Discussing Three Different Cases of Interest

Tom Kerr's detailed explanation of Ralph A. William's original presentation

November 1, 2019

1 3 Cases for Sensitivity Analysis

1.1 Definition of Terms:

Assumption: There are no $\underline{x}_1$ (i.e., time-varying states) present within this analysis so they are not considered any further here below. Evidently, the nomenclature that has evolved to denote time-varying states within "sensitivity analyses" within this Trident group is $\underline{x}_1$.

$$\overset{(n+9)\times 1}{\underline{x}} = \begin{bmatrix} \overset{(n\times 1)}{\underline{x}_o} \\ \overset{(9\times 1)}{\delta \underline{S}_1} \end{bmatrix} \text{; Being constant biases so :} \dot{\underline{x}}_o \equiv 0 \text{ for all time,} \tag{1}$$

$$\overset{9\times(n+9)}{H} = \begin{bmatrix} \overset{(9\times n)}{E} & | \; I_{9\times 9} \end{bmatrix}, \tag{2}$$

$$\overset{(n+9)\times(n+9)}{\Phi} = \begin{bmatrix} I_{n\times n} & \overset{(n\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{\phi_S} \end{bmatrix}, \tag{3}$$

$$\overset{(n+9)\times(n+9)}{Q} = \begin{bmatrix} \overset{(n\times n)}{0} & \overset{(n\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{Q_S} \end{bmatrix}, \tag{4}$$

where, in the above, the states being investigated to determine their effect, by convention, are denoted with a δ appearing in front of them, as:

$$\delta \overset{(9\times 9)}{\underline{S}_1} = \delta \overset{(9\times 9)}{\underline{S}_a} + \delta \overset{(9\times 9)}{\underline{S}_b}, \tag{5}$$

and their corresponding noise covariances (for these White Gaussian Noise terms being present and to be accounted for) are:

$$\overset{(9\times 9)}{Q_S} = \overset{(9\times 9)}{Q_a} + \overset{(9\times 9)}{Q_b}, \tag{6}$$

and the general matrix E appearing on the Left Hand Side (LHS) of what is the so-called "Observation Matrix", H, appearing in Eq. 2 above, is identical, in some sense, to the role of the actual observation matrix arising in the state variable representation of the underlying system, as will be invoked for use in several different situations, as indicated, where, E is given to start with and the interpretation throughout the above is that E_v corresponds to the 3 *velocity* states (the top left $3 \times n$ block) and E_p corresponds to the 3 *position* states (the middle $3 \times n$ block) [and so-far unidentified or suppressed at the corresponding place in the earlier presentation, perhaps because it is of less direct interest in what follows, the $3 \times n$ block in the lower right corner corresponds to the 3 *"gyro tilts"* or *attitude* states]. The symbol E in Eq. 2 represents the **Sensitivity Matrix**, which is a "given" within the three exercises focused upon as Cases A, B, and C, as our primary goal here, as pursued, respectively, in Secs. 2, 3, and 4, with computational results reported in the concluding table at the very end.

1.2 Computational Evaluation Algorithms:

The following sequence of 8 familiar equations (Eqs. 7 to 13, below, constituting the calculations performed within a Kalman filter implementation) are to be repeatedly iterated to convergence:

$$\underline{\hat{x}}^{-} = \phi \underline{\hat{x}}^{0} \tag{7}$$

$$P^{-} = \phi P^{0} \phi^{T} + Q \tag{8}$$

$$K \triangleq P^{-} H^{T} \left[H P^{-} H^{T} + R \right]^{-1} \tag{9}$$

$$\underline{z} = \text{next sequentially available sensor measurement,} \tag{10}$$

$$\underline{\hat{x}}^{+} = \{I - KH\} \underline{\hat{x}}^{-} + K \underline{z} \tag{11}$$

$$P^{+} = \{I - KH\} P^{-} \{I - KH\}^{T} + KRK^{T} \tag{12}$$

$$\underline{\hat{x}}^{o} = \underline{\hat{x}}^{+} \tag{13}$$

$$P^{o} = P^{+} \tag{14}$$

The above items in red can be omitted entirely to leave only those involved in computing covariances remaining. However, to merely perform a **covariance analysis**, the actual measurement realization, $\underline{z}$, and the state estimation equations involving $\underline{\hat{x}}$ are superfluous, unnecessary, and need not be present; so skip or delete Eqs. 7, 10, 11, and 13 from the above iteration equation but store the associated computed Kalman gains, K, sequentially ordered according to their respective time-stamps (since the time sequence of K's are saved to be used later).

