**Exercise 5 Dimensionality Reduction Introduction**

Matthis Brocheton

## Probleme 1

The general form of data mapping from high to low-dimensionality space according to PCA:

**𝑿 ∈ ℝⁿˣᵈ** : the initial data matrix with **n** samples and **d** variables.

**𝝁 ∈ ℝᵈ** : the mean vector of the columns of **X**. This vector is subtracted to center the data.

**𝑿 − 𝝁** : the centered data, meaning the data without the bias introduced by the mean.

**𝑾 ∈ ℝᵈˣᵏ** : the projection matrix, composed of the **k** eigenvectors (called principal components) associated with the largest eigenvalues of the covariance matrix.

**𝒁 ∈ ℝⁿˣᵏ** : the new projected data, now in a reduced **k**-dimensional space (with **k < d**).

*Orthogonality*

Each vector points in a unique direction.

No principal component "absorbs" or overlaps the information of another.

Matrix:

*Mean data :*

*Center data:*

*Calculation of the covariance matrix:*

*Covariance matrix :*

≈

*Now I want to find Eigenvalues:*

Quadratic formula

*result for Eigenvalues λ*

*Now I want to find the first eigenvector:*

For

*We are looking for vectors:*

*Calculation*

*If then:*

*Normalization:*

*Result for the first eigenvector*

*Now I want to find the second eigenvector:*

For

*We are looking for vectors:*

*Calculation*

*If then:*

*Normalization:*

Result second *eigenvector*:

Result we can see that the are orthogonal:

## Probleme 2

**a)**

Between-class scatter

Within-class scatter

**b)**

**A black and white image of a mathematical equation

AI-generated content may be incorrect.**

maximized for:

**A black text with a white background

AI-generated content may be incorrect.**

*The goal:*

Maximize the distance between the projected averages

Minimize the total variance after projection

Formula projection:

Then:

The projected mean is

The projected variance is

So the Fisher criterion becomes:

*This is a Rayleigh quotient*

-> Between-class scatter

-> Within-class scatter

*Maximize this quotient*

We want:

We have:

So:

*Derive and cancel the gradient*

We are maximizing a quadratic ratio function → we use Lagrange method.

the maximum of:

Is:

Because the derivative of with respect to 𝜔 gives:

*Conclusion:*

We have shown that the optimal projection vector ω is:

It is the direction that maximizes class separation while minimizing projected total variance.