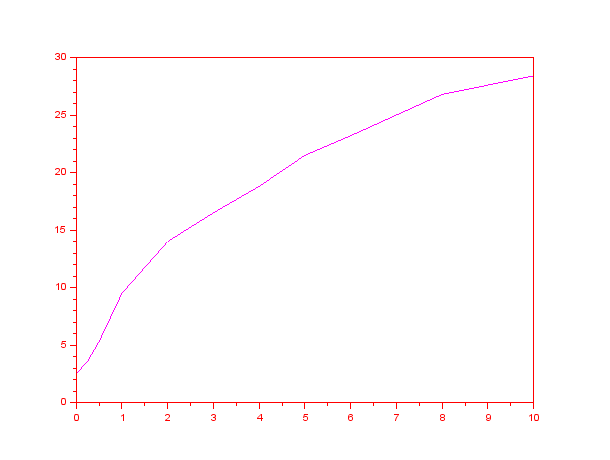
**Exercises Curve Fitting**

Fahmy et al measured the total amount of drug released from a nanoparticle over time. Some of their results are provided in the table below:

|  |  |
| --- | --- |
| Day | Micrograms drug released |
| 0 | 2.5 |
| 0.25 | 3.6 |
| 0.5 | 5.3 |
| 1 | 9.5 |
| 2 | 14.0 |
| 3 | 16.5 |
| 4 | 18.8 |
| 5 | 21.5 |
| 6 | 23.2 |
| 8 | 26.8 |
| 10 | 28.4 |



1. Answer the following questions
2. Create a linear model . Show the design matrix. Calculate the sum of squares and the determination coefficient (correlation coefficient R2)

**Answer:**

-->m=size(x,1)

m = 11.

-->X=[ones(m,1) x]

-->a=inv(X'\*X)\*(X'\*y);

-->a

a =

5.8796123

2.6521828

-->yf=a(1)+a(2).\*x;

-->e=y-yf

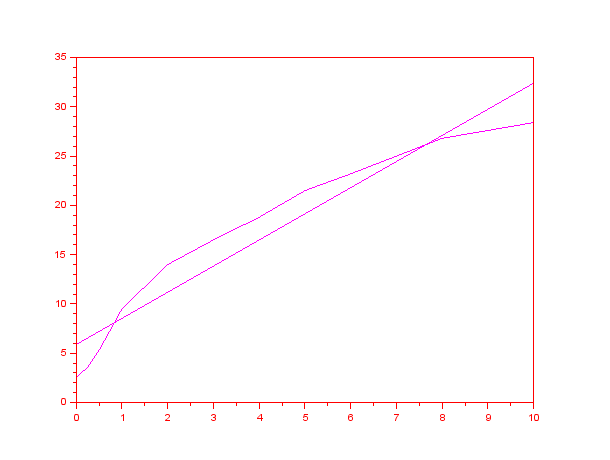
-->SSE=e'\*e

-->ybar=mean(y)

-->SST=sum((y-ybar).^2)

-->R2 = 1-SSE/SST;

R2 = 0.9196089



1. Create a quadratic model . Show the design matrix. Calculate the sum of squares and the determination coefficient (correlation coefficient R2)

**Answer:**

-->X=[ones(m,1) x x.^2]

-->a=inv(X'\*X)\*(X'\*y);

-->a

a =

3.3104404

4.9909061

- 0.2534281

-->yf=a(1)+a(2).\*x +a(3).\*x.^2;

-->e=y-yf

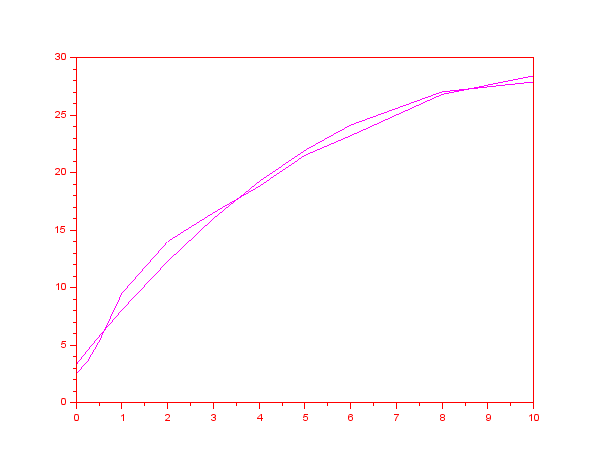
-->SSE=e'\*e

-->ybar=mean(y)

