SOD Exam April 2015 - Elaborations and solutions

Frank de Veld

October 25, 2019

1 Part 1

a.

i.

A probability density function can be obtained by making a fit through the residual data; this yields a function which represents the probability of event occurrence in a certain interval, i.e. that a found residual not part of this data set is in this interval. Note that in order to do this, the integral of the PDF over the real line must equal 1. Another way to find the PDF is to calculate several relevant statistical values like mean, median, variance, kurtosis and skewness, and compare these with theoretical values of known distributions.

ii.

A quick way of checking normality is to calculate the sample skewness and the sample kurtosis of the data set. For a normal distribution, the skewness is 0 and the kurtosis is 3. If the values for this data set differ significantly from these values, normality is unlikely. The tricky part is to determine what difference is 'significant'. A way to make this more concrete is to do a p-value test on the data set (beyond the scope of the course).

iii.

Outliers are data points or observations that lie in an interval where the amount of observations does not represent the probability of occurrence. This means that there are too many or too few observations in a specific interval than expected from statistics. This is difficult to determine without knowing the PDF, but if for example the mean and the median differ significantly, this is due to extremes in the data set. These extremes can then be outliers, since they are so far away from the other observations, and more observations in between would be expected.

b.

i.

A general first order polynomial is of the form:

$$y = ax + b \tag{1}$$

Here, a and b are the parameters to be estimated (a is the rate, b is the bias). However, we want to make use of two different slopes; one before and one after the discontinuity. To avoid problems with complex notation of such a polynomial (for example, with indicator variables), it is more sensible to perform two separate Least-Square Methods for the two regions. For the first interval between 0 and 2.4 years, we then get:

$$\vec{y} = H\vec{x} + \vec{\epsilon} \tag{2}$$

With $\vec{y} = [y_1, y_2, ..., y_n]^t$, n the amount of measurements in the interval t=0y to t=2.4y, $\vec{x} = [a, b]^t$, and

$$H = \begin{vmatrix} t_1 & 1 \\ t_2 & 1 \\ ... & ... & ... & ... \\ t_n & 1 \end{vmatrix}$$

Application of the LSQ algorithm then gives $\hat{x} = (H^t P_1^{-1} H)^{-1} H^t P_1^{-1} \vec{y}$, with P the covariance matrix of the measurements in the interval t=0y to t=2.4y.
Similarly, for the second interval t=2.4y to t=5y we have

$$\vec{y} = H\vec{x} + \vec{\epsilon} \tag{3}$$

With $\vec{y} = [y_1, y_2, ..., y_n]^t$, n the amount of measurements in the interval t=2.4y to t=5y, $\vec{x} = [c, d]^t$, and

$$H = \begin{vmatrix} t_1 & 1 \\ t_2 & 1 \\ ... & ... & ... & ... \\ t_n & 1 \end{vmatrix} \tag{4}$$

Application of the LSQ algorithm then gives $\hat{x} = (H^t P_2^{-1} H)^{-1} H^t P_2^{-1} \vec{y}$, with P the covariance matrix of the measurements in the interval t=2.4y to t=5y.

ii.

From the application of this method described above, a covariance matrix P_x is obtained for both parts of the graph. The entry on the first row and first column then is the standard deviation σ of the slopes in the first and the second half of the residual graph. Together with the found slopes in the two parts, the two slopes with their two standard deviations can be compared with each other to see if the rate of change differs more than one sigma before and after the discontinuity.

c.

This information is about the uncertainty of the measurements done. Since the residuals are the difference between the observations and the predictions, this noise information also has influence on the uncertainty of the residuals. This information is used in the covariance matrix P. Total errors are computed by taking the square root of the sum of the individual errors. If we assume no other errors, we can simply use diagonal entries of $1cm$ in the matrix P_1 and diagonal entries of $1.5cm$ in the matrix P_2. Since this is about observation noise, there will be no correlation between measurements from that part (noise is random), which explains the zero entries in the off-diagonal terms. If there are other errors, the values of these have to be known as well before filling in the covariance matrices.

d.

This information is trickier to implement, since we used two separate methods for both parts of the plot. Now, it is advisable to use one LSQ method after all and one large covariance matrix. Three parameters need to be estimated; two slopes and one bias. The covariance matrix is on the diagonal filled with the values $1cm$ up to t=2.4 year, afterward it is filled with values 1.5 cm up to t=5 year. However, now some off-diagonal terms have to be filled as well. In the covariance matrix of the parameters $P_x = (H^t P_y^{-1} H)^{-1}$, the off-diagonal elements on row 1, column 2 and on row 2, column 1 need to have values now (if the first two columns indicate the two slopes). The entries in the covariance matrix P_y need to be such that this is indeed the case.

e.

i.

