



M Ű E G Y E T E M 1 7 8 2

MATEMATIKA INTÉZET

Regressziós modellek bizonytalansága

Készítette

Czakó Adrienn

Matematika

Témavezető

Dr. Csáji Balázs Csanád

SZTAKI és ELTE

BUDAPEST, 2023

Kivonat

A regressziós modellek, azon belül is a legkisebb négyzetes becslés, sok területen alkalmazott módszer. Emiatt tartottam fontosnak, hogy én is és az olvasó is mélyebb betekintést nyerjen a témába. Szakdolgozatomat is ennek megfelelően építettem fel.

Első sorban a probléma algebrai eredetét vezettem be, hogy egy túlhatározott egyenletrendszer miképp lehet megoldani vagy annak megoldását becsülni. Analitikus bizonyításokkal támasztottam alá, hogy a legkisebb négyzetekkel történő becsléssel tudjuk minimalizálni a hibát, melynek során az eredeti adatainkat szeretnénk közelíteni x függvényében. Ehhez meg kell határoznunk, hogy x milyen függvényét és milyen paramétereket használjunk.

Erre példaképpen zajos adatokat becsültem először konstans taggal, majd egyenes-sel végül x magasabb fokú polinomjával.

A determinisztikus esetet követően az adatok sztochasztikus viselkedését vizsgáltam. Feltételeztük, hogy a zaj rendelkezik valamilyen határeloszlással és ezt kihasználva tudjuk becsülni a keresett paramétert. A határeloszlást kihasználva konfidencia ellipszist szerkesztettünk a paraméter, majd konfidenciasávot a becsült függvény köré. Ezek a konfidencia szintek segítenek eldönteni, hogy a modellünk mennyire megbízható. Megmutatják milyen tartományban mozoghatunk, ha a célunk nem csak az eredeti függvény vagy érték becslése de mellette például a kockázat csökkentése. Továbbá segítenek eldönteni, hogy a becslésünk amellet, hogy csökkenti a hibát tényleg megbízható eredményt ad.