ B_{p\times m} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pm} \end{bmatrix}$$

$$AB_{kj} = \sum_{i}^{p} a_{ki} b_{ij}$$



(예제) AB는 계산 적합하나 AB'는 계산 적합하지 않다.

```
> A%*%B
     [,1] [,2]
[1,]   -4   17
[2,]   -7   39
[3,]   -4   59
> A%*%t(B)
Error in A %*% t(B) : 적합한 인자들이 아닙니다
```

(성질)

1) $AB \neq BA$ (일반적으로)

2) $(AB)' = B'A'$

3) $A'A$은 대칭행렬이다.

역행렬

행렬식 Determinant

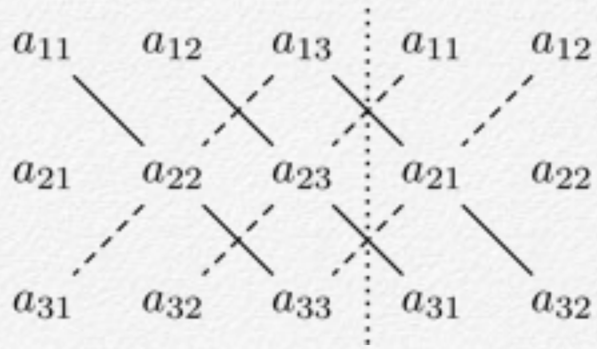
정방행렬 A의 행렬식은 다음과 같이 계산되며, 결과는 스칼라

기호 : $|A|$ 혹은 $\det(A)$

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a \cdot d - c \cdot b$$

$$\det\begin{bmatrix} 8 & 3 \\ 4 & 2 \end{bmatrix} = 8 \cdot 2 - 4 \cdot 3 = 16 - 12 = 4$$

(2×2 행렬)



(3×3 행렬)

```
> det(A)
[1] 7
```

소인자 Minor

$$M_{23} = \det\begin{bmatrix} 1 & 4 & \square \\ \square & \square & \square \\ -1 & 9 & \square \end{bmatrix} = \det\begin{bmatrix} 1 & 4 \\ -1 & 9 \end{bmatrix} = (9 - (-4)) = 13$$

$$|A| = \sum_{i=1}^{n} a_{ij} (-1)^{i+j} |M_{ij}| = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} |M_{ij}|$$

차수가 3일 경우: $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ 8 & 9 & 10 \end{bmatrix}$ 의 행렬식은

$$|A| = 1*(-1)^{1+1}\begin{vmatrix} 5 & 7 \\ 9 & 10 \end{vmatrix} + 2*(-1)^{1+2}\begin{vmatrix} 4 & 7 \\ 8 & 10 \end{vmatrix} + 3*(-1)^{1+3}\begin{vmatrix} 4 & 5 \\ 8 & 9 \end{vmatrix} = 7 \text{ (1번째 행 이용)}$$

$$|A| = 4*(-1)^{2+1}\begin{vmatrix} 2 & 3 \\ 9 & 10 \end{vmatrix} + 5*(-1)^{2+2}\begin{vmatrix} 1 & 3 \\ 8 & 10 \end{vmatrix} + 7*(-1)^{2+3}\begin{vmatrix} 1 & 2 \\ 8 & 9 \end{vmatrix} = 7 \text{ (2번째 행 이용)}$$

(행렬식 성질)

- $|A'| = |A|$, $|AB| = |BA|$, $|AB| = |A||B|$
- 행렬 A의 두 행이 같으면 행렬식은 0이다.
- 한 행(열)의 상수를 곱하여 다른 행에 더해도 행렬식 값은 변하지 않는다.
- 한 행(열)을 다른 행들의 선형 결합으로 표현할 수 있으면 행렬식 0이다.

