Tutorial on fitting with gnuplot

1. Put the data to be fit in a data file with three columns. We are going to be fitting voltage vs time data. The first column is time, the second column is the voltage (in mV) and the third column is the estimated error in the voltage. The columns are separated by blanks or any other white space, eg., several blanks or tabs.

   The error bars can be estimated by running the experiment without a signal and observing the recorded values. The standard deviation of this “null signal” gives an estimate of the uncertainties on each point. Alternately, in this case, since the signal dies out, one could analyse the standard deviation of the final 10 data points with the assumption that most of the variations are due to noise rather than signal. Once a reasonable fit is obtained this estimate can be corrected.

   Here are the first few lines of the data file “dsin.dat”:

```
#Damped oscillation
#time(s) voltage(mV) error
 0  1.31377  0.03
 0.10101  1.50314  0.03
 0.20202 -1.8479  0.03
 0.30303 -0.419889  0.03
 0.40404  1.95928  0.03
 0.505051 -0.47184  0.03
 0.606061 -1.48892  0.03
 0.707071  1.15167  0.03
 0.808081  0.855747  0.03
 0.909091 -1.43855  0.03
 1.0101 -0.122895  0.03
 1.11111  1.38982  0.03
```

2. Define the function

   In gnuplot define the function to which you wish to fit. For example, a damped oscillation follows the formula:

```
v(x) = v0*exp(-gamma*x)*sin(omega*x+phi)
```

   I’ve left the dummy variable to be x. If you wish you can change it to t.

```
set dummy t
```

   v0 is the amplitude, gamma the damping coefficient. omega is the oscillation frequency in radians per second and phi is the phase shift in radians.

3. Estimate initial values:

    The oscillation near the beginning is between ±2 mV so choose
    ```
    v0=2.
    ```

    The oscillation falls from 2 to 2/e in about 2 seconds. So
    ```
    gamma=1./2
    ```

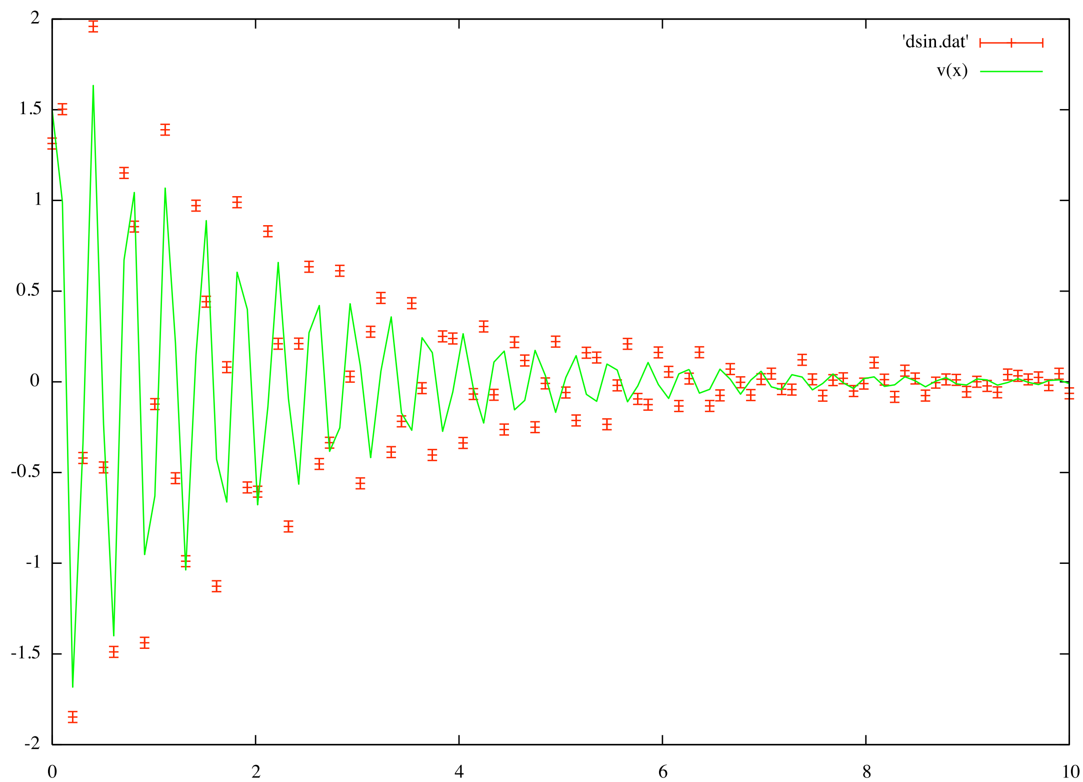
    (You have to type the decimal point or else it does an integer divide to give zero!)

