FILTRO DE KALMAN: CONCEITO, DEMONSTRAÇÃO MATEMÁTICA E APLICAÇÃO

Felipe Machado Mello (felipemachadomello@gmail.com); Manuella de Souza Gonçalves Guimarães (manuellasgg@outlook.com); Prof. MSc. Felipe Santana Santos (felipe.ssantos@souunit.com.br)

RESUMO

O filtro de Kalman é uma coleção de equações matemáticas que propicia a predição do estado futuro de um sistema dinâmico por meio da estimação do estado imediatamente anterior. Perante o objetivo de integrar a didática e o rigor matemático na apresentação e explicação do filtro de Kalman, este artigo possui uma seção de demonstração matemática, a qual contém a dedução detalhada de cada equação que compõe este filtro e o esclarecimento sobre as etapas parciais de desenvolvimento, e uma seção de aplicação que clarifica o funcionamento deste filtro a partir de uma situação-problema de engenharia. Os resultados obtidos comprovam que o filtro de Kalman é um eficiente algoritmo de otimização de estimação de estados.

Palavras-chave: *Kalman. Filtro. Algoritmo. Otimização. Estimação.*

ABSTRACT

The Kalman filter is a collection of mathematical equations that propitiates the prediction of the future state of a dynamical system by estimating the immediately previous state. In view of the objective of integrating didactics and mathematical rigor in the presentation and explanation of the Kalman filter, this article has a mathematical demonstration section, which contains a detailed deduction of each equation that makes up this filter and clarification on the partial steps of development, and an application section that clarifies how this filter works from an engineering problem situation. The obtained results prove that the Kalman filter is an efficient state estimation optimization algorithm.

Keywords: *Kalman. Filter. Algorithm. Optimization. Estimation.*

1 INTRODUÇÃO

A engenharia acontece quando algum ser em obscuro contexto se embasa em conhecimentos propiciados por outros seres para criar algo que ilumine a todos. Esta frase denota que o processo de criação de uma solução em engenharia necessita de um obstáculo e de um arcabouço de conhecimentos como pré-requisitos. Quando ambos são satisfeitos, inicia-se a etapa de estudo do problema. Nesta, a partir de informações conhecidas, desenvolve-se um modelo matemático da situação no qual as variáveis de interesse são elencadas.

As informações sobre o comportamento destas variáveis são coletadas pelos pesquisadores através de medições. Somente após a obtenção de medições é possível averiguar o comportamento do elemento de estudo, seja este um simples objeto ou um complexo sistema. A atitude de medir incansavelmente o comportamento de um evento é responsável pela concepção de notáveis formulações científicas. Por exemplo, a Lei de Coulomb a partir dos estudos de Charles Augustin de Coulomb, em 1783, sobre o experimento da balança de torção, a Lei de Biot-Savart a partir dos estudos de Jean-Baptiste Biot e Félix Savart, em 1820, alicerçados nos experimentos de Hans Christian Ørsted sobre o efeito da eletricidade em agulhas magnéticas, e a medição da carga elétrica do elétron a partir dos estudos de Robert Andrews Millikan e Harvey Fletcher, em 1909, sobre o experimento da gota de óleo. Como observado, esta ação é imprescindível para conhecer os fenômenos naturais, entretanto por residir no mundo sensível, é suscetível a imperfeições.

Desde que se tornou exequível a análise de sistemas dinâmicos a partir da criação do cálculo diferencial e integral, no século XVII, por Isaac Newton e Gottfried Wilhelm Leibniz, uma adversidade impactava a ciência. Como determinar o valor verdadeiro dos estados de um sistema dinâmico, sabendo-se que os equipamentos de medição disponíveis fornecem informações corrompidas por incertezas matemáticas? Alguns cientistas, que se ocuparam na tentativa de responder esta pergunta, conseguiram encontrar algumas soluções particulares. Contudo, somente em 1960 foi publicado o artigo que continha a solução geral desta problemática. Nesta obra, o inventor, engenheiro eletricista e matemático Rudolf Emil Kalman conceitua e demonstra matematicamente o denominado posteriormente à publicação e em homenagem, filtro de Kalman. Esta teoria foi, originalmente, desenvolvida para sistemas dinâmicos lineares discretos, todavia, devido ao elevado número de aplicabilidades, estudos permitiram estendê-lo para o domínio de tempo contínuo, e até mesmo para sistemas dinâmicos não lineares.

Apesar do quantitativo de trabalhos publicados sobre aplicações do filtro de Kalman ser vasto e permear diversas áreas da ciência, muitos autores utilizam esta ferramenta sem embasamento matemático ou provocam lacunas teóricas que dificultam a aprendizagem dos leitores. Este fato é verificado em He e Shiyi (2021), Jinah, Jeseon e Kideok (2020) e Singh et al. (2020). Dessa forma, o presente artigo objetiva apresentar e explicar detalhadamente o conceito, a demonstração matemática e uma aplicação do filtro de Kalman para sistemas dinâmicos lineares discretos. A perspectiva fundamental de escrita é o equilíbrio entre didática e rigor matemático.

2 FILTRO DE KALMAN

Nesta seção apresenta-se a demonstração matemática do filtro de Kalman e o arcabouço de conhecimentos que a fundamenta.

2.1 Arcabouço matemático

Esta subseção abarca definições, corolários, propriedades e teoremas abordados nos livros de Carreira e Pinto (1999), Steinbruch e Winterle (1987) e Montgomery e Ruger (2018), além de demonstrações matemáticas desenvolvidas pelos autores deste artigo. O conhecimento destes enunciados é primordial para compreender matematicamente o filtro de Kalman.

Definição 1: Esperança matemática: Seja x um vetor aleatório de componentes vetoriais $x_1, \dots, x_n$ cujas probabilidades são $p_{x_1}, \dots, p_{x_n}$, respectivamente, $n \in \mathbb{N}$. Define-se a esperança matemática (ou média aritmética ponderada) de x, $E[x]$, como:

$$E[x] \triangleq \sum_{i=1}^{n} x_i p_{x_i}$$

Corolário 1: Se x e y são vetores aleatórios independentes, ou seja, a distribuição de probabilidade de um destes não influência na distribuição de probabilidade do outro, então $E[xy] = E[x]E[y]$.

Demonstração:

Definição 1,

$$E[xy] = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j p_{x_i y_j}$$

Como os vetores são independentes,

$$E[xy] = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j p_{x_i} p_{y_j}$$

Leis de comutatividade e associatividade da multiplicação,

$$E[xy] = \left(\sum_{i=1}^{n} x_i p_{x_i} \right) \left(\sum_{j=1}^{m} y_j p_{y_j} \right)$$

Definição 1,

$$E[xy] = E[x]E[y]$$

$\blacksquare$

Definição 2: Transformação linear: Sejam V e W espaços vetoriais sobre o mesmo corpo K. Define-se a transformação linear $\mathcal{T}$ de V em W como a função $\mathcal{T}: V \rightarrow W$ que preserva o princípio de superposição (aditividade e homogeneidade), ou seja, $\forall x \in V, \forall \gamma \in \mathbb{R}$:

$$\mathcal{T}\left[\sum_{i=1}^{n} \gamma_i x_i \right] = \sum_{i=1}^{n} \gamma_i \mathcal{T}[x_i]$$

Propriedade 1: O operador esperança matemática é uma transformação linear, ou seja, $E[\sum_{i=1}^{n} \gamma_i x_i] = \sum_{i=1}^{n} \gamma_i E[x_i]$.

Demonstração:

Definição 1,

$$E\left[\sum_{i=1}^{n} \gamma_i x_i \right] = \sum_{j=1}^{n} \sum_{i=1}^{n} \gamma_i x_{i_j} \, p_{(\sum_{i=1}^{n} x_i)_j}$$

Expansão dos somatórios,

$$E\left[\sum_{i=1}^{n} \gamma_i x_i\right] = \sum_{a=1}^{n} \dots \sum_{b=1}^{n} \left(\gamma_1 x_{1_a} + \cdots + \gamma_n x_{n_b}\right) p_{x_{1_a} + \cdots + x_{n_b}}$$

Lei de distributividade da multiplicação em relação à adição e lei de associatividade da adição,

$$E\left[\sum_{i=1}^{n} \gamma_i x_i\right] = \sum_{a=1}^{n} \gamma_1 x_{1_a} p_{x_{1_a} + \cdots + x_{n_b}} + \cdots + \sum_{b=1}^{n} \gamma_n x_{n_b} p_{x_{1_a} + \cdots + x_{n_b}}$$

Como cada vetor é unicamente influenciado pela própria distribuição de probabilidades,

$$E\left[\sum_{i=1}^{n} \gamma_i x_i\right] = \sum_{a=1}^{n} \gamma_1 x_{1_a} p_{x_{1_a}} + \cdots + \sum_{b=1}^{n} \gamma_n x_{n_b} p_{x_{n_b}}$$

Como $\gamma_1, \dots, \gamma_n$ são constantes,

$$E\left[\sum_{i=1}^{n} \gamma_i x_i\right] = \gamma_1 \sum_{a=1}^{n} x_{1_a} p_{x_{1_a}} + \cdots + \gamma_n \sum_{b=1}^{n} x_{n_b} p_{x_{n_b}}$$

Definição 1,

$$E\left[\sum_{i=1}^{n} \gamma_i x_i\right] = \gamma_1 E[x_1] + \cdots + \gamma_n E[x_n]$$

Definição 2,

$$E\left[\sum_{i=1}^{n} \gamma_i x_i\right] = \sum_{i=1}^{n} \gamma_i E[x_i]$$

■

Definição 3: Dimensão matricial: Seja X uma matriz aleatória de número l de linhas e c de colunas, $l, c \in \mathbb{N}$. Define-se a dimensão de X, $\dim[X]$, como:

$$\dim[X] \triangleq l \times c$$

Notação estrutural 1: A notação matemática $X_{\dim[X]}$ indica que $\dim[X]$ é a dimensão da matriz X.

Definição 4: Elemento matricial típico: Seja X uma matriz aleatória tal que $\dim[X] = l \times c$. Define-se o elemento matricial típico de X, $x_{i,j}$, como o elemento localizado na posição de linha $i \in \{1, \dots, l\}$ e coluna $j \in \{1, \dots, c\}$ da matriz X.

Notação estrutural 2: A notação matemática $X = \left[x_{i,j}\right]$ indica que $x_{i,j}$ é o elemento típico da matriz X.

Definição 5: Adição matricial: Sejam X e Y matrizes aleatórias tais que $X_{l\times c} =$

$[x_{i,j}]$, $Y_{l\times c} = [y_{i,j}]$, $i \in \{1, \dots, l\}$ e $j \in \{1, \dots, c\}$. A adição matricial existe se, e somente se, as dimensões de ambas as matrizes forem iguais. Define-se a adição matricial de X e Y, $X + Y$, como:

$$X + Y \triangleq x_{i,j} + y_{i,j}$$

Definição 6: Multiplicação matricial: Sejam X e Y matrizes aleatórias tais que $X_{l\times n} = [x_{i,j}]$, $Y_{n\times c} = [y_{i,j}]$, $i \in \{1, \dots, l\}$ e $j \in \{1, \dots, c\}$. A multiplicação matricial existe se, e somente se, o número de colunas do multiplicando for igual ao número de linhas do multiplicador. Define-se a multiplicação matricial de X e Y, XY, como:

$$XY \triangleq \sum_{k=1}^{n} x_{i,k}\, y_{k,j}$$

Definição 7: Matriz transposta: Seja X uma matriz aleatória cuja dimensão é igual a $l \times c$, ou $\dim[X] = l \times c$. Define-se a matriz transposta de X, X^T, como a matriz cuja dimensão é igual a $c \times l$, ou $\dim[X^T] = c \times l$.

Propriedade 2: Se $X_{l\times c} = [x_{i,j}]$ e $Y_{l\times c} = [y_{i,j}]$, então $(X + Y)^T = X^T + Y^T$.

Demonstração:

Definição 5,

$$(X + Y)^T = \left(x_{i,j} + y_{i,j}\right)^T$$

Satisfeita a condição para adição matricial,

$$(X + Y)^T = (x + y)_{i,j}^T$$

Definição 7,

$$(X + Y)^T = (x + y)_{j,i}$$

Definição 5,

$$(X + Y)^T = x_{j,i} + y_{j,i}$$

Definição 7,

$$(X + Y)^T = x_{i,j}^T + y_{i,j}^T$$

Definição 4,

$$(X + Y)^T = X^T + Y^T$$

■

Propriedade 3: Se $X_{l\times n} = [x_{i,j}]$ e $Y_{n\times c} = [y_{i,j}]$, então $(XY)^T = Y^T X^T$.

