LEAST SQUARES - **Prof. Richard B. Goldstein**

DISCRETE DATA

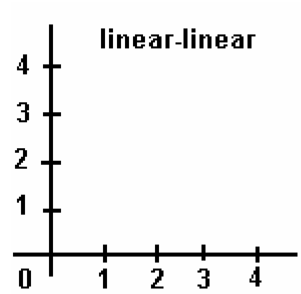
<u>GIVEN</u>: (x_i, y_i) for $i = 1,2,\ldots,n$

MODEL #1 $y = a_0 + a_1 x + \ldots + a_m x^m$

OBJECTIVE: minimize $\sum_{i=1}^{n} \left(y_i - a_0 - a_1 x_i - \ldots - a_m x_i^m\right)^2$

NORMAL EQUATIONS:

$$a_0 n + a_1 \sum x_i + \ldots + a_m \sum x_i^m = \sum y_i$$
$$a_0 \sum x_i + a_1 \sum x_i^2 + \ldots + a_m \sum x_i^{m+1} = \sum x_i y_i$$
$$\cdots$$
$$a_0 \sum x_i^m + a_1 \sum x_i^{m+1} + \ldots + a_m \sum x_i^{2m} = \sum x_i^m y_i$$

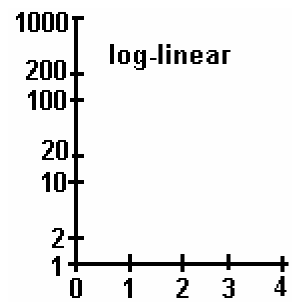


MODEL #2 $y = be^{ax}$ or $\ln y = \ln b + ax$

OBJECTIVE[1] minimize $\sum \left(y_i - be^{ax_i}\right)^2$

NORMAL EQUATIONS:

$$n \ln b + a \sum x_i = \sum \ln y_i$$
$$\ln b \sum x_i + a \sum x_i^2 = \sum x_i \ln y_i$$

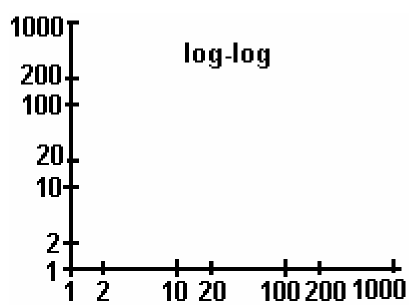


MODEL #3 $y = bx^a$ or $\ln y = \ln b + a \ln x$

OBJECTIVE[1] minimize $\sum \left(y_i - bx_i^a\right)^2$

NORMAL EQUATIONS:

$$n \ln b + a \sum \ln x_i = \sum \ln y_i$$
$$\ln b \sum \ln x_i + a \sum \left(\ln x_i\right)^2 = \sum \left(\ln x_i\right)\left(\ln y_i\right)$$



---

[1] This objective is only approximately met. The minimum takes place on log-linear or log-log scaled graphs.

EXAMPLE

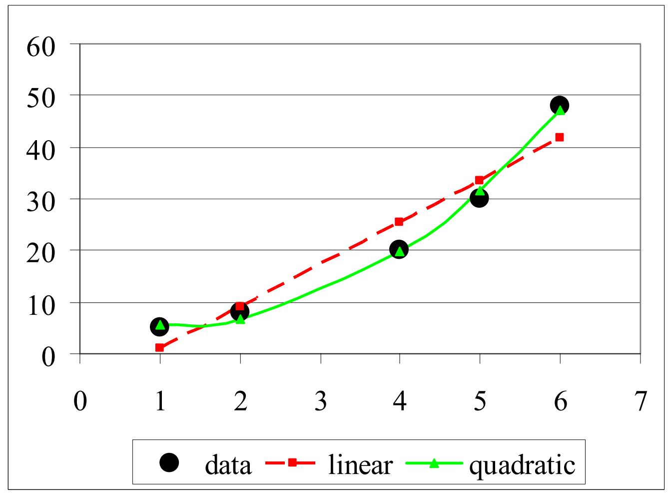
x	y	x^2	x^3	x^4	xy	x^2y	ln y	x ln y	ln x	$(\ln x)^2$	(ln x)(ln y)
1	5	1	1	1	5	5	1.60944	1.60944	0.00000	0.00000	0.00000
2	8	4	8	16	16	32	2.07944	4.15888	0.69315	0.48045	1.44136
4	20	16	64	256	80	320	2.99573	11.98293	1.38629	1.92181	4.15297
5	30	25	125	625	150	750	3.40120	17.00599	1.60944	2.59029	5.47402
6	48	36	216	1296	288	1728	3.87120	23.22721	1.79176	3.21040	6.93626
18	111	82	414	2194	539	2835	13.95701	57.98444	5.48064	8.20296	18.00460

