Derivative Free Optimization

Lecture 1: Introduction

Warren Hare

Textbook

https://link-springer-com.ezproxy.library.ubc.ca/book/10.1007/978-3-319-68913-5

What is DFO/BBO?

Background and Notation

Students expected to understand

- Multivariate Calculus (MATH 200)
- Linear Algebra (MATH 221)
- Mathematical Proofs (MATH 220)
- Basic Computer Programming (e.g., COSC 111 or 121, etc)

Notation

Some basic notation

- $f : \mathbb{R}^n \mapsto \mathbb{R}$ – f is a function from $\mathbb{R}^n$ to $\mathbb{R}$
- $f \in \mathcal{C}^0$, $f \in \mathcal{C}^{0+}$, $f \in \mathcal{C}^k$, etc, – describes the smoothness of f
- ∇f, $\nabla^2 f$ – gradients and Hessian*
- All points are vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- $x \in [\ell, u]$ – Vector-based interval notation

$*$ we will review Hessian's when required

Optimization

Optimization is the *mathematical study* of the problem

$$\min\{f(x) : x \in \Omega\}$$

and

$$\operatorname{argmin}\{f(x) : x \in \Omega\}$$

Optimization

We are equally interested in the problem

$$\max\{f(x) : x \in \Omega\}$$

and

$$\operatorname{argmax}\{f(x) : x \in \Omega\}$$

Important Remark

$$\min\{f(x) : x \in \Omega\} = -\max\{-f(x) : x \in \Omega\}$$

and

$$\mathrm{argmin}\{f(x) : x \in \Omega\} = \mathrm{argmax}\{-f(x) : x \in \Omega\}$$

Derivative-Free Optimization

Definition: **Derivative-Free Optimization** (DFO) is the *mathematical study* of optimization algorithms that do not *use* derivatives

Definition: **Blackbox Optimization** (BBO) is the study of design and analysis of optimization algorithms that assume the *objective and/or constraint functions* are *given by blackboxes*

Example: Simulating Home Care

Home and Community Care (HCC)

HCC is (very loosely) the branch of healthcare which deals with long term health services outside of a hospital
It consists of both *publicly funded* and *privately funded* options

In 2006, the BC Ministry of Health inquired:

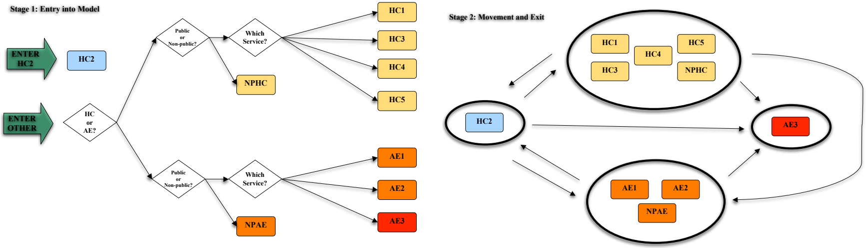
- How will publicly funded HCC look in the year 2011, 2016, 2021, 2026, 2031?

To answer the questions they helped create a model that incorporated

- Changing Population Age Demographics
- Changing Relationship of Age to Health
- Changing Wealth Demographics
- Relationship of Wealth to Non-public HCC options

The Model

Deterministic Multi-state Markov Model



The Data

Changing Population Age Demographics

Changing Relationship of Age to Health

Changing Wealth Demographics

Relationship of Wealth to Non-public HCC options

The Data

Changing Population Age Demographics

- Population projection data provided by BC Stats.

Changing Relationship of Age to Health

Changing Wealth Demographics

Relationship of Wealth to Non-public HCC options

The Data

Changing Population Age Demographics

- Population projection data provided by BC Stats.

Changing Relationship of Age to Health

- Historical HCC service data provided by BC Ministry of Health.

Changing Wealth Demographics

Relationship of Wealth to Non-public HCC options

The Data

Changing Population Age Demographics

- Population projection data provided by BC Stats.

Changing Relationship of Age to Health

- Historical HCC service data provided by BC Ministry of Health.

Changing Wealth Demographics

- Income tax data provided by Stats Canada.

Relationship of Wealth to Non-public HCC options

The Data

Changing Population Age Demographics

- Population projection data provided by BC Stats.

Changing Relationship of Age to Health

- Historical HCC service data provided by BC Ministry of Health.

Changing Wealth Demographics

- Income tax data provided by Stats Canada.

Relationship of Wealth to Non-public HCC options

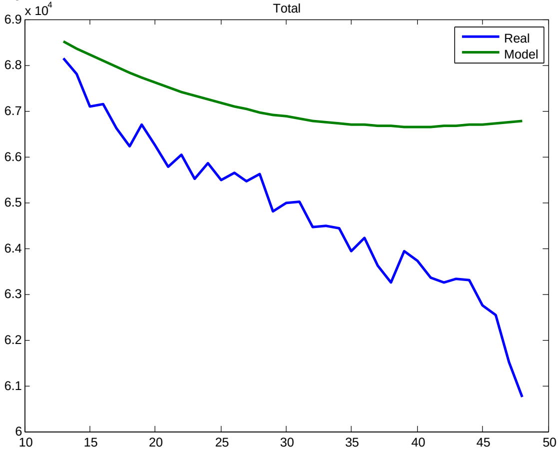
- **No data available!**

Modelling Challenge

- The result was 13 unknown parameters values that could not be approximated using data.
- These parameter control how Population Wealth Distribution and Private HCC impact HCC service usage
- If these parameters are set to zero, the model assumes that nobody ever uses private HCC

