State Transition Matrix

Gene Brownd

Summary

Three methods for computing state transition matrices for use with the ground based GPS Kalman filter are compared. Of the three methods the most accurate method of integrating the state transition matrix was chosen. The reason for the choice is that it is accurate, sufficiently fast, simple to design, and lends itself for switching to high speed propagation during periods of no data (i.e. it lends itself to using large step sizes).

The three methods compared are:

- First order (trivial integration) of the dynamics coefficient matrix. This method is fastest and least accurate.
- Integration of the dynamics coefficient matrix and low order Taylor's expansion. This method has improved accuracy over the first method and requires more processing time.
- Integration of the state transition matrix. This method is the slowest of the three and the most accurate.

Chain Rule

The state transition matrix, $\mathbf{F}$, is a Jacobean (i.e. a set of partial derivatives) that relate changes in a state vector from one point in time to another. A change in the state at time t_2 is related to the change in the state at time t_1 through the equation:

$$\mathbf{DX}(t_2) = \mathbf{F}(t_2, t_1)\mathbf{DX}(t_1) \qquad 1.$$

Since Equation 1 applies for any pair of times, the relationship over the next interval is

$$\mathbf{DX}(t_3) = \mathbf{F}(t_3, t_2)\mathbf{DX}(t_2) \qquad 2.$$

By substituting equation 1 into equation 2 the state transition matrix is chained to go from time, t_1, to time, t_3, such that

$$\mathbf{DX}(t_3) = \mathbf{F}(t_3, t_2)\mathbf{F}(t_2, t_1)\mathbf{DX}(t_1) \qquad 3.$$

And it follows that

$$\mathbf{F}(t_3, t_1) = \mathbf{F}(t_3, t_2)\mathbf{F}(t_2, t_1) \qquad 4.$$

By taking the derivative of equation 4 at the point where $t_3 = t_2$, we get

$$\dot{\mathbf{F}}(t_2, t_1) = \dot{\mathbf{F}}(t_2) \times \mathbf{F}(t_2, t_1) \qquad 5.$$

Equation 5 is in a form that can be directly integrated. Method 3 in the summary section above solves this equation. Any integrator that can solve first order differential equations can be used. This method for computing the state transition matrix is generally what is used in least-squares orbit determination applications. For Kalman filters, faster methods are generally used.

A general solution for equation 5 is

$$\mathbf{F}(t_2, t_1) = e^{\int_{t_1}^{t_2} \dot{\mathbf{F}}(t)\,dt} \qquad 6.$$

In order to transform from time t_1 to time t_3, Equation 6 can be expanded such that

$$\mathbf{F}(t_3, t_1) = e^{\int_{t_2}^{t_3} \dot{\mathbf{F}}(t)\,dt + \int_{t_1}^{t_2} \dot{\mathbf{F}}(t)\,dt} \qquad 7.$$

It is a trivial problem to transform Equation 7 into Equation 4. Also, by substituting t for t_3 and taking the derivative of Equation 7 with respect to t, we get Equation 5.

The matrix, F, is sometimes called the dynamics coefficient matrix. It is the result of the integration

$$\mathbf{F}(t_2, t_1) = \int_{t_1}^{t_2} \dot{\mathbf{F}}(t)\,dt \qquad 8.$$

Gene Brownd
United Space Alliance
(281) 282-3915

The state transition matrix, $\mathbf{F}$, can be estimated by using a Taylor's expansion of the exponential function, e^F.

$$\mathbf{F}(t_2, t_1) = \mathbf{I} + \mathbf{F}(t_2, t_1) + \frac{1}{2}\mathbf{F}^2(t_2, t_1) + \frac{1}{6}\mathbf{F}^3(t_2, t_1) + \frac{1}{24}\mathbf{F}^4(t_2, t_1) + \ldots \qquad 9.$$

For very short intervals the series can be truncated at the first order. However, it may be advantageous to use the third order expansion for intervals as short as 1 second in time.

