Preliminaries:

$$N(x|m,P) = \frac{1}{(2\pi)^{\frac{N}{2}}|P|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(x-m)^T P^{-1}(x-m)\right\}$$

$$P(x) = N(x|m,P) \quad , \quad m_i = E[x_i] \, , \; P_{ij} = Cov(x_i, x_j)$$

$$P(y|x) = N(y|Hx,R) \quad , \quad E[y_i|x] = (Hx)_i = \sum_k H_{ik} x_k$$

$\Rightarrow E[x] = m \, , \; Cov(x) = P \, , \; E[y] = ? \, , \; Cov(x,y) = ? \, , \; Cov(y) = ?$

$$
\begin{aligned}
E[y_i] &= \int y_i P(y_i)\, dy_i = \iint y_i P(y_i|x) P(x)\, dy_i\, dx \\
&= \int E[y_i|x] P(x)\, dx \\
&= \int \sum_k H_{ik} x_k N(x|m,P)\, dx \\
&= \sum_k H_{ik} \int x_k N(x|m,P)\, dx \\
&= \sum_k H_{ik} E[x_k] \\
&= \sum_k H_{ik} m_k \quad \Rightarrow E[y] = Hm
\end{aligned}
$$

$$
\begin{aligned}
Cov(x_i, y_j) &= E\Big[(x_i - m_i)(y_j - \sum_k H_{jk} m_k)\Big] \\
&= E[x_i y_j] - m_i \sum_k H_{jk} m_k \\
&= \iint x_i y_j P_{x_i y_j}(x_i, y_j)\, dx_i\, dy_j - m_i \sum_k H_{jk} m_k \\
&= \iint x_i y_j N(x|m,P) N(y|Hx,R)\, dx\, dy - m_i \sum_k H_{jk} m_k \\
&= \int x_i N(x|m,P) \int y_j N(y|Hx,R)\, dy\, dx - m_i \sum_k H_{jk} m_k \\
&= \int x_i N(x|m,P) \sum_k H_{jk} x_k\, dx - m_i \sum_k H_{jk} m_k \\
&= \sum_k H_{jk} \int x_i x_k N(x|m,P)\, dx - m_i \sum_k H_{jk} m_k \\
&= \sum_k H_{jk} E[x_i x_k] - m_i \sum_k H_{jk} m_k \\
&= \sum_k H_{jk} \left(E[x_i x_k] - m_i m_k \right) \\
&= \sum_k H_{jk} Cov(x_i, x_k) = \sum_k H_{jk} P_{ki} = (HP)_{ji} \quad \Rightarrow Cov(x,y) = P^T H^T = PH^T \quad : (N, M)
\end{aligned}
$$

$x : N \times 1$

$y : M \times 1$

Similarly, $Cov(y,x) = HP$

$$
\begin{aligned}
Cov(y_i, y_j) &= \iint y_i y_j P_{y_i y_j}(y_i, y_j)\, dy_i\, dy_j - \sum_k H_{ik} m_k \sum_k H_{jk} m_k \\
&= \iint y_i y_j N(x|m,P) N(y|Hx,R)\, dx\, dy - \sum_k H_{ik} m_k \sum_k H_{jk} m_k \\
&= \int N(x|m,P) \int y_i y_j N(y|Hx,R)\, dy\, dx - \sum_k H_{ik} m_k \sum_k H_{jk} m_k \\
&= \int N(x|m,P) \Big(R_{ij} + \sum_k H_{ik} x_k \sum_k H_{jk} x_k \Big) dx - \sum_k H_{ik} m_k \sum_k H_{jk} m_k \\
&= R_{ij} + \sum_{k_1} \sum_{k_2} H_{ik_1} H_{jk_2} \int N(x|m,P)\, x_{k_1} x_{k_2}\, dx - \sum_k H_{ik} m_k \sum_k H_{jk} m_k
\end{aligned}
$$

$$= R_{ij} + \sum_{k_1}\sum_{k_2} H_{ik_1} H_{jk_2}\left(P_{k_1k_2} + \cancel{m_{k_1} m_{k_2}}\right) - \cancel{\sum_k H_{ik} m_k \sum_k H_{jk} m_k}$$

$$= R_{ij} + \sum_{k_1}\sum_{k_2} H_{ik_1} H_{jk_2} P_{k_1} P_{k_2} \Rightarrow Cov(y, y) = R + HPH^T$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N\left(\begin{bmatrix} m \\ Hm \end{bmatrix}, \begin{bmatrix} P & PH^T \\ HP & HPH^T + R \end{bmatrix} \right)$$