Demonstração:

Definição 6,

$$(XY)^T = \left(\sum_{k=1}^{n} x_{i,k} y_{k,j}\right)^T$$

Satisfeita a condição para multiplicação matricial,

$$(XY)^T = (xy)_{i,j}^T$$

Definição 7,

$$(XY)^T = (xy)_{j,i}$$

Definição 6,

$$(XY)^T = \sum_{k=1}^{n} x_{j,k} y_{k,i}$$

Definição 7,

$$(XY)^T = (y^T x^T)_{i,j}$$

Definição 4,

$$(XY)^T = Y^T X^T$$

■

Propriedade 4: Se x é um vetor aleatório, então $E^T[x] = E[x^T]$.

Demonstração:

Definição 1,

$$E^T[x] = \left(\sum_{i=1}^{n} x_i p_{x_i}\right)^T$$

Propriedade 1,

$$E^T[x] = \sum_{i=1}^{n} \left(x_i p_{x_i}\right)^T$$

Propriedade 3,

$$E^T[x] = \sum_{i=1}^{n} p_{x_i}^T x_i^T$$

Como p_{x_i} é um escalar ($\dim[p_{x_i}] = 1 \times 1$),

$$E^T[x] = \sum_{i=1}^{n} p_{x_i} x_i^T$$

Lei de comutatividade da multiplicação,

$$E^T[x] = \sum_{i=1}^{n} x_i^T p_{x_i}$$

Definição 1,

$$E^T[x] = E[x^T]$$

■

Propriedade 5: Se $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, então $E[x] = \begin{bmatrix} E[x_1] \\ \vdots \\ E[x_n] \end{bmatrix}$.

Demonstração:

Como $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$,

$$E[x] = E\left[\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right]$$

Definição 5,

$$E[x] = E\left[\begin{bmatrix} x_1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vdots \\ x_n \end{bmatrix}\right]$$

Propriedade 1,

$$E[x] = E\left[\begin{bmatrix} x_1 \\ \vdots \\ 0 \end{bmatrix}\right] + \cdots + E\left[\begin{bmatrix} 0 \\ \vdots \\ x_n \end{bmatrix}\right]$$

Definição 1,

$$E[x] = \begin{bmatrix} E[x_1] \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vdots \\ E[x_n] \end{bmatrix}$$

Definição 5,

$$E[x] = \begin{bmatrix} E[x_1] \\ \vdots \\ E[x_n] \end{bmatrix}$$

■

Definição 8: Variância: Seja x um vetor aleatório de componentes vetoriais $x_1, \dots, x_n$, $n \in \mathbb{N}$. A variância de x, $var[x]$, é definida como:

$$var[x] \triangleq E[(x - E[x])(x - E[x])^T]$$

Corolário 2: Se x_i, $i \in \{1, \dots, n\}$, é uma das variáveis aleatórias do vetor x, então $var[x_i] = E[(x_i - E[x_i])^2]$.

Demonstração:

Definição 9,

$$var[x_i] = E[(x_i - E[x_i])(x_i - E[x_i])^T]$$

Para uma variável aleatória, $(x_i - E[x_i])^T = (x_i - E[x_i])$,

$$var[x_i] = E[(x_i - E[x_i])^2]$$

■

Definição 9: Desvio padrão: Seja x um vetor aleatório de componentes vetoriais $x_1, \dots, x_n$, $n \in \mathbb{N}$. Define-se o desvio padrão de x, σ_x, como:

$$\sigma_x \triangleq \sqrt{E[(x - E[x])(x - E[x])^T]}$$

Corolário 3: Se x_i, $i \in \{1, \dots, n\}$, é uma das variáveis aleatórias do vetor x, então $\sigma_{x_i} = \sqrt{E[(x_i - E[x_i])^2]}$.

Demonstração:

Definição 9,

$$\sigma_{x_i} = \sqrt{E[(x_i - E[x_i])(x_i - E[x_i])^T]}$$

Para uma variável aleatória, $(x_i - E[x_i])^T = (x_i - E[x_i])$,

$$\sigma_{x_i} = \sqrt{E[(x_i - E[x_i])^2]}$$

■

Definição 10: Covariância: Sejam x e y vetores aleatórios. Define-se a covariância de x e y, $cov[x, y]$, como:

$$cov[x, y] \triangleq E[(x - E[x])(y - E[y])^T]$$

Corolário 4: Se x e y são vetores aleatórios, então $cov[x, y] = E[xy^T] - E[x]E[y^T]$.

Demonstração:

Definição 10,

$$cov[x, y] = E[(x - E[x])(y - E[y])^T]$$

Propriedade 2,

$$cov[x, y] = E[(x - E[x])(y^T - E^T[y])]$$

Lei de distributividade da multiplicação em relação à adição,

$$cov[x, y] = E[xy^T - xE^T[y] - E[x]y^T + E[x]E^T[y]]$$

Propriedade 4,

$$cov[x, y] = E\left[xy^T - xE[y^T] - E[x]y^T + E[x]E[y^T]\right]$$

Propriedade 1,

$$cov[x, y] = E[xy^T] - E[x]E[y^T] - E[x]E[y^T] + E[x]E[y^T]$$

Lei do corte da adição,

$$cov[x, y] = E[xy^T] - E[x]E[y^T]$$

■

Corolário 5: Se a esperança matemática do vetor aleatório x ou y é nula, então $cov[x, y] = E[xy^T]$.

Demonstração:

Corolário 4,

$$cov[x, y] = E[xy^T] - E[x]E[y^T]$$

Lei de anulamento do produto ($E[x] = 0 \vee E[y] = 0 \Leftrightarrow E[x]E[y] = 0$),

$$cov[x, y] = E[xy^T]$$

■

Corolário 6: Se os vetores aleatórios x e y são independentes, então $cov[x, y] = 0$.

Demonstração:

Corolário 4,

$$cov[x, y] = E[xy^T] - E[x]E[y^T]$$

Corolário 1,

$$cov[x, y] = E[x]E[y^T] - E[x]E[y^T]$$

Lei do corte da adição,

$$cov[x, y] = 0$$

■

Corolário 7: Se x_i, $i \in \{1, \dots, n\}$, é uma das variáveis aleatórias do vetor x, então $cov[x_i, x_i] = E[(x_i - E[x_i])^2]$.

Demonstração:

Definição 10,

$$cov[x_i, x_i] = E[(x_i - E[x_i])(x_i - E[x_i])^T]$$

Para uma variável aleatória, $(x_i - E[x_i])^T = (x_i - E[x_i])$,

$$cov[x_i, x_i] = E[(x_i - E[x_i])^2]$$

■

Corolário 8: Se x_i, $i \in \{1, \dots, n\}$, é uma das variáveis aleatórias do vetor x, então $cov[x_i, x_i] = var[x_i] = \sigma_{x_i}^2$.

Demonstração:

Corolário 7,

$$cov[x_i, x_i] = E[(x_i - E[x_i])^2]$$

Corolário 2,

$$cov[x_i, x_i] = var[x_i]$$

Corolário 3,

$$cov[x_i, x_i] = var[x_i] = \sigma_{x_i}^2$$

■

Definição 11: Distribuição gaussiana: A distribuição gaussiana, ou distribuição normal, é uma distribuição de probabilidade contínua e parametrizada através da esperança matemática e da covariância.

Notação estrutural 3: A notação matemática $x \sim N(E[x], cov[x, x])$ indica que um vetor aleatório x de esperança matemática $E[x]$ e covariância $cov[x, x]$, segue uma distribuição gaussiana.

Definição 12: Traço matricial: Seja X uma matriz aleatória quadrada tal que $\dim[X] = n \times n$. Define-se o traço matricial X, $Tr[X]$, como:

$$Tr[X] \triangleq \sum_{i=1}^{n} x_{i,i}$$

Propriedade 6: Se X é uma matriz aleatória e $F(X)$ é uma função matricial diferenciável em cada elemento de X cuja função primitiva é $f(X)$, então $\frac{\partial Tr[F(X)]}{\partial X} = f^T(X)$.

Demonstração:

Definição 12,

$$\frac{\partial Tr[F(X)]}{\partial X} = \frac{\partial}{\partial X}\left(\sum_{i=1}^{n} F(X)_{i,i}\right)$$

Formato matricial,

$$\frac{\partial Tr[F(X)]}{\partial X} = \begin{bmatrix} \frac{\partial \sum_{i=1}^{n} F(X)_{i,i}}{\partial X_{1,1}} & \cdots & \frac{\partial \sum_{i=1}^{n} F(X)_{i,i}}{\partial X_{1,n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \sum_{i=1}^{n} F(X)_{i,i}}{\partial X_{n,1}} & \cdots & \frac{\partial \sum_{i=1}^{n} F(X)_{i,i}}{\partial X_{n,n}} \end{bmatrix}$$

Executa-se as diferenciações,

$$\frac{\partial Tr[F(X)]}{\partial X} = \begin{bmatrix} f(X)_{1,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(X)_{n,n} \end{bmatrix}$$

Pela notação de denominador de diferenciação matricial,

$$\frac{\partial Tr[F(X)]}{\partial X} = f^T(X)$$

■

Propriedade 7: Se X e Y são matrizes aleatórias, então $\frac{\partial Tr[XYX^T]}{\partial X} = X(Y + Y^T)$.

Demonstração:

Definição 12,

$$\frac{\partial Tr[XYX^T]}{\partial X} = \frac{\partial}{\partial X}\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} x_{i,j}y_{j,k}x_{i,k}\right)$$

O operador traço matricial e o operador diferencial são transformações lineares, destarte, analogamente à Propriedade 1,

$$\frac{\partial Tr[XYX^T]}{\partial X} = \frac{\partial}{\partial x_{i,j}}\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} x_{i,j}y_{j,k}x_{i,k}\right) + \frac{\partial}{\partial x_{i,k}}\left(\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} x_{i,j}y_{j,k}x_{i,k}\right)$$

Propriedade 6 (diferenciação),

$$\frac{\partial Tr[XYX^T]}{\partial X} = \sum_{k=1}^{c} y_{j,k} x_{i,k} + \sum_{j=1}^{b} x_{i,j} y_{j,k}$$

Propriedade 6 (transposta da função primitiva) e Definição 6,

$$\frac{\partial Tr[XYX^T]}{\partial X} = (YX^T)^T + (Y^T X^T)^T$$

Propriedade 3,

$$\frac{\partial Tr[XYX^T]}{\partial X} = XY^T + XY$$

Fatora-se, à esquerda, X,

$$\frac{\partial Tr[XYX^T]}{\partial X} = X(Y + Y^T)$$

■

Teorema 1: Matriz de covariância: Seja $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $x \sim N(E[x] = 0, cov[x,x])$, um vetor aleatório de componentes vetoriais independentes, $n \in \mathbb{N}$. A matriz de covariância de x, $cov[x,x]$, é igual a $\begin{bmatrix} \sigma_{x_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{x_n}^2 \end{bmatrix}$.

Demonstração:

Corolário 5,

$$cov[x,x] = E[xx^T]$$

Como $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$,

$$cov[x,x] = E\left[\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ n_n \end{bmatrix}^T\right]$$

Visto que $x_i = \sigma_{x_i} \xi_{x_i}$, $\xi_{x_i} \sim N\left(E\left[\xi_{x_i}\right] = 0, cov\left[\xi_{x_i}, \xi_{x_i}\right] = 1\right)$, $i \in [1, \dots, n]$,

$$cov[x,x] = E\left[\begin{bmatrix} \sigma_{x_1}\xi_{x_1} \\ \vdots \\ \sigma_{x_n}\xi_{x_n} \end{bmatrix} \begin{bmatrix} \sigma_{x_1}\xi_{x_1} \\ \vdots \\ \sigma_{x_n}\xi_{x_n} \end{bmatrix}^T\right]$$

Definição 7,

$$cov[x,x] = E\left[\begin{bmatrix}\sigma_{x_1}\xi_{x_1}\\ \vdots \\ \sigma_{x_n}\xi_{x_n}\end{bmatrix}\begin{bmatrix}\left(\sigma_{x_1}\xi_{x_1}\right)^T & \cdots & \left(\sigma_{x_n}\xi_{x_n}\right)^T\end{bmatrix}\right]$$

1 Propriedade 3,

$$cov[x,x] = E\left[\begin{bmatrix}\sigma_{x_1}\xi_{x_1}\\ \vdots \\ \sigma_{x_n}\xi_{x_n}\end{bmatrix}\begin{bmatrix}\xi_{x_1}^T\sigma_{x_1}^T & \cdots & \xi_{x_n}^T\sigma_{x_n}^T\end{bmatrix}\right]$$

2 Como σ_{x_i} é um escalar ($\dim\left[\sigma_{x_i}\right] = 1 \times 1$),

$$cov[x,x] = E\left[\begin{bmatrix}\sigma_{x_1}\xi_{x_1}\\ \vdots \\ \sigma_{x_n}\xi_{x_n}\end{bmatrix}\begin{bmatrix}\xi_{x_1}^T\sigma_{x_1} & \cdots & \xi_{x_n}^T\sigma_{x_n}\end{bmatrix}\right]$$

