

Classificação de Refrigerantes por Parâmetros Físico-Químicos e PCA

Sérgio B. de Oliveira, Raquel O. Pereira, Keisy O. Ribeiro e Anselmo E. de Oliveira

PCA é empregada como metodologia de reconhecimento de padrões de 18 marcas de refrigerantes dos tipos cola, laranja e guaraná, de acordo com as características pH, sacarose, acidez total, índice de refração, viscosidade e os teores de ácido sórbico e benzoico. Gráficos dos *scores* das duas componentes principais explicam cerca de 60% da variância total dos dados. Sacarose, ácido benzoico e a acidez total são as características principais para PC1, enquanto que os *scores* para a PC2 resultaram das características de viscosidade relativa e do teor de ácido sórbico.

Palavras-chave: *PCA; refrigerantes; reconhecimento de padrões.*

PCA has been applied as a pattern recognition methodology on 18 commercial carbonated soft drink brands. Cola, orange and guaraná flavors were classified according to pH, sucrose, total acidity, refraction index, viscosity and sorbic and benzoic acid characteristics. Score graphs of the first two principal components explained about 60% of total data variance. Sucrose, benzoic acid and total acidity were the main characteristics for PC1, while PC2 scores were result of sorbic acid and viscosity..

Keywords: *PCA; carbonated soft drinks; pattern recognition.*

Introdução

O comércio de refrigerantes, bebida não fermentada, adoçada e aromatizada artificialmente, cujo teor de água representa em torno de 98% da sua composição, tem crescido rapidamente, tornando-se um excelente negócio para as empresas. No estado de Goiás esse comércio vem crescendo sensivelmente e produziu, em 2002, cerca de 275 milhões de litros¹. Sabe-se que a água representa o principal veículo de contaminação dessa bebida. Outros fatores intrínsecos são derivados da natureza do produto e da sua formulação, os quais não são fáceis de mudar, nem controlar. Dentre os fatores intrínsecos podem ser citados: atividade em água, nutrientes disponíveis, inibidores, acidez, carbonatação, potencial de oxidação e agentes estabilizantes. Destes, o pH é um dos principais fatores capazes de determinar o crescimento, sobrevivência e destruição dos microorganismos presentes nos refrigerantes.

O Brasil se destaca no setor de bebidas por ser o terceiro produtor mundial de refrigerantes, produzindo mais de 10 bilhões de litros. Três países competem com seus produtos regionais na disputa da liderança com o sabor cola das multinacionais: França (laranja), Coréia do Sul (ginseng) e Brasil (guaraná). Para avaliar o binômio qualidade/quantidade do produto fabricado, determinações analíticas podem ser associadas com tratamentos estatísticos de dados, para melhor classificar a qualidade de gêneros alimentícios, empregando metodologias como o Reconhecimento de Padrões (RP), através da Análise de Componentes Principais (PCA).

Métodos de RP são bastante utilizados para a classificação de produtos alimentícios^{2,3}, e vários são os fatores ambientais inter-relacionados que determinam a susceptibilidade de um produto à contaminação e à deterioração, daí a necessidade em se monitorar a qualidade dos produtos que são consumidos pela população. O presente trabalho emprega métodos analíticos específicos para refrigerantes, coleta e conservação de amostras de três diferentes sabores de refrigerantes (cola, guaraná e laranja) comercializados no estado de Goiás, realização de análises químicas das amostras e tratamento quimiométrico dos dados pelo método de PCA. Com esses resultados foi possível estabelecer padrões que detectam prováveis adulterações nos refrigerantes.

