

Verbetes Metodológico

Análise de correspondências múltiplas

Edison Bertoncelo¹ 

¹Professor da Universidade de São Paulo, USP, São Paulo, SP, Brasil.

E-mail: edison.bertoncelo@usp.br

DOI: [10.17666/ebcs.2026.044](https://doi.org/10.17666/ebcs.2026.044)

PREÂMBULO

A análise de correspondências múltiplas (ACM) é uma técnica de análise de dados categóricos, isto é, de variáveis que registram características, qualidades ou atributos, como gênero, escolaridade, voto ou gostos, e integra o conjunto de técnicas conhecido como análise geométrica de dados. A partir de uma tabela de indivíduos descritos por muitas variáveis categóricas, a ACM constrói mapas em que categorias e indivíduos aparecem como pontos, de modo que distâncias e proximidades expressam afinidades e oposições. A técnica responde a perguntas sobre a estrutura de um conjunto de relações: que práticas, atitudes e propriedades tendem a aparecer juntas, quais se opõem e como os grupos sociais se distribuem nesse espaço. Por isso, tem sido amplamente usada nas ciências sociais em pesquisas sobre classes sociais, gostos e estilos de vida, trajetórias sociais e educacionais e preferências políticas. O verbete apresenta a história da técnica e seus fundamentos epistemológicos, seus principais procedimentos, suas vantagens e limites, três exemplos de aplicação e os debates atuais sobre análise de subgrupos, combinação com técnicas de classificação e de redes e inferência estatística.

SUMÁRIO

1. Histórico e fundamentos epistemológicos
2. Procedimentos principais
3. Vantagens, limites de validade e críticas conhecidas
4. Exemplos de aplicação em ciências sociais
5. Debates contemporâneos e variações
6. Referências bibliográficas

Aprovado em: 01/10/2026

Editor(a) responsável: Adriano Codato



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/>), que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Enciclopédia Brasileira de Ciências Sociais (EBCS). ANPOCS, 044 - 2026

1. Histórico e fundamentos epistemológicos

A análise de correspondências múltiplas (ACM) é uma ferramenta da análise de dados quantitativos e integra um conjunto mais amplo de técnicas denominado análise geométrica de dados (AGD). A AGD é uma abordagem da estatística multivariada que possibilita representar conjuntos de dados multivariados sob a forma de nuvens de pontos e baseia a interpretação dos dados na observação dessas nuvens (Le Roux & Rouanet, 2004).

Em comum, as técnicas incluídas sob essa denominação se caracterizam pela construção de espaços geométricos, nos quais as linhas e as colunas de uma matriz são representadas como pontos em um espaço euclidiano de alta dimensionalidade, e pela prioridade da descrição sobre a modelagem probabilística. Seus resultados básicos são estatísticas descritivas e, portanto, não dependem de pressupostos sobre a distribuição dos dados nem de testes de significância.

Como abordagem estatística, a AGD se desenvolveu nos trabalhos do matemático francês Jean-Paul Benzécri (1973) a partir da construção de uma ferramenta capaz de representar, sob a forma de nuvens de pontos, as relações entre as categorias em uma tabela de contingência. A construção dessa abordagem buscava contrabalançar a ênfase que a estatística dedicava então aos modelos confirmatórios. O princípio de Benzécri, segundo o qual o modelo deve ajustar-se aos dados e não os dados ao modelo, sintetiza a orientação indutiva dessa abordagem.

A AGD compreende três técnicas centrais. A análise de correspondências simples (ACS), como é designada atualmente, que permite analisar tabelas de contingência, representando graficamente a relação entre as categorias de duas variáveis (por exemplo: a classe social e as preferências políticas dos indivíduos). As outras duas aplicam o mesmo princípio geométrico a outros tipos de matriz de dados e de variáveis. A análise de correspondências múltiplas e a análise de componentes principais (ACP) são aplicadas a matrizes de dados do tipo *indivíduo por variável*. Diferem quanto à natureza do dado: a ACM analisa as relações entre muitas variáveis categóricas, enquanto a ACP foca nas relações entre os valores de variáveis contínuas. Em comum a essas duas técnicas (ACP e ACM), reside a possibilidade de construir duas nuvens distintas de pontos: a nuvem de modalidades (ACM) ou de correlações (ACP) e a nuvem de indivíduos. Nessas vertentes da AGD, são os indivíduos que carregam toda a informação analisada (Lebaron, 2006/2026).

