Лекция 2

Случайные величины и случайные векторы

Случайной величиной называется скалярная величина $\mathbf{x}$, значение которой заранее не известно. В тоже время можно указать некую числовую меру (вероятность) того, что это значение будет принадлежать той или иной заранее определенной области значений. Здесь и далее любое значение случайной величины (СВ) $\mathbf{x}$ будем обозначать как x.

Свойства СВ определяются функцией распределения вероятностей $F_{\mathbf{x}}(x)$ (интегральная функция распределения вероятностей), являющаяся действительной функцией скалярного аргумента x и определяющей вероятность:

$$F_{\mathbf{x}}(x) = \mathbf{Pr}(\mathbf{x} \le x).$$

Функция $F_{\mathbf{x}}(x)$ является неотрицательной неубывающей функцией, причем $F_{\mathbf{x}}(-\infty) = 0$, $F_{\mathbf{x}}(+\infty) = 1$.

Другой характеристикой, однозначно определяющей свойства СВ, является плотность распределения вероятностей (дифференциальная функция распределения):

$$f_{\mathbf{x}}(x) = \lim_{dx \to 0} \frac{F_{\mathbf{x}}(x+dx) - F_{\mathbf{x}}(x)}{dx} = \lim_{dx \to 0} \frac{\mathbf{Pr}(x \le \mathbf{x} < x+dx)}{dx} = \frac{dF_{\mathbf{x}}(x)}{dx}.$$

Индекс снизу у функций $F_{\mathbf{x}}(x), f_{\mathbf{x}}(x)$ указывает на случайную величину, которой он соответствует и может опускаться. Далее также будем обозначать $f_{\mathbf{x}}(x) = p(x)$.

С учетом представленных ранее соотношений можно записать, что

$$F_{\mathbf{x}}(x) = \int_{-\infty}^{x} f_{\mathbf{x}}(x)dx, \quad \int_{-\infty}^{+\infty} f_{\mathbf{x}}(x)dx = 1.$$

Случайным вектором (СВк) называется вектор, каждая компонента которого является СВ $\mathbf{x} = (\mathbf{x}_1,\ldots,\mathbf{x}_n)^T$. Значение СВк $\mathbf{x}$ будем обозначать как $x = (x_1,\ldots,x_n)^T$.

Свойства СВк задаются совместной многомерной функцией распределения вероятностей и плотностью распределения вероятностей

$$F_{\mathbf{x}}(x) = \mathbf{Pr}(\mathbf{x}_1 \le x_1, \ldots, \mathbf{x}_n \le x_n),$$

$$f_{\mathbf{x}}(x) = \lim_{dx_i \to 0} \frac{\mathbf{Pr}(x_1 \le \mathbf{x}_1 < x_1 + dx_1, \ldots, x_n \le \mathbf{x}_n < x_n + dx_n)}{dx_1 \ldots dx_n} = \frac{\partial^n F_{\mathbf{x}}(x)}{dx_1 \ldots dx_n},$$

$$F_{\mathbf{x}}(x) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f_{\mathbf{x}}(u) du_1 \ldots du_n, \quad \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} f_{\mathbf{x}}(x) dx_1 \ldots dx_n = 1$$

С учетом общности исходного статистического описания случайных величин и случайных векторов дальнейшее изложение будем вести по отношению к более общему случаю – случайных векторов, который при $n = 1$ переходит к случаю рассмотрения случайной величины.

Пусть $\omega_i, i = \overline{1, M}$ обозначают совокупность некоторых событий, появляющихся с вероятностями $p(\omega_i) = p_i, \ i = \overline{1, M}$, таких что $\sum_{i=1}^{M} p(\omega_i) = 1$. В качестве каждого такого события может рассматриваться факт появления объекта, принадлежащего классу, имеющему соответствующий индекс.

Вероятности $p(\omega_i) = p_i, \ i = \overline{1, M}$ в этом случае трактуются как **априорные вероятности** событий или классов.

Вводятся также условные плотности распределения $\mathbf{x}$: $f_{\mathbf{x}}(x / \omega_i) = p(x / \omega_i), \ \mathrm{i} = \overline{1, M}$, которые трактуются как **апостериорные плотности распределения**, каждая из которых описывает распределение $\mathbf{x}$ при условии появления соответствующего события.

