# Chapter #3

## EEE 8007

## Digital Control

- **Kalman Filter**

## 1. Derivation of the KF

The deterministic discrete time state space model of a system is given by:

$$\mathbf{X}(k+1) = \mathbf{\Phi}(k)\mathbf{X}(k) + \mathbf{\Gamma}(k)\mathbf{U}(k)$$
$$\mathbf{Y}(k) = \mathbf{H}(k)\mathbf{X}(k)$$

The above representation of the system assumes that there is "perfect" knowledge of the system and the sensors have a perfect response.

Unfortunately the model (which is a mathematical expression of the actual system) will merely be an approximation and this only under special assumptions/conditions. For example for the cage rotor IM it is assumed that there is an equivalent three phase winding on the rotor which may be only a crude assumption. Furthermore the nominal values of the system may vary during operation. For example the rotor resistances of the IM may vary up to 50%. Theoretically all these changes and uncertainties can be model exactly by differential equations. But this is a very tedious and in some cases impossible task. One solution is to assume that all these changes will have the same effect to the system as one random signal that is assumed to be a white noise Gaussian signal with zero mean. Here this random signal will be called the "model uncertainty noise" and will be denoted as **W**. The number of these noisy signals will depend on the order of the system. Hence for a fourth order system, four different signals will be needed. These noisy signals can be coupled or not, depending on the system. Their variance will be depended on the level of the uncertainty; hence a large variance will mean low knowledge of the system's parameters and vice versa.

The second problem is the fact that some states are not accessible and the ones that can be measured will be corrupted by the sensor noise. Again in order to include those noises into the total model, white noise Gaussian signals with zero mean will be used. These signals will be denoted as **V** and their number will depend upon number of the system's outputs. Once again the variance of these signals will describe the amount of noise that corrupts the outputs.

The state space model will now be called stochastic and will look like that of Fig. 1

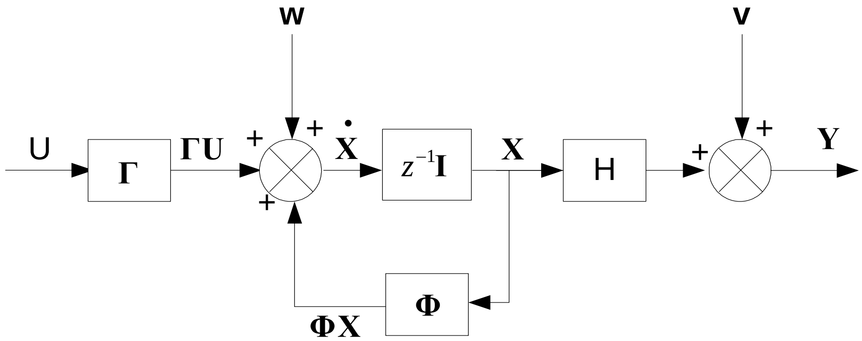


**Fig. 1** Stochastic state space model

The final problem is that some states are inaccessible. For example the rotor currents at a cage rotor IM cannot be measured easily and if they are needed then expensive probes will have to be inserted in the motor and hence all the characteristics that make the IM popular will vanish.

To overcome this problem estimators have to be used. But the well known deterministic estimation methods are designed for non-stochastic models. In the presence of noise they will not be able to produce the same satisfactory results as before. For this reason another estimator that can cope with stochastic models has to be applied. This estimator is the Kalman filter.

Before the full derivation of the KF some assumptions are needed. The noisy signals are white, with zero mean and normal distribution, totally uncorrelated. Also their autocorrelation for non-zero shift is zero, so:

$$\mu_{\mathrm{w}} = E\left[\mathbf{W}(k)\right] = 0 \qquad (1)$$

$$\mu_{\mathrm{v}} = E\left[\mathbf{V}(k)\right] = 0 \qquad (2)$$

$$\mathbf{Q}_{ww}(j) = E\left[\mathbf{W}(k)\mathbf{W}^{T}(k+j)\right] = 0, \text{ for any j} \neq 0 \qquad (3)$$

$$\mathbf{Q}_{vv}(j) = E\left[\mathbf{V}(k)\mathbf{V}^{T}(k+j)\right] = 0, \text{ for any j} \neq 0 \qquad (4)$$

