Задание 17

x_1	4	0	-1	3	4
x_2	2	-3	-2	1	2
x_3	3	2	2	1	-3

Найти гл. напр-я и дисп по гл. комп-м. Изобразить точки и гл. направл.

Итак, имеем

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

$$X_c = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{pmatrix}$$ – центрированные данные.

$$C = X_c^T X_c \quad \begin{pmatrix} 2 & -2 & -3 & 1 & 2 \\ 2 & -3 & -2 & 1 & 2 \\ 2 & 1 & 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{pmatrix} = \begin{pmatrix} 22 & 21 & -9 \\ 21 & 22 & -9 \\ -9 & -9 & 22 \end{pmatrix}$$

$$\frac{1}{N-1} C = \begin{pmatrix} 5.5 & 5.25 & -2.25 \\ 5.25 & 5.5 & -2.25 \\ -2.25 & -2.25 & 5.5 \end{pmatrix}$$ – выборочная матрица ковариации

Находим с.ч. и с.в-ры матр C:

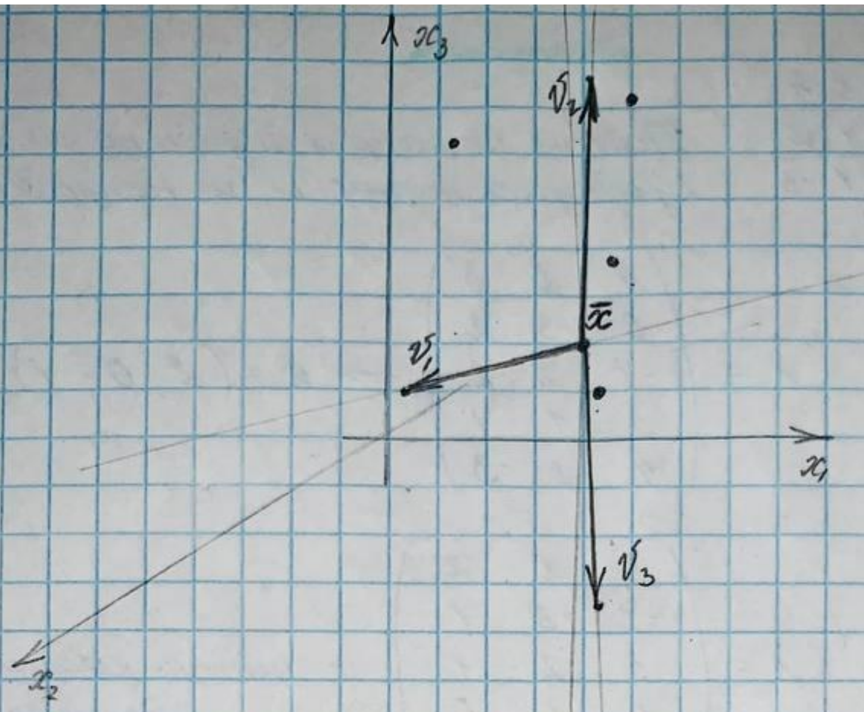
$$\begin{pmatrix} 22-\lambda & 21 & -9 \\ 21 & 22-\lambda & -9 \\ -9 & -9 & 22-\lambda \end{pmatrix} \longrightarrow \lambda_1 = 1 \quad \lambda_2 = 16 \quad \lambda_3 = 49$$

$$v_1 = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad v_2 = \frac{\sqrt{11}}{11}\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \quad v_3 = \frac{\sqrt{22}}{22}\begin{pmatrix} 3 \\ 3 \\ -2 \end{pmatrix}$$

это главн. компоненты

Дисперс. по гл. комп-м:

$$\frac{1}{N-1}\lambda_1 = 0.25 \qquad \frac{1}{N-1}\lambda_2 = 4 \qquad \frac{1}{N-1}\lambda_3 = 12.25$$



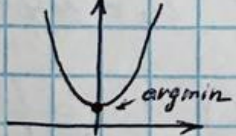
Задание 35

$L(y'; y) = (y' - y)^2$

Док-ть, что если $f^*(x) = \arg\min_c M((Y-c)^2 / X = x)$, то
$f^*(x) = M(Y / X = x)$.
Чему равен средний риск $R^*(f)$?

$\arg\min_c M((Y-c)^2 / X = x) \in \{c : \forall k \; M((Y-k)^2 / X = x) \geq$
$\geq M((Y-c)^2 / X = x)\}$

Для квадр. ф-ции min дости-ся в верш. параболы
→ при фикс. Y



$\arg\min_c M((Y-c)^2 / X = x) = M(Y / X = x)$

$((Y-c)^2)'_c = 0 \qquad (-2(Y-c))'_c > 0$

$-2(Y-C) = 0 \qquad 2 > 0 \to \min$

$c = Y$

Средний риск равен среднеквадр. ошибке:

$R(f^*) = M((y' - y)^2)$

Задание 36

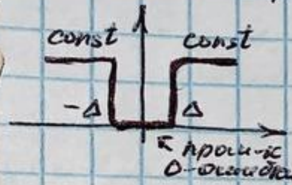
$L(y'; y) = |y' - y|$

Док-ть, что min ср. риску дост-т $f(x) = median\,(Y \mid X = x)$.

Известно, что $|y-y'|^2 = (y-y')^2$. Данная ф-ция симм. относит. центра тяжести → ср. риск будет равен серединному значению ф-ции → и есть медиана распределения.

Задание 37

Чтобы min ср. риску давала усл. мода, нужно, чтобы потери от ошибок были нулевыми в окр-ти некот. числа, а для всей остальной прямой были равными (т.к. мода характеризует наиб. часто встреч-ся знач-е).



Тогда ф-цию потерь можем задать инт-но:

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

Упражнение 4.1

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x^{(i)}$$

$H_k = v_0 + L(v_1 \dots v_k)$ – k-мерное лин. многообразие

Будем искать v_0 как решение задачи minimize H_0:

$$v_0 = \underset{a_0 \in \mathbb{R}^n}{\arg\min} \left(\sum_{i=1}^{N} dist^2(x_i, H_0) \right) = \underset{a_0 \in \mathbb{R}^n}{\arg\min} \left(\sum_{i=1}^{N} \| x^{(i)} - a_0 \|^2 \right) =$$

это и есть выборочное среднее

$$= \frac{1}{N} \sum_{i=1}^{N} x^{(i)} = \bar{x}.$$