Spectral Analysis of Real Weighted Graphs

Ryan Greenup & James Guerra

December 5, 2020

Contents

Headline	Time		
Total time	**8:31**		
Tasks	8:31		
Derive the Sigmoid Curve			4:00
Derive and explain the equation on...		4:31	

Table 1: Clock summary at *[2020-12-05 Sat 19:40]*

Project Topic	*Spectral Analysis of Real Weighted Graphs*
Adviser	Assoc. Prof. Laurence Park

```
zotero & disown
kitty -e 'cd /home/ryan/Sync/Studies/2020ResearchTraining/spectral_analysis_graphs/' &
↪ disown
mattermost & disown
```

1 Tasks

1.0.1 Derive the Sigmoid Curve

Just make sure that you understand logistic regression so there is a ground-truth classification method.

1.1 Derive and explain the equation on page 3

RankNet is concerned with ranking things, the most obvious application of this is information retrieval, for example if a search returns a variety of results that are very similar to the query they will need to be ordered in some way to make them more useful. This is traditionally done with *PageRank* or by simply using TF-IDF weighting, it is not immediately clear what advantages this approach has, presumably it ranks results better.

The output of Ranknet is concerned with a probability of one object being ranked higher than another and implements a Neural Network in order to model this. It is not clear why RankNet is used for this rather than traditional classification techniques.

Take two objects U_i, U_j (e.g. an article, document or anything else) that are described by a feature vector $\mathbf{X}_i$ and $\mathbf{X}_j$. RankNet maps this feature vector to some number:

$$f : \mathbb{R}^n \to \mathbb{R} : \mathbf{X} \mapsto s$$

If U_i is ranked higher than U_j this is denoted by:

$$U_i \triangleright U_j$$

The two outputs (s_i, s_j) are mapped to a probability that U_i is ranked higher than U_j via a sigmoid function:

$$p_{ij} \equiv P\left(U_i \triangleright U_j\right) \equiv \frac{1}{1 + e^{-\sigma(s_i - s_j)}}$$

The model is dependent on the σ value as shown in figure **??**.

`media/sigmoid_animation.gif`

In order to fit this curve a penalty term needs to be introduced to measure how well the curve fits the data. In regression analysis (\RMSE=($\sum_{\text{i=1}}^{\text{n}}$ [$(xi - \hat{x})$ 2]\) is often used and could be implemented here.

Assume this is a binary classification problem (i.e. they won't match ranks) and let p describe the probability of an object belonging to class 1. If an observation belongs to class 1 we can measure the *badness of fit* by $1 - p$, conversely if an object truly belongs to class 0 we can measure the badness of fit by p, this is illustrated in figure 1.

Actual Class	Residual	Cost
$p = 0$	p	$\ln \frac{1}{1-p}$
$p = 1$	$1 - p$	$\ln \frac{1}{p}$
Either	$p^{(1-\overline{p}) \cdot (1-p)^{p}}$	$\ln \left(p^{-p} \cdot (1-p)^{-(\overline{p}-1)}\right)$

Table 2: Residuals and costs function for

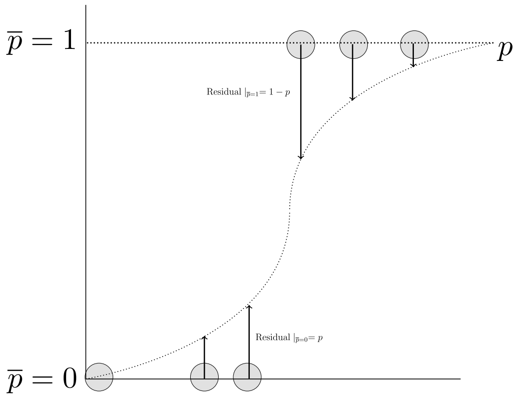

Figure 1: Residual from Classified points

These residuals could be combined to account for either situation:

$$C = p^{1-\overline{p}} \cdot (1-p)^{\overline{p}} \tag{1}$$

A log transform would give a more convenient function (i.e. no exponents), because the transform is monotone this will still work as a cost function and so this could be implemented:

$$C = (1 - \overline{p}) \cdot \ln(p) + \overline{p} \cdot \ln(1 - p) \tag{2}$$

This however is not implemented, instead the cost functions implemented correspond to table 2, presumably this is due to the fact that this cost function performs well:

$$e^C = P_{ij}^{-\overline{P}_{ij}} \cdot (1 - P_{ij})^{\left(\overline{P}_{ij} - 1\right)} \tag{3}$$

$$\implies C = -\overline{P}_{ij} \ln (P_{ij}) - \left(1 - \overline{P}_{ij}\right) \ln (1 - P_{ij}) \tag{4}$$

Let the actual status of the ranking be defined by $S_{i,j} \in \{-1, 0, 1\}$ like so:

S_{ij}	Status
1	$U_i \rhd U_j$
-1	$U_i \lhd U_j$
0	$U_i \ \square \ U_j$

This provides that:

$$\overline{P}_{ij} = \frac{1}{2} (1 + S_{ij}) \tag{5}$$

```
c=1; p = 0.7; L(x) = 1-p
c=0; p = 0.7; L(x) = p
p^c * (1-p)^(1-c)
```

$$p^c (1-p)^{(1-c)}$$

It is however more convenient to log transform this, a log transform is monotone and so we can still use it as a loss function. so we are dealing with sums and products rather than exponentials:

$$\log \left(p^c (1-p)^{(1-c)}\right) = \dots$$

The equation on page 3 we get by plugging in the the preceeding two equations (value of p and value of $\overline{p}$)

1.2 Watch and summarise the Neural Networks Video

Create a Citation for this: `RankNet/RankNet_To_LambdaRank.pdf`

1.3 Create a Neural Network to Classify binary data

- Get the Data from UCI
- Use three softwares to get an idea for it:
  - RTorch
  - pytorch
  - JuliaTorch?

Compare the usability and performance of the different OS.

1.4 Set up a Wiki and put the working in there

And/or put the working into MkDocs Pages.

1.5 Set up the wiki and comment on Mattermost

1.6 Add to the CDRMS Repo

1.7 Set up Kanboard Page

Share that page with Laurence.

This means that U_i is ranked higher than U_j

1.8 Derive the sigmoid function for logistic regression

1.9 Clarify the Question

Is there any benefit to using Ranknet rather than simply using Classification?

2 Ranknet

The Ranknet uses a Sigmoid Curve curve:

$$U_i \triangleright U_j$$

3 Summary

3.1 What is spectral Graph Theory

Spectral graph theory starts by associating matrices to graphs, mostly the:

- Adjacency Matrix
- Laplacian Matrix

3.2 What do we mean by real weighted graphs

Like non-integer values.

3.3 What do I want to look at researching

3.4 What is the Research Outline

3.4.1 This is what I put together in Planner

4 Research

-

4.1 Research Papers

- Tu, E., Zhang, Y., Zhu, L., Yang, J., & Kasabov, N., A Graph-Based Semi-Supervised k Nearest-Neighbor Method for Nonlinear Manifold Distributed Data Classification (). [4]
  - PDF
- Chang, S. Y., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., & Leskovec, J., Mobility network modeling explains higher SARS-CoV-2 infection rates among disadvantaged groups and informs reopening strategies (). [2]
  - PDF

4.2 Books

- Bondy, J. A., & Murty, U. S. R., Graph theory with applications (), : North Holland. [1]
  - PDF (local, absolute path)
- Nicodemi, O., Sutherland, M. A., & Towsley, G. W., An introduction to abstract algebra with notes to the future teacher (), : Pearson Prentice Hall. [3]
  - Books on Spectral Graph Theory

