Cost Analysis of Square Root Information Filtering and Smoothing with a Mixed Real-Integer State

December 28, 2013

The following analysis provides an analysis of the transformation undergone by the maximum *a posteriori* cost function as *a priori* information, process noise, and measurements are ingested by a square root information filter and smoother with a mixed real-integer state. The analysis supports claims made in [1].

1 Filtering Analysis

The initial cost function at time i, embedded with square root information equations that encode the *a priori* information about the real-valued state and the process noise and with square root information equations relating the new measurements to the real- and integer- valued state can be written as follows:

$$\begin{aligned}
J \quad = \quad & \underbrace{\|\mathbf{R}_{\mathrm{xx}i}\mathbf{x}_i - \mathbf{z}_{\mathrm{xi}}\|^2}_{\textit{A priori}\text{ real-valued state information}} + \underbrace{\|\mathbf{R}_{nni}\mathbf{n}_i - \mathbf{z}_{\mathrm{ni}}\|^2}_{\textit{A priori}\text{ integer-valued state information}} \\
& + \underbrace{\sum_{k=i}^{K-1} \|\mathbf{R}_{\mathrm{ww}}\mathbf{w}_k - 0\|^2}_{\textit{A priori}\text{ process noise information}} + \underbrace{\sum_{k=i}^{K} \|\mathbf{H}_{\mathrm{x}k}\mathbf{x}_k + \mathbf{H}_{\mathrm{n}k}\mathbf{n}_k - \mathbf{z}_k\|^2}_{\text{Measurements}}
\end{aligned}$$

Note that the nonhomogeneous measurement term for the *a priori* process noise is set to 0 since the process noise is assumed to be zero-mean.

Stacking the a priori terms at time i we can rewrite the cost as:

$$\begin{aligned}
J \quad = \quad & \left\| \begin{bmatrix} \mathbf{R}_{ww} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{xxi} & \mathbf{0} \\ & & \mathbf{R}_{nni} \end{bmatrix} \begin{bmatrix} \mathbf{w}_i \\ \mathbf{x}_i \\ \mathbf{n}_i \end{bmatrix} - \begin{bmatrix} 0 \\ \mathbf{z}_{xi} \\ \mathbf{z}_{ni} \end{bmatrix} \right\|^2 \\
& + \sum_{k=i+1}^{K} \|\mathbf{H}_{\mathrm{x}k}\mathbf{x}_k + \mathbf{H}_{\mathrm{n}k}\mathbf{n}_k - z_k\|^2 + \sum_{k=i+1}^{K-1} \|\mathbf{R}_{\mathrm{ww}}\mathbf{w}_k\|^2
\end{aligned}$$

To form the big block matrix above, we used the following identity:

$$\|a\|^2 + \|b\|^2 = \left\| \begin{bmatrix} a \\ b \end{bmatrix} \right\|^2 \tag{1}$$

$\mathbf{x}_i$and $\mathbf{n}_i$ can be written in terms of $\mathbf{x}_{i+1}$ and $\mathbf{n}_{i+1}$according to their respective dynamics models:

$$\begin{aligned}\mathbf{x}_i &= \mathbf{\Phi}^{-1}\left[\mathbf{x}_{i+1}-\mathbf{w}_i\right] \\ \mathbf{n}_i &= \begin{cases}\mathbf{n}_{i+1}(1:m) & \text{if a new integer becomes present between time } i \text{ and } i+1 \\ \mathbf{n}_{i+1} & \text{otherwise}\end{cases}\end{aligned}$$

where$m = \mathrm{numel}(\mathbf{n}_{i+1})-1$. Substituting $\mathbf{x}_i$with $\mathbf{x}_{i+1}$ and $\mathbf{n}_i$ with $\mathbf{n}_{i+1}$into the above cost function, we arrive at:

$$\begin{aligned}J &= \left\| \underbrace{\begin{bmatrix}\mathbf{R}_{ww} & \mathbf{0} & [\mathbf{0}|\mathbf{0}] \\ -\mathbf{R}_{xxi}\mathbf{\Phi}^{-1} & \mathbf{R}_{xxi}\mathbf{\Phi}^{-1} & [\mathbf{R}_{xni}|\mathbf{0}] \\ \mathbf{0} & \mathbf{0} & [\mathbf{R}_{nni}|\mathbf{0}]\end{bmatrix}}_{\text{Big block matrix}} \begin{bmatrix}\mathbf{w}_i \\ \mathbf{x}_{i+1} \\ \mathbf{n}_{i+1}\end{bmatrix} - \begin{bmatrix}0 \\ \mathbf{z}_{xi} \\ \mathbf{z}_{ni}\end{bmatrix} \right\|^2 \\ &\quad + \sum_{k=i+1}^{K} \left\|\mathbf{H}_{\mathrm{x}k}\mathbf{x}_k + \mathbf{H}_{\mathrm{n}k}\mathbf{n}_k - z_k\right\|^2 + \sum_{k=i+1}^{K-1}\left\|\mathbf{R}_{\mathrm{ww}}\mathbf{w}_k\right\|^2\end{aligned}$$

where $[x|\mathbf{0}]$ represents the augmentation of the square-root information matrices with a column of $0's$ to represent the introduction of a new integer between time i and $i+1$. Zeros are used to indicate that no knowledge is known about this new integer.

By QR factorizing the big block matrix and left multiplying the insides of the first term by Q^{T}, one can propagate the state elements forward to $i+1$ and arrive at the following equivalent cost function:

$$\begin{aligned}J &= \left\| \underbrace{\begin{bmatrix}\overline{\mathbf{R}}_{wwi} & \overline{\mathbf{R}}_{wx,i+1} & \overline{\mathbf{R}}_{wn,i+1} \\ \mathbf{0} & \overline{\mathbf{R}}_{xx,i+1} & \overline{\mathbf{R}}_{xn,i+1} \\ \mathbf{0} & \mathbf{0} & \overline{\mathbf{R}}_{nn,i+1}\end{bmatrix}}_{\text{Big block matrix}} \begin{bmatrix}\mathbf{w}_i \\ \mathbf{x}_{i+1} \\ \mathbf{n}_{i+1}\end{bmatrix} - \begin{bmatrix}\overline{\mathbf{z}}_{wi} \\ \overline{\mathbf{z}}_{x,i+1} \\ \overline{\mathbf{z}}_{n,i+1}\end{bmatrix} \right\|^2 \\ &\quad + \sum_{k=i+1}^{K} \left\|\mathbf{H}_{\mathrm{x}k}\mathbf{x}_k + \mathbf{H}_{\mathrm{n}k}\mathbf{n}_k - z_k\right\|^2 + \sum_{k=i+1}^{K-1}\left\|\mathbf{R}_{\mathrm{ww}}\mathbf{w}_k\right\|^2\end{aligned}$$

Next, by unpacking the process noise data equation from the big block matrix (see (1)) and replacing it with the latest measurement at time $i+1$, one can rewrite the cost as:

$$\begin{aligned} J &= \left\| \underbrace{\begin{bmatrix} \overline{\mathbf{R}}_{xx,i+1} & \overline{\mathbf{R}}_{xn,i+1} \\ \mathbf{0} & \overline{\mathbf{R}}_{nn,i+1} \\ \mathbf{H}_{i+1} & \mathbf{H}_{n,i+1} \end{bmatrix}}_{\text{Big block matrix}} \begin{bmatrix} \mathbf{x}_{i+1} \\ \mathbf{n}_{i+1} \end{bmatrix} - \begin{bmatrix} \overline{\mathbf{z}}_{x,i+1} \\ \overline{\mathbf{z}}_{n,i+1} \\ z_{i+1} \end{bmatrix} \right\|^2 \\ &\quad + \left\| \overline{\mathbf{R}}_{wwi}\mathbf{w}_i + \overline{\mathbf{R}}_{wx,i+1}\mathbf{x}_{k+1} + \overline{\mathbf{R}}_{wn,i+1}\mathbf{n}_{i+1} - \overline{\mathbf{z}}_{wi} \right\|^2 \\ &\quad + \sum_{k=i+2}^{K} \left\| \mathbf{H}_{\mathrm{x}k}\mathbf{x}_k + \mathbf{H}_{\mathrm{n}k}\mathbf{n}_k - z_k \right\|^2 + \sum_{k=1}^{K-1} \left\| \mathbf{R}_{\mathrm{ww}}\mathbf{w}_k \right\|^2 \end{aligned}$$

