Do Lessons from Metric Learning Generalize to Image-Caption Retrieval?

Maurits Bleeker and Maarten de Rijke

University of Amsterdam, Amsterdam, The Netherlands
{m.j.r.bleeker, m.derijke}@uva.nl

A Derivative of the gradient of SmoothAP w.r.t. q

In this section, we give an analyses and derivation of the gradient of SmoothAP [1] w.r.t. query $\mathbf{q}$. We start with Eq. 1, the definition of SmoothAP:

$$AP_{\mathbf{q}} = \frac{1}{|\mathcal{S}^{\mathbf{q}}_{\mathcal{P}}|} \sum_{i \in \mathcal{S}^{\mathbf{q}}_{\mathcal{P}}} \frac{1 + \sum_{j \in \mathcal{S}^{\mathbf{q}}_{\mathcal{P}}, j \neq i} \mathcal{G}(D_{ij};\tau)}{1 + \sum_{j \in \mathcal{S}^{\mathbf{q}}_{\mathcal{P}}, j \neq i} \mathcal{G}(D_{ij};\tau) + \sum_{j \in \mathcal{S}^{\mathbf{q}}_{\mathcal{N}}} \mathcal{G}(D_{ij};\tau)}. \tag{1}$$

Here, $\mathcal{G}$ is a smooth approximation of an indicator/step function:

$$\mathcal{G}(f(x);\tau) = \frac{1}{1 + e^{-\frac{f(x)}{\tau}}}. \tag{2}$$

The derivative of $\mathcal{G}$ w.r.t. a function $f(x)$ has the following form:

$$\frac{\partial \mathcal{G}(f(x);\tau)}{\partial x} = \mathcal{G}(f(x);\tau)(1 - \mathcal{G}(f(x);\tau)) \frac{1}{\tau} \frac{\partial f(x)}{\partial x}. \tag{3}$$

Note that $f(x)$ in the case of SmoothAP is D_{ij}:

$$D_{ij} = s_i - s_j = \mathbf{q}\mathbf{v}_i - \mathbf{q}\mathbf{v}_j, \tag{4}$$

where both $\mathbf{v}_i$, $\mathbf{v}_j$ and $\mathbf{q}$ are normalized on the unit-sphere. The gradient of D_{ij} w.r.t. query $\mathbf{q}$ has the following form:

$$\frac{\partial D_{ij}}{\partial \mathbf{q}} = \mathbf{v}_i - \mathbf{v}_j. \tag{5}$$

If we plug in D_{ij} into Eq. 3 and take the gradient w.r.t. query $\mathbf{q}$, we get

$$\begin{aligned} \frac{\partial \mathcal{G}(D_{ij};\tau)}{\partial \mathbf{q}} &= \mathcal{G}(D_{ij};\tau)(1 - \mathcal{G}(D_{ij};\tau)) \frac{1}{\tau} (\mathbf{v}_i - \mathbf{v}_j) \\ &= sim(D_{ij},\tau)(\mathbf{v}_i - \mathbf{v}_j), \end{aligned} \tag{6}$$

where $sim(D_{ij},\tau)$ is a function that gives an indication of how close the similarity scores are of candidate i and j are w.r.t. query $\mathbf{q}$, scaled by τ:

$$sim(D_{ij},\tau) = \mathcal{G}(D_{ij};\tau)(1 - \mathcal{G}(D_{ij};\tau)) \frac{1}{\tau}. \tag{7}$$

Now we define $\mathcal{R}(i,\mathcal{S}^{\mathbf{q}}_{\Omega})$ and $\mathcal{R}(i,\mathcal{S}^{\mathbf{q}}_{\mathcal{P}})$. $\mathcal{R}(i,\mathcal{S}^{\mathbf{q}}_{\Omega})$ gives the ranking of candidate i within the full candidate set $\mathcal{S}^{\mathbf{q}}_{\Omega}$. $\mathcal{R}(i,\mathcal{S}^{\mathbf{q}}_{\mathcal{P}})$. gives the rank of candidate i within the positive candidate set $\mathcal{S}^{\mathbf{q}}_{\mathcal{P}}$.