$$\mathbf{Q}_{ww}(0) = E\left[\mathbf{W}(k)\mathbf{W}^T(k)\right] = \mathbf{Q}(k) \tag{5}$$

$$\mathbf{Q}_{vv}(0) = E\left[\mathbf{V}(k)\mathbf{V}^T(k)\right] = \mathbf{R}(k) \tag{6}$$

$$\mathbf{Q}_{wv}(j) = E\left[\mathbf{W}(k)\mathbf{V}^T(k+j)\right] = 0 \text{ for any } j \tag{7}$$

$$\mathbf{Q}_{vw}(j) = E\left[\mathbf{V}(k)\mathbf{W}^T(k+j)\right] = 0 \text{ for any } j \tag{8}$$

The stochastic state space model is given:

$$\mathbf{X}(k+1) = \mathbf{\Phi}(k)\mathbf{X}(k) + \mathbf{\Gamma}(k)\mathbf{U}(k) + \mathbf{W}(k) \tag{9}$$

$$\mathbf{Y}(k) = \mathbf{H}(k)\mathbf{X}(k) + \mathbf{V}(k) \tag{10}$$

The KF has a similar structure to the classical posteriori estimator:

$$\hat{\mathbf{X}}(k) = \tilde{\mathbf{X}}(k) + \mathbf{K}(k)\left[\mathbf{Y}(k) - \mathbf{H}(k)\tilde{\mathbf{X}}(k)\right]$$

where $\hat{\mathbf{X}}(k)$ is the "posteriori" estimation and $\tilde{\mathbf{X}}(k)$ is the "priori" or predicted estimation of the state.

$$\hat{\mathbf{X}}(k) = \left[1 - \mathbf{K}(k)\mathbf{H}(k)\right]\tilde{\mathbf{X}}(k) + \mathbf{K}(k)\mathbf{Y}(k) \tag{11}$$

Or: $\hat{\mathbf{X}}(k+1) = \left[1 - \mathbf{K}(k+1)\mathbf{H}(k+1)\right]\tilde{\mathbf{X}}(k+1) + \mathbf{K}(k+1)\mathbf{Y}(k+1)$ (12)

Where: $\tilde{\mathbf{X}}(k+1) = \mathbf{\Phi}(k)\hat{\mathbf{X}}(k) + \mathbf{\Gamma}(k)\mathbf{U}(k)$ (13)

Another notation, [90], is $\hat{\mathbf{X}}(k) = \hat{\mathbf{X}}(k/k)$ and $\tilde{\mathbf{X}}(k) = \hat{\mathbf{X}}(k/k-1)$, where $\hat{\mathbf{X}}(j/i)$ is the estimation of the vector state at the moment *j* by having information up to *i*.

From eqn. A.12 with the use of A.13:

$$\hat{\mathbf{X}}(k+1/k+1)=(1-\mathbf{K}(k+1)\mathbf{H}(k+1))\left(\mathbf{\Phi}(k)\hat{\mathbf{X}}(k/k)+\mathbf{\Gamma}(k)\mathbf{U}(k)\right)+$$
$$+\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{X}(k+1)+\mathbf{K}(k+1)\mathbf{V}(k+1) \qquad (14)$$

Notice that the gain matrix **K** is variable to have optimal estimation.

Since there is noise in the signal, it is assumed that $\hat{\mathbf{X}}(k/k)$ is the best estimation of $\mathbf{X}(k)$. The error between the actual state vector and the posteriori estimated is:

$$\overline{\mathbf{X}}(k+1/k+1)=\mathbf{X}(k+1)-\hat{\mathbf{X}}(k+1/k+1) \qquad (15)$$

By using eqns. 9 and 14:

$$\overline{\mathbf{X}}(k+1/k+1)=\mathbf{\Phi}(k)\mathbf{X}(k)+\mathbf{\Gamma}(k)\mathbf{U}(k)+\mathbf{W}(k)-(1-\mathbf{K}(k+1)\mathbf{H}(k+1))\left(\mathbf{\Phi}(k)\hat{\mathbf{X}}(k/k)+\mathbf{\Gamma}(k)\mathbf{U}(k)\right)+$$
$$-\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{X}(k+1)-\mathbf{K}(k+1)\mathbf{V}(k+1)$$

