Supplementary Material: Predicting Multiple Structured Visual Interpretations

Debadeepta Dey Varun Ramakrishna Martial Hebert J. Andrew Bagnell
Carnegie Mellon University
Pittsburgh, PA-15213, USA
{debadeep, vramakri, hebert, dbagnell}@ri.cmu.edu

1. Derivation of Equation 15

Note that to achieve a tight upper bound, candidate lines must pass through the point $(\xi^j_{i-1}, 0)$. Substituting $(\xi^j_{i-1}, 0)$ in 15 we get

$$0 = a\xi^j_{i-1} + b \tag{21}$$
$$b = -a\xi^j_{i-1} \tag{22}$$

To obtain the tightest upper bound (in a L_2 sense) we minimize the L_2 distance between l_{Actual} and $l_{\text{SeqNBest2}}$ to obtain a. The L_2 distance between l_{Actual} and $l_{\text{SeqNBest2}}$ is

$$A = \int \|l_{\text{SeqNBest2}} - l_{\text{Actual}}\|_2^2 \, dl \tag{23}$$
$$= \int_0^{\xi^j_{i-1}} [(l^j_i - \xi^j_{i-1}) - (al^j_i + b)]^2 dl \tag{24}$$
$$+ \int_{\xi^j_{i-1}}^1 (al^j_i + b)^2 dl \tag{25}$$

Differentiating the expression for A with respect to a (the slope of the line), setting it to 0, and using the constraint that the line must pass through $(\xi^j_{i-1}, 0)$, we get

$$a = \frac{(\xi^j_{i-1})^3}{3(\xi^j_{i-1})^2 - 3\xi^j_{i-1} + 1} \tag{26}$$

This gives the optimal slope of the line $l_{\text{SeqNBest2}}$ which minimizes the gap between it and l_{Actual}.

2. Proof of monotone submodularity of quality function

Consider the quality function which scores a set of predictions for an input image by the loss of the best prediction in the set. Using the same notation in the paper, the quality function is reproduced as:

$$f(Y_S(I), \mathbf{y}_{gt}) = \max_{i \in 1,\ldots,N} \{\mathbf{g}(\pi_i(I), \mathbf{y}_{gt})\}, \tag{27}$$
$$= 1 - \min_{i \in 1,\ldots,N} \{l(\pi_i(I), \mathbf{y}_{gt})\} \tag{28}$$

The above equation scores the sequence of structured predictions $Y_S(I) = [\![\pi_i(I)]\!]_{i \in 1 \ldots N}$ by the score of the best prediction produced by the predictors $S = [\![\pi_1, \pi_2, \ldots, \pi_N]\!]$. Such a function f was proved to be monotone, submodular by Dey et al. in [1]. We reproduce the proof here for convenience while adapting the exposition to the specific usage in our case:

A set function f which maps subsets $\mathcal{A} \subseteq \mathcal{V}$ of a finite sequence $\mathcal{V}$ to the real numbers. f is called submodular if, for all $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$ and $\mathcal{S} \in \mathcal{V} \setminus \mathcal{B}$ it holds that

$$f(\mathcal{A} \oplus \mathcal{S}) - f(\mathcal{A}) \geq f(\mathcal{B} \oplus \mathcal{S}) - f(\mathcal{B}) \tag{29}$$

where $\oplus$ is the concatenation operator. Such a function is monotone if it holds that for any sets $\mathcal{S}_1, \mathcal{S}_2 \in \mathcal{V}$, we have

$$f(\mathcal{S}_1) \leq f(\mathcal{S}_1 \oplus \mathcal{S}_2) \tag{30}$$
$$f(\mathcal{S}_2) \leq f(\mathcal{S}_1 \oplus \mathcal{S}_2)$$

We want to prove that f (28) is monotone, submodular. We make 28 more general by replacing the loss of a particular prediction $l(\pi_i(I))$ with cost(a_i) where a_i is a particular item. The simplified equation is:

$$f \equiv 1 - \min_{a_i \in A} \{\text{cost}(a_i)\} \tag{31}$$

where A is the set of allowed items.

