Chapter 6

Appendix

6.1 Appendix-1

Derivation of Bayes Filter using Bayes theorem and Markov's Assumptions:

$$P(B|A_1) = \frac{P(A_1 \cap B)}{P(A_1)} \Rightarrow \quad P(A_1 \cap B) = P(B|A_1) * P(A_1)$$

similarly

$$P(B|A_2) = \frac{P(A_2 \cap B)}{P(A_2)} \Rightarrow P(A_2 \cap B) = P(B|A_2) * P(A_2)$$

Thus finally we see

$$\Rightarrow \qquad P(B) = \sum P(A_i) * P(A_i \cap B)$$

$$\Rightarrow \qquad P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(B|A_1) * P(A_1)}{\sum P(A_i) * P(A_i \cap B)}$$

Probability of area shown in figure 6.1 can be represented as

$$P(x) = \int P(x, y)dy \text{ ; or } = \sum P(x_i, y)$$

but we know $P(x, y) = P(x|y).P(y)$

$$\Rightarrow \qquad P(x) = \int P(x|y).P(y) \quad -1$$

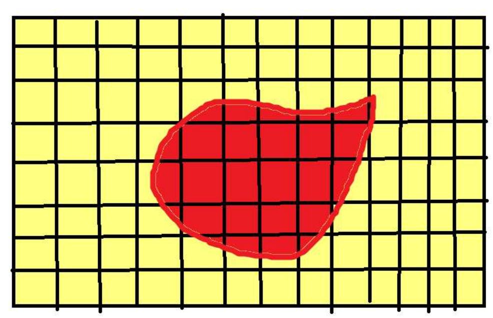

Figure-6.1 y spread in x domain

On extending the result -

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

$$P(A|B,C) = \frac{P(B|A,C) * P(A|C)}{P(B|C)} \text{ , } P(B|C) \text{ is usually } \eta - 2$$

$$P(A,B) = \sum P(A,B,C_i) = \sum P(A|B,C_i) * P(B,C_i)$$

$$\Rightarrow \qquad = \sum P(A|B,C_i) * P(C_i|B) * P(B) - 3$$

Thus finally

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{\sum P(A|B,C_i) * P(C_i|B) * P(B)}{P(B)}$$

$$\Rightarrow \qquad = \sum P(A|B,C_i) * P(C_i|B) - 4$$

Now on combining the above result and Markov Assumption we will get-

Bel- denotes belief of being at

$$Bel(X_t) = P(X_t|U_1, Z_1, ..., U_t, Z_t)$$

we may assume X_t as A and $U_1, Z_1, ...,$ as B and U_t, Z_t as C

then apply eqn. 2

$$Bel(X_t) = \eta P(Z_t|X_t, U_1, Z_1, ..., U_t) * P(X_t|U_1, Z_1, ..., U_t)$$

then apply eqn. 5

$$Bel(X_t) = \eta P(Z_t|X_t) * P(X_t|U_1, Z_1, ..., U_t)$$

then apply eqn. 4

$$Bel(X_t) = \eta P(Z_t|X_t) * \int P(X_t|U_1, Z_1, ..., U_t, X_{t-1}) * P(X_{t-1}|U_1, Z_1, ..., U_t) dX_{t-1}$$

then apply eqn. 6

$$Bel(X_t) = \eta P(Z_t|X_t) * \int P(X_t|X_{t-1}, U_t) * P(X_{t-1}|U_1, Z_1, ..., U_t) dX_{t-1}$$

$$Bel(X_t) = \eta P(Z_t|X_t) * \int P(X_t|X_{t-1}, U_t) * P(X_{t-1}|U_1, Z_1, ..., U_{t-1}, Z_{t-1}) dX_{t-1}$$

thus we see Be(Xt-1) as second term in integration hence reaching t-1 state shows recursion

Thus we will obtain the final equation of as

$$Bel(X_t) = \eta P(Z_t|X_t) * \int P(X_t|X_{t-1}, U_t) * Bel(X_{t-1})\, dX_{t-1}$$

