MA 5701 Optimización no lineal. 2021-1

Integrantes: Sergio López, Juan Pablo Cabeza

Tarea 1

4 de julio 2021

P1. Los infectados acumulados ('casos totales' en la base de datos) y recuperados acumulados ('casos confirmados recuperados' en la base de datos) en función del día se muestran en el siguiente gráfico

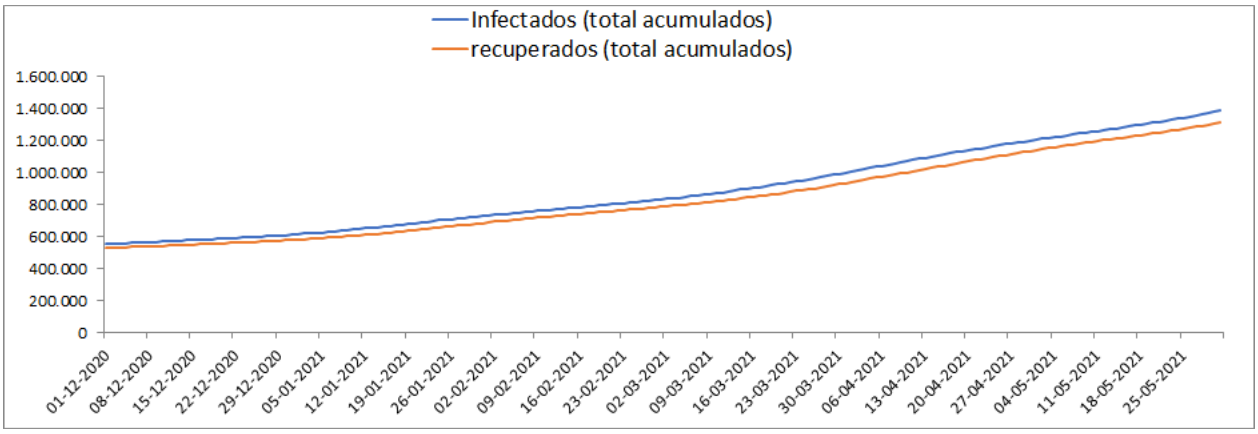

Figura 1: Numero de infectados y recuperados acumulados desde el 1 de diciembre 2020 hasta el 31 de mayo 2021.

Observamos que, en cualquier intervalo de tiempo, la tasa de crecimiento del total de infectados acumulados (i.e. la pendiente de la curva azul de la Figura 1) es muy similar a la del total de recuperados acumulados. Para determinar la variación de dicha tasa de crecimiento a través del tiempo, graficamos el numero de nuevos infectados ('casos nuevos totales' en la base de datos) y recuperados por día:

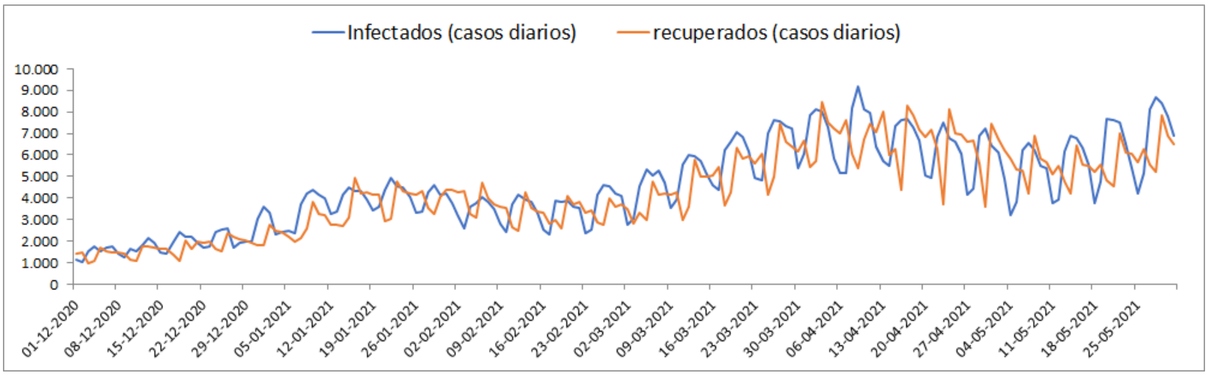

Figura 2: Numero de infectados y recuperados por día desde el 1 de diciembre 2020 hasta el 31 de mayo 2021.

