

Теоретическое задание 2

Кириленко Елена

19 ноября 2018 г.

1 Задача 1

1)

Рассматривается модель смеси гауссовских распределений

$$p(x) = w_1 N(x|\mu_1, \sigma_1^2) + w_2 N(x|\mu_2, \sigma_2^2).$$

В этой модели много параметров, а именно - $\theta = (\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, w_1, w_2)$. Хотя на самом деле можно избавиться от параметра w_2 так как для того, что $p(x)$ было плотностью необходимо, чтобы $w_2 = 1 - w_1$.

Рассмотрим правдоподобие данного распределения, построенное по выборке $X = (X_1, X_2, \dots, X_n)$.

$$\begin{aligned} L_X(\theta) &= \prod_{i=1}^n (w_1 p_{\mu_1, \sigma_1^2}(X_i) + w_2 p_{\mu_2, \sigma_2^2}(X_i)) = \\ &= w_1 p_{\mu_1, \sigma_1^2}(X_1) w_2^{n-1} \prod_{i=2}^n p_{\mu_2, \sigma_2^2}(X_i) + \underbrace{\dots\dots\dots}_{\text{неотрицательные слагаемые}} \end{aligned}$$

В предыдущем выражении через $p_{\mu_j, \sigma_j^2}(X_i)$ я обозначала плотность нормального с параметрами μ_j, σ_j^2 в точке X_i . (такое обозначение больше похоже на плотность, чем $N(x|\mu_j, \sigma_j^2)$).

Давайте теперь возьмем $\mu_1 = X_1, w_1 = w_2 = \frac{1}{2}$, а также зафиксируем μ_2, σ_2^2 . Тогда рассмотрим правдоподобие $L_X(\theta)$ при $\sigma_1 \rightarrow 0$.

Будем рассматривать только первое слагаемое ранее выписанной суммы с троеточием.

При стремлении σ_1 к нулю $p_{\mu_1, \sigma_1^2}(X_1)$ будет стремиться к бесконечности.

Также рассмотрим $p_{\mu_2, \sigma_2^2}(X_i)$ при $i \geq 2$. При фиксированных μ_2, σ_2^2 эта штука является ограниченной снизу. Возьмем, например, самый далекий по модулю от μ_2 элемент выборки: $X_k : |X_k - \mu_2| = \max_j (|X_j - \mu_2|)$. В точке X_k $p_{\mu_2, \sigma_2^2}(X_k)$ является самой маленькой среди всех элементов выборки. Тогда давайте обозначим $A = p_{\mu_2, \sigma_2^2}(X_k)$. Тогда получим, что $\forall i \ p_{\mu_2, \sigma_2^2}(X_i) \geq A$.

Отсюда получим, что $\prod_{i=2}^n p_{\mu_2, \sigma_2^2}(X_i) \geq A^{n-1}$.

В итоге отсюда получаем, что все первое слагаемое стремиться к бесконечности:

$$\begin{aligned} w_1 p_{\mu_1, \sigma_1^2}(X_1) w_2^{n-1} \prod_{i=2}^n p_{\mu_2, \sigma_2^2}(X_i) &= \left(\frac{1}{2}\right)^n p_{\mu_1, \sigma_1^2}(X_1) \prod_{i=2}^n p_{\mu_2, \sigma_2^2}(X_i) \geq \\ &\geq \left(\frac{1}{2}\right)^n A^{n-1} p_{\mu_1, \sigma_1^2}(X_1) \rightarrow \infty \end{aligned}$$

. Осталось вспомнить, что многочлен - неотрицательные слагаемые и понять, что тогда все правдоподобие стремиться к бесконечности при $\sigma_1 \rightarrow 0$. 2) Как я говорила у пункте 1 : мы можем избавиться от одного параметра w_2 и переписать нашу плотность в более простом виде. Решение от этого не измениться, то так будет чуть проще и понятней. Тогда в нашем случае $= (\mu_1, \mu_2, w)$. Итак :

$$p(x) = w p_{\mu_1, \sigma_1^2}(x) + (1 - w) p_{\mu_2, \sigma_2^2}(x)$$

Введем фиктивные переменные - δ_i , которая равна 1 если объект x_i взят из первой гауссианы, и $= 0$, если из второй.

