

MA3711 Optimización Matemática. Semestre Otoño 2020

Profesor: Héctor Ramírez C. Auxiliar: Rolando Rogers, María José Alfaro, Javier Madariaga

Tarea 1

Fecha entrega: Viernes 8 de Mayo de 2020

1. Problemas

El mundo enfrenta en estos días su peor pandemia desde la gripe española. Un virus del tipo SARS, conocido como **Covid-19**, se ha extendido rápidamente en el mundo desde su primera aparición en Wuhan, China, en diciembre del 2019. En Chile, el primer caso se confirmó el 3 de marzo de 2020, cuando un médico proveniente de Singapur fue internado en el Hospital Regional de Talca, Región del Maule. A la fecha, todas las regiones de Chile ya reportan contagiados, siendo la Región Metropolitana la que reporta la gran mayoría de ellos. Esto ha llevado al cierre de colegios y universidades, así como a cuarentenas localizadas por zonas con una alta cantidad de casos confirmados por superficie. Dada esta rápida evolución, se vuelve fundamental monitorear la dinámica de este brote, y proponer medidas para su mitigación.

En esta tarea, se le pide realizar un análisis matemático simple del crecimiento de los contagios en *Jupyter*. Además, se replicará un estudio hecho en Italia donde se resuelve el problema de encontrar el número de tests que se deben realizar por región para maximizar la capacidad general de detección de la pandemia [3].

P1. Modelando la pandemia en Chile

Uno de los factores más peligrosos del **Covid-19** es su alta tasa de contagiosidad, lo cual se refleja en una alta tasa de infectados. En esta pregunta estudiaremos tres modelos de crecimiento para los contagiados (detectados): el modelo exponencial y el modelo logístico.

El **modelo exponencial** viene dado por la ecuación

$$f(t, a, b, c) = a \cdot \exp(b(t - c))$$

donde la variable t corresponde al tiempo (en días), y los parámetros a, b, c son los ajustes de la curva.

El **modelo logístico** viene dado por la ecuación

$$f(t, a, b, c) = \frac{a}{1 + \exp\left(\frac{-(t-c)}{b}\right)}$$

donde la variable t corresponde al tiempo (en días), y los parámetros a, b, c son los ajustes de la curva.

El **modelo de Gompertz** viene dado por la ecuación

$$f(t, a, b, c) = a \cdot \exp(-b \cdot e^{-ct})$$

donde la variable t corresponde al tiempo (en días), y los parámetros a, b, c son los ajustes de la curva.

En lo siguiente usted deberá:

- 1.- Estudie las bases de datos del Minsal, disponibles como planilla de cálculo en [1]. Obtenga el número de contagiados en Santiago y de otras dos regiones de su elección.
- 2.- Escriba el **problema de optimización** asociado a ajustar la curva de contagios acumulados mediante mínimos cuadrados. Sus variables son $(a, b, c) \in \mathbb{R}^3$.
- 3.- Calibre las distintas curvas en *Jupyter* para los modelos propuesto. Grafique las curvas asociadas tanto para los días usados para calibrar como para un horizonte de 50 días de predicción.
- 4.- Comente los resultados observados. ¿Cuál modelo presentó menor error cuadrático medio? ¿Qué modelo considera que se ajusta mejor a la realidad chilena? Comente las limitaciones que observa para los modelos.

- 5.- A partir de los datos para España disponibles en [2], replique lo anterior para Madrid, con al menos 50 días de datos. ¿Son distintos los resultados?
- 6.- Pregunta extra: Averigüe el significado de conceptos como *número reproductivo básico* (R_0) y el *número reproductivo efectivo* (R_e) y las formas de estimarlos. Estímelos para las comunas que haya elegido en las preguntas anteriores.

P2. Gestión eficaz de los tests de diagnóstico.

