

Regressions- och tidsserieanalys

Föreläsning 6 - Sammanfattning av enkel och multipel linjär regression

Mattias Villani

Statistiska institutionen
Stockholms universitet

Institutionen för datavetenskap
Linköpings universitet



mattiasvillani.com




[@matvil](https://twitter.com/matvil)



[mattiasvillani](https://github.com/mattiasvillani)

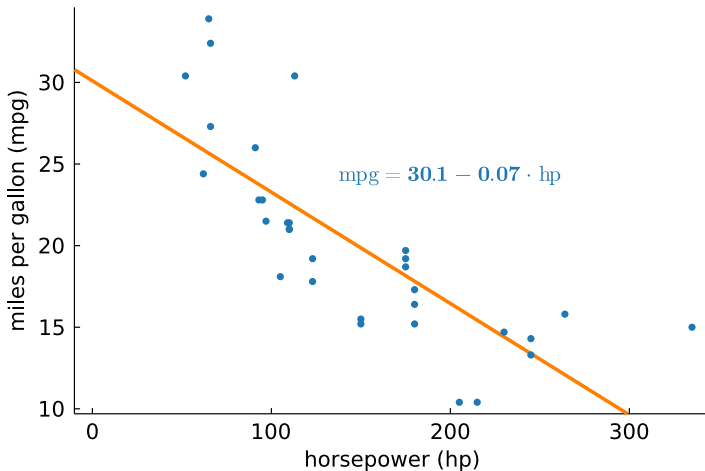
- Analys av datamaterialet mtcars

- `mtcars.csv` (F6 webbsida).  data `mtcars`
- Observationer: 32 bilar av olika märken.
- Responsvariabel: bensinförbrukning (mpg, miles per gallon).

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3
Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4
Lincoln Continental	10.4	8	460.0	215	3.00	5.424	17.82	0	0	3	4
Chrysler Imperial	14.7	8	440.0	230	3.23	5.345	17.42	0	0	3	4
Fiat 128	32.4	4	78.7	66	4.08	2.200	19.47	1	1	4	1
Honda Civic	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
Toyota Corolla	33.9	4	71.1	65	4.22	1.835	19.90	1	1	4	1
Toyota Corona	21.5	4	120.1	97	3.70	2.465	20.01	1	0	3	1
Dodge Challenger	15.5	8	318.0	150	2.76	3.520	16.87	0	0	3	2
AMC Javelin	15.2	8	304.0	150	3.15	3.435	17.30	0	0	3	2
Camaro Z28	13.3	8	350.0	245	3.73	3.840	15.41	0	0	3	4
Pontiac Firebird	19.2	8	400.0	175	3.08	3.845	17.05	0	0	3	2
Fiat X1-9	27.3	4	79.0	66	4.08	1.935	18.90	1	1	4	1
Porsche 914-2	26.0	4	120.3	91	4.43	2.140	16.70	0	1	5	2
Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.90	1	1	5	2
Ford Pantera L	15.8	8	351.0	264	4.22	3.170	14.50	0	1	5	4
Ferrari Dino	19.7	6	145.0	175	3.62	2.770	15.50	0	1	5	6
Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.60	0	1	5	8
Volvo 142E	21.4	4	121.0	109	4.11	2.780	18.60	1	1	4	2

mpg	Miles/(US) gallon
cyl	Number of cylinders
disp	Displacement (cu.in.)
hp	Gross horsepower
drat	Rear axle ratio
wt	Weight (1000 lbs)
qsec	1/4 mile time
vs	Engine (0 = V-shaped, 1 = straight)
am	Transmission (0 = automatic, 1 = manual)
gear	Number of forward gears

miles per gallon vs horsepower



Cars data - enkel linjär regression

```
> library(regkurs)
> lmfit = lm(mpg ~ hp, data = mtcars)
> regsummary(lmfit)
```

Analysis of variance - ANOVA

```
-----
              df      SS      MS      F      Pr(>F)
Regr      1  678.37 678.373 45.46 1.7878e-07
Error    30  447.67  14.922
Total    31 1126.05
```

