Maximum Inner-Product Search

For a given set S of N points and a query q, efficiently find a point p in S such that

\[ \langle q, p \rangle = \max_{r \in S} \langle q, r \rangle \]

Applications

- Retrieval of Recommendation after Matrix Factorization
- Document Matching
- Max-kernel Operation
- Greedy Gradient Descent Optimization
- SMO algorithm for SVM learning

Tree-based Branch-and-Bound Algorithm

Maximum inner-product bound for a tree node.

For a tree node indexed with a ball B(C, R) centered at C with radius R, the maximum inner-product possible between the query q and any point in the tree node is

\[ \max_{p \in B(C, R)} \langle q, p \rangle \leq \langle q, C \rangle + R \| q \| \]

Branch-and-Bound Algorithm.

1. Select child @ root
2. Recurse to best child till 1st leaf
3. Obtain best candidate @ leaf
4. Try to prune other children
4a. Explore other children if pruning not possible
4b. Save computation if pruning possible

Dual-tree Branch-and-Bound Algorithm

Maximum inner-product bound between tree nodes.

For a pair of tree nodes indexed with balls B(C_p, R_p) and B(C_q, R_q), the maximum inner-product possible between the query q in B(C_p, R_p) and any point in B(C_q, R_q) is

\[ \max_{p \in B(C_p, R_p) \ q \in B(C_q, R_q)} \langle q, p \rangle \leq \langle C_p, C_q \rangle + R_p \| C_p \| + R_q \| C_q \| + R_p R_q \]

Branch-and-Bound Algorithm.

Things to take home

- Maximum Inner-Product Search is different from Nearest-Neighbor Search in Metric Spaces
- Maximum Inner-Product Search has many applications in Machine Learning & Data Mining
- Code will be shortly available at http://mlpack.org

Maximum Inner-Product Search Using Cone Trees

Parikshit Ram
Georgia Institute of Technology
p.ram@gatech.edu

Alexander G. Gray
Georgia Institute of Technology
agray@cc.gatech.edu

Dual-tree with Cone Trees

Cone Trees

- Maximum inner-product candidate does not depend on the query norm.
- Index the queries only on their angles into cones.

Maximum inner-product bound between tree nodes.

For a ball-tree node indexed with ball B(C_p, R_p) centered around C_p with radius R_p and a cone-tree node indexed with a cone C(A_p, W_p) around the unit direction A_p with aperture W_p, the maximum inner-product possible between the query q in C(A_p, W_p) and any point in B(C_p, R_p) is

\[ \max_{p \in B(C_p, R_p)} \langle q, p \rangle \leq \| C_p \| \cos \left( \cos^{-1} \left( \frac{\langle A_p, C_q \rangle}{\| C_q \|} \right) - W_p \right) + R_p \]

Speedup over Naïve Search

Level 1 interaction

Level 2 interaction

Paper ID: rt216-ram