В качестве нелинейной функции для моделирования выбрана: $(x+a)^2 + b * x^3$
С параметрами: $a = 1,\ b\ =\ 2,\ eps = 4$

2. Оценим параметры линейной модели численно

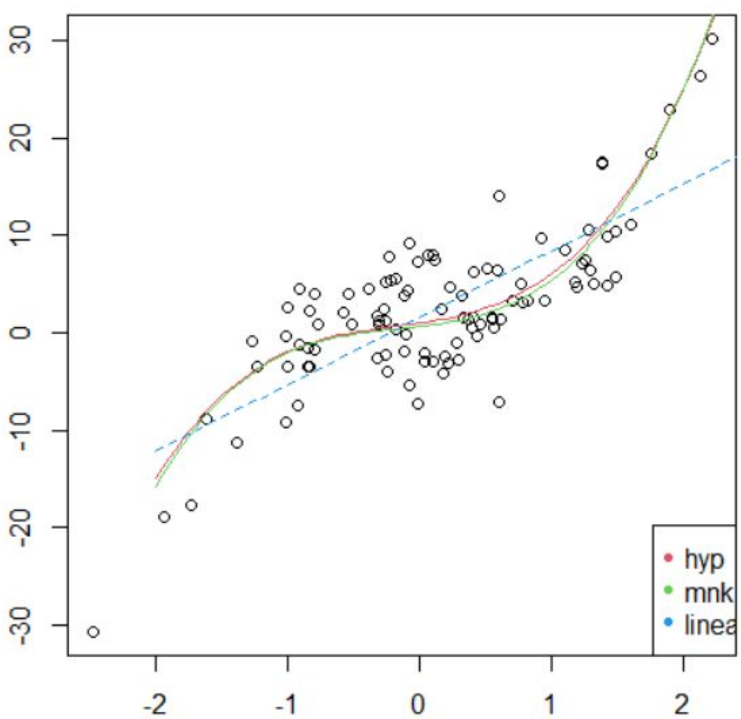
```
1.532381 6.858575
```

Вычисленные параметры линейной модели

Оценим параметры нелинейной модели

ab.	ab
0.7992796	1
2.1467934	2

Вычисленные параметры нелинейной модели (ab.) и оригинальные (ab)



Графики моделей

Q.linear	Q.model	Q.model.hat
2729.516	1687.947	1669.944

Невязки моделей

Предсказуемо, линейная модель из всех представленных показывает худший результат, поскольку плохо подходит для моделирования нелинейно распределённых данных с нормальным шумом. В то же время, стоит обратить внимание, что нелинейная модель с оценёнными параметрами показывает лучший результат при применении к данным с нормально распределённым шумом. Причиной тому является тот факт, что изначальные параметры нелинейной модели не учитывают шум, а функция потерь применяется именно на таких данных.

3. Дисперсионный анализ

```
> QT<-sum(( Y-mean(Y) )^2);QT
[1] 6673.484
> QR<-sum((Y.-mean(Y))^2);QR
[1] 3943.968
> QE<-sum((Y-Y.)^2);QE
[1] 2729.516
> R2<-QR/QT;R2
[1] 0.5909909
```

Вычисляем источники вариации (общий, обусл. регрессией и невязка)
и коэффициент детерминации

Проверим гипотезы о значимости прогноза и оценки коэффициентов регресии.

```
[1] -3.070413e-17
> F.<-QR/QT*(N-2);F.
[1] 57.91711
> Pf<-1-pf(F.,2,N-2);Pf
[1] 0
> Ta<-AB[1]/sqrt(S2a);Ta
       X
2.876625
> Tb<-AB[2]/sqrt(S2b);Tb

11.89973
> Pa<-2*(1-pt(abs(Ta),N-2));Pa
          X
0.004932683
> Pb<-2*(1-pt(abs(Tb),N-2));Pb

0
```

Убеждаемся, что оценкам коэффициентов регрессии и значимости можно доверять (p-values < 0.05). Можно сделать вывод, что оценка прошла успешно.

