Анализ свойств моделей в задачах обучения с экспертом

***А. И. Базарова*[1], *А. В. Грабовой*[1], *В. В. Стрижов*[1]**
bazarova.ai@phystech.edu, grabovoy.av@phystech.edu, strijov@phystech.edu
[1]Московский физико-технический институт

В данной работе решается задача поиска заданного набора фигур на изображении в предположении, что фигуры являются кривыми второго порядка. Построение моделей машинного обучения основывается на информации о виде этих кривых и множестве их возможных преобразований. Такую информацию называют *экспертными знаниями*, а метод машинного обучения, основанный на *экспертных знаниях*, называют *обучением с экспертом*. В работе предлагается отобразить кривые второго порядка в новое признаковое пространство, в котором каждая локальная модель является линейной моделью. Таким образом, распознавание кривых высших порядков может быть осуществлено при помощи композиции линейных моделей. В работе поставлена и решена задача оптимизации для поиска оптимальных параметров мультимодели. Качество работы предложенного метода сравнивается на синтетических данных и датасетах с реальными изображениями, которые включают в себя кривые второго порядка.

Ключевые слова: *смесь моделей, обучение с экспертом, линейные модели.*

1 Введение

В работе предлагается метод *обучения с экспертом* [1, 2]. Предметные знания экспертов используются в этом исследовании для повышения качества моделей машинного обучения. Данные предметные знания назовем *экспертными знаниями*. В данной работе в качестве экспертных знаний рассматриается информация о виде распознаваемой кривой и множестве ее допустимых преобразований. Метод машинного обучения, который учитывает *экспертные знания* при построении моделей, назовем *обучением с экспертом*.

Работа сконцентрирована на распознавании изображений кривых второго и третьего порядка: квадрик, коник, кубик и поверхностей второго порядка. Данные фигуры выбраны для анализа, так как они являются простыми для аппроксимации линейными моделями фигурами. При этом эти фигуры требуется восстановить в прикладных задачах, таких как задача распознавания радужки глаза [4], задача описания трека частицы в Большом адронном коллайдере [5, 6]. Коэффициенты уравнений, описывающих данные кривые, аналитически выражаются через оптимальные параметры построенных в работе моделей.

Для каждого класса перечисленных кривых предлагается согласно экспертным данным отобразить точки на изображении в новое признаковое пространство и затем оптимально аппроксимировать кривую. Отображение строится таким образом, чтобы в новом признаковом описании каждая кривая описывалась линейной моделью. В данной работе рассматривается задача поиска фигур на изображении в предположении, что число и тип фигур на изображении заданы, а также заданы дополнительные преобразования фигур и их взаимное расположение. Для каждого типа фигур требуется построить признаковое описание, в котором фигура задается линейно.

При распознавании нескольких кривых на одном изображении строится мультимодель, называемая смесью экспертов. Смесь экспертов — это мультимодель, которая линейно взвешивает локальные модели, аппроксимирующие выборку. Значения весовых коэффициентов зависят от того объекта, для которого производится предсказание.

Примерами мультимоделей являются беггинг, градиентный бустинг [7] и случайный лес решающих деревьев [8]. Подход к мультимоделированию предполагает, что вклад каждого эксперта в ответ зависит от рассматриваемого объекта. Для восстановления этой зависимости смесь экспертов использует шлюзовую функцию. Она определяет значимость предсказания каждого эксперта — отдельной модели, входящей в смесь. Таким образом, ансамбль моделей включает два типа параметров: параметры локальных моделей и параметры шлюзовой функции.

Для оптимизации параметров ансамбля моделей вводится функция ошибки. Она состоит из двух слагаемых: функции регуляризации, вид которой основан на экспертной информации, и суммы квадратичных функций ошибки локальных линейных моделей. Качество мультимодели оценивается с помощью интерпретируемого критерия качества.

Для иллюстрации предложенного подхода решения задач обучения с экспертом поставлен вычислительный эксперимент.

2 Постановка задачи нахождения параметров кривых второго порядка на изображении

Задано бинарное изображение. Эксперт предполагает, что на нем изображен конечный набор кривых второго порядка:

$$\mathbf{M} \in \{0, 1\}^{m_1 \times m_2},$$

где 1 отвечает черной точке изображения, а 0 — белой точке фона. По изображению $\mathbf{M}$ строится выборка $\mathbf{C}$, элементами которой являются координаты (x_i, y_i) черных точек:

$$\mathbf{C} \in \mathbb{R}^{N \times 2}.$$

Построение признакового описания на основе экспертной информации. Пусть для набора точек $\mathbf{C} \in \mathbb{R}^{N\times 2}$, образующих кривую Ω, задано экспертное описание $E(\Omega)$ — набор ограничений на коэффициенты ограничений вида равенство (24), (25), (44) и вида неравенство (21), (36), (38). $E(\Omega)$ состоит из ожидаемого экспертом вида фигуры Ω и множества ее допустимых преобразований. На основе экспертного описания введем отображения

$$K_x\bigl(E(\Omega)\bigr) : \mathbb{R}^2 \to \mathbb{R}^n, \quad K_y\bigl(E(\Omega)\bigr) : \mathbb{R}^2 \to \mathbb{R}, \tag{1}$$

