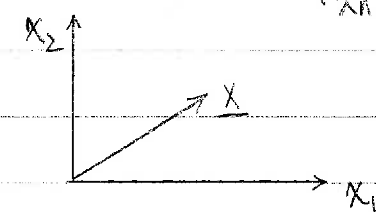
Chapter 4 The Regression Problem in Matrix Notation

4.1 Vecters

Def. 4.1: An n-tuple of numbers $(x_1, \cdots, x_n)$ arranged in a column is called an n-dimensional vector $\underline{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Geometrical interpretation for $n=2$



Two important vectors: $\underline{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, $\underline{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

- Operations

(1) Scalar multiplication: $c\underline{X} = c\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ \vdots \\ cx_n \end{pmatrix}$.

(2) Vector addition: $\underline{X} + \underline{Y} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix}$.

(3) Vector subtraction: $\underline{X} - \underline{Y} = \underline{X} + (-\underline{Y}) = \begin{pmatrix} x_1-y_1 \\ \vdots \\ x_n-y_n \end{pmatrix}$.

(4) Transpose: $\underline{X}' = (x_1, \cdots, x_n)$.

(5) Inner product of $\underline{X}$ and $\underline{Y}$: $\underline{X}'\underline{Y} = (x_1, \cdots, x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + \cdots + x_ny_n$.

4.2 Matrices

Def. 4.2: An <u>$n \times m$ matrix</u>, generally denoted by a boldface uppercase letter, is a rectangular array of numbers set out in n rows and m columns. For example,

$$\underset{n\times m}{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} = \{a_{ij}\}.$$

$n \times m$ is called the dimension of the matrix, n is the row dimension and m is the column dimension. The number in the i^{th} row, j^{th} column of a matrix A is referred to as the $(i,j)^{th}$ element of A, often a_{ij}.

If we consider each row (column) of a matrix as a vector, it is a row (column) vector.

Note: A vector is a special matrix, a matrix having only one column.

<u>Ex.</u> $A = \begin{pmatrix} 1 & 2 \\ 7 & 3 \\ 9 & 16 \end{pmatrix}$. Dimension of A: 3×2. $a_{31} = 9$.

- Operations

(1) Scalar multiplication

$$\underset{(n\times m)}{cA} = c\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1m} \\ ca_{21} & ca_{22} & \cdots & ca_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nm} \end{pmatrix} = \{ca_{ij}\}.$$

Ex. $A = \begin{pmatrix} 3 & 0 \\ 7 & 5 \\ -2 & 10 \\ 1 & 2 \end{pmatrix}$, $c = -2$, $cA = (-2)\begin{pmatrix} 3 & 0 \\ 7 & 5 \\ -2 & 10 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -6 & 0 \\ -14 & -10 \\ 4 & -20 \\ -2 & -4 \end{pmatrix}$.

(2) Matrix addition

Let the matrices A and B both be of dimension $n\times m$. Then

$$\underset{(n\times m)}{A} + \underset{(n\times m)}{B} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \cdots & \cdots & \cdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1m}+b_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1}+b_{n1} & \cdots & a_{nm}+b_{nm} \end{pmatrix} = \{a_{ij}+b_{ij}\}.$$

Note: To add matrices, # rows & columns must match.

Ex. $A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \\ 5 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & -2 \\ 1 & 1 \\ 0 & 3 \end{pmatrix}$. Then $C = A+B = \begin{pmatrix} 3 & 0 \\ 2 & 5 \\ 5 & 3 \end{pmatrix}$.

(3) Matrix subtraction

$$\underset{(n\times m)}{A} - \underset{(n\times m)}{B} = A + (-1)B = \{a_{ij} - b_{ij}\} = \begin{pmatrix} a_{11}-b_{11} & \cdots & a_{1m}-b_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1}-b_{n1} & \cdots & a_{nm}-b_{nm} \end{pmatrix}.$$

(4) Transpose: The transpose of a matrix $\underset{(n\times m)}{A}$, denoted by A', is obtained from A by interchanging the rows and columns, that is,

$$\underset{(m\times n)}{A'} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} = \{a_{ji}\}.$$

<u>Ex</u>: $B = \begin{pmatrix} 1 & 2 & 9 \\ 3 & 4 & 1 \end{pmatrix}$, $B' = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 9 & 1 \end{pmatrix}$.

(5) Matrix multiplication: The product AB of an $n\times m$ matrix $A = \{a_{ij}\}$ and an $m\times k$ matrix $B = \{b_{ij}\}$ is the $n\times k$ matrix C whose elements are

$$c_{ij} = \sum_{\ell=1}^{m} a_{i\ell} b_{\ell j}, \quad i = 1, \cdots, n, \quad j = 1, 2, \cdots, k.$$

$$\underset{(n\times m)}{A} \; \underset{(m\times k)}{B} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ \boxed{a_{i1} \; \cdots \; a_{im}} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1k} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mj} & \cdots & b_{mk} \end{pmatrix}$$

a_i'

b_j

must match

$$= \{ \underset{(n\times k)}{a_i' b_j} \} = \{ a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{im} b_{mj} \}.$$

Note: (i) The product AB is defined only when the number of columns of A is equal to # of rows of B.

