

# Embedding Layer

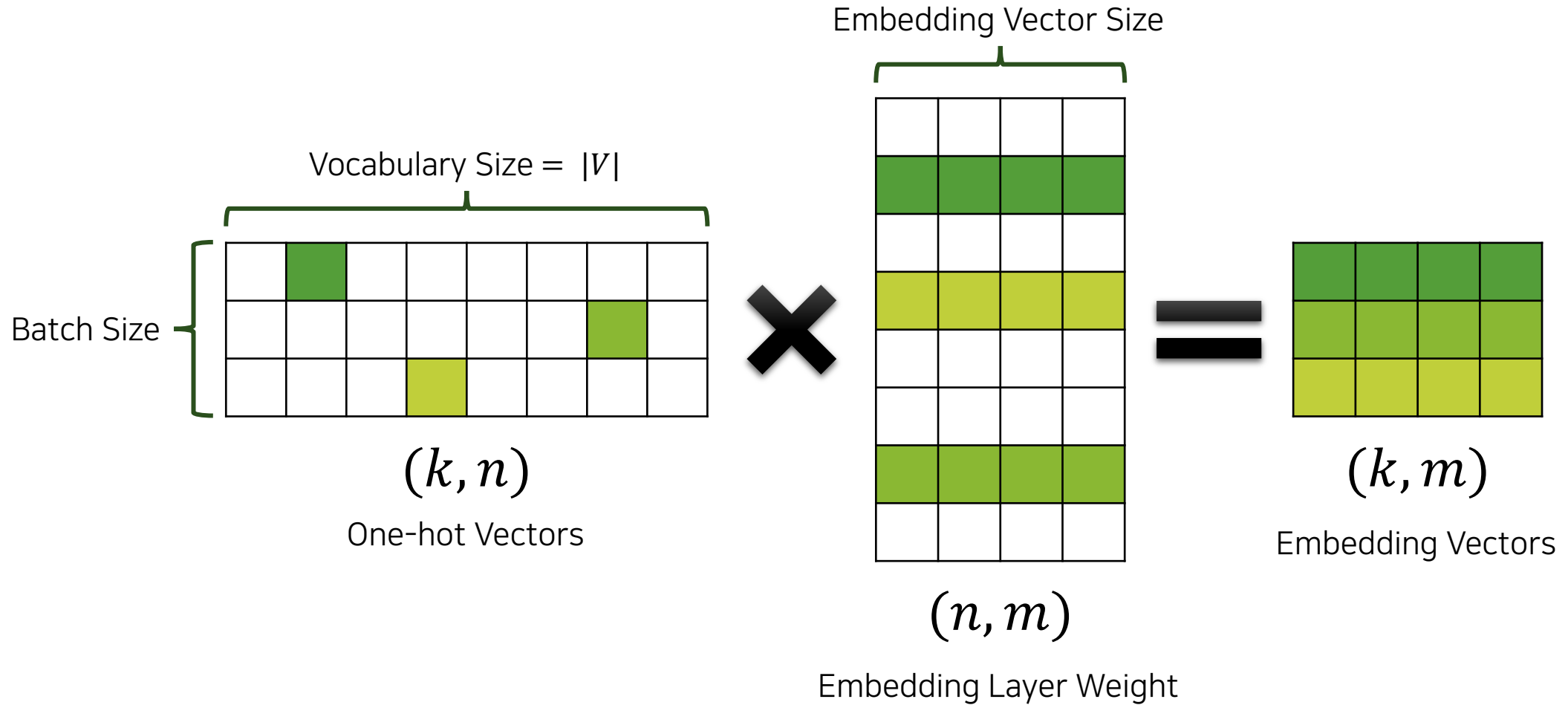
Ki Hyun Kim

[nlp.with.deep.learning@gmail.com](mailto:nlp.with.deep.learning@gmail.com)

