1 September 6th, 2019

1.1 Language Models

Definition 1.1. A **statistical language model** is a model specifying probability distribution over word sequences. It can be regarded as a probabilistic mechanism for "generating" text, thus it's also called a "generative" model.

> **Remark 1.2** — In other words, if you have a language model, you can generate text.

LM's are useful as they provide a principled way to quantify the uncertainties associated with natural language, helping us with **speech recognition**, **text categorization**, and **information retrieval**.

We can also use conditional probability, e.g. given we've seen the word "basketball", it's likely we're talking about sports.

1.2 Review On Probability

$$P(x|y) = \frac{P(x,y)}{P(y)} = \frac{P(y|x)P(x)}{P(y)}.$$

Thus,

$$\max_x P(x|y) = \max_x P(y|x)P(x).$$