Possible atmospheric corrections are to calculate ionospheric delays by the formulae $d_1 + r_0 + \frac{\alpha}{f_1^2}$ and $d_2 + r_0 + \frac{\alpha}{f_2^2}$, calculate refraction delays in the troposhere by $\Delta s = \int_A^B (n(s) - 1)ds$ and estimating the effects of concentration of gasses as oxygen and nitrogen.

ii.

A notable difference between laser and radar measurements is that laser light is not affected by the wet troposphere, so the humidity in the air as no effect on the measurements. This is different for radar measurements, since radio waves are refracted by water gas in the air. Additionally, radio waves can be reflected from the ionosphere, complicating measurements. These two effects also need to be taken into account when making atmospheric corrections.

2 Part 2

a.

i.

Positions and velocities of the satellite with a known initial state vector are usually obtained via numerical integration; thus propagation in time with a known force model. If some kinds of measurements are available, after this propagation the positions and velocities can be updated using this information. The force model is usually obtained from the gradient of the gravitational potential and the inclusion of other forces like drag or radiation pressure.

ii.

The satellite trackers are ground-based, and thus are rotating with Earth. We prefer to calculate in an inertial, Earth fixed frame, rather than a co-rotating frame. A Coriolis-pseudo force is thus present with an accompanying Coriolis acceleration. This is present in the dynamic force model. Additionally, the gravitational acceleration is present in the dynamic force model, as well as optional drag acceleration and radiation pressure accelerations.

b.

DORIS measurements are radio-wave based Doppler effect measurements, from which the velocity of the satellite can be obtained. The way this works in orbit determination is the following: the satellite sends out radio waves in a fixed frequency in fixed time intervals. These radio waves are then caught by several beacons on the ground. Due to the Doppler effect in the line-of-sight, frequencies will be shifted, but in a different way for each beacon. If four beacons are present (three for all velocity components, one for time errors), all information can be obtained on the velocity of the satellite. When an initial guess is taken and a force model is used, the velocity of the satellite can be computed with this model and then it can be compared to the actual observed velocity. Combining these dynamics and statistics yields an orbit determination procedure.

c.

Parameters need to be adjusted (to measurements) since errors are always present in every force model, observation or parameter. Furthermore, a crucial part of precise orbit determination is propagation where errors easily accumulate. Propagation is done numerically, not analytically, and this always introduces an error. After time, these errors become too large for practical use. Relating calculated orbits to new observations always is beneficial for POD. Additionally, the physical parameters can change as well during the process of POD. There are slight gravitational variations in the orbit, and radiation pressure is significantly lower on the dark side of the Earth, for example.

d.

i.

For many POD methods, it is the case that the found parameters and orbits are not close enough to actual values and orbit after one iteration. This is since the initial state vector has a large influence on the procedure, and if the first guess for the initial state vector is far of from the actual initial state vector, more iterations are needed to negate this influence. In other words, the POD methods need time to converge to a solution before the results are reliable.

ii.

For a batch LSQ algorithm in POD, it is the initial state vector that changes continuously in the model. The aim for such an algorithm is to find a good initial state vector and then from the observations also automatically state vectors for all epochs. After each iteration, the initial state vector is calculated again, and then also the observation data vector. The model itself is not changed. For sequential LSQ algorithms as the Kalman filter, the variables are updated on the fly, which means that the state vector itself is updated each time step. However, no iterations are used here.

e.

Batch parameter estimation during POD separates arc parameters and common parameters and uses optimal weight factors for combining different observation equations, but most importantly propagates an initial state vector fully to the end epoch before comparing the orbit to actual observations and updating the initial state vector to these observations. In sequential POD methods, the adjustment procedure and the state vector transition mechanism are two separate steps, and during transition the state vector and the covariance matrix are advanced to the next time step, not after the orbit has been calculated. Thus, observations are processed in blocks, and the calculated state vector is updated to the observations each time step, or every other time step.

3 Part 3

a.

The power of GPS signals is usually extremely low and the random noise present in measurements usually has a comparable magnitude as the GPS signal. However, the GPS signal can still be retrieved. First of all, the GPS signal is modulated on top of a 1.6 GHz signal. In this GPS signal, a pseudo-random noise code is present (PRN), which serves as identification. Every GPS satellite has its own unique PRN code. The receiver of the GPS signal then compares the signal with modulated PRN information with a series of known PRN codes by means of replication in the receiver station. The correlation is calculated using shifting of signals, multiplication of signals and then integration of signals. If the GPS signal matches one of the known codes, we know which GPS satellite sent this data and more importantly, what the data actually is, as this can be read of the same way. In this way, the power consumption of GPS satellites can be very low without problems.

b.

