Technika masek znakowych w obliczaniu odległości Manhattan w wielu wymiarach.

Dawid Kot

26 stycznia 2025

1 Wprowadzenie.

W wielu problemach optymalizacyjnych i geometrycznych pojawia się tzw. odległość Manhattan (inaczej odległość L_1), zdefiniowana dla dwóch punktów

$$P = (p_1, p_2, \dots, p_m), \quad Q = (q_1, q_2, \dots, q_m)$$

w przestrzeni m-wymiarowej jako

$$\text{Manhattan}(P, Q) \;=\; \sum_{k=1}^{m} |p_k - q_k|.$$

Często w zadaniach zachodzi potrzeba wielokrotnego obliczania takiej odległości, np. w uogólnionych problemach klastrowania, w algorytmach z programowaniem dynamicznym, albo w różnych wariantach zapytań geometrycznych.

Ciekawostka: odległość Czebyszewa.

W niektórych kontekstach rozpatrywana jest również odległość *Czebyszewa* (oznaczana też L_∞), zdefiniowana przez:

$$\text{Czebyszew}(P, Q) = \max_{1 \le k \le m} |p_k - q_k|.$$

O ile w metryce Manhattan optymalna „transformacja znaków" dotyczy sum wartości bezwzględnych, o tyle w odległości Czebyszewa kluczowe jest *maximum* spośród wartości bezwzględnych. W pewnych zadaniach – zwłaszcza tych związanych z większą tolerancją na odchylenia w pojedynczej współrzędnej – odległość Czebyszewa okazuje się bardziej przydatna.

2 Dlaczego odległość Manhattan daje się sprowadzić do maksimum różnic?

Kluczowa obserwacja polega na tym, że

$$\sum_{k=1}^{m} |p_k - q_k| \;=\; \max_{(\sigma_1,\sigma_2,\dots,\sigma_m)\in\{+1,-1\}^m} \sum_{k=1}^{m} \sigma_k\, p_k \;-\; \sum_{k=1}^{m} \sigma_k\, q_k.$$

Innymi słowy, odległość Manhattan da się zapisać jako maksimum (po wszystkich kombinacjach znaków $+1$ i -1) pewnej różnicy:

$$\Big(\sigma_1\, p_1 + \sigma_2\, p_2 + \cdots + \sigma_m\, p_m\Big) \;-\; \Big(\sigma_1\, q_1 + \sigma_2\, q_2 + \cdots + \sigma_m\, q_m\Big).$$

Logika stojąca za tym wzorem.

- Każdy składnik $|p_k - q_k|$ może być odczytany jako

$$|p_k - q_k| \;=\; \max\big\{(p_k - q_k),\; (q_k - p_k)\big\}.$$

- W zapisie sumy $\sum_{k=1}^{m}(p_k - q_k)$ możemy decydować, czy dany składnik wziąć z plusem (czyli $p_k - q_k$) czy z minusem (co jest równoważne braniu $q_k - p_k$ w sumie), tak aby ostatecznie uzyskać wszystkie wartości bezwzględne.
- Wyborowi plusa lub minusa przy współrzędnej k odpowiada $\sigma_k \in \{+1,-1\}$. W ten sposób szukamy takiej kombinacji $(\sigma_1,\dots,\sigma_m)$, która w sumie da dokładnie $\sum_{k=1}^{m}|p_k - q_k|$.
- Maksimum po $\{+1,-1\}^m$ „wyłapuje” właściwą kombinację znaków, która dostarcza sumy wartości bezwzględnych.

3 Reprezentacja za pomocą masek bitowych.

W praktycznej implementacji przyjmuje się zwykle, że każdemu ciągowi $(\sigma_1,\sigma_2,\dots,\sigma_m)$, gdzie $\sigma_i \in \{+1,-1\}$, odpowiada pewna *maska bitowa* (np. dla $\sigma_i = +1$ ustawiamy bit i w liczbie całkowitej, a dla $\sigma_i = -1$ bit pozostaje nieustawiony – lub odwrotnie). Jeśli m jest niewielkie (np. do 8 czy 10), liczba wszystkich masek 2^m pozostaje stosunkowo mała, co pozwala iterować po nich w rozsądnym czasie.

Definicja przekształcenia punktu przez maskę.

Niech maska będzie podzbiorem współrzędnych $\{1,2,\dots,m\}$. Dla punktu

$$P = (p_1, p_2, \dots, p_m)$$

definiujemy:

$$\mathit{warto\_maska}(P, \text{maska}) = \sum_{k \in \text{maska}} p_k \;-\; \sum_{k \notin \text{maska}} p_k.$$

Z poprzednich rozważań wiemy, że

$$\text{Manhattan}(P, Q) = \max_{\text{maska}} \Big(\mathit{warto\_maska}(P, \text{maska}) \;-\; \mathit{warto\_maska}(Q, \text{maska}) \Big).$$

Inaczej mówiąc, aby szybko obliczyć $\text{Manhattan}(P, Q)$, wystarczy znać wartości $\mathit{warto\_maska}(P, \cdot)$ oraz $\mathit{warto\_maska}(Q, \cdot)$ dla wszystkich masek, a następnie wziąć maksimum różnicy dla odpowiedniej maski.