Central force orbit dynamics example

For this example the state vector consists of simple Cartesian coordinates for position and velocity. The first order expansion of the equations is simply

$$\begin{aligned} \mathbf{x}(t_2) &= \mathbf{x}(t_1) + \mathbf{v}(t_1)\Delta t \\ \mathbf{v}(t_2) &= \mathbf{v}(t_1) + \mathbf{a}(t_1)\Delta t \end{aligned} \qquad 10.$$

The derivative of the dynamics coefficient matrix is then

$$\dot{\mathbf{F}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \dfrac{\partial \mathbf{a}}{\partial \mathbf{x}} & \dfrac{\partial \mathbf{a}}{\partial \mathbf{v}} \end{bmatrix} \qquad 11.$$

Since the acceleration, a, is a function of position, i.e.

$$\mathbf{a} = -\frac{\mu}{r^3}\mathbf{x} \qquad 12.$$

Then

$$\frac{\partial \mathbf{a}}{\partial \mathbf{v}} = \mathbf{0}, \quad \frac{\partial \mathbf{a}}{\partial \mathbf{x}} = -\frac{\mu}{r^3}\mathbf{I} + \frac{3\mu}{r^5}\mathbf{x}\mathbf{x}^T \qquad 13.$$

The partial derivative of acceleration with respect to position is referred to as the acceleration gradient. The 3x3 matrices, $\mathbf{G}_1$ and $\mathbf{G}_2$ are then

$$\mathbf{G}_1 = \frac{\partial \mathbf{a}}{\partial \mathbf{x}}, \quad \mathbf{G}_2 = \frac{\partial \mathbf{a}}{\partial \mathbf{v}} = \mathbf{0} \qquad 14.$$

And Equation 11 becomes

$$\dot{\mathbf{F}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{G}_1 & \mathbf{G}_2 \end{bmatrix} \qquad 15.$$

It follows then that

$$\mathbf{F} = \begin{bmatrix} \mathbf{0} & \mathbf{I}\Delta t \\ \int_{t_1}^{t_2} \mathbf{G}_1 dt & \int_{t_1}^{t_2} \mathbf{G}_2 dt \end{bmatrix} \qquad 16.$$

And the intermediate partitions of **F** can be defined, where

$$\mathbf{F}_{2,1} = \int_{t_1}^{t_2} \mathbf{G}_1 dt, \quad \mathbf{F}_{2,2} = \int_{t_1}^{t_2} \mathbf{G}_2 dt \qquad 16.$$

For short intervals, e.g. 1 second, the approximation that **G** is constant is made, so that

$$\mathbf{F}_{2,1} = \mathbf{G}_1 \Delta t, \quad \mathbf{F}_{2,2} = \mathbf{G}_2 \Delta t \qquad 17.$$

Assuming that $(\mathbf{G}\,\Delta t)^2 << \mathbf{G}\,\Delta t$, then the state transition matrix can be computed using a third order expansion of the Taylor's series, dropping all terms involving second order elements of **G**. I.e.

$$\mathbf{F} \approx \mathbf{I} + \mathbf{F} + \frac{1}{2}\mathbf{F}^2 + \frac{1}{6}\mathbf{F}^3 \qquad 18.$$

Where

$$\frac{1}{2}\mathbf{F}^2 \cong \begin{bmatrix} \mathbf{F}_{2,1}\dfrac{\Delta t}{2} & \mathbf{F}_{2,2}\dfrac{\Delta t}{2} \\ \mathbf{0} & \mathbf{F}_{2,1}\dfrac{\Delta t}{2} \end{bmatrix} \qquad 19.$$

And

$$\frac{1}{6}\mathbf{F}^3 \cong \begin{bmatrix} \mathbf{0} & \mathbf{F}_{2,1}\dfrac{\Delta t^2}{6} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad 20.$$

The expansion of Equation 18 then becomes (after dropping the $\mathbf{G}^2$ terms)

$$\mathbf{F} = \begin{bmatrix} \mathbf{I} + \mathbf{F}_{2,1}\dfrac{\Delta t}{2} & \mathbf{I}\Delta t + \mathbf{F}_{2,2}\dfrac{\Delta t}{2} + \mathbf{F}_{2,1}\dfrac{\Delta t^2}{6} \\ \mathbf{F}_{2,1} & \mathbf{I} + \mathbf{F}_{2,2} + \mathbf{F}_{2,1}\dfrac{\Delta t}{2} \end{bmatrix} \qquad 21.$$

The combination of using equation 17 and Equation 21 is the method 1 described above in the summary section. This method uses a sometimes overly simplified method of integration. It is also dependent on keeping the time interval, Δt, short. One property of Equation 21 is that it is independent of how **F** is computed. Normally, acceleration is a function of position and not velocity, therefore $\mathbf{F}_{2,2} = \mathbf{G}_2 = \mathbf{0}$.

Integrating The Dynamics Coefficient Matrix

The numerical integration of Equations 16 combined with Equation 21 is the method 2 described in the summary section. This method should be more accurate than method 1, but it has the same inherent properties of being limited to small time intervals. Again, $\mathbf{F}_{2,2}$ is normally **0**.

