MODELING AND FORECASTING STOCK MARKET PRICES
WITH SIGMOIDAL CURVES

A Thesis

Presented to
The Faculty of the Department of Mathematics
California State University, Los Angeles

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

in
Mathematics

By
Daniel Tran
May 2017
The thesis of Daniel Tran is approved.

Dr. Melisa Hendrata, Committee Chair

Dr. Debasree Raychaudhuri

Dr. Xiaohan Zhang

Dr. Grant Fraser, Department Chair

California State University, Los Angeles

May 2017
ABSTRACT

Modeling and Forecasting Stock Market Prices
with Sigmoidal Curves

By

Daniel Tran

Pricing stock market data is difficult because it is inherently noisy and prone to unexpected events. However, stock market data generally exhibits trends in the medium and long term. A typical successful stock index exhibits an initiation phase, rapid growth, and then saturation whereby the price plateaus. Sigmoidal curves can effectively model and forecast stock market data because it can represent nonlinear stock behavior within confidence interval bounds. This thesis surveys various members of the sigmoidal family of curves and determines which curves best fit stock market data. We explore several techniques to filter our data, such as the moving average, single exponential smoothing, and the Hodrick-Prescott filter. We fit the sigmoidal curves to raw data using the Levenberg-Marquardt algorithm. This thesis aggregates these analysis techniques and apply them towards gauging the opportune time point to sell stocks.
ACKNOWLEDGMENTS

The combination of support from family, friends, and colleagues all culminated towards the completion of my thesis.

First and foremost I would like to express gratitude and appreciation towards my mother Kelly Tran, my father David Hao Tran, and my sister Tina Tran for their support.

I would like to thank my graduate advisor Dr. Melisa Hendrata for guiding and mentoring me. Without her mentorship and encouragement, completion of this thesis would not have been possible. I would also thank members of my committee, Dr. Xiaohan Zhang for providing an economics perspective for my thesis, and Dr. Debasree Raychaudhuri for evaluating my thesis.

I would also like to thank everyone else who I may not have mentioned. The random conversations, quick insight and answers all added nuance to my thesis.
# TABLE OF CONTENTS

Abstract ................................................................................................................. iv

Acknowledgments .................................................................................................. v

List of Tables ......................................................................................................... ix

List of Figures ........................................................................................................ xiii

Chapter

1. Introduction to Stock Market Behavior and Sigmoidal Curves...................... 1

2. Various Members of the Sigmoidal Family of Curves................................. 4

   2.1. The Logistic Model ........................................................................... 5

   2.2. The Gompertz Model ..................................................................... 7

   2.3. The Generalized Logistic Equation ................................................. 10

   2.4. The Chapman-Richards Equation .................................................. 14

   2.5. The Weibull Equation ..................................................................... 19

3. Filtering Noise ........................................................................................... 24

   3.1. Moving Average Filtering ............................................................... 24

   3.2. Single Exponential Smoothing ........................................................ 24

   3.3. The Hodrick-Prescott Filter ........................................................... 26

   3.4. Comparison of Various Smoothing techniques ................................ 28

4. Fitting Data and The Levenberg-Marquardt Algorithm ............................... 33

   4.1. Polynomial Interpolation ................................................................. 33

   4.2. Nonlinear Least Square Problems ................................................... 36

   4.3. Line Search Algorithms ................................................................. 38

      4.3.1. Gradient descent method ..................................................... 40
4.3.2. The Gauss-Newton algorithm .............................................. 40

4.4. Trust-Region Methods (TRM)......................................................... 41

4.4.1. Trust-Region Method Algorithm........................................... 42

4.5. The Levenberg-Marquardt Algorithm......................................... 49

4.5.1. Motivation behind Levenberg-Marquardt Algorithm .......... 49

4.5.2. Trust-Region Subproblem Algorithm ................................. 50

4.5.3. Implementation of Levenberg-Marquardt Algorithm .......... 52

4.5.4. The Levenberg-Marquardt Algorithm .................................. 53

4.5.5. Convergence of The Levenberg-Marquardt Algorithm ....... 54

4.5.6. Computational Example ...................................................... 57

4.6. Results of Fit .................................................................................. 65

5. Forecasting Data ....................................................................................... 68

5.1. Methodology ................................................................................. 68

5.2. Results ............................................................................................ 70

5.3. Future Research ............................................................................ 80

5.4. Data ................................................................................................ 81

5.4.1. Raw Data ............................................................................. 81

5.4.2. Fit of Various Sigmoidal Curves ........................................ 82

5.4.3. Forecast Difference with 1000 Prior Known Days .......... 87

5.4.4. Forecast Difference with 5000 Prior Known Days .......... 90

5.4.5. Forecast Difference with 7000 Prior Known Days .......... 92

5.4.6. MSE with 1000 Prior Known Days ...................................... 94

5.4.7. MSE with 5000 Prior Known Days ...................................... 96
References.................................................................................................................................100

Appendices

A. The Logistic Model .............................................................................................................103
B. The Gompertz Model........................................................................................................105
C. The Generalized Logistic Equation ................................................................................106
D. The Chapman-Richards Model .......................................................................................109
  D.1. Data.................................................................................................................................111
    D.1.1. No filter ..................................................................................................................111
    D.1.2. Hodrick-Prescott Filter .......................................................................................115
    D.1.3. Exponential Smoothing .......................................................................................119
    D.1.4. Moving average .....................................................................................................123
E. MATLAB Code....................................................................................................................127
  E.1. Filters .............................................................................................................................127
    E.1.1. Moving Average ....................................................................................................127
  E.2. Exponential Filter.........................................................................................................128
    E.2.1. The Hodrick-Prescott Filter ................................................................................129
  E.3. Fitting .............................................................................................................................132
    E.3.1. Polynomial Fit .......................................................................................................132
    E.3.2. The Levenburg-Marquart Algorithm ..................................................................132
  E.4. MSE and Difference of Forecast ................................................................................132
# LIST OF TABLES

Table

3.1. MSE of moving average filtering .............................................................. 30
3.2. MSE of single exponential filtering .......................................................... 30
3.3. MSE of Hodrick-Prescott filtering ............................................................ 31

4.1. California State University, Los Angeles full-time student enrollment data from 2005-2015 ................................................................................. 58
4.2. LM algorithm of various sigmoidal curves and their respective MSE ..... 65
4.3. Polynomial algorithms of various degrees and their respective mean square average (MSE) .............................................................................. 67

5.1. Composition of VGENX Mutual Fund .................................................... 69
5.2. Average of Forecast Differences ............................................................... 75
5.3. Standard Deviation of Forecast Differences ............................................. 75
5.4. Histogram of Skews of Forecast Differences ............................................. 75
5.5. Kurtosis .................................................................................................... 76

D.1. MSE with 1000 Prior Known Days .......................................................... 111
D.2. MSE with 2000 Prior Known Days .......................................................... 111
D.3. MSE with 3000 Prior Known Days .......................................................... 111
D.4. MSE with 4000 Prior Known Days .......................................................... 112
D.5. MSE with 5000 Prior Known Days .......................................................... 112
D.6. MSE with 6000 Prior Known Days .......................................................... 112
D.7. MSE with 7000 Prior Known Days .......................................................... 112
D.8. Forecast Difference with 1000 Prior Known Days .................................... 113
D.9. Forecast Difference with 2000 Prior Known Days.................................113
D.10. Forecast Difference with 3000 Prior Known Days.................................113
D.11. Forecast Difference with 4000 Prior Known Days.................................114
D.12. Forecast Difference with 5000 Prior Known Days.................................114
D.13. Forecast Difference with 6000 Prior Known Days.................................114
D.14. Forecast Difference with 7000 Prior Known Days.................................114
D.15. MSE with 1000 Prior Known Days......................................................115
D.16. MSE with 2000 Prior Known Days......................................................115
D.17. MSE with 3000 Prior Known Days......................................................115
D.18. MSE with 4000 Prior Known Days......................................................116
D.19. MSE with 5000 Prior Known Days......................................................116
D.20. MSE with 6000 Prior Known Days......................................................116
D.21. MSE with 7000 Prior Known Days......................................................116
D.22. Forecast Difference with 1000 Prior Known Days.................................117
D.23. Forecast Difference with 2000 Prior Known Days.................................117
D.24. Forecast Difference with 3000 Prior Known Days.................................117
D.25. Forecast Difference with 4000 Prior Known Days.................................118
D.26. Forecast Difference with 5000 Prior Known Days.................................118
D.27. Forecast Difference with 6000 Prior Known Days.................................118
D.28. Forecast Difference with 7000 Prior Known Days.................................118
D.29. MSE with 1000 Prior Known Days......................................................119
D.30. MSE with 2000 Prior Known Days......................................................119
D.31. MSE with 3000 Prior Known Days......................................................119
D.32. MSE with 4000 Prior Known Days .......................................................... 120
D.33. MSE with 5000 Prior Known Days .......................................................... 120
D.34. MSE with 6000 Prior Known Days .......................................................... 120
D.35. MSE with 7000 Prior Known Days .......................................................... 120
D.36. Forecast Difference with 1000 Prior Known Days .................................. 121
D.37. Forecast Difference with 2000 Prior Known Days .................................. 121
D.38. Forecast Difference with 3000 Prior Known Days .................................. 121
D.39. Forecast Difference with 4000 Prior Known Days .................................. 122
D.40. Forecast Difference with 5000 Prior Known Days .................................. 122
D.41. Forecast Difference with 6000 Prior Known Days .................................. 122
D.42. Forecast Difference with 7000 Prior Known Days .................................. 122
D.43. MSE with 1000 Prior Known Days .......................................................... 123
D.44. MSE with 2000 Prior Known Days .......................................................... 123
D.45. MSE with 3000 Prior Known Days .......................................................... 123
D.46. MSE with 4000 Prior Known Days .......................................................... 124
D.47. MSE with 5000 Prior Known Days .......................................................... 124
D.48. MSE with 6000 Prior Known Days .......................................................... 124
D.49. MSE with 7000 Prior Known Days .......................................................... 124
D.50. Forecast Difference with 1000 Prior Known Days .................................. 125
D.51. Forecast Difference with 2000 Prior Known Days .................................. 125
D.52. Forecast Difference with 3000 Prior Known Days .................................. 125
D.53. Forecast Difference with 4000 Prior Known Days .................................. 126
D.54. Forecast Difference with 5000 Prior Known Days .................................. 126
D.55. Forecast Difference with 6000 Prior Known Days .................................. 126

D.56. Forecast Difference with 7000 Prior Known Days .................................. 126
LIST OF FIGURES

Figure

2.1. Phase diagram of logistic curve with parameters \( \beta = 5, 6, 7, Y_\infty = 100 \). ................................................................. 5

2.2. Instantaneous growth rate with logistic curve with parameters \( \beta = 5, 6, 7, Y_\infty = 100 \). ................................................................. 6

2.3. Phase diagram of Gompertz model with parameters \( \beta = 5, 6, 7, Y_\infty = 100 \). ................................................................. 8

2.4. Instantaneous growth rate of Gompertz model with parameters \( \beta = 5, 6, 7, Y_\infty = 100 \). ................................................................. 9

2.5. Phase diagram of generalized logistic with parameters \( \beta = 7, r = 0.5, 1.5, 2, Y_\infty = 100 \). ................................................................. 11

2.6. Phase diagram of generalized logistic with parameters \( \beta = 5, 6, 7, r = 1.5, Y_\infty = 100 \). ................................................................. 11

2.7. Instantaneous growth rate of generalized logistic with parameters \( \beta = 7, r = 0.5, 1.5, 2, Y_\infty = 100 \). ................................................................. 12

2.8. Instantaneous growth rate of generalized logistic with parameters \( \beta = 5, 6, 7, r = 1.5, Y_\infty = 100 \). ................................................................. 13

2.9. Chapman–Richards phase diagram with \( m = -1, \lambda = .01, .1, 1, Y_\infty = 100 \). ................................................................. 15

2.10. Chapman–Richards phase diagram with \( m = -1, -.1, -.01, \lambda = .1, Y_\infty = 100 \). ................................................................. 16

2.11. Chapman–Richards instantaneous growth rate with \( m = -1, \lambda = .01, .1, 1, Y_\infty = 100 \). ................................................................. 18
2.12. Chapman–Richards instantaneous growth rate with \( m = -1, -1, -0.1, \lambda = 0.1, Y_\infty = 100 \). ........................................................................................................ 18

2.13. Weibull phase diagram with parameters \( \alpha = .1, .01, .001, \beta = 7, \gamma = 1/5, Y_\infty = 100 \). ........................................................................................................ 19

2.14. Weibull phase diagram with parameters \( \alpha = .001, \beta = 5, 6, 7, \gamma = 1/5, Y_\infty = 100 \). ........................................................................................................ 20

2.15. Weibull phase diagram with parameters \( \alpha = .001, \beta = 7, \gamma = 1/3, 1/5, 1/7, Y_\infty = 100 \). ........................................................................................................ 20

2.16. Weibull instantaneous growth rate with parameters \( \alpha = .1, .01, .001, \beta = 7, \gamma = 1/5, Y_\infty = 100 \). ........................................................................................................ 22

2.17. Weibull instantaneous growth rate with parameters \( \alpha = .001, \beta = 5, 6, 7, \gamma = 1/5, Y_\infty = 100 \). ........................................................................................................ 22

2.18. Weibull instantaneous growth rate with parameters \( \alpha = .001, \beta = 7, \gamma = 1/3, 1/5, 1/7, Y_\infty = 100 \). ........................................................................................................ 23

3.1. Example of single exponential smoothing filter........................................ 25

3.2. Plot of moving average filter with various \( k \) days. ................................... 29

3.3. Plot of single exponential filter with various \( \alpha \). ....................................... 30

3.4. Plot of Hodrick-Prescott filter with various \( \lambda \). ........................................ 31

4.1. LM Algorithm fitting on Annual Cal State LA Full-Time Enrollment

Data from 2005 - 2015 .................................................................................. 63

4.2. LM Algorithm fitting on Annual Cal State LA Full-Time Enrollment

Data from 2005 - 2015 .................................................................................. 64
4.3. LM Algorithm of various sigmoidal curves and their respective mean square error (MSE). ................................................................. 65

4.4. Polynomial algorithms of various degrees and their respective mean square error (MSE). ................................................................. 66
CHAPTER 1

Introduction to Stock Market Behavior and Sigmoidal Curves

The stock market is a system that connects buyers and sellers of stock. Stock is partial ownership of a company in exchange for a certain amount of cash. The owner of stock hopes that the value of stock increases in the future in order to sell stock for cash profit. One may guess that the value of a stock is directly tied to the profits a company can generate, but market exchanges announce the price of a stock through a black box algorithm that depends upon buyer’s and seller’s bids and offers. This allows for human psychology and market speculation to be priced into stocks. For instance, suppose there exists stock of a company that sells poultry. If a rumor of avian flu speculates drop in profits, the panic may cause owners of the stock to worry and assume a drop in stock price, even though the outbreak may not infect any chickens. Owners of stock may irrationally sell all their shares before the spread of avian flu takes place.

This thesis will not attempt to forecast stock prices in the short term because human psychology and geopolitical events that can affect stock market prices in unpredictable ways. Stock prices with time frames that are less than a year generally exhibit a random walk. Professor Jeremy J. Siegel generated stock market data with a random walk algorithm and asked stock brokers to identify real data mixed with simulated data. Aside from the October 19th, 1987 crash, none of the brokers could distinguish which was real data [18].

Instead, this thesis will explore long term trends, or time scales of at least one year with daily data. Long term prices of stock indices show a positive correlation.
Recall that a stock index is the sum of the price of every unique stock price. The Dow-Jones Industrial Average (DJIA) is a price-weighted index, meaning the prices of 30 large major US industries are summed together, then divided by the number of firms in the index [18]. Siegel fits a best fit line onto 1997 dollars adjusted data and shows the DJIA increases 1.70% per annum. Notice that this time period covers major events in US history, including The Great Depression, World War 2, oil shortages, and many other unpredictable geopolitical events.

Sigmoidal curves were first used for modeling population dynamics. Sigmoidal curves assume that a population will grow at an increasing rate until it passes an inflection point, then the curves approaches a certain limit, called the carrying capacity. In terms of demographics, this carrying capacity might be the average mortality of a species or the maximum population a given ecosystem can sustain.

In a similar vein, the economy has finite resources and labor for goods and services, so the growth of any particular company will also have a carrying capacity in an economic environment. This paper will demonstrate that sigmoidal curves may be utilized as a tool to predict long term stock market prices.

Stock market data is noisy because of market volatility and general uncertainty about future market conditions. This thesis will follow assumptions outlined by Choliz (2007). Choliz characterizes stock market values following three phases: emergent, inflection, and saturation. The emergent phase is when a stock is initially accelerating in growth, the inflection phase is when the growth rate becomes linear, and the saturation phase is when growth decelerates. Stocks have a lower bound of zero because stock prices cannot be negative. Stocks also have a rapid phase of growth
with an inflection point that defines a decrease in the rate of stock market growth. Stocks also have an upper bound once it saturates the market.

Our sigmoidal growth curve models need to have variable growth rates and asymmetry [2]. Schumpeter observes in advanced economies over two centuries suggest that periods of expansion are generally longer than periods of decline. In this thesis, we will use the Logistic, Gompertz, Weibull, Generalized Logistic, and Chapman-Richards equations as the models to fit stock market data. All of these curves have a positive horizontal asymptotes which define the carrying capacity and a horizontal asymptote that defines a stock market price of minimum of $0. All of these sigmoidal curves exhibit an emergent, inflection, and saturation phase. The inflection points of each of these sigmoidal curves can vary, allowing for asymmetric fits. The Logistic and Gompertz equation have inflection points that are multiplied by a constant. The Weibull, Generalized Logistic, and Chapman-Richards are multiplied by a variable, so these three sigmoidal curves provide flexibility when fitting and forecasting stock market data. This thesis will show that the last three sigmoidal provide better fits and forecasts than the classical Logistic and Gompertz equations.
CHAPTER 2

Various Members of the Sigmoidal Family of Curves

Sigmoidal curves have initially been used to model the growth of biological species populating a given ecosystem with limited resources. The economy similarly has finite resources for goods and services, so the growth of any particular company must have a carrying capacity in unconstrained economic environment. This metaphor motivates the use of sigmoidal curves to model stock market prices. We need to find a function that accelerates initially as it grows, then decelerates as the size of a stock approaches a limit. The sigmoidal curves exhibit this pattern. The term ”sigmoidal” literally means s-shaped.

The inflection point is the turning point where the rate of growth starts to decrease. The Logistic and Gompertz equations are classic examples of sigmoidal curves. The problem with these functions is that the inflection point, \( Y_{inflection} \), is a fixed product between the carrying capacity and a constant. The Generalized Logistic, Chapman-Richards and Weibull equations have inflection points that are dependent upon some variables, so the inflection point is adjustable along the \( x \)-axis and \( y \)-axis.

This chapter will explore the phase diagram and instantaneous growth rate for each type of curves. The phase diagram is the derivative of the closed form solution, \( \frac{dY_t}{dt} \), whose unit is \([\text{amount}] / [\text{unit time}]\). The inflection point occurs at the maximum value of the phase diagram. In all of our graphs, when \( Y_t \) is at the carrying capacity \( Y_\infty \), the growth rate must necessarily be zero. Growth does not occur past the carrying capacity for sigmoidal curves.

The instantaneous growth rate divides \( \frac{dY_t}{dt} \) with \( Y_t \), with units of \([1] / [\text{unit time}]\). This
can be interpreted as the percentage change of $Y_t$ per unit time forward.

## 2.1 The Logistic Model

Given the closed form of the logistic model:

$$Y(t) = Y_t = \frac{Y_\infty}{1 + \alpha e^{-\beta t}}, \quad t \geq 0$$

(2.1)

where $\alpha, \beta$ are constant growth parameters, with $\beta$ being the maximum growth rate, and $Y_\infty$ is the carrying capacity. The derivatives for the logistic models are given by

$$\frac{dY_t}{dt} = \frac{\beta}{Y_\infty} Y_t(Y_\infty - Y_t)$$

(2.2)

$$\frac{d^2 Y_t}{dt^2} = \frac{\beta}{Y_\infty} (Y_\infty - 2Y_t) \frac{dY_t}{dt}.$$  

(2.3)

![Figure 2.1: Phase diagram of logistic curve with parameters $\beta = 5, 6, 7, Y_\infty = 100$.](image)

Due to symmetry, the maximum of $\frac{dY_t}{dt}$ occurs at the midpoint between 0 and $Y_\infty$, as shown in the phase diagram in Figure 2.1. Even though the height of the
maximum can change with $\beta$, the inflection point $t_{\text{inflection}}$ is fixed. The $y$-value of the inflection point occurs at $Y_t = \frac{Y_\infty}{2}$, that is when $\frac{d^2Y_t}{dt^2} = 0$. Substituting this value into the closed form of the logistic equation (2.1) gives $t = \frac{1}{\beta} \ln(\alpha)$. Hence, the inflection point occurs at

$$(t_{\text{inflection}}, Y_{\text{inflection}}) = \left(\frac{1}{\beta} \ln(\alpha), \frac{Y_\infty}{2}\right).$$

(2.4)

The instantaneous growth rate is

$$\frac{dY_t}{Y_t} = \frac{\beta}{Y_\infty} (Y_\infty - Y_t).$$

(2.5)

Figure 2.2: Instantaneous growth rate with logistic curve with parameters $\beta = 5, 6, 7, Y_\infty = 100$.

