

Inteligencia predictiva a través de modelos de Espacio de los Estados y otras consideraciones

Diego J. Pedregal

UNIVERSIDAD DE CASTILLA-LA MANCHA

ESCUELA TÉCNICA SUPERIOR INGENIEROS INDUSTRIALES (CIUDAD REAL)

Diego.Pedregal@uclm.es

Universidad Complutense de Madrid

18 de Mayo de 2018

Esquema:

- Inteligencia predictiva a través de modelos de Espacio de los Estados
- Otras consideraciones: Investigación, Big Data, etc.

Inteligencia predictiva a través de modelos de Espacio de los Estados

- El Modelo EE general (flexible):

Ecuación de Estados : $\mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t$

Ecuación de Observación : $\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t$

- Forma muy general, aplicable a:
 - Modelos con parámetros cambiantes.
 - Modelos con heterocedasticidad.
 - Modelos en tiempo continuo.
 - Series con diferentes períodos de muestreo.
 - Agregación temporal.
 - Algunas familias de modelos no lineales.
 - Modelos no - Gaussianos
 - ...

- El Modelo EE general (flexible):

Ecuación de Estados: $\mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t$

Ecuación de Observación: $\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t$

- Cuatro características del modelo:
 - Dos sistemas de ecuaciones.
 - Todos los elementos que aparecen son matrices o vectores (negrita).
 - Todo tiene subíndices de tiempo.
 - La ecuación de transición es dinámica, la de observación estática.

- El Modelo EE general (flexible):

Ecuación de Estados: $\mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t$

Ecuación de Observación: $\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t$

Información P. estocástica

$$\left\{ \begin{array}{l} \mathbf{w}_t : \text{vector de ruidos } v \times 1 \quad N(\mathbf{0}; \mathbf{Q}_t) \\ \mathbf{v}_t : \text{vector de ruidos observables } h \times 1 \quad N(\mathbf{0}; \mathbf{R}_t) \\ \mathbf{S}_t = E(\mathbf{v}_t \mathbf{w}_t') \quad \text{matriz de covarianzas de ruidos } h \times v \end{array} \right.$$

$$\left\{ \begin{array}{l} \mathbf{u}_t : \text{Vector de inputs o entradas } k \times 1 \\ \mathbf{z}_t : \text{Vector de outputs o salidas } m \times 1 \end{array} \right.$$

Vector de estados: $\mathbf{x}_t : N \times 1$

- El Modelo EE general (flexible):

Ecuación de Estados: $\mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t$

Ecuación de Observación: $\mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t$

M. del sistema P. estocástica

$$\left\{ \begin{array}{l} \mathbf{w}_t : \text{vector de ruidos } v \times 1 \quad N(\mathbf{0}; \mathbf{Q}_t) \\ \mathbf{v}_t : \text{vector de ruidos observables } h \times 1 \quad N(\mathbf{0}; \mathbf{R}_t) \\ \mathbf{S}_t = E(\mathbf{v}_t \mathbf{w}_t') \quad \text{matriz de covarianzas de ruidos } h \times v \end{array} \right.$$

$$\left\{ \begin{array}{lll} \Phi_t : N \times N & \Gamma_t : N \times k & \mathbf{E}_t : N \times v \\ \mathbf{H}_t : m \times N & \mathbf{D}_t : m \times k & \mathbf{C}_t : m \times h \\ \mathbf{Q}_t : v \times v & \mathbf{R}_t : h \times h & \mathbf{S}_t : h \times v \end{array} \right.$$

– Ejemplo 1: AR(1)

$$z_{t+1} = \phi z_t + v_t$$

$$\begin{cases} x_{t+1} = \phi x_t + w_t \\ z_t = x_t \end{cases}$$

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

– Ejemplo 2: ARMA(1, 1)

$$z_t = \phi z_{t-1} + v_t + \mathcal{I}v_{t-1}$$

$$\begin{cases} x_{t+1} = \phi x_t + (\phi + \mathcal{I})v_t \\ z_t = x_t + v_t \end{cases}$$

- Dos cuestiones fundamentales:

- 1. Estimación óptima de estados:

- Filtro de Kalman (KF):
 - Algoritmo de Suavizado de Intervalo Fijo (FIS):

- 2. Estimación óptima de parámetros desconocidos (hiper-parámetros):

- Máxima Verosimilitud en el dominio del tiempo, utilizando el Filtro de Kalman.
 - Otras posibilidades.

- Dos cuestiones fundamentales:

- 1. Estimación óptima de estados:

- Filtro de Kalman (KF):

- Ecuaciones de predicción:

- » Predicción óptima para el vector de estados

- » Predicción óptima para su matriz de covarianzas

- Ecuaciones de adaptación:

- » Estimación óptima del vector de estados.

- » Estimación óptima de su matriz de covarianzas

- Innovaciones y su matriz de varianzas y covarianzas.

- Observaciones ausentes y predicción.

- Inicialización, intervenciones de varianza y optimalidad.

- Algoritmo de Suavizado de Intervalo Fijo (FIS):

- Estimación óptima del vector de estados.

- Estimación óptima de su matriz de covarianzas.