A abordagem construída por Benzécri nos anos 1970 foi considerada bastante inovadora no campo intelectual francês ao longo dessa década, mas só veio a obter reconhecimento internacional a partir das décadas seguintes, período em que a AGD passou a ser discutida com maior frequência em publicações de língua inglesa e em manuais metodológicos, com destaque para os trabalhos de Michael Greenacre e Jörg Blasius (Blasius & Greenacre, 2014; Greenacre & Blasius, 2006). Há também manuais sobre os usos da AGD nas ciências sociais (Bertoncello, 2022; Hjelbrekke, 2019). Atualmente, as técnicas da AGD estão disponíveis em diversos programas de análise de dados quantitativos.

Quanto aos fundamentos epistemológicos, os usos da AGD privilegiam uma orientação indutiva da análise, uma leitura relacional dos dados e uma concepção estrutural de

causalidade. Em vez de construir modelos e compará-los em função do ajuste aos dados, parte-se da observação das nuvens de pontos para apreender as principais dimensões que estruturam as afinidades e oposições entre os agentes e suas propriedades. Em lugar de separar variáveis dependentes e independentes e isolar efeitos líquidos, busca-se reconstruir um espaço de relações, em que a causalidade social consiste nos efeitos globais do conjunto de inter-relações entre as variáveis. Essa orientação possui uma “afinidade eletiva” com a sociologia de Pierre Bourdieu, sobretudo com os conceitos de campo, espaço social e espaço simbólico. Conforme o autor, a análise de correspondências “pensa” em relações: uma propriedade qualquer de um agente (por exemplo: possuir um piano em casa) não possui sentido isoladamente, mas apenas em contraste ou afinidade com outras práticas e com as posições sociais dos agentes (ter um piano em casa é frequentemente um indicador bastante preciso do pertencimento à burguesia e da incorporação do gosto burguês na França estudada por Bourdieu).

2. Procedimentos principais

A ACM é uma técnica de análise de variáveis categóricas, que reúne dados que permitem mensurar características, qualidades ou atributos da unidade de observação (o gênero de uma pessoa, o partido político em que um eleitor vota, o nível de instrução, gostos etc.). É uma ferramenta, portanto, que serve para descrever, explorar, resumir e visualizar a informação contida em uma tabela de N indivíduos descritos por Q variáveis categóricas, representando as relações entre categorias e indivíduos sob a forma de distâncias e proximidades relativas em um espaço geométrico (Husson & Josse, 2014).

Matrizes de dados desse tipo são frequentemente o resultado da aplicação de questionários estruturados a uma amostra de indivíduos, em que cada célula da matriz contém a resposta do indivíduo a uma pergunta específica. Para cada questão, há um número finito de categorias (ou modalidades) e cada indivíduo escolhe uma única categoria de resposta.

Com a ACM, é possível produzir duas nuvens de pontos, representando as relações entre categorias e indivíduos sob a forma de distâncias e proximidades relativas em um espaço multidimensional. Nesse sentido, a ACM possibilita uma análise descritiva, explorando as relações ou associações entre as categorias das variáveis retidas, e uma análise explicativa, na medida em que permite apreender os possíveis determinantes das tomadas de posição (comportamentos ou atitudes), examinando como os indivíduos com diferentes características (em termos de classe social, renda, escolaridade, gênero, raça, idade etc.) se distribuem no espaço.

As nuvens de modalidades e de indivíduos são construídas a partir da definição das variáveis ativas, cujas categorias são como “bolas magnetizadas” que possuem força de atração e repulsão. Se a pesquisadora deseja investigar práticas culturais, deve incluir as variáveis que mensuram aspectos desse objeto (gosto, conhecimento, interesse, visita a equipamentos culturais) como ativas para a construção dos espaços (Nicolau Netto et al., 2024). As distâncias entre dois indivíduos no espaço são produzidas pelas questões em que escolhem categorias diferentes.

