

## CAPÍTULO 6

### ESTIMAÇÃO DE PARÂMETROS

#### 6.1 Introdução

A estimação de parâmetros procura determinar os futuros parâmetros de um processo qualquer a partir de dados observados ao longo de um determinado período de tempo. O início da técnica de estimação de parâmetros foi com Kolmogorov (1941) em 1941 e Norbert Wiener (1949) em 1942. O objetivo era prever a posição de aviões inimigos para poder derrubá-los. A partir disto, se desenvolveram uma série de métodos como: Processos auto-regressivos, mínimos quadrados, filtro de Kalman, etc.

O método que usaremos para fazer nossa previsão dos mapas globais ionosférico do conteúdo total de elétrons (TEC), é o método de *Maximum Entropy Method* (MEM) ou Método de Máxima Entropia (Kantor e Trivedi., 1977e 1980), baseado em um processo auto-regressivo (AR), o qual tem uma série de aplicações em espectro de potência, espectro cruzado, cálculo da função auto-correlação, previsão linear, filtragem de dados, transmissão de sinais e cancelamento de ruído.

Um estudo completo do MEM envolve campos diversos como entropia, teoria linear de previsão de Wiener, matrizes positivas definidas, processos auto-regressivos, matrizes Toeplitz, espaços de Hilbert e transformada Z, além de um conhecimento de processos estocásticos.

Este método se apresenta como uma ferramenta útil para poder estimar parâmetros que têm muita variação, como é nosso caso em que os coeficientes não seguem um padrão determinado.

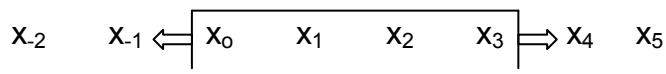
## 6.2 Processos Auto-Regresivos (AR)

Um processo auto-regressivo é baseado na informação dos eventos anteriores, que proporcionam dados do processo, que são armazenados e processados para obter informação do próximo evento.

Vamos supor um navio deslocando-se no mar. Sua posição em relação a uma direção  $x$  em intervalos fixos de tempo, gera uma série de tempo  $(x_0, x_1, x_2, x_3, x_4, \dots)$ . Se a velocidade for constante é muito fácil determinar a próxima posição, a partir de dois valores anteriores. Se o navio tiver uma aceleração constante serão necessários três valores para fazer a predição de sua posição, indicada por:

$$\hat{x}_3 = 3x_2 - 3x_1 + x_0 \quad (6.1)$$

Mesmo que a velocidade ou aceleração não sejam constantes, a posição pode ser prevista aproximadamente usando os 3 pontos anteriores. Neste caso estaríamos usando um filtro preditor com três pontos ou coeficientes. Este número de coeficientes na realidade pode ser qualquer, ou seja, o ajuste dependerá de quanta informação anterior disponhamos, qual será a dimensão da amostra que queremos processar para obter o próximo ponto e da previsibilidade da serie. Este processo também pode ser usado para determinar um ponto anterior, como é apresentado no diagrama abaixo.



Outra maneira de entender um processo auto-regressivo é através de uma expansão em série de Taylor no ponto  $x$  ( $t = 0$ ) de uma função  $x(t)$ , onde pode-se usar valores anteriores de  $t = -1, -2, -3, \dots$  para estimar as derivadas em  $t = 0$ . A primeira derivada seria dada pela diferença  $x_0 - x_{-1}$ . A segunda derivada vai depender dos três últimos pontos, a terceira derivada dos quatro últimos pontos, etc. A expansão em série de Taylor para calcular o valor  $x_1$ , daria a estimativa de  $x_1$  em termos de uma função

linear dos pontos anteriores. Se a série for truncada em N-1 termos teremos (Kantor e Trivedi, 1977):

$$\hat{x}_1 = \sum_{k=0}^N \frac{x^{(k)}(0)}{k!} = \sum_{k=0}^N g_k x_{-k} \quad (6.2)$$

Com isto tentamos mostrar que para prever cada ponto futuro como uma combinação linear dos pontos anteriores é o mesmo que expandir a função em série de Taylor, truncada na n-ésima derivada, correspondendo à ordem do filtro preditor.