Notice that $Y_{\text{inflection}}$ is dependent only on the carrying capacity $Y_\infty$, sometimes referred to as the ceiling value. To realistically model stock prices, we need functions that are more malleable where we can adjust the inflection points, and whose curves that are not necessarily symmetric.
2.2 The Gompertz Model

The closed form of the Gompertz model is:

\[ Y_t = Y_\infty e^{-\alpha e^{-\beta t}}, \quad t \geq 0 \]  \hspace{1cm} (2.6)

where \( \alpha \) and \( \beta \) are constant growth parameters, and \( Y_\infty > 0 \).

Manipulation of the closed form solution (2.6) will be useful for understanding the derivatives of the Gompertz equations. Note that

\[
\begin{align*}
Y_t &= Y_\infty e^{-\alpha e^{-\beta t}} \\
\frac{Y_t}{Y_\infty} &= e^{-\alpha e^{-\beta t}} \\
\frac{Y_\infty}{Y_t} &= e^{\alpha e^{-\beta t}} \\
\ln \left( \frac{Y_\infty}{Y_t} \right) &= \alpha e^{-\beta t} \\
e^{-\beta t} &= \frac{1}{\alpha} \ln \left( \frac{Y_\infty}{Y_t} \right)
\end{align*}
\]

The derivatives of the Gompertz equation are:

\[
\begin{align*}
\frac{dY_t}{dt} &= \alpha \beta e^{-\beta t}Y_t = \beta Y_t \ln \left( \frac{Y_\infty}{Y_t} \right) \hspace{1cm} (2.7) \\
\frac{d^2Y_t}{dt^2} &= \alpha \beta^2 e^{-\beta t}(\alpha e^{-\beta t} - 1)Y_t = \beta^2 \ln \left( \frac{Y_\infty}{Y_t} \right) \left( \ln \left( \frac{Y_\infty}{Y_t} \right) - 1 \right)Y_t \hspace{1cm} (2.8)
\end{align*}
\]
The phase diagram shows that the inflection point occurs at a fixed point on the $x$-axis, the same characteristic as the logistic equation.

The instantaneous growth rate is:

$$\frac{dY_t}{dt} = \alpha \beta e^{-\beta t} = \beta (\ln Y_\infty - \ln Y_t).$$  \hspace{1cm} (2.9)
Figure 2.4: Instantaneous growth rate of Gompertz model with parameters $\beta = 5, 6, 7, Y_\infty = 100$.

The instantaneous growth rate has a vertical asymptote at $Y_t = 0$. This is no matter for applications towards the stock market because a stock price is de-listed at zero. Our sigmoidal curves assume that stock will always be greater than zero.

To calculate the inflection point:

\[
0 = \alpha e^{-\beta t} - 1
\]

\[
1 = \alpha e^{-\beta t}
\]

\[
\frac{1}{\alpha} = e^{-\beta t}
\]

\[
\alpha = e^{\beta t}
\]

\[
\beta t = \ln(\alpha)
\]

\[
t_{inflection} = \frac{\ln(\alpha)}{\beta}
\]
Substituting this value into the closed form solution (2.6), we obtain

\[ Y_t = Y_\infty e^{-\alpha e^{-\beta \ln(\alpha)}} \]

\[ Y_t = Y_\infty e^{-\alpha e^{-\ln(\alpha)}} \]

\[ Y_t = Y_\infty e^{-\alpha^{\frac{1}{2}}} \]

\[ Y_{infection} = Y_\infty e^{-1}. \]

So the inflection point occurs at:

\[ (t_{infection}, Y_{infection}) = \left( \frac{\ln(\alpha)}{\beta}, Y_\infty e^{-1} \right). \] (2.10)

2.3 The Generalized Logistic Equation

As derived in Appendix C, the closed form solution of the generalized logistic equation is given by:

\[ Y_t = \frac{Y_\infty}{(1 + \alpha e^{-\beta rt})^{\frac{1}{r}}}, \quad \text{for } t \geq 0 \text{ and } \alpha = \frac{Y_r}{Y_0} - 1. \] (2.11)

Note that the derivatives are:

\[ \frac{dY_t}{dt} = \beta Y_t \left[ 1 - \left( \frac{Y_t}{Y_\infty} \right)^r \right] \] (2.12)

\[ \frac{d^2Y_t}{dt^2} = \beta^2 Y_t \left[ 1 - \left( \frac{Y_t}{Y_\infty} \right)^r \right] \left[ 1 - (r + 1) \left( \frac{Y_t}{Y_\infty} \right)^r \right] \] (2.13)
Figure 2.5: Phase diagram of generalized logistic with parameters $\beta = 7, r = 0.5, 1.5, 2, Y_\infty = 100$.

Figure 2.6: Phase diagram of generalized logistic with parameters $\beta = 5, 6, 7, r = 1.5, Y_\infty = 100$.

The phase diagrams for the generalized logistic equation show it is possible to
shift the maximum along the $x$-axis. The value of the $r$ parameter allows for change of the inflection point to correspond to various values of $Y_t$.

The instantaneous growth rate is:

\[
\frac{dY_t}{dt} = \beta \left[ 1 - \left( \frac{Y_t}{Y_\infty} \right)^r \right]
\]  

(2.14)

Figure 2.7: Instantaneous growth rate of generalized logistic with parameters $\beta = 7, r = 0.5, 1.5, 2, Y_\infty = 100$. 
Figure 2.8: Instantaneous growth rate of generalized logistic with parameters $\beta = 5, 6, 7$, $r = 1.5$, $Y_\infty = 100$.

We can change the concavity of the instantaneous growth rate. When $r > 1$, the instantaneous growth rate decreases at an increasing rate. When $r < 1$, the instantaneous growth rate decreases at a decreasing rate. When $r = 1$, we get back the logistic equation.

To calculate the inflection point:

\[
0 = 1 - (r + 1) \left( \frac{Y_t}{Y_\infty} \right)^r
\]

\[
1 = (r + 1) \left( \frac{Y_t}{Y_\infty} \right)^r
\]

\[
\frac{1}{r+1} = \left( \frac{Y_t}{Y_\infty} \right)^r
\]

\[
\frac{1}{(r + 1)^{1/r}} = \frac{Y_t}{Y_\infty}
\]

\[
Y_{inflection} = \frac{Y_\infty}{(r + 1)^{1/r}}
\]
To calculate \( t \), substitute \( Y_{\text{inflection}} \) into the closed form solution (2.11):

\[
\frac{Y_{\infty}}{(r + 1)^{1/r}} = \frac{Y_{\infty}}{(1 + \alpha e^{-\beta rt})^{1/r}}
\]

\[
(r + 1)^{1/r} = (1 + \alpha e^{-\beta rt})^{1/r}
\]

\[
r = \alpha e^{-\beta rt}
\]

\[
\frac{r}{\alpha} = e^{-\beta rt}
\]

\[
\ln \left( \frac{\alpha}{r} \right) = \beta rt
\]

\[
t_{\text{inflection}} = \frac{1}{\beta r} \ln \left( \frac{\alpha}{r} \right).
\]

So the inflection point for this curve is:

\[
(t_{\text{inflection}}, Y_{\text{inflection}}) = \left( \frac{1}{\beta r} \ln \left( \frac{\alpha}{r} \right), \frac{Y_{\infty}}{(r + 1)^{1/r}} \right).
\]

(2.15)

### 2.4 The Chapman-Richards Equation

The closed form solution of the Chapman–Richards equation is [13]:

\[
Y_t = Y_{\infty}[1 - ae^{-\lambda t}]^m, \quad t \geq 0.
\]

(2.16)

Before calculating the derivatives, we will need the following equations from the closed form solution.

\[
\left( \frac{Y_t}{Y_{\infty}} \right) = [1 - ae^{-\lambda t}]^m
\]

(2.17)

\[
\left( \frac{Y_t}{Y_{\infty}} \right)^{1/m} = 1 - ae^{-\lambda t}
\]

(2.18)

The first and second derivatives are:
\[
\frac{dY_t}{dt} = Y_\infty a \lambda e^{-\lambda t} (1 - ae^{-\lambda t})^{m-1} \\
= m \lambda ae^{-\lambda t} Y_\infty (1 - ae^{-\lambda t})^m \\
= m \lambda Y_t \frac{ae^{-\lambda t}}{(1 - ae^{-\lambda t})^{m-1}} \\
= m \lambda Y_t \left(1 - \left(\frac{Y_t}{Y_\infty}\right)^{1/m}\right) \left(\frac{Y_\infty}{Y_t}\right)^{1/m} \\
= m \lambda Y_t \left(\left(\frac{Y_\infty}{Y_t}\right)^{1/m} - 1\right) \\
= m \lambda Y_t \left(\left(\frac{Y_\infty}{Y_t}\right)^{1/m} - 1\right) \left[(m-1)\left(\frac{Y_\infty}{Y_t}\right)^{1/m} - m\right] 
\]

\[
\frac{d^2Y_t}{dt^2} = m \lambda^2 Y_t \left(\left(\frac{Y_\infty}{Y_t}\right)^{1/m} - 1\right) \left[(m-1)\left(\frac{Y_\infty}{Y_t}\right)^{1/m} - m\right] 
\]

Figure 2.9: Chapman–Richards phase diagram with \( m = -.1, \lambda = .01, .1, 1, Y_\infty = 100 \).
Figure 2.10: Chapman–Richards phase diagram with $m = -1, -0.1, -0.01, \lambda = 0.1, Y_\infty = 100$.

To calculate the inflection point:

$$0 = (m - 1) \left( \frac{Y_\infty}{Y_t} \right)^{1/m} - m$$

$$\left( \frac{Y_\infty}{Y_t} \right)^{1/m} = \frac{m}{m - 1}$$

$$\left( \frac{Y_\infty}{Y_t} \right) = \left( \frac{m}{m - 1} \right)^m$$

$$\frac{Y_t}{Y_\infty} = \left( \frac{m - 1}{m} \right)^m$$

$$Y_{inflection} = Y_\infty \left( \frac{m - 1}{m} \right)^m$$
Direct substitution of $Y_{inflection}$ to the closed form (2.16) gives:

$$Y_\infty \left( \frac{m-1}{m} \right)^m = Y_\infty [1 - ae^{-\lambda t}]^m$$

$$\frac{m-1}{m} = 1 - ae^{-\lambda t}$$

$$1 - \frac{1}{m} = 1 - ae^{-\lambda t}$$

$$\frac{1}{am} = e^{-\lambda t}$$

$$am = e^{\lambda t}$$

$$\ln(am) = \lambda t$$

$$t_{inflection} = \frac{\ln(am)}{\lambda}$$

So the inflection point for this curve is:

$$(t_{inflection}, Y_{inflection}) = \left( \frac{\ln(am)}{\lambda}, Y_\infty \left( \frac{m-1}{m} \right)^m \right) \quad (2.21)$$

The instantaneous growth rate from equation (2.19) gives:

$$\frac{dY_t}{dt} = m\lambda \left( \frac{Y_\infty}{Y_t} \right)^{1/m} \left( \frac{Y_\infty}{Y_t} \right)^{1/m} - 1 \right) \quad (2.22)$$
Figure 2.11: Chapman–Richards instantaneous growth rate with $m = -1, \lambda = .01, 1, Y_\infty = 100$.

Figure 2.12: Chapman–Richards instantaneous growth rate with $m = -1, -.1, -.01, \lambda = .1, Y_\infty = 100$.

Since the Chapman-Richards equation is of similar form to the generalized
logistic equation, we have the same patterns for parameter adjustments.

2.5 The Weibull Equation

The closed form solution of the Weibull equation is [13]:

\[
Y_t = Y_\infty - \alpha e^{-\beta t^\gamma}, \quad t \geq 0
\]  \hspace{1cm} (2.23)

Its first and second derivatives are

\[
\frac{dY_t}{dt} = \beta \gamma t^{\gamma-1}(Y_\infty - Y_t)
\]  \hspace{1cm} (2.24)

\[
\frac{d^2Y_t}{dt^2} = \beta \gamma (\gamma - 1) t^{\gamma-2}(Y_\infty - Y_t) - \beta \gamma t^{\gamma-1}\frac{dY_t}{dt}
\]  \hspace{1cm} (2.25)

Figure 2.13: Weibull phase diagram with parameters \( \alpha = .001, .01, .1, \beta = 7, \gamma = 1/5, Y_\infty = 100. \)
Figure 2.14: Weibull phase diagram with parameters $\alpha = .001, \beta = 5, 6, 7, \gamma = 1/5, Y_\infty = 100$.

Figure 2.15: Weibull phase diagram with parameters $\alpha = .001, \beta = 7, \gamma = 1/3, 1/5, 1/7, Y_\infty = 100$. 
To calculate the inflection point:

\[
0 = \beta \gamma t^{\gamma-1} \left[ (\gamma - 1)t^{-1}(Y_\infty - Y_t) - \frac{dY_t}{dt} \right]
\]

\[
0 = (\gamma - 1)t^{-1}(Y_\infty - Y_t) - \frac{dY_t}{dt}
\]

\[
\frac{dY_t}{dt} = (\gamma - 1)t^{-1}(Y_\infty - Y_t)
\]

\[
\beta \gamma t^{\gamma-1}(Y_\infty - Y_t) = (\gamma - 1)t^{-1}(Y_\infty - Y_t)
\]

\[
t^\gamma = \frac{\gamma - 1}{\beta \gamma}
\]

\[
t_{\text{inflection}} = \left( \frac{\gamma - 1}{\beta \gamma} \right)^{1/\gamma}
\]

By direct substitution of \( t_{\text{inflection}} \) into the closed form solution (2.23), we get:

\[
Y_{\text{inflection}} = Y_\infty - \alpha e^{-\frac{\gamma - 1}{\gamma}}
\]  \hspace{1cm} (2.26)

So the inflection point for this curve is:

\[
(t_{\text{inflection}}, Y_{\text{inflection}}) = \left( \left( \frac{\gamma - 1}{\beta \gamma} \right)^{1/\gamma}, Y_\infty - \alpha e^{-\frac{\gamma - 1}{\gamma}} \right). \hspace{1cm} (2.27)
\]

The instantaneous growth rate derived from equation (2.24) is given by

\[
\frac{dY_t}{dt} = \beta \gamma t^{\gamma-1} \left( \frac{Y_\infty}{Y_t} - 1 \right)
\]  \hspace{1cm} (2.28)
Figure 2.16: Weibull instantaneous growth rate with parameters $\alpha = .1, .01, .001, \beta = 7, \gamma = 1/5, Y_\infty = 100$.

Figure 2.17: Weibull instantaneous growth rate with parameters $\alpha = .001, \beta = 5, 6, 7, \gamma = 1/5, Y_\infty = 100$. 
Figure 2.18: Weibull instantaneous growth rate with parameters $\alpha = .001, \beta = 7, \gamma = 1/3, 1/5, 1/7, Y_\infty = 100$. 
CHAPTER 3

Filtering Noise

Before attempting to fit our models into the raw data, we need to smooth out the noise from the data to reduce forecasting error.

3.1 Moving Average Filtering

The simplest smoothing function is the moving average [9]:

\[ F_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^{t} Y_i \]  

(3.1)

where \( Y_i \) is raw data point, \( F_{t+1} \) is the smoothed data, and \( k \) is the number of previous data points to average.

The function takes the arithmetic average of its previous \( k \) data points. If we assume time is initialized at \( t = 0 \), the output of the moving average function starts at \( t = k \). The output needs a minimum of \( k \) input points. This function places equal weight for each previous \( k \) data point.

3.2 Single Exponential Smoothing

The single exponential smoothing function [9] is:

\[ F_{t+1} = F_t + \alpha(Y_t - F_t), \]  

(3.2)

where \( \alpha \) is constant such that \( 0 < \alpha < 1 \), \( F_t \) is the smoothed data, and \( Y_t \) is the raw data. The difference \( Y_t - F_t \) can be regarded as the forecast error for time period \( t \). In this interpretation, the new forecast \( F_{t+1} \) is the previous forecast \( F_t \) plus an adjustment for the error that occurred in the last forecast.
We initialize the smoothing function by either letting $F_1 = Y_2$ or taking the arithmetic average of $k - 1$ terms. The constant $\alpha$ applies a weight on the difference between smoothed data point and raw data at a given time point $t$. An $\alpha$ close to 0 has a small adjustment from the previous forecast error, while an $\alpha$ close to 1 has a large adjustment. Here is a graph that illustrates the single exponential smoothing filter with an arbitrary set of data.

![Single Exponential Smoothing of Arbitrary Data over Time](image.png)

**Figure 3.1: Example of single exponential smoothing filter.**

Notice for this data, a high $\alpha$ looks almost like a transposition of the raw data, shifted to the right on the $x$-axis. On the other extreme, the trend line barely increases relative to the shape of the raw data. Also, low $\alpha$ has very low small fluctuations in
slope in comparison to high alpha.

3.3 The Hodrick-Prescott Filter

The Hodrick-Prescott filter [6] is a technique for finding correlations in economic data by separating raw data into a trend function and a cyclic function.

Kim [7] summarizes the Hodrick-Prescott filter as follows.

Suppose a given set of raw data $y_t$ can be decomposed as follows:

$$y_t = \tau_t + c_t, \quad t = 1, 2, \ldots, T,$$

(3.3)

where $\tau_t$ is the trend component and $c_t$ is the cyclical component. The Hodrick-Prescott filter isolates $c_t$ by minimizing the function

$$f(\tau_1, \tau_2, \ldots, \tau_T) = \left[ \sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} (\tau_{t+1} - 2\tau_t + \tau_{t-1})^2 \right],$$

(3.4)

where $\lambda$ is called the penalty parameter. We want to minimize changes in the growth rate, thereby producing a curve with minimal sudden changes in acceleration. This parameter can be estimated by square rooting the quotient of the percent fluctuation of the cyclic component with the percentage growth rate of one quarter. Quarterly data typically assumes $\lambda = 1600$ because Hodrick and Prescott assumes 5% fluctuation for the cyclical component, and 1/8 % growth for a fiscal quarter. When $\lambda$ approaches 0, the trend component $\tau_t$ matches the raw data, and when $\lambda$ approaches infinity, $\tau_t$ becomes linear, or zero acceleration.

The objective function (3.4) shows two summations. The summation on the left is the variance between raw data and the trend component. The right summation is the variance of the acceleration of the trend component.
To minimize $f$, we set

$$\frac{\partial f}{\partial \tau_1} = \frac{\partial f}{\partial \tau_2} = \ldots = \frac{\partial f}{\partial \tau_T} = 0 \quad (3.5)$$

Note that

$$\frac{\partial f}{\partial \tau_1} = -2(y_1 - \tau_1) + 2\lambda(\tau_3 - 2\tau_2 + \tau_1) = 0$$

This implies

$$y_1 = (1 + \lambda)\tau_1 - 2\lambda\tau_2 + \lambda\tau_3$$

$$= \lambda(\tau_1 - 2\tau_2 + \tau_3) + \tau_1$$

For $\tau_2$:

$$\frac{\partial f}{\partial \tau_2} = -2(y_2 - \tau_2) + 2\lambda(\tau_3 - 2\tau_2 + \tau_1)(-2) + 2\lambda(\tau_4 - 2\tau_3 + \tau_2) = 0$$

This implies

$$y_2 = (-2\lambda)\tau_1 + (1 + 4\lambda + \lambda)\tau_2 + (-2\lambda - 2\lambda)\tau_3 + \lambda\tau_4$$

$$= \lambda(-2\tau_1 + 5\tau_2 - 4\tau_3 + \tau_4) + \tau_2$$

In general,

$$\frac{\partial f}{\partial \tau_k} = -2(y_k - \tau_k) + 2\lambda(\tau_k - 2\tau_{k-1} + \tau_{k-2}) + 2\lambda(\tau_{k+1} - 2\tau_k + \tau_{k-1})(-2)$$

$$+ 2\lambda(\tau_{k+2} - 2\tau_{k+1} + \tau_k) = 0$$

This implies

$$y_k = \lambda\tau_{k+2} + (-2\lambda - 2\lambda)\tau_{k+1} + (1 + \lambda + 4\lambda + \lambda)\tau_k + (-2\lambda - 2\lambda)\tau_{k-1} + \lambda\tau_{k-2}$$

$$= \lambda(\tau_{k+2} - 4\tau_{k+1} + 6\tau_k - 4\tau_{k-1} + \tau_{k-2}) + \tau_k$$
We can now rewrite the minimization function in matrix notation as:

\[ y_T = (\lambda F + I_T)\tau_T \] (3.6)

where \( y_T = (y_1, y_2, \ldots, y_T)^T \) is a \( T \times 1 \) vector of the raw data, \( I_T \) is \( T \times T \) identity matrix, \( \tau_T = (\tau_1, \tau_2, \ldots, \tau_T)^T \) is the \( T \times 1 \) trend component vector and \( F \) is a pentadiagonal symmetric matrix given by

\[
F = \begin{bmatrix}
1 & -2 & 1 & 0 & \ldots & 0 \\
-2 & 5 & -4 & 1 & 0 & \ldots & 0 \\
1 & -4 & 6 & -4 & 1 & 0 & \ldots \\
0 & 1 & -4 & 6 & -4 & 1 & 0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 1 & -4 & 6 & -4 & 1 & 0 & \ldots \\
\ldots & 0 & 1 & -4 & 6 & -4 & 1 & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\end{bmatrix}
\]

From (3.6), the trend component vector can be isolated

\[ \tau_T = (\lambda F + I_T)^{-1}y_T. \] (3.7)

The equation (3.7) has some computational advantages. The only unknown parameter needed to smooth raw data is a single real number \( \lambda \). Since we are smoothing daily data, Ravn and Uhlig [16] shows that \( \lambda = 1600 \left( \frac{\text{365}}{4} \right)^4 = 110930628906.250 \). The pentadiagonal symmetric matrix \( F \) can be easily inverted with fewer flops. The Hodrick-Prescott filter was implemented with MATLAB code given in the appendix [5].

