UPDATE STEP

$$p(x_k \mid z_{1:k}) \propto \underbrace{p(z_k \mid x_k)}_{N(C_k x_k,\, R_k)} \; \underbrace{p(x_k \mid z_{1:k-1})}_{N(\bar{\mu}_k,\, \bar{\Sigma}_k)}$$

$$\Rightarrow p(x_k \mid z_{1:k}) \propto e\left\{-\tfrac{1}{2}(z_k - C_k x_k)^T R_k^{-1}(z_k - C_k x_k) - \tfrac{1}{2}(x_k - \bar{\mu}_k)^T \bar{\Sigma}_k^{-1}(x_k - \bar{\mu}_k)\right\}$$

which is quadratic in $x_k \Rightarrow p(x_k \mid z_{1:k})$ is a gaussian.

$$J \triangleq \tfrac{1}{2}(z_k - C_k x_k)^T R_k^{-1}(z_k - C_k x_k) - \tfrac{1}{2}(x_k - \bar{\mu}_k)^T \bar{\Sigma}_k^{-1}(x_k - \bar{\mu}_k)$$

$$\frac{\partial J}{\partial x_k} = -C_k^T R_k^{-1}(z_k - C_k x_k) + \bar{\Sigma}_k^{-1}(x_k - \bar{\mu}_k)$$

[quadratic form - see "Matrix Cookbook SS 2.4.2 online]

$$\frac{\partial^2 J}{\partial x_k^2} = C_k^T R_k^{-1} C_k + \bar{\Sigma}_k^{-1}$$

so second derivative is the inverse of covariance of the posterior

$$\Rightarrow \boxed{\Sigma_k = \left(C_k^T R_k^{-1} C_k + \bar{\Sigma}_k^{-1}\right)^{-1}}$$

the mean is where $\frac{\partial J}{\partial x_k} = 0$

$$\Rightarrow \underbrace{C_k^T R_k^{-1}(z_k - C_k \mu_k)} = \bar{\Sigma}^{-1}(\mu_k - \bar{\mu}_k)$$

[substituted $x_k = \mu_k$ as this is the maximum]

$$= C_k^T R_k^{-1}(z_k - C_k \mu_k + C_k \bar{\mu}_k - C_k \bar{\mu}_k)$$

$$= C_k^T R_k^{-1}(z_k - C_k \bar{\mu}_k) - C_k^T R_k^{-1} C_k(\mu_k - \bar{\mu}_k)$$

$$\Rightarrow C_k^T R_k^{-1}(z_k - C_k \mu_k) = \underbrace{\left(C_k^T R_k^{-1} C_k + \bar{\Sigma}^{-1}\right)}_{\Sigma_k^{-1}}(\mu_k - \bar{\mu}_k)$$

$$\Rightarrow \Sigma_k C_k R_k^{-1}(z_k - C_k \mu_k) = \mu_k - \bar{\mu}_k$$

UPDATE STEP (2)

define the "KALMAN GAIN"

$$\boxed{K_k = \Sigma_k C_k^T R_k^{-1}}$$

$$\Rightarrow \boxed{\mu_k = \bar{\mu}_k + \underbrace{K_k}_{\text{"GAIN"}} \underbrace{(z_k - C_k \bar{\mu}_k)}_{\text{"INNOVATION"}}} \qquad \& \quad \Sigma_k = \left(C_k^T R_k^{-1} C_k + \bar{\Sigma}_k^{-1}\right)^{-1}$$

So - this current form is inconvenient as need to take inverse of $n \times n$ matrix twice. If the state is large (like in SLAM!) very expensive.

We can reformulate K_k & Σ_k to reduce computational cost.

