Proximal point methods for inverse problems

Marc Paulhamus

Follow this and additional works at: http://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
Proximal Point Methods for Inverse Problems

by

Marc Paulhamus

A thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in Applied Mathematics
from the School of Mathematical Sciences
Rochester Institute of Technology

26 August 2011

Advisor: Dr. Akhtar Khan
Committee members: Dr. Patricia Clark
                   Dr. Baasansuren Jadamba
                   Dr. Miguel Sama
Dedicated to my parents
Larry and Connie Paulhamus
with deep love.
Proximal Point Methods for Inverse Problems

by

Marc Paulhamus

Abstract

Numerous mathematical models in applied mathematics can be expressed as a partial differential equation involving certain coefficients. These coefficients are known and they describe some physical properties of the model. The direct problem in this context is to solve the partial differential equation. By contrast, an inverse problem asks for the identification of the variable coefficients when a certain measurement of a solution of the partial differential equation is available. One of the most commonly used approaches for solving this inverse problem is by posing a constrained minimization problem which can be written as a variational inequality. This paper investigates the inverse problem of identifying certain material parameters in the fourth-order partial differential equations representing the beam and plate models. This inverse problem has attracted a great deal of attention in recent years and has found numerous applications. Since the numerical treatment of the fourth-order problems is rather challenging, the first part of the paper describes in detail the finite element approach for solving the direct problem. The inverse problem is solved by posing an optimization problem whose solution is an approximation of the parameters sought. The optimization problem is solved by gradient based approaches, and in this setting, the most challenging aspect is the computations of the gradient of the objective function. We present a detailed treatment of three approaches to compute the gradient, namely, the so-called adjoint method, adjoint stiffness based approach, and the classical gradient computation. We also present a comparison among these different ways of gradient computation. We use different proximal point methods to solve the inverse problem. It is known that proximal point method is a regularization method which has a significantly different behavior than the well-know Tikhonov regularization method. We present a detailed comparative analysis of the numerical efficiency of several proximal point methods.
Acknowledgement

Words cannot express my gratitude for my advisor, Dr. Akhtar Khan for his professional advice and assistance in reviewing this thesis. He genuinely cares for his students and his love of mathematics makes for an astounding combination to lead his student to great accomplishments. Bless him and his family.

I would like to thank the School of Mathematical Sciences at RIT. The faculty here is truly exceptional. They go beyond the expectation of sharing their knowledge by demonstrating how mathematics should be studied at a higher level.

Lastly, the encouragement of my parents cannot go unnoticed. They have been there for me emotionally, morally, and financially. It makes a world of a difference knowing your parents are behind you.

Special thanks to my committee members: Prof. Patricia Clark, Dr. Baasansuren Jadamba, Dr. Miguel Sama, and the director of graduate programs, Prof. Tamas Wiandt.
# Contents

List of Figures vii

List of Tables x

1 Introduction to Inverse Problems 1
   1.1 Inverse Problems ........................................ 1
   1.2 Optimization Formulation ............................... 2
   1.3 Organization ............................................. 3

2 Fourth-Order Boundary Value Problems 4
   2.1 The Weak Form ........................................... 4
   2.2 The Finite Element Method .............................. 5
   2.3 Implementation Issues ................................... 12
       2.3.1 Load Vector ........................................ 12
       2.3.2 Stiffness Matrix .................................... 14
   2.4 Numerical Experiments .................................. 20

3 Optimization Formulation of Inverse Problems 23
   3.1 Problem Formulation ..................................... 23
   3.2 Optimization Formulation: Finite Dimensional Case .... 27
   3.3 Output Least-Squares (OLS) .............................. 27
   3.4 Modified Output Least-Squares (MOLS) .................. 30
   3.5 Regularization .......................................... 32

4 Identification by Derivative Free Methods 34
   4.1 Nelder-Mead Method ..................................... 34
## CONTENTS

4.2 Nelder-Mead Algorithm .................................................. 35  
4.3 Numerical Results .......................................................... 36  

5 Computation of the Derivatives 39  
5.1 Adjoint Stiffness Matrix Based Approach .............................. 39  

6 Proximal Point Method 66  
6.1 Optimization Formulation .................................................. 66  
6.2 Test Examples .............................................................. 68  
6.3 Proximal-Like Methods Using Least Squares ............................ 69  
6.4 Proximal-like Methods Using $\varphi$-Divergence ....................... 75  
6.5 Proximal-like Methods Using Bregman Functions ..................... 81  
6.6 Proximal-like Methods Using Modified $\varphi$-Divergence ............... 88  
6.7 Han Proximal-Point Method ............................................. 94  
6.8 Li-Yuan Proximal Point Method ....................................... 99  

7 Comparative Performance Analysis 104  
7.1 Numerical Results for Example 1 ....................................... 104  
7.2 Numerical Results for Example 2 ....................................... 108  
7.3 Numerical Results for Example 3 ....................................... 112  
7.4 Numerical Results for Example 4 ....................................... 115  

Bibliography 118
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Basis Function $\phi_j$</td>
<td>7</td>
</tr>
<tr>
<td>2.2</td>
<td>Basis Function $\psi_j$</td>
<td>8</td>
</tr>
<tr>
<td>2.3</td>
<td>Example 1</td>
<td>20</td>
</tr>
<tr>
<td>2.4</td>
<td>Example 2</td>
<td>21</td>
</tr>
<tr>
<td>2.5</td>
<td>Example 3</td>
<td>21</td>
</tr>
<tr>
<td>2.6</td>
<td>Example 4</td>
<td>22</td>
</tr>
<tr>
<td>4.1</td>
<td>Example 1 using Nelder-Mead Simplex Method</td>
<td>36</td>
</tr>
<tr>
<td>4.2</td>
<td>Example 2 using Nelder-Mead Simplex Method</td>
<td>37</td>
</tr>
<tr>
<td>4.3</td>
<td>Example 3 using Nelder-Mead Simplex Method</td>
<td>37</td>
</tr>
<tr>
<td>4.4</td>
<td>Example 4 using Nelder-Mead Simplex Method</td>
<td>38</td>
</tr>
<tr>
<td>5.1</td>
<td>Basis function $\phi_j$</td>
<td>41</td>
</tr>
<tr>
<td>5.2</td>
<td>Basis function $\psi_j$</td>
<td>42</td>
</tr>
<tr>
<td>5.3</td>
<td>Basis function $a_0$</td>
<td>42</td>
</tr>
<tr>
<td>5.4</td>
<td>Basis function $a_1$</td>
<td>42</td>
</tr>
<tr>
<td>5.5</td>
<td>Basis function $a_2$</td>
<td>44</td>
</tr>
<tr>
<td>5.6</td>
<td>Basis function $a_3$</td>
<td>45</td>
</tr>
<tr>
<td>5.7</td>
<td>Basis function $a_{n-1}$</td>
<td>47</td>
</tr>
<tr>
<td>5.8</td>
<td>Basis function $a_n$</td>
<td>48</td>
</tr>
<tr>
<td>5.9</td>
<td>Basis functions $a_{n+1}$</td>
<td>49</td>
</tr>
<tr>
<td>6.1</td>
<td>Example 1: Coefficient by Hager-Zhang Method using Least Squares</td>
<td>71</td>
</tr>
<tr>
<td>6.2</td>
<td>Example 1: Solution by Hager-Zhang Method using Least Squares</td>
<td>71</td>
</tr>
<tr>
<td>6.3</td>
<td>Example 2: Coefficient by Hager-Zhang Method using Least Squares</td>
<td>72</td>
</tr>
<tr>
<td>6.4</td>
<td>Example 2: Solution by Hager-Zhang Method using Least Squares</td>
<td>72</td>
</tr>
<tr>
<td>Example (Section)</td>
<td>Method Description</td>
<td>Page</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------</td>
<td>------</td>
</tr>
<tr>
<td>6.5</td>
<td>Example 3: Coefficient by Hager-Zhang Method using Least Squares</td>
<td>73</td>
</tr>
<tr>
<td>6.6</td>
<td>Example 3: Solution by Hager-Zhang Method using Least Squares</td>
<td>73</td>
</tr>
<tr>
<td>6.7</td>
<td>Example 4: Coefficient by Hager-Zhang Method using Least Squares</td>
<td>74</td>
</tr>
<tr>
<td>6.8</td>
<td>Example 4: Solution by Hager-Zhang Method using Least Squares</td>
<td>74</td>
</tr>
<tr>
<td>6.9</td>
<td>Example 1: Coefficient by Hager-Zhang Method using (\varphi)-divergence</td>
<td>77</td>
</tr>
<tr>
<td>6.10</td>
<td>Example 1: Solution by Hager-Zhang Method using (\varphi)-divergence</td>
<td>77</td>
</tr>
<tr>
<td>6.11</td>
<td>Example 2: Coefficient by Hager-Zhang Method using (\varphi)-divergence</td>
<td>78</td>
</tr>
<tr>
<td>6.12</td>
<td>Example 2: Solution by Hager-Zhang Method using (\varphi)-divergence</td>
<td>78</td>
</tr>
<tr>
<td>6.13</td>
<td>Example 3: Coefficient by Hager-Zhang Method using (\varphi)-divergence</td>
<td>79</td>
</tr>
<tr>
<td>6.14</td>
<td>Example 3: Solution by Hager-Zhang Method using (\varphi)-divergence</td>
<td>79</td>
</tr>
<tr>
<td>6.15</td>
<td>Example 4: Coefficient by Hager-Zhang Method using (\varphi)-divergence</td>
<td>80</td>
</tr>
<tr>
<td>6.16</td>
<td>Example 4: Solution by Hager-Zhang Method using (\varphi)-divergence</td>
<td>80</td>
</tr>
<tr>
<td>6.17</td>
<td>Example 1: Coefficient by Hager-Zhang Method using Bregman function</td>
<td>84</td>
</tr>
<tr>
<td>6.18</td>
<td>Example 1: Solution by Hager-Zhang Method using Bregman function</td>
<td>84</td>
</tr>
<tr>
<td>6.19</td>
<td>Example 2: Coefficient by Hager-Zhang Method using Bregman function</td>
<td>85</td>
</tr>
<tr>
<td>6.20</td>
<td>Example 2: Solution by Hager-Zhang Method using Bregman function</td>
<td>85</td>
</tr>
<tr>
<td>6.21</td>
<td>Example 3: Coefficient by Hager-Zhang Method using Bregman function</td>
<td>86</td>
</tr>
<tr>
<td>6.22</td>
<td>Example 3: Solution by Hager-Zhang Method using Bregman function</td>
<td>86</td>
</tr>
<tr>
<td>6.23</td>
<td>Example 4: Coefficient by Hager-Zhang Method using Bregman function</td>
<td>87</td>
</tr>
<tr>
<td>6.24</td>
<td>Example 4: Solution by Hager-Zhang Method using Bregman function</td>
<td>87</td>
</tr>
<tr>
<td>6.25</td>
<td>Example 1: Coefficient by Hager-Zhang Method using quadratic function</td>
<td>90</td>
</tr>
<tr>
<td>6.26</td>
<td>Example 1: Solution by Hager-Zhang Method using quadratic function</td>
<td>90</td>
</tr>
<tr>
<td>6.27</td>
<td>Example 2: Coefficient by Hager-Zhang Method using quadratic function</td>
<td>91</td>
</tr>
<tr>
<td>6.28</td>
<td>Example 2: Solution by Hager-Zhang Method using quadratic function</td>
<td>91</td>
</tr>
<tr>
<td>6.29</td>
<td>Example 3: Coefficient by Hager-Zhang Method using quadratic function</td>
<td>92</td>
</tr>
<tr>
<td>6.30</td>
<td>Example 3: Solution by Hager-Zhang Method using quadratic function</td>
<td>92</td>
</tr>
<tr>
<td>6.31</td>
<td>Example 4: Coefficient by Hager-Zhang Method using quadratic function</td>
<td>93</td>
</tr>
<tr>
<td>6.32</td>
<td>Example 4: Solution by Hager-Zhang Method using quadratic function</td>
<td>93</td>
</tr>
<tr>
<td>6.33</td>
<td>Example 1: Coefficient by Han Method</td>
<td>95</td>
</tr>
<tr>
<td>6.34</td>
<td>Example 1: Solution by Han Method</td>
<td>95</td>
</tr>
<tr>
<td>6.35</td>
<td>Example 2: Coefficient by Han Method</td>
<td>96</td>
</tr>
<tr>
<td>6.36</td>
<td>Example 2: Solution by Han Method</td>
<td>96</td>
</tr>
<tr>
<td>6.37</td>
<td>Example 3: Coefficient by Han Method</td>
<td>97</td>
</tr>
</tbody>
</table>
6.38 Example 3: Solution by Han Method ........................................ 97
6.39 Example 4: Coefficient by Han Method .................................... 98
6.40 Example 4: Solution by Han Method ....................................... 98
6.41 Example 1: Coefficient by Li-Yuan Method ................................. 100
6.42 Example 1: Solution by Li-Yuan Method .................................... 100
6.43 Example 2: Coefficient by Li-Yuan Method ................................. 101
6.44 Example 2: Solution by Li-Yuan Method .................................... 101
6.45 Example 3: Coefficient by Li-Yuan Method ................................. 102
6.46 Example 3: Solution by Li-Yuan Method .................................... 102
6.47 Example 4: Coefficient by Li Method ....................................... 103
6.48 Example 4: Solution by Li-Yuan Method .................................... 103
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Results of Example 1</td>
<td>105</td>
</tr>
<tr>
<td>7.2</td>
<td>Example 1. Using Hager and Zhang-LS</td>
<td>106</td>
</tr>
<tr>
<td>7.3</td>
<td>Example 1. Using Hager and Zhang-(\varphi)</td>
<td>106</td>
</tr>
<tr>
<td>7.4</td>
<td>Example 1. Using Hager and Zhang-Bregman Function</td>
<td>107</td>
</tr>
<tr>
<td>7.5</td>
<td>Example 1 Using Hager and Zhang-Quadratic Function</td>
<td>107</td>
</tr>
<tr>
<td>7.6</td>
<td>New Results of Example 1</td>
<td>108</td>
</tr>
<tr>
<td>7.7</td>
<td>Results of Example 2</td>
<td>108</td>
</tr>
<tr>
<td>7.8</td>
<td>Example 2. Using Hager and Zhang-NR</td>
<td>109</td>
</tr>
<tr>
<td>7.9</td>
<td>Example 2. Using Hager and Zhang-LS</td>
<td>109</td>
</tr>
<tr>
<td>7.10</td>
<td>Example 2. Using Hager and Zhang-(\varphi)</td>
<td>110</td>
</tr>
<tr>
<td>7.11</td>
<td>Example 2. Using Hager and Zhang-Bregman Function</td>
<td>110</td>
</tr>
<tr>
<td>7.12</td>
<td>Example 2. Using Hager and Zhang-Quadratic Function</td>
<td>111</td>
</tr>
<tr>
<td>7.13</td>
<td>New Results of Example Two</td>
<td>111</td>
</tr>
<tr>
<td>7.14</td>
<td>Results of Example 3</td>
<td>112</td>
</tr>
<tr>
<td>7.15</td>
<td>Example 3. Using Hager and Zhang-NR</td>
<td>112</td>
</tr>
<tr>
<td>7.16</td>
<td>Example 3. Using Hager and Zhang-LS</td>
<td>113</td>
</tr>
<tr>
<td>7.17</td>
<td>Example 3. Using Hager and Zhang-(\varphi)</td>
<td>113</td>
</tr>
<tr>
<td>7.18</td>
<td>Example 3. Using Hager and Zhang-Bregman Function</td>
<td>114</td>
</tr>
<tr>
<td>7.19</td>
<td>Example 3. Using Hager and Zhang-Quadratic Function</td>
<td>114</td>
</tr>
<tr>
<td>7.20</td>
<td>New Results of Example 3</td>
<td>115</td>
</tr>
<tr>
<td>7.21</td>
<td>Results of Example Four</td>
<td>115</td>
</tr>
<tr>
<td>7.22</td>
<td>Example 4. Using Hager and Zhang-NR</td>
<td>116</td>
</tr>
<tr>
<td>7.23</td>
<td>Example 4. Using Hager and Zhang-LS</td>
<td>116</td>
</tr>
<tr>
<td>7.24</td>
<td>Example 4. Using Hager and Zhang-(\varphi)</td>
<td>117</td>
</tr>
<tr>
<td>7.25</td>
<td>Example 4. Using Hager and Zhang-Bregman Function</td>
<td>117</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>7.26</td>
<td>Example 4. Using Hager and Zhang-Quadratic</td>
<td>118</td>
</tr>
<tr>
<td>7.27</td>
<td>New Results of Example 4</td>
<td>118</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction to Inverse Problems

This chapter gives a general description of inverse problems and introduces the specific inverse problem that is the subject of this thesis. We conclude this chapter by giving a detailed outline and organization of this thesis.

1.1 Inverse Problems

The field of inverse problems has grown tremendous interest in the realm of researchers. This is justified by the importance of the applications they lead to. With the advancement of fast computers and stable numerical methods, inverse problems are leading the way in medical imaging, computer vision, physics, and many other fields.

In this work, we deal with an inverse problem of coefficient identification in boundary value problems. In direct problems, we are given physical quantities (or the model parameters), and need to find the solution. On the other hand, in inverse problems, we have a measurement of the solution, the data, and the objective is to find the model parameters.

Our primary objective is to study the inverse problem of the identification of a variable coefficient in the beam equation. To be specific, we focus on the following static fourth-order boundary value problem in $\Omega = (0, 1)$:

\[
\begin{align*}
(a(x)u'')' &= f(x), 0 < x < 1, \\
u(0) &= u'(0) = 0, \\
u(1) &= u'(1) = 0,
\end{align*}
\]

(1.1a) (1.1b) (1.1c)

where $a(x)$ is a variable coefficient and $f$ is a suitable function.
The boundary conditions (1.1b)-(1.1c) are the so-called clamped boundary conditions. However, our approach can easily be carried over to other types of boundary conditions as well. This inverse problem has attracted a great deal of attention in recent years and has found numerous applications. Some details can be found in [7], [10], [11], [21], [27], [28], [35], [38], [39], [40], [41], [42], [43], [44], [45], [46], [48] and the cited references therein. Referring to (3.1), the goal of the direct problem for this BVP is to find \( u(x) \), given that \( a(x) \) and \( (f) \) are known. In the inverse problem, the goal is to find the coefficient \( a(x) \) when a measurement \( z \) of the solution \( u(x) \) is known. Inverse problems are challenging primarily due to the fact that they are ill-posed. To explain this concept, we recall that a problem is well-prosed if it satisfies the following three conditions:

1. A solution exist.
2. The solution is unique.
3. The solution depends continuously on the data.

If one of the conditions is not met, then the problem is considered to be ill-posed. It can be shown that direct problems are well-posed, but inverse problems are ill posed. If in (1.1a), we integrate both sides twice and solve for \( a(x) \), we obtain

\[
a(x) = \frac{1}{u''} \int \int_{\Omega} f(x).
\]

Clearly, the coefficient is not defined in the regions where \( u'' = 0 \).

### 1.2 Optimization Formulation

The coefficient \( a(x) \) in (3.1) can be obtained by solving the BVP for the coefficient. However, it has been shown by explicit examples for the simpler BVPs that this approach is not very fruitful. Therefore, instead of solving for \( a(x) \) directly, a commonly adopted approach for solving inverse problems is to pose an equivalent optimization problem whose solution is an approximation of the coefficient to be identified. The idea is to minimize the norm of the difference between the computed solution and the measured solution. Several variants of this idea are available.
In the context of our problem, one possibility is that given a measurement $z$, the coefficient $a(x)$ should be chosen to minimize the output least-squares (OLS) functional
\[
J(a) = \frac{1}{2} \| u(a) - z \|^2, \tag{1.2}
\]
where $\| \cdot \|$ is a suitable norm and $u(x)$ is the solution obtained by solving the direct problem for the coefficient $a(x)$.

A common remedy to the ill-posedness is by using regularization. That is, instead of the OLS, we indeed solve the minimization problem
\[
\min_{a \in \tilde{A}} \{ J(a) + \varepsilon R(a) \}, \tag{1.3}
\]
where $R(a)$ is called a regularization term, $\varepsilon > 0$ is the regularization parameter, and $\tilde{A}$ is the set of all feasible coefficients.

The variational inequality of finding $a^* \in \tilde{A}$
\[
\langle J'(a^*) + R'(a), a - a^* \rangle \geq 0 \quad \forall a \in \tilde{A}, \tag{1.4}
\]
then turns out to be the necessary optimality condition for the OLS optimization problem.

Other optimization formulations will be discussed in the following chapters.

1.3 Organization

The contents of this thesis are organized into seven chapters. In Chapter 2 we discuss the fourth-order boundary value problem into significant details. We describe the suitable function spaces in which the functions and the solution should reside. We derive the weak form associated to the boundary value problem. We develop a complete finite element based framework for obtaining a numerical solution. Implementation issues are also discussed in detail. Numerical examples are given to show the efficiency of the developed numerical approach. In Chapter 3 we pose the inverse problems as a minimization problem. Several results concerning the differentiability of the coefficient-to-solution map and the objective functional are given. In Chapter 4 we solve the inverse problem by using the Nelder-Mead method, an iterative derivative-free method. In order to use derivative based methods, the adjoint stiffness matrix based approach is developed in Chapter 5 for the numerical computation of the gradient of the objective functional. Chapter 6 employs several proximal point methods for the inverse problem. Each method is tested by means of several examples. A comparative study of the proximal point methods in terms of the over all computational cost involved and their numerical efficiency is given in Chapter 7.
Chapter 2

Fourth-Order Boundary Value Problems

In this chapter we describe the details of the fourth-order boundary value problem that is the theme of this thesis. We obtain the weak form of the BVP. The finite element method is used to discretize the weak form. Implementation issues are discussed. Several examples are given to show the performance of the finite element method.

2.1 The Weak Form

As already mentioned, our primary objective is to study the inverse problem of the identification of variable coefficients in the beam equation. To be specific, we focus on the following static fourth-order boundary value problem in $\Omega = (0,1)$:

\[
\begin{align*}
(a(x)u'')'' &= f(x), \quad 0 < x < 1, \\
u(0) &= u'(0) = 0, \\
u(1) &= u'(1) = 0,
\end{align*}
\]

where $a(x)$ is a variable coefficient and $f$ is a suitable function.

The function space convenient for our setting is then defined by:

\[
V = \{ u \in H^2(\Omega) | u(0) = u'(0) = u(1) = u'(1) = 0 \}.
\]

To obtain the weak form of (2.1), we multiply (2.1a) by a test function $v \in V$,

\[
(a(x)u'')'' v = f v.
\]
Integrating over the interval $(0, 1)$ yields
\[ \int_0^1 (a(x)u'')' \, v \, dx = \int_0^1 f \, v \, dx. \]

By performing integration by parts twice and using the boundary conditions, we obtain
\[ \int_0^1 a(x)u''v'' \, dx = \int_0^1 f \, v \, dx. \]

Since \(v\) is arbitrary, we have obtained
\[ \langle a(x)u'', v'' \rangle = \langle f, v \rangle, \quad \text{for all } v \in V, \tag{2.2} \]
where
\[ \langle f, g \rangle = \int_0^1 f \, g \, dx. \]

Equation (2.2) is often referred to as the weak form or the variational form of (2.1).

### 2.2 The Finite Element Method

To define the finite element space, we start with a partition of \(\Omega\):
\[ 0 = x_0 < x_1 < \cdots < x_j < \cdots < x_N < x_{N+1} = 1. \]

We then define
\[ I_j = [x_{j-1}, x_j], \quad \text{for } j = 1, \ldots, N + 1 \]
and the corresponding step-length
\[ h_j = x_j - x_{j-1}. \]

The finite element space \(V_h\) consists of elements \(v\) that satisfy the following three conditions:

- \(v\) and \(v'\) are continuous.
- \(v\) is a polynomial of degree 3 on each subinterval \(I_j\).
- The boundary conditions (2.1b) and (2.1b) hold for \(v\).

Since \(V_h \subset V\), we can consider the following finite-dimensional weak form: Find \(u_h \in V_h\) such that
\[ \langle a(x)u_h'', v'' \rangle = \langle f, v \rangle, \quad \text{for all } v \in V_h. \tag{2.3} \]
2.2. The Finite Element Method

Since a three degree polynomial has four degrees of freedom, an element \( v \in V_h \) on any interval \( I_j \) can be uniquely determined by four values, namely, \( v(x_{j-1}) \), \( v(x_j) \), \( v'(x_{j-1}) \) and \( v'(x_j) \). Therefore, at every point of the mesh, any \( v \in V_h \) has two degrees of freedom, namely, the function value \( v \) and its derivative value \( v' \). To define a bases for \( V_h \) we will define two basis functions for every node.

Let \( \phi_{j-1} \) be the basic function that corresponds to \( v(x_{j-1}) \) satisfying:

\[
\begin{align*}
  v(x_{j-1}) &= 1, \\
  v(x_j) &= 0, \\
  v'(x_{j-1}) &= 0, \\
  v'(x_j) &= 0.
\end{align*}
\]

(2.4)

Since \( \phi_{j-1} \in V_h \), it is a third degree polynomial, and hence its general form is,

\[
\phi_{j-1}(x) = ax^3 + bx^2 + cx + d, \quad \text{where } 0 \neq a, b, c, d \in \mathbb{R}.
\]

Using the constraints (2.4), we obtain the following system of equations:

\[
\begin{pmatrix}
  \phi_{j-1}(x_{j-1}) \\
  \phi_{j-1}(x_j) \\
  \phi'_{j-1}(x_{j-1}) \\
  \phi'_{j-1}(x_j)
\end{pmatrix} =
\begin{pmatrix}
  ax^3_{j-1} + bx^2_{j-1} + cx_{j-1} + d \\
  ax^3_j + bx^2_j + cx_j + d \\
  3ax^2_{j-1} + bx_{j-1} + c \\
  3ax^2_j + bx_j + c
\end{pmatrix} =
\begin{pmatrix}
  1 \\
  0 \\
  0 \\
  0
\end{pmatrix}.
\]

The above system can be written as

\[
\begin{pmatrix}
  x^3_{j-1} & x^2_{j-1} & x_{j-1} & 1 \\
  x^3_j & x^2_j & x_j & 1 \\
  3x^2_{j-1} & x_{j-1} & 1 & 0 \\
  3x^2_j & x_j & 1 & 0
\end{pmatrix}
\begin{pmatrix}
  a \\
  b \\
  c \\
  d
\end{pmatrix} =
\begin{pmatrix}
  1 \\
  0 \\
  0 \\
  0
\end{pmatrix}.
\]

By solving this system, we get

\[
\begin{pmatrix}
  a \\
  b \\
  c \\
  d
\end{pmatrix} = \frac{1}{h^2_j}
\begin{pmatrix}
  2 \\
  -3(x_{j-1} + x_j) \\
  -6x_{j-1}x_j \\
  (x_j - 3x_{j-1})x^2_j
\end{pmatrix},
\]

where \( h_j = x_j - x_{j-1} \).
2.2. The Finite Element Method

Consequently,

\[
\phi_{j-1}(x) = \frac{1}{h_j^3} \left[ 2x^3 - 3(x_{j-1} + x_j)x^2 + 6x_{j-1}x_jx + (x_j - 3x_{j-1})x_j^2 \right] \quad x \in I_j. \tag{2.5}
\]

To get a general idea about the shape of the basis function, see Figure 2.1.

![Figure 2.1: Basis Function \(\phi_j\).](image)

Indeed this equation will represent the part of the basic function from \(x_j\) to \(x_{j+1}\) due to the conditions in (2.4). To get the part of the function from \(x_{j-1}\) to \(x_j\), we will use the conditions:

\[
\begin{align*}
  v(x_{j-1}) &= 0, \\
  v(x_j) &= 1, \\
  v'(x_{j-1}) &= 0, \\
  v'(x_j) &= 0. \tag{2.6}
\end{align*}
\]

We obtain the following resulting system of equations

\[
\phi_{j-1}(x) = \frac{1}{h_j^3} \left[ -2x^3 + 3(x_{j-1} + x_j)x^2 - 6x_{j-1}x_jx + (3x_j - x_{j-1})x_j^2 \right] \quad x \in I_j. \tag{2.7}
\]

Using both the \(\phi_{j-1}\) and \(\phi_j\), a description for \(\phi_j\) can be given for all \(x \in [0, 1]\) and \(j = 1, \ldots, N\)

\[
\phi_j(x) = \begin{cases} 
  \frac{1}{h_j^3} \left[ -2x^3 + 3(x_{j-1} + x_j)x^2 - 6x_{j-1}x_jx + (3x_j - x_{j-1})x_j^2 \right] & x \in I_j \\
  \frac{1}{h_{j+1}^3} \left[ 2x^3 - 3(x_j + x_{j+1})x^2 + 6x_jx_{j+1}x + (x_{j+1} - 3x_j)x_{j+1}^2 \right] & x \in I_{j+1} \\
  0 & \text{otherwise.}
\end{cases}
\]
For the later use, we also compute the second derivative of $\phi_j$ for $j = 1, \ldots, N$

\[
\phi_j''(x) = \begin{cases}
\frac{1}{h_j^3} [-12x + 6(x_{j-1} + x_j)] & x \in I_j \\
\frac{1}{h_{j+1}} [12x - 6(x_j + x_{j+1})] & x \in I_{j+1} \\
0 & \text{otherwise}
\end{cases}
\] (2.8)

Let the basis function that corresponds to $v'(x_{j-1})$ be $\psi_j$. Using the conditions from (2.9) and (2.11), we can expect to get a picture like shown in Figure 2.2.

