

UNIVERSIDADE PRESBITERIANA MACKENZIE

ISMAR FERREIRA BARBOZA CAMINHA

RASTREAMENTO DE ALVOS EM SISTEMAS RADAR

São Paulo

2012

ISMAR FERREIRA BARBOZA CAMINHA

RASTREAMENTO DE ALVOS EM SISTEMAS RADAR

Trabalho de Conclusão de
Curso apresentado ao Curso de Engenharia
de Telecomunicações, da Universidade
Presbiteriana Mackenzie, como requisito
parcial à obtenção do grau de Especialista

ORIENTADOR: Prof. Dr. Laercio Nogueira

São Paulo

2012

AGRADECIMENTOS

Agradeço a Deus pelo dom da inteligência, e principalmente pelo desejo de usá-la em prol de todos.

Agradeço a minha mãe Mercedes, pois sem todo o seu esforço, devido a perda precoce de meu pai, nada teria conquistado.

Agradeço também a minha esposa Margarete, por me prover de todo o apoio necessário a conquista de mais um passo em minha formação.

Agradeço a meu filho Rafael, por compreender a minha ausência neste período da sua jovem vida, sem me cobrar nada por isso.

Um fato é uma verdade simples em que todos acreditam. Nasce inocente até que se prove seu erro. Uma hipótese é uma nova sugestão que ninguém ousa acreditar. Nasce culpada, até que se prove sua eficiência (Edward Teller).

RESUMO

Este trabalho aborda o tema rastreamento de alvos em sistemas radar; sua importância no sistema de proteção ao voo e as ferramentas usadas em sua concepção. O trabalho começa com um histórico do radar, desde seu desenvolvimento até a sua implantação no Brasil. A seguir fala-se sobre filtros com ênfase em filtros digitais e em seguida um capítulo dedicado exclusivamente a filtro de Kalman, pois é a ferramenta mais utilizada hoje em dia na área de rastreamento digital de sinais. O trabalho se encerra com a discussão sobre rastreamento de alvos em si e a utilização de filtros digitais com este fim, seguido da conclusão final.

Palavras-chave: Rastreamento. Radar. Filtros. Kalman.

ABSTRACT

This work is about target tracking in radar systems; its importance in the flight protection system and the tools used in its conception. The work begins with the history of radar, from its development till its arrival at Brazil. Just next, talks about filters, with especial highlights in digital filters and next a chapter dedicate exclusively to Kalman filter, because it is the the most common use tool in digital tracking nowadays. This work finishes talking about target tracking itself and the use of digital filters with this purpose, followed by the final conclusion.

Keywords: Tracking. Radar. Filters. Kalman.

Sumário

1 INTRODUÇÃO.....	7
2 NOÇÕES GERAIS.....	10
2.1 HISTÓRICO.....	10
2.2 TIPOS DE RADAR.....	17
2.2.1 Sistemas de transmissão	17
2.2.2 Classificação por aplicação	18
3 FILTROS.....	22
3.1 FILTROS PASSIVOS.....	22
3.2 FILTROS ATIVOS.....	23
3.3 FILTROS DIGITAIS.....	23
3.3.1 Filtros digitais FIR	25
3.3.2 Filtros digitais IIR	27
3.3.3 Filtros digitais adaptativos.....	29
4 FILTRO DE KALMAN.....	31
4.1 PRESUPOSIÇÕES BÁSICAS.....	35
4.2 UM EXEMPLO SIMPLES	38
4.3 O FILTRO DISCRETO DE KALMAN.....	46
4.4 O ALGORITMO DO FILTRO DISCRETO DE KALMAN.....	49
5 RASTREAMENTO DE ALVOS.....	53
5.1 O QUE É RASTREAMENTO	53
5.2 INTRODUÇÃO AO RASTREAMENTO DE ALVOS.....	54
5.3 RASTREAMENTO DE ALVOS EM UM SISTEMA RADAR	56
5.3.1 O filtro $\alpha - \beta$	58
5.3.2 O filtro de Kalman	64
6 CONCLUSÃO	73
REFERÊNCIAS BIBLIOGRÁFICAS.....	75

1 INTRODUÇÃO

Atualmente, devido ao constante avanço e crescimento da malha aeroviária mundial, torna-se necessário que a segurança apresentada pelos sistemas de proteção ao vôo acompanhe esta evolução.

Os sistemas de radar são os responsáveis pelo rastreamento das aeronaves em seu trajeto. Desta maneira, estes sistemas apresentam grande destaque na segurança aeroviária, e a informação da trajetória dos alvos deve ser apresentada com a máxima precisão, pois toda a segurança do tráfego aéreo depende destas informações.

Um sistema de rastreamento radar deve ser capaz de medir coordenadas de um determinado alvo e fornecer dados que sejam capazes de determinar a trajetória do mesmo, além de predeterminar sua posição futura.

A aplicação de técnicas de processamento digital de sinais no tratamento de dados obtidos em um sistema radar, juntamente com a aplicação de algoritmos de filtros adaptativos, com ênfase especial em filtros de Kalman, possibilitaram o desenvolvimento de sistemas mais eficientes que os antigos sistemas analógicos, aumentando a confiabilidade da navegação aérea.

Um alvo deve ser apresentado de forma correta na tela de um controlador de voo. Uma apresentação incorreta induz o indivíduo a erros que podem resultar em acidentes de grandes proporções, inclusive com perda de vidas humanas.

O correto rastreamento dos alvos é então indispensável para o bom funcionamento do sistema de navegação aérea em uma determinada região.

A aplicação de filtros de Kalman, em sistemas de rastreamento radar, reduz consideravelmente a probabilidade de associações incorretas entre os alvos

detectados, o que reduz de forma significativa o erro humano, e torna a navegação aérea mais segura e confiável.

O presente trabalho tem por objetivo demonstrar a aplicação de filtros adaptativos por intermédio de explicações matemáticas, com ênfase em filtros de Kalman, no rastreamento de alvos em sistemas RADAR.

Com o crescente aumento da demanda por melhores condições de segurança e modernização do sistema aeroportuário, principalmente por parte dos usuários do sistema, se torna imperativo que novos estudos sejam realizados de modo a tornar isto viável.

Como os sistemas radar são responsáveis por uma parte significativa desta segurança, o estudo de métodos que tornem estes sistemas mais confiáveis serão sempre uma constante.

A função de rastreamento e predição da posição de um alvo é indispensável na composição de um panorama de segurança dentro de um sistema de navegação aérea em qualquer parte do mundo.

A introdução dos filtros adaptativos representou um grande avanço nos métodos de rastreamento e predição de alvos radar. Em particular, o filtro de Kalman que apresenta grande confiabilidade tanto no rastreamento, quanto na predição de alvos.

Estudos sobre aplicações do referido recurso devem sempre ser implementados, a fim de se possibilitar um melhor aproveitamento das características funcionais destes filtros, conduzindo assim a um aumento da eficiência do sistema como um todo, que na verdade se resume a reduzir a probabilidade de acidentes aéreos.

Na primeira parte deste estudo, será realizado um levantamento histórico dos sistemas radar, seus tipos e utilização, desde sua concepção até os sistemas em uso na atualidade, levando-se em conta sua evolução tecnológica.

Na segunda parte serão abordados os filtros: conceitos, tipos e formas.

Na terceira parte, será discutido o filtro de Kalman em suas várias formas, sua descrição física e matemática além de seu algoritmo.

Na quarta parte será abordado o rastreamento de alvos, de que forma é processado e sua aplicação em sistemas radar.

A quinta e última parte será referente a conclusão, baseada no estudo matemático e na bibliografia utilizada.

2 NOÇÕES GERAIS

Radar, palavra derivada da compressão do termo em inglês *Radio Detection and Ranging*, é um sistema eletromagnético para a detecção e localização de objetos. Ele opera no princípio da transmissão de um tipo particular de forma de onda, uma onda seno, por exemplo, modulada em forma de pulso e detecta o eco do sinal transmitido. O radar é usado para ampliar os sentidos de um observador que se propõe a observar o meio ambiente que o cerca, especialmente o sentido da visão. O valor do radar se encontra não no fato de ser um substituto da visão, mas sim por ser capaz de fazer o que os olhos não podem. Um radar pode não ser capaz de distinguir cores, como os olhos humanos, contudo pode ser projetado de modo a enxergar sob condições que o olho não consegue, como na escuridão, neblina, chuva ou até mesmo neve. E ainda mais, o radar tem a capacidade de medir a distância de um determinado objeto. Este é provavelmente seu maior atributo (SKOLNIK, 1980).

2.1 HISTÓRICO

De acordo com Skolnik, embora o desenvolvimento completo da tecnologia radar não tenha ocorrido até a II Guerra Mundial, o princípio básico da detecção radar é quase tão antigo quanto o próprio eletromagnetismo. Em 1886, Heinrich Hertz testou experimentalmente as teorias de Maxwell e demonstrou a similaridade entre as ondas de rádio e a luz. Hertz mostrou que as ondas de rádio poderiam ser refletidas por corpos dielétricos e metálicos.

Marconi reconheceu o potencial das ondas curtas para detecção por ondas de rádio e em 1922 tentou pressionar a comunidade científica, para este uso das ondas de rádio. Em um discurso diante de uma plateia no Institute of Radio Engineers, ele se pronunciou:

Como já demonstrado por Hertz, as ondas elétricas podem ser refletidas completamente por corpos condutores. Em alguns dos meus testes, notei os efeitos de reflexão e detecção destas ondas em objetos metálicos à milhas de distância.

Parece-me que seria possível o desenvolvimento de um dispositivo de modo que um navio pudesse irradiar ou projetar um feixe divergente destes raios em qualquer direção desejada, e caso estes raios se encontrassem com outro objeto metálico, como um vapor ou outro navio, seriam refletidos de volta ao local de transmissão no navio transmissor e revelariam imediatamente a presença e a direção deste outro navio, mesmo em nevoeiro ou tempo muito ruim (MARCONI, 1922 apud SKOLNIK, 1980).

Segundo Skolnik, embora Marconi tenha predito e demonstrado com sucesso a comunicação via rádio entre continentes, aparentemente não conseguiu apoio para essa outra ideia.

A primeira aplicação da técnica de pulso como meio de se medir distâncias, foi por intermédio da investigação científica de Breit e Tuve em 1925, quando mediram a altura da ionosfera. Entretanto, mais de uma década se passou para que a detecção de uma aeronave por um pulso radar fosse demonstrada (SKOLNIK,1980).

O princípio por trás de um sistema de radar se baseia na propagação de ondas eletromagnéticas, ou mais precisamente, ondas de rádio frequência, como foi descrito por Hugo Gernsback em 1911. Em 1934, o cientista francês Pierre David usou o princípio radar pela primeira vez para detectar com sucesso uma aeronave. Em 1935 Maurice Ponte e Henri Gutton durante testes realizados a bordo do navio *Orégon*, parte da frota da Compagnie Générale Transatlantique, conseguiram detectar icebergs utilizando-se de ondas com comprimento de 16 cm. Em 1936 Professor Kunhold da Alemanha, também conseguiu detectar aeronaves (LACOMME et al, 2001).

O primeiro radar experimental operava em onda contínua (CW) e sua detecção se baseava na interferência proveniente entre o sinal vindo direto do transmissor e o sinal deslocado por intermédio de efeito Doppler refletido em um alvo móvel. Este tipo de radar foi chamado de Radar CW de interferência de onda. Atualmente este tipo de radar tem o nome de Radar CW bi estático. As primeiras detecções de aeronaves usaram este princípio, ao invés de um radar de pulso, pois era uma tecnologia largamente usada na época. Os radares de pulso tiveram de aguardar o desenvolvimento de novas tecnologias antes de serem aplicados de fato (SKOLNIK,1980).

Ainda de acordo com Skolnik, a primeira detecção de uma aeronave utilizando o conceito de interferência de onda se deu em junho de 1930 por H. A. Hyland do Naval Research Laboratory (NRL). A detecção se deu por acaso enquanto Hyland trabalhava em um aparato indicador de direção de uma aeronave

no solo. O transmissor operava a uma frequência de 33 MHz e o feixe de ondas atingiu um campo de pouso a duas milhas de distância. Hyland percebeu que o sinal recebido pelo receptor aumentava consideravelmente quando uma aeronave atravessava o feixe de ondas. Este fato estimulou novas pesquisas pelo pessoal do NRL, mas estas caminhavam a passos lentos por falta de apoio oficial. Em 1932, o NRL fez a apresentação de um equipamento de CW capaz de detectar uma aeronave a distância de 50 milhas do transmissor. O trabalho do NRL foi considerado como confidencial pelo Governo americano, até que em 1933 vários engenheiros do Bell Laboratories relataram a detecção de aeronaves durante a realização de um outro experimento. O trabalho do NRL foi revelado e patenteado por Taylor, Young e Hyland. O tipo de radar descrito nesta patente é do tipo CW de interferência de onda.

Os primeiros radares CW de interferência de onda se limitavam somente a detecção de presença de um alvo. O problema de se extrair a informação de posição do alvo com este tipo de radar era extremamente difícil e não podia ser solucionado por meio da tecnologia existente no período. A limitada capacidade do equipamento CW que era comparado a uma armadilha de fio esticado, cujo acionamento depende da interferência direta do alvo, não acrescentava nenhum entusiasmo a procura de novas soluções ao problema (SKOLNIK,1980).

Era fato que a utilização de transmissão em forma de pulsos resolveria o problema. Entretanto, havia muito ceticismo na época sobre a técnica de radar pulsado. Mesmo assim o NRL, na primavera de 1934, iniciou um projeto de radar de pulsos (SKOLNIK,1980).

A primeira experiência com um radar de pulsos ocorreu em dezembro de 1934. O radar operava na frequência de 60 MHz. Os testes se mostraram um fracasso e nenhum alvo foi detectado. A principal razão apontada para a falha era que os receptores utilizados haviam sido projetados para a aplicação em comunicações em onda contínua. O projeto então passou por correções e as primeiras detecções utilizando-se da tecnologia de radar de pulsos foram obtidas em abril de 1936, com um radar operando na frequência de 28,3 MHz e uma largura de

pulso de cinco μ s atingindo um alcance de 2,5 milhas. Em julho do mesmo ano, este alcance era de 25 milhas (SKOLNIK,1980).

O NRL desejava que o radar operasse em frequências mais altas, pois sua principal utilização seria a bordo de navios, onde não seria possível a utilização de grandes antenas. O sucesso do modelo de 28,3 MHz permitiu o desenvolvimento de um modelo de 200 MHz. Apenas três meses após o início do projeto o radar de 200 MHz estava operacional. Este radar também foi o primeiro a se utilizar de duplexação, onde o transmissor e o receptor se utilizavam da mesma antena para a operação (SKOLNIK,1980).

Com o desenvolvimento de válvulas de alta potência, o radar de 200 MHz pode ser melhorado, melhorando seu alcance atingindo a marca de 50 milhas, o limite permitido por esta frequência. Em outubro de 1939, foi desenvolvida uma versão comercial deste radar, e em 1941 foram instalados 19 destes em navios de guerra americanos (SKOLNIK,1980).

Na Inglaterra o desenvolvimento da tecnologia de radares foi mais lento que nos Estados Unidos inicialmente, mas com a iminência de uma guerra à sua porta, os ingleses aplicaram grande parte de seus esforços ao desenvolvimento do radar. Quando os americanos entraram na guerra, os ingleses já tinham uma larga experiência com equipamentos de radar (SKOLNIK,1980).

Em junho de 1935, os britânicos demonstraram a técnica de medida de distância de uma aeronave por intermédio de pulsos. Foi quase um ano mais cedo que a experiência bem sucedida com radar de pulso do NRL. Em setembro, distâncias de mais de 40 milhas foram atingidas em um equipamento instalado em um bombardeiro. Neste mesmo mês, foi realizada a primeira medida da altura de uma aeronave por meio da medida do ângulo de elevação do sinal refletido. Em março de 1936, o alcance obtido chegou a 90 milhas (SKOLNIK,1980).

Uma série de estações radar, operando a frequência de 25 MHz, foram instaladas com sucesso em abril de 1937. Em setembro de 1938 praticamente todas as estações se encontravam operacionais, e rastream o avião de Neville

Chamberlain, o Primeiro Ministro inglês na época, até Munique para uma conferência com Hitler e Mussolini. No mesmo mês, as estações passaram operar em regime de vigilância de 24 horas ininterruptas e continuaram desta forma até o fim da guerra (SKOLNIK, 1980).

Os britânicos chegaram logo à conclusão de que somente as estações de terra não apresentavam precisão suficiente para guiar um caça a seu alvo de interceptação em condições noturnas ou de mau tempo. Em consequência, eles desenvolveram um radar de detecção de aeronaves montado em uma aeronave, com o intuito de detecção e interceptação de aeronaves hostis. Este radar operava com a frequência de 200 MHz. Durante o desenvolvimento, notou-se que o radar também podia ser utilizado para a interceptação de navios do ar e que a característica do eco recebido variava de acordo com o tipo de terreno. Logo esta característica foi utilizada para interceptação de submarinos. O segundo efeito não foi inicialmente explorado, mas usado posteriormente em radares de mapeamento e cartografia (SKOLNIK, 1980).