$x \sim N(m, P)$

$y \sim N(Hm, HPH^T + R)$

If r.v.'s x, y have the joint probability density

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N\left(\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} A & C \\ C^T & B \end{bmatrix} \right)$$

$\Rightarrow x \sim N(a, A)$, $y \sim N(b, B)$, $x|y \sim ?$, $y|x \sim ?$

Let $z = x - CB^{-1}y = x + My$

$$Cov(z, y) = Cov(x, y) - CB^{-1} Var(y)$$

$= C - CB^{-1}B = 0 \Rightarrow$ z and y are independent

$$E[x|y] = E[z - My | y]$$

$$= E[z|y] - M\, E[y|y]$$

$$= E[z] + My$$

$$= E[x] + M\, E[y] - My$$

$$= a + M(b - y)$$

$$= a + CB^{-1}(y - b)$$

$$Var(x|y) = Var(z - My | y)$$

$$= Var(z|y) + \underbrace{Var(My|y)}_{=0} - \underbrace{2Cov(z, My|y)}_{=0}$$

$$= Var(z|y)$$

$= Var(z)$ (get rid of "given y")

$$= Var(x + My)$$

$$= Var(x) + M\,Var(y) M^T + Cov(x, y) M^T + M\,Cov(y, x)$$

$$= A + CB^{-1}BB^{-1}C^T - CB^{-1}C^T - CB^{-1}C^T$$

$$= A - CB^{-1}C^T$$

$\Rightarrow$ $x|y \sim N(a + CB^{-1}(y - b), A - CB^{-1}C^T)$

$Var(x+y)_{ij}$

$$= E[(x_i + y_i - \bar{x}_i - \bar{y}_i)(x_j + y_j - \bar{x}_j - \bar{y}_j)]$$

$$= E[(x_i - \bar{x}_i)(x_j - \bar{x}_j)] + E[(y_i - \bar{y}_i)(y_j - \bar{y}_j)] + E[(x_i - \bar{x}_i)(y_j - \bar{y}_j)] + E[(x_j - \bar{x}_j)(y_i - \bar{y}_i)]$$

$$= Var(x)_{ij} + Var(y)_{ij} + Cov(x, y)_{ij} + Cov(x, y)_{ji}$$

$$\Rightarrow Var(x+y) = Var(x) + Var(y) + Cov(x, y) + Cov(x, y)^T$$

$Cov(Ax, y)$. $A: m \times n$. $x: n \times 1$, $y: m \times 1$

$$Cov(Ax, y)_{ij} = Cov((Ax)_i, y_j)$$

$$= Cov\left(\sum_k A_{ik} x_k, y_j\right)$$

$$= \sum_k A_{ik}\, Cov(x_k, y_j)$$

$$\Rightarrow Cov(Ax, y) = A\, Cov(x, y)$$

$$Cov(y, Ax)_{ij} = Cov(y_i, (Ax)_j)$$

$$= Cov\left(y_i, \sum_k A_{jk} x_k\right)$$

$$= \sum_k A_{jk}\, Cov(y_i, x_k)$$

$$\Rightarrow Cov(y, Ax) = Cov(y, x) A^T$$

$y|x$. Similarly, find a z such that $z \perp\!\!\!\perp x$

$\Rightarrow z = y - C^T A^{-1} x = y + Mx$. $Cov(z,x) = Cov(y,x) - C^T A^{-1} Var(x) = C^T - C^T = 0$ ✓

$$\begin{aligned} E[y|x] &= E[z - Mx \mid x] \\ &= E[z] - Mx \\ &= b + M(a - x) \\ &= b + C^T A^{-1}(x - a) \end{aligned}$$

$$\begin{aligned} Var(y|x) &= Var(z - Mx \mid x) \\ &= Var(z|x) + \underbrace{Var(Mx|x)}_{=0} - 2\underbrace{Cov(z, Mx|x)}_{=0} \\ &= Var(z) \\ &= Var(y + Mx) \\ &= Var(y) + M\, Var(x) M^T + Cov(y, Mx) + Cov(Mx, y) \\ &= B + C^T \cancel{A^{-1}} \cancel{A} A^{-1} C - C^T A^{-1} C - C^T A^{-1} C \\ &= B - C^T A^{-1} C \end{aligned}$$