Metodologias

ANALÍTICAS

Para caracterizar os três tipos de refrigerantes (classes) entre as 18 marcas escolhidas, em função do grau de aceitabilidade pela população, volume de produção e variedade de marcas, foram feitas as análises de ácido benzóico, ácido sórbico, pH, acidez, teor de sacarose, viscosidade relativa e índice de refração⁴⁻⁷. Das 18 marcas de refrigerantes analisadas, 7 são comuns aos três sabores, sendo que as demais se subdividem em: 4 marcas do sabor cola, 5 guaraná e outras duas marcas do sabor laranja. No total, foram analisadas 11 marcas de refrigerantes sabor cola, 12 sabor laranja e 9 sabor guaraná. Com base nesses resultados foi feita uma análise estatística multivariada dos dados, determinando-se quais as variáveis mais representativas para caracterizar as marcas de refrigerantes, relativas à sua qualidade. Alguns métodos utilizados foram baseados em determinações já consolidadas, que são realizadas em laboratórios de rotina de indústrias, com o objetivo de simular de forma mais realista o modelo de controle de qualidade experimentado pelas empresas da área.

As determinações de ácido benzóico (AB) e ácido sórbico (AS) foram realizadas pelo método espectrofotométrico com extração prévia⁸. As determinações do pH e da acidez total (AT) foram feitas por titulação potenciométrica⁹. O teor de sacarose (S) foi determinado pelo método de Lane-Enyon¹⁰, que é um método quantitativo, redutométrico e titulométrico, empregando-se o reagente de Fehling. A medida da viscosidade relativa (VR) foi baseada na determinação das viscosidades cinemáticas (método de Ostwald) das amostras dos refrigerantes em relação à da água, numa temperatura de $30,00 \pm 0,02$ °C. Para essa determinação utilizou-se o viscosímetro automático SCHOTT CT52 adaptado à bomba AVS 350, com capilar SCHOTT Gerate nº 52010 App. 1010072. A medida do índice de refração relativo (IR), que é a razão entre o índice de refração da amostra e o índice de refração da água, foi executada no refratômetro Analytik Jane, termostatizado à temperatura de 25 °C.

ESTATÍSTICAS

Os dados obtidos para cada um dos três tipos de

refrigerantes foram pré-processados com a modificação das variáveis originais em novas variáveis, através da PCA, com propriedades matemáticas mais convenientes para análise. A metodologia estatística, PCA, consiste em uma mudança de base dos dados multivariados, representados pela matriz X , contendo os resultados das análises físico-químicas para cada categoria, encontrando o vetor \vec{v} , chamado autovetor, e um escalar, τ , chamado autovalor, para o qual a igualdade $\text{cov}(X) \cdot \vec{v} = \tau \cdot \vec{v}$ é satisfeita. $\text{cov}(X)$ é o auto-sistema da covariância dos dados multivariados, de maneira que transforme um espaço multidimensional em um subespaço menor, como por exemplo, de dimensão um, gerado por um autovetor¹¹. Ou seja, X (multivariável representada por muitos vetores) passa a ter as informações relevantes, representadas por um ou mais autovetores. Esse subsistema deve reter as informações relevantes e representar, de modo mais adequado, cada variável individualmente. Desse modo, PCA descreve, em termos geométricos das componentes principais, a covariância das variáveis (no presente caso, as medidas físico-químicas) usando o menor número de autovetores. O novo subsistema consiste, então, dos eixos das componentes principais (Figura 1) sendo as coordenadas das amostras nesses novos eixos denominados de *scores*. Como cada componente principal é uma combinação linear das variáveis originais, os coeficientes dessas variáveis (ou seja, os autovetores) são chamados de *loadings*¹².

Além da PCA, também foram empregados os pesos de Fisher¹³. Esses pesos são calculados pela razão da distância entre os valores médios de duas distribuições (classes/sabores dos refrigerantes) e a soma das variâncias, para determinar o poder de discriminação das variáveis.