Notamos que, desde el 1 de diciembre hasta el 22 de enero aproximadamente, la tasa de crecimiento de infectados y recuperados acumulados aumenta, lo que implica que las curvas de la Fig. 1 son convexas en el periodo de tiempo mencionado anteriormente. Luego, dicha tasa disminuye levemente hasta el día 17 de febrero aproximadamente, es decir, las curvas de la Fig. 1 son cóncavas en dicho periodo. Posteriormente, la tasa de crecimiento de infectados acumulados aumenta hasta llegar a su valor máximo en el día 9 de abril, mientras que la tasa de crecimiento de recuperados acumulados alcanza su valor máximo en el día 4 de abril. Después, la tasa de infectados y recuperados acumulados disminuye hasta el día 4 de mayo aproximadamente y finalmente aumenta hasta el 31 de mayo.

Dado que el numero de muertos tiene un orden bastante menor que el numero de infectados y recuperados, estudiamos su comportamiento por separado.

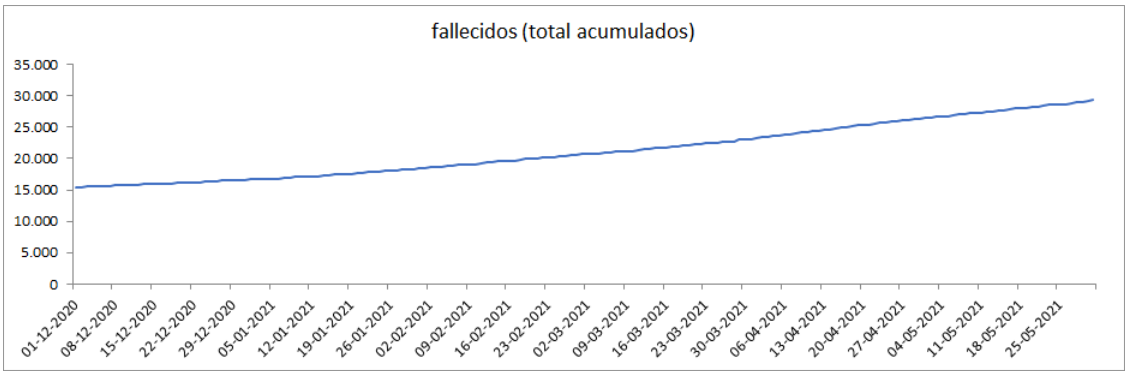

Figura 3: Numero de muertos acumulados desde el 1 de diciembre 2020 hasta el 31 de mayo 2021.

En este caso, el numero de muertos por día se ve como sigue:

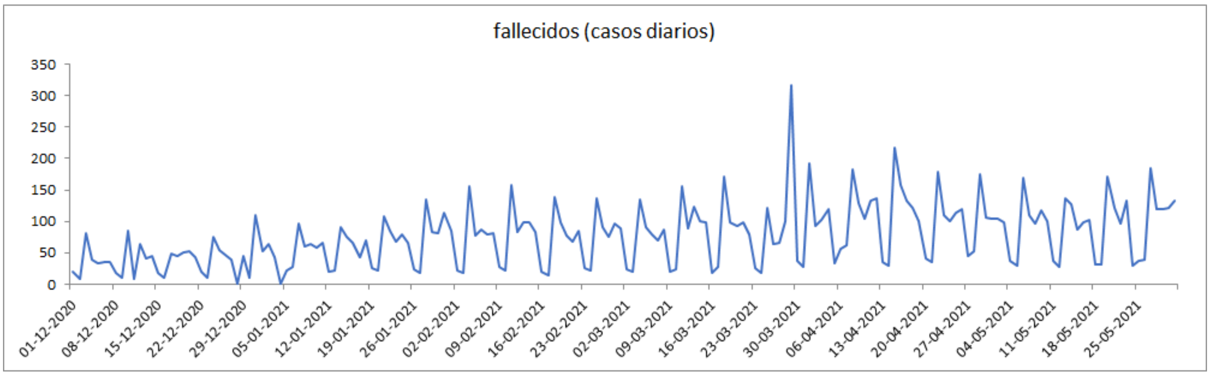

Figura 4: Numero de muertos por día desde el 1 de diciembre 2020 hasta el 31 de mayo 2021.

Vemos que la cantidad de muertos por día no posee variaciones considerables durante todo el periodo de tiempo estudiado, excepto el día 29 marzo, en donde el numero de muertos (316) es bastante mas alto que lo otros días (el segundo día con mas muertos posee 218 muertos).

