

Introdução ao Processamento Estatístico de Sinais (V1.0)

Muitos sinais reais nem são periódicos, nem têm energia finita, e por isso não possuem transformadas Z ou de Fourier. Outros sinais têm processos de gênese tão complexos que é impossível descrevê-los adequadamente por modelos determinísticos. Em suma, muitos sinais apresentam uma componente *estatística*, *aleatória* ou *estocástica* associada à incerteza dos valores que podem tomar.

A chave para a sua modelação matemática é a descrição estatística em termos de “médias”. Adicionalmente, em muitos casos, as suas propriedades mais importantes podem ser reveladas por sequências de energia finita, obtidas a partir do sinal (ou sinais) original, denominadas *correlação* e *covariância* que, frequentemente, possuem transformadas Z ou de Fourier finitas. Estas sequências relacionam-se com o espectro de frequências do sinal.

O uso da covariância e da sua transformada de Fourier tem a vantagem de proporcionar uma descrição matemática concisa do processamento de um sinal de energia infinita por um sistema linear discreto.

Em suma, os mencionados sinais de energia infinita são estudados como **processos estocásticos** (PEs) e caracterizados por um conjunto de *funções de densidade* ou *funções de distribuição de probabilidade* localizadas (no mesmo PE e/ou instante de tempo) ou cruzadas (entre PEs e/ou instantes de tempo distintos).

Objectivos da aula prática

Nesta aula pretende-se exercitar alguns comandos do Scilab que permitem estudar sinais (ou processos) estocásticos. Inclui-se a estimação espectral e a modelação de processos estocásticos. Incide-se em técnicas de modelação ligadas ao modelo de Box-Jenkins de ordem p , denominado $AR(p)$. O Scilab possui funções dedicadas a estas tarefas; na parte final deste documento são apresentadas algumas.

Correlação e covariância

Para processos estocásticos genéricos¹, $\{x_n\}$ e $\{y_n\}$, a **correlação** (e estimativa) é definida por

$$\phi_{XY}[n, m] = E\{x_n y_m\} \approx \frac{1}{N} \sum_{k=0}^{N-1} x^{(k)}[n] y^{(k)}[m]$$

(onde (k) indexa as realizações do PE) e a **covariância** (e estimativa) é definida por

$$\gamma_{XY}[n, m] = E\{(x_n - m_X)(y_m - m_Y)\}$$

¹Daqui em diante, as sequências correspondem a uma realização de um processo estocástico, pelo que “sequência” e “processo estocástico” serão, de certo modo, sinónimos.

$$\approx \frac{1}{N} \sum_{k=0}^{N-1} (x^{(k)}[n] - \bar{x}_n)(y^{(k)}[m] - \bar{y}_m)$$

onde $\bar{x}_n = N^{-1} \sum_{k=0}^{N-1} x^{(k)}[n]$ e $\bar{y}_m = N^{-1} \sum_{k=0}^{N-1} y^{(k)}[m]$ são as **médias** pontuais (estimadas). Quando $\{x_n\}$ é o mesmo que $\{y_n\}$, i.e. para um mesmo PE, aquelas quantidades são denominadas *auto-correlação* e *auto-covariância*, respectivamente. Quando as médias \bar{x}_n e \bar{y}_m são ambas zero, a correlação e a covariância são idênticas. Note que para PEs gerais ϕ e γ são sequências bidimensionais.

Para processos estocásticos estacionários (onde as estatísticas pontuais não variam com n e as estatísticas cruzadas só dependem da separação m) e ergódicos (onde as estatísticas no tempo são iguais às estatísticas nas realizações), com as² sequências $x[n]$ e $y[n]$ centradas em $n = 0$ e de dimensão finita $2K + 1$, teremos as seguintes **aproximações** da média, da correlação e da covariância dessa sequência (a estacionariedade garante a invariância com o tempo n)

$$m_X \approx \frac{1}{2K + 1} \sum_{n=-K}^K x[n] \quad (1)$$

$$\phi_{XY}[m] \approx \frac{1}{2K + 1} \sum_{n=-K}^K x[n] y^*[n + m] \quad (2)$$

$$\gamma_{XY}[m] \approx \frac{1}{2K + 1} \sum_{n=-K}^K (x[n] - m_X)(y^*[n + m] - m_Y) \quad (3)$$

onde $\phi_{xy}[m]$ é a estimativa da correlação e $\gamma_{xy}[m]$ é a estimativa da covariância, para $m = 0, 1, 2, \dots$.