иными словами, поэлементно:

$$K_x\bigl(E(\Omega), \mathbf{c}\bigr) = \mathbf{x}, \quad K_y\bigl(E(\Omega), \mathbf{c}\bigr) = y. \tag{2}$$

Здесь n — число признаков, $\mathbf{c} = (x_i, y_i)$ — точка из выборки $\mathbf{C}$. Рассмотрим линейную модель

$$g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \tag{3}$$

$$g(\mathbf{x}, \mathbf{w}) = \mathbf{x}^\mathsf{T}\mathbf{w}. \tag{4}$$

Вектор $\hat{\mathbf{w}} = [w_0, w_1, \ldots, w_n]$ является решением следующей оптимизационной задачи:

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}\in\mathbb{R}^n} \|g(\mathbf{x}, \mathbf{w}) - y\|, \tag{5}$$

Применяя отображения (2) к исходному набору точек $\mathbf{C}$, получим выборку

$$\mathfrak{D} = \bigl\{(\mathbf{x}, y) \mid \forall \mathbf{c} \in \mathbf{C}\ \mathbf{x} = K_x(\mathbf{c}),\ y = K_y(\mathbf{c})\bigr\}. \tag{6}$$

Таким образом, задача определения коэффициентов уравнения исходной фигуры сводится к решению задачи линейной регрессии, т. е. нахождения компонент вектора $\hat{\mathbf{w}}$, связывающего полученные $\mathbf{x}$ и y.

Мультимодель. В случае, когда на изображении K кривых второго порядка $\Omega_1, \ldots, \Omega_K$, для каждой из которых имеется экспертная информация $E_k = E(\Omega_k)$, $k \in \{1, \ldots, K\}$, ставится задача построения мультимодели, называемой смесью K экспертов.

Definition 1. Назовем мультимодель f смесью K экспертов

$$f = \sum_{k=1}^{K} \pi_k(\mathbf{x}, \mathbf{V}) g_k(\mathbf{w}_k), \quad \pi_k(\mathbf{x}, \mathbf{V}) : \mathbb{R}^{2\times|\mathbf{V}|} \to [0,\, 1], \quad \sum_{k=1}^{K} \pi_k(\mathbf{x}, \mathbf{V}) = 1, \tag{7}$$

где g_k — локальная модель, называемая экспертом, $\mathbf{x}$ — признаковое описание объекта, π_k — шлюзовая функция, вектор $\mathbf{w}_k$ — параметры локальной модели, вектор $\mathbf{V}$ — параметры шлюзовой функции. В данной работе g_k — линейная модель.

Для каждой кривой второго порядка заданы отображения (1). Тогда, используя локальные линейные модели, построим универсальную мультимодель, описывающую все множество кривых $\Omega_1, \ldots, \Omega_K$ на изображении $\mathbf{M}$:

$$f = \sum_{(\mathbf{x}, y) \in \mathfrak{D}} \sum_{k=1}^{K} \pi_k(\mathbf{x}, \mathbf{V}) g_k(\mathbf{x}, \mathbf{w}_k), \tag{8}$$

где π_k — шлюзовая функция:

$$\pi_k(\mathbf{x}, \mathbf{V}) : \mathbb{R}^{2\times|\mathbf{V}|} \to [0,\, 1], \quad \sum_{k=1}^{K} \pi_k(\mathbf{x}, \mathbf{V}) = 1, \tag{9}$$

$\mathbf{V}$ — параметры шлюзовой функции, а g_k — локальная линейная модель вида (4).

В данной работе

$$\boldsymbol{\pi}(\mathbf{x}, \mathbf{V}) = \mathrm{softmax}\bigl(\mathbf{V}_1^{\mathsf{T}} \boldsymbol{\sigma}(\mathbf{V}_2^{\mathsf{T}} \mathbf{x})\bigr), \tag{10}$$

где $\mathbf{V} = \{\mathbf{V}_1,\, \mathbf{V}_2\}$ — параметры шлюзовой функции, $\mathbf{V}_1 \in \mathbb{R}^{p\times k}$, $\mathbf{V}_2 \in \mathbb{R}^{n\times p}$.

