Motliuk Oleksander

Kovalevskyi Oleksander

Informatyka - II SPS

1. Jak działa metoda PCA?
2. Przykładowe obrazy przed i po kompresji.
3. Jaki stopień kompresji udało się uzyskać?
4. Principal Component Analysis (PCA) to procedura statystyczna stosowana w celu zmniejszenia liczby funkcji w zbiorze danych. Tutaj używamy PCA, aby zmniejszyć rozmiar zdjęć twarzy.

Analiza głównych składników znajduje liniowe kombinacje zmiennych, które wyjaśniają największą zmienność danych i mają najmniejszy błąd rekonstrukcji. Istnieje wiele głównych składników, a każdy główny składnik wyjaśnia pewien procent zmian w danych. Pierwszy główny składnik zawsze wyjaśnia największą zmienność, a następnie drugi itd. Zasadniczo można znacznie zmniejszyć rozmiar danych, zachowując większość informacji, stosując wystarczającą liczbę głównych składników, aby wyjaśnić 99% zmienności w Twoje dane.

Mając zbiór danych zawierający 100 przypadków (100 osób) charakteryzowanych przez 5 zmiennych (np. wzrost, waga, wiek, dochód, powierzchnia mieszkania) można przypuszczać, że zmienne „wzrost” i „waga” będą ze sobą silnie dodatnio skorelowane (gdyż im ktoś wyższy, tym więcej waży). Po to żeby uzyskać większą przejrzystość danych lub uniknąć powielania się danych (np. przy segmentacji klientów) czasami warto jest zastąpić dwie zmienne jedną zmienną – tak zwaną składową, którą można nazwać na przykład „wielkość”. Podobnie skorelowane będą ze sobą zmienne „dochód” i „powierzchnia mieszkania”, które być może można zastąpić czynnikiem „zamożność”.

2.



3. 99 %. Jak I było w przykładzie