Optimal Measurement Filtering and Motion Prediction Taking Into Account the Atmospheric Perturbations

Andrey I. NAZARENKO
Space Observation Center, Moscow, Russia, e-mail: anazarenko@iki.rssi.ru

Vasiliy S. YURASOV
Institute for Precision Instrument Engineering, Russia
e-mail: vyurasov@mail.ru

Kyle T. ALFRIEND
Texas A&M University, e-mail: alfriend@aero.tamu.edu

Paul J. CEFOLAN
MIT/Consultant, e-mail: cefola@mit.edu
Outline

1. Statement of the Problem
3. Estimation and Prediction of Orbits Allowing for Atmospheric Disturbances at Joint Processing of Measurements
4. Forecasting of the Gaussian Random Process
5. Modeling the Algorithm and Software for Optimal Measurement Filtering
6. Conclusions
1. Statement of the Problem
Correlation function for normalized ballistic factor variations for SO 6350

- Current indices for solar flux and geomagnetic activity
Correlation function for normalized ballistic factor variations for SO 7852

SO number 7852

SD(dKb/Kb) = 0.133

E(Kb) = 0.0484 + 0.0000897*(h-200), sq. m/kg

Current indices for solar flux and geomagnetic activity
Individual Character of SO Auto Correlation Functions

- Long period changes in the envelop of the ballistic factor variations
- Period of the variations
- Magnitude of the variations

How can we improve the SO motion prediction techniques on the basis of a more complete accounting for the statistical characteristics of the atmosphere drag?
Linearized Differential Equation for Prediction Errors

\[ \frac{dE}{dt} = B(E,t) \cdot [q_0(E,t) + q(E,t)] \]

\[ \frac{d\tilde{E}}{dt} = B(\tilde{E},t) \cdot q_0(\tilde{E},t) \]

\[ x(t) = E(t) - \tilde{E}(t) \]

\[ A(t) = \frac{\partial [B(\tilde{E},t) \cdot q_0(\tilde{E},t)]}{\partial \tilde{E}} \]

\[ B(t) = B(\tilde{E},t) \]

\[ \frac{dx}{dt} = A(t) \cdot x + B(t) \cdot q(t) \]

- True SO motion
- SO motion with known disturbing forces
- Prediction errors
- Definitions A & B
- Assumption:
  \[ |q_0(t)| \gg |q(t)| \]
- Linearized ODE
Colored Noise Estimation Problem

- Assume the system given by the linearized ODE
  \[
  \frac{dx}{dt} = A(t) \cdot x + B(t) \cdot q(t)
  \]
- Assume measurements given by
  \[
  z_i = h_i \cdot x(t_i) + v_i
  \]
- Assume a priori known statistical characteristics:
  \[
  M[q(t)]_0 = 0, \\
  M[q(t) \cdot q^T(\tau)]_0 = K_q(t, \tau)_0, \\
  M[v_i \cdot q(\tau)]_0 = 0, \\
  M[v_i \cdot v_j^T]_0 = R_i \cdot \delta_{ij}
  \]

**Initial Data**

Suppose, that the following characteristics are obtained after sequential processing of measurements at the \((k-1)\)-th step:

\[
M\left[x(t_{k-1})\right] = M[x_{k-1}] = \hat{x}_{k-1|k-1},
\]

\[
M\left[\left(x_{k-1} - \hat{x}_{k-1|k-1}\right) \cdot \left(x_{k-1} - \hat{x}_{k-1|k-1}\right)^T\right] = P_{k-1|k-1},
\]

\[
M[q(t)] = \hat{q}(t)_{k-1},
\]

\[
M\left\{q(t) - \hat{q}(t)_{k-1}\right\} \cdot \left[q(\tau) - \hat{q}(\tau)_{k-1}\right]^T = K_q(t, \tau)_{k-1},
\]