$$K_k = \Sigma_k C_k^T R_k^{-1} \underbrace{(C_k \bar{\Sigma}_k C_k^T + R_k)(C_k \bar{\Sigma}_k C_k^T + R_k)^{-1}}_{I}$$

$$= \Sigma_k \left(C_k^T R_k^{-1} C_k \bar{\Sigma}_k C_k^T + C_k^T \cancel{R_k^{-1} R_k}\right)\left(C_k \bar{\Sigma}_k C_k^T + R_k\right)^{-1}$$

$$= \underbrace{\Sigma_k \left(C_k^T R_k^{-1} C_k \bar{\Sigma}_k C_k^T + \bar{\Sigma}_k^{-1} \bar{\Sigma}_k C_k^T\right)}\left(C_k \cdots\right)^{-1}$$

$$= \Sigma_k \underbrace{\left(C_k^T R_k^{-1} C_k + \bar{\Sigma}_k^{-1}\right)} \bar{\Sigma}_k C_k^T \quad (\cdots)$$

$$\cancel{= \Sigma_k \Sigma_k^{-1} \Sigma_k} \quad \Sigma_k^{-1}$$

$$\boxed{K_k = \bar{\Sigma}_k C_k^T \left(C_k \bar{\Sigma}_k C_k^T + R_k\right)^{-1}}$$

Now the inverse is of measurement vector dimension!

likewise the covariance Σ_k can be reformulated: using the matrix inversion lemma

$$\Sigma_k = \left(\bar{\Sigma}_k^{-1} + C_k^T R_k^{-1} C_k\right)^{-1} = \bar{\Sigma}_k - \bar{\Sigma}_k C_k^T \left(R_k + C_k \bar{\Sigma}_k C_k^T\right)^{-1} C_k \bar{\Sigma}_k$$

$$= \left[I - \bar{\Sigma}_k C_k^T \left(C_k \bar{\Sigma}_k C_k^T + R_k\right)^{-1} C_k\right] \bar{\Sigma}_k$$

$$\boxed{\Sigma_k = I - K_k C_k \bar{\Sigma}_k}$$

PREDICTION STEP.

Look in "Probabilistic Robotics" for a similar derivation that does not use Gaussian Marginal Distributions.

This derivation is from "Pattern Recognition and Machine Learning" C. Bishop SS 2.3.3

$$p(x_k \mid z_{1:k-1}) = \int \underbrace{p(x_k \mid x_{k-1})}_{N(A x_{k-1} + B,\, Q_k)} \underbrace{p(x_{k-1} \mid z_{1:k-1})}_{N(M_{k-1},\, \Sigma_{k-1})} \, d x_{k-1}$$

outline of proof.

① find joint distribution $p(x_k, x_{k-1} \mid z_{1:k-1}) = p(x_k \mid x_{k-1})\, p(x_{k-1} \mid z_{1:k-1})$

② use marginal distribution $p(x_k \mid z_{1:k-1}) = \int p(x_k, x_{k-1} \mid z_{1:k-1})\, d x_{k-1}$ identity

define $\mathbb{X} = \begin{bmatrix} x_{k-1} \\ x_k \end{bmatrix}$, $\Lambda = \Sigma_{k-1}^{-1}$, $L = Q_k^{-1}$

Now $$p(x_k \mid x_{k-1})\, p(x_{k-1} \mid z_{1:k-1}) \propto e^{\overbrace{\left\{ -\frac{1}{2}(x_{k-1} - M_{k-1})^T \Lambda (x_{k-1} - M_{k-1}) - \frac{1}{2}(x_k - A x_{k-1} - B)^T L (x_k - A x_{k-1} - B) \cancel{\text{const}} \right\}}^{\text{"}J\text{"}}}$$

this is a quadratic function in components of $\mathbb{X}$ $\Rightarrow$ $p(x_k, x_{k-1} \mid z_{1:k-1})$ is Gaussian!

To find covariance of this function, expand J second order terms.

$$J = -\frac{1}{2} x_{k-1}^T (\Lambda + A^T L A) x_{k-1} - \frac{1}{2} x_k^T L x_k + \frac{1}{2} x_k^T L A x_{k-1} + \frac{1}{2} x_{k-1} A^T L x_k + \text{lower order terms.}$$

$$= -\frac{1}{2} \begin{bmatrix} x_{k-1} \\ x_k \end{bmatrix}^T \begin{pmatrix} \Lambda + A^T L A & -A^T L \\ -LA & L \end{pmatrix} \begin{bmatrix} x_{k-1} \\ x_k \end{bmatrix} + \text{lower order terms} \quad \text{— eq } (1)$$

$$\Rightarrow \operatorname{cov}[\mathbb{X}] = \begin{pmatrix} \Lambda + A^T L A & -A^T L \\ -LA & L \end{pmatrix}^{-1}$$