Again by using eqn. 9 to eliminate the term $\mathbf{X}(k+1)$:

$$\overline{\mathbf{X}}(k+1/k+1)=\mathbf{\Phi}(k)\mathbf{X}(k)+\mathbf{\Gamma}(k)\mathbf{U}(k)+\mathbf{W}(k)-(1-\mathbf{K}(k+1)\mathbf{H}(k+1))\left(\mathbf{\Phi}(k)\hat{\mathbf{X}}(k/k)+\mathbf{\Gamma}(k)\mathbf{U}(k)\right)+$$
$$-\mathbf{K}(k+1)\mathbf{H}(k+1)(\mathbf{\Phi}(k)\mathbf{X}(k)+\mathbf{\Gamma}(k)\mathbf{U}(k)+\mathbf{W}(k))-\mathbf{K}(k+1)\mathbf{V}(k+1)$$

After the multiplications and by selecting the terms that contain $\mathbf{X}(k)$ and $\hat{\mathbf{X}}(k/k)$ the right hand side is:

$$(\mathbf{\Phi}(k)-\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{\Phi}(k))\mathbf{X}(k)-(\mathbf{\Phi}(k)-\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{\Phi}(k))\hat{\mathbf{X}}(k/k)+\mathbf{\Gamma}(k)\mathbf{U}(k)+$$
$$-(1-\mathbf{K}(k+1)\mathbf{H}(k+1))\mathbf{\Gamma}(k)\mathbf{U}(k)-\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{\Gamma}(k)\mathbf{U}(k)+$$
$$+\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1) \qquad (16)$$

The terms with $\mathbf{\Gamma}(k)\mathbf{U}(k)$ will be self-cancelled. Also since the error at the sample *k+1* is $\overline{\mathbf{X}}(k+1/k+1)=\mathbf{X}(k+1)-\hat{\mathbf{X}}(k+1/k+1)$ at $k$ is $\overline{\mathbf{X}}(k/k)=\mathbf{X}(k)-\hat{\mathbf{X}}(k/k)$

$$(\mathbf{\Phi}(k)-\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{\Phi}(k))\overline{\mathbf{X}}(k/k)+\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{H}(k+1)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1) \qquad (17)$$

By selecting now the terms that contain $\mathbf{W}(k)$ and $\mathbf{V}(k+1)$:

$$\left(\mathbf{I}-\mathbf{K}(k+1)\mathbf{H}(k+1)\right)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)+\left(\mathbf{I}-\mathbf{K}(k+1)\mathbf{H}\right)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1) \tag{18}$$

By substituting $\mathbf{F}(k+1)$ where $\left(\mathbf{I}-\mathbf{K}(k+1)\mathbf{H}(k+1)\right)$:

$$\overline{\mathbf{X}}(k+1/k+1)=\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)+\mathbf{F}(k+1)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1) \tag{19}$$

The covariance of the error is

$$E\left(\overline{\mathbf{X}}(k+1/k+1)\overline{\mathbf{X}}^T(k+1/k+1)\right)=E[\left(\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)+\mathbf{F}(k+1)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1)\right)$$
$$\left(\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)+\mathbf{F}(k+1)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1)\right)^T]$$

$$E[\left(\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)+\mathbf{F}(k+1)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1)\right)\times$$
$$\left(\left(\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\right)^T+\left(\mathbf{F}(k+1)\mathbf{W}(k)\right)^T-\left(\mathbf{K}(k+1)\mathbf{V}(k+1)\right)^T\right)] \tag{20}$$

$$E[\left(\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)+\mathbf{F}(k+1)\mathbf{W}(k)-\mathbf{K}(k+1)\mathbf{V}(k+1)\right)\times$$
$$\left(\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)+\mathbf{W}^T(k)\mathbf{F}^T(k+1)-\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right)] \tag{21}$$

By doing the multiplications:

$$E[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)+\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)-$$
$$\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)+\mathbf{F}(k+1)\mathbf{W}(k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)$$
$$+\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)-\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)+$$
$$-\mathbf{K}(k+1)\mathbf{V}(k+1)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)-\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{W}^T(k)\mathbf{F}^T(k)+$$
$$+\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)] \tag{22}$$

