AMTH250 Assignment 6 Solutions

October 17, 2011

Question 1

This is an interpolation problem. It solved by constructing the basis matrix and solving a linear system.

```
octave:> x = [-2 -1 0 1 2]';
octave:> y = [125.948 7.673 -4 -14.493 -103.429]';
octave:> aa = [exp(-2*x) exp(-x) ones(5,1) exp(x) exp(2*x)];
octave:> coeffs = aa\y
coeffs =
   3.0000
  -4.9998
  -1.0004
   1.0003
  -2.0000
```

Thus the interpolating function is

$$y \approx 3e^{-2x} - 5e^{-x} - 1 + e^{x} - 2e^{2x}.$$

Question 2

First we compute the interpolating functions:

```
octave:> x = [0 0.01 0.04 0.09 0.16 0.25]';
octave:> y = [0 0.1 0.2 0.3 0.4 0.5]';
octave:> p = polyfit(x, y, 5);
octave:> cs = interp1(x, y, 'spline', 'pp');
octave:> pc = interp1(x, y, 'pchip', 'pp');
```

Plotting the errors, i.e. the difference between the interpolating function and $\sqrt{x}$ we see that the pchip cubic is the most accurate over most domain.

Restricting attention to the interval $[0, 0.01]$ shows that they are all about equally inaccurate over that range. The reason for this is that the derivative of $\sqrt{x}$ is infinite at $x = 0$ making approximation by a function with a finite derivative at $x = 0$ difficult near $x = 0$.

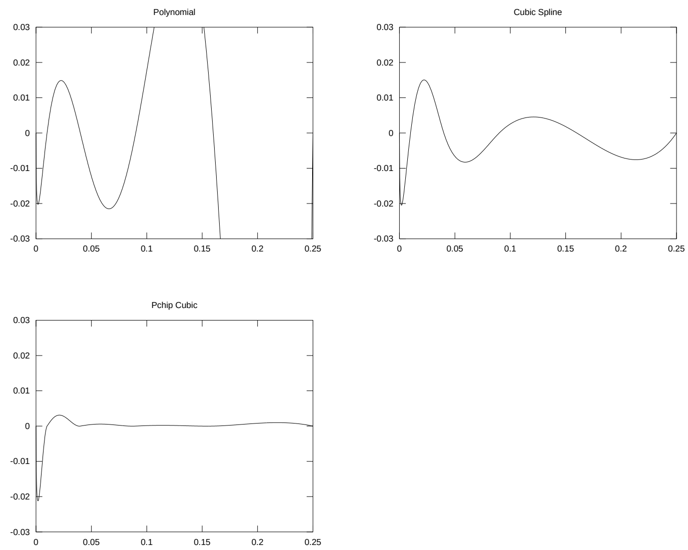


Figure 1: Errors in interpolating functions.

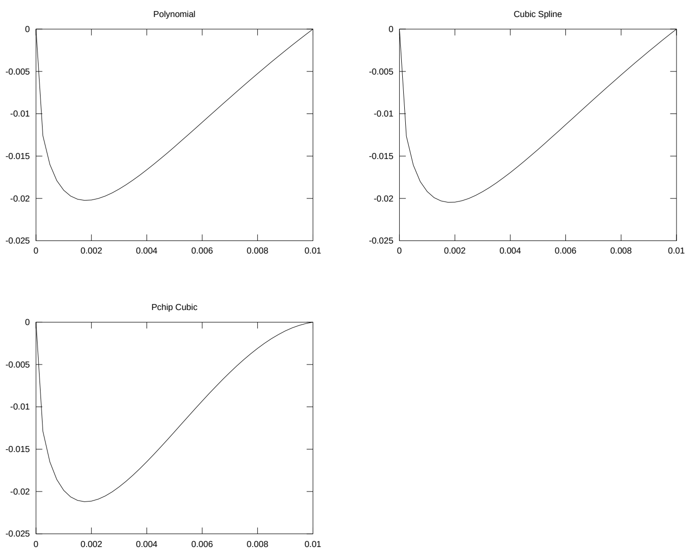


Figure 2: Errors in interpolating functions in $[0, 0.01]$.