(ii) $AB \neq BA$ in general, may not even be defined.

Ex: $A = (1\ 6\ 8)$ $B = \begin{pmatrix} 0 & 2 \\ 1 & 1 \\ -1 & 3 \end{pmatrix}$.

$AB = (1\times 0 + 6\times 1 + 8\times(-1),\ 1\times 2 + 6\times 1 + 8\times 3) = (-2, 32)$

BA is not defined.

• Some basic concepts

(i) $\underset{(n\times m)}{A}$ is called a square matrix if $n=m$, that is, A has the same number of rows and columns.

Diagonal elements of a square matrix $A_{n\times n}$: elements a_{ii}.

Ex: $A = \begin{pmatrix} 1 & 7 & 9 & 4 \\ 3 & 2 & 1 & 0 \\ 0 & 7 & 5 & 0 \\ 1 & 1 & 2 & 1 \end{pmatrix}$.

Important square matrices:

(i) $n \times n$ diagonal matrices: $\text{diag}(d_{11}, \cdots, d_{nn}) = \begin{pmatrix} d_{11} & & 0 \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}$.

(ii) The $n \times n$ identity matrix: $\underset{(n\times n)}{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ - & - & - & - & \cdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$, the $n \times n$ square matrix with ones on the main diagonal and zeros elsewhere.

(2) A square matrix $\underset{(n\times n)}{A}$ is said to be <u>symmetric</u> if $A' = A$, that is, $a_{ij} = a_{ji}$.

<u>Ex</u>: $A = \begin{pmatrix} 3 & 1 & 7 \\ 2 & 4 & 4 \\ 1 & 5 & 9 \end{pmatrix}$. $A' = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 4 & 5 \\ 7 & 4 & 9 \end{pmatrix} \neq A$. So A is not symmetric.

$B = \begin{pmatrix} 3 & -3 & 8 \\ -3 & 4 & 1 \\ 8 & 1 & 9 \end{pmatrix}$. $B' = \begin{pmatrix} 3 & -3 & 8 \\ -3 & 4 & 1 \\ 8 & 1 & 9 \end{pmatrix} = B$. So B is symmetric.

(3) Let A be a square matrix. If there exists a matrix B such that

$$\underset{(n\times n)}{A}\ \underset{(n\times n)}{B} = \underset{(n\times n)}{B}\ \underset{(n\times n)}{A} = \underset{(n\times n)}{I},$$

then B is called the <u>inverse of A</u> and is denoted by A^{-1}.

(4) The <u>determinant</u> of a square matrix $\underset{(n\times n)}{A} = \{a_{ij}\}$, denoted

by $|A|$, is the scalar

$$|A| = a_{11} \quad \text{if } n=1$$

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad \text{if } n=2$$

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{32}a_{23}a_{11} \quad \text{if } n=3$$

$$|A| = \sum_{j=1}^{n} a_{ij} |A_{ij}| (-1)^{i+j} \quad \text{if } n>1$$

where A_{ij} is the $(n-1)\times(n-1)$ matrix obtained by deleting the i^{th} row and j^{th} column of A, called <u>cofactor</u> of a_{ij}.

Note: In general, use computer to find determinant for $n \geq 3$.

<u>Ex</u>: $A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$. $|A| = 3\times4 - 1\times2 = 10$.

(5) A square matrix $\underset{(n\times n)}{A}$ is <u>nonsingular</u> if $|A| \neq 0$. Otherwise, it is called <u>singular</u>.

(6) A square matrix $\underset{(n\times n)}{A}$ is said to be <u>idempotent</u> if $A^2 = A$.

- Some important results

1. Results about matrix operations

Suppose that all operations in the following equalities are defined.

(a) $(A+B)+C = A+(B+C)$ (Associative law)
(b) $A+B = B+A$ (Commutative law)
(c) $a(A+B) = aA+aB$ (Distributive law)
(d) $(a+b)A = aA+bA$
(e) $(ab)A = a(bA)$
(f) $a(AB) = (aA)B$
(g) $A(BC) = (AB)C$
(h) $A(B+C) = AB+AC$
(i) $(B+C)A = BA+CA$
(j) $(A+B)' = A'+B'$
(k) $(aA)' = aA'$
(l) $(AB)' = B'A'$

Ex: $A = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. $AB = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}$, $(AB)' = (-1 \ 3 \ 0)$.