- Dos cuestiones fundamentales:

- 1. Estimación óptima de estados:

- Filtro de Kalman (KF; con $\mathbf{S}=\mathbf{0}$):

- Ecuaciones de predicción:

$$\hat{\mathbf{x}}_{t|t-1} = \boldsymbol{\phi}_t \hat{\mathbf{x}}_{t-1} + \boldsymbol{\Gamma}_t \mathbf{u}_t$$

$$\hat{\mathbf{P}}_{t|t-1} = \boldsymbol{\phi}_t \hat{\mathbf{P}}_{t-1} \boldsymbol{\phi}_t^T + \mathbf{E}_t \mathbf{Q}_t \mathbf{E}_t^T$$

- Ecuaciones de adaptación:

$$\hat{\mathbf{v}}_t = \mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1} - \mathbf{D} \mathbf{u}_t$$

$$\hat{\mathbf{F}}_t = \left[\mathbf{H}_t \hat{\mathbf{P}}_{t|t-1} \mathbf{H}_t^T + \mathbf{C}_t \mathbf{R}_t \mathbf{C}_t^T \right]$$

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \hat{\mathbf{P}}_{t|t-1} \mathbf{H}_t^T \hat{\mathbf{F}}_t^{-1} \hat{\mathbf{v}}_t$$

$$\hat{\mathbf{P}}_t = \hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1} \mathbf{H}_t^T \hat{\mathbf{F}}_t^{-1} \mathbf{H}_t \hat{\mathbf{P}}_{t|t-1}$$

- Dos cuestiones fundamentales:

- 1. Estimación óptima de estados:

- Algoritmo de Suavizado de Intervalo Fijo (FIS):

$$\hat{\mathbf{x}}_{t|N} = \hat{\mathbf{x}}_{t|t-1} + \hat{\mathbf{P}}_{t|t-1} \mathbf{s}_{t-1}$$

$$\hat{\mathbf{P}}_{t|N} = \hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1} \mathbf{S}_{t-1} \hat{\mathbf{P}}_{t|t-1}$$

$$\mathbf{s}_{t-1} = \mathbf{H}_t^T \mathbf{F}_t^{-1} (\mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1}) + \overline{\mathbf{\Phi}}_t^T \mathbf{s}_t$$

$$\text{with } \mathbf{s}_N = \mathbf{0}$$

$$\mathbf{S}_{t-1} = \mathbf{H}_t^T \mathbf{F}_t^{-1} \mathbf{H}_t + \overline{\mathbf{\Phi}}_t^T \mathbf{S}_t \overline{\mathbf{\Phi}}_t$$

$$\text{with } \mathbf{S}_N = \mathbf{0}$$

$$\overline{\mathbf{\Phi}}_t = \mathbf{\Phi}_t - \mathbf{\Phi}_t \hat{\mathbf{P}}_{t|t-1} \mathbf{H}_t^T \mathbf{F}_t^{-1} \mathbf{H}_t$$

- Dos cuestiones fundamentales:
 - 2. Estimación óptima de parámetros desconocidos (hiper-parametros):
 - Máxima Verosimilitud en el dominio del tiempo, utilizando el Filtro de Kalman.

$$\ln L = -\frac{mT}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^T \ln |\hat{\mathbf{F}}_t| - \frac{1}{2} \sum_{t=1}^T \hat{\mathbf{v}}_t' \hat{\mathbf{F}}_t^{-1} \hat{\mathbf{v}}_t$$

$$\hat{\mathbf{v}}_t = \mathbf{z}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1} - \mathbf{D} \mathbf{u}_t$$

$$\hat{\mathbf{F}}_t = \left[\mathbf{H}_t \hat{\mathbf{P}}_{t|t-1} \mathbf{H}_t^T + \mathbf{C}_t \mathbf{R}_t \mathbf{C}_t^T \right]$$

- **Otras posibilidades.**

Ejemplo 1: Nivel medio cambiante en el tiempo.

$$z_t = T_t + v_t = \frac{1}{(1-B)} w_t + v_t \quad \begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$\begin{cases} x_{1t+1} = x_{1t} + w_t \\ z_t = x_{1t} + v_t \end{cases}$$

R y Q estimadas por MV

Ejemplo 2: Caso anterior Multivariante.

$$\begin{cases} \mathbf{x}_{1t+1} = \mathbf{x}_{1t} + \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{x}_t + \mathbf{v}_t \end{cases}$$

\mathbf{R} y \mathbf{Q} no diagonales estimadas por MV

Ejemplo 3: Filtro de Hodrick-Prescott (IRW).

$$z_t = T_t + v_t = \frac{1}{(1-B)^2} w_t + v_t \quad \begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$\begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{t+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t + \begin{pmatrix} 0 \\ 1 \end{pmatrix} w_t \\ z_t = (1 \quad 0) \mathbf{x}_t + v_t \end{cases} \quad \lambda = \frac{\text{var}(v_t)}{\text{var}(w_t)} = \frac{R}{Q} = 1600$$

Ejemplo 4: HP Multivariante.

$$\begin{cases} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}_{t+1} = \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}_t + \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix} \mathbf{w}_t \\ \mathbf{z}_t = (\mathbf{I} \quad \mathbf{0}) \mathbf{x}_t + \mathbf{v}_t \end{cases}$$

R y **Q** no diagonales estimadas por MV

Ejemplo 5: Local Linear Trend (LLT).

$$z_t = T_t + v_t$$

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$\begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{t+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}_t \\ z_t = (1 \quad 0) \mathbf{x}_t + v_t \end{cases}$$

$$\mathbf{Q} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \quad R = \sigma^2$$

– Ejemplo 6: Componentes NO Observables.

$$z_t = T_t + C_t + S_t + v_t \quad \begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