Assim, se um conjunto de indivíduos gosta de Gustavo Lima, Aline Barros ou Ana Castela e da arte de Romero Britto e, diferentemente, outro conjunto deles rejeita esses cantores e o artista e também gosta de literatura e lê “Ensaio sobre a cegueira”, de José Saramago, então os pontos que representam essas categorias tenderão a ocupar posições opostas no espaço: de um lado, por exemplo, o gosto por esses artistas e, de outro, a rejeição a eles e a leitura da obra mencionada. O mesmo raciocínio vale para os indivíduos: quanto mais diferentes são seus perfis de resposta, maior tenderá a ser a distância entre os pontos que os representam no espaço. Além disso, por definição, quanto menos frequentes são as categorias escolhidas nas questões em que dois indivíduos divergem (categorias de discordância ou de desacordo), maior a distância entre eles, e quanto mais um indivíduo escolhe categorias pouco frequentes, mais distante do centro da nuvem estará o ponto que o representa. Ou seja, o centro da nuvem é o *tipo médio*, agrupando categorias mais frequentes e indivíduos com propriedades ou tomadas de posição mais frequentes. Os pontos mais distantes do centro são, em certo sentido, mais “distintos”: categorias ou indivíduos com propriedades ou tomadas de posição estatisticamente mais raras.

Essa característica da ACM, que sublinha a importância das categorias de desacordo menos frequentes para a construção de distâncias entre os indivíduos, é desejável, mas pode levar a distorções. Por exemplo: podemos caracterizar um grupo de indivíduos como sendo muito diferente de outro com base em um conjunto de categorias de resposta pouco frequentes (digamos, quando 5% dos indivíduos relatam gostar de um certo tipo de filme ou de um diretor de cinema), ainda que se assemelhem em termos de outras tomadas de posição mais comuns. Por isso, é recomendável que categorias pouco frequentes sejam excluídas da análise, recodificadas ou incluídas como passivas (ver seção 5).

A escolha de quais (e quantas) são as variáveis e categorias ativas é informada pelas questões que orientam a pesquisa. O pesquisador deve buscar um equilíbrio no número de categorias por variável e de variáveis por “tópico”. Voltando ao exemplo anterior, o pesquisador incluiria um número similar de variáveis e categorias para mensurar, por exemplo, gostos, conhecimento, interesse ou frequência de equipamentos em diferentes domínios culturais (música, TV, teatro, alimentação etc.). É importante considerar os critérios da *homogeneidade* e *exaustividade* na construção das variáveis ativas, o que significa que, para cada dimensão ou domínio do fenômeno investigado, o pesquisador incluirá variáveis construídas de forma similar e que mensuram os aspectos pertinentes do objeto.

Uma vez construído o espaço sob a forma de nuvens de pontos com base na escolha das variáveis ativas, o passo seguinte consiste em redefinir as dimensões do espaço de modo que as principais capturem a maior parte possível da variância, o que permite descrever os dados em menor número de dimensões (Greenacre & Blasius, 2006). Ocorre, assim, uma decomposição da variância da nuvem ao longo dos eixos, de modo que a variância projetada no primeiro eixo é máxima, decrescendo ao longo dos eixos seguintes. A parte da variância da nuvem projetada em cada eixo é dada pelos valores próprios. A determinação dos eixos que serão retidos para interpretação depende de critérios estatísticos e de ganhos analíticos. Em termos dos critérios estatísticos, deve-se observar especialmente a proporção da variância de cada eixo, que é a relação entre a variância da nuvem projetada no eixo e a variância total, considerando geralmente as taxas modificadas de Benzécri, que permitem corrigir a

subestimação da taxa de variância dos eixos iniciais. Retêm-se os primeiros eixos até que a taxa modificada acumulada seja suficiente (por exemplo: 70% a 80%). Além disso, é importante considerar os ganhos analíticos decorrentes da interpretação de cada eixo, ou seja, quanto a análise das estruturas de afinidades e oposições e de seus fatores estruturantes em cada eixo contribui para reconstruir empiricamente o objeto de estudo. Após a escolha de quantos eixos reter, cada eixo é interpretado considerando as localizações relativas das categorias com contribuições acima da média para a variância do eixo. Observando como essas categorias se distribuem ao longo do eixo (esquerda/direita; abaixo/acima), podemos apreender a estrutura de afinidades e oposições entre elas.

A interpretação das nuvens privilegia, assim, uma leitura relacional: o sentido das categorias é apreendido em função de sua localização no espaço em relação a outras (por exemplo: entende-se melhor o que significa gostar de Aline Barros quando sabemos que essa categoria de gosto se opõe, no mesmo eixo, à leitura de “Ensaio sobre a cegueira”). A leitura relacional das afinidades e contrastes em cada eixo resulta na criação de um rótulo, que sintetiza essas informações. Voltando ao exemplo anterior, é possível interpretar o eixo, considerando o conjunto das categorias pertinentes, em termos de uma oposição entre a adesão e a rejeição à cultura popular comercial.

