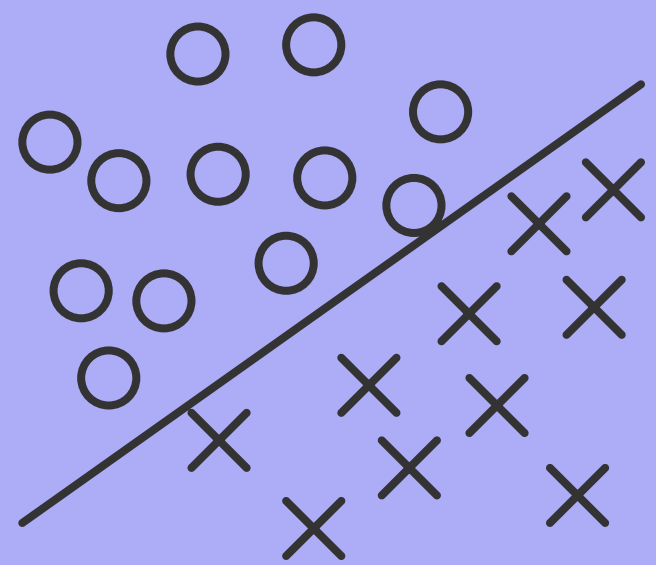


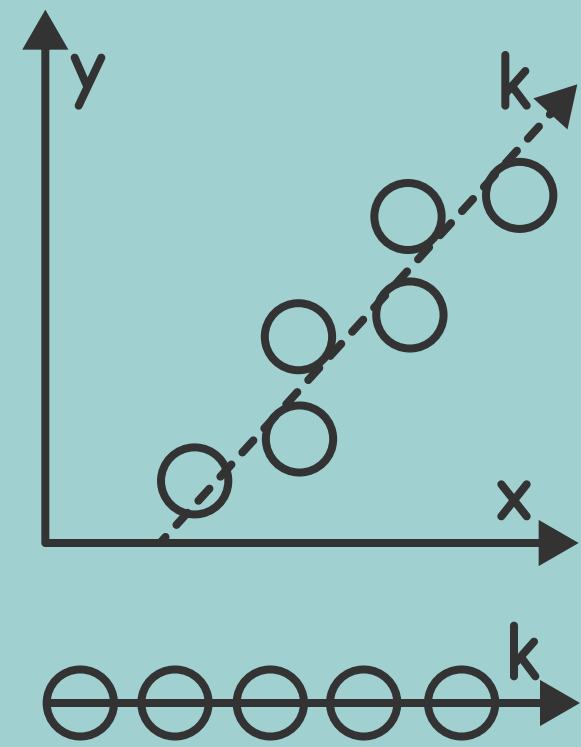
# Линейный дискриминантный анализ

Выполнили:  
Костюхин Алексей  
Тасаков Антон  
Студенты гр. 5030102/10201

# Применение LDA



Классификация объектов в зависимости от признаков



Понижение размерности без применения дополнительных методов

# Условия применимости LDA

1

Объекты каждого класса  
распределены по нормальному  
закону

2

Матрицы ковариации одинаковы  
для всех классов

# Алгоритм

$$S_W = \sum_{k=1}^K \sum_{i \in C_k} (x_i - \mu_k)^T (x_i - \mu_k)$$

Ковариационная матрица внутри  
класса

$$\mu_k = \frac{1}{n_k} \sum_{i \in C_k} x_i$$

Выборочное среднее  
по классам

$$S_W^{-1} S_B v = \lambda$$

Правило поиска  
с. в. и с. ч.

