Revista Colombiana de Estadística
Nºs 23-24, 1991.

ALGUNOS COMENTARIOS SOBRE LAS SERIES HIDROLOGICAS

Darío Valencia Restrepo
Profesor Emérito
Universidad Nacional

Resumen. Se presentan algunos atributos de las series de tiempo que son de interés para el hidrólogo. Se discute la modelación, y se destacan algunos modelos significativos que se utilizan para generar series sintéticas muy útiles para simular el comportamiento de sistemas de recursos hidráulicos. En particular se analizan los modelos de *desagregación*. Finalmente, en el contexto de la predicción, se presenta una nueva visión del filtro de Kalman.

1. La mirada del hidrólogo.

De acuerdo con el campo escogido para analizar y sintetizar las series de tiempo, se tendrá unos intereses específicos que corresponden a la disciplina del analista. En lo que sigue, se mostrará algunos aspectos que suelen ser importantes para quienes trabajan en el campo de la hidrología.

2. Un ejemplo clásico: caudal o lluvias anuales.

Supóngase que hay interés por estudiar los caudales anuales en un cierto punto de un río, de los cuales se dispone información entre 1950 y 1990. El caudal en un determinado año puede provenir de promediar sobre un registro de caudales instantáneos o de promediar sobre caudales observados cada cierto tiempo (días, por ejemplo). En la Figura 1 se muestra la traza hidrológica obtenida para los años mencionados, denotándose q_t como el caudal anual en el año t. Se aclara que los puntos obtenidos se han unido por una línea continua.

Para el hidrólogo, el futuro aparece incierto y considera que q_{1991} es una variable aleatoria. Pero además supone que el caudal ocurrido en 1970 es una muestra de la también variable aleatoria p_{1970}.

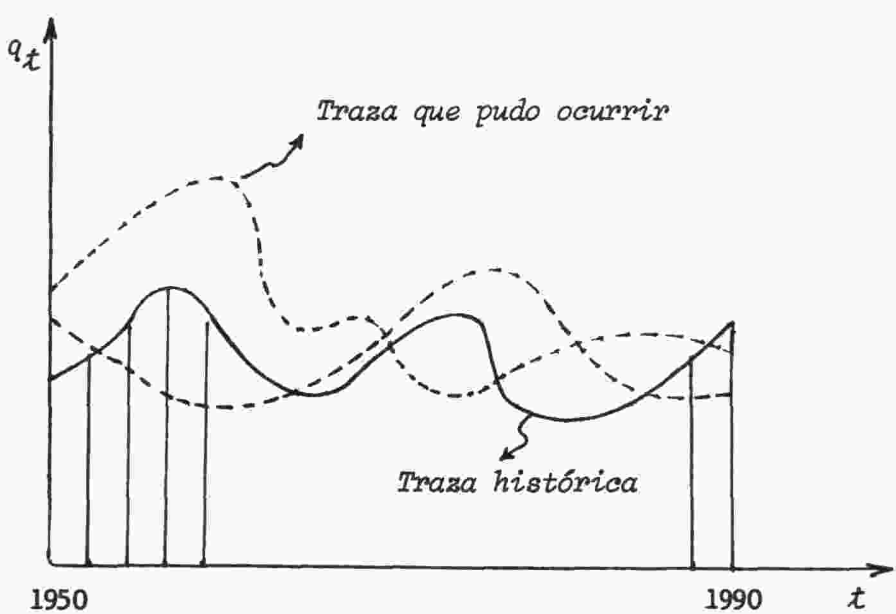


Figura 1

Ello lleva a señalar que el espacio muestral entre 1950 y 1990 está compuesto por la traza histórica y todas aquellas que "pudieron haber ocurrido". Entonces q_t puede verse como un proceso estocástico de argumento temporal discreto.

Aleatoriedad de procesos como el mencionado proviene de una estructura inherente al proceso, de la falta de información, o de la escala de observación.

Para el hidrólogo es importante la función de densidad de probabilidad de q_t, así como la función de distribución conjunta de (q_t, q_{t+1}), por ejemplo.

En primer lugar, hay interés por la media y la varianza de q_t. Se supone

$\mu_{q_t} = \mu$ para todo t, puesto que bajo ciertas condiciones (no hay trasvases aguas arriba del sitio en cuestión, ni se cambia en forma significativa los usos del suelo en la cuenca, etc.) puede suponerse que el proceso es estacionario en la media.