<u>model #1 (m =1 / linear)</u>:

$5a_0 + 18a_1 = 111$ → $a_0 = -600/86 = -6.977$

$18a_0 + 82a_1 = 539$ $a_1 = 697/86 = 8.105$

$\hat{y} = -6.977 + 8.105x$



<u>model #1 (m = 2 / quadratic)</u>:

$5a_0 + 18a_1 + 82a_2 = 111$ $a_0 = 8.318$

$18a_0 + 82a_1 + 414a_2 = 539$ → $a_1 = -4.375$

$82a_0 + 414a_1 + 2194a_2 = 2835$ $a_2 = 1.807$

$\hat{y} = 8.318 - 4.375x - 1.807x^2$

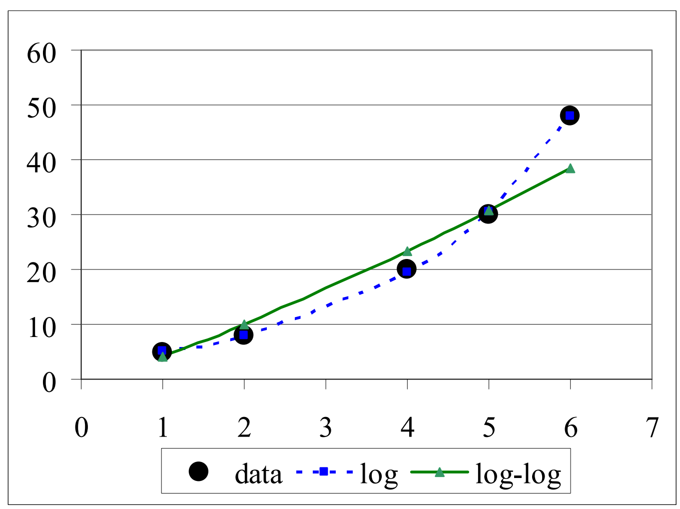
<u>model #2 (log / exponential)</u>:

$5 \ln b + 18 a = 13.95701$ → $\ln b = 1.172$

→ $b = e^{1.172} = 3.227$

$18 \ln b + 82 a = 57.98444$ $a = 0.450$

$\hat{y} = 3.227e^{0.450x}$



<u>model #3 (log-log / power)</u>:

$5 \ln b + 5.481 a = 13.95701$ → $\ln b = 1.440$

→ $b = e^{1.440} = 4.222$

$5.481 \ln b + 8.203 a = 18.00460$ $a = 1.232$

$\hat{y} = 4.222x^{1.232}$

x	y	linear	quadratic	exponential	power
1	5	1.128	5.750	5.061	4.222
2	8	9.233	6.796	7.937	9.917
4	20	25.443	19.730	19.522	23.295
5	30	33.548	31.618	30.617	30.665
6	48	41.653	47.120	48.017	38.388
Sum Sqs.		99.012	5.477	0.646	107.970

CONTINUOUS DATA

GIVEN: f(x) on [a, b]

MODEL: $y = a_0 + a_1 x + \ldots + a_m x^m$

OBJECTIVE: Minimize $\int_a^b \left(f(x) - a_0 - a_1 x - \ldots a_m x^m\right)^2 dx$

NORMAL EQUATIONS:

$$a_0 \int_a^b 1\,dx + a_1 \int_a^b x\,dx + \ldots + a_m \int_a^b x^m\,dx = \int_a^b f(x)\,dx$$

$$a_0 \int_a^b x\,dx + a_1 \int_a^b x^2\,dx + \ldots + a_m \int_a^b x^{m+1}\,dx = \int_a^b x f(x)\,dx$$

$$\cdots$$

$$a_0 \int_a^b x^m\,dx + a_1 \int_a^b x^{m=1}\,dx + \ldots + a_m \int_a^b x^{2m}\,dx = \int_a^b x^m f(x)\,dx$$

EXAMPLE

$f(x) = \sqrt{x} = x^{1/2}$ on [1, 4] with m = 1

$3 \quad a_0 + \ \ 7.5\ a_1 = 14/3 = \ 4.667$ → $a_0 = \ \ 5/6.75 = 0.741$
$7.5\ a_0 + 21.0\ a_1 = 62/5 = 12.400$ $\quad a_1 = 2.2/6.75 = 0.326$

$f(x) = \sqrt{x} \approx 0.741 + 0.326x$ on $[1,4]$

x	f(x)	0.741 + 0.326x
1	1	1.067
2	1.414	1.393
3	1.732	1.719
4	2	2.045