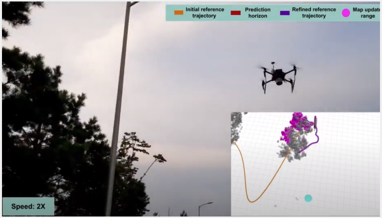
Autonomous Mobile Robotics

Kalman Filter

Geesara Kulathunga

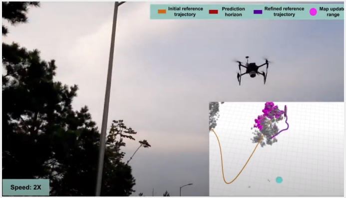
February 20, 2023



Kalman Filter

Contents

- Gaussian Distribution
- One Dimensional Kalman Filter
- Multivariate Density Function
- Marginal Density Function
- Multivariate Normal Function
- Two Dimensional Gaussian
- Multiple Random Variable
- Multidimensional Kalman Filter
- Sensor Fusion
- Linearization, Taylor Series Expansion, Linear Systems
- Extended Kalman Filter (EKF)
- Comparison between KF and EKF



Example 00

Gaussian comes under continuous probability distribution which describes with two parameters: the mean μ and the variance σ^2.

$$f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{1}$$

- Let's plot the probability density function (pdf) of $\mathcal{N}(10, 1) \times \mathcal{N}(10, 1)$. What can you say about this result (Fig. 1)?
- Similarly what about the result pdf with $\mathcal{N}(10.2, 1) \times \mathcal{N}(9.7, 1)$ and $\mathcal{N}(8.5, 1.5) \times \mathcal{N}(10.2, 1.5)$? (Fig. 2 and Fig. 3)

Example oo

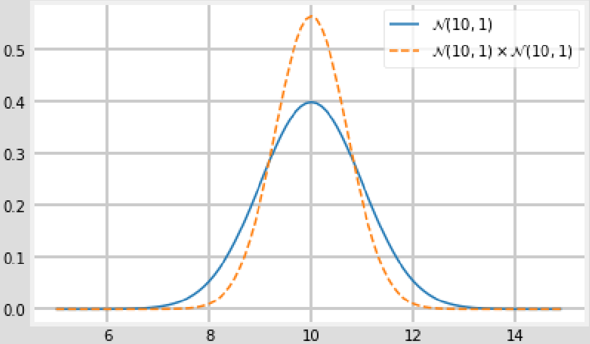


Figure: $\mathcal{N}(10, 1) \times \mathcal{N}(10, 1)$

Example oo

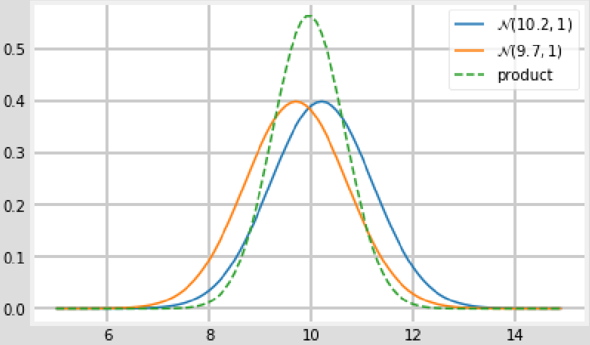


Figure: $\mathcal{N}(10.2,\ 1) \times \mathcal{N}(9.7,\ 1)$

Example 00

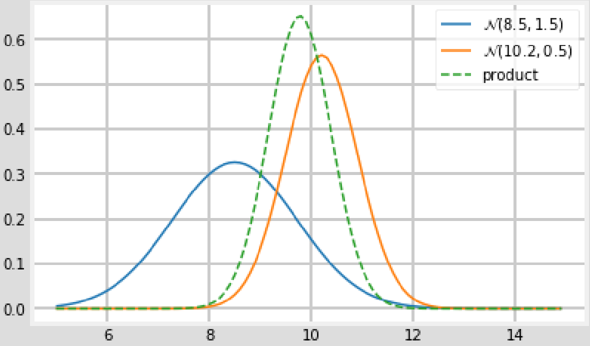


Figure: $\mathcal{N}(10.2,\ 1.5) \times \mathcal{N}(8.5,\ 1.5)$

Example 01

There are two independent estimates of the variable $x_1 \sim N(E(x_1), \sigma_1^2)$ and , $x_2 \sim N(E(x_2), \sigma_2^2)$. How can we find the optimal linear combination of these two estimates that represents the corresponding optimal state estimate $\hat{x}$?