Question 3

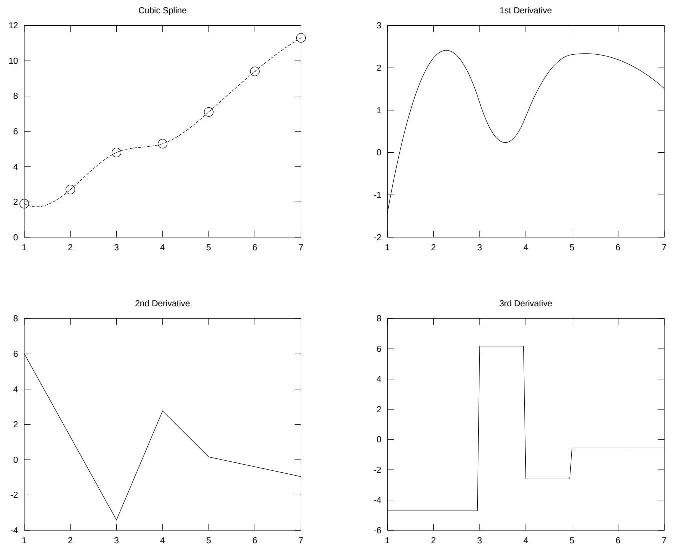


Figure 3: Cubic spline and its derivatives.

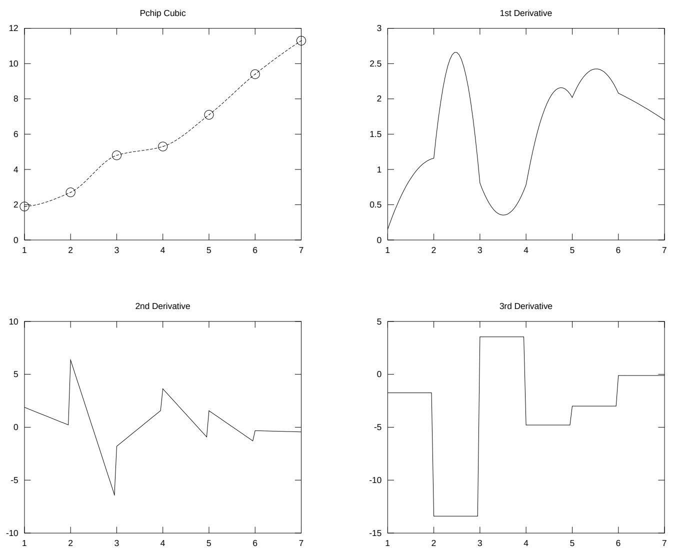


Figure 4: Pchip cubic and its derivatives.

From the graphs we see the defining conditions for a cubic spline are satisfied:

1. From the first plot the function interpolates the data.
2. From the second and third plots the function has continuous first and second derivatives.
3. From the fourth plot the function has constant third derivative on each subinterval. It follows that the function is a cubic polynomial on each subinterval.
4. The 'not-a-knot' condition shows up in the third derivative is the same on the first and second subintervals and on the second last and last subintervals.

The defining conditions for a pchip cubic are also satisfied:

1. From the first plot the function interpolates the data.
2. From the second plot the function has continuous first derivative. The third plot shows that the second derivative is discontinuous.
3. From the fourth plot the function has constant third derivative on each subinterval. It follows that the function is a cubic polynomial on each subinterval.
4. In this example the data is monotonic, a property shared by the pchip cubic. This implies that the derivative of the pchip cubic does not change sign. The second plot shows that derivative is everywhere positive.