Em seguida, projetam-se na nuvem de modalidades as categorias das variáveis suplementares (fatores estruturantes), que não servem para construir os espaços, mas contribuem para sua interpretação. Observamos como as categorias dessas variáveis se distribuem no espaço construído a partir dos eixos retidos, especialmente se tais categorias se agrupam no centro do espaço ou se elas se distribuem ao longo dos eixos. O aprofundamento da análise dos fatores que estruturam os eixos ocorre pela interpretação da nuvem de indivíduos: mais precisamente, pela construção de subnuvens a partir da decomposição da nuvem global em função das categorias das variáveis suplementares. Examina-se, por exemplo, como se distribuem os indivíduos conforme seu nível de escolaridade ou seu gênero no espaço por meio das elipses de concentração e da observação de seu formato e dispersão.

3. Vantagens, limites de validade e críticas conhecidas

A principal vantagem da ACM é representar simultaneamente relações complexas entre muitas variáveis categóricas, sem exigir qualquer diferenciação entre variáveis dependentes e independentes. Ela permite construir espaços de relações nos quais distâncias e proximidades revelam contrastes e afinidades entre práticas, atitudes, disposições e propriedades dos agentes. Com isso, reduz-se o risco de uma interpretação substancialista dos fenômenos investigados, que atribuiria a uma prática um significado social fixo independentemente das relações com outras práticas e de como se distribui socialmente.

A leitura conjunta da nuvem de modalidades e da nuvem de indivíduos permite passar das relações entre categorias aos padrões de resposta dos casos. Por meio das variáveis suplementares e das elipses de concentração, é possível investigar os fatores estruturantes do espaço examinando as relações entre conjuntos de variáveis (por exemplo: variáveis sociodemográficas e atitudinais) sem a necessidade de diferenciá-las previamente entre

dependentes e independentes. É possível, ainda, examinar se os efeitos dos fatores estruturantes variam conforme os subgrupos considerados (por exemplo: se a relação entre gênero e práticas de leitura varia conforme o nível de capital escolar do agente).

As críticas feitas à ACM se referem à ênfase descritiva da técnica, à impossibilidade de generalização dos resultados por não se basearem em estatísticas inferenciais e à instabilidade dos resultados. Quanto à primeira, a abordagem estatística subjacente à análise geométrica de dados coloca a descrição em primeiro lugar, o que é uma vantagem se considerarmos que só faz sentido submeter a testes de inferência descrições que se mostraram consistentes. Quanto à segunda, há ferramentas no âmbito da ACM para examinar a possibilidade de generalização dos resultados, como os testes de tipicidade, que nos permitem responder se uma determinada subnuvem de indivíduos (por exemplo: os idosos) é típica ou atípica em relação à nuvem global considerando um eixo ou um plano fatorial, e intervalos de confiança para as coordenadas das modalidades ou pontos médios das subnuvens (ver seção 5).

Quanto à terceira crítica, a estabilidade dos resultados pode ser examinada comparando soluções obtidas com conjuntos alternativos de variáveis e categorias ativas. A estabilidade depende, evidentemente, de escolhas operacionais adequadas. No caso de pesquisas que envolvam a produção de dados primários, a construção dos instrumentos de coleta de dados deve estar ajustada à escolha das técnicas de análise e, no caso do uso da ACM, o pesquisador deve incluir um número razoável de variáveis categóricas nominais (e não apenas ordinais) para mensurar as várias dimensões ou aspectos do objeto empírico, buscando um equilíbrio entre o número de categorias por variável e de variáveis por “tópico”. Há um trabalho prévio de codificação que não pode ser negligenciado pelo pesquisador. O uso de fontes secundárias implica algumas dificuldades adicionais para que se possa satisfazer os critérios da homogeneidade e exaustividade, pois a pesquisadora não tem controle sobre a construção das variáveis e das categorias disponíveis. Nuvens com formatos específicos, como o de uma ferradura ou de um triângulo, podem revelar alguns problemas na construção e escolha das variáveis: nuvens em formato de ferradura indicam uma certa unidimensionalidade dos dados, comum quando se empregam muitas variáveis ordinais; aquelas em formato de triângulo indicam que o segundo eixo diferencia os indivíduos apenas em uma das extremidades do primeiro, enquanto na outra extremidade eles se concentram. Isso sugere que as variáveis ativas captam mal as diferenças em uma parte do espaço.