Até meados de 1940, o desenvolvimento de sistemas de radar da Inglaterra e dos Estados Unidos, ocorreram de forma independente. Em setembro de 1940, uma missão técnica dos ingleses visitaram os Estados Unidos com o intuito de trocar informações sobre o desenvolvimento de sistemas de radar nos dois países. Os britânicos perceberam o ganho de se operar com uma melhor resolução angular obtido em frequências de micro-ondas, principalmente quando se tratava da área naval e a bordo de aeronaves. Foi sugerido aos americanos a criação de um radar de micro-ondas para uso embarcado em aeronaves e um radar de micro-ondas para controle de dispositivos antiaéreos (SKOLNIK, 1980).

A missão apresentou aos americanos a válvula de potência magnetron, desenvolvida por Randell e Boot, e também forneceu informações técnicas para que os americanos pudessem fabricá-la. A magnetron desenvolvida por Randell e Boot operava com um comprimento de onda de 10 cm com uma potência de saída de 1 kW. A válvula apresentava um ganho de 100 vezes em relação a outros dispositivos de micro-ondas existentes. O desenvolvimento da magnetron foi uma das contribuições mais importantes na evolução dos radares (SKOLNIK, 1980).

Além do desenvolvimento obtido por Estados Unidos e Inglaterra, o radar também foi desenvolvido de forma independente por outros países, como Alemanha, França, Rússia, Itália e Japão no final da década de trinta. Todos esses países fizeram experimentos com radares CW de interferência de onda, e apesar de França e Japão torná-los operacionais, mostraram-se de uso limitado. Todos esses países acabaram progredindo para a forma pulsada de radar pelas vantagens inerentes ao sistema. Embora se reconhecesse a vantagem de se operar com frequências elevadas, durante a guerra, nenhum radar foi desenvolvido com frequências maiores do que 600 MHz, com exceção de Estados Unidos e Inglaterra (SKOLNIK, 1980).

O radar foi desenvolvido simultaneamente e independentemente em vários países, principalmente durante a 2ª Guerra Mundial. Não é possível apontar um único indivíduo como criador; o radar tem vários colaboradores em comum. Foi criado não somente pelo momento favorável da tecnologia de radiocomunicação, mas sim pela maturação da aviação no mesmo período e pelo reconhecimento da ameaça militar que representava e a necessidade de se defender desta ameaça (SKOLNIK, 1980).

No Brasil, após a sua participação e o término da 2ª Guerra Mundial, as Forças Armadas Brasileiras obtiveram equipamentos de radar, o Exército recebeu equipamento de campanha transportável, para uso da artilharia e a Marinha um conjunto de radares e sonares, que vieram já instalados nos primeiros navios incorporados a frota (JUCEWICZ, 1997).

A aviação no Brasil até a época da guerra não dispunha de nenhum auxílio à navegação por rádio. A primeira aeronave que deveria ser equipada com um rádio goniômetro foi adquirido pela Varig em 1942. Era um bimotor de fabricação inglesa, que inaugurou a primeira linha internacional, ligando Porto Alegre a Montevideu. Além de rádio telegrafia, fazia parte do equipamento, um acessório de navegação aérea, um rádio goniômetro, que não chegou a ser instalado, pois não podia ser utilizado por falta de equipamentos de solo coadjuvantes nesta rota. Somente em 1943, com a compra dos bimotores americanos Electra das séries 10A

e 10E, pela Varig, iniciaram-se os voos por instrumento, utilizando-se o rádio goniômetro (JUCEWICZ,1997).

O primeiro radar de terra, destinado ao controle de tráfego aéreo do tipo ASR (Airport Surveillance Radar), radar de vigilância de aeroporto, foi instalado em 1959, no Aeroporto de Congonhas, em São Paulo, onde se concentrava o maior volume de tráfego aéreo na época. Era utilizado no controle de aproximação com alcance de 60 milhas náuticas. Esse radar era operado por técnicos civis e militares do Ministério da Aeronáutica, e sua manutenção estava a cargo de especialistas treinados na fábrica do equipamento (JUCEWICZ,1997).

A vetoração pelo radar, naquela época, era experimental ficando a cargo do comandante a aeronave que entrasse na área de cobertura do equipamento, utilizar-se ou não deste novo auxílio a navegação (JUCEWICZ,1997).

O segundo radar de terra, do tipo ASR, foi instalado no Aeroporto do Galeão, no Rio de Janeiro, no início de 1961, mas só veio a operar em fevereiro de 1964, por falta de especialistas em manutenção do equipamento. Seguiu-se a instalação, em 1966, de um modelo mais moderno de radar ASR no Aeroporto de Congonhas e no Aeroporto Salgado Filho em Porto Alegre, no ano de 1967. Esses radares operavam dentro de um raio de 60 milhas náuticas (padrão utilizado até hoje) e os ecos apresentados eram proporcionais ao tamanho da aeronave detectada (JUCEWICZ,1997).

Os anos 60 introduziram no Brasil duas evoluções paralelas, que influenciaram o transporte aéreo: o jato puro, e o transistor nos equipamentos eletrônicos. Os equipamentos de radar transistorizados passaram a ter menor tamanho, peso e menor consumo de energia, permitindo a sua expansão no país (JUCEWICZ,1997).

2.2 TIPOS DE RADAR

Não existe uma forma única de se classificar um radar. Os radares podem ser classificados de acordo com várias de suas características, como por exemplo, o seu princípio de operação ou então pela forma de utilização do equipamento.

Neste trabalho serão consideradas as características predominantes mais utilizadas na classificação dos equipamentos de radar.

2.2.1 Sistemas de transmissão

Os radares podem ser classificados do seguinte modo de acordo com a forma de transmissão utilizada:

1. Radar de onda contínua (CW);
2. Radar de FM (Frequência Modulada);
3. Radar de pulsos.

Os radares de onda contínua são usados basicamente para determinar a velocidade e aceleração de objetos móveis. Este tipo de radar utiliza uma forma de onda do tipo seno, de forma quase contínua, o equipamento é desligado por um intervalo de tempo muito pequeno em relação ao período de operação, a fim de salvaguardar o instrumento, operando então praticamente de forma contínua, como sugere a sua classificação. Quase sempre se utiliza da variação da frequência com a velocidade, conhecido como efeito Doppler, para medir a velocidade relativa de um alvo.

Sua principal aplicação é nas rodovias, no controle de velocidade dos automóveis a fim de diminuir os acidentes. Não é utilizado na área de navegação aérea.

Os radares de FM utilizam a variação de frequência com o intuito de medir a distância de um alvo.

Suas principais aplicações são a medida de altura de uma aeronave em relação ao solo (rádio altímetro) e no levantamento topográfico de uma determinada região.

Os radares de pulso são um tipo que emite uma série cíclica de pulsos quase retangulares. É capaz de fornecer a posição angular de um alvo em um determinado instante, assim como sua velocidade, por intermédio do efeito Doppler, sendo largamente utilizados nos sistemas de proteção ao voo, graças a estas características.

2.2.2 Classificação por aplicação

De acordo com sua aplicação os radares podem ser classificados do seguinte modo:

1. Navegação aeronáutica
2. Radar naval
3. Radar militar
4. Radar meteorológico
5. Controle de tráfego aéreo

Os radares de navegação aeronáutica são os do tipo utilizado nas aeronaves a fim de se evitar colisões com outras aeronaves, radares do tipo meteorológico, com o intuito de monitorar as condições climáticas, ou qualquer outro tipo de radar de uso embarcado na aeronave.

Sua principal diferença entre os outros tipos de radar é seu porte reduzido, pois em termos de operação, funciona de forma semelhante a outros sistema de radar de maior porte e potência.

Os radares navais são os utilizados em embarcações, com objetivos e aplicações semelhantes aos de navegação aeronáutica, sendo que, a diferença básica é que se encontram em navios, sejam comerciais, sejam militares.

Os radares militares, são qualquer tipo de radar de uso das Forças Armadas. Na verdade, o radar nasceu com o intuito de ser um equipamento de uso militar, mas devido a sua versatilidade foi incorporado a outras atividades de uso cotidiano, se adaptando as mais diversas tarefas.

Todo e qualquer radar que tem por objetivo a monitoração das condições climáticas é considerado um radar meteorológico. Podem ser utilizados no estudo da climatologia de uma determinada região, assim como podem ser utilizados em uma aeronave ou um navio, de modo a se obter informações sobre o tempo em um determinado trajeto.

Os radares de controle de tráfego aéreo são utilizados na navegação aérea em todos os níveis. Desde o controle da torre, para o pouso e a decolagem das aeronaves, até a vigilância das mesmas em todo o território nacional e internacional. Teoricamente, aonde quer que uma aeronave esteja, ela está sendo monitorada por um radar qualquer. Na prática isto não funciona muito bem devido ao custo de instalação e manutenção do sistema. Nos países menos desenvolvidos, a cadeia de radares se resume aos grandes centros e regiões, deixando algumas brechas em áreas mais isoladas.

Os radares de controle de tráfego aéreo podem ainda ser divididos em dois grupos:

1. Radar primário
2. Radar secundário

Os radares primários recebem esta nomenclatura por não necessitarem do auxílio de um equipamento externo para apresentar um alvo. Uma aeronave ao alcance de um radar primário será detectada independentemente da anuência da mesma.

O radar secundário tem esta denominação pelo fato de necessitar que um equipamento a bordo da aeronave seja ativado, o transponder, para que o alvo seja apresentado na tela de um monitor.

A forma de transmissão do radar secundário difere da do radar primário. Enquanto um radar primário emite um pulso de alta frequência e curta duração e espera o retorno deste pulso para indicar ou não a presença de um alvo, o radar secundário emite três pulsos, com uma codificação específica e recebe do transponder um trem de doze pulsos, com informações da aeronave, como por exemplo, sua identificação, se ela é civil ou militar e a sua velocidade.

Caso uma aeronave não queira ser identificada, ela pode simplesmente desligar o equipamento de bordo e então se tornará “invisível” ao radar secundário, mas ainda assim será captada pelo radar primário, sem as informações fornecidas pelo transponder. Desta maneira não existe operação de navegação aérea baseada somente no radar secundário, pois caso o transponder esteja desligado, não haverá presença de alvo na tela de um controlador de tráfego aéreo.

Os radares primários de navegação aérea podem ainda ser subdivididos em:

1. Radar de rota
2. Radar de aeroporto
3. Radar de superfície ou radar de solo
4. Radar PAR

O radar de rota, fornece informações sobre uma aeronave em rota, como o próprio nome já diz , é instalado afastado dos grandes centros e oferece cobertura de aeronaves a nível nacional, pois tem grande alcance, de cerca de 120 milhas náuticas de raio, que corresponde a aproximadamente 200 quilômetros de raio.

O radar de aeroporto, como o próprio nome sugere, é instalado nas proximidades de um aeroporto, oferecendo informações sobre aeronaves a nível

regional, tem alcance médio de cerca de 60 milhas náuticas, ou aproximadamente 100 quilômetros.

O radar de superfície, é instalado dentro das dependências do aeroporto, e fornece informações das aeronaves que se encontram dentro das pistas de rolamento do aeródromo. Sua principal utilidade é auxiliar o controle da torre em condições adversas, como nevoeiro, chuva ou qualquer outro fator que dificulte a visualização das pistas, de modo a se evitar um incidente.

O radar PAR (sigla da expressão em inglês Precision Approach Radar), é um tipo diferente de radar com duas antenas posicionadas perpendicularmente, que tem por objetivo informar o melhor ângulo de descida e o centro da pista a uma aeronave. Esta informação é passada de um controlador em terra para o piloto da aeronave. Atualmente, este tipo de radar é mais utilizado nos aeroportos civis, mas ainda é utilizado em algumas bases aérea da Força Aérea, e principalmente em porta aviões que por terem pouco espaço disponível, precisam de um equipamento de proporções não muito grandes.

3 FILTROS

Tipicamente, o conteúdo espectral de um sinal portador de informações ocupa uma faixa de frequência com extensão finita. Por exemplo, um sinal de fala essencial para comunicações telefônicas situa-se na faixa de frequência de 300 a 3100 Hz (HAYKIN, 2002).

Para extrair o conteúdo de informação fundamental de um sinal de áudio para este tipo de aplicação, necessitamos de um sistema seletivo de frequências, ou seja, um filtro que limite o espectro do sinal à faixa desejada de frequência. De fato os filtros são fundamentais pra o estudo de sinais e sistemas em termo de que todo sistema usado para processar sinais contem um filtro de algum tipo em sua composição (HAYKIN, 2002).

3.1 FILTROS PASSIVOS

Diz-se que um filtro é passivo quando a sua composição é feita inteiramente de elementos passivos de circuito, isto é, indutores, capacitores e resistores. Entretanto, o projeto de filtros passivos altamente seletivos da frequência, baseia-se exclusivamente em elementos reativos, ou seja, indutores e capacitores (HAYKIN, 2002).

Os elementos resistivos entram no projeto tão somente como uma resistência de fonte ou como uma resistência de carga ou ambas. A ordem N do filtro é determinada normalmente pela quantidade de elementos reativos existentes no filtro (HAYKIN, 2002).

Os filtros passivos ocuparam um papel predominante no projeto de sistemas de comunicação e de outros sistemas até o surgimento dos filtros ativos e dos filtros digitais na década de 60 (HAYKIN, 2002).

3.2 FILTROS ATIVOS

A combinação dos elementos que compõem um filtro passivo com um dispositivo amplificador, como por exemplo, um amplificador operacional, acrescenta à configuração de filtro passivo utilizada uma nova propriedade que é a amplificação de um sinal em uma faixa de frequência específica.

Isto leva a um novo tipo de filtro que proporciona um ganho no sinal que está sendo filtrado. Este tipo de filtro é denominado filtro ativo.

Graças às características de amplificação de um filtro ativo, vários estágios podem ser acoplados sem uma perda significativa de sinal. Da mesma forma, devido à alta impedância de entrada e baixa impedância de saída do circuito amplificador, existe pouca ou nenhuma interação entre os circuitos de entrada com a carga na saída.

Outra grande vantagem do filtro digital é a possibilidade de implementação de filtros sem o uso de indutores. Esta característica permite que os mesmos possam ser utilizados em circuitos integrados, tornando-os perfeitamente compatíveis com as modernas técnicas industriais.

Os filtros ativos são facilmente sintonizáveis e permitem que, de acordo com a configuração utilizada, variações na atenuação de saída que vão da casa de 6 dB até pelo menos 50 dB.

3.3 FILTROS DIGITAIS

No processamento analógico de sinais, tomamos um sinal que varia continuamente, representando uma quantidade física que varia continuamente, e o passamos por um sistema que modifica o sinal com certo propósito. Essa modificação também é, em geral, continuamente variável por natureza, isto é, pode ser descrita por equações diferenciais (DINIZ, 2004).

Alternativamente, no processamento digital de sinais, processamos sequências de números usando algum tipo de hardware digital. Estas sequências de números são chamadas de sinais digitais ou sinais no tempo discreto. O poder do processamento digital de sinais decorre do fato de que, uma vez que uma sequência de números esteja disponível para o hardware digital apropriado, podemos efetuar qualquer forma de processamento numérico sobre eles (DINIZ, 2004).

Em várias aplicações do processamento digital de sinais deseja-se projetar um dispositivo ou um algoritmo que realize uma operação predeterminada em um sinal de tempo discreto. Este dispositivo ou algoritmo é chamado de sistema de tempo discreto. De forma mais específica, um sistema de tempo discreto é um dispositivo ou algoritmo que opera sobre um sinal de tempo discreto chamado de entrada ou excitação, de acordo com alguma regra bem definida, a fim de gerar outro sinal de tempo discreto chamado de saída ou resposta do sistema. Em geral, um sistema é visto como uma operação ou conjunto de operações realizadas no sinal de entrada $x(n)$ de forma a se obter o sinal de saída $y(n)$. Diz-se então que o sinal de entrada $x(n)$ foi transformado no sinal de saída $y(n)$ e expressa à relação geral entre $x(n)$ e $y(n)$ como:

$$y(n) = T [x(n)]$$

onde o símbolo T denota a transformação (também chamada de operador), ou processamento realizado pelo sistema em $x(n)$ para produzir $y(n)$ (PROAKIS, 1996).

Um filtro digital trabalha com sinais de tempo discreto, enquanto que os outros tipos de filtros citados anteriormente trabalham com sinais de tempo contínuo.

Os filtros digitais são desta forma, filtros que atuam sobre sinais representados de forma digital. Irão processar um determinado sinal, produzindo um novo sinal na saída do filtro, modificando o sinal original, de acordo com um programa predefinido. Podem ser utilizado, por exemplo, na separação de dois sinais distintos, na remoção de ruídos, equalização de sinais e em uma gama enorme de aplicações distintas.

Os filtros digitais apresentam algumas vantagens em relação aos analógicos. Podem ser implementados e testados sem a necessidade de uma complexa manufatura de circuitos. São extremamente flexíveis, podem ser modificados sem a necessidade de troca de hardware. Um mesmo conjunto de circuitos pode realizar diferentes funções apenas com a mudança do software inserido na máquina. Em alguns casos, a implementação digital é mais barata do que a sua contraparte analógica. O método de processamento digital de sinais também permite a implementação de algoritmos de sinais mais complexos. Geralmente é muito difícil se conseguir que um circuito analógico realize operações matemáticas de alta precisão.