- **Modelo de Estacionalidad (o ciclo):**

$$\begin{cases} \begin{pmatrix} S \\ S' \end{pmatrix}_{t+1} = \rho \begin{pmatrix} \cos \omega & \text{sen } \omega \\ -\text{sen } \omega & \cos \omega \end{pmatrix} \begin{pmatrix} S \\ S' \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}_t \\ S_t = (1 \quad 0) \mathbf{x}_t; \quad Q = \begin{pmatrix} q & 0 \\ 0 & q \end{pmatrix} \end{cases}$$

- **Modelo completo:**

$$\begin{cases} \begin{pmatrix} T \\ D \\ S \\ S' \end{pmatrix}_{t+1} = \left(\begin{array}{cc|cc} 1 & 1 & & \\ 0 & 1 & & \\ \hline & & \rho \cos \omega & \rho \text{sen } \omega \\ \mathbf{0} & & -\rho \text{sen } \omega & \rho \cos \omega \end{array} \right) \begin{pmatrix} T \\ D \\ S \\ S' \end{pmatrix}_t + \left(\begin{array}{c|cc} 0 & & \\ \hline 1 & & \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{array} \right) \begin{pmatrix} w_0 \\ w_1 \\ w_2 \end{pmatrix}_t \\ z_t = (1 \quad 0 \mid 1 \quad 0) \mathbf{x}_t + v_t \end{cases}$$

– Ejemplo 7: Dynamic Harmonic Regression

$$z_t = T_t + C_t + S_t + v_t \quad \begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

- **Modelo de Estacionalidad (o ciclo):**

$$\left\{ \begin{array}{l} \begin{pmatrix} a \\ a' \end{pmatrix}_{t+1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ a' \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}_t \\ S_t = (\cos(\omega_1 t) \quad \sin(\omega_1 t)) \mathbf{x}_t; \quad Q = \begin{pmatrix} q & 0 \\ 0 & q \end{pmatrix} \end{array} \right.$$

- **Modelo completo:**

$$\left\{ \begin{array}{l} \begin{pmatrix} T \\ D \\ a \\ a' \end{pmatrix}_{t+1} = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 0 \\ & & 0 & 1 \end{pmatrix} \begin{pmatrix} T \\ D \\ a \\ a' \end{pmatrix}_t + \begin{pmatrix} 0 & & & \\ 1 & & & \\ & & 1 & 0 \\ & & 0 & 1 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \end{pmatrix}_t \\ z_t = (1 \quad 0 \mid \cos(\omega_1 t) \quad \sin(\omega_1 t)) \mathbf{x}_t + v_t \end{array} \right.$$

Ejemplo 8: Regresión lineal

$$\begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$z_t = \mathbf{X}_t \boldsymbol{\beta} + v_t$$

Versión 1: $z_t = \boldsymbol{\beta}' \mathbf{X}'_t + v_t$

Versión 2:
$$\begin{cases} \boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t + \mathbf{w}_t \\ z_t = \mathbf{X}_t \boldsymbol{\beta}_t + v_t \end{cases}$$

- Si $\mathbf{Q} = \mathbf{0}$ el Filtro de Kalman proporciona la estimación recursiva del sistema (útil cuando existen observaciones ausentes).
- Si $\mathbf{Q} > \mathbf{0}$ se puede realizar la estimación de una regresión lineal con parámetros cambiantes en el tiempo.

– Ejemplo 9: AR(2)

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + v_t$$

Versión 1:
$$\begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{t+1} = \begin{pmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t + \begin{pmatrix} 1 \\ 0 \end{pmatrix} v_t \\ z_t = (1 \quad 0) \mathbf{x}_t \end{cases}$$

Versión 2:
$$\begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{t+1} = \begin{pmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t + \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} v_t \\ z_t = (1 \quad 0) \mathbf{x}_t + v_t \end{cases}$$

Versión 3:
$$z_t = (\phi_1 \quad \phi_2) \begin{pmatrix} z_{t-1} \\ z_{t-2} \end{pmatrix} + v_t$$

Versión 4:
$$\begin{cases} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}_{t+1} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\ z_t = (z_{t-1} \quad z_{t-2}) \mathbf{x}_t + v_t \end{cases}$$

– Ejemplo 9: AR(2)

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + v_t$$

- En todos los casos, excepto en las versiones 3 y 4 se pueden estimar los coeficientes sin problemas de valores ausentes.
- Si $\mathbf{Q} = \mathbf{0}$ el Filtro de Kalman proporciona la estimación recursiva del sistema.
- Si $\mathbf{Q} > \mathbf{0}$ se puede realizar la estimación del modelo con parámetros cambiantes en el tiempo.

Versión 3:
$$z_t = (\phi_1 \quad \phi_2) \begin{pmatrix} z_{t-1} \\ z_{t-2} \end{pmatrix} + v_t$$

Versión 4:
$$\begin{cases} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}_{t+1} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}_t + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\ z_t = (z_{t-1} \quad z_{t-2}) \mathbf{x}_t + v_t \end{cases}$$

– Ejemplo 10: ARMA(1, 1).