Тогда распишем правдоподобие используя эту переменную. А лучше сразу лог-правдоподобие, чтобы избавиться от произведения.

$$\begin{aligned} l_X(\theta) &= \sum_{i=1}^n \log[w p_{\mu_1, \sigma_1^2}(X_i) + (1 - w) p_{\mu_2, \sigma_2^2}(X_i)] = \\ &= \sum_{i=1}^n \log[\delta_i w p_{\mu_1, \sigma_1^2}(X_i) + (1 - \delta_i) (1 - w) p_{\mu_2, \sigma_2^2}(X_i)] = \\ &= \sum_{i=1}^n \delta_i \log[w p_{\mu_1, \sigma_1^2}(X_i)] + (1 - \delta_i) \log[(1 - w) p_{\mu_2, \sigma_2^2}(X_i)] = \\ &= \sum_{i=1}^n \delta_i \log p_{\mu_1, \sigma_1^2}(X_i) + (1 - \delta_i) \log p_{\mu_2, \sigma_2^2}(X_i) + \sum_{i=1}^n \delta_i \log(w) + (1 - \delta_i) \log(1 - w) \end{aligned}$$

Перейдем собственно к ЕМ-алгоритму.

E-step.

Будем рассматривать мат. ожидание δ_i , то есть ожидаемое значение δ_i .

$$\begin{aligned} \gamma_i &= E(\delta_i | X, \theta) = P(\delta_i = 1 | X, \theta) = |Bayes equation| = \\ &= \frac{P(X_i | \delta_i = 1, \theta) P(\delta_i = 1 | \theta)}{P(X_i | \theta)} = \frac{P(X_i | \delta_i = 1, \theta) P(\delta_i = 1 | \theta)}{P(X_i | \delta_i = 0, \theta) P(\delta_i = 0 | \theta) + P(X_i | \delta_i = 1, \theta) P(\delta_i = 1 | \theta)} = \\ &= \frac{p_{\mu_1, \sigma_1^2}(X_i) w}{p_{\mu_1, \sigma_1^2}(X_i) w + p_{\mu_2, \sigma_2^2}(X_i) (1 - w)} \end{aligned}$$

Где последнее равенство верно так как $P(\delta_i = 1 | \theta) = w$ и $P(X_i | \delta_i = 1, \theta) = p_{\mu_1, \sigma_1^2}(X_i)$. (Здесь немного плохо говорить о вероятности $P(X_i | \delta_i = 1, \theta)$ так как распределение абсолютно непрерывное и такая вероятность $= 0$. Здесь под этим обозначением стоит

понимать именно условную плотность).

M-step

На этом шаге нам нужно максимизировать правдоподобие, предполагая, что мы знаем δ_i . На E шаге мы нашли ожидаемое значение для $\delta_i - \gamma_i$. Поэтому будем рассматривать именно это значение. Тогда лог-правдоподобие выглядит следующим образом:

$$l_X(\theta) = \sum_{i=1}^n \gamma_i \log p_{\mu_1, \sigma_1^2}(X_i) + (1 - \gamma_i) \log p_{\mu_2, \sigma_2^2}(X_i) + \\ + \sum_{i=1}^n \gamma_i \log(w) + (1 - \gamma_i) \log(1 - w)$$

Заметим, что мы можем максимизировать полученные 2 суммы отдельно. Первую по μ_1, μ_2 , а вторую по w , так как на этом шаге мы полагаем известными γ_i .

Попробуем максимизировать второе слагаемое по w .