Adding Dynamic Parameters

Dynamic parameters, such as drag, venting, solar radiation pressure, attitude misalignment, accelerometer bias, and etc. are parameters that do not necessarily change with time, but they cause the position and velocity parameter error to change in time. When solving for the dynamic parameters, the state can be partitioned into two sets of components, the position and velocity set and the dynamic parameter set. The dynamics coefficient matrix, F, is then partitioned into 9 components, where

$$\mathbf{F} = \begin{bmatrix} \mathbf{0} & \mathrm{I}\Delta t & \mathbf{0} \\ \int_{t_1}^{t_2} \mathbf{G}_1 dt & \int_{t_1}^{t_2} \mathbf{G}_2 dt & \int_{t_1}^{t_2} \mathbf{G}_3 dt \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad 22.$$

Therefore,

$$\mathbf{F} = \begin{bmatrix} \mathbf{0} & \mathrm{I}\Delta t & \mathbf{0} \\ \mathbf{F}_{2,1} & \mathbf{F}_{2,2} & \mathbf{F}_{2,3} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad 23.$$

In Equation 22, $\mathbf{G}_3$ is the set of partials that relate changes in dynamic parameters to changes in the position and velocity (through time). In order to derive $\mathbf{G}_3$, the component of acceleration due to the dynamic parameter (or set of parameters) p_3 is modeled.

$$\mathbf{a}_3 = \mathbf{f}(\mathbf{p}_3) \qquad 24.$$

The derivative of velocity is then computed.

$$\mathbf{dv} = \mathbf{f}(\mathbf{p}_2)dt \qquad 25.$$

Then take the derivative with respect to p_3.

$$d\frac{\partial \mathbf{v}}{\partial p_3} = \frac{\partial \mathbf{f}(p_3)}{\partial p_3} dt \qquad 26.$$

The derivative of **x** with respect to p_3 is assumed to be zero.

$$\frac{d\mathbf{x}}{dp_3} = \mathbf{0} \qquad 27.$$

Then

$$\mathbf{G}_3 = \frac{\partial \mathbf{f}(p_3)}{\partial p_3} \qquad 28.$$

Direct Integration of the State Transition Matrix

The derivative of the state transition matrix is given in Equation 5. This equation can be expanded into partitions in order to take advantage of the sparse nature of the elements. Equation 5 is then expanded to

$$\dot{\mathbf{F}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{G}_1 & \mathbf{G}_2 \end{bmatrix} \begin{bmatrix} \mathbf{F}_{1,1} & \mathbf{F}_{1,2} \\ \mathbf{F}_{2,1} & \mathbf{F}_{2,2} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{2,1} & \mathbf{F}_{2,2} \\ \mathbf{G}_1 \mathbf{F}_{1,1} + \mathbf{G}_2 \mathbf{F}_{2,1} & \mathbf{G}_1 \mathbf{F}_{1,2} + \mathbf{G}_2 \mathbf{F}_{2,2} \end{bmatrix} \quad 29.$$

This method (method 3) requires special integrators that can solve first or second order differential equations.

Higher-order gravity potential

The gravity potential function is sometimes represented by a spherical harmonic series of coefficients. For the Earth, the coefficient, J_2, is dominant, second only to the central force term itself. It is more than 400 times larger than the next larger coefficient, J_3. Since J_2 is approximately 1000^{th} the magnitude of the central force, no other terms are significant to the computation of G.

Using J_2, the Equation for acceleration is

$$\vec{\mathbf{a}} = -\frac{\mu}{r^3}\vec{\mathbf{x}} - \frac{\mu J_2 a_e^2}{r^5}\left(\frac{3}{2} - \frac{15z^2}{2r^2}\right)\vec{\mathbf{x}} - \frac{3\mu J_2 a_e^2 z}{r^5}\hat{\mathbf{z}} \quad 30.$$

The gradient, **G**, (**G**$_1$) is then

$$\begin{aligned} \mathbf{G} = \mathbf{I}&\left[-\frac{\mu}{r^3} - \frac{3\mu J_2 a_e^2}{2r^5} + \frac{15\mu J_2 a_e^2 z^2}{2r^7}\right] \\ &+ \mathbf{x}\mathbf{x}^T\left[\frac{3\mu}{r^5} + \frac{15\mu J_2 a_e^2}{2r^7} - \frac{105\mu J_2 a_e^2 z^2}{2r^9}\right] \\ &+ (\mathbf{x}\hat{\mathbf{z}}^T + \hat{\mathbf{z}}\mathbf{x}^T)\left[\frac{15\mu J_2 a_e^2 z}{r^7}\right] \\ &- \hat{\mathbf{z}}\hat{\mathbf{z}}^{\mathbf{T}}\left(\frac{3\mu J_2 a_e^2}{r^5}\right) \end{aligned} \quad 31.$$