$\Rightarrow y|x \sim N(\, b + C^T A^{-1}(x-a),\ B - C^T A^{-1} C\,)$

$$\begin{aligned} &Var(Ax)_{ij} \\ &= Cov\left(\sum_{k_1} A_{ik_1} x_{k_1}, \sum_{k_2} A_{jk_2} x_{k_2}\right) \\ &= \sum_{k_1}\sum_{k_2} A_{ik_1} A_{jk_2} Cov(x_{k_1}, x_{k_2}) \\ &= \sum_{k_1} A_{ik_1} \sum_{k_2} Cov(x_{k_1}, x_{k_2}) A_{jk_2} \\ &= \sum_{k_1} A_{ik_1} \left(Var(x) A^T\right)_{k_1 j} \\ &= \left(A\, Var(x)\, A^T\right)_{ij} \\ &\Rightarrow Var(Ax) = A\, Var(x) A^T \end{aligned}$$

Kalman Filter : Prediction $\qquad x_k = F_k x_{k-1} + V_k$

$P(x_{k-1}|y_{k-1}) = N(m_{k-1}, P_{k-1})$, $\qquad x_k = F_k x_{k-1} + V_k \qquad z_k = H_k x_k + W_k$

$P(x_{k-1}|z_{k-1}) = N(m_{k-1}, P_{k-1})$, $P(x_k|x_{k-1}) = N(F_k x_{k-1}, Q_k)$, $P(z_k|x_k) = N(H_k x_k, R_k)$

$$\begin{bmatrix} x_{k-1} \\ x_k \end{bmatrix}_{|z_{k-1}} \sim N\left(\begin{bmatrix} m_{k-1} \\ F_k m_{k-1} \end{bmatrix}, \begin{bmatrix} P_{k-1} & P_{k-1} F_k^T \\ F_k P_{k-1} & Q_k + F_k P_{k-1} F_k^T \end{bmatrix} \right)$$

Derivation :

$$\begin{aligned} E[x_k|z_{k-1}] &= \int x_k\, P(x_k|z_{k-1})\, dx_k \\ &= \iint x_k\, P(x_k, x_{k-1}|z_{k-1})\, dx_k\, dx_{k-1} \\ &= \iint x_k\, P(x_k|x_{k-1}) P(x_{k-1}|z_{k-1})\, dx_k\, dx_{k-1} \\ &= \int P(x_{k-1}|z_{k-1}) \int x_k\, P(x_k|x_{k-1})\, dx_k\, dx_{k-1} \\ &= \int N(m_{k-1}, P_{k-1}) \int x_k\, N(F_k x_{k-1}, Q_k)\, dx_k\, dx_{k-1} \\ &= \int F_k x_{k-1}\, N(m_{k-1}, P_{k-1})\, dx_{k-1} \\ &= \underline{F_k m_{k-1}} \end{aligned}$$

$Cov(x_{k-1}, x_k|z_k)_{ij} = Cov(x_{k-1}^{(i)}, x_k^{(j)}|z_k) \qquad$ where $x_k = \begin{bmatrix} \vdots \\ x_{k-1}^{(i)} \\ \vdots \end{bmatrix} \leftarrow i = \begin{bmatrix} \vdots \\ (x_{k-1})_i \\ \vdots \end{bmatrix} \leftarrow i$

$$= E[x_{k-1}^{(i)} x_k^{(j)} | z_k] - (m_{k-1})_i \cdot (F_k m_{k-1})_j$$

$$= \iint x_{k-1}^{(i)} x_k^{(j)} \; P(x_{k-1}^{(i)} x_k^{(j)} | z_k) \, dx_{k-1}^{(i)} dx_k^{(j)} - (m_{k-1})_i \cdot (F_k m_{k-1})_j$$

$$= \iint x_{k-1}^{(i)} x_k^{(j)} N(x_{k-1}; m_{k-1}, P_{k-1}) N(x_k; F_k x_{k-1}, Q_k) \, dx_{k-1} \, dx_k - (m_{k-1})_i (F_k m_{k-1})_j$$

$$= \int x_{k-1}^{(i)} N(x_{k-1}; m_{k-1}, P_{k-1}) \int x_k^{(j)} N(x_k; F_k x_{k-1}, Q_k) \, dx_k \, dx_{k-1} - (m_{k-1})_i (F_k m_{k-1})_j$$

$$= \int x_{k-1}^{(i)} N(x_{k-1}; m_{k-1}, P_{k-1}) \, (F_k x_{k-1})_j \, dx_{k-1} - (m_{k-1})_i (F_k m_{k-1})_j$$

