Ортогональные проекции нормальных векторов

Элементы линейной алгебры. Евклидово пространство

Напомним несколько понятий из линейной алгебры. *Действительное n-мерное евклидово пространство* $\mathcal{R}^n$ – это множество элементов (которые называются векторами) вида

$$x = \langle x_1, \dots, x_n \rangle, \qquad x_1, \dots, x_n \in \mathbb{R},$$

с линейными операциями

$$ax + by = \langle ax_1 + by_1, \dots, ax_n + by_n \rangle, \qquad a, b \in \mathbb{R},$$

скалярным произведением и нормой

$$(x, y) = \sum_{k=1}^{n} x_k y_k, \qquad \|x\| = \sqrt{(x, x)} = \left(\sum_{k=1}^{n} x_k^2 \right)^{1/2},$$

определёнными для любых $x = \langle x_1, \dots, x_n \rangle, y = \langle y_1, \dots, y_n \rangle \in \mathcal{R}^n$.

Линейные операции, скалярное произведение и норма обладают многими свойствами, из которых отметим следующие.

Разность $x - y$ элементов $x, y \in \mathcal{R}^n$ по определению задаётся как $y - z = x + (-1) \cdot y$; при этом $x - y = z$ тогда и только тогда, когда $x = y + z$, и $x = z$ тогда и только тогда, когда $x - z = \vec{0}$, где $\vec{0} = \langle 0, \dots, 0 \rangle$ есть так называемый нулевой элемент пространства $\mathcal{R}^n$.

Скалярное произведение симметрично, $(x, y) = (y, x)$, и линейно в том смысле, что $(ax + bz, y) = a(x, y) + b(z, y)$, для любых $x, y, z \in \mathcal{R}^n$ и любых $a, b \in \mathbb{R}$.

Для нормы имеем следующее свойство: $\|x\| = 0$ тогда и только тогда, когда $x = \vec{0}$. Отсюда следует, что $x = z$, если и только если $(x, y) = (z, y)$ для любого $y \in \mathcal{R}^n$. В самом деле, если $(x, y) = (z, y)$ для любого $y \in \mathcal{R}^n$, то, выбрав $y = x - z$, мы получаем $0 = (x - z, y) = \|x - y\|^2$, что эквивалентно $x - z = \vec{0}$, т. е. $x = z$.

Все эти свойства проверяются непосредственной подстановкой определений вектора, линейных операций, скалярного произведения и нормы.

Ортонормированным базисом (ОНБ) в пространстве $\mathcal{R}^n$ называется любой набор векторов $e_1, \dots, e_n$, удовлетворяющих условию

$$(e_i, e_j) = \delta_{ij} = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases} \qquad i, j = 1, \dots, n.$$

Это равносильно тому, что для любого $x \in \mathcal{R}^n$

$$x = \sum_{k=1}^{n} (x, e_k) e_k, \qquad \|x\|^2 = \sum_{k=1}^{n} (x, e_k)^2. \tag{1}$$

Числа (x, e_k), $k = 1, 2, \ldots, n$, называются *координатами вектора* x в ОНБ $e_1, e_2, \ldots, e_n$. В пространстве $\mathcal{R}^n$ существует бесконечно много ОНБ, среди которых имеется ОНБ, который мы назовем *естественным*: составляющие его векторы имеют вид

$$e_k^0 = \langle \underbrace{0, \ldots, 0, 1}_{k}, 0, \ldots, 0 \rangle, \qquad k = 1, \ldots, n. \tag{2}$$

Если набор чисел $x_1, \ldots, x_n$ образует вектор $x = \langle x_1, \ldots, x_n \rangle$, то эти числа суть координаты этого вектора в естественном базисе: $x_k = (x, e_k^0)$ для каждого $k = 1, \ldots, n$.

