

Estimación en procesos estocásticos “temporales”

(o sea, series temporales, procesos unidimensionales)

Antonio Sala Piqueras

Identificación de sistemas complejos

Dept. Ing. Sistemas y Automática (DISA)

Universitat Politècnica de València (UPV)

Video-presentación disponible en:

<http://personales.upv.es/asala/YT/V/krigt.html>



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Presentación

Motivación:

En muchas aplicaciones se tienen muestras de una señal en ciertos instantes y se desea estimar su valor en otros (interpolación, extrapolación).

Objetivos:

Comprender la aplicación de las fórmulas estadísticas de mejor predicción lineal a partir de matriz de varianzas-covarianzas a una serie temporal.

Contenidos:

Revisión de predicción lineal. Procesos gaussianos en el tiempo. Construcción de matrices de varianzas-covarianzas. Ecuación de predicción. Caso estacionario, interpretación frecuencial. Conclusiones.



Predicción lineal estadística: preliminares

Supongamos conocida la matriz de varianzas-covarianzas entre unos datos " $a_{m \times 1}$ " (a predecir) y " $b_{r \times 1}$ " (observados, información disponible):

$$\Sigma_{(m+r) \times (m+r)} = \begin{pmatrix} \Sigma_{a, m \times m} & \Sigma_{ab, m \times r} \\ \Sigma_{ba, r \times m} & \Sigma_{b, r \times r} \end{pmatrix}$$

La mejor predicción lineal (mínima vza. error pred.) de unos datos de media cero " $a_{m \times 1}$ " a partir de otros " $b_{r \times 1}$ ", es

$$\hat{a} = \underbrace{\Sigma_{ab, m \times r} \cdot \Sigma_{b, r \times r}^{-1}}_{\Xi_{m \times r}} \cdot b_{r \times 1}$$



Procesos estocásticos temporales

Proceso gaussiano: Supongamos señal generada por un proceso estocástico (serie temporal), **conjunto infinito de variables aleatorias** $y(t) \sim N(0, \sigma^2)$.

- **Modelo estadístico:** Supongamos que $\text{cov}(y(t_1), y(t_2)) = \kappa(t_1, t_2)$, siendo $\kappa : \mathbb{R}^2 \mapsto \mathbb{R}$ una función de covarianza conocida.

Obviamente, la varianza de $y(t)$ sería $\kappa(t, t)$ y la correlación entre $y(t_1)$ e $y(t_2)$ será $\rho_{y(t_1), y(t_2)} = \frac{\kappa(t_1, t_2)}{\sqrt{\kappa(t_1, t_1)\kappa(t_2, t_2)}}$

- **Observaciones:** Supongamos que disponemos de muestras medidas del valor de una realización del proceso en un conjunto de puntos $T := \{t_1, t_2, \dots, t_N\}$, esto es $y_i := y(t_i)$, $i = 1, \dots, N$. Denotemos $y(T) := (y_1, \dots, y_N)^T$.



Predicción: construcción matriz de varianzas-covarianzas

- 1 La matriz VC de la información disponible (y_i) es

$$\Sigma_{y(T)} = \kappa(T, T) := \begin{pmatrix} \kappa(t_1, t_1) & \kappa(t_1, t_2) & \dots & \kappa(t_1, t_N) \\ \kappa(t_2, t_1) & \kappa(t_2, t_2) & \dots & \kappa(t_2, t_N) \\ \vdots & \ddots & & \vdots \\ \kappa(t_N, t_1) & \kappa(t_N, t_2) & \dots & \kappa(t_N, t_N) \end{pmatrix}_{N \times N}$$

- 2 La matriz de covarianza entre los $y(t_i)$ y un cierto $y(t)$ que se desea predecir es

$$\kappa(t, T) := \left(\kappa(t, t_1) \quad \kappa(t, t_2) \quad \dots \quad \kappa(t, t_N) \right)_{1 \times N}$$



Ecuación de predicción

- 1 La mejor predicción lineal (mínima vza.) de $y(t)$ dado el conjunto de muestras $y(T)$ es:

$$\hat{y}(t) = \kappa(t, T) \cdot (\kappa(T, T))^{-1} \cdot y(T)$$

- 2 La varianza* del error de predicción es:

$$E \left((y(t) - \hat{y}(t))^2 \right) = \kappa(t, t) - \kappa(t, T) \cdot (\kappa(T, T))^{-1} \cdot \kappa(t, T)^T$$

* Si el proceso es **gaussiano** (todas las vbles. **distribución normal**), esa fórmula sirve para determinar **intervalos de confianza**.



Procesos estacionarios

- En caso estacionario, la función $\kappa(t_1, t_2)$ depende sólo de la distancia temporal $\kappa(t_1, t_2) = \tilde{\kappa}(|t_2 - t_1|)$.
- La **transf. Fourier** de $\tilde{\kappa}$ sería la **power spectral density** (PSD) del proceso, su factorización espectral sería la FdT que, tras pasar un ruido blanco por ella, generaría dicha **PSD(ω)**.

*El concepto de **PSD(ω)** permite dar una interpretación "frecuencial" (**filtrado**) a la estimación/interpolación estadística.

Ejemplo: $\kappa(t_1, t_2) = e^{-\beta|t_1-t_2|}$ está asociada a **PSD(ω) = $\frac{1}{\omega^2+\beta^2}$** , generada por el paso de un ruido blanco por el filtro **$G(s) = \frac{1}{s+\beta}$** para convertirlo en coloreado.

*fact. espectral: $PSD(\omega) = G(j\omega)G(-j\omega) = \frac{1}{j\omega+\beta} \cdot \frac{1}{-j\omega+\beta} = \frac{1}{\omega^2+\beta^2}$

Autocovarianza exponencial \Leftrightarrow filtrado paso-bajo primer orden.

Ruido de medida

Si suponemos que a cada muestras de la señal se le añade un "ruido de medida", de sensor, **no** correlado con el resto de muestras ni con la variable a predecir, de desviación típica σ_m , entonces la matriz VC de la información disponible debe cambiarse a:

$$\tilde{\kappa}(T, T) = \kappa(T, T) + \sigma_m^2 \cdot I$$

La covarianza entre variable a predecir e información, $\kappa(t, T)$, se mantiene inalterada.

Si v es no correlada con a , entonces $E(a(b + v)) = E(ab) + E(av) = E(ab)$.

Equivalentemente, definir desde el principio $\kappa(T, T)$ con suma de componente que decrezca a cero en tiempo muy pequeño (\approx ruido de medida) y otro con correlación temporal a más largo plazo (\approx señal).



Conclusiones

- La predicción lineal óptima en sentido estadístico puede generalizarse a la estimación de series temporales en instantes no medidos.
- Conjunto infinito de variables aleatorias $y(t)$, proceso estocástico, de las que se dispone de un conjunto de medidas con ruido en distintos puntos, $y(T)$ y de un modelo de la covarianza o correlación en función de la distancia temporal, o cualquier $\kappa(t_1, t_2)$ que se suponga que describe la función de covarianza de la señal a estimar/filtrar.
- Permite, suponiendo una estructura subyacente de tipo “Gaussian Process”, estimar intervalos de confianza de las variables predichas.
- El caso estacionario (frecuente en aplicaciones) permite dar una interpretación frecuencial (filtrado).