Оптимизация. Для нахождения оптимальных параметров мультимодели необходимо решить следующую задачу:

$$\mathcal{L} = \sum_{(\mathbf{x}, y) \in \mathfrak{D}} \sum_{k=1}^{K} \pi_k(\mathbf{x}, \mathbf{V}) (y - \mathbf{w}_k^{\mathsf{T}} \mathbf{x})^2 + R\bigl(\mathbf{V}, \mathbf{W}, E(\Omega)\bigr) \to \min_{\mathbf{V}, \mathbf{W}}. \tag{11}$$

Здесь $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_k]$, $R\bigl(\mathbf{V}, \mathbf{W}, E(\Omega)\bigr)$ — регуляризация параметров, основанная на экспертной информации. Функция R для каждого случая экспертного описания $E(\Omega)$ описана в пунктах (45), (47), (49).

Интерпретируемый критерий качества. Качество мультимодели оценивается с помощью внешнего критерия качества. Так, точность работы алгоритма при распознавании K окружностей на изображении оценивает функция S:

$$S = \sum_{k=1}^{K} (x_0^k - x_{pr}^k)^2 + (y_0^k - y_{pr}^k)^2 + (r_0^k - r_{pr}^k)^2, \tag{12}$$

где x_0^k, y_0^k, r_0^k — истинные значения параметров распознаваемых кривых, x_{pr}^k, y_{pr}^k, r_{pr}^k — предсказанные мультимоделью значения.

3 Построение признакового описания фигур

Окружность. Пусть (x_0, y_0) — центр окружности, которую необходимо найти на бинарном изображении $\mathbf{M}$, а r — ее радиус. Элементы выборки $(x_i, y_i) \in \mathbf{C}$ являются геометрическим местом точек, которое аппроксимируется уравнением окружности:

$$(x_i - x_0)^2 + (y_i - y_0)^2 = r^2. \tag{13}$$

Раскрыв скобки, получим:

$$(2x_0) \cdot x_i + (2y_0) \cdot y_i + (r^2 - x_0^2 - y_0^2) \cdot 1 = x_i^2 + y_i^2. \tag{14}$$

Тогда отображения (2) примут вид:

$$K_x(\mathbf{c}_i) = [x_i,\, y_i,\, 1] = \mathbf{x},\, K_y(\mathbf{c}_i) = x_i^2 + y_i^2 = y. \tag{15}$$

Поставим задачу линейной регрессии (6). Компоненты вектора $\mathbf{w} = [w_0,\, w_1,\, w_2]^\mathsf{T}$, связывающего $\mathbf{x}$ и y, восстанавливают параметры окружности:

$$x_0 = \frac{w_0}{2},\; y_0 = \frac{w_1}{2},\; r = \sqrt{w_3 + x_0^2 + y_0^2}. \tag{16}$$

Эллипс. Элементы выборки $(x_i, y_i) \in \mathbf{C}$ являются ГМТ, которое аппроксимируется уравнением общим уравнением эллипса:

$$Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F = 0,\; B^2 - 4AC < 0. \tag{17}$$

Из условия на коэффициенты A, B, C следует, что $A \neq 0$, $C \neq 0$. Перепишем уравнение:

$$B'x_iy_i + C'y_i^2 + D'x_i + E'y_i + F' = x_i^2. \tag{18}$$

В этом случае отображения (2) имеют вид:

$$K_x(\mathbf{c}_i) = [x_iy_i,\, y_i^2,\, x_i,\, y_i,\, 1] = \mathbf{x},\, K_y(\mathbf{c}_i) = x_i^2 = y. \tag{19}$$

Поставим задачу линейной регрессии (6). Оптимальные параметры линейной регрессии и коэффициенты уравнения связаны следующим образом:

$$w_0 = B' = -\frac{B}{A},\, w_1 = C' = -\frac{C}{A},\, w_2 = D' = -\frac{D}{A},\, w_3 = E' = -\frac{E}{A},\, w_4 = F' = -\frac{F}{A}, \tag{20}$$

с учетом

$$B^2 - 4AC < 0, \tag{21}$$

то есть

$$w_0^2 + 4w_1 < 0. \tag{22}$$

Парабола. Элементы выборки $\mathbf{C}$ аппроксимируются общим уравнением параболы:

$$Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F = 0, \tag{23}$$

где
$$B^2 = 4AC. \tag{24}$$
Рассмотрим различные варианты расположения параболы относительно координатных осей.