$\sigma_{q_t} = \sigma$ para todo t, pues bajo ciertas condiciones hay estacionaridad en la varianza.

De otro lado, al estudiar los caudales anuales de muchos ríos en el mundo, se ha observado una tendencia de caudales altos a ser seguidos por caudales altos, e idem para los bajos (excursiones del río), lo que constituye el fenómeno de persistencia. Una medida de éste puede ser el coeficiente de correlación serial de rezago uno

$$\rho = \frac{\mathrm{Cov}\left[q_{t+1}, q_t\right]}{\sigma^2}$$

el cual se supone también que no depende del tiempo.

Bajo los supuestos anteriores, se dice que el proceso de los caudales anuales es débilmente estacionario.

Los anteriores estadísticos son susceptibles de interpretación física.

μ: se relaciona con el balance hidrológico de la cuenca tributaria.

σ: se relaciona con la regulación natural de la cuenca.

ρ: una medida de la persistencia.

3. Modelación.

Inicialmente, podría pensarse que $q_{t+1} = f(q_t, \omega_{t+1})$ siendo ω_{t+1} un término aleatorio. Pero se propone una estructura de dependencia lineal:

$$q_{t+1} = aq_t + b + \omega_{t+1}$$

en donde

$$E[\omega_{t+1}] = 0 \quad \text{para todo } t$$

$$\text{Cov}[\omega_{t+1}, q_t] = 0 \quad \text{para todo } t$$

$$\text{Cov}[\omega_{t+1}, \omega_t] = 0 \quad \text{para todo } t$$

Utilizando la técnica de los valores esperados, es fácil hallar a y b, de modo que la ecuación queda

$$q_{t+1} = \mu + \rho(q_t - \mu) + \omega_{t+1} \, .$$

En lo que sigue, sin pérdida de generalidad se supondrá series de media nula:

$$q_{t+1} = \rho q_t + \omega_{t+1} \; .$$

Elevando al cuadrado ambos términos y tomando valores esperados,

$$E[q_{t+1}^2] = \rho^2 E[q_t^2] = 2\rho E[q_t \, \omega_{t+1}] + E[\omega_{t+1}^2]$$

$$\sigma^2 = \rho^2 \, \sigma^2 + 0 + \mathrm{Var}[\omega_{t+1}]$$

$$\therefore \quad \mathrm{Var}[\omega_{t+1}] = \sigma^2(1-\rho^2).$$

Con frecuencia se supone que ω_{t+1} es gaussiana, y como ya conocemos su media y su varianza, aquella queda completamente determinada.

A veces

$$\omega_{t+1} = b \; \nu_{t+1} \quad \text{con } \nu_{t+1} \sim N\;(0,1) \text{ para todo } t$$

$$\mathrm{Var}[\omega_{t+1}] = E[\omega_{t+1}^2] = b^2 \; .$$

Finalmente el modelo resulta ser

$$q_{t+1} = \rho q_t + \sigma\sqrt{1-\rho^2} \; \nu_{t+1} \; .$$

Se tiene entonces una aproximación al comportamiento condicional de q_{t+1} dado q_t. La media y la varianza condicionales son

$$E[q_{t+1}/q_t] = \rho q_t$$

$$\mathrm{Var}[q_{t+1}/q_t] = \sigma^2(1-\rho^2) \; .$$

Obsérvese la reducción de varianza, tanto mayor cuanto más se acerque ρ a +1 ó a -1.

Puede verse que se trata de un modelo autorregresivo de orden uno, que también es un modelo markoviano, y que en hidrología se conoce como el modelo de Thomas y Fiering.

En la práctica hay que hallar estimados de μ, σ y ρ.

$$q_{t+1} = \hat{\rho} q_t + \hat{\sigma} \sqrt{1-\hat{\rho}^2}\, \nu_{t+1} \, .$$

Como sólo se dispone de una sola traza (la histórica) del espacio muestral, el hidrólogo supone que el proceso es ergódico, lo que significa que promediar a través del espacio muestral es equivalente a promediar a lo largo de la traza histórica. Así

$$\hat{\mu} = \frac{q_{1950} + q_{1951} + \ldots + q_{1990}}{41}$$

y análogamente para $\hat{\sigma}$ y para $\hat{\rho}$.

