

Modelos de Previsão – Conceitos Básicos

O modelo estatístico proposto, adotado na dissertação é baseado no modelo Holt-Winters com múltiplos ciclos proposto por Taylor em 2002 (Taylor, 2002[27]). No entanto, antes que seja apresentado o modelo proposto, faz-se necessário a introdução a alguns conceitos estatísticos e de previsão. Após uma breve definição de séries temporais, serão apresentados e conceituados os mais importantes modelos estatísticos de previsão, e também modelos mais recentes, oriundos da inteligência computacional.

2.1

Séries Temporais

Podemos definir uma série temporal como sendo um grupo de dados observados em determinado momento, sendo esse espaço de tempo entre os dados disponíveis equidistante (horários, diário, semanal, mensal, trimestral, anual, e etc..). Para que uma determinada série seja classificada como uma série temporal, é necessário que preencha outro pré-requisito : os dados também devem apresentar uma dependência serial entre eles. Por exemplo : os dados de uma variável aleatória Z (consumo de energia) no instante t , com t variando de 1 até N , possa, de certa maneira, conter informações necessárias para que seja determinado o valor dessa variável no instante $t+1$. Cabe mencionar, que N representa o número de observações da série temporal em questão. As séries temporais podem ser classificadas como discretas, contínuas, determinísticas, estocásticas, multivariadas e multidimensionais.

Existem duas formas de se analisar uma série temporal : a análise no domínio do tempo, e a análise no domínio da frequência. No modelo proposto (Holt-Winters multiplicativo com dupla sazonalidade) a análise será feita no domínio do tempo, no

entanto, no modelo comparativo (amortecimento direto) a análise será feita no domínio da frequência. Essas duas formas de análise não são alternativas, mas sim complementares. No entanto, sabemos que a primeira se adequa mais a análise de processos não determinísticos, enquanto a segunda, a processos determinísticos.

2.1.1

Análise no Domínio do Tempo

Na análise no domínio do tempo considera-se a evolução da série temporal do processo que está sendo estudado, tendo como objetivo a determinação da magnitude de cada evento nos diversos instantes da série. As ferramentas utilizadas para essa análise são duas funções : a função de autocorrelação e a função de autocorrelação parcial. A função de autocorrelação mede a relação entre os eventos em diferentes instantes e as suas magnitudes. A análise é baseada, em geral, por modelos paramétricos.

2.1.2

Análise no Domínio da Frequência

Na análise no domínio da frequência, o interesse está em verificar a frequência que alguns eventos ocorrem em determinado período de tempo. A ferramenta utilizada para essa análise é o análise espectral, nela são estabelecidas as características de um processo estocástico em termos de frequências, podendo, no caso das séries temporais determinar as periodicidades existentes na mesma. Como o espectro de um processo não é conhecido, ele precisa ser estimado. Em geral, é estimado através do periodograma de janelas espectrais, por possuir boas propriedades estatísticas.

2.1.3

Comportamentos Cíclicos e Sazonalidade

As séries temporais podem também apresentar comportamentos cíclicos, variações periódicas. Ou seja, alguma dependência cíclica. É conveniente que se faça a medição da mesma a partir dos dados disponíveis e inclua essa sazonalidade dentro do modelo de previsão.

2.1.4

Previsão de Séries Temporais

A previsão de séries temporais é somente o estabelecimento de valores futuros para a série, e é feita com base tanto na informação atual quanto na passada. O horizonte de previsão é o comprimento do tempo, contado a partir de uma origem predeterminada (origem da previsão).

2.2

Modelos de Previsão

Para que sejam estabelecidos valores futuros para a série em estudo, é necessário que, de alguma forma, se possa captar e formular um modelo matemático capaz de representar o comportamento e as características da série temporal que se deseja prever. Essas informações são extraídas dos dados disponíveis. Existe uma grande quantidade de modelos de previsão, na literatura estatística adequados para exercer tal tarefa, e eles são classificados da seguinte forma :

Modelos Univariados

Os valores futuros de uma série são explicados somente pelos valores passados dessa mesma série. Isso acontece, por exemplo, em geral, com os dados de consumo de

carga de energia elétrica. Os métodos de decomposição, métodos de amortecimento exponencial e os modelos Box&Jenkins estão enquadrados dentro dessa classificação.

Modelos Causais ou Modelos de Função de Transferência

Os valores futuros de uma série são explicados não somente pelos valores passados da mesma, mas também por séries que de alguma forma possuam relação com ela. No caso do consumo de carga de energia elétrica, o preço relativo poderia ser uma série capaz de ajudar a explicar o mesmo.

Modelos Multivariados

São modelos capazes de realizar várias previsões ao mesmo tempo, um modelo único capaz de prever o futuro de diversas séries. Mantendo o foco sob a energia elétrica, tema de estudo, um exemplo de modelo multivariado seria um modelo que fosse capaz de prever ao mesmo tempo, o consumo de energia em diversas concessionárias prestadoras do serviço no país.

2.2.1

Modelo de Amortecimento Exponencial

Os métodos de amortecimento exponencial são classificados como sendo modelos automáticos e de validade local. Nesse breve resumo, iremos somente abordar os modelos de amortecimento exponencial, simples, duplo e triplo. As suas variantes sazonais, o modelo de amortecimento direto e o modelo de Holt-Winters, serão estudados mais a fundo nos capítulos 3 e 4 respectivamente. Existem duas formas de se obter as equações do modelo de amortecimento : de forma heurística, e por um procedimento de mínimos quadrados com desconto, sugerido por Brown (1963). Essa técnica foi, posteriormente, generalizada para abranger um maior tipo de funções (exponenciais, trigonométricas, e etc.).

O modelo de amortecimento exponencial simples é um modelo constante, apropriado para séries que são caracterizadas localmente por seu nível, acrescentado de uma variação aleatória desprezível. Ou seja, as mudanças ocorridas numa série nesse formato são bastante lentas. No amortecimento exponencial duplo, a série possui uma tendência aditiva. Diferente daquele, no amortecimento simples, a média não é constante ao longo do tempo, ocorre, sim, uma mudança linear. No caso do amortecimento triplo, a tendência existente é quadrática, sendo seu efeito multiplicativo. Abaixo estão dispostos exemplos de modelos série x_t para cada uma das formas apresentadas.

Amortecimento Exponencial Simples

$$x_t = b + a_t \quad (2.1)$$

Amortecimento Exponencial Duplo

$$x_t = b_1 + b_2 t + a_t \quad (2.2)$$

Amortecimento Exponencial Triplo

$$x_t = b_1 + b_2 t + \frac{1}{2} b_3 t^2 + a_t \quad (2.3)$$

com $t = 1, \dots, n$

De forma geral, assume-se que o nível médio das observações pode ser escrito, para cada instante de tempo, através de uma função conhecida. O objetivo desse método é estimar parâmetros que caracterizem a função $f(t)$. Para uma dada série $x_t, t = 1, 2, \dots$ sendo $\mu(t) = E(x_t)$ será considerado o seguinte modelo estocástico

$$x_t = \mu(t) + \varepsilon_t \quad (2.4)$$

dado que $\mu(t) = f(t)$, então

$$x_t = f(t) + \varepsilon_t \quad (2.5)$$

sendo ε_t o ruído do sistema no instante t com média nula e variância constante.