Экспертные данные: ось симметрии параболы параллельна Ox. Из экспертных данных следует, что коэффициенты общего уравнения параболы
$$A = B = 0. \tag{25}$$
Тогда общее уравнение параболы, аппроксимирующее выборку $\mathbf{C}$, приобретает вид
$$Cy_i^2 + Dx_i + Ey_i + F = 0. \tag{26}$$
Перепишем:
$$y_i^2 = D'x_i + E'y_i + F'. \tag{27}$$
Тогда вид преобразований (2):
$$K_x(\mathbf{c}_i) = [x_i, y_i, 1] = \mathbf{x},\ K_y(\mathbf{c}_i) = y_i^2 = y. \tag{28}$$
Поставим задачу линейной регрессии (6). Параметры параболы выражаются через оптимальные параметры линейной регрессии:
$$w_0 = D' = -\frac{D}{C},\ w_1 = E' = -\frac{E}{C},\ w_2 = F' = -\frac{F}{C}. \tag{29}$$
Экспертные данные: ось симметрии параболы параллельна Oy. Аналогично предыдущему случаю, общее уравнение приобретет вид:
$$x_i^2 = D'x_i + E'y_i + F'. \tag{30}$$
Преобразования (2) будут иметь вид:
$$K_x(\mathbf{c}_i) = [x_i, y_i, 1] = \mathbf{x},\ K_y(\mathbf{c}_i) = x_i^2 = y. \tag{31}$$
Экспертные данные: ось симметрии параболы не параллельна ни одной из координатных осей. В таком случае $A \neq 0$, $B \neq 0$, $C \neq 0$.
Тогда общее уравнение параболы, аппроксимирующее выборку, имеет вид:
$$x_i^2 = B'x_iy_i + C'y_i^2 + D'x_i + E'y_i + F'. \tag{32}$$
Преобразования (2):
$$K_x(\mathbf{c}_i) = [x_iy_i,\ y_i^2,\ x_i,\ y_i,\ 1] = \mathbf{x},\ K_y(\mathbf{c}_i) = x_i^2 = y. \tag{33}$$
Поставим задачу линейной регрессии (6). Исходные параметры параболы восстанавливаются по вектору оптимальных параметров линейной регрессии $\mathbf{w}$следующим образом:
$$w_0 = B' = -\frac{B}{A},\ w_1 = C' = -\frac{C}{A},\ w_2 = D' = -\frac{D}{A},\ w_3 = E' = -\frac{E}{A},\ w_4 = F' = -\frac{F}{A}. \tag{34}$$
Гипербола. Элементы выборки $\mathbf{C}$ аппроксимируются общим уравнением гиперболы:
$$Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + F = 0, \tag{35}$$

где

$$B^2 - 4AC > 0. \tag{36}$$

Экспертные данные: полуоси гиперболы параллельны осям координат. В таком случае уравнение гиперболы имеет вид:

$$Ax_i^2 + Cy_i^2 + Dx_i + Ey_i + F = 0, \tag{37}$$

где

$$AC < 0. \tag{38}$$

Перепишем уравнение:

$$C'y_i^2 + D'x_i + E'y_i + F' = x_i^2. \tag{39}$$

Преобразования (2):

$$K_x(\mathbf{c}_i) = [y_i^2, x_i, y_i, 1] = \mathbf{x},\ K_y(\mathbf{c}_i) = x_i^2 = y. \tag{40}$$

Экспертные данные: полуоси гиперболы непараллельны осям координат. Тогда $B \neq 0$. Перепишем уравнение гиперболы:

$$A'x_i^2 + C'y_i^2 + D'x_i + E'y_i + F' = x_iy_i. \tag{41}$$

Преобразования (2):

$$K_x(\mathbf{c}_i) = [x_i^2, y_i^2, x_i, y_i, 1] = \mathbf{x},\ K_y(\mathbf{c}_i) = x_iy_i = y. \tag{42}$$

Композиция фигур.

Экспертные данные: на изображении n концентрических окружностей. В задаче оптимизации используются векторы $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$. Так как центры окружностей совпадают, $w_{10} = w_{20} = \cdots = w_{n0}$ и $w_{11} = w_{21} = \cdots = w_{n1}$. Отображения (2) приобретут вид:

$$K_x(\mathbf{c}_i) = [x_i,\ y_i,\ 1],\ K_y(\mathbf{c}_i) = [x_i^2 + y_i^2]. \tag{43}$$

Ограничения на w_{i0} и w_{i1} учитываются согласно экспертной информации подстановкой одинаковых значений w_0 и w_1 во все векторы $\mathbf{w}_k$.

Экспертные данные: на изображении две окружности, известно расстояние между центрами. В этом случае расстояние

$$\sqrt{(w_{10} - w_{20})^2 + (w_{11} - w_{21})^2} = \rho \tag{44}$$

задано экспертом. K_x и K_y имеют вид (15). Функция регуляризации R:

$$R = \frac{\mu}{2}\left(\sqrt{(w_{10} - w_{20})^2 + (w_{11} - w_{21})^2} - \rho\right)^2. \tag{45}$$

Экспертные данные: на изображении три окружности, известны соотношения между сторонами треугольника центров. Экспертная информация: треугольник равносторонний. Тогда

$$\sqrt{(w_{10} - w_{20})^2 + (w_{11} - w_{21})^2} = \sqrt{(w_{10} - w_{30})^2 + (w_{11} - w_{31})^2} = \sqrt{(w_{20} - w_{30})^2 + (w_{21} - w_{31})^2}. \tag{46}$$