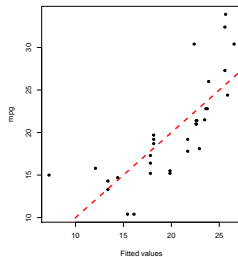
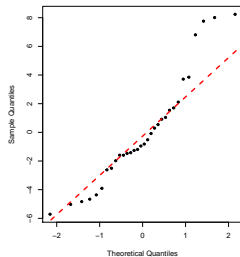
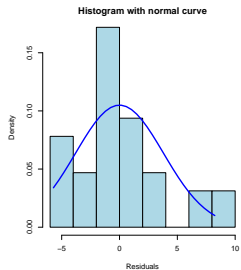
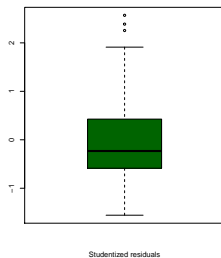
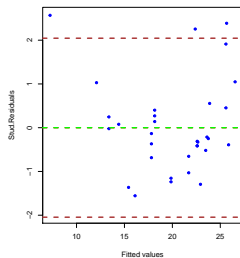
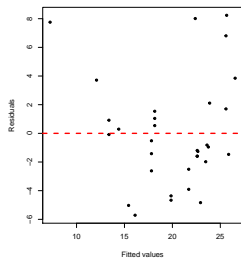
Measures of model fit

```
-----
Root MSE      R2    R2-adj
  3.86296  0.60244  0.58919
```

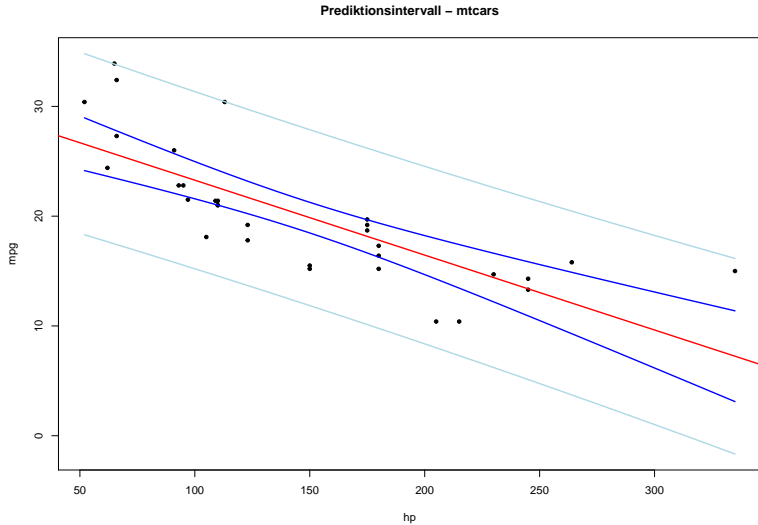
Parameter estimates

```
-----
              Estimate Std. Error t value  Pr(>|t|)
(Intercept) 30.098861   1.633921 18.4212 6.6427e-18
hp          -0.068228   0.010119 -6.7424 1.7878e-07
```

Cars data - residual analysis



Cars data - prediktionsintervall



Cars data - multipel regression

```
> lmfit = lm(mpg ~ hp + wt + am, data = mtcars)
> regsummary(lmfit)
```

Analysis of variance - ANOVA

```
-----
              df      SS      MS      F      Pr(>F)
Regr      3   945.76 315.252 48.96 2.9079e-11
Error    28   180.29   6.439
Total    31  1126.05
```

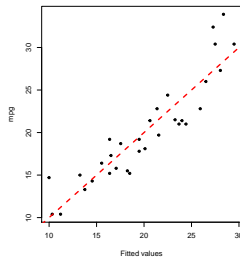
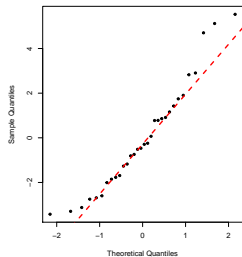
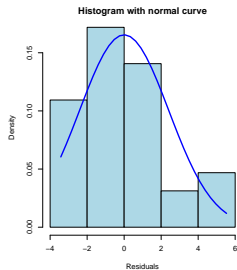
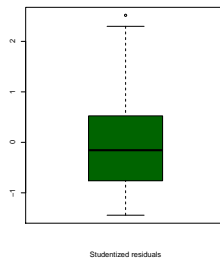
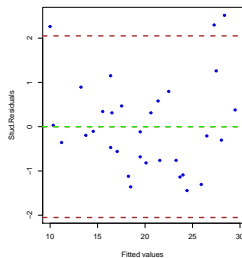
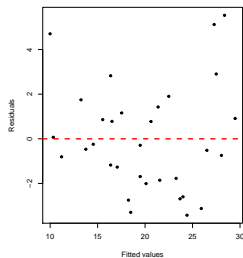
Measures of model fit

```
-----
Root MSE      R2    R2-adj
2.53751  0.83989  0.82274
```

Parameter estimates

```
-----
              Estimate Std. Error t value  Pr(>|t|)
(Intercept) 34.002875  2.6426593 12.8669 2.8240e-13
hp          -0.037479  0.0096054 -3.9018 5.4640e-04
wt          -2.878575  0.9049705 -3.1808 3.5740e-03
am           2.083710  1.3764202  1.5139 1.4127e-01
```


Cars data - multipel regression



Cars data - multipel regression

