Heriot-Watt University

Masters Thesis

Bayesian Reconsruction and Regression over Networks

Author:
John Smith

Supervisor:
Dr. James Smith

A thesis submitted in fulfilment of the requirements for the degree of MSc.

in the

School of Mathematical and Computer Sciences

January 2023

Declaration of Authorship

I, John Smith, declare that this thesis titled, 'Bayesian Reconsruction and Regression over Networks' and the work presented in it is my own. I confirm that this work submitted for assessment is my own and is expressed in my own words. Any uses made within it of the works of other authors in any form (e.g., ideas, equations, figures, text, tables, programs) are properly acknowledged at any point of their use. A list of the references employed is included.

Signed: _____

Date: _____

"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."

Dave Barry

Abstract

The Thesis Abstract is written here (and usually kept to just this page).

Acknowledgements

The acknowledgements and the people to thank go here, don't forget to include your project advisor :)

Contents

Declaration of Authorship i

Abstract iii

Acknowledgements iv

Contents v

List of Figures vii

List of Tables viii

Abbreviations ix

Symbols x

1 Introduction 1
- 1.1 Background and Definitions . . . 1
- 1.2 Thesis overview . . . 2

2 Outline and Fundamentals 3
- 2.1 Graph Signal Processing . . . 3
  - 2.1.1 A broad overview of the field . . . 3
  - 2.1.2 The graph Laplacian . . . 3
  - 2.1.3 Graph filters . . . 3
- 2.2 Regression and Reconstruction . . . 3
  - 2.2.1 Graph Signal Reconstruction . . . 3
  - 2.2.2 Kernel Graph Regression . . . 3
  - 2.2.3 Regression with Network Cohesion . . . 3

3 Kernel Generalized Least Squares Regression for Network Data 4
- 3.1 Kernel Graph Regression with Missing Values . . . 4
- 3.2 GLS Kernel Graph Regression . . . 4
  - 3.2.1 A Gauss-Markov estimator . . . 4
  - 3.2.2 AR(1) processes . . . 4

3.2.3 Experiments 4

4 Regression and Reconstruction on Cartesian Product Graphs 5
4.1 Product Graphs 5
4.2 Graph Signal Reconstruction on Cartesian Product Graphs 7
4.2.1 A stationary iterative method 7
4.2.2 A conjugate gradient method 7
4.2.3 Convergence properties 7
4.2.4 Image processing experiments 7
4.3 Kernel Graph Regression with Unrestricted Missing Data Patterns 7
4.3.1 Cartesian product graphs and KGR 8
4.3.2 Convergence properties 8
4.4 Regression with Network Cohesion 8
4.4.1 Regression with node-level covariates 8
4.4.2 Convergence properties 8
4.5 Multi-Dimensional Cartesian Product Graphs 8
4.5.1 Fast computation with d-dimensional Kronecker products 8
4.5.2 Signal reconstruction 8
4.5.3 Kernel Graph Regression 8
4.5.4 Regression with Network Cohesion 9

5 Signal Uncertainty: Estimation and Sampling 10
5.1 Introduction 10
5.2 Posterior Estimation 10
5.2.1 Log-variance prediction 10
5.2.2 Estimation models 10
5.2.3 Query strategies 10
5.2.4 Comparison and analysis 10
5.3 Posterior Sampling 10
5.3.1 Perturbation optimization 10
5.4 Estimation vs Sampling 10
5.4.1 Experiments 10

6 Working with Binary-Valued Graph Signals 11
6.1 Logistic Graph Signal Reconstruction 11
6.2 Logistic Kernel Graph Regression 11
6.3 Logistic Regression with Network Cohesion 11
6.4 Approximate Sampling via the Laplace Approximation 11

7 Conclusions 12
7.1 Main Section 1 12

A Appendix Title Here 13

List of Figures

1.1 A graphical depiction of a graph signal. Here, the nodes are represented by circles, the edges as dotted lines, and the value of the signal at each node is represented by the height of its associated bar. 2

List of Tables

Abbreviations

GSP	Graph Signal Processing
GSR	Graph Signal Reconstruction
KGR	Kernel Graph Regression
RNC	Regression with Network Cohesion
GLS	Generalised Least Squares

Symbols

Unless otherwise specified, the following naming conventions apply.