3 Definição 6,

$$cov[x,x] = E\left[\begin{bmatrix}\sigma_{x_1}\xi_{x_1}\xi_{x_1}^T\sigma_{x_1} & \cdots & \sigma_{x_1}\xi_{x_1}\xi_{x_n}^T\sigma_{x_n}\\ \vdots & \ddots & \vdots \\ \sigma_{x_n}\xi_{x_n}\xi_{x_1}^T\sigma_{x_1} & \cdots & \sigma_{x_n}\xi_{x_n}\xi_{x_n}^T\sigma_{x_n}\end{bmatrix}\right]$$

4 Leis de comutatividade e associatividade da multiplicação,

$$cov[x,x] = E\left[\begin{bmatrix}\sigma_{x_1}^2\left(\xi_{x_1}\xi_{x_1}^T\right) & \cdots & \sigma_{x_1}\sigma_{x_n}\left(\xi_{x_1}\xi_{x_n}^T\right)\\ \vdots & \ddots & \vdots \\ \sigma_{x_n}\sigma_{x_1}\left(\xi_{x_n}\xi_{x_1}^T\right) & \cdots & \sigma_{x_n}^2\left(\xi_{x_n}\xi_{x_n}^T\right)\end{bmatrix}\right]$$

5 Propriedade 5 e Propriedade 1,

$$cov[x,x] = \begin{bmatrix}\sigma_{x_1}^2 E\left[\xi_{x_1}\xi_{x_1}^T\right] & \cdots & \sigma_{x_1}\sigma_{x_n}E\left[\xi_{x_1}\xi_{x_n}^T\right]\\ \vdots & \ddots & \vdots \\ \sigma_{x_n}\sigma_{x_1}E\left[\xi_{x_n}\xi_{x_1}^T\right] & \cdots & \sigma_{x_n}^2 E\left[\xi_{x_n}\xi_{x_n}^T\right]\end{bmatrix}$$

6 Corolário 5,

$$cov[x,x] = \begin{bmatrix}\sigma_{x_1}^2 cov\left[\xi_{x_1},\xi_{x_1}\right] & \cdots & \sigma_{x_1}\sigma_{x_n}cov\left[\xi_{x_1},\xi_{x_n}\right]\\ \vdots & \ddots & \vdots \\ \sigma_{x_n}\sigma_{x_1}cov\left[\xi_{x_n},\xi_{x_1}\right] & \cdots & \sigma_{x_n}^2 cov\left[\xi_{x_n},\xi_{x_n}\right]\end{bmatrix}$$

7 Corolário 6 e Corolário 8,

$$cov[x,x] = \begin{bmatrix}\sigma_{x_1}^2 & \cdots & 0\\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{x_n}^2\end{bmatrix}$$

■

Teorema 2: Teorema central do limite: Seja $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $x \sim N(E[x], cov[x,x])$, um vetor aleatório de componentes vetoriais independentes e identicamente distribuídas, $S_n = \sum_{i=1}^{n} x_i$ e $z_n = \frac{S_n - E[x]}{\sigma_x \sqrt{n}}$, $n \in \mathbb{N}$. A variável aleatória z_n converge para uma distribuição normal, $z_n \sim N(E[z_n] = 0, cov[z_n, z_n] = 1)$, quando n tende ao infinito.

Demonstração:

Parte I: $E[Z_n] = 0$:

Como $z_n = \frac{S_n - E[x]}{\sigma_x \sqrt{n}}$,

$$E[z_n] = E\left[\frac{S_n - E[x]}{\sigma_x \sqrt{n}}\right]$$

Como $S_n = \sum_{i=1}^{n} x_i$

$$E[z_n] = E\left[\frac{\left(\sum_{i=1}^{n} x_i\right) - E[x]}{\sigma_x \sqrt{n}}\right]$$

Propriedade 1,

$$E[z_n] = \frac{\left(\sum_{i=1}^{n} E[x_i]\right) - E[x]}{\sigma_x \sqrt{n}}$$

Corolário 1,

$$E[z_n] = \frac{E[x] - E[x]}{\sigma_x \sqrt{n}}$$

Lei do corte da adição,

$$E[z_n] = 0$$

Parte II: $cov[z_n, z_n] = 1$:

Corolário 7,

$$cov[z_n, z_n] = E[(z_n - E[z_n])^2]$$

Como $E[z_n] = 0$,

$$cov[z_n, z_n] = E[(z_n)^2]$$

Como $z_n = \frac{S_n - E[x]}{\sigma_x \sqrt{n}}$,

$$cov[z_n, z_n] = E\left[\left(\frac{S_n - E[x]}{\sigma_x\sqrt{n}}\right)^2\right]$$

1 Propriedade 1,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}E[(S_n - E[x])^2]$$

2 Como $S_n - E[x] = \sum_{i=1}^{n}(x_i - E[x_i])$,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}E\left[\left(\sum_{i=1}^{n}(x_i - E[x_i])\right)^2\right]$$

3 Desenvolvimento do quadrado do somatório,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}E[(x_1 - E[x_1])^2 + 2(x_1 - E[x_1])(x_2 - E[x_2]) + \cdots + 2(x_{n-1} - E[x_{n-1}])(x_n - E[x_n]) + (x_n - E[x_n])^2]$$

4 Propriedade 1,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}(E[(x_1 - E[x_1])^2] + 2E[(x_1 - E[x_1])(x_2 - E[x_2])] + \cdots + 2E[(x_{n-1} - E[x_{n-1}])(x_n - E[x_n])] + E[(x_n - E[x_n])^2])$$

5 Para uma variável aleatória, $(x_i - E[x_i])^T = (x_i - E[x_i])$,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}(E[(x_1 - E[x_1])^2] + 2E[(x_1 - E[x_1])(x_2 - E[x_2])^T] + \cdots + 2E[(x_{n-1} - E[x_{n-1}])(x_n - E[x_n])^T] + E[(x_n - E[x_n])^2])$$

6 Corolário 2 e Corolário 7,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}(var[x_1] + 2cov[x_1, x_2] + \cdots + 2cov[x_{n-1}, x_n] + var[x_n])$$

7 Corolário 6 e Corolário 8,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}\left(\sigma_{x_1}^2 + \cdots + \sigma_{x_n}^2\right)$$

8 Como $\sigma_x^2 n = \sigma_{x_1}^2 + \cdots + \sigma_{x_n}^2$,

$$cov[z_n, z_n] = \frac{1}{\sigma_x^2 n}\sigma_x^2 n$$

9 Lei do corte da multiplicação,

$$cov[z_n, z_n] = 1$$

1 **Parte III:** $\phi_{z_n}(t) \sim N(E[z_n] = 0, cov[z_n, z_n] = 1)$:

2 Função característica de argumento t,

$$\phi_{z_n}(t) = E[\exp(jtz_n)]$$

3 Como $z_n = \frac{S_n - E[x]}{\sigma_x \sqrt{n}}$,

$$\phi_{z_n}(t) = E\left[\exp\left(jt\frac{S_n - E[x]}{\sigma_x\sqrt{n}}\right)\right]$$

4 Como $S_n - E[x] = \sum_{i=1}^{n}(x_i - E[x_i])$,

$$\phi_{z_n}(t) = E\left[\exp\left(\frac{jt}{\sigma_x\sqrt{n}}\sum_{i=1}^{n}(x_i - E[x_i])\right)\right]$$

5 O exponencial de um somatório é igual ao produtório dos exponenciais,

$$\phi_{z_n}(t) = E\left[\prod_{i=1}^{n}\exp\left(\frac{jt}{\sigma_x\sqrt{n}}(x_i - E[x_i])\right)\right]$$

6 As componentes vetoriais de x são independentes,

$$\phi_{z_n}(t) = \prod_{i=1}^{n}E\left[\exp\left(\frac{jt}{\sigma_x\sqrt{n}}(x_i - E[x_i])\right)\right]$$

7 As componentes vetoriais de x são identicamente distribuídas,

$$\phi_{z_n}(t) = \left(E\left[\exp\left(\frac{jt}{\sigma_x\sqrt{n}}(x - E[x])\right)\right]\right)^n$$

8 Série de Taylor da exponencial centrada em zero,

$$\phi_{z_n}(t) = \left(E\left[1 + \frac{jt}{\sigma_x\sqrt{n}}(x - E[x]) + \frac{1}{2}\left(\frac{jt}{\sigma_x\sqrt{n}}\right)^2\left((x - E[x])\right)^2 + \cdots\right]\right)^n$$

9 Truncamento nas três primeiras parcelas,

$$\phi_{z_n}(t) = \left(E\left[1 + \frac{jt}{\sigma_x\sqrt{n}}(x - E[x]) - \frac{t^2}{2\sigma_x^2 n}(x - E[x])^2\right]\right)^n$$

10 Propriedade 1,

$$\phi_{z_n}(t) = \left(1 + \frac{jt}{\sigma_x\sqrt{n}}(E[x] - E[x]) - \frac{t^2}{2\sigma_x^2 n}E[(x - E[x])^2]\right)^n$$

Lei do corte da adição, Corolário 7 e Corolário 8,

$$\phi_{z_n}(t) = \left(1 - \frac{t^2}{2\sigma_x^2 n}\sigma_x^2\right)^n$$

Lei do corte da multiplicação,

$$\phi_{z_n}(t) = \left(1 - \frac{t^2}{2n}\right)^n$$

Limite da função característica quando n tende ao infinito,

$$\lim_{n\to\infty} \phi_{z_n}(t) = \lim_{n\to\infty}\left(1 - \frac{t^2}{2n}\right)^n$$

Limite exponencial fundamental,

$$\lim_{n\to\infty} \phi_{z_n}(t) = e^{-\frac{t^2}{2}}$$

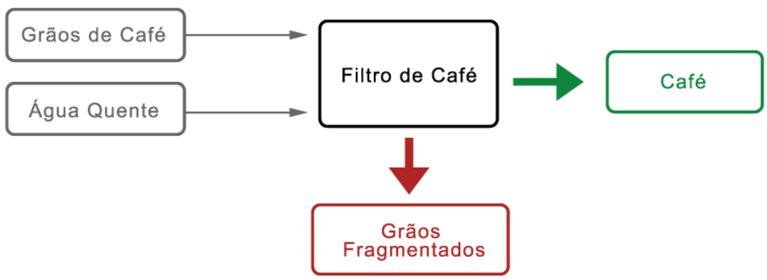
Como $E[z_n] = 0$, $cov[z_n, z_n] = 1$ e $e^{-\frac{t^2}{2}}$ é a função característica de uma distribuição normal padrão, demonstra-se que $z_n \sim N(E[z_n] = 0, cov[z_n, z_n] = 1)$.

■

2.2 Conceito e demonstração matemática

Esta subseção abrange a conceituação do filtro de Kalman, escrita primeiramente através de uma abordagem didática e posteriormente pelo formalismo teórico. Outrossim, apresenta e explica detalhadamente a demonstração matemática do filtro de Kalman.

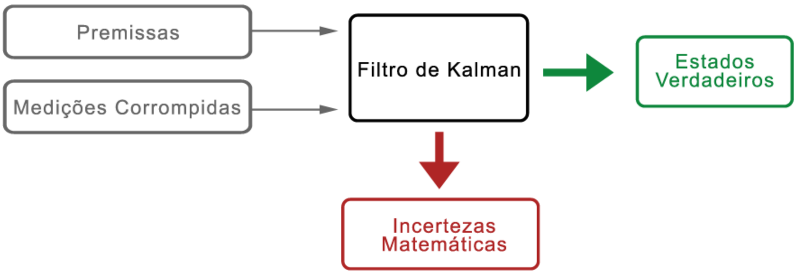
Figura 1: Fluxograma simplificado de sistema dinâmico com filtro de café



Fonte: Autoria própria (2021)

Na teoria de controle de sistemas, ruído significa uma incerteza sistemática ou aleatória, e filtragem é um método utilizado para reduzir a intensidade dos ruídos de um sistema. Didaticamente, pretende-se, apresentar o funcionamento do filtro de Kalman simplificadamente para facilitar a compreensão deste. Portanto, através de fluxogramas, compara-se um filtro de café ao filtro de Kalman. No primeiro, Figura 1, os grãos de café e a água quente são as entradas do sistema, o filtro de café impede a passagem de grãos fragmentados, e o café é o produto desejado. No segundo, Figura 2, as premissas e as medições corrompidas são as entradas do sistema, o filtro de Kalman impede a passagem de incertezas matemáticas, e os estados verdadeiros são as informações desejadas.

Figura 2: Fluxograma simplificado de sistema dinâmico com filtro de Kalman



Fonte: Autoria própria (2021)

Formalmente, o filtro de Kalman é um algoritmo matemático de otimização de estimações que objetiva a partir de premissas e medições corrompidas, minimizar incertezas matemáticas e predizer o valor verdadeiro dos estados de um sistema dinâmico (QIANG *et al*., 2015). Dentre as aplicações deste filtro, algumas ocorrem em situações nas quais é necessário estimar um parâmetro indiretamente, exemplificadas pela constatação da temperatura dos propulsores de um foguete, porque, se aferida diretamente, o dispositivo de medição seria danificado (JIHYOUNG *et al*., 2019). Outras aplicações figuram em cenários nos quais é necessário fundir uma coleção de medições, pois, a informação requerida se trata de uma estimação combinada, exemplificados pela determinação da localização de um automóvel através do sistema de posicionamento global (RAHEMI; MOSAVI, 2021).

