ASSIGNMENT 2

**Name: Jain Samkitkumar Hasmukhlal**

**Roll No.: 20 (BEIT)**

> sample <- read.csv("C:/Users/Admin/Desktop/R/sample.txt", row.names=1, sep="")

> View(sample)

> sample

p.id y x1 x2 x3 x4 x5 x6 x7

1 297497 10.24 9.47 10.70 10.54 9.01 10.10 10.30 9.86

2 218250 10.47 11.46 8.87 9.90 10.31 9.71 11.32 11.28

3 808994 11.61 10.00 11.69 11.31 9.39 8.88 10.01 9.58

4 352628 9.57 10.09 8.95 8.49 9.63 8.15 9.99 9.56

5 371676 9.90 9.37 10.91 11.63 11.77 10.69 10.84 11.28

6 342721 9.65 9.13 10.28 10.11 8.74 9.84 10.30 10.51

7 936561 11.83 11.77 10.93 8.87 10.14 9.91 10.50 9.95

8 688596 9.95 10.89 9.59 8.78 10.50 10.80 10.89 11.43

9 617893 11.83 12.59 9.18 11.18 10.72 11.35 9.80 9.56

10 892466 10.08 9.44 11.35 9.81 10.64 10.28 10.47 10.13

11 583189 11.70 11.52 10.87 9.32 10.96 9.26 9.15 9.64

12 884235 9.12 10.49 9.17 7.54 11.33 9.55 10.14 10.79

13 247372 10.85 11.43 9.74 11.07 10.60 12.19 10.14 10.71

14 845347 11.47 10.72 11.31 10.21 8.86 10.51 8.77 9.31

15 584813 9.81 11.29 8.97 8.25 10.23 10.53 11.86 12.21

16 331417 9.30 7.65 11.13 10.87 9.69 7.66 7.41 7.08

17 420750 9.87 9.40 10.35 9.36 8.39 9.66 9.35 9.57

18 848683 8.88 8.97 9.47 10.35 9.82 10.14 11.12 11.22

19 432153 10.36 10.72 9.66 10.57 10.92 10.28 11.48 11.17

20 628985 10.63 9.87 11.09 10.98 10.61 9.27 9.52 9.31

21 402889 10.02 10.64 9.69 9.22 11.51 9.34 9.98 10.45

22 545527 9.15 9.20 10.14 9.90 10.30 10.56 10.65 9.93

23 651801 10.53 11.11 10.40 8.39 10.60 11.21 9.75 10.44

24 128116 10.42 10.53 9.73 9.80 9.76 10.48 9.19 9.26

25 962372 8.65 8.56 10.07 10.20 10.61 10.02 10.40 10.38

26 40027 10.82 9.48 10.80 12.01 8.82 9.03 8.87 8.24

27 331741 10.18 10.09 9.22 11.47 10.01 9.49 10.46 10.25

28 688675 11.43 11.57 10.41 8.90 11.45 9.22 6.98 6.90

29 185088 9.38 10.40 8.97 9.04 10.47 11.07 9.74 9.57

30 149812 10.69 10.60 10.40 11.54 10.50 11.14 12.44 12.40

31 448176 9.99 9.46 11.53 9.98 11.66 10.70 12.11 11.75

32 260936 9.97 9.89 10.11 9.69 9.97 10.55 8.54 8.80

33 91951 11.46 12.22 10.02 8.47 10.70 9.59 10.35 10.60

34 894025 9.90 8.97 10.73 11.90 10.69 9.60 9.65 9.92

35 373101 11.17 11.12 11.34 7.98 8.70 10.87 9.51 10.47

36 305071 9.38 10.35 9.27 9.00 11.47 8.69 12.63 11.59

37 179629 11.00 9.90 11.70 10.67 10.28 9.79 9.65 10.06

38 300820 9.91 10.88 8.15 10.09 9.81 9.47 10.03 9.48

39 103049 9.36 10.69 7.63 10.14 10.41 9.14 10.17 9.96

40 915820 9.79 9.64 10.35 9.54 9.51 10.31 10.43 10.96

41 953096 8.79 9.54 9.16 9.78 11.03 11.11 9.89 9.15

42 446925 9.68 9.81 8.86 10.42 8.89 8.88 9.17 9.11

43 233823 7.88 9.40 6.61 9.84 7.82 10.82 9.91 10.13

44 323472 9.14 9.13 9.78 10.17 10.62 10.19 11.23 11.18

45 675515 10.06 10.02 10.19 9.91 9.35 10.36 12.60 12.26

46 558965 10.84 10.35 10.07 10.61 10.45 8.22 8.81 8.85

47 291438 10.61 11.38 9.80 9.50 8.96 12.03 10.62 11.18

48 441567 10.91 10.82 10.64 10.69 10.59 10.59 10.27 10.16

49 33985 10.24 9.20 12.00 9.16 9.94 9.61 10.80 10.74

50 729983 10.84 10.29 10.91 9.92 8.31 10.32 10.52 11.23

51 183190 9.53 9.80 9.90 9.59 9.51 11.18 10.53 11.15

52 767464 9.17 8.80 9.75 10.12 9.71 8.82 9.68 9.03

53 144760 10.75 10.58 11.02 11.19 12.29 11.26 12.83 11.74

54 704173 11.02 10.56 11.19 11.27 10.58 11.48 10.90 10.93

55 679179 11.03 10.49 11.16 10.45 10.89 9.89 8.58 9.00

56 382746 9.85 10.83 8.32 10.05 10.09 10.11 11.50 11.06

57 302919 9.51 9.01 10.05 10.11 9.81 8.76 9.77 8.60

58 594967 10.63 9.80 11.26 10.73 9.17 10.74 9.27 9.90

59 702793 10.29 9.83 11.30 8.99 9.72 10.54 9.60 10.73

60 974842 8.86 9.51 8.16 11.69 9.51 10.93 10.10 10.43

61 755316 11.59 10.40 12.02 10.15 9.95 9.71 8.42 7.55

62 275564 9.18 9.73 8.90 10.65 10.35 10.04 9.10 9.64

63 474705 8.97 9.70 9.17 9.07 9.71 10.26 9.47 10.02

64 158827 9.46 9.12 9.22 11.43 9.46 9.17 10.94 9.51

65 225003 11.70 12.31 10.01 9.33 10.61 10.26 10.06 10.97

66 978778 10.87 11.04 10.26 9.72 9.61 10.26 11.18 11.47

67 345114 10.36 10.33 10.26 10.35 10.59 10.95 10.98 11.14

68 331464 10.72 11.16 10.12 10.31 11.46 11.10 9.87 9.76

69 60028 9.23 9.54 9.02 10.19 9.86 10.15 9.57 9.35

70 346826 12.02 10.99 11.28 11.93 10.07 9.83 11.10 10.82

71 376464 9.74 10.46 8.23 10.49 10.29 8.13 10.45 9.53

72 889119 9.42 9.75 8.73 11.49 10.79 9.97 9.69 9.02

73 471537 11.22 11.63 10.39 9.24 9.19 12.03 10.07 10.91

74 701368 9.42 8.76 10.41 10.33 10.45 8.88 8.16 8.50

75 91061 10.10 10.17 9.32 11.83 10.62 9.57 11.20 10.45

76 876968 10.60 10.71 10.10 9.28 8.94 11.12 10.42 10.14

77 938896 9.30 9.07 9.78 9.55 9.88 9.33 9.26 9.54

78 672091 8.23 8.52 9.10 9.43 8.90 9.74 10.03 10.77

79 760473 9.81 10.11 10.34 8.96 11.39 9.93 9.87 10.06

80 772694 10.35 10.05 11.23 9.81 12.33 10.11 8.74 8.98

81 643786 10.71 10.69 10.71 7.90 10.19 8.69 9.72 9.75

82 323282 7.91 8.64 9.05 8.61 10.85 8.98 9.54 9.49

83 657369 7.77 8.83 8.88 8.47 11.56 9.21 9.61 10.37

84 234114 8.98 8.85 9.61 9.90 8.81 9.66 8.28 9.15

85 979113 10.86 11.10 10.10 9.88 10.72 10.68 8.68 8.56

86 770556 8.84 9.87 9.24 9.13 9.94 12.39 10.38 10.87

87 678377 10.77 11.58 9.18 9.48 10.81 8.79 9.58 9.49

88 747496 9.75 10.11 10.05 10.52 10.76 12.10 10.33 10.89

89 679210 11.15 11.68 10.39 9.71 10.82 10.62 10.86 11.31

90 671885 10.73 10.48 9.92 11.81 10.23 9.38 11.69 11.14

91 460454 10.96 11.47 8.87 10.24 9.67 9.47 9.43 9.89

92 675481 11.13 9.60 12.69 10.14 10.96 9.39 8.67 9.59

93 913708 8.82 9.20 9.81 8.78 10.33 10.88 8.75 9.20

94 448100 10.98 12.20 8.62 8.65 9.86 8.89 10.14 10.30

95 608282 11.01 10.51 10.31 10.85 9.79 8.65 10.62 10.31

96 216595 8.36 8.67 9.12 9.63 9.63 10.26 9.69 9.87

97 40556 9.93 10.08 9.56 11.72 10.65 10.15 11.37 11.24

98 810820 9.16 9.32 9.05 11.29 10.92 9.51 9.64 9.07

99 341012 9.72 9.96 9.07 10.32 9.29 9.14 9.98 10.14

100 238466 9.43 8.79 10.67 10.83 10.58 9.23 11.25 10.85

>p <- lm(y ~ sample$x1 , data=sample, x=T, y=T)

>p

Call:

lm(formula = y ~ sample$x1, data = sample, x = T, y = T)

Coefficients:

(Intercept) sample$x1

2.7821 0.7196

>dim(sample)

[1] 100 9

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

9.596326 11.028242 9.977691 10.042451 9.524370 9.351677 11.251304 10.618095

9 10 11 12 13 14 15 16

11.841340 9.574739 11.071415 10.330273 11.006655 10.495771 10.905918 8.286735

17 18 19 20 21 22 23 24

9.545957 9.236548 10.495771 9.884148 10.438206 9.402046 10.776398 10.359055

25 26 27 28 29 30 31 32

8.941530 9.603522 10.042451 11.107393 10.265513 10.409424 9.589131 9.898539

33 34 35 36 37 38 39 40

11.575104 9.236548 10.783593 10.229535 9.905735 10.610900 10.474184 9.718651

41 42 43 44 45 46 47 48

9.646695 9.840975 9.545957 9.351677 9.992082 10.229535 10.970678 10.567726

49 50 51 52 53 54 55 56

9.402046 10.186362 9.833779 9.114224 10.395033 10.380642 10.330273 10.574922

57 58 59 60 61 62 63 64

9.265330 9.833779 9.855366 9.625108 10.265513 9.783411 9.761824 9.344482

65 66 67 68 69 70 71 72

11.639864 10.726029 10.215144 10.812375 9.646695 10.690051 10.308686 9.797802

73 74 75 76 77 78 79 80

11.150567 9.085441 10.100015 10.488575 9.308504 8.912748 10.056842 10.013668

81 82 83 84 85 86 87 88

10.474184 8.999095 9.135810 9.150201 10.769202 9.884148 11.114589 10.056842

89 90 91 92 93 94 95 96

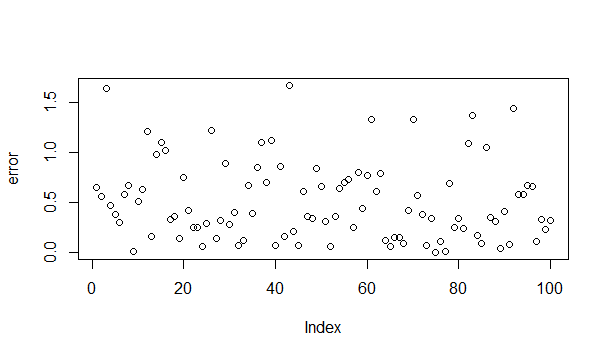
11.186544 10.323077 11.035438 9.689868 9.402046 11.560713 10.344664 9.020681