3.3.1 Filtros digitais FIR

Os filtros digitais de resposta ao impulso finita FIR (da expressão em inglês Finite-duration Impulse Response) são filtros digitais não recursivos ou causais, pois o valor do sinal de saída depende única e exclusivamente do sinal aplicado na entrada. São filtros em que a resposta ao impulso atinge o valor zero após determinado período de tempo, como o próprio nome sugere.

Os filtros FIR são lineares e invariantes no tempo – LTI (da expressão em inglês Linear Time Invariant).

Diz-se que um sistema é linear se ele satisfizer o princípio da superposição. Ou seja, a resposta de um sistema linear a uma soma ponderada de sinais de entrada é igual à mesma soma ponderada de sinais de saída, sendo que cada sinal de saída é associado a um sinal de entrada particular que age no sistema independentemente de todos os outros sinais de entrada (HAYKIN, 2002).

Diz-se que um sistema é invariante no tempo se um retardo de tempo ou um avanço no tempo do sinal de entrada levar a um deslocamento de tempo idêntico no sinal de saída. Isto implica que um sistema invariante no tempo reage de maneira idêntica, não importando quando o sinal de entrada seja aplicado. Dizendo

com outras palavras, as características de um sistema invariante no tempo não se modificam com o tempo (HAYKIN, 2002).

Em geral, um sistema FIR é descrito pela seguinte equação de diferença:

$$y(n) = \sum_{k=0}^{M-1} b_k x(n-k)$$

Onde o coeficiente b_k se relaciona diretamente com a resposta ao impulso do sistema, isto é, $b_k = h(k)$. Podemos reescrever a equação acima como:

$$y(n) = \sum_{k=0}^{M-1} h(k)x(n-k)$$

Aplicando a transformada z a esta equação, chegamos a seguinte relação entrada/saída:

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{k=0}^M b_k z^{-k} = \sum_{k=0}^M h(k)z^{-k}$$

Na prática, esta equação pode ser implementada de várias formas distintas usando como blocos básicos atrasos, multiplicadores e somadores (DINIZ, 2004).

A função de transferência tem um termo constante, todos os outros termos têm uma potência negativa de z , e todos os polos se originam no plano z . Desta forma a estabilidade do filtro está garantida. Sua resposta ao impulso tem um número finito de termos (TAN, 2008).

As operações do filtro FIR envolvem somente multiplicações das entradas do filtro por seus respectivos coeficientes e o armazenamento dos

resultados; a implementação deste tipo de filtro em aplicações de tempo real é direta (TAN, 2008).

Em alguns casos, filtros FIR são considerados ineficientes no sentido de requererem uma função de transferência de ordem alta (M) para satisfazer as exigências do sistema, quando comparada a ordem requerida por filtros digitais com resposta ao impulso de duração infinita. Contudo, filtros digitais FIR apresentam algumas vantagens quanto à implementação, tais como a possibilidade de terem fase linear exata e o fato de serem intrinsecamente estáveis, quando realizados de forma não recursiva. Além disso, a complexidade computacional dos filtros FIR pode ser reduzida se eles são implementados por meio de algoritmos numéricos rápidos, tais como a transformada rápida de Fourier (DINIZ, 2004).

3.3.2 Filtros digitais IIR

Os filtros de resposta ao impulso de duração infinita IIR (da expressão em inglês Infinite duration Impulse Response), são filtros digitais recursivos ou não causais, pois o valor do sinal de saída não depende somente das entradas passada e presente, mas depende também de entradas futuras.

Fica claro que em um sistema de processamento de sinais em tempo real, não se pode observar valores futuros do sinal, e deste modo um sistema recursivo é fisicamente impossível de ser implementado. Mas por outro lado, se o sinal for gravado de modo que o processamento ocorra em um tempo que não o real (off-line), é possível a implementação de um sistema não causal, desde que todos os valores do sinal estão disponíveis no momento do processamento. Este caso é muito comum no processamento de imagens (PROAKIS, 1996).

A equação de diferença recursiva geral apresenta o seguinte formato:

$$y(n) = \sum_{k=0}^M b_k x(n-k) - \sum_{k=1}^N a_k y(n-k)$$

O inteiro N é chamado de ordem da equação de diferença ou ordem do sistema (PROAKIS, 1996).

Esta equação expressa a saída do sistema no tempo n diretamente como a soma ponderada das saídas passadas $y(n-1)$, $y(n-2)$, ..., $y(n-N)$, assim como amostras dos sinais de entrada passado e presente. Observa-se que a fim de determinar $y(n)$ para $n \geq 0$ se faz necessário a entrada $x(n)$ para todos os $n \geq 0$, e as condições iniciais $y(-1)$, $y(-2)$, ..., $y(-N)$. em outras palavras, as condições iniciais sumarizam todo o que precisa ser conhecido sobre a história passada do sistema a fim de computar as saídas presente e futura (PROAKIS, 1996).

Aplicando a transformada z a equação temos a função de transferência do sistema dado por:

$$H(z) = \frac{N(z)}{D(z)} = \frac{\sum_{k=0}^M b_k z^{-k}}{\sum_{k=1}^N a_k z^{-k}}$$

Podemos considerar que $H(z)$ na forma acima resulta na cascata de dois filtros separados, com funções de transferência $N(z)$ e $\frac{1}{D(z)}$. O polinômio $N(z)$

pode ser reconhecido como a forma básica do filtro FIR, como foi mostrada no tópico anterior (DINIZ, 2004).

Como a função de transferência apresenta um denominador polinomial, os polos do filtro IIR devem se situar dentro do círculo unitário no plano z para se assegurar a sua estabilidade (TAN, 2008).

Comparado com o filtro FIR, o filtro IIR se apresenta em um formato bem menor. Desta maneira o filtro necessita de bem menos passos computacionais para ser definido, mas a fase linear não é facilmente obtida. Portanto, um filtro IIR é utilizado quando um filtro de tamanho reduzido se faz necessário, mas a aplicação não exige uma fase linear (TAN, 2008).

Em geral os filtros IIR são capazes de aproximar uma resposta na frequência prescrita com um número de multiplicações menor que os filtros FIR. Por este motivo, filtros IIR podem ser mais adequados a algumas aplicações práticas, especialmente aquelas envolvendo processamento de sinais de tempo real (DINIZ, 2004).

3.3.3 Filtros digitais adaptativos

A não linearidade num sistema físico é uma causa de afastamento em relação ao modelo idealizado de sistema linear invariante no tempo (LTI). Outro caso de afastamento é a variação dos parâmetros do sistema no decorrer do tempo. Essa variação pode ser devida a uma variedade de fatores físicos. Também pode haver mudanças imprevistas nas propriedades estatísticas das entradas externas e das perturbações aplicadas ao sistema, cujos efeitos podem ser considerados como mudanças no ambiente que o sistema opera. As ferramentas das teorias dos sistemas lineares invariantes no tempo convencionais, que produzem um projeto de sistema com parâmetros fixos, normalmente são inadequadas para lidar com estas situações da vida real. Para produzir um desempenho satisfatório ao longo de toda a gama de variações nos parâmetros, a abordagem preferida é usar um filtro adaptativo (HAYKIN, 2002).

Um filtro adaptativo é definido como um sistema variável no tempo que é provido com mecanismo iterativo para ajustar gradativamente seus parâmetros de modo a operar de maneira ótima de acordo com algum critério especificado (HAYKIN, 2002).

Desta forma, um filtro adaptativo é um tipo de filtro que apresenta características de auto ajuste. É capaz de ajustar seus coeficientes de forma automática a fim de se ajustar ao sinal de entrada aplicado, por meio de um algoritmo adaptativo. Os filtros adaptativos desempenham um importante papel no processamento digital de sinais moderno em áreas tão variadas quanto cancelamento de eco de sinais telefônicos, cancelamento de ruídos, equalização de canais de comunicação, sistemas de controle adaptativo, aperfeiçoamento de sinais biomédicos. Os filtros adaptativos trabalham geralmente na adaptação das

mudanças de um sinal causadas pelo ambiente, pela superposição espectral de um ruído sobre um sinal, um sinal de ruído variante no tempo ou desconhecido. Por exemplo, quando um sinal de interferência é muito forte e seu espectro se sobrepõe sobre o do sinal desejado, uma aproximação convencional irá falhar no momento de preservar as características originais do sinal e uma tentativa de se remover o ruído com a aplicação de filtros convencionais se mostrará infrutífera. Contudo, um filtro adaptativo irá obter êxito nesta empreitada (TAN, 2008).

O uso de filtros adaptativos, e suas aplicações, têm mais de duas décadas de uso na comunidade científica, e ainda se encontra em expansão nesta mesma comunidade (TAN, 2008).

4 FILTRO DE KALMAN

A primeira pergunta a ser feita é: o que é um filtro de Kalman? Um filtro de Kalman simplesmente é um algoritmo de processamento de dados recursivos ótimo. Existem várias maneiras de se definir o que vem a ser ótimo, dependendo de qual critério será escolhido para avaliar o desempenho. Será mostrado que, o filtro de Kalman é ótimo no que diz respeito a praticamente qualquer critério que faça sentido. Um aspecto dessa qualidade de ser ótimo é que o filtro de Kalman incorpora todas as informações que possam ser fornecidas a ele. O filtro processa todas as medidas disponíveis, não importando a precisão, de modo a estimar o valor corrente das variáveis que interessam, com o uso do conhecimento do sistema e da dinâmica dos dispositivos de medida, da descrição estatística dos ruídos do sistema, medida dos erros, das incertezas dos modelos dinâmicos, e de qualquer informação disponível sobre as condições iniciais das variáveis de interesse. Por exemplo, para medir a velocidade de uma aeronave, poderia ser usado um radar Doppler, ou as indicações de velocidade de um sistema de navegação inercial, ou um pitot e a pressão estática e a informação do vento relativa do sistema de dados do ar. Ao invés de ignorar alguma dessas informações, um filtro de Kalman poderia ser construído de forma a combinar todos esses dados e conhecimento das várias dinâmicas do sistema de modo a gerar a melhor estimativa global de velocidade (MAYBECK, 1979).

A palavra recursivo na descrição prévia significa que apesar de certos conceitos de processamento de dados, o filtro de Kalman não necessita que todos os dados anteriores estejam memorizados e sejam reprocessados toda vez que uma nova medida é feita. Isto é de vital importância para a implementação do filtro de forma prática (MAYBECK, 1979).

O “filtro” é na realidade um algoritmo de processamento de dados. Apesar de se usar a típica conotação de filtro como sendo uma “caixa preta” que contém circuitos eletrônicos, o fato é que na maioria das aplicações práticas, o “filtro” é somente um programa de computador em um processador central. Desta feita, o

filtro incorpora inerentemente amostras de medidas de tempo discreto ao invés de entradas de tempo contínuo (MAYBECK, 1979).

A figura 4.1 apresenta uma típica situação na qual um filtro de Kalman poderia ser usado de forma vantajosa. Um sistema de um tipo qualquer é guiado por controles conhecidos, e dispositivos de medida fornecem o valor de certas quantidades pertinentes. O conhecimento destas entradas e saídas do sistema é tudo que está explicitamente disponível do sistema físico para propósitos de estimativa (MAYBECK, 1979).

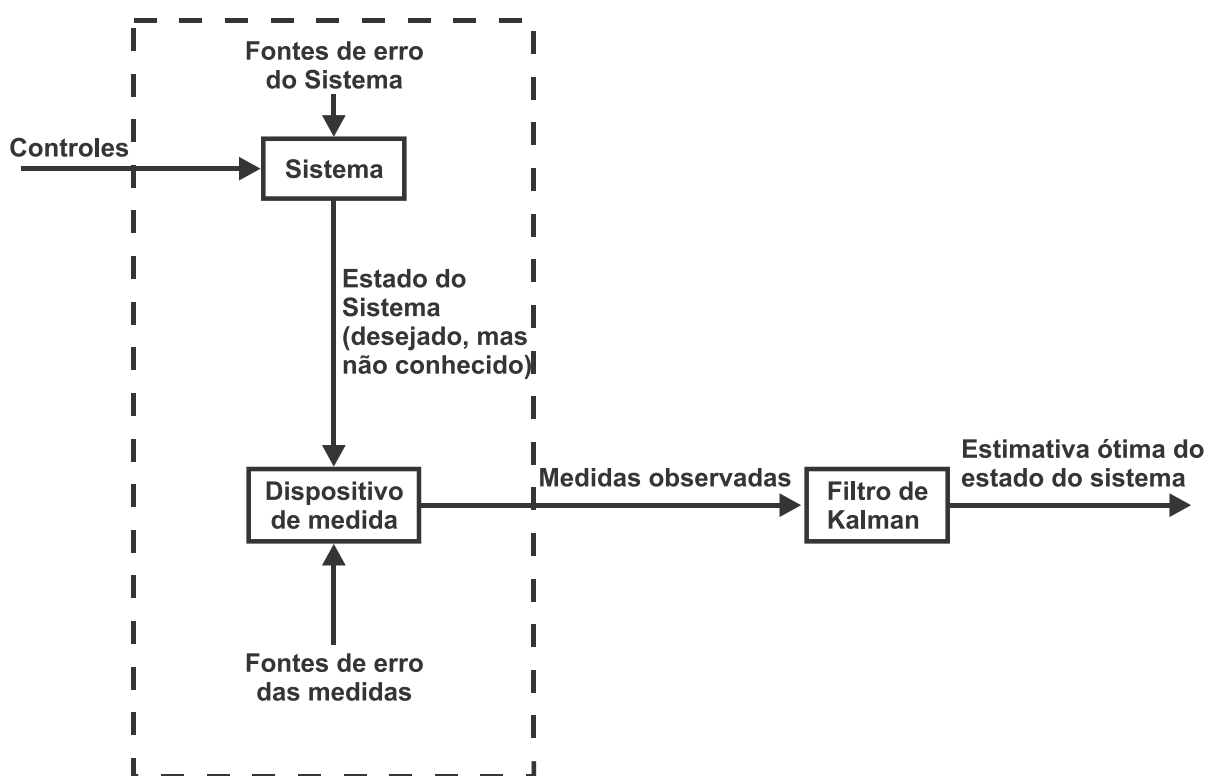


Figura 4.1

A necessidade de um filtro então se torna aparente. Com frequência as variáveis de interesse, alguns números finitos de quantidades que descrevem o “estado” do sistema, não podem ser medidas diretamente, e algum meio de se inferir esses valores por meio dos dados disponíveis devem ser gerados. Por exemplo, um sistema de dados do ar fornece diretamente a pressão estática e a de pitot, das quais medidas, a velocidade deve ser gerada. Esta geração de dados se complica pelo fato de que o sistema é tipicamente mais influenciado por suas entradas do que pelos controles conhecidos e as relações entre as várias variáveis de “estado” e as

saídas medidas são conhecidas com certo grau de incerteza. Além disso, qualquer medida realizada será corrompida por um grau de ruído, distorção, e imprecisão dos dispositivos, e desta maneira um meio de se conseguir extrair informações válidas por meio de uma fonte ruidosa deve ser fornecido igualmente. Podem existir ainda vários dispositivos de medidas, cada qual com sua dinâmica e erro característico, que fornecem alguma informação sobre uma variável particular, e seria desejável a combinação dessas saídas de uma maneira sistemática e ótima. Um filtro de Kalman combina todos os dados medidos disponíveis, com o conhecimento do sistema e dos dispositivos de medidas, a fim de fornecer uma estimativa das variáveis desejadas de tal maneira que o erro é minimizado estatisticamente. Em outras palavras, se fossem rodados um certo número de filtros candidatos várias vezes para a mesma aplicação, a média dos resultados do filtro de Kalman seria melhor do que o resultado médio de qualquer outro filtro (MAYBECK, 1979).

Conceitualmente, o que qualquer tipo de filtro faz é tentar obter uma estimativa “ótima” de quantidades desejadas por meio de dados fornecidos por um ambiente ruidoso, com “ótima” significando que o filtro minimiza os erros em algum aspecto. Existem várias maneiras de se atingir este objetivo. Se for adotado um ponto de vista Bayesiano, então o que se deseja é que o filtro propague a densidade de probabilidade condicional da quantidade desejada, condicionado ao conhecimento dos dados reais que vem dos dispositivos de medida. Para se entender esse conceito, considere a figura 4.2, uma representação de uma densidade de probabilidade condicional do valor de uma quantidade escalar x no instante de tempo i ($x(i)$), condicionada ao conhecimento de que o vetor de medição $z(1)$ no instante de tempo 1 assume o valor z_1 ($z(1) = z_1$) e de forma similar para os instantes 2 até i , plotados como uma função de possíveis valores de $x(i)$. Isto é indicado como $f_{x(i)|z(1),z(2),\dots,z(i)}(x|z_1, z_2, \dots, z_i)$. Por exemplo, assumindo que $x(i)$ seja a posição unidimensional de um veículo no instante de tempo 1, e assumindo que $z(j)$ seja um vetor bidimensional que apresenta as medidas de posição obtidas por dois radares distintos no tempo j . Esta densidade de probabilidade condicional contém todas as

informações disponíveis sobre $x^{(i)}$: isto indica, para o valor dado de todas as medidas assumidas no instante de tempo i , quais seriam as probabilidades de $x^{(i)}$ assumindo qualquer valor particular ou intervalo de valores (MAYBECK, 1979).

