

4 Filtro de Kalman

Este capítulo trata da apresentação resumida do filtro de Kalman. O filtro de Kalman tem sua origem na década de sessenta, dentro da área da engenharia elétrica relacionado à teoria do controle de sistemas. Posteriormente, esta metodologia foi sendo incorporada a outras áreas como a estatística. Sua aplicação na área econômica e financeira é vasta. Inúmeros artigos e trabalhos são publicados rotineiramente fazendo uso do filtro de Kalman. Este capítulo está baseado em Harvey (1989). Outra referência relevante e recente sobre o filtro de Kalman e modelos não lineares e não Gaussianos é Durbin e Koopman (2002). As referências clássicas sobre o tema são Anderson e Moore (1979) e Jazwinski (1970).

O capítulo está assim organizado: a primeira seção apresenta uma breve introdução do assunto; a segunda seção mostra a forma espaço-estado; a seguir são apresentados os aspectos computacionais e probabilísticos; a quinta seção apresenta o algoritmo do filtro; a sexta seção mostra a estimação dos parâmetros; em seguida são apresentados os conceitos de previsão e alisamento e, por fim, o conceito do filtro de Kalman estendido.

4.1 Introdução

Em 1960 Rudolph Emil Kalman publicou um famoso artigo descrevendo um processo recursivo para solucionar problemas lineares relacionados à filtragem de dados discretos. Sua pesquisa proporcionou contribuições relevantes ajudando a estabelecer bases teóricas sólidas em várias áreas da engenharia de sistemas. Em 1960-1961 Kalman desenvolveu, com colaboração de Richard S. Bucy, a versão em tempo contínuo do filtro de Kalman, que se tornou conhecida como o filtro de Kalman-Bucy. Com o avanço computacional, o filtro de Kalman e suas extensões a problemas não lineares representam o produto mais largamente utilizado dentro da moderna teoria de controle.

O filtro de Kalman é um conjunto de equações matemáticas que constitui um processo recursivo eficiente de estimação, uma vez que o erro quadrático é minimizado. Através da observação da variável denominada “variável de observação” outra variável, não observável, denominada “variável de estado” pode ser estimada eficientemente. Podem ser estimados os estados passados, o estado presente e mesmo previstos os estados futuros.

O filtro de Kalman é um procedimento aplicável quando os modelos estão escritos sob a forma espaço-estado. Além disso, o filtro de Kalman permite a estimação dos parâmetros desconhecidos do modelo através da maximização da verossimilhança via decomposição do erro de previsão.

4.2

Definição do modelo na forma espaço-estado

Seja y_t uma série temporal multivariada com N elementos. Estas variáveis são denominadas variáveis observáveis e constituem um vetor $N \times 1$, $\mathbf{y}_t \in \mathbb{R}^N$. As variáveis observáveis estão relacionadas às variáveis de estado \mathbf{x}_t através da equação de medição (ou observação):

$$\mathbf{y}_t = \mathbf{Z}_t \mathbf{x}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t \quad t = 1, 2, \dots, T \quad (44)$$

onde \mathbf{Z}_t é uma matriz $N \times m$, \mathbf{d}_t é um vetor $N \times 1$, $\boldsymbol{\varepsilon}_t$ é um vetor serialmente não correlacionado com média zero e matriz de covariância \mathbf{H}_t e \mathbf{x}_t é um vetor $m \times 1$ que contém as variáveis de estado não observáveis.

As variáveis de estado são geradas por um processo Markoviano de primeira ordem e sua equação é denominada equação de transição:

$$\mathbf{x}_t = \mathbf{T}_t \mathbf{x}_{t-1} + \mathbf{c}_t + \mathbf{R}_t \boldsymbol{\eta}_t \quad t = 1, 2, \dots, T \quad (45)$$

onde \mathbf{T}_t é uma matriz $m \times m$, \mathbf{c}_t é um vetor $m \times 1$, \mathbf{R}_t é uma matriz $m \times g$ e $\boldsymbol{\eta}_t$ é um vetor $g \times 1$ serialmente não correlacionado com média zero e matriz de covariância \mathbf{Q}_t .