При таком описании плотность распределения $\mathbf{x}$ равна

$$f_{\mathbf{x}}(x) = p(x) = \sum_{i=1}^{M} p(x / \omega_i) p(\omega_i) = \sum_{i=1}^{M} p(x, \omega_i),$$

где $p(x, \omega_i)$ называется совместной плотностью величины $\mathbf{x}$ и события ω_i.

Формула Байеса:

$$p(x/\omega_i) = \frac{p(\omega_i/x)p(x)}{p(\omega_i)},\ \ p(\omega_i/x) = \frac{p(x/\omega_i)p(\omega_i)}{p(x)},\ \ i=\overline{1,M}.$$

Вероятность $p(\omega_i/x)$ называется **апостериорной вероятностью события.**

Аналогичным образом вводятся условные плотности распределения двух случайных векторов (случайных величин) $\mathbf{x}$ и $\mathbf{y}$

$$f_{\mathbf{x}}(x) = p(x) = \int p(x/y)p(y)dy = \int p(x,y)dy\ ,$$

$$p(x/y) = \frac{p(y/x)p(x)}{p(y)},\ \ p(y/x) = \frac{p(x/y)p(y)}{p(x)}\ ,$$

где $p(x,y), p(x/y)$ ($p(y/x)$) определяют соответственно совместную плотность случайных векторов $\mathbf{x}$ и $\mathbf{y}$ (фактически плотность составного вектора $\mathbf{z} = (\mathbf{x}^T, \mathbf{y}^T)^T$) и плотность распределения вектора $\mathbf{x}$ ($\mathbf{y}$) при фиксированном значении другого вектора y (x).

Здесь и далее интегралы понимаются как многократные, если область интегрирования не указана конкретно, то пределы по каждой компоненте случайного вектора предполагаются заданными от $-\infty$ **до** $+\infty$.

Если $\mathbf{x}$ и $\mathbf{y}$ статистически независимые случайнее векторы, то

$$p(x,y) = p(x)p(y),\ p(x/y) = p(x),\ p(y/x) = p(y).$$

Для СВк вводятся статистические характеристики, являющиеся параметрами безусловных и условных плотностей распределения. Из них основными являются моменты случайного вектора и, прежде всего, **вектор математического ожидания** и **матрица ковариации** (дисперсия для одномерной случайной величины):

$$m = M[\mathbf{x}] = \int xp(x)dx\ ,$$

$$m_i = M[\mathbf{x}/\omega_i] = \int xp(x/\omega_i)dx,\ i=\overline{1,M}\ ,$$

$$C = M[(\mathbf{x}-m)(\mathbf{x}-m)^T] = \int (x-m)(x-m)^T p(x)dx,$$

$$C_i = M[(\mathbf{x}-m_i)(\mathbf{x}-m_i)^T] = \int (x-m_i)(x-m_i)^T p(x/\omega_i)dx,\ i=\overline{1,M}.$$

Элементы любой матрицы ковариации можно представить в виде

$$C = \|c_{kt}\|, c_{kt} = \boldsymbol{\sigma}_k r_{kt} \boldsymbol{\sigma}_t,\ t=\overline{1,n},\ k=\overline{1,n}\ ,$$

где σ_k, σ_t – средние квадратичные отклонения (СКО), определяемые как корни квадратные от дисперсий $D_k = \sigma_k^2, D_t = \sigma_t^2$ соответствующих компонентов СВк; r_{kt}, $0 \le |r_{kt}| \le 1$ – являются коэффициентами корреляции компонентов случайного вектора.

Значения r_{kt} при $k = t,\ t, k = \overline{1,n}$ равны единице, а это означает, что диагональные элементы матрицы ковариации соответствуют дисперсиям компонентов случайного вектора $diag C = \{\sigma_1^2, \ldots, \sigma_n^2\}$.

При $n = 1$ для описании случайной величины вместо матрицы фактически используется скалярная величина, называемая дисперсией $D = \sigma^2$. Всё сказанное аналогичным образом распространяются и на условные матрицы ковариации $C_i, i = \overline{1,M}$.

Важным частным случаем СВк является **гауссовский случайный вектор** (случайная величина) (ГСВ).

Существенно, что его распределение определяется только двумя параметрами математическим ожиданием m и матрицей ковариации C.