Максимизация этого слагаемого равносильна максимизации $\prod_{i=1}^n w^{\gamma_i} (1 - w)^{1 - \gamma_i} = w^{\sum_{i=1}^n \gamma_i} (1 - w)^{n - \sum_{i=1}^n \gamma_i}$. Для этого возьмем производную и приравняем к 0:

$$\sum_{i=1}^n \gamma_i w^{\sum_{i=1}^n \gamma_i - 1} (1 - w)^{n - \sum_{i=1}^n \gamma_i} - w^{\sum_{i=1}^n \gamma_i} (n - \sum_{i=1}^n \gamma_i) (1 - w)^{n - \sum_{i=1}^n \gamma_i - 1} = 0 \\ \Rightarrow \frac{\sum_{i=1}^n \gamma_i}{w} + \frac{n - \sum_{i=1}^n \gamma_i}{1 - w} = 0 \\ \Rightarrow w = \frac{\sum_{i=1}^n \gamma_i}{n}$$

Полученное значение даже кажется понятным - ожидаемое значение кол-ва элементов, попавших в первую гауссиану.

Теперь рассмотрим первое слагаемое. Это первое слагаемое можно переписать в виде :

$$\sum_{i=1}^n \gamma_i \log p_{\mu_1, \sigma_1^2}(X_i) + (1 - \gamma_i) \log p_{\mu_2, \sigma_2^2}(X_i) = \\ = \sum_{i=1}^n \gamma_i \left[-\frac{1}{2} \log(2\pi\sigma_1^2) - \frac{(X_i - \mu_1)^2}{2\sigma_1^2} \right] + (1 - \gamma_i) \left[-\frac{1}{2} \log(2\pi\sigma_2^2) - \frac{(X_i - \mu_2)^2}{2\sigma_2^2} \right]$$

Для того, чтобы найти новую оценку для μ_1, μ_2 возьмем производные данного выражения.

По μ_1 :

$$\sum_{i=1}^n \gamma_i \frac{X_i - \mu_1}{\sigma_1^2} = 0 \\ \Rightarrow \sum_{i=1}^n \gamma_i X_i - \mu_1 \sum_{i=1}^n \gamma_i = 0 \\ \Rightarrow \mu_1 = \frac{\sum_{i=1}^n \gamma_i X_i}{\sum_{i=1}^n \gamma_i}$$

Аналогично для производной по μ_2 получаем:

$$\sum_{i=1}^n (1 - \gamma_i) \frac{X_i - \mu_2}{\sigma_2^2} = 0$$
$$\Rightarrow \mu_2 = \frac{\sum_{i=1}^n (1 - \gamma_i) X_i}{\sum_{i=1}^n (1 - \gamma_i)}$$

Вот, собственно и все. Повторяем шаги Е и М до стабильности.

2 Задача 2

В алгоритме кластеризации k-means мы минимизируем L2 расстояние между объектами в одном кластере: $D(x, y) = \sum_{i=1}^n (x_i - y_i)^2$. Однако рассматривать L2 расстояние для категориальных - плохо. Давайте подумаем какую кластеризацию мы хотим получить. Хотим, чтобы в одном кластере лежали объекты с примерно похожими вещественными признаками и с совпадающими категориальными признаками. Пусть первые k признаков являются вещественными, а последние d - k - категориальными. Давайте тогда рассматривать другое расстояние между объектами :

$$D(x, y) = \sum_{i=1}^k (x_i - y_i)^2 + \sum_{i=k+1}^d I(x_i \neq y_i)$$

. При минимизации данного расстояния между объектами в кластере мы, очевидно, также будем стараться поместить в один кластер объекты с совпадающими категориальными признаками. Давайте теперь поймем как измениться наш алгоритм.

- 1) Выбираем начальные центры кластеров случайно.
- 2) Перестраиваем кластеры. Каждый объект определяем к тому кластеру, к чьему центру он ближе по введенному выше расстоянию.
- 3) Изменяем центры кластеров. Для каждого кластера находим новый центр кластера. Для этого для каждого вещественного признака находим его среднее значение среди объектов данного кластера (как известно, имеено среднее минимизирует сумму L2 расстояний). Для каждого категориального признака сделаем голосование : j-ый категориальный признак для центра выбираем как самое популярное значение j-ого признака в этом кластере. Очевидно, что именно такой выбор центров кластеров будет минимизировать сумму введенных ранее расстояний от каждого объекта до соответствующего центра при фиксированных центрах.
- 4) пока не центры кластеров не стабилизируются повторяем 2) и 3).