Na técnica de amortecimento exponencial, os estimadores são obtidos via médias móveis de tamanho N . O tamanho das médias móveis é determinado pela quantidade de parâmetros a serem estimados. A fórmula geral para a determinação dos mesmos seria a seguinte :

$$M_T = \alpha x_t + (1 - \alpha)M_{T-1} \quad (2.6)$$

$$M_T^i = \alpha M_T^{i-1} + (1 - \alpha)M_{t-1}^i \quad (2.7)$$

A equação de previsão do modelo para τ passos-à-frente é a seguinte

$$x_T(\tau) = f(T + \tau) \quad (2.8)$$

onde

$$x_T(\tau) = E[x_{T+\tau} | x^T] \quad (2.9)$$

Para séries não sazonais, a função frequentemente empregada é do tipo polinomial

$$f(t) = \sum_i a_{i+1} t^i + \varepsilon_t \quad (2.10)$$

passando a equação do modelo a ser a seguinte

$$x_t = \sum_{i=0}^n a_{i+1} t^i + \varepsilon_t \quad (2.11)$$

e sua equação de previsão

$$x_T(\tau) = \sum_{i=0}^n a_{i+1} (T + \tau)^i \quad (2.12)$$

Para séries sazonais, entenda por séries sazonais, séries que possuem repetição periódicas, é necessário, que de alguma maneira se inclua essa informação no nível médio da série. Isso pode ser feito através de uma modelagem via fatores sazonais (variáveis *dummy*) ou via função trigonométrica (combinação de senos e cosenos). A modelagem via fatores sazonais pode ser feita de forma aditiva ou multiplicativa, como visto abaixo :

Modelo aditivo :

$$x_t = \mu(t) + \rho_t + \varepsilon_t \quad (2.13)$$

Modelo multiplicativo :

$$x_t = \mu(t) \times \rho_t \times \varepsilon_t \quad (2.14)$$

A equação de previsão do modelo multiplicativo τ passos-à-frente será

$$x_T(\tau) = \sum_{i=0}^n \hat{a}_{i+1} (T + \tau)^i \rho(T + \tau) \quad (2.15)$$

onde ρ_t são os fatores sazonais. Tanto \hat{a} quanto $\hat{\rho}_t$ são estimados seguindo o mesmo procedimento seqüencial, demonstrado anteriormente.

2.2.2

Modelos Box&Jenkins

Interpreta a série temporal como uma realização de um vetor aleatório multivariado, sendo sua dimensão o tamanho da séries disponível, seguindo o princípio da parcimônia (representação dos modelos matemáticos, com a menor quantidade de parâmetros possível) e da construção de modelos através de um ciclo iterativo. Os modelos Box&Jenkins determinam o processo estocástico adequado para representar uma dada série temporal através da passagem de um ruído branco por um filtro linear. A representação do processo utilizado por Box&Jenkins pode ser observado na Figura 2.1.



Figura 2.1 –Processo Previsor de Box & Jenkins

Como na vida cotidiana nem todos os processos são estacionários, busca-se um operador de retardo, para que seja permitida a construção de modelos para séries com comportamentos seriais, através da descrição delas por processos estacionários ou processos estacionários homogêneos. O processo não estacionário é transformado em estacionário homogêneo, através de diferenciações sucessivas, feitas, pelo uso do operador de retardo mencionado anteriormente.

$$x_t = \mu + \sum_{k=0}^{\infty} \Psi_k a_{t-k} \quad (2.16)$$

fazendo $\tilde{x} = x_t - \mu$, temos que

$$\tilde{x}_t = \sum_{k=0}^{\infty} \Psi_k a_{t-k} \quad (2.17)$$

onde

μ - nível da série.

a_t - ruído branco.

Ψ_k - função de transferência ou filtro linear definido como $\frac{\theta_q(B)}{\phi_p(B)}$.

B - operador retardo que representa um atraso de um período de tempo.

A equação 2.16 é a formulação geral do modelo Box&Jenkins. Primeiramente foram desenvolvidos os modelos ARMA, com o objetivo de se modelar somente séries estacionárias, como pode ser observado na equação abaixo.

$$\phi(B)X_t = \theta(B)a_t \quad (2.18)$$

No entanto, sabe-se que no mundo real, a maioria das séries são não estacionárias. Então, os modelos ARIMA foram a forma encontrada para solucionar esse problema, podendo assim modelar também séries não estacionárias homogêneas. Entende-se por séries não estacionárias homogêneas, aquelas que não são estacionárias na média ou no nível, sendo as mesmas transformadas em séries estacionárias através de sucessivas diferenças. Se mesmo depois de terem sido feitas as sucessivas diferenças, a série não se tornar estacionária, convém fazer a transformação logarítmica da série em questão. Mas é preciso estar atento, pois a introdução de transformações na previsão pode causar tendenciosidade (no caso da transformação logarítmica geram estimadores tendenciosos do anti-log), sendo necessário a aplicação de correções para eliminá-la. Com a inclusão do operador de diferença, a formula geral, é a que segue.

$$\phi(B)\nabla^d x_t = \theta(B)a_t \quad (2.19)$$

E finalmente, os modelos SARIMA são uma extensão dos modelos ARIMA para que seja incluída a sazonalidade dentro da modelagem. Ou seja, inclui-se na modelagem a correlação serial dentro e entre períodos sazonais.

$$\varphi(B)\Phi(B^S)\nabla_S^D\nabla^d x_t = \theta(B)\Theta(B^S)a_t \quad (2.20)$$

onde

$\varphi(B)$ - operador auto-regressivo não sazonal.

$\nabla^d = (1 - B)^d$ - operador de diferença não sazonal de ordem d.

$\Phi(B^S)$ - operador auto-regressivo sazonal.

$\nabla_S^D = (1 - B^S)^D$ - operador de diferença sazonal de ordem D.

$\theta(B)$ - operador de médias móveis não sazonal.

$\Theta(B^S)$ - operador médias móveis sazonal.

A modelagem geral de Box&Jenkins segue na equação acima, que pode ser usada tanto para modelos sazonais (SARIMA(p,d,q)x(P,D,Q)s) quanto para os modelos não sazonais (ARIMA(p,d,q)).

A estimação dos parâmetros do modelo é feita por máxima verossimilhança. No entanto, existem dois problemas na estimação da máxima verossimilhança : o primeiro problema, seria o estabelecimento dos valores iniciais do modelo, e o segundo a minimização da soma dos quadrados dos resíduos, que seria o equivalente a maximizar a verossimilhança, pode talvez não conduzir a uma função linear dos parâmetros.

2.2.2.1

Testes de Adequação do Modelo

Existem diversos testes para verificar se o modelo estimado é adequado. Os sugeridos por Box&Jenkins são os testes de sobredefinição e o teste dos resíduos gerados, que pode ser feito através do teste de Portmanteau, e o teste de periodograma acumulado.

No teste de sobredefinição, elaboram-se modelos com número de parâmetros superior ao do modelo escolhido, para que seja testada a significância dos parâmetros adicionais. Caso sejam encontrados parâmetros significativos, ficará claro que o modelo foi subidentificado.

No teste dos resíduos, o teste de Portmanteau estima a autocorrelação dos mesmos e calcula-se a estatística

$$Q = n \sum_{j=1}^k r_j^2(\hat{a}) \quad (2.21)$$

A hipótese nula é a de que os resíduos são aleatórios, e para que isso ocorra, o valor de Q deve ser menor que o valor da abscissa da função Qui-Quadrado com (k-p-q) graus de liberdade. No teste do periodograma acumulado, compara-se o periodograma acumulado da série dos resíduos com o de um ruído branco, com o intuito de se encontrar componentes periódicos (Montgomery, 1976[17]).

Nos estudos usando esse método são, normalmente, utilizadas séries de carga semanal ou mensal, visto que o modelo como proposto inicialmente, somente comporta a sazonalidade, não tendo como modelarmos os ciclos. Também, os softwares disponíveis para a modelagem do método, tais como Forecast Pro e Autobox, somente comportam até 500 dados, não podendo, portanto, modelar fazendo uso de múltiplos ciclos. O software Autobox está atualmente desenvolvendo uma nova versão, onde poderão ser incluídos mais dados, existindo, conseqüentemente, a possibilidade de se modelar os múltiplos ciclos existentes numa série de carga horária, ou de ½ em ½ hora. Ou seja, com essa evolução será

possível utilizar os modelos Box&Jenkins para a modelagem de múltiplos ciclos. O SPSS também é um software capaz de modelar o métodos Box&Jenkins, mas também não é capaz de tratar múltiplos ciclos. O PREVCAR também é um sistema de previsão de séries mensais de carga, utilizado dentro do setor de energia elétrica brasileiro. Nele são combinadas previsões obtidas através abordagens univariadas de previsão de séries temporais. São utilizados dentro do sistema, dois modelos de natureza estocástica, o de Box&Jenkins e o de Holt-Winters, e dois provenientes da inteligência computacional, o de Redes Neurais e de Lógica Nebulosa. Os resultados obtidos com essa ferramenta foram bastante satisfatórios.(Souza, 2003[26]).