Além disso, o vetor inicial de estado \mathbf{x}_0 tem média $\hat{\mathbf{x}}_0$ e matriz de covariância \mathbf{P}_0 . Os ruídos $\boldsymbol{\varepsilon}_t$ e $\boldsymbol{\eta}_t$ são não correlacionados entre si e não correlacionados com estado inicial.

4.3 As origens computacionais do filtro de Kalman

Vamos definir por $\hat{\mathbf{x}}_t^- \in \mathfrak{R}^m$ a estimativa do estado anterior (a *priori*) no tempo t dado que se conhece todo o processo anterior a t . Por estado anterior refere-se ao estado antes do conhecimento da variável de observação em t , \mathbf{y}_t . Da mesma forma, $\hat{\mathbf{x}}_t \in \mathfrak{R}^m$ é a estimativa do estado posterior (a *posteriori*) em t dado que conhece a medição (ou observação), \mathbf{y}_t .

Definem-se os erros de medição anterior e posterior como:

$$\mathbf{e}_t^- = \mathbf{x}_t - \hat{\mathbf{x}}_t^- \quad (46)$$

$$\mathbf{e}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t \quad (47)$$

A matriz de covariância do erro anterior \mathbf{P}_t^- é $m \times m$ e dada por:

$$\mathbf{P}_t^- = E(\mathbf{e}_t^- \mathbf{e}_t^{-'}) \quad (48)$$

A matriz de covariância do erro posterior \mathbf{P}_t é $m \times m$ e dada por:

$$\mathbf{P}_t = E(\mathbf{e}_t \mathbf{e}_t') \quad (49)$$

Deseja-se encontrar uma equação que relaciona o estado posterior $\hat{\mathbf{x}}_t$ como uma combinação linear do estado anterior $\hat{\mathbf{x}}_t^-$ com a ponderação da diferença entre a observação \mathbf{y}_t e a previsão $\mathbf{Z}_t \hat{\mathbf{x}}_t^- + \mathbf{d}_t$ ou seja:

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t^- + \mathbf{K}_t (\mathbf{y}_t - \mathbf{Z}_t \hat{\mathbf{x}}_t^- - \mathbf{d}_t) \quad (50)$$

O termo entre parênteses reflete a diferença entre o previsto $\mathbf{Z}_t \hat{\mathbf{x}}_t^- + \mathbf{d}_t$ e a observação \mathbf{y}_t . A matriz \mathbf{K}_t ($m \times N$) é denominada ganho de Kalman e é tal que minimiza a matriz de covariância do erro \mathbf{P}_t dada pela eq. (49).

A minimização da covariância do erro é obtida substituindo a eq. (50) na eq. (47) e obtendo-se uma expressão para \mathbf{e}_t em termos de \mathbf{K}_t . Levando este resultado na eq. (49), tomando-se então os valores esperados, derivando-os com relação a \mathbf{K}_t e igualando a equação a zero; tem-se a condição de primeira ordem. Resolvida esta equação para \mathbf{K}_t , resulta:

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{Z}_t' (\mathbf{Z}_t \mathbf{P}_t^- \mathbf{Z}_t' + \mathbf{H}_t)^{-1} \quad (51)$$

Na equação acima quando o erro da equação de medição aproxima-se de zero ($\mathbf{H}_t \rightarrow 0$), a ponderação da matriz ganho aumenta. Ou ainda

$$\lim_{H_t \rightarrow 0} \mathbf{K}_t = \mathbf{Z}_t^{-1}$$

4.4 As origens probabilísticas do filtro

A derivação do filtro de Kalman apóia-se no fato de que tanto os ruídos das equações de medição e transição como o vetor inicial de estado, são normalmente distribuídos. Isto significa que apenas os dois primeiros momentos são suficientes para descrever todos os estados em qualquer instante de $t = 1$ a $t = T$. Assim sendo escreve-se:

$$\hat{\mathbf{x}}_t = E(\mathbf{x}_t) \quad \text{e} \quad \mathbf{P}_t = E\{(\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)'\}$$

