

Dimensionality Reduction with Principal Component Analysis



Name: Trinh Tran

Matriculation: 1105425

Instructor: Damir Dobric

Course: Software Engineering

Date: 2019 - 03 - 12

Layout

Introduction

Which reduce?
Why reduce?

Methods

How reduce ?
PCA

Implementation

How reduce ?
.NET Core

Test & Discussion

Unit Test

Integration Test

Operational Test



Introduction

Methods

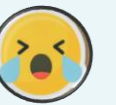
Implementation

Test & Discussion



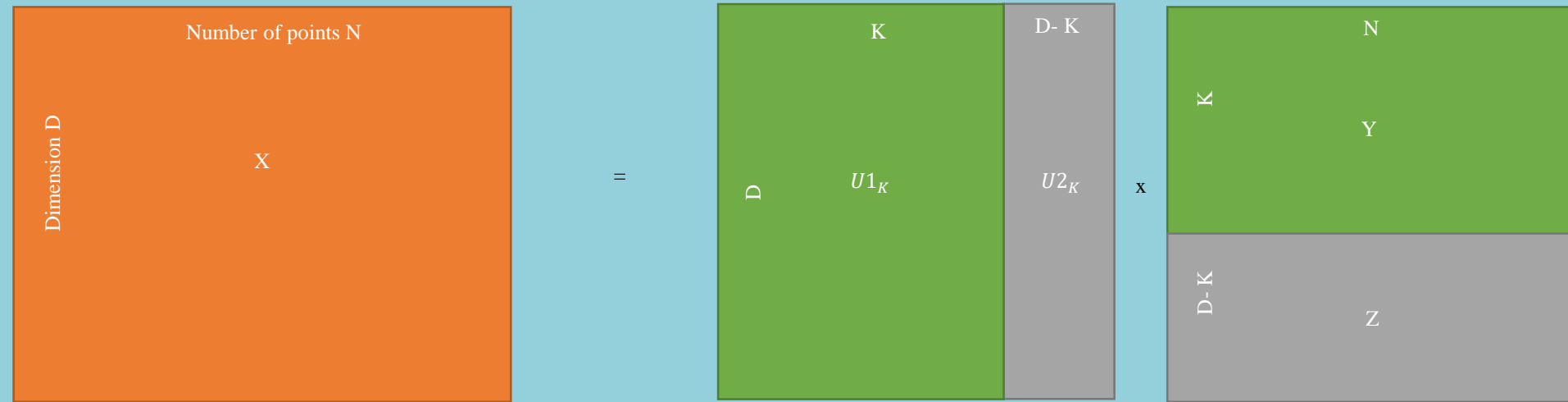
“An image is worth a thousand words”

➔ But it's too much!



Introduction

Methods



$$X = U_1 Y + U_2 Z$$

Implementation

Main idea of PCA:

- System U could be divided into two sub-spaces :

