Algorithms for Graph-Based Supervised Learning

Brian Dolan

September 3, 2017

Hello

1 Modified Adsorption

This routine is described in Talukdar below.

1.1 Prerequisites

I'm looking for a good parallelization strategy for this function. I'm going to expand it as far as I can to see if something presents itself. We are aiming to find c_v and d_v for each vertex in the graph. First, let's define

$$m(x) = \sum_u W_{u,x} \tag{1}$$

$$l(x) = \sum_u W_{u,x} \log W_{u,x} \tag{2}$$

Okay, check my math here...

$$
\begin{aligned}
p(a|b) &= \frac{W_{a,b}}{\sum_u W_{u,b}} && (3)\\
&= \frac{W_{a,b}}{m(b)} && (4)\\
H(x) &= -\sum_y p(y|x)\log p(y|x) && (5)\\
&= -\sum_y \left(\frac{W_{y,x}}{\sum_u W_{u,x}}\right)\log\left(\frac{W_{y,x}}{\sum_u W_{u,x}}\right) && (6)\\
&= -\sum_y \left(\frac{W_{y,x}}{m(x)}\right)\log\left(\frac{W_{y,x}}{m(x)}\right) && (7)\\
&= \frac{-1}{m(x)}\sum_y W_{y,x}\left[\log W_{y,x} - \log m(x)\right] && (8)\\
&= \frac{-1}{m(x)}\left[\sum_y W_{y,x}\log W_{y,x} - \sum_y W_{y,x}\log m(x)\right] && (9)\\
&= \frac{-1}{m(x)}\left[\sum_y W_{y,x}\log W_{y,x} - m(x)\log m(x)\right] && (10)\\
&= \log m(x) - \frac{1}{m(x)}\sum_y W_{y,x}\log W_{y,x} && (11)\\
&= \log m(x) - \frac{l(x)}{m(x)} && (12)
\end{aligned}
$$

Next we have the smoothing function for a given β

$$
\begin{aligned}
f(x) &= \frac{\log\beta}{\log(\beta+e^x)} && (13)\\
c_x &= f(H(x)) && (14)\\
&= \log\beta\left[\log(\,\beta+e^{H(x)}\,)\right]^{-1} && (15)\\
d_x &= (1-c_x)\sqrt{H(x)} && (16)\\
z_x &= \max(c_x+d_x, 1) && (17)
\end{aligned}
$$

Given these values, the authors define

$$
p_v^{cont} = \frac{c_v}{z_v}, p_v^{inj} = \frac{d_v}{z_v}, p_v^{abnd} = 1 - p_v^{cont} - p_v^{inj}
$$

1.2 Algorithm 3: Modified Adsoprtion

Taken from the book reference below. This will optimize

$$C(\hat{Y}) = \sum_l \left[\mu_1 \left(Y_l - \hat{Y}_l \right)^T S \left(Y_l - \hat{Y}_l \right) + \mu_2 \hat{Y}_l^T L \hat{Y}_l + \mu_3 \|\hat{Y}_l - R_l\|^2 \right]$$

where $\mu_{1,2,3}$ are "hyperparameters that determine the relative importance of each term" in C. So Dealer's Choice on μ. Y_l and R_l are the l^{th} columns of Y and R. The definition of R and the choice of convergence criteria are unclear at the moment.

I've asked Dr. Talukdar for guidance on this selection and I'm awaiting his response.

Algorithm 2 Modified Adsoprtion

procedure INPUT:
- $G = (V, E, W), M, R$
- Labels $= Y_v \in \mathcal{R}^{m+1}$ for $v \in V$
- Probabilities $p_v^{inj}, p_v^{cont}, p_v^{abnd}$ for $v \in V$
- Constants μ_1, μ_1, μ_3

procedure OUTPUT:
- $\hat{Y}_v$ for $v \in V$

procedure ITERATE
- $\hat{Y}_v \leftarrow Y_v$ for $v \in V$ [Initialization]
- $M_{vv} \leftarrow \mu_1 \times p_v^{inj} + \mu_2 \times p_v^{cont} \times \sum_u W_{vu} + \mu_3$
- **while** Not Converged **do**
  - $D_v \leftarrow \sum_x \left(p_v^{cont} W_{vx} + p_x^{cont} W_{xv}\right) \hat{Y}_x$
  - **for all** $v \in V$ **do**
    - $\hat{Y} \leftarrow \frac{1}{M_{vv}} \left(\mu_1 \times p_v^{inj} \times Y_v + \mu_2 \times D_v + \mu_3 \times p_v^{abnd} \times R_v \right)$

2 References

BibTeX is a pain, so for right now I'm going to do

```
@article{doi:10.2200/S00590ED1V01Y201408AIM029,
author = {  Amarnag
  Subramanya  and   Partha Pratim
  Talukdar },
title = {Graph-Based Semi-Supervised Learning},
journal = {Synthesis Lectures on Artificial Intelligence and Machine Learning},
volume = {8},
number = {4},
pages = {1-125},
year = {2014},
doi = {10.2200/S00590ED1V01Y201408AIM029},
```

```
URL = {
        https://doi.org/10.2200/S00590ED1V01Y201408AIM029

},
eprint = {
        https://doi.org/10.2200/S00590ED1V01Y201408AIM029

}

}
```