The reason why in the polar regions the GPS determined position is less accurate (in general, not only for airplanes) is that there is less GPS coverage, as GPS satellites prefer an orbit closer to the equator. Very large inclinations of the orbits are needed to fly over the poles regularly, and there are simply less users near the Arctic then in southern regions. There are still GPS satellites in view on the poles, but they are relatively low in the horizon. This leads to more influence of the troposphere and ionosphere, making the position determination less accurate.

c.

This is a matter of geometry; in the horizontal plane, there is a view of 360 degrees. For orbit determination, it is good if the satellites are far away from each other, since this leads to a low position dilution of precision. If GPS satellite are close to each other, the error overlap is larger, which is a poor satellite geometry. In a view of 360 degrees, there is more space for separate satellites to transmit information. For the vertical plane, there is only a view of 180 degrees, since GPS signals are not accurately transmitted through the Earth itself. Thus, there is less space for spacing out, which usually leads to a lower accuracy in position determination in height than in the horizontal plane.

d.

A double difference error mitigation method is based on observations from two satellites and two receivers. In this case, the receivers are the airplane and the landing strip. The main idea behind this method is that errors can be removed if they are correlated in time and space. If the two receivers are close together, the errors in their received signals are more correlated. Atmospheric errors are the main errors that can be removed better if the two receivers are closer together. On the other end, there are no significant sources that produce errors if the two receivers are closer together. Thus, the position determination becomes better as the airplane approaches the airport's landing strip.

e.

Multipath errors are caused by the fact that there are multiple pathways possible for GPS signals if they are reflected from other surfaces. It becomes then more difficult to distinguish the true signal from the reflected signals. Thus, the multipath error becomes larger if there are more reflective surfaces. In this case, position B is preferable above position A, as a GPS antenna in position A would also receive GPS signals reflected from the aircraft's fuselage. An antenna in position B would be integrated in the aircraft's fuselage, which makes the amount of reflections from the fuselage to the antenna lower.

f.

The formula for single difference measurements is:

$$SD(t_P, t^A, t^B) = \rho(t_P, t^A) - \rho(t_P, t^B) \tag{5}$$

Here, P is the receiver and A, B are two GPS satellites. $\rho(t_P, t^A)$ and $\rho(t_P, t^B)$ are uncorrelated. A covariance matrix P would in this case look like:

$$P_l = \begin{vmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{vmatrix}$$

The matrix A to get the single difference from the pseudo-range measurements is simply:

$$A = \begin{vmatrix} 1 \\ -1 \end{vmatrix}$$

This means that we have:

$$P_{SD} = A P_l A^T = [1 - 1] \cdot \begin{vmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{vmatrix} \cdot \begin{vmatrix} 1 \\ -1 \end{vmatrix} = [1 - 1] \cdot \begin{vmatrix} \sigma_1^2 \\ -\sigma_2^2 \end{vmatrix} = \sigma_1^2 + \sigma_2^2 \tag{6}$$

As expected, the noise error increases when single differencing is done.

4 Part 4

a.

$\bar{P}_i$ is the covariance matrix of the initial state vector. In other words, it consists of the expected standard deviation or error in the elements of the calculated state vector on the diagonal, and the correlation or mutual influence between the elements of the calculated state vector on the off-diagonal spots in the matrix.
$\bar{R}_i$ the observation data covariance matrix. In other words, it consists of the expected standard deviation or error in the individual (GPS) observations on the diagonal of the matrix and the correlation or mutual influence between the individual (GPS) observations on the off-diagonal spots in the matrix.

b.

The $\tilde{H}$-matrix is by definition equal to $\frac{\partial \bar{y}}{\partial \bar{X}}$. There is only one observation, namely ρ, thus $\tilde{H}$ has one row. However, $\bar{X}$ has six elements, thus $\tilde{H}$ has six columns. We can write $\tilde{H}$ as:

$$\tilde{H} = \left| \begin{matrix} \frac{\partial \rho}{\partial X} & \frac{\partial \rho}{\partial Y} & \frac{\partial \rho}{\partial U} & \frac{\partial \rho}{\partial V} & \frac{\partial \rho}{\partial C_D} & \frac{\partial \rho}{\partial C_L} \end{matrix} \right|$$

This leads to:

$$\tilde{H} = \left| \begin{matrix} \frac{X-X_S}{\rho} & \frac{Y-Y_S}{\rho} & 0 & 0 & 0 & 0 \end{matrix} \right|$$

c.

