

Aplicação da Análise Multidimensional ao Mapeamento Cognitivo de Conceitos Físicos*†

CARLOS ALBERTO DOS SANTOS

Departamento de Física Teórica e Experimental, UFRN, 59000 Natal, RN

e

MARCO ANTONIO MOREIRA

Instituto de Física, UFRGS, 90000 Porto Alegre, RS

Recebido em 25 de Maio de 1979

A Multidimensional Scaling (MDS) algorithm is described and some examples are given concerning the application of this technique to similarity measures among physical concepts provided by concept association tests. The purpose of MDS is to establish an Euclidean configuration, using the method of least squares, which best fits to the experimentally obtained similarity measures. Assuming that these measures do reflect aspects of the student's cognitive structure, the Euclidean configuration given by MDS provides a "mapping" of this structure. The examples given are concerned with the concepts involved in the study of Thermodynamics at introductory college level.

É descrito um algoritmo de Análise Multidimensional (AMD) e são apresentados alguns exemplos de sua aplicação às medidas de similaridade entre conceitos físicos fornecidos por testes de associação numérica e escrita de conceitos. AMD tem como objetivo estabelecer uma

* Trabalho parcialmente financiado pela FINEP

† Versão preliminar e reduzida deste trabalho foi apresentada ao IV Simpósio Nacional de Ensino de Física, Rio, 8 a 12 de janeiro de 1978, sob a forma de comunicação intitulada "Mapeamento Cognitivo: ANÁLISE MULTIDIMENSIONAL".

configuração euclidiana, pelo método dos mínimos quadrados, que melhor se ajuste às medidas de similaridade obtidas experimentalmente. Partindo-se da suposição de que tais medidas de similaridade, ou proximidade, refletem aspectos da estrutura cognitiva do aluno, a configuração euclidiana obtida através da AMD provê um "mapeamento" dessa estrutura. Os exemplos apresentados referem-se a conceitos envolvidos no estudo de Termodinâmica ao nível de Física Geral.

1. INTRODUÇÃO

O objetivo deste trabalho é apresentar a análise multidimensional (AMD) como uma técnica potencialmente útil para o "mapeamento cognitivo" de conceitos físicos, i.e., o mapeamento da organização, ou estrutura, dada ao conjunto de conceitos básicos de um certo corpo de conhecimento por aquele que aprende.

Após uma breve descrição da AMD e de um dos algoritmos mais utilizados são apresentados exemplos do tipo de resultado ou "mapeamento" que se obtêm com a aplicação da AMD. Concluindo o trabalho são discutidas algumas possibilidades da AMD e do mapeamento cognitivo para a pesquisa em ensino de Física.

2. A ANÁLISE MULTIDIMENSIONAL

Análise multidimensional [T58], [She62a], [She62b], [K64a], [K64b], [LR77], é uma técnica estatística capaz de fornecer uma representação espacial de um conjunto de estímulos psicológicos, a partir das medidas de similaridade entre eles.

Esses estímulos psicológicos podem, em princípio, ser de qualquer natureza. Podem, por exemplo, ser conceitos de uma determinada área do conhecimento, tais como conceitos de Eletricidade e Magnetismo, Termodinâmica, Mecânica, etc.

Nesse caso, as medidas de similaridade podem ser obtidas com a aplicação de testes de associação de conceitos [SM79a].

Sob o ponto de vista de análise numérica, a AMD consiste basicamente em ajustar uma configuração de n pontos num espaço n -dimensional tal que as distâncias nesse espaço e as medidas de similaridade sejam monotonicamente relacionadas.

Apresentaremos a seguir um dos mais usados algoritmos de AMD. Uma revisão de outros algoritmos encontra-se na referência [578].

3. O ALGORITMO DE GUTTMAN E LINGOES

Para se obter a mencionada relação monotônica entre distância e similaridade, necessita-se apenas conhecer a ordenação das medidas de similaridade. Isto é, dada apenas a ordem crescente ou decrescente das similaridades, o algoritmo determina a configuração que melhor se ajusta, no sentido da monotonicidade. Quer dizer, conceitos semelhantes (muito relacionados) correspondem a pontos próximos, e conceitos diferentes (pouco relacionados) correspondem a pontos distantes.

O problema é ajustar uma configuração inicial qualquer, até que a ordenação das distâncias seja tão próxima quanto possível da ordenação das similaridades.

Intuitivamente, o problema pode ser visto assim: Inicia-se com uma configuração qualquer. Como o objetivo é estabelecer a ordenação das similaridades, tenta-se aproximar os pontos que estão "muito longe", e afastar aqueles que estão "muito próximos"; "muito longe" e "muito próximo", no sentido de violação da ordem. Assim, se a similaridade S_{12} , entre os conceitos 1 e 2 ocupa a primeira posição da ordenação das similaridades, e a distância d_{12} , entre os pontos 1 e 2 ocupa, por exemplo, a décima posição na ordenação das distâncias, então os pontos 1 e 2 estão "muito afastados" um do outro. Nesse caso o algoritmo deve ser capaz de aproximá-los.

