

# word2vec

---

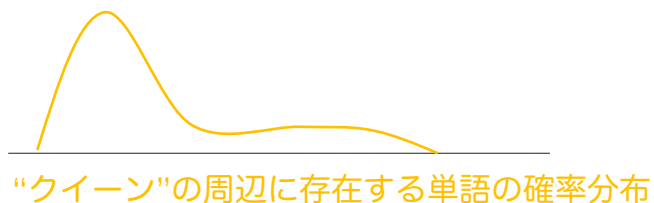
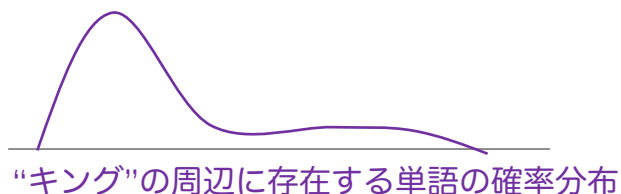
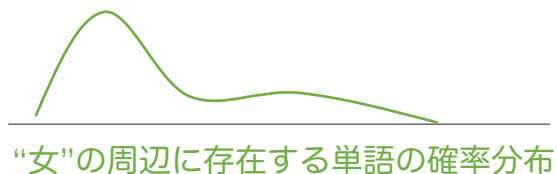
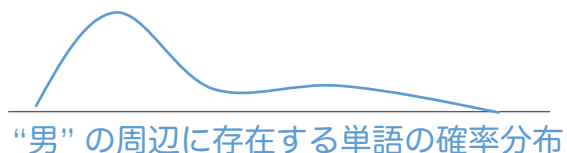
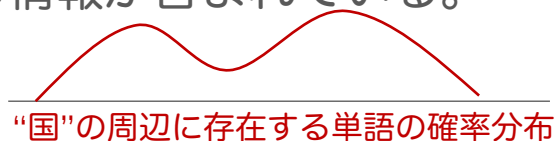
# 単語の加法構成性

---

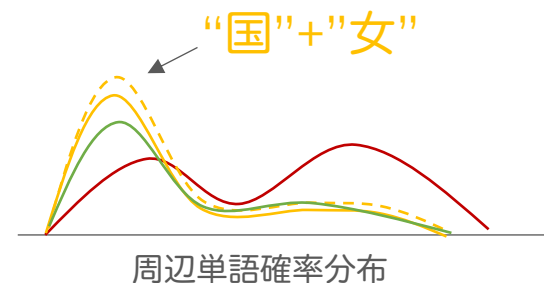
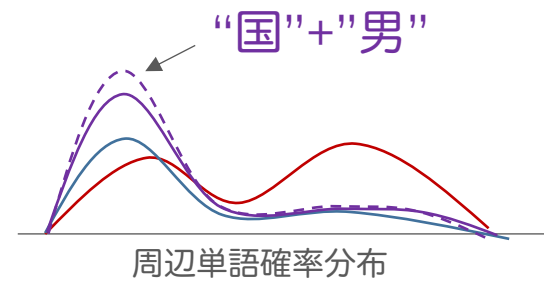
- 原著論文では、skip-gramにより、単語の加法構成性(additive compositionality)が得られたと報告されている。
  - Tomas Mikolov et al. Distributed Representations of Words and Phrases and their Compositionality. <https://arxiv.org/pdf/1310.4546.pdf>
- 単語の加法構成性とは、王 - 男 + 女 = 女王 のような計算ができることを指している。
- 単語の加法構成性が得られることの理由として、原著論文の5章では次のような直感的な説明がなされている。
  - Skip-gramで得られる単語埋め込みベクトルには、その単語の周辺単語の確率分布に関する情報が含まれており、単語埋め込みベクトルを足すということは、2つの周辺単語確率分布を掛けることに相当する。
  - この意味を次頁で図示する。

# 単語の加法構成性

Skip-gramで得られる単語埋め込みベクトルには、その単語の周辺単語の確率分布に関する情報が含まれている。



単語埋め込みベクトルを足すということは、2つの周辺単語確率分布を掛けることに相当する。



“国”+“男”という単語ベクトルは、ある文章で“国”と“男”が同時に出現していることに相当し、その場合、両方の単語で高い確率になる周辺単語は強め合い、両方の単語で低い確率になる周辺単語は弱め合う。