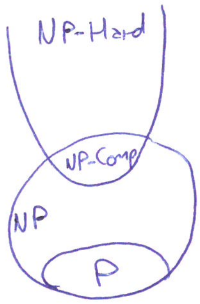
APPROXIMATION ALGORITHMS

1

"find efficient (not necessarily optimal) solutions to intractable problems"

$A \in \mathbb{P}$ A is solvable in poly time in the input size

intractable problem: Assuming $\mathbb{P} \neq NP$, $A \notin P$ is intractable



NP-complete: the hardest problems in NP.
intractable!

opt problem Π: $\boxed{\begin{array}{ll} \min & f(x) \\ s.t & x \in \Omega \end{array}}$ obj.

$\Omega \subseteq \mathbb{R}^n$

feasible domain.

Approx:

1. from intractable to tractable.
   How? perturbing the → input value
   → obj function
   → feasible domain
2. solve the tractable exactly
3. if necessary, convert solution to solution of original

all these steps must be completed in polynomial time

I input instance

Def: performance ration of approx algorithm A:

$$r(A) = \sup_I \frac{A(I)}{opt(I)}$$

OBS $r(A) \geq 1$ and the smaller the better.

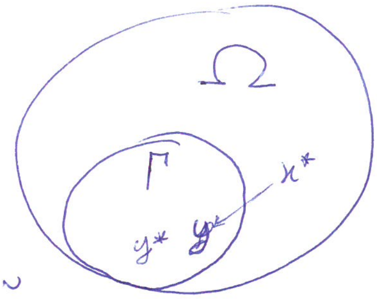
Perturbing the **OBJ. FUNCTION**: ex. greedy algorithm:

A. define potential function $f(A)$ on potential solution sets

B. start: $A \leftarrow \emptyset$

polynomial [C. until "possible"
add element $\{x\}$ such that $\{x\} = \operatorname{argmax}_{y \in \dots} f(A \cup \{x\})$
$A \leftarrow A \cup \{x\}$

Perturbing the **FEASIBLE DOMAIN**: Restriction and Reduction



RESTRICTION

$$\min_{x \in \Omega} f(x)$$

- restrict feasible region to $x \in \Gamma \subseteq \Omega$
- find optimal $y^* \in \Gamma$

ANALYSIS:

OBS: $f(x^*) \le f(y^*)$

- compute $x^* \in \Omega$ (intractable)
- modify it to $y \in \Gamma$ (the closer to y^* the better the bound)

$$\frac{f(y^*)}{f(x^*)} \le \frac{f(y)}{f(x^*)} = 1 + \frac{f(y) - f(x^*)}{f(x^*)}$$

RELAXATION

$$\min_{x \in \Gamma} f(x) \quad \text{wlog}$$

- relax feasible region to $\Omega \supseteq \Gamma$
- find optimal $x^* = \operatorname{argmin}_{x \in \Omega} f(x)$
- round x^* to y and return y:

poly: so also step 3 must be poly.

ANALYSIS:

$$\frac{f(y)}{f(y^*)} \le \frac{f(y)}{f(x^*)}$$

So; ~~the~~ analysis techniques developed for relaxation algorithms are extendable to the analysis of restriction algorithms.

The converse is not true as the rounding might not be done in polynomial time!

GREEDY STRATEGY

Greedy algorithms are very popular.

How do we design and/or analyze a greedy algorithm?

~~BASIC IDEA BEHIND A GREEDY ALGORITHM:~~

~~1) define potential function $f(A)$ on potential solution sets A returned by function~~

Start from defining an independent system

Def $(E, \mathcal{I})$ where E is finite set and $\mathcal{I}$ is a family of subsets of E. (i.e. $\mathcal{I} \subseteq 2^E$)

$(E, \mathcal{I})$ ind system if $I \in \mathcal{I}$ and $I' \subseteq I \Rightarrow I' \in \mathcal{I}$

Consider non-negative function $c: E \longrightarrow \mathbb{R}^+$, and for every subset $F \subseteq E$ define $c(F) = \sum_{e \in F} c(e)$.

MAXIMUM INDEPENDENT SUBSET (MAX-ISS) INPUT: $(E, \mathcal{I})$, c

$\max \; c(I)$

subject to $I \in \mathcal{I}$

Intractable because $\mathcal{I}$ has in general exponential size.

But: usually "easy" to check whether $F \in \mathcal{I}$ (polynomial).

