№ 17

Выборка:

X_1	4	0	-1	3	4
X_2	2	-3	-2	1	2
X_3	3	2	2	1	-3

$$\longrightarrow \begin{pmatrix} 4 & 2 & 3 \\ 0 & -3 & 2 \\ -1 & -2 & 2 \\ 3 & 1 & 1 \\ 4 & 2 & -3 \end{pmatrix} \Rightarrow \bar{x} = (2;0;1) \rightarrow$$

1) центрируем данные:

$$\rightarrow \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{pmatrix} = X_c$$

2) найдем:

$$C = X_c^T X_c =$$

$$= \begin{pmatrix} 22 & 21 & -9 \\ 21 & 22 & -9 \\ -9 & -9 & 22 \end{pmatrix}$$

$$\frac{1}{N-1} C = \begin{pmatrix} \frac{11}{2} & \frac{21}{4} & -\frac{9}{4} \\ \frac{21}{4} & \frac{11}{2} & -\frac{9}{4} \\ -\frac{9}{4} & -\frac{9}{4} & \frac{11}{2} \end{pmatrix}$$

3) Наход. собств. числа

$$\det(C - \lambda I) =$$

$$= \begin{vmatrix} 22-\lambda & 21 & -9 \\ 21 & 22-\lambda & -9 \\ -9 & -9 & 22-\lambda \end{vmatrix} =$$

$$= -x^3 + 66x^2 - 849x + 784$$

Корни:

$x = 1$

$x = 16$

$x = 49$

4) Нормир собств. векторы:

$$\lambda_1 = 1 \Rightarrow v_1 = \frac{\sqrt{2}}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$\lambda_2 = 16 \Rightarrow v_2 = \frac{\sqrt{11}}{11} \begin{pmatrix} 1 \\ 1 \\ 8 \end{pmatrix} \text{ -- }$$

главные компоненты

$$\lambda_3 = 49 \Rightarrow v_3 = \frac{\sqrt{22}}{22} \begin{pmatrix} 3 \\ 3 \\ -8 \end{pmatrix} \text{ -- }$$

а) $\frac{1}{N-1}\lambda_1 = \frac{1}{4}\delta_1^2 = \frac{1}{4} = 0{,}25$

б) $\frac{1}{N-1}\lambda_2 = \frac{16}{4} = 4$ — " — дисперсии по главным компонентам.

в) $\frac{1}{N-1}\lambda_3 = \frac{49}{4} = 12{,}25$

№35

Квадр. ф-я потер: $L(y', y) = (y'-y)^2$.

Доказать, что если $f^*(x) = \arg\min E((Y-c)^2 | X=x)$, то

$f^*(x) = E(Y | X=x)$.

Чему равен средний риск $R(f^*)$?

1) $\arg\min E((Y-c)^2 | X=x) = E(Y^2 - 2Yc + c^2 | X=x) =$

$$= \underbrace{E(Y^2 | X=x) - 2cE(Y | X=x) + c^2}_{\text{функция } g(c)} \to \min_c$$

$g'(c) = -2E(Y | X=x) + 2c = 0 \to$

$\to c_{opt} = E(Y|X=x) \to = E(Y | X=x)$

2) $R(f^*) = \int E((f(x) - Y)^2 \,|\, x)\, p(x)\, dx =$

$= \int E((E(Y|X=x) - Y)^2 \,|\, x)\, p(x)\, dx = E_x D_y (Y | X=x)$

$\qquad \text{"}(Y - \langle Y | X=x\rangle)^2$

№ 36

Ф-я потерь: $L(y', y) = |y' - y|$

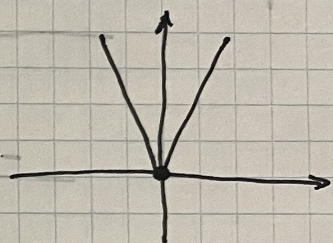
Док-ть: $R(f) \to \min$ при $f(x) = \text{median}(Y|X=x)$.

Решение:

$$R(f) = \int E(|f(x) - Y| \,/\, x)\, p(x)\, dx = \int E(f(x) - Y | x)\, p(x)\, dx +$$

$$+ \int E((f(x - Y) | x) \cdot p(x)\, dx \to \min$$

- можно оценить геометрически:



Тогда:

$$\int E(f(x) - Y | x)\, p(x)\, dx + \int E(Y - f(x) | x)\, p(x)\, dx \geq 0$$

минимум будет при равенстве, следовательно

$$\int E(f(x) - Y | x)\, p(x)\, dx = \int E(Y - f(x) | x)\, p(x)\, dx$$

из этого следует, что $\underline{f(x) = \text{median}(Y|X=x)}$.

№ 37

Как выглядит ф-я потерь, если минимум ср. риску дает условная мода?

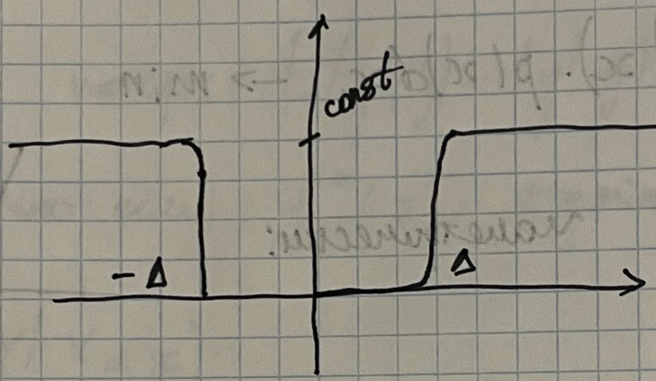
$$f^*(x) = \text{moda}(Y|X=x) = \max P_r(Y|X=x)$$

$$R(f^*(x)) \to \min$$

Если в окружении какого-то числа ошибки нулевые, а для всей остальной прямой — равные, то условная мода дает минимум ср. риска, т.е. мода хар-ется наиболее часто встречающимся значением.

$$L(y', y) = \begin{cases} 0, & y' - y \leq \Delta \\ const, & |y' - y| > \Delta \end{cases}$$

График:



Упражнение 4.1

1) Док-ть, что можно положить $v_0 = \bar{x}$.
2) Описать множество всех возможных v_0, дост. минимум суммы квадратов расстояний до искомого многообразия.