Пользуясь линейностью скалярного произведения, мы получаем

$$\begin{aligned}(x, z) = \left(\sum_{k=1}^{n} (x, e_k) e_k, \sum_{j=1}^{n} (z, e_j) e_j \right) = \sum_{k,j=1}^{n} (x, e_k)(z, e_j)(e_k, e_j) = \\ = \sum_{k,j=1}^{n} (x, e_k)(z, e_j) \delta_{ij} = \sum_{k=1}^{n} (x, e_k)(z, e_k). \end{aligned} \tag{3}$$

Это равенство можно интерпретировать следующим образом. Соберём координаты векторов x и z в векторы

$$\tilde{x} = \langle (x, e_1), \ldots, (x, e_n) \rangle = \langle \tilde{x}_1, \ldots, \tilde{x}_n \rangle, \quad \tilde{z} = \langle (z, e_1), \ldots, (z, e_n) \rangle = \langle \tilde{z}_1, \ldots, \tilde{z}_n \rangle.$$

Тогда по определению скалярного произведения

$$(\tilde{x}, \tilde{z}) = \sum_{k,j=1}^{n} \tilde{x}_k \tilde{z}_k = \sum_{k=1}^{n} (x, e_k)(z, e_k)$$

и равенство (3) доказывает, что $(x, z) = (\tilde{x}, \tilde{z})$. Другими словами, скалярное произведение не зависит от того, в каком ОНБ мы вычисляем координаты векторов. Выбрав $x = z$, мы получаем $\|x\|^2 = \|\tilde{x}\|^2$, и норма также не зависит от ОНБ.

Имеет место следующий факт: если задана ортономированная система (ОНС) $e_1, \ldots, e_r$, т. е. $(e_j, e_j) = \delta_{ij}$ для всех $i, j = 1, \ldots, r$, где $1 \leqslant r < n$, то в пространстве $\mathcal{R}^n$ найдутся элементы $e_{r+1}, \ldots, e_n$, такие что $e_1, \ldots, e_n$ образуют ОНБ.

Линейные операторы

Линейным оператором A, действующим из пространства $\mathcal{R}^n$ в пространство $\mathcal{R}^n$, называется правило, которое единственным образом сопоставляет каждому вектору $x \in \mathcal{R}^n$ вектор $Ax \in \mathcal{R}^n$ и удовлетворяет условию линейности

$$A(ax + bz) = aAx + bAz, \qquad x, z \in \mathcal{R}^n, \quad a, b \in \mathbb{R}.$$

Для любых линейных операторов A и B, действующих из $\mathcal{R}^n$ в $\mathcal{R}^n$, определены их линейная комбинация и произведение (композиция) операторов:

$$(aA + bB)x = aAx + bBx, \qquad (AB)x = A(Bx), \qquad x \in \mathcal{R}^n, \quad a, b \in \mathbb{R},$$

при этом $aA + bB$ и AB, очевидно также являются линейными операторами, действующими из $\mathcal{R}^n$ в $\mathcal{R}^n$.

Если задан ОНБ $e_1, \ldots, e_n$ в пространстве $\mathcal{R}^n$, то можно взаимно однозначно связать линейный оператор A и *матрицу оператора* размера $n \times n$, элементы которой равны $A_{ij} = (\tilde{e}_i, Ae_j)$, $i, j = 1, \ldots, n$. При этом операторам $aA + bB$ и AB сопоставляются соответственно линейная комбинация и произведение их матриц. Теория матриц в нашем курсе использоваться не будет, за исключением понятия определителя матрицы. Но можно показать, что определитель не зависит от ОНБ, поэтому мы будем говорить об определителе оператора A и обозначать его как $\det A$.

Сопряжённым к линейному оператору A называется линейный оператор A^*, такой что $(Ax, y) = (x, A^*y)$ для любых векторов $x, y \in \mathcal{R}^n$.

Матрица сопряженного оператора получается как результат транспонирования матрицы оператора A: $A^*_{ji} = A_{ij}$, $i, j = 1, \ldots, n$.

Унитарные (ортогональные) операторы

Пусть $\{e_k\}$ и $\{g_k\}$ – два ОНБ пространства $\mathcal{R}^n$. Линейный оператор U, действующий по правилу

$$Ue_k = g_k, \quad k = 1, \ldots, n, \quad \Longleftrightarrow \quad Ux = U\left(\sum_{k=1}^{n}(x, e_k)e_k\right) = \sum_{k=1}^{n}(x, e_k)g_k, \quad x \in \mathcal{R}^n,$$

называется *унитарным* (также используется термин *ортогональный оператор*). Мы видим, что унитарный оператор – это оператор перехода от ОНБ $\{e_k\}$ к ОНБ $\{g_k\}$.