Integer constants

N	The number of nodes in a graph
T	The number of time points considered
M	The number of explanatory variables
Q	The number of queries

Integer variables

n	The index of a specific node in a graph
t	The index of a specific time point
m	The index of a specific explanatory variable
q	The index of a specific query
i, j, k	Generic indexing variables

Scalar variables

α	An autocorrelation regularisation parameter
β	A hyperparameter characterising a graph filter
γ	A precision parameter
λ	An eigenvalue *or* ridge regression penalty parameter
μ	The mean of a random variable
θ	AR(1) autocorrelation parameter
σ^2	The variance of a random variable

Matrices

$\mathbf{A}$	The graph adjacency matrix
$\mathbf{D}$	A diagonal matrix
$\mathbf{E}$	The prediction residuals
$\mathbf{F}$	A predicted graph signal
$\mathbf{G}$	A graph filter matrix
$\mathbf{H}$	A Hessian matrix
$\mathbf{I}_N$	The $(N \times N)$ identity matrix
$\mathbf{J}_N$	An $(N \times N)$ matrix of ones
$\mathbf{K}$	A kernel (Gram) matrix
$\mathbf{L}$	The graph Laplacian
$\mathbf{S}$	A binary selection matrix
$\mathbf{U}$	Laplacian eigenvector matrix
$\mathbf{V}$	Kernel eigenvector matrix
$\mathbf{X}$	Data matrix of explanatory variables
$\mathbf{Y}$	(Partially) observed graph signal
$\mathbf{\Lambda}$	A diagonal eigenvalue matrix
$\mathbf{\Sigma}$	A covariance matrix
$\mathbf{\Phi}$	A generic eigenvector matrix
$\mathbf{\Psi}$	A generic eigenvector matrix
$\mathbf{\Omega}$	Log marginal variance matrix

Vectors/tensors

$\mathbf{e}$	The prediction residuals
$\mathbf{f}$	The predicted graph signal
$\mathbf{s}$	A binary selection vector/tensor
$\mathbf{x}$	A vector of explanatory variables
$\mathbf{y}$	The observed graph signal
$\boldsymbol{\alpha}$	A flexible intercept vector/tensor
$\boldsymbol{\beta}$	A graph filter parameter vector *or* vector of regression coefficients
$\boldsymbol{\theta}$	A aggregated coefficient vector $[\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top]^\top$

Functions

$g(\cdot)$	A graph filter function
$p(\text{statement})$	The probability that a statement is true
$\pi(\cdot)$	A probability density function
$\xi(\cdot)$	Optimisation target function
$\kappa(\cdot, \cdot)$	A kernel function

Operations

$(\cdot)^\top$	Transpose of a matrix/vector
$\|\| \cdot \|\|_\mathrm{F}$	The Frobenius norm
$\mathrm{tr}\big(\cdot \big)$	The trace of a square matrix
$\mathrm{vec}(\cdot)$	Convert a matrix to a vector in column-major order
$\mathrm{vec}_\mathrm{RM}(\cdot)$	Convert a matrix to a vector in row-major order
$\mathrm{mat}(\cdot)$	Convert a vector to a matrix in column-major order
$\mathrm{mat}_\mathrm{RM}(\cdot)$	Convert a vector to a matrix in row-major order
$\mathrm{diag}(\cdot)$	Convert a vector to a diagonal matrix
$\mathrm{diag}^{-1}(\cdot)$	Convert the diagonal of a matrix into a vector
$\otimes$	The Kronecker product
$\oplus$	The Kronecker sum
$\circ$	The Hadamard product

Graphs

$\mathcal{G}$	A graph
$\mathcal{V}$	A vertex/node set
$\mathcal{E}$	An edge set

Miscellaneous

$\hat{(\cdot)}$	The estimator of a matrix/vector/tensor
$O(\cdot)$	The runtime complexity
x_i	A vector element
$\mathbf{X}_i$	A matrix column
$\mathbf{X}_{ij}$	A matrix element

For/Dedicated to/To my...