4 Zastosowanie w zadaniach z programowaniem dynamicznym.

Rozważmy zadanie, w którym:

- Mamy ciąg punktów $\{P_1, P_2, \ldots, P_n\}$ w m-wymiarowej przestrzeni.
- Szukamy pewnego wyniku wyrażonego wzorem

$$\text{dp}[i] = \max_{j<i} \Big(\text{dp}[j] + \text{Manhattan}(P_j, P_i) \Big) \quad \text{(lub wariantu tego typu).}$$

Jeśli po każdej aktualizacji $\text{dp}[j]$ lub podczas obliczania $\text{dp}[i]$ musimy liczyć $\text{Manhattan}(P_j, P_i)$ w prost ($O(m)$ czasowo), to przy dużej liczbie punktów n może się to okazać nieefektywne.

Zamiast tego, korzystamy z faktu, że:

$$\text{Manhattan}(P_j, P_i) = \max_{\text{maska}} \Big(\mathit{warto\_maska}(P_i, \text{maska}) \;-\; \mathit{warto\_maska}(P_j, \text{maska}) \Big).$$

Możemy więc przepisać wyrażenie:

$$\text{dp}[i] = \max_{j<i} \big(\text{dp}[j] + \max_{\text{maska}} \big(\mathit{warto\_maska}(P_i, \text{maska}) - \mathit{warto\_maska}(P_j, \text{maska}) \big) \big).$$

Z kolei $\max_{\text{maska}}$ można „rozdzielić”:

$$\text{dp}[i] = \max_{\text{maska}} \Big(\mathit{warto\_maska}(P_i, \text{maska}) + \max_{j<i} \big(\text{dp}[j] - \mathit{warto\_maska}(P_j, \text{maska}) \big) \Big).$$

Jeśli dla każdej maski maska będziemy w strukturze $\mathit{pom}[\text{maska}]$ przechowywać dotychczas najwyższą wartość $\text{dp}[j] - \mathit{warto\_maska}(P_j, \text{maska})$ (pośród wszystkich $j < i$), to wyliczenie $\text{dp}[i]$ sprowadzi się do przejrzenia jedynie wszystkich masek 2^m, a nie wszystkich j *oraz* masek. Otrzymujemy wówczas złożoność $O(n \cdot 2^m)$ (pomijając koszt wcześniejszego przygotowania wartości $\mathit{warto\_maska}(P_j, \text{maska})$).

5 Dalsze korzyści i przykładowe zastosowania.

- **Zadania typu klastrowanie lub maksymalna różnica:** Jeśli potrzebujemy szybko znajdować parę punktów maksymalnie odległych w metryce Manhattan, możemy przechowywać $warto\_maska(P, \text{maska})$ dla różnych punktów P. Potem największa różnica $\max_{P,Q}[warto\_maska(P, \text{m}) - warto\_maska(Q, \text{m})]$ łatwo daje wynik w metryce L_1.
- **Zapytania dynamiczne:** W pewnych problemach (np. struktury danych do zapytań $\max_{j \in S}\{\text{Manhattan}(P_j, Q)\}$ przy zmieniającym się zbiorze S) można rozważyć przechowywanie maksimum $warto\_maska(P_j, \text{maska})$ w każdym węźle drzewa przedziałowego czy innej struktury, co pozwala w logarytmicznym czasie odpowiadać na zapytania.
- **Inne metryki i geometria obliczeniowa:** Podobne sztuczki bywają używane, jeśli da się „zredukować" inną miarę odległości do pewnej liczby obrotów lub przekształceń. Przykładowo, w odległości *Czebyszewa* (metryka L_∞):

$$\text{Czebyszew}(P, Q) \;=\; \max_{1 \le k \le m} |p_k - q_k|,$$

transformacje dotyczą głównie maksymalnych różnic w pojedynczych współrzędnych.

6 Podsumowanie.

Technika masek znakowych (*sign masks*) wykorzystuje fakt, że metrykę Manhattan można wyrazić jako maksimum różnic przekształconych współrzędnych. Dzięki temu, zamiast każdorazowo sumować wartości bezwzględne, wystarczy:

1. Obliczyć i przechować $warto\_maska(P, \text{maska})$ dla wszystkich punktów P i wszystkich masek 2^m.
2. Każdorazowo przy obliczaniu $\text{Manhattan}(P, Q)$ wziąć $\max_{\text{maska}}[warto\_maska(P, \text{maska}) - warto\_maska(Q, \text{maska})]$.

W zadaniach o małym wymiarze m, 2^m jest wystarczająco małe, by iteracja po wszystkich maskach była efektywna. Podejście to przyspiesza różne algorytmy dynamiczne i pozwala budować struktury danych odpowiadające szybkim zapytaniom na odległość w metryce L_1. Logika działania opiera się na tym, że dla każdej współrzędnej możemy swobodnie wybrać znak $+1$ lub -1, aby oddać wartość bezwzględną $|p_k - q_k|$ w sumie, a wybór ten można reprezentować właśnie jako tzw. maskę bitową.