



RAD5053 – Análise de Dados Multivariados I

Prof. Dr. Evandro Marcos Saidel Ribeiro

Lista 02 – Componentes Principais

1 Produto Estadual Bruto, nos estados americanos. (Gross State Product – GSP).

Dados sobre produto estadual bruto estão disponíveis em dois arquivos: **gsp_raw** (expresso em milhões de dólares) e **gsp_share** (expresso como uma proporção do total do GSP para cada estado). Cada arquivo consiste em 50 linhas (uma para cada estado) e 14 colunas (a primeira é a identidade do estado e as 13 restantes correspondendo às diferentes áreas de atividade econômica).

Utilize estes arquivos para estudar a técnica de Análise de Componentes Principais.

1.1 Obtenha matrizes de correlações e de covariâncias.

1.1.1 Descreva a variância total dos dados.

1.2 Faça uma análise dos autovalores:

1.2.1 Gráfico screeplot.

1.2.2 Percentual de variância explicada.

1.3 Considere apenas os dois primeiros autovalores,

1.3.1 Informe o percentual de variação explicada

1.3.2 Apresente gráficos com os valores obtidos para as primeiras duas componentes. (no R os valores são obtidos pelo comando “predict”)

1.4 Quais as diferenças em utilizar uma base de dados ou a outra?

1.5 O que acontece quando se remove as observações correspondentes ao Alasca e Wyoming?