Determinants for 3×3 Matrices

Nobuyuki TOSE

December 06, 2016

Definition

We are given three 3-dimensional column vectors

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \ \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \ \vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

and a 3×3 matrix

$$A = (\vec{a}\ \vec{b}\ \vec{c})$$

We define the determinant of A by

$$\begin{aligned} \det(A) = |A| &= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \\ &= a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \end{aligned}$$

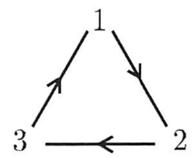
Sarrus Rule

We expand the three determinants of 2×2 matrices to get

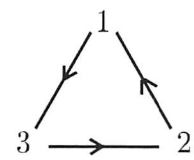
$$|A| = a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1$$

The signs in the head of $a_i b_j c_k$ are determined by the following rule.

$$\varepsilon(i, j, k) = \begin{cases} 1 & ((i, j, k) \text{ has positive orientation}) \\ -1 & ((i, j, k) \text{ has negative orientation}) \end{cases}$$



Positive Orientation



Negative Orientation

Sarrus Rule(2)

$$\det(A) = \sum_{i \neq j,\ j \neq k, k \neq i} \varepsilon(i\ j\ k) \cdot a_i b_j c_k \tag{1}$$

Cofactor Expansion

Laplace Expansion

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$ ← Definition

$$= -b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

$$= c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

Basic Property I

Basic Property I – Multi-linearity

$$\left|(s_1\vec{\alpha} + s_2\vec{\beta})\ \vec{b}\ \vec{c}\right| = s_1\left|\vec{\alpha}\ \vec{b}\ \vec{c}\right| + s_2\left|\vec{\beta}\ \vec{b}\ \vec{c}\right|$$

$$\left|\vec{a}\ (s_1\vec{\alpha} + s_2\vec{\beta})\ \vec{c}\right| = s_1\left|\vec{a}\ \vec{\alpha}\ \vec{c}\right| + s_2\left|\vec{a}\ \vec{\beta}\ \vec{c}\right|$$

$$\left|\vec{a}\ \vec{b}\ (s_1\vec{\alpha} + s_2\vec{\beta})\right| = s_1\left|\vec{a}\ \vec{b}\ \vec{\alpha}\right| + s_2\left|\vec{a}\ \vec{b}\ \vec{\beta}\right|$$

Basic Property I

$$c_1\vec{\alpha} + c_2\vec{\beta} = \begin{pmatrix} c_1\alpha_1 + c_2\beta_1 \\ c_1\alpha_2 + c_2\beta_2 \\ c_1\alpha_3 + c_2\beta_3 \end{pmatrix}$$

The basic Property I is based on the following theorem.

Theorem

$F: \mathbf{R}^3 \to \mathbf{R}$ is defined by

$$F(\vec{x}) = b_1x_1 + b_2x_2 + b_3x_3$$

Then we have

$$F(c_1\vec{\alpha} + c_2\vec{\beta}) = c_1F(\vec{\alpha}) + c_3F(\vec{\beta})$$

$$= b_1(c_1\alpha_1 + c_2\beta_1) + b_2(c_1\alpha_2 + c_2\beta_2) + b_3(c_1\alpha_3 + c_2\beta_3)$$

$$= c_1(b_1\alpha_1 + b_2\alpha_2 + b_3\alpha_3) + c_2(b_1\beta_1 + b_2\beta_2 + b_3\beta_3)$$

$$= c_1F(\vec{\alpha}) + c_2F(\vec{\beta})$$

Basic Properties II+III

Basic Property II – Interchanging two colums

$$|\vec{a}\ \vec{b}\ \vec{c}| = -|\vec{b}\ \vec{a}\ \vec{c}| = -|\vec{a}\ \vec{c}\ \vec{b}| = -|\vec{c}\ \vec{b}\ \vec{a}|$$

$$\begin{aligned} |\vec{a}\ \vec{b}\ \vec{c}| &= a_1\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \\ &= -a_1\begin{vmatrix} c_2 & b_2 \\ c_3 & b_3 \end{vmatrix} + a_2\begin{vmatrix} c_1 & b_1 \\ c_3 & b_3 \end{vmatrix} - a_3\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \\ &= -|\vec{a}\ \vec{c}\ \vec{b}| \end{aligned}$$

Basic Property III

$$|I_3| = 1$$

Basic Property IV

Basic Property III

$$|A| = |{}^tA| \quad i.e. \quad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

To prove this, develop the RHS:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \cdots$$

$$a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}$$

$$= \cdots \qquad \cdots = |\vec{a}\ \vec{b}\ \vec{c}|$$

$\uparrow$

Exercise.

