Computer Controlled Systems

Random variables, stochastic processes
Discrete time stochastic LTI models

Katalin Hangos

University of Pannonia
Faculty of Information Technology
Department of Electrical Engineering and Information Systems

hangos.katalin@virt.uni-pannon.hu

March 2017

Overview

Scalar-valued random variables

The random variable ξ has a *normal or Gaussian distribution*, in notation

$$\xi \quad \sim \quad \mathbb{N}(m, \sigma^2) \tag{1}$$

if its probability density function (**p.d.f.**) f_ξ

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \tag{2}$$

where m is its *mean value* and σ^2 is its *variance*.

The *mean value* and *variance* of the random variable ξ with its p.d.f. f_ξ

$$E\{\xi\} = \int x f_\xi(x) dx \quad , \quad \sigma^2\{\xi\} = \int (x - E\{\xi\})^2 f_\xi(x) dx$$

Covariance

The *covariance* of two scalar-valued random variables ξ és θ

$$COV\{\xi, \theta\} = E\{(\xi - E\{\xi\})(\theta - E\{\theta\})\}$$

The *variance* of a scalar-valued random variable ξ is *the covariance of* ξ *with itself*:

$$\sigma^2\{\xi\} = COV\{\xi, \xi\} = E\{(\xi - E\{\xi\})^2\}$$

Correlation (normed covariance):

$$\rho\{\xi, \theta\} = \frac{E\{(\xi - E\{\xi\})(\theta - E\{\theta\})\}}{\sigma\{\xi\}\sigma\{\theta\}}$$

Vector-valued random variables

Given a vector valued random variable ξ

$$\xi \ : \ \ \xi(\omega), \quad \omega \in \Omega, \quad \xi(\omega) \in \mathbb{R}^{\mu}$$

Its ***mean value*** $m \in R^{\mu}$ *is a real vector.*

Its ***variance*** $COV\{\xi\}$ is a square real matrix, the *covariance matrix*:

$$COV\{\xi\} = E\{(\xi - E\{\xi\})(\xi - E\{\xi\})^T\}$$

Covariance matrices are positive definite symmetric matrices:

$$z^T COV\{\xi\} z \geq 0 \quad , \quad \forall z \in \mathbb{R}^{\mu}$$

Joint probability density functions

The joint probability density function $f(x_1, ..., x_n)$ of the scalar-valued random variables $\xi_1, ..., \xi_n$ is an n-variable non-negative function such that

$$P[a_1 \leq x_1 \leq b_1,\ \ldots,\ a_n \leq x_n \leq b_n] = \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} f(x_1, ..., x_n) dx_1 ... dx_n$$

Two dimensional special case

$$f_{\xi_1,\xi_2}(x_1, x_2)$$

Independence (in the two dimensional case)

$$f_{\xi_1,\xi_2}(x_1, x_2) = f_{\xi_1}(x_1) \cdot f_{\xi_2}(x_2)$$

Multi-dimensional Gaussian distribution

A vector-valued random variable ξ has a normal or Gaussian distribution with mean value m and covariance matrix Σ

$$\xi \sim N(m, \Sigma)$$

if its elements $\xi_i, i = 1, ..., \mu$ are all normally distributed scalar-valued random variables.

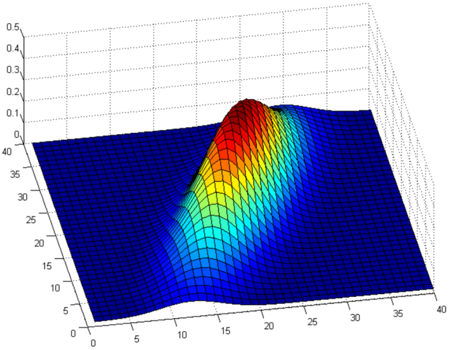
The *probability density function* of a vector-valued Gaussian random variable: with R being a determinant composed from the correlation coefficients ρ_{ij}

$$f(x_1, ..., x_\mu) = \frac{1}{\sqrt{2\pi}\sigma_1 ... \sigma_\mu \sqrt{R}} e^{-\frac{1}{2R}\left(\sum_{i=1}^{\mu}\sum_{j=1}^{\mu}\rho_{ij}\frac{(x_i - m_i)(x_j - m_j)}{\sigma_1 \sigma_2}\right)}$$