![Figure 2.2: Basis Function $\psi_j$](image)

We will need to use the conditions from (2.9):

\[
v(x_{j-1}) = 0, \\
v(x_j) = 0, \\
v'(x_{j-1}) = 1, \\
v'(x_j) = 0.
\] (2.9)

The resulting system of equations reads:

\[
\begin{pmatrix}
\psi_{j-1}(x_{j-1}) \\
\psi_{j-1}(x_j) \\
\psi'_{j-1}(x_{j-1}) \\
\psi'_{j-1}(x_j)
\end{pmatrix} =
\begin{pmatrix}
ax_{j-1}^3 + bx_{j-1}^2 + cx_{j-1} + d \\
ax_j^3 + bx_j^2 + cx_j + d \\
3ax_{j-1}^2 + bx_{j-1} + c \\
3ax_j^2 + bx_j + c
\end{pmatrix} =
\begin{pmatrix}
0 \\
0 \\
1 \\
0
\end{pmatrix}.
\]
Therefore,

\[
\begin{pmatrix}
x_{j-1}^3 & x_{j-1}^2 & x_{j-1} & 1 \\
x_j^3 & x_j^2 & x_j & 1 \\
3x_{j-1}^2 & x_{j-1} & 1 & 0 \\
3x_j^2 & x_j & 1 & 0
\end{pmatrix}
\begin{pmatrix}
a \\ b \\ c \\ d
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
1 \\
0
\end{pmatrix}.
\]

Solving the above system, we get

\[
\begin{pmatrix}
a \\ b \\ c \\ d
\end{pmatrix}
= \frac{1}{h_j^2}
\begin{pmatrix}
1 \\ -(x_{j-1} + x_j) \\ x_j(2x_{j-1} + x_j) \\ -x_{j-1}x_j^2
\end{pmatrix}
\]

where \( h_j = x_j - x_{j-1} \).

Consequently

\[
\psi_{j-1}(x) = \frac{1}{h_j^2} \left[ x^3 - (x_{j-1} + 2x_j)x^2 + x_j(2x_{j-1} + x_j)x - x_{j-1}x_j^2 \right] \text{ for } x \in I_j. \tag{2.10}
\]

This equation will represent the part of the basic function in Figure 2.2, from \( x_j \) to \( x_{j+1} \) due to the conditions in (2.9). To get the part of the function from \( x_{j-1} \) to \( x_j \) in Figure 2.2, we will use the conditions:

\[
\begin{align*}
v(x_{j-1}) &= 0, \\
v(x_j) &= 0, \\
v'(x_{j-1}) &= 0, \\
v'(x_j) &= 1.
\end{align*}
\tag{2.11}
\]

This leads to

\[
\psi_j(x) = \frac{1}{h_j^2} \left[ x^3 - (2x_{j-1} + x_j)x^2 + x_{j-1}(x_{j-1} + 2x_j)x - x_{j-1}x_j^2 \right] \text{ for } x \in I_j \tag{2.12}
\]

Using both the \( \psi_{j-1} \) and \( \psi_j \), a description for \( \psi_j \) can be given for all \( x \in [0, 1] \) and \( j = 1, \ldots, N \)

\[
\psi_j(x) = \begin{cases}
\frac{1}{h_j^2} \left[ x^3 - (2x_{j-1} + x_j)x^2 + x_{j-1}(x_{j-1} + 2x_j)x - x_{j-1}x_j^2 \right] & x \in I_j \\
\frac{1}{h_{j+1}^2} \left[ x^3 - (x_j + 2x_{j+1})x^2 + x_{j+1}(2x_j + x_{j+1})x - x_jx_{j+1}^2 \right] & x \in I_{j+1} \\
0 & \text{otherwise.}
\end{cases}
\]
The second derivative will be of importance to me later, so we have derive the second derivative of $\psi_j$ for $j = 1, \ldots, N$

$$
\psi''_j(x) = \begin{cases} 
\frac{1}{h_j} \left[ 6x - 2(2x_{j-1} + x_j) \right] & x \in I_j \\
\frac{1}{h_{j+1}} \left[ 6x - 2(x_j + 2x_{j+1}) \right] & x \in I_{j+1} \\
0 & \text{otherwise.}
\end{cases} 
$$

We have now constructed a set of basis functions $\{\phi_1, \ldots, \phi_N, \psi_1, \ldots, \psi_N\}$ or $\{\phi_j, \psi_j\}_{j}^{N}$ for $V_h$.

By the definition of $V_h$, any element $v \in V_h$ can be uniquely written as:

$$
v = \sum_{j=1}^{N} [v_j \phi_j + \hat{v}_j \psi_j], 
$$

where $v_j = v(x_j)$ and $\hat{v}_j = v'(x_j)$.

Let $u_h \in V_h$ be the solution of the finite-dimensional weak form. Using (2.14), we obtain

$$
u_h = \sum_{j=1}^{N} [u_j \phi_j + \hat{u}_j \psi_j], 
$$

To obtain a matrix form for the weak form we proceed as follows: By setting $v = \phi_j$ in (2.2), we get

$$
\langle a(x)u''_j, \phi''_j \rangle = \langle f, \phi_j \rangle,
$$

which further implies that

$$
\left\langle a(x) \sum_{j=1}^{N} [u_j \phi''_j + \hat{u}_j \psi''_j], \phi''_j \right\rangle = \langle f, \phi_j \rangle.
$$

By using the linearity of the inner product, we obtain

$$
\sum_{j=1}^{N} u_j \langle a(x)\phi''_j, \phi''_j \rangle + \sum_{j=1}^{N} \hat{u}_j \langle a(x)\psi''_j, \phi''_j \rangle = \langle f, \phi_j \rangle. \tag{2.16}
$$

Analogously, by taking $v = \psi_j$, we obtain

$$
\sum_{j=1}^{N} u_j \langle a\phi''_j, \psi''_j \rangle + \sum_{j=1}^{N} \hat{u}_j \langle a\psi''_j, \psi''_j \rangle = \langle f, \psi_j \rangle. \tag{2.17}
$$

Equations (2.16) and (2.17) now give me a system of $2N$ equations in $2N$ unknowns, namely $\{u_1, \ldots, u_N, \hat{u}_1, \ldots, \hat{u}_N\}$. 

We will now write this system in terms of a matrix equation. For \( i, j = 1, \ldots, N \), we define \( N \times N \) matrices \( A, B, \) and \( C \) by

\[
A_{ij} = \langle a(x) \phi_j'', \phi_i'' \rangle, \\
B_{ij} = \langle a(x) \psi_j'', \phi_i'' \rangle, \\
C_{ij} = \langle a(x) \psi_j'', \psi_i'' \rangle.
\]

Furthermore, defining the vectors \( F_\phi, F_\psi \in \mathbb{R}^N \) by

\[
F_\phi = (\langle f, \phi_1 \rangle, \ldots, \langle f, \phi_N \rangle)^T, \\
F_\psi = (\langle f, \psi_1 \rangle, \ldots, \langle f, \psi_N \rangle)^T.
\]

Then \( U \in \mathbb{R}^{2N} \) is given by

\[
U = (u_1, \ldots, u_N, \hat{u}_1, \ldots, \hat{u}_N)^T.
\]

From equations (2.16) and (2.17), we obtain the system,

\[
KU = F
\]

where the stiffness matrix \( K \) has the form

\[
K = \begin{pmatrix}
A & B \\
B^T & C
\end{pmatrix}
\]

and the load vector \( F \) is given by

\[
F = \begin{pmatrix}
F_\phi \\
F_\psi
\end{pmatrix}.
\]

Notice that when \( |i - j| > 1 \) then from the definition of \( \phi \) and \( \psi \),

\[
\phi_j'' \phi_j'' = \psi_j'' \psi_j'' = \phi_j'' \psi_j'' = 0 \text{ when } |i = j| > 1.
\]

Therefore this leads to,

\[
\langle a\phi_j'', \phi_j'' \rangle = \langle a\psi_j'', \psi_j'' \rangle = \langle a\phi_j'', \psi_j'' \rangle = 0 \text{ when } |i = j| > 1.
\]

From this, it can be shown that the \( A, B, \) and \( C \) are tridiagonal.
2.3 Implementation Issues

Knowing how to write the fourth order boundary value problem as system of equations, we now switch to the implementation aspects. First task is to numerically approximate the integrals. We will do this by using Simpson’s rule. When integrating over the interval \( I_j = [x_{j-1}, x_j] \), Simpson’s rule states,

\[
\int_{x_{j-1}}^{x_j} f(x)dx = \frac{h_j}{6} [f(x_{j-1}) + 4f(x_{j-1}/2) + f(x_j)] \quad (2.18)
\]

where \( x_{j-1/2} \) is the midpoint of the interval \( I_j \). Similarly, on \( I_{j+1} \),

\[
\int_{x_j}^{x_{j+1}} f(x)dx = \frac{h_{j+1}}{6} [f(x_j) + 4f(x_{j+1}/2) + f(x_{j+1})] \quad (2.19)
\]

where \( x_{j+1/2} \) is the midpoint of the interval \( I_{j+1} \).

2.3.1 Load Vector

We now construct the load vector, \( F = (F_\phi, F_\psi)^T \) where \( F_\phi = (f, \phi_j)_{j=1,...,N} \). The \( j \)th element of \( F_\phi \) can be expressed as,

\[
\int_0^1 f \phi_j dx = \int_{x_{j-1}}^{x_{j+1}} f \phi_j dx + \int_{x_{j-1}}^{x_j} f \phi_j dx + \int_{x_j}^{x_{j+1}} f \phi_j dx.
\]

This is justified since anything outside of the range of \( I_j \) and \( I_{j+1} \) will be zero. Using the Simpson’s rule from (2.18)

\[
\int_{x_{j-1}}^{x_j} f \phi_j dx = \frac{h_j}{6} [f(x_{j-1})\phi_j(x_{j-1}) + 4f(x_{j-1}/2)\phi_j(x_{j-1}/2) + f(x_j)\phi_j(x_j)].
\]

We know that \( \phi_j(x_{j-1}) = 0 \) and \( \phi_j(x_j) = 1 \). It can be shown that \( \phi_j(x_{j-1}/2) = \phi_j(x_{j+1}/2) = 1/2 \).

Therefore, the above reduces to

\[
\int_{x_{j-1}}^{x_j} f \phi_j dx = \frac{h_j}{6} [2f(x_{j-1}/2) + f(x_j)].
\]

Similarly, the integral on \((x_j, x_{j+1})\) simplifies to

\[
\int_{x_j}^{x_{j+1}} f \phi_j dx = \frac{h_{j+1}}{6} [f(x_j) + 2f(x_{j+1}/2)].
\]
Therefore the integral for the load vector can be given by

\[
\int_0^1 f \phi_j dx = \frac{h_j}{6} \left[ 2f(x_{j-1/2}) + f(x_j) \right] + \frac{h_j}{6} \left[ f(x_j) + 2f(x_{j+1/2}) \right] = \frac{1}{3} \left[ h_j f(x_{j-1/2}) + \frac{1}{2} (h_j + h_{j+1}) f(x_j) + h_{j+1} f(x_{j+1/2}) \right].
\]

If \( N \) is large, then we can approximate \( f(x_{j-1/2}) \) and \( f(x_{j+1/2}) \) by \( f(x_j) \). In this case, we will get,

\[
\langle f, \phi_j \rangle = \frac{h_j + h_{j+1}}{2} f(x_j), \quad j = 1, \ldots, N. \tag{2.20}
\]

For \( F \psi \), we follow the same procedure to find \( \langle f, \psi_j \rangle \).

\[
\int_{x_{j-1}}^{x_j} f \psi_j dx = \frac{h_j}{6} \left[ f(x_{j-1}) \psi_j(x_{j-1}) + 4f(x_{j-1/2}) \psi_j(x_{j-1/2}) + f(x_j) \psi_j(x_j) \right]. \tag{2.21}
\]

We know that \( \psi_j(x_{j-1}) = 0 \) and \( \psi_j(x_j) = 0 \). It can be shown that \( \phi_j(x_{j-1/2}) = -h_j/8 \), and \( \phi_j(x_{j+1/2}) = h_j/8 \).

\[
\psi_j(x_{j-1/2}) = \frac{1}{h_j^2} \left[ x_j^3 - (2x_{j-1} + x_j)x_j^2 + x_j(x_{j-1} + 2x_j)x_{j-1} - x_{j-1}x_j \right]
\]

\[
= \frac{1}{h_j^2} \left[ \left(x_j - \frac{h_j}{2}\right)^3 - (2x_{j-1} + x_j)\left(x_j - \frac{h_j}{2}\right)^2 + x_{j-1}(x_{j-1} + 2x_j)\left(x_j - \frac{h_j}{2}\right) - x_{j-1}x_j \right]
\]

\[
= \frac{1}{h_j^2} \left[ \left(x_j - \frac{h_j}{2}\right)^3 - (3x_j - 2h_j)\left(x_j - \frac{h_j}{2}\right)^2 + (x_j + h_j)(3x_j - h_j)\left(x_j - \frac{h_j}{2}\right) - (x_j - h_j)^2x_j \right]
\]

\[
= \frac{1}{h_j^2} \left[ \frac{1}{8}h_j^3 \right]
\]

\[
\psi_j(x_{j+1/2}) = \frac{1}{h_j^2} \left[ x_{j+1}^3 - (x_j + 2x_{j+1})x_{j+1}^2 + x_{j+1}(2x_j + x_{j+1})x_{j+1} - x_jx_{j+1}^2 \right]
\]

\[
= \frac{1}{h_j^2} \left[ \left(x_j + \frac{h_j}{2}\right)^3 - (x_j + 2x_{j+1})\left(x_j + \frac{h_j}{2}\right)^2 + x_{j+1}(2x_j + x_{j+1})\left(x_j + \frac{h_j}{2}\right) - x_jx_{j+1}^2 \right]
\]

\[
= \frac{1}{h_j^2} \left[ \left(x_j + \frac{h_j}{2}\right)^3 - (3x_j + 2h_j)\left(x_j + \frac{h_j}{2}\right)^2 + (x_j + h_j)(3x_j + h_j)\left(x_j + \frac{h_j}{2}\right) - x_j(x_j + h_j)^2 \right]
\]

\[
= \frac{1}{h_j^2} \left[ \frac{1}{8}h_j^3 \right]
\]

\[
= \frac{h_j}{8}.
\]

2.3. Implementation Issues

13
This integral from (2.21) gets reduced to
\[
\int_{x_{j-1}}^{x_j} f\psi_j \, dx = \frac{h_j}{6} \left[ \frac{h_j}{2} f(x_{j-1/2}) \right] = -\frac{h_j^2}{12} [f(x_{j-1/2})].
\]

Similarly, the integral on \((x_j, x_{j+1})\) simplifies to
\[
\int_{x_j}^{x_{j+1}} f\psi_j \, dx = \frac{h_{j+1}}{6} \left[ \frac{h_j}{2} f(x_{j-1/2}) \right] = \frac{h_{j+1}h_j}{12} [f(x_{j+1/2})].
\]

Therefore the integral for the load vector can be given by
\[
\int_{0}^{1} f\psi_j \, dx = -\frac{h_j^2}{12} [f(x_{j-1/2})] + \frac{h_{j+1}h_j}{12} [f(x_{j+1/2})]
= \frac{h_j}{12} [-h_jf(x_{j-1/2}) + h_{j+1}f(x_{j+1/2})].
\]

### 2.3.2 Stiffness Matrix

For the computation of the stiffness matrix, we will set the interval size \(I_j\), to be the same length. That is, \(h_j = h\) for all \(j = 1, \cdots, N + 1\). This causes the (2.8) and (2.13) to become

\[
\phi_j''(x) = \begin{cases} 
\frac{1}{h^3} \left[ -12x + 12x_j - 6h \right] & x \in I_j \\
\frac{1}{h^3} \left[ 12x - 12x_j - 6h \right] & x \in I_{j+1} \\
0 & \text{otherwise}
\end{cases} \tag{2.22}
\]

and

\[
\psi_j''(x) = \begin{cases} 
\frac{2}{h^2} \left[ 3x - 3x_j + 2h \right] & x \in I_j \\
\frac{2}{h^2} \left[ 3x - 3x_j - 2h \right] & x \in I_{j+1} \\
0 & \text{otherwise}
\end{cases} \tag{2.23}
\]
2.3. Implementation Issues

We will be using Simpson’s method to approximate the integrals, and therefore it is necessary to compute the values of $φ''_j$ and $ψ''_j$ at the points $x_{j-1}, x_{j-1/2}, x_j, x_{j+1/2}, x_{j+1}$.

\[
\begin{align*}
φ''_j(x_{j-1}) &= \frac{6}{h^2}, & φ''_j(x_{j+1}) &= \frac{6}{h^2} \\
φ''_j(x_{j-1/2}) &= 0, & φ''_j(x_{j+1/2}) &= 0 \\
φ''_j(x_j) &= -\frac{6}{h^2}, & ψ''_j(x_{j-1}) &= -\frac{2}{h}, & (2.24) \\
ψ''_j(x_{j+1}) &= \frac{2}{h}, & ψ''_j(x_{j-1/2}) &= \frac{1}{h}, \\
ψ''_j(x_{j+1/2}) &= -\frac{1}{h}, & \lim_{x \to x_j^-} ψ''_j(x_j) &= \frac{4}{h}, \\
\lim_{x \to x_j^+} ψ''_j(x_j) &= -\frac{4}{h}. & (2.25)
\end{align*}
\]

We have

\[
\begin{align*}
A_{j,j} &= \langle aφ''_j, φ''_j \rangle \\
&= \int_{x_{j-1}}^{x_{j+1}} a(φ''_j)^2 \, dx \\
&= \int_{x_{j-1}}^{x_j} a(φ''_j)^2 \, dx + \int_{x_j}^{x_{j+1}} a(φ''_j)^2 \, dx.
\end{align*}
\]

Since

\[
\begin{align*}
\int_{x_{j-1}}^{x_j} a(φ''_j)^2 \, dx &\approx \frac{h}{6} \left[ a_{j-1} \left( φ''_j(x_{j-1}) \right)^2 + 4a_{j-1/2} \left( φ''_j(x_{j-1/2}) \right)^2 + a_j \left( φ''_j(x_j) \right)^2 \right] \\
&= \frac{h}{6} \left[ a_{j-1} \frac{36}{h^4} + 4a_{j-1/2}(0) + a_j \right] \\
&= \frac{6}{h^3} \left[ a_{j-1} + a_j \right] \\
\int_{x_j}^{x_{j+1}} a(φ''_j)^2 \, dx &\approx \frac{h}{6} \left[ a_j \left( φ''_j(x_j) \right)^2 + 4a_{j+1/2} \left( φ''_j(x_{j+1/2}) \right)^2 + a_{j+1} \left( φ''_j(x_{j+1}) \right)^2 \right] \\
&= \frac{h}{6} \left[ a_j \frac{36}{h^4} + 4a_{j+1/2}(0) + a_{j+1} \right] \\
&= \frac{6}{h^3} \left[ a_j + a_{j+1} \right]
\end{align*}
\]
we obtain

\[
A_{j,j} = \langle a\phi''_j, \phi''_j \rangle \\
= \int_{x_{j-1}}^{x_{j+1}} a (\phi''_j)^2 \, dx \\
= \frac{6}{h^3} [a_j - 1 + a_j] + \frac{6}{h^3} [a_j + a_{j+1}] \\
= \frac{6}{h^3} [a_{j-1} + 2a_j + a_{j+1}].
\]

Now for the second diagonal, we have

\[
A_{j-1,j} = \langle a\phi''_{j-1}, \phi''_{j-1} \rangle \\
= \int_{x_{j-1}}^{x_{j}} a\phi''_j\phi''_{j-1} \, dx \\
\approx \frac{h}{6} \left[ a_{j-1}\phi''_j(x_{j-1})\phi''_{j-1}(x_{j-1}) + 4a_{j-1/2}\phi''_j(x_{j-1/2})\phi''_{j-1}(x_{j-1/2}) + \right. \\
\left. a_j\phi''_j(x_j)\phi''_{j-1}(x_j) \right] \\
= \frac{h}{6} \left[ a_{j-1} \frac{-36}{h^4} + 4a_{j-1/2}(0) + a_j \frac{-36}{h^4} \right] \\
= \frac{-6}{h^3} [a_{j-1} + a_j].
\]

We know that the matrix \( A \) is symmetric,

\[
A_{j,j-1} \approx \frac{-6}{h^3} [a_{j-1} + a_j].
\]

Now using (2.24), we can come up with the entries of matrix \( C \) :

\[
C_{j,j} = \langle a\psi''_j, \psi''_j \rangle \\
= \int_{x_{j-1}}^{x_{j+1}} a (\psi''_j)^2 \, dx \\
= \int_{x_{j-1}}^{x_{j}} a (\psi''_j)^2 \, dx + \int_{x_{j}}^{x_{j+1}} a (\psi''_j)^2 \, dx.
\]
2.3. Implementation Issues

Since

\[
\int_{x_{j-1}}^{x_j} a(x) \psi''(x) \, dx \approx \frac{h}{6} \left[ a_{j-1} \left( \psi''(x_{j-1}) \right)^2 + 4a_{j-1/2} \left( \psi''(x_{j-1/2}) \right)^2 + a_j \left( \psi''(x_j) \right)^2 \right]
\]

\[
= \frac{h}{6} \left[ a_{j-1} \frac{4}{h^2} + 4a_{j-1/2} \frac{1}{h^2} + a_j \frac{16}{h^2} \right]
\]

\[
= \frac{2}{3h} \left[ a_{j-1} + a_{j-1/2} + 4a_j \right]
\]

\[
\int_{x_j}^{x_{j+1}} a(x) \psi''(x) \, dx \approx \frac{h}{6} \left[ a_j \left( \psi''(x_j) \right)^2 + 4a_{j+1/2} \left( \psi''(x_{j+1/2}) \right)^2 + a_{j+1} \left( \psi''(x_{j+1}) \right)^2 \right]
\]

\[
= \frac{h}{6} \left[ a_j \frac{16}{h^2} + 4a_{j+1/2} \frac{1}{h^2} + a_{j+1} \frac{4}{h^2} \right]
\]

\[
= \frac{2}{3h} \left[ 4a_j + a_{j+1/2} + a_{j+1} \right]
\]

we deduce

\[
C_{j,j} = \langle a \psi'', \psi'' \rangle \approx \int_{x_{j-1}}^{x_j} a(x) \psi''(x) \, dx
\]

\[
\approx \frac{2}{3h} \left[ a_{j-1} + a_{j-1/2} + 4a_j \right] + \frac{2}{3h} \left[ 4a_j + a_{j+1/2} + a_{j+1} \right]
\]

\[
= \frac{2}{3h} \left[ a_{j-1} + a_{j-1/2} + 8a_j + a_{j+1/2} + a_{j+1} \right].
\]

Notice that the values of \( a_{j-1/2} \) and \( a_{j+1/2} \) are not known to me due to the fact that I’m doing this by computational methods. We shall approximate these values,

\[
a_{j-1/2} = \frac{1}{2} (a_{j-1} + a_j),
\]

\[
a_{j+1/2} = \frac{1}{2} (a_j + a_{j+1}).
\]

With these approximations, we shall get

\[
C_{j,j} = \langle a \psi'', \psi'' \rangle \approx \frac{2}{3h} \left[ a_{j-1} + \frac{1}{2} (a_{j-1} + a_j) + 8a_j + \frac{1}{2} (a_j + a_{j+1}) + a_{j+1} \right]
\]

\[
\approx \frac{2}{3h} \left[ \frac{3}{2} a_{j-1} + 9a_j + \frac{3}{2} a_{j+1} \right]
\]

\[
\approx \frac{1}{h} \left[ a_{j-1} + 6a_j + a_{j+1} \right].
\]
Now to obtain the second diagonal

\[ C_{j-1,j} = \langle a\psi_j'', \psi_{j-1}' \rangle = \int_{x_{j-1}}^{x_j} a\psi_j'' \psi_{j-1}'' \, dx \approx \frac{h}{6} \left[ a_{j-1} \psi_j''(x_{j-1}) \psi_{j-1}''(x_{j-1}) + 4a_{j-1/2} \psi_j''(x_{j-1/2}) \psi_{j-1}''(x_{j-1/2}) + a_j \psi_j''(x_j) \psi_{j-1}''(x_j) \right] \]

\[ = \frac{h}{6} \left[ a_{j-1} \frac{8}{h^2} + 4a_{j-1/2} \frac{-1}{h^2} + a_j \frac{8}{h^2} \right] \]

\[ = \frac{h}{6} \left[ a_{j-1} \frac{8}{h^2} + \frac{1}{2}(a_{j-1} + a_j) \frac{-4}{h^2} + a_j \frac{8}{h^2} \right] \]

\[ = \frac{1}{h} [a_{j-1} + a_j] . \]

We know that the \( C \) matrix is symmetric,

\[ C_{j,j-1} \approx \frac{1}{h} [a_{j-1} + a_j] . \]

Now using the information from (2.24) and (2.25), we can compute the entries of matrix \( B \):

\[ B_{j,j} = \langle a\psi_j'', \phi_j'' \rangle = \int_{x_{j-1}}^{x_j} a\psi_j'' \phi_j'' \, dx \]

\[ = \int_{x_{j-1}}^{x_j} a\psi_j'' \phi_j'' \, dx + \int_{x_j}^{x_{j+1}} a\psi_j'' \phi_j'' \, dx \]

and we know,

\[ \int_{x_{j-1}}^{x_j} a\psi_j'' \phi_j'' \, dx \approx \frac{h}{6} \left[ a_{j-1} \psi_j''(x_{j-1}) \phi_j''(x_{j-1}) + 4a_{j-1/2} \psi_j''(x_{j-1/2}) \phi_j''(x_{j-1/2}) + a_j \psi_j''(x_j) \phi_j''(x_j) \right] \]

\[ = \frac{h}{6} \left[ a_{j-1} \frac{-12}{h^3} + 4a_{j-1/2}(0) + a_j \frac{-24}{h^3} \right] \]

\[ = \frac{2}{h^2} [a_{j-1} - 2a_j] . \]

\[ \int_{x_j}^{x_{j+1}} a\psi_j'' \phi_j'' \, dx \approx \frac{h}{6} \left[ a_j \psi_j''(x_j) \phi_j''(x_j) + 4a_{j+1/2} \psi_j''(x_{j+1/2}) \phi_j''(x_{j+1/2}) + a_{j+1} \psi_j''(x_{j+1}) \phi_j''(x_{j+1}) \right] \]

\[ = \frac{h}{6} \left[ a_j \frac{24}{h^3} + 4a_{j+1/2}(0) + a_{j+1} \frac{12}{h^3} \right] \]

\[ = \frac{2}{h^2} [2a_j + a_{j+1}] . \]
Therefore,

\[ B_{j,j} = \langle a\psi''_j, \phi''_j \rangle \]
\[ \approx \int_{x_{j-1}}^{x_{j+1}} a\psi''_j \phi''_j \, dx \]
\[ = \frac{2}{h^2} [-a_{j-1} - 2a_j] + \frac{2}{h^2} [2a_j + a_{j+1}] \]
\[ = \frac{2}{h^2} [-a_{j-1} + a_{j+1}] \]

\[ B_{j-1,j} = \langle a\psi''_j, \phi''_{j-1} \rangle \]
\[ = \int_{x_{j-1}}^{x_j} a\psi''_j \phi''_{j-1} \, dx \]
\[ \approx \frac{h}{6} [a_{j-1} \psi''_j(x_{j-1}) \phi''_{j-1}(x_{j-1}) + 4a_{j-1/2} \psi''_j(x_{j-1/2}) \phi''_{j-1}(x_{j-1/2}) + a_j \psi''_j(x_j) \phi''_{j-1}(x_j)] \]
\[ = \frac{h}{6} \left[ a_{j-1} \frac{12}{h^3} + 4a_{j-1/2}(0) + a_j \frac{24}{h^3} \right] \]
\[ = \frac{2}{h^2} [a_{j-1} + 2a_j] \]

\[ B_{j,j-1} = \langle a\psi''_{j-1}, \phi''_j \rangle \]
\[ = \int_{x_{j-1}}^{x_j} a\psi''_{j-1} \phi''_j \, dx \]
\[ \approx \frac{h}{6} \left[ a_{j-1} \psi''_{j-1}(x_{j-1}) \phi''_j(x_{j-1}) + 4a_{j-1/2} \psi''_{j-1}(x_{j-1/2}) \phi''_j(x_{j-1/2}) + a_j \psi''_{j-1}(x_j) \phi''_j(x_j) \right] \]
\[ = \frac{h}{6} \left[ a_{j-1} \frac{24}{h^3} + 4a_{j-1/2}(0) + a_j \frac{12}{h^3} \right] \]
\[ = \frac{2}{h^2} [-2a_{j-1} - a_j] . \]

\[ F_{\Phi,j} \approx \frac{h}{12} f(x_j), \quad j = 1, \ldots, N \]
\[ F_{\Psi,j} \approx \frac{h^2}{12} [-f(x_{j-1/2}) + f(x_{j+1/2})], \quad j = 1, \ldots, N \]
\[ A_{j,j} \approx \frac{6}{h^3} [a_{j-1} + 2a_j + a_{j+1}], \quad j = 1, \ldots, N \]
\[ A_{j,j-1} = A_{j-1,j} \approx \frac{-6}{h^3} [a_{j-1} + a_j], \quad j = 2, \ldots, N \]
\[ B_{j,j} \approx \frac{2}{h^2} [-a_{j-1} + a_{j+1}], \quad j = 1, \ldots, N \]
\[ B_{j-1,j} \approx \frac{2}{h^2} [a_{j-1} + 2a_j], \quad j = 2, \ldots, N \]
\[ B_{j,j-1} \approx \frac{2}{h^2} [2a_{j-1} + a_j], \quad j = 2, \ldots, N \]
\[ C_{j,j} \approx \frac{1}{h} [a_{j-1} + 6a_j + a_{j+1}], \quad j = 1, \ldots, N \]
\[ C_{j,j-1} = C_{j-1,j} \approx \frac{1}{h} [a_{j-1} + a(x_j)], \quad j = 2, \ldots, N . \]
2.4 Numerical Experiments

To demonstrate that the above method works properly, we test it for some examples. There are four examples, in each example we provide the function $f$, the coefficient $a(x)$, and the compute for the solution $u$. We also included images comparing the exact solution and the computed solution.