Проверим оценку при помощи встроенных функций.

```
              [,1]         [,2]
Pf   0.000000e+00 0.000000e+00
Pa.X 4.596642e-07 4.596642e-07
Pb   0.000000e+00 1.552455e-28
```

Таблица сравнения p-values ручной оценкой ([, 1]) и встроенной функции ([, 2])

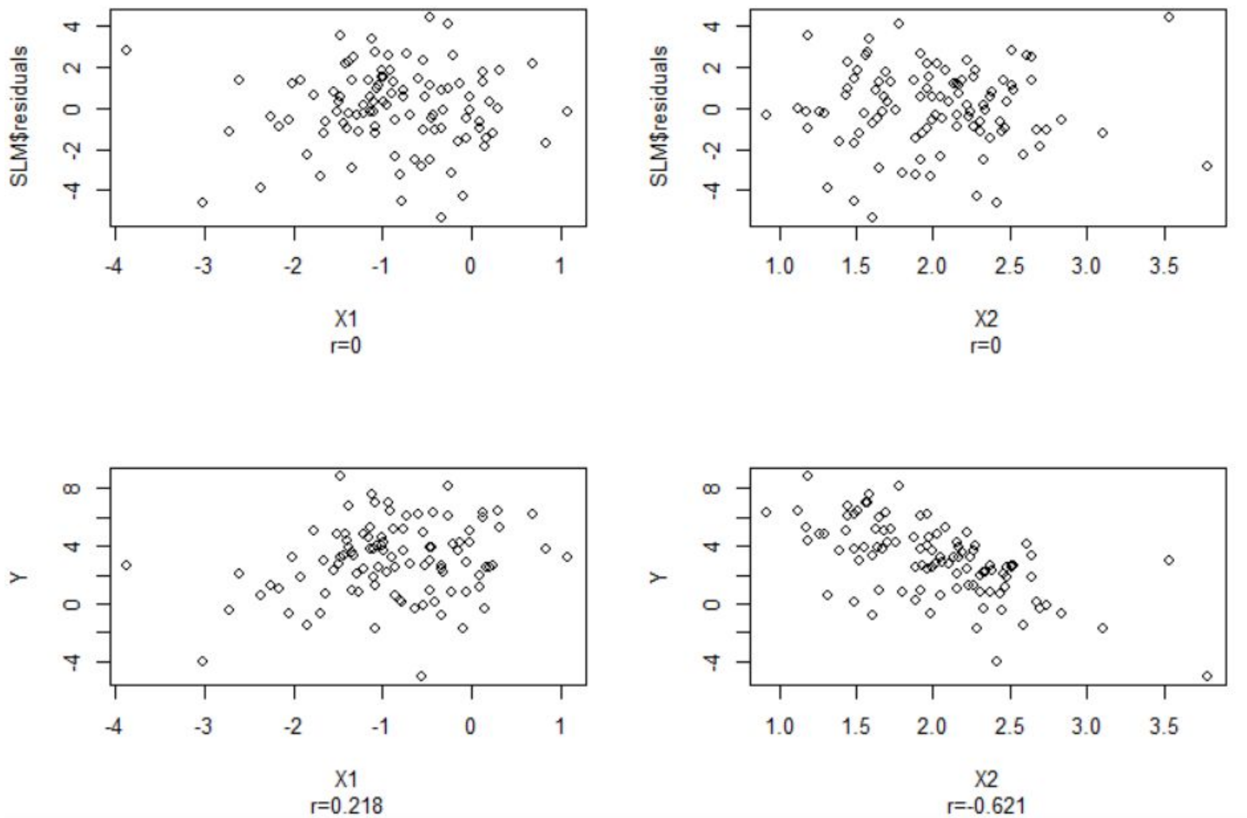
По таблице сравнения можно убедиться, что оценка в обоих случаях позволяет сделать одинаковый вывод, при чём, значения получены с большой точностью (ошибка не более чем 1.e-13).

4.

```
Call:
lm(formula = Y ~ X1 + X2)

Residuals:
    Min      1Q  Median      3Q     Max 
-5.3255 -1.0445 -0.0133  1.3133  4.4684 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   9.6933     0.8365  11.588  < 2e-16 ***
X1            0.5287     0.2383   2.219   0.0288 *  
X2           -3.0865     0.3953  -7.808 6.84e-12 ***
---
Signif. codes:  
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.955 on 97 degrees of freedom
Multiple R-squared:  0.4152,	Adjusted R-squared:  0.4032 
F-statistic: 34.44 on 2 and 97 DF,  p-value: 4.996e-12
```

Результаты применения lm-функции

Как видно на результатах, p-value по распр. Фишера принимает значение < 0.05, что позволяет судить о том, что существенного различия между выборочным коэффициентом детерминации и коэффициентом детерминации генеральной совокупности нет.



Представление коррелляции между остатком и независ. Переменными

По результатам оценки корреляции можно судить о слабой связи (корреляции) между свободными переменными и остатками.