Modelagem Matemático-computacional
Aula 6

Nesta aula, apresentaremos um método de otimização baseado em teoria da evolução biológica, conhecido como algoritmo genético.

I. MOTIVAÇÃO

Uma das formas de se otimizar um problema que tem propriedades X_i e parâmetros p_i consiste em definir uma função $\mu = \mu(\vec{X}, \vec{p})$, através da qual é possível encontrar um conjunto de X_i e p_i tal que μ seja extrema (máxima ou mínima). Um problema tradicional que ilustra esta ideia é conhecido como *Caixeiro viajante*. Imagine um caixeiro que deve distribuir produtos desde a central onde trabalha até N cidades (veja a figura 1). Considerando que a central está na posição $(0, 0)$ e que as distâncias entre duas cidades são l_{ij}, existem diversos possíveis caminhos que o caixeiro pode fazer. Para economizar no trajeto entre as sucessivas cidades, gostaríamos de traçar uma rota ideal para a qual o caixeiro percorre a menor distância possível, desde a saída até à chegada na central. Desse modo, o problema do caixeiro reduziu-se a encontrar uma configuração de l_{ij} e X_i onde a distância total percorrida seja mínima. Então definimos a função mérito $\mu = \sum_{i,j \neq i} l_{ij}$ como sendo a soma de todas as distâncias entre as cidades. A representação de um trajeto particular seria um vetor como $\{|\vec{l}_{2,0}|, |\vec{l}_{5,2}|, |\vec{l}_{10,5}|, ...|\vec{l}_{0,3}|\}$, de modo que qualquer um dos demais possíveis trajetos seriam permutações dos vetores l_{ij}. Um destes trajetos está ilustrado na figura 2.

Existem diversos problemas onde este procedimento de otimização é requerido. Em muitos deles, é necessário considerar-se todas as possíveis combinações de configurações possíveis, de modo que a otimização acaba sendo custosa quando o número $N \to \infty$.

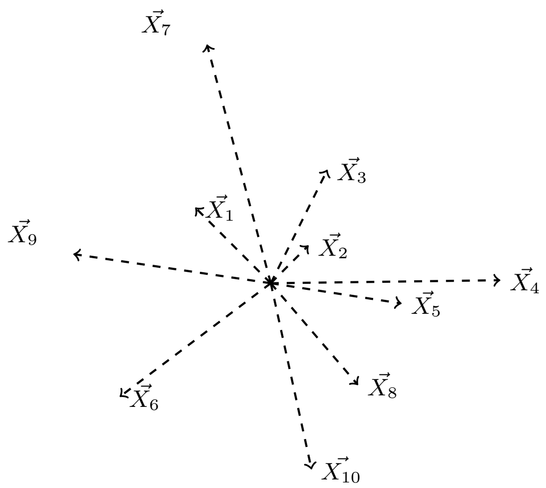

Figura 1: Posições das cidades em relação à central de distribuição do caixeiro.

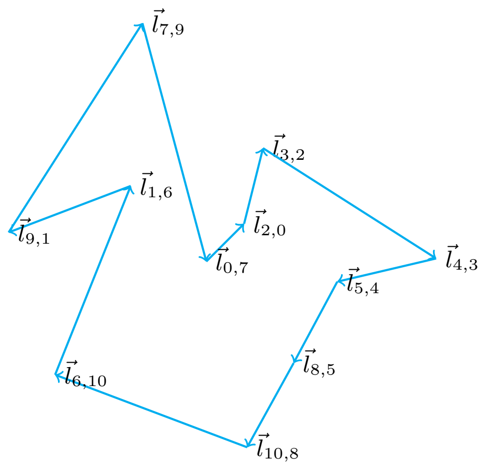

Figura 2: Um possível caminho que o caixeiro poderia fazer saindo e retornando à central e visitando todas as cidades.

II. ALGORITMO GENÉTICO

No *algoritmo genético* a idéia de otimização é utilizada para se reproduzir a evolução natural de organismos vivos. Podemos considerar que o processo de evolução (tanto do ser humano quanto de outras espécies) é dado pelas leis darwinianas: filhos herdam misturas de DNA dos pais e o ambiente atua sobre estes indivíduos, de modo que somente os mais aptos sobreviverão e darão origem a novos indivíduos. Uma vez que variações progressivas do ambiente poderiam levar espécies à extinção, a variabilidade é um mecanismo que assegura a sobrevivência.