Resultados e Discussão

PESOS DE FISHER

A habilidade das três classes foi testada usando os pesos de Fisher para a classificação dos refrigerantes. Uma análise visual é suficiente para separar uma amostra (objeto) entre as classes. Porém, como o objetivo do presente trabalho é a proposição de um modelo de classificação, esses pesos não são inadequados, uma vez

que a determinação de características estatisticamente importantes deve ser somada à análise intuitiva. Na Tabela 1 encontram-se os valores obtidos para os pesos de Fisher para a separação das três classes. Como o maior valor do peso significa maior capacidade de separação entre as classes, a separação entre cola e guaraná se dá, principalmente, com base no pH (peso = 2,473), além da acidez total (peso = 1,706). Esse resultado decorre do fato que, apesar da maior diferença entre os valores médios obtidos para AT, os valores de pH apresentam uma menor variância. Um gráfico entre essas duas características, AT e pH, para as classes cola e guaraná deve resultar uma boa separação como pode ser verificado na figura 2. Essas duas características são igualmente importantes entre as classes cola e laranja, Tabela 1 e Figura 3. Já para as classes laranja e guaraná, as características AT e IR são as mais importantes, embora seus valores sejam relativamente menores que nas outras classificações. Conseqüentemente, um gráfico entre os valores de AT e IR para as classes guaraná e laranja não deve resultar numa boa separação, conforme a Figura 4. Nesse caso, uma pequena diferenciação pode ser percebida para valores de AT acima de 140 mg/100 mL, característica de algumas marcas dos refrigerantes sabor guaraná.

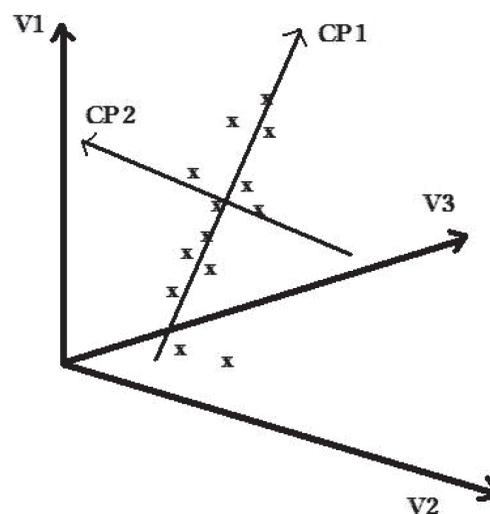


Figura 1: Representação dos eixos das duas primeiras componentes principais (PC1 e PC2) resultantes da mudança de base dos eixos das variáveis originais (V1, V2 e V3).

Tabela 1: Pesos de Fisher para a separação das três classes.

Característica	cola x guaraná	cola x laranja	guaraná x laranja
pH	2,473	2,694	0,086
AB	0,360	0,767	0,040
AS	0,292	0,560	0,243
AT	1,706	4,509	0,765
S	0,002	0,172	0,117
VR	0,062	0,290	0,482
IR	0,021	0,556	0,960

AB, AS, AT e S em mg/100 mL.

ANÁLISE DE COMPONENTES PRINCIPAIS

A redução da dimensionalidade da matriz de dados originais em um gráfico de duas dimensões, para estudar a importância das sete características e, então, poder classificar cada um dos tipos de refrigerantes, separadamente, foi feita pelo método da Análise de Componentes Principais. Os dados foram auto-escalados por classe. Na Tabela 2 são mostrados os resultados da transformação Karhunen-Loeve. Para a classe guaraná as duas primeiras componentes principais explicam cerca de 70 % da variância total dos dados. Os resultados obtidos para as outras duas classes (cola e laranja) para as duas primeiras componentes principais (PC) explicam cerca de 60 % da variância total dos dados. Os gráficos das duas primeiras componentes principais (PC1 x PC2), para cada uma das três classes analisadas, constam das Figuras 5, 6 e 7.

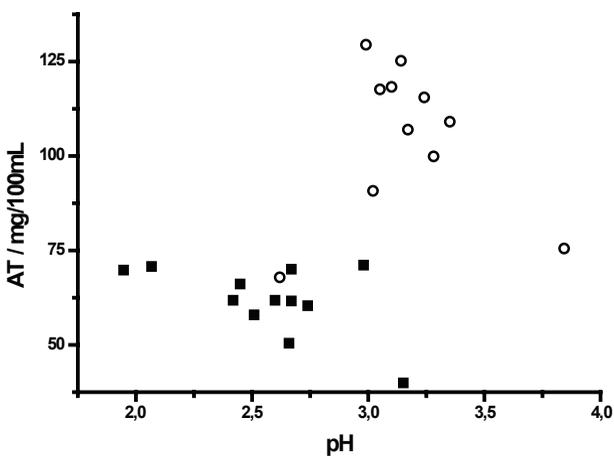


Figura 2: Gráfico do pH versus a acidez total, AT, em mg/100 mL para as classes cola (■) e guaraná (○).