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$z_t = \phi z_{t-1} + v_t + \mathcal{G}v_{t-1}$$

Versión 1:

$$\begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{t+1} = \begin{pmatrix} \phi & \mathcal{G} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t + \begin{pmatrix} 1 \\ 1 \end{pmatrix} v_t \\ z_t = (1 \quad 0) \mathbf{x}_t \end{cases}$$

Versión 2:

$$\begin{cases} x_{t+1} = \phi x_t + (\phi + \mathcal{G})v_t \\ z_t = x_t + v_t \end{cases}$$

– Ejemplo 11: VARMA(p, p).

$$\begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$\mathbf{z}_t = \varphi_1 \mathbf{z}_{t-1} + \dots + \varphi_p \mathbf{z}_{t-p} + \mathbf{v}_t + \Theta_1 \mathbf{v}_{t-1} + \Theta_p \mathbf{v}_{t-p}$$

$$\mathbf{x}_{t+1} = \begin{pmatrix} \varphi_1 & \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \varphi_2 & \mathbf{0} & \mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varphi_{p-1} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{I} \\ \varphi_p & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix} \mathbf{x}_t + \begin{pmatrix} \varphi_1 + \Theta_1 \\ \varphi_2 + \Theta_2 \\ \vdots \\ \varphi_{p-1} + \Theta_{p-1} \\ \varphi_p + \Theta_p \end{pmatrix} \mathbf{v}_t$$

$$\mathbf{z}_t = (\mathbf{I} \quad \mathbf{0} \quad \mathbf{0} \quad \dots \quad \mathbf{0}) \mathbf{x}_t + \mathbf{v}_t$$

– Ejemplo 12: MCNO Multivariantes.

$$\mathbf{z}_t = \mathbf{T}_t + \mathbf{C}_t + \mathbf{S}_t + \mathbf{v}_t$$

$$\begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$\begin{pmatrix} \mathbf{T} \\ \mathbf{D} \\ \mathbf{S} \\ \mathbf{S}' \end{pmatrix}_{t+1} = \left(\begin{array}{cc|cc} \mathbf{I} & \mathbf{I} & & \\ \mathbf{0} & \mathbf{I} & & \\ \hline & & \rho \cos \omega \mathbf{I} & \rho \text{sen} \omega \mathbf{I} \\ & & -\rho \text{sen} \omega \mathbf{I} & \rho \cos \omega \mathbf{I} \end{array} \right) \begin{pmatrix} \mathbf{T} \\ \mathbf{D} \\ \mathbf{S} \\ \mathbf{S}' \end{pmatrix}_t + \left(\begin{array}{cc|cc} \mathbf{0} & & & \\ \mathbf{I} & & & \\ \hline & & \mathbf{I} & \mathbf{0} \\ & & \mathbf{0} & \mathbf{I} \end{array} \right) \begin{pmatrix} \mathbf{w}_0 \\ \mathbf{w}_1 \\ \mathbf{w}_2 \end{pmatrix}_t$$

$$\mathbf{z}_t = (\mathbf{I} \quad \mathbf{0} \mid \mathbf{I} \quad \mathbf{0}) \mathbf{x}_t + \mathbf{v}_t$$

R y Q no diagonales estimadas por MV

– Ejemplo 14: Función de transferencia

$$z_t = \frac{w_1 B + w_2 B^2}{1 + \delta B} u_t + v_t$$

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$z_t = -\delta z_{t-1} + w_1 u_{t-1} + w_2 u_{t-2} + v_t + \delta v_{t-1}$$

$$\begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{t+1} = \begin{pmatrix} -\delta & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} u_t \\ z_t = (1 \quad 0) \mathbf{x}_t + v_t \end{cases}$$

– Ejemplo 15: Concatenación de sistemas.

Función de transferencia con ruido

$$z_t = \frac{w_1 B + w_2 B^2}{1 + \delta B} u_t + \frac{\mathcal{G}_1 B + \mathcal{G}_2 B^2}{1 + \phi B} v_t$$

$$\begin{cases} \mathbf{x}_{t+1} = \Phi_t \mathbf{x}_t + \Gamma_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$\begin{cases} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}_{t+1} = \left(\begin{array}{cc|cc} -\delta & 1 & & \\ 0 & 0 & & \\ \hline & & \phi_1 & 1 \\ & & \phi_2 & 0 \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}_t + \begin{pmatrix} w_1 \\ w_2 \\ 0 \\ 0 \end{pmatrix} u_{t+} + \begin{pmatrix} 0 \\ 0 \\ \phi + \mathcal{G}_1 \\ \mathcal{G}_2 \end{pmatrix} v_t \end{cases}$$

$$z_t = (1 \quad 0 \mid 1 \quad 0) \mathbf{x}_t + v_t$$

– Ejemplo 16: Modelo TAR.

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

$$z_t = \begin{cases} 1 + \phi_1^1 B + \phi_2^1 B^2 + v_{1,t} & z_{t-2} > 0 \\ 1 + \phi_1^2 B + \phi_2^2 B^2 + v_{2,t} & z_{t-2} \leq 0 \end{cases}$$

$$\begin{cases} z_t = \begin{pmatrix} \phi_1^1 & \phi_2^1 \end{pmatrix} \begin{pmatrix} z_{t-1} \\ z_{t-2} \end{pmatrix} + v_{1,t} & z_{t-2} > 0 \\ z_t = \begin{pmatrix} \phi_1^2 & \phi_2^2 \end{pmatrix} \begin{pmatrix} z_{t-1} \\ z_{t-2} \end{pmatrix} + v_{2,t} & z_{t-2} \leq 0 \end{cases}$$

– Ejemplo 17: Agregación temporal.

Supongamos una variable observada a espacios de tiempo irregular.

