MODELOS EM SÉRIES TEMPORAIS

I Séries Temporais

A descrição de sinais que dependem do tempo é em geral realizada na forma discreta no tempo. Supõe-se que os sinais são amostrados em intervalos regulares. A estes sinais dá-se o nome também de séries temporais. A modelagem de tais sinais corresponde a uma atividade importante da identificação de sistemas.

II Modelos de Ruído

A hipótese de independência do ruído é um hipótese que em geral não se verifica na prática, para séries temporais. Uma maneira de modelar a correlação do ruído entre dois pontos (ou mais) i e i-1 é através de um modelo de correlação. Os modelos mais comuns são:

II.1 Modelo Auto-Regressivo

para ambos os modelos que serão apresentados aqui, supõe-se a existência de uma "semente" normal (média zero e variância σ^2) e independente, dito ruído branco. Notaremos ε este ruído branco. As propriedades do ruído branco são tais que para o mesmo instante:

$$\sigma_{ii}^2 = E\left(\varepsilon_i^2\right)$$

e para dois instantes diferentes (duas amostras diferentes):

$$\sigma_{ij}^2 = E(\varepsilon_i \varepsilon_j) = 0$$

Um modelo de erro auto-regressivo (AR) é dado por, por exemplo:

$$\eta_j = -a_1 \eta_{j-1} + \varepsilon_j$$

Desta forma o ruído η é auto-correlacionado no tempo. Estes modelos podem ser dados de maneira mais genérica utilizando-se a notação com polinômios do operador atraso q^{-1}. A aplicação do operador atraso a uma determinada variável é definida por:

$$q^{-1} y_j = y_{j-1}$$

De maneira mais genérica define-se um polinômio mônico A(q):

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \ldots + a_n q^{-n}$$

e portanto um ruído auto-correlacionado segundo um modelo AR pode ser expresso por:

$$A(q)\,\eta = \varepsilon$$

II.1 Modelo de Média Móvel

Este tipo de modelo denominado comumente de MA (moving-average) pode ser dado por exemplo por:

$$\eta_j = \varepsilon_j + c_1 \varepsilon_{j-1}$$

ou de maneira mais genérica por:

$$\eta_j = C(q)\,\varepsilon_j$$

onde C(q) é um polinômio mônico.

Pode-se ainda definir modelos ARMA que combinam ambos os tipos de auto-correlação, genericamente através de:

$$A(q)\,\eta_j = C(q)\,\varepsilon_j$$

III Modelos Determinísticos e Estocásticos

Estes modelos permitem modelar a relação entre uma variável deterministíca, u(t), chamada de variável exógena ou variável controlada, e uma variável estocástica, y(t), que tem uma contribuição, tanto da variável exógena quanto de um ruído branco.

O modelo mais simples é o chamado modelo ARX (combinação de um modelo AR, com uma contribuição exógena). Neste modelo, os sinais y(t) e u(t) são relacionados através da seguinte relação matemática:

$$y(t) + a_1 y(t-1) + ... + a_{na}\, y(t-na) = b_1 u(t-n_r) + ... + b_{nb}\, u(t-nb-nr+1) + \varepsilon(t)$$

Este modelo pode ser colocado em notação polinomial:

$$A(q)\,y(t) = B(q)\,u(t) + \varepsilon(t)$$

Os parâmetros deste modelo são a_1, a_2, …, a_{na}, b_1, b_2, …, b_{nb}.

Outro modelo deste tipo é o chamado ARMAX, que inclui além da parcela AR, da variável exógena (X), a contribuição de um ruído MA. Os sinais y(t) e u(t) são relacionados através de:

$$y(t) + a_1 y(t-1) + ... + a_{na}\, y(t-na) = b_1 u(t-n_r) + ... + b_{nb}\, u(t-nb-nr+1) + \varepsilon(t) + c_1\varepsilon(t-1) + ... + c_{nc}\varepsilon(t-nc)$$

Que pode ser notado através de polinômios no operador atraso:

$$A(q)\,y(t) = B(q)\,u(t) + C(q)\,\varepsilon(t)$$

Os parâmetros deste modelo são a_1, a_2, …, a_{na}, b_1, b_2, …, b_{nb}, c_1, c_2, …, c_{nc}

Nesta linha, podem ser geradas as seguintes estruturas:

ARARMAX:

$$A(q)\,y(t) = B(q)\,u(t) + \frac{C(q)}{D(q)}\varepsilon(t)$$

Output Error:

$$y(t) = \frac{B(q)}{F(q)}\,u(t) + \varepsilon(t)$$

Box-Jenkins:

$$y(t) = \frac{B(q)}{F(q)}\,u(t) + \frac{C(q)}{D(q)}\varepsilon(t)$$

V Metodologia de Ajuste

A metodologia mais comum para o ajuste de modelos em séries temporais é a metodologia do erro de predição. Ela é baseada na análise de uma série de dados no tempo:

[y(n) y(n-1) y(n-2) … y(1)]

[u(n) u(n-1) u(n-2) … u(1)]

Para o modelo ARX é proposto o seguinte preditor :

$$\hat{y}(t,\theta) = -a_1 y(t-1) - \ldots - a_{na}\, y(t-na) + b_1 u(t-n_r) + \ldots + b_{nb}\, u(t-nb-nr+1)$$

Neste caso:

$$y(t) - \hat{y}(t,\theta)$$

é um resíduo que deve ser minimizado para obter a estimativa de θ. O vetor θ é composto pelos parâmetros:

$$\theta = (-a_1, -a_2, ..., -a_{na}, b_1, b_2, ..., b_{nb})^T$$

O ajuste deste modelo gera um problema linear em relação aos parâmetros.

