UTFPR - Especialização em Métodos Matemáticos Aplicados

Trabalho MATLAB

Cintia Izumi Shinoda

# 1 Introdução

Este trabalho tem como objetivo realizar uma Análise de Componentes Principais (PCA) utilizando MATLAB como ferramenta a fim de demonstrar o domínio desta.

O *dataset* analisado é referente ao consumo de 9 tipos de proteínas em 25 países europeus (WEBER, 1973).

Foram utilizados os conhecimentos adquiridos na disciplina Análise Multivariada da Especialização em Métodos Matemáticos Aplicados.

A análise PCA permitirá destacar os principais componentes (redução da complexidade dos dados) que explicam a maior parte da variação no consumo de proteínas entre os países, oferecendo *insights* sobre hábitos alimentares e agrupamentos de países com padrões de consumo de proteína semelhantes.

# 2 PCA – Análise de Componentes Principais

## 2.1 Importação dos dados

A table of food items

Description automatically generated

Figura 1: Tabela com os dados importados.

## 2.2 Matriz de Covariância

A matriz de covariância é utilizada para encontrar as direções de maior variância dos dados.

A number and numbers on a white background

Description automatically generated

Figura 2: Tabela com a matriz de covariâncias.

## 2.3 Matriz de Correlação

A matriz de correlação, é a matriz de covariância normalizada.

A group of numbers on a white background

Description automatically generated

Figura 3: Tabela com a matriz de correlações.

Figura 4: Gráfico de calor para a matriz de correlação.

## 2.4 PCA

## A number of numbers on a white background Description automatically generated

A black text with numbers

Description automatically generated

Figura 5: Porcentagem de variância explicada por cada componente.

A number of numbers on a white background

Description automatically generated

Figura 6: Variância explicada por cada componente.

# 3 Referências

SCOTT, J. G. Protein data set. Disponível em: <https://raw.githubusercontent.com/jgscott/STA380/master/data/protein.csv>. Acesso em: 30 nov. 2024.