2.2.2.2

Extensão do Modelo Box&Jenkins

Taylor fez uma extensão dos modelos Box&Jenkins, com o objetivo de fornecer uma forma de se incluir no método original, ciclos, tal como feito no modelo proposto na dissertação. O autor utilizou a extensão a fim de modelar 2 ciclos sazonais, um ciclo diário e um ciclo semanal, existentes numa série de carga de $\frac{1}{2}$ em $\frac{1}{2}$ hora de energia elétrica. Darbellay e Slama usaram também a mesma modelagem, só que para prever dados de carga horária. A extensão do modelo para os dois ciclos sazonais é a que está disposta na equação 2.22. A expressão desse modelo pode ser escrita da seguinte forma, SARIMA(p,d,q)x(P₁,D₁,Q₁)s₁x(P₂,D₂,Q₂)s₂.

$$\varphi(B)\Phi(B^{S_1})\Omega(B^{S_2})\nabla_{S_1}^{D_1}\nabla_{S_2}^{D_2}\nabla^d x_t = \theta(B)\Theta(B^{S_1})\Psi(B^{S_2})a_t \quad (2.22)$$

onde

$\varphi(B)$ - operador auto-regressivo não sazonal.

$\nabla^d = (1 - B)^d$ - operador de diferença não sazonal de ordem d.

$\Phi(B^{S_1})$ - operador auto-regressivo do ciclo diário.

$\Omega(B^{S_2})$ - operador auto-regressivo do ciclo semanal.

$\nabla_{S_1}^{D_1} = (1 - B^{S_1})^{D_1}$ - operador de diferença do ciclo diário de ordem D_1 .

$\nabla_{S_2}^{D_2} = (1 - B^{S_2})^{D_2}$ - operador de diferença do ciclo semanal de ordem D_2 .

$\theta(B)$ - operador de médias móveis não sazonal.

$\Theta(B^{S_1})$ - operador médias móveis do ciclo diário.

$\Psi(B^{S_2})$ - operador médias móveis do ciclo semanal.

O modelo pode facilmente ser estendido para 3 ou mais ciclos sazonais, para isso, somente faz-se necessário a inclusão de funções polinomiais adicionais do operador auto-regressivo, e operadores de diferença, equivalentes a quantidades de ciclos que se deseja incluir no modelo. (Taylor, 2002[27]). Então, com isso, pode-se incluir, a sazonalidade (influência das 4 estações do ano) dentro do modelo, caso o investigador tenha interesse em fazê-lo. Contudo, é comum assumir que essa componente não é significativa para intervalos de dados inferiores a um dia. Ou seja, somente se considera interessante a inclusão da sazonalidade para séries, diárias, semanais, trimestrais, semestrais ou anuais. No caso se séries horárias e de $\frac{1}{2}$ em $\frac{1}{2}$ hora torna-se desnecessária a inclusão, pois se acredita que a transição é suave, podendo ser captada nos dados.

2.2.3

Modelo de Espaço Estado e Filtro de Kalman

O filtro de Kalman é um método de processamento de sinais que fornece estimativas ótimas do estado atual de um sistema dinâmico. Consiste em um grupo de equações para estimar recursivamente o estado corrente do sistema e para encontrar a variância dessas estimativas.

2.2.3.1

Modelo de Espaço Estado

Nos modelos de espaço estado o sinal é tido como sendo uma combinação linear de variáveis, chamada variáveis de espaço estado, que constituem o vetor de estado no tempo

t. Esse vetor descreve o estado de um sistema no tempo t, e é também conhecido como “estado da natureza”. Denota-se o vetor de estado (mx1) por θ_t , que pode ser calculado seguindo a equação das observações disposta abaixo

$$x_t = h_t^T \theta_t + n_t \quad (2.23)$$

onde o vetor h_t (mx1) é conhecido e n_t denota o erro das observações, que são, por convenção, assumidos como sendo descorrelacionados.

O vetor de θ_t não pode ser observado diretamente, então usam-se as observações de x_t para que sejam feitas inferências a respeito de θ_t . Assumindo que θ_t muda ao longo do tempo, denotamos a equação de transição como

$$\theta_t = G_t \theta_{t-1} + w_t \quad (2.24)$$

onde G_t é uma matriz (mxm) conhecida e w_t representa o vetor dos desvios. Os dois erros do modelo, n_t e w_t , são assumidos, geralmente, como sendo independentes e normalmente distribuídos com média zero e variâncias σ_n^2 e σ_w^2 , respectivamente. A razão entre essas duas variâncias é conhecida como razão do ruído do sinal.

A equação 2.26 constitui a fórmula geral do modelo de espaço-estado univariado. O modelo pode, facilmente ser generalizado para o caso no qual x_T é um vetor, passando o modelo a ser multivariado, fazendo h_t uma matriz do tamanho apropriado e, n_t um vetor de comprimento apropriado. Uma das características mais importantes dos modelos estruturais, e em especial dos modelos de espaço-estado, é que a equação das observações envolve uma função linear das variáveis de estado e obriga o modelo a ser constante ao longo do tempo.

A partir da equação 2.26, podemos obter alguns casos particulares do modelo. Tais como as que serão apresentadas a seguir.

Modelo Estacionário

Supondo a equação das observações como sendo

$$x_t = \mu_t + n_t \quad (2.25)$$

onde μ_t o nível corrente, uma variável não observada, e segue um passeio aleatório dado por

$$\mu_t = \mu_{t-1} + w_t \quad (2.26)$$

No caso apresentado através das equações 2.27 e 2.28, θ_t é considerada somente uma única variável de estado, μ_t , enquanto G_t e h_t são constantes escalares unitárias. O modelo apresentado é chamado dessa forma por não conter tendência incluída. Caso, $w_t = 0 \quad \forall t$, pode ser verificado que o modelo estacionário se transforma em um modelo estático.

Modelo de Crescimento Linear

O modelo é representado através das seguintes 3 equações

$$x_t = \mu_t + n_t \quad (2.27)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + w_{1,t} \quad (2.28)$$

$$\beta_t = \beta_{t-1} + w_{2,t} \quad (2.29)$$

Nesse modelo o vetor de estado θ_t possui dois componentes, o nível, μ_t , que varia linearmente ao longo do tempo, e a tendência local, β_t , que pode evoluir no tempo. No entanto, é improvável que a situação apresentada pelo modelo, onde a tendência é constante no tempo, ocorra. Por isso, usa-se um modelo onde a tendência possa variar.

Modelo Estrutural Básico

Nesse modelo a componente de sazonalidade é incorporada ao modelo, tal como pode ser observado nas equações (2.32) a (2.34). Nele além das componentes de nível, μ_t , e tendência, β_t , é incluída uma componente sazonal, i_t , que assume-se como sendo aditiva.

$$x_t = \mu_t + i_t + n_t \quad (2.30)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + w_{1,t} \quad (2.31)$$

$$\beta_t = \beta_{t-1} + w_{2,t} \quad (2.32)$$

$$i_t = -\sum_{j=1}^{s-1} i_{t-j} + w_{3,t} \quad (2.33)$$

2.2.3.2

Filtro de Kalman

O filtro de Kalman fornece uma forma geral de estimar o sinal com a presença de ruído, objetivo principal da modelagem de espaço-estado. O método é formado por um grupo de equações que permitem ao investigador atualizar a estimativa de θ_t quando se torna disponível uma nova observação. O procedimento de atualização é composto por dois estágios, o estágio de previsão e o de atualização das estimativas do modelo.