By breaking the eqn. 22 to its parts:

$$E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]+E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]-$$
$$+E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right]+E\left[\mathbf{F}(k+1)\mathbf{W}(k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]$$
$$+E\left[\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]-E\left[\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right]+$$
$$-E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]-$$
$$+E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]+E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right] \tag{23}$$

$$E\left(\overline{\mathbf{X}}(k+1/k+1)\overline{\mathbf{X}}^T(k+1/k+1)\right)=E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]+$$
$$+E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]+$$
$$-E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right]+$$
$$+E\left[\mathbf{F}(k+1)\mathbf{W}(k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]+E\left[\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]+$$
$$-E\left[\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right]-E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]+$$
$$-E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]+E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right] \quad (24)$$

So now the expectations of the above 9 terms must be found and calculated:

$E\left(\overline{\mathbf{X}}(k+1/k+1)\overline{\mathbf{X}}^T(k+1/k+1)\right)=\mathbf{P}(k+1/k+1)$, where $\mathbf{P}$ is the covariance matrix of the error.

$$E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]=$$
$$=\mathbf{F}(k+1)\mathbf{\Phi}(k)E\left[\overline{\mathbf{X}}(k/k)\overline{\mathbf{X}}^T(k/k)\right]\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)=$$
$$=\mathbf{F}(k+1)\mathbf{\Phi}(k)\mathbf{P}(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1) \quad (25)$$

$$+\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]\overset{\text{Not Correlated}}{=}0 \quad (26)$$

$$-E\left[\mathbf{F}(k+1)\mathbf{\Phi}(k)\overline{\mathbf{X}}(k/k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right]\overset{\text{Not Correlated}}{=}0 \quad (27)$$

$$E\left[\mathbf{F}(k+1)\mathbf{W}(k)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]\overset{\text{Not Correlated}}{=}0 \quad (28)$$

$$E\left[\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]=\mathbf{F}(k+1)E\left[\mathbf{W}(k)\mathbf{W}^T(k)\right]\mathbf{F}^T(k+1)\overset{(5)}{=}$$
$$\mathbf{F}(k+1)\mathbf{Q}(k)\mathbf{F}^T(k+1) \quad (29)$$

$$-E\left[\mathbf{F}(k+1)\mathbf{W}(k)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right]=-\mathbf{F}(k+1)E\left[\mathbf{W}(k)\mathbf{V}^T(k+1)\right]\mathbf{K}^T(k+1)\overset{(7)}{=}0 \quad (30)$$

$$-E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\overline{\mathbf{X}}^T(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1)\right]\overset{\text{Not Correlated}}{=}0 \quad (31)$$

$$-E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{W}^T(k)\mathbf{F}^T(k+1)\right]=-\mathbf{K}(k+1)E\left[\mathbf{V}(k+1)\mathbf{W}^T(k)\right]\mathbf{F}^T(k+1)\overset{(8)}{=}0 \quad (32)$$

$$E\left[\mathbf{K}(k+1)\mathbf{V}(k+1)\mathbf{V}^T(k+1)\mathbf{K}^T(k+1)\right] = \mathbf{K}(k+1)E\left[\mathbf{V}(k+1)\mathbf{V}^T(k+1)\right]\mathbf{K}^T(k+1) \overset{(6)}{=}$$
$$\mathbf{K}(k+1)\mathbf{R}(k+1)\mathbf{K}^T(k+1)$$

So the outcome is:

$$\mathbf{P}(k+1/k+1) = \mathbf{F}(k+1)\mathbf{\Phi}(k)\mathbf{P}(k/k)\mathbf{\Phi}^T(k)\mathbf{F}^T(k+1) + \mathbf{F}(k+1)\mathbf{Q}(k)\mathbf{F}^T(k+1) +$$
$$+\mathbf{K}(k+1)\mathbf{R}(k+1)\mathbf{K}^T(k+1) =$$

$$\mathbf{P}(k+1/k+1) = \mathbf{F}(k+1)\left(\mathbf{\Phi}(k)\mathbf{P}(k/k)\mathbf{\Phi}^T(k) + \mathbf{Q}(k)\right)\mathbf{F}^T(k+1) + \mathbf{K}(k+1)\mathbf{R}(k+1)\mathbf{K}^T(k+1) \quad (33)$$

