Otimização Combinatória - INF 2912
Prof. Marcus Vinicius Soledade Poggi de Aragão

Ciro Cavani
BigData / Globo.com

Learning in Combinatorial Optimization: What and How to Explore

Sajad Modaresi, University of Pittsburgh
Denis Saure, University of Pittsburgh
Juan Pablo Vielma, MIT Sloan School of Management

June 20, 2013

Seminário, 25/11/2015

Agenda

1. Visão Geral
2. Motivação
3. Formulação
4. Política Simples
5. Política Adaptável
6. Implementação

Visão Geral

O *paper* apresenta um algoritmo para aprendizado em tempo real de soluções para problemas combinatoriais com garantia teórica de performance.

- Sequência de decisões em um modelo com incerteza
- Otimização de um problema combinatorial em que a função objetivo muda a cada período
- Implementação de determinada solução implica em ganho de conhecimento e custo

Esse trabalho mostra como equilibrar entre a exploração do desconhecido e o aproveitamento das melhores soluções conhecidas (*exploration and exploitation*).

Motivação

Problemas em que uma solução off-line não são possíveis, por exemplo, falta dados históricos ou eles são parciais.

Aplicações de Rede

O uso constante do caminho mais curto revela apenas o custo associado aos arcos desse caminho e não outros arcos do grafo. Explorando os custos associados a esses arcos desconhecidos é possível resolver um problema combinatório cuja solução tem um custo menor do que a conhecida inicialmente.

O mesmo princípio pode ser usado em Planejamento de Oferta e Demanda, Testes Clínicos com combinação de medicamentos.

Formulação

$n = 1$ corresponds to the first instance, and $n = N$ to the last full instance is defined as $f(B_n)$, where

$$f(B) : z^*(B) := \min \left\{ \sum_{a \in S} b_a : S \in \mathcal{S} \right\} \quad B \in \mathbb{R}^{|A|}$$

$B_n := (b_{a,n} : a \in A) \in \mathbb{R}^{|A|}$

b_a is the *cost* associated with a ground element $a \in A$

$b_{a,n} \in B_n$ is a random variable

$l_a \leq b_{a,n} \leq u_a$ a.s., for all $a \in A$ and $n \in \mathbb{N}$.

$F(\cdot)$ common distribution of B_n for $n \in \mathbb{N}$ initially *unknown*.

$\mathcal{S}$ is a family of subsets of elements of a given ground set A

let $\mathcal{S}^*(B)$ be the set of optimal solutions

$z^*(B)$ be its optimal objective value (cost)

$$\mathcal{F}_n = \sigma(\{b_{a,m} : a \in S_m\,,\, m < n\})$$

$\pi := (S_n)_{n=1}^{\infty}$ denote an admissible policy

expectedcumulative cost incurred by policy π is given by

$$J^\pi(F,N) := \sum_{n=1}^{N} \mathbb{E}_F \left\{ \sum_{a \in S_n} b_{a,n} \right\}$$

regret is defined as

$$R^\pi(F,N) := J^\pi(F,N) - N\ z^*\left(\mathbb{E}_F\left\{B_n\right\}\right)$$

e expected optimality gap associated with implementing S.

$$\Delta_S^F := \sum_{a \in S} \mathbb{E}_F\left\{b_{a,n}\right\} - z^*\left(\mathbb{E}_F\left\{B_n\right\}\right)$$

number of times that the agent has implemented solution $S_m = S$ prior to instance n

$$T_n(S) := |\{m < n \, : \, S_m = S\}|$$

number of times that the agent has selected element a prior to instance n

$$T_n(a) := |\{m < n \, : \, a \in S_m\}|$$

Using this notation we have that

$$R^\pi(F, N) = \sum_{S \in \mathcal{S}} \Delta_S^F \, \mathbb{E}_F \left\{ T_{N+1}(S) \right\}$$

