

Задачи МОТП v 0.2

by Александра Астахова, Юрий Бердников, Николай Леонов, Михаил Нокель, Данила Потапов

Задача 1

Вывести формулы векторного дифференцирования:

$$\nabla_x a^T x = a$$

$$\nabla_x \|Ax - b\|^2 = 2A^T Ax - 2A^T b$$

$$\nabla_x^2 \|Ax - b\|^2 = 2A^T A$$

Решение:

Будем считать все матрицы и векторы вещественнозначными. Тогда:

$$1. \nabla_x a^T x = \nabla_x (a_1 x_1 + \dots + a_n x_n) = \left(\frac{\partial \sum a_i x_i}{\partial x_1}, \dots, \frac{\partial \sum a_i x_i}{\partial x_n} \right) = (a_1, \dots, a_n) = a$$

Заметим также, что $a^T x = \langle a, x \rangle$.

$$2. \nabla_x \|Ax - b\|^2 = \nabla_x \langle Ax - b, Ax - b \rangle = \nabla_x \langle Ax, Ax \rangle - \nabla_x \langle Ax, b \rangle - \nabla_x \langle b, Ax \rangle + \nabla_x \langle b, b \rangle = \nabla_x \langle A^T Ax, x \rangle - 2\nabla_x \langle A^T b, x \rangle + \nabla_x \langle b, b \rangle = \nabla_x \langle A^T Ax, x \rangle - 2A^T b$$

Рассмотрим подробнее член $\nabla_x \langle A^T Ax, x \rangle$:

$$\begin{aligned} \nabla_x \langle A^T Ax, x \rangle &= \nabla_x ((A^T Ax)^T x) = \nabla_x ((Ax)^T Ax) = \nabla_x \sum_{i=1}^n \sum_{j=1}^n x_i x_j A_i^T A_j \Rightarrow \frac{\partial (Ax)^T Ax}{\partial x_i} \\ &= 2 \sum_{j=1}^n x_j A_i^T A_j \Rightarrow \nabla_x \langle A^T Ax, x \rangle = 2A^T Ax \end{aligned}$$

где A_k - k-й столбец матрицы A .

Таким образом:

$$\nabla_x \langle A^T Ax, x \rangle = 2A^T Ax \Rightarrow \nabla_x \|Ax - b\|^2 = 2A^T Ax - 2A^T b$$

$$3. \nabla_x^2 \|Ax - b\|^2 = \nabla_x (\nabla_x \|Ax - b\|^2) = \nabla_x (2A^T Ax - 2A^T b) = 2A^T A.$$

Задача 2

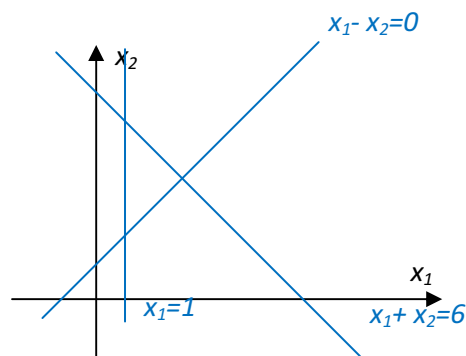
Найти нормальное псевдорешение для системы линейных уравнений.

Решение:

Рассмотрим переопределенную нерешаемую систему линейных уравнений:

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} x = \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix}$$

Данная система задает в двумерном пространстве три прямые, не пересекающиеся в одной точке:



Таким образом она не решается. Найдем ее псевдорешение, для этого воспользуемся формулой для его нахождения:

$$x = (A^T A)^{-1} A^T b$$

Вычислим соответствующее произведение по порядку:

$$A^T = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

$$A^T A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$$

$$(A^T A)^{-1} = \frac{1}{6} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

$$(A^T A)^{-1} A^T = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

$$(A^T A)^{-1} A^T b = \begin{pmatrix} 2\frac{1}{3} \\ 3 \end{pmatrix}$$

Заметим, что полученное решение соответствует центру треугольника, образованного прямыми.

Задача 3

Даны N точек в двумерном пространстве. Найти с помощью метода главных компонент первую главную компоненту и проекцию выборки на одномерное пространство.

Решение:

Рассмотрим 4 точки в двумерном пространстве ($N = 4$):

$$x_1 = (1,2), x_2 = (2,1), x_3 = (4,2), x_4 = (6,3)$$

Вычислим выборочное среднее \bar{x} :

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{4} (13,8) = \left(3\frac{1}{4}, 2\right)$$

Затем вычислим выборочную матрицу ковариаций S как:

$$S = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^T (x_i - \bar{x}) = \frac{1}{4} \left(\begin{pmatrix} 81 & 0 \\ 16 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 25 & 5 \\ 16 & 4 \\ 5 & 1 \end{pmatrix} + \begin{pmatrix} 9 & 0 \\ 16 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 121 & 11 \\ 16 & 4 \\ 11 & 1 \end{pmatrix} \right) = \frac{1}{4} \begin{pmatrix} 14 & 3 \\ 4 & 4 \\ 4 & 2 \end{pmatrix}$$

Найдем собственные вектора матрицы S , решая характеристическое уравнение $\det(S - \lambda I) = 0$:

$$\det \begin{pmatrix} \frac{59}{16} - \lambda & 1 \\ 1 & \frac{1}{2} - \lambda \end{pmatrix} = \lambda^2 - \frac{67}{16} \lambda + \frac{59}{32} = 0$$

Получаем $\lambda_{1,2} = \frac{\frac{67}{16} \pm \frac{51}{16}}{2} = \left[\frac{59}{2} \right]$. Вектор, соответствующий главной компоненте, отвечает

максимальному собственному значению выборочной матрицы ковариации. Получим его, решая однородную систему уравнений $(S - \lambda_{max} I) \hat{d} = 0$ и нормируя:

$$\hat{d} = (0.99, 0.018)^T, \|\hat{d}\| = 1$$

Спроецируем точки выборки на полученную прямую:

$$T = XW = X\hat{d}$$

где W -матрица из собственных векторов, X -матрица из точек выборки, центрированных относительно выборочного среднего, а T -матрица проекций:

$$\begin{pmatrix} -2\frac{1}{4} & 0 \\ -1\frac{1}{4} & -1 \\ \frac{3}{4} & 0 \\ 2\frac{3}{4} & 1 \end{pmatrix} \begin{pmatrix} 0.99 \\ 0.018 \end{pmatrix} = \begin{pmatrix} -2.22 \\ -1.25 \\ 0.74 \\ 2.74 \end{pmatrix}$$

Задача 4

Дана выборка $X = \{x_1, \dots, x_N\}$ из некоторого распределения $p(x)$. Требуется оценить по выборке с помощью метода максимального правдоподобия значения параметров этого распределения.

Например, у распределения Лапласа вида $p(x) = \frac{1}{2b} e^{-\frac{|x-\mu|}{2b}}$ оценить мат. ожидание μ при известном b или оценить b при известном μ .