97 98 99 100

10.035255 9.488393 9.948908 9.107028

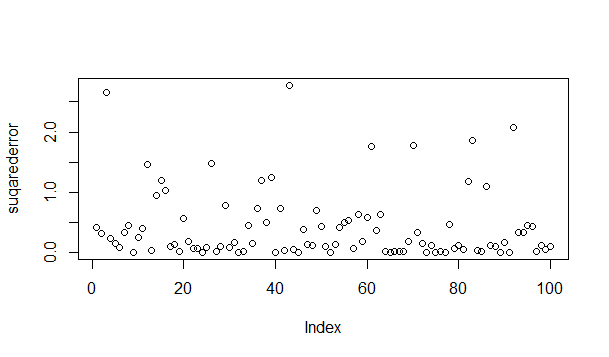
>error <- abs(sample$y - predictedY)

>plot(error)



>suqarederror<- error \* error

>plot(suqarederror)



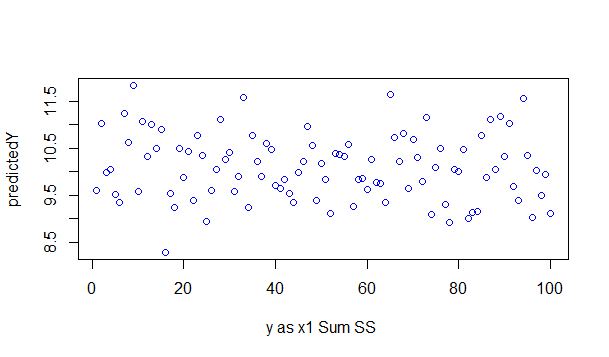
>sumSS<-1:15

>sumSS[1]<-sum(suqarederror)

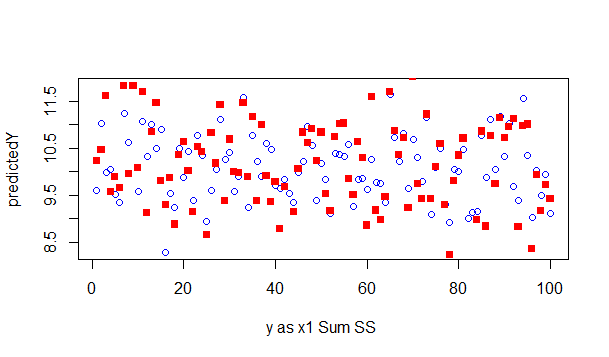
>sumSS[1]

[1] 40.99029

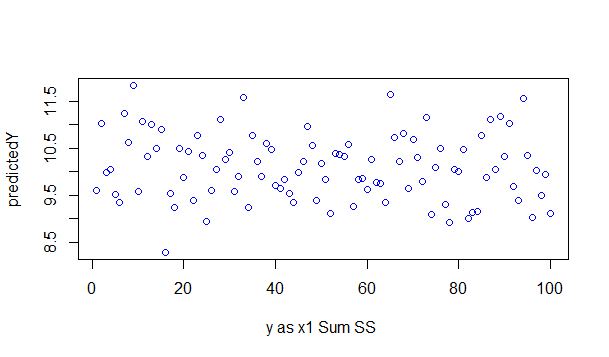
>plot(predictedY,xlab="y as x1 Sum SS " ,col="blue")



>points(sample$y,pch=15, col="red")



>plot(predictedY,xlab="y as x1 Sum SS " ,col="blue")



>p <- lm(y ~ sample$x2 , data=sample, x=T, y=T)

>p

Call:

lm(formula = y ~ sample$x2, data = sample, x = T, y = T)

Coefficients:

(Intercept) sample$x2

4.943 0.516

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.464726 9.520389 10.975598 9.561671 10.573093 10.247993 10.583414 9.891931

9 10 11 12 13 14 15 16

9.680358 10.800147 10.552452 9.675198 9.969336 10.779506 9.571992 10.686620

17 18 19 20 21 22 23 24

10.284115 9.830008 9.928054 10.665979 9.943535 10.175749 10.309917 9.964176

25 26 27 28 29 30 31 32

10.139627 10.516330 9.701000 10.315077 9.571992 10.309917 10.893033 10.160268

33 34 35 36 37 38 39 40

10.113825 10.480207 10.794987 9.726801 10.980758 9.148846 8.880510 10.284115

41 42 43 44 45 46 47 48

9.670038 9.515228 8.354157 9.989977 10.201550 10.139627 10.000298 10.433764

49 50 51 52 53 54 55 56

11.135567 10.573093 10.051901 9.974496 10.629856 10.717582 10.702101 9.236571

57 58 59 60 61 62 63 64

10.129306 10.753704 10.774345 9.154006 11.145888 9.535870 9.675198 9.701000

65 66 67 68 69 70 71 72

10.108665 10.237672 10.237672 10.165428 9.597793 10.764025 9.190128 9.448144

73 74 75 76 77 78 79 80

10.304757 10.315077 9.752603 10.155107 9.989977 9.639076 10.278955 10.738223

81 82 83 84 85 86 87 88

10.469887 9.613274 9.525549 9.902252 10.155107 9.711320 9.680358 10.129306

89 90 91 92 93 94 95 96

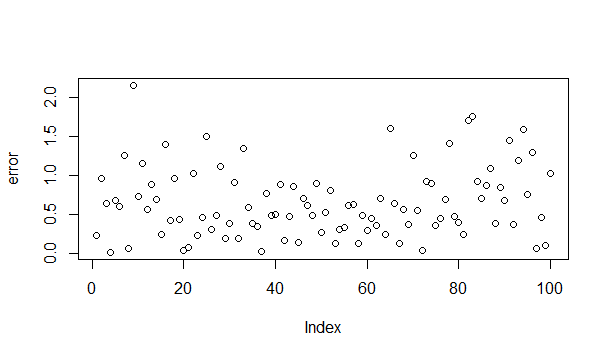
10.304757 10.062222 9.520389 11.491629 10.005458 9.391381 10.263474 9.649397

97 98 99 100

9.876450 9.613274 9.623595 10.449245

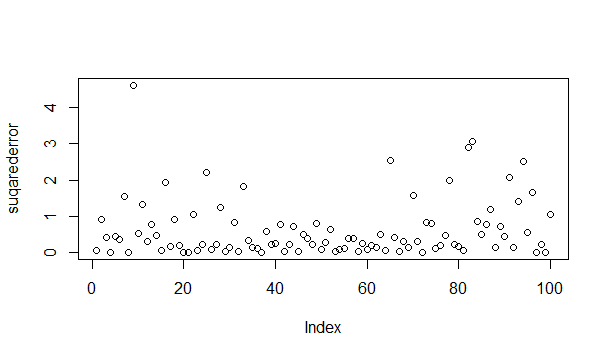
>error <- abs(sample$y - predictedY)

>plot(error)



>suqarederror<- error \* error

>plot(suqarederror)

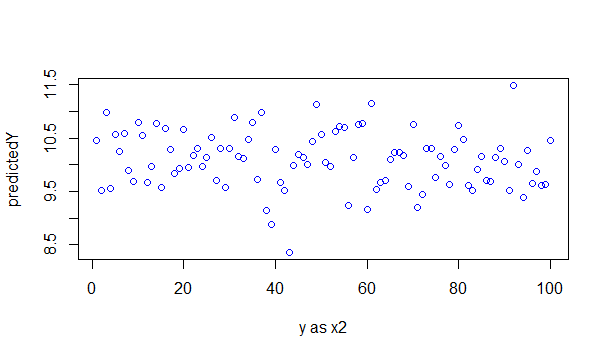


>sumSS[2]<-sum(suqarederror)

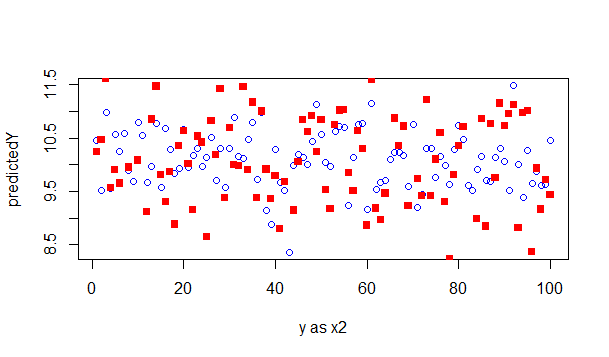
>sumSS[2]

[1] 61.723

> plot(predictedY,xlab="y as x2" ,col="blue")



> points(sample$y,pch=15, col="red")



>p <- lm(y ~ sample$x3 , data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.155823 10.076387 10.251394 9.901379 10.291112 10.102452 9.948544 9.937373

9 10 11 12 13 14 15 16

10.235259 10.065216 10.004398 9.783466 10.221606 10.114863 9.871590 10.196782

17 18 19 20 21 22 23 24

10.009362 10.132240 10.159546 10.210435 9.991986 10.076387 9.888967 10.063975

25 26 27 28 29 30 31 32

10.113622 10.338277 10.271253 9.952268 9.969644 10.279942 10.086316 10.050322

33 34 35 36 37 38 39 40

9.898897 10.324624 9.838078 9.964680 10.171958 10.099969 10.106175 10.031704

41 42 43 44 45 46 47 48

10.061492 10.140928 10.068939 10.109899 10.077628 10.164511 10.026739 10.174440

49 50 51 52 53 54 55 56

9.984539 10.078869 10.037910 10.103693 10.236500 10.246429 10.144652 10.095004

57 58 59 60 61 62 63 64

10.102452 10.179405 9.963438 10.298559 10.107416 10.169476 9.973368 10.266288

65 66 67 68 69 70 71 72

10.005639 10.054045 10.132240 10.127275 10.112381 10.328348 10.149617 10.273736

73 74 75 76 77 78 79 80

9.994468 10.129758 10.315936 9.999433 10.032945 10.018051 9.959715 10.065216

81 82 83 84 85 86 87 88

9.828149 9.916273 9.898897 10.076387 10.073904 9.980815 10.024257 10.153340

89 90 91 92 93 94 95 96

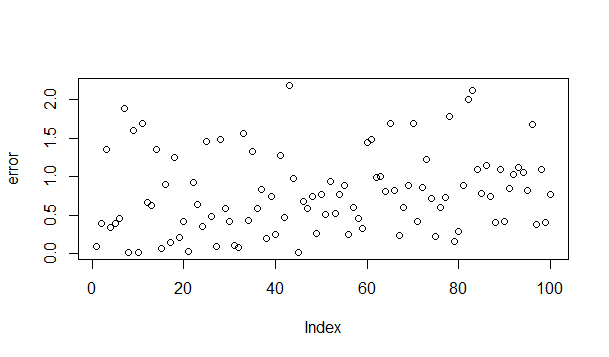
10.052804 10.313454 10.118587 10.106175 9.937373 9.921238 10.194299 10.042874