Two dimensional Gaussian distribution

Probability density function:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left(\frac{(x_1-m_1)^2}{\sigma_1^2} - 2r\frac{(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \frac{(x_2-m_2)^2}{\sigma_2^2}\right)}$$



Linearly transformed random variables

Let us transform the vector-valued random variable $\xi(\omega) \in R^n$ using the non-singular square transformation matrix $T \in R^{n \times n}$:

$$\eta = T\xi$$

The properties of the vector-valued random variable η:

$$E\{\eta\} = TE\{\xi\} \quad , \quad COV\{\eta\} = TCOV\{\xi\}T^T$$

If the random variable ξ has a Gaussian distribution $N(m_\xi, \Delta_\xi)$ with mean value m_ξ and covariance matrix Δ_ξ, then the transformed random variable η will also be Gaussian $N(m_\eta, \Delta_\eta)$, where

$$m_\eta = Tm_\xi \quad , \quad \Delta_\eta = T\Delta_\xi T^T$$

Overview

Stochastic processes – 1

Stochastic processes are used for describing random disturbances in systems and control theory.

Stochastic process

family of random variables $x(.,.)$ where

$$x : T \times \Omega \to \mathbb{R}^p$$

The set T is called *time*.

- *continuous time* process: $T \subseteq \mathbb{R}$
- *discrete time process*: $T \subseteq \mathbb{N}$

Stochastic processes – 2

- Realization
  the (deterministic) function $x(., \omega_0)$ with ω_0 being fixed
- Fixed-time value
  $x(t_0, .)$ with t_0 is being fixed is a random variable
- Notation
  $x(t, .) = x(t)$ for the random variable generated from the stochastic process x by fixing the time at t

Distribution functions

A stochastic process can be specified by describing all of its finite dimensional distribution functions

Definition

A finite dimensional distribution function of a stochastic process is defined by the formulae

$$F(\zeta_1, ..., \zeta_n; t_1, ..., t_n) = P\{x(t_1) \leq \zeta_1, ..., x(t_n) \leq \zeta_n\}$$

Gaussian or normal process all finite dimensional distribution functions of the process are Gaussian.

Mean value and covariance

Definition (mean value function)

The mean-value function of the stochastic process x is as follows

$$m_x(t) = Ex(t) = \int_{-\infty}^{\infty} \zeta dF(\zeta, t)$$

Note that $m_x(t)$ is an ordinary (deterministic) function of time t.

Definition (covariance function)

The (auto)covariance function of the stochastic process x is defined as

$$r_{xx}(s,t) = cov\ [x(s), x(t)] = E\{\ [x(s) - m(s)][x(t) - m(t)]^T\ \}$$

The covariance function is a deterministic two-variate function.

Stationary stochastic processes

Definition (stationary stochastic process)

A stochastic process x is termed stationary if all of its finite dimensional distribution functions on $x(t_1), ..., x(t_n)$ are identical to that on $x(t_1 + \tau), ..., x(t_n + \tau)$ for all τ.
The process is termed *weakly stationary* if the two first moments of the distribution functions are the same for all τ, i.e.

$$m(t) = const \quad , \quad r_{xx}(s,t) = r_{xx}(t-s)$$

.

White noise processes

Definition (discrete time white noise, e)

A stochastic process $e = \{e(\theta)\}_{\theta=-\infty}^{\infty}$ is a discrete time white noise process if it is a sequence of identically distributed, independent random variables.