O termo “condicional” em densidade de probabilidade condicional é denominado desta forma porque sua forma e localização no eixo x depende dos valores medidos. Sua forma transporta a quantidade de certeza que se tem no conhecimento do valor de x . Se a densidade plotada for um pico estreito, então o “peso” da probabilidade está concentrado em uma faixa estreita de valores de x . Por outro lado, se a forma plotada for gradual, o “peso” da probabilidade está espalhado por uma larga faixa de valores de x , indicando que existe uma maior incerteza do valor (MAYBECK, 1979).

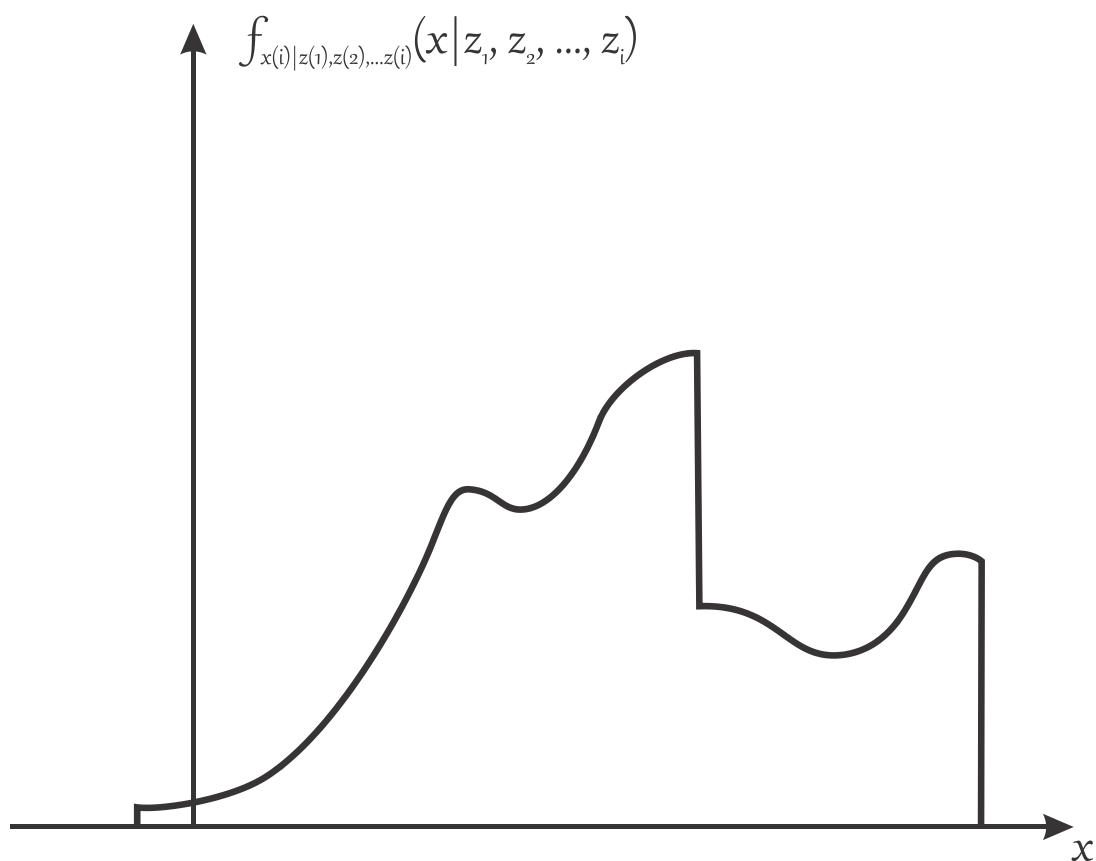


FIGURA 4.2

4.1 PRESUPOSIÇÕES BÁSICAS

Neste ponto, é necessário se observar as três pressuposições básicas na formulação do filtro de Kalman. Em uma primeira inspeção, podem parecer excessivamente restritivas e não realistas. A fim de dirimir qualquer receio, será discutido de forma breve as implicações físicas destas pressuposições (MAYBECK, 1979).

Um modelo de sistema linear se justifica por várias razões. Frequentemente, este modelo é o mais adequado para um propósito imediato, e embora existam não linearidades, a abordagem típica da Engenharia é linearizar em algum ponto nominal ou trajetória, gerando um modelo de perturbação ou modelo de erro. Sistemas lineares são desejáveis, pois são muito mais amigáveis e mais facilmente manipulados pelas ferramentas da Engenharia, e a teoria de sistemas lineares (equações diferenciais) é muito mais completa e prática do que a não linear. O fato é que não existem meios de se estender o conceito de filtro de Kalman para aplicações não lineares ou desenvolver filtros não lineares de forma direta, mas isso seria levado em consideração se um modelo linear se provasse inadequado (MAYBECK, 1979).

Densidade de ruído branco implica que o valor do ruído não está correlacionado no tempo. Expondo de forma mais simples, se o valor instantâneo do ruído é conhecido, este conhecimento não influi na predição de qual será seu valor em qualquer outro instante. Densidade de ruído branco também significa que o ruído tem uma potencia igual em todas as frequências. Como isso implica em um ruído com potencia infinita, um ruído branco não pode realmente existir. Porque então se considerar este conceito se ele realmente não existe? A resposta tem duas partes. Primeiro, qualquer sistema físico tem uma certa frequência, uma banda passante, uma faixa de frequências de entrada na qual o sistema irá responder. Fora desta faixa ou a entrada não tem efeito, ou será tão severamente atenuada que praticamente é como se não existisse. Na figura 4.3 uma típica curva da banda passante de um sistema na forma da densidade espectral de potência (interpretada como a quantidade de potência debitada em uma certa frequência) pela própria frequência. Tipicamente um sistema será acionado por um ruído de banda larga, que tem potencia em frequências muito além da banda do sistema, e potencia

praticamente constante em toda a banda passante do sistema. Neste mesmo gráfico, um ruído branco meramente extenderia este nível de potencia constante à todas as frequências. Dentro da banda passante do sistema de interesse, o fictício ruído branco se apresenta de forma idêntica ao ruído de banda larga. Então qual foi o ganho? Esta é a segunda parte da resposta porque um modelo de ruído branco é utilizado. Acontece que a matemática utilizada no filtro é amplamente simplificada (na verdade, se torna maleável) na substituição do ruído de banda larga pelo ruído branco, que pelo ponto de vista do sistema funciona de forma idêntica. Portanto, o modelo de ruído branco é o usado (MAYBECK, 1979).

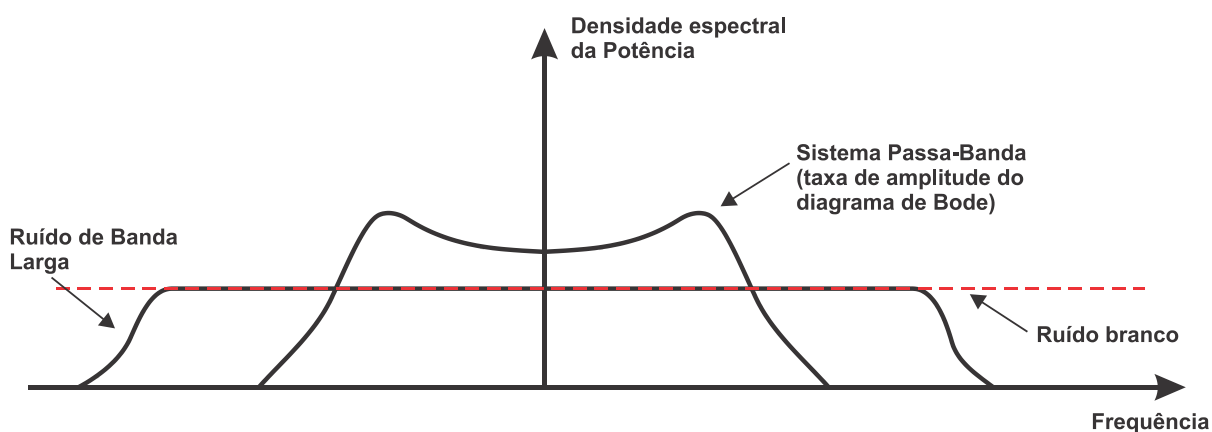


Figura 4.3

Pode-se argumentar que existem casos em que a potencia do ruído não é constante em todas as frequências da banda passante do sistema, ou que o ruído é de fato, correlacionado no tempo. Um pequeno sistema chamado de “filtro de modelagem”, é então adicionado ao sistema original, de modo a se obter um sistema linear global guiado pelo ruído branco novamente (MAYBECK, 1979).

Enquanto o ruído branco relaciona-se com tempo e frequência, o Gaussiano se relaciona com a amplitude. Portanto, em qualquer ponto no tempo, a distribuição de densidade de probabilidade da amplitude de um ruído Gaussiano assume a forma de uma curva normal em forma de sino. Esta suposição pode ser justificada fisicamente pelo fato que um ruído, medido ou de sistema, é tipicamente causado por várias pequenas fontes. Pode ser demonstrado matematicamente

que quando várias variáveis aleatórias independentes são adicionadas, o resultado da soma pode ser descrito de forma muito similar por uma densidade de probabilidade Gaussiana, não importando o formato das densidades individuais (MAYBECK, 1979).

Também existe uma justificativa prática para o uso de densidades Gaussianas. De forma similar a do ruído branco, também torna a matemática maleável. Mas além disso, um engenheiro conhecerá as estatísticas de primeira e segunda ordem (média e variância ou desvio padrão) do processo de ruído. Na ausência de qualquer estatística de mais alta ordem, não existe maneira melhor de simular que a densidade Gaussiana. As estatísticas de primeira e segunda ordem definem completamente uma densidade Gaussiana, ao contrário da maioria das densidades que requerem um número sem fim de ordens estatísticas para definir a sua forma completamente. Portanto, o filtro de Kalman, que propaga estatísticas de primeira e a segunda ordem inclui toda a informação contida na densidade de probabilidade condicional, ao invés de somente parte da informação, como aconteceria no caso de outra forma de densidade diferente (MAYBECK, 1979).

As hipóteses particulares que são feitas, são regidas pelos objetivos, e pelas motivações fundamentais, do modelo em desenvolvimento. Se o objetivo era meramente construir um bom modelo descritivo, a atenção não se restringiria somente a modelos de sistemas lineares regidos por ruído branco Gaussiano. Muito pelo contrário, se procuraria um modelo, de qualquer formato que fosse, que melhor se ajustasse aos dados gerados pelo “mundo real”. O desejo é construir estimadores e controladores baseados nesses modelos de sistema o que leva as seguintes suposições: outras suposições geralmente não produzem uma formulação de problemas de estimativa ou de controle maleáveis. Felizmente, a classe de modelos que fornecem uma matemática maleável, também fornece uma representação adequada à várias aplicações de interesse (MAYBECK, 1979).

4.2 UM EXEMPLO SIMPLES

Para demonstrar o funcionamento do filtro de Kalman, um simples exemplo será desenvolvido. Qualquer exemplo de um único dispositivo que forneça dados de uma única variável seria suficiente, mas a determinação de posição foi escolhida porque a probabilidade da exata localização de alguma coisa é um conceito familiar que facilmente permite que a dinâmica seja incorporada ao problema (MAYBECK, 1979).

Supondo que uma pessoa α está perdida no mar durante a noite e não faça a menor ideia de sua localização. Então α procura por uma estrela para estabelecer a sua posição (por questão de simplicidade, considera-se uma posição unidimensional). Em algum momento t_1 a localização será z_1 . Contudo, por causa de erros inerentes aos dispositivos de medida, erro humano, ou qualquer outro tipo de erro, o resultado da medida é incerto. Deste modo o desvio padrão envolvido é σ_{z_1} (ou de forma equivalente, a variância, ou estatística de segunda ordem, é $\sigma_{z_1}^2$). Portanto, pode-se estabelecer a probabilidade condicional de $x(t_1)$, a posição no tempo t_1 , condicionada ao valor de medida observado z_1 , como mostrado na figura 4.4. Este é um gráfico de $f_{x(t_1)|z(t_1)}(x|z_1)$ como função da localização x : isto informa a probabilidade de estar em qualquer lugar, baseado na medida obtida. Nota-se que σ_{z_1} é uma medida direta da incerteza: quanto maior σ_{z_1} , mais achatado o pico de probabilidade, espalhando o “peso” da probabilidade por uma faixa maior de valores de x . Para uma densidade Gaussiana, 68,3% do “peso” da probabilidade se encontra σ unidades dentro da banda de cada lado da média (MAYBECK, 1979).

Baseado nessa densidade de probabilidade condicional, a melhor estimativa da posição é :

$$\hat{x}(t_1) = z_1$$

E a variância do erro na estimativa é :

$$\sigma_x^2(t_1) = \sigma_{z_1}^2$$

Nota-se que \hat{x} é a moda (pico) e a mediana (valor com 1/2 do peso da probabilidade para cada lado), assim como a média (centro de massa) (MAYBECK, 1979).

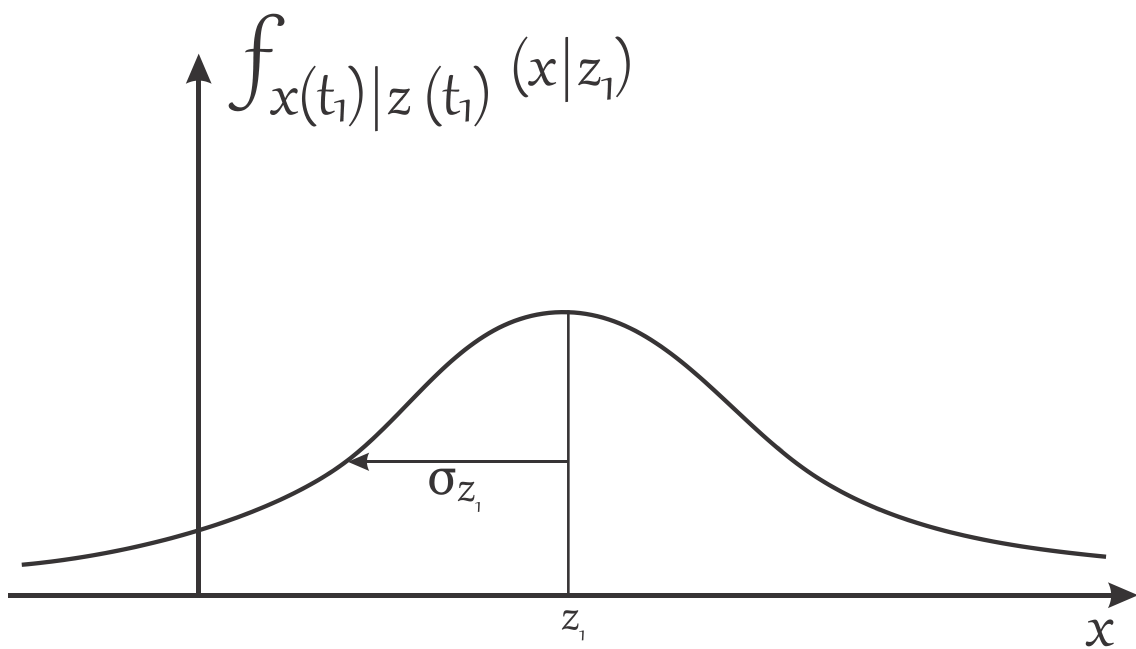


Figura 4.4

Agora uma segunda pessoa, β um navegador treinado faz uma correção independente, no momento $t_2 \cong t_1$ (de modo que a posição real não se altera), e obtém a medida z_2 com a variância σ_{z_2} . Como β é mais habilidoso, supõe-se que a variância da sua media é menor do que a de α . A figura 4.5 apresenta a densidade condicional da posição no momento t_2 , baseado somente no valor medido z_2 . Nota-se então um pico mais estreito, devido a variância menor, indicando um grau de certeza maior em relação à posição (MAYBECK, 1979).