No caso de um modelo ARMAX, o preditor é não linear em relação aos parâmetros:

$$\hat{y}(t,\theta) = -a_1 y(t-1) - ... - a_{na} y(t-na) + b_1 u(t-n_r) + ... + b_{nb} u(t-nb-nr+1)$$
$$+ c_1 (y(t-1) - \hat{y}(t-1,\theta)) + c_2 (y(t-2) - \hat{y}(t-2,\theta)) + ... + c_{nc} (y(t-nc) - \hat{y}(t-nc,\theta))$$

V Independência do resíduo

A independência do resíduo é uma das hipóteses mais importantes na formulação do critério de estimação para séries temporais (ε(t) é suposto ruído branco).

As duas rotas mais usuais são a análise da função de auto-correlação e os testes não parâmétricos. O cálculo da função de autocorrelação (Walter e Lecourtier, 1997) envolve o cálculo da estimativa da funçao de autocorrelação:

$$\hat{c}_\varepsilon(k) = \frac{1}{n-k} \sum_{i=1}^{n-k} (\varepsilon_i - \bar{\varepsilon})(\varepsilon_{i+k} - \bar{\varepsilon}) \qquad (17)$$

A função de auto-correlação reflete a correlação da variável, no tempo, e por isso depende do parâmetro k, que é o número de intervalos de tempo, ou lag. A autocorrelação normalizada é estimada por:

$$\hat{\rho}(k) = \frac{\hat{c}_\varepsilon(k)}{\hat{c}_\varepsilon(0)} \qquad (18)$$

Fazendo-se inúmeras hipóteses pode se mostrar que $\hat{\rho}(k)$ tem distribuição normal quando n é grande com média -1/(n-1) e variância 1/n. Um teste de hipóteses aproximado de que $\hat{\rho}(k)$ é nulo pode ser feito, com um tamanho de 5% (95% de significância) tomando-se a região crítica:

$$\left]-\infty, -2/\sqrt{n}\right] \cup \left[2/\sqrt{n}, +\infty\right[\qquad (19)$$

Segundo Walter e Lecourtier (1997), esta região crítica pode ser corrigida ainda por:

$$\left]-\infty, -2\sqrt{n}/\sqrt{(n+2)(n-k)}\right] \cup \left[2\sqrt{n}/\sqrt{(n+2)(n-k)}, +\infty\right[\qquad (20)$$

Os testes não paramétricos são baseados na descrição qualitativa dos dados através de sequências binárias, que podem ser testadas frente a hipóteses. Um deste métodos é o de corridas acima e abaixo da mediana. Para tal define-se uma função auxiliar, u_i:

$$u_i = \begin{cases} 1 \text{ se } \varepsilon_i \succ \nu \\ 0 \text{ se } \varepsilon_i \prec \nu \end{cases}$$

onde ν é a mediana da amostra de ε_i. Se ε_i for igual a ν é só usar o bom senso (por exemplo, sortear com uma moeda se o valor de u_i associado é 1 ou 0). A partir disto gera-se uma sequência de u_i´s:

u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}	...
0	1	1	1	0	0	1	0	1	0	0	

a partir do quê definem-se as corridas como uma sequência em que há repetição do mesmo número, desta forma na sequência acima pode-se separar as seguintes corridas:

0 | 1 1 1 | 0 0 | 1 | 0 | 1 | 0 0 ...

ou seja que neste conjunto de 11 pontos temos 7 corridas. O teste de hipóteses serve para decidir se este resultado deve ou não ser aceito como um resultado de uma variável aleatória independente. Para dados p (número de resultados 1) e n, o número de corridas, p, segue uma distribuição aproximadamente normal com média:

$$\frac{2p(n-p)}{(n+1)}$$

com variância:

$$\frac{2p(n-p)[2p(n-p)-n]}{n^2(n-1)}$$

Siegel (1956) tabelou valores para a região crítica de r (número de corridas) para diferentes números de n e p (Mandansky, 1988, p. 109). Outra opção de teste não paramétrico é o teste de corridas para cima e para baixo. Nesta caso calcula-se uma sequência de variáveis auxiliar:

$$v_i = \begin{cases} 1 \text{ se } \varepsilon_{i+1} \succ \varepsilon_i \\ 0 \text{ se } \varepsilon_{i+1} \prec \varepsilon_i \end{cases}$$

Mostra-se que a média para o número de corridas é:

$$\frac{2n-1}{3}$$

e a variância:

$$\frac{16n-29}{90}$$

sendo que para n muito grande a distribuição se aproxima de uma distribuição normal.

Um último teste de independência é o chamado "rank von Neumann ratio". Para este teste associa-se a cada ε_i o seu rango, isto é, a sua posição quando os resíduos são ordenados do menor ao maior. O rango pode variar de 1 a n. Cria-se um série de rangos, r_1, r_2, r_3, ..., r_n. E calcula-se a estatística:

$$v = \frac{\sum_{i=1}^{n} (r_i - r_{i-1})^2}{n(n^2 - 1)/12} \qquad (21)$$

chamada de razão de rank von Newmann. Valores críticos para $c = [n\ (n^2-1)/12]\ v$ são tabelados em Mandansky (1988, p. 117).

Bilbliografia

Ljung, L., System Identification, Prentice-Hall, 1999

Mandansky, A., Prescriptions for Working Statisticians, Springer Texts in Statistics, 1988

Walter, E., Pronzato, L., Identification of Parametric Models from Experimental Data, Springer, 1997