6.2 Appendix-2

The linear state may be represented as-

$$x_{k+1} = A_k x_k + B_k u_k + G_k v_k$$

$$z_k = H_k x_k + w_k$$

$\rightarrow$ Relation between two states x_{k-1}, x_k is linear

$\rightarrow$ Gaussian distribution

$\rightarrow$ for one dimensional $1-D \rightarrow \quad \mu\,(mean) \quad ; \quad \sigma^2\,(is\,variance)$

$\rightarrow$ for N dimensional $N-D \rightarrow \quad \vec{\mu}\,[matrix]\,(mean) \quad ; \quad \sum(is \quad covariance \quad matrix)$

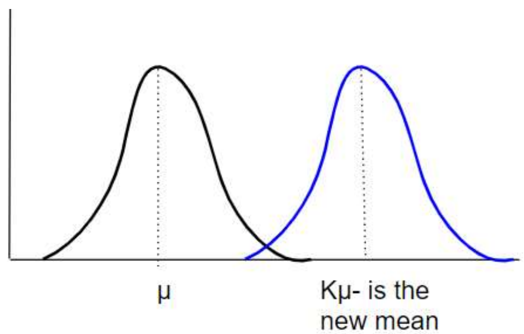

Figure-6.2- state multiplied by a scalar K

$$\mu' = K\mu \quad ; \quad \sigma'^2 = K^2\sigma^2$$

while for N-Dimensional gaussian multiplied by F matrix

$$\vec{\mu}' = F\vec{\mu} \quad ; \quad \sum' = F \sum F^T$$

Now on modeling the belief as a Gaussian distribution-

→ belief is modeled as Gaussian

$$N(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\Pi}} exp(\frac{-(x-\mu)^2}{2\sigma^2})$$

→ multiplication of two beliefs for a state :
$N(x,\mu_0,\sigma_0) * N(x,\mu_1,\sigma_1) = N(x,\mu',\sigma')$

$$\mu' = \mu_0 + \frac{\sigma_0^2(\mu_1-\mu_0)}{\sigma_0^2+\sigma_1^2}$$

$$\sigma'^2 = \sigma_0^2 - \frac{\sigma_0^4}{\sigma_0^2+\sigma_1^2}$$

$$if \quad K = \frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2} \quad then \quad \mu' = \mu_0 + K(\mu_1-\mu_0) \quad ;$$

$$and \quad \sigma'^2 = \sigma_0^2 - K\sigma_0^2$$

→ K is called as gain

⇒ if $\sigma_0 \to \infty$; $\mu' \to \mu_1$ this means higher uncertainity in μ_0

⇒ if $\sigma_1 \to \infty$; $\mu' \to \mu_0$ this means higher uncertainity in μ_1

→ for N-Dimensional Gaussian

$$K = \sum_0 (\sum_0 + \sum_1)^{-1} \quad \text{-A}$$

$$\vec{\mu'} = \vec{\mu_0} + K(\vec{\mu_1} - \vec{\mu_0}) \quad \text{-B}$$

$$\sum' = \sum_0 - K \sum_0 \quad \text{-C}$$

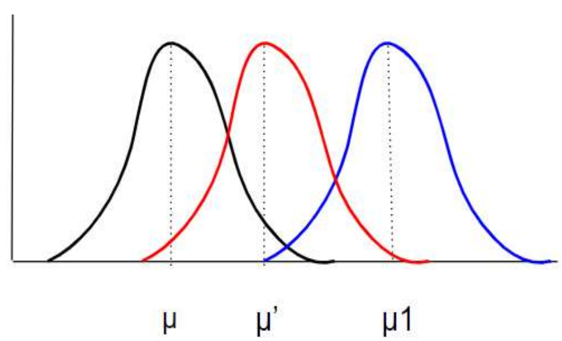

Figure-6.3-represents gaussian multiplication

6.2.1 Prediction Step

if we are to consider acceleration then the equation becomes

$$X_k = X_{k-1} + \triangle t + \frac{1}{2} a \triangle t^2$$

$$\widehat{x_k} = F_k \widehat{x_{k-1}} + \begin{bmatrix} \frac{\triangle t^2}{2} \\ \triangle t \end{bmatrix} a$$

