

Todos os itens da prova têm o mesmo valor: 1.0 ponto cada (total de 10 pontos). Tempo de prova: 2 horas.

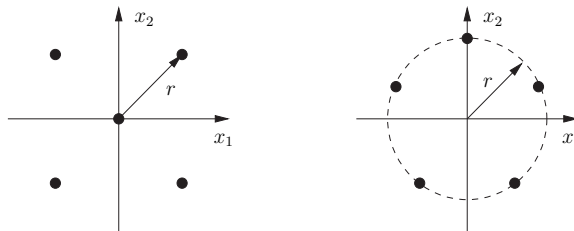
1. (*Algoritmo de Metropolis*) Nesta questão, consideramos o problema da descrição de todas as configurações possíveis de um sistema com 5 partículas em duas dimensões. A posição de cada partícula é definida por um vetor \mathbf{x}_i , $i = 1, 2, \dots, 5$ e o estado do sistema é definido por um vetor \mathbf{x} contendo todas as 10 coordenadas. Duas configurações particulares são mostradas na figura do item (b). A função custo na qual estamos interessados é uma combinação linear entre a soma das normas dos vetores \mathbf{x}_i e a soma das “repulsões eletrostáticas” entre as partículas, imaginando o caso em que todas são positivas:

$$J(\mathbf{x}) = \sum_i \|\mathbf{x}_i\|^2 + \frac{1}{\sum_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|^2} \quad (1)$$

- a) Escreva, utilizando pseudo-código, uma implementação do algoritmo de Metropolis que, para temperatura T e a partir de uma configuração inicial qualquer, permita a geração de estados seguindo uma distribuição de Boltzmann em função dos seus custos J .
- b) Para as duas soluções locais a seguir, uma expressão simplificada para J pode ser calculada em função da variável escalar positiva r :

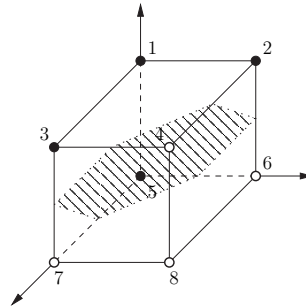
$$J_1(r) = 4r^2 + \frac{6.5}{r^2} \quad \text{e} \quad J_2(r) = 5r^2 + \frac{5}{r^2} \quad (2)$$

Assumindo $T = 0.1$, calcule a proporção entre as probabilidades de um estado que tem a configuração da direita com $r = 1.0000$ e outro estado que tem a configuração da esquerda com $r = 1.1291$.



- c) Considere a definição da variável aleatória “distância média à origem”: $L(\mathbf{x}) = (1/5) \sum_i \|\mathbf{x}_i\|$. Explique como o algoritmo do item (a) é modificado, de forma que possamos calcular o valor médio de L a uma temperatura T arbitrária.
- d) (0.5 ponto extra) À medida em que T diminui, o que acontece com o valor médio de L ?
2. (*Simulated Annealing*) Considere a função custo $J(x) = x^2$, avaliada sobre os estados $x = -2, -1, 0, +1$ e $+2$. Considere a minimização desta função utilizando o algoritmo SA básico.
- a) Assumindo temperatura fixa T , calcule a matriz de transição \mathbf{M} associada ao processo de Markov que se obtém com a execução do algoritmo de Metropolis para esta função custo.
- b) Mostre que o vetor contendo os fatores de Boltzmann associados a cada um dos estados é igual a um autovetor da matriz \mathbf{M} e que o autovalor associado a este autovetor é igual a 1.0.

3. (*Simulated Annealing*) Considere uma situação em que gostaríamos de dividir os oito vértices de um cubo unitário em dois agrupamentos, de modo que o erro quadrático total entre os centros dos agrupamentos e os membros dos agrupamentos seja minimizado. Considere que os vértices são numerados da seguinte forma:



e também que os agrupamentos são definidos através de um vetor \mathbf{x} binário de comprimento oito. A i -ésima componente de \mathbf{x} é igual a zero se o vértice i pertencer ao agrupamento zero. E é igual a um no caso contrário. Por exemplo, o estado $\mathbf{x} = 11101000$ corresponde ao caso da figura, em que os centróides são $(3/4, 3/4, 1/4)$ (agrupamento 0) e $(1/4, 1/4, 3/4)$ (agrupamento 1) e o custo $J(\mathbf{x})$ é 4.50.

- Baseando-se em um esquema de perturbação que consiste em sortear uma posição do vetor \mathbf{x} e invertê-la, escreva um algoritmo SA básico para a minimização do erro quadrático total. Defina quaisquer parâmetros que você julgar necessários.
- Assumindo $T = 1.0$ e o mesmo esquema de perturbação do item (a), calcule a probabilidade de transição do estado 11101000 para o estado 11100000, e também a probabilidade de transição do estado 11100000 para o estado 11110000.
- A contagem dos estados conforme os seus custos é dada pela tabela a seguir:

Custo	4.00	4.50	4.53	4.67	5.00	5.14	5.33	5.50	5.60	6.00
Número de Estados	6	8	48	24	24	16	24	24	64	18

Calcule qual é a probabilidade de $\mathbf{x} = 00001111$ ser gerado, quando o SA básico atinge $T = 0.1$.

4. (*Deterministic Annealing*) Considere um problema de *soft clustering* com cinco vetores de dados $\mathbf{x} = (x_1, x_2)$ e três centróides iniciais $\mathbf{y} = (y_1, y_2)$, definidos segundo a tabela a seguir. Considere também que a distância entre dois vetores é quadrática, ou seja, $d(\mathbf{x}, \mathbf{y}) = (x_1 - y_1)^2 + (x_2 - y_2)^2$.

\mathbf{x}_1	5	4
\mathbf{x}_2	4	5
\mathbf{x}_3	5	5
\mathbf{x}_4	-5	-4
\mathbf{x}_5	-4	-5
\mathbf{y}_1	0	0
\mathbf{y}_2	1	1
\mathbf{y}_3	-1	-1

- Calcule a matriz de probabilidades $p(\mathbf{y}|\mathbf{x})$ que minimiza $J = D - TH$ com $T = 10$. Calcule também os valores dos centróides atualizados segundo esta matriz.
- Repita o item (a) para $T = 0.1$ e comente sobre qual é a diferença.