

NÚMERO ÍNDICE DE INDICADORES DE DESEMPENHO EMPRESARIAL

NICACIO, Joaquim Eduardo de Moura¹

RESUMO: Nessa pesquisa objetivou-se determinar um número índice a partir de indicadores de desempenho empresarial publicados no *Anuário do Agronegócio* de 2012. Foram utilizados dois métodos da estatística multivariada: a análise de agrupamento pelo critério de Ward com a distância de Mahalanobis para ordenar um conjunto de n objetos segundo critério definido por um conjunto de m variáveis correlacionadas entre si e a transformada de Hotelling para fornecer as ponderações ou pesos das componentes principais. Foi considerada apenas uma componente representando 84% da variância total. A metodologia é aplicável em empresas que geram informações mensais ou em diversos períodos de um ano com forte benefício para a análise de indicadores.

PALAVRAS CHAVE: Número índice, análise de agrupamento, transformada de Hotelling, indicadores de desempenho empresarial.

ABSTRACT: This research aimed to determine an index number from business performance indicators published in *Agribusiness Yearbook* 2012. We used two statistical methods of multivariate cluster analysis by Ward criterion with Mahalanobis distance to sort a set of n objects with criteria defined by a set of m variables correlated and Hotelling transform to provide the weights or weights principal components. Was considered only a component representing 84% of the total variance. The methodology is applicable to companies that generate a monthly or several times a year with strong benefit for the analysis of indicators.

KEYWORDS: Number index, cluster analysis, Hotelling transform, business performance indicators.

1. INTRODUÇÃO

O protocolo de Kyoto, em 1997, alertou o mundo para os problemas climáticos do planeta e, com eles, estava o problema da sustentabilidade ambiental. Em se tratando de clima, o impacto imediato dele afeta de forma contundente o problema agrícola. Em sua esteira vem o problema da fome e, com ela, o caos social.

Observando a aptidão agrícola do Brasil e analisando o crescimento populacional do planeta, torna-se claro a grande demanda por alimentos de origem agropecuária e também pelos produtos florestais. Com isto observa-se neste começo de século a expansão das fronteiras agrícolas, principalmente no Centro Oeste onde se encontra parte do bioma cerrado no planalto central.

Para suprir essa demanda, empresas rurais e agroindústrias se desenvolveram para dar suporte à manutenção das empresas rurais e cadeias de suprimentos de seus produtos. Surgem, então, as indústrias de carne, de rações, de óleo comestível, fertilizantes, laticínios, enfim, uma gama imensa de empresas vinculadas à produção rural.

Com a globalização da economia, com a escassez de terras agricultáveis na Europa, com o surgimento de nações em desenvolvimento como China, Índia e Rússia, o Brasil torna-se um polo de investimento estrangeiro além de ótimo para os

¹ Professor Adjunto do departamento Acadêmico de Contabilidade na Universidade Federal de Mato Grosso

naturais. Com isso é necessário conhecer um pouco mais sobre a situação econômica financeira das empresas, principalmente às ligadas as atividades rurais. Para isso é preciso ir além dos indicadores econômicos financeiros fundamentados pelas demonstrações contábeis de forma individualizada.

Sob essa ótica é possível efetuar uma ordenação de empresas do setor agropecuário, notadamente algumas do setor de *Raçoes*, sobre as quais são observadas variáveis relacionadas ao seu desempenho para determinar um número índice que meça o seu *grau de desempenho conjunto*? A hipótese é que mediante o emprego das estatísticas multivariadas da análise de agrupamento e das componentes principais pode-se obter tal índice.

Diante dessa necessidade conjuntural, esse artigo tem o objetivo de encontrar um índice que permita encontrar o grau de desempenho conjunto como uma combinação linear tal que: $y = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n$ das variáveis sendo estas os indicadores econômicos financeiros previamente determinados.

A importância dessa pesquisa é observada pela aplicação da estatística multivariada na análise simultânea de indicadores econômicos financeiros de um conjunto de empresas atuantes no agronegócio apresentando uma metodologia robusta de análise contábil.