A estimativa posterior dada em (50) é Gaussiana. A matriz de covariância posterior em (49) reflete a variância da distribuição das variáveis de estado. Então,

$$p(\mathbf{x}_t | \mathbf{y}_t) \sim N(\hat{\mathbf{x}}_t, \mathbf{P}_t)$$

Até o momento foi visto que o filtro de Kalman é um procedimento recursivo que permite determinar o estimador ótimo do vetor de estado dadas as informações disponíveis até o tempo t , inclusive as variáveis de observação \mathbf{y}_t . Esta é a tradução do que está escrito na eq. (50). O estimador é dito ótimo no sentido de que a matriz de ganho é tal que a variância do erro das variáveis de estado é mínima. Esta é a tradução do resultado obtido na eq. (51). Quando a hipótese da normalidade não se verifica, o filtro de Kalman não fornece o valor esperado das variáveis de estado. Entretanto, o filtro continua sendo o estimador ótimo, isto é minimiza a variância do erro.

4.5 O algoritmo do filtro de Kalman

As equações do filtro de Kalman podem ser agrupadas em dois tipos distintos: equações de atualização do tempo e equações de atualização da medição. Estes dois grupos de equações funcionam conjuntamente como um sistema com retroalimentação.

As equações de atualização do tempo são responsáveis pelo avanço das variáveis de estado e das covariâncias no tempo para se obter, desta forma, as estimativas anteriores (*a priori*) para o próximo instante.

As equações de atualização das medições são responsáveis pela retroalimentação, ou seja, incorporam uma nova informação da variável observável nas estimativas anteriores para obter um ganho (ou melhoria) na estimação posterior.

As equações de atualização do tempo são denominadas equações de previsão. As equações de atualização das medições são denominadas equações de correção.

Seja então o modelo especificado em (44) e (45). Seja $\hat{\mathbf{x}}_{t-1}$ o estimador ótimo de \mathbf{x}_{t-1} baseado em informações até $t-1$ incluindo \mathbf{y}_{t-1} . Dados $\hat{\mathbf{x}}_{t-1}$ e \mathbf{P}_{t-1} , o estimador ótimo de \mathbf{x}_t é dado por:

$$\hat{\mathbf{x}}_t^- = \mathbf{T}_t \hat{\mathbf{x}}_{t-1} + \mathbf{c}_t \quad (52)$$

A matriz de covariância dos erros das variáveis de estado é dada por:

$$\mathbf{P}_t^- = \mathbf{T}_t \mathbf{P}_{t-1} \mathbf{T}_t' + \mathbf{R}_t \mathbf{Q}_t \mathbf{R}_t' \quad (53)$$

As equações (52) e (53) constituem o grupo denominado de equações de atualização do tempo ou equações de previsão. Estas equações representam um avanço no tempo de $t-1$ para t . Quando uma nova observação \mathbf{y}_t é verificada, o estimador $\hat{\mathbf{x}}_t^-$ de \mathbf{x}_t pode ser melhorado ou corrigido. As equações de atualização das medições são:

$$\mathbf{K}_t = \mathbf{P}_t^- \mathbf{Z}_t' (\mathbf{Z}_t \mathbf{P}_t^- \mathbf{Z}_t' + \mathbf{H}_t)^{-1} \quad (54)$$

