Stochastic Signals and Systems

State space representation

2020. december 11.

Multivariate AR(1), I.

From spectral representation to state space representation

Recall the simplest non-trivial example: a stable AR(1) process.

Extension: consider a multivariate AR(1) process given by

$$x_{n+1} = Ax_n + Bv_n,$$

with $x_n \in \mathbb{R}^s$. Here (v_n) is an $\mathbb{R}^t$-valued w.s.st. orthogonal process.

Let $\mathrm{E} v_n v_n^T = \Sigma_{vv}$. The matrices A and B are $s \times s$ and $s \times t$, resp.

Quest.: under what condition is there a unique stationary solution (x_n) ?

Multivariate AR(1), II.

Extending the condition $|a| < 1$. Recall the definition os spectral radius:

$$\rho(A) = \max_{i=1,\dots s} |\lambda_i(A)|,$$

where $\lambda_i(A)$, $i = 1, \dots s$ denote the eigenvalues of A.

Definition

A square matrix A matrix is <u>*stable*</u> *(in discrete sense) if* $\rho(A) < 1$*. Equivalently, A is stable if all the roots of the polynomial equation*

$$|zI - A| = 0$$

are in $D = \{z \ : \ |z| < 1\}$*.*

In other words: the matrix A is stable (in discrete sense) if $\rho(A) < 1$. if its spectral radius is less than 1.

Multivariate AR(1), III.

Recall the following result of linear algebra:

$$\rho(A) = \lim_{n\to\infty} \|A^n\|^{1/n}.$$

Hence letting $r := \rho(A)$ we have for any $\varepsilon > 0$, with some $C = C(\varepsilon) > 0$,

$$\|A^n\| \leq C(r+\varepsilon)^n.$$

Thus if $r = \rho(A) < 1$ then $\|A^n\|$ is exponentially decaying.

Multivariate AR(1), IV.

Proposition

If A is a stable, then there is a unique w.s.st. solution (x_n), given by

$$x_{n+1} = \sum_{k=0}^{\infty} A^k B v_{n-k}.$$

Exercise. Existence: prove that the r.h.s. converges in $L_2(\Omega, \mathcal{F}, \mathcal{P})$, and it satisfies the AR(1) dynamics.

Multivariate AR(1), V.

Proving uniqueness: iterate $x_{n+1} = Ax_n + Bv_n$ forward in time:

$$x_{n+\tau} = A^\tau x_n + \sum_{k=0}^{\tau-1} A^k B v_{n+\tau-1-k}, \qquad \tau \geq 1.$$

Letting $\tau \to \infty$ we get the x_n must be as given above.

The proposition implies that x is a causal linear function of v, thus:

$$H^x_{n+1} \subset H^v_n \quad \forall n.$$

Remark. Note the shift in the time index, due to the fact that v_n effects x_{n+1}. Not aligned with notation for ARMA processes.

Multivariate AR(1), VI.

An operator form: letting q^{-1} denote the backward shift operator equation (1) can be written as

$$(qI - A)x = Bv. \qquad (1)$$

Question: under what condition does a w.s.st. solution exist, which is not necessarily causal ? Using spectral methods we ‘easily get:

Proposition

Assume, that $e^{i\omega} I - A$ is not singular for all $\omega \in [0, 2\pi]$. Then (1) has a unique w.s.st. solution.

Partial observation

State space equations with **partial observation**:

$$
\begin{aligned}
x_{n+1} &= Ax_n + Bv_n \\
y_n &= Cx_n + Dw_n,
\end{aligned}
$$

where y is called *observation*.

The dimension of is y typically much smaller than the dimension of x.

$$\dim x \gg \dim y.$$

The *observation noise* is Dw_n, the matrix D is square.

$$\dim w = \dim y.$$

Partial observation, I.

Extension to state-space systems with partial observation.

Mathematically speaking we consider the dynamics given by the set of equations

$$\begin{aligned} x_{n+1} &= Ax_n + Bv_n \\ y_n &= Cx_n + Dw_n, \end{aligned}$$

The dimension of the observed process y (simply called observation) is typically much smaller than the dimension of the state process x.

Condition

The joint noise process $(v_n, w_n), -\infty < n < \infty$ is a w.s.st. orthogonal process with covariance matrix

$$\begin{pmatrix} \Sigma_{vv} & \Sigma_{vw} \\ \Sigma_{wv} & \Sigma_{ww} \end{pmatrix}. \tag{2}$$

Partial observation, II.

