

Modelos *single-index* en el contexto de la regresión cuantil

Álvaro Arrojo Vázquez

Curso 2022-2023

Por motivos de confidencialidad no es posible la publicación de la memoria completa del Trabajo Fin de Máster titulado “Modelos *single-index* en el contexto de la regresión cuantil”, por lo que en el presente documento se incluye un resumen del mismo.

Resumen

Los modelos de regresión en media precisan de condiciones más estrictas para poder ser aplicados y tienen un peor comportamiento ante datos atípicos que los modelos de regresión cuantil. Además, gracias a los modelos de regresión cuantil podemos hacer un estudio más pormenorizado de la relación existente entre las variables explicativas y la variable respuesta, sin estar limitados al estudio de una tendencia central como sucede en el contexto de la regresión en media. Por otro lado, puede que en ocasiones los modelos de regresión paramétricos no sean lo suficientemente flexibles para recoger toda la información de la muestra y en este contexto surgen los modelos de regresión cuantil *single-index*.

En este trabajo se ha realizado una revisión de los modelos de regresión cuantil paramétricos, para posteriormente, en el Capítulo 2, pasar a presentar los modelos de regresión cuantil *single-index*. Dichos modelos se definen de la siguiente manera:

$$Y = m_{\tau}(\theta'_{\tau}X) + \varepsilon_{\tau},$$

con $\mathbb{P}(\varepsilon_{\tau} \leq 0|X) = \tau$. Donde Y representa la variable respuesta, X es un vector de dimensión d que contiene todas las variables explicativas, τ un valor en el intervalo $(0, 1)$, θ_{τ} el vector de parámetros *single-index* y m_{τ} la función link. Notemos que esta función m_{τ} es la que dota de mayor flexibilidad a este tipo de modelos de regresión frente a los modelos de regresión cuantil paramétrico, pues en este tipo de modelos la función m_{τ} es una función desconocida y que por lo tanto deberá ser estimada. En cambio, en los modelos de regresión cuantil paramétrico se presupone que dicha función es conocida. Además de presentar el modelo de regresión cuantil *single-index*, en el Capítulo 2 se detalla el algoritmo de estimación para la obtención de estimadores asociados a dicho modelo y sus propiedades asintóticas junto con los resultados de convergencia de dicho algoritmo. Finalmente se muestra un ejemplo de aplicación modelo de regresión *single-index* con datos simulados.

Una vez hemos presentado el ajuste del modelo de regresión cuantil tipo *single-index*, resulta crucial disponer de procedimientos de contraste de dichos modelos. Dado que en la literatura no existe ninguna propuesta en esta línea, en el Capítulo 3 se presenta un nuevo contraste de bondad de ajuste para modelos de regresión cuantil tipo *single-index* basado en la función de regresión integrada presentada por Stute(1997) y usando la idea de proyecciones introducida por Escanciano (2006) puesto que se supone que estamos considerando múltiples variables explicativas. En particular la nueva propuesta es una extensión del contraste propuesto por Conde-Amboage et al (2015) para el caso de modelos *single-index*. Para finalizar este Capítulo 3 se presentará un pequeño estudio de simulación para ilustrar el buen compor-

tamiento del nuevo contraste de bondad de ajuste en la práctica.

Una vez presentado tanto el modelo de regresión cuantil *single-index* como los contrastes de bondad de ajuste, en el Capítulo 4 llevamos a la práctica toda la teoría explicada a lo largo de los capítulos previos. Primeramente ajustamos un modelo de regresión cuantil lineal para cada uno de los valores de $\tau = 0.25$, $\tau = 0.5$ y $\tau = 0.75$. Con el objetivo de ver si dichos modelos son válidos para cada uno de los cuantiles considerados, aplicamos el contraste de bondad de ajuste visto en la Sección 3.1. Dado que los resultados de dicho contraste rechazaban la hipótesis nula según la cual el modelo de regresión cuantil lineal sería correcto, el siguiente paso realizado consistió en ajustar un modelo de regresión cuantil *single-index*. Con el objetivo de ver si dicho modelo es correcto, aplicamos el contraste de bondad de ajuste visto en la Sección 3.2 para cada uno de los valores τ considerados. Tras los resultados arrojados por dicho contraste, estamos en condiciones de aceptar la hipótesis nula según la cual el modelo de regresión cuantil *single-index* es correcto para dos de los tres cuantiles considerados. De esta manera hemos obtenido un modelo satisfactorio para varios de los cuantiles considerados, ahorrando la complejidad y tiempo de computación requerida a la hora de ajustar modelos no paramétricos. Este proceso lo hemos realizado de manera similar para los datos de `boston` y `weather` de .

A modo de conclusión, los modelos de regresión cuantil *single-index* deben ser vistos como una simplificación de los modelos de regresión cuantil no paramétricos, constituyendo una buena estrategia como paso previo a considerar otras alternativas completamente no paramétricas. Tal y como hemos visto en los dos ejemplos con datos reales mostrados en el Capítulo 4, para algunos cuantiles estaríamos en condiciones de aplicar este tipo de modelos de regresión, sin necesidad de optar por alternativas completamente no paramétricas con mayor coste computacional.

Para terminar, en el Apéndice A se incluye el código de  empleado para

la aplicación de las diferentes técnicas presentadas a lo largo del trabajo.