### 3.4 Comparison of Various Smoothing techniques

To see which smoothing technique is best for sigmoidal curve fitting, this paper will use the mean square error as a metric for the best fitting technique. The
The following data set is the daily closing price of Chipotle’s stock price from its initial public offering date, January 26th, 2006, to June 17th, 2016 [30].

The equation for the mean square error [26] is:

\[
\text{MSE} = \frac{1}{T} \sum_{t=1}^{T} (S_t - R_t)^2, \quad (3.8)
\]

where \(S_t\) is the smoothed data and \(R_t\) is the raw data.

Figure 3.2: Plot of moving average filter with various \(k\) days.
Table 3.1: MSE of moving average filtering

<table>
<thead>
<tr>
<th>$k$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>57.796012692925402</td>
</tr>
<tr>
<td>30</td>
<td>480.525902866916</td>
</tr>
<tr>
<td>100</td>
<td>1760.21885253903</td>
</tr>
<tr>
<td>300</td>
<td>4227.2094117137203</td>
</tr>
</tbody>
</table>

Figure 3.3: Plot of single exponential filter with various $\alpha$.

Table 3.2: MSE of single exponential filtering

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>2.60E+02</td>
</tr>
<tr>
<td>0.2</td>
<td>1.33E+02</td>
</tr>
<tr>
<td>0.3</td>
<td>93.3976601</td>
</tr>
<tr>
<td>0.4</td>
<td>74.56602442</td>
</tr>
<tr>
<td>0.5</td>
<td>63.73344863</td>
</tr>
<tr>
<td>0.6</td>
<td>56.91617951</td>
</tr>
<tr>
<td>0.7</td>
<td>56.91617951</td>
</tr>
<tr>
<td>0.8</td>
<td>49.67360402</td>
</tr>
<tr>
<td>0.9</td>
<td>48.08265613</td>
</tr>
</tbody>
</table>
Figure 3.4: Plot of Hodrick-Prescott filter with various $\lambda$.

**Table 3.3: MSE of Hodrick-Prescott filtering**

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>160</td>
<td>44.51601627</td>
</tr>
<tr>
<td>800</td>
<td>63.33548743</td>
</tr>
<tr>
<td>1600</td>
<td>74.14633128</td>
</tr>
<tr>
<td>3200</td>
<td>88.30237894</td>
</tr>
<tr>
<td>16000</td>
<td>1.38E+02</td>
</tr>
<tr>
<td>160000</td>
<td>2.52E+02</td>
</tr>
</tbody>
</table>

The MSE can only measure the extent to which the smoothed data deviates from the raw data. After we explore fitting algorithms used in this paper, the MSE will reveal how well sigmoidal curves fit with the raw data and how well sigmoidal curves forecast data.

For moving average filtering, the choice of using 5, 30, 100, and 300 days is used to approximate the average of a fiscal week, fiscal month, fiscal quarter, and
fiscal year, respectively. The deviation in moving average filtering increases as the number of days averaged increases. For single exponential smoothing, the smoothing deviation decreases as $\alpha$ increases. For the Hodrick-Prescott filter, the MSE increases as $\lambda$ increases.
CHAPTER 4
Fitting Data and The Levenberg-Marquardt Algorithm

This chapter starts with the discussion of polynomial interpolation as one of the basic techniques for curve fitting. Next we look into nonlinear least squares problems that arise in the context of fitting a more general parameterized function to a set of data points by minimizing the sum of the squares of the errors between the data points and the function. The Levenberg-Marquardt algorithm is a standard technique for solving nonlinear least squares problems. We present the derivation of the Levenberg-Marquardt algorithm along with its convergence theorem. A computational example is also presented to illustrate the algorithm.

4.1 Polynomial Interpolation

One of the most common and simplest ways to fit data is by fitting polynomial functions into a given data set. Given a data set \{(x_i, y_i), i = 1, 2 \ldots n\}, we aim to find a \(k\)-th order polynomial, where \(k < n\):

\[
y = a_0 + a_1 x + \cdots + a_k x^k.
\]  

(4.1)

The error \(r\), also called \textit{residual}, is defined to be the difference between the fitted function and the data points. The sum of the square error can be written as

\[
R(a_0, a_1, \ldots, a_k) = r^2 = \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i + \cdots + a_k x_i^k)]^2.
\]  

(4.2)
Note that $R$ is a function of $k + 1$ variables $a_0, a_1, \ldots, a_k$. To minimize $R$, we take the partial derivative with respect to each $a_k$ and set it equal to zero:

$$
\frac{\partial R}{\partial a_0} = -2 \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i + \ldots + a_k x_i^k)] = 0
$$

$$
\frac{\partial R}{\partial a_1} = -2 \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i + \ldots + a_k x_i^k)] x_i = 0
$$

$$
\vdots
$$

$$
\frac{\partial R}{\partial a_k} = -2 \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i + \ldots + a_k x_i^k)] x_i^k = 0
$$

By dividing both sides by the constants and distributing terms we get:

$$
\frac{\partial R}{\partial a_0} = \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i + \ldots + a_k x_i^k)] = 0
$$

$$
\frac{\partial R}{\partial a_1} = \sum_{i=1}^{n} [x_i y_i - (a_0 x_i + a_1 x_i^2 + \ldots + a_k x_i^{k+1})] = 0
$$

$$
\vdots
$$

$$
\frac{\partial R}{\partial a_k} = \sum_{i=1}^{n} [x_i^k y_i - (a_0 x_i^k + a_1 x_i^{k+1} + \ldots + a_k x_i^{2k})] = 0.
$$
We now separate each summation term and move all terms containing $y$ to one side, we get:

$$a_0n + a_1 \sum_{i=1}^{n} x_i + \ldots + a_k \sum_{i=1}^{n} x_i^k = \sum_{i=1}^{n} y_i$$

$$a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 + \ldots + a_k \sum_{i=1}^{n} x_i^{k+1} = \sum_{i=1}^{n} x_i y_i$$

(4.3)

$$\vdots$$

$$a_0 \sum_{i=1}^{n} x_i^k + a_1 \sum_{i=1}^{n} x_i^{k+1} + \ldots + a_k \sum_{i=1}^{n} x_i^{2k} = \sum_{i=1}^{n} x_i^k y_i$$

The above system of equations is called the \textit{normal equations} and can be written in the following matrix form

$$
\begin{bmatrix}
\sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \ldots & \sum_{i=1}^{n} x_i^k \\
\sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \ldots & \sum_{i=1}^{n} x_i^{k+1} \\
\vdots & \vdots & \ddots & \vdots \\
\sum_{i=1}^{n} x_i^k & \sum_{i=1}^{n} x_i^{k+1} & \ldots & \sum_{i=1}^{n} x_i^{2k}
\end{bmatrix}
\begin{bmatrix}
a_0 \\
a_1 \\
\vdots \\
a_k
\end{bmatrix}
= 
\begin{bmatrix}
\sum_{i=1}^{n} y_i \\
\sum_{i=1}^{n} x_i y_i \\
\vdots \\
\sum_{i=1}^{n} x_i^k y_i
\end{bmatrix}.
$$

(4.4)

A \textit{Vandermonde matrix} is a matrix with the terms of a geometric progression in each row. The matrix

$$V = 
\begin{bmatrix}
1 & x_1 & \ldots & x_1^k \\
1 & x_2 & \ldots & x_2^k \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_n & \ldots & x_n^k
\end{bmatrix}
$$

(4.5)

is a Vandermonde matrix. Note that (4.4) can be decomposed in terms of Vandermonde matrix $V$ as shown below:

$$
\begin{bmatrix}
1 & 1 & \ldots & 1 \\
x_1 & x_2 & \ldots & x_n \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{k} & x_2^{k} & \ldots & x_n^{k}
\end{bmatrix}
\begin{bmatrix}
1 & x_1 & \ldots & x_1^k \\
1 & x_2 & \ldots & x_2^k \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_n & \ldots & x_n^k
\end{bmatrix}
\begin{bmatrix}
a_0 \\
a_1 \\
\vdots \\
a_k
\end{bmatrix}
= 
\begin{bmatrix}
1 & 1 & \ldots & 1 \\
x_1 & x_2 & \ldots & x_n \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{k} & x_2^{k} & \ldots & x_n^{k}
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix},
$$

(4.6)

that is,

$$V^T V a = V^T y,$$

(4.7)
where \( \mathbf{a} = [a_0, a_1, \ldots, a_k]^T \) and \( \mathbf{y} = [y_1, y_2, \ldots, y_n]^T \). Therefore, the coefficients \( \mathbf{a} \) can be written as

\[
\mathbf{a} = (V^T V)^{-1} V^T \mathbf{y}.
\] (4.8)

Note that the dimension of \( V \) is \( n \times (k + 1) \) and it easily becomes very large as the number of data points is large. Solving for coefficients \( \mathbf{a} \) from the system (4.7) takes \( O((k + 1)^3) \) using Gaussian elimination. Moreover, the behavior of polynomial functions as \( t \) increases approaches \( \pm \infty \), which is impractical for modeling a carrying capacity. In the next section, we will look at the least square problems that arise from fitting parameterized functions, such as the sigmoidal curves, to a set of data points.

### 4.2 Nonlinear Least Square Problems

Given a set of data points \( \{(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)\} \), the nonlinear least square problem is a problem of finding a function \( p(t, x_1, x_2, \ldots, x_n) \) of \( n \) parameters \( x_1, x_2, \ldots, x_n \) that best fits the data. We want to find the parameter values \( \mathbf{x} = (x_1, x_2, \ldots, x_n) \) through iterative improvement that minimizes the sum of the squares of the errors between the data points and the function. The problem can be formulated as follows:

\[
\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}),
\] (4.9)

where

\[
f(\mathbf{x}) = \frac{1}{2} \sum_{j=1}^{m} r_j^2(\mathbf{x}),
\] (4.10)

\( r_j \) are residuals, or more specifically \( r_j = |\text{raw data} - \text{fitted function}| = |y_j - p(t_j, \mathbf{x})|, j = 1, \ldots, m \). We assume that \( m \geq n \).
The minimization function can be rewritten as:

\[ f(x) = \frac{1}{2} ||r(x)||^2 = \frac{1}{2} r(x)^T r(x), \]

(4.11)

where \( r(x) = (r_1(x), r_2(x), \ldots, r_m(x))^T \).

Recall that the Jacobian \( J(x) \) of \( r \) is the \( m \times n \) matrix of the first partial derivatives, that is,

\[
J = \begin{bmatrix}
\frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \cdots & \frac{\partial r_1}{\partial x_n} \\
\frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & \cdots & \frac{\partial r_2}{\partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial r_m}{\partial x_1} & \frac{\partial r_m}{\partial x_2} & \cdots & \frac{\partial r_m}{\partial x_n}
\end{bmatrix}.
\]

(4.12)

Recall also

\[
\nabla f(x) = \begin{bmatrix}
\frac{\partial f}{\partial x_1} \\
\frac{\partial f}{\partial x_2} \\
\vdots \\
\frac{\partial f}{\partial x_n}
\end{bmatrix} = \begin{bmatrix}
\frac{1}{2} (2r_1(x) \frac{\partial r_1}{\partial x_1} + 2r_2(x) \frac{\partial r_2}{\partial x_1} + \cdots + 2r_m(x) \frac{\partial r_m}{\partial x_1}) \\
\frac{1}{2} (2r_1(x) \frac{\partial r_1}{\partial x_2} + 2r_2(x) \frac{\partial r_2}{\partial x_2} + \cdots + 2r_m(x) \frac{\partial r_m}{\partial x_2}) \\
\vdots \\
r_1 \frac{\partial r_1}{\partial x_n} + r_2 \frac{\partial r_2}{\partial x_n} + \cdots + r_m \frac{\partial r_m}{\partial x_n}
\end{bmatrix}
\]

37
We can now rewrite $\nabla f(x)$ as
\[
\nabla f(x) = \begin{bmatrix}
\frac{\partial r_1}{\partial x_1} & \frac{\partial r_2}{\partial x_1} & \cdots & \frac{\partial r_m}{\partial x_1} \\
\frac{\partial r_1}{\partial x_2} & \frac{\partial r_2}{\partial x_2} & \cdots & \frac{\partial r_m}{\partial x_2} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial r_1}{\partial x_n} & \frac{\partial r_2}{\partial x_n} & \cdots & \frac{\partial r_m}{\partial x_n}
\end{bmatrix}
\begin{bmatrix}
r_1 \\
r_2 \\
\vdots \\
r_m
\end{bmatrix}
= J^T r
\]

\[
= \begin{bmatrix}
r_1, r_2, \ldots, r_m
\end{bmatrix}
\begin{bmatrix}
\nabla r_1 \\
\nabla r_2 \\
\vdots \\
\nabla r_m
\end{bmatrix}, \text{ where } \nabla r_j = \begin{bmatrix}
\frac{\partial r_j}{\partial x_1} \\
\frac{\partial r_j}{\partial x_2} \\
\vdots \\
\frac{\partial r_j}{\partial x_n}
\end{bmatrix}
\]

\[
= r_1 \nabla r_1 + r_2 \nabla r_2 + \cdots + r_m \nabla r_m
\]

\[
= \sum_{j=1}^m r_j \nabla r_j.
\]

The derivatives of $f$ can be expressed in terms of the Jacobian matrix $J(x) = \begin{bmatrix} \frac{\partial r_i}{\partial x_j} \end{bmatrix}$, $1 \leq i \leq m$, $1 \leq j \leq n$, as follows
\[
\nabla f(x) = \sum_{j=1}^m r_j(x) \nabla r_j(x) = J(x)^T r(x) \tag{4.13}
\]
\[
\nabla^2 f(x) = J(x)^T J(x) + \sum_{j=1}^m r_j(x) \nabla^2 r_j(x) \tag{4.14}
\]

In the vicinity of a solution, $r(x)$ is usually small, so the summation in the second term of (4.14) is negligible and $J(x)^T J(x)$ can be taken as an approximation for the Hessian:
\[
\nabla^2 f(x) \approx J(x)^T J(x). \tag{4.15}
\]

4.3 Line Search Algorithms

A general procedure of line search algorithms for function minimization is as follows. We start with an initial guess, $x_0 \in \mathbb{R}^n$, and produce a sequence of points
\{x_k\} that, under appropriate conditions, will converge to a minimizer \(x^*\). At each iteration \(k\), the next iterate \(x_{k+1}\) is determined from the current iterate \(x_k\) as:

\[
x_{k+1} = x_k + \alpha_k p_k
\]

(4.16)

where \(p_k \in \mathbb{R}^n\) is a suitably chosen direction and \(\alpha_k\) is a suitably chosen step size.

In line search algorithms, we first determine the direction \(p_k\), then compute the step size \(\alpha_k\) to determine how far we need to move along that direction. The search direction \(p_k\) can be written in the form

\[
p_k = -B_k^{-1}\nabla f_k,
\]

(4.17)

where \(B_k = B(x_k)\) is an \(n \times n\) matrix and \(\nabla f_k = \nabla f(x_k)\) is the gradient of \(f\) at the current iterate \(x_k\). There are many choices for \(p_k\), but in most line search algorithms, \(p_k\) is chosen to be a descent direction.

**Definition:** Let \(f : \mathbb{R}^n \rightarrow \mathbb{R}\). A vector \(p \in \mathbb{R}^n\) is a descent direction for \(f\) at \(x\) if \(p^T \nabla f(x) < 0\).

Using Taylor’s theorem one can show that if we move in sufficiently small step along the descent direction \(p\), then the function value is reduced. Moreover, since \(p\) is a descent direction, we also have from (4.17)

\[
p^T \nabla f(x) < 0 \Leftrightarrow (-B^{-1}\nabla f(x))^T \nabla f(x) < 0
\]

(4.18)

\[
\Leftrightarrow -\nabla f(x)^T B^{-T} \nabla f(x) < 0
\]

(4.19)

\[
\Leftrightarrow \nabla f(x)^T B^{-T} \nabla f(x) > 0
\]

(4.20)
which implies that $B^{-T}$ is a positive definite matrix and so is $B$.

Two commonly used methods in the family of line search algorithms are the gradient descent and Gauss-Newton methods, which will be described next.

4.3.1 Gradient descent method

In gradient descent method, the direction $p_k$ is chosen to obtain the greatest decrease in $f$. For any direction $p$ with $\|p\| = 1$ we have

$$\nabla f(x)^T p = \|\nabla f(x)\| \|p\| \cos \theta,$$

(4.21)

where $\theta$ is the angle between $p$ and $\nabla f(x)$. Since $-1 \leq \cos \theta \leq 1$, this implies that

$$-\|\nabla f(x)\| \leq \nabla f(x)^T p \leq \|\nabla f(x)\|$$

(4.22)

and hence the greatest decrease of $f$ occurs when

$$\nabla f(x)^T p = -\|\nabla f(x)\|$$

(4.23)

that is,

$$p = \frac{-\nabla f(x)}{\|\nabla f(x)\|}.$$  

(4.24)

This direction $p$ is known as the steepest descent direction. In the form of equation (4.17), the matrix $B = I$, the $n \times n$ identity matrix.

In spite of its simplicity, slow convergence of gradient descent method is one of its major disadvantages, especially for functions with long and narrow valley structures.

4.3.2 The Gauss-Newton algorithm

In Gauss-Newton algorithm, the sum of the square errors is reduced by assuming that the objective function $f$ is locally quadratic and finding the minimum of
the quadratic approximation.

Let $m_k(p_k)$ be the quadratic approximation to $f(x_k + p_k)$ at the point $x_k$. From Taylor’s theorem we have

$$m_k(p_k) = f(x_k) + p_k^T \nabla f_k + \frac{1}{2} p_k^T \nabla^2 f_k p_k.$$  \hfill (4.25)

We seek to find $p_k$ that minimizes $m_k$. Taking the derivative of (4.25) with respect to $p_k$ and setting it equal to 0, we obtain

$$\nabla m_k(p_k) = \nabla f_k + \nabla^2 f_k p_k = 0,$$  \hfill (4.26)

which gives us the Newton’s direction

$$p_k = -(\nabla^2 f_k)^{-1} \nabla f_k.$$  \hfill (4.27)

Gauss-Newton method takes advantage of the special structure of the least square problems. Rather than using the complete second-order Hessian matrix for the quadratic model, the Gauss-Newton method uses an approximation (4.15). Hence, the search direction for Gauss–Newton method is given by:

$$p_k = -(J_k^T J_k)^{-1} \nabla f_k,$$  \hfill (4.28)

where $J_k = J(x_k)$. In the form of equation (4.17), the matrix $B_k = J_k^T J_k$.

### 4.4 Trust-Region Methods (TRM)

Another approach for solving minimization problem is by using the trust region methods. Line search methods calculate a direction towards the minimizer, then figure out the appropriate step size. Trust region methods take the opposite approach. The
trust region algorithm defines a region around an iterate and constructs a model function that approximates the objective function in that region. The algorithm finds the minimizer of the model function and then takes an iterative step.

In other words, for every $k$-th iterate, given the model function $m_k$ of a trust region within $p$ of the current position $x_k$, the algorithm minimizes $m_k(x_k + p)$ with respect to $p$. If sufficient reduction in the function value $f$ is obtained, then $m_k$ is accepted to be a good representation of $f$ in that region. Otherwise the trust region needs to be adjusted accordingly. The goal of the trust region method is to find an approximate trust region radius to arrive at the minimizer $x^*$.

The algorithm for the trust region method is as follows [12]:

4.4.1 Trust-Region Method Algorithm

Given $\hat{\Delta} > 0$, $\Delta_0 \in (0, \hat{\Delta})$, and $\eta \in [0, \frac{1}{4})$

for $k = 0, 1, 2, ...$

(1) Approximate $p_k$ by solving:

$$\min_{p \in \mathbb{R}^n} m_k(p) = f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T \nabla^2 f(x_k) p, \quad ||p|| \leq \Delta_k \quad (4.29)$$

(2) Evaluate:

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)} \quad (4.30)$$

(3) Determine how to change trust region radius for the next iteration:

if $\rho_k < \frac{1}{4}$

$$\Delta_{k+1} = \frac{1}{4} \Delta_k$$

else

42
if $\rho_k > \frac{3}{4}$ and $||p_k|| = \Delta_k$

$$\Delta_{k+1} = \min(2\Delta_k, \hat{\Delta})$$

else

$$\Delta_{k+1} = \Delta_k$$

(4) Determine the next iterate:

if $\rho_k > \eta$

$$x_{k+1} = x_k + p_k$$

else

$$x_{k+1} = x_k.$$  

(End of algorithm)

Letting $g_k = \nabla f(x_k)$ and using $B_k$ as an approximation to $\nabla^2 f(x_k)$, we can rewrite (4.29) as

$$m_k(p) = f_k + g_k^T p + \frac{1}{2} p^T B_k p.$$  \hspace{1cm} (4.31)

The following theorems from [12] will be useful in proving the convergence of the Levenberg-Marquardt algorithm in later section.

**Theorem 4.1.** Let $m$ be the quadratic function defined by

$$m(p) = g^T p + \frac{1}{2} p^T B p,$$  \hspace{1cm} (4.32)

where $B$ is any symmetric matrix. Then

(1) a minimizer of $m$ exists if and only if $B$ is positive semidefinite and $g$ is in the range of $B$. If $B$ is positive semidefinite, then every $p$ satisfying $Bp = -g$ is a global minimizer of $m$.  

43
(2) \( m \) has a unique minimizer if and only if \( B \) is positive definite.

Proof. For statement (1), assuming \( B \) is positive semidefinite and \( g \) is in the range of \( B \), we want to show there exists some \( p^* \) that minimizes \( m(p) \).

