

Projekt 3: Najtańsze drzewo l_1 -zrównoważone (opis)

PROBLEM. Dany jest zbiór S zawierający punkty płaszczyzny z metryką l_1 . Należy znaleźć najkrótsze drzewo zrównoważone dla S , tzn. takie ukorzenione drzewo T zanurzone na płaszczyźnie z metryką l_1 , że punkty ze zbioru S są liśćmi w T , a wszystkie ścieżki w T łączące korzeń z liściem mają tę samą długość.

Przez **koszt** drzewa rozumiemy sumę długości jego krawędzi; analogicznie, przez **długość** ścieżki rozumiemy sumę jej krawędzi.

Niech T będzie drzewem zrównoważonym i niech R' będzie długością ścieżki od liścia do korzenia. Dla dowolnego $r \geq 0$ oznaczmy przez $T(r)$ liczbę punktów w odległości $R' - r$ od korzenia.

LEMAT 1.

$$\text{koszt}(T) = \int_0^{R'} T(r) \, dr.$$

Dowód. Zauważmy, że funkcja $T(r)$ jest nierosnącą funkcją schodkową, stałą na przedziałach $[0, r_1), [r_1, r_2), \dots, [r_k, R')$, ponadto $T(R') = 1$, oraz $T(r) = 0$ dla $r \in (R', +\infty)$. A zatem, mając na uwadze definicję całki funkcji przedziałami dyskretnej, otrzymujemy

$$\text{koszt}(T) = r_1 \cdot T(0) + |r_2 - r_1| \cdot T(r_1) + \dots + |R' - r_k| \cdot T(r_k) = \int_0^{R'} T(r) \, dr. \quad \square$$

Do analizy współczynnika aproksymacji potrzebować będziemy oszacowania dolnego na koszt OPT rozwiązania optymalnego. Zdefiniujmy **domkniętą l_1 -kulę o promieniu r i środku p** jako zbiór wszystkich punktów, których odległość od p w metryce l_1 wynosi co najwyżej r . Następnie, niech R będzie promieniem najmniejszej domkniętej l_1 -kuli zawierającej wszystkie punkty z S . Dla $r \geq 0$, niech $N(r)$ będzie minimalną liczbą domkniętych l_1 -kul o promieniu r , wystarczającą do pokrycia wszystkich punktów z S ; $N(R) = 1$.

LEMAT 2.

$$\int_0^R N(r) \, dr \leq \text{OPT}.$$

Dowód. Niech T będzie dowolnym drzewem zrównoważonym i niech R' będzie długością ścieżki od liścia do korzenia. Zauważmy, że $R \leq R'$ oraz — z definicji drzewa zrównoważonego — dla dowolnego $r \in [0, R]$ mamy, że $T(r)$ l_1 -kul o promieniu r pokrywa wszystkie punkty z S , a tym samym, $N(r) \leq T(r)$. Z monotoniczności całki otrzymujemy, że (przypomnijmy, że funkcja $T(r)$ przyjmuje wartości nieujemne):

$$\int_0^R N(r) \, dr \leq \int_0^R T(r) \, dr \leq \int_0^{R'} T(r) \, dr.$$

Jako że T jest dowolnym drzewem zrównoważonym, w tym także tym o optymalnym koszcie, teza lematu zachodzi. $\square$

OPIS ALGORYTMU

Zaczynamy od obliczenia najmniejszej l_1 -kuli o promieniu R zawierającej wszystkie punkty z S ; środek tej kuli stanowi korzeń drzewa. Następnie kulę dzielimy na cztery mniejsze kule, zwane **ćwiartkami**, każdą o promieniu $R/2$. Korzeń drzewa można połączyć ze środkami tych kul za pomocą krawędzi długości $R/2$. Każdą z kul/ćwiartek można teraz podzielić na 4 mniejsze kule w ten sam sposób, itd. (Rysunek 1.)