Example 01

- The optimal value of the variable x is assumed to be a linear combination of two estimates

$$\hat{x} = w_1 x_1 + w_2 x_2,\ w_1 + w_2 = 1$$

Example 01

- The optimal value of the variable x is assumed to be a linear combination of two estimates

$$\hat{x} = w_1 x_1 + w_2 x_2,\ w_1 + w_2 = 1$$

- Let us find out variance of $\hat{x}$

$$\begin{aligned} \sigma^2 &= E\{(\hat{x} - E\{\hat{x}\})^2\} \\ &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2E\{(x_1 - E\{x_1\})(x_2 - E\{x_2\})\} \\ &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 = (1-w)^2\sigma_1^2 + w^2\sigma_2^2 \end{aligned} \tag{2}$$

Example 01

- The optimal value of the variable x is assumed to be a linear combination of two estimates

$$\hat{x} = w_1 x_1 + w_2 x_2,\ w_1 + w_2 = 1$$

- Let us find out variance of $\hat{x}$

$$\begin{aligned} \sigma^2 &= E\{(\hat{x} - E\{\hat{x}\})^2\} \\ &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2E\{(x_1 - E\{x_1\})(x_2 - E\{x_2\})\} \\ &= w_1^2\sigma_1^2 + w_2^2\sigma_2^2 = (1-w)^2\sigma_1^2 + w^2\sigma_2^2 \end{aligned} \tag{2}$$

- Let's try to minimize the variance

$$\frac{\partial \sigma^2}{\partial w} = -2(1-w)\sigma_1^2 + 2w\sigma_2^2 = 0 \tag{3}$$

Example 01

- The optimal value for w

$$w = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

Example 01

- The optimal value for w

$$w = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

- The optimal value for σ^2

$$\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

Example 01

- The optimal value for w

$$w = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

- The optimal value for σ^2

$$\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

- The optimal value $\hat{x}$

$$\hat{x} = \frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2}$$

Example 02

Consider example 01, if control input $u(k)$, which is having $\sigma_u^2(k)$ uncertainty, applied to the estimated state $\hat{x}(k)$ determine the uncertainty of the updated state, i.e., $\hat{x}(k+1) = \hat{x}(k) + u(k)$

Example 02

Consider example 01, if control input $u(k)$, which is having $\sigma_u^2(k)$ uncertainty, applied to the estimated state $\hat{x}(k)$ determine the uncertainty of the updated state, i.e., $\hat{x}(k+1) = \hat{x}(k) + u(k)$

$$\begin{aligned}\sigma^2(k+1) &= E\{(\hat{x}(k+1) - E\{\hat{x}(k+1)\})^2\} \\ &= E\{(\hat{x}(k) - E\{\hat{x}(k)\} + \hat{u}(k) - E\{\hat{u}(k)\})^2\} \\ &= \sigma^2(k) + \sigma_u^2(k)\end{aligned} \tag{4}$$

One Dimensional Kalman Filter

Let’s start off with this example:
`http://david.wf/kalmanfilter/`

Multivariate Density Function

When there are more than one random variable, i.e. $\mathbf{x}, \mathbf{y}$, we stack them into a vector, i.e. $[\mathbf{x}, \mathbf{y}]$, and let new random variable be $\mathbf{z} \in \mathbb{R}^n$. Thus, probability density function is called the **Joint Density Function**, which we can express as:

$$p_Z(\mathbf{z}) : \mathbb{R}^n \rightarrow \mathbb{R}^+ \tag{5}$$

Marginal Density Function

Subsequently, when $\mathbf{z}$ takes values with a range, e.g., a and b, the corresponding probability is given by

$$P_z(a \leq \mathbf{z} \leq b) = \int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} p_z(\mathbf{z}) dz_1 .. dz_n, \tag{6}$$

that is called the **Marginal Density Function**. So what is the condition should hold for a given join density function to be valid?

$$P_{XY}(x, y) \geq 0, \quad \sum_{\forall x} \sum_{\forall y} P_{XY}(x, y) = 1$$

Moreover, what is the relationship between joint density function and marginal density function?
Consider the following example:

$$P(x, y) = 4x^2 y, \quad 0 < x < 4, \quad 0 < y < 3 \tag{7}$$

How can you find the $P_X(x)$?