O estudo de um sistema real cujo comportamento evidencia as características de um sistema dinâmico linear discreto, inicia-se pela criação de um modelo por intermédio das equações de espaço de estados discretos. Logo,

$$x_k = A_k x_{k-1} + B_k u_k \tag{1}$$

$$y_k = C_k x_k + D_k u_k \tag{2}$$

Ou, matricialmente,

$$\begin{bmatrix} x_{1_k} \\ \vdots \\ x_{n_k} \end{bmatrix}_{n\times 1} = \begin{bmatrix} A_{1,1_k} & \cdots & A_{1,n_k} \\ \vdots & \ddots & \vdots \\ A_{n,1_k} & \cdots & A_{n,n_k} \end{bmatrix}_{n\times n} \begin{bmatrix} x_{1_{k-1}} \\ \vdots \\ x_{n_{k-1}} \end{bmatrix}_{n\times 1} + \begin{bmatrix} B_{1,1_k} & \cdots & B_{1,p_k} \\ \vdots & \ddots & \vdots \\ B_{n,1_k} & \cdots & B_{n,p_k} \end{bmatrix}_{n\times p} \begin{bmatrix} u_{1_k} \\ \vdots \\ u_{p_k} \end{bmatrix}_{p\times 1}$$

$$\begin{bmatrix} y_{1_k} \\ \vdots \\ y_{q_k} \end{bmatrix}_{q\times 1} = \begin{bmatrix} C_{1,1_k} & \cdots & C_{1,n_k} \\ \vdots & \ddots & \vdots \\ C_{q,1_k} & \cdots & C_{q,n_k} \end{bmatrix}_{q\times n} \begin{bmatrix} x_{1_k} \\ \vdots \\ x_{n_k} \end{bmatrix}_{n\times 1} + \begin{bmatrix} D_{1,1_k} & \cdots & D_{1,p_k} \\ \vdots & \ddots & \vdots \\ D_{q,1_k} & \cdots & D_{q,p_k} \end{bmatrix}_{q\times p} \begin{bmatrix} u_{1_k} \\ \vdots \\ u_{p_k} \end{bmatrix}_{p\times 1}$$

Neste formato, as equações 1 e 2 representam o estado e a resposta do sistema modelado determinístico, respectivamente. Onde x é o vetor de estado ($x \in \mathbb{R}^n$), y é o vetor de resposta ($y \in \mathbb{R}^q$), u é o vetor de controle ($u \in \mathbb{R}^p$), A é a matriz de estado ($\dim[A] = n \times n$), B é a matriz de entrada ($\dim[B] = n \times p$), C é a matriz de saída ($\dim[C] = q \times n$), D é a matriz de alimentação ($\dim[D] = q \times p$) e k é a variável temporal discreta ($k \in \mathbb{N}$).

Entretanto, este modelo matemático não considera parâmetros estocásticos inerentes ao sistema real. Pois, incertezas matemáticas estão presentes em cada um dos estados que se deseja conhecer e nas medições executadas para aferi-los. A equação de estado expressa o comportamento dinâmico do sistema modelado e possui imprecisões inatas que são dirimidas unicamente através de reconstrução material. Alternativamente, a equação de medição revela a resposta do sistema modelado devido ao comportamento dinâmico, afinal, trata-se da representação matemática de um sensor ou equipamento similar que fornece uma medida numérica a partir de uma observação, e possui imprecisões que corrompem as medições e, consequentemente, impedem a coleta do valor verdadeiro dos estados do sistema modelado. As incertezas da equação de medição são minimizadas matematicamente pelo filtro de Kalman.

Portanto, adicionam-se vetores aleatórios discretos independentes de esperança matemática nula e covariância não nula que representam as incertezas matemáticas do sistema real, $\omega \sim N(E[\omega] = 0, Q = E[\omega\omega^T])$ e $\vartheta \sim N(E[\vartheta] = 0, R = E[\vartheta\vartheta^T])$, ou seja, ruídos aleatórios, às equações de estado e medição do sistema modelado idealmente, respectivamente (WELCH; BISHOP, 2006). Estes vetores, conforme o Teorema 2, seguem uma distribuição gaussiana. Ademais, considera-se para este processo de otimização que inexista uma fonte de alimentação do sistema, isto significa que D é uma matriz nula. Logo,

$$x_k = A_k x_{k-1} + B_k u_k + \omega_k \tag{3}$$

$$y_k = C_k x_k + \vartheta_k \tag{4}$$

Ou, matricialmente,

$$\begin{bmatrix} x_{1_k} \\ \vdots \\ x_{n_k} \end{bmatrix}_{n\times 1} = \begin{bmatrix} A_{1,1_k} & \cdots & A_{1,n_k} \\ \vdots & \ddots & \vdots \\ A_{n,1_k} & \cdots & A_{n,n_k} \end{bmatrix}_{n\times n} \begin{bmatrix} x_{1_{k-1}} \\ \vdots \\ x_{n_{k-1}} \end{bmatrix}_{n\times 1} + \begin{bmatrix} B_{1,1_k} & \cdots & B_{1,p_k} \\ \vdots & \ddots & \vdots \\ B_{n,1_k} & \cdots & B_{n,p_k} \end{bmatrix}_{n\times p} \begin{bmatrix} u_{1_k} \\ \vdots \\ u_{p_k} \end{bmatrix}_{p\times 1} + \begin{bmatrix} \omega_{1_k} \\ \vdots \\ \omega_{n_k} \end{bmatrix}_{n\times 1}$$

$$\begin{bmatrix} y_{1_k} \\ \vdots \\ y_{q_k} \end{bmatrix}_{q\times 1} = \begin{bmatrix} C_{1,1_k} & \cdots & C_{1,n_k} \\ \vdots & \ddots & \vdots \\ C_{q,1_k} & \cdots & C_{q,n_k} \end{bmatrix}_{q\times n} \begin{bmatrix} x_{1_k} \\ \vdots \\ x_{n_k} \end{bmatrix}_{n\times 1} + \begin{bmatrix} \vartheta_{1_k} \\ \vdots \\ \vartheta_{n_k} \end{bmatrix}_{q\times 1}$$

Ou ainda, consoante o Teorema 1,

$$\begin{bmatrix} x_{1_k} \\ \vdots \\ x_{n_k} \end{bmatrix}_{n\times 1} = \begin{bmatrix} A_{1,1_k} & \cdots & A_{1,n_k} \\ \vdots & \ddots & \vdots \\ A_{n,1_k} & \cdots & A_{n,n_k} \end{bmatrix}_{n\times n} \begin{bmatrix} x_{1_{k-1}} \\ \vdots \\ x_{n_{k-1}} \end{bmatrix}_{n\times 1} + \begin{bmatrix} B_{1,1_k} & \cdots & B_{1,p_k} \\ \vdots & \ddots & \vdots \\ B_{n,1_k} & \cdots & B_{n,p_k} \end{bmatrix}_{n\times p} \begin{bmatrix} u_{1_k} \\ \vdots \\ u_{p_k} \end{bmatrix}_{p\times 1} + \sqrt{\begin{bmatrix} \sigma^2_{\omega_{1_k}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2_{\omega_{n_k}} \end{bmatrix}_{n\times n}} \begin{bmatrix} \xi_{\omega_{1_k}} \\ \vdots \\ \xi_{\omega_{n_k}} \end{bmatrix}_{n\times 1}$$

$$\begin{bmatrix} y_{1_k} \\ \vdots \\ y_{q_k} \end{bmatrix}_{q\times 1} = \begin{bmatrix} C_{1,1_k} & \cdots & C_{1,n_k} \\ \vdots & \ddots & \vdots \\ C_{q,1_k} & \cdots & C_{q,n_k} \end{bmatrix}_{q\times n} \begin{bmatrix} x_{1_k} \\ \vdots \\ x_{n_k} \end{bmatrix}_{n\times 1} + \sqrt{\begin{bmatrix} \sigma^2_{\vartheta_{1_k}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2_{\vartheta_{n_k}} \end{bmatrix}_{q\times q}} \begin{bmatrix} \xi_{\vartheta_{1_k}} \\ \vdots \\ \xi_{\vartheta_{n_k}} \end{bmatrix}_{q\times 1}$$

Dessa forma, as equações 3 e 4 representam um sistema modelado estocástico, e fornecem o comportamento dinâmico dos estados do sistema e a resposta desde comportamento, corrompida por incertezas matemáticas, respectivamente. Deseja-se, por conseguinte, analisar a influência destas imprecisões de medição no sistema e minimizá-las. Posto isto, por hipótese, supõe-se que a esperança matemática do sistema estocástico contenha os verdadeiros valores dos estados desejados. A verificação deste raciocínio ocorre por intermédio da aplicação do operador esperança matemática ao sistema em questão. Logo,

$$E[x_k] = E[A_k x_{k-1} + B_k u_k + \omega_k]$$

$$E[y_k] = E[C_k x_k + \vartheta_k]$$

Propriedade 1,

$$E[x_k] = A_k E[x_{k-1}] + B_k E[u_k] + E[\omega_k]$$

$$E[y_k] = C_k E[x_k] + E[\vartheta_k]$$

O vetor de controle, u_k, não é aleatório,

$$E[x_k] = A_k E[x_{k-1}] + B_k u_k + E[\omega_k]$$

$$E[y_k] = C_k E[x_k] + E[\vartheta_k]$$

Como $\omega \sim N(E[\omega] = 0, Q = E[\omega\omega^T])$ e $\vartheta \sim N(E[\vartheta] = 0, R = E[\vartheta\vartheta^T])$,

$$E[x_k] = A_k E[x_{k-1}] + B_k u_k$$

$$E[y_k] = C_k E[x_k]$$

Percebe-se na esperança do sistema estocástico que a distribuição de probabilidade condicional de um estado futuro depende unicamente do estado imediatamente anterior a este, ou seja, preserva-se a propriedade de perda de memória ou propriedade de Markov. Assim, atribui-se a notação de dependência de estado e reescrevem-se as esperanças, $E[x_k] = \hat{x}_{k|k-1}$, $E[x_{k-1}] = \hat{x}_{k-1|k-1}$ e $E[y_k] = \hat{y}_k$, para explicitar que a esperança de estado na etapa temporal k está condicionada à ocorrência da esperança de estado na etapa temporal $k-1$. Logo,

$$\hat{x}_{k|k-1} = A_k \hat{x}_{k-1|k-1} + B_k u_k \tag{5}$$

$$\hat{y}_k = C_k \hat{x}_{k|k-1} \tag{6}$$

Compara-se o sistema estocástico, equações 3 e 4, e a esperança deste, equações 5 e 6. Observa-se que o segundo contém as informações do primeiro, porém, excetuam-se os vetores aleatórios de incertezas. Dessa forma, demonstra-se a conjectura. Esta conclusão é necessária, mas não suficiente para a análise da influência das incertezas matemáticas no sistema. Posto isto, por hipótese, supõe-se que exista uma variação na evolução temporal dos estados do sistema estocástico em relação aos estados da esperança deste. Caso esta suposição seja falsa, a esperança do sistema estocástico informará a esperança do valor verdadeiro dos estados em qualquer etapa temporal, pois não possui vetores de incertezas. A verificação deste raciocínio ocorre por intermédio da criação de um sinal de erro e do cálculo da covariância deste. Portanto, define-se o sinal de erro de estado, obtido mediante a subtração entre a equação de estado e a esperança desta, $\tilde{x}_{k|k-1}$, como:

$$\tilde{x}_{k|k-1} \triangleq x_k - \hat{x}_{k|k-1}$$

Segundo as equações 3 e 5,

$$\tilde{x}_{k|k-1} = (A_k x_{k-1} + B_k u_k + \omega_k) - \left(A_k \hat{x}_{k-1|k-1} + B_k u_k\right)$$

Leis de distributividade da multiplicação em relação à adição, comutatividade da adição e corte da adição,

$$\tilde{x}_{k|k-1} = A_k x_{k-1} - A_k \hat{x}_{k-1|k-1} + \omega_k$$

Fatora-se A_k,

$$\tilde{x}_{k|k-1} = A_k\left(x_{k-1} - \hat{x}_{k-1|k-1}\right) + \omega_k$$

Visto que $\tilde{x}_{k-1|k-1} = x_{k-1} - \hat{x}_{k-1|k-1}$,

$$\tilde{x}_{k|k-1} = A_k \tilde{x}_{k-1|k-1} + \omega_k \tag{7}$$

Após a definição e simplificação deste sinal de erro, deseja-se calcular o quão distante o valor verdadeiro dos estados está em relação à esperança destes durante a evolução temporal. Esta possível disparidade trata-se de uma dispersão estatística relativa à interdependência destas variáveis, ou seja, a covariância do sinal de erro.