Решение:

Запишем логарифм функции правдоподобия:

$$L(X|b, \mu) = n \log \frac{1}{2b} - \sum_{i=1}^n \frac{|x_i - \mu|}{2b},$$

Тогда

$$\frac{\partial L}{\partial \mu} = \sum_{i=1}^n \operatorname{sgn}(x_i - \mu).$$

Таким образом, μ_{ML} должно удовлетворять уравнению $\sum_{i=1}^n \operatorname{sgn}(x_i - \mu_{ML}) = 0$. Очевидно, что $\sum_{i=1}^n \operatorname{sgn}(x_i - \mu)$ обращается в 0, когда $\mu = \operatorname{med}(x_1, \dots, x_n)$, так как тогда количество положительных слагаемых будет совпадать с количеством отрицательных.

Обозначим $\lambda = \frac{1}{b}$ и перепишем логарифм функции правдоподобия:

$$L(X|b, \lambda) = n \log \lambda - \frac{1}{2} \lambda \sum_{i=1}^n |x_i - \mu| - n \log 2.$$

Требуется найти такое λ , что:

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} - \frac{1}{2} \sum_{i=1}^n |x_i - \mu| = 0.$$

Следовательно, ответ:

$$b_{ML} = \frac{1}{\lambda} = \frac{1}{2n} \sum_{i=1}^n |x_i - \mu|$$

$$\mu_{ML} = \operatorname{med}(x_1, \dots, x_n).$$

Задача 5

Дана выборка из N точек в двумерном пространстве. Первая координата – это x , вторая – t . С помощью метода наименьших квадратов построить линейную регрессию вида $\hat{t} = kx + b$, т.е. найти коэффициенты k и b .

Решение

Метод наименьших квадратов состоит в нахождении коэффициентов регрессии путем минимизации квадратичной функции потерь $S(t, \hat{t}) = (t - \hat{t})^2$. Для выборки из N точек $\{(x_1, t_1), \dots, (x_N, t_N)\}$ и заданного вида регрессии функция потерь будет иметь вид:

$$\sum_{i=1}^N (t_i - kx_i - b)^2 \rightarrow \min_{k,b}$$

Условия минимума на S будут:

$$\frac{\partial S}{\partial k} = -2 \sum_{i=1}^N (t_i - kx_i - b)x_i = 0$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^N (t_i - kx_i - b) = 0$$

Их можно переписать в виде:

$$\sum_{i=1}^N t_i x_i = k \sum_{i=1}^N x_i^2 + b \sum_{i=1}^N x_i$$

$$\sum_{i=1}^N t_i = k \sum_{i=1}^N x_i + Nb$$

Таким образом, выражая k и b по заданной выборке, получаем коэффициенты линейной регрессии.

Задача 6

Решить задачу условной оптимизации выпуклой функции при выпуклых ограничениях, например:

$$-5x^2 + 2xy - 3y^2 \rightarrow \max_{x,y}$$

$$x = y - 1$$

Решение:

Для решения задачи оптимизации воспользуемся правилом множителей Лагранжа. Для этого приведем ограничение к стандартному виду и запишем функцию Лагранжа $L(x, y, \lambda)$:

$$x = y - 1 \Leftrightarrow g(x, y) = x - y + 1 = 0$$

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y) = -5x^2 + 2xy - 3y^2 + \lambda(x - y + 1)$$

Условия экстремума функции $f(x, y)$ тогда примут вид:

$$\frac{\partial L}{\partial x} = -10x + 2y + \lambda = 0$$

$$\frac{\partial L}{\partial y} = 2x - 6y - \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = x - y + 1 = 0$$

Для нахождения экстремума функции необходимо решить систему:

$$\begin{pmatrix} -10 & 2 & 1 \\ 2 & -6 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

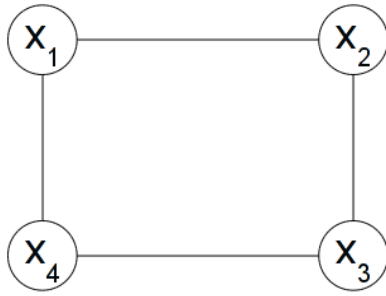
Решением данной системы будет вектор:

$$(x_0, y_0, z_0)^T = \left(-\frac{1}{3}, \frac{2}{3}, -\frac{14}{3}\right)$$

Значит, максимум искомой функции достигается в точке $\left(-\frac{1}{3}, \frac{2}{3}\right)$.

Задача 7

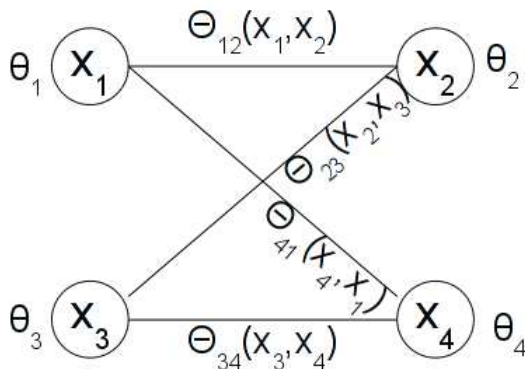
Дана марковская сеть вида:



с бинарными переменными x_1, x_2, x_3, x_4 . Для этой модели заданы все значения унарных функций $\theta_1(x_1), \theta_2(x_2), \theta_3(x_3), \theta_4(x_4)$ и бинарных функций $\theta_{12}(x_1, x_2), \theta_{23}(x_2, x_3), \theta_{34}(x_3, x_4), \theta_{41}(x_4, x_1)$. С помощью репараметризации построить граф, минимальный разрез которого соответствует минимуму энергии вида $\sum_{i=1}^4 \theta_i(x_i) + \sum_{(i,j) \in E} \theta_{ij}(x_i, x_j)$.

Решение:

Приведём марковскую сеть к стандартному виду:



Теперь проведём репараметризацию данной сети, сделав веса горизонтальных рёбер нулевыми, а веса диагональных рёбер - равными. Получим следующую систему уравнений (обозначив через a_i - величину, вычитаемую из рёбер соответствующей вершины x_i и прибавляемую к весу этой вершины):

$$\begin{cases} \theta_{12} - a_1 - a_2 = 0 \\ \theta_{34} - a_3 - a_4 = 0 \\ \theta_{41} - a_1 - a_4 = \theta_{23} - a_2 - a_3 \end{cases}$$

Решая данную систему, придём к следующей:

$$\begin{cases} a_2 = \theta_{12} - a_1 \\ a_4 = \theta_{34} - a_3 \\ a_3 = a_1 + \frac{-\theta_{41} + \theta_{34} + \theta_{23} - \theta_{12}}{2} \end{cases}$$

$$\begin{cases} a_2 = & \theta_{12} - a_1 \\ a_3 = & a_1 + \frac{-\theta_{41} + \theta_{34} + \theta_{23} - \theta_{12}}{2} \\ a_4 = & \frac{\theta_{34} - \theta_{23} + \theta_{12} + \theta_{41}}{2} - a_1 \end{cases}$$

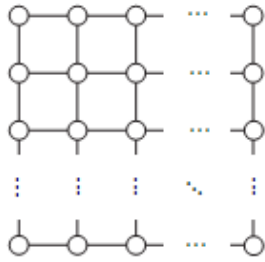
Отсюда веса диагональных рёбер будут

$$\theta_{41} - a_1 - a_4 = \frac{\theta_{41} - \theta_{34} - \theta_{12} + \theta_{23}}{2}$$

Веса горизонтальных же рёбер равны 0 (по определению репараметризации). Веса же самих вершин определены с точностью до неизвестной. Важно лишь, чтобы эти веса были неотрицательны.