Lo anterior permite afirmar que los parámetros del modelo son el resultado de preservar propiedades de primer y segundo momentos (estadísticos de primer y segundo orden).

En virtud de ello, si las distribuciones subyacentes son normales, el modelo es una representación exacta del comportamiento condicional de q_{t+1} dado q_t.

4. Generación sintética de caudales.

Si se genera números al azar provenientes de una distribución normal con media cero y varianza unitaria (los ν_{t+1}), y se usa recursivamente el modelo

$$q_{t+1} = \hat{\rho} q_t + \hat{\sigma} \sqrt{1-\hat{\rho}^2} \quad \nu_{t+1}$$

empezando con el último valor observado como el primer q_t, se obtiene tantas series como se quiera y de duración arbitraria. Estas series se denominan series sintéticas, y son estadísticamente indistinguibles de la serie histórica en términos de media, varianza y correlación serial de rezago uno. Ver Figura 2.

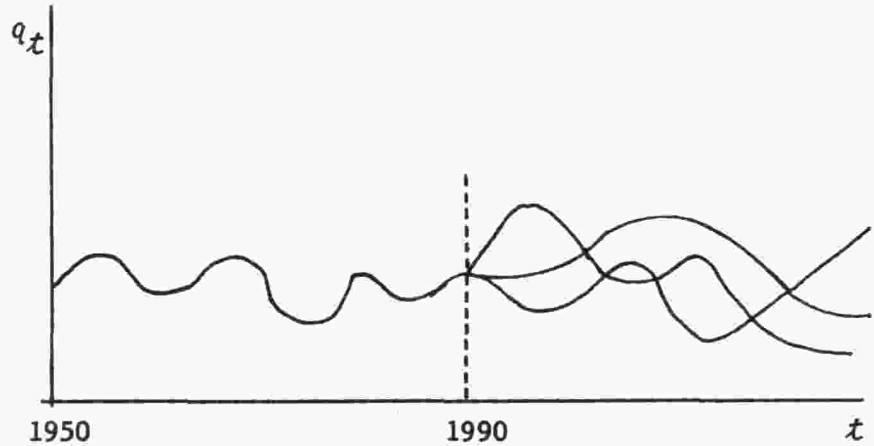


Figura 2

Las series sintéticas se emplean para simular el comportamiento de sistemas de recursos hidráulicos, los cuales pueden incluir embalses, plantas hidroeléctricas, áreas de riego, derivaciones, etc. La simulación permite analizar el comportamiento del sistema, no sólo frente a la serie histórica, sino frente a una gama de series sintéticas (que en cierto sentido despliegan la información histórica) cuya variabilidad como estímulo enriquece las respuestas del sistema.

En la figura 3 se muestran las respuestas del sistema (valores obtenidos por venta de energía, por ejemplo) ante

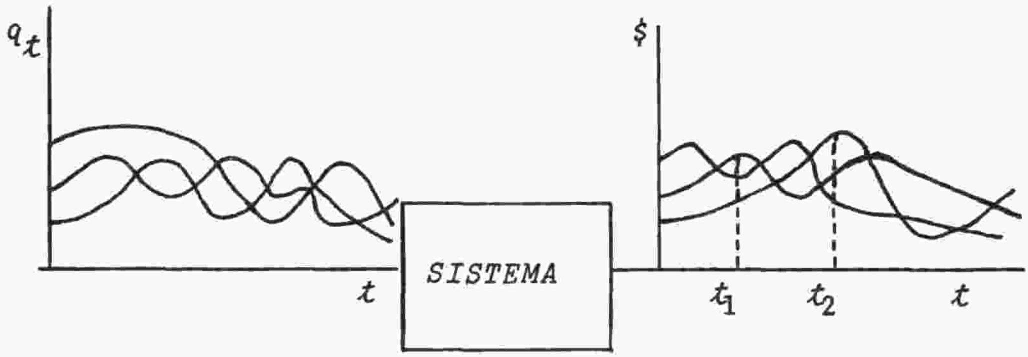


Figura 3

los estímulos del sistema. Puede entonces hallarse estimación de funciones de distribución derivadas.