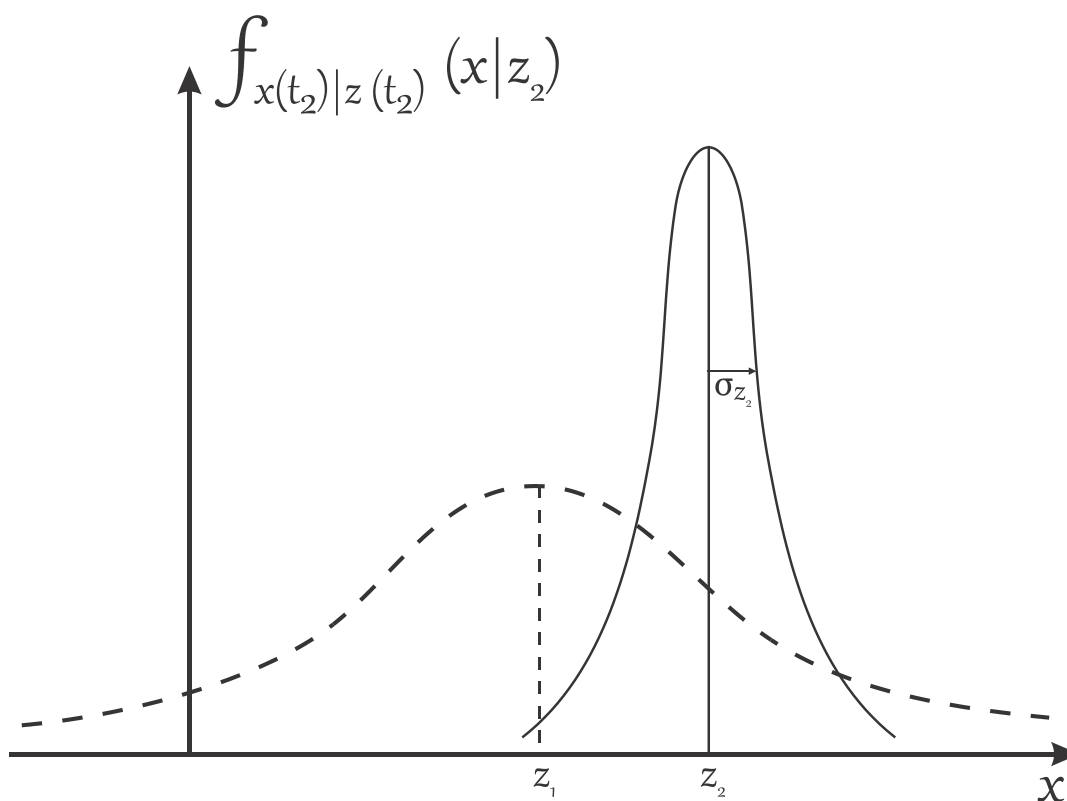


Figura 4.5

Neste ponto, existem duas medidas disponíveis para avaliar a posição. A pergunta é, como combinar esses dados? Baseado nas suposições feitas, a densidade condicional da posição no tempo $t_2 \cong t_1$, $x(t_2)$, dados tanto z_1 quanto z_2 , é uma densidade Gaussiana com média μ e variância σ^2 como indicado na figura 4.6, com:

$$\mu = \left[\frac{\sigma_{z_2}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_1 + \left[\frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_2$$

e

$$\frac{1}{\sigma^2} = \left(\frac{1}{\sigma_{z_1}^2} \right) + \left(\frac{1}{\sigma_{z_2}^2} \right)$$

Nota-se que σ é menor que σ_{z_1} ou σ_{z_2} , o que indica que a incerteza da posição diminuiu com a combinação das duas informações (MAYBECK, 1979).

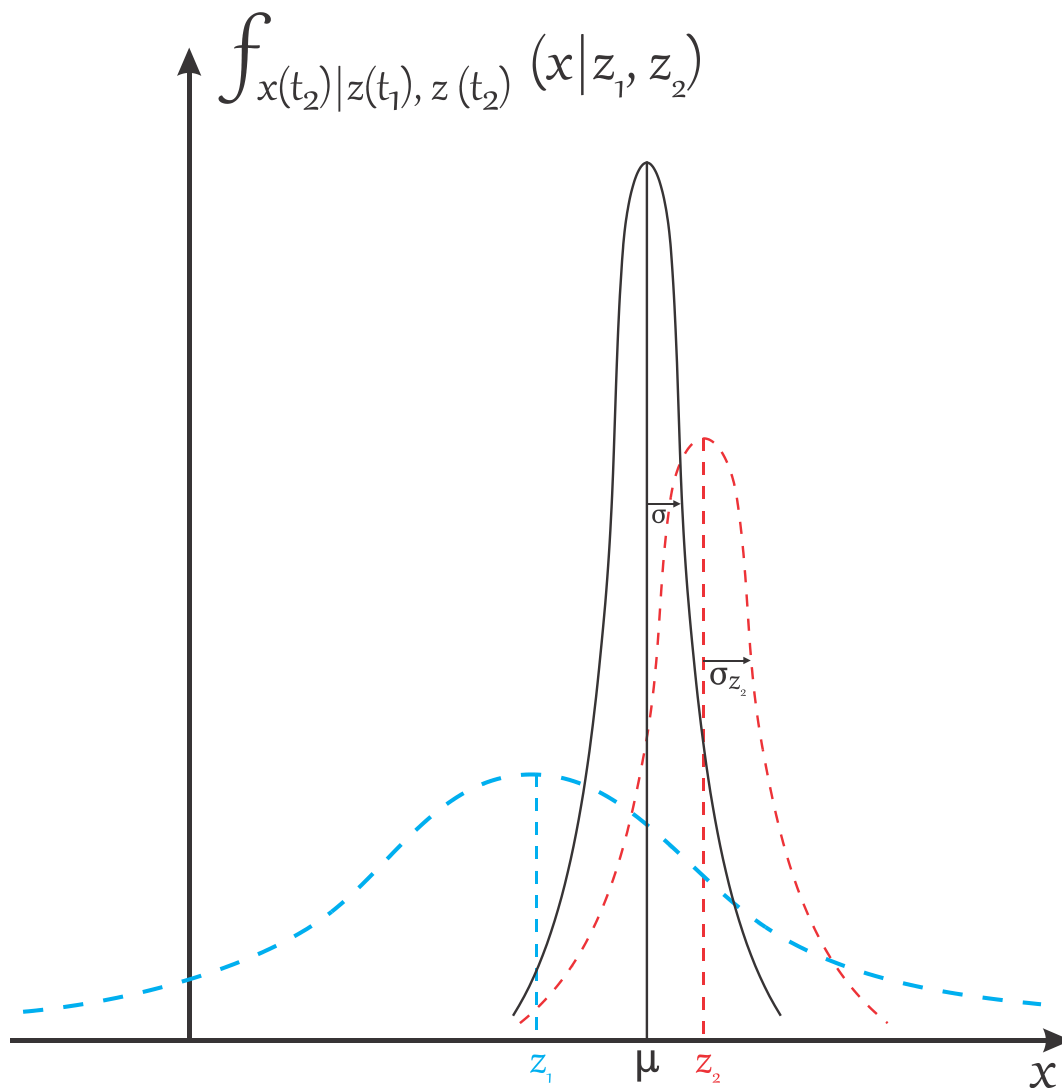


Figura 4.6

Dada esta densidade, a melhor estimativa é:

$$\hat{x}(t_2) = \mu$$

com uma variância de erro associada σ^2 . Isto representa a moda e a média (como é uma média de uma densidade condicional, também é chamada de média condicional). Ainda mais, também é a máxima verossimilhança estimada, a estimativa dos mínimos quadrados ponderada, e a estimativa linear cuja variância é menor do que qualquer outra estimativa linear desbalanceada. Em outras palavras, é o melhor que se pode conseguir dentro de um critério razoável (MAYBECK, 1979).

Após algum estudo, a forma de μ dada pela equação, faz bastante sentido. Se σ_{z_1} fosse igual a σ_{z_2} , diria-se que as medidas apresentavam a mesma precisão, e a equação diria que a estimativa ideal da posição seria simplesmente a média entre as duas medidas, com era de esperar. Se σ_{z_1} fosse maior que σ_{z_2} , significaria dizer que a incerteza envolvida na medida de z_1 é maior do que na medida de z_2 , e a equação exigiria que z_2 tivesse um peso maior que z_1 . Finalmente, a variância da estimativa é menor que σ_{z_1} , mesmo se σ_{z_2} for muito grande: mesmo dados de baixa qualidade são capazes de fornecer alguma informação, e devem portanto aumentar a precisão da saída do filtro (MAYBECK, 1979).

A equação para $\hat{x}(t_2)$ pode ser reescrita como:

$$\begin{aligned}\hat{x}(t_2) &= \left[\frac{\sigma_{z_2}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_1 + \left[\frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] z_2 \\ &= z_1 + \left[\frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \right] [z_2 - z_1]\end{aligned}$$

ou, na forma final, como é realmente usada na implementação do filtro de Kalman (observe que $\hat{x}(t_1) = z_1$)

$$\hat{x}(t_2) = \hat{x}(t_1) + K(t_2)[z_2 - \hat{x}(t_1)]$$

onde,

$$K = \frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2}$$

estas equações dizem que a estimativa ideal no tempo t_2 , $\hat{x}(t_2)$, é igual a melhor previsão de seus valores antes de z_2 ser tomado, $\hat{x}(t_1)$, mais um termo de correção

de ponderação ideal ajusta a diferença entre z_2 e a melhor predição de seu próprio valor, antes que seja realmente tomado, $\hat{x}(t_1)$. Vale a pena entender esta estrutura de “predição-correção” do filtro. Baseado em todas as informações prévias, haverá uma predição do valor das variáveis e medidas desejadas, no momento em que novas medidas forem adquiridas. Então quando as próximas medidas forem feitas, a diferença entre as novas medidas e os valores previstos será usada para “corrigir” a previsão das variáveis desejadas (MAYBECK, 1979).

Usando $K(t_2)$ na equação de variância, a equação será reescrita como:

$$\sigma_x^2(t_2) = \sigma_x^2(t_1) - K(t_2)\sigma_x^2(t_1).$$

Nota-se que os valores de $\hat{x}(t_2)$ e de $\sigma_x^2(t_2)$ incorporam toda a informação contida em $f_{x(t_2)|z(t_1),z(t_2)}(x|z_1,z_2)$. Expondo de outra forma, pela propagação destas duas variáveis, a densidade condicional da posição no instante t_2 , dados z_1 e z_2 , é completamente especificada (MAYBECK, 1979).

Portanto, foi solucionado o problema de estimativa estática. Agora, será incorporada dinâmica ao sistema.

Supondo que α tenha viajado por algum tempo antes de fazer nova medida. Assumindo que o melhor modelo que se tem para o movimento é a forma simples;

$$dx/dt = u + w$$

onde u é a velocidade nominal e w é um termo de ruído, usado para representar a incerteza no conhecimento da velocidade real devido a perturbações, efeitos não levados em consideração pela equação simples de primeira ordem, e assim por

diante. O “ruído” w será modelado como um ruído branco Gaussiano com média zero e variância σ_w^2 (MAYBECK, 1979).

A figura 4.7 apresenta graficamente o que acontece à densidade de posição condicional, dados z_1 e z_2 . Com o passar do tempo, a densidade viaja através do eixo x na velocidade nominal u , enquanto que simultaneamente espalha a sua média. Desta forma, a densidade da probabilidade começa na sua melhor estimativa, se move de acordo com o modelo nominal de dinâmica, e se espalha no tempo porque a certeza do correto posicionamento vai diminuindo devido a constante adição de incerteza com o passar do tempo. No tempo t_3^- , um instante antes que a medida seja tomada no tempo t_3 , a densidade $f_{x(t_3^-)|z(t_1), z(t_2)}(x|z_1, z_2)$, se apresenta como na figura 4.7, e pode ser expressa matematicamente como uma densidade Gaussiana com média e variância dadas por:

$$\hat{x}(t_3^-) = \hat{x}(t_2) + u[t_3 - t_2]$$

$$\sigma_x^2(t_3^-) = \sigma_x^2(t_2) + \sigma_w^2[t_3 - t_2]$$

portanto, $\hat{x}(t_3^-)$ é a predição ideal de qual é o valor de x em t_3^- , antes que a medida seja feita em t_3 , e $\sigma_x^2(t_3^-)$ é a variância esperada nesta predição (MAYBECK, 1979).

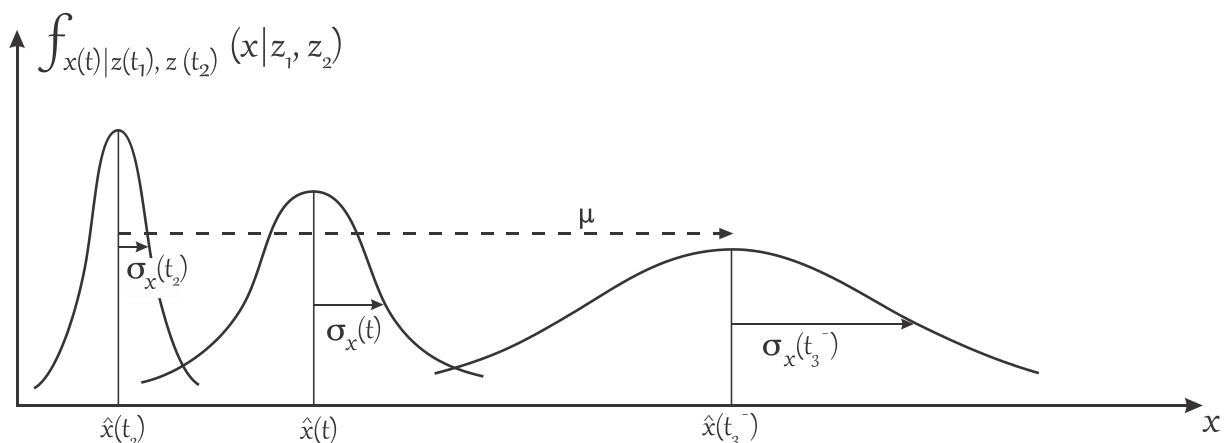


FIGURA 4.7

Neste momento, uma nova medida foi feita, e seu valor é z_3 , e sua variância passa a ser $\sigma_{z_3}^2$. Como anteriormente, existem agora duas densidades Gaussianas disponíveis que contém informação sobre a posição, uma contendo toda a informação disponível antes da medida e a outra com a informação da própria medida. Pelo mesmo processo descrito anteriormente, a densidade com média $\hat{x}(t_3^-)$ e a variância $\sigma_x^2(t_3^-)$ é combinada com a densidade de média z_3 e variância $\sigma_{z_3}^2$, e resulta na densidade Gaussiana com média:

$$\hat{x}(t_3) = \hat{x}(t_3^-) + K(t_3)[z_3 - \hat{x}(t_3^-)]$$

e variância:

$$\sigma_x^2(t_3) = \sigma_x^2(t_3^-) - K(t_3)\sigma_x^2(t_3^-)$$

onde o ganho $K(t_3)$ é dado por:

$$K(t_3) = \sigma_x^2(t_3^-) / [\sigma_x^2(t_3^-) + \sigma_{z_3}^2].$$

A estimativa ideal, $\hat{x}(t_3)$, cumpre a mesma forma da equação vista anteriormente. A melhor predição do valor antes de z_3 ser medido, é corrigida pelo valor ponderado ótimo e ajusta a diferença entre z_3 e a predição de seu valor (MAYBECK, 1979).

Observe a forma da equação de $K(t_3)$. Se $\overline{\sigma_{z_3}}$, a medida da variância do ruído, é grande, então $K(t_3)$ é pequeno; isto simplesmente diz que existe uma tendência de se confiar pouco em uma medição muito ruidosa e o peso seria leve. No limite com $\sigma_{z_3}^2 \rightarrow \infty$, $K(t_3)$ se torna zero, e $\hat{x}(t_3)$ é igual $\hat{x}(t_3^-)$; uma medida infinitamente ruidosa é completamente ignorada. Se a variância do ruído do sistema dinâmico σ_w^2 é grande, então $\sigma_x^2(t_3^-)$ será grande assim como $K(t_3)$; nesse caso, não

se tem certeza da saída do modelo do sistema dentro da estrutura do filtro e desta forma irá aplicar um peso alto à medida. Nota-se que quando $\sigma_w^2 \rightarrow \infty$, $\sigma_x^2(t_3^-) \rightarrow \infty$ e $K(t_3) \rightarrow 1$, teremos o seguinte:

$$\hat{x}(t_3) = \hat{x}(t_3^-) + 1[z_3 - \hat{x}(t_3^-)] = z_3.$$

Deste modo, no limite, quando não existe nenhuma confiança na saída no modelo do sistema, a política ideal é ignorar a saída e usar a nova medida como a estimativa ideal. Finalmente, se $\sigma_x^2(t_3^-)$ em algum momento se tornar zero, o mesmo acontecerá a $K(t_3)$, isto faz sentido como $\sigma_x^2(t_3^-) = 0$, se está absolutamente certo sobre a estimativa e desta forma pode descartar a medida de z_3 (MAYBECK, 1979).

Embora não se tenha chegado a este resultado por intermédio da matemática, foi demonstrado a razoabilidade da estrutura do filtro (MAYBECK, 1979).

4.3 O FILTRO DISCRETO DE KALMAN

Em 1960, Rudolf Emil Kalman publicou seu famoso artigo que descrevia uma solução recursiva para o problema de filtragem linear de dados discretos. Desde aquela época, devido principalmente aos grandes avanços na computação digital, o filtro de Kalman em sido objeto de ampla pesquisa e aplicação, principalmente nas áreas de sistema autônomos e navegação assistida (WELCH, 2006).

O filtro de Kalman diz respeito ao problema geral de se tentar estimar o estado $x \in \mathfrak{R}^n$ de um processo controlado de tempo discreto que é governado pela equação linear de diferenças estocástica

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$$

com a equação de medida

$$z_k = Hx_k + v_k.$$

As variáveis aleatórias w_k e v_k representam os ruídos do processo e da medidas respectivamente. São arbitrados como independentes, brancos e de distribuição de probabilidades normal

$$p(w) \sim N(0, Q),$$

$$p(v) \sim N(0, R).$$

Na prática, as matrizes de covariância do ruído do processo Q e covariância do ruído de medidas R podem variar a cada passo de tempo ou de medida, mas nesse caso deve ser consideradas constantes (WELCH, 2006).

A matriz A de ordem $n \times m$ na equação de diferenças relaciona o estado de tempo anterior $k - 1$ com o estado atual k . Nota-se que na prática A pode variar a cada período de tempo, mas nesse estudo vai se considerar que é constante. A matriz B de ordem $n \times l$ relaciona a entrada de controle ótima $u \in \mathfrak{R}$ com o estado x . A matriz H de ordem $m \times n$ na equação de medida relaciona o estado a medida z_k . Na prática, H pode variar a cada período de tempo ou cada medida, mas nesse caso considera-se como constante (WELCH, 2006).