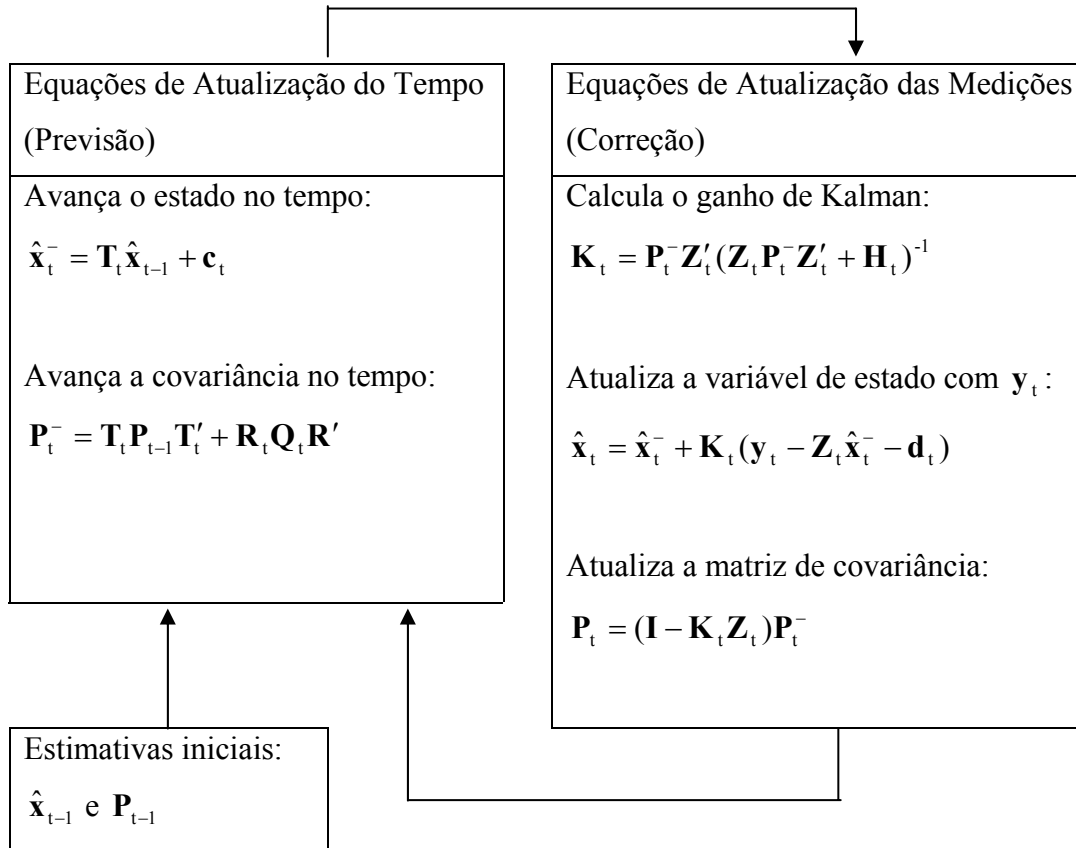
$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t^- + \mathbf{K}_t (\mathbf{y}_t - \mathbf{Z}_t \hat{\mathbf{x}}_t^- - \mathbf{d}_t) \quad (55)$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{Z}_t) \mathbf{P}_t^- \quad (56)$$

O primeiro passo é determinar o ganho \mathbf{K}_t dado pela eq. (54). Posteriormente, a nova informação observada \mathbf{y}_t é incorporada à previsão anterior $\hat{\mathbf{x}}_t^-$ juntamente com a matriz ganho \mathbf{K}_t através da eq. (55), gerando a estimação posterior $\hat{\mathbf{x}}_t$. O último passo é obter matriz de covariância dos erros através da eq. (56). O ciclo do algoritmo se repete para o instante de tempo $t+1$ sendo $\hat{\mathbf{x}}_t$ e \mathbf{P}_t dados de entrada nas equações (52) e (53), respectivamente. Esta natureza recursiva do modelo o torna um instrumento de atualização de medidas em tempo real, daí o seu grande uso em sistemas de controle e rastreamento no campo da engenharia. Em Finanças tem um forte apelo nos mercados financeiros

que produzem informações a cada instante. Desta forma, variáveis não observáveis podem ser estimadas a partir de preços obtidos no mercado à medida que uma nova informação é produzida.

O Quadro 3 é uma representação esquemática das etapas recursivas do filtro de Kalman. O algoritmo exige que haja uma operação de inversão de matriz para o cálculo do ganho de Kalman.