Qualquer processo estacionário, no qual a média é constante e, em particular nula ( $E x_j = 0$ ) e a auto-correlação ( $E x_j x_k = \phi(j-k)$ ), é estacionário (Yule, 1927). A seqüência  $x_t$  pode ser expressada por:

$$x_t = \sum_{k=1}^{\infty} g_k x_{t-k} + \varepsilon_k \quad (6.3)$$

onde  $\varepsilon_k$  é o resíduo. Isto implica que toda série resultante de um processo estacionário, pode ser decomposta numa parte previsível ou determinística e numa parte aleatória.

Para prever nosso ponto previsto  $\hat{x}_t$ , em função linear dos M pontos do passado, devemos otimizar os coeficientes  $g_k$ . Vamos estabelecer que a esperança do quadrado da diferença entre o predito e o real, ou erro quadrático médio da predição seja mínimo:

$$E^2 = E(x_t - \hat{x}_t)^2 = E(x_t x_t) - 2 \sum g_k E(x_t x_{t-k}) + \sum \sum g_k g_l E(x_{t-k} x_{t-l}) \quad (6.4)$$

$\Downarrow$   
 $\phi(0)$

$\Downarrow$   
 $\phi(k)$

$\Downarrow$   
 $\phi(k-l)$

O mínimo pode ser encontrado igualando a zero as derivadas parciais em relação a cada coeficiente,  $g_k$ , ou seja:

$$\partial E^2 / \partial g_m = 0 \quad (m=1,2,3,\dots,M) \quad (6.5)$$

Encontramos assim um sistema de equações chamado de Wiener- Hopf:

$$\sum_{k=1}^M g_k \phi(k-m) = \phi(m) \quad (m = 1, 2, \dots, M) \quad (6.6)$$

Conhecendo a auto-correlação para diversos “lags” podemos resolver o sistema e obter os coeficientes  $g_k$ .

Conhecidos os  $g_k$ , que minimizam o erro, podemos calcular este erro mínimo quadrático, que chamaremos de potência do erro de previsão,  $P_M$ . Substituindo o sistema de Wiener- Hopf na expressão do erro temos:

$$P_M = E^2_{\text{MIN}} = \phi(0) - \sum_{k=1}^M g_k \phi(k) \quad (6.7)$$

Para uma visão melhor das equações usaremos vetores e matrizes, pois simplifica a notação, permitindo escrever as equações de forma mais compacta. Construindo o vetor “filtro do erro de predição”, de dimensão  $M+1$

$$\Gamma_M = \begin{pmatrix} \Gamma_0 \\ \Gamma_1 \\ \vdots \\ \Gamma_M \end{pmatrix} \quad \text{onde} \quad \begin{cases} \Gamma_0 = 1 \\ \Gamma_k = -g_k \quad k = 1, 2, \dots, M \end{cases} \quad (6.8)$$

a matriz de auto-correlação é composta dos  $\phi(k)$ , e é simétrica em relação à diagonal:

$$\Phi_M = \begin{pmatrix} \phi(0) & \phi(1) & \dots & \phi(M) \\ \phi(1) & \phi(0) & \dots & \phi(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \phi(M) & \phi(M-1) & \dots & \phi(0) \end{pmatrix} \quad (6.9)$$

As matrizes deste tipo são chamadas matrizes Toeplitz. Finalmente temos o “vetor potência do erro da predição”, dado por:

$$\mathbf{P}_M = \begin{pmatrix} P_{0M} \\ 0 \\ \vdots \\ \vdots \end{pmatrix} \quad (6.10)$$

