Large-scale Evaluation of Neighborhood Weights and Sizes

The Experiment

We investigate the quality of weighted neighborhoods \( N_i \subset [n] \) with different sizes, see [1] for complete results. The weights are determined by the normal similarity between points and are given by a sigmoid function modeling both continuous and sharp increases. Our large-scale analysis consists of more than 1,000 point sets \( P = \{ p_i \mid i \in [n] \} \), with neighborhoods \( N_i \subset [i] \) and normals \( n_i \in \mathbb{S}^2 \).

Sigmoid Weights

The sigmoid weight function for \( (a, b) \) is given by

\[
\sigma_{a,b}(x) = \frac{1}{1+\exp(-bx)}
\]

For \( x = \frac{(n_i - n_j) + 1}{2} \) and \( j \in N_i \) assign weights for \( p_j \) from \( p_i \) as

\[
w_{ij} = \sigma_{a,b}(x) = \begin{cases} 
0 & x \in (-\infty, a), \\
\frac{1}{2} \cos \left( \frac{b(x - a)}{1-a} \right) + \frac{1}{2} & x \in [a, a'], \\
1 & x \in [a', +\infty),
\end{cases}
\]

with \( a \in [0, 1], b \in \mathbb{R}_+ \cup \{ \infty \} \) and \( a' = (1-a)b^{-1} + a \).

Figure: Plots of the sigmoid \( \sigma_{a,b}(x) \) for three parameter choices.

Evaluation Model

Following [2,3], consider the covariance matrix of \( N_i \)

\[
C_i := \sum_{j \in N_i} w_{ij}(p_j - \bar{p}_i)(p_j - \bar{p}_i)^T.
\]

The eigenvalues \( \lambda_1 > \lambda_2 \geq \lambda_3 \geq 0 \) of \( C_i \) give quantities

\[
L_i^1 = (\lambda_1^1 - \lambda_2^1) / \lambda_1^1 \quad \text{(linearity)},
\]

\[
P_i^1 = (\lambda_1 - \lambda_2^1) / \lambda_1^1 \quad \text{(planarity)},
\]

\[
S_i^1 = \lambda_3^1 / \lambda_1^1 \quad \text{(scattering)}.
\]

For each point \( p_i \), evaluate these via a Shannon-type error

\[
E_i^{\text{dim}} = -L_i^1 \ln(L_i^1) - P_i^1 \ln(P_i^1) - S_i^1 \ln(S_i^1)
\]

and for a range of possible neighborhood sizes \( k \in \mathbb{N} := \{ 6, \ldots, 20 \} \), determine the optimal weighting parameter pair \((a^*, b^*)\) via

\[
(a^*, b^*) = \arg \min_{(a,b) \in \mathbb{P}} \frac{1}{|P|} \sum_{i=1}^{|P|} \min_{k \in \mathbb{N}} E_i^{\text{dim}}. \quad (1)
\]

Optimal Parameter \((a^*, b^*)\) analysis

We run the evaluation from Equation (1) on 1,000 clean [4] and nine noisy models [5].

(a) 1,000 clean models [4].

(b) nine noisy models [5].

Figure: Distribution of optimal parameters \((a^*, b^*)\) on the two model sets.

Could the values be increased \((a+, b+)?\), were they maximal \((a = .9, b = \infty)?\), or would an increase lead to a faulty evaluation, e.g. via empty neighborhoods \((-a+, -b+)\).

<table>
<thead>
<tr>
<th>clean models [4]</th>
<th>a = .9</th>
<th>a+</th>
<th>−a+</th>
<th>b = \infty</th>
<th>b+</th>
<th>−b+</th>
</tr>
</thead>
<tbody>
<tr>
<td>248</td>
<td>20</td>
<td>732</td>
<td>32</td>
<td>968</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>noisy models [5]</th>
<th>0</th>
<th>1</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: Distribution into maximal parameters, possible increase, and failure when increasing.

Conclusions:

- Set \( b \) small, which takes all neighbors into account.
- Set \( a \) as large as possible to only weight similar neighbors highly.
- Have to choose \( a \) small enough to actually obtain a non-empty neighborhood.
- Equal-weights \((a = 0, b = \infty)?\) or sharp cut-off weights \((b = \infty)?\) as widely used in the literature are rarely optimal.

Optimal parameter \( k \) analysis

Plot \( k \) that achieve optimality in Equation (1).

Figure: \( k \) distributed over 1,000 clean models [4].

Figure: \( k \) distribution over nine noisy models [5].

Conclusions:

- Use small \( k \) for clean models, have to vary \( k \) for noisy models.
- Optimality is only achieved with varying neighborhood size \( k \), but literature usually fixes some \( k \).

References