Portanto, define-se a covariância do sinal de erro de estado, $P_{\tilde{x}_{k|k-1}}$, como:

$$P_{\tilde{x}_{k|k-1}} \triangleq E\left[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T\right]$$

Segundo a equação 7,

$$P_{\tilde{x}_{k|k-1}} = E\left[(A_k\tilde{x}_{k-1|k-1} + \omega_k)\left(A_k\tilde{x}_{k-1|k-1} + \omega_k\right)^T\right]$$

Propriedade 2,

$$P_{\tilde{x}_{k|k-1}} = E\left[(A_k\tilde{x}_{k-1|k-1} + \omega_k)\left(\left(A_k\tilde{x}_{k-1|k-1}\right)^T + \omega_k^T\right)\right]$$

Propriedade 3,

$$P_{\tilde{x}_{k|k-1}} = E\left[(A_k\tilde{x}_{k-1|k-1} + \omega_k)\left(\tilde{x}_{k-1|k-1}^T A_k^T + \omega_k^T\right)\right]$$

Lei de distributividade da multiplicação em relação à adição,

$$P_{\tilde{x}_{k|k-1}} = E\left[A_k\tilde{x}_{k-1|k-1}\tilde{x}_{k-1|k-1}^T A_k^T + A_k\tilde{x}_{k-1|k-1}\omega_k^T + \omega_k\tilde{x}_{k-1|k-1}^T A_k^T + \omega_k\omega_k^T\right]$$

Propriedade 1,

$$P_{\tilde{x}_{k|k-1}} = A_k E\left[\tilde{x}_{k-1|k-1}\tilde{x}_{k-1|k-1}^T\right]A_k^T + A_k E\left[\tilde{x}_{k-1|k-1}\omega_k^T\right] + E\left[\omega_k\tilde{x}_{k-1|k-1}^T\right]A_k^T + E[\omega_k\omega_k^T]$$

Os vetores $\tilde{x}_{k-1|k-1}$ e ω_k são independentes, destarte, pelo Corolário 1,

$$P_{\tilde{x}_{k|k-1}} = A_k E\left[\tilde{x}_{k-1|k-1}\tilde{x}_{k|k}^T\right]A_k^T + A_k E\left[\tilde{x}_{k-1|k-1}\right]E[\omega_k^T] + E[\omega_k]E\left[\tilde{x}_{k-1|k-1}^T\right]A_k^T + E[\omega_k\omega_k^T]$$

Como $\omega \sim N(E[\omega] = 0, Q = E[\omega\omega^T])$,

$$P_{\tilde{x}_{k|k-1}} = A_k E\left[\tilde{x}_{k-1|k-1}\tilde{x}_{k-1|k-1}^T\right]A_k^T + Q_k$$

Visto que $P_{\tilde{x}_{k-1|k-1}} = E\left[\tilde{x}_{k-1|k-1}\tilde{x}_{k-1|k-1}^T\right]$,

$$P_{\tilde{x}_{k|k-1}} = A_k P_{\tilde{x}_{k-1|k-1}} A_k^T + Q_k \tag{8}$$

A equação 8 revela que a covariância $P_{\tilde{x}_{k|k-1}}$ depende da covariância do sinal de erro de estado $\tilde{x}_{k-1|k-1}$, $P_{\tilde{x}_{k-1|k-1}}$, e da conhecida matriz de estado, A_k. Se estes fossem os únicos termos de dependência, então a covariância do sinal de erro de estado da iteração atual convergiria a zero ($\lim_{k\to\infty} P_{\tilde{x}_{k|k-1}} = 0$) desde que a covariância do sinal de erro de estado da iteração anterior convergisse a zero ($\lim_{k\to\infty} P_{\tilde{x}_{k-1|k-1}} = 0$). Isto implicaria em inexistência de dispersão estatística entre o sistema estocástico e a esperança deste.

Entretanto, a covariância do vetor de incertezas da equação de estado, Q, é um termo aditivo na equação 8. Este fato provoca o incremento de $P_{\tilde{x}_{k|k-1}}$ a cada iteração ($\lim_{k\to\infty} P_{\tilde{x}_{k|k-1}} = \infty$) e, consequentemente, acarreta divergência entre o sistema estocástico e a esperança deste. Dessa forma, demonstra-se a conjectura.

Após a constatação destas duas conjecturas, o problema de minimizar as incertezas matemáticas do sistema ainda existe e teve a dificuldade agravada. Este fato ocorre, pois, apesar de conhecer-se tanto o sistema estocástico quanto a esperança deste, a dispersão estatística que os acomete impede que as informações livres de incertezas matemáticas advindas da esperança do sistema estocástico apresentem o valor verdadeiro dos estados do sistema. A solução para este contexto é denominada filtro de Kalman.

A solução proposta por Kalman foi adicionar uma parcela corretiva, Ψ_k, à atual esperança de estado, $\hat{x}_{k|k-1}$, para obter a esperança de estado corrigida, $\hat{x}_{k|k}$ (KALMAN, 1960). Logo,

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \Psi_k \tag{9}$$

Esta adição faria a atual esperança de estado convergir à esperança corrigida de estado ($\lim_{k\to\infty} \hat{x}_{k|k-1} = \hat{x}_{k|k}$), a qual, consequentemente, convergiria ao valor verdadeiro dos estados do sistema estocástico ($\lim_{k\to\infty} \hat{x}_{k|k} = x_k$). Posto isto, o problema resume-se a calcular a parcela de correção Ψ_k. Percebe-se que as únicas informações disponíveis para que esta convergência se perfaça são fornecidas pela equação de medição e pela esperança desta, pois y_k contém x_k, apesar de corrompida por incertezas matemáticas, e $\hat{y}_k$ contém $\hat{x}_{k|k-1}$. Dessa forma, Ψ_k será proporcional à subtração entre y_k e $\hat{y}_k$, ou seja, ao sinal de erro de medição. O fator de proporcionalidade ou, em homenagem, ganho de Kalman, K_k, será responsável pela atenuação das incertezas matemáticas de medição e por assegurar que a convergência desejada se concretize (KALMAN, 1960). Logo,

$$\Psi_k = K_k(y_k - \hat{y}_k) \tag{10}$$

Conclui-se, a partir das equações 9 e 10, que a esperança corrigida de estado será:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_k) \tag{11}$$

Portanto, define-se o atualizado sinal de erro de estado, obtido mediante a subtração entre a equação de estado e a esperança corrigida desta, $\tilde{x}_{k|k}$, como:

$$\tilde{x}_{k|k} \triangleq x_k - \hat{x}_{k|k}$$

Conforme a equação 11,

$$\tilde{x}_{k|k} = x_k - \left(\hat{x}_{k|k-1} + K_k(y_k - \hat{y}_k)\right)$$

Consoante as equações 4 e 6,

$$\tilde{x}_{k|k} = x_k - \left(\hat{x}_{k|k-1} + K_k\left(C_k x_k + \vartheta_k - C_k\hat{x}_{k|k-1}\right)\right)$$

Lei de distributividade da multiplicação em relação à adição,

$$\tilde{x}_{k|k} = x_k - \hat{x}_{k|k-1} - K_k C_k x_k - K_k\vartheta_k + K_k C_k\hat{x}_{k|k-1}$$

Multiplica-se x_k e $\hat{x}_{k|k-1}$ pela matriz identidade de ordem n, I,

$$\tilde{x}_{k|k} = I x_k - I\hat{x}_{k|k-1} - K_k C_k x_k - K_k\vartheta_k + K_k C_k\hat{x}_{k|k-1}$$

Fatora-se $(I - K_k C_k)$,

$$\tilde{x}_{k|k} = (I - K_k C_k)\left(x_k - \hat{x}_{k|k-1}\right) - K_k\vartheta_k$$

Visto que $\tilde{x}_{k|k-1} = x_k - \hat{x}_{k|k-1}$,

$$\tilde{x}_{k|k} = (I - K_k C_k)\tilde{x}_{k|k-1} - K_k\vartheta_k \qquad (12)$$

Após a definição e simplificação deste sinal de erro, deseja-se avaliar a dispersão estatística atualizada entre o valor verdadeiro dos estados e a esperança corrigida destes. Portanto, define-se a covariância do atualizado sinal de erro de estado, $P_{\tilde{x}_{k|k}}$, como:

$$P_{\tilde{x}_{k|k}} \triangleq E\left[\tilde{x}_{k|k}\tilde{x}_{k|k}^T\right]$$

Segundo a equação 12,

$$P_{\tilde{x}_{k|k}} = E\left[\left((I - K_k C_k)\tilde{x}_{k|k-1} - K_k\vartheta_k\right)\left((I - K_k C_k)\tilde{x}_{k|k-1} - K_k\vartheta_k\right)^T\right]$$

Propriedade 2,

$$P_{\tilde{x}_{k|k}} = E\left[\left((I - K_k C_k)\tilde{x}_{k|k-1} - K_k\vartheta_k\right)\left(\left((I - K_k C_k)\tilde{x}_{k|k-1}\right)^T - (K_k\vartheta_k)^T\right)\right]$$

Propriedade 3,

$$P_{\tilde{x}_{k|k}} = E\left[\left((I - K_k C_k)\tilde{x}_{k|k-1} - K_k\vartheta_k\right)\left(\left(\tilde{x}_{k|k-1}^T(I - K_k C_k)^T\right) - (K_k\vartheta_k)^T\right)\right]$$

Propriedade 2,

$$P_{\tilde{x}_{k|k}} = E\left[\left((I - K_k C_k)\tilde{x}_{k|k-1} - K_k\vartheta_k\right)\left(\left(\tilde{x}_{k|k-1}^T(I^T - (K_k C_k)^T)\right) - (K_k\vartheta_k)^T\right)\right]$$

Propriedade 3 e como $I = I^T$,

$$P_{\tilde{x}_{k|k}} = E\left[\left((I - K_k C_k)\tilde{x}_{k|k-1} - K_k\vartheta_k\right)\left(\tilde{x}_{k|k-1}^T(I - C_k^T K_k^T) - \vartheta_k^T K_k^T\right)\right]$$

Lei de distributividade da multiplicação em relação à adição,

$$P_{\tilde{x}_{k|k}} = E\big[(I - K_k C_k)\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T(I - C_k^T K_k^T) - (I - K_k C_k)\tilde{x}_{k|k-1}\vartheta_k^T K_k^T \\ - K_k\vartheta_k\tilde{x}_{k|k-1}^T(I - C_k^T K_k^T) + K_k\vartheta_k\vartheta_k^T K_k^T\big]$$

Propriedade 1,

$$P_{\tilde{x}_{k|k}} = (I - K_k C_k)E\big[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T\big](I - C_k^T K_k^T) - (I - K_k C_k)E\big[\tilde{x}_{k|k-1}\vartheta_k^T\big]K_k^T \\ - K_k^T E\big[\vartheta_k\tilde{x}_{k|k-1}^T\big](I - C_k^T K_k^T) + K_k E[\vartheta_k\vartheta_k^T]K_k^T$$

Os vetores $\tilde{x}_{k|k-1}$ e ϑ_k são independentes, destarte, pelo Corolário 1,

$$P_{\tilde{x}_{k|k}} = (I - K_k C_k)E\big[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T\big](I - C_k^T K_k^T) - (I - K_k C_k)E\big[\tilde{x}_{k|k-1}\big]E[\vartheta_k^T]K_k^T \\ - K_k^T E[\vartheta_k]E\big[\tilde{x}_{k|k-1}^T\big](I - C_k^T K_k^T) + K_k E[\vartheta_k\vartheta_k^T]K_k^T$$

Como $\vartheta \sim N(E[\vartheta] = 0, R = E[\vartheta\vartheta^T])$,

$$P_{\tilde{x}_{k|k}} = (I - K_k C_k)E\big[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T\big](I - C_k^T K_k^T) + K_k E[\vartheta_k\vartheta_k^T]K_k^T$$

Visto que $P_{\tilde{x}_{k|k-1}} = E\big[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T\big]$,

$$P_{\tilde{x}_{k|k}} = (I - K_k C_k)P_{\tilde{x}_{k|k-1}}(I - C_k^T K_k^T) + K_k R_k K_k^T$$

Lei de distributividade da multiplicação em relação à adição,

$$P_{\tilde{x}_{k|k}} = IP_{\tilde{x}_{k|k-1}}I - IP_{\tilde{x}_{k|k-1}}C_k^T K_k^T - K_k C_k P_{\tilde{x}_{k|k-1}}I + K_k C_k P_{\tilde{x}_{k|k-1}}C_k^T K_k^T + K_k R_k K_k^T$$

Como I é o elemento neutro da multiplicação matricial,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - P_{\tilde{x}_{k|k-1}}C_k^T K_k^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k C_k P_{\tilde{x}_{k|k-1}}C_k^T K_k^T + K_k R_k K_k^T$$