By defining as

$$\mathbf{P}^*(k) = \mathbf{P}(k+1/k) =$$

$$= \mathbf{E}\left(\left(\mathbf{X}(k) - \mathbf{X}(k+1/k)\right)\left(\mathbf{X}(k) - \mathbf{X}(k+1/k)\right)^T\right) = \mathbf{\Phi}(k)\mathbf{P}(k/k)\mathbf{\Phi}^T(k) + \mathbf{Q}(k) \quad (34)$$

$$\mathbf{P}(k+1/k+1) = \mathbf{F}(k+1)\mathbf{P}(k+1/k)\mathbf{F}^T(k+1) + \mathbf{K}(k+1)\mathbf{R}(k+1)\mathbf{K}^T(k+1) \quad (35)$$

By substituting **F**($k$+1) again:

$$\mathbf{P}(k+1/k+1) = \left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)\right)\mathbf{P}(k+1/k)\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)\right)^T +$$

$$+\mathbf{K}(k+1)\mathbf{R}(k+1)\mathbf{K}^T(k+1) \quad (36)$$

The goal is now to minimise **P**($k$+1/$k$+1) by the optimal choice of **K**($k$+1). To do the above Riccatti Differential Equation (**DE**) must be solved. (Eqn. 36 is also called Joseph form of the covariance update equation). This is a very difficult task and requires numerical methods; hence the answer is given without any proof:

$$\mathbf{K}(k+1) = \mathbf{P}(k+1/k)\mathbf{H}^T(k+1)\left[\mathbf{H}(k+1)\mathbf{P}(k+1/k)\mathbf{H}^T(k+1) + \mathbf{R}(k+1)\right]^{-1} \quad (37)$$

and: $\mathbf{P}(k+1/k+1) = \left[\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)\right]\mathbf{P}(k+1/k)$ (38)

To summarise, the next equations form a recursive set to implement the KF:

$$\mathbf{P}(k+1/k)=\mathbf{\Phi}(k)\mathbf{P}(k/k)\mathbf{\Phi}^T(k)+\mathbf{Q}(k) \tag{39}$$

$$\tilde{\mathbf{X}}(k+1)=\hat{\mathbf{X}}(k+1/k)=\mathbf{\Phi}(k)\hat{\mathbf{X}}(k)+\mathbf{\Gamma}(k)\mathbf{U}(k)=\mathbf{\Phi}(k)\hat{\mathbf{X}}(k/k)+\mathbf{\Gamma}(k)\mathbf{U}(k) \tag{40}$$

$$\mathbf{K}(k+1)=\mathbf{P}(k+1/k)\mathbf{H}^T(k+1)\left[\mathbf{H}(k+1)\mathbf{P}(k+1/k)\mathbf{H}^T(k+1)+\mathbf{R}(k+1)\right]^{-1} \tag{41}$$

$$\hat{\mathbf{X}}(k+1)=\hat{\mathbf{X}}(k+1/k+1)=(1-\mathbf{K}(k+1)\mathbf{H}(k+1))\hat{\mathbf{X}}(k+1/k)+\mathbf{K}(k+1)\mathbf{Y}(k+1) \tag{42}$$

$$\mathbf{P}(k+1/k+1)=[\mathbf{I}-\mathbf{K}(k+1)\mathbf{H}(k+1)]\mathbf{P}(k+1/k) \tag{43}$$

If there was coupling between the noisy signals that represent the parameters sensitivity then a new matrix **G** must be introduced, Fig. 2.

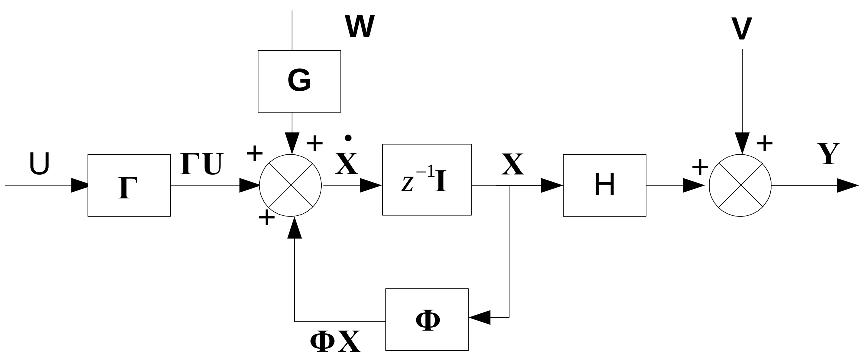


**Fig. 2** Stochastic state space model with coupled parameter uncertainty noise

And the above equations would be exactly the same except the first one:

$$\mathbf{P}(k+1/k)=\mathbf{\Phi}(k)\mathbf{P}(k/k)\mathbf{\Phi}^T(k)+\mathbf{G}(k)\mathbf{Q}(k)\mathbf{G}^T(k) \tag{44}$$

By using all the above equations the block diagram of the KF is shown in Fig. 3.