Тогда в качестве функции регуляризации R:

$$R = \frac{\mu}{2} \sum_{\substack{1 \leqslant i,j \leqslant 3 \\ i<j}} (\sqrt{(w_{i0} - w_{j0})^2 + (w_{i1} - w_{j1})^2} - \sqrt{(w_{i0} - w_{j0})^2 + (w_{i1} - w_{j1})^2})^2. \quad (47)$$

Экспертные данные: на изображении n концентрических эллипсов. Координаты центра симметрии эллипса выражаются из коэффициентов его общего уравнения следующим образом:

$$y_0 = \frac{4\frac{E}{A} - 2\frac{DB}{A^2}}{\frac{B^2}{A^2} - 4\frac{C}{A}} = \frac{-4w_3 - 2w_0w_2}{w_0^2 + 4w_1},\ x_0 = -\frac{D}{A} - \frac{B}{A}y_0 = w_2 - w_0 \cdot \frac{4w_3 + 2w_0w_2}{w_0^2 + 4w_1}. \quad (48)$$

Экспертные данные задают одинаковые значения x_0 и y_0 для всех эллипсов на изображении. Функцию регуляризации R в таком случае выберем:

$$\begin{aligned} R = \frac{\mu}{2} \sum_{\substack{1 \leqslant i,j \leqslant n \\ i<j}} \left(\frac{-4w_{i3} - 2w_{i0}w_{i2}}{w_{i0}^2 + 4w_{i1}} - \frac{-4w_{j3} - 2w_{j0}w_{j2}}{w_{j0}^2 + 4w_{j1}} \right)^2 + \\ + \left(\left(w_{i2} - w_{i0} \cdot \frac{4w_{i3} + 2w_{i0}w_{i2}}{w_{i0}^2 + 4w_{i1}} \right) - \left(w_{j2} - w_{j0} \cdot \frac{4w_{j3} + 2w_{j0}w_{j2}}{w_{j0}^2 + 4w_{j1}} \right) \right)^2. \end{aligned} \quad (49)$$

4 Вычислительный эксперимент

Проведен вычислительный эксперимент для анализа качества моделей кривых второго порядка на изображении. В эксперименте использован подход с заданием априорных распределений, указывающих на типы кривых на изображении

Синтетические данные с разным типом шума на изображении. Для проведения эксперимента на синтетических данных сгенерированы 2 типа выборок (точки принадлежащие соотвествующих кривым): выборка Synthetic 1 — выборка без шума, Synthetic 2 — выборка с шумом вблизи кривых.

Для сравнения качества моделей с разными априорными распределениями использована интерпретируемая функция ошибки S:

$$S = \sum_{k=1}^{K} \|\mathbf{w}_k - \mathbf{w}_k^0\|^2, \quad (50)$$

где $\mathbf{w}_k^0$ — истинные значения параметров распознаваемых кривых, $\mathbf{w}_k$ — предсказанные мультимоделью значения.

Анализ моделей в зависимости от разных априорных предположений. На рис.1 результат работы мультимоделей $\mathfrak{M}_1$ (с заданием априорного распределения) и $\mathfrak{M}_2$ (без задания). На всех изображениях обе модели обучались 50 итераций EM-алгоритма. Качество прогноза, вычисленное по формуле 50, в таблице 1:

	$S_{\mathfrak{M}_1}$	$S_{\mathfrak{M}_2}$
Synthetic1	10^{-5}	10^{-4}
Synthetic2	10^{-3}	0.13

