ПРИМЕНЕНИЕ МЕТРИЧЕСКИХ ПРОСТРАНСТВ

Метрические пространства играют основную роль в большинстве разделов математики. Рассмотрим некоторые из них.

1. Пусть А – некоторое множество, A^* – семейство из всех конечных последовательностей элементов из А, допускающее пустую последовательность. Будем интерпретировать А как алфавит некоторого языка. Тогда элементы из А естественно называть *буквами*, а элементы из A^* *словами*. *Редакторской операцией* на A^* называется преобразование слова, состоящее из исключения одной буквы, либо вставки буквы в слово, или в замене одной буквы на другую. Наименьшее число редакторских операций, необходимых для перехода от одного слова к другому, называется *расстоянием Левинштейна* и задает метрику на A^*.

2. Рассмотрим задачу о передаче информации по коммуникациям. В компьютерной технике информация передается с помощью последовательности битов, принимающих два значения 0 или 1. Биты будем рассматривать в качестве букв. Рассмотрим всевозможные упорядоченные последовательности длины N. Обозначим это множество через I^N. Мощность этого множества 2^N. В этом множестве выделим подмножество A^N, состоящее из значащих слов. Множество I^N является метрическим пространством относительно метрики

$$\rho(x,y) = \sum_{i=1}^{N} |x_i - y_i|,$$

где $x = (x_1, x_2, \dots, x_N)$, $y = (y_1, y_2, \dots, y_N)$, при этом x_i, y_i принимают два значения 0 или 1. Данная метрика называется *расстоянием Хемминга*.

Расстояние Хемминга используется для построения кодов с исправлением ошибок. Предположим, что в процессе передачи информации некоторые биты были искажены, т.е. произошла замена 1 на 0 или наоборот. Это означает, что произошла ошибка передачи.

Например, нам необходимо передать два слова, закодированных 1 и 0. В этом случае $N = 1$. При передаче информации такой код ничем не защищен. Можно усложнить код и дублировать каждое слово – 11 и 00. Тогда при однократной ошибке при передаче мы сможем констатировать ошибку, если получили сообщение вида 10 или 01. Если кодировать сообщение тремя битами, т.е. передавать 111 или 000, то ошибку можно исправить.

Пусть для элементов множества A^N выполняется неравенство $\rho(x,y) > d, x \neq y$. Тогда для того, чтобы код мог корректировать n ошибок при передаче, достаточно выполнения условия $d \geq 2n+1$. Из этого соотношения следует, что при кодировании слов нужно выбирать коды как можно дальше отстоящие друг от друга.

3. Пусть X – произвольное метрическое пространство. Для каждого $x \in X$ и не пустого $A \subset X$ положим

$$\rho(x,A) = \inf_{a \in A} \rho(x,a).$$

Для произвольных множеств $A, B \subset X$ положим

$$d_H(A,B) = \max\{\sup_{a \in A} \rho(a,B), \sup_{b \in B} \rho(b,A)\}.$$

Если рассмотреть множество $H(X)$ всех замкнутых ограниченных подмножеств в X, то на $H(X)$ функция $d_H(A,B)$ определяет метрику, которая называется *расстоянием Хаусдорфа*. Расстояние Хаусдорфа используется в задачах распознавания изображений цифровых сигналов.

Пусть задан алфавит изображений-образцов и одно тестовое изображение. Задача состоит в том, чтобы для тестового изображения найти наиболее похожее на него изображение-образец. Степень схожести определяется метрикой Хаусдорфа. Применение метрики Хаусдорфа в задаче распознавания изображений основано на сопоставлении изображениям замкнутых ограниченных множеств. Заметим, что каждому растровому черно-белому изображению соответствует матрица $A = (a_{ij})$ размера $n \times m$ такая, что $a_{ij} = 1$, если пиксель изображения с координатами (i,j) отличен от фона или $j = 1$. В противном случае $a_{ij} = 0$. Каждую матрицу можно соотнести с дискретным множеством точек на плоскости.

Метрика Хаусдорфа также применяется для анализа сходства траекторий в таких областях как экология, биология, геоинформатика, телемеханика. Анализ траекторий в телемеханике позволяет решить задачу об оптимизации маршрутов, навигации.

Рассмотрим еще одну задачу на метрику Хаусдорфа. Пусть имеются две дороги (непрерывные кривые, содержащие концы) Γ_1 и Γ_2. По дороге движется машина (точка), поливающая их. Будем считать, что машина поливает вокруг себя площадь замкнутого круга, причем радиус полива можно поменять. Найдем наименьший радиус, при котором при движении машины по одной из дорог будут политы обе.

Если точка x движется по Γ_1, то чтобы полить точку $y \in \Gamma_2$, нужно выбрать радиус полива $R(y) = \min\limits_{x \in \Gamma_1} \rho(x, y)$, а чтобы полить все точки $y \in \Gamma_2$, нужно выбрать радиус

$$R_1(\Gamma_1, \Gamma_2) = \max_{y \in \Gamma_2} \min_{x \in \Gamma_1} \rho(x, y).$$

Аналогично, если точка x движется по Γ_2, то чтобы полить всю Γ_1, нужно выбрать радиус

$$R_2(\Gamma_1, \Gamma_2) = \max_{y \in \Gamma_1} \min_{x \in \Gamma_2} \rho(x, y).$$

Обе дороги будут политы, если

$$R(\Gamma_1, \Gamma_2) = \max\{R_1(\Gamma_1, \Gamma_2), R_2(\Gamma_1, \Gamma_2)\}.$$

Расстояние $R(\Gamma_1, \Gamma_2) = d_H(\Gamma_1, \Gamma_2)$ является метрикой Хаусдорфа.

4. В задачах кластерного анализа и классификации используется статистическое расстояние или *расстояние Махаланобиса*, с помощью которого определяется сходство образов и классов. Оно отличается от расстояния Евклида тем, что учитывает дисперсии переменных и определяется по формуле

$$d_M(A, B) = \sqrt{(x - y)^T S^{-1} (x - y)},$$

где $x = (x_1, x_2, \dots, x_N), y = (y_1, y_2, \dots, y_N)$ элементы пространства $\mathbb{R}^N$, S – матрица ковариации. Расстояние Махаланобиса можно определить как меру несходства двух случайных величин x и y из одного распределения вероятностей с матрицей ковариации S.