Basic Property IR

Basic Property IR – Multi-linearity

$$\begin{vmatrix} \mathbf{a} \\ s_1\mathbf{p} + s_2\mathbf{q} \\ \mathbf{c} \end{vmatrix} = s_1 \begin{vmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{c} \end{vmatrix} + s_2 \begin{vmatrix} \mathbf{a} \\ \mathbf{q} \\ \mathbf{c} \end{vmatrix}$$

$$\begin{aligned} LHS &= |{}^t\mathbf{a}\ {}^t(s_1\mathbf{p} + s_2\mathbf{q})\ {}^t\mathbf{c}| \\ &= |{}^t\mathbf{a}\ s_1{}^t\mathbf{p} + s_2{}^t\mathbf{q}\ {}^t\mathbf{c}| \\ &= s_1|{}^t\mathbf{a}\ {}^t\mathbf{p}\ {}^t\mathbf{c}| + s_2|{}^t\mathbf{a}\ {}^t\mathbf{q}\ {}^t\mathbf{c}| = RHS \end{aligned}$$

$$= s_1 \begin{vmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{c} \end{vmatrix} + s_2 \begin{vmatrix} \mathbf{a} \\ \mathbf{q} \\ \mathbf{c} \end{vmatrix}$$

Basic Property IIR

Basic Property IIR–Interchanging two rows

$$\begin{vmatrix}\mathbf{a}\\ \mathbf{b}\\ \mathbf{c}\end{vmatrix} = -\begin{vmatrix}\mathbf{b}\\ \mathbf{a}\\ \mathbf{c}\end{vmatrix} = -\begin{vmatrix}\mathbf{a}\\ \mathbf{c}\\ \mathbf{b}\end{vmatrix} = -\begin{vmatrix}\mathbf{c}\\ \mathbf{b}\\ \mathbf{a}\end{vmatrix}$$

$= |{}^t\mathbf{a}\ {}^t\mathbf{b}\ {}^t\mathbf{c}|$

$= -|{}^t\mathbf{b}\ {}^t\mathbf{a}\ {}^t\mathbf{c}| = -\begin{vmatrix}\mathbf{b}\\ \mathbf{a}\\ \mathbf{c}\end{vmatrix}$

BP II for Colums

Basic Properties V+VR and VIR

Basic Properties V+VR

If A has two idential colums (rows):

$$|\vec{a}\ \vec{a}\ \vec{c}| = 0, \quad \begin{vmatrix}\mathbf{a}\\ \mathbf{a}\\ \mathbf{c}\end{vmatrix} = 0$$

$|\vec{a}\ \vec{a}\ \vec{c}| = -|\vec{a}\ \vec{a}\ \vec{c}|$

$\rightarrow |\vec{a}\ \vec{a}\ \vec{c}| = 0$

Basic Property VIR

Adding a multiple of a row to another:

$$\begin{vmatrix}\mathbf{a}\\ \mu\mathbf{a}+\mathbf{b}\\ \mathbf{c}\end{vmatrix} = \begin{vmatrix}\mathbf{a}\\ \mathbf{b}\\ \mathbf{c}\end{vmatrix} \tag{2}$$

$= \mu\underbrace{\begin{vmatrix}\mathbf{a}\\ \mathbf{a}\\ \mathbf{c}\end{vmatrix}}_{=0} + \begin{vmatrix}\mathbf{a}\\ \mathbf{b}\\ \mathbf{c}\end{vmatrix} = \begin{vmatrix}\mathbf{a}\\ \mathbf{b}\\ \mathbf{c}\end{vmatrix}$

example

$$
\begin{array}{r}
2 \quad 4 \quad 6 \\
-\ 2 \quad 8 \quad 18 \\
\hline
0 \quad -4 \quad -12
\end{array}
$$

$$
\begin{vmatrix} 0 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 4 & 9 \end{vmatrix} = -\begin{vmatrix} 1 & 4 & 9 \\ 2 & 4 & 6 \\ 0 & 2 & 3 \end{vmatrix} = -\begin{vmatrix} 1 & 4 & 9 \\ 0 & \boxed{-4} & \boxed{-12} \\ 0 & \boxed{2} & \boxed{3} \end{vmatrix}
$$

$$
= -\begin{vmatrix} -4 & -12 \\ 2 & 3 \end{vmatrix} = -(-4)\begin{vmatrix} 1 & 3 \\ 2 & 3 \end{vmatrix} = 4(3-6) = -12
$$

