№17

$$X = \begin{bmatrix} 4 & 2 & 3 \\ 0 & -3 & 2 \\ -1 & -2 & 2 \\ 3 & 1 & 1 \\ 4 & 2 & -3 \end{bmatrix} \qquad N = 5 \qquad \bar{x} = (2, 0, 1)$$

$$X_c = X - \begin{bmatrix} \bar{x} \\ \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{bmatrix}$$

$$C = X_c^T X_c = \begin{bmatrix} 22 & 21 & -9 \\ 21 & 22 & -9 \\ -9 & -9 & 22 \end{bmatrix}$$

Найдем собственные числа и векторы C.

$$\det(C - \lambda I) = -(\lambda - 1)(\lambda - 16)(\lambda - 49) = 0$$

$$\lambda_1 = 1 \qquad v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \cdot \frac{\sqrt{2}}{2}$$

$$\lambda_2 = 16 \qquad v_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1 \end{bmatrix} \cdot \frac{3\sqrt{11}}{11} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} \cdot \frac{\sqrt{11}}{11}$$

$$\lambda_3 = 49 \qquad v_3 = \begin{bmatrix} -3/2 \\ -3/2 \\ 1 \end{bmatrix} \cdot \frac{2\sqrt{22}}{22} = \begin{bmatrix} -3 \\ -3 \\ 2 \end{bmatrix} \cdot \frac{\sqrt{22}}{22}$$

Векторы v_1, v_2, v_3 – главные компоненты

$$\frac{1}{N-1}\lambda_1 = \frac{1}{N-1}\sigma_1^2 = \frac{1}{4} \qquad \frac{1}{N-1}\lambda_2 = \frac{1}{N-1}\sigma_2^2 = 4$$

$$\frac{1}{N-1}\lambda_3 = \frac{1}{N-1}\sigma_3^2 = \frac{49}{4}$$ – дисперсии по гл-м компонентам

Упражнение 4.1.

Сначала опишем мн-во всех возможных v_0 доставляющих минимум сумме квадратов расстояний до искомого многообразия:

$$\left\{ v_0 : \underset{v_0}{\operatorname{argmin}} \sum_{i=1}^{N} dist^2(x^{(i)}, v_0) \right\}$$

Докажем, что это мн-во состоит из 1^го значения $v_0 = \bar{x}$:

$$\underset{v_0}{\operatorname{argmin}} \sum_{i=1}^{N} dist^2(x^{(i)}, v_0) = \underset{v_0}{\operatorname{argmin}} \sum_{i=1}^{N} \|x^{(i)} - v_0\|^2 =$$

$$= \underset{v_0}{\operatorname{argmin}} \sum_{i=1}^{N} \left[(x_1^{(i)} - v_{0_1})^2 + (x_2^{(i)} - v_{0_2})^2 + \dots \right]$$

$$v_{0_1} = \underset{v_{0_1}}{\operatorname{argmin}} \sum_{i=1}^{N} (x_1^{(i)} - v_{0_1})^2 =$$

$$= \underset{v_{0_1}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} (x_1^{(i)})^2 - 2 v_{0_1} \sum_{i=1}^{N} x_1^{(i)} + N v_{0_1}^2 \right\}$$

$$= \underset{v_{0_1}}{\operatorname{argmin}} \left\{ N v_{0_1}^2 - 2 v_{0_1} \sum_{i=1}^{N} x_1^{(i)} \right\}$$

Видим, что выражение в скобках является параболой, чьи ветви направлены вверх, значит минимум будет достигаться в единственной точке, равной:

$$v_{0_1} = \frac{2 \sum_{i=1}^{N} x_1^{(i)}}{2N}$$

Относительно др-х компонент v_0 действуют такие же рассуждения, значит:

$$\exists!\ v_0 = \frac{\sum_{i=1}^{N} x^{(i)}}{N} = \bar{x}$$