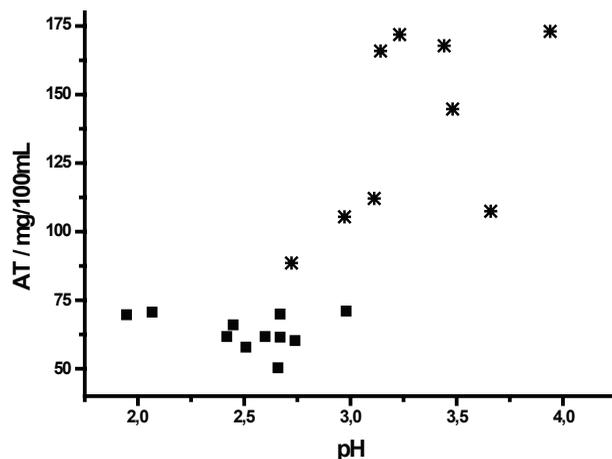


Figura 3: Gráfico do pH versus a acidez total, AT, em mg/100 mL, para as classes cola (■) e laranja (*).

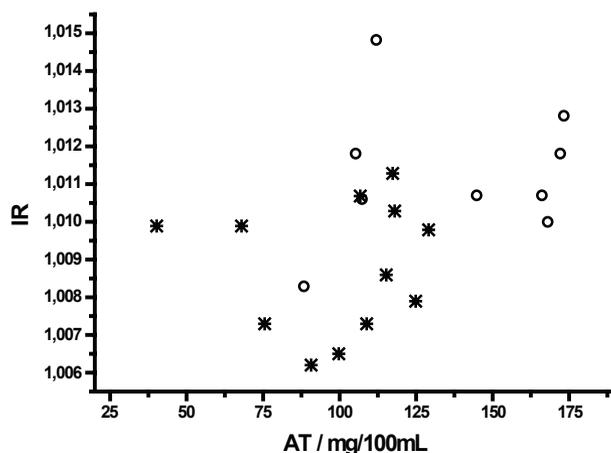


Figura 4: Gráfico da acidez total, AT, em mg/100 mL versus o índice de refração relativo, IR, para as classes laranja (*) e guaraná (○).

A classe cola apresenta PC1 com cerca de 30 % da variância explicada e é caracterizada, principalmente, pelas variáveis AB, AS e S. Desse modo, as amostras 4, 6, 7, 9 e 11, as quais apresentam os maiores teores de sacarose (acima de 11 g/100 mL) ocupam o lado direito da figura 5. Por outro lado, a amostra 5 apresentou o maior valor de AB (≈ 32 mg/100 mL) estando mais à esquerda. Já PC2 é caracterizada pelas variáveis pH, VR e IR. Como o autovetor/peso da variável AT é negativo nessa componente principal, as marcas que se encontram na parte inferior dessa figura são as que apresentam menores índices de acidez total. Algumas marcas apresentam boa

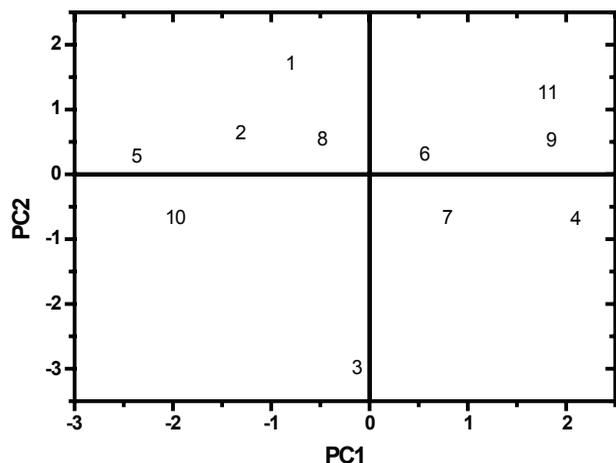


Figura 5: Gráfico dos scores das duas primeiras componentes principais, PC1 e PC2, para a categoria cola. Esse gráfico explica 56,81 % da variância total dos dados.