Quadro 3 – Representação esquemática do filtro de Kalman

4.6 Estimação por máxima verossimilhança

O filtro de Kalman retorna o valor de todas as variáveis de $t = 1$ até $t = T$. Resulta então todos os valores da variável de observação \mathbf{y}_t . Os valores observados de \mathbf{y}_t são conhecidos. Se as observações $\{\mathbf{y}_t\}_{t=1}^{t=T}$ são independentes e identicamente distribuídas a função de densidade conjunta é dada por:

$$L(\mathbf{y}; \Theta) = \prod_{t=1}^T p(\mathbf{y}_t) \tag{57}$$

onde $p(\cdot)$ representa a função densidade de probabilidade em um instante t e Θ é o conjunto de hiperparâmetros que fazem parte do sistema de matrizes. A eq. (57) representa a função de verossimilhança. O estimador de máxima verossimilhança é obtido maximizando (57) em relação a Θ . Quando as observações não são independentes pode-se usar a definição de função densidade conjunta condicional:

$$L(\mathbf{y}; \Theta) = \prod_{t=1}^T p(\mathbf{y}_t | \mathbf{Y}_{t-1})$$

onde $P(\cdot | \mathbf{Y}_{t-1})$ representa a função densidade de probabilidade condicional em $t-1$. Se os ruídos em (44) e (45) e se o vetor inicial de estado são Gaussianos, a distribuição de \mathbf{y}_t condicional a \mathbf{Y}_{t-1} também é Gaussiana. Pode-se escrever que condicionalmente a \mathbf{Y}_{t-1} , \mathbf{x}_t é tal que:

$$p(\mathbf{x}_t | \mathbf{Y}_{t-1}) \sim N(\hat{\mathbf{x}}_t^-, \mathbf{P}_t^-)$$

Da equação (42) pode ser escrito:

$$\mathbf{y}_t = \mathbf{Z}_t \mathbf{x}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t$$

que pode ser reescrita como

$$\mathbf{y}_t = \mathbf{Z}_t \hat{\mathbf{x}}_t^- - \mathbf{Z}_t \hat{\mathbf{x}}_t^- + \mathbf{Z}_t \mathbf{x}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t$$

ou ainda

$$\mathbf{y}_t = \mathbf{Z}_t \hat{\mathbf{x}}_t^- + \mathbf{Z}_t (\mathbf{x}_t - \hat{\mathbf{x}}_t^-) + \mathbf{d}_t + \boldsymbol{\varepsilon}_t$$

Então a distribuição condicional de \mathbf{y}_t é normal com média:

$$E(\mathbf{y}_t | \mathbf{Y}_{t-1}) = \tilde{\mathbf{y}}_t^- = \mathbf{Z}_t \hat{\mathbf{x}}_t^- + \mathbf{d}_t$$

com matriz de covariância dada por:

$$\mathbf{F}_t = \mathbf{Z}_t \mathbf{P}_t^- \mathbf{Z}_t' + \mathbf{H}_t$$

Conseqüentemente para um modelo Gaussiano o logaritmo da função verossimilhança é:

$$\ln L(\mathbf{y} | \Theta) = -\frac{NT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln |\mathbf{F}_t| - \frac{1}{2} \sum_{t=1}^T \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t \quad (58)$$

onde $\mathbf{v}_t = \mathbf{y}_t - \mathbf{y}_{t|t-1}$ para $t = 1, \dots, T$, é o vetor dos erros de previsão. Então a eq. (58) é a decomposição dos erros de previsão.

A cada rodada do filtro (dados os parâmetros do modelo) são determinados os estados e as observações. A função de verossimilhança também pode ser

calculada a cada rodada. Pode-se então implementar um algoritmo para maximização da função de verossimilhança.