Define-se $\hat{x}_k^- \in \mathfrak{R}^n$, como sendo a estimativa de estado *a priori* de um passo k , sabend- se o processo anterior de k , e $\hat{x}_k \in \mathfrak{R}^n$ como sendo a estimativa de estado *a posteriori* de um passo k dada uma medida z_k . Define-se então os erro *a priori* e *a posteriori* como sendo:

$$e_k^- \equiv x_k - \hat{x}_k^-$$

e

$$e_k \equiv x_k - \hat{x}_k.$$

A covariância do erro estimado *a priori* é então:

$$P_k^- = E[e_k^- e_k^{-T}]$$

e a covariância do erro estimado *a posteriori* é:

$$P_k = E[e_k e_k^T].$$

A fim de obter as equações para o filtro de Kalman, inicia-se procurando uma equação que calcule o estado estimado *a posteriori* \hat{x}_k como uma combinação linear da estimativa *a priori* \hat{x}_k^- e uma diferença ponderada entre uma medida real z_k e uma predição da medida $H\hat{x}_k^-$, como pode ser visto abaixo;

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-).$$

A diferença $(z_k - H\hat{x}_k^-)$ é chamada de inovação da medida ou resíduo. O resíduo reflete a discrepância entre a medida prevista $H\hat{x}_k^-$ e a medida real z_k . Um resíduo com valor zero significa que os dois valores estão em completo acordo (WELCH, 2006).

A matriz K de ordem $n \times m$ é escolhida de modo a ser o ganho, ou fator de mistura que minimiza a covariância do erro *a posteriori*. Esta minimização pode ser conseguida substituindo a equação anterior na definição de e_k , substituindo o resultado na equação da covariância do erro estimado, realizando-se as

operações necessárias, derivando-se o resultado em função de K , igualando-se esse resultado a zero (ponto crítico da função), e resolvendo para K . Uma forma obtida do resultado de K , que minimiza a covariância do erro estimado é o seguinte:

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} = \frac{P_k^- H^T}{(H P_k^- H^T + R)}$$

observando-se o resultado final, percebe-se que quanto mais a covariância do erro medida R se aproxima de zero, mais o ganho K influencia o resíduo.

Especificamente,

$$\lim_{R_k \rightarrow 0} K_k = H^{-1}$$

por outro lado, quando a covariância do erro estimado *a priori* P_k^- se aproxima de zero, o ganho K pesa menos no resíduo. Especificamente,

$$\lim_{P_k^- \rightarrow 0} K_k = 0$$

uma outra forma de se visualizar a importância de K , é que quando a covariância de erro medida R se aproxima de zero, a medida real z_k se torna cada vez mais confiável, enquanto que a medida prevista $H\hat{x}_k^-$ se torna cada vez menos confiável. Por outro lado, quando o erro estimado *a priori* P_k^- se aproxima de zero, a medida real z_k se torna menos confiável, enquanto que a medida prevista $H\hat{x}_k^-$ se torna mais e mais confiável (WELCH, 2006).

4.4 O ALGORITMO DO FILTRO DISCRETO DE KALMAN

O filtro de Kalman estima um processo usando um tipo de controle de realimentação: o filtro estima o estado do processo em algum período de tempo e

então é realimentado na forma de medidas (ruidosas). Desta maneira as equações do filtro de Kalman se enquadram em dois grupos; equações de tempo atualizado e equações de medidas atualizadas. As equações de tempo atualizado são responsáveis por projetar adiante (no tempo) o estado corrente e o erro de covariância estimado para obter estimativas *a priori* para o próximo passo de tempo. As equações de medidas atualizadas são responsáveis pela realimentação, quer dizer, por incorporar uma nova medida a estimativa *a priori* de modo a obter uma estimativa *a posteriori* aprimorada (WELCH, 2006).

As equações de tempo atualizado podem ser vistas como as equações de predição, enquanto que as equações de medidas atualizadas, como equações de correção. Na realidade, o algoritmo final de estimação se assemelha ao algoritmo predictor-corrector utilizado na resolução de problemas numéricos, como mostrado na figura abaixo.

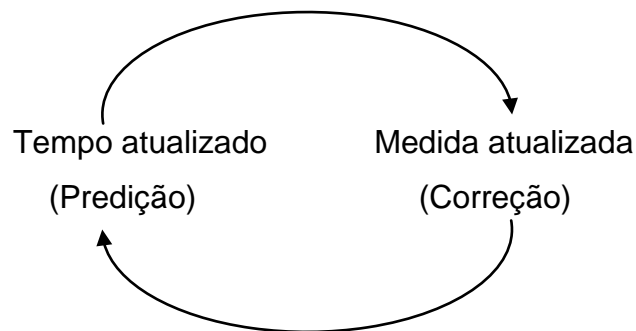


Figura 4.8 – O ciclo contínuo do filtro de Kalman

As equações de tempo atualizado são as seguintes:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \quad (4.1)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (4.2)$$

As equações de medidas atualizadas são as seguintes:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (4.3)$$

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-) \quad (4.4)$$

$$P_k = (1 - K_k H)P_k^- \quad (4.5).$$

A primeira tarefa durante a atualização da medida é computar o ganho de Kalman K . O próximo passo é medir o processo realmente para obter z_k , e então gerar uma estimativa de estado *a posteriori* pela incorporação da medida a equação 4.4. o passo final é obter a covariância de erro estimada *a posteriori* por meio da equação 4.5 (WELCH, 2006).

Após cada par de atualizações de tempo e medida, o processo é repetido com as estimativas *a posteriori* prévias usadas para projetar ou prever as novas estimativas *a priori*. Esta característica recursiva é um dos aspectos mais atrativos do filtro de Kalman, isto torna as aplicações práticas do filtro de Kalman muito mais fáceis de se implementar do que por exemplo um filtro de Wiener, que é projetado para operar em todos os dados diretamente a cada estimativa. O filtro de Kalman, ao contrário, determina recursivamente a estimativa corrente em todas as medidas passadas. A figura 4.9 oferece um quadro completo da operação do filtro, combinado o diagrama da figura 4.8, com as equações de tempo e medidas atualizadas.

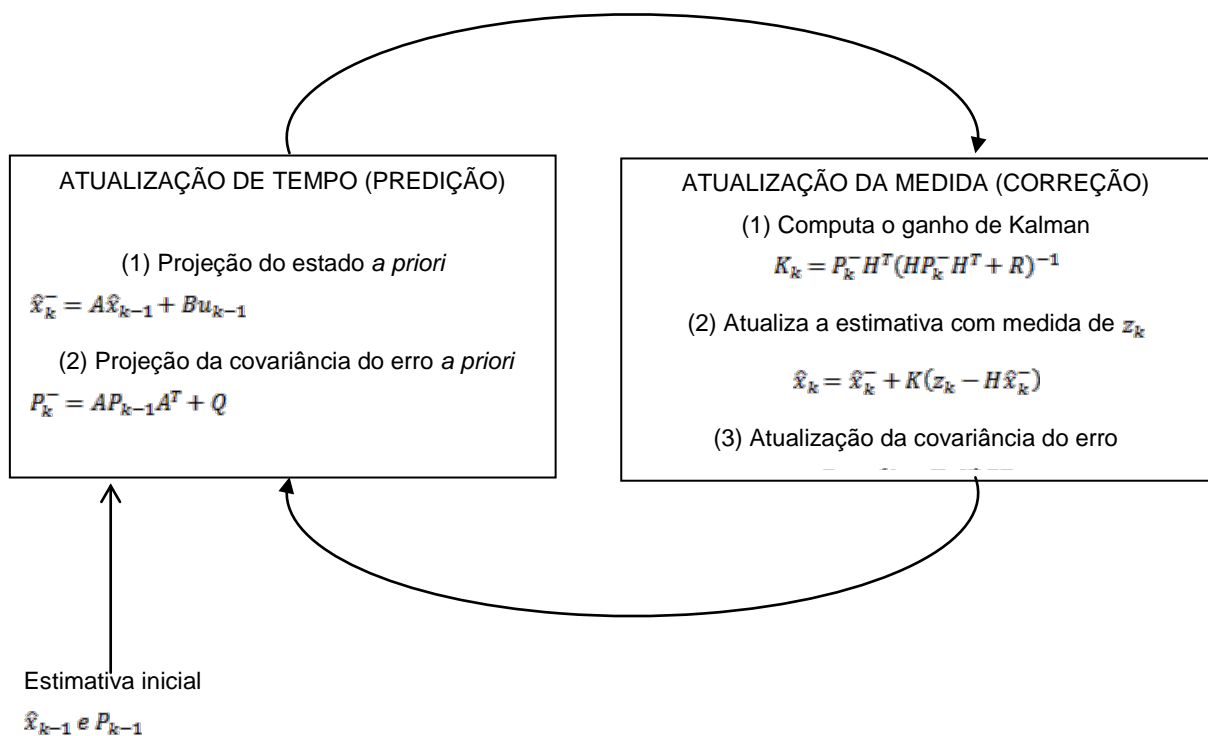


Figura 4.9 – Algoritmo do filtro de Kalman

Quando da implementação real do filtro, a covariância do ruído medida R é usualmente medida antes da operação do filtro. Avaliar a covariância do erro medida R geralmente é possível porque o processo deve ser medido de alguma forma (durante a operação do filtro), então, geralmente é possível conseguir algumas amostras de medida *off-line* para determinar a variância do ruído medido (WELCH, 2006).

A determinação da covariância do ruído do processo Q é geralmente mais difícil, pois não existe a habilidade de se observar diretamente o processo que está sendo estimado. Algumas vezes um relativamente simples modelo do processo é capaz de produzir resultados aceitáveis se for inserido incertezas suficientes no processo pela seleção de Q . Certamente neste caso, espera-se que as medidas do processo sejam confiáveis (WELCH, 2006).

Em ambos os casos, se existe, ou não, uma base racional para a escolha de parâmetros, na maioria das vezes é possível se obter um desempenho superior do filtro (estatisticamente falando), ajustando-se os parâmetros Q e R do filtro. O ajuste é geralmente realizado *off-line*, frequentemente com o auxílio de outro filtro de Kalman (distinto) em um processo chamado de identificação do sistema (WELCH, 2006).

Encerrando, nota-se que sob as condições em que Q e R são de fato constantes, tanto a estimativa da covariância do erro P_k , quanto o ganho de Kalman K_k irão se estabilizar rapidamente se manter constantes. Se esse for o caso, estes parâmetros podem ser pré computados seja rodando o filtro *off-line*, seja pela determinação do valor estável de P_k (WELCH, 2006).

5 RASTREAMENTO DE ALVOS

Desde a invenção do radar, operadores tem detectado e rastreado alvos com o auxílio de uma gama de mostradores. Embora os operadores sejam capazes de realizar esta tarefa de forma bastante precisa, os operadores logo se saturam e se cansam facilmente. Vários estudos tem mostrado que um operador consegue rastrear manualmente uma pequena quantidade de alvos. De modo a corrigir esta situação, sistemas automáticos de detecção e rastreamento tem sido agregados a muitos radares. Conforme o processamento digital aumenta de velocidade e o hardware diminui em custo e tamanho, os sistemas de detecção e rastreamento automático tem sido associados com praticamente todos os tipos de radares (SKOLNIK, 1990).

5.1 O QUE É RASTREAMENTO

O rastreamento de alvos se resume ao problema de usar sensores de medida de modo a determinar a localização, rota e características de objetos de interesse. Um sensor pode se referir a qualquer dispositivo que realize uma medida, radar, sonar, câmera, sensor infravermelho, microfone, ou qualquer outro sensor que possa ser usado para coletar informações sobre objetos em um ambiente qualquer (CHALLA, 2011).

Rastreamento é o processo de determinação da velocidade e direção de alvos, que possibilita o monitoramento do alvo através da área de cobertura. Tecnicamente, rastreamento é uma estimativa de parâmetros. Contudo, a observação de rastros, estágio por estágio em tempo real, estabelecidos por retornos de um alvo, na forma de picos de detecção, separa 'técnica de estimativa de parâmetros' do 'rastreamento'. A fim de esclarecer, é necessário fazer a distinção entre um alvo e um rastro. Um alvo é um objeto físico, capaz de excitar um sensor e produzir medidas, enquanto que um rastro é uma representação simbólica do alvo, formado por sucessivas posições detectadas. De forma geral, determinar um rastro significa o desenvolvimento de um modelo matemático que representa a estrutura do

alvo, os valores dos parâmetros, e se necessário, os valores das variáveis dependentes como por exemplo, as variáveis de estado (KOLAWOLE, 2002).

5.2 INTRODUÇÃO AO RASTREAMENTO DE ALVOS

O objetivo básico do rastreamento de alvos é a determinação de um certo número de objetos, sua identidades, e seus estados, como por exemplo, posição velocidade, e em alguns casos suas características físicas. Um exemplo típico de rastreamento de alvos é o rastreamento radar de uma aeronave. O rastreamento de alvos neste contexto se resume a tentativa de determinação do número de aeronaves em uma determinada região sobre vigilância, seus tipos, militar, civil ou de recreação, suas identificações e suas velocidade e posição, tudo baseado nas medidas do radar (CHALLA, 2011).

Supondo que um radar de vigilância, que possui uma antena rotativa, com período de rotação de 10 segundos (cerca de 6 rotações por minuto), fornece uma informação bidimensional sobre um determinado alvo. A primeira dimensão é a distância do alvo (o tempo que o pulso transmitido leva para ir da antena até o alvo e voltar); a segunda dimensão é azimute (ângulo formado pela direção na qual a antena aponta em relação ao norte magnético) (BROOKNER, 1998).

Supondo que em um tempo $t = t_1$ o radar está apontando para o ângulo θ e dois alvos são detectados nas distâncias R_1 e R_2 . Supondo que na próxima varredura em um tempo $t = t_1 + T$, novamente dois alvos são detectados. A questão que surge é a seguinte: os alvos detectados na segunda varredura são os mesmos dois alvos ou são alvos novos? A resposta a esta pergunta é de extrema importância para radares de controle de tráfego aéreo e radares militares. No caso dos radares de tráfego aéreo, o correto conhecimento do número de alvos a serem controlados é imprescindível a fim de se evitar colisões. No caso dos radares militares é importante para uma avaliação apropriada em uma situação de ameaça e no caso de uma possível interceptação (BROOKNER, 1998).

Supondo que os dois alvos detectados na segunda varredura foram identificados de forma correta como os mesmos dois alvos detectados na primeira varredura. A nova questão que surge é como associar corretamente o eco do alvo 1 da segunda varredura com o eco do alvo 1 da primeira varredura, o mesmo acontecendo para o alvo 2 (BROOKNER, 1998).

Se uma for feita associação incorreta, então uma velocidade incorreta será vinculada ao alvo. Por exemplo, se o eco do alvo 1 na segunda varredura for associado com o eco do alvo 2 da primeira varredura, então o alvo 2 aparentará ter uma velocidade bem maior do que a real. No controle de tráfego aéreo, este tipo de erro muito provavelmente acarretará numa colisão; em um radar militar, ocasionará um erro no caso de uma interceptação com possível perda do alvo a ser interceptado (BROOKNER, 1998).

As chances de uma associação incorreta podem ser reduzidas drasticamente, se fosse possível prever, com precisão, onde os alvos 1 e 2 estariam na segunda varredura. Tal previsão é facilmente obtida com o conhecimento da velocidade estimada e da posição dos alvos em questão na primeira varredura. Desta forma, estimaria-se o quanto os alvos se movimentariam a cada varredura e como resultado obteria-se uma estimativa das futuras posições dos alvos. Como a exata velocidade e posição não são conhecidas na primeira varredura, esta estimativa não é exata. Se a incorreção for conhecida, pode-se estipular uma janela de $\pm 3\sigma$ (ou $\pm 2\sigma$) sobre o valor esperado, onde σ é o valor médio quadrático, ou de forma equivalente, o desvio padrão da soma da previsão mais o valor médio quadrático da medida da distância (BROOKNER, 1998,).

Existe um grande número de incertezas no problema de rastreamento de alvos, que o torna uma tarefa não trivial. Por exemplo, o movimento de um objeto está frequentemente sujeito a uma gama de perturbações aleatórias, alguns objetos podem passar despercebidos pelo sensor e o número de objetos no campo de visão do sensor pode variar de forma incerta. As medidas do sensor estão sujeitas a ruídos aleatórios e a quantidade de medidas recebidas pelo sensor de uma varredura a outra, pode se alterar e se tornar imprevisíveis. Os objetos podem estar

muito próximos uns dos outros e a medida recebida pode não tê-los distinguido separadamente. Algumas vezes os sensores podem enviar dados quando nenhum objeto se encontra no campo de visão do sensor (CHALLA, 2011).

5.3 RASTREAMENTO DE ALVOS EM UM SISTEMA RADAR

Um importante problema de rastreamento é o rastreamento de alvos em um sistema radar, como por exemplo, no controle de tráfego aéreo. O rastreamento radar também é usado em sistemas de segurança militar, onde o problema envolve a identificação da aeronave, sua velocidade, seu tipo, sua localização e suas intenções, de forma a se determinar se é, ou não, uma ameaça. O radar usa ondas de rádio refletidas para medir a direção, a distância e a velocidade radial de um objeto detectado.. Um transmissor radar emite ondas eletromagnéticas, que são refletidas por um objeto e detectadas por um receptor. Os dados medidos são usados para se extrair um rastro, que é apresentado junto com a reflexão do objeto em uma tela (CHALLA, 2011).