Example 1:

We have the following data:

$$f_1 = 8\pi^2((2\pi^2(x^2 + 1) - 1)\cos(2\pi x) + 4\pi x\sin(2\pi x))$$

$$a(x)_1 = x^2 + 1$$

$$u_1 = \cos(2\pi x) - 1.$$  

Example 2:

We have the following data:

$$f_2 = 4(90x^4 - 60x^3 + 42x^2 - 18x + 13)$$

$$a(x)_2 = x^4 + x^2 + 2$$

$$u_2 = x^2(x - 1)^2.$$
Example 3:

We have the following data:

\[ f_3 = 32\pi^2(-631.655(x+0.400059)(x^2-1.15006x+0.6095)\cos(4\pi x)-301.593(x-0.25)^2\sin(4\pi x)) \]
\[ a(x)_3 = (2x-.5)^3 + 2 \]
\[ u_3 = \sin(4\pi x - 0.5\pi) + 1. \]
Example 4:

We have the following data:

\[ f_4 = 1764(x - 0.773513)(x - 0.5)(x - 0.226487)(x^2 - x + 0.687621) \]
\[ a(x)_4 = (x - 0.5)^2 + 1 \]
\[ u_4 = x^2(x - 1)^2(x - 0.5)^3. \]

Figure 2.6: Example 4
Chapter 3

Optimization Formulation of Inverse Problems

This chapter investigates the inverse problem as an optimization problem and studies the basic features of the objective functionals. Computable forms of the objective functions are studied in detail.

3.1 Problem Formulation

Recall that we have been dealing with the following static fourth-order boundary value problem in $\Omega = (0, 1)$:

\begin{align*}
(a(x)u'')'' &= f(x), 0 < x < 1, \\
u(0) &= u'(0) = 0, \\
u(1) &= u'(1) = 0,
\end{align*}

where $a(x)$ is a variable coefficient and $f$ is a suitable function.

The weak form for the above boundary problem reads: Find $u(x) \in V := H_0^2(0, 1)$ such that

\begin{align*}
a(u, v) = \langle f, v \rangle \quad \text{for all} \quad v \in V,
\end{align*}

where

\begin{align*}
a(u, v) &= \int_0^1 a(x)u''v'' \, dx, \\
\langle f, v \rangle &= \int_0^1 f(x)v(x) \, dx.
\end{align*}
3.1. Problem Formulation

We assume that the coefficient $a(x)$, in equation (3.2) is strictly positive and constrained between two constants $k_0, k_1$. That is the set of all admissible coefficients is defined as

$$ A := \{a(x) \mid 0 \leq a_0 \leq a(x) \leq a_1 \}. \quad (3.3) $$

Since in the analysis the exact role of the coefficient is of vital importance, it is convenient to define a trilinear form $T : A \times V \times V \to \mathbb{R}$ by

$$ T(a, u, v) = \int_0^1 a(x)u''v'' \, dx. $$

The inner product in $V$ is given by:

$$ \langle u, v \rangle = \int_0^1 (uv + u'v' + u''v'') \, dx. $$

The corresponding norm is then defined by

$$ \|u\|_V^2 = \langle u, v \rangle = \int_0^1 (u^2 + (u')^2 + (u'')^2) \, dx. $$

The trilinear form $T$ satisfies the following conditions for all $u, v \in V$, and $a \in \tilde{A}$

$$ T(a, u, v) \leq \|a\|_{\infty} \|u\|_V \|v\|_V \quad (3.4a) $$

$$ T(a, u, u) \geq \alpha \|u\|_V^2 \quad \text{where } \alpha > 0 \quad (3.4b) $$

Therefore, we have the variational form: Find $u \in V$ such that

$$ T(a, u, v) = m(v) \quad \text{for all } v \in V, \quad (3.5) $$

where

$$ m(v) = \int_0^1 f(x)v(x) \, dx. $$

The Lax-Milgram lemma, in view of the above mentioned properties of the trilinear form, ensures that for every $a \in A$, the variation problem (3.5) is uniquely solvable. We can therefore define the coefficient-to-solution map $F : A \to V$ by the condition that $u = F(a)$ is the unique solution of the variational problem.

We recall the following results for the better understanding of the variational form. More details on the following theoretical results can be found in [14] (see also [12], [13], [15], [16], [17] and the cited references therein).

**Lemma 3.1.1.** For each $a \in A$, $u = F(a)$ satisfies $\|u\|_V \leq \alpha^{-1} \|m\|_V$. 

3.1. Problem Formulation

**Theorem 3.1.1.** The operator $F$ is continuous an
\[
\|F(a) - F(b)\|_V \leq \frac{1}{\alpha} \min\{\|F(a)\|_V, \|F(b)\|_V\} \|b - a\|_\infty \text{ for all } a, b \in A. \tag{3.6}
\]

**Proof.** Let $u = F(a)$ and $w = F(b)$, then
\[
T(a, u, v) = m(v) = T(b, w, v), \quad \forall v \in V.
\]
Substituting $v = u - w$ and manipulating yields
\[
T(a, u - w, u - w) = -T(a - b, w, u - w).
\]
Then applying conditions (3.4a) and (3.4b), we obtain
\[
\alpha \|u - w\|^2_V \leq \|a - b\|_\infty \|w\|_V \|u - w\|_V
\]
and hence
\[
\|u - w\|_V \leq \frac{1}{\alpha} \|w\|_V \|a - b\|_\infty = \frac{1}{\alpha} \|F(b)\|_V \|a - b\|_\infty.
\]
We can change the roles of $a$ and $b$, and by bounding $\|F(b)\|_V$ to receive the other minimum bound. \hfill \Box

The following important result discusses the differentiability of $F(a)$.

**Theorem 3.1.2.** For each $a \in A$ $F$ is differentiable at $a$, then $\delta u = DF(a)\delta a$ is an unique solution to the variational equation
\[
T(a, \delta u, v) = -T(\delta a, u, v) \text{ for all } v \in V, \tag{3.7}
\]
where $u = F(a)$.

**Proof.** We know that $u = F(a)$ is the solution of
\[
T(a, u, v) = m(v) \text{ for all } v \in V. \tag{3.8}
\]
Because of the linearity We can say
\[
T(a + \delta a, u + \delta w, v) = m(v) \text{ for all } v \in V \tag{3.9}
\]
Now using equations (3.8) and (3.9),
\[
T(a, u, v) = T(a + \delta a, u + \delta w, v)
\]
\[
T(a, u, v) = T(a + \delta a, u, v) + T(a + \delta a, \delta w, v)
\]
\[
T(a, u, v) = T(a, u, v) + T(\delta a, u, v) + T(a + \delta a, \delta w, v).
\]
Therefore,
\[
0 = -T(a, \delta u, v) + T(a + \delta a, \delta w, v) \\
0 = -T(a, \delta u, v) + T(a, \delta w, v) + T(\delta a, \delta w, v) \\
0 = T(a, \delta w - \delta u, v) + T(\delta a, \delta w, v).
\]

This is true for all \(v \in V\), therefore we can choose \(v = \delta w - \delta u\) and obtain
\[
0 = T(a, \delta w - \delta u, \delta w - \delta u) + T(\delta a, \delta w, v).
\]

From (3.4), this implies that
\[
\alpha \|\delta w - \delta u\|_V^2 \leq T(a, \delta w - \delta u, \delta w - \delta u) \\
T(a, \delta w - \delta u, \delta w - \delta u) \leq \|\delta a\|_{\infty} \|\delta w\|_V \|\delta w - \delta u\|_V \\
\alpha \|\delta w - \delta u\|_V^2 \leq \|\delta a\|_{\infty} \|\delta w\|_V \|\delta w - \delta u\|_V \\
\alpha \|\delta w - \delta u\|_V \leq \|\delta a\|_{\infty} \|\delta w\|_V \\
\|\delta w - \delta u\|_V \leq \frac{1}{\alpha} \|\delta a\|_{\infty} \|\delta w\|_V. 
\] (3.10)

This is where we can use the definition of \(\delta w\) and equation (3.6)
\[
\|\delta w\|_V = \|F(a + \delta a) - F(a)\|_V \\
\leq \frac{1}{\alpha} \|F(a)\|_V \|(a + \delta a) - a\|_{\infty} \\
= \frac{1}{\alpha} \|F(a)\|_V \|\delta a\|_{\infty}.
\]

We can now use this result with equation (3.10),
\[
\frac{\|\delta w - \delta u\|_V}{\|\delta a\|_{\infty}} \leq \frac{1}{\alpha^2} \|F(a)\|_V \|\delta a\|_{\infty}^2 \\
\frac{\|\delta w - \delta u\|_V}{\|\delta a\|_{\infty}} \leq \frac{1}{\alpha^2} \|F(a)\|_V \|\delta a\|_{\infty} \\
\frac{\|F(a + \delta a) - F(a) - \delta u\|_V}{\|\delta a\|_{\infty}} \leq \frac{1}{\alpha^2} \|F(a)\|_V \|\delta a\|_{\infty} \\
= O(\|\delta a\|_{\infty}).
\]

Therefore, \(F\) is differentiable at \(a\) and \(DF(a) \delta a = \delta u\)
\[\square\]
3.2 Optimization Formulation: Finite Dimensional Case

We have already discussed the discretized space $V_h$ of the space $V$. Let $A_h$ be the finite dimensional subspace of $B$. We consider

$$a \in \mathcal{A} = \{ A_h : 0 < a_1 \leq a \leq a_2 \},$$

which is set of admissible coefficients.

The finite dimensional optimization formulation for the inverse problem of identifying $a(x)$ then takes the form:

$$\min_{a \in \mathcal{A}} J(a) + \varepsilon R(a) \tag{3.11}$$

where $J : A_h \to \mathbb{R}$ and $R : A_h \to \mathbb{R}$ are suitable functionals and $\varepsilon > 0$ a positive parameter.

The functional $R$ is the so-called regularization functional and the parameter $\varepsilon$ is the so-called regularization parameter.

In the next section, we study the case when the map $J$ is the output least squares functional.

3.3 Output Least-Squares (OLS)

We will now consider the variational problem of (3.7). We first want to choose $a$ to minimize the least output-squares

$$\min_{a \in A} \frac{1}{2} \| u(a) - z \|^2, \tag{3.12}$$

where $z$ is the data and $u(a)$ solves the variational problem

$$\int_0^1 a(x) (u'')(v') dx = \int_0^1 f v dx \quad \forall v \in V. \tag{3.13}$$

For the finite dimensional analogue of the above functional, we have

$$J_1(A) = \frac{1}{2} \langle u - z, u - z \rangle$$

$$= \frac{1}{2} \| u(a) - z, u(a) - z \|$$

$$= \frac{1}{2} \int_0^1 (u(a) - z)(u(a) - z) dx$$

where $A$ is the vector of the nodal values for $a$. 
We have \( z \in V_h \) and \( u \in V_h \):

\[
\begin{align*}
z &= \left[ \sum_{i=1}^{n} Z_i \phi_i, \sum_{i=1}^{n} Z_i \psi_i \right]^T, \\
u &= \left[ \sum_{i=1}^{n} U_i \phi_i, \sum_{i=1}^{n} U_i \psi_i \right]^T.
\end{align*}
\]

By letting, \( v = u - z \) or \( V_i = U_i - Z_i \), we have

\[
\begin{align*}
\|u(a) - z\|^2 &= \langle v, v \rangle \\
&= \left\langle \sum_{i}^{n} V_i \phi_i, \sum_{j}^{n} V_j \phi_j \right\rangle \\
&= \langle v_1 \phi_i + \ldots + v_n \phi_n, v_1 \phi_j + \ldots + v_n \phi_n \rangle \\
&= V^T A V \\
&= \left\langle \sum_{i}^{n} V_i \psi_i, \sum_{j}^{n} V_j \psi_j \right\rangle \\
&= \langle v_1 \psi_i + \ldots + v_n \psi_n, v_1 \psi_j + \ldots + v_n \psi_n \rangle \\
&= V^T C V \\
&= \left\langle \sum_{i}^{n} V_i \phi_i, \sum_{j}^{n} V_j \phi_j \right\rangle \\
&= \langle v_1 \phi_i + \ldots + v_n \phi_n, v_1 \phi_j + \ldots + v_n \phi_n \rangle \\
&= V^T B V.
\end{align*}
\]

We can put these four submatrices into one cumulative matrix by letting

\[
M = \begin{pmatrix}
A & B \\
B^T & C
\end{pmatrix}.
\]

This is the so-called mass matrix.

To summarize:

Entries of \( A : A_{ik} = \int_0^1 \phi_i \phi_j \, dx \)
3.3. Output Least-Squares (OLS)

Entries of $C: C_{ik} = \int_0^1 \psi_i \psi_j \, dx$

Entries of $B: B_{ik} = \int_0^1 \phi_i \psi_j \, dx$

To find the mass matrix,

$$A_{i,i} = \int_{x_{j-i}}^{x_j} \phi_j^2 \, dx + \int_{x_j}^{x_{j+1}} \phi_j^2 \, dx = \frac{13}{35} h^7 + \frac{13}{35} h^6 = \frac{26}{35} h$$

$$A_{i-1,i} = a_{i,i-1} = \int_{x_{j-i}}^{x_j} \phi_{j-1} \phi_j \, dx = \frac{9}{70} h^7 = \frac{9}{70} h$$

$$B_{i,i} = \int_{x_{j-i}}^{x_j} \phi_j \psi_j \, dx + \int_{x_j}^{x_{j+1}} \phi_j \psi_j \, dx = \frac{11}{210} h^5 - \frac{11}{210} h^5 = 0$$

$$B_{i-1,i} = \int_{x_{j-i}}^{x_j} \phi_{j-1} \psi_j \, dx = \frac{-13}{420} h^5 = \frac{-13}{420} h^2$$

$$B_{i,i-1} = \int_{x_{j-i}}^{x_j} \phi_j \psi_{j-1} \, dx = \frac{13}{420} h^5 = \frac{13}{420} h^2$$

$$C_{i,i} = \int_{x_{j-i}}^{x_j} \psi_j^2 \, dx + \int_{x_j}^{x_{j+1}} \psi_j^2 \, dx = \frac{1}{105} h^4 + \frac{1}{105} h^4 = \frac{2}{105} h^3$$

$$C_{i-1,i} = c_{i,i-1} = \int_{x_{j-i}}^{x_j} \psi_{j-1} \psi_j \, dx = \frac{-1}{140} h^4 = \frac{-1}{140} h^3$$

We now want to take the gradient of the objective function. We first need to compute

$$\delta V = DF(A) \delta A.$$

From $K(A)V = F$ where $K(A)$ is the stiffness matrix and $F$ is the load vector, we take the derivative with respect to $A$:

$$DK(A) \delta AV + K(A) \delta V = 0$$

$$K(A) \delta V = -DK(\delta A) \delta AV$$

$$K(A) \delta V = -K(\delta A)V$$

$$\delta V = -K(A)^{-1}DK(\delta A)V$$

We next define the adjoint stiffness matrix $L(V)$ by the condition

$$L(V)A = K(A)V \text{ for all } A \in R^{n+2}, V \in R^n \quad (3.14)$$

We can now say

$$\delta V = -K(A)^{-1}L(V) \delta A. \quad (3.15)$$
3.4 Modified Output Least-Squares (MOLS)

We will compute the gradient of \( J_1 \)

\[
J_1(a) = \frac{1}{2}(u - z) \cdot M(u - z)
\]

\[
DJ_1(a)(\delta a) = \frac{1}{2}(\delta u - 0) \cdot M(u - z) + \frac{1}{2}(u - z) \cdot M(\delta u - 0)
\]

\[
= \frac{1}{2} \langle \delta u, M(u - z) \rangle + \frac{1}{2} \langle (u - z), M \delta u \rangle
\]

\[
= \langle \delta u, M(u - z) \rangle \quad \text{where} \quad \delta u = -K(A)^{-1}K(\delta A)V
\]

\[
= -K(A)^{-1}K(\delta A)V \cdot M(u - z)
\]

\[
= -K(A)^{-1}L(V)\delta A \cdot M(u - z) \quad \text{where} \quad L(V)\delta A = K(\delta A)V
\]

\[
= -\delta A \cdot (K(A)^{-1}L(V))^TM(u - z)
\]

\[
= -\delta A \cdot L(V)^TK(A)^{-1}M(u - z)
\]

\[
\nabla J_0 = -L(V)^TK(A)^{-1}M(u - z).
\]

The adjoint stiffness matrix will be explained more in chapter 4.

### 3.4 Modified Output Least-Squares (MOLS)

The modified OLS is caused by replacing the \( L^2 \) norm by the coefficient-dependent energy norm, and is how we define the objective function, \( J_2 \)

\[
J_2(A) = \frac{1}{2}T(a,u(a) - z,u(a) - z)
\]

\[
= \frac{1}{2}(U - Z)K(A)(U - Z).
\]

Now to show the gradient of \( J_2 \),

\[
J_2(a) = \frac{1}{2}T(a,u(a) - z,u(a) - z)
\]

\[
DJ_2(a)(\delta a) = \frac{1}{2}T(\delta a,u(a) - z,u(a) - z) + \frac{1}{2}T(a,DU(a)(\delta a,u(a) - z) + \frac{1}{2}T(a,DU(a)(\delta a,u(a) - z)
\]

\[
= \frac{1}{2}T(\delta a,u(a) - z,u(a) - z) + T(a,\delta u,u(a) - z)
\]

\[
= \frac{1}{2}T(\delta a,u(a) - z,u(a) - z) - T(\delta a,u,a(a) - z)
\]

\[
= \frac{1}{2}T(\delta a,u(a) + z,u(a) - z).
\]
For the discrete form:

\[
J_2(A) = \frac{1}{2}(U - Z)K(A)(U - Z)
\]

\[
DJ_2(A)(\delta A) = \frac{1}{2}\delta \cdot K(A) + \frac{1}{2}(U - Z) \cdot K(A)(\delta U) + \frac{1}{2}(U - Z) \cdot DK(A)(\delta A)(U - Z)
\]

\[
= \delta U \cdot K(A)(U - Z) + \frac{1}{2}(U - Z) \cdot K(\delta A)(U - Z)
\]

\[
= -(K(A)^{-1}L(A)\delta A) \cdot K(A)(U - Z) + \frac{1}{2}(U - Z) \cdot K(\delta A)(U - Z)
\]

\[
= -\delta A \cdot L(U)^T(U - Z) + \frac{1}{2}(U - Z) \cdot L(U - Z) \delta A
\]

\[
= -\delta A \cdot L(U)^T(U - Z) + \frac{1}{2}\delta A \cdot L(U - Z)^T(U - Z)
\]

\[
= -\frac{1}{2}\delta A \cdot L(U + Z)^T(U - Z)
\]

\[
= -\frac{1}{2}\delta A \cdot L(U)^T U + \frac{1}{2}\delta A \cdot L(Z)^T Z.
\]

Later on in chapter five, we will need the second derivative or the Hessian matrix of the objective functional.

\[
J_2(A) = \frac{1}{2}(U - Z)K(A)(U - Z)
\]

\[
DJ_2(A)(\delta A) = -\frac{1}{2}\delta A \cdot L(U + Z)^T(U - Z)
\]

\[
= -\frac{1}{2}\delta A \cdot L(U + Z)^T(U - Z)
\]

\[
D^2J_2(A)(\delta A, \delta A) = -\frac{1}{2}\delta A \cdot L(\delta U)^T U - \frac{1}{2}\delta A \cdot L(U)^T \delta U.
\]

Take notice that \( T_{ijk} = T_{jik} \) so therefore we can say

\[
L(\delta U)^T U = L(U)^T \delta U
\]

\[
\delta U = K(A)^{-1}L(U)\delta A.
\]

Therefore,

\[
D^2J_2(A)(\delta A, \delta A) = -\frac{1}{2}\delta A \cdot L(\delta U)^T U - \frac{1}{2}\delta A \cdot L(U)^T \delta U
\]

\[
= -\delta A \cdot L(U)^T \delta U
\]

\[
= -\delta A \cdot L(U)^T K(A)^{-1}L(U)\delta A.
\]
and hence,

\[ D^2 J_2(A) = L(U)^T K(A)^{-1} L(U). \]

Summarizing

\[ J_2(A) = \frac{1}{2} (U - Z) K(A) (U - Z); \]
\[ DJ_2(A) = -\frac{1}{2} L(U + Z)^T (U - Z); \]
\[ D^2 J_2(A) = L(U)^T K(A)^{-1} L(U). \]

### 3.5 Regularization

From (3.11), we need to minimize a regularization function. The regularization norm can be one of three norms. Each norm represents the Euclidean norm, \( \| \cdot \|_2 \).

- **L₂-norm (square)**: \( R_1(a) = \| a \|^2 = \int_0^1 a(x)^2 dx \)

- **H₁-norm**: \( R_2(a) : \| a \|^2 + \left\| \frac{da}{dx} \right\|^2 = \int_0^1 a(x)^2 + \left( \frac{da}{dx} \right)^2 dx \)

- **H₁-semi-norm**: \( R_3(a) = \left\| \frac{da}{dx} \right\|^2 = \int_0^1 \left( \frac{da}{dx} \right)^2 dx \)

The discretized regularized norms read:

\[ \| a \|^2 = \left\langle \sum_{i=0}^{n+1} A_i \phi_j, \sum_{i=0}^{n+1} A_j \phi_j \right\rangle = A^T \tilde{M} A \]
\[ \left\| \frac{da}{dx} \right\|^2 = \left\langle \sum_{i=0}^{n+1} A_i \frac{d\phi_j}{dx}, \sum_{i=0}^{n+1} A_j \frac{d\phi_j}{dx} \right\rangle = A^T \tilde{K} A \]

where \( \tilde{M} \) is a \((N+2) \times (N+2)\) matrix and \( \tilde{K} \) is a \((N+2) \times (N+2)\) matrix due to not knowing the
boundary values of the coefficient. The regularization term \( R(a) \) will be one of three matrices

\[
\begin{align*}
L_2\text{-norm (square)} & : \quad R_1(a) = A^T \bar{M} A \\
H_1\text{-norm} & : \quad R_2(a) = A^T \bar{M} A + A^T \bar{K} A \\
H_1\text{-simi-norm} & : \quad R_3(a) = A^T \bar{K} A.
\end{align*}
\]

Throughout this paper we shall use the \( H_1 \)-norm as the regularization term.
Chapter 4

Identification by Derivative Free Methods

In this chapter we use a derivative free Nelder-Mead method to solve the inverse problem of identifying variable parameters in fourth-order boundary value problem. Derivative free methods are very easy to use as they by pass the stringent requirement of computing the derivatives. However their accuracy is rather limited to small dimension problem.

4.1 Nelder-Mead Method

We have shown that the direct method works, and now we want to show that an iterative derivative free method will work. We chose to use the Nelder-Mead method. This method works by stating with an initial guess of $n$ variables and a radius. Then from the initial guess, $n+1$ vectors are form to creat an $n$-dimensional simplex. The $n+1$ vertices of this simplex are tested and are arranged in ascending order. The goal is to replace the last value, $y_h$, this corresponds to the vector $a_h = a_{n+1}$. Then the center of the face of the simplex that omits the $a_h$ vertex is computed, $a_c$. Then the reflection point is computed, $a_r = a_c - a_h$. If $y_r$ lies in the range of $y_1$ and $y_n$, then we replace the worst point, $a_h$ with $a_r$, and sort the vertices by their functional values and repeat the step.

If $y_r$ is less than $y_1$, then we expand by setting $a_e = 3a_c - 2a_h$, replace with the better of $y_e$, or $y_r$. If $y_r$ is greater than $y_n$, then we make an outside contraction by setting $a_{oc} = 1.5a_c - 0.5a_h$ or an inside contraction by setting $a_{ic} = 0.5a_c + 0.5a_h$, if the contraction is less then $y_h$ then replace $a_h$ with that value.

If $y_h$ is better than both $a_{ic}$ and $a_{oc}$, then shrink the simplex by a factor of two in the direction of the minimum $a_1$, then repeat the steps.
4.2 Nelder-Mead Algorithm

**Initialization Step:** Choose some radius, and an n-dimensional $a^0$ initial column vector. Let $e$ be a n-vector size of ones, and $e_i$ be a vector of zeros with a one in the $i$th position. Let $y_i = J(a_i)$

**Step 1:** Create vertices $\sum [(a_0 + \text{radius} \ast e_i) \ast e]$

**Step 2:** Order $y_1 < y_2 < \cdots < y_{n+1} = y_h$

**Step 3:** Reflection $a_r = a_c - a_h$

If $(y_1 < y_r < y_n)$, then let $a_h = a_r$ go to Step 2.
else if $(y_r < y_1)$, then go to Step 4
else $(y_n < y_r)$, then go to Step 5

**Step 4:** Expansion $x_e = 3x_c - 2x_h$

If $(y_e < y_h)$, then let $x_h = x_e$ go to step 2.
else $(y_h < y_e)$, then let $x_h = x_r$ go to step 2.

**Step 5:** Contraction

If $(y_r < y_h)$, then go to 5a.
else $(y_h < y_r)$, then go to 5b.

5a. Outside Contraction $a_{oc} = 1.5a_c - 0.5a_h$

If $(y_{oc} < y_h)$, then let $a_h = a_{oc}$ go to Step 2.
else $(y_h < y_{oc})$, then go to Step 6.

5b. Inside Contraction $x_{oc} = 0.5a_c + 0.5a_h$

If $(y_{ic} < y_h)$, then let $a_h = x_{ic}$ go to Step 2.
else $(y_h < y_{ic})$, then go to Step 6.

**Step 6:** Reduction

Shrink
4.3 Numerical Results

In the following results, OLS and MOLS was used along with a regularization norm of $H_1$-norm. We have used the same four examples that we used for the direct method in Chapter 2. To obtain the optimal regularization parameter $\varepsilon$ in the following results, we tested a ranged of values for each example and then picked the one most reliable. The results will show two methods for each result, (a) using OLS, and (b) using MOLS. The top figures in each method will display two lines, the experimented or computed coefficient and the real coefficient. Then the bottom figure is putting the computed coefficient line into (3.2) directly and seeing how closely it matches up to the real data and we’ll see it’s very close most of the time. In all the cases, the MOLS performs better and converges quickly. Using the iterative Nelder-Mead algorithm for the OLS method took roughly 15 to 30 minutes to run completely, while for the MOLS method for the algorithm took about 2 minutes to complete.