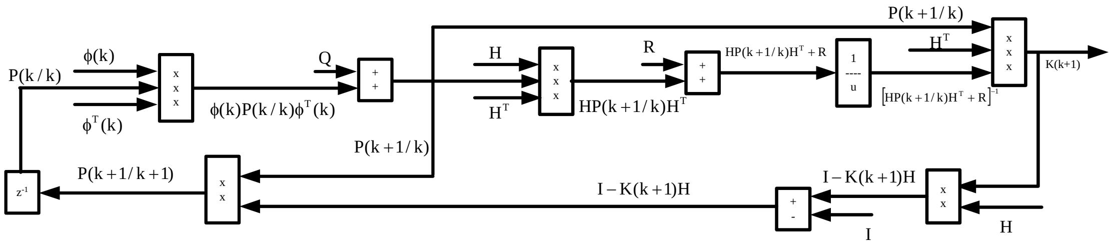


**Fig. 3** Block diagram of the discrete time Kalman filter

## 2 Intuitive Approach to the KF

The Kalman Filter works exactly as a deterministic posteriori estimator. The only difference is that the goal is not to achieve some error dynamics criteria but to minimise the error covariance matrix. Hence it can be said that it consists of two parts. In the first part the prediction of the new state and the new error covariance matrix are taking place, eqns. 39 & 40 respectively. Also at this stage the new KF gain matrix is calculated from the previous predicted values eqn. 41. Here the estimator is using the previous posteriori estimation of the state and error covariance matrix.

In the second part the KF corrects the previous measurements. First it corrects the previous prediction of the states by using current information from the output. The correction is accomplished by adding to the previous state prediction a weighted difference between the actual and the predicted output. The weight is nothing more than the previous KF gains, eqns. 42. At this stage the filter is also correcting the error covariance matrix, eqn. A.43, again by using the KF gain.

As can be seen from the previous equations, to calculate the KF gains the state space model of the system and the characteristics of the noise signals are needed. Also it is assumed that the noise signals are represented completely by their mean values (zero in this case) and covariances. If the characteristic of those signals change then the KF will not know this and will produce wrong results. Thus the KF must be provided with the accurate **Q** and **R**. Also from eqn. 42 can be seen that the error is influenced by the matrix **H**. Fortunately this is not a big problem since the output matrix is usually the identity matrix. The uncertainty of **A** and **B** (or **Φ** and **Γ**) can be described by **Q** and hence they do not need to be very accurate.

The covariance matrix **Q** represents the uncertainty of the model of the system. So if **Q** decreases at some point then the KF will assume that the model of the system is more accurate and hence will focus more on the stage of predicting and not at the stage of correcting. Therefore the KF gains will decrease. Theoretically if **Q**=0 then the matrix gains will converge to zero. Obviously this is will produce poor results since the estimator tends towards an open loop.

On the other hand if **Q** increases then the KF will assume that the model of the system is very inaccurate and hence will increase its gains. The same results will appear if **R** is increased in the first case or **R** is decreased in the second. The problem that appears here if **R**=0 is that the calculation of eqn. 41 might not be possible (even numerically) since it is not guaranteed that the matrix that is inversed is not singular. On the other hand if **R** is very big then the numerical solution of the equation will diverge.

Since the noise signals **W** and **V** are totally uncorrelated the error covariance matrix must be diagonal. Big values of the elements of the main diagonal of **P** indicate large error in the estimation process.

Furthermore the initial estimation of error must be such to indicate big error; hence the elements in **P** must be very big. Also **P** must be positive semi - definite since its elements represent the square of the variances, which they must be zero or positive.

Finally if the system is LTI then gains of the KF will converge very fast to a specific value. Then the so-called Steady State Kalman Filter **(SSKF)** can be used where the gains have been pre-calculated.