More important U_1 and less important U_2

➔ Keep U_1 and remove U_2

Test & Discussion

Introduction

Methods

Implementation

Test & Discussion

1. Find the mean vector of the input data:

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_N$$

2. Transform to the zero-corrected data:

$$\widehat{x}_N = x_N - \bar{x}$$

3. Find covariance matrix:

$$S = \frac{1}{N} \widehat{X} \widehat{X}^T$$

4. Compute eigenvalues and eigenvectors of S

5. Sort the eigenvalues descending and choose only K most important elements.

6. Project the input data to the new subspace:

$$Y = U_K^T \widehat{X}$$

Test Driven Development

Define components

1. Main component – PCA
2. Linear Algebra Operations (Abstract)
➔ PCA calls LAO

Define test cases

- Based on the LAO (abstract) – define test cases
➔ Test calls LAO

Implement functions

- Based on the LAO (abstract) – implement functions IF
➔ IF plug to LAO

Azure Pipelines

- Build + Run Test + Package

Introduction

Methods

Implementation

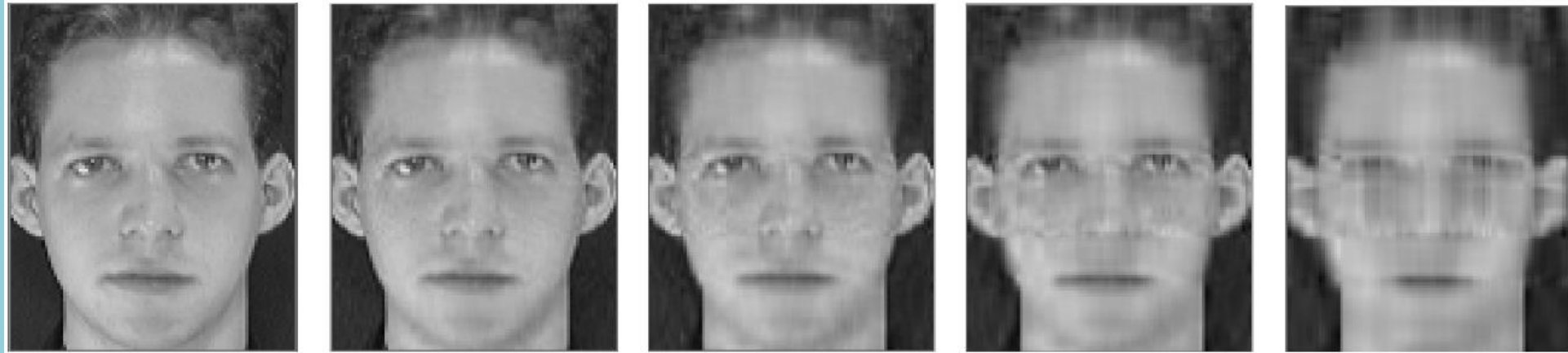
Test & Discussion

DimensionalityReduction.PCA (32 tests)	
Passed Tests (32)	11 sec
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.CalculateCovarianceMatrix_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.CalculateEigenVectors_RR	2 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.CalculateMeanVector_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.DotProductM2V_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.DotProductM2V_WD	3 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.DotProductM2V_WR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.MatrixAbs_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.MatrixAbs_WD	2 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.MatrixAbs_WR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.MatrixMatrixDotProduct_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.MatrixMatrixDotProduct_WD	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.MatrixMatrixDotProduct_WR	55 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SVD_RR	117 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.Subtract2Vectors_RR	3 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.Subtract2Vectors_WD	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.Subtract2Vectors_WR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractMatrixScala_RR	2 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractMatrixScala_WD	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractMatrixScala_WR	2 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractMatrixVector_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractMatrixVector_WD	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractMatrixVector_WR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractVectorScala_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.SubtractVectorScala_WR	53 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.Transpose_RR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.Transpose_WD	2 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.Transpose_WR	1 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.VectorMatrixDotProduct_RR	21 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.VectorMatrixDotProduct_WD	4 ms
DimensionalityReduction.PCA.Tests.LinearAlgebraUtilsTest.VectorMatrixDotProduct_WR	1 ms
DimensionalityReduction.PCA.Tests.PCAPipelineExtensionsTest.Test_PCAPipelineModuleIntergra...	369 ms
DimensionalityReduction.PCA.Tests.PCAPipelineModuleTest.Test_PCAPipelineWithCSVImage	10 sec

- Linear Functions such as covariance matrix computation, matrix vector dot product, etc.
- Each function should be tested once with a true expectation, a wrong expectation and a wrong dimension.
- In summary, there are 32 test cases developed and run successfully with XUnit.

Introduction

Methods



Implementation

Apply PCA - Reduce dimension to compress an image.

From left to right: input image, loss 1%, loss 3%, loss 5%, loss 10%.