$$= \sum_\alpha (F_k)_{j\alpha} \int x_{k-1}^{(\alpha)} x_{k-1}^{(i)} N(x_{k-1}; m_{k-1}, P_{k-1}) \, dx_{k-1} - (m_{k-1})_i (F_k m_{k-1})_j$$

$$= \sum_\alpha (F_k)_{j\alpha} \left[(P_{k-1})_{\alpha i} + (m_{k-1})_\alpha (m_{k-1})_i \right] - (m_{k-1})_i (F_k m_{k-1})_j$$

$$= (F_k P_{k-1})^T_{ij} \Rightarrow Cov(x_{k-1}, x_k | z_{k-1}) = (F_k P_{k-1})^T = P_{k-1}^T F_k^T = \underline{P_{k-1} F_k^T}$$

$$Var(x_k | z_{k-1}) = E[x_k x_k^T | z_{k-1}] - F_k m_{k-1} m_{k-1}^T F_k^T$$

$$= \int x_k x_k^T \; P(x_k | z_{k-1}) \, dx_k - F_k m_{k-1} m_{k-1}^T F_k^T$$

$$= \iint x_k x_k^T \; N(x_{k-1}; m_{k-1}, P_{k-1}) N(x_k; F_k x_{k-1}, Q_k) \, dx_{k-1} dx_k - F_k m_{k-1} m_{k-1}^T F_k^T$$

$$= \int N(x_{k-1}; m_{k-1}, P_{k-1}) \int x_k x_k^T \; N(x_k; F_k x_{k-1}, Q_k) \, dx_k \, dx_{k-1} - F_k m_{k-1} m_{k-1}^T F_k^T$$

$$= \int (Q_k + F_k x_{k-1} x_{k-1}^T F_k^T) N(x_{k-1}; m_{k-1}, P_{k-1}) \, dx_{k-1} - F_k m_{k-1} m_{k-1}^T F_k^T$$

$$= Q_k + F_k P_{k-1} F_k^T + F_k m_{k-1} m_{k-1}^T F_k^T - F_k m_{k-1} m_{k-1}^T F_k^T$$

$$= \underline{Q_k + F_k P_{k-1} F_k^T}$$

$$P(x_k | z_{k-1}) = N(F_k m_{k-1}, \; Q_k + F_k P_{k-1} F_k^T) = N(x_{k|k-1}, P_{k|k-1})$$

Kalman Filter : Update $\qquad z_k = H_k x_k + w_k$

$$P(x_k, z_k | z_{k-1}) = P(x_k, z_k | z_{k-1}) = P(z_k | x_k) P(x_k | z_{k-1}) \quad \text{where } P(z_k | x_k) = N(H_k x_k, R_k)$$

$$\begin{bmatrix} x_k \\ z_k \end{bmatrix}_{|z_{k-1}} \sim N\left(\begin{bmatrix} x_{k|k-1} \\ H_k x_{k|k-1} \end{bmatrix}, \begin{bmatrix} P_{k|k-1} & P_{k|k-1} H_k^T \\ H_k P_{k|k-1} & R_k + H_k P_{k|k-1} H_k^T \end{bmatrix} \right)$$

$$E[z_k | z_{k-1}] = \int z_k \, P(z_k | z_{k-1}) \, dz_k$$

$$= \iint z_k \; P(x_k, z_k | z_{k-1}) \, dx_k \, dz_k$$

$$= \iint z_k \; N(x_k; x_{k|k-1}, P_{k|k-1}) N(z_k; H_k x_k, R_k) \, dx_k \, dz_k$$

$$= \int N(x_k; x_{k|k-1}, P_{k|k-1}) \int z_k N(z_k; H_k x_k, R_k) \, dz_k \, dx_k$$

$$= \int H_k x_k \; N(x_k; x_{k|k-1}, P_{k|k-1}) \, dx_k$$

$$= \underline{H_k x_{k|k-1}}$$

$$Cov(x_k, z_k | z_{k-1}) = E[x_k z_k^T | z_{k-1}] - x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= \int N(x_k; x_{k|k-1}, P_{k|k-1}) \int x_k z_k^T N(z_k; H_k x_k, R_k) \, dz_k \, dx_k - x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= \int x_k x_k^T H_k^T N(x_k; x_{k|k-1}, P_{k|k-1})\, dx_k - x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= P_{k|k-1} H_k^T + x_{k|k-1} x_{k|k-1}^T H_k^T - x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= \underline{P_{k|k-1} H_k^T}$$