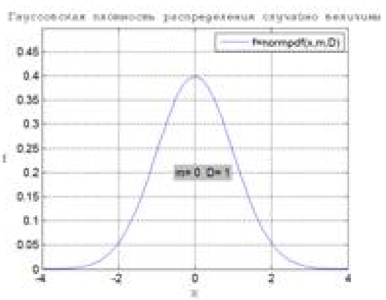
Плотность распределения ГСВ выражается следующим образом:

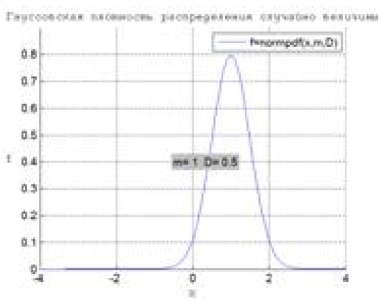
$$p(x) = N(x, m, C) = \frac{1}{(2\pi)^{n/2}|C|^{1/2}} \exp[-\frac{1}{2}(x - m)^T C^{-1}(x - m)],$$

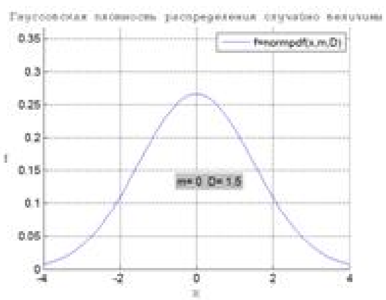
где $|C|$ – определитель матрицы.

Величина $\rho = [(x - m)^T C^{-1}(x - m)]^{1/2}$ называется расстоянием Махалонобиса. Если матрица ковариации является единичной (случай, когда все компоненты случайного вектора независимы и имеют дисперсии, равные единице), то расстояние Махаланобиса переходит в расстояние Евклида $\rho = [(x - m)^T (x - m)]^{1/2}$.

В качестве примера можно провести визуализацию графиков гауссовской плотности распределения вероятностей

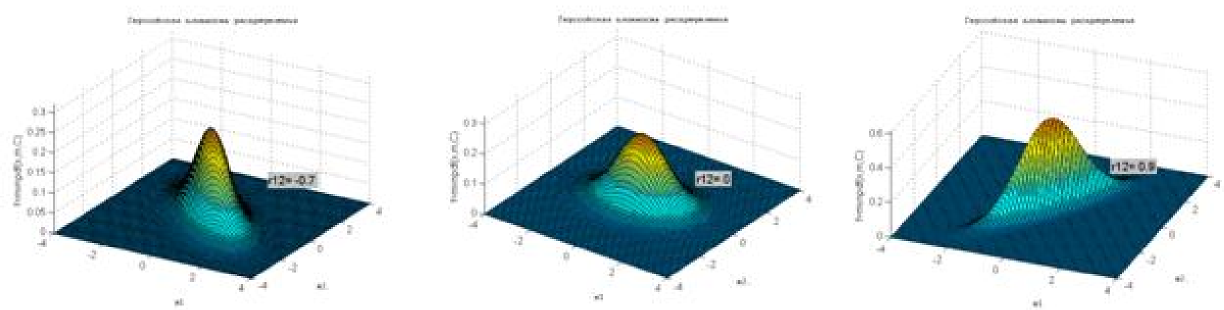






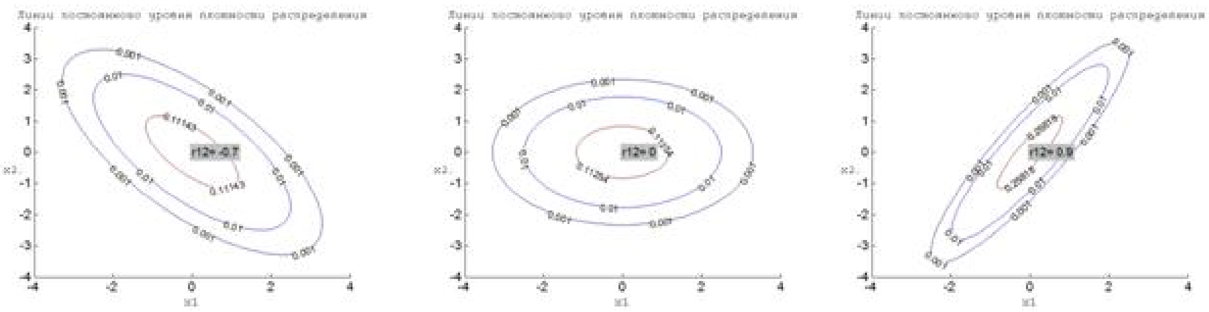
а) б) с)