Suponha que observamos uma série temporal até o período $t-1$, e que $\hat{\theta}_{t-1}$ é o melhor estimador de θ_{t-1} com base nas informações até o período mencionado. Suponha também que foi avaliada a matriz variância-covariância de $\hat{\theta}_{t-1}$. No estágio de previsão, o

interesse está concentrado na previsão de θ_t para o tempo $t-1$. As equações (2.34) e (2.35) são as equações de previsão do modelo. O estimador de θ_t é dado por

$$\hat{\theta}_{t|t-1} = G_t \hat{\theta}_{t-1} \quad (2.34)$$

e a matriz de variância-covariância é dada por

$$P_{t|t-1} = G_t P_{t-1} G_t^T + W_t \quad (2.35)$$

Quando novas observações se encontram disponíveis no tempo t , o estimador de θ_t pode ser modificado, passando assim, a levar em consideração as informações trazidas por essa nova observação. Esse processo compreende o segundo estágio do filtro de Kalman, a etapa de atualização. As equações de previsão são dadas por

$$\hat{\theta}_t = \hat{\theta}_{t|t-1} + K_t e_t \quad (2.36)$$

e

$$P_t = P_{t|t-1} - K_t h_t^T P_{t|t-1} \quad (2.37)$$

onde

$$K_t = P_{t|t-1} h_t / [h_t^T P_{t|t-1} h_t + \sigma_n^2] \quad (2.38)$$

A maior vantagem prática do filtro de Kalman são que os cálculos são recursivos, então apesar da estimativa estar sendo baseada na série histórica toda, não existe a necessidade de uma memória que esteja sempre em expansão. Outra vantagem reside no fato de convergir rapidamente quando o modelo implícito é constante.

Para inicializar o filtro de Kalman, precisa-se calcular os valores iniciais de θ_t e P_t . Esses valores podem ser encontrados de duas formas: por suposição, baseado no fato do filtro atualizar rapidamente esses valores fazendo que as escolhas iniciais sejam dominadas pelos dados; e pode-se também estimar o vetor ($m \times 1$) de θ_t no tempo $t = m$ por mínimos quadrados para as primeiras m observações.

O filtro de Kalman é utilizado para modelos espaço-estado que são lineares nos parâmetros. Para séries não lineares, pode-se aplicar um filtro, conhecido como filtro de Kalman estendido. O filtro de Kalman estendido é feito através de uma aproximação linear local para o modelo em questão.

Em estudo realizado por Gordon (Gordon, 1996[7]), a autora modela dados diários de carga de energia elétrica através de uma abordagem estrutural. Sabemos que os dados diários apresentam dois tipos de ciclo, sendo o primeiro um ciclo semanal, e o segundo um anual (sazonalidade). O ciclo semanal pode ser modelado por um modelo estrutural básico. No entanto, o ciclo anual, se fosse modelado pelos meios tradicionais, utilizaria uma grande quantidade de parâmetros, violando assim, o princípio da parcimônia. O que a autora faz para contornar esse problema é utilizar a técnica de splines para a modelagem do ciclo anual. Essa técnica tem como principal vantagem a simplicidade da sua implementação e a praticidade de permitir que um efeito não linear se transforme em uma regressão múltipla, facilitando, assim a sua estimação. A autora também trata os feriados de forma diferenciada, por causa da sua grande influência nas variações do consumo de energia elétrica.

2.2.4

Modelos Bayesianos de Previsão

O conhecimento se encontra disponível de diversas maneiras, sendo uma classificação útil classificá-lo como proveniente de informações históricas ou conhecimento profissional. Para a análise das informações históricas podemos derivar um modelo de previsão. No entanto, não somente esse conhecimento é importante, tendo também muita

importância um outro conhecimento que não pode ser explicado por esse modelo formal. O paradigma bayesiano fornece uma ferramenta lógica, racional e coerente de se combinar informações. O que o paradigma faz é criar uma rotina de modelos de previsão capaz de ser ajustada por intervenção subjetiva, refletindo assim, a informação subjetiva mencionada anteriormente. (West,1989[28]). A linha de pensamento bayesiana é disciplinada pelo senso comum. (Smith, 1991[22]). A disciplina do pensamento bayesiano e o mecanismo formal de aprendizado do Teorema de Bayes evita que ocorram problemas lógicos garantindo, assim, a validade das conclusões obtidas a partir das premissas feitas.

O método bayesiano de previsão tem a capacidade de incorporar informações, além das já existentes na série histórica estudada. Ou seja, ocorre no modelo a inclusão de informações subjetivas, além de permitir a adaptabilidade dos parâmetros do modelo. No método bayesiano de previsão obtém resultados muito mais satisfatórios quando não existem estatísticas suficientes para resolver e/ou modelar o problema em questão. As outras teorias de inferência produzem bons resultados somente quando existem hipóteses feitas sobre o problema, tais como, normalidade e independência dos erros. Quanto não se tem nenhuma hipótese por trás da modelagem, normalmente, eles geram resultados insatisfatórios e confusos.

O modelo linear dinâmico é o que estrutura analiticamente o modelo bayesiano de previsão. Abaixo iremos apresentar o modelo linear dinâmico univariado.

Equação das Observações

$$x_t = F_t' \theta_t + v_t \quad v_t \sim N[0, V_t] \quad (2.39)$$

onde

x_t - série de dados no tempo t .

F_t - vetor das constantes conhecidas (vetor de regressão).

θ_t - vetor dos parâmetros do modelo de estado.

v_t - termo do erro estocástico, tendo a mesma distribuição normal com média zero e variância V_t .

Equação do Sistema

$$\theta_t = G_t \theta_{t-1} + \omega_t \quad \omega_t \sim N[0, W_t] \quad (2.40)$$

onde

G_t - matriz dos coeficientes conhecidos que definem a evolução sistemática do vetor de estado ao longo do tempo.

ω_t - termo do erro estocástico tendo a mesma distribuição normal com média zero e matriz de variância W_t .

As duas séries estocásticas, $\{v_t\}$ e $\{\omega_t\}$, são assumidas como sendo independentes e mutuamente exclusivos. Ou seja, as covariâncias $Cov[v_t, v_s]$, $Cov[\omega_t, \omega_s]$ para todo $t \neq s$, e $Cov[v_t, \omega_s]$ para todo t, s tem valor zero.

O aprendizado bayesiano acontece através da combinação de informações provenientes das observações de uma dada série através da função de verossimilhança. O mecanismo de combinação é o Teorema de Bayes, que nada mais é que um simples teorema inverso de probabilidade.

2.2.4.1

Teorema de Bayes

Supondo que $x' = (x_1, \dots, x_n)$ é um vetor de n observações com uma distribuição de probabilidade $p(x|\theta)$ que depende dos valores dos k parâmetros $\theta' = (\theta_1, \dots, \theta_k)$.

Suponha também que o θ tenha uma distribuição de probabilidade $p(\theta)$. Então, teorema de Bayes diz que

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} \quad (2.41)$$

2.2.4.2

Informação a Priori

A distribuição *a priori* que representa o conhecimento a priori ou uma relativa ignorância. Supõe-se que essa distribuição é capaz de representar tudo o que é conhecido, naquele dado momento, sobre os parâmetros, antes de estarmos em posse dos dados. A determinação dessa distribuição tem um papel crucial na modelagem bayesiana. A utilidade da teoria se dá por causa da tentativa de ser adaptar o modelo, no caso, a distribuição a priori, considerando se as conseqüências concordam com o senso comum ou não, e se mostram, dão uma noção de onde o senso comum falha.

A informação a priori para um vetor de estado no tempo $(t + 1)$ pode ser resumido como uma distribuição normal com média a_{t+1} e covariância R_{t+1} ,

$$\theta_{t+1}|D_t \sim N[a_{t+1}, R_{t+1}] \quad (2.42)$$

onde D_t o estado de conhecimento no tempo t .