Desde el punto de vista de la agregación temporal, las variables pueden ser de dos tipos:

1. **Stock:** las observaciones son estáticas.
2. **Flujo:** las observaciones son “dinámicas”, una observación anual equivale a la suma (o el promedio) de todos los meses; una observación trimestral equivale a la suma de los tres trimestres; etc.

La agregación temporal de variables stock es interpolación.

– Ejemplo 17: Agregación temporal.

Variables flujo:

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}\mathbf{x}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{C}_t \mathbf{v}_t \end{cases} \quad \begin{cases} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{H}\mathbf{\Phi} \\ \mathbf{0} & \mathbf{\Phi} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{t-1} \\ \mathbf{x}_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{H} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{v}_t \\ \mathbf{w}_t \end{bmatrix} \\ \mathbf{z}_t = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} \end{cases}$$

$$\mathbf{C}_t = \begin{cases} 0, & t = \text{cuando existe el dato} \\ 1, & \text{en el resto} \end{cases}$$

$$\begin{cases} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} \mathbf{C}_t \otimes \mathbf{I} & \mathbf{H}\mathbf{\Phi} \\ \mathbf{0} & \mathbf{\Phi} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{t-1} \\ \mathbf{x}_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{H} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{v}_t \\ \mathbf{w}_t \end{bmatrix} \\ \mathbf{z}_t = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} \end{cases}$$

- SSpace:
 - Toolbox de series temporales que implementa un modelo EE con alto grado de flexibilidad.
 - Villegas, M.A., Pedregal, D.J. (2018), SSpace: A toolbox for State Space modelling, Journal of Statistical Software, in press.
 - <https://bitbucket.org/predilab/sspace-matlab/>

-
- Otros programas relacionados:
 - E4: Terceiro *et al.*, www.ucm.es/info/icae/e4
 - CAPTAIN: Young, PC *et al.*,
ww.es.lancs.ac.uk/cres/captain
 - Ox Metrics: www.oxmetrics.net
 - Ssfpack for Ox: www.ssfpack.com
 - STAMP: www.stamp-software.com
 - BATS: www.stat.duke.edu/~mw/bats.html
 - SEATS y TRAMO:
www.bde.es/servicio/software/econom.html
 - ...

- Cuatro pasos para la modelización SS:
 1. Escribir en un papel el modelo y su forma de espacio de los estados.
 2. Crear una función que traduzca el modelo en código MATLAB.
 3. Estimar los parámetros desconocidos del modelo.
 4. Correr el Filtro de Kalman y/o FIS para obtener estimación óptima de estados y sus covarianzas.

1. Escribir en un papel el modelo y su forma de espacio de los estados

$$\begin{cases} x_{1t+1} = x_{1t} + w_t \\ z_t = x_{1t} + v_t \end{cases}$$

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

2. Crear una función MATLAB que represente el modelo:

$$\begin{cases} x_{1t+1} = x_{1t} + w_t \\ z_t = x_{1t} + v_t \end{cases}$$

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{\Phi}_t \mathbf{x}_t + \mathbf{\Gamma}_t \mathbf{u}_t + \mathbf{E}_t \mathbf{w}_t \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{D}_t \mathbf{u}_t + \mathbf{C}_t \mathbf{v}_t \end{cases}$$

```
function model= SampleSS(p)
model.Phi= [];
model.Gam= [];
model.E= [];
model.H= [];
model.D= [];
model.C= [];
model.Q= [];
model.R= [];
model.S= [];
```

```
function model = example1(p)
model.Phi = 1;
model.Gam = [];
model.E = 1;
model.H = 1;
model.D = [];
model.C = 1;
model.Q = 10.^p(1);
model.R = 10.^p(2);
model.S = 0;
```

3. Estimar los parámetros desconocidos del modelo :

```
% Build SSpace model
sys = SSmodel('y', y, 'model', @example1);
```

```
% Estimate model by exact ML
sys = SSestim(sys);
```

```
% Model table output et
sys = SSvalidate(sys);
```

