

Релаксационный подход в задаче структурного обучения по слаборазмеченным данным

Дмитрий Кондрашкин,
научный руководитель: к.ф.-м.н. Ветров Д. П.

2 июня 2014 г.

Пример: задача семантической сегментации

Определить, к какому классу принадлежит каждый пиксель изображения.

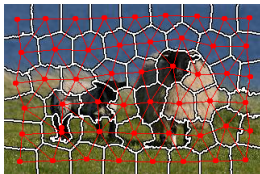
Обучающая выборка $\{(\mathbf{x}^j, \mathbf{y}^j)\}_{j=1}^J$ состоит из признаков \mathbf{x} и пиксельной разметки \mathbf{y} :



Отдельные переменные y_i не являются независимыми.

Марковское случайное поле

Пусть задан граф $G = (V, E)$, $\mathbf{y} \in \{1, \dots, K\}^{|V|}$ — разметка.



Зададим линейную параметризацию энергии:

$$\begin{aligned} E(\mathbf{x}, \mathbf{y}, \mathbf{w}) &= - \sum_{v \in V} \mathbf{w}^\top \psi_v(y_v, \mathbf{x}_v) - \sum_{(v,u) \in E} \mathbf{w}^\top \psi_{vu}(y_v, y_u, \mathbf{x}_{vu}) \\ &= -\mathbf{w}^\top \Psi(\mathbf{y}, \mathbf{x}). \end{aligned}$$

Задача вывода $\mathbf{y}^* = \operatorname{argmin}_{\mathbf{y}} E(\mathbf{y})$.

Структурный метод опорных векторов (SSVM)

Оптимизационная задача:

$$\min_{\mathbf{w}} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{j=1}^J \left[\max_{\bar{\mathbf{y}} \in \mathcal{Y}} \{ \mathbf{w}^T \Psi(\bar{\mathbf{y}}, \mathbf{x}^j) + \Delta(\bar{\mathbf{y}}, \mathbf{y}^j) \} - \mathbf{w}^T \Psi(\mathbf{y}^j, \mathbf{x}^j) \right].$$

Задача $\max_{\bar{\mathbf{y}}} \{ \mathbf{w}^T \Psi(\bar{\mathbf{y}}, \mathbf{x}^j) + \Delta(\bar{\mathbf{y}}, \mathbf{y}^j) \}$ называется *выводом, дополненной функцией потерь*.

В качестве функции потерь часто используется расстояние Хэмминга $\Delta(\bar{\mathbf{y}}, \mathbf{y}) = \sum_{v \in \mathbf{v}} [\bar{y}_v \neq y_v]$.

Слабая аннотация

Можно дополнить обучающую выборку объектами со слабой аннотацией $\{(\mathbf{x}^i, \mathbf{z}^i)\}_{i=1}^I$.



sky tree plane grass

Набор данных PASCAL VOC 2012:

Всего	Полная разметка	Список меток классов
11540	2913 (25%)	11540 (100%)

В работе в качестве слабой аннотации рассматривается список меток классов, присутствующих на изображении.

Структурное обучение с учетом слабой аннотации

Модифицируем задачу SSVM:

$$\min_{\mathbf{w}} \frac{1}{2} \mathbf{w}^\top \mathbf{w} + C \sum_{j=1}^J \left[\max_{\bar{\mathbf{y}} \in \mathcal{Y}} \{ \mathbf{w}^\top \Psi(\bar{\mathbf{y}}, \mathbf{x}^j) + \Delta(\bar{\mathbf{y}}, \mathbf{y}^j) \} - \mathbf{w}^\top \Psi(\mathbf{y}^j, \mathbf{x}^j) \right] \\ + \alpha C \sum_{i=1}^I \left[\max_{\bar{\mathbf{y}}} \{ \mathbf{w}^\top \Psi(\bar{\mathbf{y}}, \mathbf{x}^i) + K(\bar{\mathbf{y}}, \mathbf{z}^i) \} - \max_{\mathbf{y}: \mathbf{y} \in L(\mathbf{z}^i)} \mathbf{w}^\top \Psi(\mathbf{y}, \mathbf{x}^i) \right].$$

- ▶ $L(\mathbf{z})$ — множество разметок, совместных со слабой аннотацией,
- ▶ $K(\bar{\mathbf{y}}, \mathbf{z})$ — слабая функция потерь.

Функция потерь

В качестве функции потерь используется расстояние Хэмминга:

$$\Delta(\bar{\mathbf{y}}, \mathbf{y}) = \sum_{v \in \mathbf{V}} [\bar{y}_v \neq y_v].$$

Слабая функции потерь (Шаповалов 2012)

$$K(\bar{\mathbf{y}}, \mathbf{z}) = \sum_{k \notin \mathbf{z}} \sum_{v \in V} c_v [\bar{y}_i = k] + \sum_{k \in \mathbf{z}} c'_k \prod_{v \in V} [\bar{y}_v \neq k]$$

является оценкой для расстояния Хэмминга.