Da mesma forma, se S_{38} ocupa a oitava posição na ordenação das similaridades, e d_{38} a primeira na ordenação das distâncias, então os pontos 3 e 8 estão "muito próximos", e o algoritmo deve afastá-los.

Desse modo, deve haver algum tipo de "interação" entre os pontos. Isto é, deve haver uma "força de repulsão" ou de "atração" entre os pontos, capaz de afastá-los ou aproximá-los, conforme o caso.

Essa idéia, desenvolvida por Shepard [She62a], foi usada por Guttman e Lingoes [LR77] para definir uma função do tipo χ^2 , que mede a qualidade do ajuste.

$$\phi^* = \sum_{i,j} (d_{ij} - d_{ij}^*)^2, \quad (1)$$

onde

$$d_{ij} = \left[\sum_{a=1}^n (x_{ia} - x_{ja})^2 \right]^{1/2}. \quad (2)$$

e d_{ij}^* é a "imagem em ordenação" (rank image) de d_{ij} , satisfazendo às seguintes propriedades

$$\sum_{i,j} d_{ij}^* = \sum_{i,j} d_{ij} \quad (3)$$

$$\sum_{i,j} d_{ij}^{*2} = \sum_{i,j} d_{ij}^2 \quad (4)$$

$$d_{ij}^* < d_{kl}^*, \quad \text{se } S_{ij} < S_{kl} \quad (5)$$

os vetores D^* e D são relacionados através da transformação T^*

$$D^* = T^*D, \quad (6)$$

onde

$$T^* = E(E^*)^1, \quad (7)$$

onde o sobrescrito 1 representa a operação de transposição.

As matrizes E e E^* , de ordem $m = n(n-1)/2$ (para n conceitos), são tais que

$$e_{ijk} = \begin{cases} 1, & \text{se } S_{ij} \text{ ocupa a posição } k \text{ em } S. \\ 0, & \text{caso contrário} \end{cases}$$

$$e_{ijk}^* = \begin{cases} 1, & \text{se } d_{ij} \text{ ocupa a posição } k \text{ em } D. \\ 0, & \text{caso contrário} \end{cases}$$

Então, dadas as similaridades S_{ij} , o problema resume-se em minimizar a função ϕ^* .

O processo de minimização desenvolve-se através de dois algoritmos, denominados por Guttman e Lingoes de algoritmo de uma e duas fases.

O algoritmo de duas fases consiste em minimizar a função

$$Q^* = \sum_{i,j} (d_{ij} - \delta_{ij})^2, \quad (\text{primeira fase})$$

onde δ_{ij} é um elemento da matriz

$$\Delta(s+1) = T^*D(s) \quad (\text{segunda fase})$$

Quer dizer, a matriz Δ na iteração $(s+1)$ é determinada a partir da matriz D da iteração s .

A convergência do algoritmo de duas fases tem como resultado fundamental a obtenção de T^* em sua forma quase final, de modo que o algoritmo de uma fase se constitui num ajuste fino do processo total.

Esse ajuste fino consiste em minimizar ϕ^* , onde D e D^* são variáveis.

Fazendo

$$\frac{\partial \phi^*}{\partial x_{ia}} = 0,$$

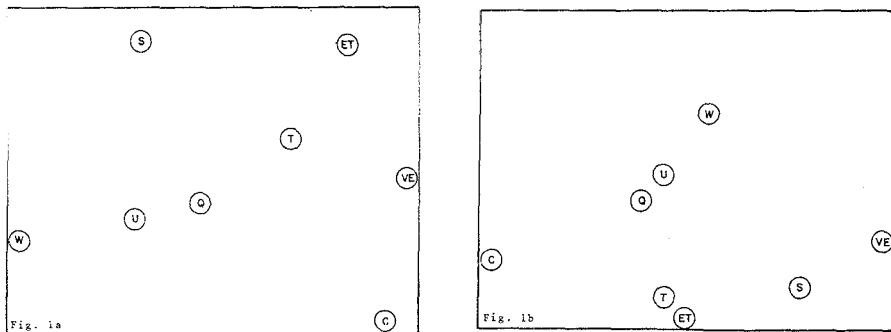
mostra-se [S78] que a configuração é incrementada através da relação

$$x_{ia}^{(s+1)} = x_{ia}^{(s)} - \frac{1}{4n} \left[\frac{\partial \phi^*}{\partial x_{ia}} \right]^{(s)}$$

4. ALGUNS EXEMPLOS

A seguir, são apresentados alguns exemplos da aplicação da AMD ao mapeamento cognitivo de conceitos de Termodinâmica. Tais exemplos foram tirados de um experimento realizado por Santos [S78] com estudantes de Física Geral do Instituto de Física da UFRGS, no primeiro semestre de 1978. O programa de Guttman e Lingoês descrito na referência [L73] foi o utilizado.