\[
M\left\{\left(x_{k-1} - \hat{x}_{k-1|k-1}\right) \cdot \left[q(t) - \hat{q}(t)_{k-1}\right]\right\} = K_{xq}(t_{k-1}, t)_{k-1}.
\]
The solution of considered task is as follows:

\[
\hat{x}_k|_k = \hat{x}_k|_{k-1} + P_k|_k \cdot h_k^T \cdot R_k^{-1} \cdot (z_k - \hat{z}_k|_{k-1}),
\]

\[
P_k|_k = \left( P_k|_{k-1} + h_k^T \cdot R_k^{-1} \cdot h_k \right)^{-1},
\]

\[
\hat{q}(t)_k = \hat{q}(t)|_{k-1} + K_{xq}(t_k,t)|_{k-1} \cdot P_k|_{k-1} \cdot P_k|_k \cdot h_k^T \cdot R_k^{-1} \cdot (z_k - \hat{z}_k|_{k-1}),
\]

\[
K_{xq}(t_k,t)|_k = P_k|_k \cdot P_k|_{k-1} \cdot K_{xq}(t_k,t)|_{k-1},
\]

\[
K_q(t,\tau)|_k = K_q(t,\tau)|_{k-1} - K_{xq}(t_k,t)|_{k-1} \cdot h_k^T \cdot \left[ R_k^{-1} - h_k \cdot P_k|_k \cdot h_k^T \cdot R_k^{-1} \right] \cdot h_k \cdot K_{xq}(t_k,\tau)|_{k-1}.
\]

Comments: The basic distinction of the algorithm for estimation and prediction of a state vector from the Kalman filter consists in including three new functional relations: a) for estimating the noise, b) for the function of mutual correlation, and c) for the a posteriori auto correlation function. These functional relations take into account the effect of the system’s noise \( q(t) \) and, on this basis, allow us to determine the optimum (in accuracy) estimates of the state vector at future time instants.
3. Estimation and Prediction of Orbits
Allowing for Atmospheric Disturbances at
Joint Processing of Measurements

\[
\hat{x}_j = \left( X_j^T \cdot P_j \cdot X_j \right)^{-1} \cdot X_j^T \cdot P_j \cdot Z_k, \quad \text{or} \quad \hat{x}_j = U(t_j, t_k) \cdot \hat{x}_k + \sum_{i=1}^{k} Q^{(j)}_{ij} \cdot H^T \cdot P_k \cdot (Z_k - X_k \cdot \hat{x}_k)
\]

\[
\text{where} \quad P_j = \left( H \cdot K_{wj} \cdot H^T + R_\Sigma \right)^{-1}, \quad K_{wj} = \begin{bmatrix} Q_{11}^{(j)} & Q_{12}^{(j)} & \cdots & Q_{1k}^{(j)} \\ Q_{21}^{(j)} & Q_{22}^{(j)} & \cdots & Q_{2k}^{(j)} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{k1}^{(j)} & Q_{k2}^{(j)} & \cdots & Q_{kk}^{(j)} \end{bmatrix} = \{ Q_{il}^{(j)} \} \quad i, l = 1, \ldots, k,
\]

\[
|Q_{il}^{(j)}| = \int_{t_j}^{t_j} \int_{t_j}^{t_j} U(t_i, \xi) \cdot B(\xi) \cdot K_q(\xi, \eta) \cdot B^T(\eta) \cdot U^T(t_i, \eta) \cdot d\eta \cdot d\xi,
\]

\[
Z_k = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix}, \quad H = \begin{bmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_k \end{bmatrix}, \quad R_\Sigma = \begin{bmatrix} R_1 & 0 & \cdots & 0 \\ 0 & R_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_k \end{bmatrix}, \quad U_j = \begin{bmatrix} U(t_1, t_j) \\ \vdots \\ U(t_k, t_j) \end{bmatrix}, \quad H \cdot U_j = X_j.
\]
This solution has a traditional form. The distinction takes place in the formation of the weight matrix \( P_j \), which takes into account the effect of not only the measurement errors (matrix \( R \Sigma \)), but the system’s noise (power density matrix \( K_{wij} \)) as well.