Since \( g \) is in the range of \( B \), there exists some \( p^* \) such that \( Bp^* = -g \). For any \( w \in \mathbb{R}^n \):

\[
m(p^* + w) = g^T(p^* + w) + \frac{1}{2}(p^* + w)^TB(p^* + w) \\
= g^Tp^* + g^Tw + \frac{1}{2}(p^* + w)^T(Bp^* + Bw) \\
= g^Tp^* + g^Tw + \frac{1}{2}(p^*)^TBp^* + \frac{1}{2}(p^*)^TBw + \frac{1}{2}w^TBp^* + \frac{1}{2}w^TBw
\]

(4.33)

Since \( B \) is symmetric, \( B^T = B \), which implies \( (p^*)^TBw = (Bp^*)^Tw \), and

\[
w^TBp^* = w^T(Bp^*) = (Bp^*)^Tw = (p^*)^TBw
\]

(4.34)

Hence, (4.33) becomes:

\[
m(p^* + w) = (g^Tp^* + \frac{1}{2}(p^*)^TBp^*) + g^Tw + (Bp^*)^Tw + \frac{1}{2}w^TBw \\
= m(p^*) + \frac{1}{2}w^TBw
\]

(4.35)

\[
\geq m(p^*).
\]

The last inequality is due to the fact that \( B \) is positive semidefinite and thus \( w^TBw \geq 0 \). Hence, \( p^* \) is a minimizer of \( m(p) \).

Now assume \( p^* \) is a minimizer of \( m \). It follows that \( \nabla m(p^*) = 0 \) and \( \nabla^2 m(p^*) \) is positive semidefinite. From (4.32), note that \( \nabla m(p^*) = Bp^* + g = 0 \), which implies that \( g \) is in the range of \( B \). Moreover, \( \nabla^2 m(p^*) = B \), so \( B \) is positive semidefinite.

For statement (2), assume that \( B \) is positive definite. Also assume \( p \) and \( q \) are both minimizers of \( m \). We want to show that \( p = q \). Using statement (1), since
\( p \) and \( q \) are minimizers,

\[
Bp = Bq = -g. \tag{4.36}
\]

Since \( B \) is positive definite, \( B \) is invertible. So this leads to \( B^{-1}Bp = B^{-1}Bq \) and therefore \( p = q \). Therefore, \( m \) has a unique minimizer.

Now assume \( m \) has a unique minimizer, call it \( p^* \). We want to show that \( B \) is positive definite. Suppose \( B \) is not positive definite. Then there exists some \( w \neq 0 \) such that \( w^T B w = 0 \). From (4.35), \( m(p^* + w) = m(p^*) \), indicating that both \( p^* \) and \( p^* + w \) are minimizers of \( m \), which is a contradiction. Therefore \( B \) must be positive definite.

The following theorem [12] gives the conditions to the solution of trust region problem.

**Theorem 4.2.** The vector \( p^* \) is a global solution to the trust region problem

\[
\min_{p \in \mathbb{R}^n} m(p) = f + g^T p + \frac{1}{2} p^T B p, \quad \|p\| \leq \Delta \tag{4.37}
\]

if and only if \( p^* \) is feasible and there exists some \( \lambda \geq 0 \) such that the following conditions are satisfied:

1. \( (B + \lambda I) p^* = -g \)
2. \( \lambda(\Delta - \|p^*\|) = 0 \)
3. \( (B + \lambda I) \) is positive semidefinite.

**Proof.** \((\Leftarrow)\) Assume there exists \( \lambda \geq 0 \) satisfying the three conditions above. We want to show that \( p^* \) is a global minimizer of \( m(p) \). By Theorem 4.1, \( p^* \) is the global
minimizer of the quadratic function:

\[
\hat{m}(p) = g^T p + \frac{1}{2} p^T (B + \lambda I) p = m(p) + \frac{\lambda}{2} p^T p
\]  

(4.38)

Since \( \hat{m}(p) \geq \hat{m}(p^*) \) for any \( p \),

\[
m(p) \geq m(p^*) + \frac{\lambda}{2} \left[ (p^*)^T p^* - p^T p \right]
\]  

(4.39)

From condition (2), \( \lambda(\Delta - \|p^*\|) = 0 \) implies

\[
\lambda(\Delta - \|p^*\|)(\Delta + \|p^*\|) = \lambda(\Delta^2 - \|p^*\|^2) = \lambda(\Delta^2 - (p^*)^T p^*) = 0.
\]  

(4.40)

Thus,

\[
m(p) \geq m(p^*) + \frac{\lambda}{2} \left[ (p^*)^T p^* + p^T p \right]
\]

\[
= m(p^*) + \frac{\lambda}{2} \left[ (p^*)^T p^* - \Delta^2 + \Delta^2 - p^T p \right]
\]

\[
= m(p^*) + \frac{\lambda}{2} (\Delta^2 - p^T p)
\]

Since \( \lambda \geq 0 \), \( m(p) \geq m(p^*) \) for all \( p \) satisfying \( \|p\| \leq \Delta \). Therefore, \( p^* \) is a global minimizer.

(\( \Rightarrow \)) Assume \( p^* \) is a global solution to \( m(p) \). We want to show there exists \( \lambda \geq 0 \) satisfying the three conditions.

Case 1: \( \|p^*\| < \Delta \), that is, \( p^* \) is an unconstrained minimizer of \( m \).

Note that \( \nabla m(p^*) = Bp^* + g = 0 \). It follows that \( \lambda = 0 \) satisfies condition (1). Also \( \nabla^2 m(p^*) = B \), where \( B \) is positive semidefinite. The choice \( \lambda = 0 \) satisfies condition (3). Condition (2) is automatically satisfied when \( \lambda = 0 \).
Case 2: \( \|p^*\| = \Delta \).

Note that condition (2) is immediately satisfied and the minimizer is within the trust region radius. Moreover, \( p^* \) also solves the constraint problem (4.37). Define the Lagrangian function:

\[
L(p, \lambda) = m(p) + \frac{\lambda}{2}(p^T p - \Delta^2).
\]

By the optimality conditions for constrained optimization, there exists some \( \lambda \) for which \( p^* \) is a stationary point. Setting the partial derivative \( \nabla_p L \) of \( L \) with respect to \( p \) to 0, we obtain

\[
\nabla_p L(p, \lambda) = g + Bp + \lambda p = 0,
\]

and it follows that

\[
g + Bp^* + \lambda p^* = 0 \implies (B + \lambda I)p^* = -g.
\]

So condition (1) is satisfied.

Since \( p^* \) is the minimizer of \( m(p) \), \( m(p) \geq m(p^*) \) for any \( p \) with \( p^T p = (p^*)^T p^* = \Delta^2 \) and \( p \neq p^* \). We can write

\[
m(p) \geq m(p^*) + \frac{\lambda}{2}((p^*)^T p^* - p^T p).
\]

From (4.37),

\[
m(p) - m(p^*) = (f + g^T p + \frac{1}{2}p^T Bp) - (f + g^T p^* + \frac{1}{2}(p^*)^T Bp^*)
\]

and from (4.42),

\[
g^T = -(p^*)^T (B + \lambda I)^T = -(p^*)^T (B + \lambda I),
\]
where \((B + \lambda I) = (B + \lambda I)^T\) because it is symmetric. Thus, combining (4.43) and (4.44),

\[
m(p) - m(p^*)
\]

\[
= -(p^*)^T(B + \lambda I)p + \frac{1}{2}p^T(B + \lambda I)p + (p^*)^T(B + \lambda I)(p^*) - \frac{1}{2}(p^*)^T(B + \lambda I)(p^*)
\]

\[
= -(p^*)^T B p - (p^*)^T \lambda p + \frac{1}{2}p^T B p + \frac{1}{2}p^T \lambda p + (p^*)^T B(p^*) + (p^*)^T \lambda (p^*)
\]

\[
- \frac{1}{2}(p^*)^T B(p^*) - \frac{1}{2}(p^*)^T \lambda (p^*)
\]

Collect terms of \(B\) and \(\lambda\):

\[
= \frac{1}{2}p^T B p - (p^*)^T B p + \frac{1}{2}(p^*)^T B p^* + \frac{1}{2}p^T \lambda p - (p^*)^T \lambda p + \frac{1}{2}(p^*)^T \lambda p^*
\]

\[
= \frac{1}{2}p^T (B + \lambda I)p - (p^*)^T (B + \lambda I)p + \frac{1}{2}(p^*)^T (B + \lambda I)p^*
\]

\[
= \frac{1}{2}(p - p^*)^T (B + \lambda I)p - \frac{1}{2}(p^*)^T (B + \lambda I)(p^* - p)
\]

\[
= \frac{1}{2}(p - p^*)^T (B + \lambda I)p + \frac{1}{2}(p^* - p)^T (B + \lambda I)(p^*)
\]

\[
= \frac{1}{2}(p - p^*)^T (B + \lambda I)p - \frac{1}{2}(p - p^*)^T (B + \lambda I)(p^*)
\]

So,

\[
\frac{1}{2}(p - p^*)^T (B + \lambda I)(p - p^*) \geq 0
\]

which implies \((B + \lambda I)\) is positive semidefinite.

All three conditions are satisfied when \(p^*\) is a global minimizer. Now we need to show that \(\lambda \geq 0\). We will show this by proof of contradiction. Suppose to the contrary that \(\lambda < 0\) and satisfies conditions (1) and (2). Since \(p^*\) minimizes \(m\), by
Theorem 4.1, $B$ is positive semidefinite and $Bp^* = -g$. This implies $\lambda = 0$ in our theorem. This contradicts our supposition. Hence, $\lambda \geq 0$. 

4.5 The Levenberg-Marquardt Algorithm

4.5.1 Motivation behind Levenberg-Marquardt Algorithm

Before delving into the full details of the Levenberg-Marquardt (LM) algorithm, reviewing the motivation behind the algorithm will add clarity to how the algorithm works. The Gauss-Newton method, just like Newton’s method, has rapid convergence, but is sensitive to the initial position. On the other hand, the gradient descent method is not sensitive to initial position even though convergence may be slow. Levenberg combines the advantages of gradient descent and Gauss-Newton by taking $B_k$ in equation (4.17) as:

$$B_k = \nabla^2 f_k + \lambda I \quad (4.46)$$

where $\lambda$ is a damping factor that is adjusted at each iteration.

As in the Gauss-Newton method, the approximation $J_k^T J_k$ is used instead of the actual Hessian $\nabla^2 f_k$, that is,

$$B_k = J_k^T J_k + \lambda I \quad (4.47)$$

and

$$x_{k+1} = x_k - (J_k^T J_k + \lambda I)^{-1} J_k^T r_k \quad (4.48)$$
Recall that the Hessian of $f$ is
\[
\nabla^2 f = \begin{bmatrix}
\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\
\frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}
\end{bmatrix}
\] (4.49)

Along with the equation (4.48), Levenburg [10] defined the following rule to determine the damping factor $\lambda$ at each iteration:

1. Perform one iteration.
2. Evaluate error at the given iterate.
3. If error increases, increase $\lambda$. If error decreases, decrease $\lambda$.

A more precise algorithm for calculating $\lambda$ in the LM algorithm can be given in trust-region framework and is often called the trust-region subproblem [12]:

4.5.2 Trust-Region Subproblem Algorithm

Given $\lambda_1$ and $k$-th time step of the LM algorithm.

for $n = 1, 2, 3, \ldots$

1. Conduct a Cholesky factorization:
\[
J_{k+1}^T J_{k+1} + \lambda_{k,n} I = L_n L_n^T,
\] (4.50)

where $L_n$ is an $n \times n$ lower triangular matrix.

2. Solve $p_{n}^{(\lambda)}$ and $q_{n}^{(\lambda)}$ in the following equations in sequence:
\[
L_n L_n^T p_{n}^{(\lambda)} = -J_{k+1}^T r_{k+1}
\] (4.51)
\[ L_n q_n^{(\lambda)} = p_n^{(\lambda)} \]  

(4.52)

(3) Solve the equation:

\[
\lambda_{n+1} = \lambda_n + \left( \frac{\|p_n^{(\lambda)}\|}{\|q_n^{(\lambda)}\|} \right)^2 \left( \frac{\|p_n^{(\lambda)}\| - \Delta_k}{\Delta_k} \right)
\]  

(4.53)

end

Given \( \lambda_1 = 1 \) as an initial guess. For \( k > 1 \), we calculate \( \lambda \) using the trust-region subproblem algorithm (Algorithm 4.2). For practical purposes, the algorithm will not be implemented until convergence is obtained because it is computationally expensive. Most will define a finite number of iterations \( n \) for the algorithm, or define a tolerance for \( |\lambda_{n+1} - \lambda_n| \) and stop the algorithm.

Marquardt [11] noticed that if \( \lambda \) becomes too large, the term \( J_k^T J_k \) becomes negligible and the algorithm (4.48) behaves similarly to the gradient descent algorithm. The gradient drop towards the minimum becomes very small for a given path \( p_k \). We want movement along smaller gradients to be larger, and vice versa. Marquardt eliminates this issue by replacing the identity matrix with the diagonal of \( J_k^T J_k \) as follows

\[
x_{k+1} = x_k - \left[ J_k^T J_k + \lambda \, \text{diag}(J_k^T J_k) \right]^{-1} J_k^T r_k.
\]  

(4.54)

The above equation is the Levenberg-Marquardt algorithm.
4.5.3 Implementation of Levenberg-Marquardt Algorithm

Using the trust region framework, the goal of the LM algorithm is to solve the following minimization problem:

$$\min_{p} \frac{1}{2} \| J_k p + r_k \|^2, \text{ subject to } \| p \| \leq \Delta_k,$$

where $\Delta_k > 0$ is the trust-region radius. We define the model function $m$ to be:

$$m_k(p) = \frac{1}{2} \| r_k \|^2 + p^T J_k^T r_k + \frac{1}{2} p^T J_k^T J_k p.$$

If the Gauss-Newton direction $p^{GN}$ obtained from solving $J_k^T J_k p^{GN} = -J_k^T r_k$ satisfies the constraint $\| p^{GN} \| < \Delta$, then $p^{GN}$ also solves the trust-region subproblem. If this is not the case, then there exists $\lambda > 0$ for which $p^{LM}_k$ solves

$$(J_k^T J_k + \lambda I)p^{LM}_k = -J_k^T r_k = -\nabla f_k,$$

and $\| p^{LM} \| = \Delta$.

The following lemma [12] gives the conditions for the solution of minimization problem (4.55).

**Lemma 4.3.** The vector $p^{LM}$ is the solution to the minimization problem (4.55) if and only if $p^{LM}$ is feasible and there exists $\lambda \geq 0$ such that

$$(J_k^T J_k + \lambda I)p^{LM} = -J_k^T r_k$$

and

$$\lambda (\Delta - \| p^{LM} \|) = 0.$$

**Proof.** Condition (3) in Theorem 4.2 is satisfied automatically since $J_k^T J_k$ is positive semidefinite and $\lambda \geq 0$. Equations (4.58) and (4.59) follow from condition (1) and condition (2) of Theorem 4.2. □
4.5.4 The Levenberg-Marquardt Algorithm

Given $\hat{\Delta} > 0$, $\Delta_1 \in (0, \hat{\Delta})$, and $\eta \in [0, \frac{1}{4})$

for $k = 1, 2, ...$

(1) If $k = 1$, calculate $p_k^{GN}$:

$$p_k^{GN} = -(J_k^T J_k)^{-1} J_k^T r_k$$ (4.60)

if $p_k^{GN} < \Delta_1$

Use the Gauss–Newton method to obtain convergence

else

Initiate the LM algorithm.

(2) Calculate $\lambda_k$ using the trust-region subproblem (Algorithm 4.5.2).

(3) Approximate $p_k$ by:

$$p_k^{LM} = -(J_k^T J_k + \lambda I)^{-1} J_k^T r_k$$ (4.61)

(4) Evaluate $\rho_k$ using equation (4.56) for $m_k(x)$:

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)}$$ (4.62)

(5) Determine how to change trust region radius for the next iteration:

if $\rho_k < \frac{1}{4}$

$$\Delta_{k+1} = \frac{1}{4} \Delta_k$$

else

if $\rho_k > \frac{3}{4}$ and $\|p_k\| = \Delta_k$
\[ \Delta_{k+1} = \min(2\Delta_k, \hat{\Delta}) \]

else

\[ \Delta_{k+1} = \Delta_k \]

(6) Determine if after the step direction \( p_k \), \( \rho_k \) is small enough to reach an acceptable tolerance \( \eta \).

if \( \rho_k > \eta \)

\[ x_{k+1} = x_k + p_k \]

else

\[ x_{k+1} = x_k. \]

4.5.5 Convergence of The Levenberg-Marquardt Algorithm

Before proving the convergence of the LM algorithm, we have to prove the convergence of the trust region algorithm.

**Theorem 4.4.** Let \( \eta \in (0, \frac{1}{4}) \) in the trust region algorithm (Algorithm 4.4.1). Suppose that \( \|B_k\| \leq \beta \) for some constant \( \beta \). Let \( g \) be bounded below on the set level set \( S \) defined by:

\[ S(R_0) = \{ x \mid \|x - y\| < R_0, \text{ for some } y \in S \}, \]  

(4.63)

where \( R_0 > 0 \). Let \( g \) be a Lipschitz continuous function in \( S(R_0) \) with Lipschitz constant \( \beta_1 \), that is \( g \in LC_{\beta_1}(S(R_0)) \). Suppose all approximate solution \( p_k \) in trust-region algorithm satisfies

\[ m_k(0) - m_k(p) \geq c_1\|g_k\| \min(\Delta_k, \frac{\|g_k\|}{B_k}) \]  

(4.64)

and \( \|p_k\| \leq \gamma \Delta_k \) for some constant \( \gamma \geq 0, c_1 > 0 \). Then \( \{g_k\} \to 0 \).
Proof. We consider a particular positive index \( m \) such that \( g(x_m) \neq 0 \). Since \( g \in LC_{\beta_1}(S) \), we have:

\[
\| g(x) - g(x_m) \| \leq \beta_1 \| x - x_m \|, \quad \forall x, x_m \in S(R_0). \tag{4.65}
\]

We define scalars \( \epsilon = \frac{1}{2} \| g_m \| \) and \( R = \min \left( \frac{\epsilon}{\beta_1}, R_0 \right) \). Notice the \( R \)-ball around \( x_m \)

\[
B(x_m, R) = \{ x \mid \| x - x_m \| \leq R \} \tag{4.66}
\]

is contained in \( S(R_0) \), so Lipschitz continuity of \( g \) holds inside \( B(x_m, R) \), that is,

\[
\| g(x) - g(y) \| \leq \beta_1 \| x - y \|, \quad \forall x, y \in B(x_m, R).
\]

In particular,

\[
\| g(x) - g(x_m) \| \leq \beta_1 \| x - x_m \|
\]

\[
\leq \beta_1 R \leq \beta_1 (\epsilon/\beta_1) = \epsilon = \frac{1}{2} \| g(x_m) \|.
\]

From the triangle inequality

\[
\| g(x_m) \| - \| g(x) \| \leq \| g(x_m) - g(x) \| \leq \frac{1}{2} \| g(x_m) \| \tag{4.67}
\]

which implies

\[
\| g(x) \| \geq \frac{1}{2} \| g(x_m) \| = \epsilon. \tag{4.68}
\]

Let \( \{ x_k \} \) be a sequence generated by trust-region algorithm. If \( \{ x_k \}_{k \geq m} \subset B(x_m, R) \), then \( \| g(x_k) \| \geq \epsilon \) for all \( k \geq m \). Hence, \( \{ g(x_k) \} \not\to 0 \). Therefore, there must exist some index \( l \geq m \) such that \( \{ x_{l+1}, x_{l+2}, \ldots \} \) lie outside the ball \( B(x_m, R) \), that is, \( x_{l+1} \) is the first iterate that escapes \( B(x_m, R) \). Note that \( \| g(x_k) \| \geq \epsilon \) for
\[ k = m, m + 1, \ldots, l. \] Thus,

\[ f(x_m) - f(x_{l+1}) = f(x_m) - f(x_{m+1}) + f(x_{m+1}) - \ldots - f(x_{l+1}) \quad (4.69) \]

\[ = \sum_{k=m}^l f(x_k) - f(x_{k+1}). \quad (4.70) \]

If \( x = x_{k+1} \), then \( f(x_k) - f(x_{k+1}) = 0 \). If \( x \neq x_{k+1} \), then \( x_{k+1} = x_k + p_k \) for some \( p_k \neq 0 \) and this happens when \( \rho_k > \eta \), that is,

\[ \rho_k = \frac{f(x_k) - f(x_{k+1})}{m_k(0) - m_k(p_k)} > \eta \]

\[ \Rightarrow f(x_k) - f(x_{k+1}) > \eta(m_k(0) - m_k(p_k)) \]

From (4.70), we have

\[ f(x_k) - f(x_{l+1}) \geq \sum_{k=m, x_k \neq x_{k+1}}^l \eta(m_k(0) - m_k(p_k)) \]

\[ \geq \sum_{k=m, x_k \neq x_{k+1}}^l \eta c_1 \|g_k\| \min \left( \Delta_k, \frac{\|g_k\|}{B_k} \right) \quad \text{(by assumption)} \]

\[ \geq \sum_{k=m, x_k \neq x_{k+1}}^l \eta c_1 \epsilon \min \left( \Delta_k, \frac{\epsilon}{\beta} \right). \]

The last inequality comes from the fact that \( \|g_k\| \geq \epsilon \) for all \( k \geq m \) and \( \|B_k\| \leq \beta \).

