18. Матрица ковариаций. Многомерное нормальное распределение и его свойства.

Status	Completed

Ковариационной матрицей вектора $\vec{X}$ называется $C(\vec{X}) = E(\vec{X} - E\vec{X})^T(\vec{X} - E\vec{X})$.

Аналог дисперсии вводится только для случайных векторов. На протяжении этого и следующего параграфов мы будем изображать векторы в виде столбцов

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{pmatrix}.$$

Определение. *Матрицей ковариаций* случайного вектора X называется матрица $\mathbf{C}(X)$, у которой на месте с номером (i, j) стоит $c_{i,j} = Cov(X_i,\, X_j)$, $i, j = 1, \dots, n$.

Матрица ковариаций есть аналог дисперсии. При $n = 1$ она совпадает с дисперсией. В общем случае на главной диагонали у нее стоят дисперсии $\mathbf{D}X_1, \dots, \mathbf{D}X_n$, матрица симметрична относительно главной диагонали: $c_{i,j} = c_{j,i}$.

Мы знаем, что $\mathbf{D}(AX+B) = A^2\mathbf{D}X$ в одномерном случае, если A и B — константы. Аналогом этого свойства для случайных векторов является следующее утверждение.

Теорема. *Пусть A — матрица из констант, имеющая m строк и n столбцов, а B — вектор из констант размерности m. Тогда*

$$\mathbf{C}(AX + B) = A\mathbf{C}(X)A^T,$$

где верхний индекс T соответствует транспонированной матрице.

Доказательство. Обозначим для краткости $m_i = \mathbf{E}X_i$, $i = 1, \dots, n$, и воспользуемся следующим свойством произведения матриц: если умножить вектор-столбец на вектор-строку, то в итоге получим матрицу:

$$\begin{pmatrix} X_1 - m_1 \\ X_2 - m_2 \\ \dots \\ X_n - m_n \end{pmatrix} \cdot (X_1 - m_1, X_2 - m_2, \dots, X_n - m_n) =$$

$$= \begin{pmatrix} (X_1 - m_1)^2 & (X_1 - m_1)(X_2 - m_2) & \dots & (X_1 - m_1)(X_n - m_n) \\ (X_2 - m_2)(X_1 - m_1) & (X_2 - m_2)^2 & \dots & (X_2 - m_2)(X_n - m_n) \\ \dots & \dots & \dots & \dots \\ (X_n - m_n)(X_1 - m_1) & (X_n - m_n)(X_2 - m_2) & \dots & (X_n - m_n)^2 \end{pmatrix}.$$

Взяв теперь математическое ожидание от обеих частей, получим

$$\mathbf{E}(X - \mathbf{E}X)(X - \mathbf{E}X)^T = \mathbf{C}(X),$$

и аналогично

$$\begin{aligned} \mathbf{C}(AX + B) &= \mathbf{E}(AX + B - \mathbf{E}(AX + B))(AX + B - \mathbf{E}(AX + B))^T = \\ &= \mathbf{E}(AX + B - \mathbf{E}AX - B)(AX + B - \mathbf{E}AX - B)^T = \\ &= \mathbf{E}(A(X - \mathbf{E}X)(X - \mathbf{E}X)^T A^T) = \\ &= A(\mathbf{E}(X - \mathbf{E}X)(X - \mathbf{E}X)^T)A^T = A\mathbf{C}(X)A^T. \end{aligned}$$

Теорема доказана.

Многомерное нормальное распределение (которого не было на лекциях)

Мы уже рассматривали ранее в качестве примера частный случай плотности многомерного нормального закона, она соответствовала стандартному многомерному нормальному распределению. Сейчас введем этот закон распределения в общей форме.

Будем действовать по аналогии с одномерным случаем.

Пусть Y — одномерная случайная величина, имеющая стандартное нормальное распределение, $Y \Subset \Phi_{0,1}$. Взяв произвольные числа α и $\sigma > 0$, образуем случайную величину $X = \sigma Y + \alpha$, которая уже будет распределена по закону Φ_{α,σ^2} с плотностью

$$\varphi_{\alpha,\sigma^2}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-\alpha)^2/2\sigma^2}, \quad -\infty < t < \infty.$$

Тем самым мы получили общий вид плотности нормального распределения с помощью линейных преобразований над случайной величиной Y, имеющей стандартное нормальное распределение.