![Figure 4.1](image)

Figure 4.1: Example 1 using Nelder-Mead Simplex Method
4.3. Numerical Results

Figure 4.2: Example 2 using Nelder-Mead Simplex Method

Figure 4.3: Example 3 using Nelder-Mead Simplex Method
Figure 4.4: Example 4 using Nelder-Mead Simplex Method
Chapter 5

Computation of the Derivatives

In this chapter we describe the details of the computation of the adjoint stiffness matrix which plays the most fundamental role in the calculation of the derivatives of the objective functionals.

5.1 Adjoint Stiffness Matrix Based Approach

We recall that the adjoint stiffness matrix is defined by the condition

\[ L(\tilde{V})\tilde{A} = K(\tilde{A})\tilde{V}, \quad \forall \tilde{A} \in \mathbb{R}^{N+2}, \quad \forall \tilde{V} \in \mathbb{R}^{2N}, \quad (5.1) \]

where \( \tilde{V} = (V, \tilde{V})^T \).

Since

\[ K(\tilde{A})_{ij} = \int_0^1 a(x) \bar{\phi}_j'' \bar{\phi}_i'' dx \]

where

\[ a(x) = \sum_{k=0}^{n+1} A_k a_k \]

and

\[ \bar{\phi} = (\phi, \psi)^T. \]

Therefore, with

\[ K(\tilde{A})_{ij} = \sum_{k=0}^{n+1} T_{ijk} \tilde{A}_k \]

where

\[ T_{ijk} = \int_0^1 a_k \bar{\phi}_j'' \bar{\phi}_i'' dx \]
we obtain the formula
\[(K(\tilde{A})\tilde{V})_i = \sum_{j=1}^{n} T_{ijk} \tilde{A}_k \tilde{V}_j.\]

Now using (5.1), we obtain
\[(L(\tilde{V}))_{ik} = \sum_{j=1}^{2n} T_{ijk} \tilde{V}_j \]
\[= \sum_{j=1}^{2n} \left( \int_{0}^{1} a_k \phi_j'' \phi_i'' \, dx \right) V_j \text{ where } \phi = \begin{pmatrix} \phi \\ \psi \end{pmatrix} \text{ and} \]
\[(L(\tilde{V}))_{ik} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_k \phi_j'' \phi_i'' \, dx \right) V_j + \sum_{j=1}^{n} \left( \int a_k \psi_j'' \phi_i'' \, dx \right) \hat{V}_j \]
where \(i = 1, \ldots, 2n\) and \(k = 0, \ldots, n + 1\).

Notice that the \(L(V)\) matrix has a size of \(2n \times (n + 2)\). We break the adjoint stiffness matrix \(L\) into two blocks of size \(n \times (n + 2)\) matrices. We will use the notation
\[L = \begin{pmatrix} A + B \\ D + C \end{pmatrix}\]
where
\[A_{ik} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_k \phi_j'' \phi_i'' \, dx \right) V_j \]
\[B_{ik} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_k \psi_j'' \phi_i'' \, dx \right) \hat{V}_j \]
\[D_{ik} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_k \phi_j'' \psi_i'' \, dx \right) V_j \]
\[C_{ik} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_k \psi_j'' \psi_i'' \, dx \right) \hat{V}_j.\]
The basis functions for \(\phi\) and \(\psi\) are as defined in Chapter 2.
Also from Chapter 2, we need the second derivatives at different $x$-values which are:

\[
\begin{align*}
\phi_j''(x_{j-1}) &= \phi_j''(x_{j+1}) = \frac{6}{h^2} \\
\phi_j''(x_{j-1/2}) &= \phi_j''(x_{j+1/2}) = 0 \\
\phi_j''(x_j) &= -\frac{6}{h^2} \\
\psi_j''(x_{j-1}) &= -\frac{2}{h} \\
\psi_j''(x_{j-1/2}) &= -\frac{1}{h} \\
\psi_j''(x_{j+1/2}) &= -\frac{2}{h} \\
\psi_j''(x_{j+1}) &= -\frac{2}{h} \\
\lim_{x \to x_j^-} \psi_j''(x) &= \frac{4}{h} \\
\lim_{x \to x_j^+} \psi_j''(x) &= -\frac{4}{h}.
\end{align*}
\]

We will denote $I_i$ to be the $i$-th subinterval $[x_{i-1}, x_i]$. It is important to note that support of both $\phi_i$ and $\psi_i$ is $I_i \cup I_{i+1}$.

![Figure 5.1: Basis function $\phi_j$](image)

We first construct the matrix $A$.

\[
A_{ik} = \sum_{j=1}^{n} \left( \int a_k \phi_j'' \phi_i'' \, dx \right) V_j \text{ for } i = 1, \ldots, n \text{ and } k = 0, \ldots n + 1.
\]

We start with the first row ($i = 1$) and the index $k$ varies from 0 to $n + 1$. When $k = 0$, notice that the basis function for $w_0$ has non-zero values only on $I_1$. 

5.1. Adjoint Stiffness Matrix Based Approach

Figure 5.2: Basis function $\psi_j$

Figure 5.3: Basis function $a_0$

Figure 5.4: Basis function $a_1$
5.1. Adjoint Stiffness Matrix Based Approach

We have

$$A_{10} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{0} \phi_{j}'' \phi_{1}'' dx \right) V_{j}$$

$$= V_{1} \int_{0}^{1} a_{0} (\phi_{1}'')^{2} dx$$

$$= V_{1} \int_{0}^{1} a_{0} (\phi_{1}'')^{2} dx$$

$$= V_{1} \frac{h}{2} \left[ a_{0}(x_{0}) (\phi_{1}'')(x_{0})^{2} + 4a_{0}(x_{1}/2) (\phi_{1}'')(x_{1}/2)^{2} + a_{0}(x_{1}) (\phi_{1}'')(x_{1})^{2} \right]$$

$$= V_{1} \frac{h}{6} \left[ 1 \cdot \left( \frac{6}{h^{2}} \right)^{2} + 4 \cdot \frac{1}{2} \cdot 0 + 0 \cdot \left( -\frac{6}{h^{2}} \right)^{2} \right]$$

$$= \frac{6}{h^{3}} V_{1}.$$ 

where we used Simpson’s Rule for the integration. For $k = 1$ we notice that the basis function $a_{1}$ has non-zero values only on $I_{1}$ and $I_{2}$. We are collecting all nonzero terms involving basis functions $\phi_{1}$ and $\phi_{2}$. Then

$$A_{11} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{1} \phi_{j}'' \phi_{2}'' dx \right) V_{j}$$

$$= V_{1} \int_{0}^{1} a_{1} (\phi_{1}'')^{2} dx + V_{2} \int_{1}^{a} a_{1} \phi_{2}'' \phi_{1}'' dx$$

$$= V_{1} \int_{0}^{1} a_{1} (\phi_{1}'')^{2} dx + V_{1} \int_{I_{2}} a_{1} (\phi_{1}'')^{2} dx + V_{2} \int_{I_{2}} a_{1} \phi_{2}'' \phi_{1}'' dx$$

$$= V_{1} \frac{h}{6} \left[ a_{1}(x_{0}) (\phi_{1}'')(x_{0})^{2} + 4a_{1}(x_{1}/2) (\phi_{1}'')(x_{1}/2)^{2} + a_{1}(x_{1}) (\phi_{1}'')(x_{1})^{2} \right]$$

$$+ V_{1} \frac{h}{6} \left[ a_{1}(x_{1}) (\phi_{1}'')(x_{1})^{2} + 4a_{1}(x_{3}/2) (\phi_{1}'')(x_{3}/2)^{2} + a_{1}(x_{2}) (\phi_{1}'')(x_{2})^{2} \right]$$

$$+ V_{2} \frac{h}{6} \left[ a_{1}(x_{1}) \phi_{2}''(x_{1}) \phi_{1}'(x_{1}) + 4a_{1}(x_{3}/2) \phi_{2}''(x_{3}/2) \phi_{1}'(x_{3}/2) + a_{1}(x_{2}) \phi_{2}''(x_{2}) \phi_{1}'(x_{2}) \right]$$

$$= V_{1} \frac{h}{6} \left( -\frac{6}{h^{2}} \right)^{2} + V_{1} \frac{h}{6} \left( -\frac{6}{h^{2}} \right)^{2} + V_{2} \frac{h}{6} \left( \frac{6}{h^{2}} \right) \left( -\frac{6}{h^{2}} \right)$$

$$= \frac{6}{h^{3}} (2V_{1} - V_{2}).$$

When $k = 2$, we notice that the basis function $a_{2}$ has non-zero values only on $I_{2}$ and $I_{3}$. Similarly,
5.1. Adjoint Stiffness Matrix Based Approach

\[ A_{12} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_2 \phi_j'' \phi_1'' dx \right) V_j \]

\[ = V_1 \int_{0}^{1} a_2 (\phi_1'')^2 dx + V_2 \int a_2 \phi_2'' \phi_1'' dx \]

\[ = V_1 \int_{I_2} a_2 (\phi_1'')^2 dx + V_2 \int_{I_2} a_2 \phi_2'' \phi_1'' dx \]

\[ = V_1 \frac{h}{6} \left[ a_2(x_1) (\phi_1''(x_1))^2 + 4 a_2(x_{3/2}) (\psi_1''(x_{3/2}))^2 + a_2(x_2) (\phi_1''(x_2))^2 \right] \]

\[ + V_2 \frac{h}{6} \left[ a_2(x_1) \phi_2''(x_1) \phi_1''(x_1) + 4 a_2(x_{3/2}) \phi_2''(x_{3/2}) \phi_1''(x_{3/2}) + a_2(x_2) \phi_2''(x_2) \phi_1''(x_2) \right] \]

\[ = V_1 \frac{h}{6} \left( \frac{6}{h^2} \right)^2 + V_2 \frac{h}{6} \left( - \frac{6}{h^2} \right) \left( \frac{6}{h^2} \right) \]

\[ = \frac{6}{h^3} (V_1 - V_2). \]

When \( k = 3 \), we notice that the basis function \( a_3 \) has non-zero values only on \( I_3 \) and \( I_4 \), and these subintervals have no overlapping with the support of \( \phi_k \) if \( k \geq 3 \). Therefore \( A_{1k} = 0 \) for \( k \geq 3 \).
For the second row of matrix $A$ we have

$$A_{20} = 0$$

$$A_{21} = \sum_{j=1}^{n} \left( \int_0^1 a_1 \phi_j'' \phi_2'' \, dx \right) V_j$$

$$= V_1 \int_0^1 a_1 \phi_1'' \phi_2'' \, dx + V_2 \int_0^1 a_1 (\phi_2'')^2 \, dx$$

$$= V_1 \int_{I_2} a_2 \phi_1'' \phi_2'' \, dx + V_2 \int_{I_2} a_2 (\phi_2'')^2 \, dx$$

$$= V_1 \frac{h}{6} \left[ a_2(x_1) \phi_1''(x_1) \phi_2''(x_1) + 4a_2(x_{3/2}) \phi_1''(x_{3/2}) \phi_2''(x_{3/2}) + a_2(x_2) \phi_1''(x_2) \phi_2''(x_2) \right]$$

$$+ V_2 \frac{h}{6} \left[ a_2(x_1)(\phi_2''(x_1))^2 + 4a_2(x_{3/2})(\phi_2''(x_{3/2}))^2 + a_2(x_2)(\phi_2''(x_2))^2 \right]$$

$$= V_1 \frac{h}{6} \left( \frac{6}{h^2} \right) \left( - \frac{6}{h^2} \right) + V_2 \frac{h}{6} \left( - \frac{6}{h^2} \right)^2$$

$$= \frac{6}{h^3} (-V_1 + V_2).$$
5.1. Adjoint Stiffness Matrix Based Approach

\[ A_{22} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_2 \phi_j'' \phi_2'' \, dx \right) V_j \]

\[ = V_1 \int_{0}^{1} a_2 \phi_1'' \phi_2'' \, dx + V_2 \int_{1/2}^{1} a_2 \left( \phi_2''(x) \right)^2 \, dx + V_3 \int_{1/2}^{1} a_2 \phi_2'' \phi_2'' \, dx \]

\[ = V_1 \int_{j_2}^{1} a_2 \phi_1'' \phi_2'' \, dx + V_2 \int_{j_2}^{1} a_2 \left( \phi_2''(x) \right)^2 \, dx + V_2 \int_{j_3}^{1} a_2 \phi_2'' \phi_2'' \, dx \]

\[ = V_1 \frac{h}{6} \left[ a_2(x_1) \phi_1''(x_1) \phi_2''(x_1) + 4a_1(x_{3/2}) \phi_1''(x_{3/2}) \phi_2''(x_{3/2}) + a_2(x_2) \phi_1''(x_2) \phi_2''(x_2) \right] \]

\[ + V_2 \frac{h}{6} \left[ a_2(x_1) \left( \phi_2''(x_1) \right)^2 + 4a_1(x_{3/2}) \left( \phi_2''(x_{3/2}) \right)^2 + a_2(x_2) \left( \phi_2''(x_2) \right)^2 \right] \]

\[ + V_3 \frac{h}{6} \left[ a_2(x_3) \phi_2''(x_3) \phi_2''(x_3) + 4a_2(x_{5/2}) \phi_3''(x_{5/2}) \phi_2''(x_{5/2}) + a_2(x_3) \phi_3''(x_3) \phi_2''(x_3) \right] \]

\[ = V_1 \frac{h}{6} \left( \frac{6}{h^2} \right) \left( - \frac{6}{h^2} \right) + V_2 \frac{h}{6} \left( - \frac{6}{h^2} \right)^2 + V_3 \frac{h}{6} \left( \frac{6}{h^2} \right) \left( - \frac{6}{h^2} \right) \]

\[ = \frac{6}{h^3} (-V_1 + 2V_2 - V_3). \]

\[ A_{23} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_3 \phi_j'' \phi_2'' \, dx \right) V_j \]

\[ = V_2 \int_{0}^{1} a_3 \left( \phi_2''(x) \right)^2 \, dx + V_3 \int_{1/2}^{1} a_3 \phi_2'' \phi_2'' \, dx \]

\[ = V_2 \int_{j_2}^{1} a_3 \left( \phi_2''(x) \right)^2 \, dx + V_2 \int_{j_2}^{1} a_3 \phi_2'' \phi_2'' \, dx \]

\[ = V_2 \frac{h}{6} \left[ a_3(x_2) \left( \phi_2''(x_2) \right)^2 + 4a_3(x_{5/2}) \left( \phi_2''(x_{5/2}) \right)^2 + a_3(x_3) \left( \phi_2''(x_3) \right)^2 \right] \]

\[ + V_3 \frac{h}{6} \left[ a_3(x_2) \phi_3''(x_2) \phi_2''(x_2) + 4a_2(x_{5/2}) \phi_3''(x_{5/2}) \phi_2''(x_{5/2}) + a_2(x_3) \phi_3''(x_3) \phi_2''(x_3) \right] \]

\[ = V_2 \frac{h}{6} \left( \frac{6}{h^2} \right)^2 + V_3 \frac{h}{6} \left( - \frac{6}{h^2} \right) \left( \frac{6}{h^2} \right) \]

\[ = \frac{6}{h^3} (V_2 - V_3). \]

Notice that

\[ A_{2k} = 0 \text{ for } k \geq 4. \]
This pattern will continue up to the row $N - 1,$

\[ A_{n-1,k} = 0 \text{ for } k \leq n - 3 \]
\[ A_{n-1,n-2} = \frac{6}{h^3} (-V_{n-2} + V_{n-1}) \]
\[ A_{n-1,n-1} = \frac{6}{h^3} (-V_{n-2} + 2V_{n-1} - V_n) \]
\[ A_{n-1,n} = \frac{6}{h^3} (V_{n-1} - V_n) \]
\[ A_{n-1,n+1} = 0. \]

For the last row, the $n$-th,

\[ A_{n,k} = 0 \text{ for } k \leq n - 2. \]

When $k = n - 1$ the basis function $a_{n-1}$ will have non-zero values only on of $I_{n-1}$ and $I_n,$ so

\[ \text{Figure 5.7: Basis function } a_{n-1} \]
5.1. Adjoint Stiffness Matrix Based Approach

\[ A_{n,n-1} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n-1}\phi_j'' \phi_n'' \,dx \right) V_j \]

\[ = V_{n-1} \int_{I_n} a_{n-1}\phi_n'' \,dx + V_n \int_{I_n} a_{n-1} (\phi_n'')^2 \,dx \]

\[ = V_{n-1} \frac{h}{6} \left[ a_{n-1}(x_{n-1})\phi_{n-1}''(x_{n-1})\phi_n''(x_{n-1}) + 4a_{n-1}(x_{n-1}/2)\phi_{n-1}''(x_{n-1}/2)\phi_n''(x_{n-1}/2) \right. \]

\[ + a_{n-1}(x_{n})\phi_{n-1}''(x_{n})\phi_n''(x_{n}) \left] + V_n \frac{h}{6} \left[ a_{n-1}(x_{n-1}) (\phi_n''(x_{n-1}))^2 \right. \]

\[ + 4a_{n-1}(x_{n-1}/2) (\phi_n''(x_{n-1}/2))^2 + a_{n-1}(x_{n}) (\phi_n''(x_{n}))^2 \left] \right. \]

\[ = V_{n-1} \frac{h}{6} \left[ \left( -\frac{6}{h^2} \right) \left( \frac{6}{h^2} \right) \right] + V_n \frac{h}{6} \left[ \left( \frac{6}{h^2} \right)^2 \right] \]

\[ = \frac{6}{h^3} (-V_{n-1} + V_n). \]

When \( k = n \), notice that the basis function \( a_n \) has non-zero values only on \( I_n \) and \( I_{n+1} \).

Figure 5.8: Basis function \( a_n \)

Then

\[ A_{n,n} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_n\phi_j'' \phi_n'' \,dx \right) V_j \]

\[ = V_{n-1} \int_{I_n} a_n\phi_{n-1}'' \phi_n'' \,dx + V_n \int_{I_n} a_n(\phi_n'')^2 \,dx + V_n \int_{I_{n+1}} a_n(\phi_n'')^2 \,dx \]

\[ = V_{n-1} \frac{h}{6} \left[ \left( \frac{6}{h^2} \right) \left( -\frac{6}{h^2} \right) \right] + V_n \frac{h}{6} \left[ \left( \frac{6}{h^2} \right)^2 \right] + V_n \frac{h}{6} \left[ \left( \frac{6}{h^2} \right)^2 \right] \]

\[ = \frac{6}{h^3} (-V_{n-1} + 2V_n). \]
For $k = n + 1$ the basis function $a_{n+1}$ has non-zero values only on $I_{n+1}$.

Therefore

$$A_{n,n+1} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n+1} \phi_j'' \phi_n'' \, dx \right) V_j = V_n \int_{I_{n+1}} a_{n+1} (\phi_n'')^2 \, dx$$

$$= V_n \frac{h}{6} \left[ a_{n+1}(x_n) \left( \phi_n''(x_n) \right)^2 + 4a_{n+1}(x_{n+1}/2) \left( \phi_n''(x_{n+1}/2) \right)^2 + a_{n+1}(x_{n+1}) \left( \phi_n''(x_{n+1}) \right)^2 \right]$$

$$= V_n \frac{h}{6} \left[ \left( \frac{6}{h^2} \right)^2 \right]$$

$$= \frac{6}{h^3} V_n.$$  

Putting together the entries of we get the following matrix:

$$A = \frac{6}{h^3} \begin{pmatrix} V_1 & 2V_1 - V_2 & V_1 - V_2 & 0 & 0 & \ldots & 0 \\ 0 & -V_1 + V_2 & -V_1 + 2V_2 - V_3 & V_2 - V_3 & 0 & \ldots & 0 \\ 0 & \ldots & -V_2 + V_3 & -V_2 + 2V_3 - V_4 & V_3 - V_4 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & -V_{n-2} + V_{n-1} & -V_{n-2} + 2V_{n-1} - V_n & V_{n-1} - V_n & 0 \\ 0 & \ldots & 0 & 0 & -V_{n-1} + V_n & -V_{n-1} + 2V_n & V_n \end{pmatrix}.$$  

We evaluate the entries of matrix $B$ in a similar manner. Recall that the entries of matrix $B$ are defined by

$$B_{ik} = \sum_{j=1}^{n} \left( \int a_k \psi_j'' \phi_i'' \, dx \right) \hat{V}_j \text{ for } i = 1, \ldots, n \text{ and } k = 0, \ldots, n + 1.$$
5.1. Adjoint Stiffness Matrix Based Approach

We start with the first row.

\[ B_{10} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_0 \psi_j'' \phi_i'' \, dx \right) \hat{V}_j \]

\[ = \hat{V}_1 \int_{0}^{1} a_0 \psi_j'' \phi_i'' \, dx \]

\[ = \hat{V}_1 \int_{0}^{1} a_0 \psi_j'' \phi_i'' \, dx \]

\[ = \hat{V}_1 \frac{h}{2} \left[ a_0(x_0) \psi_1''(x_0) \phi_1''(x_0) + 4a_0(x_{1/2}) \psi_1''(x_{1/2}) \phi_1''(x_{1/2}) + a_0(x_1) \psi_1''(x_1) \phi_1''(x_1) \right] \]

\[ = \hat{V}_1 \frac{h}{6} \left[ \left( -\frac{2}{h} \right) \left( \frac{6}{h^3} \right) + 0 + 0 \right] \]

\[ = -\frac{2}{h^2} \hat{V}_1. \]

\[ B_{11} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_1 \psi_j'' \phi_i'' \, dx \right) \hat{V}_j \]

\[ = \hat{V}_1 \int_{0}^{1} a_1 \psi_j'' \phi_i'' \, dx + \hat{V}_2 \int_{0}^{1} a_1 \psi_j'' \phi_i'' \, dx \]

\[ = \hat{V}_1 \int_{h}^{1} a_1 \psi_j'' \phi_i'' \, dx + \hat{V}_2 \int_{l_2}^{1} a_1 \psi_j'' \phi_i'' \, dx \]

\[ = \hat{V}_1 \frac{h}{6} \left[ a_1(x_0) \psi_1''(x_0) \phi_1''(x_0) + 4a_1(x_{1/2}) \psi_1''(x_{1/2}) \phi_1''(x_{1/2}) + a_1(x_1) \psi_1''(x_1) \phi_1''(x_1) \right] + \]

\[ \hat{V}_1 \frac{h}{6} \left[ a_1(x_1) \psi_1''(x_1) \phi_1''(x_1) + 4a_1(x_{3/2}) \psi_1''(x_{3/2}) \phi_1''(x_{3/2}) + a_1(x_2) \psi_1''(x_2) \phi_1''(x_2) \right] + \]

\[ \hat{V}_2 \frac{h}{6} \left[ a_1(x_1) \psi_2''(x_1) \phi_1''(x_1) + 4a_1(x_{3/2}) \psi_2''(x_{3/2}) \phi_1''(x_{3/2}) + a_1(x_2) \psi_2''(x_2) \phi_1''(x_2) \right] \]

\[ = \hat{V}_1 \frac{h}{6} \left( -\frac{4}{h} \right) \left( -\frac{6}{h^2} \right) + \hat{V}_1 \frac{h}{6} \left( -\frac{4}{h} \right) \left( -\frac{6}{h^2} \right) + \hat{V}_2 \frac{h}{6} \left( -\frac{2}{h} \right) \left( -\frac{6}{h^2} \right) \]

\[ = \frac{2}{h^2} \hat{V}_2. \]
5.1. Adjoint Stiffness Matrix Based Approach

\[ B_{12} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_2 \psi_j'' \phi_1'' \, dx \right) \hat{V}_j \]

= \hat{V}_1 \int_{0}^{1} a_2 \psi_j'' \phi_1'' \, dx + \hat{V}_2 \int_{0}^{1} a_2 \psi_j'' \phi_1'' \, dx

= \hat{V}_1 \int_{I_2} a_2 \psi_j'' \phi_1'' \, dx + \hat{V}_2 \int_{I_2} a_2 \psi_j'' \phi_1'' \, dx

= \hat{V}_1 \frac{h}{6} \left[ a_2(x_1) \psi_j''(x_1) \phi_1''(x_1) + 4a_2(x_{3/2}) \psi_j''(x_{3/2}) \phi_1''(x_{3/2}) + a_2(x_2) \psi_j''(x_2) \phi_1''(x_2) \right] + \\
\hat{V}_2 \frac{h}{6} \left[ a_2(x_1) \psi_j''(x_1) \phi_1''(x_1) + 4a_2(x_{3/2}) \psi_j''(x_{3/2}) \phi_1''(x_{3/2}) + a_2(x_2) \psi_j''(x_2) \phi_1''(x_2) \right]

= \hat{V}_1 \frac{h}{6} \left( \frac{2}{h} \right) \left( \frac{6}{h^2} \right) \hat{V}_2 \frac{h}{6} \left( \frac{4}{h} \right) \left( \frac{6}{h^2} \right)

= \frac{2}{h^2} (\hat{V}_1 + 2\hat{V}_2).

\[ B_{1k} = 0 \quad \text{for } k \geq 3. \]

Now let’s look at the 2nd row. We have

\[ B_{20} = 0. \]

\[ B_{21} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_1 \psi_j'' \phi_2'' \, dx \right) \hat{V}_j \]

= \hat{V}_1 \int_{0}^{1} a_1 \psi_j'' \phi_2'' \, dx + \hat{V}_2 \int_{0}^{1} a_1 \psi_j'' \phi_2'' \, dx

= \hat{V}_1 \int_{I_2} a_1 \psi_j'' \phi_2'' \, dx + \hat{V}_2 \int_{I_2} a_1 \psi_j'' \phi_2'' \, dx

= \hat{V}_1 \frac{h}{6} \left[ a_1(x_1) \psi_j''(x_1) \phi_2''(x_1) + 4a_1(x_{3/2}) \psi_j''(x_{3/2}) \phi_2''(x_{3/2}) + a_1(x_2) \psi_j''(x_2) \phi_2''(x_2) \right] + \\
+ \hat{V}_2 \frac{h}{6} \left[ a_1(x_1) \psi_j''(x_1) \phi_2''(x_1) + 4a_1(x_{3/2}) \psi_j''(x_{3/2}) \phi_2''(x_{3/2}) + a_1(x_2) \psi_j''(x_2) \phi_2''(x_2) \right]

= \hat{V}_1 \frac{h}{6} \left( \frac{4}{h} \right) \left( \frac{6}{h^2} \right) + \hat{V}_2 \frac{h}{6} \left( \frac{2}{h} \right) \left( \frac{6}{h^2} \right)

= \frac{2}{h^2} (\hat{V}_1 - \hat{V}_2). \]
\[ B_{22} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_2 \psi_j'' \psi_2'' dx \right) \hat{V}_j \]

\[ = \hat{V}_1 \int_{0}^{1} a_2 \psi_1'' \phi_2'' dx + \hat{V}_2 \int_{0}^{1} a_2 \psi_2'' \phi_2'' dx + \hat{V}_3 \int_{0}^{1} a_2 \psi_3'' \phi_2'' dx \]

\[ = \hat{V}_1 \int_{I_2} a_2 \psi_1'' \phi_1'' dx + \hat{V}_2 \int_{I_2} a_2 \psi_2'' \phi_2'' dx + \hat{V}_3 \int_{I_2} a_2 \psi_3'' \phi_2'' dx \]

\[ = \hat{V}_1 \frac{h}{6} \left[ a_2(x_1) \psi_1''(x_1) \phi_2''(x_1) + 4a_2(x_{3/2}) \psi_1''(x_{3/2}) \phi_2''(x_{3/2}) + a_2(x_2) \psi_2''(x_2) \phi_2''(x_2) \right] + \]

\[ + \hat{V}_2 \frac{h}{6} \left[ a_2(x_2) \psi_2''(x_2) \phi_2''(x_2) + 4a_2(x_{5/2}) \psi_2''(x_{5/2}) \phi_2''(x_{5/2}) + a_2(x_3) \psi_3''(x_3) \phi_2''(x_3) \right] + \]

\[ + \hat{V}_3 \frac{h}{6} \left[ a_2(x_3) \psi_3''(x_3) \phi_2''(x_3) + 4a_2(x_{7/2}) \psi_3''(x_{7/2}) \phi_2''(x_{7/2}) + a_2(x_4) \psi_4''(x_4) \phi_2''(x_4) \right] \]

\[ = \hat{V}_1 \frac{h}{6} \left( -\frac{6}{h} \right) + \hat{V}_2 \frac{h}{6} \left( -\frac{4}{h} \right) + \hat{V}_3 \frac{h}{6} \left( -\frac{2}{h} \right) \]

\[ = \frac{2}{h^2} \left( -\hat{V}_1 + \hat{V}_3 \right) . \]

\[ B_{23} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_3 \psi_j'' \phi_2'' dx \right) \hat{V}_j \]

\[ = \hat{V}_2 \int_{0}^{1} a_3 \psi_2'' \phi_2'' dx + \hat{V}_3 \int_{0}^{1} a_3 \psi_3'' \phi_2'' dx \]

\[ = \hat{V}_2 \int_{I_3} a_3 \psi_2'' \phi_2'' dx + \hat{V}_3 \int_{I_3} a_3 \psi_3'' \phi_2'' dx \]

\[ = \hat{V}_2 \frac{h}{6} \left[ a_3(x_2) \psi_2''(x_2) \phi_2''(x_2) + 4a_3(x_{5/2}) \psi_2''(x_{5/2}) \phi_2''(x_{5/2}) + a_3(x_3) \psi_3''(x_3) \phi_2''(x_3) \right] + \]

\[ + \hat{V}_3 \frac{h}{6} \left[ a_3(x_3) \psi_3''(x_3) \phi_2''(x_3) + 4a_3(x_{7/2}) \psi_3''(x_{7/2}) \phi_2''(x_{7/2}) + a_3(x_4) \psi_4''(x_4) \phi_2''(x_4) \right] \]

\[ = \hat{V}_2 \frac{h}{6} \left( \frac{6}{h} \right) + \hat{V}_3 \frac{h}{6} \left( \frac{6}{h} \right) \]

\[ = \frac{2}{h^2} \left( \hat{V}_2 + \hat{V}_3 \right) . \]

\[ B_{2k} = 0 \quad \text{for } k \geq 4. \]
Furthermore, it can be shown that this pattern goes on until the row \( n-1 \):

\[
\begin{align*}
B_{n-1,k} &= 0 \text{ for } k \leq n-3 \\
B_{n-1,n-2} &= \frac{2}{h^2} (-2\hat{V}_{n-2} - \hat{V}_{n-1}) \\
B_{n-1,n-1} &= \frac{2}{h^2} (-\hat{V}_{n-2} + \hat{V}_n) \\
B_{n-1,n} &= \frac{2}{h^2} (V_{n-1} + 2\hat{V}_n) \\
B_{n-1,n+1} &= 0.
\end{align*}
\]