Question 4

Write the equations in matrix form and solve the least squares problem with the backslash operator.

```
octave:> A = [1 0 0 0
>0 1 0 0
>0 0 1 0
>0 0 0 1
>1 -1 0 0
>1 0 -1 0
>1 0 0 -1
>0 1 -1 0
>0 1 0 -1
>0 0 1 -1];
octave:> y = [2.95 1.74 -1.45 1.32 1.23 4.45 1.61 3.21 0.45 -2.75]';
octave:> x = A\y
x =
   2.9600
   1.7460
  -1.4600
   1.3140
```

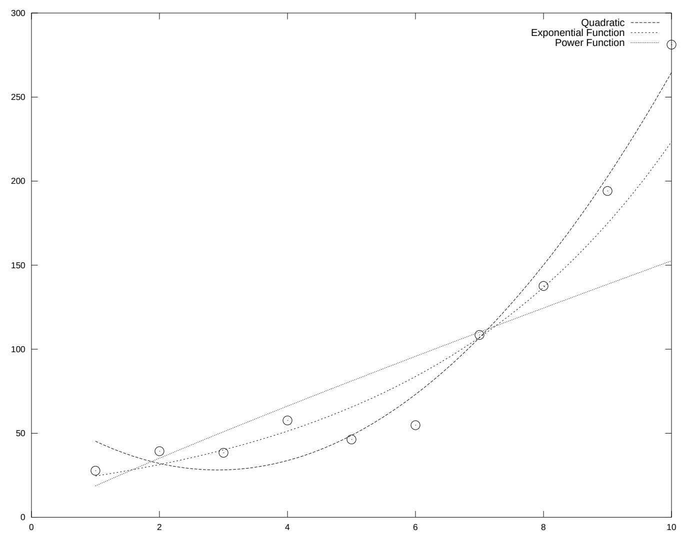
These values are within ± 0.01 of the direct measurements.

Question 5

Here is a script to compute and plot the fitted functions:

```
x = 1:10;
y = [27.7 39.3 38.4 57.6 46.3 54.8 108.5 137.6 194.1 281.2];
lx =log(x);
ly =log(y);
p = polyfit(x, y, 2);
ce = polyfit(x, ly, 1);
ae = exp(ce(2));
ke = ce(1);
fe = @(x) ae*exp(ke*x);
cp = polyfit(lx, ly, 1);
ap = exp(cp(2));
kp = cp(1);
fp = @(x) ap*x.^kp;

clf()
xx = linspace(1, 10, 201);
plot(x,y,'or')
hold on
plot(xx, polyval(p,xx))
plot(xx, fe(xx),'k')
plot(xx, fp(xx), 'g')
hold off
```



From this we see that the power function gives a poor fit, but visually there is little to chose between the quadratic polynomial and the exponential function.

Part of the reason the power function $y = ax^p$ gives such a poor fit is that when $p > 0$ the graph of the function *must* pass through the origin.

Plotting the residuals

```
res1 = y - polyval(p,x);
res2 = y - fe(x);
res3 = y - fp(x);
```

Shows that the quadratic polynomial has the smallest residuals. This is confirmed by the norms

```
octave:> norm(res1)
ans =  43.364
octave:> norm(res2)
ans =  70.894
octave:> norm(res3)
ans =  151.69
```

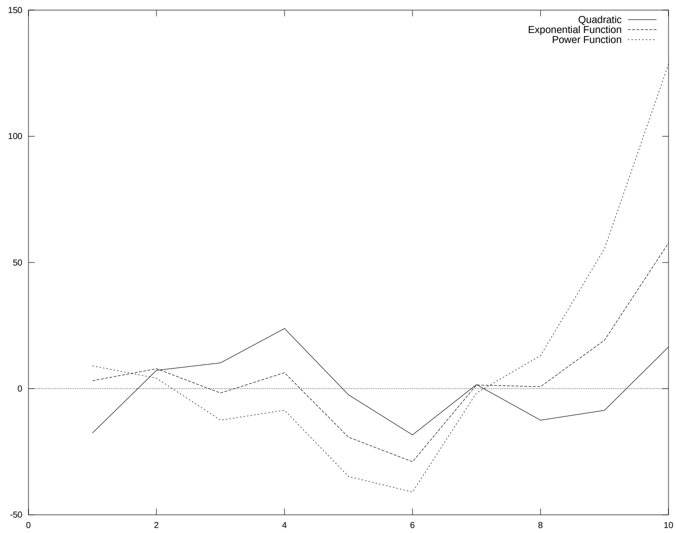


Figure 5: Residuals of fitted functions.