$$\mathcal{R}(i,\mathcal{S}^{\mathbf{q}}_{\Omega}) = \left(\overbrace{1 + \sum_{j \in \mathcal{S}^{\mathbf{q}}_{\mathcal{P}}, j \neq i} \mathcal{G}(D_{ij};\tau)}^{A} + \overbrace{\sum_{j \in \mathcal{S}^{\mathbf{q}}_{\mathcal{N}}} \mathcal{G}(D_{ij};\tau)}^{C} \right) \tag{8}$$

$$\mathcal{R}(i,\mathcal{S}_{\mathcal{P}}^{\mathbf{q}}) = \left(\overbrace{1+\sum_{j\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}},j\neq i}\mathcal{G}(D_{ij};\tau)}^{A}\right) \tag{9}$$

The gradient of $\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})$ w.r.t. to $\mathbf{q}$ has the following form.

$$\frac{\partial\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})}{\partial\mathbf{q}} = \left(\overbrace{\sum_{j\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}},j\neq i} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)}^{B}+\overbrace{\sum_{j\in\mathcal{S}_{\mathcal{N}}^{\mathbf{q}}} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)}^{D}\right). \tag{10}$$

Using all the detentions above, we can write the full gradient of $AP_{\mathbf{q}}$ w.r.t. $\mathbf{q}$:

$$\frac{\partial AP_{\mathbf{q}}}{\partial\mathbf{q}}$$

$$= \frac{1}{|\mathcal{S}_{P}^{\mathbf{q}}|}\sum_{i\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}}} \frac{\mathcal{R}(i,\mathcal{S}_{\mathcal{P}}^{\mathbf{q}})\frac{\partial\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})}{\partial\mathbf{q}} - \mathcal{R}(i,\mathcal{S}_{\Omega}^{q})\left(\sum_{j\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}},j\neq i} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)\right)}{\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})^2} \tag{11}$$

$$= \frac{1}{|\mathcal{S}_{\mathcal{P}}^{\mathbf{q}}|}\sum_{i\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}}} \frac{1}{\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})^2}\left(\left(\overbrace{\mathcal{R}(i,\mathcal{S}_{\mathcal{P}}^{\mathbf{q}})}^{A}\overbrace{\frac{\partial\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})}{\partial\mathbf{q}}}^{B+D}\right) - \left(\overbrace{\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})}^{A+C}\left(\overbrace{\sum_{j\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}},j\neq i} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)}^{D}\right)\right)\right). \tag{12}$$

When looking at eq. 12 it becomes clear that we have a function in the following form $A(B+D)-(A+C)B$. This can be rewritten to: $AB+AD-AB-CB=AD-CB$. If we apply this to Eq. 12, we end up with the following form.

$$\frac{\partial AP_{\mathbf{q}}}{\partial\mathbf{q}}$$

$$= \frac{1}{|\mathcal{S}_{\mathcal{P}}^{\mathbf{q}}|}\sum_{i\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}}} \frac{1}{\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})^2}\left(\overbrace{\mathcal{R}(i,\mathcal{S}_{\mathcal{P}}^{\mathbf{q}})}^{A}\left(\overbrace{\sum_{j\in\mathcal{S}_{\mathcal{N}}^{\mathbf{q}}} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)}^{D}\right) - \overbrace{\sum_{j\in\mathcal{S}_{\mathcal{N}}^{\mathbf{q}}}\mathcal{G}(D_{ij};\tau)}^{C}\left(\overbrace{\sum_{j\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}},j\neq i} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)}^{B}\right)\right) \tag{13}$$

$$= \frac{1}{|\mathcal{S}_{P}^{\mathbf{q}}|}\sum_{i\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}}} \frac{1}{\mathcal{R}(i,\mathcal{S}_{\Omega}^{\mathbf{q}})^2}\left(\mathcal{R}(i,\mathcal{S}_{\mathcal{P}}^{\mathbf{q}})\left(\sum_{j\in\mathcal{S}_{\mathcal{N}}^{q}} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)\right) - (\mathcal{R}(i,\mathcal{S}_{\mathcal{N}}^{q})-1)\left(\sum_{j\in\mathcal{S}_{\mathcal{P}}^{\mathbf{q}},j\neq i} sim(D_{ij})(\mathbf{v}_i-\mathbf{v}_j)\right)\right) \tag{14}$$

Bibliography

[1] Brown A, Xie W, Kalogeiton V, Zisserman A (2020) Smooth-AP: Smoothing the path towards large-scale image retrieval. In: European Conference on Computer Vision (ECCV), Springer, pp 677–694