$$\Rightarrow \widehat{x_k} = F_k \widehat{x_{k-1}} + B_k \vec{U}_k$$

while the new covariance matrix becomes :$P_k = F_k P_{k-1} F_k^T + Q_k$

B_k - is the control matrix

$\vec{U}_k$ - is the control vector or command input

Q_k - is the Process noise matrix

6.2.2 Updation Step

$$\vec{\mu_0} = H_k X_k \quad ; \quad \sum_0 = H_k P_k H_k^T$$

Sensor data : $Z_k \quad ; \quad R_k \qquad ; \sum_k = H_k P_k {H_k}^T + R_k$

$$(\mu_0, \sum_0) = (H_k \widehat{x_k}, H_k P_k H_k^T)$$

$$(\mu_1, \sum_1) = (Z_k, R_k)$$

$$H_k \widehat{x_k'} = H_k \widehat{x_k} + K(\vec{Z_k} - H_k \widehat{x_k})$$

$$H_k {P_k}' {H_k}^T = H_k P_k {H_k}^T - K(H_k P_k {H_k}^T)$$

$$K = H_k P_k {H_k}^T (H_k P_k {H_k}^T - R_k)^{-1}$$

'Thus the final optimum equation obtained is-

$$\widehat{x_k'} = \widehat{x_k} + K'(Z_k - H_k \widehat{x_k})$$

$$P_k' = P_k - K' H_k P_k$$

$$K' = P_k H^T (H_k P_k {H_k}^T + R_k)^{-1}$$

6.3 Appendix-3

Taking a one-dimensional random variable x having a gaussian distribution.

- Mean $\bar{x}$ (μ)and variance σ^2 characterize the equation as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- Three sample points called as Sigma Points are selected

$$\tilde{x}^0 = \bar{x} \qquad W_0 = \frac{\kappa}{1+\kappa}$$

$$\tilde{x}^1 = \bar{x} + \sqrt{1+\kappa} \cdot \sigma$$

$$\tilde{x}^2 = \bar{x} - \sqrt{1+\kappa} \cdot \sigma \qquad W_1 = W_2 = \frac{1}{2(1+\kappa)}$$

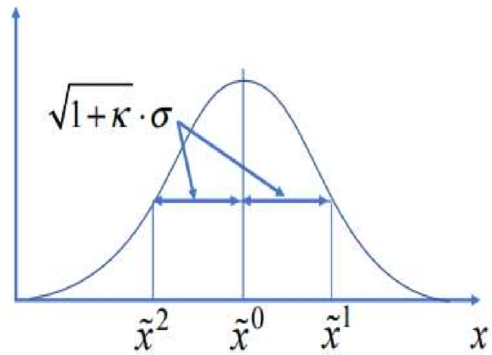

Figure 6.4 - One dimensional gaussian distribution showing sigma points. [5]

- Here κ is a parameter of sigma points to be tuned, and Wi is the weight of the i th sigma point used for computing mean and variance
- The weighted mean of the three Sigma points agrees with the true mean of the

$$\sum_{i=0}^{2} W_i \tilde{x}^i = \frac{\kappa}{1+\kappa}\bar{x} + \frac{1}{2(1+\kappa)}\left\{(\bar{x} + \sqrt{1+\kappa} \cdot \sigma) + (\bar{x} - \sqrt{1+\kappa} \cdot \sigma)\right\}$$

$$= \frac{\kappa}{1+\kappa}\bar{x} + \frac{2}{2(1+\kappa)}\bar{x} = \bar{x}$$

- The weighted variance of the three Sigma points agrees with the true variance

$$\sum_{i=0}^{2} W_i (\tilde{x}^i - \bar{x})^2 = \frac{\kappa}{1+\kappa}(\bar{x} - \bar{x}) + \frac{1}{2(1+\kappa)}\left\{(\bar{x} + \sqrt{1+\kappa} \cdot \sigma - \bar{x})^2 + (\bar{x} - \sqrt{1+\kappa} \cdot \sigma - \bar{x})^2\right\}$$

$$= \frac{2}{2(1+\kappa)}(\sqrt{1+\kappa} \cdot \sigma)^2 = \sigma^2$$

6.3.1 Unscented Transform

- Consider a non-linear transformation of x to y by g(x).
- Here the function is

$$y = g(x) = \sum_{k=0}^{\infty} \frac{g^{(k)}(\overline{x})}{k!}(x-\overline{x})^k$$