For the last row, the \( n \)th row,

\[
\begin{align*}
B_{n,k} &= 0 \text{ for } k \leq n-2 \\
B_{n,n-1} &= \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n-1} \psi_j'' \phi_n'' \, dx \right) \psi_j' \\
&= \hat{V}_{n-1} \int_{0}^{1} a_{n-1} \psi_j'' \phi_n'' \, dx + \hat{V}_n \int_{0}^{1} a_{n-1} \psi_j'' \phi_n'' \, dx \\
&= \hat{V}_{n-1} \int_{0}^{l} a_{n-1} \psi_j'' \phi_n'' \, dx + \hat{V}_n \int_{0}^{l} a_{n-1} \psi_j'' \phi_n'' \, dx \\
&= \hat{V}_{n-1} \frac{h}{6} \left[ a_{n-1} (x_{n-1}) \psi_n'' (x_{n-1}) \psi_j'' (x_{n-1}) + 4a_{n-1} (x_{n-1}/2) \psi_j'' (x_{n-1}/2) \right. \\
&\quad \left. + a_{n-1} (x_j) \phi_n'' (x_j) \right] + \\
&\hat{V}_n \frac{h}{6} \left[ a_{n-1} (x_{n-1}) \psi_n'' (x_{n-1}) \psi_j'' (x_{n-1}) + 4a_{n-1} (x_{n-1}/2) \psi_j'' (x_{n-1}/2) \right. \\
&\quad \left. + a_{n-1} (x_j) \phi_n'' (x_j) \right] \\
&= \hat{V}_{n-1} \frac{h}{6} \left( -\frac{4}{h} \right) \left( \frac{6}{h^3} \right) + \hat{V}_n \frac{h}{6} \left( -\frac{2}{h^2} \right) \left( \frac{6}{h^2} \right) \\
&= \frac{2}{h^2} (-2\hat{V}_{n-1} - \hat{V}_n)
\end{align*}
\]
\[ B_{n,n} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_n \psi_j'' \psi_n'' \, dx \right) V_j' \]

\[ = \hat{V}_{n-1} \int_{0}^{1} a_n \psi_{n-1}'' \phi_n'' \, dx + \hat{V}_n \int_{0}^{1} a_n \psi_n'' \phi_n'' \, dx \]

\[ = \hat{V}_{n-1} \int_{I_n} a_n \psi_{n-1}'' \phi_n'' \, dx + \hat{V}_n \int_{I_n} a_n \psi_n'' \phi_n'' \, dx + \hat{V}_n \int_{I_{n+1}} a_n \psi_n'' \phi_n'' \, dx \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ a_n(x_{n-1}) \psi_{n-1}''(x_{n-1}) \phi_n''(x_{n-1}) + 4a_n(x_{n-1}/2) \psi_{n-1}''(x_{n-1}/2) \phi_n''(x_{n-1}/2) \right. \]

\[ + a_n(x_n) \psi_n''(x_n) \phi_n''(x_n) \left. \right] \]

\[ + \hat{V}_n \frac{h}{6} \left[ a_n(x_{n-1}) \psi_{n-1}''(x_{n-1}) \phi_n''(x_{n-1}) + 4a_n(x_{n-1}/2) \psi_{n-1}''(x_{n-1}/2) \phi_n''(x_{n-1}/2) \right. \]

\[ + a_n(x_n) \psi_n''(x_n) \phi_n''(x_n) \left. \right] \]

\[ + \hat{V}_n \frac{h}{6} \left[ a_n(x_{n+1}) \psi_{n+1}''(x_{n+1}) \phi_n''(x_{n+1}) + 4a_n(x_{n+1}/2) \psi_{n+1}''(x_{n+1}/2) \phi_n''(x_{n+1}/2) \right. \]

\[ + a_n(x_{n+1}) \psi_n''(x_{n+1}) \phi_n''(x_{n+1}) \left. \right] \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left( \frac{2}{h} \right) \left( -\frac{6}{h^2} \right) + \hat{V}_n \frac{h}{6} \left( \frac{4}{h} \right) \left( -\frac{6}{h^2} \right) + \hat{V}_n \frac{h}{6} \left( \frac{4}{h} \right) \left( -\frac{6}{h^2} \right) \]

\[ = -\frac{2}{h^2} \hat{V}_{n-1} \]

\[ B_{n,n+1} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n+1} \psi_j'' \psi_n'' \, dx \right) V_j' \]

\[ = \hat{V}_n \int_{0}^{1} a_{n+1} \psi_n'' \phi_n'' \, dx \]

\[ = \hat{V}_n \int_{I_{n+1}} a_{n+1} \psi_n'' \phi_n'' \, dx \]

\[ = \hat{V}_n \frac{h}{6} \left[ a_{n+1}(x_n) \psi_n''(x_n) \phi_n''(x_n) + 4a_{n+1}(x_{n+1}/2) \psi_n''(x_{n+1}/2) \phi_n''(x_{n+1}/2) \right. \]

\[ + a_{n+1}(x_{n+1}) \psi_n''(x_{n+1}) \phi_n''(x_{n+1}) \left. \right] \]

\[ = \hat{V}_n \frac{h}{6} \left( \frac{2}{h} \right) \left( \frac{6}{h^2} \right) \]

\[ = \frac{2}{h^2} \hat{V}_n \]
5.1. Adjoint Stiffness Matrix Based Approach

This will produce the following matrix:

\[
B = \frac{2}{\hat{h}^2} \begin{pmatrix}
-\hat{V}_1 & \hat{V}_2 & \hat{V}_1 + 2\hat{V}_2 & 0 & 0 & \ldots & 0 \\
0 & -2\hat{V}_1 - \hat{V}_2 & -\hat{V}_1 + \hat{V}_3 & \hat{V}_2 + 2\hat{V}_3 & 0 & \ldots & 0 \\
0 & 0 & -2\hat{V}_2 - \hat{V}_3 & -\hat{V}_2 + \hat{V}_4 & \hat{V}_3 + 2\hat{V}_4 & \ldots & 0 \\
0 & \ldots & 0 & -2\hat{V}_{n-2} - \hat{V}_{n-1} & -\hat{V}_{n-2} + \hat{V}_n & \hat{V}_{n-1} + 2\hat{V}_n & 0 \\
0 & \ldots & 0 & 0 & -2\hat{V}_{n-1} - \hat{V}_n & -\hat{V}_{n-1} + \hat{V}_n & 
\end{pmatrix}
\]

Now for the entries of the matrix \( D \)

\[
D_{ik} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_k \phi_j'' \psi_i'' \, dx \right) V_j \quad \text{for} \quad i = 1, \ldots, n \quad \text{and} \quad k = 0, \ldots, n + 1.
\]

\[
D_{10} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_0 \phi_0'' \psi_i'' \, dx \right) V_j 
\]

\[
= V_1 \int_{l_1}^{h/2} a_0 \phi_0'' \psi_1'' \, dx 
\]

\[
= V_1 \frac{h/2}{3} \left[ a_0 (x_0) \phi_0''(x_0) \psi_1''(x_0) + 4a_0 (x_{1/2}) \phi_1''(x_{1/2}) \psi_1''(x_{1/2}) + a_0 (x_1) \phi_1''(x_1) \psi_1''(x_1) \right] 
\]

\[
= V_1 \frac{h}{6} \left[ \left( \frac{6}{\hat{h}^2} \right) \left( -\frac{2}{\hat{h}} \right) \right] 
\]

\[
= \frac{2}{\hat{h}^2} V_1. 
\]

\[
D_{11} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_1 \phi_1'' \psi_1'' \, dx \right) V_j 
\]

\[
= V_1 \int_{0}^{1} a_1 \phi_1'' \psi_1'' \, dx + V_2 \int_{0}^{1} a_1 \phi_2'' \psi_1'' \, dx 
\]

\[
= V_1 \int_{l_1}^{h/2} a_1 \phi_1'' \psi_1'' \, dx + V_1 \int_{l_2}^{1/2} a_1 \phi_1'' \psi_1'' \, dx + V_2 \int_{l_2}^{1} a_1 \phi_2'' \psi_1'' \, dx 
\]

\[
= V_1 \frac{h}{6} \left[ a_1 (x_0) \phi_0''(x_0) \psi_1''(x_0) + 4a_1 (x_{1/2}) \phi_1''(x_{1/2}) \psi_1''(x_{1/2}) + a_1 (x_1) \phi_1''(x_1) \psi_1''(x_1) \right] + 
\]

\[
V_1 \frac{h}{6} \left[ a_1 (x_1) \phi_1''(x_1) \psi_1''(x_1) + 4a_1 (x_{3/2}) \phi_1''(x_{3/2}) \psi_1''(x_{3/2}) + a_1 (x_2) \phi_1''(x_2) \psi_1''(x_2) \right] + 
\]

\[
V_2 \frac{h}{6} \left[ a_1 (x_1) \phi_2''(x_1) \psi_1''(x_1) + 4a_1 (x_{3/2}) \phi_2''(x_{3/2}) \psi_1''(x_{3/2}) + a_1 (x_2) \phi_2''(x_2) \psi_1''(x_2) \right] 
\]

\[
= V_1 \frac{h}{6} \left( -\frac{6}{\hat{h}^2} \right) \left( \frac{4}{\hat{h}^2} \right) + V_1 \frac{h}{6} \left( -\frac{6}{\hat{h}^2} \right) \left( -\frac{4}{\hat{h}^2} \right) + V_2 \frac{h}{6} \left( \frac{6}{\hat{h}^2} \right) \left( -\frac{4}{\hat{h}^2} \right) 
\]

\[
= -\frac{4}{\hat{h}^2} V_2. 
\]
5.1. Adjoint Stiffness Matrix Based Approach

\[ D_{12} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_2 \phi_j'' \psi_1'' dx \right) V_j \]

\[ = V_1 \int_{0}^{1} a_2 \phi_1'' \psi_1'' dx + V_2 \int_{0}^{1} a_2 \phi_2'' \psi_1'' dx \]

\[ = V_1 \int_{l_2}^{l_1} a_2 \phi_1'' \psi_1'' dx + V_2 \int_{l_2}^{l_1} a_2 \phi_2'' \psi_1'' dx \]

\[ = V_1 \frac{h}{6} \left[ a_2 (x_1) \phi_1''(x_1) \psi_1''(x_1) + 4a_2 (x_{3/2}) \phi_1''(x_{3/2}) \psi_1''(x_{3/2}) + a_2 (x_2) \phi_1''(x_2) \psi_1''(x_2) \right] + \]

\[ V_2 \frac{h}{6} \left[ a_2 (x_1) \phi_2''(x_1) \psi_1''(x_1) + 4a_2 (x_{3/2}) \phi_2''(x_{3/2}) \psi_1''(x_{3/2}) + a_2 (x_2) \phi_2''(x_2) \psi_1''(x_2) \right] \]

\[ = V_1 \frac{h}{6} \left( \frac{6}{h^2} \right) \left( \frac{2}{h} \right) + V_2 h \left( -\frac{6}{h^2} \right) \left( \frac{2}{h} \right) \]

\[ D_{1k} = 0 \text{ for } k \geq 3. \]

Now let's look at the 2nd row:

\[ D_{20} = 0. \]

\[ D_{21} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_1 \phi_j'' \psi_2'' dx \right) V_j \]

\[ = V_1 \int_{0}^{1} a_1 \phi_1'' \psi_2'' dx + V_2 \int_{0}^{1} a_1 \phi_2'' \psi_2'' dx \]

\[ = V_1 \int_{l_2}^{l_1} a_1 \phi_1'' \psi_2'' dx + V_2 \int_{l_2}^{l_1} a_1 \phi_2'' \psi_2'' dx \]

\[ = V_1 \frac{h}{6} \left[ a_1 (x_1) \phi_1''(x_1) \psi_2''(x_1) + 4a_1 (x_{3/2}) \phi_1''(x_{3/2}) \psi_2''(x_{3/2}) + a_1 (x_2) \phi_1''(x_2) \psi_2''(x_2) \right] + \]

\[ V_2 \frac{h}{6} \left[ a_1 (x_1) \phi_2''(x_1) \psi_2''(x_1) + 4a_1 (x_{3/2}) \phi_2''(x_{3/2}) \psi_2''(x_{3/2}) + a_1 (x_2) \phi_2''(x_2) \psi_2''(x_2) \right] \]

\[ = V_1 \frac{h}{6} \left[ \left( -\frac{6}{h^2} \right) \left( \frac{2}{h} \right) \right] + V_2 \frac{h}{6} \left[ \left( \frac{6}{h^2} \right) \left( -\frac{2}{h} \right) \right] \]

\[ = \frac{2}{h^2} (V_1 - V_2). \]
\[ D_{22} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_2 \phi_j'' \psi_j'' dx \right) V_j \]

\[ = V_1 \int_{0}^{1} a_2 \phi_1'' \psi_1'' dx + V_2 \int_{0}^{1} a_2 \phi_2'' \psi_2'' dx + V_3 \int_{0}^{1} a_2 \phi_3'' \psi_3'' dx \]

\[ = V_1 \int_{I_2} a_2 \phi_1'' \psi_1'' dx + V_2 \int_{I_2} a_2 \phi_2'' \psi_2'' dx + V_2 \int_{I_3} a_2 \phi_2'' \psi_2'' dx + V_3 \int_{I_3} a_2 \phi_3'' \psi_3'' dx \]

\[ = V_1 \frac{h}{6} \left[ a_2(x_1) \phi_1''(x_1) \psi_2''(x_1) + 4a_2(x_3/2) \phi_1''(x_3/2) \psi_2''(x_3/2) + a_2(x_2) \phi_1''(x_2) \psi_2''(x_2) \right] + \]

\[ V_2 \frac{h}{6} \left[ a_2(x_1) \phi_2''(x_1) \psi_2''(x_1) + 4a_2(x_3/2) \phi_2''(x_3/2) \psi_2''(x_3/2) + a_2(x_2) \phi_2''(x_2) \psi_2''(x_2) \right] + \]

\[ V_2 \frac{h}{6} \left[ a_2(x_2) \phi_2''(x_2) \psi_2''(x_2) + 4a_2(x_5/2) \phi_2''(x_5/2) \psi_2''(x_5/2) + a_2(x_3) \phi_2''(x_3) \psi_2''(x_3) \right] + \]

\[ V_3 \frac{h}{6} \left[ a_2(x_2) \phi_3''(x_2) \psi_2''(x_2) + 4a_2(x_5/2) \phi_3''(x_5/2) \psi_2''(x_5/2) + a_2(x_3) \phi_3''(x_3) \psi_2''(x_3) \right] \]

\[ = V_1 \frac{h}{6} \left( \frac{6}{h^2} \right) \left( \frac{4}{h^2} \right) + V_2 \frac{h}{6} \left( - \frac{6}{h^2} \right) \left( \frac{4}{h^2} \right) + V_2 \frac{h}{6} \left( - \frac{6}{h^2} \right) \left( - \frac{4}{h^2} \right) + V_3 \frac{h}{6} \left( \frac{6}{h^2} \right) \left( - \frac{4}{h^2} \right) \]

\[ = \frac{2}{h^2} (2V_1 - 2V_3). \]

\[ D_{23} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_3 \phi_j'' \psi_j'' dx \right) V_j \]

\[ = V_2 \int_{0}^{1} a_3 \phi_2'' \psi_2'' dx + V_3 \int_{0}^{1} a_3 \phi_3'' \psi_2'' dx \]

\[ = V_2 \int_{I_3} a_3 \phi_2'' \psi_2'' dx + V_3 \int_{I_3} a_3 \phi_3'' \psi_2'' dx \]

\[ = V_2 \frac{h}{6} \left[ a_3(x_2) \phi_2''(x_2) \psi_2''(x_2) + 4a_3(x_5/2) \phi_2''(x_5/2) \psi_2''(x_5/2) + a_3(x_3) \phi_2''(x_3) \psi_2''(x_3) \right] + \]

\[ V_3 \frac{h}{6} \left[ a_3(x_2) \phi_3''(x_2) \psi_2''(x_2) + 4a_3(x_5/2) \phi_3''(x_5/2) \psi_2''(x_5/2) + a_3(x_3) \phi_3''(x_3) \psi_2''(x_3) \right] \]

\[ = V_2 \frac{h}{6} \left( \frac{6}{h^2} \right) \left( \frac{2}{h} \right) + V_3 \frac{h}{6} \left( - \frac{6}{h^2} \right) \left( \frac{2}{h} \right) \]

\[ = \frac{2}{h^2} (V_2 - V_3). \]

\[ D_{2k} = 0 \text{ for } k \geq 4. \]
5.1. Adjoint Stiffness Matrix Based Approach

It can be shown that this pattern goes on until the $N - 1$-st row,

\[ D_{n-1,k} = 0 \text{ for } k \leq n - 3 \]
\[ D_{n-1,n-2} = \frac{2}{h^2} (V_{n-2} - V_{n-1}) \]
\[ D_{n-1,n-1} = \frac{2}{h^2} (2V_{n-2} - 2V_n) \]
\[ D_{n-1,n} = \frac{2}{h^2} (V_{n-1} - V_n) \]
\[ D_{n-1,n+1} = 0. \]

For the last row, the $n$-th row,

\[ D_{n,k} = 0 \text{ for } k \leq n - 2 \]
\[ D_{n,n-1} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n-1} \phi''_j \psi''_n dx \right) V_j \]
\[ = V_{n-1} \int_{0}^{1} a_{n-1} \phi''_{n-1} \psi''_n dx + V_n \int_{0}^{1} a_{n-1} \phi''_n \psi''_n dx \]
\[ = V_{n-1} \int_{I_n} a_{n-1} \phi''_{n-1} \psi''_n dx + V_n \int_{I_n} a_{n-1} \phi''_n \psi''_n dx \]
\[ = V_{n-1} \frac{h}{6} \left[ a_{n-1} (x_{n-1}) \phi''_{n-1}(x_{n-1}) \psi''_n(x_{n-1}) + 4a_{n-1} (x_{n-1}/2) \phi''_{n-1}(x_{n-1}/2) \psi''_n(x_{n-1}/2) \right. \]
\[ + a_{n-1} (x_n) \phi''_{n-1}(x_n) \psi''_n(x_n) \] 
\[ + V_n \frac{h}{6} \left[ a_{n-1} (x_{n-1}) \phi''_{n}(x_{n-1}) \psi''_n(x_{n-1}) \right. \]
\[ + 4a_{n-1} (x_{3/2}) \phi''_{n}(x_{3/2}) \psi''_n(x_{3/2}) + a_{n-1} (x_n) \phi''_{n}(x_n) \psi''_n(x_n) \] 
\[ = V_{n-1} \frac{h}{6} \left[ -\frac{6}{h^2} \left( -\frac{2}{h} \right) \right] + V_n \frac{h}{6} \left[ \frac{6}{h^2} \left( -\frac{2}{h} \right) \right] \]
\[ = \frac{2}{h^2} (V_{n-1} - V_n). \]
5.1. Adjoint Stiffness Matrix Based Approach

\[ D_{n,n} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_n \phi_j'' \psi_n'' \, dx \right) V_j \]

\[ = V_{n-1} \int_{0}^{1} a_n \phi_{n-1}' \psi_n' \, dx + V_n \int_{0}^{1} a_n \phi_n'' \psi_n'' \, dx \]

\[ = V_{n-1} \int_{I_n} a_n \phi_n'' \psi_n'' \, dx + V_n \int_{I_n} a_n \phi_n'' \psi_n'' \, dx + V_n \int_{I_{n+1}} a_n \phi_n'' \psi_n'' \, dx + \]

\[ = V_{n-1} \frac{h}{6} \left[ a_n (x_{n-1}) \phi_n'' (x_{n-1}) \psi_n'' (x_{n-1}) + 4a_n (x_{n-1/2}) \phi_n'' (x_{n-1/2}) \psi_n'' (x_{n-1/2}) + a_n (x_{n}) \phi_n'' (x_{n}) \psi_n'' (x_{n}) \right] \]

\[ + V_n \frac{h}{6} \left[ a_n (x_{n-1}) \phi_n'' (x_{n-1}) \psi_n'' (x_{n-1}) + 4a_n (x_{n-1/2}) \phi_n'' (x_{n-1/2}) \psi_n'' (x_{n-1/2}) + a_n (x_{n}) \phi_n'' (x_{n}) \psi_n'' (x_{n}) \right] \]

\[ + V_n \frac{h}{6} \left[ a_n (x_{n}) \phi_n'' (x_{n}) \psi_n'' (x_{n}) + 4a_n (x_{n+1/2}) \phi_n'' (x_{n+1/2}) \psi_n'' (x_{n+1/2}) + a_n (x_{n+1}) \phi_n'' (x_{n+1}) \psi_n'' (x_{n+1}) \right] \]

\[ = V_{n-1} \frac{h}{6} \left( \frac{6}{h^2} \right) \left( \frac{4}{h^2} \right) + V_n \frac{h}{6} \left( -\frac{6}{h^2} \right) \left( \frac{4}{h^2} \right) + V_n \frac{h}{6} \left( -\frac{6}{h^2} \right) \left( -\frac{4}{h^2} \right) \]

\[ = \frac{2}{h^2} \cdot 2V_{n-1}. \]

\[ D_{n,n+1} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n+1} \phi_j'' \psi_n'' \, dx \right) V_j \]

\[ = V_n \int_{0}^{1} a_{n+1} \phi_n'' \psi_n'' \, dx \]

\[ = V_n \int_{I_{n+1}} a_n \phi_n'' \psi_n'' \, dx \]

\[ = V_n \frac{h}{6} \left[ a_{n+1} (x_n) \phi_n'' (x_n) \psi_n'' (x_n) + 4a_{n+1} (x_{n+1/2}) \phi_n'' (x_{n+1/2}) \psi_n'' (x_{n+1/2}) + a_{n+1} (x_{n+1}) \phi_n'' (x_{n+1}) \psi_n'' (x_{n+1}) \right] \]

\[ = V_n \frac{h}{6} \left( \frac{6}{h^2} \right) \left( \frac{2}{h} \right) \]

\[ = \frac{2}{h^2} \cdot V_n. \]
5.1. Adjoint Stiffness Matrix Based Approach

Putting the entries together we get

\[
D = \frac{2}{h^2} \begin{pmatrix}
-V_1 & -2V_2 & V_1 - V_2 & 0 & 0 & \ldots & 0 \\
0 & V_1 - V_2 & 2V_1 - 2V_3 & V_2 - V_3 & 0 & \ldots & 0 \\
0 & \ldots & V_2 - V_3 & -V_2 - 2V_4 & V_3 - V_4 & \ldots & 0 \\
0 & \ldots & 0 & V_{n-2} - V_{n-1} & 2V_{n-2} - 2V_n & V_{n-1} - V_n & 0 \\
0 & \ldots & 0 & 0 & V_{n-1} - V_n & 2V_{n-1} & V_n \\
\end{pmatrix}.
\]

Now for the entries of the \(C\) matrix we have the formula

\[
C_{ik} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_k \psi_j'' \psi_i'' dx \right) \hat{V}_j \text{ for } i = 1, \ldots, n \text{ and } k = 0, \ldots, n + 1.
\]

Therefore, \(C_{10} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_0 \psi_j'' \psi_1'' dx \right) \hat{V}_j\)

\[
= \hat{V}_1 \int_{l_1} a_0 \left( \psi_1'' \right)^2 dx
= \hat{V}_1 \frac{h}{2} \left[ a_0(x_0) \left( \psi_1''(x_0) \right)^2 + 4a_0(x_{1/2}) \left( \psi_1''(x_{1/2}) \right)^2 + a_0(x_1) \left( \psi_1''(x_1) \right)^2 \right]
= \hat{V}_1 \frac{h}{6} \left[ \left( \frac{2}{h} \right)^2 + 4 \left( \frac{1}{2} \right) \left( \frac{1}{h} \right)^2 + 0 \right]
= \hat{V}_1 \frac{h}{6} \left[ \frac{4}{h^2} + \frac{2}{h^2} \right]
= \frac{1}{h} \hat{V}_1.
\]
5.1. Adjoint Stiffness Matrix Based Approach

\[ C_{11} = \sum_{j=1}^{n} \left( \int_0^1 a_1 \psi_j''(x) \psi_j''(x) \, dx \right) V_j' 
= \hat{V}_1 \int_0^1 a_1 (\psi_j')^2 \, dx + \hat{V}_2 \int_0^1 a_1 \psi_j'' \psi_j' \, dx 
= \hat{V}_1 \int_{l_1} a_1 (\psi_j')^2 \, dx + \hat{V}_2 \int_{l_2} a_1 \psi_j'' \psi_j' \, dx 
= \hat{V}_1 \frac{h}{6} \left[ a_1(x_0) \left( \psi_j''(x_0) \right)^2 + 4a_1(x_{1/2}) \left( \psi_j''(x_{1/2}) \right)^2 + a_1(x_1) \left( \psi_j''(x_1) \right)^2 \right] + 
\hat{V}_1 \frac{h}{6} \left[ a_1(x_1) \left( \psi_j''(x_1) \right)^2 + 4a_1(x_{3/2}) \left( \psi_j''(x_{3/2}) \right)^2 + a_1(x_2) \left( \psi_j''(x_2) \right)^2 \right] + 
\hat{V}_2 \frac{h}{6} \left[ a_1(x_1) \psi_j'(x_1) \psi_j''(x_1) + 4a_1(x_{3/2}) \psi_j'(x_{3/2}) \psi_j''(x_{3/2}) + a_1(x_2) \psi_j'(x_2) \psi_j''(x_2) \right] 
= \hat{V}_1 \frac{h}{6} \left[ 4 \left( \frac{1}{2} \right) \left( \frac{1}{h} \right)^2 + \left( \frac{4}{h} \right)^2 \right] + \hat{V}_1 \frac{h}{6} \left[ -4 \left( \frac{1}{h} \right)^2 + 4 \left( \frac{1}{2} \right) \left( \frac{1}{h} \right)^2 \right] + 
\hat{V}_2 \frac{h}{6} \left[ 0 \frac{2}{h^2} + \left( \frac{16}{h^2} \right) + \frac{2}{h^2} \right] + \hat{V}_2 \frac{h}{6} \left[ \frac{8}{h^2} + \frac{2}{h^2} \right] 
= \hat{V}_1 \frac{h}{6} \left[ 36 \frac{h^2}{h^2} \right] + \hat{V}_2 \frac{h}{6} \left[ 6 \frac{h^2}{h^2} \right] 
= \frac{1}{h} (6\hat{V}_1 + \hat{V}_2) . 
\]

\[ C_{12} = \sum_{j=1}^{n} \left( \int_0^1 a_2 \psi_j''(x) \psi_j'(x) \, dx \right) V_j 
= \hat{V}_1 \int_0^1 a_2 (\psi_j')^2 \, dx + \hat{V}_2 \int_0^1 a_2 \psi_j'' \psi_j' \, dx 
= \hat{V}_1 \int_{l_1} a_2 (\psi_j')^2 \, dx + \hat{V}_2 \int_{l_2} a_2 \psi_j'' \psi_j' \, dx 
= \hat{V}_1 \frac{h}{6} \left[ a_2(x_1) \left( \psi_j''(x_1) \right)^2 + 4a_2(x_{3/2}) \left( \psi_j''(x_{3/2}) \right)^2 + a_2(x_2) \left( \psi_j''(x_2) \right)^2 \right] + 
\hat{V}_2 \frac{h}{6} \left[ a_2(x_1) \psi_j'(x_1) \psi_j''(x_1) + 4a_2(x_{3/2}) \psi_j'(x_{3/2}) \psi_j''(x_{3/2}) + a_2(x_2) \psi_j'(x_2) \psi_j''(x_2) \right] 
= \hat{V}_1 \frac{h}{6} \left[ 0 \frac{2}{h^2} + \left( \frac{4}{h^2} \right) \right] + \hat{V}_2 \frac{h}{6} \left[ 0 \frac{2}{h^2} + \frac{8}{h^2} \right] 
= \frac{1}{h} (\hat{V}_1 + \hat{V}_2) . 
\]

\[ C_{1k} = 0 \text{ for } k \geq 3. \]
5.1. Adjoint Stiffness Matrix Based Approach

