Proofs Related to Kalman Filtering

Anirban Sinha

December 10, 2019

1 Proofs related to Kalman filter derivation

click reference link.

1.1 Proof of the rule of iterated expectation

Proof 1: $\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]$ First notice that, $\mathbb{E}[\mathbb{E}[X|Y]]$ is a random variable because the expected value changes based on the realization of Y. If X and Y are discrete random variables, then the proof goes as follows.

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[X|Y]] &= \sum_{y\in R_Y} \mathbb{E}[X|Y=y]p_Y(y) \qquad \text{(by the definition of expected value)} \\
&= \sum_{y\in R_Y}\sum_{x\in R_X} x p_{X|Y=y}(x)p_Y(y) \qquad \text{(by the definition of conditional expectation)} \\
&= \sum_{x\in R_X}\sum_{y\in R_Y} x p_{XY}(x,y) \qquad \text{(where } p_{XY} \text{ is the joint probability mass function} \\
&= \sum_{x\in R_X} x \sum_{y\in R_Y} p_{XY}(x,y) \\
&= \sum_{x\in R_X} x p_X(x) \qquad \text{(marginalization of the joint pmf)} \\
&= \mathbb{E}[X] \qquad \text{(by the definition of expected value)} \qquad (1)
\end{aligned}$$

If X and Y are continuous random variables, then the proof goes as follows.

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[X|Y]] &= \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y] f_Y(y) dy \qquad \text{(by the definition of expected value)} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y=y}(x) dx f_Y(y) dy \qquad \text{(by the definition of conditional expectation)} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y=y}(x) f_Y(y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{XY}(x,y) dy dx \\
&= \int_{-\infty}^{\infty} x f_X(x) dx \\
&= \mathbb{E}[X] \qquad (2)
\end{aligned}$$

1.2 Prediction error has zero mean

Proof 2: $\mathbb{E}[\tilde{x}^-] = 0$

$$\begin{aligned}
\mathbb{E}[\tilde{x}^-] &= \mathbb{E}[x_k - \hat{x}_k] \\
&= \mathbb{E}[x_k] - \mathbb{E}[\hat{x}_k] \\
&= \mathbb{E}[x_k] - \mathbb{E}[\mathbb{E}[\hat{x}_k|\mathbb{Z}_{k-1}]] \\
&= \mathbb{E}[x_k] - \mathbb{E}[x_k] \qquad \text{(using rule of iterated expectation)} \\
&= 0 \qquad (3)
\end{aligned}$$

1.3 Measurement error/ Innovation has zero mean

Proof 3: $\mathbb{E}[\tilde{z}] = 0$

$$\begin{aligned}
\mathbb{E}[\tilde{z}] &= \mathbb{E}[z_k - \hat{z}_k] \\
&= \mathbb{E}[z_k] - \mathbb{E}[\hat{z}_k] \\
&= \mathbb{E}[z_k] - \mathbb{E}[\mathbb{E}[\hat{z}_k|\mathbb{Z}_{k-1}]] \\
&= \mathbb{E}[z_k] - \mathbb{E}[z_k] \qquad \text{(using rule of iterated expectation)} \\
&= 0 \qquad (4)
\end{aligned}$$

1.4 Find conditional PDF of two random variables

Here we are interested in finding $p(a|b)$ when a and b are jointly Gaussian vectors.

Step 1: Define the augmented vector, mean and covariance

We first combine a and b into an augmented vector Y where,

$$Y = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and,} \quad \mathbb{E}[Y] = \begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix} = \bar{Y} \qquad (5)$$

The covariance of the joint distribution is,

$$\Sigma_Y = \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix} \tag{6}$$

Step 2: Use the rule of joint pdf

From the rule of conditional pdf we know that $p(a,b) = p(a|b)p(b)$. Therefore,

$$\begin{aligned}
p(a|b) &= \frac{p(a,b)}{p(b)} \\
&\propto \frac{\exp\left(-\frac{1}{2}(Y - \mathbb{E}[Y])^T \Sigma_Y^{-1} (Y - \mathbb{E}[Y])\right)}{\exp\left(-\frac{1}{2}(b - \mathbb{E}[b])^T \Sigma_b^{-1} (b - \mathbb{E}[b])\right)} \\
&= \mathcal{C} \frac{\exp\left(-\frac{1}{2}(Y - \bar{Y})^T \Sigma_Y^{-1} (Y - \bar{Y})\right)}{\exp\left(-\frac{1}{2}(b - \bar{b})^T \Sigma_b^{-1} (b - \bar{b})\right)} \quad \text{(where } \mathcal{C} \text{ is the normalizing constant)} \\
&= \mathcal{C} \exp\left\{-\frac{1}{2}(Y - \mathbb{E}[Y])^T \Sigma_Y^{-1} (Y - \mathbb{E}[Y])) + \frac{1}{2}(b - \bar{b})^T \Sigma_b^{-1} (b - \bar{b})\right\} \qquad (7)
\end{aligned}$$

The covariance Σ_Y in equation (7), can be factorized as,

$$\begin{aligned}
\Sigma_Y &= \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{I} & \Sigma_{ab}\Sigma_{bb}^{-1} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba} & \mathbf{0} \\ \mathbf{0} & \Sigma_{bb} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \Sigma_{bb}^{-1}\Sigma_{ba} & \boldsymbol{I} \end{bmatrix} \qquad (8)
\end{aligned}$$

Then the inverse of Σ_Y can be found as,

$$\begin{aligned}
\Sigma_Y^{-1} &= \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \Sigma_{bb}^{-1}\Sigma_{ba} & \boldsymbol{I} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba} & \mathbf{0} \\ \mathbf{0} & \Sigma_{bb} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{I} & \Sigma_{ab}\Sigma_{bb}^{-1} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ -\Sigma_{bb}^{-1}\Sigma_{ba} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{bb}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\Sigma_{ab}\Sigma_{bb}^{-1} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} \qquad (9)
\end{aligned}$$

Substituting, the expression of Σ_Y^{-1} from Equation (9) into equation (7), we get

$$\begin{aligned}
p(a|b) &= \mathcal{C} \exp\left\{-\frac{1}{2}\left(Y - (\bar{Y} + \Sigma_{ab}\Sigma_b^{-1}(b - \bar{b}))\right)^T \left(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right)^{-1} \left(Y - (\bar{Y} + \Sigma_{ab}\Sigma_b^{-1}(b - \bar{b}))\right)\right\} \\
&\sim \mathcal{N}\left(\bar{Y} + \Sigma_{ab}\Sigma_b^{-1}(b - \bar{b}), \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}\right) \qquad (10)
\end{aligned}$$