Если U – унитарный оператор, то для любых $x, z \in \mathcal{R}^n$ в силу ортонормированности базиса $\{g_k\}$

$$(Ux, Uz) = \left(\sum_{k=1}^{n}(x, e_k)g_k, \sum_{j=1}^{n}(z, e_j)g_j\right) = \sum_{k=1}^{n}(x, e_k)(z, e_k) = (x, z),$$

где мы воспользовались равенством (3). Таким образом, унитарный оператор сохраняет неизменным скалярное произведение элементов, на которые действует, и, следовательно, норму: $\|Ux\| = \|x\|$ для любого $x \in \mathcal{R}^n$.

Перепишем равенство $(Ux, Uz) = (x, z)$ как

$$(Ux, Uz) = (U^*Ux, z) = (x, z), \qquad x, z \in \mathcal{R}^n,$$

следовательно, $U^*Ux = x$ для любого $z \in \mathcal{R}^n$ и $U^*U = I$, где I – единичный оператор ($Ix = x$ для любого $x \in \mathcal{R}^n$). Другими словами, для унитарного оператора его обратный совпадает с его сопряжённым. При этом обратный оператор U^{-1}, очевидно, действует как $U^{-1}g_k = e_k$, т.е. является оператором перехода от ОНБ $\{g_k\}$ к ОНБ $\{e_k\}$ и, следовательно, унитарным оператором. Из этих рассуждений и известных свойств определителя имеем

$$1 = \det I = \det U^*U = \det U^* \cdot \det U = (\det U)^2 \quad \Longrightarrow \quad |\det U| = 1.$$

Пусть $x = \langle x_1, \ldots, x_n \rangle$ и $\tilde{x} = \langle (x, e_1), \ldots, (x, e_n) \rangle = \langle \tilde{x}_1, \ldots, \tilde{x}_n \rangle$ – тот же вектор, но записанный в другом ОНБ. Тогда существует унитарный оператор U, такой, что $\tilde{x} = Ux$. Чтобы это показать, введём оператор U, действующий по правилу $Ue_k = e_k^0$, $k = 1, \ldots, n$, т.е. U переводит ОНБ $\{e_k\}$ в естественный ОНБ $\{e_k^0\}$. Его сопряжённый (или обратный) оператор действует как $U^*e_k^0 = e_k$, $k = 1, \ldots, n$, поэтому

$$\tilde{x} = \sum_{k=1}^{n} \tilde{x}_k e_k^0 = \sum_{k=1}^{n} (x, e_k) e_k^0 = \sum_{k=1}^{n} (x, U^*e_k^0) e_k^0 = \sum_{k=1}^{n} (Ux, e_k^0) e_k^0,$$

тем самым, в естественном ОНБ $\{e_k^0\}$ координаты $\tilde{x}_1, \ldots, \tilde{x}_n$ вектора $\tilde{x}$ в точности равны координатам $(Ux)_1, \ldots, (Ux)_n$ вектора Ux, и векторы совпадают (представляют собой одинаковые упорядоченные наборы вещественных чисел).

Операторы ортогонального проецирования (ортогональные проекторы)

Пусть $e_1, \ldots, e_r$ – ОНС в $\mathcal{R}^n$, $1 \leqslant r < n$. Линейный оператор Π из $\mathcal{R}^n$ в $\mathcal{R}^n$, действующий по правилу

$$\Pi x = \sum_{k=1}^{r} (x, e_k) e_k, \qquad x \in \mathcal{R}^n,$$

называется *оператором ортогонального проецирования* (*ортогональным проектором* на множество

$$\mathcal{L}(e_1, \ldots, e_r) = \left\{ x = \sum_{k=1}^{r} a_k e_k,\ a_1, \ldots, a_r \in \mathbb{R} \right\},$$

которое называется линейной оболочкой элементов $e_1, \dots, e_r$. Число r называется рангом или размерностью проектора.

Дополним ОНС элементами $e_{r+1}, \dots, e_n$ до ОНБ, тогда

$$\Pi e_k = \begin{cases} e_k & \text{для } k = 1, \dots, r, \\ \vec{0} & \text{для } k = r+1, \dots, n. \end{cases}$$

Очевидно, что для любого $x \in \mathcal{R}^n$

$$(I - \Pi)x = \sum_{k=1}^{n} (x, e_k)e_k - \sum_{k=1}^{r} (x, e_k)e_k = \sum_{k=r+1}^{n} (x, e_k)e_k,$$

