부분공간 근사화를 통한 빠르고 확장성있는 커널 주성분 분석법

(Fast and Scalable Kernel PCA via subspace approximation)

Woosang Lim CS.KAIST
Byungkon Kang CS.KAIST
Kyomin Jung CS.KAIST

01.30.2013

동기(Motivation)

- 이전 연구
  - Nystrom method 기반의 고유벡터(eigenvector) 근사 방법이 있다.
    - 시간 복잡도$O(k^2n)$, 공간 복잡도$O(kn)$

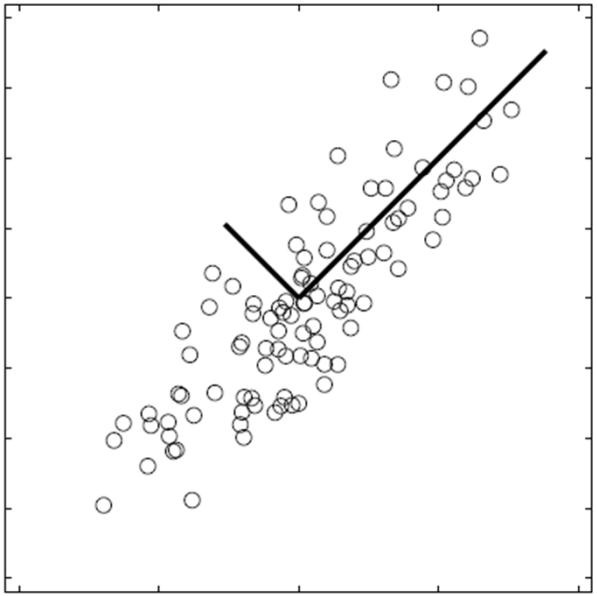
- 향상된 Nystrom method기반의 고유벡터 근사 방법 제시 (KPCA-S 알고리즘)
  - 튜닝 파라미터(tuning parameter)를 도입
    - 튜닝 파라미터 수치에 따라서 보다 정확한 근사 가능
  - 튜닝 파라미터에 따른 정확도 이론적인 보장

개요(Outline)

- 동기
- 백그라운드 정보
- 메인 알고리즘
- 실험

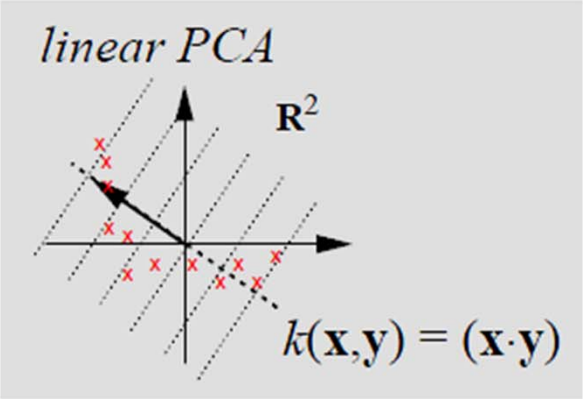
주성분 분석법(PCA)

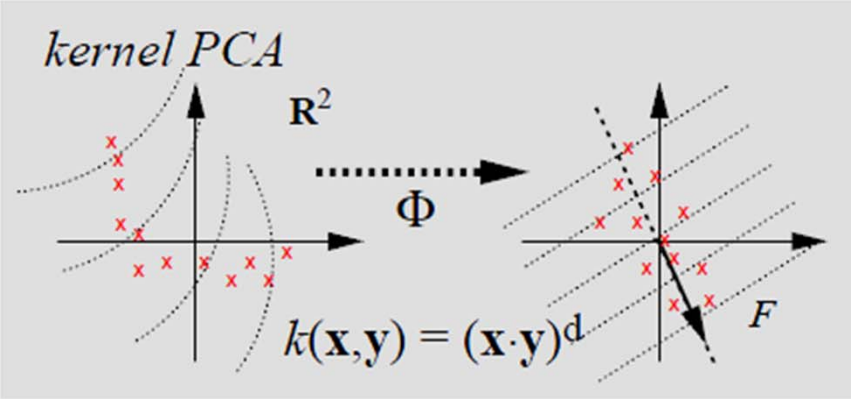
- 주성분 분석법(PCA)
  - 주성분 분석법은 데이터의 차원을 줄이는 방법
  - PCA 고차원 데이터를 낮은 차원으로 사영(projection)시킨다.
  - PCA는 데이터 분포를 가장 잘 대표하는 정규직교(orthonormal) 벡터를 계산한다.