- The distribution of y is no longer Gaussian, but its mean E[y] and variance E[(y-E[y])2] can characterize the distribution

$$\overline{y}_{sample} \triangleq \sum_{i=0}^{2} W_i \tilde{y}^i = E[y]$$

$$\sigma_{sample}^2 \triangleq \sum_{i=0}^{2} W_i (\tilde{y}^i - \overline{y}_{sample})^2 = E[(y-E[y])^2]$$

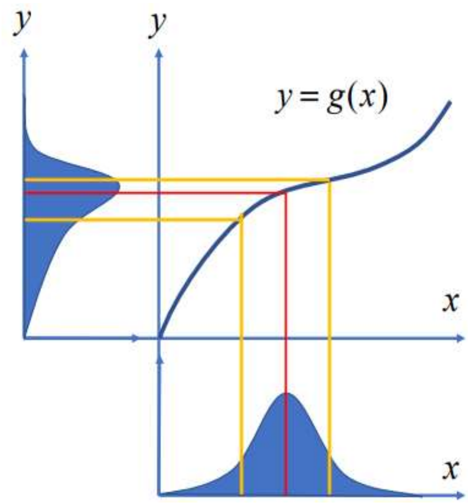

Figure 6.5 -Graph showing sigma points . [5]

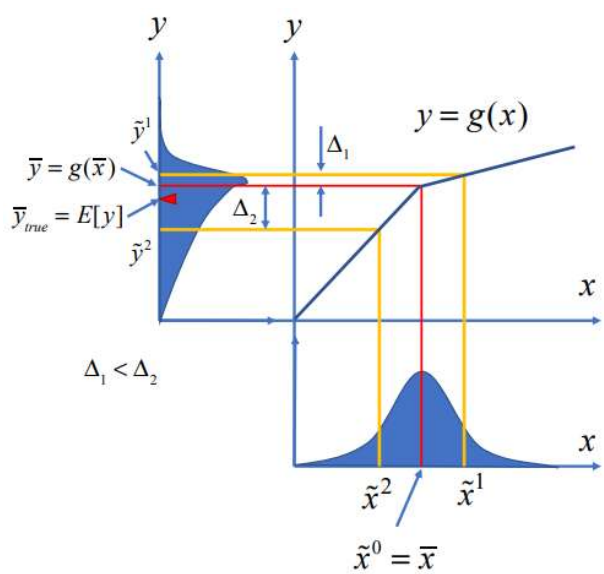

Figure 6.6 -Graph showing sigma points and their propagation . [5]

Thus we see that weighted mean of sigma points gives better mean value.

$$\overline{y}_{sample} \triangleq \sum_{i=0}^{2} W_i \tilde{y}^i = \frac{\kappa}{1+\kappa}\tilde{y}^0 + \frac{1}{2(1+\kappa)}(\tilde{y}^1 + \tilde{y}^2)$$

$$= \frac{\kappa}{1+\kappa}\overline{y} + \frac{1}{2(1+\kappa)}(\overline{y} + \Delta_1 + \overline{y} - \Delta_2) = \overline{y} + \frac{1}{2(1+\kappa)}(\Delta_1 - \Delta_2) < \overline{y}$$

6.3.2 Unscented Transform in a multidimensional Gaussian distribution

- For an n-dimensional Gaussian distribution, we use (2n + 1) Sigma points

- The covariance matrix Px is real, symmetric, and positive-definite. Therefore, it can be diagonalized

$$P_X = VDV^T$$

$$\mathbf{V} = (\mathbf{v}_1 \cdots \mathbf{v}_n), \mathbf{D} = diag(\sigma_1^2 \cdots \sigma_n^2)$$

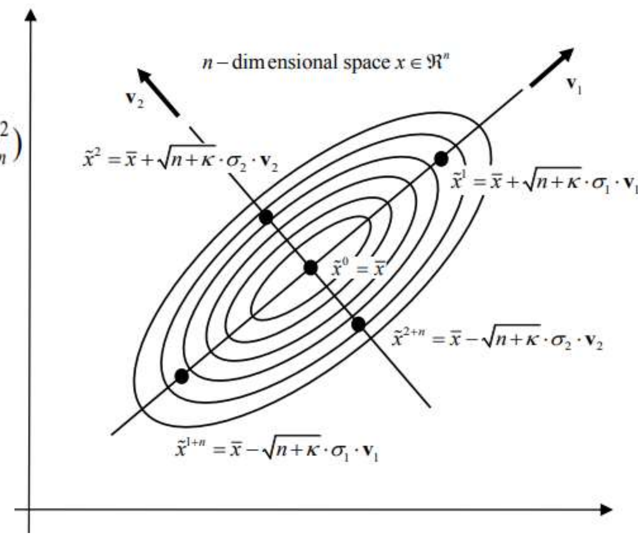

Figure 6.7- N dimensional distribution [5]