A semilog plot of the data indicates that an exponential function function may give a good description of the data. When working with the semilog data, the appropriate residuals are the differences $\log y_i - \log f(x_i)$

```
lres1 = ly - log(polyval(p,x));
lres2 = ly - log(fe(x));
lres3 = ly - log(fp(x));
```

Using the semilog scale that the exponential function has the smallest residuals. This is confirmed by the norms

```
octave:> norm(lres1)
ans =  0.87342
octave:> norm(lres2)
ans =  0.66707
octave:> norm(lres3)
ans =  1.1785
```

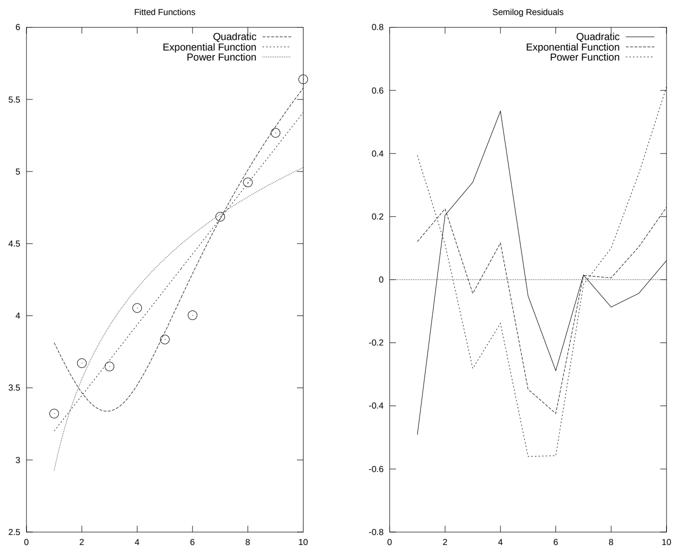


Figure 6: Semilog plot of fitted functions and semilog residuals.

A log-log plot of the data indicates that a power function would not be expected to give a good fit to the data.

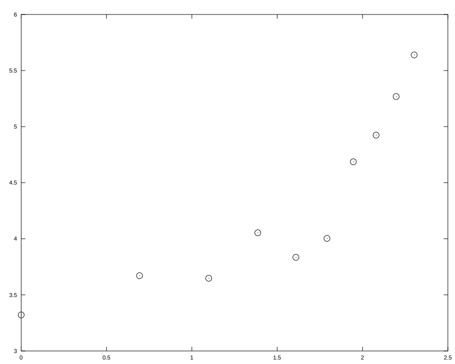


Figure 7: Log-log plot of the data

In summary

1. The power function does not give a good representation of the data.
2. The quadratic polynomial and exponential function give reasonable representations of the data.
3. Examining the residuals would indicate that the quadratic gives the best representation.
4. A semilog plot indicates that an exponential function may give a good representation of the data. Another reason for preferring the exponential is that it is monotonic and we may believe from examining the data that the underlying trend is monotonic. The exponential function is monotonic, the quadratic polynomial is not.