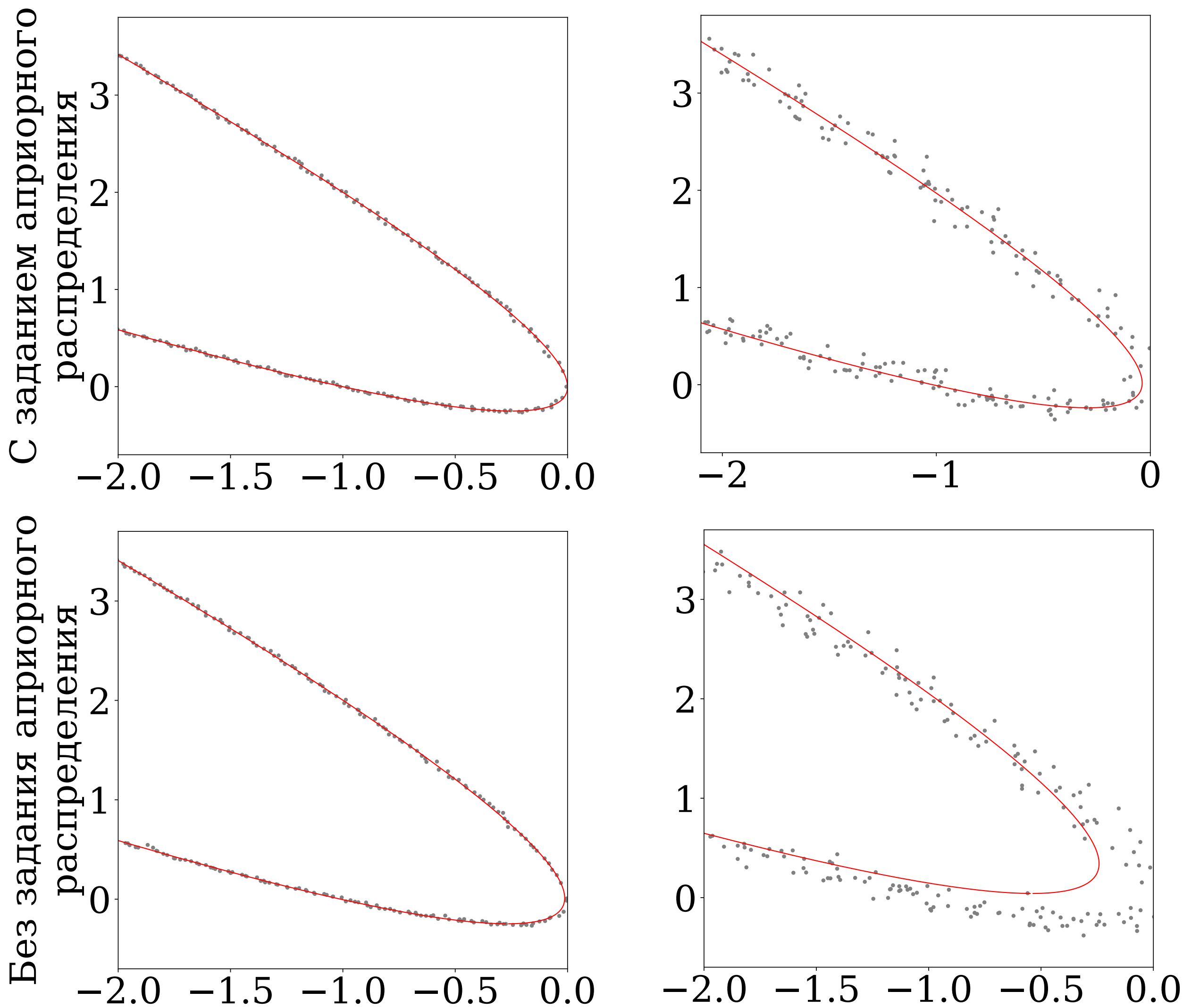

Рис. 1 Мультимодель в зависимости от разных априорных предположений и уровня шума.