This can be proved if $\min_{a_i \in A}\{\text{cost}(a_i)\}$ is monotone, supermodular. A function f is supermodular if it holds that

$$f(\mathcal{A} \oplus \mathcal{S}) - f(\mathcal{A}) \leq f(\mathcal{B} \oplus \mathcal{S}) - f(\mathcal{B}) \tag{32}$$

Theorem 1. *The function* $\min_{a_i \in A}\{cost(a_i)\}$ *is monotone, supermodular where* a_i *are predictions.*

Proof. **Submodularity:** Assume that we are given sets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$, $\mathcal{S} \in \mathcal{V} \setminus \mathcal{B}$. We want to prove the inequality in 32. Let $\mathcal{R} = \mathcal{B} \setminus \mathcal{A}$, the set of elements that are in $\mathcal{B}$ but not in $\mathcal{A}$. Since $\mathcal{A} \oplus \mathcal{R} = \mathcal{B}$ we can now rewrite 32 as

$$f(\mathcal{A} \oplus \mathcal{S}) - f(\mathcal{A}) \leq f(\mathcal{A} \oplus \mathcal{R} \oplus \mathcal{S}) - f(\mathcal{A} \oplus \mathcal{R}) \quad (33)$$

We refer to the left and right sides of 33 as LHS and RHS respectively. Define a^* as the prediction which has the least cost. Hence there can be three cases:

- Case 1: $a* \in \mathcal{A}$ In this case $LHS = RHS = 0$
- Case 2: $a* \in \mathcal{R}$ In this case $RHS \geq LHS$
- Case 3: $a* \in \mathcal{S}$ In this case $RHS \geq LHS$

Since in all possible cases it can be seen that RHS is greater than or equal to LHS it is proved that $\min_{a_i \in A}\{\text{cost}(a_i)\}$ is supermodular. Note that if there are multiple predictions which have the same minimum cost as a^* then similar arguments still hold and even in the worst case when they are distributed across $\mathcal{S}$, $\mathcal{R}$ and $\mathcal{A}$, Case 1 holds.

Monotonicity Consider two sequences $\mathcal{S}_1$ and $\mathcal{S}_2$. Define a^* as the predictions which has the least cost. We want to prove that $\min_{a_i \in A}\{\text{cost}(a_i)\}$ is monotone decreasing, i.e.

$$\begin{aligned} f(\mathcal{S}_1) &\geq f(\mathcal{S}_1 \oplus \mathcal{S}_2) \\ f(\mathcal{S}_2) &\geq f(\mathcal{S}_1 \oplus \mathcal{S}_2) \end{aligned} \quad (34)$$

There are three possible cases:

- Case 1: $a^* \in \mathcal{S}_1 \implies f(\mathcal{S}_1) = f(\mathcal{S}_1 \oplus \mathcal{S}_2)$ and $f(\mathcal{S}_2) \geq f(\mathcal{S}_1 \oplus \mathcal{S}_2)$
- Case 2: $a^* \in \mathcal{S}_2 \implies f(\mathcal{S}_1) \geq f(\mathcal{S}_1 \oplus \mathcal{S}_2)$ and $f(\mathcal{S}_2) = f(\mathcal{S}_1 \oplus \mathcal{S}_2)$
- Case 3: $a^* \in \mathcal{S}_1 \oplus \mathcal{S}_2 \implies f(\mathcal{S}_1) = f(\mathcal{S}_1 \oplus \mathcal{S}_2)$ and $f(\mathcal{S}_2) = f(\mathcal{S}_1 \oplus \mathcal{S}_2)$

Since in all possible cases the conditions in 34 are satisfied $\min_{a_i \in A}\{\text{cost}(a_i)\}$ is monotone decreasing. □

Corollary 1. *The function f of Equation 4 in the paper is monotone, submodular due to $\min_{a_i \in A}\{cost(a_i)\}$ being monotone, supermodular by Theorem 1.*

Corollary 2. *The function $F(S, \mathcal{D}) = \mathbb{E}_{(I, \mathbf{y}_{gt}) \sim \mathcal{D}} \; [f(Y_S(I), \mathbf{y}_{gt})]$ (Equation 5 in the paper) is also monotone submodular since non-negative sums of monotone submodular functions is also monotone submodular.*

□

References

[1] D. Dey, T. Y. Liu, B. Sofman, and J. A. Bagnell. Efficient optimization of control libraries. In *AAAI*, 2012.