Next QR factorize this new big matrix and left multiply by Q^{T} to perform a measurement update to get :

$$\begin{aligned} J &= \left\| \begin{bmatrix} \mathbf{R}_{xx,i+1} & \mathbf{R}_{xn,i+1} \\ \mathbf{0} & \mathbf{R}_{nn,i+1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{i+1} \\ \mathbf{n}_{i+1} \end{bmatrix} - \begin{bmatrix} \mathbf{z}_{x,i+1} \\ \mathbf{z}_{n,i+1} \\ \mathbf{z}_{r,i+1} \end{bmatrix} \right\|^2 \\ &\quad + \left\| \overline{\mathbf{R}}_{wwi}\mathbf{w}_i + \overline{\mathbf{R}}_{wx,i+1}\mathbf{x}_{k+1} + \overline{\mathbf{R}}_{wn,i+1}\mathbf{n}_{i+1} - \overline{\mathbf{z}}_{wi} \right\|^2 \\ &\quad + \sum_{k=i+2}^{K} \left\| \mathbf{H}_{\mathrm{x}k}\mathbf{x}_k + \mathbf{H}_{\mathrm{n}k}\mathbf{n}_k - z_k \right\|^2 + \sum_{k=1}^{K-1} \left\| \mathbf{R}_{\mathrm{ww}}\mathbf{w}_k \right\|^2 \end{aligned}$$

The cost function can now be unpacked from its big block matrix form and written as follows:

$$\begin{aligned} J &= \left\| \mathbf{R}_{xx,i+1}\mathbf{x}_{i+1} + \mathbf{R}_{xn,i+1}\mathbf{n}_{i+1} - \mathbf{z}_{x,i+1} \right\|^2 \\ &\quad + \left\| \mathbf{R}_{nn,i+1}\mathbf{n}_{i+1} - \mathbf{z}_{n,i+1} \right\|^2 + \left\| \mathbf{z}_{r,i+1} \right\|^2 \\ &\quad + \left\| \overline{\mathbf{R}}_{wwi}\mathbf{w}_i + \overline{\mathbf{R}}_{wx,i+1}\mathbf{x}_{i+1} + \overline{\mathbf{R}}_{wn,i+1}\mathbf{n}_{i+1} - \overline{\mathbf{z}}_{w,i} \right\|^2 \\ &\quad + \sum_{k=i+2}^{K} \left\| \mathbf{H}_k\mathbf{x}_k + \mathbf{H}_{nk}\mathbf{n}_k - z_k \right\|^2 + \sum_{k=i+1}^{K-1} \left\| \mathbf{R}_{ww} w_k \right\|^2 \end{aligned}$$

The above steps can be reiterated for $i \leftarrow i+1$.

After incorporating all measurements up to time $i = K$, the cost function can be written as:

$$\begin{aligned} J &= \left\| \mathbf{R}_{xx,K}\mathbf{x}_K + \mathbf{R}_{xn,K}\mathbf{n}_K - \mathbf{z}_{x,K} \right\|^2 + \left\| \mathbf{R}_{nn,K}\mathbf{n}_K - \mathbf{z}_{n,K} \right\|^2 \\ &\quad + \sum_{k=0}^{K-1} \left\| \overline{\mathbf{R}}_{wwk}\mathbf{w}_k + \overline{\mathbf{R}}_{wx,k+1}\mathbf{x}_{k+1} + \overline{\mathbf{R}}_{wn,k+1}\mathbf{n}_{k+1} - \overline{\mathbf{z}}_{w,k} \right\|^2 \qquad (2) \\ &\quad + \sum_{k=1}^{K} \left\| \mathbf{z}_{r,k} \right\|^2 \\ &= 0 + \left\| \mathbf{R}_{nn,K}\hat{\mathbf{n}}_K - \mathbf{z}_{n,K} \right\|^2 + 0 + \sum_{k=1}^{K} \left\| \mathbf{z}_{r,k} \right\|^2 \\ &\quad \text{when minimized over } \mathbf{n}_K,\ \mathbf{x}_k\text{'s and } \mathbf{w}_k\text{'s} \end{aligned}$$

