Sparse Principal Component Analysis for Frequency Data

Tobias Bork

Institute for Numerical Simulation

December 6, 2019





Sparse Principal Component Analysis

Tobias Bork

ntroduction

PCA

Fundamentals

pharse r CA

Application

кетегепсеs

Dimensionality Reduction

Sparse Principal Component Analysis

Tobias Bork

Introduction

PCA

Fundamental

parse PCA

Application

- ► What are DR methods?
- Why are DR methods important / what is the goal of DR methods?
- ▶ When to use DR methods?

Central Idea

- ► Reduce dimensionality of a data set while retaining as much information as possible
- extract the most important information from the data table;
- compress the size of the data set by keeping only this important information;
- simplify the description of the data set; and
- analyze the structure of the observations and the variables
- ▶ first component has most variance

Sparse Principal Component Analysis

Tobias Bork

ntroduction

FCA

Idea

Mathematical Formulation

Limits of Usability

Fundamentals

parse PCA

pplication

Figure: Some description of the plots

Gewicht [kg]

Sparse Principal Component Analysis

Tobias Bork

troduction

CA

Idea

Mathematical Formulation

imits of Usability

undamentais

Mathematical Formulations

Let $\mathbf{X} \in \mathbb{R}^{n \times p}$ be a centered data matrix with n samples and p variables. We find the first principal axis by

$$v_1 = \underset{\|v\|_2 = 1}{\operatorname{arg \, max}} \operatorname{Var}[\mathbf{X}v] = \underset{\|v\|_2 = 1}{\operatorname{arg \, max}} v^T \mathbf{\Sigma} v$$

where $\mathbf{\Sigma} = \mathbf{X}^T \mathbf{X}$ is the sample covariance matrix. Then we can find the following principal axis successively

$$v_{k+1} = rg \max_{\|v\|=1} v^T \mathbf{\Sigma} v$$

subject to
$$v_{k+1}^T v_l = 0 \quad \forall 1 \le l \le k$$

The new, transformed variables are defined by $Z_i = \mathbf{X}v_i$

The principal axis can also be computed via the eigendecomposition of Σ .

$$\pmb{\Sigma} = \pmb{\mathsf{VLV}}^{\mathsf{T}}$$

where ${\bf L}$ is a diagonal matrix with eigenvalues λ_i and ${\bf V}$ is the matrix of eigenvectors. Closely related is the Singular Value Decomposition (SVD)

$$X = UDV^T$$

where **D** is a diagonal matrix with singular values d_1, \ldots, d_p , **U** a $n \times p$ and **V** a $p \times p$ orthogonal matrix.

Sparse Principal Component Analysis

Tobias Bork

ntroduction

PCA

Idea

Mathematical Formulations

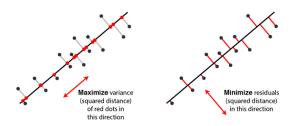
Theorems

Application

unuamentais

......

PCA as a regression problem



Suppose we want to extract the first k principal axis.

$$\hat{\mathbf{V}}_k = \arg\min_{\mathbf{V}_k} \sum_{i=1}^n \left\| x_i - \mathbf{V}_k \mathbf{V}_k^T x_i \right\|^2 + \lambda \sum_{j=1}^k \|\beta_j\|^2$$
subject to $\mathbf{V}_k^T \mathbf{V}_k = I_{k \times k}$

Sparse Principal Component Analysis

Tobias Bork

Introduction

PCA

Idea

Mathematical Formulations

Limits of Usabil

undamental

. .

Mathematical Formula

Theorems

Application

Fundamentals

oparse . e. .

eferences

Succes of PCA is due to the following two important optimal properties

- 1. Principal Components sequentially capture the maximum variability (among the columns of X, thus guaranteeing minimal information loss)
- Principal Components are uncorrelated, (so we can talk about one principal component without referring to others)

Sparse Principal Component Analysis

Tobias Bork

Limits of Usability

Linear Relationship between variables

Correlation of variables

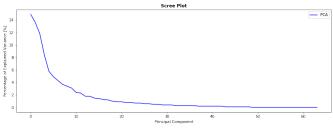
Completeness of data set

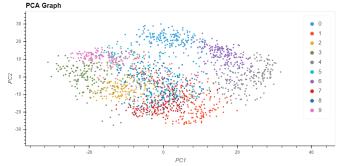
Outliers

► Number of variables p ; Number of Samples n (Inconsistency Theorem)

Interpreation of principal axis

Application to handwritten digits





Sparse Principal Component Analysis

Tobias Bork

ntroduction

PCA

Idos

Mathematical Formulation

Theorems

Limits of Usability

Application

Poforoncos



- 1: procedure SPCA(A, B)
- 2: $\mathbf{A} \leftarrow \mathbf{V}[1:k]$, the loadings of the first k ordinary principal components
- 3: while not converged do ▷ Definiere Abbruchkriterium
- 4: Given a fixed $\mathbf{A} = [\alpha_1, \dots, \alpha_k]$, solve the elastic net problem

$$\beta_{j} = \operatorname*{arg\,min}_{\beta} \left\| \mathbf{X} \alpha_{j} - \mathbf{X} \beta \right\|^{2} + \lambda \left\| \beta \right\|^{2} + \lambda_{1,j} \left\| \beta \right\|_{1}$$

5: For a fixed $\mathbf{B} = [\beta_1, \dots, \beta_k]$, compute the SVD of

$$\mathbf{X}^T \mathbf{X} \mathbf{B} = \mathbf{U} \mathbf{D} \mathbf{V}^T$$

- 6: $\mathbf{A} \leftarrow \mathbf{U}\mathbf{V}^T$
- 7: end while
- 8: $\hat{V}_j = \frac{\beta_j}{\|\beta_i\|}$ for $j = 1, \dots, k$
- 9: end procedure

Sparse Principal Component Analysis

Tobias Bork

troduction

PCA

Fundamental

parse PCA

Mathematical Formulat

Numerical Solution Adjusted Variances

p ¿¿ n case

- Sparse Principal Component Analysis
 - Tobias Bork

- References

- Beamer Paket http://latex-beamer.sourceforge.net/
- User's Guide to the Beamer
- DANTE e.V. http://www.dante.de