These same 8 equations (but reduced by 1 now by merely incrementing the time index k instead) will now be viewed in more detail below using the nomenclature of another prevalent standard convention for a Kalman filter. (There is yet another convention that only looks somewhat similar to this but is, in fact, also correct but is **not** addressed at all in the 1974 TASC textbook but *is* addressed in R. Grover Brown's Book, 4th edition, 2012. When I hear of competent people or organizations noticing or complaining about discrepancies between two different filter implementations, I usually suspect apparent contradictions between standard TASC formulation and that alternate version reported on by R. Grover Brown! Both formulations give **identical outputs** from **identical inputs**.) For reassurance that Eqs. 7 to 14 are, in fact, correct and consistent with the standard TASC convention below:

$$\hat{\underline{x}}(k+1|k) = \phi(k+1,k)\hat{\underline{x}}(k|k)$$

$$P(k+1|k) = \phi(k+1,k)P(k|k)\phi^T(k+1,k) + Q(k)$$

$$K(k) \triangleq P(k+1|k)H^T(k)\left[H(k)P(k+1|k)H^T(k) + R(k)\right]^{-1}$$

$$\underline{z}(k) = \text{next sequentially available sensor measurement at time index k,}$$

$$\hat{\underline{x}}(k+1|k+1) = \{I - K(k)H(k)\}\hat{\underline{x}}(k+1|k) + K(k)\underline{z}(k)$$

$$P(k+1|k+1) = \{I - K(k)H(k)\}P(k+1|k)\{I - K(k)H(k)\}^T + K(k)R(k)K^T(k)$$

$$k = k+1 \ \text{(increment the time index so that 2 earlier equations are now omitted)}$$

Again, to merely perform a **covariance analysis**, the actual measurement realization, $\underline{z}(k)$, and the state estimation equations involving $\hat{\underline{x}}(k+1|k)$, $\hat{\underline{x}}(k|k)$, and $\underline{z}(k)$ are superfluous, unnecessary, and need not be present; so skip or delete all equations depicted above in red from the above iteration equation but store the associated computed Kalman gains, $K(k)$ (that still correspond to when a sensor measurement was taken without needing the actual measurement itself), sequentially ordered according to their respective time-stamps or time index k (since the sequence of $K(k)$'s are saved to be used later). Deleting these three lines above in calculating these covariances is really not so strange when one recalls that for these Gaussians present throughout linear systems, the mean is independent of the variance and, likewise, the conditional mean [i.e., the optimal estimate] is independent of the conditional variance of estimation error.

2 Case A: "TRUTH!"

Beginning with Eq. 7, the corresponding discrete-time representations of state estimator, state evolution itself, and state error-in-estimation being, respectively, Eqs. 15, 16, and 17:

$$\underline{\hat{x}}^- = \phi \underline{\hat{x}}^+_{\text{old}}, \tag{15}$$

$$\underline{x}^- = \phi \underline{x}^+_{\text{old}} + \underline{q} \tag{16}$$

$$\underline{\tilde{x}} \stackrel{\triangle}{=} \underline{\hat{x}} - \underline{x} \quad \Rightarrow \quad \underline{\tilde{x}}^- = \phi \underline{\tilde{x}}^+_{\text{old}} - \underline{q} \tag{17}$$

from which we have the general discrete-time solution evolution described by:

$$\underline{\tilde{x}}_n = N \underline{\tilde{x}}_o + \sum_{i=1}^{n} A_i \underline{\tilde{x}}_i - \sum_{i=1}^{n} B_i \underline{q}_{i-1}, \tag{18}$$

where, from the above, more structural insight can be gleaned by collecting the following summarizing terms:

$$N \stackrel{\triangle}{=} \left(\prod_{j=1}^{n} \underleftarrow{[I - K_j H_j]\phi_{j-1}} \right) \equiv \pi_o, \tag{19}$$

$$A_i \stackrel{\triangle}{=} \left(\prod_{j=i+1}^{n} \underleftarrow{[I - K_j H_j]\phi_{j-1}} \right) K_i \equiv \pi_i K_i, \tag{20}$$