Структурное обучение с учетом слабой аннотации

Задача SSVM, учитывающего слабоаннотированные объекты:

$$\min_{\mathbf{w}} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{j=1}^J \left[\max_{\bar{\mathbf{y}} \in \mathcal{Y}} \{ \mathbf{w}^T \Psi(\bar{\mathbf{y}}, \mathbf{x}^j) + \Delta(\bar{\mathbf{y}}, \mathbf{y}^j) \} - \mathbf{w}^T \Psi(\mathbf{y}^j, \mathbf{x}^j) \right] \\ + \alpha C \sum_{i=1}^I \left[\max_{\bar{\mathbf{y}}} \{ \mathbf{w}^T \Psi(\bar{\mathbf{y}}, \mathbf{x}^i) + K(\bar{\mathbf{y}}, \mathbf{z}^i) \} - \max_{\mathbf{y}: \mathbf{y} \in L(\mathbf{z}^i)} \mathbf{w}^T \Psi(\mathbf{y}, \mathbf{x}^i) \right].$$

Цель работы

Цель работы: исследовать поведение верхней оценки целевой функции обобщенного SSVM.

Задачи:

1. Построить верхнюю оценку целевой функцию структурного SVM, использующего слабую аннотацию.
2. Разработать и реализовать алгоритм для минимизации полученной верхней оценки.
3. Провести экспериментальное сравнение с существующими методами.

При минимизации верхней оценки мы можем быть уверены в том, что оптимальное значение целевой функции ограничено некоторой величиной.

Вложенные оптимизационные задачи

В функционале обобщенного SSVM возникают три вложенных задачи:

Оракул	Оценка	Метод
$\operatorname{argmax}_{\bar{y}} \{ \mathbf{w}^\top \Psi(\bar{y}, \mathbf{x}) + \Delta(\bar{y}, \mathbf{y}) \}$	Сверху	Релаксационный
$\operatorname{argmax}_{\bar{y}} \{ \mathbf{w}^\top \Psi(\bar{y}, \mathbf{x}) + K(\bar{y}, \mathbf{z}) \}$	Сверху	Релаксационный
$\operatorname{argmax}_{\mathbf{y}: \mathbf{y} \in L(\mathbf{z})} \mathbf{w}^\top \Psi(\mathbf{y}, \mathbf{x})$	Снизу	α -расширение

- ▶ Задачи являются NP-трудными, поэтому мы используем приближенные методы для их решения.
- ▶ *Релаксационный подход* — подход к решению оптимизационных задач, при котором допустимое множество значений переменных расширяется.

Вывод, дополненный функцией потерь

Воспользуемся двойственным разложением и декомпозицией на деревья (Komodakis 2007):

$$\max_{\bar{\mathbf{y}}} \Phi(\bar{\mathbf{y}}, \mathbf{x}_j, \mathbf{y}_j, \mathbf{w}) \leq \min_{\lambda_j: \sum_{\tau} \lambda_j^{\tau} = 0} \sum_{\tau \in \mathcal{T}} \max_{\bar{\mathbf{y}}^{\tau}} \Phi(\bar{\mathbf{y}}^{\tau}, \mathbf{x}_j^{\tau}, \mathbf{y}_j^{\tau}, \mathbf{w} + \lambda_j^{\tau}),$$

где

$$\Phi(\bar{\mathbf{y}}, \mathbf{x}_j, \mathbf{y}_j, \mathbf{w}) = \mathbf{w}^{\top} \Psi(\bar{\mathbf{y}}, \mathbf{x}_j) + \Delta(\bar{\mathbf{y}}, \mathbf{y}_j).$$

Вывод, дополненный слабой функцией потерь

Воспользуемся двойственным разложением:

$$\begin{aligned} & \max_{\bar{\mathbf{y}}} \{ \mathbf{w}^\top \Psi(\bar{\mathbf{y}}, \mathbf{x}) + K(\bar{\mathbf{y}}, \mathbf{z}) \} \leq \\ & \min_{\mu} \left\{ \max_{\bar{\mathbf{y}}} (\mathbf{w}^\top \Psi(\bar{\mathbf{y}}, \mathbf{x}) - \mu^\top \bar{\mathbf{y}}) + \max_{\bar{\mathbf{y}}} (K(\bar{\mathbf{y}}, \mathbf{z}) + \mu^\top \bar{\mathbf{y}}) \right\}. \end{aligned}$$

Первое слагаемое ограничим сверху с помощью декомпозиции на деревья.

Для того чтобы получить верхнюю оценку на второе слагаемое, надо знать точный вид функции $K(\bar{\mathbf{y}}, \mathbf{z})$.