The following expression is deduced for the required derivative of the state vector:

\[
\frac{dx_j}{dt_j} = A(t_j) \cdot \hat{x}_j + B(t_j) \cdot W_k^T(t_j) \cdot H^T \cdot P_k \cdot (Z_k - X_k \cdot \hat{x}_k), \quad (t_j \geq t_k),
\]

where the following designations are applied:

\[
W_k(t_j) = \begin{bmatrix}
K_{wq}(t_1, t_j) \\
K_{wq}(t_2, t_j) \\
\vdots \\
K_{wq}(t_k, t_j)
\end{bmatrix}, \quad K_{wq}(t_i, t_j) = \int_{t_k}^{t_i} U(t_i, \xi) \cdot B(\xi) \cdot K_q(\xi, t_j) \cdot d\xi.
\]

So, the differential equation for estimating the maximum likelihood value \( \hat{x}_j \) at an arbitrary time that takes into account the effect of the system’s noise is constructed. From the derivative of the state vector it is obvious that

\[
\hat{q}(t_j) = W_k^T(t_j) \cdot H^T \cdot P_j \cdot (Z_k - X_j \cdot \hat{x}_j)
\]

is the estimate of the system’s noise at the arbitrary time instant \( t_j \geq t_k \).
4. Forecasting of the Gaussian Random Process

Consider the particular case of the problem. Assume that there are direct measurements of the Gaussian random process $q(t)$:

$$z_i = q(t_i) + v_i, \ i=1,2,...,k.$$  

The random values $v_i$, have the same characteristics, as those mentioned above.

As a result, the problem solution under the successive measurement processing algorithm is reduced to two functional recurrence relations:

$$\hat{q}(t)_k = \hat{q}(t)_{k-1} + K_q(t,t_k)_{k-1} \cdot [K_q(t_k,t_k)_{k-1} + R]^{-1} \cdot [z_k - \hat{q}(t_k)_{k-1}], \ t \geq t_k,$$

$$K_q(t,\tau)_k = K_q(t,\tau)_{k-1} - K_q(t,t_k)_{k-1} \cdot [K_q(t_k,t_k)_{k-1} + R]^{-1} \cdot K_q^T(\tau,t_k)_{k-1}, \ t,\tau \geq t_k.$$

As a result, the problem solution under joint processing of measurement algorithm is reduced to:

$$\hat{q}_j = K_{qZ}(t_j) \cdot K_{ZZ}^{-1} \cdot Z_k,$$

$$K_q(t,\tau)_k = M \begin{bmatrix} q(t) - \hat{q}(t) \cdot [q(\tau) - \hat{q}(\tau)]^T \cdot Z_k \end{bmatrix} = K_q(t,\tau)_0 - K_{qZ}(t) \cdot K_{ZZ}^{-1} \cdot K_{qZ}^T(\tau), \ t,\tau > t_k,$$

where

$$K_{qZ}(t_j) = M \begin{bmatrix} q_j \cdot Z_k^T \end{bmatrix} = M \begin{bmatrix} q_j \cdot [z_{1}^T, z_{2}^T, ..., z_{k}^T] \end{bmatrix} =$$

$$= \begin{bmatrix} K_q(t_j,t_1)_0 & K_q(t_j,t_2)_0 & ... & K_q(t_j,t_k)_0 \end{bmatrix}$$

$$K_{ZZ}(t_j) = M \begin{bmatrix} Z_k \cdot Z_k^T \end{bmatrix} =$$

$$\begin{bmatrix} K_q(t_1,t_1)_0 & K_q(t_1,t_2)_0 & ... & K_q(t_1,t_k)_0 \ K_q(t_2,t_1)_0 & K_q(t_2,t_2)_0 & ... & K_q(t_2,t_k)_0 \ ... & ... & ... & ... \ K_q(t_k,t_1)_0 & K_q(t_k,t_2)_0 & ... & K_q(t_k,t_k)_0 \end{bmatrix} +$$

$$\begin{bmatrix} R_1 & 0 & ... & 0 \ 0 & R_2 & ... & 0 \ ... & ... & ... & ... \ 0 & 0 & ... & R_k \end{bmatrix}$$