As previsões do modelo bayesiano são geradas a partir dessa informação a priori, juntamente com a equação das observações.

2.2.4.3

Previsão k passos a frente

Para que possa ser feita a previsão k passos a frente, é necessário que a informação a priori seja projetada para o futuro, aplicando o sistema de equações repetidas vezes, até que sejam obtidos todos os valores que se deseja prever. Abaixo segue a equação de previsão para k passos a frente, supondo que essa previsão tem uma distribuição normal.

$$\theta_{t+k} | D_t \sim N[a_t(k), R_t(k)] \quad (2.43)$$

onde, para $k \geq 2$, a média e a variância são dadas por

$$a_t(k) = G^{k-1} a_{t+1} \quad (2.44)$$

$$R_t(k) = G^{k-1} R_{t+1} (G^{k-1})' + \sum_{j=2}^k G^{k-j} W_{t+j} (G^{k-j})' \quad (2.45)$$

Cabe mencionar que a soma começa a partir de $j = 2$ porque a evolução estocástica da variância W_{t+1} já está incluída na variância da priori R_{t+1} . Em posse dessa previsão para o estado a previsão associada para as observações da série em questão é obtida a partir da equação de previsão como

$$x_{t+k} | D_t \sim N[f_t(k), Q_t(k)] \quad (2.46)$$

onde os momentos são definidos nos termos familiares

$$f_t(k) = F'_{t+k} a_t(k) \quad (2.47)$$

$$Q_t(k) = F'_{t+k} R_t(k) F_{t+k} + V_{t+k} \quad (2.48)$$

2.2.4.4

Componentes do Modelo Linear Dinâmico

A modelagem de uma série temporal é extremamente facilitada quando se determinam as estruturas internas do modelo. Ou seja, quando se identificam as componentes dessa série temporal, tais como tendência, sazonalidade, fator de desconto. A modelagem é feita seguindo a mesma estrutura dos modelos de espaço-estado. Em baixo seguem exemplos das componentes principais do modelo linear dinâmico. Atenção especial será dada a componente de sazonalidade.

A componente de tendência podem ser exemplificadas pelas equações 2.25 e 2.25, no caso de uma modelagem simples de polinômio de primeira ordem, e pelas equações de 2.27 a 2.29, onde o polinômio apresentado permite a incorporação de um aumento ou redução sistemática do nível dentro da modelagem. Podemos observar que as equações mencionadas, tratam essa componente como sendo um passeio aleatório ao longo do tempo. Apesar de pouco usados, polinômios de ordens superiores podem ser generalizados, se estendendo os dois polinômios descritos (West , 1989[28]).

Quanto aos padrões sazonais, sabemos que modelá-los em uma dada série temporal requer que a componente tenha uma forma que seja periódica. A forma mais comum, mais utilizada de se representar esses padrões sazonais é a modelagem via fatores sazonais, no qual são determinados diferentes parâmetros para cada ponto no ciclo. No entanto, existe uma forma alternativa de se representar esses padrões sazonais. Nela se separa a tendência implícita do movimento que ocorre sobre essa tendência. Durante um ciclo completo esse efeito soma zero dado que a tendência, que é a média dos fatores ao longo de todo o ciclo, contém todo o movimento da série durante essa amplitude de tempo. Esse modelo de efeito sazonal define parâmetros para medir as partes sazonais da tendência. Na prática, essa forma de modelar é preferível se comparada com a via fatores sazonais, porque ela facilita a divisão entre o que na série representa a tendência, e o que representa a variação sazonal. Por exemplo, se temos dados trimestrais, para um ciclo anual, o efeito sazonal no sistema de espaço estado do modelo linear dinâmico compreende 4 parâmetros, um para cada

trimestre. E a medida que o tempo vai passando, ou melhor, os trimestres vão passando, o vetor simplesmente roda um elemento a cada período.

2.2.4.5

Fatores de Desconto

Os fatores de desconto são usados, em dados provenientes de séries temporais, são utilizados como uma ferramenta para se demonstrar o grau de importância que determinado dado possui na explicação daquela série. Ou seja, os dados mais recentes possuem um conteúdo informativo muito mais importante do que os dados mais antigos, e isso precisa ser, de alguma forma, modelado. Os fatores de desconto exercem essa função. O grande problema é encontrar a taxa para esse fator, de forma que ele seja capaz de descontar a perda de informação (perda multiplicativa) de forma correta. Ou seja, essa taxa de desconto deve ser equivalente ao aumento de incerteza dos dados, e portanto perda de precisão. O fator de desconto é um valor definido entre zero e um. Nos modelos de previsão bayesianos a incorporação dos fatores de desconto no modelo, permite com que os mesmos sejam utilizados para possíveis intervenções no sistema. Como são parâmetros livres, possuem uma certa parcimônia. Esse é um dos fatores mais importantes do seu uso, dado que é muito mais simples estabelecer o valor do fator de desconto do que determinar a matriz de covariância do sistema (Reinaldo, notas de aula,2001).

Atualmente, existem diversos trabalhos onde se utiliza os modelos bayesianos para a previsão de carga de energia elétrica. No entanto, tal como em todos os métodos estudados até o presente momento, não se encontra disponível a modelagem de ciclos, nem uma extensão do método capaz de fazê-lo, tal como apresentando nos modelos Box&Jenkins. Tão pouco o software disponível para a modelagem, o BATS, é capaz de fornecer uma alternativa a esse problema.

Dentre os trabalhos realizados utilizando a modelagem bayesiana, podemos citar um trabalho onde o autor (Fernandes, 1985[9]) utiliza o método de previsão de séries temporais que apresentam descontinuidades bruscas. Esse método é conhecido como Modelo de

Crescimento Linear de Estados Múltiplos (MCL-EM), tendo sido o mesmo desenvolvido por Harrison e Stevens. No estudo, Fernandes (1985[9]) aplica esse modelo em séries da economia brasileira, tais como, de demanda de energia elétrica, índices de preços setoriais e taxa de câmbio, comparando os resultados obtidos com outros métodos disponíveis. O método de MCL-EM é o único modelo existente que admite variações bruscas nos próprios parâmetros, sendo capaz, de se adaptar com mais rapidez e facilidade as discontinuidades existentes nas séries. Essa característica, em conjunto, com a praticidade do uso do filtro de Kalman, faz com que seu uso seja bastante eficiente e eficaz em sistemas automáticos destinados a prever rapidamente um grande número de séries (Fernandes, 1985[9]). Também na dissertação de mestrado de Mendes (2002[16]), o autor utiliza os modelos bayesianos de previsão no tratamento dos dados do período do racionamento de energia elétrica.

2.2.5

Modelo de Decomposição de Gupta

Nesse item da dissertação descreveremos, brevemente, um modelo diferente dos métodos de previsão citados anteriormente, conhecido como modelo de decomposição de Gupta (Bunn, 1985[3]).

O procedimento de previsão é feito da seguinte forma : combina-se dois modelos de previsão; um modelo estocástico, que relaciona, as cargas futuras com as cargas passadas; e um segundo o modelo, que é o modelo que estabelece a relação entre tempo e carga, com o objetivo de representar a influência das variáveis de tempo nas cargas futuras. O modelo é atualizado a medida que as previsões são produzidas.

O sistema de carga horária é composto por 3 componentes, que estão representadas na equação abaixo

$$Z(i, j) = T(i, j) + WC(i, j) + X(i, j) \quad (2.49)$$

onde

$Z(i, j)$ - a medida do sistema de carga em horas (MWH) na hora j do dia i .

$T(i, j)$ - componente de nível da carga na hora j do dia i .

$WC(i, j)$ - componente de ciclo semanal (efeito do dia da semana) na hora j do dia i .

$X(i, j)$ - componente residual contendo o efeito da variação do tempo na hora j do dia i .