We consider two cases:

Case 1: If \( \Delta_k > \epsilon/\beta \), then

\[ f(x_m) - f(x_{l+1}) \geq \eta c_1 \epsilon \frac{\epsilon}{\beta}, \quad (4.71) \]

Case 2: If \( \Delta_k \leq \epsilon/\beta \) for \( k = m, m + 1, \ldots, l \), then

\[ f(x_m) - f(x_{l+1}) \geq \eta c_1 \epsilon \sum_{k=m, x_k \neq x_{k+1}}^l \Delta_k \quad (4.72) \]

\[ \geq \eta c_1 \epsilon R \quad (4.73) \]

\[ = \eta c_1 \epsilon \min \left( \frac{\epsilon}{\beta}, R_0 \right). \quad (4.74) \]
Since \( \{f(x_k)\} \) is decreasing and bounded below, \( \{f(x_k)\} \to f(x^*) \) and \( f(x^*) > -\infty \).

Hence, combining both cases we obtain

\[
f(x_m) - f(x^*) \geq f(x_m) - f(x_{t+1}) \quad (\text{since } f(x^*) \leq f(x_{t+1}))
\]

\[
\geq \eta c_1 \epsilon \min \left( \frac{\epsilon}{\beta^1}, R_0 \right)
\]

\[
= \frac{1}{2} \eta c_1 \|g(x_m)\| \min \left( \frac{\|g(x_m)\|}{2\beta}, \frac{\|g(x_m)\|}{2\beta^1}, R_0 \right).
\]

But as \( m \to \infty \), \( f(x_m) - f(x^*) \to 0 \), and this forces \( \|g(x_m)\| \to 0 \) as well. \( \square \)

Now we use this theorem to show that the Levenberg-Marquardt algorithm converges [12].

**Theorem 4.5.** Let \( \eta \in (0, \frac{1}{4}) \) in the trust region algorithm. Suppose the set level \( L \) as defined by (4.63) is bounded and the residual functions \( r_j \), where \( j = 1, \ldots, m \) are Lipschitz continuous and differentiable in neighborhood \( N \) of \( L \). Assume that for each \( k \), the approximate solution for \( p_k \) in 4.55 satisfies:

\[
m_k(0) - m_k(p_k) \geq c_1 \|J_k^T r_k\| \min \left( \Delta_k, \frac{\|J_k^T r_k\|}{\|J_k^T J_k\|} \right)
\]

(4.75)

for some constant \( c_1 > 0 \). In addition, \( \|p_k\| \leq \gamma \Delta_k \) for some constant \( \gamma \geq 1 \). Then

\[
\lim_{k \to \infty} J_k^T r_k = 0
\]

(4.76)

**Proof.** From the smoothness of \( r_j \), i.e. \( r_j \) is infinitely differentiable. We can choose \( M > 0 \) such that \( \|J_k^T J_k\| \leq M \) for all \( k \). \( f \) is bounded is bounded below by zero. Thus, Theorem 4.4 is satisfied. \( \square \)

4.5.6 Computational Example

This example will illustrate the Levenberg-Marquardt (LM) algorithm 4.5.4.

The following table shows the annual full-time student enrollment data from Californ-
nia State University, Los Angeles from 2005-2015 [21].

Table 4.1: California State University, Los Angeles full-time student enrollment data from 2005-2015

<table>
<thead>
<tr>
<th>Year</th>
<th>Full-Time Student Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>15936</td>
</tr>
<tr>
<td>2006</td>
<td>16251</td>
</tr>
<tr>
<td>2007</td>
<td>16687</td>
</tr>
<tr>
<td>2008</td>
<td>16297</td>
</tr>
<tr>
<td>2009</td>
<td>15967</td>
</tr>
<tr>
<td>2010</td>
<td>16151</td>
</tr>
<tr>
<td>2011</td>
<td>17262</td>
</tr>
<tr>
<td>2012</td>
<td>17952</td>
</tr>
<tr>
<td>2013</td>
<td>18796</td>
</tr>
<tr>
<td>2014</td>
<td>20445</td>
</tr>
<tr>
<td>2015</td>
<td>23252</td>
</tr>
</tbody>
</table>

We fit the following nonlinear model function

\[ p(t, x) = x_2 \ln(x_1 t) + x_3 \]  

(4.77)

using the LM algorithm 4.5.4.

The parameter vector changes after each \( k \)-th iterate:

\[ x_k = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{bmatrix} \]  

(4.78)

Our initial guess for \( x_1 \) after a rough estimate will be:

\[ x_1 = \begin{bmatrix} 100 \\ 50 \\ 100 \end{bmatrix} \]  

(4.79)

The first step of the LM algorithm is to use the Gauss–Newton method.
\[
\begin{pmatrix}
50 \ln(100) + 100 - 15936 \\
50 \ln(200) + 100 - 16251 \\
50 \ln(300) + 100 - 16687 \\
50 \ln(400) + 100 - 16297 \\
50 \ln(500) + 100 - 15967 \\
50 \ln(600) + 100 - 16151 \\
50 \ln(700) + 100 - 16834 \\
50 \ln(800) + 100 - 17952 \\
50 \ln(900) + 100 - 18356 \\
50 \ln(1000) + 100 - 20445 \\
50 \ln(1100) + 100 - 23252
\end{pmatrix}
= \begin{pmatrix}
15606 \\
15886 \\
16302 \\
15897 \\
15556 \\
15731 \\
16834 \\
17518 \\
18356 \\
20000 \\
22802
\end{pmatrix}
\]

(4.80)

\[||r(x_1)||^2 = 3.3510 \times 10^9, \text{ so } f(x_1) = 1.6755 \times 10^{13}\]

Recall that the residual is defined as \( r_j = |y_j - p(t_j, x)| \). Since the absolute function is not smooth, to ensure positivity by re-writing the residual \( r_j \) as a square function:

\[ r_j^2 = (y_j - x_2 \ln(x_1 t_j) - x_3)^2 \]  

(4.81)

The Jacobian is calculated:
Combining equation (4.13) and equation (4.28) from the Gauss–Newton method (GN), we get:

\[ p_{k}^{GN} = - (J_{k}^{T} J_{k})^{-1} J_{k}^{T} r_{k} \]  

\[ (4.83) \]

Substituting our calculated values we get \( p_{1}^{GN} = (-36.9018, -2.6100, 0.4891) \)

Once we go through one step of the GN algorithm, we compare \( p_{1}^{GN} \) to \( \Delta_{1} \). The trust regions acts as an indicator to see if we are within an acceptable range of the minimum of the minimization function \( f(x) \) from equation (4.10). For illustrative purposes, let \( \Delta_{1} = 0.1 \). In this case, ||\( p_{1}^{GN} \)|| = 36.9972 > 0.1. Because of this, we switch to the LM algorithm.

We can now initialize the LM algorithm. Going back to our initial guess \( x_{1} \),
$||r(x_1)||^2 = 3.3510 \times 10^9$, so $f(x_1) = 1.6755 \times 10^{13}$, same as the initialization step of GN.

Let $\lambda_1 = 1$ as an initial guess. For the purposes of this illustration, we will use this algorithm only once.

So using $\lambda_1 = 1$ and equation (4.61), $p_1^{LM} = (0.0050, -5.5109 \times 10^{-10}, 0.5000)$.

Following the trust region algorithm (4.4.1), we now calculate $\rho_k$ (4.84).

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)} \tag{4.84}$$

(1) $f(x_1) = 1.6755 \times 10^{13}$ \tag{4.85}

(2) $f(x_1 + p_1) = f(x_2) = \frac{1}{2} ||r(x_2)||^2 = 1.6754 \times 10^9$ \tag{4.86}

(3) $m_1(0) = \frac{1}{2} ||r(x_1)||^2 = f(x_1) = 1.6755 \times 10^{13}$ \tag{4.87}

(4) $m_k(p_k) = 8.7897 \times 10^{25}$ \tag{4.88}

Combining terms, we end up with:

$$\rho_1 = \frac{f(x_1) - f(x_1 + p_1)}{m_1(0) - m_1(p_1)} = 5.2461 \times 10^{16} \tag{4.89}$$
For the purpose of illustration, let $\Delta_1 = 0.1$ and $\eta = 0.001$. From the trust algorithm (4.4.1), we keep the same trust region value, so $\Delta_2 = \Delta_1$. Since $\rho_1 > \eta$, $x_2 = x_1 + p_1$

We can now update our parameter values:

$$x_2 = x_1 + p_1^{LM} = \begin{bmatrix} 100 + .0050 \\ 50 + (-5.5109 \times 10^{-10}) \\ 100 + .5000 \end{bmatrix} = \begin{bmatrix} 100.0050 \\ 50.0000 \\ 100.5000 \end{bmatrix}$$ (4.90)

For $k = 2$, we need to calculate $\lambda_2$ first with the trust region subproblem (4.4.1). When $k > 1$, $\lambda$ in equation (4.61) is calculated using the trust region subproblem algorithm (4.5.2):

$$J_2^T J_2 + \lambda_1 I = \begin{bmatrix} 1340177.876 & 847098339.3 & 134024387.8 \\ 847098339.3 & 5.41941 \times 10^{11} & 84714069006 \\ 134024387.8 & 84714069006 & 13403108838 \end{bmatrix}$$ (4.91)

We take the Cholesky Decomposition to get:

$$L_1 L_1^T = \begin{bmatrix} 1157.6605 & 0 & 0 \\ 73173.9440 & 80673.0616 & 0 \\ 115771.7532 & 0.7835 & 100.0069 \end{bmatrix} \begin{bmatrix} 1157.6605 & 73173.9440 & 115771.7532 \\ 0 & 80673.0616 & 0.7835 \\ 0 & 0 & 100.0069 \end{bmatrix}$$ (4.92)

Solving $p_1^{(\lambda)}$ from equation (4.51):

$$p_1^{(\lambda)} = \begin{bmatrix} 3350.1457 \\ -5.5320 \times 10^{-5} \\ -32.9994 \end{bmatrix}$$ (4.93)

Solving $q_1^{(\lambda)}$ from equation (4.52):

$$q_1^{(\lambda)} = \begin{bmatrix} 2.8939 \\ -26.2486 \\ -3350.2041 \end{bmatrix}$$ (4.94)

Using the equation (4.53) we get:

$$\lambda_2 = \lambda_1 + \left( \frac{||p_1^{(\lambda)}||}{||q_1^{(\lambda)}||} \right)^2 \left( \frac{||p_1^{(\lambda)}|| - \Delta}{\Delta} \right)$$

$$= 1 + \left( \frac{3.3503 \times 10^4}{3.3503 \times 10^4} \right)^2 \left( \frac{3.3503 \times 10^4 - 0.1}{0.1} \right) = 3.3503 \times 10^4$$ (4.95)
Using to calculate (4.61) to calculate $\mathbf{p}_2^{LM}$, we end up with:

$$
\mathbf{p}_2 = \begin{bmatrix}
0.0050 \\
1.6264 \times 10^{-5} \\
0.5000
\end{bmatrix}
$$

(4.96)

This implies:

$$
\mathbf{x}_3 = \mathbf{x}_2 + \mathbf{p}_2^{LM} = \begin{bmatrix}
100.0100 \\
50 \\
100.9998
\end{bmatrix}
$$

(4.97)

The following graph provides an illustration of the LM algorithm after a successive number of iterations:

Figure 4.1: LM Algorithm fitting on Annual Cal State LA Full-Time Student Enrollment Data from 2005 - 2015
The LM algorithm ends once \( \rho_k < \eta \).
4.6 Results of Fit

Figure 4.3: LM Algorithm of various sigmoidal curves and their respective mean square error (MSE).

Table 4.2: LM algorithm of various sigmoidal curves and their respective MSE

<table>
<thead>
<tr>
<th>Curve Name</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>4835.38127595731</td>
</tr>
<tr>
<td>Gompertz</td>
<td>5409.55782739912</td>
</tr>
<tr>
<td>Weibull</td>
<td>4548.42018423027</td>
</tr>
<tr>
<td>Generalized</td>
<td>4060.9265664517</td>
</tr>
<tr>
<td>Chapman-Richards</td>
<td>4005.64641784122</td>
</tr>
</tbody>
</table>
Figure 4.4: Polynomial algorithms of various degrees and their respective mean square error (MSE).
Table 4.3: Polynomial algorithms of various degrees and their respective mean square average (MSE)

<table>
<thead>
<tr>
<th>Polynomial Degree</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7362.6697517347902</td>
</tr>
<tr>
<td>2</td>
<td>6168.5780648502696</td>
</tr>
<tr>
<td>3</td>
<td>4615.8348964957704</td>
</tr>
<tr>
<td>4</td>
<td>3407.5441301470801</td>
</tr>
<tr>
<td>5</td>
<td>3107.53868716131</td>
</tr>
<tr>
<td>6</td>
<td>2235.1476172573798</td>
</tr>
<tr>
<td>7</td>
<td>2070.1495434897602</td>
</tr>
<tr>
<td>8</td>
<td>1433.1560713026099</td>
</tr>
<tr>
<td>9</td>
<td>1257.1509207751301</td>
</tr>
<tr>
<td>10</td>
<td>1191.5658148058201</td>
</tr>
<tr>
<td>11</td>
<td>1179.1434984611301</td>
</tr>
<tr>
<td>12</td>
<td>1178.4457355050699</td>
</tr>
<tr>
<td>13</td>
<td>1006.92989762918</td>
</tr>
<tr>
<td>14</td>
<td>924.50777729245601</td>
</tr>
<tr>
<td>15</td>
<td>868.82744941962801</td>
</tr>
<tr>
<td>16</td>
<td>833.82532793095197</td>
</tr>
<tr>
<td>17</td>
<td>829.35627632649903</td>
</tr>
<tr>
<td>18</td>
<td>823.47416471310203</td>
</tr>
<tr>
<td>19</td>
<td>822.90489838668702</td>
</tr>
<tr>
<td>20</td>
<td>780.12966874691404</td>
</tr>
</tbody>
</table>
This chapter will demonstrate the use of the Levenberg-Marquardt (LM) algorithm to fit data and forecast stock market prices. We filter the data with the Hodrick-Prescott (HP), exponential smoothing, and moving average techniques. Data without a filter applied is our standard of comparison. We will use the Logistic, Gompertz, Weibull, Chapman-Richards, and the Generalized Logistic equations after application of each respective filter.

All data fitted starts at the closing price of the initial public offering (IPO) to variable amounts of days chosen forward in time. The raw data is the daily closing prices of Vanguard Energy Fund Investor Shares (VGENX) [31]. It starts from May 23rd, 1984 to November 11th, 2016. The fund invests in US energy and foreign securities. The composition of the fund as of December 31st, 2016 is shown in this data table:
Table 5.1: Composition of VGENX Mutual Fund

<table>
<thead>
<tr>
<th>Energy Fund Investor</th>
<th>as of 12/31/2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal &amp; Consumable Fuels</td>
<td>0.00%</td>
</tr>
<tr>
<td>Consumer Discretionary</td>
<td>0.10%</td>
</tr>
<tr>
<td>Consumer Staples</td>
<td>0.10%</td>
</tr>
<tr>
<td>Financials</td>
<td>0.20%</td>
</tr>
<tr>
<td>Health Care</td>
<td>0.10%</td>
</tr>
<tr>
<td>Industrials</td>
<td>0.20%</td>
</tr>
<tr>
<td>Information Technology</td>
<td>0.20%</td>
</tr>
<tr>
<td>Integrated Oil &amp; Gas</td>
<td>36.10%</td>
</tr>
<tr>
<td>Oil &amp; Gas Drilling</td>
<td>1.60%</td>
</tr>
<tr>
<td>Oil &amp; Gas Equipment &amp; Services</td>
<td>9.00%</td>
</tr>
<tr>
<td>Oil &amp; Gas Exploration &amp; Production</td>
<td>37.90%</td>
</tr>
<tr>
<td>Oil &amp; Gas Refining &amp; Marketing</td>
<td>7.20%</td>
</tr>
<tr>
<td>Oil &amp; Gas Storage &amp; Transportation</td>
<td>3.70%</td>
</tr>
<tr>
<td>Utilities</td>
<td>3.50%</td>
</tr>
</tbody>
</table>

From this data set, we start with the IPO to a certain number of days we assume to be known data. We call this ”prior data.” The prior data consists of 1000, 2000, 3000, 4000, 5000, 6000, and 7000 data points. From the prior data, we attempt to forecast a set number of days after the last prior data point. We attempt to forecast stock prices 100, 300, 1000, and 3000 trading days into the future. Prior to fitting the data with the LM algorithm, we either leave the prior data unfiltered, apply the Hodrick-Prescott filter, the moving average filter, or the exponential smoothing filter. The moving average filter is arbitrarily 300 trading days, which approximates one year’s worth of trading. The weight factor $\alpha$ for the exponential average was chosen by taking the lowest mean square error value between the prior data and filtered data set in 0.1 intervals between 0 and 1. The forecast difference is defined as the actual data at the forecasted time point minus the fitted data at the forecast time point.
Positive values correspond to forecast underestimates, and negative values correspond to forecast overestimates.

5.2 Results

Since the raw data set is large, only 1000, 5000, and 7000 prior data points are provided with more detailed analysis. Their respective forecast plots, forecast difference bar graph, and MSE bar graph are shown in section 5.4. The reason for these choices is because 1000 prior data points is representative of initial behavior of a sigmoidal curve, 5000 prior points is representative of behavior immediately before inflection behavior, and 7000 prior data points is representative of behavior of a sigmoidal curve inclusive of the inflection point. In other words, these prior data points are representative of emergent, inflection, and saturation phases. The inflection point occurs roughly between 5000 - 6000 days after IPO. Histograms of forecast differences display all prior data sets from 1000 - 7000 prior data points. Data tables of each forecast differences and their mean square error (MSE) are located in appendix D.1.

From section 5.4, the data set shows the MSE and forecast difference magnitude increases as the number of forecast days increases. For 1000 prior data points, the all MSE are less than 100 $^2$, which implies the mean error is within the square root of the MSE, or $10$. But if we look at 1000 forecast days or less, the MSE is generally less than 10$^2$, or error that is roughly $3$.

For 5000 prior data points, the MSE are generally less than 200 $^2$ for 100 and 300 forecast days, and range from 800 - 1800 $^2$ for 1000 to 3000 forecast days, which
implies a mean error of about $10 – $40. Forecasting near or past the inflection point produces higher uncertainty. The forecast differences 300 and 1000 forecast days were generally high, with a range of approximately $10 – $50. The HP filter was able to keep 100 to 1000 forecast days all under $20. But forecasting past the inflection point at 3000 forecast days, the forecast difference drops, but the MSE remains relatively high. This suggests there is much more volatility after the inflection point.

For 7000 prior data points, the forecast values are generally negative, meaning the sigmoidal curves generally overestimates the actual stock value. But once the maximum value for the stock price is known and the approximate location of the inflection point, the MSE are all below 1000 $^2$. The behavior of stock prices near the carrying capacity is less volatile than it is near the inflection point.
When plotting the histogram of forecast differences, positive values correspond to underestimates, and negative values correspond to overestimates. Ideally, we would like to have our difference curves skew to the right and be as close to zero as possible. This translates to accurate forecasts that are slightly underestimated. The standard deviation is generally within $30$ for all filtration techniques and curves. The exponential filter appears to be worse than our control. A filter should provide a greater contrast in moment values so that it is easier to distinguish how each sigmoidal curve member is behaving. By inspection of minimums and maximums of each statistical moment with each type of filtration and curve, the difference between the minimums and maximums of the exponential filter is the smallest. The minimum and maximum differences are smaller than unfiltered data. So the exponential filter performs worse than the control.

A normal distribution itself still does not demonstrate a particular sigmoidal curve is better at forecasting than another. A normal distribution near a mean of zero simply implies that the forecasts have a 50% chance of forecasting above or below the actual price. We want the normal distributions to have a mean past zero and have a positive skew to that the probability of underestimating accurate data is greater than 50%.
Table 5.2: Average of Forecast Differences

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Unfiltered</th>
<th>Hodrick-Prescott</th>
<th>Exponential</th>
<th>Moving Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.966</td>
<td>1.263</td>
<td>1.757</td>
<td>1.508</td>
</tr>
<tr>
<td>Gompertz</td>
<td>1.982</td>
<td>1.604</td>
<td>1.931</td>
<td>2.331</td>
</tr>
<tr>
<td>Weibull</td>
<td>4.930</td>
<td>1.911</td>
<td>3.468</td>
<td>3.607</td>
</tr>
<tr>
<td>Chapman-Richards</td>
<td>9.161</td>
<td>8.542</td>
<td>5.312</td>
<td>6.072</td>
</tr>
<tr>
<td>Generalized</td>
<td>10.406</td>
<td>4.207</td>
<td>5.179</td>
<td>10.163</td>
</tr>
</tbody>
</table>

Note: Red highlighted are the lowest values of a column, and blue values are the highest.

Table 5.3: Standard Deviation of Forecast Differences

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Unfiltered</th>
<th>Hodrick-Prescott</th>
<th>Exponential</th>
<th>Moving Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>22.059</td>
<td>22.832</td>
<td>22.706</td>
<td>22.498</td>
</tr>
<tr>
<td>Gompertz</td>
<td>22.837</td>
<td>22.496</td>
<td>22.860</td>
<td>22.020</td>
</tr>
<tr>
<td>Weibull</td>
<td>22.313</td>
<td>22.051</td>
<td>22.482</td>
<td>22.102</td>
</tr>
</tbody>
</table>

Note: Red highlighted are the lowest values of a column, and blue values are the highest.

Table 5.4: Histogram of Skews of Forecast Differences

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Unfiltered</th>
<th>Hodrick-Prescott</th>
<th>Exponential</th>
<th>Moving Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>0.056</td>
<td>0.035</td>
<td>0.161</td>
<td>0.224</td>
</tr>
<tr>
<td>Gompertz</td>
<td>0.036</td>
<td>0.025</td>
<td>0.032</td>
<td>0.181</td>
</tr>
<tr>
<td>Weibull</td>
<td>0.247</td>
<td>0.255</td>
<td>0.179</td>
<td>0.296</td>
</tr>
<tr>
<td>Chapman-Richards</td>
<td>1.252</td>
<td>1.600</td>
<td>0.732</td>
<td>1.374</td>
</tr>
<tr>
<td>Generalized</td>
<td>1.259</td>
<td>1.279</td>
<td>0.718</td>
<td>1.274</td>
</tr>
</tbody>
</table>

Note: Red highlighted are the lowest values of a column, and blue values are the highest.