Sabemos que os organismos são constituídos por células cujo núcleo contém os genes herdados dos pais. Em princípio, um *gene* é um trecho de DNA que codifica uma certa proteína. O conjunto de genes compõem o genoma e definem as características do indivíduo. Desse modo, podemos associar a um dado indivíduo um mérito para sobreviver no ambiente, conforme um dado conjunto de genes.

O conjunto de passos do algoritmo genético que vamos considerar é ilustrado abaixo:

1. Representação do problema
   Consideramos que os indivíduos são representados

por vetores de parâmetros ao qual podemos associar uma função mérito. Exemplos são as distâncias às cidades no caso do caixeiro ou uma sequência de genes do DNA.

2. 1ª geração
   Inicia-se o conjunto evolutivo com uma população aleatória de N indivíduos. É preciso considerar que a forma como a primeira geração será implementada depende do problema. No caso do melhor caminho do caixeiro, não é permitido visitar duas vezes a mesma cidade. Para o DNA, dependendo de como um gene será implementado, ele poderá ter componentes iguais. Associa-se a cada indivíduo uma função mérito.

3. Nova geração
   Uma nova geração é obtida através de mutações, inversões, combinações (crossing over) das propriedades dos indivíduos (genes, no caso do DNA, ou distâncias às cidades para o caixeiro). Ordena-se a população por mérito, escolhendo-se uma data percentagem de indivíduos mais eficientes (por exemplo 30%). Mutações são implementadas escolhendo-se aleatoriamente duas posições e trocando suas componentes. Ex: $\vec{v} = [3, 10, 2, 1, ...] \rightarrow [3, 1, 2, 10, ...]$. Inversões consistem na troca de um trecho pelo seu inverso. Ex: $\vec{v} = [3, 10, 2, 1, 8, 7...] \rightarrow [3, 7, 8, 1, 2, 10...]$. Combinações consistem na troca de trechos ou componentes de dois indivíduos. Ex: $\vec{v_1} = [3, 10, 2, 1, 8, 7...]$ e $\vec{v_2} = [5, 1, 2, 10, 9, 4...]$, levando a um indivíduo $\vec{v_3} = [3, 10, 2, 10, 9, 4, ...]$. Note que neste último exemplo, a troca levaria à repetições de cidades para o caixeiro viajante, o que não pode acontecer (o caixeiro deve passar por cada cidade uma única vez).

4. Repetir sucessivamente o último passo, fazendo-se um gráfico do mérito em função da geração. O critério de parada será definido quando a maximização ou minimização é alcançada em uma geração (assintoticamente).

Seguindo a idéia descrita nos passos anteriores, pode-se calibrar o algoritmo genético para o problema de interesse. Uma forma de se fazer isto consiste em especificar uma situação particular cuja solução seja previamente conhecida (ou intuitiva). Por exemplo, no caso do caixeiro viajante, se as distâncias entre todas as cidades forem iguais, a solução ótima é um círculo

Referências

[1] ...

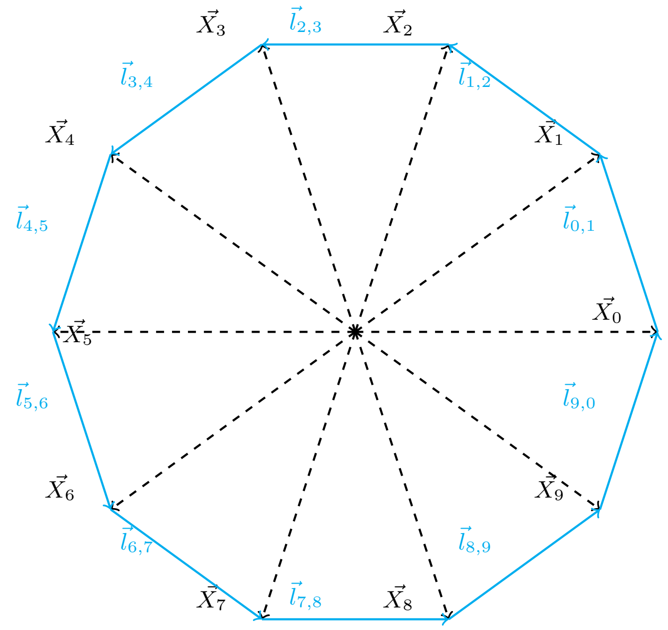

Figura 3: Solução ótima para o problema do caixeiro viajante quando todas as cidades estão igualmente espaçadas.