97 98 99 100

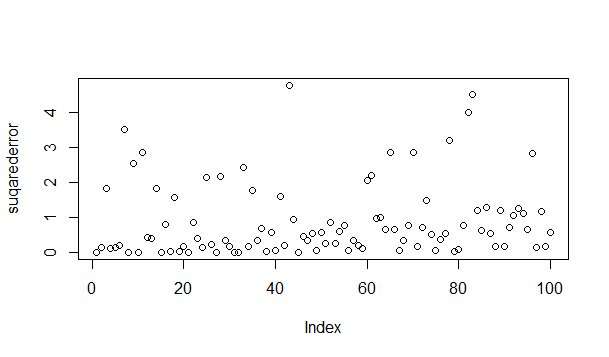
10.302283 10.248912 10.128516 10.191817

>error <- abs(sample$y - predictedY)

>plot(error)

  
>suqarederror<- error \* error

>plot(suqarederror)



|  |
| --- |
| >sumSS[3]<-sum(suqarederror)  >sumSS[3]  [1] 88.20018  > plot(predictedY,xlab="y as x3" ,col="blue")  C:\Users\Admin\Desktop\ass2\12.png |
| >points(sample$y,pch=15, col="red")  C:\Users\Admin\Desktop\ass2\13.png |
| |  | | --- | |  | |

>p <- lm(y ~ sample$x4 , data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.008260 10.101775 10.035595 10.052860 10.206800 9.988838 10.089546 10.115443

9 10 11 12 13 14 15 16

10.131268 10.125514 10.148533 10.175149 10.122636 9.997470 10.096020 10.057176

17 18 19 20 21 22 23 24

9.963661 10.066527 10.145655 10.123356 10.188097 10.101056 10.122636 10.062211

25 26 27 28 29 30 31 32

10.123356 9.994593 10.080195 10.183781 10.113285 10.115443 10.198887 10.077317

33 34 35 36 37 38 39 40

10.129830 10.129110 9.985960 10.185219 10.099617 10.065808 10.108969 10.044227

41 42 43 44 45 46 47 48

10.153568 9.999628 9.922658 10.124075 10.032718 10.111846 10.004663 10.121917

49 50 51 52 53 54 55 56

10.075159 9.957906 10.044227 10.058614 10.244206 10.121198 10.143497 10.085950

57 58 59 60 61 62 63 64

10.065808 10.019770 10.059334 10.044227 10.075879 10.104653 10.058614 10.040631

65 66 67 68 69 70 71 72

10.123356 10.051421 10.121917 10.184500 10.069405 10.084511 10.100337 10.136304

73 74 75 76 77 78 79 80

10.021208 10.111846 10.124075 10.003225 10.070843 10.000347 10.179465 10.247083

81 82 83 84 85 86 87 88

10.093143 10.140620 10.191694 9.993873 10.131268 10.075159 10.137743 10.134146

89 90 91 92 93 94 95 96

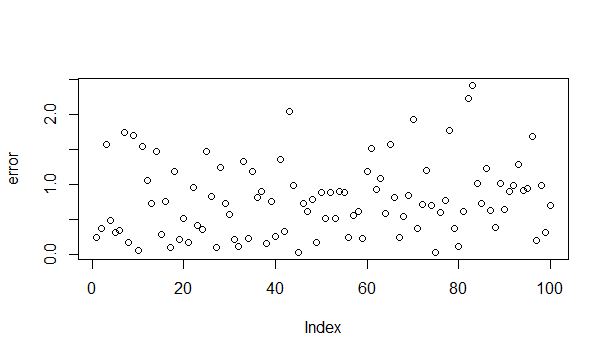
10.138462 10.096020 10.055737 10.148533 10.103214 10.069405 10.064369 10.052860

97 98 99 100

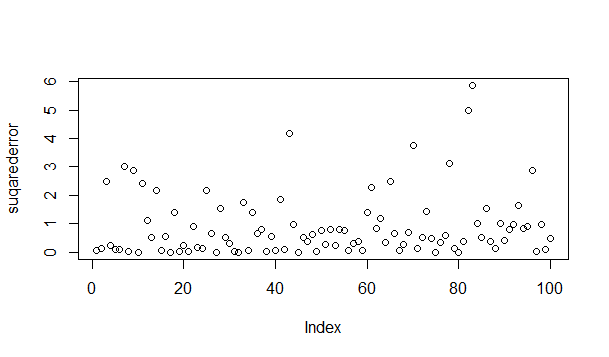
10.126233 10.145655 10.028402 10.121198

>error <- abs(sample1$y - predictedY)

>plot(error)

>suqarederror<- error \* error

>plot(suqarederror)

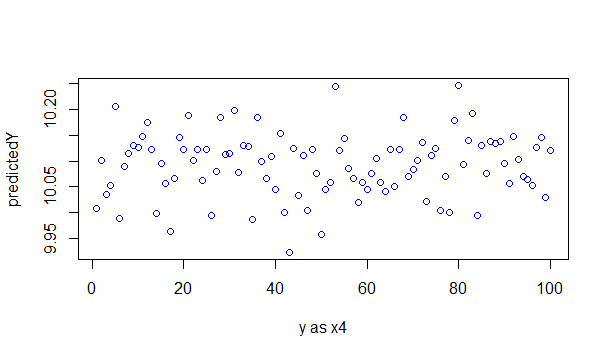


>sumSS[4]<-sum(suqarederror)

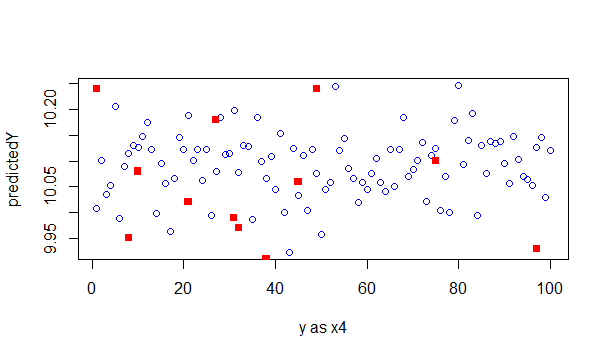
>sumSS[4]

[1] 89.35453

>plot(predictedY,xlab="y as x4" ,col="blue")



>points(sample$y,pch=15, col="red")



>p <- lm(y ~ sample$x5 , data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.098342 10.069768 10.008956 9.955471 10.141570 10.079293 10.084422 10.149629

9 10 11 12 13 14 15 16

10.189926 10.111530 10.036798 10.058045 10.251471 10.128382 10.129847 9.919571

17 18 19 20 21 22 23 24

10.066105 10.101273 10.111530 10.037531 10.042659 10.132045 10.179669 10.126184

25 26 27 28 29 30 31 32

10.092481 10.019947 10.053649 10.033867 10.169411 10.174540 10.142303 10.131313

33 34 35 36 37 38 39 40

10.060976 10.061709 10.154758 9.995036 10.075630 10.052184 10.028006 10.113728

41 42 43 44 45 46 47 48

10.172342 10.008956 10.151095 10.104936 10.117392 9.960600 10.239748 10.134243

49 50 51 52 53 54 55 56

10.062441 10.114461 10.177471 10.004560 10.183332 10.199451 10.082956 10.099075

57 58 59 60 61 62 63 64

10.000164 10.145233 10.130580 10.159154 10.069768 10.093946 10.110065 10.030204

65 66 67 68 69 70 71 72

10.110065 10.110065 10.160619 10.171609 10.102006 10.078560 9.954006 10.088818

73 74 75 76 77 78 79 80

10.239748 10.008956 10.059511 10.173075 10.041927 10.071966 10.085887 10.099075

81 82 83 84 85 86 87 88

9.995036 10.016283 10.033135 10.066105 10.140837 10.266124 10.002362 10.244877

89 90 91 92 93 94 95 96

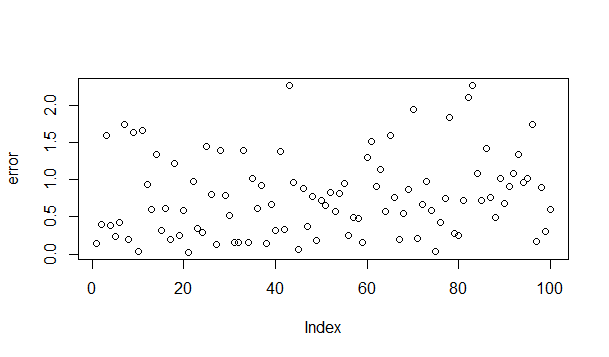
10.136441 10.045590 10.052184 10.046323 10.155491 10.009689 9.992105 10.110065

97 98 99 100

10.102006 10.055115 10.028006 10.034600

>error <- abs(sample$y - predictedY)

>plot(error)

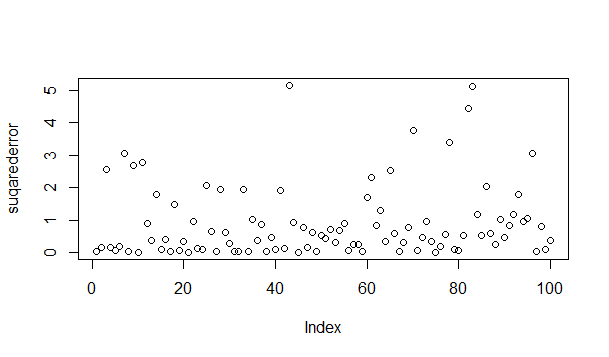


>error <- abs(sample$y - predictedY)

>plot(error)

>suqarederror<- error \* error

>plot(suqarederror)

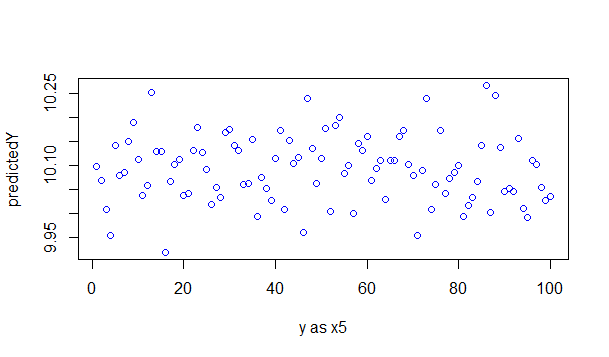


>sumSS[5]<-sum(suqarederror)