Задача 8

Дана марковская сеть с бинарными переменными вида “решетка”:



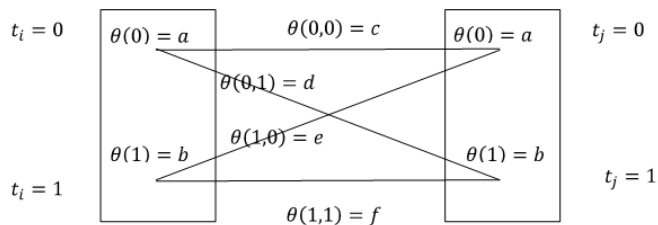
Пусть все унарные энергии совпадают для всех вершин $\theta_i(x_i) = \theta(x)$ и равны $\theta(0) = a, \theta(1) = b$. Аналогично все бинарные энергии совпадают между собой $\theta_{ij}(x_i, x_j) = \theta(x, y)$ и равны $\theta(0,0) = c, \theta(0,1) = d, \theta(1,0) = e, \theta(1,1) = f$. Требуется выполнить репараметризацию в этом графе так, чтобы все энергии $\theta_{ij}(0,0) = \theta_{ij}(1,1) = 0$.

Решение:

Рассмотрим две соседние вершины t_i и t_j :

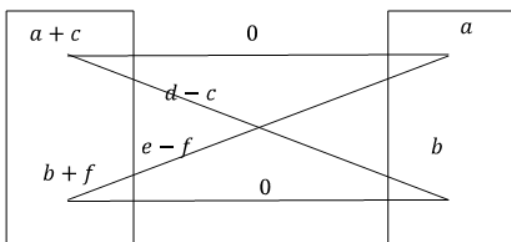


Для них:



Вычитая одинаковые значения из двух ребер, сходящихся в одну вершину и прибавляя это значение к весу самой вершины, мы получаем эквивалентный функционал энергии.

Таким образом, вычитая из ребер, сходящихся в левой нижней вершине f и прибавляя к ее весу f , получим $\theta'(1,1) = 0$. Аналогично с левой верхней вершиной. Получим:



В случае если требуется, чтобы так же были равными диагональные энергии $\theta'(0,1) = \theta'(1,0)$, решим систему:

$$c - \delta_1 - \delta_2 = 0$$

$$f - \delta_3 - \delta_4 = 0$$

$$d - \delta_1 - \delta_4 = e - \delta_2 - \delta_3$$

Где δ_1 прибавляем к левой верхней вершине и вычитаем из входящих в нее ребер, аналогично δ_2 прибавляем к правой верхней, δ_3 соответствует левой нижней, а δ_4 - правой нижней.

В итоге получим:

$$\delta_1 = \delta_3 + \frac{d - e + c - f}{2}.$$

Положив, например, $\delta_3 = 0$, получим

$$\delta_1 = \frac{d - e + c - f}{2}$$

$$\delta_2 = \frac{c - d + e + f}{2}$$

$$\delta_4 = f$$

Тогда после репараметризации энергии будут равны:

$$\theta'(0,1) = d - \delta_1 - \delta_4 = \frac{d + e - c - f}{2}$$

$$\theta'(1,0) = e - \delta_2 - \delta_3 = \frac{d + e - c - f}{2}$$

$$\theta'(0,0) = c - \delta_1 - \delta_2 = 0$$

$$\theta'(1,1) = f - f = 0$$

Таким образом, получили репараметризацию, при которой горизонтальные ребра графа имеют нулевые веса, а веса диагональных ребер совпадают.

Задача 9

Дана следующая вероятностная модель:

$$p(X, T | \mu, a_0, b_0) = \prod_{n=1}^N p(x_n, t_n | \mu, a_0, b_0) = \prod_{n=1}^N p(x_n | t_n, \mu) p(t_n | a_0, b_0),$$

$$p(x_n | t_n, \mu) = N(x_n | \mu, t_n^{-1}) = \sqrt{\frac{t_n}{2\pi}} e^{-\frac{t_n}{2}(x_n - \mu)^2},$$

$$p(t_n | a_0, b_0) = G(t_n | a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} t_n^{a_0-1} e^{-b_0 t_n}, \quad E t_n = \frac{a_0}{b_0}.$$

Требуется выписать формулы EM-алгоритма для максимизации правдоподобия:

$$p(X | \mu, a_0, b_0) \rightarrow \max_{\mu}$$

при фиксированных a_0, b_0 .

Решение:

Выписываем E-шаг.

На E-шаге необходимо вычислить $\log(p(X, T | \theta_{old}))$ и найти условное математическое ожидание $E_{T|X, \theta_{old}} \log(p(X, T | \theta_{old}))$. Вычислим их:

$$\begin{aligned} \log(p(X, T | \theta_{old})) &= \sum_{n=1}^N \log(p(x_n | t_n, \mu)) + \sum_{n=1}^N \log(p(t_n | a_0, b_0)) = \\ &= \sum_{n=1}^N \left(\log\left(\sqrt{\frac{t_n}{2\pi}}\right) - \frac{t_n}{2}(x_n - \mu)^2 + \log\left(\frac{b_0^{a_0}}{\Gamma(a_0)}\right) + (a_0 - 1)\log(t_n) - b_0 t_n \right) \end{aligned}$$

Учитывая, что максимизация правдоподобия будет проводиться только лишь по μ , перепишем найденную величину следующим образом:

$$\log(p(X, T | \mu_{old})) = -\sum_{n=1}^N \frac{t_n}{2}(x_n - \mu)^2 + f(t_n)$$

Тогда вычислим условное математическое ожидание, учитывая, что $E t_n = \frac{a_0}{b_0}$:

$$E_{T|X, \theta_{old}} \log(p(X, T | \theta_{old})) = \frac{N a_0}{2 b_0} \sum_{n=1}^N (x_n - \mu)^2 + E f(t_n)$$

Выписываем M-шаг.

