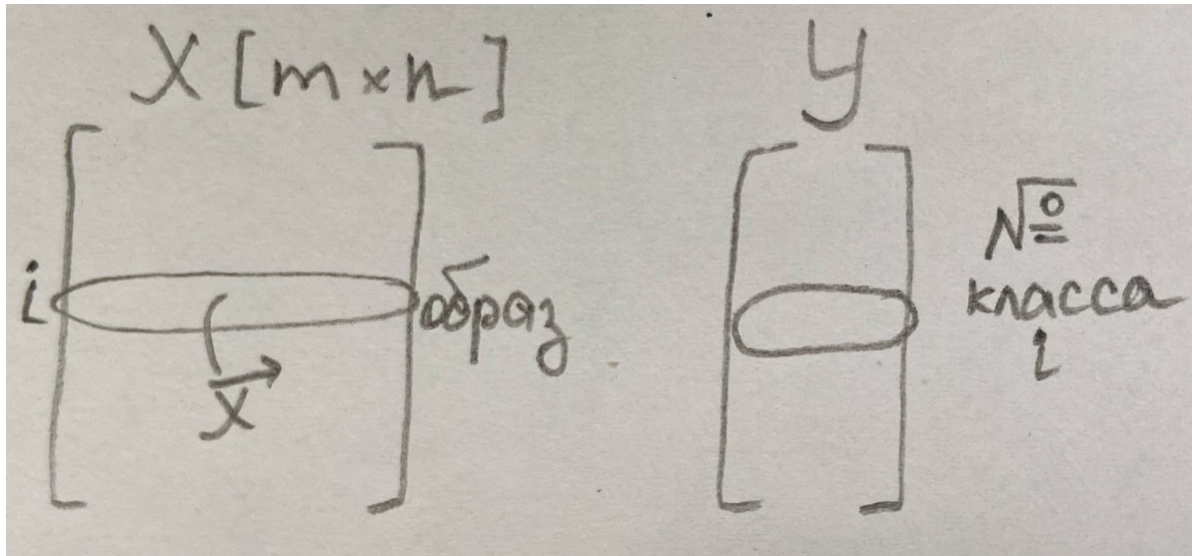


Распознавание образов или классификация

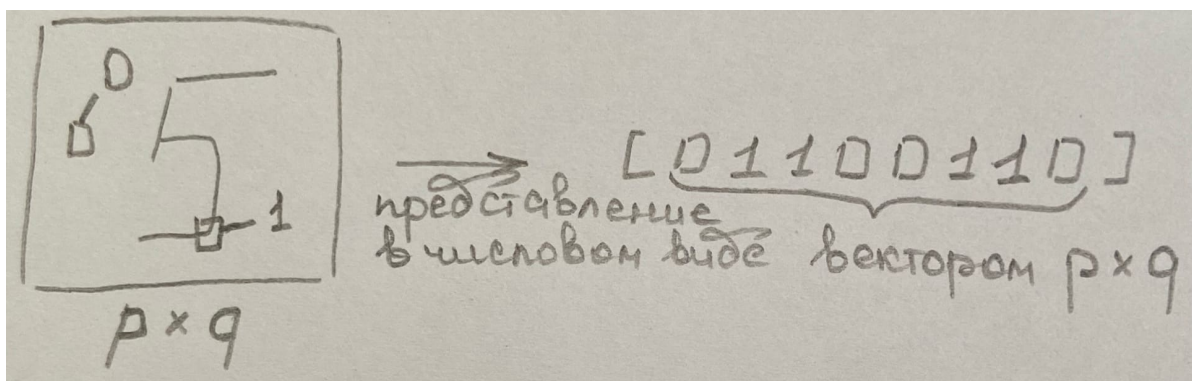


$Y \in [1, 2, \dots, k]$, k -число классов. Y представлен в номинальной шкале.

Ищем ф-цию $Y=f(x)$ это представление какого-то i -большого или i -образца. Имея обучающую выборку, т.е. набор объектов, которые подлежат распознаванию с известной принадлежностью каждого объекта классу какому-то.

Мы пытаемся построить ф-цию $f(x)$, которая будет выдавать нам номер класса на выходе. Для того, чтобы научиться на обучающей выборке, построим такую ф-цию $f(x)$, т.е. запрограммируем и шлем куда-то.

Одна из классических задач распознавая образа – это распознавание рукописных цифр.