Política Simples: $\pi_s(\mathcal{E})$

$\overline{B}_n := (\bar{b}_{a,n}, a \in A)$

$\bar{b}_{a,n} := \frac{1}{T_n(a)} \sum_{m<n\,:\,a \in S_m} b_{a,m}$

$\mathcal{E}$ be a cover of A

$n_i := \max\left\{\lfloor e^{i/H} \rfloor, n_{i-1} + 1\right\}$

$\Phi := \{n_i : i \in \mathbb{N}\}$

Algorithm 1 Simple policy $\pi_s(\mathcal{E})$

Set $i = 0$, and $\mathcal{E}$ a minimal cover of A
for $n = 1$ to N **do**
  if $n \in \Phi$ **then**
    Set $i = i + 1$
    Set $S^* \in \mathcal{S}^*\left(\overline{B}_n\right)$
  end if
  if $T_n(a) < i$ for some $a \in S$, for some solution $S \in \mathcal{E}$ **then**
    Implement such a solution, i.e., set $S_n = S$
  else
    Implement $S_n = S^*$
  end if
end for

Theorem 4.3. *There exists a positive finite constant* H_s *such that for any* $H \geq H_s$, $N > 0$, *and cover* $\mathcal{E}$, *the regret of* $\pi_s(\mathcal{E})$ *admits the following bound*

$$\frac{R^{\pi_s(\mathcal{E})}(F,N)}{\ln N} \leq \min\left\{|\mathcal{E}|, |A|\right\} \Delta^F_{max} H + o(1),$$

where $\Delta^F_{max} := \max\left\{\Delta^F_S \,:\, S \in \mathcal{S}\right\}$.

Let $\mathcal{D}$ contain all subsets D of suboptimal ground elements
such that they become part of every optimal solution if their costs are the lowest possible.

$$\mathcal{D} := \left\{ D \subseteq A \setminus S^*\left(\mathbb{E}_F\left\{B_n\right\}\right) : D \subseteq \bigcap_{S \in \mathcal{S}^*(B'_D)} S \right\} \; B'_D := (b'_a : a \in A) \;\; b'_a := \begin{cases} l_a & a \in D \\ \mathbb{E}_F\left\{b_{a,n}\right\} & a \notin D \end{cases}$$

Proposition 5.2. *For any consistent policy π, regular F, and $D \in \mathcal{D}$ we have that*

$$\lim_{N \to \infty} \mathbb{P}_F \left\{ \frac{\max\left\{T_{N+1}(a) : a \in D\right\}}{\ln N} \geq K_D \right\} = 1,$$

where K_D is a positive finite constant depending on F.

Proposition 5.2 gives an answer to *what* needs to be explored: a critical subset.

Critical Subsets.

$\mathcal{C} := \{C \subseteq A \; : \; \forall D \in \mathcal{D}, \exists \; a \in C \text{ s.t. } a \in D\}$

Integer Lower Bound Problem (ILBP).

$$
\begin{aligned}
ILBP: \quad z(F,N) := \min \quad & \sum_{S \in \mathcal{S}} \Delta_S^F \, y_S && \text{(10a)} \\
\text{s.t.} \quad & \max\{x_a : a \in D\} \geq K_D \ln N, \quad D \in \mathcal{D} && \text{(10b)} \\
& x_a \leq \sum_{S \in \mathcal{S} : a \in S} y_S, \quad a \in A && \text{(10c)} \\
& x_a,\, y_S \in \mathbb{N}_0, \quad a \in A, S \in \mathcal{S}. && \text{(10d)}
\end{aligned}
$$

In this formulation, y_S and x_a are meant to represent $T_{N+1}(S)$ and $T_{N+1}(a)$, respectively. Here, the a's with non-zero x_a correspond to the critical subset and the S's with non-zero y_S correspond to the cover of the critical subset. Indeed, constraints (10b) enforce exploration conditions (9) on the critical subset and constraints (10c) enforce the cover of the critical subset.