На M-шаге максимизируется условное математическое ожидание, найденное на E-шаге:

$$\frac{\delta E_{T|X, \theta_{old}} \log(p(X, T | \theta_{old}))}{\delta \mu} = -\frac{2Na_0}{2b_0} \sum_{n=1}^N (x_n - \mu) = 0$$

Отсюда следует:

$$\mu = \frac{1}{N} \sum_{n=1}^N x_n$$

Задача 10

Рассматривается вероятностная смесь двух дискретных распределений вида:

$$p(x) = \gamma p_1(x) + (1-\gamma)p_2(x).$$

Величина x может принимать значения (1,2,3). При этом параметры распределений равны:

$$\begin{array}{ccc} p1: & 1 & 2 & 3 & p2: & 1 & 2 & 3 \\ & \alpha & 1-\alpha & 0 & & 0 & 1-\beta & \beta \end{array}$$

Выборка X состоит из 30 единиц, 20 двоек и 60 троек. Требуется провести первые две итерации EM-алгоритма для восстановления параметров смеси (α, β, γ) для начального приближения $\alpha_0 = \beta_0 = \gamma_0 = 0.5$.

Решение:

Первая итерация

E-шаг

Вычисляем значения $g_{ij} = \frac{w_j * p_j(x_i)}{\sum_j w_j * p_j(x_i)}$:

$$g_{11} = \frac{0.5 * 0.5}{0.5 * 0.5 + 0 * 0.5} = 1, \quad g_{12} = \frac{0 * 0.5}{0.5 * 0.5 + 0 * 0.5} = 0,$$

$$g_{21} = \frac{0.5 * 0.5}{0.5 * 0.5 + 0.5 * 0.5} = 0.5, \quad g_{22} = \frac{0.5 * 0.5}{0.5 * 0.5 + 0.5 * 0.5} = 0.5,$$

$$g_{31} = \frac{0 * 0.5}{0.5 * 0.5 + 0 * 0.5} = 0, \quad g_{32} = \frac{0.5 * 0.5}{0.5 * 0.5 + 0 * 0.5} = 1.$$

M-шаг

Вычисляем новые значения параметров, максимизируя "взвешенное" правдоподобие:

$$30 g_{11} \log(p_1(x=1)) + 20 g_{21} \log(p_1(x=2)) \rightarrow \max_{\alpha},$$

$$60 g_{32} \log(p_2(x=3)) + 20 g_{22} \log(p_2(x=2)) \rightarrow \max_{\beta}$$

Отсюда, подставив значения:

$$30 \log(\alpha) + 10 \log(1-\alpha) \rightarrow \max_{\alpha},$$

$$60 \log(\beta) + 10 \log(1-\beta) \rightarrow \max_{\beta}$$

Взяв логарифм и продифференцировав первое выражение по α , второе - по β , получим:

$$\alpha = \frac{3}{4}, \beta = \frac{6}{7}$$

Пересчитаем γ :

$$\gamma = \frac{30g_{11} + 20g_{21}}{60 + 30 + 20} = \frac{4}{11}$$

Вторая итерация

E-шаг

Вычисляем значения $g_{ij} = \frac{w_j * p_j(x_i)}{\sum_j w_j * p_j(x_i)}$:

$$g_{11} = \frac{\frac{4}{11} * \frac{3}{4}}{\frac{4}{11} * \frac{3}{4} + 0 * \frac{7}{11}} = 1, \quad g_{12} = \frac{\frac{7}{11} * 0}{\frac{4}{11} * \frac{3}{4} + 0 * \frac{7}{11}} = 0,$$

$$g_{21} = \frac{\frac{4}{11} * \frac{1}{4}}{\frac{4}{11} * \frac{1}{4} + \frac{7}{11} * \frac{1}{7}} = 0.5, \quad g_{22} = \frac{\frac{7}{11} * \frac{1}{7}}{\frac{4}{11} * \frac{1}{4} + \frac{7}{11} * \frac{1}{7}} = 0.5,$$

$$g_{31} = \frac{\frac{4}{11} * 0}{\frac{4}{11} * 0 + \frac{7}{11} * \frac{6}{7}} = 0, \quad g_{32} = \frac{\frac{7}{11} * \frac{6}{7}}{\frac{7}{11} * \frac{6}{7} + 0 * \frac{7}{11}} = 1.$$

Заметим, что значения g_{ij} не поменялись. А значит, не поменяются и значения параметров на M-шаге.

M-шаг

Действительно в формулах для "взвешенного" правдоподобия ничего не изменилось:

$$30g_{11} \log(p_1(x=1)) + 20g_{21} \log(p_1(x=2)) \rightarrow \max_{\alpha},$$

$$60g_{32} \log(p_2(x=3)) + 20g_{22} \log(p_2(x=2)) \rightarrow \max_{\beta}$$

Следовательно:

$$\alpha = \frac{3}{4}, \beta = \frac{6}{7}$$

Аналогично и для γ :

$$\gamma = \frac{30g_{11} + 20g_{21}}{60 + 30 + 20} = \frac{4}{11}$$

Как видим, значения параметров на второй итерации не изменились.

Задача 11

Пусть имеется три бинарных переменных $a, b, c \in 0,1$, совместное распределение которых задаётся следующей таблицей:

a	b	c	$p(a,b,c)$
0	0	0	0.192
0	0	1	0.144
0	1	0	0.048
0	1	1	0.216
1	0	0	0.192
1	0	1	0.064
1	1	0	0.048
1	1	1	0.096

Требуется показать, что переменные a и b не являются независимыми, но при этом являются условно независимыми как при $c = 0$, так и при $c = 1$.

Решение:

Вначале покажем зависимость этих переменных. Покажем, например, что $p(a = 0, b = 0) \neq p(a = 0)p(b = 0)$: Имеем:

$$p(a = 0) = \sum_{a=0} p(a, b, c) = 0.192 + 0.144 + 0.048 + 0.216 = 0.6,$$

$$p(b = 0) = \sum_{b=0} p(a, b, c) = 0.192 + 0.144 + 0.192 + 0.064 = 0.592,$$

$$p(a = 0, b = 0) = \sum_{a=0, b=0} p(a, b, c) = 0.192 + 0.144 = 0.336$$

Очевидно:

$$0.336 \neq 0.592 * 0.6 = 0.3552$$

Значит, переменные зависимые.

Далее покажем условную независимость. Для этого составим таблички, в которые занесём необходимые вероятности: $p(a, c)$, $p(b, c)$, $p(c)$:

a	c	$p(a, c)$
0	0	0.24
0	1	0.36
1	0	0.24
1	1	0.16

b	c	$p(b,c)$
0	0	0.384
0	1	0.208
1	0	0.096
1	1	0.312

c	$p(c)$
0	0.48
1	0.52

Теперь проверим условную независимость переменных a и b при условии c , воспользовавшись следующей формулой, вытекающей из определения условной независимости двух величин:

$$\frac{p(a,b,c)}{p(c)} = \frac{p(a,c)}{p(c)} * \frac{p(b,c)}{p(c)}.$$

• При $c = 0$.

$$\text{При } a = 0, b = 0: \frac{0.192}{0.48} = \frac{0.24}{0.48} * \frac{0.384}{0.48} - \text{верно.}$$

$$\text{При } a = 0, b = 1: \frac{0.048}{0.48} = \frac{0.24}{0.48} * \frac{0.096}{0.48} - \text{верно.}$$

$$\text{При } a = 1, b = 0: \frac{0.192}{0.48} = \frac{0.24}{0.48} * \frac{0.384}{0.48} - \text{верно.}$$

$$\text{При } a = 1, b = 1: \frac{0.048}{0.48} = \frac{0.24}{0.48} * \frac{0.096}{0.48} - \text{верно.}$$