De particular interés es el caso de un sistema no estacionario (por ejemplo la disponibilidad de agua para una planta hidroeléctrica decrece con el tiempo), pues ahora no podría invocarse ergodicidad en las respuestas, siendo los estadísticos en t_1 diferentes de los de t_2. Pero ahora sí se dispone de varias muestras en el espacio muestra sintético.

Con base en las series sintéticas puede encontrarse para dicho sistema:

- Rendimiento (inclusive en el largo plazo)
- Confiabilidad
- Vulnerabilidad
- Capacidad de recuperación.

5. Modelos autorregresivos de mayor orden.

Puede pensarse que

$$q_{t+1} = aq_t + cq_{t-1} + \omega_{t+1}$$

o sea, un modelo autorregresivo de orden dos; y así sucesivamente podría pensarse en órdenes mayores. Surge, sin embargo, el problema de estimar coeficientes de correlación serial con rezago alto, ya que en el mejor de los casos suele disponer de una serie histórica no muy larga, y bien se sabe que los estimadores de dichos coeficientes son de alta varianza.

Sin embargo, para reproducir memorias muy largas se suele usar otro tipo de modelos: ARIMA, ruido fraccional gaussiano, línea quebrada (Poveda y Mesa, (1987)).

Finalmente, conviene señalar que con modelos como los vistos y los que se verán más adelante, es posible extender o rellenar vacíos en registros históricos.

6. Generación multivariada.

Si se contempla dos sitios sometidos a regímenes hidrológicos similares, no es adecuado utilizar separadamente el modelo de Thomas y Fiering para los sitios 1 y 2

$$q_{1,t+1} = \rho_1 \, q_{1t} + \sigma_1 \sqrt{1-\rho_1^2} \, \nu_{1,t+1}$$

$$q_{2,t+1} = \rho_2 \, q_{2t} + \sigma_2 \sqrt{1-\rho_2^2} \, \nu_{2,t+1}$$

pues no se reproduce la correlación espacial que sin duda está presente.

En vista de lo anterior, se recurre a

$$\begin{bmatrix} q_{1,t+1} \\ q_{2,t+1} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} q_{1t} \\ q_{2t} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} v_{1,t+1} \\ v_{2\ t+1} \end{bmatrix}$$

Y así sucesivamente para más de dos sitios. Se llega al modelo de Matalas

$$Q_{t+1} = A\ Q_t + B\ V_{t+1}\ .$$

Empleando valores esperados y teniendo en cuenta que la matriz de covarianza entre los vectores aleatorios R y S es

$$S_{RS} = E\ [(R-\bar{R})(S-\bar{S})^T]$$

se llega a:

$$A = S_{Q_{t+1}Q_t}\ S^{-1}_{Q_tQ_t}$$

$$BB^T = S_{Q_{t+1}Q_{t+1}} - A\ S_{Q_tQ_t} A^T\ .$$

En la práctica se trabaja con estimados a partir de registros históricos:

$$\hat{A} = \widehat{S_{Q_{t+1}\ Q_t}}\ \widehat{S^{-1}_{Q_tQ_t}}$$

$$\widehat{BB^T} = \widehat{S_{Q_{t+1}Q_{t+1}}} - \hat{A}\ \widehat{S_{Q_tQ_t}}\ \widehat{A^T} = C$$

para que la ecuación matricial $\widehat{BB^T} = C$ tenga solución, se requiere que C sea definida positivamente, o al menos semidefinida positivamente.

El modelo de Matalas preserva

- Media y varianza anuales en cada sitio.
- Covarianza entre pares de sitio para el mismo año.
- Covarianza entre pares de sitios para años consecutivos (y para cada sitio consigo mismo, o sea, la covarianza serial de rezago uno).

El modelo de Matalas proporciona el comportamiento condicional de Q_{t+1} dado Q_t.

$$Q_{t+1} = A\, Q_t + B\, V_t$$

$$E\,[Q_{t+1}/Q_t] = \hat{A}\, Q_t$$

Matriz de covarianza $[Q_{t+1}, Q_{t+1}/Q_t] = BB^T$

Si las distribuciones subyacentes son gaussianas, el modelo reproduce exactamente el comportamiento condicional de Q_{t+1} dado Q_t (se conoce el vector medio condicional y la matriz de covarianza condicional).