Question 6

(a)

(ii) Octave can compute the least squares polynomial up to about degree 36 without complaining about conditioning. The norms of the residuals for each degree are:

```
Columns 1 through 8:
  86.356   55.122   54.979   54.056   54.054   53.687   53.620   53.551
Columns 9 through 16:
  53.358   53.354   53.293   53.263   53.263   53.194   53.167   53.081
Columns 17 through 24:
  53.056   52.925   52.912   52.883   52.772   52.765   52.761   52.753
Columns 25 through 32:
  52.751   52.724   52.708   52.676   52.651   52.466   52.466   52.463
Columns 33 through 36:
  52.458   52.370   52.361   52.466
```

After the large drop from degree one to degree two – the reason for which was discussed in the notes – the decrease in norm is gradual and small. On this basis there is good reason to select the degree two polynomial.

For comparison purposes I will also look other degrees at which there is largest decrease in norm over the previous degree. These are degrees 2, 4, 6, 9 and 30.

The first thing to look is plot of the polynomials themselves. Two interesting things to note are:

1. The degree 9 polynomial is showing the rise above the general trend around 1990 apparent in the original data.
2. The degree 30 polynomial shows a small oscillation at each end. If you look at the residual plot for this polynomial you will see that it is trying to follow the annual oscillation in the first and last years.

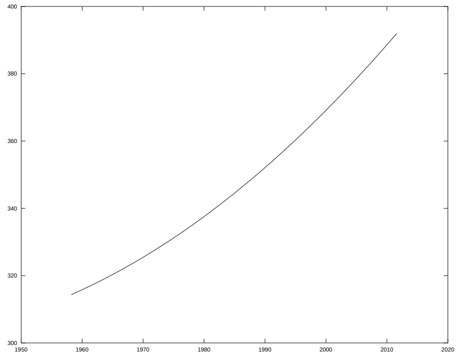


Figure 8: Degree 2 polynomial

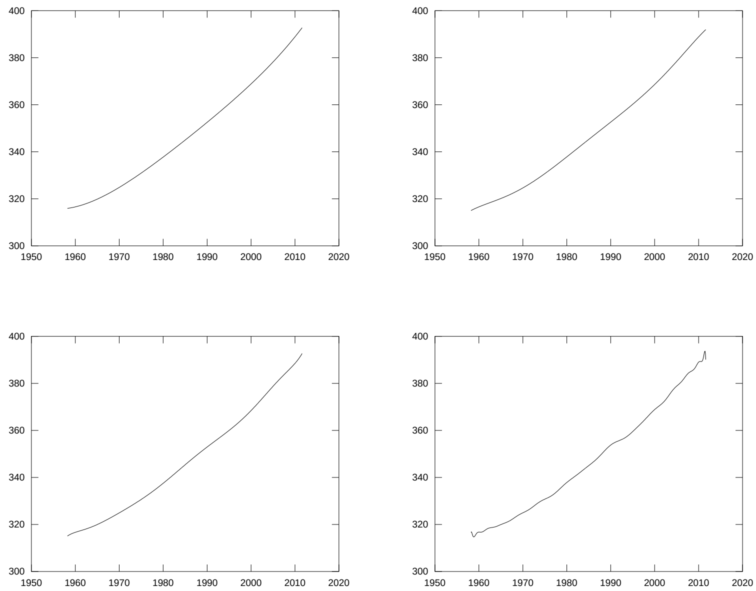


Figure 9: Polynomials of degree 4, 6, 9 and 30

The other thing to look at is residuals. Although the size of the residuals remains nearly constant in magnitude they become more regular as the polynomial degree increases.