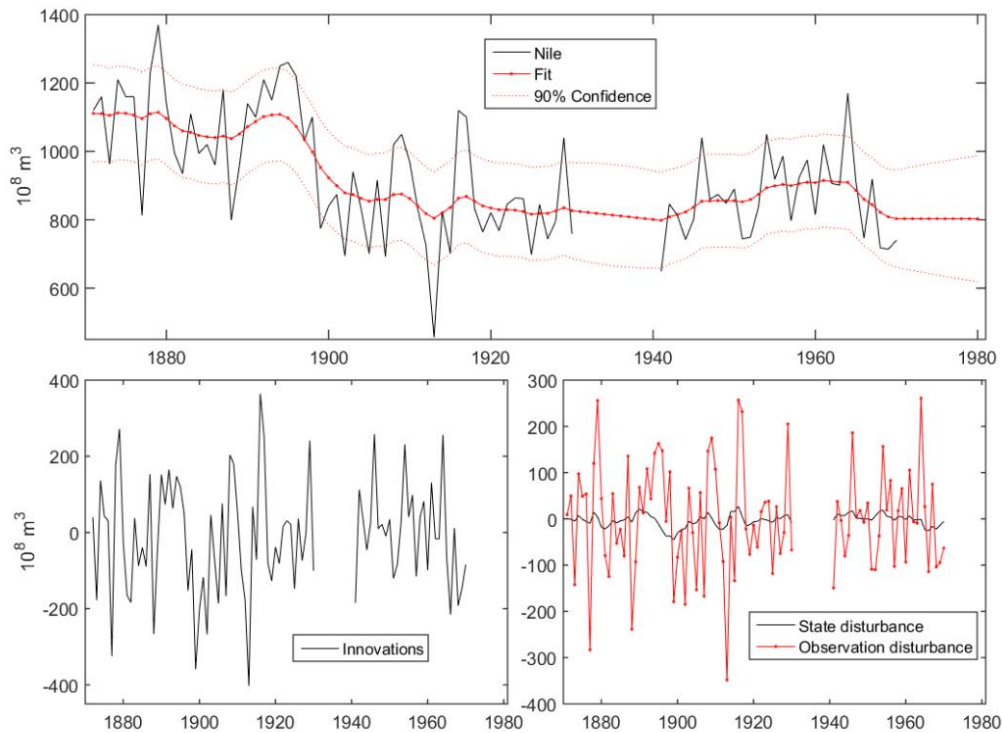
```
-----
Linear Gaussian model: example1.
Objective function: llik
System colapsed at observation 1 of 100.
Exact gradient used.
-----
          Param      S.E.    T-test    P-value |Gradient|
-----
p(1)    3.1404    0.3595    8.7362    0.0000    0.000008
p(2)    4.2084    0.0855   49.2198    0.0000    0.000051
-----
                    AIC: 12.906
                    SBC: 12.9938
                    HQC: 12.9398
Log-likelihood: -571.3177
Corrected R2: 0.2669
```

4. Correr el Filtro de Kalman y/o FIS para obtener estimación óptima de estados y sus covarianzas

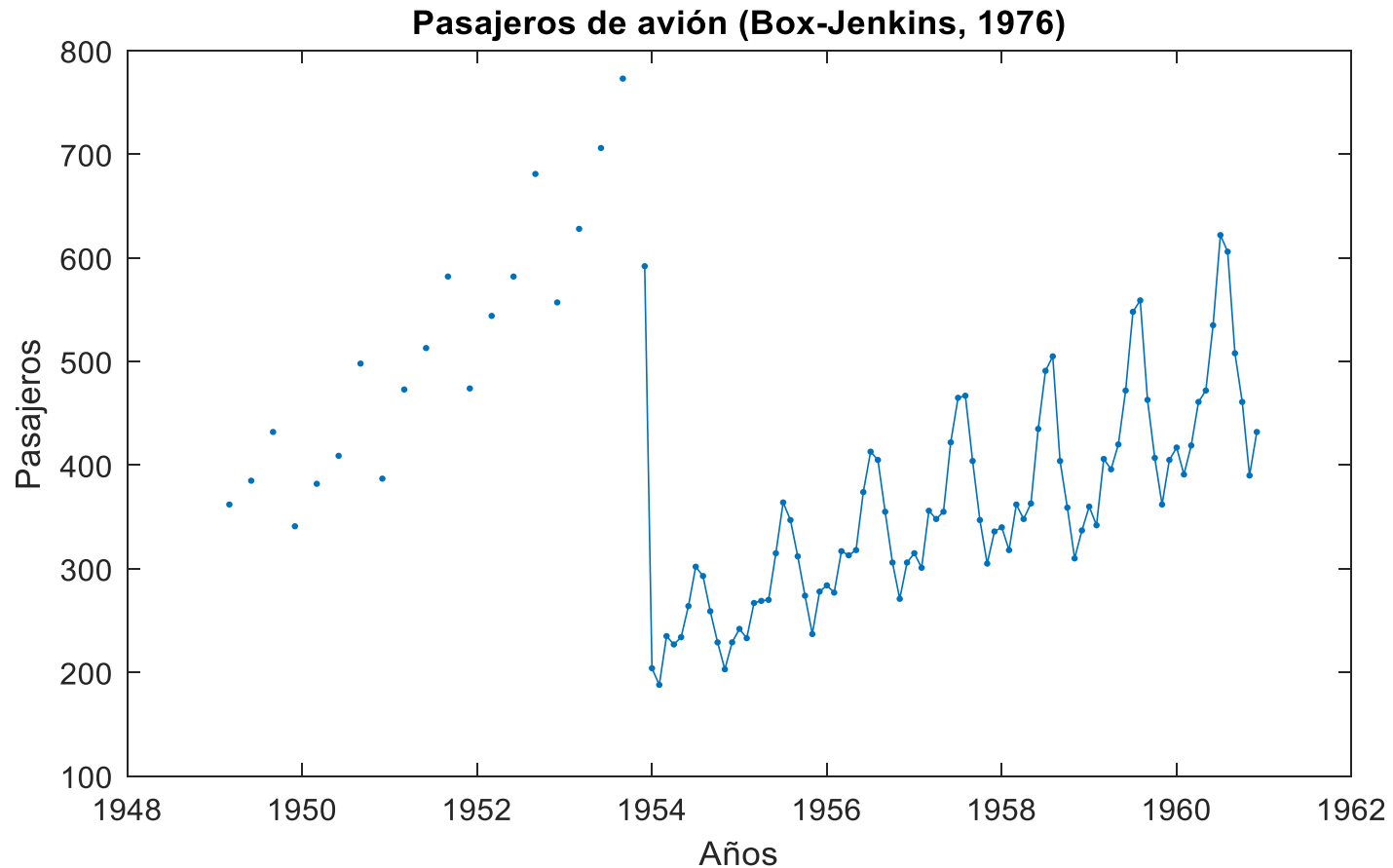
% Smoothing

```
sys = SSsmooth(sys);
```