Substituindo os novos coeficientes  $\Gamma_k$ , no sistema de Wiener-Hopf na fórmula de  $P_M$ , obtemos:

$$\sum_{k=0}^M \Gamma_k \phi(k-m) = 0$$

$$\sum_{k=0}^M \Gamma_k \phi(k) = P_M \quad (6.11)$$

O sistema pode ser escrito na forma matricial

$$\begin{pmatrix} \phi(0) & \phi(1) & \dots & \phi(M) \\ \phi(1) & \phi(0) & \dots & \phi(M-1) \\ \vdots & \vdots & \cdot & \vdots \\ \vdots & \vdots & \cdot & \vdots \\ \phi(M) & \phi(M-1) & \dots & \phi(0) \end{pmatrix} \begin{pmatrix} \Gamma_0 \\ \Gamma_1 \\ \vdots \\ \vdots \\ \Gamma_M \end{pmatrix} = \begin{pmatrix} P_{0M} \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix} \quad (6.12)$$

o qual pode ser expresso de forma compacta:

$$\Phi_M \Gamma_M = P_M \quad (6.13)$$

Isto permite encontrar os coeficientes do erro de previsão e o valor da potência do erro de predição, sendo conhecidas as auto-correlações. O cálculo das equações cuja matriz é Toeplitz, pode ser bastante simplificado usando o método de recursão de Levinson. Este

método utiliza uma previsão para frente e para atrás, calculando sucessivamente  $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_{M+1}$  sem a necessidade da inversão da matriz  $\Phi_M$ . A potência do erro  $P_M$  vai decrescer a cada ordem sucessiva.

### 6.3 Cálculo dos Coeficientes GIM

Para calcular nosso GIM, usamos a representação do TEC global proposta por Schaer et al., (1995). Esta alternativa é uma expansão em harmônicos esféricos (Equação 4.4). Os coeficiente  $\check{C}_{nm}$  e  $\check{S}_{nm}$  são proporcionados pelo CODE (Figura 6.1) para cada dia, e são válidas para um período de duas horas, gerando um total de 12 por dia.

Para calcular a expressão 4.4, nos baseamos na subrotina proporcionada pelo CODE, Comunicação pessoal (Schaer, 2002) e programas do *Numerical Recipes in Fortran*, (Press et al., 1992), com um número de grau igual a 15, ou seja que trabalhamos com 256 coeficientes para gerar a matriz 71 x 73 (latitude, longitude) que representa todo o mundo. O programa Kantor-Dulanto (PKD) foi desenvolvido em *Interactive Data Language* (IDL 5.3) pela facilidade de recursos que ele proporciona. O programa completo é apresentado no anexo B.

A matriz gerada 71 x 73 é uma representação global do TEC em 0,1TECU ( $10^{15}$  elétrons/m<sup>2</sup>) desde uma latitude de 87.5°N até 87.5°S, de 2.5° em 2.5° e desde uma longitude 180°W até 180°E, de 5° em 5°.

Na Figura 6.1 apresentamos um trecho do arquivo de coeficientes obtidos da CODE pela Internet para o dia 15 de março de 2002, para o primeiro período compreendido entre 00:00 a 02:00 UT. Por dia são gerados 12 destes arquivos e são armazenados juntos, no mesmo formato.

Em nossa predição usamos o método de máxima entropia, tomando como base de dados 30 dias para cada período específico de duas horas (a influência horária é marcante), pois é um período que vai compreender 1 dia solar (27 dias) e dois ciclos lunares. Este

pode ser o menor lapso do tempo em que esperamos ter um comportamento com alguma tendência, pois o próximo período seria analisar 1 ano ou 11 anos, que equivale ao período de rotação da Terra e a um ciclo solar.

O número de coeficientes do filtro que usamos para a predição foi 15, que representa à metade de nossa base de dados e foi a mesma escolha que faz Schaer (1997), que também considerou a influência que exerce um ciclo lunar 14,77 dias. Isto nos permite ter suficiente informação inicial (anterior) para poder estimar os próximos valores dos coeficientes,

Nas Figuras 6.2, 6.3, 6.4 e 6.5 são apresentados os coeficiente  $C_{00}$ ,  $C_{11}$ ,  $S_{22}$  (Para diferenciar no arquivo este coeficiente do  $C_{22}$  é representado com número de ordem negativo) e  $C_{15,15}$  do CODE em diferentes horários para os dias 14 de fevereiro a 15 de março e a predição destes coeficientes para o 16 a 21 de Março de 2002. A linha contínua representa o valor real processado e os asteriscos representam nossa predição considerando o dia 15 de março (último dia da base de dados) para verificar se partem do mesmo ponto.