$(-4 \ \ -12) = (-4)(1 \ \ 3)$

Cramer's Rule

We study the following system of equations:

$$
\begin{cases} a_1x + b_1y + c_1z = \alpha_1 & \cdots (1) \\ a_2x + b_2y + c_2z = \alpha_2 & \cdots (2) \\ a_3x + b_3y + c_3z = \alpha_3 & \cdots (3) \end{cases}
$$

Cramer's Rule (2)

To elimeinate z, we consider $(1) \times c_2 - (2) \times c_1$:

$$
\begin{array}{rrcll}
 & a_1c_2x + b_1c_2y + c_1c_2z & = & \alpha_1 c_2 & \cdots(1)\times c_2 \\
-) & a_2c_1x + b_2c_1y + c_1c_2z & = & \alpha_2 c_1 & \cdots(2)\times c_1 \\
\hline
 & \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} x + \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} y & = & \begin{vmatrix} \alpha_1 & c_1 \\ \alpha_2 & c_2 \end{vmatrix} & \cdots(\mathrm{I})
\end{array}
$$

We also consider $(1) \times c_3 - (3) \times c_1$ and $(2) \times c_3 - (3) \times c_2$to get

$$
\begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} x + \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} y = \begin{vmatrix} \alpha_1 & c_1 \\ \alpha_3 & c_3 \end{vmatrix} \quad \cdots(\mathrm{II}) = (1)\times c_3 - (3)\times c_1
$$

$$
\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} x + \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} y = \begin{vmatrix} \alpha_2 & c_2 \\ \alpha_3 & c_3 \end{vmatrix} \quad \cdots(\mathrm{III}) = (2)\times c_3 - (3)\times c_2
$$

use the cofactor expansion of the 2nd Colum

Cramer's Rule (3)

We consider $-b_1 \times (III) + b_2 \times (II) - b_3 \times (I)$ to get

$$|\vec{a}\ \vec{b}\ \vec{c}|x + |\vec{b}\ \vec{b}\ \vec{c}|y = |\vec{\alpha}\ \vec{b}\ \vec{c}|$$

It follows from $|\vec{b}\ \vec{b}\ \vec{c}| = 0$ that

$$|\vec{a}\ \vec{b}\ \vec{c}|x = |\vec{\alpha}\ \vec{b}\ \vec{c}|$$

Accordingly if $D := |\vec{a}\ \vec{b}\ \vec{c}| \neq 0$, we find

$$x = \frac{1}{D}|\vec{\alpha}\ \vec{b}\ \vec{c}|$$

In the same way as above, we can derive

$$y = \frac{1}{D}|\vec{a}\ \vec{\alpha}\ \vec{c}|, \quad z = \frac{1}{D}|\vec{a}\ \vec{b}\ \vec{\alpha}|$$

m

Multi-variate Data (1)

We are given a set of nulti-variate data:

x	y	z
x_1	y_1	z_1
$\vdots$	$\vdots$	$\vdots$
x_n	y_n	z_n

We introduce a model

$$z = ax + by + c$$

and try to fit it to the given data. In this situation x and y are called the explanetary variables and z the objective variable.
We define a new variable

$$\varepsilon = z - (ax + by + c)$$

Then the given data entails the data for the variable ε:

$$\varepsilon_j = z_j - (ax_j + by_j + c) \quad (j = 1, \ldots, n)$$

z_j : observed data

$ax_j + by_j + c$: theoretical value based on the model

Multi-variate Data (2)

The variable ε is considered the errors of the data based on the model. We set up the coefficients a, b and c so that

- **(i)** $\varepsilon = 0$
- **(ii)** $V(\varepsilon)$ is minimized.