следовательно, оператор $I - \Pi$ есть ортогональный проектор на $\mathcal{L}(e_{r+1}, \dots, e_n)$ ранга $n - r$. При этом для любых $x, z \in \mathcal{R}^n$

$$\big((I - \Pi)x, \Pi z\big) = \left(\sum_{k=1}^{r} (x, e_k)e_k, \sum_{j=r+1}^{n} (z, e_j)e_j \right) = \sum_{k=1}^{r} \sum_{j=r+1}^{n} (x, e_k)(z, e_j)(e_k, e_j) = 0,$$

другими словами, векторы $(I - \Pi)x$ и Πz ортогональны. Отсюда

$$\|x\|^2 = \|\Pi x + (I - \Pi)x\|^2 = \|\Pi x\|^2 + 2(\Pi x, (I - \Pi)x)\|(I - \Pi)x\|^2 = \|\Pi x\|^2 + \|(I - \Pi)x\|^2.$$

Нормальные случайные векторы

Рассмотрим n-мерную с.в. $(\xi_1, \dots, \xi_n)$. Если рассматривать любую её реализацию $\xi_1 = x_1, \dots, \xi_n = x_n$ как вектор $x = \langle x_1, \dots, x_n \rangle \in \mathcal{R}^n$, то естественно назвать эту с.в. случайным вектором, $\xi = \langle \xi_1, \dots, \xi_n \rangle$. На случайные векторы можно перенести все операции с обычными векторами: если все с.в. $\xi_1, \dots, \xi_n$ заданы на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, P)$, то $\xi \colon \Omega \mapsto \mathcal{R}^n$, при этом скалярное произведение и норма случайных векторов – это с.в. (ξ, η), $\|\xi\|$, задающиеся как

$$(\xi, \eta)(\omega) = (\xi(\omega), \eta(\omega)), \quad \|\xi\|(\omega) = \|\xi(\omega)\|, \qquad \omega \in \Omega.$$

Линейный оператор A действует на случайный вектор ξ по правилу

$$(A\xi)(\omega) = A(\xi(\omega)), \qquad \omega \in \Omega,$$

при этом $A\xi$ – случайный вектор со значениями в $\mathcal{R}^n$. Все формулы, которые мы писали для векторов из $\mathcal{R}^n$, справедливы и для случайных векторов.

Далее мы будем предполагать, что с.в. $\xi_1, \dots, \xi_n$ независимы и имеют одно и то же нормальное распределение $\mathbf{N}(\mu, \sigma^2)$ с параметрами $\mu = \mathrm{M}\xi_k$ и $\sigma^2 = \mathrm{D}\xi_k$ для всех $k = 1, \dots, n$. Тогда совместная плотность вероятности этих с.в. записывается как

$$p_\xi(x_1, \dots, x_n) = \prod_{k=1}^n p_{\xi_k}(x_k) = \prod_{k=1}^n \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x_k-\mu)^2}{2\sigma^2}} = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}}.$$

для $x_1, \dots, x_n \in \mathbb{R}$. Если ввести в $\mathcal{R}^n$ векторы $x = \langle x_1, \dots, x_n \rangle$ и $\vec{\mu} = \langle \mu, \dots, \mu \rangle$, то плотность вероятности принимает более компактный вид:

$$p_\xi(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{\|x-\vec{\mu}\|^2}{2\sigma^2}}, \qquad x \in \mathcal{R}^n.$$

Совместная функция распределения с.в. $\xi_1, \dots, \xi_n$ равна

$$F_\xi(y_1, \dots, y_n) = \prod_{k=1}^n F_{\xi_k}(y_k) = \prod_{k=1}^n \int_{-\infty}^{y_k} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x_k-\mu)^2}{2\sigma^2}}\, dx_k, \qquad y_1, \dots, y_n \in \mathbb{R}. \tag{4}$$

Теорема 1.
Пусть $\xi = \langle \xi_1, \dots, \xi_n \rangle$ – случайный вектор, в котором с.в. $\xi_1, \dots, \xi_n$ независимы и имеют нормальное распределение $\mathbf{N}(0, \sigma^2)$. Тогда координаты этого вектора в любом ОНБ $\{e_k\}$ евклидова пространства $\mathcal{R}^n$ имеют такое же распределение, т. е. $\eta_1 = (\xi, e_1), \dots, \eta_n = (\xi, e_n)$ – н.с.в. и $\eta_k \sim \mathbf{N}(0, \sigma^2)$ для всех $k = 1, \dots, n$.