역행렬

$AB = I$ 만족하는 행렬 B를 행렬 A의 역행렬이라 하고 A^{-1}로 표기한다.

$$A^{-1} = \frac{1}{|A|} adjA = \frac{1}{|A|}$$

(역행렬 성질)

- 역행렬은 unique하다.
- $|A^{-1}| = 1/|A|$, $(A^{-1})^{-1} = A$, $(A')^{-1} = (A^{-1})'$, $(AB)^{-1} = B^{-1}A^{-1}$

계수 Rank

(정의) 행렬의 계수는 행렬에서 선형 독립인 행(그리고 열)의 수, $rank(A)$

정의(LIN: linearly independent vector)

$a_1\underline{x}_1 + a_2\underline{x}_2 + \ldots + a_p\underline{x}_p = \underline{0}$ 가 성립하려면 모든 $a_i = 0$ 일 때만 만족한다면 열벡터 ($\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_p$) 는 선형 독립(linearly independent) 벡터라 하고, 0이 아닌 a_i에 대해서 만족한다면 선형 종속(linearly dependent)인 벡터

상호 종속인 벡터는 하나의 벡터를 다른 벡터들의 선형 결합으로 표시할 수 있다는 것을 의미한다.

정의(full rank)

(nxn)정방 행렬에서 선형 독립인 행(열)의 개수( )가 행렬의 차수 n와 같다면 이 행렬은 full-rank 행렬이라 한다. 즉 $rank(A_{n\times n}) = n$ 이면 full-rank이다.

▪ 역행렬이 존재한다. ▪ full-rank이다. rank(A)=n ▪ A는 non-singular이다. ▪ $|A| \neq 0$ ▪ Ax=b 의 해가 존재한다.	▪ 역행렬이 존재하지 않는다 ▪ full-rank아니다. rank(A)<n ▪ A는 singular이다. ▪ $|A|=0$ ▪ Ax=b 의 해가 존재하지 않는다.

```
> rankMatrix(A)
[1] 3
```

```
> solve(A)
           [,1] [,2]       [,3]
[1,] -1.8571429    1 -0.1428571
[2,]  2.2857143   -2  0.7142857
[3,] -0.5714286    1 -0.4285714
```

```
> A%*%solve(A)
              [,1] [,2]
[1,]  1.000000e+00    0
[2,] -8.881784e-16    1
```

연립방정식 풀이

$2x - 2y - z = 5$
$x + y - 2z = 1$ => $A\underline{x} = \underline{b}$, $\underline{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ => $\underline{x} = A^{-1}\underline{b}$
$x - z = 4$

```
> A=matrix(c(2,-2,-1,1,1,-2,1,0,-1),byrow=T,nrow=3,ncol=3)
> b=matrix(c(5,1,4),byrow=T,nrow=3,ncol=1)
> rankMatrix(A)
[1] 3
```

```
> solve(A)%*%b
     [,1]
[1,]   13
[2,]    6
[3,]    9
```

$x + y + z = 1$
$2x + y + 2z = 4$
$4x + 3y + 4z = 7$

```
> A=matrix(c(1,1,1,2,1,2,4,3,4),byrow=T,nrow=3,ncol=3)
> b=matrix(c(1,4,7),byrow=T,nrow=3,ncol=1)
> rankMatrix(A)
[1] 2
```

```
> solve(A)
Error in solve.default(A) :
  Lapack routine dgesv: system is exactly singular: U[3,3] = 0
```

행렬의 미분

상수 벡터 $\underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}$, 확률변수 벡터 $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$ 라 하면

(1) $\dfrac{\partial}{\partial \underline{x}}(\underline{a}'\underline{x}) = \underline{a}$

(2) $\dfrac{\partial}{\partial \underline{x}}(\underline{x}'\underline{a}) = \underline{a}$

(3) $\dfrac{\partial}{\partial \underline{x}}(\underline{x}'A\underline{x}) = A\underline{x} + A'\underline{x}$ (A는 정방 행렬)

=> 만약 A가 대칭 행렬이면 $\dfrac{\partial}{\partial \underline{x}}(\underline{x}'A\underline{x}) = 2A\underline{x}$