• При $c = 1$.

$$\text{При } a = 0, b = 0: \frac{0.144}{0.52} = \frac{0.36}{0.52} * \frac{0.208}{0.52} - \text{верно.}$$

$$\text{При } a = 0, b = 1: \frac{0.216}{0.52} = \frac{0.36}{0.52} * \frac{0.312}{0.52} - \text{верно.}$$

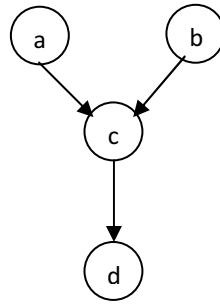
$$\text{При } a = 1, b = 0: \frac{0.064}{0.52} = \frac{0.16}{0.52} * \frac{0.208}{0.52} - \text{верно.}$$

$$\text{При } a = 1, b = 1: \frac{0.096}{0.52} = \frac{0.16}{0.52} * \frac{0.312}{0.52} - \text{верно.}$$

Отсюда следует, что данное равенство: $p(a,b|c) = p(a|c)p(b|c)$ и переменные условно независимы.

Задача 12

Пусть имеется байесовская сеть с графом следующего вида:



Требуется показать, что переменные a и b являются независимыми, но при этом не являются условно независимыми от d .

Решение:

Покажем, что a и b – независимы:

$p(a, b) = \int \int p(a, b, c, d) \mathbf{d}c \mathbf{d}d = \int \int p(d|c) * p(c|a, b) * p(a) * p(b) \mathbf{d}c \mathbf{d}d = p(a) * p(b) * \int p(c|a, b) \int p(d|c) \mathbf{d}d \mathbf{d}c = p(a) * p(b) * \int p(c|a, b) \mathbf{d}c = p(a) * p(b)$. Значит случайные величины a и b независимы по определению.

Докажем отсутствие условной независимости случайных величин a и b от d .

$$p(a|d) = \frac{p(a, d)}{p(d)} = \frac{p(a)p(d|c)}{\int p(d|c)p(c)dc}$$

$$p(b|d) = \frac{p(b, d)}{p(d)} = \frac{p(b)p(d|c)}{\int p(d|c)p(c)dc}$$

В то время как

$$p(a, b|d) = \frac{p(a, b, d)}{p(d)} = \frac{p(a)p(b)p(d|c)}{\int p(d|c)p(c)dc}$$

Следовательно $p(a, b|d) \neq p(a|d)p(b|d)$, что означает отсутствие условной независимости по d .

Задача 13

Имеется скрытая марковская модель с двумя скрытыми состояниями и бинарными наблюдаемыми переменными. Пусть наблюдаемая последовательность имеет вид $X = (1011001110001 \dots)$, т.е. идет группа из n единиц, потом группа из n нулей, потом группа из $n + 1$ единиц, $n + 1$ нулей и так далее. При этом наблюдаемая последовательность состояний $T = (1111222211112222 \dots)$. С помощью метода максимального правдоподобия требуется оценить вектор априорных вероятностей π и матрицу перехода A , если длина последовательности равна 200.

Решение:

Вспользуемся формулами, полученными для оценки параметров методом максимального правдоподобия:

$$\pi_j = t_{1j}, \quad j = 1, \dots, K$$

$$A_{ij} = \frac{\sum_{n=2}^N t_{n-1,i} t_{nj}}{\sum_{n=2}^N t_{n-1,i}}, \quad i, j = 1, \dots, K$$

Где $K = 2, N = 200$. Тогда $\pi = (1,0)$, так как $t_1 = (1,0)$, так как в первый момент модель находится в первом состоянии. Для матрицы перехода A получим:

$$A_{11} = \frac{\sum_{n=2}^N t_{n-1,1} t_{n1}}{\sum_{n=2}^N t_{n-1,1}} = \frac{25 * 3}{25 * 4} = 0.75$$

Числитель дроби соответствует количеству переходов из первого состояния во второе, а знаменатель обозначает общее количество переходов из первого состояния. Аналогично:

$$A_{12} = \frac{\sum_{n=2}^N t_{n-1,1} t_{n2}}{\sum_{n=2}^N t_{n-1,1}} = \frac{25 * 1}{25 * 4} = 0.25$$

$$A_{21} = \frac{\sum_{n=2}^N t_{n-1,2} t_{n1}}{\sum_{n=2}^N t_{n-1,2}} = \frac{24}{4 * 24 + 3} = 0. (24)$$

$$A_{22} = \frac{\sum_{n=2}^N t_{n-1,2} t_{n2}}{\sum_{n=2}^N t_{n-1,2}} = \frac{24 * 3 + 3}{4 * 24 + 3} = 0. (75)$$

Таким образом матрица A равна:

$$A = \begin{pmatrix} 0.75 & 0.25 \\ 0. (24) & 0. (75) \end{pmatrix}$$

Ответ: $\pi = (1,0), A = \begin{pmatrix} 0.75 & 0.25 \\ 0. (24) & 0. (75) \end{pmatrix}$.

Задача 14

Имеется скрытая марковская модель с двумя скрытыми состояниями. Вектор априорных вероятностей $\pi = (0.5, 0.5)$, матрица перехода $A = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$. Наблюдаемая переменная является бинарной, в первом состоянии значение 0 выпадает с вероятностью 0.8, а во втором состоянии – с вероятностью 0.2. Требуется с помощью алгоритма Витерби найти наиболее правдоподобную последовательность скрытых состояний для наблюдаемой последовательности $X = \{0, 0, 1, 0, 0, 1, 1\}$.

Решение:

Воспользуемся формулами для функции Беллмана:

$$V_{1j} = \log \pi_j$$

$$V_{nj} = \max_i (V_{n-1,i} + \log A_{ij} + \log p(x_n | \phi_j))$$

$$S_{nj} = \arg \max_i (V_{n-1,i} + \log A_{ij} + \log p(x_n | \phi_j)).$$

Тогда:

$$V_{11} = \log \pi_1 = \log 0.5$$

$$V_{12} = \log \pi_2 = \log 0.5$$

$$V_{21} = \max(\log 0.5 + \log 0.9 + \log 0.8, \log 0.5 + \log 0.2 + \log 0.8) = \max(\mathbf{\log 0.36}, \log 0.08)$$

$$V_{22} = \max(\log 0.5 + \log 0.1 + \log 0.2, \log 0.5 + \log 0.8 + \log 0.2) = \max(\log 0.01, \mathbf{\log 0.08})$$

$$V_{31} = \max(\log 0.36 + \log 0.9 + \log 0.2, \log 0.08 + \log 0.2 + \log 0.2) = \max(\mathbf{\log 0.0648}, \log 0.0032)$$