Os valores iniciais \mathbf{x}_0 podem ser tratados como parâmetro iniciais e a maximização da eq. (58) tem como base o conjunto (\mathbf{x}_0, Θ) . Entretanto, isto é um complicador considerável para o problema de maximização. Uma solução prática é considerar que condicionalmente a Θ o estimador de máxima verossimilhança de \mathbf{x}_0 é uma função linear das observações. Neste caso, a maximização seria somente em relação a Θ . Em alguns modelos estacionários as condições de regime permanente permitem que sejam obtidas soluções para \mathbf{x}_0 e \mathbf{P}_0 . Em outras situações são necessárias as condições iniciais de \mathbf{x}_0 e \mathbf{P}_0 .

4.7 Previsão

Após a obtenção da estimação dos estados simultaneamente à estimação dos parâmetros, os estados futuros e as observações futuras em $T+1, T+2, \dots, T+\ell$ podem ser obtidos. Assim, para $T+1$:

$$\mathbf{x}_{T+1|T} = \mathbf{T}_{T+1} \hat{\mathbf{x}}_T + \mathbf{c}_{T+1}$$

$$\mathbf{y}_{T+1|T} = \mathbf{Z}_{T+1} \hat{\mathbf{x}}_{T+1|T} + \mathbf{d}_{T+1}$$

Para a previsão de $T+2$, $T+3$, etc deve-se usar sucessivamente as equações (44) e (45). Desta forma obtém-se $\mathbf{x}_{T+\ell}$ e $\mathbf{y}_{T+\ell}$. A previsão será então:

$$\hat{\mathbf{x}}_{T+\ell|T} = \mathbf{E}_t(\mathbf{x}_{T+\ell})$$

$$\hat{\mathbf{y}}_{T+\ell|T} = \mathbf{E}_t(\mathbf{y}_{T+\ell})$$

onde $\mathbf{E}_t(\cdot)$ representa o valor esperado condicional dadas as informações até o instante t .

4.8 Alisamento

No processo de filtragem obtém-se o valor filtrado em t que é o valor esperado da variável de estado condicional às informações disponíveis até o tempo t , ou seja:

$$\hat{\mathbf{x}}_t = E(\mathbf{x}_t | \mathbf{Y}_t)$$

O alisamento permite obter o valor condicional acima baseado em informações posteriores a t . Assim, a estimativa alisada é dada por:

$$\hat{\mathbf{x}}_{t|T} = E(\mathbf{x}_t | \mathbf{Y}_T)$$

$$\text{onde } \mathbf{Y}_T = \{y_t\}_{t=1}^{t=T}.$$

A matriz de covariância suavizada (ou alisada) é dada por:

$$\mathbf{P}_{t|T} = E\left\{(\mathbf{x}_t - \hat{\mathbf{x}}_{t|T})(\mathbf{x}_t - \hat{\mathbf{x}}_{t|T})' | \mathbf{Y}_T\right\}$$

Em geral, o erro da estimativa alisada é menor que o erro da estimativa filtrada. Isto porque o alisamento leva em consideração um maior número de informações que a filtragem.

4.9

O filtro de Kalman estendido

As principais características do filtro de Kalman estão relacionadas à linearidade das equações de medição e transição e a Gaussianidade dos resíduos como visto em (44) e (45). Existem vários aspectos referentes à não adequabilidade de aplicação do filtro de Kalman que violam as propriedades acima:

(i) Quando as matrizes em (44) e (45) $\mathbf{Z}_t, \mathbf{T}_t, \mathbf{d}_t, \mathbf{c}_t, \mathbf{H}_t, \mathbf{R}_t$ e \mathbf{Q}_t são estocásticas, isto é dependem da informação disponível em $t-1$, o modelo é dito condicionalmente Gaussiano.

(ii) Quando o modelo não é mais Gaussiano, isto é quando os resíduos não são distribuições normais, a obtenção da estimativa ótima do vetor de estado não é mais alcançada usando-se o modelo do filtro de Kalman na sua forma clássica. Ainda assim, dadas as hipóteses das novas distribuições, pode-se alcançar a solução teórica de minimização da matriz de covariância dos erros dos vetores de estado.