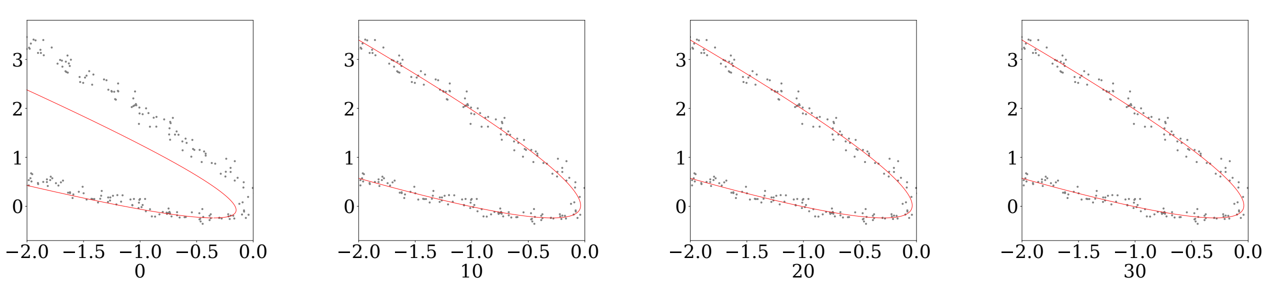

Рис. 2 Визуализация процесса обучения мультимодели в течение 30 итераций

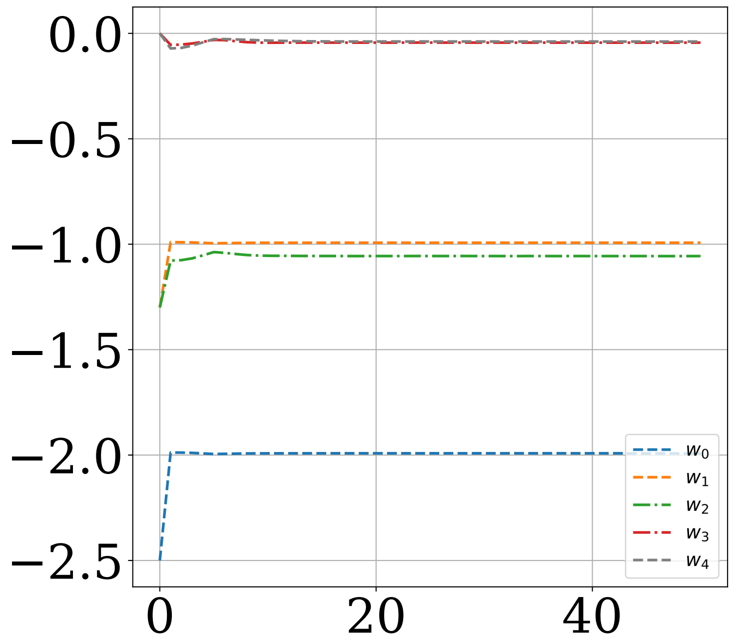

Рис. 3 График зависимости w_i от номера итерации

На рис. 2 показан процесс обучения мультимодели $\mathfrak{M}_1$ на выборке Synthetic 2 в течение 30 итераций.

На рис. 3 изображены графики зависимости коэффициентов w_i от номера итерации при работе модели $\mathfrak{M}_2$ на выборке Synthetic 2.

В ходе эксперимента показано, что задание априорного распределения улучшает качество распознавания изображений.

Анализ мультимодели в зависимости от шума. Для анализа свойств мультимоделей $\mathfrak{M}_1$ и $\mathfrak{M}_2$ в зависимости от зашумленности изображения проведен вычислительный эксперимент на выборках Synthetic 3, Synthetic 4 и Synthetic 5. Минимальный уровень шума равен 0, когда на изображении нет шумовых точек, а максимальный равен 0.15, когда на изображении число шумовых точек равно $\frac{1}{6}$ от числа точек всех окружностей. Работа мультимодели на данных выборках с заданным априорным распределением и без задания априорного распределения показана на рис. 4.

На рис. 5 показан график зависимости радиуса r и центра (x_0, y_0) от номера итерации для каждой окружности. Нетрудно видеть, что модель $\mathfrak{M}_1$ с заданием априорного распределения сходится быстрее, чем модель $\mathfrak{M}_2$ без задания, и является более устойчивой к шуму. Качество работы обеих моделей снижается при работе на равномерно зашумленных данных. При этом $\mathfrak{M}_1$ разделяет точки исходных окружностей точнее, чем $\mathfrak{M}_2$, однако из-за наличия шумовых точек параметры окружностей, найденные $\mathfrak{M}_1$, не совпадают с истинными параметрами.

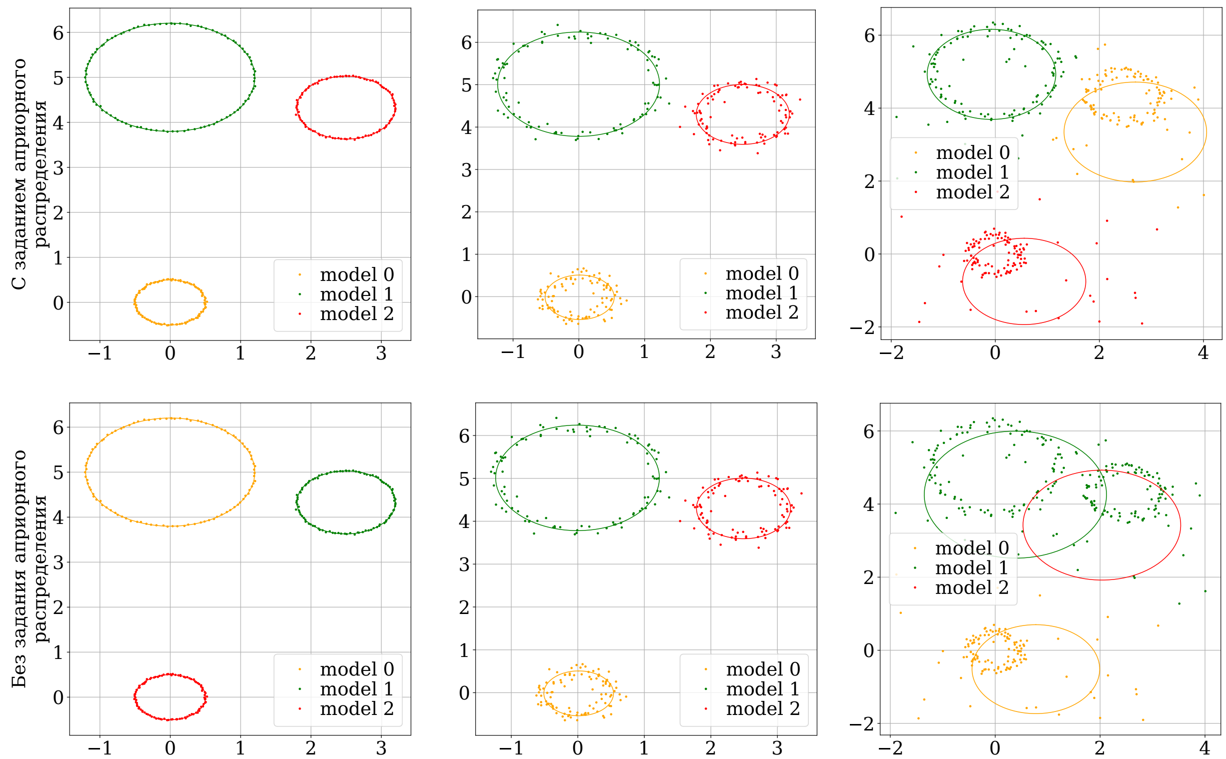

Рис. 4 Мультимодель в зависимости от разных априорных предположений и уровня шума

