1 Recursive Least Squares Estimation

We have seen how to calculate an estimate of a constant with the least squares solution. We build up a model matrix H from the different measurements and solve a matrix equation. This requires all the measurements to be available at the time we calculate the solution, what happens if we want to update the estimate as the measurements become available? Do we keep storing all the past measurements and add new rows to the H matrix? A better way is to recursive update the estimate every time a new measurement is available, doing it this way would mean that we don't have to keep a history of all the previous measurements, we would only have to keep the best estimate from last time.

Suppose $\hat{x}_k$ is the estimated constant n-dimensional vector that includes all the measurement information for all the measurements up to and including the k-th measurement. Let y_k be the k-th noisy measurement vector that is a linear combination of x via the model matrix H_k and is corrupted with some random additive measurement noise v_k.

We can form a linear recursive estimator with the form:

$$
\begin{aligned}
y_k &= H_k x + v_k && (1)\\
\hat{x}_k &= \hat{x}_{k-1} + K_k(y_k - H_k \hat{x}_{k-1}) && (2)
\end{aligned}
$$

We compute the latest estimate for $\hat{x}_k$ based on the previous estimate $\hat{x}_{k-1}$ and the new measurement y_k. The amount that the estimate $\hat{x}$ changes from the previous estimate is based on the error between the current measurement y_k and the estimate measurement calculated from $H_k \hat{x}_{k-1}$. The error is multiplied by an gain matrix K_k to calculate the amount to update the estimate. This gain matrix K_k will be selected to be a matrix which is optimum, to make the solution minimize a least squares cost function.

Lets have a look at the estimation error $\epsilon_k = x - \hat{x}_k$ and write it out in a recursive form:

$$
\begin{aligned}
\epsilon_k &= x - \hat{x}_k && (3)\\
&= x - \hat{x}_{k-1} - K_k(y_k - H_k \hat{x}_{k-1}) && (4)\\
&= \epsilon_{k-1} - K_k(H_k x + v_k - H_k \hat{x}_{k-1}) && (5)\\
&= \epsilon_{k-1} - K_k H_k (x - \hat{x}_{k-1}) - K_k v_k && (6)\\
&= (I - K_k H_k)\epsilon_{k-1} - K_k v_k && (7)
\end{aligned}
$$

So the latest estimation error is a function of the previous estimation error and the current measurement noise. Lets define a estimation error covariance matrix P_k which is a measure of the error in the estimation:

$$
P_k = \epsilon_k \epsilon_k^T \quad (8)
$$

Now lets expand the covariance matrix P_k into a recursive form:

$$
\begin{aligned}
P_k &= \epsilon_k \epsilon_k^T && (9)\\
&= [(I - K_k H_k)\epsilon_{k-1} - K_k v_k][\dots]^T && (10)\\
&= (I - K_k H_k)\epsilon_{k-1}\epsilon_{k-1}^T (I - K_k H_k)^T + K_k v_k v_k^T K_k^T && (11)\\
&\quad - (I - K_k H_k)\epsilon_{k-1} v_k^T K_k^T - K_k v_k \epsilon_{k-1}^T (I - K_k H_k)^T
\end{aligned}
$$

Now we know $P_{k-1} = \epsilon_{k-1}\epsilon_{k-1}^T$ is the estimation error covariance matrix for the previous estimate and the measurement noise covariance is $R_k = v_k v_k^T$ from the noise properties of the measurement. We also know that the estimation error for $k-1$ is independent of the noise on measurement k, so that $E(\epsilon_{k-1} v_k^T) = E(v_k \epsilon_{k-1}^T) = 0$, therefore we can simplify the recursive covariance matrix equation to:

$$
P_k = (I - K_k H_k) P_{k-1} (I - K_k H_k)^T + K_k R_k K_k^T \tag{12}
$$

So now lets define a cost function J_k to minimize the sum of all the estimation errors:

$$
\begin{aligned}
J_k &= (x_1 - \hat{x}_1)^2 + (x_2 - \hat{x}_2)^2 + \cdots + (x_k - \hat{x}_k)^2 && (13)\\
&= \epsilon_1^2 + \epsilon_2^2 + \cdots + \epsilon_k^2 && (14)\\
&= \epsilon_k^T \epsilon_k && (15)
\end{aligned}
$$

So we can use the relationship $Tr(P_k) = \epsilon_k^T \epsilon_k$, so that minimizing the cost function J_k is the same as minimizing the trace of the estimation error covariance matrix P_k.

We have the cost function J that we would like to minimize, so we can simply follow the method we have used in the past, but this time instead of calculating $\hat{x}$ to minimize the solution, we will calculate the optimum gain matrix K_k which will minimize the estimation error and hence drive $\hat{x}_k$ towards x.

The derivative of the cost function is:

$$
\frac{\partial J_k}{\partial K_k} = -2(I - K_k H_k) P_{k-1} H_k^T + 2 K_k R_k \tag{16}
$$

Setting the derivative of the cost function to zero and solving for K_k gives:

$$
\begin{aligned}
K_k R_k &= (I - K_k H_k) P_{k-1} H_k^T && (17)\\
K_k (R_k + H_k P_{k-1} H_k^T) &= P_{k-1} H_k^T && (18)\\
K_k &= P_{k-1} H_k^T (H_k P_{k-1} H_k^T + R_k)^{-1} && (19)
\end{aligned}
$$

So to use the recursive least squares you first initialize the estimator:

$$
\hat{x}_0 = E(x) \tag{20}
$$

$$
P_0 = E[(x - \hat{x}_0)(x - \hat{x}_0)^T] \tag{21}
$$

So that $\hat{x}_0$ is the best estimate or initial guess and P_0 is the current uncertainty of the estimate.

Next every time a new measurement y_k is made then you update the solution using:

$$K_k = P_{k-1} H_k^T (H_k P_{k-1} H_k^T + R_k)^{-1} \tag{22}$$

$$\hat{x}_k = \hat{x}_{k-1} + K_k (y_k - H_k \hat{x}_{k-1}) \tag{23}$$

$$P_k = (I - K_k H_k) P_{k-1} (I - K_k H_k)^T + K_k R_k K_k^T \tag{24}$$