O rastreamento de uma aeronave utilizando-se os dados do radar, é particularmente difícil por causa das incertezas na origem das medidas. Pássaros, animais, vegetação, terreno, nuvens, mar, chuva, neve e sinais gerados por outros radares, geram um ruído no sinal radar conhecido como *clutter*. O *clutter* consiste em qualquer detecção que não seja o objeto de interesse (CHALLA, 2011).

O rastreamento de um alvo em um radar de vigilância pode ser determinado pelas coordenadas medidas do alvo varredura por varredura. A qualidade deste rastro irá depender do tempo entre as observações, da precisão da localização a cada observação e da quantidade de *clutter* que pode estar presente nas proximidades do alvo sendo rastreado (SKOLNIK, 1990).

Um método de se obter rastreamento em um radar de vigilância é pedir que o operador marque manualmente com uma caneta na tela de apresentação, a localização do alvo a cada varredura. A simplicidade deste processo contrasta com a

pouca precisão do método. A precisão pode ser melhorada com a utilização de um computador sendo alimentado com os dados fornecidos pelo operador. Contudo, um operador humano consegue atualizar os dados a uma taxa máxima de apenas um alvo a cada dois segundos. E ainda mais, um único operador não é capaz de lidar com mais de seis rastros de alvos em um radar de baixa velocidade de varredura (cerca de 5 rotações de antena por minuto). E para finalizar, a capacidade de detecção de novos alvos por esse mesmo operador diminui rapidamente com o passar do tempo. Todas essas limitações humanas podem ser eliminadas com a utilização de um sistema de detecção e rastreamento automático, ADT (abreviatura da expressão em inglês Automatic Detection and Tracking). Um sistema automático do tipo ADT, é capaz de realizar as funções de detecção do alvo, inicialização do rastreamento, associação de alvos, atualização dos rastros, suavização dos rastros e término do rastreamento (SKOLNIK, 1981).

Um detector automático, parte do sistema ADT, realiza a quantização da distância em intervalos iguais a resolução de distância do radar (resolução de distância é a distância mínima no qual dois alvos próximos podem ser considerados alvos independentes). A cada intervalo de distância o detector integra n pulsos, onde n é o número de pulsos de retorno esperados de um determinado alvo a cada varredura de antena. Os pulsos integrados são comparados com um nível a fim de indicar a presença ou ausência de alvo. Um exemplo usado de forma comum é a janela móvel de detecção, que examina as últimas n amostras dentro do intervalo de distância quantizado e indica a presença de um alvo se m de cada n amostras atingirem o nível adequado. Localizando-se o centro dos n pulsos, uma estimativa da direção angular do alvo pode ser obtida. Esta técnica é chamada divisão do feixe (SKOLNIK, 1981).

Se houver pelo menos um alvo presente dentro da cobertura radar, a detecção em duas varreduras é tudo que se necessita para estabelecer um rastreamento do alvo e estimar a sua velocidade. Porém, existem muito mais alvos, e a presença de ecos de *clutter*, de modo que pelo menos três ou mais detecções são necessárias de modo a se obter um rastreamento confiável, sem a geração de rastros falsos ou espúrios. Embora possa se programar o sistema de modo a rejeitar alvos falsos, uma quantidade muito alta de alvos falsos pode sobrecarregar o sistema, resultando

em informações não confiáveis. Um mapa de *clutter* gerado pelo radar é usado para reduzir a carga do sistema de rastreamento, por meio do apagamento de áreas de *clutter* e da remoção de detecções associadas com grandes áreas de *clutter*. Alvos lentos, com velocidade muito abaixo da esperado por uma aeronave, também podem ser removidos pelo mapa de *clutter*, embora o correto processamento desta informação é capaz de separar a informação de nuvens que podem ser úteis no monitoramento das condições meteorológicas. A disponibilidade de algumas características do alvo, como a altitude por exemplo, ajudam na associação correta do alvo a ser rastreado. A qualidade do sistema ADT depende muito da habilidade de rejeição de sinais indesejados do radar (SKOLNIK, 1981).

Quando uma nova detecção é recebida, a primeira tentativa é associar com um alvo já existente. Isto é realizado por meio do estabelecimento de uma pequena região de procura, ou janela, dentro da qual uma nova detecção é prevista baseado na velocidade e direção do alvo. O tamanho da janela deve ser a menor possível para evitar que mais de um alvo se encontre dentro da janela quando a densidade de tráfego for muito alta, mas também deve ser a maior possível para capturar todas as manobras repentinas do alvo. Desta forma, para se superar este dilema, mais de um tamanho de janela deve ser utilizado. O tamanho da janela menor será determinado pela precisão do rastro. Quando um alvo não aparecer na janela menor, uma janela maior deve ser utilizada, cuja área de atuação deve ser determinada pela aceleração máxima estimada do alvo (SKOLNIK, 1981).

Baseado nas detecções passadas, o radar deve realizar uma estimativa regular da posição atual e da velocidade de um alvo, assim como uma previsão da posição e da velocidade. Um método para se computar esta informação é usar um filtro $\alpha - \beta$ (SKOLNIK, 1981).

5.3.1 O filtro $\alpha - \beta$

Por questão de simplicidade, vamos supor apenas uma dimensão de coordenada. Supõe-se então que um alvo se move radialmente em direção a um

radar, com x_n representando a distância do alvo no tempo n . Supõe-se, ainda por questão de simplificar, que a velocidade do alvo é constante; deste modo a previsão da posição do alvo e sua velocidade na próxima varredura pode ser feita usando-se as simples equações de movimento do alvo:

$$x_{n+1} = x_n + T\dot{x}_n \quad (5.1)$$

$$\dot{x}_{n+1} = \dot{x}_n \quad (5.2)$$

Onde x_n é a posição do alvo na varredura n , \dot{x}_n é a velocidade do alvo na varredura n , e T é o período entre as varreduras. Estas equações de movimento são chamadas de modelo dinâmico do sistema. Estas equações permitem uma previsão do tempo n para o tempo $n + 1$. É necessário mostrar como melhorar a estimativa da posição do alvo e sua velocidade depois que uma observação do alvo no tempo n e em tempos sucessivos, for feita. Supõe-se que no momento tem-se uma estimativa da posição do alvo e sua velocidade no tempo $n - 1$ (BROOKNER, 1998).

Supondo que a velocidade estimada do alvo no tempo $n - 1$ é de 360 Km/h e que o período entre as varreduras é de 10 segundos. Utilizando-se a equação 5.1, estima-se que o alvo estará 1 Km mais longe no tempo n , se arbitrado que o mesmo está se afastando do radar. Supondo que no tempo n , o radar enxerga o alvo na posição y_n , a uma distância 50 m além do esperado. Onde o alvo está na realidade? Em x_n , em y_n , ou em algum lugar entre essas medidas? Novamente, supondo que a medida em $n - 1$ foi muito precisa e que sua precisão é de 1 m. Neste caso, a visualização do alvo a uma distância 50 m ou mais que a distância prevista indica que o alvo está se movendo mais rápido do que o foi originariamente estimado no tempo $n - 1$, viajando 50 m mais em 10 segundos, o alvo está $50 \text{ m}/10 \text{ s} = 5 \text{ m/s}$ mais rápido do que se estimava. Então a velocidade corrigida do alvo é:

$$\text{Velocidade corrigida} = 100 \text{ m/s} + \frac{50 \text{ m}}{10 \text{ s}} = 105 \text{ m/s} = 378 \text{ Km/h}$$

Isto é verdade em um sistema muito preciso, mas novamente supondo-se que a precisão é de 40 m. O desvio de 50 m da localização estimada do alvo poderia ser somente em função de um erro da medida do radar e não devido a uma velocidade maior do alvo. Porém, o alvo poderia realmente estar mais veloz do que o esperado, e esta hipótese não poderia ser descartada. Este problema pode ser resolvido se for debitado apenas uma fração do aumento de velocidade aparente. Nesse caso usa-se um décimo do acréscimo de velocidade aparente, isto leva a :

$$\text{Velocidade corrigida} = 100 \text{ m/s} + 1/10(50 \text{ m}/10 \text{ s}) = 100,5 \text{ m/s} = 361,8 \text{ Km/h}$$

Desta forma a velocidade não é corrigida no seu valor total. Se o alvo estiver indo realmente mais rápido, então em observações sucessivas, a visualização do alvo, na média, tenderá a ser mais corrigida na distância do que as posições prevista para o alvo. Se em varreduras sucessivas a velocidade for aumentada de $0,5 \text{ m/s}$ na média, em 10 varreduras a velocidade tenderá ao valor exato. Por outro lado, se a velocidade do alvo for realmente 360 Km/h, a posição medida do alvo estaria com uma probabilidade igual de estar na frente ou atrás da posição prevista, logo na média, a velocidade não se alteraria e manteria o valor estimado e equação terá o seguinte formato:

$$\dot{x}_{n,n}^* = \dot{x}_{n,n-1}^* + \beta_n \left(\frac{y_n - \dot{x}_{n,n-1}^*}{T} \right)$$

A fração $1/10$ é representada pelo parâmetro β_n . O subscrito n indica que o parâmetro β dependerá do tempo. O segundo subscrito, $n - 1$ na velocidade estimada $\dot{x}_{n,n-1}^*$ aponta uma estimativa da velocidade do alvo no tempo n baseada em uma medida no tempo $n - 1$ e tempos anteriores. O segundo subscrito n na

velocidade estimada $\dot{x}_{n,n}^*$, indica que esta velocidade estimada usa a medida de distância do tempo n , ou seja, y_n . O asterisco sobrescrito significa que os parâmetros são estimativas. Sem o asterisco, os parâmetros representam os valores reais da posição e da velocidade (BROOKNER, 1998).

Agora se faz necessário uma equação para atualização da posição do alvo. Como anteriormente, supondo-se que no tempo $n - 1$ o alvo se encontra a 10 milhas náuticas (unidade de medida utilizada para indicar distância em sistemas de radar, onde uma milha náutica equivale a 1852 m) e no tempo n , $T = 10$ s, o alvo tem uma velocidade radial de 360 Km/h e se encontra a uma distância de 1 Km mais longe que no tempo $n - 1$. Supondo-se que no tempo n o alvo esteja sendo realmente visualizado 50 m a frente do previsto. Novamente questiona-se onde é a posição real; em $\dot{x}_{n,n-1}^*$, em y_n , ou em algum lugar entre esse valores. Novamente supõe-se que um radar muito preciso está sendo usado para as medidas em $n - 1$ e n . Pode se então concluir que o alvo se encontra na posição em que está sendo visualizado no tempo n , ou seja, a 50 m a frente do previsto. Logo,

$$\text{Posição atualizada} = 10 \text{ NM} + 1 \text{ Km} + 50 \text{ m}$$

Se, contudo, assume-se que o radar tem um desvio padrão de 40 m, então o alvo que aparece a 50 m além do esperado, pode ter esta medida apenas por um erro de medição do radar. Desta maneira não podemos garantir com segurança que o alvo se encontra na posição medida y_n no tempo n . Por outro lado, garantir que o alvo se encontra na posição prevista é igualmente inseguro. Pode-se assumir que provavelmente o alvo se encontra em uma posição um pouco a frente da posição prevista por uma fração do valor de 50 m. Então supõe-se que o alvo se encontra a um quinto dos 50 m, ou seja, 10 m além. Logo a posição em distância a tualizada é:

$$\text{Posição atualizada} = 10 \text{ NM} + 1 \text{ Km} + \frac{1}{5}(50 \text{ m})$$

Colocando-se em uma forma paramétrica temos:

$$\dot{x}_{n,n}^* = \dot{x}_{n,n-1}^* + \alpha_n (y_n + \dot{x}_{n,n-1}^*)$$

onde a fração $1/5$ é representada pelo parâmetro α_n , que é dependente do tempo n .

Esta equação representa a forma desejada para atualizar a posição do alvo (BROOKNER, 1998)

As equações de atualização de velocidade e atualização da posição em um tempo n após uma medida da distância do alvo y_n foram assim encontradas.

Se faz conveniente escrever estas equações juntas como as equações de atualização (filtragem) de rastreamento posição presente e de velocidade $\alpha - \beta$:

$$\begin{aligned} \dot{x}_{n,n}^* &= \dot{x}_{n,n-1}^* + \alpha_n (y_n + \dot{x}_{n,n-1}^*) \\ \dot{x}_{n,n}^* &= \dot{x}_{n,n-1}^* + \beta_n \left(\frac{y_n - \dot{x}_{n,n-1}^*}{T} \right) \end{aligned}$$

Estas equações fornecem uma estimativa atualizada da posição do alvo e de sua velocidade, baseado na medida atual da distância do alvo y_n , assim como em medidas prévias. Estas equações são chamadas equações de filtragem. A estimativa $\dot{x}_{n,n}^*$ é chamada de estimativa filtrada, uma estimativa de x_n no tempo presente baseado no uso da medida atual y_n , assim como no uso medidas passadas. Esta estimativa contrasta com a estimativa de predição $\dot{x}_{n,n-1}^*$ que é uma estimativa de x_n baseada em medições passadas. O termo suavizado é algumas vezes usado no lugar de filtrado. O termo suavizado também é usado para indicar uma estimativa de uma posição ou uma velocidade de um alvo em um tempo passado entre a primeira e a última medida, por exemplo, uma estimativa $x_{\beta,n}^*$, onde $n_0 < \beta < n$, sendo que n_0 é o tempo da primeira medida e n o tempo da última (BROOKNER, 1998).

Agora se torna possível o uso das equações 5.1 e 5.2 para prever qual será a posição e a velocidade do alvo no tempo $n + 1$ e repetir todo o processo de atualização da velocidade e da posição do tempo $n + 1$ quando a medida de y_{n+1} no tempo $n + 1$ tenha sido feita. Para este propósito as equações 5.1 e 5.2 serão reescritas usando-se as notações de $\alpha - \beta$:

$$\begin{aligned}\dot{x}_{n+1,n}^* &= \dot{x}_{n,n}^* \\ x_{n+1,n}^* &= x_{n,n}^* + T\dot{x}_{n+1,n}^*\end{aligned}$$

Estas equações permitem a transição da velocidade e da posição no tempo n para o tempo $n + 1$ e são chamadas de equações de transição. Nota-se que a velocidade estimada no tempo $n + 1$, $\dot{x}_{n+1,n}^*$ é igual ao valor $\dot{x}_{n,n}^*$, porque foi suposto um modelo de velocidade constante. Estas equações combinadas com as equações de filtragem, permitem o rastreamento de um alvo. Geralmente em um radar de rastreamento, não se está interessado na posição atual do alvo x_n , e sim na posição prevista do alvo x_{n+1} para se ajustar a janela de previsão de distância. Logo combinando as equações de filtragem com as de transição temos:

$$\begin{aligned}\dot{x}_{n+1,n}^* &= \dot{x}_{n,n-1}^* + \frac{\beta_n}{T}(y_n - x_{n,n-1}^*) \\ x_{n+1,n}^* &= x_{n,n-1}^* + T\dot{x}_{n+1,n}^* + \alpha_n(y_n - x_{n,n-1}^*)\end{aligned}$$

Estas equações representam as equações de filtro e rastreamento $\alpha - \beta$. Estas equações são muito usadas em sistemas radar. Em contraste as equações de filtragem, as equações de filtro e rastreamento também são chamadas de equações de predição, pois prevêm a posição e a velocidade do alvo na próxima varredura (BROOKNER, 1998).

5.3.2 O filtro de Kalman

O filtro de Kalman é similar ao filtro de rastreamento $\alpha - \beta$ clássico, com exceção que é capaz de fornecer inerentemente o modelo dinâmico de alvos manobrantés. No filtro de Kalman um modelo para a medida do erro foi incluído, assim como um modelo para a trajetória do alvo e para as perturbações ou incertezas da trajetória. Estas perturbações da trajetória podem ser devido a um abandono das derivadas de mais alta ordem no modelo dinâmico, movimentos aleatórios devido a turbulência atmosférica, ou manobras deliberadas do alvo. O filtro de Kalman pode, em princípio, utilizar uma grande variedade de modelos para a medida da perturbação da trajetória e do ruído; contudo, supõe-se com frequência que são do tipo ruído branco e com média zero. Nem sempre um alvo manobrando se encaixa em um modelo ideal, pois é muito provável que produzam observações correlatas. A inclusão própria de um modelo dinâmico realista, aumenta a complexidade dos cálculos. É difícil de descrever *a priori*, a precisa natureza das perturbações. É necessário alguma forma de se adaptar as manobras. O filtro de Kalman é sofisticado e preciso, mas é mais complicado de se montar que muitos outros métodos usados para a predição de dados de rastreamento. Porém, se o filtro de Kalman fosse restrito ao modelo de trajetória de alvo em linha reta e a medida do ruído e do ruído de perturbação da trajetória fossem modelados como ruído branco Gaussiano de média zero, as equações do filtro de Kalman se reduziriam as equações do filtro $\alpha - \beta$ com os parâmetros α e β sendo computados sequencialmente pelo filtro de Kalman (SKOLNIK, 1981).