Para terminar, veamos el comportamiento de la tasa de mortalidad (muertos acumulados/infectados acumulados) en función del día:

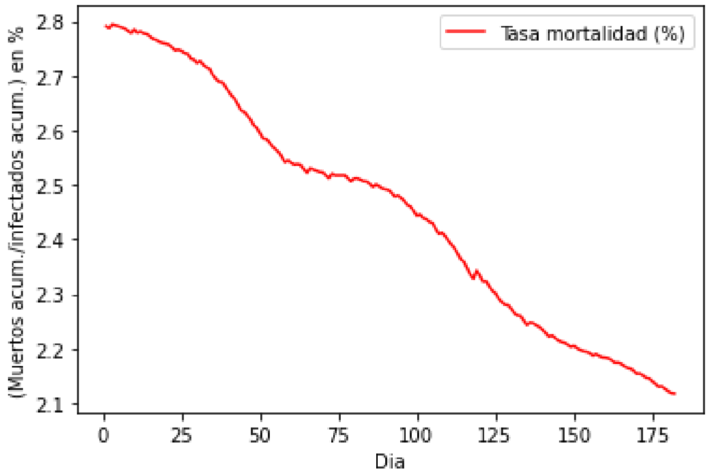

Figura 5: Tasa de mortalidad en porcentaje en cada día, desde el 1 de diciembre 2020 (Día 0) hasta el 31 de mayo 2021 (Día 181).

Observamos un decrecimiento en la tasa de mortalidad en el periodo de tiempo estudiado, llegando hasta un $2,12\,\%$ en el día 31 de mayo.

P2. *a*) Para discretizar el modelo, particionamos el intervalo de tiempo estudiado en la cantidad de dias ($n := 182$) de tal manera que para cada $j \in \{0, \dots, n-1\}$, definimos

$$\begin{aligned} I_j &:= \text{N° de infectados en el día } j \\ R_j &:= \text{N° de recuperados en el día } j \\ D_j &:= \text{N° de muertos en el día } j \\ A_j &:= \text{N° de asintomáticos en el día } j \\ E_j &:= \text{N° de expuestos en el día } j \\ S_j &:= \text{N° de susceptibles en el día } j \\ \alpha_a(j) &:= \text{tasa de distanciamiento social en el día } j \\ \alpha_i(j) &:= \text{tasa de cuarentena en el día } j \\ \kappa(j) &:= \text{Proporción de test en el día } j \end{aligned}$$

donde el día 0 es el 1 de diciembre 2020 y el día 181 es el día 31 de mayo 2021. Luego, $\dot{S} = S_{j+1} - S_j$ y así para cada variable. Así, considerando que $(\eta, \rho, \gamma) = \left(\frac{1}{2}, \frac{1}{10}, 0\right)$, el modelo SEAIRD discretizado queda

$$(\forall j \in \{0, \dots, n-2\}) \quad \left\{ \begin{array}{c} S_{j+1} = \left(1 - \dfrac{\alpha_a(j)A_j + \alpha_i(j)I_j}{N}\right) S_j \\ E_{j+1} = 0,5E_j + \left(\dfrac{\alpha_a(j)A_j + \alpha_i(j)I_j}{N}\right) S_j \\ A_{j+1} = (0,9 - \kappa(j))A_j + 0,5E_j \\ I_{j+1} = (1 - \beta - \mu)I_j + \kappa(j)A_j \\ R_{j+1} = R_j + 0,1A_j + \beta I_j \\ D_{j+1} = D_j + \mu I_j \end{array} \right. \tag{0.1}$$

donde $N = 19107000$ es la población total.

b) Haremos las siguientes suposiciones de acuerdo a las recomendaciones para la calibración:

- Para cada $j \in \{0, \ldots, n-1\}$ consideramos $\kappa(j) = \dfrac{\hat{A}_j}{\hat{I}_j}$ donde $\hat{A}$ son los casos activos asintomáticos (infectados no confirmados) y $\hat{I}$ son los casos activos confirmados, obtenidos de la base de datos. Es decir, consideramos al parámetro κ como un valor conocido.
- Las funciones α_a y α_i son constantes en cada periodo de 7 días. Es decir, como $n/7 = 26$, los vectores $\{\alpha_a(j)\}_{j=0}^{n-1}$ y $\{\alpha_i(j)\}_{j=0}^{n-1}$ deben satisfacer que

$$(\forall k \in \{0, \ldots, 25\})(\forall j \in \{1, \ldots, 6\}) \quad \begin{cases} \alpha_a(7k+j) = \alpha_a(7k) \\ \alpha_i(7k+j) = \alpha_i(7k) \end{cases}$$

Usaremos $\hat{I}$, $\hat{R}$ y $\hat{D}$ como los 'casos activos confirmados', 'casos confirmados recuperados' y 'fallecidos' (acumulados) en la base de datos, respectivamente.