2. REFERENCIAL TEÓRICO

A exigência legal das publicações das demonstrações financeiras de acordo com suas estruturas padronizadas facilita sua análise pelos usuários que dela necessitam. Delas surgem os *indicadores* que vão indicar diversas relações de desempenho e entre elas as econômicas financeiras principalmente.

É basicamente um processo de análise temporal, desenvolvido por meio de números índices, sendo seus cálculos processados de acordo com a seguinte expressão: $Ni = \frac{V_d}{V_b} \times 100$. Pela identidade, revela-se que o número índice é a relação existente entre o valor de uma conta contábil (ou grupo de contas) em determinada data (V_d) e seu valor obtido na data base (V_b), (NETO, 2002).

Um dos trabalhos pioneiros na construção de índices foi o *Best Linear Index Numbers of Prices and Quantities* de (THEIL, 1960, p. 1) que utilizou a análise das componentes principais de forma diferenciada, em problemas de preços e quantidades de commodities agrícolas.

O modelo atual de gerenciamento de fazendas baseado em aplicações uniformes de fertilizantes e defensivos agrícolas produz uma elevação muito grande dos custos de produção apresenta baixa automação e pouca utilização de tecnologia voltada para a produção agrícola.

No entanto, com a disponibilidade de tecnologia e o uso de sistemas de posicionamento global (GPS), o avanço dos sistemas de informações geográficas

(GIS) e o sensoriamento remoto (SR) fornecem um sistema de gerência de ótimo desempenho que leva o produtor ter um aumento considerável de produtividade, redução de custos de produção e menor impacto ambiental.

Com a evolução constante dos sistemas de informações econômicas rurais e tendo o Brasil grandes desenvolvimentos na última década do agronegócio surgem os *Anuários* que são publicações técnicas com respaldo governamental na sua fonte de informações básicas. Nesse setor se destaca o *Anuário do Agronegócio* elaborado pelo grupo da revista Globo Rural e pela empresa Serasa Experian.

Nesse *Anuário* são apresentados alguns *indicadores* cuja definição é dada pela empresa privada que faz análises e pesquisas de informações econômicas financeiras para apoiar decisões de crédito, denominada Serasa Experian. Foi criada pelos bancos com o objetivo de centralizar informações e reduzir custos. Também a revista Globo Rural adota a mesma definição.

Os indicadores definidos pelo Anuário do Agronegócio que serão analisados são:

- Rentabilidade do Patrimônio Líquido (RPL) = Lucro Líquido / Patrimônio Líquido (%)
- Giro do Ativo (GA) = Receitas Líquidas / Ativo Total (Pontos)
- Margem Líquida (ML) = Lucro Líquido / Receita Líquida (%)
- Margem da Atividade (MA) = Lucro da Atividade / Receita Líquida (%)
- Receita Líquida (RL) = Vendas Líquidas anuais (R\$ milhões)
- Evolução da Receita Líquida (ERL) = Receita Líquida atual / Receita Líquida anterior (%)

Dentre os indicadores apresentados no setor de rações, esses estão correlacionados, pois quaisquer alterações no valor da Receita Líquida todos refletirão o efeito dessa alteração.

2.1 Definição do Índice

Dado um conjunto de objetos (empresas) O_1, \dots, O_n que se deseja ordenar, segundo características associadas a um conjunto de variáveis X_1, \dots, X_p . Assim a cada objeto O_j associamos um valor

$$I_j = \sum a_i x_{ij} \quad (1)$$

onde x_{ij} é o valor da i -ésima variável observada para o j -ésimo objeto e a_i é o peso da i -ésima variável que reflete a importância da variável na construção do índice I .

Nessa construção é necessário primeiramente selecionar as variáveis e depois ponderá-las. (Kubrusly, 2001, p. 107). Esses dois aspectos do problema serão

abordados aqui, com a utilização de duas técnicas da estatística multivariada que são nessa ordem: a análise de agrupamento e a transformada de Hotelling.

2.2 A análise de agrupamento ou análise de clusters

É um conjunto de procedimentos estatísticos que tem por finalidade classificar *objetos* segundo sua medida de *semelhança* (distância, associação) ou *dissemelhança*. Essa medida está afeta as variáveis que cada *objeto* contém.

