Proposal for Summer Undergraduate Research Award 2023

Data Envelopment Analysis for the Performance of Various Geographical Regions in Fighting COVID-19

Facilitator:Prof.Aparna Mehra

Department Head:Prof.Aparna Mehra

Authors

Rakshitha
2021MT10904
CGPA:8.78
6360845582
mt1210904@iitd.ac.in

Ayush Gupta
2021MT10697
CGPA:9.53
9903899866
mt1210697@iitd.ac.in

Introduction

With the November 2019 outbreak of **COVID-19**, there was a considerable surge in hospital admissions worldwide. The symptoms included fever, cough, difficulty breathing, and other respiratory symptoms that were routinely treated in intensive care units (ICUs). The majority of governments were driven to respond fast in response to a dramatic surge in demand for the medical systems and infrastructure of the nations. **Due to disparities in political power, legislative framework, and health system capacity, different nations responded differently to this circumstance (both demographic and economic).**

Based on their lasting health infrastructure, resource distribution, and availability of resources, many geographical regions responded differently to the circumstance. In our nation, the central government handled India's campaign against COVID-19, while state governments had great leeway in allocating resources and addressing the situation on their own fronts. **Owing to the lack of recent similar experiences and the short reaction time, the bulk of responses were disorganised and ineffectual, leaving us with a wealth of valuable lessons.**

The pandemic resulted in several research advancements. Some of them were based on the mathematical study of the pandemic's dynamics and the optimization of resource use during the COVID era.

Data Envelopment Analytics(DEA) is a mathematical method that may be used to examine and optimise the relative efficiency of similar **Decision-Making Units(DMUs)** utilising a collection of inputs and outputs categorised as performance indicators. **DEA creates an efficient frontier by analysing the inputs and outcomes.** If a DMU falls on the efficiency frontier, it is efficient; otherwise, it is inefficient. Two fundamental DEA models are **Charnes Cooper and Rhodes (1978) (CCR) and Banker Chames and Cooper (1984) (BCC)**. DEA is widely used in the banking industry because it can deal with many outputs of varying types and does not need an a priori production function. It has also been used to a number of other sectors (e.g., insurance, agriculture, supply chain, transportation, public policy and mutual funds).

Inverse data envelopment analysis (InvDEA) is utilised to estimate output levels corresponding to changing input levels or vice versa while maintaining a predetermined degree of efficiency. **Whereas DEA calculates the (inputs and relative) technical efficiency of each project when the outcomes are known, InvDEA estimates the inputs required for similar projects to maintain DEA efficiency when only the (intended) results are known.**

Objectives

- Some research and analysis has already been done to construct models for examining the efficiency of government systems during the pandemic. The following input and output parameters have been taken into consideration in the models that we have come across(citations provided in references):

  1. **Input Parameters**:
     a. Costs of medical equipment- includes COVID-19 diagnostic testing, disinfection and sterilisation products, oxygen therapy equipment
     b. Number of affected people(number of infections)
     c. Health infrastructure - includes number of hospitals, number of hospital beds, ICU units
     d. Health professionals - including doctors, nurses
     e. Population density
  2. **Output Parameters**:
     a. Recoveries - including home recoveries, hospital recoveries and in specific ICU recoveries
     b. Number of deaths(negative output)

  Alongside these parameters, we notice that a few more parameters can be included which might make a significant difference to the analysis of data.

  1. **Additional input parameters**:
     a.Domestic and international funds received by the states
  2. **Additional output parameters**:
     a.Utilisation and allocation of the received funds
     b.Medical delays due to COVID-19 - number of non-COVID deaths during the interval (We notice that a lot of non-COVID affected patients did not receive appropriate and required medical attention due to unavailability of medical equipment and medical practitioners, which led to further loss of lives.)

  Our first objective would be to try and incorporate these modifications into the existing models.

  Remark: The additional parameters are not exhaustive, and we hope to find more such parameters as we work through the project.

- Our second objective would be to **collect data and analyse the available data and rank the Indian states on the basis of efficiency** of their response to the COVID-19 outbreak. Simultaneously, we would try to provide suggestions based on the mathematics of our analysis to indicate what sectors need more work for better efficiency.

  Remark: We also suspect that there is an uneven distribution of resources among the states of India. Hence, based on our analysis, we will also try to **look into the possibility of mutual exchange of resources between states to attain a more efficient working system**.

- Our third objective would be to reverse the process. In the case that we are faced with a similar pandemic in future, and we have resources more or less stagnant at the present numbers, **we wish to analyse the amount of change in input that would be required to attain a higher efficiency for a given output.**

  Remark: There is also a possibility to include uncertainty and errors in our analysis by adopting Stochastic Models for the same dataset.

Methodology

Step 1: **Data Collection and Data Cleaning**

The data for various countries around the world and for all the Indian States will be collected and any unnecessary information will be erased. There will be a need to normalise this data to account for discrepancies in the order of magnitude of the various parameters.

Remark: Existing Models of Optimisation:

(a) **Slack Based Model**: Slack Based Model is a non-radial additive DEA model which defines the additive combination of inputs and outputs to achieve virtual outputs and input data.

Let J={1,2,...,n}, I={1,2,...n} and R={1,2,...s} stand for the index set of observation(DMUs), input and output respectively. In other words, we assume that there are n DMUs that each DMU_j produces s different positive outputs $y_{rj}(r \in R)$ using m positive inputs $x_{ij}(i \in I)$. The non-negative decision variables s_i^- and s_r^+ are the **ith input slack(excess)** and **rth output slack(shortfall)** respectively. λ_j is the intensity variable corresponding to DMU_j.