$$B_i \stackrel{\triangle}{=} \left(\prod_{j=i+1}^{n} \underleftarrow{[I - K_j H_j]\phi_{j-1}} \right) [I - K_i H_i] \equiv \pi_i \left[I - K_i H_i \right], \tag{21}$$

and then :

$$P \stackrel{\triangle}{=} E\left[\underline{\tilde{x}}_n \underline{\tilde{x}}_n^T \right] = N P_o N^T + \sum_{i=1}^{n} A_i R_i A_i{}^T + \sum_{i} B_i Q_{i-1} B_i{}^T, \tag{22}$$

where the above Eq. 22 final result is obtained by multiplying Eq. 18 term-by-term by it's transpose and then taking expectations throughout (on both sides) and using the fact that uncorrelated Gaussian entities are independent of the other terms and are of zero mean in order that several of the requisite intermediate terms drop out (i.e., go to zero when multiplied by a zero valued term) and the concluding Eq. 22 result is valid for *any* linear discrete-time Kalman filter with *any* gain. Within the above described derivation, we tacitly utilized the property and relationship between total expectations and conditional expectations, since the optimal estimate is, in fact, the conditional estimate given the measurements (i.e., $\underline{\hat{x}} \equiv E\left[\underline{x}|\underline{Z}\right]$) then $E\left[\underline{\hat{x}}\right] = E\left[E\left[\underline{x}|\underline{Z}\right]\right] = E\left[\underline{x}\right]$. Since above Eq. 22 is true for <u>any</u> set of gains K_i, proceed by holding the K_i $fixed$, then $\partial P_{ii}/\partial P^o_{jj} = N^2_{ij}$. There is, in fact, a precedent for using partial derivatives in this manner [**w**ith **r**espect **t**o (wrt) some states but not wrt other elements of the same covariance matrix P], while investigating the general expression for the solution to the covariance evolution (i.e., "Variations of the Sensitivity Matrix" due to scalar parameters on pages 21 to 26 for a discrete Kalman filter, and also in Appendix A for EKF's in Eqs. A.6 and A.7) that arise, as reported from SRI Information and Control Laboratory by B. L. Ho, **Sensitivity of Kalman Filter with Respect to Parameter Variations**, SRI Project 5188-305, Memorandum 33, March 1968. Their list of cited references from both the open and closed literature is also familiar (with very familiar authors such as: R. E. Larson, David Luenberger, B. Friedland, T. Nishimura).

*Since $P(P^o)$ is linear in P^o and the above result is exact for finite $\Delta P_{ii} + \Delta P^o_{jj}$! (Therefore use the above result and calculate Δr_i, as $r_i \equiv \sigma_i/\sigma_i^o$ in nonlinear manner.)

*Furthermore, $\Delta P_{ii} = N_{ij}^2 \Delta P_{ij}^o$ is always an upper bound to the optimum results obtained by letting $K_i \rightarrow K_{i,optimal}$.

*Similar arguments apply to evaluating the effect of changes to the $R_i + Q_i$.

*These results are general and DO NOT assume that $\dot{\underline{\tilde{x}}} = 0$. For that case, $N = PP_o^{-1}$ if you start with the optimum K_i.

Another fresh authoritative discussion of sensitivity analysis that is different yet consistent with the TASC approach is found in Sec. 7.8 **Error Budgets and Sensitivity Analysis** (pp. 313-317) by Mohinder S. Grewal and Angus P. Andrews, **Kalman Filtering: Theory and Practice**, Prentice-Hall Information and System Sciences Series, Simon & Schuster/A Viacom Company, Upper Saddle River, NJ, 1993.

Ralph Williams obtained computed tabular evaluation results "using only ARW and the 2[nd] run used gains obtained via the iteration of Sec. 1.2 first but then with other noises zeroed out and using proper initial conditions as being also totally zeroed out (i.e., $p^o \equiv 0$)". While it is true that the initial condition for the error in estimation, $\tilde{x}$, (being the mean) should be the zero vector $\underline{0}$; however, the error vector $\tilde{x}$ is Gaussian and, as such, is a two parameter family with both a mean vector and a covariance matrix. Theoretically, the initial condition covariance matrix should never be the zero matrix, $\overset{(n\times n)}{0}$. Theoretically, the initial condition covariance of estimation error matrix must be positive *definite* (unlike the situation for the covariance of *plant* or *process* noise, which can be positive *semi-definite* without incurring any problem). No element along the principal diagonal of P_o can be zero let alone the entire covariance matrix.