Доказательство. Запишем совместную функцию распределения случайных величин $\eta_1, \dots, \eta_n$ как n-кратный интеграл:

$$F_\eta(y) = P((\xi, e_1) < y_1, \dots, (\xi, e_n) < y_n) = \int\limits_{x \in \mathcal{R}^n:\ (x,e_1)<y_1, \dots, (x,e_n)<y_n} \!\!\!\! \cdots \int \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{\|x\|^2}{2\sigma^2}}\, dx,$$

где $dx = dx_1 \dots dx_n$. Как мы знаем, существует унитарный оператор U, такой что $(x, e_k) = (Ux)_k$ для всех $k = 1, \dots, n$. Сделаем замену переменных $u = Ux$ в пространстве $\mathcal{R}^n$, тогда обратное преобразование переменных записывается как $x = U^{-1}u$. При этом $U^{-1} = U^*$ также является унитарным оператором, следовательно,

$$\|x\|^2 = \|U^{-1}u\|^2 = \|u\|^2,$$

а якобиан перехода, который при любом линейном преобразовании переменных равен модулю определителя матрицы преобразования, в данном случае равен $|\det U^{-1}| = 1$,

$$dx = dx_1 \dots dx_n = du_1 \dots du_n = du.$$

В результате такой замены переменных имеем для любых $y_1, \ldots, y_n \in \mathbb{R}$

$$F_\eta(y) = \underset{u\colon u_1<y_1,\ldots,u_n<y_n}{\int \cdots \int} \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{\|u\|^2}{2\sigma^2}}\, du = \prod_{k=1}^{n} \int_{-\infty}^{y_k} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{u_k^2}{2\sigma^2}}\, du_k.$$

Таким образом, $F_\eta(\,\cdot\,)$ с точностью до несущественной замены обозначений переменных совпадает с функцией распределения (4) случайного вектора ξ.

Распределения хи-квадрат и Стьюдента.

Определение. Случайная величина χ_n^2 имеет *распределение Пирсона* (или, в другой терминологии, *распределение хи-квадрат*) *с n степенями свободы*, если

$$\chi_n^2 = \sum_{k=1}^{n} \nu_k^2, \qquad (5)$$

где $\nu_1, \ldots, \nu_n$ – н.с.в., имеющие стандартное нормальное распределение $\mathbf{N}(0,1)$. Тогда мы пишем $\chi_n^2 \sim \mathbf{X}_n$.

Определение. Случайная величина τ_n имеет *распределение Стьюдента* (или, в другой терминологии, *T-распределение*) *с n степенями свободы*, если

$$\tau_n = \frac{\xi_0}{\sqrt{\frac{1}{n}\sum\limits_{k=1}^{n} \nu_k^2}} = \frac{\xi_0}{\sqrt{\frac{1}{n}\chi_n^2}},$$

где $\xi_0, \xi_1, \ldots, \xi_n$ – н.с.в., имеющие нормальное распределение $\mathbf{N}(0,1)$, или, эквивалентно, ξ_0 и χ_n^2 – н.с.в., имеющие соответственно распределение $\mathbf{N}(0,1)$ и распределение хи-квадрат с n степенями свободы. Тогда мы пишем $\tau_n \sim \mathbf{T}_n$.

Для случайного вектора $\xi = \langle \xi_1, \ldots, \xi_n \rangle$ введем с.в. (функции от $\xi_1, \ldots, \xi_n$)

$$\bar{\xi} = \frac{1}{n}\sum_{k=1}^{n} \xi_k, \qquad \widehat{\sigma}^2(\xi) = \frac{1}{n-1}\sum_{k=1}^{n} (\xi_k - \bar{\xi})^2.$$

С.в. $\bar{\xi}$ – это так называемое выборочное среднее, а с.в. $\widehat{\sigma}^2(\xi)$ – выборочная дисперсия с.в. $\xi_1, \ldots, \xi_n$. Легко видеть, что выборочное среднее обладает линейным свойством, аналогичным свойству математического ожидания: для любых вещественных чисел a, b и $x_1, \ldots, x_n$

$$\overline{ax+b} \overset{\text{def}}{=\!=} \frac{1}{n}\sum_{k=1}^{n}(ax_k + b) = \frac{a}{n}\sum_{k=1}^{n} x_k + \frac{1}{n}\sum_{k=1}^{n} b = a\bar{x} + b, \qquad \bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k. \qquad (6)$$

Теорема 2.