-->SST=sum((y-ybar).^2)

-->R2 = 1-SSE/SST;

R2 = 0.9899185



1. Create a cubic model . Show the design matrix. Calculate the sum of squares and the determination coefficient (correlation coefficient R2)

**Answer:**

-->X=[ones(m,1) x x.^2 x.^3]

-->a=inv(X'\*X)\*(X'\*y);

a =

2.5821752

6.520698

- 0.6864861

0.0295764

-->yf=a(1)+a(2).\*x +a(3).\*x.^2+a(4).\*x.^3;

-->e=y-yf

-->SSE=e'\*e

-->ybar=mean(y)

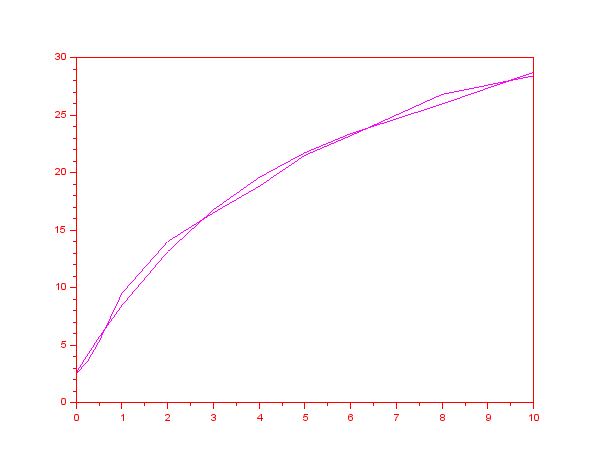
-->SST=sum((y-ybar).^2)

-->R2 = 1-SSE/SST;

-->a

-->R2

R2 = 0.9953975



1. What model would be the best fit? Why?

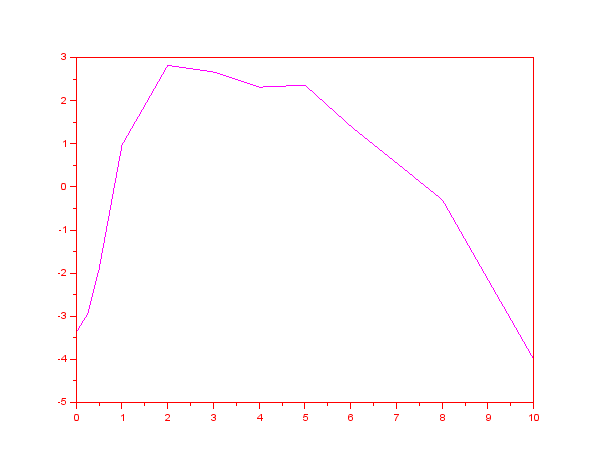
**Answer:**

Cubic fit is best because it has the highest R2

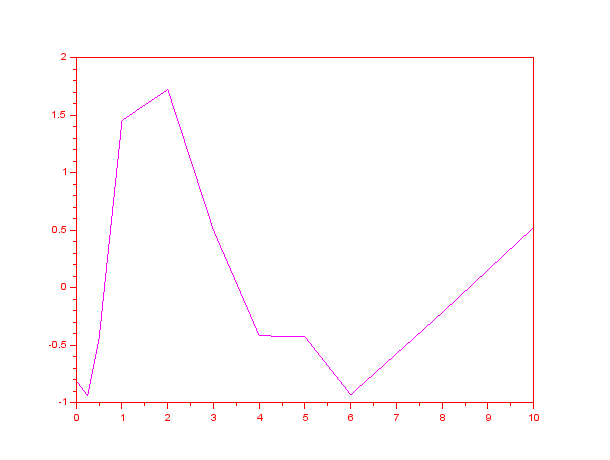
1. Definition of Residual. For each data point, its residual is the difference between the actual data point and the value of the equation used to model the data (in this case a polynomial). That is, the residual is . The residuals should be normally distributed. Make a plot of x (Day) vs e. If the residuals are “s” shaped or “u” the fit is not a good one. Are any of the models created a good fit?