É necessário ressaltar que a componente de nível varia muito pouco de um dia para o outro, podendo até, em certos casos, se apresentar constante. A componente de ciclo semanal também apresenta mudança lenta, evidenciando o padrão semanal das cargas horárias. Já a componente residual muda rapidamente, pois, a mesma, contém a variação da carga, hora a hora, motivada por uma série de fatores, tal como condições de tempo experimentadas. Um modelo adequado para modelar essa componente aleatória, seria um do tipo auto-regressivo

$$X(i) = AX(i-1) + W(i-1) \quad (2.50)$$

onde

$$X(i) = \begin{bmatrix} X(i,1) \\ X(i,2) \\ \vdots \\ X(i,24) \end{bmatrix} - (24 \times 1) \text{ vetores coluna das componentes residuais para o } i\text{-ésimo dia}$$

A - matriz (24×24) dos coeficientes

$W(i-1)$ - vetor coluna dos erros do modelo.

Uma importante característica da equação do modelo (2.50) é que cada elemento de $X(i)$ é uma função linear de todos os elementos de $X(i-1)$. Os erros são assumidos como tendo média zero e como sendo estatisticamente independentes. Ou seja, o erro de previsão

de um dado dia, não possui nenhuma relação com o erro do dia anterior a ele. A matriz Q de covariância (24×24) é desconhecida é definida de modo que :

$$E(W(i)W(i)^T) = Q \quad (2.51)$$

onde $W(i)^T$ denota a transposta de $W(i)$. Uma forma equivalente da equação 4.3 pode ser escrita para o (j, k) -ésimo elemento de Q como

$$Q(j, k) = E(W(i, j)W(i, k)) \quad (2.52)$$

onde $W(i, j)$ é o erro do modelo para hora j do i -ésimo dia.

2.2.5.1

Componente de Carga-clima

O procedimento de previsão, tal como mencionado anteriormente, requer um modelo de carga-clima, onde o pico de carga diário seja representada em termos de variáveis climáticas. O modelo tem a seguinte forma

$$Y(i) = B(i) + S(i) + W(i) + \xi(i) \quad (2.53)$$

onde

$Y(i)$ - pico de carga no i -ésimo dia

$B(i)$ - componente de nível de carga do pico de carga no i -ésimo dia.

$S(i)$ - componente de padrão semanal do pico de carga no i -ésimo dia.

$W(i)$ - componente de tempo do pico de carga no i -ésimo dia.

$\xi(i)$ - componente randômica do pico de carga no i -ésimo dia.

É preciso ser dito que $B(i)$ e $S(i)$ devem ser definidos como elementos de $T(i, j)$ e $WC(i, j)$. No entanto, melhores estimativas podem ser esperadas quando $B(i)$ e $S(i)$ são estimados independentemente dessas duas variáveis. A componente de tempo, $W(i)$, é assumida como sendo uma função linear dos valores transformados das variáveis meteorológicas, tal como temperatura e velocidade do vento. Com isso temos que a forma geral dessa componente pode ser descrita por

$$W(i) = \sum_{j=1}^k C_j WV_j(i) \quad (2.54)$$

onde

k - número de variáveis meteorológicas no modelo.

$WV_j(i)$ - j -ésima variável meteorológica (ou um valor transformado da variável) no i -ésimo dia.

C_j - j -ésimo coeficiente desconhecido do modelo.

Quanto ao padrão semanal, sabemos que existirão, sete valores, cada um deles estando relacionado a um dia da semana. Logo, conclui-se que

$$S(i) = \sum_{j=1}^7 S_j P_j(i) \quad (2.55)$$

onde

S_j - j -ésimo valor do padrão semanal ($j=1$ correspondendo ao Domingo).

$P_j(i) = \begin{cases} 1 \\ 0 \end{cases}$ - se o i -ésimo dia corresponde ao j -ésimo dia da semana.

A identificação das matrizes A e Q é baseada na componente residual $X(\cdot)$ dos dados históricos de carga. Dado que a componente residual não se encontra disponível de forma direta, convém que, primeiramente, se estime os valores de $T(i, j)$ e $WC(i, j)$ de um dado de carga. Cada um dos elementos de $T(i, j)$ é obtido pela filtragem do padrão semanal e da componente residual dos dados de carga :

$$T(i, j) = \frac{1}{7} \sum_{k=i-6}^i Z(i, j) \quad \begin{array}{l} j = 1, 2, \dots, 24 \\ i = 7, \dots, N \end{array} \quad (2.56)$$

onde N é o número de dias para qual os dados de carga se encontram disponíveis.

A componente de padrão semanal, $WC(i, j)$, é obtida por um filtro de amortecimento exponencial :

$$WC(i, j) = WC(i - 7, j) + \alpha \{Z(i, j) - T(i, j) - WC(i - 7, j)\} \quad \begin{array}{l} j = 1, 2, \dots, 24 \\ i = 8, 9, \dots, N \end{array} \quad (2.57)$$

onde α é uma constante de filtro. De forma geral, convém determinar essa constante de filtro com valores compreendidos entre 0.2 e 0.5.

$$WC(k, j) = Z(k, j) - T(7, j) \quad k = 1, 2, \dots, 7. \quad (2.58)$$

Os parâmetros desconhecidos do modelo tempo-carga são identificados, minimizando a soma exponencialmente ponderada dos quadrados dos erros de previsão. Ou seja,

$$E = \sum_{n=1}^N \left[Y(n) - B(n) - \sum_{i=1}^7 S_i P_i(n) - \sum_{i=1}^k C_i W V_i(n) \right]^2 \beta^{N-n} \quad (2.59)$$

A previsão de carga para o $(i+1)$ -ésimo dia, $\hat{Z}(i+1, j)$, $j=1, 2, \dots, 24$, é feita nas bases de medidas de $Z(i, j)$, $j=1, 2, \dots, 24$, e da previsão de tempo $W\tilde{V}_i(i+1)$, $j=1, 2, \dots, k$. O procedimento utilizado para a atualização das componentes $T(i, j)$ e $WC(i, j)$ já foi mencionada acima. A componente residual $X(i)$ é determinada da seguinte forma :

$$X(i, j) = Z(i, j) - T(i, j) - WC(i, j) \quad k = 1, 2, \dots, 24. \quad (2.60)$$

As matrizes A e Q são atualizadas pela re-estimação das matrizes de covariância Γ_0 e Γ_1 .

$$\Gamma_0^i = \Gamma_0^{i-1} + \frac{1}{i} [X(i)X^T(i) - \Gamma_0^{i-1}]. \quad (2.61)$$

$$\Gamma_1^i = \Gamma_1^{i-1} + \frac{1}{i-1} [X(i)X^T(i-1) - \Gamma_1^{i-1}]. \quad (2.62)$$

onde Γ_0^i e Γ_1^i são estimativas das matrizes de covariância Γ_0 e Γ_1 obtidas pelo processamento de dados do i -ésimo dia. O algoritmo acima é baseado somente nas prévias estimativas e no novo produto dos termos $X(i)X^T(i)$ e $X(i)X^T(i-1)$.

Para maiores detalhes de formulação teórica ver Bunn (1985[3]), onde este modelo é descrito em detalhes. Esta formulação foi usada por Sobral (1999[23]), em sua dissertação de mestrado. Na dissertação é desenvolvido um modelo de previsão de carga horária de curto prazo, fazendo uso de informações climáticas, além das informações de carga. O modelo proposto por Gupta é combinado com as metodologias de redes neurais e lógica nebulosa. A autora identifica os curvas típicas de consumo de cargas, e inclui elas no modelo de Gupta, como uma forma alternativa de se identificar esses perfis. A forma proposta por Gupta, é a de se classificar a curva de consumo por dia da semana. Os resultados obtidos com essa adaptação do modelo, foram bastante satisfatórios, obtendo erros de previsão relativamente baixos. Também na tese de doutorado de Quadrelli (1998[20]), o modelo proposto por Gupta (1985[3]), é utilizado para previsões horárias e

diárias, e seu desempenho é comparado com o de um modelo de amortecimento direto, modelado por senos e cosenos (funções trigonométricas).