2 Smoothing Analysis

Next, we can derive the above mixed integer-real cost function after smoothing. We first define the smoother's initial nonhomogeneous term $\mathbf{z}^{\star}_{x,K}$, initial square-root information matrix $\mathbf{R}^{\star}_{xx,K}$, and the integer-valued vector estimate $\hat{\mathbf{n}}_K$ at time K. The new quantities are derived from the variables output after running the filter:

$$\begin{aligned}
\mathbf{z}^{\star}_{x,K} &= \mathbf{z}_{x,K} - \mathbf{R}_{xn,K}\hat{\mathbf{n}}_K \\
\mathbf{R}^{\star}_{xx,K} &= \mathbf{R}_{xx,K} \\
\hat{\mathbf{n}}_K &= \min_{\mathbf{n}_i \in \mathbb{Z}^K} \left\|\mathbf{R}_{nn,K}\mathbf{n}_i - \mathbf{z}_{n,K}\right\|^2
\end{aligned}$$

These variables can be substituted into the cost function in (2) to arrive at:

$$\begin{aligned}
J &= \left\|\mathbf{R}^{\star}_{xx,K}\mathbf{x}_K - \mathbf{z}^{\star}_{x,K}\right\|^2 + \left\|\mathbf{R}_{nn,K}\hat{\mathbf{n}}_K - \mathbf{z}_{n,K}\right\|^2 \\
&\quad \sum_{k=0}^{K-1} \left\|\overline{\mathbf{R}}_{wwk}\mathbf{w}_k + \overline{\mathbf{R}}_{wx,k+1}\mathbf{x}_{k+1} + \overline{\mathbf{R}}_{wn,k+1}\hat{\mathbf{n}}_{k+1} - \overline{\mathbf{z}}_{w,k}\right\|^2 \\
&\quad \sum_{k=1}^{K} \left\|\mathbf{z}_{r,k}\right\|^2
\end{aligned} \tag{3}$$

We can use the following dynamics equation to eliminate $\mathbf{x}_k$ in favor of $\mathbf{x}_{K-1}$ and $\hat{\mathbf{n}}_K$ in favor of $\hat{\mathbf{n}}_{K-1}$:

$$\begin{aligned}
\mathbf{x}_k &= \mathbf{\Phi}\mathbf{x}_{K-1} + \mathbf{w}_{K-1} \\
\hat{\mathbf{n}}_K &= \begin{cases} \begin{bmatrix} \hat{\mathbf{n}}_{K-1} & \hat{n}_{i_K} \end{bmatrix} & \text{if a new integer becomes present between time } K-1 \text{ and } K \\ \hat{\mathbf{n}}_{K-1} & \text{otherwise} \end{cases}
\end{aligned}$$

where $\hat{n}_{i_K}$ is the last component of $\hat{\mathbf{n}}_K$, the integer that arose between time $K-1$ and K, if there was one. Stacking the first term and part of the third term of (3) and applying these substitutions, we arrive at