Chapter 1

Introduction

1.1 Background and Definitions

Graph Signal Processing (GSP) is a rapidly evolving field that sits at the intersection between spectral graph theory, statistics and data science [Shuman et al., 2013]. In this context, a graph is an abstract collection of objects in which any pair may be, in some sense, "related". These objects are referred to as vertices (or nodes) and their connections as edges [Newman, 2018]. GSP is concerned with the mathematical analysis of signals that are defined over the nodes of a graph, referred to simply as *graph signals*.

A graph signal can be thought of as a value that is measured simultaneously at each node in a graph. In practice, it is represented as a vector where each element corresponds to a single node. For example, consider a social network where each node represents an individual and presence of an edge between two nodes indicates that the two individuals have met. An example of a graph signal in this context could be the age of each person in the network. Figure 1.1 shows a graphical depiction of a signal defined over a network.

Graphs and graph signals have proven a useful way to describe data across a broad range of applications owing to their flexibility and relative simplicity. They are able to summarise the of properties large, complex systems within a single easily-digestible structure. Much of the data

The GSP community, in particular, is focused on generalising tools designed for traditional signal processing tasks to irregular graph-structured domains.

[Ortega et al., 2018]

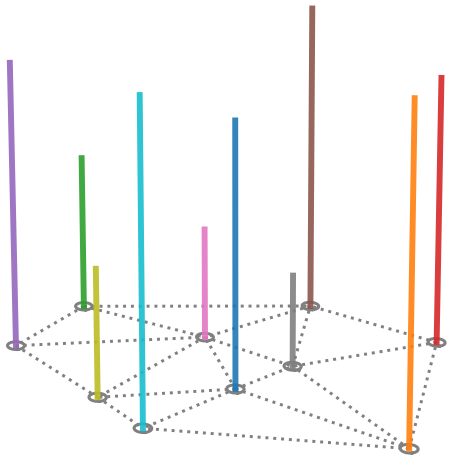

FIGURE 1.1: A graphical depiction of a graph signal. Here, the nodes are represented by circles, the edges as dotted lines, and the value of the signal at each node is represented by the height of its associated bar.

1.2 Thesis overview

Chapter 2

Outline and Fundamentals

2.1 Graph Signal Processing

2.1.1 A broad overview of the field

2.1.2 The graph Laplacian

2.1.3 Graph filters

2.2 Regression and Reconstruction

2.2.1 Graph Signal Reconstruction

Introduce the known work on GSR

2.2.2 Kernel Graph Regression

Introduce the known work on KGR and GPoG

2.2.3 Regression with Network Cohesion

Introduce the known work on RNC

Chapter 3

Kernel Generalized Least Squares Regression for Network Data

3.1 Kernel Graph Regression with Missing Values

3.2 GLS Kernel Graph Regression

3.2.1 A Gauss-Markov estimator

3.2.2 AR(1) processes

3.2.3 Experiments

Chapter 4

Regression and Reconstruction on Cartesian Product Graphs

4.1 Product Graphs

In this chapter, we turn our attention to the topic of signal processing on *Cartesian product graphs*. This special class of graph finds applications in numerous areas, such as video, hyper-spectral image processing and network time series problems. However, the Cartesian product is not the only way to consistently define a product between two graphs. In this section we formally introduce the concept of a graph product, examine some prominent examples, and explain why we choose to look specifically at the Cartesian graph product.