2.2.6

Redes Neurais

As Redes Neurais são sistemas inspirados na estrutura de funcionamento do cérebro e dos neurônios biológicos. Esse interesse foi principalmente motivado pela observação da facilidade e eficácia com o que o cérebro realiza tarefas difíceis e complexas. As redes neurais resolvem problemas onde é difícil criar modelos adequados a realidade ou, então, situações que mudam muito (problemas não lineares), sem a necessidade de se definir regras ou modelos explícitos.

Devido a similaridade de uma rede neural com a estrutura de um cérebro, elas, também, acabam por exibir características semelhantes, tais como :

- Aprendizado : aprende-se por experiência;
- Associação : faz associações entre padrões diferentes;
- Generalização : são capazes de generalizar o conhecimento adquirido a partir de experiências passadas.
- Abstração : extrai a essência de um conjunto de informações, retirando os ruídos.

Um bom exemplo da topologia de rede neural encontra-se na figura 2.2, que, por convenção, é um formato de rede bastante usado, a Rede *Feedforward*. Nela, podem existir uma ou mais camadas de processamento. Na figura mencionada, a rede possui 3 camadas : a camada de entrada, a camada escondida, e a camada de saída. As rede neural possui também unidades de processamento da informação. Essas unidades são denominadas neurônios, e são conectadas por pesos sinápticos. Vale a pena mencionar que as redes *feedforward* não possuem realimentação.

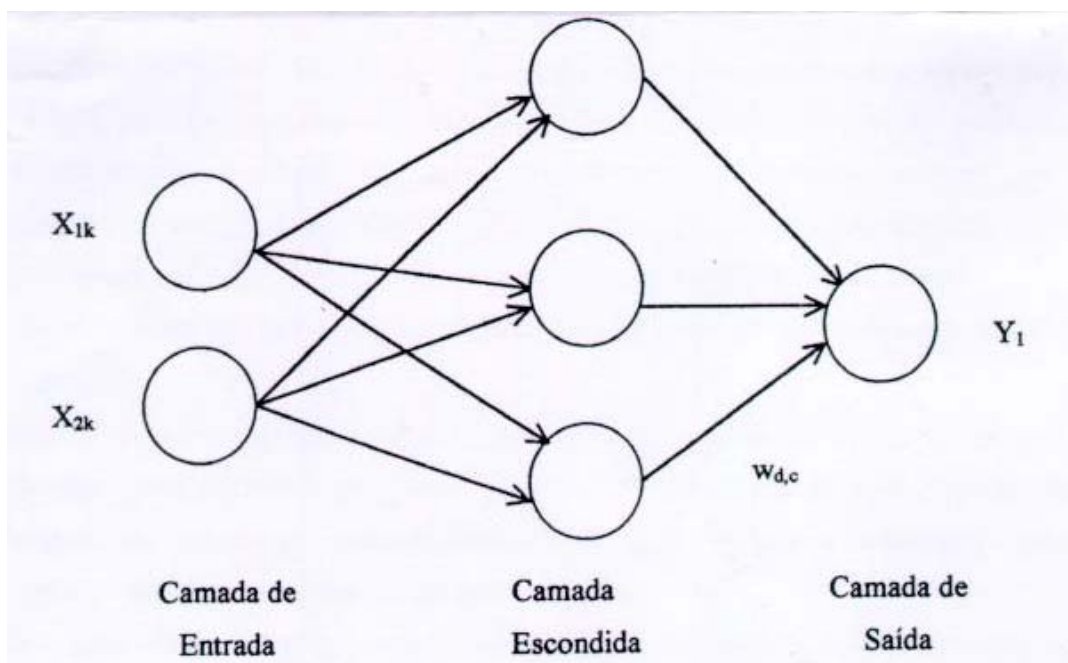


Figura 2.2 – Rede *Feedforward*

Existe também um outro tipo de rede, conhecida como rede recorrente. Nesse tipo de topologia existe uma conexão entre os processadores não só de uma mesma camada como também de camadas diferentes, existindo realimentação na rede.

As redes neurais tem várias fases até que se obtenha o resultado final. A fase de treinamento, onde é retirado o conhecimento do ambiente. A fase de generalização, onde o conhecimento adquirido na fase anterior, é testado, para verificar se o que foi aprendido pode ser utilizado para o fim desejado. No entanto, existem, dentre várias, diversas formas muito importantes de uma rede aprender. O aprendizado pode acontecer de forma supervisionada ou de forma não supervisionada. No primeiro, trabalha-se com conjuntos de pares de entrada e saída, ambos previamente conhecidos e representantes da realidade. Já no aprendizado não supervisionado, não se trabalha com conjuntos previamente conhecidos. Estabelece-se uma medida que represente a qualidade da representação da rede, e os parâmetros são modificados de forma a otimizá-la.

Em uma etapa do estudo realizado por Sobral (1999[23]), a autora utiliza a metodologia das redes neurais para identificar os grupos de cargas típicas, agrupando, então

as cargas de acordo com seus padrões de comportamento, rompendo a rigidez da classificação via estações do ano. Nesse caso, as redes neurais estão sendo usadas como um classificador de padrões, agrupando os dados de acordo com padrões de comportamentos semelhantes. Essa seria uma forma, segundo Sobral, de se evitar o surgimento de sazonalidades múltiplas. No estudo, a autora faz uso do mapa auto-organizável de Kohonen para efetuar tal tarefa. Somente após a classificação das cargas típicas, foi utilizado o modelo de previsão proposto por Gupta. Os resultados obtidos para utilização dessa ferramenta foram bastante satisfatórios (Sobral, 1999[23]).

Em um outro estudo, Hippert (2001[11]) testa a viabilidade do uso da metodologia de redes neurais aplicada na previsão de perfis de carga. O objetivo do estudo é verificar se o uso dessa técnica é mais vantajoso do que o uso da modelagem tradicional. Esse estudo é feito baseado na idéia de que apesar da existência de diversas vantagens, as redes neurais ainda não conseguiram se mostraram convincentes. As vantagens das redes neurais no que diz respeito a previsão de perfis de carga, diz respeito a possibilidade da construção de um modelo com saídas multivariadas, possibilitando a previsão de vários pontos de um perfil de forma simultânea. Através do estudos de 39 artigos sobre o tema, o autor chegou a conclusão que, essa falta de credibilidade das redes neurais, ocorre devido a escolha de modelos super-ajustados e a hiper-parâmetrização dos modelos, e pela falha dos autores no momento da validação dos mesmos, (falta uma análise comparativa com os modelos de referência) (Hippert, 2001[11]). No estudo, Hippert examina a possibilidade da ocorrência do super-ajuste do modelo, e também o processo de validação. Os resultados obtidos mostraram que as simulações feitas utilizando redes neurais obtiveram previsões de perfis melhores que a feita através do métodos lineares usuais, e os modelos se mostraram como não sendo super-ajustados.

No entanto, é preciso ser enfatizado, em se tratando do uso de rede neurais para o tratamento de problemas lineares, é fácil ser verificado, que a maioria dos métodos estatísticos tradicionais, obtém resultados mais eficientes. As redes neurais, tal como a lógica nebulosa, somente oferece resultados mais adequados, em casos onde o investigador

lida com a busca de solução para problemas não lineares, e quando se tem como objetivo a modelagem de relações complicadas.

2.2.7

Lógica Nebulosa

Os modelos matemáticos, estatísticos e de previsão são formulados seguindo o conhecimento objetivo. No entanto, existe também o conhecimento subjetivo, aquele adquirido através de informações lingüísticas, e, por esta razão, muito difíceis de ser quantificado.

O sistema de lógica nebulosa tenta unir a precisão dos modelos matemáticos com a imprecisão do mundo real. Essa formulação surgiu da motivação de se conhecer mais profundamente o funcionamento do raciocínio humano.