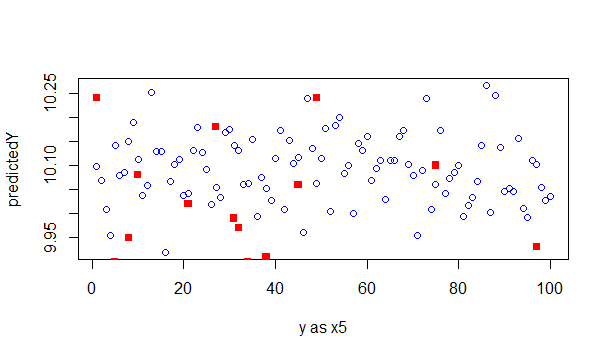
>sumSS[5]

[1] 89.25842

>plot(predictedY,xlab="y as x5" ,col="blue")



>points(sample$y,pch=15, col="red")



>p <- lm(y ~ sample$x6 , data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8 9

10.08993 10.07975 10.09282 10.09302 10.08454 10.08993 10.08793 10.08404 10.09492

10 11 12 13 14 15 16 17 18

10.08823 10.10141 10.09153 10.09153 10.10520 10.07436 10.11878 10.09941 10.08174

19 20 21 22 23 24 25 26 27

10.07815 10.09772 10.09312 10.08644 10.09542 10.10101 10.08893 10.10421 10.08833

28 29 30 31 32 33 34 35 36

10.12308 10.09552 10.06856 10.07186 10.10750 10.08943 10.09642 10.09782 10.06667

37 38 39 40 41 42 43 44 45

10.09642 10.09263 10.09123 10.08863 10.09402 10.10121 10.09382 10.08064 10.06697

46 47 48 49 50 51 52 53 54

10.10481 10.08673 10.09023 10.08494 10.08773 10.08763 10.09612 10.06467 10.08394

55 56 57 58 59 60 61 62 63

10.10710 10.07795 10.09522 10.10021 10.09692 10.09193 10.10870 10.10191 10.09822

64 65 66 67 68 69 70 71 72

10.08354 10.09233 10.08114 10.08314 10.09422 10.09722 10.08194 10.08843 10.09602

73 74 75 76 77 78 79 80 81

10.09223 10.11129 10.08094 10.08873 10.10031 10.09263 10.09422 10.10550 10.09572

82 83 84 85 86 87 88 89 90

10.09752 10.09682 10.11010 10.10610 10.08913 10.09712 10.08963 10.08434 10.07605

91 92 93 94 95 96 97 98 99

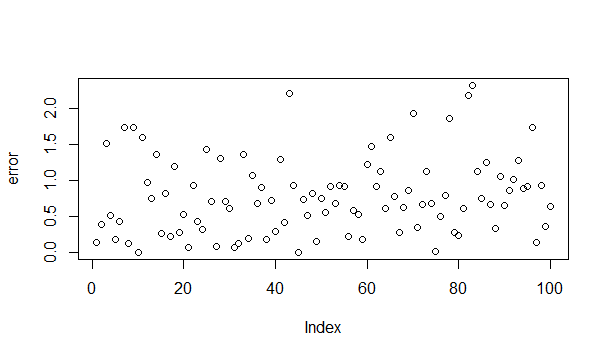
10.09862 10.10620 10.10540 10.09153 10.08673 10.09602 10.07925 10.09652 10.09312

100

10.08045

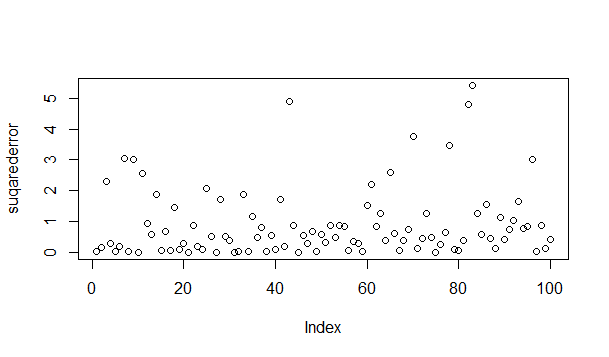
>error <- abs(sample$y - predictedY)

>plot(error)

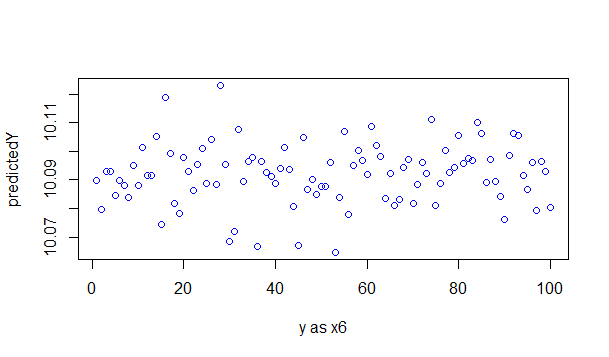


>suqarederror<- error \* error

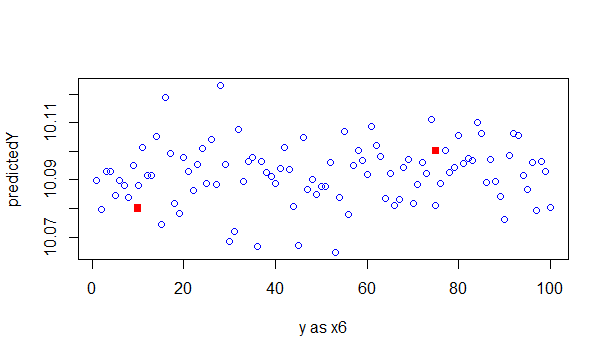
>plot(suqarederror)



>plot(predictedY,xlab="y as x6" , col="blue")



>points(sample$y,pch=15, col="red")



>p <- lm(y ~ sample$x1+sample$x2 , data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

9.973081 10.455380 10.902319 9.483832 10.012826 9.493553 11.801611 10.423379

9 10 11 12 13 14 15 16

11.460525 10.303175 11.583887 9.899403 10.904721 11.229731 10.383645 8.857879

17 18 19 20 21 22 23 24

9.731514 8.935980 10.335384 10.480796 10.292380 9.469526 11.025401 10.232572

25 26 27 28 29 30 31 32

8.957465 10.034692 9.630180 11.371594 9.724324 10.647588 10.415557 9.964424

33 34 35 36 37 38 39 40

11.641729 9.618937 11.542316 9.849892 10.833658 9.635449 9.212840 9.909308

41 42 43 44 45 46 47 48

9.190213 9.227623 7.704325 9.222539 10.104092 10.283516 10.900202 10.940653

49 50 51 52 53 54 55 56

10.477700 10.694371 9.783925 8.961810 10.968830 11.046158 10.978041 9.690554

57 58 59 60 61 62 63 64

9.279989 10.521085 10.564990 8.625960 11.377513 9.190039 9.314163 8.911594

65 66 67 68 69 70 71 72

11.702982 10.897660 10.371685 10.910673 9.114329 11.413489 9.367671 9.112711

73 74 75 76 77 78 79 80

11.405202 9.289917 9.743648 10.566468 9.178090 8.402065 10.252069 10.690026

81 82 83 84 85 86 87 88

10.882290 8.463860 8.512469 8.922967 10.855384 9.478043 10.712306 10.094881

89 90 91 92 93 94 95 96

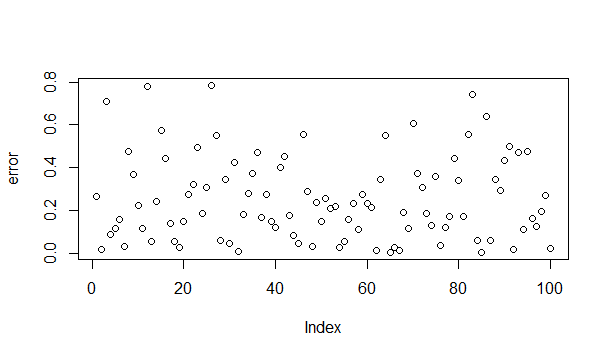
11.442243 10.298517 10.462788 11.148024 9.290656 10.868072 10.532132 8.524027

97 98 99 100

9.807062 8.967611 9.452571 9.453069

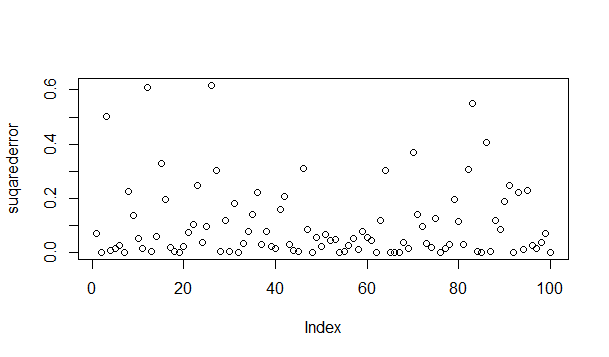
>error <- abs(sample$y - predictedY)

>plot(error)



>suqarederror<- error \* error

>plot(suqarederror)

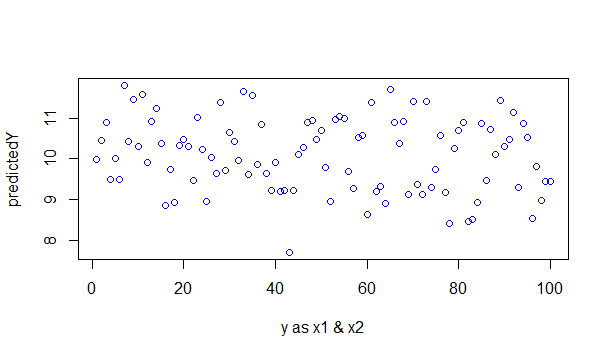


>sumSS[7]<-sum(suqarederror)

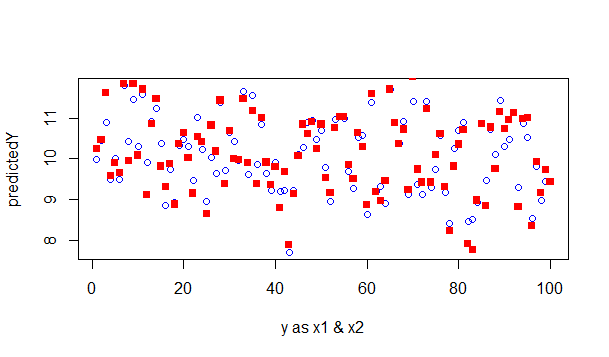
>sumSS[7]

[1] 10.12563

>plot(predictedY,xlab="y as x1 & x2" , col="blue")



>points(sample$y,pch=15, col="red")



>p <- lm(y ~ sample$x1+sample$x2+sample$x3, data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.044930 10.508544 11.157376 9.146924 10.326450 9.458507 11.593803 10.178725

