

MAE5776 – Análise Multivariada LISTA 3 1º Semestre/2022

Alunos:

Fernando F. Paulos Vieira - n° USP: 13492870

Leandro Alves da Silva - n° USP: 11868023

Thiago Ferreira Miranda - n° USP: 11925711

1 - A partir de uma matriz de dados normalizados $Y_{n \times p}^*$, considere a matriz de covariâncias $nS_{p \times p} = Y^{*'}Y^* = V\Lambda V'$, tal que $V_{p \times p} = (V_1, \dots, V_p)$ e $\Lambda = \text{diag}(\lambda_j)$ são matrizes de autovetores (das colunas de $Y_{n \times p}^*$) e autovalores, respectivamente, e a matriz de distâncias $D_{n \times n}$, tal que seus elementos são função dos elementos de $B_{n \times n} = Y^*Y^{*'} = U\Lambda U'$, com $U_{n \times n} = (U_1, \dots, U_n)$ matriz de autovetores (das linhas de $Y_{n \times p}^*$). Três pesquisadores realizaram análises estatísticas e chegaram à seguinte redução de dimensionalidade de Y^* .

Pesquisador 1: $Y_{n \times p}^* \rightarrow \tilde{Y}_{n \times 2} = Y^*(V_1 \ V_2)$

Pesquisador 2: $Y_{n \times p}^* \rightarrow \tilde{Y}_{n \times 2} = Y^*\left(\frac{V_1}{\sqrt{\lambda_1}} \ \frac{V_2}{\sqrt{\lambda_2}}\right)$

Pesquisador 3: $Y_{n \times p}^* \rightarrow \tilde{Y}_{n \times 2} = Y^*(V_1\sqrt{\lambda_1} \ V_2\sqrt{\lambda_2})$

- Qual análise estatística cada pesquisador realizou? Que propriedades dos dados estão preservadas em cada caso? Eles partiram do mesmo objetivo? Faça suposições necessárias.
- Simule dados e realize as análises dos três pesquisadores. Interprete os resultados.
- Para os dados simulados, obtenha uma representação Biplot. Como esse gráfico é construído?

2 - Considere os dados “bodyfat” disponíveis na biblioteca TH.data do R. Neste caso, a matriz de trabalho contém 71 observações avaliadas em 10 variáveis. Gere 5 observações (para tanto, adote um critério) e considere seu novo conjunto de dados “bodyfat_new”. Com base na matriz de trabalho resultante realize as seguintes análises:

2.1 - Componentes Principais.

2.2 - Escalonamento Multidimensional (ou Coordenadas Principais) – Compare as soluções métricas e não-métricas.

2.3 - Análise Fatorial Exploratória – Solução de MVS (rotacionar, se for de interesse). Em cada caso, que proporção da variância total dos dados pode ser explicada por 2 componentes? Quais variáveis mais influenciaram na redução de dimensionalidade? Represente os dados em eixos bidimensionais, identifique as observações e compare os resultados das três análises.

2.4 - Escolha uma das variáveis do banco de dados e obtenha uma tabela de contingência categorizando esta variável de acordo com faixas etárias das observações. Realize uma Análise de Correspondência e comente sobre o padrão de associação presente nesses dados.