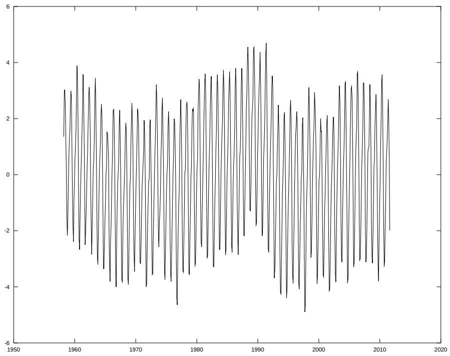


Figure 10: Residuals for the degree 2 polynomial

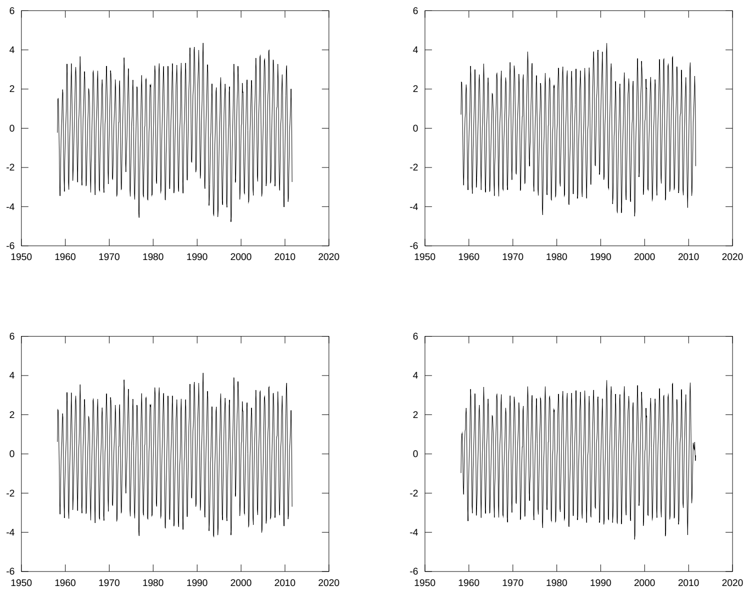


Figure 11: Residuals for polynomials of degree 4, 6, 9 and 30

The residuals for degree 30 are the most regular, but are as noted earlier there are problems with this polynomial at the ends of the time interval. The residuals for degrees 4, 6 and 9 look fairly similar and are slightly more regular than for the degree 2 polynomial.

In summary, the polynomials of degree 2 and 4 can be taken as giving the 'best' representation of the overall trend of the data.

(ii) The polynomial coefficients look quite different depending on whether or not time is rescaled. The coefficients computed with time rescaled are shown below.

	Coefficient									
Degree	0	1	2	3	4	5	6	7	8	9
1	347.34	22.47								
2	344.41	22.47	2.94							
3	344.41	22.83	2.94	-0.20						
4	344.85	22.83	1.45	-0.20	0.58					
5	344.85	22.89	1.45	-0.30	0.58	0.03				
6	345.13	22.89	-0.51	-0.30	2.54	0.03	-0.48			
7	345.13	23.41	-0.51	-1.84	2.54	1.17	-0.48	-0.23		
8	345.25	23.41	-1.97	-1.84	5.23	1.17	-2.03	-0.23	0.28	
9	345.25	24.52	-1.97	-7.29	5.23	8.25	-2.03	-3.62	0.28	0.53

Some points to note:

1. The constant and linear terms change only a little with polynomial degree. The other trems exhibit large changes, the quadratic term even changes sign.

2. For an odd degree polynomial only the odd powers change compared to the polynomial of one degree less. For an even degree polynomial only the even powers change compared to the polynomial of one degree less. This can be understood in terms of even and odd functions. (This does pattern does not occur if time is not rescaled; rescaling centres the data.)

The reason this can occurs is due to a property of polynomials — polynomials with different degree and with different coefficients can have similar graphs. Another way of saying the same thing is that low degree polynomials can give good approximations to higher degree polynomials.

(b)

(i) and **(ii)**

The following data shows the norms of the residuals for polynomial degrees 2 to 16 (rows) and trigonometric degrees 0 to 6 (columns).

```
55.122   22.826   18.266   18.205   18.167   18.163   18.162
54.979   22.482   17.818   17.756   17.715   17.710   17.708
54.056   20.589   15.356   15.291   15.244   15.237   15.236
54.054   20.582   15.350   15.285   15.238   15.231   15.230
53.687   19.171   13.406   13.324   13.269   13.263   13.263
53.620   19.008   13.150   13.066   13.009   13.002   13.002
53.551   18.958   13.075   12.992   12.934   12.927   12.927
53.358   18.328   12.193   12.104   12.047   12.039   12.039
53.354   18.311   12.169   12.079   12.021   12.014   12.014
53.293   18.193   11.968   11.876   11.815   11.808   11.807
53.263   18.191   11.963   11.872   11.810   11.803   11.803
53.263   18.180   11.952   11.861   11.800   11.793   11.793
53.194   18.137   11.879   11.789   11.728   11.721   11.721
```