$$\begin{aligned}
J &= \left\| \begin{bmatrix} \overline{\mathbf{R}}_{ww,K-1} + \overline{\mathbf{R}}_{wx,K} & \overline{\mathbf{R}}_{wx,K}\mathbf{\Phi} \\ \mathbf{R}^{\star}_{xx,K} & \overline{\mathbf{R}}^{\star}_{xx,K}\mathbf{\Phi} \end{bmatrix} \begin{bmatrix} \mathbf{w}_{K-1} \\ \mathbf{x}_{K-1} \end{bmatrix} - \begin{bmatrix} \overline{\mathbf{z}}_{w,K-1} - \overline{\mathbf{R}}_{wn,K}\hat{\mathbf{n}}_K \\ \overline{\mathbf{z}}^{\star}_{x,K} \end{bmatrix} \right\|^2 \\
&\quad + \sum_{k=0}^{K-2} \left\|\overline{\mathbf{R}}_{ww,k}\mathbf{w}_k + \overline{\mathbf{R}}_{wx,k+1}\mathbf{x}_{k+1} - \overline{\mathbf{z}}_{w,k}\right\|^2 + \left\|\mathbf{R}_{nn,K}\hat{n}_K - \mathbf{z}_{n,K}\right\|^2 + \sum_{k=1}^{K}\left\|\mathbf{z}_{r,k}\right\|^2
\end{aligned}$$

After QR factorization and left multiplying, we propagate the smoother from time K to $K-1$:

$$\begin{aligned}
J &= \left\| \begin{bmatrix} \mathbf{R}^{\star}_{ww,K-1} & \mathbf{R}^{\star}_{wx,K-1} \\ \mathbf{0} & \mathbf{R}^{\star}_{xx,K-1} \end{bmatrix} \begin{bmatrix} \mathbf{w}_{K-1} \\ \mathbf{x}_{K-1} \end{bmatrix} - \begin{bmatrix} \mathbf{z}^{\star}_{w,K-1} \\ \mathbf{z}^{\star}_{x,K-1} \end{bmatrix} \right\|^2 \\
&\quad + \sum_{k=0}^{K-2} \left\|\overline{\mathbf{R}}_{ww,k}\mathbf{w}_k + \overline{\mathbf{R}}_{wx,k+1}\mathbf{x}_{k+1} + \overline{\mathbf{R}}_{wn,k+1}\hat{n}_{k+1} - \overline{\mathbf{z}}_{w,k}\right\|^2 \\
&\quad + \left\|\mathbf{R}_{nn,K}\hat{n}_K - \mathbf{z}_{n,K}\right\|^2 + \sum_{k=1}^{K}\left\|\mathbf{z}_{r,k}\right\|^2
\end{aligned}$$

Unpacking the cost function, decrementing i by 1, we can repeat the last two steps of subsitution and QR factorization until $i = 0$ and the cost function looks as follows:

$$\begin{aligned} J &= \left\|\mathbf{R}^{\star}_{xx,0}x_0 - \mathbf{z}^{\star}_{x,0}\right\|^2 + \sum_{k=0}^{K-1} \left\|\mathbf{R}^{\star}_{ww,k}\mathbf{w}_k + \mathbf{R}^{\star}_{wx,k}\mathbf{x}_k - \mathbf{z}^{\star}_{w,k}\right\|^2 \\ &\quad + \left\|\mathbf{R}_{nn,K}\hat{n}_K - \mathbf{z}_{n,K}\right\|^2 + \sum_{k=1}^{K} \left\|\mathbf{z}_{r,k}\right\|^2 \\ &= 0 + 0 + \left\|\mathbf{R}_{nn,K}\hat{n}_K - \mathbf{z}_{n,K}\right\|^2 + \sum_{k=1}^{K} \left\|\mathbf{z}_{r,k}\right\|^2 \\ &\quad \text{when minimized over } \mathbf{x}_k\text{'s and } \mathbf{w}_k\text{'s} \end{aligned}$$

From this, it is clear that the minimum cost for both the filter and the smoother are the same, however the minimizing values of $\mathbf{x}'_k s$ and $\mathbf{w}'_k s$ will be different. The only difference is that the dynamics are enforced going backward. A good analogy here is that between the filtering and smoothing stage, the minimization landscape is changing, however, the minimum value remains the same.

References

[1] K. M. Pesyna Jr., Z. M. Kassas, R. W. Heath Jr., and T. E. Humphreys, "A phase-reconstruction technique for low-power centimeter-accurate mobile positioning," *IEEE Transactions on Signal Processing*, 2013, submitted for review.