The skew is the measure of asymmetry about the mean. Positive skew values mean the data is more spread out to the right of the mean, and negative skew values mean the data is more spread out to the left of the mean. Here is a general schematic of skew [29]:

75
The following is the equation to calculate skew [23]:

\[
s = \frac{E(x - \mu)^3}{\sigma^3} = \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^3}{\left(\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}\right)^3}
\] (5.1)

For the applications of this thesis, positive skew values correspond to higher frequency of positive forecast differences. The Chapman-Richards and Generalized Logistic equations have the largest positive skew values, meaning forecast values tend to underestimate actual values. The Logistic and Gompertz equations have the lowest skew values for all filtration techniques, implying that their distributions are close to symmetry about the mean.

<table>
<thead>
<tr>
<th></th>
<th>Unfiltered</th>
<th>Hodrick-Prescott</th>
<th>Exponential</th>
<th>Moving Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>5.856</td>
<td>5.259</td>
<td>4.961</td>
<td>5.543</td>
</tr>
<tr>
<td>Gompertz</td>
<td>5.262</td>
<td>5.492</td>
<td>5.283</td>
<td>5.850</td>
</tr>
<tr>
<td>Weibull</td>
<td>4.352</td>
<td>5.003</td>
<td>4.874</td>
<td>5.448</td>
</tr>
<tr>
<td>Chapman-Richards</td>
<td>4.947</td>
<td>6.131</td>
<td>5.120</td>
<td>5.636</td>
</tr>
<tr>
<td>Generalized</td>
<td>4.970</td>
<td>5.725</td>
<td>5.091</td>
<td>4.383</td>
</tr>
</tbody>
</table>

Note: Red highlighted are the lowest values of a column, and blue values the highest.

The kurtosis is the measure of "tailedness" of a normal distribution, not the sharpness nor relative height of a normal distribution’s peak [8]. In other words,
kurtosis is the measure outliers normalized to the standard deviation. The kurtosis does not directly imply the shape of a peak for a given distribution.

The following is the equation to calculate kurtosis [22]:

\[
k = \frac{E(x - \mu)^4}{\sigma^4} = \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^4}{\left(\sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}\right)^4}
\] (5.2)

The forecast distributions should ideally show low kurtosis, which would imply the data would have few outliers and be close to the mean. Low kurtosis implies that the data would have forecasts far from the mean. The Logistic equation generally has the lowest kurtosis, which implies fewest outliers relative to the mean. But our data also shows that for the logistic equation, the mean is close to zero, and the standard deviation is very high, so we end up with a low kurtosis. The Gompertz has high kurtosis because the it has many outliers, so the numerator has a high value. The Chapman-Richards with the Hodrick–Prescott filter have high kurtosis because it has the lowest standard deviation. The Generalized Logistic with the moving average has a low kurtosis because is contains few outliers relative to the mean.

Another important feature of the data is the value of the fitted inflection point. The inflection point defines where the growth rate transitions from acceleration to deceleration, meaning the change of growth rate goes from positive to negative. From visual inspection, the inflection point occurs between 5000-6000 points after IPO. When trying to fit data with prior data near the inflection point, the fits forecast an extremely high carrying capacity, generally to the order of $10^3 - 10^4$. This is because the fitting algorithm assumes that the high growth rate will last for a long time. The growth rate is the highest near the inflection point because the growth
rate is approaching an extrema value with respect to the second derivative of the closed form solution. Since we are fitting sigmoidal curves, the curve to the left of the inflection point exhibits positive concavity, and the curve to the right of the inflection point exhibits negative concavity. For all forecast differences, most of the differences were positive before encountering the inflection point and negative after the inflection point. This implies the sigmoidal curve generally underestimates the actual data prior to the inflection point, and overestimates after the inflection point. If the inflection point is known, an investor can take advantage of this behavior and determine whether to buy or sell stock.

The MSE before and after the inflection point increases dramatically. If we focus our attention to 1000 and 3000 forecast days, for 1000 prior data points, all the MSE were less than 40 \$^2$. 3000 forecast days implies 4000 days after IPO, so this time point is not past the inflection point. The increase is apparent for 5000 prior data points. The MSE for 1000 and 3000 forecast days are around 1000\$^2$. This implies stock prices have greater volatility after the inflection point.

Macroeconomic reasons can explain why stock prices have greater volatility. Baumeister and Peersman [1] and Robe and Wallen [17] show fluctuations in physical crude oil inventories are the biggest factor that determine oil prices. Crude oil production reported by the OECD has been fairly stable since VGENX went into IPO [27]:

\[ \text{MSE before inflection point} < 40 \text{ \$}^2 \]
\[ \text{MSE after inflection point} > 1000 \text{ \$}^2 \]
But what has changed is the oil price elasticity. Price elasticity of is percentage change of the quantity demanded/supplied to the percent change in price [3]. Baumeister et al. [1] shows the price elasticity of crude oil decreases from 1985 to 2010. Low price elasticity implies small fluctuations in supply results in large changes in price.

Based on the data and analysis results, the Hodrick-Prescott (HP) filter used in tandem with the Chapman-Richards equation provide the best forecast results for this particular data set. The HP filter provides the largest difference in extreme values for the standard deviation, skew, and kurtosis. This allows for a bigger contrast when looking at the differences in moments between each type of equation. The Chapman-Richards provided consistent high average forecast differences for 2 out of 3 filtering techniques, which demonstrates its tendency to underestimate forecasts.
The Chapman-Richards consistently demonstrated the smallest standard deviation for all filtering techniques, so most differences are least likely to produce extreme values. The Chapman-Richards produces the highest positive skew values, so the forecast differences tend to underestimate actual values. The combination of the HP filter and Chapman-Richards has the highest kurtosis, which implies most of the data is concentrated near the mean.

5.3 Future Research

This paper used a deterministic approach to forecasting stock market data. Another path to explore could explore correlations with other stock market indicators, such as oil rig counts or the VIX indicator, to see how it would influence forecasting with sigmoidal curves as suggested by Baumeister et al [1]. One can also explore how to incorporate sigmoidal curves into agent based modeling [4]. One can also explore applications to mean reversion theory and how to determine the best conditions to sell stock [18].
5.4 Data

5.4.1 Raw Data

Daily Closing Prices of VGENX from IPO to November 11th, 2016
5.4.2 Fit of Various Sigmoidal Curves

1000 Prior Data Points
5000 Prior Data Points

Various Sigmoidal Curves Fitted on 5000 Prior Data Points on Unfiltered Data

Various Sigmoidal Curves Fitted on 5000 Prior Data Points on Hodrick-Prescott Filtered Data
Various Sigmoidal Curves Fitted on 5000 Prior Data Points on Exponential Filtered Data

Price of VGENX in USD

Trading Days Since IPO

Various Sigmoidal Curves Fitted on 5000 Prior Data Points on Moving Average (300) Filtered Data

Price of VGENX in USD

Trading Days Since IPO
7000 Prior Data Points

Various Sigmoidal Curves Fitted on 7000 Prior Data Points on Unfiltered Data

Various Sigmoidal Curves Fitted on 7000 Prior Data Points on Hodrick-Prescott Filtered Data
5.4.3 Forecast Difference with 1000 Prior Known Days

Note: Negative correspond to overestimation in forecasts, while positive numbers correspond to underestimation of forecasts. Units are in U.S. dollars.
5.4.4 Forecast Difference with 5000 Prior Known Days

5000 Prior Data Points: Forecast Difference with No Filter

5000 Prior Data Points: Forecast Difference with Hodrick Prescott Filter
5000 Prior Data Points: Forecast Difference with Exponential Smoothing Filter

Forecast Difference in [$]

Forecast Days

5000 Prior Data Points: Forecast Difference with Moving Average Filter

Forecast Difference in [$]

Forecast Days

Logistic
Gompertz
Weibull
Chapman Richards
Generalized Logistic
5.4.5 Forecast Difference with 7000 Prior Known Days

7000 Prior Data Points: Forecast Difference with No Filter

7000 Prior Data Points: Forecast Differences with HP Filter
5.4.6 MSE with 1000 Prior Known Days

1000 Prior Data Points: Mean Square Error with No Filter

1000 Prior Data Points: Mean Square Error with Hodrick Prescott Filter
5.4.7 MSE with 5000 Prior Known Days

![Chart 1: MSE with No Filter](image1)

![Chart 2: MSE with Hodrick Prescott Filter](image2)
5000 Prior Data Points: Mean Square Error with Exponential Smoothing Filter

MSE in $[\text{ft}^2]$ vs Forecast Days

5000 Prior Data Points: Mean Square Error with Moving Average Filter

MSE in $[\text{ft}^2]$ vs Forecast Days
5.4.8 MSE with 7000 Prior Known Days

7000 Prior Data Points: Mean Square Error with No Filter

7000 Prior Data Points: Mean Square Error with Hodrick-Prescott Filter
REFERENCES


[26] https://en.wikipedia.org/wiki/Mean_squared_error (Visited November 9, 2016)


APPENDIX A

The Logistic Model

We start with definition of the logistic differential equation [13]:

\[
\frac{dY_t}{dt} = \beta Y_t \left(1 - \frac{Y_t}{Y_\infty}\right), \tag{A.1}
\]

where \( \beta \) is the maximal growth rate when \( Y_t \) is much smaller than \( Y_\infty \), and \( Y_\infty \) is the upper limit for the sigmoidal curve, also known as the *carrying capacity*.

We nondimensionalize (A.1) by dividing \( Y_t \) by its max value \( Y_\infty \):

\[
\frac{dY_t/\ dt}{Y_\infty} = \beta \frac{Y_t}{Y_\infty} \left(1 - \frac{Y_t}{Y_\infty}\right)
\]

and then substitute \( x = Y_t/Y_\infty \) to get

\[
\frac{dx}{dt} = \beta x(1 - x). \tag{A.2}
\]

Using the method of separation of variables, we have

\[
\frac{dx}{x(1-x)} = \beta dt. \tag{A.3}
\]

Now we use partial fractions decomposition to write

\[
\frac{1}{x(1-x)} = \frac{A}{x} + \frac{B}{(1-x)}
\]

\[A(1-x) + Bx = 1 \Rightarrow A - Ax + Bx = 1\]

which gives us a system of equations

\[
\begin{cases}
-A + B = 0 \\
A = 1
\end{cases}
\]

and it follows that \( B = 1 \). Integrating the left hand side of (A.3), we obtain

\[
\int \frac{1}{x(1-x)}\ dx = \int \frac{1}{x}\ dx + \int \frac{1}{1-x}\ dx = \ln(x) - \ln(1 - x)
\]

103
and therefore

\[ \ln(x) - \ln(1-x) = \beta t + C \]

\[ \ln \left( \frac{1-x}{x} \right) = -\beta t - C \]

\[ \frac{1}{x} - 1 = \alpha e^{-\beta t}, \text{ where } \alpha = e^{-C}. \]

Solving for \( x \),

\[ x = \frac{1}{1 + \alpha e^{-\beta t}}. \]

When we substitute \( x = \frac{Y_t}{Y_\infty} \), we finally obtain

\[ Y_t = \frac{Y_\infty}{1 + \alpha e^{-\beta t}}, \quad t \geq 0. \]  

(A.4)
APPENDIX B

The Gompertz Model

We start with definition of the Gompertz differential equation from [13]:

\[
\frac{dY_t}{dt} = \beta Y_t \ln \left( \frac{Y_\infty}{Y_t} \right),
\]  

(B.1)

where \( \beta \) is a positive constant, and \( Y_\infty \) is the upper limit for the curve. We multiply \( 1/Y_\infty \) on both sides:

\[
\frac{dY_t}{dt} Y_\infty = \beta Y_t \ln \left( \frac{Y_\infty}{Y_t} \right)
\]

and then substitute \( x = Y_t/Y_\infty \) to get

\[
\frac{dx}{dt} = \beta x \ln \left( \frac{1}{x} \right).
\]  

(B.2)

Using the separation of variables, we solve for \( x \) as follows:

\[
\int \frac{dx}{x \ln \left( \frac{1}{x} \right)} = \int \beta dt
\]

\[
-\ln \left( \ln \frac{1}{x} \right) = \beta t + C, \text{ where } C \text{ is a constant}
\]

\[
\ln \left( \ln \frac{1}{x} \right) = -\beta t - C
\]

\[
\ln \frac{1}{x} = \alpha e^{-\beta t}, \text{ and } \alpha = e^{-C}
\]

\[
\ln x = -\alpha e^{-\beta t}
\]

\[
x = e^{-\alpha e^{-\beta t}}
\]

Finally, the substitution \( x = Y_t/Y_\infty \) gives the closed form solution of the Gompertz equation:

\[
Y_t = Y_\infty e^{-\alpha e^{-\beta t}}.
\]  

(B.3)
APPENDIX C

The Generalized Logistic Equation

The Generalized Logistic Equation (or sometimes called the Richard’s model) is given by:

\[
\frac{dY_t}{dt} = \beta Y_t \left[ 1 - \left( \frac{Y_t}{Y_\infty} \right)^r \right] = \beta Y_t - \frac{\beta}{Y_\infty} Y_t^{r+1}
\]  

(C.1)

Let \( \eta = -\frac{\beta}{Y_\infty} \). Then

\[
\frac{dY_t}{dt} \left( \frac{1}{Y_t^{r+1}} \right) = \frac{\beta}{Y_t^r} + \eta
\]

\[
\frac{dY_t}{dt} \left( \frac{1}{Y_t^{r+1}} \right) - \frac{\beta}{Y_t^r} = \eta
\]

We use the substitution \( w = \frac{1}{Y_t^r} = Y_t^{-r} \) to get

\[
w' = -rY_t^{r-1} \frac{dY_t}{dt} = \frac{-r \frac{dY_t}{dt}}{(Y_t^{(r+1)})}
\]

Note that

\[
-\frac{w'}{r} - \beta w = \eta
\]

\[
w' + r \beta w = -r \eta
\]

Using the method of integrating factor, we let \( \mu(t) = e^{\int r \beta dt} = e^{r \beta t} \) and multiply both sides of the equation with \( \mu(t) \):

\[
e^{r \beta t} w' + r \beta e^{r \beta t} w = -r \eta e^{r \beta t}
\]
Note that the left hand side of the above equation is the derivative of $e^{r\beta t}w$. Integrating both sides of the equation with respect to $t$ gives us

$$e^{r\beta t}w = -\int r\eta e^{r\beta t}dt$$

$$we^{r\beta t} = -\frac{r\eta}{r\beta}e^{r\beta t} + k,$$ where $k$ is the integration constant

$$we^{r\beta t} + \frac{\eta}{\beta}e^{r\beta t} = k$$

$$e^{r\beta t}\left(w + \frac{\eta}{\beta}\right) = k$$

$$w = \frac{k}{e^{r\beta t}} - \frac{\eta}{\beta}$$

$$\frac{1}{Y_{t}^r} = \frac{k}{e^{r\beta t}} - \frac{\eta}{\beta} = \frac{k\beta - \eta e^{r\beta t}}{e^{r\beta t}\beta}$$

We then obtain

$$Y_{t}^r = \frac{e^{r\beta t}\beta}{k\beta - \eta e^{r\beta t}} \quad \text{(C.2)}$$

Letting $Y(0) = Y_0$ we have

$$Y_{0}^r = \frac{\beta}{k\beta - \eta}$$

$$k\beta - \eta = \frac{\beta}{Y_{0}^r}$$

$$k\beta = \frac{\beta}{Y_{0}^r} + \eta$$

$$k = \frac{1}{Y_{0}^r} + \frac{\eta}{\beta} = \frac{1}{Y_{0}^r} - \frac{1}{Y_{\infty}^r}.$$
Substituting these values into (C.2):

\[
Y_t^r = \frac{1}{\left(\frac{e^{-r\beta t}}{\beta}\right) (k\beta - \eta e^{r\beta t})}
\]

\[
= \frac{1}{e^{-r\beta t} k - \left(\frac{\eta}{\beta}\right)}
\]

\[
= \frac{1}{e^{-r\beta t} \left(\frac{1}{Y_0^r} - \frac{1}{Y_\infty^r}\right) + \frac{1}{Y_\infty^r}}
\]

\[
= \frac{Y_\infty^r}{e^{-r\beta t} \left(\frac{Y_\infty^r}{Y_0^r}\right) - e^{-r\beta t} + 1}
\]

\[
= \frac{Y_\infty^r}{1 + \alpha e^{-\beta rt}}, \text{ where } \alpha = \frac{Y_\infty^r}{Y_0^r} - 1.
\]


By taking the \(r\)th root of both sides, we obtain the generalized logistic growth function:

\[
Y_t = \frac{Y_\infty^r}{(1 + \alpha e^{-\beta rt})^\frac{1}{r}}, \ t \geq 0 \quad \text{(C.3)}
\]
APPENDIX D

The Chapman-Richards Model

The von Bertalanffy proposed the following ODE to model the growth of a biological organism [13]:

\[
\frac{dY_t}{dt} = \eta Y_t^k - \gamma Y_t
\]  \hspace{1cm} (D.1)

Letting \( x = Y_t^{1-k} \), its derivative is given by

\[
\frac{dx}{dt} = (1-k)Y_t^{-k}\frac{dY_t}{dt}
\]

which gives

\[
\frac{dY_t}{dt} = \frac{1}{1-k} Y_t^k \frac{dx}{dt}.
\]  \hspace{1cm} (D.2)

By equating both forms of \( \frac{dY_t}{dt} \) in (D.1) and (D.2), and dividing by \( Y_t^k \), we solve for the closed form of von Bertalanffy equation:

\[
\frac{1}{1-k} \frac{dx}{dt} = \eta - \gamma Y_t^{1-k}.
\]

Substituting \( x = Y_t^{1-k} \), we obtain

\[
\int \frac{dx}{\eta - \gamma x} = \int (1-k)dt
\]

\[
-\gamma^{-1} \ln(\eta - \gamma x) = (1-k)t + C
\]

\[
\ln(\eta - \gamma x) = -\gamma(1-k)t - \gamma C
\]

\[
\eta - \gamma x = \alpha e^{-\gamma(1-k)t}, \text{ where } \alpha = e^{-\gamma C}
\]

\[
x = \frac{\eta}{\gamma} - \frac{\alpha}{\gamma} e^{-\gamma(1-k)t}.
\]
We then substitute $Y_t^{1-k}$ for $x$ to get

$$Y_t = \left[ \frac{\eta}{\gamma} - \frac{\alpha}{\gamma} e^{-\gamma(1-k)t} \right]^{1/(1-k)}. \tag{D.3}$$

For $t = 0$,

$$Y_0 = \left[ \frac{\eta}{\gamma} - \frac{\alpha}{\gamma} \right]^{1/(1-k)} \Rightarrow \frac{\alpha}{\gamma} = \frac{\eta}{\gamma} - Y_0^{1-k} \tag{D.4}$$

Substituting (D.4) into (D.3) we get:

$$Y_t = \left[ \frac{\eta}{\gamma} - \left( \frac{\eta}{\gamma} - Y_0^{1-k} \right) e^{-\gamma(1-k)t} \right]^{1/(1-k)}.$$

Since $\lim_{t \to \infty} Y_t = Y_\infty = \left( \frac{\eta}{\gamma} \right)^{1/(1-k)}$,

$$Y_t = \left[ Y_\infty^{1-k} - (Y_\infty^{1-k} - Y_0^{1-k}) e^{-\gamma(1-k)t} \right]^{1/(1-k)}.$$

We arrive at the von Bertalanffy growth equation:

$$Y_t = Y_\infty [1 - \beta e^{-\gamma(1-k)t}]^{1/(1-k)}, \text{ where } \beta = 1 - \left( \frac{Y_0}{Y_\infty} \right)^{1-k}. \tag{D.5}$$

By rewriting the parameters of von Bertalanffy, Chapman and Richards arrive at the final equation:

$$Y_t = Y_\infty [1 - ae^{-\lambda t}]^m. \tag{D.6}$$
### D.1 Data

#### D.1.1 No filter

Mean Square Error (MSE)

#### Table D.1: MSE with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.61135103</td>
<td>1.938620929</td>
<td>1.831250179</td>
<td>15.33841508</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.70122206</td>
<td>2.034266155</td>
<td>1.96325777</td>
<td>10.77792298</td>
</tr>
<tr>
<td>Weibull</td>
<td>2.120946758</td>
<td>1.503765706</td>
<td>2.5956849</td>
<td>31.20901334</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>1.293579949</td>
<td>1.095170952</td>
<td>4.522710583</td>
<td>39.18503707</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>1.299437909</td>
<td>1.09723158</td>
<td>4.515052626</td>
<td>39.15932587</td>
</tr>
</tbody>
</table>

#### Table D.2: MSE with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>1.800483147</td>
<td>2.08533946</td>
<td>1.995083273</td>
<td>54.33160091</td>
</tr>
<tr>
<td>Gompertz</td>
<td>1.910437072</td>
<td>2.117080353</td>
<td>1.869612992</td>
<td>50.10425501</td>
</tr>
<tr>
<td>Weibull</td>
<td>1.194105121</td>
<td>2.018420201</td>
<td>3.063977076</td>
<td>68.50349107</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>0.156830314</td>
<td>2.662841411</td>
<td>5.957570279</td>
<td>82.734415</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>0.886154736</td>
<td>5.149701411</td>
<td>10.71258562</td>
<td>98.6432342</td>
</tr>
</tbody>
</table>