Now we look at the second row:

\[ C_{20} = 0. \]
\[ C_{21} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{1} \psi''_{j} \psi''_{2} \, dx \right) V_{j}' \]
\[ = V_{1} \int_{0}^{1} a_{1} \psi''_{1} \psi''_{2} \, dx + V_{2} \int_{0}^{1} a_{1} (\psi''_{j})^{2} \, dx \]
\[ = V_{1} \int_{I_{2}} a_{1} \psi''_{1} \psi''_{2} \, dx + V_{2} \int_{I_{2}} a_{1} (\psi''_{j})^{2} \, dx \]
\[ = \check{V}_{1} \frac{h}{6} [a_{1}(x_{1}) \psi''_{1}(x_{1}) \psi''_{2}(x_{1}) + 4a_{1}(x_{3/2}) \psi''_{1}(x_{3/2}) \psi''_{2}(x_{3/2}) + a_{1}(x_{2}) \psi''_{1}(x_{2}) \psi''_{2}(x_{2})] \]
\[ + \check{V}_{2} \frac{h}{6} [a_{1}(x_{1}) [\psi''_{1}(x_{1})]^{2} + 4a_{1}(x_{3/2}) [\psi''_{2}(x_{3/2})]^{2} + a_{1}(x_{2}) [\psi''_{2}(x_{2})]^{2}] \]
\[ = \check{V}_{1} \frac{h}{6} \left[ \left( \frac{-4}{h} \right) \left( \frac{-2}{h} \right) + 4 \left( \frac{1}{2} \right) \left( \frac{-1}{h} \right) \left( \frac{1}{h} \right) \right] + \check{V}_{2} \frac{h}{6} \left[ \left( \frac{-2}{h} \right)^{2} + 4 \left( \frac{1}{2} \right) \left( \frac{1}{h} \right)^{2} \right] \]
\[ = \check{V}_{1} \frac{h}{6} \left[ \frac{8}{h^{2}} + \frac{-2}{h^{2}} \right] + \check{V}_{2} \frac{h}{6} \left[ \frac{4}{h^{2}} + \frac{2}{h^{2}} \right] \]
\[ = \frac{1}{h} (\check{V}_{1} + \check{V}_{2}) \cdot \check{V}_{3} \]
\[ C_{22} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{2} \psi''_{j} \psi''_{2} \, dx \right) V_{j} \]
\[ = \check{V}_{1} \int_{0}^{1} a_{2} \psi''_{1} \psi''_{2} \, dx + \check{V}_{2} \int_{0}^{1} a_{2} (\psi''_{j})^{2} \, dx + \check{V}_{3} \int_{0}^{1} a_{2} \psi''_{3} \psi''_{2} \, dx \]
\[ = \check{V}_{1} \int_{I_{2}} a_{2} \psi''_{1} \psi''_{2} \, dx + \check{V}_{2} \int_{I_{2}} a_{2} (\psi''_{j})^{2} \, dx + \check{V}_{2} \int_{I_{3}} a_{2} \psi''_{3} \psi''_{2} \, dx \]
\[ = \check{V}_{1} \frac{h}{6} \left[ a_{2}(x_{1}) \psi''_{2}(x_{1}) \psi''_{2}(x_{1}) + 4a_{2}(x_{3/2}) \psi''_{1}(x_{3/2}) \psi''_{2}(x_{3/2}) + a_{2}(x_{2}) \psi''_{1}(x_{2}) \psi''_{2}(x_{2}) \right] + \]
\[ \check{V}_{2} \frac{h}{6} \left[ a_{2}(x_{1}) [\psi''_{2}(x_{1})]^{2} + 4a_{2}(x_{3/2}) [\psi''_{2}(x_{3/2})]^{2} + a_{2}(x_{2}) [\psi''_{2}(x_{2})]^{2} \right] + \]
\[ \check{V}_{3} \frac{h}{6} \left[ a_{2}(x_{2}) [\psi''_{2}(x_{2})]^{2} + 4a_{2}(x_{3/2}) [\psi''_{2}(x_{3/2})]^{2} + a_{2}(x_{3}) [\psi''_{2}(x_{3})]^{2} \right] \]
\[ = \check{V}_{1} \frac{h}{6} \left[ 4 \left( \frac{1}{2} \right) \left( \frac{-1}{h} \right) \left( \frac{1}{h} \right) + \left( \frac{2}{h} \right) \left( \frac{4}{h} \right) \right] + \check{V}_{2} \frac{h}{6} \left[ 4 \left( \frac{1}{2} \right) \left( \frac{1}{h} \right)^{2} + \left( \frac{4}{h} \right)^{2} \right] \]
\[ + \check{V}_{3} \frac{h}{6} \left[ \left( \frac{-4}{h} \right)^{2} + 4 \left( \frac{1}{2} \right) \left( \frac{-1}{h} \right)^{2} \right] \]
\[ = \check{V}_{1} \frac{h}{6} \left[ \frac{2}{h^{2}} + \frac{8}{h^{2}} \right] + \check{V}_{2} \frac{h}{6} \left[ \frac{2}{h^{2}} + \frac{16}{h^{2}} + \frac{16}{h^{2}} + \frac{2}{h^{2}} \right] + \check{V}_{3} \frac{h}{6} \left[ \frac{8}{h^{2}} - \frac{2}{h^{2}} \right] \]
\[ = \check{V}_{1} \frac{h}{6} \left[ \frac{6}{h^{2}} \right] + \check{V}_{2} \frac{h}{6} \left[ \frac{36}{h^{2}} \right] + \check{V}_{3} \frac{h}{6} \left[ \frac{6}{h^{2}} \right]. \]
\[
C_{23} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_3 \psi_j'' \psi_2'' \, dx \right) \hat{V}_j
\]
\[
= \hat{V}_2 \int_{0}^{1} a_3 (\psi_2'')^2 \, dx + \hat{V}_3 \int_{0}^{1} a_3 \psi_3'' \psi_2'' \, dx
\]
\[
= \hat{V}_2 \int_{I_3} a_3 (\psi_2'')^2 \, dx + \hat{V}_3 \int_{I_3} a_3 \psi_3'' \psi_2'' \, dx
\]
\[
= \hat{V}_2 \frac{h}{6} \left[ a_3(x_2) \psi_2''(x_2)^2 + 4a_3(x_{5/2}) \psi_2''(x_{5/2})^2 + a_3(x_3) \psi_2''(x_3)^2 \right] + \hat{V}_3 \frac{h}{6} \left[ a_3(x_2) \psi_3''(x_2) \psi_2''(x_2) + 4a_3(x_{5/2}) \psi_3''(x_{5/2}) \psi_2''(x_{5/2}) + a_3(x_3) \psi_3''(x_3) \psi_2''(x_3) \right]
\]
\[
= \hat{V}_2 \frac{h}{6} \left[ \frac{1}{2} - \frac{1}{h} \right]^2 + \hat{V}_3 \frac{h}{6} \left[ \frac{1}{2} \frac{1}{h} \frac{1}{h} \frac{4}{h} \frac{2}{h} \right]
\]
\[
= \hat{V}_2 \frac{h}{6} \left[ \frac{2}{h^2} + \frac{4}{h^2} \right] + \hat{V}_3 \frac{h}{6} \left[ -\frac{2}{h^2} + \frac{8}{h^2} \right]
\]
\[
= \frac{1}{h} (\hat{V}_2 + \hat{V}_3)
\]
\[
C_{2k} = 0 \text{ for } k \geq 4.
\]

It can be shown that this pattern goes on until the \(n-1\) row,

\[
C_{n-1,k} = 0 \text{ for } k \leq n-3
\]
\[
C_{n-1,n-2} = \frac{1}{h} (\hat{V}_{n-2} + \hat{V}_{n-1})
\]
\[
C_{n-1,n-1} = \frac{1}{h} (\hat{V}_{n-2} + 6\hat{V}_{n-1} + \hat{V}_n)
\]
\[
C_{n-1,n} = \frac{1}{h} (\hat{V}_{n-1} + \hat{V}_n)
\]
\[
C_{n-1,n+1} = 0.
\]

For the last row, the \(n^{th}\) row,

\[
C_{n,k} = 0 \text{ for } k \leq n-2.
\]
\[ C_{n,n-1} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n-1} \psi''_{j} \psi'_{n} \, dx \right) \hat{V}_{j} \]

\[ = \hat{V}_{n-1} \int_{0}^{1} a_{n-1} \psi''_{n-1} \psi'_{n} \, dx + V_{n} \int_{0}^{1} a_{n-1} \left( \psi''_{n} \right)^{2} \, dx \]

\[ = \hat{V}_{n-1} \int_{I_{n}} a_{n-1} \psi''_{n-1} \psi'_{n} \, dx + V_{n} \int_{I_{n}} a_{n-1} \left( \psi''_{n} \right)^{2} \, dx \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ a_{n-1} (x_{n-1}) \psi''_{n-1} (x_{n-1}) \psi''_{n} (x_{n-1}) \right] \]

\[ + 4 a_{n-1} (x_{n-1}/2) \psi''_{n-1} (x_{n-1}/2) \psi''_{n} (x_{n-1}/2) + a_{n-1} (x_{n}) \psi''_{n-1} (x_{n}) \psi''_{n} (x_{n}) \]

\[ + \hat{V}_{n-1} \frac{h}{6} \left[ a_{n-1} (x_{n-1}) \left( \psi''_{n} (x_{n-1}) \right)^{2} + 4 a_{n-1} (x_{n-1}/2) \left( \psi''_{n} (x_{n-1}/2) \right)^{2} + a_{n-1} (x_{n}) \left( \psi''_{n} (x_{n}) \right)^{2} \right] \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ \left( -\frac{4}{h} \right) \left( -\frac{2}{h} \right) + 4 \left( \frac{1}{2} \right) \left( -\frac{1}{h} \right) \left( \frac{1}{h} \right) \right] + \hat{V}_{n} \frac{h}{6} \left[ \left( -\frac{2}{h} \right) + 4 \left( \frac{1}{2} \right) \left( \frac{1}{h} \right)^{2} \right] \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ \frac{8}{h^{2}} - \frac{2}{h^{2}} \right] + \hat{V}_{n-1} \frac{h}{6} \left[ \frac{4}{h^{2}} + \frac{2}{h^{2}} \right] \]

\[ = \frac{1}{h} (\hat{V}_{n-1} + \hat{V}_{n}). \]

\[ C_{n,n} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{2} \psi''_{j} \psi''_{n} \, dx \right) \hat{V}_{j} \]

\[ = \hat{V}_{n-1} \int_{0}^{1} a_{n} \psi''_{n-1} \psi''_{n} \, dx + \hat{V}_{n} \int_{0}^{1} a_{n} \left( \psi''_{n} \right)^{2} \, dx \]

\[ = \hat{V}_{n-1} \int_{I_{n}} a_{n} \psi''_{n-1} \psi''_{n} \, dx + \hat{V}_{n} \int_{I_{n+1}} a_{n} \left( \psi''_{n} \right)^{2} \, dx + V_{n} \int_{I_{n+1}} a_{n} \left( \psi''_{n} \right)^{2} \, dx \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ a_{n} (x_{n-1}) \psi''_{n-1} (x_{n-1}) \psi''_{n} (x_{n-1}) + 4 a_{n} (x_{n-1}/2) \psi''_{n-1} (x_{n-1}/2) \psi''_{n} (x_{n-1}/2) \right. \]

\[ + a_{n} (x_{n}) \psi''_{n-1} (x_{n}) \psi''_{n} (x_{n}) \left. \right] + \hat{V}_{n-1} \frac{h}{6} \left[ a_{n} (x_{n-1}) \left( \psi''_{n} (x_{n-1}) \right)^{2} \right] \]

\[ + 4 a_{n} (x_{n-1}/2) \left( \psi''_{n} (x_{n-1}/2) \right)^{2} + a_{n} (x_{n}) \left( \psi''_{n} (x_{n}) \right)^{2} \]

\[ + \hat{V}_{n} \frac{h}{6} \left[ a_{n} (x_{n}) \left( \psi''_{n} (x_{n}) \right)^{2} \right] \]

\[ + 4 a_{n} (x_{n+1}/2) \left( \psi''_{n} (x_{n+1}/2) \right)^{2} + a_{n} (x_{n+1}) \left( \psi''_{n} (x_{n+1}) \right)^{2} \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ 4 \left( \frac{1}{2} \right) \left( -\frac{1}{h} \right) \left( \frac{1}{h} \right) + \left( \frac{2}{h} \right) \left( \frac{4}{h} \right) \right] + \hat{V}_{n} \frac{h}{6} \left[ 4 \left( \frac{1}{2} \right) \left( \frac{1}{h} \right)^{2} + \left( \frac{4}{h} \right)^{2} \right] \]

\[ + \hat{V}_{n} \frac{h}{6} \left[ \left( -\frac{4}{h} \right)^{2} + 4 \left( \frac{1}{2} \right) \left( -\frac{1}{h} \right)^{2} \right] \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ -\frac{2}{h^{2}} + \frac{8}{h^{2}} \right] + \hat{V}_{n} \frac{h}{6} \left[ \frac{2}{h^{2}} + \frac{16}{h^{2}} + \frac{16}{h^{2}} + \frac{2}{h^{2}} \right] \]

\[ = \hat{V}_{n-1} \frac{h}{6} \left[ \frac{6}{h^{2}} \right] + \hat{V}_{n} \frac{h}{6} \left[ \frac{36}{h^{2}} \right] \]

\[ = \frac{1}{h} (\hat{V}_{n-1} + 6\hat{V}_{n}). \]
5.1. Adjoint Stiffness Matrix Based Approach

\[ C_{n,n+1} = \sum_{j=1}^{n} \left( \int_{0}^{1} a_{n+1} \psi_j'' \psi_n'' \, dx \right) \hat{V}_j \]

\[ = \hat{V}_n \int_{0}^{1} a_{n+1} (\psi_n'')^2 \, dx \]

\[ = \hat{V}_n \int_{h+1}^{n} a_{n+1} (\psi_n'')^2 \, dx \]

\[ = \hat{V}_n \frac{h}{6} \left[ a_{n+1}(x_n) (\psi_n'(x_n))^2 + 4a_{n+1}(x_{n+1/2}) (\psi_n'(x_{n+1/2}))^2 + a_{n+1}(x_{n+1}) (\psi_n''(x_{n+1}))^2 \right] \]

\[ = \hat{V}_n \frac{h}{6} \left[ \frac{1}{2} \left( -\frac{1}{h} \right)^2 + \left( \frac{2}{h} \right)^2 \right] \]

\[ = \hat{V}_n \frac{h}{6} \left[ \frac{2}{h^2} + \frac{4}{h^2} \right] \]

\[ = \frac{1}{h} \hat{V}_n. \]

Collecting the entries of the matrix we get

\[
C = \frac{1}{h} \begin{pmatrix}
\hat{V}_1 & 6\hat{V}_1 + \hat{V}_2 & \hat{V}_1 + \hat{V}_2 & 0 & 0 & \ldots & 0 \\
0 & \hat{V}_1 + \hat{V}_2 & \hat{V}_1 + 6\hat{V}_2 + \hat{V}_3 & \hat{V}_2 + \hat{V}_3 & 0 & \ldots & 0 \\
0 & \ldots & \hat{V}_2 + \hat{V}_3 & \hat{V}_2 + 6\hat{V}_3 + \hat{V}_4 & \hat{V}_3 + \hat{V}_4 & \ldots & 0 \\
0 & \ldots & 0 & \hat{V}_{n-2} + \hat{V}_{n-1} & \hat{V}_{n-2} + 6\hat{V}_{n-1} + \hat{V}_n & \hat{V}_{n-1} + \hat{V}_n & 0 \\
0 & \ldots & 0 & \hat{V}_{n-1} + \hat{V}_n & \hat{V}_{n-1} + 6\hat{V}_n & \hat{V}_n \\
\end{pmatrix}.
\]
Chapter 6

Proximal Point Method

In this chapter we discuss various proximal point methods and employ them to solve the inverse problem of parameter identification in the fourth-order beam equation.

6.1 Optimization Formulation

We recall the optimization problem that we consider reads as:

\[
\min_{a \in A} J(a). \tag{6.1}
\]

We emphasize on two cases. The first case is of the output least squares (OLS) functional given by

\[
J(a) = J_1(a) + R(a), \tag{6.2}
\]

where

\[
J_1(a) = \frac{1}{2} \| u(a) - z \|^2.
\]

The second case is of the modified output least squares (OLS) functional given by

\[
J(a) = J_2(a) + R(a), \tag{6.3}
\]

where

\[
J_2(a) = \frac{1}{2} T(a, u - z, u - z).
\]

Recall the in the above \( u(a) \) is a solution of the weak form, \( z \) is the data, and \( R \) is the regularization functional.
We consider the following three choices for $R$:

\begin{align}
R(a) &= R_1(a) \ (L_2\text{-norm}) \quad (6.4a) \\
R(a) &= R_2(a) \ (H_1\text{-norm}) \quad (6.4b) \\
R(a) &= R_3(a) \ (H_1\text{semi-norm}). \quad (6.4c)
\end{align}

We remark that the functional $J_1$ is not convex and the convexity can only be achieved by choosing a sufficiently large regularization parameter. On the other hand, $J_2$ is always convex. Due to the convexity, our most arguments are valid for $J_2$ only. Furthermore, all the numerical experiments are done for $J_2$ only. We choose the constraint set $\tilde{A}$ to be closed and convex.

We now describe the proximal point method:

\begin{align}
\alpha^{k+1} &= \min_{\alpha \in \tilde{A}} \left\{ J(\alpha) + \frac{1}{2\lambda_k} \| \alpha - y^k \|_2^2 \right\} \\
y^{k+1} &= \alpha^{k+1}
\end{align}

or equivalently

\begin{equation}
\alpha^{k+1} = \min_{\alpha \in \tilde{A}} \left\{ J(\alpha) + \frac{1}{2\lambda_k} \| \alpha - \alpha^k \|_2^2 \right\} \quad (6.5)
\end{equation}

where $\lambda^k$ is a sequence of positive numbers with $\lim_{k \to \infty} \lambda_k > 0$. An important part of the proximal point algorithm is solving the subproblem (6.5). The term

\begin{equation}
\frac{1}{2\lambda_k} \| \alpha - \alpha^k \|_2^2
\end{equation}

is known as the regularization term and is strictly convex. This will guarantee that the subproblem has a unique minimizer for each $k$. We set

\begin{equation}
\alpha^{k+1} = \min_{\alpha \in \tilde{A}} \{ \mathcal{J}_P(\alpha) \} \quad (6.6)
\end{equation}

where

\begin{equation}
\mathcal{J}_P(\alpha) = J(\alpha) + \frac{1}{2\lambda_k} \| \alpha - \alpha^k \|_2^2 \quad (6.7)
\end{equation}

The necessary and sufficient optimality condition for the above optimization problem is then the following variational inequality of finding $\alpha^* \in \tilde{A}$:

\begin{equation}
\langle \nabla \mathcal{J}_P(\alpha^*), \alpha - \alpha^* \rangle \geq 0 \ \forall \alpha \in \tilde{A}. \quad (6.8)
\end{equation}

In this chapter, we will look at the proximal point method described by Hager and Zhang [18]. Then using Kanzow [22]. We will incorporate different strategies to solve the subproblem and
combine them to method proposed by Hager and Zhang [18]. We will look at two more proximal point methods suggested by Han and Li. In each method, we will describe the algorithm, and test the method on four examples which are given on the next section. Some details on the proximal point methods are given in [2], [5], [19], [20], [23], [24], [32], [33], [34], [36], [50] and the cited differences therein.

6.2 Test Examples

We will test every method on the following four examples for each method:

Example 1.

\[
\begin{align*}
  f_1 &= 8\pi^2(2\pi^2(x^2 + 1) - 1)\cos(2\pi x) + 4\pi x \sin(2\pi x) \\
  k_1 &= x^2 + 1 \\
  z_1 &= \cos(2\pi x) - 1.
\end{align*}
\]

Example 2.

\[
\begin{align*}
  f_2 &= 4(90x^4 - 60x^3 + 42x^2 - 18x + 13) \\
  k_2 &= x^4 + x^2 + 2 \\
  z_2 &= x^2(x - 1)^2.
\end{align*}
\]

Example 3.

\[
\begin{align*}
  f_3 &= 32\pi^2(-631.65(x+0.4)(x^2-1.15x+0.6)\cos(4\pi x) - 301.6(x-0.25)^2 \sin(4\pi x)) \\
  k_3 &= (2x - .5)^3 + 2 \\
  z_3 &= \sin(4\pi x - 0.5\pi) + 1.
\end{align*}
\]

Example 4.

\[
\begin{align*}
  f_4 &= 1764(x - 0.773513)(x - 0.5)(x - 0.226487)(x^2 - x + 0.687621) \\
  k_4 &= (x - 0.5)^2 + 1 \\
  z_4 &= x^2(x - 1)^2(x - 0.5)^3.
\end{align*}
\]
6.3 Proximal-Like Methods Using Least Squares

The first method we use to solve the inverse problem was using the strategy suggested by Hager and Zhang. This involves the gradient of the proximal function, the gradient of the original function, and the difference between iterates. Proximal point method generates the iterative scheme by solving a minimization problem

$$a^{k+1} = \min_{a \in A} J_P(a)$$  \hspace{1cm} (6.9)

where

$$J_P(a) = J(a) + \frac{1}{2} \mu_k \|a - a^k\|^2.$$  \hspace{1cm} (6.10)

The subproblem is solved according to the following two criteria:

$$J_P(a^{k+1}) \leq J(a^k) \hspace{1cm} (6.11a)$$

$$\|\nabla J_P(a^{k+1})\| \leq \mu_k \|\nabla J(a^k)\|. \hspace{1cm} (6.11b)$$

This results in a fast convergence without using the Hessian of $J(a^k)$. Notice that in this scheme the least-square regularization term is used (cf. (6.10)). This strictly convex regularization term guarantees that the subproblem has a unique minimizer for each $a$. Hence this proximal point method is well-defined. The regularization parameter has the form

$$\mu^k = \beta \|\nabla J(a^k)\|^\eta,$$

where $\mu \in [0, 2)$ and $\beta > 0$ is a constant to obtain convergence.

**Algorithm 5.1**

*Initialization Step:* Choose an initial guess $a^0$

Let $\mu_k = \beta \|\nabla J(a^k)\|^\eta$ and let $\gamma = 1$

Initialize $\beta$ and $\eta$

**Step 1:** Find an $a^{k+1}$ satisfying

$$\|\nabla J_P(a^{k+1})\| \leq \mu_k \gamma \|\nabla J(a^k)\| \hspace{1cm} (6.12)$$

**Step 2:** If $a^{k+1}$ satisfies

$$J_P(a^{k+1}) \leq J(a^k) \hspace{1cm} (6.13)$$
6.3. Proximal-Like Methods Using Least Squares

Go to Step 3.

Else,

Set $\gamma = \gamma \times 0.1$ and go to Step 1,

End

**Step 3:** Let

$$a^k = a^{k+1}.$$  

**Step 4:** Set $k = k + 1$ and go to Step 1.

More specifically, in Step 1, the subproblem of (6.12) used a conjugate gradient method to find $a^{k+1}$.

In our experiments, we set

$$\mu_k = \beta \| \nabla J(a^k) \|^\eta = 0.5 \| \nabla J(a^k) \|^1.$$
6.3. Proximal-Like Methods Using Least Squares

Example 1: Hager-Zhang Method using Least Squares

Figure 6.1: Example 1: Coefficient by Hager-Zhang Method using Least Squares

Figure 6.2: Example 1: Solution by Hager-Zhang Method using Least Squares
Example 2: Hager-Zhang Method using Least Squares

Figure 6.3: Example 2: Coefficient by Hager-Zhang Method using Least Squares

Figure 6.4: Example 2: Solution by Hager-Zhang Method using Least Squares
6.3. Proximal-Like Methods Using Least Squares

Example 3: Hager-Zhang Method using Least Squares

Figure 6.5: Example 3: Coefficient by Hager-Zhang Method using Least Squares

Figure 6.6: Example 3: Solution by Hager-Zhang Method using Least Squares
Example 4: Hager-Zhang Method using Least Squares

Figure 6.7: Example 4: Coefficient by Hager-Zhang Method using Least Squares

Figure 6.8: Example 4: Solution by Hager-Zhang Method using Least Squares
6.4 Proximal-like Methods Using $\varphi$-Divergence

This method is a variant of the classical proximal point method. Instead of using the strict convex quadratic term in the subproblem (6.7), this method uses $\varphi$-divergence functions to be the regularization term. In order to be a $\varphi$ functions it must hold the following properties,

1. $\varphi$ is twice continuously differentiable on $\text{int}(\Omega) = (0, +\infty)$.
2. $\varphi$ is strictly convex on its domain.
3. $\lim_{t \to 0^+} \frac{d\varphi(x)}{dx} = -\infty$.
4. $\varphi(1) = \frac{d\varphi(1)}{dx} = 0$ and $\frac{d^2\varphi(1)}{dx^2} > 0$.
5. There exists $v \in \left(\frac{1}{2} \frac{d^2\varphi(1)}{dx^2}, \frac{d^2\varphi(1)}{dx^2}\right)$ such that

$$
\left(1 - \frac{1}{t}\right) \left(\frac{d^2\varphi(1)}{dx^2} + v(t - 1)\right) \leq \frac{d\varphi(t)}{dx} \leq \frac{d^2\varphi(1)}{dx^2}(t - 1) \quad \forall t > 0.
$$

Instead of (6.6), the subproblem will be

$$
a^{k+1} = \min \left\{ J_\varphi(a) \mid a \in \Omega \right\} \quad (6.14)
$$

where

$$
J_\varphi(x) = J(a) + \mu_k d_\varphi(a, a^k) \quad (6.15)
$$

and

$$
d_\varphi(x, y) = \sum_{i=1}^{n} y_i \varphi \left( \frac{x_i}{y_i} \right). \quad (6.16)
$$

A few examples of $\varphi$ functions are,

$$
\varphi_1(t) = t \log t - t + 1 \\
\varphi_2(t) = -\log t + t - 1 \\
\varphi_3(t) = (\sqrt{t} - 1)^2.
$$

In our examples, we chose to use $\varphi_3(t)$ function and therefore the definition of $d_\varphi$ is

$$
d_\varphi(x, y) = \sum_{i=1}^{n} y_i \varphi \left( \sqrt{\frac{x_i}{y_i}} - 1 \right)^2. \\
d_\varphi(a, a^k) = \sum_{i=1}^{n} a^k_i \varphi \left( \sqrt{\frac{a_i}{a^k_i}} - 1 \right)^2.
$$
6.4. Proximal-like Methods Using $\varphi$-Divergence

and solving the subproblem

$$a^{k+1} = \min \left\{ J_2(a) + \mu_k \sum_{i=1}^{n} \varphi \left( \sqrt{\frac{a_i}{a_i^k}} - 1 \right)^2 \mid a \in \Omega \right\}. \quad (6.17)$$

**Algorithm 5.2**

**Initialization Step:** Choose an initial $x^0$.

Let $\mu_k = \beta \| \nabla J(a^k) \|^\eta$ and let $\gamma = 1$.

Initialize $\beta$ and $\eta$

**Step 1:** Find an $a^{k+1}$ satisfying

$$\| \nabla J_\varphi(a^{k+1}) \| \leq \mu_k \gamma \| \nabla J(a^k) \| \quad (6.18)$$

**Step 2:** If $a^{k+1}$ satisfies

$$J_\varphi(a^{k+1}) \leq J(a^k) \quad (6.19)$$

Go to Step 3.

Else,

Set $\gamma = \gamma \times 0.1$ and go to step 1.

End

**Step 3:** Let

$$a^k = a^{k+1}.$$

**Step 4:** Set $k = k + 1$ and go to step 1.

More specifically, in Step 1, the subproblem of (6.18) used a conjugate gradient method to find $a^{k+1}$.

