18. Взаимная информация и ее свойства. Примеры вычисления. Информационная емкость и пропускная способность. Теоремы кодирования. Симметричные каналы. Пропускная способность гауссовского канала.

http://books.ifmo.ru/file/pdf/723.pdf

Для заданного произведения $XY = \{(x, y), p(x, y)\}$ дискретных ансамблей X и Y нужно количественно измерить информацию об элементах $x \in X$ входного ансамбля, содержащуюся в выходных символах $y \in Y$. Подходящей мерой такой информации является взаимная информация, определяемая для любых пар $(x, y) \in XY$ соотношением $I(x; y) = I(x) - I(x|y)$ (где $I(x) = -\log p(x)$ – собственная информация, $H(X) = \mathsf{M}[-\log p(x)] = -\sum_{x \in X} p(x) \log p(x)$ – энтропия, $I(x|y) = -\log p(x|y)$ – условная собственная информация).

Уменьшаемое $I(x)$ представляет собой количество собственной информации, содержащейся в сообщении x. Вычитаемое – условная собственная информация при известном y, иными словами, это количество информации, оставшейся в x после получения y. Разность $I(x; y)$ представляет собой изменение информации в x благодаря получению y.

Свойства взаимной информации:

1. Симметричность: $I(x; y) = I(y; x)$.

2. Если x и y независимы, то $I(x; y) = 0$.

Средней взаимной информацией между ансамблями X и Y называется величина $I(X; Y) = \mathsf{M}[I(x; y)]$, где $I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x|y)}{p(y)}$.

Свойства:

1. Симметричность $I(X; Y) = I(Y; X)$

2. Неотрицательность $I(X; Y) \geq 0$

3. Тождество $I(X; Y) = 0$ тогда и только тогда, когда ансамбли X и Y независимы.

4. $I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(XY)$

5. $I(X; Y) \leq \min\{H(X), H(Y)\}$

6. $I(X;Y) \leq \min\{\log|X|, \log|Y|\}$
7. $I(X;Y)$ – выпуклая вверх функция $p(x)$
8. $I(X;Y)$ – выпуклая вних функция $p(y|x)$

Информационная емкость и пропускная способность

Рассмотрим дискретный стационарный канал. Мы знаем, что количество информации о входных символах X канала, содержащееся в выходных символах Y определяется средней взаимной информацией $I(X;Y)$. Это верно при передаче одного символа канала. При использовании кодов длины n количество информации, получаемой декодером при передаче доного слова в среднем составит $I(X^n;Y^n)$ бит, что соответствует скорости передачи информации $\frac{1}{n}I(X^n;Y^n)$ бит/символ канала. Эта величина зависит от переходных вероятностей $\{p(y|x), y \in Y^n, x \in X^n\}$ и от распределения вероятностей на входе канала $\{p(x), x \in X^n\}$. Результирующая скорость получается $\max\limits_{\{p(x)\}} \frac{1}{n}I(X^n;Y^n)$ бит/символ канала. Величина $C_0 = \sup\limits_{n} \max\limits_{\{p(x)\}} \frac{1}{n}I(X^n;Y^n)$ называется *информационной емкостью* канала.

Допустим, у нас есть код канала длины n, количеством последовательностей M (*мощность* кода), $R = \frac{\log M}{n}$ – скорость кода. Число C называется *пропускной способностью канала*, если при любой скорости кода $R < C$ существуют коды, обеспечивающие сколь угодно малую вероятность ошибки и, напротив, при $R > C$ существует константа $\varepsilon > 0$, что вероятность ошибки любого кода ограничена снизу величиной ε.

Обратная теорема кодирования утверждает, что информационная ёмкость C_0 ограничивает сверху скорость, при которой достижима сколь угодно малая вероятность ошибки. Прямая теорема кодирования утверждает, что при скорости, сколь угодно близкой к C_0 достижима сколь угодно малая вероятность ошибки.

Теоремы кодирования

Обратная теорема кодирования. Для дискретного стационарного канала с информационной емкостью C_0 для любого $\delta > 0$ существует число $\varepsilon > 0$, такое, что для любого кода со скоростью $R > C_0 + \delta$ средняя вероятность ошибки удовлетворяет неравенству $\overline{P_e} \geq \varepsilon$, где $\overline{P_e} = \frac{1}{N}\sum\limits_{i=1}^{N} P_{ei}$ (то есть среднее по N сообщениям).

В частности, для дискретных постоянных каналов ёмкость вычисляется по формуле $C_0 = \max\limits_{\{p(x)\}} I(X;Y)$, а для двоичного симметричного канала информационная емкость вычисляется по формуле $C_0 = 1 - \eta(p)$

Прямая теорема кодирования. Для дискретного постоянного канала с информационной емкостью C_0 для любых $\varepsilon, \delta > 0$ существует достаточно большое число n_0 такое, что для любого натурального числа $n > n_0$ существует код длины n со скоростью $R \geq C_0 - \delta$, средняя вероятность ошибки которого $P_e \leq \varepsilon$.

Непрерывные каналы дискретного времени

Пусть модель канала задается условными плотностями распределения вида $f(\overline{y}|\overline{x}) = f(y_1, \ldots, y_n|x_1, \ldots, x_n)$. Пусть это канал стационарный и без памяти. Также допустим, что у нас *аддитивный* шум, то есть $\overline{y} = \overline{x} + \overline{z}$, $f(\overline{y}|\overline{x}) = f(\overline{z})$. Такая модель наызвается *каналом с аддитивным шумом.*

Энергия входной последовательности – $\sum\limits_{i=1}^{n} x_i^2$. Мощность на отсчет – $E(\overline{x}) = \frac{\sum\limits_{i=1}^{n} x_i^2}{n}$. Информационная емкость непрерывного стационарного канала дискретного времени с ограничением E на мощность входных сигналов – $C_0 = \sup\limits_{n} \max\limits_{f(\overline{x}):\mathsf{M}[E(\overline{x})]\leq E} \frac{1}{n} I(X^n;Y^n)$. В частности, для канала без памяти $C_0 = \max\limits_{f(x):\mathsf{M}[E(x)]\leq E} I(X;Y)$.

Канал называется *гауссовским*, если шум в канале имеет гауссовское распределение, в общем случае многомерное. Для одномерного случая $f(z) = \frac{1}{\sqrt{2\pi N_0}} e^{-\frac{z^2}{2N_0}}$. Для такого канала информационная ёмкость $C_0 = \frac{1}{2}\log(1 + \frac{E}{N_0})$.