

# MeanShiftR

Jonathan Lisic

September 9, 2015

## 1 Normal Newton Shift

For Gaussian kernels there is a simple relationship between Newton's Method and mean shift iterations, namely for element  $k$ ,

$$v_k^{(i+1)} = v_k^{(i)} - \frac{\sum_{j=1}^{N_R} (v_k^{(i)} - x_k) \phi \left( (v^{(i)} - x_j)^T H^{-2} (v^{(i)} - x_j) \right)}{\sum_{j=1}^{N_R} \left( 1 - \left( h_k^{-1} (v_k^{(i)} - x_k) \right)^2 \right) \phi \left( (v^{(i)} - x_j)^T H^{-2} (v^{(i)} - x_j) \right)} \quad (1)$$

can be rewritten as

$$v_k^{(i+1)} = v_k^{(i)} - \frac{\sum_{j=1}^{N_R} (v_k^{(i)} - x_k) \phi \left( (v^{(i)} - x_j)^T H^{-2} (v^{(i)} - x_j) \right)}{\sum_{j=1}^{N_R} \left( 1 - \alpha^{(i)} \left( (v_k^{(i)} - x_k) \right)^2 \right) \phi \left( (v^{(i)} - x_j)^T H^{-2} (v^{(i)} - x_j) \right)} \quad (2)$$

for  $\alpha^{(i)} = 1$ . Setting  $\alpha^{(i)} = 0$  will produce the mean shift iterator. This relationship is missing in previous comparisons of mean shift and Newton's algorithm in ?, ? and ?, and will be referred to as NNS (Normal Newton Shift).

In application the choice of  $\alpha$ -sequence  $\{\alpha_i\}_i^\infty$  should allow for convergence to the local mode for all values in the support, while providing for accelerated convergence near the mode. Furthermore the iterator for choice of  $\alpha$ -sequence should be bijective over the support, this allows for a "memoryless" mapping from  $X \rightarrow X$  independent of the iteration index. For simplicity only constant valued  $\alpha$ -sequences, will be considered in this dissertation. This last requirement both significantly simplifies the convergence proof

## 2 Dual Tree Merge-Path Algorithm

Unlike the dual tree mean shift implementation of ? based on the more general ? where a reference tree and a query tree are used to perform high-dimension binning for kernel density estimates. The dual-tree approach presented here instead develops a mapping between  $\mathbb{R}^d \rightarrow \mathbb{R}^d$  by exploiting the fact that means shift mapping is independent of the current iteration, and the observation that on approach to the mode the mean shift iterations follow similar paths. The approach is rather simple, and the algorithm is outlined in 2. The general ideas is to merge sequences that are sufficiently close to each other, where the distance between sequences at iteration  $m$  is defined by

$$d(x_m, y_m) = \min \{d(x_i, y_j) : i \in \{1, \dots, m\}, j \in \{1, \dots, m\}\} \quad (3)$$

for sequences  $x_m = \{x_i\}_{i=1}^m$  and  $y_m = \{y_i\}_{i=1}^m$ . If sampling is used, the final  $T_Q$  tree may be used to classify the non-sampled units through nearest neighbor interpolation.

In the algorithm below the query and reference points, Q and R are within an array.

---

```

 $T_Q \leftarrow \text{tree}(Q)$  {Build k-d tree from Q}
 $T_R \leftarrow \text{tree}(R)$  {Build k-d tree from R}
 $Q^{(0)} \leftarrow Q$ 
while  $i \leq m$  do
   $Q_{\text{KNN}} \leftarrow \text{query}(T_R, Q^{(i-1)}, k)$  {Get k-nearest neighbors.}
   $Q^{(i)} \leftarrow \hat{f}(Q^{(i-1)}, Q_{\text{KNN}})$  {Perform NNS.}
   $Q_{\text{1NN}} \leftarrow \text{query}(T_Q, Q^{(i)}, 1)$  {Get one nearest neighbor.}
  while  $j \leq \text{length}(Q_{\text{1NN}})$  do
     $q^* \leftarrow Q_{\text{1NN}}[j]$ 
    if  $(\|q - q^*\| < \epsilon)$  and  $(\hat{f}(q) < \hat{f}(q^*))$  then
       $Q_{\text{MERGE}} \leftarrow Q_{\text{MERGE}} \cup \{q\}$ 
    end if
     $j = j + 1$ 
  end while
   $Q^{(i)} \leftarrow Q^{(i)} \cup Q_{\text{MERGE}}$ 
   $T_Q \leftarrow \text{tree}(\cup_{j=1}^i Q^{(j)})$ 
   $i = i + 1$ 
end while

```

---