Multivariate means multiple variables. Main intuition here is to represent multiple variables with normal distribution. Let's recap a few basic concepts. The covariance between x and y is

$$cov(\mathbf{x}, \mathbf{y}) = \sigma_{xy} = \mathbb{E}\left[(\mathbf{x} - \mu_x)(\mathbf{y} - \mu_y)^T\right]$$

where $\mathbb{E}[\mathbf{x}]$ is the **expected value** of $\mathbf{x}$ is given by

$$\mathbb{E}[\mathbf{x}] = \begin{cases} \sum_{i=1}^{n} p_i x_i & \text{discrete} \\ \int_{-\infty}^{\infty} f(x)\, x & \text{continuous} \end{cases}$$

If the each data point is equally likely, so the probability of each event is $\frac{1}{N}$. Then the expectation can be express as follows for the discrete case.

$$\mathbb{E}[\mathbf{x}] = \frac{1}{N} \sum_{i=1}^{n} x_i$$

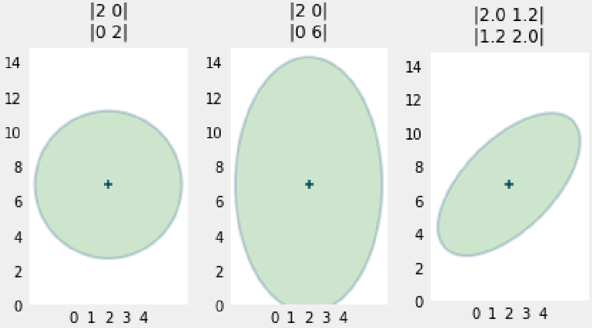
Compare covariance with variance

$$var(\mathbf{x}) = \sigma_x^2 = \mathbb{E}[(\mathbf{x} - \mu)^2]$$
$$cov(\mathbf{x}, \mathbf{y}) = \sigma_{xy} = \mathbb{E}\big[(\mathbf{x} - \mu_x)(\mathbf{y} - \mu_y)\big] = \mathbb{E}[\mathbf{xy}] - \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{y}]$$

What if $\mathbb{E}[\mathbf{xy}] == \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{y}]$? When there is **no covariance between two random variables**, e.g., $\mathbf{x}, \mathbf{y}$, we can say that those two random variables are uncorrelated, i.e, no linear dependency, but does not say anything about independence. However, if two random variables are independent, then those are uncorrelated, i.e., $E[\mathbf{xy}] == \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{y}]$ as well. Covariance matrix (Σ) denotes covariances of a multivariate normal distribution.

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

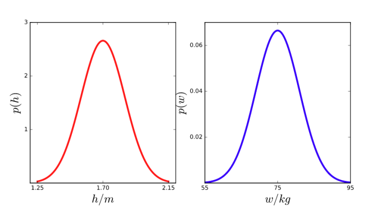
Multivariate Normal Function



The correlation between two variables can be given a numerical value with **Pearson's Correlation Coefficient** ρ_{xy}. It is defined as:

$$\rho_{xy} = \frac{cov(\mathbf{x}, \mathbf{y})}{\sigma_x \sigma_y} \tag{8}$$

Consider the students' weight and height distributions. The marginal density of heights is given by 1.7*m* and variance is 0.025 approximately. Similarly, the marginal density of weights is given by 75kg and variance is 36.



Let's assume there is **no correlation between weight and height distributions**. In other words, student height does not depend on her/his weight. Whats is the **joint probability distribution** p(w, h) ?

Two Dimensional Gaussian

Let's assume there is no correlation between weight and height distributions. In other words, student height does not depend on her/his weight.

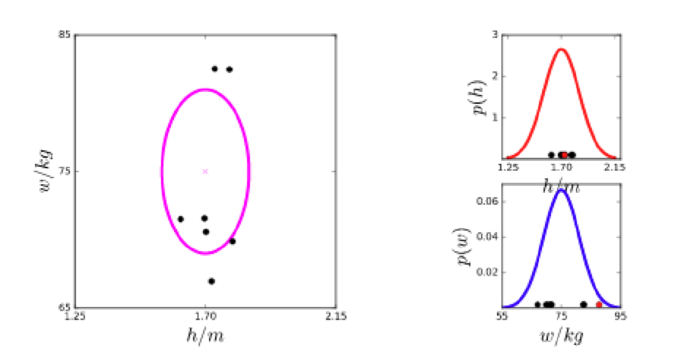
$$p(\mathbf{z}) = p(w, h) = p(w)p(h)$$

$$p(w, h) = \frac{1}{\sqrt{2\pi\sigma_1^2}\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{1}{2}\left(\frac{(w-\mu_1)^2}{\sigma_1^2} + \frac{(h-\mu_2)^2}{\sigma_2^2}\right)\right)$$

$$= \frac{1}{\sqrt{2\pi\sigma_1^2 2\pi\sigma_2^2}} \exp\left(-\frac{1}{2}\left(\begin{bmatrix} w \\ h \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)^\top \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} w \\ h \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)\right)$$

$$p(\mathbf{z}) = \frac{1}{\det 2\pi \mathbf{D}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{z}-\mu)^\top \mathbf{D}^{-1}(\mathbf{z}-\mu)\right) \tag{9}$$

