

Projekt 5: Regressionsanalyse von CPU-Leistungen

Der Datensatz `COMPUTER.DAT` enthält technische Merkmale und zwei relative Leistungsmaße von 209 CPU's (central processing units) von Computern verschiedener Herstellern. Dabei wird die Leistung im Verhältnis zu der Leistung eines IBM 370/158-3 angegeben. Es soll untersucht werden, wie die technischen Merkmale eines Computers dessen relative CPU-Leistung (Perf für Performance) und dessen geschätzte relative CPU-Leistung (EstPerf) beeinflussen. Dabei enthält der Datensatz `COMPUTER.DAT` folgende technische Merkmale

Cycle = Cycle time in ns

MinMem = Minimum Memory in kb

MaxMem = Maximum Memory in kb

Cache = Cache Size in kb

MinChan = Minimum Channels

MaxChan = Maximum Channels

und der Datensatz hat die Form

Cycle	MinMem	MaxMem	Cache	MinChan	MaxChan	Perf	EstPerf
125	256	6000	256	16	128	198	199
29	8000	32000	32	8	32	269	253
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Aufgaben

1. Welche technischen Merkmale haben einen signifikanten Einfluss auf die relative und die geschätzte relative CPU-Leistung? Untersuchen Sie dazu, ob eine Kollinearität der Regressoren vorliegt. Beachten Sie dabei auch, dass der Versuchsplan nicht balanziert ist.
2. Bestimmen Sie unter Verwendung von Variablenselektion- sowie Modelldiagnoseverfahren ein möglichst „gutes“ Modell, welches die relative CPU-Leistung anhand der ausgewählten Variablen erklärt.
3. Sagen Sie mit dem in 2. bestimmten Modell die relative CPU-Leistung bei Cycle Time = 300 ns, Minimum Memory = 500 kb, Maximum Memory = 4000 kb, Cache-Größe = 6 kb, Minimum Channels = 16, Maximum Channels = 64 voraus.

Literatur

Elpelt, B., Hartung, J. (2004). *Grundkurs Statistik. Lehr- und Übungsbuch der angewandten Statistik.*, Oldenbourg, München.

Sachs, L. und Hedderich, J. (2009). *Angewandte Statistik: Methodensammlung mit R.* Springer, Berlin.

Faraway, J.J. (2006). *Extending the Linear Model with R.* Chapman & Hall/CRC.

Groß, J. (2003). *Linear Regression.* Springer, Berlin.

http://en.wikipedia.org/wiki/Akaike_information_criterion, 12.2.2010

Literatur zu den Daten

Ein-Door, P. und Feldmesser, J. (1987). Attributes of the performance of central processing units: a relative performance prediction model. *Communications of the ACM* **30**, 308-317.

aus

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. und Ostrowski, E. (1994). *A Handbook of Small Data Sets.* Chapman & Hall, London, Seite 366.

Abgabetermin

Abgabe bis spätestens **Montag, den 18. Juni 2012**, in der Veranstaltung.