Como $P_{\tilde{x}_{k|k-1}}^T = P_{\tilde{x}_{k|k-1}}$ (vide Teorema 1),

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - P_{\tilde{x}_{k|k-1}}^T C_k^T K_k^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k C_k P_{\tilde{x}_{k|k-1}}C_k^T K_k^T + K_k R_k K_k^T$$

Fatora-se K_k e K_k^T nas duas últimas parcelas,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - P_{\tilde{x}_{k|k-1}}^T C_k^T K_k^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k\left(C_k P_{\tilde{x}_{k|k-1}}C_k^T + R_k\right)K_k^T$$

Propriedade 3,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - \left(K_k C_k P_{\tilde{x}_{k|k-1}}\right)^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k\left(C_k P_{\tilde{x}_{k|k-1}}C_k^T + R_k\right)K_k^T \qquad (13)$$

Percebe-se, em conformidade com o Teorema 1, que a matriz de covariância do atualizado sinal de erro, $P_{\tilde{x}_{k|k}}$, equação 13, é uma matriz diagonal cujos elementos da diagonal principal são os erros quadráticos médios entre os correspondentes elementos dos vetores x_k e $\hat{x}_{k|k}$, ou seja, uma função matricial quadrática cuja variável é o ganho de Kalman, K_k, o qual é o fator responsável pela minimização desejada. A ferramenta matemática que relaciona estes erros em uma única equação é o traço matricial. Logo, pela Definição 12,

$$Tr\left[P_{\tilde{x}_{k|k}}\right] = Tr\left[P_{\tilde{x}_{k|k-1}} - \left(K_k C_k P_{\tilde{x}_{k|k-1}}\right)^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) K_k^T\right]$$

O operador traço matricial é uma transformação linear, dessarte, analogamente à Propriedade 1,

$$Tr\left[P_{\tilde{x}_{k|k}}\right] = Tr\left[P_{\tilde{x}_{k|k-1}}\right] - Tr\left[\left(K_k C_k P_{\tilde{x}_{k|k-1}}\right)^T\right] - Tr\left[K_k C_k P_{\tilde{x}_{k|k-1}}\right] + Tr\left[K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) K_k^T\right]$$

Como $Tr\left[\left(K_k C_k P_{\tilde{x}_{k|k-1}}\right)\right] = Tr\left[\left(K_k C_k P_{\tilde{x}_{k|k-1}}\right)^T\right]$ (vide Definição 12),

$$Tr\left[P_{\tilde{x}_{k|k}}\right] = Tr\left[P_{\tilde{x}_{k|k-1}}\right] - 2Tr\left[K_k C_k P_{\tilde{x}_{k|k-1}}\right] + Tr\left[K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) K_k^T\right] \tag{14}$$

A minimização desejada é obtida por intermédio do método de mínimos quadrados desenvolvido por Carl Friedrich Gauss, em 1795. Este método minimiza a soma dos quadrados de uma função de erro, ou seja, a equação 14. Logo, aplica-se o operador diferencial a $Tr\left[P_{\tilde{x}_{k|k}}\right]$ com respeito a K_k,

$$\frac{\partial\left(Tr\left[P_{\tilde{x}_{k|k}}\right]\right)}{\partial K_k} = \frac{\partial\left(Tr\left[P_{\tilde{x}_{k|k-1}}\right] - 2Tr\left[K_k C_k P_{\tilde{x}_{k|k-1}}\right] + Tr\left[K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) K_k^T\right]\right)}{\partial K_k}$$

O operador diferencial é uma transformação linear, dessarte, analogamente à Propriedade 1,

$$\frac{\partial\left(Tr\left[P_{\tilde{x}_{k|k}}\right]\right)}{\partial K_k} = \frac{\partial\left(Tr\left[P_{\tilde{x}_{k|k-1}}\right]\right)}{\partial K_k} - 2\frac{\partial\left(Tr\left[K_k C_k P_{\tilde{x}_{k|k-1}}\right]\right)}{\partial K_k} + \frac{\partial\left(Tr\left[K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) K_k^T\right]\right)}{\partial K_k}$$

Propriedade 6 e Propriedade 7,

$$\frac{\partial\left(Tr\left[P_{\tilde{x}_{k|k}}\right]\right)}{\partial K_k} = 0 - 2\left(C_k P_{\tilde{x}_{k|k-1}}\right)^T + K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)^T + K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)$$

Como $\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)$ é a covariância do sinal de erro de medição,

$\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) = \left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)^T$ (vide Teorema 1),

$$\frac{\partial\left(Tr\left[P_{\tilde{x}_{k|k}}\right]\right)}{\partial K_k} = 2\left(\left(C_k P_{\tilde{x}_{k|k-1}}\right)^T - K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)\right) \tag{15}$$

Para minimizar a diferenciação, deve-se igualar a equação 15 a zero, $\frac{\partial Tr\left[P_{\tilde{x}_{k|k}}\right]}{\partial K_k} = 0$. Logo,

$$K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) = \left(C_k P_{\tilde{x}_{k|k-1}}\right)^T$$

Propriedade 3 e como $P_{\tilde{x}_{k|k-1}} = P_{\tilde{x}_{k|k-1}}^T$ (vide Teorema 1),

$$K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) = P_{\tilde{x}_{k|k-1}} C_k^T$$

Multiplica-se, à direita, ambos os membros da equação por $\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)^{-1}$,

$$K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)^{-1} = P_{\tilde{x}_{k|k-1}} C_k^T\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)^{-1}$$

Como $\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)^{-1}$ é igual ao elemento neutro da multiplicação matricial, I,

$$K_k = P_{\tilde{x}_{k|k-1}} C_k^T\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right)^{-1} \tag{16}$$

Dessa forma, a equação 16 fornece o ganho de Kalman, ou seja, a solução que minimiza as incertezas matemáticas de medição e garante que a esperança corrigida de estado convirja ao valor verdadeiro dos estados do sistema. Ademais, a formulação obtida para o ganho de Kalman simplifica algebricamente e interpretativamente a equação de covariância do atualizado sinal de erro de estado. Logo, pela equação 13,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - \left(K_k C_k P_{\tilde{x}_{k|k-1}}\right)^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k\left(C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k\right) K_k^T$$

Atribui-se $S_k = C_k P_{\tilde{x}_{k|k-1}} C_k^T + R_k$,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - \left(K_k C_k P_{\tilde{x}_{k|k-1}}\right)^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k S_k K_k^T$$

Propriedade 3,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - P_{\tilde{x}_{k|k-1}} C_k^T K_k^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k S_k K_k^T$$

Como $K_k S_k = P_{\tilde{x}_{k|k-1}} C_k^T$,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - K_k S_k K_k^T - K_k C_k P_{\tilde{x}_{k|k-1}} + K_k S_k K_k^T$$

Lei de comutatividade e corte da adição,

$$P_{\tilde{x}_{k|k}} = P_{\tilde{x}_{k|k-1}} - K_k C_k P_{\tilde{x}_{k|k-1}}$$

Multiplica-se a primeira parcela pela matriz identidade, I, e fatora-se $P_{\tilde{x}_{k|k-1}}$,

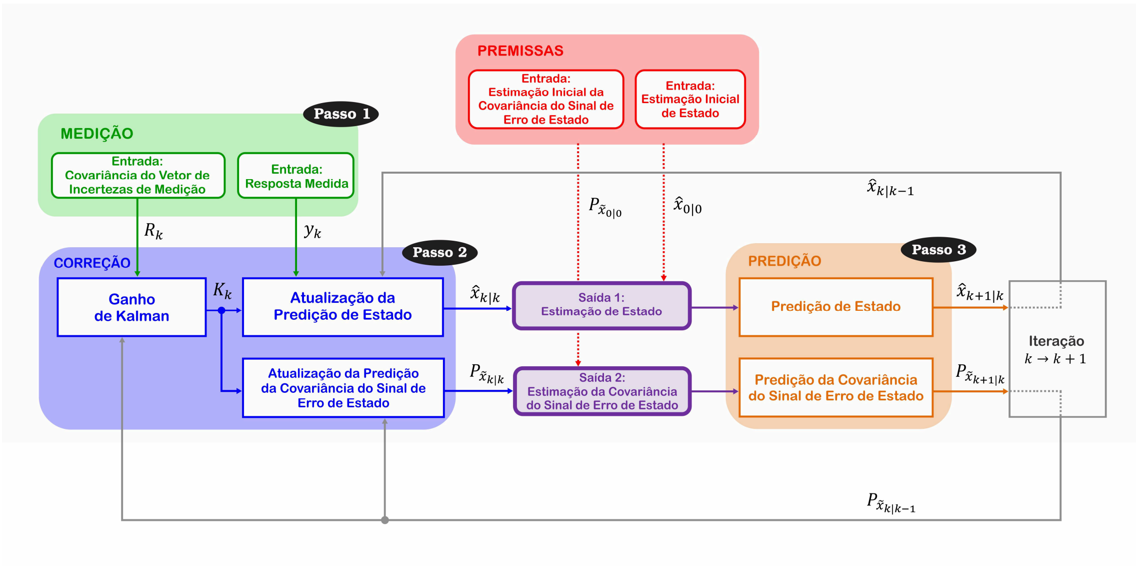
$$P_{\tilde{x}_{k|k}} = (I - K_k C_k) P_{\tilde{x}_{k|k-1}} \tag{17}$$

A equação 17 comprova que a dispersão estatística da etapa temporal k convergirá a zero ($\lim_{k\to\infty} P_{\tilde{x}_{k|k}} = 0$) desde que a dispersão estatística da etapa temporal $k-1$ convirja a zero ($\lim_{k\to\infty} P_{\tilde{x}_{k|k-1}} = 0$). Esta é a única e suficiente condição para que a convergência da esperança corrigida ao valor verdadeiro de estados ($\lim_{k\to\infty} \hat{x}_{k|k} = x_k$) seja indubitável. Logo, torna-se conhecido o predicado do sistema real.

2.3 Resumo teórico

Esta subseção engloba um completo fluxograma de funcionamento do algoritmo de otimização, Figura 3, e dois quadros, Quadro 1 e Quadro 2, que resumem as cinco equações do filtro de Kalman. Utilizou-se este resumo para a resolução de um problema de engenharia, apresentado na seguinte seção deste artigo.

Figura 3: Fluxograma de funcionamento do filtro de Kalman



Fonte: Adaptado de Kalman Filter Tutorial (2021)

Quadro 1: Equações da etapa de predição

Predição de Estado	$\hat{x}_{k|k-1} = A_k \hat{x}_{k-1|k-1} + B_k u_k$

Predição da Covariância do Sinal de Erro de Estado	$P_{\tilde{x}_{k\|k-1}} = A_k P_{\tilde{x}_{k-1\|k-1}} A_k^T + Q_k$

Fonte: Autoria própria (2021)

Quadro 2: Equações da etapa de correção

Ganho de Kalman	$K_k = P_{\tilde{x}_{k\|k-1}} C_k^T \left(C_k P_{\tilde{x}_{k\|k-1}} C_k^T + R_k \right)^{-1}$
Atualização da Predição de Estado	$\hat{x}_{k\|k} = \hat{x}_{k\|k-1} + K_k (y_k - \hat{y}_k)$
Atualização da Predição da Covariância do Sinal de Erro de Estado	$P_{\tilde{x}_{k\|k}} = (I - K_K C_K) P_{\tilde{x}_{k\|k-1}}$

Fonte: Autoria própria (2021)

3 METODOLOGIA

Nesta seção apresenta-se didaticamente uma aplicação do filtro de Kalman para estimação de parâmetros, simulada por intermédio do *software* computacional MATLAB.

3.1 Apresentação do problema

Realizada em 1969, a expedição espacial Apollo 11 foi responsável pelo primeiro pouso humano em solo lunar. A espaçonave tripulada pelos astronautas Neil Alden Armstrong, Edwin Eugene Aldrin Junior e Michael Collins foi lançada pelo foguete Saturno V. Durante o processo de decolagem, a sala de controle da NASA (*National Aeronautics and Space Administration*), localizada no terceiro andar do Centro Espacial Johnson, em Houston, nos Estados Unidos da América, que estava no comando desta expedição, tinha como um dos objetivos medir a altitude do foguete em relação ao solo terrestre. De acordo com Macedo e Serra (2020), a compatibilidade entre a trajetória percorrida e a trajetória prevista, é imprescindível para a segurança de voo aeroespacial.