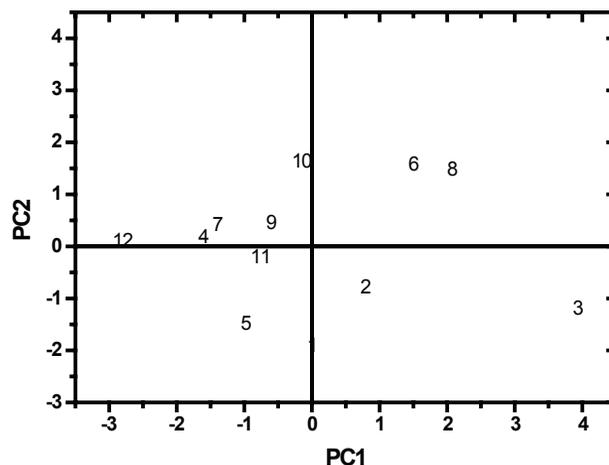


Figura 6: Gráfico dos scores das duas primeiras componentes principais, PC1 e PC2, para a categoria guaraná. Esse gráfico explica 67,83 % da variância total dos dados.

similaridade, já outras ocupam espaços bem definidos no gráfico das componentes principais.

Já para a classe guaraná, os autovetores para PC1 referentes às variáveis pH, AB e AS apresentam sinais opostos aos das outras quatro variáveis - similar ao verificado para a classe cola -, como verificado na Tabela 2. A amostra 3 apresentou o pH menos ácido (3,84); as amostras 6 e 8 apresentam valores próximos para as características pH, AS, S, VR e IR; as amostras 4 e 7 apresentam similaridades em AS, VR e IR; as amostras 9 e 11 apresentam similaridades decorrentes, principalmente, do pH e VR. Em PC2, as amostras abaixo desse eixo decorrem da combinação dos valores de IR (altos) com baixos valores para as outras características. O gráfico

das duas primeiras componentes principais para a classe cola consta da Figura 6.

Para a classe laranja, Figura 7 e Tabela 2, as características AB e S são as de menor significância para a classificação das marcas dentro dessa classe. O maior peso é dado pela variável AT, e a amostra 4 (AT = 173 mg/100 mL) é a que se apresenta mais à direita no gráfico das componentes principais, com PC1 explicando cerca de 40 % da variância dos resultados experimentais para essa classe. Por conseguinte, a amostra 3 é a que apresenta o menor valor (AT = 88 mg/100 mL) para essa característica, situando-se mais à esquerda. Em PC2 AS e AB apresentam os maiores pesos, porém indicam tendências opostas. Portanto, a amostra 7, a

Tabela 2: Informações obtidas pela transformação de Karhunen-Loeve para as classes cola, guaraná e laranja.

Classe	Autovalores	Variância explicada (%)	Variância acumulada (%)	Autovetoresa						
				pH	AB	AS	AT	S	VR	IR
cola	(PC1) 4,90	34,33	34,33	-0,25	-0,51	-0,44	0,08	0,58	0,17	0,34
	(PC2) 3,97	22,48	56,81	0,54	0,04	0,04	-0,33	-0,08	0,65	0,40
guaraná	(PC1) 6,06	47,69	47,69	0,40	0,35	0,23	-0,20	-0,48	-0,45	-0,44
	(PC2) 3,94	20,14	67,83	0,14	0,22	0,64	0,62	0,24	0,25	-0,15
laranja	(PC1) 4,58	37,50	37,50	0,45	-0,13	-0,02	0,52	0,38	0,49	0,34
	(PC2) 3,64	23,64	61,14	-0,13	-0,49	0,51	0,28	0,40	-0,43	-0,24

^a conforme rodapé da tabela 1.