7. Series estacionales (por ejemplo, mensuales)

En muchas aplicaciones es de interés observar comportamiento de sistemas durante períodos intraanuales (trimestres, bimestres, meses). Considérese el caso de una serie mensual

$$q_{1t}, q_{2t}, \ldots, q_{12,t}, q_{1,t+1}, q_{2,t+1}, \ldots, q_{12,t+1}, \ldots$$

Ahora no es posible suponer estacionaridad. Pero si se estandariza las variables

$$z_{it} = \frac{q_{it} \ \mu_{qi}}{\sigma_{qi}} \quad \text{para todo } i$$

hay ya estacionaridad en la media (valor nulo) y en la varianza (valor unitario).

Así se llega a

$$z_{i+1,t} = \rho_i \, z_{it} + \sqrt{1-\rho^2} \, v_t \qquad i = 1,2,\ldots,11$$

$$z_{1,t+1}, = \rho_{12} \, z_{12,t} + \sqrt{1-\rho_{12}^2} \, v_{t+1}$$

Hay dos caminos para definir a ρ_i

- Suponer que existen 12 coeficientes de correlación serial de rezago uno.
- Suponer que existe un ρ único, que es cierto "promedio" de los ρ_i.

En el segundo caso (hipótesis operacional) se dispone de muchos más datos.

Como puede verse, no es difícil extender el caso univariado al caso multivariado.

Los valores mensuales que se generen sintéticamente pueden agregarse para formar años sintéticos. No es difícil deducir que estos años sintéticos preservan la media anual histórica pero no otros estadísticos anuales de especial importancia (como la varianza anual histórica, por ejemplo).

Surgen entonces los modelos de desegregación, introducidos por Valencia y Schaake, (1973). Pero antes, analicemos un modelo particular.

8. Un modelo lineal muy práctico.

Sea

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{21} & a_{22} \quad a_{2n} \\ \cdots & \\ a_{m1} & a_{m2} \quad a_{mn} \end{bmatrix} \begin{bmatrix} x \\ x \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \cdots b_{1m} \\ b_{21} & b_{22} \quad b_{2m} \\ \cdots & \\ b_{m1} & b_{m2} \quad b_{mm} \end{bmatrix} \begin{bmatrix} v \\ v \\ \vdots \\ v_m \end{bmatrix}$$

$$\underset{m\times 1}{Y} = \underset{m\times n}{A}\ \underset{n\times 1}{X} + \underset{m\times m}{B}\ \underset{m\times 1}{V}$$

órdenes

Este tipo de modelo se propone cuando los vectores aleatorios Y y X estan correlacionados y se desea una expresión para el comportamiento condicional de Y dado X.

Empleando valores esperados y suponiendo

$$S_{VV} = I \quad \text{y} \quad S_{VX} = 0$$

se alcanza

$$A = S_{YX}\, S_{XX}^{-1} \qquad \text{(preservación de } S_{YX}\text{)}$$

$$BB^T = S_{YY} - S_{YX}\, S_{XX}^{-1}\, S_{XY} \qquad \text{(preservación de } S_{YY}\text{)}$$

Es fácil ver que a partir del modelo en cuestión se lle ga al modelo de Matalas si

$$X = Q_t \qquad\qquad X = q_t$$
$$Y = Q_t \qquad\qquad Y = q_{t+1}$$

$\left.\begin{matrix} X = q_t \\ Y = q_{t+1} \end{matrix}\right\}$ Thomas y Fiering

O a un modelo autorregresivo de orden 2 en el caso multivariado si

$$X = \begin{bmatrix} Q_t \\ \cdots \\ Q_{t-1} \end{bmatrix}$$

$$Y = Q_{t+1}$$

El modelo lineal propuesto se utilizará para desarrollar los modelos de desagregación.

9. El modelo de desagregación.

Sea

y_{ijt} : caudal mensual en el sitio i, mes j, año t.

x_{it} : caudal anual en el sitio i, año t.