Nas figuras 1a e 1b são apresentados os mapeamentos cognitivos obtidos para um grupo de alunos antes e após a instrução, i.e., antes e após terem estudado o conteúdo de Termodinâmica. Nessas figuras



Figuras 1a e 1b - Mapeamento cognitivo obtido através da aplicação da AMD às medidas de similaridade fornecidas pelo teste de associação numérica de conceitos; a) antes da instrução; b) após a instrução.

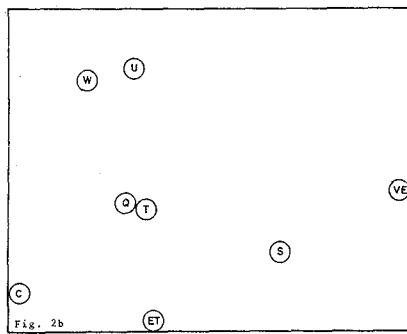
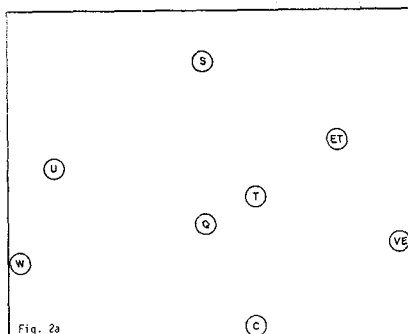
os conceitos estão representados pelos respectivos símbolos: temperatura (T), calor (Q), trabalho (W), energia interna (U), equilíbrio térmico (ET), calor específico (C), variável de estado (VE), entropia (S).

Os dados que deram origem a esses mapeamentos foram obtidos através de um teste de associação numérica de conceitos [SM79a]. Nesse tipo de teste, para um dado conjunto de conceitos são formados todos os pares possíveis, os quais são, então, listados em ordem aleatória ao lado de uma escala numérica na qual deve ser assinalado um número que reflita o grau de relacionamento entre os conceitos de cada par segundo a percepção de quem responde o teste. Obtém-se assim uma matriz de similaridades a qual é, então, aplicada a AMD.

Comparando as figuras 1a e 1b observa-se uma sensível modificação na estrutura ao final da instrução. Na figura 1a os conceitos estão espalhados, não havendo qualquer agrupamento predominante. Ao contrário, a figura 1b apresenta três grupos distintos. Um com calor (Q), energia (U) e trabalho (W), outro com temperatura (T) e equilíbrio térmico (ET), e um terceiro com entropia (S) e variável de estado (VE). A figura 1b mostra que ao final da instrução os conceitos foram associados de forma a apresentar uma estrutura lógica e coerente com o conteúdo da Termodinâmica, visto que os conceitos envolvidos nas três leis da Termodinâmica (incluindo a "lei zero") formaram grupos distintos.

Nas figuras 2a e 2b são apresentados os mapeamentos cognitivos para outro grupo de alunos com outro tipo de teste de associação de conceitos, o teste de associação escrita de conceitos. Nesse tipo de teste [SM79a] as medidas de similaridade são obtidas indiretamente: a cada conceito de um dado conjunto são associados tantos outros conceitos ou palavras quanto possível num certo tempo; a partir das listas de palavras associadas a cada conceito do conjunto determina-se um coeficiente de relacionamento para cada par de conceitos possível. Obtém-se assim uma matriz de coeficientes, i.e., uma matriz de similaridades cuja estrutura, se houver, pode ser analisada através da AMD.

Comparando-se as figuras 2a e 2b pode-se também dizer que a instrução produziu algum efeito, pois após a instrução certos conceitos estão mais próximos como é o caso de W e U , e Q e T . Em termos de orga-



Figuras 2a e 2b - Mapeamento cognitivo obtido através de aplicação da AMD às medidas de similaridade fornecidas pelo teste de associação escrita de conceitos; a) antes da instrução; b) após a instrução.

nização conceitual, porém, a alteração não foi relevante: se girarmos um pouco a figura 2b (após a instrução) no sentido anti-horário obter-se-á quase a mesma configuração da figura 2a (antes da instrução). Cabe, no entanto, repetir que este exemplo refere-se a outro grupo de alunos que estudou o mesmo conteúdo sob outras condições e respondeu um outro tipo de teste.