Basados en el trabajo de Lorenzo Lampariello y Simone Sagratella en regiones italianas [3], replicaremos su modelo matemático en Chile. Para ello seguiremos los siguientes pasos:

- 1.- Recopile desde [1] los datos necesarios para realizar esta pregunta. Precisar hasta que fecha tomarán datos y realizarán el estudio.
- 2.- Para cada región i de Chile deberá investigar/calcular/estimar los siguientes parámetros:
 - $d_i \in \mathbb{R}_+$ que corresponde al número total de muertes por la enfermedad, acumuladas desde el comienzo del brote.
 - $c_i \in \mathbb{R}_+$, es el número total de casos, es decir, el número total de personas diagnosticadas con la enfermedad, acumulada desde el comienzo del brote.
 - $p_i \in \mathbb{R}_+$ es el tamaño de la población.
 - $\sigma_i \in [0, 1]$ es la tasa de positividad para el test. Por ejemplo, si se hicieron 100 test y 2 fueron positivos, entonces $\sigma_i = 2/100 = 0,02$.
- 3.- Considere $\varphi \in [0, 1]$, que corresponde a la proporción de muertes por la enfermedad y el número total de personas diagnosticadas con la enfermedad durante su brote a la fecha, en el mundo. Basados en los datos disponibles para Chile y considerando un subreporte de casos contagiados de 46 %, calculados en base a la metodología usada en [4], este parámetro tiene un valor de $\varphi = 0,0097$.
- 4.- Por supuesto, hacer grandes cantidades de test implica grandes cantidades de gastos en salud. Es poco realista que se realice un test por habitante. Por ello, estime $b \in \mathbb{R}_+$ el número máximo de test realizables en Chile. Así mismo, estime cotas inferiores l_i y superiores s_i para la cantidad de test realizables en cada región.
- 5.- Calcule para cada región i la cantidad

$$u_i := \max \left\{ 0, \frac{d_i}{\varphi} - c_i \right\},$$

que corresponde a una estimación de las personas enfermas aún no detectadas a partir de los datos disponibles.

- 6.- La variable que queremos calcular corresponde a x_i , para cada región i . Este valor corresponde a la cantidad de tests asignados a cada región.
- 7.- Considere para cada región i las funciones $\mu_i : \mathbb{R} \rightarrow \mathbb{R}$ definidas por:

$$\mu_i(x_i) := \frac{u_i}{p_i}x_i - \frac{\sigma_i}{2p_i}x_i^2.$$

Esta función proporciona una medida del resultado positivo para el diagnóstico de los tests x_i asignadas a la región i .

- 8.- Resuelva en *Jupyter* el problema

$$\max_x \sum_i \mu_i(x_i); \sum_i x_i \leq b, l_i \leq x_i \leq s_i, i = 1, 2, \dots$$

- 9.- El problema anterior, según lo estudiado por [3], nos entrega la cantidad óptima de test a realizar por ciudad. Reporte y comente los resultados obtenidos.

2. Entregables

Deberá subir a U-cursos dos archivos:

- a) Un reporte en formato *PDF* que contenga los modelos propuestos, incluyendo las suposiciones y los desarrollos teóricos requeridos en cada problema. El reporte NO debe ser escrito a mano si no que usando alguna editor de texto apropiado (LaTeX, Word, etc.) y debe ser bien escrito y conciso (no más de 5 páginas). No se deben explicar los códigos en este reporte.
- b) Un archivo comprimido (usando *.zip*) que contenga los archivos *.ipynb* utilizados en la resolución numérica. Debe asegurarse que estos archivos se ejecuten apropiadamente.

- [1] Base de datos de COVID19 en Chile.
<https://github.com/jorgeperezrojas/covid19-data>
- [2] Base de datos de COVID19 en España.
<https://github.com/datadista/datasets>
- [3] Lampariello, L., & Sagratella, S. (2020). *Effectively managing diagnostic tests to monitor the COVID-19 outbreak in Italy*.
- [4] Russell T, Hellewell J, Abbott S, Golding N, Gibbs H, Jarvis CI, et al. Using a delay-adjusted case fatality ratio to estimate under-reporting. CMMID; London School of Hygiene & Tropical Medicine. CMMID Repository. April 2020; cmmid.github.io/topics/covid19/global_crf_estimates.html.