(iii) Quando a equação de medição não é uma função linear dos estados e por outro lado, quando na equação o vetor de estado não é uma função linear do vetor de estado no instante anterior; os modelos saem da condição de linearidade clássica propriamente dita. Em tais situações os modelos são ditos funcionalmente não-lineares. A solução é a linearização dos modelos.

Para os itens (i) e (ii) veja o capítulo 3, seção 3.7, de Harvey (1989). Com relação à não funcionalidade linear, os aspectos mais relevantes são apresentados a seguir.

Considere o modelo não-linear com as equações de transição e medição:

$$\mathbf{x}_t = \mathbf{h}_t(\mathbf{x}_{t-1}) + \mathbf{R}_t(\mathbf{x}_{t-1})\boldsymbol{\eta}_t \quad (59)$$

$$\mathbf{y}_t = \mathbf{f}_t(\mathbf{x}_t) + \boldsymbol{\varepsilon}_t \quad (60)$$

onde $\mathbf{h}(\cdot)$ e $\mathbf{f}(\cdot)$ são vetores $m \times 1$ e $N \times 1$, respectivamente e seus elementos são funções não lineares. A matriz \mathbf{R}_t é $m \times g$ e é função do vetor de estado em $t-1$.

As condições apresentadas em (59) e (60) fogem das condições clássicas de linearidade do filtro de Kalman. Sob a hipótese de que $\mathbf{f}(\cdot)$, $\mathbf{h}(\cdot)$ e \mathbf{R}_t são suaves pode-se usar a série de Taylor linearizando-as, centradas nas médias condicionais $\tilde{\mathbf{x}}_t^-$ e $\tilde{\mathbf{x}}_{t-1}$. Isto pode ser escrito assim:

$$\mathbf{h}_t(\hat{\mathbf{x}}_{t-1}) \sim \mathbf{h}_t(\tilde{\mathbf{x}}_{t-1}) + \mathbf{T}_t(\mathbf{x}_{t-1} - \tilde{\mathbf{x}}_{t-1}) \quad (61)$$

$$\mathbf{f}_t(\mathbf{x}_t) \sim \mathbf{f}_t(\tilde{\mathbf{x}}_t^-) + \mathbf{Z}_t(\mathbf{x}_t - \tilde{\mathbf{x}}_t^-) \quad (62)$$

$$\mathbf{R}_t(\mathbf{x}_{t-1}) \sim \mathbf{R}_t \quad (63)$$

onde \mathbf{T}_t é matriz Jacobiana dada por:

$$\mathbf{T}_t = \left. \frac{\partial(\mathbf{h}_t(\mathbf{x}_{t-1}))}{\partial \mathbf{x}'_{t-1}} \right|_{\mathbf{x}_{t-1} = \tilde{\mathbf{x}}_{t-1}} \quad (64)$$

Da mesma forma, \mathbf{Z}_t é a matriz Jacobiana dada por:

$$\mathbf{Z}_t = \left. \frac{\partial(\mathbf{f}_t(\mathbf{x}_{t-1}))}{\partial \mathbf{x}'_{t-1}} \right|_{\mathbf{x}_t = \tilde{\mathbf{x}}_t} \quad (65)$$

Ainda pode ser escrito que:

$$\mathbf{R}_t = \mathbf{R}_t(\tilde{\mathbf{x}}_t^-) \quad (66)$$

Levando (61) e (66) em (59):

$$\mathbf{x}_t = \mathbf{h}_t(\tilde{\mathbf{x}}_{t-1}) + \mathbf{T}_t(\mathbf{x}_{t-1} - \tilde{\mathbf{x}}_{t-1}) + \mathbf{R}_t\boldsymbol{\eta}_t$$

ou ainda:

$$\mathbf{x}_t = \mathbf{T}_t\mathbf{x}_{t-1} + \mathbf{c}_t + \mathbf{R}_t\boldsymbol{\eta}_t \quad (67)$$

$$\text{onde } \mathbf{c}_t = \mathbf{h}_t(\tilde{\mathbf{x}}_{t-1}) - \mathbf{T}_t\tilde{\mathbf{x}}_{t-1}$$

