UFRJ / COPPE / Programa de Engenharia Elétrica – Primeiro Período de 2019
CPE-723 – Otimização Natural (Parte I - *Simulated Annealing*)
Prova Parcial #1 – 11 de abril de 2019

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*) Resolva os três problemas a seguir:

   a) Para a avaliação da integral $\int_0^\infty xe^{-2x}dx$, duas pessoas procederam de duas formas diferentes. Uma utilizou um gerador de números aleatórios com distribuição exponencial e variância igual a 1.0 para gerar um grande número de amostras x, depois calculou valores xe^{-x} usando as amostras e fez a média destes valores. A outra pessoa utilizou um gerador de números aleatórios com distribuição exponencial e variância igual a 0.5 para gerar um grande número de amostras x, depois calculou valores $x/2$ usando as amostras e fez a média destes valores. Ambas as pessoas obtêm resultados corretos? Comente.

   b) Escreva um pseudo-código para a avaliação da integral $\int_0^3 f(x)e^{-x}dx$ a partir do algoritmo de Metropolis, usando função energia $J(x) = x$.

   c) Escreva um pseudo-código para a avaliação da integral de $\prod_{i=1}^N x_i e^{-x_i}$ dentro do hipercubo $[0,1]^N$.

2. (*Simulated Annealing*) Considere a função-custo $J(x_1, x_2)$ definida pela tabela a seguir, onde $a > b > c > d$:

x_1	x_2	$J(x_1, x_2)$
0	0	a
0	1	b
1	1	c
1	0	d

   a) A aplicação do algoritmo de Metropolis a um vetor inicial $\mathbf{x}_0$ qualquer, perturbando-se uma componente (x_1 ou x_2) por vez, define um processo de Markov com duas matrizes de transição, $\mathbf{M}_1$ e $\mathbf{M}_2$, ambas com dimensão 4×4. Calcule essas matrizes, considerando $T = 1$.

   b) Calcule, também para temperatura $T = 1$, a distribuição de Boltzmann/Gibbs do vetor aleatório (X_1, X_2). Verifique que esta distribuição de probabilidades corresponde a um vetor invariante para ambas as matrizes de transição calculadas no item (a).

   c) Utilizando pseudo-código, descreva um algoritmo de *simulated annealing* para minimizar funções-custo semelhantes à do enunciado, ou seja, funções $J(x_1, x_2, \ldots, x_N)$, onde x_i, com $i = 1, 2, \ldots, N$, são variáveis discretas (pode-se assumir que são binárias). Os valores assumidos por $J(x_1, x_2, \ldots, x_N)$ são arbitrários.

   d) Começando em $\mathbf{x}_0 = (0, 0, \ldots, 0)$, e usando quaisquer números aleatórios dos quais você precise (escolha livremente), execute manualmente as duas primeiras iterações do algoritmo do item (c).

   e) Quando um número suficientemente grande de iterações do algoritmo do item (c) tiver sido executado à temperatura $T = 0.1$, qual é a probabilidade do evento $J = d$?

3. (*Deterministic Annealing*) Considere a separação de dados escalares $x_1 = 3$, $x_2 = 1$, $x_3 = -1$ e $x_4 = -3$ em dois agrupamentos (classes, ou *clusters*) com centróides y_1 e y_2.

   a) Considerando $T = 5$ e centróides inicialmente localizados em $y_1 = 2$ e $y_2 = -2$, calcule a partição $p(y|x)$ e, a partir dela, as posições atualizadas dos centróides y_1 e y_2.

   b) Repita o item (a), considerando $T = 0.5$ e centróides inicialmente localizados em $y_1 = 0.5$ e $y_2 = -0.5$.

Boa prova!