The above set of equation for modelling a multivariate time series is called a state-space model or linear stochastic system.

Foundations has been laid down by the Kyoto prize laureate Hungarian scientist R. Kalman.

This theory revolutionized the area of w.s.st. processes, in particular especially by allowing a very effective solution of the so-called filtering problem.

Partial observation, III.

Question: what is the innovation process of the observed process (y_n) ?

An amazingly simple and elegant answer: the innovation process (ν_n) satisfies:

$$\begin{aligned} x_{n+1} &= Ax_n + K\nu_n \\ y_n &= Cx_n + \nu_n, \end{aligned}$$

with some matrix K, called the Kalman-gain. Note that $\dim y = \dim \nu$, as expected.
In operator form:

$$\begin{aligned} (qI - A)x &= K\nu \\ y &= C + \nu, \end{aligned}$$

Partial observation, IV.

Solving the above we can write formally:

$$y = (C(qI - A)^{-1}K + I)\nu,$$

Getting ν: invert the above. We may use linear algebra, or simply (3), (3):

Write $\nu = y - Cx$, and substitute to get

$$(qI - A)x = K(y - Cx)$$

from which

$$(qI - A + KC)x = Ky$$

and hence

$$x = (qI - A + KC)^{-1}Ky$$

Partial observation, V.

Finally:

$$\nu = y - Cx = (I - C(qI - A + KC)^{-1}K)y$$

The yet open issue: how doe we get K ?

State space model, partial observation

The joint noise process (v_n, w_n) is a w.s.st. orthogonal process.

The covariance matrix of the noise is:

$$\mathrm{E}(v_n,\ w_n)\begin{pmatrix} v_n^T \\ w_n^T \end{pmatrix} = \begin{pmatrix} \Sigma_{vv} & \Sigma_{vw} \\ \Sigma_{wv} & \Sigma_{ww} \end{pmatrix}.$$

This set of equation is called a **state-space model** or

linear stochastic system.

$$\begin{aligned} x_{n+1} &= Ax_n + Bv_n \\ y_n &= Cx_n + Dw_n, \end{aligned}$$

Covariance of x_n

Take the dyadic product of $x_{n+1} = Ax_n + Bv_n$ with itself:

$$x_{n+1}x_{n+1}^T = Ax_nx_n^TA^T + Bv_nv_n^TB^T + Ax_nv_n^TB^T + Bv_nx_n^TA^T.$$

Now $x_{n+1} = \sum_{k=0}^{\infty} A^kBv_{n-k} \quad \Rightarrow \quad x_{n+1} \perp v_{n+1}.$

Then taking expectation on both sides we get:

$$\mathrm{E}x_{n+1}x_{n+1}^T = A\left(\mathrm{E}x_nx_n^T\right)A^T + B\left(\mathrm{E}v_nv_n^T\right)B^T.$$

Then $P = \mathrm{E}x_nx_n^T$ satisfies the equation

$$P = APA^T + B\Sigma_{vv}B^T.$$

This equation is called a (discrete-time) **Lyapunov equation**.

Auto-covariance function of x_n

Iterate $x_{n+1} = Ax_n + Bv_n$ forward in time $\tau \geq 1$ times:

$$x_{n+\tau} = A^\tau x_n + \sum_{k=0}^{\tau-1} A^k B v_{n+\tau-1-k}.$$

Multiply the equation by x_n^T from left, and take expectation, we get:

$$\mathrm{E} x_{n+\tau} x_n^T = A^\tau \mathrm{E} x_n x_n^T.$$

The auto-covariance function of x is $R(\tau) = \mathrm{E} x_{n+\tau} x_n^T$. If A stable, then

$$R(\tau) = A^\tau P \quad for \quad \tau \geq 0, \quad R(\tau) = P(A^T)^\tau \quad for \quad \tau \leq 0.$$

Auto-covariance function of y_n

Let us now consider a general linear stochastic system given by

$$\begin{aligned} x_{n+1} &= Ax_n + Bv_n \\ y_n &= Cx_n + Dw_n, \end{aligned}$$

Assume, that A stable. Then the auto-covariance function of (y_n) is:

$$\begin{aligned} R^y(0) &= \mathrm{E}(y_n y_n^T) = CPC^T + D\Sigma_{ww}D^T \qquad \text{and} \\ R^y(\tau) &= \mathrm{E}(y_{n+\tau} y_n^T) = CA^\tau PC^T \qquad \text{for } \tau \geq 1. \end{aligned}$$