Properties

- stationary process (usually $m(t) = 0$ is assumed)
- the covariance function in *real-valued case* is

$$r_{ee}(t) = cov\ [e(s), e(s-t)] = \begin{cases} \sigma^2 & t = 0 \\ 0 & t = \pm 1, \pm 2, \ldots \end{cases}$$

- *A white noise process is **not** necessarily a Gaussian process.*

MA processes

Definition (moving average process (MA process))

Let $e = \{ \ e(k) \ , \ k = ..., -1, 0, 1, 2, ...\}$ be a white noise process with variance σ^2. Then the related process $y = \{y(t)\}_{t=-\infty}^{\infty}$ which fulfils

$$y(k) = e(k) + b_1 e(k-1) + ... + b_n e(k-n) = B^*(q^{-1})e(k)$$

is termed a MA process.

Mean value and auto-covariance function of a MA process

$$m_y(t) = 0, \ r_{yy}(0) = \sigma^2(1+b_1^2+...+b_n^2), \ r_{yy}(1) = \sigma^2(b_1+b_1 b_2+...+b_{n-1}b_n)$$

AR and ARMAX processes

Definition (autoregressive process (AR process))

With the white noise process $e = \{e(t)\}_{t=-\infty}^{\infty}$ above an AR process is defined as follows

$$y(k) + a_1 y(k-1) + \ldots + a_n y(k-n) = A^*(q^{-1})y(k) = e(k)$$

Definition (ARMAX process)

An autoregressive-moving average process with an exogeneous signal (ARMAX process) is a linear combination an AR and MA process extended with an exogeneous signal $u = \{u(t)\}_{t=-\infty}^{\infty}$:

$$A^*(q^{-1})y(k) = B^*(q^{-1})u(k) + C^*(q^{-1})e(k)$$

with $A^*(q^{-1}) = 1 + a_1 q^{-1} + a_n q^{-n}$, $B^*(q^{-1}) = b_0 + b_1 q^{-1} + b_m q^{-m}$, $C^*(q^{-1}) = 1 + c_1 q^{-1} + c_n q^{-n}$ and $m < n$.

General representation theorem

Theorem

Every stationary discrete time stochastic process $x = \{x(k)\}_{-\infty}^{\infty}$ with finite 1st and 2nd momenta can be represented in ARMA form as

$$A^*(q^{-1})x(k) = B^*(q^{-1})e(k)$$

where $\{e(k)\}_{-\infty}^{\infty}$ is a white noise process (not necessarily Gaussian!) and $A^*(z)$ is a stable, $B^*(z)$ is a stable or not unstable polynomial.

Interpretation: *Every stationary discrete time stochastic process $\{x(k)\}_{-\infty}^{\infty}$ can be viewed as the output of a stable discrete time LTI system with pulse transfer operator $H(z) = \frac{B^*(z)}{A^*(z)}$ and with white noise input.*

Overview

DT-LTI state-space models

$$x(k+1) = \Phi x(k) + \Gamma u(k) \qquad (state\ equation)$$
$$y(k) = Cx(k) + Du(k) \qquad (output\ equation)$$

with given initial condition $x(0)$ and

$$x(k) \in \mathbb{R}^n\ ,\ y(k) \in \mathbb{R}^p\ ,\ u(k) \in \mathbb{R}^r$$

being vectors of finite dimensional spaces and

$$\Phi \in \mathbb{R}^{n\times n}\ ,\ \Gamma \in \mathbb{R}^{n\times r}\ ,\ C \in \mathbb{R}^{p\times n}\ ,\ D \in \mathbb{R}^{p\times r}$$

being matrices

DT-LTI SISO I/O system models

Discrete difference equation models: for SISO systems

- Forward difference form

$$y(k+n_a)+a_1y(k+n_a-1)+...+a_{n_a}y(k) = b_0u(k+n_b)+...+b_{n_b}u(k)$$

  with $n_a \geq n_b$ (proper).