In our experiments, we set $\mu_k = \beta \| \nabla J(a^k) \|^\eta = 0.5 \| \nabla J(a^k) \|^1$. 
Example 1: Hager-Zhang Method using $\varphi$-divergence

Figure 6.9: Example 1: Coefficient by Hager-Zhang Method using $\varphi$-divergence

Figure 6.10: Example 1: Solution by Hager-Zhang Method using $\varphi$-divergence
Example 2: Hager-Zhang Method using $\varphi$-divergence

Figure 6.11: Example 2: Coefficient by Hager-Zhang Method using $\varphi$-divergence

Figure 6.12: Example 2: Solution by Hager-Zhang Method using $\varphi$-divergence
Example 3: Hager-Zhang Method using $\phi$-divergence

Figure 6.13: Example 3: Coefficient by Hager-Zhang Method using $\phi$-divergence

Figure 6.14: Example 3: Solution by Hager-Zhang Method using $\phi$-divergence
6.4. Proximal-like Methods Using $\phi$-Divergence

Example 4: Hager-Zhang Method using $\phi$-divergence

Figure 6.15: Example 4: Coefficient by Hager-Zhang Method using $\phi$-divergence

Figure 6.16: Example 4: Solution by Hager-Zhang Method using $\phi$-divergence
6.5 Proximal-like Methods Using Bregman Functions

This method uses Bregman functions which are more general convex functions. It involves replacing the least squares or $\varphi$-divergence term by another strictly convex function

$$D_\psi(x, y) = \psi(x) - \psi(y) - \nabla \psi(y)^T (x - y),$$  \hspace{1cm} (6.20)

where $\psi$ is the so-called Bregman function.

Let $S$ be an open and convex set and a mapping $\psi : \bar{S} \to \mathbb{R}$. In order to be a Bregman function, it must hold the following properties:

1. $\psi$ is strictly convex and continuous on $\bar{S}$.

2. $\psi$ is continuously differentiable in $S$.

3. The partial level set
   $$L_\alpha = \{y \in \bar{S} | D_\psi(x, y) \leq \alpha\}$$
   is bounded for every $x \in \bar{S}$.

4. If $\{y^k\} \subset S$ converges to $x$, then $\lim_{k \to \infty} D_\psi(x, y^k) = 0$.

Instead of (6.6), the subproblem will be

$$a^{k+1} = \min \{ J_\psi(a) \mid a \in \Omega \},$$  \hspace{1cm} (6.21)

where

$$J_\psi(x) = J(a) + \mu_k D_\psi(a, a^k)$$  \hspace{1cm} (6.22)

and

$$D_\psi(x, y) = \psi(x) - \psi(y) - \nabla \psi(y)^T (x - y).$$  \hspace{1cm} (6.23)

A few examples of Bregman functions are,

$$\psi_1(x) = \frac{1}{2} \|x\|^2$$

$$\psi_2(x) = \sum_{i=1}^{n} x_i \log x_i - x_i$$

$$\psi_3(x) = - \sum_{i=1}^{n} \log x_i,$$
and there respective convex-like functions

\[
D_1(x, y) = \frac{1}{2} \|x - y\|^2 \\
D_2(x, y) = \sum_{i=1}^{n} x_i \log \frac{x_i}{y_i} + y_i - x_i \\
D_3(x, y) = -\sum_{i=1}^{n} \frac{x_i}{y_i} \log \frac{x_i}{y_i} - 1.
\]

In our problem, we know the coefficient must be within a lower and upper limit otherwise known as a box constraints, \(0 < l_i < a_i < u_i < \infty\). A Bregman function that satisfies the box constraints and the one that we use our example is,

\[
\psi_4(a^k) = \sum_{i=1}^{n} (a_i - l_i) \log (a_i - l_i) + (u_i - a_i) \log (u_i - a_i^k)
\]

(6.24)

\[
D_4(a^k, y) = \sum_{i=1}^{n} (a_i - l_i) \log \left( \frac{a_i - l_i}{a_i^k - l_i} \right) + (u_i - a_i) \log \left( \frac{u_i - a_i}{u_i - a_i^k} \right).
\]

(6.25)

We are solving the subproblem

\[
da^{k+1} = \min \left\{ J_2(a) + \mu_k \sum_{i=1}^{n} (a_i - l_i) \log \left( \frac{a_i - l_i}{a_i^k - l_i} \right) + (u_i - a_i) \log \left( \frac{u_i - a_i}{u_i - a_i^k} \right) \mid a \in \Omega \right\}.
\]

Algorithm 5.3

**Initialization Step:** Choose an initial \(a^0\). Let \(\mu_k = \beta \|\nabla J(a^k)\|\eta\) and let \(\gamma = 1\). Initialize \(\beta\) and \(\eta\). **Step 1.** Find an \(a^{k+1}\) satisfying

\[
\|\nabla \mathcal{J}_\psi(a^{k+1})\| \leq \mu_k \gamma \|\nabla J_2(a^k)\| \quad (6.26)
\]

**Step 2.** If \(a^{k+1}\) satisfies

\[
\mathcal{J}_\psi(a^{k+1}) \leq J(a^k) \quad (6.27)
\]

Go to Step 3.

Else,

Set \(\gamma = \gamma \times 0.1\) and go to Step 1.

End

**Step 3.** Let

\[
a^k = a^{k+1}.
\]
Step 4. Set $k = k + 1$ and go to Step 1.

More specifically, in Step 1, the subproblem of (6.26) used a conjugate gradient method to find $a^{k+1}$.

In our experiments, we set $\mu_k = \beta \|\nabla J(a^k)\|^\eta = 0.5 \|\nabla J(a^k)\|^1$. 
Example 1: Hager-Zhang Method using Bregman Function

Figure 6.17: Example 1: Coefficient by Hager-Zhang Method using Bregman function

Figure 6.18: Example 1: Solution by Hager-Zhang Method using Bregman function
Example 2: Hager-Zhang Method using Bregman Function

![Graphs showing exact and estimated coefficients, errors in coefficients, data, final simulation, and errors in simulated data.]

Figure 6.19: Example 2: Coefficient by Hager-Zhang Method using Bregman function

Figure 6.20: Example 2: Solution by Hager-Zhang Method using Bregman function
6.5. Proximal-like Methods Using Bregman Functions

Example 3: Hager-Zhang Method using Bregman Function

Figure 6.21: Example 3: Coefficient by Hager-Zhang Method using Bregman function

Figure 6.22: Example 3: Solution by Hager-Zhang Method using Bregman function
Example 4: Hager-Zhang Method using Bregman Function

Figure 6.23: Example 4: Coefficient by Hager-Zhang Method using Bregman function

Figure 6.24: Example 4: Solution by Hager-Zhang Method using Bregman function
6.6 Proximal-like Methods Using Modified $\phi$-Divergence

In the $\phi$-divergence method, we used a conjugate gradient method to solve the subproblem. Now we would like to solve the subproblem by using Newton’s method. This method involves knowing the second order derivative and the Hessian matrix of the objective functional. With $\phi$-divergence, we have (6.32)

$$d_\phi(x,y) = \sum_{i=1}^{n} y_i \varphi \left( \frac{x_i}{y_i} \right).$$

The second derivative of this gives us

$$\nabla_{xx}^2 d_\phi(x,y) = \sum_{i=1}^{n} \frac{1}{y_i} \frac{d^2 \varphi}{dx^2} \left( \frac{x_i}{y_i} \right) e_i e_i^T,$$

where $e_i$ is the $i$th unit vector.

This is a problem since the factor $\frac{1}{y_i}$ goes to infinity during the iteration process for all indices for which a constraint like $x_i \geq 0$ is a solution. This will lead to an ill-condition Hessian matrix. To avoid this, the modification comes by replacing the $y_i$ term in the $\varphi$ function to $y_i^2$. Then the factor of $\frac{1}{y_i}$ will not exist in the second derivative. Now the modified $\phi$-divergence function can be represented by

$$d_\tilde{\phi}(x,y) = \sum_{i=1}^{n} y_i^2 \varphi \left( \frac{x_i}{y_i} \right), \quad (6.28)$$

and the second derivative by

$$\nabla_{xx}^2 d_\tilde{\phi}(x,y) = \sum_{i=1}^{n} \frac{1}{y_i} \frac{d^2 \varphi}{dx^2} \left( \frac{x_i}{y_i} \right) e_i e_i^T. \quad (6.29)$$

Instead of (6.6), the subproblem will be

$$a^{k+1} = \min \left\{ \tilde{J}_\varphi(a) \mid a \in \Omega \right\} \quad (6.30)$$

where

$$\tilde{J}_\varphi(x) = J(a) + \frac{1}{\lambda_k} d_\varphi(a,a^k) \quad (6.31)$$

and

$$d_\varphi(x,y) = \sum_{i=1}^{n} y_i \varphi \left( \frac{x_i}{y_i} \right). \quad (6.32)$$
In our examples, we chose to use $\varphi_3(t) = (\sqrt{t} - 1)^2$ function and therefore the definition of $\bar{d}_\varphi$ is

$$
\bar{d}_\varphi(x, y) = \sum_{i=1}^{n} y_i^2 \varphi\left(\sqrt{\frac{x_i}{y_i}} - 1\right)^2.
$$

and solving the subproblem

$$
a^{k+1} = \min \left\{ J_2(a) + \frac{1}{\lambda_k} \sum_{i=1}^{n} (a_i^k)^2 \varphi\left(\sqrt{\frac{a_i}{a_i^k}} - 1\right)^2 \mid a \in \Omega \right\}
$$

**Algorithm 5.4**

**Initialization:** Choose an initial $a^0$.

Let $\mu_k = \beta \|\nabla J(a^k)\|^\eta$ and let $\gamma = 1$.

Initialize $\beta$ and $\eta$

**Step 1.** Find an $a^{k+1}$ satisfying

$$
\|\nabla \tilde{J}_\varphi(a^{k+1})\| \leq \mu_k \gamma \|\nabla J(a^k)\|. \quad (6.33)
$$

**Step 2.** If $a^{k+1}$ satisfies

$$
\tilde{J}_\varphi(a^{k+1}) \leq J(a^k) \quad (6.34)
$$

Go to Step 3.

Else,

Set $\gamma = \gamma \times 0.1$ and go to step 1,

End

**Step 3.** Let

$$
a^k = a^{k+1}.
$$

**Step 4.** Set $k = k + 1$ and go to step 1.

More specifically, in Step 1, the subproblem of (6.33) used a conjugate gradient method to find $a^{k+1}$. In our experiments, we set $\mu_k = \beta \|\nabla J(a^k)\|^\eta = 0.5 \|\nabla J(a^k)\|^1$.
6.6. Proximal-like Methods Using Modified $\varphi$-Divergence

Example 1: Hager-Zhang Method using Quadratic function

Figure 6.25: Example 1: Coefficient by Hager-Zhang Method using quadratic function

Figure 6.26: Example 1: Solution by Hager-Zhang Method using quadratic function
Example 2: Hager-Zhang Method using Quadratic function

Figure 6.27: Example 2: Coefficient by Hager-Zhang Method using quadratic function

Figure 6.28: Example 2: Solution by Hager-Zhang Method using quadratic function
Example 3: Hager-Zhang Method using Quadratic function

Figure 6.29: Example 3: Coefficient by Hager-Zhang Method using quadratic function

Figure 6.30: Example 3: Solution by Hager-Zhang Method using quadratic function
Example 4: Hager-Zhang Method using Quadratic function

Figure 6.31: Example 4: Coefficient by Hager-Zhang Method using quadratic function

Figure 6.32: Example 4: Solution by Hager-Zhang Method using quadratic function
6.7 Han Proximal-Point Method

This method features a Bregman function-based proximal point algorithm for solving the vari-ational inequality problem. The first method we adapted for the inverse problem was using the stopping criteria from Hager-Zhang. Recall that the two criteria are

\[
\mathcal{J}_P(a^{k+1}) \leq J(a^k) \quad (6.35a)
\]
\[
\|\nabla \mathcal{J}_P(a^{k+1})\| \leq \mu_k \|\nabla J(a^k)\|. \quad (6.35b)
\]

Algorithm 5.5

Initialization Step: Choose an appropriate Bregman function.
Choose an initial \(a^0\)
Initialize \(\sigma \in (0, 1)\) and \(0 < c_k < \infty\)
Step 1. Find an \(y^k\) satisfying

\[
\|r^k\| \leq \sigma \|a^k - y^k\| \quad (6.36)
\]

where

\[
\|r^k\| = c_k (\nabla J_P(y^k) + \nabla h(y^k) - \nabla h(a^k)) + (y^k - a^k)
\]

Step 2. Compute \(a^{k+1}\),

\[
a^{k+1} = (1 - t) \text{Proj}[a^k - c_k \alpha_k \nabla J_P(y^k)] + t a^k \quad (6.37)
\]

where

\[
\alpha_k = \frac{\langle \nabla J_P(y^k), a^k - y^k \rangle}{c_k \|\nabla J(y^k)\|^2}
\]

Step 3. Set \(k = k + 1\) and go to Step 1.

More specifically, in Step 1, the subproblem of (6.36) used a Newton method to find \(a^{k+1}\).
In our experiments, we set \(c_k = \frac{1}{\|\nabla(a^k)\|}\).
Example 1: Han Method

Figure 6.33: Example 1: Coefficient by Han Method

Figure 6.34: Example 1: Solution by Han Method
Example 2: Han Method

Figure 6.35: Example 2: Coefficient by Han Method

Figure 6.36: Example 2: Solution by Han Method
6.7. Han Proximal-Point Method

Example 3: Han Method

Figure 6.37: Example 3: Coefficient by Han Method

Figure 6.38: Example 3: Solution by Han Method
Example 4: Han Method

Figure 6.39: Example 4: Coefficient by Han Method

Figure 6.40: Example 4: Solution by Han Method
The next method we employ is by Li and Yuan. Their aim was to make better adjustments on the subproblem and to optimize the step-size in the gradient step.

**Algorithm 5.6**

**Initialization Step.** Choose an initial $a^0$

Initialize $\sigma \in (0, 1)$, $v \in (0, 1)$, and $1 < \gamma_k < 2$. **Step 1.** Find an $y^k$ satisfying

$$
\Delta(y^k) \leq v(||a^k - y^k||^2 + ||a^k - y^k||)
$$

where

$$
\Delta(y^k) = \langle 2(y^k - \tilde{y}^k), \nabla J(y^k) \rangle - ||y^k - \tilde{y}^k||^2
$$

$$
\tilde{y}^k = \text{Proj}[a^k - \lambda_k \nabla J(y^k)]
$$

**Step 2.** Correct the approximation $y^k$,

$$
\tilde{a}^k = \text{Proj}[a^k - \alpha_k \lambda_k \nabla J_P(y^k)]
$$

where

$$
\alpha_k = \frac{||a^k - \tilde{a}^k||^2 - \langle \lambda_k(y^k - \tilde{y}^k), \nabla J_P(y^k) \rangle}{||a^k - \tilde{y}^k||^2}
$$

**Step 3.** Compute $a^{k+1}$,

$$
\tilde{a}^{k+1} = \text{Proj}[a^k - \gamma_k \tau_k(a^k - \tilde{a}^k)]
$$

where

$$
\tau_k = 1 + \frac{\langle \alpha_k \lambda_k(\tilde{a}^k - y^k), \nabla J_P(y^k) \rangle}{||a^k - \tilde{a}^k||^2}
$$

**Step 4.** Set $k = k + 1$ and go to Step 1.

More specifically, in Step 1, the subproblem of (6.38) used a Newton method to find $a^{k+1}$. 
Example 1: Li-Yuan Method

Figure 6.41: Example 1: Coefficient by Li-Yuan Method

Figure 6.42: Example 1: Solution by Li-Yuan Method
Example 2: Li Method

Figure 6.43: Example 2: Coefficient by Li-Yuan Method

Figure 6.44: Example 2: Solution by Li-Yuan Method
Example 3: Li-Yuan Method

Figure 6.45: Example 3: Coefficient by Li-Yuan Method

Figure 6.46: Example 3: Solution by Li-Yuan Method
Example 4: Li-Yuan Method

Figure 6.47: Example 4: Coefficient by Li Method

Figure 6.48: Example 4: Solution by Li-Yuan Method
Chapter 7

Comparative Performance Analysis

In this chapter we compare the various proximal point algorithms that we have employed in this work.

7.1 Numerical Results for Example 1

We would like to compare the results of Hager and Zhang’s proximal point algorithm. We make this discussion in terms of Example 1. Table 7.1 contains the results for this example. Each of Hager and Zhang’s first derivative each have roughly the same numbers, while Hager and Zhang’s quadratic has the best performance as far as time is concerned. Recall that Hager and Zhang’s quadratic Method uses the Newton method to solve the subproblem, unlike the conjugate gradient method. In each of Hager and Zhang’s method, we set $\beta = 0.05$, and $\mu = 1$. Then testing the method on many epsilon’s from $1e-1, 5e-2, 1e-2, 5e-3, 1e-3, ..., 1e-10$. The best results occurred when epsilon=$1e-6$, and $N = 50$.

In the following, let LS stand for the method involving Least Squares, let $\varphi$ stand for the method involving $\varphi$-divergence, and the other two of Hanger and Zhang’s method be the Bregman function, and the quadratic method.

We want to look at two parameters, the number of functions evaluation, and the time to complete the computation. In general, Hager and Zhang’s proximal point method gives very good results, being computed in a very reasonable time. The proximal point method of Han’s, and Li and Yuan’s lack in the number of functions and on time. However the stopping criteria of the subproblem were different for Han, and Li and Yaun. We are more interested in the algorithms proposed by Hager and Zhang. In each of Hager and Zhang’s method the initial value of $\beta$ was
set to 0.05 and Hager-Zhang note that $\beta$ should be a sufficiently small. In his numerical results, he uses $\beta = 0.05$, which is what lead me to use it for my own numerical results. Since Hager and Zhang don’t specify exactly what $\beta$ is, we wanted to explore this a bit more. We then ran each of Hager and Zhang’s method for example one with the $\beta$ value ranging from 0.01 to 0.1 by steps of 0.01 and the results are shown in tables 7.2 to table 7.5.

It is important to note that Hager and Zhang-quadratic method can only be used when the Hessian form is available for use. If it is not available the best first derivative method would be Hager and Zhang- $\varphi$, but either of the other two will work just as good.

We rearrange table 7.1 by using the best $\beta$ values from each method in table ?? Now we are able to see that the already reasonably good results are now even better. Notice the time’s are all roughly the same with Hager-Zhang quadratic method still being a bit better in time, but also now the number of function’s evaluated are much better. All the methods showed much better results when $\beta$ values were lower than 0.05 (the original initial value).

Other examples experience the same performance of proximal point methods and hence we only give the tables containing the numerical results.

<table>
<thead>
<tr>
<th>Method</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-LS</td>
<td>865</td>
<td>783</td>
<td>NA</td>
<td>5</td>
<td>2.2138</td>
</tr>
<tr>
<td>Hager and Zhang-$\varphi$</td>
<td>869</td>
<td>792</td>
<td>NA</td>
<td>5</td>
<td>2.2394</td>
</tr>
<tr>
<td>Hager and Zhang-Bregman</td>
<td>856</td>
<td>775</td>
<td>NA</td>
<td>5</td>
<td>2.2044</td>
</tr>
<tr>
<td>Hager and Zhang-Quadratic</td>
<td>3609</td>
<td>3496</td>
<td>85</td>
<td>5</td>
<td>0.7286</td>
</tr>
<tr>
<td>Han</td>
<td>181844</td>
<td>179171</td>
<td>1739</td>
<td>127</td>
<td>29.9905</td>
</tr>
<tr>
<td>Li and Yuan</td>
<td>5122</td>
<td>3852</td>
<td>NA</td>
<td>1270</td>
<td>29.9202</td>
</tr>
</tbody>
</table>

Table 7.1: Results of Example 1.
### 7.1. Numerical Results for Example 1

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>521</td>
<td>496</td>
<td>NA</td>
<td>3</td>
<td>1.3721</td>
</tr>
<tr>
<td>0.02</td>
<td>909</td>
<td>831</td>
<td>NA</td>
<td>4</td>
<td>2.3844</td>
</tr>
<tr>
<td>0.03</td>
<td>864</td>
<td>797</td>
<td>NA</td>
<td>5</td>
<td>2.2442</td>
</tr>
<tr>
<td>0.04</td>
<td>888</td>
<td>809</td>
<td>NA</td>
<td>5</td>
<td>2.2990</td>
</tr>
<tr>
<td>0.05</td>
<td>865</td>
<td>783</td>
<td>NA</td>
<td>5</td>
<td>2.2601</td>
</tr>
<tr>
<td>0.06</td>
<td>851</td>
<td>782</td>
<td>NA</td>
<td>5</td>
<td>2.1912</td>
</tr>
<tr>
<td>0.07</td>
<td>897</td>
<td>814</td>
<td>NA</td>
<td>5</td>
<td>2.3321</td>
</tr>
<tr>
<td>0.08</td>
<td>620</td>
<td>590</td>
<td>NA</td>
<td>4</td>
<td>1.9805</td>
</tr>
<tr>
<td>0.09</td>
<td>926</td>
<td>850</td>
<td>NA</td>
<td>5</td>
<td>2.3916</td>
</tr>
<tr>
<td>0.10</td>
<td>839</td>
<td>774</td>
<td>NA</td>
<td>4</td>
<td>2.4310</td>
</tr>
</tbody>
</table>

**Table 7.2: Example 1. Using Hager and Zhang-LS**

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>591</td>
<td>563</td>
<td>NA</td>
<td>3</td>
<td>1.5696</td>
</tr>
<tr>
<td>0.02</td>
<td>525</td>
<td>500</td>
<td>NA</td>
<td>3</td>
<td>1.4301</td>
</tr>
<tr>
<td>0.03</td>
<td>504</td>
<td>480</td>
<td>NA</td>
<td>4</td>
<td>1.3420</td>
</tr>
<tr>
<td>0.04</td>
<td>812</td>
<td>744</td>
<td>NA</td>
<td>5</td>
<td>2.1116</td>
</tr>
<tr>
<td>0.05</td>
<td>869</td>
<td>792</td>
<td>NA</td>
<td>5</td>
<td>2.2390</td>
</tr>
<tr>
<td>0.06</td>
<td>826</td>
<td>747</td>
<td>NA</td>
<td>5</td>
<td>2.1499</td>
</tr>
<tr>
<td>0.07</td>
<td>843</td>
<td>763</td>
<td>NA</td>
<td>5</td>
<td>2.18530</td>
</tr>
<tr>
<td>0.08</td>
<td>844</td>
<td>777</td>
<td>NA</td>
<td>4</td>
<td>2.2062</td>
</tr>
<tr>
<td>0.09</td>
<td>811</td>
<td>746</td>
<td>NA</td>
<td>4</td>
<td>2.1657</td>
</tr>
<tr>
<td>0.10</td>
<td>784</td>
<td>714</td>
<td>NA</td>
<td>4</td>
<td>2.0441</td>
</tr>
</tbody>
</table>

**Table 7.3: Example 1. Using Hager and Zhang-$\phi$**
### 7.1. Numerical Results for Example 1

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>591</td>
<td>563</td>
<td>NA</td>
<td>3</td>
<td>1.5738</td>
</tr>
<tr>
<td>0.02</td>
<td>939</td>
<td>852</td>
<td>NA</td>
<td>4</td>
<td>2.4800</td>
</tr>
<tr>
<td>0.03</td>
<td>921</td>
<td>847</td>
<td>NA</td>
<td>5</td>
<td>2.4224</td>
</tr>
<tr>
<td>0.04</td>
<td>894</td>
<td>817</td>
<td>NA</td>
<td>5</td>
<td>2.3555</td>
</tr>
<tr>
<td>0.05</td>
<td>856</td>
<td>775</td>
<td>NA</td>
<td>5</td>
<td>2.2461</td>
</tr>
<tr>
<td>0.06</td>
<td>810</td>
<td>745</td>
<td>NA</td>
<td>5</td>
<td>2.1094</td>
</tr>
<tr>
<td>0.07</td>
<td>879</td>
<td>799</td>
<td>NA</td>
<td>5</td>
<td>2.3223</td>
</tr>
<tr>
<td>0.08</td>
<td>778</td>
<td>741</td>
<td>NA</td>
<td>4</td>
<td>2.0573</td>
</tr>
<tr>
<td>0.09</td>
<td>666</td>
<td>634</td>
<td>NA</td>
<td>4</td>
<td>1.7689</td>
</tr>
<tr>
<td>0.10</td>
<td>688</td>
<td>655</td>
<td>NA</td>
<td>4</td>
<td>1.8227</td>
</tr>
</tbody>
</table>

Table 7.4: Example 1. Using Hager and Zhang-Bregman Function

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>4040</td>
<td>3934</td>
<td>77</td>
<td>3</td>
<td>0.7290</td>
</tr>
<tr>
<td>0.02</td>
<td>479</td>
<td>466</td>
<td>9</td>
<td>3</td>
<td>0.1476</td>
</tr>
<tr>
<td>0.03</td>
<td>430</td>
<td>417</td>
<td>8</td>
<td>4</td>
<td>0.1595</td>
</tr>
<tr>
<td>0.04</td>
<td>431</td>
<td>417</td>
<td>8</td>
<td>4</td>
<td>0.1923</td>
</tr>
<tr>
<td>0.05</td>
<td>4307</td>
<td>4193</td>
<td>82</td>
<td>5</td>
<td>0.7452</td>
</tr>
<tr>
<td>0.06</td>
<td>4256</td>
<td>4142</td>
<td>81</td>
<td>5</td>
<td>0.7388</td>
</tr>
<tr>
<td>0.07</td>
<td>4308</td>
<td>4193</td>
<td>82</td>
<td>5</td>
<td>0.9739</td>
</tr>
<tr>
<td>0.08</td>
<td>538</td>
<td>520</td>
<td>10</td>
<td>4</td>
<td>0.2139</td>
</tr>
<tr>
<td>0.09</td>
<td>538</td>
<td>520</td>
<td>10</td>
<td>4</td>
<td>0.1845</td>
</tr>
<tr>
<td>0.10</td>
<td>538</td>
<td>520</td>
<td>10</td>
<td>4</td>
<td>0.1891</td>
</tr>
</tbody>
</table>

Table 7.5: Example 1 Using Hager and Zhang-Quadratic Function
### 7.2. Numerical Results for Example 2

<table>
<thead>
<tr>
<th>Method</th>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-LS</td>
<td>0.01</td>
<td>521</td>
<td>496</td>
<td>NA</td>
<td>3</td>
<td>1.3721</td>
</tr>
<tr>
<td>Hager and Zhang-$\phi$</td>
<td>0.03</td>
<td>504</td>
<td>480</td>
<td>NA</td>
<td>4</td>
<td>1.3420</td>
</tr>
<tr>
<td>Hager and Zhang-Breg</td>
<td>0.01</td>
<td>591</td>
<td>563</td>
<td>NA</td>
<td>3</td>
<td>1.5739</td>
</tr>
<tr>
<td>Hager and Zhang-Quad</td>
<td>0.03</td>
<td>430</td>
<td>417</td>
<td>8</td>
<td>4</td>
<td>0.1595</td>
</tr>
</tbody>
</table>

Table 7.6: New Results of Example 1.