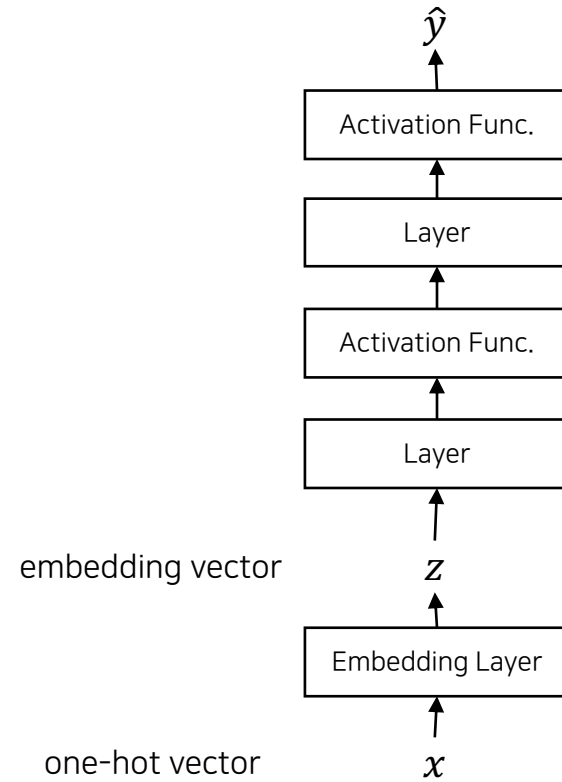
# Myth

- 미리 skip-gram등을 통해 계산한 word embedding vector를 네트워크에 넣어 주는 것이 아니다.
- 보통은 one-hot vector 자체를 넣어줘야 한다.
  - 즉, 단어 자체를 넣어주는 것
  - N개의 단어가 있다면, N+@차원의 벡터가 네트워크의 입력이 된다.