$$Var(z_k|Z_{k-1}) = E[z_k z_k^T | Z_{k-1}] - H_k x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= \int N(x_k; x_{k|k-1}, P_{k|k-1}) \int z_k z_k^T N(z_k; H_k x_k, R_k)\, dz_k\, dx_k - H_k x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= \int (R_k + H_k x_k x_k^T H_k^T) N(x_k; x_{k|k-1}, P_{k|k-1})\, dx_k - H_k x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= R_k + H_k P_{k|k-1} H_k^T + H_k x_{k|k-1} x_{k|k-1}^T H_k^T - H_k x_{k|k-1} x_{k|k-1}^T H_k^T$$

$$= \underline{R_k + H_k P_{k|k-1} H_k^T}$$

$$P(x_k | z_k) = ?$$

Let $a = A x_k + B z_k \perp\!\!\!\perp z_k$

$$\Rightarrow A P_{k|k-1} H_k^T + B \, R_k + H_k P_{k|k-1} H_k^T = 0$$

$$\Rightarrow a = x_k - P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} z_k = x_k + M z_k$$

$$E[x_k | z_k] = E[a - M z_k | z_k]$$

$$= E[a] - M z_k$$

$$= E[x_k] + M(E[z_k] - z_k)$$

$$= x_{k|k-1} + P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} (z_k - H_k x_{k|k-1})$$

$$Var(x_k | z_k) = Var(a - M z_k | z_k)$$

$$= Var(a)$$

$$= Var(x_k) + M\, Var(z_k)\, M^T$$

$$= P_{k|k-1} - P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} \cancel{(R_k + H_k P_{k|k-1} H_k^T)} \cancel{(R_k + H_k P_{k|k-1} H_k^T)^{-1}} H_k P_{k|k-1}$$

$$= P_{k|k-1} - P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} H_k P_{k|k-1}$$

Summary

Model: $x_k = F_k x_{k-1} + v_k$, $v_k \sim N(0, Q_k)$

$z_k = H_k x_k + w_k$, $w_k \sim N(0, R_k)$

Notation

State vector at time step k:	$\mathbf{x}_k$
Measurement vector at time step k:	$\mathbf{z}_k$
Set of measurements up to step k:	$Z^k = \{\mathbf{z}_0, \ldots, \mathbf{z}_k\}$
Estimate of $\mathbf{x}_k$ given measurements up to step j:	$\mathbf{x}_{k\|j}$
Covariance associated with $\mathbf{x}_{k\|j}$	$\mathbf{P}_{k\|j}$
Observation/Measurement matrix for time time k:	$\mathbf{H}_k$
State transition matrix for time time k:	$\mathbf{F}_k$

Algorithm :

Predict :

$$x_{k|k-1} = F_k x_{k-1|k-1}$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

Update :

$$x_{k|k} = x_{k|k-1} + P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} (z_k - H_k x_{k|k-1})$$

$$P_{k|k} = P_{k|k-1} - P_{k|k-1} H_k^T (R_k + H_k P_{k|k-1} H_k^T)^{-1} H_k P_{k|k-1}$$

Introduce Kalman Gain and some variables :

$$x_{k|k} = x_{k|k-1} + \underbrace{P_{k|k-1} H_k^T (\overbrace{R_k + H_k P_{k|k-1} H_k^T}^{S_k})^{-1}}_{K_k} (\underbrace{z_k - \overbrace{H_k x_{k|k-1}}^{z_{k|k-1}}}_{\nu_k})$$

$\Rightarrow$ $S_k = R_k + H_k P_{k|k-1} H_k^T$ = expected measurement covariance

$z_{k|k-1} = H_k x_{k|k-1}$ = expected measurement

$\nu_k = z_k - z_{k|k-1}$ = measurement residual

$K_k = P_{k|k-1} H_k^T S_k^{-1}$ = Kalman Gain.

Rewrite the algorithm with introduced variables

Predict:

$$x_{k|k-1} = F_k x_{k-1|k-1}$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

Update :

$$x_{k|k} = x_{k|k-1} + K_k \nu_k$$

$$P_{k|k} = P_{k|k-1} - K_k S_k^{-1} K_k^T$$

Reference :

Särkkä, S. 2013. Bayesian Filtering and Smoothing. Cambridge University Press.