#### Table D.3: MSE with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>10.69470413</td>
<td>24.91988161</td>
<td>37.49973957</td>
<td>664.5941422</td>
</tr>
<tr>
<td>Gompertz</td>
<td>10.25766667</td>
<td>24.11311424</td>
<td>36.03035963</td>
<td>652.1851722</td>
</tr>
<tr>
<td>Weibull</td>
<td>10.28566453</td>
<td>24.15709457</td>
<td>36.08489618</td>
<td>652.3824135</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>15.51644259</td>
<td>32.73336354</td>
<td>48.97689301</td>
<td>716.9550356</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>31.33035885</td>
<td>53.89153253</td>
<td>73.56245124</td>
<td>785.2451826</td>
</tr>
</tbody>
</table>
### Table D.4: MSE with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>3.726135938</td>
<td>8.74613049</td>
<td>8.077473188</td>
<td>428.4271285</td>
</tr>
<tr>
<td>Gompertz</td>
<td>4.742461287</td>
<td>10.85126113</td>
<td>7.297735</td>
<td>516.4951064</td>
</tr>
<tr>
<td>Weibull</td>
<td>111.8297312</td>
<td>136.6688715</td>
<td>84.47549707</td>
<td>866.0758109</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>33.23430693</td>
<td>55.81688673</td>
<td>45.82072379</td>
<td>1181.878319</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>51.71422851</td>
<td>79.05724671</td>
<td>66.48795532</td>
<td>1266.361606</td>
</tr>
</tbody>
</table>

### Table D.5: MSE with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>19.92395364</td>
<td>100.0127934</td>
<td>906.3739881</td>
<td>736.3018321</td>
</tr>
<tr>
<td>Gompertz</td>
<td>18.56187854</td>
<td>96.21082464</td>
<td>881.786391</td>
<td>676.3090507</td>
</tr>
<tr>
<td>Weibull</td>
<td>25.48713249</td>
<td>115.2253061</td>
<td>999.8426897</td>
<td>946.5744979</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>42.9318586</td>
<td>156.6388477</td>
<td>1190.154764</td>
<td>1248.220562</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>40.51928163</td>
<td>151.8599956</td>
<td>1175.75574</td>
<td>1232.410253</td>
</tr>
</tbody>
</table>

### Table D.6: MSE with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>691.8738066</td>
<td>450.8928624</td>
<td>305.0963175</td>
</tr>
<tr>
<td>Gompertz</td>
<td>397.0122517</td>
<td>501.646422</td>
<td>524.9902366</td>
</tr>
<tr>
<td>Weibull</td>
<td>2986.086057</td>
<td>1348.620977</td>
<td>1031.69712</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>381.4951379</td>
<td>425.3962578</td>
<td>198.942821</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>748.0353833</td>
<td>423.6319938</td>
<td>157.6057573</td>
</tr>
</tbody>
</table>

### Table D.7: MSE with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>290.3010775</td>
<td>301.8115534</td>
<td>867.3404451</td>
</tr>
<tr>
<td>Gompertz</td>
<td>316.3715881</td>
<td>327.0463339</td>
<td>895.4991004</td>
</tr>
<tr>
<td>Weibull</td>
<td>287.4328534</td>
<td>271.0701173</td>
<td>562.2934959</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>58.34814661</td>
<td>32.12436709</td>
<td>105.1386588</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>59.39813492</td>
<td>33.11909998</td>
<td>106.9589824</td>
</tr>
</tbody>
</table>
Forecast Difference

Note: Negative correspond to overestimation in forecasts, while positive numbers correspond to underestimation of forecasts. Units are in U.S. dollars.

Table D.8: Forecast Difference with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-2.066310904</td>
<td>0.662142318</td>
<td>-2.480770687</td>
<td>6.740436053</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-2.103747199</td>
<td>0.569037486</td>
<td>-2.910795642</td>
<td>5.194857981</td>
</tr>
<tr>
<td>Weibull</td>
<td>-1.847781119</td>
<td>1.18507066</td>
<td>-0.801632266</td>
<td>9.647463173</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-1.438128878</td>
<td>1.851865219</td>
<td>0.10186521</td>
<td>10.56186421</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-1.440795843</td>
<td>1.849181744</td>
<td>0.09918166</td>
<td>10.55918066</td>
</tr>
</tbody>
</table>

Table D.9: Forecast Difference with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-0.727687614</td>
<td>2.075221929</td>
<td>3.222525593</td>
<td>15.40425759</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-0.778357031</td>
<td>1.988727232</td>
<td>3.00033098</td>
<td>14.93770252</td>
</tr>
<tr>
<td>Weibull</td>
<td>-0.417377295</td>
<td>2.557194625</td>
<td>4.13370569</td>
<td>16.64318908</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>0.51198514</td>
<td>3.52198614</td>
<td>5.12198414</td>
<td>17.63198414</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>1.1075219</td>
<td>4.1175229</td>
<td>5.7175209</td>
<td>18.2275209</td>
</tr>
</tbody>
</table>

Table D.10: Forecast Difference with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>3.357168478</td>
<td>6.86100949</td>
<td>6.2245945</td>
<td>62.65587074</td>
</tr>
<tr>
<td>Gompertz</td>
<td>3.282331626</td>
<td>6.757179044</td>
<td>6.017809144</td>
<td>62.24497855</td>
</tr>
<tr>
<td>Weibull</td>
<td>3.28659424</td>
<td>6.762092192</td>
<td>6.022852113</td>
<td>62.25121457</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>4.066483983</td>
<td>7.706484637</td>
<td>7.3964826</td>
<td>64.1364866</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>5.72866624</td>
<td>9.36866724</td>
<td>9.05866524</td>
<td>65.79866924</td>
</tr>
</tbody>
</table>
Table D.11: Forecast Difference with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.385771094</td>
<td>4.771539117</td>
<td>1.542554833</td>
<td>19.74543194</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.713315502</td>
<td>5.229037776</td>
<td>2.640628169</td>
<td>25.04711562</td>
</tr>
<tr>
<td>Weibull</td>
<td>11.3661223</td>
<td>13.7656101</td>
<td>10.89390101</td>
<td>34.15105328</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>6.4856222</td>
<td>9.993897312</td>
<td>11.12329382</td>
<td>46.33328735</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>8.209253432</td>
<td>11.71551963</td>
<td>12.84401299</td>
<td>48.05399221</td>
</tr>
</tbody>
</table>

Table D.12: Forecast Difference with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>5.191962884</td>
<td>14.41322567</td>
<td>46.62142119</td>
<td>-3.500302616</td>
</tr>
<tr>
<td>Gompertz</td>
<td>5.015369782</td>
<td>14.15218463</td>
<td>45.91605572</td>
<td>-7.214083317</td>
</tr>
<tr>
<td>Weibull</td>
<td>5.859695626</td>
<td>15.41079745</td>
<td>49.12448701</td>
<td>5.245876076</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>7.450869562</td>
<td>17.61110106</td>
<td>53.15909336</td>
<td>11.9986534</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>7.339379866</td>
<td>17.4825779</td>
<td>53.01503836</td>
<td>11.85341717</td>
</tr>
</tbody>
</table>

Table D.13: Forecast Difference with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>16.44554473</td>
<td>-15.67002212</td>
<td>-20.19462441</td>
</tr>
<tr>
<td>Weibull</td>
<td>8.044992968</td>
<td>-24.5737186</td>
<td>-28.60646251</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>10.10905934</td>
<td>-18.38267036</td>
<td>-3.107529918</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>17.90783962</td>
<td>-11.93527626</td>
<td>0.069921092</td>
</tr>
</tbody>
</table>

Table D.14: Forecast Difference with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-18.94603776</td>
<td>-17.3070845</td>
<td>-60.31517522</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-19.70018713</td>
<td>-17.95053452</td>
<td>-60.76854887</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-7.84916078</td>
<td>-1.55461871</td>
<td>-26.35174228</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-7.942010079</td>
<td>-1.697978502</td>
<td>-26.56022142</td>
</tr>
</tbody>
</table>
D.1.2 Hodrick-Prescott Filter

Mean Square Error (MSE)

Table D.15: MSE with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>3.14</td>
<td>2.482836339</td>
<td>2.687134216</td>
<td>6.986980062</td>
</tr>
<tr>
<td>Gompertz</td>
<td>3.16812579</td>
<td>2.522960154</td>
<td>2.960802858</td>
<td>6.409337644</td>
</tr>
<tr>
<td>Weibull</td>
<td>3.15</td>
<td>2.51E+00</td>
<td>2.985237236</td>
<td>6.603449025</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>1.45</td>
<td>1.15E+00</td>
<td>4.326874982</td>
<td>38.51581804</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>3.14</td>
<td>2.488612783</td>
<td>2.725313696</td>
<td>6.789343938</td>
</tr>
</tbody>
</table>

Table D.16: MSE with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>3.268255972</td>
<td>2.841676677</td>
<td>2.073973497</td>
<td>29.27206475</td>
</tr>
<tr>
<td>Gompertz</td>
<td>3.358160819</td>
<td>2.90571996</td>
<td>2.226282854</td>
<td>24.10895201</td>
</tr>
<tr>
<td>Weibull</td>
<td>3.341454951</td>
<td>2.894599571</td>
<td>2.226008255</td>
<td>23.27260654</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>0.145267343</td>
<td>2.799319487</td>
<td>6.285097605</td>
<td>84.00143871</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>3.286002181</td>
<td>2.854144311</td>
<td>2.101803306</td>
<td>28.23959813</td>
</tr>
</tbody>
</table>

Table D.17: MSE with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>8.46968511</td>
<td>21.06393166</td>
<td>31.43812992</td>
<td>623.8057854</td>
</tr>
<tr>
<td>Gompertz</td>
<td>8.202362778</td>
<td>20.51400416</td>
<td>30.3008658</td>
<td>608.0868897</td>
</tr>
<tr>
<td>Weibull</td>
<td>8.06225778</td>
<td>20.21942417</td>
<td>29.67089433</td>
<td>597.6596448</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>25.30586286</td>
<td>46.02906226</td>
<td>64.51944499</td>
<td>761.2256188</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>8.24400015</td>
<td>20.59944641</td>
<td>30.4766176</td>
<td>610.6552474</td>
</tr>
</tbody>
</table>
Table D.18: MSE with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>3.982988865</td>
<td>9.283457336</td>
<td>7.780115081</td>
<td>453.0562864</td>
</tr>
<tr>
<td>Gompertz</td>
<td>4.872692766</td>
<td>11.09661809</td>
<td>7.287253988</td>
<td>523.0280666</td>
</tr>
<tr>
<td>Weibull</td>
<td>5.117064623</td>
<td>11.7049181</td>
<td>7.161920067</td>
<td>575.2289679</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>3.799803948</td>
<td>8.893864838</td>
<td>7.965383031</td>
<td>532.8350959</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>4.814877291</td>
<td>10.98188876</td>
<td>7.297876108</td>
<td>518.9677351</td>
</tr>
</tbody>
</table>

Table D.19: MSE with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>19.49550529</td>
<td>98.88928381</td>
<td>900.1640677</td>
<td>724.0282654</td>
</tr>
<tr>
<td>Gompertz</td>
<td>18.43511692</td>
<td>95.86178546</td>
<td>879.721134</td>
<td>672.2705621</td>
</tr>
<tr>
<td>Weibull</td>
<td>20.71078937</td>
<td>102.1319052</td>
<td>918.0005185</td>
<td>753.8243247</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>29.51304598</td>
<td>127.9355016</td>
<td>1091.71416</td>
<td>1134.78763</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>18.47832367</td>
<td>95.98176214</td>
<td>880.4812841</td>
<td>674.030009</td>
</tr>
</tbody>
</table>

Table D.20: MSE with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>381.4076452</td>
<td>516.9280736</td>
<td>602.749032</td>
</tr>
<tr>
<td>Gompertz</td>
<td>468.706295</td>
<td>472.1715354</td>
<td>418.3656329</td>
</tr>
<tr>
<td>Weibull</td>
<td>360.7075314</td>
<td>514.0727826</td>
<td>523.4484477</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>619.0322864</td>
<td>439.1713901</td>
<td>228.3569864</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>503.6905178</td>
<td>453.9548338</td>
<td>290.8366148</td>
</tr>
</tbody>
</table>

Table D.21: MSE with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>334.5253618</td>
<td>345.9903777</td>
<td>914.6566041</td>
</tr>
<tr>
<td>Gompertz</td>
<td>332.4030947</td>
<td>345.7409865</td>
<td>941.0796625</td>
</tr>
<tr>
<td>Weibull</td>
<td>316.0754505</td>
<td>314.7064835</td>
<td>749.4005854</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>17.95971655</td>
<td>11.3953223</td>
<td>81.94574406</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>125.2538798</td>
<td>90.66611107</td>
<td>176.5710753</td>
</tr>
</tbody>
</table>

116
Forecast Difference

Note: Negative correspond to overestimation in forecasts, while positive numbers correspond to underestimation of forecasts. Units are in U.S. dollars.

Table D.22: Forecast Difference with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-0.008415298</td>
<td>-0.014301092</td>
<td>0.013383698</td>
<td>6.995860441</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-0.015109631</td>
<td>-0.034434975</td>
<td>-0.127789896</td>
<td>5.921448366</td>
</tr>
<tr>
<td>Weibull</td>
<td>-0.004551068</td>
<td>-0.011185207</td>
<td>-0.043801143</td>
<td>6.089955726</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-0.001395194</td>
<td>0.007408208</td>
<td>0.167018255</td>
<td>7.940645832</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>0.004006005</td>
<td>0.024305323</td>
<td>0.285719445</td>
<td>8.524301828</td>
</tr>
</tbody>
</table>

Table D.23: Forecast Difference with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>0.091456871</td>
<td>0.263383777</td>
<td>1.764616031</td>
<td>22.5438809</td>
</tr>
<tr>
<td>Gompertz</td>
<td>0.065465285</td>
<td>0.212582267</td>
<td>1.573874101</td>
<td>21.71695979</td>
</tr>
<tr>
<td>Weibull</td>
<td>0.066122699</td>
<td>0.210534936</td>
<td>1.543841054</td>
<td>21.45731051</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>0.083246532</td>
<td>0.247250255</td>
<td>1.704444771</td>
<td>22.29852757</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>0.077709916</td>
<td>0.236388956</td>
<td>1.663513139</td>
<td>22.12318071</td>
</tr>
</tbody>
</table>

Table D.24: Forecast Difference with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>0.616459482</td>
<td>1.073309839</td>
<td>4.055010712</td>
<td>35.56383512</td>
</tr>
<tr>
<td>Gompertz</td>
<td>0.606327059</td>
<td>1.048157958</td>
<td>3.925942575</td>
<td>34.56101197</td>
</tr>
<tr>
<td>Weibull</td>
<td>0.709596252</td>
<td>1.180538704</td>
<td>4.163899581</td>
<td>34.99633086</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>0.606763492</td>
<td>1.05022314</td>
<td>3.93958177</td>
<td>34.68263727</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>0.607685884</td>
<td>1.052200636</td>
<td>3.94908127</td>
<td>34.75365463</td>
</tr>
</tbody>
</table>
Table D.25: Forecast Difference with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>1.743689131</td>
<td>3.024696582</td>
<td>11.5614518</td>
<td>17.87814296</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.025912464</td>
<td>3.410911831</td>
<td>12.45724241</td>
<td>22.10166844</td>
</tr>
<tr>
<td>Weibull</td>
<td>1.765377871</td>
<td>3.116396744</td>
<td>12.134019</td>
<td>22.97501495</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>1.754728259</td>
<td>3.039064065</td>
<td>11.63594002</td>
<td>26.73289831</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>10.07404053</td>
<td>12.43016539</td>
<td>25.12719578</td>
<td>47.57826876</td>
</tr>
</tbody>
</table>

Table D.26: Forecast Difference with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>6.705109637</td>
<td>9.274409414</td>
<td>13.38242541</td>
<td>-29.64868752</td>
</tr>
<tr>
<td>Gompertz</td>
<td>7.395251896</td>
<td>10.19510614</td>
<td>15.44813183</td>
<td>-19.49862191</td>
</tr>
<tr>
<td>Weibull</td>
<td>4.234554445</td>
<td>6.225852605</td>
<td>8.681303607</td>
<td>-27.60409232</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>1.754728259</td>
<td>3.039064065</td>
<td>11.63594002</td>
<td>26.73289831</td>
</tr>
</tbody>
</table>

Table D.27: Forecast Difference with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-0.188315936</td>
<td>-3.62324445</td>
<td>-22.92033175</td>
</tr>
<tr>
<td>Gompertz</td>
<td>1.892872138</td>
<td>-0.821910946</td>
<td>-16.40615778</td>
</tr>
<tr>
<td>Weibull</td>
<td>-1.777275786</td>
<td>-4.880709853</td>
<td>-20.09416812</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>0.13842415</td>
<td>-2.965114941</td>
<td>-17.39558626</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>0.739168347</td>
<td>-2.021867697</td>
<td>-13.8706468</td>
</tr>
</tbody>
</table>

Table D.28: Forecast Difference with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>4.822888331</td>
<td>2.809086553</td>
<td>-9.24708424</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-12.9152326</td>
<td>-18.52542613</td>
<td>-42.92180762</td>
</tr>
<tr>
<td>Weibull</td>
<td>-3.114465025</td>
<td>-5.227639673</td>
<td>-14.25369377</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-2.978530658</td>
<td>-4.350017956</td>
<td>-11.59722922</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-3.210986067</td>
<td>-4.647619835</td>
<td>-11.98960708</td>
</tr>
</tbody>
</table>
D.1.3 Exponential Smoothing

Mean Square Error (MSE)

Table D.29: MSE with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.669944362</td>
<td>1.994311526</td>
<td>1.859967218</td>
<td>14.23897108</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.762401552</td>
<td>2.094862134</td>
<td>2.036553555</td>
<td>9.714757291</td>
</tr>
<tr>
<td>Weibull</td>
<td>1.94167264</td>
<td>1.366310537</td>
<td>3.85164855</td>
<td>36.77015075</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>1.312027382</td>
<td>1.101692763</td>
<td>4.499461711</td>
<td>39.10669887</td>
</tr>
</tbody>
</table>

Table D.30: MSE with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.365195256</td>
<td>2.316449902</td>
<td>1.748999339</td>
<td>44.45182348</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.010209922</td>
<td>2.156101118</td>
<td>1.825256389</td>
<td>48.563441111</td>
</tr>
<tr>
<td>Weibull</td>
<td>1.250578778</td>
<td>2.024860438</td>
<td>2.999704891</td>
<td>68.05853326</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>1.111565046</td>
<td>2.042198386</td>
<td>3.303800886</td>
<td>70.08249952</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>1.22667632</td>
<td>2.054372709</td>
<td>3.182756702</td>
<td>69.30544503</td>
</tr>
</tbody>
</table>

Table D.31: MSE with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>10.73385612</td>
<td>24.97979318</td>
<td>37.5764782</td>
<td>664.9106837</td>
</tr>
<tr>
<td>Gompertz</td>
<td>10.29517381</td>
<td>24.17068621</td>
<td>36.1037297</td>
<td>652.5044372</td>
</tr>
<tr>
<td>Weibull</td>
<td>10.36463391</td>
<td>24.29087237</td>
<td>36.29633479</td>
<td>653.9414169</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>10.374575404</td>
<td>24.31837733</td>
<td>36.37521187</td>
<td>654.9505867</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>10.35296661</td>
<td>24.27712385</td>
<td>36.29729237</td>
<td>654.2356193</td>
</tr>
</tbody>
</table>
Table D.32: MSE with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>3.745947608</td>
<td>8.779146568</td>
<td>8.076462059</td>
<td>428.3595572</td>
</tr>
<tr>
<td>Gompertz</td>
<td>4.798205752</td>
<td>10.93308602</td>
<td>7.319144661</td>
<td>514.916165</td>
</tr>
<tr>
<td>Weibull</td>
<td>4.726824011</td>
<td>10.92323258</td>
<td>7.153879283</td>
<td>550.1415442</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>3.838510766</td>
<td>8.984984786</td>
<td>7.922412924</td>
<td>442.433408</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>3.822858374</td>
<td>8.949865613</td>
<td>7.948129622</td>
<td>440.0709937</td>
</tr>
</tbody>
</table>

Table D.33: MSE with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>20.03521194</td>
<td>100.2625705</td>
<td>907.3350164</td>
<td>737.6272655</td>
</tr>
<tr>
<td>Gompertz</td>
<td>18.67207357</td>
<td>96.46006556</td>
<td>882.7315455</td>
<td>677.5116702</td>
</tr>
<tr>
<td>Weibull</td>
<td>25.80770342</td>
<td>115.9158129</td>
<td>1002.485035</td>
<td>950.6577789</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>41.97263962</td>
<td>154.7956104</td>
<td>1184.825706</td>
<td>1242.432728</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>40.71034937</td>
<td>152.2234075</td>
<td>1176.802899</td>
<td>1233.548805</td>
</tr>
</tbody>
</table>

Table D.34: MSE with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>308.4063305</td>
<td>564.8859472</td>
<td>769.7007274</td>
</tr>
<tr>
<td>Gompertz</td>
<td>398.0235938</td>
<td>501.2133996</td>
<td>523.7149649</td>
</tr>
<tr>
<td>Weibull</td>
<td>348.3991981</td>
<td>521.6327645</td>
<td>546.9904746</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>308.7844532</td>
<td>564.5352978</td>
<td>768.229178</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>306.4900621</td>
<td>566.0357431</td>
<td>772.4602557</td>
</tr>
</tbody>
</table>

Table D.35: MSE with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>280.6887896</td>
<td>279.3134455</td>
<td>715.4719307</td>
</tr>
<tr>
<td>Gompertz</td>
<td>324.4042263</td>
<td>335.9050986</td>
<td>912.850504</td>
</tr>
<tr>
<td>Weibull</td>
<td>320.3188007</td>
<td>316.8772496</td>
<td>730.27608</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>667.8927667</td>
<td>409.4925154</td>
<td>153.720457</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>74.66805199</td>
<td>45.49988629</td>
<td>121.2489878</td>
</tr>
</tbody>
</table>

120
Forecast Difference

Note: Negative correspond to overestimation in forecasts, while positive numbers correspond to underestimation of forecasts. Units are in U.S. dollars.

