Aplicación de Bayesian Optimization over Permutation Spaces with Heuristics para determinar la estrategia de vacunación óptima contra el COVID-19 a partir de que se dispusieran dosis de la vacuna en EEUU

Gilberto González-Parra
Department of Mathematics
New Mexico Tech
Socorro, NM, USA
Gilberto.GonzalezParra@nmt.edu

F.J. Navarro-González
Departament of Applied Mathematics
University of Alicante
Alicante, Spain
francisco.navarro@ua.es

Juan Carlos Cortés, Rafael J. Villanueva
Instituto de Matemática Multidisciplinar
Universitat Politècnica de València
Valencia, Spain
jccortes@imm.upv.es,rjvillan@imm.upv.es,

Josu Ceberio
Department of Computer Science and Artificial Intelligence
Universidad del País Vasco UPV/EHU
Donostia, Spain
josu.ceberio@ehu.eus

Abstract

Esta Key Work es parte del trabajo publicado en González-Parra et al. [2024] donde se construye un modelo basado en agentes para el estudio de la dinámica de transmisión del SARS-CoV-2 en Estados Unidos al principio de declararse la pandemia de COVID-19. Una vez diseñado el modelo, se calibra con datos de casos reportados y muertes de antes de que estuviera disponible la vacuna. A partir de aquí, se utiliza la técnica Bayesian Optimization over Permutation Spaces with Heuristics (BOPS-H) para determinar cuál hubiera sido la mejor estrategia de vacunación para evitar el máximo número de muertes.

1 Presentación del problema

En González-Parra et al. [2024] se presentó un modelo basado en agentes para EEUU que describía la dinámica de transmisión del SARS-CoV-2, donde se discriminaba a los individuos en 5 grupos de edad y en cada grupo de edad, los que tenían comorbilidades o los que no tenían, total 10 grupos (tabla 1). Asimismo, considerábamos a los individuos que aceptaban vacunarse y a los que no.

Una vez calibrado el modelo, teníamos que evaluar qué estrategia de vacunación podría haber sido la que evitara más muertes, teniendo en cuenta que, al principio, el suministro de vacunas era limitado. Una estrategia de vacunación consiste en asignar un orden a los diferentes grupos indicando la precedencia en el proceso de vacunación. Así, cada estrategia puede identificarse mediante una permutación del mismo orden que el número de grupos.

ID	Grupo de edad	ID	Grupo de edad
1	0-39 años, no comorb	6	0-39 años, si comob
2	40-59 años, no comorb	7	40-59 años, si comob
3	60-69 años, no comorb	8	60-69 años, si comob
4	70-79 años, no comorb	9	70-79 años, si comob
5	80+ años, no comorb	10	80+ años, si comob

Table 1: Grupos de edad y su codificación.

De esta forma, planteamos la vacunación por grupos de edad y por comorbilidad, de donde el espacio de búsqueda es $10! = 3\ 628\ 800$ combinaciones posibles. Disponemos de un ordenador con 64 núcleos en 8 procesadores Xeon Sandy Bridge E5-4620 funcionando a 2,2 GHz, con 16 MB de memoria caché y 512 GB de RAM y sistema operativo Ubuntu Server 20.04.6 LTS. Cada simulación del modelo es costosa desde el punto de vista computacional, ya que cada una necesita aproximadamente 3 minutos. Por lo tanto, realizar sólo una para cada una de las $3\ 628\ 800$ posibles estrategias de vacunación costaría unos $17\ 640$ días en un ordenador estándar y con nuestro clúster (64 núcleos) funcionando en modo paralelo como el que disponemos, tardaría aproximadamente 276 días.

La optimización bayesiana (BO) es un método probabilístico utilizado para optimizar funciones complejas y, por lo general, costosas de evaluar. Utiliza un modelo probabilístico sustitutivo de la función objetivo que se actualiza a medida que se realizan nuevas evaluaciones. BO se aplica de forma natural a problemas en dominios continuos. Sin embargo, existen investigaciones dedicadas a extender su aplicabilidad a dominios discretos, como los trabajos de Hutter et al. [2011],Wang et al. [2016], Baptista and Poloczek [2018] sobre el dominio binario y en particular Deshwal et al. [2022] sobre espacios de permutaciones utilizando núcleos de Mallows Jiao and Vert [2015]. La presente investigación utiliza una adaptación del algoritmo BOPS-H de Deshwal et al. [2022].

2 Resultados

Tomando una población inicial de 20 individuos y limitando el número de iteraciones máximas a 100, se realizan varias repeticiones del algoritmo de optimización (96 o 128, para aprovechar la arquitectura de 64 núcleos disponible y correr las repeticiones en paralelo), pudiendo estudiarse la evolución de diferentes parámetros estadísticos de las ejecuciones y su evolución a lo largo de las iteraciones. Los principales resultados se muestran a continuación.