Рис.1. Графики гауссовской плотности распределения вероятностей случайной величины с различными параметрами



а) б) с)

Рис.2. Графики гауссовской плотности распределения вероятностей случайного вектора с различными параметрами



а) б) с)

Рис.3. Графики линий постоянного уровня плотности распределения вероятностей случайного вектора (рис.2) с различными параметрами

Линейные преобразования и квадратичные формы при описании СВ

Любое линейное преобразование ГСВ с параметрами m_x, C_x, задаваемое прямоугольной матрицей H преобразует его к новому ГСВ $\mathbf{y} = H\mathbf{x}$ с параметрами $m_y = Hm_x$, $C_y = HC_xH^T$.

Квадратичной формой матрицы C (необязательно матрицы ковариации) относительно произвольного вектора x является **скалярная величина**

$$x^T C x = \sum_{k=1}^{n}\sum_{t=1}^{n} x_k c_{kt} x_t .$$

Положительной (неотрицательно определенной) матрицей называется такая, что для любого $x \neq 0$ выполняется

$$x^T C x > 0 \ (x^T C x \geq 0).$$

Любая матрица ковариации случайного вектора является, как минимум, неотрицательно определенной.

Заметим, что выражение для квадратичной формы обратной матрицы ковариации C^{-1} относительно вектора $(x - m)$ присутствует под знаком экспоненты в выражении для плотности ГСВ.

Рассмотрим для простоты СВк с нулевым математическим ожиданием (центрированный случайный вектор).

При преобразовании системы координат компонентов такого случайного вектора с использованием ортонормированного линейного преобразования получается новый СВк

$$\mathbf{y} = H\mathbf{x}, \ HH^T = H^T H = I .$$

При этом квадратичная форма матрицы ковариации преобразуется следующим образом:

$$x^T C x = x^T H^T H C H^T H x = y^T L y, \ L = HCH^T,$$

где L – матрица ковариации вектора $\mathbf{y} = H\mathbf{x}$.

Пусть столбцы матрицы H^T: $h_j = (h_{j1},\dots h_{jn})^T, \ j = \overline{1,n}$ состоят из собственных векторов матрицы, т.е. являются решениями уравнений $Ch = \boldsymbol{\lambda} h$, $|C - \boldsymbol{\lambda} I| = 0$ в виде

$$Ch_j = \boldsymbol{\lambda}_j h_j, \quad h_j = (h_{j1},\dots h_{jn})^T, \ j = \overline{1,n},$$

где $\boldsymbol{\lambda}_j, \ j = \overline{1,n}$ – собственные числа матрицы C.

Тогда матрица ковариации полученного после преобразования вектора $\mathbf{y} = H\mathbf{x}$ имеет вид диагональной матрицы с диагональными элементами, равными собственным числам матрицы ковариации исходного СВк

$$L = HCH^T = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}.$$

При этом исходная квадратичная форма преобразуется к виду

$$x^T Cx = \sum_{k=1}^{n}\sum_{t=1}^{n} x_k c_{kt} x_t = y^T Ly = \sum_{k=1}^{n} \lambda_k y_k^2 .$$

Для обращенной матрицы ковариации применение подобного преобразования приводит к выражению

$$x^T C^{-1} x = \sum_{k=1}^{n}\sum_{t=1}^{n} x_k \bar{c}_{kt} x_t = y^T L^{-1} y = \sum_{k=1}^{n} \lambda_k^{-1} y_k^2, \quad C^{-1} = \|\bar{c}_{kt}\|.$$

Очевидно, что применение ортонормирующего преобразования существенно упрощает выражение для плотности распределения ГСВ

$$p(y = Hx) = N(y, m_y, L) = \frac{1}{(2\pi)^{n/2} \prod_{k=1}^{n} \lambda_k^{1/2}} \exp\left[-\frac{1}{2}\sum_{k=1}^{n} \frac{y_k^2}{\lambda_k}\right],$$

поскольку полученный при этом вектор имеет статистически независимые компоненты.

[◄ Список сданных лаб]
Перейти на...
Перейти на...

[Лекция 3 ►]