Outro problema conhecido resulta da definição incorreta da lista de variáveis ativas. Quando queremos explorar as relações entre dois conjuntos diferentes de variáveis (por exemplo: práticas culturais e perfil sociodemográfico), em geral um deles é incluído como ativo e o outro como suplementar. Se ambos forem inseridos como ativos, o pesquisador inscreve a relação hipotetizada na própria construção do espaço. Isso não impede que variáveis de posição sejam ativas quando o objeto é o próprio espaço social, como no estudo de Pereira (2005) discutido adiante.

4. Exemplos de aplicação em ciências sociais

Uma das temáticas mais exploradas nas ciências sociais com o uso da ACM tem sido a das relações entre classes sociais, gostos e estilos de vida, com o objetivo de colocar à prova empírica a hipótese sobre as homologias entre o espaço social (o sistema de posições relacionalmente definidas conforme a estrutura de distribuição dos recursos socialmente eficazes em um universo social, que pode ser “recortado” para propósitos analíticos, possibilitando a delimitação de classes e frações de classe) e o espaço simbólico (o conjunto de estilos de vida, que exprimem a mesma lógica nos diferentes domínios da vida social). A ACM é uma técnica bastante útil a essa temática porque a hipótese das homologias supõe correspondências entre dois sistemas de relações, o do espaço social e o do espaço simbólico, e a ACM permite representá-los e compará-los relacionalmente.

Em “Classes e culturas de classe das famílias portuenses”, Pereira (2005) debruça-se sobre um contexto social e espacial específico, a cidade do Porto, tendo como objeto empírico as relações entre classes sociais e estilos de vida. O enfoque sobre a produção do cotidiano é fundamental para dar conta da relação entre estrutura, práticas e representações: as classes sociais condicionam os usos do tempo e do espaço por meio das representações incorporadas e se reproduzem socialmente por intermédio das práticas. O autor faz inicialmente uma reconstrução do espaço das propriedades sociais das freguesias (bairros) do Porto, por meio da ACM, com base em indicadores de renda, escolaridade, ocupação, situação do domicílio etc., construídos a partir de dados estatísticos oficiais. Interessava ao autor caracterizar a objetivação do espaço social no espaço físico da cidade, evidenciando a imbricação entre os processos histórico-sociais de formação da cidade e dos grupos sociais. Em seguida, analisando dados de um questionário estruturado aplicado a uma amostra da população portuense, Pereira reconstrói, com o uso da ACM, o espaço das classes sociais na cidade e, em combinação com uma técnica de classificação, produz uma tipologia das classes sociais. Para tanto, o autor utilizou indicadores de diversos tipos de capital (econômico, cultural e social), informações de ocupação e do tipo de habitação como variáveis ativas. Ao projetar como suplementares, no espaço assim construído, as variáveis referentes às vivências cotidianas, o estudo evidenciou a existência de padrões de classe na distribuição das práticas cotidianas. Por exemplo, a sociabilidade das classes populares mostrou-se fortemente restrita a contextos domésticos e locais (idas à igreja, saídas para compras), enquanto a sociabilidade das classes superiores incluía mais intensamente práticas de lazer cultural no contexto urbano (deslocamentos a restaurantes, teatros ou cinemas).

Em “*Upper Middle Class Social Reproduction: Wealth, Schooling, and Residential Choice in Chile*”, Méndez e Gayo (2019) tematizam a reprodução das classes médias altas em sociedades contemporâneas, delimitando o universo empírico à cidade de Santiago, no Chile. O estudo sublinha a heterogeneidade das classes médias em suas frações superiores e a complexidade do processo de reprodução dessas posições, que depende da acumulação e mobilização de recursos econômicos, culturais e sociais, por meio de estratégias residenciais e educativas. O estudo opera em dois níveis analíticos: no nível macrosocial, investigam-se as mudanças estruturais na sociedade chilena, com especial atenção às transformações na estrutura de classes; no nível mesossocial, examinam-se as práticas familiares de reprodução de classe, com base em dados primários coletados a partir de um questionário estruturado aplicado a 400 indivíduos das classes médias superiores com filhos (em idade escolar ou que já haviam cursado