커널 방법(kernel methods)

- 커널 방법(Kernel methods)
  - 커널 방법은 비선형(nonlinear) 데이터를 선형(linear)적으로 다룰 수 있게 고차원 유클리디안(Euclidean)공간으로 맵핑(mapping)한다.
  - 커널 방법은 양함수(explicit function) 대신에 커널이라고 부르는 유사도(similarity)함수를 쓴다.





k-평균 군집화(k-means clustering)

- k-평균 군집화(k-means clustering)
  - 서로 유사한 점들을 같은 클러스터(cluster)로 그룹을 만든다
    - 다른 클러스터에 속한 점들끼리는 유사성이 적게
  - 매 단계 마다 클러스터의 무게중심 점(centroids)을 계산
  - 매 단계 마다 점들을 가장 가까운 무게중심에 할당

개요(Outline)

- 동기
- 백그라운드 정보
- 메인 알고리즘
- 실험

KPCA-S(Kernel PCA via Subspace Approximation)

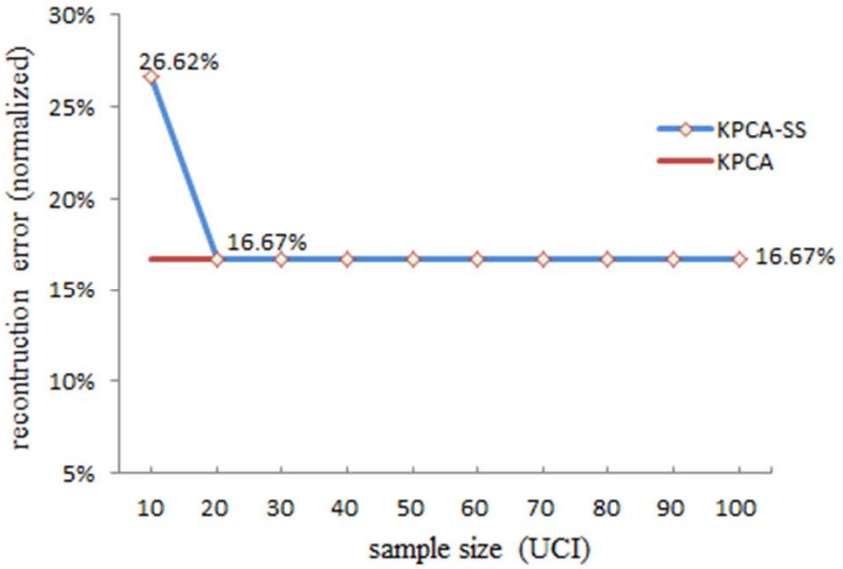
- 기저(basis)의 종류 따른 고유공간 근사

- 기저가 무게중심으로 이루어져 있을 때
  - 무게중심 근사 튜닝 파라미터 ℓ을 사용
  - ℓ 을 조절함에 따라 근사 정확도를 조절 할 수 있다.
    - ℓ에 따른 정확도 이론적 보장

- 기저가 샘플된 데이터 벡터로 이루어져 있을 때
  - 샘플 개수 조절 튜닝 파라미터 ρ 를 사용
  - ρ 를 조절함에 따라 근사 정확도를 조절 할 수 있다.
    - ρ 에 따른 정확도 이론적 보장