Levando (62) em (60):

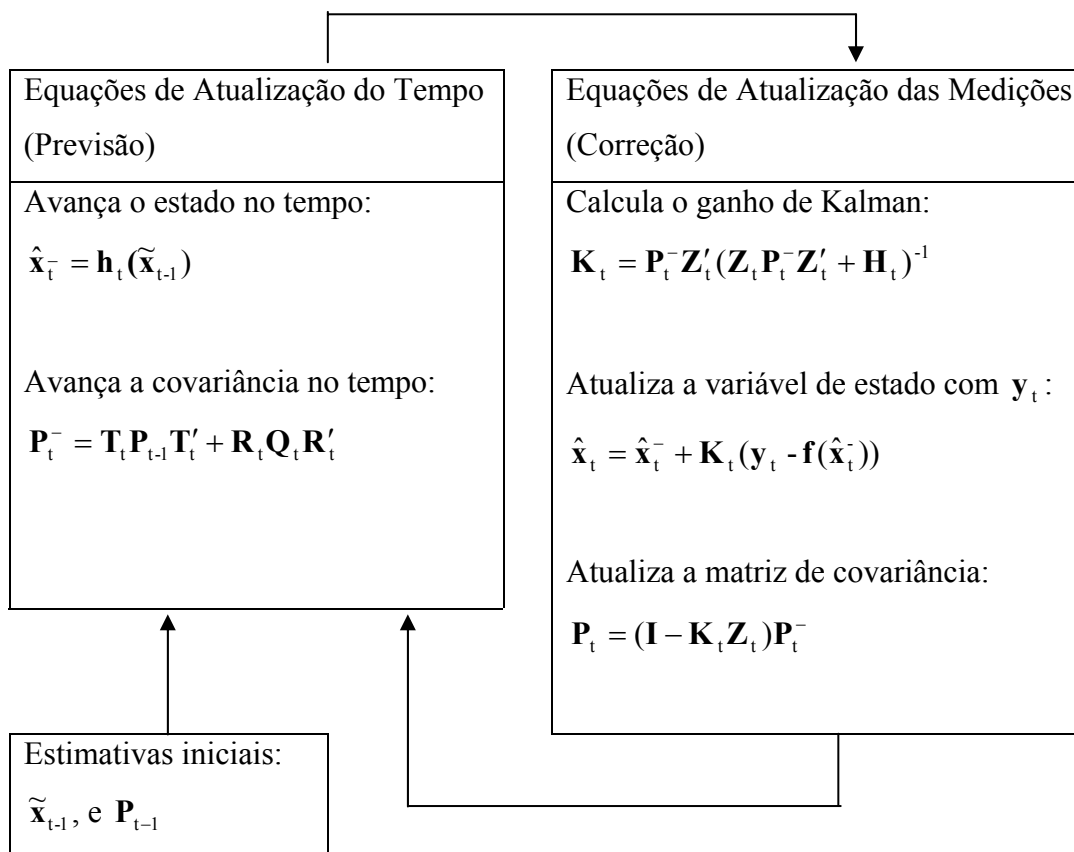
$$y_t = f_t(\tilde{x}_t^-) + Z_t(x_t - \tilde{x}_t^-) + \varepsilon_t$$

ou ainda:

$$y_t = Z_t x_t + d_t + \varepsilon_t \tag{68}$$

onde $d_t = f_t(\tilde{x}_t^-) + Z_t \tilde{x}_t^-$

Agora as equações (67) e (68) constituem uma aproximação linear para o problema descrito em (59) e (60). A nova equação de transição (67) é análoga a equação (45). Por outro lado, a nova equação de medição (68) é análoga a equação (44). O filtro de Kalman clássico agora pode ser aplicado seguindo o algoritmo de equações de atualização do tempo e de equações de atualização das medições conforme o Quadro 4.



Quadro 4 – Representação esquemática do filtro de Kalman estendido