53.167	18.123	11.847	11.757	11.695	11.687	11.687
53.081	17.588	11.055	10.949	10.883	10.876	10.876

The most important thing to note is that there is very little reduction in the norm of the residuals after trigonometric degree 2 for any polynomial degree. Another, less obvious point, is that with fixed trigonometric degree the residuals decreases more with polynomial degree at trigonometric degree 2 than at trigonometric degree 0 (a plain polynomial).

The conclusion from the first point is that trigonometric degree 2 gives the best model. The second point indicates that we should consider models with polynomial degree greater than 2.

The largest decreases in residual occur at degrees 2, 4, 6, 9 and 16.

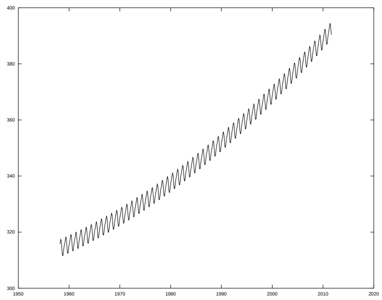


Figure 12: Polynomial degree 2, trigonometric degree 2

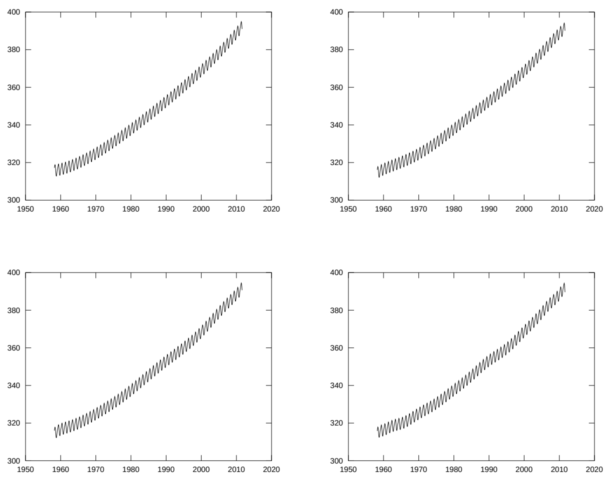


Figure 13: Polynomial degrees 4, 6, 9 and 16, trigonometric degree 2

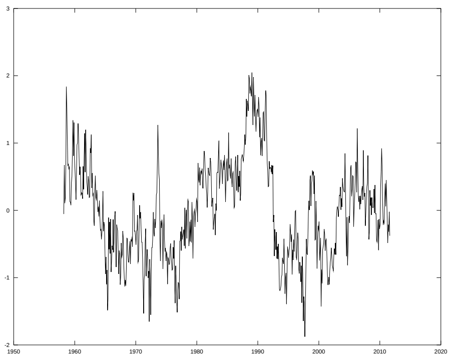


Figure 14: Residuals: polynomial degree 2, trigonometric degree 2

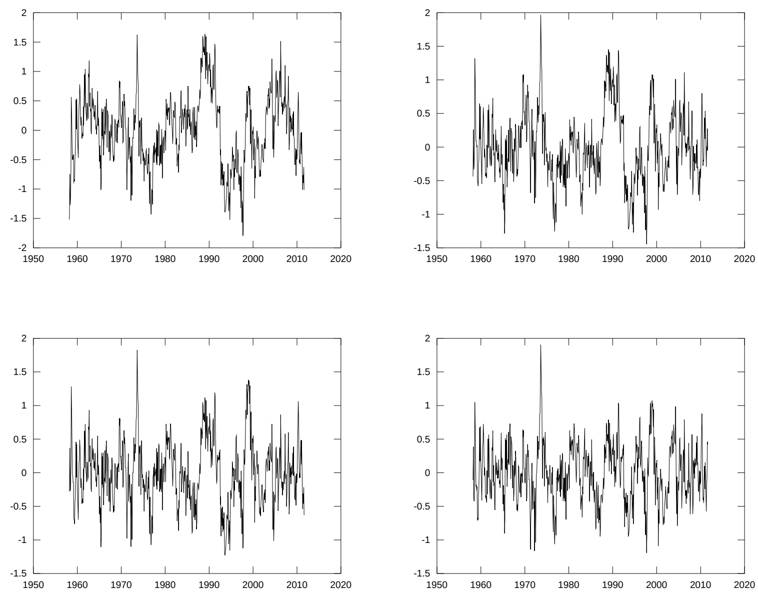


Figure 15: Residuals: polynomial degrees 4, 6, 9 and 16, trigonometric degree 2

Looking at the residual plots, the residuals for polynomial degree 16 are clearly the most regular, especially around 1990.

In summary, the model with trigonometric degree 2 and polynomial degree 16 gives the best fit to the data. But given its relative simplicity, there is a strong case for the model with polynomial degree 2 and trigonometric degree 2.

(iii) There is very little change in the coefficients when successive trigonometric terms are added. The following table shows the coefficients for polynomial degree 4 and trigonometric degrees 0 to 5. The results are similar for other degrees.

Trig.	Polynomial Coefficients					Trig. Coefficients				
Degree	0	1	2	3	4	1	2	3	4	5
0	344.85	22.83	1.45	-0.20	0.58					
1	344.83	22.82	1.57	-0.20	0.52	1.74				