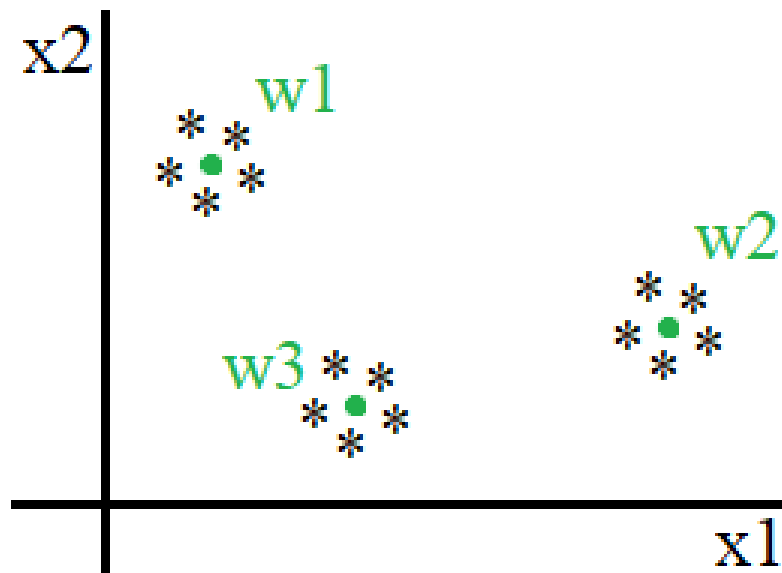


Если через какой-то пиксель прошло то, что написал клиент, то получаем 1, если пиксель в пустом поле, то будет 0.

Когда распознаем цифры у нас десять классов.

Наиболее сложная задача – распознавание текстов рукописных. Но и эта задача успешно решается. Если распознаем слова из латинских букв, то получаем 26 классов.

Методов распознавания куда и они продолжают появляться.



Эталон - среднее по всем точкам каждого класса (зеленая точка на рисунке)

Среднее по всем классам

$$\vec{m}_2 = \begin{bmatrix} \frac{1}{9} \sum_{l=1}^{16} x_{1e} \\ \frac{1}{9} \sum_{l=1}^{16} x_{2e} \end{bmatrix}$$

т.е. взяли все точки принадлежащие первому классу, например их 7, сложили их (сложить все первые и вторые координаты) и делим на количество точек.