$A' = \begin{pmatrix} 1 & 1 & 1 \\ 3 & -1 & 2 \end{pmatrix}$, $B' = (2 \ -1)$. $B'A' = (-1 \ 3 \ 0)$. ✓

(m) $(A^{-1})' = (A')^{-1}$
(n) $(AB)^{-1} = B^{-1}A^{-1}$. Pf. $AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I$.

2. Results about determinant

Let A and B be $n \times n$ square matrices and c be a scalar.

(a) $|A'| = |A|$

(b) $|A^{-1}| = 1/|A|$

(c) $|cA| = c^n |A|$

(d) $|AB| = |A||B|$

3. Results about inverse matrix

(a) If $\underset{(n\times n)}{A}$ is nonsingular, that is $|A| \neq 0$, then A has a unique inverse A^{-1}, whose $(i,j)^{\text{th}}$ entry is $\dfrac{(-1)^{i+j}|A_{ji}|}{|A|}$, i.e.,

$$A^{-1} = \frac{1}{|A|}\begin{pmatrix} (-1)^{1+1}|A_{11}| & (-1)^{2+1}|A_{21}| & \cdots & (-1)^{n+1}|A_{n1}| \\ (-1)^{1+2}|A_{12}| & (-1)^{2+2}|A_{22}| & \cdots & (-1)^{n+2}|A_{n2}| \\ \cdots & \cdots & \cdots & \cdots \\ (-1)^{1+n}|A_{1n}| & (-1)^{2+n}|A_{2n}| & \cdots & (-1)^{n+m}|A_{nn}| \end{pmatrix},$$

where A_{ji} is the cofactor of a_{ji}.

If $n=1$, $A = (a_{11})$, $A^{-1} = (a_{11}^{-1})$.

If $n=2$, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $A^{-1} = \dfrac{1}{|A|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

If $n=3$, $A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$A^{-1}=\frac{1}{|A|}\begin{pmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & A_{33} \end{vmatrix} & -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix}$$

Note. Generally use computer to find inverses for $n \geq 3$.

Ex. $A=\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$. $|A| = 3\times 4 - 2\times 1 = 10$.

Since $|A| \neq 0$, A^{-1} exists.

$$A^{-1}=\frac{1}{10}\begin{pmatrix} 4 & -2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 0.4 & -0.2 \\ -0.1 & 0.3 \end{pmatrix}$$

Verify $AA^{-1}=\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 0.4 & -0.2 \\ -0.1 & 0.3 \end{pmatrix} = \begin{pmatrix} 1.2-0.2 & -0.6+0.6 \\ 0.4-0.4 & -0.2+1.2 \end{pmatrix}$

$$=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underset{(2\times 2)}{I}$$

Check $A^{-1}A$ by yourself.

(b) For a nonsingular diagonal matrix $\underset{(n\times n)}{A} = \text{diag}(a_{11}, \cdots, a_{nn})$

$A^{-1} = \text{diag}(a_{11}^{-1}, \ldots, a_{nn}^{-1})$.

4. Results about transpose

For any $n\times m$ matrix A,

(i) $A'A$ is a symmetric $m\times m$ matrix.

Pf: Clearly $\underset{(m\times n)}{A'}\ \underset{(n\times m)}{A}$ is an $m\times m$ matrix.

Symmetric? $(A'A)' = A'(A')' = A'A$. ✓

(ii) AA' is a symmetric $n\times n$ matrix.

5. Partial derivative with respect to a matrix.

Def. 4.3: Let $\underset{(n\times m)}{X} = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$.

$f(X) = f(x_{11}, \cdots, x_{1m}, x_{21}, \cdots, x_{2m}, \cdots, x_{n1}, \cdots, x_{nm})$

$\frac{\partial f}{\partial X}$ is defined as

$$\frac{\partial f}{\partial X} = \begin{pmatrix} \frac{\partial f}{\partial x_{11}} & \cdots & \frac{\partial f}{\partial x_{1m}} \\ \cdots & \cdots & \cdots \\ \frac{\partial f}{\partial x_{n1}} & \cdots & \frac{\partial f}{\partial x_{nm}} \end{pmatrix}.$$

Results: Let $\underset{(p\times 1)}{\underline{c}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{pmatrix}$, $\underset{(p\times 1)}{\underline{X}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$, $\underset{(p\times p)}{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ - & - & - & - \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{pmatrix}$. Then

(1) $\dfrac{\partial(\underline{c}'\underline{X})}{\partial \underline{X}} = \dfrac{\partial(\underline{X}'\underline{c})}{\partial \underline{X}} = \underline{c}$.

(2) $\dfrac{\partial(\underline{X}'A\underline{X})}{\partial \underline{X}} = (A+A')\underline{X}$.

Note: If A is symmetric, $\dfrac{\partial(\underline{X}'A\underline{X})}{\partial \underline{X}} = 2A\underline{X}$.