Two Dimensional Gaussian



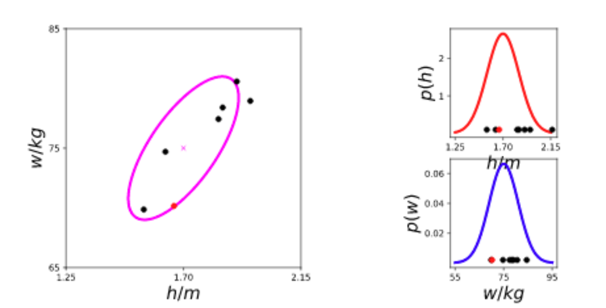
Two Dimensional Gaussian

But, are weight and height uncorrelated each other in reality? Answer is no, relation between weight and height is given by $BMI = \frac{w}{h^2}$. How can we incorporate this correlation into the model? Correlation can be added by using the following trick.

$$
\begin{aligned}
p(\mathbf{z}) &= \frac{1}{\det 2\pi \mathbf{D}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{z}-\mu)^\top \mathbf{D}^{-1}(\mathbf{z}-\mu)\right) \\
p(\mathbf{z}') &= \frac{1}{\det 2\pi \mathbf{D}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{R}^\top \mathbf{z}-\mathbf{R}^\top\mu)^\top \mathbf{D}^{-1}(\mathbf{R}^\top \mathbf{z}-\mathbf{R}^\top\mu)\right) \\
p(\mathbf{z}') &= \frac{1}{\det 2\pi \mathbf{D}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{z}-\mu)^\top \mathbf{R}\mathbf{D}^{-1}\mathbf{R}^\top(\mathbf{z}-\mu)\right)
\end{aligned}
\tag{10}
$$

where inverse of the covariance matrix is given by $C^{-1} = \mathbf{R}\mathbf{D}^{-1}\mathbf{R}^T$.

$$p(\mathbf{z}) = \frac{1}{\det 2\pi \mathbf{C}^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{z}-\mu)^\top \mathbf{C}^{-1}(\mathbf{z}-\mu)\right) \tag{11}$$



Recall the equation for the normal distribution:

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}(x-\mu)^2/\sigma^2\right] \tag{12}$$

Moreover, multivariate normal function is given by

$$f(\mathbf{z},\ \mu,\ \Sigma) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}}\ \exp\left[-\frac{1}{2}(\mathbf{z}-\mu)^\mathsf{T}\Sigma^{-1}(\mathbf{z}-\mu)\right] \tag{13}$$

Let $\mathbf{z} = \begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix}$ be independent random vector of two random variables and their means are given by

$$m_z = \begin{bmatrix}m_x\\ m_y\end{bmatrix} \tag{14}$$

Can you define the covariance matrix whose $\mathbf{z}$?

$$\Sigma = E[(\mathbf{z}-m_z)(\mathbf{z}-m_z)^T] = \begin{bmatrix}\Sigma_{xx} & \Sigma_{x,y}\\ \Sigma_{yx} & \Sigma_{yy}\end{bmatrix} \tag{15}$$

Now assume **z** is a Gaussian random variable. Thus, how can we define the joint probability density function of **z**?

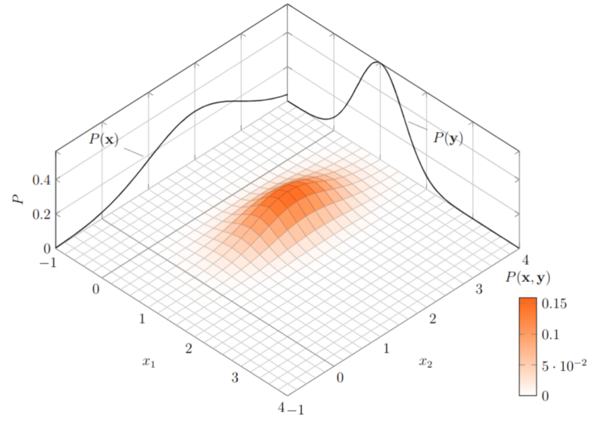
$$p(\mathbf{z}) = p(\mathbf{x}, \mathbf{y}) = \frac{1}{\sqrt{2\pi}^n \sqrt{det\Sigma}} exp(\frac{-1}{2} \begin{bmatrix} \mathbf{x} - m_x \\ \mathbf{y} - m_y \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} \mathbf{x} - m_x \\ \mathbf{y} - m_y \end{bmatrix}) \tag{16}$$