ω_1 : $n_1=7$

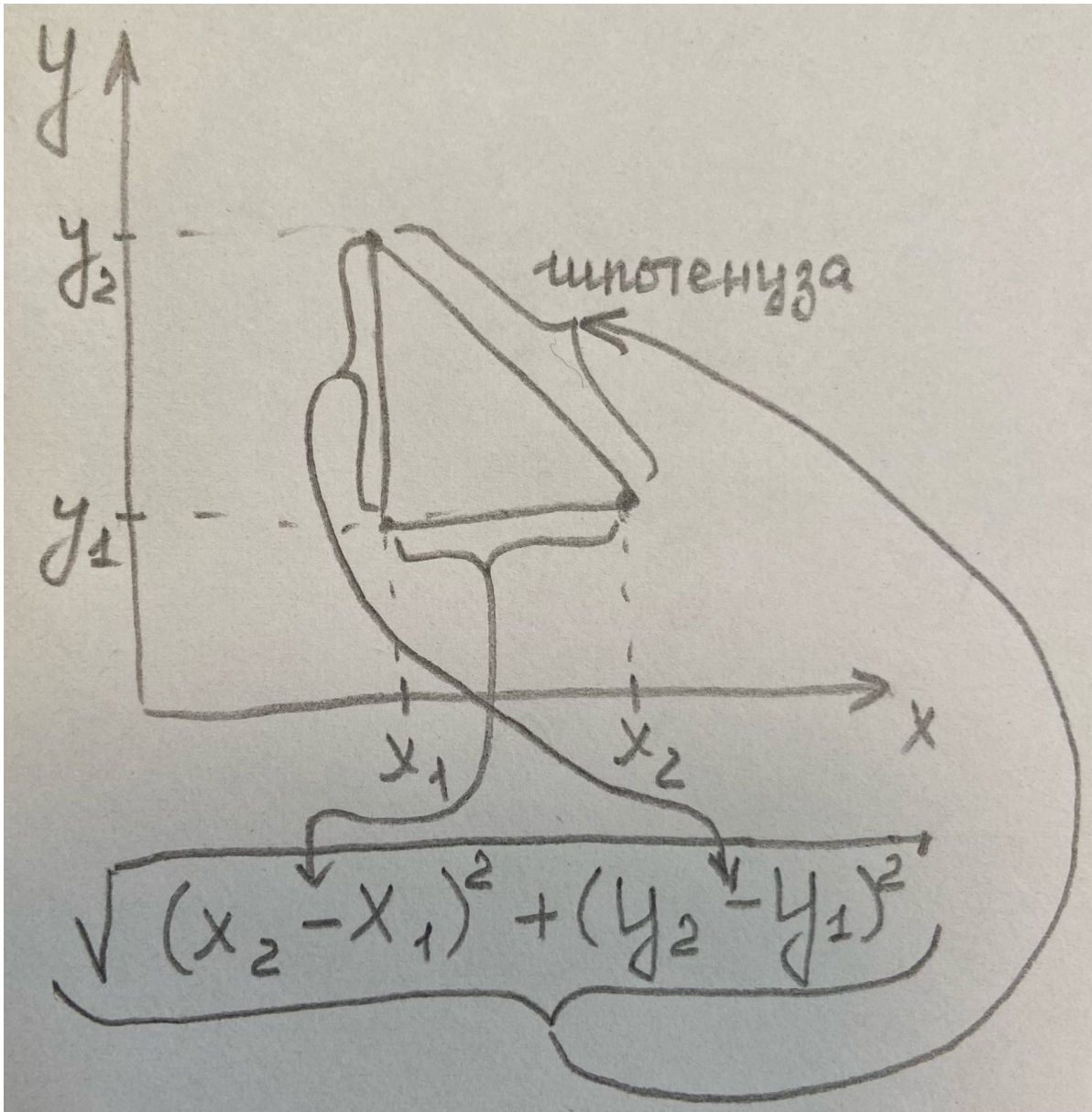
ω_2 : $n_2=9$

ω_3 : $n_3=8$

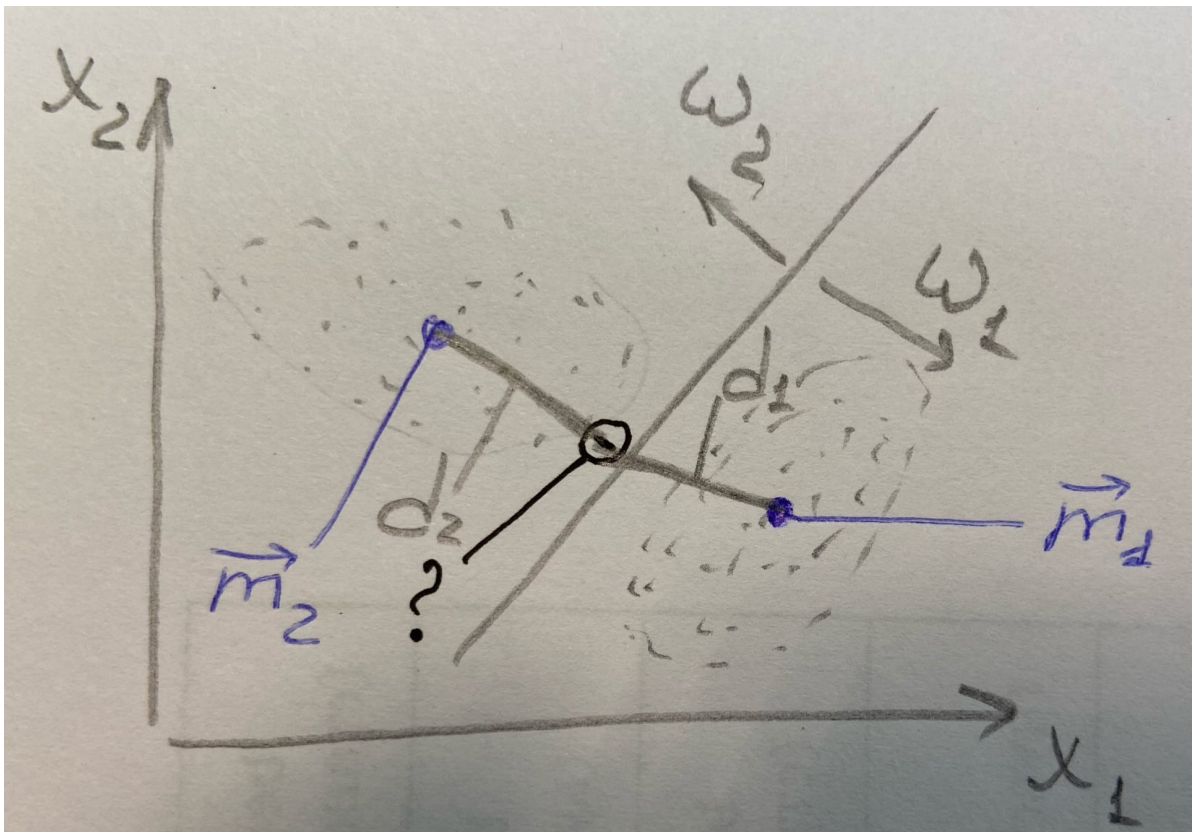
Только по этим точкам мы производим суммирование и делим. Например, у нас упорядоченные вектора, которые относятся к каждому классу, т.е.

$$\underbrace{1-7}_{\omega_1}, \underbrace{8-16}_{\omega_2}, \underbrace{17-25}_{\omega_3}$$

Теорема Пифагора для расстояний



Эталонами не удовлетворимся, т.к. у них есть слабые места:



Если у нас $d_2 > d_1$, и точку мы отнесем по рисунку к ω_2 , то получим ошибку, т.к. она должна быть в классе до центра которого от нее меньшее расстояние, т.е. ω_1 .