- Thus all points propagate through the function

- The weighted mean is $\bar{y}_{sample} = \sum_{i=0}^{2n} W_i \, \tilde{y}^i$
- While weighted covariance is

$$P_{y,sample} = \sum_{i=0}^{2n} W_i (\tilde{y}^i - \bar{y}_{sample})(\tilde{y}^i - \bar{y}_{sample})^T$$

- weighted mean can approximate the true mean to the third order, and the weighted covariance to the second order.

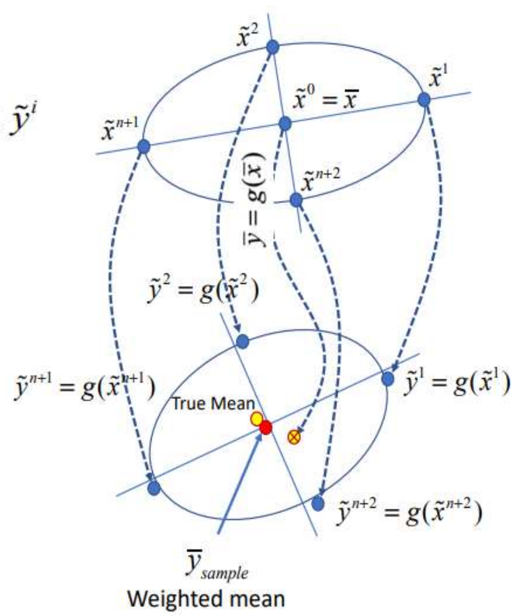

Figure 6.8 -Propagation via function. [5]

6.3.3 State Space and Prediction

- Let us consider a dynamic system represented as

$$u_t = 0$$

$$x_{t+1} = f(x_t, \cancel{u_t}, t) + w_t$$

$$y_t = h(x_t, t) + v_t$$

- Process noise and measurement noise are zero-mean, uncorrelated (white) noise with covariance Qt and Rt $\quad w_t \sim N(0, Q_t), \quad v_t \sim N(0, R_t)$
- At time t-1 the state is x_{t-1} and the covariance is given by

$$P_{t-1} = E[(\hat{x}_{t-1} - x_{t-1})(\hat{x}_{t-1} - x_{t-1})^T]$$

- Then Using the eigen values and eigen vector of state Pt-1 is found along with the sigma points

$$\tilde{x}^0_{t-1} = \hat{x}_{t-1} : \quad \bar{x}$$

$$\tilde{x}^i_{t-1} = \hat{x}_{t-1} + \sqrt{n+\kappa} \cdot \sigma_i \cdot \mathbf{v}_i, \quad \tilde{x}^{i+n}_{t-1} = \hat{x}_{t-1} - \sqrt{n+\kappa} \cdot \sigma_i \cdot \mathbf{v}_i$$

$$i = 1, \cdots, n$$

- Sigma points are propagated through the state equation

$$\tilde{x}^{i*}_{t|t-1} = f(\tilde{x}^i_{t-1}, t-1) + \cancelto{0}{w_{t-1}}, \quad i = 0, \cdots, 2n$$

- The weighted mean is given as

$$\hat{x}_{t|t-1,sample} = \sum_{i=0}^{2n} W_i \hat{x}^{i*}_{t|t-1}$$

- The covariance is propagated as

$$P_{t|t-1,sample} = \sum_{i=0}^{i=2n} W_i (\tilde{x}^i_{t|t-1} - \hat{x}_{t|t-1,sample})(\tilde{x}^i_{t|t-1} - \hat{x}_{t|t-1,sample})^T + Q_{t-1}$$