In the general case, consider two undirected graphs $\mathcal{G}_N = (\mathcal{V}_N, \mathcal{E}_N)$ and $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$ with vertex sets given by $\mathcal{V}_N = \{n \in \mathbb{N} \,|\, n \leq N\}$ and $\mathcal{V}_T = \{t \in \mathbb{N} \,|\, t \leq T\}$ respectively. (In this context we do not regard zero to be an element of the natural numbers). A new graph $\mathcal{G}$ can be constructed by taking the product between $\mathcal{G}_N$ and $\mathcal{G}_T$. This can be generically written as follows.

$$\mathcal{G} = \mathcal{G}_N \diamond \mathcal{G}_T = (\mathcal{V}, \mathcal{E}) \tag{4.1}$$

For all definitions of a graph product, the new vertex set $\mathcal{V}$ is given by the Cartesian product of the vertex sets of the factor graphs, that is

$$\mathcal{V} = \mathcal{V}_N \times \mathcal{V}_T = \{(n, t) \in \mathbb{N}^2 \,|\, n \leq N \text{ and } t \leq T\} \tag{4.2}$$

Each consistent rule for constructing the new edge set $\mathcal{E}$ corresponds to a different definition of a graph product. In general, there are eight possible conditions for deciding whether two nodes (n, t) and (n', t') are to be connected in the new graph.

$$
\begin{array}{lccc}
1. & [n,\, n'] \in \mathcal{E}_N & \text{and} & t = t' \\
2. & [n,\, n'] \notin \mathcal{E}_N & \text{and} & t = t' \\
3. & [n,\, n'] \in \mathcal{E}_N & \text{and} & [t,\, t'] \in \mathcal{E}_T \\
4. & [n,\, n'] \notin \mathcal{E}_N & \text{and} & [t,\, t'] \in \mathcal{E}_T \\
5. & [n,\, n'] \in \mathcal{E}_N & \text{and} & [t,\, t'] \notin \mathcal{E}_T \\
6. & [n,\, n'] \notin \mathcal{E}_N & \text{and} & [t,\, t'] \notin \mathcal{E}_T \\
7. & n = n' & \text{and} & [t,\, t'] \in \mathcal{E}_T, \\
8. & n = n' & \text{and} & [t,\, t'] \notin \mathcal{E}_T
\end{array}
$$

Each definition of a graph product corresponds to the union of a specific subset of these conditions, thus, there exist 256 different types of graph product [Barik et al., 2015]. Of these, the Cartesian product (conditions 1 or 7), the direct product (condition 3), the strong product (conditions 1, 3 or 7) and the lexicographic product (conditions 1, 3, 5 or 7) are referred to as the standard products and are well-studied [Imrich and Klavžar, 2000]. In each of these four cases, the adjacency and Laplacian matrices of the product graph can be describes in terms of matrices relating to the factor graphs [Barik et al., 2015, 2018].

	Adjacency	Laplacian
Cartesian	$\mathbf{A}_N \oplus \mathbf{A}_T$	$\mathbf{L}_N \oplus \mathbf{L}_T$
Direct	$\mathbf{A}_N \otimes \mathbf{A}_T$	$\mathbf{D}_N \otimes \mathbf{A}_T + \mathbf{A}_N \otimes \mathbf{D}_T - \mathbf{L}_N \otimes \mathbf{L}_T$
Strong	$\mathbf{A}_N \otimes \mathbf{A}_T + \mathbf{A}_N \oplus \mathbf{A}_T$	$\mathbf{D}_N \otimes \mathbf{A}_T + \mathbf{A}_N \otimes \mathbf{D}_T$
Lexicographic	$\mathbf{I}_N \otimes \mathbf{A}_T + \mathbf{A}_N \oplus \mathbf{J}_N$	$\mathbf{I}_N \oplus \mathbf{L}_T + \mathbf{L}_N \otimes \mathbf{J}_N + \mathbf{D}_N \otimes (N\mathbf{I}_N - \mathbf{J}_N)$

For the Cartesian product, which is commonly denoted as $\mathcal{G}_N \,\square\, \mathcal{G}_T$, the new edge set is constructed as follows.