From information pertaining to numerical properties of the solution to the Ricatti equation that I obtained in the mid- to late 1970's, I know that when the system is *Observable* and *Controllable*, then the Riccati equation can have n^2 possible solutions but can have only one unique solution when P_o is positive definite (within this context, bad behavior of attempted Riccati equation computed solutions is more understandable when accumulated computational roundoff error or truncation error intefere and set things off in a wrong direction). Just what value should be used for P_o? I don't know. Perhaps, if one used the *Information Matrix* formulation instead for these Kalman filter covariance calculations, one could use it with very large yet finite initial values on the main diagonal and zeros elsewhere (to correspond to very small values indicated to be desired for P_o) yet still have the computed results be positive definite for both!

3 Case B: CSDL and JHU/APL

$$\overset{(n+9+9)\times 1}{\underline{\mathrm{x}}} = \begin{bmatrix} \overset{(n\times 1)}{\underline{\mathrm{x}}_o} \\ \overset{(9\times 1)}{\delta\underline{\mathrm{S}}_a} \\ \overset{(9\times 1)}{\delta\underline{\mathrm{S}}_b} \end{bmatrix} ;\ \text{Being constant biases}:\ \dot{\underline{\mathrm{x}}}_o \equiv 0 \text{ for all time,} \tag{23}$$

$$\overset{9\times(n+9+9)}{H} = \left[\overset{(9\times n)}{E} \mid I_{9\times 9} \mid I_{9\times 9} \right], \tag{24}$$

$$\overset{(n+9+9)\times(n+9+9)}{\Phi} = \begin{bmatrix} I_{n\times n} & \overset{(n\times 9)}{0} & \overset{(n\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{\phi_S} & \overset{(9\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{0} & \overset{(9\times 9)}{\phi_S} \end{bmatrix}, \tag{25}$$

$$\overset{(n+9+9)\times(n+9+9)}{Q} = \begin{bmatrix} \overset{(n\times n)}{0} & \overset{(n\times 9)}{0} & \overset{(n\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{Q_a} & \overset{(9\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{0} & \overset{(9\times 9)}{Q_b} \end{bmatrix}, \tag{26}$$

where, in the above, the states being investigated to determine their effect, by convention, are denoted with a δ appearing in front of them. However, here

$$\delta \overset{(9\times 9)}{\underline{\mathrm{S}}_1} = \delta \overset{(9\times 9)}{\underline{\mathrm{S}}_a} + \delta \overset{(9\times 9)}{\underline{\mathrm{S}}_b}, \tag{27}$$

since each constituent component is called out and identified separately in Eq. 23 and, likewise, their corresponding associated noise covariances (for these White Gaussian Noise terms being present) are already properly accounted for separately in Eq. 24. However, the appropriate general E still appears on the LHS of the Observation Matrix, H, in Eq. 24, as it should.

The corresponding computed tabular evaluation results for Case B: CSDL JHU/APL were obtained, according to Ralph Williams, "using the same approach with indicated changes in H and indicated changes in state definitions corresponding to Eq. 1 and 3 and, lastly, changes in Q matrix corresponding to the indicated changes in Eq. 4".

4 Case C: CSDL (with ARW iteration contributions)

To evaluate for **A**ngle **R**andom **W**alk (ARW) associated with the gyro attitude, again assume that there are no $\underline{\mathrm{x}}_1$ states present and also assume that the effect on the system is only due to the presence of process noise, as represented by:

$$\overset{(n+9+9)\times 1}{\underline{\mathrm{x}}} = \begin{bmatrix} \overset{(n\times 1)}{\underline{\mathrm{x}}_o} \\ \overset{(9\times 1)}{\delta \underline{\mathrm{S}}_1} \\ \overset{(9\times 1)}{\delta \underline{\mathrm{S}}_a} \end{bmatrix} \text{; Being constant biases : } \dot{\underline{\mathrm{x}}}_o \equiv 0 \text{ for all time,} \tag{28}$$

$$\overset{9\times(n+9+9)}{H} = \left[\overset{(9\times n)}{E} \mid I_{9\times 9} \mid \overset{(9\times 9)}{0} \right], \tag{29}$$