$$V_{32} = \max(\log 0.36 + \log 0.1 + \log 0.8, \log 0.08 + \log 0.8 + \log 0.8) = \max(\log 0.0288, \mathbf{\log 0.0512})$$

$$\begin{aligned} V_{41} &= \max(\log 0.0648 + \log 0.9 + \log 0.8, \log 0.0512 + \log 0.2 + \log 0.8) \\ &= \max(\mathbf{\log 0.0466}, \log 0.0082) \end{aligned}$$

$$\begin{aligned} V_{42} &= \max(\log 0.0648 + \log 0.1 + \log 0.2, \log 0.0512 + \log 0.8 + \log 0.2) \\ &= \max(\log 0.0013, \mathbf{\log 0.0082}) \end{aligned}$$

$$\begin{aligned} V_{51} &= \max(\log 0.00466 + \log 0.9 + \log 0.8, \log 0.0082 + \log 0.2 + \log 0.8) \\ &= \max(\mathbf{\log 0.0335}, \log 0.0013) \end{aligned}$$

$$\begin{aligned} V_{52} &= \max(\log 0.00466 + \log 0.1 + \log 0.2, \log 0.0082 + \log 0.8 + \log 0.2) \\ &= \max(\log 0.001, \mathbf{\log 0.0013}) \end{aligned}$$

$$\begin{aligned} V_{61} &= \max(\log 0.0335 + \log 0.9 + \log 0.2, \log 0.0013 + \log 0.2 + \log 0.2) \\ &= \max(\mathbf{\log 0.006}, \log 0.000\dots) \end{aligned}$$

$$V_{62} = \max(\log 0.0335 + \log 0.1 + \log 0.8, \log 0.0013 + \log 0.8 + \log 0.8)$$

$$= \max(\mathbf{\log 0.0027}, \log 0.000832)$$

$$V_{71} = \max(\log 0.006 + \log 0.9 + \log 0.2, \log 0.0027 + \log 0.2 + \log 0.2)$$

$$= \max(\mathbf{\log 0.001}, \log 0.0001)$$

$$V_{72} = \max(\log 0.006 + \log 0.1 + \log 0.8, \log 0.0027 + \log 0.8 + \log 0.8)$$

$$= \max(\log 0.00048, \mathbf{\log 0.00172}).$$

Таким образом, $\max(V_{71}, V_{72}) = i_7 = 2$. Рассматривая получившиеся значения S_{nj} , получим

$$i_6 = S_{72} = 2$$

$$i_5 = S_{62} = \mathbf{1}$$

$$i_4 = S_{51} = 1$$

$$i_3 = S_{41} = 1$$

$$i_2 = S_{31} = 1$$

$$i_1 = S_{21} = 1.$$

Значит, последовательность крытых состояний имеет вид: (1,1,1,1,1,2,2).

Ответ: (1,1,1,1,1,2,2).

Задача 15

Имеется скрытая марковская модель с двумя скрытыми состояниями. Вектор априорных вероятностей $\pi = (0.5, 0.5)$, матрица перехода $A = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$. Наблюдаемая переменная является бинарной, в первом состоянии значение 0 выпадает с вероятностью 0.8, а во втором состоянии – с вероятностью 0.2. Требуется с помощью алгоритма «вперед-назад» вычислить все маргинальные распределения вида $p(t_n|X)$ для наблюдаемой последовательности $X = (0, 0, 1)$.

Решение:

Пользуясь свойствами условной независимости в для скрытых марковских сетей, запишем

$$p(t_n|X) = \frac{\alpha(t_n) \cdot \beta(t_n)}{p(X)}, \text{ где } p(X) = \sum_{t_N} \alpha(t_N).$$

Вычислим $\alpha(t_n)$, пользуясь формулами:

$$\alpha(t_1) = \prod_{j=1}^K (\pi_j p(x_1|\phi_j))^{t_{1j}}$$

$$\alpha(t_n) = p(x_n|t_n) \sum_{t_{n-1}} \alpha(t_{n-1}) p(t_n|t_{n-1}).$$

Тогда

$$\alpha(t_1) = \alpha \begin{pmatrix} 0.8 & 0 \\ 0 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} = \alpha \begin{pmatrix} 0.4 \\ 0.1 \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix},$$

Где $\begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} = A$ – матрица перехода, $\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} = \pi$ – априорные вероятности, α – нормировочная константа. Аналогично

$$\alpha(t_2) = \alpha \begin{pmatrix} 0.8 & 0 \\ 0 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.8 \\ 0.2 \end{pmatrix} = \alpha \begin{pmatrix} 0.592 \\ 0.064 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix}$$

$$\alpha(t_3) = \alpha \begin{pmatrix} 0.2 & 0 \\ 0 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.9 \\ 0.1 \end{pmatrix} = \alpha \begin{pmatrix} 0.164 \\ 0.208 \end{pmatrix} = \begin{pmatrix} 0.44 \\ 0.56 \end{pmatrix}.$$

Рассчитываем $\beta(t_n)$ по формулам:

$$\beta(t_n) = \sum_{t_{n+1}} \beta(t_{n+1}) p(x_{n+1}|t_{n+1}) p(t_{n+1}|t_n)$$

$$\beta(t_N) = 1.$$

Тогда в нашей задаче:

$$\beta(t_3) = \beta \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.2 & 0 \\ 0 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \beta \begin{pmatrix} 0.188 \\ 0.68 \end{pmatrix} = \begin{pmatrix} 0.22 \\ 0.78 \end{pmatrix}$$

$$\beta(t_2) = \beta \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.8 & 0 \\ 0 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} 0.22 \\ 0.78 \end{pmatrix} = \beta \begin{pmatrix} 0.174 \\ 0.16 \end{pmatrix} = \begin{pmatrix} 0.52 \\ 0.48 \end{pmatrix}$$

$$\beta(t_1) = \beta \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 0.8 & 0 \\ 0 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} 0.52 \\ 0.48 \end{pmatrix} = \beta \begin{pmatrix} 0.384 \\ 0.16 \end{pmatrix} = \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix},$$

Где β – нормировочная константа.

Тогда итоговые маргинальные распределения:

$$\gamma_0(t_3) = p(t_3|X) = \frac{0.7 \cdot 0.44}{0.3 \cdot 0.56 + 0.7 \cdot 0.44} = 0.373$$

$$\gamma_0(t_2) = p(t_2|X) = \frac{0.52 \cdot 0.9}{0.1 \cdot 0.48 + 0.52 \cdot 0.9} = 0.906$$

$$\gamma_0(t_3) = p(t_3|X) = \frac{0.22 \cdot 0.8}{0.78 \cdot 0.2 + 0.22 \cdot 0.8} = 0.831$$

Задача 16

Рассматривается игра «Морской бой». В квадрате размера 3×3 возможны 2 ситуации: один двухпалубный корабль и два двухпалубных корабля. С помощью построения тупикового теста найти минимальное число ходов, необходимое для гарантированного определения того, какая из двух ситуаций имеет место.