Assim, com a estacionariedade a correlação e a covariância tornaram-se sequências unidimensionais.

Note que nos anteriores somatórios não houve a preocupação de garantir que $n + m$ não extravasa as dimensões do vector de amostras, $y[n]$; na aplicação real das aproximações esse aspecto não pode ser descurado (veja a eq. (8)). As anteriores expressões prevêm o tratamento de processos estocásticos no espaço complexo, ao conjugar (com “*”) o factor $y^*[n]$.

As equações (1), (2) e (3) oferecem-nos *estimativas*. A obtenção de estimativas da média, variância e da covariância é um problema de estatística.

Estimativas da estatística de processos estocásticos

Em geral, um processo estocástico $\{x_n\}$ (neste caso, uma sequência discreta corresponderá a uma possível realização desse PE) tem um número infinito de possíveis realizações. Se for **estacionário**, as propriedades estatísticas das variáveis aleatórias x_n (a variável aleatória (VA) associada ao instante n) não variam com n . Se for **ergódico**, essas propriedades estatísticas podem ser inferidas (ou estimadas) a partir de uma *amostra finita* retirada de *uma única realização* do processo.

²A adaptação das expressões (1,2,3) para serem aplicadas a sequências indexadas com $n = 0, 1, 2, \dots, N - 1$ é trivial.

Assim, num PE ergódico as estimativas da estatística da VA x_n não necessitam de mais do que uma realização do processo.

As estimativas que têm a maior probabilidade de “estar certas” são denominadas, em Estatística, *estimativas de máxima verosimilhança*. No caso da média e da variância de processos aleatórios Gaussianos, elas são:

$$\hat{m}_x = \frac{1}{N} \sum_{n=0}^{N-1} x_n \quad \hat{\sigma}_x^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x_n - \hat{m}_x)^2$$

e são estimativas *consistentes*, pois convergem para as verdadeiras média e variância quando $N \rightarrow \infty$.

Considere agora um processo estocástico estacionário real $\{x_n\}$ que, por conveniência, supomos ter média nula ($m_x = 0$), o que implica auto-covariância igual à auto-correlação. Um **estimador da auto-covariância** é (para $|m| < N$)

$$c'_{xx}[m] = \frac{1}{N - |m|} \sum_{n=0}^{N-|m|-1} x[n] x[n+m] \quad (4)$$

Esta estimativa é de máxima verosimilhança se a sequência $x[n] x[n+m]$ for Gaussiana.

Um outro estimador alternativo é

$$c_{xx}[m] = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x[n] x[n+m] \quad (5)$$

Cada um destes estimadores apresenta problemas relativamente à variância das estimativas quando $|m| \rightarrow N$, pois próximo deste limite há poucos pontos para calcular a média de $x[n] x[n+m]$. Na prática, é sugerido o uso de $c_{xx}[m]$ que “atenua” a ponderação dos valores afastados. De facto, este é o estimador utilizado na função `corr()` do Scilab. Veja a eq. (8).

O espectro de potência e o periodograma

Se $\Gamma_{xx}(z)$ for a transformada Z da auto-covariância $\gamma_{xx}[m]$ de $\{x_n\}$, o **espectro de potência** (ou somente **espectro**), $P_{xx}(\omega)$, de $\{x_n\}$ é definido por

$$P_{xx}(\omega) = \Gamma_{xx}(e^{j\omega})$$

ou seja, o **espectro é a resposta na frequência da auto-covariância**. Na prática usa-se a sua estimativa calculada a partir da sequência $x[n]$ retirada de uma realização do PE $\{x_n\}$. Infelizmente, as transformadas de Fourier dos estimadores da covariância nas eqs. (4) e (5) não são estimadores consistentes, pois a variância do espectro assim estimado não tende para zero quando N aumenta. No entanto, se for feito o alisamento (*smoothing*) das transformadas obtêm-se boas estimativas do espectro de potência.