Se propone:

$$\begin{bmatrix} y_{11t} \\ y_{12t} \\ \vdots \\ y_{1,12,t} \\ y_{21t} \\ y_{22t} \\ \vdots \\ y_{2,12,t} \\ \vdots \\ y_{n1t} \\ y_{n2t} \\ \vdots \\ y_{n,12,t} \end{bmatrix} = \underset{12n\times n}{A} \begin{bmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{bmatrix} + \underset{12n\times 12n}{B} \begin{bmatrix} v_{1t} \\ v_{2t} \\ \vdots \\ v_{12n,t} \end{bmatrix}$$

A y B se estiman utilizando las ecuaciones genéricas antes vistas. En la práctica, se trabajará con $\hat{A}$ y $\hat{B}$.

Como se nota, el modelo se usa para desagregar valores anuales en valores mensuales.

En la desagregación se preserva:

- Media y varianza de cualesquiera mes y sitio.
- Covarianza entre meses del mismo sitio o sitios distintos (S_{yy}).
- Covarianza entre cualquier valor mensual y cualquier valor anual (S_{yX}).
- Además, una propiedad muy agradable

$$\frac{1}{12}\sum_{j=1}^{12} y_{ijt} = x_{it} \quad \text{para todo } i\,.$$

O sea, que la desagregación preserva relaciones lineales entre meses y el año correspondiente (el promedio de caudales mensuales da el valor anual; la suma de lluvias mensuales da la lluvia anual. Y ello para cualquier sitio).

Si A y B se estiman con ciertos estimadores, las relaciones anteriores siguen menteniéndose.

Surge una propiedad muy fuerte: si los X_{it} son generadores mediante modelos anuales que preservan propiedades especiales, entonces la desagregación respeta dichas propiedades.

Es posible efectuar desegregación por pasos.

Por ejemplo:

De años y trimestres.

De trimestre a meses.

O también:

De años a estaciones (una seca y otra húmeda).

De estaciones a meses.

La desagregación reiterada preserva estadísticos en todos los niveles de agregación. Obsérvese que este tipo de modelo no es parsimonioso.

10. Desagregación espacial.

Es posible desagregar un valor x, que corresponde a un cierto espacio, en valores correspondientes a subespacios.

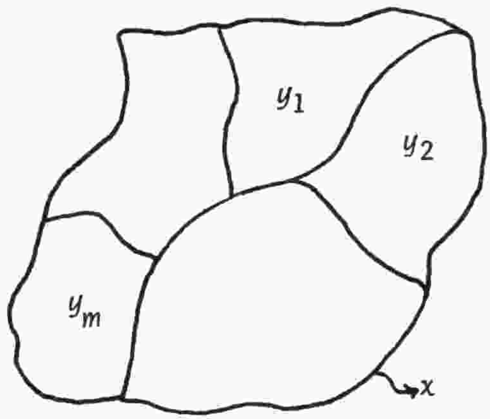
Por ejemplo:

x : escorrentía total de una cuenca.

y_i : escorrentía en la subcuenca i.

Y el modelo resulta ser:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \underset{m\times 1}{A} X + \underset{m\times m}{B} V$$



11. Otras aproximaciones al modelo $Y = AX + BV$.

Una primera aproximación (ya vista) resulta de utilizar valores esperados preservando estadísticos de primer y segundo orden. Los resultados que se obtienen son independientes de los tipos subyacentes de distribución.

La segunda aproximación proviene de un teorema de la teoría de la distribución multinormal:

Si el vector $\begin{bmatrix} Y \\ \vdots \\ X \end{bmatrix}$ se distribuye según la multinormal, entonces el comportamiento condicional de Y dado X es también multinormal con

$$E[Y/X] = S_{YX} S_{XX}^{-1} X$$

$$\text{MATRIZ DE } COV[Y,Y/X] = S_{YY} - S_{YX} S_{XX}^{-1} S_{XY}$$

Obsérvese que la media condicional es equivalente a

AX, y la matriz de covarianza condicional es igual a BB^T, con A y B los parámetros del modelo lineal. En ese caso el modelo es una representación exacta (en términos probabilísticos) del comportamiento condicional de Y dado X, pues una distribución multinormal está plenamente definida si se conoce su vector medio y la matriz de covarianza del vector aleatorio consigo mismo.