Na análise de clusters é fundamental ter cuidado particular na seleção das variáveis de partida que vão caracterizar cada indivíduo ou caso, e determinar, em última instância, qual o grupo em que deve ser inscrito. Nesta análise não existe qualquer tipo de interdependência entre as variáveis, isto é, os grupos confirmam-se por si mesmo sem necessidade de ser definida uma relação causal entre as variáveis utilizadas. (REIS, 1991).

Uma das medidas de distância muito utilizada em análise de cluster é a *distância euclidiana* que é o comprimento de reta que une dois pontos A (x_0, y_0) e B(x_1, y_1). É a raiz quadrada do somatório dos quadrados das diferenças entre os pontos A e B. Generalizando temos para $p = 2$ pontos:

$$d(p_1, p_2) = \sqrt{|x_{p1} - x_{p2}|^2 + |y_{p1} - y_{p2}|^2} \quad (2)$$

Quando a desigualdade métrica $d_{ij} + d_{jk} \geq d_{ik}$ se constata, torna-se trivial a construção do triângulo euclidiano (GOWER *et al.*, 1986, p.6). Essa distância possui as seguintes propriedades:

- a) $d(AB) > 0$ exceto se $A = B = 0$. Isso significa que a distância entre dois pontos distintos é sempre positiva;
- b) $d(AB) = d(BA)$. Isso significa que a distância do ponto A ao ponto B é a mesma que a distância de B a A;
- c) $d(AB) + d(BC) \geq d(AC)$. Desigualdade triangular, isto é, a soma dos dois lados de um triângulo é maior ou igual do que o outro lado.

Uma quarta propriedade é: $D(AB) = 0 \Leftrightarrow A = B$ (WEINBERG *et al.*, 2009, p. 209).

Uma das medidas de associação utilizadas na análise de cluster é o coeficiente de correlação linear de Karl Pearson. É a medida de associação entre duas variáveis quantitativas. É representada pela letra r e é calculada segundo a fórmula (TRIOLA, 1999, p. 236):

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \cdot \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (3)$$

O pressuposto da correlação linear de Karl Pearson consiste em: "O teste de significância de r não pode ser aplicado se as distribuições das variáveis não forem normais, isto é: seguem a lei de Gauss (distribuição simétrica em relação à média)".

Para validar o cálculo do coeficiente de correlação, aplica-se o teste de hipóteses pela distribuição t como segue:

a) Elaboração das hipóteses

$H_0: \rho = 0$. O coeficiente de correlação é zero;

$H_a: \rho \neq 0$. O coeficiente de correlação é diferente de zero.

b) Escolha do nível de significância

$\alpha = 0,05$.

c) Determinação do valor crítico do teste

$t_{\alpha, gl} = t_{0,05;6} = 2,447$ (gl = n - 2, onde n é o número de pares de valores x,y)

d) Determinação do valor calculado de t

$$t_{calc}^2 = \frac{r^2}{\frac{1-r^2}{n-2}}$$

e) Se $t_{calc} > t_{tab}$ rejeita-se H_0 .

Um teste apresentado por (KACHIGAN, 1991, p. 296) tem as seguintes etapas:

a) Elaboração das hipóteses

$H_0: \rho = 0$ (Não há correlação linear significativa)

$H_a: \rho \neq 0$. (Há correlação linear significativa).

b) Escolha do nível de significância

$\alpha = 0,05$.

c) Determinação do valor crítico do teste

$r_{calc} > r_{tab}$ Rejeita-se a hipótese de nulidade.

d) Determinação do valor calculado de r

$$r = \frac{t_{calc}}{\sqrt{t_{calc}^2 + n - 2}} \quad (4)$$

Uma das medidas estatísticas de grande importância é a variância. Um índice que é *combinação linear de variáveis* deve conter o máximo de informação contida nelas. Isso equivale a dizer que esse índice tem a propriedade de ter a *máxima variância* que é obtida pelo emprego da análise de componentes principais.

Um método de aglomeração que leva em conta a variância é o método de Ward ou de Variância Mínima. É o método de agregação baseada na *variância* que se aplica no caso de dados tabelados em uma tabela do tipo T(n, p) de variáveis quantitativas. Busca minimizar a soma dos quadrados de quaisquer dois (hipotéticos) clusters que podem ser formados em cada passo.

coordenada, elas terão o mesmo peso ou a mesma importância para o cálculo da distância.