Seja $x[n]$ uma sequência real finita, com transformada de Fourier $X(e^{j\omega}) = \sum_{n=0}^{N-1} x[n] e^{-j\omega n}$, e seja

$$I_N(\omega) = \sum_{m=-(N-1)}^{N-1} c_{xx}[m] e^{-j\omega m}$$

a transformada de Fourier do respectivo estimador $c_{xx}[m]$. Ao estimador espectral $I_N(\omega)$ chama-se o **periodograma** e pode mostrar-se que

$$I_N(\omega) = \frac{1}{N} |X(e^{j\omega})|^2$$

Este estimador do espectro não é consistente (diz-se que é *enviesado*). Por sua vez, a transformada de Fourier de $c'_{xx}[m]$, $P_N(\omega) = \sum_{m=-(N-1)}^{N-1} c'_{xx}[m] e^{-j\omega m}$, também é um estimador enviesado do espectro. A razão, em ambos os casos, deve-se a que a variância do periodograma aumenta com N e, assim, as oscilações no espectro aumentam à medida que mais termos de $c'_{xx}[m]$ (com $|m| \rightarrow N$) são usados na Tr. de Fourier. Como já tínhamos avisado, $I_N(\omega)$ e $P_N(\omega)$ não são bons estimadores do espectro.

Estimadores espectrais por alisamento

A redução da variância dos estimadores pode ser feita de várias formas.

Uma delas, atribuída a **Bartlett**, consiste em fatiar a sequência $x[n]$, de dimensão N , em K segmentos de dimensão M (i.e., $N = KM$), criando assim K sub-sequências $x^{(i)}[n]$ de dimensão M , calcular de seguida os correspondentes K periodogramas, $I_M^{(i)}(\omega)$, e obter um estimador espectral pela média destes:

$$B_{xx}(\omega) = \frac{1}{K} \sum_{i=1}^K I_M^{(i)}(\omega)$$

Uma outra abordagem consiste em convolver o periodograma com uma janela espectral $W(e^{j\omega})$ apropriada para conseguir o alisamento. O periodograma alisado, $S_{xx}(\omega)$, será dado por

$$S_{xx}(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} I_N(\theta) W(e^{j(\omega-\theta)}) d\theta$$

Sendo dada a janela no tempo $w[m] = (1/2\pi) \int_{-\pi}^{\pi} W(e^{j\omega}) e^{j\omega m} d\omega$, com largura $2M - 1$, ter-se-á (invocando a dualidade convolução na frequência/produto no tempo):

$$S_{xx}(\omega) = \sum_{m=-(M-1)}^{M-1} c_{xx}[m] w[m] e^{-j\omega m}$$

As janelas $w[m]$ utilizadas com esta técnica são habitualmente a triangular, de Bartlett, de Hamming ou de Hanning, mas estas duas últimas podem conduzir a estimativas negativas (obviamente erradas) no espectro de potência.

Para terminar, acrescenta-se apenas que a FFT é utilizada para calcular muitos destes estimadores, permitindo assim diminuir o tempo do seu cálculo, como já acontecia no caso da convolução e dos filtros FIR.

Trabalho prático

Após ter lido esta introdução, onde se terá apercebido de que o problema da estimação espectral não é trivial, experimente várias estimativas espectrais para diversos sinais e com comandos do Scilab apropriados (veja a sua lista no final.) Concretamente, compare as transformadas de Fourier dos sinais, efectuadas com a FFT, com a transformada de Fourier da autocovariância e com os comandos-estimadores do Scilab. Use $N < 1000$, dependendo dos casos, some sinais determinísticos (sinusóides, ondas quadradas, etc...) com ruído Gaussiano ou uniforme, e observe os respectivos espectros.

Modelação Box-Jenkins

Muitos modelos práticos de sistemas ou sinais estocásticos usam a metodologia introduzida por G. Box e G. Jenkins (BJ) para modelar as séries temporais correspondentes às respectivas realizações.

O modelo BJ mais popular é o modelo $ARMA(p,q)$, com p pólos e q zeros, com $p + q + 1$ parâmetros (os a_k e os b_k) e cuja função de transferência $H(z)$ é:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{B(z)}{A(z)} = \frac{\sum_{k=0}^q b_k z^{-k}}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (6)$$

$ARMA$ é a sigla de **Auto Regressive Moving Average**. Introduzindo na entrada de $H(z)$ **ruído branco Gaussiano**, $x[n] = w[n] = \mathcal{W}_n$, e tendo em conta as propriedades do processo estocástico Y_n gerado na saída, é possível modelar muitos fenómenos e sistemas estocásticos.

O modelo no tempo do tipo $ARMA(p,q)$, equivalente à eq. (6), é a equação às diferenças com ruído branco, $w[n]$, na entrada e saída $y[n]$

$$y[n] = - \sum_{k=1}^p a_k y[n-k] + \sum_{k=0}^q b_k w[n-k] + \epsilon_n \quad (7)$$

O termo ϵ_n é um resíduo, ou “ruído do modelo”.