Até agora, foi usado um modelo determinístico para o movimento do alvo. Especificamente foi suposto que o alvo possui uma velocidade constante dada por:

$$x_{n+1} = x_n + T\dot{x}_n$$

$$\dot{x}_{n+1} = \dot{x}_n$$

No mundo real, o alvo não tem velocidade constante o tempo todo. Existe uma incerteza na trajetória do alvo, a aceleração do alvo ou mudança de direção em um dado momento. Kalman inseriu esta incerteza no movimento do alvo por meio da adição de um componente aleatório a dinâmica do alvo. Por exemplo, um

componente u_n deve ser adicionado a velocidade de alvo como indicado pelas seguintes equações de dinâmica do alvo:

$$x_{n+1} = x_n + T\dot{x}_n$$

$$\dot{x}_{n+1} = \dot{x}_n + u_n$$

onde u_n é uma variação aleatória na velocidade do tempo n para o tempo $n + 1$. Supõe-se que u_n é independente de n até $n + 1$ pra todo n e tem variância σ_n^2 . Fisicamente, u_n representa um salto aleatório da velocidade que ocorreu um instante antes da observação de $n + 1$ (BROOKNER, 1998).

O componente de velocidade aleatória u_n , é ajustado de forma a considerar uma possível aceleração do alvo, ou um movimento brusco. O modelo dinâmico aleatório tem o nome de ruído do processo, ruído de planta, ruído dinâmico ou ruído do sistema (BROOKNER, 1998).

Supondo que x_{n+1} representa a localização real do alvo no tempo $n + 1$. Supondo ainda que $x_{n+1,n}^*$ representa a posição prevista estimada do alvo no tempo $n + 1$ baseado nas medidas feitas até o momento, incluindo o tempo n . Kalman abordou a questão de como encontrar a estimativa ótima entre todas as classes de estimativas, lineares e não lineares, que minimizassem o erro médio quadrático

$$(x_{n+1,n}^* - x_{n+1})^2$$

Após muito esforço, Kalman descobriu que o filtro ótimo é dado pelas seguintes equações:

$$\dot{x}_{n+1,n}^* = \dot{x}_{n,n-1}^* + \frac{\beta_n}{T}(y_n - x_{n,n-1}^*)$$

$$x_{n+1,n}^* = x_{n,n-1}^* + T\dot{x}_{n+1,n}^* + \alpha_n(y_n - x_{n,n-1}^*)$$

Mas estas equações são idênticas as equações do filtro $\alpha - \beta$ especificado anteriormente. Mas para o filtro de Kalman, o peso das variáveis α e β dependem de n . Ainda mais, α_n e β_n são funções da variância da medida de posição do radar. Estas constantes de filtro também estão em função do quão preciso é o conhecimento da posição e da velocidade antes que qualquer medida seja feita, ou seja, o conhecimento anterior da trajetória do alvo, dados como a variância *a priori* da posição e da velocidade do alvo. No estado estacionário, as constantes de filtro α_n e β_n são dadas por:

$$\beta = \frac{\alpha^2}{2 - \alpha}$$

Portanto, o filtro de Kalman de estado estacionário, é idêntico ao filtro de Benedict-Bordner. O filtro de Kalman em sua forma geral foi desenvolvido antes do filtro de Benedict-Bordner. O filtro de Kalman foi apresentado pela primeira vez em 1960, enquanto que o filtro de Benedict-Bordner foi publicado em 1962 (BROOKNER, 1998).

O filtro de Kalman se utiliza de forma ótima das medidas do alvo, ajustando os coeficientes do filtro α e β de forma a considerar a precisão da n -ésima medida. Por exemplo, se na n -ésima medida a taxa da relação sinal/ruído for muito boa, mais precisa será a medição da posição do alvo, então α_n e β_n serão ajustados automaticamente para levar este fato em consideração. Especificamente, eles terão seu valor aumentado de forma a dar mais peso a esta medida mais precisa (BROOKNER, 1998).

O filtro de Kalman faz ótimo uso das informações *a priori*. Estas informações *a priori*, podem vir de outro radar que já venha rastreando o alvo a algum tempo e o está transferindo, como por exemplo, um radar de rota que está transferindo o alvo para a cobertura de um radar de aeroporto. Os dados do outro

radar podem ser usados para ajustar de maneira ótima os parâmetros α e β do novo radar. O filtro de Kalman escolhe automaticamente os pesos dos parâmetros α e β de forma que sejam de valor alto, como é necessário para uma ótima inicialização de rastreamento. Os pesos vão se ajustando lentamente de forma a atingirem um conjunto de pequenos valores constantes após a inicialização do rastreamento. O modelo dinâmico incorporado pelo filtro de Kalman permite a direta determinação da taxa de atualização do filtro por meio de:

$$T^2 \frac{\sigma_u^2}{\sigma_x^2} = \frac{\alpha^2}{(2 - \alpha)^2(1 - \alpha)} = \frac{\beta^2}{1 - \alpha}$$

Fianlmente, a da adição da variável de velocidade aleatória u_n força o filtro de Kalman ser sempre estável (BROOKNER, 1998).

Agora o filtro de Kalman será apresentado em sua forma matricial. Na forma matricial, o filtro de Kalman pode ser aplicado de forma mais abrangente que o caso unidimensional já apresentado.

Primeiro, o modelo dinâmico do sistema será apresentado na forma matricial. Em seguida, o modelo do sistema dinâmico aleatório será apresentado na forma matricial. A forma matricial do modelo do sistema dinâmico é a seguinte:

$$X_{n+1} = \Phi X_n,$$

onde:

$$X_n = \begin{bmatrix} x_n \\ \dot{x}_n \end{bmatrix} = \text{vetor de estados}$$

e

$\Phi = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$ = matriz do estado de transição em trajetória de velocidade constante

Para provar que a forma matricial é idêntica a forma apresentada pelas equações 5.1 e 5.2, basta substituir os valores e o resultado será:

$$\begin{bmatrix} x_{n+1} \\ \dot{x}_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_n \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} x_n + T\dot{x}_n \\ \dot{x}_n \end{bmatrix}$$

que tem a forma idêntica as equações 5.1 e 5.2.

O vetor X_n representa o estado da trajetória do alvo. Este vetor é representado por uma matriz coluna, e representa as quantidades sendo rastreadas. Para o filtro que está sendo considerado, estas quantidades representam a posição do alvo e sua velocidade no tempo n . Este vetor é denominado vetor de dois estados, pois representa dois estados do alvo, sua posição e sua velocidade. A matriz Φ , é a matriz de transição de estados. Esta matriz representa a transição do vetor de estados X_n no tempo n , para o vetor de estados X_{n+1} no tempo $n + 1$ um período T depois (BROOKNER, 1998).

Agora se torna simples a transformação do modelo de sistema dinâmico aleatório em sua forma matricial. A equação se torna:

$$X_{n+1} = \Phi X_n + U_n$$

onde;

$$U_n = \begin{bmatrix} 0 \\ u_n \end{bmatrix} = \text{vetor do modelo dinâmico de ruído do sistema}$$

Para novamente provar que as equações e a forma matricial são idênticas, faz-se novamente a substituição e temos:

$$\begin{bmatrix} x_{n+1} \\ \dot{x}_{n+1} \end{bmatrix} = \begin{bmatrix} x_n + T\dot{x}_n \\ \dot{x}_n \end{bmatrix} + \begin{bmatrix} 0 \\ u_n \end{bmatrix} = \begin{bmatrix} x_n + T\dot{x}_n \\ \dot{x}_n + u_n \end{bmatrix}$$

Que prova que ambas as formas são idênticas.

Agora a equação de medida do radar será posta na forma matricial, e será dada por:

$$Y_n = MX_n + N_n,$$

onde:

$$M = [1 \ 0] = \text{matriz de observação}$$

$$N_n = [v_n] = \text{erro da observação}$$

$$Y_n = [y_n] = \text{matriz de medidas}$$

A equação $Y_n = MX_n + N_n$ é chamada de equação de observação do sistema, porque relaciona as quantidades sendo estimadas com os parâmetros sendo observados.

Ao invés de colocar as equações de rastreamento na forma matricial, serão colocadas as equações de filtragem e predição, pois estas combinadas gerarão as equações de rastreamento. Colocando-se a equação de predição na forma matricial temos:

$$X_{n+1,n}^* = \Phi X_{n,n}^*$$

onde;

$$X_{n+1,n}^* = \begin{bmatrix} x_{n+1,n}^* \\ \dot{x}_{n+1,n}^* \end{bmatrix}$$

$$X_{n,n}^* = \begin{bmatrix} x_{n,n}^* \\ \dot{x}_{n,n}^* \end{bmatrix}$$

Esta forma tem o nome de predição porque prevê a posição e a velocidade do alvo no tempo $n + 1$ baseado na velocidade e na posição do alvo no tempo n , e a posição e a velocidade previstas são dadas pelo vetor de estados. Colocando-se a equação de filtragem na forma matricial temos:

$$X_{n,n}^* = X_{n,n-1}^* + H_n(Y_n - MX_{n,n-1}^*)$$

Esta equação é chamada de equação de filtragem porque fornece a estimativa atualizada da posição e velocidade presente do alvo. A matriz H_n é a matriz responsável por fornecer as constantes de filtragem-rastreamento α_n e β_n e é dada por:

$$H_n = \begin{bmatrix} \alpha_n \\ \beta_n \\ T \end{bmatrix}$$

Esta forma contudo, não diz como obter o valor das constantes. A forma seguinte nos mostra como obter estes valores:

$$H_n = \Phi S_{n,n-1}^* M^T [R_n + M S_{n,n-1}^* M^T]^{-1}$$

onde;

$$S_{n,n-1}^* = \Phi S_{n-1,n-1}^* \Phi^T + Q_n \text{ (equação do preditor)}$$

e

$$Q_n = COV[U_n] = E[U_n U_n^T] \text{ (modelo dinâmico de ruído de covariância)}$$

$$S_{n,n-1}^* = COV(X_{n,n-1}^*) = E[X_{n,n-1}^* X_{n,n-1}^{*T}]$$

$$R_n = COV(N_n) = E[N_n N_n^T] \text{ (covariância do ruído de observação)}$$

$$S_{n-1,n-1}^* = COV(X_{n-1,n-1}^*) = [I - H_{n-1}M]S_{n-1,n-2}^* \text{ (equação do corretor)}$$

As covariâncias em $Q_n = COV[U_n] = E[U_n U_n^T]$ e $R_n = COV[N_n] = E[N_n N_n^T]$ se aplicam enquanto as entradas da matrizes coluna U_n e N_n tenham média zero. Senão, U_n e N_n , tem de ser substituídas por $U_n - E[U_n]$ e $N_n - E[N_n]$, respectivamente. Fisicamente a matriz $S_{n,n-1}^*$ é uma estimativa da precisão em prever a posição e a velocidade do alvo no tempo n baseado nas medidas em $n-1$ e em tempos anteriores. Neste caso, $S_{n,n-1}^*$ é a matriz de covariância do vetor de estados $X_{n,n-1}^*$. A matriz R_n fornece a precisão das medidas do radar. É a matriz da covariância da matriz de erros de medida N_n . A matriz Q_n , fornece a magnitude da incerteza da trajetória do alvo, ou de forma equivalente, a capacidade de manobra do alvo. É a matriz de covariância do modelo dinâmico de ruído do sistema, ou seja, o componente de velocidade aleatório do alvo (BROOKNER,1998).

A equação $S_{n,n-1}^* = \Phi S_{n-1,n-1}^* \Phi^T + Q_n$ permite a obtenção da matriz de covariância da predição $S_{n,n-1}^*$, por meio da matriz de covariância da estimativa filtrada do vetor de estados do alvo no tempo $n-1$ dado por $S_{n-1,n-1}^*$. A matriz de covariância da estimativa filtrada, por sua vez, é obtida da matriz de previsões prévias $S_{n-1,n-2}^*$. As equações $H_n = \Phi S_{n,n-1}^* M^T [R_n + M S_{n,n-1}^* M^T]^{-1}$, $S_{n,n-1}^* = \Phi S_{n-1,n-1}^* \Phi^T + Q_n$ e $S_{n-1,n-1}^* = COV(X_{n-1,n-1}^*) = [I - H_{n-1}M]S_{n-1,n-2}^*$, permitem obter os pesos do filtro H_n após sucessivos intervalos de observação (BROOKNER, 1998).

O mais interessante da forma matricial do filtro de Kalman, é que as equações apresentadas para o caso unidimensional (distância) de dois estados (posição e velocidade), é que pode ser estendida para o caso geral, ou seja, pode ser utilizada para o rastreamento com qualquer número de dimensões, sendo necessário apenas a identificação apropriada do vetor de estados, da matriz de observação, da matriz de transição, do modelo dinâmico e da matriz de covariância das medidas (BROOKNER, 1998).

6 CONCLUSÃO

A quantidade de aeronaves que cruzam os céus hoje em dia, tem aumentado de forma acentuada a cada ano.

A precisão dos equipamentos de proteção ao vôo tem de acompanhar este crescimento de modo a não permitir que incidentes possam ocorrer abalando a confiança neste sistema.

Os sistemas de radar são parte integrante e muito importante do sistema de proteção ao vôo. Aonde quer que uma aeronave esteja voando, com certeza estará sob a cobertura de pelo menos um equipamento de radar.

O rastreamento correto dos alvos pelos sistemas de radar aumentam muito a segurança, reduzindo a possibilidade de falhas, tanto do próprio sistema quanto humanas.

Com o volume crescente de aeronaves, a precisão dos sistemas de rastreamento deve aumentar, pois a separação entre aeronaves em uma determinada região tem diminuído a cada dia, forçando o desenvolvimento de novas técnicas que permitam uma correta identificação e separação desses alvos.

As modernas técnicas de integração tem permitido que cada vez mais circuitos lógicos estejam alojados nos circuitos integrados, aliados a um aumento da velocidade de processamento das máquinas e redução dos custos, vem permitindo enormes avanços na área de processamento digital de sinais, área na qual os filtros digitais se encontram.

Há duas décadas atrás, a implementação de um filtro de Kalman em um equipamento de radar aumentava muito o custo final do equipamento, sendo utilizados outros tipos de filtros digitais mais simples como os do tipo $\alpha - \beta$, ao custo de um desempenho menor do sistema. Hoje os filtros do tipo $\alpha - \beta$ são utilizados como pré filtros em um sistema de rastreamento com filtro de Kalman. O filtro de

Kalman é capaz de uma resposta mais precisa e mais eficiente que outros tipos de filtros utilizados para rastreamento de alvos. Mas com a evolução dos sistemas novas soluções na área dos filtros deve ser encontrada de modo a melhorar ainda mais a eficiência do sistema.

Atualmente, se discute um projeto no qual as torres de controles dos aeroportos venham a ser totalmente automáticas, sem a presença de operadores, o que acabaria com a principal fonte de erros do sistema, o homem. Mas para que tal projeto possa seguir adiante, a qualidade dos equipamentos tem de se elevar. O rastreamento das aeronaves tem de ser praticamente perfeito, pois o menor erro pode colocar o avião fora da pista no momento do pouso, ou na própria pista, um radar de superfície poderia desviar a aeronave para fora da pista de rolamento.

A aplicação da forma estendida do filtro de Kalman no rastreamento de aeronaves já é uma realidade, e a aplicação de filtros Bayesianos de forma mais complexa que o de Kalman também já tem sido experimentados.

A conclusão que se pode chegar é que, graças a grande quantidade de pesquisa e desenvolvimento de novas tecnologias no setor, em um breve intervalo de tempo, novas soluções estarão disponíveis para aplicação na área de rastreamento de alvos, permitindo que novos horizontes dentro da proteção ao voo sejam alcançados.

REFERÊNCIAS BIBLIOGRÁFICAS

BROOKNER, Eli. *Tracking and Kalman filtering made easy*. New York: John Wiley & Sons. Inc. 1998. 463 p.

CHALLA, Subhash et al. *Fundamentals of object tracking*. San Diego: Cambridge University Press. 2011. 576 p.

DINIZ, Paulo Sérgio Ramirez; SILVA, Eduardo Antônio da; NETTO, Sérgio Lima. *Processamento digital de sinais: projeto e análise de sistemas*. Porto Alegre: Bookman. 2004. 590 p.

HAYKIN, Simon; VEEN, Barry Van. *Sinais e sistemas*. Porto Alegre: Bookman. 2002. 968 p.

JUCEWICZ, Stefan. *Radar*. São Paulo: Asa – Edições e artes gráficas. 1997. 144 p.

KOLAWOLE, Michael O. *Radar systems, peak detection and tracking*. Oxford: Newnes. 2002. 364 p.

LACOMME, Philippe et al. *Air and spaceborn radar systems: An introduction*. Norwich: Willian Andrews publishing. 2001. 504 p.

MAYBECK, Peter S. *Stochastic models, estimation, and control*. New York: Academic Press Inc. 1979. 423 p.

PROAKIS, John G. MANOLAKIS, Dimitris G. *Digital signal processing: principles, algorithms and applications*. 3rd. ed. New Jersey: Prentice Hall. 1996. 968 p.

SKOLNIK, Merril Ivan. *Introduction to radar systems*. 2nd ed. London: McGraw Hill. 1981. 581 p.

SKOLNIK, Merril Ivan. *Radar handbook*. 2nd ed. New York: McGraw Hill. 1990. 1244 p.

TAN, Li. *Digital signal processing. Fundamentals and applications*. San Diego: Academic press. 2008. 880 p.

WELCH, Greg. BISHOP, Gary. An introduction to Kalman filter. Chapel Hill: University of North Carolina. 2006. 16 p.