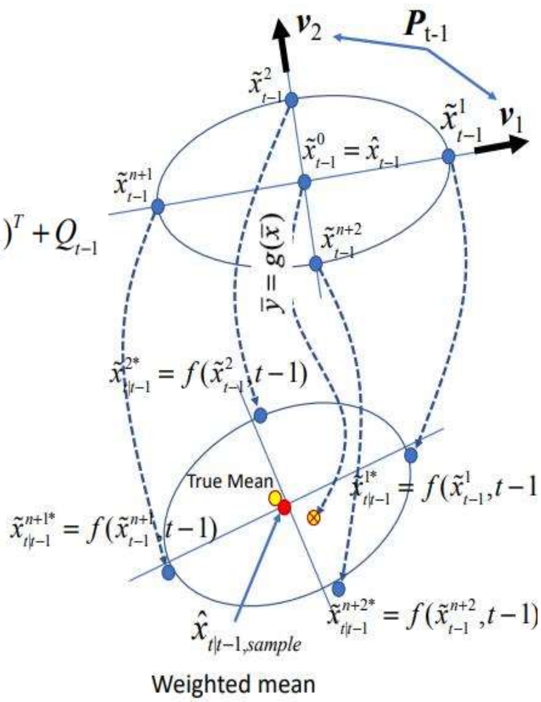

Figure 6.9 -Propagation of State. [5]

6.3.4 State Update

- State and covariance can be updated by the Unscented Transform on $P_{t|t-1}$
- We first obtain the innovation covariance by examining the distribution of the output created through the measurement equation
- We compute the eigenvalues and eigen-vectors of the a priori covariance,
- Then, we generate (2n+1) Sigma points, and estimate the mean and covariance of the distribution of output y with the Sigma points
- Using the deterministic part of the measurement function, the Sigma points are mapped

$$\tilde{y}_t^i = h(\tilde{x}_{t|t-1}^i, t), \quad i = 0, \cdots, 2n$$

- The weighted mean of the Sigma points is given by

$$\hat{y}_{t,sample} = \sum_{i=0}^{2n} W_i \tilde{y}_t^i$$

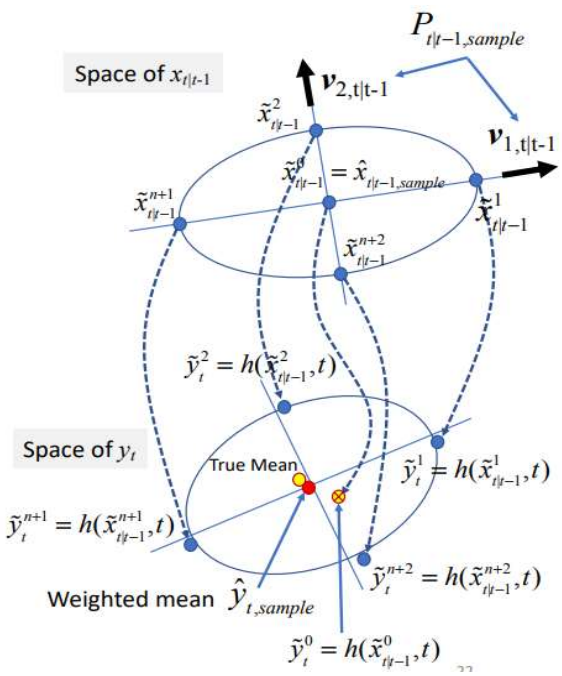

Figure 6.10 -Updation of State. [5]

- The next step is computing the first term and cross variance

$$P_{xy} = \sum_{i=0}^{2n} W_i (\tilde{x}_{t|t-1}^i - \hat{x}_{t|t-1,sample})(\tilde{y}_t^i - \hat{y}_{t,sample})^T \quad P_y = \sum_{i=0}^{2n} W_i (\tilde{y}_t^i - \hat{y}_{t,sample})(\tilde{y}_t^i - \hat{y}_{t,sample})^T + R_t$$

- Kalman gain K can be computed by $K_t = P_{xy} P_y^{-1}$

- State update given as $\hat{x}_t = \hat{x}_{t|t-1,sample} + K_t [y_t - \hat{y}_{t,sample}]$

- Finally new covariance is given by $P_t \cong P_{t|t-1,sample} - K_t P_y K_t^T$