    There are about 11 oscillations in the first 4 seconds to
    ```
    omega=11*2*pi/4.
    ```

    The first point, t=0, is 1.5 mV compared to the estimated amplitude of 2 mV. Therefore
    ```
    phi=sin-1(1.5/2)= 0.85.
    ```

4. Plot the theoretical curve with these starting values along with the data.

```
plot v(x), 'dsin.dat' with errorbars
```



As you can see. the first guess isn't too great. Let's see what happens when we fit with these initial parameters.

5. Do the fitting.

```
fit v(x) 'dsin.dat' using 1:2:3 via v0,gamma,omega,phi
```

(There is no comma between v(x) and the file name.)

The "using 1:2:3" tells it to use column 1 for the independent variable, column 2 for the dependent variable and column 3 for the error estimates. At the end of the fitting process it will print the following:

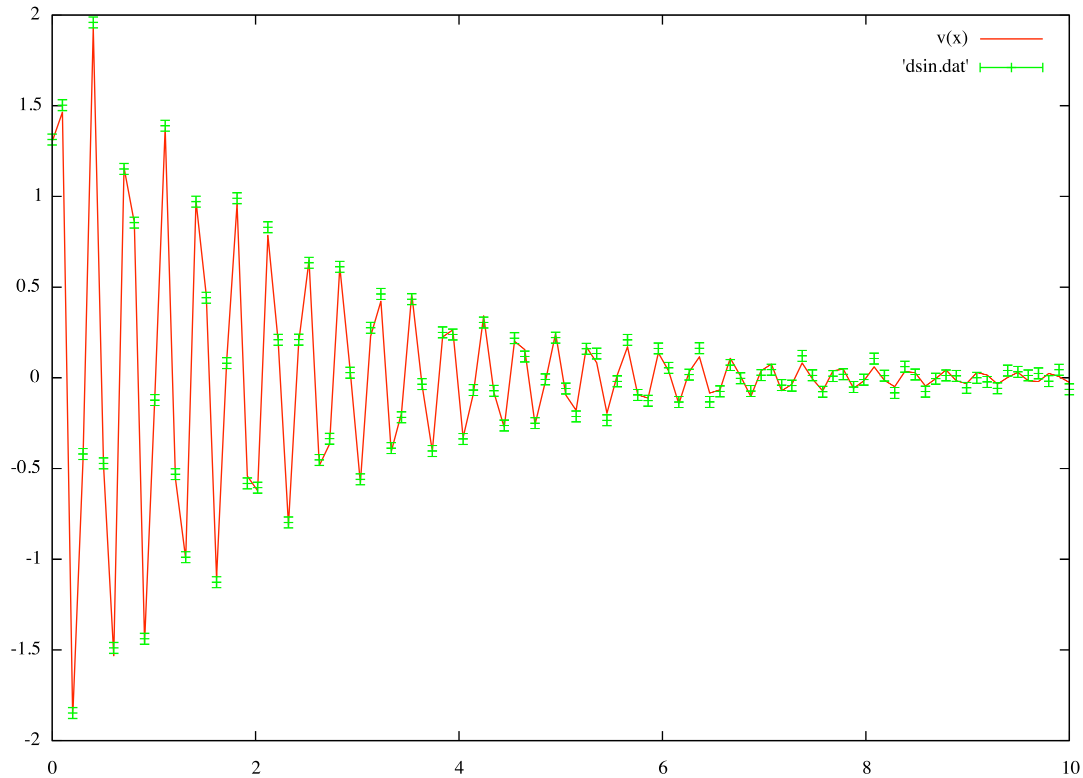
```
After 8 iterations the fit converged.
final sum of squares of residuals : 81.3434
rel. change during last iteration : -7.85339e-08

degrees of freedom (ndf) : 96
rms of residuals      (stdfit) = sqrt(WSSR/ndf)      : 0.920504
variance of residuals (reduced chisquare) = WSSR/ndf : 0.847327

Final set of parameters            Asymptotic Standard Error
=======================            ==========================

v0              = 2.31149          +/- 0.01645      (0.7116%)
gamma           = 0.450897         +/- 0.004587     (1.017%)
omega           = 18.0021          +/- 0.00447      (0.02483%)
phi             = 0.598476         +/- 0.006749     (1.128%)


correlation matrix of the fit parameters:

               v0     gamma  omega  phi
v0              1.000
gamma           0.702  1.000
omega           0.041  0.001  1.000
phi            -0.069 -0.030 -0.683  1.000
```

The statistics have some utility. The "reduced chisquare" should be near one for a reasonable fit. If it is much greater than one we might pose two questions: (1) are the estimated standard-deviations of the points unrealistically small? or (2) is the theoretical model correct? In this case a value of 0.85 assures us that the assumed standard deviations and the theoretical model may be correct.