Greedy for MAX-ISS INPUT: $(E, \mathcal{I})$, c

1) sort all elements of E in decreasing order of $c(\cdot)$

   w.l.o.g. $c(e_1) \geq c(e_2) \geq \dots c(e_n)$

2) $I \leftarrow \emptyset$

3) For $i \leftarrow 1$ to n do

   if $I \cup \{e_i\} \in \mathcal{I}$, then $I \leftarrow I \cup \{e_i\}$

4) Output $I_G \leftarrow I$

Pause:

This is a greedy algo for the specific MAX-ISS problem. The potential function is the same as the objective function. Maybe we prove an approx?

Why should we look at the MAX-ISS problem?

If you look at the structure of MAX-ISS is quite general. In fact Two famous probs: MAX-HC and MAX-DHP are special cases of MAX-ISS.

It is really not uncommon to have the feasible domain that is an independent set.

Let's look at this algorithm. What are the possible solutions?

Example

$$\mathcal{I} = \{ \underbrace{\{1,2\}, \{2,3,4\}}_{\text{possible solutions}}, \{1\}, \{2\}, \{3\}, \{4\}, \emptyset, \{2,3\} \{2,4\} \{3,4\} \}$$

possible solutions = max ind subsets of E

Consider $(E, \mathcal{I})$

$\forall F \subseteq E$, a set $I \in \mathcal{I}$ and $I \subseteq F$ is a maximal ind. subset of F, if no ind. subset of F contains I as a proper subset.

<u>Def</u>:

$$u(F) = \min \{ |I| \mid I \text{ maximal ind subset of } F \}$$

$$v(F) = \max \{ |I| \mid I \text{ an ind subset of } F \}$$

ex. $u(E) = 2 \qquad v(E) = 3$

what do we do with these functions?

Theorem 2.1

$\forall (E, \mathcal{I})$ and $c: E \to \mathbb{R}^+$:

$$1 \leq \frac{c(I^*)}{c(I_G)} \leq \max_{F \subseteq E} \frac{v(F)}{u(F)}$$

I^* optimal MAX-ISS
I_G greedy MAX-ISS

calculated the approx ratio of greedy!

and it does not depend on c! Only on the structure of $\mathcal{I}$!

Proof 2.1 $\longrightarrow$ EXTERNAL $\square$

How to apply such result?

Def: Hamiltonian Circuit HC: $G=(V,E)$, simple cycle that passes through each vertex $v \in V$ exactly once.

MAX-HC: given $G=(V,E)$ COMPLETE and a distance function, find a hamiltonian circuit with maximum total distance.

Define $\mathcal{I}$: family of subsets of E:

$I \in \mathcal{I}$ if $\to$ (I Hamiltonian Circuit)
$\to$ I union of disjoint paths (no common vertex)

check in poly time!

Greedy for MAX-ISS works. ① Analyze $\mathcal{I}$ for $v(\cdot)$ and $u(\cdot)$, ② Theorem 2.1 $\Longrightarrow$ greedy is a 2-approx!

MAX-DHP (Hamiltonian Path): $G=(V,E)$ COMPLETE and DIRECTED, and a distance function, find Hamiltonian path (cycle or not) with maximum total distance.

$I \in \mathcal{I} \iff I$ union of disjoint paths. $\longrightarrow$ poly time to check!

Lemma 2.4, Consider $(E, \mathcal{I})$, $F \subseteq E$. Suppose that I and J are two maximal independent subsets of F. Then $|J| \leq 3|I|$.

Proof ... $\longrightarrow$ external $\square$

<u>Theo</u> Greedy is a 3 approximation.

Possible to show that 3 is also lowerbound for greedy!

We have seen two applications for theorem 2.1. Also, we have seen two particular cases of MAX-ISS.

Now we add a condition on the structure of $\mathcal{I}$.

MATROIDS

<u>Def</u>: $(E, \mathcal{I})$ is a MATROID if

- $I \in \mathcal{I}$ and $I' \subseteq I \Rightarrow I' \in \mathcal{I}$ (ind system)
- $\forall F \subseteq E,\ u(F) = v(F)$ "all maximal independent subsets of F have the same cardinality"

$\Rightarrow$ Greedy produces optimal solution! from Theo 2.1.

Actually it is possible to prove that the result holds in both directions!

<u>Theo 2.7</u> $(E, \mathcal{I})$ is a MATROID <u>iff</u> for every cost $c: E \to \mathbb{R}^+$, the greedy algorithm produces an optimal solution for the instance $(E, \mathcal{I}, c)$ of MAX-ISS.

<u>Proof 2.7</u> $\longrightarrow$ external $\square$

What is the relationship between matroids and ind systems?

We might use such relationship to prove some results on approx.