5 Ring Theory ATTACH

A rign is a set that has two operations (see [3, §§2.4-2.6]):

- Addition $(+)$
- Multiplication $\cdot$

And Satisfies the axioms of a ring:

1. Associativity of Addition
   $(\forall a, b, c \in \mathcal{R})\, (a+b)+c = a+(b+c)$

1. Commutativity of Addition
   $(\forall a, b \in \mathcal{R})\, a+b = b+a$

2. Additive Elements Exist
   $(\forall a \in \mathcal{R}) \land (\exists_0 \in \mathcal{R})\, a+0 = 0+a = a$

3. Additive Inverse Exists
   $(\forall a \in \mathcal{R}) \land (\exists b \in \mathcal{R})\, a+b = b+a = 0$

   - This can be equivalently expressed:

$(\forall a \in \mathcal{R}) \wedge (\exists\, (-a) \in \mathcal{R})\, a + (-a) = (-a) + a = 0$

4. Associativity of Multiplication
   $(\forall a, b, c, \in \mathcal{R})\, (a \cdot b) \cdot c = a \cdot (b \cdot c)$

1. Distributivity of Multiplication over Addition
   - $(\forall a, b, c, \in \mathcal{R})\, (a \cdot (b + c) = (a \cdot b) + (a \cdot c))$, AND
   - $(\forall a, b, c, \in \mathcal{R})\, (a + b) \cdot c = (a \cdot c) + (b \cdot c)$

5.0.1 Further Axioms

Other conditions to have special classes of rings exist:

1. Commutativity of Multiplication
   - A ring that satisfies this property is called a **commutative ring**
     $(\forall a, b \in \mathcal{R})\, a \cdot b = b \cdot a$
2. Existence of a Multiplicative Identity Element (A ring with Unity)
   - A ring that satisfies this property is called a **ring with identity** or

   equivalently a **ring with unity** (the multiplicative identity, often denoted by 1, is called the **unity** of the ring.
   $(\exists 1 \in \mathcal{R})\, (\forall a \in \mathcal{R})\, 1 \cdot a = a \cdot 1 = a$

5.1 Integral Domain

An integral domain is a ring that:

1. is commutative
2. With identity/unity
3. Has no Zero Divisors

In an integral domain we can cancel values:

$$(c \neq 0) \wedge (ac = bc) \implies a = b$$

5.2 Fields

A field is:

1. An Integral Domain
2. In which every non-zero element is a unit

A unit in a ring is an element of a ring that always has a multiplicative identity.

References

[1] J. A. Bondy and U. S. R. Murty. *Graph Theory with Applications*. Includes index. New York: North Holland, 1976. 264 pp. ISBN: 978-0-444-19451-0 (cit. on p. 6).

[2] Serina Y Chang et al. *Mobility Network Modeling Explains Higher SARS-CoV-2 Infection Rates among Disadvantaged Groups and Informs Reopening Strategies*. preprint. Epidemiology, June 17, 2020. DOI: 10.1101/2020.06.15.20131979. URL: http://medrxiv.org/lookup/doi/10.1101/2020.06.15.20131979 (visited on 12/02/2020) (cit. on p. 6).

[3] Olympia Nicodemi, Melissa A. Sutherland, and Gary W. Towsley. *An Introduction to Abstract Algebra with Notes to the Future Teacher*. Includes bibliographic references (S. 391-394) and index. Upper Saddle River, NJ: Pearson Prentice Hall, 2007. 436 pp. ISBN: 978-0-13-101963-8 (cit. on p. 6).

[4] Enmei Tu et al. *A Graph-Based Semi-Supervised k Nearest-Neighbor Method for Nonlinear Manifold Distributed Data Classification*. Comment: 32 pages, 12 figures, 7 tables. June 3, 2016. arXiv: 1606.00985 [cs, stat]. URL: http://arxiv.org/abs/1606.00985 (visited on 12/02/2020) (cit. on p. 6).