Can you try to define Σ_{xy} and Σ_{yy}? Also, what is the different between covariance and variance?

$$\Sigma_{xy} = E[(\mathbf{x} - m_x)(\mathbf{y} - m_y)^T] = E[(\mathbf{y} - m_y)(\mathbf{x} - m_x)^T]^T = \Sigma_{yx}^T \tag{17}$$

Now let's define the conditional mean and variance of x and y.

$$
\begin{aligned}
m_{X|y} &= m_X + \Sigma_{XY}\Sigma_{YY}^{-1}[y - m_Y] \\
\Sigma_{X|y} &= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}
\end{aligned}
\tag{18}
$$



Let's assume Σ^{-1} and μ are given as

$$\Sigma^{-1} = \begin{bmatrix} 3.5 & 2.5 \\ 2.5 & 4.0 \end{bmatrix}, \quad \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{19}$$

To visualize Σ^{-1} we have to perform the eigen decomposition of the covariance matrix Σ^{-1}. Eigen decomposition can be seen as a connection between linear transformation and the covariance matrix. An eigen vector is a vector whose direction remains unchanged when a linear transformation is applied $Av = \lambda v$, where v and λ are the the eigenvector and eigenvalue of A. In this example, we have 2 eiegenvectors V and corresponding eigenvalues as digonal matrix L, we can form the $\Sigma^{-1} = C$

$$\begin{aligned} CV &= VL \\ C &= VLV^{-1} \end{aligned} \tag{20}$$

We can use SVD to find out V and L

Multidimensional Kalman Filter

The optimization criterion used in one dimensional Kalman filter is minimization of least square error of the random variable x. For a given random process, relationship between current and next state and the observation model:

$$\begin{aligned} x_{k+1} &= \Phi_k x_k + w_k \\ z_k &= H_k x_k + v_k \end{aligned} \tag{21}$$

where x_k = (nx1) process state vector at time t_k
Φ_k = (nxn) is the state transition matrix from state t_k to t_{t+1}
w_k = (nx1) vector, a white noise sequence with known covariance structure
z_k = (mx1) measurment vector at time t_k
H_k = (mxn) transition matrix between measurement and state vector at time t_k
v_k = (mx1) measurment error (can you guess) and having zero crosscorrelation with the w_k

Corresponding covariance matrices for the w_k and v_k are given by

$$E[w_k w_i^T] = \begin{cases} Q_k & \text{if } i = k \\ 0 & i \neq k \end{cases} \tag{22}$$

$$E[v_k v_i^T] = \begin{cases} R_k & \text{if } i = k \\ 0 & i \neq k \end{cases} \tag{23}$$

$$E[w_k v_i^T] = 0 \text{ for all k and i} \tag{24}$$

- A normal multivariate Gaussian distribution

$$p(\mathbf{x}) = f(\mathbf{x},\ \mu,\ \Sigma) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}}\ \exp\left[-\frac{1}{2}(\mathbf{x}-\mu)^\mathsf{T}\Sigma^{-1}(\mathbf{x}-\mu)\right]$$

- Kalman filter equations can be written as

$$p(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{u}_k) = \frac{1}{\sqrt{(2\pi)^n|Q|}}$$

$$\exp\left[-\frac{1}{2}(\mathbf{x}_k-(A\mathbf{x}_{k-1}+B\mathbf{u}_{k-1}))^\mathsf{T}Q^{-1}(\mathbf{x}_k-(A\mathbf{x}_{k-1}+B\mathbf{u}_{k-1}))\right]$$

$$p(\mathbf{z}_k|\mathbf{x}_k) = \frac{1}{\sqrt{(2\pi)^n|R|}}\ \exp\left[-\frac{1}{2}(\mathbf{z}_k-\mathbf{x}_k)^\mathsf{T}R^{-1}(\mathbf{z}_k-\mathbf{x}_k)\right]$$

where discrete system dynamic is given by
$\mathbf{x}_k = A\mathbf{x}_{k-1} + B\mathbf{u}_{k-1}$