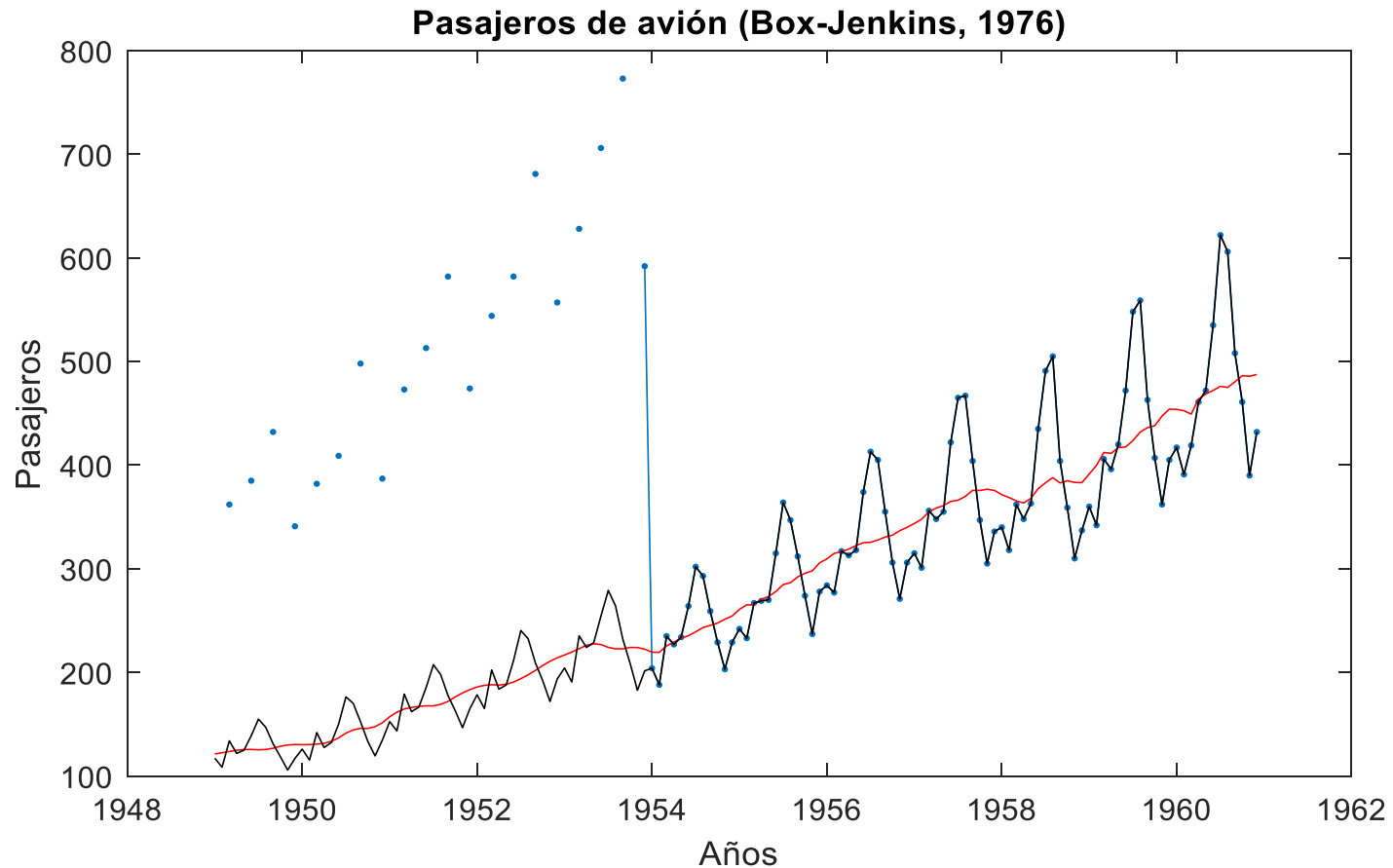
```
plot(t, y, 'k', t, sys.yfit, 'r.-', t, ...  
confband(sys.yfit, sys.F, 1), 'r:')
```



Ejemplo agregación temporal:



Ejemplo agregación temporal:



Otras consideraciones

(opiniones desordenadas)