Пусть $\xi = \langle \xi_1, \ldots, \xi_n \rangle$ *– вектор, координаты которого независимы и имеют нормальное распределение* $\mathbf{N}(\mu, \sigma^2)$*. Тогда:*

$$1)\ \ T_1(\xi) = \frac{1}{\sqrt{n\sigma^2}} \sum_{k=1}^{n} (\xi_k - \mu) \sim \mathbf{N}(0,1);$$

$$2)\ \ T_2(\xi) = \frac{1}{\sigma^2} \sum_{k=1}^{n} (\xi_k - \mu)^2 \sim \mathbf{X}_n;$$

$$3)\ \ T_3(\xi) = \frac{1}{\sigma^2} \sum_{k=1}^{n} (\xi_k - \bar{\xi})^2 \sim \mathbf{X}_{n-1};$$

$$4)\ \ T_4(\xi) = \frac{1}{\sqrt{n\widehat{\sigma}^2(\xi)}} \sum_{k=1}^{n} (\xi_k - \mu) \sim \mathbf{T}_{n-1}.$$

Доказательство.

1. Введём случайный вектор ξ° как

$$\xi^\circ = \langle \xi_1^\circ, \ldots, \xi_n^\circ \rangle, \qquad \xi_k^\circ = \frac{\xi_k - \mu}{\sqrt{\sigma^2}}, \quad k = 1, \ldots, n.$$

Координаты этого вектора независимы и имеют нормальное распределение $\mathbf{N}(0,1)$. Тогда по свойствам нормального распределения .

$$T_1(\xi) = \frac{1}{\sqrt{n\sigma^2}} \sum_{k=1}^{n} (\xi_k - \mu) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \xi_k^\circ$$

распределена нормально со средним и дисперсией

$$\mathrm{M}T_1(\xi) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \mathrm{M}\xi_k^\circ = 0, \qquad \mathrm{D}T_1(\xi) = \frac{1}{n} \sum_{k=1}^{n} \mathrm{D}\xi_k^\circ = 1.$$

2. Перепишем с.в. $T_2(\xi)$ как

$$T_2(\xi) = \frac{1}{\sigma^2} \sum_{k=1}^{n} (\xi_k - \mu)^2 = \sum_{k=1}^{n} (\xi_k^\circ)^2,$$

откуда немедленно получаем $T_2(\xi) \sim \mathbf{X}_n$ по определению распределения хи-квадрат.

3. Для доказательства третьего и четвёртого утверждения зададим определённый случайный вектор e_1 единичной нормы и стандартным образом определим одномерный проектор Π_1 на $\mathcal{L}(e_1)$:

$$e = \left\langle \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}} \right\rangle, \qquad \Pi_1 x = (x, e_1) e_1, \qquad x \in \mathcal{R}^n.$$

В явном виде

$$(x, e_1) = \sum_{k=1}^{n} x_k \frac{1}{\sqrt{n}} = \sqrt{n}\bar{x}, \qquad \Pi_1 x = \sqrt{n}\bar{x}\left\langle \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}} \right\rangle = \langle \bar{x}, \dots, \bar{x} \rangle. \tag{7}$$

Если дополнить вектор e_1 векторами $e_2, \dots, e_n$ до ОНБ, то

$$(I - \Pi_1)x = x - \Pi_1 x = \sum_{k=1}^{n} (x, e_k)e_k - (x, e_1)e_1 = \sum_{k=2}^{n} (x, e_k)e_k,$$

а в явном виде

$$(I - \Pi_1)x = x - \Pi_1 x = \langle x_1 - \bar{x}, \dots, x_k - \bar{x} \rangle.$$

Введём выборочное среднее для вектора ξ°,

$$\overline{\xi^\circ} = \frac{1}{n} \sum_{k=1}^{n} \xi_k^\circ.$$

Тогда с учётом того, что $\xi_k - \bar{\xi} = (\xi_k - \mu) - \overline{(\xi - \mu)}$ в силу свойства (6), с.в. $T_3(\xi)$ можно записать как

$$T_3(\xi) = \frac{1}{\sigma^2} \sum_{k=1}^{n} (\xi_k - \bar{\xi})^2 = \sum_{k=1}^{n} (\xi_k^\circ - \overline{\xi^\circ})^2 = \|(I - \Pi_1)\xi^\circ\|^2 = \sum_{k=2}^{n} (\xi^\circ, e_k)^2.$$

По теореме 1 с.в. $(\xi^\circ, e_2), \dots, (\xi^\circ, e_n)$ независимы и стандартно нормально распределены, следовательно, $T_3(\xi) \sim \mathbf{X}_{n-1}$.