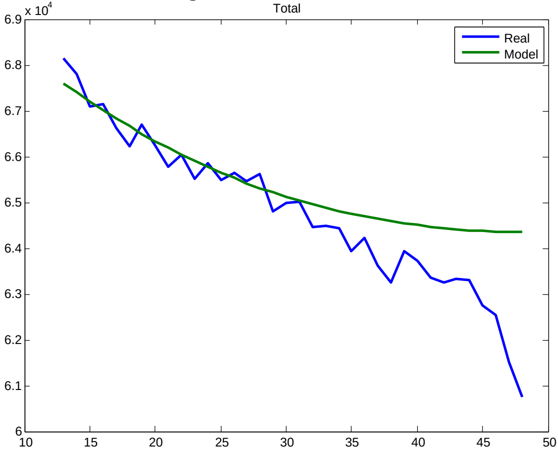
Baseline Model Fit

Model Fit if all parameters are zero



Best Guess Model Fit

Model Fit based on educated guess



Math

Let us consider the problem as follows

- Let $x_1, x_2, ...x_{13}$ represent the 13 unknown parameters
- Let $f(x)$ map a parameter selection to its resulting fit

Then we seek to locate a solution to

$$\max_{x}\{f(x)\}$$

such that

$$x = (x_1, x_2, ...x_{13}) \text{ are acceptable parameters}$$

Math

Let us consider the problem as follows

- Let $x_1, x_2, ...x_{13}$ represent the 13 unknown parameters
- Let $f(x)$ map a parameter selection to its resulting fit

Then we seek to locate a solution to

$$\max_{x}\{f(x)\}$$

such that

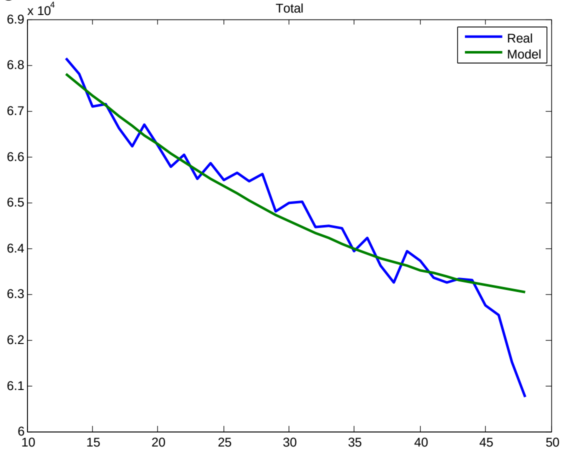
$$x = (x_1, x_2, ...x_{13}) \text{ are acceptable parameters}$$

In this case:

$$\begin{aligned}
&x_i \geq 0 \text{ for all } i\\
&x_3 \leq 21000\\
&x_4 \leq x_5 \leq x_6 \leq x_7 \leq x_8 \leq 1\\
&x_9 \leq 0.75\\
&x_{10}, x_{11}, x_{12}, x_{13} \leq 1
\end{aligned}$$

DFO Model Fit

Model Fit using DFO



Other examples

Other examples

DFO has been used in

- *Finance* to Optimize financial trading strategies

  [De Santis, Dellepiane, Lucidi, Renzi, 2020]
- *Computer science* to Develop worst case analysis of circuit design

  [Latorre, Habal, Graeb, Lucidi, 2018]
- *Electrical engineering* to Optimize on-site renewable energy storage

  [González-Mahecha, Lucena, Szklo, Ferreira, Vaz, 2018]
- *Biology* to Understand dynamic energy budget models

  [Morais, Custódio, Marques, 2019]
- *Hydrology* to Calibrate high detail models

  [Pierre-Luc, Poulin, Audet, Alarie, 2019]

and much more

Types of Optimization Problems

Classifications

Optimization problems fall into categories

- helps researchers think about how to design algorithms
- helps users determine what algorithms to use

Level 1: Continuous or Discrete

Definition: An optimization problem is **continuous** if the *constraint space* allows for continuous *selection of variables*
Otherwise, the optimization problem is **discrete**

Example: Let $f(x) = \begin{cases} 1 & x \in \mathcal{Z} \\ 0 & \text{otherwise} \end{cases}$
Then $\min\{f(x) : x \in \mathbb{R}\}$ is *continuous*

Example: Let $f(x) = x^2$
Then $\min\{f(x) : x \in \mathcal{Z}\}$ is *discrete*

Level 2: Structure attributes

Continuous optimization problems are further subdivided

- Convex or nonconvex
- Smooth or nonsmooth
- Stochastic or deterministic
- etc...

Some structures are so common they get special names

- Least Squares
- Linear Programming
- ℓ_1-Regularization
- etc...

DFO

Most DFO methods assume *continuous* optimization problems

Assignment 1

Assignment 1

MATH 462, COSC 419K, MATH 562

1. Express the following system of linear equations in matrix form and solve the system

$$\begin{array}{rcl} 7x_1 + 2x_2 - x_3 & = & 5 \\ -x_1 - x_2 + 5x_3 & = & 9 \\ 2x_1 + 2x_2 - 9x_3 & = & -16 \end{array}$$

2. Let $f(x) = \sin(x_1 x_2) - (x_1)^3 \exp((x_2)^2)$. Find the gradient f at $(\pi, 0)$.

MATH 462

- Textbook #1.2, 1.3, 1.4

COSC 419K

- Textbook # 1.2(abc), 1.8

MATH 562

- All MATH 462 and COSC 419K questions
- Textbook # 1.5, 1.9