The list of final parameter values is useful, However, the "asymptotic standard errors" listed appear to be much too small. One should estimate the standard error of a parameter by finding a value of the parameter at which the parameter can be fixed while yielding a chi-squared roughly twice the minimum value (2 x 0.85) when the data are fit with the remaining parameters allowed to vary. In this case I was able to fix omega at either 18.05 or 17.95 and a refit to the data, allowing v0, gamma and phi to vary, gave a chi-squared around 2. Therefore, I would list the value of omega as 18.00±0.05 radians/s.

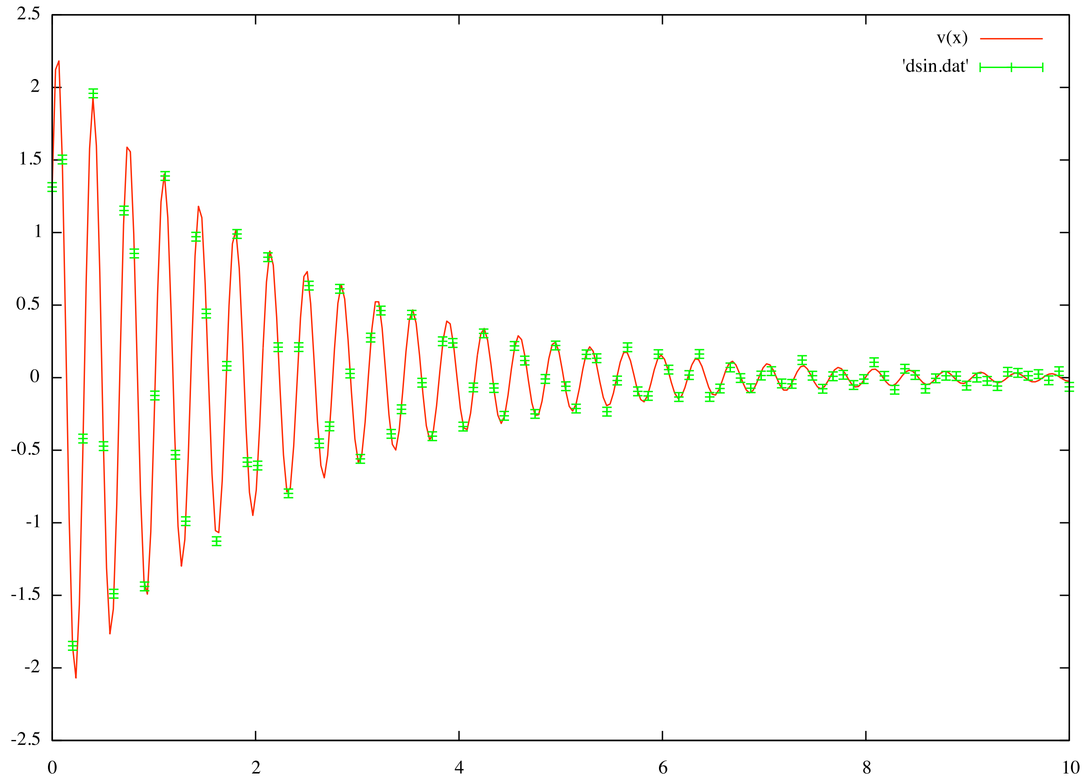
The covariance matrix shows which parameters are correlated. For example, omega and phi are correlated at –0.68. This means that if omega were fixed at a larger value, the goodness-of-fit can be partially restored by decreasing phi. Similarly v0 and gamma are positively correlated. Increasing v0 can be compensated by an increase of gamma.

Below is a plot of the data and theoretical curve at the best-fit parameter values.



To get a little smoother graph, increase the number of points for the theoretical curve:

```
set samples 300; replot
```



NA: 2007.02.26.