Разумова Д. гр. 533
ВД_НИД-1

Условие:

$$x_i = \begin{pmatrix} x_i^1 \\ \vdots \\ x_i^m \end{pmatrix}, \quad i = \overline{1,n}$$

Вектор выборочного среднего:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Выборочная ковариационная матрица:

$$V = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})(x_i - \bar{X})^T$$

Решение:

Каноническая форма:

$$V = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})(x_i - \bar{X})^T =$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i x_i^T - \bar{X} x_i^T - x_i \bar{X}^T + \bar{X}\bar{X}^T\right) =$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i x_i^T - \left(\sum_{i=1}^{n} x_i\right)\bar{X}^T - \bar{X}\sum_{i=1}^{n} x_i^T + n\bar{X}\bar{X}^T\right)$$

NOTES

Сделаем следующие замены:

$$S = \sum_{i=1}^{n} X_i \text{ , т.е. построчная сумма } (m \times 1)$$

$$\sum_{i=1}^{n} X_i X_i^T = R$$

$$\bar{X} \sum_{i=1}^{n} X_i^T = \frac{S}{n} \cdot S^T$$

$$\left(\sum_{i=1}^{n} X_i\right) \bar{X}^T = S \cdot \frac{S^T}{n}$$

$$n \cdot \bar{X}\bar{X}^T = \frac{SS^T}{n}$$

Следовательно, канонической информацией будет (n, S, R), т.к. для расчета V нужны только n, S, R.

$$V = \frac{1}{n-1}\left(R - S \cdot \frac{S^T}{n} - \frac{S}{n} S^T + \frac{SS^T}{n}\right) =$$

$$= \frac{1}{n-1}\left(R - \frac{SS^T}{n}\right)$$

Теперь нужно убедиться в том, что выполнены все св-ва канонической информации:

а) существование и единственность.

NOTES

Существование следует из вида (n, S, R).
n - единственно, S и R - также единственны, т.к. не зависят от порядка x_i (явл. суммами).

b) полнота.
Были проведены равносильные преобразования при выводе формулы для V.

c) Элементарная каноническая информация.

$$n = 1 \Rightarrow S_i = x_i, \; R_i = x_i x_i^T$$

d) пустая каноническая информация.

$n = 0 \Rightarrow S = 0$ (формально вектор), $R = 0$ (формально матрица)

e) объединение.

$$(n_1, S_1, R_1) \oplus (n_2, S_2, R_2) = (n_1 + n_2, S_1 + S_2, R_1 + R_2)$$

Это следует из ассоциативности операции суммирования.

f) обновление.

NOTES

Добавление нового $i+1$ столбца:

$(n, S, R) \to (n + n_{i+1}, S + S_{i+1}, R + R_{i+1}) =$
$= (n+1, S + x_{i+1}, R + x_{i+1}x_{i+1}^T)$

g) минимальное число наблюдений.

Т.к. в знаменателе формулы для V стоит $n-1$, то $n=1$ не подходит $\Rightarrow n_{min} = 2$.

h) компактность и эффективность.

В случае канонической информации мы храним S (столбец, $m \times 1$), R (матрица, $m \times m$) и число столбцов n вместо n столбцов длиной m (если бы мы хранили всю информацию).

Выберем в качестве критерия компактности кол-во хранимых чисел. Тогда
$m + m^2 + 1 < mn \Rightarrow m(m+1) + 1 < mn \Rightarrow$
$m(m+1) < mn \Rightarrow n > m+1$ – при таком n будет выигрыш.

NOTES

Кол-во шагов: выигрыш, кот. будет максимальным при распараллеливании операции объединения.

NOTES