Monterey 2007 7th US/Russian Space Surveillance Workshop 12
As random processes $q(t)$, consider the deviations of current values $F10.7$ indices from averaged ones

$$t_{\text{prediction}1} = t_k + t,$$
$$t_{\text{prediction}2} = t_k + \tau$$

Normalized auto correlation functions (1 step=1 day). Data for years 2000-2003

Values of components of matrix $Kq(t,\tau)$ according to the results of joint processing of measurements
The a posteriori SD estimates of the normalized errors versus the forecasting interval were calculated (designated as “real errors of least-squares technique”). Below, figure presents these estimates and their comparison with the results of calculations by Kq (t,τ) matrixes. In total, 1488 determinations of the random process parameters and the corresponding number of forecasts were performed.

Comments:
1. The SD of the real forecasting errors and their calculated estimates agree rather well.
2. From the presented results the conclusion can be drawn on the identity of techniques considered in Sections 2 and 3. Their basic distinctions relate to calculation aspects.
3. Calculations, using the recurrence functional equation, show that two types of results of the iterative process take place: either the process converges to the steady value of an a posteriori auto correlation function, or the pseudo steady mode is only achieved, in which the solution oscillates in the vicinity of some average value. The possible cause of the aforementioned principal distinction in the process of convergence of the iterative process consists in the properties of the initial auto correlation functions.
4. The formulation of necessary and sufficient conditions of correctness of the initial auto correlation function is topical problem.
5. Modeling the Algorithm and Software for Optimal Measurement Filtering (OMF)

We take the angular motion of a satellite in revolutions $N$ as an argument, and as the state vector components we use the following three elements: the time instant $t_i$ corresponding to the trajectory point with the given latitude argument ($u=0$), the period of one revolution $T_i$, and the variation of period during one revolution $\Delta T_i$. The following equations are valid for these variables:

$$t_{i+1} = t_i + T_i,$$
$$T_{i+1} = T_i + \Delta T_i,$$
$$\Delta T_i = \Delta T_m \cdot (1 + q_i).$$

Quantity $\Delta T_i$ is supposed to be associated with the atmospheric drag only. The data sources are supposed to be the direct measurements of time of the trajectory points $N_i$:

$$z_i = t_i + v_i \cdot (i=1,2,...),$$

Example of linear correlation function $K_q(\ldots)$

Parameters: $\sigma_q$, $\Delta$
Basic parameters:

\( \sigma_q \) – RMS of color noise,

\( \Delta \) – interval of correlation (revolutions),

\( \Delta N \) – interval between measurements (revolutions),

\( n_z \) – number of measurements for fit span,

\( \sigma_z \) – RMS of measurement error (time),

\( \Delta T_m \) – mean value of \( \Delta T_i \) (time),

\( S_n = \sigma_q \cdot \Delta T_m / \sigma_z \) – signal to noise merit.
Modeling by parameters: $Sn=1/5$, $\Delta N= 2$, $\Delta = 50$, $nz=\text{var}$

Techniques: OMF and LST. In all cases the forecasting was carried out for 60 revolutions with a step of 5 revolutions. The modeling results indicate, that:

• In the calculations for cases “LST, nz=8” and “OMF, nz=15” the forecasting errors virtually coincide with the optimum RMS deviations calculated for case “Optimal RMS”. This gives rise to the important conclusion, that under the considered conditions the improvement of the forecasting accuracy with respect to achieved one is impossible.

• In using the classical least squares technique the increase of a fit span and application of mean values of parameter at forecasting intervals up to 30 revolutions results in an increase in the errors by 30-40 % or more. With a further increase in the forecasting interval the relative distinction between the RMS deviations of the errors for various calculation cases decreases.