Uniqueness?

$$\begin{aligned} x_{n+1} &= Ax_n + Bv_n \\ y_n &= Cx_n + Dw_n. \end{aligned}$$

The state-space description is far from being unique. From (v, w) to y:

Let T be a non-singular linear transformation, define $x' = Tx$. Then

$$\begin{aligned} x'_{n+1} &= TAT^{-1}x'_n + TBv_n \\ y_n &= CT^{-1}x'_n + Dw_n. \end{aligned}$$

The two systems *generate the same input-output mapping*:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{is equivalent to} \quad \begin{pmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{pmatrix}.$$

The stochastic realization problem

Now, look for a representation of the process (y_n) without specifying the driving noise process (v, w).

Problem: realize a given auto-covariance sequence $R^y(.)$ in the form

$$\begin{aligned} R^y(0) &= CPC^T + D\Sigma_{ww}D^T \qquad \text{and} \\ R^y(\tau) &= CA^\tau PC^T \qquad \text{for } \tau \geq 1. \end{aligned}$$

We have to find appropriate $A,\ C$ matrices.

This is called the **stochastic realization problem**.

Initialization at time 0

Assume, the state-space equation is initialized at $n = 0$, rather than $-\infty < n < +\infty$.

Let us assume that $\mathrm{E} x_0 = 0$, and $\mathrm{E} x_0 x_0^T = P_0$.

Then $P_n = \mathrm{E} x_n x_n^T$, satisfies

$$P_{n+1} = A P_n A^T + B B^T$$

with initial condition P_0.

Exercise. (HW) Show that if A is stable, then P_n converges to the unique solution of

$$P = A P A^T + B \Sigma_{vv} B^T.$$

Non-singular covariance matrix of x

$P = \mathrm{E} x_n x_n^T$ can be written as $P = \sum_{k=0}^{\infty} A^k B \Sigma_{vv} B^T (A^T)^k$

If Σ_{vv} is non-singular, we may assume $\mathrm{E} v_n v_n^T = I$. Then we can write P:

$$P = \mathcal{C}_\infty \mathcal{C}_\infty^T, \qquad \text{with} \quad \mathcal{C}_\infty = (B,\ AB,\ A^2 B,\ \dots).$$

Now $\mathrm{rank}(P) = s \iff \mathrm{rank}(\mathcal{C}_\infty) = s$.

Define the **controllability matrix** by

$$\mathcal{C} = (B,\ AB,\ A^2 B,\ \dots\ A^{s-1} B).$$

$$\mathcal{C} = (B,\ AB,\ A^2B,\ \ldots\ A^{s-1}B).$$

Since, by the Cayley-Hamilton theorem

$$\sum_{k=0}^{s} \alpha_k A^k = 0, \quad \alpha_s = 1,$$

all the columns of $A^m B$, $m \geq 0$, can be expressed via the columns of $\mathcal{C}$.

Thus it follows that

$$\mathrm{rank}\ \mathcal{C}_\infty = \mathrm{rank}\ \mathcal{C}.$$

Then we get:

$P = \mathrm{E} x_n x_n^T$ is non-singular $\Longleftrightarrow$ the *controllability matrix* $\mathcal{C}$ has full rank.

State space representation of AR processes

Let (y_n) be a w.s.st. AR(p) process defined by

$$A(q^{-1})y = e, \quad a_0 = 1, \; a_p \neq 0.$$

Define the state vector $x_n = (y_{n-1}, \ldots, y_{n-p})$. Then

$$\begin{aligned} x_{n+1} &= \widetilde{A} x_n + b e_n \\ y_n &= b^T x_{n+1}, \end{aligned}$$

with $$\tilde{A} = \begin{pmatrix} -a_1 & \ldots & & -a_p \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$\tilde{A}$ is the **companion matrix** associated with the polynomial $A(z^{-1})$.

The process $A(q^{-1})y = e$ can be realized by the state-space system

$$\begin{aligned} x_{n+1} &= \widetilde{A}x_n + be_n \\ y_n &= b^T x_{n+1}, \end{aligned}$$

We assumed that $A(z^{-1})$ is a stable polynomial, hence e is the innovation process of y.

The the eigenvalues of $\widetilde{A}$ are identical with the roots of $A(z^{-1})$.

Thus, if $A(z^{-1})$ is stable, then $\widetilde{A}$ is also stable.