$$S_B = \sum_{k=1}^K \pi_k (\mu_k - \mu)^T (\mu_k - \mu)$$

Ковариационная матрица между  
классами

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

Выборочное среднее

# Алгоритм

$$b = \log \pi_k - \frac{1}{2} \text{diag}(\mu_k w^T)$$

Смещение модели

$$w = \mu_k d_v d_v^T$$

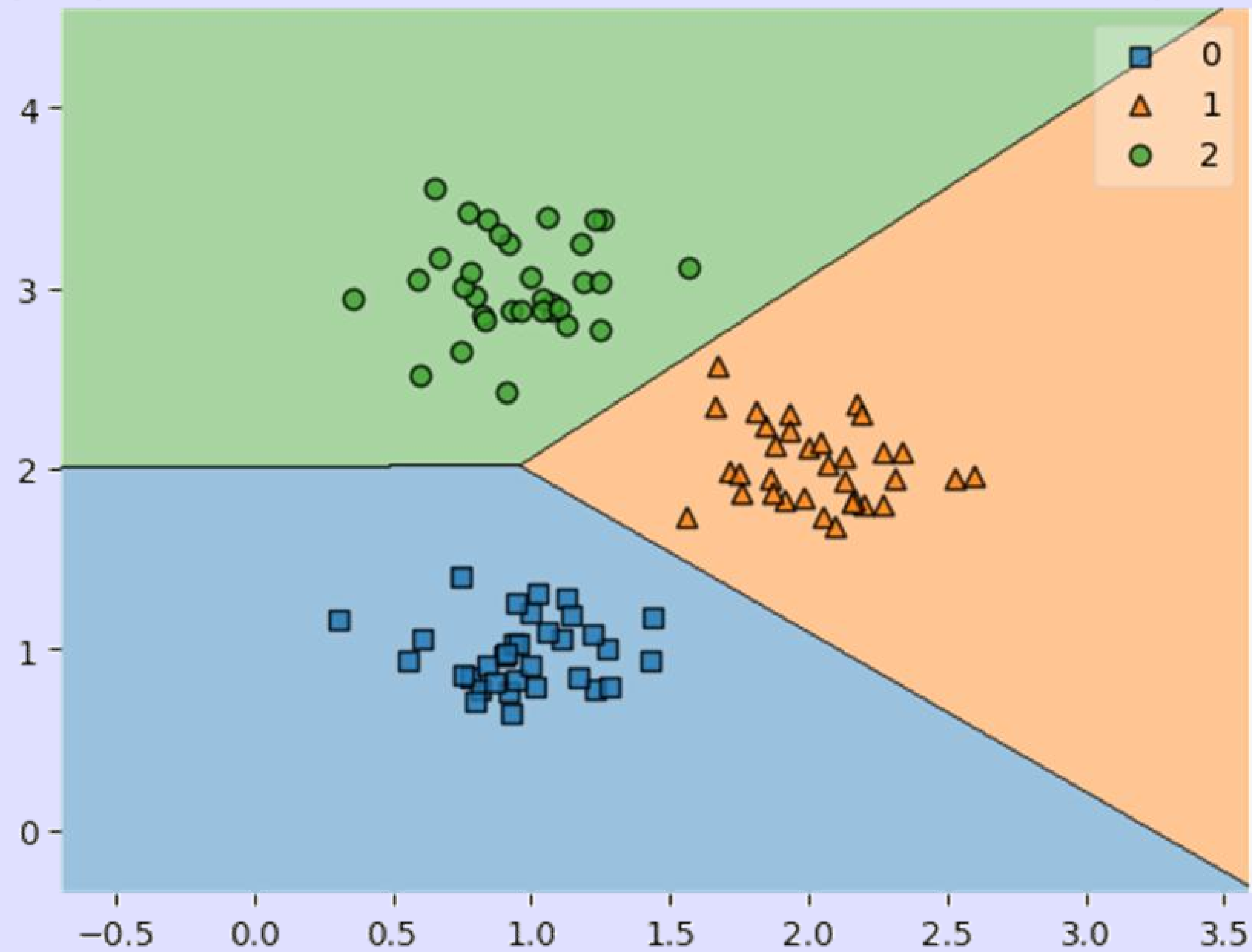
Веса модели

$$\delta(x) = \arg \max_{by\ class} (wx + b)$$

Итоговый прогноз

# Пример

Кластеры, разделённые вычисленными с помощью LDA гиперплоскостями



# Преимущества и недостатки LDA

## Плюсы:

- Простота реализации.
- Быстрая работа на больших наборах данных.
- Хорошо работает при нормальном распределении данных.

## Минусы:

- Ограничен предположением о нормальности данных и равенстве ковариационных матриц.
- Не подходит для нелинейно разделимых данных.
- Чувствительность к выбросам.

# Модификации

```
graph TD; QDA --- KDA; KDA --- FDA; FDA --- QDA;
```

**QDA**

Квадратичный  
дискриминантный  
анализ

**KDA**

Ядерный  
дискриминантный  
анализ

**FDA**

Гибкий  
дискриминантный  
анализ