Los parámetros iniciales para resolver el problema serán:

$$\mu_0 = \frac{\hat{D}_1 - \hat{D}_0}{\hat{I}_0}$$

$$\beta_0 = \frac{\hat{R}_1 - \hat{R}_0 - 0{,}1\hat{A}_0}{\hat{I}_0}$$

$$\alpha_a(0) = \alpha_i(0) = 0,5$$

Además, para las variables A, I, R y D, los valores iniciales para el modelo (0.1) serán $I_0 = \hat{I}_0$, $R_0 = \hat{R}_0$, $D_0 = \hat{D}_0$ y $A_0 = \hat{A}_0$. El valor inicial para la variable E estará dado por

$$E_0 = \frac{\hat{A}_1 - (0,9 - \kappa(0))\hat{A}_0}{0,5}.$$

Definamos $\delta_0 = \dfrac{\alpha_a(0)\hat{A}_0 + \alpha_i(0)\hat{I}_0}{N}$ y $\hat{E}_1 = \dfrac{\hat{A}_2 - (0,9 - \kappa(1))\hat{A}_1}{0,5}$. Entonces el valor inicial para la variable S estará dado por $S_0 = \dfrac{\hat{E}_1 - 0{,}5E_0}{\delta_0}$.

Por ultimo, dado que la idea es tener la mayor cantidad de recuperados posibles consideramos $P_R = 0$. Por otro lado, dado que es mejor tener menos muertos, consideramos $P_D = 2/3$ y $P_I = 1/3$. Resolveremos el problema usando dos metodos: la funcion `dual_annelaing` y la funcion `minimize` con el metodo 'trust-constr'. Ambas funciones usan el paquete `scipy.optimize` de `python`.

c) Considerando los parámetros iniciales dados en el ítem anterior, los valores de β y μ encontrados son los siguientes:

Método	μ	β
trust-constr	0,002779	0.07234
dual_annelaing	0,002777	0.06722

Cuadro 1: soluciones del problema asociado a la calibración para cada método usado.

La implementación del método 'trust-constr' se basa en un algoritmo para optimización con restricciones de igualdad a gran escala.

El método de optimización 'dual annelaing' es un método de optimización global que es una versión modificada del método 'simulated annealing', el cual utiliza un algoritmo de búsqueda local.

Los resultados en la Tabla 1, muestran que hay una tasa de fatalidad (muertos diarios/casos activos confirmados) del $0,278\,\%$ aproximadamente.

La comparación de los casos activos confirmados de la base de datos con respecto a las soluciones encontradas por cada método se resumen en el siguiente gráfico:

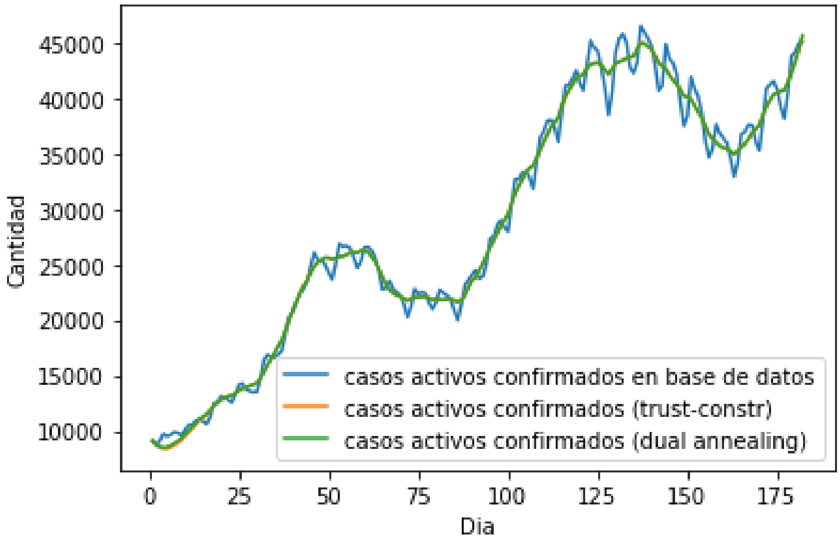

Figura 6: casos activos confirmados por día, desde el 1 de diciembre 2020 (Día 0) hasta el 31 de mayo 2021 (Día 181).