Posto isto, considerou-se que o processo de decolagem do foguete detinha as características de um movimento retilíneo uniformemente variado, logo, modelou-se matematicamente o problema como um sistema dinâmico linear discreto. Dessa forma, a altitude, x_s, e a velocidade, x_v, do foguete são descritas recursivamente pelas leis de cinemática para esta categoria de movimento. Assim, obtêm-se as equações 18 e 19. Onde t é o tempo de amostragem, k é a variável temporal discreta, a é o vetor de controle que representa a constante aceleração do foguete, $\omega_s = \sigma_{\omega_s}\xi_{\omega_s}$ é o vetor de incertezas de posição, $\omega_v = \sigma_{\omega_v}\xi_{\omega_v}$ é o vetor de incertezas de velocidade, y_s representa o sensor de posição, acometido pelo vetor de incertezas de medição de posição $\vartheta_s = \sigma_{\vartheta_s}\xi_{\vartheta_s}$, e y_v representa o sensor de velocidade, acometido pelo vetor de incertezas de medição de velocidade $\vartheta_v = \sigma_{\vartheta_v}\xi_{\vartheta_v}$. As unidades estabelecidas para altitude, velocidade, aceleração e tempo, respeitam o Sistema Internacional de Unidades, ou seja, são, respectivamente, metros (m), metros por segundo ($\frac{\mathrm{m}}{\mathrm{s}}$), metros por segundo ao quadrado ($\frac{\mathrm{m}}{\mathrm{s}^2}$) e segundos ($\mathrm{s}$).

$$\begin{bmatrix} x_{s_k} \\ x_{v_k} \end{bmatrix} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s_{k-1}} \\ x_{v_{k-1}} \end{bmatrix} + \begin{bmatrix} \frac{t^2}{2} \\ t \end{bmatrix} [a_k] + \sqrt{\begin{bmatrix} \sigma^2_{\omega_{s_k}} & 0 \\ 0 & \sigma^2_{\omega_{v_k}} \end{bmatrix}} \begin{bmatrix} \xi_{\omega_{s_k}} \\ \xi_{\omega_{v_k}} \end{bmatrix} \tag{18}$$

$$\begin{bmatrix} y_{s_k} \\ y_{v_k} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s_k} \\ x_{v_k} \end{bmatrix} + \sqrt{\begin{bmatrix} \sigma^2_{\vartheta_{s_k}} & 0 \\ 0 & \sigma^2_{\vartheta_{v_k}} \end{bmatrix}} \begin{bmatrix} \xi_{\vartheta_{s_k}} \\ \xi_{\vartheta_{v_k}} \end{bmatrix} \tag{19}$$

3.2 Aplicação do filtro de Kalman ao problema

3.2.1 Premissas e incertezas matemáticas

Os parâmetros físicos utilizados nesta aplicação foram obtidos mediante o histórico de dados de decolagem da Apollo 11, disponibilizado no *website* de NASA History Division (2021). A aceleração do foguete é de 14,22 $\frac{\mathrm{m}}{\mathrm{s}^2}$, o desvio padrão de altitude e de velocidade são iguais a 12 m e 4 $\frac{\mathrm{m}}{\mathrm{s}}$, respectivamente, e o desvio padrão de medição de altitude e de medição de velocidade são iguais a 180 m e 60 $\frac{\mathrm{m}}{\mathrm{s}}$, respectivamente. Escolheu-se um tempo de amostragem igual a 0,1 s. Substituíram-se estes parâmetros nas equações 18 e 19 e obtiveram-se as equações 20 e 21.

$$\begin{bmatrix} x_{s_k} \\ x_{v_k} \end{bmatrix} = \begin{bmatrix} 1 & 0{,}1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s_{k-1}} \\ x_{v_{k-1}} \end{bmatrix} + \begin{bmatrix} \frac{0{,}1^2}{2} \\ 0{,}1 \end{bmatrix} [14{,}22] + \sqrt{\begin{bmatrix} 12^2 & 0 \\ 0 & 4^2 \end{bmatrix}} \begin{bmatrix} \xi_{\omega_{s_k}} \\ \xi_{\omega_{v_k}} \end{bmatrix} \tag{20}$$

$$\begin{bmatrix} y_{s_k} \\ y_{v_k} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s_k} \\ x_{v_k} \end{bmatrix} + \sqrt{\begin{bmatrix} 180^2 & 0 \\ 0 & 60^2 \end{bmatrix}} \begin{bmatrix} \xi_{\vartheta_{s_k}} \\ \xi_{\vartheta_{v_k}} \end{bmatrix} \tag{21}$$

Após a obtenção de todas as informações necessárias para a aplicação do filtro de Kalman, iniciou-se a criação do código fonte da simulação computacional do sistema no MATLAB. Este código estruturou-se em etapa de predição, na qual calcularam-se a predição de estados e a covariância do sinal de erro de estados, e em etapa de correção, na qual calculou-se o ganho de Kalman e atualizou-se a predição de estados (MAIA, 2009).

Por tratar-se de uma decolagem, admitiu-se para a primeira iteração que os vetores de estado, o vetor de controle e o vetor de incertezas de estado são nulos, e que a covariância do sinal de erro de estados é igual a matriz de covariância das incertezas dos estados. Após isto, alterou-se a equação de medição para aferir apenas o estado desejado, ou seja, a altitude do foguete. Para o vetor aleatório discreto gaussiano, $\xi \sim N(E[\xi] = 0, cov[\xi, \xi] = 1)$, utilizou-se a função *randn* do MATLAB, a qual retorna uma matriz com elementos pseudoaleatórios.

As equações 22 e 23 representam a primeira iteração ($k = 1$) do sistema estocástico.

$$\begin{bmatrix} x_{s_1} \\ x_{v_1} \end{bmatrix} = \begin{bmatrix} 1 & 0{,}1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{0{,}1^2}{2} \\ 0{,}1 \end{bmatrix} [0] + \sqrt{\begin{bmatrix} 12^2 & 0 \\ 0 & 4^2 \end{bmatrix}} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{22}$$

$$y_{s_1} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \sqrt{\begin{bmatrix} 180^2 & 0 \end{bmatrix}} \begin{bmatrix} \xi_{\vartheta_{s_1}} \\ \xi_{\vartheta_{v_1}} \end{bmatrix} \tag{23}$$

As equações 24 e 25 retratam a esperança deste sistema nesta iteração.

$$\begin{bmatrix} \hat{x}_{s_{1|0}} \\ \hat{x}_{v_{1|0}} \end{bmatrix} = \begin{bmatrix} 1 & 0{,}1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{0{,}1^2}{2} \\ 0{,}1 \end{bmatrix} [0] \tag{24}$$

$$\hat{y}_{s_1} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{25}$$

3.2.2 Etapa de predição de estados

Descreve-se o sistema estocástico, conforme equações 3 e 4, pelas equações 26 e 27 nas demais iterações.

$$\begin{bmatrix} x_{s_k} \\ x_{v_k} \end{bmatrix} = \begin{bmatrix} 1 & 0{,}1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s_{k-1}} \\ x_{v_{k-1}} \end{bmatrix} + \begin{bmatrix} \frac{0{,}1^2}{2} \\ 0{,}1 \end{bmatrix} [14{,}22] + \sqrt{\begin{bmatrix} 12^2 & 0 \\ 0 & 4^2 \end{bmatrix}} \begin{bmatrix} \xi_{\omega_{s_k}} \\ \xi_{\omega_{v_k}} \end{bmatrix} \tag{26}$$

$$y_{s_k} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_{s_k} \\ x_{v_k} \end{bmatrix} + \sqrt{\begin{bmatrix} 180^2 & 0 \end{bmatrix}} \begin{bmatrix} \xi_{\vartheta_{s_k}} \\ \xi_{\vartheta_{v_k}} \end{bmatrix} \tag{27}$$

Retrata-se a esperança do sistema estocástico, consoante equações 5 e 6, pelas equações 28 e 29 nas demais iterações.

$$\begin{bmatrix} \hat{x}_{s_{k|k-1}} \\ \hat{x}_{v_{k|k-1}} \end{bmatrix} = \begin{bmatrix} 1 & 0{,}1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_{s_{k-1|k-1}} \\ \hat{x}_{v_{k-1|k-1}} \end{bmatrix} + \begin{bmatrix} \frac{0{,}1^2}{2} \\ 0{,}1 \end{bmatrix} [14{,}22] \tag{28}$$

$$\hat{y}_{s_k} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{s_{k|k-1}} \\ \hat{x}_{v_{k|k-1}} \end{bmatrix} \tag{29}$$

A verificação da existência de dispersão estatística entre o sistema estocástico e a esperança deste ocorre mediante as equações 7 e 8. Expressam-se os sinais de erro dos estados pela equação 30. Representa-se a covariância do sinal de erro de posição e do sinal de erro de velocidade pela equação 31.

$$\begin{bmatrix} \tilde{x}_{s_{k|k-1}} \\ \tilde{x}_{v_{k|k-1}} \end{bmatrix} = \begin{bmatrix} 1 & 0{,}1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{x}_{s_{k-1|k-1}} \\ \tilde{x}_{v_{k-1|k-1}} \end{bmatrix} + \sqrt{\begin{bmatrix} 12^2 & 0 \\ 0 & 4^2 \end{bmatrix}} \begin{bmatrix} \xi_{\omega_{s_k}} \\ \xi_{\omega_{v_k}} \end{bmatrix} \tag{30}$$

$$\begin{bmatrix} P_{\tilde{x}_{s_{k|k-1}}} & 0 \\ 0 & P_{\tilde{x}_{v_{k|k-1}}} \end{bmatrix} = \begin{bmatrix} 1 & 0{,}1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P_{\tilde{x}_{s_{k-1|k-1}}} & 0 \\ 0 & P_{\tilde{x}_{v_{k-1|k-1}}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0{,}1 & 1 \end{bmatrix} + \begin{bmatrix} 12^2 & 0 \\ 0 & 4^2 \end{bmatrix} \tag{31}$$

Realiza-se a minimização da divergência constatada pela aplicação do filtro de Kalman.

3.2.3 Etapa de correção de estimações

Nesta etapa calculou-se o ganho de Kalman de cada estado, equação 32, com base na equação 16.

$$\begin{bmatrix} K_{s_k} \\ K_{v_k} \end{bmatrix} = \begin{bmatrix} P_{\tilde{x}_{s_{k|k-1}}} & 0 \\ 0 & P_{\tilde{x}_{v_{k|k-1}}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} P_{\tilde{x}_{s_{k|k-1}}} & 0 \\ 0 & P_{\tilde{x}_{v_{k|k-1}}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 180^2 & 0 \end{bmatrix} \right)^{-1} \tag{32}$$

Após este cálculo, obtiveram-se as esperanças corrigidas de estados, equação 33, com base na equação 11.

$$\begin{bmatrix} \hat{x}_{s_{k|k}} \\ \hat{x}_{v_{k|k}} \end{bmatrix} = \begin{bmatrix} \hat{x}_{s_{k|k-1}} \\ \hat{x}_{v_{k|k-1}} \end{bmatrix} + \begin{bmatrix} K_{s_k} \\ K_{v_k} \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_{s_k} \\ x_{v_k} \end{bmatrix} + \sqrt{\begin{bmatrix} 180^2 & 0 \end{bmatrix}} \begin{bmatrix} \xi_{\vartheta_{s_k}} \\ \xi_{\vartheta_{v_k}} \end{bmatrix} - \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_{s_k} \\ \hat{x}_{v_k} \end{bmatrix} \right) \tag{33}$$

Enfim, atualizaram-se as covariâncias dos sinais de erro de estados, equação 34, com base na equação 17.

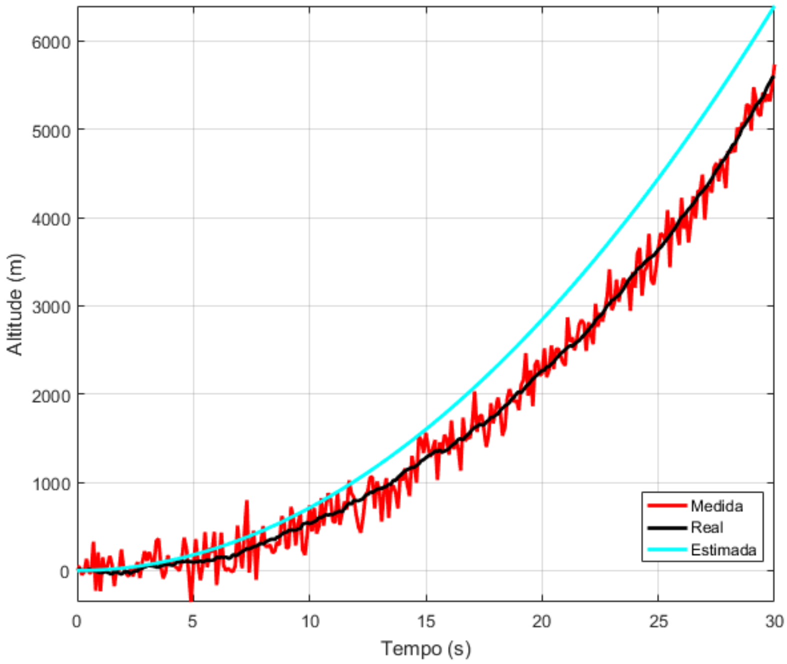
$$\begin{bmatrix} P_{\tilde{x}_{s_{k|k}}} & 0 \\ 0 & P_{\tilde{x}_{v_{k|k}}} \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} K_{s_k} \\ K_{v_k} \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \right) \begin{bmatrix} P_{\tilde{x}_{s_{k|k-1}}} & 0 \\ 0 & P_{\tilde{x}_{v_{k|k-1}}} \end{bmatrix} \tag{34}$$

Por caracterizar-se como um processo recursivo de estimação, as etapas de predição e correção mencionadas ocorrem iterativamente até que se atinja o tempo máximo predeterminado para a simulação do sistema, o qual parametrizou-se em 30 segundos (DOS SANTOS, 2019). Isto ocasionou a atualização dos valores da esperança dos estados e da covariância do sinal de erro de cada estado. Por derradeiro, coletaram-se os resultados de simulação computacional que serão expostos e discutidos na próxima seção deste artigo.

