

稳健 (robust) 线性回归

稳健线性回归适用于普通线性回归的任何情况，当做回归诊断时，可能发现数据存在一些异常值，而且没有理由去除这些异常值进行分析，此时可用稳健线性回归。回归方程中预测值和实际观测值的差值是残差，应变变量 Y 异常的值残差比较大，自变量 X 异常的值表现为其离均值较远。此时做稳健线性回归时，系统将根据变量值残差大小赋予不同的权重，残差越大，权重越小。

例：DEMO 数据对 SBP 进行稳健回归输入界面如下：

稳健(Robust)线性回归 ?

标题:

选择分析对象:

应变变量

变量名	分布类型	联系函数
Systolic BP, mmhg	Gaussian	Identity

自变量

变量
Age, years
Height, m
Weight, kg
SEX
SMOKE
Alcohol
Occupation
Education

分层变量

研究对象编号

其中研究对象编号是用来识别有异常值的观测记录。

输出结果：

结果变量: Systolic BP, mmhg

变量分布与联系函数: gaussian

模型: $SBP \sim AGE+HEIGHT+WEIGHT+SEX+SMOKE+ALH+OCCU.NEW+factor(EDU.NEW)$

稳健回归 (迭代重加权最小二乘法)

	Value	Std. Error	t value	95%CI low	95%CI upp
(Intercept)	142.7495	21.1815	6.7393	101.2337	184.2652
AGE	0.6561	0.0517	12.6843	0.5547	0.7575
HEIGHT	-21.6669	13.5887	-1.5945	-48.3007	4.9668
WEIGHT	0.233	0.1063	2.192	0.0247	0.4413
SEX	-6.5231	2.1123	-3.0881	-10.6631	-2.383
SMOKE	-3.1107	1.6938	-1.8365	-6.4306	0.2092
ALH	-3.0637	1.8783	-1.6311	-6.7451	0.6178
OCCU.NEW	-4.3884	1.2538	-3.5001	-6.8458	-1.931
factor(EDU.NEW) 2	0.2867	1.6438	0.1744	-2.9351	3.5085
factor(EDU.NEW) 3	1.2338	1.9749	0.6247	-2.637	5.1045

Log Likelihood:	-3455.6405 , df= 11
sigma:	14.6848
Number of observations used:	784

普通最小二乘回归

	Estimate	Std. Error	t value	Pr(> t)	95%CI low	95%CI upp
(Intercept)	147.8383	24.944	5.9268	0	98.948	196.7286
AGE	0.8057	0.0609	13.2278	0	0.6863	0.9251
HEIGHT	-29.5941	16.0024	-1.8493	0.0648	-60.9589	1.7707
WEIGHT	0.3088	0.1252	2.467	0.0138	0.0635	0.5541
SEX	-6.5138	2.4875	-2.6186	0.009	-11.3893	-1.6383
SMOKE	-3.5055	1.9947	-1.7574	0.0792	-7.4152	0.4042
ALH	-1.8426	2.2119	-0.833	0.4051	-6.178	2.4928
OCCU.NEW	-4.3717	1.4765	-2.9609	0.0032	-7.2656	-1.4778
factor(EDU.NEW) 2	0.3192	1.9358	0.1649	0.8691	-3.4749	4.1133
factor(EDU.NEW) 3	0.5288	2.3257	0.2274	0.8202	-4.0295	5.0871

sigma:	19.7411
Log Likelihood:	-3445.8529 , df= 11
R-squared:	0.2417
Adj. r-squared:	0.2329
Number of observations used:	784

被分配权重最小的 10 条记录

Cooks.distance	std.residual	weight	SBP	AGE	HEIGHT	WEIGHT	SEX	SMOKE	ALH	OCCU.NEW	EDU.NEW	SUBJ
0.0396377027791032	4.90165528794603	0.190195239656303	255	68.6	1.46	53.5	2	0	0	1	1	378
0.0272284372562517	4.48782888285521	0.212756742944988	234	60.5	1.66	55	1	1	0	2	2	508
0.0335020755372733	4.0904860037293	0.228949780947317	226	63.9	1.72	58	1	1	1	2	2	289
0.0250505296953744	3.61090765838018	0.260224368727607	220	59.4	1.67	59.5	1	0	0	2	1	743
0.0103355328296951	3.60732683052625	0.261036691757827	210	54.4	1.48	42	2	0	0	2	1	52

0.0134598488528394	3.4983224795797	0.264620694933585	219	63.7	1.48	40	2	0	0	1	1	117
0.0117418205275415	3.48919803447704	0.269945473646378	211	52.1	1.43	40	2	0	0	1	1	663
0.0159989477486614	3.30573194349291	0.283284985967438	208	57.4	1.55	51.5	2	1	0	1	1	155
0.0136113482478162	3.29478664748086	0.293397575402698	196	33.9	1.57	65	2	0	0	1	1	391
0.0191479507044136	3.28913396251554	0.276522289932103	213	64.7	1.56	50	1	1	1	2	1	347

