Илија Вишинов 161078

Линеарна Алгебра

Principal Component Analysis – PCA

1. Алгоритам

Principal Component Analysis (PCA) е метод кој се користи за   
дефинирање на координатен систем (линеарна трансформација?) кои помага во

намалување на димензионалноста на многу димезионални податоци и генерирање на нивна покорисна визуелизација. Сето ова, со цел да се задржи количеството информација кое се содржи во множеството. Задржувањето на количеството информација се прави преку анализирањето на коваријансата на м-вото.

Потребата од ПЦА се истакнува во проблеми во кои сакаме да избегнеме curse of dimensionality кога потребата од бројот на примероци расте експоненцијално како што се зголемува димензијата на податоците.

1. Имплементација

За имплементација на алгоритмот е даден линк со кодот на github.

1. Услови за конвергенција
2. Теориска комплексност

Алгоритмот може да се сегментира на неколку чекори, чија комплексност се анализира.

Бидејќи PCA се состои од:

* Нормализација на податоците
  + Пресметување на средни вредности
  + Пресметување на стандардни девијации
  + Одземање на средните вредности
  + Делење со стандрадната девијација
* Пресметување на матрицата на коваријанса
  + Пресметување на производот на матрицата со податоци со нејзината транспонирана матрица (во обратен редослед).
  + Делење на секој елемент со бројот на редици
* Пресметување на сопствени вредности и сопствени вектори

1. Анализа над симулирани податоци

Time plot for different simulated data

1. Анализа над реални податоци