a escola), residentes nas cinco regiões mais ricas da cidade, e a outras 800 pessoas, escolhidas aleatoriamente na região metropolitana de Santiago, também com filhos. O questionário buscava reunir informações sobre diversos domínios da vida social, pertinentes ao estudo da reprodução de classe: estratégias educativas (por exemplo: escolha da escola e participação na vida escolar do filho), práticas culturais, expectativas de mobilidade social, estratégias residenciais (escolha do local de moradia, mobilidade residencial) e tomadas de posição política (ideologia, ativismo). A análise desses dados com o uso da ACM possibilitou apreender as relações complexas entre mobilidade residencial (real ou aspiracional) e mobilidade social, entre estratégias residenciais (a busca dos agentes por estarem no lugar “certo” no momento “certo” de suas trajetórias) e educativas, entre expectativas de mobilidade social e estratégias familiares de transmissão de capital cultural (por exemplo: participação nas atividades escolares dos filhos), e entre orientações político-ideológicas, expectativas educativas e repertórios culturais.

Em “Classe social e alimentação: padrões de consumo alimentar no Brasil contemporâneo”, Bertonecello (2019) examina as homologias entre práticas alimentares e as posições de classe dos indivíduos. Parte-se da hipótese de que as classes moldam as práticas alimentares em função do volume e da composição dos capitais (econômico e cultural) possuídos pelos agentes, que conferem graus desiguais de distanciamento em relação às injunções estruturais, e das representações construídas sobre os alimentos e sobre a forma de ingeri-los, além de seus efeitos sobre o corpo e suas propriedades pertinentes (forma, peso, força etc.). A análise se baseou nos dados de registros de consumo alimentar de uma amostra de mais de 30 mil indivíduos coletados no âmbito da Pesquisa de Orçamentos Familiares (2008 – 2009). Com o uso da ACM, foram construídos mapas de consumo alimentar, evidenciando oposições entre o “caro”, o “leve” e o “diverso”, de um lado, e o “pesado”, o “gorduroso” e o “barato”, de outro (por exemplo: entre o grelhado e o cru, de um lado, e o frito, de outro; entre o adoçante e o açúcar). O uso combinado da ACM com a técnica de classificação hierárquica ascendente possibilitou a diferenciação de seis agrupamentos, cuja análise evidenciou diferentes padrões de consumo alimentar, suas características e distribuição social.

5. Debates contemporâneos e variações

É comum diferenciar duas variantes da ACM em função da codificação das variáveis retidas para construir os espaços, ou seja, as variáveis ativas. Na ACM padrão, todas as categorias das variáveis ativas entram na construção das nuvens. Em outros casos, algumas categorias dessas variáveis podem ser incluídas como passivas, como se fossem “desmagnetizadas”, de modo a não exercerem qualquer força de atração e repulsão sobre outras categorias. Nesse sentido, assemelham-se às categorias suplementares, pois não contam para a construção das nuvens e dos eixos. No entanto, têm funções distintas. As categorias suplementares são projetadas na nuvem de modalidades com o objetivo de contribuir para a interpretação das estruturas de afinidades e oposições. Diferentemente, as categorias passivas compõem as variáveis ativas e são incluídas dessa forma por não conterem informação pertinente ao problema (por exemplo: “não sei”), por se referirem a dado ausente

ou por serem pouco frequentes (por exemplo: menos de 5% na amostra), entre outras possibilidades. Quando as variáveis ativas usadas para construir as nuvens possuem categorias passivas, a literatura costuma empregar o termo “específica” para se referir a essa variante da ACM, ou seja, ACM específica. Embora a lógica e o procedimento de interpretação sejam os mesmos, há pequenas variações nas fórmulas para calcular os parâmetros pertinentes (número de eixos e variância da nuvem).

Os debates contemporâneos em relação à técnica incluem a possibilidade de aprofundar a análise de subgrupos, de combinar a ACM com técnicas de classificação e de análise de redes, e de compatibilizá-la com procedimentos de inferência estatística, entre outras.