Remark that the condition (i) is equivalent to

$$\bar{\varepsilon} = \bar{z} - (a\bar{x} + b\bar{y} + c) = 0$$

We look into the condition (ii) by introducing a vector

$$\vec{\varepsilon} = \frac{1}{\sqrt{n}} \begin{pmatrix} \varepsilon_1 - \bar{\varepsilon} \\ \vdots \\ \varepsilon_n - \bar{\varepsilon} \end{pmatrix}$$

$V(\varepsilon) = \|\vec{\varepsilon}\|^2$

Multi-variate Data (3)

We have

$$\frac{1}{\sqrt{n}}(\varepsilon_j - \bar{\varepsilon}) = \frac{1}{\sqrt{n}}\{(z_j - ax_j - by_j - c) - (\bar{z} - a\bar{x} - b\bar{y} - c)\}.$$

$$= \frac{1}{\sqrt{n}}(z_j - \bar{z}) - a \cdot \frac{1}{\sqrt{n}}(x_j - \bar{x}) - b \cdot \frac{1}{\sqrt{n}}(y_j - \bar{y}) \quad (j = 1, ..$$

which gives

$$\vec{\varepsilon} = \vec{z} - a\vec{x} - b\vec{y} = \vec{z} - D\begin{pmatrix} a \\ b \end{pmatrix}$$

with

$$D = (\vec{x}\ \vec{y})$$

The method of minimum square

Theorem

Let $A = (\vec{p}\ \vec{q})$ a $n \times 2$ matrix and $\vec{c} \in \mathbf{R}^n$. If

$$\left(\vec{c} - A\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, A\begin{pmatrix} x \\ y \end{pmatrix}\right) = 0 \quad \left(\text{for any } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2\right) \qquad (1)$$

then

$$\left\|\vec{c} - A\begin{pmatrix} x \\ y \end{pmatrix}\right\|^2 \geq \left\|\vec{c} - A\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}\right\|^2 \quad \left(\text{for any } \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2\right) \qquad (2)$$

Moreover the condition (1) is equivalent to l

$${}^tAA\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = {}^tA\vec{c} \qquad (3)$$

Normal equation

holds when

$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$

Multi-variate data (4)

We apply the result in the last slide and find that

$$V(\varepsilon) = ||\vec{\varepsilon}||^2 = \left\| \vec{z} - D\begin{pmatrix} a \\ b \end{pmatrix} \right\|^2$$

taken the minimum value when

$${}^tDD\begin{pmatrix} a \\ b \end{pmatrix} = {}^tD\vec{z}$$

namely

$$\begin{pmatrix} V(x) & C_{xy} \\ C_{yx} & V(y) \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} C_{xz} \\ C_{yz} \end{pmatrix}$$

The matrix

$$V = \begin{pmatrix} V(x) & C_{xy} \\ C_{yx} & V(y) \end{pmatrix}$$

is called the variance matrix of x and y.

${}^tDD$ $\cancel{\begin{pmatrix} {}^t\vec{a} \\ {}^t\vec{b} \end{pmatrix}(\vec{a}\ \vec{b})}$

$$= \begin{pmatrix} {}^t\vec{x} \\ {}^t\vec{y} \end{pmatrix}(\vec{x}\ \vec{y}) = \begin{pmatrix} {}^t\vec{x}\vec{x} & {}^t\vec{x}\vec{y} \\ {}^t\vec{y}\vec{x} & {}^t\vec{y}\vec{y} \end{pmatrix} = \begin{pmatrix} \|\vec{x}\|^2 & (\vec{x},\vec{y}) \\ (\vec{y},\vec{x}) & \|\vec{y}\|^2 \end{pmatrix}$$

$${}^tD\vec{z} = \begin{pmatrix} {}^t\vec{x} \\ {}^t\vec{y} \end{pmatrix}\vec{z} = \begin{pmatrix} {}^t\vec{x}\vec{z} \\ {}^t\vec{y}\vec{z} \end{pmatrix} = \begin{pmatrix} (\vec{x},\vec{z}) \\ (\vec{y},\vec{z}) \end{pmatrix} /\!/ \begin{pmatrix} C_{xz} \\ C_{yz} \end{pmatrix}$$