Observamos que la soluciones entregadas por cada método son muy similares y ambas dan una aproximación bastante certera a los datos reales.

Con respecto a los recuperados confirmados acumulados, tenemos el siguiente comportamiento por día:

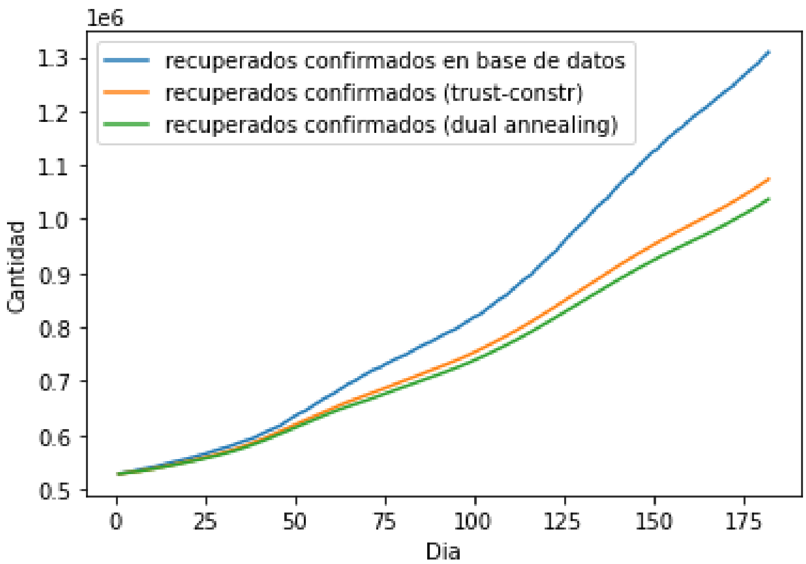

Figura 7: casos confirmados recuperados por día, desde el 1 de diciembre 2020 (Día 0) hasta el 31 de mayo 2021 (Día 181).

En este caso, vemos que los métodos tienen una diferencia bastante considerable con respecto a los datos reales a partir del día 50 aproximadamente. Por otro lado, el método dual annealing entrega una menor cantidad de recuperados que el método trust-constr.

Finalmente, mostramos el comportamiento de los muertos acumulados en función del tiempo:

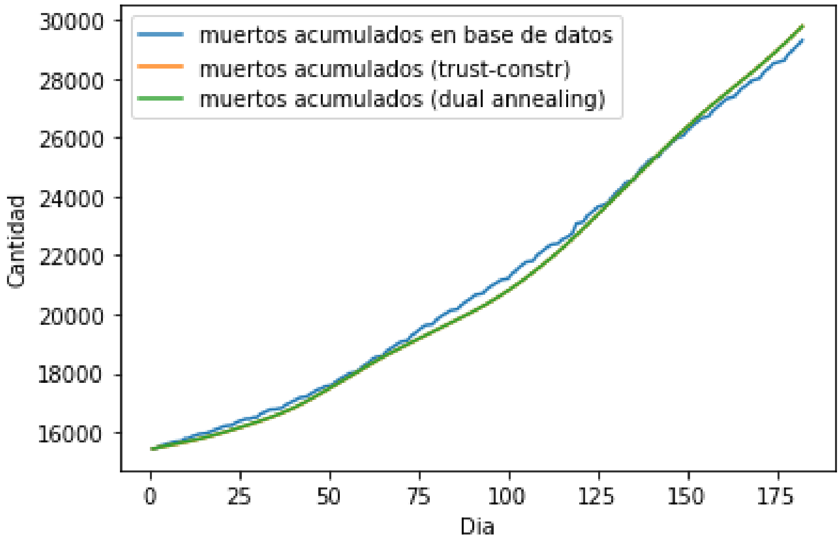

Figura 8: muertos acumulados desde el 1 de diciembre 2020 (Día 0) hasta el 31 de mayo 2021 (Día 181).

Notamos que sucede algo similar al caso de los activos confirmados, es decir, ambos métodos entregan prácticamente la misma solución y existe una leve diferencia de los métodos con respecto a los datos de la base de datos.

P3. *a*) Para discretizar el problema, particionaremos el intervalo $[t_0, t_f]$ en la cantidad de días, donde consideramos el periodo comprendido entre el día 01 de mayo y el día 30 de septiembre (es decir, $n := 153$ días). Luego, definimos $t_0 = 0$, $t_f = n - 1 = 152$ y para cada $j \in \{1, \ldots, n-1\}$

$$\alpha_a(j) =$$