Решение:

Для начала запишем таблицы всех возможных положений кораблей каждого из классов K1 (1 двухпалубный) и K2 (2 двухпалубных). Поле 9×9 будем представлять как вектор длины 9, соответствующий следующей нумерации клеток:

```
1 2 3
4 5 6
7 8 9
```

Найдём таблицу попарных различий строк из таблиц первого и второго классов.

Далее применяем стандартный метод. Для каждой строки вида $[z_1 z_2 \dots z_9]$ записываем соответствующую дизъюнкт: $D_k = x_{i_1} V x_{i_2} V \dots V x_{i_{n_k}}$, в которую включается переменная $x_i \leftrightarrow z_i = 1$. Рассматриваем функцию $F = \bigwedge_k D_k$, в которой получаем 96 дизъюнктов. Упрощение с использованием преобразований $D * (D V D') = D$ производилось с использованием компьютера.

В итоге получилось $F = (x_8 V x_9)(x_7 V x_8)(x_1 V x_2)(x_2 V x_3)(x_3 V x_6)(x_6 V x_8)(x_1 V x_4)(x_4 V x_7) = (x_8 V x_7 x_9)(x_2 V x_1 x_3)(x_6 V x_3 x_9)(x_4 V x_1 x_7) = (x_2 x_4 x_6 x_8) V (x_1 x_3 x_7 x_9) V F'$, где F' очевидно содержит только конъюнкты длины > 4 . Значит кратчайшими тупиковыми тестами будут 2-4-6-8 и

1-3-7-9, что соответствует: * * * и * * *. Очевидно, в первом случае при 2-х

попаданиях получаем 2 корабля, при одном – один. Во втором – 1 корабль при 0 и 1 попадании, 2 корабля – при 2.

Код на Матлабе для исключения лишних дизъюнктов прилагается:

```
a = [
1 1 0 0 0 0 0 0 0
0 1 1 0 0 0 0 0 0
0 0 0 1 1 0 0 0 0
0 0 0 0 1 1 0 0 0
0 0 0 0 0 0 1 1 0
0 0 0 0 0 0 0 1 1
1 0 0 1 0 0 0 0 0
0 0 0 1 0 0 1 0 0
0 1 0 0 1 0 0 0 0
0 0 0 0 1 0 0 1 0
0 0 1 0 0 1 0 0 0
0 0 0 0 0 1 0 0 1];
```

```
b = [
1 1 0 0 0 0 1 1 0
1 1 0 0 0 0 0 1 1
0 1 1 0 0 0 0 1 1
0 1 1 0 0 0 1 1 0
1 0 1 1 0 1 0 0 0
1 0 0 1 0 1 0 0 1
0 0 1 1 0 1 1 0 0
0 0 0 1 0 1 1 0 1];
```

```
na = size(a, 1);
nb = size(b, 1);
```

```
m = size(a, 2);
c = zeros(na*nb, m, 'uint8');
```



```

for i = 1 : size(a, 1)
    for j = 1 : size(b, 1)
        c((i-1)*nb + j, :) = (a(i,:) ~= b(j, :));
    end
end

finish = false;
while ~finish
    finish = true;
    n = size(c, 1);
    for i = 1 : n
        for j = 1 : i-1
            if (c(i, :) <= c(j, :))
                c(j, :) = [];
                finish = false;
                break;
            elseif (c(j, :) <= c(i, :))
                c(i, :) = [];
                finish = false;
                break;
            end
        end
        end
        if finish == false
            break;
        end
    end
end
end

```

Задача 17

Найти результирующую ДНФ для системы тестовых уравнений

$$\begin{cases} x_1 \vee x_2 = 1 \\ x_2 \vee x_3 = 1 \\ x_3 \vee x_4 = 1 \\ \dots \\ x_{n-1} \vee x_n = 1 \end{cases}$$

Решение:

Рассмотрим д. н. ф. $\bigvee_k x_{i_1} x_{i_2} \dots x_{i_{n_k}}$, где индексы i_1, \dots, i_{n_k} – всевозможные комбинации, удовлетворяющие условиям:

1) $0 < i_1 < \dots < i_{n_k} < n$

2) Среди i_j нет трёх подряд идущих чисел (считая, что индексы i_0 и i_{n_k+1} присутствуют всегда).

3) $i_j - i_{j-1} \leq 2, j = 0, 1, \dots, n_k + 1, i_0 = 0, i_{n_k+1} = n + 1$.

Очевидно, если д.н.ф. = 1, то все уравнения системы будут выполняться. Действительно, в д.н.ф. найдётся конъюнкт $K = x_{i_1} x_{i_2} \dots x_{i_{n_k}} = 1$, у которого индексы i_1, \dots, i_{n_k} удовлетворяют условиям 1)-3). Из условия 3) сразу же следует выполнимость каждого уравнения системы.

Осталось показать, что д.н.ф. – тупиковая. Поскольку правило склейки неприменимо (нет отрицаний переменных), все конъюнкты различны и правило поглощения не применимо, то по теореме Блейка-Квайна д.н.ф. является максимальной. Невозможность применить правила поглощения ($K \vee K * K' \Rightarrow K$) следует из того, что характеристические вектора наборов индексов различных конъюнктов несравнимы между собой. Под характеристическим вектором конъюнкта понимается булев вектор, где 1 стоят на тех (и только тех) местах, которые соответствуют номерам индексов переменных, входящих в этот конъюнкт. А несравнимы они потому, что при добавлении хотя бы одной новой единицы в характеристический вектор какого-либо конъюнкта приведёт к нарушению свойства 2).

Значит построенная таким образом д.н.ф. является тупиковой д.н.ф., ч.т.д

Задача 18

В обучающей таблице класс K_1 состоит из всех векторов, принадлежащих шару радиуса 3 с центром в $(0, 0, \dots, 0)$, а класс K_2 состоит из всех векторов, принадлежащих шару радиуса 4 с центром в $(1, 1, \dots, 1)$. К какому классу будет отнесен объект $(0, 1, 0, \dots, 1)$ алгоритмом «Кора», если n – четное?

Решение:

Утверждение: Для любого набора из 3 столбцов (i_1, i_2, i_3) и любой тройки чисел (k_1, k_2, k_3) можно найти такой вектор U из K_1 и V из K_2 , что $U_{i_1} = V_{i_1} = k_1$, $U_{i_2} = V_{i_2} = k_2$, $U_{i_3} = V_{i_3} = k_3$.

Доказательство:

1) Для класса K_1 : берём вектор U : $U_i = k_j$ для $i = i_j$ и $U_i = 0$ для остальных i . В нем не больше трёх единиц (по построению), следовательно, U отстоит от $(0, 0, \dots, 0)$ не более, чем на 3 и принадлежит K_1 .

2) Для класса K_2 : берём вектор V : $V_i = k_j$ для $i = i_j$ и $U_i = 1$ для остальных i . В нем не больше трёх нулей (по построению), следовательно, V отстоит от $(1, 1, \dots, 1)$ не более чем на 3 и принадлежит K_2 .