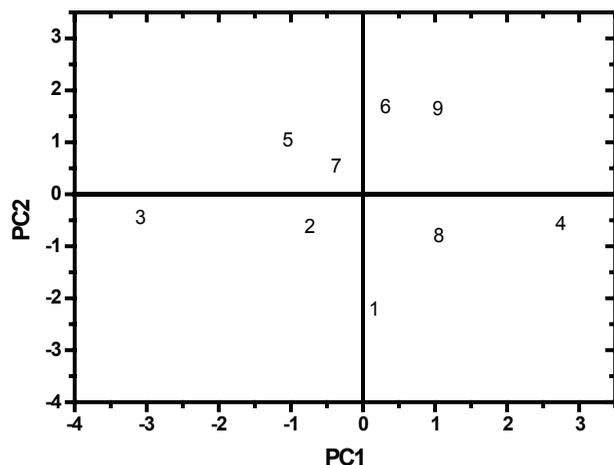


Figura 7: Gráfico dos scores das duas primeiras componentes principais, PC1 e PC2, para a categoria laranja. Esse gráfico explica 61,14 % da variância total dos dados.

qual apresenta alto valor para AB, e também para AS, se encontra quase que no centro desse eixo. Já amostra 1, com alto teor de AB e baixo de AS se encontra bem abaixo. Essa classe apresenta uma boa separação entre os tipos de refrigerantes.

Conclusões

Com base nos gráficos das componentes principais é possível estabelecer modelos de classificação de modo a identificar amostras adulteradas de refrigerantes. O teor de sacarose é importante para classificar refrigerantes dos sabores cola e guaraná. Para as colas, os teores dos ácidos sórbico e benzóico também devem ser considerados. Já para os guaranás e as laranjas, essas variáveis têm menor peso na classificação.

REFERÊNCIAS

1. Jornal Diário da Manhã, Caderno Universidade, Goiás, 13 de Outubro de 2002.
2. Togari, N.; Kobayashi, A.; Aishima, T.; *Food Research International* **1995**, 495, 28.
3. Martín, M.J.; Pablos, F.; González, A.G.; *Analytica Chimica Acta* **1996**, 191, 320.

4. Germain, L.; Colas, L.; Rouquet, J.; Tratamento de Águas, Editora Polígono S.A.: São Paulo, **1972**.
5. Ewing, G.W.; Métodos Instrumentais de Análise Química, vol.1, Editora Edgard Blücher Ltda: São Paulo, **1972**.
6. Vogel, A.; Bassett, J.; Jeffery, G.H.; Mendhan, J.; Denneey, R.C.; *Análise Química Quantitativa*, Editora Guanabara Koogan: Rio de Janeiro, **1992**.
7. AOAC Official Method of Analysis, **1995**.
8. Grupo Antártica laboratório central, Instrução Técnica: LC.4404,027, Data de aprovação: **15/03/1995**.
9. Grupo Antártica laboratório central, Instrução Técnica: LC.4404,031, Data de aprovação: **19/12/1997**.
10. Lane, J.H.; Eynon, L.; *Determination of reducing sugars by Fehling's solution with methylene blue indicator*, Normam Rodge: London, **1934**.
11. Noble, B.; Daniel, J.W.; Álgebra Linear Aplicada, Prentice-Hall do Brasil: Rio de Janeiro, **1986**.
12. Ferreira, M.M.C.; Antunes, A.M.; Melgo, M.S.; Volpe, P.L.O.; *Quim. Nova* **1999**, 22, 724.
13. Sharaf, M.A.; Illman, D.L.; Kowalski, B.R.; *Chemometrics*, Wiley: New York, **1986**.

Sérgio B. de Oliveira¹, Raquel O. Pereira¹, Keisy O. Ribeiro¹, Anselmo E. de Oliveira*²

¹ CEFET, Setor Central, 74055-110, Goiânia, GO

² Instituto de Química, UFG, CP 131, 74001-970, Goiânia, GO.

*E-mail: elcana@quimica.ufg.br