Также стоит отметить, что перед тем, как применять данный алгоритм стоит отмасштабировать вещественные признаки (что также нужно и при обычном алгоритме) так как иначе мы, возможно, будем минимизировать расстояние только для одного признака, который просто имеет большой масштаб. Также стоит сказать, что даже после масштабирования вещественных признаков нам возможно все равно будет не выгодно

минимизировать $\sum_{i=k+1}^d I(x_i = y_i)$, поэтому перед этой суммой стоит добавить гиперпараметр λ , позволяющий регулировать то, насколько сильно мы хотим опираться на категориальные признаки при кластеризации. То есть итоговое расстояние нужно взять $D(x, y) = \sum_{i=1}^k (x_i - y_i)^2 + \lambda \sum_{i=k+1}^d I(x_i = y_i)$

3 Задача 3

Введем обозначение: $M = \max [r_1, \dots, r_t]$. Предположим, что мы сначала сгенерировали вектор $[r_1, r_2, \dots, r_n]$, (который мы не знаем), соответствующий рекламам $[a_1, a_2, \dots, a_n]$. Будем идти по этому вектору последовательно (Случайный выбор рекламы на каждом шаге соответствует последовательному прохождению при другом порядке сгенерированных элементов, который равновероятен тому, что мы имеем)

1)

$$P(r_{t+1} > \alpha M) = (1 - P(r_{t+1} \leq \alpha M)) = 1 - \int_0^{\alpha M} p(x) dx$$

2)

$$\begin{aligned} P(\text{хотя бы один из } r_{t+1}, r_{t+2}, \dots, r_n > \alpha M) &= 1 - P(\text{все из } r_{t+1}, r_{t+2}, \dots, r_n \leq \alpha M) = \\ &= 1 - P(\forall t+1 \leq s \leq n : r_s \leq \alpha M) = \\ &= |\text{элементы } r_{t+1}, r_{t+2}, \dots, r_n \text{ были сгенерированы независимо и равновероятно}| = \\ &= 1 - P(r_{t+1} \leq \alpha M)^k = 1 - \left(\int_0^{\alpha M} p(x) dx \right)^n \end{aligned}$$

3) Обозначим $I = \int_0^{\alpha M} p(x) dx$. Также введем обозначение r_{sum} , равное прибыли.

$$\begin{aligned} &E(\text{прибыль в такой стратегии, начиная с } r_{t+1}) = \\ &= \sum_{k=t+1}^n E(r_{sum} | \text{на шаге } k \text{ нашли элемент } > \alpha M) P(\text{на шаге } k \text{ нашли элемент } > \alpha M) = \\ &= \sum_{k=t+1}^n E(r_{sum} | \text{на шаге } k \text{ нашли элемент } > \alpha M) I^{k-t+1} (1 - I) = \\ &|\text{Обозначим } A = (\text{на шаге } k \text{ нашли элемент } > \alpha M)| \\ &= \sum_{k=t+1}^n [E(r_{t+1} + r_{t+2} + \dots + r_{t+k-1} | A) + E(r_k + r_k + \dots + r_k | A)] I^{k-t+1} (1 - I) = \\ &= \sum_{k=t+1}^n [(k-1) E(r_{t+1} | A) + (n-t-k+1) E(r_{t+1} | A)] I^{k-t+1} (1 - I) = \\ &\sum_{k=t+1}^n [(k-1) E(r_{t+1} | r_{t+1} \leq \alpha M) + (n-t-k+1) E(r_{t+1} | r_{t+1} > \alpha M)] I^{k-t+1} (1 - I) \end{aligned}$$

4 Задача 4

Условие не совсем ясно.

Если при выборе примера с какого-то источника у нас нет возможности как-то сказать, что пример должен быть из определенного места пространства, то тут тогда, наверно, лучше действовать согласно максимальной сумме полезностей. То есть выбрать элементы из источников так, чтобы набрать максимальную полезность и потратить меньше денег.