						-2.18				
2	344.83	22.83	1.57	-0.20	0.52	1.74	0.36			
						-2.18	0.68			
3	344.83	22.83	1.57	-0.20	0.52	1.74	0.36	0.03		
						-2.18	0.68	0.07		
4	344.83	22.83	1.57	-0.20	0.52	1.74	0.36	0.03	-0.06	
						-2.18	0.68	0.07	-0.02	
5	344.83	22.83	1.57	-0.20	0.52	1.74	0.36	0.03	-0.06	-0.02
						-2.18	0.68	0.07	-0.02	0.02

The reason for this is that as we add higher frequency trigonometric terms these cannot be well approximated by lower frequency trigonometric terms or polynomials and so pick up features in the data which cannot be captured by those terms.

(c)

Without rescaling time we already have conditioning trouble with a polynomial of degree 4:

```
octave:> cond(vander(t,5))
ans =  6.1811e+21
```

With rescaled time there is little trouble with a polynomial of degree 16:

```
octave:329> cond(vander(tt,17))
ans =  2.7944e+06
```

When we add trigonometric terms an interesting thing happens. Here are the condition numbers for polynomial degree 16 and trigonometric degrees 0 to 10:

```
Columns 1 through 6:
 2.7944e+06   2.7952e+06   2.7952e+06   2.7954e+06   2.7954e+06   2.7956e+06
Columns 7 through 11:
 1.7538e+16   1.8468e+16   1.8476e+16   1.8664e+16   1.9040e+16
```

The conditions number remains nearly unchanged for trigonometric degrees 0 to 6 and then makes a large jump at trigonometric degree 7. The same jump in condition number at trigonometric degree 7 occurs for all polynomial degrees.

[The jump at trigonometric degree 7 can be explained: all the trigonometric terms in our model perform a whole number of cycles in one year and so take only 12 different values for the monthly time data. There are 12 trigonometric

terms with degrees 1 to 6 and these take up all the available degrees of freedom in the data — mathematically we have a 12 dimensional vector space of function values spanned by the trigonometric terms with degrees 1 to 6 — making the higher frequency terms redundant.]