Observando uma amostra de n indivíduos, vemos que essas variáveis são independentes, pois, não se pode prever uma dependendo do conhecimento da medida da outra. Além disso, quando colocadas em um sistema de eixos ortogonais, facilmente se verifica que a variável comprimento possui maior variabilidade do que a variável largura para um peixe comum no pantanal mato grossense que é a piraputanga, *Brycon orbygnianus*. A figura 1 apresenta um diagrama de dispersão que simula essa condição:

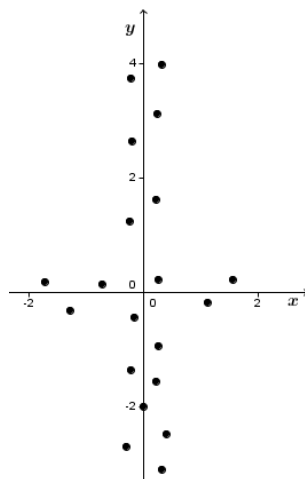


Figura 1 Diagrama de dispersão mostrando a maior variabilidade na direção de y do que na direção de x .
Fonte: construção do autor.

Uma forma de compensar essa diferença de variabilidade quando as distâncias da origem de coordenadas forem calculadas, é dividir cada *coordenada* pelo desvio padrão *amostral*. Esse procedimento é uma *transformação homotética*.

Para exemplificar, as coordenadas da figura 1 são:

$x = [0.32; -0.22; -0.72; -1.28; 0.24; -0.2; 0.22; -0.24; 0.26; 1.12; 1.56; -1.72; -0.16; 0.26; -0.22; 0.22; 0; 0.4; -0.3; 0.32]$ são os valores das abscissas.

$y = [3.98; 3.74; 0.14; -0.32; 3.12; 2.64; 1.62; 1.24; 0.22; -0.18; 0.22; 0.18; -0.44; -0.94; -1.36; -1.56; -2; -2.48; -2.7; -3.1]$ são os valores das ordenadas.

As variâncias são: $S_{XX} = 0.5140$ e $S_{YY} = 4.3781$. Dessa forma são eliminadas as diferenças de variabilidade. Agora a distância da origem $O(0,0)$ a um ponto genérico $P(x,y)$ hipotético pela distância euclidiana padrão, com distância quadrada conhecida torna-se:

$$d(O, P) = \sqrt{x^2 + y^2} = \sqrt{\frac{x^2}{0.5140} + \frac{y^2}{4.3781}} = 0.23 \quad (7)$$

A figura 2 mostra a elipse de centro O (0,0) com o eixo maior coincidente com o eixo das abscissas e o eixo menor coincidente com o eixo das ordenadas. A metade do maior semieixo é $\sqrt{0.23} \cdot \sqrt{0.5140} = 0.34$ e a metade do menor semieixo é $\sqrt{0.23} \cdot \sqrt{4.3781} = 1.00$.

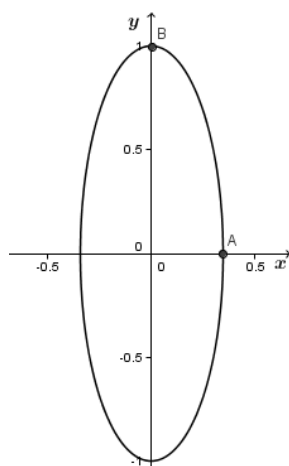


Figura 2 – Elipse de uma distância estatística quadrática

Fonte: construção do autor.