La tercera aproximación proviene de la regresión múltiple multivariada. En efecto

$$A = S_{YX}\, S_{XX}^{-1}$$

es la matriz de la regresión. Entonces

$$y_i = \sum_{j=1}^{n} a_{ij}\, x_j = \sum_{j=1}^{m} b_{ij}\, v_j \qquad i = 1,2,\ldots,m.$$

Las ecuaciones

$$y_i = \sum_{j=1}^{n} a_{ij}\, x_j$$

representan m hiperplanos en el espacio euclídeo de n+1 dimensiones, hiperplanos que corresponden al ajuste mínimocuadrático. La novedad es que los residuos $\Sigma b_{ij}, v_j$ se correlacionan para preservar la estructura de correlación de los y_i dada por la matriz de covarianza condicional de Y consigo mismo dado X.

12. Desarrollos ulteriores de los modelos de desagregación.

Si se recuerda el modelo de desagregación de valores anuales en valores mensuales, se concluye que no se preserva el coeficiente de correlación serial entre diciembre de un año y enero el siguiente. Para subsanar esta deficiencia, Mejía y Rouselle (1976) propusieron una variación del modelo básico:

$$Y_t = AX_t + C \begin{bmatrix} y_{1,12,t-1} \\ y_{2,12,t-1} \\ \vdots \\ y_{n,12,t-1} \end{bmatrix} + B\, V_t$$

O sea, que se hace depender los elementos de Y_t (valores mensuales en los diversos sitios para el año t) de los valores de diciembre en los diferentes sitios y correspondientes a t-1.

Infortunadamente, el modelo tiene una inconsistencia teórica, la cual fue corregida en el trabajo de Valencia, Berdugo y García (1983).

Durante la década de los años ochentas, hubo diversos trabajos encaminados a reducir el alto número de parámetros en los modelos de desagregación. Para el efecto, se mencionan los trabajos de Santos y Salas, (1983) y Pachón y Valencia (1987).

13. Predicción.

Modelos como los vistos antes pueden usarse para hacer pronósticos. Por ejemplo, con el modelo markoviano

$$q_{t+1} = \rho q_t + \sigma\sqrt{1-\rho^2}\; \nu_{t+1}$$

se tiene que

$$E\left[q_{t+1}/q_t\right] = \rho q_t$$

$$Var\left[q_{t+1}/q_t\right] = \sigma^2(1-\rho^2)\ .$$

Si los comportamientos son gaussianos, se tendría una situación como la de la figura 4, en donde se muestran las funciones de densidad de probabilidad para q_{t+1} (la marginal) y la condicional de q_{t+1} dado q_t.

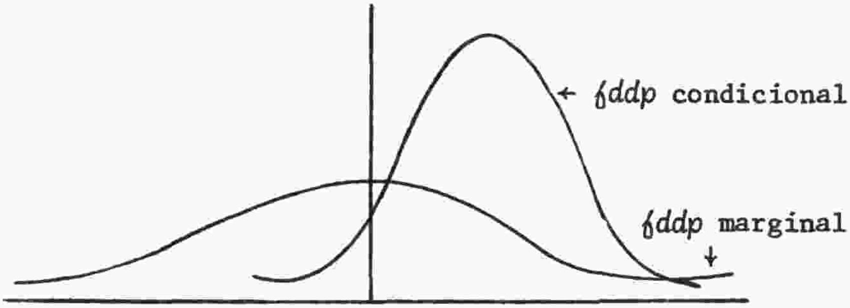


Figura 4

Se llama la atención sobre la reducción de varianza si se usa como estimador

$$\widehat{q_{t+1}} = \rho q_t$$

Si se tiene modelos de memoria más larga

$$q_{t+1} = A \begin{bmatrix} q_t \\ q_{t-1} \\ \vdots \\ q_{t-m} \end{bmatrix} + BV$$

Y si se conoce $q_{t-1}, q_{t-2}, \ldots, q_{t-m}$, así como un estimado de q_t, igual a $\hat{q}_t$, podría hallarse un estimador con adelanto de dos períodos

$$\widehat{q_{t+1}} = A \begin{bmatrix} \hat{q}_t \\ q_{t-1} \\ q_{t-} \\ \vdots \\ q_{t-m} \end{bmatrix}$$

A veces en el vector X se incluyen sitios del río aguas arriba o sitios en otras cuencas vecinas.

También se utilizan modelos del tipo ARIMA para hacer pronósticos.