  - *Compact form*

$$A(q)y(k) = B(q)u(k) \ ,$$
$$A(q) = q^{n_a} + a_1q^{n_a-1} + ... + a_{n_a} \ , \ B(q) = b_0q^{n_b} + b_1q^{n_b-1} + ... + b_{n_b}$$

- Backward difference form

$$y(k)+a_1y(k-1)+...+a_{n_a}y(k-n_a) = b_0u(k-d)+...+b_{n_b}u(k-d-n_b)$$

  where $d = n_a - n_b > 0$ is the *pole excess (time delay)*.

  - *Compact form*

$$A^*(q^{-1})y(k) = B^*(q^{-1})u(k-d) \ , \ A(q) = q^{n_a}A^*(q^{-1})$$

Overview

DT-LTI stochastic SISO I/O model

Definition (discrete time stochastic LTI input-output model)

The general form of the input-output model of discrete time stochastic LTI SISO systems is the following canonical ARMAX process:

$$A(q)y(k) = B(q)u(k) + C(q)e(k) \tag{3}$$

with the polynomials

$$A(q) = q^n + a_1 q^{n-1} + \ldots + a_n \quad , \quad C(q) = q^n + c_1 q^{n-1} + \ldots + c_n$$

$$B(q) = b_0 q^m + b_1 q^{m-1} + \ldots + b_m$$

where $C(q)$ is assumed to be a stable polynomial.

DT-LTI stochastic state-space model

$$x(k+1) = \Phi x(k) + \Gamma u(k) + v(k)$$
$$y(k) = Cx(k) + e(k)$$

$$\Phi \in \mathbb{R}^{n \times n} \ , \ \Gamma \in \mathbb{R}^{n \times r} \ , \ C \in \mathbb{R}^{p \times n}$$

and with independent discrete time zero mean Gaussian white noise processes $\{v(k)\}_0^\infty$ *and* $\{e(k)\}_0^\infty$

$$E[v(k)v^T(k)] = R_1 \quad , \quad E[v(k)v^T(j)] = 0 \ , \ \forall \ k \neq j$$
$$E[v(k)e^T(j)] = 0 \quad , \quad \forall \ k, j$$
$$E[e(k)e^T(k)] = R_2 \quad , \quad E[e(k)e^T(j)] = 0 \ , \ \forall \ k \neq j$$

Initial conditions

$$Ex(0) = m_0 \quad , \quad cov[x(0)] = R_0$$

Parameters:

$$(\Phi, \Gamma, C \ ; \ R_1, R_2 \ ; \ m_0, R_0)$$

DT-LTI stochastic difference equations

Definition (linear stochastic difference equation)

is in the form

$$x(k+1) = \Phi x(k) + v(k)$$

where $\{v(k)\}_0^\infty$ is a discrete time white noise process and $v(k)$ is independent of $x(k)$.

The solution of the equation above is a stochastic process $\{x(k)\}_0^\infty$ itself.

Solution of the state equation

$$x(k+1) = \Phi x(k) + v(k)$$

- Mean value function $m(k$ is the solution of

$$m(k+1) = \Phi m(k) \quad , \quad m(0) = m_0$$

- Covariance function:

$$P(k) = cov[x(k), x(k)] = E\{\overline{x}(k)\overline{x}^T(k)\} \quad , \quad \overline{x}(k) = x(k) - m(k)$$

$$\overline{x}(k+1)\overline{x}^T(k+1) = [\Phi\overline{x}(k) + v(k)][\Phi\overline{x}(k) + v(k)]^T =$$
$$= \Phi\overline{x}(k)\overline{x}^T(k)\Phi^T + \Phi\overline{x}(k)v^T(k) + v(k)\overline{x}^T(k)\Phi^T + v(k)v^T(k)$$

$$P(k+1) = \Phi P(k)\Phi^T + R_1 \quad , \quad P(0) = R_0$$

The output process

We associate the output stochastic process $\{y(k)\}_0^\infty$ *to the solution of the linear stochastic difference equation (the state equation) by the equation*

$$y(k) = Cx(k)$$

where C *is a constant matrix then*

$$m_y(k) = Cm(k) \quad , \quad r_{yy} = CP(k)C^T$$