Outra distância utilizada é a distância ao quadrado de Mahalanobis definida por (MACHLAN, 1999, p. 21) como:

$$\Delta^2 = (\mu_1 - \mu_2)^T \Sigma^{-1} (\mu_1 - \mu_2) \quad (8)$$

Essa distância também é denominada *distância generalizada*. Suas variáveis são correlacionadas (MAHALANOBIS, 1936, p. 50). Colocada na forma de fração

$$\Delta^2 = \frac{(\mu_1 - \mu_2)^T (\mu_1 - \mu_2)}{\Sigma} \quad (9)$$

Para testar a significância desta distância, é aplicado o teste F, cuja fórmula (VALENTIN, 2000, p. 101) é a seguinte:

$$F = \frac{D^2 \cdot n_1 \cdot n_2 \cdot (n_1 + n_2 - 3)}{2(n_1 + n_2)(n_1 + n_2 - 2)} \quad (10)$$

e os graus de liberdade do numerador é $\nu_1 = 2$ e do denominador $\nu_2 = n_1 + n_2 - 3$, sendo n_1 e n_2 os números de observações em cada grupo.

Observamos que Σ , que é a matriz de covariância, pondera os parâmetros de acordo com a sua importância. No caso de $\Sigma = I$, distância Mahalanobis é o mesmo que a distância Euclidiana:

$$d_{Mahalanobis}(x, y) = \sqrt{(x - y)^T I^{-1} (x - y)} = \sqrt{(x - y)^2} \\ = d_{Euclidiana}(x, y) \quad (11)$$

A distância de Mahalanobis é uma distância em sentido geométrico, porque a matriz de covariância e sua inversa são matrizes definidas positivas, isto é, tem autovalores, que são as raízes, positivos. Essas matrizes possuem as seguintes propriedades (LEON, 1999, p. 265).

- É matriz simétrica
- É matriz invertível
- Seu determinante é maior do que zero
- Suas submatrizes principais são todas positivas definidas.

Os autovalores possuem as seguintes propriedades:

1. $\det(C) = \lambda_1 \times \lambda_2 \times \dots \times \lambda_n$. (produto dos autovalores)
2. $\text{tr}(C) = \lambda_1 + \lambda_2 + \dots + \lambda_n$. (soma dos autovalores)
3. C é matriz não singular se, e somente se, todos os autovalores $\neq 0$.

Os autovalores de C e de C^T são os mesmos.

2.4 A Transformada de Hotelling

A transformada de Hotelling também conhecida como componentes principais uma vez que foi HOTELLING (1933, p. 1) que utilizou essa expressão em seu artigo "*Analysis of a complex of statistical variable into principal components*" que derivou a partir de um problema de autovalores e autovetores.

A estatística multivariada das componentes principais é aplicada aos dados coletados, pois, permite transformar um conjunto de variáveis iniciais *correlacionadas entre si*, em outro conjunto de variáveis *não correlacionadas e ortogonais*, as chamadas componentes principais, que resultam de *combinações lineares* do conjunto inicial.

Sobre o cálculo das componentes principais (ANDERSON, 1974, p. 272) assim explica: "*São calculadas por ordem decrescente de importância, isto é: a primeira explica o máximo possível da variância dos dados originais, a segunda o máximo possível da variância ainda não explicada, e assim por diante*".

O principal efeito da Transformada de Hotelling é o alinhamento do eixo principal dos dados com o maior autovalor encontrado em um novo sistema de

coordenadas cuja origem é o centroide da população. Dessa forma essa transformada alinha os dados com os autovetores da matriz de covariância.

A tabela 1 apresenta o conjunto de variáveis denotadas por $X = (X_1, \dots, X_n)$ que foram observadas sobre as $n = 9$ empresas. As componentes principais C_i são definidas como:

$$C_i = \sum a_{ij} X_j, \quad \text{sujeito à:} \quad (12)$$

$$\text{var}(C_i) = \text{máxima e } \sum a_{ij}^2 = 1.$$

Quanto maior for a proporção da variância total contida na C_i melhor.

2.5 A Decomposição em Valores Singulares (SVD)

A decomposição em valores singulares está calcada na seguinte propriedade da diagonalização usual que pode ser aplicada em matrizes retangulares: os valores absolutos dos autovalores de uma matriz simétrica A medem as quantidades que A *estica* ou *contrai* os autovalores. (LAY, 2007, p.428) apresenta a seguinte comprovação: Se $Ax = \lambda x$ e $\|x\| = 1$, então

$$\|Ax\| = \|\lambda x\| = |A| \|x\| = |\lambda| \quad (13)$$

2.6 O comando $[U, S, V] = \text{svd}(C_x)$ do software Matlab

Retorna uma matriz diagonal, denominada S , da mesma dimensão de C_x , matriz de covariância, com elementos diagonais não negativos e em ordem decrescente. Produz ainda matrizes unitárias U e V de modo que $C_x = U*S*V^T$. A matriz V retorna os vetores próprios associados aos valores próprios da matriz diagonal S .