### 7.2 Numerical Results for Example 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-NR</td>
<td>157</td>
<td>146</td>
<td>NA</td>
<td>2</td>
<td>0.4800</td>
</tr>
<tr>
<td>Hager and Zhang-LS</td>
<td>323</td>
<td>205</td>
<td>NA</td>
<td>4</td>
<td>0.9358</td>
</tr>
<tr>
<td>Hager and Zhang-$\phi$</td>
<td>276</td>
<td>154</td>
<td>NA</td>
<td>3</td>
<td>0.8297</td>
</tr>
<tr>
<td>Hager and Zhang-Bregman</td>
<td>284</td>
<td>153</td>
<td>NA</td>
<td>3</td>
<td>0.8421</td>
</tr>
<tr>
<td>Hager and Zhang-Quadratic</td>
<td>4282</td>
<td>4140</td>
<td>81</td>
<td>4</td>
<td>0.8565</td>
</tr>
<tr>
<td>Han</td>
<td>66839</td>
<td>64859</td>
<td>1271</td>
<td>20</td>
<td>17.5445</td>
</tr>
<tr>
<td>Li and Yuan</td>
<td>2028</td>
<td>1597</td>
<td>NA</td>
<td>432</td>
<td>11.4972</td>
</tr>
</tbody>
</table>

Table 7.7: Results of Example 2.
### 7.2. Numerical Results for Example 2

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>67</td>
<td>64</td>
<td>NA</td>
<td>1</td>
<td>0.2237</td>
</tr>
<tr>
<td>0.02</td>
<td>67</td>
<td>62</td>
<td>NA</td>
<td>1</td>
<td>0.2249</td>
</tr>
<tr>
<td>0.03</td>
<td>67</td>
<td>62</td>
<td>NA</td>
<td>1</td>
<td>0.2260</td>
</tr>
<tr>
<td>0.04</td>
<td>157</td>
<td>146</td>
<td>NA</td>
<td>2</td>
<td>0.4731</td>
</tr>
<tr>
<td>0.05</td>
<td>157</td>
<td>146</td>
<td>NA</td>
<td>2</td>
<td>0.4710</td>
</tr>
<tr>
<td>0.06</td>
<td>141</td>
<td>131</td>
<td>NA</td>
<td>2</td>
<td>0.4190</td>
</tr>
<tr>
<td>0.07</td>
<td>141</td>
<td>131</td>
<td>NA</td>
<td>2</td>
<td>0.4202</td>
</tr>
<tr>
<td>0.08</td>
<td>141</td>
<td>131</td>
<td>NA</td>
<td>2</td>
<td>0.4251</td>
</tr>
<tr>
<td>0.09</td>
<td>141</td>
<td>131</td>
<td>NA</td>
<td>2</td>
<td>0.4280</td>
</tr>
<tr>
<td>0.10</td>
<td>141</td>
<td>131</td>
<td>NA</td>
<td>2</td>
<td>0.4182</td>
</tr>
</tbody>
</table>

Table 7.8: Example 2. Using Hager and Zhang-NR

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>241</td>
<td>153</td>
<td>NA</td>
<td>3</td>
<td>0.6960</td>
</tr>
<tr>
<td>0.02</td>
<td>274</td>
<td>191</td>
<td>NA</td>
<td>3</td>
<td>0.8317</td>
</tr>
<tr>
<td>0.03</td>
<td>259</td>
<td>191</td>
<td>NA</td>
<td>3</td>
<td>0.7582</td>
</tr>
<tr>
<td>0.04</td>
<td>367</td>
<td>249</td>
<td>NA</td>
<td>4</td>
<td>1.0540</td>
</tr>
<tr>
<td>0.05</td>
<td>323</td>
<td>205</td>
<td>NA</td>
<td>4</td>
<td>0.9552</td>
</tr>
<tr>
<td>0.06</td>
<td>361</td>
<td>229</td>
<td>NA</td>
<td>4</td>
<td>1.0626</td>
</tr>
<tr>
<td>0.07</td>
<td>310</td>
<td>185</td>
<td>NA</td>
<td>4</td>
<td>1.0313</td>
</tr>
<tr>
<td>0.08</td>
<td>346</td>
<td>232</td>
<td>NA</td>
<td>4</td>
<td>1.0185</td>
</tr>
<tr>
<td>0.09</td>
<td>345</td>
<td>203</td>
<td>NA</td>
<td>4</td>
<td>1.0135</td>
</tr>
<tr>
<td>0.10</td>
<td>317</td>
<td>207</td>
<td>NA</td>
<td>4</td>
<td>0.9097</td>
</tr>
</tbody>
</table>

Table 7.9: Example 2. Using Hager and Zhang-LS
### 7.2. Numerical Results for Example 2

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>211</td>
<td>110</td>
<td>NA</td>
<td>2</td>
<td>0.5789</td>
</tr>
<tr>
<td>0.02</td>
<td>284</td>
<td>145</td>
<td>NA</td>
<td>2</td>
<td>0.8319</td>
</tr>
<tr>
<td>0.03</td>
<td>123</td>
<td>86</td>
<td>NA</td>
<td>2</td>
<td>0.3443</td>
</tr>
<tr>
<td>0.04</td>
<td>170</td>
<td>100</td>
<td>NA</td>
<td>2</td>
<td>0.4905</td>
</tr>
<tr>
<td>0.05</td>
<td>276</td>
<td>154</td>
<td>NA</td>
<td>3</td>
<td>0.8264</td>
</tr>
<tr>
<td>0.06</td>
<td>223</td>
<td>119</td>
<td>NA</td>
<td>3</td>
<td>0.6589</td>
</tr>
<tr>
<td>0.07</td>
<td>290</td>
<td>156</td>
<td>NA</td>
<td>3</td>
<td>0.8716</td>
</tr>
<tr>
<td>0.08</td>
<td>186</td>
<td>116</td>
<td>NA</td>
<td>3</td>
<td>0.5654</td>
</tr>
<tr>
<td>0.09</td>
<td>222</td>
<td>128</td>
<td>NA</td>
<td>3</td>
<td>0.6604</td>
</tr>
<tr>
<td>0.10</td>
<td>284</td>
<td>159</td>
<td>NA</td>
<td>3</td>
<td>0.8317</td>
</tr>
</tbody>
</table>

Table 7.10: Example 2. Using Hager and Zhang-$\phi$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>228</td>
<td>107</td>
<td>NA</td>
<td>2</td>
<td>0.6300</td>
</tr>
<tr>
<td>0.02</td>
<td>244</td>
<td>129</td>
<td>NA</td>
<td>2</td>
<td>0.7546</td>
</tr>
<tr>
<td>0.03</td>
<td>272</td>
<td>139</td>
<td>NA</td>
<td>3</td>
<td>0.7686</td>
</tr>
<tr>
<td>0.04</td>
<td>229</td>
<td>123</td>
<td>NA</td>
<td>3</td>
<td>0.6872</td>
</tr>
<tr>
<td>0.05</td>
<td>284</td>
<td>153</td>
<td>NA</td>
<td>3</td>
<td>0.8595</td>
</tr>
<tr>
<td>0.06</td>
<td>268</td>
<td>146</td>
<td>NA</td>
<td>3</td>
<td>0.8018</td>
</tr>
<tr>
<td>0.07</td>
<td>278</td>
<td>148</td>
<td>NA</td>
<td>3</td>
<td>0.8166</td>
</tr>
<tr>
<td>0.08</td>
<td>243</td>
<td>136</td>
<td>NA</td>
<td>3</td>
<td>0.7295</td>
</tr>
<tr>
<td>0.09</td>
<td>264</td>
<td>140</td>
<td>NA</td>
<td>3</td>
<td>0.7802</td>
</tr>
<tr>
<td>0.10</td>
<td>276</td>
<td>147</td>
<td>NA</td>
<td>3</td>
<td>0.8131</td>
</tr>
</tbody>
</table>

Table 7.11: Example 2. Using Hager and Zhang-Bregman Function
### 7.2. Numerical Results for Example 2

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>3451</td>
<td>3322</td>
<td>65</td>
<td>2</td>
<td>0.7665</td>
</tr>
<tr>
<td>0.02</td>
<td>3195</td>
<td>3067</td>
<td>60</td>
<td>2</td>
<td>0.7287</td>
</tr>
<tr>
<td>0.03</td>
<td>767</td>
<td>721</td>
<td>14</td>
<td>3</td>
<td>0.3272</td>
</tr>
<tr>
<td>0.04</td>
<td>2279</td>
<td>2149</td>
<td>42</td>
<td>3</td>
<td>0.6296</td>
</tr>
<tr>
<td>0.05</td>
<td>4282</td>
<td>4140</td>
<td>81</td>
<td>4</td>
<td>0.8793</td>
</tr>
<tr>
<td>0.06</td>
<td>883</td>
<td>825</td>
<td>16</td>
<td>4</td>
<td>0.3217</td>
</tr>
<tr>
<td>0.07</td>
<td>830</td>
<td>774</td>
<td>15</td>
<td>4</td>
<td>0.3223</td>
</tr>
<tr>
<td>0.08</td>
<td>3107</td>
<td>2967</td>
<td>58</td>
<td>4</td>
<td>0.7920</td>
</tr>
<tr>
<td>0.09</td>
<td>2698</td>
<td>2559</td>
<td>50</td>
<td>4</td>
<td>0.7106</td>
</tr>
<tr>
<td>0.10</td>
<td>2545</td>
<td>2406</td>
<td>47</td>
<td>4</td>
<td>0.9290</td>
</tr>
</tbody>
</table>

Table 7.12: Example 2. Using Hager and Zhang-Quadratic Function

<table>
<thead>
<tr>
<th>Method</th>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-NR</td>
<td>0.02</td>
<td>67</td>
<td>64</td>
<td>NA</td>
<td>1</td>
<td>0.2249</td>
</tr>
<tr>
<td>Hager and Zhang-LS</td>
<td>0.01</td>
<td>241</td>
<td>153</td>
<td>NA</td>
<td>3</td>
<td>0.6960</td>
</tr>
<tr>
<td>Hager and Zhang-$\phi$</td>
<td>0.03</td>
<td>123</td>
<td>86</td>
<td>NA</td>
<td>2</td>
<td>0.3443</td>
</tr>
<tr>
<td>Hager and Zhang-Breg</td>
<td>0.01</td>
<td>228</td>
<td>107</td>
<td>NA</td>
<td>2</td>
<td>0.6300</td>
</tr>
<tr>
<td>Hager and Zhang-Quad</td>
<td>0.03</td>
<td>767</td>
<td>721</td>
<td>14</td>
<td>3</td>
<td>0.3272</td>
</tr>
</tbody>
</table>

Table 7.13: New Results of Example Two
## 7.3 Numerical Results for Example 3

<table>
<thead>
<tr>
<th>Method</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-NR</td>
<td>348</td>
<td>295</td>
<td>NA</td>
<td>5</td>
<td>0.8531</td>
</tr>
<tr>
<td>Hager and Zhang-LS</td>
<td>361</td>
<td>300</td>
<td>NA</td>
<td>6</td>
<td>0.8809</td>
</tr>
<tr>
<td>Hager and Zhang-ϕ</td>
<td>324</td>
<td>282</td>
<td>NA</td>
<td>6</td>
<td>0.9049</td>
</tr>
<tr>
<td>Hager and Zhang-Bregman</td>
<td>280</td>
<td>267</td>
<td>NA</td>
<td>5</td>
<td>0.7294</td>
</tr>
<tr>
<td>Hager and Zhang-Quadratic</td>
<td>4458</td>
<td>4351</td>
<td>85</td>
<td>6</td>
<td>0.7085</td>
</tr>
<tr>
<td>Han</td>
<td>64538</td>
<td>62360</td>
<td>1222</td>
<td>22</td>
<td>19.1101</td>
</tr>
<tr>
<td>Li and Yuan</td>
<td>4989</td>
<td>3831</td>
<td>NA</td>
<td>1136</td>
<td>27.2631</td>
</tr>
</tbody>
</table>

Table 7.14: Results of Example 3.

<table>
<thead>
<tr>
<th>β</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>351</td>
<td>298</td>
<td>NA</td>
<td>6</td>
<td>0.8701</td>
</tr>
<tr>
<td>0.02</td>
<td>352</td>
<td>299</td>
<td>NA</td>
<td>7</td>
<td>0.8981</td>
</tr>
<tr>
<td>0.03</td>
<td>347</td>
<td>294</td>
<td>NA</td>
<td>5</td>
<td>0.8730</td>
</tr>
<tr>
<td>0.04</td>
<td>349</td>
<td>296</td>
<td>NA</td>
<td>6</td>
<td>0.8752</td>
</tr>
<tr>
<td>0.05</td>
<td>348</td>
<td>295</td>
<td>NA</td>
<td>5</td>
<td>0.8720</td>
</tr>
<tr>
<td>0.06</td>
<td>348</td>
<td>295</td>
<td>NA</td>
<td>5</td>
<td>0.8940</td>
</tr>
<tr>
<td>0.07</td>
<td>351</td>
<td>298</td>
<td>NA</td>
<td>6</td>
<td>0.9015</td>
</tr>
<tr>
<td>0.08</td>
<td>351</td>
<td>298</td>
<td>NA</td>
<td>6</td>
<td>0.9348</td>
</tr>
<tr>
<td>0.09</td>
<td>351</td>
<td>298</td>
<td>NA</td>
<td>6</td>
<td>0.9046</td>
</tr>
<tr>
<td>0.10</td>
<td>351</td>
<td>298</td>
<td>NA</td>
<td>6</td>
<td>0.9113</td>
</tr>
</tbody>
</table>

Table 7.15: Example 3. Using Hager and Zhang-NR
### 7.3. Numerical Results for Example 3

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>380</td>
<td>311</td>
<td>NA</td>
<td>5</td>
<td>0.9410</td>
</tr>
<tr>
<td>0.02</td>
<td>325</td>
<td>282</td>
<td>NA</td>
<td>6</td>
<td>0.8062</td>
</tr>
<tr>
<td>0.03</td>
<td>353</td>
<td>283</td>
<td>NA</td>
<td>5</td>
<td>0.8806</td>
</tr>
<tr>
<td>0.04</td>
<td>402</td>
<td>337</td>
<td>NA</td>
<td>6</td>
<td>0.9975</td>
</tr>
<tr>
<td>0.05</td>
<td>361</td>
<td>300</td>
<td>NA</td>
<td>6</td>
<td>0.9101</td>
</tr>
<tr>
<td>0.06</td>
<td>404</td>
<td>331</td>
<td>NA</td>
<td>6</td>
<td>1.0350</td>
</tr>
<tr>
<td>0.07</td>
<td>411</td>
<td>341</td>
<td>NA</td>
<td>6</td>
<td>1.0531</td>
</tr>
<tr>
<td>0.08</td>
<td>370</td>
<td>320</td>
<td>NA</td>
<td>6</td>
<td>0.9740</td>
</tr>
<tr>
<td>0.09</td>
<td>380</td>
<td>311</td>
<td>NA</td>
<td>7</td>
<td>0.9960</td>
</tr>
<tr>
<td>0.10</td>
<td>370</td>
<td>297</td>
<td>NA</td>
<td>7</td>
<td>0.9615</td>
</tr>
</tbody>
</table>

Table 7.16: Example 3. Using Hager and Zhang-LS

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>F</td>
<td>F</td>
<td>NA</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>0.02</td>
<td>361</td>
<td>295</td>
<td>NA</td>
<td>7</td>
<td>0.9852</td>
</tr>
<tr>
<td>0.03</td>
<td>391</td>
<td>346</td>
<td>NA</td>
<td>5</td>
<td>1.0641</td>
</tr>
<tr>
<td>0.04</td>
<td>591</td>
<td>392</td>
<td>NA</td>
<td>9</td>
<td>1.5917</td>
</tr>
<tr>
<td>0.05</td>
<td>324</td>
<td>282</td>
<td>NA</td>
<td>6</td>
<td>0.8934</td>
</tr>
<tr>
<td>0.06</td>
<td>411</td>
<td>348</td>
<td>NA</td>
<td>6</td>
<td>1.1268</td>
</tr>
<tr>
<td>0.07</td>
<td>335</td>
<td>273</td>
<td>NA</td>
<td>6</td>
<td>0.9551</td>
</tr>
<tr>
<td>0.08</td>
<td>351</td>
<td>294</td>
<td>NA</td>
<td>6</td>
<td>1.0124</td>
</tr>
<tr>
<td>0.09</td>
<td>342</td>
<td>292</td>
<td>NA</td>
<td>6</td>
<td>0.9878</td>
</tr>
<tr>
<td>0.10</td>
<td>390</td>
<td>331</td>
<td>NA</td>
<td>5</td>
<td>1.0826</td>
</tr>
</tbody>
</table>

Table 7.17: Example 3. Using Hager and Zhang-$\phi$
### 7.3. Numerical Results for Example 3

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>342</td>
<td>293</td>
<td>NA</td>
<td>5</td>
<td>0.8882</td>
</tr>
<tr>
<td>0.02</td>
<td>366</td>
<td>301</td>
<td>NA</td>
<td>7</td>
<td>0.9787</td>
</tr>
<tr>
<td>0.03</td>
<td>399</td>
<td>336</td>
<td>NA</td>
<td>6</td>
<td>1.0468</td>
</tr>
<tr>
<td>0.04</td>
<td>361</td>
<td>295</td>
<td>NA</td>
<td>6</td>
<td>0.9211</td>
</tr>
<tr>
<td>0.05</td>
<td>280</td>
<td>267</td>
<td>NA</td>
<td>5</td>
<td>0.7284</td>
</tr>
<tr>
<td>0.06</td>
<td>366</td>
<td>296</td>
<td>NA</td>
<td>6</td>
<td>0.9368</td>
</tr>
<tr>
<td>0.07</td>
<td>397</td>
<td>331</td>
<td>NA</td>
<td>6</td>
<td>1.0432</td>
</tr>
<tr>
<td>0.08</td>
<td>367</td>
<td>317</td>
<td>NA</td>
<td>6</td>
<td>0.9604</td>
</tr>
<tr>
<td>0.09</td>
<td>386</td>
<td>316</td>
<td>NA</td>
<td>6</td>
<td>0.9845</td>
</tr>
<tr>
<td>0.10</td>
<td>370</td>
<td>297</td>
<td>NA</td>
<td>7</td>
<td>0.9615</td>
</tr>
</tbody>
</table>

Table 7.18: Example 3. Using Hager and Zhang-Brgman Function

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>0.02</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>0.03</td>
<td>3793</td>
<td>3686</td>
<td>72</td>
<td>5</td>
<td>0.6623</td>
</tr>
<tr>
<td>0.04</td>
<td>4407</td>
<td>4300</td>
<td>84</td>
<td>6</td>
<td>0.7142</td>
</tr>
<tr>
<td>0.05</td>
<td>4458</td>
<td>4351</td>
<td>85</td>
<td>6</td>
<td>0.7405</td>
</tr>
<tr>
<td>0.06</td>
<td>380</td>
<td>373</td>
<td>7</td>
<td>6</td>
<td>0.1414</td>
</tr>
<tr>
<td>0.07</td>
<td>433</td>
<td>425</td>
<td>8</td>
<td>6</td>
<td>0.1513</td>
</tr>
<tr>
<td>0.08</td>
<td>3950</td>
<td>3843</td>
<td>75</td>
<td>6</td>
<td>0.7862</td>
</tr>
<tr>
<td>0.09</td>
<td>434</td>
<td>426</td>
<td>8</td>
<td>6</td>
<td>0.1701</td>
</tr>
<tr>
<td>0.10</td>
<td>4565</td>
<td>4457</td>
<td>87</td>
<td>7</td>
<td>0.8098</td>
</tr>
</tbody>
</table>

Table 7.19: Example 3. Using Hager and Zhang-Quadratic Function
7.4 Numerical Results for Example 4

<table>
<thead>
<tr>
<th>Method</th>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-NR</td>
<td>0.03</td>
<td>347</td>
<td>294</td>
<td>NA</td>
<td>5</td>
<td>0.8730</td>
</tr>
<tr>
<td>Hager and Zhang-LS</td>
<td>0.02</td>
<td>325</td>
<td>282</td>
<td>NA</td>
<td>6</td>
<td>0.8062</td>
</tr>
<tr>
<td>Hager and Zhang-$\phi$</td>
<td>0.05</td>
<td>324</td>
<td>282</td>
<td>NA</td>
<td>6</td>
<td>0.8934</td>
</tr>
<tr>
<td>Hager and Zhang-Breg</td>
<td>0.05</td>
<td>280</td>
<td>267</td>
<td>NA</td>
<td>5</td>
<td>0.7284</td>
</tr>
<tr>
<td>Hager and Zhang-Quad</td>
<td>0.06</td>
<td>380</td>
<td>373</td>
<td>7</td>
<td>6</td>
<td>0.1414</td>
</tr>
</tbody>
</table>

Table 7.20: New Results of Example 3.

7.4 Numerical Results for Example 4

<table>
<thead>
<tr>
<th>Method</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-NR</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.4864</td>
</tr>
<tr>
<td>Hager and Zhang-LS</td>
<td>404</td>
<td>224</td>
<td>NA</td>
<td>3</td>
<td>1.4113</td>
</tr>
<tr>
<td>Hager and Zhang-$\phi$</td>
<td>198</td>
<td>132</td>
<td>NA</td>
<td>2</td>
<td>0.6776</td>
</tr>
<tr>
<td>Hager and Zhang-Bregman</td>
<td>359</td>
<td>217</td>
<td>NA</td>
<td>3</td>
<td>1.2298</td>
</tr>
<tr>
<td>Hager and Zhang-Quadratic</td>
<td>15028</td>
<td>14751</td>
<td>146</td>
<td>3</td>
<td>2.1026</td>
</tr>
<tr>
<td>Han</td>
<td>386426</td>
<td>382367</td>
<td>3785</td>
<td>41</td>
<td>55.9540</td>
</tr>
<tr>
<td>Li and Yuan</td>
<td>3571</td>
<td>2773</td>
<td>NA</td>
<td>786</td>
<td>24.2631</td>
</tr>
</tbody>
</table>

Table 7.21: Results of Example Four
### Table 7.22: Example 4. Using Hager and Zhang-NR

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.4990</td>
</tr>
<tr>
<td>0.02</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.5012</td>
</tr>
<tr>
<td>0.03</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.4989</td>
</tr>
<tr>
<td>0.04</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.5053</td>
</tr>
<tr>
<td>0.05</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.5181</td>
</tr>
<tr>
<td>0.06</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.6047</td>
</tr>
<tr>
<td>0.07</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.5391</td>
</tr>
<tr>
<td>0.08</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.5017</td>
</tr>
<tr>
<td>0.09</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.5194</td>
</tr>
<tr>
<td>0.10</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.4994</td>
</tr>
</tbody>
</table>

Table 7.23: Example 4. Using Hager and Zhang-LS

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>234</td>
<td>165</td>
<td>NA</td>
<td>2</td>
<td>0.8231</td>
</tr>
<tr>
<td>0.02</td>
<td>349</td>
<td>250</td>
<td>NA</td>
<td>3</td>
<td>1.1807</td>
</tr>
<tr>
<td>0.03</td>
<td>349</td>
<td>250</td>
<td>NA</td>
<td>3</td>
<td>1.2101</td>
</tr>
<tr>
<td>0.04</td>
<td>386</td>
<td>226</td>
<td>NA</td>
<td>3</td>
<td>1.3014</td>
</tr>
<tr>
<td>0.05</td>
<td>404</td>
<td>224</td>
<td>NA</td>
<td>3</td>
<td>1.3848</td>
</tr>
<tr>
<td>0.06</td>
<td>302</td>
<td>243</td>
<td>NA</td>
<td>3</td>
<td>1.0854</td>
</tr>
<tr>
<td>0.07</td>
<td>243</td>
<td>199</td>
<td>NA</td>
<td>3</td>
<td>0.8474</td>
</tr>
<tr>
<td>0.08</td>
<td>294</td>
<td>220</td>
<td>NA</td>
<td>3</td>
<td>1.0695</td>
</tr>
<tr>
<td>0.09</td>
<td>302</td>
<td>244</td>
<td>NA</td>
<td>3</td>
<td>1.0911</td>
</tr>
<tr>
<td>0.10</td>
<td>271</td>
<td>191</td>
<td>NA</td>
<td>3</td>
<td>0.9510</td>
</tr>
</tbody>
</table>
### 7.4. Numerical Results for Example 4

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>80</td>
<td>74</td>
<td>NA</td>
<td>1</td>
<td>0.3298</td>
</tr>
<tr>
<td>0.02</td>
<td>244</td>
<td>179</td>
<td>NA</td>
<td>2</td>
<td>0.8964</td>
</tr>
<tr>
<td>0.03</td>
<td>213</td>
<td>164</td>
<td>NA</td>
<td>2</td>
<td>0.7726</td>
</tr>
<tr>
<td>0.04</td>
<td>204</td>
<td>136</td>
<td>NA</td>
<td>2</td>
<td>0.7447</td>
</tr>
<tr>
<td>0.05</td>
<td>198</td>
<td>132</td>
<td>NA</td>
<td>2</td>
<td>0.6955</td>
</tr>
<tr>
<td>0.06</td>
<td>302</td>
<td>243</td>
<td>NA</td>
<td>2</td>
<td>0.7074</td>
</tr>
<tr>
<td>0.07</td>
<td>243</td>
<td>199</td>
<td>NA</td>
<td>2</td>
<td>0.7181</td>
</tr>
<tr>
<td>0.08</td>
<td>294</td>
<td>220</td>
<td>NA</td>
<td>2</td>
<td>0.9658</td>
</tr>
<tr>
<td>0.09</td>
<td>302</td>
<td>244</td>
<td>NA</td>
<td>2</td>
<td>1.0016</td>
</tr>
<tr>
<td>0.10</td>
<td>271</td>
<td>191</td>
<td>NA</td>
<td>2</td>
<td>0.8200</td>
</tr>
</tbody>
</table>

Table 7.24: Example 4. Using Hager and Zhang-$\varphi$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>228</td>
<td>160</td>
<td>NA</td>
<td>2</td>
<td>0.8111</td>
</tr>
<tr>
<td>0.02</td>
<td>185</td>
<td>127</td>
<td>NA</td>
<td>2</td>
<td>0.6477</td>
</tr>
<tr>
<td>0.03</td>
<td>182</td>
<td>122</td>
<td>NA</td>
<td>2</td>
<td>0.6384</td>
</tr>
<tr>
<td>0.04</td>
<td>316</td>
<td>230</td>
<td>NA</td>
<td>3</td>
<td>1.1214</td>
</tr>
<tr>
<td>0.05</td>
<td>359</td>
<td>217</td>
<td>NA</td>
<td>3</td>
<td>1.2792</td>
</tr>
<tr>
<td>0.06</td>
<td>280</td>
<td>215</td>
<td>NA</td>
<td>3</td>
<td>0.9896</td>
</tr>
<tr>
<td>0.07</td>
<td>310</td>
<td>236</td>
<td>NA</td>
<td>3</td>
<td>1.0853</td>
</tr>
<tr>
<td>0.08</td>
<td>252</td>
<td>189</td>
<td>NA</td>
<td>3</td>
<td>0.8773</td>
</tr>
<tr>
<td>0.09</td>
<td>207</td>
<td>205</td>
<td>NA</td>
<td>3</td>
<td>1.0186</td>
</tr>
<tr>
<td>0.10</td>
<td>272</td>
<td>195</td>
<td>NA</td>
<td>3</td>
<td>0.9625</td>
</tr>
</tbody>
</table>

Table 7.25: Example 4. Using Hager and Zhang-Bregman Function
### 7.4. Numerical Results for Example 4

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>10506</td>
<td>10306</td>
<td>102</td>
<td>2</td>
<td>1.5502</td>
</tr>
<tr>
<td>0.02</td>
<td>11303</td>
<td>11115</td>
<td>110</td>
<td>2</td>
<td>1.6678</td>
</tr>
<tr>
<td>0.03</td>
<td>10693</td>
<td>10509</td>
<td>104</td>
<td>2</td>
<td>1.6954</td>
</tr>
<tr>
<td>0.04</td>
<td>16145</td>
<td>15864</td>
<td>157</td>
<td>3</td>
<td>2.1762</td>
</tr>
<tr>
<td>0.05</td>
<td>15028</td>
<td>14751</td>
<td>146</td>
<td>3</td>
<td>2.9400</td>
</tr>
<tr>
<td>0.06</td>
<td>16039</td>
<td>15763</td>
<td>156</td>
<td>3</td>
<td>2.4297</td>
</tr>
<tr>
<td>0.07</td>
<td>15633</td>
<td>15359</td>
<td>152</td>
<td>3</td>
<td>2.3839</td>
</tr>
<tr>
<td>0.08</td>
<td>16338</td>
<td>16065</td>
<td>159</td>
<td>3</td>
<td>2.3227</td>
</tr>
<tr>
<td>0.09</td>
<td>15933</td>
<td>15661</td>
<td>155</td>
<td>3</td>
<td>2.3171</td>
</tr>
<tr>
<td>0.10</td>
<td>15831</td>
<td>15560</td>
<td>154</td>
<td>3</td>
<td>2.4638</td>
</tr>
</tbody>
</table>

Table 7.26: Example 4. Using Hager and Zhang-Quadratic

<table>
<thead>
<tr>
<th>Method</th>
<th>( \beta )</th>
<th>Num Func</th>
<th>Num Grads</th>
<th>Num Hess</th>
<th>Num Iters</th>
<th>Time (secs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hager and Zhang-NR</td>
<td>-</td>
<td>143</td>
<td>81</td>
<td>NA</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Hager and Zhang-LS</td>
<td>0.01</td>
<td>234</td>
<td>165</td>
<td>NA</td>
<td>2</td>
<td>0.8231</td>
</tr>
<tr>
<td>Hager and Zhang-( \phi )</td>
<td>0.01</td>
<td>80</td>
<td>74</td>
<td>NA</td>
<td>1</td>
<td>0.3298</td>
</tr>
<tr>
<td>Hager and Zhang-Breg</td>
<td>0.03</td>
<td>182</td>
<td>122</td>
<td>NA</td>
<td>2</td>
<td>0.6384</td>
</tr>
<tr>
<td>Hager and Zhang-Quad</td>
<td>0.01</td>
<td>10506</td>
<td>10306</td>
<td>102</td>
<td>2</td>
<td>1.5502</td>
</tr>
</tbody>
</table>

Table 7.27: New Results of Example 4
Bibliography