Следствие: Алгоритм «Кора» не сможет дать ответ ни для какого входного вектора, т.к. не сможет найти ни одной тройки столбцов и тройки чисел, позволяющих отличить один класс от другого.

Задача 19

В обучающей таблице класс K_1 представлен объектами $(0,0,\dots,0,0)$ и $(1,1,\dots,1,1)$, а класс K_2 – объектами $(1,0,1,0,\dots)$ и $(0,1,0,1,\dots)$. Тестовый объект имеет вид $\left(\underbrace{1,1,\dots,1}_k, \underbrace{0,0,\dots,0}_k \right)$. К какому классу будет отнесен этот объект алгоритмом «Кора» при четном и нечетном n ?

[тут сомнительное место в условии. Мы его подправили, сообразуясь с логикой]

Решение:

Считаем, что $n = 2k$. В тестовом объекте k нулей и k единиц.

В случае $k = 2t$ получаем:

- количество троек, совпадающих с K_1 : $\Gamma_1 = (C_t^2 * C_t^1 + C_t^1 * C_t^2) * 2$. При этом выбирается 1 чётная и 2 нечётных позиции из 1-й половины вектора; 2 чётных и 1 нечётная из первой; умножаем на 2, поскольку можем выбрать также из второй половины.
- количество троек, совпадающих с K_2 : $\Gamma_2 = (C_t^2 * C_t^1 + C_t^1 * C_t^2) * 2$. При этом выбирается 1 чётная позиция из первой половины, 2 чётных – из второй; 2 чётных из первой, 1 – из второй; умножаем на 2, поскольку можем заменить чётные позиции на нечётные.

В итоге, получаем равенство голосов и объект не будет отнесён ни к какому классу.

В случае $k = 2t+1$ получаем:

- количество троек, совпадающих с K_1 : $\Gamma_1 = (C_{t+1}^2 * C_t^1 + C_{t+1}^1 * C_t^2) * 2$. (Здесь выбираем индексы как и в случае $k = 2t$).
- количество троек, совпадающих с K_2 : $\Gamma_2 = (C_{t+1}^2 * C_{t+1}^1 + C_{t+1}^1 * C_{t+1}^2) + (C_t^2 * C_t^1 + C_t^1 * C_t^2)$. (Здесь выбираем аналогично случаю $k = 2t$, только при выборе из второй половины вместо чётных позиций берём нечётные (поскольку $k = 2t+1$)).

$$\Gamma_1 = (t+1)t^2 + (t+1)t(t-1)$$

$$\Gamma_2 = (t+1)^2t + (t-1)t^2$$

В итоге, $\Gamma_1 - \Gamma_2 = -(t+1)t + t(t-1) = -2t$. То есть выигрывает второй.

Задача 20

Написать формулу для числа голосов в алгоритме вычисления оценок, если функция близости определяется параметрами $\epsilon_1, \dots, \epsilon_n$, допустимое число невыполняющихся неравенств $q = 3$, а совокупность характеристических векторов опорных множеств образует интервал конъюнкции $x_1 \& \dots \& x_r \& \overline{x_{r+1}} \& \dots \& \overline{x_{r+k}}$.

Решение:

[теория – см. лекции Журавлёва, 2008, стр 16..18. Обозначения взяты оттуда же]
Предположим (иное не указано в условии), что веса всех признаков и экземпляров классов одинаковы ($=1$). Количество выполнившихся неравенств:

$$N'(\omega S, \omega S_i) = \sum_{j=1}^{|\omega|} (|a_{u_j} - a_{j u_i}| \leq \epsilon_j)$$

Индикатор близости векторов по заданному множеству столбцов:

$$N(\omega S, \omega S_i) = \begin{cases} 1, & N'(\omega S, \omega S_i) \geq |\omega| - q \\ 0, & \text{иначе} \end{cases}$$

Кол-во голосов для заданной строки:

$$\Gamma_{ij}(S) = \sum_{\omega \in \Omega} N(\omega S, \omega S_i)$$

Здесь Ω – множество векторов вида $(u_1 \dots u_l)$ – подмножеств векторов номеров столбцов, в которые входят столбцы $1..r$ и не входят $(r+1)..k$

Итоговое количество голосов за класс K_j :

$$\Gamma_j(S) = \sum_{i: S_i \in K_j} N(\omega S, \omega S_i)$$

Задача 21

В алгоритме вычисления оценок $x_{11} = 1$; $x_{10} = x_{01} = x_{00} = 0$. Написать формулу для числа голосов, если функция близости определяется параметрами $\varepsilon_1, \dots, \varepsilon_n$, допустимое число невыполняющихся неравенств равно q , а система опорных множеств состоит из всех подмножеств мощности $2q$.

Решение:

[теория – см. лекции Журавлёва, 2008, стр 16..18. На стр 17 есть странная фраза про «легко видеть» и $q/2$, что заставляет подозревать здесь где-то подколку]

Фактически, задача сводится к предыдущей с минимальными изменениями.

Предположим (иное не указано в условии), что веса всех признаков и экземпляров классов одинаковы ($=1$).

Количество выполнившихся неравенств:

$$N'(\omega S, \omega S_i) = \sum_{j=1}^{2q} (|a_{u_j} - a_{j u_i}| \leq \varepsilon_j)$$

Индикатор близости векторов по заданному множеству столбцов:

$$N(\omega S, \omega S_i) = \begin{cases} 1, & N'(\omega S, \omega S_i) \geq 2q - q \\ 0, & \text{иначе} \end{cases}$$

Кол-во голосов для заданной строки:

$$\Gamma_{ij}(S) = \sum_{\omega \in \Omega} N(\omega S, \omega S_i)$$

Здесь Ω – множество векторов вида $(u_1 \dots u_l)$ – подмножеств векторов номеров столбцов, каждый вектор длины $2q$

Итоговое количество голосов за класс K_j : так как x_{ij} для $(i, j) \neq (1, 1)$ равно нулю, лишние слагаемые убираются, и остается

$$\Gamma_j(S) = \sum_{i: S_i \in K_j} N(\omega S, \omega S_i)$$

Упражнение 1

Доказать, что матрица $\Phi^T \Phi \in \mathbb{R}^{m \times m}$ при $m > n$.

Решение:

Дополним матрицу $\Phi \in \mathbb{R}^{n \times m}$ до квадратной нулевыми строками:

$$\hat{\Phi} = \begin{pmatrix} \Phi \\ \bar{0} \end{pmatrix}$$

Тогда:

$$\hat{\Phi}^T = (\Phi^T \bar{0})$$

Произведение полученных матриц в точности совпадает с исходным:

$$\hat{\Phi}^T \hat{\Phi} = \Phi^T \Phi$$

однако они являются квадратными, значит, можем воспользоваться свойством определителя:

$$\det AB = \det A \cdot \det B$$

В силу вырожденности матриц $\hat{\Phi}$ и $\hat{\Phi}^T$ матрица $\hat{\Phi}^T \hat{\Phi}$ также является вырожденной.