Relación entre inteligencia artificial, Machine learning y Deep learning

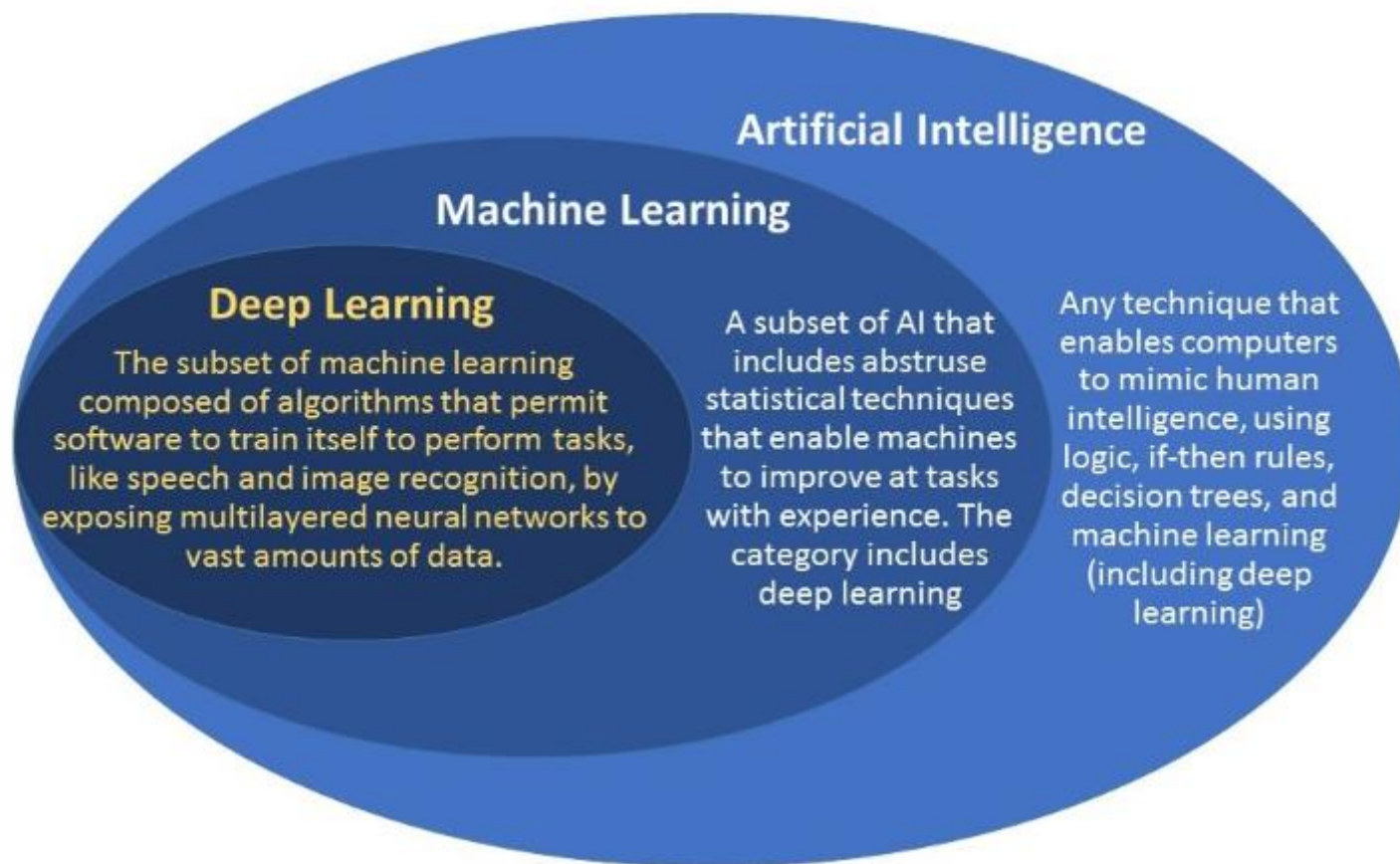


Imagen tomada de <http://www.devacademy.es>

- Big Data y tecnología causa transformación profunda de la sociedad y la economía:
 - Trabajo hecho por robots.
 - 2 de cada 3 operaciones en Bolsa las hacen robots.
 - Coches autónomos.
 - Facebook te conoce mejor que tu madre y que tú mismo. Delegar decisiones. Google también.
 - Ejemplo de marketing analytics.
 - Tecnología blockchain (seguridad en redes con información compartida).

- Programación por reinforcement learning de una máquina capaz de aprender a jugar al ajedrez en 72 horas. (Lai, 2017, Giraffe: Using Deep Reinforcement Learning to Play Chess)
- AlphaGo Zero es el mejor jugador de GO, entrenado en días (Gibney, E. (2017). Self-taught AI is Best Yet at Strategy Game Go: Artificial-Intelligence Program, AlphaGo Zero Trained in Just Days, Without Any Human Input, Nature, International Weekly Journal of Science , Oct., 2017)
- Google ha lanzado auriculares con traductor simultáneo en 40 idiomas.
- Amazon Go (<https://www.cnet.com/pictures/photos-inside-amazon-go-store-no-cashiers-seattle/>).

- ¿Transforma el Big Data la forma de hacer ciencia económica?
 - Una parte ya se hace así, puesto que Big Data incluye la estadística clásica. Modelos de panel, factoriales, etc.
 - Sí, porque se pueden analizar problemas nuevos por la abundancia y detalle de la información o de formas nuevas.
 - Anécdota de función de demanda clásica.
 - Demanda intermitente.
 - Demanda jerárquica.
 - Estudiar implicaciones de la revolución digital.
 - Etc.

- ¿Transforma el Big Data la forma de hacer ciencia económica?
 - Punto filosófico: gente friki que se sale de moldes, muy eficiente, informal, y cambian la forma de publicar resultados...
 - Competiciones.
 - Software libre.
 - WEB 3.0.
 - Necesario estar despiertos a innovaciones y ser flexibles. Formación y adaptación constante.
Aprender Big Data.

- ¿Significa que Big Data va a mejorar los resultados de la Economía como disciplina?
 - Optimistas (u oportunistas) dicen que Sí. Hay muchos papers de predicción que así lo dicen.
 - Serias dudas: Un problema puede ser difícil por dos razones, al menos:
 - Técnicamente difícil pero con estructura bien definida que detectar. Detección de imágenes, lenguaje, etc.
<https://www.youtube.com/watch?v=z8RVC7VmY8k>
 - Técnicamente difícil y además con una estructura muy laxa (economía).