Os sistemas de lógica nebulosa, as redes neurais, e os algoritmos genéticos foram criados a partir dessa motivação. Esse sistema tem a capacidade de capturar informações imprecisas, descritas em linguagem natural, e convertê-las para o formato numérico. Ele mapea vetores numéricos de entrada em valores numéricos de saída, como pode ser observado no esquema abaixo.

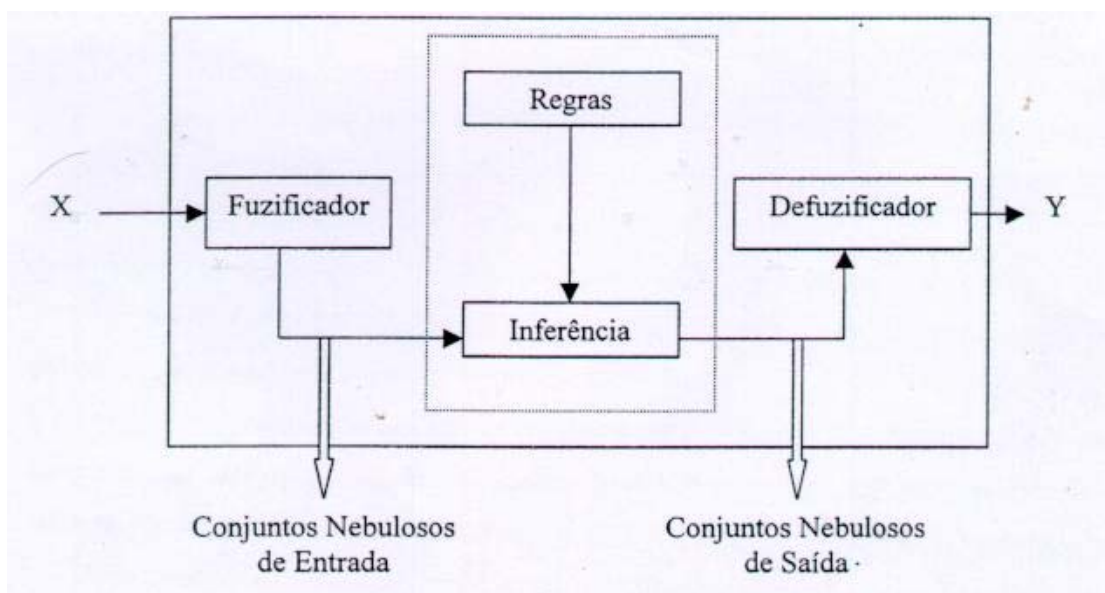


Figura 2.3 – Sistema de Lógica Nebulosa

Mas como funciona um sistema de lógica nebulosa? Primeiramente, entra-se com os valores precisos (dados numéricos) no sistema. Esses dados passam por uma ferramenta que tem a função de transformar esses valores precisos em variáveis lingüísticas e criar regras que são associadas a conjuntos nebulosos. Os conjuntos nebulosos são funções que mapeam um valor escalar em um número entre 0 e 1, que indique o grau de pertinência a um determinado conjunto. Seria mais ou menos um valor que representasse o quanto esse valor escalar pertence a um determinado conjunto. Ou seja, o grau de pertinência é o nível de compatibilidade de um elemento com o conceito de um conjunto.

A ferramenta utilizada para a transformação de desses valores se chama fuzificador. Existem, na verdade, duas formas de se criar regras para um sistema nebuloso .

- conhecimento extraído a partir dos dados;
- conhecimento fornecido por um especialista da área.

Depois de extraídas e definidas as regras, é determinado como essas regras serão combinadas e as que serão ativadas e/ou não. Isso é feito por um sistema de inferência, gerando conjuntos nebulosos de saída que passarão por um processo inverso ao de

fuzificação, sendo as variáveis lingüísticas transformadas em variáveis numéricas pelo desfuzificador.

No caso da aplicação da lógica nebulosa para a previsão de séries temporais, as regras são estabelecidas pela série histórica analisada, estabelecendo os conjuntos nebulosos do antecedente (entradas), e do conseqüente (saídas). Outra forma de estabelecer essas regras seria especificando previamente os conjuntos e associando os dados a esses conjuntos predefinidos. Se trabalha com janelas, no caso da previsão como pode ser observado na figura 2.5. Cada janela é formada pela quantidade de dados capazes de explicar o valor a ser previsto.

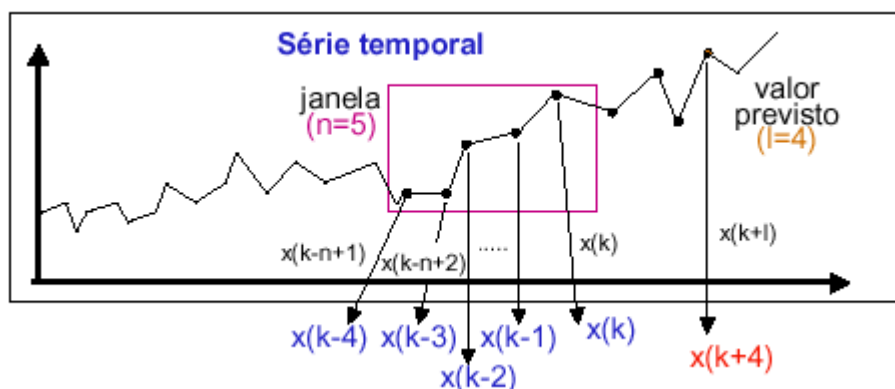


Figura 2.4 – Previsão de Séries Temporais na Lógica Nebulosa

Supondo que x seja uma série temporal de tamanho k , sendo $k=1,2,\dots$. São criadas janelas. Mas como se determina o tamanho da janela? O tamanho de uma janela é determinado, na maioria das vezes, pelo próprio especialista, ele determina quantos dados são suficientes para explicar o dado a ser previsto. Então, dado uma janela com n medidas, de $x(k)$, variando de $x(k-n+1)$, $x(k-n+2)$, \dots , $x(k)$, determina-se o valor de $x(k+1)$. Nessa etapa são estipulados o horizonte de previsão e os graus de pertinência dos elementos de X . A cada uma das variáveis é atribuído o conjunto com maior grau, obtendo, por conseguinte, uma regra para cada par de entrada e saída.

Sobral (1999[23]) utilizou a lógica nebulosa para poder fazer uso das possibilidades de se trabalhar com a relação entre carga e temperatura. No estudo, a autora cria atributos meteorológicos e não meteorológicos capazes de identificar o grau de pertinência do dia a ser previsto em cada um dos grupos de carga que foram encontrados utilizando as redes neurais, como mencionado no item anterior. Ela utiliza esses atributos são utilizados ao invés do modelo carga-clima proposto por Gupta. Os resultados obtidos após a finalização de todos os procedimentos descritos nos itens 2.2.5 e 2.2.6 foram bastante interessantes, ocorrendo uma melhora nas previsões. Também foi feita, a parte, a mesma modelagem para os dias onde tiveram jogos da Copa do Mundo. Essa é somente, um exemplo das diversas formas onde podemos fazer uso dos conceitos e métodos provenientes da inteligência computacional dentro da área de previsão de carga de energia elétrica. No sistema PREVCAR, a lógica nebulosa é utilizada de forma combinada com outros modelos, de Redes Neurais, Box&Jenkins e Holt-Winters, como já mencionado anteriormente, com o objetivo de se obter, previsões com o menor erro possível. Também existem, no sistema 3 algoritmos, que tem a função de combinar os dois melhores métodos, desagregar as previsões obtidas para o primeiro mês em semanas elétricas e dias, e desagregar todas as previsões das séries de energia, em patamares de carga (leve, média e pesada). Existe a intenção de se incluir ainda mais um 4 algoritmo, que tratar as possíveis mudanças nos modelos e nos seus parâmetros, também abre a possibilidade do tratamento das discontinuidades similares que possam vir a ocorrer.