(1) $E(\underline{x}) = \underline{\mu} = \begin{bmatrix} \mu_1 \\ \cdots \\ \mu_p \end{bmatrix}$

(2) $V(\underline{x}) = \Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1p}^2 \\ \sigma_{11}^2 & \sigma_{22}^2 & \cdots & \sigma_{2p}^2 \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{p1}^2 & \sigma_{p2}^2 & \cdots & \sigma_{pp}^2 \end{bmatrix}$

일변량 확률변수 기대값 행렬 표현

확률변수 X의 n개 관측치 벡터를 $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$이라 하자.

- $E(X) = \dfrac{\sum x_i}{n}$ => (행렬 표현) $E(X) = \dfrac{1}{n}(\underline{1}_n{}'\underline{x})$, where $\underline{1}_n$은 차수가 n인 1벡터
- $V(X) = E(X^2) - E(X)^2$ => (행렬 표현) => $E(X^2) = (\underline{x}'\underline{x})/n$

확률벡터의 기대값

p개 확률변수 벡터의 원소 $x_i \sim (E(X_i) = \mu_i, V(X_i) = \sigma_{ii})$이고 확률변수 (x_i, x_j)의 공분산을 $COV(X_i, X_j) = \sigma_{ij}$ 이라 정의하자.

$$E(\underline{x}) = E\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \underline{\mu} \quad V(\underline{x}) = V\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \sigma_{2p} \\ \sigma_{p1} & \sigma_{p2} & \sigma_{pp} \end{pmatrix} = \Sigma$$

$\underline{a}'\underline{x}$ 기대값

$E(\underline{a}'\underline{x}) = \underline{a}'\underline{\mu}$ -> 결과는 스칼라 ---- (1)

$V(\underline{a}'\underline{x}) = \underline{a}\Sigma\underline{a}'$ -> 결과는 (pXp) 행렬 ---- (2)

```
> bvn
                 X            Y
         [,1]        [,2]
[1,]  8.213868  0.79394993
[2,]  9.268818  1.15076144
[3,]  9.671887  1.70516050
[4,] 14.014720  4.45184985
```

```
Sigma <- matrix(c(5,3,3,2),2,2) #covariance matrix
library(MASS) #mvnorm function package
bvn=as.matrix(mvrnorm(n = 100, c(10, 2), Sigma))
mean(bvn[,1]); mean(bvn[,2]) #mean 10, 2
var(bvn[,1]); var(bvn[,2])  #var 5, 2
cov(bvn[,1], bvn[,2])   #covariance = 3
```

```
> Sigma <- matrix(c(5,3,3,2),2,2) #covariance matrix
> library(MASS) #mvnorm function package
> bvn=as.matrix(mvrnorm(n = 100, c(10, 2), Sigma))
> mean(bvn[,1]); mean(bvn[,2]) #mean 10, 2
[1] 10.1535
[1] 2.015747
> var(bvn[,1]); var(bvn[,2])  #var 5, 2
[1] 6.086475
[1] 2.110198
> cov(bvn[,1], bvn[,2])   #covariance = 3
[1] 3.402672
```

평균, 분산, 공분산은 이론적인 값이므로 이론 값과 다르다.

*) (3X+7Y)의 평균과 분산을 구하자. => $\underline{a} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$

```
a=matrix(c(3,7))
mu=matrix(c(mean(bvn[,1]),mean(bvn[,2])))
sigma=matrix(c(var(bvn[,1]),cov(bvn[,1], bvn[,2]),
             cov(bvn[,1], bvn[,2]),var(bvn[,2])),
           ncol=2,nrow=2)
t(a)%*%mu
t(a)%*%sigma%*%a
```

```
> t(a)%*%mu
         [,1]
[1,] 44.57073
> t(a)%*%sigma%*%a
         [,1]
[1,] 301.0902
```