Cabe também registrar que nem sempre obtêm-se uma configuração bidimensional como as apresentadas nas figuras 1a, 1b, 2a e 2b. A AMD tem por objetivo estabelecer uma configuração espacial que melhor se ajuste às medidas de similaridade obtidas experimentalmente e pode ocorrer que esta configuração tenha três ou mais dimensões. Obviamente, a análise dos resultados torna-se, nesse caso, ainda mais difícil, principalmente no que concerne à interpretação do significado das dimensões e à visualização da configuração obtida.

Outro problema é que mesmo para uma matriz aleatória sempre se obtém uma configuração. Torna-se então necessário o estabelecimento de critérios de significância estatística a fim de se ter segurança, dentro de certos limites, de que a configuração obtida não provém de uma matriz aleatória [S78].

5. CONCLUSÃO

A AMD foi sugerida neste trabalho como uma técnica potencialmente útil para o mapeamento cognitivo de conceitos físicos. (Em outro trabalho, SM79b, os autores sugerem a análise de agrupamentos hierárquicos como uma técnica alternativa para o mapeamento cognitivo e em um terceiro, SM79a, discutem a idéia de mapeamento em si e apresentam alguns instrumentos de medida.) Foram apresentados mapeamentos cognitivos obtidos para grupos de alunos, porém, é claro que a técnica se aplica também a casos individuais. Isto é, pode-se obter o mapeamento cognitivo de cada aluno antes, durante e após a instrução. Obtém-se assim um outro tipo de informação acerca da aprendizagem do aluno. Por exemplo, como ele relaciona, diferencia e estrutura os conceitos de um certo corpo de conhecimento.

Obviamente, considerando o estágio relativamente recente e pouco desenvolvido da pesquisa nessa área (mapeamento cognitivo) os "mapas cognitivos" devem ser utilizados com cautela. Entretanto, no caso de "mapas" individuais ter-se-ia, pelo menos, um instrumento diagnóstico que poderia ser discutido com o aluno. Talvez muitas das dificuldades em "aprender Física" estejam relacionadas com a dificuldade em adquirir uma estrutura conceitual adequada. Se assim for, o mapeamento cognitivo poderia ser um instrumento útil para detectar falhas nessa área.

No caso de "mapas" para grupos, os mesmos podem fornecer informações acerca da influência do tipo de instrução na "estrutura cognitiva do grupo". Por exemplo, se antes e após a instrução o mapeamento cognitivo do grupo permanece o mesmo, isso pode significar alguma falha no processo instrucional. Podem também ser utilizados como instrumentos de medida para comparar o efeito de diferentes organização sequenciais de um mesmo conteúdo sobre a estrutura cognitiva do aluno.

Estas são algumas das possibilidades de aplicação da AMD e do mapeamento cognitivo à pesquisa em ensino de Física, mas somente a efetiva realização de experimentos com essas técnicas é que determinará sua validade e utilidade.

REFERÊNCIAS

- K64a Kruskal, J.B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, Chicago, 29 (1) 1-27, Mar. 1964.
- K64b Kruskal, J.B. Nonmetric multidimensional scaling; a numerical method. *Psychometrika*, Chicago, 29(2):115-29, June, 1964.
- L73 Lingoes, J.C. *The Guttman-Lingoes nonmetric program series*. Michigan, Mathesis, 1973. p.39-79.
- LR77 Lingoes, J.C. & Roskam, E.E. A mathematical and empirical analysis of two multidimensional scaling algorithms. In: ____ *Geometric representations of relational data*. Michigan, Mathesis, 1977. p.371-3.

- S78 Santos, C.A. Aplicação da análise multidimensional e da análise de agrupamentos hierárquicos ao mapeamento cognitivo de conceitos físicos. Porto Alegre, Instituto de Física da UFRGS, 1978. Dissertação de Mestrado.
- She62a Shepard, R.N. The analysis of proximities; multidimensional scaling with an unknown distance function. I. Psychometrika. Chicago, 27(2):125-40, June, 1962.
- She62b Shepard, R.N. II. Psychometrika, Chicago, 27(3):219-46, Sept. 1962.
- SM79a Santos, C.A. & Moreira, M.A. *Mapeamento Cognitivo: INSTRUMENTOS DE MEDIDA*. Comunicação apresentada ao IV Simpósio Nacional de Ensino de Física, Rio de Janeiro, 08-12 de janeiro de 1979.
- SM79b Santos, C.A. & Moreira, M.A. *Mapeamento Cognitivo: ANÁLISE DE AGRUPAMENTOS HIERÁRQUICOS*. Comunicação apresentada no IV Simpósio Nacional de Ensino de Física, Rio de Janeiro, 08-12 de janeiro de 1979.
- T58 Torgerson, W.S. *Theory and methods of scaling*. New York, John Wiley, 1958. cap.II.