Table D.36: Forecast Difference with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-2.262918868</td>
<td>0.262658636</td>
<td>-3.874716821</td>
<td>2.694651597</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-2.274469883</td>
<td>0.226360757</td>
<td>-4.149266786</td>
<td>0.450379975</td>
</tr>
<tr>
<td>Weibull</td>
<td>-2.269373774</td>
<td>0.233682636</td>
<td>-4.182065623</td>
<td>-0.518034138</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-1.508252049</td>
<td>1.781515189</td>
<td>0.031514465</td>
<td>10.49151346</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-2.264591926</td>
<td>0.257344768</td>
<td>-3.914986438</td>
<td>2.404411214</td>
</tr>
</tbody>
</table>

Table D.37: Forecast Difference with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-1.267141611</td>
<td>1.296587923</td>
<td>1.601835493</td>
<td>12.19785897</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-1.298707999</td>
<td>1.234099357</td>
<td>1.359918685</td>
<td>11.09998192</td>
</tr>
<tr>
<td>Weibull</td>
<td>-1.29375167</td>
<td>1.239142102</td>
<td>1.34697859</td>
<td>10.85696325</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>0.593925679</td>
<td>3.603923635</td>
<td>5.203921536</td>
<td>17.71392154</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-1.273431493</td>
<td>1.284056306</td>
<td>1.553519077</td>
<td>11.99542495</td>
</tr>
</tbody>
</table>

Table D.38: Forecast Difference with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.970745734</td>
<td>6.377748151</td>
<td>5.433157246</td>
<td>61.34159653</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.917629416</td>
<td>6.295763722</td>
<td>5.226217406</td>
<td>60.74600871</td>
</tr>
<tr>
<td>Weibull</td>
<td>2.889211568</td>
<td>6.25069238</td>
<td>5.103922385</td>
<td>60.32833679</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>5.160097401</td>
<td>8.799282862</td>
<td>8.489150847</td>
<td>65.22915472</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>2.925944574</td>
<td>6.308553712</td>
<td>5.258774108</td>
<td>60.84562596</td>
</tr>
</tbody>
</table>
### Table D.39: Forecast Difference with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.473843683</td>
<td>4.893205285</td>
<td>1.838675494</td>
<td>21.42038939</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.750559053</td>
<td>5.276063558</td>
<td>2.731944806</td>
<td>25.3788115</td>
</tr>
<tr>
<td>Weibull</td>
<td>2.826655137</td>
<td>5.415999675</td>
<td>3.224445145</td>
<td>28.25414299</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>2.40966313</td>
<td>4.80448294</td>
<td>1.677645871</td>
<td>30.21645533</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>2.733772539</td>
<td>5.253137994</td>
<td>2.680286686</td>
<td>25.16638851</td>
</tr>
</tbody>
</table>

### Table D.40: Forecast Difference with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>5.138138884</td>
<td>14.33941019</td>
<td>46.45675916</td>
<td>-4.096021363</td>
</tr>
<tr>
<td>Gompertz</td>
<td>4.998690285</td>
<td>14.12833484</td>
<td>45.85929082</td>
<td>-7.445213018</td>
</tr>
<tr>
<td>Weibull</td>
<td>5.290685464</td>
<td>14.55338452</td>
<td>46.9179543</td>
<td>-3.004556288</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>6.327918276</td>
<td>16.26147526</td>
<td>51.48929617</td>
<td>10.26514451</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>5.004385544</td>
<td>14.1366081</td>
<td>45.88111895</td>
<td>-7.327423602</td>
</tr>
</tbody>
</table>

### Table D.41: Forecast Difference with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>9.256473656</td>
<td>-23.76144198</td>
<td>-32.31856511</td>
</tr>
<tr>
<td>Gompertz</td>
<td>11.64395473</td>
<td>-20.48367338</td>
<td>-24.29960552</td>
</tr>
<tr>
<td>Weibull</td>
<td>8.778315117</td>
<td>-23.6875907</td>
<td>-26.99182023</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>15.11971468</td>
<td>-16.08340281</td>
<td>-9.501655126</td>
</tr>
</tbody>
</table>

### Table D.42: Forecast Difference with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-20.22548227</td>
<td>-18.46927024</td>
<td>-60.87177879</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-20.18462503</td>
<td>-18.55216347</td>
<td>-61.92458382</td>
</tr>
<tr>
<td>Weibull</td>
<td>-19.55873347</td>
<td>-17.04703306</td>
<td>-55.28220878</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-1.451166873</td>
<td>4.865345005</td>
<td>-19.9106586</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-11.97192188</td>
<td>-6.13246367</td>
<td>-31.49059789</td>
</tr>
</tbody>
</table>
### Moving average

Mean Square Error (MSE)

#### Table D.43: MSE with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.783553787</td>
<td>2.479959358</td>
<td>6.512169118</td>
<td>86.52237742</td>
</tr>
<tr>
<td>Gompertz</td>
<td>2.538221027</td>
<td>2.18670505</td>
<td>4.823457284</td>
<td>35.82214713</td>
</tr>
<tr>
<td>Weibull</td>
<td>9.278385675</td>
<td>12.00970097</td>
<td>17.55032478</td>
<td>15.8358527</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>2.781715323</td>
<td>2.477492061</td>
<td>6.491756386</td>
<td>85.34189969</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>2.777375901</td>
<td>2.471967544</td>
<td>6.441968335</td>
<td>79.92021265</td>
</tr>
</tbody>
</table>

#### Table D.44: MSE with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>4.412226784</td>
<td>3.711932122</td>
<td>3.627867348</td>
<td>14.43022009</td>
</tr>
<tr>
<td>Gompertz</td>
<td>4.569694085</td>
<td>3.856002982</td>
<td>4.177404381</td>
<td>9.543539458</td>
</tr>
<tr>
<td>Weibull</td>
<td>3.337214839</td>
<td>2.866321033</td>
<td>1.862662718</td>
<td>45.71176139</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>2.043808258</td>
<td>2.26078458</td>
<td>2.596196972</td>
<td>64.80701341</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>2.121283323</td>
<td>2.287190281</td>
<td>2.548615182</td>
<td>64.37108156</td>
</tr>
</tbody>
</table>

#### Table D.45: MSE with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>12.16433469</td>
<td>26.98542795</td>
<td>39.71706267</td>
<td>670.0284523</td>
</tr>
<tr>
<td>Gompertz</td>
<td>11.63760329</td>
<td>26.02763744</td>
<td>37.96275961</td>
<td>655.1270863</td>
</tr>
<tr>
<td>Weibull</td>
<td>13.73913945</td>
<td>29.72743468</td>
<td>44.38069737</td>
<td>698.7482607</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>13.75559166</td>
<td>29.77766389</td>
<td>44.49996193</td>
<td>699.1445167</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>13.75537277</td>
<td>29.77725868</td>
<td>44.49923213</td>
<td>699.1406427</td>
</tr>
</tbody>
</table>
Table D.46: MSE with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>5.142166985</td>
<td>11.01945988</td>
<td>8.061311933</td>
<td>433.0551156</td>
</tr>
<tr>
<td>Gompertz</td>
<td>6.265782064</td>
<td>13.16580912</td>
<td>7.729422718</td>
<td>514.7802866</td>
</tr>
<tr>
<td>Weibull</td>
<td>5.665775838</td>
<td>12.20221223</td>
<td>7.552815115</td>
<td>528.3374692</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>5.166742234</td>
<td>11.07522786</td>
<td>8.026588796</td>
<td>437.5257159</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>46.75435489</td>
<td>72.82795307</td>
<td>60.83293425</td>
<td>1244.298544</td>
</tr>
</tbody>
</table>

Table D.47: MSE with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>26.2453239</td>
<td>112.7899884</td>
<td>944.7882963</td>
<td>772.9001994</td>
</tr>
<tr>
<td>Gompertz</td>
<td>24.78999089</td>
<td>108.9637432</td>
<td>920.4360407</td>
<td>711.3290176</td>
</tr>
<tr>
<td>Weibull</td>
<td>31.56192177</td>
<td>126.5662578</td>
<td>1027.877097</td>
<td>964.5900265</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>47.76848652</td>
<td>164.98888151</td>
<td>1211.615384</td>
<td>1270.993694</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>169.9131162</td>
<td>351.0156976</td>
<td>1639.711449</td>
<td>1725.989964</td>
</tr>
</tbody>
</table>

Table D.48: MSE with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>691.8738066</td>
<td>450.8928624</td>
<td>305.0963175</td>
</tr>
<tr>
<td>Gompertz</td>
<td>807.7989969</td>
<td>446.1882789</td>
<td>216.6611512</td>
</tr>
<tr>
<td>Weibull</td>
<td>653.4003697</td>
<td>447.9109943</td>
<td>285.4299857</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>769.1785331</td>
<td>439.856513</td>
<td>185.4372621</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>740.814953</td>
<td>433.2490506</td>
<td>174.6319294</td>
</tr>
</tbody>
</table>

Table D.49: MSE with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>290.3010775</td>
<td>301.8115534</td>
<td>867.3404451</td>
</tr>
<tr>
<td>Gompertz</td>
<td>276.7096807</td>
<td>287.5778911</td>
<td>854.2818288</td>
</tr>
<tr>
<td>Weibull</td>
<td>285.1606003</td>
<td>287.815095</td>
<td>750.71414</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>70.31857887</td>
<td>42.19105593</td>
<td>118.0810847</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>22.82271923</td>
<td>23.42658721</td>
<td>93.45590279</td>
</tr>
</tbody>
</table>
Forecast Difference

Note: Negative correspond to overestimation in forecasts, while positive numbers correspond to underestimation of forecasts. Units are in U.S. dollars.

Table D.50: Forecast Difference with 1000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-2.21992966</td>
<td>-0.040314981</td>
<td>-6.460394281</td>
<td>-19.48797949</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-2.127300056</td>
<td>0.145172059</td>
<td>-5.616302599</td>
<td>-10.86211149</td>
</tr>
<tr>
<td>Weibull</td>
<td>-3.90147552</td>
<td>-2.790527773</td>
<td>-6.972131628</td>
<td>3.473705776</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-2.219209188</td>
<td>-0.038662229</td>
<td>-6.450667057</td>
<td>-19.29816221</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-2.21758517</td>
<td>-0.035083929</td>
<td>-6.425831193</td>
<td>-18.24393338</td>
</tr>
</tbody>
</table>

Table D.51: Forecast Difference with 2000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-1.611344652</td>
<td>0.761044237</td>
<td>0.325405529</td>
<td>8.97215896</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-1.6578925</td>
<td>0.670213133</td>
<td>-0.047254275</td>
<td>6.722920973</td>
</tr>
<tr>
<td>Weibull</td>
<td>-1.266574622</td>
<td>1.410430382</td>
<td>2.374244468</td>
<td>14.65125837</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-0.756759184</td>
<td>2.251666205</td>
<td>3.851551279</td>
<td>16.36155127</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>-0.787265749</td>
<td>2.218110837</td>
<td>3.817564271</td>
<td>16.32756407</td>
</tr>
</tbody>
</table>

Table D.52: Forecast Difference with 3000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>3.57237535</td>
<td>7.064050431</td>
<td>6.395321345</td>
<td>62.78925595</td>
</tr>
<tr>
<td>Gompertz</td>
<td>3.487488344</td>
<td>6.944972134</td>
<td>6.15322809</td>
<td>62.29488477</td>
</tr>
<tr>
<td>Weibull</td>
<td>3.813352741</td>
<td>7.384042533</td>
<td>6.954058839</td>
<td>63.65207079</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>3.815514235</td>
<td>7.390558355</td>
<td>6.967696968</td>
<td>63.66090836</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>3.815480143</td>
<td>7.390509924</td>
<td>6.967612977</td>
<td>63.66079905</td>
</tr>
</tbody>
</table>
Table D.53: Forecast Difference with 4000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>2.792258756</td>
<td>5.164854658</td>
<td>1.871579972</td>
<td>19.66136388</td>
</tr>
<tr>
<td>Gompertz</td>
<td>3.102266284</td>
<td>5.595267605</td>
<td>2.897681306</td>
<td>24.65748832</td>
</tr>
<tr>
<td>Weibull</td>
<td>2.930810017</td>
<td>5.413530447</td>
<td>2.768938023</td>
<td>25.82768393</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>2.799558686</td>
<td>5.177488234</td>
<td>1.914514119</td>
<td>20.0053434</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>7.847033254</td>
<td>11.34706546</td>
<td>12.47274528</td>
<td>47.68267735</td>
</tr>
</tbody>
</table>

Table D.54: Forecast Difference with 5000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>5.873120348</td>
<td>15.10541143</td>
<td>47.34919267</td>
<td>-2.693802561</td>
</tr>
<tr>
<td>Gompertz</td>
<td>5.701733247</td>
<td>14.84984782</td>
<td>46.64949736</td>
<td>-6.41807983</td>
</tr>
<tr>
<td>Weibull</td>
<td>6.446343308</td>
<td>15.97029499</td>
<td>49.56036761</td>
<td>5.264815663</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>7.898629238</td>
<td>18.05334892</td>
<td>53.59800879</td>
<td>12.43745346</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>14.08575591</td>
<td>24.26514699</td>
<td>59.82483848</td>
<td>18.66482709</td>
</tr>
</tbody>
</table>

Table D.55: Forecast Difference with 6000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>16.44548575</td>
<td>-15.67008809</td>
<td>-20.19472133</td>
</tr>
<tr>
<td>Gompertz</td>
<td>18.79869837</td>
<td>-12.52372989</td>
<td>-12.94842988</td>
</tr>
<tr>
<td>Weibull</td>
<td>15.70444051</td>
<td>-16.16262837</td>
<td>-17.3620608</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>18.08976838</td>
<td>-12.91439742</td>
<td>-5.731746716</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>17.61827939</td>
<td>-13.00579795</td>
<td>-3.356993845</td>
</tr>
</tbody>
</table>

Table D.56: Forecast Difference with 7000 Prior Known Days

<table>
<thead>
<tr>
<th>Forecast Days:</th>
<th>100</th>
<th>300</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic</td>
<td>-18.94607399</td>
<td>-17.30712349</td>
<td>-60.31522455</td>
</tr>
<tr>
<td>Gompertz</td>
<td>-18.56316153</td>
<td>-16.93658111</td>
<td>-60.36809333</td>
</tr>
<tr>
<td>Weibull</td>
<td>-18.73304387</td>
<td>-16.52367769</td>
<td>-56.1086117</td>
</tr>
<tr>
<td>Chapman Richards</td>
<td>-8.768779743</td>
<td>-2.632819952</td>
<td>-27.59889037</td>
</tr>
<tr>
<td>Generalized Logistic</td>
<td>0.760684731</td>
<td>7.132579374</td>
<td>-17.59355243</td>
</tr>
</tbody>
</table>
APPENDIX E

MATLAB Code

E.1 Filters

E.1.1 Moving Average

If the time point has less data points than the moving average length, take
the average of the prior data points.

data: Raw data. MALength: Length of moving average.

function [ result ] = meanAvg( data, MALength )

t = length(data);
for i = 1:t
    if i == 1
        result(1) = data(1);
    end
    if i <= MALength
        A = data(1:i);
        result(i) = 1/i * sum(A);
    end
    if i > MALength
        k = i-MALength;
...
A = data(k:i);
result(i) = 1/MALength * sum(A);
end
end

*Note: MATLAB has its own built in function, "movmean."

E.2 Exponential Filter

This MATLAB code outputs columns of various weight parameters. The number of columns and the size of the weight parameter can easily be changed as per the user’s specifications.

function [ result ] = alpha( data, PriorDataLength)

for i = 1:PriorDataLength
for k = 1:9
if i == 1
result(1, k)= data(1);
end

if i ~= 1
result(i, k) = result(i-1, k) + k * 0.1* (data(i-1) - result(i-1, k)) ;
end
end
end
*Note: MATLAB has its own built in function, ”tsmovavg” in the Financial Toolbox.

E.2.1 The Hodrick-Prescott Filter

\[ y : \text{Unfiltered data} \quad w: \text{lambda constant for Hodrick-Prescott Filter} \]

function \([s, \text{desvabs}] = \text{hpfilter}(y,w,\text{plotter})\)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Author: Wilmer Henao wi-henao@uniandes.edu.co
% Department of Mathematics
% Universidad de los Andes
% Colombia
%
% Hodrick-Prescott filter extracts the trend of a time series, the
% output is not a formula but a new filtered time series. This
% trend can be adjusted with parameter w; values for w lie usually
% in the interval \([100,20000]\), and it is up to you to use the one
% you like, As w approaches infty, H-P will approach a line. If
% the series doesn’t have a trend p.e.White Noise, doing H-P is
% meaningless
% \([s] = \text{hpfilter}(y,w)\)
% w = Smoothing parameter (Economists advice: "Use w = 1600 for
% quarterly data")
% y = Original series
\% s = Filtered series
\%
\% This program can work with several series at a time, as long
\% as the number of series you are working with doesn’t exceed
\% the number of elements in the series + it uses sparse matrices
\$
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\%
\[ [m,n] = \text{size} (y); \]

if \( m < n \)

\[ y = y'; \quad m = n; \]

end

\[ d = \text{repmat}([w -4*w \ ((6*w+1)/2)], m, 1); \]
\[ d(1,2) = -2*w; \quad d(m-1,2) = -2*w; \]
\[ d(1,3) = (1+w)/2; \quad d(m,3) = (1+w)/2; \]
\[ d(2,3) = (5*w+1)/2; \quad d(m-1,3) = (5*w+1)/2; \]

\[ B = \text{spdiags}(d, -2:0, m, m); \quad \% \text{I use a sparse version of } B, \]

because when \( m \) is large, \( B \) will have many zeros

\[ B = B+B'; \]

\[ s = B\backslash y; \]

if nargin == 3

\[ t = \text{size}(y,2); \]

for \( i = 1:t \)

\[ \text{figure}(i) \]
\[ \text{plot}(s(:,i),'r'); \quad \text{grid} \text{ on}; \quad \text{hold} \text{ on}; \quad \text{plot}(y(:,i)); \]
\[ \text{title}(['Series #',\text{num2str}(i)]); \]

end

end

if nargout == 2

\[ \text{desvabs} = \text{mean}(\text{abs}(y-s)./s); \]

end
E.3 Fitting

E.3.1 Polynomial Fit

   parameter = polyfit(x value,y value raw data,polynomial order)

y values of fit = polyval( parameter, x value)

E.3.2 The Levenburg-Marquart Algorithm

   [beta,R,J] = nlinfit(days,raw data, function ,test parameters);

E.4 MSE and Difference of Forecast

This function calculates the mean square error (MSE) and difference between
the forecasted final time point and actual data point.

   a : parameters from LM algorithm func: function PriorDataLength: Number
of known time points ydata: Entire raw data set

function [ mse, diff ] = results( a, func, PriorDataLength , ydata)

for i = 1:4

   if i == 1
      MSE and difference calculation for 100 future points

end


toc

*Note: MATLAB has its own built in function, ”hpfilter” in the Econometrics Toolbox.
A1 = ydata(1: (PriorDataLength + 100));

B2 = transpose (func(a,(1: (PriorDataLength + 100))));

A = A1( (PriorDataLength + 1):(PriorDataLength + 100) );
B = B2( (PriorDataLength + 1):(PriorDataLength + 100) );

C = (minus(A,B)).^2;

mse(i) = sum( C )/100;

diff(i) = ydata(PriorDataLength + 100) - func(a,(PriorDataLength + 100)) ;
end

if i == 2 \%MSE and difference calculation for 300 future points

A1 = ydata(1: (PriorDataLength + 300));

B2 = transpose (func(a,(1: (PriorDataLength + 300))));

A = A1( (PriorDataLength + 1):(PriorDataLength + 300) );
B = B2( (PriorDataLength + 1):(PriorDataLength + 300) );

C = (minus(A,B)).^2;

mse(i) = sum( C )/300;

diff(i) = ydata(PriorDataLength + 300) - func(a,(PriorDataLength + 300)) ;
end

if i == 3 \%MSE and difference calculation for 1000 future points

A1 = ydata(1: (PriorDataLength + 1000));
B2 = transpose (func(a,(1: (PriorDataLength + 1000))));

A = A1( (PriorDataLength + 1):(PriorDataLength + 1000) );

B = B2( (PriorDataLength + 1):(PriorDataLength + 1000) );

C = (minus(A,B)).^2;

mse(i) = sum( C )/1000;

diff(i) = ydata(PriorDataLength + 1000)
- func(a,(PriorDataLength + 1000)) ;
end

if i == 4  %%MSE and difference calculation for 3000 future points
A1 = ydata(1: (PriorDataLength + 3000));

B2 = transpose (func(a,(1: (PriorDataLength + 3000))));

A = A1( (PriorDataLength + 1):(PriorDataLength + 3000) );

B = B2( (PriorDataLength + 1):(PriorDataLength + 3000) );

C = (minus(A,B)).^2;

mse(i) = sum( C )/3000;

diff(i) = ydata(PriorDataLength + 3000)
- func(a,(PriorDataLength + 3000)) ;
end

end
mse = transpose(mse);

diff = transpose(diff);

end