**Answer:**

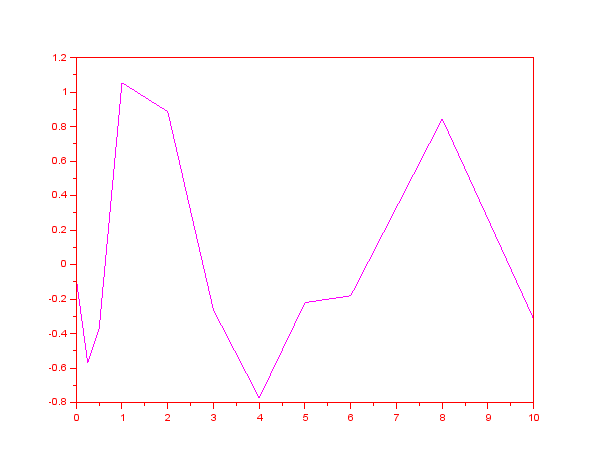
For linear model



For the quadratic



For the cubic



Cubic has the most even distribution about the mean error, which is zero.

1. The given data can also be modelled using the following exponential equation:

This equation models a situation where a system goes from a zero value to the maximum value C with a rate constant a. This time, the parameters C and a have a conceptual meaning (maximum drug released and rate constant for the release). Make a fit for the new model

1. Since this model cannot be linearized, solve using MS Solver

**Answer:**

deff('y=f(c,x,y)','y=c(1)\*(1-exp(c(2)\*x))-y');

[fopt,copt,gopt]=leastsq(list(f,x,y),[1,1])

fopt =

20.365437

copt =

28.685858 -0.3042985

gopt =

3.275D-12 -3.438D-09

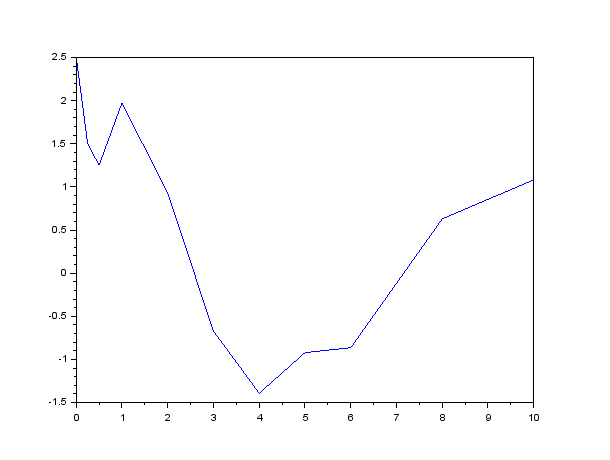
1. Show and plot the residuals. Is this model a good fit?

**Answer:**

deff('y=g(c,x)','y=c(1)\*(1-exp(-c(2)\*x))')

ypred=g(copt,x)

correl(y,ypred)^2

**

Hard to tell if it is a good fit. The residuals start very badly and then improve.

1. One of the most valuable things a model can do is to predict the behaviour of the system in conditions not originally measured. We held out on you until now, giving you only data through 10 days. However, the published data goes through 21 days:

|  |  |
| --- | --- |
| Day | Micrograms |
| 12 | 28.4 |
| 16 | 28.5 |
| 21 | 29.5 |

Your models should be good at predicting what will happen between days 10 and 21. Without refitting the models, calculate the values of your 4 models at days 12 to 21.

1. Plot the residuals only for days 12 to 21

**Answer:**

ynewpred=g(copt,xnew)

ynewpred =

27.941459

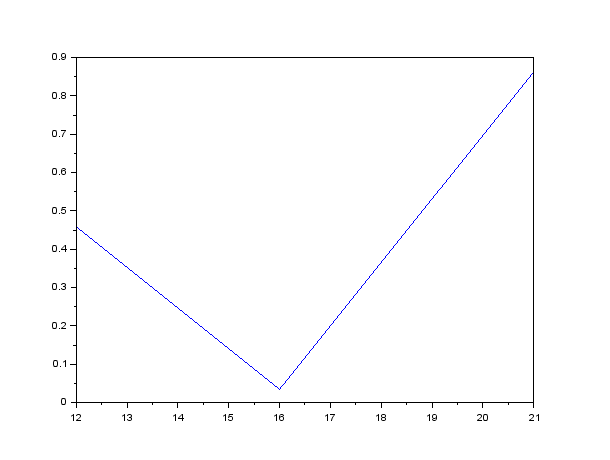
28.465472

28.637729

correl(ynew,ynewpred)^2

0.5601420

plot(xnew,ynew-ynewpred)



1. Plot the fit for all data points from 0 to 21. Calculate . Plot residuals.

**Answer:**

1. Plot micrograms vs the prediction. A good fit would be a straight line

**Answer:**

1. Answer the questions. Which of the models is the most predictive through 21 days?

**Answer:**