O comando $[\text{COEFF}, \text{SCORE}, \text{LATENT}, \text{TSQUARED}] = \text{princomp}(X)$ do Matlab

Retorna os coeficientes das componentes principais e as raízes latentes (autovalores) da matriz X entre outros valores. Esses coeficientes são os autovetores da matriz de covariância C_x .

3. MATERIAIS E MÉTODOS

Foi utilizado o *software* Matlab. A base de dados é proveniente do *Anuário do Agronegócio* de 2012. Os dados que foram analisados referem-se ao setor de alimentos e dentre esses as rações para animais. A tabela 1 apresenta as empresas (objetos) e as variáveis (indicadores). A empresa Nutrisal não foi considerada nos cálculos e a variável Evolução da Receita Líquida também não foi considerada nos cálculos.

Tabela 1 – Indicadores de Desempenho

Empresas	Ren.PL	G. Ativo	Ma.L	Ma.Ativ	Rec.Liq.	E.R.Liq.
Nutron Alimentos	46,0	2,73	8,3	12,3	487,17	6,0
Amireia Pajoara	58,8	2,65	15,6	15,3	7,79	22,1
Marinho Sebo	34,2	1,30	22,8	26,3	13,94	44,9
Rações Guabi	15,3	3,05	1,5	3,2	328,44	17,9
Premix	18,2	2,05	7,4	13,4	133,77	17,8
Agropecuária Carrilho	3,7	0,02	153,7	83,3	0,13	11,4
Nutrisal	27,6	7,49	0,1	-	276,61	759,2
Rações Total	5,8	1,71	1,9	2,8	237,88	26,0
Supra	5,9	1,82	1,3	1,9	246,08	10,2
Sebo Mariense	200,3	0,90	15,6	15,0	35,42	31,8

Fonte: Anuário do Agronegócio 2012.

O emprego da análise de agrupamento pelo critério de Ward e utilizando a distância métrica de Mahalanobis foi o método adotado para selecionar as variáveis com maior correlação cofenética bem como ordenar as empresas segundo seus indicadores de desempenho. Na construção do Índice pela ponderação das variáveis levando em conta o valor dos pesos de ponderação foi utilizada a *Transformada de Hotelling* também chamada de componentes principais.

Considerando que os indicadores são as amostras de uma população aleatória, os vetores de observação dos indicadores são expressos como:

$$X = [x_1, x_2, x_3; \dots; x_n]^T \quad (14)$$

O vetor *médio das amostras* é calculado como segue:

$$m_x = \frac{1}{n} \sum_{i=1}^n x_i \quad (15)$$

A matriz de covariância dessas amostras é calculada assim:

$$C_x = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_x)(x_i - m_x)^T \quad (16)$$

Os autovalores e autovetores da matriz quadrada C_x , real e simétrica, são determinados pela equação característica $(C_x - \lambda I)x = 0$. Como consequência os autovetores são ortogonais. A matriz **A** tem suas linhas formadas por esses autovetores sendo esses ordenados em correspondência com os autovalores associados de forma descendente, ou seja, do maior para o menor.

$$A = \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix} \quad (17)$$

A transformada de Hotelling é determinada pela equação (CATTIN, 2009, p. 85)

$$y = A(x - m_x) \quad (18)$$

onde A são os autovetores ou vetores próprios da matriz de covariância de X .

O efeito de usar essa equação é o de estabelecer um novo sistema de coordenadas, cuja origem está no centro de gravidade da população e cujos eixos estão na direção dos vetores próprios de C_x .

