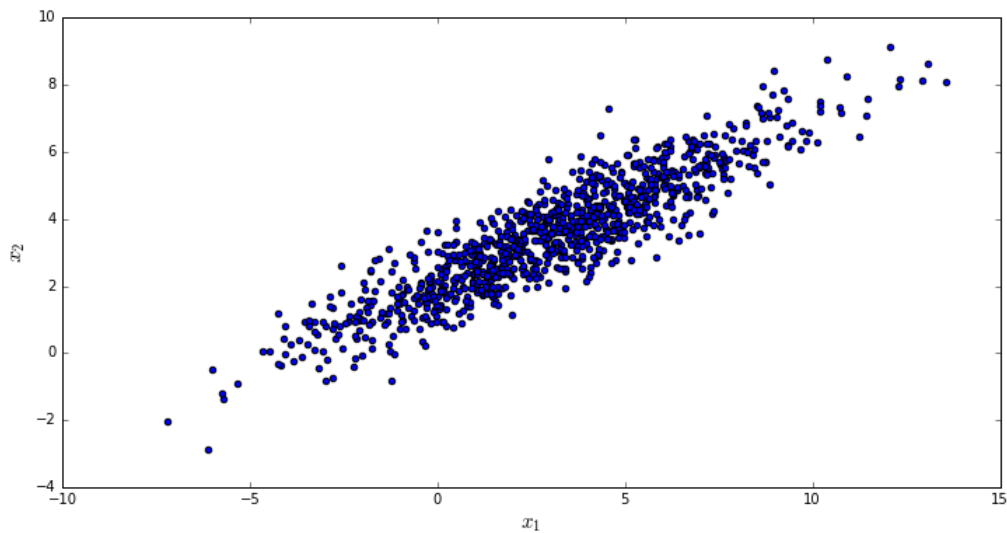


# Примеры задач для экзамена

## 1 Методы понижения размерности

**Задача 1.** У вас имеется следующая двумерная выборка:



Изобразите на рисунке линии, параллельные 2 главным компонентам, получающимся в результате применения метода главных компонент (по одной линии на каждую компоненту).

**Задача 2.** У вас имеется двумерная выборка из 3 объектов:

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

и найденная в результате применения метода главных компонент матрица  $W = [-0.7, -0.7]$ . Найдите признаковое описание приведенных объектов в новом признаковом пространстве.

## 2 Рекомендательные системы

**Задача 3.** Рассмотрим задачу построения рекомендаций фильмов пользователям по имеющимся оценкам. Вы обучили модель со скрытыми переменными, которая предполагает, что предсказание оценки  $\hat{r}_{ui}$  фильму  $i$  пользователем  $u$  строится следующим образом:

$$\hat{r}_{ui} = p_u^T q_i, \quad p_u \in \mathbb{R}^r, \quad q_i \in \mathbb{R}^r,$$

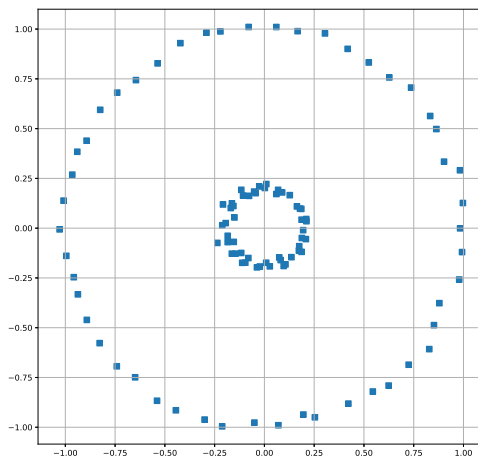
и получили для фильмов векторы

$$\begin{aligned} q_1 &= (4, -2), \\ q_2 &= (-2, 6), \\ q_3 &= (5, -3), \\ q_4 &= (-4, 10). \end{aligned}$$

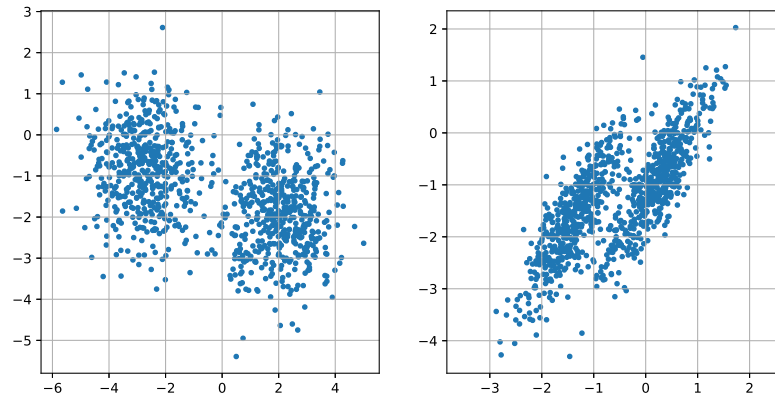
Пусть теперь появился новый пользователь  $u$ , который поставил первому фильму оценку 3, а второму фильму — оценку 1. Найдите скрытый вектор  $p_u$  для этого пользователя, а также предскажите оценки для третьего и четвертого фильмов. Регуляризацию и поправку на сдвиги использовать не нужно.

## 3 Обучение без учителя

**Задача 4.** Обозначьте расположение центроидов и границ кластеров после применения метода  $K$ -Means с  $K = 2$  на следующих данных:



**Задача 5.** Обозначьте расположение центроидов и границ кластеров после применения метода  $K$ -Means с  $K = 2$  на следующих данных:



**Задача 6.** Для данных на рисунке приведите по 3 различных набора параметров  $minPts$  и  $\varepsilon$  метода DBSCAN, при которых:

- все буквы объединятся в один кластер;
- каждая буква будет представлена отдельным кластером;
- все точки будут помечены как выбросы.