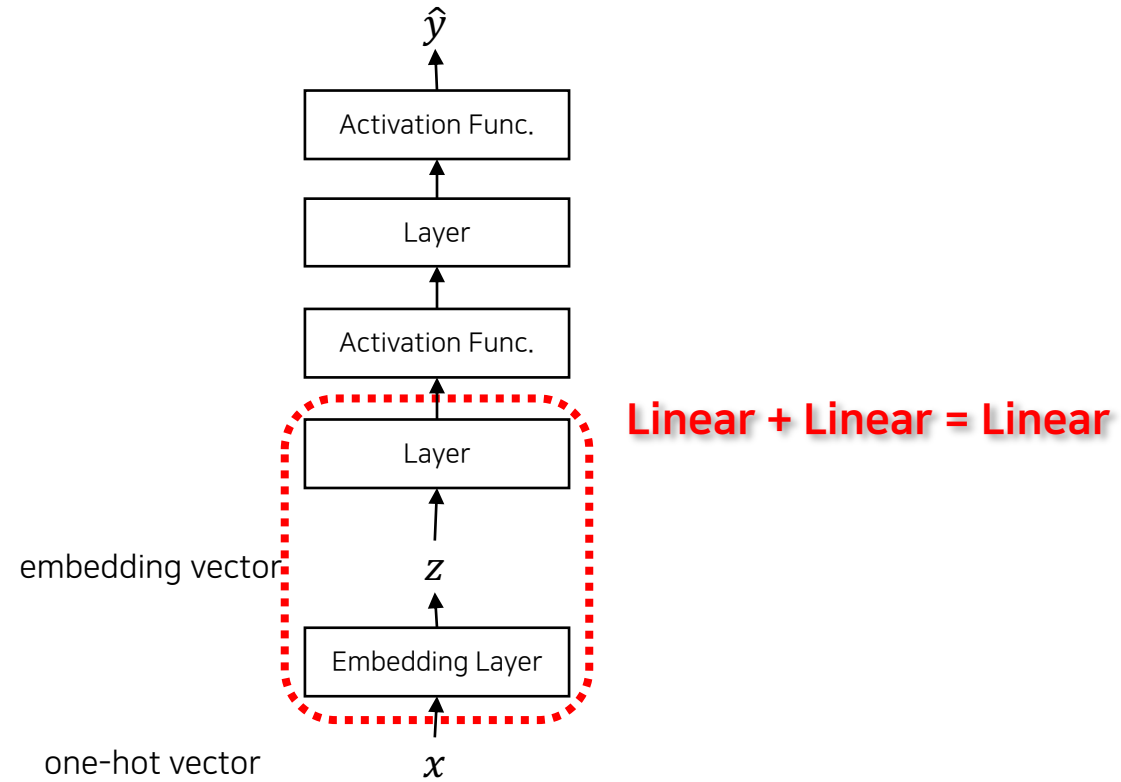
# Embedding Layer



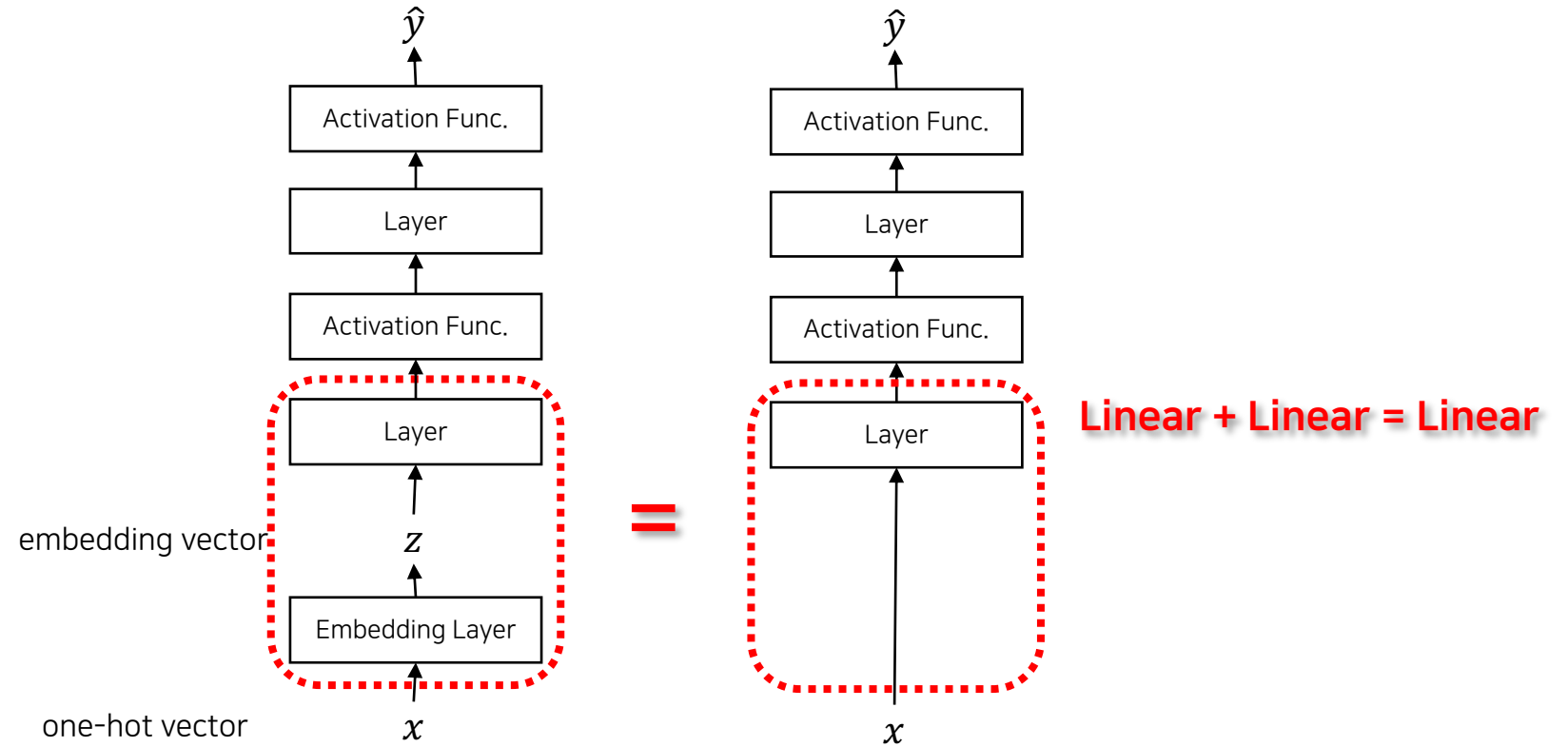
# Why Embedding Layer



# Why Embedding Layer

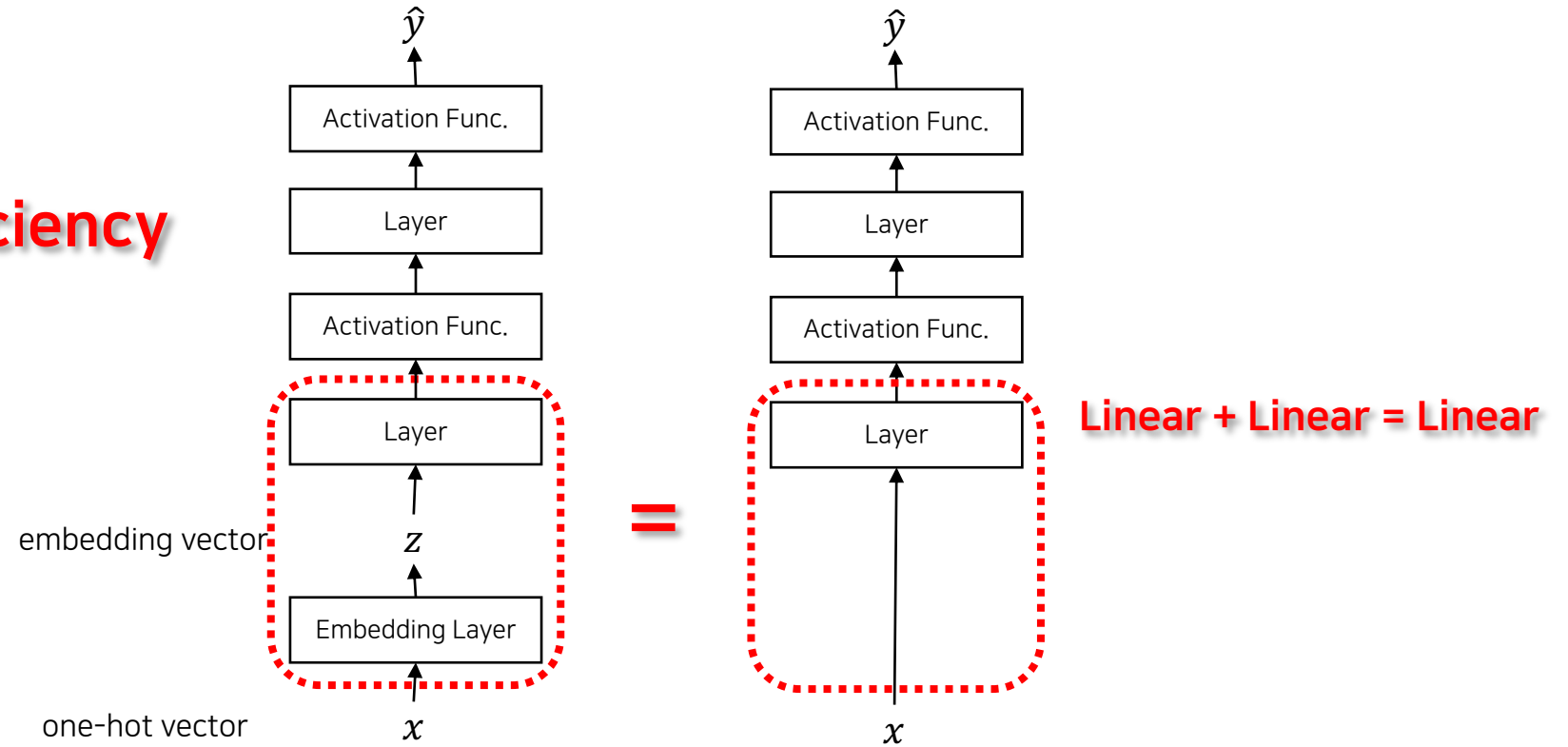


# Why Embedding Layer



# Why Embedding Layer

Why?  
Computation Efficiency



# Summary

- Embedding Layer를 통해 One-hot vector를 각 task에 맞는 dense representation으로 바꿀 수 있음
  - 단순히 fix된 word embedding vector 자체(e.g. word2vec)를 넣어줄 경우, 해당 알고리즘의 objective에 최적화된 representation을 넣어주는 것임
- 특수한 경우에 한하여,
  - fixed word representation을 사용하거나,
  - word embedding vector들로 구성된 weight parameter를 seed로 사용하여 최적화
- 하지만, baseline은 one-hot vector를 넣어주는 것!
- 그럼 word embedding algorithm은 언제 쓰나요?