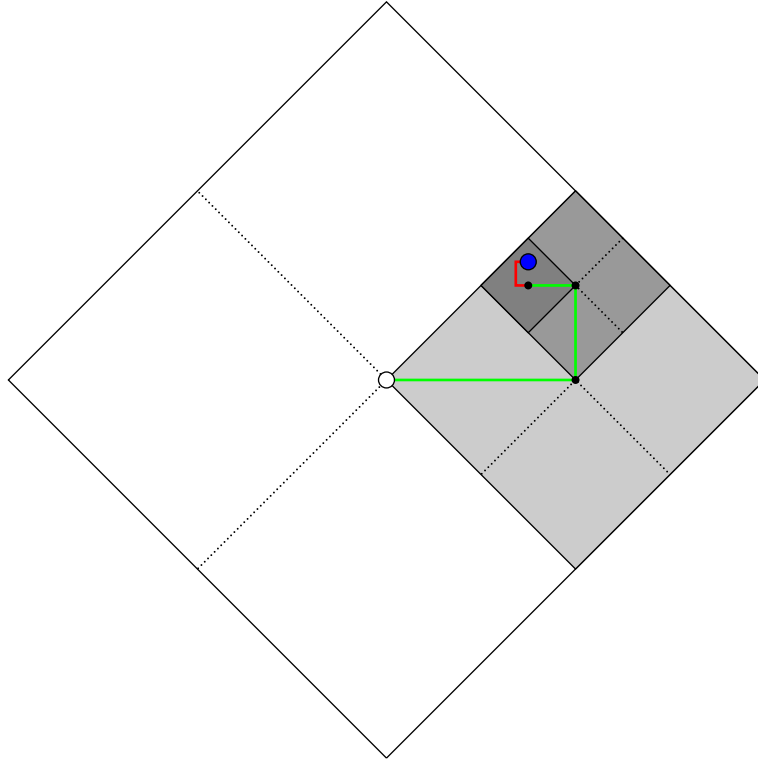


Figure 1: Idea algorytmu dla problemu zrównoważonego l_1 -drzewa.

Drzewo zrównoważone T konstruujemy rekurencyjnie, zaczynając od kuli o promieniu R . Środek aktualnej kuli łączymy ze środkami tych ćwiartek, które zawierają punkty z S , a następnie wykonujemy algorytm dla każdej z tych ćwiartek (Rysunek 1). Jeżeli aktualna kula B o promieniu $r_B = R/2^i$ zawiera dokładnie jeden punkt $p \in S$, to nie dzielimy jej już na ćwiartki, a jedynie łączymy p ze środkiem c kuli B ścieżką długości dokładnie r_B składającą się z odcinków pionowych i poziomych, niekoniecznie najkrótszą (odległość między c i p w metryce l_1 wynosi co najwyżej r_B).

LEMAT 3.

Skonstruowane drzewo T jest zrównoważonym drzewem, w którym odległość od liścia do korzenia wynosi R .

Dowód. Załóżmy, że punkt $p \in S$ został połączony ze środkiem kuli o promieniu $\frac{R}{2^i}$ dla pewnego $i \geq 1$ ($|S| \geq 2$). Wówczas z konstrukcji drzewa mamy, że jego odległość od korzenia wynosi $(\frac{R}{2^1} + \frac{R}{2^2} + \dots + \frac{R}{2^i}) + \frac{R}{2^i} = R$. $\square$

Pozostaje oszacować koszt otrzymanego rozwiązania. Zauważmy, że w najgorszym wypadku,

punkty, które pokrywane są przez którąś z $N(r)$ kul, pokryte są przez dziewięć z $T(r)$ kul. A zatem, dla każdego $0 \leq r \leq R$, zachodzi $T(r) \leq 9N(r)$. W konsekwencji, jako że funkcje $T(r)$ i $N(r)$ są funkcjami nieujemnymi i nierosnącymi oraz ponadto mając na uwadze Lematy 1 oraz 2, otrzymujemy

$$\int_0^R N(r) \, dr \leq \text{koszt}(T) = \int_0^R T(r) \, dr \leq 9 \int_0^R N(r) \, dr \leq 9\text{OPT},$$

gdzie OPT jest kosztem optymalnego zrównoważonego l_1 -drzewa. W konsekwencji, jako że powyższy algorytm może być zaimplementowany tak, aby jego czas działania był wielomianowy, omawiany algorytm ma współczynnik aproksymacji równy 9.

TWIERDZENIE. Powyższy wielomianowy algorytm dla problemu minimalnego zrównoważonego l_1 -drzewa jest algorytmem 9-przybliżonym.