4 RESULTADOS E DISCUSSÕES

Nesta seção, realiza-se a ilustração e análise dos resultados obtidos, através de uma coleção de figuras e quadros gerados em concordância com a simulação computacional da problemática no MATLAB. Inicialmente, para um tempo de simulação igual a 30 segundos, ilustra-se, na Figura 4, a simulação do sistema anteriormente à aplicação do filtro de Kalman. Percebe-se que houve dispersão estatística entre a altitude real e a altitude estimada, e que esta divergência se agravou no decorrer do tempo.

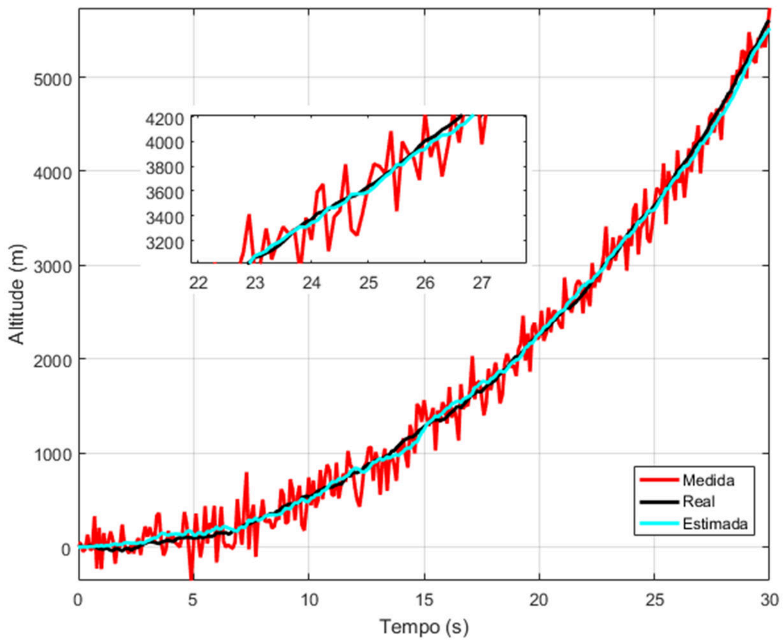
Figura 4: Simulação sem a aplicação do filtro de Kalman



Fonte: Autoria própria (2021)

Na Figura 5, observa-se que a implementação do filtro de Kalman no sistema propiciou a correção da altitude estimada e, consequentemente, a convergência desta à altitude real do foguete, pois minimizou as incertezas matemáticas de medição. O trecho ampliado evidencia a superioridade de acurácia da altitude estimada em relação à altitude medida.

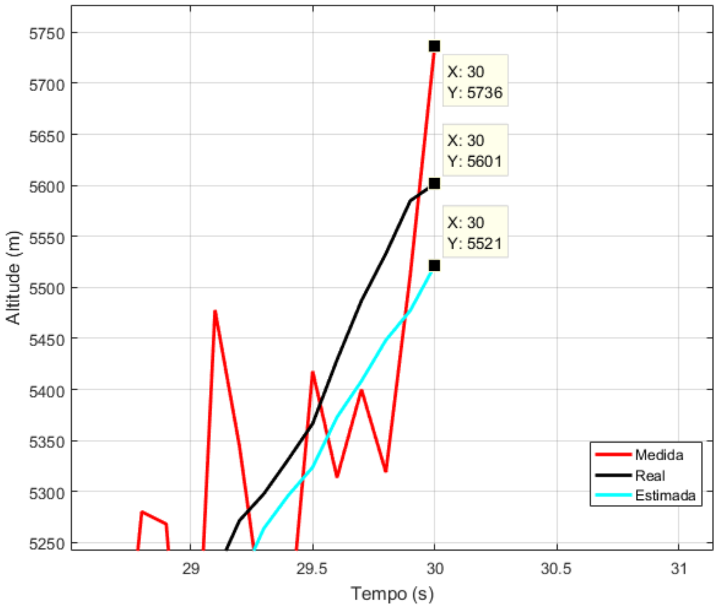
Figura 5: Simulação com a aplicação do filtro de Kalman



Fonte: Autoria própria (2021)

Na Figura 6, nota-se uma altitude estimada de 5521 metros, uma altitude medida de 5736 metros e uma altitude real de 5601 metros ao final do tempo de simulação predeterminado. Constata-se a proximidade numérica entre as altitudes estimada e real.

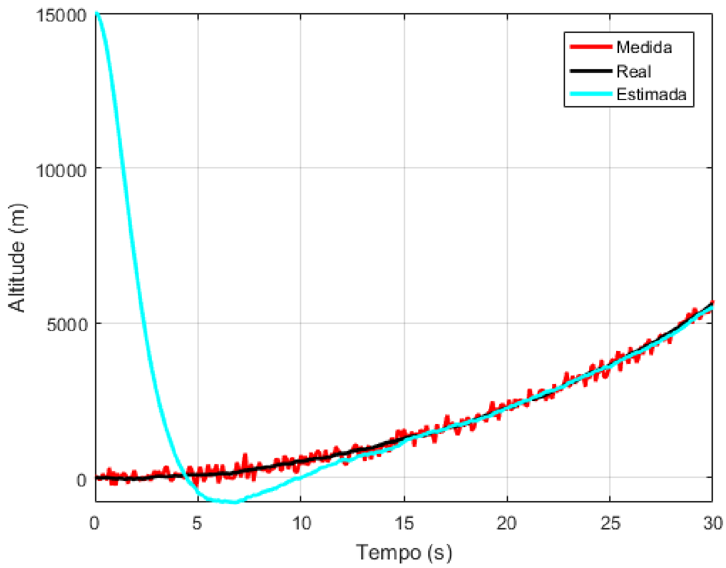
Figura 6: Altitudes obtidas na última etapa de simulação



Fonte: Autoria própria (2021)

Na Figura 7, desejou-se verificar a possibilidade de convergência entre as altitudes estimada e real na condição de uma premissa extrapolada. Para isto, supôs-se uma estimação inicial de altitude igual a 15000 metros. Depreende-se que mesmo em casos onde a estimação inicial apresenta ampla divergência em relação ao valor verdadeiro de estado inicial do sistema, a altitude estimada sempre covergirá à altitude real do foguete. Neste caso, o erro de estimação tornou-se irrisório a partir de 15 segundos de simulação.

Figura 7: Simulação com a aplicação do filtro de Kalman e extrapolação de premissa



Fonte: Autoria própria (2021)

Posteriormente, aumentou-se o tempo da simulação para 60 segundos e coletaram-se a cada 5 segundos os valores referentes às altitudes real, medida e estimada, conforme o Quadro 3. Pois, almejou-se ampliar a coleção de informações referentes ao sistema.

Quadro 3: Altitudes obtidas nas etapas de simulação

Tempo (segundos)	Altitude (metros)			Tempo (segundos)	Altitude (metros)		
	Real	Medida	Estimada		Real	Medida	Estimada
5	93,29	48,29	125,5	35	8154	8192	8184
10	531,3	738,5	486	40	10880	11030	10960
15	1277	1561	1247	45	13780	13700	13790
20	2265	2234	2266	50	17090	16950	17050
25	3639	3647	3594	55	20700	20760	20700
30	5601	5736	5521	60	24560	24810	24670

Fonte: Autoria própria (2021)

Finalmente, executou-se uma análise estatística dos dados elencados no Quadro 3. Dessa forma, verificou-se o erro quadrático médio (MSE, *Mean Square Error*) ou variância, definição 8, e a raiz do erro quadrático médio (RMSE, *Root Mean Square Error*) ou desvio padrão, definição 9, entre altitude medida e real, e entre altitude estimada e real para tempos de simulação iguais a 30 e 60 segundos. Percebe-se, no Quadro 4, que independentemente do tempo de simulação, o MSE da comparação entre altitude estimada e real é consideravelmente inferior ao MSE da comparação entre altitude medida e real, e, consequentemente, o mesmo ocorre para os valores do RMSE das comparações.

Quadro 4: Comparação estatística entre altitudes

Tempo (segundos)	MSE e RMSE entre altitudes			
	Medida e real		Estimada e real	
30	MSE	24143,81	MSE	2069,26
	RMSE	155,38	RMSE	45,49
60	MSE	21742,24	MSE	2792,96
	RMSE	147,45	RMSE	52,85

Fonte: Autoria própria (2021)

5 CONCLUSÃO

O presente artigo dissemina o conhecimento através da demonstração matemática e aplicação do filtro de Kalman. Isto fomenta a replicação de utilização desta ferramenta em diversos contextos da ciência. Os resultados obtidos na etapa de simulação confirmam a eficácia do filtro de Kalman em situações nas quais deseja-se estimar o comportamento de um sistema dinâmico cuja resposta medida é acometida por incertezas matemáticas. A minimização das imprecisões de medição proporcionou a covergência do estado estimado ao estado real do sistema. Por esta razão, comprova-se que o filtro de Kalman é um eficiente algoritmo de otimização de estimação de estados.

Sugere-se para trabalhos futuros a realização da demonstração matemática do filtro de Kalman para sistemas dinâmicos lineares contínuos e para sistemas dinâmicos não lineares. Além disto, aconselha-se a aplicação do filtro de Kalman em problemáticas relacionadas à fusão de sensores e reconciliação de dados. Ambas as recomendações devem respeitar a perspectiva fundamental de escrita que equilibra a didática e o rigor matemático.

REFERÊNCIAS

CARREIRA, A; PINTO, G. **Cálculo matricial I**. 1. ed. Lisboa: Instituto Piaget, 1999.

CARREIRA, A. **Cálculo matricial II**. 1. ed. Lisboa: Instituto Piaget, 1999.

CARREIRA, A; PINTO, G. **Cálculo matricial III**. 1. ed. Lisboa: Instituto Piaget, 1999.

DOS SANTOS, K. M. **Método de determinação de atitude utilizando o filtro de Kalman para a missão SPORT**. Rio Grande do Sul: UFSM, 2019.

HE, L.; SHIYI, C. **Video stabilization algorithm for tunnel robots based on improved Kalman filter**. Vancouver: Journal of Physics, 2020.

INTRODUCTION TO KALMAN FILTER. **Kalman Filter Tutorial**, 2021. Disponível em: https://www.kalmanfilter.net. Acesso em: 02 de novembro de 2021.

JIHYOUNG, C.; SANGHO, K.; SOON-YOUNG, P.; EUNHWAN, J. **Fault detection and diagnosis algorithms for transient state of an open-cycle liquid rocket engine using nonlinear Kalman filter methods**. Amsterdã: Elsevier, 2019.

JINAH, K.; JESEON, Y.; KIDEOK, D. **Wave data assimilation to modify wind forcing using an ensemble Kalman filter**. Berlim: Springer, 2020.

KALMAN, R. E. **A new approach to linear filtering and prediction problems**. Nova Iorque: ASME, 1960.

MACEDO, A. M.; SERRA, G. L. de O. **Localização de veículos aeroespaciais por triangulação de antenas**. São Paulo: SIGE, 2020.

MAIA, N. F. L. **Estimação de incertezas em modelos dinâmicos pelo método de filtragem de Kalman-Bucy**. Covilhã: UBI, 2009.

MONTGOMERY, D. C.; RUGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 6. ed. Rio de Janeiro: LTC, 2018.

QIANG, L.; RANYANG, L.; KAIFAN, J.; WEI, D. **Kalman filter and its application**. Kunming: KUST, 2015.

RAHEMI, N.; MOSAVI, M. R. **Positioning accuracy improvement in high-speed GPS**

receivers using sequential extended Kalman filter. Stevenage: IET Signal Processing, 2021.

SINGH, K. K.; KUMAR, S.; DIXIT, P.; BAJPAI, M. K. **Kalman filter based short term prediction model for COVID-19 spread**. Berlim: Springer, 2020.

STEINBRUCH, A.; WINTERLE, P. **Álgebra linear**. 2. ed. São Paulo: Pearson, 1987.

THE APOLLO 11 FLIGHT JOURNAL. **NASA History Division**, 2019. Disponível em: https://www.history.nasa.gov/afj/ap11fj/index.html. Acesso em: 02 de novembro de 2021.

WELCH, G.; BISHOP, G. **An introduction to the Kalman filter**. Chapel Hill: UNC, 2006.