Quanto à análise de subgrupos, vimos que interpretar as duas nuvens, de modalidades e de indivíduos, possibilita fazer uma análise ao mesmo tempo descritiva e explicativa. Ou seja, construindo o espaço a partir das tomadas de posição dos agentes (por exemplo: práticas de consumo musical) e sabendo como as propriedades que os caracterizam se distribuem no espaço, podemos determinar onde provavelmente estão localizados conhecendo sua escolaridade, gênero, raça etc. Ao mesmo tempo, sabendo de suas propriedades pertinentes, podemos antecipar suas prováveis tomadas de posição. A construção das elipses de concentração nos fornece indícios empíricos de como diferentes subgrupos se distribuem no espaço social e no espaço de tomadas de posição. O aprofundamento do estudo de subgrupos de indivíduos pode ser feito no âmbito da ACM por meio da análise de classe específica (ACE). Essa técnica nos permite responder se as estruturas de afinidades e contrastes observadas para a nuvem global se reproduzem internamente a subgrupos construídos a partir das categorias das variáveis suplementares. Com a ACE, podemos examinar subgrupos como subespaços dentro do espaço global, preservando sua estrutura, distâncias e proximidades. Por exemplo: podemos perguntar se as oposições em termos de práticas de consumo musical observadas para um conjunto de indivíduos são as mesmas quando consideramos os jovens e os idosos, ou quando consideramos os indivíduos de classes superiores, médias ou inferiores. Se não, como diferem e quais são as implicações analíticas? É importante notar, a partir do exemplo anterior, que fazer uma ACE do subgrupo dos mais jovens não é o mesmo que fazer uma ACM restrita a esse grupo. Na ACE, as distâncias são calculadas com base nas frequências das categorias na amostra total, de modo que o subespaço dos jovens pode ser comparado com o espaço global. Em uma ACM feita apenas com os jovens, as frequências são calculadas com base nessa amostra restrita, e o espaço resultante deixa de ser comparável ao global.

A possibilidade de combinar a análise geométrica de dados e técnicas de classificação não é uma preocupação recente. Contornar áreas de maior densidade nas nuvens agrupando indivíduos e modalidades animou os esforços iniciais dos pesquisadores preocupados em produzir classificações. Desenvolveram-se, posteriormente, métodos de classificação ajustados aos princípios da AGD, como a classificação hierárquica ascendente ou aglomerativa (CHA). Partindo das coordenadas dos indivíduos em todos os eixos da nuvem global, são feitas sucessivas agregações até que todos sejam agrupados em uma única classe. Há diversos critérios de agregação, sendo o critério da variância o que mais se harmoniza com a AGD (Le Roux & Rouanet, 2004). Por esse critério, os agrupamentos são construídos buscando minimizar a variância interna (ou a dispersão dos indivíduos). Define-se um número de agrupamentos ou *clusters* considerando a relação entre a homogeneidade interna aos

agrupamentos e a diferenciação entre eles (por meio da relação entre variância interna e variância total). Um dos resultados da CHA é um dendrograma, que ajuda a visualizar onde estão as possíveis linhas de corte. Examina-se, então, como esses grupos se distribuem no espaço e quais são as propriedades que os caracterizam. A combinação entre ACM e CHA possibilita construir agrupamentos de modo indutivo, sem impor definições *ex ante* de onde estão as possíveis fronteiras que separam os grupos. Nos estudos de classe, o uso conjunto dessas técnicas é muito útil para reconstruir as fronteiras entre classes e frações de classe.

Nas últimas décadas, buscou-se desenvolver modos de compatibilizar a análise dos laços que conectam os indivíduos, típica da análise de redes, com a análise dos espaços sociais hierarquizados onde atuam. A combinação pode partir das redes para o espaço social, reconstruindo, inicialmente, as linhas de força que estruturam um dado universo social a partir dos vínculos e interações entre os agentes socialmente situados. Ao revés, a combinação pode ir do espaço social às redes, possibilitando responder se e como as polarizações que estruturam os campos moldam os laços que conectam os agentes (Klüber, 2017).