4. С учётом полученного выше мы можем записать выборочную дисперсию как

$$\begin{aligned}\widehat{\sigma}^2(\xi) &= \frac{1}{n-1} \sum_{k=1}^{n} (\xi_k - \bar{\xi})^2 = \frac{\sigma^2}{n-1} \sum_{k=1}^{n} \frac{((\xi_k - \mu) - (\bar{\xi} - \mu))^2}{\sigma^2} = \\ &= \frac{\sigma^2}{n-1} \sum_{k=1}^{n} (\xi_k^\circ - \overline{\xi^\circ})^2 = \frac{\sigma^2}{n-1} \|(I - \Pi_1)\xi^\circ\|^2 = \frac{\sigma^2}{n-1} \chi_{n-1}^2,\end{aligned}$$

где

$$\chi_{n-1}^2 = \|(I - \Pi_1)\xi^\circ\|^2 = \sum_{k=2}^{n} (\xi^\circ, e_k)^2 \sim \mathbf{X}_{n-1}.$$

Тогда

$$T_4(\xi) = \frac{1}{\sqrt{\widehat{\sigma}^2(\xi)}} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (\xi_k - \mu) = \frac{\frac{\sigma^2}{\sqrt{n}} \sum_{k=1}^{n} \xi_k^\circ}{\sqrt{\frac{\sigma^2}{n-1} \chi_{n-1}^2}} = \frac{(\xi^\circ, e_1)}{\sqrt{\frac{1}{n-1} \chi_{n-1}^2}},$$

где мы использовали первое равенство в (7) для (ξ°, e_1). Из теоремы 1 следует, что с.в. (ξ°, e_1) распределена стандартно нормально и не зависит от $(\xi^\circ, e_2), \dots, (\xi^\circ, e_n)$, поэтому $(\xi^\circ, e_1) \sim \mathbf{N}(0,1)$ не зависит от $\chi^2_{n-1} \sim \mathbf{X}_{n-1}$. Отсюда напрямую получаем, что $T_4(\xi) \sim \mathbf{T}_{n-1}$. Теорема доказана.

Замечание 1.
Мы видим, что $T_3(\xi)$ получается из $T_2(\xi)$ заменой μ на выборочное среднее $\bar{\xi}$, а $T_4(\xi)$ получается из $T_1(\xi)$ заменой σ^2 на выборочную дисперсию $\widehat{\sigma}^2(\xi)$. При этом

$$\mathrm{M}\bar{\xi} = \frac{1}{n}\sum_{k=1}^{n} \mathrm{M}\xi_k = \mu, \qquad \bar{\xi} = \frac{1}{n}\sum_{k=1}^{n} \xi_k \underset{n\to\infty}{\longrightarrow} \mu,$$

где сходимость (по вероятности или почти наверное) вытекает из закона больших чисел. Аналогично

$$\mathrm{M}\widehat{\sigma}^2(\xi) = \frac{\sigma^2}{n-1}\sum_{k=2}^{n} \mathrm{M}(\xi^\circ, e_k)^2 = \sigma^2, \qquad \widehat{\sigma}^2(\xi) \underset{n\to\infty}{\longrightarrow} \sigma^2,$$

где, опять же в силу закона больших чисел, мы имеем сходимость по вероятности или почти наверное. Таким образом, выборочные средние и дисперсия могут служить приближениями (оценками) для соответствующих параметров μ и σ^2 нормального распределения.

Приложение

- Распределение Пирсона (хи-квадрат) с n степенями свободы. Плотность вероятности
$$p_{\chi^2_n}(x) = \frac{1}{\Gamma(n/2)2^{n/2}}\, x^{(n-2)/2} e^{-x/2}, \qquad x > 0;$$
$p_{\chi^2_n}(x) = 0$, если $x < 0$. Моменты распределения: $\mathrm{M}\chi^2_n = n$, $\mathrm{D}\chi^2_n = 2n$.

- Распределение Стьюдента с n степенями свободы. Плотность вероятности
$$p_{\tau_n}(x) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{\pi n}}\left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, \qquad -\infty < x < +\infty.$$
Моменты распределения: $M\tau_n = 0$, $D\tau_n = \frac{n}{n-2}$ для $n > 2$, при $n = 1, 2$ дисперсия не существует.