Uma especialização do anterior modelo é o modelo **Auto-Regressivo** de ordem p , $AR(p)$, que tem as seguintes representações no tempo e na frequência:

$$y[n] = - \sum_{k=1}^p a_k y[n-k] + b_0 w[n] + \epsilon_n$$

$$H_{AR}(z) = \frac{Y(z)}{W(z)} = \frac{b_0}{A(z)} = \frac{b_0}{1 + \sum_{k=1}^p a_k z^{-k}}$$

O modelo de **Média Móvel** $MA(q)$ tem q zeros “interessantes” e q pólos na origem, independentemente do valor dos b_k (logo é *sempre estável*) e tem as seguintes representações no tempo e na frequência:

$$y[n] = \sum_{k=0}^q b_k w[n-k] + \epsilon_n$$

$$H_{MA}(z) = \frac{Y(z)}{W(z)} = B(z) = \sum_{k=0}^q b_k z^{-k}$$

Por comparação com a definição da resposta impulsiva de um SLIT, $h[n]$, é imediato concluir que

$$h[n] = \{b_k\}$$

A resposta impulsiva do modelo $AR(p)$ é, em geral, de dimensão infinita. Por isso, um modelo $AR(p)$ pode ser sempre substituído por um modelo $MA(\infty)$.

Na aplicação prática da modelação BJ, o caso mais comum corresponde ao modelo auto-regressivo $AR(p)$. Em geral, a ordem p necessária para modelar uma dada série é bastante menor que a ordem q , associada ao modelo $MA(q)$, se este fosse o escolhido para a modelação.

Nesta aula prática iremos apostar precisamente nos modelos $AR(p)$ pois estes, embora menos gerais do que os modelos $ARMA(p,q)$, permitem ilustrar integralmente a metodologia deste tipo de modelação.

Equações de Yule-Walker no modelo $AR(p)$

O processo mais utilizado para calcular os parâmetros a_k do modelo $AR(p)$ que melhor descrevem uma determinada série temporal é a resolução das equações de Yule-Walker³ (YW) que usam a estimativa da autocovariância/autocorrelação calculada para a série. No caso do modelo $AR(p)$, e presumindo que $b_0 = h[0]$ é uma dada constante, as equações de YW são:

$$\begin{bmatrix} r_y[0] & r_y[-1] & \cdots & \cdots & r_y[-p] \\ r_y[1] & r_y[0] & r_y[-1] & \cdots & \vdots \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ r_y[p] & r_y[p-1] & \cdots & \cdots & r_y[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma_x^2 b_0^2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

onde $r_y[k]$ designa o estimador $c_{yy}[k]$. Quando a série é real, a covariância é uma sequência par, i.e., $r_y[k] = r_y[-k]$, $\forall k$.

Se removermos a primeira equação do sistema, o termo $\sigma_x^2 b_0^2$ desaparece (de facto, é somente uma constante multiplicativa) e torna-se claro que a partir de um vector de covariâncias de dimensão $p+1$ se obtêm os p coeficientes a_k do modelo $AR(p)$. O sistema reduzido será:

$$\begin{bmatrix} r_y[0] & r_y[-1] & \cdots & r_y[-p+1] \\ r_y[1] & r_y[0] & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ r_y[p-1] & r_y[p-2] & \cdots & r_y[0] \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} -r_y[1] \\ -r_y[2] \\ \vdots \\ -r_y[p] \end{bmatrix}$$

ou, atendendo à simetria da covariância:

$$\begin{bmatrix} r_y[0] & r_y[1] & \cdots & r_y[p-1] \\ r_y[1] & r_y[0] & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ r_y[p-1] & r_y[p-2] & \cdots & r_y[0] \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} -r_y[1] \\ -r_y[2] \\ \vdots \\ -r_y[p] \end{bmatrix}$$

³Podem usar-se também as equações de Yule-Walker com modelos mais gerais $ARMA(p,q)$.

Esta é uma *matriz de Toeplitz* e, por esta razão, o sistema pode ser resolvido eficientemente pelo *algoritmo de Levinson* ou outros aparentados.

“Detrending” (remoção de tendências)

Muitas vezes, em aplicações práticas, os processos estocásticos não são estacionários. Um exemplo comum está ligado a fenômenos sazonais (e.g. a temperatura do ar, a quantidade de vendas de gelados, etc...) ou, por exemplo, no caso das vendas, a associação com a intensidade das campanhas de publicidade. No entanto, a modelação BJ assume ruído branco na entrada do sistema, um processo estocástico estacionário.