Os valores previstos dos coeficientes  $\tilde{C}_{nm}$  e  $\tilde{S}_{nm}$  para os próximos seis dias, ou seja os 256 coeficientes que geram o GIM serão analisados e comparados com outros mapas centros internacionais do IGS para o mesmo dia e o mesmo período do dia (cada período é de duas horas) no próximo Capítulo.

CODE'S GLOBAL IONOSPHERE MAPS FOR DAY 074, 2002  
22:47

18-MAR-02

```
-----  
MODEL NUMBER / STATION NAME : 0740-01  
MODEL TYPE (1=LOCAL,2=GLOBAL,3=STATION) : 2  
MAXIMUM DEGREE OF SPHERICAL HARMONICS : 15  
MAXIMUM ORDER : 15  
DEVELOPMENT WITH RESPECT TO  
  GEOGRAPHICAL (=1) OR GEOMAGNETIC (=2) FRAME : 2  
  MEAN (=1) OR TRUE (=2) POSITION OF THE SUN : 1  
MAPPING FUNCTION (0=NONE,1=COSZ,2=MSLM,3=ESM) : 2  
HEIGHT OF SINGLE LAYER AND ITS RMS ERROR (KM) : 450.00 0.00  
COORDINATES OF EARTH-CENTERED DIPOLE AXIS  
  LATITUDE OF NORTH GEOMAGNETIC POLE (DEGREES) : 79.55  
  EAST LONGITUDE (DEGREES) : -71.92  
PERIOD OF VALIDITY  
  FROM EPOCH / REFERENCE EPOCH (Y,M,D,H,M,S) : 2002 03 15 00 00 00  
  TO EPOCH : 2002 03 15 02 00 00  
LATITUDE BAND COVERED  
  MINIMUM LATITUDE (DEGREES) : -88.66  
  MAXIMUM LATITUDE (DEGREES) : 86.65  
ADDITIONAL INFORMATION  
  NUMBER OF CONTRIBUTING STATIONS : 153  
  NUMBER OF CONTRIBUTING SATELLITES : 27  
  ELEVATION CUT-OFF ANGLE (DEGREES) : 10  
  MAXIMUM TEC AND ITS RMS ERROR (TECU) : 144.91 0.91  
COMMENT / WARNING :  
COEFFICIENTS  
DEGREE ORDER VALUE (TECU) RMS (TECU)  
  0 0 54.67370497 0.0831  
  1 0 1.71368352 0.0747  
  1 1 17.33254065 0.0745  
  1 -1 17.25376914 0.0730  
  2 0 -14.44517686 0.0775  
  2 1 -1.62258123 0.0668  
  : : : :  
  : : : :  
 15 15 0.21273672 0.0492  
 15 -15 0.11673741 0.0494
```