Por fim, destaco desenvolvimentos recentes quanto aos modos de incorporar a inferência estatística no âmbito da AGD, com o objetivo de corroborar, quando possível, conclusões descritivas, evitando depender de hipóteses gratuitas ou dificilmente verificáveis. Na AGD, a descrição vem primeiro; a inferência, depois. Há diferentes testes que podem ser feitos: i) testes de tipicidade combinatória, que buscam responder se um grupo observado é típico ou atípico em relação a uma população de referência; ii) testes de tipicidade geométrica, que buscam responder se o desvio (distância) do ponto médio de uma subnuvem em relação a um ponto de referência (centro da nuvem) é real ou devido ao acaso; iii) testes de homogeneidade, que permitem responder se dois grupos podem ser tratados como homogêneos em relação a uma propriedade estudada (Le Roux et al., 2019). Em suma, a lógica da combinação dessas abordagens prevê, inicialmente, a análise descritiva dos dados utilizando as ferramentas da AGD, a identificação dos eixos, polarizações e afinidades pertinentes e, a seguir, a realização de testes combinatórios ou permutacionais para perguntar se um subgrupo é típico ou atípico da população de referência, se o desvio do ponto médio de uma subnuvem em relação a um ponto de referência é “genuíno” e/ou se os grupos podem ser tratados como equivalentes ou são heterogêneos.

6. Referências bibliográficas

Textos fundadores e genealogia

Benzécri, Jean-Paul. (1973). *L'analyse des données* (Vol. 1, La taxinomie; Vol. 2, L'analyse des correspondances). Dunod.

Le Roux, Brigitte, & Rouanet, Henry. (2004). *Geometric data analysis: From correspondence analysis to structured data analysis*. Kluwer Academic Publishers. <https://doi.org/10.1007/1-4020-2236-0>.

Manuais e guias de referência

Bertoncello, Edison. (2022). *Construindo espaços relacionais com a análise de correspondências múltiplas: Aplicações nas ciências sociais*. Enap.

Blasius, Jörg, & Greenacre, Michael (Eds.). (2014). *Visualization and verbalization of data*. CRC Press. <https://doi.org/10.1201/b16741>.

- Greenacre, Michael, & Blasius, Jörg (Eds.). (2006). *Multiple correspondence analysis and related methods*. Chapman & Hall/CRC. <https://doi.org/10.1201/9781420011319>.
- Hjellbrekke, Johs. (2019). *Multiple correspondence analysis for the social sciences*. Routledge. <https://doi.org/10.4324/9781315516257>.
- Husson, François, & Josse, Julie. (2014). Multiple correspondence analysis. In Jörg Blasius & Michael Greenacre (Eds.), *Visualization and verbalization of data* (pp. 165–183). CRC Press. <https://doi.org/10.1201/b16741-17>.
- Lebaron, Frédéric. (2006/2026). *Pesquisa quantitativa nas ciências sociais: Coleta e análise de dados* (Ana Maria Fiorini, trans.). Editora da Universidade de São Paulo.

Crítica e debates metodológicos

- Klüger, Elisa. (2017). Espaço social e redes: Contribuições metodológicas à sociologia das elites. *Tempo Social*, 29(3), 83–110. <https://doi.org/10.11606/0103-2070.ts.2017.125961>.
- Le Roux, Brigitte, Bienaise, Solène, & Durand, Jean-Luc. (2019). *Combinatorial inference in geometric data analysis*. CRC Press. <https://doi.org/10.1201/9781315155289>.
- Nicolau Netto, Michel, Michetti, Miqueli, & Bertonecelo, Edison. (2024). Desafios e potencialidades do uso de questionário estruturado em pesquisas sociológicas sobre distinção e classes sociais. *Revista Brasileira de Ciências Sociais*, 39, e39037. <https://doi.org/10.1590/39037/2024>.

Aplicações empíricas

- Bertonecelo, Edison. (2019). Classe social e alimentação: Padrões de consumo alimentar no Brasil contemporâneo. *Revista Brasileira de Ciências Sociais*, 34 (100), e3410005. <https://doi.org/10.1590/3410005/2019>.
- Méndez, María Luisa, & Gayo, Modesto. (2019). *Upper middle class social reproduction: Wealth, schooling, and residential choice in Chile*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-319-89695-3>.
- Pereira, José Virgílio Borges. (2005). *Classes e culturas de classe das famílias portuenses: Classes sociais e “modalidades de estilização da vida” na cidade do Porto*. Edições Afrontamento.

VER TAMBÉM:

[Análise geométrica de dados](#) | [Análise de correspondências simples](#) | [Análise de componentes principais](#) | [Análise de redes sociais](#) | [Classes sociais](#) | [Espaço social](#) | [Estilos de vida](#) | [Elites sociais](#)

COMO CITAR ESTE VERBETE:

Bertonecelo, Edison. (2026). Análise de correspondências múltiplas. In *Enciclopédia Brasileira de Ciências Sociais (EBCS)*. ANPOCS. 10.17666/ebcs.2026.044.