

Tópicos Especiais em Estatística Aplicada – 2020-2

EPC11

Sejam as 1000 amostras de 7 variáveis fornecidas no arquivo `epc11dat.txt` na forma $Y=[y_1 \dots y_7]$.

1. Use as primeiras 500 amostras de Y para fazer a análise por componentes principais e plote um gráfico mostrando a variância (0 a 100% da variância) retida nos dados ao se adicionar gradativamente 1 a 7 componentes principais.
2. Use apenas 2 componentes principais para estimar as 7 variáveis, plotando y_1 a y_7 e os valores obtidos com suas estimativas usando apenas as 2 componentes. Também neste caso, use apenas as primeiras 500 amostras.
3. Usando apenas as primeiras 500 amostras, calcule a estatística T^2 usando PCA. Aplique então a estatística às 1000 amostras e plote seu gráfico junto com o limiar de controle. Verifique se esse monitoramento multivariado poderia ser feito sem o uso de PCA, diretamente com a matriz de covariância.