More loss – Less components involved – Less quality output

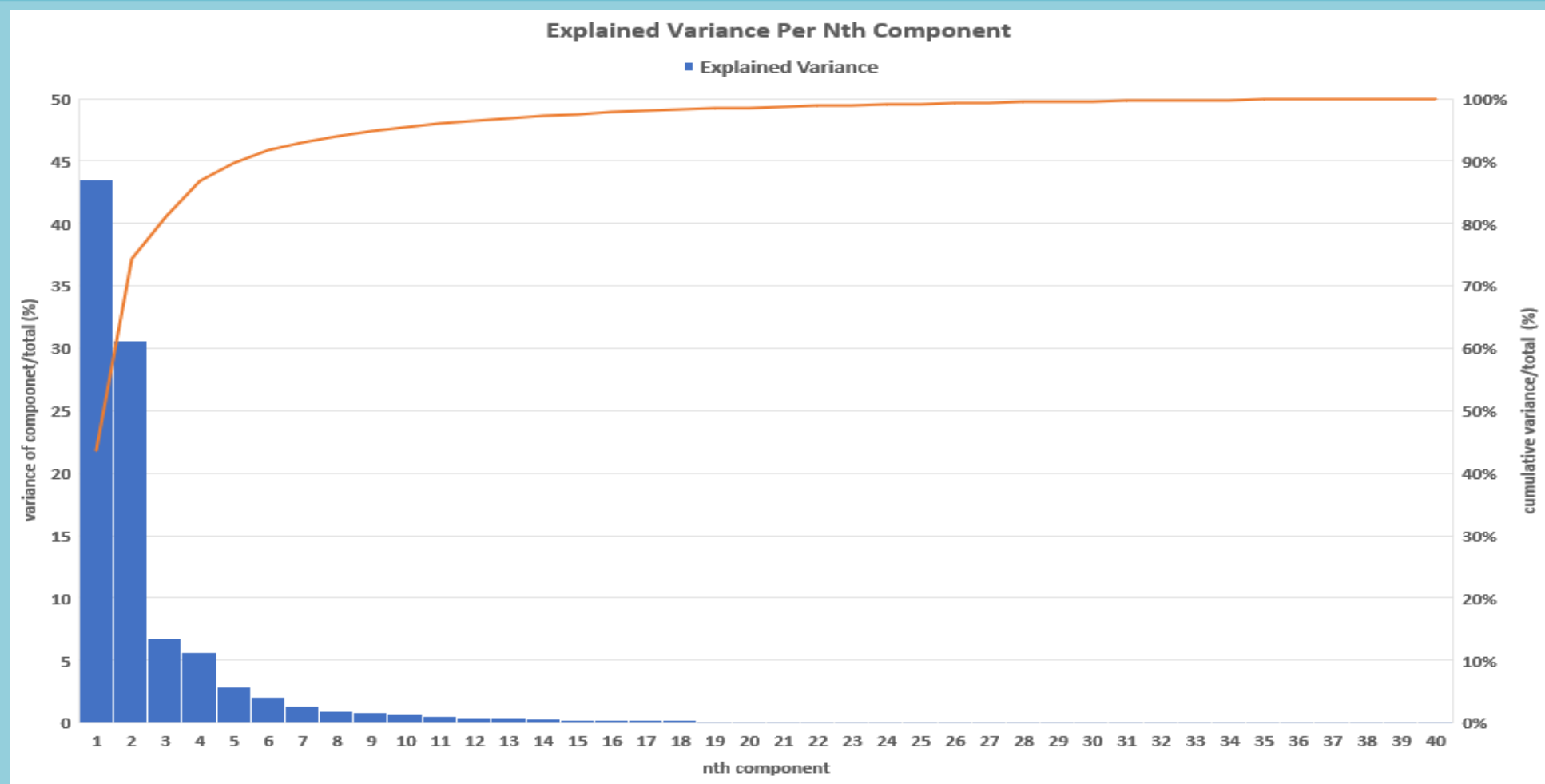
– Lower dimensional data set - More efficient in processing

Test & Discussion

Introduction

Methods

Implementation



- No. components ~ their importance.
- The blue column is the importance of nth component (descending).
- The red line is the level of output quality when cumulating from 1st component to nth component.

Test & Discussion



Thank you for your time!

Name: Trinh Tran

Email: phrong@stud.fra-uas.de

LinkedIn: <https://www.linkedin.com/in/trinh-nguyen-phuong-tran-2a0553142/>

Date: 2019 - 03 - 12