[MATRIX INVERSION LEMMA]

$$\operatorname{cov}[\mathbb{X}] = \begin{pmatrix} \Lambda^{-1} & \Lambda^{-1} A^T \\ A \Lambda^{-1} & L^{-1} + A \Lambda^{-1} A^T \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma_{k-1} & \Sigma_{k-1} A^T \\ A \Sigma_{k-1} & Q_k + A \Sigma_{k-1} A^T \end{pmatrix} \triangleq \Sigma_{\mathbb{X}}$$

PREDICTION STEP (2)

Now Expand 'J' in terms of the FIRST ORDER TERMS

$$J = \text{second order terms} + x_{k-1}^T \Lambda \mu_{k-1} - x_{k-1}^T A^T L B + x_k^T L B \quad + \text{lower order terms}$$

$$= \begin{bmatrix} x_{k-1} \\ x_k \end{bmatrix} \begin{pmatrix} \Lambda \mu_{k-1} - A^T L B \\ L B \end{pmatrix} + \text{2nd \& 0th order terms.} \qquad - \text{eq} (2)$$

$$E[X] = \Sigma_x \begin{pmatrix} \Lambda \mu_{k-1} - A^T L B \\ L B \end{pmatrix} \stackrel{\Delta}{=} \mu_X$$

(we can see this by comparing terms to the gaussian

$$p(X) \propto e\left\{ -\frac{1}{2} (X - \mu_X)^T \Sigma_X^{-1} (X - \mu_X) \right\}$$

$$\propto e\left\{ \underbrace{-\frac{1}{2} X^T \Sigma_x^{-1} X} + \underbrace{\mu_X^T \Sigma_X^{-1} X}_{= X^T \Sigma_X^{-1} \mu_X} - \frac{1}{2} \mu_X^T \Sigma_X \mu_X \right\}$$

compare to terms in J (eq 1 + eq 2))

$$E[x] = \begin{pmatrix} \Lambda^{-1} & \Lambda^{-1} A^T \\ A \Lambda^{-1} & L^{-1} + A \Lambda^{-1} A^T \end{pmatrix} \begin{pmatrix} \Lambda \mu_{k-1} - A^T L B \\ L B \end{pmatrix}$$

$$= \begin{bmatrix} \mu_{k-1} - \Lambda^{-1} A^T L B + \Lambda^{-1} A^T L B \\ A \Lambda^{-1} \Lambda \mu_{k-1} - A \Lambda^{-1} A^T L B + L^{-1} L B + A \Lambda^{-1} A^T L B \end{bmatrix}$$

$$= \begin{bmatrix} \mu_{k-1} \\ A \mu_{k-1} + B \end{bmatrix}$$

$$\Rightarrow \cancel{p(x_k \mid z_{1:k-1})} \; p(x_{k-1}, x_k) = N\left(\begin{bmatrix} x_{k-1} \\ x_k \end{bmatrix} ; \begin{bmatrix} \mu_{k-1} \\ A \mu_{k-1} + B \end{bmatrix}, \begin{bmatrix} \Sigma_{k-1} & \Sigma_{k-1} A^T \\ A \Sigma_{k-1} & Q_k + A \Sigma_{k-1} A^T \end{bmatrix} \right)$$

~~⇒ Using marginal~~

~~⇒ $p(x_k \mid z_{1:k-1}) =$~~

PREDICTION STEP (3)

so $$p(x_{k-1}, x_k \mid z_{1:k-1}) = N\left(\begin{bmatrix} \mu_{k-1} \\ A\mu_{k-1} + B \end{bmatrix}, \begin{bmatrix} \Sigma_{k-1} & \Sigma_{k-1} A^T \\ A\Sigma_{k-1} & Q_k + A\Sigma_{k-1}A^T \end{bmatrix}\right)$$

$$\Rightarrow \quad p(x_k \mid z_{1:k-1}) = \int p(x_{k-1}, x_k \mid z_{1:k-1}) \, dx_{k-1}$$

[by Gaussian identity]

$$\Rightarrow \quad p(x_k \mid z_{1:k-1}) = N\left(\left[A\mu_{k-1} + B\right], \left[A\Sigma_{k-1}A^T + Q_k\right]\right)$$