Вывод, дополненный слабой функцией потерь

Если использовать слабую функцию потерь

$$K(\bar{\mathbf{y}}, \mathbf{z}) = \sum_{k \notin \mathbf{z}} \sum_{v \in V} c_v [\bar{y}_i = k] + \sum_{k \in \mathbf{z}} c'_k \prod_{v \in V} [\bar{y}_v \neq k],$$

то задачу, возникающую во втором слагаемом, можно решить точно с помощью небольшого перебора:

$$\begin{aligned} & \max_{\bar{\mathbf{y}}} (K(\bar{\mathbf{y}}, \mathbf{z}) + \boldsymbol{\mu}^\top \bar{\mathbf{y}}) = \\ & \max_{l \in \mathcal{P}(\mathbf{z})} \left(\max_{\bar{\mathbf{y}} \in \mathcal{Y} \setminus L(l)} (\mathbf{u} + \boldsymbol{\mu})^\top \bar{\mathbf{y}} + \sum_{k \in l} c'_k \right), \end{aligned}$$

где $\mathcal{P}(\mathbf{z})$ — множество всех подмножеств множества \mathbf{z} .

Метод оптимизации

Полученная оценка недифференцируема, является суммой выпуклой и вогнутой функций.

Для ее минимизации воспользуемся блочно-координатным спуском по группам переменных:

- ▶ w — субградиентный спуск,
- ▶ λ — метод проекции субградиента,
- ▶ μ — субградиентный спуск.

Для возникающих подзадач:

- ▶ Вывод, согласованный с аннотацией — алгоритм α -расширения.
- ▶ Вывод на деревьях — метод передачи сообщений.

Описание наборов данных

Используются два набора данных:

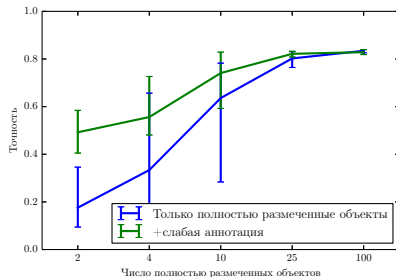
- ▶ Синтетический набор данных, 10 классов, 100 объектов в обучающей выборке, 400 — в тестовой.
- ▶ Набор данных MSRC (реальные данные), 23 класса, 276 объектов в обучающей выборке, 256 — в тестовой.

Исследуются два подхода:

- ▶ Стандартный: в качестве всех оракулов используется алгоритм α -расширения.
- ▶ Релаксационный подход, основанный на минимизации верхней оценки.

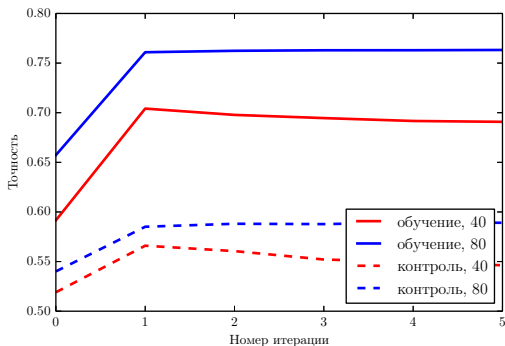
Синтетический набор данных, стандартный подход

Зависимость точности от числа полностью размеченных объектов, $C = 10, \alpha = 0.1$.



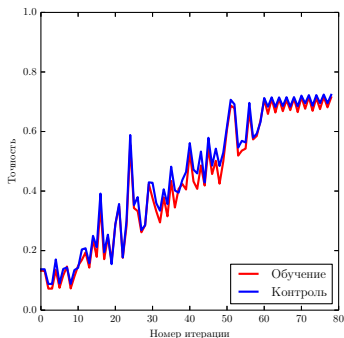
Набор данных MSRC, стандартный подход

Графики точности в зависимости от номера итерации при
разном числе полностью размеченных объектов,
 $C = 100, \alpha = 0.1$.



Синтетический набор данных, релаксационный подход

График точности в зависимости от номера итерации,
синтетические данные, 10 полностью размеченных объектов,
 $C = 100, \alpha = 0.1$.



Сравнение методов без учета слабой аннотации

	Релаксационный подход	Подход с α -расширением
Обучение	0.7979	0.8192
Контроль	0.7832	0.8122

Синтетический набор данных.

	Релаксационный подход	Подход с α -расширением
Обучение	0.5164	0.6518
Контроль	0.5671	0.8430

Набор данных MSRC.

Сравнение методов с учетом слабой аннотации

	Релаксационный	Релаксационный + α	$\alpha + \alpha$
Обучение	0.7144	0.7423	0.8395
Контроль	0.7232	0.7595	0.8254

Синтетический набор данных.

- ▶ Релаксационный + α : для задачи вывода, дополненного слабой функцией потерь, используется алгоритм α -расширения.
- ▶ $\alpha + \alpha$: для решения всех подзадач используется алгоритм α -расширения.

Выводы

Одной из причин падения точности является то, что метод α -расширения часто находит решение близкое к оптимальному. А при использовании двойственного разложения для задачи вывода, дополненного слабой функцией потерь, получается большой зазор и слишком неплотная верхняя оценка.

Результаты работы

- ▶ Построена верхняя оценка на функционал обобщенного структурного метода опорных векторов.
- ▶ Предложен алгоритм для ее минимизации.
- ▶ Результаты экспериментов, которые показали, что метод, основанный на релаксационном подходе, несмотря на свою теоретическую обоснованность, по точности проигрывает методу, где для решения всех подзадач используется алгоритм α -расширения.