For efficient computation, the following intermediate variables are defined.

$$b_0 = -\frac{\mu}{r^3} \quad 32.$$

$$q = \frac{J_2 a_e^2}{2r^2} \quad 33.$$

$$u = \frac{z}{r} \quad 34.$$

$$b_1 = b_0[1 + 3q - 15q \cdot u^2] \quad 35.$$

$$b_2 = -b_0[3 + 15q - 105q \cdot u^2] \quad 36.$$

$$b_4 = 6b_0 q \quad 37.$$

$$b_3 = -5b_4 u \quad 38.$$

Then

$$\mathbf{G} = \mathbf{I}b_1 + \hat{\mathbf{x}}\hat{\mathbf{x}}^T b_2 + (\hat{\mathbf{x}}\hat{\mathbf{z}}^T + \hat{\mathbf{z}}\hat{\mathbf{x}}^T) b_3 + \hat{\mathbf{z}}\hat{\mathbf{z}}^T b_4 \quad 39.$$

Simulation Results

A simulation was setup to measure the relative accuracy's of the three methods and compare their processing times. The simulation was written in Visual Basic ™ in an Excel ™ spreadsheet. In order to make the comparisons, a fourth order Runge-Kutta-Nystrom integrator was selected. This integrator is very accurate for step sizes of 10 seconds and less, in that after 15 hours of propagation, less than 1 meter of error in the spacecraft's position vector was due to integration error.

The timing test was straight forward, since logic is inserted just prior to and after the loop that integrated the satellite's state (position and velocity) and then computed the state transition matrix. Since the state vector has to be integrated along with the state transition matrix, its contribution to the overall time is subtracted by creating a run that did not compute a state transition matrix.

Results of the timing test are in Table 1. These results are from integrating the three methods for 15 hours. In order to properly compare the methods for use in a Kalman filter, delta state transition matrices were computed. (I.e. the state transition over the 10 second time interval used as the step size.) For methods 1 and 2 this reduced the computation time considerably.

	Full STM Integration	Dynamics Integration	Dynamics = G Dt
8x8 State	53 sec	53 sec	53 sec
J2 State	9 sec	9 sec	9 sec
Delta STM	16 sec	10 sec	3 sec
Full STM	16 sec	14 sec	6 sec

Table 1. Time for 15.0 hour integration using a 10-sec step Runge Kutta Nystrom with J_2 Gravity Model

The interesting result of the timing test is that the state transition matrix computation is very fast compared to time required to integrate the state itself. It is especially fast compared to using a spherical harmonic gravity model (in this case of degree and order 8). Even though the method 1 is over 5 times faster than method 3, speed is not a major issue because all of the methods are extremely fast.

The accuracy test was run over 90 minutes of propagation time. The test was to predict state errors after that amount of time using the state transition matrix and the initial error at the beginning time. In order to fairly assess that the error is due to the error in the state transition matrix small perturbations were made in the state vectors. The problem with using large perturbations is that the problem is non-linear and a perfect state transition matrix would not predict the resultant error.

The procedure used to test the accuracy was to run a truth vector in order to get results at the end of the propagation period. Then an error in one component of the state vector was added one at a time to get six propagated state vectors resulting from six different component errors.

Table 2 provides results of the accuracy tests based on using a central force model. Figure 1 is a plot of these results. The results are shown into four different bins for each method. The first bin, labeled “Vel to pos” is the error in position due to an initial error in velocity. Likewise, “Vel to vel” is the error in velocity due to initial error in velocity, “Pos to vel” is the error in velocity due to an initial error in position, and “Pos to pos” is the error in position due to an initial error in position.

	Method 1	Method 2	Method 3
Vel to pos	4.20E+01	4.62E+00	7.25E-02
Vel to vel	5.53E-02	5.59E-03	8.05E-05
Pos to pos	5.61E-02	4.90E-03	8.60E-06
Pos to vel	6.07E-05	5.41E-06	1.04E-08

Table 2. Comparisons of standard deviations computed from 1.5-hour arc lengths using a central force model. The errors are differences between state transition matrix projections of state vector changes minus integrated state vector results with the same changes.