En primer lugar, en la Figura 1 se puede comprobar el comportamiento de los errores absolutos en cada paso (el menor de los candidatos encontrados hasta ese momento) para las diferentes repeticiones.

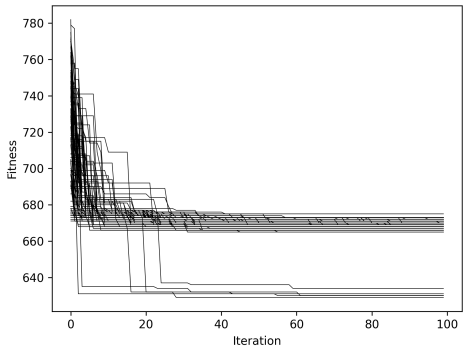
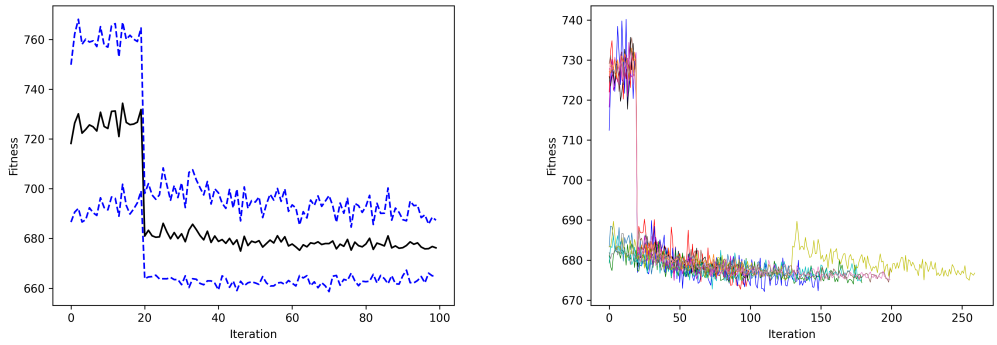


Figure 1: Evolución de los mínimos en cada iteración para las repeticiones BOPSH.

Por otro lado, en la Figura 2a se puede observar que la disminución del valor medio de la función objetivo se produce básicamente a partir de la población inicial (salto en la iteración 20), por lo que las evoluciones posteriores tienen, en promedio, un comportamiento prácticamente estacionario (lo

que no quiere decir que en algún paso no se pueda encontrar un mínimo inferior a cualquiera de los encontrados anteriormente).



(a) Promedio, máximos y mínimos en cada iteración.

(b) Promedio en cada iteración.

Figure 2: Evolución de promedios y otros estadísticos en cada iteración.

La magnitud de las oscilaciones de los promedios para las diferentes repeticiones, en la Figura 2b nos da una idea de la "amplitud" de la búsqueda realizada en cada paso por el algoritmo BOPSH en la búsqueda del óptimo.

Las 4 mejores estrategias de vacunación encontradas mediante el algoritmo BOPS-H, es decir, la que menos muertes procurarían, corresponden a los vectores $[7, 9, 1, 5, 4, 10, 2, 3, 8, 6]$, $[7, 9, 1, 5, 4, 10, 2, 3, 6, 8]$, $[7, 9, 1, 5, 4, 10, 2, 6, 3, 8]$, $[7, 9, 1, 5, 4, 10, 3, 6, 2, 8]$.

Como puede verse, los 6 primeros grupos coinciden en las 4 estrategias, esto es, la prioridad de vacunación debe darse a los grupos de edad de 40-59 y 60-69 años, ambos con comorbilidades. Esto se debe al hecho de que estas personas tienen comorbilidades y al mismo tiempo una cantidad significativa de contactos al día. Se trata de personas en edad laboral, lo que se ha considerado importante en otros estudios relacionados con la vacunación óptima El siguiente grupo prioritario es el grupo de 0-39 años sin comorbilidades, ya que tienen un gran número de contactos.

References

Ricardo Baptista and Matthias Poloczek. Bayesian optimization of combinatorial structures. In *International conference on machine learning*, pages 462–471. PMLR, 2018.

Aryan Deshwal, Syrine Belakaria, Janardhan Rao Doppa, and Dae Hyun Kim. Bayesian optimization over permutation spaces. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 6515–6523, 2022.

Gilberto González-Parra, Javier Villanueva-Oller, Francisco J Navarro-González, Josu Ceberio, and Giulia Luebben. A network-based model to assess vaccination strategies for the covid-19 pandemic by using bayesian optimization. *Chaos, Solitons & Fractals*, 181:114695, 2024.

Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *Learning and intelligent optimization: 5th international conference, LION 5, rome, Italy, January 17-21, 2011. selected papers 5*, pages 507–523. Springer, 2011.

Yunlong Jiao and Jean-Philippe Vert. The kendall and mallows kernels for permutations. In *International Conference on Machine Learning*, pages 1935–1944. PMLR, 2015.

Ziyu Wang, Frank Hutter, Masrour Zoghi, David Matheson, and Nando De Feitas. Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016.