Question 1: Using the data from problem set 3, estimate using your own estimator (not polyfit) the coefficients of the quadratic polynomial that best fit the data in a least squares sense.

(a) Write the equation for the polynomial. What are the observations and unknown parameters in the polynomial (5-points).

Denoting the measured 2*(Sun Elevation angle) as DE(t) (for “double elevation”) where t is UT, we can write the polynomial as

\[ DE(t) = a_0 + a_1 t + a_2 t^2 \]

The observations are the 14 measured values of DE (actual values are in degrees and minutes and we will need to convert these to decimal degrees). The unknown parameters we seek to estimate are \(a_0, a_1, a_2\).

(b) Write the above set of equations in matrix form (5-points).

In matrix form we can write these equations as:

\[
\begin{bmatrix}
DE(t_1) \\
DE(t_2) \\
\vdots \\
DE(t_{n-1}) \\
DE(t_n)
\end{bmatrix} =
\begin{bmatrix}
1 & t_1 & t_1^2 \\
1 & t_2 & t_2^2 \\
\vdots & \vdots & \vdots \\
1 & t_{n-1} & t_{n-1}^2 \\
1 & t_n & t_n^2
\end{bmatrix}
\begin{bmatrix}
a_0 \\
a_1 \\
a_2
\end{bmatrix}
\]

where there are 14 rows in the left hand side vector. In standard matrix form the above equation can be written as

\[ DE = A \mathbf{X} \]

(c) Form the least squares estimator and solve for the coefficients. You may use the Matlab matrix inversion routine inv and/or a calculator matrix inversion to solve the system of equations (10 points).

The least squares estimate of \(\mathbf{X}\) will be given by

\[ \hat{\mathbf{X}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{DE} \]

The Sextant_LSQ.m Matlab program that goes with the solution generates the following results.

Solution for polynomial coefficients
Offset \(-858.71273\)
Question 2: Estimate the standard deviation of the errors in the measurements using the differences between the observed values and the polynomial fit. (20-points)

Assuming the same noise process acts on all of the measurements, we can generate the differences between the observed values and the estimates from the model fit with

\[ \mathbf{R} = \mathbf{DE} - \mathbf{A} \hat{\mathbf{X}} \]

The estimate of the standard deviation of the noise process is then given by:

\[ \sigma^2 = \mathbf{R}^T \mathbf{R} / (n - 3) \]

where \( n \) is the number of observations and 3 is the number of parameters estimated.

From the MatLab program we obtain

Standard Deviation of Fit to polynomial
Sigma (RMS fit)  0.0616 degs

Question 3: What is the probability that the 9th measurement (time 16h 19m 35.0s Measurement 74° 4.80') differs from the polynomial fit due to random error assuming that the noise in the measurements is Gaussianly distributed and the data standard deviation computed in question 2 (20-points).

The 9th measurement has a residual of –0.103 degrees, and given the above sigma estimate of 0.0616 degrees, the residual deviates by -1.685\( \sigma \). From a table of Standard Normal Distributions we find the probability of deviating by this amount or more due to random error is 4.61%. Hence there is about 5% change we would have a residual of this size. Had we made 20 measurements, then one value of this magnitude should have been expected. (The 9th measurement turns out not to be the largest residual. The 2nd residual is –1.72s from the polynomial fit. This is not so clear in the figure shown in the solution to HW3 because of the rapid rate of change at that time.)

Question 4: Estimate the standard deviation of the peak in the polynomial (i.e., its maximum value and the time at which the maximum occurs) based on the least squares estimate in Question 1 (20-points).

To solve this problem we use propagation of variances and we need to write a linear equations between the time the maximum DE is reached, denoted by \( T_m \), and the value of DE at the time, denoted \( \text{DE}_m \). From HW3 Solution, we have an equation for \( T_m \) and
substituting this value into the equation for $DE_m$ yield non-linear equations for $T_m$ and $DE_m$:

$$T_m = -\frac{a_1}{2a_2}$$

$$DE_m = a_0 + a_1T_m + a_2T_m^2$$

$$\therefore DE_m = a_0 - \frac{a_1^2}{4a_2^2}$$

We can linearize these equations which generates:

$$\begin{bmatrix} \Delta T_m \\ \Delta DE_m \end{bmatrix} \approx \begin{bmatrix} 0 & -\frac{1}{2a_2} \\ 1 & \frac{a_1}{2a_2} \\ \frac{1}{4a_2} & \frac{a_3}{4a_2} \\ \frac{1}{4a_2} & \frac{a_3}{4a_2} \end{bmatrix} \begin{bmatrix} \Delta a_0 \\ \Delta a_1 \\ \Delta a_2 \end{bmatrix}$$

The variance covariance matrix of the parameter estimates is obtained from by scaling the inverse of $A^TA$ by the observation variance computed in Question 2. We denote this matrix by $C_{xx}$. The covariance matrix of the maximum values is then given by

$$C_m = BC_{xx}B^T$$

These equations are solved in the Matlab program and the results converted in the errors in longitude and latitude (for the latter we divide the sigma by 2 since we use divide $DE_m$ by 2 to obtain the latitude.

The results from the Matlab program are:

- Estimate of sigma of $T_m$ (converted to minutes of arc) 11.81 min
- Estimate of sigma of Latitude (minutes of arc) 0.68 min
- Correlation between values is 0.055

We can compare these standard deviations with the actual differences from HW3 of:

$\Delta \lambda = 4.6$ min., and $\Delta \phi = 1.2$ min. The observed differences are comparable to the estimated standard deviations.