CODE'S GLOBAL IONOSPHERE MAPS FOR DAY 074, 2002  
22:47

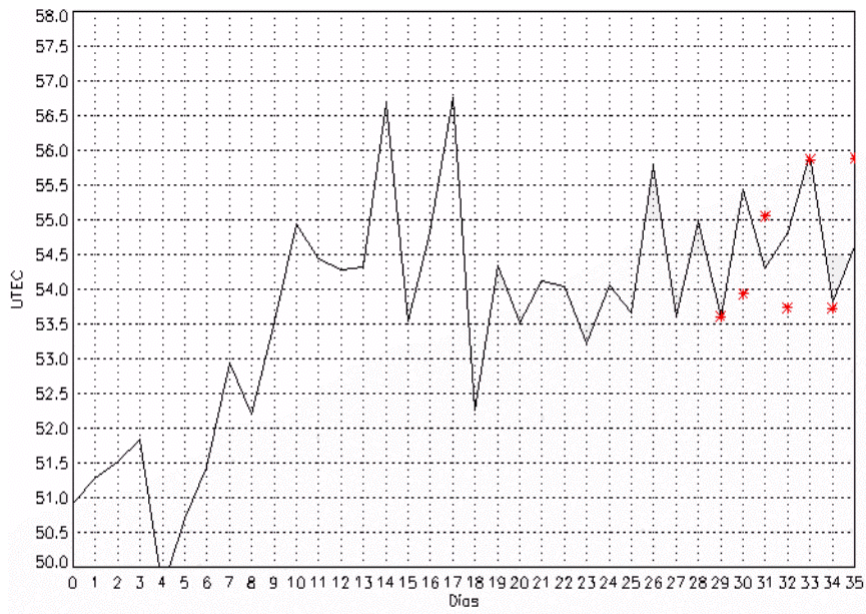
18-MAR-02

```
-----  
MODEL NUMBER / STATION NAME : 0740-02  
MODEL TYPE (1=LOCAL,2=GLOBAL,3=STATION) : 2  
MAXIMUM DEGREE OF SPHERICAL HARMONICS : 15  
MAXIMUM ORDER : 15
```

FIGURA 6.1 - Partes do arquivo de coeficientes (15-03-2002) do CODE.

FONTE: CODE (2002).

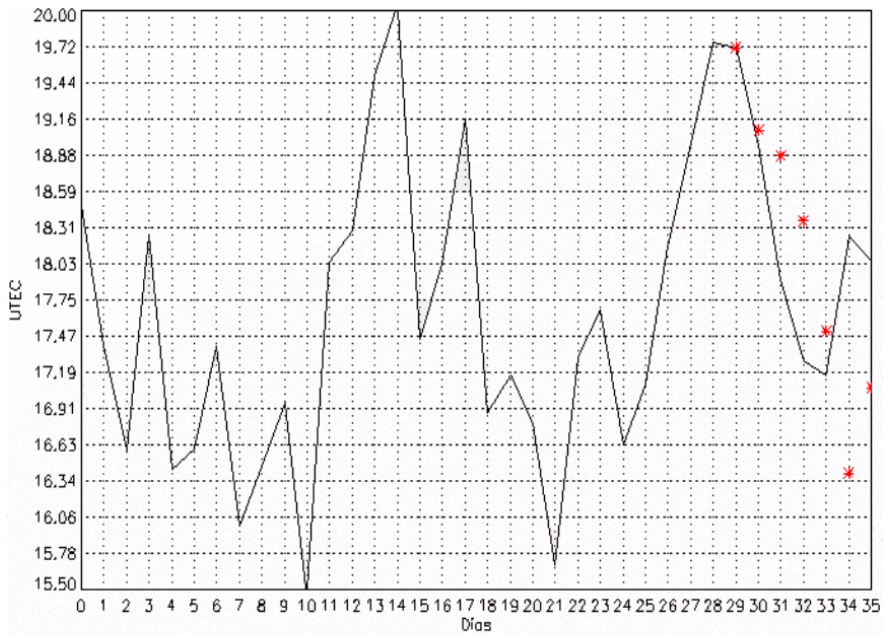




Coeficientes processados pela CODE \_\_\_\_\_ Coeficientes previstos PKD \*\*\*

FIGURA 6.2 - Predição do coeficiente  $C_{00}$  (16-03 ao 21-03 de 2002) para 09:00UT.

FONTE: Kantor e Dulanto (2002).



Coeficientes processados pela CODE \_\_\_\_\_ Coeficientes previstos PKD \*\*\*

FIGURA 6.3 - Predição do coeficiente  $C_{11}$  (16-03 ao 21-03 de 2002) para 15:00 UT.

FONTE: Kantor e Dulanto (2002).