실험: UCI 데이터(n=10992, r=16), k=10

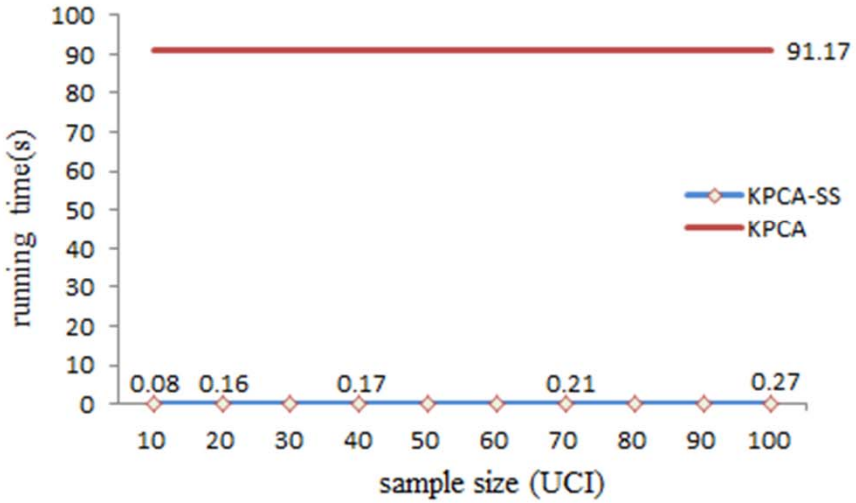
정확도



x 축: 샘플 개수 ρk

y 축: reconstruction error

러닝 타임

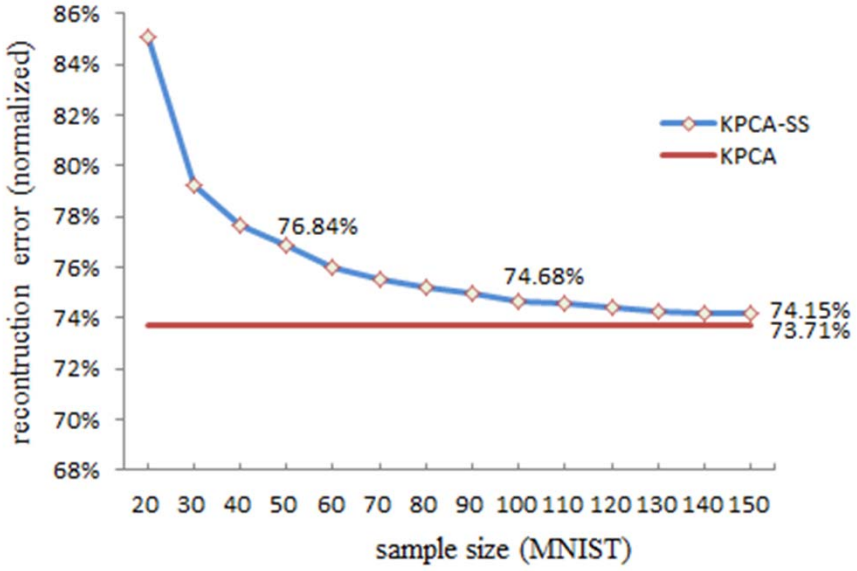


x 축: 샘플 개수 ρk

y 축: reconstruction error

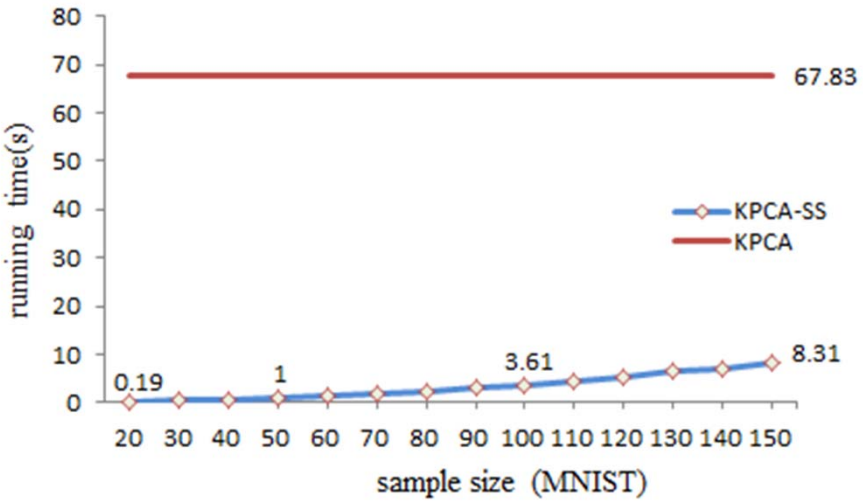
실험: MNIST 데이터(n=10000, r=784), k=10

정확도



x 축: 샘플 개수 ρk
y 축: reconstruction error

러닝 타임



x 축: 샘플 개수 ρk
y 축: reconstruction error

실험: ℓ에 따른 결과 비교, k=10

for $k = 10$. The results of KPCA-SC with parameter

UCI	$\ell = 30$	$\ell = 50$	Time
$E_{KPCA-SC}$	0.2517	0.2405	0.24s, 0.25s
E_{KPCA}	0.1667	0.1667	91.17s
MNIST	$\ell = 30$	$\ell = 50$	Time
$E_{KPCA-SC}$	0.7688	0.7590	0.75s, 0.77s
E_{KPCA}	0.7371	0.7371	67.83s

Table 1. NREs of KPCA-SC for UCI and MNIST dataset for $k = 9$ and $\ell = 30, 50$.

Thank you