Para mitigar o problema de forma a ser possível recorrer à metodologia BJ mesmo em casos não estacionários, aplica-se *detrending* à série temporal, ou seja, retira-se-lhe a “tendência”. Isto corresponde, grosso modo, a remover a sua componente determinística. Se este passo for bem executado, o processo estocástico ficará (quase) estacionário, pelo menos no que respeita à média e, assim, em condições de poder ser correctamente modelado por um processo $AR(p)$ (ou por outro modelo ARMA).

O Scilab tem a função `detrend()` para remover tendências dos tipos linear ou linear por troços. Outras tendências, cuja forma geral seja conhecida (e. g., sinusoidais ou exponenciais), poderão ser removidas por aproximação (*fitting*) usando técnicas de mínimos quadrados, por exemplo; a função `leastsq()` pode ser usada nestes casos.

Trabalho prático

Crie processos estocásticos estacionários Gaussianos e uniformes para aplicar nas entradas dos modelos. Crie outros com tendências (e.g. do tipo linear, exponencial ou sinusoidal) adicionadas. Especifique alguns modelos $AR(p)$, com um valor razoável de p (digamos 3 ou 4, para não exagerar, e tenha cuidado com a localização dos pólos para obter modelos estáveis) e use-os para filtrar os processos de entrada. Observe as respostas na frequência, i.e. os espectros, dos vários processos que vai criando. Usando a função `lev()` do Scilab estime os parâmetros dos filtros que criou e avalie a justeza dos resultados. Faça experiências com e sem *detrending*, verificando que se não aplicar este passo de processamento prévio obtém resultados errados (i.e., não calcula os coeficientes correctos dos filtros). Resolva explicitamente as equações de Yule-Walker para um exemplo e confirme os resultados obtidos com `lev()`.

Funções Scilab relacionadas com estimação espectral e modelação ARMA(p,q)

O Scilab dispõe de funções dedicadas à estimação espectral e à modelação Box-Jenkins. Apresenta-se um breve sumário das mais importantes, deixando os detalhes da sua utilização para a consulta do *help* da ferramenta. Recorde

que a função `corr()` na realidade *calcula a covariância* e não a correlação.

- `[cov,Mean]=corr(x,[y],nlags)` or `[cov,Mean]=corr('fft',xmacro,[ymacro],n,sect)`: computes (or better, estimates) the covariance between $x[n]$ and $y[n]$, given by

$$\phi_{xy}[m] = \frac{1}{N} \sum_{k=1}^{N-m} \{x[k] - \bar{x}\} \{y[m+k] - \bar{y}\} \quad (8)$$

for $m = 0, 1, \dots, nlags - 1$ and two vectors with size N (starting at $n = 1$, as usual). The short form, `[cov,Mean]=corr(x,[y],nlags)`, returns the first $nlags$ correlation coefficients and `Mean = mean(x)` (mean of $[x,y]$ if y is an argument). The sequence x (resp. y) is assumed real, and x and y are of the same dimension N .

- `[sm,cwp]=pspect(sec_step,sec_leng,wtype,x,y,wpar)`: cross-spectral estimate between x and y if both are given, and auto-spectral estimate of x otherwise. The spectral estimate is obtained using the modified periodogram method. Some of these parameters are optional.
- `[sm,cwp]=cspect(nlags,ntp,wtype,x,y,wpar)`: spectral estimation using the correlation method. Cross-spectral estimate of x and y is calculated when both x and y are given. Auto-spectral estimate of x is calculated if y is not given.
- `ar=lev(r)`: solve the Yule-Walker equations using Levinson’s algorithm. r is the vector with correlation/covariance coefficients, and ar is the vector with the $AR(p)$ model parameters. The size of r is $N_r = p + 1$, that is, giving a certain length, N_r , of the covariance vector, we get an $AR(N_r - 1)$ model in ar .
- `y=detrend(x)`: this function removes the constant, linear, or piecewise linear trend from a vector x . In general this can be useful before Fourier analysis. If x is a matrix, this function removes the trend columnwise.
- `[fopt,[xopt],[grdopt]]=leastsq(fun,x0)` or `[fopt,[xopt],[grdopt]]=leastsq(fun,dfun,x0)`: solves non-linear least squares problems. Tries to minimize $f(x) = ||\text{fun}(x)||^2 = \sum_i \text{fun}_i^2(x)$ which is the sum of the squares of the components of `fun()`. Bound constraints may be imposed on x . The user must define `fun(x)`.