$$\overset{(n+9+9)\times(n+9+9)}{\Phi} = \begin{bmatrix} I_{n\times n} & \overset{(n\times 9)}{0} & \overset{(n\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{\phi_S} & \overset{(9\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{0} & \overset{(9\times 9)}{\phi_S} \end{bmatrix}, \tag{30}$$

$$\overset{(n+9+9)\times(n+9+9)}{Q} = \begin{bmatrix} \overset{(n\times n)}{0} & \overset{(n\times 9)}{0} & \overset{(n\times 9)}{0} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{Q_S} & \overset{(9\times 9)}{Q_a} \\ \overset{(9\times n)}{0} & \overset{(9\times 9)}{Q_a} & \overset{(9\times 9)}{Q_a} \end{bmatrix}, \tag{31}$$

where, in the above, the states being investigated to determine their effect, by convention, are denoted with a δ appearing in front of them, as:

$$\delta \overset{(9\times 9)}{\underline{\mathrm{S}}_1} = \delta \overset{(9\times 9)}{\underline{\mathrm{S}}_a} + \delta \overset{(9\times 9)}{\underline{\mathrm{S}}_b}, \tag{32}$$

and their corresponding noise covariances (for these White Gaussian Noise terms being present and to be accounted for) since the constituent components are uncorrelated Gaussians (and so independent) and of zero-mean (therefore cross-terms drop out when Eq. 32 is multiplied by its transpose and expectations taken throughout) to yield:

$$\overset{(9\times 9)}{Q_S} = \overset{(9\times 9)}{Q_a} + \overset{(9\times 9)}{Q_b}, \tag{33}$$

and the general E appears on the left hand side (LHS) of the Observation Matrix, H, in Eq. 29.

To properly account for inherent cross-correlations present in the system formulation of Eqs. 28 to 33, now forming the appropriate matrix and taking expectations (also denoted by an upper case E)throughout yields:

$$E\left[\overset{(n+9+9)\times 1}{\underline{\mathrm{x}}}\overset{1\times(n+9+9)}{\underline{\mathrm{x}}^T}\right] = \begin{bmatrix} E\left[\overset{(n\times 1)}{\underline{\mathrm{x}}_o}\overset{(1\times n)}{\underline{\mathrm{x}}_o{}^T}\right] & E\left[\overset{(n\times 1)}{\underline{\mathrm{x}}_o}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_1^T}\right] & E\left[\overset{(n\times 1)}{\underline{\mathrm{x}}_o}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_a^T}\right] \\ E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_1}\overset{(1\times n)}{\underline{\mathrm{x}}_o{}^T}\right] & E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_1}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_1^T}\right] & E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_1}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_a^T}\right] \\ E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_a}\overset{(1\times n)}{\underline{\mathrm{x}}_o{}^T}\right] & E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_a}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_1^T}\right] & E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_a}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_a^T}\right] \end{bmatrix}$$

$$= \begin{bmatrix} E\left[\overset{(n\times 1)}{\underline{\mathrm{x}}_o}\overset{(1\times n)}{\underline{\mathrm{x}}_o{}^T}\right] & E\left[\overset{(n\times 1)}{\underline{\mathrm{x}}_o}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_1^T}\right] & E\left[\overset{(n\times 1)}{\underline{\mathrm{x}}_o}\overset{(1\times 9)}{\delta\underline{\mathrm{S}}_a^T}\right] \\ E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_1}\overset{(1\times n)}{\underline{\mathrm{x}}_o{}^T}\right] & [Q_a + Q_b] & Q_a \\ E\left[\overset{(9\times 1)}{\delta\underline{\mathrm{S}}_a}\overset{(1\times n)}{\underline{\mathrm{x}}_o{}^T}\right] & Q_a & Q_a \end{bmatrix} \quad (34)$$

The corresponding computed tabular evaluation results for Case C: CSDL (with ARW iterations) were obtained by Ralph Williams using the same approach as for the Case B JHU/APL evaluation approach, with indicated changes in H and indicated changes in state definitions corresponding to Eq. 1 and 3 and, lastly, changes in Q matrix corresponding to the indicated changes in Eq. 4", which confirms the correctness of Eq. 31 (just as Ralph Williams had originally presented) but now also displays slightly more internal structural detail here to allow easier, more direct reader confirmation.

Ralph William's computational evaluation results are found in the following table, corresponding, respectively, to Cases A, B, and C above.