Если же мы можем как-то выбирать примеры (или то, где они примерно лежат в пространстве), то оптимальной стратегией кажется следующее: Возьмем сначала несколько дешевых примеров, которые будут разбросаны более-менее равномерно по всему рассматриваемому пространству. После этого обучим классификатор или несколько классификаторов и посмотрим на их предсказания. Найдем в нашем пространстве место (объекты) в которых наши классификаторы наименее уверены или вообще ошибаются. Уверенность здесь можно понимать по-разному : принцип наименьшей достоверности, принцип разности отступов, максимума энтропии и прочее. Таким образом нашли области, на которых наш классификатор работает плохо. Теперь из источника с дорогими и хорошими примерами возьмем пример из этой области и обучим классификатор заново. Опять найдем область, где классификатор работает плохо и возьмем дорогой пример из этой области. И так далее.

Такая стратегия кажется оптимальной так как сначала мы немного исследуем пространство и начинаем понимать примерную картину происходящего, совсем не точную. После этого этапа мы потратили не слишком много денег и при этом научились строить хоть какой-то, не самый плохой классификатор. Покупая дорогие примеры из областей неувернности мы уточняем наши знания в самых плохих областях и тем самым картина становится гораздо более ясной.

5 Задача 5

Рассмотрим фиксированный X_i из исходной выборки X .

$$\begin{aligned} P(X_i \text{ не попадет в } \hat{X}) &= P(X_i \neq \hat{X}_1, X_i \neq \hat{X}_2, \dots, X_i \neq \hat{X}_l) = \\ &= |\text{элементы бутстрепной выборки генерируются независимо}| = \\ &= P(X_i \neq \hat{X}_1) * P(X_i \neq \hat{X}_2) \dots * P(X_i \neq \hat{X}_l) = \\ &= |\text{элементы бутстрепной выборки одинаково распределены}| = \\ &= P(X_i \neq \hat{X}_l)^l = (1 - P(X_i = \hat{X}_l))^l = (1 - \frac{1}{l})^l \xrightarrow{l \rightarrow \infty} e^{-1} \approx 0.63 \approx \frac{2}{3} \end{aligned}$$

6 Задача 6

Поймем как проходят итерации алгоритма Q-обучения. На итерации t мы :

- выбираем $a_t \sim \pi_t(a|s_t)$:
 $a_t := \operatorname{argmax}_a Q(s_t|a)$
- среда генерирует премию и новое состояние:
 $r_t \sim p(r|s_t)$
 $s_{t+1} \sim p(s|a_t, s_t)$
- пересчитываем функцию ценности действия.
 $Q(s_t, a_t) = Q(s_t, a_t) + \alpha_t(r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t))$

Тогда понимаем, что в нашем случае мы сначала находились в состоянии S_1 после выбрали действия в соответствии со стратегией $a_t := \operatorname{argmax}_a Q(S_1|a)$, что в нашем случае равно a_0 . После этого среда генерирует премию и новое состояние. В данном случае мы имеем только 3 вероятности : с вероятностью 0.7 переходим в состояние S_0 , получая премию +5, с вер-тью 0.1 переходим в S_1 с нулевой премией и с вер-тью 0.2 переходим в S_2 также с нулевой премией.

По условию нам известно, что мы перешли в состояние S_0 .

Теперь нам осталось только пересчитать функция полезности действия и тем самым изменить нашу стратегию. Подставим известные нам значения : $s_t = S_1, s_{t+1} = S_0, a_t = a_0, r_t = +5, \gamma = 0.5$ в формулу для пересчета Q и получим:

$$\begin{aligned} Q(S_1, a_0) &= Q(S_1, a_0) + \alpha_t(r_t + \gamma \max_{a'} Q(S_0, a') - Q(S_1, a_0)) = \\ &= Q(S_1, a_0) + \alpha_t(+5 + \gamma Q(S_0, a_1) - Q(S_1, a_0)) = 4 + \alpha_t(5 + 0.5 * 3 - 4) = 4 + 2.5\alpha_t \end{aligned}$$

где α_t - фактор обучения, который не дан нам в условии.

Таким образом, получили, что после данного шага в матрице Q изменится значение $Q(S_1, a_0)$ на $4 + 2.5\alpha_t$. Тем самым и изменится наша стратегия, так как наша стратегия π заключается в том, что при известном текущем состоянии выбирать действие с максимальной полезностью : $a_t := \operatorname{argmax}_a Q(s_t|a)$.