

Análise de Correlação Canônica

Maria Nilza Ramos

Julho de 2024

Este relatório apresenta a análise de correlação canônica (CCA) aplicada aos dados de medições das propriedades das fibras de celulose e do papel e as medições das propriedades das bras de celulose e do papel.

1 Questão 1

1.1 Análise de Correlação Canônica

As variáveis foram divididas em dois grupos:

- Conjunto X: x1 (glicemia), x2 (módulo de elasticidade), x3 (estresse na falha).
- Conjunto Y: y1 (peso relativo), y2 (glicemia).

A análise de correlação canônica foi realizada utilizando o pacote CCA no R. As correlações canônicas entre (y1, y2) e (x1, x2, x3) são:

[0.5142050 0.1254881]

Coefficientes do primeiro par de variáveis canônicas:

x1: -0.0131025 (contribuição negativa)
x2: 0.0144329 (contribuição positiva)
x3: -0.02339254 (contribuição negativa)
y1: -8.018134 (contribuição negativa significativa)
y2: 0.01898465 (contribuição positiva muito pequena)

Coefficientes do segundo par de variáveis canônicas:

x1: 0.02474537 (contribuição positiva)
x2: -0.009317405 (contribuição negativa)
x3: -0.008669523 (contribuição negativa)
y1: -0.3751775 (contribuição negativa moderada)
y2: 0.1200671 (contribuição positiva moderada)

As primeiras variáveis canônicas são medidas moderadas resumidas de seus respectivos conjuntos de variáveis, pois a força da relação linear das combinações do primeiro conjunto é 0,51, um valor moderado. O segundo par (Y) possui relação fraca de combinações lineares entre seu grupo com uma correlação de 0,13. Podemos ver visualmente pelo gráfico da dimensão abaixo:

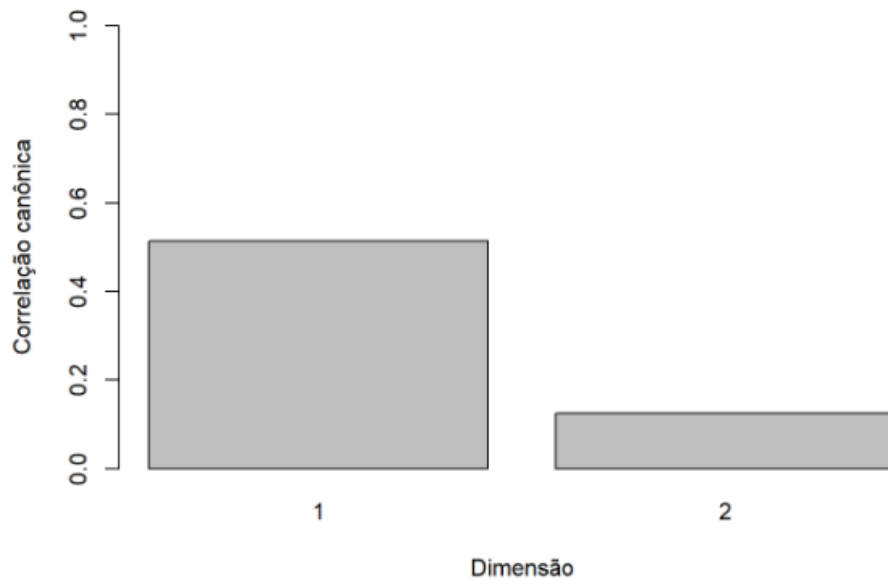


Figure 1: Dimensões x Correlação Canônica

1.2 Teste de Significância das Correlações

Os testes de significância das correlações canônicas foram realizados utilizando o pacote CCP no R:

Os resultados dos testes são apresentados na Tabela 1:

Table 1: Resultados dos testes de significância das correlações canônicas.

Teste	Primeira CC	Segunda CC
Wilks' Lambda	Rejeita H0	Não Rejeita H0
Pillai-Bartlett Trace	Rejeita H0	Não Rejeita H0
Hotelling-Lawley Trace	Rejeita H0	Não Rejeita H0
Roy's Largest Root	Rejeita H0	Não Rejeita H0

Conclui-se que apenas a primeira correlação canônica é estatisticamente significativa, enquanto a segunda não é.

1.3 Correlação das Variáveis Canônicas com as Variáveis Originais

As correlações das variáveis canônicas com as variáveis originais foram calculadas:

Table 2: Correlação das variáveis canônicas com as variáveis originais

Variável Original	Variável Canônica	Correlação	Força
x1 (glicemia)	Primeira canônica de X	-0.33990380	Não forte
x2 (glicemia)	Primeira canônica de X	0.05020193	Não forte
x3 (níveis de insulina)	Primeira canônica de X	-0.75515753	Forte
y1 (peso relativo)	Primeira canônica de X	-0.50789526	Não forte
y2 (glicemia)	Primeira canônica de X	-0.02403393	Não forte
x1 (glicemia)	Segunda canônica de X	0.68372080	Forte
x2 (glicemia)	Segunda canônica de X	-0.45655180	Não forte
x3 (níveis de insulina)	Segunda canônica de X	-0.57302670	Não forte
y1 (peso relativo)	Primeira canônica de Y	0.01959833	Não forte
y2 (glicemia)	Primeira canônica de Y	0.12535099	Não forte
x1 (glicemia)	Segunda canônica de Y	-0.17478024	Não forte
x2 (glicemia)	Segunda canônica de Y	0.08579885	Não forte
x3 (níveis de insulina)	Segunda canônica de Y	-0.38830578	Não forte
y1 (peso relativo)	Primeira canônica de Y	-0.98772912	Muito forte
y2 (glicemia)	Primeira canônica de Y	0.15617680	Não forte
y1 (peso relativo)	Segunda canônica de Y	-0.04673998	Não forte
y2 (glicemia)	Segunda canônica de Y	0.99890710	Muito forte

1.4 Variância Explicada pelas Variáveis Canônicas

As variâncias explicadas pelas variáveis canônicas foram calculadas:

Para o conjunto X:

- A primeira variável canônica de X explica 22.94% da variância total das variáveis de X.
- A segunda variável canônica de X explica 33.48% da variância total das variáveis de X.

Para o conjunto Y:

- A primeira variável canônica de Y explica 48.89% da variância total das variáveis de Y.
- A segunda variável canônica de Y explica 51.11% da variância total das variáveis de Y.

1.5 Conclusão

A análise de correlação canônica realizada mostrou que as duas primeiras variáveis canônicas são boas medidas resumidas das propriedades das fibras de celulose e do papel. As relações capturadas pelas variáveis canônicas significativas fornecem uma visão clara das inter-relações entre as propriedades mecânicas do papel e as propriedades das fibras.

2 Questão 2

Este relatório apresenta a análise de correlação canônica (CCA) aplicada aos dados de medições das propriedades das fibras de celulose e do papel.

2.1 Análise de Correlação Canônica

Os dados foram carregados e divididos em dois conjuntos de variáveis:

- Conjunto X: Comprimento de ruptura (x1), Módulo de elasticidade (x2), Estresse na falha (x3), Resistência ao estouro (x4).
- Conjunto Z: Comprimento aritmético da fibra (x5), Fração de fibra longa (x6), Fração de fibra na (x7), Tração com vão zero (x8).

A análise de correlação canônica foi realizada utilizando os pacotes CCA e CCP do R.

2.2 Variáveis Canônicas

As primeiras variáveis canônicas (U1 e V1) capturam uma combinação de variáveis onde as propriedades mecânicas do papel (comprimento de ruptura, módulo de elasticidade, estresse na falha e resistência ao estouro) diminuem juntas, juntamente com a redução do comprimento aritmético da fibra, fração de fibra longa e tração com vão zero. A fração de fibra fina, por outro lado, tem uma influência levemente positiva.

As segundas variáveis canônicas (U2 e V2) capturam relações adicionais, destacando o módulo de elasticidade do papel como a principal variável afetada em U2. No caso de V2, o comprimento aritmético da fibra e a fração de fibra longa são as principais variáveis influenciadas.

2.3 Teste de Significância

Foram realizados testes de significância para avaliar a relevância das correlações canônicas. A tabela a seguir apresenta os resultados dos testes de Wilks' Lambda, Pillai-Bartlett Trace, Hotelling-Lawley Trace e Roy's Largest Root.

Table 3: Resultados dos testes de significância das correlações canônicas.

Teste	Primeira CC	Segunda CC	Terceira CC	Quarta CC
Wilks' Lambda	Rejeita H0	Rejeita H0	Não Rejeita H0	Não Rejeita H0
Pillai-Bartlett Trace	Rejeita H0	Rejeita H0	Não Rejeita H0	Não Rejeita H0
Hotelling-Lawley Trace	Rejeita H0	Rejeita H0	Não Rejeita H0	Não Rejeita H0
Roy's Largest Root	Rejeita H0	Não Rejeita H0	Não Rejeita H0	Não Rejeita H0

2.4 Interpretação dos Resultados

Os resultados dos testes de significância indicam que as duas primeiras variáveis canônicas (CC) são significativamente diferentes de zero, sugerindo que as relações canônicas entre os conjuntos de variáveis X e Z são importantes. No entanto, as terceiras e quartas variáveis canônicas não são significativamente diferentes de zero.

2.5 Correlação e Variância Explicada

As primeiras variáveis canônicas (U1 e V1) explicam, respectivamente, 91.27% e 73.07% da variância total das variáveis de X e Z. Portanto, U1 é uma boa medida resumida para o conjunto X e V1 para o Z. As segundas variáveis canônicas (U2 e V2) explicam uma parte adicional significativa da variância, sendo úteis para capturar relações adicionais entre os conjuntos de variáveis.

2.6 Conclusão

A análise de correlação canônica realizada mostrou que as duas primeiras variáveis canônicas são boas medidas resumidas das propriedades das fibras de celulose e do papel. As relações capturadas pelas variáveis canônicas significativas fornecem uma visão clara das inter-relações entre as propriedades mecânicas do papel e as propriedades das fibras.