Так же поступим и в многомерном случае. Пусть Y — случайный вектор с координатами $Y_1, \ldots, Y_n$, имеющий многомерное стандартное нормальное распределение с плотностью

$$f_Y(t) = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} t_i^2\right\} = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2} t^T t\right\} = \prod_{i=1}^{n} \varphi_{0,1}(t_i).$$

Последнее означает, что координаты вектора $Y_1, \ldots, Y_n$ независимы и одинаково распределены в соответствии со стандартным нормальным законом. Здесь $\mathbf{C}(Y) = E$ — единичная матрица, $t = (t_1, \ldots, t_n)^T$ — вектор-столбец.

Возьмем произвольную невырожденную матрицу A размерности $n \times n$, состоящую из констант, и постоянный вектор $\alpha = (\alpha_1, \ldots, \alpha_n)^T$ и образуем новый случайный вектор

$$X = AY + \alpha.$$

Распределение получившегося вектора X и будем называть *многомерным нормальным* распределением.

Теорема. *Плотность многомерного нормального распределения задается формулой*

$$f_X(t) = \frac{\sqrt{\det Q}}{(2\pi)^{n/2}} \exp\left\{-(t-\alpha)^T Q (t-\alpha)/2\right\},$$

где $t = (t_1, \ldots, t_n)^T$, $Q = (\mathbf{C}(X))^{-1} = (A\mathbf{C}(Y)A^T)^{-1} = (AA^T)^{-1}$.

Заметим, что при $n = 1$ и $A = \sigma$ получаем $Q = 1/\sigma^2$, $\sqrt{\det Q} = 1/\sigma$.

Следствие 1. *Пусть случайный вектор* $X = (X_1, X_2, \ldots, X_n)^T$ *имеет многомерное нормальное распределение и все его компоненты попарно некоррелированны. Тогда они независимы.*

Доказательство. Вследствие некоррелированности компонент заключаем, что матрица ковариаций $\mathbf{C}(X)$ имеет диагональный вид: на главной диагонали стоят дисперсии $\mathbf{D}X_1, \ldots, \mathbf{D}X_n$, а все остальные элементы равны нулю. Обозначим для краткости $\sigma_i^2 = \mathbf{D}X_i$, $i = 1, \ldots, n$. Тогда матрица $Q = (\mathbf{C}(X))^{-1}$ также будет диагональной, у нее на главной диагонали будут стоять числа $\sigma_1^{-2}, \ldots, \sigma_n^{-2}$. По этой причине плотность распределения вектора X приобретает вид

$$f_X(t) = \frac{1}{\sigma_1 \ldots \sigma_n (2\pi)^{n/2}} \exp\left\{ -\sum_{i=1}^{n} \frac{(t_i - \alpha_i)^2}{2\sigma_i^2} \right\} = \prod_{i=1}^{n} \varphi_{\alpha_i \sigma_i^2}(t_i),$$

Следствие 2. *Пусть случайный вектор* $X = (X_1, X_2, \ldots, X_n)^T$ *имеет многомерное стандартное нормальное распределение (напомним: это соответствует тому, что все компоненты вектора независимы и имеют распределение* $\Phi_{0,1}$*). Образуем новый вектор* $Y = AX$*, где* A *— ортогональная матрица. Тогда вектор* Y *также будет иметь многомерное стандартное нормальное распределение.*

Доказательство. Ортогональная матрица, по определению, обладает свойством $A^T = A^{-1}$. По этой причине $\mathbf{C}(Y) = A\mathbf{C}(X)A^T = AA^T = E$ и, следовательно,

$$f_Y(t) = \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2} t^T t \right\} = \prod_{i=1}^{n} \varphi_{0,1}(t_i) = f_X(t),$$

что и требовалось доказать.