Поиск вектора, максимизирующего скалярное произведение с заданным запросом

Задача 1. *По множеству точек* $X \subset \mathbb{R}^d$ *и запросу* $q \in \mathbb{R}^d$ *найти точку* $p \in X$*, т.ч.:*

$$p = argmax_X < q, x >$$

Приближенная переформулировка для случая известного множества части запросов

$X = \{x_1, ...x_n\}$ — выборка. $x_i \in \mathbb{R}^d$
$Q = \{q_1, ...q_m\}$ — примеры запросов. $q_i \in \mathbb{R}^d$
Требуется найти разбиение $X = X_1 \sqcup X_2$, которое бы в среднем неплохо отделяло вектора с маленьким скалярным произведением от векторов с большим.

$$\sum_{q \in Q} \left(\frac{1}{||X_1||} \sum_{x_1 \in X_1} < q, x_1 > - \frac{1}{||X_2||} \sum_{x_2 \in X_2} < q, x_2 > \right)^2 \longrightarrow \max_{X_1, X_2}$$

Добавляя условие сбалансированности разделения ($||X_1|| = ||X_2||$), можем перейти к следующей задаче

$$\sum_{Q} \left(\sum_{X_1} < q, x_1 > - \sum_{X_2} < q, x_2 > \right)^2 \longrightarrow \max_{X_1, X_2, ||X_1|| = ||X_2||}$$

Упростим немного выражение слева

$$
\begin{aligned}
&\sum_{Q}\left(\sum_{X_1} < q, x_1 > - \sum_{X_2} < q, x_2 >\right)^2 = \\
=&\sum_{Q}\left(< q, \sum_{X_1} x_1 - \sum_{X_2} x_2 >\right)^2 = \\
=&\sum_{Q}(\sum_{X_1} x_1 - \sum_{X_2} x_2)^T q q^T (\sum_{X_1} x_1 - \sum_{X_2} x_2) = \\
=&(\sum_{X_1} x_1 - \sum_{X_2} x_2)^T Q^T Q (\sum_{X_1} x_1 - \sum_{X_2} x_2) = \\
=&(\sum_{X_1} x_1 - \sum_{X_2} x_2)^T W^T W (\sum_{X_1} x_1 - \sum_{X_2} x_2) = \\
=&(\sum_{X_1} W x_1 - \sum_{X_2} W x_2)^T (\sum_{X_1} W x_1 - \sum_{X_2} L x_2) = \\
=&|| \sum_{X_1} W x_1 - \sum_{X_2} W x_2 ||^2
\end{aligned}
$$

$W \in \mathbb{R}^{r \times d}$, где $r \leq d$.
Если $m \leq d$, то, в качестве W можно взять саму Q.
Если $m \gg d$, то сойдет разложение Холецкого.

$$
\boxed{|| \sum_{X_1} W x_1 - \sum_{X_2} W x_2 ||^2 \longrightarrow \max_{\{X_1, X_2 | X_1 \sqcup X_2 = X, ||X_1|| = ||X_2||\}}}
$$