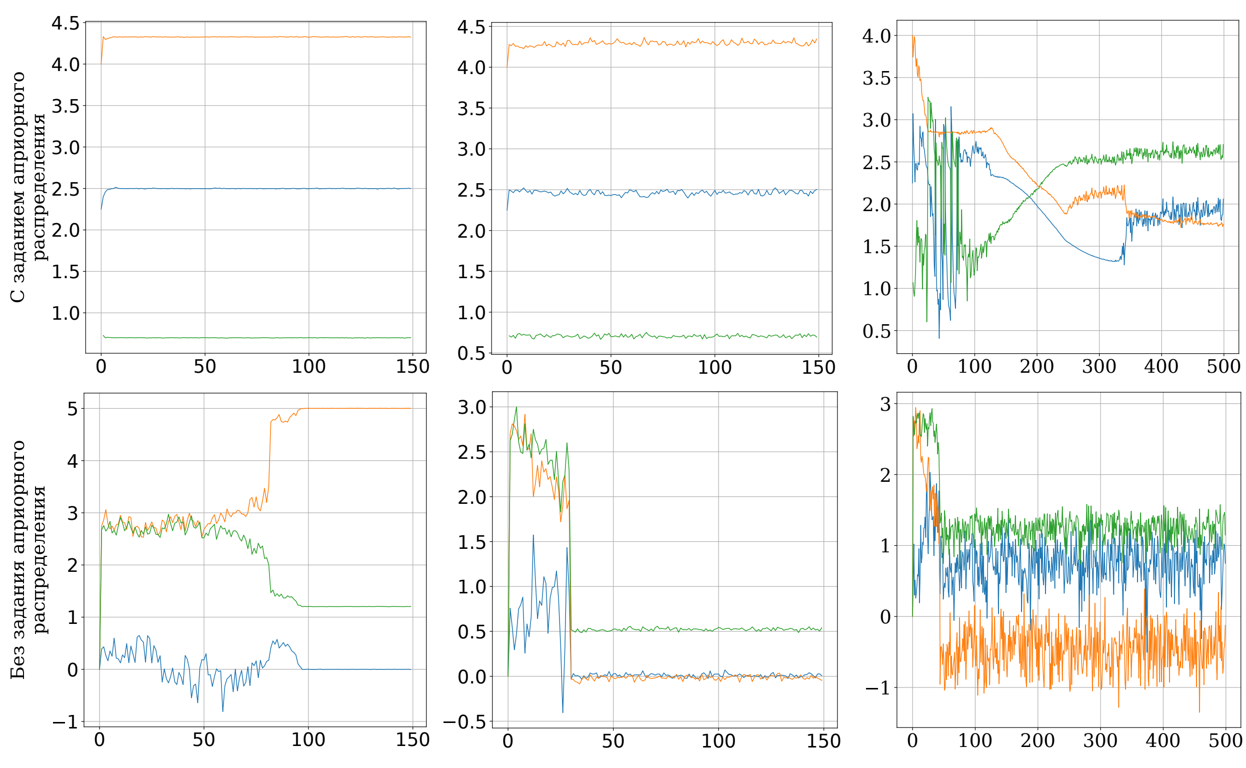

Рис. 5 Зависимость параметров r, x_0 и y_0 от номера итерации при разных априорных распределениях

Данный эксперимент демонстрирует, что модель $\mathfrak{M}_1$ с заданием априорного распределения более устойчива к шуму. $\mathfrak{U}$

5 Список литературы

[1] *Graboviy A. V., Strijov V. V.* Analysis of prior distributions for a mixture of experts // Computational Mathematics and Mathematical Physics, to appear in 2020.

[2] *Scheres S. H. W.* A Bayesian view on Cryo-EM structure determination. // Journal of Molecular Biology. 2012. Vol. 415. № 2. P. 406–418.

[3] *Yuksel S. E., Wilson J. N., Gader P. D.* Twenty years of mixture of experts // IEEE Transactions on Neural Networks and Learning Systems. 2012. Vol. 23, № 8, P. 1177–1193.

[4] *Matveev I. A.* Detection of iris in image by interrelated maxima of brightness gradient projections // Applied and Computational Mathematics. 2010. Vol. 9. № 2. P. 252–257.