9 10 11 12 13 14 15 16

11.855209 10.193595 11.468827 9.358906 11.208874 11.273151 10.047776 8.911643

17 18 19 20 21 22 23 24

9.535540 8.963839 10.492693 10.665234 10.135153 9.392170 10.686202 10.202675

25 26 27 28 29 30 31 32

8.920198 10.442894 9.972904 11.171075 9.528027 11.008144 10.342581 9.871906

33 34 35 36 37 38 39 40

11.380413 9.979146 11.091624 9.636380 10.936756 9.718533 9.308280 9.766007

41 42 43 44 45 46 47 48

9.119824 9.322684 7.689295 9.210724 10.066741 10.424830 10.840116 11.111736

49 50 51 52 53 54 55 56

10.195174 10.658444 9.668155 8.923519 11.236309 11.327866 11.068599 9.758878

57 58 59 60 61 62 63 64

9.243631 10.641567 10.286321 9.011919 11.378650 9.333422 9.087985 9.199405

65 66 67 68 69 70 71 72

11.643704 10.863409 10.448726 11.020291 9.140792 11.865417 9.521357 9.453260

73 74 75 76 77 78 79 80

11.286013 9.285525 10.170973 10.418583 9.020986 8.204308 9.997840 10.611466

81 82 83 84 85 86 87 88

10.404848 8.084289 8.112896 8.839126 10.863748 9.272343 10.668764 10.204580

89 90 91 92 93 94 95 96

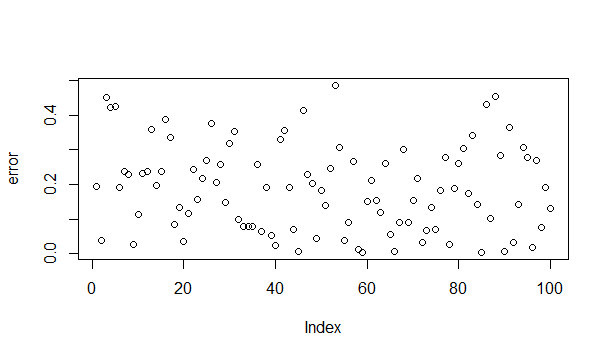
11.433405 10.724523 10.594547 11.096498 8.962186 10.673620 10.731601 8.378918

97 98 99 100

10.200335 9.235892 9.527758 9.560067

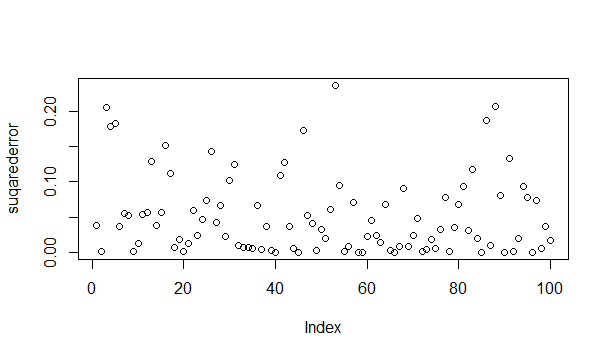
> error <- abs(sample$y - predictedY)

>plot(error)



>suqarederror<- error \* error

>plot(suqarederror)

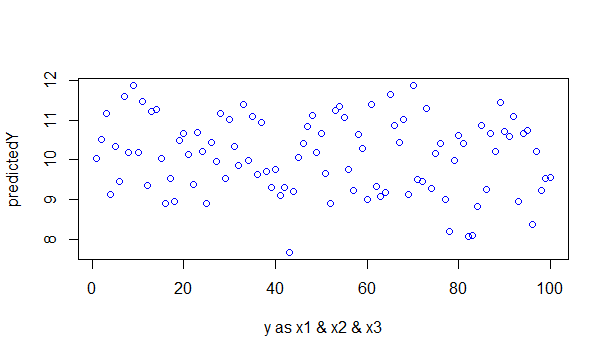


>sumSS[8]<-sum(suqarederror)

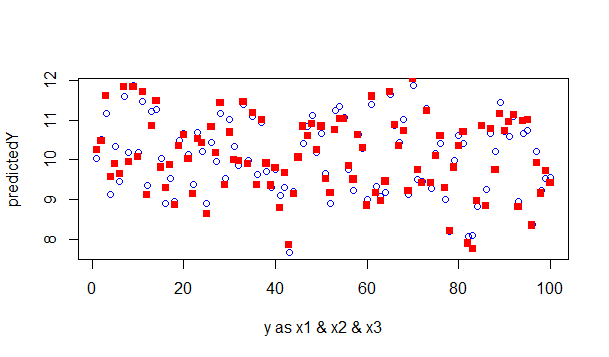
>sumSS[8]

[1] 5.113856

>plot(predictedY,xlab="y as x1 & x2 & x3" , col="blue")



>points(sample1$y,pch=15, col="red")



>p <- lm(y ~ sample$x1+sample$x2+sample$x3+sample$x4, data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.227419 10.500284 11.312818 9.206974 10.077698 9.666370 11.648667 10.132401

9 10 11 12 13 14 15 16

11.821322 10.123689 11.388801 9.157813 11.174186 11.518441 10.041189 8.950350

17 18 19 20 21 22 23 24

9.802846 8.984372 10.385927 10.612389 9.926374 9.350894 10.640716 10.272457

25 26 27 28 29 30 31 32

8.814428 10.664976 9.992439 11.004824 9.465587 10.982211 10.115185 9.898343

33 34 35 36 37 38 39 40

11.340474 9.890339 11.361133 9.418796 10.945426 9.764505 9.246119 9.862453

41 42 43 44 45 46 47 48

8.954366 9.501706 7.988047 9.112439 10.197171 10.390486 11.057279 11.076495

49 50 51 52 53 54 55 56

10.237760 10.972368 9.761392 8.960746 10.934306 11.299851 10.984976 9.761792

57 58 59 60 61 62 63 64

9.275026 10.816952 10.366895 9.080254 11.454525 9.280708 9.132371 9.282178

65 66 67 68 69 70 71 72

11.624512 10.977800 10.393465 10.844880 9.160753 11.933737 9.484246 9.332763

73 74 75 76 77 78 79 80

11.481180 9.216302 10.099197 10.626271 9.035432 8.348932 9.804431 10.286459

81 82 83 84 85 86 87 88

10.416667 7.917470 7.834593 9.016924 10.801716 9.285872 10.587063 10.114376

89 90 91 92 93 94 95 96

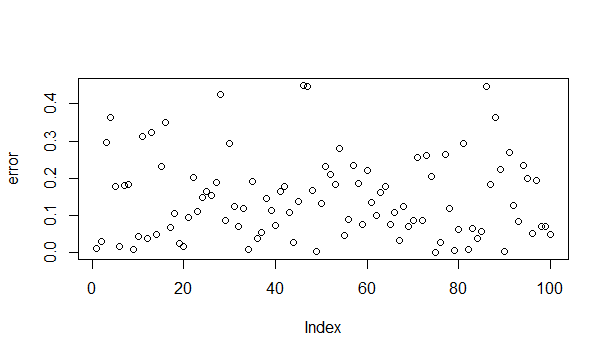
11.373635 10.731981 10.689812 11.003469 8.904966 10.744372 10.811064 8.412774

97 98 99 100

10.124968 9.088447 9.649941 9.477802

>error <- abs(sample$y - predictedY)

>plot(error)



>suqarederror<- error \* error

>plot(suqarederror)

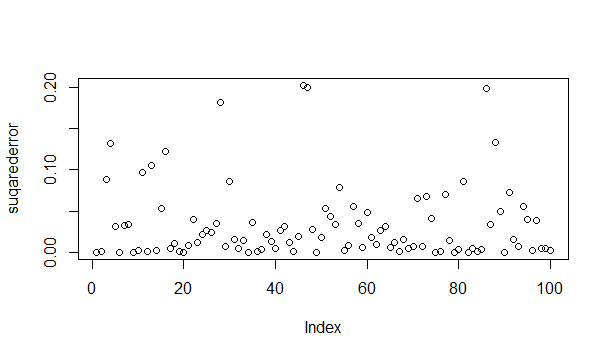
>sumSS[9]<-sum(suqarederror)

>

>

>sumSS[9]

[1] 3.337781

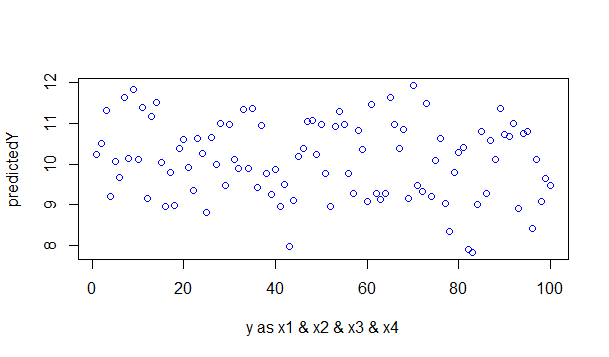


>sumSS[9]<-sum(suqarederror)

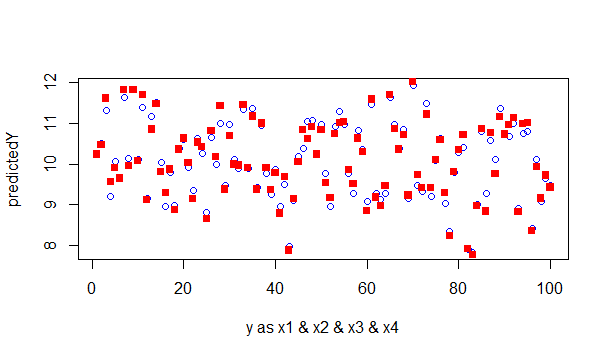
>sumSS[9]

[1] 3.337781

>plot(predictedY,xlab="y as x1 & x2 & x3 & x4" , col="blue")



>points(sample$y,pch=15, col="red")



>p <- lm(y ~ sample$x1+sample$x2+sample$x3+sample$x4+sample$x5, data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.202615 10.595451 11.532778 9.511657 9.930979 9.667144 11.743494 10.016755

9 10 11 12 13 14 15 16

11.688821 10.057423 11.579113 9.223222 10.853113 11.484476 9.982517 9.267034

17 18 19 20 21 22 23 24

9.847000 8.909296 10.357264 10.741152 10.045997 9.215286 10.470813 10.208189

25 26 27 28 29 30 31 32

8.741768 10.833849 10.078628 11.193164 9.276256 10.816073 9.971747 9.796150

33 34 35 36 37 38 39 40

11.493453 9.921958 11.274963 9.632363 10.994280 9.872386 9.391874 9.796591

41 42 43 44 45 46 47 48

8.721154 9.681527 7.796435 9.031907 10.140365 10.709330 10.769266 11.013185

49 50 51 52 53 54 55 56

10.286920 10.951111 9.547698 9.108825 10.737408 11.080584 11.030288 9.756657

57 58 59 60 61 62 63 64

9.445179 10.702154 10.276495 8.885922 11.543234 9.246486 9.060685 9.385459

65 66 67 68 69 70 71 72

11.670208 10.980695 10.240937 10.693851 9.102760 12.025604 9.803270 9.309418

73 74 75 76 77 78 79 80

11.207834 9.354569 10.172700 10.468418 9.103991 8.322642 9.806514 10.260823

81 82 83 84 85 86 87 88

10.667494 8.007255 7.884662 9.027376 10.724529 8.853740 10.843645 9.750976

89 90 91 92 93 94 95 96

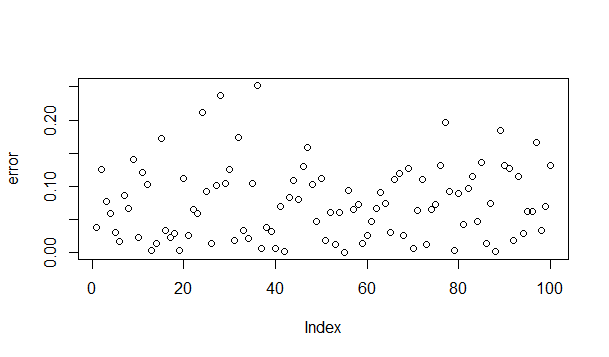
11.333922 10.861761 10.833488 11.111559 8.705522 11.008848 11.072010 8.298469

