2024/11/27 Compare Elements in Extended Fields

$\mathbb{Q}\langle \xi_1, \xi_2 \rangle$ algorithm : replace all $\xi_1 \xi_2$ with α_0

then compare $p(\xi_1, \xi_2) = p_1(\xi_1) + p_2(\xi_2)$ separately.

$\mathbb{Q}\langle \xi_1, \xi_2, \xi_3 \rangle$ algorithm :

$$\xi_1 \xi_2 \xi_3 = \alpha_0 \qquad \tau \xi_3 = \alpha_0$$

$$\xi_1\xi_2 + \xi_2\xi_3 + \xi_3\xi_1 = \alpha_1 \qquad \tau + \xi_3(\alpha_2 - \xi_3) = \alpha_1$$

$$\xi_1 + \xi_2 + \xi_3 = \alpha_2 \qquad \tau + \alpha_0\tau^{-1}(\alpha_2 - \alpha_0\tau^{-1}) = \alpha_1$$

$$\tau^3 + \alpha_0(\tau\alpha_2 - \alpha_0) = \alpha_1\tau^2$$

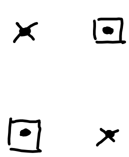
$$\tau^3 - \alpha_1\tau^2 + \alpha_2\alpha_0\tau - \alpha_0^2 = 0$$

2024/11/28 NN Search via Iteration

Iteration : Can be SGD or modification of problem

Reduce Problem to Single Dimensional ...

□ must have higher energy than x

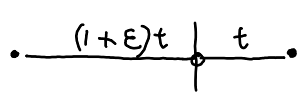


If energy is defined by $-\sum_{i,j} \frac{1}{|q_j - k_{ij}|}$, then it is possible for something to land at a non-point.

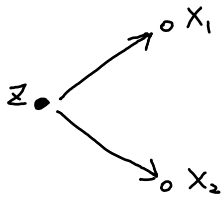
There must be some form of entangle.

$$-\sum_{i,j} \frac{1}{|q_j - k_{ij}| + |q_{j+1} - k_{i(j+1)}|} \quad \text{(when } j = d\text{, } j+1 \text{ wraps back to } 1\text{)}$$

Because it is only 2-dimensional, it is easy to determine which parts are significant.



2024/11/28 Normalize Gravity



• moving mass pt
∘ fixed mass pt

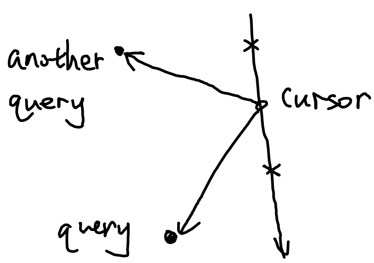
$$g(z) = (x_1 - z) \cdot \frac{1}{\|x_1 - z\|_2^3} + (x_2 - z) \cdot \frac{1}{\|x_2 - z\|_2^3}$$

$$f(z) = g(z) / \|g(z)\|_2$$

$f(z)$: the force on •

2024/11/28 Another Intuition

For a random direction $\vec{u}$, it is very unlikely that $\vec{u} \perp \vec{v}$ ($\vec{v}$: cursor → query)



Therefore, there should almost always possible to improve by moving the cursor.
(if the opposite point always exists in the dataset)

However, the bad thing is that the opposite point is not always available.

Suppose u is v's nearest neighbor, construct w s.t. w is opposite to u and can only be the nearest neighbor of at most 2 nodes.
(v itself and another)

Q: how to make sure the final point is not a constructed point?

2024/11/30 Select Exponential

$f_{\gg} : \mathbb{R}^d \to \mathbb{R}^d$ $\quad n := |D|$

Q : query set $\quad \Psi(q) = \sum_{x \in D} c^{-\|q-x\|}$

When algorithm find x as NN distance,
then real nn distance is at least
$x - \log n / \log c$.

Suppose $\min_{\substack{x \neq y \\ x,y : D}} \|x-y\| = \delta$, then to achieve $(1+\varepsilon)$-approximation

$$\frac{\varepsilon\delta}{2} = \log n / \log c \qquad c = n^{\frac{2}{\varepsilon\delta}}$$

If we set $\Psi(q) = \sum_{x \in D} \frac{1}{c^{|q-x|} - h}$ bigger h, better performance

$$\frac{1}{c^{d_0} - h} = \frac{n}{c^{d_1} - h}$$

$$c^{d_1} - h = n c^{d_0} - nh$$

$$d_1 = \frac{\log n + \log(c^{d_0} - h + \frac{1}{n} h)}{\log c}$$