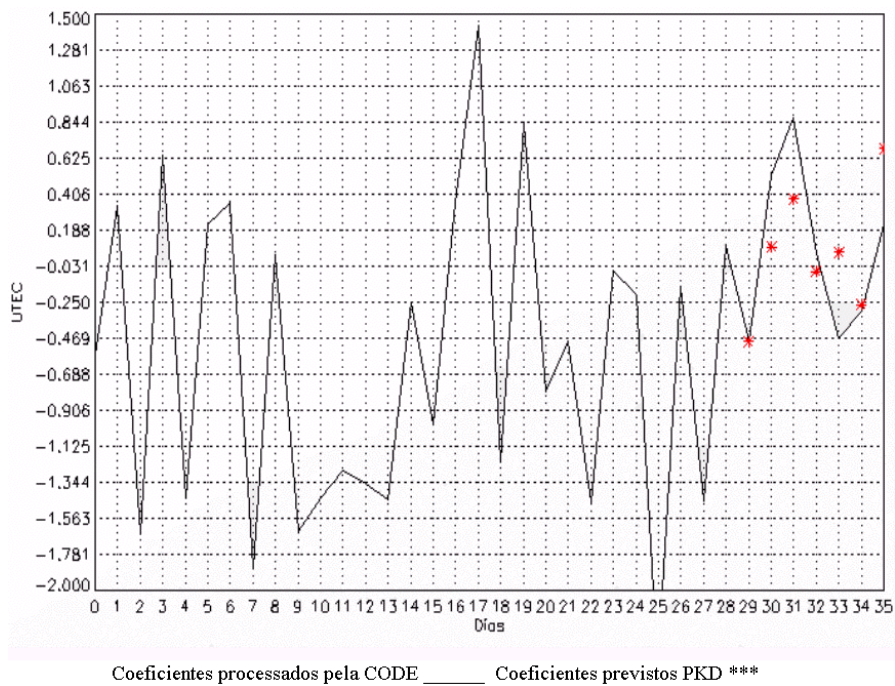


FIGURA 6.4 - Predição do coeficiente  $S_{22}$  (16-03 ao 21-03 de 2002) para 23:00 UT.

FONTE: Kantor e Dulanto (2002).

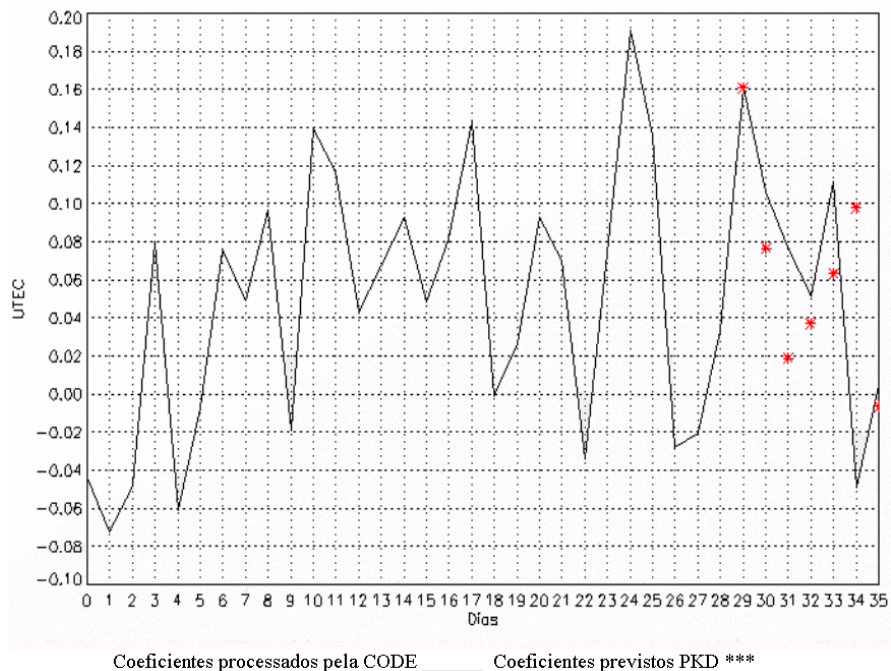


FIGURA 6.5 - Predição do coeficiente  $C_{1515}$  (16-03 ao 21-03 de 2002) para 17:00 UT.

FONTE: Kantor e Dulanto (2002).