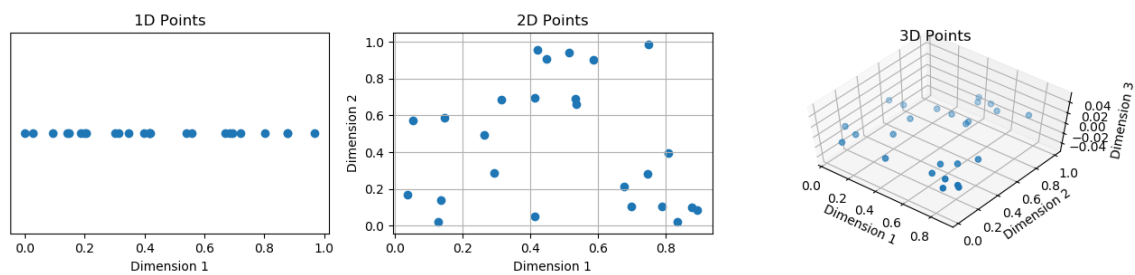


8장 공부 및 추가했으면 좋을 부분

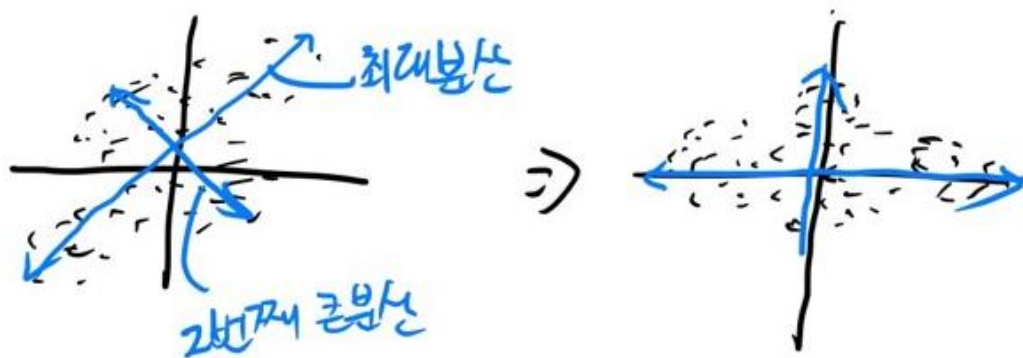
맨 앞에 넣으면 좋을거같아요

차원 축소가 왜 필요할까요?



위와 같이 차원이 늘어남에 따라 데이터 간의 거리가 Sparse해지는 것을 파악할 수 있습니다. 즉 특징이 늘어남에 따라, 필요한 데이터 수가 증폭되기 때문에 충분히 학습이 되지 않아 오버피팅 위험이 있습니다. 이에 따라, 차원의 수를 줄임으로써, 학습의 정확도를 늘릴 수 있습니다.

매니폴드랑 투영을 좀 더 어떻게 설명할까?



pca부분 그림 - 별론데 이해하기 좋은하기 좋은게 필요할 듯,,

kernel pca부분 설명이 앞에 있으면 좋을거 같아요. 물론 그림도인데 적합한거 찾기가 어렵네여

앞서 pca와 kernel pca의 차이는 무엇일까?

Pca는 기존 데이터의 차원에서 분산이 큰 축을 새롭게 지정하여 가져오는 것과 비슷하다. 하지만 완벽한 비선형으로 데이터의 원래 차원 d 에서 분산이 최대인 적절한 축을 발견할 수 없는 경우에는 사용할 수 없는 문제를 가진다. 이에 경우 kernel trick을 이용하여 데이터의 차원을 더 높은 공간으로 보내는 것이 kernel pca이다. 이 경우, 원 데이터의 특성에서 찾을 수 없던 분산이 큰 축을 더 높은 고차원 공간에서 발견할 수 있으며 이의 경우에 앞서 대입한 pca를 대입하는 것이다.