

Thema: Singulärwertzerlegung - geometrische Intuition

Schwerpunkte und Vorkenntnisse

Schwerpunkte: Visualisierung der SVD und der PCA, deren Zusammenhang

Vorkenntnisse: Eigenzerlegung; Varianz; Kovarianz; mathematische Herleitung der Singulärwertzerlegung; der Zusammenhang zwischen der Singulärwertzerlegung einer Matrix A und der Eigenzerlegung der Matrix $A \cdot A^T$.

Konzeptbeschreibung

Der Konzept des Videos ist dadurch entstanden, dass ich selbst schon mehrfach aus verschiedenen Quellen über PCA gehört habe und gedacht habe, dass ich das Thema verstanden habe. Auf der anderen Seite hatte ich auch gedacht, dass ich Singulärwertzerlegung verstanden habe. Das Problem dabei war: ich konnte die beiden Konzepte nicht in Verbindung setzen. Sie werden häufig als zusammenhängend, ja manchmal fast schon synonym genannt (daher auch eine der Verständnisfragen). Aber einen Klick hat es bei mir persönlich nie gemacht. Ich habe vermutet, dass wenn ich diese Schwierigkeiten hatte, werde ich wahrscheinlich nicht die Einzige sein. Und dabei schien es mir natürlicher und logischer, die Präsentation so aufzubauen, dass man zuerst die gängigen Darstellungen der beiden Themen gezeigt bekommt. Dann entsteht vielleicht die Frage nach deren Zusammenhang von alleine und wird am Ende des Videos beantwortet. Dabei setze ich Kenntnisse der Begriffe Varianz und Kovarianz voraus, weil ich leider nicht genügend Zeit dafür hatte.

Auf dem Weg zu dem Verständnis des Zusammenhangs von PCA und SVD wollte ich unbedingt beide Konzepte in 3d Visualisieren (und zwar schrittweise): 2d sieht immer zu einfach und realitätsfremd aus und mehr als 3d können wir, Menschen, schlecht wahrnehmen. Ich finde, dass das, was man einmal gesehen hat, wird sich viel besser in dem Gedächtnis einprägen.

Die andere Verständnisfrage (#1) stelle ich, weil ich es wichtig fand, Grenzen zu setzen: was ist PCA und was sie nicht ist.

Beantwortung der Verständnisfragen

1. **Kann man PCA als Datenreduktionstechnik bezeichnen?** Ja. An sich reduziert PCA nichts. Der Zweck der Analyse ist es, eine Basis zu finden, in der die Daten so dargestellt werden können, dass entlang jeder weiteren Axe die Varianz sinkt. Diese Darstellung kann man natürlich danach dafür verwenden, die Dimensionalität der Daten zu reduzieren.
2. **Kann man SVD und PCA synonym verwenden?** Nein, das sind verschiedene Sachen. Das eine (SVD) ist eine allgemeine Methode der Zerlegung von Matrizen, die weitere Anwendungen hat, als PCA. PCA kann wiederum nicht nur durch SVD kalkuliert werden, z.B. durch Berechnung der Kovarianzmatrix und deren Eigenzerlegung.

Inhaltsübersicht

- Auf dem Weg zu SVD 00:00
- **Python Beispiel 1** 0:16, Funktion `slideEVD()`
- **Python Beispiel 1** 0:57, Funktion `slideEVnonSquareMatr()`
- **Schwerpunkt 1** 01:21
- **Python Beispiel 2** 01:48, Funktion `slideSVD2d()`
- **Python Beispiel 3** 02:13, Funktion `slideSVD3d()`
- **Schwerpunkt 2** 02:44
- **Python Beispiel 4** 03:55, Funktion `slidePCA()`
- **Verständnisfragen** 07:03

Quellen und Referenzen

- WR Skript Seiten 55-70
- [arXiv:1404.1100v1](https://arxiv.org/abs/1404.1100v1)
- <https://www.youtube.com/watch?v=rYz83XPxiZo>
- <https://www.youtube.com/watch?v=DG7YTIGnCEo>

Veröffentlichung

Ich bin einverstanden, dass mein Video als Unterrichtsmaterial für Studierende der TU im Kurs Wissenschaftliches Rechnen in der Zukunft verwendet wird: ✓

Falls das Video verwendet wird, soll die Verwendung anonym (ohne Nennung des Namens) erfolgen: ✗