

统计序列模型

2018 年 5 月 21 日

这篇文章翻译了 Mark Gales 教授的对统计序列建模的报告 (Statistical Sequence Modelling)。

原始的报告可以从 <http://mi.eng.cam.ac.uk/~mjfg/sequence17-draft.pdf> 下载。

语音识别和语音合成都可以被分别视为机器学习中的分类和回归问题。语音处理之所以能成为一个独立的研究方向主要是因为需要对观测序列 $\mathbf{O}_{1:T}$ 和词序列 $\mathbf{w}_{1:L}$ 的序列性质进行研究。这篇报告探讨了如何从机器学习的框架来进行语音处理以及各种用来处理数据的时序性的方法。

1 语言模型：一种序列模型

本报告主要关注语音识别和语音合成中的序列模型，两者都是序列到序列的映射模型。首先测试了语言模型，它是一种单序列任务，从某些程度描绘了所使用的序列模型的性质。

在语言模型中，建模目标是将一个长度为 L 的词序列 $\mathbf{w}_{1:L} = w_1, \dots, w_L$ 映射为一个标量用来表示该语言模型在使用参数 $\mathbf{\Lambda}^1$ 生成词序列的概率。

¹为了公式的简洁，后文除了特殊需求，一律将参数省去