The equations of motion can be written in first order form, in the form

$$\dot{\bar{X}} = F(\bar{X}, t) \tag{7}$$

This leads to the following:

$$\left| \begin{matrix} \dot{X} & \dot{Y} & \dot{U} & \dot{V} & \dot{C}_D & \dot{C}_L \end{matrix} \right| = \left| \begin{matrix} U & V & T - \frac{1}{2}C_D\rho U^2 S & \frac{1}{2}C_L\rho U^2 S - W & 0 & 0 \end{matrix} \right|$$

Thrust (T) and weight (W) are taken as general variables, as no expressions can be made for them with the given variables. As can be seen, $\bar{F}$ indeed only depends on $\bar{X}$, t and constants (like S and ρ).

d.

To find the derivative of Φ, the matrix $A(t)$ must be known. For the state transition matrix as we know it, this matrix $A(t)$ is equal to $\frac{\partial F(\bar{X}, t)}{\partial \bar{X}(t)}$. Since both F and $\bar{X}$ have six elements, this is a six-by-six matrix, consisting of:

$$A(t) = \left| \begin{matrix} \frac{\partial U}{\partial X} & \frac{\partial U}{\partial Y} & \frac{\partial U}{\partial U} & \frac{\partial U}{\partial V} & \frac{\partial U}{\partial C_D} & \frac{\partial U}{\partial C_L} \\ \frac{\partial V}{\partial X} & \frac{\partial V}{\partial Y} & \frac{\partial V}{\partial U} & \frac{\partial V}{\partial V} & \frac{\partial V}{\partial C_D} & \frac{\partial V}{\partial C_L} \\ \frac{\partial (T-\frac{1}{2}C_D\rho U^2 S)}{\partial X} & \frac{\partial (T-\frac{1}{2}C_D\rho U^2 S)}{\partial Y} & \frac{\partial (T-\frac{1}{2}C_D\rho U^2 S)}{\partial U} & \frac{\partial (T-\frac{1}{2}C_D\rho U^2 S)}{\partial V} & \frac{\partial (T-\frac{1}{2}C_D\rho U^2 S)}{\partial C_D} & \frac{\partial (T-\frac{1}{2}C_D\rho U^2 S)}{\partial C_L} \\ \frac{\partial (\frac{1}{2}C_L\rho U^2 S-W)}{\partial X} & \frac{\partial (\frac{1}{2}C_L\rho U^2 S-W)}{\partial Y} & \frac{\partial (\frac{1}{2}C_L\rho U^2 S-W)}{\partial U} & \frac{\partial (\frac{1}{2}C_L\rho U^2 S-W)}{\partial V} & \frac{\partial (\frac{1}{2}C_L\rho U^2 S-W)}{\partial C_D} & \frac{\partial (\frac{1}{2}C_L\rho U^2 S-W)}{\partial C_L} \\ \frac{\partial 0}{\partial X} & \frac{\partial 0}{\partial Y} & \frac{\partial 0}{\partial U} & \frac{\partial 0}{\partial V} & \frac{\partial 0}{\partial C_D} & \frac{\partial 0}{\partial C_L} \\ \frac{\partial 0}{\partial X} & \frac{\partial 0}{\partial Y} & \frac{\partial 0}{\partial U} & \frac{\partial 0}{\partial V} & \frac{\partial 0}{\partial C_D} & \frac{\partial 0}{\partial C_L} \end{matrix} \right|$$

Calculating these gives:

$$A(t) = \left| \begin{matrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -C_D\rho U S & 0 & -\frac{1}{2}\rho U^2 S & 0 \\ 0 & 0 & C_L\rho U S & 0 & 0 & \frac{1}{2}\rho U^2 S \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \right|$$

e.

The traces of $\bar{P}_i$ and K_i will change over time. After propagation, errors increase. Thus, the trace of $\bar{P}_i$ increases, as this indicates the standard deviation of the state vector. From the given equation, the trace of K_i will then also increase. However, when GPS observations are available again after being blocked or jammed, the trace of $\bar{P}_i$ decreases with a sudden jump, and the same happens for the trace of K_i. If the observations become regular again, over time the trace of both matrices decrease, but in the same zig-zag fashion as described above.

f.

The advantage of the Kalman filter is that it does the POD 'on-the-fly', so without iterations. After a certain number of epochs, the Kalman filter then converges to an accurate solution. The problem is that in the first few epochs, the Kalman filter is still far off from this accurate solution, and the filter does not solve this problem in itself. An intuitive and useful fix for this issue is to also apply a backward Kalman filter after convergence, and then combine the found state vectors and their covariances. This reverses the front and the back of the orbit, and makes the found orbit much smoother and more accurate.

5 Comments and remarks

These solutions have been made by me and are in no way the 'official' solutions. Personally, I am not sure of my answers for questions 1bii and 3f. Please post a comment somewhere if you have remarks, questions or corrections (I am not sure how that works in GitHub).