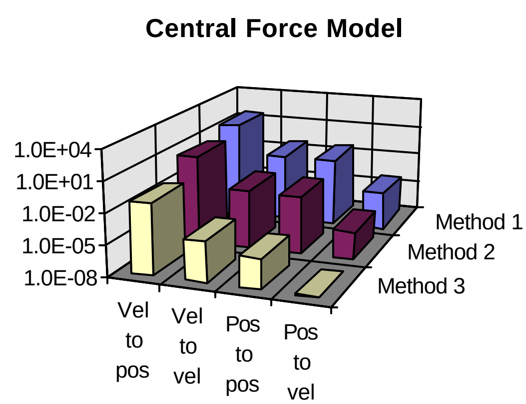

Figure 1. Standard deviation of state transition matrix projection error after 1.5 hours of propagation using a central force model. (See Table 2.)

As can be seen in Table 2, the method 3 is much more accurate than the first two methods, by two orders of magnitude. This is a significant trait, especially in terms of long term prediction of the state error (covariance).

Table 3 is only slightly different from Table 2 in terms of results. It is however the results of extending the computation to the use of the J2 (oblatness) gravity term. Figure 2 is a plot of the same results.

	Method 1	Method 2	Method 3
Vel to pos	4.19E+01	4.71E+00	7.28E-02
Vel to vel	5.55E-02	5.70E-03	8.01E-05
Pos to pos	5.64E-02	4.97E-03	8.63E-06

Pos to vel	6.06E-05	5.54E-06	1.05E-08

Table 3. Comparisons of standard deviations computed from 1.5-hour arc lengths using a central force model plus J_2 gravity term.

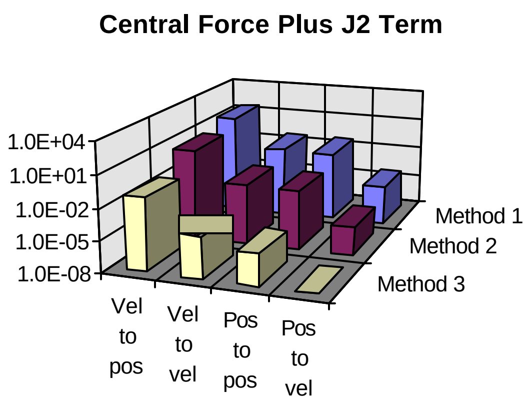

Figure 2. Standard deviation of state transition matrix projection error after 1.5 hours of propagation using a central force plus J2 gravity model. (See Table 3.)

Table 4 (and Figure 3) are the results of using an 8x8 spherical harmonic gravity model for integrating the state vector and using the J_2 model for computing the gradient for the state transition matrix. There is a loss of accuracy by two orders of magnitude due to truncating the gradient to using J_2 only.

	Method 1	Method 2	Method 3
Vel to pos	4.20E+01	6.11E+00	1.42E+00
Vel to vel	5.48E-02	7.37E-03	1.66E-03
Pos to pos	5.66E-02	6.35E-03	1.40E-03
Pos to vel	6.11E-05	6.96E-06	1.46E-06

Table 4. Comparisons of standard deviations computed from 1.5-hour arc lengths using a central force model plus 8x8 gravity terms.

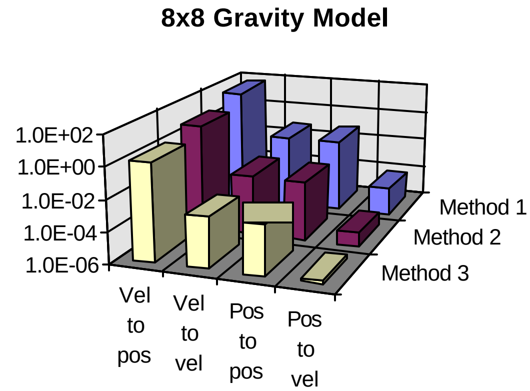

Figure 3. Standard deviation of state transition matrix projection error after 1.5 hours of propagation using a central force plus 8x8 gravity model. (See Table 4.)

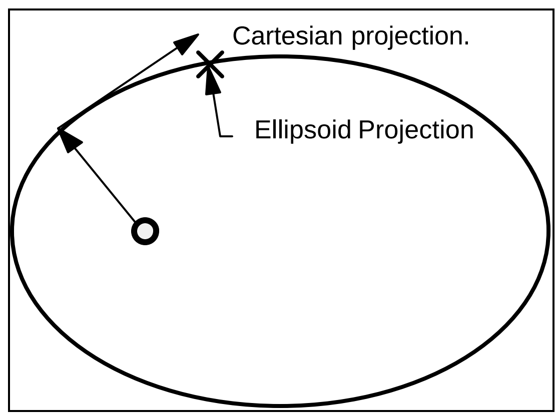

Figure 4. Differences between Cartesian and ellipsoid projections create second and higher order error when using Cartesian updates.