A transformada de Hotelling tem as seguintes propriedades:

- A média dos vetores resultantes y é igual a zero, isto é:

$$m_y = 0 \quad (19)$$

- A matriz de covariância dos y 's pode ser obtida por

$$C_y = VC_xV^T \quad (20)$$

- C_y é uma matriz diagonal onde os elementos são os autovalores de C_x .

$$C_y = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \quad (21)$$

- C_x e C_y tem os mesmos valores próprios e os mesmos vetores próprios.
- Como as linhas de V são vetores ortogonais a inversa de V é igual à transposta dele.
- Qualquer vetor x pode ser recuperado a partir do seu correspondente y pelo uso da seguinte equação:

$$x = V^T y + m_x \quad (22)$$

4. RESULTADOS

As análises de clusters foram executadas com todas as variáveis e também com a eliminação de cada uma delas para determinar os valores dos coeficientes de correlação cofenéticos.

O maior coeficiente de correlação cofenético obtido foi $r_c = 0.78$. Todas as variáveis exceto a Evolução da Receita Líquida (ERL) foram consideradas. A figura 1 apresenta o dendrograma pelo critério de Ward com emprego da distância métrica de Mahalanobis.

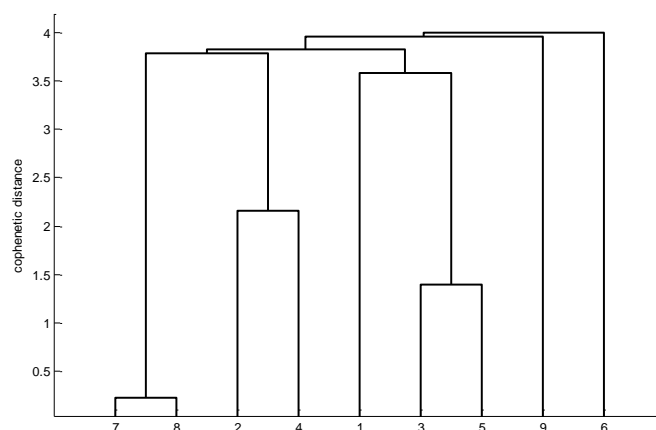


Figura 1: Dendrograma
Fonte: *output* do Matlab

A matriz **S** dos valores próprios é:

$$S = \begin{bmatrix} 30377.37 & & & & \\ & 4077.42 & & & \\ & & 1747.98 & & \\ & & & 12.23 & \\ & & & & 0.29 \end{bmatrix}$$

O traço da matriz diagonal $S = 36215.29$. A contribuição de λ_1 : $\frac{30377.37}{36215.29} = 0.84$. A matriz dos vetores próprios é:

$$V = \begin{bmatrix} -0.11 & -0.87 & -0.48 & 0.01 & -0.00 \\ 0.00 & -0.00 & 0.01 & -0.02 & -1.00 \\ -0.14 & 0.44 & -0.77 & 0.45 & -0.02 \\ -0.08 & 0.21 & -0.38 & -0.89 & 0.02 \\ 0.98 & -0.02 & -0.19 & -0.01 & 0.00 \end{bmatrix}$$

A tabela 2 apresenta os coeficientes da 1ª componente principal.

Tabela 2.- Coeficiente da C_1

Variáveis	Coeficientes
Rentabilidade do P. Líquido	-0.11
Giro do Ativo	0.00
Margem Líquida	-0.14
Margem da Atividade	-0.08
Receita Líquida	0.98

A componente principal é: $y = -0.11RPL + 0.00GA - 0.14ML - 0.08MA + 0.98RL$.

5. CONCLUSÕES

Os coeficientes negativos dos indicadores de Rentabilidade do Patrimônio Líquido, Margem Líquida e Margem da Atividade reduzem o valor do índice, mas, esse é positivamente expressivo para algumas empresas mas também é negativamente expressivo para outras.

Levando em conta que as empresas analisadas tem estruturas administrativas diferentes, e operam em condições de mercado também diferentes, o indicador Receita Líquida do tempo T_1 deve, como em qualquer empresa, superar sempre o indicador T_0 .

Quando há possibilidade de efetuar essa análise mensalmente ou em tempo mais curto, no âmbito da unidade empresarial, com certeza mostraria a tendência da gestão quanto ao lucro, receita, custos, ativo total e patrimônio líquido permitindo correção de rumo caso haja necessidade e otimização do lucro como consequência.