97 98 99 100

10.096618 9.127290 9.790186 9.561051

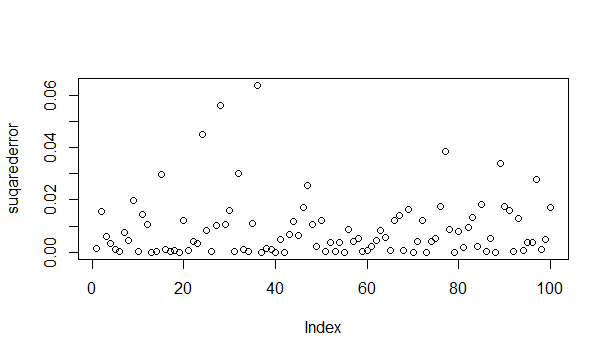
>error <- abs(sample$y - predictedY)

>plot(error)



>suqarederror<- error \* error

>plot(suqarederror)



>suqarederror<- error \* error

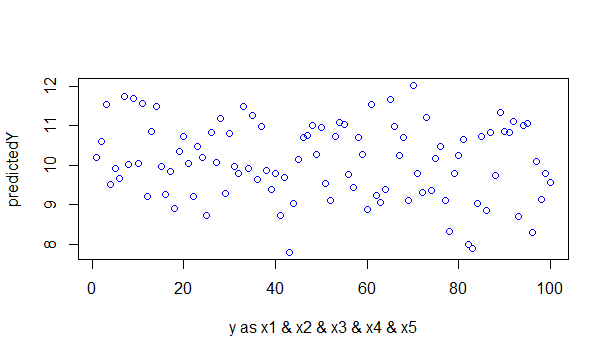
>plot(suqarederror)

>sumSS[10]<-sum(suqarederror)

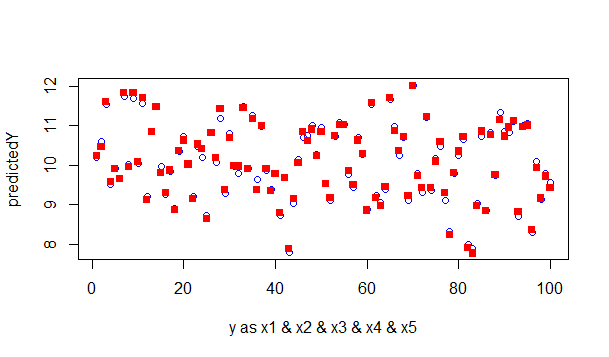
>sumSS[10]

[1] 0.8723916

>plot(predictedY,xlab="y as x1 & x2 & x3 & x4 & x5" , col="blue")



>points(sample$y,pch=15, col="red")



>p <- lm(y ~ sample$x1+sample$x2+sample$x3+sample$x4+sample$x5+sample$x6, data=sample, x=T, y=T)

>predictedY<-predict(p,sample)

>predictedY

1 2 3 4 5 6 7 8

10.179443 10.546297 11.512742 9.480965 9.918352 9.635997 11.711922 9.985629

9 10 11 12 13 14 15 16

11.754200 10.029026 11.614672 9.208576 10.900465 11.539312 9.900366 9.348409

17 18 19 20 21 22 23 24

9.851306 8.858021 10.308550 10.760710 10.049890 9.188157 10.492959 10.256653

25 26 27 28 29 30 31 32

8.722046 10.877797 10.070692 11.338272 9.309399 10.730894 9.876759 9.871618

33 34 35 36 37 38 39 40

11.475194 9.946580 11.278584 9.497345 11.004332 9.883641 9.400087 9.768205

41 42 43 44 45 46 47 48

8.752538 9.711747 7.817878 8.978888 10.011799 10.755690 10.763199 11.019210

49 50 51 52 53 54 55 56

10.215787 10.911685 9.530734 9.104575 10.640441 11.067057 11.104128 9.704599

57 58 59 60 61 62 63 64

9.435180 10.738213 10.281962 8.917984 11.605292 9.308141 9.085452 9.340017

65 66 67 68 69 70 71 72

11.682716 10.925479 10.214218 10.735344 9.132904 11.984146 9.780787 9.353787

73 74 75 76 77 78 79 80

11.224813 9.427816 10.137251 10.452057 9.124347 8.304418 9.812380 10.332169

81 82 83 84 85 86 87 88

10.645514 8.011641 7.895223 9.095139 10.808910 8.864768 10.865786 9.774621

89 90 91 92 93 94 95 96

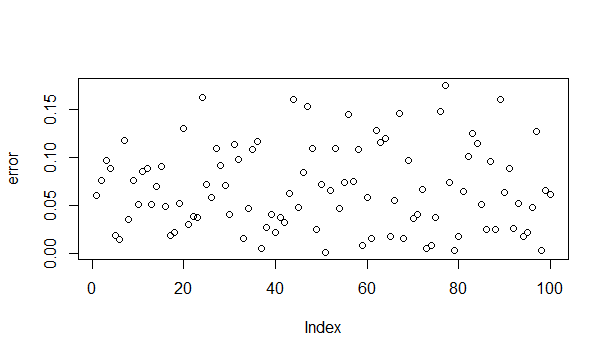
11.310800 10.793553 10.871903 11.155403 8.768227 10.997608 11.031472 8.311889

97 98 99 100

10.057378 9.162426 9.785520 9.491326

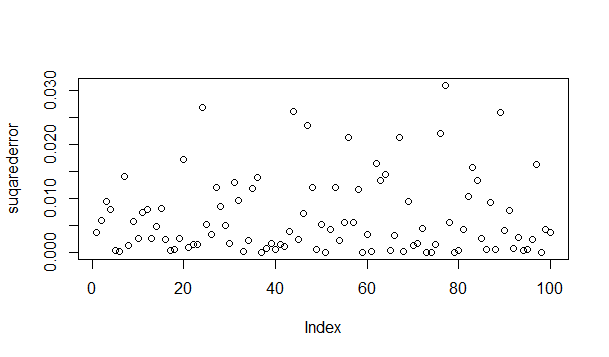
>error <- abs(sample$y - predictedY)

>plot(error)



>suqarederror<- error \* error

>plot(suqarederror)

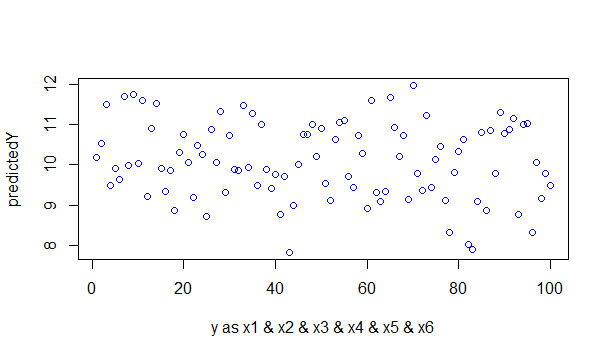


>sumSS[11]<-sum(suqarederror)

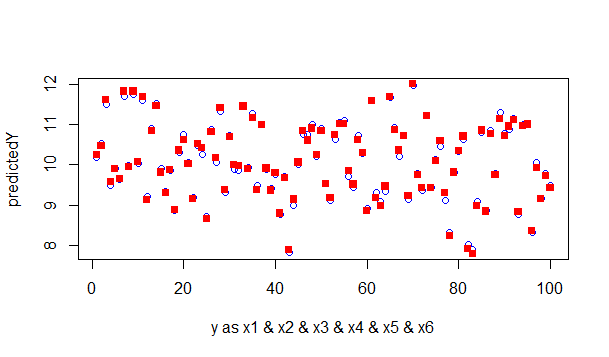
>sumSS[11]

[1] 0.6432416

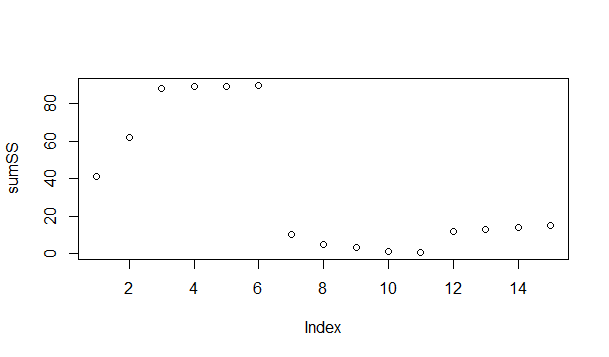
>plot(predictedY,xlab="y as x1 & x2 & x3 & x4 & x5 & x6" , col="blue")



>points(sample$y,pch=15, col="red")



>plot(sumSS)



>TrsumSS<-0

>TssumSS<-0

>TrMSE<-0

>TsMSE<-0

>TsMSE<-0

>p<-0

>i<-1

>for ( i in 1:10)

+ {

+ p=10\*i

+ p

+ dim(sample)

+ TrainingData<- head(sample,p)

+ TrainingData

+ dim(TrainingData)

+ lmodel\_trainig\_data<- lm(y ~ x1 + x2+ x3 + x4 + x5 + x6 + x7, data=TrainingData, x=T, y=T)

+ predictedY<-predict(lmodel\_trainig\_data,TrainingData)

+ predictedY

+ TrainingData$y

+ TestData<- head(sample,-p )

+ Trerror<- abs(TrainingData$y - predictedY)

+ Trerror

+ Trsuqarederror<- Trerror \* Trerror

+ TrsumSS[i]<-sum(Trsuqarederror)

+ TrsumSS[i]

+ TrMSE[i]<-mean(Trsuqarederror)

+ predictedY<-predict(lmodel\_trainig\_data,TestData)

+ predictedY

+ TestData$y

+ Tserror<- abs(TestData$y - predictedY)

+ Tserror

+ Tssuqarederror<- Tserror \* Tserror

+ TssumSS[i]<-sum(Tssuqarederror)

+ TssumSS[i]

+ TsMSE[i]<-mean(Tssuqarederror)

+ }

>TrsumSS

[1] 0.001145062 0.065073632 0.126748120 0.181695975 0.271780808 0.310115347

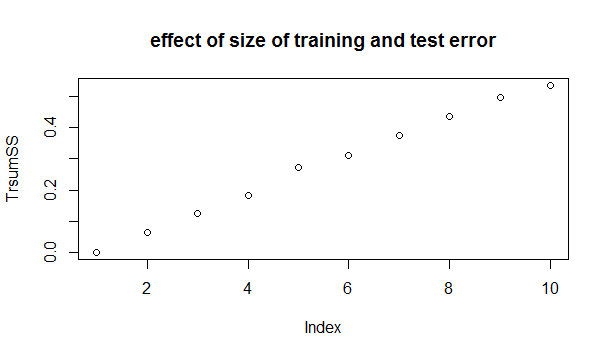
[7] 0.375726569 0.435561134 0.496046827 0.534387207