Here we present the envelopment form of SBM:

$\theta_o^{SBM} = min\rho - \frac{1}{m} * \Sigma_{i\in I}\frac{s_i^-}{x_{io}}$
such that
$\rho + \frac{1}{s} * \Sigma_{r\in R}\frac{s_r^+}{y_{ro}} = 1$
$\Sigma_{j\in J}\lambda_j * x_{ij} + s_i^- = \rho * x_{io} \qquad \forall j$
$\Sigma_{j\in J}\lambda_j * y_{rj} - s_r^+ = \rho * y_{ro} \qquad \forall r$
$\Sigma_{j\in J}\lambda_j = \rho$
$\lambda_j, s_r^+, s_i^- \geq 0 \qquad \forall j, \forall r, \forall i$

Note: DMU_o is *SBM-efficient* if $\theta_o^{SBM} = 1$;otherwise it is *SBM-inefficient*.

(b) **Range Directional Based Model**: There are two variants of this Model. One is used for cases where targets are sought to improve those variables where the DMU is furthest from best attainable levels while the second, is for cases where improvement is prioritised for variables where the DMU is closest to best attainable levels.

Step 2: **Study of Existing Models and Modification to Account for Negative Parameters**

The existing optimisation models deal majorly with non-negative data parameters. In effect, most of these models focus on maximising output while minimising input to increase efficiency. However, in our case, the most important output parameter is the death rate, which is an undesirable output and is required to be minimised. For the same, we will have to make appropriate changes to the optimisation model. There is no standard model for handling negative data and we wish to discover the best possible one as we explore this topic in depth during our first 2 weeks.

Step 3: **Running data sets through the Modified Models**

We plan to use the slack-based model primarily for our analysis. The reason is, that this model not only helps in analysing the efficiency of the system, but also provides an indication as to what exactly is the parameter causing any reduction in efficiency, and how it can be mathematically improved. All the data sets will be run through this model with appropriate tweaks using a programming-based approach(majorly using Python and MATLAB).

Step 4: **Analysis of Data**

Obtained results will be analysed to progress with objective-1. The regions will be ranked on the basis of their efficiency and performance during COVID-19. Here we plan to conduct analysis on two data sets, one for the comparison of states of India and another for the comparison of countries around the world. We also intend to provide mathematical suggestions to improve the efficiency of performance for any future use.

Step 5: **Running the Inverse DEA Process**

Given we are put up against a similar situation in future and the other conditions remain more or less stagnant, we would like to provide mathematical suggestive measures to attain higher desirable efficiency(performance).

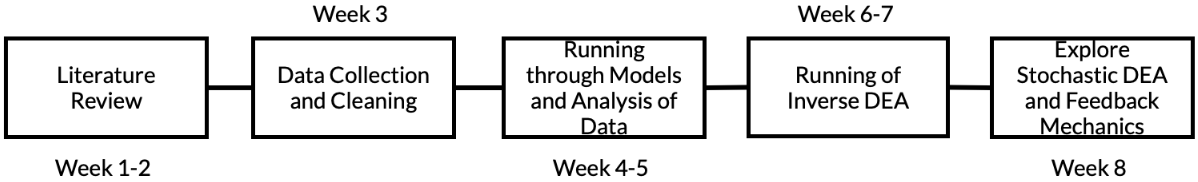
(a) Input given: Optimise output so as to attain desirable higher efficiency.

(b) Output given: Optimise input so as to attain desirable higher efficiency.

Step 6: **Include Stochastic Modeling**

If time permits, we intend to incorporate uncertainty into our data and use stochastic DEA to predict the DMU's efficiency on a weekly or monthly basis.

Timeline



Budget

Our project will require a well-configured workstation and a good deal of health sector database. We will not require any extra funding for the purpose of this project.

References

1. **Data Envelopment Analysis**

   - *Kaoru Tone* A slacks-based measure of efficiency in data envelopment analysis(European Journal of Operational Research 130 (2001) 498-509)
   - *Kshitish Kumar Mohanta, Deena Sunil Sharanappa, Abha Aggarwal* Efficiency analysis in the management of COVID-19 pandemic in India based on data envelopment analysis(Current Research in Behavioral Sciences 2 (2021) 100063)
   - *Nahia Mourada, Ahmed Mohamed Habibb and Assem Tharwata* Appraising healthcare systems' efficiency in facing COVID-19 through data envelopment analysis(Decision Science Letters 10 (2021) 301–310)

2. **Inverse Data Envelopment Analysis**

   - *Sabri Boubaker, Tu D.Q. Le and Thanh Ngo* Managing bank performance under COVID-19: A novel inverse DEA efficiency approach(International Transactions in Operational Research (2022) 1–17)
   - *Ali Ghomi, Saeid Ghobadi, Mohammad Hassan Behzadi and Mohsen Rostamy-Malkhalifeh* Inverse Data Envelopment Analysis with Stochastic Data(RAIRO Operations Research 55 (2021) 2739–2762)

3. **Stochastic Data Envelopment Analysis**

   - *Susan X. Li* Stochastic models and variable returns to scales in data envelopment analysis(European Journal of Operational Research 104 (1998) 532-548)
   - *Jati K. Sengupta* Data Envelopment Analysis for Efficiency Measurement in the Stochastic Case(Computers & Operations Research 14 (1987) 117-129)