$$
\mathcal{E} = \Big\{ \big[(n,\, t),\, (n', t')\big] \;\Big|\; \big(n = n' \text{ and } [t, t'] \in \mathcal{E}_T\big) \text{ or } \big(t = t' \text{ and } [n, n'] \in \mathcal{E}_N\big) \Big\} \tag{4.3}
$$

This can be described more succinctly by considering the respective adjacency matrices of the factor graphs. If $\mathcal{G}_N$ and $\mathcal{G}_T$ have adjacency matrices given by $\mathbf{A}_N$ and $\mathbf{A}_T$, then

the adjacency matrix of their Cartesian product is given by the Kronecker sum of their respective adjacency matrices [Fiedler, 1973].

$$\mathbf{A} = \mathbf{A}_N \oplus \mathbf{A}_T \tag{4.4}$$

Note that while, given this definition, it may seem that the Cartesian product of graphs is non-commutative since, in general, $\mathbf{A}_N \oplus \mathbf{A}_T \neq \mathbf{A}_T \oplus \mathbf{A}_N$, the graphs $\mathcal{G}_N \square \mathcal{G}_T$ and $\mathcal{G}_T \square \mathcal{G}_N$ are isomorphically identical [Imrich and Klavžar, 2000]. The only key difference arises from the ordering of the node sets.

4.2 Graph Signal Reconstruction on Cartesian Product Graphs

One task that is of particular interest

4.2.1 A stationary iterative method

Hello

4.2.2 A conjugate gradient method

Hello

4.2.3 Convergence properties

Hello

4.2.4 Image processing experiments

Hello

4.3 Kernel Graph Regression with Unrestricted Missing Data Patterns

Hello

4.3.1 Cartesian product graphs and KGR

Hello

4.3.2 Convergence properties

Hello

4.4 Regression with Network Cohesion

Hello

4.4.1 Regression with node-level covariates

Hello

4.4.2 Convergence properties

Hello

4.5 Multi-Dimensional Cartesian Product Graphs

Hello

4.5.1 Fast computation with d-dimensional Kronecker products

Hello

4.5.2 Signal reconstruction

Hello

4.5.3 Kernel Graph Regression

Hello

4.5.4 Regression with Network Cohesion

Chapter 5

Signal Uncertainty: Estimation and Sampling

5.1 Introduction

5.2 Posterior Estimation

5.2.1 Log-variance prediction

5.2.2 Estimation models

5.2.3 Query strategies

5.2.4 Comparison and analysis

5.3 Posterior Sampling

5.3.1 Perturbation optimization

5.4 Estimation vs Sampling

5.4.1 Experiments

Chapter 6

Working with Binary-Valued Graph Signals

6.1 Logistic Graph Signal Reconstruction

6.2 Logistic Kernel Graph Regression

6.3 Logistic Regression with Network Cohesion

6.4 Approximate Sampling via the Laplace Approximation

Chapter 7

Conclusions

7.1 Main Section 1

Appendix A

Appendix Title Here

Write your Appendix content here.

Bibliography

Barik, S., Bapat, R. B., and Pati, S. (2015). On the laplacian spectra of product graphs. *Applicable Analysis and Discrete Mathematics*, 9:39–58.

Barik, S., Kalita, D., Pati, S., and Sahoo, G. (2018). Spectra of graphs resulting from various graph operations and products: a survey. *Special Matrices*, 6:323 – 342.

Fiedler, M. (1973). Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23:298–305.

Imrich, W. and Klavžar, S. (2000). *Product Graphs: Structure and Recognition*. A Wiley-Interscience publication. Wiley.

Newman, M. (2018). *Networks*. Oxford University Press.

Ortega, A., Frossard, P., Kovačević, J., Moura, J. M. F., and Vandergheynst, P. (2018). Graph signal processing: Overview, challenges, and applications. *Proceedings of the IEEE*, 106(5):808–828.

Shuman, D. I., Narang, S. K., Frossard, P., Ortega, A., and Vandergheynst, P. (2013). The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3):83–98.