>TssumSS

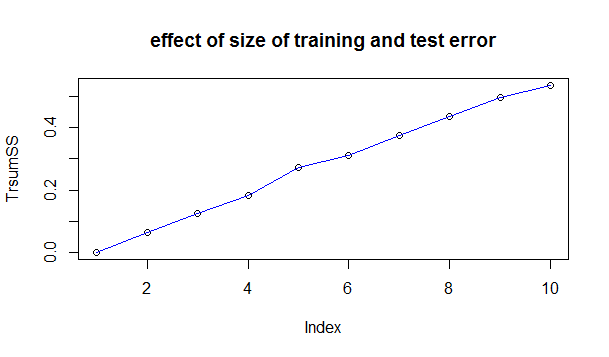
[1] 1.80001297 0.51670065 0.41019613 0.33922189 0.27178081 0.20249314 0.14423624

[8] 0.07960180 0.02276229 0.00000000

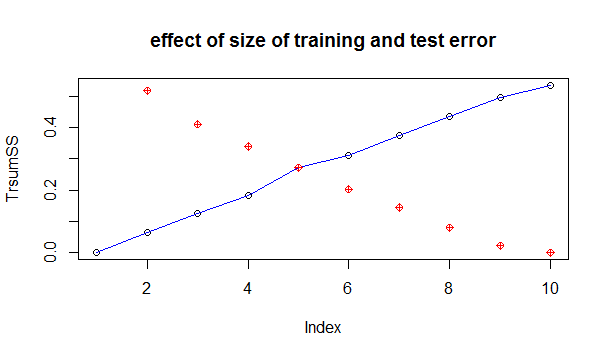
>plot(TrsumSS, main="effect of size of training and test error")



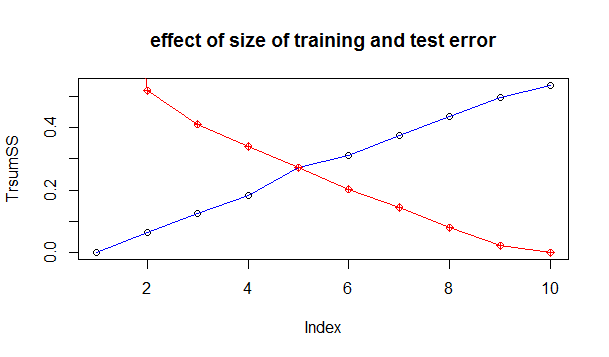
>lines(x=TrsumSS, y = NULL, type = "l", col="blue")



>points(TssumSS,pch=10,col="red")



>lines(x=TssumSS, y = NULL, type = "l", col="red")



>TrMSE

[1] 0.0001145062 0.0032536816 0.0042249373 0.0045423994 0.0054356162 0.0051685891

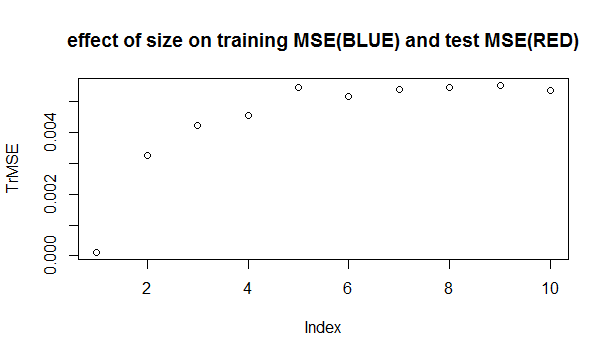
[7] 0.0053675224 0.0054445142 0.0055116314 0.0053438721

>TsMSE

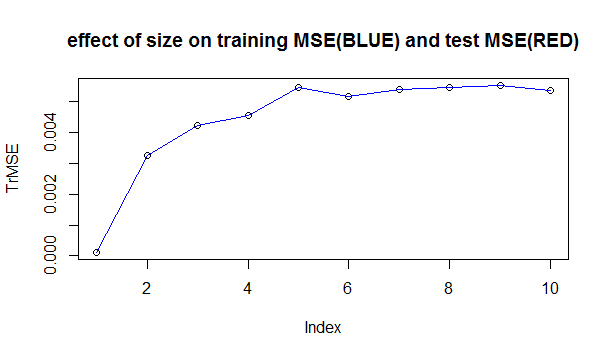
[1] 0.020000144 0.006458758 0.005859945 0.005653698 0.005435616 0.005062328

[7] 0.004807875 0.003980090 0.002276229 NaN

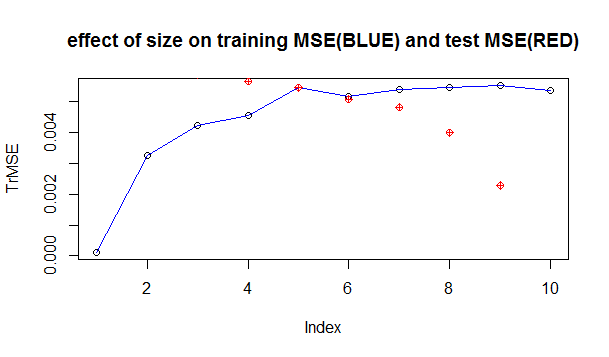
>plot(TrMSE, main="effect of size on training MSE(BLUE) and test MSE(RED)")



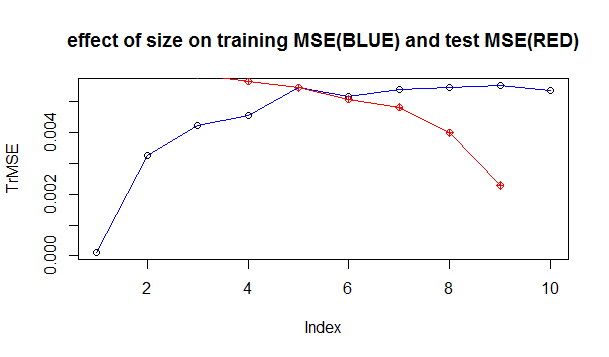
>lines(x=TrMSE, y = NULL, type = "l", col="blue")



>points(TsMSE,pch=10,col="red")



>lines(x=TsMSE, y = NULL, type = "l", col="red")



# Cross Validation using the boot method

# val.boot<- validate(lm1)

# Using the cross validation method

# val.cv <- validate(lm1, method = â€œcrossvalidationâ€)

# Load the Data Analysis And Graphic (DAAG) library

> library(DAAG)

Loading required package: lattice

> library(ISLR)

>val.daag<- CVlm(df=sample1.df, m=2,

+ form.lm=formula(y ~ x1 + x2 +

+ x3 + x4 + x5 + x6 + x7))

Analysis of Variance Table

Response: y

Df Sum Sq Mean Sq F value Pr(>F)

x1 1 48.7 48.7 8631.8 < 2e-16 \*\*\*

x2 1 30.8 30.8 5471.0 < 2e-16 \*\*\*

x3 1 5.1 5.1 900.1 < 2e-16 \*\*\*

x4 1 1.8 1.8 312.4 < 2e-16 \*\*\*

x5 1 2.5 2.5 438.3 < 2e-16 \*\*\*

x6 1 0.2 0.2 41.2 6.0e-09 \*\*\*

x7 1 0.1 0.1 19.4 2.8e-05 \*\*\*

Residuals 92 0.5 0.0

---

Signif. codes: 0 â€˜\*\*\*â€™ 0.001 â€˜\*\*â€™ 0.01 â€˜\*â€™ 0.05 â€˜.â€™ 0.1 â€˜ â€™ 1

fold 1

Observations in test set: 50

1 2 4 7 9 10 15 17 19 20 21 23

Predicted 10.219 10.527 9.5063 11.771 11.7877 10.0604 9.8778 9.8580 10.3109 10.76 10.0062 10.5

cvpred 10.214 10.542 9.4772 11.715 11.7451 10.0342 9.8660 9.8468 10.3202 10.77 10.0032 10.4

y 10.239 10.468 9.5650 11.827 11.8287 10.0821 9.8055 9.8689 10.3614 10.63 10.0176 10.5

CV residual 0.025 -0.074 0.0879 0.112 0.0836 0.0479 -0.0605 0.0222 0.0412 -0.14 0.0144 0.1

24 25 26 28 29 30 31 32 33 38 41 43 47

Predicted 10.278 8.7167 10.9 11.385 9.3635 10.7064 9.8947 9.896 11.4702 9.9149 8.8308 7.8204 10.77

cvpred 10.235 8.7206 10.9 11.275 9.2829 10.7622 9.8968 9.844 11.4411 9.8857 8.7428 7.8016 10.74

y 10.417 8.6497 10.8 11.431 9.3787 10.6913 9.9853 9.969 11.4598 9.9063 8.7942 7.8846 10.61

CV residual 0.182 -0.0709 -0.1 0.156 0.0958 -0.0709 0.0885 0.125 0.0187 0.0206 0.0514 0.0829 -0.13

48 50 52 55 58 59 64 67 68 69 71 72

Predicted 11.030 10.8745 9.1436 11.0908 10.712 10.2285 9.4175 10.210 10.75851 9.155 9.8031 9.39

cvpred 11.020 10.9154 9.1271 11.0780 10.726 10.2335 9.4161 10.215 10.71084 9.124 9.8044 9.37

y 10.913 10.8368 9.1711 11.0314 10.626 10.2909 9.4606 10.357 10.71996 9.230 9.7438 9.42

CV residual -0.107 -0.0786 0.0439 -0.0466 -0.101 0.0574 0.0445 0.142 0.00912 0.106 -0.0607 0.05

75 76 78 80 81 82 84 85 90 91 93 95

Predicted 10.151 10.51 8.2525 10.3325 10.6557 8.0183 9.0478 10.8508 10.790 10.834 8.783 11.019

cvpred 10.181 10.46 8.2804 10.2864 10.6157 7.9794 9.0546 10.7765 10.849 10.846 8.715 11.061

y 10.102 10.60 8.2295 10.3483 10.7054 7.9086 8.9810 10.8582 310.734 10.955 8.821 11.013

CV residual -0.079 0.14 -0.0509 0.0619 0.0898 -0.0708 -0.0735 0.0817 -0.115 0.109 0.106 -0.048

97

Predicted 10.029

cvpred 10.080

y 9.934

CV residual -0.146

Sum of squares = 0.4 Mean square = 0.01 n = 50

fold 2

Observations in test set: 50

3 5 6 8 11 12 13 14 16 18 22 27

Predicted 11.522 9.861 9.6256 9.9660 11.585 9.1776 10.87834 11.536 9.3647 8.836 9.257 10.054

cvpred 11.491 9.797 9.6098 9.9758 11.576 9.1944 10.86223 11.531 9.3339 8.812 9.282 10.016

y 11.607 9.905 9.6473 9.9490 11.698 9.1206 10.85438 11.466 9.3040 8.883 9.154 10.182

