Linear Data Analysis
PCA And The Rayleigh Quotient

Cain Susko

Queen's University
School of Computing

March 16, 2022

a PCA, Scores, and the SVD

This lesson will look at the relations between the Rayleigh Quotient and PCA. The main concepts are we will view PCA as an optimization process. And the Rayleigh Quotient is just one way we can represent this optimization.

Recall given the orginal data matrix $A \in \mathbb{R}^{m \times n}$, we can create a zero mean data matrix $M \in \mathbb{R}^{m \times n}$. The scatter matrix for this system would be:

$$S = M^\top M$$

Loading vectors $\vec{v}_j$ are the eigenvectors of S, which are the columns of SVD factor $M = U\Sigma V^\top$. Thus, the first score vector is:

$$\vec{z}_1 = M\vec{v}_1$$

b PCA Maximizes a Linear Transformation

we will explore PCA as a process of maximizing a mapping. Given the zeor mean matrix $M \in \mathbb{R}^{m \times n}$, we can maximize its L_2 vector norm $||M\vec{u}||^2$. Note: we square the norm cause its better for calculations and will give us the same result in computations. Thus:

$$||M\vec{u}||^2 = [M\vec{u}]^\top [M\vec{u}] = \vec{u}^\top M^\top M\vec{u} = \vec{u}^\top S\vec{u}$$

$$||M\vec{u}||^2 = \vec{u}^\top S\vec{u}$$

This is what is known as the Max Quadratic Form.

Solving the Max Quadratic Form

Using Spectral Decomposition, we an solve a Max Quadratic Form equation $\vec{u}^\top S\vec{u}$ like so:

$$\vec{w} = \arg \max_{\vec{u} \in \mathbb{R}:||\vec{u}||=1} \vec{u}^\top S\vec{u}$$

$$\vec{w} = \vec{v}_{max} \in S$$

Note: $\vec{u}$ is constrained to a unit vector such that $||\vec{u}|| = 1$ Thus, one could see that the maximum value is achived when $\vec{u}$ is the largest loading vector in S.

Therefore, maximizing $M\vec{u}$ is the same as finding the largest eigenvalue and corresponding eigenvector within the scatter matrix S, where the found vector will be the principle loading vector for our data as well as the **Principal Axis** of the system.

c PCA and Rayleigh Quotient

this section will explore using PCA and the Rayleigh Quotient.
Consider the vector $\vec{w}$:

$$\vec{w} = \arg \max_{t \in \mathbb{R}^n : ||\vec{t}||=1} \vec{t}^\top S \vec{t}$$

$\vec{w}$ is only a maximization of a unit normal vector. to make this equaton more general, we can set $\vec{t}$ to be:

$$\vec{t} = \frac{\vec{u}}{||\vec{u}||}$$

such that $\vec{w}$ will now be:

$$\vec{w} = \arg \max_{\vec{u} \in \mathbb{R}^n} \frac{\vec{u}^\top S \vec{u}}{\vec{u}^\top \vec{u}}$$

The Rayleigh Quotient Thus, the rayliegh quotient, used to maximize a mapping, is:

$$R(S, \vec{u}) =^{def} \frac{\vec{u}^\top S \vec{u}}{\vec{u}^\top \vec{u}}$$

Where **if** $\vec{u} = \vec{0}$**, then** $R(S, \vec{0}) = 0$

Recall For the sake of notation, recall the following symbols:

- λ_{max} is the largest eigenvalue
- $\vec{v}_{max}$ is the eigenvector associated with the largest eigenvalue.

Fact Let us remeber that:

$$\lambda_{max} \in B = \max_{\vec{u} \in \mathbb{R}^n} R(B, \vec{u})$$

$$\vec{v}_{max} \in B = \arg \max_{\vec{u} \in \mathbb{R}^n} R(B, \vec{u})$$

We can now see that not only is the PCA related to the covariance matrix and eigenvalues of the scatter matrix, but also the Rayleigh Quotient. The Principal Axis of PCA is the result of finding the maximum input argument ($\vec{v}_{max}$) for the Rayleigh Quotient. Additionally, the largest eigenvalue λ_{max} is the latent variable for the PCA.

PCA Summary The principal loading vector in PCA:

- is the main way to describe covariance matrix
- provides scores of how data vary from the mean
- is the solution to an optimization problem
- described by the Rayleigh Quotient of the Scatter Matrix

Learning Summary

Students should now be able to:

- Relate PCA scores to singular vectors
- interperet PCA as a maximization process
- describe Principal PCA Axis as Rayleigh Quotient