Intuitively, we expect the regret of any consistent policy to grow at least as fast as $z(F, N)$. To transform this intuition into a fundamental performance limit, we consider the following Lower Bound Problem (LBP), which characterizes the rate of growth of $z(F, N)/\ln N$.

$$
\begin{aligned}
LBP: \quad \kappa(F) := \min \quad & \sum_{S \in \mathcal{S}} \Delta_S^F \, y_S && \text{(11a)} \\
s.t. \quad & \max\{x_a : a \in D\} \geq K_D, \quad D \in \mathcal{D} && \text{(11b)} \\
& x_a \leq \sum_{S \in \mathcal{S}: a \in S} y_S, \quad a \in A && \text{(11c)} \\
& x_a,\, y_S \in \mathbb{R}_+, \quad a \in A, S \in \mathcal{S}. && \text{(11d)}
\end{aligned}
$$

Note that $\kappa(F)$ does not change if we refine $\mathcal{D}$ in (11b) to include only subsets that are minimal with respect to inclusion. One can show that

$$\lim_{N \to \infty} \frac{z(F, N)}{\ln N} = \kappa(F).$$

Theorem 5.4. *The regret of any consistent policy π is such that for any regular F we have*

$$\liminf_{N \to \infty} \frac{R^{\pi}(F,N)}{\ln N} \geq \kappa(F),$$

where $\kappa(F)$ is the optimal objective value of formulation LBP in (11).

Consider a special case of formulation (11) where $K_D = H$ for all $D \in \mathcal{D}$: one can show that $\kappa(F)$ is homogeneous in H. Thus, without loss of generality, one can take $H = 1$ and interpret LBP as the problem of finding a set of solutions with minimum regret that covers at least one critical subset $C \in \mathcal{C}$. For a given cost-coefficient vector B, such a formulation, which we denote as the *Optimality Cover Problem* (henceforth, OCP), solves for a minimum additional cost solution set whose feedback suffices to guarantee the optimality of $\mathcal{S}^*(B)$.

$$
\begin{aligned}
OCP(B): \quad \min \quad & \sum_{S \in \mathcal{S}} \Delta_S^F(B)\, y_S && \text{(12a)}\\
s.t. \quad & x_a \leq \sum_{S \in \mathcal{S}: a \in S} y_S, \quad a \in A && \text{(12b)}\\
& \sum_{a \in S} \left(l_a(1 - x_a) + b_a x_a\right) \geq z^*(B), \quad S \in \mathcal{S} && \text{(12c)}\\
& x_a,\, y_S \in \{0, 1\}, \quad a \in A, S \in \mathcal{S}, && \text{(12d)}
\end{aligned}
$$

where with a slight abuse of notation, we make the dependence of Δ_S^F on B explicit. By construction, a feasible solution (x, y) to this problem corresponds to incidence vectors of a set $C \subseteq A$ and a cover $\mathcal{E}$ of such a set[4]. In what follows we refer to a solution (x, y) to OCP and the induced pair of sets $(C, \mathcal{E})$ interchangeably.

[4] That is, $(x, y) := (x^C, y^{\mathcal{E}})$ where $x_a^C = 1$ if $a \in C$ and zero otherwise and $y_S^{\mathcal{E}} = 1$ if $S \in \mathcal{E}$ and zero otherwise.

$$
\begin{aligned}
OCP(B): \quad \min \quad & \sum_{S \in \mathcal{S}} \Delta_S^F(B)\; y_S && \text{(12a)}\\
s.t. \quad & x_a \leq \sum_{S \in \mathcal{S}: a \in S} y_S, \quad a \in A && \text{(12b)}\\
& \sum_{a \in S} \left(l_a(1-x_a) + b_a x_a\right) \geq z^*(B), \quad S \in \mathcal{S} && \text{(12c)}\\
& x_a,\, y_S \in \{0,1\}, \quad a \in A, S \in \mathcal{S}, && \text{(12d)}
\end{aligned}
$$

Constraints (12c) guarantee the optimality of $\mathcal{S}^*(B)$ even if costs of elements outside C are set to their lowest possible values (i.e., $b_a = l_a$ for all $a \notin C$), and constraints (12b) guarantee that $\mathcal{E}$ covers C (i.e., $a \in S$ for some $S \in \mathcal{E}$, for all $a \in C$). One can show that the former constraints imply that C is in fact a critical subset, i.e., $C \in \mathcal{C}$. On the other hand, while the converse does not hold in general (i.e., not all incidence vectors of critical subsets satisfy (12c)), all critical subsets covering optimal elements of A do satisfy (12c). In addition, note that when solving (11), one can impose $y_S = 1$ for all $S \in \mathcal{S}^*(\mathbb{E}_F\{B_n\})$ without affecting the objective function, thus one can restrict attention to critical subsets that cover optimal elements of A.

