

6.5 Vícerozměrné lineární regresní modely

Vícenásobná lineární regrese se týká skupiny technik sloužících ke studiu lineární závislosti mezi dvěma či více proměnnými. Určuje odhady parametrů β v regresním modelu

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_m x_{i,m} + \epsilon_i,$$

kde x jsou nezávisle proměnné a y je závisle proměnná. Index i značí pořadové číslo měření a β jsou neznámé regresní parametry a b jejich odhady o počtu m . Pro $m = 1$ se regresní model zjednoduší na jednoduchou lineární regresi. Absolutní člen β_0 je průsečíkem regresní nadroviny s osou y . Odhady b_j jsou směrnice regresní nadroviny ze směru x_j a jsou nazvány parciálními regresními parametry (nebo parciálními regresními koeficienty). Každý takový parciální regresní parametr představuje síť efektů j -té proměnné působící na závisle proměnnou, když ostatní x jsou v regresním modelu drženy na konstantních hodnotách.

Lineární regresní model je schopen změřit pouze lineární, přímkový vztah. Jsou-li body v kruhu, regresní analýza nebude detekovat lineární vztah. Doporučuje se vynášet parciální regresní grafy pro všechny proměnné a vyšetřovat přímkový charakter grafu. Tak se odhalí nepřímkový tvar, vybočující body, míra rozptýlení bodů okolo přímky a řada dalších anomálií přímkového grafu. Nejdůležitějším kritériem lineariry je *Pearsonův korelační koeficient* r . Blíží-li se jeho hodnota +1 nebo -1, jde o přímkový vztah; blíží-li se však nule, nejde o lineární (přímkový) vztah. Perfektní přímka má r rovno +1 (vzestupná přímka) nebo r rovno -1 (sestupná přímka).