<u>Theo 2.11</u>: $\forall$ $(E, \mathcal{I})$ independent system, there exist a finite number of matroids $(E, \mathcal{I}_i)$, $1 \leq i \leq k$ such that $\mathcal{I} = \bigcap_{i=1}^{k} \mathcal{I}_i$

<u>Proof 2.11</u>

Now a theorem useful for proving approx results.

Theo 2.12 suppose ind system $(E, \mathcal{I})$ is the intersection of k matroids $(E, \mathcal{I}_i)$ that is $\mathcal{I} = \bigcap_{i=1}^{k} \mathcal{I}_i$. Then

$$\max_{F \subseteq E} \frac{v(F)}{u(F)} \leq k .$$

Proof 2.12 $\longrightarrow$ external $\square$

How to use it?

Remember MAX-DHP? we proved that greedy is a 3 approx ratio.

We can prove the same result using this new technique, showing that

the ind system $(E, \mathcal{I})$ for MAX-DHP is the intersection of the following three matroids

1) $(E, \mathcal{I}_1)$ all subgraphs with out-degree at most 1 at each vertex

2) $(E, \mathcal{I}_2)$ all subgraphs with in-degree at most 1 at each vertex

3) $(E, \mathcal{I}_3)$ all subgraphs that do not contain a cycle when the edge direction is ignored.

Now we analyze greedy algorithms for minimization problems instead!

Very famous minimization problem:

<u>MIN-SC</u> (set cover): Given a set S and a collection $\mathcal{C}$ of subsets of S such that $\bigcup_{C \in \mathcal{C}} C = S$, find a subcollection $\mathcal{A} \subseteq \mathcal{C}$ with the minimum cardinality, such that $\bigcup_{C \in \mathcal{A}} C = S$.

i.e. $\arg\min \{ |\mathcal{A}| \mid \bigcup_{C \in \mathcal{A}} C = S, \ \mathcal{A} \subseteq \mathcal{C} \}$

Can we find an approx algo for MIN-SC. We use a procedure similar to MAX-ISS. Also MIN-SC is a very general problem!

New definition. Consider E finite set, and a function $f: 2^E \to \mathbb{Z}$.

<u>Def</u>:

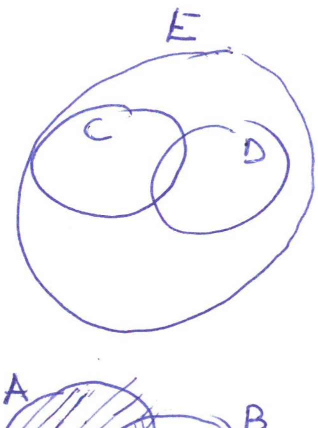
f is SUBMODULAR if $\forall A, B \in 2^E$

$$f(A) + f(B) \geq f(A \cap B) + f(A \cup B)$$

Intuitively? Alternative definition.

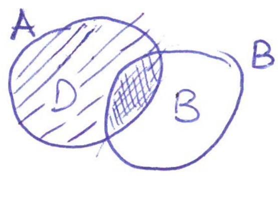
Define
$$\Delta_D f(C) = f(C \cup D) - f(C)$$

$C, D \subseteq E$ Marginal increase due to adding D to C



$$\Delta_D f(A \cap B) \geq \Delta_D f(B)$$

$D = A \setminus B$



Now let us define a potential function for a greedy algorithm for the problem MIN-SC.

$\forall$ subcollection $\mathcal{A} \subseteq \mathcal{C}$:

$$f(\mathcal{A}) = \left| \bigcup_{C \in \mathcal{A}} C \right|$$

f is submodular and also monotone increasing, i.e.

$$\forall A, B \subseteq E \ , \quad A \subseteq B \Rightarrow f(A) \leq f(B).$$

GREEDY for MIN-SC : input S and $\mathcal{C}$

- $A \leftarrow \emptyset$
- while $f(\mathcal{A}) < |S|$ do
  - Select a set $C \in \mathcal{C}$ to maximize $f(\mathcal{A} \cup \{C\})$
  - $\mathcal{A} \leftarrow \mathcal{A} \cup \{C\}$
- Output A

Theorem 2.22 GREEDY for MIN-SC is a polynomial-time $(1 + \ln \gamma)$-approximation for MIN-SC, where $\gamma = \max_{C \in \mathcal{C}} |C|$.

Proof Theo 2.22 $\longrightarrow$ external $\square$

It is possible to prove a similar resul for a problem that is more general than MIN-SC, provided that it is possible to provide a potential function that is submodular.