Política Adaptável: π_a.

the exploration set $\mathcal{E}$ is recomputed at the beginning of each cycle i and equaled to the solution to $OCP(\bar{B}_{n_i})$.

$\Gamma(B)$ denote the set of feasible solutions $(C, \mathcal{E})$ to OCP

$\Gamma^*(B)$ denote the set of optimal solutions to $OCP(B)$

Theorem 6.2. *There exists a positive finite constant H_a such that for any $H \geq H_a$ and $N > 0$, the regret of π_a admits the following bound*

$$\frac{R^{\pi_a}(F,N)}{\ln N} \leq G\ \Delta^F_{max} H + o(1),$$

where $G := \max\{|\mathcal{E}| : (C, \mathcal{E}) \in \Gamma(\mathbb{E}_F\{B_n\})\}$.

Algorithm 2 Adaptive policy π_a

Set $i = 0$, $C = A$, and $\mathcal{E}$ a minimal cover of A
for $n = 1$ to N **do**
  if $n \in \Phi$ **then**
    Set $i = i + 1$
    Set $S^* \in \mathcal{S}^* \left(\overline{B}_n\right)$ [Update exploitation set]
    if $(C, \mathcal{E}) \notin \Gamma(\overline{B}_n)$ **then**
      Set $(C, \mathcal{E}) \in \Gamma^*(\overline{B}_n)$ [Update exploration set]
    end if
  end if
  if $T_n(a) < i$ for some $a \in C$ **then**
    Try such an element, i.e., set $S_n = S$ with $S \in \mathcal{E}$ such that $a \in S$ [Exploration]
  else
    Implement $S_n = S^*$ [Exploitation]
  end if
end for

Implementação

Algorithm 3 essentially applies the simple policy in the transient period and eventually implements the adaptive policy with one cycle delay (once the cycles are long enough, the policy essentially implements the exploration and exploitation solutions that the adaptive policy would have implemented in the previous cycle). This one cycle delay does not affect the asymptotic analysis of the policy and hence the performance guaranty of the adaptive policy is preserved. In addition, the short-term performance of this policy should be similar to that of the simple policy.

Algorithm 3 Basic Time-Constrained Asynchronous Policy

Set $i = 0$, $C = A$, and $\mathcal{E}$ a minimal cover of A
Let $S^* \in \mathcal{S}$ be an arbitrary solution and $B_f = B_{OCP}$ be an initial cost estimate
Asynchronously begin solving $f\left(B_f\right)$ and $OCP\left(B_{OCP}\right)$
for $n = 1$ to N **do**
  if $n \in \Phi$ **then**
    Set $i = i + 1$
    if Asynchronous solution to $f\left(B_f\right)$ has finished **then**
      Set $S^* \in \mathcal{S}^*\left(B_f\right)$ [Update exploitation set]
      Set $B_f = \overline{B}_n$
      Asynchronously begin solving $f\left(B_f\right)$
    end if
    if Asynchronous solution to $OCP\left(B_{OCP}\right)$ has finished **then**
      if $(C, \mathcal{E}) \notin \Gamma\left(B_{OCP}\right)$ **then**
        Set $(C, \mathcal{E}) \in \Gamma^*\left(B_{OCP}\right)$ [Update exploration set]
      end if
      Set $B_{OCP} = \overline{B}_n$
      Asynchronously begin solving $OCP\left(B_{OCP}\right)$
    end if
  end if
  if $T_n(a) < i$ for some $a \in C$ **then**
    Try such an element, i.e., set $S_n = S$ with $S \in \mathcal{E}$ such that $a \in S$ [Exploration]
  else
    Implement $S_n = S^*$ [Exploitation]
  end if
end for