Como rara vez el sistema es directa o perfectamente observable, se recurre a los filtros como el de Kalman.

14. Una nueva visión del filtro de Kalman.

En la teoría de los sistemas lineales, la ecuación de estado viene dada por

$$x_{k+1} = A_k\, x_k + B_k\, \mu_k$$

en donde:

x_k: vector de estado (atributos) del sistema en el tiempo k.

μ_k: vector de control en el tiempo k.

A veces la ecuación de estado es más compleja

$$x_{k+1} = f_k(x_k, \mu_k) .$$

Se recurre en estos casos a una expansión en serie de Taylor en el entorno de x_k^* y μ_k^*, denominados estado nominal y control nominal en k. Si se define las perturbaciones de estado y control como

$$\delta x_k = x_k - x_k^*$$

$$\delta \mu_k = \mu_k - \mu_k^*$$

el truncamiento de la serie de Taylor conduce a la llamada ecuación de estado en forma pertubacional

$$\delta x_{k+1} = A_k \, \delta x_k + B_k \, \delta \mu_k$$

Como hay errores (de modelación, de estimación) y puede haber impulsos aleatorios distintos al control, suele agregarse un vector aleatorio:

$$\delta x_{k+1} = A_k \, \delta x_k + B_k \, \delta \mu_k + \alpha_k .$$

Análogamente existirá una ecuación de medición en forma perturbacional

$$\delta y_k = C_k \, \delta x_k + \beta_k$$

Una posible presentación del filtro de Kalman es

$$\delta x_{k+1/k+1} = \delta x_{k+1/k} + H_{k+1}(\delta y_{k+1} - \delta y_{k+1/k})$$

O sea, la perturbación de estado en k+1 considerando toda la información hasta la llegada de la observación δy_{k+1}, que se denota por $\delta x_{k+1/k+1}$, es igual a la perturbabación predicha con base en la ecuación de estado en forma perturbacional, $\delta x_{k+1/k}$, más una corrección tanto más importante cuanto más difiera la medición predicha, $\delta y_{k+1/k}$, de la medición hecha, δy_{k+1}.

El filtro se ha hallado de diversas maneras:

- Suponiendo comportamientos gaussianos.
- Hallando el mejor estimador lineal con ayuda de las proyecciones ortogonales en un espacio euclídeo de dimensión finita.
- Usando un estimador mínimo-cuadrático ponderado.

Etc.

En el artículo de Valencia (1988) se deduce el filtro con ayuda exclusivamente del modelo

$$Y = AX + BV$$

lo que demuestra que el filtro proviene de preservar estadísticos de primer y segundo orden, y que no se requiere que los comportamientos sean gaussianos. También se demuestra allí que el estimador de Kalman es mínimo-cuadrático.

Pero si los comportamientos son gaussianos, el filtro de Kalman es también

- el de mínima varianza,
- el de máxima verosimilitud.

*

BIBLIOGRAFIA

Mejía y Rouselle, (1976) Disaggregation models in hidrology revisited. Water Resources Research, Vol. 12, # 2.

Pachón y Valencia, (1987) Un Modelo de Degregación en Hidrología. Tesis de Magíster, Posgrado en Aprovechamiento de Recursos Hídráulicos. Facultad de Minas, Medellín.

Poveda y Mesa. (1987) El Fenómeno de Hurst. Tesis de Magíster Posgrado en Aprovechamiento de Recursos Hidráulicos. Facultad de Minas, Medellín.

Santos y Salas. (1983) A parsimonious step diseggregation model for operational hydrology. Presentado en la reunión de otoño de la American Geophysical Union. San Francisco, California.

Valencia y Shaake. (1973) Disaggregation processes in stochastic hydrology. Water Resources Research. Vol. 9 # 3.

Valencia, Berdugo y García (1983) Disaggregation models in hydrology - An evaluation. Conferencia por invitación presentada en la reunión de otoño de la American Geophysical Union. San Francisco, California. Ver también Berdugo, García y Valencia (1983) Evaluación crítica de los modelos de desagregación en hidrología. Trabajo dirigido de grado, Facultad de Minas. Medellín.

Valencia, (1988) Una nueva visión del estimador de Kalman, revista DYNA Nº 111, Facultad de Minas. Medellín.

*