REFERÊNCIAS

- ANDERSON, T. W. **An Introduction to Multivariate Statistical Analysis**. New York: John Wiley & Sons, 1974. 374 p.
- CATTIN, Philippe. **Signalprocessing Biomedical Image Analysis**. Basileia: MIAC, University of Basel.
- GLOBO RURAL. **ANUÁRIO DO AGRONEGÓCIO 2012**. São Paulo: Editora Globo, 2012.
- GOWER, J. C.; LEGENDRE, P. **Metric and Euclidean Properties of Dissimilarity Coefficients**. Journal of Classification, New York. Springer-Verlag, v. 3, p. 5-48, 1986.
- HOTELLING, Harold. **ANALYSIS OF A COMPLEX OF STATISTICAL VARIABLES INTO PRINCIPAL COMPONENTS**. The journal of Educational Psychology. 1933.
- KACHIGAN, S. K. **Multivariate Statistical Analysis: a conceptual introduction**. 2. ed. New York: Radius Press, 1991.
- KOPP, M. M., DE SOUZA, V. Q., COIMBRA, J. L. M, LUZ, V. K., MARINI, N. & OLIVEIRA, A.C. 2007. Melhoria da correlação cofenética pela exclusão de unidades experimentais na construção de dendrogramas. Rev. Fac. Zoo. Vet. e Agr. 14(2):46-53.
- KUBRUSLY, Lucia Silva. **Um procedimento para calcular índices a partir de uma base de dados multivariados**. Pesquisa Operacional. Rio de Janeiro. v. 21, n. 1, p. 107-117, jun. 2001.
- LAY, David C. **Álgebra Linear e suas aplicações**. Tradução Ricardo Camelier, Valéria de Magalhães Iório. 2. ed. Rio de Janeiro: LTC, 2007. LEON, Steven J. Tradução Valéria de Magalhães Iório. 4. ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos, 1999.

MACHLAN, Geoffrey John. **Mahaanobis Distance**. RESONANCE, june 1999. p. 20-26. Disponível em: < <https://etd.lib.metu.edu.tr/upload/12612852/index.pdf> >. Acesso em 05 jan 2013.

MAHALANOBIS, P. C. **ON THE GENERALIZED DISTANCE IN STATISTICS**. Proceedings of the National Institute of Sciences of India. Calcutta. v. 2, 49-55, 1936.

MITTERSTEIN. Mônica R.; SEVERO, Dirceu L. **Análise de variabilidade intrassazonal e interanual da precipitação no vale do Itajaí com a transformada de ondaletas**. Dynamis revista técnico-científica. Santa Catarina. v. 13, n. 1, p. 01-10, out-dez.2007.

NETO, Alexandre Assaf. **Estrutura e Análise de Balanços: um enfoque econômico-financeiro**. 7. ed. São Paulo: Atlas, 2002.

REIS, E. **Estatística Multivariada Aplicada**. 2. ed. Lisboa: Silabo, 2001.

ROMESBURG, H.C. 1985. Cluster analysis for researchers. Lifetime Learning Publications, Belmont.

THEIL, H. **Best Linear Index Numbers of Prices and Quantities**. *Econometrica*. v. 28, n.2, 464-480, abril 1960.

TRIOLA, Mario F. **Introdução à Estatística**. Tradução Alfredo Alves de Farias. 7. ed. Rio de Janeiro: Livros Técnicos e Científicos, 1999.

VALENTIN, Jean Louis. **Ecologia Numérica: uma introdução à análise multivariada de dados ecológicos**. Rio de Janeiro: Interciência, 2000.

XIMENES, Lidiani Queli Lubas; MATEUS, Lúcia Aparecida de Fatima e PENHA, *Jerry Magno Ferreira*. **Varição temporal e espacial na composição de guildas alimentares da ictiofauna em lagoas marginais do rio Cuiabá, Pantanal Norte**. *Biota Neotrópica*. v. 11, n.1, 205-206, 2011.

WEINBERGER, Kilian Q. ; SAUL, Lawrence K. **Distance Metric for Large Margin Nearest Neighbor Classification**. *Journal of Machine Learning Research*. v.10, 207-244, 2009.