CV residual 0.116 0.108 0.0375 -0.0268 0.122 -0.0738 -0.00784 -0.065 -0.0299 0.071 -0.128 0.166

34 35 36 37 39 40 42 44 45 46 49 51

Predicted 9.9002 11.27 9.531 10.9752 9.3924 9.7372 9.7063 8.972 10.0225 10.727 10.2241 9.510

cvpred 9.8336 11.30 9.546 10.9381 9.3834 9.7264 9.6919 8.954 10.0222 10.692 10.2264 9.507

y 9.9035 11.17 9.382 11.0038 9.3647 9.7899 9.6782 9.136 10.0593 10.845 10.2423 9.533

CV residual 0.0699 -0.13 -0.165 0.0657 -0.0187 0.0635 -0.0137 0.182 0.0371 0.153 0.0158 0.026

53 54 56 57 60 61 62 63 65 66 70 73

Predicted 10.7014 11.0693 9.715 9.5098 8.8863 11.709 9.2612 9.068 11.6247 10.9038 11.9772 11.2151

cvpred 10.6996 11.0490 9.721 9.5307 8.8468 11.744 9.2290 9.074 11.6112 10.8960 11.9326 11.2352

y 10.7534 11.0216 9.847 9.5148 8.8619 11.593 9.1784 8.966 11.6992 10.8703 12.0192 11.2197

CV residual 0.0538 -0.0274 0.126 -0.0159 0.0151 -0.151 -0.0506 -0.108 0.0881 -0.0258 0.0866 -0.0155

74 77 79 83 86 87 88 89 92 94 96

Predicted 9.4024 9.115 9.80963 7.8388 8.8744 10.8578 9.75766 11.279 11.0994 10.9846 8.3152

cvpred 9.3706 9.105 9.81672 7.8226 8.9040 10.8586 9.74563 11.269 11.0522 10.9988 8.3173

y 9.4207 9.301 9.81473 7.7698 8.8425 10.7678 9.75071 11.155 11.1285 10.9844 8.3574

CV residual 0.0501 0.196 -0.00199 -0.0529 -0.0615 -0.0908 0.00508 -0.115 0.0763 -0.0143 0.0402

98 99 100

Predicted 9.181804 9.759 9.4798

cvpred 9.161371 9.734 9.4447

y 9.162332 9.717 9.4293

CV residual 0.000961 -0.017 -0.0154

Sum of squares = 0.37 Mean square = 0.01 n = 50

Overall (Sum over all 50 folds)

ms

0.0077

Warning message:

In CVlm(df = sample1.df, m = 2, form.lm = formula(y ~ x1 + x2 + :

As there is >1 explanatory variable, cross-validation

predicted values for a fold are not a linear function

of corresponding overall predicted values. Lines that

are shown for the different folds are approximate

> for (i in 1:10)

+ {

+ val.daag<- CVlm(df=sample1.df, m = i, form.lm=formula(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7))

+ cv.error.10[i]

+ }

Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :

0 (non-NA) cases

>cverr = val.daag$y - val.daag$cvpred

> plot(x = sample1.df$s.id, y = cverr, type ='l')

>val.daag<- CVlm(df=sample1.df, m=4, form.lm=formula(y ~ x1 + x2 + x3 + x4 + x5 + x6 + x7))

Analysis of Variance Table

Response: y

Df Sum Sq Mean Sq F value Pr(>F)

x1 1 48.7 48.7 8631.8 < 2e-16 \*\*\*

x2 1 30.8 30.8 5471.0 < 2e-16 \*\*\*

x3 1 5.1 5.1 900.1 < 2e-16 \*\*\*

x4 1 1.8 1.8 312.4 < 2e-16 \*\*\*

x5 1 2.5 2.5 438.3 < 2e-16 \*\*\*

x6 1 0.2 0.2 41.2 6.0e-09 \*\*\*

x7 1 0.1 0.1 19.4 2.8e-05 \*\*\*

Residuals 92 0.5 0.0

---

Signif. codes: 0 â€˜\*\*\*â€™ 0.001 â€˜\*\*â€™ 0.01 â€˜\*â€™ 0.05 â€˜.â€™ 0.1 â€˜ â€™ 1

fold 1

Observations in test set: 25

2 4 9 15 17 19 23 25 26 30 32 38

Predicted 10.5271 9.506 11.788 9.8778 9.85803 10.3109 10.4858 8.717 10.9107 10.7064 9.8956 9.91493

cvpred 10.5310 9.494 11.790 9.8762 9.86243 10.3177 10.4782 8.722 10.9239 10.7305 9.8913 9.91134

y 10.4677 9.565 11.829 9.8055 9.86895 10.3614 10.5308 8.650 10.8242 10.6913 9.9687 9.90626

CV residual -0.0633 0.071 0.039 -0.0708 0.00652 0.0436 0.0525 -0.072 -0.0997 -0.0392 0.0774 -0.00508

47 48 58 69 71 75 76 80 81 84 85 91

Predicted 10.767 11.030 10.712 9.1547 9.8031 10.1506 10.5089 10.3325 10.6557 9.0478 10.8508 10.834

cvpred 10.777 11.036 10.727 9.1546 9.7997 10.1629 10.5106 10.3212 10.6411 9.0535 10.8413 10.837

y 10.612 10.913 10.626 9.2300 9.7438 10.1024 10.5950 10.3483 10.7054 8.9810 10.8582 10.955

CV residual -0.165 -0.123 -0.101 0.0754 -0.0559 -0.0605 0.0844 0.0271 0.0643 -0.0725 0.0169 0.118

95

Predicted 11.0194

cvpred 11.0295

y 11.0132

CV residual -0.0163

Sum of squares = 0.14 Mean square = 0.01 n = 25

fold 2

Observations in test set: 25

5 6 8 11 22 27 35 39 40 42 45 46

Predicted 9.8614 9.626 9.9660 11.585 9.257 10.054 11.267 9.39241 9.7372 9.7063 10.023 10.727

cvpred 9.8317 9.625 9.9643 11.570 9.279 10.034 11.292 9.37117 9.7345 9.6966 10.028 10.699

y 9.9049 9.647 9.9490 11.698 9.154 10.182 11.169 9.36474 9.7899 9.6782 10.059 10.845

CV residual 0.0732 0.022 -0.0154 0.128 -0.125 0.148 -0.123 -0.00643 0.0554 -0.0184 0.031 0.146

49 53 54 57 60 61 62 73 77 79 86 96

Predicted 10.2241 10.7014 11.0693 9.50982 8.886 11.709 9.2612 11.2151 9.115 9.8096 8.8744 8.3152

cvpred 10.2304 10.7096 11.0760 9.52392 8.873 11.745 9.2406 11.2390 9.105 9.8042 8.8997 8.3176

y 10.2423 10.7534 11.0216 9.51479 8.862 11.593 9.1784 11.2197 9.301 9.8147 8.8425 8.3574

CV residual 0.0118 0.0438 -0.0544 -0.00913 -0.011 -0.152 -0.0622 -0.0192 0.196 0.0105 -0.0572 0.0399

99

Predicted 9.7588

cvpred 9.7412

y 9.7173

CV residual -0.0239

Sum of squares = 0.18 Mean square = 0.01 n = 25

fold 3

Observations in test set: 25

1 7 10 20 21 24 28 29 31 33 41 43

Predicted 10.219 11.7714 10.060 10.763 10.00623 10.278 11.3849 9.3635 9.8947 11.47018 8.83075 7.8204

cvpred 10.205 11.7517 10.047 10.770 10.01643 10.262 11.3700 9.3349 9.8893 11.46772 8.79882 7.7984

y 10.239 11.8271 10.082 10.632 10.01757 10.417 11.4315 9.3787 9.9853 11.45981 8.79424 7.8846

CV residual 0.034 0.0754 0.035 -0.138 0.00114 0.154 0.0614 0.0438 0.0961 -0.00791 -0.00458 0.0861

50 52 55 59 64 67 68 72 78 82 90 93

Predicted 10.8745 9.1436 11.0908 10.229 9.417 10.210 10.7585 9.3912 8.2525 8.018 10.7896 8.7830

cvpred 10.8749 9.1358 11.0952 10.227 9.407 10.206 10.7465 9.3859 8.2534 8.011 10.8041 8.7601

y 10.8368 9.1711 11.0314 10.291 9.461 10.357 10.7200 9.4207 8.2295 7.909 10.7338 8.8210

CV residual -0.0381 0.0353 -0.0638 0.064 0.054 0.151 -0.0265 0.0348 -0.0239 -0.102 -0.0703 0.0608

97

Predicted 10.029

cvpred 10.041

y 9.934

CV residual -0.108

Sum of squares = 0.14 Mean square = 0.01 n = 25

fold 4

Observations in test set: 25

3 12 13 14 16 18 34 36 37 44 51 56

Predicted 11.522 9.178 10.87834 11.5357 9.3647 8.836 9.9002 9.531 10.9752 8.972 9.5098 9.715

cvpred 11.505 9.222 10.86283 11.5177 9.3649 8.836 9.8894 9.566 10.9675 8.979 9.5069 9.723

y 11.607 9.121 10.85438 11.4660 9.3040 8.883 9.9035 9.382 11.0038 9.136 9.5326 9.847

CV residual 0.102 -0.102 -0.00844 -0.0517 -0.0609 0.047 0.0142 -0.184 0.0362 0.157 0.0257 0.124

63 65 66 70 74 83 87 88 89 92 94 98

Predicted 9.068 11.6247 10.9038 11.9772 9.4024 7.839 10.858 9.75766 11.28 11.0994 10.9846 9.1818

cvpred 9.080 11.6367 10.9032 11.9554 9.4106 7.880 10.881 9.74904 11.28 11.1013 11.0103 9.1851

y 8.966 11.6992 10.8703 12.0192 9.4207 7.770 10.768 9.75071 11.15 11.1285 10.9844 9.1623

CV residual -0.115 0.0626 -0.0329 0.0638 0.0101 -0.111 -0.113 0.00167 -0.13 0.0273 -0.0259 -0.0228

100

Predicted 9.4798

cvpred 9.4809

y 9.4293

CV residual -0.0516

Sum of squares = 0.17 Mean square = 0.01 n = 25

Overall (Sum over all 25 folds)

ms

0.00634

Warning message:

In CVlm(df = sample1.df, m = 4, form.lm = formula(y ~ x1 + x2 + :

As there is >1 explanatory variable, cross-validation

predicted values for a fold are not a linear function

of corresponding overall predicted values. Lines that

are shown for the different folds are approximate

> cverr1 = val.daag$y - val.daag$cvpred

> z=cverr1

> plot(x = sample1.df$s.id, y = cverr, type ='l')

> x

> y

1 2 3 5 9 10 11 12 14 15 19 20 21 22 23 24 26 27 29 31 35 37 40 42 43 44

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

45 47 49 52 59 64 68 69 70 74 75 76 77 80 82 84 85 88 90 92 94 96 97 100

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0