• In comparison with case “OMF, nz=15”, prediction with q estimates”, disregarding the system’s noise (OMF, nz=15”, prediction without q estimates) results in about a 10 % increase in the forecasting errors.
Modeling by parameters: $\text{Sn}=1/5, \Delta N=2, \Delta=50, nz=\text{var}$

RMS deviations of residual discrepancies (sec) on the processing interval

RMS at the instant of updating (the last measurement) are lowest for calculation cases “LST, nz=8” and “OMF, nz=15”. The application of the OMF algorithm is characterized by an increase in the residual discrepancies at the fit span beginning.

Example of weighting matrix components, nz=30

This Figure shows that for OMF the weighting matrix essentially differs from the identity matrix. The shape of the weighting matrix is similar to a fine-pointed roof with variable height.
Modeling by parameters: $S_n=1$, $\Delta N=2$, $\Delta=50$, $nz=\text{var}$

RMS estimates of the normalized temporary errors were calculated:

$$RMS_{\Delta t}(N)_{\text{norm}} = RMS_{\Delta t}(N) \sqrt{\frac{\sigma_z \cdot \sqrt{0.5 \cdot S_n \cdot N^2}}{1}}$$

Here $N$ is the forecasting interval. Corresponding results are presented in Figure below.
Modeling by parameters: $S_n=1$, $\Delta N=2$, $\Delta=50$, $n_z=\text{var}$

Comments

• The forecasting errors for case “OMF $n_z=30$” coincide with the expected optimum RMS deviations calculated for case “A priori $n_z=30$”. This gives rise to the important conclusion, that the application of the optimum algorithm for estimation and prediction of orbits taking into account atmospheric disturbances at joint processing of measurements provides achievement of maximally possible forecasting accuracy under considered conditions.

• Using the classical least squares technique for the optimum processing interval (case “LST $n_z=8$) results, at forecasting intervals of about 1 day, in $\approx 30$-$35\%$ increase of errors as compared to the expected optimum accuracy. Subsequently this difference decreases down to $15\%$.

• Using the classical least squares technique for a fit span increasing up to $\approx 2$ days (case “LST $n_z=15$”) results (at forecasting intervals up to 1 day) in about $(2$-$4)$-fold increase in errors as compared to the maximally achievable accuracy.

• With a further increase in the forecasting intervals the relative distinction between the RMS deviations of errors for the various cases decreases. In all cases the application of the optimum measurement processing algorithm with using the system’s noise estimates at extrapolation (case “OMF $n_z=30$”) provides achievement of the best accuracy.
Modeling the OMF technique

Thus, with increasing the “Signal-to-noise merit” up to unity the application of the optimum algorithm for estimation and prediction of orbits taking into account atmospheric disturbances at joint processing of measurements provides essentially better accuracy of the short-term forecast as compared to the application of the classical least squares technique. In this case the best possible forecasting accuracy is achieved. The mentioned effect is provided as a result of the correct "weighing" of the measurements over the processing interval and accounting for the system’s noise estimates over the forecasting interval.

With using OMF the necessity in optimizing the processing interval disappears, i.e. in this case the increase of a processing interval does not cause worsening the forecasting accuracy.
6. Conclusions

Modeling the algorithm and software for the optimal measurement filtering has demonstrated the efficiency of the method.

Though the OMF basis was developed rather long ago, its wide practical application was restrained by a series of circumstances. One of the circumstances was insufficient knowledge of the atmospheric density variations and ballistic characteristics of the SOs. Another circumstance was associated with insufficient characteristics of computer technology (memory, speed, word length). The latter circumstance is no longer an insufficiency.

There is a reason to believe, that the additional constraining circumstance is the certain "inertness" of specialists, who are used to the least squares method and believe, that after the appearance of the Kalman filter all methodological problems in the considered area are already solved.